

Modulating Synthetic Pathways in Megasyntases

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich 14
Biochemie, Chemie und Pharmazie
der Johann Wolfgang Goethe–Universität
in Frankfurt am Main

von
Franziska Stegemann
aus Frankfurt am Main

Frankfurt 2021
(D30)

vom Fachbereich Biochemie, Chemie und Pharmazie der
Johann Wolfgang Goethe–Universität als Dissertation angenommen.

Dekan: Prof. Dr. Clemens Glaubitz

Gutachter: Prof. Dr. Martin Grininger
Prof. Dr. Stefan Knapp

Datum der Disputation: 21.09.2021

Vorwort

Teile der vorliegenden Arbeit wurden vorab zur Publikation akzeptiert in: F. Stegemann, M. Grininger, "Transacylation Kinetics in Fatty Acid and Polyketide Synthases and its Sensitivity to Point Mutations", *ChemCatChem* (2021).^[1] Eine detaillierte Auflistung der Inhalte, mein eigener Anteil, Copyright sowie Creative-Commons-Lizenzen sind im Anhang unter dem Abschnitt "Statement of Personal Contributions" dargestellt.

In die vorliegende Arbeit sind die Ergebnisse verschiedener Bachelor- und Master-Abschlussarbeiten des FB14 eingegangen, die von mir betreut wurden. Eine detaillierte Auflistung findet sich ebenfalls im Anhang unter dem Abschnitt "Statement of Personal Contributions". Zusätzlich ist an geeigneten Stellen auf entsprechende Arbeiten verwiesen.

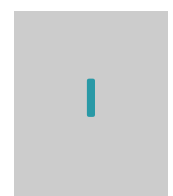
Zusammenfassung

Polyketide sind eine Klasse wichtiger Naturstoffe, die aufgrund ihrer nützlichen Eigenschaften in vielen Bereichen der Medizin als Pharmazeutika Anwendung finden.^[2] Ihr Wirkspektrum reicht von antibakteriellen und antimykotischen über immunsuppressive bis hin zu antitumoralen Eigenschaften.^[3] Diese funktionelle Vielfalt ist auf ihre strukturelle Diversität zurückzuführen. Polyketide können äußerst komplexe Strukturen aufweisen und eine Vielzahl verschiedener funktioneller Gruppen tragen. Inspiriert von der Natur und angetrieben vom Wirkverlust zahlreicher Pharmazeutika sowie auftretenden Resistenzen wird häufig versucht, modifizierte oder neue Naturstoffe zu generieren. Doch wie werden diese komplexen und vielfältigen Moleküle in der Natur hergestellt?

Große Enzymkomplexe, die sogenannten Polyketidsynthasen (PKSs), führen die Biosynthese von Polyketiden durch.^[4] Auch wenn ihre Komplexität anderes erwarten ließe, werden Polyketide aus einfachen Acyleinheiten schrittweise aufgebaut. Verschiedene Enzyme der PKSs sind für Kondensation, Modifikation und Transport der Substrate und Intermediate zuständig. Abhängig von der Anordnung dieser Enzyme wird zwischen verschiedenen PKS-Typen unterschieden. Die vorliegende Thesis befasst sich ausschließlich mit Typ I-PKS-Systemen, die ihre enzymatischen Funktionen als Domänen auf einer oder mehreren Polypeptidketten tragen. Der Begriff PKS bezieht sich im Folgenden auf Typ I-PKSs.

In allen PKSs schleust die Domäne Acyltransferase (AT) die kleinen Acylsubstrate in den Enzymkomplex ein. Genauer gesagt belädt die AT das Acyl-Carrier-Protein (ACP) mit Substraten, das diese wiederum zu den weiteren katalytischen Domänen zur Prozessierung transportiert. Manche PKSs nutzen ihre Domänen mehrfach in repetitiven Zyklen und werden deshalb iterative PKSs genannt. Andere PKSs enthalten mehrere Domänensätze und nutzen jede Domäne nur genau ein Mal für die Herstellung eines Polyketids. Diese Domänensätze werden als Module bezeichnet und die zugehörigen PKSs deshalb modulare PKSs genannt. PKSs oder PKS-Module bestehen aus mindestens drei Domänen, die zur Kondensation der Acylsubstrate nötig sind – AT, ACP und β -Ketoacylsynthase (KS).^[5] Drei weitere Domänen können diesen Minimalsatz erweitern und das Substrat modifizieren. Die Domänenzusammensetzung bestimmt den Grad der Substratmodifikation (keine, partielle oder vollständige Reduktion) und somit die funktionelle Vielfalt im resultierenden Polyketid.

Polyketide und Fettsäuren ähneln sich auf den ersten Blick nur wenig. Dennoch sind ihre Biosynthesewege eng miteinander verwandt.^[3,6] Fettsäuren werden von Fettsäuresynthasen (FASs) hergestellt, die viele strukturelle sowie funktionelle Eigenschaften mit PKSs teilen.^[3]



FASs bestehen aus den gleichen Domänen, die die gleichen Reaktionen durchführen, um aus kleinen Acylsubstraten Fettsäuren herzustellen. Im Gegensatz zu PKSs enthalten FASs jedoch immer einen kompletten Domänensatz, der iterativ und vollständig reduzierend arbeitet.

FASs sind insgesamt gut in ihrer Struktur und Arbeitsweise verstanden,^[7–9] während über PKSs verhältnismäßig wenig bekannt ist. Die vorliegende Thesis beschäftigt sich mit der detaillierten Charakterisierung der AT verschiedener PKSs in ihrer Substratselektivität, ihrer AT-ACP Domänen-Domänen-Interaktion und ihren enzymkinetischen Eigenschaften. Der Vergleich mit enzymkinetischen Daten zu FASs liefert die folgenden Erkenntnisse:

1. ATs der PKSs erscheinen grundsätzlich langsamer als ATs der FASs. Dies könnte auf deren unterschiedliche Biosynthesewege zurückzuführen sein. Fettsäuren sind für den Organismus essentiell und erfüllen wichtige physiologische Aufgaben, u. a. als Bausteine für Membranen, als Energiequelle oder als Botenstoffe in Signalkaskaden.^[10–12] Diese vielfältigen Funktionen erfordern hohe Mengen an Fettsäuren, während Polyketide als Sekundärstoffe schon bei niedrigeren Konzentrationen wirken können.
2. Die langsameren ATs der modularen PKSs zeigen sogar in Abwesenheit der nativen Substrate keine Aktivität gegenüber den getesteten nicht-nativen Substraten. Dies steht im Gegensatz zu den schnelleren ATs von iterativen PKSs und FASs, die auch nicht-native Substrate auf ihr jeweiliges ACP laden. ATs modularer PKSs zeigen demnach als einzige eine hohe Substratspezifität in ihrer Beladungsfunktion und erfüllen wohl die Rolle eines Selektivitätsfilters in PKSs.
3. Die detaillierte Analyse der AT-vermittelten Reaktionsfolge zeigt, dass die Substratselektivität grundsätzlich in beiden Schritten der ACP-Beladung erfolgen kann. Die Beladung mit dem korrekten Substrat ist nicht über eine hydrolytische Korrekturfunktion sichergestellt, wie es beispielsweise für Aminoacyl-tRNA-Synthetasen berichtet wird.^[13]

Darüber hinaus gibt eine Mutationsstudie weitere hilfreiche Aufschlüsse über das Zusammenspiel von AT und ACP. Als Testsystem dient Modul 5 (M5) der wohl bestuntersuchten PKS, 6-Desoxyerythronolid B-Synthase (DEBS) (im Folgenden DEBSM5 genannt). Punktmutationen in der AT:ACP-Schnittstelle haben komplexe Effekte auf die AT-vermittelten Reaktionen. Die Sensitivität gegenüber Punktmutationen ist auf eine hohe Plastizität der AT-ACP-Interaktion zurückzuführen, wie sie auch vor kurzem für eine Typ II-FAS beschrieben wurde.^[14]

Diese detaillierten Erkenntnisse über die kinetischen Eigenschaften von ATs sind höchst relevant für Engineering-Ansätze. Aus dem Satz analysierter Domänen weist besonders die AT der Säuger-FAS Eigenschaften eines primordialen Enzyms auf (schnell, polyspezifisch, plastisch) und kann vermutlich leichter neue oder veränderte Funktionen einführen als spezialisierte ATs.

Somit scheint diese Domäne für Engineering-Ansätze gut geeignet zu sein, in denen neue oder nicht-natürliche Substrate über die AT geladen werden sollen. In anderen Ansätzen, die auf den gezielten Einbau eines spezifischen Substrats abzielen, eignet sich besonders die Verwendung von ATs der modularen PKSs für einen Domänenaustausch. Mit ihren intrinsisch strikten Substratspezifitäten können sie leicht einen hochselektiven Substrateinbau gewährleisten.

Aufbauend auf diesen Ergebnissen beschäftigt sich die vorliegende Thesis mit dem Engineering der Säuger-FAS zu PKS-ähnlichen Fertigungsstraßen mit dem Ziel, maßgeschneiderte Produkte effizient herzustellen. FASs sind grundsätzlich deutlich schneller und effizienter in ihrer Produktsynthese als PKSs und deshalb vielversprechende Kandidaten.^[15,16] Jahrelange Erfahrung im Bereich der FAS-Charakterisierung und des FAS-Engineerings fließt in dieses ambitionierte Projekt mit ein und beleuchtet besonders die Säuger-FAS als geeignetes System. Die offene Architektur bietet die Möglichkeit, Domänen zu entfernen,^[17] was Voraussetzung für nicht- und partiell-reduzierende FAS-Module ist. Folgende Engineering-Aspekte müssen beachtet werden, um die iterative Säuger-FAS in eine vektorielle Synthesestraße umzuwandeln:

1. In der Säuger-FAS gibt es nur eine AT, die für die Beladung mit Starter- und Elongationssubstrat zuständig ist. In der FAS-basierten Fertigungsstraße muss diese Beladungsfunktion von Starter- und Elongationssubstrat getrennt werden. Es muss sichergestellt werden, dass jedes Modul über die eigene AT gezielt ein Substrat einbaut und gegen das oder die anderen Substrate selektiert.
2. Die Prozessierung der Substrate innerhalb der FAS-Module (Kettenelongation) muss sichergestellt werden. Es müssen verschiedene FAS-Module entworfen werden, die über ihre Domänenzusammensetzung den Erhalt der verschiedenen funktionellen Gruppen ermöglichen. Für maximale Variabilität müssen vier FAS-Module erstellt werden.
3. Die Interaktion und Kommunikation zwischen den Modulen (Kettentranslokation) müssen sichergestellt werden. In PKSs kommen hierfür kovalente und nicht-kovalente Linker zum Einsatz. Geeignete Linker müssen für die FAS-Fertigungsstraße entwickelt werden.

Die vorliegende Thesis widmet sich vor allem dem ersten Aspekt, der Installation einer spezifischen Ladefunktion. Hierzu gibt es prinzipiell zwei Möglichkeiten:^[18]

1. Zum einen bietet gezielte Mutagenese die Möglichkeit, die Substratspezifität der AT zu beeinflussen. Diese Herangehensweise erzielt zwar eine veränderte, aber oftmals keine strikte Substratspezifität. Mutagenese der AT der Säuger-FAS identifizierte die Mutation R606A, die eine drastische Verschiebung von der Beladung mit Elongationssubstrat zur Beladung mit Startersubstrat hervorrief.^[9,19] Diese punktmutierte AT-Domäne kann für ein FAS-basiertes Startermodul interessant sein.

2. Wie bereits erwähnt, kann durch den Domänen austausch mit einer intrinsisch spezifischen AT eine selektive Ladefunktion in ein unspezifisches System eingeführt werden. Diese Vorgehensweise stellt zwar grundsätzlich eine hohe Spezifität in Aussicht, ist aber oft mit starken Stabilitäts- und Aktivitätseinbußen aufgrund von künstlichen permanenten und nicht-permanenten Protein-Protein-Interaktionen verbunden.

Die vorliegende Thesis setzt den Fokus auf die Installation einer spezifischen Ladefunktion von Elongationssubstraten in der Säuger-FAS mittels Domänen austausch. Hierzu werden substratspezifische ATs von modularen PKSs verwendet. Erste Experimente in der Vollängen-FAS zeigen, dass der Domänen austausch mit deutlichen Expressionsausbeute- und Stabilitätsverlusten einhergeht. Um zunächst stabile FAS/PKS-Chimären zu etablieren, wird die FAS auf das Subkonstrukt KS-AT reduziert. Nach erfolgreicher Installation der Ladefunktion in der Didomäne soll die stabile und aktive Chimäre in ein FAS-Elongationsmodul implementiert werden.

Zunächst liegt der Fokus auf der Erstellung einer FAS/DEBSM5-Chimäre. Struktur- und Sequenzalignments werden herangezogen, um geeignete Domänenschnittstellen zu finden. Diverse Grenzen werden experimentell getestet – inklusive in der Literatur für modulare PKSs beschriebene universelle Schnittstellen.^[20] Alle acht Chimären zeigen deutliche Ausbeute- und Stabilitätseinbußen und auch die in anderen Systemen etablierten Grenzen zeigen keine Verbesserung. Dies wirft die folgenden Fragen auf: Wurden die passenden Domänengrenzen noch nicht gefunden? Sind die beiden Proteine überhaupt kompatibel für den AT-Austausch? Sind die beiden Proteine grundsätzlich geeignet für einen AT-Austausch?

Um diese Fragen zu beantworten, werden die beiden Wildtyp-Proteine in ihrer Robustheit gegenüber Mutationen in den Domänenschnittstellen untersucht. Hierzu werden verschiedene Bioinformatikprogramme verwendet, um für die Proteinstabilität relevante Aminosäuren zu identifizieren und die Effekte einer Mutation dieser Reste vorherzusagen. Die Säuger-FAS-Didomäne stellt sich hierbei als äußerst robust heraus und erlaubt die Mutation vieler Reste in den Schnittstellen. Allerdings identifizieren *In silico*-Vorhersage und experimentelle Daten eine Aminosäure, die für die Stabilität der Didomäne essentiell zu sein scheint. Grundsätzlich scheint sich die Säuger-FAS unter Bedacht des gefundenen Hotspots sehr gut für einen AT-Austausch zu eignen. DEBSM5 hingegen zeigt keinerlei Toleranz gegenüber Punktmutationen in den Schnittstellen und jede Mutation führt experimentell zu drastischen Stabilitätseinbrüchen, die in diesem Ausmaß nicht für alle Reste von den Bioinformatikprogrammen vorhergesagt werden. Dies kann auf die verhältnismäßig schlechte Qualität der Strukturlösung zurückzuführen sein, da diese die Grundlage aller Vorhersagen darstellt. Die Studie zeigt deutlich, dass DEBSM5 nicht für einen AT-Austausch geeignet und somit höchstwahrscheinlich für die Instabilität der FAS/DEBSM5-Chimären verantwortlich ist. Eine weitere Analyse der Schnittstellen zeigt zudem, dass die beiden Proteine nicht kompatibel für einen AT-Austausch sind.

Da die Säuger-FAS mit ihrer Robustheit und Plastizität an sich aber äußerst gut für diesen Engineering-Ansatz geeignet ist, wird in weiteren Experimenten ihre AT durch Elongations-substrat-spezifische ATs aus *Escherichia coli* (*E. coli*) FAS und DEBSM6 ersetzt. Auch hier zeigen sich teilweise Stabilitätsprobleme, die ebenfalls auf Inkompatibilitäten der Systeme zurückzuführen sind. Eine der Säuger-FAS/DEBSM6-Chimären zeigt jedoch in ersten Untersuchungen hohe Stabilität und Qualität. Die weitere Charakterisierung der AT-Domäne lässt erkennen, dass die Kinetiken in der Chimäre im Vergleich zum Wildtyp-DEBSM6 zwar verändert sind, aber die Chimäre immer noch 70 % der Wildtyp-Protein-Umsatzrate aufweist. Dies stellt den ersten Meilenstein in der Etablierung der chimären Fertigungsstraße dar: die Etablierung geeigneter Domänengrenzen für den AT-Austausch in der Säuger-FAS. Leider ist die DEBSM6 AT an sich sehr langsam und deshalb nur bedingt für den Einsatz in der FAS-basierten Fertigungsstraße geeignet. Nun müssen geeignete Donor-ATs gefunden werden.

Darauf aufbauend werden im Minimalmodul der Säuger-FAS (KS-AT-ACP) andere spezifische ATs von verschiedenen PKSs ohne vorherige Experimente mit denselben Domänengrenzen eingebaut. Nur eines der neun Konstrukte zeigt akzeptable Expressionsausbeuten und wird weiter analysiert. Experimente zeigen, dass dieses FAS-basierte Elongationsmodul tatsächlich nicht mehr Startersubstrat, sondern nur noch Elongationssubstrat lädt. Dies demonstriert das Potential des Engineering-Ansatzes und dessen grundsätzliche Machbarkeit. Im Zusammenspiel mit einem PKS-Startermodul werden jedoch leider nur geringe Mengen des Produkts detektiert. Dies kann auf alle drei oben genannten Engineering-Aspekte zurückzuführen sein. Um ein intrinsisches Problem der Wildtyp-AT oder ein Problem in der chimären Ladefunktion ausschließen zu können, ist die Analyse des Subkonstrukts KS-AT der Wildtyp-PKS und der Chimäre nötig. Falls die Chimäre AT-Aktivität zeigt, muss das Design des Elongationsmoduls überdacht und die intermodulare Kommunikation innerhalb der Fertigungsstraße optimiert werden.

Insgesamt trägt die vorliegende Thesis maßgeblich zum detaillierten mechanistischen Verständnis der AT in FASs und PKSs bei. Sie beleuchtet Unterschiede der verschiedenen Klassen und Typen und bietet Einblicke in die essentielle AT-ACP-Interaktion einer PKS. Darüber hinaus werden erste wichtige Schritte zur Etablierung einer chimären Fertigungsstraße unternommen. Die Ergebnisse des Protein-Engineerings zeigen, dass die Säuger-FAS für die Etablierung einer Synthesestraße prädestiniert ist. Geeignete Domänenschnittstellen für einen AT-Austausch werden identifiziert und erste aktive Chimären erfolgreich erstellt. Weiterhin zeigt sich, dass sich jedes System individuell verhält und minimale Änderungen in den Grenzflächen einen großen Einfluss auf Stabilität und Aktivität haben können. Zukünftige Engineering-Ansätze sollten auf einer Analyse der KS-AT-Schnittstellen des Wildtyp-Proteins – unterstützt durch den Einsatz von Bioinformatikprogrammen – und der Charakterisierung der Donor-AT im Wildtyp- und im chimären Subkonstrukt KS-AT aufbauen, um eine spezifische Ladefunktion in FAS-basierten Modulen erfolgreich zu installieren.

Contents

List of Tables	VII
List of Figures	VIII
List of Abbreviations	X
1 Abstract	1
2 Introduction	3
2.1 Natural Products in Modern Medicine	3
2.2 Fatty Acid and Polyketide Synthases	4
2.3 Biosynthesis of Fatty Acids and Polyketides	5
2.4 Evolution of Fatty Acid and Polyketide Synthases	8
2.5 Structural Comparison of Fatty Acid and Polyketide Synthases	10
2.6 Key Players in Fatty Acid and Polyketide Synthesis	13
2.6.1 The Acyl Carrier Protein	13
2.6.2 The Acyl Transferase	14
2.6.3 The β -Ketoacyl Synthase	17
2.6.4 Important Domain-Domain Interactions	20
2.7 Generation of New Natural Products	21
2.8 Motivation Behind This Thesis	22
2.9 Aim	24
3 Results	25
3.1 Preliminary Work	25
3.2 Kinetic Description of AT-mediated Reactions	25
3.2.1 Transacylation	26
3.2.2 AT-mediated Hydrolysis	28
3.3 Kinetic Analysis of Acyl Transferases	30
3.3.1 Selection of Acyl Transferases	30
3.3.2 Isolation of Acyl Transferases	32
3.3.3 Isolation of Acyl Carrier Proteins	32
3.3.4 Initial Substrate Screening	33
3.3.5 AT-mediated Hydrolysis	35
3.3.6 AT-mediated Transacylation	37
3.3.7 AT:ACP Interface Mutation Study	41
3.4 Generation of a FAS-Based Chimeric Elongation Module	46
3.4.1 Chimeric mFAS with Swapped DEBS AT5 and RAPS AT14	46
3.4.2 Generation of mFAS/DEBS3M5 KS-AT Chimeras	52
3.4.3 Analysis of mFAS KS-AT Wild Type	55
3.4.4 Analysis of DEBS3M5 KS-AT Wild Type	62
3.4.5 Correlation of Experimental and <i>in silico</i> Data	65
3.4.6 Generation of mFAS/FabD KS-AT Chimeras	71
3.4.7 Generation of mFAS/DEBS3M6 KS-AT Chimeras	73
3.4.8 Analysis of mFAS/DEBS3M6 KS-AT Chimeras	75

4 Discussion	79
4.1 AT Activity Assay	79
4.2 Kinetic Analysis of Acyl Transferases	81
4.3 Implications of Kinetic Properties on the AT Reaction Mechanism	83
4.4 Implications of Kinetic Properties on FAS/PKS Function and Evolution	84
4.5 Impact of Interface Mutations on Transacylation Kinetics	85
4.6 Implications of Interface Mutation Study on PKS Engineering	87
4.7 Towards the Generation of a Functional Chimeric mFAS Module	88
4.8 Experimental Outlook: Implementation of Chimeras in mFAS Modules	96
4.9 Strategies for Creating Functional Chimeras	99
4.10 Bioinformatics Tools in Chimeric Protein Design	102
4.11 Conclusion and Outlook	103
5 Experimental Procedures	106
5.1 Material and Methods	106
5.1.1 Molecular Cloning	106
5.1.2 Protein Expression	107
5.1.3 Purification of AT Constructs and mFAS Modules	107
5.1.4 Purification of ACPs	108
5.1.5 Protein Concentration	108
5.1.6 SDS-PAGE Analysis	111
5.1.7 Mass Spectrometric Analysis of ACPs	112
5.1.8 CoA-488 Assay	112
5.1.9 HPLC-SEC Analysis	112
5.1.10 Thermal Shift Assay	113
5.1.11 AT Activity Assay	113
5.1.12 TAL Assay	114
5.1.13 Design of DEBS3M5 AT:ACP Interface Mutations	115
5.1.14 Bioinformatics Tools	115
5.2 Synthesis of 2-Methylbutyryl-CoA	117
References	118
Appendix	140
Statement of Personal Contributions	206
Eidesstattliche Erklärung	212
Lebenslauf	213
Danksagung	214

List of Tables

3.1	Hydrolysis parameters of FAS and PKS ATs.	36
3.2	Transacylation parameters of FAS and PKS ATs.	38
3.3	Kinetic parameters of DEBS M5 AT:ACP interface mutants.	43
3.4	Expression yields of chimeric mFAS with DEBS AT5 and RAPS AT14.	50
3.5	Expression yields of mFAS/DEBS M5 KS-AT chimeras.	54
3.6	Results of bioinformatic analysis of mFAS KS-AT.	57
3.7	Expression yields of mFAS KS-AT mutants.	58
3.8	Results of bioinformatic analysis of DEBS M5 KS-AT.	63
3.9	Expression yields of DEBS M5 KS-AT mutants.	64
3.10	Expression yields of mFAS/FabD KS-AT chimeras.	71
3.11	Expression yields of mFAS/DEBS M6 KS-AT chimeras.	73
3.12	Average expression yields of mFAS/DEBS M6 KS-AT chimeras.	75
3.13	Average expression yields of mFAS and DEBS M6 ACPs.	75
3.14	Transacylation parameters of DEBS M6 wild type and chimera.	78
5.1	Absorbance of wild type proteins.	109
5.2	Absorbance of DEBS M5 AT interface mutants.	109
5.3	Absorbance of mFAS KS-AT mutants.	110
5.4	Absorbance of DEBS M5 KS-AT mutants.	110
5.5	Absorbance of chimeras.	111
S1	Expression yields of FAS and PKS ATs.	164
S2	TSA of FAS and PKS ATs and ACPs.	164
S3	Expression yields of FAS and PKS ACPs.	165
S4	Mass spectrometric analysis of FAS and PKS ACPs.	166
S5	Initial substrate screening of FAS and PKS ATs.	167
S6	Hydrolysis parameters of FAS and PKS ATs after transacylation.	168
S7	Expression yields of DEBS M5 AT:ACP interface mutants.	169
S8	TSA of DEBS M5 AT:ACP interface mutants.	169
S9	Mass spectrometric analysis of mFAS and DEBS M6 ACPs.	170
S10	Transacylation parameters of DEBS M6 wild type and chimera (replicates). . .	171
S11	NADH calibration factor over time.	172

List of Figures

2.1	Scheme of the different FAS types.	4
2.2	Scheme of the different type I PKSs.	5
2.3	Biosynthesis of FASs and PKSs.	6
2.4	Phylogenetic relationship between FASs and PKSs.	8
2.5	Structure and domain organization of type I FAS systems.	11
2.6	Structural models for modular PKSs.	12
2.7	Structure of the ACP and AT.	14
2.8	Reaction catalyzed by the AT.	15
2.9	Reaction catalyzed by the KS.	19
2.10	Idea of the FAS-based assembly line.	23
3.1	Scheme of AT-mediated reactions.	25
3.2	Phylogenetic tree of AT domains.	31
3.3	Sequence alignment of AT domains.	31
3.4	Transacylation and hydrolysis screening of FAS and PKS ATs.	34
3.5	Transacylation parameters of FAS and PKS ATs.	40
3.6	Modeled DEBS M5 AT:ACP interface.	41
3.7	Transacylation and hydrolysis screening of DEBS M5 AT:ACP interface mutants.	42
3.8	Transacylation parameters of DEBS M5 AT:ACP interface mutants.	44
3.9	Transacylation transition state energy of DEBS M5 AT:ACP interface mutants.	45
3.10	Structural similarity of mFAS and DEBS M5 KS-AT.	47
3.11	Domain boundaries for mFAS/DEBS AT5 and RAPS AT14 chimeras.	49
3.12	Analysis of chimeric mFAS with DEBS AT5 and RAPS AT14.	51
3.13	Design of mFAS/DEBS M5 KS-AT chimeras.	53
3.14	Analysis of mFAS/DEBS M5 KS-AT chimeras.	55
3.15	Mutation sites in mFAS KS-AT.	56
3.16	SDS-PAGE analysis of mFAS KS-AT mutants.	59
3.17	TSA of mFAS KS-AT mutants.	59
3.18	HPLC-SEC of mFAS KS-AT mutants.	60
3.19	Interfacial salt bridge in mFAS KS-AT.	61
3.20	Mutation sites in DEBS M5 KS-AT.	62
3.21	Analysis of DEBS M5 KS-AT mutants.	64
3.22	Correlation of Rosetta predictions with mFAS KS-AT yields.	66
3.23	Correlation of Rosetta predictions with mFAS KS-AT melting temperatures.	67
3.24	Correlation of Rosetta predictions with mFAS KS-AT oligomeric states.	68
3.25	Correlation of Rosetta predictions with DEBS M5 KS-AT yields.	70
3.26	Domain boundaries for mFAS/FabD KS-AT chimeras.	71
3.27	Analysis of mFAS/FabD KS-AT chimeras.	72
3.28	Domain boundaries for mFAS/DEBS AT6 KS-AT chimeras.	73
3.29	Analysis of mFAS/DEBS M6 chimeric KS-AT.	74
3.30	Analysis of mFAS/DEBS M6 KS-AT chimeras and wild type DEBS M6 KS-AT.	76
3.31	Transacylation screening of mFAS/DEBS M6 KS-AT chimeras.	77

3.32	Transacylation curves of DEBS M6 KS-AT wild type and chimera.	78
4.1	Plasticity of transacylation kinetics of DEBS M5 AT:ACP interface mutants.	86
4.2	Contact potential of mFAS KS-AT wild type.	93
4.3	Contact potential of DEBS M5 KS-AT wild type.	94
4.4	Structure and contact potential of FabD.	95
4.5	Idea of assembly line with chimeric elongation module.	97
S1	SDS-PAGEs of FAS and PKS ATs.	173
S2	Size exclusion chromatograms of FAS and PKS ATs.	175
S3	TSA melting curves of FAS and PKS ATs.	177
S4	Analytical SDS-PAGEs of FAS and PKS ACPs.	177
S5	Mass spectrometric analysis of FAS and PKS ACPs.	179
S6	TSA melting curves of FAS and PKS ACPs.	182
S7	Titration curves for AT-mediated hydrolysis.	184
S8	Titration curves for AT-mediated transacylation.	187
S9	Detailed model of DEBS M5 AT:ACP interface.	188
S10	SDS-PAGEs of DEBS M5 AT:ACP interface mutants.	189
S11	Size exclusion chromatograms of DEBS M5 AT:ACP interface mutants.	191
S12	TSA melting curves of DEBS M5 AT:ACP interface mutants.	193
S13	Hydrolysis curves for DEBS M5 AT:ACP interface mutants.	194
S14	Transacylation curves for DEBS M5 AT:ACP interface mutants.	195
S15	ACP titration curve of DEBS M5 AT:ACP mutant R850E.	196
S16	Purification of chimeric mFAS with DEBS AT5.	197
S17	Correlation of bioinformatics tools predictions with mFAS KS-AT yields.	198
S18	TSA melting curves of mFAS/DEBS M6 KS-AT chimeras.	199
S19	Mass spectrometric analysis of mFAS and DEBS M6 ACPs.	200
S20	Transacylation curves for mFAS/DEBS M6 KS-AT chimera (replicates).	201
S21	Transacylation curves for DEBS M6 KS-AT (replicates).	202
S22	Electron density map of DEBS M5 AT residue R850.	203
S23	NADH calibration.	204
S24	NMR of 2-methylbutyryl-CoA.	205

List of Abbreviations

2-MB	2-methylbutyryl
Ac	acetyl
ACP	acyl carrier protein
AT	acyl transferase
AVES	avermectin polyketide synthase
BSA	bovine serum albumin
CMN-bacteria	<i>Corynebacteria</i> , <i>Mycobacteria</i> and <i>Nocardia</i>
CoA	coenzyme A
DEBS	6-deoxyerythronolide B synthase
DD	docking domain
DH	dehydratase
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
DTT	1,4-dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EcPKS	<i>Elysia chlorotica</i> polyketide synthase
EDTA	ethylenediaminetetraacetic acid
EM	electron microscopy
ER	β -enoyl reductase
FAS	fatty acid synthase
GMQE	Global model quality estimation
HPLC	high performance liquid chromatography
IPTG	isopropyl β -D-1-thiogalactopyranoside
K_M	Michaelis-Menten constant
KAL	KS-AT linker
KR	β -ketoacyl reductase
KS	β -ketoacyl synthase
LB	liquid broth
LD	linker domain
Mal	malonyl
MAS	mycocerosic acid synthase
MAT	malonyl/acetyl transferase
MBP	maltose binding protein
mFAS	murine FAS
MMal	methylmalonyl
MS	mass spectrometry

MSAS	6-methylsalicylic acid synthase
M x	module x
NAD	nicotinamide adenine dinucleotide
NADPH	nicotinamide adenine dinucleotide phosphate
NMR	nuclear magnetic resonance
Npt	4'-phosphopantetheinyl transferase
NTA	nitrilotriacetic acid
OD	optical density
PAL	post-AT linker
PCR	polymerase chain reaction
PikA	pikromycin synthase
PKS	polyketide synthase
PPT	phosphopantetheinyl transferase
Prop	propionyl
QMEAN	quality model energy analysis
QM/MM	quantum mechanics/molecular mechanics
RAPS	rapamycin synthase
REU	Rosetta energy unit
RFU	relative fluorescence unit
RSRZ	real-space R-value Z-score
SAXS	small-angle X-ray scattering
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
<i>S. erythraea</i>	<i>Saccharopolyspora erythraea</i>
Sfp	4'-phosphopantetheinyl transferase
SOC	super optimal broth with catabolite repression
SPNS	spinosyn synthase
TAL	triacetic acid lactone
TB	terrific broth
TE	thioesterase
TPP	thiamine pyrophosphate
TSA	thermal shift assay

1 Abstract

Polyketides are highly valuable natural products, which are widely used as pharmaceuticals due to their beneficial characteristics, comprising antibacterial, antifungal, immunosuppressive, and antitumor properties, among others.^[3] Their biosynthesis is performed by large and complex multiproteins, the polyketide synthases (PKSs).^[4] This study solely focuses on the class of type I PKSs, which arrange all their enzymatic domains on one or more polypeptides. Despite their high medical value, little is known about mechanistic details in PKSs.

One central domain is the acyl transferase (AT), which is present in all PKSs and channels small acyl substrates into the enzyme. More precisely, the AT loads the substrates onto the essential acyl carrier protein (ACP), which subsequently shuttles the substrates and all intermediates for condensation and modification to additional domains to build the final polyketide.

Some PKSs use their domains several times during biosynthesis and work iteratively – these are called iterative PKSs. Others feature several sets of domains, each being used only once during biosynthesis – these PKSs are called modular PKSs. All PKSs or PKS modules consist of minimum three essential domains to connect the acyl substrates.^[5] Three modifying domains are optional and can enlarge the minimal set. According to the domain composition, the acyl substrate is fully reduced, partly reduced, or not reduced at all. This variation of modifying domains accounts for the huge structural and therefore functional variety of polyketides.

Even though the structure of fatty acids is not exactly reminiscent of polyketides, their biosynthetic pathways are closely related. Fatty acid biosynthesis is carried out by fatty acid synthases (FASs), which share many similarities with PKSs.^[3,6] Both megasynthases feature the same domains, performing the same reactions to connect and modify small acyl substrates. In contrast to PKSs, FASs always contain one full set of modifying domains which is used iteratively, leading to fully reduced fatty acids.

The present thesis extensively analyzes the AT of different PKSs in its substrate selectivity, AT-ACP domain-domain interaction, and enzymatic kinetic properties. The following key findings are revealed through comparison: 1.) ATs of PKSs appear slower than the ones of FASs, which may reflect the different scopes of biosynthetic pathways. Fatty acids as essential compounds in all organisms are needed in high amounts for physiological functions, whereas polyketides as secondary metabolites only require basal concentrations to take effect. 2.) The slower ATs from modular PKSs do not load non-native substrates even in absence of the native substrates. This is different to the faster ATs from iterative PKSs and FASs, which indicates high substrate

specificity solely for the ATs from modular PKSs and emphasizes their role as gatekeepers in polyketide synthesis. 3.) The substrate selectivity can emerge in either the first or the second step of the AT-mediated ACP loading and is not assured by a hydrolytic proofreading function.

Moreover, a mutational study on the AT-ACP interaction in the modular PKS 6-deoxyerythronolide B synthase (DEBS) shows that single surface point mutations can influence AT-mediated reactions in a complex manner. Data reveals high enzyme kinetic plasticity of the AT-ACP interaction, which was also recently demonstrated for the interaction in a type II FAS.^[14]

Based on these findings, the mammalian FAS is engineered towards a modular PKS-like assembly line with the long-term goal to rationally synthesize new products. Basically, three important aspects need to be considered: 1.) AT's loading needs to be splitted in specific loading of a priming substrate by a priming AT and in specific loading of an elongation substrate by an elongation AT. 2.) FAS-based elongation modules need to be designed with varying domain compositions for introducing functional groups in the product. 3.) Covalent and non-covalent linkers need to be designed for connection of priming and elongation modules.

This study focuses on the first aspect, splitting loading of priming and elongation substrates. An elongation substrate-specific AT is installed in the mammalian FAS via domain swapping. Since ATs from modular PKSs were proven to be substrate specific, these are used to exchange the mammalian FAS AT. This work demonstrates that it is extremely challenging to create stable and functional chimeras, but first essential steps are taken. Proper domain boundaries for AT swapping are established and a stable chimera with 70 % wild type AT activity is created. However, this chimera is only of limited value for application in an elongation module due to the intrinsic slow turnover rate of the wild type AT. Using another PKS AT, a stable elongation module is designed and analyzed in its activity in combination with a priming module. These experiments demonstrate that the loading of priming substrates are successfully suppressed in the elongation module, but nonetheless only minor turnover rates are detected in the assembly line. This can have various causes: the AT domain itself, intra-, or intermodular communication problems. Further experiments on subconstructs, including characterization of the wild type, are needed to identify the real problem.

Overall, the present thesis contributes significantly to the mechanistic understanding of AT domains in FASs and PKSs. Furthermore, although challenging, it is possible to create stable and functional chimeras via AT swapping to generate mammalian FAS-based elongation modules as shown in this thesis. However, each system behaves individually and small adaptations in domain boundaries can have a huge impact on stability and activity. Further engineering approaches should be based on detailed analysis of wild type donor ATs to be successful.

2 Introduction

2.1 Natural Products in Modern Medicine

Natural products are highly valuable compounds being key players in the treatment of human diseases.^[21] Almost two thirds of all pharmaceuticals used today involve natural products.^[2] Many drugs are natural products themselves, some are derivatives created semisynthetically, others are chemically synthesized using natural products as lead structures.


Natural products show immense chemical diversity resulting in a large variety of medically relevant properties as antitumor,^[22] immunosuppressive,^[23] and antimalarial^[24] functions.^[25] Even though natural products are widely used for treatment of various diseases, the most relevant aspect is presumably the treatment of bacterial infections.^[26–28] Those are one of the most prevalent causes of death worldwide, including developed countries.^[2,29] The excessive use of antibiotics and other drugs has resulted in resistances, particularly in hospitals, which requires the discovery and development of new pharmaceuticals.^[30] In order to generate new natural products, it is crucial to understand their natural synthesis.

During the metabolism of living entities, thousands of chemical reactions, catalyzed by enzymes, take place. The metabolic pathway of all essential compounds – such as amino acids, lipids, sugars, fatty acids – was found to be evolutionary conserved among all organisms.^[31,32] However, certain species – as plants, bacteria, and fungi – evolved distinct expanded metabolic pathways for the production of natural products, adapted to the needs of the host.^[33] Despite their chemical diversity, natural products are generated from basic building blocks – acetate, mevalonate, and shikimate. During biosynthesis, these building blocks are transformed and combined in many different ways leading to numerous classes of natural products, including polyketides, steroids, and flavonoids. Polyketides are closely related to fatty acids, both being products of the acetate pathway.^[29]

2.2 Fatty Acid and Polyketide Synthases

Fatty acids play a vital role in living organisms. They fulfill a variety of important functions as central building blocks of biological membranes and major energy source.^[11,34] They act as modulator and messenger, involving positive and negative regulation of signaling cascades.^[10,35] Thus, the fatty acid biosynthesis is essential for all organisms except archaea.^[36] It is performed by fatty acid synthases (FASs), multienzymes, which consist of a certain set of domains to catalyze the conserved chemical reactions.^[3] FASs are classified according to their domain arrangement: Type I FASs are large multifunctional polypeptides as found in fungi, animals, and some bacteria.^[7,37,38] In contrast, type II FASs consist of dissociated monofunctional proteins as found in bacteria, plants, and eukaryotic mitochondria (Fig. 2.1).^[39,40] In both FAS types, the same set of domains catalyzes all biosynthetic reactions in an iterative manner.^[3,41]

Type I FAS

Animal FAS: 
murine FAS

Type II FAS



Bacterial FAS: 
E. coli FAS
FabB FabD FabA FabI FabG AcpP
FabF FabZ FabK
FabH FabL

Figure 2.1: Scheme of the different FAS types. Type I FAS systems of fungi and animals are large multifunctional polypeptide chains, whereas type II FAS systems of bacteria and plants are individual monofunctional proteins. Abbreviations: β -ketoacyl synthase (KS), acyl transferase (AT), dehydratase (DH), β -enoyl reductase (ER), β -ketoacyl reductase (KR), acyl carrier protein (ACP), thioesterase (TE). Adapted from reference.^[42]

The same set of domains is also found in polyketide synthases (PKSs), which are responsible for the *de novo* synthesis of polyketides.^[5,43] They carry out the same chemical reactions and are classified as follows: Type I PKS systems found in bacteria, fungi, and plants are multifunctional polypeptide chains similar to type I FASs, whereas type II PKSs found in prokaryotes consist of freestanding monofunctional proteins like type II FAS systems.^[3,4] Like FASs, type I PKSs of fungi contain one set of catalytic domains which is used iteratively, whereas the majority of bacterial type I PKS systems consists of multiple sets of catalytic domains, so-called modules, each being used only one time during synthesis (Fig. 2.2).^[44] These modular, assembly line-like PKSs contain either an integrated (*cis*) or a distinct (*trans*) acyl transferase (AT) domain.^[45] Their modules are either covalently linked or bridged non-covalently by α -helical linkers, the so-called docking domains (DDs).^[46,47] Besides type I and type II PKSs, there is a third type of PKSs, which are individual iterative condensing enzymes.^[3] In contrast to type I and II PKSs, the polyketide intermediate and product are never directly attached to type III PKSs.^[48] In addition to these three types of PKSs, there are also hybrid systems found, including hybrid forms of type I/type II PKSs as well as type I/nonribosomal peptide synthases.^[49,50]

ⁱthe *Corynebacterium-Mycobacterium-Nocardia* (CMN) group of bacteria

Type I iterative PKS

Bacterial PKS: 
MSAS

Type I modular PKS

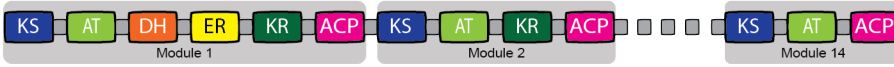
Bacterial PKS: 
RAPS

Figure 2.2: Scheme of the different type I PKSs. Type I PKS systems of fungi and some bacteria are monomeric iteratively-acting enzymes like the 6-methylsalicylic acid synthase (MSAS). The majority of bacterial type I PKSs evolved multimodular assembly line-like PKSs like the rapamycin synthase (RAPS) with either integrated (*cis*) or distinct (*trans*) AT domains. Abbreviations: β -ketoacyl synthase (KS), acyl transferase (AT), dehydratase (DH), β -enoyl reductase (ER), β -ketoacyl reductase (KR), acyl carrier protein (ACP), thioesterase (TE). Adapted from reference.^[42]

2.3 Biosynthesis of Fatty Acids and Polyketides

The megasynthases FASs and PKSs produce important products – fatty acids (most commonly palmitic (C16) and stearic acid (C18)^[41]) and polyketides, respectively. Whereas FASs always produce fully reduced fatty acids, polyketides show a large structural and therefore functional variety.^[3] Despite this large diversity, they are all made up of small building blocks, acyl extender units. Every FAS and PKS domain is responsible for one specific reaction, connecting and modifying these extender units.

As mentioned above, fatty acid synthesis is an iterative process and the catalytic cycle can be divided into three parts: initiation, chain elongation, and chain termination.^[7,10,15,34,51] The initial priming step is the loading of an acetyl (Ac) moiety, which is then successively elongated (Fig. 2.3A). In each round of chain elongation, a C2 unit, originated from a malonyl (Mal) moiety, is added to the growing acyl chain until the desired chain length is obtained. The acyl chain is then released either as free fatty acid or as coenzyme A (CoA) derivative.^[41]

The AT domain catalyzes the loading of acyl units of CoA esters onto the acyl carrier protein (ACP).^[41,52,53] This small protein is post-translationally modified with a phosphopantetheinyl prosthetic group, which binds acyl moieties covalently upon thioester formation.^[54,55] After AT-mediated loading, the ACP shuttles the acyl substrates and intermediates to all enzymatic stations throughout the whole catalytic cycle.^[39,56,57] This central role of the ACP domain highlights the importance of domain-domain interactions during biosynthesis, ensuring specificity allowing for catalytic action on the acyl chain and simultaneously facilitating high turnover rates.^[41]

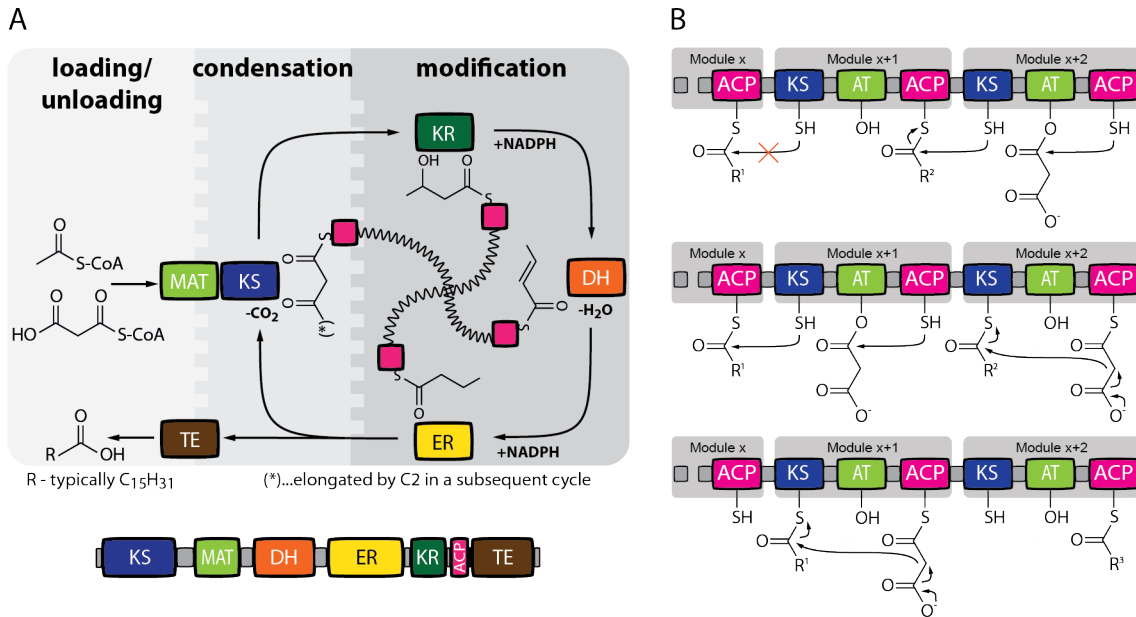


Figure 2.3: (A) Schematic cycle of fatty acid synthesis catalyzed by the mammalian FAS. Fatty acids are stepwise built from priming substrate Ac and elongation substrates Mal. After condensation, the β -ketoacyl moiety is fully reduced, before the next elongation substrate is added to the chain. After seven cycles, the final product is released by the TE domain. (B) Turnstile mechanism of modular PKSs. The KS domain cannot be acylated by the next polyketide chain, until the current product of chain elongation and subsequent modification is translocated to the downstream module (shown) or released by the TE (not shown here). Adapted from reference.^[6]

The loaded substrates are connected via C-C bond formation through decarboxylative Claisen condensation, catalyzed by the β -ketoacyl synthase (KS) yielding a β -ketoacyl intermediate.^[41,58–60] Subsequently, this intermediate is successively modified by an NADPH-dependent β -ketoacyl reductase (KR),^[61–63] a dehydratase (DH)^[64–66] and an NADPH-dependent β -enoyl reductase (ER)^[67] resulting in a fully saturated acyl intermediate. After several rounds of chain elongation, the product is released. In fungal and CMN-bacterial type I systems, AT mediates the transfer of the acyl moiety from ACP to CoA.^[68] Animal type I FASs and type II plant systems use a TE domain to release the free fatty acid.^[69,70] Bacterial and mitochondrial type II systems forward the acyl-ACP to other biosynthetic pathways.^[71,72]

Comparing FAS and PKS systems, a main difference obviously lies in the catalytic efficiency of the biosynthesis: FASs produce fatty acids in less than a second,^[15] whereas PKSs are much slower. The 6-deoxyerythronolide B synthase (DEBS) for example takes around two minutes to produce 6-deoxyerythronolide B.^[16] This difference in turnover rates might be a result of the respective function in the organisms. Fatty acids are essential for cellular processes, whereas polyketides are secondary metabolites.

The overall reaction series catalyzed by type I PKSs is very similar to FAS systems.^[3] Substrate loading, chain elongation, and chain termination of both megasynthases are related and all reaction intermediates remain covalently attached. The priming substrate for PKSs is typically either an Ac or propionyl (Prop) moiety and the elongation substrate is a Mal or methylmalonyl (MMal) moiety.^[5] In iterative type I PKSs, one set of catalytic domains is used repetitively like in FAS systems, whereas in modular PKSs each module contains one set of catalytic domains and is used only once, being responsible for the incorporation of one extender unit. Each set of domains in PKSs comprises three essential catalytic domains connecting the extender units: KS, AT, and ACP. The three domains (KR, DH, ER) responsible for further modification of the growing acyl chain are optional and can enlarge this minimal set. Whether the extender unit is fully or partly reduced or not reduced at all, depends on the specific domain composition. Based on the degree of reductive behavior, fungal type I PKSs are further divided into highly reducing (HR-), partially reducing (PR-), and non-reducing (NR-) PKSs.^[5] This diverse domain arrangement accounts for the huge structural and functional variety of compounds found within the class of PKSs.^[73] FASs, on the other hand, always use a full set of catalytic domains and, hence, can be seen as fully reducing PKSs within this nomenclature.^[74]

Most modular PKSs with integrated AT domains, so-called *cis*-AT PKSs, follow the principle of collinearity: The modules and domains operate in the same order in which they are encoded on the DNA.^[18] This property allows for conclusions on the domain composition by analyzing the product and vice versa. Additionally to chain elongation, an intramodular process present in all FASs and PKSs, modular PKSs involve also the intermodular process of chain translocation. The ACP of one module must pass on its β -ketoacyl product to the downstream module before another extender unit can be loaded onto the ACP, which the KS domain can act on. While it remains unclear how exactly the back-transfer of one module's polyketide intermediate to its own KS is prevented, one could assume kinetic control by coupling of catalytic steps, leading to a vectorial synthesis.^[75] An alternative scenario was proposed, in which chain elongation is energetically coupled to chain translocation via a turnstile mechanism: The β -ketoacyl-ACP product of one module must be passed on to the downstream module before the KS domain can act on the next upstream β -ketoacyl-ACP product (Fig. 2.3B).^[76] This mechanism avoids iterative chain elongation and is assumed to be evolutionary acquired.^[6]

2.4 Evolution of Fatty Acid and Polyketide Synthases

Fatty acid synthesis is an essential biosynthetic pathway among all groups of organisms and therefore likely to have evolved early. Based on the remarkable similarities between FASs and PKSs, it is reasonable to assume shared evolutionary processes. Phylogenetic analysis has revealed a long joint evolution of these two biosynthetic pathways.^[6,42,77] In the phylogenetic analysis of the highly conserved KS domain of FASs and PKSs published by H. Jenke-Kodama and co-workers, the archaeobacterial KS homologs and the bacterial KS of FabH type were used as outgroup (Fig. 2.4).^[42] The FabF type KS of bacterial type II FASs was found to evolve in parallel to the KS of bacterial type II PKSs. The two type I FAS groups of fungi and CMN bacteria were found to be evolutionary close and to predate the third FAS I group of animals, to which they are apparently only distantly related. As oldest type I PKSs, some bacterial iterative systems were found which evolved in parallel to glycolipid synthases. Modular type I PKSs then evolved in bacteria either with *cis*- or *trans*-AT domains in parallel to iterative type I PKSs of fungi and bacteria and the animal type I FASs.^[42]

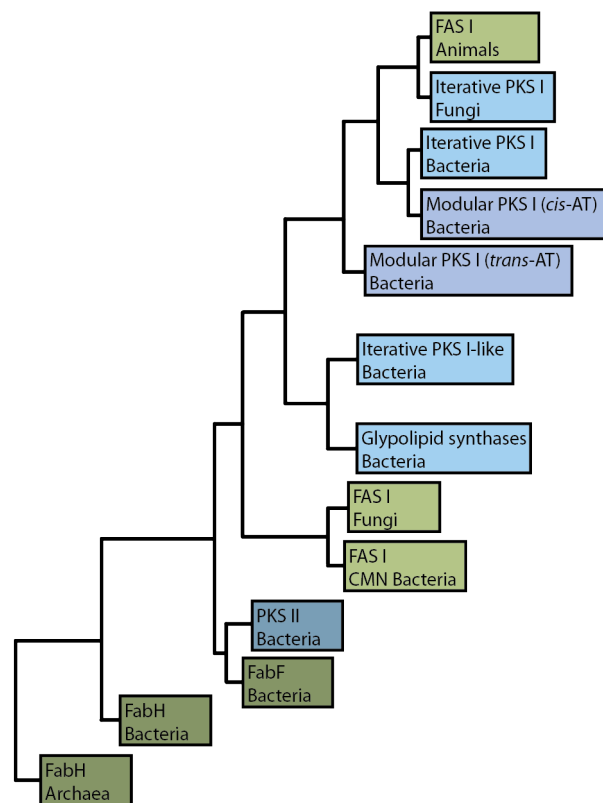


Figure 2.4: Schematic representation of the phylogenetic relationship between FASs and PKSs based on the alignment of KS domains, inferred by Bayesian statistics. Type II FAS and PKS systems have predated the multidomain and multimodular type I systems. Type I systems shown in light, type II systems shown in dark colors. FAS systems shown in green, PKS systems shown in blue. Adapted from references.^[6,42,78]

To date, there is scientific consensus on the evolution of the multifunctional type I systems from the distinct proteins of type II systems.^[6,34,42] The dominant doctrine in this context is the gene fusion hypothesis, which states that duplication and successive fusion of primordial genes occurred early in the evolution resulting in type I systems prior to the divergence of prokaryotes and eukaryotes.^[6,34] An alternative scenario is the evolution of multifunctional type I FAS systems first in eukaryotes and later via horizontal gene transfer (HGT) in prokaryotes.

However, there are two separate developments involved in the evolution of type I FASs: Fungal and CMN-bacterial type I FASs evolved together, whereas the animal type I FAS evolved in parallel to type I PKS systems.^[41] The latter systems share a common ancestor, an iterative fully methylating PKS.^[41,42] The evolution of PKSs is even more complex. Although there is evidence that all PKSs are evolutionary related, the precise evolutionary relationships between different PKSs remain unclear.^[6] However, the outstanding evolution of assembly line-like systems is only found in bacterial type I PKSs. Their evolution from iterative PKSs involves certainly several genetic processes like mutation, gene fusion, and gene duplication to yield multimodular PKSs and also further recombination, gene conversion, and mutation, leading to today's diversity of assembly line-like PKSs.^[79] Furthermore, it is unclear whether *cis*- and *trans*-AT PKSs evolved together or independently.^[6] However, phylogenetic analysis of KS domains yields distinct clades of *cis*- and *trans*-AT PKSs.^[80] *cis*-AT domains from bacterial type I PKSs were analyzed further to determine their phylogenetic relationship.^[42] They were found to form two major clades, grouped by their substrate specificity: One earlier clade includes four groups of Mal-CoA-activating ATs, the other, higher evolved clade consists of one AT-group activating Mal-CoA besides three AT-groups activating MMal-CoA and rarer substrates. Thus, it can be deduced that bacterial PKS I ATs first evolved from the Mal-CoA-activating ancestor involved in fatty acid synthesis and later functional specialization led to the evolution of the second clade with modified/expanded substrate specificity.^[42]

2.5 Structural Comparison of Fatty Acid and Polyketide Synthases

Distinct evolutionary pathways of different types of FASs and PKSs are also well reflected in structural differences. It took several decades to solve large and complex structures like type I FASs from initial low-resolution structures obtained by electron microscopy (EM)^[81] to high-resolution structures determined by X-ray crystallography.^[38] Today, the structures of fungal,^[82,83] CMN-bacterial,^[84] and animal type I FASs^[85] confirm their different evolutionary development (Fig. 2.5). As discussed above, the first two types evolved together and share many structural features: Both show a barrel-shaped, hexameric-likeⁱⁱ architecture with a rigid central wheel and two domes forming the reaction chambers for biosynthesis. To form such a complex structure, about 50 % and 35 % of the polypeptide are used as scaffolding elements in fungal and bacterial type I FASs, respectively.^[86] Although the CMN-bacterial type is encoded on single genes and does not include a C-terminally attached phosphopantetheinyl transferase (PPT), the overall architecture is very similarⁱⁱⁱ. On the contrary, animal type I FASs, which originated from another evolutionary pathway than fungal and CMN-bacterial types, show an open architecture forming highly flexible, X-shaped homodimers.^[41,69] Due to the high conformational flexibility, the ACP and TE domains remain unresolved in the crystal structure from Maier *et al.*^[85] The homodimers contain a condensing and a modifying wing connected by a waist region, the condensing part consisting of KS and AT domains, and the modifying wing carrying the three domains KR, DH, and ER responsible for modification. In contrast to the other two type I FAS systems, all enzymatic domains are exposed to the environment. The X-shaped dimer forms two reaction chambers with high distances between the single domains, making large-scale conformational changes necessary for efficient biosynthesis.^[8]

For PKSs, especially the modular PKSs, structural elucidation is even more challenging. Due to their large size, isolation of complete modules, let alone full-length proteins, is extremely difficult. Crystallization efforts are hampered by their inherent high flexibility.^[87] Since the first high-resolution structure of a PKS domain in 2001,^[88] it took more than a decade to gain high-resolution structural information on all domains from *cis*-AT PKSs.^[89] Structures of single domains (KR,^[90,91] ER,^[92] DH,^[93,94] TE,^[88,95] ACP,^[96–98]) or didomains (KS-AT^[99,100], KR-ER^[101]) from modular PKSs overall share a remarkable similarity with their counterparts in the mammalian FAS.^[85,102,103] Due to the high similarity and the lack of structural infor-

ⁱⁱBacterial type I FASs form homohexameric barrels, whereas fungal type I FASs form heterododecameric barrels consisting of two polypeptide chains forming an $\alpha_6\beta_6$ complex.

ⁱⁱⁱBacterial type I FASs are relatively conserved, whereas fungal systems show further diversification resulting from a late evolutionary development:^[41,86] Fungal types show variations on gene level being either encoded on a single gene or on genes splitted at different positions. Nevertheless, the overall architecture and assembly pathway of these variants is very similar.

mation on larger PKS units, modular PKSs were often imagined as several mammalian FAS structures lined up with its segregation into the condensing and the modifying wing (Fig. 2.6: mammalian FAS).^[85,89,91] However, some structural studies revealed certain differences, questioning the role of the mammalian FAS structure as universal model.^[85,87,88,93,101,104]

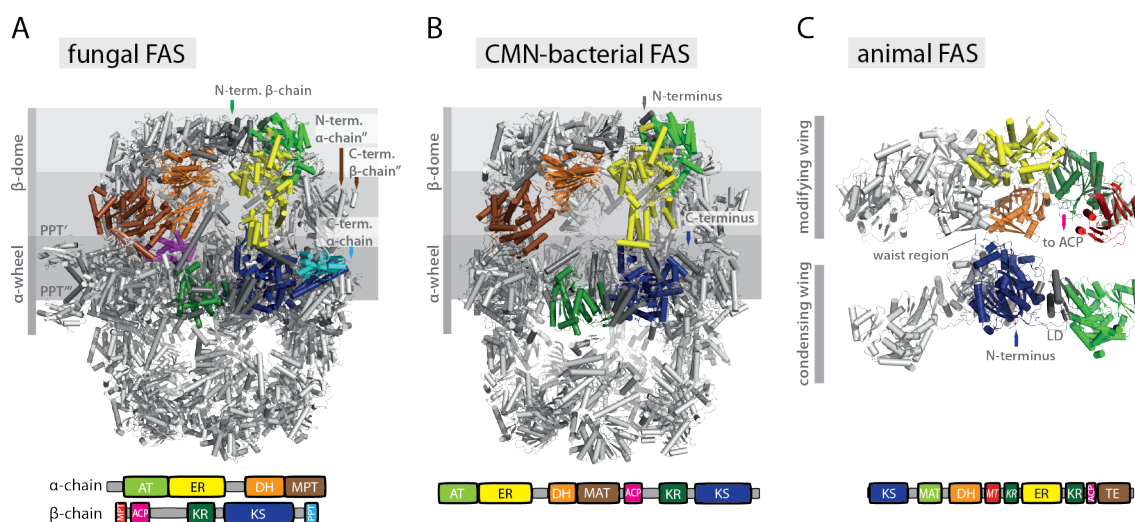


Figure 2.5: Structure and domain organization of fungal, CMN-bacterial, and mammalian type I FAS systems. (A) X-ray crystallographic model of baker's yeast FAS (PDB: 3HMJ^[83]). (B) X-ray crystallographic model of *M. smegmatis* FAS (PDB: 4V8L^[84]). (C) X-ray crystallographic model of porcine FAS (PDB: 2VZ8^[85]). Adapted from reference.^[41]

Over time, different models for the architecture in modular PKSs emerged (Fig. 2.6). The DEBS M3 and pikromycin synthase (PikA) M5 model are based on low resolution small-angle X-ray scattering (SAXS)^[105] and medium resolution cryo-EM data,^[106,107] respectively. These two models differ most in the relative orientation of KS and AT, leading to completely different architectures. For the DEBS M3 model, the KS-AT has the same orientation as in KS-AT didomain structures solved before,^[85,99,100,103] but was treated as rigid body during modeling and refinement based on KS-AT data, so its configuration was invariable. Even though the shape of the rigid KS-AT body fits well with the experimental data set, it would be interesting to treat both as individual domains inspecting the actual domain orientation.^[87] In the PikA M5 model, the AT domain is positioned below the KS and rotated by approximately 120° compared to other KS-AT structures, leading to an overall arch-shaped dimer with one single central reaction chamber.^[87] This module showed major discrepancies to previous models and solved structures, speaking against its role as universal model for modular PKSs.^[108] It might as well be possible that the structure represents an inactive conformation or an artifact caused by sample handling and preparation of the cryo-EM sample, which can be challenging for macromolecular complexes.^[109]

Introduction

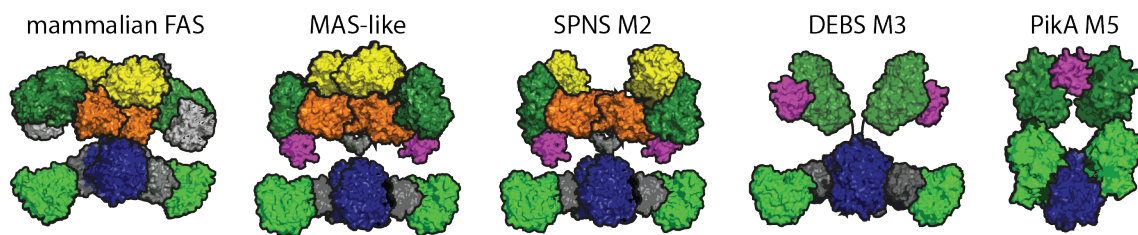


Figure 2.6: Structural models for modular PKSs. Surface representation of the X-ray structure from the mammalian FAS (PDB: 2VZ8^[85]), and from a MAS-like PKS assembled from condensing and modifying wing (PDB: 5BP1 and 5BP4^[110]), of the SPNS M2 model^[101] assembled from PDBs 5BP1 (KS-AT), 5BP4 (DH), and 3SLK (ER-KR),^[101] of the SAXS data from the partially reducing module DEBS M3,^[105] and of the cryo-EM data from the partially reducing module PikAIII M5.^[106] Adapted from reference.^[75]

Two other models were created by assembly of high-resolution crystal structures of the mycocerosic acid synthase (MAS)-like PKS from *Mycobacterium tuberculosis*^[110] and of the spinosyn synthase (SPNS)^[101] to full modules. Even though the MAS-like PKS is a fully reducing iterative PKS, it is more similar to modular PKSs (27-35 % sequence identity) than to fungal iterative PKSs (20-22 % sequence identity) and FASs (19 % sequence identity).^[38] Both models are overall similar and the overall fold is related to the mammalian FAS, but they differ especially in the domains DH and ER. The dimeric DH domain in both modules is rather linear and resembles the dimeric DH domain of modular PKSs (interdomain angle of 167-203 °),^[93,94,111-114] whereas the DH dimer in the mammalian FAS is V-shaped.^[85] Looking at the ER, this domain is dimeric in the MAS-like PKS^[110] like in the mammalian FAS^[85] and some modular PKSs.^[92,115] In contrast, the ER from the SPNS and the majority of modular PKSs were found to be monomeric.^[101,115-117] The dimerization motif in ER domains and the ER-KR linker from modular PKSs are significantly shorter than those from iterative PKSs hindering ER dimerization.^[115] The relevance of ER's oligomeric state in modular PKSs remains unclear, but the MAS-like PKS and the SPNS structures together may represent a relevant model for modular PKSs' scaffold.

Most recently, a high-resolution cryo-EM structure of the iterative lovastatin nonaketide synthase LovB was published (PDB: 7CPX).^[118] Like the majority of the other models, it reveals an overall X-shaped dimeric architecture and strongly resembles the mammalian FAS structure with V-shaped dimeric DH domains. The KS-AT forms a dimer with a rather linear conformation, which is slightly rotated compared to the mammalian FAS structure. This is most likely caused by differences in the linker domain (LD), since single domains KS and AT reveal a high structural similarity to domains in the mammalian FAS and DEBS M3. This structure supports the significance of the mammalian FAS structure in the context of type I PKS structures.

While all these structures give important insights into the architecture of standalone PKS modules, structural information beyond that is even rarer. Indeed, several structures of non-covalent connections between modules, the DDs, were solved^[46,47,119–121] and SAXS analysis gave low-resolution structures of a bimodular construct.^[105] However, high-resolution structures from full modules^{iv} and multimodular constructs are still missing. Nevertheless, every PKS structure solved complements the picture of PKSs even if each can just be seen as one possible out of many conformations. The same also applies to protein-protein interactions which are the key to catalytic activity in megasynthases. Central domains and their interaction will be discussed in the next section.

2.6 Key Players in Fatty Acid and Polyketide Synthesis

KS, AT, and ACP are essential domains/enzymes and part of every FAS and PKS system or module.^[3] As mentioned above, the AT domain loads the substrates into the system onto the ACP domain.^[9,52] The ACP domain shuttles the substrate and the intermediate towards all catalytic sites.^[57] The KS domain is responsible for the condensation of the substrates to the final product.^[58] In the following, enzymatic reactions and properties of these three domains are illuminated.

2.6.1 The Acyl Carrier Protein

The ACP is a small, highly dynamic protein.^[41,57] It shuttles the substrates and all intermediates throughout the whole biosynthesis of fatty acids and polyketides, interacting with every catalytic enzyme/domain.^[57] This protein enables the biosynthesis in the first place. Overall, ACPs from different systems share a common fold.^[3]

Type I and II ACP domains from FASs were found to form helical bundles consisting of four α -helices stabilized by inter-helical hydrophobic interactions (Fig. 2.7A).^[3,122,123] Predicted and solved structures for PKSs' ACPs show a similar fold.^[124–126] The inactive *apo*-form of the protein needs to be post-translationally modified: The conserved active site serine (S38 in rat FAS) is phosphopantetheinylated, leading to an active thiol group.^[54,55] This prosthetic group with its length and high conformational flexibility is crucial for ACP's role in substrate and intermediate shuttling.^[8]

^{iv}LovB is an iterative HR-PKS, which contains a pseudo-ER, but interacts with the *trans*-ER LovC. The complex structure was also solved by Wang *et al.* (PDB: 7XPY^[118]).

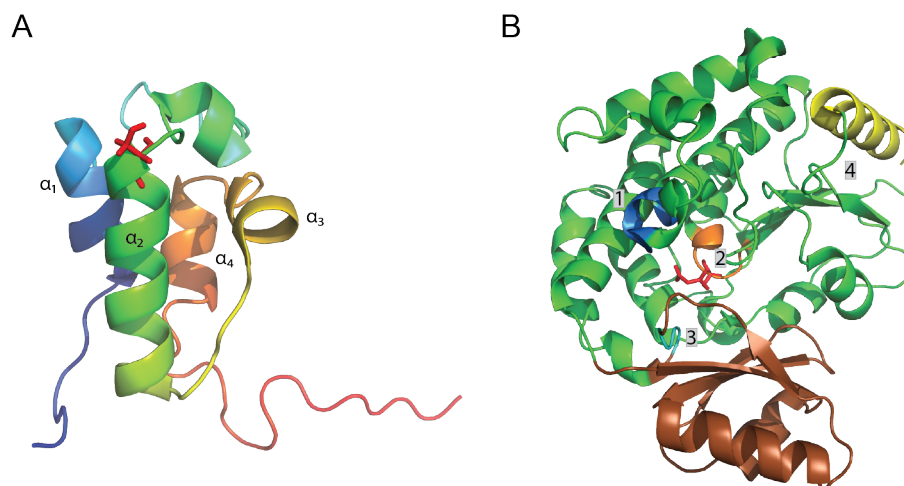


Figure 2.7: (A) Structure of the *apo*-ACP domain from rat FAS (PDB: 2PNG^[102]) as cartoon representation. Coloring from blue to red corresponds to N- and C-termini, respectively. ACPs consist of four helices α_1 to α_4 . The active site S38 is highlighted in red sticks. (B) Structure of the AT domain from murine FAS (mFAS, PDB: 5MY0^[9]) as cartoon representation. The core domain and the subdomain are shown in green and brown, respectively. The active site S581 is highlighted in red sticks. Motifs 1 to 4 for substrate selectivity are highlighted in blue, orange, cyan, and yellow.

2.6.2 The Acyl Transferase

The AT is responsible for the loading of substrates into the system.^[52] This domain/enzyme selects the correct substrates for fatty acid and polyketide biosynthesis out of a vast variety of compounds within the cell. In animal type I FASs, there is only one AT responsible for loading both, priming and elongation substrates.^[3] Modular PKS systems, on the other hand, contain specialized ATs which discriminate between the two substrates.

More specifically, ATs catalyze the acyl transfer from CoA to the ACP. This reaction is termed transacylation and follows a ping-pong bi-bi catalytic mechanism (Fig. 2.8) catalyzed by a Ser-His dyad.^[52,127–129] In the first step, the ping step, the conserved active site serine (Ser581 in murine FAS (mFAS)) attacks the carbonyl group of the substrate acyl-CoA, leading to an AT-bound acyl and releasing free CoA.^[41] In the pong step, the ACP attacks the ester carbonyl, resulting in acylated ACP and the regenerated AT domain/enzyme. The nucleophilicity of the active serine is increased by a conserved histidine (His683 in mFAS) via hydrogen bonding.^[41] Elongation substrates like Mal and MMal are directly interacting with a conserved arginine (Arg606 in mFAS) within the AT, facilitating its enclosure into the binding pocket. The tetrahedral intermediates of both steps are stabilized by an oxyanion hole at the active site, formed by two backbone amides (Met499 and Leu582 in mFAS).^[130] As side reaction, AT-mediated hydrolysis can occur after the ping step.^[131–133]

Using steady state assumptions, the reaction velocity is given by

$$v = \frac{v_{\max} [X\text{-CoA}] [ACP]}{K_m^{\text{ACP}} [X\text{-CoA}] + [X\text{-CoA}] [ACP] + K_m^{\text{X-CoA}} [ACP]},$$

with $v_{\max} = k_{\text{cat}} [\text{AT}]_0$.^[129]

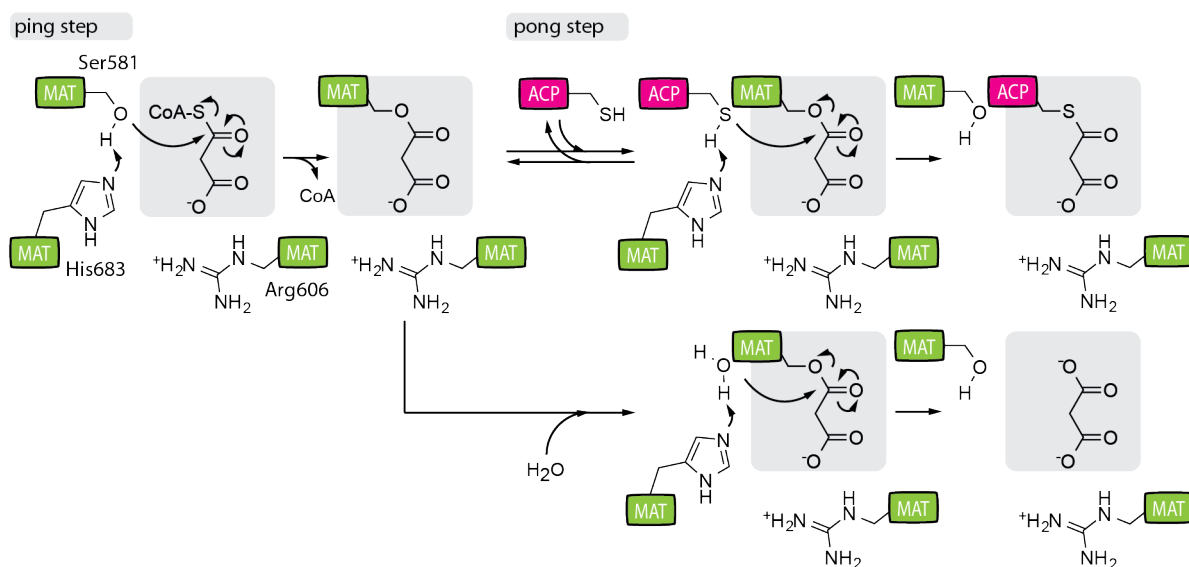


Figure 2.8: Ping-pong bi-bi reaction catalyzed by the MAT domain from mFAS: During transacylation, the acyl moiety X is first loaded onto the AT, then transferred to the ACP domain. As side reaction AT-mediated hydrolysis occurs. Adapted from reference.^[41]

Recently, P. Paiva and co-workers analyzed the reactions mediated by the human MAT in mechanistic detail using quantum mechanics/molecular mechanics (QM/MM) calculations.^[134,135] They showed that both steps, the AT-loading ping and the ACP-loading pong step, occur in two sequential steps. The first step consists of the concerted deprotonation of Ser581 by His683 and the nucleophilic attack of the substrate thioester carbonyl carbon by the deprotonated Ser581, leading to a tetrahedral intermediate.^[134] Like proposed before,^[130] they found that this intermediate is stabilized by the interaction of the negatively charged carbonyl oxygen with the oxyanion hole formed by the backbone amides of Met499 and Leu582. The second step is the breakdown of the intermediate releasing free CoA which deprotonates the protonated His683, resulting in the acyl-loaded MAT. The third step, the first step of the pong step, consists of the concerted deprotonation of the *holo*-ACP thiol group by His683 and the nucleophilic attack of the MAT-bound substrate carbonyl carbon by the deprotonated thiol group.^[135] The resulting tetrahedral intermediate is stabilized by the oxyanion hole. The fourth step is the breakdown of the intermediate and the deprotonation of the protonated His683 by the MAT active site Ser581 yielding acyl-loaded ACP. The computational analysis identified the first step as rate-limiting for the full MAT-mediated reaction.^[135]

Type I and II ATs from FASs and PKSs show similar folds.^[3,41,69,99,136,137] They comprise a core domain with an α/β -hydrolase-like structure and a smaller subdomain with a ferredoxin-like structure (Fig. 2.7B).^[3] Even though the overall structure is very similar, there are differences in the primary sequence, which affect the substrate selectivity of ATs: Animal type I FAS AT domains show a broad substrate tolerance for priming and elongation substrates.^[9,138] Likewise, priming AT domains from type I modular PKSs are strikingly substrate tolerant.^[3,139,140] In contrast, elongation ATs exhibit a remarkable substrate selectivity and act as filter in the polyketide biosynthesis.^[133,141–143] This selectivity can be predicted from certain amino acid motifs and regions.^[3]

1. The “RVDVVQ” motif: This motif is located 30 amino acids upstream from the active serine and indicates substrate specificity for MMal. Later on, it was extended to [RQSED]V[DE]VVQ, in which the first and third amino acid can vary between R, Q, S, E, D and D, E, respectively. The motif for Mal specificity is ZTX\$[AT][QE], in which Z is any hydrophilic residue, \$ is any aromatic amino acid, and the fifth and sixth amino acid vary between A,T and Q, E, respectively.^[3,144,145]
2. The “GHSXG” motif: This motif is found at the so-called “nucleophilic elbow” between a β -sheet and an α -helix.^[146] The substrate selectivity depends on the amino acid X next to the active serine. Mal-specific ATs are characterized by a branched hydrophobic amino acid like valine or isoleucine at this position. Other ATs feature a less bulky residue like glutamine or methionine due to steric restrictions of larger substrates.^[3,144] In FAS ATs which load Mal discriminating against MMal, X is leucine or valine.^[3]
3. The “YASH” motif: This motif is located 100 amino acids downstream from the active serine and is characteristic for MMal-specific ATs.^[147,148] Mal-specific ATs are indicated by a “HAFH” motif. These motifs were refined to [YVW]ASH and [HTVY]AFH for MMal and Mal specificity, respectively.^[3] In FAS, this motif is [MILV]AFH.
4. The C-terminal region: This region is part of an α - β - α motif. The helix is part of the binding pocket entrance, but no specific sequence is defined for this motif.^[3] A domain swapping study showed its role in determining the substrate specificity most likely due to its involvement in the binding pocket.^[3,149]

Overall, priming ATs differ from elongation ATs mostly in the third motif. Additionally, the characteristic site arginine found in Mal- and MMal-transferring elongation ATs which interacts with the carboxylate of the elongation substrate is usually missing in priming ATs.^[3] For DEBS, the priming AT shows the following amino acid sequences: 1. RVEVVQ, 2. GHSIG, and 3. MAAH. The C-terminal region is well conserved.

2.6.3 The β -Ketoacyl Synthase

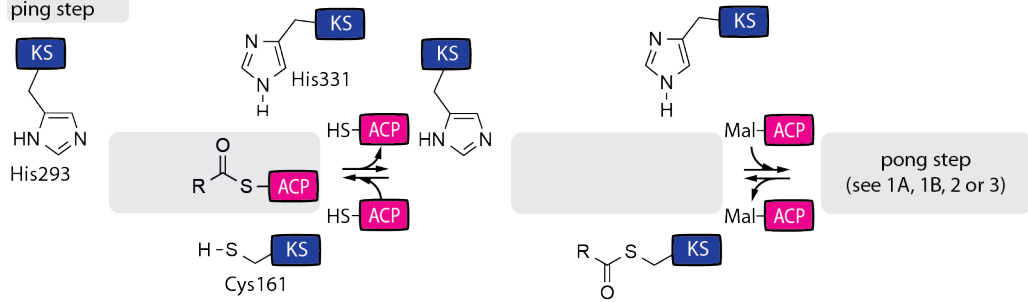
The KS domain catalyzes the key step in fatty acid and polyketide biosynthesis: the formation of carbon-carbon bonds.^[150] This domain enables the extension of the growing chain. Besides its major role in catalysis, the KS domain is essential for formation of the dimeric interface in FAS and PKS systems and, hence, is crucial for the architecture of these megasynthases.^[85,110] Overall, the KS domain of FASs and PKSs shows a relatively conserved primary sequence, which makes them predestined for phylogenetic analysis.^[6,42]

FASs' and PKSs' KS domains share a common fold, which resembles the characteristic fold of the thiolase-superfamily.^[3] The KSs form homodimers; each monomer consists of two sub-domains with alternating layers of α -helices and β -sheets forming an $\alpha/\beta/\alpha/\beta/\alpha$ sandwich motif.^[151] The dimer interface is stabilized by antiparallel β -strands, which exhibit hydrogen bonds.^[3,103] N-terminal KS domains of modular PKSs contain an additional α -helix which forms a coiled-coil structure in the dimeric interface.

The KS catalyzes a decarboxylative Claisen condensation using a catalytic triade consisting of the catalytic cysteine and two histidines.^[152] Similar as in the AT, the active site is also located at a "nucleophilic elbow" increasing its nucleophilicity. The reaction catalyzed by the KS follows a ping-pong mechanism (Fig. 2.9).^[150] In the ping step, an acyl moiety is loaded into the KS' active site. This acyl moiety is either a priming substrate or a fatty acid/polyketide processed intermediate. In the pong step, the carbon-carbon bond between the KS-bound acyl and an elongation substrate is formed. For this step, there are several possible sequences of events and catalytic mechanisms:^[41,153] In the sequential decarboxylation and carbon-carbon bond formation, the elongation substrate is decarboxylated to an enolate nucleophile, which attacks the KS-bound acyl, leading to a β -ketoacyl-ACP. This sequence may be accompanied by the release of carbon dioxide or bicarbonate.^[60] Other possible scenarios are the sequential carbon-carbon bond formation and decarboxylation or the concerted reactions. To the present knowledge, the first mechanism is the predominantly preferred one.^[41]

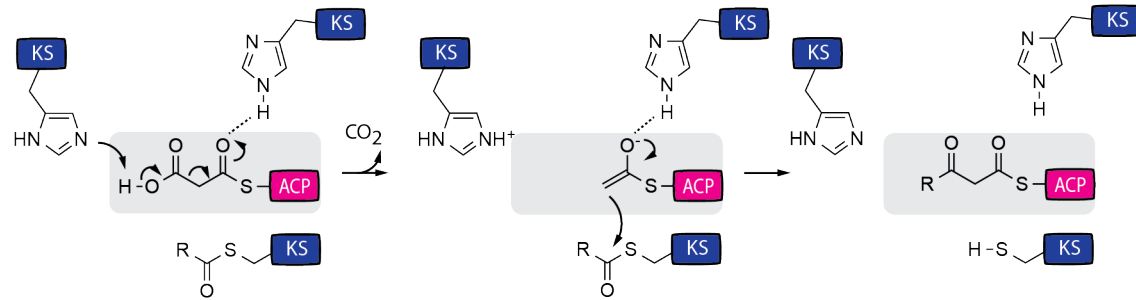
Introduction

ping step

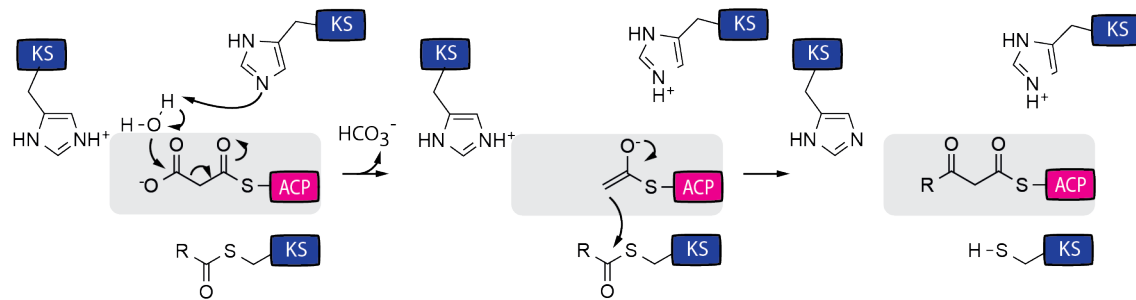


pong step

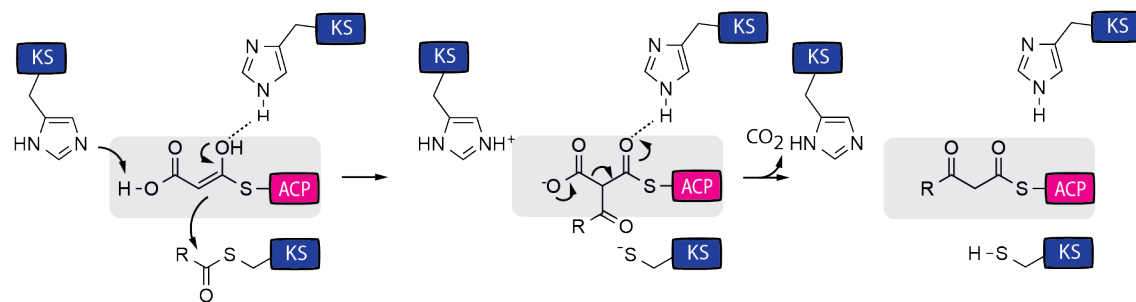
1A: Sequential decarboxylation & C-C bond formation (under CO₂ release)



1B: Sequential decarboxylation & C-C bond formation (under bicarbonate release)



2: Sequential C-C bond formation & decarboxylation



3: Concerted decarboxylation & C-C bond formation

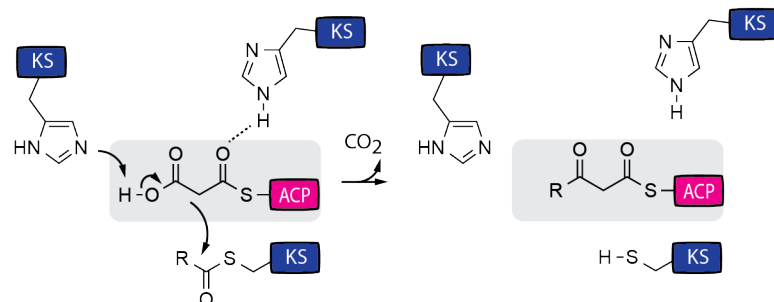


Figure 2.9: Ping-pong reaction catalyzed by the KS domain from mFAS: During transthioesterification, the acyl moiety X (either priming substrate or intermediate) is first loaded onto the KS. In the pong step, the carbon-carbon bond between the acyl X and an elongation substrate is formed. The detailed mechanism of the pong step is still not fully understood. Adapted from reference.^[41]

KSs in iterative systems accept acyl moieties of different lengths from priming substrates to different processed intermediates.^[41] Since the products of fatty acid/polyketide biosynthesis show an accurately defined length, certain chain length regulation processes have to be involved. With the KS being the essential domain/enzyme responsible for carbon-carbon bond formation, it is evident to attribute to it some kind of gatekeeper function. And indeed, structures show limited space in the KS's binding pocket.^[41] Only intermediates of certain chain lengths fit in for carbon-carbon bond formation, regulating the chain length of the final product.^[154] However, it remains unclear whether this control takes place in the ping or the pong step.^[41] Besides this chain length selectivity in iterative systems, it was shown that KSs of FAS systems exhibit a certain specificity for fully reduced acyl chains.^[59] However, the KS is also able to accept β -keto groups for condensation to produce triacetic acid lactone (TAL).^[155] Overall, it was shown that the KS domains from mammalian FASs accept other priming and elongation substrates besides their native substrates Ac and Mal for condensation.^[138,156,157] The substrate permissiveness of KSs from PKSs is an intrinsic property of PKS systems as they do not always use the full set of processing domains/enzymes. Indeed, it was shown that the KSs from the modular PKS DEBS tolerate a variety of substrates.^[158]

2.6.4 Important Domain-Domain Interactions

Besides the catalytic domains/enzymes themselves, proper domain-domain interactions are important for functional fatty acid/polyketide biosynthesis.^[133,159] Due to the transient nature of these interactions, only few of them are structurally solved. For the human FAS, an intensive computational docking study of the ACP to its partner domains has been performed, which enables mapping its trajectory and gives structural insights into catalytic mechanisms.^[159] However, this section will focus solely on the most important interfaces for the present work, namely the KS:ACP and the AT:ACP.

Recently, KS:ACP interfaces of the *Escherichia coli* (*E. coli*) type II FAS were solved.^[160,161] A mutation study gives insights into the mechanistic details of chain elongation. For the AT:ACP interface, there are structures of two PKS systems available.^[125,126] Both structures as well as the human FAS model show different binding motifs of the ACP towards the AT domain. A recent study on the AT:ACP interface of *E. coli* FAS supports the hypothesis of different AT:ACP binding modes by a mutation study, showing multiple, productive conformations.^[14]

Domains lined up on a polypeptide chain of type I FASs and PKSs do not interact directly in protein folding, but are covalently connected via discrete LDs. In modular PKSs, modules of the same polypeptide are also connected by covalent linkers. Non-covalent DDs connect modules of different polypeptide chains, forming coiled-coil structures.^[3]

The essential KS and AT domains are connected via a defined LD, the KS-AT linker (KAL). It contains an α - β motif, which interacts with the α - β - α motif of the AT domain. This rigid LD seems to be important for the KS-ACP interaction.^[162] Downstream of the AT domain, there is the post-AT linker (PAL), which interacts with the KAL and the KS domain. Both linker regions are relevant in the context of protein engineering, the next section will focus on.^[163,164] In sections 2.2 and 2.3, the original module boundaries of modular PKSs were shown. Each module starts with the KS domain and ends with the ACP. A recent study on four *cis*-AT PKSs shows the evolutionary co-migration of the ACP with the downstream, not the upstream KS.^[165] A study on *trans*-AT PKSs supports this observation.^[166] Accordingly, a new module definition, starting with the AT, ending with the KS, has been established.^[167] While these new module boundaries may reflect the evolution of PKSs, they do not consider the functional and architectural unit described by the classic module boundaries. This will be particularly interesting for different protein engineering approaches for directed synthesis of new products as discussed in the next section.

2.7 Generation of New Natural Products

In the context of resistant bacteria and drugs losing their effect, the development of new pharmaceuticals is indispensable.^[30] Nature, having generated the class of polyketides of diverse functions, has always set an example and inspired scientists of different disciplines. For a long time, chemists attempted to efficiently synthesize natural products in their laboratories.^[168] In 1981, R. B. Woodward and co-workers successfully synthesized and isolated Erythromycin A. They established a 50-step synthesis for four parts forming the desired compound, at least 13 people were involved in the practical part, and ended up with an overall yield of 0.00642%.^[169–171] This example shows how time-consuming and inefficient chemical synthesis can be, especially in the context of these complex natural products.

Concurrently, researchers came up with a different idea. They had the vision of exploiting the natural systems producing these complex compounds in order to create new derivatives with enhanced bioactivity or bioavailability or completely new functions.^[168] After the initial cloning of PKS gene clusters in the 1980s,^[172] different hybrid PKSs were constructed producing hybrid molecules.^[173,174] After the discovery of modular PKSs in the 1990s,^[26,175] these systems seemed to be predestined for this purpose due to their assembly line-like architecture.^[176,177] Following the principle of collinearity, distinct changes in modular PKSs promise to introduce predictable changes in the product. Different mix-and-match engineering approaches emerged from their modularity and are still being followed today: rational engineering approaches like reordering,^[178] deletion,^[179,180] inactivation,^[181,182] exchanging,^[183–185] or insertion^[186] of single domains or complete modules, substrate specificity engineering,^[140,187,188] extended to high-throughput combinatorial approaches.^[189,190] Different approaches succeeded in generating modified products: Specific manipulation of domains led to modified substrate selection,^[148,191] modified chain length of the product,^[192,193] and modified state of reduction of the β -position.^[194] New products were generated by PKSs with swapped modules or subunits.^[195] Despite these promising results, all these strategies often suffered from drastically low activity due to various reasons.^[6,18,196] New substrates, tolerated by a specific domain/module (swapping, insertion, specificity engineering) of the assembly line, or new intermediates were not tolerated by other domains/modules.^[197,198] Domains or modules were not able to communicate due to artificial intra- and intermodular protein-protein interactions.^[178,199] Rational control of stereochemistry was problematic,^[200] just to name a few challenges.

Regarding these obstacles, it became obvious that classical lab techniques cannot even rudimentarily keep up with nature's capability to evolve new PKSs for gaining new functions as response to evolutionary pressure. Thus, the idea of evolution-inspired engineering arose in the late 2000s.^[6] Since changes in domain specificity in nature appear to result from

domain swapping by recombination, one approach attempts to use these natural splice points for engineering with traditional cloning techniques.^[201–203] Other approaches generate new PKSs by natural recombination.^[204,205] Both strategies rely on the identification of conserved regions, which most likely indicate natural splice points. Regions of high sequence similarity were found in the KAL and PAL, which may explain early successes in AT swapping.^[20,79]

In this context, I. Abe and co-workers investigated the evolutionary comigration of PKSs (see section 2.6.4).^[165] This study supports the role of the PAL as natural splice point. Based on this, new module boundaries were defined and used to create hybrid assembly lines with limited success.^[167,206] Three hybrid assembly lines were tested and showed increased activity (23–50 times) using the updated module boundaries compared to the traditional boundaries, but overall product formation was significantly decreased to 40, 30, and 2.4 % of wild type activity. More studies are needed in order to reveal whether these new module boundaries also reflect the functional not only the evolutionary unit.

From this short outline it becomes evident that generation of new PKSs in order to synthesize modified products in a controlled and efficient manner is still challenging. This is mainly due to limited knowledge of the highly dynamic protein-protein interactions. Extensive studies on structure and function of these complex megasynthases were not able to illuminate biochemical details and the mechanistic basis of domain interplay remains obscure.

2.8 Motivation Behind This Thesis

The overall goal of this research project is the modification of biosynthetic pathways in megasynthases. More precisely, the project aims for converting the mFAS into an assembly line, like modular PKSs, in order to efficiently produce designer compounds with certain functions. In this context, some questions arise.

1. If nature uses modular PKSs as assembly lines, why should we even try to use the iterative mFAS for vectorial synthesis?
2. Why should the mFAS of all FASs be used in this approach?

First of all, FASs feature higher turnover numbers in their biosynthesis than PKSs, which might be due to their involvement in primary metabolism. This is important in the context of product yields. Furthermore, FASs are not as complex as PKSs and are well characterized in structure and function. This project is based on years of experience with FAS characterization and engineering. Previous investigation revealed the special suitability of the mFAS for this research approach due to its open architecture: Domains can easily be deleted, which is required for partly reducing or non-reducing modules.^[17]

Converting the iterative mFAS into an assembly line is challenging and several aspects need to be addressed (Fig. 2.10):

1. The loading function of the AT domain needs to be splitted into specific loading of priming and specific loading of elongation substrates. This can be achieved either by direct mutagenesis of the AT domain or by AT domain swapping. Mutant R606A was shown to be very effective in priming the ACP, loading the priming substrate Ac three times faster than the elongation substrate Mal.^[9,19] This transferase mutant can be used in a priming module.
2. Elongation modules with varying domain compositions need to be designed. As mentioned above, the mFAS fold allows for deletion of domains.^[17]
3. The interaction and communication between different modules need to be ensured. Covalent and non-covalent linkers need to be designed. First bimodular constructs with covalent linkers were already successfully designed.^[17] Recent studies have shown the suitability of natural DDs from PKS DEBS and of synthetic coiled-coils, termed SYNZIPs,^[207,208] for non-covalently bridging non-native interfaces in engineered PKSs.^[163,209] These linkers could be installed in the intended mFAS assembly line.

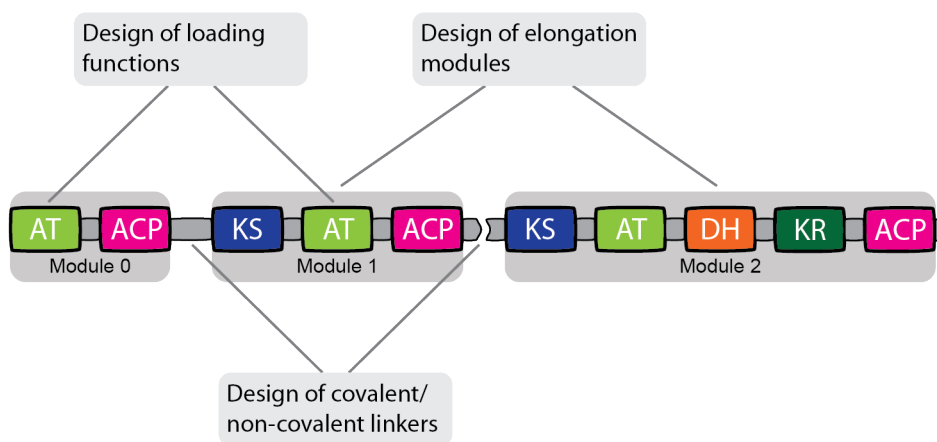


Figure 2.10: Schematic representation of the mFAS-based assembly line consisting of a priming and several exemplary elongation modules. Engineering aspects of the research plan are highlighted.

2.9 Aim

The present thesis is embedded in the previously described research project with a focus mainly on installation of the loading function in the elongation module via AT domain swapping and partly the design of an elongation module. Prior to AT swapping approaches, it is essential to understand this central domain in more detail. At the beginning of my doctoral studies, little was known about kinetics in PKSs, whereas the mFAS AT domain had been characterized in depth in our group.^[9]

My doctoral studies can basically be divided into two subprojects. The first is dedicated to the kinetic characterization of the AT domain from different PKSs. A fluorescence assay for kinetic analysis of ATs was already established in my master's thesis and should be further optimized during my doctoral studies. This comparative study should give deep insight into the kinetics of the complex and essential mechanism of substrate loading in PKSs. Furthermore, it should shed light on the relationship of FAS and PKS systems by revealing similarities and differences in kinetics. This first subproject is essential for FAS/PKS engineering and lays the foundation for the second subproject.

The second subproject focuses on the design of an elongation module for the artificial FAS assembly line. It mainly addresses one engineering aspect, the installation of a loading function in the mFAS elongation module via domain swapping. Based on the findings gained in the first subproject, potential candidates for domain swapping approaches should be identified. AT domains should be swapped in the mFAS, creating an elongation module. This module is of reduced complexity as it is allowed to keep its iterative working mode and does not have to interact with downstream modules, yet. Its functionality should be probed and, ideally, the interplay between priming and elongation module should be tested in first experiments. This last part is especially challenging, since it addresses all three engineering aspects mentioned above.

3 Results

Parts of this chapter have been previously accepted for publication in: F. Stegemann, M. Grininger, "Transacylation Kinetics in Fatty Acid and Polyketide Synthases and its Sensitivity to Point Mutations", *ChemCatChem* (2021).^[1] For individual contributions, copyright, and Creative Commons license, please see the statement of personal contributions.

3.1 Preliminary Work

As described above, the ATs as central domains in all FAS and PKS systems are attractive targets for protein engineering. They are gatekeepers for loading these protein complexes and control product synthesis as selectivity filters. Key for successful engineering strategies is detailed mechanistic understanding. During my master's thesis,^[210] we established an enzyme-coupled assay to quantify transferase activity, based on previous work.^[133,211] In the course of my doctoral studies, we further optimized this assay and subsequently used it to compare the AT domains of different FAS and PKS systems.

3.2 Kinetic Description of AT-mediated Reactions

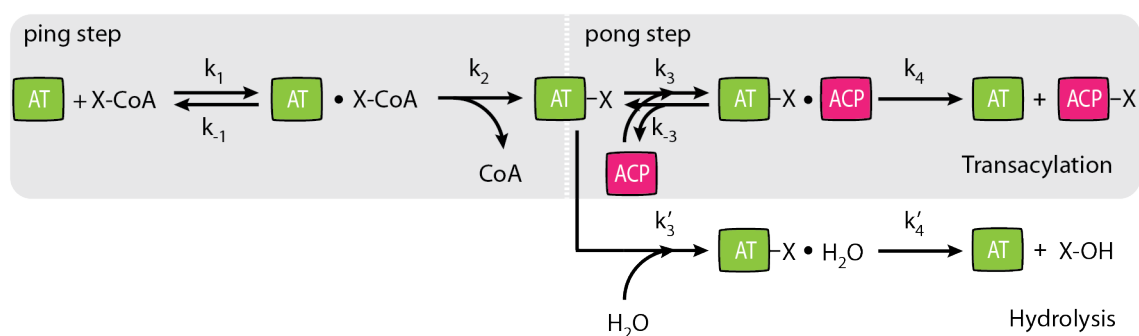


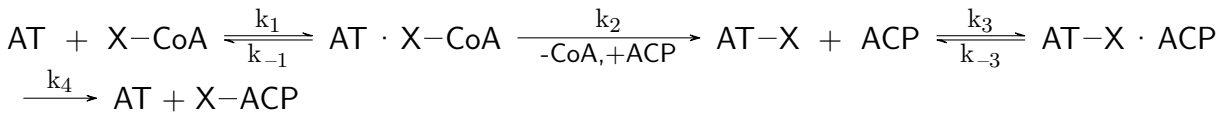
Figure 3.1: Schematic representation of AT-mediated reactions. After formation of the first enzyme-substrate complex AT-X-CoA, free CoA is released and AT-X is formed. The substrate X is either transacylated via formation of the second enzyme-substrate complex AT-X-ACP (main branch) or hydrolyzed during AT-mediated hydrolysis (side branch). The kinetic constants k_1 to k_4 as well as k_{-1} , k_{-3} , k'_3 , and k'_4 describe the AT-mediated reactions.

Results

The AT domain catalyzes the transfer of the acyl moiety X of the substrate X-CoA onto the *holo*-ACP domain (Fig. 3.1). This transacylation reaction follows a double displacement mechanism, called ping-pong bi-bi mechanism.^[129] Two substrates (X-CoA and ACP) are involved leading to two products (CoA and X-ACP). The double displacement mechanism starts with the ping step, in which the acyl moiety X is first transferred to the AT, resulting in release of product CoA. The pong step denotes the subsequent transfer of the acyl moiety X to the ACP after binding the second substrate ACP to AT-X, leading to the final product ACP-X. Hydrolysis occurs as side reaction, releasing the free acid “X-OH”. The mathematical description of these reactions is given in the following.

3.2.1 Transacylation

AT mediates the loading of the acyl moiety X onto the ACP domain via a double displacement reaction:



Notation:

$$c_1 := [\text{AT}], \quad \tilde{c}_1 := [\text{X-CoA}], \quad c_2 := [\text{AT} \cdot \text{X-CoA}], \quad c_3 := [\text{AT-X}], \\ \tilde{c}_3 := [\text{ACP}], \quad c_4 := [\text{AT-X} \cdot \text{ACP}], \quad c_{1,0} := [\text{AT}]_0, \quad \dot{c}_i := \frac{dc_i}{dt}.$$

Please note that c_i is a shorthand notation for $c_i(t)$ where t denotes time. $c_{1,0}$ is a shorthand notation for $c_1(0)$.

The following kinetic equations describe the chemical reactions following steady state assumptions:

$$v := \frac{d[\text{X-ACP}]}{dt} = k_4 c_4, \quad (3.1)$$

$$0 = \dot{c}_4 = k_3 c_3 \tilde{c}_3 - k_{-3} c_4 - k_4 c_4 \quad \Rightarrow \quad k_3 c_3 \tilde{c}_3 = (k_{-3} + k_4) c_4, \quad (3.2)$$

$$0 = \dot{c}_3 = k_2 c_2 - k_3 c_3 \tilde{c}_3 + k_{-3} c_4 \quad \Rightarrow \quad k_3 c_3 \tilde{c}_3 = k_2 c_2 + k_{-3} c_4, \quad (3.3)$$

$$0 = \dot{c}_2 = k_1 c_1 \tilde{c}_1 - k_{-1} c_2 - k_2 c_2 \quad \Rightarrow \quad k_1 c_1 \tilde{c}_1 = (k_{-1} + k_2) c_2, \quad (3.4)$$

$$c_1 = c_{1,0} - c_2 - c_3 - c_4. \quad (3.5)$$

Mathematical transformations:

$$(3.2) = (3.3) \quad (k_{-3} + k_4) c_4 = k_2 c_2 + k_{-3} c_4 \Rightarrow c_2 = \frac{k_4}{k_2} c_4, \quad (3.6)$$

$$c_2 \text{ into (3.4)} \quad k_1 c_1 \tilde{c}_1 = (k_{-1} + k_2) \frac{k_4}{k_2} c_4, \quad (3.7)$$

$$c_2 \text{ into (3.5)} \quad c_1 = c_{1,0} - c_3 - \frac{k_2 + k_4}{k_2} c_4, \quad (3.8)$$

$$\begin{aligned} c_1 \text{ into (3.7)} \quad k_1 \tilde{c}_1 \left(c_{1,0} - c_3 - \frac{k_2 + k_4}{k_2} c_4 \right) &= (k_{-1} + k_2) \frac{k_4}{k_2} c_4 \\ \Rightarrow \tilde{c}_1 c_3 &= c_{1,0} \tilde{c}_1 - \left(\frac{k_2 + k_4}{k_2} \tilde{c}_1 + \frac{k_4 (k_{-1} + k_2)}{k_1 k_2} \right) c_4, \end{aligned} \quad (3.9)$$

$$(3.2) \times \tilde{c}_1 \quad k_3 \tilde{c}_1 c_3 \tilde{c}_3 = (k_{-3} + k_4) \tilde{c}_1 c_4, \quad (3.10)$$

$$\begin{aligned} (3.9) \text{ into (3.10)} \quad k_3 \tilde{c}_3 \left[c_{1,0} \tilde{c}_1 - \left(\frac{k_2 + k_4}{k_2} \tilde{c}_1 + \frac{k_4 (k_{-1} + k_2)}{k_1 k_2} \right) c_4 \right] &= (k_{-3} + k_4) \tilde{c}_1 c_4 \\ \Rightarrow k_3 c_{1,0} \tilde{c}_3 \tilde{c}_1 &= c_4 \\ &\times \left(\frac{k_3 k_4 (k_{-1} + k_2)}{k_1 k_2} \tilde{c}_3 + \frac{k_3 (k_2 + k_4)}{k_2} \tilde{c}_3 \tilde{c}_1 + (k_{-3} + k_4) \tilde{c}_1 \right) \\ &= \frac{k_3 (k_2 + k_4)}{k_2} c_4 \\ &\times \left(\frac{k_4 (k_{-1} + k_2)}{k_1 (k_2 + k_4)} \tilde{c}_3 + \tilde{c}_3 \tilde{c}_1 + \frac{k_2 (k_{-3} + k_4)}{k_3 (k_2 + k_4)} \tilde{c}_1 \right) \end{aligned} \quad (3.11)$$

$$\Rightarrow c_4 = \frac{\frac{k_2}{k_2 + k_4} c_{1,0} \tilde{c}_3 \tilde{c}_1}{\frac{k_4 (k_{-1} + k_2)}{k_1 (k_2 + k_4)} \tilde{c}_3 + \tilde{c}_3 \tilde{c}_1 + \frac{k_2 (k_{-3} + k_4)}{k_3 (k_2 + k_4)} \tilde{c}_1}, \quad (3.12)$$

$$c_4 \text{ into (3.1)} \quad v = \frac{\frac{k_2 k_4}{k_2 + k_4} c_{1,0} \tilde{c}_3 \tilde{c}_1}{\frac{k_4 (k_{-1} + k_2)}{k_1 (k_2 + k_4)} \tilde{c}_3 + \tilde{c}_3 \tilde{c}_1 + \frac{k_2 (k_{-3} + k_4)}{k_3 (k_2 + k_4)} \tilde{c}_1}. \quad (3.13)$$

For a double displacement reaction, the velocity equation is defined as follows:

$$v = \frac{v_{\max} \tilde{c}_3 \tilde{c}_1}{K_m^{\tilde{c}_1} \tilde{c}_3 + \tilde{c}_3 \tilde{c}_1 + K_m^{\tilde{c}_3} \tilde{c}_1} \text{ with } v_{\max} = k_{\text{cat}} c_{1,0}. \quad [129] \quad (3.14)$$

Results

This yields the following kinetic parameters describing the transacylation:

$$k_{\text{cat}} = \frac{k_2 k_4}{k_2 + k_4}, \quad (3.15)$$

$$K_m^{\text{X-CoA}} = \frac{k_4(k_{-1} + k_2)}{k_1(k_2 + k_4)} = \frac{k_{\text{cat}}}{k_1} + \frac{k_{-1}k_4}{k_1(k_2 + k_4)}, \quad (3.16)$$

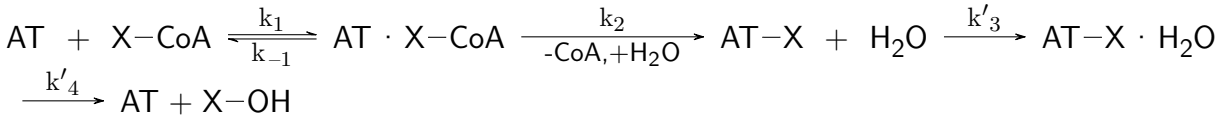
$$K_m^{\text{ACP}} = \frac{k_2(k_{-3} + k_4)}{k_3(k_2 + k_4)} = \frac{k_{\text{cat}}}{k_3} + \frac{k_2 k_{-3}}{k_3(k_2 + k_4)}, \quad (3.17)$$

$$\frac{k_{\text{cat}}}{K_m^{\text{X-CoA}}} = \frac{k_1 k_2}{k_{-1} + k_2}, \quad (3.18)$$

$$\frac{k_{\text{cat}}}{K_m^{\text{ACP}}} = \frac{k_3 k_4}{k_{-3} + k_4}. \quad (3.19)$$

3.2.2 AT-mediated Hydrolysis

As side reaction to transacylation, AT-mediated hydrolysis occurs. Both reaction pathways include the acyl transfer, during which free CoA is released. After loading the acyl moiety X onto the AT domain, the former is either transferred to the ACP or hydrolyzed. This AT-mediated hydrolysis follows these chemical reactions:



Notation:

$$c_1 := [\text{AT}], \quad \tilde{c}_1 := [\text{X-CoA}], \quad c_2 := [\text{AT} \cdot \text{X-CoA}], \quad c_3 := [\text{AT-X}], \\ \tilde{c}'_3 := [\text{H}_2\text{O}], \quad c'_4 := [\text{AT-X} \cdot \text{H}_2\text{O}], \quad c_{1,0} := [\text{AT}]_0, \quad \dot{c}_i := \frac{dc_i}{dt}.$$

The derivation is done analogously, the following parameters change from transacylation to hydrolysis:

$$\tilde{c}_3 \rightarrow \tilde{c}'_3, \quad c_4 \rightarrow c'_4, \quad k_3 \rightarrow k'_3, \quad k_4 \rightarrow k'_4, \quad k_{-3} \rightarrow 0.$$

This gives the following reaction velocity:

$$v = \frac{\frac{k_2 k'_4}{k_2 + k'_4} c_{1,0} \tilde{c}'_3 \tilde{c}_1}{\frac{k'_4 (k_{-1} + k_2)}{k_1 (k_2 + k'_4)} \tilde{c}'_3 + \tilde{c}'_3 \tilde{c}_1 + \frac{k_2 k'_4}{k'_3 (k_2 + k'_4)} \tilde{c}_1}. \quad (3.20)$$

Setting $\tilde{c}'_3 \approx 1$, since working in aqueous solution, gives v as follows:

$$v = \frac{\frac{k_2 k'_4}{k_2 + k'_4} c_{1,0} \tilde{c}_1}{\frac{k'_4 (k_{-1} + k_2)}{k_1 (k_2 + k'_4)} + \left[1 + \frac{k_2 k'_4}{k'_3 (k_2 + k'_4)} \right] \tilde{c}_1} = \frac{\frac{k_2 k'_3 k'_4}{k_2 k'_4 + k'_3 (k_2 + k'_4)} c_{1,0} \tilde{c}_1}{\frac{k'_3 k'_4 (k_{-1} + k_2)}{k_1 [k_2 k'_4 + k'_3 (k_2 + k'_4)]} + \tilde{c}_1}. \quad (3.21)$$

This yields the following kinetic parameters describing hydrolysis. Notably, $\frac{k_{\text{cat}}}{K_m^{\text{X-CoA}}}$ is the same as for transacylation:

$$k_{\text{cat}} = \frac{k_2 k'_3 k'_4}{k_2 k'_4 + k'_3 (k_2 + k'_4)}, \quad (3.22)$$

$$K_m^{\text{X-CoA}} = \frac{k'_3 k'_4 (k_{-1} + k_2)}{k_1 [k_2 k'_4 + k'_3 (k_2 + k'_4)]} = \frac{k_{\text{cat}}}{k_1} + \frac{k_{-1} k'_3 k'_4}{k_1 [k_2 k'_4 + k'_3 (k_2 + k'_4)]}, \quad (3.23)$$

$$\frac{k_{\text{cat}}}{K_m^{\text{X-CoA}}} = \frac{k_1 k_2}{k_{-1} + k_2}. \quad (3.24)$$

3.3 Kinetic Analysis of Acyl Transferases

As part of my master's thesis, we established a fluorescence assay, which was further optimized during my doctoral studies. The transacylation involves two substrates and can be seen as two sequential Michaelis-Menten reactions, which can be kinetically analyzed by titration of substrates. Titration of the acyl substrate to a series of fixed ACP concentrations (ACP titration) allows for quantification of transacylation activity. Via the velocity equation (3.13) describing this reaction, the data can be plotted globally,^[212] determining the kinetic parameters described in equations (3.15) to (3.19). For quantification of AT-mediated hydrolysis, X-CoA is titrated in the absence of ACP. The corresponding Michaelis-Menten plot gives access to the kinetic parameters described in equations (3.22) to (3.24).

3.3.1 Selection of Acyl Transferases

To understand similarities and differences between types of FASs and PKSs, the following AT domains were selected for analysis: Representing elongation ATs of modular PKSs, the AT of DEBS M5 from *Saccharopolyspora erythraea* (*S. erythraea*) (termed DEBS3M5), of PikA M5 from *Streptomyces venezuelae* (PikAIIIIM5), and of rapamycin synthase (RAPS) module 14 from *Streptomyces hygroscopicus* (RAPS3M14) were used. The first two ATs transfer MMal, while the latter transfers Mal.^[4,87] Priming ATs of DEBS and avermectin PKS (DEBS1M0 and AVES1M0) were used, loading Prop and 2-methylbutyryl (2-MB). For the iterative PKSs, we worked with the AT of Pks7 from *S. erythraea*. Pks7 AT is phylogenetically similar to 6-methylsalicylic acid synthase (MSAS) AT from *Penicillium patulum*^[213] and is likely involved in methylsalicylic acid (MSA) synthesis (Fig. 3.2).^[214] Thus, it is expected to load Ac and Mal onto the ACP for MSA production.^[3,215] The specificity for Mal, and not MMal, is supported by its amino acid fingerprint (Fig. 3.3).^[3] Data will be compared to the Mal-loading type II FAS AT from *E. coli* (AT termed FabD) and to the previously analyzed type I mFAS AT (mMAT), which was found to be promiscuous in transferring a large variety of substrates.^[9] If not explicitly mentioned otherwise, the terms FAS and PKS will refer to type I FAS and type I PKS in the following.

Results

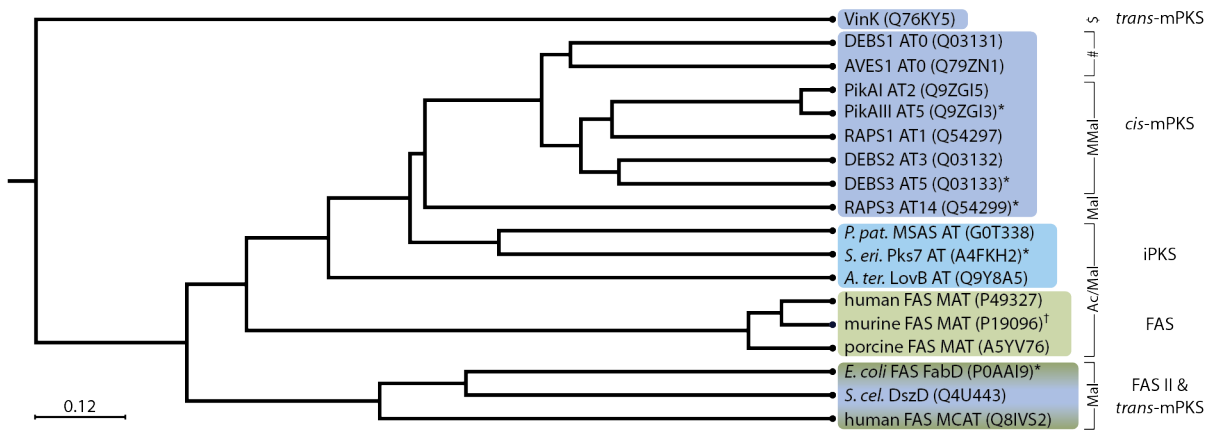


Figure 3.2: Schematic representation of the phylogenetic relationship between AT domains from FASs and PKs (UniProt IDs in brackets). ATs from modular PKs form a distinct clade with a subclade for the priming ATs (# indicates priming substrates: Prop for DEBS1 AT0 and 2-MB for AVES1 AT0). As expected, RAPS3 AT14 accepting Mal is separated from the MMal-transferring ATs. ATs from iterative PKs fall into a distinct clade. FAS II ATs form a distinct clade with the *trans*-acting AT DszD from disorazole PKS. All three ATs transfer Mal as substrate. The *trans*-acting AT VinK from vicenistatin PKS transfers an unusual substrate – a dipeptidyl moiety (indicated by \$)^[125] – and falls into a different clade. FAS systems depicted in green (type I light green, type II dark green), iterative and modular PKS systems (iPKS and mPKS) shown in blue and violet, respectively. ATs analyzed in this study are indicated by *. The MAT analyzed in a previous study and discussed in this study is marked with †.^[9] Phylogenetic tree created using CLC Main Workbench 6.9.1 (tree construction method: UPGMA, protein distance measure: Jukes-Cantor, bootstrap value=100).

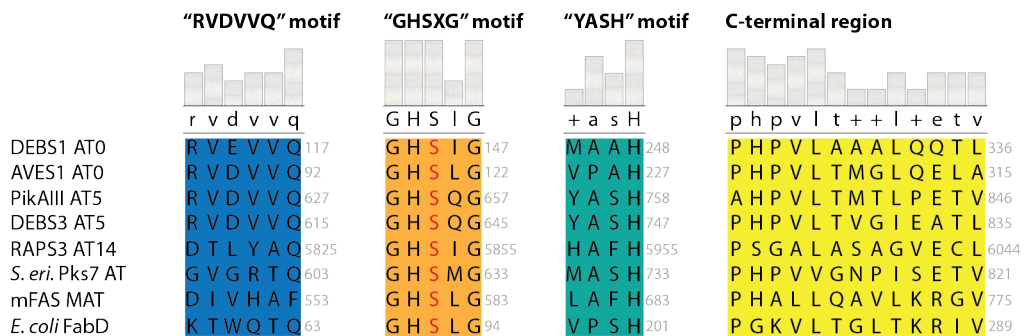


Figure 3.3: Excerpt of the sequence alignment of AT domains showing the characteristic motifs and regions for substrate specificity with the active site serine depicted in red. Primary sequences from priming ATs DEBS1 AT0 (UniProt: Q03131) and AVES1 AT0 (Q79ZN1), from specific elongation ATs PikAIII AT5(Q9ZGI3), DEBS3 AT5 (Q03133), RAPS3 AT14 (Q54299), and *E. coli* FabD (P0AAI9), and from bifunctional ATs *S. eri.* Pks7 AT (A4FKH2) and mFAS MAT (P19096). Alignment created using CLC Main Workbench 6.9.1.

3.3.2 Isolation of Acyl Transferases

Recent functional and structural characterization revealed that type I FAS and PKS systems form stable KS-AT dimers.^[9,99] Therefore, all elongation AT domains were expressed as KS-AT didomains. A KS knockout, leading to constructs denoted as KS⁰-AT, allowed for selectively inspecting the AT domain in transacylation properties. Priming AT domains and FabD were expressed as standalone proteins. Different DEBS1 AT0 constructs did not result in soluble protein. Thus, a construct with C-terminal maltose binding protein (MBP) fusion was designed. Recombinant expression of all ATs in *E. coli* gave sufficient yields of soluble, high-quality proteins (Tab. S1). Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) confirmed high purity after tandem affinity chromatography (Fig. S1). Analytical size exclusion chromatography (SEC) (Fig. S2) revealed stable dimers for Pks7 and PikAIIIM5 KS⁰-AT as well as tetrameric and dimeric oligomers for DEBS3M5 KS⁰-AT. Surprisingly, RAPS3M14 KS⁰-AT eluted from SEC mainly as monomer, besides a subfraction of tetrameric species. We note that the prevalence for a monomeric state should not impact AT activity, since the dimeric interface is formed by the KS domain, while the AT domain is monomeric in its active form. As expected, the standalone ATs *E. coli* FabD, DEBS1M0 AT-MBP, and AVES1M0 AT were solely monomeric. Thermal shift assay (TSA) in two buffers validated protein quality, giving melting points within a range of 45-65 °C and 36-61 °C in storage and assay buffer, respectively (Tab. S2 and Fig. S3). All AT domains were prepared in biological triplicates, produced from single clones of one or more independent plasmid transformations.

3.3.3 Isolation of Acyl Carrier Proteins

Complete kinetic characterization of AT domains requires high yield and high quality of the corresponding ACPs. In order to generate *holo*-ACP, co-expression of ACPs with the 4'-PPT from *Bacillus subtilis* (Sfp) was performed, which, however, resulted in insufficient yields of most ACPs. Using codon-optimized sequences, the *holo*-ACP yields increased significantly. ACPs were purified by affinity chromatography and SEC of which the latter allowed selecting soluble ACPs. For typical expression yields and SDS-PAGE analyses, see Table S3 and Figure S4. Quantitative *holo*-ACP formation was confirmed by mass spectrometry (MS; Tab. S4 and Fig. S5). Protein integrity was confirmed by TSA giving melting temperatures within a range of 53-64 °C and 50-65 °C in storage and assay buffer, respectively (Tab. S2 and Fig. S6).

3.3.4 Initial Substrate Screening

At the outset of the functional characterization of AT domains, we performed an initial substrate screen. In doing so, three substrates were used for the elongation ATs: Mal and MMal, as typical elongation units, and Ac, as typical priming unit, and likewise a negative control for the ATs of modular PKSs. For the priming ATs, two priming substrates (the native and a non-native typical priming unit) and Mal as a negative control were used. For the initial substrate screen, AT domains were probed at fixed ACP and X-CoA concentrations.

With our initial screen, we essentially confirmed the substrate specificity of the elongation ATs from modular PKSs (DEBS3M5, PikAIIIM5, and RAPS3M14) as expected from the chemical structure of the product (Fig. 3.4A and Tab. S5): The AT domains of DEBS3M5 and PikAIIIM5 transferred solely MMal, and RAPS3M14 solely Mal. The AT domains of the iterative Pks7 tolerated MMal besides its natural elongation substrate Mal, as well as the priming substrate Ac at slower rate. Since MMal-CoA is present in the native host *S. erythraea*, the loose specificity implies that both elongation substrates can be loaded for condensation. We analyzed Pks7 in transacylation of both substrates for deeper insight in the following (see section 3.3.6). For the mMAT, broad substrate tolerance was already shown previously, including the fast transfer of the priming substrate Ac and elongation substrate Mal with similar rates as well as the transfer of MMal at lower rate.^[9] The *E. coli* AT FabD tolerated MMal in addition to its native elongation unit Mal, as published previously.^[216]

Probing AT-mediated hydrolysis, we observed that most domains hydrolyze their native substrates with higher rates than non-native substrates (Fig. 3.4B).^v This overlapping substrate specificity of transacylation and hydrolysis was also reported before.^[133,143] Only PikAIII AT5 showed higher hydrolysis rates for the non-native elongation substrate Mal, which was also reported for PikAIV AT6.^[132] Overall, this data confirms that there is no general hydrolysis-based proofreading mechanism present in AT domains for clearing the active site from wrongly loaded acyl moieties, as previously suggested.^[217,218]

^vWe note that for determination of transacylation rates, self-acylation of ACP and not AT-mediated hydrolysis was subtracted as background. Accordingly, in this initial screening, the PKSs' transacylation rates may contain hydrolysis rates to some extent.

Results

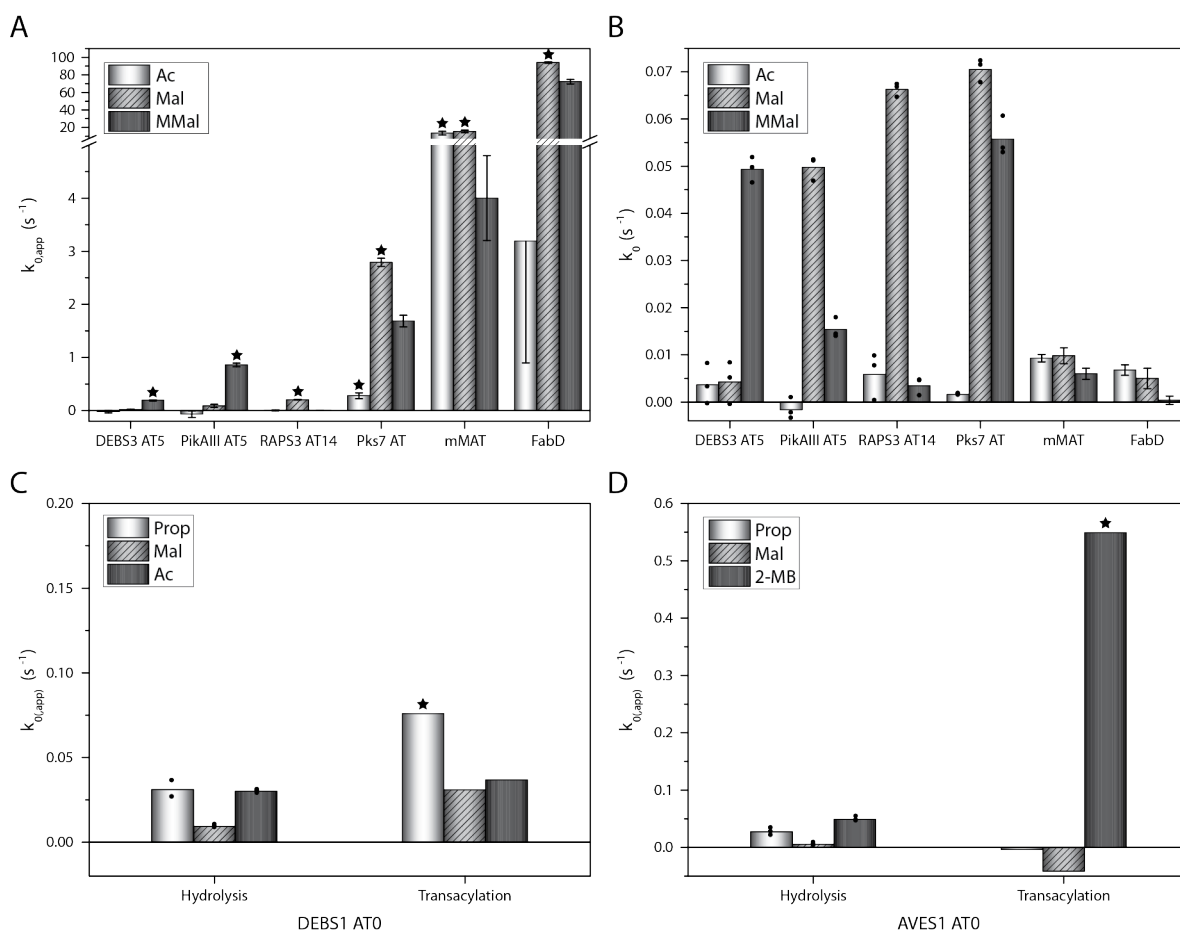


Figure 3.4: Initial screening of FAS and PKS systems. (A) Transacylation screening. Loading of Ac- (white), Mal- (light gray, hatched), and MMal-CoA (dark gray) was measured in presence of ACP. Substrate concentrations (X-CoA and ACP) were fixed at $50 \mu\text{M}$. Bars indicate the average substrate turnover. A star indicates the native substrates for transacylation. Error bars show the standard deviation of biological triplicates. mMAT data from a previous study, referring to $k_{cat, app}$ at saturated X-CoA concentration and $60 \mu\text{M}$ ACP.^[9] (B) Hydrolysis screening. Loading of Ac- (white), Mal- (light gray, hatched), and MMal-CoA (dark gray) was measured in absence of ACP. Substrate concentration (X-CoA) was fixed at $50 \mu\text{M}$. Bars indicate the average substrate turnover. Black dots correspond to the average of technical triplicates of one biological replicate. Error bars show the standard deviation of biological triplicates. mMAT data from a previous study, referring to k_{cat} at saturated X-CoA concentration.^[9] (C) Initial hydrolysis and transacylation screening of DEBS1M0 AT. Loading of Prop- (white), Mal- (light gray, hatched), and Ac-CoA (dark gray) was measured in absence and presence of ACP, respectively. Substrate concentrations (X-CoA and ACP) were fixed at $50 \mu\text{M}$. Hydrolysis bars indicate the average substrate turnover. Black dots correspond to the average of technical triplicates of one biological replicate. Transacylation bars are exemplary single well measurements. The star indicates the native substrate for transacylation. (D) Initial hydrolysis and transacylation screening of AVES1M0 AT. Loading of Prop- (white), Mal- (light gray, hatched), and 2-MB-CoA (dark gray) was measured in absence and presence of ACP, respectively. Substrate concentrations (X-CoA and ACP) were fixed at $50 \mu\text{M}$. Hydrolysis bars indicate the average substrate turnover. Black dots correspond to the average of technical triplicates of one biological replicate. Transacylation bars are exemplary single well measurements. The star indicates the native substrate for transacylation.

We faced difficulties measuring the transacylation rates mediated by the priming AT of DEBS. Large signal fluctuations occurred in numerous measurements indicating that there was no significant acyl transfer. Comparing hydrolysis rates and presumed transacylation rates confirmed this assumption: Transacylation rates were in the same range and not significantly higher compared to hydrolysis rates (Fig. 3.4C). The priming AT of AVES, on the other hand, showed significantly higher transacylation rates than hydrolysis rates like all the other systems analyzed (Fig. 3.4D). This indicates that the MBP fusion is interfering with the DEBS1M0 AT-ACP interaction. Nevertheless, the acyl substrate seemed to be loaded onto the AT domain, but transfer onto the ACP domain is not possible anymore. Both priming ATs accepted the native and the non-native priming substrate in hydrolysis.

Based on this initial screening, full kinetic analysis of AT-mediated hydrolysis and transacylation was performed with the substrates that are preferentially transferred. For the priming ATs, only AT-mediated hydrolysis was analyzed.

3.3.5 AT-mediated Hydrolysis

To analyze AT-mediated hydrolysis, the substrates X-CoA were titrated in absence of ACP. We faced difficulties in conducting the hydrolysis assay, originating from extreme differences in K_m^{X-CoA} , and we were just able to determine kinetic parameters precisely for K_m^{X-CoA} values lying within suited conditions for recording Michaelis-Menten kinetics.

1. For high Michaelis-Menten constants, high substrate concentrations have to be used. This can lead to experimental problems caused for example by substrate inhibition effects, and only rough values for K_m^{X-CoA} and k_{cat} can be given, leading also to catalytic efficiencies k_{cat}/K_m^{X-CoA} that are error-prone.
2. For low Michaelis-Menten constants K_m^{X-CoA} , low substrate concentrations have to be used, causing poor signal-to-noise ratios in the assay. Increasing enzyme concentrations leads to disregard of Michaelis-Menten concentration requirement. Thus, for AT domains with low K_m^{X-CoA} , just k_{cat} is determined precisely, whereas K_m^{X-CoA} and therefore also k_{cat}/K_m^{X-CoA} are approximate values.

Results

Table 3.1: Kinetic parameters determined for AT-mediated hydrolysis. Measurement performed in technical triplicates of biological triplicates (DEBS3 AT5, DEBS1 AT0 with Ac), technical triplicates of one replicate (PikAIII AT5, RAPS3 AT14, DEBS1 AT0 with Prop, and AVES1 AT0 with 2-MB and Prop), or biological triplicates (Pks7 with Mal and MMal). Abbreviations: n. d.: not determinable. ^[a] This system has a very low K_m^{X-CoA} . Only k_{cat} is determined precisely. ^[b] This system has a very high K_m^{X-CoA} . Measurement gave large signal fluctuations (Fig. S7). Corresponding kinetic values with substantial errors are indicated by constants in italic letters. ^[c] Measurement gave large signal fluctuations (Fig. S7). Corresponding kinetic values with substantial errors are indicated by constants in italic letters. ^[d] mMAT from previous study.^[9]

Protein X-CoA	k_{cat} (s^{-1})	K_m^{X-CoA} (μM)	k_{cat}/K_m^{X-CoA} ($s^{-1}M^{-1}$)
DEBS3 AT5 MMal	$5.31 \times 10^{-2} \pm 9.62 \times 10^{-4}$	$7.88 \pm 3.98 \times 10^{-1}$	6.7×10^3
PikAIII AT5 MMal ^[a]	$1.94 \times 10^{-2} \pm 5.03 \times 10^{-4}$	$< 8.91 \times 10^{-1}$	$> 2.2 \times 10^4$
PikAIII AT5 Mal ^[b]	<i>$1.95 \times 10^{-1} \pm 1.54 \times 10^{-2}$</i>	<i>99.9 ± 28.8</i>	n. d.
RAPS3 AT14 Mal ^[a]	$6.51 \times 10^{-2} \pm 1.89 \times 10^{-3}$	$< 5.45 \times 10^{-1}$	$> 1.2 \times 10^5$
Pks7 AT Mal ^[a]	$9.93 \times 10^{-2} \pm 3.17 \times 10^{-3}$	$< 4.74 \times 10^{-1}$	$> 2.1 \times 10^5$
Pks7 AT MMal ^[a]	$7.34 \times 10^{-2} \pm 1.51 \times 10^{-3}$	$< 3.98 \times 10^{-1}$	$> 1.8 \times 10^5$
DEBS1 AT0 Prop ^[a]	$3.08 \times 10^{-2} \pm 2.37 \times 10^{-3}$	$< 5.46 \times 10^{-1}$	$> 5.6 \times 10^4$
DEBS1 AT0 Ac	$3.65 \times 10^{-2} \pm 1.65 \times 10^{-3}$	$7.73 \pm 9.70 \times 10^{-1}$	4.7×10^3
AVES1 AT0 2-MB ^[a]	$4.86 \times 10^{-2} \pm 3.75 \times 10^{-4}$	$< 2.12 \times 10^{-1}$	$> 2.3 \times 10^5$
AVES1 AT0 Prop ^[c]	<i>$5.19 \times 10^{-1} \pm 1.08 \times 10^{-2}$</i>	<i>25.2 ± 18.5</i>	n. d.
mMAT Mal ^[d]	$(9.8 \pm 1.7) \times 10^{-3}$	n. d.	n. d.
mMAT MMal ^[d]	$(6.0 \pm 1.2) \times 10^{-3}$	n. d.	n. d.
mMAT Ac ^[d]	$9.3 \times 10^{-3} \pm 1.2 \times 10^{-4}$	n. d.	n. d.

Given these constraints, only DEBS3M5 (with MMal) and DEBS1M0 (with Ac) AT-mediated hydrolysis was eventually characterized in K_m^{X-CoA} and k_{cat} at high confidence, while for the other AT domains of PKSs partly just rough values were determined (Tab. 3.1). Kinetic parameters determined for DEBS3M5 AT-mediated hydrolysis are very similar to the data on DEBS2M3 AT-mediated hydrolysis previously published by Dunn *et al.*^[133] Overall, AT-mediated hydrolysis rates for the native substrates are similar for all analyzed PKS systems. Hydrolysis mediated by mMAT is very slow compared to the PKS systems. Hydrolysis by FabD was not measurable at all. For hydrolysis titration curves, see Figure S7.

3.3.6 AT-mediated Transacylation

To characterize the transacylation reaction, both substrates X-CoA and ACP were titrated. Specifically, we titrated X-CoA substrates to five different, fixed ACP concentrations and globally fitted all titration curves. This approach is robust to measurement errors as well as delivers absolute kinetic constants and is in its quality superior to other approaches titrating one substrate while keeping the other at a fixed saturated concentration.^[14,216] For all ATs, absolute kinetic parameters were received, although the kinetic parameters for systems with high K_m^{ACP} were determined less accurately, simply because the vast amounts of ACP needed to cover an appropriate substrate range could not always be provided. This was particularly problematic for Pks7. Here, the ACP concentration was varied only within a range of $0.16-0.80 \times K_m^{ACP}$ for the substrate MMal-CoA.

Overall, data reveals that PKS systems feature slower turnover rates than found in AT of FAS type I and type II (Tab. 3.2 and Fig. 3.5),^[9] with the AT domain of the iterative Pks7 transferring the substrates with significantly higher rates than AT domains of modular PKSs. Catalytic efficiencies k_{cat}/K_m^{X-CoA} and k_{cat}/K_m^{ACP} show that mMAT and FabD transacylate with highest efficiency, followed by the iterative Pks7 AT (about one order of magnitude less efficient than mMAT/FabD) and modular PKSs (another order of magnitude less efficient than Pks7 AT), indicating that the iterative systems feature low transition state energies for the transacylation reaction. Particularly, the catalytic efficiencies of the AT-acylating ping step k_{cat}/K_m^{X-CoA} is outstandingly high for mMAT and FabD compared to ATs of modular PKSs; e. g., Mal is transferred by mMAT and FabD with catalytic efficiency of $1.4 \times 10^7 \text{ M}^{-1}\text{s}^{-1}$ compared to RAPS3 AT14 with $4.8 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$ (Tab. 3.2).

Results

Table 3.2: Kinetic parameters determined for AT-mediated transacylation. Measurement performed in biological triplicates. [a] mMAT from previous study.[9]

Protein X-CoA	k_{cat} (s^{-1})	K_m^{X-CoA} (μM)	K_m^{ACP} (μM)	k_{cat}/K_m^{X-CoA} ($s^{-1}M^{-1}$)	k_{cat}/K_m^{ACP} ($s^{-1}M^{-1}$)	k_{cat} norm (%)	K_m^{X-CoA} norm (%)	K_m^{ACP} norm (%)	$k_{cat}/K_m^{X-CoA}/$ (k_{cat}/K_m^{ACP})
DEBS3 AT5 MMal	$1.25 \pm$ 7.11×10^{-2}	209 ± 17.5	72.0 ± 8.21	6.0×10^3	1.7×10^4	1.1	4.3×10^{-2}	3.5	3.4×10^{-1}
PikAIII AT5 MMal	$7.47 \pm$ 5.16×10^{-1}	96.2 ± 8.69	330 ± 44.1	7.8×10^4	2.2×10^4	6.3	5.6×10^{-1}	4.5	3.4
RAPS3 AT14 Mal	$7.10 \times 10^{-1} \pm$ 3.84×10^{-2}	14.8 ± 1.33	66.0 ± 10.8	4.8×10^4	1.1×10^4	6.0×10^{-1}	3.4×10^{-1}	2.2	4.6
Pks7 AT Mal	23.4 ± 1.50	$9.48 \pm 8.84 \times 10^{-1}$	286 ± 30.3	2.5×10^6	8.2×10^4	20	18	17	30
Pks7 AT MMal	20.5 ± 2.03	21.1 ± 2.46	505 ± 73.8	9.7×10^5	4.1×10^4	17	6.9	8.4	24
mMAT Mal[a]	119	8.8	245	1.4×10^7	4.9×10^5	100	100	100	28
mMAT Ac[a]	99.2	12	265	8.3×10^6	3.7×10^5	83	59	76	22
FabD Mal	268 ± 15.3	19.7 ± 1.88	44.3 ± 5.59	1.4×10^7	6.0×10^6	225	224	18	2.2
FabD MMal	273 ± 15.6	99.0 ± 8.78	45.2 ± 7.76	2.8×10^6	6.0×10^6	229	1125	18	0.46

Comparing the AT-loading (ping) step with the ACP-loading (pong) step in catalytic efficiencies by forming the ratio $(k_{\text{cat}}/K_{\text{m}}^{\text{X-CoA}})/(k_{\text{cat}}/K_{\text{m}}^{\text{ACP}})$ shows that the iterative type I systems feature efficiency ratios of up to 30 for native elongation substrates, while ratios for the slower modular PKS systems are significantly lower (Tab. 3.2). We interpret a high ratio as representing an AT domain that is particularly efficient in the initial acylation ping step (AT-X formation) and, in doing so, supports high overall transacylation rates due to a higher probability of a productive AT-X-ACP interaction in the pong step. Following this interpretation, the high catalytic efficiency of Pks7 AT and mMAT in their ping step is key to overall high turnover rates. Vice versa, the lower efficiency in the ping step of ATs of modular PKSs contributes to their slower transacylation rates. The type II FabD, on the other hand, shows an outstandingly high efficiency also for the pong step, which leads to a reduced ratio of efficiencies. This property might be explained by the dissociated enzymes in type II systems. Both substrates, acyl substrate and ACP, need to be selected out of the vast variety of molecules in the environment (efficient ping and pong step required). Whereas in type I systems, only the acyl substrate needs to be channeled into the assembly line (efficient ping step).

The $K_{\text{m}}^{\text{X-CoA}}$ values for Ac-CoA and Mal-CoA are in the range of bacterial metabolite levels^[219] and are similar for most AT substrate pairs. Interestingly, the $K_{\text{m}}^{\text{MMal-CoA}}$ ($K_{\text{m}}^{\text{X-CoA}}$ for MMal-CoA) values for DEBS3 AT5 and PikAIII AT5 are ten- to twentyfold higher than $K_{\text{m}}^{\text{X-CoA}}$ for other systems tested (Tab. 3.2). This either indicates an adaptation to higher cellular MMal-CoA levels or a control mechanism for regulating MMal uptake depending on its availability. The regulation of polyketide biosynthesis by AT-mediated transacylation would be particularly efficient, when a limited metabolic flux of MMal into the polyketide biosynthetic pathway upon shortage of MMal-CoA (possible by the high $K_{\text{m}}^{\text{MMal-CoA}}$) becomes overall rate limiting. We note that a similar regulation of polyketide biosynthesis would not be plausible with Mal-CoA, which is dedicated mainly to the central metabolism of any bacterial cell and regulated in concentration to its need as precursor in fatty acid biosynthesis.^[220]

Our data further shows that Michaelis-Menten constants $K_{\text{m}}^{\text{ACP}}$ vary moderately for the different ATs (Tab. 3.2). Considering molar concentrations within the FAS/PKS compartment of about 1 mM (rough calculation based on a cylinder volume of dimensions taken from MAS-like PKS; i. e., radius 10 nm, height 15 nm; 2 molecules) as well as freely diffusing domains within the multidomain compartment, all AT domains of type I systems operate at high (Pks7) to saturated (other systems tested) ACP conditions so that the pong step runs at maximum rate (v_{max}).

Results

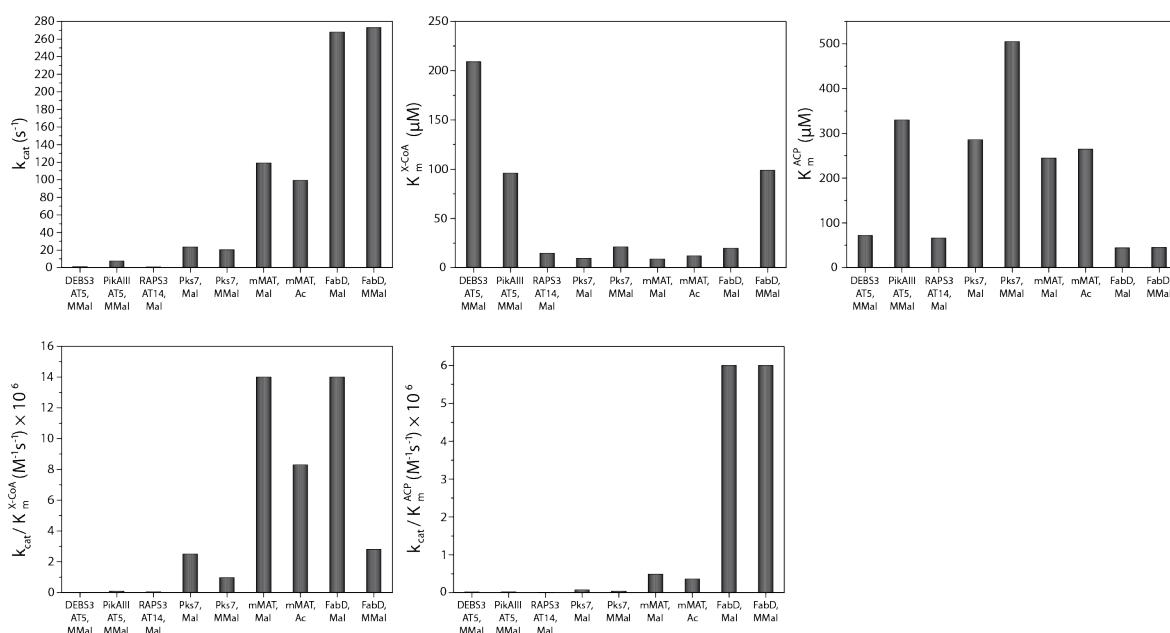


Figure 3.5: Plots of transacylation parameters. Turnover rate, Michaelis-Menten constants, and catalytic efficiencies determined for transacylation for the FAS and PKS systems. mMAT data from previous study.^[9]

The analysis of the AT domain of the iterative Pks7 in transferring Mal and MMal gives further interesting insight into the AT-mediated transacylation. Both Michaelis-Menten constants K_m^{X-CoA} and K_m^{ACP} differ for the two tested elongation substrates Mal-CoA and MMal-CoA (Tab. 3.2), which implies that the interaction of AT and ACP is modulated by the substrate to be transferred, similarly as recently suggested for *E. coli* type II FAS.^[14] Catalytic efficiencies differ for the two substrates, indicating a higher efficiency for Mal in both steps, the AT-loading ping and the ACP-loading pong step. This data suggests a double selection for Mal over MMal to ensure the specific condensation with Mal in the organism. For transacylation titration curves, see Figure S8. We note that, based on the catalytic efficiency of the ping step, the Michaelis-Menten constants for the AT-mediated hydrolysis can be calculated (Tab. S6).

3.3.7 AT:ACP Interface Mutation Study

Some of the molecular details to the ACP-mediated shuttling of substrates and intermediates to the catalytic domains in FASs and PKSs are being currently revealed. For example, interactions are guided by weak electrostatic interactions,^[221] and, as shown for the AT-ACP interaction in the *E. coli* type II system (FabD:AcpP interface), do not necessarily involve prevalent interfaces.^[14] In order to help understanding this critical domain-domain interaction in type I complexes, the AT:ACP interface of DEBS3M5 was mutated and analyzed with the transacylation assay in enzyme kinetic detail. Three available AT:ACP complex structures^[125,126,159] were used for modeling the interaction (Fig. 3.6 and S9). Based on that, different bioinformatics tools were used to identify interfacial residues. Eventually two residues, A539 and R850, were selected that are likely involved in the AT-ACP interaction (residue numbering following UniProt entry Q03133). Overall, nine point mutations were introduced in the AT domain of the KS⁰-AT didomain construct of DEBS3M5; A539S, A539D, A539E, A539F, R850K, R850A, R850E, R850F, R850S. All mutants showed wild type-like properties during preparation (Tab. S7 and S8 and Fig. S10, S11, and S12).

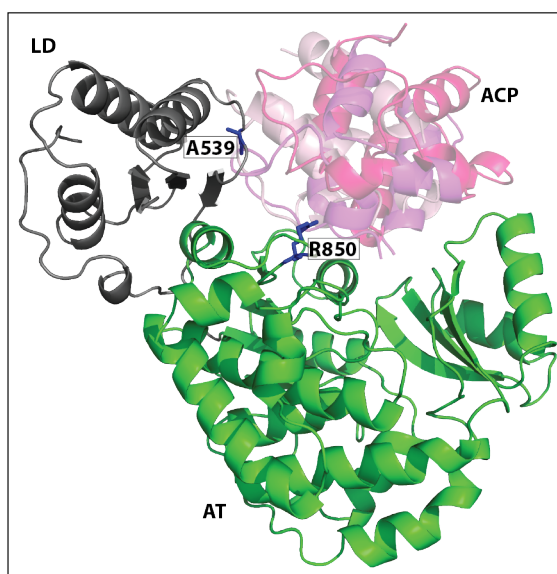


Figure 3.6: Cartoon representation of the modeled DEBS3M5 AT:ACP interface. AT domains of known AT:ACP interfaces were used to align DEBS3 AT5 (PDB: 2HG4,^[99] depicted in green) structurally. The LD in the KS-AT didomain substructure is depicted in gray. DSZS ACP (PDB: 5ZK4^[126]), VinL (PDB: 5CZD^[125]), and hFAS ACP^[159] are depicted in light pink, hotpink, and magenta, respectively. Despite their differences, all models locate amino acids A539 and R850 (highlighted in blue) in the AT:ACP interfaces. Residue numbering following UniProt entry Q03133.

Results

In an initial screen, the mutants were tested in transacylation activity at fixed MMal-CoA and fixed ACP concentrations. We additionally screened the hydrolysis rates at fixed MMal-CoA concentrations. Since AT-mediated hydrolysis does not involve ACP, hydrolysis rates are able to report any non-local effect of the surface mutation. My initial screen revealed that some mutations affect both transacylation and hydrolysis, although the impact on hydrolysis was comparably lower than on transacylation (Fig. 3.7). This data shows that indeed some surface point mutations are invasive to structural and conformational properties that determine the enzymatic reaction. This phenomenon was pronounced for mutations introduced at R850, and not for A539, as depicted in Figure 3.7.

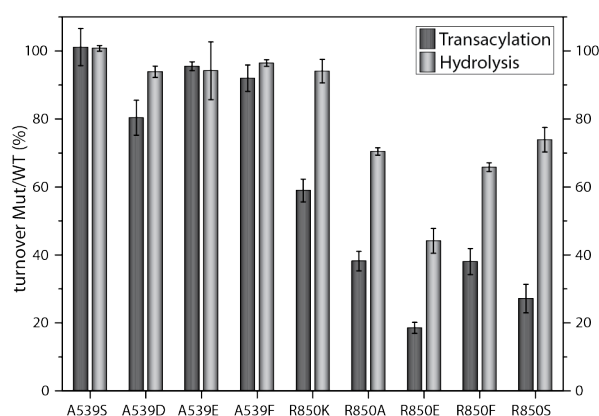


Figure 3.7: Transacylation (dark gray) and hydrolysis (light gray) screening of the different DEBS3M5 AT:ACP point mutants. The transacylation activity was measured in technical triplicates of one biological replicate; hydrolysis was measured in biological (A539E, R850K, R850E, R850S) or technical triplicates (A539S, A539D, A539F, R850A, R850F). The average activity of each mutant (Mut) is divided by wild type's (WT) activity and is given in %. Error bars indicate the standard deviation of technical or biological triplicates. A539 mutants seem to have no influence on transacylation and hydrolysis, whereas R850 mutants decrease both transacylation and hydrolysis rates significantly.

Based on the initial screening, four mutations were chosen for full kinetic characterization to determine absolute kinetic constants; i.e., A539E, which showed nearly no effect in screening, and the mutations R850K, R850E, and R850S, which showed moderate to strong effects, preserving, inverting, and neutralizing the charge of the mutated residue. For protein quality control of proteins in biological replicates, see Table S8 and Figure S10, S11, and S12. We note that the ACP concentrations for the enzymatic analysis of mutant R850E were within a non-ideal range of $1-22 \times K_m^{ACP}$ due to a low K_m^{ACP} value and the kinetic transacylation parameters for this mutant were just determined at lower confidence.

Table 3.3: Transacylation and hydrolysis kinetic parameters determined for DEBS3 AT5 wild type and mutants with MMal-CoA. Measurement performed in biological triplicates. Overall, transacylation and hydrolysis kinetic parameters are in good accordance and the catalytic efficiencies give similar values for each mutant, as it should be. No mutation is wild type AT (DEBS3 AT5).

Transacylation		Hydrolysis						
Mutation	k_{cat} (s^{-1})	$K_{\text{m}}^{\text{MMal-CoA}}$ (μM)	$K_{\text{m}}^{\text{ACP}}$ (μM)	$k_{\text{cat}}/K_{\text{m}}^{\text{MMal-CoA}}$ ($\text{s}^{-1}\text{M}^{-1}$)	$k_{\text{cat}}/K_{\text{m}}^{\text{ACP}}$ ($\text{s}^{-1}\text{M}^{-1}$)	k_{cat} (s^{-1})	$K_{\text{m}}^{\text{MMal-CoA}}$ (μM)	$k_{\text{cat}}/K_{\text{m}}^{\text{MMal-CoA}}$ ($\text{s}^{-1}\text{M}^{-1}$)
-	$1.25 \pm 7.11 \times 10^{-2}$	209 ± 17.5	72.0 ± 8.21	6.0×10^3	1.7×10^4	$5.31 \times 10^{-2} \pm 9.62 \times 10^{-4}$	$7.88 \pm 3.98 \times 10^{-1}$	6.7×10^3
A539E	$1.02 \pm 5.19 \times 10^{-2}$	164 ± 13.5	74.3 ± 7.96	6.2×10^3	1.4×10^4	$5.16 \times 10^{-2} \pm 1.15 \times 10^{-3}$	$9.99 \pm 6.06 \times 10^{-1}$	5.2×10^3
R850K	$7.10 \times 10^{-1} \pm 5.82 \times 10^{-2}$	121 ± 16.3	73.2 ± 12.7	5.9×10^3	9.7×10^3	$5.38 \times 10^{-2} \pm 1.59 \times 10^{-3}$	$11.2 \pm 8.75 \times 10^{-1}$	4.8×10^3
R850E	$6.82 \times 10^{-2} \pm 3.04 \times 10^{-3}$	75.5 ± 7.65	$2.38 \pm 4.08 \times 10^{-1}$	9.0×10^2	2.9×10^4	$3.59 \times 10^{-2} \pm 1.79 \times 10^{-3}$	37.7 ± 5.37	9.5×10^2
R850S	$2.82 \times 10^{-1} \pm 1.41 \times 10^{-2}$	140 ± 11.7	46.0 ± 4.55	2.0×10^3	6.1×10^3	$5.32 \times 10^{-2} \pm 2.94 \times 10^{-3}$	25.9 ± 3.60	2.1×10^3

Results

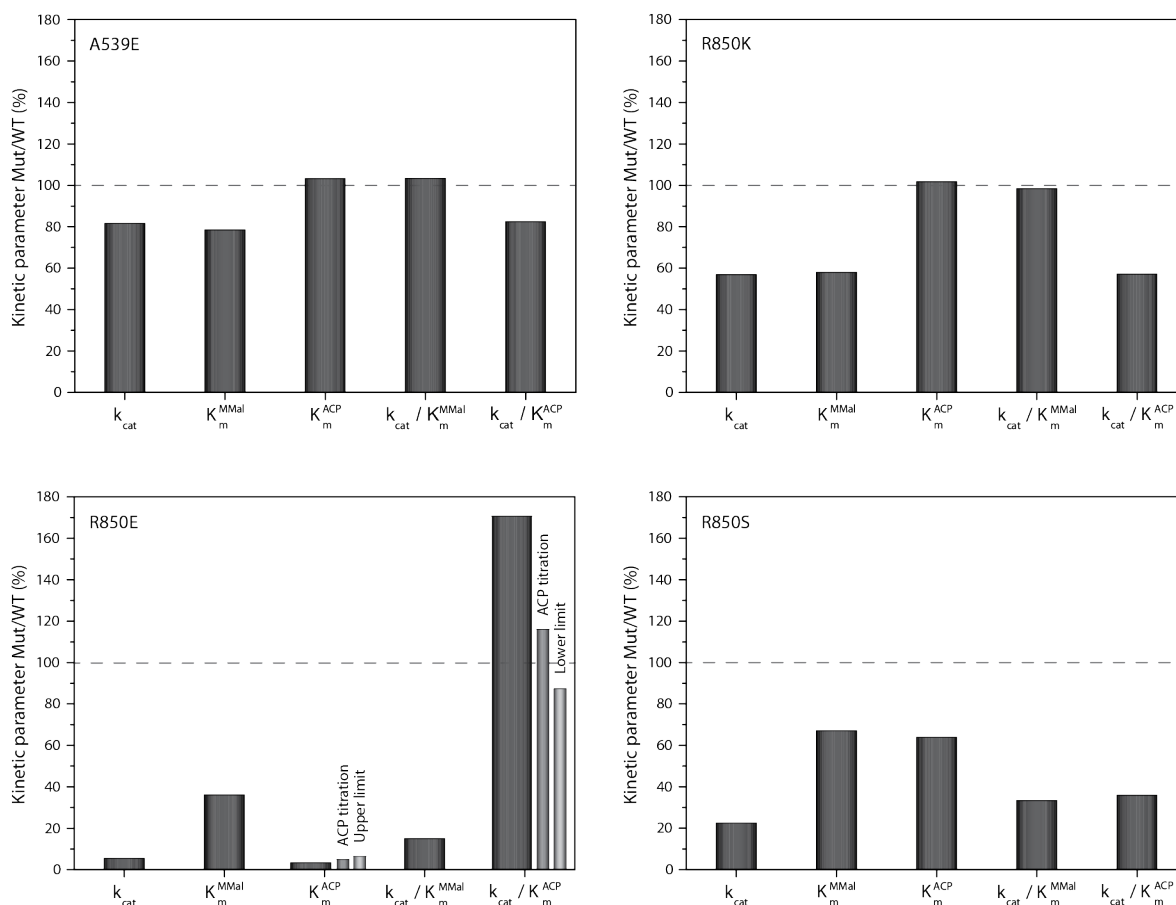


Figure 3.8: Ratio of kinetic constants determined for transacylation catalyzed by mutants (Mut) A539E, R850K, R850E, and R850S to wild type DEBS3M5 AT (WT) kinetic constants (in %). R850E shows three ratios of K_m^{ACP} Mut/WT corresponding to the different K_m^{ACP} received for this mutant: value determined via global fit (ratio shown in dark gray), determined via ACP titration without lowest ACP concentration (ratio shown in gray), and the upper limit of the value (ratio shown in light gray). Accordingly, this gives three corresponding ratios for the catalytic efficiency k_{cat}/K_m^{ACP} .

Trends in transacylation and hydrolysis turnover rates were in good accordance with the findings of the initial screening (Tab. 3.3 and Fig. 3.8). For hydrolysis and transacylation titration curves, see Figure S13 and S14. Mutant A539E showed comparable transacylation rates to the wild type protein (82 % of wild type). Overall, kinetic characterization indicates that the site A539 is rather not involved in the AT-ACP and the AT-MMal-CoA interaction or mutations are not invasive to the AT-ACP interaction. The kinetics of the transacylation reaction is essentially unaltered compared to the wild type, as can also be seen in transition state energies for the ping step and pong step remaining at wild type level (Fig. 3.9). In contrast, R850 mutants were significantly compromised in transacylation activity (R850K with 57 % of wild type, R850E 5.5 %, and R850S 23 %).^{vi} Hydrolysis rates remained unaltered by the mutations A539E, R850K, and R850S, but dropped to 68 % of wild type rates for mutation R850E. We note (again) that the transacylation rate always include the hydrolysis rate to some extent.

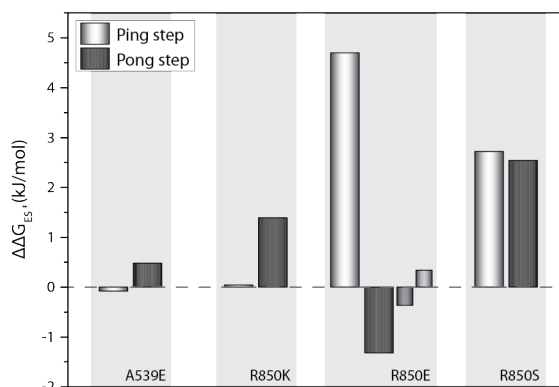


Figure 3.9: Difference in transition state energy (in kJ/mol) between mutants A539E, R850K, R850E, and R850S and wild type determined for the ping step (white bars) and the pong step (dark gray bars) of transacylation. Mutant R850E gives three values for K_m^{ACP} : one determined via global fit, one via ACP titration (gray), and the upper limit (light gray). Differences in transition state energy to wild type are shown in dark gray, gray, and light gray, respectively.

$$\Delta\Delta G_{ES}^{\ddagger} = -RT \ln \left[\frac{(k_{cat}/K_m)^{Mut}}{(k_{cat}/K_m)^{WT}} \right]$$

with the gas constant R , $T=298$ K and the ratio of catalytic efficiencies between mutant (Mut) and wild type (WT).

^{vi}For R850E, kinetic parameters of the pong step were only collected with lower confidence due to the low K_m^{ACP} . For a reliable estimate on K_m^{ACP} , we extracted data for saturated substrate MMal-CoA concentration titrated with ACP and received a K_m^{ACP} of 3.46 μ M and an upper boundary of 4.6 μ M (Fig. S15).

3.4 Generation of a FAS-Based Chimeric Elongation Module

Based on kinetic analysis of ATs, we selected ATs for construction of a FAS-based elongation module via domain swapping. The most important criterion was the strict specificity for a single elongation substrate. Only the three ATs from modular PKSs fulfill this requirement. For more variability, we aimed to design two elongation modules that selectively load Mal and MMal, respectively. Among the three elongation ATs from modular PKSs, only RAPS3 AT14 transfers Mal. For the MMal-specific AT, we had to decide between DEBS3 AT5 and PikAIII AT5. The advantage of the available crystal structure of the first outweighs the faster transacylation rates of the latter. Thus, in a first domain swapping attempt, we used RAPS3 AT14 and DEBS3 AT5.

3.4.1 Chimeric mFAS with Swapped DEBS AT5 and RAPS AT14

To identify proper domain boundaries for the AT swap, we used the crystal structures of mFAS and DEBS3M5 KS-AT (Fig. 3.10A and B) and created a structural alignment. Additionally, we created a sequence alignment to analyze and consider conserved areas and to transfer structural insights on RAPS3 AT14, for which no structural information is available. DEBS3M5 and RAPS3M14 KS-AT have high sequence identities of 31 % and 30 % with mFAS KS-AT, respectively. DEBS3M5 and RAPS3M14 KS-AT exhibit an even higher sequence identity of 48 %.

As described before, KS and AT are connected via an LD (see section 2.6.4). Both analyzed didomain structures are remarkably similar, although the overall structural alignment gives a high root-mean-square deviation (RMSD) of 4.824 Å (721 atoms). Alignment of solely the KS and solely the AT domains reveals the high similarity of the single domains, giving an RMSD of 1.221 Å (318 atoms) and 1.428 Å (243 atoms), respectively (Fig. 3.10C and D). Besides, also the orientation of the AT towards the KS is almost the same. In contrast, the LD differs strongly in both structures. Structural alignment (not shown) results in a rotation of around 120° and an RMSD of 10.111 Å (41 atoms). The overall high similarity of the condensing part in both megasynthases is promising for the domain swapping approach.

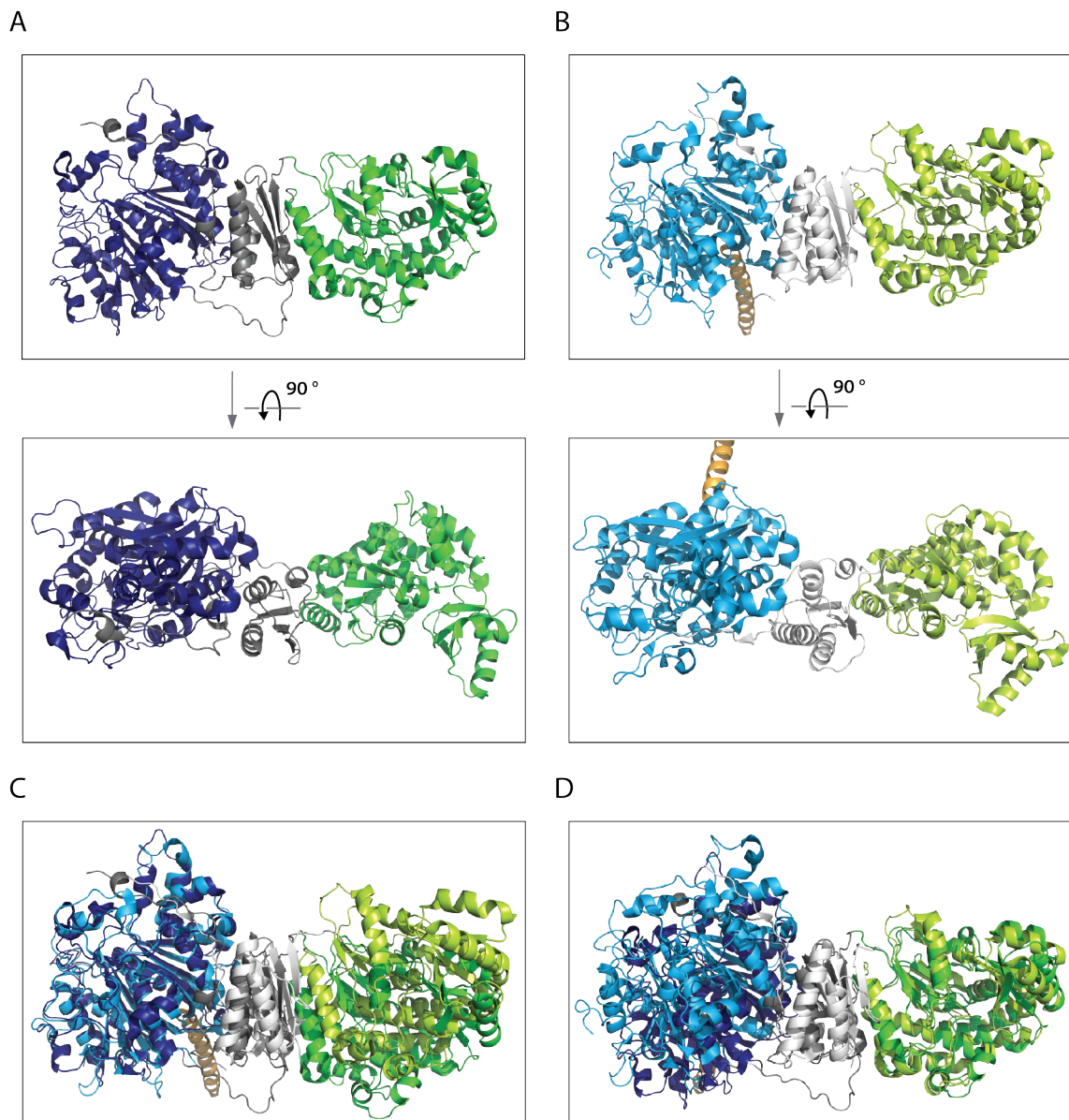
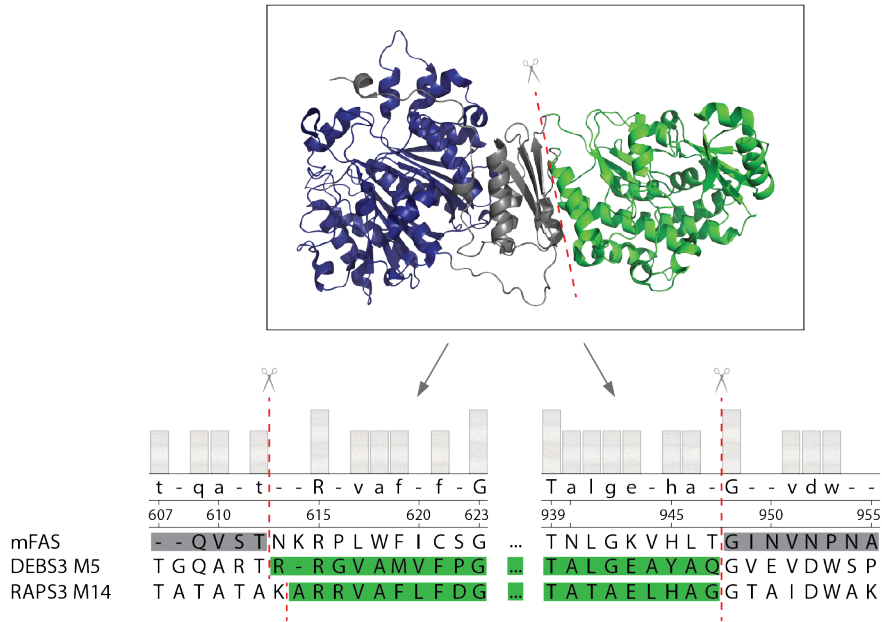


Figure 3.10: Structural similarity between the KS-AT didomain from mFAS (PDB: 5MY0^[9]) and DEBS3M5 (PDB: 2HG4^[99]). (A) Cartoon representation of the KS-AT didomain from mFAS linked by the LD. KS, LD, and AT shown in dark blue, dark gray, and green, respectively. (B) Cartoon representation of the KS-AT didomain from DEBS3M5 linked by the LD. KS, LD, and AT depicted in light blue, light gray, and lemon, respectively. (C) Structural alignment of solely the KS domains reveals their high similarity (RMSD = 1.221 Å). mFAS and DEBS3M5 KS-AT shown in dark and bright colors, respectively. (D) Structural alignment of solely the AT domains shows their high resemblance (RMSD = 1.428 Å). mFAS and DEBS3M5 KS-AT shown in dark and bright colors, respectively.

Results

Regarding the structures of the KS-AT didomains, one can identify two possible interfaces for an AT swap: The first approach exchanges solely the AT domain, creating an artificial LD:AT interface (boundaries between KAL and AT and between AT and PAL). The second strategy is to swap the AT domain with its LD, resulting in a new KS:LD interface (boundaries between KS and KAL and within PAL). A recent study showed, that the interface between the KS and LD is better suited for AT swapping approaches.^[20] However, every system has its own characteristic properties and might respond differently to changes in certain interfaces. Thus, we decided to design constructs with swapped AT and with swapped LD-AT, based on the structural alignment and the sequence alignment of the three KS-AT didomains (Fig. 3.11). For the AT swap, we chose the same domain boundaries based on both alignment types. For the LD-AT swap, the boundaries we chose based on the structure and sequence alignment differed considerably in the second domain boundary within the PAL (PKS PAL:mFAS PAL). Hence, it seemed reasonable to design two different constructs with modified boundaries – one following the structural alignment, one following the sequence alignment.

A



B

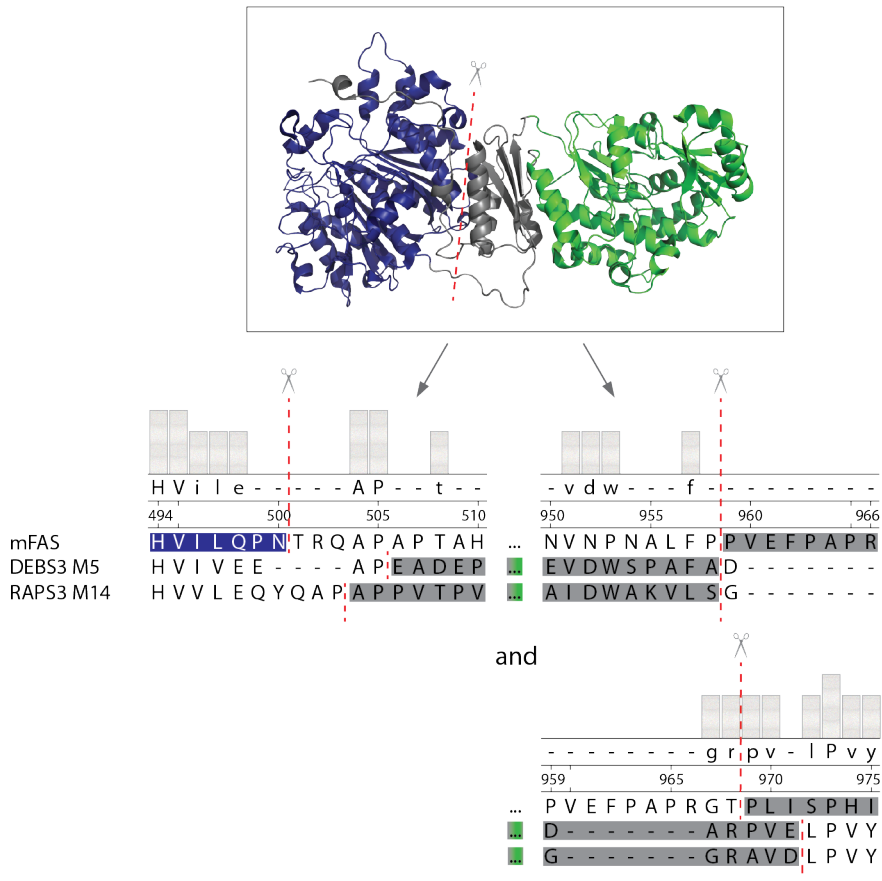


Figure 3.11: Domain boundaries for the AT swapping in mFAS. Cartoon representation of mFAS (PDB: 5MY0^[9]) schematically shows the artificial interfaces introduced in domain swapping. KS, LD, and AT depicted in blue, gray, and green, respectively. Excerpt of alignment depicts the exact artificial domain boundaries. (A) Domain boundaries for the exchange of the AT domain in mFAS with DEBS3 AT5 and RAPS3 AT14. Primary sequences of KS-AT wild types show boundaries between KAL

Results

and AT as well as between AT and PAL. (B) Domain boundaries for the exchange of the LD-AT domain in mFAS with DEBS3 (LD-AT)5 and RAPS3 (LD-AT)14. Primary sequences of KS-AT wild types show boundaries between KS and KAL as well as within the PAL. Two alternative boundaries were chosen for the latter – the upper boundaries are based on the sequence alignment, the lower ones on the structure alignment.

Table 3.4: Test expression yields of chimeric mFAS with DEBS3 AT5 and RAPS3 AT14 after tandem affinity chromatography of 1 L expression culture.

Domain Swapped	Pellet's Weight (g)	Total Yield (μg)	Yield (μg) per g Pellet
DEBS AT	10.9	286	26.2
DEBS (LD-AT)1	10.3	16.6	1.61
DEBS (LD-AT)2	10.4	79.9	7.68
RAPS AT	12.0	56.4	4.70
RAPS (LD-AT)1	10.8	14.0	1.30
RAPS (LD-AT)2	15.2	86.6	5.70

The chimeric constructs were expressed in *E. coli* and purified by tandem affinity chromatography. All proteins showed drastically reduced yields (Tab. 3.4) compared to the wild type mFAS (wild type yield around 1-2 mg/L^[17,222]). Proteins were analyzed by SDS-PAGE (Fig. 3.12A). The first construct with swapped LD-AT from DEBS3 AT5 could not be isolated at all in full-length (exp. mass of 277 kDa), but showed some other protein bands most likely due to protein degradation. All other proteins were successfully isolated (exp. masses of 276-278 kDa) as indicated by the protein bands around 280 kDa. All elution fractions contained other bands of lower molecular weight proteins. Since purification via tandem affinity chromatography usually results in pure protein,^[1,17] these are most likely degraded proteins. The band of 170 kDa could correspond to the protein without KS-AT, the band at 60 kDa to KS-LD, and the band of 45 kDa to KS or LD-AT domains. Western blot analysis could be used to give additional information. The construct with swapped DEBS3 AT5 showed an additional protein band of around 100 kDa after the second affinity chromatography, which was not present in the elution fraction of the first (Fig. S16). This was definitely old protein bound to the second column and most probably a KS-AT construct. Therefore, the protein yield for this chimeric construct could not be determined precisely, but should lie within the double-digit μg range.

Even though the chimeras were only isolated once, I observe the following: Proteins with swapped LD-AT seem to demonstrate the value of structural information in domain swapping approaches. Constructs designed on the basis of the structural alignment gave higher yields than constructs with domain boundaries chosen based on the sequence alignment.

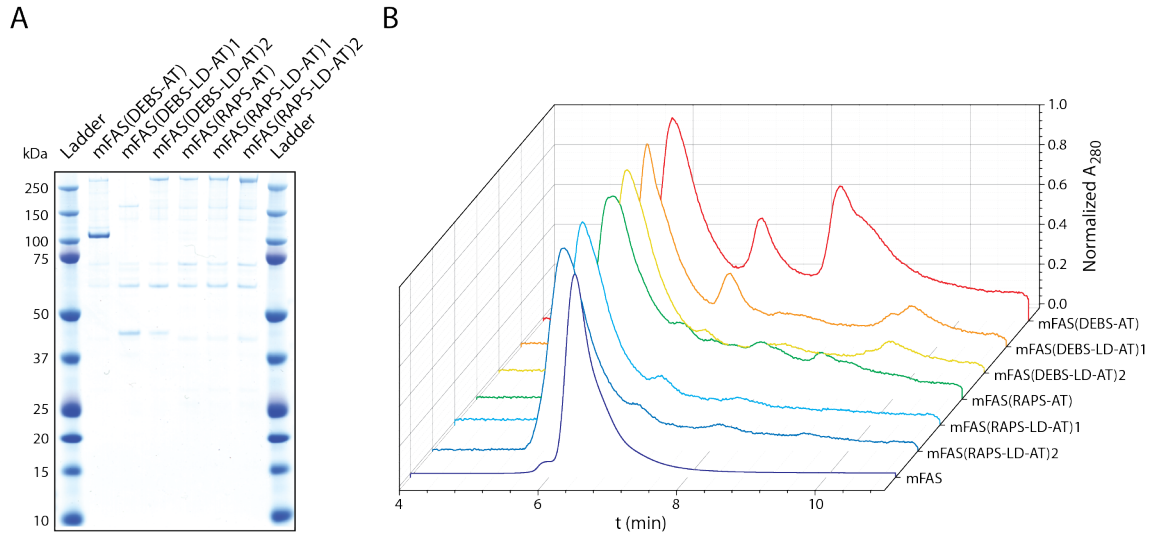


Figure 3.12: Analysis of chimeric mFAS constructs. (A) SDS-PAGE (NuPAGE Bis-Tris 4-12 %) of chimeric full-length mFAS. Elution fractions after tandem affinity chromatography. (B) Normalized size exclusion chromatogram of wild type mFAS and chimeric constructs.

Proteins were further analyzed regarding their oligomeric states via high performance liquid chromatography (HPLC)-SEC (Fig. 3.12B). The wild type mFAS solely formed dimeric species. The chimeric constructs, however, showed mainly higher oligomeric states and only partly dimeric states. Impurities of lower molecular weight proteins and degradation were observed in their chromatograms.

Overall, this first attempt to create a chimeric mFAS module illustrates the complexity of domain swapping and shows that fine-tuning might be needed to gain a functional protein in good yield and quality. Thus, we decided to focus on only one construct – the domain swapping with the AT domain from DEBS3M5, since structural information is available. Furthermore, we decided to reduce the complexity by working first on the didomain construct KS-AT, which is easier to modify, isolate, and analyze. After successful design of a stable chimeric didomain, this construct can be expanded to the full-length mFAS elongation module.

3.4.2 Generation of mFAS/DEBS3M5 KS-AT Chimeras

Analogously to the full-length chimeras, two chimeric KS-AT didomains were designed consisting of mFAS KS-LD with DEBS3 AT5 (FS37) and mFAS KS with DEBS3 (LD-AT)5 (FS38), respectively. Since the first LD-AT swap was not successful in full-length mFAS, only the second domain boundaries were used. It became obvious, that slight changes in domain boundaries can totally change the outcome of the chimera. Thus, we chose to design alternative constructs. The two mentioned chimeras (FS37 and FS38) served as starting point for the design of six others. Following Yuzawa *et al.*'s strategy to design active chimeras,^[20] we modified the domain boundaries as follows, which is described schematically in Figure 3.13A:

1. One approach modifies the second boundary and keeps (only) the first part of the DEBS PAL (PAL1). In the AT only swap, this leads to a slightly shortened mFAS PAL and addition of the DEBS PAL1 (FS57). In the AT-LD swap, this involves a slightly lengthened mFAS PAL and a shortened DEBS PAL (FS60).
2. The second approach consists of modifying the first boundary. For the AT only swap, the mFAS KAL was slightly lengthened (FS58). For the LD-AT swap, the DEBS KAL was slightly lengthened (FS61).
3. The third approach is the combination of both (FS59 and FS62).

Sequence information of modified domain boundaries is shown in Figure 3.13B.

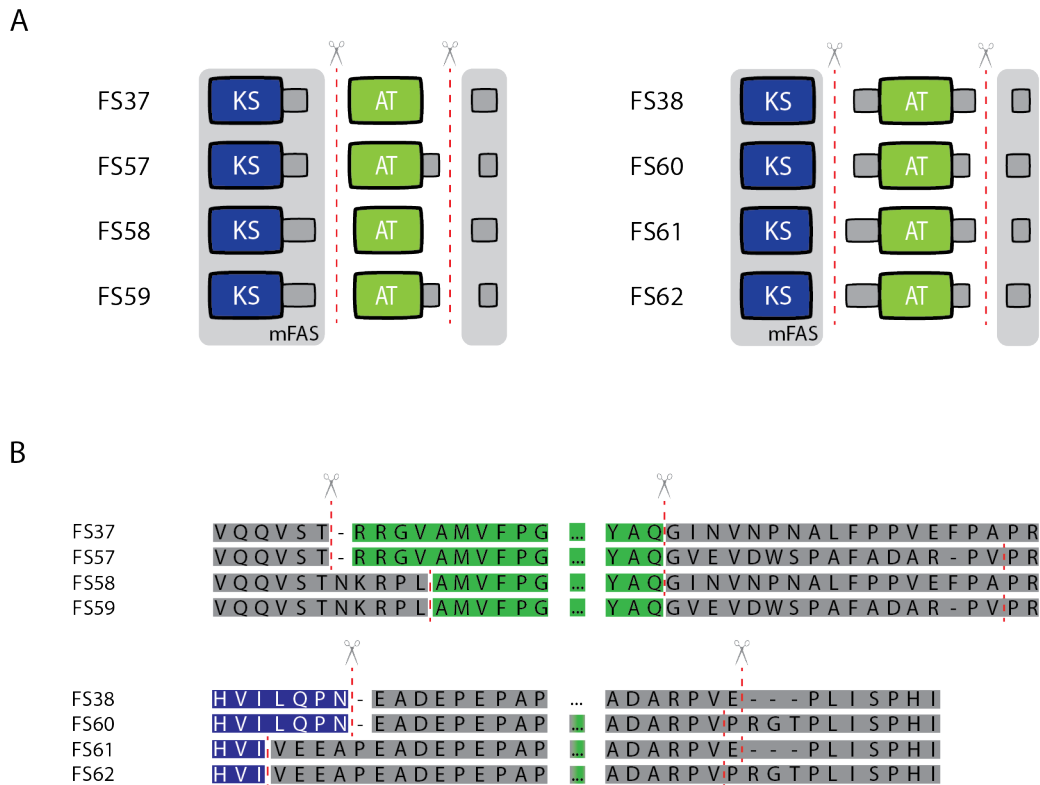


Figure 3.13: Design of mFAS/DEBS3M5 KS-AT chimeras. (A) Schematic representation of the KS-AT chimeras with modified domain boundaries. The linkers KAL and PAL are adapted in length to optimize the artificial interactions. (B) Domain boundaries in the KS-AT chimeras. Excerpt of the primary sequence of KS-AT chimeras shows the variation in linker regions.

Cloning, expression, purification, and analysis of chimeras were performed by Maren Berlinghof in the course of her master's thesis under my supervision.^[223] All eight chimeras were successfully isolated, but yields were drastically reduced compared to wild type yields (Tab. 3.5, mFAS KS-AT 16 mg/L, DEBS3M5 KS-AT 14 mg/L^[1]). Analytical SDS-PAGE confirmed successful isolation of all chimeras (Fig. 3.14), but also showed some degraded proteins similar to the chimeric full-length mFAS. The band of 60 kDa could correspond to KS-LD and the band of 45 kDa could correspond to KS. Judging from this test expression, domain swapping of AT with LD seemed to be more effective than solely swapping the AT domain, which is in agreement to Yuzawa *et al.*'s findings.^[20] In case of the AT only swap, keeping the first part of PAL from the AT system did not improve expression yields and indicates that there are almost no important stabilizing interactions between AT and PAL. In case of LD-AT swap, the yield was significantly reduced by shortening the PAL to PAL1 from the AT system. This indicates that there are also important interactions between LD and PAL2. In case of single AT swapping, the modified domain boundaries between LD and AT seemed to be better suited, but conclusions cannot be drawn without uncertainty due to only single protein preparations and to the low significance of yields. In case of LD-AT swapping, the original boundaries between KS and LD seemed to be best suited.

Results

Table 3.5: Test expression yields of KS-AT chimeras after tandem affinity chromatography of 1 L expression culture.

Protein	Pellet's Weight (g)	Total Yield (μg)	Yield (μg) per g Pellet
FS37	8.11	73.3	9.04
FS57	8.62	89.1	10.3
FS58	7.60	131	17.2
FS59	7.54	106	14.0
FS38	8.11	506	62.4
FS60	6.82	101	14.8
FS61	6.36	164	25.8
FS62	7.58	171	22.6

Without further purification, proteins were analyzed via TSA. Impurities may have affect the outcome, but trends in protein stability were still observed. Due to low yields, only the proteins FS37, FS38, FS58, and FS59 resulted in significant and analyzable melting curves (Fig. 3.14B). The fluorescence signal of all chimeras was already high from the start, indicating unfolded or degraded protein.^[224] Surprisingly, FS37 seemed to be significantly more stable than even the wild type proteins according to the determined melting temperatures. This can be explained by the fact that protein degradation may lead to stable fragments.^[225] Interestingly, the other chimeric proteins seemed to reflect the wild type with higher portion of composition. FS58 and FS59 with solely AT exchanged consist mainly of mFAS parts. Both proteins showed similar melting behavior as the wild type mFAS KS-AT. FS38, on the other hand, with LD-AT from DEBS3M5 resembled its wild type behavior. More data on the melting behavior of the other chimeras would be preferable in order to evaluate if this is an actual consistent trend.

Furthermore, proteins were analyzed in their oligomeric states via HPLC-SEC (Fig. 3.14C and D). Due to low yields, only a limited amount of protein could be analyzed leading to a suboptimal signal-to-noise ratio for proteins FS57 and FS60-FS62. For these chimeras, mainly aggregates and higher oligomeric states were detected, besides the desired dimeric state. FS37 and FS38 showed aggregates, higher oligomeric, and monomeric states. From these two, only FS38 showed some dimeric species. Regarding SEC, only proteins FS58 and FS59 showed no aggregates and mainly dimeric species, as well as almost no higher oligomeric states. These two proteins seemed to be the only chimeras forming proper quaternary structures.

Overall, we observed a drastic decrease in expression yield. The introduction of artificial interfaces had serious effects on folding and formation of quaternary structures. Impurities may have contributed in our measurements. However, proteins were found to be quite stable.

Fine-tuning might be needed in order to address these issues and improve protein-protein interactions. This requires a detailed understanding of both proteins. Thus, the KS-AT wild types were studied by mutagenesis.

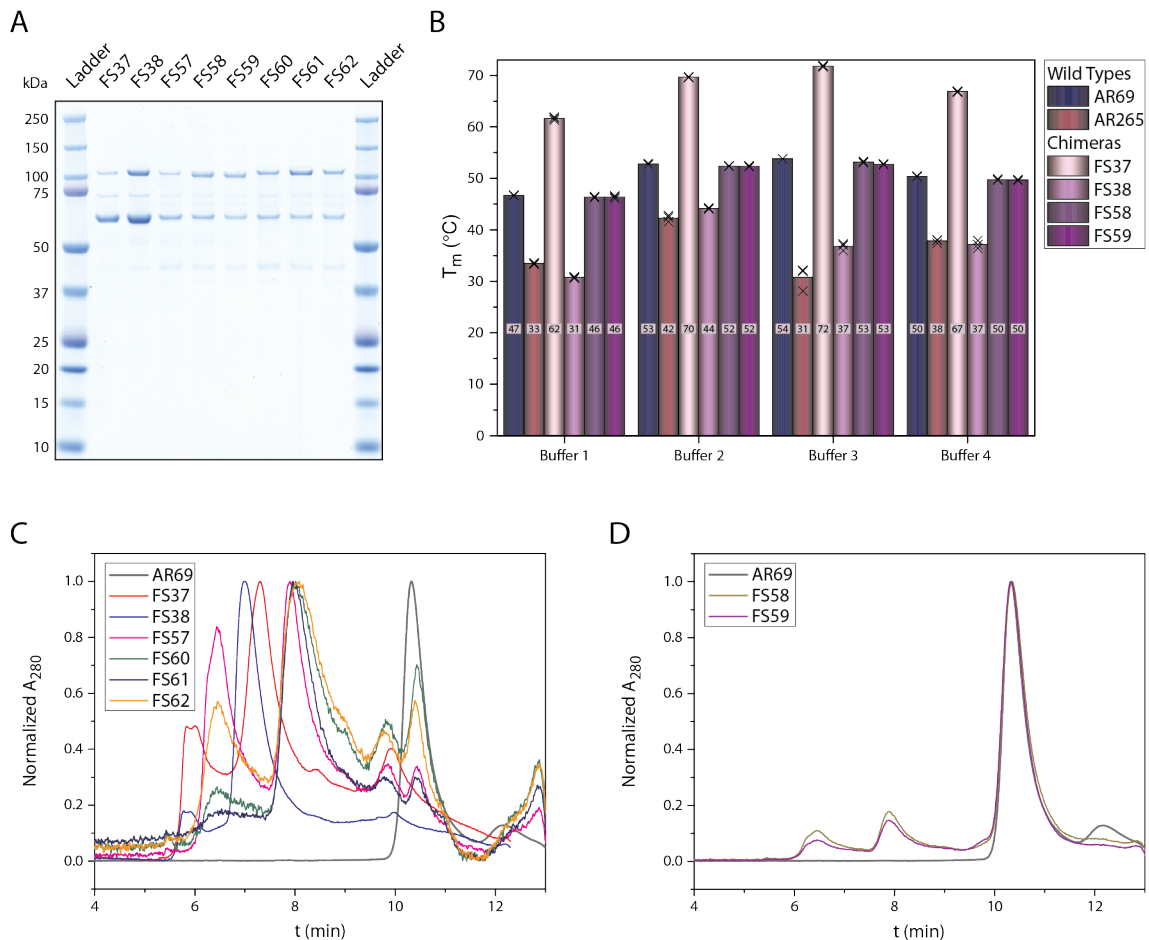


Figure 3.14: Analysis of KS-AT chimeras. (A) SDS-PAGE (NuPAGE Bis-Tris 4-12 %) of mFAS/DEBS3M5 chimeric KS-AT. Elution fractions after tandem affinity chromatography. (B) Average melting temperatures (in °C) of mFAS and DEBS3M5 KS-AT wild types (AR69 and AR265) and of chimeric KS-AT constructs (FS37, FS38, FS58, FS59) in different buffers. Measured in triplicates. Crosses correspond to one biological replicate. Buffer 1: water, Buffer 2: Strep wash buffer pH 7.6, Buffer 3: Strep wash buffer pH 6.0, Buffer 4: 40 mM Tris, pH 7.6, 1 mM EDTA, 10 % glycerol. (C) and (D) Exemplary normalized size exclusion chromatogram of wild type AR69 and chimeric KS-AT constructs FS37, FS38, and FS57-FS62.

3.4.3 Analysis of mFAS KS-AT Wild Type

In order to identify hotspot residues which contribute to the formation of stable domain:domain interfaces in the KS-AT construct, single point mutations were introduced and their effects on protein's stability were analyzed. Expression yield, melting behavior, and oligomeric states served as readout.^[17,20,225] Different bioinformatics tools were used to predict interfacial hotspots using the mFAS KS-AT structure (chains A and/or D, PDB: 5MY0, see

Results

section 5.1.14). Based on these predictions, the following residues were chosen as alanine mutation sites: D61 and N96 (residue numbering following UniProt entry P19096) located in the KS, interacting with the LD (Fig. 3.15). H425, R429, I457, R468, Y470, V479, E481, and Q483 located in the LD and interacting with KS or AT. H804 and L805 in the AT domain, interacting with the LD. Predicted effects are summarized in Table 3.6. Dr. Karthik Paithankar performed the Rosetta calculations on the mFAS KS-AT wild type using chains A and D of the solved structure.^[9]

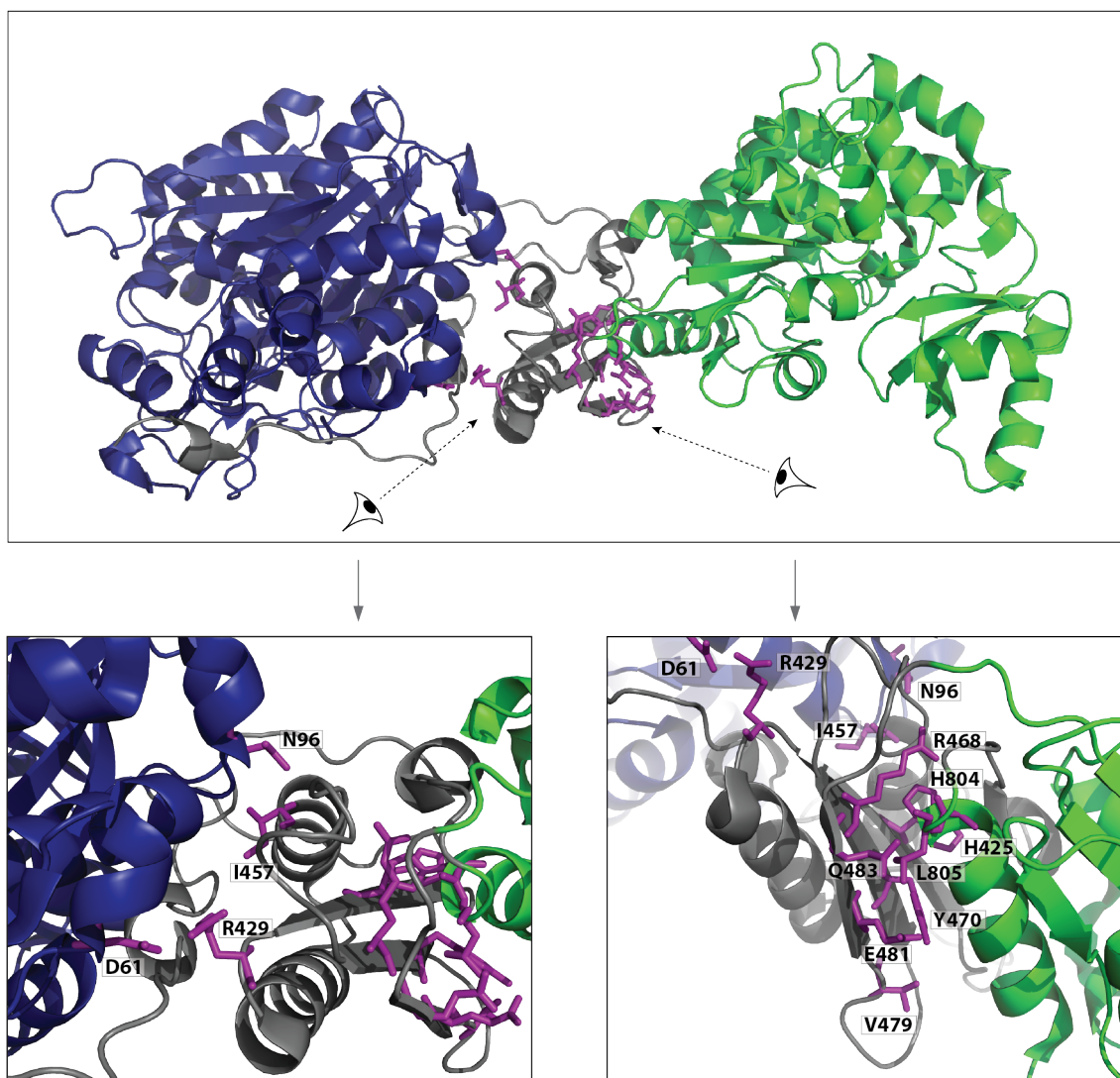


Figure 3.15: Sites for mutagenesis in mFAS KS-AT. Cartoon representation (PDB: 5MY0^[9]) with KS, LD, and AT depicted in blue, gray, and green, respectively. Interfacial residues between KS and LD as well as LD and AT chosen for mutagenesis are highlighted in purple. Residue numbering following UniProt entry P19096.

Table 3.6: Results of bioinformatic analysis of mFAS KS-AT. Difference in the Gibbs free energy ($\Delta\Delta G$) between mutant and wild type (in Rosetta energy unit (REU) and kcal/mol) predicted by different bioinformatics tools. For better comparison of results from different bioinformatics tools, values of mCSM are handled as described in section 5.1.14. SpotOn predicts mutations to be a Nullspot (not relevant for the interaction) or a Hotspot (relevant for the interaction). Residue numbering following UniProt entry P19096.

Program	Rosetta				Robetta	DrugScore	mCSM	SpotOn
Chain	A		D		A			A
PAL	w	w/o	w	w/o	w/o		w	w/o
Protein	$\Delta\Delta G$ (REU)				$\Delta\Delta G$ (kcal/mol)			Hot/ Null
FS39 (D61A)	7.34	4.59	6.63	3.84	1.52	1.98	1.05	Null
FS40 (N96A)	5.71	5.71	6.20	6.20	3.26	2.11	0.181	Null
FS41 (R429A)	3.34	3.64	3.60	3.92	3.02	1.04	0.263	Null
FS42 (I457A)	4.86	3.10	5.19	3.44	1.61	2.17	0.424	Hot
FS43 (R468A)	3.34	3.68	3.53	3.53	1.70	2.07	0.987	Null
FS44 (Y470A)	2.11	2.11	3.51	3.51	2.75	4.28	0.537	Hot
FS51 (H425A)	0.778	0.778	2.18	2.18	1.11	0.30	0.53	Null
FS52 (V479A)	3.34	3.43	3.48	3.48	0.32	0.42	0.49	Hot
FS53 (E481A)	3.17	3.17	4.16	4.16	0.59	0.09	0.68	Hot
FS54 (Q483A)	-4.62	-4.62	3.22	-2.28	0.96	0.18	0.159	Hot
FS55 (H804A)	0.730	0.731	0.797	0.797	3.97	0.45	1.433	Null
FS56 (L805A)	-2.50	-2.49	0.243	-0.243	2.69	2.21	0.42	Null

Results

Maren Berlinghof cloned the mutants and isolated the wild type AR69 and mutants FS39-FS44 as well as FS51-FS56 in biological triplicates during her master's thesis under my supervision.^[223] Overall, introducing alanine mutations decreased expression yields (Tab. 3.7). Mutations at R468, Y470, and L805 (FS43, FS44, and FS56) showed only minor effects and gave wild type-like yields. Especially mutations introduced at D61 and I457 (FS39 and FS42) seemed to strongly affect protein expressibility. All other proteins showed around 50 % protein yield compared to the wild type. Overall, protein expression was clearly influenced, but yields of most proteins were still in a good range.

Table 3.7: Average expression yields of mFAS KS-AT wild type (AR69) and mutants (FS39-FS44, FS51-FS56) after tandem affinity chromatography of 1 L expression culture. Yields determined from three independent expressions.

Protein	Pellet's Weight (g)	Total Yield (mg)	Yield (mg) per g Pellet
AR69	9.17	15.7	1.71
FS39	8.02	0.332	0.0414
FS40	8.06	7.33	0.909
FS41	7.73	6.05	0.783
FS42	8.86	2.56	0.289
FS43	9.11	12.2	1.34
FS44	7.53	10.1	1.34
FS51	8.23	7.74	0.940
FS52	7.97	7.89	0.990
FS53	7.44	8.51	1.14
FS54	8.03	8.32	1.04
FS55	7.25	8.46	1.17
FS56	7.43	9.83	1.32

SDS-PAGE analysis showed a single distinct protein band for full-length protein (Fig. 3.16). FS39 (D61A) showed an intense additional band of around 60 kDa, which was also found for FS42 (I457A) to a much lesser extent. Mutation of D61 (and I457 to a lesser extent) seemed to destabilize the protein, leading to protein degradation, as observed in SDS-PAGE. A band of the same size was also observed in SDS-PAGE of chimeric KS-AT and chimeric full-length mFAS constructs and could correspond to KS-LD. The other mFAS KS-AT mutants showed wild type-like behavior in SDS-PAGE. This indicates that amino acids D61 and I457 might be crucial for the KS-AT fold, whereas mutation of the other residues does not show drastic effects on protein stability. Interestingly, the same protein degradation was observed in chimeras when swapping solely the AT domain, preserving the natural KS-LD interaction.

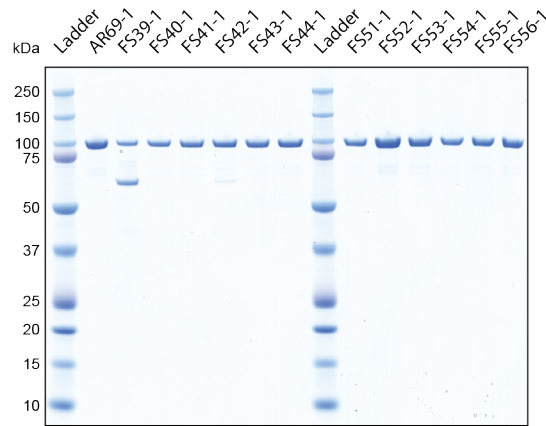


Figure 3.16: Exemplary SDS-PAGE analysis (NuPAGE Bis-Tris 4-12 %) of one biological replicate of mFAS KS-AT wild type (AR69) and mutants (FS39-FS44, FS51-FS56) of concentrated elution fractions after tandem affinity chromatography.

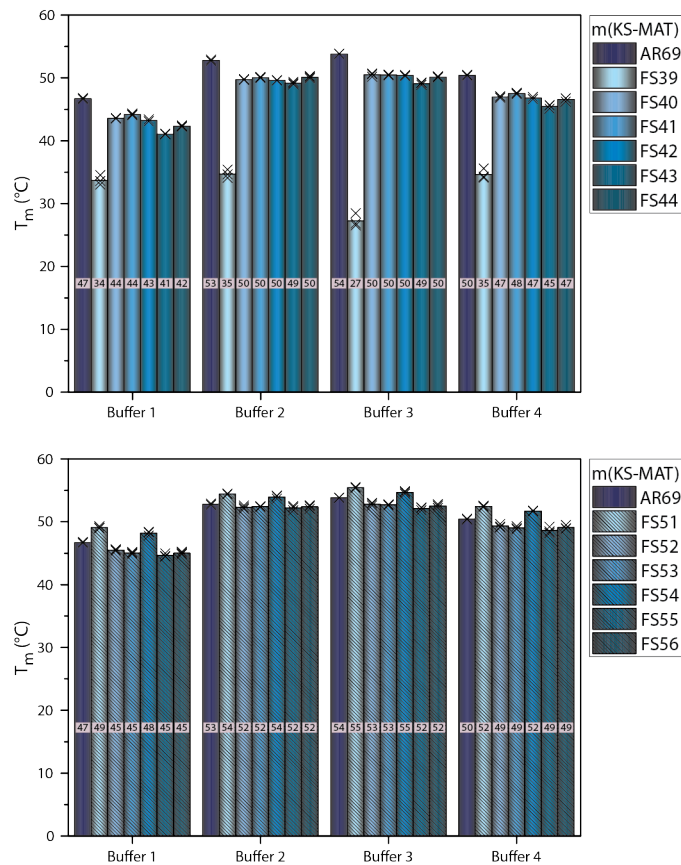


Figure 3.17: Average melting temperatures (in °C) of mFAS KS-AT wild type (AR69) and mutants (FS39-FS44, FS51-FS56) in different buffers. Measured in triplicates. Crosses correspond to one biological replicate. Buffer 1: water, Buffer 2: Strep wash buffer pH 7.6, Buffer 3: Strep wash buffer pH 6.0, Buffer 4: 40 mM Tris, pH 7.6, 1 mM EDTA, 10 % glycerol.

Results

TSA further fortified the role of site D61. While all other mutants showed wild type-like melting behavior in all buffers tested, the melting temperatures were significantly decreased for mutant FS39 (D61A), indicating drastically reduced protein stability (Fig. 3.17). This different melting behavior could be (in parts) caused by degraded protein in the sample. However, in case of the KS-AT chimeras, protein degradation did not lower the melting temperature at all or at least not to this extent. Thus, the stability of the protein seemed to be significantly lowered by this mutation (D61A). Notably, FS42 (I457A) showed wild type-like behavior in the TSA.

Finally, SEC confirmed that all tested amino acids are quite robust to mutation, except for D61 (Fig. 3.18): Only mutant FS39 (D61A) showed different oligomeric states than the wild type, forming mainly higher oligomeric states and monomers, almost no dimeric species was found. In contrast, all other eleven proteins formed mainly the dimeric, partly the monomeric state, just like the wild type AR69. Notably, also FS42 (I457A) shows wild-type like behavior. Apparently the protein is overall quite stable and well folded. Like the SDS-PAGE showed, protein degradation occurred only to a minor extent.

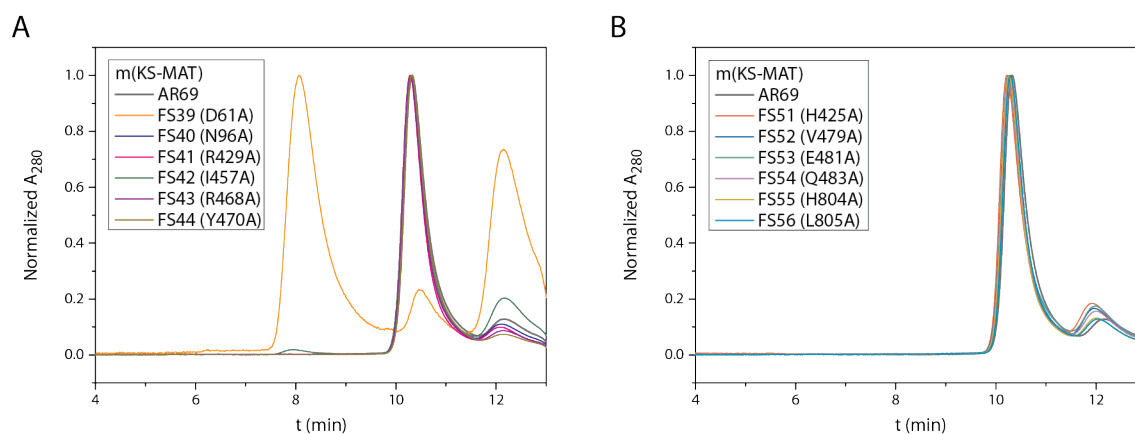


Figure 3.18: Exemplary normalized size exclusion chromatogram of mFAS KS-AT wild type (AR69) and mutants FS39-FS44 (A) and mutants FS51-FS56 (B).

All experiments showed that residue D61 is important for correct folding of mFAS KS-AT. This residue was selected for mutagenesis due to its involvement in the salt bridge between KS and LD domain (Fig. 3.19). Interestingly, mutation of its counterpart R429 to alanine (FS41) did not have the same destabilizing effect.

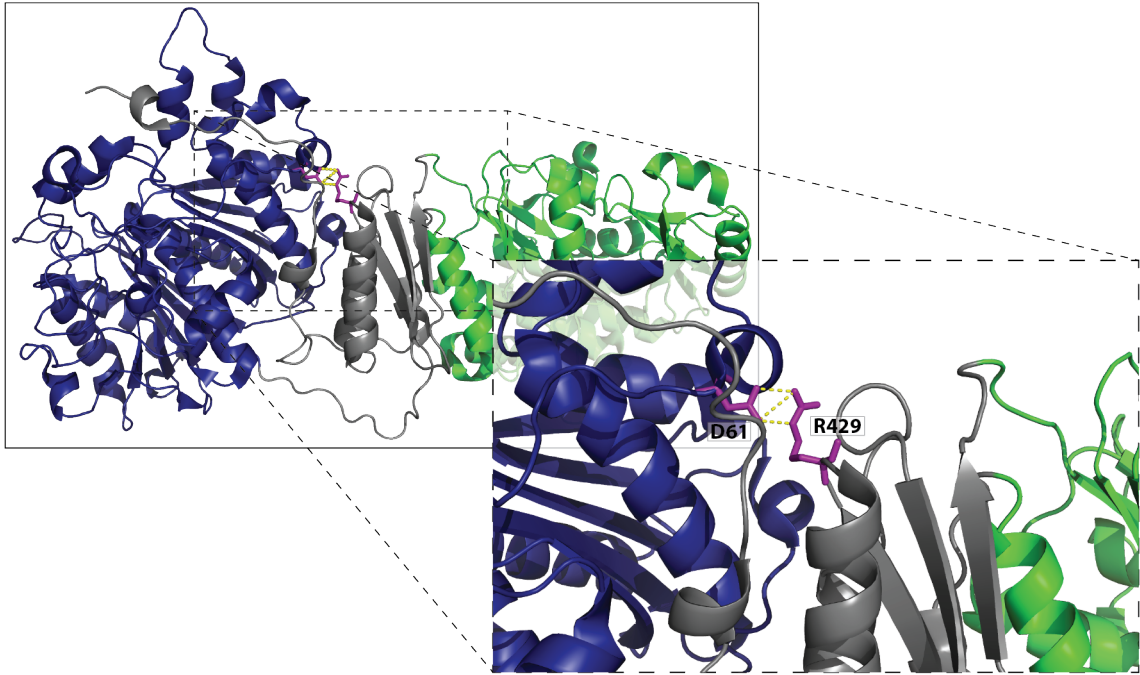


Figure 3.19: Cartoon representation of the interfacial salt bridge formed between D61 and R429 in mFAS KS-AT (PDB: 5MY0^[9]). KS, LD, and AT shown in blue, gray, and green, respectively. Residues D61 and R429 depicted as purple sticks. Salt bridge depicted in yellow and calculated using PyMOL. Residue numbering following UniProt entry P19096.

Overall, the mFAS KS-AT construct seemed to be quite robust to mutations in its interfaces and therefore is a promising candidate for AT swapping approaches. Only mutation D61A destabilized the didomain severely, which was reflected in drastically reduced yield and melting temperatures as well as changes in oligomeric states. Additionally, mutant I457A showed some destabilizing effects, manifested in significantly reduced yields and minor protein degradation, but showed wild type-like behavior in TSA and SEC. Although some other mutants decreased the expression yield, proteins still seemed to be folded correctly and behaved like the wild type in TSA and SEC. Even mutation of R429, D61's counterpart in the salt bridge stabilizing the KS-LD interaction, led to stable protein. These results demonstrate the high robustness of mFAS KS-AT and prove that AT and LD-AT swapping should be possible within the mFAS.

Results

3.4.4 Analysis of DEBS3M5 KS-AT Wild Type

Six different point mutations were chosen within DEBS3M5 KS-AT (Fig. 3.20). Effects of mutations predicted by the different bioinformatics tools using the DEBS3M5 KS-AT structure (chains A and/or D, PDB: 2HG4) are summarized in Table 3.8. Dr. Karthik Paithankar performed the Rosetta calculations on the DEBS3M5 KS-AT wild type using chains A and D of the solved structure.^[99]

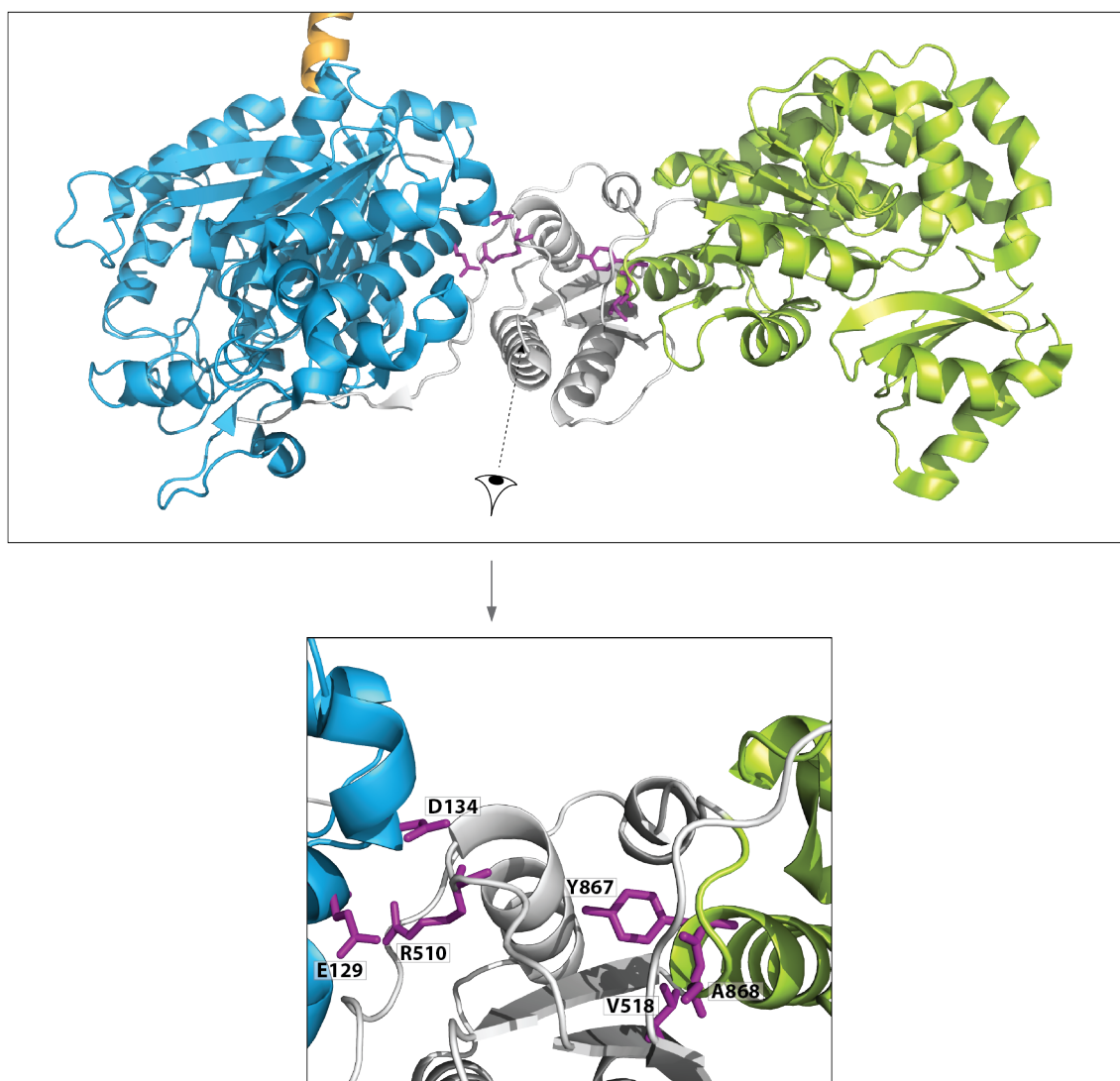


Figure 3.20: Sites for mutagenesis in DEBS3M5 KS-AT wild type. Cartoon representation (PDB: 2HG4^[99]) with KS, LD, and AT depicted in light blue, light gray, and lemon, respectively. Interfacial residues between KS and LD as well as LD and AT chosen for mutagenesis are highlighted in purple. Residue numbering following UniProt entry Q03133.

Table 3.8: Results of bioinformatic analysis of DEBS3M5 KS-AT. Difference in the Gibbs free energy ($\Delta\Delta G$) between mutant and wild type (in Rosetta energy unit (REU) and kcal/mol) predicted by different bioinformatics tools. For better comparison of results from different bioinformatics tools, values of mCSM are handled as described in section 5.1.14. Mutations are predicted to be a Nullspot (not relevant for the interaction) or a Hotspot (relevant for the interaction) via SpotOn. Since Robetta and DrugScore are limited to alanine mutations, no prediction could be made for FS50 (A868R). n. a.: not available. Residue numbering following UniProt entry Q03133.

Program	Rosetta				Robetta	DrugScore	mCSM	SpotOn
Chain	A		D		A			A
PAL	w	w/o	w	w/o	w/o		w	w/o
Protein	$\Delta\Delta G$ (REU)				$\Delta\Delta G$ (kcal/mol)			Hot/ Null
FS45 (E129A)	4.27	0.856	4.70	1.26	0.48	-0.18	1.42	Null
FS46 (D134A)	0.550	0.592	3.96	3.96	-0.50	3.73	3.23	Null
FS47 (R510A)	2.56	2.42	2.46	1.08	0.40	0.54	1.57	Null
FS48 (V518A)	6.65	5.64	6.59	5.64	0.90	0.84	1.00	Hot
FS49 (Y867A)	4.51	3.90	2.18	1.76	1.90	3.29	1.31	Null
FS50 (A868R)	100.5	78.4	17.3	18.0	n. a.	n. a.	3.24	Null

Maren Berlinghof cloned the mutants and isolated the wild type AR265 and mutants FS45-FS50 in biological triplicates during her master's thesis under my supervision.^[223] Introducing any point mutations decreased expression yields drastically for all sites (Tab. 3.9). While differences between mutant and wild type yields were very pronounced, all mutant yields were very low and did not show significant differences between each other.

Results

Table 3.9: Average expression yields of DEBS3M5 KS-AT wild type (AR265) and mutants (FS45-FS50) after tandem affinity chromatography of 1 L expression culture. Yields determined from three independent expressions.

Protein	Pellet's Weight (g)	Total Yield (mg)	Yield (mg) per g Pellet
AR265	8.83	11.3	1.28
FS45	8.67	0.0269	0.00310
FS46	7.48	0.0224	0.00299
FS47	7.73	0.0192	0.00248
FS48	6.63	0.0637	0.00961
FS49	9.92	0.0985	0.00993
FS50	8.00	0.101	0.0126

Despite the low expression yield, mutants FS45-FS48 were pure after tandem affinity chromatography (Fig. 3.21). Elution fractions of mutants FS49 and FS50 showed a band shift and also impurities of higher molecular weight proteins. These have already been observed in the cell lysate and most likely were *E. coli* proteins. Depending on protein handling, some replicates showed these impurities, others did not. Almost no protein degradation was observed. Due to the very low yields and impurities, further analysis of the proteins was difficult and gave no reliable results.

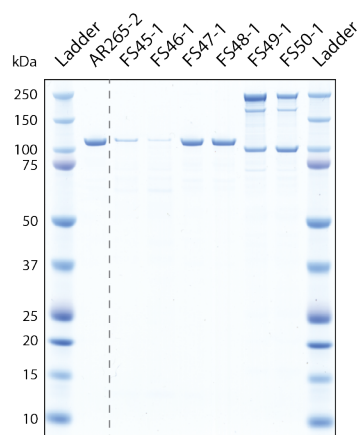


Figure 3.21: Analysis of DEBS3M5 KS-AT wild type (AR265) and mutants (FS45, FS48-FS50). Exemplary SDS-PAGE (NuPAGE Bis-Tris 4-12%) of one biological replicate. Concentrated elution fractions after tandem affinity chromatography.

Since the expression yield is a measure for stability,^[20,225] the drastically reduced mutant yields compared to the wild type indicate that single point mutations introduced at interfaces within the DEBS3M5 KS-AT didomain led to severe protein destabilization. In contrast to the mFAS, DEBS3M5 was not robust at all to mutagenesis and seemed to be unsuitable for AT

domain swapping approaches. This might be the reason why the design of mFAS/DEBS3M5 chimeras was not successful. In the following, our results were used to correlate the outcome of experiments to the prediction of the different bioinformatics tools.

3.4.5 Correlation of Experimental and *in silico* Data

We compared the experimental data on mFAS and DEBS3M5 KS-AT wild type with the prediction of the different bioinformatics tools to see which of these tools are helpful or whether they are useful at all in engineering approaches of complex interfaces and to see which of our analysis methods can serve as reliable readout.

This section will solely focus on the bioinformatics tool Rosetta. Within the set of tools used in this study, it is the one which uses the most complex algorithm to predict effects of mutations on the stability of proteins. I expected the predictions of this elaborate tool to be more reliable than the simpler web-based programs. Correlation of the experimental output yields, melting temperatures, and oligomeric states confirmed this expectation: Overall, only predictions of Rosetta showed correlations to experimental output, which will be shown in the following. Predictions of the other bioinformatics tools showed no or only partly correlation tendencies. This is exemplarily shown in Figure S17. Some bioinformatics tools (Robetta, DrugScore, SpotOn) could only be used without the PAL. Thus, we performed Rosetta calculations additionally without PAL to see whether this impairs the correlations. Furthermore, we wanted to analyze effects of different chains on the prediction. All bioinformatics tools were used for analysis of chain A. We performed analysis with Rosetta additionally for chain D, which shows the largest and second largest RMSD to chain A in DEBS3M5 KS-AT (PDB: 2HG4,^[99] 0.852 Å) and in mFAS KS-AT (PDB: 5MY0,^[9] 0.480 Å), respectively.

Correlation of Experimental and *in silico* Data for mFAS KS-AT Wild Type

Overall, the predictions for chains A and D gave different values for the change in Gibbs free energy (Tab. 3.6). However, the discrepancy between different mutations is similar and, thus, similar correlation tendencies should occur for chains A and D. Regarding predictions with and without PAL for chain A and chain D, most mutants remained unaffected. Solely mutants D61A and I457A showed a significant shift in predicted energy change: Without PAL, these mutants were predicted to effect stability to a lesser extent than with PAL, indicating that sites D61 and I457 interact with this linker. This should clearly influence correlation tendencies.

Expression yields served as first readout for effects on protein stability.^[20,225] To see, whether there is a correlation, the expression yield was plotted against the predicted change in Gibbs

Results

free energy (Fig. 3.22A). Positive changes correspond to destabilizing effects, while negative $\Delta\Delta G$ values indicate stabilizing effects. Mutations with higher change in the Gibbs free energy are expected to influence the stability to a higher degree. Most proteins showed wild type-like expression levels, which can be explained by small changes in the Gibbs free energy of ± 3.5 Rosetta energy units (REU). For higher energy changes, there was a tendency towards lower expression yields. Overall, predictions without PAL showed a similar correlation trend (Fig. 3.22B), although the predicted values differed slightly. However, the predictions of the two mutants with the lowest yields (D61A and I457A) were significantly shifted towards lower energy changes, impairing the correlation. This can be explained by interactions of the sites D61 and I457 with the PAL. Experimental and *in silico* analysis showed their importance in the didomain. Mutation of D61 led to severe destabilization, mutation of I457 showed partly destabilization. Chain D with and without PAL (Fig. 3.22C and D) showed similar predictions and, hence, similar correlation tendencies as chain A.

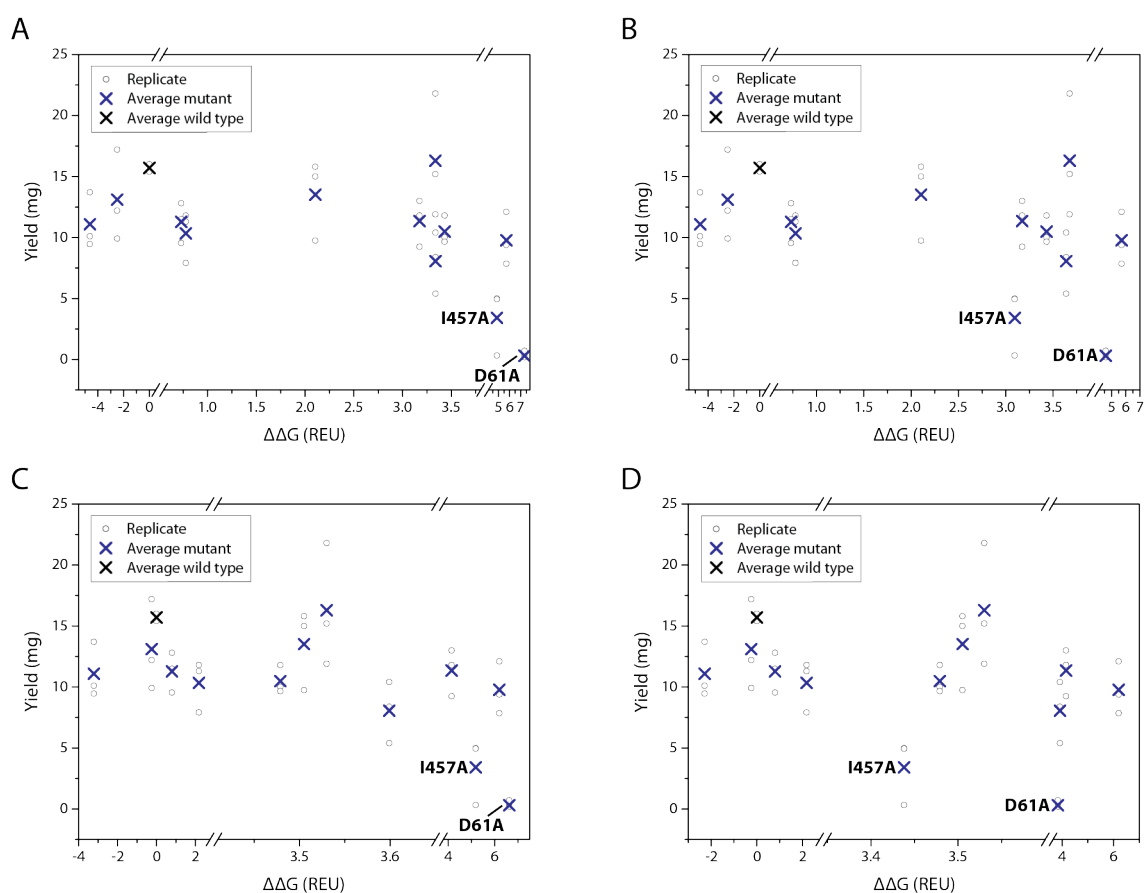


Figure 3.22: Yields (in mg per 1 L expression culture) of mFAS KS-AT wild type and mutants correlated to changes in the Gibbs free energy ($\Delta\Delta G$ in REU) predicted by Rosetta. (A) Chain A with PAL. (B) Chain A without PAL. (C) Chain D with PAL. (D) Chain D without PAL. Black circles indicate the yields of biological replicates. Average yields of biological triplicates of wild type and mutants are indicated by black and blue crosses, respectively. Residues D61 and I457 are labeled for clarity. Please mind the scales of the different x-axis sectors.

Melting temperatures served as second experimental readout. Differences in melting temperatures ($\Delta T_m = T_m(\text{Wild Type}) - T_m(\text{Mutant})$) were plotted against changes in the Gibbs free energy (Fig. 3.23A). Negative energy changes have stabilizing effects and are expected to increase melting temperatures. ΔT_m should be positive and corresponding mutants should be located in the second quadrant. Positive energy changes should be destabilizing and should lower melting temperatures leading to a negative ΔT_m . Mutants should be located in the fourth quadrant. Most mutants behaved as expected and there is a clear correlation between the difference in melting temperature and the predicted energy change. Overall, predictions without PAL (Fig. 3.23B) with shifted predictions for mutants D61A and I457A impaired this correlation due to their interaction with PAL. Correlations with predictions for chain D with and without PAL are shown in Figure 3.23C and D.

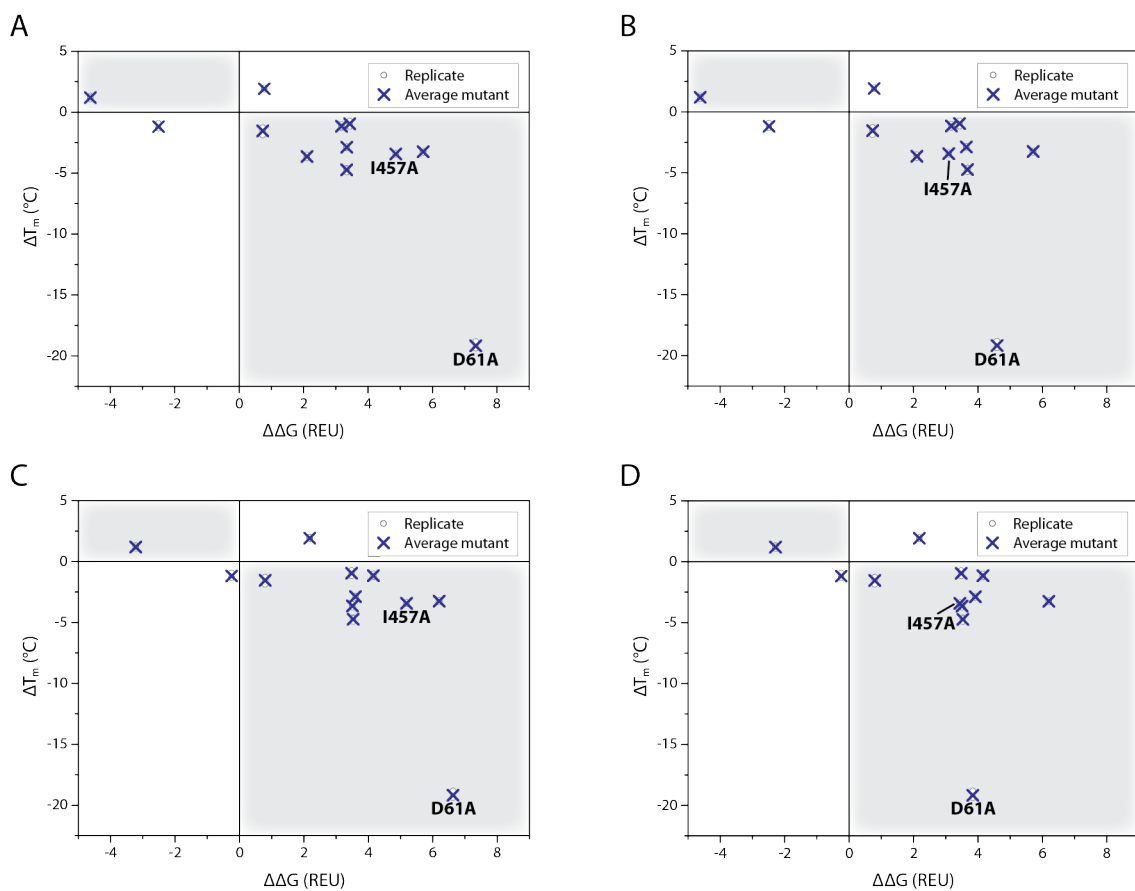


Figure 3.23: Changes in melting temperatures (ΔT_m in $^{\circ}\text{C}$) of mFAS KS-AT mutants correlated to changes in the Gibbs free energy ($\Delta\Delta G$ in REU) predicted by Rosetta. (A) Chain A with PAL. (B) Chain A without PAL. (C) Chain D with PAL. (D) Chain D without PAL. Gray areas indicate quadrants, which would show a correlation between experimental and predicted output. Black circles indicate the melting temperatures of biological replicates. Average of biological triplicates is indicated by blue crosses. Residues D61 and I457 are labeled for clarity.

Results

Oligomeric states served as third readout. The mFAS KS-AT wild type showed mainly dimeric and partly monomeric species (Fig. 3.18). Most of the mutants showed a similar behavior. In order to evaluate correlation, the ratio of dimer to monomer was plotted against changes in the Gibbs free energy (Fig. 3.24A). Only mutant D61A showed mainly higher oligomeric species. These results were not correlated, but the average ratio of higher oligomeric states to the sum of dimeric and monomeric species was 0.9. This indicates strong differences compared to the wild type. Overall, the ratio of dimeric to monomeric species of most mutants is very similar to the wild type and is even increased. Similar to the correlation of expression yields, this ratio dropped for high positive energy changes and a correlation tendency emerged. This trend was impaired for correlation without the PAL (Fig. 3.24B) due to D61's and I457's interactions with the linker. Correlations with predictions for chain D with and without PAL are shown in Figure 3.24C and D.

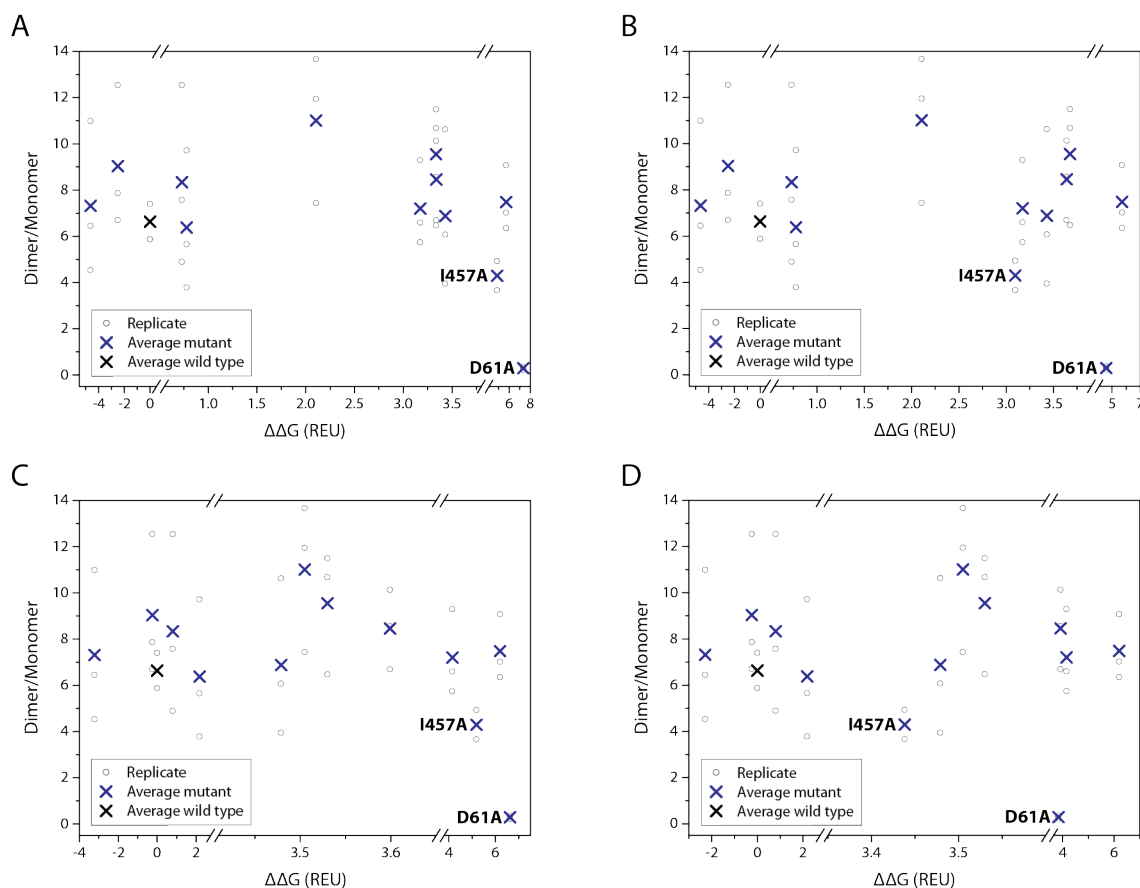


Figure 3.24: Ratio of monomeric to dimeric species of mFAS KS-AT wild type and mutants correlated to changes in the Gibbs free energy ($\Delta\Delta G$ in REU) predicted by Rosetta. (A) Chain A with PAL. (B) Chain A without PAL. (C) Chain D with PAL. (D) Chain D without PAL. Positive and negative $\Delta\Delta G$ correspond to destabilizing and stabilizing effects, respectively. Black circles indicate the yields of biological replicates. Average yields of biological triplicates of wild type and mutants are indicated by black and blue crosses, respectively. Residues D61 and I457 are labeled for clarity. Please mind the scales of the different x-axis sectors.

Overall, experimental results were found to correlate with the predicted effects irrespectively of which chain was used for the prediction. As expected, predictions without PAL influenced solely residues interacting with the linker. Since the number of those residues is limited, correlation tendencies were still observed. Both experimental and *in silico* data showed that residues D61 and I457 are important for protein stability. Each experimental output used for estimation of protein stability was overall suited. All three methods clearly identified mutant D61A as highly destabilizing.

Correlation of Experimental and *in silico* Data for DEBS3M5 KS-AT Wild Type

For the DEBS3M5 KS-AT, only six mutants were created, leading to less data points for correlation. Overall, predictions in energy change for chain A and D are quite similar (Tab. 3.8). Mutants D134A and Y867A gave slightly different values. Highly remarkable are the predicted strong effect of mutant A868R as well as the high discrepancy between the predicted values for the two chains. Overall, chains A and D are expected to show different correlation tendencies. Calculations without the PAL lowered only the predicted effect of mutant E129A, which indicates that the site E129 is interacting with this linker.

Results

All mutations introduced led to a severe drop in expression yields. No correlation between the predicted changes in the Gibbs free energy was observed (Fig. 3.25): Mutants predicted to have higher destabilizing effects showed rather higher yields than mutants with lower predicted effects. This might be caused by the drastically low yields, which make a precise determination of these effects impossible. Mutations introduced at the interfaces might destabilize the protein to an extent that the experimental readout “expression yield” is insufficiently sensitive towards the subtle differences between mutants and delivers no distinguishable results.

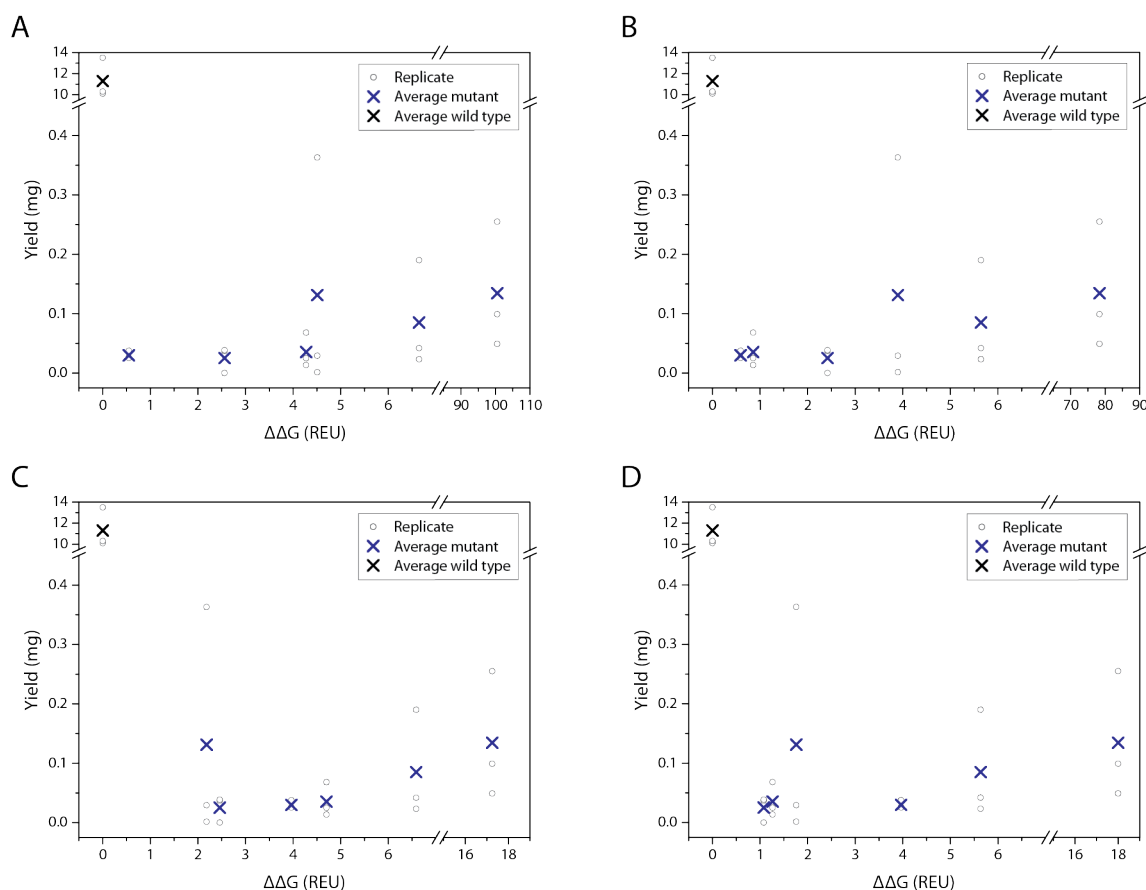


Figure 3.25: Yields (in mg per 1L expression culture) of DEBS3M5 KS-AT wild type and mutants correlated to changes in the Gibbs free energy ($\Delta\Delta G$ in REU) predicted by Rosetta. (A) Chain A with PAL. (B) Chain A without PAL. (C) Chain D with PAL. (D) Chain D without PAL. Positive and negative $\Delta\Delta G$ correspond to destabilizing and stabilizing effects, respectively. Black circles indicate the yields of biological replicates. Average yields of biological triplicates of wild type and mutants are indicated by black and blue crosses, respectively. Please mind the scales of the different x- and y-axis sectors.

3.4.6 Generation of mFAS/FabD KS-AT Chimeras

Since the AT domain from DEBS3M5 was not suited for domain swapping, we designed new chimeras with the AT (FabD) from *E. coli* FAS, which shows a specificity for elongation substrates (Mal, MMal) over the priming substrate Ac (section 3.3.6 and reference^[216]). FabD and mMAT are overall similar: They share a sequence identity of 23% and their structural alignment gave an RMSD of 2.823 Å. In conformity with the DEBS3M5 chimera design, we created two chimeras with slightly different domain boundaries (Fig. 3.26).

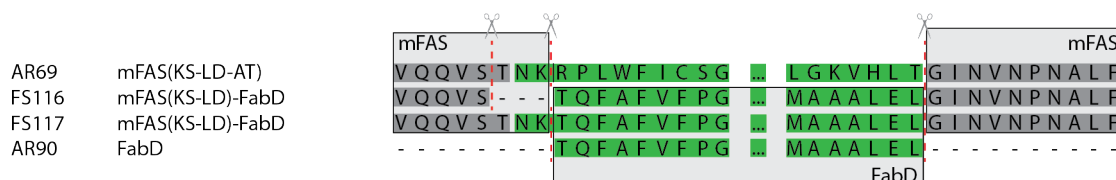


Figure 3.26: Domain boundaries for AT swapping with FabD. Two alternative boundaries were chosen for the first interface (mFAS(KAL):FabD), leading to the two chimeras FS116 and FS117.

Table 3.10: Test expression yields of KS-AT chimeras after tandem affinity chromatography of 1 L expression culture.

Protein	Pellet's Weight (g)	Total Yield (μg)	Yield (μg) per g Pellet
FS116	21.5	828	38.5
FS117	24.5	973	39.7

Test expression of both proteins resulted in comparable yields (Tab. 3.10). SDS-PAGE analysis of the concentrated elution fractions showed mainly full-length protein (FS116: 94 kDa, FS117: 95 kDa), but also degraded proteins (Fig. 3.27A). Proteins were further analyzed by TSA. Both chimeras showed high start fluorescences, which indicates unfolded protein.^[224] Melting temperatures were significantly decreased compared to the wild type mFAS KS-AT and *E. coli* FabD (Fig. 3.27B). HPLC analysis showed no defined dimeric species and rather higher oligomeric states (Fig. 3.27C). Since no mFAS KS-AT wild type was available, the DEBS3M5 KS-AT served as reference. All these results pointed out poor stability of the mFAS/FabD chimeras. Nevertheless, FabD's activity in the chimeric didomain was analyzed using the established AT activity assay (Fig. 3.27D). As expected, FabD preferred loading the elongation substrates Mal and MMal. Although turnover rates were significantly decreased to about 30% of the wild type activity, the absolute numbers of 25 to 30 molecules/s were still outstandingly high compared to PKS systems.

Results

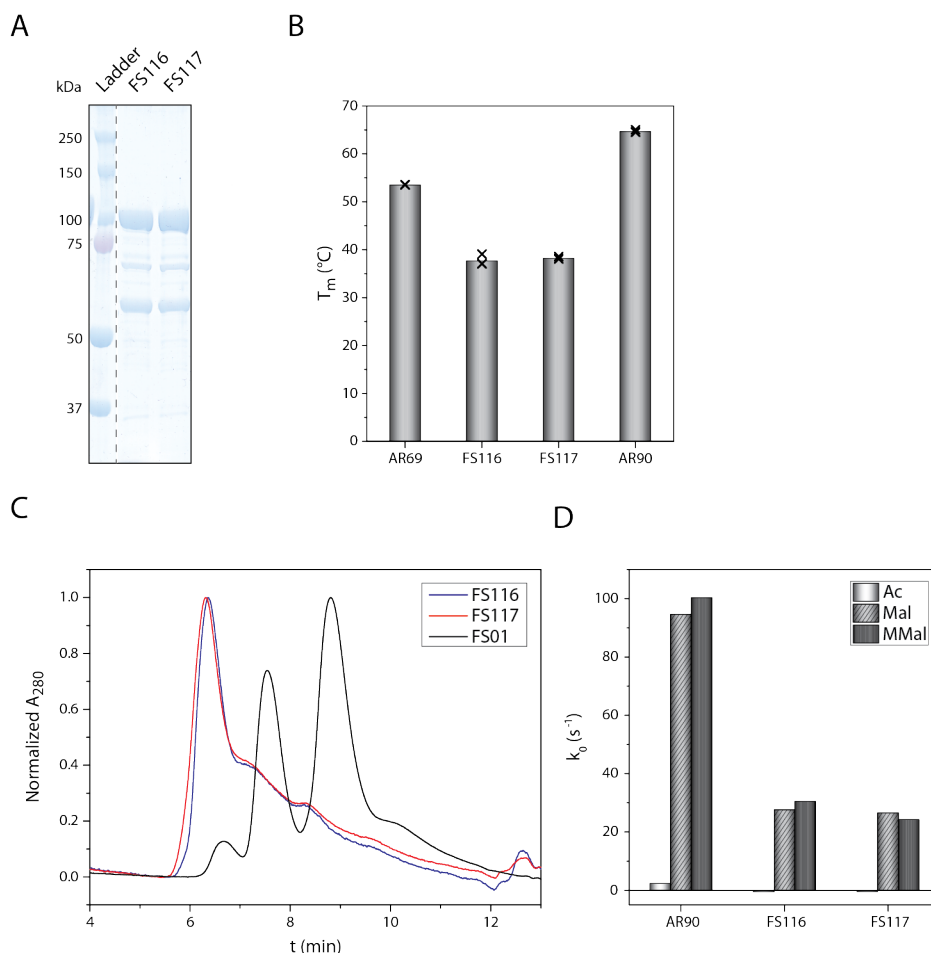


Figure 3.27: Analysis of KS-AT chimeras. (A) SDS-PAGE (10% polyacrylamide) of mFAS/FabD chimeric KS-AT FS116 and FS117. Concentrated elution fractions after tandem affinity chromatography. (B) Average melting temperatures (in °C) of mFAS KS-AT and *E. coli* FabD wild types (AR69 and AR90) and of chimeric KS-AT constructs (FS116, FS117) in storage buffer (Strep wash buffer pH 7.6). Measured in triplicates. Crosses correspond to one technical replicate (FS116, FS117) or one biological replicate (AR69, AR90). (C) Normalized size exclusion chromatogram of FS116 (blue) and FS117 (red) with DEBS3M5 KS-AT wild type FS01 (black) as reference. (D) Substrate screening of AT activity assay. Loading of substrates Ac (white), Mal (light gray, hatched), and MMal (dark gray) onto *E. coli* ACP mediated by wild type FabD (AR90) and by mFAS/FabD KS-AT chimeras.

Overall, analysis of proteins revealed their high potential in domain swapping approaches. Even though the chimeras suffered from artificial interfaces, leading to less stable and degraded protein, the transferase activity was still high. For our application as elongation module, interfaces could be optimized. Furthermore, I expect the integration of the typeII ACP in the typeI system to be challenging. The design of a suitable linker will be key for productive domain-ACP interactions.

3.4.7 Generation of mFAS/DEBS3M6 KS-AT Chimeras

Additionally to the new hybrids with FabD, we designed new chimeras with the AT domain from DEBS3M6, which is expected to be highly substrate specific for MMal-CoA. Overall, the first domain boundaries used seemed to be most promising, according to structural and sequential alignments, despite the low success with domain swapping of DEBS3 AT5. The same domain boundaries were used in a study of my colleague Mirko Joppe for domain swapping of mFAS AT in DEBS3M6. He showed that the exchange of solely the AT worked better than the exchange of LD-AT. Based on this data, we designed mFAS/DEBS3M6 KS-AT chimeras.

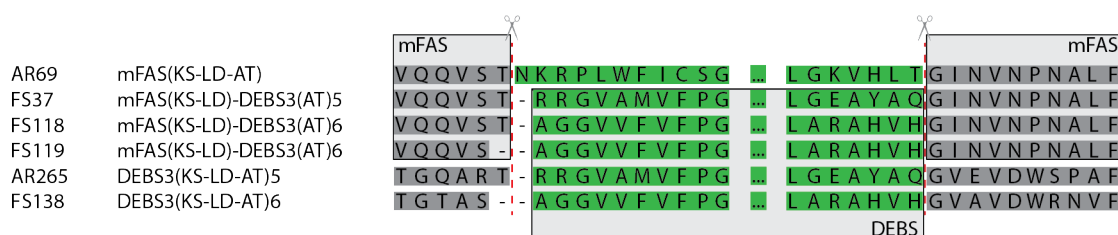


Figure 3.28: Domain boundaries for AT swapping with DEBS AT6. Two alternative boundaries were chosen for the first interface (mFAS(KAL):DEBS3M6(AT)), leading to two chimeras FS118 and FS119.

We created a sequence alignment in order to find the corresponding domain boundaries in DEBS3M6 (Fig. 3.28). For one construct (FS118), we used the exact same domain boundaries as for swapping with DEBS3 AT5 (FS37). Since the wild type DEBS3M6 KS-AT does not contain the amino acid T at the C-terminus of the LD, we created a second chimera without this residue (FS119).

Table 3.11: Test expression yields of KS-AT chimeras after tandem affinity chromatography of 1 L expression culture.

Protein	Pellet's Weight (g)	Total Yield (μg)	Yield (μg) per g Pellet
FS118	17.4	1270	73.0
FS119	22.3	356	16.0

Results

Test expression of both proteins showed significantly higher yields for FS118, which contains the original domain boundaries for AT swapping (Tab. 3.11). SDS-PAGE analysis of the concentrated elution fractions of FS118 and FS119 showed mainly full-length protein (95 kDa), but also degraded proteins (Fig. 3.29A). Proteins were further analyzed by TSA (Fig. 3.29B). FS118 showed low start fluorescence and gave a melting temperature of 41 °C in storage buffer, which is significantly decreased compared to wild type mFAS KS-AT but comparable to wild type DEBS3M5 KS-AT (see section 3.4.2). FS119, on the other hand, showed high start fluorescence, which indicates unfolded protein,^[224] but gave a similar melting temperature of 40 °C. HPLC analysis of both constructs (Fig. 3.29C) fortified this finding. FS118 showed defined monomeric, dimeric, and higher oligomeric states (like the DEBS3M5 KS-AT wild type), whereas FS119 showed no defined dimeric species like the mFAS/FabD chimeric constructs FS116 and FS117 (see section 3.4.6). Overall, these results indicated that FS118 is a stable chimera and has great potential for its application in an mFAS assembly line, whereas FS119 is in general poorly stable and not suited. Even though proteins were not isolated in replicates, these preliminary results point towards the sensitivity of chimeras towards minimal changes (difference in one amino acid) in the artificial domain-domain interaction.

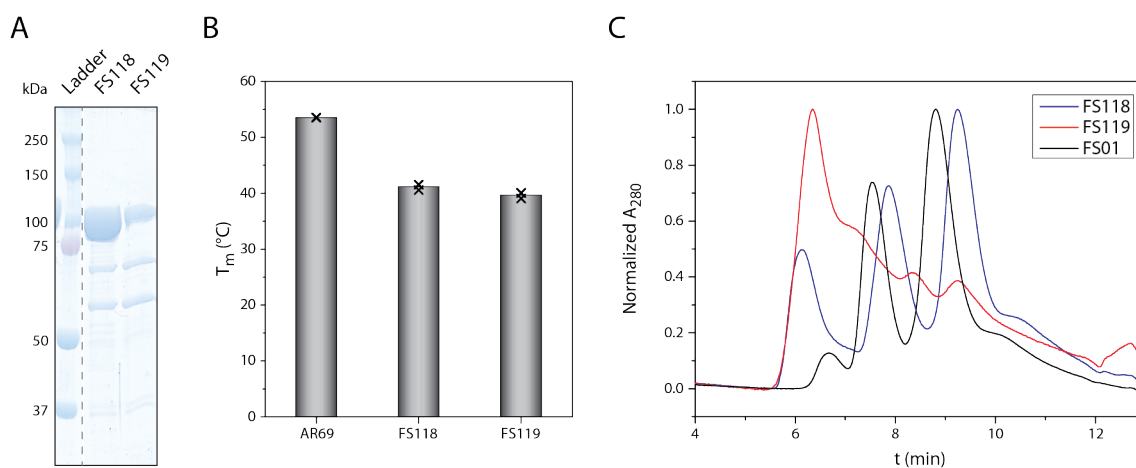


Figure 3.29: Analysis of chimeric KS-AT. (A) SDS-PAGE (10% polyacrylamide) of mFAS/DEBS3M6 chimeric KS-AT FS118 and FS119. Concentrated elution fractions after tandem affinity chromatography. (B) Average melting temperatures (in °C) of mFAS KS-AT wild type (AR69) and of chimeric KS-AT constructs (FS118, FS119) in storage buffer (Strep wash buffer pH 7.6). Measured in triplicates. Crosses correspond to one technical replicate. (C) Normalized size exclusion chromatogram of FS118 (blue) and FS119 (red) with DEBS3M5 KS-AT wild type FS01 (black) as reference.

3.4.8 Analysis of mFAS/DEBS3M6 KS-AT Chimeras

Based on the promising results on expression and stability of the chimeric mFAS/DEBS3M6 KS-AT FS118, we started the master's thesis project of Alicia Just. In this project, we wanted to compare the AT activity of the chimera with the AT activity of the wild type system (FS138, DEBS3M6 KS-AT). Additionally to FS118, consisting of mFAS KS and DEBS3M6 AT domain, we designed the corresponding "inverse" chimera, consisting of DEBS3M6 KS and mFAS AT domain (FS139). The AT activity of the wild type mFAS KS-AT was characterized in kinetic detail before.^[9] Analysis of the three KS-AT constructs in their AT activity should give insights into the compatibility of mFAS/DEBS3M6 in domain swapping approaches.

Table 3.12: Average expression yields of KS-AT chimeras after tandem affinity chromatography of 1 L expression culture. Average of eight, five, and seven expressions for FS118, FS138, and FS139, respectively.

Protein	Pellet's Weight (g)	Total Yield (mg)	Yield (mg) per g Pellet
FS118	9.14	0.913	0.0999
FS138	8.77	4.96	0.566
FS139	9.61	0.568	0.0568

Alicia Just expressed all proteins as biological replicates under my supervision.^[226] Yields of both chimeras were drastically reduced compared to wild type yields, but still sufficient for AT characterization (Tab. 3.12). SDS-PAGE analysis showed mainly full-length protein of 95-98 kDa (Fig. 3.30A), but revealed some degraded protein for the chimeras of around 70 and 60 kDa. TSA showed slightly decreased melting temperatures for both chimeras in storage buffer (FS118: 41 °C, FS138: 43 °C, FS139: 40 °C, Fig. 3.30B and S18). The wild type FS138 showed mainly dimeric species, whereas both chimeras formed monomers, dimers, and different higher oligomers in different ratios for biological replicates (Fig. 3.30C).

Table 3.13: Average expression yields of ACPs after affinity chromatography of 1 L expression culture. Yields determined from twelve independent expressions.

Protein	Pellet's Weight (g)	Total Yield (mg)	Yield (mg) per g Pellet
mFAS ACP	8.61	24.7	2.87
DEBS3M6 ACP	6.60	20.0	3.03

Results

For AT characterization, mFAS and DEBS3M6 ACP were isolated in high yields (Tab. 3.13). ACP was further purified via SEC, resulting in highly pure protein (Fig. 3.30A). Different expression batches were pooled and full phosphopantetheinylation was confirmed by MS analysis (Tab. S9 and Fig. S19).

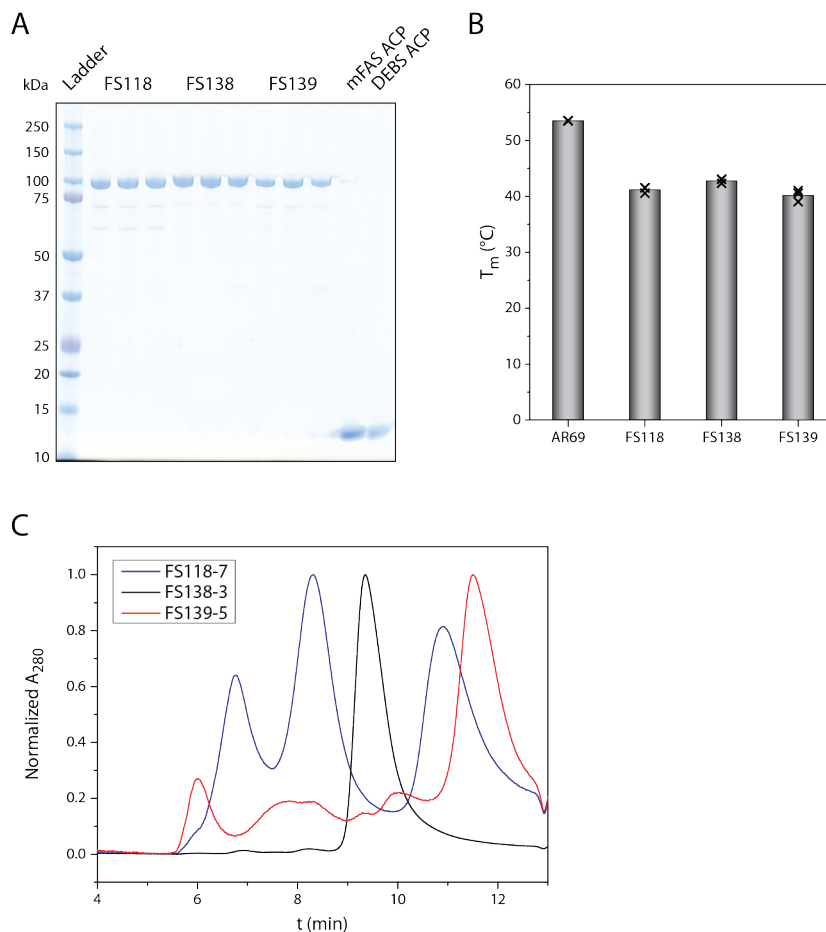


Figure 3.30: Analysis of chimeric KS-AT. (A) SDS-PAGE (NuPAGE Bis-Tris 4-12 %) of mFAS/DEBS3M6 chimeric KS-AT FS118 and FS139 and of wild type DEBS3M6 KS-AT FS138. Concentrated elution fractions after tandem affinity chromatography. (B) Average melting temperatures (in °C) of mFAS and DEBS3M6 KS-AT wild types (AR69 and FS138) and of chimeric KS-AT constructs (FS118, FS139) in storage buffer (Strep wash buffer pH 7.6). Measured in triplicates. Crosses correspond to one technical (AR69 and FS118) and one biological replicate (FS138 and FS139). (C) Exemplary normalized size exclusion chromatograms of chimeras FS118 (blue) and FS139 (red) and DEBS3M6 KS-AT wild type FS138 (black).

In a first substrate screening, the loading of Ac, Mal, and MMal onto mFAS and DEBS3M6 ACP mediated by FS118, FS138, and FS139 was tested at fixed substrate concentrations (X-CoA and ACP) of 100 μ M (Fig. 3.31A-C). As expected, the wild type FS138 loaded mainly MMal onto the native DEBS3M6 ACP, but was also able to load MMal onto the non-native mFAS ACP with slightly reduced turnover rates (Fig. 3.31B). The chimera FS118 (containing DEBS3M6 AT) also exhibited a loading preference of MMal and showed similar transacylation

rates for MMal onto the DEBS3M6 ACP as the wild type and even higher rates for loading onto the mFAS ACP (Fig. 3.31A). The chimera FS139 (containing mFAS AT) overall showed the highest turnover rates, as expected, and loaded all three substrates onto its native ACP with higher rates for Ac and Mal than for MMal, like the wild type mFAS KS-AT (Fig. 3.31C).^[9] This chimera was also able to load the three substrates onto the non-native DEBS3M6 ACP with only slightly reduced rates.

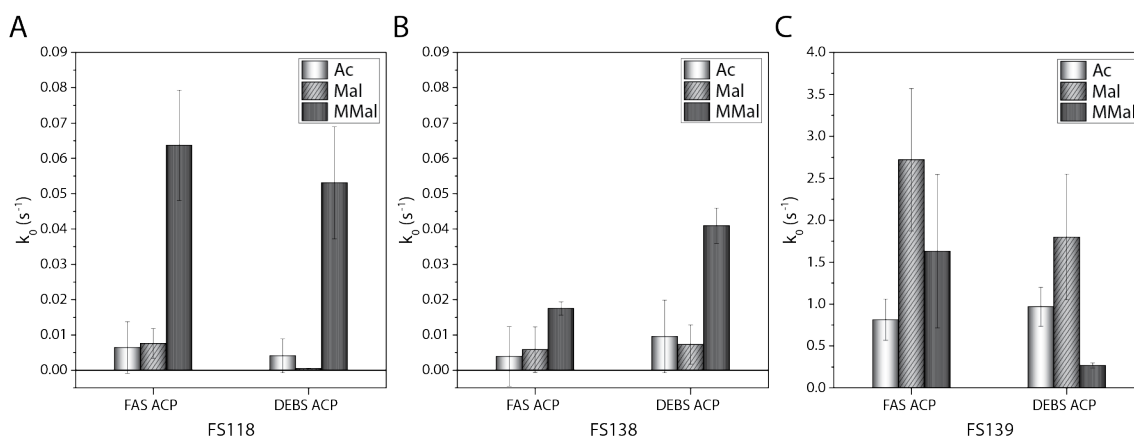


Figure 3.31: Transacylation substrate screening of KS-AT chimera FS118 (A), DEBS3M6 KS-AT wild type FS138 (B), and KS-AT chimera FS139 (C). Substrates Ac (white), Mal (light gray, hatched), and MMal (dark gray) as well as mFAS and DEBS3M6 ACP were used at fixed concentrations of 100 μ M. ATs measured in biological and technical triplicates. Bars show the average, error bars the standard deviation of each replicate.

Based on these first results, DEBS3M6 AT activity was analyzed in the chimeric construct FS118 and the wild type FS138 with its native substrates MMal-CoA and DEBS3M6 ACP (Fig. 3.32A and B). Biological triplicates were measured in technical triplicates and gave similar titration curves and kinetic parameters (Tab. S10 and Fig. S20 and S21). Kinetic parameters determined via global Michaelis-Menten fit for the average of each replicate are given in Table 3.14. Due to the relatively high value of $K_m^{MMal-CoA}$ for FS138, the ACP titration range was not ideal and the parameters are defective. Nevertheless, the data reveals that DEBS3M6 AT is relatively slow compared to other modular PKSs and features low catalytic efficiencies in ping and pong step. Overall, all kinetic parameters are decreased in the chimeric protein (turnover rate 70% of wild type and Michaelis-Menten constants $K_m^{MMal-CoA}$ 40% and K_m^{ACP} 37% of wild type parameters), leading to relatively similar catalytic efficiencies. Transition state energies are lowered by 1.3 kJ/mol and 1.4 kJ/mol during ping and pong step, respectively, in the chimeric protein compared to the wild type. This might simply be a consequence of the large calculation errors of the wild type kinetic parameters caused by a high value of K_m^{ACP} . Overall, the data shows that we successfully designed a functional chimera, which shows only minor loss in AT activity compared to the wild type.

Results

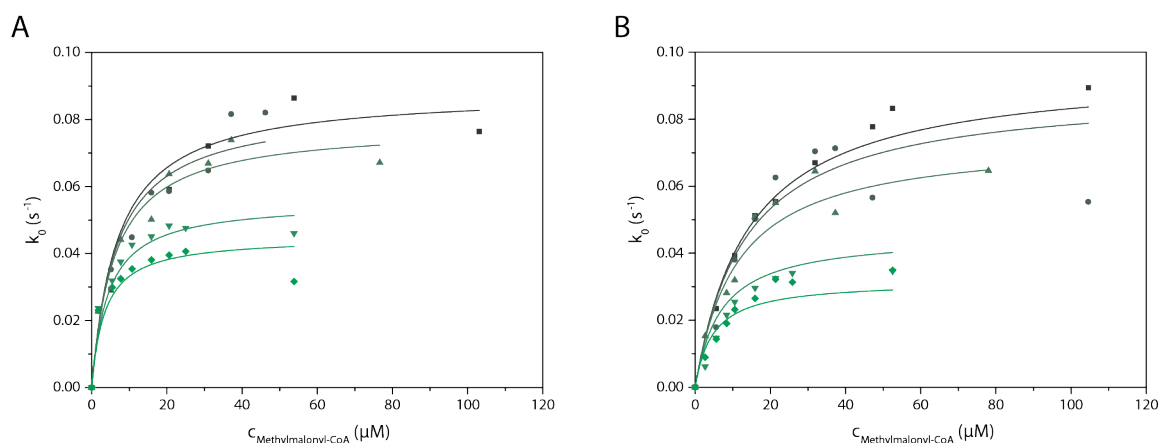


Figure 3.32: Titration curves of transacylation mediated by DEBS3M6 AT. Global fits of average measurements determined for biological triplicates measured in technical triplicates. (A) Chimera FS118. (B) Wild type FS138.

Table 3.14: Average parameters of transacylation of MMal-CoA onto DEBS3M6 ACP mediated by DEBS3M6 AT chimera (FS118) and wild type (FS138). Determined for biological triplicates measured in technical triplicates.

Protein	k_{cat} (s^{-1})	$K_m^{\text{MMal-CoA}}$ (μM)	K_m^{ACP} (μM)	$k_{\text{cat}}/K_m^{\text{MMal-CoA}}$ ($\text{M}^{-1}\text{s}^{-1}$)	$k_{\text{cat}}/K_m^{\text{ACP}}$ ($\text{M}^{-1}\text{s}^{-1}$)
FS118	$1.13 \times 10^{-1} \pm 6.93 \times 10^{-3}$	8.86 ± 1.28	103 ± 16.6	1.3×10^4	1.1×10^3
FS138	$1.64 \times 10^{-1} \pm 2.28 \times 10^{-2}$	24.6 ± 4.91	257 ± 66.2	6.7×10^3	6.4×10^2

We were not only interested in the functional, but also in the structural differences of wild type and chimera. In the course of her master's thesis, Alicia Just tried to solve the structures of FS118, FS138, and FS139 via X-ray crystallography under supervision of Dr. Apirat Chaikwad, Dr. Karthik Paithankar, and me. Unfortunately, in our numerous crystallization experiments, we were not able to obtain any protein crystals.^[226] As further KS-AT crystallization experiments of master student Matthias Zeug showed, phosphate seems to form salt crystals and should be avoided in further experiments.

4 Discussion

Parts of this chapter have been previously accepted for publication in: F. Stegemann, M. Grininger, "Transacylation Kinetics in Fatty Acid and Polyketide Synthases and its Sensitivity to Point Mutations", *ChemCatChem* (2021).^[1] For individual contributions, copyright, and Creative Commons license, please see the statement of personal contributions.

4.1 AT Activity Assay

The core of the present thesis is the established AT activity assay, which is a powerful tool to analyze ATs in enzymatic detail. Our work demonstrates that this assay is extremely useful, especially for comparative studies on related proteins and for mutation studies. Using this assay, we were able to determine absolute kinetic parameters for both AT-mediated reaction steps (ping and pong) and to gain detailed mechanistic insights into functional differences of proteins from different classes or of different protein mutants.

This elaborate assay was established based on earlier work^[133,211] and was optimized during the course of my doctoral studies. Previous studies on AT activity performed single titrations for both substrates (X-CoA and ACP), titrating only one substrate while keeping the other substrate at saturated conditions.^[14,133,211,227] Separate Michaelis-Menten fits of both titrations gave access to all absolute kinetic parameters. This approach has the following disadvantage: ATs with a high value of K_m cannot be analyzed by this assay, simply because substrate saturation cannot be reached. In our assay, we wanted to circumvent this problem and came up with the following solution: We simultaneously titrated both substrates by titrating the acyl substrate to a series of fixed ACP concentrations. During my master's thesis and in previous studies,^[9,228] these titration curves were fitted separately, leading to apparent kinetic parameters for each ACP concentration. These apparent kinetic parameters themselves were then fitted in double-reciprocal plots, giving access to global kinetic parameters. This approach has the following disadvantage: The absolute kinetic parameters are rather error-prone, since the data is not weighted equally. In the double-reciprocal fits, there is an antiproportional relationship between weight and ACP concentration, meaning that kinetic parameters determined for low ACP concentrations have higher weight than those determined for high ACP concentrations. This is particularly problematic, since the first are much more error-prone than the latter. In the present study, this problem was solved by globally fitting all data from all ACP concentrations via a global Michaelis-Menten fit (Eq. (3.14)).^[212] This way, all data is weighted equally. Furthermore, this method is very robust to measurement errors and outliers and directly grants access to absolute kinetic parameters.

Overall, the AT activity assay is very valuable. However, some aspects need to be considered using this assay. It requires thorough handling of samples and is sensitive to even minimal handling differences. Usage of the same substrate batch (ACP and acyl substrate) for each system and for all mutants in our measurements reduced the variation caused by differences in substrate quality. Correct measurement requires careful inspection of raw data, ensuring the analyzed time span falls within the range of initial reaction velocity v_0 . This is usually guaranteed for substrate consumption below 10%.^[129] Thus, for each measurement, substrate consumption was always examined. Additionally, dust and bubbles disturb proper measurement and working in a clean environment is essential. Wells were thoroughly inspected after careful and slow pipetting. Raw data was immediately inspected for outliers and artifacts and measurements were repeated, if necessary. Overall, the AT activity assay provides a large set of kinetic parameters, describing the AT-mediated reactions, with reasonable effort in the assay itself. Preceding preparation of stock solutions and high-quality protein, however, is quite time-consuming (biological triplicates of ATs, high amounts of high-quality ACPs) and acyl substrates are expensive. Adjustment of the originally 96-well format to the 384-well format downsizes the reaction volume to a fifth and by this reduces resources significantly (reduced usage of acyl substrates, enzymes, AT, and ACP, less protein preparations, more wells measured simultaneously).

Extensive usage of the AT activity assay in the present work also revealed general limitations. Kinetic parameters can only be determined precisely for systems with Michaelis-Menten constants lying within a certain range. For reliable values, the substrate (X-CoA and ACP) concentration range ideally should be varied within $0.2-5 \times K_m$. For extreme values, this range cannot be covered.

1. For exceptionally low Michaelis-Menten constants, low substrate (X-CoA or ACP) concentrations have to be used. This is problematic, since the substrate concentration needs to be significantly higher than the enzyme concentration (at least ten times).^[129] Decrease of the enzyme concentration is not infinitely possible, since poor signal-to-noise ratios impede proper data analysis.
2. On the other hand, exceptionally high Michaelis-Menten constants require high substrate (X-CoA or ACP) concentrations. This can lead to experimental problems, e.g. caused by substrate inhibition effects. Often the vast amount of ACP can simply not be provided and too high concentrated stock solutions can cause instability of ACP.

Nevertheless, our study demonstrates that even for systems with extreme Michaelis-Menten constants, kinetic parameters can still be estimated using this assay.

4.2 Kinetic Analysis of Acyl Transferases

Using this assay, we characterized different ATs in their transacylation and hydrolysis activity. The initial substrate screening is a suitable and fast method to gain first insights into substrate tolerance and relative transfer velocities of different systems. Even though the absolute maximum velocities differ, the screening shows the same velocity trends: ATs from modular PKSs feature slower transacylation rates than those from iterative PKSs, which are slower than those from FASs (Fig. 3.4A and B). Hydrolysis mediated by FASs is significantly slower than hydrolysis mediated by PKSs.

Unfortunately, we were not able to express the standalone DEBS1 AT0 for analysis. The priming AT with C-terminal MBP fusion did not show any significant transacylation activity in numerous test measurements (Fig. 3.4C). Nevertheless, the hydrolysis for the native and the non-native priming substrates Prop and Ac were measurable. This indicates that the C-terminal MBP impedes productive AT-ACP interactions, but does not or at least not completely inhibit the acyl substrate loading onto the AT domain. AVES1 AT0, on the other hand, showed a significantly higher transacylation than hydrolysis rate already in the substrate screening (Fig. 3.4D). This system was chosen as a reference priming AT to DEBS and was not analyzed further due to the inactivity of the latter. In the following, only the elongation ATs will be compared regarding their activity.

Although limitations occurred in hydrolysis measurements for many systems due to extreme Michaelis-Menten constants, the maximal turnover rate for all PKS systems was determined at least for the native substrate (Tab. 3.1 and Fig. S7). All PKS systems feature similar hydrolysis rates, which is also in accordance with previous studies on other DEBS and PikA modules.^[132,133] Hydrolysis of FabD was not measurable.

Transacylation showed the highest turnover rates and efficiencies in both steps, ping and pong, for the FAS systems (Tab. 3.2). The iterative PKS is faster and more efficient in both steps than the modular PKSs. Although only few kinetic studies on transacylation exist, the determined kinetic parameters are in good accordance with Michaelis-Menten constants and overall turnover rates determined for PKS systems.^[133,143,206] For FabD, there are numerous studies available, giving various and partly divergent kinetic parameters.^[14,211,216,227] These studies determined even higher transacylation rates of up to 2000 molecules per second. Buffers with different phosphate concentrations and additives, pH values, and temperatures may influence the outcome of the parameters.^[206] Presence and absence of purification tags may also contribute to differences. Nevertheless, FabD was the fastest system in our setup with a turnover rate of almost 300 molecules per second (Tab. 3.2). The Michaelis-Menten constants we de-

Discussion

terminated are similar to the ones published by Misson *et al.*^[14] In the context of our study on FAS's and PKS's ATs, the kinetic parameters of the systems are comparable, since the measurements were performed under the same conditions. The only difference is that *E. coli* ACP contains solely a C-terminal His-tag for purification, whereas all other ACPs contain an additional N-terminal StrepII-tag. This additional tag could potentially affect the AT-ACP interaction, but would explain lower turnover rates for the other systems.

4.3 Implications of Kinetic Properties on the Acyl Transferase Reaction Mechanism by Comparing Hydrolysis and Transacylation

Our enzyme kinetic analyses reveal that hydrolysis mediated by ATs of PKS systems is just one to two orders of magnitude slower than transacylation (Tab. 3.1 and 3.2), which is in accordance with previous studies.^[133,143] This rate difference is much less than for FAS systems for which four orders of magnitude difference were recorded.^[9] We observe AT-mediated hydrolytic activity for all substrates – if native or non-native – transferred onto the ACP during transacylation and in most cases hydrolysis is fastest for the native substrates (Fig. 3.4). The purpose of the relatively high hydrolysis rates for PKSs remains elusive, but our data rules out any quality control by hydrolysis, as e. g. reported for some aminoacyl-tRNA synthetases in which hydrolytic active sites evolved in addition to their catalytic active sites.^[13]

Comparing AT-mediated transacylation and AT-mediated hydrolysis can give insight into the substrate specificity of enzymes.

1. The PikAIII AT5 is the only enzyme from our set of proteins for which we recorded a higher hydrolysis rate for the non-native substrate, although at low rate, speaking also in the specific case of PikAIII AT5 against a proofreading function. Different specificities for hydrolysis and transacylation are interesting from a mechanistic point of view, because they can just arise after the substrates have been loaded to the enzyme (after the ping step). Data implies that PikAIII AT5 can forward the native moiety MMal to the ACP-loading pong step, while redirecting the non-native moiety Mal into hydrolysis. In this regard, PikAIII AT5 seems capable of differentiating between the two substrates in the pong step.
2. In case of the other modular PKSs (DEBS3M5 and RAPS3M14), the absence of hydrolytic activity for non-native substrates, with respect to the detection limit of the assay, indicates a more stringent selection for substrates in the ping step.
3. We note that, based on the catalytic efficiency of the ping step, we can calculate the Michaelis-Menten constants for the AT-mediated hydrolysis (Tab. S6). This is possible, because the efficiencies for the ping step in transacylation and for hydrolysis are identical (Eq. (3.18) and (3.24)). We observe decreasing K_m^{X-CoA} values for hydrolysis from modular PKSs to the iterative PKS to the FAS system. This trend, as well as the trend in catalytic efficiencies (increased for iterative proteins), is overall similar to the ping step of transacylation.

4.4 Implications of Kinetic Properties on FAS/PKS Function and Evolution

A. Bar-Even and E. Noor *et al.*^[229] have mined enzyme databases for a systematic study on kinetic parameters of enzymes and substrates. As one of their key findings, they revealed that enzyme substrate pairs associated with primary metabolism tend to have higher turnover numbers than pairs of secondary metabolism and also average k_{cat}/K_m values are higher in primary than in secondary metabolism. We observe a very similar trend in our study on the transacylation function in FASs and PKSs; i. e., high turnover rates by FabD^[14,216,227] and MAT of FASs^[9,230] and low turnover rates by ATs of modular PKSs (Tab. 3.2). The lower turnover rates and catalytic efficiencies can be explained by serving the purpose of PKSs in providing sufficient efficiency to produce molecules at a low basal level or when necessary without draining vital metabolic pools (particularly in case of using Mal-CoA as elongation substrate). In contrast, the MAT domain of FAS needs to be fast and was under selection pressure for providing high rates for the high metabolic fluxes that are required by the central metabolic fatty acid biosynthesis. In conclusion, we observe in our dataset that the AT domains have evolved differently, serving the different purposes in enzyme function in the metabolic context.^[231]

In previous studies, the lab of S. Smith and we described MAT from rat and mouse FAS as fast, substrate polyspecific, and plastic,^[9,212,230] which are typical properties of primordial enzymes. We speculate that in MAT primordial enzyme properties kept preserved due to its role in fatty acid synthesis and due to the KS-mediated condensation step ensuring the fidelity of the fatty acid biosynthesis.^[212] Compared to MAT of FAS, the ATs of modular PKSs appear more specialized. Although ATs of modular PKSs with substrate tolerance towards mainly non-native elongation substrates have been reported,^[132,141,143,232] they have the ability to selectively transfer the native elongation substrates Mal and MMal, as we show in this study. Here, we note that we have not tested the transacylation of the native substrate ethylmalonyl-CoA.

Given that AT domains are integral part of monomodular iterative PKSs and modules of modular PKSs, their course of evolution is likely closely connected to the evolution of PKSs in general. Current models on the evolution of PKSs consider modular PKSs as successors of simpler systems that have likely been (also) iterative PKSs (and FASs or their precursor).^[6,42,44,78,233] Accordingly, it is tempting to speculate that a primordial AT enzyme, to which MAT of FAS is still reminiscent in functional properties, has developed into more selective and slower modular AT domains (as they appear from the dataset of this study) as part of PKS evolution.

4.5 Impact of Interface Mutations on Transacylation Kinetics

The critical domain-domain interaction between AT and ACP was further analyzed in a mutation study. Like for the wild type ATs, the initial substrate screening gave a first reliable insight into the effects of the mutations on reaction kinetics and revealed that the residue A539 is most likely not involved in the AT-ACP interaction (Fig. 3.7). Detailed kinetic analysis of mutant A539E confirmed this finding featuring wild type-like kinetic parameters (Tab. 3.3). In contrast to A539E, mutations in R850 led to intricate effects on both the AT-loading ping and the ACP-loading pong step.

1. The mutation R850K, preserving the positive charge, has almost no influence on the AT-MMal-CoA interaction, as indicated by the ping step remaining unaltered in catalytic efficiency and transition state energy compared to wild type DEBS3M5 AT (Fig. 3.9). However, the pong step drops in efficiency and overall reduces turnover rates for transacylation (57 % of wild type). The K_m^{ACP} of the pong step is similar to wild type AT, which indicates unchanged AT-X:ACP complex stability. Since the catalytic rate of the pong step (k_4) is decreased, K_m^{ACP} can just remain at wild type level when compensated by an increase of ratio k_{-3}/k_3 (see appendix “Mathematical Examination of Changes in Transacylation Kinetic Parameters Caused by Point Mutations”), indicating that the main effect of the mutation R850K is decelerating the formation (k_3) and/or accelerating the (non-productive) dissociation of AT-X:ACP (k_{-3}) (Fig. 4.1).
2. The mutant R850E inverses the charge from positive to negative and we expected severe changes in the kinetics of the transacylation reaction. Indeed, we observed lowest turnover rates within our set of mutants (Tab. 3.3). We note that the kinetic parameters of the pong step could only be collected with lower confidence due to the low K_m^{ACP} . For a reliable estimate on K_m^{ACP} , we therefore extracted data for saturated substrate MMal-CoA concentration titrated with ACP and received a K_m^{ACP} of 3.46 μM and an upper boundary of 4.6 μM (Fig. S15). Given this data, we conclude that the pong step remains essentially unaltered in catalytic efficiency and transition state energy. In contrast, the ping step is severely affected by this mutation and responsible for the overall decrease of turnover rates. Since the catalytic rate of the ping step (k_2) is contributing to K_m^{ACP} , the drop of K_m^{ACP} is “just” a direct consequence of the slow ping step. Under this assumption, the kinetics of the pong step would be entirely unaffected in spite of exchanging the positively charged arginine with the negatively charged glutamate (Fig. 4.1). In conclusion, this mutant, designed for a severe impact on the domain-domain interplay during the pong step, mainly (or entirely) takes effect on the initial ping step, which demonstrates the intricate effect surface mutations can have.

Discussion

3. In the mutant R850S, both ping step and pong step are affected by the amino acid exchange, as indicated by changed catalytic efficiencies and transition state energies compared to wild type DEBS3M5 AT, and effects cumulate to a decrease in turnover rates to 23 % compared to wild type (Tab. 3.3 and Fig. 3.9). The decrease in the catalytic rates of the ping step (k_2) and the pong step (k_4) accounts for the increased stabilities of the complexes AT:X-CoA and AT-X:ACP (lowered K_m^{X-CoA} and K_m^{ACP} values), but ratios k_{-1}/k_1 and k_{-3}/k_3 may contribute here (Fig. 4.1). Overall, the effect of mutation R850S is not constrained to the pong step, but also invasive to AT-X formation; i. e., not showing the specific properties expected of an AT:ACP interface mutation.

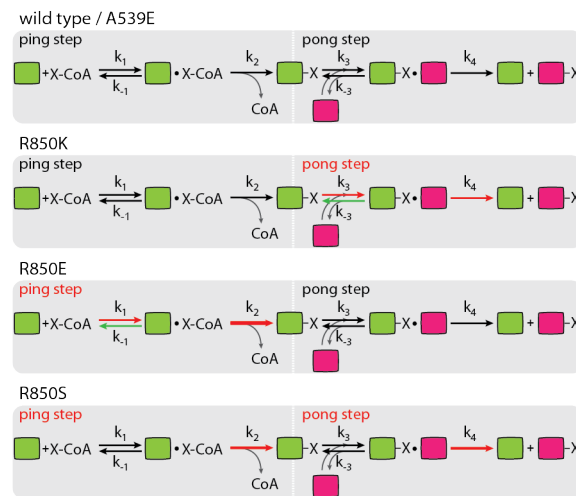


Figure 4.1: Plasticity of transacylation kinetics. Fluxes through the transacylation pathway in DEBS3M5 AT wild type and mutants. Green and red arrows indicate accelerated and decelerated reactions, respectively. The thicker the arrow, the more pronounced the effect. Only the most probable effects according to our interpretation are shown. For the mutant R850S, no statements can be given about changes in the kinetic constants k_1 , k_{-1} , k_3 , and k_{-3} due to complex effects not covered by our mathematical examination (see appendix). Headers "ping step" and "pong step" are in red when decelerated. Domains AT and ACP are depicted as green and pink boxes, respectively.

4.6 Implications of Interface Mutation Study on PKS Engineering

Electrostatic networks have previously been shown to dominate interactions of ACP with catalytic domains.^[14,221] Accordingly, we assumed that R850 with its cationic guanidinium headgroup is likely responsible for a such a key electrostatic interaction at the AT:ACP interface. Further support for R850 being involved in AT-ACP interaction was received from X-ray data in which residue R850 is poorly resolved in electron density in the KS-AT didomain structure (Fig. S22),^[99] indicating that R850 is not part of an electrostatic network of AT, but involved in AT-ACP interaction. Our kinetic data reflects the importance of the positive charge by the rather moderate effect of R850K mutation on the overall rate of transacylation compared to mutations R850S and R850E erasing and inverting the charge at this amino acid position, respectively. However, with help of the quantitative data, we likewise revealed that the mutations are not constrained to the AT-ACP interplay, but able to change kinetics of the entire transacylation reaction sequence (AT-loading ping step and ACP-loading pong step).

The incorporation of non-native elongation substrates to polyketides has previously been achieved by exchanging AT domains; i. e., replacing the native domain with a domain of interest.^[133,168,234] While such strategies can be successful in delivering the desired compound, the chimeric PKSs inherently suffer from non-cognate (permanent) interfaces in the KS-AT didomain as well as non-cognate (transient) interfaces during substrate shuttling (AT:ACP), which generally leads to compromised activity.^[20] Engineering domain-domain interplay can be useful when non-cognate domains are to be integrated in chimeric PKSs. Beyond the boundaries of the DEBS3M5 AT that has served as a model system in this study, the intricate impact of mutations on the catalytic properties suggests that simple engineering strategies based on local effects of amino acid exchanges are prone to fail. For example, engineering interfaces for stabilizing enzyme-substrate complexes (low K_m values) will not necessarily lead to high turnover rates and high catalytic efficiencies. However, the design of chimeric PKSs that perform product synthesis at high efficiency can then be successful, when including detailed enzyme kinetic analysis, ideally complemented by computational methods (as recently demonstrated with FAS^[235]) that eventually allows setting up highly effective enzyme kinetic processes.

4.7 Towards the Generation of a Functional Chimeric mFAS Module

In order to convert the iterative mFAS into an assembly line, several engineering aspects need to be considered (Fig. 2.10).

1. The design of the loading function: One needs to ensure, that the ATs of priming and elongation modules are highly specific for their substrates and discriminate against the respective other substrate so that there is no crosstalk between priming and elongation.
2. Design of an elongation module: Proper scaffold design needs to ensure intramodular communication between the different domains and processing of the substrate throughout the whole module (chain elongation and processing).
3. Design of covalent and/or non-covalent linkers between the modules: Proper scaffold design needs to enable intermodular communication (chain translocation), favoring the vectorial over the iterative working mode.

Previous studies have shown that domain swapping approaches often result in unstable and inactive proteins or proteins with reduced activity.^[6,20,206] Problems may occur regarding all three engineering aspects. In theory, I would expect that the implementation of a new loading function and design of a chimeric elongation module work best with similar systems. KS-AT structures from mFAS and DEBS3M5 revealed a high similarity.^[9,99] Additionally, DEBS3 AT5 met the criteria for donor ATs. It showed a high specificity for MMal, discriminating against priming substrate Ac and elongation substrate Mal, with sufficient transacylation rates of 1.3 molecules/s (75 molecules/min; Tab. 3.2). Thus, within our set of characterized ATs, this domain seemed to be ideally suited for swapping into mFAS to create an elongation module. To increase chances of successful chimera design and to ideally create two FAS-based modules specifically loading different elongation substrates, we additionally used the Mal-loading RAPS3 AT14, which shows a high sequence identity to mFAS and DEBS3M5 AT.

Even though structural information on mFAS and DEBS3M5 KS-AT as well as sequence alignments were used to find proper domain boundaries (Fig. 3.11), problems in scaffold design occurred. Chimeric FAS constructs with DEBS3 (LD-)AT5 and RAPS3 (LD-)AT14 gave drastically low yields and showed protein degradation (Tab. 3.4 and Fig. 3.12). Although proteins were only isolated in single replicates, these preliminary results point towards the following trends. For the LD-AT swap, we designed two different chimeras – one following the sequence alignment, the other following the structure alignment. In both chimeras (with DEBS and with RAPS), the proteins with domain boundaries chosen based on structure alignment resulted in significantly higher yields than the proteins with domain

boundaries based on the sequence alignment. Additionally, these constructs gave higher yields than the constructs with only the AT swapped. This might imply that the LD-AT swap creating an artificial KS:LD interface is better suited for domain swapping than the AT swap leading to an artificial LD:AT interface, which would be in accordance with a previous study.^[20]

These initial results confirmed that the design of a chimeric elongation module is quite challenging. We decided to first focus on the proper design of the loading function. For this, we reduced the complexity of the chimera by working on the KS-AT subconstruct to establish suitable artificial domain:domain interfaces. In the following, we focused on the DEBS3M5 construct, since structural information is beneficial to identify problems in the chimeric constructs. We designed eight different KS-AT chimeras with adapted domain boundaries after Yuzawa *et al.*'s strategy to design active chimeras (Fig. 3.13).^[20] They describe the boundaries of KAL-AT-PAL1 as universally applicable domain boundaries for AT swapping in most type I PKS modules. Expression and analysis of the chimeric didomains showed that these cannot be applied to mFAS (Tab. 3.5 and Fig. 3.14). Modified domain boundaries did not significantly improve protein stability and led to rather lower yields. In general, the constructs with swapped LD-AT resulted in higher expression yields than constructs with swapped AT, which was already observed in test expression of the full-length mFAS. TSA and SEC analysis showed that the artificial interfaces introduced in the didomain have serious effects on protein stability, folding, and formation of quaternary structures.

These rather disappointing results made us question the suitability of both wild types for the construction of chimeras. To understand the native interactions within the wild type and to estimate their robustness, we decided to analyze the KS-AT wild types via a mutation study. We used different bioinformatics tools to identify important residues in the KS:LD and LD:AT interfaces. 12 mutants were analyzed in the wild type mFAS KS-AT and revealed an exceptionally high robustness of this protein towards mutations (Tab. 3.7 and Fig. 3.16, 3.17, and 3.18). The residue D61 was found to be crucial for the KS:LD interaction in mFAS and alanine mutant D61A showed a drastical loss in stability compared to the wild type. Mutant I457A, located in the KS:LD interface, showed only minor destabilization effects, while all other mutants behaved like the wild type. Structural analysis revealed D61's role in forming a salt bridge between the KS and LD. Electrostatic interactions have been shown to influence protein folding and stability.^[236,237] To date, there is a controversy whether close-range electrostatic interactions like salt bridges are stabilizing or destabilizing proteins.^[236] However, our analysis reveals that only mutation of D61 and not mutation of its counterpart R429 cause drastic destabilization. This demonstrates that D61 has to be part of a complex electrostatic network not limited to the salt bridge, which dominates the intrinsic stability of the protein fold, just like shown above for protein-protein interactions.

Overall, these results demonstrate a high robustness of the mFAS KS-AT fold and imply that mFAS is in general highly suited for AT swapping approaches. DEBS3M5 KS-AT wild type, on the other hand, was very sensitive to any point mutations. All mutations introduced in the KS:LD and LD:AT interfaces led to drastically destabilized proteins (Tab. 3.9 and Fig. 3.21). This analysis revealed that this protein is not robust at all and not well suited for AT swapping. Differences in the robustness of the KS-AT fold of mFAS and DEBS towards mutations might again be explained by their evolutionary stage. Our findings support the above-mentioned hypothesis that in mFAS primordial enzyme properties were preserved which is reflected in its high robustness (shown in the present thesis for the KS-AT fold) and plasticity (shown for the AT in ref^[9]).^[238]

Based on that, we decided to use different ATs for swapping into mFAS. Like our analysis of FabD showed, it is a highly efficient enzyme discriminating against the mFAS priming substrate Ac (Tab. 3.2 and Fig. 3.4). We designed two chimeras (differing in length by three amino acids), which behaved very similar in protein expression and analysis (Tab. 3.10 Fig. 3.27). Although the chimeras showed destabilization, the AT activity substrate screening still showed approximately a third of wild type FabD activity in transferring Mal and MMal. As seen in the SDS-PAGE gel, the analyzed protein samples contained different species of degraded proteins that have been dragged along in the tandem affinity chromatography (Fig. 3.27A). Since FabD does not contain a Strep-tag, the samples should not include full-length FabD. Indeed, no band of the size of solely FabD was observed in the gel. The band of 60 kDa could correspond to KS-LD, stemming from purification of heterodimers formed between full-length and truncated protein chains. Thus, I expect the degraded proteins to not contribute to the measured AT activity. Taking degraded protein of around 30 % in the sample (as judged from SDS-PAGE, Fig. 3.27A) into account, one could assume around 43 % wild type activity of the pure chimera in the substrate screening, which is outstandingly high for destabilized proteins. This illustrates the high potential of FabD in domain swapping approaches.

In parallel, we designed a chimeric KS-AT construct with DEBS3M6 AT. My colleague Mirko Joppe showed that the mMAT swap into DEBS3M6 works better than the m(LD-MAT) swap. Our mutation study of the mFAS KS-AT wild type supports this finding. The results of the mFAS/DEBS3M5 chimera were not further considered. Due to the inherent fragility of the DEBS3M5 KS-AT wild type, chimeras were highly instable and gave drastically low yields which do not allow any conclusions regarding the general swapping strategy in mFAS. Thus, we designed two constructs for the AT only swap, which differed in length by only one amino acid (Fig. 3.28). According to test expression, this small variation in the domain boundary had a huge impact on the protein yield, which was significantly lower for the shorter construct (Tab. 3.11). Both proteins showed some degradation, but the longer construct showed defined

oligomeric states in SEC, which was not the case for the shorter chimera (Fig. 3.29). The stable chimera contains the original domain boundaries, but results in twentyfold increased expression yield compared to the mFAS/DEBS3M5 KS-AT chimera. This illustrates that evolutionary close ATs from the same PKS can feature fundamentally different properties and that even small changes in domain boundaries can have a huge impact on the stability of chimeras.

Based on these promising results on protein stability, we wanted to analyze the chimera in its AT activity. For comparison, we also cloned and expressed the wild type DEBS3M6 (KS-LD-AT) and the “inverse” chimera DEBS3M6(KS-LD)-m(MAT). The initial substrate screening showed comparable transacylation rates for the DEBS3M6 AT-containing chimera and DEBS3M6 wild type onto the corresponding ACP domain (Fig. 3.31A and B). Transacylation onto the mFAS ACP was also observed for both constructs with even higher rates for the chimeric protein. This could be explained by the active mFAS KS in the chimera contributing to the rate. On the other hand, this could be caused by the large measurement errors or it could be that only the apparent transfer rates differ, caused by differences in Michaelis-Menten constants. Apparent and absolute k_{cat} could still be similar, which can only be determined in full kinetic analysis. As expected, the “inverse”, mMAT-containing chimera featured overall higher transacylation rates onto the native ACP (Fig. 3.31C). Transfer rates onto the non-cognate DEBS3M6 ACP were in a similar range. This data shows that usage of both ACPs, cognate to KS or cognate to AT, might be possible in the full-length mFAS module. Tolerance of non-cognate ACPs was also reported for PKS's ATs before,^[20,142] whereas another study showed drastically decreased transacylation activity of DEBS2M3 AT with DEBS3M6 ACP.^[133] A study on DEBS2M3 KS showed that some ACPs from different modules are comparable substrates to ACP3, while others are not.^[239] This could also apply to mFAS KS.

A detailed kinetic analysis was performed for DEBS3M6 KS-AT wild type and m(KS-LD)-DEBS3M6(AT) chimera with the substrates MMal and DEBS3M6 ACP (Tab. 3.14 and Fig. 3.32). Transacylation rate of the chimera was still at 70% wild type rate, which represents a huge success. Interestingly, both Michaelis-Menten constants are lowered in the chimeric construct. This means that for ping and pong step, the ratio of non-productive dissociation/formation of the enzyme-substrate complex is lowered (accelerating the formation and/or decelerating the non-productive dissociation). Self-acylation of mFAS KS with MMal (as demonstrated for the m(KS-AT-ACP) construct) may contribute to kinetic parameters.

Discussion

Comparing kinetic parameters of DEBS3M5 and M6, transacylation rates in M6 were ten times slower than in M5 (Tab. 3.14 and 3.2). These slower transacylation rates are realistic, since the full-length DEBS was described with an overall slow production rate of 0.5 molecules/min.^[16] The Michaelis-Menten constant $K_m^{\text{MMal-CoA}}$ for M6 is in the range for bacterial metabolite level,^[219] but tenfold higher in M5. The data on M6 indicates a control mechanism for regulating MMal uptake in M5. The Michaelis-Menten constant K_m^{ACP} , on the other hand, is 3.5 times higher in M6 than in M5. This moderate variation was also observed for other modular PKSs and the pong step should still run at maximum rate (saturated ACP conditions). Overall, kinetic parameters determined for M6 are in good accordance with parameters determined for M3.^[133] This study showed that ATs of the same PKS may differ significantly in their kinetics.

Unfortunately, we were not able to solve the structures of chimeras and DEBSM6 KS-AT wild type. It would be very interesting to evaluate whether structural differences occur and if so, whether they can explain the variation in activity between the chimera and wild type. Furthermore, it would be helpful to compare structures of DEBS3M5 and M6 KS-AT, especially since the same domain boundaries were suitable for the design of chimeras with M6 but not with M5.

To understand differences in the success of chimera design with different donor ATs in more detail, I calculated the interfacial contact potential of the available wild type structures using PyMOL. Electrostatic interactions are important for the stability of proteins.^[236,237] In this discussion, non-specific interactions like hydrophobic interactions are not included.

The contact potential of mFAS and DEBS3M5 in their KS:LD and LD:AT interfaces are shown in Figure 4.2 and 4.3. For both systems, the contact between LD and AT domain seems to be tighter than the one between KS and LD. While a wide cleft can be seen between the KS and LD which seems to be mainly held together by only some selective interactions, the LD-AT contact seems to hold more close interactions across a wider area. This may point out that LD-AT swaps are in general more suited than solely AT swaps. Comparing both interfaces from both systems, it becomes obvious that the main regions of interaction do not fit to the respective other system (repelling interactions). There are more contacts of high electrostatic potential involved in the domain-domain interactions of DEBS3M5 than in the ones of mFAS. This may explain the finding that DEBS3M5 is not robust to any point mutation.

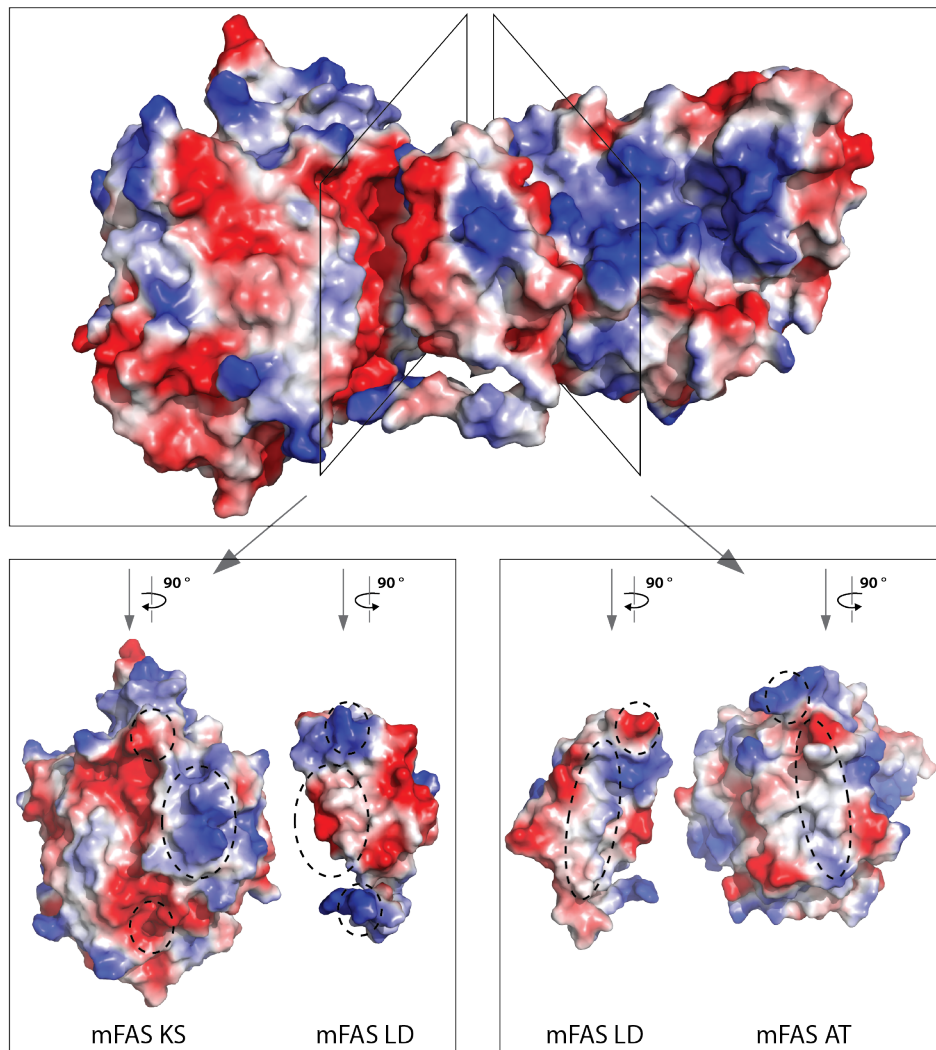


Figure 4.2: Contact potential of mFAS KS-AT wild type (PDB: 5MY0^[9]). KS:LD and LD:AT interfaces are shown on the bottom. Red indicates negative potential, blue indicates positive potential. White indicates a neutral potential. Regions of interactions are highlighted.

To analyze differences in domain swapping with DEBS3M5 and DEBS3M6 AT, I tried to identify the residues in DEBS3M5 disturbing the m(LD):DEBS3M5(AT) interface. Especially the upper interaction region of negative contact potential in DEBS3 AT5 (depicted in red) is problematic concerning the interaction with mFAS LD (Fig. 4.3). The lower region of negative contact potential may impair the interaction as well. I identified residue D814 in the upper region and residue D840 as well as the backbone formed by residue A857 in the lower region. Sequence alignment of DEBS3M5 and M6 showed the conservation of sites D840 (D2302 in M6) and A857 (A2314 in M6), but D814 of M5 is replaced by A2281 in M6. This residue variation could explain the significantly different behavior in domain swapping of DEBS3M5 and M6 AT. Introducing the mutation D814A in the mFAS/DEBS3M5 chimera could confirm or disprove this hypothesis.

Discussion

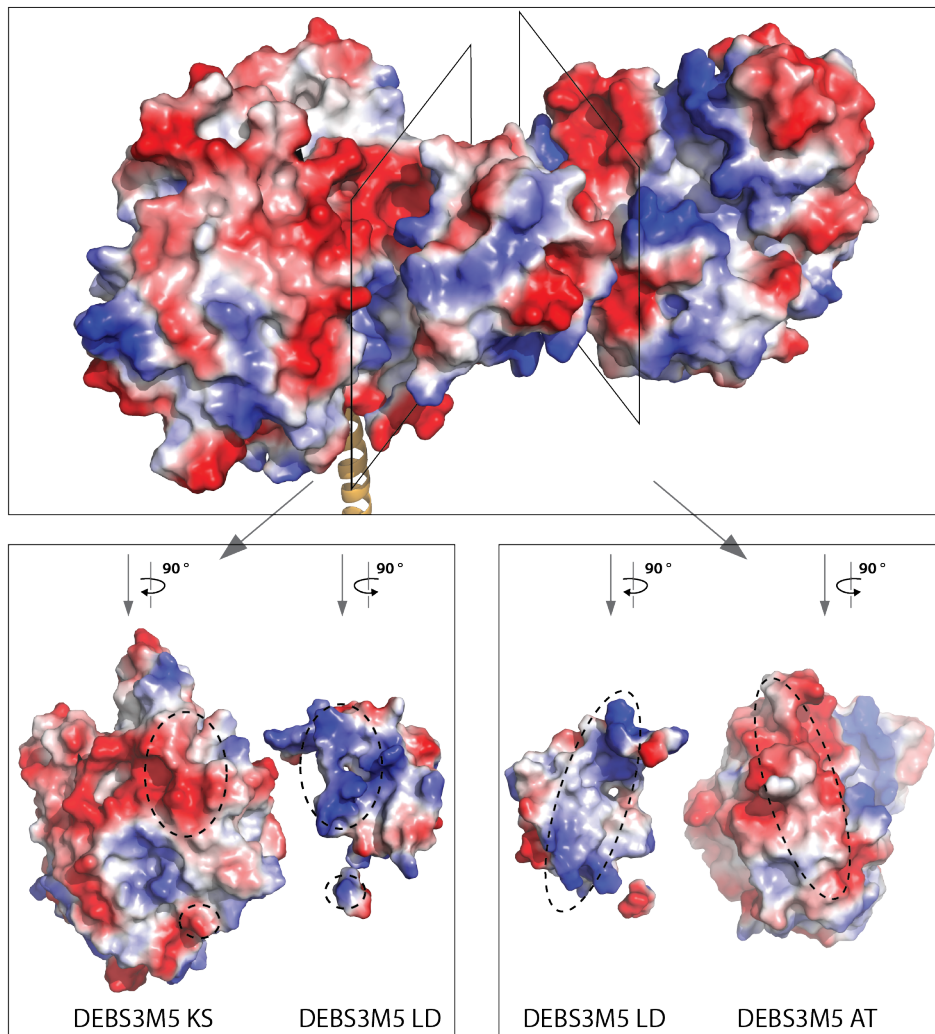


Figure 4.3: Contact potential of DEBS3M5 KS-AT wild type (PDB: 2HG4^[99]). KS:LD and LD:AT interfaces are shown on the bottom. Red and blue indicate negative and positive potentials, respectively. White indicates a neutral potential. Regions of interactions are highlighted.

FabD/mFAS chimeras showed destabilizing effects, but nonetheless AT activity (Fig. 3.27). Since FabD features outstandingly high transferase efficiencies (Tab. 3.2), these chimeras have a high potential for usage in an mFAS-based assembly line. Even though FabD and mMAT are overall similar (Fig. 4.4A), regarding the interfacial contact regions, it became obvious that mFAS LD and FabD are not compatible (Fig. 4.4B). Especially the lower region of FabD featuring the negative potential might hamper the interaction with LD leading to a loss of stability. This may be a good starting point for mutagenesis in order to improve the artificial domain-domain interaction. For introducing FabD in an mFAS module, the design of a proper AT-ACP linker is key to success.

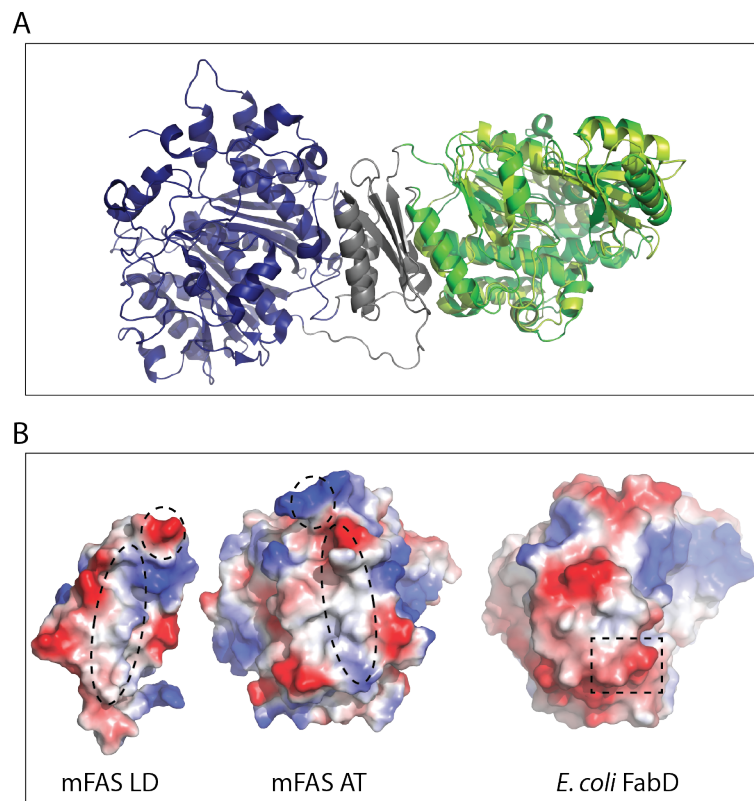


Figure 4.4: Analysis of FabD (A) Cartoon representation of mFAS KS-AT (PDB: 5MY0^[9]) and *E. coli* FabD (PDB: 1MLA^[136]). Alignment of mMAT (green) with *E. coli* FabD (lemon) demonstrates their structural similarity (RMSD = 2.823 Å). mFAS KS and LD shown in blue and gray, respectively. (B) Contact potential of mFAS LD and AT and FabD. Regions of LD:AT interactions are highlighted in mFAS. Region of repelling interaction is highlighted in FabD. Red and blue indicate negative and positive potentials, respectively. White indicates neutral potential.

4.8 Experimental Outlook: Implementation of Chimeras in mFAS Modules

At the end of my doctoral studies, Lynn Buyachuihan and I tried in a collaborative project to implement the active chimeric mFAS/DEBS3M6 KS-AT in an elongation module. To reduce complexity, we decided to use the minimal elongation module consisting of KS-AT-ACP. Activity of stable chimeras could then be analyzed in a minimal assembly line by following the production of a TAL derivative. The wild type full-length mFAS produces around 0.10 TAL/s (6.4 TAL/min), whereas the m(KS-MAT-ACP) construct is faster with around 0.28 TAL/s (16.5 TAL/min).^[17] Given the slow transacylation rate of the chimeric KS-AT of around 0.11 MMal/s (6.6 MMal/min; Tab.3.14), the transacylation reaction will probably limit the overall TAL derivative production rate significantly (maximal rate possible is around 0.055 TAL derivative/s = 3.3 TAL derivative/min).

Nevertheless, to see if the chimera can be installed in the elongation module (scaffold design), we designed two FAS modules with DEBS3M6 AT, based on the active chimera FS118 – one containing mFAS ACP, the other one containing DEBS3M6 ACP. Since the slow DEBS3M6 AT might not be well suited for efficient product formation, we decided to create further chimeric modules with different elongation substrate-specific ATs. Functional analysis of the modular venemycin PKS revealed relatively high turnover rates, which contains Mal-transferring ATs.^[206] Another study found FAS-like PKSs in sacoglossans and showed the substrate specificity for MMal of ATs from *Elysia chlorotica* PKS1 (EcPKS1).^[240] With its close evolutionary relationship to FASs, we expected its AT to be faster and more efficient than those from modular PKSs. Thus, we designed minimal FAS modules with swapped AT and swapped LD-AT domains from VemG, VemH, and EcPKS1 with the same domain boundaries as in FS37/FS118 (AT swapping) and FS38 (LD-AT swapping).

In a first approach, we decided to use the ACP from the PKSs and not from mFAS in the elongation module. This is especially promising for the venemycin chimeric constructs, since VemG and VemH naturally contain only the minimal set of domains. In doing so, only one artificial interface is installed in the elongation module. We tested their expression and isolation via tandem affinity chromatography. Our results showed that the stable chimeric subconstruct cannot be easily installed in the elongation module. Both mFAS/DEBS3M6 constructs showed low expression yields. Only one of the in total nine different elongation modules, mFAS(KS)-EcPKS1(LD-AT-ACP) (FS172), showed good stability and quality in our initial experiments (proper scaffold design).

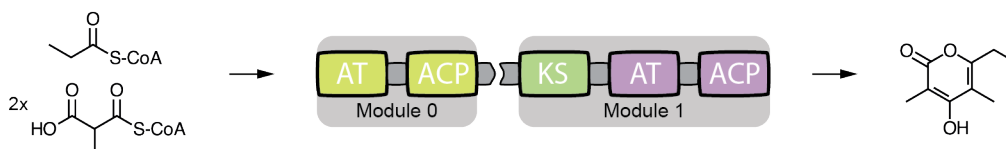


Figure 4.5: Assembly line consisting of a Prop-loading priming module (Module0) and the MMal-loading chimeric elongation module (Module1). Module0 (depicted in yellow) is the DEBS priming module. The elongation module consists of mFAS KS (green) and of AT-ACP from EcPKS1 (purple). Modules are non-covalently linked using DEBS3M4/5 DDs.

This module with the MMal-loading AT from EcPKS1 was analyzed with an established priming module^[209] in a minimal assembly line (Fig. 4.5). After loading of one priming and two elongation substrates, spontaneous cyclization/lactonization occurs leading to the release of a TAL derivative, which can be followed and quantified by absorption at 298 nm. Our initial experiments showed that we successfully eliminated the loading of the priming substrate Prop by the chimeric elongation module, but only minor TAL derivative production rates were detected. Different control experiments proved that the non-native substrates Prop and MMal are tolerated. All three above-mentioned engineering aspects could be cause of the low production rates. Further experiments on different parts of the assembly line will be needed to distinguish between the mentioned possibilities and to identify the cause of the slow production rate.

1. To analyze if the AT domain itself is limiting the TAL derivative production rate, the wild type EcPKS1 and chimera KS-AT should be cloned, isolated, and analyzed in the AT activity assay with mFAS and EcPKS1 ACP. Even the substrate screening is sufficient for first insights into transfer rates.
2. An intramodular communication problem could arise through the non-cognate (permanent) m(KS):EcPKS1(AT) interface as well as through the non-cognate (transient) m(KS):EcPKS1(ACP) interface. Experiments on a chimeric KS-AT construct as well as exchange of the EcPKS1 ACP domain by the mFAS ACP could be performed.
3. Intermodular communication problems could be caused by an additional StrepI-tag upstream to the DD5, which might impede the formation of the coiled-coil structure to link the priming and elongation module. Previous studies showed the acceptance of the N-terminal His-tag downstream to DD4.^[163,199,209] To ensure proper linkage of the modules, the purification tag could be removed, the DDs could be replaced by another pair of non-covalent linkers (SYNZIPs),^[163] or the two modules could be fused together on one polypeptide chain.
 - a) Removal of the purification tag: Previous experiments showed problems with an alternative purification strategy via His-tag affinity, ion exchange chromatography,

and SEC. Thus, the StrepI-tag needs to be removed after isolation, adding another purification step to the protocol.

- b) Replacement of DDs: Since a previous study has demonstrated the efficient linkage of two modules via the DDs,^[209] I would not expect them to be the bottleneck in intermodular module communication. However, different non-covalent linkers like SYNZIPS could be tested.
- c) Fusion of modules: This approach has the advantage that the two modules are forced to interact with each other. Nevertheless, there are two big disadvantages: First, the size of the fusion construct is quite problematic and even lower (or zero) yields would be received. Second, by fusing the two modules we will no longer be able to perform different control measurements with only one of the two modules. We would no longer be able to distinguish between different reactions and to trace back experimental problems to their root.

Overall, we made tremendous effort to realize the idea of an mFAS-based elongation module. We followed different approaches and used different strategies to tackle various problems. Even though we did not achieve this challenging aim yet, we have taken essential steps towards an mFAS-based assembly line. Self-priming by the chimeric elongation module was significantly suppressed and minor amounts of the expected TAL derivative were produced. I will discuss the different strategies in the following.

4.9 Strategies for Creating Functional Chimeras

Our initial strategy for the design of a functional FAS module was to use deeply characterized ATs for domain swapping in full-length mFAS. Key steps of this strategy are

1. identification of suitable donor ATs and
2. comparison of stability and activity of wild type and chimera.

A suitable donor AT is necessarily specific with respect to the elongation substrate, ideally efficient, and related to the acceptor system. Following this strategy, we made use of our extensive kinetic study on ATs. Based on structural information, we designed chimeras with DEBS3M5 and RAPS3M14 AT, which unfortunately all showed poor stability (Tab. 3.4 and Fig. 3.12).

After the limited success of the initial chimeric full-length mFAS, we chose to modify the strategy by reducing the complexity of the protein to the essential KS-AT subconstruct. This construct is much easier to modify and shows significantly higher yields. Thus, effects on protein stability can be identified more sensitively by analyzing differences in protein yields. After establishing proper artificial domain boundaries in the subconstruct, the chimera can in principle be implemented in the full-length module. The first experiments confirmed that domain swapping is challenging and fine-tuning might be needed. Thus, we decided to investigate numerous domain boundaries with focus on the DEBS3M5 AT, since structural information can support finding suitable interfaces. However, using the less complex subconstruct did not improve domain swapping success rates. We were not able to gain stable chimeras (Tab. 3.5 and Fig. 3.14). Thus, we had to take a step back and analyze the wild types to evaluate their compatibility for domain swapping. After detailed analysis via a mutation study to improve the domain boundaries, we finally came to the conclusion that the donor AT is not suited for domain swapping into mFAS due to its fragility towards mutations.

After this time-consuming and elaborate approach, we decided to create chimeric KS-AT constructs using known substrate-specific ATs as donor without their prior analysis. To save time, the donor AT was analyzed in kinetic detail after the design of stable chimeras in both wild type and chimera in parallel (Tab. 3.14). Following this strategy, we were able to create a stable and functional chimeric KS-AT with DEBS3 AT6 which showed only minor loss of transacylation activity, but overall low kinetic rates. Unfortunately, the donor AT itself features low transacylation activity. Thus, the active chimera was of limited value and not suited for the efficient biosynthesis of designer products.

Nevertheless, we tried to implement this active chimera as proof of principle in a minimal assembly line consisting only of a priming and an elongation module. We combined this experiment with the strategy of directly creating minimal FAS-modules with other donor ATs. We used the original domain boundaries for AT and LD-AT swap and chose AT domains from 1. venemycin synthase and from 2. EcPKS1.

1. We decided for two ATs from the venemycin synthase due to product rates of 0.6 molecules/s (36 molecules/min), which is quite high for modular PKSs.^[206] The venemycin PKS contains only two elongation modules. Assuming that transacylation is the slowest process in biosynthesis, transacylation rates have to be at least around 1.2 molecules/s (72 molecules/min).
2. EcPKS1, a sacoglossan PKS, is evolutionary close to FASs.^[240] No information on production rates is available, but our kinetic analysis and fundamental findings from A. Bar-Even and E. Noor *et al.*^[229] suggest that this close relationship might accompany with higher turnover rates. The chance of successful design of a stable chimera might even be increased by the similarity of the systems.

Indeed, this strategy yielded a stable chimera with EcPKS1 (FS172) (proper scaffold design of elongation module). Three of the eight other chimeras were only expressible with low yields (among them two mFAS/DEBS3M6 chimeras), the others were not isolated at all (poor scaffold design). Among the three expressible chimeras, there was also one chimera which differed only in one amino acid in the KAL (one amino acid less than in FS172). This, again, points towards the fragility of proper domain boundaries in chimeras. In order to evaluate the design of the loading function, the stable chimera was tested in a minimal assembly line consisting of a priming and the chimeric elongation module. The low activity in producing the TAL derivative in this setup can be caused by all three engineering aspects (design of the loading function, of the elongation module, of intermodular linkers) and one cannot clearly distinguish between them. I assume that the problem is most likely caused by the installation of the loading function. Analysis of the elongation module showed good stability and we made use of an established priming module with established intermodular linkers in our setup. The additional purification tag might impede the linkage. To get to the bottom of this, kinetic analysis of the AT activity in the wild type and the chimera is essential.

All three strategies are quite elaborate and demonstrate that problems can occur in each step of the assembly line design. Problems on later stages without prior establishment of the earlier steps often require to take a step back in order to identify the root of the problem. In my opinion, the best approach would be the stepwise increase of complexity – starting from analysis of donor ATs in robustness and activity,^{vii} over creating and investigating chimeric subconstructs prior to installation in elongation modules to the final assembly line. This strategy has the huge advantage that the cause of each problem can be identified more easily and each problem can be tackled separately. Changes on early levels are less time-consuming and resources are not wasted on ATs which are simply not suited for our approach. In fact, the last strategy of directly creating a chimeric elongation module was most successful, but only because it was build on the extensive groundwork of the others. However, it forced us to go back to the most basic level of wild type characterization.

^{vii}Even the initial substrate screening is able to provide valuable information and allows to draw conclusions on relative activity, while saving time and cost compared to full kinetic analysis.

4.10 Bioinformatics Tools in Chimeric Protein Design

We used different simple bioinformatics tools for the identification of interfacial residues and for estimation of mutational effects on the protein's energy. As expected, Rosetta, being the most elaborate tool among these, gave the most reliable output. Clear correlation between predictions and experimental data on the mFAS KS-AT wild type was recognizable for chain A and chain D (Fig. 3.22, 3.23, and 3.24). We chose both chains as they featured a high RMSD, but both had the highest quality in the structural solution (PDB: 5MY0^[9]). In retrospect, it would have been useful to analyze chain C (the chain with the lowest quality) as well, using Rosetta. For DEBS3M5 KS-AT, on the other hand, no correlation tendencies were observed (Fig. 3.25). This can be explained by the poor quality of the structure, manifesting itself in a low R_{free} value and high percentage of Ramachandran, sidechain, and real-space R-value Z-score (RSRZ) outliers (PDB: 2HG4^[99]). Rosetta analysis of mFAS KS-AT chains without PAL were especially helpful, as we identified two residues interacting with the linker simply by comparing the predicted differences in energy with and without PAL (Tab. 3.6). Residue D61 was predicted to interact with PAL to a greater extent than I457. And indeed, our experimental data supports this prediction. Overall, the importance of D61's interaction in the KS-AT can explain the stability problems in LD-AT swapping approaches in/with mFAS.

The tools with simpler algorithms than Rosetta (Robetta, Drug Score, mCSM, SpotOn) showed no correlation tendencies at all (Fig. S17). This discrepancy cannot be explained by program limitations to the usage of structures without PAL (Robetta, DrugScore, SpotOn) as the corresponding Rosetta predictions still showed clear correlation tendencies in mFAS KS-AT. Thus, our study indicates limitations of the programs and their algorithms themselves.

In conclusion, the simpler bioinformatics tools tested were not helpful for *in silico* analysis of the wild type. Rosetta, on the other hand, was shown to be useful for stability predictions of mutations under the premise that high quality structural information is available. Our protein engineering attempts have shown that the characterization of wild types may be key for successful strategies. Suitable *in silico* tools like Rosetta can save a lot of time by reliably predicting stability and robustness of the proteins involved. Furthermore, Rosetta's application in more complex *in silico* approaches on the chimeric level would be conceivable, since this method is not limited to alanine mutations. Proper modeling of chimeras might be combined with subsequent Rosetta calculations to estimate their stability. Promising candidates could then be experimentally analyzed regarding their stability and activity, finally paving the way to functional chimeras.

4.11 Conclusion and Outlook

FASs and PKSs are evolutionary closely related enzymes. The AT domains of both systems are key players in their biosynthesis and responsible for loading acyl substrates onto the ACP domains of these multidomain complexes. The set of AT domains we focused on in this thesis comprises high transacylation rates of 15 molecules/s for iterative PKSs and up to several hundred molecules per second for type I and II FASs for both priming and elongation substrates. The tolerance and polyspecificity of ATs of iterative PKSs/FASs towards natural and non-natural substrates was already reported in previous studies.^[9,230,241] In contrast, the ATs of modular PKSs appear to have significantly decreased turnover rates (0.7-7 molecules/s) and substrate tolerance and to be adjusted to their function in selecting the specific elongation substrates for the synthesis of intricate secondary metabolites. Confirming the findings of previous studies, the specificity of AT domains analyzed in this study is in general not assured by hydrolytic proofreading.^[133,143]

In addition to a basic characterization of AT kinetic properties, we showed that mutations at the interface between AT and ACP can lead to complex kinetic outputs, particularly when affecting the residues involved in the electrostatic network that guides and tunes enzyme-substrate interactions. We showed that replacing an arginine that is likely involved in a key electrostatic interaction at the AT:ACP interface by lysine, serine, or glutamate (R850K, R850S, R850E) leads to a severe drop in enzyme-substrate complex stability, catalytic efficiencies, and turnover rates. The impact of mutations can be explained by changing conformational properties that include those of active site and/or binding tunnel distant to the point mutation. The broadly changing kinetic parameters upon single mutations point towards a plasticity in the interaction of AT with acyl-CoA and ACP, as recently suggested for the AT-ACP interaction in the *E. coli* type II FAS (FabD and AcpP).^[14]

Our findings on the kinetic properties of AT domains are highly relevant for AT-centered PKS engineering approaches. The present study can provide guidance in PKS engineering for different strategies to modify the product spectrum. In cases where AT domains are engineered for loading non-natural elongation substrates, the promiscuous ATs of iterative systems may be best suited. Particularly, the MAT domain of the mammalian FAS seems ideal for such approaches, as it appears in functional properties and phylogeny as evolutionary close to a FAS/PKS common ancestor. As a primordial-like enzyme with polyspecificity and rapid turnover rates, it can be regarded as better suited to acquire novel or altered functions by mutation than ATs from modular PKSs.^[238,242] In other cases where AT domains are engineered to achieve the transacylation of a particular acyl compound of interest, two approaches are essentially possible: the introduction of specific mutations in the

AT binding site or the swap-in of an entire AT domain natively transferring the compound of interest.^[75] Both approaches have been performed successfully (for recent examples, see references^[20,243]). While the first approach might be hampered by partially unspecific loading of substrates, the latter suffers from protein instability and loss of activity. We recommend harnessing the intrinsic specificities of AT domains from the pool of modular PKSs whenever possible, because those domains will likely provide the required selectivity.

Following this strategy, we tried to design a functional chimeric mFAS-based elongation module. Analysis of the wild type mFAS KS-AT revealed an outstandingly high robustness of the fold. Showing this characteristic of a primordial enzyme, mFAS was proven to be highly suited for domain swapping approaches. In our study, we identified one amino acid as essential binding hotspot in the KS-AT, which seems to be part of an electrostatic network stabilizing the fold. Based on this, we were able to identify proper domain boundaries in mFAS for an AT swap. This lays the foundation for the design of a stable and active chimeric elongation module.

In further experiments we tried to swap different substrate-specific ATs from other systems in mFAS to create stable and active chimeras. The ATs – even different ATs from the same PKS – behaved differently in these experiments using the same domain boundaries. In contrast to a previous study in which universally applicable domain boundaries for AT swapping in type I PKS modules were published,^[20] we found that every particular system behaves differently and small variations in domain boundaries can have a huge impact on protein stability and activity.

Many of our chimeric constructs showed poor stability. Eventually, we designed a stable chimera using DEBS3 AT6 which showed 70% wild type donor AT activity. Unfortunately, the donor AT itself features slow transacylation rates and is of limited value for our application. As proof of principle, we implemented this chimera in a minimal elongation module, but received only low expression yields. Using the established domain boundaries, we were able to create another chimeric elongation module with EcPKS1 AT which featured high protein quality. Installing this stable chimeric elongation module in a minimal assembly line proved that priming by the elongation module was successfully suppressed. However, only minor product rates were detected.

Several parts of the assembly line might cause these limitations. The AT itself could feature slow transacylation rates, the artificial intramodular interactions in the chimeric elongation module might impede chain elongation, or the artificial intermodular interaction between priming and elongation module might hamper chain translocation. Identifying the root of the problem is key for optimizing the chimeric assembly line. Among other experiments, we will

have to take a step back and analyze the wild type AT. Analysis of the chimeric subconstruct KS-AT in the AT activity assay would provide information regarding the transacylation ability. The non-covalent intermodular linkers could be exchanged or replaced by a covalent linkage just like in modular PKSs. Overall, these findings strongly emphasize the value of detailed analysis and kinetic characterization of wild type donor ATs and chimeric subconstructs prior to installation in even a simple assembly line. For successful engineering approaches, detailed knowledge on the wild types should be considered in the design of chimeric assembly lines, ideally complemented by *in silico* approaches. Rosetta and the initial AT substrate screening could be used to estimate the robustness of the protein fold and the AT activity, respectively. This approach could enable selecting promising candidates for AT swapping in mFAS.

5 Experimental Procedures

Parts of this chapter have been previously accepted for publication in: F. Stegemann, M. Grininger, "Transacylation Kinetics in Fatty Acid and Polyketide Synthases and its Sensitivity to Point Mutations", *ChemCatChem* (2021).^[1] For individual contributions, copyright, and Creative Commons license, please see the statement of personal contributions.

5.1 Material and Methods

5.1.1 Molecular Cloning

Genes encoding for FAS and PKS constructs were cloned into pET-22b vector (Novagen), containing a pBR322 origin, ampicillin resistance, and *lacI* gene, which controls expression via the T7 polymerase system. Vectors encode an optional N-terminal Strep-tag and C-terminal His(6/8)-tag. Vectors were linearized by polymerase chain reaction (PCR) or enzymatic restriction. Phusion[®] High-Fidelity DNA Polymerase (NEB) and CloneAmp[™] HiFi PCR Premix (Clontech) were used to amplify genes of interest from template DNA by PCR with specific primers, generating complementary overhangs. Melting temperatures were calculated using oligocalc basic.^[244] Genes were inserted into the linearized vector via ligation independent cloning using the In-Fusion[®] HD Cloning Kit (Takara Bio). Multiple gene fragments were assembled by a MegaPrimer PCR to a single insert. For large PCR products (> 5,000 bp), template DNA was digested with DpnI (NEB). DNA was extracted from agarose gels using the NucleoSpin 3[®] Gel and PCR Clean-up (Macherey-Nage) or the Wizard SV Gel and PCR Clean-Up System (Promega).

Cloning products were transformed in *E. coli* Stellar[™] competent cells (Clontech) by heat shock. Cells were cultivated on liquid broth (LB)-agar (Lennox; Carl Roth) plates supplemented with 100 µg/mL ampicillin and 1% glucose at 37 °C overnight or at room temperature over the weekend. Plates were stored in the dark up to several weeks at 4 °C. Single colonies were used to inoculate 2-5 mL LB-media (Lennox; Carl-Roth) containing 100 µg/mL ampicillin and 1% glucose. Cells were cultivated at 37 °C overnight or at 20 °C over the weekend. Plasmids were isolated using the PureYield[™] Plasmid Miniprep System (Promega) and the GeneJET[™] Plasmid Miniprep Kit (Thermo Fisher Scientific). Plasmids were controlled by Sanger sequencing (Microsynth SeqLab).

An excerpt of vectors used in this study is listed in the appendix. Plasmids pAR and pMJD were kindly provided by Dr. Alexander Rittner and Mirko Joppe.

5.1.2 Protein Expression

Constructs were transformed in *E. coli* BL21-Gold (DE3) competent cells (Agilent Technologies) by heat shock. Constructs containing an ACP were co-transformed with a pETcoco vector (Novagen) encoding the 4'-PPT Sfp from *B. subtilis* or with a pCDF-1b vector (Novagen) encoding the 4'-PPT Npt from *Streptomyces platensis* for *in vivo* phosphopantetheinylation of the ACP. For mFAS ACP, a bicistronic vector encoding the ACP and Sfp from *B. subtilis* was used. Cells were cultivated on LB-agar (Lennox; Carl Roth) plates supplemented with 1% glucose and the respective selection antibiotics: 100 µg/mL ampicillin for single transformations, 50 µg/mL ampicillin and 17 µg/mL chloramphenicol for co-transformation with the pETcoco vector, and 50 µg/mL ampicillin and 25 µg/mL spectinomycin for co-transformations with the pCDF vector. Plates were incubated at 37 °C overnight or at room temperature over the weekend and were stored at 4 °C up to several days, subsequently.

20 mL LB-media (Lennox; Carl Roth) supplemented with the respective antibiotics and 1% glucose was inoculated with a randomly picked single clone. The pre-culture was incubated at 130-180 rpm and at 37 °C overnight or at 20 °C over the weekend. 1 L terrific broth (TB)-media was inoculated with the pre-culture and incubated at 37 °C and 130-180 rpm until an OD₆₀₀ of 0.6-0.8 was reached. The culture was cooled down for 15 min at 4 °C, prior to expression induction with 0.25 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Expression was carried out at 20 °C and 130-180 rpm overnight. Cells were harvested by centrifugation at 4,000 rpm for 20 min. Cell pellets were immediately purified or stored at -80 °C for several days.

5.1.3 Purification of AT Constructs and mFAS Modules

Cell pellets were resuspended in corresponding Ni Buffer (450 mM NaCl, 50 mM NaH₂PO₄, 10 mM imidazole, 20% v/v glycerol, pH 7.6) containing DNase I and 1 mM EDTA. Cells were mechanically disrupted at 1,000 bar using a French pressure cell press and centrifuged at 50,000 rcf and 4 °C for 1 h. After centrifugation, 2 mM MgCl₂ were added. Constructs were purified by tandem affinity chromatography. Lysate was loaded onto 5 mL Ni-NTA columns, unbound protein was washed off with 5 column volumes (CV) Ni Buffer. Bound protein was eluted with 2.5 CV Ni Buffer containing 300 mM imidazole. Subsequently, Ni elution fractions were loaded onto 5 mL Strep-Tactin columns (Iba), unbound protein was washed off with 6 CV Strep Buffer (500 mM NaCl, 50 mM NaH₂PO₄, 20% v/v glycerol, pH 7.6). Bound protein was eluted with 2.5 CV Strep Buffer containing 2.5 mM desthiobiotin. Elution fractions were analyzed by SDS-PAGE. Proteins were concentrated using Amicon Ultra Centrifugal Filters (Merck Millipore), flash frozen in liquid nitrogen, and stored in aliquots at -80 °C.

5.1.4 Purification of ACPs

Cell pellets were resuspended in corresponding Ni Buffer (200 mM NaCl, 50 mM NaH₂PO₄, 20 mM imidazole, 10 % v/v glycerol, pH 7.4) containing DNase I and 1 mM EDTA. Cells were mechanically disrupted at 1,000 bar using a French pressure cell press and centrifuged at 50,000 rcf and 4 °C for 1 h. After centrifugation, 2 mM MgCl₂ were added. *E. coli* ACP was activated *in vitro* with Sfp prior to purification. All ACPs were purified by Ni-affinity chromatography. Lysate was loaded onto 5 mL Ni-NTA columns, unbound protein was washed off with 5 CV Ni Buffer. Bound protein was eluted with 2.5 CV Ni Buffer containing 300 mM imidazole. ACPs were further purified by SEC using HiLoad 16/600 Superdex 200 pg column (200 mM NaCl, 50 mM NaH₂PO₄, 10 % v/v glycerol, pH 7.4, filtered and degassed). The column was calibrated with the Gel Filtration LMW and HMW Calibration Kits (GE Healthcare). Elution fractions were analyzed by SDS-PAGE. ACP was pooled after SEC to create one stock of each ACP as substrate used for all measurements of the AT activity assay. This should minimize errors in assay resulting from inconsistent protein expression/quality. Proteins were concentrated using Amicon Ultra Centrifugal Filters (Merck Millipore), flash frozen in liquid nitrogen, and stored in aliquots at -80 °C. Complete kinetic characterization of AT domains requires high yield and high quality of the corresponding ACPs. Expression resulted in insufficient yields of most ACPs. Using codon-optimized sequences for expression in *E. coli*, the *holo*-ACP yields increased significantly.

5.1.5 Protein Concentration

Protein concentration was determined from the absorbance at 280 nm using a NanoDrop 2000c (Thermo Fisher Scientific). Extinction coefficient and absorbance of 1 g/L at 280 nm (10 mm) were calculated from the primary sequence with CLC Main Workbench (QIAGEN Bioinformatics). Values are given in tables 5.1 to 5.5 for *E. coli* ACP with, for all other proteins without the N-terminal methionine.

Table 5.1: Absorbance of 1 g/L wild type proteins at 280 nm (10 mm).

Protein	Species	Abs.
ACP	mouse	0.475
	<i>E. coli</i>	0.131
	AVES1M0	0.650
	DEBS1M0	1.367
	DEBS3M5	0.842
	DEBS3M6	0.469
	PikAIIIM5	0.857
	RAPS3M14	0.917
	Pks7	1.500
(KS ⁽⁰⁾ -)AT	<i>E. coli</i>	1.171
	AVES1M0	1.187
	DEBS1M0	1.456
	DEBS3M5	1.021
	DEBS3M6	0.994
	PikAIIIM5	1.124
	RAPS3M14	1.208
	Pks7	1.266

Table 5.2: Absorbance of 1 g/L DEBS3M5 AT interface mutants at 280 nm (10 mm).

Construct	Mutation	Abs.
KS ⁰ -AT	A539S	1.021
	A539D	1.021
	A539E	1.021
	A539F	1.021
	R850K	1.022
	R850A	1.022
	R850E	1.022
	R850F	1.022
	R850S	1.022

Table 5.3: Absorbance of 1 g/L mFAS KS-AT mutants at 280 nm (10 mm).

Construct	Mutation	Abs.
mFAS (KS ⁰ -AT)	D61A	1.051
	N96A	1.051
	R429A	1.052
	I457A	1.051
	R468A	1.052
	Y470A	1.036
	H425A	1.052
	V479A	1.051
	E481A	1.051
	Q483A	1.051
	H804A	1.051
	L805A	1.051

Table 5.4: Absorbance of 1 g/L DEBS3M5 KS-AT mutants at 280 nm (10 mm).

Construct	Mutation	Abs.
DEBS3M5 (KS ⁰ -AT)	E129A	1.022
	D134A	1.021
	R510A	1.022
	V518A	1.021
	Y867A	1.009
	A868R	1.020

Table 5.5: Absorbance of 1 g/L chimeric proteins at 280 nm (10 mm).

Construct	Protein number	Swapped domain	Abs.
KS-AT	FS37	DEBS3AT5	0.900
	FS38	DEBS3(LD-AT)5	0.993
	FS57	DEBS3(AT)5	0.960
	FS58	DEBS3(AT)5	0.899
	FS59	DEBS3(AT)5	0.959
	FS60	DEBS3(LD-AT)5	0.990
	FS61	DEBS3(LD-AT)5	0.992
	FS62	DEBS3(LD-AT)5	0.989
	FS116	FabD	0.942
	FS117	FabD	0.939
	FS118	DEBS3(LD-AT)6	0.894
FS119	DEBS3(LD-AT)6	0.895	
KS-AT-ACP	FS172	EcPKS1(LD-AT-ACP)	0.963
FAS	FS31	DEBS3(AT)5	0.840
	FS32	DEBS3(LD-AT)5	0.871
	FS33	DEBS3(LD-AT)5	0.873
	FS34	RAPS3(AT)14	0.838
	FS35	RAPS3(LD-AT)14	0.919
	FS36	RAPS3(LD-AT)14	0.921

5.1.6 SDS-PAGE Analysis

Proteins were analyzed in their quality by SDS-PAGE. After incubation with Laemmli buffer at 95 °C for 5 min, a 10 µL sample of about 0.05 mg/mL protein dissolved in water was loaded. Self-cast 10 % (for 30-300 kDa proteins), 15 % (for 10-30 kDa proteins), and NuPAGE™ 4-12 % Bis-Tris Protein Gels (Invitrogen) were used. Gel electrophoresis was performed in Tris/Glycine/SDS running buffer (Bio-Rad) or NuPAGE™ MES SDS running buffer (Invitrogen) at 200 V for 45-60 min. InstantBlue™ Coomassie Protein Stain (Expediton) was used for staining.

5.1.7 Mass Spectrometric Analysis of ACPs

Full phosphopantetheinylation of ACPs was controlled by MS. Proteins were precipitated and resolved in sterile water. An ultimate 3000 RSLC system (Dionex) coupled to an impact II (Bruker, for non-codon optimized ACPs) or to a micrOTOF-Q II (Bruker, for codon optimized ACPs) equipped with an electrospray ionization source was used to perform HPLC-MS of ACPs. For chromatographic separation an RP-C3 column (Zorbax 300-SB, 300 Å, 150 mm × 3.0 mm × 3.5 μm, Agilent) was used with a mobile phase system consisting of solvent A (water with 0.1 % (v/v) formic acid) and solvent B (acetonitrile with 0.1 % (v/v) formic acid). After equilibration with 15 % solvent B and a flow of 0.6 mL/min for 1.5 min, the concentration of solvent B was linearly increased to 65 % over 22.5 min. This was followed by a linear increase of solvent B to 95 % over 4 min prior to re-equilibration of 15 % solvent B for 2 min. Proteins were detected by absorbance at 280 nm. ACPs were found to elute at 3.5-16 min. MS data was acquired in positive mode in the range of 50-2000 m/z and analyzed using DataAnalysis 4.3 software (Bruker Dalton GmbH).

5.1.8 CoA-488 Assay

The CoA-488 assay was used as fast in-house method for analysis of ACP activation. ACP was incubated with 0.5 equivalents purified Sfp, 5 equivalents CoA-488 (NEB), and 5,000 equivalents MgCl₂ protected from light at 37 °C and 350 rpm for 1 h. Laemmli buffer was added and the sample was incubated at 95 °C for 5 min. 10 μL sample was loaded onto a NuPAGE™ Bis-Tris 4-12 % Protein Gel (Invitrogen). Electrophoresis was carried out in NuPAGE™ MES SDS running buffer (Invitrogen) at 200 V for 1 h. In-gel fluorescence was detected at the Fusion SL Fluorescence Imaging System (VilberLourmat) at 488 nm, prior to gel staining with InstantBlue™ Coomassie Protein Stain (Expedeon).

5.1.9 HPLC-SEC Analysis

Quality and oligomeric state of all AT wild types and chimeric proteins were analyzed by HPLC-SEC (500 mM NaCl, 50 mM NaH₂PO₄, 5 % v/v glycerol, pH 7.6) at room temperature with a flow of 0.3 ml/min. An UltiMate 3000 RSLC (Dionex) equipped with an UV-vis array detector was used with SEC columns bioZen 1.8 μm SEC-3, Yarra-SEC-X150, and Yarra-SEC-X300. Proteins were detected by absorbance at 280 nm. Buffers and samples were filtered through 0.1 μm. 8 μL protein sample was injected. Columns were calibrated with the Gel Filtration LMW and HMW Calibration Kits (GE Healthcare).

5.1.10 Thermal Shift Assay

TSA was performed to analyze all constructs in their stability using a CFX96 Touch Real-Time PCR Detection System (BioRad) with excitation and emission wavelength set to 450-490 and 560-580 nm, respectively. Fluorescence of the dye SYPRO Orange was measured from 5 to 95 °C with 0.5 °C steps per minute in white Multiplate™ 96-well PCR Plates (BioRad). Melting curves of all wild type AT constructs and ACPs were measured in storage buffer and AT activity assay buffer. DEBS3M5 AT:ACP mutants were measured in storage buffer. mFAS and DEBS3M5 KS-AT wild types and mutants were measured in four different buffers. Chimeric constructs were analyzed in four different buffers or in storage buffer. Analysis was performed using OriginPro 9.1.

5.1.11 AT Activity Assay

The AT activity assay was adapted from references.^[9,133,211] It couples formation of free CoA with the reduction of nicotinamide adenine dinucleotide (NAD⁺), which can then be followed fluorimetrically, allowing the quantification of transferase activity. The acyl transfer is measured in presence and absence of ACP, giving access to transacylation and hydrolysis rates, respectively. Kinetic hydrolysis parameters are calculated by titration of X-CoA. To determine absolute kinetic transacylation parameters, X-CoA is titrated to a series of fixed ACP concentrations.

The AT activity assay buffer (50 mM NaH₂PO₄, 1 mM EDTA, 1 mM DTT added directly before usage, 10% v/v glycerol, pH 7.6) was used to set up all other solutions for this assay. For all proteins and substrates, fresh aliquots were used. For each system, the same batch of substrates (ACP and acyl-CoAs) was used. Four master mixes in 4 × concentration were prepared. For the acyltransferase solutions the proteins were diluted to 0.1-400 nM final assay concentration with assay buffer supplemented with 0.1 mg/mL BSA. The ACP solution was prepared by diluting highly pure and concentrated ACP to 2.5-530 μM final concentration with the assay buffer. Substrate acyl-CoA solution was also prepared with the assay buffer in different concentrations ranging from 1 μM to 1 mM final concentration. The read out solution contained 2 mM α-ketoglutaric acid, 0.4 mM NAD⁺, 0.4 mM TPP and 5 to 15 mU/100 μL αKGDH as final concentrations. All solutions were pre-heated/incubated at 25 °C for at least 5 min. Assays were performed in 96-well f-bottom microtiter plates (Greiner Bio-one). For transacylation, 25 μL of each solution was added, ACP was added via the dispenser to initiate the reaction. For hydrolysis, assay buffer was used to compensate the volume of ACP. For background measurements the AT solution was replaced by 0.1 mg/mL BSA in assay buffer. NADH fluorescence was monitored for 5 min at 25 °C, taking equidistant measurements every 20 s using a ClarioStar microplate reader (BMG labtech). The following

Experimental Procedures

settings were used: excitation: 348-20 nm, emission: 476-20 nm, gain: 1900, focal height: 5.2 mm, flashes: 70, orbital averaging: 4 mm. For kinetic characterization of the chimeras FS118 and FS139 as well as the wild type FS139, the assay was performed in 384 Well Small Volume™ HiBase Microplates (Greiner Bio-One) with a total volume of 20 μ L and the following settings: excitation: 348-20 nm, emission: 476-20 nm, gain: 1500, focal height: 5.2 mm, flashes: 17.

To meet Michaelis-Menten requirements, the substrate concentration was always kept significantly higher than the enzyme concentration, at least ten times.^[129] For hydrolysis and for each ACP curve in transacylation measurements, initial experiments were performed to estimate the (apparent) K_m^{X-CoA} . Based on these, the acyl-CoA substrate concentration was varied to calculate kinetic parameters: (0.2; 0.3; 0.5; 0.75; 1.25; 2; 3; 5) $\times K_m^{X-CoA}$. In order to ensure that the measurement is performed in the initial quasi-linear range of substrate consumption, the product concentration has to be significantly lower than the substrate concentration.^[129] The increase during measurement was always linear and substrate consumption was ensured to be below 10% for the majority of assays performed. Some measurements showed significantly higher substrate consumption. In such cases, those measurements were included, as long as the increase was shown to be linear and the Michaelis-Menten fit as well as the kinetic parameters do not change considerably. If not mentioned otherwise, transacylation was performed in biological triplicates. Hydrolysis was performed in technical triplicates of biological triplicates or technical/biological triplicates. Background values were subtracted from the average of technical/biological measurement values. Occasional dust or air bubbles caused strong outliers. These were omitted for further analysis after careful analysis. The time span of 40-120 and 20-120 s was analyzed for transacylation and hydrolysis, respectively. For high substrate consumption, this time span was adapted. In order to convert relative fluorescence units (RFU) into concentration, an NADH calibration curve was measured (Fig. S23). Hydrolysis kinetic parameters were calculated by Michaelis-Menten plots using OriginPro 9.1. Transacylation parameters were calculated by global Michaelis-Menten fits using OriginPro 9.1.

5.1.12 TAL Assay

For quantification of TAL derivative production, the TAL assay was performed in glass cuvettes Suprasil type no. 117.104-QS 10 mm (Hellma) at NanoDrop 2000c (Thermo Fisher Scientific) in cuvette mode, following the absorbance at 298 nm and 25 °C. Priming and elongation module as well as priming (Prop or Ac) and elongation substrate (MMal) were pipetted together in assay buffer (400 mM NaPi, 4% glycerol, 5 mM TCEP, pH 7.2). Reaction was started with elongation substrate. Final concentrations were 0 or 100 μ M substrates, 0-4 μ M priming module, 1-3 μ M elongation module. The molar extinction coefficient at 298 nm $\epsilon=2540 \text{ M}^{-1}\text{cm}^{-1}$ for TAL was used for conversion of absorbance into concentration.^[245]

5.1.13 Design of DEBS3M5 AT:ACP Interface Mutations

Since the AT:ACP interface of DEBS3M5 and the structure of its ACP is not known, Swiss-Model was used to model the ACP domain.^[246] The sequence was obtained from the Uniprot database (Q03133, residues 1392-1467). As structural template served the ACP from type I modular PKS from *Mycobacterium ulcerans* solved by nuclear magnetic resonance (NMR) spectroscopy (PDB: 5HVC, identity percentage: 47 %, similarity: 0.40). Global model quality estimation (GMQE) by SwissModel was 0.73, and the quality model energy analysis (QMEAN) was -1.93. Subsequently, three AT:ACP interfaces were used as template for structural alignment of the modeled ACP and the structure of AT domain (PDB: 2HG4^[99]). The *in silico* docking study on human FAS as well as crystal structures of the AT:ACP complex of vicanistatin PKS VinK:VinL and of disorazole synthase (DSZS) served as templates.^[125,126,159] Based on these three alignments, PISA was used to identify hotspots in the AT:ACP interface.^[247] All three models gave the amino acids A539 and R850 as possible interfacial hotspots.

5.1.14 Bioinformatics Tools

Different bioinformatics tools were used to predict hotspots in the mFAS and DEBS3M5 wild type KS:LD and LD:AT interfaces, namely Prodigy,^[248] Rosetta,^[249] Robetta,^[250] DrugScore,^[251] mCSM,^[252] and SpotOn.^[253] These programs use different algorithms to predict interactions between chains or effects of mutations on protein stability based on structural information (PDBs: 5MY0 and 2HG4). For this purpose, some programs require splitting of crystal structures into chains for the three domains KS, LD, and AT. For these programs the PAL (Prodigy, Robetta, DrugScore, SpotOn) was neglected, which potentially can influence the predictions, since interactions with the PAL are not considered.

The program Prodigy was used to identify interactions within the two interfaces. This program gives a list of interacting residues of two separate chains with the classification in weak and strong interactions. This list of predicted interactions was analyzed in the crystal structure. Potential interesting residues were further analyzed with the other bioinformatics tools.

Experimental Procedures

The program mCSM performs a random mutation scan on a single residue, resulting in a list of all possible amino acid mutations with predicted changes in Gibbs free energy $\Delta\Delta G = \Delta G(\text{WT}) - \Delta G(\text{Mut})$ in kcal/mol, with the Gibbs free energy of wild type (WT) and mutant (Mut). For destabilizing effects of the mutation introduced, the predicted $\Delta\Delta G$ is negative, for stabilizing effects, the predicted $\Delta\Delta G$ is positive. All other programs define the change in Gibbs free energy as $\Delta\Delta G = \Delta G(\text{Mut}) - \Delta G(\text{WT})$. For better comparability, the sign was changed for mCSM predictions. The higher (positive) the change in Gibbs free energy, the more destabilizing a mutation.

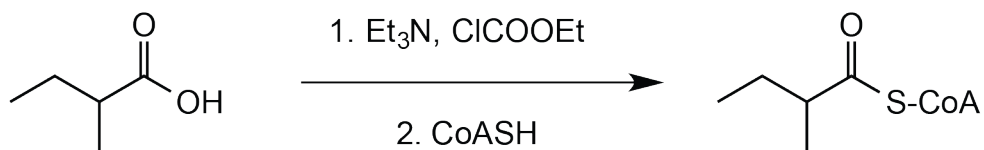
The program Rosetta performs a random point mutation scan on all residues. Similar to mCSM the output is a list of all possible amino acid mutations with predicted changes in Gibbs free energy in REU, which is a generic energy unit.

The programs Robetta and DrugScore are limited to an alanine scan. Predictions are made for alanine mutation of all residues. Both tools give a list of mutations with specific predicted changes in Gibbs free energy in kcal/mol. The program DrugScore provides additional information on degree of buriedness and identifies possible saltbridges.

The program SpotOn is analyzing two chains in their interfacial contacts, classifying them in so-called Nullspots and Hotspots. Based on alanine mutations, the change in Gibbs free energy is predicted. Nullspots are defined as mutations with predicted changes < 2 kcal/mol, hotspots as mutations with predicted changes ≥ 2 kcal/mol. A list of predicted Nullspots and Hotspots is provided without specific $\Delta\Delta G$ values.

Informations of all programs were compared. Those mutations were introduced, which were most likely to have a huge impact on proteins stability according to the different tools.

5.2 Synthesis of 2-Methylbutyryl-CoA



2-Methylbutyric acid (34.5 mg, 98 %, 0.33 mmol, 6.6 equiv.) was dissolved in THF (2 mL) and cooled to 0 °C. Et₃N (40 μL) and ethyl chloroformate (27 μL, 0.28 mmol, 5.6 equiv.) were added and the solution was stirred for 45 minutes at 0 °C. The reaction was centrifuged for 4 minutes at 5,000 × g for removal of insoluble salts. 1.5 mL of supernatant were added to CoA Li₃ (40.1 mg, 0.051 mmol, 1.0 equiv.) and NaHCO₃ (32.1 mg, 0.38 mmol, 7.6 equiv.) dissolved in water (2 mL). After stirring for 1 hour at room temperature, the solvents were removed using a SpeedVac. The raw product was dissolved in water (2 mL) and purified by preparative HPLC. Solvents were removed by lyophilization overnight, yielding 2-methylbutyryl-CoA (38.0 mg, 0.045 mmol, 88 %) as white powder. Synthesis adopted from references.^[254,255]

¹H-NMR (500 MHz, MeOH-d₄) δ (ppm) = 8.63 (s, 1H), 8.42 (s, 1H), 6.20 (d, ³J = 5.00 Hz, 1H), 4.93-4.85 (m, 2H), 4.60-4.59 (m, 1H), 4.33-4.23 (m, 2H), 4.01 (s, 1H), 3.91-3.85 (m, 1H), 3.66-3.61 (m, 1H), 3.44 (t, ³J = 5.00 Hz, 2H), 3.32 (t, ³J = 5.00 Hz, 2H), 2.98 (t, ³J = 5.00 Hz, 2H), 2.68-2.61 (m, 1H), 2.42 (t, ³J = 5.00 Hz, 2H), 1.64-1.55 (m, 1H), 1.48-1.38 (m, 1H), 1.11-1.08 (m, 3H), 0.95-0.93 (m, 3H), 0.84-0.81 (m, 6H).

MALDI-MS (m/z): [M+H]⁺ calc. for C₂₆H₄₅N₇O₁₇P₃S: 852.18000, found: 852.18009.

References

- [1] Stegemann, F. & Grninger, M. Transacylation Kinetics in Fatty Acid and Polyketide Synthases and its Sensitivity to Point Mutations. *ChemCatChem* (2021, accepted article). URL <https://doi.org/10.1002/cctc.202002077>.
- [2] Nussbaum, F. V., Brands, M., Hinzen, B., Weigand, S. & Dieter, H. Antibacterial Natural Products in Medicinal Chemistry — Exodus or Revival? *Angewandte Chemie - International Edition* **45**, 5072–5129 (2006).
- [3] Smith, S. & Tsai, S.-C. The Type I Fatty Acid and Polyketide Synthases: A Tale of Two Megasyntases. *Natural Product Reports* **24**, 1041–1072 (2007).
- [4] Staunton, J. & Weissman, K. J. Polyketide biosynthesis: A millennium review. *Natural Product Reports* **18**, 380–416 (2001).
- [5] Hertweck, C. The biosynthetic logic of polyketide diversity. *Angewandte Chemie - International Edition* **48**, 4688–4716 (2009).
- [6] Nivina, A., Yuet, K. P., Hsu, J. & Khosla, C. Evolution and Diversity of Assembly-Line Polyketide Synthases. *Chemical Reviews* **119**, 12524–12547 (2019).
- [7] Maier, T., Leibundgut, M., Boehringer, D. & Ban, N. Structure and function of eukaryotic fatty acid synthases. *Quarterly Reviews of Biophysics* **43**, 373–422 (2010).
- [8] Brignole, E. J., Smith, S. & Asturias, F. J. Conformational flexibility of metazoan fatty acid synthase enables catalysis. *Nature Structural and Molecular Biology* **16**, 190–197 (2009).
- [9] Rittner, A., Paithankar, K. S., Huu, K. V. & Grninger, M. Characterization of the Polyspecific Transferase of Murine Type I Fatty Acid Synthase (FAS) and Implications for Polyketide Synthase (PKS) Engineering. *ACS Chemical Biology* **13**, 723–732 (2018).
- [10] Schweizer, E. Microbial Type I Fatty Acid Synthases (FAS): Major Players in a Network of Cellular FAS Systems. *Microbiology and Molecular Biology Reviews* **68**, 501–517 (2004).
- [11] Tehlivets, O., Scheuringer, K. & Kohlwein, S. D. Fatty acid synthesis and elongation in yeast. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids* **1771**, 255–270 (2007).
- [12] Erdbrügger, P. & Fröhlich, F. The role of very long chain fatty acids in yeast physiology and human diseases. *Biological Chemistry* **402**, 25–38 (2020).

- [13] Fersht, A. *Structure and mechanism in protein science . A guide to enzyme catalysis and protein* (W. H. Freeman and Company, New York, 1999), 2nd print edn.
- [14] Misson, L. E. *et al.* Interfacial plasticity facilitates high reaction rate of *E. coli* FAS malonyl-CoA:ACP transacylase, FabD. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 24224–24233 (2020).
- [15] Chang, S. I. & Hammes, G. G. Structure and Mechanism of Action of a Multifunctional Enzyme: Fatty Acid Synthase. *Accounts of Chemical Research* **23**, 363–369 (1990).
- [16] Pieper, R., Ebert-Khosla, S., Cane, D. & Khosla, C. Erythromycin biosynthesis: Kinetic studies on a fully active modular polyketide synthase using natural and unnatural substrates. *Biochemistry* **35**, 2054–2060 (1996).
- [17] Rittner, A., Paithankar, K. S., Drexler, D. J., Himmler, A. & Grininger, M. Probing the modularity of megasynthases by rational engineering of a fatty acid synthase Type I. *Protein Science* **28**, 414–428 (2019).
- [18] Weissman, K. J. & Leadlay, P. F. Combinatorial biosynthesis of polyketides. *Nature Reviews Microbiology* **3**, 925–936 (2005).
- [19] Rangan, V. S. & Smith, S. Alteration of the substrate specificity of the malonyl-CoA/acetyl-CoA:acyl carrier protein S-acyltransferase domain of the multifunctional fatty acid synthase by mutation of a single arginine residue. *Journal of Biological Chemistry* **272**, 11975–11978 (1997).
- [20] Yuzawa, S. *et al.* Comprehensive in vitro analysis of acyltransferase domain exchanges in modular polyketide synthases and its application for short-chain ketone production. *ACS Synthetic Biology* **6**, 139–147 (2017).
- [21] Newman, D. J. & Cragg, G. M. Natural Products as Sources of New Drugs from 1981 to 2014. *Journal of Natural Products* **79**, 629–661 (2016).
- [22] Wani, M. C., Taylor, H. L., Wall, M. E., Coggon, P. & McPhail, A. T. Plant antitumor agents. VI. Isolation and structure of taxol, a novel antileukemic and antitumor agent from *Taxus brevifolia*. *J. Am. Chem. Soc.* **93**, 2325–2327 (1971).
- [23] Schwecke, T. *et al.* The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 7839–7843 (1995).
- [24] Ansdell, V. E., Boosey, C. M., Geddes, A. M. & Morgan, H. V. Malaria in Birmingham 1968–73. *British Medical Journal* **2**, 206–208 (1974).

- [25] Wright, G. D. Unlocking the potential of natural products in drug discovery. *Microbial Biotechnology* **12**, 55–57 (2019).
- [26] Cortes, J., Haydock, S. F., Roberts, G. A., Bevitt, D. J. & Leadlay, P. F. An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of *Saccharopolyspora erythraea*. *Nature* **348**, 176–178 (1990).
- [27] Brockmann, H. & Henkel, W. Pikromycin, ein bitter schmeckendes Antibioticum aus Actinomyceten. *Chemische Berichte* **84**, 284–288 (1951).
- [28] Margalith, P. & Pagani, H. Rifomycin. XIV. Production of rifomycin B. *Applied Microbiology* **9**, 325–334 (1961).
- [29] Civjan, N. (ed.) *Natural Products in Chemical Biology* (Wiley, Hoboken, NJ, 2012).
- [30] Spellberg, B. *et al.* The epidemic of antibiotic-resistant infections: A call to action for the medical community from the infectious diseases society of America. *Clinical Infectious Diseases* **46**, 155–164 (2008).
- [31] Peregrín-Alvarez, J. M., Tsoka, S. & Ouzounis, C. A. The phylogenetic extent of metabolic enzymes and pathways. *Genome Research* **13**, 422–427 (2003).
- [32] Peregrín-Alvarez, J. M., Sanford, C. & Parkinson, J. The conservation and evolutionary modularity of metabolism. *Genome Biology* **10**: R63 (2009).
- [33] Dias, D. A., Urban, S. & Roessner, U. A Historical overview of natural products in drug discovery. *Metabolites* **2**, 303–336 (2012).
- [34] Smith, S. The animal fatty acid synthase: one gene, one polypeptide, seven enzymes. *The FASEB Journal* **8**, 1248–1259 (1994).
- [35] Iburguren, M., López, D. J. & Escribá, P. V. The effect of natural and synthetic fatty acids on membrane structure, microdomain organization, cellular functions and human health. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1838**, 1518 – 1528 (2014).
- [36] Dibrova, D. V., Galperin, M. Y. & Mulkidjanian, A. Y. Phylogenomic reconstruction of archaeal fatty acid metabolism. *Environmental Microbiology* **16**, 907–918 (2014).
- [37] Beld, J., Lee, D. J. & Burkart, M. D. Fatty acid biosynthesis revisited: structure elucidation and metabolic engineering. *Molecular BioSystems* **11**, 38–59 (2015).
- [38] Herbst, D. A., Townsend, C. A. & Maier, T. The architectures of iterative type I PKS and FAS. *Natural Product Reports* **35**, 1046–1069 (2018).

- [39] White, S. W., Zheng, J., Zhang, Y. M. & Rock, C. O. The structural biology of type II fatty acid biosynthesis. *Annual Review of Biochemistry* **74**, 791–831 (2005).
- [40] Hiltunen, J. K. *et al.* Mitochondrial fatty acid synthesis and respiration. *Biochimica et Biophysica Acta - Bioenergetics* **1797**, 1195–1202 (2010).
- [41] Heil, C. S., Wehrheim, S. S., Paithankar, K. S. & Grininger, M. Fatty Acid Biosynthesis: Chain-Length Regulation and Control. *ChemBioChem* **20**, 2298–2321 (2019).
- [42] Jenke-Kodama, H., Sandmann, A., Müller, R. & Dittmann, E. Evolutionary implications of bacterial polyketide synthases. *Molecular Biology and Evolution* **22**, 2027–2039 (2005).
- [43] Hopwood, D. A. Genetic contributions to understanding polyketide synthases. *Chemical Reviews* **97**, 2465–2497 (1997).
- [44] Wang, B., Guo, F., Huang, C. & Zhao, H. Unraveling the iterative type I polyketide synthases hidden in *Streptomyces*. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 8449–8454 (2020).
- [45] Helfrich, E. J. N. & Piel, J. Biosynthesis of polyketides by *trans*-AT polyketide synthases. *Natural Product Reports* **33**, 231–316 (2016).
- [46] Broadhurst, R. W., Nietlispach, D., Wheatcroft, M. P., Leadlay, P. F. & Weissman, K. J. The Structure of Dockin Domains in Modular Polyketide Synthases. *Chemistry & Biology* **10**, 723–731 (2003).
- [47] Dorival, J. *et al.* Characterization of Intersubunit Communication in the Virginiamycin *trans*-Acyl Transferase Polyketide Synthase. *Journal of the American Chemical Society* **138**, 4155–4167 (2016).
- [48] Funa, N. *et al.* A new pathway for polyketide synthesis in microorganisms. *Nature* **400**, 897–899 (1999).
- [49] Paitan, Y., Alon, G., Orr, E., Ron, E. Z. & Rosenberg, E. The first gene in the biosynthesis of the polyketide antibiotic TA of *Myxococcus xanthus* codes for a unique PKS module coupled to a peptide synthetase. *Journal of Molecular Biology* **286**, 465–474 (1999).
- [50] Shen, B. Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Current Opinion in Chemical Biology* **7**, 285–295 (2003).
- [51] Wakil, S. J. Fatty acid synthase, a proficient multifunctional enzyme. *Biochemistry* **28**, 4523–4530 (1989).

References

- [52] Joshi, V. & Wakil, S. J. Studies on the mechanism of fatty acid synthesis. XXVI. Purification and properties of malonyl-coenzyme A-acyl carrier protein transacylase of *Escherichia coli*. *Archives of Biochemistry and Biophysics* **143**, 493–505 (1971).
- [53] McCarthy, A. D. & Grahame Hardie, D. Evidence that the acyl-O-esters are intermediates in the catalysis. The mechanism of rabbit mammary fatty acid synthase. *FEBS Letters* **150**, 181–184 (1982).
- [54] Majerus, P. W., Alberts, A. W. & Vagelos, P. R. Acyl Carrier Protein, IV. the Identification of 4'-Phosphopantetheine As the Prosthetic Group of the Acyl Carrier Protein. *Proceedings of the National Academy of Sciences* **53**, 410–417 (1965).
- [55] Elovson, J. & Vagelos, P. R. Acyl carrier protein. X. Acyl carrier protein synthetase. *Journal of Biological Chemistry* **243**, 3603–3611 (1968).
- [56] Smith, S., Witkowski, A. & Joshi, A. K. Structural and functional organization of the animal fatty acid synthase. *Progress in Lipid Research* **42**, 289–317 (2003).
- [57] Crosby, J. & Crump, M. P. The structural role of the carrier protein - Active controller or passive carrier. *Natural Product Reports* **29**, 1111–1137 (2012).
- [58] Arnstadt, K., Lynen, F. & Schindlbeck, G. Zum Mechanismus der Kondensationsreaktion der Fettsäurebiosynthese. *European Journal of Biochemistry* **55**, 561–571 (1975).
- [59] Witkowski, A., Joshi, A. K. & Smith, S. Characterization of the interthiol acyltransferase reaction catalyzed by the β -ketoacyl synthase domain of the animal fatty acid synthase. *Biochemistry* **41**, 16338–16344 (1997).
- [60] Witkowski, A., Joshi, A. K. & Smith, S. Mechanism of the β -ketoacyl synthase reaction catalyzed by the animal fatty acid synthase. *Biochemistry* **41**, 10877–10887 (2002).
- [61] Price, A. C., Zhang, Y. M., Rock, C. O. & White, S. W. Structure of β -ketoacyl-[acyl carrier protein] reductase from *Escherichia coli*: Negative cooperativity and its structural basis. *Biochemistry* **40**, 12772–12781 (2001).
- [62] Price, A. C., Zhang, Y. M., Rock, C. O. & White, S. W. Cofactor-induced conformational rearrangements establish a catalytically competent active site and a proton relay conduit in FabG. *Structure* **12**, 417–428 (2004).
- [63] Medina, F. E., Neves, R. P., Ramos, M. J. & Fernandes, P. A. A QM/MM study of the reaction mechanism of human β -ketoacyl reductase. *Physical chemistry chemical physics* **19**, 347–355 (2016).

- [64] Leesong, M., Henderson, B. S., Gillig, J. R., Schwab, J. M. & Smith, J. L. Structure of a dehydratase-isomerase from the bacterial pathway for biosynthesis of unsaturated fatty acids: Two catalytic activities in one active site. *Structure* **4**, 253–264 (1996).
- [65] Zhang, L. *et al.* Crystal structure of FabZ-ACP complex reveals a dynamic seesaw-like catalytic mechanism of dehydratase in fatty acid biosynthesis. *Cell Research* **26**, 1330–1344 (2016).
- [66] Medina, F. E., Neves, R. P., Ramos, M. J. & Fernandes, P. A. QM/MM Study of the Reaction Mechanism of the Dehydratase Domain from Mammalian Fatty Acid Synthase. *ACS Catalysis* **8**, 10267–10278 (2018).
- [67] Heath, R. J. & Rock, C. O. Enoyl-acyl carrier protein reductase (*fabI*) plays a determinant role in completing cycles of fatty acid elongation in *Escherichia coli*. *Journal of Biological Chemistry* **270**, 26538–26542 (1995).
- [68] 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* **10**, 377–387 (1969).
- [69] Maier, T., Jenni, S. & Ban, N. Architecture of Mammalian Fatty Acid Synthase at 4.5 Å Resolution. *Science* **311**, 1258–1262 (2006).
- [70] Pollard, M. R., Anderson, L., Fan, C., Hawkins, D. J. & Davies, H. M. A specific acyl-ACP thioesterase implicated in medium-chain fatty acid production in immature cotyledons of *Umbellularia californica*. *Archives of Biochemistry and Biophysics* **284**, 306–312 (1991).
- [71] Lu, Y. J. *et al.* Acyl-Phosphates Initiate Membrane Phospholipid Synthesis in Gram-Positive Pathogens. *Molecular Cell* **23**, 765–772 (2006).
- [72] Van Vranken, J. G. *et al.* ACP Acylation Is an Acetyl-CoA-Dependent Modification Required for Electron Transport Chain Assembly. *Molecular Cell* **71**, 567–580 (2018).
- [73] McDaniel, R., Ebert-Khosla, S., Hopwood, D. A. & Khosla, C. Engineered biosynthesis of novel polyketides. *Science* **262**, 1546–1550 (1993).
- [74] Keatinge-Clay, A. T. The structures of type I polyketide synthases. *Natural Product Reports* **29**, 1050–1073 (2012).
- [75] Klaus, M. & Grninger, M. Engineering strategies for rational polyketide synthase design. *Natural Product Reports* **35**, 1070–1081 (2018).

References

- [76] Lowry, B., Li, X., Robbins, T., Cane, D. E. & Khosla, C. A turnstile mechanism for the controlled growth of biosynthetic intermediates on assembly line polyketide synthases. *ACS Central Science* **2**, 14–20 (2016).
- [77] Kroken, S., Glass, N. L., Taylor, J. W., Yoder, O. C. & Turgeon, B. G. Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 15670–15675 (2003).
- [78] Jenke-Kodama, H. & Dittmann, E. Evolution of metabolic diversity: Insights from microbial polyketide synthases. *Phytochemistry* **70**, 1858–1866 (2009).
- [79] Ridley, C. P., Ho, Y. L. & Khosla, C. Evolution of polyketide synthases in bacteria. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 4595–4600 (2008).
- [80] Piel, J., Hui, D., Fusetani, N. & Matsunaga, S. Targeting modular polyketide synthases with iteratively acting acyltransferases from metagenomes of uncultured bacterial consortia. *Environmental Microbiology* **6**, 921–927 (2004).
- [81] Lynen, F. Nobel Lecture – The pathway from "activated acetic acid" to the terpenes and fatty acids. <http://biotools.nubic.northwestern.edu/OligoCalc.html> (1964). NobelPrize.org. Nobel Media AB 2021. accessed 06.05.2021.
- [82] Jenni, S. *et al.* Structure of Fungal Fatty Acid Synthase and Implications for Iterative Substrate Shuttling. *Science* **316**, 254–261 (2007).
- [83] Johansson, P. *et al.* Multimeric Options for the Auto-Activation of the *Saccharomyces cerevisiae* FAS Type I Megasyntase. *Structure* **17**, 1063–1074 (2009).
- [84] Boehringer, D., Ban, N. & Leibundgut, M. 7.5-Å cryo-EM structure of the mycobacterial fatty acid synthase. *Journal of Molecular Biology* **425**, 841–849 (2013).
- [85] Maier, T., Leibundgut, M. & Ban, N. The Crystal Structure of a Mammalian Fatty Acid Synthase. *Science* **321**, 1315–1322 (2008).
- [86] Grininger, M. Perspectives on the evolution, assembly and conformational dynamics of fatty acid synthase type I (FAS I) systems. *Current Opinion in Structural Biology* **25**, 49–56 (2014).
- [87] Weissman, K. J. Uncovering the structures of modular polyketide synthases. *Natural Product Reports* **32**, 436–453 (2015).

- [88] Tsai, S. C. *et al.* Crystal structure of the macrocycle-forming thioesterase domain of the erythromycin polyketide synthase: Versatility from a unique substrate channel. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 14808–14813 (2001).
- [89] Weissman, K. J. The structural biology of biosynthetic megaenzymes. *Nature Chemical Biology* **11**, 660–670 (2015).
- [90] 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* **14**, 737–748 (2006).
- [91] Zheng, J., Fage, C. D., Demeler, B., Hoffman, D. W. & Keatinge-Clay, A. T. The missing linker: A dimerization motif located within polyketide synthase modules. *ACS Chemical Biology* **8**, 1263–1270 (2013).
- [92] Gogos, A., Mu, H. & Shapiro, L. Putative enoyl reductase domain of polyketide synthase. (PDB: 1PQW) (2003). URL <https://www.rcsb.org/structure/1pqw>.
- [93] Keatinge-Clay, A. Crystal Structure of the Erythromycin Polyketide Synthase Dehydratase. *Journal of Molecular Biology* **384**, 941–953 (2008).
- [94] Akey, D. L. *et al.* Crystal Structures of Dehydratase Domains from the Curacin Polyketide Biosynthetic Pathway. *Structure* **18**, 94–105 (2010).
- [95] Tsai, S. C., Lu, H., Cane, D. E., Khosla, C. & Stroud, R. M. Insights into channel architecture and substrate specificity from crystal structures of two macrocycle-forming thioesterases of modular polyketide synthases. *Biochemistry* **41**, 12598–12606 (2002).
- [96] Alekseyev, V. Y., Liu, C. W., Cane, D. E., Puglisi, J. D. & Khosla, C. Solution structure and proposed domain domain recognition interface of an acyl carrier protein domain from a modular polyketide synthase. *Protein Science* **16**, 2093–2107 (2007).
- [97] Tran, L., Broadhurst, R. W., Tosin, M., Cavalli, A. & Weissman, K. J. Insights into protein-protein and enzyme-substrate interactions in modular polyketide synthases. *Chemistry and Biology* **17**, 705–716 (2010).
- [98] Busche, A. *et al.* Characterization of molecular interactions between ACP and halogenase domains in the curacin A polyketide synthase. *ACS Chemical Biology* **7**, 378–386 (2012).
- [99] Tang, Y., Kim, C.-Y., Mathews, I. I., Cane, D. E. & Khosla, C. The 2.7-Å crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase. *PNAS* **103**, 11124–11129 (2006).

References

- [100] Tang, Y., Chen, A. Y., Kim, C. Y., Cane, D. E. & Khosla, C. Structural and Mechanistic Analysis of Protein Interactions in Module 3 of the 6-Deoxyerythronolide B Synthase. *Chemistry and Biology* **14**, 931–943 (2007).
- [101] Zheng, J., Gay, D. C., Demeler, B., White, M. A. & Keatinge-Clay, A. T. Divergence of multimodular polyketide synthases revealed by a didomain structure. *Nature Chemical Biology* **8**, 615–621 (2012).
- [102] Płoskoń, E. *et al.* A mammalian type I fatty acid synthase acyl carrier protein domain does not sequester acyl chains. *Journal of Biological Chemistry* **283**, 518–528 (2008).
- [103] Pappenberger, G. *et al.* Structure of the Human Fatty Acid Synthase KS-MAT Didomain as a Framework for Inhibitor Design. *Journal of Molecular Biology* **397**, 508–519 (2010).
- [104] Chakravarty, B., Gu, Z., Chirala, S. S., Wakil, S. J. & Quioco, F. A. Human fatty acid synthase: Structure and substrate selectivity of the thioesterase domain. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 15567–15572 (2004).
- [105] 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* **426**, 2229–2245 (2014).
- [106] Dutta, S. *et al.* Structure of a modular polyketide synthase. *Nature* **510**, 512–517 (2014).
- [107] Whicher, J. R. *et al.* Structural rearrangements of a polyketide synthase module during its catalytic cycle. *Nature* **510**, 560–564 (2014).
- [108] Rittner, A. & Grninger, M. Modular polyketide synthases (PKSs): A new model fits all? *ChemBioChem* **15**, 2489–2493 (2014).
- [109] D'Imprima, E. *et al.* Protein denaturation at the air-water interface and how to prevent it. *eLife* **8**, e42747 (2019).
- [110] Herbst, D. A., Jakob, R. P., Zähringer, F. & Maier, T. Mycocerosic acid synthase exemplifies the architecture of reducing polyketide synthases. *Nature* **531**, 533–537 (2016).
- [111] Gay, D., You, Y. O., Keatinge-Clay, A. & Cane, D. E. Structure and stereospecificity of the dehydratase domain from the terminal module of the rifamycin polyketide synthase. *Biochemistry* **52**, 8916–8928 (2013).

- [112] Fiers, W. D., Dodge, G. J., Sherman, D. H., Smith, J. L. & Aldrich, C. C. Vinylogous Dehydration by a Polyketide Dehydratase Domain in Curacin Biosynthesis. *Journal of the American Chemical Society* **138**, 16024–16036 (2016).
- [113] Faille, A. *et al.* Insights into Substrate Modification by Dehydratases from Type I Polyketide Synthases. *Journal of Molecular Biology* **429**, 1554–1569 (2017).
- [114] Jakob, R., Herbst, D., Muller, R. & Maier, T. Crystal Structures of Dehydratase Domains from *trans*-AT Polyketide Biosynthetic Pathway. (PDBs: 5J6O, 5IL5, 5IL6, 5HU7, 5HQW) (2017). URL <https://www.rcsb.org/structure/5J6O>.
- [115] Keatinge-Clay, A. T. The Structural Relationship between Iterative and Modular PKSs. *Cell Chemical Biology* **23**, 540–542 (2016).
- [116] Ames, B. D. *et al.* Crystal structure and biochemical studies of the *trans*-acting polyketide enoyl reductase LovC from lovastatin biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 11144–11149 (2012).
- [117] Khare, D. *et al.* Structural Basis for Cyclopropanation by a Unique Enoyl-Acyl Carrier Protein Reductase. *Structure* **23**, 2213–2223 (2015).
- [118] Wang, J. *et al.* Structural basis for the biosynthesis of lovastatin. *Nature Communications* **12**: **867** (2021).
- [119] Richter, C. D., Nietlispach, D., Broadhurst, R. W. & Weissman, K. J. Multienzyme docking in hybrid megasynthetases. *Nature Chemical Biology* **4**, 75–81 (2008).
- [120] Buchholz, T. J. *et al.* Structural basis for binding specificity between subclasses of modular polyketide synthase docking domains. *ACS Chemical Biology* **4**, 41–52 (2009).
- [121] Whicher, J. R. *et al.* Cyanobacterial polyketide synthase docking domains: A tool for engineering natural product biosynthesis. *Chemistry and Biology* **20**, 1340–1351 (2013).
- [122] Reed, M. A. *et al.* The type I rat fatty acid synthase ACP shows structural homology and analogous biochemical properties to type II ACPs. *Organic and Biomolecular Chemistry* **1**, 463–471 (2003).
- [123] Andrec, M., Blake Hill, R. & Prestegard, J. H. Amide exchange rates in *Escherichia coli* acyl carrier protein: Correlation with protein structure and dynamics. *Protein Science* **4**, 983–993 (1995).
- [124] Weissman, K. J., Hong, H., Popovic, B. & Meersman, F. Evidence for a Protein-Protein Interaction Motif on an Acyl Carrier Protein Domain from a Modular Polyketide Synthase. *Chemistry and Biology* **13**, 625–636 (2006).

- [125] Miyanaga, A., Iwasawa, S., Shinohara, Y., Kudo, F. & Eguchi, T. Structure-based analysis of the molecular interactions between acyltransferase and acyl carrier protein in vicenistatin biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 1802–1807 (2016).
- [126] Miyanaga, A. *et al.* Structural Basis of Protein–Protein Interactions between a *trans*-Acting Acyltransferase and Acyl Carrier Protein in Polyketide Disorazole Biosynthesis. *Journal of the American Chemical Society* **140**, 7970–7978 (2018).
- [127] Cleland, W. W. The Kinetics of Enzyme-Catalyzed Reactions with Two or More Substrates or Products. I. Nomenclature and Rate Equations. *Biochimica et Biophysica Acta* **67**, 104–137 (1963).
- [128] Cleland, W. W. The Kinetics of Enzyme-Catalyzed Reactions with Two or More Substrates or Products. III. Prediction of Initial Velocity and Inhibition Patterns by Inspection. *Biochimica et Biophysica Acta* **67**, 188–196 (1963).
- [129] Copeland, R. A. *Enzymes - A practical Introduction to Structure, Mechanism, and Data Analysis* (Wiley-VCH, New York/ Chichester/ Weinheim/ Brisbane/ Singapore/ Toronto, 2000), second edn.
- [130] Bunkoczi, G. *et al.* Structural Basis for Different Specificities of Acyltransferases Associated with the Human Cytosolic and Mitochondrial Fatty Acid Synthases. *Chemistry and Biology* **16**, 667–675 (2009).
- [131] Cognet, J. A. & Hammes, G. G. Elementary Steps in the Reaction Mechanism of Chicken Liver Fatty Acid Synthase: Acetylation-Deacetylation. *Biochemistry* **22**, 3002–3007 (1983).
- [132] Bonnett, S. A. *et al.* Acyl-CoA subunit selectivity in the pikromycin polyketide synthase PikAIV: Steady-state kinetics and active-site occupancy analysis by FTICR-MS. *Chemistry and Biology* **18**, 1075–1081 (2011).
- [133] Dunn, B. J., Cane, D. E. & Khosla, C. Mechanism and Specificity of an Acyltransferase Domain from a Modular Polyketide Synthase. *Biochemistry* **52**, 1839–1841 (2013).
- [134] Paiva, P., Sousa, S. F., Ramos, M. J. & Fernandes, P. A. Understanding the Catalytic Machinery and the Reaction Pathway of the Malonyl-Acetyl Transferase Domain of Human Fatty Acid Synthase. *ACS Catalysis* **8**, 4860–4872 (2018).
- [135] Paiva, P., Sousa, S. F., Fernandes, P. A. & João Ramos, M. Human Fatty Acid Synthase: A Computational Study of the Transfer of the Acyl Moieties from MAT to the ACP Domain. *ChemCatChem* **11**, 3853–3864 (2019).

- [136] Serre, L., Verbree, E. C., Dauter, Z., Stuitje, A. R. & Derewenda, Z. S. The *Escherichia coli* malonyl-CoA:acyl carrier protein transacylase at 1.5-Å resolution: Crystal structure of a fatty acid synthase component. *Journal of Biological Chemistry* **270**, 12961–12964 (1995).
- [137] Keatinge-Clay, A. T. *et al.* Catalysis, specificity, and ACP docking site of *Streptomyces coelicolor* malonyl-CoA:ACP Transacylase. *Structure* **11**, 147–154 (2003).
- [138] Smith, S. & Stern, A. The Effect of Aromatic CoA Esters on Fatty Acid Synthetase: Biosynthesis of ω -Phenyl Fatty Acids. *Archives of Biochemistry and Biophysics* **222**, 259–265 (1983).
- [139] Liou, G. F., Lau, J., Cane, D. E. & Khosla, C. Quantitative analysis of loading and extender acyltransferases of modular polyketide synthases. *Biochemistry* **42**, 200–207 (2003).
- [140] Wang, F. *et al.* Structural and functional analysis of the loading acyltransferase from avermectin modular polyketide synthase. *ACS Chemical Biology* **10**, 1017–1025 (2015).
- [141] Koryakina, I., McArthur, J. B., Draelos, M. M. & Williams, G. J. Promiscuity of a modular polyketide synthase towards natural and non-natural extender units. *Organic & Biomolecular Chemistry* **11**, 4449–4458 (2013).
- [142] Dunn, B. J., Watts, K. R., Robbins, T., Cane, D. E. & Khosla, C. Comparative Analysis of the Substrate Specificity of *trans*- versus *cis*-Acyltransferases of Assembly Line Polyketide Synthases. *Biochemistry* **53**, 3796–3806 (2014).
- [143] Geyer, K., Sundaram, S., Sušnik, P., Koert, U. & Erb, T. J. Understanding Substrate Selectivity of Phoslactomycin Polyketide Synthase by Using Reconstituted in Vitro Systems. *ChemBioChem* **21**, 2080 – 2085 (2020).
- [144] Haydock, S. F. *et al.* Divergent sequence motifs correlated with the substrate specificity of (methyl)malonyl-CoA:acyl carrier protein transacylase domains in modular polyketide synthases. *FEBS Letters* **374**, 246–248 (1995).
- [145] Yadav, G., Gokhale, R. S. & Mohanty, D. Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases. *Journal of Molecular Biology* **328**, 335–363 (2003).
- [146] Nardini, M. & Dijkstra, B. W. α/β hydrolase fold enzymes: The family keeps growing. *Current Opinion in Structural Biology* **9**, 732–737 (1999).

References

- [147] Reeves, C. D. *et al.* Alteration of the substrate specificity of a modular polyketide synthase acyltransferase domain through site-specific mutations. *Biochemistry* **40**, 15464–15470 (2001).
- [148] Del Vecchio, F. *et al.* Active-site residue, domain and module swaps in modular polyketide synthases. *Journal of Industrial Microbiology and Biotechnology* **30**, 489–494 (2003).
- [149] Lau, J., Fu, H., Cane, D. E. & Khosla, C. Dissecting the role of acyltransferase domains of modular polyketide synthases in the choice and stereochemical fate of extender units. *Biochemistry* **38**, 1643–1651 (1999).
- [150] Heath, R. J. & Rock, C. O. The Claisen condensation in biology. *Natural Product Reports* **19**, 581–596 (2002).
- [151] Huang, W. *et al.* Crystal structure of β -ketoacyl-acyl carrier protein synthase II from *E.coli* reveals the molecular architecture of condensing enzymes. *EMBO Journal* **17**, 1183–1191 (1998).
- [152] Von Wettstein-Knowles, P., Olsen, J. G., McGuire, K. A. & Henriksen, A. Fatty acid synthesis: Role of active site histidines and lysine in Cys-His-His-type β -ketoacyl-acyl carrier protein synthases. *FEBS Journal* **273**, 695–710 (2006).
- [153] Blaquiere, N., Shore, D. G., Rousseaux, S. & Fagnou, K. Decarboxylative ketone aldol reactions: Development and mechanistic evaluation under metal-free conditions. *Journal of Organic Chemistry* **74**, 6190–6198 (2009).
- [154] Cui, W., Liang, Y., Tian, W., Ji, M. & Ma, X. Regulating effect of β -ketoacyl synthase domain of fatty acid synthase on fatty acyl chain length in de novo fatty acid synthesis. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids* **1861**, 149–155 (2016).
- [155] Yalpani, M., Willecke, K. & Lynen, F. Triacetic Acid Lactone, a Derailment Product of Fatty Acid Biosynthesis. *European Journal of Biochemistry* **8**, 495–502 (1969).
- [156] Seyama, Y., Otsuka, H., Kawaguchi, A. & Yamakawa, T. Fatty acid synthetase from the harderian gland of Guinea pig: Biosynthesis of methyl-branched fatty acids. *Journal of Biochemistry* **90**, 789–798 (1981).
- [157] Dewulf, J. P. *et al.* The synthesis of branched-chain fatty acids is limited by enzymatic decarboxylation of ethyl- and methylmalonyl-CoA. *Biochemical Journal* **476**, 2427–2447 (2019).

- [158] Khosla, C., Gokhale, R. S., Jacobsen, J. R. & Cane, D. E. Tolerance and Specificity of Polyketide Synthases. *Annual Review of Biochemistry* **68**, 219–253 (1999).
- [159] Viegas, M. F., Neves, R. P., Ramos, M. J. & Fernandes, P. A. Modeling of Human Fatty Acid Synthase and *in Silico* Docking of Acyl Carrier Protein Domain and Its Partner Catalytic Domains. *Journal of Physical Chemistry B* **122**, 77–85 (2018).
- [160] Milligan, J. C. *et al.* Molecular basis for interactions between an acyl carrier protein and a ketosynthase. *Nature Chemical Biology* **15**, 669–671 (2019).
- [161] Mindrebo, J. T. *et al.* Gating mechanism of elongating β -ketoacyl-ACP synthases. *Nature Communications* **11**: 1727 (2020).
- [162] Kapur, S., Chen, A. Y., Cane, D. E. & Khosla, C. Molecular recognition between ketosynthase and acyl carrier protein domains of the 6-deoxyerythronolide B synthase. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 22066–22071 (2010).
- [163] Klaus, M., D'Souza, A. D., Nivina, A., Khosla, C. & Grininger, M. Engineering of Chimeric Polyketide Synthases Using SYNZIP Docking Domains. *ACS Chemical Biology* **14**, 426–433 (2019).
- [164] Peng, H., Ishida, K., Sugimoto, Y., Jenke-Kodama, H. & Hertweck, C. Emulating evolutionary processes to morph aureothin-type modular polyketide synthases and associated oxygenases. *Nature Communications* **10**: 3918 (2019).
- [165] Zhang, L. *et al.* Characterization of Giant Modular PKSs Provides Insight into Genetic Mechanism for Structural Diversification of Aminopolyol Polyketides. *Angewandte Chemie - International Edition* **56**, 1740–1745 (2017).
- [166] Vander Wood, D. A. & Keatinge-Clay, A. T. The modules of *trans*-acyltransferase assembly lines redefined with a central acyl carrier protein. *Proteins: Structure, Function and Bioinformatics* **86**, 664–675 (2018).
- [167] Keatinge-Clay, A. T. Polyketide Synthase Modules Redefined. *Angewandte Chemie - International Edition* **56**, 4658–4660 (2017).
- [168] Weissman, K. J. Genetic engineering of modular PKSs: from combinatorial biosynthesis to synthetic biology. *Natural Product Reports* **33**, 203–230 (2016).
- [169] Woodward, R. B. *et al.* Asymmetric Total Synthesis of Erythromycin. 1. Synthesis of an Erythronolide A Seco Acid Derivative via Asymmetric Induction. *Journal of the American Chemical Society* **103**, 3210–3213 (1981).

References

- [170] Woodward, R. B. *et al.* Asymmetric Total Synthesis of Erythromycin. 3. Total Synthesis of Erythromycin. *Journal of the American Chemical Society* **103**, 3215–3217 (1981).
- [171] Woodward, R. B. *et al.* Asymmetric Total Synthesis of Erythromycin. 2. Synthesis of an Erythronolide A Lactone System. *Journal of the American Chemical Society* **103**, 3213–3215 (1981).
- [172] Malpartida, D. & Hopwood, D. A. Molecular cloning of the whole biosynthetic pathways of a *Streptomyces* antibiotic and its expression in a heterologous host. *Nature* **309**, 462–464 (1984).
- [173] Bartel, P. L. *et al.* Biosynthesis of anthraquinones by interspecies cloning of actinorhodin biosynthesis genes in streptomycetes: Clarification of actinorhodin gene functions. *Journal of Bacteriology* **172**, 4816–4826 (1990).
- [174] McDaniel, R., Kao, C. M., Hwang, S. J. & Khosla, C. Engineered intermodular and intramodular polyketide synthase fusions. *Chemistry and Biology* **4**, 667–674 (1997).
- [175] Donadio, S., Staver, M. J., McAlpine, J. B., Swanson, S. J. & Katz, L. Modular organization of genes required for complex polyketide biosynthesis. *Science* **252**, 675–679 (1991).
- [176] Carreras, C. W. & Santi, D. V. Engineering of modular polyketide synthases to produce novel polyketides. *Current Opinion in Biotechnology* **9**, 403–411 (1998).
- [177] Verdine, G. L. The combinatorial chemistry of nature. *Nature* **384**, 11–13 (1996).
- [178] Menzella, H. G. *et al.* Combinatorial polyketide biosynthesis by de novo design and rearrangement of modular polyketide synthase genes. *Nature Biotechnology* **23**, 1171–1176 (2005).
- [179] Kao, C. M., Khosla, C., Luo, G., Cane, D. E. & Katz, L. Engineered biosynthesis of a triketide lactone from an incomplete modular polyketide synthase. *Journal of the American Chemical Society* **116**, 11612–11613 (1994).
- [180] Cortes, J. *et al.* Repositioning of a domain in a modular polyketide synthase to promote specific chain cleavage. *Science* **268**, 1487–1489 (1995).
- [181] Donadio, S., McAlpine, J. B., Sheldon, P. J., Jackson, M. & Katz, L. An erythromycin analog produced by reprogramming of polyketide synthesis. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 7119–7123 (1993).
- [182] Jacobsen, J. R., Hutchinson, C. R., Cane, D. E. & Khosla, C. Precursor-directed biosynthesis of erythromycin analogs by an engineered polyketide synthase. *Science* **277**, 367–369 (1997).

- [183] Oliynyk, M., Brown, M. J., Cortes, J., Staunton, J. & Leadlay, P. F. A hybrid modular polyketide synthase obtained by domain swapping. *Chemistry and Biology* **3**, 833–839 (1996).
- [184] Kuhstoss, S., Huber, M., Turner, J. R., Paschal, J. W. & Rao, R. N. Production of a novel polyketide through the construction of a hybrid polyketide synthase. *Gene* **183**, 231–236 (1996).
- [185] Liu, L., Thamchaipenet, A., Fu, H., Betlach, M. & Ashley, G. Biosynthesis of 2-nor-6-deoxyerythronolide B by rationally designed domain substitution. *Journal of the American Chemical Society* **119**, 10553–10554 (1997).
- [186] McDaniel, R. *et al.* Gain-of-Function Mutagenesis of a Modular Polyketide Synthase. *Journal of the American Chemical Society* **119**, 4309–4310 (1997).
- [187] Bravo-Rodriguez, K. *et al.* Substrate Flexibility of a Mutated Acyltransferase Domain and Implications for Polyketide Biosynthesis. *Chemistry and Biology* **22**, 1425–1430 (2015).
- [188] Murphy, A. C., Hong, H., Vance, S., Broadhurst, R. W. & Leadlay, P. F. Broadening substrate specificity of a chain-extending ketosynthase through a single active-site mutation. *Chemical Communications* **52**, 8373–8376 (2016).
- [189] Xue, Q., Ashley, G., Hutchinson, C. R. & Santi, D. V. A multiplasmid approach to preparing large libraries of polyketides. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 11740–11745 (1999).
- [190] Kim, B. S., Sherman, D. H. & Reynolds, K. A. An efficient method for creation and functional analysis of libraries of hybrid type I polyketide synthases. *Protein Engineering, Design and Selection* **17**, 277–284 (2004).
- [191] Long, P. F. *et al.* Engineering specificity of starter unit selection by the erythromycin-producing polyketide synthase. *Molecular Microbiology* **43**, 1215–1225 (2002).
- [192] Rowe, C. J. *et al.* Engineering a polyketide with a longer chain by insertion of an extra module into the erythromycin-producing polyketide synthase. *Chemistry and Biology* **8**, 475–485 (2001).
- [193] Martin, C. J. *et al.* Heterologous expression in *Saccharopolyspora erythraea* of a pentaketide synthase derived from the spinosyn polyketide synthase. *Organic and Biomolecular Chemistry* **1**, 4144–4147 (2003).

- [194] Gaisser, S. *et al.* Direct production of ivermectin-like drugs after domain exchange in the avermectin polyketide synthase of *Streptomyces avermitilis* ATCC31272. *Organic and Biomolecular Chemistry* **1**, 2840–2847 (2003).
- [195] Reeves, C. D. *et al.* Production of Hybrid 16-Membered Macrolides by Expressing Combinations of Polyketide Synthase Genes in Engineered *Streptomyces fradiae* Hosts. *Chemistry and Biology* **11**, 1465–1472 (2004).
- [196] Barajas, J. F., Blake-Hedges, J. M., Bailey, C. B., Curran, S. & Keasling, J. D. Engineered polyketides: Synergy between protein and host level engineering. *Synthetic and Systems Biotechnology* **2**, 147–166 (2017).
- [197] Mcdaniel, R. *et al.* Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel "unnatural" natural products. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 1846–1851 (1999).
- [198] Kushnir, S. *et al.* Minimally invasive mutagenesis gives rise to a biosynthetic polyketide library. *Angewandte Chemie - International Edition* **51**, 10664–10669 (2012).
- [199] Klaus, M. *et al.* Protein-protein interactions, not substrate recognition, dominate the turnover of chimeric assembly line polyketide synthases. *Journal of Biological Chemistry* **291**, 16404–16415 (2016).
- [200] Holzbaur, I. E. *et al.* Molecular basis of Celmer's rules: Role of the ketosynthase domain in epimerisation and demonstration that ketoreductase domains can have altered product specificity with unnatural substrates. *Chemistry and Biology* **8**, 329–340 (2001).
- [201] Kellenberger, L. *et al.* A polylinker approach to reductive loop swaps in modular polyketide synthases. *Chembiochem : a European journal of chemical biology* **9**, 2740–2749 (2008).
- [202] Sugimoto, Y., Ding, L., Ishida, K. & Hertweck, C. Rational design of modular polyketide synthases: Morphing the aureothin pathway into a luteoreticuliculin assembly line. *Angewandte Chemie - International Edition* **53**, 1560–1564 (2014).
- [203] Awakawa, T. *et al.* Reprogramming of the antimycin NRPS-PKS assembly lines inspired by gene evolution. *Nature Communications* **9**: 3534 (2018).
- [204] Chemler, J. A. *et al.* Evolution of Efficient Modular Polyketide Synthases by Homologous Recombination. *Journal of the American Chemical Society* **137**, 10603–10609 (2015).
- [205] Wlodek, A. *et al.* Diversity oriented biosynthesis via accelerated evolution of modular gene clusters. *Nature Communications* **8**: 1206 (2017).

- [206] Miyazawa, T., Hirsch, M., Zhang, Z. & Keatinge-clay, A. T. An in vitro platform for engineering and harnessing modular polyketide synthases. *Nature Communications* **11**: 80 (2020).
- [207] Reinke, A. W., Grant, R. A. & Keating, A. E. A synthetic coiled-coil interactome provides heterospecific modules for molecular engineering. *Journal of the American Chemical Society* **132**, 6025–6031 (2010).
- [208] Thompson, K. E., Bashor, C. J., Lim, W. A. & Keating, A. E. Synzip protein interaction toolbox: In vitro and in vivo specifications of heterospecific coiled-coil interaction domains. *ACS Synthetic Biology* **1**, 118–129 (2012).
- [209] Klaus, M., Buyachuihan, L. & Grninger, M. Ketosynthase Domain Constrains the Design of Polyketide Synthases. *ACS Chemical Biology* **15**, 2422–2432 (2020).
- [210] Stegemann, F. *Fluorescence Studies on Dynamic Domains of the Mammalian Fatty Acid Synthase and Different Polyketide Synthases*. Master's thesis, Goethe-University Frankfurt (2016).
- [211] Molnos, J., Gardiner, R., Dale, G. E. & Lange, R. A continuous coupled enzyme assay for bacterial malonyl-CoA:acyl carrier protein transacylase (FabD). *Analytical Biochemistry* **319**, 171–176 (2003).
- [212] Rittner, A., Paithankar, K. S., Himmler, A. & Grninger, M. Type I fatty acid synthase trapped in the octanoyl-bound state. *Protein Science* **29**, 589–605 (2020).
- [213] Dimroth, P., Rlingelmann, E. & Lynen, F. 6-Methylsalicylic Acid Synthetase from *Penicillium patulum* Some Catalytic Properties of the Enzyme and Its Relation to Fatty Acid Synthetase. *European Journal of Biochemistry* **68**, 591–596 (1976).
- [214] Oliynyk, M. *et al.* Complete genome sequence of the erythromycin-producing bacterium *Saccharopolyspora erythraea* NRRL23338. *Nature Biotechnology* **25**, 447–453 (2007).
- [215] Moriguchi, T., Kezuka, Y., Nonaka, T., Ebizuka, Y. & Fujii, I. Hidden function of catalytic domain in 6-methylsalicylic acid synthase for product release. *Journal of Biological Chemistry* **285**, 15637–15643 (2010).
- [216] Marcella, A. M. & Barb, A. W. The R117A variant of the *Escherichia coli* transacylase FabD synthesizes novel acyl-(acyl carrier proteins). *Applied Microbiology and Biotechnology* **101**, 8431–8441 (2017).
- [217] Liew, C. W. *et al.* Crystal structure of the acyltransferase domain of the iterative polyketide synthase in enediyne biosynthesis. *Journal of Biological Chemistry* **287**, 23203–23215 (2012).

References

- [218] Wang, Y.-Y. *et al.* Biochemical Characterization of a Malonyl-Specific Acyltransferase Domain of FK506 Biosynthetic Polyketide Synthase. *Protein & Peptide Letters* **22**, 2–7 (2015).
- [219] Takamura, Y. & Nomura, G. Changes in the intracellular concentration of acetyl-CoA and malonyl-CoA in relation to the carbon and energy metabolism of *Escherichia coli* K12. *Journal of general microbiology* **134**, 2249–2253 (1988).
- [220] Davis, M. S., Solbiati, J. & Cronan, J. E. Overproduction of acetyl-CoA carboxylase activity increases the rate of fatty acid biosynthesis in *Escherichia coli*. *Journal of Biological Chemistry* **275**, 28593–28598 (2000).
- [221] Rossini, E., Gajewski, J., Klaus, M., Hummer, G. & Grninger, M. Analysis and engineering of substrate shuttling by the acyl carrier protein (ACP) in fatty acid synthases (FASs). *Chemical Communications* **54**, 11606–11609 (2018).
- [222] Heil, C. S., Rittner, A., Goebel, B., Beyer, D. & Grninger, M. Site-Specific Labelling of Multidomain Proteins by Amber Codon Suppression. *Scientific Reports* **8**: **14864** (2018).
- [223] Berlinghof, M. *Analyse von Domänen-Domänen-Interaktionen in Megasyntasen*. Master's thesis, Goethe-University Frankfurt (2018).
- [224] Chari, A. *et al.* ProteoPlex: Stability optimization of macromolecular complexes by sparse-matrix screening of chemical space. *Nature Methods* **12**, 859–865 (2015).
- [225] Mayer, S., Rüdiger, S., Ang, H. C., Joerger, A. C. & Fersht, A. R. Correlation of Levels of Folded Recombinant p53 in *Escherichia coli* with Thermodynamic Stability *in Vitro*. *Journal of Molecular Biology* **372**, 268–276 (2007).
- [226] Just, A. *Analyse von Fettsäure- und Polyketidsynthase-Hybriden in Struktur und Funktion*. Master's thesis, Goethe-University Frankfurt (2020).
- [227] Szafranska, A. E., Hitchman, T. S., Cox, R. J., Crosby, J. & Simpson, T. J. Kinetic and mechanistic analysis of the malonyl CoA:ACP transacylase from *Streptomyces coelicolor* indicates a single catalytically competent serine nucleophile at the active site. *Biochemistry* **41**, 1421–1427 (2002).
- [228] Lowe, P. N. & Rhodes, S. Purification and characterization of [acyl-carrier-protein] acetyltransferase from *Escherichia coli*. *The Biochemical journal* **250**, 789–796 (1988).
- [229] Bar-Even, A. *et al.* The moderately efficient enzyme: Evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry* **50**, 4402–4410 (2011).

- [230] Rangan, V. S. & Smith, S. Expression in *Escherichia coli* and refolding of the malonyl-/acetyltransferase domain of the multifunctional animal fatty acid synthase. *Journal of Biological Chemistry* **271**, 31749–31755 (1996).
- [231] Albery, W. J. & Knowles, J. R. Evolution of Enzyme Function and the Development of Catalytic Efficiency. *Biochemistry* **15**, 5631–5640 (1976).
- [232] Lowry, B. *et al.* *In vitro* reconstitution and analysis of the 6-deoxyerythronolide B synthase. *Journal of the American Chemical Society* **135**, 16809–16812 (2013).
- [233] Grininger, M. The role of the iterative modules in polyketide synthase evolution. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 8680–8682 (2020).
- [234] Musiol-Kroll, E. M. & Wohlleben, W. Acyltransferases as tools for polyketide synthase engineering. *Antibiotics* **7**: **62** (2018).
- [235] Gajewski, J. *et al.* Engineering fatty acid synthases for directed polyketide production. *Nature Chemical Biology* **13**, 363–365 (2017).
- [236] Kumar, S. & Nussinov, R. Close-range electrostatic interactions in proteins. *Chem-BioChem* **3**, 604–617 (2002).
- [237] Zhou, H. X. & Pang, X. Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation. *Chemical Reviews* **118**, 1691–1741 (2018).
- [238] Khersonsky, O. & Tawfik, D. S. Enzyme promiscuity: A mechanistic and evolutionary perspective. *Annual Review of Biochemistry* **79**, 471–505 (2010).
- [239] Kim, C. Y. *et al.* Reconstituting modular activity from separated domains of 6-deoxyerythronolide B synthase. *Biochemistry* **43**, 13892–13898 (2004).
- [240] Torres, J. P., Lin, Z., Winter, J. M., Krug, P. J. & Schmidt, E. W. Animal biosynthesis of complex polyketides in a photosynthetic partnership. *Nature Communications* **11**: **2882** (2020).
- [241] Richardson, M. T., Pohl, N. L., Kealey, J. T. & Khosla, C. Tolerance and specificity of recombinant 6-methylsalicylic acid synthase. *Metabolic Engineering* **1**, 180–187 (1999).
- [242] Jensen, R. A. Enzyme recruitment in evolution of new function. *Annual review of microbiology* **30**, 409–425 (1976).
- [243] Kalkreuter, E., Crowetipton, J. M., Lowell, A. N., Sherman, D. H. & Williams, G. J. Engineering the Substrate Specificity of a Modular Polyketide Synthase for Installation

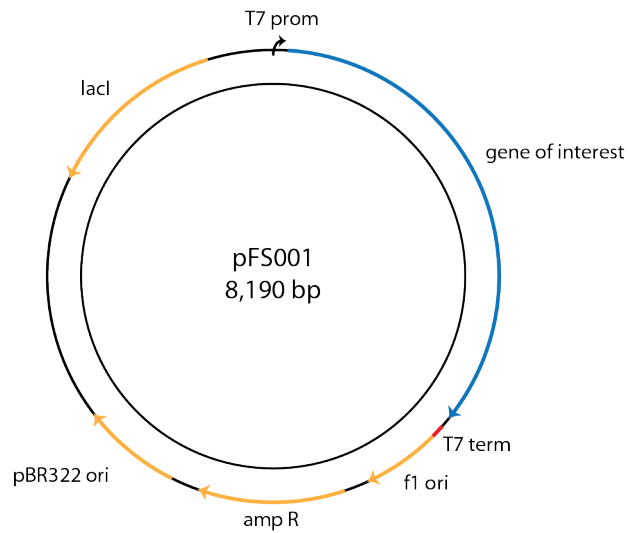
References

- of Consecutive Non-Natural Extender Units. *Journal of the American Chemical Society* **141**, 1961–1969 (2019).
- [244] Kibbe, W. A. Oligocalc calculator version 3.27. <http://biotools.nubic.northwestern.edu/OligoCalc.html>.
- [245] Xie, D. *et al.* Microbial synthesis of triacetic acid lactone. *Biotechnology and Bioengineering* **93**, 727–736 (2006).
- [246] Waterhouse, A. *et al.* SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research* **46**, W296–W303 (2018).
- [247] Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology* **273**, 774–797 (2007).
- [248] Xue, L. C., Rodrigues, J. P., Kastiris, P. L., Bonvin, A. M. & Vangone, A. PRODIGY: A web server for predicting the binding affinity of protein-protein complexes. *Bioinformatics* **32**, 3676–3678 (2016).
- [249] Barlow, K. A. *et al.* Flex ddG: Rosetta Ensemble-Based Estimation of Changes in Protein-Protein Binding Affinity upon Mutation. *Journal of Physical Chemistry B* **122**, 5389–5399 (2018).
- [250] Kortemme, T., Kim, D. E. & Baker, D. Computational alanine scanning of protein-protein interfaces. *Science Signaling* **2004**, pl2–pl2 (2004).
- [251] Dittrich, J., Schmidt, D., Pflieger, C. & Gohlke, H. Converging a Knowledge-Based Scoring Function: DrugScore 2018. *Journal of Chemical Information and Modeling* **59**, 509–521 (2019).
- [252] Pires, D. E., Ascher, D. B. & Blundell, T. L. MCSM: Predicting the effects of mutations in proteins using graph-based signatures. *Bioinformatics* **30**, 335–342 (2014).
- [253] Moreira, I. S. *et al.* SpotOn: High Accuracy Identification of Protein-Protein Interface Hot-Spots. *Scientific Reports* **7**: 8007 (2017).
- [254] Valenzano, C. R. *et al.* Stereospecificity of the dehydratase domain of the erythromycin polyketide synthase. *Journal of the American Chemical Society* **132**, 14697–14699 (2010).
- [255] Peter, D. M. *et al.* Screening and Engineering the Synthetic Potential of Carboxylating Reductases from Central Metabolism and Polyketide Biosynthesis. *Angewandte Chemie - International Edition* **54**, 13457–13461 (2015).

- [256] Fahim, A. *Modulation einer AT:ACP-Interaktionsfläche der 6-Desoxyerythronolid B-Synthase*. Bachelor's thesis, Goethe-University Frankfurt (2019).

Supplementary Information

Exemplary Plasmid Map



Plasmid map of pFS001 representing all pET22b vector constructs. Vectors contain an f1 and a pBR322 origin, an ampicillin resistance, the gene of interest encoding the respective protein, and the gene *lacI*, which controls protein expression via the T7 polymerase system.

Excerpt of Plasmids

Plasmids pAR and pMJD were kindly provided by Dr. Alexander Rittner and Mirko Joppe. The gene encoding for EcPKS1 was kindly provided by Eric Schmidt Lab (University of Utah). The gene encoding for Npt from *Streptomyces platensis* was kindly provided by Kyra Geyer, who is part of the Tobias Erb Lab (MPI Marburg).

List of plasmids for wildtype AT characterization.

#	Name
–	Sfp_Bsub_pETcoco
pAR090	Strepl_Ecoli(FabD)_H8_pET22b
pAR225	Strepl_Pks7(KS(C161G)-AT)_H8_pET22b
pAR318	Strepl_AVES1(ACP)0_H8_pET22b
pAR319	Strepl_AVES1(AT)0_H8_pET22b
pAR331	Strepl_DEBS3(ACP)5_H8_pET22b
pAR332	Strepl_PikAIII(ACP)5_H8_pET22b
pAR333	Strepl_RAPS3(ACP)14_H8_pET22b
pAR357	Sfp_Bsub_pCDF-1b
pFS001	Strepl_DEBS3(KS(C233G)-AT)5_H8_pET22b
pFS004	Strepl_PikAIII(KS(C243G)-AT)5_H8_pET22b
pFS007	Strepl_RAPS(KS(C226G)-AT)14_H8_pET22b
pFS011	Ecoli(ACP)_H8_pET22b
pFS025	Strepl_DEBS1(ACP)0_H8_pET22b
pFS026	Strepl_Pks7(ACP)_H8_pET22b
pFS030	Strepl_DEBS1(AT)0_MBP_H8_pET22b
pFS120	Strepl_DEBS3(ACP)5_COP_H8_pET22b
pFS121	Strepl_PikAIII(ACP)5_COP_H8_pET22b
pFS122	Strepl_RAPS3(ACP)14_COP_H8_pET22b
pFS123	Strepl_Pks7(ACP)_COP_H8_pET22b
pFS124	Strepl_Ecoli(ACP)_COP_H8_pET22b
pFS125	Strepl_DEBS1(ACP)0_COP_H8_pET22b
pFS126	Strepl_AVES1(ACP)0_COP_H8_pET22b
pFS129	Strepl_DEBS3(KS(C233G)-AT(A539S))5_H8_pET22b
pFS130	Strepl_DEBS3(KS(C233G)-AT(A539D))5_H8_pET22b
pFS131	Strepl_DEBS3(KS(C233G)-AT(A539E))5_H8_pET22b
pFS132	Strepl_DEBS3(KS(C233G)-AT(A539F))5_H8_pET22b
pFS133	Strepl_DEBS3(KS(C233G)-AT(R850K))5_H8_pET22b
pFS134	Strepl_DEBS3(KS(C233G)-AT(R850A))5_H8_pET22b
pFS135	Strepl_DEBS3(KS(C233G)-AT(R850E))5_H8_pET22b
pFS136	Strepl_DEBS3(KS(C233G)-AT(R850F))5_H8_pET22b
pFS137	Strepl_DEBS3(KS(C233G)-AT(R850F))5_H8_pET22b

List of plasmids for construction of FAS chimeras.

#	Name
pAR069	Strepl_m(KS-LD-AT)_H8_pET22b
pAR265	Strepl_DEBS3(KS-LD-AT)5_H8_pET22b
pAR352B	Strepl_m(ACP)_H8_RBS_Sfp_Bsub_pET22b
pFS031	Strepl_m(KS-LD)-DEBS3(AT)5-m(DH-ER-KR-ACP)_H8_pET22b
pFS032	Strepl_m(KS)-DEBS3(LD-AT)5-m(DH-ER-KR-ACP)_H8_pET22b
pFS033	Strepl_m(KS)-DEBS3(LD-AT)5-m(DH-ER-KR-ACP)_H8_pET22b
pFS034	Strepl_m(KS-LD)-RAPS3(AT)14-m(DH-ER-KR-ACP)_H8_pET22b
pFS035	Strepl_m(KS)-RAPS3(LD-AT)14-m(DH-ER-KR-ACP)_H8_pET22b
pFS036	Strepl_m(KS-LD)-RAPS3(LD-AT)14-m(DH-ER-KR-ACP)_H8_pET22b
pFS037	Strepl_m(KS-LD)-DEBS3(AT)5_H8_pET22b
pFS038	Strepl_m(KS)-DEBS3(LD-AT)5_H8_pET22b
pFS039	Strepl_m(KS(D61A)-LD-AT)_H8_pET22b
pFS040	Strepl_m(KS(N96A)-LD-AT)_H8_pET22b
pFS041	Strepl_m(KS-LD(R429A)-AT)_H8_pET22b
pFS042	Strepl_m(KS-LD(I457A)-AT)_H8_pET22b
pFS043	Strepl_m(KS-LD(R468A)-AT)_H8_pET22b
pFS044	Strepl_m(KS-LD(Y470A)-AT)_H8_pET22b
pFS045	Strepl_DEBS3(KS(E129A)-LD-AT)5_H8_pET22b
pFS046	Strepl_DEBS3(KS(D134A)-LD-AT)5_H8_pET22b
pFS047	Strepl_DEBS3(KS-LD(R510A)-AT)5_H8_pET22b
pFS048	Strepl_DEBS3(KS-LD(V518A)-AT)5_H8_pET22b
pFS049	Strepl_DEBS3(KS-LD-AT(Y867A))5_H8_pET22b
pFS050	Strepl_DEBS3(KS-LD-AT(A868R))5_H8_pET22b
pFS051	Strepl_m(KS-LD(H425A)-AT)_H8_pET22b
pFS052	Strepl_m(KS-LD(V479A)-AT)_H8_pET22b
pFS053	Strepl_m(KS-LD(E481A)-AT)_H8_pET22b
pFS054	Strepl_m(KS-LD(Q483A)-AT)_H8_pET22b
pFS055	Strepl_m(KS-LD-AT(H804A))_H8_pET22b
pFS056	Strepl_m(KS-LD-AT(L805A))_H8_pET22b
pFS057	Strepl_m(KS-LD)-DEBS3(AT)5_H8_pET22b
pFS058	Strepl_m(KS-LD)-DEBS3(AT)5_H8_pET22b
pFS059	Strepl_m(KS-LD)-DEBS3(AT)5_H8_pET22b
pFS060	Strepl_m(KS)-DEBS3(LD-AT)5_H8_pET22b
pFS061	Strepl_m(KS)-DEBS3(LD-AT)5_H8_pET22b
pFS062	Strepl_m(KS)-DEBS3(LD-AT)5_H8_pET22b
pFS116	Strepl_m(KS-LD)-Ecoli(FabD)_H8_pET22b
pFS117	Strepl_m(KS-LD)-Ecoli(FabD)_H8_pET22b
pFS118	Strepl_m(KS-LD)-DEBS3(AT)6_H8_pET22b
pFS119	Strepl_m(KS-LD)-DEBS3(AT)6_H8_pET22b
pFS138	Strepl_DEBS3(KS-LD-AT)6_H8_pET22b
pFS139	Strepl_DEBS3(KS-LD)6-m(AT)_H8_pET22b
pFS163	Strepl_DEBS(DD)5-m(KS-LD)-VemG(AT-ACP)-SZ3_H6_pET22b
pFS164	Strepl_DEBS(DD)5-m(KS)-VemG(LD-AT-ACP)-SZ3_H6_pET22b
pFS165	Strepl_DEBS(DD)5-m(KS-LD)-VemH(AT-ACP)-SZ3_H6_pET22b
pFS166	Strepl_DEBS(DD)5-m(KS)-VemH(LD-AT-ACP)-SZ3_H6_pET22b
pFS167	Strepl_DEBS(DD)5-m(KS-LD)-DEBS3(AT)6-m(L-ACP)-SZ3_H6_pET22b
pFS168	Strepl_DEBS(DD)5-m(KS-LD)-DEBS3(AT)6-m(L)-DEBS3(ACP)6-SZ3_H6_pET22b
pFS170	Strepl_DEBS(DD)5-m(KS-LD)-EcPKS1(AT-L-ACP)-SZ3_H6_pET22b
pFS171	Strepl_DEBS(DD)5-m(KS)-EcPKS1(LD-AT-L-ACP)-SZ3_H6_pET22b
pFS172	Strepl_DEBS(DD)5-m(KS)-EcPKS1(LD-AT-L-ACP)-SZ3_H6_pET22b
pFS173	Strepl_DEBS(DD)5-EcPKS1(KS-LD-AT-L-ACP)-SZ3_H6_pET22b
pMJD091	Npt_Splat_pCDF-1b
pMJD094	H6_DEBS3(ACP)6_pET28a(+)

Amino Acid Sequences of Chimeras

FS031 StrepI-m(KS-LD)-**DEBS3(AT)5**-m(DH-ER-KR-ACP)-H8

MSAWSHPPQFEKGGGSGGGSSGSAWHPQFEKAAAEVVIAGMSGKLPESENLQEFWANLIGGVDMVTDDD
 RRWKAGLYGLPKRSGKLDLSKFDASFFGVHPKQAHTMDPQLRLLLEVSYEAIVDGGINPASLRGTNTGVW
 VGVSGEASEALSRDPETLLGYSMVGCGRAMMANRLSFFDFKGPSIALDTACSSLLALQNAVQAIRSGEC
 PAALVGGINLLLKPNTSVQFMKLGMLSPDGTCRSFDSSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTN
 TDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPESLEYIEAHGTGTVKGDPELNGITRSLCAFRQAPLLIGST
 KSNMGHPEPASGLAALTKVLLSLEHGVWAPNLHFHNPNEIPALLDGRLQVDRPLPVRGGNVGINSFGGG
 SNVHVILQPNTQAPAPTAHAALPHLLHASGRTLEAVQDLLEQGRQHSQDLAFVSMINDIAATPTAAMPFR
 GYTVLGVVEGRVQEVQQVSTRRGVAMVFPGGQAQWQGMARDLLRESQVFADSIRD CERALAPHVDW
SLTDLLSGARPLDRVDVVQPALFAMVSLAALWRSHGVPEAAVVGHSQGEIAAAHVAGALTLEDAA
KLVAVRSRVLRRLLGGQGGMASFGLGTEQAERIGRFAGALSIVNGPRSVVAGESGPLDELIAEC
EAEGITARRIPVDYASHSPQVESLREELLTELAGISPVSAVALYSTTTGQPIDTATMDTAYWYANL
REQVRFQDATRQLAEAGFADFVEVSPHPVLTVGIEATLDSALPADAGACVVGTLRRDRGGLADFH
TALGEAYAQGINVNPNALFPPVEFPAPRGTP LISPHIKWDHSQTDVDPVAEDFPNGSSSSSATVYSIDASPE
 SPDHYLVDHCIDGRVIFPGTYLCLVWKT LARSLGSLLEETPVVFENVSFHQATILPKTGTVALEVRLLLEASH
 AFEVSDTGNLIVSGKVYLWEDPNSKLFHDHPEVPTPPESASVSRLTQGEVYKELRLRGYDYGPQFQGICEATL
 EGEQGKLLWKDNWVTFMDTMLQVSILGSSQQSLQLPTRVTAIYIDPATHRQKVYRLKEDTQVADVTTSRC
 LGITVSGGIHISRLQTTATSRRQEQQLVPTLEKVFVTPHMEAECLSESTALQKELQLCKGLARALQTKATQQ
 GLKAAMLGQEDPPQHGLPRLAAACQLQLNGNLQLELGEALAQERLLL PEDPLISGLLNSQALKACVDTALE
 NLSTLKMKVAEVLAGEGHLYSRIPALLNTQPMLQLEYTATDRHPQALKDVQTKLQQHDVAQQGWNPSDPA
 PSSLGALDLLVCNCALATLGD PALALDNMVAALKEGGFLLVHTVLKGHALGETLACL PSEVQPAPSLLSQEE
 WESLFSRKALHLVGLKRSFYGTALFLCRRRAIPQEKPIFLSVEDTSFQWVDSLKSTLATSSSQPVWLTAMDCP
 TSGVVGLVNC LRKEPGGHRIRCILLSNLSNTSHAPKLDPGSPELQQVLKHDLMNVYRDGAWGAFRHFQLE
 QDKPKEQTAHAFVNLTRGD LASIRWVSSPLKHTQPSSSGAQLCTVYYASLNFRDIMLATGKLS PDAIPGK
 WASRDCMLGMEFSGRDRCGRRVMGLVPAEGLATSVLLSSDFLWDVPSSWTLEEAASVPVYTTAYYSLVV
 RGRIQRGETVLIHSGSGVGQA AISIALSLGCRVFTTVGSAEKRAYLQARFPQLDDTSFANSRDT SFEQHVLL
 HTGGKGVLDLVLNSLAEEKLQASVRCLAQHGRFLEIGKFDLSNNHPLGMAIFLKNVTFHGILLDALFEEANDS
 WREVAALLKAGIRDGVVKPLKCTVFPKAQVEDAFRYMAQGKHIGKVLVQVREEEPEAVLPGAQPTLISAISK
 TFCPAHKSYYITGGLGGFLELARWLVL RGAQRLVLT SRSGIRTGYQAKHIREWRRQGIQVLVSTSNVSSLEGA
 RALIAEATKLG PVGGVFNLAMVLRDAMLENQTPELFQDVNKPKNYNGTLNLD RATREACPELDYFVAFSSVS
 CGRGNAGQTNYGFANSTMERICEQRRHDGLPGLAVQWGAIGDVGIVLEAMGTNDTVIGGTL PQRISSCMEV
 LDLFLNQPHAVLSSFVLA EKKAVAHGDGDTQRDLVKAVAHILGIRDLAGINLDSTLADLGLDSL MGVEVRQIL
 EREHDLVLPMEVRQLTLRKLQEMSSKTDSATDTTAPKRSRDTSLKQNLNLSTLLVNPEGPTLTQLNSVQ
 SSERPLFLVHPIEGSTTVFHSLAAKLSVPTYGLQCTQAAPLDSIPNLAAYYIDCIKQVQPEGPYRIAGYSFGAC
 VAFEMCSQLQAQQGPAPTHNNLFLFDGSHTYVLAYTQSYRAKMT PGCEAEAEAEALCFFIKQFLDVEHSKV
 LEALLPLKSLED RVAASVDLITKSHHSLDRRELSFAAVSFYHKLRAADQYKPKAKYHGNVTL LRAKTGGTY
 GEDLGADYNLSQVCDGKVSVHIIEGDHRTLLESGSLESIIIIHSSLAEP RVS VREGLEHHHHHHHH

FS032 StrepI-m(KS)-**DEBS3(LD-AT)5**-m(DH-ER-KR-ACP)-H8

MSAWSHPQFEKGGGSGGGSSAWSHPQFEKAAAEVVIAGMSGKLPESENLQEFWANLIGGVDMVTDDD
 RRWKAGLYGLPKRSGKLDLSKFDASFFGVHPKQAHTMDPQLRLLLEVSYEIVDGGINPASLRGTNTGVW
 VGVSGSEASEALSRDPETLLGYSMVGCQRAMMANRLSFFDFKGPSIALDTACSSLLALQNAYQAIRSGEC
 PAALVGGINLLLKPNNTSVQFMKLGMLSPDGTCSRFDSDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTN
 TDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPESLEYIEAHGTGTKVGDQPQELNGITRSLCAFRQAPLLIGST
 KSNMGHPEPASGLAALTKVLLSLEHGVWAPNLHFHNPNEIPALLDGRLQVVDRLPVRGGNVGINSFGFGG
 SNVHVILQPNEADEPEPAPDSGPVPLVLSGRDEQAMRAQAGRLADHLAREPRNSLRDTGFTLATRR
SAWEHRAVVVGDRRDDALAGLRAVADGRIADRTATGQARTRRGVAMVFPQGAQWQGMARDLLR
ESQVFADSIRDCERALAPHVDWSLTDLLSGARPLDRVDVVQPALFAMVSLAALWRSHGVEPAAV
VGHSQGEIAAAHVAGALTLEDAAKLVAVRSRVLRRLLGGQGGMASFGLGTEQAAERIGRFAGALSIA
VNGPRSVVVAGESGLDELIAECEAEGITARRIPVDYASHSPQVESLREELLELAGISPVSAVALYS
TTTGQPIDTATMDTAYWYANLREQVRFQDATRQLAEAGFDAFVEVSPHPVLTVGIEATLDSALPA
DAGACVVGTLRDRGLADFH**TALGEAYAQ**GVEVDW**SPAFAPVEFP**APRGTPLISPHIKWDHSQTWD
 VPVAEDFPNGSSSSSATVYSIDASPEPDHYLVDHCIDGRVIFPGTGYLCLVWKTLARSLGLSLEETPVVFEN
 VSFHQATILPKTGTVALEVRLLEASHAFEVSDTGNLIVSGKVYLWEDPNSKLFDPHPEVPTPPESASVSRLTQ
 GEVYKELRLRGYDYGPFQFGICEATLEGEQKLLWKNWVTFMDTMLQVSILGSSQQSLQLPTRVTAIYID
 PATHRQKVYRLKEDTQVADVTTSRCLGITVSGGIHISRLQTTATSRRQQEQLVPTLEKVFVTPHMEAECLSE
 STALQKELQLCKGLARALQTKATQQGLKAAMLGQEDPPQHGLPRLLAAACQLQLNGLQLELGEALAQER
 LLLPEDPLISGLLNSQALKACVDTALENLSTLKMKVAEVLAGEGHLYSRIPALLNTQPMLQLEYTATDRHPQ
 ALKDVQTKLQQHDVAQGGQWNPSPAPSSGALDLLVCNCALATLGDPALALDNMVAALKEGGFLLVHTVL
 KGHALGETLACLPEVQPAPSSLSQEEWESLFSRKALHLVGLKRSFYGTALFLCRAIPQEKPIFLSVEDTSFQ
 WVDSLKSTLATSSSQPVWLTAMDCPTSGVVGLVNLKRKEPGGHRIRCILLSNLSNTSHAPKLDPGPELQQV
 LKHDLVMNVYRDGAWGAFRHFQLEQDKPKEQTAHAFVNVLTRGDLASIRWVSSPLKHTQPSSSGAQLCTV
 YYASLNFRDIMLATGKLSIPAIGKASRDCMLGMEFSGRDRCGRRVMGLVPAEGLATSULLSSDFLWDVP
 SSWTLEEAASVPVVYTTAYYSLVVRGRIQRGETVLIHSGSGGVGQAISIALSLGCRVFTTVGSAEKRAYLQA
 RFPQLDDTSFANSRDTSEFQHVLLHTGGKGVLDLVLNSLAEELQASVRCLAQHGRFLEIGKFDLSNNHPLG
 MAIFLKNVTFHGILLDALFEEANDSWREVAALLKAGIRDGVVKPLKCTVFPKAQVEDAFRYMAQKGHIGKV
 LVQVREEEPEAVLPGAQPTLISAISKTFCPAHKSYIITGGLGGFLELARWLVRGAQRLVLTSSRGIRTGYQA
 KHIREWRRQGIQVLVSTSNVSSLEGARALIAEATKLGPGVGVFNLAMVLRDAMLENQTPELFQDVNPKPKYN
 GTLNLDRA**TREACPELDYFVAFSSV**SCGRGNAGQTNYGFANSTMERICEQRRHDGLPGLAVQWGAIGDVGI
 VLEAMGTNDTVIGGTL**PQR**ISSCMEVLDLFLNQPHAVLSSFVLAEKKA**VAHGDGDTQRDLV**KAVAHILGIRDL
 AGINLDSTLADLGLDSLMGVEVRQILEREHDLVLPMEVRQTLRKLQEMSSKTDSATDTTAPKSRSDTSLK
 QNQLNLSTLLVNPEGPTLTQLNSVQSSERPLFLVHPIEGSTTVFHSLAAKLSVPTYGLQCTQAAPLDSIPNLA
 AYYIDCIKQVQPEGPYRIAGYSFGACVAFEMCSQLQAQGGPAPTHNNLFLFDGSHTYVLAYTQSYRAKMTP
 GCEAEAEAEALCFFIKQFLDVEHSKVLEALLPLKSLEDRVAASVDLITKSHSLDRRELSFAAVSFYHKLRAA
 DQYKPKAKYHGNVTL**LRAKTGGTYGEDLGADYNLSQVCDGKVS**VHIIEGDHRTLLEGGSGLESIIIIHSSLAE
 PRVSVREGLEHHHHHHHH

FS033 StrepI-m(KS)-**DEBS3(LD-AT)5**-m(DH-ER-KR-ACP)-H8

MSAWSHQPFEKGGGSGGGSSAWSHQPFEKAAAEVVIAGMSGKLPESENLQEFWANLIGGVDMTDDD
 RRWKAGLYGLPKRSGKLDLSKFDASFFGVHPKQAHTMDPQLRLLLEVSYEIVDGGINPASLRGTNTGVW
 VGVSGSEASEALSRDPETLLGYSMVGCQRAMMANRLSFFDFKGPSIALDTACSSSLLALQNAVQAIRSGEC
 PAALVGGINLLLKPNTSVQFMKLGMLSPDGTCRSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTN
 TDGSKEQGVTFPSPGEVQEQLICSLYQPAGLAPESLEYIEAHGTGTKVGDPELNGITRSLCAFRQAPLLIGST
 KSNMGHPEPASGLAALTKVLLSLEHGVWAPNLHFHNPNPEIPALLDGRLQVVDRLPVRGGNVGINSFGFGG
 SNVHVILQPNEADEPEPAPDSGPVPLVLSGRDEQAMRAQAGRLADHLAREPRNSLRDTGFTLATRR
SAWEHRAVVVGDRDALAGLRAVADGRIADRTATGQARTRRGVAMVFPQGGAQWQGMARDLLR
ESQVFADSIRDCERALAPHVDWSLTDLLSGARPLDRVDVQPALFAVMVSLAALWRSHGVEPAAV
VGHSQGEIAAAHVAGALTLEDAAKLVAVRSRVLRRLLGGQGGMASFGLGTEQAAERIGRFAGALSIA
VNGPRSVVVAGESGPLDELIAECEAEGITARRIPVDYASHSPQVESLREELLTELAGISPVSAADVLYS
TTTGQPIDTATMDTAYWYANLREQVRFQDATRQLAEAGFDAFVEVSPHPVLTVGIEATLDSALPA
DAGACVGTLRDRRGLADFHTALGEAYAQGEVDWSPAFADARPEPLISPHIKWDHSQTWDVVPV
 AEDFPNGSSSSSATVYSIDASPESPDHYLVDHCIDGRVIFPGTGYLCLWKTLLARSLGLSLEETPVVFENVSF
 HQATILPKTGTVALEVRLLEASHAFEVSDTGNLIVSGKVYLWEDPNSKLFDPHEVPTPPESASVSRLTQGEV
 YKELRLRGYDYGPPQFQGICEATLEGEQKLLWKNWVTFMDTMLQVSILGSSQQSLQLPTRVTAIYIDPAT
 HRQKVYRLKEDTQVADVTTSRCLGITVSGGIHISRLQTATSRRQQEQLVPTLEKFVFTPHMEAECLSESTA
 LQKELQLCKGLARALQTKATQQGLKAAMLGQEDPPQHGLPRLAAACQLQLNGNLQLELGEALAQERLLL
 PEDPLISGLLNSQALKACVDTALENLSTLKMKVAEVLAGEGHLYSRIPALLNTQPMLQLEYTATDRHPQALK
 KDVTQKLQHQDVAQGWNPSPDPAPSSLGALDLLVCNCALATLGDPALALDNMVAALKEGGFLLVHTVLKG
 HALGETLACLPEVQPAPELLSQEEWESLFSRKALHLVGLKRSFYGTALFLCRAIPQEKPIFLSVEDTSFQW
 VDSLKSTLATSSSQPVWLTAMDCPTSGVVGLVNLCKRKEPGGHRIRCILLSNLSNTSHAPKLDPGSPELQQL
 KHDLMNVYRDGAWGAFRHFQLEQDKPKEQTAHAFVNVLTRGDLASIRWVSSPLKHTQPSSSGAQLCTVY
 YASLNFRDIMLATGKLSPPAIPGKWASRDCMLGMEFSGRDRCGRRVMGLVPAEGLATSULLSSDFLWDVPS
 SWTLEEAASVPVVYTTAYYSLVVRGRIQRGETVLIHSGGGVQAAISIALSLGCRVFTTVGSAEKRAYLQAR
 FPQLDDTSFANSRDTSEFQHVLLHTGGKGVLDLNLNLAEEKLQASVRCLAQHGRFLEIGKFDLSNNHPLGM
 AIFLKNVTFHGILLDALFEEANDSWREVAALLKAGIRDGVVKPLKCTVFPKAQVEDAFRYMAQKGKHKVVLV
 QVREEEPEAVLPGAQPTLISAISKTFCPAHKSYIITGGLGGFLELARWLVLVLRGAQRLVLTSSRSGIRTGYQAKHI
 REWRRQGIQVLVSTSNVSSLEGARALIAEATKLGPPGGVFNLMVLRDAMLENQTPELFQDVNPKPYNGTL
 NLDTRATREACPELDYFVAFSSVSCGRGNAGQTNYGANSTMERICEQRRHDGLPGLAVQWGAIGDVGIVLEA
 MGTNDTVIGGTLPPQRISSCMEVLDLFLNQPHAVLSSFVLAEKKAHAHGDGDTQRDLVKAVAHILGIRDLGIN
 LDSTLADLGLDSLGMGEVRQILEREHDLVLPMEVRQLTLRKLQEMSSKTDSATDTTAPKSRSDTSLKQNNQ
 LNLSTLLVNPEGPTLTQLNSVQSSERPLFLVHPIEGSTTVFHSLAAKLSVPTYGLQCTQAAPLDSIPNLAAYY
 IDCIKQVQPEGPYRIAGYSFGACVAFEMCSQLQAQQGPAPTHNNLFLFDGSHTYVLAYTQSYRAKMTPGCE
 AEAEAEALCFFIKQFLDVEHSKVLEALLPLKSLEDRVAASVDLITKSHSLDRRELSFAAVSFYHKLRAADQY
 KPKAKYHGNVTLLRAKTGGTYGEDLGADYNLSQVCDGKVSVHIIEGDHRTLLEGGSGLESIIIIHSSLAEPV
 SVREGLEHHHHHHHH

FS034 StrepI-m(KS-LD)-**RAPS3(AT)14**-m(DH-ER-KR-ACP)-H8

MSAWSHPQFEKGGGSGGGSSAWSHPQFEKAAAEVVIAGMSGKLPESENLQEFWANLIGGVDMVTDDD
 RRWKAGLYGLPKRSGKLDLSKFDASFFGVHPKQAHTMDPQLRLLLEVSYEAIVDGGINPASLRGTNTGVW
 VGVSGSEASEALSRDPETLLGYSMVGCQRAMMANRLSFFDFKGPSIALDTACSSLLALQNAQAIRSSEC
 PAALVGGINLLKPNNTSVQFMKLGMLSPDGTCSRFDSDSGSYCRSEAVVAVLLTKKSLARRVYATILNAGTN
 TDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPESLEYIEAHGTGTKVGDQPQELNGITRSLCAFRQAPLLIGST
 KSNMGHPEPASGLAALTKVLLSLEHGVWAPNLHFHNPNEIPALLDGRQLQVDRPLPVRGGNVGINSFGFGG
 SNVHVILQPNTRQAPAPATAHAALPHLLHASGRTLEAVQDLLEQGRQHSQDLAFVSMNDIAATPTAAMPFR
 GYTVLGVVEGRVQEVQVST**ARRVAFLFDGQGTQRLGMGKELYDSYPAFARAWDTVSAGFDKHLDH**
SLTDVCFGEGGSTTAGLVDDTLYAQAGIFAMEAALFGLLEDWGVPRPDFVAGHSIGEAATAAYASGML
SLENTTTLIVARGRALRTTPPGAMVALRAGEEEVREFLSRTGAALDLAAVNSPEAVVVSAGEPEPVA
DFAAWTASGREARKLKVRRHAFHSRHEAVLDEFRTALESKFRAPALPVVSTVTGRLIDQDEMG
TPEYWLRQVRRPVRFQDAVRELAEQGVGTFVEVGPSSGALASAGVECLGGDASFHAVLRPRSPEDV
CLMTAIAELHAGGINVNPALFPPVEFPAPRGTPPLISPHIKWDHSQTWDVPVAEDFPNGSSSSSATVYSID
 ASPESPDHYLVDHCIDGRVIFPGTGYLCLVWKT LARSGLSLEETPVVFENVSFHQATILPKTGTVALEVRL
 EASHAFEVSDTGNLIVSGKVYLWEDPNSKLFDPHEVPTPPESASVSRLTQGEVYKELRLRGYDYGPFQGC
 EATLEGEQGKLLWKDNWVTFMDTMLQVSILGSSQSLQLPTRVTAIYIDPATHRQKVYRLKEDTQVADVT
 TSRCLGITVSGGIHSRLQTTATSRRQQEQLVPTLEKFVFTPHMEAECLSESTALQKELQLCKGLARALQTK
 ATQQGLKAAMLGQEDPPQHGLPRLAAACQLQLNGNLQLELGEALAQERLLLPEPLISGLLNSQALKACV
 DTALENLSTLKMKVAEVLAGEGHLYSRIPALLNTQPMLQLEYTATDRHPQALKDVQTKLQVQHDVAQGW
 PSDPAPSSLGALDLLVCNCALATLGDPALALDNMVAALKEGGFLLVHTVLKGHALGETLACLPEVQPAPSSL
 SQEWEWESLFSRKALHLVGLKRSFYGTALFLCRAIPQEKPILSVEDTSFQWVDSLKSTLATSSSQPVWLTA
 MDCPTSGVVGLVNCLRKEPGGHRIRICLLSNLSNTSHAPKLDPGSPELQQVLKHDLMNVYRDGAWGAFRH
 FQLEQDKPKEQTAHAFVNVLTRGDLASIRWVSSPLKHTQPSSSGAQLCTVYYASLNFRDIMLATGKLSPTAI
 PGKWASRDCMLGMEFSGRDRCGRRVMGLVPAEGLATSULLSSDFLWDVPSSWTLEEAASVPVYTTAYYS
 LVVRGRIQRGETVLIHSGSGVQAASIALSLGCRVFTTVGSAEKRAYLQARFPQLDDTSFANSRDTSFQEH
 VLLHTGGKGVLDLVLNSLAEEKLQASVRCLAQHGRFLEIGKFDLSNNHPLGMAIFLKNVTFHGILLDALFEEAN
 DSWREVAALLKAGIRDGVVPLKCTVFPKAQVEDAFRYMAQGKHIGKVLVQVREEEPEAVLPGAQPTLISAI
 SKTFCPAHKSYIITGGLGGFLELARWLVLRGAQRLVLTSRSGIRTGYQAKHIREWRRQGIQVLVSTSNVSSLE
 GARALIAEATKLGPPGGVFNLAMVLRDAMLENQTPELFQDVNPKPKYNGTLNLDTRATREACPELDYFVAFSS
 VSCGRGNAGQTNYGFANSTMERICEQRRHDGLPGLAVQWGAIGDVGIVLEAMGTNDTVIGGTLPQRISCM
 EVLDLFLNQPHAVLSSFVLAEKKAHAHGDGDTQRDLVKAVAHILGIRDLAGINLDSTLADLGLDSLGMGEVRQ
 ILEREHDLVLPMEVRQLTLRKLQEMSSKTDSATDTTAPKSRSDTSLKQNLNLSTLLVNPEGPTLTQLNSV
 QSSERPLFLVHPIEGSTTVFHSLAAKLSVPTYGLQCTQAAPLDSIPNLAAYYIDCIKQVQPEGPYRIAGYSFGA
 CVAFEMCSQLQAQQGPAPTHNNLFLFDGSHTYVLAYTQSYRAKMTPGCEAEAEALCFFIKQFLDVEHSK
 VLEALLPLKSLEDRVAASVDLITKSHSLDRRELSFAAVSFYHKLRAADQYKPKAKYHGNVTLLRAKTGGTY
 GEDLGADYNLSQVCDGKVSVHIIEGDHRTLLEGGSGLESIINIIHSSLAEPVSVREGLEHHHHHHHH

FS035 StrepI-m(KS)-**RAPS3(LD-AT)14**-m(DH-ER-KR-ACP)-H8

MSAWSHPPQFEKGGGSGGGSSAWSHPPQFEKAAAEVVIAGMSGKLPESENLQEFWANLIGGVDMMTDDD
 RRWKAGLYGLPKRSGKLDLSKFDASFFGVHPKQAHTMDPQLRLLLEVSYEAIVDGGINPASLRGTNTGVW
 VGVSGSEASEALS RDPETLLGYSMVGCQRAMMANRLSFFDFKGPSIALDTACSSSLLALQNAQAIIRSGEC
 PAALVGGINLLLKPNTSVQFMKLGMLSPDGT CRSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTN
 TDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPESLEYIEAHGTGTKVGDPELNGITRSLCAFRQAPLLIGST
 KSNMGHPEPASGLAALT KVLLSLEHGVWAPNLHFHNPNEIPALLDGRLQVVD RPLPVRGGNVGINSFGFGG
 SNVHVILQPNA**PPVTPVTPVTPNEPGLPWVLSAQSPKALREQAGRLYASLAGDSEWNSLDIGYSL**
ATTRSDFAHRAVAVGSGREDFLRALS KLADGAPWPGLTTATATAKARRVAFLFDGQGTQRLGMGK
ELYDSYPA FARAWDTVSAGFDKHL D HSLTDVCFEGEGSTTAGLVDDTLYAQAGIFAMEAALFGLLE
DWGVRPDFVAGHSIG EATAAYASGMLSLENTTLIVARGRALRTTPPGAMVALRAGEEEVREFLSR
TGAALD LAAVNSPEAVVVS GEPEPVADFEAAWTASGREARKLKVRHAFHSRHEAVLDEFRTALE
SLKFRAPALPVVSTVTGRLIDQDEMGTP EYWLRQVRRPVRFQDAVRELAEQGVGT FVEVGP SGA
LASAGVECLGGDASFHAVLRPRSPEDVCLMTAIAELHAGGTAIDWAKVLS PVEFPAPRGTPLISPHIK
 WDHSQTWDVPVAEDFPNGSSSSSATVYSIDASPESPDHYLV DHCIDGRVIFPGTGYLCLVWKT LARSLGLSL
 EETPVVFENV SFHQATILPKTGTVALEVRLL EASHAFEVSDTGNLIVSGKVYLWEDPNSKLF DHPVPTPPE
 SASVSRLTQGEVYKELRLRGYDYGPFQFGICEATLEGEQKLLWKNWVT FMDTMLQVSILGSSQQSLQLP
 TRVTAIYIDPATHRQKVYRLKEDTQVADVTTSRCLGITVSGGIHISRLQT TATSRRQEQLVPTLEKFVFTPH
 MEAECLSESTALQKELQLCKGLARALQTKATQQGLKAAMLGQEDPPQHGLPRLAAACQLQLNGNLQLEL
 GEALAQERLLL PEDPLISGLLNSQALKACVDTALENLSTLKMKVAEVLAGEGHLYSRIPALLNTQPMLQLEYT
 ATDRHPQALKDVQTKLQQHDVAQGWNPSPD PAPSSGALDLLVCNCALATLGDPALALDNMVAALKEGG
 FLLVHTVLKGHALGETLACLPSEVQPAP SLLSQEWESELSFRKALHLVGLKRSFYGTALFLCRRRAIPQEKPIFL
 SVEDTSFQWVDSLKSTLATSSSQPVWLTAMDCPTSGVVGLVNCLRKEPGGHRIRCILLSNLSNTSHAPKLDP
 GSPQLQVLKHDLMNVYRDGAWGAFRHFQLEQDKPKEQTAHAFVNVLTRGD LASIRWVSSPLKHTQPSS
 SGAQLCTVYYASLNFRDIMLATGKLS PDAIPGKWASRDCMLGMEFSGRDRRCRRVMGLVPAEGLATSVLLS
 SDFLWDVPSSWTL EEAASVPVYTTAYYSLVVRGRIQRGETVLIHSGSGGVQAAISIALSLGCRVFTTVGSA
 EKRAYLQARFPQLDDTSFANSRDT SFEQHVLLHTGGKGVLDLVNLSLAEKQLQASVRCLAQHGRFLEIGKFDL
 SNNHPLGMAIFLKNVTFHGILLDALFEEANDSWREVAALLKAGIRDGVV KPLKCTVFPKAQVEDAFRYMAQ
 GKHIGKVLVQVREEEPEAVLPGAQPTLISAISKTF CPAHKSYIITGGLGGFLELARWLVL RGAQRLVLT SRGI
 RTGYQAKHIREWRRQGIQVLVSTSNVSSLEGARALIAEATKLG PVGGVFNLAMVLRDAMLENQTPELFQDV
 NPKPYNGTLNLD RATREACPELDYFVAFSSVSCGRGNAGQTNYGFANSTMERICEQRRHDGLPGLAVQWG
 AIGDVGIVLEAMGTNDTVIGGTL PQRISSCMEVLDLFLNQPHAVLSSFVLA EKKAVAHGDGDTQRDLV KAVA
 HILGIRDLAGINLDSTLADLGLDSL MGVEVRQILEREHDLVLPREVRQLTLRKLQEMSSKTD SATD TTAPKS
 RSDTSLKQNQLNLSLTLVNPEGPTLTQLNSVQSSERPLFLVHPIEGSTTVFHSLAAKLSVPTYGLQCTQAAP
 LDSIPNLAAYYIDCIKQVQPEGPYRIAGYSFGACVAFEMCSQLQAQQGPAPTHNNLFLFDGSHTYVLAYTQS
 YRAKMTPGCEAEAEALCFFIKQFLDVEH SKVLEALLPLKSLEDRVAASVDLITKSHHSLDRRELSFAAVSF
 YHKLRAADQYKPKAKYHGNVTL LRAKTGGTYGEDLGADYNLSQVCDGKVSVHIIEGDHRTLLEGSGLSII
 NIIHSSLAEPRVSVREGLEHHHHHHHH

FS036 StrepI-m(KS-LD)-**RAPS3(LD-AT)14**-m(DH-ER-KR-ACP)-H8

MSAWSHPQFEKGGGSGGGSSAWSHPQFEKAAAEVVIAGMSGKLPESENLQEFWANLIGGVDMTDD
 RRWKAGLYGLPKRSGKLDLSKFDASFFGVHPKQAHTMDPQLRLLLEVSYEIVDGGINPASLRGTNTGVW
 VGVSGSEASEALSRDPETLLGYSMVGCQRAMMANRLSFFDFKGPSIALDTACSSLLALQNAQAIRSGEC
 PAALVGGINLLLKPNNTSVQFMKLGMLSPDGTCRSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTN
 TDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPESLEYIEAHGTGTKVGDQPQELNGITRSLCAFRQAPLLIGST
 KSNMGHPEPASGLAALTKVLLSLEHGVWAPNLHFHNPNEIPALLDGRQLQVDRPLPVRGGNVGINSFGFGG
 SNVHVILQPNAPPVTPVTPVTPNEPGPLPWVLSAQSPKALREQAGRLYASLAGDSEWNSLDIGYSL
ATTRSDFAHRAVAVGSGREDFLRALSKLADGAPWPGLTTATATAKARRVAFLFDGQGTQRLGMGK
ELYDSYPAFARAWDTVSAGFDKHLHDHSLTDVCFEGEGSTTAGLVDDTLYAQAGIFAMEAALFGLLE
DWGVRPDFVAGHSIGEATAAYASGMLSLENTTLLIVARGRALRTTPPGAMVALRAGEEEVREFLSR
TGAALDLAAVNSPEAVVVSAGEPEPVADFEAAWTASGREARKLKVRHAFHSRHEAVLDEFRTALE
SLKFRAPALPVVSTVTGRLIDQDEMGTPEYWLRQVRRPVRFQDAVRELAEQGVGTFVEVGPSSGA
LASAGVECLGGDASFHAVLRPRSPEDVCLMTAIAELHAGGTAWKVLSSGGRAVDPLISPHIKWDH
 SQTWDVPVAEDFPNGSSSSSATVYSIDASPEPDHYLVDHCIDGRVIFPGTGYLCLVWKTLARSLGLSLEETP
 VVFENVSFHQATILPKTGTVALEVRLLLEASHAFEVSDTGNLIVSGKVYLWEDPNSKLFDPHEVPTPPESASV
 SRLTQGEVYKELRLRGYDYGPFQGICEATLEGEQGKLLWKDNWVTFMDTMLQVSILGSSQSLQLPTRV
 TAIYIDPATHRQKVYRLKEDTQVADVTTSRCLGITVSGGIHISRLQTTATSRRQEQQLVPTLEKFVFTPHMEA
 ECLSESTALQKELQLCKGLARALQTKATQQGLKAAMLGQEDPPQHGLPRLLAACQLQLNGNLQLELGEA
 LAQERLLLPEPLISGLLNSQALKACVDTALENLSTLKMKVAEVLAGEGHLYSRIPALLNTQPMLQLEYTATD
 RHPQALKDVQTKLQQHDVAQQQWNPSDPAPSSLGALDLLVCNCALATLGDPALALDNMVAALKEGGFLL
 VHTVLKGHALGETLACLPSVQPAPSLLSQEEWESLFSRKALHLVGLKRSFYGTALFLCRAIPQEKPIFLSVE
 DTSFQWVDSLKSTLATSSSQPVWLTAMDCPTSGVVGLVNCLRKEPGGHRIRCILLSNLSNTSHAPKLDPGSP
 ELQQVLKHDLMNVYRDGAWGAFRHFQLEQDKPKEQTAHAFVNVLTRGDLASIRWVSSPLKHTQPSSSGA
 QLCTVYYASLNFRDIMLATGKLSPPDAIPGKWASRDCMLGMEFSGRDRCGRRVMGLVPAEGLATSVLLSSDF
 LWDVPSSWTLEEAASVPVVYTTAYYSLVVRGRIQRGETVLIHSGSGGVQAAISIALSLGCRVFTTVGSAEKR
 AYLQARFPQLDDTSFANSRDTSEFQHVLLHTGGKGVLDLVLNSLAEEKLQASVRCLAQHGRFLEIGKFDLSNN
 HPLGMAIFLKNVTFHGILLDALFEEANDSWREVAALLKAGIRDGVVKPLKCTVFPKAQVEDAFRYMAQGKH
 IGKVLVQVREEEPEAVLPGAQPTLISAISKTFCPAHKSYIITGGLGGFLELARWLVRGAQRLVLTSSRGIRTG
 YQAKHIREWRRQGIQVLVSTSNVSSLEGARALIAEATKLGVPVGGVFNLAMVLRDAMLENQTPELFDVKNP
 KYNGTLNLDTRATREACPELDYFVAFSSVSCGRGNAGQTNYGFANSTMERICEQRRHDGLPGLAVQWGAIG
 DVGIVLEAMGTNDTVIGGTLPQRISSCMEVLDLFLNQPHAVLSSFVLAEEKKAVAHGDGDTQRDLVKAVAHIL
 GIRDLAGINLDSTLADLGLDSLGMVEVRQILEREHDLVLPREVRQLTLRKLQEMSSKTDSATDTTAPKSRS
 DTSLKQNQLNLSTLLVNPEGPTLTQLNSVQSSERPLFLVHPIEGSTTVFHSLAAKLSVPTYGLQCTQAAPLD
 SIPNLAAYYIDCIKQVQPEGPYRIAGYSFGACVAFEMCSQLQAQQGPAPTHNNLFLFDGSHTYVLAYTQSYR
 AKMTPGCEAEAEALCFFIKQFLDVEHSKVEALLPLKSLEDRVAASVDLITKSHHSLDRRELSFAAVSFYH
 KLRAADQYKPKAKYHGNVTLLRAKTGGTYGEDLGADYNLSQVCDGKVSVHIIIEGDHRTLLEGGSGLESIIII
 HSSLAEPVSVREGLEHHHHHHHHH

FS037 StrepI-m(KS-LD)-**DEBS3(AT)5**-H8

MSAWSHQPFEKGGGSGGGSGGSAWSHQPFEKGAGSEEVVIAGMSGKLPESENLQEFWANLIGGVDMTDDD
 RRWKAGLYGLPKRSGKLDLSKFDASFFGVHPKQAHTMDPQLRLLLEVSYEAIVDGGINPASLRGTNTGVW
 VGVSGEASEALSRDPETLLGYSMVGCQRAMMANRLSFFFDFKGPSIALDTACSSSLLALQNAYQAIRSGEC
 PAALVGGINLLLKPNTSVQFMKLGMLSPDGTCRSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTN
 TDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPESLEYIEAHGTGTKVGDQPQLNGITRSLCAFRQAPLLIGST
 KSNMGHPEPASGLAALTKVLLSLEHGVWAPNLHFHNPNEIPALLDGRLQVVDRLPLVRGGNVGINSFGFGG
 SNVHVILQPNTSRQAPAPTAHAALPHLLHASGRTLEAVQDLLEQGRQHSQDLAFVSMINDIAATPTAAMPFR
 GYTVLGVVEGRVQEVQQVSTRRGVAMVFPQGGAQWQGMARDLLRESQVFADSIRD CERALAPHVDW
**SLTDLLSGARPLDRVDVVQPALFAVMVSLAALWRSHGVEPAAVVGHSSQGEIAAAHVAGALTLEDA
 A KLVAVRSRVLRRLLGGQGGMASFGLGTEQAAERIGRFAGALSIVNGPRSVVAGESGPLDELIAEC
 EAEGITARRIPVDYASHSPQVESLREELLTELAGISPVSAADVLYSTTTGQPIDTATMDTAYWYANL
 REQVRFQDATRQLAEAGFDAFVEVSPHPVLTVGIEATLDSALPADAGACVVGTLRRDRGGLADFH
 TALGEAYAQGINVNPALFPPVEFPAPRGTP LISPHIKWDHSQTWDVPAEDFPNGSGSPSAHHHHHHHH**

FS038 StrepI-m(KS)-**DEBS3(LD-AT)5**-H8

MSAWSHQPFEKGGGSGGGSGGSAWSHQPFEKGAGSEEVVIAGMSGKLPESENLQEFWANLIGGVDMTDDD
 RRWKAGLYGLPKRSGKLDLSKFDASFFGVHPKQAHTMDPQLRLLLEVSYEAIVDGGINPASLRGTNTGVW
 VGVSGEASEALSRDPETLLGYSMVGCQRAMMANRLSFFFDFKGPSIALDTACSSSLLALQNAYQAIRSGEC
 PAALVGGINLLLKPNTSVQFMKLGMLSPDGTCRSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTN
 TDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPESLEYIEAHGTGTKVGDQPQLNGITRSLCAFRQAPLLIGST
 KSNMGHPEPASGLAALTKVLLSLEHGVWAPNLHFHNPNEIPALLDGRLQVVDRLPLVRGGNVGINSFGFGG
 SNVHVILQPNEADEPEPAPDSGPVPLVLSGRDEQAMRAQAGRLADHLAREPRNSLRDTGFTLATRR
**SAWEHRAVVVGDRDDALAGLRAVADGRIADRTATGQARTRRGVAMVFPQGGAQWQGMARDLLR
 ESQVFADSIRD CERALAPHVDWSLTDLLSGARPLDRVDVVQPALFAVMVSLAALWRSHGVEPAAV
 VGHSSQGEIAAAHVAGALTLEDAAKLVAVRSRVLRRLLGGQGGMASFGLGTEQAAERIGRFAGALSIV
 NGPRSVVAGESGPLDELIAECEAEGITARRIPVDYASHSPQVESLREELLTELAGISPVSAADVLYS
 TTTGQPIDTATMDTAYWYANLREQVRFQDATRQLAEAGFDAFVEVSPHPVLTVGIEATLDSALPA
 DAGACVVGTLRRDRGGLADFHTALGEAYAQGV EVDWSPAFADARVPELISPHIKWDHSQTWDVPA
 AEDFPNGSGSPSAHHHHHHHHH**

FS057 StrepI-m(KS-LD)-**DEBS3(AT)5**-H8

MSAWSHQPFEKGGGSGGGSGGSAWSHQPFEKGAGSEEVVIAGMSGKLPESENLQEFWANLIGGVDMTDDD
 RRWKAGLYGLPKRSGKLDLSKFDASFFGVHPKQAHTMDPQLRLLLEVSYEAIVDGGINPASLRGTNTGVW
 VGVSGEASEALSRDPETLLGYSMVGCQRAMMANRLSFFFDFKGPSIALDTACSSSLLALQNAYQAIRSGEC
 PAALVGGINLLLKPNTSVQFMKLGMLSPDGTCRSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTN
 TDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPESLEYIEAHGTGTKVGDQPQLNGITRSLCAFRQAPLLIGST
 KSNMGHPEPASGLAALTKVLLSLEHGVWAPNLHFHNPNEIPALLDGRLQVVDRLPLVRGGNVGINSFGFGG
 SNVHVILQPNTSRQAPAPTAHAALPHLLHASGRTLEAVQDLLEQGRQHSQDLAFVSMINDIAATPTAAMPFR
 GYTVLGVVEGRVQEVQQVSTRRGVAMVFPQGGAQWQGMARDLLRESQVFADSIRD CERALAPHVDW
**SLTDLLSGARPLDRVDVVQPALFAVMVSLAALWRSHGVEPAAVVGHSSQGEIAAAHVAGALTLEDA
 A KLVAVRSRVLRRLLGGQGGMASFGLGTEQAAERIGRFAGALSIVNGPRSVVAGESGPLDELIAEC
 EAEGITARRIPVDYASHSPQVESLREELLTELAGISPVSAADVLYSTTTGQPIDTATMDTAYWYANL
 REQVRFQDATRQLAEAGFDAFVEVSPHPVLTVGIEATLDSALPADAGACVVGTLRRDRGGLADFH
 TALGEAYAQGV EVDWSPAFADARVPRGTPLISPHIKWDHSQTWDVPAEDFPNGSGSPSAHHHHHHH
 HH**

pFS058 StrepI-m(KS-LD)-**DEBS3(AT)5**-H8

MSAWSHPQFEKGGGSGGGSSGSAWSHPQFEKGAGSEEVVIAGMSGKLPESENLQEFWANLIGGVDMVTDDD
RRWKAGLYGLPKRSGKLDLSKFDASFFGVHPKQAHTMDPQLRLLLEVSYEIVDGGINPASLRGTNTGVW
VGVSGSEASEALSRDPETLLGYSMVGCQRAMMANRLSFFDFKGPSIALDTACSSLLALQNAYQAIRSGEC
PAALVGGINLLLKPNTSVQFMKLGMLSPDGTCRSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTN
TDGSKEQGVTFFPSGEVQEQLICSLYQPAGLAPESLEYIEAHGTGTKVGDQPQELNGITRSLCAFRQAPLLIGST
KSNMGHPEPASGLAALTKVLLSLEHGVWAPNLHFHNPNEIPALLDGRLQVDRPLPVRGGNVGINSFGFGG
SNVHVILQPNTRQAPAPTAHAALPHLLHASGRTLEAVQDLLEQGRQHSQDLAFVSMNDIAATPTAAMPFR
GYTVLGVVEGRVQEVQVSTNKRPL**AMVFPQGAQWQGMARDLLRESQVFADSIRD**CERALAPHVD
WSLTDLLSGARPLDRVDVVQPALFAMVSLAALWRSHGVEPAAVVGHSQGEIAAAHVAGALTLED
AAKLVAVRSRVLRRLLGGQGGMASFGLGTEQAAERIGRFAGALSIASVNGPRSVVAGESGPLDELIA
ECEAEGITARRIPVDYASHSPQVESLREELLELAGISPVSAADVLYSTTTGQPIDTATMDTAYWYA
NLREQVRFQDATRQLAEAGFDAFVEVSPHPVLTVGIEATLDSALPADAGACVVGT**LRRDRGGLAD**
FHTALGEAYAQGINVNPALFPPVEFPAPRGTP**LISPHIKWDHSQ**TWDVPVAEDFPNGSGSPSAHHHHH
HH

FS059 StrepI-m(KS-LD)-**DEBS3(AT)5**-H8

MSAWSHPQFEKGGGSGGGSSGSAWSHPQFEKGAGSEEVVIAGMSGKLPESENLQEFWANLIGGVDMVTDDD
RRWKAGLYGLPKRSGKLDLSKFDASFFGVHPKQAHTMDPQLRLLLEVSYEIVDGGINPASLRGTNTGVW
VGVSGSEASEALSRDPETLLGYSMVGCQRAMMANRLSFFDFKGPSIALDTACSSLLALQNAYQAIRSGEC
PAALVGGINLLLKPNTSVQFMKLGMLSPDGTCRSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTN
TDGSKEQGVTFFPSGEVQEQLICSLYQPAGLAPESLEYIEAHGTGTKVGDQPQELNGITRSLCAFRQAPLLIGST
KSNMGHPEPASGLAALTKVLLSLEHGVWAPNLHFHNPNEIPALLDGRLQVDRPLPVRGGNVGINSFGFGG
SNVHVILQPNTRQAPAPTAHAALPHLLHASGRTLEAVQDLLEQGRQHSQDLAFVSMNDIAATPTAAMPFR
GYTVLGVVEGRVQEVQVSTNKRPL**AMVFPQGAQWQGMARDLLRESQVFADSIRD**CERALAPHVD
WSLTDLLSGARPLDRVDVVQPALFAMVSLAALWRSHGVEPAAVVGHSQGEIAAAHVAGALTLED
AAKLVAVRSRVLRRLLGGQGGMASFGLGTEQAAERIGRFAGALSIASVNGPRSVVAGESGPLDELIA
ECEAEGITARRIPVDYASHSPQVESLREELLELAGISPVSAADVLYSTTTGQPIDTATMDTAYWYA
NLREQVRFQDATRQLAEAGFDAFVEVSPHPVLTVGIEATLDSALPADAGACVVGT**LRRDRGGLAD**
FHTALGEAYAQGV**VDWSPAFADARPV**PRGTP**LISPHIKWDHSQ**TWDVPVAEDFPNGSGSPSAHHHH
HHHH

FS060 StrepI-m(KS)-**DEBS3(LD-AT)5**-H8

MSAWSHQPFEKGGGSGGGSGGSAWSHQPFEKGAGSEEVVIAGMSGKLPESENLQEFWANLIGGVDMTDDD
 RRWKAGLYGLPKRSGKLDLSKFDASFFGVHPKQAHTMDPQLRLLLEVSYEAIVDGGINPASLRGTNTGVW
 VGVSGEASEALSRDPETLLGYSMVGCQRAMMANRLSFFDFKGPSIALDTACSSSLLALQNAYQAIRSGEC
 PAALVGGINLLLKPNTSVQFMKLGMLSPDGTCRSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTN
 TDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPESLEYIEAHGTGTKVGDQPQLNGITRSLCAFRQAPLLIGST
 KSNMGHPEPASGLAALTKVLLSLEHGVWAPNLHFHNPPEIPALLDGRLQVDRPLPVRGGNVGINSFGFGG
 SNVHVILQPNEADEPEPAPDSGPVPLVLSGRDEQAMRAQAGRLADHLAREPRNSLRDTGFTLATRR
SAWEHRAVVVGDRRDDALAGLRAVADGRIADRTATGQARTRRGVAMVFPQGGAQWQGMARDLLR
ESQVFADSIRDCERALAPHVDWSLTDLLSGARPLDRVDVVPALFAVMVSLAALWRSHGVEPAAV
VGHSQGEIAAAHVAGALTLEDAAKLVAVRSRVLRRLLGGQGGMASFGLGTEQAAERIGRFAGALSIA
VNGPRSVVVAGESGPLDELIAECEAEGITARRIPVDYASHSPQVESLREELLTELAGISPVSAVALYS
TTTGQPIDTATMDTAYWYANLREQVRFQDATRQLAEAGFDAFVEVSPHPVLTVGIEATLDSALPA
DAGACVVGTLRRDRGGLADFHTALGEAYAQQVEVDWSPAFADARPVPRGTPLISPHIKWDHSQTWD
 VPVAEDFPNGSGSPSAHHHHHHHH

FS061 StrepI-m(KS)-**DEBS3(LD-AT)5**-H8

MSAWSHQPFEKGGGSGGGSGGSAWSHQPFEKGAGSEEVVIAGMSGKLPESENLQEFWANLIGGVDMTDDD
 RRWKAGLYGLPKRSGKLDLSKFDASFFGVHPKQAHTMDPQLRLLLEVSYEAIVDGGINPASLRGTNTGVW
 VGVSGEASEALSRDPETLLGYSMVGCQRAMMANRLSFFDFKGPSIALDTACSSSLLALQNAYQAIRSGEC
 PAALVGGINLLLKPNTSVQFMKLGMLSPDGTCRSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTN
 TDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPESLEYIEAHGTGTKVGDQPQLNGITRSLCAFRQAPLLIGST
 KSNMGHPEPASGLAALTKVLLSLEHGVWAPNLHFHNPPEIPALLDGRLQVDRPLPVRGGNVGINSFGFG
GTNAHVIVEEAPEADEPEPAPDSGPVPLVLSGRDEQAMRAQAGRLADHLAREPRNSLRDTGFTLA
TRRSAWEHRAVVVGDRRDDALAGLRAVADGRIADRTATGQARTRRGVAMVFPQGGAQWQGMAR
DLLRESQVFADSIRDCERALAPHVDWSLTDLLSGARPLDRVDVVPALFAVMVSLAALWRSHGVE
PAAVVGHSQGEIAAAHVAGALTLEDAAKLVAVRSRVLRRLLGGQGGMASFGLGTEQAAERIGRFAGA
LSIASVNNGPRSVVAGESGPLDELIAECEAEGITARRIPVDYASHSPQVESLREELLTELAGISPVSA
VALYSTTTTGQPIDTATMDTAYWYANLREQVRFQDATRQLAEAGFDAFVEVSPHPVLTVGIEATLD
SALPDAGACVVGTLRRDRGGLADFHTALGEAYAQQVEVDWSPAFADARPVPEPLISPHIKWDHSQT
 WDVPVAEDFPNGSGSPSAHHHHHHHH

FS062 StrepI-m(KS)-**DEBS3(LD-AT)**5-H8

MSAWSHQPFEKGGGSGGGSSGSAWSHPQFEKGAGSEEVVIAGMSGKLPESENLQEFWANLIGGVDMVTDDD
 RRWKAGLYGLPKRSGKLDLSKFDASFFGVHPKQAHTMDPQLRLLLEVSYEIVDGGINPASLRGTNTGVW
 VGVSGSEASEALSRDPETLLGYSMVGCQRAMMANRLSFFDFKGPSIALDTACSSLLALQNAYQAIRSGEC
 PAALVGGINLLLPNTSVQFMKLGMLSPDGTCSRFDSDSGSYCRSEAVVAVLLTKKSLARRVYATILNAGTN
 TDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPESLEYIEAHGTGTKVGDQPQLNGITRSLCAFRQAPLLIGST
 KSNMGHPEPASGLAALTKVLLSLEHGVWAPNLHFHNPNEIPALLDGRQLQVDRPLPVRGGNVGINSFGFG
GTNAHVIVEEAPEADEPEPAPDSGPVPLVLSGRDEQAMRAQAGRLADHLAREPRNSLRDTGFTLA
TRRSAWEHRVVVGDREDDALAGLRAVADGRIADRTATGQARTRRGVAMVFPQGAQWQGMAR
DLLRESQVFADSIRDCERALAPHVDWSLTDLLSGARPLDRVDVVQPALFAMVMSLAALWRSHGVE
PAAVVGHSQGEIAAAHVAGALTLEDAAKLVAVRSRVLRRLLGGQGGMASFGLGTEQAAERIGRFAGA
LSIASVNGPRSVVAGESGPLDELIAECEAEGITARRIPVDYASHSPQVESLREELLTELAGISPV SAD
VALYSTTTGQPIDTATMDTAYWYANLREQVRFQDATRQLAEAGFDAFVEVSPHPVLTVGIEATLD
SALPADAGACVVGTLRRDRGGLADFHTALGEAYAQGVEVDWSPAFADARPVPRGTPLISPHIKWDH
 SQTWDVPAEDFPNGSGSPSAHHHHHHHH

FS116 StrepI-m(KS-LD)-**Ecoli(FabD)**-H8

MSAWSHQPFEKGGGSGGGSSGSAWSHPQFEKGAGSEEVVIAGMSGKLPESENLQEFWANLIGGVDMVTDDD
 RRWKAGLYGLPKRSGKLDLSKFDASFFGVHPKQAHTMDPQLRLLLEVSYEIVDGGINPASLRGTNTGVW
 VGVSGSEASEALSRDPETLLGYSMVGCQRAMMANRLSFFDFKGPSIALDTACSSLLALQNAYQAIRSGEC
 PAALVGGINLLLPNTSVQFMKLGMLSPDGTCSRFDSDSGSYCRSEAVVAVLLTKKSLARRVYATILNAGTN
 TDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPESLEYIEAHGTGTKVGDQPQLNGITRSLCAFRQAPLLIGST
 KSNMGHPEPASGLAALTKVLLSLEHGVWAPNLHFHNPNEIPALLDGRQLQVDRPLPVRGGNVGINSFGFGG
 SNVHVILQPNTRQAPAPATAHAALPHLLHASGRTLEAVQDLLEQGRQHSQDLAFVSMNDIAATPTAAMPFR
 GYTVLGVVEGRVQEVQVST**QFAFVFPQGSQTVGMLADMAASYPIVEETFAEASAALGYDLWALTQ**
QGPAEELNKTWQTQPALLTASVALYRVWQQQGGKAPAMMAGHSLGEYSALVCAGVIDFADAVRL
VEMRGKFMQEAVPEGTGAMAAIIGLDDASIAKACEEAAEGQVSPVNFNSPGQVVIAGHKEAVER
AGAACKAAGAKRALPLPVSVPSHCALMKPAADKLAVELAKITFNAPTVPVNNVDVKCETNGDA
IRDALVRQLYNPVQWTKSVEYMAAQGVEHLYEVGPGKVLTLGLTKRIVDTLTASALNEPSAMAAAL
 ELGINVNPALFPPVEFPAPRGTPLISPHIKWDHSQTWDVPAEDFPNGSGSPSAHHHHHHHH

FS117 StrepI-m(KS-LD)-**Ecoli(FabD)**-H8

MSAWSHQPFEKGGGSGGGSSGSAWSHPQFEKGAGSEEVVIAGMSGKLPESENLQEFWANLIGGVDMVTDDD
 RRWKAGLYGLPKRSGKLDLSKFDASFFGVHPKQAHTMDPQLRLLLEVSYEIVDGGINPASLRGTNTGVW
 VGVSGSEASEALSRDPETLLGYSMVGCQRAMMANRLSFFDFKGPSIALDTACSSLLALQNAYQAIRSGEC
 PAALVGGINLLLPNTSVQFMKLGMLSPDGTCSRFDSDSGSYCRSEAVVAVLLTKKSLARRVYATILNAGTN
 TDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPESLEYIEAHGTGTKVGDQPQLNGITRSLCAFRQAPLLIGST
 KSNMGHPEPASGLAALTKVLLSLEHGVWAPNLHFHNPNEIPALLDGRQLQVDRPLPVRGGNVGINSFGFGG
 SNVHVILQPNTRQAPAPATAHAALPHLLHASGRTLEAVQDLLEQGRQHSQDLAFVSMNDIAATPTAAMPFR
 GYTVLGVVEGRVQEVQVST**NKTQFAFVFPQGSQTVGMLADMAASYPIVEETFAEASAALGYDLWA**
LTQQGPAEELNKTWQTQPALLTASVALYRVWQQQGGKAPAMMAGHSLGEYSALVCAGVIDFADA
VRLVEMRGKFMQEAVPEGTGAMAAIIGLDDASIAKACEEAAEGQVSPVNFNSPGQVVIAGHKEA
VERAGAACKAAGAKRALPLPVSVPSHCALMKPAADKLAVELAKITFNAPTVPVNNVDVKCETN
GDAIRDALVRQLYNPVQWTKSVEYMAAQGVEHLYEVGPGKVLTLGLTKRIVDTLTASALNEPSAMA
AALELGINVNPALFPPVEFPAPRGTPLISPHIKWDHSQTWDVPAEDFPNGSGSPSAHHHHHHHH

FS118 StrepI-m(KS-LD)-**DEBS3(AT)6**-H8

MSAWSHQPFEKGGGSGGGGSAWSHQPFEKGAGSEEVVIAGMSGKLPESENLQEFWANLIGGVDMTDDD
 RRWKAGLYGLPKRSGKLDLSKFDASFFGVHPKQAHTMDPQLRLLLEVSYEAIVDGGINPASLRGTNTGVW
 VGVSGSEASEALSRDPETLLGYSMVGCQRAMMANRLSFFDFKGPSIALDTACSSSLLALQNAYQAIRSGEC
 PAALVGGINLLLKPNTSVQFMKLGMLSPDGTCRSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTN
 TDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPESLEYIEAHGTGTKVGDQPQLNGITRSLCAFRQAPLLIGST
 KSNMGHPEPASGLAALTKVLLSLEHGVWAPNLHFHNPNEIPALLDGRLQVDRPLPVRGGNVGINSFGFGG
 SNVHVILQPNTAQAPAPTAHAALPHLLHASGRTLEAVQDLEQGRQHSQDLAFVSMINDIAATPTAAMPFR
 GYTVLGVVEGRVQEVQQVST**AGGVVVFVPGQGAQWEGMARGLLSVPVFAESIAECDAVLSEVAGFSAS**
EVLEQRPDAPSLERVDVVQPVLFSVMVSLARLWGACGVSPSAVIGHSQGEIAAAVVAGVLSLEDGV
RVVALRAKALRALAGKGGMVSLAAPGERARALIAPWEDRISVAAVNSPSSVVVSGDPEALAEVAR
CEDEGVRAKTLPVDYASHSRHVEEIRETILADLDGISARRAAIPLYSTLHGERRDGADMGPYRWY
DNLRSQVRFDEAVSAAVADGHATFVEMSPHPVLTAAVQEIAADAVAIGSLHRDTAEEHLIAELARA
HVHGINVNPALFPPVEFPAPRGTPPLISPHIKWDHSQTWDVPAEDFPNGSGSPSAHHHHHHHH

FS119 StrepI-m(KS-LD)-**DEBS3(AT)6**-H8

MSAWSHQPFEKGGGSGGGGSAWSHQPFEKGAGSEEVVIAGMSGKLPESENLQEFWANLIGGVDMTDDD
 RRWKAGLYGLPKRSGKLDLSKFDASFFGVHPKQAHTMDPQLRLLLEVSYEAIVDGGINPASLRGTNTGVW
 VGVSGSEASEALSRDPETLLGYSMVGCQRAMMANRLSFFDFKGPSIALDTACSSSLLALQNAYQAIRSGEC
 PAALVGGINLLLKPNTSVQFMKLGMLSPDGTCRSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTN
 TDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPESLEYIEAHGTGTKVGDQPQLNGITRSLCAFRQAPLLIGST
 KSNMGHPEPASGLAALTKVLLSLEHGVWAPNLHFHNPNEIPALLDGRLQVDRPLPVRGGNVGINSFGFGG
 SNVHVILQPNTAQAPAPTAHAALPHLLHASGRTLEAVQDLEQGRQHSQDLAFVSMINDIAATPTAAMPFR
 GYTVLGVVEGRVQEVQQVSA**AGGVVVFVPGQGAQWEGMARGLLSVPVFAESIAECDAVLSEVAGFSASE**
VLEQRPDAPSLERVDVVQPVLFSVMVSLARLWGACGVSPSAVIGHSQGEIAAAVVAGVLSLEDGVR
VVALRAKALRALAGKGGMVSLAAPGERARALIAPWEDRISVAAVNSPSSVVVSGDPEALAEVARC
EDEGVRAKTLPVDYASHSRHVEEIRETILADLDGISARRAAIPLYSTLHGERRDGADMGPYRWYD
NLRSQVRFDEAVSAAVADGHATFVEMSPHPVLTAAVQEIAADAVAIGSLHRDTAEEHLIAELARAH
VHGINVNPALFPPVEFPAPRGTPPLISPHIKWDHSQTWDVPAEDFPNGSGSPSAHHHHHHHH

FS139 StrepI-DEBS3(KS-LD)6-m(**AT**)-H8

MSAWSHQPFEKGGGSGGGGSAWSHQPFEKGAGSPIAIVGMACRFPGGVHNPGELWEIFVGGDAVTEMP
 TDRGWLDALFDPDPQRHGTSYSRHGAFLDGAADFDAFFGISPREALAMDPQQRQVLETTWELFENAGI
 DPHSLRGS DTVFLGAAQQYGGQDAVVPEDSEG YLLTGNSSAVVSGRVAYVVLGLEGPAVTVDTACSSSLLVAL
 HSACGSLRDGDCGLAVAGGVSVMAGPEVFTESFRQGGLAVDGRCKAFSAEADGFGFAEGVAVVLLQRLSDA
 RRAGRQVLGVVAGSAINQDGASNGLAAPSGVAQQRVIRKAWARAGITGADVAVVEAHGTGTRLGDPVEASA
 LLATYGKSRGSSGPVLLGSVKSNIHAQAAAGVAGVIKVVLLGNRGLVPPMLCRGERSPLIEWSSGGVELAEA
 VSPWPPAADGVRRAGVSAFGVSGTNAHVIAEPEPEPELPEPGPVGLAAANSVPVLLSARTETALAAQARL
 LESAVDDSVPLTALASALATGRAHLPRRAALLAGDHEQLRGQLRAVAEGVAAPGATTGTAS**NKRPLWFICS**
GMGTQWRGMGLSLMRLDSFRESILRSDEAVKPLGVKVS DLLLSTDERTFDDIVHAFVSLTAIQIALI
DLLTSVGLKPDGIIGHSLGEVACGYADGCLSQREAVLAAYWRGQCICKDAHLPPGSMAAVGLSWEE
CKQRCPAGVVPACHNSEDTVTISGPQAAVNEFVEQLKQEGVFAKEVRTGGLAFHSYFMEGIAPTL
LQALKKVIREPRPRSARWLSTSIPEAQWQSSLARTSSAEYNNVNLVSPVLFQEALWHIPEHAVVLEI
APHALLQAVLKRGVKSSCTIIPLMKRDHKDNLEFFLTNLGKVHLTGVAVDWRNVFPAAPPVALPNYPF
 EPQRYWLAPEVSDQLADSLEHHHHHHHH

FS163 StrepI-DEBS(DD)5-m(KS-LD)-**VemG(AT-ACP)**-SZ3-H6

MSAWSHQPFEKGGGSGGGSSGSAWSHPQFEKGAGSSGDNGMTEEKLRRLKRTVTELDVSTARLREVEHR
 AGEEVVIAGMSGKLPESENLQEFWANLIGGVDMVTDDDRRWKAGLYGLPKRSGKLDLSKFDASFFGVHP
 KQAHTMDPQLRLLLEVSYEIVDGGINPASLRGTNTGVWVGVSSEASEALSRDPETLLGYSMVGCQRAM
 MANRLSFFDFKGPSIALDTACSSLLALQNAVQAIRSGECPAALVGGINLLKPNNTSVQFMKLGMLSPDGT
 RSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTNTDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPE
 SLEYIEAHGTGTVKVGDPQELNGITRSLCAFRQAPLLIGSTKSNMGHPEPASGLAALTKVLLSLEHGWWAPNL
 HFHNPPEIPALLDGRLLQVDRPLPVRGGNVGINSFGGGSNVHVILQPNTRQAPAPATAHAALPHLLHASGR
 TLEAVQDLLEQGRQHSQDLAFVSMNDIAATPTAAMPFRGYTVLGVVEGRVQEVQVST**EPSLAIGFTGQG**
SQHPGMGRELYAAFPVFAEALDAAWAALDPHLERPLRDVMWAPDGTARAALLDRTDFTQAALFA
LEGALYRLVESWGVVPDVVLGHSVGALTAHAAGVLSLPDAASLVAARGRLMAALPPGGAMTSIE
ATEDELRTLRATDGVLAIAAVNGPRSVVVSGEEKAVRTVGEAFRARGRRVTALRVSHAFHSPLM
 DPMVEEFRAAAARVTYRPPLLPMISDLTGRPAPHLRSPDHWVRHVRET VRFADAVRALPGQG
 VTAFLLELGPDRQLTTMAAAGAPGSGPALCGGLRRGRSEVRSLLDAVAQAHVRGVPVDWREFFAGG
 PTRPVDLPVYPFEHRRYWLASAPTRTGGATSVPEAPAPPTGDTGPDRLRADLAPLTDDERERR
 LTDLVRGEIAAVAGFDGPHEIEAHRSMIDLGLDSVSAVDLSTRLGARTGLDLPASLAFDHTSTAIA
RHLLAALGGGSGGGSGNEVTTLENDAAFIENENAYLEKEIARLRKEKAALRNRLAHKKLEHHHHH

FS164 StrepI-DEBS(DD)5-m(KS)-**VemG(LD-AT-ACP)**-SZ3-H6

MSAWSHQPFEKGGGSGGGSSGSAWSHPQFEKGAGSSGDNGMTEEKLRRLKRTVTELDVSTARLREVEHR
 AGEEVVIAGMSGKLPESENLQEFWANLIGGVDMVTDDDRRWKAGLYGLPKRSGKLDLSKFDASFFGVHP
 KQAHTMDPQLRLLLEVSYEIVDGGINPASLRGTNTGVWVGVSSEASEALSRDPETLLGYSMVGCQRAM
 MANRLSFFDFKGPSIALDTACSSLLALQNAVQAIRSGECPAALVGGINLLKPNNTSVQFMKLGMLSPDGT
 RSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTNTDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPE
 SLEYIEAHGTGTVKVGDPQELNGITRSLCAFRQAPLLIGSTKSNMGHPEPASGLAALTKVLLSLEHGWWAPNL
 HFHNPPEIPALLDGRLLQVDRPLPVRGGNVGINSFGGGSNVHVILQPN**EPDAPPALPPAVHPDTMPVL**
LSARTPSALRRQAERVLDAVRADPGLPVPDLAAALATTRTAFPRRAAFVVEGRAELERRLSFAD
GEDGGAEPDEVPAEPSLAIGFTGQGSQHPGMGRELYAAFPVFAEALDAAWAALDPHLERPLRDV
MWAPDGTARAALLDRTDFTQAALFALEGALYRLVESWGVVPDVVLGHSVGALTAHAAGVLSLP
DAASLVAARGRLMAALPPGGAMTSIEATEDELRTLRATDGVLAIAAVNGPRSVVVSGEEKAVRT
 VGEAFRARGRRVTALRVSHAFHSPLMDPMVEEFRAAAARVTYRPPLLPMISDLTGRPAPHLR
 SPDHWVRHVRET VRFADAVRALPGQGVTAFLLELGPDRQLTTMAAAGAPGSGPALCGGLRRGRSE
 VRSLLDAVAQAHVRGVPVDWREFFAGGPTRPVDLPVYPFEHRRYWLASAPTRTGGATSVPEAPA
 PPTGDTGPDRLRADLAPLTDDERERRLTDLVRGEIAAVAGFDGPHEIEAHRSMIDLGLDSVSAVD
LSTRLGARTGLDLPASLAFDHTSTAIARHLLAALGGGSGGGSGNEVTTLENDAAFIENENAYLEKEIAR
 LRKEKAALRNRLAHKKLEHHHHH

FS165 StrepI-DEBS(DD)5-m(KS-LD)-**VemH(AT-ACP)**-SZ3-H6b

MSAWSHPPQFEKGGGSGGGSGGSAWSHPPQFEKAGSSGDNGMTEEKLRRLYLKRTVTELD SVTARLREVEHR
 AGEEVVIAGMSGKLPESENLQEFWANLIGGVD MVTD DRRWKAGLYGLPKRSGKLDLSKFDASFFGVHP
 KQAHTMDPQLRLLLEVS YEAIVDGGINPASLRGTNTGVWVGVS GSEASEALSRDPETLLGYSMVGCQRAM
 MANRLSFFFDFKGPSIALDTACSSLLALQ NAYQAIRSGECPAALVGGINLLLKPNTSVQFMKLGMLSPDGTC
 RSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTNTDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPE
 SLEYIEAHGTGTKVGD PQELNGITRSLCAFRQAPLLIGSTKSNMGHPEPASGLAALT KVLLSLEHG VVWAPNL
 HFHNPNPEIPALLDGRLQVVD RPLPVRGGNVGINSFGFGGNSVHVILQPNTRQAPAPTAHAALPHLLHASGR
 TLEAVQD LLEQGRQHSQDLAFVSM LNDIAATPTAAMPFRGYTVLGV EGRVQEVQQVST **PGGTALLFSGQ**
GSQRVGMGSELYETYPVFAESF DAVA EHTGLPLKDVVLGGTPDGLLDRTRYAQPALFAVEVSLFRL
VRALGLDVRAVVGHSVGEIAAAHVAGVM SMADACRLVEARGRLMDALPPGGAMVAVEVTEAEAS
AALAGLEDRVAVAVNGPASTVLSGEEGAVLKLADAWRERGV RTHRLTVSHAFHSPLMEPMVDA
FREVVAGLDLHRPTLAGLPAE VVDPEYWVRHVRRPVRFADAVARAREAGAVRWLEVGPGGVLTA
LAQRIVPDTEEHVFAAALRTDRPEPEALLVALSQVHVDGGTVDWSGLCAGGRLVDLPTYPFQRQH
YWIEDQPLPPTAPRPGTAPSGTG TAAEGAAA EVPLSERLARLTGAERLA AVRELVLA EASETLGH
TGTLITADRTRQELGFD SLTAIELNRNISR TLGVRLPPTLVFDHEDLGEIASFVDARLGGGSGGGSGNE
 VTTLENDAAFIENENAYLEKEIARLRKEKAALRNRLAHKKLEHHHHHH

FS166 StrepI-DEBS(DD)5-m(KS)-**VemH(LD-AT-ACP)**-SZ3H6

MSAWSHPPQFEKGGGSGGGSGGSAWSHPPQFEKAGSSGDNGMTEEKLRRLYLKRTVTELD SVTARLREVEHR
 AGEEVVIAGMSGKLPESENLQEFWANLIGGVD MVTD DRRWKAGLYGLPKRSGKLDLSKFDASFFGVHP
 KQAHTMDPQLRLLLEVS YEAIVDGGINPASLRGTNTGVWVGVS GSEASEALSRDPETLLGYSMVGCQRAM
 MANRLSFFFDFKGPSIALDTACSSLLALQ NAYQAIRSGECPAALVGGINLLLKPNTSVQFMKLGMLSPDGTC
 RSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTNTDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPE
 SLEYIEAHGTGTKVGD PQELNGITRSLCAFRQAPLLIGSTKSNMGHPEPASGLAALT KVLLSLEHG VVWAPNL
 HFHNPNPEIPALLDGRLQVVD RPLPVRGGNVGINSFGFGGNSVHVILQPN **EEPARDA AETRPEARS REAA**
EGRREAAGQGR TETGPGPSATPPEAPGRARPPVPWPLSGRDAGALRDQIGRLRAHLDAAPADPE
DVAHSLARRAVFRHRAVLLAAPQAPAGGSPRAVTGVARPGGTALLFSGQGSQRVGMGSELYETYP
VFAESF DAVA EHTGLPLKDVVLGGTPDGLLDRTRYAQPALFAVEVSLFRLVRALGLDVRAVVGHSV
GEIAAAHVAGVM SMADACRLVEARGRLMDALPPGGAMVAVEVTEAEASAALAGLEDRVAVAVN
GPASTVLSGEEGAVLKLADAWRERGV RTHRLTVSHAFHSPLMEPMVDAFREVVAGLDLHRPTLA
GLPAE VVDPEYWVRHVRRPVRFADAVARAREAGAVRWLEVGPGGVLTA LAQRIVPDTEEHVFAAA
LRTDRPEPEALLVALSQVHVDGGTVDWSGLCAGGRLVDLPTYPFQRQHYWIEDQPLPPTAPRPG
TAPSGTG TAAEGAAA EVPLSERLARLTGAERLA AVRELVLA EASETLGHTGTLITADRTRQELGFD
SLTAIELNRNISR TLGVRLPPTLVFDHEDLGEIASFVDARLGGGSGGGSGNEVTTLENDAAFIENENAYL
 EKEIARLRKEKAALRNRLAHKKLEHHHHHH

FS167 StrepI-DEBS(DD)5-m(KS-LD)-**DEBS3(AT)6**-m(L-ACP)-SZ3-H6

MSAWSHQPFEKGGGSGGGSSGSAWSHPQFEKGAGSSGDNGMTEEKLRRLKRTVTELD SVTARLREVEHR
 AGEEVVIAGMSGKLPESENLQEFWANLIGGVDMVTDDDRRWKAGLYGLPKRSGKLDLSKFDASFFGVHP
 KQAHTMDPQLRLLLEVS YEAIVDGGINPASLRGTNTGVWVGVS GSEASEALSRDPETLLGYSMVGCQRAM
 MANRLSFFDFKGPSIALDTACSSLLALQ NAYQAIRSGECPAALVGGINLLLKPNTSVQFMKLGMLSPDGTC
 RSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTNTDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPE
 SLEYIEAHGTGTKVGD PQELNGITRSLCAFRQAPLLIGSTKSNMGHPEPASGLAALT KVLLSLEHG VWAPNL
 HFHNP NPEIPALLDGR LQVVD RPLPVRGGNVGINSFGFGGSNVH VILQPNTRQAPAPTAHAALPHLLHASGR
 TLEAVQDLLEQGRQHSQDLAFVSM LN DIAATPTAAMPFRGYTVLGVEGRVQEVQVST **AGGVVVFVPGQ**
GAQWEGMARGLLSVPVFAESIAECD AVLSEVAGFSASEVLEQRPDAPSLERVDVVQPVLFSVMVSL
ARLWGACGVSPSAVIGHSQGEIAAAVVAGVLSLEDGVRVVALRAKALRALAGKGGMVSLAAPGERA
RALIAPWEDRISVAAVNSPSSVVVSGDPEALAE LVARCEDEGVRAKTLPVDYASHSRHVEEIRETIL
ADLDGISARRAAIPLYSTLHGERRD GADMGP RYWDNLRSQVRFDEAVSAAVADGHATFVEMSP
HPVLTA AVQEI AADAVAIGSLHRDTAE EHLIAELARAHVHGINVNP NALFPPVEFPAPRG TPLISPHIKW
 DHSQ TWDVPVAEDFPNGSSSSSATVYSIDASAEKKAVAHGDGDTQRDLVKAVAHILGIRDLAGINLDSTLAD
 LGLDSL MGVEVRQILEREHDLVLP MREVRQLTLRKLQEMSSKTDSATDTTGGGSGGGSGNEVTTLENDAAF
 IENENAYLEKEIARLRKEKAALRNRLAHKKLEHHHHHH

FS168 StrepI-DEBS(DD)5-m(KS-LD)-**DEBS3(AT)6**-m(L)-**DEBS3(ACP)6**-SZ3-H6

MSAWSHQPFEKGGGSGGGSSGSAWSHPQFEKGAGSSGDNGMTEEKLRRLKRTVTELD SVTARLREVEHR
 AGEEVVIAGMSGKLPESENLQEFWANLIGGVDMVTDDDRRWKAGLYGLPKRSGKLDLSKFDASFFGVHP
 KQAHTMDPQLRLLLEVS YEAIVDGGINPASLRGTNTGVWVGVS GSEASEALSRDPETLLGYSMVGCQRAM
 MANRLSFFDFKGPSIALDTACSSLLALQ NAYQAIRSGECPAALVGGINLLLKPNTSVQFMKLGMLSPDGTC
 RSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTNTDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPE
 SLEYIEAHGTGTKVGD PQELNGITRSLCAFRQAPLLIGSTKSNMGHPEPASGLAALT KVLLSLEHG VWAPNL
 HFHNP NPEIPALLDGR LQVVD RPLPVRGGNVGINSFGFGGSNVH VILQPNTRQAPAPTAHAALPHLLHASGR
 TLEAVQDLLEQGRQHSQDLAFVSM LN DIAATPTAAMPFRGYTVLGVEGRVQEVQVST **AGGVVVFVPGQ**
GAQWEGMARGLLSVPVFAESIAECD AVLSEVAGFSASEVLEQRPDAPSLERVDVVQPVLFSVMVSL
ARLWGACGVSPSAVIGHSQGEIAAAVVAGVLSLEDGVRVVALRAKALRALAGKGGMVSLAAPGERA
RALIAPWEDRISVAAVNSPSSVVVSGDPEALAE LVARCEDEGVRAKTLPVDYASHSRHVEEIRETIL
ADLDGISARRAAIPLYSTLHGERRD GADMGP RYWDNLRSQVRFDEAVSAAVADGHATFVEMSP
HPVLTA AVQEI AADAVAIGSLHRDTAE EHLIAELARAHVHGINVNP NALFPPVEFPAPRG TPLISPHIKW
 DHSQ TWDVPVAEDFPNGSSSSSATVYSIDASAEKKAVA **APAREMTSQELLEFT HSHVAAILGHSSPDAV**
GQDQPFTELGFDSL TAVGLRNQLQ QATGLALPATLVFEHPTVRR LADHIGQQLDTTGGGSGGGSGN
 EVTTLENDAAF IENENAYLEKEIARLRKEKAALRNRLAHKKLEHHHHHH

FS170 StrepI-DEBS(DD)5-m(KS-LD)-**EcPKS1(AT-L-ACP)**-SZ3-H6

MSAWSHPPQFEKGGGSGGGSGGSAWSHPPQFEKAGSSGDNGMTEEKLRRLKRTVTELDVSTARLREVEHR
 AGEVVVIAGMSGKLPESENLQEFWANLIGGVDVMTDDDRRWKAGLYGLPKRSGKLDLSKFDASFFGVHP
 KQAHTMDPQLRLLLEVSYEAIVDGGINPASLRGTNTGVWVGVSSEASEALSRDPETLLGYSMVGCQRAM
 MANRLSFFFDKFGPSIALDTACSSSLLALQNAYQAIRSGECPAALVGGINLLLKPNTSVQFMKLGMLSPDGTC
 RSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTNTDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPE
 SLEYIEAHGTGTKVGGDPQELNGITRSLCAFRQAPLLIGSTKSNMGHPEPASGLAALTKVLLSLEHGVWAPNL
 HFHNPNPEIPALLDGRLQVVDRPLPVRGGNVGINSFGFGGNSVHVILQPNTRQAPAPTAHAALPHLLHASGR
 TLEAVQDLLEQGRQHSQDLAFVSMINDIAATPTAAMPFRGYTVLGVGRVQEVQQVST**AREVWFIYAGM**
GSQWVGMARCLMQLDVFRHSLEKSAAVLKPHGVDLLNILSEGTEDLR**TILNPFVCITSIQVALTD**
LLWSMGIRPDGIVGHSMGEVGCAYCDGCLTAEEAVLTAYWRGKCVTDGK**VPPGKMAAVGLTWE**
AKSQCPAGVVPACHNAEDSVTISGAADVMLKFMEELKAKDV**FVREVNSSNIAYHSYF**MENIYASL
 KDSLSKVISPKPRTARWLPT**SVPEELWDSAPAQSSSAEFHANNLVSPVLFHEALQKIPPTAIAIELA**
 PHG**LLQSVIKRTLGNESVCVGLQKRNYADNLEFLFASLGKCFANGLSLNPLACYPPVEFPVPK**GT**P**
 RLS**DMVAGAWD**HSAQWL**VPKNE**DFEGRVQASGSDSSYSIDV**SAEKVKA**AVEGEETVG**QQVIKAV**
 GNV**LGLKSVSGVDPDKVFLDLGLDSLMSVEIKQMLERDLDLALG**TKDIQ**MLTFAQLQAMVGGGSG**
 GSGNEVTTLENDAAFIENENAYLEKEIARLRKEKAALRNRLAHKKLEHHHHHH

FS171 StrepI-DEBS(DD)5-m(KS)-**EcPKS1(LD-AT-L-ACP)**-SZ3-H6

MSAWSHPPQFEKGGGSGGGSGGSAWSHPPQFEKAGSSGDNGMTEEKLRRLKRTVTELDVSTARLREVEHR
 AGEVVVIAGMSGKLPESENLQEFWANLIGGVDVMTDDDRRWKAGLYGLPKRSGKLDLSKFDASFFGVHP
 KQAHTMDPQLRLLLEVSYEAIVDGGINPASLRGTNTGVWVGVSSEASEALSRDPETLLGYSMVGCQRAM
 MANRLSFFFDKFGPSIALDTACSSSLLALQNAYQAIRSGECPAALVGGINLLLKPNTSVQFMKLGMLSPDGTC
 RSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTNTDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPE
 SLEYIEAHGTGTKVGGDPQELNGITRSLCAFRQAPLLIGSTKSNMGHPEPASGLAALTKVLLSLEHGVWAPNL
 HFHNPNPEIPALLDGRLQVVDRPLPVRGGNVGINSFGFGGNSVHVILQPN**VSSSKPHPSA**HK**PRLFTYSA**
RTEHGLRAILREAQ**THAASMEFH**ALC**QASADAP**LG**SMPYRGATILNGQHDFEVVEKCKSKAREV**
WFIYAGMGSQWVGMARCLMQLDVFRHSLEKSAAVLKPHGVDLLNILSEGTE**DLR**TILN**PFVCIT**
SIQVALTDLLWSMGIRPDGIVGHSMGEVGCAYCDGCLTAEEAVLTAYWRGKCVTDGK**VPPGKMA**
AVGLTWEEAK**SQCPAGVVPACHNAEDSVTISGAADV**MLKFMEELKAKDV**FVREVNSSNIAYHSYF**
MENIYASLKDSLSKVISPKPRTARWLPT**SVPEELWDSAPAQSSSAEFHANNLVSPVLFHEALQKIP**
TAIAIELAPHG**LLQSVIKRTLGNESVCVGLQKRNYADNLEFLFASLGKCFANGLSLNPLACYPPVEFP**
VPKGTPRLSDMVAGAWD**HSAQWL**V**PKNE**DFEGRVQASGSDSSYSIDV**SAEKVKA**AVEGEETVG**Q**
QVIKAVGNVLGLK**SVSGVDPDKVFLDLGLDSLMSVEIKQMLERDLDLALG**TKDIQ**MLTFAQLQAM**
 VGGGSGGGSGNEVTTLENDAAFIENENAYLEKEIARLRKEKAALRNRLAHKKLEHHHHHH

FS172 StrepI-DEBS(DD)5-m(KS)-**EcPKS1(LD-AT-L-ACP)**-SZ3-H6

MSAWSHPQFEKGGGSGGGSSAWSHPQFEKGAGSSGDNGMTEEKLRRLKRTVTELDVSVTARLREVEHR
AGEEVVIAGMSGKLPESENLQEFWANLIGGVDMVTDDDRRWKAGLYGLPKRSGKLDLSKFDASFFGVHP
KQAHTMDPQLRLLLEVSYEIVDGGINPASLRGTNTGVWVGVSSEASEALSRDPETLLGYSMVGCQRAM
MANRLSFFDFKGPSIALDTACSSLLALQNAYQAIRSGECPAALVGGINLLKPNNTSVQFMKLGMLSPDGTC
RSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTNTDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPE
SLEYIEAHGTGTVKVGDPQELNGITRSLCAFRQAPLLIGSTKSNMGHPEPASGLAALTKVLLSLEHGWWAPNL
HFHNPNPPEIPALLDGRLLQVDRPLPVRGGNVGINSFGFGGSNVHVLQPNDVSSSKPHPSAHKPRLFTYS
ARTEHGLRAILREQTHAASMEFHALCQASADAPLGSMYPYRGATILNGQHDFEVVEKCKSKARE
VWFIYAGMGSQWVGMARCLMQLDVFHRSLEKSAVLKPHGVDLLNILSEGTEDLRILNPFVCI
TSIQVALDLLWSMGIRPDGIVGHSMGEVGCAYCDGCLTAEAVLTAYWRGKCVTDGKVPPGKM
AAVGLTWEEAKSQCPAGVVPACHNAEDSVTISGAADVMLKFMEELKAKDVFVREVNSSNIAYHS
YFMENIYASLKDSLKVISPKPRTARWLPTSVPEELWDSAPAQSSSAEFHANNLVSPVLFHEALQK
IPPTAIAIELAPHGLLQSVIKRTLGNESVCVGLQKRNYADNLEFLFASLGKCFANGLSLNPLACYPPV
EFPVPGKTPRLSDMVAGAWDHSAQWLVPKNEDFEGRVQASGSDSSYSIDVSAEKVKAAVEGEET
VGQQVIKAVGNVLGLKSVSGVDPDKVFLDLGLDSLMSVEIKQMLERDLDLALGTKDIQMLTFAQL
QAMVGGGSGGGSGNEVTTLENDAAFIENENAYLEKEIARLRKEKAALRNRLAHKKLEHHHHHH

Mathematical Examination of Changes in Transacylation Kinetic Parameters Caused by Point Mutations

Kinetic parameters without bar correspond to the wild type (WT) DEBS3M5 KS-AT. Kinetic parameters with a bar correspond to the mutant (Mut).

$$\begin{aligned} k_{\text{cat}} &:= k_{\text{cat}} (\text{WT}), & \bar{k}_{\text{cat}} &:= k_{\text{cat}} (\text{Mut}), \\ K_{\text{m}}^{\text{MMal-CoA}} &:= K_{\text{m}}^{\text{MMal-CoA}} (\text{WT}), & \bar{K}_{\text{m}}^{\text{MMal-CoA}} &:= K_{\text{m}}^{\text{MMal-CoA}} (\text{Mut}), \\ K_{\text{m}}^{\text{ACP}} &:= K_{\text{m}}^{\text{ACP}} (\text{WT}), & \bar{K}_{\text{m}}^{\text{ACP}} &:= K_{\text{m}}^{\text{ACP}} (\text{Mut}). \end{aligned}$$

Kinetic parameters are described as follows:

$$\begin{aligned} k_{\text{cat}} &= \frac{k_2 k_4}{k_2 + k_4}, & \bar{k}_{\text{cat}} &= \frac{\bar{k}_2 \bar{k}_4}{\bar{k}_2 + \bar{k}_4}, \\ K_{\text{m}}^{\text{MMal-CoA}} &= \frac{k_4 (k_{-1} + k_2)}{k_1 (k_2 + k_4)}, & \bar{K}_{\text{m}}^{\text{MMal-CoA}} &= \frac{\bar{k}_4 (\bar{k}_{-1} + \bar{k}_2)}{\bar{k}_1 (\bar{k}_2 + \bar{k}_4)}, \\ K_{\text{m}}^{\text{ACP}} &= \frac{k_2 (k_{-3} + k_4)}{k_3 (k_2 + k_4)}, & \bar{K}_{\text{m}}^{\text{ACP}} &= \frac{\bar{k}_2 (\bar{k}_{-3} + \bar{k}_4)}{\bar{k}_3 (\bar{k}_2 + \bar{k}_4)}. \end{aligned}$$

Kinetic analysis of the mutants revealed the following relationships of kinetic parameters:

$$\bar{k}_{\text{cat}} = \alpha k_{\text{cat}}, \quad (5.1)$$

$$\bar{K}_{\text{m}}^{\text{MMal-CoA}} = \beta K_{\text{m}}^{\text{MMal-CoA}}, \quad (5.2)$$

$$\bar{K}_{\text{m}}^{\text{ACP}} = \gamma K_{\text{m}}^{\text{ACP}}. \quad (5.3)$$

The following ratios of factors were found for mutants significantly changing kinetic parameters:

$$\text{R850K} : \quad \alpha = \beta \neq \gamma = 1, \quad (5.4)$$

$$\text{R850E} : \quad \alpha = \gamma \neq \beta, \quad (5.5)$$

$$\text{R850S} : \quad \alpha \neq \beta = \gamma. \quad (5.6)$$

Appendix

All three mutants showed significant changes in the turnover rate. This can be caused by changes in either k_2 , or in k_4 , or in both constants.

$$(5.1) \quad \Rightarrow \bar{k}_2 \bar{k}_4 = (\bar{k}_2 + \bar{k}_4) \alpha \frac{k_2 k_4}{k_2 + k_4}. \quad (5.7)$$

$$(5.8)$$

This equation can be solved for \bar{k}_4 or \bar{k}_2 , respectively:

$$\Rightarrow \bar{k}_4 = \frac{\alpha k_2 k_4 \bar{k}_2}{\bar{k}_2 (k_2 + k_4) - \alpha k_2 k_4}, \quad (5.9)$$

$$\bar{k}_2 = \frac{\alpha k_2 k_4 \bar{k}_4}{\bar{k}_4 (k_2 + k_4) - \alpha k_2 k_4}. \quad (5.10)$$

The mutations introduced can affect either the ping, or the pong, or both steps:

Scenario 1: Ping step is not affected ($\bar{k}_1 = k_1$, $\bar{k}_{-1} = k_{-1}$, $\bar{k}_2 = k_2$).

$$(5.9) \quad \Rightarrow \bar{k}_4 = \frac{\alpha k_2 k_4}{k_2 + (1 - \alpha) k_4}$$

$$\Rightarrow k_2 + \bar{k}_4 = \frac{[k_2 + (1 - \alpha) k_4] k_2 + \alpha k_2 k_4}{k_2 + (1 - \alpha) k_4} = \frac{k_2 (k_2 + k_4)}{k_2 + (1 - \alpha) k_4}. \quad (5.11)$$

$$\bar{k}_4 \text{ into (5.2)} \quad \frac{\bar{k}_4 (k_{-1} + k_2)}{k_1 (k_2 + \bar{k}_4)} = \beta \frac{k_4 (k_{-1} + k_2)}{k_1 (k_2 + k_4)} \quad (5.12)$$

$$\Rightarrow \frac{\alpha k_2 k_4}{k_2 (k_2 + k_4)} = \beta \frac{k_4}{k_2 + k_4} \Rightarrow \alpha = \beta. \quad (5.13)$$

For mutants R850E and R850S, this finding leads to a contradiction and, hence, it can be ruled out that these two mutations affect solely the pong step, leaving the ping step unaffected. For mutation R850K on the other hand, this scenario is consistent.

Scenario 2: Pong step is not affected ($\bar{k}_3 = k_3$, $\bar{k}_{-3} = k_{-3}$, $\bar{k}_4 = k_4$).

$$(5.10) \quad \Rightarrow \bar{k}_2 = \frac{\alpha k_2 k_4}{(1 - \alpha) k_2 + k_4}$$

$$\Rightarrow \bar{k}_2 + k_4 = \frac{[(1 - \alpha) k_2 + k_4] k_4 + \alpha k_2 k_4}{(1 - \alpha) k_2 + k_4} = \frac{k_4 (k_2 + k_4)}{(1 - \alpha) k_2 + k_4}. \quad (5.14)$$

$$\bar{k}_2 \text{ into (5.3)} \quad \frac{\bar{k}_2 (k_{-3} + k_4)}{k_3 (\bar{k}_2 + k_4)} = \gamma \frac{k_2 (k_{-3} + k_4)}{k_3 (k_2 + k_4)} \quad (5.15)$$

$$\Rightarrow \frac{\alpha k_2 k_4}{k_4 (k_2 + k_4)} = \gamma \frac{k_2}{k_2 + k_4} \Rightarrow \alpha = \gamma. \quad (5.16)$$

For mutants R850K and R850S, this finding leads to a contradiction and, hence, it can be ruled out that these two mutations affect solely the ping step, leaving the pong step unaffected. For mutation R850E on the other hand, this scenario is consistent.

In summary, the changes in kinetic parameters caused by mutant R850S cannot be explained by sole changes in either the ping or the pong step. This means that both steps are influenced by this mutation.

Scenario 3: Both steps are affected, exemplarily shown for mutant R850K.

$$\begin{aligned}
 (5.9) \quad \bar{k}_4 &= \frac{\alpha k_2 k_4 \bar{k}_2}{\bar{k}_2(k_2 + k_4) - \alpha k_2 k_4} \\
 \Rightarrow \bar{k}_2 + \bar{k}_4 &= \frac{[\bar{k}_2(k_2 + k_4) - \alpha k_2 k_4] \bar{k}_2 + \alpha k_2 k_4 \bar{k}_2}{\bar{k}_2(k_2 + k_4) - \alpha k_2 k_4} \\
 &= \frac{\bar{k}_2^2 (k_2 + k_4)}{\bar{k}_2(k_2 + k_4) - \alpha k_2 k_4} \tag{5.17}
 \end{aligned}$$

$$\Rightarrow \frac{\bar{k}_4}{\bar{k}_2 + \bar{k}_4} = \frac{\alpha k_2 k_4}{\bar{k}_2(k_2 + k_4)}. \tag{5.18}$$

$$(5.18) \text{ into } (5.2) \quad \frac{\alpha k_2 k_4 (\bar{k}_{-1} + \bar{k}_2)}{\bar{k}_1 \bar{k}_2 (k_2 + k_4)} = \beta \frac{k_4 (k_{-1} + k_2)}{k_1 (k_2 + k_4)}, \tag{5.19}$$

$$\text{with } \alpha = \beta \quad \Rightarrow \frac{k_2 (\bar{k}_{-1} + \bar{k}_2)}{\bar{k}_1 \bar{k}_2} = \frac{k_{-1} + k_2}{k_1} \tag{5.20}$$

$$\Rightarrow \frac{\bar{k}_{-1}}{\bar{k}_1 \bar{k}_2} + \frac{1}{\bar{k}_1} = \frac{k_{-1}}{k_1 k_2} + \frac{1}{k_1} \tag{5.21}$$

$$\Rightarrow \frac{1}{\bar{k}_1} \left(\frac{\bar{k}_{-1}}{\bar{k}_2} + 1 \right) = \frac{1}{k_1} \left(\frac{k_{-1}}{k_2} + 1 \right). \tag{5.22}$$

Changes of kinetic constants $k_i, i \in \{-1, 1, 2\}$ do not result in a contradiction. However, the complex and dependent change of constants does seem unplausible, especially considering that the changes can be explained by effects on solely the pong step. This can be done analogously for mutant R850E.

Overall, it is most plausible to assume that R850K and R850E solely take effect on the pong step and the ping step, respectively.

Now that we know the relationships between \bar{k}_2 and k_2 as well as between \bar{k}_4 and k_4 for the mutants R850K and R850E, it remains to investigate whether the other constants involved in the affected step (k_3 and k_{-3} in addition to k_4 for R850K, k_1 and k_{-1} in addition to k_2 for R850E) are influenced by the respective mutation. By a “proof of contradiction”, I will show that $\bar{k}_3 \neq k_3$ and $\bar{k}_{-3} \neq k_{-3}$ for R850K as well as $\bar{k}_1 \neq k_1$ and $\bar{k}_{-1} \neq k_{-1}$ for R850E.

R850K: Assume that solely k_4 changes and $\bar{k}_3 = k_3$ as well as $\bar{k}_{-3} = k_{-3}$.

$$(5.3) \quad \Rightarrow \frac{k_2 (k_{-3} + \bar{k}_4)}{k_3 (k_2 + \bar{k}_4)} = \gamma \frac{k_2 (k_{-3} + k_4)}{k_3 (k_2 + k_4)}. \quad (5.23)$$

$$\bar{k}_4 \text{ into (5.23)} \quad \frac{k_2 \left(k_{-3} + \frac{\alpha k_2 k_4}{k_2 + (1-\alpha)k_4} \right) [k_2 + (1-\alpha)k_4]}{k_2 k_3 (k_2 + k_4)} = \gamma \frac{k_2 (k_{-3} + k_4)}{k_3 (k_2 + k_4)} \quad (5.24)$$

$$\Rightarrow k_{-3} [k_2 + (1-\alpha)k_4] + \alpha k_2 k_4 = \gamma k_2 (k_{-3} + k_4) \quad (5.25)$$

$$\Rightarrow (1-\gamma) k_2 k_{-3} + (1-\alpha) k_{-3} k_4 = (\gamma-\alpha) k_2 k_4. \quad (5.26)$$

$$\text{with } \gamma = 1 \Rightarrow (1-\alpha) k_{-3} k_4 = (1-\alpha) k_2 k_4 \quad (5.27)$$

$$\Rightarrow k_2 = k_{-3}. \quad (5.28)$$

Although this is not a contradiction in a strict mathematical sense, it represents a very specific situation. From a chemical point of view, there is no reason why this relation should hold as k_2 and k_{-3} are independent constants. This leads me to the aforementioned conclusion that for the mutation R850K most likely $\bar{k}_3 \neq k_3$ and $\bar{k}_{-3} \neq k_{-3}$ in addition to $\bar{k}_4 \neq k_4$.

R850E: Assuming that $\bar{k}_1 = k_1$ and $\bar{k}_{-1} = k_{-1}$ using the above relation (5.2) gives the corresponding equation to (5.26):

$$(1-\beta) k_4 k_{-1} + (1-\alpha) k_{-1} k_2 = (\beta-\alpha) k_2 k_4. \quad (5.29)$$

Since $\beta \neq 1$, this cannot be further simplified. However, the kinetic constant k_{-1} is dependent on k_2 and k_4 :

$$\Rightarrow k_{-1} = \frac{(\beta-\alpha) k_2 k_4}{(1-\alpha) k_2 + (1-\beta) k_4}. \quad (5.30)$$

As discussed above, from a mathematical point of view it cannot be ruled out that the kinetic constants k_1 and k_{-1} remain unchanged. However, it is unplausible that the kinetic constants are dependent and, hence, I assume these two change ($\bar{k}_1 \neq k_1$ and $\bar{k}_{-1} \neq k_{-1}$) in addition to k_2 ($\bar{k}_2 \neq k_2$).

Supplementary Tables

Table S1: Average expression yields (in mg) of FAS and PKS AT domains after tandem affinity chromatography of 1 L expression culture. Yields determined from three independent expressions.

Protein	Yield (mg)
DEBS3 (KS ⁰ -AT)5	14
PikAIII (KS ⁰ -AT)5	14
RAPS3 (KS ⁰ -AT)14	3.5
Pks7 KS ⁰ -AT	22
DEBS1 AT0-BMP	7.6
AVES1 AT0	12
FabD	12

Table S2: TSA results of FAS and PKS AT and ACP domains. Average melting temperatures (in °C) of biological triplicates (ATs) and technical triplicates (ACPs). n. d.: not determinable.

Protein	AT		ACP	
	Storage Buffer	Assay Buffer	Storage Buffer	Assay Buffer
DEBS3M5	51	42	64	65
PikAIIIM5	49	42	59	56
RAPS3M14	45	36	57	50
Pks7	49	41	62	58
DEBS1M0	48	42	63	59
AVES1M0	55	49	53	55
<i>E. coli</i>	65	61	n. d.	n. d.

Table S3: Average expression yields (in mg) of non-codon optimized and codon optimized (COP) *holo*-ACP domains after SEC of 1 L expression culture. Yields determined from four to nine independent expressions.

Protein	Yield (mg)
DEBS3 ACP5	23
DEBS3 ACP5 COP	87
PikAIII ACP5	4.5
PikAIII ACP5 COP	51
RAPS3 ACP14	31
Pks7 ACP	1.0
Pks7 ACP COP	30
DEBS1 ACP0	9.4
AVES1 ACP0	11
<i>E. coli</i> ACP COP	29

Table S4: MS data of FAS and PKS ACP domains. *m/z* expected (exp) and found experimentally of *apo*- and *holo*-ACP with and without N-terminal Methionine (Met). All ACPs are fully activated (see Fig.S5).

Protein	<i>m/z</i> (-Met)		<i>m/z</i> (+Met)		State of Charge				
	<i>apo</i>		<i>holo</i>						
	exp	found	exp	found					
DEBS3 ACP5	836.21	–	855.15	855.16	843.49	–	862.43	–	18
DEBS3 ACP5 COP	836.21	836.19	855.15	855.14	843.49	–	862.43	–	18
PiKAIII ACP5 COP	822.92	–	840.82	840.81	829.20	–	848.15	–	18
RAPS3 ACP14	767.88	–	786.78	786.78	775.16	–	794.06	–	18
Pks7 ACP COP	891.29	–	910.18	910.20	898.57	–	917.46	–	18
<i>E. coli</i> ACP COP	976.98	–	1010.99	1011.03	990.09	–	1024.20	1024.18	10
DEBS1M0 ACP	746.61	–	765.55	765.56	753.89	–	772.83	–	18
AVES1M0 ACP	974.38	–	993.27	993.28	981.66	–	1000.55	–	18

Table S5: Transacylation and hydrolysis rates (in s^{-1}) of initial substrate screening (50 μ M X-CoA and ACP) of FAS and PKS AT domains. Transacylation measured in biological triplicates. Hydrolysis measured in technical triplicates of each biological triplicate. ^[a] mMAT from previous study, referring to k_{cat} for hydrolysis (saturated X-CoA) and $k_{cat, app}$ for transacylation (saturated X-CoA, 60 μ M ACP).^[9]

Protein	X-CoA	Transacylation $k_{0, app}$ (s^{-1})	Hydrolysis k_0 (s^{-1})
DEBS3 AT5	MMal	1.92×10^{-1}	4.24×10^{-2}
PikAIII AT5	Mal	8.86×10^{-2}	4.98×10^{-2}
PikAIII AT5	MMal	8.60×10^{-1}	1.54×10^{-2}
RAPS3 AT14	Mal	2.05×10^{-1}	6.63×10^{-2}
Pks7 AT	Ac	2.81×10^{-1}	1.60×10^{-3}
Pks7 AT	Mal	2.79	7.05×10^{-2}
Pks7 AT	MMal	1.69	5.58×10^{-2}
DEBS1 AT0	Prop	7.59×10^{-2}	3.11×10^{-2}
DEBS1 AT0	Ac	3.67×10^{-2}	3.00×10^{-2}
AVES1 AT0	2-MB	5.49×10^{-1}	4.89×10^{-2}
AVES1 AT0	Prop	-4.03×10^{-3}	2.72×10^{-2}
mMAT ^[a]	Ac	13.7	9.3×10^{-3}
mMAT ^[a]	Mal	15.6	9.8×10^{-3}
mMAT ^[a]	MMal	4.0	6.0×10^{-3}
FabD	Ac	3.19	6.77×10^{-3}
FabD	Mal	93.9	5.00×10^{-3}
FabD	MMal	72.4	3.71×10^{-4}

Table S6: Kinetic parameters determined for AT-mediated hydrolysis. Turnover measured in technical triplicates of biological triplicates (DEBS3 AT5), technical triplicates of one replicate (PikAIII AT5, RAPS3 AT14), or biological triplicates (Pks7 with Mal and MMal). ^[a] Value experimentally determined in hydrolysis measurement. ^[b] Value calculated from experimentally determined hydrolysis and transacylation values. ^[c] Value experimentally determined in transacylation measurement. ^[d] mMAT from previous study.^[9]

Protein X-CoA	k_{cat} (s^{-1})^[a]	$K_{\text{m}}^{\text{X-CoA}}$ (μM)^[b]	$k_{\text{cat}}/K_{\text{m}}^{\text{X-CoA}}$ ($\text{s}^{-1}\text{M}^{-1}$)^[c]
DEBS3 AT5 MMal	$5.31 \times 10^{-2} \pm 9.62 \times 10^{-4}$	8.85	6.0×10^3
PikAIII AT5 MMal	$1.94 \times 10^{-2} \pm 5.03 \times 10^{-4}$	2.49×10^{-1}	7.8×10^4
RAPS3 AT14 Mal	$6.51 \times 10^{-2} \pm 1.89 \times 10^{-3}$	1.36	4.8×10^4
Pks7 AT Mal	$9.93 \times 10^{-2} \pm 3.17 \times 10^{-3}$	3.97×10^{-2}	2.5×10^6
Pks7 AT MMal	$7.34 \times 10^{-2} \pm 1.51 \times 10^{-3}$	7.57×10^{-2}	9.7×10^5
mMAT Mal ^[d]	$(9.8 \pm 1.7) \times 10^{-3}$	7.0×10^{-4}	1.4×10^7
mMAT Ac ^[d]	$9.3 \times 10^{-3} \pm 1.2 \times 10^{-4}$	1.1×10^{-3}	8.3×10^6

Table S7: Average expression yields (in mg) of DEBS3 (KS⁰-AT)₅ interface mutants after tandem affinity chromatography of 1 L expression culture. Data from three independent protein expressions, except proteins marked with * are single expressions.

Mutation	Yield (mg)
A539S*	10
A539D*	15
A539E	9.7
A539F*	1.7
R850K	14
R850A*	5.7
R850E	7.7
R850F*	2.4
R850S	9.6

Table S8: TSA results of DEBS3 (KS⁰-AT)₅ interface mutants. Average melting temperatures (in °C) of technical/biological triplicates. DEBS3M5 ACP COP was used in the mutagenesis study.

Mutation	AT		ACP
	Storage Buffer	Storage Buffer	Assay Buffer
–	51	63	62
A539S	48	–	–
A539D	48	–	–
A539E	48	–	–
A539F	48	–	–
R850K	48	–	–
R850A	48	–	–
R850E	47	–	–
R850F	48	–	–
R850S	48	–	–

Table S9: MS data of mFAS and DEBS3M6 ACPs. *m/z* expected (*exp*) and found experimentally of *apo*- and *holo*-ACP with and without N-terminal Methionine (Met). All ACPs are fully activated (see Fig. S19).

Protein	<i>m/z</i> (-Met)				<i>m/z</i> (+Met)				State of Charge
	<i>apo</i>		<i>holo</i>		<i>apo</i>		<i>holo</i>		
	<i>exp</i>	<i>found</i>	<i>exp</i>	<i>found</i>	<i>exp</i>	<i>found</i>	<i>exp</i>	<i>found</i>	
mFAS ACP	667.12	–	686.02	686.01	674.40	–	693.30	–	18
DEBS3 ACP6	683.49	–	657.44	657.44	645.77	–	664.72	–	18

Table S10: Parameters of transacylation of MMal-CoA onto DEBS3M6 ACP mediated by DEBS3M6 AT chimera (FS118) and wild type (FS138). Parameters determined for the three replicates (measured in technical triplicates), average of parameters, and parameters determined for the average values of biological triplicates.

Protein	k_{cat} (s^{-1})	K_{m} (μM)	$K_{\text{m}}^{\text{ACP}}$ (μM)	$k_{\text{cat}}/K_{\text{m}}^{\text{ACP}}$ ($\text{M}^{-1}\text{s}^{-1}$)	$k_{\text{cat}}/K_{\text{m}}^{\text{MMal-CoA}}$ ($\text{M}^{-1}\text{s}^{-1}$)	$k_{\text{cat}}/K_{\text{m}}^{\text{ACP}}$ ($\text{M}^{-1}\text{s}^{-1}$)
FS118-1	$1.08 \times 10^{-1} \pm 7.84 \times 10^{-3}$	8.27 ± 1.47	95.2 ± 19.1	1.3×10^4	1.3×10^4	1.1×10^3
FS118-2	$1.15 \times 10^{-1} \pm 8.02 \times 10^{-3}$	9.67 ± 1.56	101 ± 18.8	1.2×10^4	1.2×10^4	1.1×10^3
FS118-3	$1.20 \times 10^{-1} \pm 8.40 \times 10^{-3}$	9.65 ± 1.53	112 ± 19.9	1.2×10^4	1.2×10^4	1.1×10^3
average	$1.14 \times 10^{-1} \pm 8.09 \times 10^{-3}$	9.20 ± 1.52	103 ± 19.3	1.2×10^4	1.2×10^4	1.1×10^3
FS118	$1.13 \times 10^{-1} \pm 6.93 \times 10^{-3}$	8.86 ± 1.28	103 ± 16.6	1.3×10^4	1.3×10^4	1.1×10^3
FS138-1	$1.55 \times 10^{-1} \pm 2.10 \times 10^{-2}$	24.3 ± 4.88	226 ± 59.7	6.4×10^3	6.4×10^3	6.8×10^2
FS138-2	$1.85 \times 10^{-1} \pm 2.84 \times 10^{-2}$	26.9 ± 5.57	324 ± 84.9	6.9×10^3	6.9×10^3	5.7×10^2
FS138-3	$1.57 \times 10^{-1} \pm 2.23 \times 10^{-2}$	23.2 ± 4.97	235 ± 65.0	6.7×10^3	6.7×10^3	6.7×10^2
average	$1.65 \times 10^{-1} \pm 2.40 \times 10^{-2}$	24.8 ± 5.14	262 ± 69.9	6.7×10^3	6.7×10^3	6.3×10^2
FS138	$1.64 \times 10^{-1} \pm 2.28 \times 10^{-2}$	24.6 ± 4.91	257 ± 66.2	6.7×10^3	6.7×10^3	6.4×10^2

Table S11: NADH calibration factor over time determined via internal standard measurements for AT activity assay in 96-well format.

Time Span	Calibration Factor (RFU/μM)
22/02/2018–27/03/2018	13450
07/11/2018–05/12/2018	11950
07/08/2019–27/09/2019	11064
31/10/2019–18/11/2019	9058
03/03/2020–06/03/2020	8596

Supplementary Figures

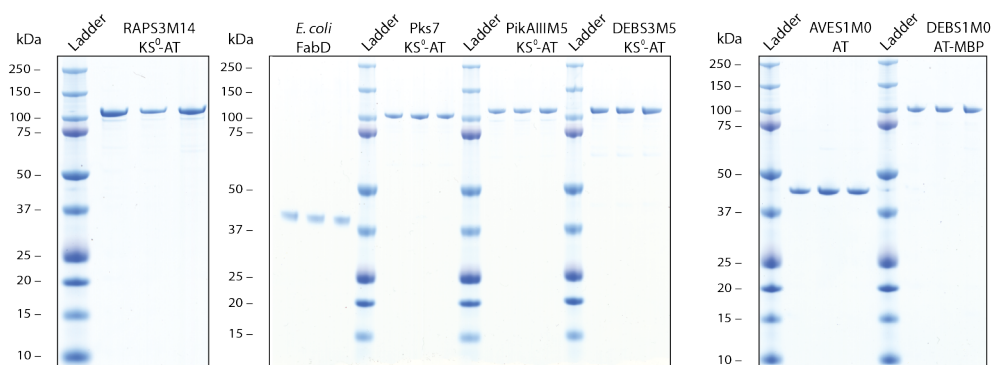


Figure S1: Analytical SDS-PAGEs (NuPAGE Bis-Tris 4-12%) of FAS and PKS AT constructs expressed in biological triplicates. All AT constructs are highly pure after tandem affinity chromatography. Protein bands migrate at expected masses: RAPS3M14 KS⁰-AT – 101 kDa, *E. coli* FabD – 36.8 kDa, Pks7 KS⁰-AT – 98.6 kDa, PikAIIIM5 KS⁰-AT – 101 kDa, DEBS3M5 KS⁰-AT – 99.9 kDa, AVES1M0 AT – 42.8 kDa, DEBS1M0 AT-MBP – 94.1 kDa.

Appendix

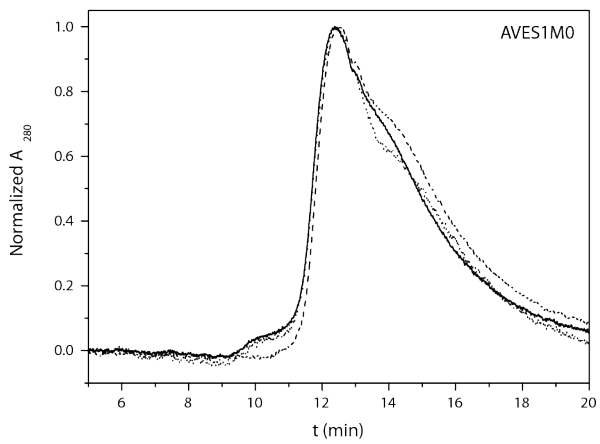
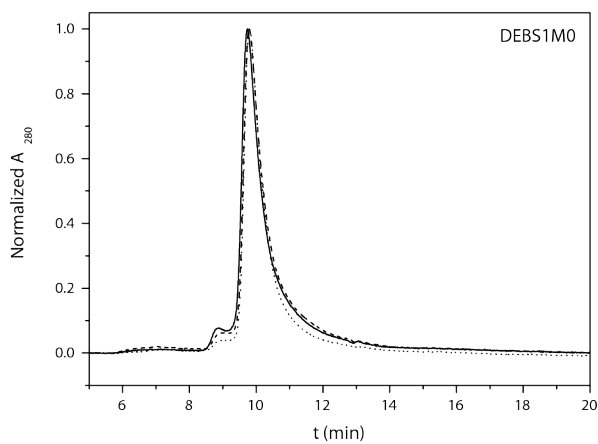
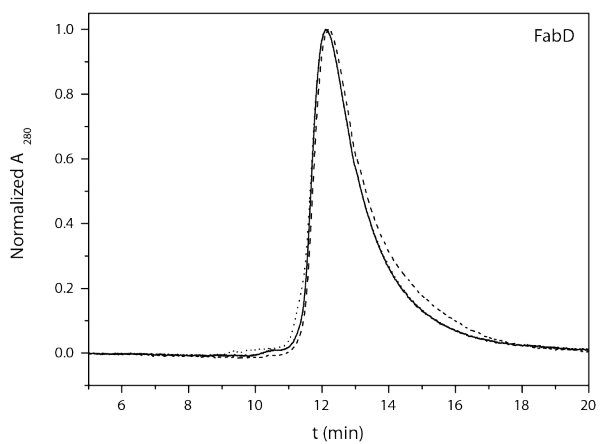
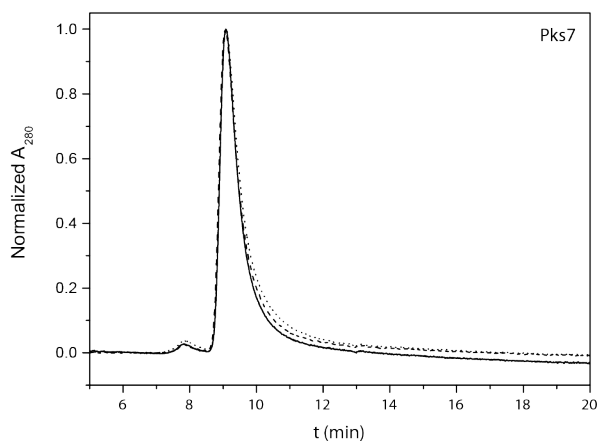
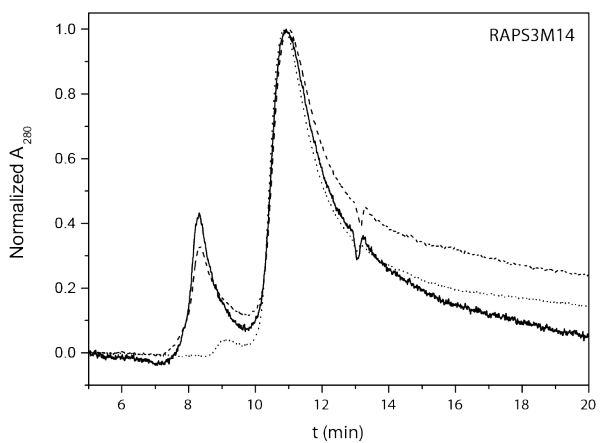
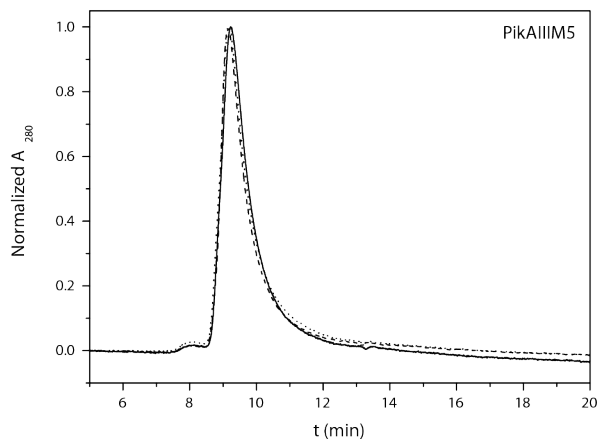
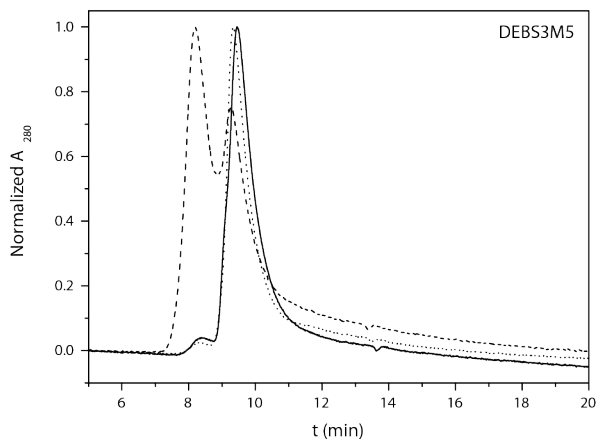


Figure S2: Normalized size exclusion chromatograms of FAS and PKS AT constructs. Each curve (solid, dotted, dashed) corresponds to one biological replicate. DEBS3M5 KS⁰-AT shows dimeric (9.4 min) and tetrameric (8.2 min) oligomers. PikAIIIM5 KS⁰-AT forms solely dimers (9.2 min). RAPS3M14 KS⁰-AT shows monomeric (11 min) and tetrameric (8.3 min) species. We note that the prevalence for a monomeric state should not impact AT activity, since the dimeric interface is formed by the KS domain, while the AT domain is monomeric in its active form. Pks7 KS⁰-AT forms solely dimers (9.1 min). *E. coli* FabD, DEBS1M0 AT-MBP, and AVES1M0 ATs are solely monomeric (12.1 min, 9.8 min, and 12.5 min).

Appendix

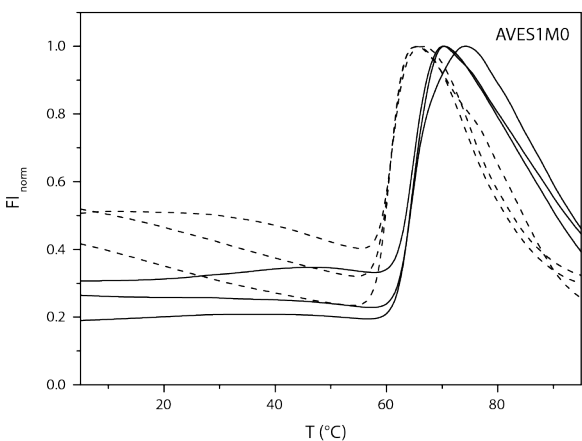
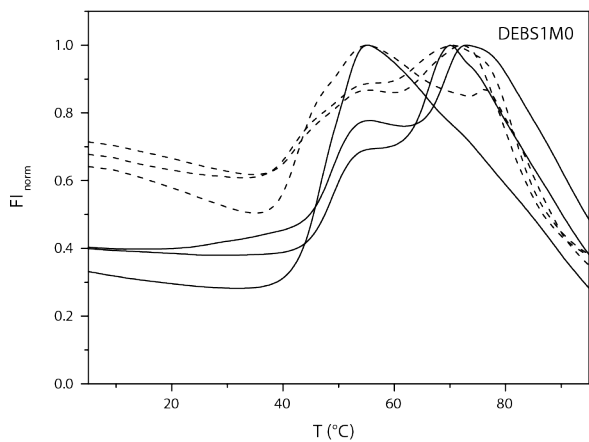
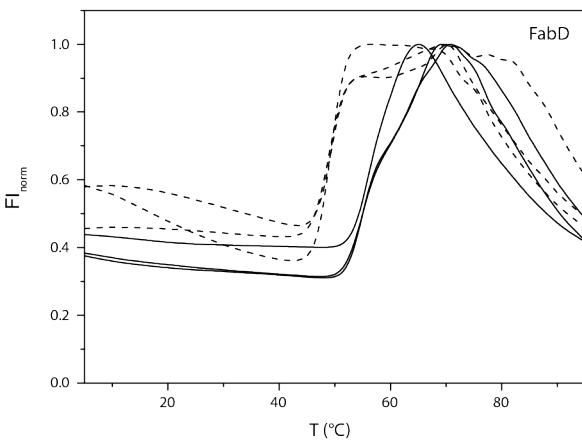
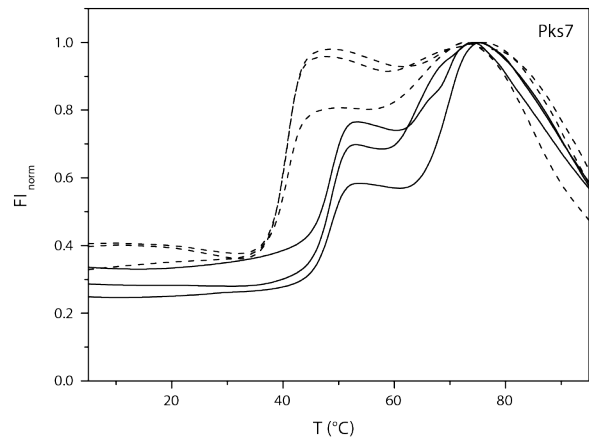
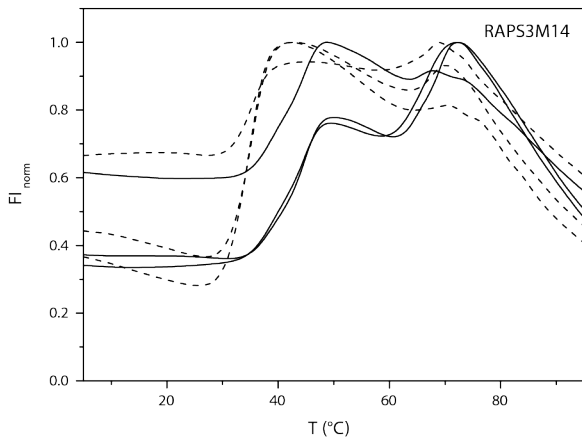
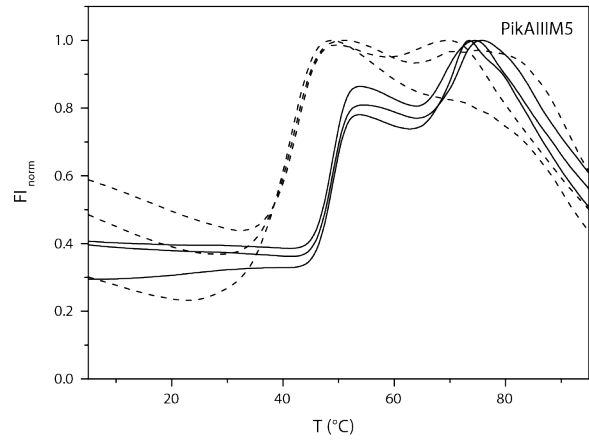
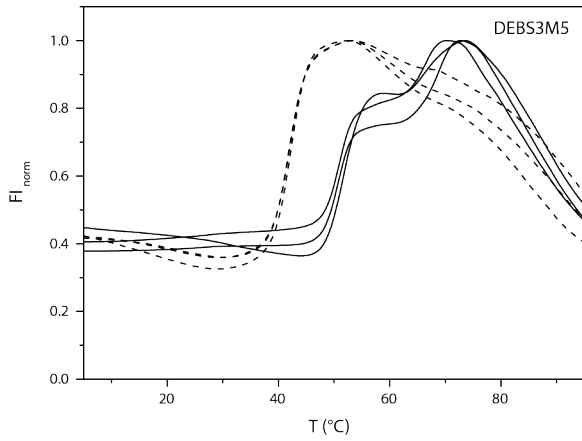


Figure S3: Normalized TSA melting curves of FAS and PKS (KS⁰-)AT domains. Solid and dashed line display data collected in storage and assay buffer, respectively. Each curve (solid, dashed) corresponds to one biological replicate.

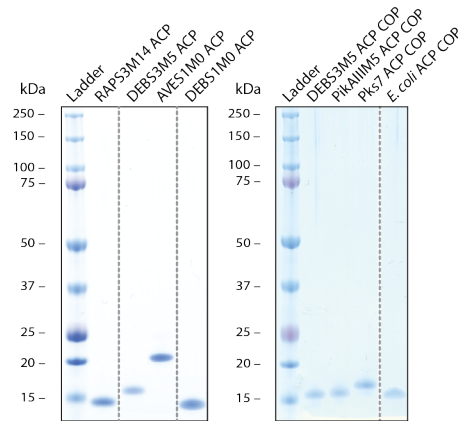
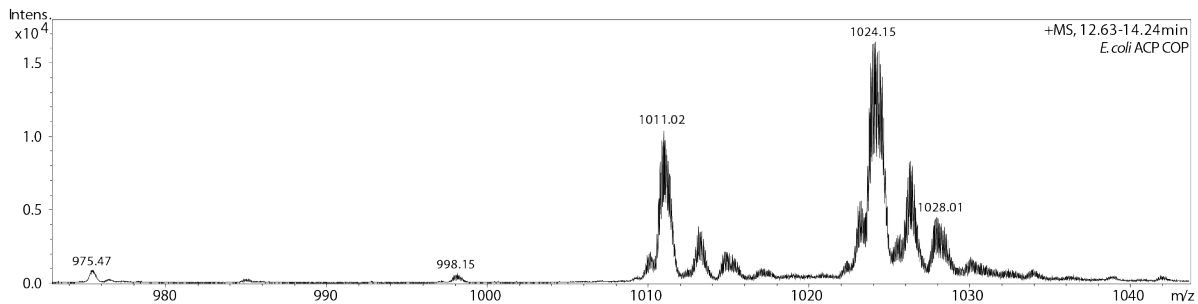
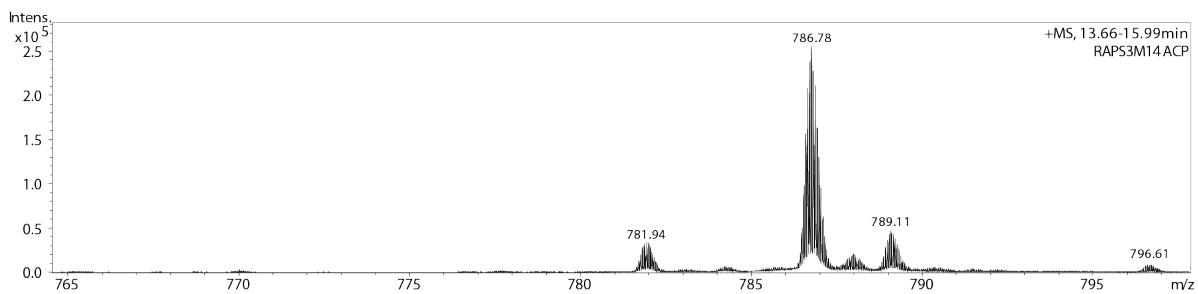
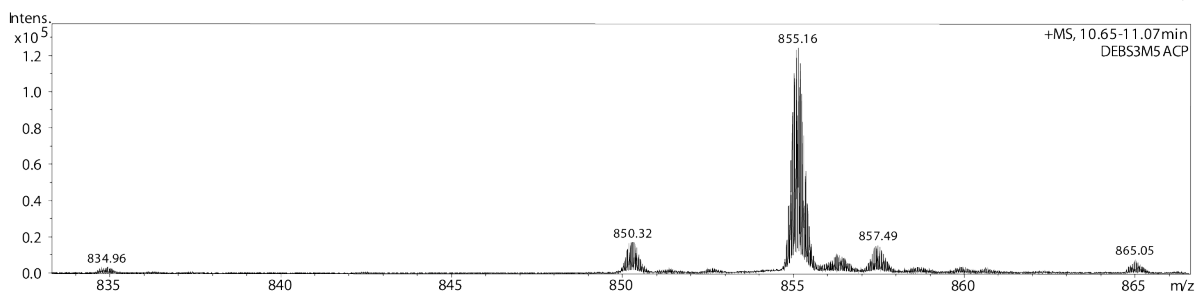
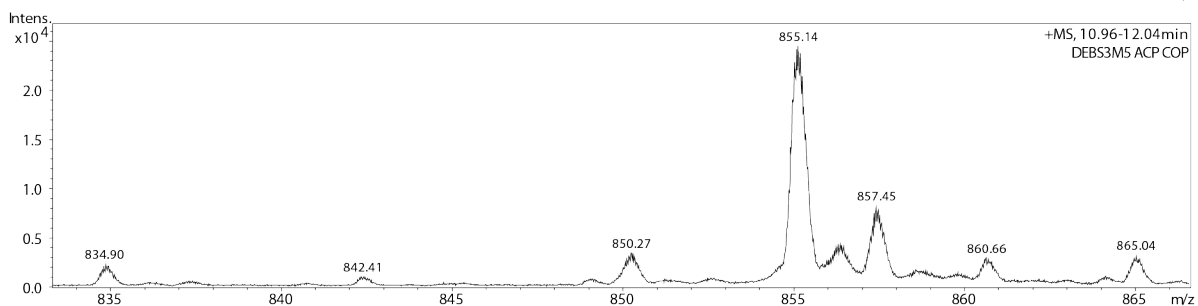
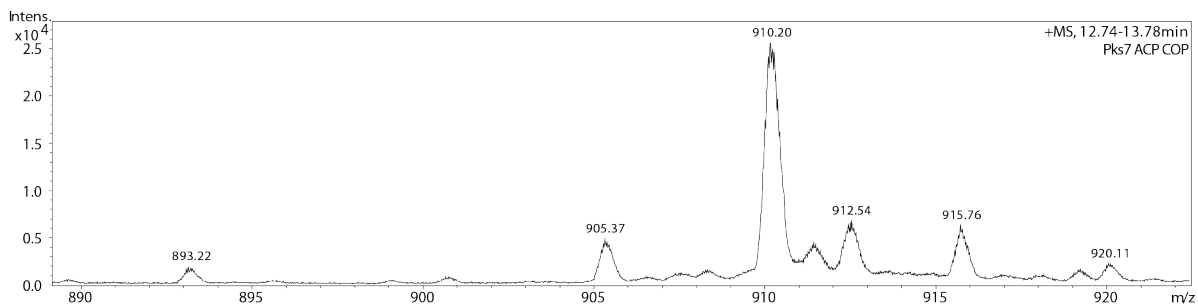
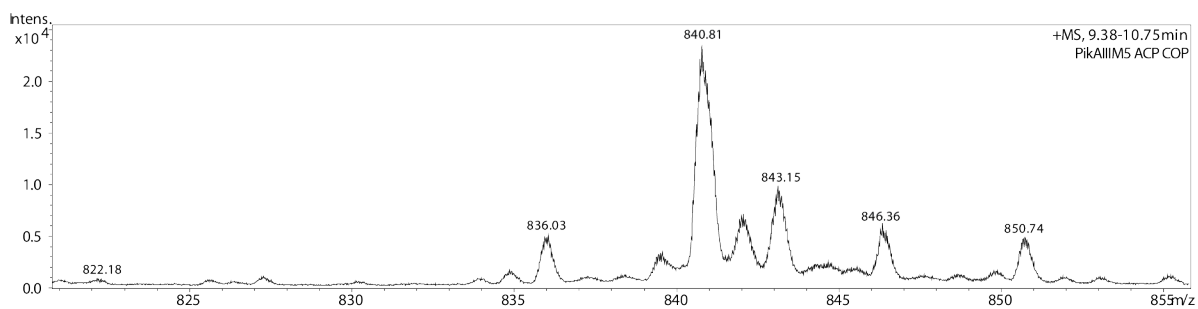


Figure S4: Analytical SDS-PAGEs (NuPAGE Bis-Tris 4-12%) of FAS and PKS ACP domains. All pooled ACPs are highly pure after SEC. Masses are as expected: RAPS3M14 ACP – 14.1 kDa, DEBS3M5 ACP (COP) – 15.4 kDa, PikAIIIM5 ACP COP – 15.2 kDa, Pks7 ACP COP – 16.4 kDa, AVES1M0 ACP – 17.7 kDa, DEBS1M0 ACP – 13.6 kDa. *E. coli* ACP is known to run in gel higher than expected from its mass of 10.2 kDa.

Appendix



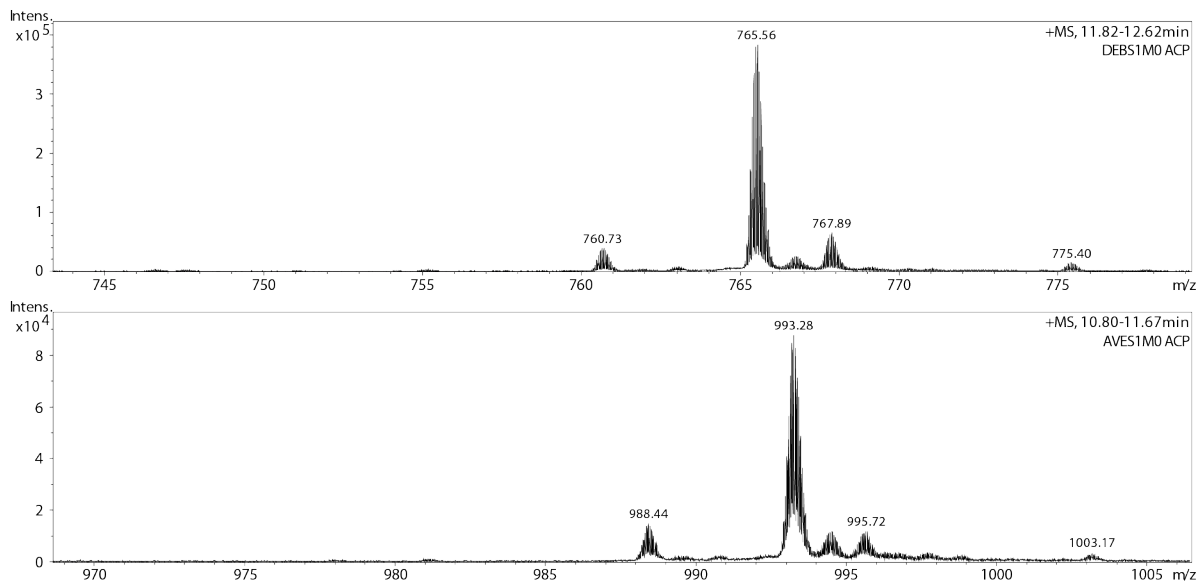


Figure S5: MS analysis of FAS and PKS ACP domains after SEC. Shown are the m/z -values of the average protein masses of the 10^+ and 18^+ charge states (MS^1) for *E. coli* and all other proteins, respectively. In none of the samples, m/z -values of the *apo* species were recorded. Main m/z -values correspond to the *holo* form. *E. coli* ACP shows two species: one with and one without N-terminal methionine. All other proteins do not include the N-terminal methionine. COP stands for codon optimized. **PikAIIIM5 ACP COP**, $z = 18$: $m/z = 840.81$ corresponds to *holo*-ACP without N-terminal methionine. $m/z = 843.15$ corresponds to acetylation ($\Delta m = 42$ Da), $m/z = 850.74$ corresponds to α -N-gluconozylation ($\Delta m = 178$ Da). $m/z = 846.36$ is an unknown modified variant ($\Delta m = 100$ Da). $m/z = 836.03$ corresponds to *holo*-ACP without N-terminal methionine and serine ($\Delta m = -86$ Da). **Pks7 ACP COP**, $z = 18$: $m/z = 910.20$ corresponds to *holo*-ACP without N-terminal methionine. $m/z = 912.54$ corresponds to acetylation ($\Delta m = 42$ Da), $m/z = 920.11$ corresponds to α -N-gluconozylation ($\Delta m = 178$ Da). $m/z = 915.76$ is an unknown modified variant ($\Delta m = 100$ Da). $m/z = 905.37$ corresponds to *holo*-ACP without N-terminal methionine and serine ($\Delta m = -87$ Da). **DEBS3M5 ACP COP**, $z = 18$, used for kinetic analysis of mutants: $m/z = 855.14$ corresponds to *holo*-ACP without N-terminal methionine. $m/z = 857.45$ corresponds to acetylation ($\Delta m = 42$ Da), $m/z = 865.04$ corresponds to α -N-gluconozylation ($\Delta m = 178$ Da). $m/z = 860.66$ is an unknown modified variant ($\Delta m = 99$ Da). $m/z = 850.27$ corresponds to *holo*-ACP without N-terminal methionine and serine ($\Delta m = -88$ Da). **DEBS3M5 ACP**, $z = 18$, used for kinetic analysis of the wild type: $m/z = 855.16$ corresponds to *holo*-ACP without N-terminal methionine. $m/z = 857.49$ corresponds to acetylation ($\Delta m = 42$ Da), $m/z = 865.05$ corresponds to α -N-gluconozylation ($\Delta m = 178$ Da). $m/z = 860.66$ is an unknown modified variant ($\Delta m = 99$ Da). $m/z = 850.32$ corresponds to *holo*-ACP without N-terminal methionine and serine ($\Delta m = -87$ Da). **RAPS3M14 ACP**, $z = 18$: $m/z = 786.78$ corresponds to *holo*-ACP without N-terminal methionine. $m/z = 789.11$ corresponds to acetylation ($\Delta m = 42$ Da), $m/z = 796.61$ corresponds to α -N-gluconozylation ($\Delta m = 177$ Da). $m/z = 781.94$ corresponds to *holo*-ACP without N-terminal methionine and serine ($\Delta m = -87$ Da). ***E. coli* ACP COP**, $z = 10$: $m/z = 1011.02$ corresponds to *holo*-ACP including the N-terminal methionine. $m/z = 1015.28$ corresponds to the acetylated form ($\Delta m = 43$ Da). $m/z = 1024.15$ corresponds to *holo*-ACP without N-terminal methionine. $m/z = 1028.01$ corresponds to acetylation ($\Delta m = 39$ Da), $m/z = 1042.00$ corresponds to α -N-gluconozylation ($\Delta m = 178$ Da). **DEBS1M0 ACP**, $z = 18$: $m/z = 765.56$ corresponds to *holo*-ACP without N-terminal methionine. $m/z = 767.89$ corresponds to acetylation ($\Delta m = 42$ Da), $m/z = 775.40$

corresponds to α -N-gluconozylation ($\Delta m = 177$ Da). $m/z = 760.73$ corresponds to *holo*-ACP without N-terminal methionine and serine ($\Delta m = -87$ Da). **AVES1M0 ACP**, $z = 18$: $m/z = 993.28$ corresponds to *holo*-ACP without N-terminal methionine. $m/z = 995.72$ corresponds to acetylation ($\Delta m = 44$ Da), $m/z = 1003.17$ corresponds to α -N-gluconozylation ($\Delta m = 178$ Da). $m/z = 988.44$ corresponds to *holo*-ACP without N-terminal methionine and serine ($\Delta m = -87$ Da).

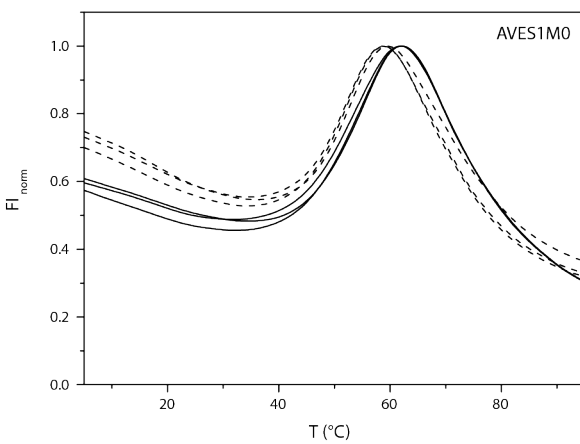
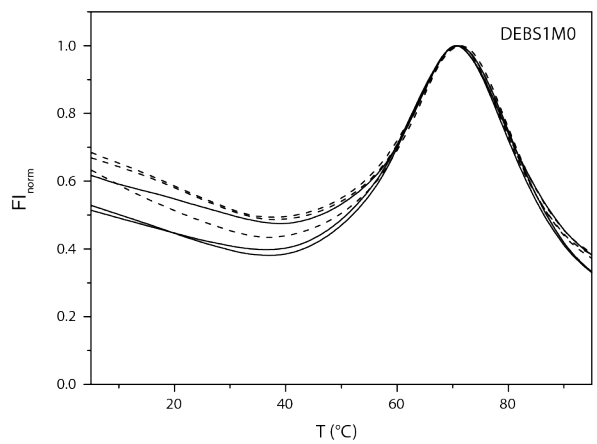
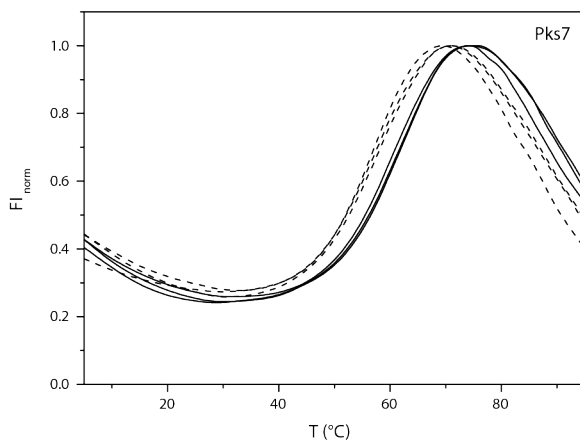
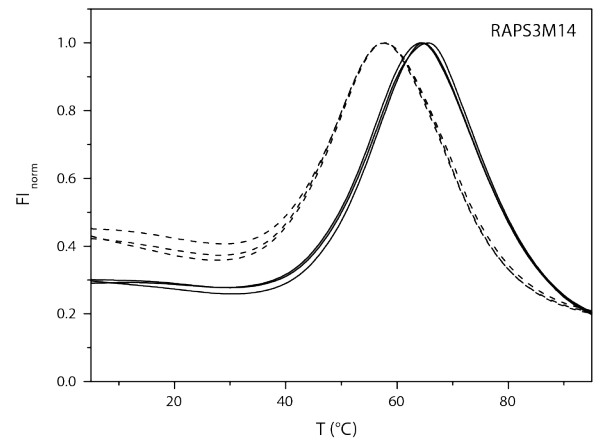
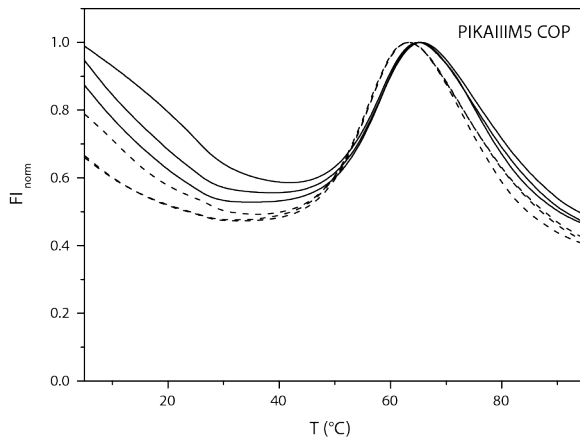
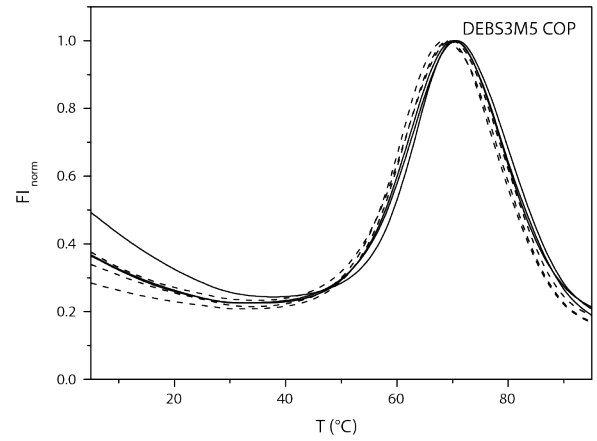
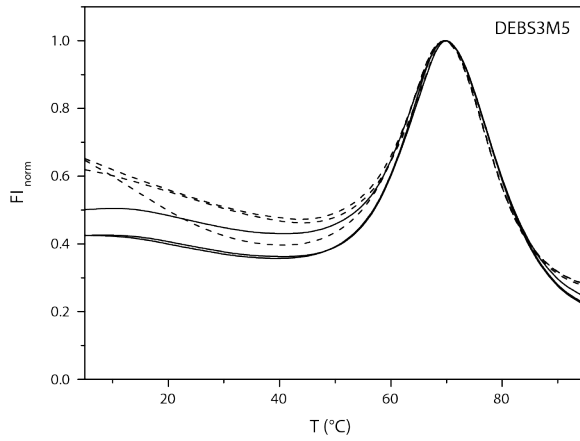
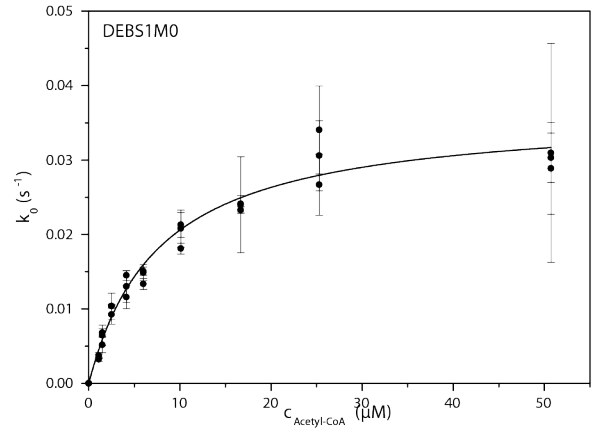
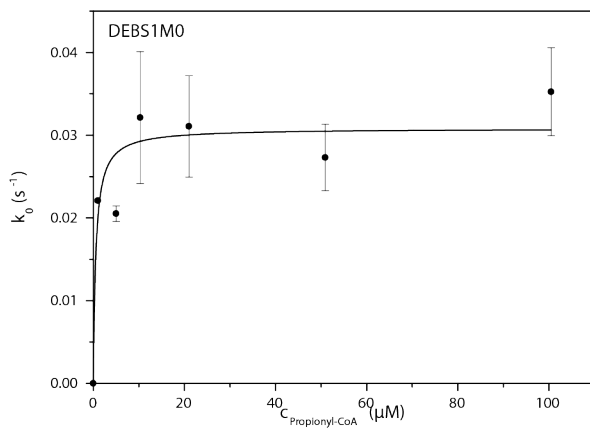
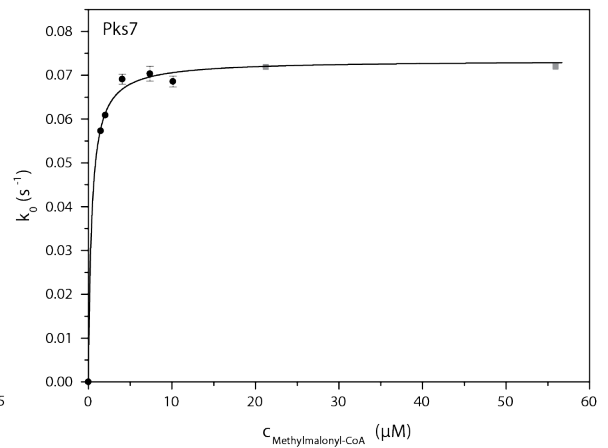
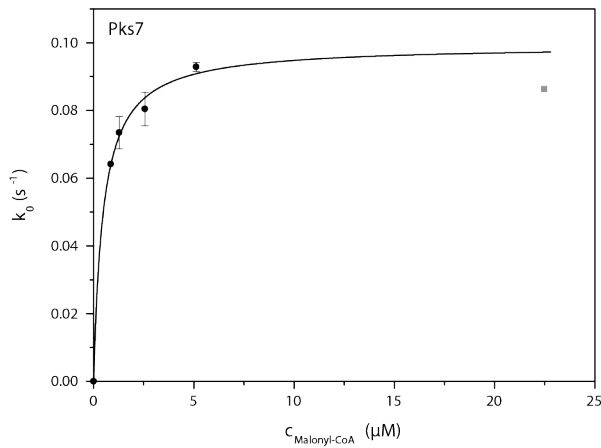
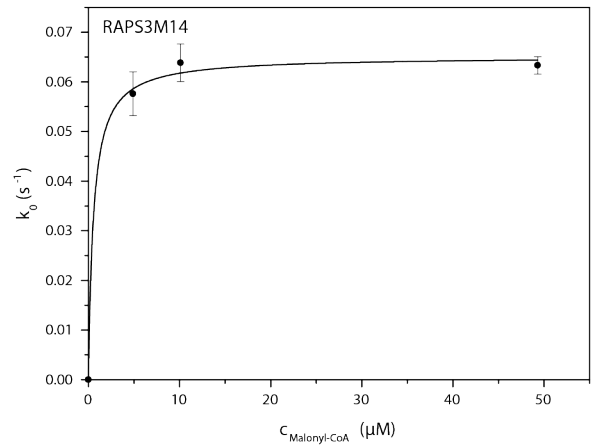
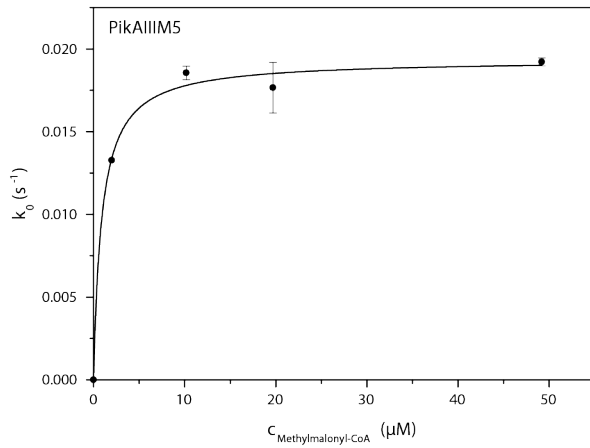
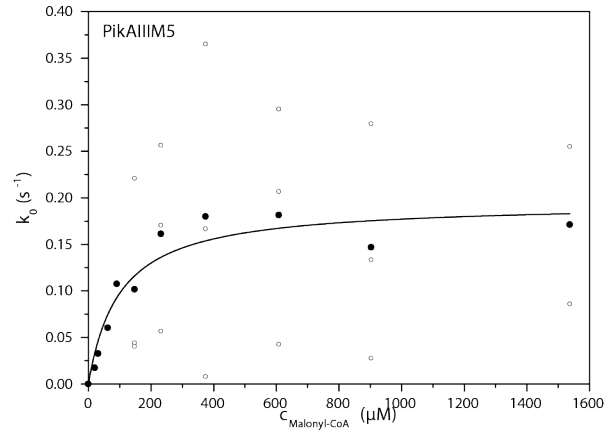
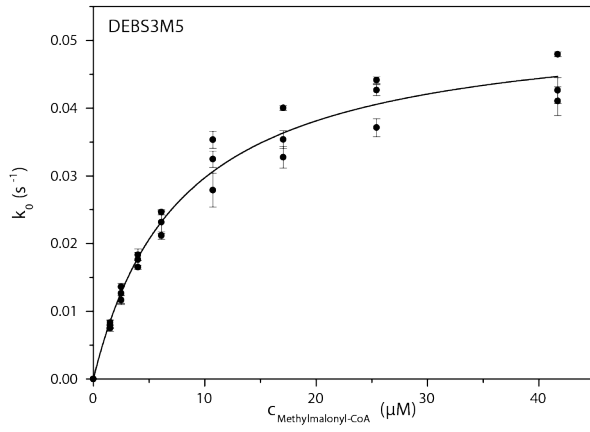


Figure S6: Normalized TSA melting curves of FAS and PKS ACP domains. Solid and dashed line display data collected in storage and assay buffer, respectively. Each curve (solid, dashed) corresponds to one technical replicate. COP stands for codon optimized.



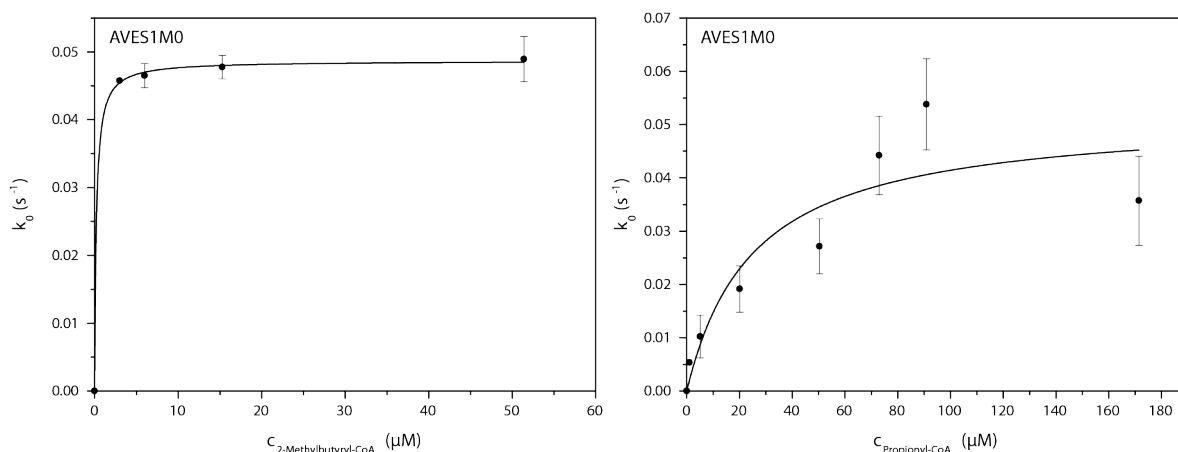


Figure S7: Titration curves for AT-mediated hydrolysis of X-CoA. **DEBS3M5, MMal:** DEBS3M5 KS^0 -AT measured in technical triplicates of biological triplicates. Black dots with error bars show the average of one biological replicate with standard deviation. Kinetic parameters were determined precisely. Substrate consumption was below 10%. MMal range was $0.19\text{--}5.3 \times K_m^{\text{MMal-CoA}}$. **PikAIIIM5, Mal:** PikAIIIM5 KS^0 -AT measured in biological triplicates. Circles show one biological replicate, black dots show their average. Problems with high Mal concentrations were observed. Kinetic parameters with substantial errors were calculated. Substrate consumption was below 5%. **PikAIIIM5, MMal:** PikAIIIM5 KS^0 -AT measured in technical triplicates. Black dots with error bars show the average with standard deviation. The lowest concentration is a one well measurement. Due to the low substrate concentrations, the maximum velocity and an upper limit of the Michaelis-Menten constant were determined for this system. Substrate consumption was up to 17%. **RAPS3M14, Mal:** RAPS3M14 KS^0 -AT measured in technical triplicates. Black dots with error bars show the average with standard deviation. Due to the low substrate concentration, the maximum velocity and an upper limit of the Michaelis-Menten constant were determined. Substrate consumption was up to 24%. **Pks7, Mal:** Pks7 KS^0 -AT measured in biological triplicates. Black dots with error bars show the average with standard deviation. The lowest concentration is a one well measurement. The gray box is a single well measurement with high substrate concentration, which was not used for the hydrolysis fit. It shows that maximum velocity was reached. Due to the low substrate concentrations, the maximum velocity and an upper limit of the Michaelis-Menten constant were determined. Substrate consumption was up to 44%. Signal increase was still constant. **Pks7, MMal:** Pks7 KS^0 -AT measured in biological triplicates. Black dots with error bars show the average with standard deviation. The two lowest concentrations are one well measurements. Grey boxes are single well measurements with high substrate concentrations, which were not used for the hydrolysis fit. These values show that maximum velocity was reached. Due to the low substrate concentration, the maximum velocity and an upper limit of the Michaelis-Menten constant were determined. Substrate consumption was up to 34%. Signal increase was still constant. **DEBS1M0, Prop:** DEBS1M0 AT-MBP measured in technical triplicates. Black dots with error bars show the average with standard deviation. The lowest concentration is a one well measurement. Due to the low substrate concentrations, the maximum velocity and an upper limit of the Michaelis-Menten constant were determined for this system. Substrate consumption was up to 20%. **DEBS1M0, Ac:** DEBS1M0 AT-MBP measured in technical triplicates of biological triplicates. Black dots with error bars show the average of one biological replicate with standard deviation. Kinetic parameters were determined precisely. Substrate consumption was up to 11%. Ac range was $0.14\text{--}5.4 \times K_m^{\text{Ac-CoA}}$. **AVES1M0, 2-MB:** AVES1M0 AT measured in technical triplicates. Black dots with error bars show the average with

standard deviation. The lowest concentration is a one well measurement. Due to the low substrate concentrations, the maximum velocity and an upper limit of the Michaelis-Menten constant were determined for this system. Substrate consumption was up to 14 %. **AVES1M0, Prop:** AVES1M0 AT measured in technical triplicates. Black dots with error bars show the average with standard deviation. The lowest concentration is a one well measurement. Due to large signal fluctuations kinetic parameters with substantial errors were calculated. Substrate consumption was up to 16 %. Prop range was $0.040-6.8 \times K_m^{\text{Prop-CoA}}$.

Appendix

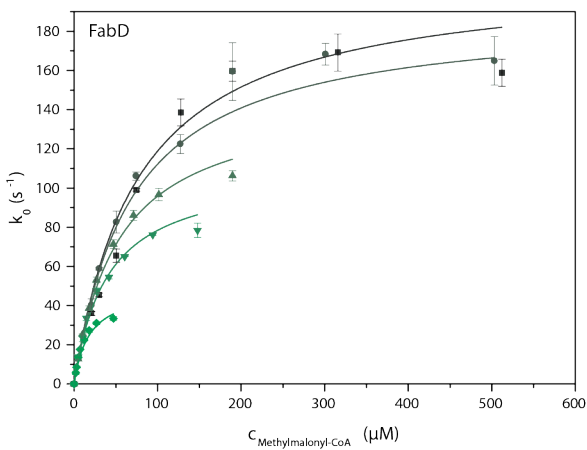
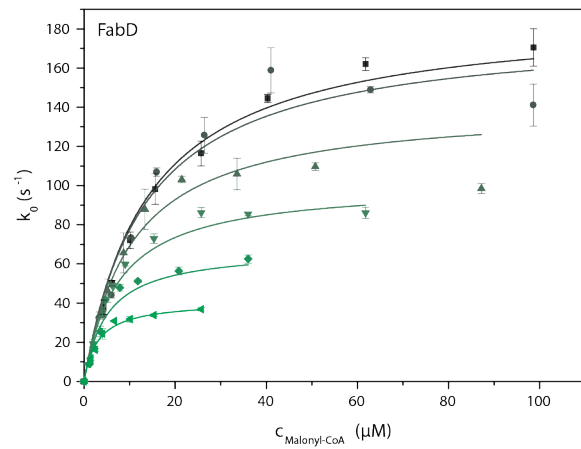
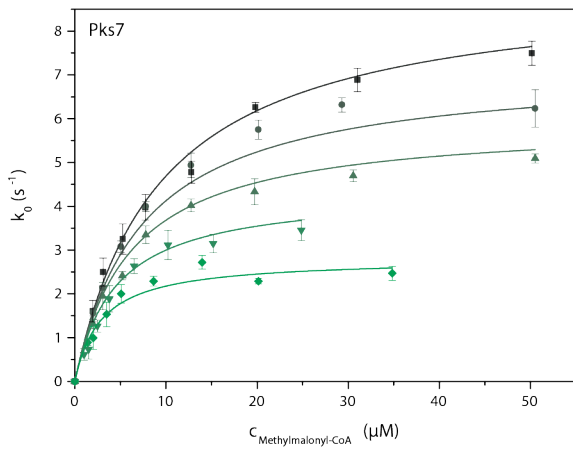
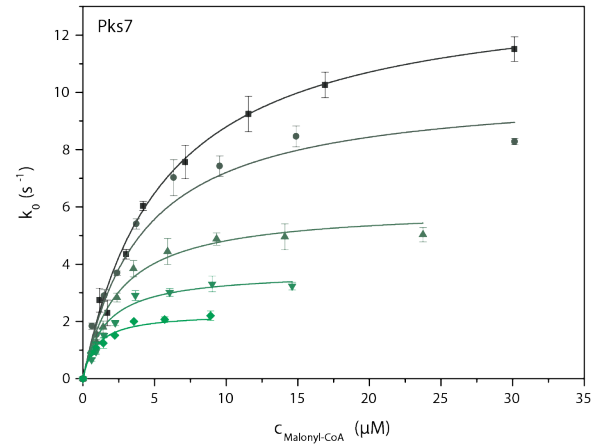
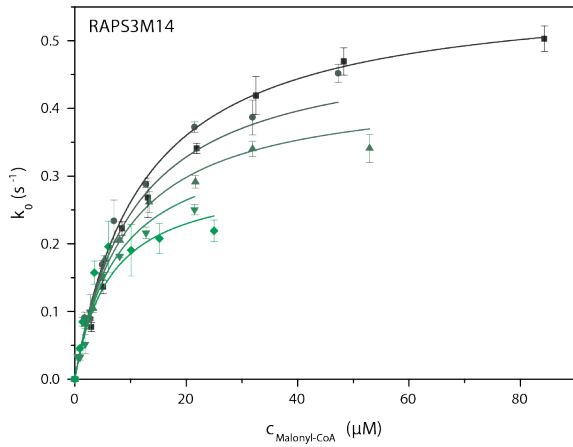
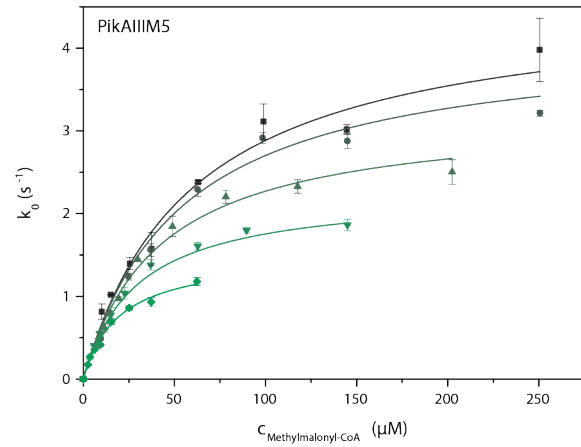
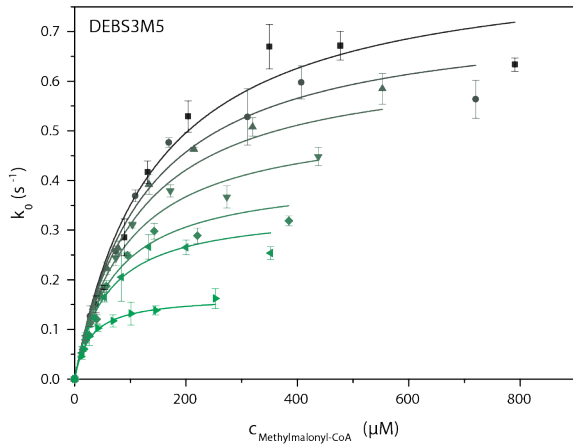


Figure S8: Global Michaelis-Menten fits of transacylation titration curves with X-CoA of AT domains. All measured in biological triplicates. Each color and symbol corresponds to one X-CoA titration curve at a fixed ACP concentration. Error bars give the standard deviation of biological triplicates. ACP concentration as follows: **DEBS3M5**: 11.1 μM , 27.2 μM , 35.4 μM , 53.1 μM , 78.2 μM , 104.9 μM and 152.1 μM , **PikAIIIM5**: 83.3 μM , 143.9 μM , 249.2 μM , 412.9 μM and 528.4 μM , **RAPS3M14**: 48.9 μM , 70.1 μM , 103.2 μM , 156.7 μM and 287.3 μM , **Pks7, Mal**: 31.2 μM , 54.5 μM , 98.6 μM , 220.7 μM and 403.6 μM , **Pks7, MMal**: 80.3 μM , 136.0 μM , 206.3 μM , 273.1 μM and 402.5 μM , **FabD, Mal**: 8.0 μM , 15.0 μM , 26.7 μM , 49.2 μM , 91.0 μM , 103.8 μM , **FabD, MMal**: 10.0 μM , 30.4 μM , 52.9 μM , 102.6 μM , 148.6 μM . ACP concentration ranges as follows: **DEBS3M5**: $0.15\text{-}2.1 \times K_m^{\text{ACP}}$, **PikAIIIM5**: $0.25\text{-}1.6 \times K_m^{\text{ACP}}$, **RAPS3M14**: $0.74\text{-}4.4 \times K_m^{\text{ACP}}$, **Pks7, Mal**: $0.11\text{-}1.4 \times K_m^{\text{ACP}}$, **Pks7, MMal**: $0.16\text{-}0.80 \times K_m^{\text{ACP}}$, **FabD, Mal**: $\times K_m^{\text{ACP}}$, **FabD, MMal**: $\times K_m^{\text{ACP}}$. Substrate consumption as follows: **DEBS3M5**: below 5%, **PikAIIIM5**: below 11%, **RAPS3M14**: up to 36%, but signal increase was still linear, **Pks7, Mal**: up to 42%, but signal increase was still linear, **Pks7, MMal**: up to 18%, **FabD, Mal**: up to 28%, but signal increase was still linear **FabD, MMal**: up to 19%. X-CoA concentration ranges as follows: **DEBS3M5**: $0.16\text{-}8.3 \times K_{m, \text{app}}^{\text{MMal-CoA}}$, **PikAIIIM5**: $0.12\text{-}5.4 \times K_{m, \text{app}}^{\text{MMal-CoA}}$, **RAPS3M14**: $0.076\text{-}8.7 \times K_{m, \text{app}}^{\text{Mal-CoA}}$, **Pks7**: $0.19\text{-}9.5 \times K_{m, \text{app}}^{\text{Mal-CoA}}$ and $0.19\text{-}12 \times K_{m, \text{app}}^{\text{MMal-CoA}}$, **FabD**: $0.23\text{-}13 \times K_{m, \text{app}}^{\text{Mal}}$ and $0.12\text{-}7.2 \times K_{m, \text{app}}^{\text{MMal-CoA}}$.

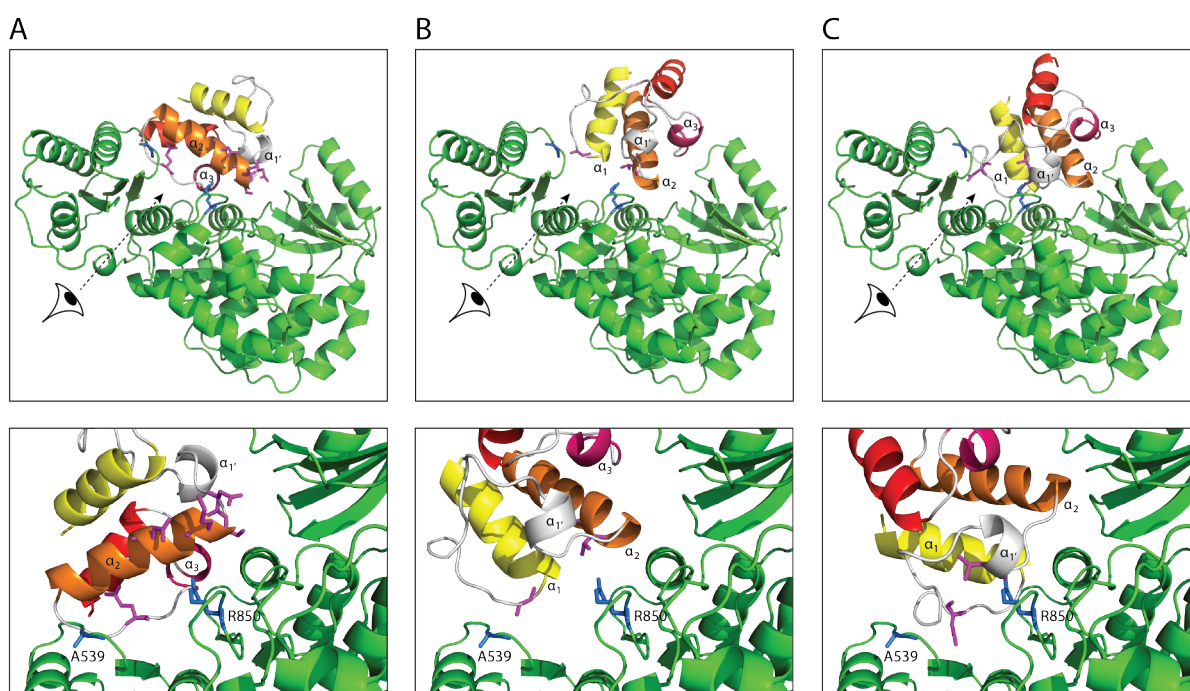


Figure S9: Detailed cartoon representation of the modeled DEBS3M5 AT:ACP interface. DEBS3M5 AT (PDB: 2HG4, shown in green) alignment with the standalone AT from disorazole synthase (PDB: 5ZK4) (A), with the standalone AT VinK from vicanistatin synthase (PDB: 5CZD) (B), and with MAT from human FAS (C).^[125,126,159] Upper panel shows an overview of the AT-ACP interaction of each alignment with perspective eyes for the detailed view showed in the lower panel. Residues A539 and R850 are highlighted in blue. Model DEBS3M5 ACP was aligned with the corresponding ACPs of templates, colored from yellow (N-terminus) to red (C-terminus). Interfacial amino acids suggested by PISA are highlighted in magenta. Despite their different AT:ACP interfaces, all models locate amino acids A539 and R850 (highlighted in blue) in the AT:ACP interfaces.

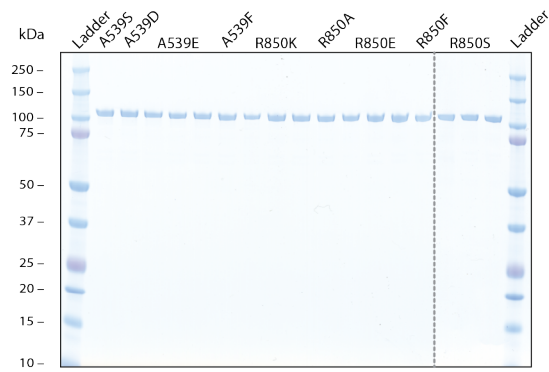
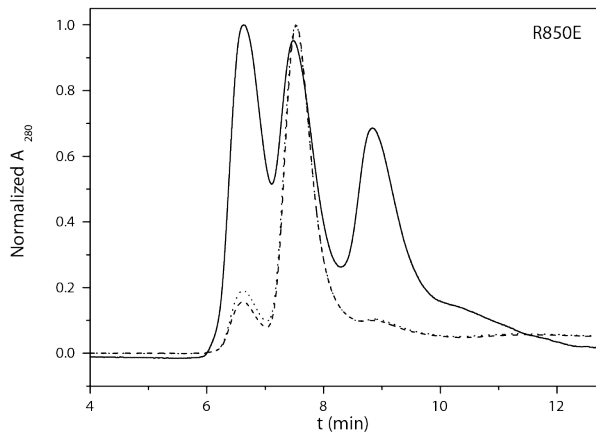
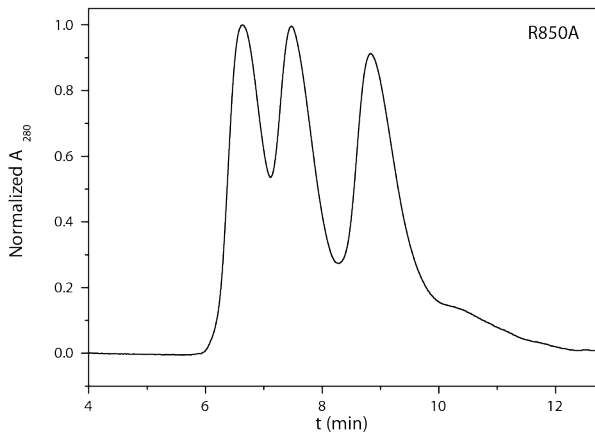
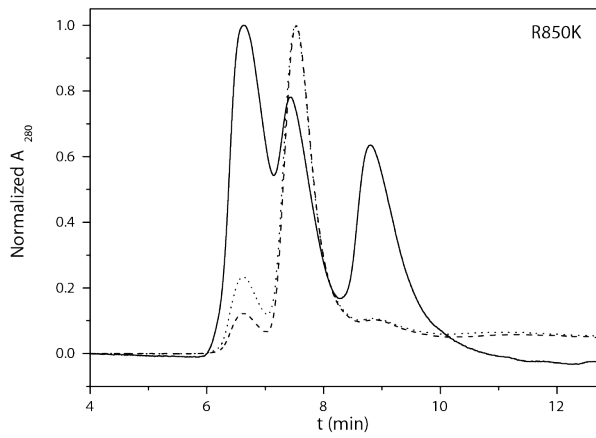
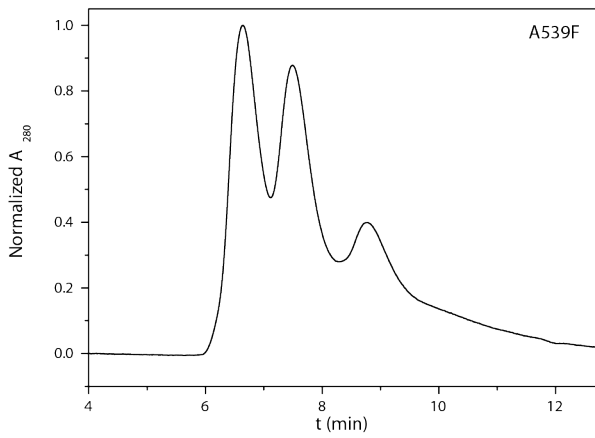
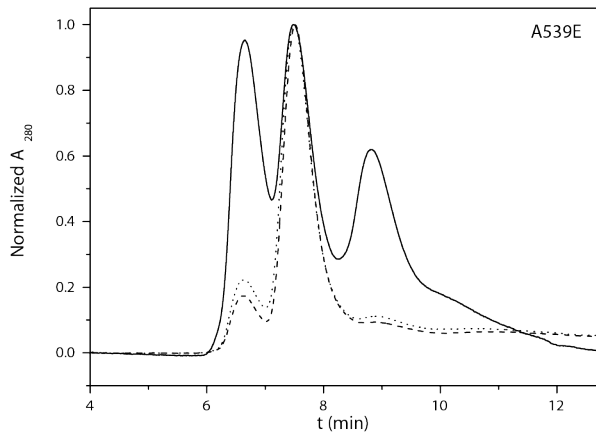
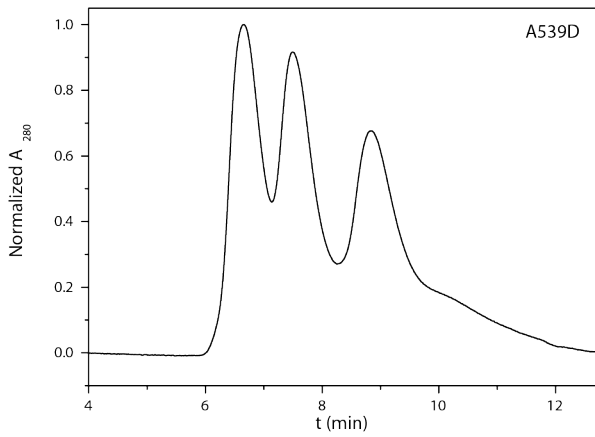
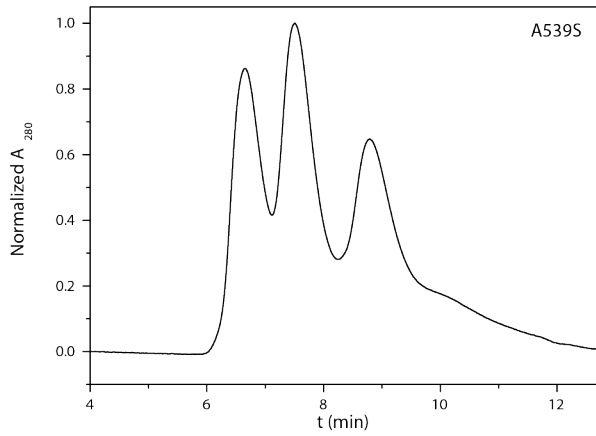
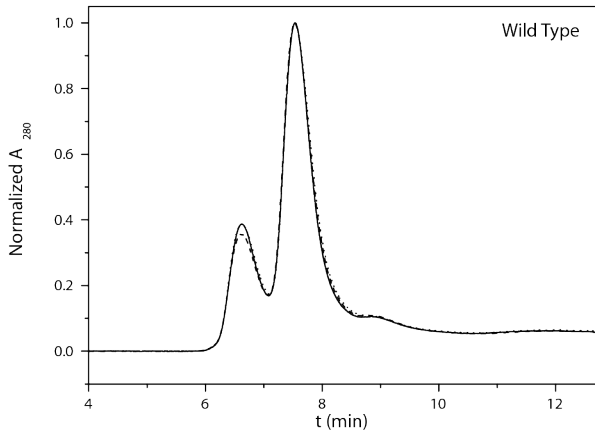


Figure S10: Analytical SDS-PAGEs (NuPAGE Bis-Tris 4-12%) of DEBS3M5 KS^0 -AT mutants. Mutants A539E, R850K, R850E, and R850S expressed in biological triplicates. All AT constructs are highly pure after tandem affinity chromatography. Protein bands migrate at expected masses: A539S – 99.9 kDa, A539D – 99.9 kDa, A539E – 99.9 kDa, A539F – 99.9 kDa, R850K – 99.8 kDa, R850A – 99.8 kDa, R850E – 99.8 kDa, R850F – 99.8 kDa, R850S – 99.8 kDa.

Appendix



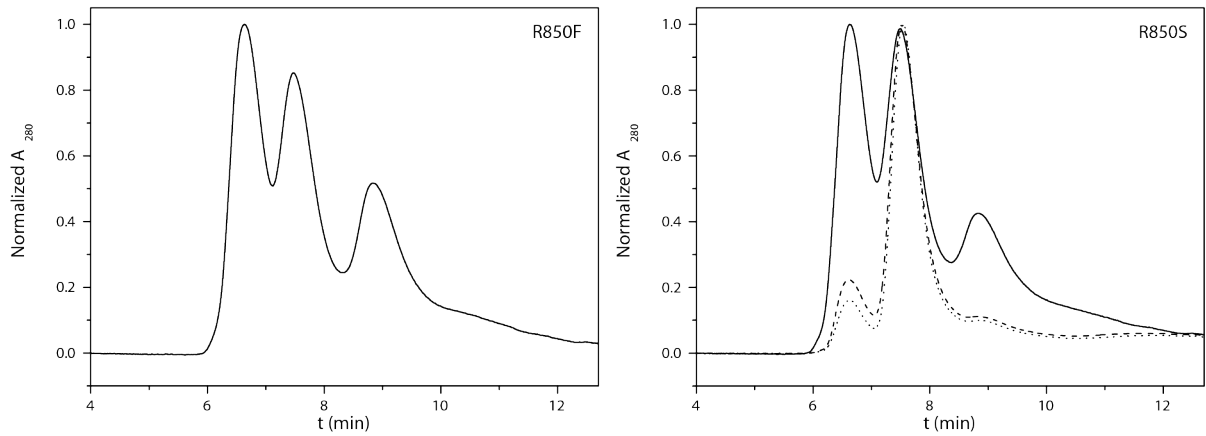
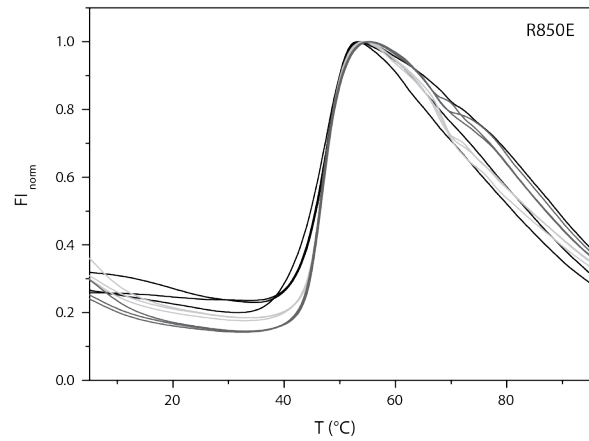
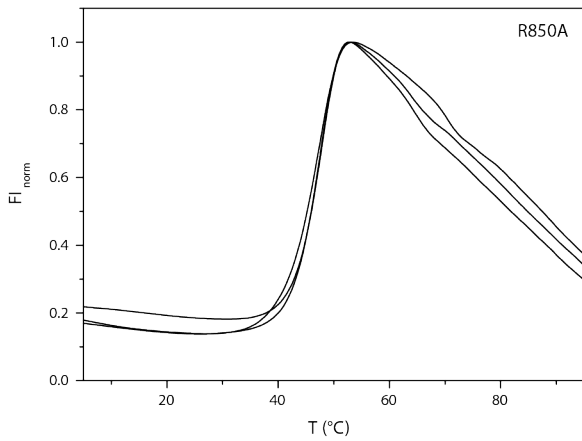
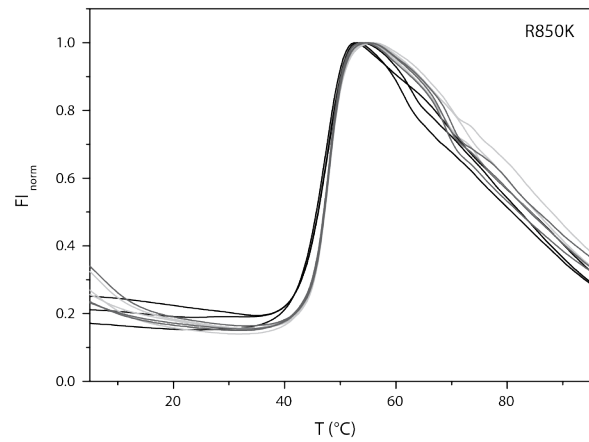
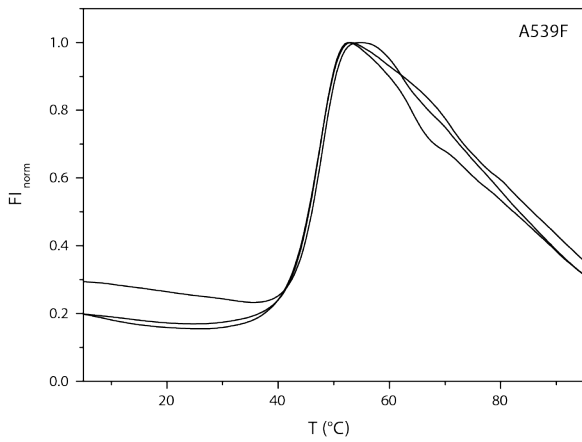
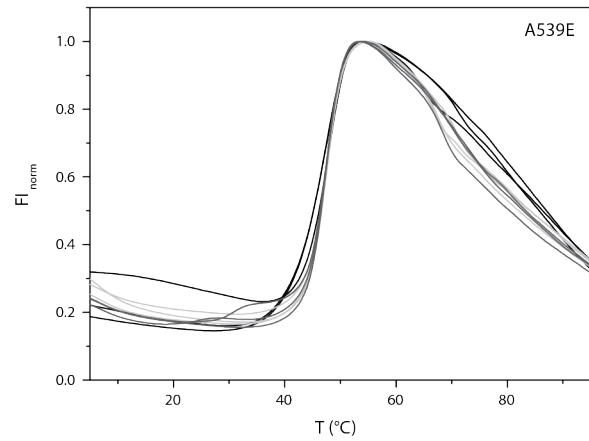
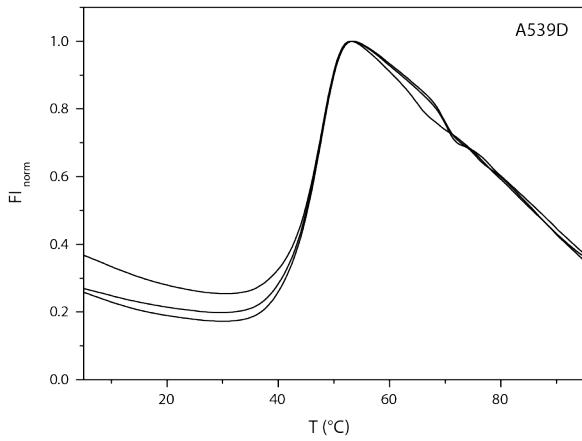
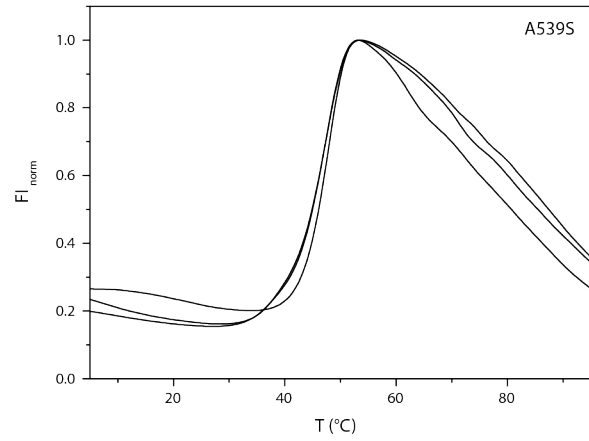
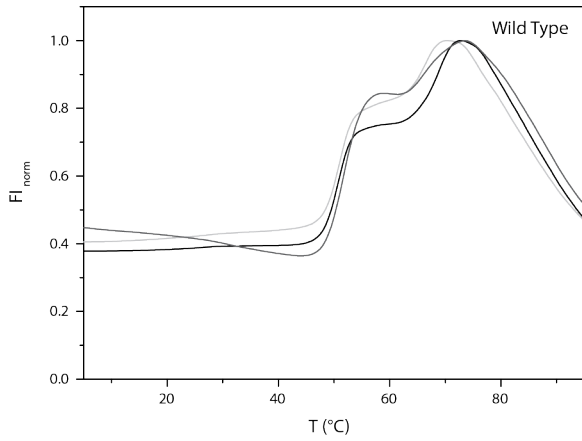


Figure S11: Normalized size exclusion chromatograms of DEBS3M5 KS⁰-AT mutants. Each curve (solid, dotted, dashed) corresponds to one biological replicate. All proteins show monomeric (8.8 min), dimeric (7.5 min), and tetrameric (6.6 min) species.

Appendix



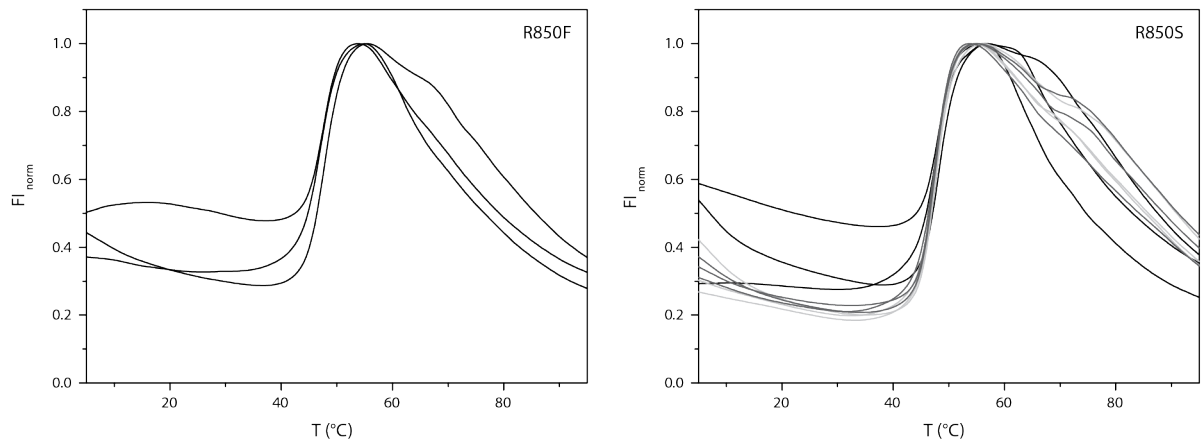


Figure S12: Normalized TSA melting curves of DEBS3M5 KS^0 -AT mutants. Each color (black, dark gray, and light gray) corresponds to one biological replicate. All biological replicates of mutants measured in technical triplicates in assay buffer.

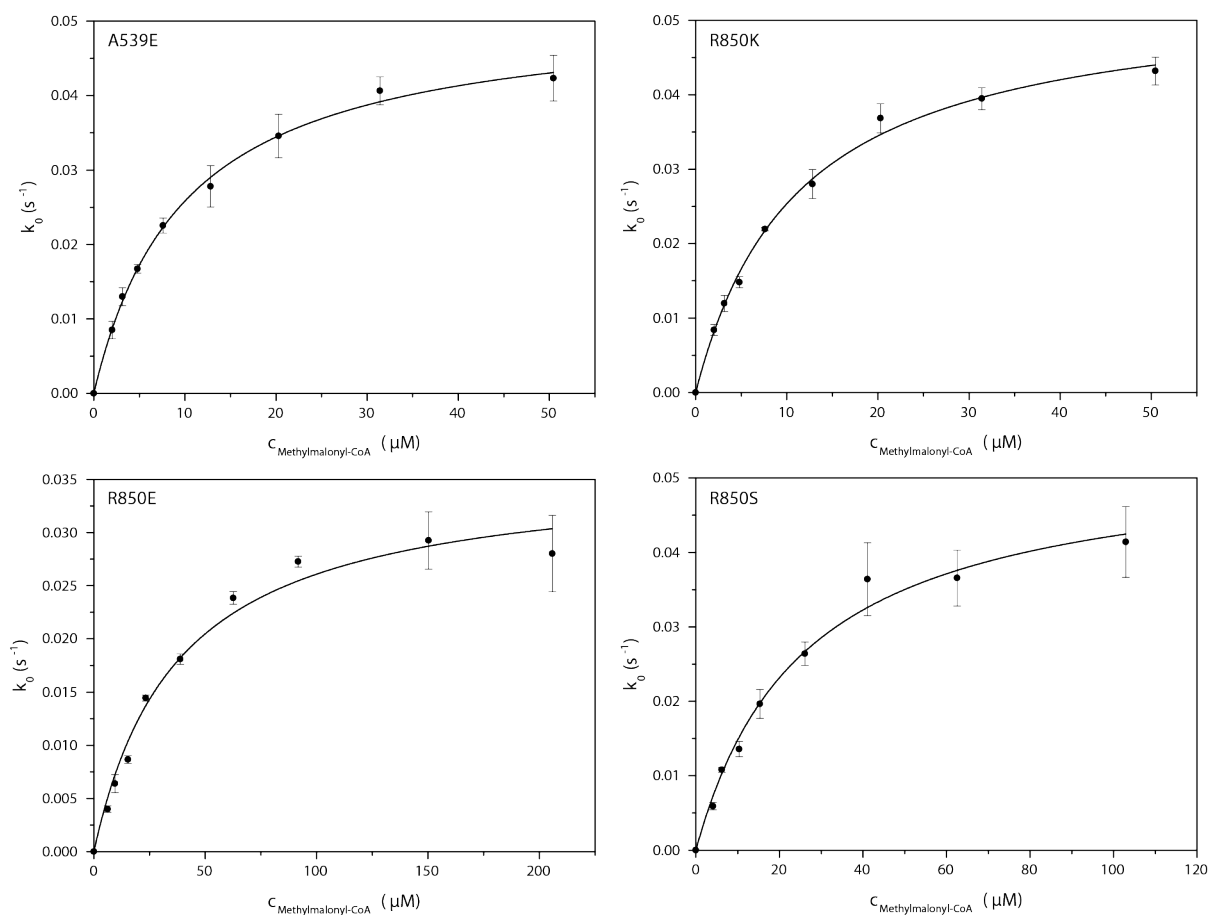


Figure S13: Titration curves for AT-mediated hydrolysis of MMal-CoA using DEBS3M5 KS⁰-AT mutants A539E, R850K, R850E, and R850S. Hydrolysis was measured in biological triplicates. Black dots show the average. Error bars correspond to standard deviation of biological triplicates. Kinetic parameters were determined precisely. Substrate consumption for A539E and R850K was below 10 %, for R850E and R850S below 5 %. MMal range was $0.16\text{-}5.5 \times K_m^{\text{MMal-CoA}}$ (A539E: 0.20-5.1; R850K: 0.18-4.5; R850E: 0.16-5.5; R850S: 0.16-4.0).

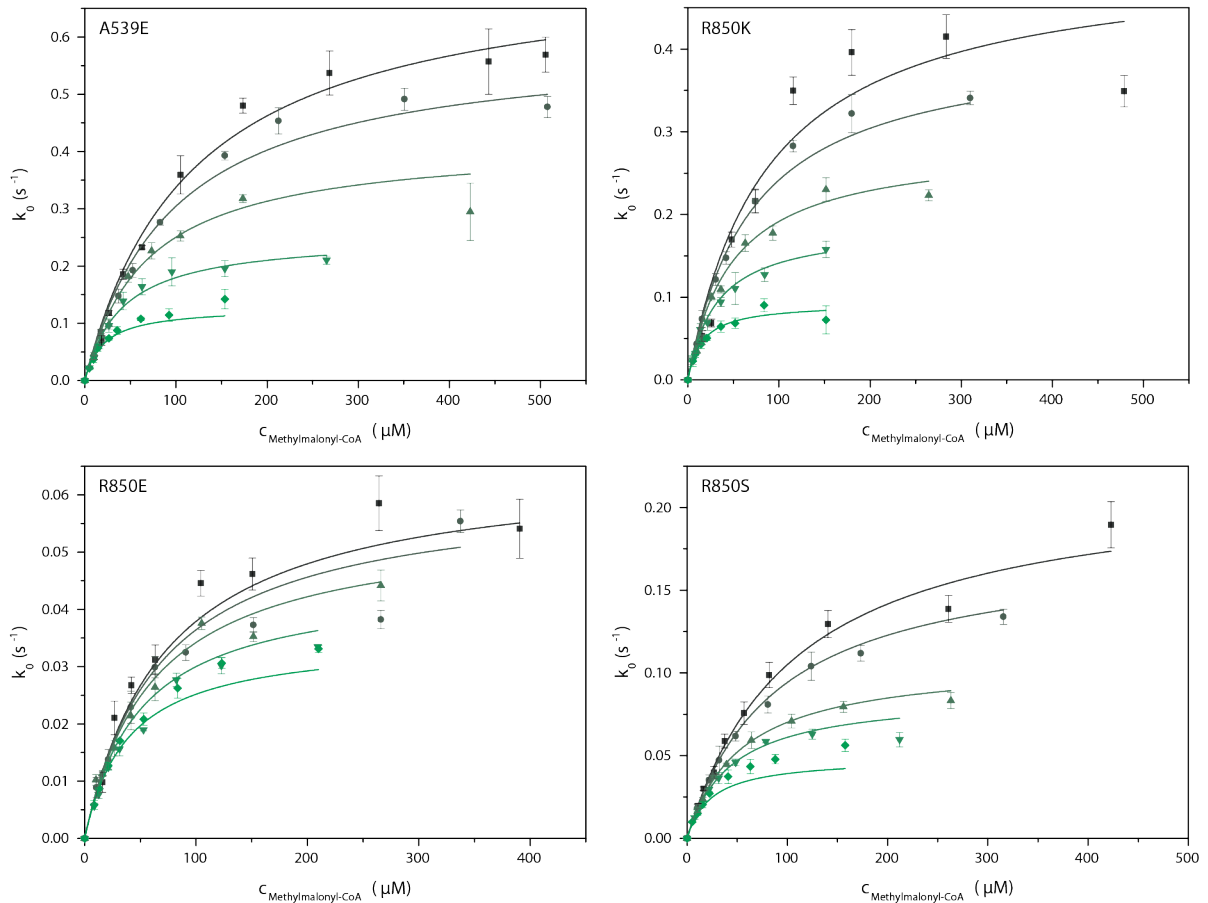


Figure S14: Global Michaelis-Menten fits of transacylation titration curves with MMal-CoA mediated by DEBS3M5 KS^0 -AT mutants A539E, R850K, R850E, and R850S. All measured in biological triplicates. Each color and symbol corresponds to one X-CoA titration curve at a fixed ACP concentration. Error bars give the standard deviation of biological triplicates. ACP concentrations as follows: A539E: 10.6 μM , 24.3 μM , 51.2 μM , 102.3 μM and 187.3 μM , R850K: 11.0 μM , 26.0 μM , 48.9 μM , 100.4 μM and 190.1 μM , R850E: 2.5 μM , 4.6 μM , 10.0 μM , 20.3 μM and 53.3 μM , R850S: 9.5 μM , 20.5 μM , 28.0 μM , 75.5 μM and 152.4 μM . ACP ranges as follows: A539E: $0.14\text{--}2.5 \times K_m^{\text{ACP}}$, R850K: $0.15\text{--}2.6 \times K_m^{\text{ACP}}$, R850E: $1.1\text{--}22 \times K_m^{\text{ACP}}$, R850S: $0.21\text{--}3.3 \times K_m^{\text{ACP}}$. Substrate consumption for R850K was below 5%, for all other proteins was below 10%. MMal-CoA concentration range as follows: A539E: $0.092\text{--}4.9 \times K_{m, \text{app}}^{\text{MMal-CoA}}$, R850K: $0.13\text{--}9.9 \times K_{m, \text{app}}^{\text{MMal-CoA}}$, R850E: $0.15\text{--}6.0 \times K_{m, \text{app}}^{\text{MMal-CoA}}$, R850S: $0.096\text{--}6.6 \times K_{m, \text{app}}^{\text{MMal-CoA}}$.

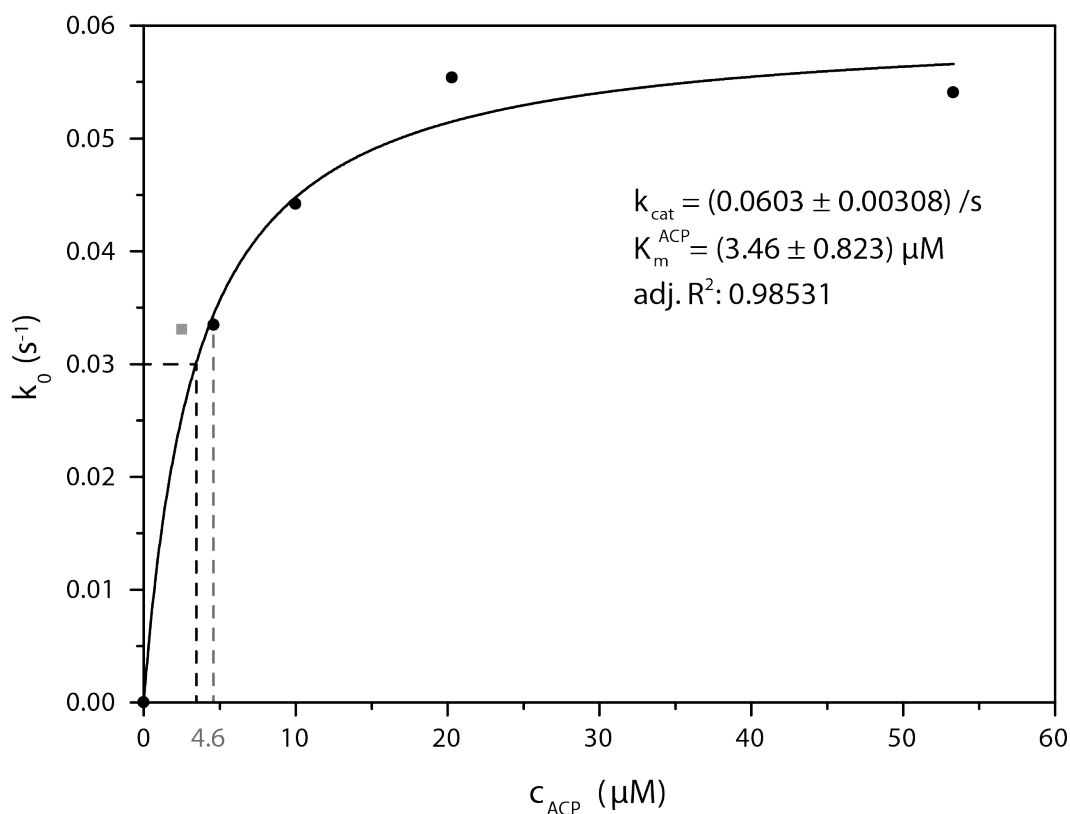


Figure S15: ACP titration curve at saturated MMal-CoA concentration of mutant R850E. Due to high measurement errors at low ACP concentrations, the kinetic parameters determined via the global fit are rather error-prone. The ACP titration curve was used to determine more reliable kinetic parameters for the AT-ACP interaction. This data was not separately determined, but data points were extracted from MMal-CoA titrations with increasing ACP concentrations (see Fig. S14). The lowest ACP concentration (gray box) is omitted to give a more accurate Michaelis-Menten fit. The new plot gives a comparable turnover rate, but a Michaelis-Menten constant of 3.46 μM, which is more reliable than determined before. This value is only 4.8 % of the wild type K_m^{ACP} , which is almost the same percentage as determined for the turnover rate k_{cat} . The second lowest ACP concentration (4.6 μM) has a k_0 above half k_{cat} . We have taken this value as upper boundary of the Michaelis-Menten constant.

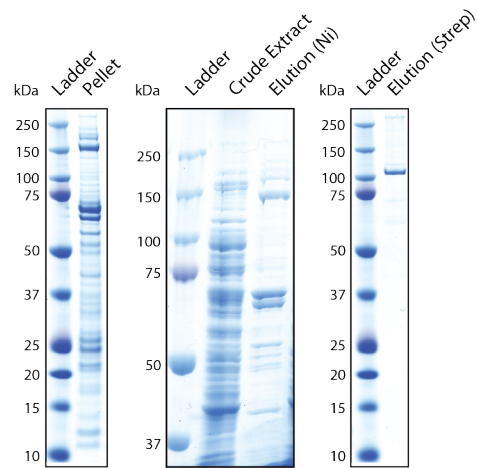


Figure S16: Purification of chimeric mFAS with DEBS3M5 AT. SDS-PAGE (NuPAGE Bis-Tris 4-12 % and 7 % polyacrylamide) of different purification fractions. In the elution fraction of the second affinity column (Strep), an additional protein band of around 100 kDa was observed, which was not present in the elution fraction of the first (Ni). This was old protein bound to the second column and most probably a KS-AT construct.

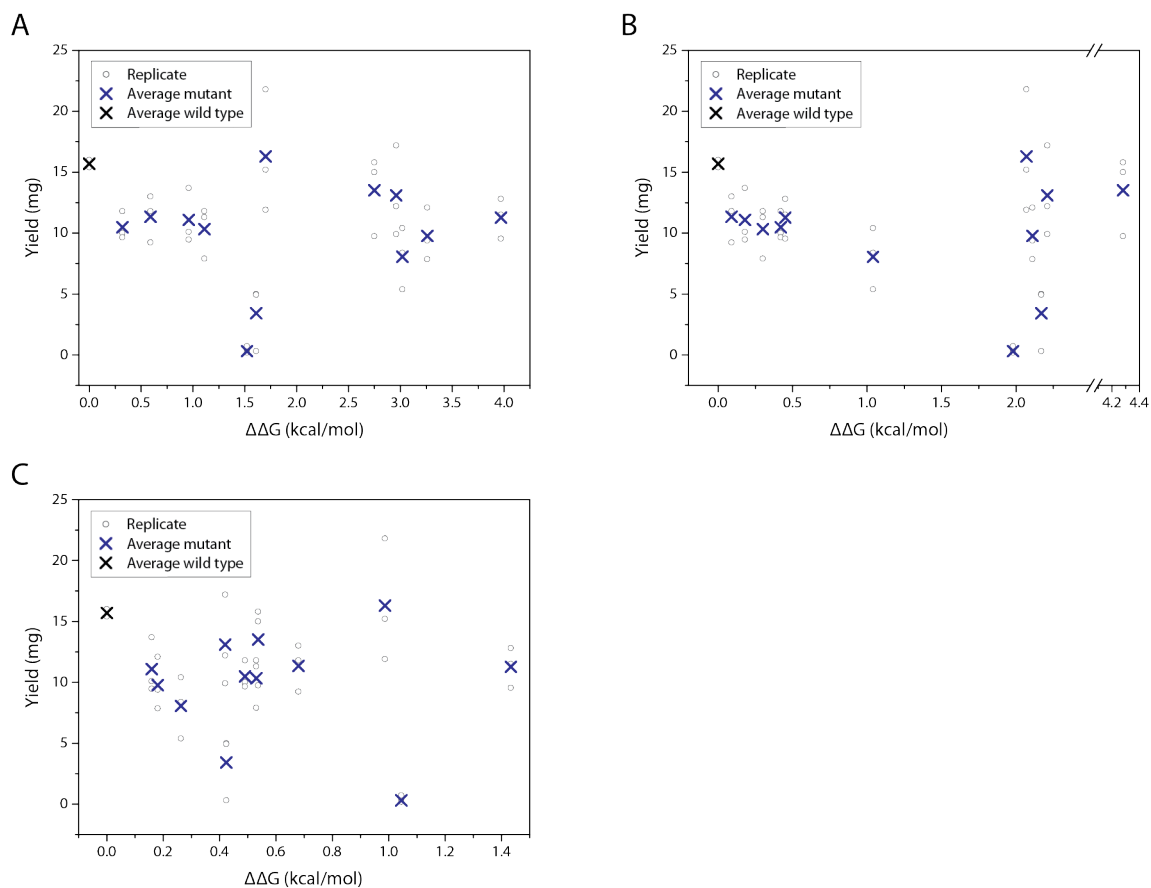


Figure S17: Yields (in mg per 1L expression culture) of mFAS KS-AT wild type and mutants correlated to predicted changes in the Gibbs free energy ($\Delta\Delta G$ in kcal/mol). (A) Chain A without PAL predicted by Robetta. (B) Chain A without PAL predicted by DrugScore. (C) Chain A with PAL predicted by mCSM. Black circles indicate the yields of biological replicates. Average yields of biological triplicates of wild type and mutants are indicated by black and blue crosses, respectively. Please mind the scales of the different x-axis sectors. Predictions of the different tools showed no or only partly correlation tendencies. Overall, they were not suitable for estimating effects of mutations on protein stability.

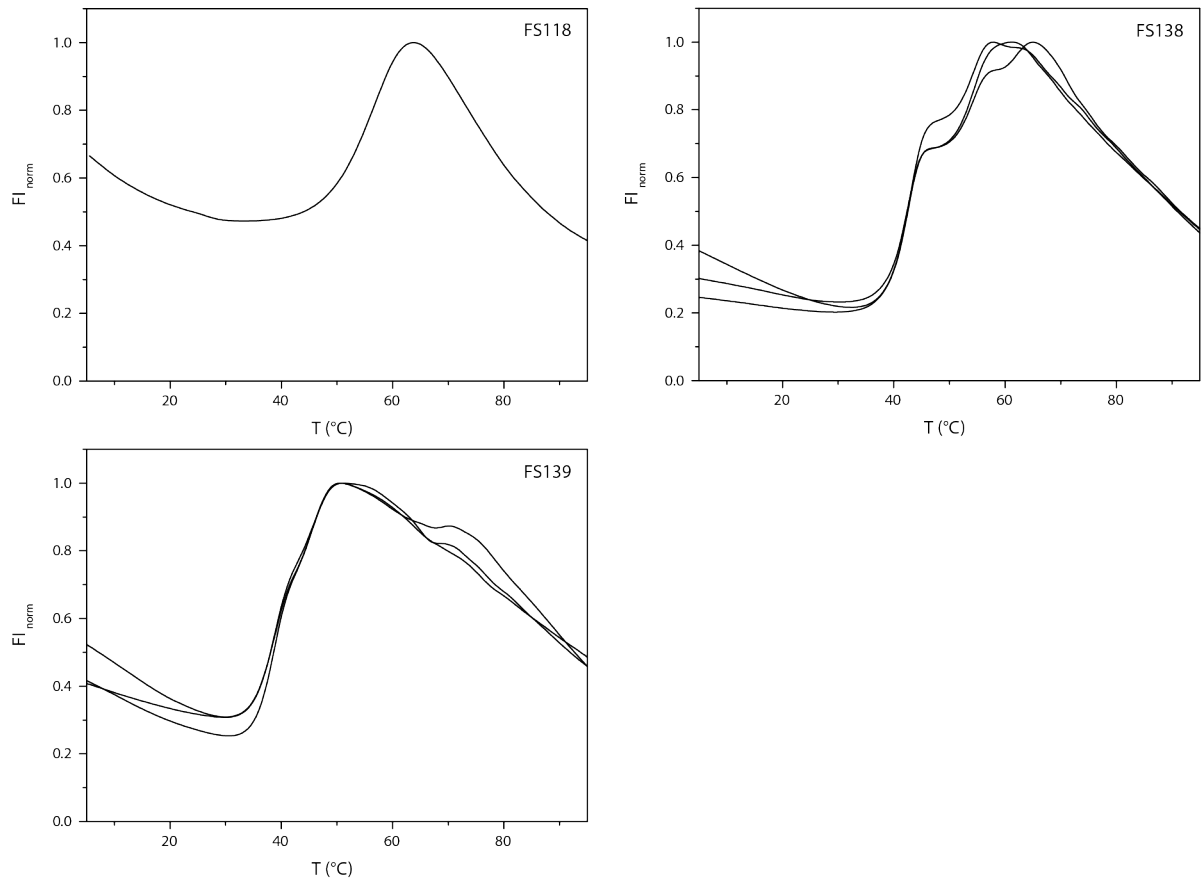


Figure S18: Normalized TSA melting curves of mFAS/DEBS3M6 KS-AT chimeras FS118 and FS139 and DEBS3M6 KS-AT FS138 in storage buffer. Each curve corresponds to one biological replicate.

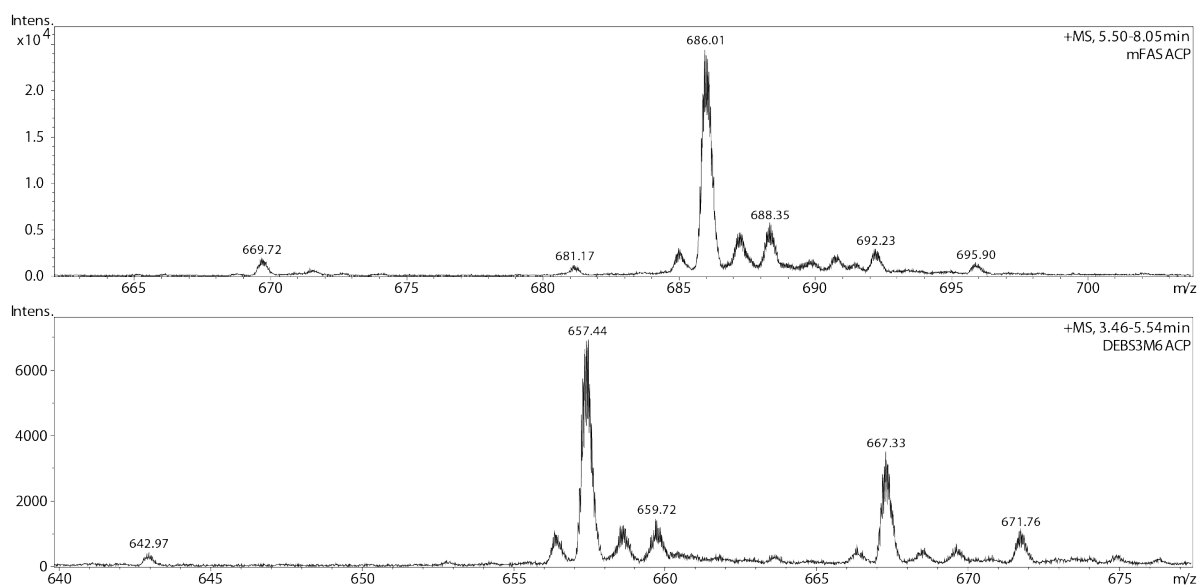


Figure S19: MS analysis of mFAS and DEBS3M6 ACP domains after SEC. Shown are the m/z -values of the average protein masses of the 18⁺ charge state (MS¹). In none of the samples, m/z -values of the *apo* species were recorded. Main m/z -values correspond to the *holo* form. Both proteins do not include the N-terminal methionine. **mFAS ACP**, $z = 18$: $m/z = 686.01$ corresponds to *holo*-ACP without N-terminal methionine. $m/z = 688.35$ corresponds to acetylation ($\Delta m = 42$ Da), $m/z = 695.90$ corresponds to α -N-gluconozylation ($\Delta m = 178$ Da). $m/z = 692.23$ is an unknown modified variant ($\Delta m = 112$ Da). $m/z = 681.17$ corresponds to *holo*-ACP without N-terminal methionine and serine ($\Delta m = -87$ Da). **DEBS3M6 ACP**, $z = 18$: $m/z = 657.44$ corresponds to *holo*-ACP without N-terminal methionine. $m/z = 659.72$ corresponds to acetylation ($\Delta m = 41$ Da), $m/z = 667.33$ and 671.76 correspond to α -N-gluconozylation ($\Delta m = 178$ Da and 258 Da). $m/z = 692.23$ is an unknown modified variant ($\Delta m = 112$ Da).

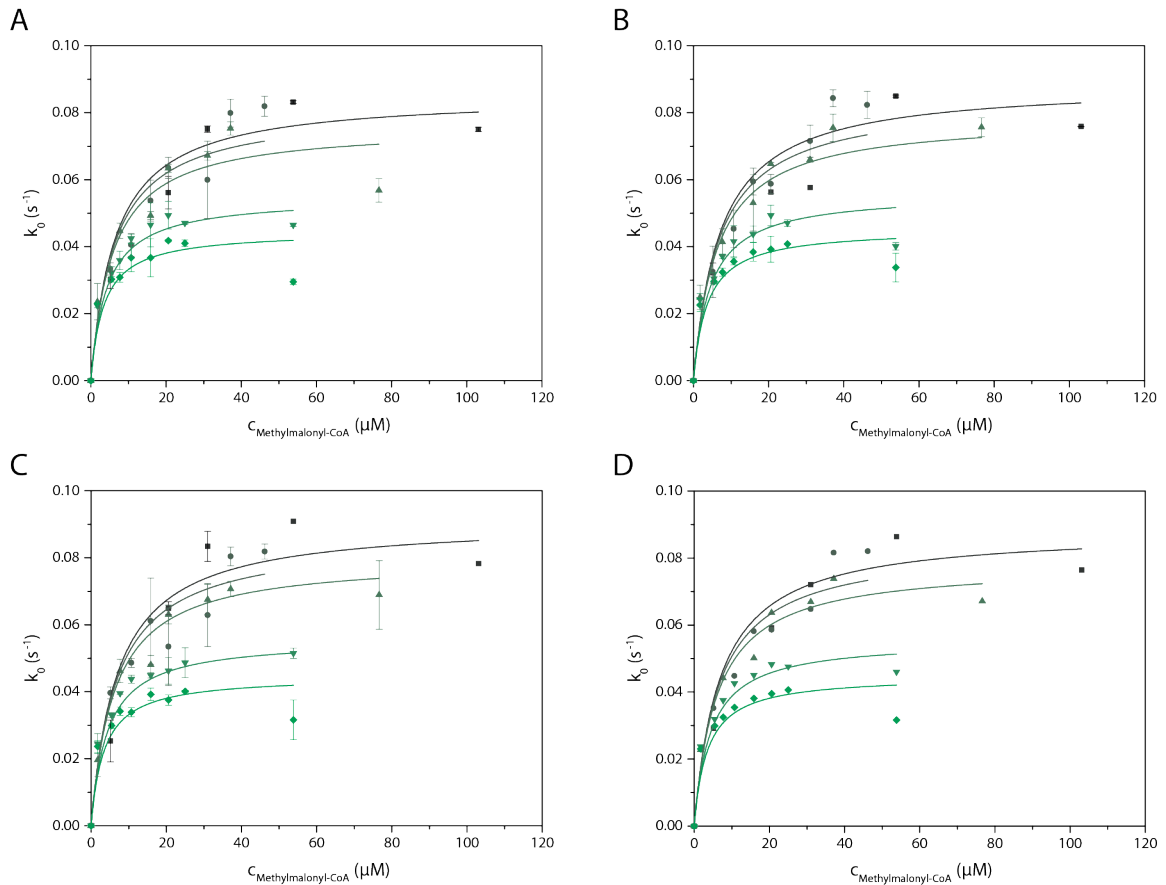


Figure S20: Global Michaelis-Menten fits of transacylation titration curves with MMal-CoA mediated by m(KS)-DEBS3M6(AT) chimera FS118. (A)-(C) Biological replicates measured in technical triplicates. Each color and symbol corresponds to one X-CoA titration curve at a fixed ACP concentration. Error bars give the standard deviation of the technical triplicates. (D) Average of biological replicates. Each color and symbol corresponds to one X-CoA titration curve at a fixed ACP concentration. ACP concentrations as follows: 67.1 μM , 98.4 μM , 227.3 μM , 290.0 μM and 360.9 μM . ACP range as follows: $0.65\text{--}3.5 \times K_m^{\text{ACP}}$. Substrate consumption was up to 26%. MMal-CoA concentration range as follows: $0.17\text{--}85 \times K_{m, \text{app}}^{\text{MMal-CoA}}$.

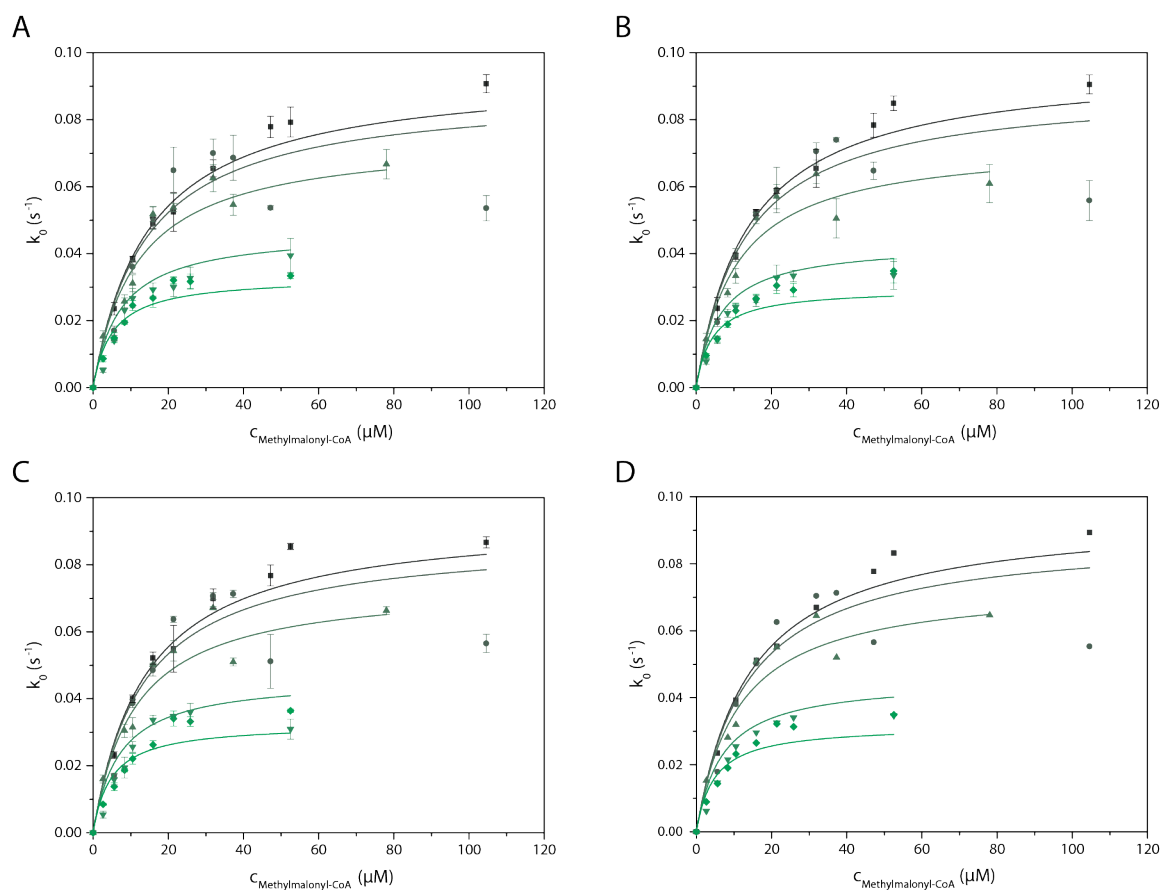


Figure S21: Global Michaelis-Menten fits of transacylation titration curves with MMal-CoA mediated by DEBS3M6(KS-AT) wild type FS138. (A)-(C) Biological replicates measured in technical triplicates. Each color and symbol corresponds to one X-CoA titration curve at a fixed ACP concentration. Error bars give the standard deviation of the technical triplicates. (D) Average of biological replicates. Each color and symbol corresponds to one X-CoA titration curve at a fixed ACP concentration. ACP concentrations as follows: 61.5 μM , 98.4 μM , 211.9 μM , 303.5 μM and 354.0 μM . ACP range as follows: $0.24-1.4 \times K_m^{\text{ACP}}$. Substrate consumption was up to 24%. MMal-CoA concentration range as follows: $0.14-11 \times K_{m, \text{app}}^{\text{MMal-CoA}}$.

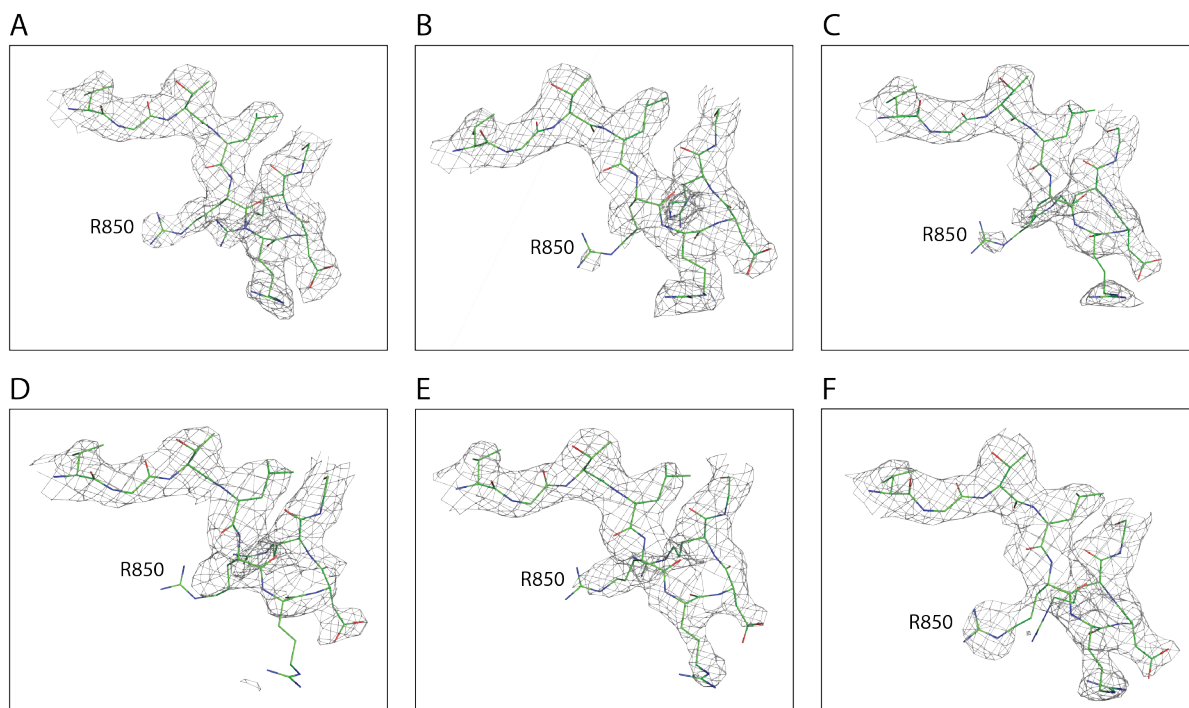


Figure S22: Electron density map of residue R850. The electron density map for the residues V846-G854 of DEBS3M5 AT (PDB: 2HG4^[9]) for all six chains (A-F) shows poor electron density for the interfacial residue R850 mutated in this study. This finding supports the flexibility of this residue. Different binding modes might be involved during catalysis. The 2fo-fc map was created using PyMOL (MacPyMOL Version 1.7.0.3, Schrödinger, LLC) (sigma=1, carve=1.8).

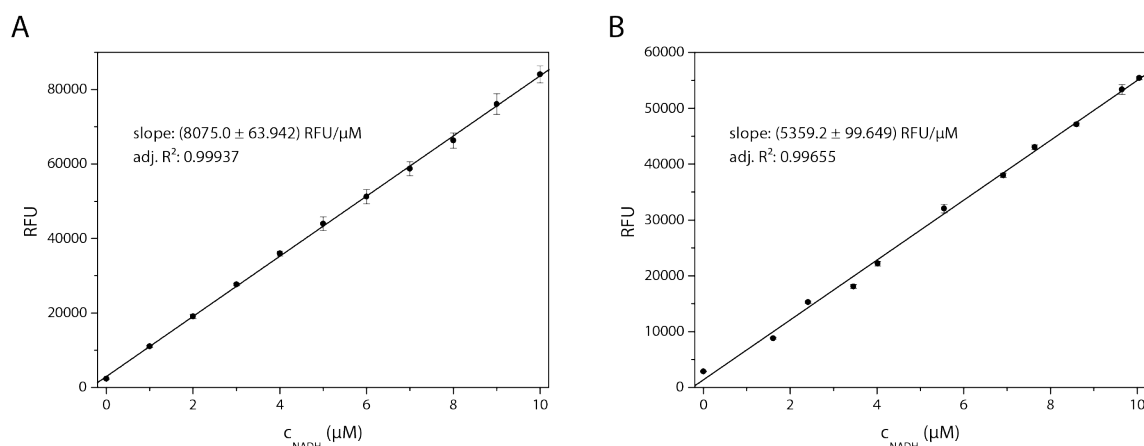
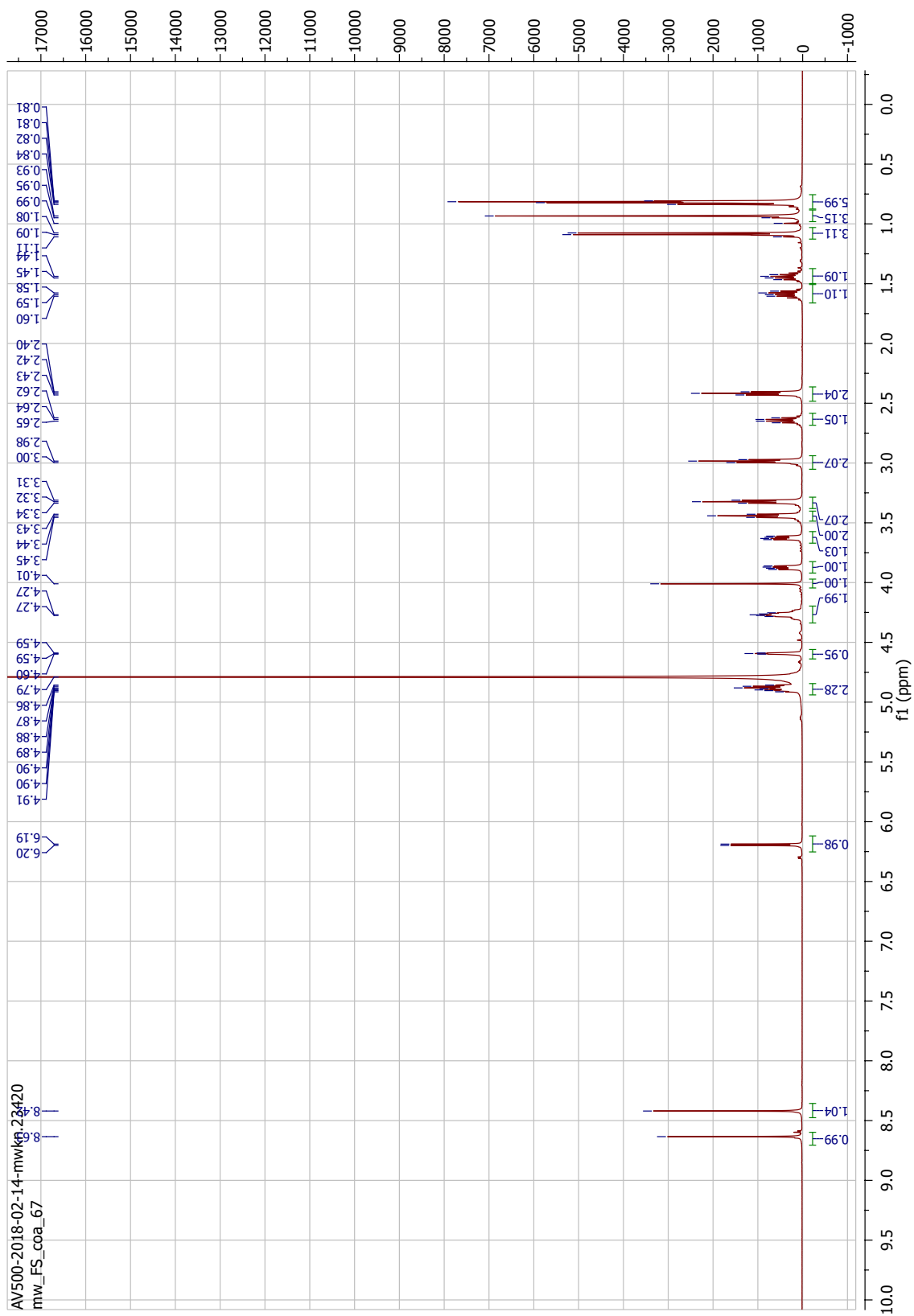


Figure S23: NADH calibration. For quantification of AT activity, an NADH calibration was performed. (A) 96-well format. 8075 relative fluorescence units (RFU) correspond to 1 μM NADH. The NADH calibration factor changes over time due to the decreasing power of the UV lamp. An internal standard (50 μM X-CoA and 50 μM ACP) was measured the same day as the NADH calibration. This internal standard was used to determine the calibration factor for each measurement time span (Tab. S11). This factor was subsequently used to convert RFU into concentrations (μM) and ranges from 13450 to 8075 RFU per 1 μM NADH. Error bars represent the standard deviation of three independent measurements that were measured in technical triplicates. (B) 384-well format. 5359 RFU correspond to 1 μM NADH. Error bars represent the standard deviation of technical triplicates.

Figure S24: $^1\text{H-NMR}$ of 2-methylbutyryl-CoA.

Statement of Personal Contributions

Except where stated otherwise by reference or acknowledgment, the work presented in this thesis was generated by myself. Projects were regularly discussed with my advisor Prof. Dr. Martin Grininger (Goethe-University Frankfurt, Institute for Organic Chemistry and Chemical Biology). Contributions from colleagues and students are explicitly referenced at respective parts in the thesis. A detailed statement of personal contributions is given below.

The following persons contributed to the results of this thesis:

Dr. Christina S. Heil (AK Grininger, Goethe-University Frankfurt, Institute for Organic Chemistry and Chemical Biology) and Dr. Alexander Rittner (AK Grininger, Goethe-University Frankfurt, Institute for Organic Chemistry and Chemical Biology) contributed through their supervision during my master's thesis and conceived together with Prof. Dr. Martin Grininger (Goethe-University Frankfurt, Institute for Organic Chemistry and Chemical Biology) the initial idea of the kinetic characterization of AT domains.

Dr. Alexander Rittner (AK Grininger, Goethe-University Frankfurt, Institute for Organic Chemistry and Chemical Biology) helped choosing the first domain boundaries for AT swapping.

My colleague Lynn Buyachuihan (AK Grininger, Goethe-University Frankfurt, Institute for Organic Chemistry and Chemical Biology) was involved in the project of creating mFAS-based elongation modules. We conceived this project together with Prof. Dr. Martin Grininger (Goethe-University Frankfurt, Institute for Organic Chemistry and Chemical Biology). Lynn Buyachuihan was involved in isolation of the elongation modules. She performed the TSA, the CoA-488 assay, activity assays, and LC-MS analyses of all modules.

Student Maren Berlinghof was involved in the construction of chimeric mFAS/DEBS3M5 KS-AT as well as in the isolation and analysis of mFAS and DEBS3M5 KS-AT wild types during her master's thesis under my supervision. Furthermore, she assisted in the isolation of AT domains for their kinetic characterization.

Student Alicia Just was involved in isolation and analysis of mFAS/DEBS3M6 KS-AT constructs during her master's thesis under my supervision.

Students Guanzhu Götze, Shengnan Li, and Yunqing Li assisted in cloning plasmids encoding the DEBS3M5 AT:ACP interface mutants.

Student Amin Fahim assisted in the DEBS3M5 AT:ACP interface mutation study during his bachelor's thesis under my supervision.

Student Robin Klimek assisted in the isolation of AT domains for their kinetic characterization.

Student Tanja Ott was involved in the construction of chimeric mFAS-based elongation modules with VemG, VemH, and EcPKS1 AT.

Student Matthias Zeug was involved in isolation and analysis of mFAS-based elongation modules during his master's thesis under supervision of Lynn Buyachuihan (AK Grininger, Goethe-University Frankfurt, Institute for Organic Chemistry and Chemical Biology).

Student Laurell Kessler was involved in analysis of mFAS-based elongation modules under supervision of Lynn Buyachuihan (AK Grininger, Goethe-University Frankfurt, Institute for Organic Chemistry and Chemical Biology).

Dr. Gina Grammbitter (AK Bode, Goethe-University Frankfurt, Institute for Molecular Bio Science) kindly performed HPLC-MS analyses of ACPs and assisted with analysis of the data.

Dr. Karthik S. Paithankar (AK Grininger, Goethe-University Frankfurt, Institute for Organic Chemistry and Chemical Biology) performed the Rosetta calculations of the mFAS and DEBS3M5 KS-AT wild types. Furthermore, he was involved in the crystallization experiments.

Dr. Apirat Chaikwad (AK Knapp, Goethe-University Frankfurt, Institute for Pharmaceutical Chemistry) was involved in the crystallization experiments.

Marek Wanior (AK Knapp, Goethe-University Frankfurt, Institute for Pharmaceutical Chemistry) kindly assisted in purification of 2-MB-CoA via HPLC.

Alexander Stegemann (AK Rischke, Goethe-University Frankfurt, Institute for Theoretical Physics) kindly discussed the derivation of kinetic parameters and assisted in the mathematical examination of changes in transacylation kinetic parameters caused by point mutations.

Martin Schwalm (AK Knapp, Goethe-University Frankfurt, Institute for Pharmaceutical Chemistry) assisted with depiction of electron density data.

Copyright and Creative Commons Licences

Whenever a figure, table, or text is identical to a previous publication, it is hereby stated explicitly that copyright permission and/or co-author agreement has been obtained.

The following figures of the thesis have been modified from previous publications:

- Fig. 2.5 “Structure and domain organization of type I FAS systems.”: Heil, Christina S., Wehrheim, S. Sophia, Paithankar, Karthik S., Grininger, Martin “Fatty Acid Biosynthesis: Chain-Length Regulation and Control” *ChemBioChem* 20, 2298-2321 (2019).^[41] Copyright Wiley-VCH GmbH. Reproduced with permission.
- Fig. 2.6 “Structural models for modular PKSs.”: Reproduced from Klaus, Maja and Grininger, Martin “Engineering strategies for rational polyketide synthase” *Natural Product Reports* 35, 1070-1081 (2018)^[75] with permission from the Royal Society of Chemistry.
- Fig. 2.8 “Reaction catalyzed by the AT.”: Heil, Christina S., Wehrheim, S. Sophia, Paithankar, Karthik S., Grininger, Martin “Fatty Acid Biosynthesis: Chain-Length Regulation and Control” *ChemBioChem* 20, 2298-2321 (2019).^[41] Copyright Wiley-VCH GmbH. Reproduced with permission.
- Fig. 2.9 “Reaction catalyzed by the KS.”: Heil, Christina S., Wehrheim, S. Sophia, Paithankar, Karthik S., Grininger, Martin “Fatty Acid Biosynthesis: Chain-Length Regulation and Control” *ChemBioChem* 20, 2298-2321 (2019).^[41] Copyright Wiley-VCH GmbH. Reproduced with permission.

The following parts of the thesis have been (in parts) previously published: Stegemann, Franziska and Grininger, Martin “Transacylation Kinetics in Fatty Acid and Polyketide Synthases and its Sensitivity to Point Mutations” *ChemCatChem* (2021).^[1]

- Chapters/Sections:

- 1 “**Abstract**”
- 3.2 “**Kinetic Description of AT-mediated Reactions**”
- 3.3 “**Kinetic Analysis of Acyl Transferases**”
- 4.2 “**Kinetic Analysis of Acyl Transferases**”
- 4.3 “**Implications of Kinetic Properties on the AT Reaction Mechanism**”
- 4.4 “**Implications of Kinetic Properties on FAS/PKS Function and Evolution**”
- 4.5 “**Impact of Interface Mutations on Transacylation Kinetics**”
- 4.6 “**Implications of Interface Mutation Study on PKS Engineering**”
- 4.11 “**Conclusion and Outlook**”
- 5 “**Experimental Procedures**”

- Figures:

Fig. 2.4 "Phylogenetic relationship between FASs and PKSs."

Fig. 3.1 "Scheme of AT-mediated reactions."

Fig. 3.2 "Phylogenetic tree of AT domains."

Fig. 3.4 "Transacylation and hydrolysis screening of FAS and PKS ATs."

Fig. 3.5 "Transacylation parameters of FAS and PKS ATs."

Fig. 3.7 "Transacylation and hydrolysis screening of DEBS M5 AT:ACP interface mutants."

Fig. 3.8 "Transacylation parameters of DEBS M5 AT:ACP interface mutants."

Fig. 3.9 "Transacylation transition state energy of DEBS M5 AT:ACP interface mutants."

Fig. 4.1 "Plasticity of transacylation kinetics of DEBS M5 AT:ACP interface mutants."

Fig. S1 "SDS-PAGEs of FAS and PKS ATs."

Fig. S2 "Size exclusion chromatograms of FAS and PKS ATs."

Fig. S3 "TSA melting curves of FAS and PKS ATs."

Fig. S4 "SDS-PAGEs of FAS and PKS ACPs."

Fig. S5 "Mass spectrometric analysis of FAS and PKS ACPs."

Fig. S6 "TSA melting curves of FAS and PKS ACPs."

Fig. S7 "Titration curves for AT-mediated hydrolysis."

Fig. S8 "Titration curves for AT-mediated transacylation."

Fig. S9 "Detailed model of DEBS M5 AT:ACP interface."

Fig. S10 "SDS-PAGEs of DEBS M5 AT:ACP interface mutants."

Fig. S11 "Size exclusion chromatograms of DEBS M5 AT:ACP interface mutants."

Fig. S12 "TSA melting curves of DEBS M5 AT:ACP interface mutants."

Fig. S13 "Hydrolysis curves for DEBS M5 AT:ACP interface mutants."

Fig. S14 "Transacylation curves for DEBS M5 AT:ACP interface mutants."

Fig. S15 "ACP titration curve of DEBS M5 AT:ACP mutant R850E."

Fig. S22 "Electron density map of DEBS M5 AT residue R850."

Fig. S23 "NADH calibration."

- Tables:

Tab. 3.1 “Hydrolysis parameters of FAS and PKS ATs.”

Tab. 3.2 “Transacylation parameters of FAS and PKS ATs.”

Tab. 3.3 “Kinetic parameters of DEBS M5 AT:ACP interface mutants.”

Tab. S1 “Expression yields of FAS and PKS ATs.”

Tab. S2 “TSA of FAS and PKS ATs and ACPs.”

Tab. S3 “Expression yields of FAS and PKS ACPs.”

Tab. S4 “Mass spectrometric analysis of FAS and PKS ACPs.”

Tab. S5 “Initial substrate screening of FAS and PKS ATs.”

Tab. S6 “Hydrolysis parameters of FAS and PKS ATs after transacylation.”

Tab. S7 “Expression yields of DEBS M5 AT:ACP interface mutants.”

Tab. S8 “TSA of DEBS M5 AT:ACP interface mutants.”

The complete draft of the manuscript accepted for publication as Stegemann, Franziska and Grininger, Martin “Transacylation Kinetics in Fatty Acid and Polyketide Synthases and its Sensitivity to Point Mutations” *ChemCatChem* (2021) was initially written by myself. Prof. Dr. Martin Grininger participated significantly with editing and phrasing to the final version in the process of publication. Copyright retains by the authors. Creative Commons Attribution 4.0 International License (CC BY 4.0).

Copyright and Creative Commons declaration of the *ChemCatChem* journal:

“Authors retain copyright of their article and are given a choice of Creative Commons license under which to publish their work.”

Source:

https://onlinelibrary.wiley.com/page/journal/18673899/homepage/2491_onlineopen.html
(accessed 29th April 2021)

The following master's and bachelor's theses, which were submitted to the department 14 – Biochemistry, Chemistry and Pharmacy of the Goethe-University Frankfurt, contributed to the results of the present thesis:

Franziska Stegemann "Fluorescence Studies on Dynamic Domains of the Mammalian Fatty Acid Synthase and Different Polyketide Synthases" Master's thesis (2016).^[210]

Maren Julia Berlinghof "Analyse von Domänen-Domänen-Interaktionen in Megasyntasen" Master's thesis (2018).^[223]

Amin Fahim "Modulation einer AT:ACP-Interaktionsfläche der 6-Desoxyerythronolid B-Synthase" Bachelor's thesis (2019).^[256]

Alicia Just "Analyse von Fettsäure- und Polyketidsynthase-Hybriden in Struktur und Funktion" Master's thesis (2020).^[226]

The following figures were used in a modified way, prepared by me, and printed with my permission in Maren Berlinghof's master's thesis.^[223]

- Fig. 3.22 "Correlation of Rosetta predictions with mFAS KS-AT yields."
- Fig. 3.23 "Correlation of Rosetta predictions with mFAS KS-AT melting temperatures."
- Fig. 3.25 "Correlation of Rosetta predictions with DEBS M5 KS-AT yields."

The following figures, prepared by me, were used in a modified way, prepared by Maren Berlinghof, in her master's thesis.^[223]

- Fig. 3.14 "Analysis of mFAS/DEBS M5 KS-AT chimeras."
- Fig. 3.17 "TSA of mFAS KS-AT mutants."
- Fig. 3.18 "HPLC-SEC of mFAS KS-AT mutants."
- Fig. 3.24 "Correlation of Rosetta predictions with mFAS KS-AT oligomeric states."
- Fig. S17 "Correlation of bioinformatics tools predictions with mFAS KS-AT yields. "

Eidesstattliche Erklärung

Ich erkläre hiermit, dass ich mich bisher keiner Doktorprüfung im mathematisch-naturwissenschaftlichen Bereich unterzogen habe. Des Weiteren erkläre ich, dass ich die vorgelegte Dissertation "Modulating Synthetic Pathways in Megasyntases" selbstständig angefertigt und mich keiner anderen 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 und nicht die Hilfe einer kommerziellen Promotionsvermittlung in Anspruch genommen zu haben.

Frankfurt am Main, den

