Structure and Conformational Dynamics of Fatty Acid Synthases

Inauguraldissertation
zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Friederike Maria Carola Benning
aus Deutschland

Basel, 2017

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. Timm Maier
Prof. Dr. Sebastian Hiller

Basel, den 20.06.2017

Prof. Dr. Martin Spiess
Dekan

I Abstract

Multistep reactions rely on substrate channeling between active sites. Carrier protein-based enzyme systems constitute the most versatile class of shuttling systems due to their capability of linking multiple catalytic centers. In eukaryotes and some bacteria, these systems have evolved to multifunctional enzymes, which integrate all functional domains involved into one or more giant polypeptide chains. The metazoan fatty acid synthase (FAS) is a key paradigm for carrier protein-based multienzymes. It catalyzes the *de novo* biosynthesis of fatty acids from carbohydrate-derived precursors in more than 40 individual reactions steps. Its seven functional domains are encoded on one polypeptide chain, which assembles into an X-shaped dimer for activity. The dimer features two lateral reaction clefts, each equipped with a full set of active sites and a flexibly tethered carrier protein. Substrate loading and condensation in the condensing region are structurally and functionally separated from the β-carbon processing domains in the modifying region.

At the beginning of this thesis, only a single crystal structure of an intact metazoan FAS was known. FAS, in particular its modifying region, displays extensive conformational variability, according to electron microscopy (EM) studies. Thus, the aim was to obtain a crystal structure of the FAS modifying region to identify a ground-state structure of the FAS modifying region and to characterize its structural heterogeneity. The second aim was to establish a method for mapping conformational changes in multienzymes at high spatiotemporal resolution. Chapter 1 introduces FAS and gives a methodological overview of studying conformational dynamics of multienzymes.

In chapter 2, the 2.7-Å crystal structure of the entire 250-kDa modifying region of insect FAS is presented. It presents a conserved ground-state conformation adopted by the most divergent member of metazoan FAS. Remarkably, even the V-shape of the central dehydratase dimer is conserved, despite a minimal interface. Structural comparison to polyketide synthases (PKSs) highlights distinct properties of FAS such as strong domain interactions and the absence of an N-terminal β-α-β-α extension of the lateral non-catalytic pseudo-ketoreductase.

Chapter 3 presents a novel approach for identifying conformational dynamics of multienzymes in solution by filming with high-speed atomic force microscopy (HS-AFM) at a spatial resolution of 2-5 nm. The temporal resolution of 10 fps correlates with the timescale of large-scale conformational changes in FAS. Varied viewing angles are provided by combining different molecular tethering strategies. Reference-free particle classification enables the quantitative characterization of conformational states and their transitions.

Chapter 4 discusses the implications of the results on multienzyme biology with respect to biological questions and biotechnological applications. The results of this thesis provide the tools to understand the role of conformational dynamics in multienzymes with respect to their function. These findings will ultimately advance the engineering of enzymatic assembly lines for tailored compound production.

II Table of Contents

I Abstract 5

II Table of Contents 7

III List of Figures 10

IV Abbreviations 11

1 Introduction 15

- **1.1 Carrier proteins enhance catalysis in multistep reactions 16**
- **1.2 Fatty acid synthase catalyzes *de novo* fatty acid biosynthesis 18**
  - 1.2.1 Proteins in bacterial fatty acid biosynthesis 19
  - 1.2.2 Two strikingly different FAS multienzyme architectures in eukaryotes 22
- **1.3 The metazoan FAS 24**
  - 1.3.1 Structural variations of domains among FAS I and II systems 26
- **1.4 Fatty acid biosynthesis is an attractive therapeutic target 27**
- **1.5 FAS and PKS are closely related multienzymes 29**
- **1.6 Structural studies of metazoan FAS 30**
  - 1.6.1 High-resolution crystallographic structure of a mammalian FAS 32
  - 1.6.2 Insect FAS is a divergent member of the metazoan FAS family 32
  - 1.6.3 FAS and PKS have distinct structural properties 33
- **1.7 Large scale conformational dynamics are characteristic for mFAS 34**
- **1.8 Studying structural dynamics 37**
  - 1.8.1 Methods to study conformational variability in multienzymes 37
  - 1.8.2 High-speed AFM filming of macromolecules 39
- **1.9 Aim of the thesis 41**

2 Structural Conservation of the Complete Modifying Region in Metazoan Fatty Acid Synthase 43

- **2.1 Introduction 44**
- **2.2 Results 47**
  - 2.2.1 Identification of divergent metazoan FAS 47
  - 2.2.2 Purification and activity of dFAS-FL and dFAS-DEK 47
  - 2.2.3 Structure determination of dFAS-DEK 48
  - 2.2.4 Catalytic domains 50
  - 2.2.5 Non-catalytic domains 52
  - 2.2.6 Integration of a non-natural expression tag region 54
  - 2.2.7 Dimeric architecture of the insect FAS modifying region 56

2.3 Discussion 59

2.4 Materials and Methods 61

2.4.1 Protein expression and purification 61

2.4.2 Enzymatic activity assay 62

2.4.3 Crystallization of dFAS-DEK 62

2.4.4 Structure determination of dFAS-DEK 63

2.4.5 Structural analysis 64

2.5 Author Contributions 65

2.6 Notes 65

2.7 Acknowledgments 65

2.8 Supplemental Information 66

3 High-Speed AFM Visualization of the Dynamics of the Multienzyme Fatty Acid Synthase 81

3.1 Abstract 82

3.2 Introduction 83

3.3 Results and Discussion 86

3.3.1 Expression and Purification of FAS 86

3.3.2 Imaging FAS Immobilized on Mica 86

3.3.3 Imaging N-/C-Terminally Tagged FAS selectively Tethered to Affinity Lipid Bilayers 88

3.3.4 Movie Frame Alignment and Analysis 90

3.3.5 Temporal Resolution of FAS Imaging and Time Scales of Biological Motions 91

3.3.6 Spatial Resolution of HS-AFM Imaging of FAS 92

3.3.7 Analysis of Conformational Transitions of Single FAS Molecules by 2D Image Classification 93

3.4 Conclusion 96

3.5 Materials and Methods 97

3.5.1 Protein Expression and Purification 97

3.5.2 Enzymatic Activity Assay 98

3.5.3 FAS Immobilization on Mica 98

3.5.4 FAS Tethering to Mica-Supported Lipid Bilayers 99

3.5.5 HS-AFM Imaging of FAS 100

3.5.6 Image Analysis 101

3.6 Associated Content 103

3.7 Author Information 103

3.7.1 Corresponding Authors 103

3.7.2 ORCID 103

3.7.3 Present Address 104

3.7.4 Author Contributions 104
3.7.5 Notes 104
3.8 Acknowledgments 104
3.9 Supporting Information 106
3.9.1 Purification of *D. melanogaster* FAS and enzymatic activity 106
3.9.2 Overview of NADPH turnover rates calculated for the biosynthesis of palmitate in metazoan FAS 107
3.9.3 Tilt correction in *x*- and *y*-direction, filtering and alignment with MATLAB.. 108
3.9.4 Generation of variance maps with MATLAB 108
3.9.5 Autocorrelation of the first and second principal component of HS-AFM images 109
3.9.6 Estimation of the minimum spatial resolution and comparison to simulated AFM images 110
3.9.7 Cantilever preparation with electron beam-deposited carbon tips 112
3.9.8 Force and impulse acting on FAS during HS-AFM imaging 113
3.9.9 Simulation of AFM images 114
3.9.10 Correlation between simulated and experimental data 115
4 Discussion and Outlook 117
4.1 Structural conservation of the modifying region in divergent metazoan FAS 117
4.2 A new approach to study multienzyme conformational dynamics 118
4.3 Hybrid approaches to study dynamics and structures of multienzymes 121
4.4 Conformational dynamics as a basis for understanding assembly lines 122
4.5 Applications of biomimetic nanomachines 124
5 Acknowledgments 129
6 References 131

III List of Figures

Figure 1.1 Conceptual differences between monofunctional and multifunctional enzymes. 17
Figure 1.2 Fatty acid biosynthesis in *E. coli*. 21
Figure 1.3 Multienzyme architectures in eukaryotes. 23
Figure 1.4 Fatty acid biosynthesis in metazoans. 25
Figure 1.5 Fatty acid biosynthesis in cancer. 28
Figure 1.6 Crystal structure of *S. scrofa* FAS at 3.2 Å. 31
Figure 1.7 Large-scale conformational dynamics in metazoan FAS. 35
Figure 1.8 Schematic representation of a HS-AFM. 40
Figure 2.1 Crystal structure of the modifying region of insect FAS. 49
Figure 2.2 Folds and catalytic centers of enzymatic domains in dFAS-DEK. 51
Figure 2.3 Comparison of domain fold and SAM-binding loop in the insect and mammalian ΨME. 54
Figure 2.4 Incorporation of a non-natural tag sequence at the KR C-terminus. 55
Figure 2.5 Comparison of the dFAS modifying region to mammalian FAS and PKS. 58
Figure 3.1 Crystal structure of metazoan FAS (Protein Data Bank (PDB) accession code: 2VZ9, from *S. scrofa*). 84
Figure 3.2 Imaging FAS immobilized on mica. 88
Figure 3.3 Imaging N-/C-terminally tagged FAS selectively anchored to mica-supported lipid bilayers (SLB). 89
Figure 3.4 Dynamics of a single FAS on mica. 91
Figure 3.5 Identification of conformational variety by reference-free 2D classification of FAS imaged on mica in the absence of substrate. 94
Figure 4.1 Identification of conformational variability by HS-AFM imaging. 119
Figure 4.2 Functional dynamics of FAS. 123
Figure 4.3 Applications of multistep reactions. 126

IV Abbreviations

ACC	Acetyl-CoA carboxylase
ACP	Acyl carrier protein
AFM	Atomic force microscopy
Asp	Aspartate
AT	Acetyl transferase
BC	Biotin carboxylase
BCCP	Biotin carboxyl carrier protein
BIIC	Baculovirus-infected insect cells
Bis-Tris propane	1,3-bis(tris(hydroxymethyl)methylamino)propane
B. subtilis	*Bacillus subtilis*
CC	Cross correlation
ChREBP	Carbohydrate-responsive element-binding protein
CMN	*Corynebacterium-Mycobacterium-Nocardia*
CoA	Coenzyme A
CHC	Cuticular hydrocarbons
CPS 1	Carbamoyl phosphate synthetase 1
CPT 1	Carnitine O-palmitoyltransferase 1
Cryo-EM	Cryo electron microscopy
CT	Carboxyl transferase
CV	Column volume
Cys	Cysteine
DEER	Double electron-electron resonance
DEK	Synonym for modifying region (although abbreviation for only DH/ER/KR domains)
dFAS	*D. melanogaster* FAS
DGRC	Drosophila Genomics Resource Center
DGS-NTA(Ni)	1,2-dioleoyl-*sn*-glycero-3[(N-(5-amino-1-carboxypentyl)iminodiacetic acid)succinyl] (nickel salt)
DH	Dehydratase
D. melanogaster	*Drosophila melanogaster*
DMPC	1,2-dimyristoyl-*sn*-glycero-3-phosphocholine

DTT	Dithiothreitol
EBD	Electron beam deposition
EBI	European Bioinformatics Institute
E. coli	*Escherichia coli*
EM	Electron microscopy
EPR	Electron paramagnetic resonance
ER	Enoylreductase
ET	Electron tomography
FAS	Fatty acid synthase
fps	Frames per second
FRET	Förster resonance energy transfer
Gln	Glutamine
HEPES	4-(2-hydroxyethyl)1-piperazineethanesulfonic acid
His	Histidine
HS-AFM	High-speed atomic force microscopy
H. sapiens	*Homo sapiens*
iPKS	Iterative polyketide synthase
kDa	Kilo dalton
Lys	Lysine
MAS	Mycocerosic acid synthase
MsMAS	*Mycobacterium smegmatis* mycocerosic acid synthase
MAT	Malonyl/acetyl transferase
MDR	Medium chain dehydrogenase/reductase family
ME	Methyltransferase
Met	Methionine
mFAS	Metazoan fatty acid synthase
modPKS	Modular polyketide synthase
NADH	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
NMR	Nuclear magnetic resonance
NRPS	Nonribosomal peptide synthetase
KR	Ketoreductase

KS	Ketosynthase
PDB	Protein data bank
PDC	Pyruvate dehydrogenase complex
PEG	Polyethylene glycol
PKS	Polyketide synthase
PP	Phosphopantetheine
PPT	Phosphopantetheine transferase
Pro	Proline
r.m.s.d	Root mean square deviation of atomic positions
SAM	S-adenosylmethionine
SAXS	Small-angle X-ray scattering
SBD	Substrate binding region
SDR	Short chain dehydrogenase/reductase
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SLS	Swiss Light Source
SREBP	Sterol regulatory element-binding protein
S. scrofa	*Sus scrofa*
S. pneumoniae	*Streptococcus pneumoniae*
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride
TE	Thioesterase
TIM	Triosephosphase isomerase
TLS	Translation/Libration/Screw
Val	Valine
ΨME	Pseudo-methyltransferase
ΨKR	Pseudo-ketoreductase

1 Introduction

Proteins are major players in biological processes. Evolution has shaped them to orchestrate a variety of highly specific functions. They perform as signaling molecules, receptors, key controllers in gene encoding and expression, channels, compartmental gatekeepers, structural components of the cell, transporters, folding assistants and force generating machines.[1]

Enzymes are proteins or RNA molecules, which function as biological catalysts. They enhance the rates of biological reactions by reducing the Gibbs free energy of the transition state of a chemical reaction. Electronic and geometric complementarity of the amino acid residues in the substrate binding site leads to a stabilization of the transition state. Highly selective substrate binding is governed by non-covalent interactions such as van-der-Waals forces, hydrogen bonds, electrostatic and hydrophobic interactions. The serine catalytic triad in hydrolases and transferases is an example for covalent catalysis involving the concerted action of active site residues.[2] Nucleophilicity of the attacking serine is enhanced through base catalysis by a nearby histidine, which in turn is stabilized by the carboxylic group of an aspartate.[3] Non-covalent contacts with the peptide backbone exclusively stabilize the tetrahedral intermediates, thereby reducing the activation energy of this reaction step.[4]

Rates of enzymatically catalyzed reactions are typically increased by a factor of 10^6-10^{14}, compared to uncatalyzed reactions.[5] A prominent example for enzyme proficiency is orotidine 5'-monophosphate decarboxylase, which catalyzes an essential step in the biosynthesis of nucleic acids, the carriers of genetic information. The reaction rate is enhanced by a factor of 10^{17} compared to the uncatalyzed reaction, rendering this enzyme crucial for the evolution of life.[6] Contrastingly to chemical catalysts, which often require harsh conditions such as high temperatures, pressures and an extreme pH, enzymes are characterized by mild reaction condition requirements, high substrate selectivity and specificity, which lead to fewer side reactions. Enzymes are regulated by the availability or abundance of reactants including cofactors, allosteric control through binding partners and posttranslational covalent modifications.[5]

1.1 Carrier proteins enhance catalysis in multistep reactions

The majority of multistep biochemical reactions are performed by multiple enzymes which may dynamically interact in transient multiprotein complexes to ensure high local concentrations of substrate at the site of reaction. In some prokaryotes and eukaryotes, these systems have evolved to multifunctional enzymes. As opposed to multienzymatic complexes, multienzymes harbor two or more enzymatic domains on giant polypeptide chains, which often form large oligomeric assemblies. A hallmark of multienzymes is the specific interplay of their domains, allowing efficient substrate transfer between active sites (Figure 1.1). Substrate channeling is associated with enhanced catalytic activity due to ensuring solubility of the unstable intermediate, protection from interaction with non-cognate enzymes and providing a high local concentration of substrate and catalytic residues.[7,8] One such example is the multienzyme carbamoyl phosphate synthetase (CPS 1), in which activation leads to the formation of a 35 Å-long substrate tunnel, enabling the migration of the unstable intermediate between two distant active sites.[9]

The most versatile class of metabolic multienzymatic complexes and multifunctional proteins relies on substrate channeling between active sites by an ancillary carrier protein.[10] Specific protein contacts between catalytic domains and the carrier protein ensure highly selective shuttling of intermediates.[8,10] Covalent substrate binding to the carrier protein occurs *via* a prosthetic group, which is attached to a conserved residue of the carrier protein. In prokaryotes and plants, the majority of carrier protein-dependent systems occur as monofunctional, dissociated enzymes. In multienzymes, however, carrier proteins form an integral part of the polypeptide chain. Flexible tethering to the multienzyme *via* only one to two peptide linkers enables extensive mobility.[10-12]

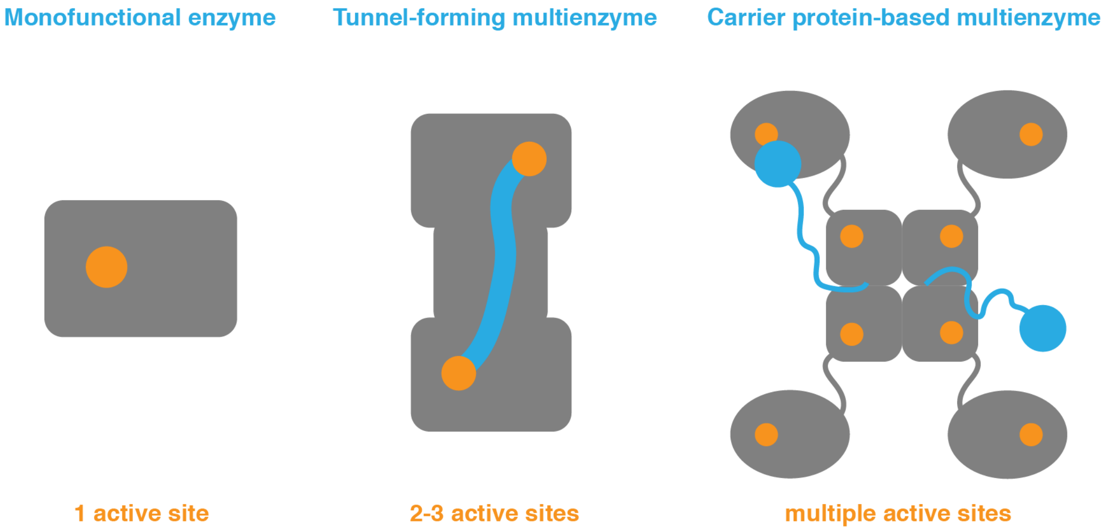


Figure 1.1 Conceptual differences between monofunctional and multifunctional enzymes.
Schematic representations of enzyme systems. Carrier protein-based multienzymes exhibit the highest degree of versatility, as the substrate can be shuttled between multiple active sites. Active sites are depicted as orange filled circles; shuttling systems such as a channel or a flexibly attached carrier protein are colored blue.

In addition to carrier protein-mediated substrate shuttling, domain motions in the ~100 Å-scale are characteristic for most carrier protein-dependent multienzymes including the animal fatty acid synthase (FAS), polyketide synthases (PKS), nonribosomal peptide synthetases, carboxylases and the pyruvate dehydrogenase multisubunit complexes (PDC).[12] This thesis encompasses the structural studies on conformational dynamics of a metazoan fatty acid synthase and the development of an approach, which films these dynamics in multienzymes and enables unbiased, quantitative identification of conformations and their transitions.

1.2 Fatty acid synthase catalyzes *de novo* fatty acid biosynthesis

Fatty acids play multiple important roles in cellular metabolism of all organisms. They (i) are key components of cellular membranes forming the hydrophobic portion of phospholipids, (ii) serve as long-term energy storage molecules in the form of triacylglycerol, (iii) function in cellular signaling as post-translational protein modifiers as well as second messengers and (iv) provide intermediates for the synthesis of vital cofactors.[13] The intracellular pool of fatty acids originates from both, uptake from exogenous sources as well as *de novo* biosynthesis. The *de novo* biosynthesis of fatty acids follows a common cyclic reaction pathway for stepwise, iterative elongation of carbohydrate-derived precursors. It is catalyzed by a collective set of enzymatic activities termed fatty acid synthase (FAS).[14]

While the underlying chemistry is highly conserved among all organisms, two remarkably different FAS systems have evolved. In bacteria, plants, eukaryotic mitochondria and parasites, fatty acid biosynthesis is performed by a highly conserved set of dissociated enzymes, termed type II FAS. Each protein is encoded by a separate gene and catalyzes a single reaction step. Contrastingly, in all other eukaryotes as well as in bacteria of the *Corynebacterium*-*Mycobacterium*-*Nocardia* (CMN) group, *de novo* fatty acid synthesis is carried out by large multifunctional enzymes, termed type I FAS.[15-18]

In the following section, the major players of the conserved reaction cycle of *de novo* fatty acid synthesis are introduced with the *E. coli* type II FAS system, followed by structural comparison of the different FAS systems with a focus on metazoan FAS.

1.2.1 Proteins in bacterial fatty acid biosynthesis

As a main characteristic, all FAS systems rely on an acyl carrier protein (ACP) for substrate shuttling between active sites. ACP is a monomeric 4-helical bundle protein, which requires post-translational modification to be active. ACP is activated by attachment of a 4'-phosphopantheteine (PP) group from Coenzyme-A (CoA) to a conserved serine residue by the 4'-PP transferase holo-ACP synthase (AcpS). Throughout fatty acid synthesis, intermediates are covalently bound to the terminal sulfhydryl group of the prosthetic 4'-PP group of ACP *via* a thioester linkage.[19]

Fatty acids are synthesized by the stepwise, iterative addition of two-carbon moieties to a growing acyl chain. These precursors consist of the priming molecules acetyl-CoA, which is freely diffusing in the cytoplasm, and the elongation substrate malonyl-CoA. Malonyl-CoA is generated by acetyl-CoA carboxylase (ACC) through biotin-dependent carboxylation of acetyl-CoA.[20] The iterative fatty acid biosynthesis cycle can be divided into initiation and elongation. In bacteria, fatty acid synthesis is initiated by malonyl-CoA:ACP transacylase (FabD), which charges the prosthetic 4'-

PP arm of ACP with malonyl-CoA, thereby yielding malonyl-ACP and CoA (Figure 1.2). Subsequently, acetyl-CoA and malonyl-ACP are condensated in a decarboxylative Claisen-type reaction to acetoacetyl-ACP by β-ketoactyl-ACP synthase III (FabH). The NADPH-dependent β-ketoacyl-ACP reductase (KR) FabG catalyzes the reduction to β-hydroxyacyl-ACP, followed by water elimination through the β-hydroxyacyl-ACP dehydratase (DH) FabZ. The resultant enoyl-ACP is further reduced by the NADH-dependent enoyl-ACP reductase (ER) FabI to a saturated acyl-ACP, which serves as a substrate for a next round of chain elongation and modification. FabI plays an important role in completing chain elongation cycles by pulling the equilibrium of the dehydratase reaction towards the β-hydroxyacyl-ACP product.[21]

Type II FAS comprises three distinct ketosynthases (KS) with FabH being only active in the initiation condensation reaction, while FabB and FabF catalyze the condensation of malonyl-ACP with a growing acyl chain during chain elongation. Unsaturated fatty acids featuring a *cis* double bond are generated by FabA, a dehydratase with the additional capability of catalyzing isomerization. Unlike other enzymes in the FAS type II system, enoylreductases are diverse in structure and mechanism. This could be explained by evolutionary adaptation to naturally occurring inhibitors, selectively targeting specific ERs. FabI from *E. coli* and FabL from *B. subtilis* belong to the superfamily of short-chain dehydrogenase/reductase (SDR), whereas the isozyme FabK in *S. pneumoniae* is a TIM-barrel flavoprotein and mitochondrial ER adopts a medium-chain dehydrogenase/reductase (MDR) fold.[22]

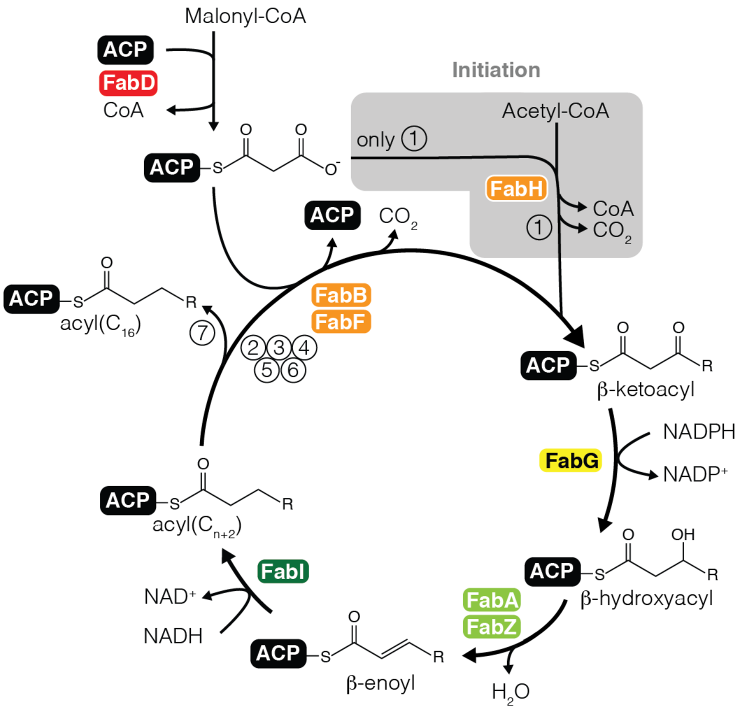


Figure 1.2 Fatty acid biosynthesis in *E. coli*.
Schematic representation of enzymes and reactions involved.

Multifunctional type I FAS exhibit a rather narrow product spectrum and predominantly produce palmitic acid. Contrastingly, dissociated type II FAS synthesizes a diverse range of fatty acid products, featuring both saturated and monounsaturated, as well as branched fatty acids with differing chain lengths. This effect can be ascribed to the carrier protein, which can access other biosynthetic pathways in type II dissociated systems.[23] Moreover, the diverse product spectrum of bacterial FAS is determined by regulation and properties of individual enzymes of the chain elongation cycle. In *E. coli*, the ratio of saturated to unsaturated fatty acids is defined by the interplay of FabA, FabB and FabI.[24,25]

1.2.2 Two strikingly different FAS multienzyme architectures in eukaryotes

Two distinct type I FAS multienzymes have evolved in animals and fungi (Figure 1.3). Metazoan FAS (mFAS) integrates all required enzymatic functions into a ~2500 amino acid polypeptide, which dimerizes to form a functional biological nanomachine. The 540-kDa homodimer adopts an X-shaped structure, with two lateral reaction clefts, each lined by a complete set of active sites. Each set comprises a ketosynthase (KS), a malonyl/acetyl transferase (MAT), a dehydratase (DH), an enoylreductase (ER) and a ketoreductase (KR).[26] In fungal and CMN-bacterial FAS, the functional domains are encoded by two polypeptide chains with a different linear domain organization than in mFAS, which assemble into a 2.6 MDa $\alpha_6\beta_6$ heterododecameric barrel. A central wheel, consisting of six α-chains, divides the barrel into two reaction chambers. Each chamber is capped by a dome of three β-chains and encloses three full sets of active sites. Every set contains an acetyl transferase (AT), an ER, a DH, a malonyl/palmitoyl transferase (MPT), a KR, a KS and a phosphopantetheine transferase (PPT). The PPT is located outside the reaction chamber at the edge of the central wheel.[27,28]

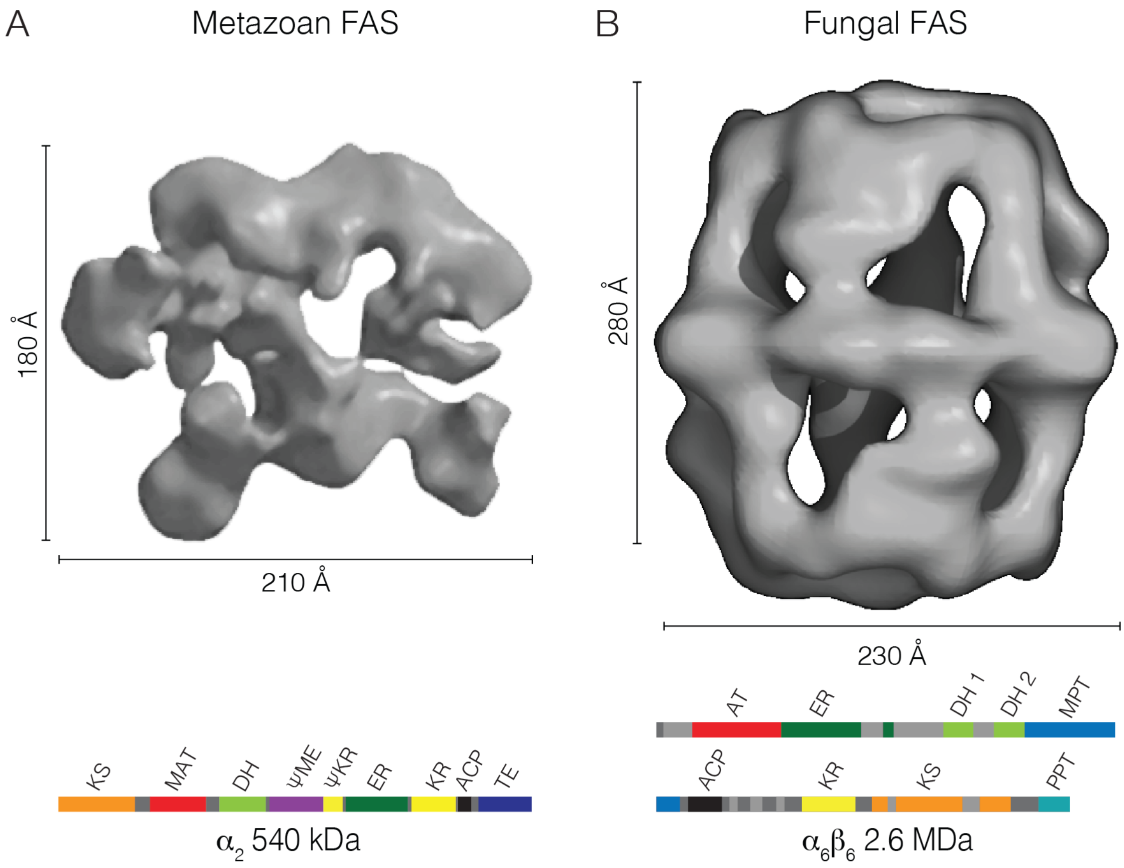


Figure 1.3 Multienzyme architectures in eukaryotes.

(A) 14 Å cryo-electron microscopy (EM) reconstruction of metazoan FAS. All functional domains are integrated into one polypeptide chain, which dimerizes to form a functional enzyme. Figure adapted by permission from Nature Publishing Group from Asturias *et al.*, *Nat. Struct. Mol. Biol.* **2005**, copyright 2005.[29] **(B)** 21 Å cryo-EM reconstruction of fungal FAS. The functional domains are distributed onto two polypeptide chains, which assemble into a heterododecamer. Cryo-electron microscopic map of *S. cerevisiae* FAS was provided by P. Pencek.[30]

With only 9% of the primary sequence serving as unstructured interdomain linkers and an additional 16% forming two lateral non-catalytic domains, mFAS is an inherently flexible multienzyme. Contrarily, the rigid barrel of fungal FAS is stiffened by a structured non-catalytic scaffold, comprising almost 50% of the primary sequence.[11]

In both metazoa and fungi, the ACP forms an integral part of the multienzyme as a flexibly tethered domain. Fungal and CMN-bacterial ACP is doubly tethered to the scaffold, counting three ACPs per reaction chamber. Metazoan FAS however, comprises a singly tethered and thereby extremely flexible ACP domain per reaction cleft. Although it is highly conserved in all organisms, eukaryotic ACP does not sequester acyl chains like its bacterial homolog.[31]

1.3 The metazoan FAS

Fatty acid biosynthesis in metazoan FAS proceeds by a derivative of the common cyclic reaction pathway, predominantly producing palmitate. Unique functional features of animal FAS include a bifunctional malonyl/acetyl transferase (MAT) for substrate loading onto the ACP and the release of the final product palmitate in its free fatty acid form by cleavage through a thioesterase (TE) domain (Figure 1.4).[32,33] Contrarily to type II FAS systems, in which three KS enzymes exist, metazoan FAS contains only one KS domain for the condensation reaction. Subsequent processing of the β-carbon of the acyl chain is carried out by a sole KR, DH and ER enzyme, respectively, whereas dissociated FAS II systems feature multiple enzyme forms with DH and ER activity.

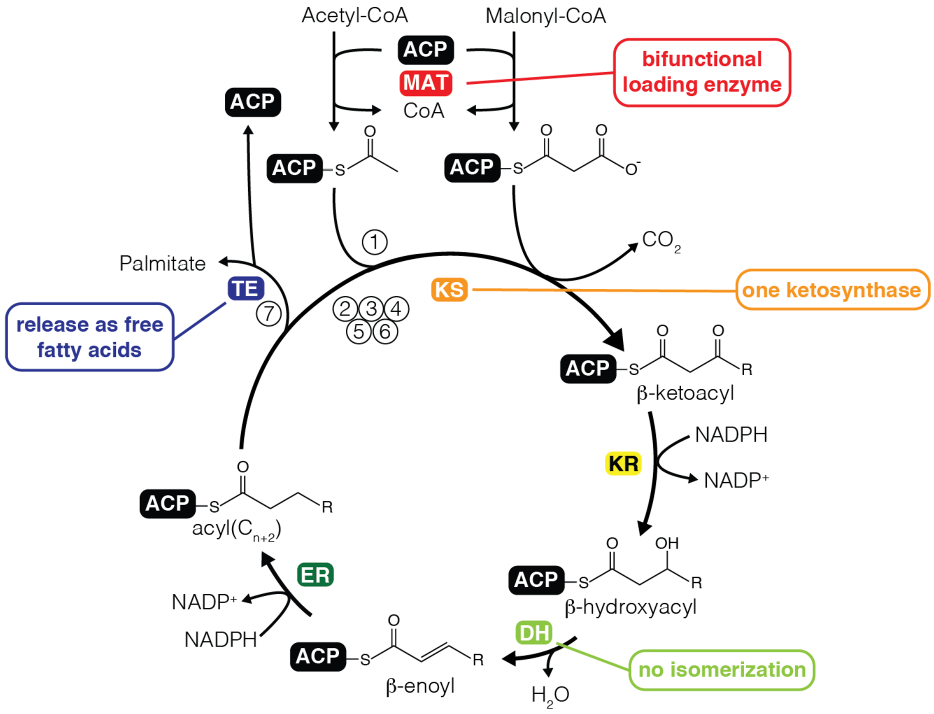


Figure 1.4 Fatty acid biosynthesis in metazoans.

Schematic representation of the reaction cycle and all functional domains involved. Major differences to bacterial fatty acid biosynthesis are highlighted.

Substrate loading by the bifunctional MAT is a random process with a rapid translocation of both acetyl and malonyl moieties between CoA thioester and the sulfhydryl group of the 4'-PP arm of ACP. The bifunctional MAT domain charges the ACP with the starter acetyl moiety from freely available acetyl-CoA. The ACP then transfers the acetyl group to the active site cysteine in the KS domain, before being loaded with the chain extender malonyl unit from malonyl-CoA. The malonyl unit is shuttled to the KS active site and undergoes decarboxylative Claisen-condensation with the KS-bound acetyl.[33]

A major determinant of chain length in mFAS, the thioesterase, selectively hydrolyses C16 or longer acyl-ACP thioesters, thereby releasing free fatty acids. Harboring three hydrophobic pockets of different lengths, it can distinguish short C8 acyl-ACPs from medium C12-acyl-ACPs and longer chains. Only C16/18 acyl-chains are positioned energetically favorable for acyl-chain transfer to the TE active site serine and subsequent hydrolysis of the ester bond. Additional control of the product spectrum is provided by the KS domain, as acyl-ACPs longer than 16 carbons present poor substrates for the elongation reaction.[19]

1.3.1 Structural variations of domains among FAS I and II systems

In contrast to all other catalytic domains, metazoan enoylreductases are structurally unrelated to their functional counterparts in bacterial and fungal FAS. While bacterial ER's (FabI, FabL, FabV)[34-36] belong to the SDR superfamily or adopt a TIM-barrel flavoprotein fold such as fungal ER and FabK, metazoan ER is a member of the medium-chain dehydrogenase/reductase (MDR) superfamily.[37] It is structurally most closely related to the bacterial quinone oxidoreductase. Characteristics of MDR enzymes are a nucleotide-binding Rossmann fold and a substrate binding domain. In mFAS, $NADP^+$ binds between both subdomains of the ER, thereby forming a part of the active site.[38]

Dehydratases in all FAS systems display a remarkable variation of the incorporation of hot-dog folds. While bacterial FabA and FabZ form a homodimer consisting of two hot-dog folds with two equivalent active sites, animal DH comprises a pseudo-dimer of two successive hot-dog fold subdomains with only a single active site. Both the N- and C-terminal hot dog fold subdomain contribute to the active site. Fungal FAS features a third, non-catalytic hot dog fold, which is inserted between the two subdomains comprising the pseudo-dimer.[11]

1.4 Fatty acid biosynthesis is an attractive therapeutic target

Metazoan FAS is highly regulated on a transcriptional level and coupled to the activity of ACC. Sterol regulatory element binding proteins (SREBPs)-1 and 2 and carbohydrate response element binding proteins (ChREBPs) are major transcriptional regulators of several fatty acid metabolic enzymes, including FAS and ACC.[39,40] They respond to glucose and insulin and lead to transcription upregulation under a carbohydrate-rich diet.[41] Cytoplasmic ACC1, which catalyzes the first committed step in fatty acid biosynthesis by generating malonyl-CoA, is allosterically activated by citrate and blocked by long-chain fatty acyl-CoA esters as well as phosphorylation.[39,42] A second homolog, ACC2, is bound to the outer mitochondrial membrane. Localized malonyl-CoA production inhibits the mitochondrial carnitine palmitoyltransferase (CPT 1), thereby preventing mitochondrial fatty acid uptake targeted for degradation by β-oxidation.[43,44]

In humans, elevated levels of FAS expression occur in lactating mammary glands, the liver, adipose tissue, as well as in proliferating fetal tissue.[13] Fatty acid biosynthesis has been closely associated with diseases such as obesity and cancer. While *de novo* lipogenesis is inconsequential in healthy human cells, proliferating cells such as in tumorigenic tissue are dependent on it (Figure 1.5).[45] FAS inhibition blocks metastasis, reduces tumor growth and induces apoptosis in several solid tumors including pancreatic, prostate, breast, renal and lung cancer cells, making fatty acid biosynthesis an attractive therapeutic target. FAS inhibitors in preclinical or clinical trial include inhibitors for the KS, KR and TE domain.[13,45]

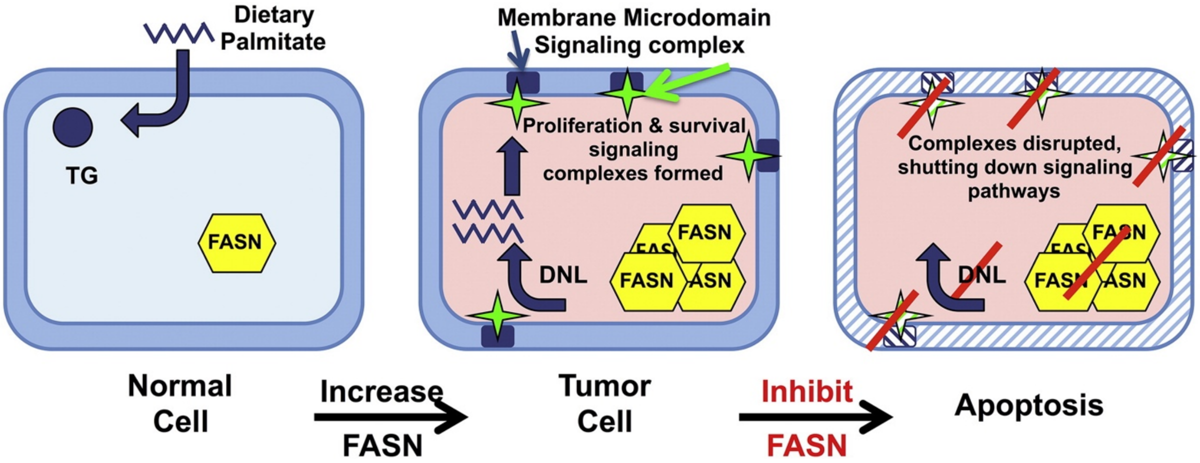


Figure 1.5 Fatty acid biosynthesis in cancer.

Model explaining how FAS overexpression contributes to establishing a tumorigenic and proliferative phenotype in tumor cells. In normal cells, the majority of fatty acids originate from dietary uptake and are stored as triglycerides (TG). FAS overexpression leads to enhanced *de novo* lipogenesis (DNL), generating predominantly saturated membrane components and thereby altering the physical properties of the cellular membrane. New membrane microdomains are formed, which promote the assembly of signaling complexes necessary for proliferation and cell survival (green stars). Inhibition of FAS decreases the amount of saturated fatty acids available. This prevents the maintenance of microdomains and signaling pathways and causes apoptosis. Reprinted with permission from Elsevier from Buckley, D.; Duke, G.; Heuer, T. S.; O'Farrell, M.; Wagman, A. S.; McCulloch, W.; Kemble, G. Fatty acid synthase - Modern tumor cell biology insights into a classical oncology target. *Pharmacol. Ther.* **2017**, *177*, 23-31, copyright 2017, available through DOI 10.1016/j.pharmthera.2017.02.021 and published under the terms of the Creative Commons Attribution-NonCommercial-No Derivatives License (CC BY NC ND, https://creativecommons.org/licenses/by-nc-nd/4.0/).

1.5 FAS and PKS are closely related multienzymes

Metazoan FAS and the type I polyketide synthases (PKS) (i) share a similar overall architecture, involving structurally homologous domains and their linear arrangement, (ii) employ a similar chemistry and (iii) utilize common precursors. Further parallels include the usage of linker regions as well as the presence of two non-catalytic domains, termed pseudo-methyltransferase (ΨME) and pseudo-ketoreductase (ΨKR), in FAS and some PKSs.[11,46,47] However, while FAS exclusively produces fatty acids, PKS systems synthesize a broad range of structurally diverse natural products such as environmental toxins, antibiotics and immunosuppressants.[48] Type I PKSs, which are widely distributed among prokaryotes and eukaryotes, can be classified into iterative PKSs (iPKS) and modular PKSs (modPKS) system. The iterative PKS is characterized by the repeated use of its enzymatic domains for a defined number of reaction cycles. It can be subdivided into non-reducing, partially-reducing and highly-reducing PKSs, based on chemical modification of the elongated chain. In modPKS each cycle of chain elongation is catalyzed by a distinct module and the polyketide chain is passed unidirectionally to a downstream module. Modular PKSs can be categorized into *cis*-AT PKSs, in which the AT domain belongs to the module, and *trans*-AT PKSs, which are characterized by free-standing AT domains, acting *in trans*.[49,50]

In contrast to mFAS, which produces a fully reduced acyl compound, a hallmark of PKSs is the variable utilization and availability of β-processing domains (DH, ER, KR), introducing ketones, hydroxyls and alkenes into the product.[50] Some PKSs contain additional functional domains such as a methyltransferase. Methylated biosynthesis products originate from either incorporation of methylated precursors or from an active methyltransferase domain (ME). Among PKSs, MEs occur in both an active form and an inactive form (pseudo-methyltransferase; ΨME). ΨMEs have been exclusively reported in some highly-reducing iterative PKSs and in FAS, while active MEs can be found in both modular and iterative PKS.[51] Despite low sequence conservation, active MEs and inactive ΨMEs share a conserved structural organization.[51] The functional role of the ΨME is unclear, though it is hypothesized to be a remnant of a common precursor of PKS and FAS which contained an active ME and a full set of DH, ER and KR domains.[11]

The stereochemistry of the β-hydroxyacyl intermediate is determined by specific sequence motifs in the KR domain, resulting in two different substrate entry modes in the KR active site. While mFAS features a so-called B-type KR, producing a *D*-β-hydroxyacyl-ACP, both A-type and B-type KRs exist in PKSs, generating either a hydroxyl-group with *L*- or *D*-orientation, respectively.[49,52]

1.6 Structural studies of metazoan FAS

The X-shaped homodimer of mFAS is functionally and structurally separated into a condensing and a modifying region (Figure 1.6). Its condensing region contains the MAT and KS domains, which are responsible for substrate loading onto the ACP and C-C bond formation. Modifications of the β-carbon of the growing acyl chain are carried out in the modifying region by the DH, ER and KR domains, which are flanked by two non-catalytic domains, ΨME and ΨKR. An 8-residue polypeptide linker serves as the sole connection between both regions. Domain organization in mFAS deviates from linear sequence arrangement. The N-terminal KS domain forms the dimerization interface in the condensing region. The polypeptide extends as a structured linker to the lateral MAT domain and reverts *via* the structured KS-MAT. After proceeding by the core KS, it forms a connection with the DH domain of the modifying region along the dimerization axis. From the central DH domain, the chain continues as a linker to the ΨME domain at the periphery of the modifying region, followed by the ΨKR domain. It then heads back to the ER domain, which forms the major dimerization interface in the core of the modifying region. The ER is followed by a linker to the KR domain, which is located between the DH/ER core of the modifying region and the lateral non-catalytic pseudo-domains. Centered above the reaction cleft, the C-terminus of the KR domain constitutes the anchor point of the linker to the ACP domain. The C-terminal thioesterase is connected to the flexibly tethered ACP *via* an extended polypeptide chain.[38]

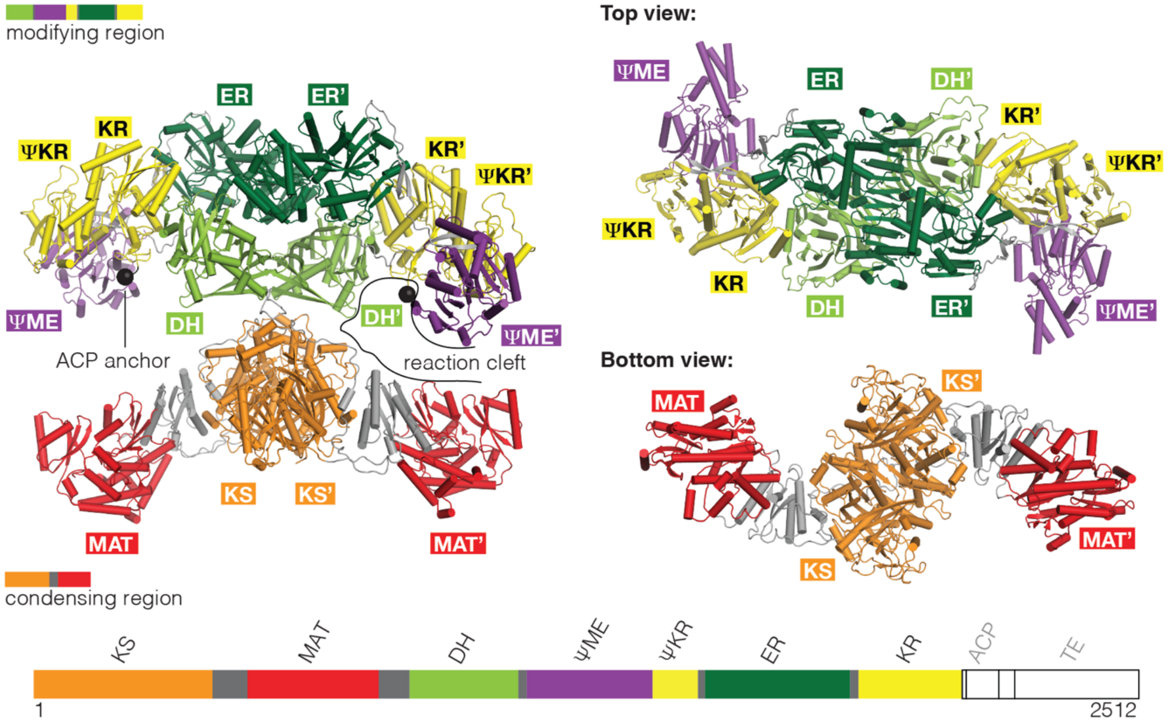


Figure 1.6 Crystal structure of *S. scrofa* FAS at 3.2 Å.

A schematic representation of the linear domain organization is displayed below. The ACP and TE were not resolved due to their inherent flexibility. Black spheres indicate the attachment point of the ACP linker. Figure created using PDB file 2VZ9.

1.6.1 High-resolution crystallographic structure of a mammalian FAS

Its inherent conformational flexibility and size render mFAS a challenging target for high-resolution structure determination. Despite advances in cloning, eukaryotic expression systems, automation of purification strategies and high-throughput crystallization screening as well as single-particle electron microscopy, only one high-resolution structure of the entire modifying and condensing region in mFAS exists today. Fundamental insights into organization of the homodimer, domain arrangement and interaction, cofactor binding as well as active site geometry were provided by the crystal structure of *S. scrofa* FAS at 3.2 Å resolution (apo and in complex with $NADP^+$).[38] The terminal TE and ACP domains are not included in the structure due to missing electron density of these flexibly attached domains. Extensive crystallization screening of natively purified mFAS of various sources, combined with a multi-crystal data collection strategy and additional reduction of solvent background scattering as well as advanced model refinement procedures were prerequisite for the elucidation of a high-resolution structure.[11] Further - albeit smaller - multidomain crystal structures include a *H. sapiens* KS-MAT didomain at 2.15 Å resolution and a portion of the human modifying region, consisting of an inhibitor-bound KR and the two non-catalytic domains at 2.7 Å with the ER being excised from the sequence.[53,54]

1.6.2 Insect FAS is a divergent member of the metazoan FAS family

Despite high sequence conservation among metazoan FAS, comparison of primary FAS sequences in public databases unveils that insect FAS shows some variations in sequence. These include large deletions in the ΨME-domain, which were not discussed so far. While mammalian FAS predominantly produces palmitic acid (C16:0) and only trace amounts of stearic acid (C18:0), the product spectrum of *D. melanogaster* FAS is evenly distributed ranging from myristic acid (C14:0) to stearic acid (C18:0).[55-57] Some studies also report the production of lauric acid (C12:0) at increased ionic strength.[58]

Shorter fatty acids are also attributed to some metazoan glands, where they arise from the action of an ancillary dissociated soluble thioesterase (TE II) with a preference for medium-chain fatty acyl-CoA.[59] However, *D. melanogaster* FAS has been shown to autonomously regulate chain length.[58,60] In insects, fatty acids play an additional key role as precursors of the essential methyl-branched cuticular hydrocarbons (CHC) in the integument, which function as sex pheromones components and protect from lethal desiccation.[61] Unlike in mammals, the *D. melanogaster* genome encodes three distinct FAS homologs. FAS 1 is ubiquitously expressed in the body mass-accumulating larvae as well as in the fat bodies of the adult fly, whereas FAS 2 and FAS 3 are only expressed in adult epithelial tissue such as cuticle, muscle and the CHC-producing oenocytes.[62] This thesis covers structural studies of *D. melanogaster* FAS 1.

1.6.3 FAS and PKS have distinct structural properties

Crystallographic studies of mFAS and PKSs revealed conservation of the condensing region in both systems.[38,53,63] Contrarily, two crystal structures of intact FAS and PKS modifying regions, respectively, uncovered considerable differences.[38,63] In mFAS, domains engage in strong interdomain contacts in the modifying region. The KR acts as a central connector of the core DH/ER domains with the lateral non-catalytic domains.[38] PKSs, however, maintain linker-based connections between the domains of the modifying region.[63]

Further major structural differences to mFAS include an N-terminal β-α-β-α-extension of the ΨKR in PKS and a linear DH-dimer arrangement in PKSs. Dimerization of the DH protomers occurs *via* a large interface, which is partly formed by 20 additional N-terminal residues, which are not included in the FAS DH domain.[11,49,63] Contrastingly, the single crystal structure of metazoan FAS features a V-shaped DH dimer with a miniscule dimerization interface.

1.7 Large scale conformational dynamics are characteristic for mFAS

In the crystal structure of porcine FAS, the mobile ACP is sufficiently long for its prosthetic PP-arm to reach all active sites within one reaction cleft but is unable to make functional contacts with active sites of the other cleft.[38] ACP is tethered *via* a flexible ~12 residue linker to the KR, with the anchor point being centered above the reaction cleft. Distances between active sites amount up to 85 Å, whereas the tethered ACP and its 4'-PP arm can extend to ~90 Å.[11,64] However, both intra- and intersubunit dibromopropanone cross-linking of the ACP 4'-PP and the KS active site serine thiols was observed in FAS dimers.[65] Additionally, mutant complementation of FAS dimers showed that the KS and MAT domains can functionally cooperate with ACP domains of either monomer.[66] Indeed, a wide range of conformations was identified by negative stain electron microscopy (nsEM), enabling the characterization of three major large-scale domain movements in the ~100-Å scale

(Figure 1.7).[29,64] These include in-plane and off-plane rotation between the condensing and the modifying region (“swinging” and “swiveling”), as well as rearrangement of the modifying region. Pendulum-like in-plane swinging of the condensing region relative to the β-processing domains revealed an opening of the reaction cleft of up to 25°. Off-plane swiveling of up to 90° around the connection between the modifying and condensing region, allowing a perpendicular configuration of these regions, was mostly observed in combination with a symmetric modifying region.[64] Offering an ancillary catalytic route, the swiveling motion was estimated by *in vivo* mutant assays to contribute to 20% of the total fatty acid synthesis rate.[66] Reorganization of the modifying region extended from a symmetric to an asymmetric conformation and was shown to require mobility of individual domains instead of one rigid unit. The distribution of conformations was influenced by the presence of substrate, as well as by catalytic mutants, suggesting a correlation to enzyme activity.[64]

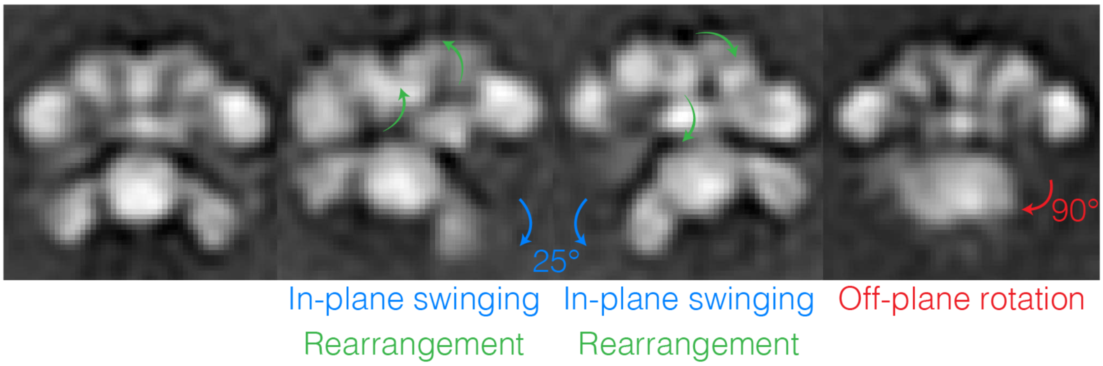


Figure 1.7 Large-scale conformational dynamics in metazoan FAS.
2D class averages depict in-plane swinging, off-plane rotation as well as domain rearrangement of the modifying region. Figure adapted with permission from Nature Publishing Group from Brignole *et al.*, *Nat. Struct. Mol. Biol.* **2009**, *16*, 190-197, copyright 2009.[64]

Large-scale domain motions are observed in most carrier protein-dependent multidomain enzymes, including polyketide synthases (PKS), CoA carboxylases and the pyruvate dehydrogenase multisubunit complex (PDC). Crystallographic structure determination of a structurally-related PKS showed conformationally coupled domain dynamics within the modifying region based on 18 different conformations in the crystal.[63] A hybrid approach of crystallography, EM and small-angle X-ray scattering (SAXS) revealed significant conformational variability in carboxylases. In fungal ACC, phosphorylation transforms the multienzyme from an extended, inactive form into a U-shaped active form, involving large-scale rotations around a central hinge of up to 160°.[67] YCC, a triangular-shaped multienzyme CoA carboxylase in prokaryotes, displays considerable variability in positioning of its lateral biotin carboxylase (BC) domains relative to the central carboxyltransferase (CT) domain.[12] Cryo-electron tomography uncovered in *E. coli* PDC that pyruvate dehydrogenase (E1) and dihydrolipoamid dehydrogenase (E3) subunits are flexibly tethered to the dihydrolipoamid acetyltransferase (E2) core, at varying distances of up to ~150 Å.[68]

Domain movement in several multienzymes is facilitated by linker regions. These are often rich in alanine and proline residues, which predominantly adopt an all-*trans* configuration and thereby provide increased stiffness.[10,69] Linker stiffness is hypothesized to allow both, substantial domain movement and prevent entanglement of domains.[10]

1.8 Studying structural dynamics

This chapter gives an overview of select techniques used for studying conformational dynamics in multienzymes. It further serves to introduce high-speed atomic force microscopy (HS-AFM), one of the two main methods employed in this thesis and mentions advantages and limits of methods required to complement HS-AFM data.

1.8.1 Methods to study conformational variability in multienzymes

Methods employed to study structural dynamics in multienzymes can be grouped into techniques generating mainly static snapshots and those providing temporal information. Static methods include single particle electron microscopy (EM), electron tomography (ET), crystallography, site-specific crosslinking and mutant complementation assays while temporal data is obtained by methods such as electron paramagnetic resonance (EPR) spectroscopy, Förster-resonance energy transfer (FRET)-measurements, super-resolution imaging and high-speed atomic force microscopy. These techniques either produce dynamic distance information (EPR, FRET), 3D structures or medium- to low-resolution films, with several methods being applicable for utilization *in situ*. Combining different techniques provides complementary information on different length and time scales.

Suitable methods for gaining dynamic distance information in multienzymes are EPR and FRET. EPR measurements of spin-labeled proteins reveal dynamic distance information ranging from 8-20 Å for continuous wave EPR and 15-160 Å in double electron-electron resonance (DEER) experiments.[70-72] These distances correspond to dimensions in FAS with domain dimensions of ~20-70 Å and distances between active sites of 85 Å.[11] Unlike in nuclear magnetic resonance, protein size does not restrict the measurements. In-cell EPR studies have been carried out in frozen eukaryotic and prokaryotic cells and include the development of more stable labeling through gadolinium (III) ions or trityl radicals compared to the standard nitroxide-spin labels.[73-75] Dynamic distances of 20-100 Å in and between fluorescently-labeled macromolecules in life cells can be obtained by single-molecule FRET with a temporal resolution in the ms to sub-ms range.[76,77] Using stimulated emission depletion (STED) super-resolution microscopy, fluorescently-tagged proteins can be imaged at a resolution of up to 300-500 Å at frame rates of 28 Hz, enabling their localization in life cells.[78,79] Combining stochastical photoswitching with localizing photon emitters using a doughnut-shaped excitation beam improved the attainable resolution barrier to 60 Å, albeit at lower temporal resolution.[80]

Besides 3D structure determination at the single-Å level, single-particle electron microscopy is a tool for visualizing conformational heterogeneity of macromolecules, as shown for mFAS, ACC and YCC.[12,64,67] Contrastingly to the dynamic methods mentioned above, particles are adsorbed onto a carbon-coated grid and either vitrified or stained with heavy metal salts. Similarly, electron tomography provides a 3D snapshot of cryo-frozen cellular landscapes and is often combined with fluorescent microscopy for localization of targeted proteins.[81]

1.8.2 High-speed AFM filming of macromolecules

In comparison to the methods mentioned in 1.8.1, high-speed atomic force microscopy (HS-AFM) fills the gap as a unique, label-free technique, filming macromolecules at high spatial (10 Å) and high temporal (30-60 ms/frame) resolution in solution.[82,83] As in conventional tapping-mode AFM, a cantilever is oscillated at or near its resonance frequency and scans a substrate surface with intermittent taps (Figure 1.8). Upon contact with the sample, the resulting cantilever deflection is translated *via* a feedback-loop into vertical adjustment of the sample stage to restore the mechanical state of the cantilever to a given set point. Tapping-mode AFM allows imaging of soft biological tissues by controlling scanning forces through modulating the oscillation amplitude of the cantilever.[84] HS-AFM is characterized by dramatically reduced image acquisition times of below 100 ms compared to the 30-60 s in conventional AFM, due to smaller cantilevers with higher resonance frequencies and faster feedback loops.[85] Improvement of the spatial resolution to ~1-2 nm is achieved by microfabricated carbon nanofibers with an apex radius <10 nm grown on top of the cantilever tips.[85] High-speed AFM is considered a non-invasive method, as the impulse exerted by the tip onto the sample lies in the femto-regime. Biological macromolecules imaged to date include myosin V walking on actin, the rotary motion of the F_1-ATPase, the nuclear pore complex and the assembly of a pore forming toxin in a lipid bilayer.[86-89]

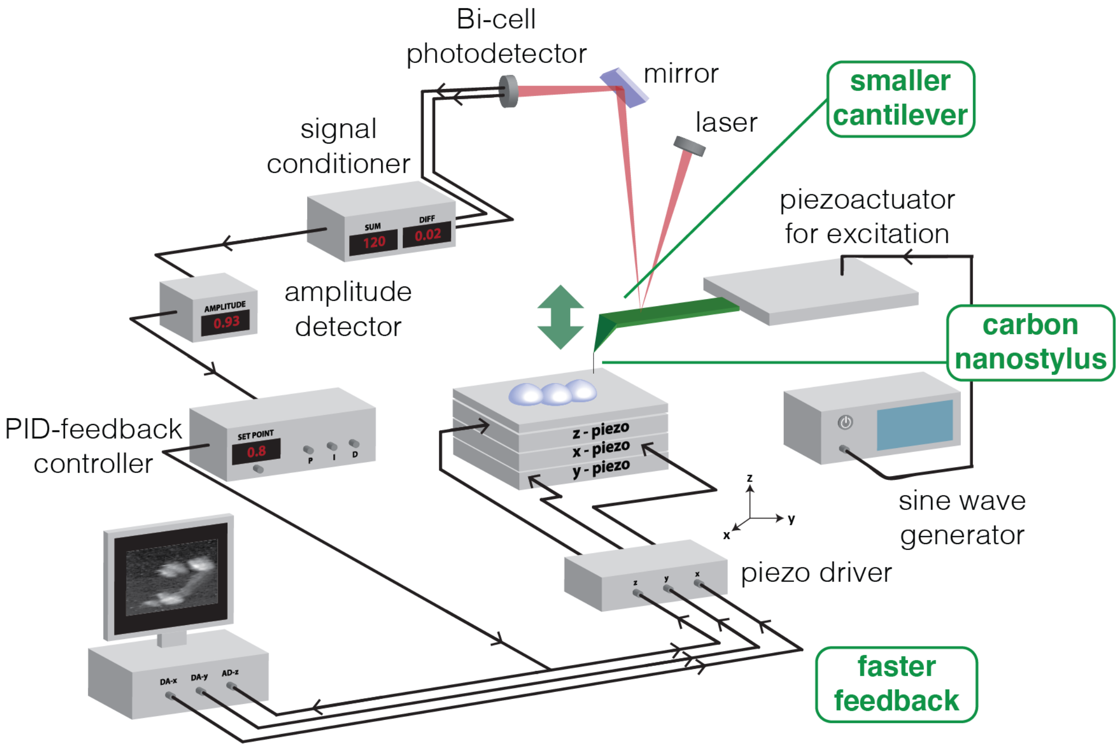


Figure 1.8 Schematic representation of a HS-AFM.

Differences to conventional tapping-mode AFM are highlighted in green. An oscillating cantilever samples the substrate surface. Topographical differences result in deflection of the cantilever. Deflection changes are registered by a laser setup and are converted into electronic signals. Cantilever oscillation is registered by a fast amplitude detector. This input is used to adjust the height of the sample stage *via* PID feedback control to restore the cantilever deflection. A topographical image of the sample is generated from the output of the feedback controller. Here, a HS-AFM 1.0 setup (RIBM, Japan) is depicted. Figure created based on figure from Ando *et al.*, 2014.[90]

1.9 Aim of the thesis

In over ten years, only a single high-resolution structure of an intact metazoan FAS was solved. A crystal structure of the condensing region indicated that this part is structurally conserved in metazoan FAS. For the modifying region, however, which is most variable according to EM and diverges from the closely related PKS modifying region, only a structural snapshot exists. Thus, the first aim consisted of identifying a ground-state structure and its variability of the entire modifying region of a divergent member of the metazoan FAS family and to show how characteristic differences to PKSs are. With the additional objective of providing a structural explanation of severe sequence variations in the ΨME-domain of insect FAS, *D. melanogaster* was chosen as the organism of choice.

Extensive large-scale domain motions and conformational variability have been observed in animal FAS by EM and crystallographic studies. However, only little high-resolution information is available on conformational diversity in metazoan FAS. Moreover, currently no method exists, which enables real-time recording of individual conformations and their transitions in a near-native environment, allowing unrestrained movement of the multienzyme. Are there more conformations than those visualized by electron microscopy? Do specific conformations follow another? Is there a preferred, low-energy conformation? Does the presence of substrate, inhibitors and mutants alter the distribution and transitions of conformations? Therefore, the second aim was to develop a methodological approach, which can provide direct answers to these questions, not only for FAS, but for multienzymes and large, soluble proteins in general.

2 Structural Conservation of the Complete Modifying Region in Metazoan Fatty Acid Synthase

Based on manuscript in preparation:

Structural Conservation of the Complete Modifying Region in Metazoan Fatty Acid Synthase

Friederike M. C. Benning, Habib S. T. Bukhari, Timm Maier

2.1 Introduction

Fatty acids play a vital role in all organisms as lipid components of cellular membranes, energy storage compounds, secondary messengers and covalent modifiers influencing protein translocation. In humans, *de novo* fatty acid biosynthesis is carried out by the cytosolic fatty acid synthase (FAS) and is typically confined to the liver, lactating mammary glands, adipose tissue, proliferating fetal tissue and solid tumors.[13] Proliferation of many solid tumors is blocked by inhibition of fatty acid biosynthesis, making it an attractive target for anti-tumor therapy.[13,45]

De novo fatty acid biosynthesis follows a pathway for iterative elongation of carbohydrate precursors in more than 40 individual reactions, which is conserved with some variations in pro- and eukaryotes. Reaction intermediates are covalently attached to the 4'-phosphopantetheine prosthetic group of an acyl carrier protein (ACP) for shuttling between active sites. Carrier protein-mediated substrate tethering is associated with enhanced catalytic activity based on protecting the hydrophobic intermediate from aggregation, preventing side-reactions with non-cognate enzymes and ensuring a high local concentration of the intermediate.

In metazoans, six different enzymatic activities are involved in fatty acid biosynthesis: For the priming step of the reaction cycle, an acetyl-moiety is loaded from coenzyme A to the thiol group of phosphopantetheinylated ACP by the bifunctional malonyl/acetyl transferase (MAT), followed by transfer to the active site of the ketosynthase (KS). MAT then charges the ACP with a malonyl elongation substrate for decarboxylative condensation to acetoacetyl-ACP in the ketoacyl synthase (KS) active site. In three subsequent reaction steps, the β-carbon moiety is processed by the ketoreductase (KR), dehydratase (DH), and enoylreductase (ER) into a saturated acyl-ACP elongated by a two-carbon unit. The product serves as a primer for the next round of elongation until a chain length of C_{16} is reached and predominantly palmitate is released as a free fatty acid by thioesterase (TE).[33] This multistep reaction is catalyzed by a multifunctional enzyme, which encodes all seven functional domains as one polypeptide chain. It forms a 270-kDa X-shaped homodimer for enzymatic activity.[26,38] The condensation and loading domains are functionally and structurally separated from the β-carbon processing domains, forming a condensing and a modifying region. They define two lateral reaction clefts, each equipped with a full set of catalytic sites and a flexibly tethered ACP at a central position above the reaction cleft (Figure 2.1).

In a static model of metazoan FAS, the tethered ACP is sufficiently long to reach all active sites within one reaction cleft but not in the opposite side cleft. Yet, site-specific crosslinking and mutant complementation studies have indicated an interplay of ACP with active sites of the condensing regions of both clefts.[65,66] Low-resolution electron microscopy (EM) identified a variety of large-scale conformational changes, detecting a weak preference for specific states linked to active site mutations and the presence of substrate.[29,64] Moreover, EM indicated major domain rearrangements in the modifying region of metazoan FAS.[64]

To date, only a single structure of an intact metazoan FAS exists.[38] Together with the structure of a KS-MAT fragment it shows that the condensing region is structurally conserved in metazoans. However, for the part which is most variable according to EM, a second structure is required to evaluate how representative the existing structure is. FAS has been a paradigm for the structure of related multienzymes of the polyketide synthase family (type I PKS; 19% sequence identity). Type I modular

PKSs synthesize a broad range of natural products by processing substrates in modular assembly lines with a distinct set of enzymatic domains for each chain elongation cycle.[91] Product diversity is generated through the variable exploitation of β-processing domains, including additional domains such methyltransferases.[46,51] Crystallographic analysis revealed that the condensing region is related and undisputed in FAS and PKSs.[63] The PKS modifying region, however, has been found to exhibit a distinct domain organization.[63]

Thus, we aimed to determine a ground state structure of the modifying region of a divergent member of the metazoan FAS family, *D. melanogaster* FAS (dFAS) to characterize its variability and to define how characteristic differences to PKSs are. Insect FAS constitutes the most diverging FAS class compared to mammals based on sequence comparison (42% sequence identity). Their product spectrum ranges from myristic acid to stearic acid, whereas mammalian FAS predominantly produces palmitic acid with only minor amounts of stearic acid.[55,57] FAS itself has been shown to regulate chain length, instead of a secondary TE acting *in trans*.[58] Here we provide a high-quality crystal structure of the entire 252-kDa dimeric insect FAS modifying region, which demonstrates a striking and unexpected level of structural conservation compared to intact porcine FAS.

2.2 Results

2.2.1 Identification of divergent metazoan FAS

Sequence alignment identified insect and nematode FAS as the most divergent subclass of metazoan FAS, as compared to the structurally characterized mammalian FAS, with an overall sequence identity of 42%. While most metazoans do not contain any large insertions or deletions relative to mammalian FAS, insects and nematodes exhibit large sequence variations in the ΨME. In insects, the ΨME is significantly truncated by ~100 residues, whereas it contains a ~30 residue insertion in nematodes. As sequence identity of the ΨME of insect FAS is only 14% compared to other subclasses of animal FAS, structural consequences of the deletion as well as its exact location are difficult to predict (Supplementary Figure 2.1).

2.2.2 Purification and activity of dFAS-FL and dFAS-DEK

The modifying region (DEK) of *D. melanogaster* FAS (residues 889-2036; dFAS-DEK), carrying a carboxy terminal PreScission cleavage site, followed by a decahistidine tag, was overexpressed in Sf21 cells and purified with a yield of 5.2 mg per liter of culture. Purification steps included immobilized nickel-affinity chromatography, followed by anion exchange chromatography and a final size exclusion chromatography step. Sample purity was monitored by SDS-PAGE analysis (Supplementary Figure 2.2A). Full-length *D. melanogaster* FAS (residues 1-2438; dFAS-FL), containing the same C-terminal expression tag as dFAS-DEK, was purified according to the same protocol with a yield of 9.6 mg per liter of culture. Formation of long-chain fatty acids in dFAS-FL upon the addition of the substrates acetyl-CoA, malonyl-CoA and NADPH was monitored by measuring NADPH depletion at 340 nm, yielding a specific activity of 926 nmol NADPH per min/mg (Supplementary Figure 2.2B).

2.2.3 Structure determination of dFAS-DEK

Diffraction of initial crystals with unit cell constants of a=75.8 Å, b=174.1 Å and c=197.4 Å was limited to 4.5 Å. Microseeding resulted in crystals in space group $P2_12_12_1$ with unit cell constants of a=75.5 Å, b=174.2 Å, c=194.5 Å. Phases were obtained by molecular replacement using the atomic coordinates of the DH, ER and KR domains (36.4%, 59.1%, 58.1% sequence identity, respectively) of the *S. scrofa* FAS structure (PDB: 2VZ9) as multiple search models. Diffraction data were processed to a resolution of 2.7 Å. 25 residues of the C-terminal non-native expression tag were found to be ordered parts of the structure. The final model thus encompasses residues 893-2055 of the expression construct and was refined to R_{work}/R_{free} of 0.22/0.25 (Supplementary Table 2.1). Residues 1480-1484, 1903-1919, 1951-1957 and 1990-2029 were disordered. Although dFAS-DEK was crystallized in the presence of 5 mM $NADP^+$, no electron density for this cofactor was observed in the ER and KR active sites. Presumably, high bromine concentrations in the crystallization conditions had a competitive effect on $NADP^+$-binding, as partially occupied bromine sites (as based on anomalous signal) overlay with expected binding sites of $NADP^+$ phosphate groups. Contrastingly to the substrate NADPH, which forms part of the substrate binding interface in the ER and KR, $NADP^+$ is known to bind FAS with low affinity. Previous crystal structures of $NADP^+$-bound FAS required soaking at very high concentrations of $NADP^+$ (20 mM).[38,54]

Domain arrangement in the modifying region of dFAS-DEK differs from linear sequence organization (Figure 2.1). In the protomer, the polypeptide chain first folds into the N-terminal DH domain to then extend as a mostly unstructured 19-residue polypeptide stretch to the peripheral ΨME domain. From here, it proceeds through the adjacent pseudo-ketoreductase (ΨKR) domain, followed by a 17-residue linker leading back to the ER domain, which is situated above the DH domain. A 19-residue linker connects the ER domain with the KR domain, which is located between the ER domain and the two lateral non-catalytic domains.

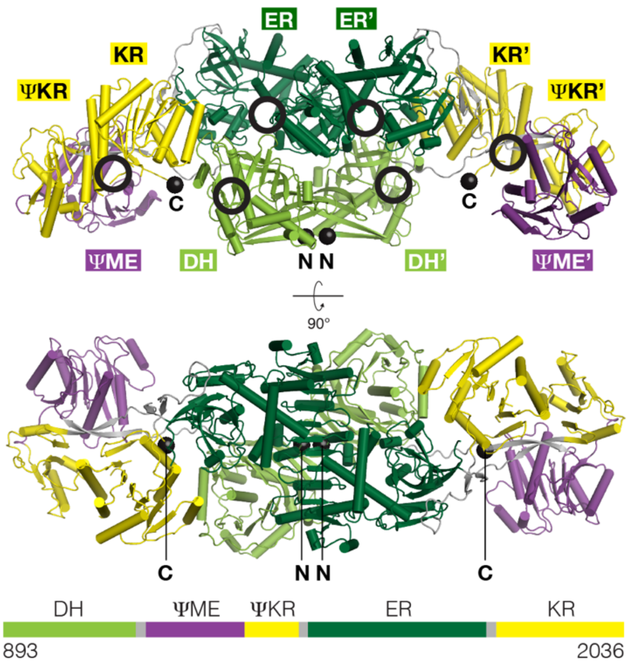


Figure 2.1 Crystal structure of the modifying region of insect FAS.
The linear domain organization is schematically displayed below. Active sites are indicated by black rings and the N- and C-termini are denoted as black spheres.

2.2.4 Catalytic domains

The DH domain in insect FAS adopts a pseudo-dimeric double hot dog fold, consisting of two six-stranded antiparallel β-sheets each wrapping around a five-turn α-helix. Each pseudo-dimer contains a single active site formed by the cooperation of His911 from the N-terminal and Asp1062 and Gln1066 from the C-terminal hot dog fold (Figure 2.2A). An extended 26-residue long linker connecting DH1 and DH2 wraps around the β-sheet back of the two hot dog folds.

The NADPH-dependent enoylreductase belongs to the medium chain dehydrogenase/reductase family (MDR).[37] It is composed of a central nucleotide-binding Rossmann fold (residues 1578-1721) and a substrate binding region (SBD) comprising the domain termini with an antiparallel β-sheet in the core. The Rossmann fold consists of a 6-stranded parallel β-sheet, which is flanked by three α-helices on each side. Both subdomains are separated by a cleft for cofactor and substrate binding including the catalytic residues (Figure 2.2B). Binding of NADPH, which provides part of the substrate binding interface, has been shown to stabilize the ER active site.[38] Loop 1480-1484 at the solvent-exposed edge of the substrate-binding subdomain is disordered in the dFAS-DEK structure.

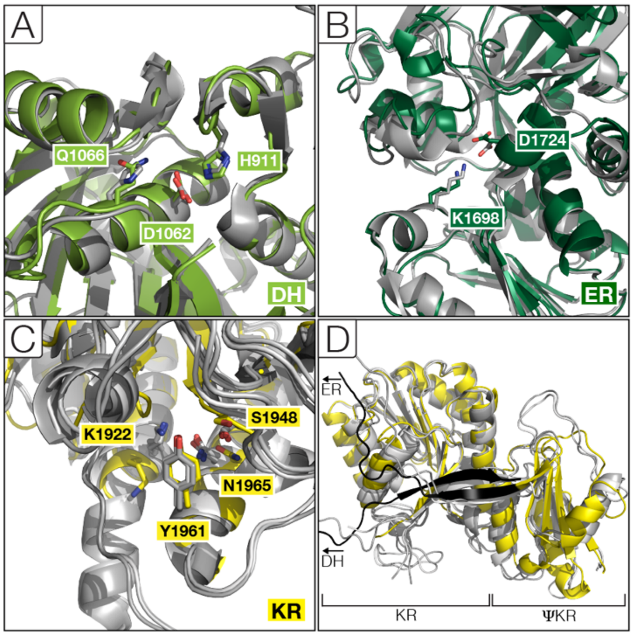


Figure 2.2 Folds and catalytic centers of enzymatic domains in dFAS-DEK.
Domains are superposed onto existing crystal structures of metazoan FAS. **(A)** Active site conservation of insect (green) and porcine DH (gray), aligned with a Cα r.m.s.d. of 1.2 (Supplementary Table 2.1). **(B)** Catalytic residues Lys1698 and Asp1724 of the ER domain are conserved in metazoan FAS (insect: dark green; pig: gray), aligned with 1.0 Cα r.m.s.d.). **(C)** KR active site residues, which are involved in proton relaying upon hydride transfer to the substrate from NADPH.[38,54,92] Structural alignment with human (light gray; 1.3 Cα r.m.s.d.) and porcine (dark gray; 1.2 Cα r.m.s.d.). **(D)** Fold conservation of the KR and ΨKR domains in insect (yellow), porcine (dark gray) and human FAS (light gray). PDB accession codes used were 2VZ9 for porcine FAS and 4PIV for the human KR/ΨKR/ΨME tridomain.

As a member of the short chain dehydrogenase/reductase (SDR) family, the NADPH-dependent ketoreductase domain harbors a nucleotide-binding Rossmann fold with a substrate binding extension preceding $\alpha 6$ and $\beta 7$. Formation of the KR active site requires binding of the nucleotide cofactor and the actual substrate.[38,54] Identified as a residue of the proton relay in previous work, Lys1922 is disordered in the dFAS-DEK structure.[38,92] Disordered residues include loops 1903-1919, 1951-1957 and 1990-2029, with the latter comprising the substrate binding extension according to structural superposition onto porcine KR. The unnatural 25-residue expression tag is incorporated as the C-terminal α6 and β7 of the Rossmann fold, despite the presence of the natural sequence (see Results and Supplemental Information).

The overall domain fold and active site geometries are conserved in the dFAS-DEK structure, compared to the metazoan FAS structure (Figure 2.2A-C). This finding was expected due to the limited differences in product spectrum. The DH, ER and KR domains are likely not involved in determining the different output, highlighting the relevance of the TE in combination with reduced KS turnover for long acyl chains for chain length determination.[93,94]

2.2.5 Non-catalytic domains

The non-catalytic ΨKR in dFAS-DEK adopts a truncated KR fold, consisting of a central five-stranded parallel β-sheet and four α-helices (Figure 2.2D). The last strand of the β-sheet is antiparallel and merges into the structured portion of the linker, which connects the ΨKR with the ER. Together with the antiparallel two-stranded β-sheet formed by the DH-ΨME and the ΨKR-ER linkers, the ΨKR mimics one of the major dimerization interfaces of tetrameric bacterial KR, thereby stabilizing the KR active site.[38]

The ΨME in metazoan FAS is structurally related to SAM-dependent methyltransferases.[11] These contain a D/ExGxGxG motif, which is involved in SAM-cofactor binding and is strictly conserved in several iterative and modular PKSs with methyltransferase activity.[38,49,95] Structural alignment of metazoan FASs indicates a conserved ExxxGxG motif in the ΨME domains of mammalian FAS on a loop connecting β1 with αA, whereas in insects, only the glutamate is conserved.[49] Moreover, unlike in other animal ΨME domains, the mutated SAM-binding motif in insect FAS contains a species-dependent 1-4 residue insertion, resulting in a minor conformational change of the loop (Figure 2.3A). These mutations correspond to the inability of the ΨME to bind SAM and to methylate substrates.[11]

The dFAS-DEK structure explains the severe sequence variations in insect ΨME in comparison to mammalian FAS. It reveals that the insect ΨME domain comprises only the core SAM-dependent methyltransferase fold, encompassing a 7-stranded β-sheet with an antiparallel β7, flanked by two and three helices on each side (Figure 2.3B).[38,95] Four additional helices (~20 Å long), protruding from the methyltransferase fold in the structures of porcine and human FAS, are replaced by a 6- and a 3-residue loop in *D. melanogaster* FAS. This leads to a significant shortening of the extension of the ΨME to 35 Å compared to 55 Å in mammalian FAS. It was speculated before that the non-catalytic part of the modifying region may provide steric guidance to the ACP.[11] However, the dFAS-DEK structure suggests that ACP movement is independent of the ΨME, thus rendering it dispensable.

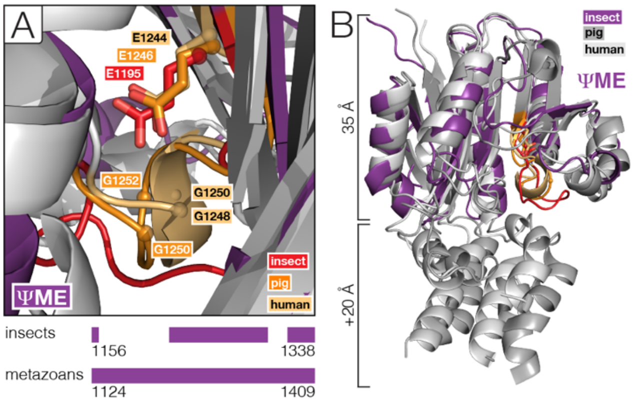


Figure 2.3 Comparison of domain fold and SAM-binding loop in the insect and mammalian ΨME.

(A) Insect ΨME is truncated by four ~20 Å helices, which are present in any other metazoan FAS (Cα r.m.s.d. of 2.0 Å for porcine ΨME and 1.8 Å for human ΨME). **(B)** While the inactive SAM-binding motif is conserved in human and porcine FAS, it adopts a different fold in insect FAS and only contains the conserved glutamate. A schematic representation of ΨME structural alignment is displayed below.

2.2.6 Integration of a non-natural expression tag region

The KR C-terminus of the dFAS-DEK construct was designed to be shortened by three terminal amino acids relative to the last ordered residues in the structure of *S. scrofa* FAS (Figure 2.4A). Unexpectedly, at the C-terminus of the KR, the non-natural expression tag is incorporated into the protein fold, rather than the last 22 residues of the natural sequence, despite its confirmed presence in the sample. Instead, the natural C-terminus up to residue Val2032 and the preceding substrate binding extension loop are disordered in this crystal structure. The last four residues of the natural sequence, Val2033-Met2036, are displaced relative to their cognate position and adopt the position of residues Met2015-Cys2018. They are followed by incorporation of the non-natural tag region, which displaces residues 2019-2036 from the natural sequence. Starting from residue Val2033, this shifted region adopts a highly similar fold as the regular C-terminus in the available structures of porcine KR and human KR (Figure 2.4, Supplementary Figures 2.3-2.5). It consists of a helix, followed by a turn and a β-strand, which constitutes part of the Rossmann fold. Contrastingly, the 6-residue turn features a more pronounced bulge in dFAS-DEK, than the 3-residue loop in pig (Figure 2.4C, Supplementary Figure 2.3).

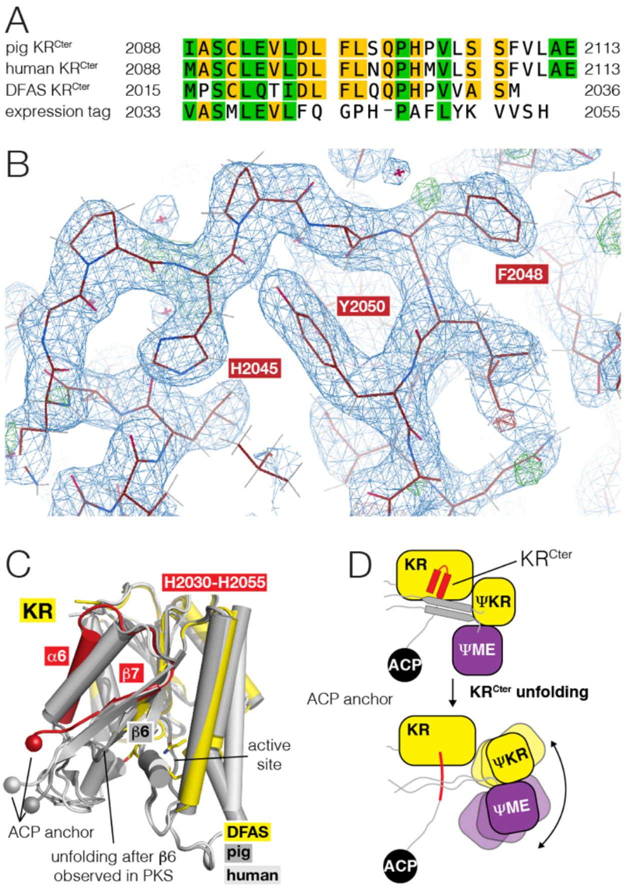


Figure 2.4 Incorporation of a non-natural tag sequence at the KR C-terminus.

(A) Structural alignment of the C-terminal 25 residues of the KR in pig, human and fruit fly, as well as the incorporated expression tag. The fruit fly construct was designed to be truncated by the last three residues relative to the terminal ordered residues in the structure of *S. scrofa* FAS. These sequences would continue into the ACP linker in full-length FAS enzymes. **(B)** Electron density map at 1.5 σ showing unambiguous density for the expression tag. Select residues are highlighted for clarity. **(C)** Structural alignment of porcine, human and insect KR indicates that the unnatural C-terminus adopts a similar fold as in mammals, containing an additional bulge after α6. **(D)** Possible consequences of C-terminal KR unfolding include an extension of the KR-ACP linker, release of the DH-ΨME and ΨKR-ER linker, potentially resulting in a collapse of the peripheral arms of the modifying region.

Sequence alignment shows the conserved position of a glutamine (Gln2042 in dFAS, Gln2101 in mammals) and a proline (Pro2046 in dFAS, Pro2101 in mammals), which are involved in formation of the turn (Figure 2.4A and Supplementary Figure 2.1). A low B-factor distribution within the KR suggests rigid incorporation (Supplementary Figure 2.4). Interface analysis of the terminal helix-turn-strand with the KR β-sheet using PDBpisa predicts a free solvation energy upon interface formation (ΔG_{sol}) of -17.4 to -17.8 kcal/mol for the non-natural C-terminus and -15.8 to -17.1 kcal/mol for a model containing the original sequence. In the human and porcine KR structure, these interactions are slightly more favorable with a ΔG_{sol} of -21.8 and -20.6 kcal/mol, respectively. Further, contacts within the terminal helix-turn-sheet fold of the shifted sequence (ΔG_{sol} of -3.5 kcal/mol for chain A and -4.2 kcal/mol for chain B) correspond to values for the KR C-terminus in the human and porcine structure (ΔG_{sol} = -3.8 kcal/mol). These interactions are predicted to be more energetically favorable than those in a model of the natural dFAS KR C-terminus (ΔG_{sol} = -1.1 kcal/mol; Supplementary Tables 2.4-2.6).

2.2.7 Dimeric architecture of the insect FAS modifying region

In the dFAS-DEK structure, the major dimerization interface in the homodimeric modifying region is formed by the central ER domains (1100 Å^2 interface area in dFAS) by homophilic interactions of the β6- and β7-strands of the Rossmann fold. Smaller contributions to the dimer interface are provided by hydrogen bonding between the β1- and β2-strands of the N-terminal hot dog folds in the central DH domains (400 Å^2) and interactions between the N-terminus of the ER and the C-terminus of the DH (Supplementary Table 2.3). The DH dimer adopts a V-shaped conformation with an angle of 96°.

While the central core of the modifying region is mainly stabilized by the dimerization interface, additional direct domain interactions occur between the DH and ER within the monomer. Fundamental contacts with all domains within the dFAS-DEK monomer are provided by the KR, with approximately 15% of the KR surface area being engaged in domain connectivity. The largest interfaces are formed with the ΨKR and DH, respectively, while the interaction to the ER is comparatively smaller. The peripheral ΨME domain is mainly connected to the ΨKR with marginal interactions to the KR.

Three polypeptide stretches connect the central core with the lateral KR/ΨKR/ΨME domains. The majority of the linkers are extended, solvent-exposed polypeptide chains. However, a small part of the DH-ΨME and the ΨKR-ER linker forms an antiparallel, two-stranded β-sheet, which is buried between the KR, ΨKR and ΨME. Two linkers, the DH-ΨME and the ER-KR linker are particularly rich in alanine and proline (26% and 32%, respectively).

Superposition of the modifying region of *S. scrofa* (PDB 2VZ9) and the dFAS-DEK structure revealed that both modifying regions were crystallized in the same conformation with an overall Cα r.m.s.d. of 1.5 Å (Figure 2.5A). However, in comparison to porcine FAS, the N-terminal part of the substrate binding domain (SBD) in insect ER is widened by 2-3 Å. The KR, directly interacting with the repositioned N-terminus of the ER-SBD, is shifted towards the periphery by ~4 Å, which in turn affects the position of the ΨKR and ΨME. The DH is rotated into the central core along an axis through the major dimerization interface of the modifying region by 2-3 Å.

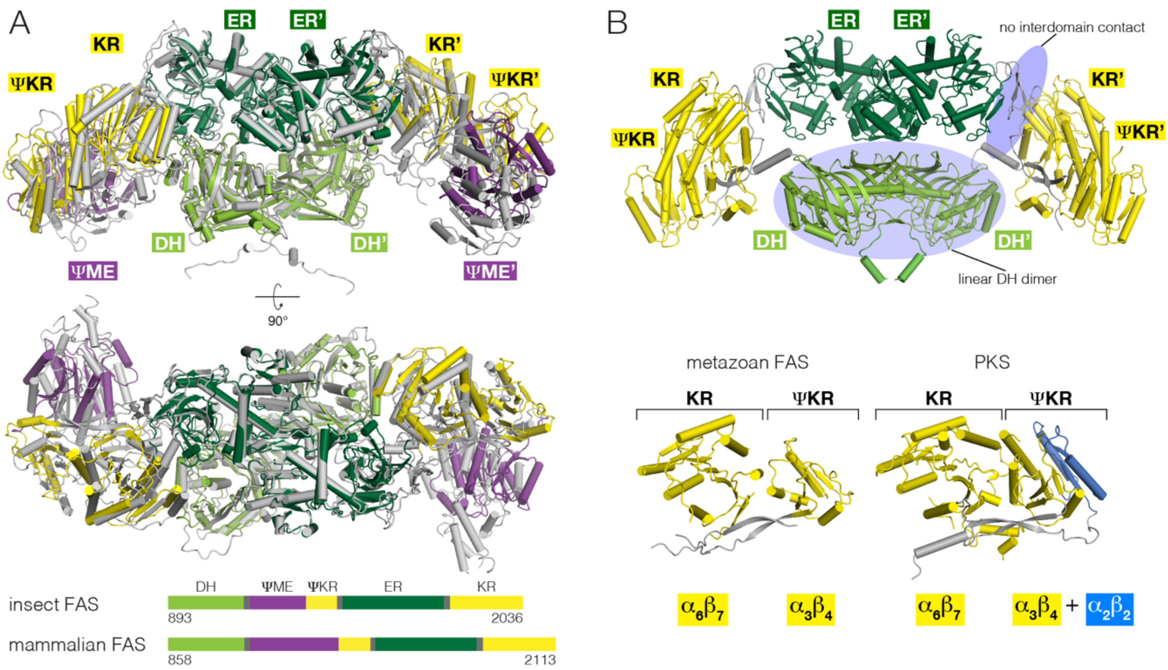


Figure 2.5 Comparison of the dFAS modifying region to mammalian FAS and PKS. **(A)** Structural alignment onto porcine FAS (Cα r.m.s.d. 1.5 Å). The domain organization of insect and mammalian FAS is schematically displayed below. **(B)** Despite a similar overall architecture and conserved monomer folds, PKSs (PDB 5BP4) exhibit a distinct domain organization in their modifying domain. Contrastingly to FAS, PKSs feature a linear DH dimer instead of a V-shaped conformation. Moreover, the PKS ΨKR is equipped with an N-terminal $\alpha_2\beta_2$-extension. While the FAS modifying region is connected by interdomain contacts, the organization of the PKS modifying domain is based on linker connections.

Individual domain folds were structurally aligned and compared to those of porcine FAS, a human ΨME/ΨKR/KR tridomain (PDB 4PIV) and a human KR domain (PDB 5C37). Domain folds and active site geometries are structurally highly conserved, except for the significantly shorter ΨME-domain of insect FAS. Deviations are smallest for dimerizing ER domains (1.0 Å Cα r.m.s.d.) and highest for the ΨKR (2.1 Å Cα r.m.s.d.) and ΨME (1.8 Å Cα r.m.s.d.) (Supplemental Information).

2.3 Discussion

Here, we present the crystal structure of the entire modifying region of *D. melanogaster FAS* at 2.7 Å. Often, multidomain proteins crystallize in distinct conformations due to their extensive domain interplay, enabling a wide range of conformational sampling.[67,96] This structure shows that even distantly related FAS modifying regions adopt almost identical ground-state structures. The absence of any random conformational variations, as observed for metazoan FAS in electron microscopy studies, suggests that the occurrence of these may possibly be coupled to distinct enzymatic states.[29,64] Remarkably, even the characteristic V-shaped DH dimer is preserved in metazoan FAS despite a miniscule dimerization interface and the absence of linkage to the dimeric condensing region. It therefore appears to be a general inherent structural property of the animal FAS modifying region.

An artificial expression tag is integrated into the β-sheet of the KR in this structure. As the incorporation of alternative regions is not usually observed for β-sheets but has now twice been seen for carrier protein-dependent multienzymes at the same location, it may indicate the presence of a dynamic hinge.[97] This structure of *D. melanogaster FAS* explains substantial sequence variations in the ΨME of insects compared to mammals. Activity assays of full-length insect FAS confirm similar turnover rates for the formation of palmitic acid as mammalian FAS. Previously thought to add spatial restraints to the reaction cleft, this structure invalidates its role as steric guidance for the ACP.

Although structurally related, the modifying regions in animal FAS and PKS display distinct properties. FAS maintains strong interdomain connections between all domains, with the KR functioning as a central connector between the core dimerizing ER and DH domains and the lateral non-catalytic domains. The two DH protomers form a V-shaped dimer with only a small dimerization interface. Contrastingly, PKSs exhibit a linear DH-dimer with a significantly larger dimerization interface including handshake interactions between the DH protomers. The PKS ΨKR contains a lateral β-α-β-α extension at its N-terminus (Figure 2.4B), which is absent in FAS ΨKR. Unlike in FAS, domains in the PKS modifying region are connected *via* a linker-based organization with minimal interdomain contacts, facilitating conformational flexibility.[63]

Here, we report the first structure of an excised intact modifying region of metazoan FAS. Although it belongs to the most divergent subclass of metazoan FAS, insect FAS adopts the same conformation as mammalian FAS, suggesting a low-energy conformational ground state. Distinct structural characteristics of FAS in comparison to the structurally-closely related PKS are conserved, including a V-shaped DH-dimer arrangement and direct interactions between the central core and the KR domain. We further provide a structural explanation for the truncated non-catalytic ΨME domain in insect FAS, questioning its role as a steric barrier for ACP. In addition, we observe the highly rare event of incorporation of an artificial expression tag in the presence of the natural sequence into the Rossmann-fold at the C-terminus of the KR domain. Being the sole attachment site of the flexible ACP tether, these findings suggest a more dynamic role for the KR-ACP anchor as anticipated and provide a basis for establishing a specific tagging system in FAS.

2.4 Materials and Methods

2.4.1 Protein expression and purification

Full-length *D. melanogaster* FAS isoform 1A (Uniprot identifier: Q9VQL7; FAS) was constructed from fragments GH17750 (GenBank: BT050523.1, obtained from DGRC, USA) and GH07627 (GenBank: AF145643.1, obtained from DGRC, USA) by Red/ET-mediated homologous recombination, serving as a PCR template for generation of the DEK construct (residues 889-2036).[98] All variants were cloned into a modified, Gateway-compatible version of the baculovirus transfer vector pACEBac1 (Geneva Biotech, CH), coding for an amino (pAB1GN-His10)- or a carboxy-terminal (pAB1GC-His10) decahistidine tag (His_{10}) to yield amino His_{10}-tagged full-length dFAS (dFAS-FL) and carboxy His_{10}-tagged dFAS-DEK.

Protein expression was carried out in Sf21 insect cells (Expression Systems, USA) in Insect-XPRESS medium (Lonza, CH) at 27°C using the baculovirus expression system.[99] Both constructs were expressed and purified individually. Virus was stored according to the titerless infected-cells preservation and scale-up method as baculovirus-infected insect cells (BIIC).[100] Cells were harvested 72-90 h post-infection by centrifugation and lysed by sonication in 50 mM HEPES pH 7.4, 200 mM NaCl, 20 mM imidazole, 10% glycerol and 10 mM β-mercaptoethanol. Lysate was cleared by ultracentrifugation. Soluble protein was purified by affinity chromatography using High Affinity Ni-Charged Resin (GenScript, USA) and eluted with a linear gradient to 300 mM imidazole. The eluate was diluted to 50 mM NaCl and subjected to anion exchange chromatography using a PL-SAX 4000Å 10 µm resin (Agilent, USA). For dFAS-FL, a 5-column volume (CV) linear gradient to 500 mM NaCl was applied, whereas dFAS-DEK was eluted by a 20-CV linear gradient to 200 mM NaCl. Both, dFAS-FL and dFAS-DEK were applied onto a Superdex 200 size exclusion column (GE Healthcare, USA), equilibrated in crystallization buffer (20 mM HEPES pH 7.4, 250 mM NaCl, 5% glycerol and 5 mM DTT). Purified dFAS-FL was concentrated to 12 mg/mL and dFAS-DEK to 20 mg/mL, respectively, flash-frozen in liquid nitrogen and stored at -80°C.

2.4.2 Enzymatic activity assay

FAS activity was measured by spectrophotometrically (Synergy H1 Hybrid Reader, BioTek, USA) monitoring the decrease in absorbance at 340 nm due to NADPH oxidation in the overall reaction of long-chain fatty acid synthesis.[101] Following incubation of 20 nM dFAS-FL with 40 µM acetyl-CoA and 150 µM NADPH in 50 mM potassium phosphate buffer at pH 7.0 at 25°C for 5 minutes, malonyl-CoA was added to a final concentration of 120 µM and recorded for 40 minutes. Measurements were conducted with four replicates per condition.

2.4.3 Crystallization of dFAS-DEK

Insect FAS modifying region, dFAS-DEK, was crystallized at 18°C by sitting-drop vapor diffusion. Initial 100 µm-long crystal plates, diffracting to ~4-5 Å resolution, appeared after two days with 10 mg/mL dFAS-DEK in 0.1 M Bis-Tris propane, 0.2 M sodium acetate and 20% PEG 3350 at pH 7 and 0.1 M Bis-Tris propane, 0.2 M sodium bromide and 22.5% PEG 3350 at pH 6.5, both in the presence of 1.5 mM $NADP^+$. After scaling-up, dFAS-DEK was crystallized at 20 mg/mL using microseeding of 100-fold diluted initial crystals and a protein to reservoir (0.1 M Bis-Tris propane pH 6.5, 0.35 M NaBr, 17.5% PEG 3350 and 10 mM $NADP^+$) to seed ratio of 2:1.6:0.4 in a total drop volume of 4 µL. Plate-shaped crystals appeared after 12 hours and continued to grow for 3 days to a size of ~400x200 µm. Crystals were dehydrated and cryo-protected in a stepwise protocol using reservoir-like stabilization solution at linearly increasing concentrations of ethylene glycol before vitrification in liquid nitrogen. The final solution contained 0.1 M Bis-Tris propane pH 6.5, 0.35 M NaBr, 22.5% PEG 3350, 20% ethylene glycol.

2.4.4 Structure determination of dFAS-DEK

X-ray diffraction data was collected at beamline X06DA (PXIII) at Swiss Light Source (SLS, Paul Scherrer Institute, Villigen, Switzerland) using a PILATUS detector at 100 K. A total of 4 sets, containing 720 frames each, were collected from two dFAS-DEK crystals at a wavelength of 1.00003 Å and a rotation angle of 0.25°. Raw data was processed and merged utilizing XDS.[102,103] Phases were obtained by molecular replacement in PHASER, using the individual atomic coordinates of the DH, ER and KR domains of the *S. scrofa* FAS structure (PDB: 2VZ9) as multiple search models. An initial round of model refinement was performed using REFMAC5 (CCP4 program suite)[104,105], followed by density modification with parrot[106] and model autobuilding by buccaneer.[107,108] The initial model was refined by iterative cycles of manual model building in coot (Version 0.8.8),[109] supported by feature enhanced maps generated by phenix.fem,[110] and refinement in BUSTER.[111] For final fine-tuning, refinement was performed with phenix.refine.[112-115] Figure preparation was carried out in PyMOL (Schrödinger LLC).

2.4.5 Structural analysis

Structure analysis was performed using the protein structure comparison service PDBeFold at European Bioinformatics Institute (EBI)[116] for structural alignments and the ‘Protein interfaces, surfaces and assemblies’ service PISA at EBI[117] for interface analysis. Sequence alignments were conducted in Geneious version 7.1.7[118]. For structural comparisons, PDB files of porcine FAS (2VZ9), human KR/ΨKR/ΨME (4PIV and 5C37), a PKS ER/KR/ΨKR dimer (3SLK) and the modifying region of a MsMAS-like PKS (5BP4) were utilized.

2.5 Author Contributions

H.S.T.B. performed initial cloning; F.M.C.B. cloned constructs, expressed, purified and characterized dFAS-FL and dFAS-DEK, crystallized dFAS-DEK, determined the structure of dFAS-DEK, performed structure analysis, and wrote the manuscript. T.M. designed and guided research, contributed to crystallographic analysis and contributed to the manuscript. All authors have given approval to the final version of the manuscript.

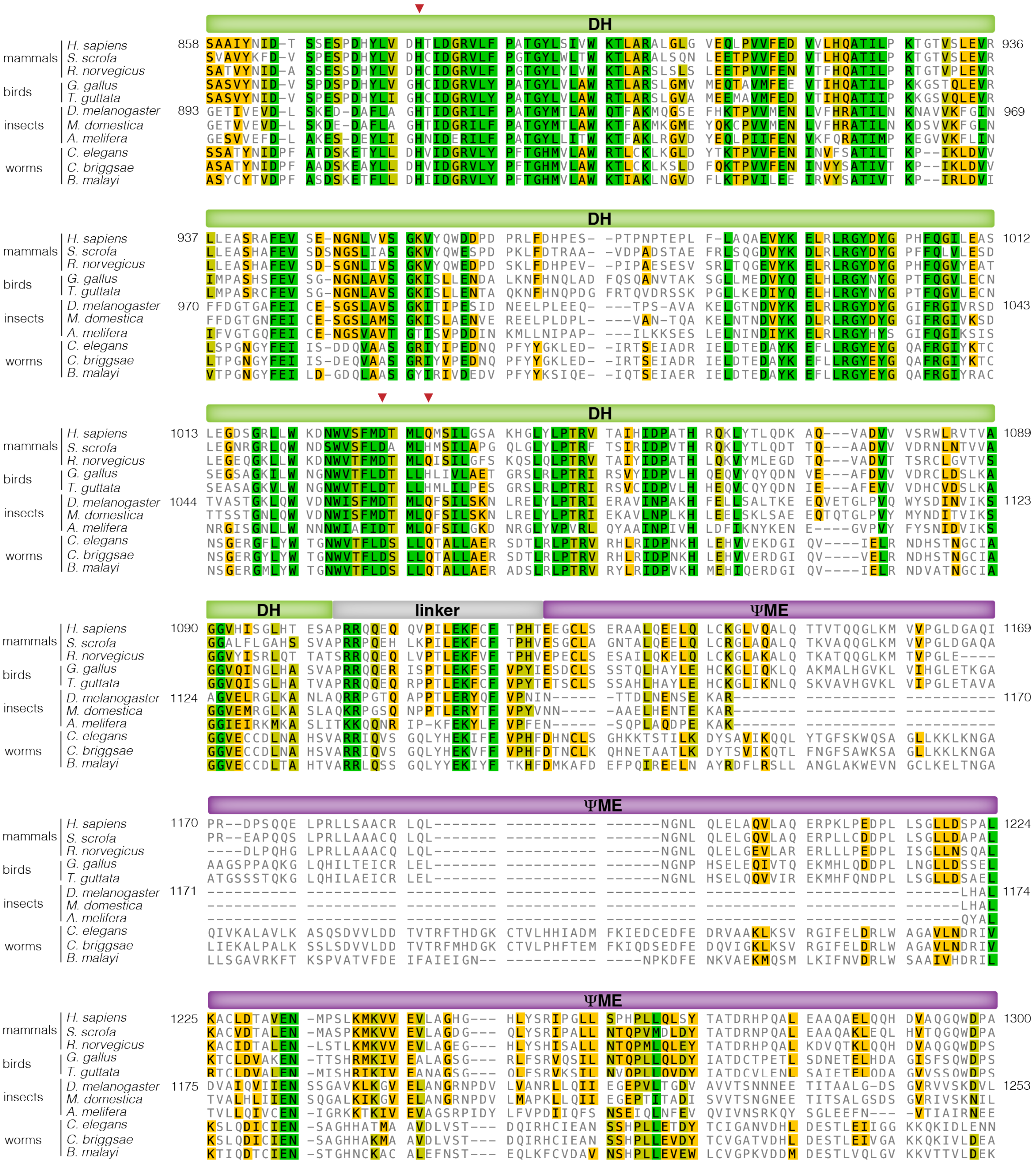
2.6 Notes

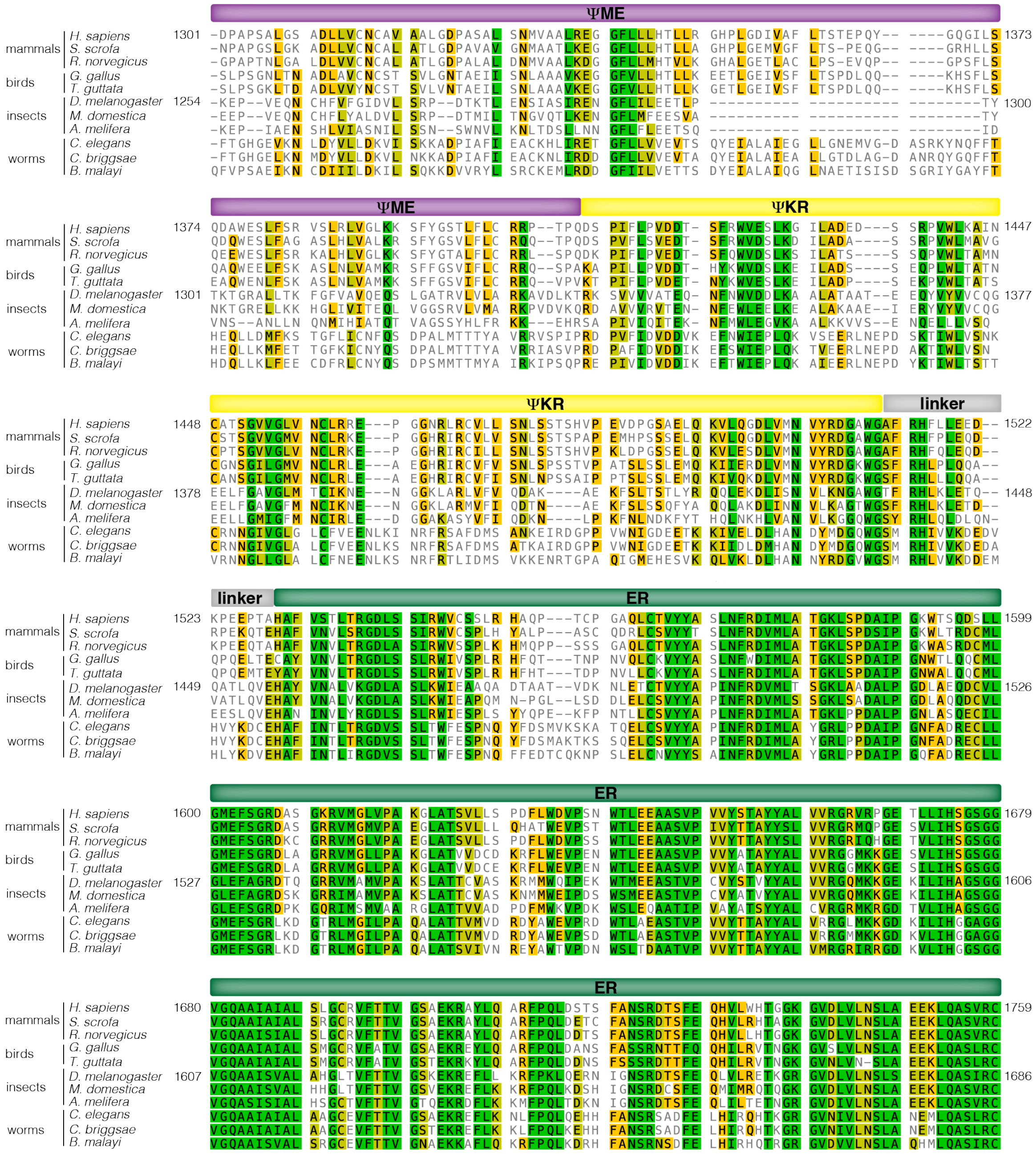
The authors declare no competing financial interest.

2.7 Acknowledgments

We acknowledge R. Jakob for support with crystallographic analysis. F.M.C.B. acknowledges a PhD fellowship from the Werner-Siemens Foundation.

2.8 Supplemental Information

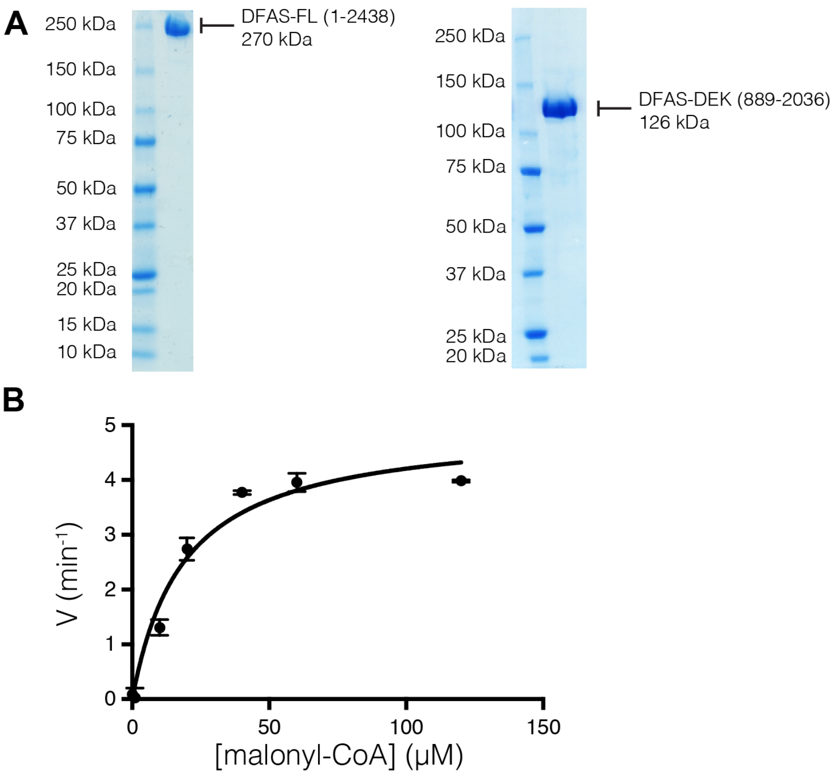




```
                                  ΨME
mammals  H. sapiens      1301 -DPAPSALGS ADLLVCNCAV AALGDPASAL SNMVAALREG GFLLLHTLLR GHPLGDIVAF LTSTEPQY-- ----GQGILS 1373
         S. scrofa            -NPAPGSLGK ADLLVCNCAL ATLGDPAVAV GNMAATLKEG GFLLLHTLLA GHPLGEMVGF LTS-PEQG-- ----GRHLLS
         R. norvegicus        -GPAPTNLGA LDLVVCNCAL ATLGDPALAL DNMVAALKDG GFLLMHTVLK GHALGETLAC LPS-EVQP-- ----GPSFLS
birds    G. gallus            -SLPSGNLTN ADLAVCNCST SVLGNTAEII SNLAAAVKEG GFVLLHTLLK EETLGEIVSF LTSPDLQQ-- ----KHSFLS
         T. guttata           -SLPSGKLTD ADLVVYNCST SVLVNTAEIL SNLAAAVKEG GFVLLHTLLK GETLGEIVSF LTSPDLQQ-- ----KHSFLS
insects  D. melanogaster 1254 -KEP--VEQN CHFVFGIDVL SRP--DTKTL ENSIASIREN GFLILEETLP ---------- ---------- --------TY 1300
         M. domestica         -EEP--VEQN CHFLYALDVL SRP--DTMIL TNGVQTLKEN GFLMFEESVA ---------- ---------- --------TY
         A. melifera          -KEP--IAEN SHLVIASNIL SSN--SWNVL KNLTDSLLNN GFLFLEETSQ ---------- ---------- --------ID
worms    C. elegans           -FTGHGEVKN LDYVLLDKVI SKKADPIAFI EACKHLIRET GFLLVVEVTS QYEIALAIEG LLGNEMVG-D ASRKYNQFFT
         C. briggsae          -FTGHGELKN MDYVLLDKVL NKKADPIAFI EACKNLIRDD GFLLVVEVTA QYEIALAIEA LLGTDLAG-D ANRQYGQFFT
         B. malayi            QFVPSAEIKN CDIIILDKIL SQKKDVVRYL SRCKEMLRDD GFIILVETTS DYEIALAIQG LNAETISISD SGRIYGAYFT

                                  ΨME                                        ΨKR
mammals  H. sapiens      1374 QDAWESLFSR VSLRLVGLKK SFYGSTLFLC RRP--TPQDS PIFLPVDDT- SFRWVESLKG ILADED---S SRPVWLKAIN 1447
         S. scrofa            QDQWESLFAG ASLHLVALKR SFYGSVLFLC RQQ--TPQDS PVFLSVEDT- SFRWVDSLKD ILADA----S SRPVWLMAVG
         R. norvegicus        QEEWESLFSR KALHLVGLKK SFYGTALFLC RRL--SPQDK PIFLPVEDT- SFQWVDSLKS ILATS----S SQPVWLTAMN
birds    G. gallus            QAQWEELFSK ASLNLVAMKR SFFGSVIFLC RRQ--SPAKA PILLPVDDT- HYKWVDSLKE ILADS----S EQPLWLTATN
         T. guttata           EAQWENLFSK ASLNLVAMKK SFFGSVIFLC RRQ--VPVKT PIFLPVDET- NYKWVESLKE ILADP----S EKPVWLTATS
insects  D. melanogaster 1301 TKTGRALLTK FGFVAVQEQS LGATRVLVLA RKAVDLKTRK SVVVVATEQ- NFNWVDDLKA ALATAAT--E EQYVYVVCQG 1377
         M. domestica         NKTGRELLKK HGLIVITEQL VGGSRVLVMA RKPVDVKQRD AVVVRVTEN- NFEWLEELKE ALAKAAE--I ERYVYVVCQG
         A. melifera          VNS--ANLLN QNMIHIATQT VAGSSYHLFR KK---EHRSA PIVIQITEK- NFMWLEGVKA ALKKVVS--E NQELLLVSQ
worms    C. elegans           HEQLLDMFKS TGFLICNFQS DPALMTTTYA VRRVSPIPRD PVFIDVDDVK EFNWIEPLQK VSEERLNEPD SKTIWLVSNK
         C. briggsae          HEQLLKMFET TGFKICNYQS DPALMTTTYA VRRIASVPRD PAFIDVDDIK EFSWIEPLQK TVEERLNEPD AKTIWLVSN
         B. malayi            HDQLLKLFEE CDFRLCNYQS DPSMMTTMYA IRKIPSQPRE PIVIDVDDIK EFTWIEPLQK AIEERLNEPD YKTIWLTSTT

                                                     ΨKR                                           linker
mammals  H. sapiens      1448 CATSGVVGLV NCLRRE---P GGNRLRCVLL SNLSSTSHVP EVDPGSAELQ KVLQGDLVMN VYRDGAWGAF RHFLLEED-- 1522
         S. scrofa            CSTSGVVGMV NCLRKE---P GGHRIRCVLV SNLSSTSPAP EMHPSSSELQ KVLQGDLVMN VYRDGAWGAF RHFPLEQD--
         R. norvegicus        CPTSGVVGLV NCLRKE---P GGHRIRCILL SNLSSTSHVP KLDPGSSELQ KVLESDLVMN VYRDGAWGAF RHFQLEQD--
birds    G. gallus            CGNSGILGMV NCLRLE---A EGHRIRCVFV SNLSPSSTVP ATSLSSLEMQ KIIERDLVMN VYRDGKWGSF RHLPLQQA--
         T. guttata           CANSGILGMV NCLRLE---A EGHRIRCVFI SNLNPSSAIP PTSLSSLEMQ KIVQKDLVMN VYRDGKWGSF RHLLLQQA--
insects  D. melanogaster 1378 EELFGAVGLM TCIKNE---N GGKLARLVFV QDAK----AE KFSLTSTLYR QQLEKDLISN VLKNGAWGTF RHLKLETQ-- 1448
         M. domestica         EELFGAVGFM NCIKNE---N GGKLARMVFI QDTN----AE KFSLSSQFYA QQLAKDLINN VLKAGTWGSF RHLKLETD--
         A. melifera          EELLGMIGFM NCIRLE---D GGAKASYVFI QDKN----LP KFNLNDKFYT HQLNKHLVAN VLKGGQWGSY RHLQLDLQN-
worms    C. elegans           CRNNGIVGLG LCFVEENLKI NRFRSAFDMS ANKEIRDGPP VWNIGDEETK KIVELDLHAN DYMDGQWGSM RHIVVKDEDV
         C. briggsae          CRNNGIVGLA LCFVEENLKS NRFRSAFDMS ATKAIRDGPP VWNIGDEETK KIIDLDMHAN DYMDGQWGSM RHIVVKDEDA
         B. malayi            VRNNGLLGLA LCFNEENLKS NRFRTLIDMS VKKENRTGPA QIGMEHESVK QLVKLDLHAN NYRDGVWGSM RHLVVKDDEM

                           linker                                    ER
mammals  H. sapiens      1523 KPEEPTAHAF VSTLTRGDLS SIRWVCSSLR HAQP---TCP GAQLCTVYYA SLNFRDIMLA TGKLSPDAIP GKWTSQDSLL 1599
         S. scrofa            RPEKQTEHAF VNVLSRGDLS SIRWVCSPLH YALP---ASC QDRLCSVYYT SLNFRDVMLA TGKLSPDSIP GKWLTRDCML
         R. norvegicus        KPEEQTAHAF VNVLTRGDLA SIRWVSSPLK HMQPP--SSS GAQLCTVYYA SLNFRDIMLA TGKLSPDAIP GKWASRDCML
birds    G. gallus            QPQELTECAY VNVLTRGDLS SLRWIVSPLR HFQT---TNP NVQLCKVYYA SLNFWDIMLA TGKLSPDAIP GNWTLQQCML
         T. guttata           QPQEMTEYAY VNVLTRGDLS SLRWIVSPLR HFHT---TDP NVLLCKVYYA SLNFRDIMLA TGKLSPDAIP GNWALQQCML
insects  D. melanogaster 1449 QATLQVEHAY VNALVKGDLA SLKWIEAAQA DTAAT--VDK NLETCTVYYA PINFRDVMLT SGKLAADALP GDLAEQDCVL 1526
         M. domestica         VATLQVEHAY VNALVKGDLA SLKWIEAPQM N-PGL--LSD DLELCSVYYA PINFRDVMLS SGKLSADALP GDLAQQDCVL
         A. melifera          EESLQVEHAN INVLYRGDLS SLRWIESPLS YYQPE--KFP NTLLCSVYYA PLNFRDIMLA TGKLPPDALP GNLASQECIL
worms    C. elegans           HVYKDCEHAF INTLTRGDVS SLTWFESPNQ YFDSMVKSKA TQELCSVYYA PINFRDIMLA YGRLPPDAIP GNFADRECLL
         C. briggsae          HVYKDCEHAF INTLTRGDVS SLTWFESPNQ YFDSMAKTKS SQELCSVYYA PINFRDIMLA YGRLPPDAIP GNFADRECLL
         B. malayi            HLYKDVEHAF INTLIRGDVS SLTWFESPNQ FFEDTCQKNP SLELCNVYYS AINFRDVMLA YGRLPPDAIP GQFADRECLL

                                                      ER
mammals  H. sapiens      1600 GMEFSGRDAS GKRVMGLVPA KGLATSVLLS PDFLWDVPSN WTLEEAASVP VVYSTAYYAL VVRGRVRPGE TLLIHSGSGG 1679
         S. scrofa            GMEFSGRDAS GRRVMGMVPA EGLATSVLLL QHATWEVPST WTLEEAASVP IVYTTAYYSL VVRGRMQPGE SVLIHSGSGG
         R. norvegicus        GMEFSGRDKC GRRVMGLVPA EGLATSVLLS PDFLWDVPSS WTLEEAASVP VVYTTAYYSL VVRGRIQHGE TVLIHSGSGG
birds    G. gallus            GMEFSGRDLA GRRVMGLLPA KGLATVVDCD KRFLWEVPEN WTLEEAASVP VVYATAYYAL VVRGGMKKGE SVLIHSGSGG
         T. guttata           GMEFSGRDLA GRRVMGLLPA KGLATVVDCE KRFLWEVPEN WTLEEAASVP VVYATAYYAL VVRGGMKKGE SILIHSGSGG
insects  D. melanogaster 1527 GLEFAGRDTQ GRRVMAMVPA KSLATTCVAS KRMMWQIPEK WTMEEASTVP CVYSTVYYAL VVRGQMKKGE KILIHAGSGG 1606
         M. domestica         GLEFAGRDSK GRRIMAMVPA KSLATTCVAS KNMMWEIPDS WSMEEASTVP CVYATVYYAL VVRGQMKKGE KILIHAGSGG
         A. melifera          GLEFSGRDPK GQRIMSMVAA RGLATTVVAD PDFMWKVPDK WSLEQAATIP VAYATSYYAL CVRGRMKRGD TVLIHAGSGG
worms    C. elegans           GMEFSGRLKD GTRLMGILPA QALATTVMVD RDYAWEVPRD WTLAEASTVP VVYTTAYYAL VRRGLMKKGD KILIHGGAGG
         C. briggsae          GMEFSGRLKD GTRLMGILPA QALATTVMVD RDYAWEVPSD WTLAEASTVP VVYTTAYYAL VRRGMMKKGD KVLIHGGAGG
         B. malayi            GMEFSGRLKD GTRLMGILPA QALATSVIVN REYAWTVPDN WSLTDAATVP VVYTTAYYAL VMRGRIRRGD KVLIHGGSGG

                                                      ER
mammals  H. sapiens      1680 VGQAAIAIAL SLGCRVFTTV GSAEKRAYLQ ARFPQLDSTS FANSRDTSFE QHVLWHTGGK GVDLVLNSLA EEKLQASVRC 1759
         S. scrofa            VGQAAIAIAL SRGCRVFTTV GSAEKRAYLQ ARFPQLDETC FANSRDTSFE QHVLRHTAGK GVDLVLNSLA EEKLQASVRC
         R. norvegicus        VGQAAISIAL SLGCRVFTTV GSAEKRAYLQ ARFPQLDDTS FANSRDTSFE QHVLLHTGGK GVDLVLNSLA EEKLQASVRC
birds    G. gallus            VGQAAIAIAL SMGCRVFATV GSAEKREYLQ ARFPQLDANS FASSRNTTFQ QHILRVTNGK GVSLVLNSLA EEKLQASLRC
         T. guttata           VGQAAIAIAL SMGCRVFATV GSTEKRKYLQ ARFPQLDDNS FSSSRDTTFE QHILRVTNGK GVNLVN-SLA EEKLQASLRC
insects  D. melanogaster 1607 VGQAAISVAL AHGLTVFTTV GSKEKREFLL KRFPKLQERN IGNSRDTSFE QLVLRETKGR GVDLVLNSLS EEKLQASIRC 1686
         M. domestica         VGQAAISVAL HHGLTVFTTV GSVEKREFLK KRFPQLKDSH IGNSRDCSFE QMIMRQTQGK GVDLVLNSLA EEKLQASVRC
         A. melifera          VGQASISIAL HSGCTVFTTV GTQEKRDFLK KMFPQLTDKN IGNSRDTSFE QLILTETNGR GVDIVLNSLA EEKLQASVRC
worms    C. elegans           VGQAAIAIAL AAGCEVFTTV GSAEKREFLK NLFPQLQEHH FANSRSADFE LHIRQHTKGR GVNIVLNSLA NEMLQASLRC
         C. briggsae          VGQAAIAIAL AAGCEVFTTV GSTEKREFLK KLFPQLKEHH FANSRSADFE LHIRQHTKGR GVNIVLNSLA NEMLQASLRC
         B. malayi            VGQAAISVAL SRGCEVFTTV GNAEKKAFLQ KRFPQLKDRH FANSRNSDFE LHIRHQTRGR GVDVVLNSLA QHMLQASIRC
```

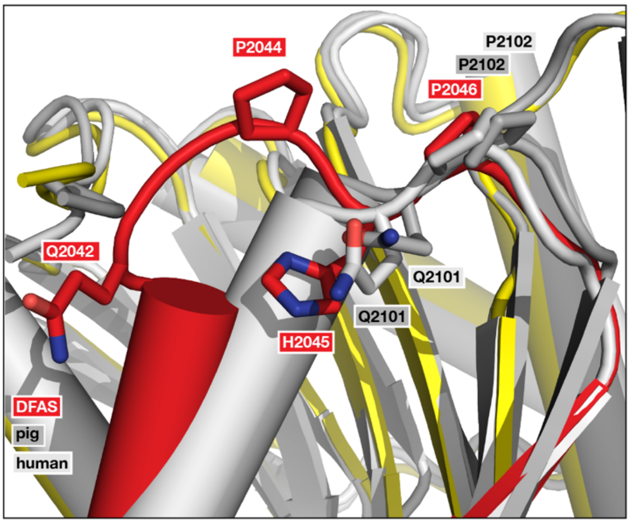
Supplementary Figure 2.1 Sequence alignment of metazoan FAS.

ClustalW sequence alignment with BLOSUM matrix, Gap open cost of 10 and Gap extent cost of 0.1. Domains are indicated above sequence. Red triangles show active site residues. Similarity is colored from green (100% similar) to yellow (60-80% similar).



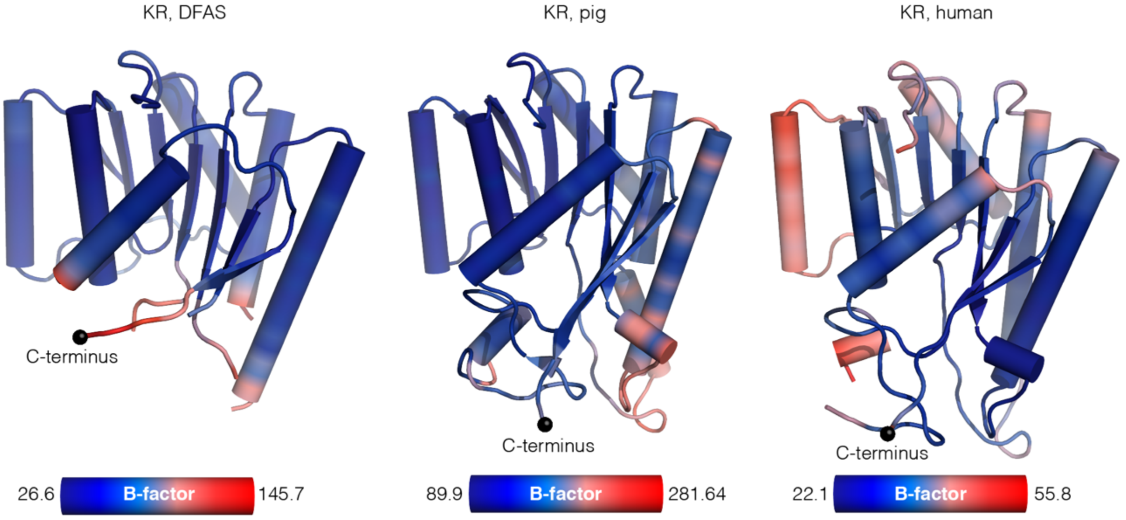
Supplementary Figure 2.2 Purification and activity of *D. melanogaster* FAS.

(A) SDS-PAGE analysis of full-length FAS (dFAS-FL; residues 1-2438) and modifying region (dFAS-DEK; residues 889-2036). **(B)** dFAS-FL catalytic activity measured by the NADPH oxidation at 340 nm, yielding a specific activity of 926 nmol NADPH per min/mg.



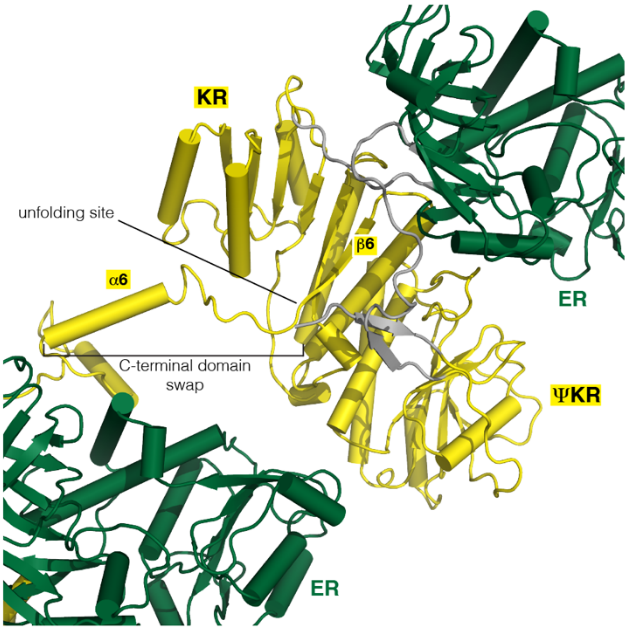
Supplementary Figure 2.3 Fold conservation in the KR C-terminus.

Structural alignment of the dFAS (red), porcine (dark gray, PDB 2VZ9) and human (light gray, PDB 4PIV) KR domain. The last 22 residues of the C-terminus in dFAS KR are colored red. In all structures, a glutamine (Gln2042 in dFAS, Gln2101 in human and pig) induces a turn after helix $\alpha6$. His2045 is structurally aligned with the Gln2101 in human and pig, suggesting a possible role in inducing a second turn in dFAS. Each turn features a proline residue in its center, with Pro2046 (dFAS) being conserved in all FAS KR structures, suggesting an important role in turn formation.



Supplementary Figure 2.4 B-factor distribution in the KR domain of dFAS, pig and human.

B-factor values are low throughout the KR domains in dFAS (left), pig (middle, PDB 2VZ9) and human (right, PDB 4PIV). Higher B-factor values appear close to and in the substrate binding extension in pig and human. In the human KR structure, higher B-factor values are shown for helices which are involved in interdomain contacts with the ER. The ER is excised in the human KR/ΨKR/ΨME FAS structure. B-factors in the C-terminal α6-turn-β7 motif are low, suggesting a rigid fold.



Supplementary Figure 2.5 Domain swap at identical site in a PKS.

Alternative refolding of the KR C-terminus in a PKS (PDB 3SLK). Following β6 of the KR Rossmann-fold (yellow), the polypeptide chain extends and interacts with an ER (green) of a different molecule in the crystal. The site of refolding corresponds to the beginning of the disordered substrate binding domain in dFAS-DEK, preceding the C-terminal α6-turn-β7 fold.

Supplementary Table 2.1 Crystallographic data collection and refinement statistics for dFAS-DEK

Wavelength (Å)	1.00003
Resolution range (Å)	129.7 – 2.7 (2.797– 2.7)*
Space group	$P\ 2_1\ 2_1\ 2_1$
Unit cell (Å)	75.45, 174.15, 194.45
α, β, γ (°)	90, 90, 90
Total reflections	1643444
Unique reflections	68282 (7006)
Multiplicity	24.06
Completeness (%)	96 (100)
Mean I/σ (I)	15.57 (2.33)
Wilson B-factor (Å^2)	54.01
R-meas	0.24
CC1/2	0.998 (0.832)
Reflections used in refinement	68240 (6989)
Reflections used for R-free	3384 (381)
R-work	0.221 (0.329)
R-free	0.255 (0.351)
Number of non-hydrogen atoms	17117
Macromolecules	16905
Protein residues	2182
RMS (bonds)	0.003
RMS (angles)	0.69
Ramachandran favored (%)	97
Ramachandran allowed (%)	3.1
Ramachandran outliers (%)	0
Rotamer outliers (%)	0
Clashscore	1.29
Average B-factor (Å^2)	70.07
Macromolecules	70.36
Solvent (%)	44.7
Number of TLS groups	8

*Statistics for the highest-resolution shell are shown in parentheses.

Supplementary Table 2.2 Structural similarity of the dFAS modifying domain with human and pig.

Cα r.m.s.d. deviations from structural alignment of dFAS-DEK onto available high-resolution structures of FAS using PDBeFOLD. Only the best matching monomer is listed. PDB accession codes are given in parentheses.

Structure 1	Structure 2	Cα r.m.s.d. [Å]	Aligned residues
dFAS DEK	pig DEK (2VZ9)	1.51	622
dFAS DH	pig DH (2VZ9)	1.18	225
dFAS ΨME	pig ΨME (2VZ9)	1.97	155
dFAS ΨME	human ΨME (4PIV)	1.79	143
dFAS ΨKR	pig ΨKR (2VZ9)	2.07	84
dFAS ΨKR	human ΨKR (4PIV)	2.10	90
dFAS ER	pig ER (2VZ9)	1.03	316
dFAS KR	pig KR (2VZ9)	1.22	177
dFAS KR	human KR (4PIV)	1.28	177
dFAS KR	human KR (5C37)	1.21	177

Supplementary Table 2.3 Interface analysis of the dFAS modifying region. Interface surface area and interaction types within the dFAS-DEK dimer obtained with PDBePISA.

Interface 1	**Interface 2**	**Area [Å^2]**	**Hydrogen bonds**	**Salt bridges**
ER, chain A	ER, chain B	1113.8	31	-
DH, chain A	DH, chain B	417.9	8	-
DH, chain A	ER, chain A	221.0	1	-
DH, chain A	ER, chain B	331.1	4	-
DH, chain A	KR, chain A	559.1	7	1
KR, chain A	ER, chain A	257.2	3	2
KR, chain A	ΨKR, chain A	591.1	8	-
KR, chain A	ΨME, chain A	27.0	-	-
ΨKR, chain A	ΨME, chain A	697.1	9	1

Supplementary Table 2.4 Interactions of the KR C-terminus with adjacent linker regions.

Interface surface area, interaction types and free solvation energy upon interface formation (ΔG) between the KR C-terminus und adjacent linker regions. dFAS $KR^{C\text{-}ter}$ encompasses residues 2033-2055 (unnatural C-terminus) and pig $KR^{C\text{-}ter}$ comprises residues 2088-2113.

Interface 1	Interface 2	Area [Å^2]	Hydrogen bonds	Salt bridges	ΔG [kcal/mol]§
dFAS					
$KR^{C\text{-}ter}$, A-chain	DH-ΨME linker	630.1	8	-	-12.1
$KR^{C\text{-}ter}$, B-chain	DH-ΨME linker	561.7	6	-	-11.1
$KR^{C\text{-}ter}$, A-chain	ΨKR-ER linker	209.4	-	-	-2.5
$KR^{C\text{-}ter}$, B-chain	ΨKR-ER linker	229.8	-	-	-2.9
Pig					
$KR^{C\text{-}ter}$, A-chain	DH-ΨME linker	490.7	6	-	-8.1
$KR^{C\text{-}ter}$, B-chain	DH-ΨME linker	489.0	5	-	-7.7
$KR^{C\text{-}ter}$, A-chain	ΨKR-ER linker	243.3	1	1	-4.0
$KR^{C\text{-}ter}$, B-chain	ΨKR-ER linker	230.4	1	1	-3.9

§difference in total solvation energy of isolated and interfacing structure

Supplementary Table 2.5 Interactions in the C-terminal helix-turn-strand fold. Interface surface area, interaction types and free solvation energy upon interface formation (ΔG) within the KR C-terminus for dFAS, porcine and human KR. The model "dFAS 2015-2036" was obtained by building the natural sequence of the C-terminus in the dFAS model, following the fold in the porcine and human KR structures.

Construct		**Area [Å^2]**	**Hydrogen bonds**	**Salt bridges**	**ΔG [kcal/mol]§**
dFAS					
2015-2036, A-chain	Natural sequence	122.9	-	1	-1.1
2015-2036, B-chain	Natural sequence	122.9	-	1	-1.1
2033-2055, A-chain	Incorporated expression tag	194.1	-	1	-4.2
2033-2055, B-chain	Incorporated expression tag	148.8	-	1	-3.5
Pig					
2088-2113, A-chain	Natural sequence	196.2	-	1	-3.8
2088-2113, B-chain	Natural sequence	194.6	-	1	-3.8
Human					
2088-2113, A-chain	Natural sequence	216.9	-	1	-3.8
2088-2113, B-chain	Natural sequence	202.8	-	1	-4.1

§difference in total solvation energy of isolated and interfacing structure

Supplementary Table 2.6 Interactions of the KR C-terminus with the KR fold. Interface surface area, interaction types and free solvation energy upon interface formation (ΔG) of the KR C-terminus with the β-sheet of the KR domain for dFAS, porcine and human KR. The model "dFAS 2015-2036" was obtained by building the natural sequence of the C-terminus in the dFAS model, following the fold in the porcine and human KR structures.

Interface 1	Interface 2	Area [Å^2]	Hydrogen bonds	Salt bridges	ΔG [kcal/mol]§
dFAS					
KR, chain A	C-ter 2033-2055 (expression tag)	935	8	1	-17.4
KR, chain B	C-ter 2033-2055 (expression tag)	903.3	9	-	-17.8
KR, chain A	C-ter 2015-2036 (natural sequence)	804.1	7	2	-17.1
KR, chain B	C-ter 2015-2036 (natural sequence)	799.0	6	1	-15.8
Pig					
KR, chain A	Natural sequence	1212.0	20	5	-20.7
KR, chain B	Natural sequence	1202.3	21	5	-20.5
Human					
KR, chain A	Natural sequence	1276.8	23	6	-21.8

KR, chain B	Natural sequence	1279.3	23	6	-21.8

[§] difference in total solvation energy of isolated and interfacing structure

3 High-Speed AFM Visualization of the Dynamics of the Multienzyme Fatty Acid Synthase

Reproduced with permission from

Benning, F. M. C.; Sakiyama, Y.; Mazur, A.; Bukhari, H. S. T.; Lim, R. Y. H.; Maier, T.

High-Speed Atomic Force Microscopy Visualization of the Dynamics of the

Multienzyme Fatty Acid Synthase. *ACS Nano*. **2017**, *11*, 10852-10859.

3.1 Abstract

Multienzymes, such as the protein metazoan fatty acid synthase (FAS), are giant and highly dynamic molecular machines for critical biosynthetic processes. The molecular architecture of FAS was elucidated by static high-resolution crystallographic analysis, while electron microscopy revealed large-scale conformational variability in FAS with some correlation to functional states in catalysis. However, little is known about time scales of conformational dynamics, the trajectory of motions in individual FAS molecules, and the extent of coupling between catalysis and structural changes. Here, we present an experimental single-molecule approach to film immobilized or selectively tethered FAS in solution at different viewing angles and high spatiotemporal resolution using high-speed atomic force microscopy. Mobility of individual regions of the multienzyme is recognized in video sequences, and correlation of shape features implies a convergence of temporal resolution and velocity of FAS dynamics. Conformational variety can be identified and grouped by reference-free 2D class averaging, enabling the tracking of conformational transitions in movies. The approach presented here is suited for comprehensive studies of the dynamics of FAS and other multienzymes in aqueous solution at the single-molecule level.

KEYWORDS: multienzyme; molecular movie; conformational flexibility; lipid bilayer; immobilization; macromolecular assembly; carrier protein

3.2 Introduction

Multienzymes and multienzymatic complexes are dynamic machines of life with crucial roles in cellular functions such as energy homeostasis, signaling, and anabolic metabolism.[96,119-122] They also serve as blueprints for the design of bioinspired nanomachines, targeted for the tailored production of biological compounds.[123-126] Multienzymes integrate several enzymatic functions into giant polypeptide chains, which may further oligomerize to form a functional biological assembly. In carrier protein-dependent multienzymes, a flexibly tethered carrier protein shuttles covalently bound intermediates between catalytic domains. Substrate tethering enhances product formation by preventing noncognate access by other cellular enzymes, providing high local concentrations of intermediates and enabling the processing of insoluble substrates.[10,127]

Metazoan fatty acid synthase (FAS) is a key paradigm for carrier protein-dependent multienzymes. It catalyzes the essential *de novo* biosynthesis of fatty acids from carbohydrate-derived precursors.[33] FAS inhibition blocks proliferation in many solid tumors, making FAS an attractive target for antitumor therapy.[13]

The FAS polypeptide comprises seven distinct functional domains, that catalyze more than 40 individual reaction steps in the cyclic biosynthesis of palmitic acid. Metazoan FAS is active as a homodimer. It adopts an X-shape featuring two lateral reaction clefts and is structurally and functionally segregated into a condensing and a modifying region (Figure 3.1).[11]

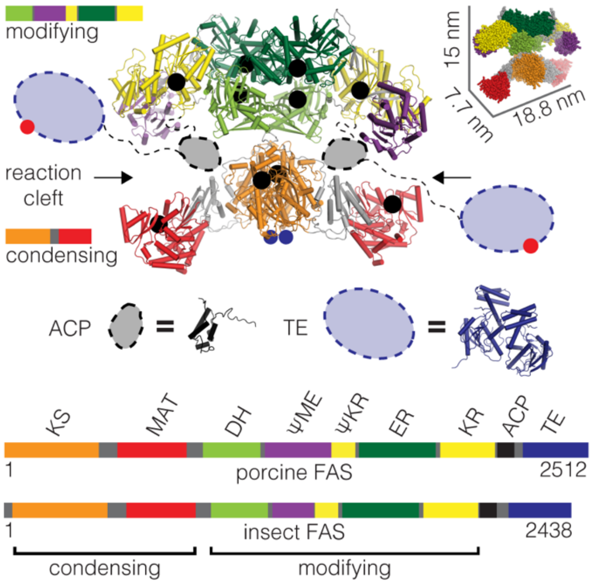


Figure 3.1 Crystal structure of metazoan FAS (Protein Data Bank (PDB) accession code: 2VZ9, from *S. scrofa*).

Functionally plausible positions of the nonresolved flexible acyl carrier protein (ACP) and thioesterase domain (TE) are schematically indicated on the basis of isolated high-resolution structures (rat ACP, black, PDB: 2PNG; human TE, blue, PDB: 2PX6). The domain organization of mammalian and insect FAS is schematically displayed below. Sequence identity to *D. melanogaster* FAS is 44% for porcine FAS, 47% for rat ACP, and 31% for human TE. Locations of the active sites (black spheres) as well as the His_{10}-tag anchor sites at the amino terminus (blue spheres) and carboxy terminus (red spheres) are indicated. Domains are abbreviated as follows: ketosynthase (KS; orange), malonyl/acetyltransferase (MAT; red), dehydratase (DH; light green), pseudomethyltransferase (ΨME; violet), pseudoketoreductase (ΨKR; yellow), enoylreductase (ER; dark green) and ketoreductase (KR; yellow). Linkers are indicated in gray.

Our crystallographic analysis of mammalian FAS revealed key insights into domain organization and provided a static representation of multienzyme architecture at near-atomic resolution.[26,38] The flexibly tethered acyl carrier protein (ACP) and thioesterase (TE) domains remained unresolved in the crystallographic analysis. However, on the basis of the extent of its flexible linker, the mobility of the phosphopantetheinylated ACP in the static model is sufficient to reach a complete set of active sites within one reaction cleft at distances between active sites of up to 85 Å, while the opposite site reaction cleft remains out of reach (Figure 3.1).[11] However, *in vitro* mutant complementation and cross-linking assays indicated the interaction of one ACP with active sites of the condensing region from both reaction clefts.[65,66] Indeed, electron microscopy (EM) analysis of metazoan FAS demonstrated large-scale conformational variability with some correlation to functional states.[29,64] A similar degree of structural variability has been observed in other carrier protein-dependent multienzymes and suggests a general role of large-scale domain movements.[12,67]

Here, we report an experimental approach based on high-speed atomic force microscopy (HS-AFM) for studying the dynamics of individual multienzymes in aqueous solution approaching a temporal and spatial resolution required to follow large-scale conformational changes during catalysis. HS-AFM has been used to visualize dynamic protein systems such as myosin V on actin,[86] the F_1-ATPase,[87] GroEL-GroES interaction[128] and the nuclear pore complex.[88] It enables direct simultaneous recording of structure and dynamics at nanometer spatial and sub-100 ms temporal resolution in aqueous buffer without the need for labeling.[90] As HS-AFM is operated in tapping mode with forces in the hundred piconewton range at microsecond pulses, disturbances to the specimen are minimized.

To explore monitoring real-time dynamics of single multienzymes and megasynthases ranging from 0.3 to >1 MDa during catalysis in solution, we visualized metazoan FAS using HS-AFM and employed quantitative approaches for unbiased image analysis. Soluble proteins are demanding objects for HS-AFM imaging and analysis due to their nonspecific surface orientation. Carrier protein-dependent multienzymes pose a particular challenge because of their pronounced conformational variability and the small size of their domains at the 1-10 nm length scale. For obtaining alternate views of the multienzyme, we employ different methods of nonspecific immobilization and selective tethering. The approach depicted here provides a basis for studying the dynamics of individual multienzymes in solution at high spatiotemporal resolution with the ultimate aim of filming their inherent conformational flexibility during enzymatic action.

3.3 Results and Discussion

3.3.1 Expression and Purification of FAS

D. melanogaster FAS, carrying a decahistidine tag (His_{10}) at either its amino (N-His_{10}-FAS) or carboxy (FAS-His_{10}-C) terminus, was overexpressed in *S. frugiperda* (Sf21) insect cells and purified with a yield of 9.6 mg per liter of culture. Immobilized nickel-affinity chromatography was applied as a first purification step; FAS was further purified by anion exchange and size exclusion chromatography. Homogeneity was confirmed by SDS-PAGE analysis (Supplementary Figure 3.1A). Formation of long-chain fatty acids through the addition of the substrates acetyl-coenzyme A (acetyl-CoA), malonyl-CoA, and NADPH was monitored by measuring NADPH oxidation at 340 nm, yielding a specific activity of 926 nmol NADPH per min/mg (Supplementary Figure 3.1B), similar to human FAS (Chapter 3.9.2).

3.3.2 Imaging FAS Immobilized on Mica

High-speed AFM measurements of FAS on both mica and mica-supported lipid bilayers are compatible with substrate addition. For imaging on mica, FAS was first nonspecifically immobilized on freshly cleaved mica at room temperature (Figure 3.2A). Imaging quality was best in 50 mM potassium phosphate at pH 7.0. At particle concentrations of 2-3 nM, single multienzyme immobilization was achieved. Higher concentrations resulted in crowded surfaces. Frames were collected at rates of 0.5-4 frames per second (fps). Frame collection was in some cases impeded by spontaneous particle detachment from the mica surface and drifting of the sample stage. Drifting was counteracted by manual position correction of the sample stage. The majority of FAS molecules showed a preferred orientation on mica, offering a side view onto the molecule resembling the X-shape view depicted in Figure 3.1. FAS dimers were identified by their size and shape. The average dimensions of the particles of 24.9 ± 3.8 nm along the long axis and 6.4 ± 1.1 nm height match the dimensions of 18.8 nm by 15 nm by 7.7 nm from the crystal structure of *S. scrofa* FAS, in which the flexibly tethered ACP and TE domains are not resolved (Figure 3.2B).[26] In consecutive video frames, two independently rotating large subsegments of FAS can be identified (supporting movie).

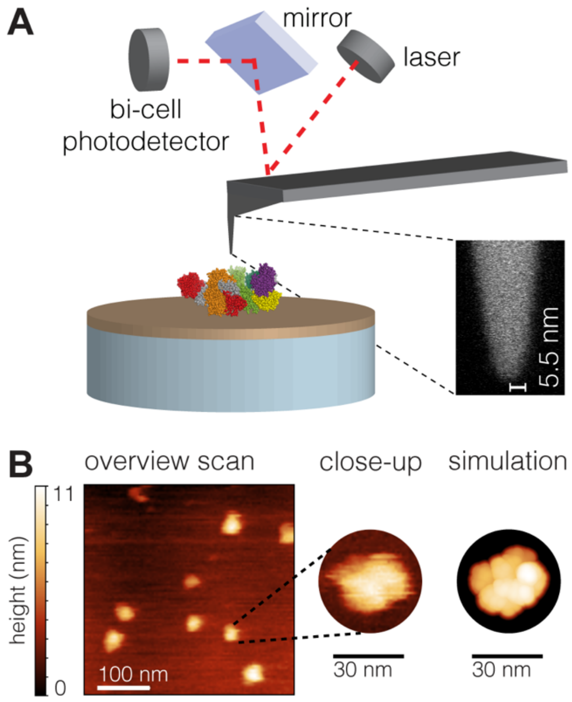


Figure 3.2 Imaging FAS immobilized on mica.

(**A**) Experimental setup of HS-AFM imaging of FAS (as shown in Figure 3.1) on mica. An electron microscopy image of the electron beam-deposited carbon tip is shown (right); carbon tip and molecular structure are not drawn to scale. (**B**) 400 nm × 400 nm overview scan of 2 nM FAS on mica, close-up scan of a single FAS molecule and matched simulated AFM-image (for details, see chapter 3.9.10). The FAS orientation used for generating the AFM simulation was determined by cross-correlation analysis (Chapter 3.9.10).

3.3.3 Imaging N-/C-Terminally Tagged FAS selectively Tethered to Affinity Lipid Bilayers

With the objective of providing different filming perspectives and to avoid immobilization-specific restraints on dynamics, a decahistidine tag plus a connecting sequence was introduced at either the amino (N-His_{10}-FAS) or the carboxy terminus (FAS-His_{10}-C) of FAS for tag-based tethering (Figure 3.3A, B; for details, see the Methods). While the two amino termini of the active homodimer are in close proximity to each other at the center of the condensing region of the molecule, preceding the ketosynthase (KS) domain, the two carboxy termini are located at the TE domain (Figure 3.1).[38] This domain has been shown to be extremely flexible in metazoan FAS[129] and is only tethered to the inherently mobile carrier protein *via* an unstructured approximately 40-residue long linker.

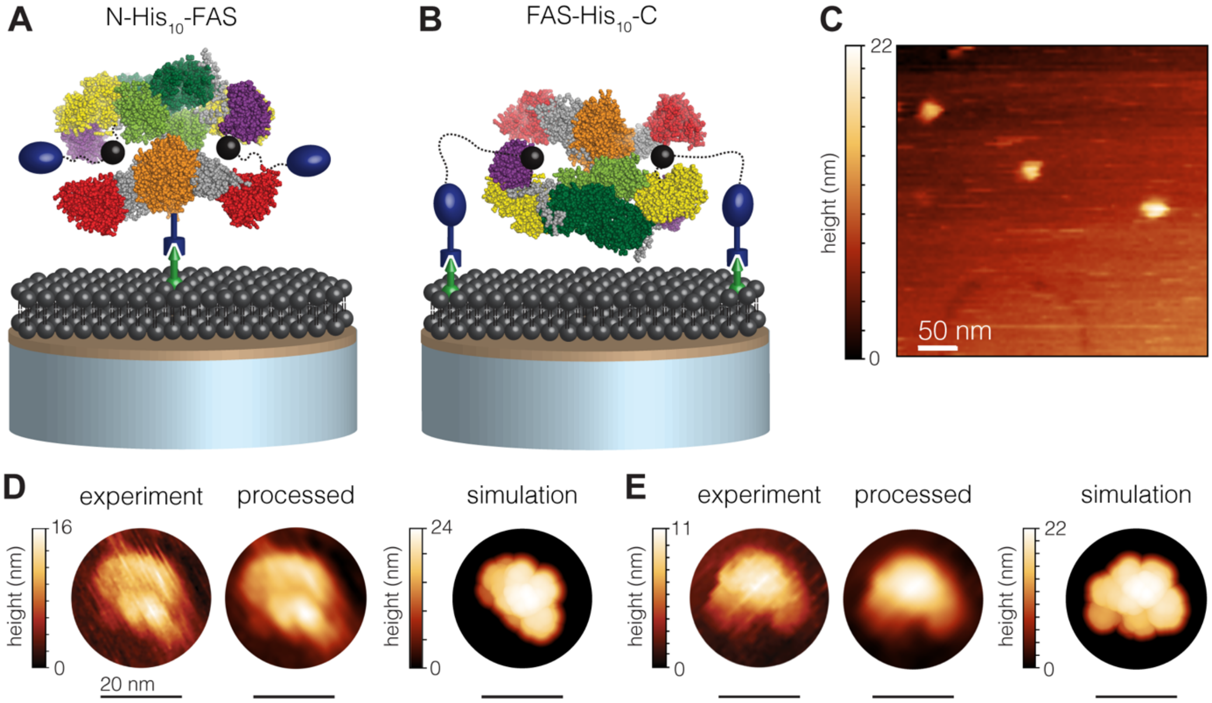


Figure 3.3 Imaging N-/C-terminally tagged FAS selectively anchored to mica-supported lipid bilayers (SLB).
(**A**, **B**) Schematic representation of N-His$_{10}$-terminally (**A**) and C-His$_{10}$-terminally (**B**) tethered FAS. (**C**) Unprocessed 400 nm × 400 nm overview scan of N-His$_{10}$-FAS on SLB. (**D**, **E**) Experimental raw HS-AFM image, processed and matched simulated AFM image of N-His$_{10}$-FAS (**D**) and FAS-His$_{10}$-C (**E**) on SLB (for details, see chapter 3.9.10).

Decahistidine-tagged FAS was tethered to mica-supported lipid bilayers (SLB) comprising 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) and 1,2-dioleoyl-*sn*-glycero-3-[(N-(5-amino-1-carboxypentyl)iminodiacetic acid)succinyl] (nickel salt) (DGS-NTA(Ni)). Planar complete bilayer patches were created by dispensing 100 µM liposomes, comprised of a 4-fold molar excess DMPC over DGS-NTA(Ni), onto freshly cleaved mica and incubation for at least 12 h at room temperature. Shorter incubation times resulted in incomplete SLBs. Prior to immobilization of FAS onto the SLBs, the bilayer quality was evaluated by initial scans. The bilayers appeared planar and continuous, containing only very few small holes. FAS was filmed at imaging rates ranging from 0.5 to 10 fps. Frame collection was in some cases obstructed by lipid particles being dragged by the carbon tip. Visual inspection did not reveal any particle detachment from the lipid bilayer, indicating sufficient affinity for tethering. As opposed to FAS on mica, all molecules remained stably attached.

FAS molecules display a different preferential orientation compared to mica-attached FAS. The average height of N-His$_{10}$-FAS molecules was 11.5 ± 0.9 nm, matching the dimensions of the atomic models of homologous enzymes and suggesting that the short axis of the particle is preferentially oriented perpendicular to the bilayer (Figure 3.3D). FAS-His$_{10}$-C molecules (Figure 3.3E) had an average height of 11.9 ± 3.1 nm. The larger standard deviation could be attributed to varying tilting of the molecule, presumably favored by tethering *via* the mobile TE domain.

3.3.4 Movie Frame Alignment and Analysis

Further analysis and processing was conducted on all movies filmed under nonspecific immobilization or selective tethering. To compensate for drift, we have aligned particles in a total of 13637 frames from 163 movies. Problems encountered using conventional alignment methods were sudden particle dislocation when analyzing FAS on mica movies, two or more molecules moving independently in one movie, as well as image tilt, dirt, and noise. These problems prevented the collection of longer frame sets than 100 frames (20-100 frames) but were partly overcome by introducing an algorithm for tilt correction and *xy*-alignment (Chapter 3.9.3). Selecting a single frame as a reference for alignment led to significantly better results than aligning an image onto the preceding frame. Structural variability as a combined result of conformational or orientational changes was quantitatively described by calculating variance maps based on the height information on the aligned frames of a movie (Figure 3.4, Chapter 3.9.4).

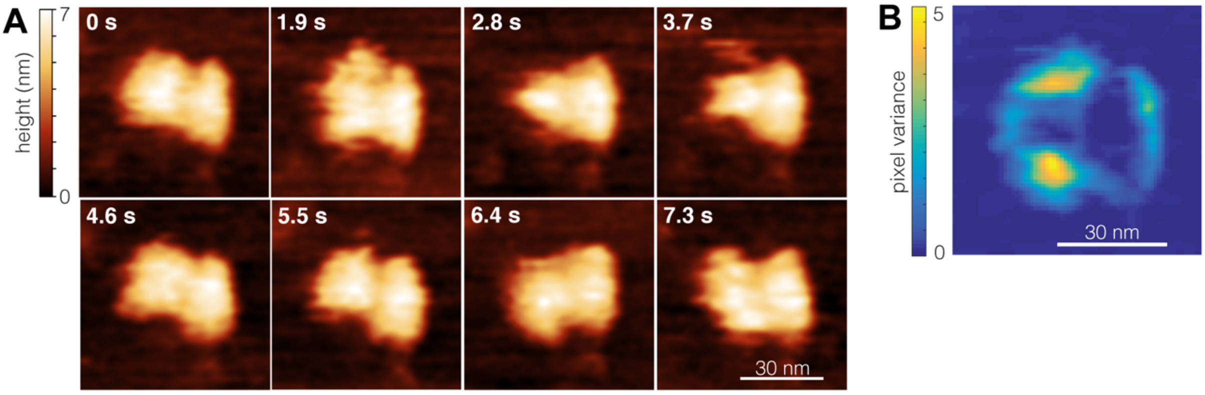


Figure 3.4 Dynamics of a single FAS on mica.

(**A**) Successive aligned and filtered frames from a movie of FAS at 0.5 fps. (**B**) Variance map calculated from all frames in (**A**) showing variance in pixel height, which can be interpreted as a relative rotation between two segregated subsegments of FAS.

3.3.5 Temporal Resolution of FAS Imaging and Time Scales of Biological Motions

One of the hallmarks of HS-AFM is its simultaneous high temporal and spatial resolution. High imaging speeds result from using small cantilevers with higher resonance frequencies compared to cantilevers used in conventional tapping-mode AFM. In this protocol, frame durations of down to 100 ms (for selectively-tethered FAS) were achieved; the maximum frame duration employed was 2000 ms.

In metazoan FAS, depending on assay conditions and source organism, the rate of product formation, calculated for palmitate, varies between 0.3 and 1.6 molecules per second with two to 11 individual enzymatic elongation cycles performed per second per FAS dimer (Chapter 3.9.2). With imaging speeds of up to 10 fps, insect FAS is sampled 2.5 times per cycle. The correlation of sequential images within a movie was addressed by calculating the autocorrelation function of the first and second principal component of the imaged particle. Autocorrelation is observed in 46% of all movies, including all scan speeds, suggesting that sequential images are not random (Chapter 3.9.5).

Synthesis of palmitic acid, the major product of FAS, requires seven cycles of chain elongation, each of which involves shuttling of intermediates between five active sites by ACP. Mutant complementation assays in rat FAS indicated that in 20% of the cycles the ACP reaches to the opposite site reaction cleft.[130] Assuming a conserved molecular architecture and synthesis mechanism in metazoan FAS, this suggests a rotation of the condensing part at a rate of 0.84 s^{-1} in our movies of insect FAS. Thus, we expect at least two regimes of motions in FAS: fast smaller scale motions during substrate shuttling by ACP and slower large-scale rotations. While the large-scale rotations can be temporally and spatially resolved, the faster conformational changes induced during substrate shuttling require a higher temporal and spatial resolution.

3.3.6 Spatial Resolution of HS-AFM Imaging of FAS

With the current setup, we achieve an estimated minimum lateral resolution of ~2-5 nm and ~1.5 nm vertical resolution. The resolution was estimated by distance measurements between two separated feature peaks in cross sections of the particles and comparison to simulated AFM images (Chapter 3.9.6). In movie series, we observe a segregation of FAS into two principal segments, which slightly differ in shape and size. While the general X-shape of the enzyme is recognizable, identifying individual domain regions remains a challenge. Plasma etching of the electron beam-deposited (EBD) carbon tip is suited for producing tips with an apex radius of 1-2 nm, albeit at low yields.[131,132] This could significantly improve the spatial resolution and enable the identification of individual domains in FAS.

3.3.7 Analysis of Conformational Transitions of Single FAS Molecules by 2D Image Classification

Large-scale motions in FAS were visualized previously by negative stain electron microscopy.[29,64] By fitting atomic structures of FAS domains into the EM reconstructions, the authors identified rotation and swinging of the condensing relative to the modifying portions as well as domain rearrangements in the modifying part of FAS. Visual inspection of our movies reveals large-scale movements of parts of the molecule in most frames (Figure 3.4A). Most notably, we observe fast rotation of one of the two segments relative to the other as well as an opening and closing of the cleft (supporting movie).

Reference-free correntropy-based 2D classification[133] of particles enabled us to identify a wide variety of conformations and to track them on a movie timeline (Figure 3.5). 2D classification is a technical term for a standard procedure in single-particle EM for analyzing individual 2D projection images, which contain 3D information from electron transmission in the form of pixel values. The technique is used to obtain homogeneous image populations with respect to molecular structure and image quality. Here, we used this procedure for grouping individual 2D particle images with pixel values encoding 3D information based on surface height, extracted from the context of consecutive movie frames, and for the bias-free identification of similar states within and across movies.

Due to a substantially higher contrast and an intrinsically high resolution of the HS-AFM images, basic parameters had to be modified compared to the application of 2D classification in single-particle EM. The images used for classification, in contrast to EM images, already have strong local contrast, inherently limited resolution and provide a surface representation only. Thus, classification does not lead to enhanced detail level or toward full 3D reconstructions.

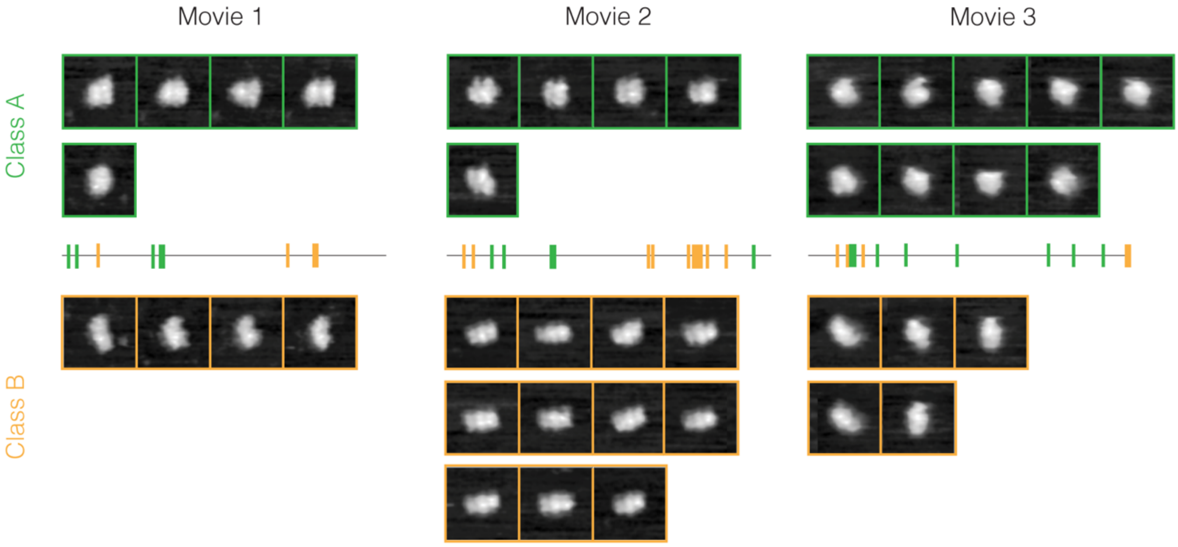


Figure 3.5 Identification of conformational variety by reference-free 2D classification of FAS imaged on mica in the absence of substrate.

Example for two out of 60 classes. Similar conformations appearing in different movies can be sorted into the same class. The temporal occurrence of a classified particle is depicted on a schematic movie timeline (green, class A; yellow, class B). However, at present no complete coverage of frames is achieved.

For analysis, between 397 and 2196 frames originating from different movies in each individual immobilization category were extracted and subjected to classification without prior information on their movie context. The number of classes was varied, and an optimum of 15 to 100 classes was found, depending on the category. Images with residual artifacts such as streaks had a minor effect on classification. The success of the classification procedure is demonstrated by the fact that classification preferentially groups successive movie frames into identical classes, in agreement with the results of the autocorrelation analysis, without considering the movie context in the classification process itself. Additionally, the classification procedure is insensitive to small variations between source movies and is able to identify related states across movies (Figure 3.5). Thus, the unbiased 2D classification procedure enabled us to assign principal conformational states of FAS as well as to obtain their lifetime, transition duration, and sequential order.

Many principal FAS conformations appear in both the absence and the presence of substrate, indicating that FAS already samples a large range of conformations in a resting state. At the current stage of data collection, mapping of principal classes is still incomplete, leaving substantial gaps between identified classes in the movie frame series. Interpretation of plausible differences in conformational dynamics between FAS imaged in the absence and presence of substrate requires a substantially higher amount of data. Based on the principal success of analysis, we can now aim for large-scale HS-AFM movie collection to obtain sufficient data for a comprehensive analysis of functional dynamics of FAS using series of substrate conditions and mutant protein variants.

3.4 Conclusion

Here, we present an approach to visualize the conformational dynamics of FAS in solution. We demonstrate that single, highly flexible multienzymes can be analyzed by HS-AFM in aqueous environments to obtain a mapping of their conformational dynamics with subsegment level spatial resolution. The combination of different molecular tethering strategies offers variable viewpoints, eliminates artifacts and is advantageous because it preserves the dynamic functional state of FAS. Autocorrelation analysis demonstrates that the temporal resolution of ~100 ms overlaps with the time scale of FAS large-scale conformational changes. Reference-free 2D classification allows a detailed analysis of conformational states and their transitions. HS-AFM imaging proves to be a suitable technique for elucidating conformational dynamics of multienzymes with high spatiotemporal resolution with further potential for optimizing spatial resolution (*e.g.*, based on sharper tips) or matching temporal resolution by slowing molecular motions using low temperature or higher viscosity media. Nevertheless, the current approach provides highly complementary data to single-molecule studies measuring individual Förster resonance energy transfer (FRET) distances between fluorescent label pairs or static 3D cryo-EM and crystallographic imaging with the additional option to directly add substrates and observe enzymatic action *in situ*. Based on the data presented here, HS-AFM imaging will strongly contribute to our understanding of functional multienzyme dynamics, as a prerequisite for the ultimate aim of re-engineering efficient biosynthetic factories for the production of complex natural products.

3.5 Materials and Methods

3.5.1 Protein Expression and Purification

Full-length *D. melanogaster* FAS (FAS) was assembled from fragments GH17750 (GenBank: BT050523.1, obtained from Drosophila Genomics Resource Center (DGRC), USA) and GH07627 (GenBank: AF145643.1, obtained from DGRC, USA) by RedET-mediated linear-linear homologous recombination.[134] N-His_{10}-FAS and FAS-His_{10}-C were generated by cloning FAS into pAB1GN-His10 and pAB1GC-His10, which are Gateway (Thermo Fisher Scientific, USA) compatible derivatives of the pACEBac1 vector (Geneva Biotech, CH). They encode an additional amino- or carboxy-terminal decahistidine tag, respectively, and a connecting sequence encompassing a protease cleavage and cloning site with a combined maximum length of 9.3 nm (tag: 3.3 nm, connecting sequence: 6 nm) in an all-extended conformation. Both full-length constructs, which only differ by the position of the affinity tag, were expressed and purified individually. Protein expression was carried out in Sf21 insect cells (Expression Systems, USA) in Insect-XPRESS medium (Lonza, CH) at 27 °C using the baculovirus expression system.[99] Virus was stored according to the titerless infected-cells preservation and scale-up method as baculovirus-infected insect cells (BIIC).[100]

Cells were harvested 72 h postinfection by centrifugation and lysed by sonication in 50 mM HEPES pH 7.4, 200 mM NaCl, 20 mM imidazole, 10% glycerol and 10 mM β-mercaptoethanol. Lysate was cleared by ultracentrifugation. Soluble protein was purified by affinity chromatography using High Affinity Ni-Charged Resin (GenScript, USA) and eluted with a linear gradient to 300 mM imidazole. The eluate was diluted to 50 mM NaCl and subjected to anion exchange chromatography using a PL-SAX

4000 Å 10 μm resin (Agilent, USA) and a linear gradient to 500 mM NaCl. FAS was concentrated to ~30 mg/mL and applied onto a Superdex 200 size exclusion chromatography column (GE Healthcare, USA). Purified FAS was concentrated to 12 mg/mL in 20 mM HEPES pH 7.4, 250 mM NaCl, 5% glycerol and 5 mM DTT, flash-frozen in liquid nitrogen and stored at -80 °C.

3.5.2 Enzymatic Activity Assay

FAS activity in the overall reaction of long-chain fatty acid formation was measured spectrophotometrically (Synergy H1 Hybrid Reader, BioTek, USA) by monitoring the decrease in absorbance of NADPH at 340 nm.[101] After 20 nM FAS was incubated with 40 μM acetyl-CoA and 150 μM NADPH in 50 mM potassium phosphate buffer at pH 7.0 at 25 °C for 5 min, malonyl-CoA was added at different concentrations up to 120 μM, as indicated in Supplementary Figure 3.1B, followed by UV recordings of 40 min.

3.5.3 FAS Immobilization on Mica

Mica disks were glued to a glass sample stage and attached to the top of the *z*-piezo on the scanner according to Research Institute of Biological Metrology (RIBM) protocols.[131] Following centrifugation, thawed protein was diluted to 2-3 nM in storage buffer (20 mM HEPES pH 7.4, 250 mM NaCl, 5% glycerol, and 5 mM DTT) at room temperature (~20 °C). A 2 μL drop of protein solution was deposited onto freshly cleaved mica and incubated for 5-10 min in a humidified chamber. Excess molecules were removed by gentle rinsing with 50 mM potassium phosphate at pH 7.0 (observation buffer).

3.5.4 FAS Tethering to Mica-Supported Lipid Bilayers

For HS-AFM imaging of affinity-tag tethered FAS, mica-supported planar lipid bilayers (SLB) were formed by deposition of large unilamellar vesicles (LUV) on a freshly cleaved mica surface, as described elsewhere.[135,136] Briefly, liposomes were prepared by mixing 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) and 1,2-dioleoyl-*sn*-glycero-3-[(N-(5-amino-1-carboxypentyl)iminodiacetic acid)succinyl] (nickel salt) (DGS-NTA(Ni)) lipids (Avanti Polar Lipids, USA) in a 4:1 ratio in chloroform, followed by solvent evaporation under nitrogen. Lipids were further dried under vacuum for at least 2 h before suspension in liposome buffer (10 mM HEPES pH 7.2, 150 mM NaCl, and 4 mM $CaCl_2$) at a lipid concentration of 1 mM DMPC. The resuspension was subjected to four thaw and freeze cycles of heating to 42 °C followed by agitating, vortexing and plunging into liquid nitrogen. After a fifth heating step at 42 °C, LUVs were generated by 15-fold extrusion at room temperature through a Whatman polycarbonate membrane (Simga-Aldrich, USA) with a pore size of 100 nm. LUVs were monodisperse and had a diameter of 116 ± 3 nm as determined from intensity-weighed size distributions by dynamic light scattering (DLS) (Zetasizer Nano ZS, Malvern, UK) at 25 °C. Mica-SLBs were formed by placing drops of 10-fold diluted LUV solution in liposome buffer onto freshly cleaved mica glued to a glass sample stage and incubation at room temperature overnight in a humidified chamber. Prior to imaging, the lipid bilayer was rinsed with observation buffer (50 mM potassium phosphate pH 7.0), followed by deposition of a 2 µL drop of 2-3 nM freshly diluted FAS in storage buffer and rinsed with observation buffer after incubation for 10 min.

3.5.5 HS-AFM Imaging of FAS

Imaging was performed with a commercial tapping-mode HS-AFM 1.0 apparatus (RIBM, Japan) equipped with a standard scanner with a maximum scan speed of 80 ms/frame.[137] BL-AC10DS-A2 cantilevers with a spring constant of 0.1 N/m, a resonance frequency of 300-700 kHz and a quality factor of ~2 in water (Olympus, Japan) were used. For all measurements, 434 ± 90 nm long amorphous carbon nanofiber tips with a tip apex radius of 5.5 ± 0.9 nm were grown onto the cantilever by electron beam deposition (EBD) (Chapter 3.9.7). The free oscillation amplitude was ~6 nm with a set point of 0.6-0.8, which corresponded to a maximum tapping force of ~190 or 250 pN, depending on FAS immobilization. The momentum transferred to the protein was minimized by short taps of 2.5 μs (Chapter 3.9.8). Both N-His$_{10}$-FAS and FAS-His$_{10}$-C were imaged in observation buffer at room temperature in over 150 FAS movies. Bilayer quality and particle dimensions were first verified in a 400 nm × 400 nm overview scan, followed by zooming in onto individual proteins in 80 nm × 80 nm to 100 nm × 100 nm scans imaging with 100 × 100 pixels to 120 × 120 pixels per image. Most videos were recorded at 2-0.5 fps.

For imaging in the presence of substrate, FAS in observation buffer was supplemented with 4 μM acetyl-CoA and 15 μM NADPH. After 5 min incubation at room temperature, 4 μM malonyl-CoA was added to the mixture followed by subsequent filming.

3.5.6 Image Analysis

Raw data was converted and *z*-aligned by custom-made Python-based software, which was described previously.[88] Movie stacks were transformed to image sequences using ImageJ.[138] Select z-aligned frames were corrected for *xy*-tilting using basic trigonometry followed by Gaussian filtering with MATLAB (R2016b, MathWorks, USA). Subsequently, the frames were aligned in *x*- and *y*-direction by normalized cross-correlation of pixel values higher than a predefined threshold using MATLAB (Chapter 3.9.3). In order to minimize bias from inaccurate alignment due to noise, each aligned movie was visually inspected, and where necessary, frames were discarded before further processing. Variance map calculations, principal component analysis and autocorrelation analysis of the first and second principal components of the particle were carried out in MATLAB (Chapters 3.9.4 and 3.9.5). HS-AFM simulations were performed with custom-made Python-based software using the atomic coordinates of the crystal structure of *S. scrofa* FAS (PDB: 2VZ9), in which the flexibly attached ACP and TE domains are not resolved (Chapter 3.9.9). Simulated AFM images were assigned to selected experimental images in a cross-correlation approach based on sampling FAS in Euler angle space using custom-made Python-based software and MATLAB (Chapter 3.9.10).

Particle dimensions were measured in ImageJ by extracting 2D cross-sectional height profiles of individual frames of *z*-aligned and tilt-corrected movies and averaging the profiles for each corresponding movie. From these averages, the difference between the maximum and minimum *z*-value was determined and averaged, giving the measured height (N-His$_{10}$-FAS: 38 frames, FAS-His$_{10}$-C: 64 frames, FAS on mica: 57 frames). For the dimensions of FAS on mica, the average diameter at 4 nm height was calculated from 2D cross-sectional height profiles. These dimensions were compared to maximum main chain distances in the homologous porcine FAS structure (PDB: 2VZ9).

For particle classification, *z*-aligned images were scaled to a 1:1 pixel to nm ratio in ImageJ. Artificial micrographs for particle picking were generated for each individual movie by assembling its associated tilt-corrected and filtered images into a chronologically tiled image using basic image processing software. Particles were semiautomatically selected and a box size of 65 nm was used in e2boxer from the EMAN2 electron microscopy data processing suite[139] (version 2.07). To prevent misclassification, particles which were located at the edge of images or close to other scanning artifacts were discarded. Particle normalization by Xmipp3 was followed by reference-free 2D classification using CL2D from Xmipp3 in Scipion.[133,140] Correntropy was chosen as the clustering algorithm.[141] Particle classification was carried out for each movie category (mica, SLB-N-His_{10}-FAS, SLB-FAS-His_{10}-C; plus or minus substrate). The number of classes was optimized based on increasing class homogeneity while decreasing class redundancy.

3.6 Associated Content

The authors declare no competing financial interest.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsnano.7b04216.

Purification of *D. melanogaster* FAS and enzymatic activity (S1); overview of NADPH turnover rates during biosynthesis of palmitic acid in metazoan FAS (S2); tilt correction in *x*- and *y*-direction, filtering and alignment with MATLAB (S3); generation of variance maps with MATLAB (S4); autocorrelation of the first and second principal component of HS-AFM images (S5); estimation of minimum spatial resolution and comparison to simulated AFM images (S6); cantilever preparation with electron beam-deposited carbon tips (S7); force and impulse acting on FAS during HS-AFM imaging (S8); simulation of AFM images (S9); correlation between simulated and experimental data (S10)(PDF).

Conformational dynamics of FAS immobilized on mica filmed at 0.5 fps in the presence of substrate (AVI).

3.7 Author Information

3.7.1 Corresponding Authors

*(R. Y. H. Lim) E-mail: roderick.lim@unibas.ch

*(T. Maier) E-mail: timm.maier@unibas.ch

3.7.2 ORCID

Timm Maier: 0000-0002-7459-1363

Roderick Y.H. Lim: 0000-0001-5015-6087

3.7.3 Present Address

#(H.S.T.B.) Janelia Research Campus, Howard Hughes Medical Institute, Ashburn,
VA 20147

3.7.4 Author Contributions

Y.S. maintained the HS-AFM and provided technical support; A.M. wrote Python-based software for *z*-alignment, AFM-simulation, and conformational sampling of FAS in Euler angle space; H.S.T.B. performed initial cloning; F.M.C.B. cloned constructs, expressed, purified, and characterized FAS, prepared lipid bilayer, performed HS-AFM imaging, analyzed data, wrote MATLAB code, and designed experiments with R.Y.H.L. and T.M.; T.M. designed and guided research and data analysis. F.M.C.B. and T.M. wrote the manuscript through contributions of all authors.

3.7.5 Notes

The authors declare no competing financial interest.

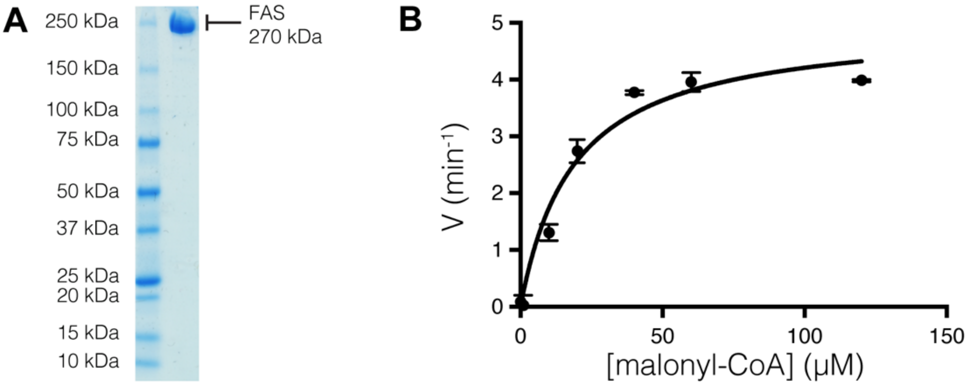
3.8 Acknowledgments

We acknowledge M. Dürrenberger, E. Bieler, and D. Mathys for support with EM and focused ion beam; H. Aragão and P. Rios Flores for advice on liposome preparation; the Biophysics Facility of the Biozentrum for support with dynamic light scattering; and D. Herbst for support with 2D classification. F.M.C.B. acknowledges a Ph.D. fellowship from the Werner Siemens Foundation, and Y.S. is supported by a

Ph.D. fellowship from the Swiss Nanoscience Institute. This work was supported by the Swiss National Science Foundation (SNF) project grant 159696.

3.9 Supporting Information

3.9.1 Purification of *D. melanogaster* FAS and enzymatic activity



Supplementary Figure 3.1 (**A**) SDS-PAGE analysis of FAS after size exclusion chromatography. (**B**) Mean FAS activity and standard deviation of four replicates plotted as reaction rate versus substrate concentration.

3.9.2 Overview of NADPH turnover rates calculated for the biosynthesis of palmitate in metazoan FAS

source	**M_w (kDa)**	**k_{cat} (s^{-1})**	**k_{cat} (s^{-1}) in dimeric FAS[†]**	**cycles (s^{-1}) of complete palmitate formation***	**reference**
Gallus gallus	275	7.3	14.6	1	[142]
Gallus gallus	275	6.8	13.7	1	[143]
Homo sapiens	273	2.1	4.2	0.3	[144]
Rattus norvegicus	273	11	22	1.6	[145]
Rattus norvegicus	273	10.3	20.5	1.5	[146]
Rattus norvegicus	273	10	20	1.4	[66]
Homo sapiens	273	2	4	0.3	[147]
Drosophila melanogaster	266	4.2	8.3	0.6	this publication

[†] FAS is only active in its dimeric form. Each dimer contains two sets of active sites.

*In one complete cycle of palmitate formation 14 NADPH molecules are consumed.

3.9.3 Tilt correction in *x*- and *y*-direction, filtering and alignment with MATLAB

For each image, the tilt angle was calculated trigonometrically and the image rotated accordingly. Noise was reduced by applying a 2D Gaussian filter with a standard deviation of 1 pixel. A reference frame was manually selected based on image quality and the normalized cross-correlation was calculated for each image above a pre-defined pixel value with respect to the reference frame. Based on the maximum of the cross-correlation function, a translation distance was calculated in x and y for frame alignment.

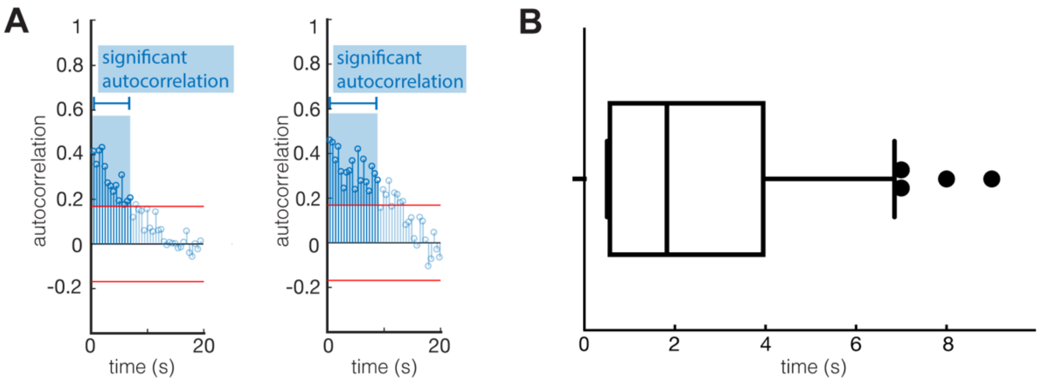
3.9.4 Generation of variance maps with MATLAB

Variance maps were calculated with the aim of visualizing height changes throughout movies through a reduced dimensionality approach with high information content. As the variance is calculated along the first dimension in an array in MATLAB, the tilt-corrected, filtered and aligned images were first concatenated, followed by permutation of the dimensions for calculation of the variance at each position $P_i(x_i,y_i)$ along z. For graphic display of the variance maps, the concatenated movie frames were permutated back to their original dimensions and the results presented as an image with scaled color information representing the amount of changes in height.

3.9.5 Autocorrelation of the first and second principal component of HS-AFM images

For each movie, shape features of FAS were calculated and autocorrelated using MATLAB. An abstracted description of observed FAS molecules was derived in two steps in the form of the first and second principal component of HS-AFM images: First, we applied a pre-defined threshold of 60% of the maximum pixel value to aligned, tilt-corrected and filtered HS-AFM images to select the particle region and set the background pixels to zero. Second, we performed a principal component analysis on the resulting background-corrected images by calculating the covariance matrix using the *cov(x)* function followed by calculating the principal components by the function *pcacov(x)* in MATLAB. The eigenvalues of the first and second principal component of successive images were subjected to autocorrelation analysis using *autocorr(x)* and the autocorrelation was plotted with a confidence interval of 95%.

Autocorrelation is observed for 46% of all movies with more than 30 consecutive frames (112 movies analyzed). 54% of these movies have a frame duration of 500 ms, autocorrelating with a median time of 1.825 s (Supplementary Figure 3.2). In all cases, the movies are sufficiently long with at least 80% of the data points following after the respective autocorrelation time. The median autocorrelation time was calculated for both selective tethering and non-selective immobilization combined. Analysis was not extended to comparing median autocorrelation times between the different immobilization methods due to an insignificant number of movies showing autocorrelation and complying with all criteria such as select frame rate and sufficient data length.



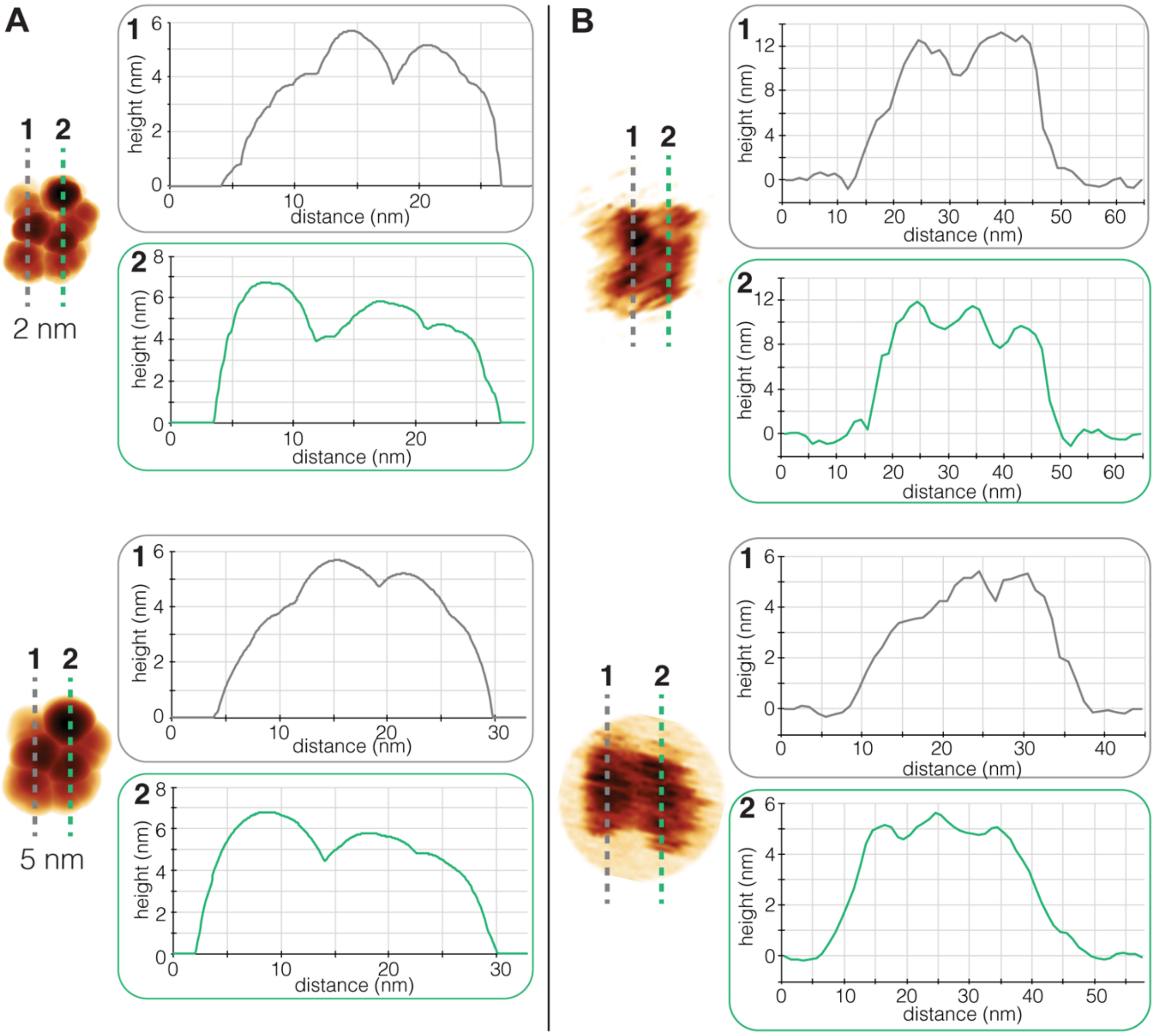
Supplementary Figure 3.2 Autocorrelation plots and box plot.

(**A**) Autocorrelation plots of the first two principal components of FAS in two HS-AFM movies (FAS on mica, 500 ms/frame). Confidence intervals (95%) are marked red. (**B**) Box plot of the autocorrelation times in movies of selectively-tethered FAS and FAS on mica with a frame duration of 500 ms (28 movies); times outside percentiles 10 to 90 are shown as individual points. The median autocorrelation time is 1.825 s.

3.9.6 Estimation of the minimum spatial resolution and comparison to simulated AFM images

Estimations of the minimum spatial resolution are based on distance measurements in image cross-sections using ImageJ. For the minimum lateral resolution, the width at half maximum between two separated feature peaks was calculated. The height difference between separated feature peaks was used as an approximation of the minimum vertical resolution.

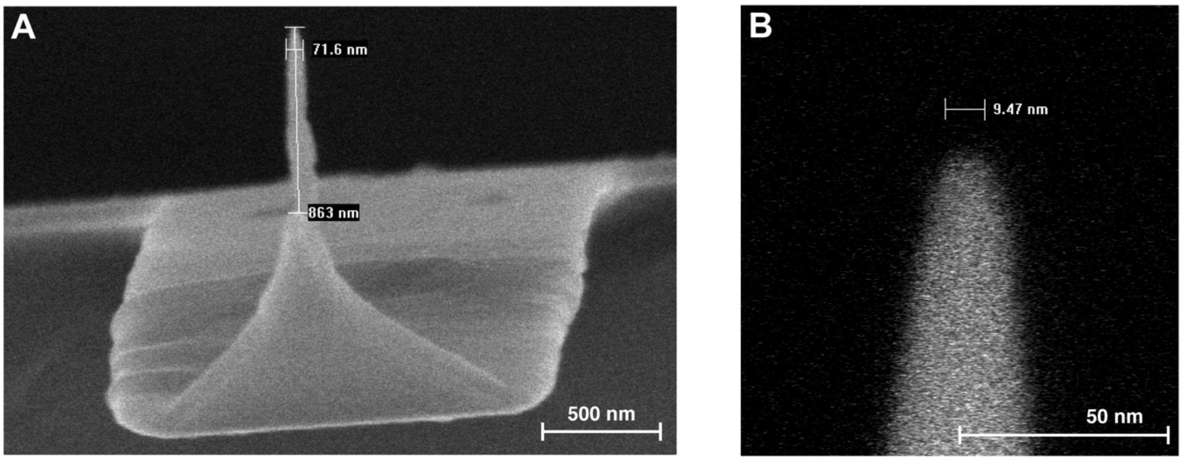
Cross-sectional profiles of experimentally obtained HS-AFM frames were compared to those of simulated AFM images (Supplementary Figure 3.3). Simulated AFM images were generated using the atomic coordinates of the crystal structure of porcine FAS (PDB: 2VZ9), sphere radii of 2-9 nm and tilt angles between 0 and 180° in 10° increments.



Supplementary Figure 3.3 Comparison of cross-sectional profiles of examples of simulated and experimental HS-AFM images.

Cross-sections are indicated by dashed, colored lines. (**A**) Cross-sections of simulated AFM images using a tip apex radius of 2 nm and 5 nm. (**B**) Baseline-corrected 3-pixel wide cross-sections of *z*-aligned experimentally obtained HS-AFM frames from two different movies of FAS on mica. Orientation of particles in (**A**) and (**B**) is not matched.

3.9.7 Cantilever preparation with electron beam-deposited carbon tips



Supplementary Figure 3.4 (**A**) Scanning electron microscopic image of a carbon nanofiber grown onto a cantilever tip. (**B**) Close-up of a microfabricated nanofiber tip.

Commercial cantilever tips were supplemented with sharp amorphous carbon nanofibers, as described elsewhere.[131] Briefly, electron beam deposition of evaporating phenol was used to grow a carbon stylus on top of the original cantilever tip (BL-AC10DS-A2, Olympus, Japan) at low pressure (10^{-6} mBar) in a scanning electron microscope (FEI XL30 FEG ESEM, Thermo Fisher Scientific, USA). By applying a small aperture of 30 µm and an acceleration voltage of 5 kV, carbon tips of 434 ± 90-nm length with a tip apex radius of 5.5 ± 0.9 nm were grown. To compensate for the tilted sample stage, tips were grown at an angle of 10°. Cantilevers were recycled by plasma etching and treatment with a focused ion beam (FEI Helios NanoLab 650, Thermo Fisher Scientific, USA).

3.9.8 Force and impulse acting on FAS during HS-AFM imaging

Force effects on the protein during tapping-mode HS-AFM imaging can be quantitatively depicted by the impulse. It describes the change of momentum over the short fraction of time during which the sample is tapped by the tip. The maximum force F_{max} exerted on a sample in AFM tapping mode can be described as:[85]

$$F_{max} = \left(\frac{k_c}{Q_c}\right) \times \left[A_0(1-r) + h_0 sin\left(\frac{\theta}{2}\right)\right]$$

where k_c is the spring constant of the cantilever, Q_c the cantilever quality factor in water, A_0 the free oscillation amplitude, r the setpoint, h_0 the sample height and θ the phase shift. With k_c = 0.1 N·m^{-1}, Q_c = 2, f_c = 400 kHz (supplied by manufacturer), A_0 = 6 nm, r = 0.6-0.8, and θ ~ 20°, for FAS on mica (h_0 = 7.7 nm), the maximum tapping force is 190 pN and for FAS immobilized on SLB (h_0 = 15 nm) it is estimated to be 250 pN. With a force acting time of 2.5 μs in these experiments, the impulse is in the femto regime and thus no significant disturbance of the protein is expected.

The energy dissipation of the oscillating cantilever in a tapping-mode HS-AFM is given by:[85]

$$E_{dis} = \frac{1}{2} k_c \left(A_0^2 - A_S^2\right)/Q_c$$

With the free oscillation amplitude A_0= 6 nm and the amplitude set point A_S= 3.6-4.8 nm, the maximum energy dissipated at each tap is 140 k_BT (T = 298.15 K). However, with the contact radius being only a few nm, most of the energy is delivered to the surrounding buffer.

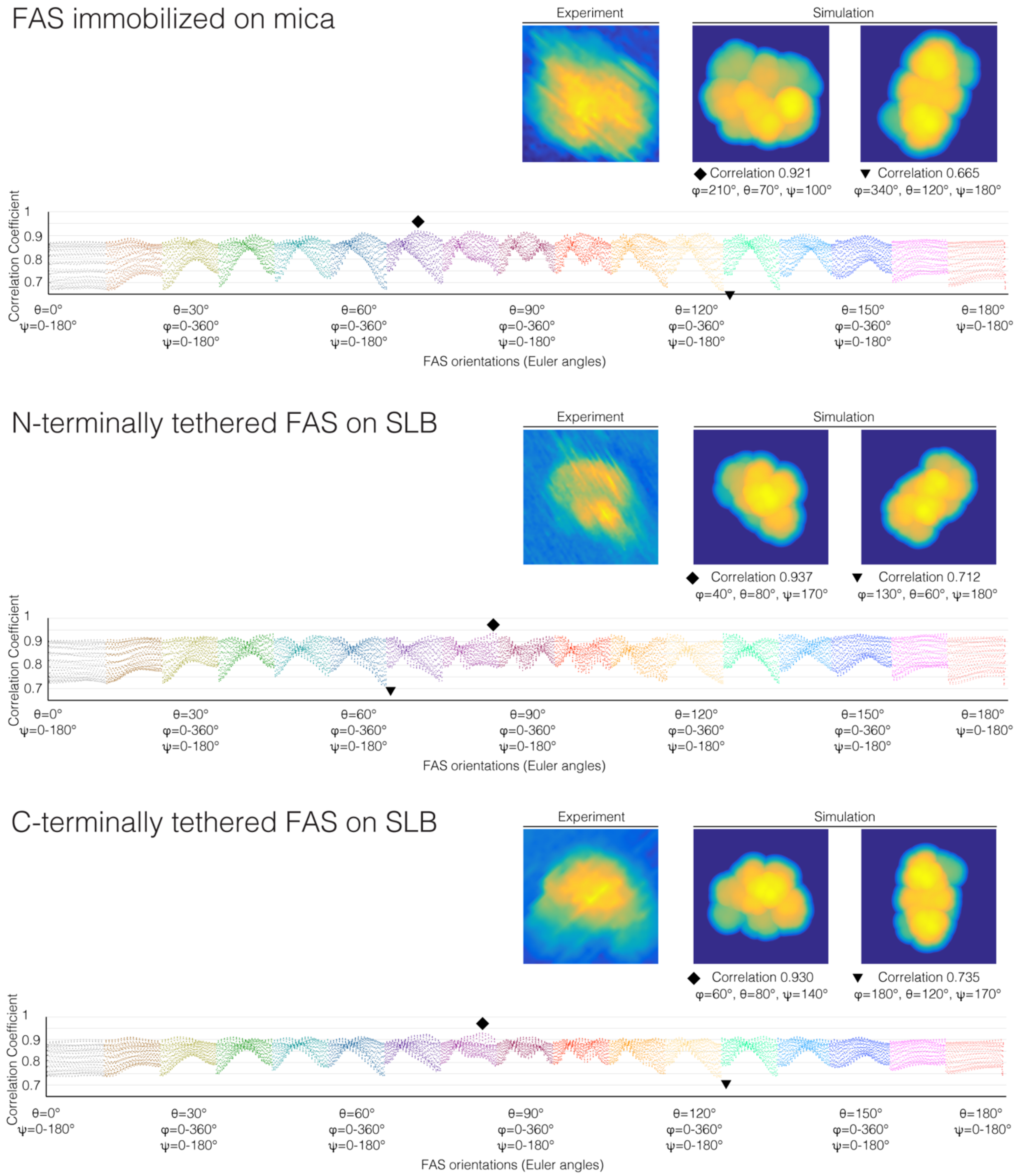
3.9.9 Simulation of AFM images

AFM images were simulated using in-house written Python software. The AFM topography is calculated based solely on the non-elastic contact between the tip and the protein surface. The tip is modelled using a sphere-cone-rod model described by three parameters: sphere radius, cone semi angle and rod radius.

3.9.10 Correlation between simulated and experimental data

To obtain a direct interpretation of selected HS-AFM images, we applied a correlation-based comparison to simulated AFM images based on atomic coordinates. The available atomic model PDB 2VZ9 includes only the core multienzyme, while images were acquired for the biologically active enzyme (see also S1B), which additionally contains the mobile ACP and TE domains (see Figure 3.1 for details). The static crystallographic model only represents a single FAS conformation, and may not be suited for comparison to diverging dynamic conformations of FAS visually observed in HS-AFM movies, *e.g.* the rotation of large subsegments.

Simulated AFM images of PDB files of FAS (original PDB: 2VZ9) rotated around its Euler angles (φ, θ, ψ) in 10° increments were cross-correlated with experimental raw images in MATLAB. In a first step, rotation of PDB file 2VZ9 around the θ and ψ angle was performed utilizing in-house written Python software. Simulated AFM images of the PDB files were generated as described in chapter 3.9.9 with a sphere radius of 2 nm, a cone semi-angle of 10° and a rod radius of 10 nm. Using MATLAB, the simulated AFM images were rotated in 10° increments, thereby performing rotation of FAS around the φ angle, and subsequently subjected to calculation of the cross-correlation.



Supplementary Figure 3.5 Correlation coefficients for all calculated FAS orientations and the respective raw image of each immobilization strategy.

Images (top to bottom) correspond to those depicted in Figure 3.2B and Figure 3.3D, E. Simulated image orientations with the highest (◆) and lowest (▼) correlation coefficient are depicted.

4 Discussion and Outlook

4.1 Structural conservation of the modifying region in divergent metazoan FAS

The first crystal structure of an excised modifying region of FAS shows that even divergent classes of metazoan FAS adopt a conserved ground-state conformation. Domains are connected *via* interdomain contacts, with the KR acting as a connector between the lateral non-catalytic domains and the core DH and ER domains. An artificial sequence is incorporated into a β-sheet at the C-terminus of the KR, despite the presence of the natural polypeptide stretch. This structure of *D. melanogaster* FAS explains large sequential variations in the ΨME of insects compared to that of mammals, thereby invalidating the proposed role of ΨME in providing steric guidance to the ACP.

Striking differences to iterative reducing PKS include a higher rigidity of the modifying region in metazoan FAS, which is maintained by conserved domain interfaces. In contrast, domains in the PKS modifying part are connected *via* linkers, facilitating conformational flexibility.[63] Particularly the DH dimers show a distinct arrangement in metazoan FAS and PKS. While the PKS DH protomers form a linear dimer with a substantial dimer interface, FAS features a V-shaped DH dimer with a miniscule interface. [63,148] Surprisingly, even without a connection to the condensing region, the V-shaped dimer arrangement is intact in this crystal structure, strongly suggesting this dimerization to be an inherent structural feature of FAS. Both dimer forms provide unique advantages for each system. A linear DH arrangement likely provides less surface area for interactions with adjacent domains. In combination with its linker-based organization, the modifying domain in PKS allows for variable conformational coupling, which is necessary for the formation of large modular assembly lines. Indeed, conformational dynamics were observed within the core of a PKS modifying domain.[63] Contrarily, the V-shaped DH dimer in FAS directly interacts with the KR and ER domains, thereby restraining the mobility of the central ER/DH core, while still allowing for flexibility of the lateral domains. A more rigid core might enhance the turnover of the enzyme by providing higher local concentrations of the active sites.

The structurally stable ground state structure of the modifying region represents metazoan FAS in a resting state. Open questions involve single-molecule observations of dynamics during turnover. Due to difficulties in labeling proteins in eukaryotic expression hosts, new methods are required.

4.2 A new approach to study multienzyme conformational dynamics

The results presented in chapter 3 describe the development of a new protocol to study multienzyme conformational dynamics by filming with HS-AFM. For the first time, conformational dynamics of single carrier-protein dependent multienzyme were visualized in solution. Since HS-AFM is a surface scanning technique, alternate views of the multienzyme require selective tethering. Combinations of different molecular tethering strategies that preserve the functional dynamics of the multienzyme were used to offer various viewing angles. The spatiotemporal resolution allows for identification of slower large-scale conformational dynamics in the range of 2-5 nm at 10 fps. In FAS, we observed rotation of the two major segments relative to each other as well as an opening and closing of a cleft, separating these segments. Reference-free 2D classification was shown to be a suitable tool for identification of conformations from HS-AFM data, enabling analysis of their transitions in a temporal context (Figure 4.1).

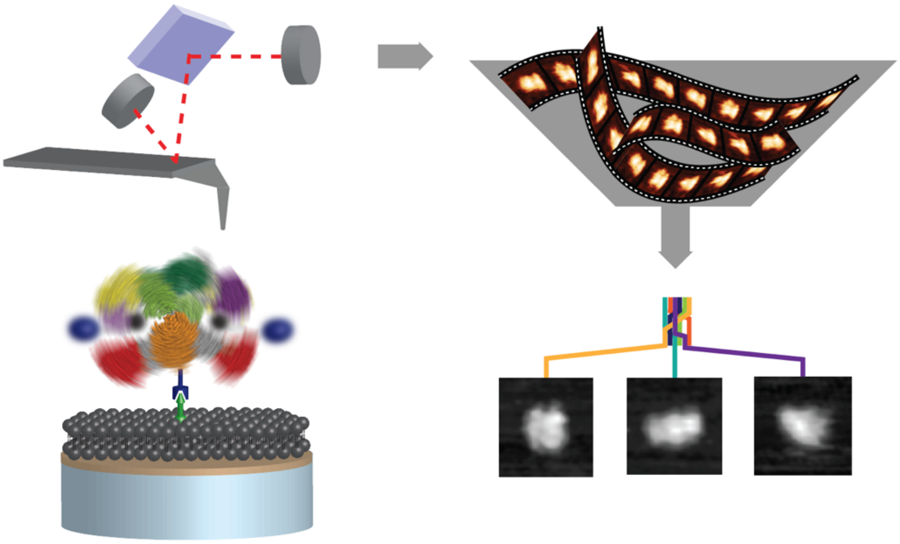


Figure 4.1 Identification of conformational variability by HS-AFM imaging. Schematic representation of experimental setup and data analysis. Individual states of the multienzyme can be identified and classified across movies by reference-free 2D class averaging. Figure adapted with permission from Benning, F. M. C.; Sakiyama, Y.; Mazur, A.; Bukhari, H. S. T.; Lim, R. Y. H.; Maier, T. High-Speed Atomic Force Microscopy Visualization of the Dynamics of the Multienzyme Fatty Acid Synthase. *ACS Nano*. **2017**, *11*, 10852-10859. Copyright 2017 American Chemical Society.

HS-AFM, while previously only applied to membrane proteins, crystalline substrate arrays and macromolecules with a rigid overall structure, with this protocol renders to be a suitable technique for studying highly flexible, soluble proteins such as multienzymes.[86,88,149,150] In the past, studies of multienzyme conformational dynamics were based on – often hybrid – approaches of electron microscopy, small angle X-ray scattering and X-ray crystallography.[12,64,67,68] All of these techniques are based on averaged information from many particles. While any of these methods provides either high resolution information outside a temporal context or dynamics data at low resolution, HS-AFM fills the gap as a method with the capabilities to record the dynamics of single multienzymes at high spatial and temporal resolution.

Brignole *et al.* visualized a wide variety of FAS conformations using negative stain electron microscopy.[64] It was shown that the distribution of conformations was affected by catalytic mutants and substrate availability. While EM provides a large number of snapshots, it does not yield any temporal correlation, which allows following successive changes. Besides, EM requires averaging over thousands of molecules, which might obscure changes to some extent. The HS-AFM approach depicted here, allows recording of unrestrained dynamics in near-native aqueous solution through selective tethering of the multienzyme, thereby diminishing measurement artefacts. It further provides access to conformational dynamics which may be hidden in other techniques due to a limited number of viewing angles. While EM studies only give a distribution of conformational states, our protocol enables the additional identification of conformational transitions with the potential to reconstitute a complete temporal succession of conformations. With this method, it is now possible to monitor real-time changes in conformational dynamics upon addition of substrates or further effectors.

4.3 Hybrid approaches to study dynamics and structures of multienzymes

Understanding the conformational variability in FAS and potentially linking specific conformations to functional states, requires comprehension of the underlying molecular interactions such as domain interplay and linker connections within the multienzyme.

High-resolution structure determination of the modifying region of insect FAS gives an explanation of how interdomain contacts and linkers stabilize the present conformation. This structure of insect FAS expands the current knowledge of this highly dynamic region by revealing structural similarities and differences in related metazoan FAS and PKS. Conformational conservation of the complete modifying region in metazoan FAS suggests the existence of a conformational low-energy ground state. Large-scale conformational dynamics in carrier protein-dependent multienzymes are too large to be studied by NMR spectroscopy.[151] Moreover, protein labeling in eukaryotic cells is expensive and difficult, with new methods just emerging.[152] However, real-time filming of multienzymes is capable of sampling the conformational space of single FAS molecules in near-native conditions. It provides real-time temporal information such as the sequence of conformational states and their transitions. It can be used to analyze insights from crystal structures such as the suggested presence of a conformational resting state and the influence of substrate on the conformational pattern of FAS. The two projects in this thesis provide opportunities for linking high-resolution static information on FAS with medium-resolution dynamic data with the objective of gaining a deeper understanding of conformational dynamics in multienzymes.

4.4 Conformational dynamics as a basis for understanding assembly lines

The next step in understanding FAS dynamics lies in elucidating the underlying structural basis of its conformational dynamics at a resolution capable of separating individual domains. This includes deciphering the influence of structural properties such as linker lengths, as well as functional features, involving substrate availability and active site mutations on dynamics.

An in-depth analysis of functional dynamics in FAS requires a large-scale data collection strategy to achieve complete mapping of all conformational states. This thesis provides a direct starting point for large-scale movie acquisition. With the objective of gaining a comprehensive overview of functional dynamics, FAS can now be imaged using varying substrate conditions, active site mutations and shortened or extended linker regions. Unambiguous differentiation of individual domains in FAS by HS-AFM requires a spatial resolution of below 2 nm. Improvement of the lateral spatial resolution to ~1-2 nm can be achieved by sharpening the carbon nanofiber stylus on the cantilever tip by plasma etching.[86] However, high-throughput cantilever preparation and plasma etching are indispensable to counteract the low yields in tip sharpening.[131,132] Timescales of FAS dynamics can be adjusted to the temporal resolution by decelerating enzymatic motion using higher viscosity media or measuring at temperatures lower than room temperature (Figure 4.2).

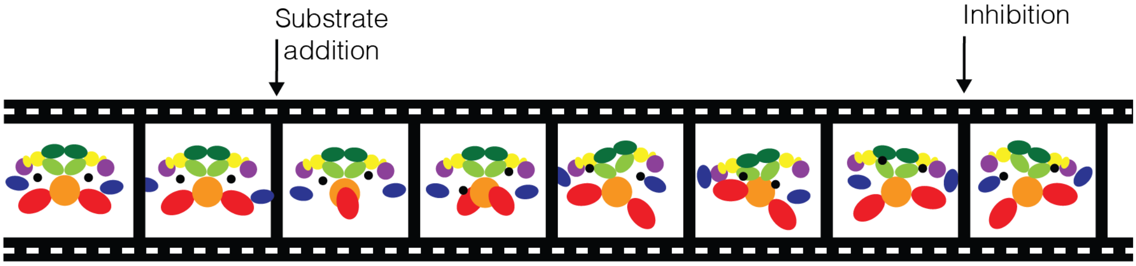


Figure 4.2 Functional dynamics of FAS.
Schematic representation of a movie from HS-AFM imaging, showing potential conformational variability of FAS. Large-scale data acquisition is required for a complete mapping of all conformational states. HS-AFM filming enables the characterization of conformational changes as well as transition times upon substrate addition or inhibition.

More comprehensive insights into functional dynamics are expected by complementing HS-AFM data with distance information obtained by single-molecule FRET. Site-specific fluorescent labels on lateral domains like the MAT in the condensing region and ΨME in the modifying region can provide temporal information on domain interactions which indicate opening and closing of the reaction cleft. Combinatorial multicolor FRET analysis of fluorescently labeled ACP can yield temporal insights into interaction with two FAS domains, with the ultimate goal of following ACP in the reaction cleft.[76]

FAS serves as a blueprint for related carrier protein-dependent biosynthetic machines such as polyketide synthases (PKS) and nonribosomal peptide synthetases (NRPS).[153] By forming large assembly lines, modular PKSs and NRPSs produce a tremendous variety of biologically and pharmacologically active natural products.[91,154] Chemical diversity is achieved by the availability and activity of select catalytic domains in each module in combination with *trans*-acting enzymes in some systems.[50] While static structural information is available for individual domains and regions, little is known about inter- and intramodular dynamics.[153] The HS-AFM approach developed in this thesis and insights from FAS dynamics may advance the understanding of assembly line dynamics, its formation and regulation.

4.5 Applications of biomimetic nanomachines

Multistep reactions require the selective transfer of reactants from one catalytic center to the next. In all kingdoms of life, natural selection has formed various strategies for efficient substrate transfer. These adaptations all share a combination of advantages such as protecting unsoluble intermediates from aggregation, shielding them from access by noncognate cellular enzymes and providing high local concentrations of substrates and catalysts. Tunnel formation was observed in some enzymes, providing substrate transfer between two to three distant active sites.[9] A higher degree of versatility is achieved by carrier proteins, which can shuttle substrates between multiple catalytic sites. Carrier protein-dependent enzyme systems rely on specific protein contacts between catalytic domains and the carrier protein. In carrier protein-dependent multifunctional enzymes such as metazoan FAS, PKS and NRPS, all functional domains are integrated into the polypeptide chain. The carrier protein constitutes an integral part of the protein and is connected *via* a flexible tether to the enzyme. Functional contacts between enzymatic domains and the carrier protein are accompanied by large-scale conformational dynamics in most carrier-protein dependent enzymes.[11,12,67] This thesis presents a basis for advancing our understanding of the functional role of conformational dynamics in multienzymes.

The next steps lie in understanding how multienzyme dynamics are influenced by the cellular environment. The HS-AFM approach depicted in this thesis may be expanded by filming in cytoplasmic extract or by adding molecular interaction partners to the tethered multienzyme. Furthermore, the lipid bilayer in the experimental setup may be altered to give a more complex composition, including lipid rafts, for the implementation of signaling cascades. Identifying molecules, which block conformational flexibility of multienzymes, may provide a way to detect the relevance of multienzyme dynamics in a cellular environment.

While HS-AFM studies may answer how FAS dynamics are affected by interaction partners, synthetic cells provide the means to understand the function of dynamics of individual FAS molecules within a cellular context. Typically lipid or polymer vesicles, synthetic cells are more realistic mimics of simple cells than a 2D surface.[155] As a surface scanning technique, HS-AFM is not capable of filming dynamics within cells. However, super-resolution microscopy (SRM) in single synthetic cells provides a way to discern individual active FAS enzymes from inactive ones, to potentially uncover the influence of specific cellular situations on the dynamics of single FAS. Interaction partners can be localized and structurally characterized at high resolution in a structural snapshot of the cellular landscape by combining SRM with cryo-electron tomography.[81]

Multistep reactions are widely distributed among industrial applications. Spatial or temporal segregation of individual reactions steps, is key for economic and sustainable production (Figure 4.3). One such approach is the exploitation of engineered PKS and NRPS assembly lines for tailored drug production (Figure 4.3A). Providing a toolbox for performing highly complex organic chemistry, these multienzymes harbor significant biotechnological potential for the rational production of novel compounds including therapeutics and biofuels.[156] The majority of bioactive natural products feature a high intrinsic potency. This is however often associated with low bioavailability or side effects. While the structural complexity of natural products renders them extremely difficult targets for chemical approaches, re-engineering PKS and NRPS systems are utilized to generate natural product analogs with improved pharmacokinetics.[157] Biosynthetic engineering efforts include the manipulation of catalytic functions within modules, altering module arrangement, as well as manipulating the number of modules in the assembly line. Efficient operation of tailored assembly lines requires not only static information of correct inter- and intramodular interactions but equally important, the preservation of conformational dynamics between *cis*- and *trans*-acting enzymes. Insights from studies on FAS conformational dynamics, as well as HS-AFM filming of assembly lines, provide a basis for the comprehensive analysis of assembly line dynamics with the ultimate goal of engineering new biosynthetic factories.

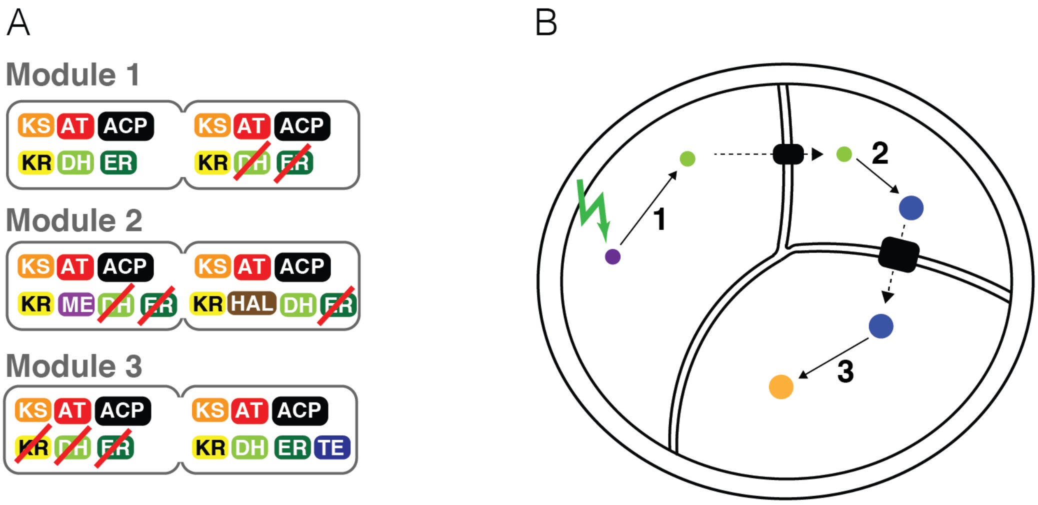


Figure 4.3 Applications of multistep reactions.

(A) Rationally designed natural products may be generated from engineered assembly lines. Schematic representation of three modules containing two sets of functional domains each. Tailored production of compounds is achieved by selectively inactivating catalytic domains (red bar), incorporating additional functional domains (halogenase, brown; methyltransferase, violet; thioesterase, blue) and altering the order of modules. **(B)** Schematic representation of a compartmentalized vesicle. Each individual reaction step is confined to a compartment. The product shuttles through a nanopore (black rectangles) to the next compartment, serving as a reactant for the following reaction. Reactions may be triggered by photoactivaton (green flash).

Besides carrier protein systems, an alternative strategy for the spatial and temporal segregation of individual reactions from multistep processes, is the use of compartmentalized synthetic cells for *in situ* chemical synthesis (Figure 4.3B).[155,158] Individual enzymatic steps are either segregated by lipid or polymer membranes, allowing substrate transfer *via* nanopores, or by controlled enzyme activation.[159] Regulation of enzyme activity may be achieved by signaling proteins in the vesicle membrane, which are activated by external stimuli, or by the photoactivation of enzymes.[157,159]

Evolution has formed biological systems such as cells, certain cellular compartments and carrier protein-dependent multienzymes into highly regulated factories. Bio-inspired nanomachines enable efficient turnover of multistep reactions by mimicking biological strategies such as compartmentalization or the exploitation of carrier protein systems. Thus, deciphering the mode of action of these molecular machines will boost the biomimetic production of complex chemicals.

5 Acknowledgments

First, I would like to thank you, Timm, for the opportunity to work on this fantastic project. Thanks for the trust and the well-balanced mixture of guidance and freedom I experienced, while expanding my own (scientific) limits.

Many thanks to Sebastian and Rod for being part of my PhD advisory committee and for helpful advice. Thanks, Rod, for a great collaboration on the HS-AFM project.

Special thanks to Yusuke for being a fun and helpful collaborator on the HS-AFM project. Thank you, Adam, for your support and good discussions.

Thanks to Anna for being a great friend and always providing me with food. One day we will organize our “Big F complexes” conference! Thanks, DJ Roman, for providing us with decent music and for sharing your scientific experience. Thank you, Moritz, for establishing the Schlumi tradition. Many thanks to Shubham, my fun and fearless conference buddy, for culinary excursions to India. Thanks, Domi, for inspiring discussions and passionate happy hour preparations. Special thanks to Imsi for being a great lab mate, bringing more Valais to Basel and fantastic skiing trips. Thanks, Eddi, for all you do for the lab and for sharing your experience. Big thanks to all current and former members of the Maier lab for making this such a great team to be part of. I would like to express my gratitude to Beat, Barbara and Adela for fantastic maintenance and glassware service on the 3rd floor.

Many thanks to Habib, my meeen, for great discussions ranging from FAS to rockets and for sharing your cloning enthusiasm and expertise. Special thanks to the ladies, Fra, Johanna, Anja and Leonie for the good times in and outside of the Biozentrum and for saving my last birthday.

Big thanks to my team mates from fencing, Nora, Nadine, Georg, Meli and Annekäthi, for your fantastic friendship and support in good and rough times.

Last but not least, I would like to thank my family Frank, Sybille, Leo, Ferdi and Louise for always supporting me.

6 References

1. Bugg, T. D. H. From Jack Beans to Designer Genes. In *Introduction to Enzyme and Coenzyme Chemistry*, 3rd ed.; John Wiley & Sons, Ltd: Chichester, 2012; pp 1-6.
2. Polgár, L. The catalytic triad of serine peptidases. *Cell. Mol. Life Sci.* **2005**, *62*, 2161-2172.
3. Blow, D. M.; Birktoft, J. J.; Hartley, B. S. Role of a Buried Acid Group in the Mechanism of Action of Chymotrypsin. *Nature* **1969**, *221*, 337-340.
4. Ekici, O. D.; Paetzel, M.; Dalbey, R. E. Unconventional serine proteases: variations on the catalytic Ser/His/Asp triad configuration. *Protein Sci.* **2008**, *17*, 2023-2037.
5. Bugg, T. D. H. Enzymes Are Wonderful Catalysts. In *Introduction to Enzyme and Coenzyme Chemistry*, 3rd ed.; John Wiley & Sons, Ltd: Chichester, 2012; pp 26-49.
6. Radzicka, A.; Wolfenden, R. A proficient enzyme. *Science* **1995**, *267*, 90-93.
7. Perham, R. N. Self-assembly of biological macromolecules. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **1975**, *272*, 123-136.
8. Reed, L. J. Multienzyme complexes. *Accounts Chem. Res.* **1974**, *7*, 40-46.
9. de Cima, S.; Polo, L. M.; Diez-Fernandez, C.; Martinez, A. I.; Cervera, J.; Fita, I.; Rubio, V. Structure of human carbamoyl phosphate synthetase: deciphering the on/off switch of human ureagenesis. *Sci. Rep.* **2015**, *5*, 16950.
10. Perham, R. N. Swinging arms and swinging domains in multifunctional enzymes: catalytic machines for multistep reactions. *Annu. Rev. Biochem.* **2000**, *69*, 961-1004.
11. Maier, T.; Leibundgut, M.; Boehringer, D.; Ban, N. Structure and function of eukaryotic fatty acid synthases. *Q. Rev. Biophys.* **2010**, *43*, 373-422.
12. Hagmann, A.; Hunkeler, M.; Stuttfeld, E.; Maier, T. Hybrid Structure of a Dynamic Single-Chain Carboxylase from Deinococcus radiodurans. *Structure* **2016**, *24*, 1227-1236.
13. Rohrig, F.; Schulze, A. The multifaceted roles of fatty acid synthesis in cancer. *Nat. Rev. Cancer* **2016**, *16*, 732-749.
14. Lynen, F. Biosynthesis of saturated fatty acids. *Fed. Proc.* **1961**, *20*, 941-951.
15. Brindley, D. N.; Matsumura, S.; Bloch, K. Mycobacterium phlei Fatty Acid Synthetase - A Bacterial Multienzyme Complex. *Nature* **1969,** *224*, 666-669.
16. Schweizer, E.; Kniep, B.; Castorph, H.; Holzner, U. Pantetheine-Free Mutants of the Yeast Fatty-Acid-Synthetase Complex. *Eur. J. Biochem.* **1973**, *39*, 353-362.
17. Stoops, J. K.; Arslanian, M. J.; Oh, Y. H.; Aune, K. C.; Vanaman, T. C.; Wakil, S. J. Presence of two polypeptide chains comprising fatty acid synthetase. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 1940-1944.
18. Enderle, M.; McCarthy, A.; Paithankar, K. S.; Grininger, M. Crystallization and X-ray diffraction studies of a complete bacterial fatty-acid synthase type I. *Acta Cryst. F* **2015**, *71*, 1401-1407.

19. Leibundgut, M.; Maier, T.; Jenni, S.; Ban, N. The multienzyme architecture of eukaryotic fatty acid synthases. *Curr. Opin. Struct. Biol.* **2008**, *18*, 714-725.
20. Wakil, S. J.; Titchener, E. B.; Gibson, D. M. Evidence for the participation of biotin in the enzymic synthesis of fatty acids. *Biochim. Biophys. Acta* **1958**, *29*, 225-226.
21. 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. *J. Biol. Chem.* **1995**, *270*, 26538-26542.
22. Massengo-Tiasse, R. P.; Cronan, J. E. Diversity in enoyl-acyl carrier protein reductases. *Cell. Mol. Life Sci.* **2009**, *66*, 1507-1517.
23. Mocibob, M.; Ivic, N.; Bilokapic, S.; Maier, T.; Luic, M.; Ban, N.; Weygand-Durasevic, I. Homologs of aminoacyl-tRNA synthetases acylate carrier proteins and provide a link between ribosomal and nonribosomal peptide synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 14585-14590.
24. Cronan, J. E.; Thomas, J. Bacterial fatty acid synthesis and its relationships with polyketide synthetic pathways. *Methods Enzymol.* **2009**, *459*, 395-433.
25. Rock, C. O.; Jackowski, S. Forty years of bacterial fatty acid synthesis. *Biochem. Biophys. Res. Commun.* **2002**, *292*, 1155-1166.
26. Maier, T.; Jenni, S.; Ban, N. Architecture of mammalian fatty acid synthase at 4.5 Å resolution. *Science* **2006**, *311*, 1258-1262.
27. Jenni, S.; Leibundgut, M.; Boehringer, D.; Frick, C.; Mikolasek, B.; Ban, N. Structure of fungal fatty acid synthase and implications for iterative substrate shuttling. *Science* **2007**, *316*, 254-261.
28. Jenni, S.; Leibundgut, M.; Maier, T.; Ban, N. Architecture of a fungal fatty acid synthase at 5 Å resolution. *Science* **2006**, *311*, 1263-1267.
29. Asturias, F. J.; Chadick, J. Z.; Cheung, I. K.; Stark, H.; Witkowski, A.; Joshi, A. K.; Smith, S. Structure and molecular organization of mammalian fatty acid synthase. *Nat. Struct. Mol. Biol.* **2005**, *12*, 225-232.
30. Kolodziej, S. J.; Penczek, P. A.; Stoops, J. K. Utility of Butvar support film and methylamine tungstate stain in three-dimensional electron microscopy: agreement between stain and frozen-hydrated reconstructions. *J. Struct. Biol.* **1997**, *120*, 158-167.
31. Ploskon, E.; Arthur, C. J.; Evans, S. E.; Williams, C.; Crosby, J.; Simpson, T. J.; Crump, M. P. A mammalian type I fatty acid synthase acyl carrier protein domain does not sequester acyl chains. *J. Biol. Chem.* **2008**, *283*, 518-528.
32. Schweizer, E.; Hofmann, J. Microbial Type I Fatty Acid Synthases (FAS): Major Players in a Network of Cellular FAS Systems. *Microbiol. Mol. Biol. Rev.* **2004**, *68*, 501-517.
33. Smith, S.; Witkowski, A.; Joshi, A. K. Structural and functional organization of the animal fatty acid synthase. *Prog. Lipid Res.* **2003**, *42*, 289-317.
34. Heath, R. J.; Su, N.; Murphy, C. K.; Rock, C. O. The enoyl-[acyl-carrier-protein] reductases FabI and FabL from Bacillus subtilis. *J. Biol. Chem.* **2000**, *275*, 40128-40133.
35. Massengo-Tiasse, R. P.; Cronan, J. E. Vibrio cholerae FabV defines a new class of enoyl-acyl carrier protein reductase. *J. Biol. Chem.* **2008**, *283*, 1308-1316.

36. Baldock, C.; Rafferty, J. B.; Stuitje, A. R.; Slabas, A. R.; Rice, D. W. The X-ray structure of Escherichia coli enoyl reductase with bound NAD+ at 2.1 Å resolution. *J. Mol. Biol.* **1998**, *284*, 1529-1546.
37. Persson, B.; Hedlund, J.; Jornvall, H. Medium- and short-chain dehydrogenase/reductase gene and protein families : the MDR superfamily. *Cell. Mol. Life Sci.* **2008**, *65*, 3879-3894.
38. Maier, T.; Leibundgut, M.; Ban, N. The crystal structure of a mammalian fatty acid synthase. *Science* **2008**, *321*, 1315-1322.
39. Wakil, S. J.; Abu-Elheiga, L. A. Fatty acid metabolism: target for metabolic syndrome. *J. Lipid Res.* **2009**, *50 Suppl*, S138-S143.
40. Pegorier, J. P.; Le May, C.; Girard, J. Control of gene expression by fatty acids. *J. Nutr.* **2004**, *134*, 2444s-2449s.
41. Davies, M. N.; O'Callaghan, B. L.; Towle, H. C. Glucose activates ChREBP by increasing its rate of nuclear entry and relieving repression of its transcriptional activity. *J. Biol. Chem.* **2008**, *283*, 24029-24038.
42. Hardie, D. G.; Pan, D. A. Regulation of fatty acid synthesis and oxidation by the AMP-activated protein kinase. *Biochem. Soc. Trans.* **2002**, 30, 1064-1070.
43. McGarry, J. D.; Leatherman, G. F.; Foster, D. W. Carnitine palmitoyltransferase I. The site of inhibition of hepatic fatty acid oxidation by malonyl-CoA. *J. Biol. Chem.* **1978**, *253*, 4128-4136.
44. Svensson, R. U.; Parker, S. J.; Eichner, L. J.; Kolar, M. J.; Wallace, M.; Brun, S. N.; Lombardo, P. S.; Van Nostrand, J. L.; Hutchins, A.; Vera, L.; Gerken, L.; Greenwood, J.; Bhat, S.; Harriman, G.; Westlin, W. F.; Harwood, H. J., Jr.; Saghatelian, A.; Kapeller, R.; Metallo, C. M.; Shaw, R. J. Inhibition of acetyl-CoA carboxylase suppresses fatty acid synthesis and tumor growth of non-small-cell lung cancer in preclinical models. *Nat. Med.* **2016**, *22*, 1108-1119.
45. Buckley, D.; Duke, G.; Heuer, T. S.; O'Farrell, M.; Wagman, A. S.; McCulloch, W.; Kemble, G. Fatty acid synthase - Modern tumor cell biology insights into a classical oncology target. *Pharmacol. Ther.* **2017**, *177*, 23-31.
46. Smith, S.; Tsai, S.-C. The type I fatty acid and polyketide synthases: a tale of two megasynthases. *Nat. Prod. Rep.* **2007**, *24*, 1041-1072.
47. Xu, W.; Qiao, K.; Tang, Y. Structural analysis of protein-protein interactions in type I polyketide synthases. *Crit. Rev. Biochem. Mol. Biol.* **2013**, *48*, 98-122.
48. Weissman, K. J.; Leadlay, P. F. Combinatorial biosynthesis of reduced polyketides. *Nat. Rev. Microbiol.* **2005**, *3*, 925-936.
49. Keatinge-Clay, A. T. The structures of type I polyketide synthases. *Nat. Prod. Rep.* **2012**, *29*, 1050-1073.
50. Helfrich, E. J.; Piel, J. Biosynthesis of polyketides by trans-AT polyketide synthases. *Nat. Prod. Rep.* **2016**, *33*, 231-316.
51. Storm, P. A.; Herbst, D. A.; Maier, T.; Townsend, C. A. Functional and Structural Analysis of Programmed C-Methylation in the Biosynthesis of the Fungal Polyketide Citrinin. *Cell Chem. Biol.* **2017**, *24*, 316-325.
52. Keatinge-Clay, A. T. A Tylosin Ketoreductase Reveals How Chirality Is Determined in Polyketides. *Chem. Biol.* **2007**, *14*, 898-908.

53. Pappenberger, G.; Benz, J.; Gsell, B.; Hennig, M.; Ruf, A.; Stihle, M.; Thoma, R.; Rudolph, M. G. Structure of the human fatty acid synthase KS-MAT didomain as a framework for inhibitor design. *J. Mol. Biol.* **2010**, *397*, 508-519.
54. Hardwicke, M. A.; Rendina, A. R.; Williams, S. P.; Moore, M. L.; Wang, L.; Krueger, J. A.; Plant, R. N.; Totoritis, R. D.; Zhang, G.; Briand, J.; Burkhart, W. A.; Brown, K. K.; Parrish, C. A. A human fatty acid synthase inhibitor binds beta-ketoacyl reductase in the keto-substrate site. *Nat. Chem. Biol.* **2014**, *10*, 774-779.
55. Joshi, A. K.; Witkowski, A.; Berman, H. A.; Zhang, L.; Smith, S. Effect of modification of the length and flexibility of the acyl carrier protein-thioesterase interdomain linker on functionality of the animal fatty acid synthase. *Biochemistry* **2005**, *44*, 4100-4107.
56. Seegmiller, A. C.; Dobrosotskaya, I.; Goldstein, J. L.; Ho, Y. K.; Brown, M. S.; Rawson, R. B. The SREBP Pathway in Drosophila: Regulation by Palmitate, Not Sterols. *Dev. Cell* **2002**, *2*, 229-238.
57. de Renobales, M.; Woodin, T. S.; Blomquist, G. J. Drosophila melanogaster fatty acid synthetase. *Insect Biochem.* **1986**, *16*, 887-894.
58. de Renobales, M.; Blomquist, G. J., Biosynthesis of medium chain fatty acids in Drosophila melanogaster. *Arch. Biochem. Biophys.* **1984**, *228*, 407-414.
59. Libertini, L. J.; Smith, S. Purification and properties of a thioesterase from lactating rat mammary gland which modifies the product specificity of fatty acid synthetase. *J. Biol. Chem.* **1978**, *253*, 1393-1401.
60. Ritchie, M. K.; Johnson, L. C.; Clodfelter, J. E.; Pemble, C. W.; Fulp, B. E.; Furdui, C. M.; Kridel, S. J.; Lowther, W. T. Crystal Structure and Substrate Specificity of Human Thioesterase 2: INSIGHTS INTO THE MOLECULAR BASIS FOR THE MODULATION OF FATTY ACID SYNTHASE. *J. Biol. Chem.* **2016**, *291*, 3520-3530.
61. Wicker-Thomas, C.; Garrido, D.; Bontonou, G.; Napal, L.; Mazuras, N.; Denis, B.; Rubin, T.; Parvy, J. P.; Montagne, J. Flexible origin of hydrocarbon/pheromone precursors in Drosophila melanogaster. *J. Lipid Res.* **2015**, *56*, 2094-2101.
62. Garrido, D.; Rubin, T.; Poidevin, M.; Maroni, B.; Le Rouzic, A.; Parvy, J. P.; Montagne, J. Fatty acid synthase cooperates with glyoxalase 1 to protect against sugar toxicity. *PLoS Genet.* **2015**, *11*, e1004995.
63. Herbst, D. A.; Jakob, R. P.; Zähringer, F.; Maier, T. Mycocerosic acid synthase exemplifies the architecture of reducing polyketide synthases. *Nature* **2016**, *531*, 533-537.
64. Brignole, E. J.; Smith, S.; Asturias, F. J. Conformational flexibility of metazoan fatty acid synthase enables catalysis. *Nat. Struct. Mol. Biol.* **2009**, *16*, 190-197.
65. Witkowski, A.; Joshi, A. K.; Rangan, V. S.; Falick, A. M.; Witkowska, H. E.; Smith, S. Dibromopropanone cross-linking of the phosphopantetheine and active-site cysteine thiols of the animal fatty acid synthase can occur both inter- and intrasubunit. Reevaluation of the side-by-side, antiparallel subunit model. *J. Biol. Chem.* **1999**, *274*, 11557-11563.

66. Rangan, V. S.; Joshi, A. K.; Smith, S. Mapping the functional topology of the animal fatty acid synthase by mutant complementation in vitro. *Biochemistry* **2001**, *40*, 10792-10799.
67. Hunkeler, M.; Stuttfeld, E.; Hagmann, A.; Imseng, S.; Maier, T. The dynamic organization of fungal acetyl-CoA carboxylase. *Nat. Commun.* **2016**, *7*, 11196.
68. Murphy, G. E.; Jensen, G. J. Electron Cryotomography of the E. coli Pyruvate and 2-Oxoglutarate Dehydrogenase Complexes. *Structure* **2005**, *13*, 1765-1773.
69. Radford, S. E.; Laue, E. D.; Perham, R. N.; Martin, S. R.; Appella, E. Conformational flexibility and folding of synthetic peptides representing an interdomain segment of polypeptide chain in the pyruvate dehydrogenase multienzyme complex of Escherichia coli. *J. Biol. Chem.* **1989**, *264*, 767-775.
70. Jeschke, G. DEER distance measurements on proteins. *Annu. Rev. Phys. Chem.* **2012**, *63*, 419-446.
71. Schmidt, T.; Walti, M. A.; Baber, J. L.; Hustedt, E. J.; Clore, G. M. Long Distance Measurements up to 160 Å in the GroEL Tetradecamer Using Q-Band DEER EPR Spectroscopy. *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 15905-15909.
72. Jeschke, G.; Polyhach, Y. Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1895-1910.
73. Hänsel, R.; Luh, L. M.; Corbeski, I.; Trantirek, L.; Dötsch, V. In-cell NMR and EPR spectroscopy of biomacromolecules. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 10300-10314.
74. Martorana, A.; Bellapadrona, G.; Feintuch, A.; Di Gregorio, E.; Aime, S.; Goldfarb, D. Probing protein conformation in cells by EPR distance measurements using Gd3+ spin labeling. *J. Am. Chem. Soc.* **2014**, *136*, 13458-13465.
75. Fielding, A. J.; Concilio, M. G.; Heaven, G.; Hollas, M. A. New developments in spin labels for pulsed dipolar EPR. *Molecules* **2014**, *19*, 16998-17025.
76. Roy, R.; Hohng, S.; Ha, T. A Practical Guide to Single Molecule FRET. *Nat. Methods* **2008**, *5*, 507-516.
77. Kim, J. Y.; Kim, C.; Lee, N. K. Real-time submillisecond single-molecule FRET dynamics of freely diffusing molecules with liposome tethering. *Nat. Commun.* **2015**, *6*, 6992.
78. Maglione, M.; Sigrist, S. J. Seeing the forest tree by tree: super-resolution light microscopy meets the neurosciences. *Nat. Neurosci.* **2013**, *16*, 790-797.
79. Westphal, V.; Rizzoli, S. O.; Lauterbach, M. A.; Kamin, D.; Jahn, R.; Hell, S. W. Video-rate far-field optical nanoscopy dissects synaptic vesicle movement. *Science* **2008**, *320*, 246-249.
80. Balzarotti, F.; Eilers, Y.; Gwosch, K. C.; Gynna, A. H.; Westphal, V.; Stefani, F. D.; Elf, J.; Hell, S. W. Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes. *Science* **2017**, *355*, 606-612.

81. Zhang, P. Correlative cryo-electron tomography and optical microscopy of cells. *Curr. Opin. Struct. Biol.* **2013**, *23*, 763-770.
82. Ando, T.; Uchihashi, T.; Fukuma, T. High-speed atomic force microscopy for nano-visualization of dynamic biomolecular processes. *Prog. Surf. Sci.* **2008**, *83*, 337-437.
83. Ando, T.; Uchihashi, T.; Kodera, N.; Yamamoto, D.; Miyagi, A.; Taniguchi, M.; Yamashita, H. High-speed AFM and nano-visualization of biomolecular processes. *Pflugers Arch.* **2008**, *456*, 211-225.
84. Eghiaian, F.; Rico, F.; Colom, A.; Casuso, I.; Scheuring, S. High-speed atomic force microscopy: imaging and force spectroscopy. *FEBS Lett.* **2014**, *588*, 3631-3638.
85. Ando, T.; Uchihashi, T.; Kodera, N. High-speed AFM and applications to biomolecular systems. *Annu. Rev. Biophys.* **2013**, *42*, 393-414.
86. Kodera, N.; Yamamoto, D.; Ishikawa, R.; Ando, T. Video imaging of walking myosin V by high-speed atomic force microscopy. *Nature* **2010**, *468*, 72-76.
87. Uchihashi, T.; Iino, R.; Ando, T.; Noji, H. High-speed atomic force microscopy reveals rotary catalysis of rotorless F(1)-ATPase. *Science* **2011**, *333*, 755-758.
88. Sakiyama, Y.; Mazur, A.; Kapinos, L. E.; Lim, R. Y. H. Spatiotemporal dynamics of the nuclear pore complex transport barrier resolved by high-speed atomic force microscopy. *Nat. Nanotechnol.* **2016**, *11*, 719-723.
89. Yilmaz, N.; Yamada, T.; Greimel, P.; Uchihashi, T.; Ando, T.; Kobayashi, T. Real-time visualization of assembling of a sphingomyelin-specific toxin on planar lipid membranes. *Biophys. J.* **2013**, *105*, 1397-1405.
90. Ando, T.; Uchihashi, T.; Scheuring, S. Filming biomolecular processes by high-speed atomic force microscopy. *Chem. Rev.* **2014**, *114*, 3120-3188.
91. Khosla, C.; Kapur, S.; Cane, D. E. Revisiting the modularity of modular polyketide synthases. *Curr. Opin. Chem. Biol.* **2009**, *13*, 135-143.
92. 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* **2004**, *12*, 417-428.
93. Pemble, C. W. 4th; Johnson, L. C.; Kridel, S. J.; Lowther, W. T. Crystal structure of the thioesterase domain of human fatty acid synthase inhibited by Orlistat. *Nat. Struct. Mol. Biol.* **2007**, *14*, 704-709.
94. Witkowski, A.; Joshi, A. K.; Smith, S. Characterization of the interthiol acyltransferase reaction catalyzed by the beta-ketoacyl synthase domain of the animal fatty acid synthase. *Biochemistry* **1997**, *36*, 16338-16344.
95. Martin, J. L.; McMillan, F. M. SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold. *Curr. Opin. Struct. Biol.* **2002**, *12*, 783-793.
96. Tong, L. Structure and function of biotin-dependent carboxylases. *Cell. Mol. Life Sci.* **2013**, *70*, 863-891.
97. Zheng, J.; Gay, D. C.; Demeler, B.; White, M. A.; Keatinge-Clay, A. T. Divergence of multimodular polyketide synthases revealed by a didomain structure. *Nat. Chem. Biol.* **2012**, *8*, 615-621.
98. Fu, J.; Bian, X.; Hu, S.; Wang, H.; Huang, F.; Seibert, P. M.; Plaza, A.; Xia, L.; Muller, R.; Stewart, A. F.; Zhang, Y. Full-length RecE enhances linear-

linear homologous recombination and facilitates direct cloning for bioprospecting. *Nat. Biotechnol.* **2012**, *30*, 440-446.

99. Fitzgerald, D. J.; Berger, P.; Schaffitzel, C.; Yamada, K.; Richmond, T. J.; Berger, I. Protein complex expression by using multigene baculoviral vectors. *Nat. Methods* **2006**, *3*, 1021-1032.

100. Wasilko, D. J.; Lee, S. E.; Stutzman-Engwall, K. J.; Reitz, B. A.; Emmons, T. L.; Mathis, K. J.; Bienkowski, M. J.; Tomasselli, A. G.; Fischer, H. D. The titerless infected-cells preservation and scale-up (TIPS) method for large-scale production of NO-sensitive human soluble guanylate cyclase (sGC) from insect cells infected with recombinant baculovirus. *Protein Expr. Purif.* **2009**, *65*, 122-132.

101. Smith, S.; Abraham, S. Fatty acid synthase from lactating rat mammary gland. *Methods Enzymol.* **1975**, *35*, 65-74.

102. Kabsch, W. XDS. *Acta Cryst. D* **2010**, *66*, 125-132.

103. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Cryst. D* **2010**, *66*, 133-144.

104. Murshudov, G. N.; Skubak, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Cryst. D* **2011**, *67*, 355-367.

105. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 suite and current developments. *Acta Cryst. D* **2011**, *67*, 235-242.

106. Zhang, K. Y.; Cowtan, K.; Main, P. Combining constraints for electron-density modification. *Methods Enzymol.* **1997**, *277*, 53-64.

107. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Cryst. D* **2006**, *62*, 1002-1011.

108. Cowtan, K. Fitting molecular fragments into electron density. *Acta Cryst. D* **2008**, *64*, 83-89.

109. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Cryst. D* **2010**, *66*, 486-501.

110. Afonine, P. V.; Moriarty, N. W.; Mustyakimov, M.; Sobolev, O. V.; Terwilliger, T. C.; Turk, D.; Urzhumtsev, A.; Adams, P. D. FEM: feature-enhanced map. *Acta Cryst. D* **2015**, *71*, 646-666.

111. Smart, O. S.; Womack, T. O.; Flensburg, C.; Keller, P.; Paciorek, W.; Sharff, A.; Vonrhein, C.; Bricogne, G. Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Cryst. D* **2012**, *68*, 368-380.

112. Afonine, P. V.; Grosse-Kunstleve, R. W.; Adams, P. D.; Urzhumtsev, A. Bulk-solvent and overall scaling revisited: faster calculations, improved results. *Acta Cryst. D* **2013**, *69*, 625-634.

113. Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. Towards automated crystallographic structure refinement with phenix.refine. *Acta Cryst. D* **2012**, *68*, 352-367.

114. Afonine, P. V.; Grosse-Kunstleve, R. W.; Urzhumtsev, A.; Adams, P. D. Automatic multiple-zone rigid-body refinement with a large convergence radius. *J. Appl. Crystallogr.* **2009**, *42*, 607-615.
115. Headd, J. J.; Echols, N.; Afonine, P. V.; Grosse-Kunstleve, R. W.; Chen, V. B.; Moriarty, N. W.; Richardson, D. C.; Richardson, J. S.; Adams, P. D. Use of knowledge-based restraints in phenix.refine to improve macromolecular refinement at low resolution. *Acta Cryst. D* **2012**, *68*, 381-390.
116. Krissinel, E.; Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Cryst. D* **2004**, *60*, 2256-2268.
117. Krissinel, E.; Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **2007**, *372*, 774-797.
118. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; Thierer, T.; Ashton, B.; Meintjes, P.; Drummond, A. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647-1649.
119. Patel, M. S.; Nemeria, N. S.; Furey, W.; Jordan, F. The pyruvate dehydrogenase complexes: structure-based function and regulation. *J. Biol. Chem.* **2014**, *289*, 16615-16623.
120. Barends, T. R.; Dunn, M. F.; Schlichting, I. Tryptophan synthase, an allosteric molecular factory. *Curr. Opin. Chem. Biol.* **2008**, *12*, 593-600.
121. Wakil, S. J.; Stoops, J. K.; Joshi, V. C. Fatty acid synthesis and its regulation. *Annu. Rev. Biochem.* **1983**, *52*, 537-579.
122. Cronan, J. E., Jr.; Waldrop, G. L. Multi-subunit acetyl-CoA carboxylases. *Prog. Lipid Res.* **2002**, *41*, 407-435.
123. Yan, J.; Yan, Y.; Madzak, C.; Han, B. Harnessing biodiesel-producing microbes: from genetic engineering of lipase to metabolic engineering of fatty acid biosynthetic pathway. *Crit. Rev. Biotechnol.* **2017**, *37*, 26-36.
124. Gajewski, J.; Buelens, F.; Serdjukow, S.; Janssen, M.; Cortina, N.; Grubmüller, H.; Grininger, M. Engineering fatty acid synthases for directed polyketide production. *Nat. Chem. Biol.* **2017**, *13*, 363-365.
125. Zhu, Z.; Zhou, Y. J.; Krivoruchko, A.; Grininger, M.; Zhao, Z. K.; Nielsen, J. Expanding the product portfolio of fungal type I fatty acid synthases. *Nat. Chem. Biol.* **2017**, *13*, 360-362.
126. Jeon, E.; Lee, S.; Lee, S.; Han, S. O.; Yoon, Y. J.; Lee, J. Improved production of long-chain fatty acid in Escherichia coli by an engineering elongation cycle during fatty acid synthesis (FAS) through genetic manipulation. *J. Microbiol. Biotechnol.* **2012**, *22*, 990-999.
127. Crosby, J.; Crump, M. P. The structural role of the carrier protein--active controller or passive carrier. *Nat. Prod. Rep.* **2012**, *29*, 1111-1137.
128. Yamamoto, D.; Ando, T. Chaperonin GroEL-GroES Functions as both Alternating and Non-Alternating Engines. *J. Mol. Biol.* **2016**, *428*, 3090-3101.
129. Mattick, J. S.; Nickless, J.; Mizugaki, M.; Yang, C. Y.; Uchiyama, S.; Wakil, S. J. The architecture of the animal fatty acid synthetase. II. Separation of the core and thioesterase functions and determination of the N-C orientation of the subunit. *J. Biol. Chem.* **1983**, *258*, 15300-15304.

130. Joshi, A. K.; Witkowski, A.; Smith, S. The malonyl/acetyltransferase and beta-ketoacyl synthase domains of the animal fatty acid synthase can cooperate with the acyl carrier protein domain of either subunit. *Biochemistry* **1998**, *37*, 2515-2523.
131. Uchihashi, T.; Kodera, N.; Ando, T. Guide to video recording of structure dynamics and dynamic processes of proteins by high-speed atomic force microscopy. *Nat. Protoc.* **2012**, *7*, 1193-1206.
132. Shibata, M.; Yamashita, H.; Uchihashi, T.; Kandori, H.; Ando, T. High-speed atomic force microscopy shows dynamic molecular processes in photoactivated bacteriorhodopsin. *Nat. Nanotechnol.* **2010**, *5*, 208-212.
133. de la Rosa-Trevin, J. M.; Oton, J.; Marabini, R.; Zaldivar, A.; Vargas, J.; Carazo, J. M.; Sorzano, C. O. Xmipp 3.0: an improved software suite for image processing in electron microscopy. *J. Struct. Biol.* **2013**, *184*, 321-328.
134. Fu, J.; Bian, X.; Hu, S.; Wang, H.; Huang, F.; Seibert, P. M.; Plaza, A.; Xia, L.; Müller, R.; Stewart, A. F.; Zhang, Y. Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. *Nat. Biotechnol.* **2012**, *30*, 440-446.
135. Yamamoto, D.; Uchihashi, T.; Kodera, N.; Yamashita, H.; Nishikori, S.; Ogura, T.; Shibata, M.; Ando, T. High-speed atomic force microscopy techniques for observing dynamic biomolecular processes. *Methods Enzymol.* **2010**, *475*, 541-564.
136. Nussio, M. R.; Voelcker, N. H.; Miners, J. O.; Lewis, B. C.; Sykes, M. J.; Shapter, J. G. AFM study of the interaction of cytochrome P450 2C9 with phospholipid bilayers. *Chem. Phys. Lipids* **2010**, *163*, 182-189.
137. Ando, T.; Kodera, N.; Takai, E.; Maruyama, D.; Saito, K.; Toda, A. A high-speed atomic force microscope for studying biological macromolecules. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 12468-12472.
138. Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W., NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **2012**, *9*, 671-675.
139. Tang, G.; Peng, L.; Baldwin, P. R.; Mann, D. S.; Jiang, W.; Rees, I.; Ludtke, S. J. EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **2007**, *157*, 38-46.
140. de la Rosa-Trevin, J. M.; Quintana, A.; Del Cano, L.; Zaldivar, A.; Foche, I.; Gutierrez, J.; Gomez-Blanco, J.; Burguet-Castell, J.; Cuenca-Alba, J.; Abrishami, V.; Vargas, J.; Oton, J.; Sharov, G.; Vilas, J. L.; Navas, J.; Conesa, P.; Kazemi, M.; Marabini, R.; Sorzano, C. O.; Carazo, J. M. Scipion: A software framework toward integration, reproducibility and validation in 3D electron microscopy. *J. Struct. Biol.* **2016**, *195*, 93-99.
141. Sorzano, C. O.; Bilbao-Castro, J. R.; Shkolnisky, Y.; Alcorlo, M.; Melero, R.; Caffarena-Fernandez, G.; Li, M.; Xu, G.; Marabini, R.; Carazo, J. M. A clustering approach to multireference alignment of single-particle projections in electron microscopy. *J. Struct. Biol.* **2010**, *171*, 197-206.
142. Cardon, J. W.; Hammes, G. G. Investigation of reduced nicotinamide adenine dinucleotide phosphate and acyl-binding sites on avian fatty acid synthase. *Biochemistry* **1982**, *21*, 2863-2870.
143. Anderson, V. E.; Hammes, G. G. Stereochemistry of the reactions catalyzed by chicken liver fatty acid synthase. *Biochemistry* **1984**, *23*, 2088-2094.

144. Jayakumar, A.; Huang, W. Y.; Raetz, B.; Chirala, S. S.; Wakil, S. J. Cloning and expression of the multifunctional human fatty acid synthase and its subdomains in Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 14509-14514.
145. Joshi, A. K.; Rangan, V. S.; Smith, S. Differential affinity labeling of the two subunits of the homodimeric animal fatty acid synthase allows isolation of heterodimers consisting of subunits that have been independently modified. *J. Biol. Chem.* **1998**, *273*, 4937-4943.
146. Rangan, V. S.; Joshi, A. K.; Smith, S. Fatty acid synthase dimers containing catalytically active beta-ketoacyl synthase or malonyl/acetyltransferase domains in only one subunit can support fatty acid synthesis at the acyl carrier protein domains of both subunits. *J. Biol. Chem.* **1998**, *273*, 34949-34953.
147. Carlisle-Moore, L.; Gordon, C. R.; Machutta, C. A.; Miller, W. T.; Tonge, P. J. Substrate recognition by the human fatty-acid synthase. *J. Biol. Chem.* **2005**, *280*, 42612-42618.
148. Akey, D. L.; Gehret, J. J.; Khare, D.; Smith, J. L. Insights from the Sea: Structural Biology of Marine Polyketide Synthases. *Nat. Prod. Rep.* **2012**, *29*, 1038-1049.
149. Ruan, Y.; Miyagi, A.; Wang, X.; Chami, M.; Boudker, O.; Scheuring, S. Direct visualization of glutamate transporter elevator mechanism by high-speed AFM. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 1584-1588.
150. Igarashi, K.; Uchihashi, T.; Koivula, A.; Wada, M.; Kimura, S.; Penttila, M.; Ando, T.; Samejima, M. Visualization of cellobiohydrolase I from Trichoderma reesei moving on crystalline cellulose using high-speed atomic force microscopy. *Methods Enzymol.* **2012**, *510*, 169-182.
151. Frueh, D. P.; Goodrich, A.; Mishra, S.; Nichols, S. NMR methods for structural studies of large monomeric and multimeric proteins. *Curr. Opin. Struct. Biol.* **2013**, *23*, 734-739.
152. Opitz, C.; Isogai, S.; Grzesiek, S. An economic approach to efficient isotope labeling in insect cells using homemade 15N-, 13C- and 2H-labeled yeast extracts. *J. Biomol. NMR* **2015**, *62*, 373-385.
153. Weissman, K. J. The structural biology of biosynthetic megaenzymes. *Nat. Chem. Biol.* **2015**, *11*, 660-670.
154. Tarry, M. J.; Haque, A. S.; Bui, K. H.; Schmeing, T. M. X-Ray Crystallography and Electron Microscopy of Cross- and Multi-Module Nonribosomal Peptide Synthetase Proteins Reveal a Flexible Architecture. *Structure* **2017**, *25*, 783-793.e4.
155. Hammer, D. A.; Kamat, N. P. Towards an artificial cell. *FEBS Lett.* **2012**, *586*, 2882-2890.
156. Wong, F. T.; Khosla, C. Combinatorial biosynthesis of polyketides--a perspective. *Curr. Opin. Chem. Biol.* **2012**, *16*, 117-123.
157. Floss, H. G. Combinatorial Biosynthesis – Potential and Problems. *J. Biotechnol.* **2006**, *124*, 242-257.
158. Buddingh', B. C.; van Hest, J. C. M. Artificial Cells: Synthetic Compartments with Life-like Functionality and Adaptivity. *Acc. Chem. Res.* **2017**, *50*, 769-777.

159. Elani, Y.; Law, R. V.; Ces, O. Vesicle-based artificial cells as chemical microreactors with spatially segregated reaction pathways. *Nat. Commun.* **2014**, *5*, 5305.