BRIEF COMMUNICATION
PUBLISHED ONLINE: 20 FEBRUARY 2017 | DOI: 10.1038/NCHEMBIO.2314

Engineering fatty acid synthases for directed polyketide production

Jan Gajewski[1,4], Floris Buelens[2,4], Sascha Serdjukow[3], Melanie Janβen[1], Niña Cortina[1], Helmut Grubmüller[2]* & Martin Grininger[1,3]*

In this study, we engineered fatty acid synthases (FAS) for the biosynthesis of short-chain fatty acids and polyketides, guided by a combined *in vitro* and *in silico* approach. Along with exploring the synthetic capability of FAS, we aim to build a foundation for efficient protein engineering, with the specific goal of harnessing evolutionarily related megadalton-scale polyketide synthases (PKS) for the tailored production of bioactive natural compounds.

Fatty acid synthases (FAS) synthesize fatty acids (FA) by repeatedly elongating and modifying a growing acyl chain until a specific length is attained. Exploiting the versatile synthetic capability of FAS, and taking advantage of the biochemical[1] and structural[2] characterization of FAS as well as access to FAS by recombinant methods[3], we aimed to engineer the 1.9-MDa *Corynebacterium ammoniagenes* FAS to produce the polyketide 6-heptyl-4-hydroxypyran-2-one (6-HHP). We designed a synthetic route consisting of two engineered FAS modules, the first (module 1) optimized to produce octanoyl-CoA (C8-CoA), and the second (module 2) to nonreductively elongate this intermediate, yielding 6-HHP (**Fig. 1**)[4]. While short-chain acyl-CoAs produced by module 1 are valuable precursors for short FA and short alkanes for biofuel production[5,6], the final lactone 6-HHP or its derivatives are interesting platform chemicals[7]. We also developed FAS into an experimental and theoretical testbed system, culminating in an *in silico* model of the FAS catalytic network.

In engineering module 1, we assumed that the kinetics of the ketoacyl synthase (KS), acetyl transferase (AT) and malonyl–palmitoyl transferase (MPT) domains are the main determinants of FA chain length. It has been reported that the acyl-CoA output spectrum of FAS is sensitive to relative concentrations of the priming acetyl-CoA and the elongating malonyl-CoA[8]. Accordingly, we posited that we could influence chain length by modulating substrate binding affinities of AT and MPT. Furthermore, the relative rates of the KS and MPT reactions determine whether the growing acyl carrier protein (ACP)-bound acyl chain is loaded in the KS, initiating another cycle of elongation, or loaded in the MPT, leading to product export[8]. As the MPT domain has been reported to be tolerant in binding acyl chains of several lengths[9], we initially focused on achieving chain-length control by engineering KS only. From the X-ray structural data of *Saccharomyces cerevisiae* FAS[10–12], M1251, positioned centrally in the KS channel, has been proposed to act as a 'gatekeeper' residue to the KS binding channel (**Supplementary Results**, **Supplementary Fig. 1**)[12,13]. In the *Corynebacterium ammoniagenes* FAS, the equivalent residue is M2600 (refs. 14,15); thus, we constructed mutants FAS1^{G2599S} and FAS1$^{G2599S\text{-}M2600W}$. These variants indeed seemed to

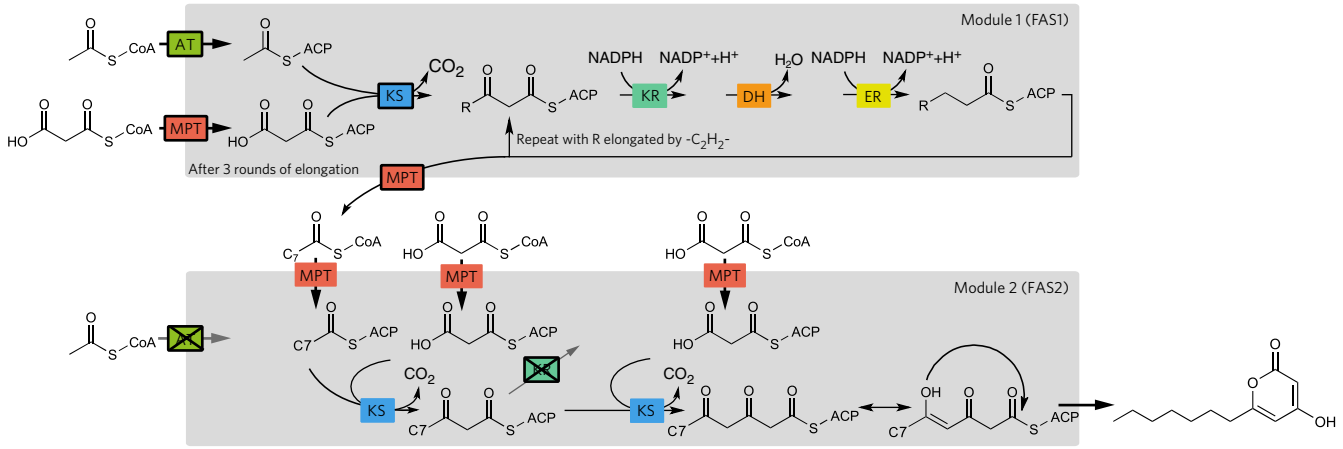

Figure 1 | Engineered reaction pathway for 6-HHP. Two engineered fatty acid synthases (FASs), termed module 1 and module 2, work in sequence to synthesize C8-CoA (FAS1) and the final 6-HHP product (FAS2). FAS1 is modulated in FA chain length regulation, shifting the FA product spectrum from the native products C16- or C18-CoA toward C8-CoA. FAS2 selectively accepts and nonreductively elongates C8-CoA to the triketide, which dissociates from ACP by spontaneous lactonization. Engineered domains are highlighted by black frames; crossed out indicates a functional knockout. AT, acetyl transferase; MPT, malonyl–palmitoyl transferase; KS, ketoacyl synthase, KR, ketoacyl reductase; ER, enoyl reductase; DH, dehydratase.

[1]Institute of Organic Chemistry and Chemical Biology, Buchmann Institute for Molecular Life Sciences, Cluster of Excellence "Macromolecular Complexes," Goethe University Frankfurt, Frankfurt am Main, Germany. [2]Max Planck Institute for Biophysical Chemistry, Department of Theoretical and Computational Biophysics, Göttingen, Germany. [3]Max Planck Institute of Biochemistry, Project Group Biological Chemistry, Martinsried, Germany. [4]These authors contributed equally to this work. *e-mail: grininger@chemie.uni-frankfurt.de or hgrubmu@gwdg.de

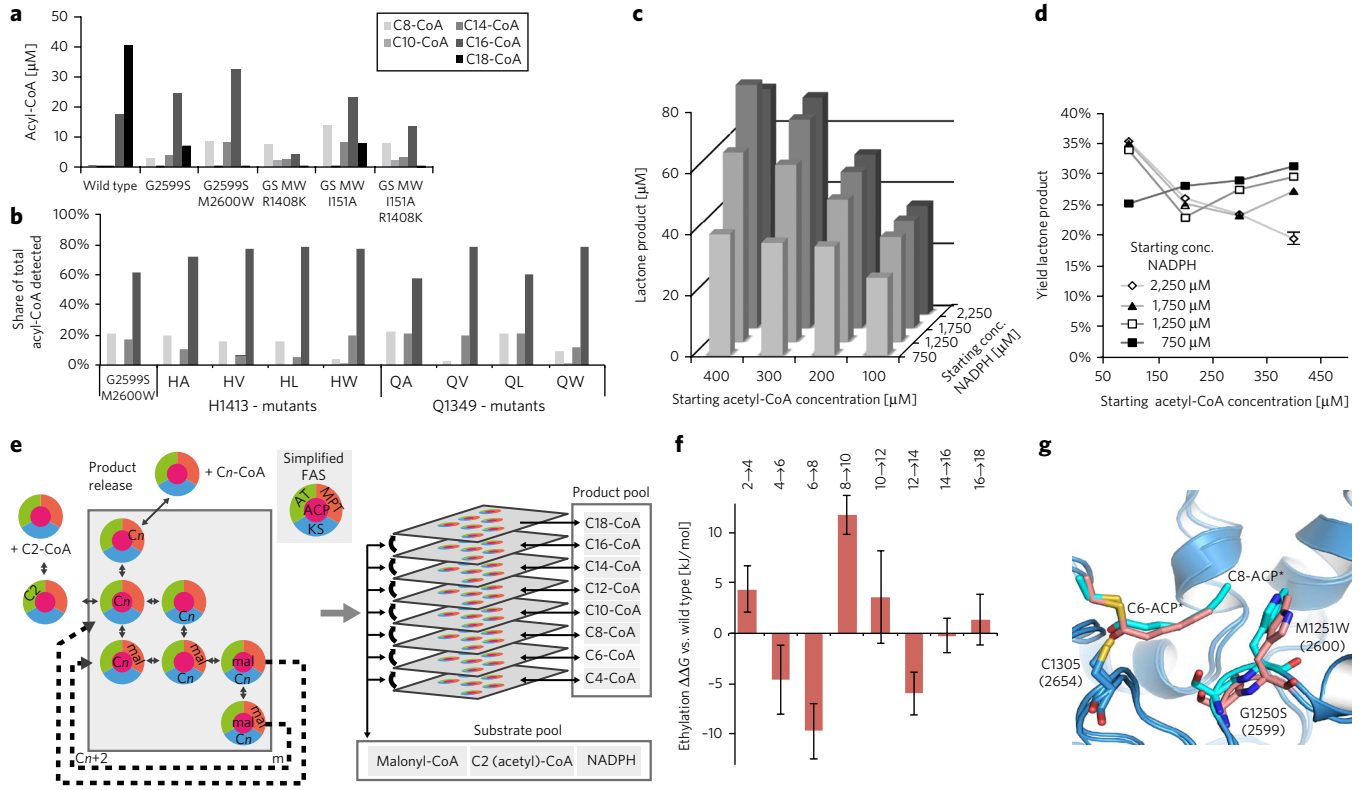

Figure 2 | FAS mediated synthesis and computational modeling. (**a**) Product distributions of engineered FAS1 mutated in KS. Data refer to means of technical replicates (n = 1; three measurements with s.d. below ± 0.76 μM for this sample set). For more information on statistics, see Online Methods. (**b**) Product distributions of FAS1 variants mutated in MPT (n = 1; four measurements with s.d. below ± 2.8% (share of total acyl-CoA detected)), except for Q1349V and Q1349L, which were determined in a single measurement owing to limited protein supply. (**c**) 6-HHP as produced in the coupled reaction by FAS1$^{G2599S\text{-}M2600W\text{-}R1408K}$ → FAS2$^{S126A\text{-}Y2227F}$ at different substrate concentrations (conc.) (n = 1). For more information on statistics, see Online Methods. (**d**) Percentage yield in reference to the limiting substrate, calculated from **c**. (**e**) Reaction network of a simplified FAS. Underlying this representation is an array of kinetic rate constants, including the elongation reaction (dashed arrows, yielding C(n+2)) (left). Abstraction of chain elongations from C4 to C18, with each layer in the stack representing the simplified model for a different chain length Cn (right). ACP, acyl carrier protein. (**f**) Binding free energy change for acyl chains C4–C16 in the mutant FAS1$^{G1250S\text{-}M1251W}$ with respect to wild-type FAS, as calculated from molecular dynamics simulations. Negative values correspond to more favorable surroundings for the introduction of respective ethyl groups in the mutant protein compared to wild type. (**g**) Snapshots from two molecular dynamics simulations, representing respectively C6- (pink) and C8-ACP (cyan) binding to the modified KS domain. Acyl-ACP was modeled in truncated form, indicated by (*).

counteract binding of acyl intermediates of C8 or longer, and exhibited increased production of shorter FA (**Fig. 2a** and **Supplementary Fig. 2**). For quality and repeatability of protein preparations, see **Supplementary Figure 3** and Online Methods.

As a second step in engineering module 1, we focused on transferase domains AT and MPT. We first aimed to engineer the MPT domain (**Supplementary Fig. 4**) to have higher C8-CoA affinities, with the intention of facilitating short acyl-CoA release. None of the constructs showed desired effects *in vitro* (**Fig. 2b**). As the MPT domain mediates the dual functions of loading malonyl and releasing the acyl product through a single catalytic serine residue, a decrease in malonyl affinity can be expected to decrease the rate of malonyl loading and to reduce competition for acyl product release (**Fig. 1**)[16,17]. Indeed, KS- and MPT-modified FAS1$^{G2599S\text{-}M2600W\text{-}R1408K}$ yielded C8-CoA as 47% of the total acyl-CoA products detected (**Fig. 2a**). In the AT domain, the I151A mutation (in addition to the KS-domain mutations; FAS1$^{G2599S\text{-}M2600W\text{-}I151A}$) led to higher absolute and relative yields of C8-CoA (**Fig. 2a**). Comparable mutations have been reported to increase acetyl throughput, matching our goal of shifting the balance toward shorter products[16,17]. Moreover, mutations have been postulated to form a novel binding channel that can accommodate longer acyl chains of up to C10, potentially introducing a product-exporting acyltransferase activity for short-length acyl chains (**Supplementary Fig. 4**). While these transferase mutations were beneficial in isolation, their effects on relative C8-CoA yield were not additive (**Fig. 2a**).

For module 2, we produced a single construct incorporating catalytic knockout mutations in both the KR and AT sites. Mutation Y2227F eliminates KR activity[4], thereby blocking the canonical reaction pathway that processes β-keto intermediates to saturated acyl chains (**Fig. 1**). We additionally blocked the canonical acetyl-CoA priming capacity with mutation S126A in the AT domain, enforcing sequential processing of the module 1 acyl-CoA products by module 2. Module 2 achieved ~5% of the specific activity of the wild-type FAS (**Supplementary Fig. 3**). To optimize yields, we first identified the module 1 construct best suited for collaboration with module 2, and then optimized reactant and cofactor concentrations, finally leading to 6-HHP yields of up to 35% (**Fig. 2c,d**; **Supplementary Figs. 5–7**).

In light of the challenge posed by the modulation of the interplay between several catalytic functions in a finely balanced reaction network, we sought to establish a quantitative description of FA synthesis built on kinetic and atomistic modeling, thus qualitatively and quantitatively linking structural and functional properties of FAS with their product outputs. The quantitative model represents a simplified canonical catalytic network of 51 distinct reactions (**Supplementary Table 1**), which we evaluated by stochastic simulation (**Fig. 2e**)[18].

We restricted the parameter space of the rate constants that describe the reaction network by enforcing agreement with known product distributions for *S. cerevisiae* FAS (**Supplementary Note 1** and Online Methods), setting aside experimental data described below for validating the predictive power of the model. From the tens of millions of evaluated parameter sets, an ensemble of the 1,000 sets best matching the input data (hereafter referred to collectively as 'the model') was retained as a representative sample of the remaining parameter space and used for the subsequent analysis and predictions. To represent the effects of active site mutations on the model, we calculated free energy changes for enzyme–substrate complexes from atomistic simulation (**Supplementary Note 2**) and sampled alternate rate constants for the affected reactions, selected in accordance with the calculated $\Delta\Delta G$ values. Applying the computational model to results from the transferase mutations in *C. ammoniagenes* FAS enabled our model to correctly reproduce experimental data that were not used in the model construction (for more information on model cross-validation, see **Supplementary Note 2** and **Supplementary Fig. 8**).

Unexpectedly, for the KS domain, atomistic simulation-based free energy calculations contradicted the 'gatekeeper rationale', which predicted weakened C8 binding. Instead, the double KS mutation was calculated to enhance the binding of C8 (**Fig. 2f** and **Supplementary Fig. 9**). A structural model of the *S. cerevisiae* $KS^{G1250S\text{-}M1251W}$–C8 complex from atomistic molecular dynamics simulation (**Fig. 2g**) supports this result, showing favorable close nonpolar van der Waals contacts of the acyl chain tail with the indole ring of W1251, with no apparent displacement of its side chain. No tested combination of parameters led to the experimentally observed increased C8-CoA release.

To reconcile the apparent contradiction between the computational model and experimental data, we first considered unproductive KS-mediated decarboxylation, which occurs naturally in FAS–PKS systems[19]. Both *in silico* and *in vitro* studies ruled out this interpretation (**Supplementary Fig. 10** and **Supplementary Note 2**). However, in our initial representation of the calculated enhanced binding affinity of the C8–KS complex, we assumed that the calculated free energy change, which was considerably lower for the mutated C8–KS state, would not yield a reduction in the forward rate of the reaction C8–ACP + KS → ACP + C8–KS. We relaxed this restriction to allow reductions in this reaction rate up to a factor of 10 (while still fixing the ratio of forward and reverse rate constants in accordance with the calculated free energy difference). In doing so, we identified parameter sets among our model ensemble for which a reduction in the forward rate of as little as 30% yielded increased (by more than 5%) C8-CoA output (**Supplementary Fig. 11** and **Supplementary Note 2**). According to this interpretation, chain-length control imposed by KS mutations is achieved not by a 'gate' in the KS substrate tunnel, which sterically restricts loading of longer FA, but rather by a kinetic barrier that steers C8-CoA away from KS-mediated elongation and toward release. This conclusion underscores the advantage of combining *in vitro* classical enzymology with *in silico* modeling, highlighting the potential to achieve insights into complex biosynthetic systems going beyond the phenomenological interpretation of engineering successes.

On the basis of these descriptions of 6-HHP and C8-CoA synthesis, mutations may now be translated into cellular systems for the fermentative production of these compounds. FAS-mediated synthesis will, however, be limited by the confined iterative reaction mode, and high versatility for the directed synthesis of chemical compounds will require engineering of vectorial natural compound biosynthesis, such as that performed by modular PKS[20,21] (**Supplementary Note 3**).

Besides bioengineering of FAS, our study was also largely encouraged by exploration of FAS as a model for evaluating an integrative approach for engineering FAS–PKS reaction networks. PKS are evolutionarily and mechanistically related to FAS and are responsible for important natural compounds[22,23]. Achievements in engineering PKS notwithstanding[24,25], progress in developing PKS into versatile tools for compound synthesis has remained slow[21]. A combined *in vitro* and *in silico* approach may serve as a blueprint for making PKS more amenable to pathway design. Likewise, the kinetic model for describing iterative FAS can be adapted to represent even complex modular PKS with each layer of the model describing the chemistry of a separate module. Development of such a model will depend on the collection of quantitative data from in-depth enzymological characterization of PKS.

Received 11 February 2016; accepted 10 January 2017;
published online 20 February 2017

Methods

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

References

1. Lynen, F. *Methods Enzymol.* **14**, 17–33 (1969).
2. Maier, T., Leibundgut, M., Boehringer, D. & Ban, N. *Q. Rev. Biophys.* **43**, 373–422 (2010).
3. Enderle, M., McCarthy, A., Paithankar, K.S. & Grininger, M. *Acta Crystallogr. F Struct. Biol. Commun.* **71**, 1401–1407 (2015).
4. Zha, W., Shao, Z., Frost, J.W. & Zhao, H. *J. Am. Chem. Soc.* **126**, 4534–4535 (2004).
5. Choi, Y.J. & Lee, S.Y. *Nature* **502**, 571–574 (2013).
6. Peralta-Yahya, P.P., Zhang, F., del Cardayre, S.B. & Keasling, J.D. *Nature* **488**, 320–328 (2012).
7. Chia, M., Schwartz, T.J., Shanks, B.H. & Dumesic, J.A. *Green Chem.* **14**, 1850–1853 (2012).
8. Sumper, M., Oesterhelt, D., Riepertinger, C. & Lynen, F. *Eur. J. Biochem.* **10**, 377–387 (1969).
9. Pirson, W., Schuhmann, L. & Lynen, F. *Eur. J. Biochem.* **36**, 16–24 (1973).
10. Leibundgut, M., Jenni, S., Frick, C. & Ban, N. *Science* **316**, 288–290 (2007).
11. Lomakin, I.B., Xiong, Y. & Steitz, T.A. *Cell* **129**, 319–332 (2007).
12. Johansson, P. *et al. Proc. Natl. Acad. Sci. USA.* **105**, 12803–12808 (2008).
13. Morisaki, N. *et al. Eur. J. Biochem.* **211**, 111–115 (1993).
14. Aritomi, K. *et al. Biosci. Biotechnol. Biochem.* **68**, 206–214 (2004).
15. Christensen, C.E., Kragelund, B.B., von Wettstein-Knowles, P. & Henriksen, A. *Protein Sci.* **16**, 261–272 (2007).
16. Rangan, V.S. & Smith, S. *J. Biol. Chem.* **272**, 11975–11978 (1997).
17. Bunkoczi, G. *et al. Chem. Biol.* **16**, 667–675 (2009).
18. Gillespie, D.T. *J. Phys. Chem.* **81**, 2340–2361 (1977).
19. Kresze, G.-B., Steber, L., Oesterhelt, D. & Lynen, F. *Eur. J. Biochem.* **79**, 191–199 (1977).
20. Zhu, Z. *et al. Nat. Chem. Biol.* **13** http://dx.doi.org/10.1038/nchembio.2301 (2017).
21. Robbins, T., Liu, Y.-C., Cane, D.E. & Khosla, C. *Curr. Opin. Struct. Biol.* **41**, 10–18 (2016).
22. Hertweck, C. *Angew. Chem. Int. Ed. Engl.* **48**, 4688–4716 (2009).
23. Carter, G.T. *Nat. Prod. Rep.* **28**, 1783–1789 (2011).
24. Cane, D.E., Walsh, C.T. & Khosla, C. *Science* **282**, 63–68 (1998).
25. Menzella, H.G. *et al. Nat. Biotechnol.* **23**, 1171–1176 (2005).

Acknowledgments

We thank A. Matijczak and M. Enderle for starting the project with us by establishing purification of *Corynebacterium ammoniagenes* FAS and cloning first constructs. M.G. and H.G. thank D. Oesterhelt for continuous support over many years and for initiating the collaboration between H.G. and M.G. This work was supported by a Lichtenberg Grant of the Volkswagen Foundation to M.G. (grant number 85 701), the German Federal Ministry of Education and Research to F.B. (grant number 0315450I), the DFG excellence program CNMBP to F.B. (DFG-EXC 171), and the Max Planck Society (H.G., F.B.).

Author contributions

J.G., S.S., M.J. and M.G. performed biochemical studies; J.G. and M.G. analyzed biochemical data; N.C. supported MS-based product analysis; F.B. performed kinetic and atomistic modeling; J.G., F.B., H.G. and M.G. designed research; J.G., F.B., H.G. and M.G. wrote the manuscript.

Competing financial interests

The authors declare competing financial interests: details accompany the online version of the paper.

Additional information

Any supplementary information, chemical compound information and source data are available in the online version of the paper. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html. Correspondence and requests for materials should be addressed to H.G. or M.G.

ONLINE METHODS

Cloning. For the initial FAS construct, *Corynebacterium ammoniagenes* genomic DNA (ordered from DSMZ; DSM 20306) was used. The construct was cloned into a pET-22b(+)vector (Novagen, Merck Millipore) with an N-terminal strep twin tag with a short linker (MSAWSHPQFEKGGGSGGGSGGSAWSHPQFEKGAGS) or the shorter version, the strep II tag (MSAWSHPQFEKGAGS). For primer information, see below. For preparation, the pET-22b(+)vector was digested with BamHI and XhoI, and the insert was cloned using the In-Fusion HD cloning kit (Clontech).

Point mutations were introduced using site-directed mutagenesis with forward and reverse primers that carried the mutation (see below). After extraction from a preparative agarose gel (0.8%), the linear fragment was ligated using the In-Fusion HD cloning kit (Clontech). The vectors were transformed into competent cells (Stellar cells, Clontech) and amplified. All constructs were Sanger sequenced. For co-transformation, the *C. ammoniagenes* acyl carrier protein synthase (AcpS) was amplified (for primers, see below) and cloned into a pETcoco vector (Novagen, Merck Millipore).

Primers module 1. Primers for the amplification of the *C. ammoniagenes* FAS for In-Fusion cloning into pET-22b(+)vector (15 bp overlap with vector ends at the 5′ end, underlined). The amplified insert was included approximately 270 bp after the actual stop codon due to an initial wrong annotation in of the gene in databases.

fwd: 5′ AAAAGGCGCCGGATCCACTATTGGCATCTCTAACCACCGC CTGG 3′

rev: 5′ GGTGATGATGCTCGAGCTGGTGGCTTGCCGTAGATCGCTT GC 3′

Primers for the introduction of point mutations were designed as complimentary pairs with the mutation at a central position (PCR was then performed covering the whole vector). As an alternative, the overlap of the primers was approximately only 15 bp. The underlined part indicates the overlap of the primers, and the mutation sites are shown in bold.

Primers covering the mutation in the AT domain (I151A)

fwd: 5′ CAGCTC**GCC**GGCGTCGCTATTTCTAAG 3′

rev: 5′ GACGCC**GGC**GAGCTGCGCCAGCGCAATAAC 3′

Primers covering the mutation in the MPT domain (R1408K)

fwd: 5′ GAAATCGTCTACGCC**AAG**GGTTTGACCATGCAC 3′

rev: 5′ GGCGTAGACGATTTCTACAACGGCTTCC 3′

Primers covering the mutation in the KS domain (G2599S)

fwd: 5′ GCTCGACCCAGGGCACG**AGT**ATGGGCGGCATGCAGTCG 3′

rev: 5′ CGACTGCATGCCGCCCAT**ACT**CGTGCCCTGGGTCGAGC 3′

Primers covering the double mutation in the KS domain (G2599S-M2600W)

fwd: 5′ CCCAGGGCACG**AGTTGG**GGCGGCATGCAGTCGATGCGC 3′

rev: 5′ GCGCATCGACTGCATGCCGCC**CCAACT**CGTGCCCTGGG 3′

Primers covering the knockout in the AT domain (S125A)

fwd: 5′ GTCGCACACATTGGCCAT**GCC**CAAGGCGCGCTTGCTAC 3′

rev: 5′ GTAGCAAGCGCGCCTTG**GGC**ATGGCCAATGTGTGCGAC 3′

Primers covering the knockout in the KR domain (Y2227F)

fwd: 5′ CGGCGGCGACGGTGCC**TTT**GGTGAGTCCAAGGCTGCC 3′

rev: 5′ GGCAGCCTTGGACTCACC**AAA**GGCACCGTCGCCGCCG 3′

Primers *C. ammoniagenes* AcpS. The overlap to the vector is underlined, and the parts binding the genomic DNA for the insert amplification are shown in bold.

fwd: 5′ AGAAGGAGATATAAGC**ATGCTCGACAACCGTGAAGCGA TGAC** 3′

rev: 5′ TCGAGTGCGGCCTAGG**TTACCGCTGGTACCGCAGCAGG** 3′

Expression. For expression, the plasmid containing the FAS construct and the plasmid containing the AcpS were co-transformed into *Escherichia coli* BL21 (DE3) gold competent cells (Agilent Technologies) according to the manufacture's protocol. Cells were plated on LB + 1.5% agar (containing 50 µg/ml ampicillin, 11 µg/ml chloramphenicol, 0.01% arabinose) and grown over night at 37 °C. Five clones were picked randomly, then united in one pre-culture (35 ml LB media, 100 µg/ml ampicillin, 34 µg/ml chloramphenicol, 0.01% arabinose) and incubated at 200 r.p.m. at 37 °C overnight. For the main culture, the pre-culture was transferred into 2 l TB media (containing 100 µg/ml ampicillin, 34 µg/ml chloramphenicol, 0.01% arabinose). The main culture was incubated at 180 r.p.m. at 37 °C until optical density at 600 nm (OD_{600}) 0.8 to 1.0. After cooling to 20 °C, expressions were induced with IPTG (final concentration, 250 µM). Expressions were performed at 20 °C overnight. Then, after centrifugation at 7,000 r.c.f. for 14 min, the supernatant was discarded and the cell pellet was used directly for protein purification or stored at −80 °C after freezing in liquid N_2.

Protein purification. For protein purification, the cell pellet from 1 L TB culture (approximately 20 g) was resuspended in buffer W (100 mM Na_2HPO_4/NaH_2PO_4 pH 7.2, 100 mM NaCl, 1 mM EDTA) to a total volume of 35 ml. DNase I (2 mg; AppliChem) and protease inhibitor (cOmplete EDTA-free, Roche) were added.

Cells were broken using French press (16,000 psi, 1,100 bar). To avoid protein degradation, all of the following steps were carried out at 4 °C. After centrifugation for 1 h at 60,000 r.c.f., the supernatant (approximately 30 ml) was put on a strep column with a 5 ml matrix volume (purchased at IBA). After washing with 8 column volumes (CV) of buffer W, the protein was eluted with 3 CV of buffer E (same as buffer W but with 2.5 mM D-Desthiobiotin). The fractions were checked for impurities by SDS–PAGE and then concentrated in a centrifugal filter with a 100,000 nominal molecular weight limit (Amicon Ultra-4, Merck Millipore). The sample was further purified by size-exclusion chromatography (column: Superose 6 10/300GL, GE Healthcare, buffer G: 100 mM Na_2HPO_4/NaH_2PO_4 pH 7.2, 100 mM NaCl) and examined for its oligomeric state. Fractions were pooled and concentrated to a final concentration of 10–20 mg/ml of protein. Overall protein yields varied from 3–18 mg per liter TB culture with a typical yield of 10 mg, on average. A loss of activity was observed after the final purification step, most likely due to loss of flavin mononucleotide (FMN). Therefore, the samples were incubated with FMN in a five-fold molecular excess for 3–5 h. By this procedure, the activity could be restored. For storage, glycerol was added to a final volume of 50% and the protein sample was stored at −20 °C until used. Protein stability was checked via UV-circular dichroism measurements (for selected samples, see **Supplementary Fig. 12** and **Supplementary Table 2**). For more information on the integrity of the protein and the repeatability of the protein purifications, see below and **Supplementary Figure 3**.

Activity assay. Activity assays on module 1, were prepared on 60 µl scale including 30 µl of buffer AB (400 mM KH_2PO_4/K_2HPO_4 pH 7.3, 3.5 mM DTT), 50 nmol acetyl-CoA, 30 nmol NADPH and 25 µg of the FAS protein. The reaction was incubated for two minutes at room temperature. After one minute of recording the absorption at 334 nm, the reaction was started by the addition of 60 nmol malonyl-CoA. For a typical graph, see **Supplementary Figure 13**. The activity assay of module 2 was performed on a 60 µl scale including 30 µl of buffer AB (see above), 40 nmol C8-CoA and 30 µg of the FAS protein. The reaction was incubated for 2 min at room temperature. After 2 min of monitoring the absorption at 298 nm (absorption of lactone ring of 6-HHP without interference of any CoA absorption), the reaction was started with the addition of 50 nmol malonyl-CoA. For a typical graph, see **Supplementary Figure 14**. Turnover numbers [min^{-1}] were calculated from specific activities under the assumption that the enzymes were probed under substrate saturation.

Product assay setup. For testing product assay repeatability, samples were probed in biological and technical replicates (for a detailed description, please also see below and **Supplementary Figs. 2** and **7**). The product assay of module 1 was prepared on a 100 µl scale including 50 µl of buffer AB (400 mM KH_2PO_4/K_2HPO_4 pH 7.3, 3.5 mM DTT), 20 nmol acetyl-CoA, 100 nmol malonyl-CoA, 225 nmol NADPH and 20 µg of the FAS protein. CoA ester analysis was conducted with proteins from 50% glycerol stock that was stored at −20 °C. Glycerol did not have an influence on the product distribution (**Supplementary Fig. 15**). Since glycerol had an influence on 6-HHP production, freshly purified enzyme was used for module 2 assays and coupled reaction assays. For product analysis of the separated module 2, the reaction solution as used in the activity assay was worked up and analyzed for 6-HHP output.

doi:10.1038/nchembio.2314

The standard product assay for a sequential one-pot synthesis was prepared on a 100 µl scale including 50 µl of buffer AB (see above), 40 nmol acetyl-CoA, 200 nmol malonyl-CoA, 225 nmol NADPH, 20 µg of the FAS protein module 1 and 30 µg of the FAS protein module 2.

For the variations in substrate concentrations, the amounts of acetyl-CoA were 10 nmol, 20 nmol, 30 nmol and 40 nmol and the amounts of NADPH 75 nmol, 125 nmol, 175 nmol and 225 nmol. The combination of these conditions gave a matrix of 16 combinations.

The assays were left to react overnight. An internal standard was added; for the analysis of CoA esters (module 1) it was isovaleryl-CoA (iC5-CoA) and *n*-heptadecanoyl-CoA (C17-CoA) (both acquired from Sigma-Aldrich), for lactones 6-octyl-4-hydroxypyran-2-one was used (for synthesis see below). In order to stop the reaction and to precipitate the protein, four volumes of acetone (cooled to −20 °C) were added directly after internal standard addition. The mixture was vortexed for 20 s and kept at −20 °C for 1 h to complete the precipitation process. After centrifugation (5 min at 20,000 r.c.f.), 4/5 of the supernatant were transferred to a new vial and the solvent evaporated to dryness at 4 °C under reduced pressure in a SpeedVac. A volume of 60 µl water was added. The samples were treated 5 min in an ultrasonication bath and then were ready to be measured by HPLC–UV–MS.

HPLC quantification. LC–MS analysis of acyl-CoAs was carried out using a Dionex UltiMate 3000 RSLC coupled to a Bruker micrOTOF-Q II equipped with an electrospray ionization source. Chromatographic separation was performed on a RP-18 column (100 × 2.1 mm, particle size 1.7 µm, Waters Acquity BEH) with a mobile-phase system consisting of buffer A (water, 10 mM triethylamine/acidic acid buffer, adjusted to pH 9.0) and buffer B (acetonitrile). A multistep gradient at a flow rate of 0.25 ml min^{-1} was used with the starting condition of buffer B at 7%, a linear increase to 60% until 6 min, then to 70% until 9.5 min and finally to 90% until 10 min runtime. For samples initially aiming at enhanced MPT cleave off (mutation sites H1413 and Q1349), a different HPLC buffer system (buffer A: water with 0.05% triethylamine, buffer B: water/ACN 10:90 with 0.05% triethylamine) was used at the time. Data were always acquired in negative mode in a scan range from 200–2,000 *m/z* and later analyzed using DataAnalysis 4.0 software (Bruker Daltonik GmbH). For the quantification, the UV trace at 260 nm was used (for a typical result, see **Supplementary Fig. 10**). Calibration was done for C6-CoA, C8-CoA, C10-CoA, C14-CoA, C16-CoA, C18-CoA as well as for the internal standards iC5-CoA and C17-CoA (for details see **Supplementary Table 3**). The concentration in the sample was determined and initial concentration in the assay was calculated to the internal standard reference. For further characterization, see below. MS traces were also scanned for other CoA ester species (any other lengths longer ≥C_4-CoA, any hydroxyl or partially reduced derivatives) but no additional species to the ones we calibrated were found.

LC–MS analysis of lactones (6-HHP) was carried out using a Dionex UltiMate 3000 RSLC coupled to a Bruker micrOTOF-Q II equipped with an electrospray ionization source. Chromatographic separation was performed on a RP-18 column (100 × 2.0 mm, particle size 2.5 µm, Phenomenex Luna HST) with a mobile-phase system consisting of water and acetonitrile (each containing 0.1% formic acid). A multi-step gradient at a flow rate of 0.40 ml min^{-1} was used with starting conditions of ACN at 5% until 2 min, then increased gradually to 95% until 8 min runtime. For identification, MS data was acquired in positive mode in the range from 100–1,500 *m/z* and later analyzed using DataAnalysis 4.0 software (Bruker Daltonik GmbH). For quantification, the UV trace at 298 nm was used (a typical result is shown in **Supplementary Figure 16**).

Additional information on quantification of CoA esters by HPLC. Initially, the recovery of the extraction procedure was tested with the internal standards (three different concentrations: 20 µM, 50 µM, 100 µM with four replicates). The obtained recovery rates for iC5-CoA were (74.5 ± 0.9)% for 20 µM, (74.8 ± 1.4)% for 50 µM; (73.1 ± 1.7)% for 100 µM. For C17-CoA the rates were (70.1 ± 1.8)% for 20 µM, (75.0 ± 2.0)% for 50 µM; (72.4 ± 1.9)% for 100 µM.

The calibration coefficients and calibration ranges are listed for each CoA species in **Supplementary Table 3**. For the shorter acyl-CoAs (C6-CoA, C8-CoA, C10-CoA), iC5-CoA served as the internal standard to calculate losses from the work up procedure, while C17-CoA was used as internal standard for the longer acyl-CoA (C14-CoA, C16-CoA, C18-CoA).

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

Additional information on quantification of 6-HHP 1 by HPLC. For 6-HHP analysis in a sequential reaction, the recovery of the extraction procedure was tested on three different concentrations (10 µM, 50 µM, 100 µM) with four replicates for each concentration. The obtained rates were (74.7 ± 5.1)% for 10 µM, (76.7 ± 6.1)% for 50 µM; (73.1 ± 2.8)% for 100 µM respectively.

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

Chemical synthesis of the internal standard 6-octyl-4-hydroxypyran-2-one. For the synthesis of 6-octyl-4-hydroxypyran-2-one, a slightly modified protocol from ref. 26 was used. 500 mg (3.96 mmol, 1 eq.) 4-hydroxy-6-methyl-2H-pyran-2-one were dried by suspending it three times in 20 ml toluene and removing the solvent under reduced pressure. Anhydrous THF (20 ml) was added to the residue, and the resulting suspension was cooled to −78 °C. Then, 5.70 ml (9.12 mmol, 2.3 eq., 1.6 M solution in hexane) *n*-butyllithium was added dropwise over 15 min. The suspension was stirred at −78 °C for 4 h before 1.30 ml (1.79 g, 7.93 mmol, 2 eq.) 1-iodoheptane were added dropwise. The reaction was slowly warmed to room temperature overnight while being stirred. A 3 M HCl solution was used to quench the reaction by slowly adjusting the pH to 2. After separation of the phases, the aqueous phase was extracted with Et_2O (3× 30 ml). The combined organic phases were washed with brine and dried over Na_2SO_4. After filtration and evaporation of the solvent under reduced pressure, the residue was redissolved in acetone and concentrated onto silica gel. Purification by flash chromatography (silica gel, gradient of petrol ether/ethyl acetate 1:0, 9:1, 5:1, 4:1, 3:2) gave the desired product as a white powder (16.0 mg, 0.07 mmol, 2% yield). The product was stored at −20 °C until utilization. The purity was confirmed by HPLC–UV at 298 nm. Chemicals were obtained from Sigma-Aldrich.

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

Experimental repeatability. *Quantitative product analysis by HPLC:* We purchased standards for each acyl-CoA ester (purity checked in lab), and calibrated with standards providing the basis for quantifications (see above and **Supplementary Table 3**). By using purchased compounds iC5-CoA and C17-CoA, as representatives of short and long chain acyl-CoA esters, we optimized the recovery rates of the reaction work-up. iC5-CoA and C17-CoA were further used as internal standards in all assays ensuring precise quantification of losses in reaction work-up. For the coupled reaction, we worked with 6-HHP 1 for optimizing the recovery rates and with 6-octyl-4-hydroxypyran-2-one as internal standard.

Repeatability of FA and 6-HHP productions: The repeatability of our data was tested in two parallel experiments. A first experiment addressed the repeatability of module 1 (FAS1) short FA production. Here, we focused on the wild-type FAS, as well as the two KS-engineered constructs $FAS1^{G2599S}$ and $FAS1^{G2599S\text{-}M2600W}$. A second experiment addressed the repeatability of the coupled reaction, producing 6-HHP. Here, we worked with proteins $FAS1^{G2599S\text{-}M2600W\text{-}R1408A}$ and $FAS1^{G2599S\text{-}M2600W\text{-}I3151A\text{-}R1408K}$, as well as $FAS2^{S126A\text{-}Y2227F}$. Both experiments were run with four biological samples; i.e. proteins, which were we prepared from independent expression cultures (separate transformation of plasmids in the *E. coli* expression host).

(1) Repeatability in FA synthesis: The biological repeatability of our data is shown in **Supplementary Figure 2**. As demonstrated by FA output spectra of three FAS1 constructs (wild-type FAS, FAS1^{G2599S} and FAS1$^{G2599S\text{-}M2600W}$), each one probed in four biological replicates, variation between biological samples is low. The data in **Supplementary Figure 2** were collected simultaneously in one experiment. We note that FA chain length distribution is slightly systematically shifted toward shorter FA, as compared to data presented in **Figure 2a** (higher yield of C8-CoA for the FAS$^{G2599S\text{-}M2600W}$ and lower yield in C18-CoA). As the sensitivity of the FA chain length distribution on acyl-CoA substrate concentrations has been observed before[8], as well as data on 6-HHP had revealed an influence on substrate concentrations on the overall 6-HHP yield (**Fig. 2c**,**d**), we speculate that, particularly, variations in substrate concentrations might account for systematically shifted spectra.

(2) Repeatability in 6-HHP synthesis: We collected data for both coupled reactions (FAS1$^{G2599S\text{-}M2600W\text{-}R1408K}$→FAS2$^{S126A\text{-}Y2227F}$ and FAS1$^{G2599S\text{-}M2600W\text{-}I151A\text{-}R1408K}$ →FAS2$^{S126A\text{-}Y2227F}$), varying module 1 as well as module 2. Similar to the described FA synthesizing module 1, we worked with four independent biological samples. We generally observed high biological repeatability in the production of 6-HHP 1 by the coupled reaction (**Supplementary Fig. 7**).

Having established good repeatability on both the biological and technical levels, and given that our findings of interest concern the relative changes in FA chain length distribution, we proceeded to largely build our experimental set-up on single biological samples. Working with single biological samples allowed us to probe a set of constructs in a single experiment, thereby minimizing experiment-to-experiment variations. Accordingly, data presented in **Figure 2a**,**b**, have been collected in single experiments (for example, proteins prepared in parallel under identical conditions, as well as reactions run with proteins of the same "age" and with identical substrate stocks), allowing the impact of the mutations on FA chain length patterns to be assessed as accurately as possible.

Repeatability of protein purification and enzymatic activates: This section refers to data presented in **Supplementary Figure 3**. For the preparation of FAS, we used highly optimized protocols (for details to the preparation of proteins see above). While yields and the purity of proteins were generally high (>95% as judged from SDS–PAGE gels), we observed preparation-to-preparation variations in enzymatic activity and oligomeric purity indicating that not all parameters in protein preparation are under strict control. These variations do not seem to be systematic. For example, construct FAS1$^{G1250S\text{-}M2600W\text{-}I151A\text{-}R1408K}$ assembled well to hexamers, while construct FAS1$^{G1250S\text{-}M2600W\text{-}R1408K}$ showed an assembly defect. We cannot think of any molecular basis that is able to account for AT-domain mutation I151A counteracting an assembly defect of FAS1$^{G1250S\text{-}M2600W\text{-}R1408K}$ (see **Supplementary Fig. 3b**). As a last step in FAS purification, we performed SEC to select for the hexameric fraction. We observed not an equilibrium of hexameric and monomeric species, but stable hexamers that slowly disassemble over time. Disassembly might be one effect that causes the decreasing activity over time, which we observed for FAS constructs. Storage of FAS in glycerol at −20 °C was intended to slow down such aging effects and to maintain protein integrity during the time window when experiments were performed.

Enzymatic activities of different biological samples, presented in **Supplementary Figure 3c**,**d**, have been observed to vary more strongly than product distribution spectra and product yields (see **Supplementary Figs. 2** and **7**). This indicates that protein quality is underlying certain influences that are not entirely controlled by our purification protocols. Enzymatic activity has not been used as a read-out or model parameter, and variations in activity therefore do not affect the claims made in this work.

Molecular simulation. The crystal structure of *S. cerevisiae* FAS with PDB code 2UV8 was used for structure-based studies. Molecular dynamics simulations were conducted using Gromacs[27], versions 4.5 and customized coding derived from version 4.0.5 (refs. 28,29). The Amber ff99SBildn[30] force field was used to represent the protein and common molecules and ions, and the General Amber[31] force field was used to parameterize substrate molecules.

In the simulation studies reported, the computationally prohibitive size of the full-sized FAS protein necessitated a truncated representation of areas of interest. Cubes of 8 nm edge lengths were constructed with the key catalytic residues of active sites of interest at their centers. The orientations of the cubes were selected to maximize the number of atoms they contained. All atoms belonging to amino acid residues not lying fully within the cubes were discarded, and distance restraints were applied to all atoms within 0.5 nm of the positions of any of the deleted atoms, in order to preserve the structure along the boundary. Including water and ions at a concentration of 150 mM, simulation systems consisted of ca. 59,000 atoms.

Free energy calculations. Free energy differences between enzyme-bound forms of related substrates were calculated using alchemical techniques. Hybrid topologies, representing either mutations between amino acids or changes in acyl chain length, were constructed using custom software according to a single-topology paradigm. Herein, in both the beginning and end states of the topology, atoms that do not form part of the respective state are present as part of the same molecule, but do not interact with the rest of the system other than through bonded interactions carefully chosen so as not to affect the end state[32]. Free energy differences were calculated using approaches described previously[28], and using CGI for changes in acyl chain length and RE/MBAR for amino acid mutations.

Modeling of KS domain active site. Modeling of enzyme–substrate complexes of the ketoacyl synthase active site was facilitated by the availability of a crystal structure of the *S. cerevisiae* FAS with the inhibitor cerulenin bound (PDB code 2VKZ), providing a template for accommodation in the KS active site of a long, hydrophobic substrate.

For the purposes of characterizing the effect of increasing chain length in enzyme–substrate complexes of acyl intermediates, substrates were represented using a truncation of the ACP-bound form, consisting of the terminal thioester atoms of the phosphopantetheine arm of the ACP covalently bound to the acyl intermediate. A representation of a pre-catalytic complex was constructed on the basis of the reported catalytic mechanism[33], with distance restraints enforcing proximity between the substrate carbonyl and the backbone amide hydrogens of the nucleophilic C1305 and F1646, which form the 'oxyanion hole', and between the sulfur atom of C1305 and the α-carbon of the substrate. These distance restraints were kept in force for the duration of molecular dynamics simulations seeking to elucidate bound states of intermediates, and for the calculation of free energy differences for successive addition of methyl groups to the growing acyl chain.

Systems representing the wild type KS domain and mutant G1250S–M1251W (G2599S–M2600W in *C. ammoniagenes*) were constructed in 8 nm cubes as detailed above. Wild type and mutant systems were identical in their composition except for the introduced mutation. Free energy changes were calculated separately for the 16 individual methylation steps in the chain length range C2–C18, in the respective contexts of the wild type and the mutant protein construct. The difference between respective free energy changes calculated in the mutant and in the wild type protein yields the contribution of the mutation to the binding free energy change associated with adding each methyl group ($\Delta\Delta G$ per methylation). Taking the C_2 complex as the reference point, successive $\Delta\Delta G$ values can be summed to yield a relative binding free energy in the mutant protein versus wild type for each chain length in the range 4–18, as per:

$$\Delta\Delta G = \sum_{i=4,6\le i\le n,i\ \text{even}} \left(\Delta G_{mut}^{i-2\to i} - \Delta G_{WT}^{i-2\to i}\right)$$

Modeling of MPT and AT domain active sites. Initial alignment of acetyl- and malonyl-CoA to the MPT and AT domain active sites was modeled on the basis of homologous PDB entry 2G2Z. Systems representing the wild type protein and mutants R1834K (MPT domain, R1408K in *C. ammoniagenes*) and I306A (AT domain, I151A in *C. ammoniagenes*) were constructed using standard techniques; hybrid topologies describing a single-topology alchemical transformation pathway between them were constructed using custom software. Free energy changes were calculated for the amino acid point mutations in the presence and absence of the respective ligands, with the difference corresponding to the change in binding free energy of each ligand resulting from the mutation.

doi:10.1038/nchembio.2314

Modeling of palmitoyl binding to the MPT domain. Unlike the KS domain active site, modeling of substrate interactions in the malonyl–palmitoyl transferase (MPT) active site did not benefit from the availability of a co-crystal structure with an acyl chain analog. Through knowledge of the catalytic mechanism of the acyltransferase activity and the crystal structure[34], a putative pre-catalytic complex could in principle be constructed analogously to the approach described above for the active site of the KS domain, with distance restraints between the substrate carbonyl and the backbone amide hydrogens of Q1669 and L1809 representing the oxyanion hole, and from the side chain hydroxyl hydrogen of S1808 to the thioester carbon of the acyl intermediate. However, in the case of MPT, there is no clear pre-formed hydrophobic channel for longer acyl chains.

Evidence suggests that longer acyl chains will interact with patch of exposed hydrophobic residues[11]. However, initial modeling based on the crystal structure showed that a 16–18-carbon acyl chain is too short to bridge over a ridge formed by M1838, T1774 and neighboring residues, indicating that non-trivial rearrangement with respect to the crystal structure are required to accommodate the presumed long-chain substrate binding mode in MPT.

Using the palmitoylate (C16) derivative as a probe substrate, molecular dynamics simulations were performed which made use of manual conformational selection to guide the exploration of the energy landscape in the direction of complexes fulfilling the experimentally determined structure-function hypothesis, without direct imposition of explicit biasing forces. Over multiple successive iterations, sets of 5–10 independent molecular dynamics simulations were launched, unbiased with the exception of the distance restraints maintaining a pre-catalytic conformation as described above, and allowed to run for up to 100 ns. The resulting conformations were then assessed individually, and a small subset of those qualitatively judged to most closely resemble a complex matching the experimental structure-function hypothesis were selected as input for a further iteration. After three such iterations, the molecular dynamics simulations arrived at a complex where both the putative pre-catalytic distance restraints and the interaction of the acyl chain with the hydrophobic patch were fulfilled.

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

As this structure was derived using the C16 derivative as a probe, and since the associated structural perturbations are non-trivial and can be expected to vary as a function of substrate acyl chain length, we do not present calculated free energy differences for successive substrate methylations based on this model, as we judge that the atomistic simulation techniques applied in this work, short of intractably exhaustive sampling, would fail to capture the details of the energy landscape representing the nontrivial 'open–closed' transition.

Kinetic model. Due to the substrate-shuttling mechanism of type I FAS, under which the physical transfer of all intermediates is mediated by a single tethered shuttle protein (ACP) per set of catalytic sites, we applied a stochastic modeling approach to generate statistically representative reaction trajectories of individual FAS units, while accounting for the discrete nature of the individual enzyme components. The algorithm of Gillespie[18] was implemented in custom-written software. Details of the simplified canonical reaction cycle of yeast FAS we represented, and of the assignment of rate constants and the subsequent filtering of candidate sets to yield an ensemble of models matching known properties of yeast FAS, are given in **Supplementary Note 1**.

Code availability. The custom-written software used to generate and analyze the kinetic model is available as **Supplementary Software**.

Data availability. All data generated or analyzed during this study are included in this published article (and its supplementary information files) or are available from the corresponding author upon reasonable request.

26. Zhang, X. *et al. Synthesis* **5**, 749–753 (2007).
27. Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. *J. Chem. Theory Comput.* **4**, 435–447 (2008).
28. Seeliger, D., Buelens, F.P., Goette, M., de Groot, B.L. & Grubmüller, H. *Nucleic Acids Res.* **39**, 8281–8290 (2011).
29. Buelens, F.P. & Grubmüller, H. *J. Comput. Chem.* **33**, 25–33 (2012).
30. Lindorff-Larsen, K. *et al. Proteins* **78**, 1950–1958 (2010).
31. Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A. & Case, D.A. *J. Comput. Chem.* **25**, 1157–1174 (2004).
32. Boresch, S., Tettinger, F., Leitgeb, M. & Karplus, M. *J. Phys. Chem. B* **107**, 9535–9551 (2003).
33. Olsen, J.G., Kadziola, A., von Wettstein-Knowles, P., Siggaard-Andersen, M. & Larsen, S. *Structure* **9**, 233–243 (2001).
34. Oefner, C., Schulz, H., D'Arcy, A. & Dale, G.E. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 613–618 (2006).