



Published in final edited form as:

*Nat Prod Rep.* 2018 October 17; 35(10): 1046–1069. doi:10.1039/c8np00039e.

## The Architectures of Iterative Type I PKS and FAS

Dominik A. Herbst<sup>a</sup>, Craig A. Townsend<sup>b</sup>, and Timm Maier<sup>a</sup>

<sup>a</sup>Department Biozentrum, University of Basel, Klingelbergstrasse 50/70, 4056 Basel, Switzerland

<sup>b</sup>Department of Chemistry, Johns Hopkins University, Baltimore, Maryland, 21218, United States

### Abstract

Type I Fatty acid synthases (FASs) are giant multienzymes catalyzing all steps of the biosynthesis of fatty acids from acetyl- and malonyl-CoA by iterative precursor extension. Two strikingly different architectures of FAS evolved in yeast (as well as in other fungi and some bacteria) and metazoans. Yeast-type FAS (yFAS) assembles into a barrel-shaped structure of more than 2 MDa molecular weight. Catalytic domains of yFAS are embedded in an extensive scaffolding matrix and arranged around two enclosed reaction chambers. Metazoan FAS (mFAS) is a 540 kDa X-shaped dimer, with lateral reaction clefts, minimal scaffolding and pronounced conformational variability. All naturally occurring yFAS are strictly specialized for the production of saturated fatty acids. The yFAS architecture is not used for the biosynthesis of any other secondary metabolite. On the contrary, mFAS is related at the domain organization level to major classes of polyketide synthases (PKSs). PKSs produce a variety of complex and potent secondary metabolites; they either act iteratively (iPKS), or are linked via directed substrate transfer into modular assembly lines (modPKSs). Here, we review the architectures of yFAS, mFAS, and iPKSs. We rationalize the evolution of the yFAS assembly, and provide examples for re-engineering of yFAS. Recent studies have provided novel insights into the organization of iPKS. A hybrid crystallographic model of a mycocerosic acid synthase-like Pks5 yielded a comprehensive visualization of the organization and dynamics of fully-reducing iPKS. Deconstruction experiments, structural and functional studies of specialized enzymatic domains, such as the product template (PT) and the starter-unit acyltransferase (SAT) domain have revealed functional principles of non-reducing iterative PKS (NR-PKSs). Most recently, a six-domain loading region of an NR-PKS has been visualized at high-resolution together with cryo-EM studies of a trapped loading intermediate. Altogether, these data reveal the related, yet divergent architectures of mFAS, iPKS and also modPKSs. The new insights highlight extensive dynamics, and conformational coupling as key features of mFAS and iPKS and are an important step towards collection of a comprehensive series of snapshots of PKS action.

---

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here].  
See DOI: 10.1039/c8np00039e

Conflicts of interest

There are no conflicts to declare.

## I. Introduction

Fatty acids are central molecules for life. They serve as core components of membrane lipids and act as energy storage compounds in the form of triacylglycerols. Fatty acids further act as signalling molecules and are attached to proteins to serve as membrane anchors. Biological synthesis of fatty acids is based on variations of a conserved pathway for the iterative elongation of carbohydrate-derived precursors by two-carbon units provided by coenzyme A (CoA)-activated carboxylic acids. In the late 1950s and early 1960s the principal mechanism of fatty acid biosynthesis was established (reviewed in ref.<sup>1</sup>): Acetyl-CoA is the initial substrate for fatty acid biosynthesis. It is iteratively elongated in multiple reaction steps by two-carbon units from malonyl-CoA, which is obtained by energy-dependent carboxylation of acetyl-CoA. The growing intermediate is tethered to a small acyl carrier protein (ACP) by covalent thioester linkage to the terminal thiol group of a phosphopantetheine (Ppant) cofactor bound to a conserved ACP serine. In the late 1960s, individual enzymes had been isolated from bacteria catalysing all reactions of fatty acid biosynthesis. In the 1970s, it was discovered that eukaryotic fatty acid biosynthesis is carried out by large, multifunctional proteins, which had been termed Type I fatty acid synthases (FASs) in contrast to bacterial dissociated Type II FASs. Sequences of eukaryotic FASs finally demonstrated in the late 1980s that yeast FAS and metazoan FAS (mFAS) have a completely different domain organisation.

Sequence analysis of the gene cluster for the biosynthesis of the antibiotic erythromycin in *Saccharopolyspora erythraea* demonstrated in 1990 that the enzymatic domains and their linear organization in microbial polyketide synthases (PKSs), which produce complex and potent secondary metabolites, is related to those of mFASs<sup>2</sup>. Here, we review the structural organization of the iterative yeast-type (yFASs) and mFASs, together with those of iterative PKSs (iPKSs), which also catalyse multiple rounds of precursor elongation by the same set of enzymatic domains in one multienzyme. The architecture of modular PKSs (modPKSs), which hand-over intermediates after a single round of elongation to a downstream module and include the erythromycin biosynthesis system will be reviewed elsewhere in this issue.

## II. Yeast-type Fatty Acid Synthase (yFAS)

In yFAS, the acetyl moiety of acetyl-CoA is loaded by a specific acetyltransferase (AT) onto ACP prior to transfer to the active site cysteine of the condensing ketosynthase (KS) (Fig. 1A). The malonyl-moiety of malonyl-CoA is transferred by the malonyl-palmitoyl-transferase (MPT) to ACP. KS catalyses the decarboxylative Claisen condensation of the acetyl and malonyl groups to yield a  $\beta$ -ketoacyl intermediate bound to the ACP. This intermediate is modified by the sequential action of an NADPH-dependent ketoreductase (KR), water elimination by a dehydratase (DH) and reduction by an FMN-containing NADPH-dependent enoylreductase (ER) to yield a fully saturated acyl intermediate elongated by two carbon atoms. This intermediate serves as the initial substrate for another round of elongation. After seven or eight iterative elongation cycles, the resulting acyl-group is transferred back to CoA by MPT to produce the products palmitoyl- and stearoyl-CoA, respectively; for an extended discussion of individual enzymatic functions see<sup>1, 3, 4</sup>.

## II.1 The architecture of yFAS

Yeast FAS was characterized at low resolution beginning in the 1960s by electron microscopy (EM) and solution scattering techniques, but detailed insights into the structure of yeast and other fungal FASs were only obtained between 2006 and 2008 by X-ray crystallography: The structure of *Thermomyces lanuginosus* FAS was first published at 5 Å resolution<sup>5</sup>, and then at 3.1 Å resolution<sup>6</sup>, simultaneously with a structure of yeast FAS at the same resolution<sup>7</sup>. Importantly, in the yeast FAS structure the ACP was fortuitously trapped under the crystallization conditions at the KS active site. A concurrent 4 Å resolution crystal structure<sup>8</sup> misinterpreted chain topology in the MPT region, but provided the localization of the integral phosphopantetheinyltransferase (PPT) domain, which catalyses the Ppant-ylation of the integral ACP domain of yFAS, in agreement with a subsequent 4 Å crystal structure of yeast FAS inhibited by the covalent KS inhibitor cerulenin<sup>9</sup>. Further insights into PPT function were obtained by the structure determination and functional characterization of the excised PPT domain<sup>10</sup>. Finally, a cryo-EM reconstruction at 5.9 Å resolution provided a solution structure of yeast FAS and an initial suggestion for the localization of ACP in various steps of the catalytic cycle<sup>11</sup>. In the following paragraph, we will discuss the architecture of yeast FAS itself based on the above, before turning to variations and evolution of yFAS.

The catalytic domains of yeast FAS are distributed over two protein chains (Fig. 1B). The FAS1 gene encodes the  $\beta$ -chain, which comprises (in sequence order) the AT, ER, DH and the majority of the split MPT domain. The  $\alpha$ -chain is encoded by the downstream FAS2 gene, and contains the smaller part of the MPT domain, and the ACP, KR, KS and PPT domains. The two chains assemble into a heterododecameric 2.6 MDa  $\alpha_6\beta_6$  complex with D3 symmetry and a barrel shape with overall dimensions of 250 Å x 270 Å. Conserved regions of enzymatic domains in fungal FAS provide only roughly 50% of the total polypeptide length. The other half of the sequence is composed of non-catalytic rigid scaffolding elements and a few flexible linking regions (Fig. 2A,B).

Six  $\alpha$ -chains form a wheel-like structure with D3 symmetry, with a central helical hub, connected via helical spokes to the rim. The dimeric KS and KR domains together with linking elements form the tire. The ACP is double tethered; a 47 amino acid (aa) long, alanine and proline rich linker connects its *N*-terminus to the MPT fragment, and the ACP *C*-terminus is bridged to a scaffolding region preceding the KR by a 27 aa long, proline-free linker. The PPT domain is connected to the outside of the tire, without access to ACP and in an inactive monomeric form, indicating that Ppant-ylation of ACP occurs prior to complete yeast FAS assembly catalysed by a PPT oligomer formed between different  $\alpha$ -chains<sup>10</sup>.

Two domes with C3 symmetry formed by three  $\beta$ -chains cap the  $\alpha$ -chain wheel on either side. The  $\beta$ -chains include the structurally and functionally related MPT and AT acyltransferases, and a double-hotdog fold DH extended by another hot-dog-fold linker element, as well as an FMN-containing TIM-barrel ER. In contrast to the KR and KS domains, all domains of the  $\beta$ -chain are monomeric. The  $\beta$ -chain domes define two large reaction chambers (Fig. 2C) at either face of the wheel, each containing three full sets of enzymatic domains required for fatty acid biosynthesis as well as three mobile ACP domains, which are double tethered to the central hub and the chamber walls. The openings

in the chamber wall are small, presumably restricting access of other proteins to the reaction chambers.

The ACP domain of yeast FAS is extended relative to the canonical bacterial four-helix bundle ACP by a second four-helix bundle. Based on NMR studies of the yeast FAS ACP domain, the canonical four-helix bundle, and not the additional subdomain of ACP, interacts with covalently attached acyl chains, however, only very weakly with about 15% sequestration of acyl chains to ACP<sup>12</sup>. In the yeast FAS crystal structure, ACP binds symmetrically to both active sites of the KS dimer. The canonical four-helix bundle of ACP primarily interacts with the KS domain, while the extension bundle mediates further interactions to the non-catalytic scaffolding spoke region. The environment for substrate shuttling by restricted ACP diffusion appears to be well adapted to the catalytic cycle of fungal FAS. KS, KR, DH, and ER are located on a closed path. MPT, which is required in each elongation cycle and for product release is located closer to the condensing KS than AT, which is only required in the starting cycle. Chain-length determination is mediated by kinetic competition between substrates and intermediates for MPT and possibly also for KS<sup>3, 13</sup>. The product spectrum depends on the substrate conditions; under physiological conditions palmitoyl-CoA and stearoyl-CoA are produced in a 2:3 ratio.

## II.2 Bacterial and secondary-metabolism yFAS

Yeast FAS is the most thoroughly studied member of the yFAS family. Nevertheless, as early as 1969, multienzyme yFAS had been discovered as co-occurring with dissociated Type II FAS systems in *Mycobacterium phlei*<sup>14</sup>. Based on the availability of genomic sequences, yFAS systems have now been discovered as components for fatty acid biosynthesis<sup>15–18</sup> across three genera of the phylum *Actinobacteria*, the *Corynebacteria*, *Mycobacteria* and *Nocardia*, also known as the CMN group of *Actinobacteria*. In *Mycobacteria*, yFAS contributes to providing precursors for the biosynthesis of complex cell wall lipids. A secondary metabolic yFAS specialized in the production of hexanoic acid (HexS) has been found in the filamentous fungus *Aspergillus parasiticus*<sup>19</sup>. Its hexanoic acid product is directly transferred as a starter unit to an iterative non-reducing PKS (NR-PKS) for the synthesis of aflatoxin B<sub>1</sub><sup>20</sup>. Due to the inherent difficulties in aligning the less-conserved non-enzymatic sequence regions, understanding of the structural and functional differences between various bacterial and fungal yFAS has been linked closely to experimental structure determination.

Structures of mycobacterial yFAS have been resolved via cryo-EM at resolutions of 7.5 Å under application of D3 symmetry<sup>21</sup> and in three distinct conformational states without applying symmetry at 17.5 Å to 27.0 Å resolution<sup>22</sup>. Interpretation of the low-resolution maps was achieved using models derived on the basis of the yeast FAS structure. The central wheel region is similar between mycobacterial and yeast FAS. The dome regions are slightly smaller and have larger lateral or top openings due to the absence of some of the non-catalytic regions. Such reduced extent of scaffolding is likely linked to the observed higher structural variability of yFAS from *Mycobacteria*. Conformational changes may well be coupled to substrate turnover<sup>22</sup>.

### II.3 The evolution of yFAS

Supported by structural insights, the evolution from early bacterial yFASs to yeast FAS can be traced along gene splitting, domain deletion and domain duplication events (Fig. 3A)<sup>17, 18</sup>. A primordial bacterial yFAS was encoded by a contiguous gene encompassing the enzymatic domains for fatty acid biosynthesis as in the yeast FAS1 and FAS2 genes, but lacking several scaffolding regions. In bacterial yFASs, the PPT domain is not part of the contiguous single gene, but it is encoded as a separate protein by the directly adjacent downstream *acpS* gene in *Mycobacteria*. The considerably more accessible reaction chambers of bacterial yFAS may provide access of an external PPT or other proteins directly interacting with ACP-bound substrates, resembling the direct substrate transfer suggested for HexS.

Early fungal FASs are encoded by a single gene, which includes the PPT encoding region. In fungal evolution, splitting of the contiguous FAS gene occurred at different positions, leading to splitting into two consecutive FAS genes. Splitting never occurs directly in the ACP linker, so that the ACP remains double-tethered. Segment duplication presumably has contributed to the extension of fungal FAS, as indicated by the occurrence of a third, non-catalytic hotdog-fold next to the catalytic double hotdog fold of the DH domain, or the duplication of the ACP domain in the oleaginous yeast *Rhodospiridium toruloides*' FAS, which has also been structurally characterized by cryo-EM at 7.8 Å resolution<sup>23</sup>.

The evolution from early bacterial to late fungal yFAS is mapped out by a large number of intermediate sequences. The origins, however, of a primordial contiguous FAS gene encoding the entire yFAS are enigmatic. Due to the low level of sequence conservation in non-enzymatic linker regions, early evolutionary intermediates are difficult to detect by sequence analysis alone. Crystallographic structure determination, however, has confirmed the presence of early precursors of docking and scaffolding elements from yFAS already in two bacterial proteins involved in Type II FASs and polyketide biosynthesis<sup>17</sup>: FabY, a non-canonical starter KS for fatty acid biosynthesis in *Pseudomonas aeruginosa* is a close homolog of yFASs KS domains. Structure determination revealed that it already encompasses structural elements contributing to the spoke and hub structure of yFASs, while in FabY they presumably are stabilizing the dimeric KS structure. The ER domain of the dual-function *trans*-acting AT-ER protein DfnA, part of the *trans*-AT PKS for difficidin biosynthesis, and likely also ER proteins in PUFA biosynthesis, contain insertion regions relative to the bacterial TIM-barrel FabK ER, which are ancestors of ER-KS interacting regions in yFASs. The regions of bacterial yFASs that presumably have evolved from the insertion regions of DfnA and FabY mediate a core contact between the FAS wheels and domes, qualifying the two proteins as potential early precursors for partial assembly of a bacterial yFAS.

### II.4 Engineering yFAS for alternate product synthesis

The chemical versatility of multienzymes renders them interesting targets for re-engineering of product formation, either via mutation of individual active sites or via deletion, addition or swapping of functional domains. In contrast to other families of multienzymes, such as PKSs, naturally occurring yFASs only produce a single type of product, saturated fatty acids,

albeit with different chain length. The extensive scaffolding and tight domain integration in yFAS massively complicate the exchange of enzymatic domains. Consistently, functional exchange of whole domains has not been reported for yFAS.

Nevertheless, two successful routes to the production of alternate products have recently been discovered (Fig. 3B-D). The first approach<sup>24</sup> exploits kinetic competition between enzymatic domains based on comprehensive simulation of product formation in yFAS and specific active site mutations (Fig. 3C). A first reaction module was engineered based on *Corynebacterium ammoniagenes* FAS and mutations in KS, MPT and AT, that increase the production of short-chain fatty acids. A second module designed on the same FAS by a mutation in KR loads the short chain products from the first module via its MPT domain and catalyses one round of condensation with malonyl-CoA, before an inactivating mutation in KR prevents normal  $\beta$ -carbon modification. The  $\beta$ -keto-intermediate is condensed in a side reaction to the normal elongation cycle one more time with malonyl-CoA. Spontaneous cyclization results in the release of a new lactone product. In the second approach<sup>25</sup>, the complications of inserting domains into the rigid yFAS barrel were elegantly circumvented (Fig. 3D). Here, the large reaction chambers and flexible nature of the ACP linkers were exploited to insert ectopic thioesterase domains with alternate specificity into the ACP linking regions in yeast FAS. This approach enabled production of short-chain fatty acids or methyl ketones by a single type of designed yeast FAS.

### III. Metazoan Fatty Acid Synthase

Fatty acid biosynthesis by mFASs uses a variation of the synthetic scheme of yFAS. The acyl groups of the starter substrate acetyl-CoA and the elongation substrate malonyl-CoA are loaded both by a single dual-specific malonyl-acetyl transferase (MAT) onto ACP (Fig. 4A). The ER of mFAS is a member of the medium-chain dehydrogenase family, different from the FMN-dependent TIM barrel ERs of yFASs<sup>6</sup> and bacterial FabK<sup>26</sup> or the canonical bacterial short-chain dehydrogenase ERs. The main product of mFAS is palmitate, which is released from the ACP by an integral C-terminal thioesterase domain.

#### III.1 The structural organization of mFAS

The FASN gene encodes the single protein chain of human mFAS (Fig. 4B), which forms a 540 kDa homodimer. Initial insights into the dimeric organization of mFASs were obtained by careful biochemical, mutational and complementation analysis<sup>1, 27, 28</sup>. Early EM analysis has been considerably more challenging than for yFAS due to the smaller size, less characteristic shape and higher conformational variability of mFAS, but yielded a shape description at a resolution of 19 Å<sup>29</sup>. Crystallographic analysis of an intact mammalian mFAS finally provided an atomic structural model at 3.2 Å resolution<sup>5, 30</sup> (Fig. 5A). The crystal structure did not resolve the flexibly tethered and mobile C-terminal ACP and TE domains, but structural data on the isolated domains are available (Fig. 5B).

The ACP domain of rat FAS has been analysed by NMR<sup>31</sup>. The ACP domain adopts a canonical four-helix bundle fold. In contrast to bacterial FAS, mFAS ACP does not sequester the covalently linked acyl moieties. The monomeric  $\alpha/\beta$  hydrolase fold TE has been resolved by X-ray crystallography<sup>32</sup>. A structure of TE in complex with the inhibitor Orlistat



further revealed a mechanism for chain-length selectivity based on distinct binding sites for acyl chains of different length<sup>33</sup>. The resulting discrimination of chain lengths together with the chain length-dependence of the mFAS KS activity provides the basis for the palmitate product specificity of mFAS.

The ACP linkers are not enriched in Ala-Pro residues as observed for  $\gamma$ FAS. The properties of the flexible linker leading into the ACP domain of mFAS have not been studied in detail. The longer linker between ACP and TE, typically 20 residues or more, is remarkably permissive for modifications: Substitution with a considerably longer linker from PKS had no effect, and mFAS variants with harsh or almost complete deletions in the linker region maintained about 30% of activity<sup>34</sup>.

The structurally resolved core of porcine mFAS adopts an X-shape structure with dimensions of approx. 200 Å x 170 Å x 110 Å (Fig. 5C, D). The structure is segregated into two subregions centrally connected via a narrow link of extended polypeptides. The dimeric condensing KS domain in the centre and laterally disposed monomeric loading MAT domains define the “condensing region”. The *N*-terminal KS domain is linked via a small  $\alpha/\beta$ -fold linker domain (LD) to the MAT, and a post-MAT linker winds through the linker domain and around the KS to the central connection of the two regions. The second subregion comprises the domains required for  $\beta$ -carbon modification, the KR, DH and ER domains, and is thus referred to as the “modifying region”. The modifying and condensing region hardly form any direct contact other than the covalent peptide linkage. Asymmetry of the crystal structure of dimeric mFAS suggests pronounced hinge-like flexibility around the central connection.

In the dimeric mFAS, the two DH domains arrange in a V-shape with few contacts between the DH domains, which is mostly stabilized by interactions with the dimeric ER domain. The KR domains are laterally linked to the ER domains. While related bacterial KR domains are true oligomers, the fold of the mFAS KR domain is completed by a non-catalytic fragment of a KR domain, termed the pseudo-KR ( $\Psi$ KR). A second non-catalytic domain is found in the modifying region. This domain resembles *S*-adenosyl-methionine (SAM) dependent C-methyltransferases (CMeTs), but lacks conserved active site residues, and is here referred to as pseudo-C-methyltransferase ( $\Psi$ CMeT).

The domains in the modifying region are linked by a minimal amount of linking sequences, either extended peptides or short  $\beta$ -strands. Overall, only 9% of the total sequence is invested into linking regions and an additional 16% in the two non-catalytic domains, the pseudo-KR ( $\Psi$ KR) and  $\Psi$ CMeT (Fig. 5C) in porcine mFAS.

The X-shaped fold of mFAS exposes all enzymatic domains to the environment. In contrast to  $\gamma$ FAS, the ACP is not confined into a specific reaction chamber, but is accessible by other proteins. The accessibility of mFAS TE is exploited by thioesterase 2 (TE2)<sup>35</sup>, a short chain-specific thioesterase, which is highly expressed in lactating mammary glands, but also in breast cancer. TE2 cleavage of intermediates from mFAS-ACP results in an enrichment of short-chain acyl products.

The X-shape of porcine mFAS defines two lateral reaction clefts each surrounded by a complete set of enzymatic domains required for fatty acid biosynthesis (Fig. 5D). Based on the static crystal structure, the tether that links ACP to KR is just sufficient to reach all active sites of the proximal reaction cleft, but too short to transfer substrates into the distal cleft. However, biochemical data indicated that also substrate transfer between the two clefts occurs albeit at a lower rate<sup>28, 36</sup>

### III.2 Structural variability of mFAS

The analysis of contacts and asymmetry in the crystal structure of mFAS already suggested the possibility for large-scale conformational changes<sup>30, 37</sup>, in particular a hinge motion around the central connection. Experimental evidence for such large-scale conformational changes was provided by EM analysis, which indeed identified a variety of modes of motions in mFAS<sup>38</sup>. Sorting of particles observed in negative stain EM at low resolution into conformational classes in 2D and 3D depicted large-scale swinging and rotation motions around the central connection, which behaved like a ball-and-socket joint. Little structural variation was observed in the condensing region, in line with a close structural agreement of the structure of this region in the entire mFAS and as an excised fragment<sup>4, 39</sup>. Considerably larger variability was detected in the modifying region. The close match between the structures of an engineered human  $\Psi$ CMeT- $\Psi$ KR-KR fragment bound to an inhibitor and the corresponding region in intact mFAS<sup>40</sup> suggest, that a large contribution to the variability in the modifying region may originate from the motion of ER- and  $\Psi$ CMeT- $\Psi$ KR-KR relative to DH.

The EM analysis of mFAS trapped by mutation at various stages of the catalytic cycle revealed some correlation between the functional state and the conformational equilibrium of mFAS<sup>38</sup>. Still, multiple conformational states were observed in each state. The structural and mechanistic basis for the coupling of conformational changes to enzymatic turnover remains unknown. To resolve this issue, novel approaches for studying single mFAS molecules during turnover may be required, *e.g.* based on single-molecule fluorescence. Most recently, it has also been demonstrated that high-speed atomic force microscopy is able to resolve large-scale conformational changes in mFAS at a spatial and temporal resolution permitting detailed analysis of the transitions of single mFAS molecules between different conformational states<sup>41</sup>. However, further work will be required to apply this technique for a comprehensive analysis of mFAS structural dynamics.

## IV. From FAS to Iterative Polyketide Synthases (PKS)

While FASs are highly optimized molecular machines specifically for the production of fatty acids, many multienzymes with an overlapping set of functional domains and a related synthetic scheme are not involved in primary fatty acid metabolism. In contrast to mFAS (Fig. 6A), these machines do not necessarily produce fully reduced products and reveal C<sub>α</sub> as well as C<sub>β</sub> modifications that originate from incomplete reduction cycles caused by the absence of specific modifying domains, by catalytically inactive domains, or by programmed domain skipping (Fig. 6B-E). Because of their chemically diverse products, these synthases are named according to their primary condensation product: Polyketide



synthases (PKSs). The principal catalytic domains of PKSs are closely related to those of mFAS. They include KS, DH, ER, KR as well as CMeT domains. Substrate loading is catalyzed by **acyl**transferases, which are specific for the particular substrate of the PKS and are commonly also abbreviated as AT, but should not be confused with the *acetyl* specific acyltransferase of yFASs.

All PKSs that make use of dissociated domains (type II and III) as well as all members of the multienzyme (type I) superfamily work iteratively<sup>42</sup>, except for modular PKSs (modPKSs) that function linearly through linked modules in an assembly line with vectorial transfer of intermediates<sup>43</sup>. The iterative type I PKSs (iPKS) include the distinct families of bacterial<sup>44</sup>, mycocerosic acid synthase-like PKSs (MAS-like PKSs)<sup>15, 45, 46</sup>, and fungal iPKSs<sup>47</sup>. Fungal iPKS are further classified by their degree of reductive behavior into highly reducing (HR-), partially reducing (PR-), and non-reducing (NR-) PKSs. In the past years, substantial progress has been made in characterization of the architecture of MAS-like, HR-, and NR-PKSs. This work has provided new concepts of PKS assembly with relevance also for other types of PKSs, in particular modPKSs, and will be reviewed in the following sections of this article. Progress towards the depiction of intact synthetic assemblies is less advanced for partially-reducing PKSs as well as for other specialized iPKSs, such as endiayne synthases<sup>48</sup>, polycyclic tetramate macrolactam synthases (PTMs)<sup>49</sup> and polyunsaturated fatty acid synthases (PUFAs)<sup>50</sup>, which are not covered in this review.

## V. Fully-Reducing Iterative PKS

MAS-like as well as HR-PKSs contain a fully reducing modifying region, such as mFASs (Fig. 6 A,B,E), and are capable of stepwise reduction of the primary  $\beta$ -carbonyl condensation product to form a fully saturated carbon-carbon bond using KR, DH, and ER activities. These single module PKSs are involved in the synthesis of secondary metabolites and often produce precursors or unusual fatty acids that are channeled to other PKSs or downstream processing enzymes. In chapters V.1 and V.2 these two enzyme families will be discussed in more detail. In contrast to HR-PKSs, NR-PKR lack all reducing/modifying domains and use the primary polyketide chain for aromatic cyclization reactions (Fig. 6D), which will be discussed in chapter VI. Despite integration into modPKS assembly lines, several modPKSs modules have been found to work iteratively<sup>51, 52</sup> (Fig. 6F), which will be discussed in the context of a common architecture with MAS-like PKSs.

### V.1. Fungal Highly Reducing PKSs (HR-PKSs)

Fungi are well known for their extraordinary repertoire of complex secondary metabolites including many toxins. Usually a fungal secondary metabolic gene cluster contains between one to two PKSs and five to twenty accessory enzymes, respectively. Table 1 lists a selection of HR-PKS pathways that often channel their products to downstream HR-PKSs, NR-PKSs, or non-ribosomal peptide synthetases for further modifications and production of fungal toxins, such as fumonisin<sup>53</sup>, important drugs like the cholesterol lowering agent lovastatin<sup>54</sup>, or the HIV integrase inhibitor Equisetin<sup>55, 56</sup> (Fig. 6B).

**Programming in HR-PKSs**—The name HR-PKS suggests a uniform substitution pattern of the product with all characteristics of fatty acids, however, their products are varied and

often highly complex. The fundamental difference to mFASs is an enigmatic programming pattern that defines which modifying domains are used in each cycle. Thus, HR-PKSs products can contain the full spectrum of all possible reduction states in one molecule, and even Diels-Alder cyclizations are known<sup>57</sup>.

Among the various patterns the most common is programmed methylation. Most HR-PKSs contain either an inactive  $\Psi$ CMeT or, in contrast to mFASs, a fully active SAM-dependent CMeT, which is capable of methylating the C $_{\alpha}$ -methylene. CMeT catalyzed methylation must precede  $\beta$ -carbonyl reduction by the KR (Fig. 7A). For the lovastatin PKS LovB, Cacho *et al.* have shown that the specific intermediate of elongation cycle three binds with exceptional specificity to the CMeT and thereby outcompetes the KR<sup>58</sup>. This programmed methylation event indicated a gatekeeping function, which is likely to result in off-loading reactions after deficient reaction cycles<sup>58</sup>. An equivalent function has also been observed for the citrinin producing NR-PKS PksCT, which relies on methylation for complete polyketide elongation and also indicated a contribution of the KS in programmed chain length control<sup>59</sup>. In other cases, such as the closely related HR-PKSs for the production of Desmethylbassianin (DMBS) and Tenellin (TENS) (Table 1), domain swapping experiments indicated chain length control by the KR<sup>60</sup>.

Special programming functions have also been observed for ERs. The *cis* acting ER of the squalenstatin HR-PKS (SQTKS) was shown to contribute to sequestering the final product and preventing further elongation<sup>61</sup>. In the lovastatin HR-PKS LovB, the inactive  $\Psi$ ER is functionally substituted by the *trans* acting LovC ER, which is only active in two out of eight elongation cycles.

In summary, current data on HR-PKSs suggest that programming is primarily controlled by kinetic competition of all enzymes with all ACP-bound intermediates, which describes the trajectory for selective domain usage in each elongation cycle<sup>58, 61</sup> (Fig. 7B).

**HR-PKSs in iterative assembly lines**—HR-PKSs are often found in clusters that contain several iterative PKS modules, which channel their substrate from one module to another (Fig. 8A-C). Such a channeling was shown first for the hexanoyl synthase (HexS)<sup>19, 62</sup>, which transfers its ACP-linked product directly to the starter-unit AT domain (SAT) of the NR-PKS PksA<sup>63</sup> (Fig. 8A) and produces the aromatic precursor to aflatoxin B<sub>1</sub>, a highly carcinogenic mycotoxin<sup>64</sup> (Fig. 6D). Similar transfer reactions have been found between the HR-PKS CazF and the SAT of the NR-PKS CazM in chaetoviridin and chaetomugilin natural product biosynthesis<sup>65</sup> (Fig. 8B). These data indicated that the starter-unit ATs (SATs) of NR-PKSs (see chapter VI.1) can serve a special adapter function and mediate substrate transfer from HR-PKSs to NR-PKSs<sup>65</sup>. This active transfer function resembles the role of passive docking domains in enabling vectorial substrate transfer in modPKSs (Fig. 8D)<sup>66, 67</sup> and non-ribosomal peptide synthetases (NRPSs)<sup>68</sup> and explains why many HR-PKS lack a C-terminal release domain. Some HR-PKSs even go one step further and occur as fused bimodules with a C-terminal NRPS module, such as the producer of equisetin (EquiS, Fig. 6B), an HIV-I integrase inhibitor<sup>55, 56</sup>, and cytochalasins (CheA), an actin filament binder<sup>55, 69, 70</sup> (Table 1, Fig. 8C). Notably, the well-studied lovastatin HR-PKS LovB contains a C-terminal NRPS condensing (C) domain<sup>54, 71</sup>, which is required for

efficient polyketide biosynthesis<sup>72</sup>, as well as a truncated adenylation (A) domain<sup>73</sup>, indicating a potential evolutionary relationship. Indirect transfer reactions have also been observed between two HR-PKSs in the lovastatin, compactin (MlcA, MlcB), and T-toxin (Pks1, Pks2) biosynthetic pathways (Table 1), which are commonly mediated by acyl-transferases.

**The architecture of HR-PKS**—HR-PKSs are related to mFASs with a sequence identity of approx. 20% and an iterative mode of catalysis. Apart from a single crystal structure of the *trans* acting LovC ER<sup>74</sup>, no specific structural information is available. However, the notable sequence identity (seq. id.), a homologous domain organization, and occurrence of (Ψ)CMeT domains suggest a similar architecture to mFAS<sup>30</sup>. The structural and functional conservation of all available condensing region structures from mFASs<sup>4, 30, 39</sup> modPKSs<sup>67, 75</sup> (except for PikAIII<sup>76</sup>), MAS-like PKS<sup>77</sup>, and NR-PKS<sup>78</sup> suggest a common architecture also in HR-PKSs.

The organization in the modifying region is likely to be particularly similar between HR-PKSs and mFASs<sup>30, 40</sup> too: While HR-PKSs and mFASs maintain degraded non-functional CMeT remnants (ΨCMeTs), NR-, MAS-like, and modPKSs mostly do not contain such an inactive domain. The conservation of mFAS architecture in HR-PKSs is further supported by the dimeric character of the *cis* acting ER domains of the squalestatin HR-PKS (SQTKS)<sup>61</sup>.

The integration of (Ψ)CMeT domains in HR-PKSs between the DH and ΨKR resembles mFAS, but low sequence conservation and deletions in inactive ΨCMeTs made detailed linking analysis difficult. For the same reason the mFAS ΨCMeT was initially annotated as unknown inter domain region two (ID2)<sup>37</sup> until higher resolution data became available<sup>30</sup>. After crystal structures of active CMeTs from the modPKS CurJ<sup>79</sup> and NR-PKS PksCT<sup>59</sup> were obtained, identification of ΨCMeT domain remnants has become more reliable.

The recent CMeT structures revealed that deletions mostly occur in the *N*-terminal subdomain, which is exposed to the lateral reaction cleft of the synthase (based on the mFAS structure), while the *C*-terminal subdomain is structurally more conserved. It remains unclear, however, why ΨCMeTs have not been totally deleted from mFASs. They may serve additional relevant functions *e.g.* in restricting ACP diffusion in the lateral reaction clefts.

The two structurally characterized CMeT domains are integrated into different sequence contexts in their full-length PKSs. CurJ CMeT shares the mFAS integration pattern, while the PksCT CMeT is inserted downstream of the ACP. Nevertheless, both CMeT structures are highly similar and exhibit a belt-helix in the *N*-terminal linker, which is not present in mFAS ΨCMeT. The linker ends in a different relative position, which might indicate a divergent CMeT organization of *cis*-AT modPKSs and NR-PKSs relative to mFAS. In summary, the organization of HR-PKS modifying regions remains unclear, but current data suggest a considerable similarity to mFAS.

## V.2. Mycobacterial mycocerosic acid synthase (MAS)-like PKSs

Mycobacteria are well known for their complex cell envelope and their huge repertoire of multienzymes including PKSs, NRPSs, and various FASs<sup>15, 46</sup>. These multienzymes

contribute to the biosynthesis of unusual cell wall lipids with very long and branched chain fatty acids<sup>46, 80</sup> as reviewed in detail by Quadri<sup>81</sup>. The prototype of mycobacterial PKSs is the mycocerosic acid synthase (MAS) from *Mycobacterium tuberculosis*, which produces mycocerosic acids. These branched chain fatty acids are the main component of dimycocerosates, which are involved in immune evasion<sup>82</sup>, contribute to cell wall impermeability<sup>45, 83</sup>, and determine lipid-specific replication of *M. tuberculosis* in lung tissue<sup>84</sup>. Due to the high sequence identity (up to ~60%) and functional similarity to the canonical MAS, related mycobacterial PKSs have been named MAS-like PKSs<sup>45</sup>.

While many HR-PKSs and mFASs use acetyl starter- and malonyl- extender units, MAS-like PKSs can use methylmalonyl extender units and larger starters, such as C<sub>12</sub>-C<sub>20</sub> fatty acids, which are presumably loaded by the AT domain or a member of the fatty acyl-AMP ligase family (FAAL)<sup>81, 85</sup>. Complex programming patterns, as described for fungal HR-PKSs, are generally not observed, although programming events cannot be entirely ruled out<sup>86</sup>. For MAS itself, a *trans*-acting polyketide associated acyltransferase (PapA5)<sup>87</sup> directly links the product mycocerosic acid to the product of the PpsA-E modPKSs to form phthiocerol or phenolphthiocerol lipids. Such fusion steps of two PKS products by *trans*-acting acyltransferase domains are common, have been proposed to be involved in related pathways<sup>88</sup>, and can also be found in the biosynthesis of Lovastatin (LovD)<sup>89</sup>.

Although MAS-like PKSs produce fatty acids and contain a fully reducing modifying region, they are distinct from mFAS and HR-PKSs. Unlike HR-PKSs and mFASs, MAS-like PKSs lack a (Ψ)CMeT domain and are more closely related to modPKSs (~27–35% seq. id.) than to HR-PKSs (20–22% seq. id.) or mFASs (~17–19% seq. id.)<sup>77</sup>, which is reflected in phylogenetic analysis of individual domains as well as larger multidomain regions<sup>77</sup> (Fig. 9). While mFASs may serve as a prototype for HR-PKS architecture, it is likely not the correct template for MAS-like or modPKSs, as confirmed by the hybrid structure of the MAS-like Pks5<sup>77</sup>, which is involved in lipooligosaccharide (LOS) biosynthesis<sup>86, 90</sup> in *M. smegmatis*. Due to its high sequence identity of 60%, Pks5 was initially annotated as MAS in databases; it is representative of the general architecture of MAS-like and modPKSs fully reducing modules.

**The architecture of MAS-like PKSs**—Key insights into the organization of MAS-like PKSs were recently provided by a hybrid crystal structure of the MAS-like Pks5<sup>77</sup>. Based on the conformational flexibility observed in mFAS around the central connection, overlapping condensing and modifying regions of MAS-like Pks5 were crystallized and re-connected *in silico* (Fig. 10A-C). The excised condensing region is in a monomer-dimer equilibrium in solution. The occurrence of a monomeric form likely is the result of reduced avidity arising from removal of linkage to the dimeric modifying region, and has similarly been observed for several condensing regions<sup>67, 76, 91</sup>.

The MAS-like PKS condensing region<sup>77</sup> revealed the same conserved architecture as all other mFAS<sup>30, 39</sup>, NR-PKS<sup>78</sup>, and modPKS<sup>67, 75</sup> condensing regions, which, however, differs from those observed in the modPKS PikAIII module<sup>76, 92</sup> at low resolution by cryo-EM. The high conservation of the condensing region (25–30% seq. id) is not surprising, because it catalyzes the universally conserved elongation step.

The architecture of the modifying region, a main evolutionary contributor to product diversity, is less conserved between mFASs and PKSs. Differences between mFASs and modPKSs modifying regions were already indicated by the first structure of a modPKS DH domain from module four of the 6-deoxyerythronolide B synthase (DEBS)<sup>93</sup>. While the DH domains in porcine mFAS form a V-shaped arrangement at an angle of 96° in the center of the modifying region (Fig. 5A, C), the DEBS DH and other PKS DHs<sup>93–95, 96</sup> (unpublished PDBs: 5J6O, 5IL5, 5IL6, 5HU7, 5HQW) or DH-like domains<sup>97</sup> are rather linear dimers at angles between 170–200°. Further support for an alternate PKS architecture arose from ΨKR/KR structures that revealed an N-terminal β-α-β-α extension of the ΨKR domain<sup>77, 98, 99, 100</sup> relative to mFAS<sup>30, 40, 101</sup>. Finally, an excised, non-naturally occurring ER-ΨKR/KR fragment of the Spinosyn modPKS SpnB, which lacked the DH domain, was crystallized as a domain-swapped monomer with a completely distinct disposition of the ER relative to the KR as compared to mFAS. This structure was the basis for the prediction of an alternative PKS modifying region architecture with a monomeric ER<sup>100</sup>.

The hybrid crystal structure of the MAS-like PKS modifying region<sup>77</sup> rationalizes most of the predicted mFAS-PKS differences, such as the ΨKR extension, the DH dimer angle, and helices in linker fragments of individual modPKSs domain structures, which connect the condensing region with the DH<sup>93, 96, 97</sup> as well as the DH<sup>94</sup> with the ΨKR/KR<sup>99</sup>. Relative to mFAS, the opening angle of the DH dimer is rather inverted at 222° (Fig. 10B) and resembles angles of recently determined *trans*-AT modPKS structures (unpublished PDBs: 5J6O, 5IL5, 5IL6, 5HU7, 5HQW).

However, the predicted PKS architecture based on the SpnB fragment with a monomeric ER was not reflected in the MAS-like PKS structure. The ER forms a similar dimer as in mFAS and rests on top of the DH platform with a small and highly variable interface. While ER dimers have been observed here and for the modPKS PpsC (PDB: 1PQW), some excised ER domains have been crystallized as monomers<sup>100, 102</sup>. Whether this difference is caused by reduced avidity due to the deletion of dimeric partner domains in the SpnB construct or indicative of an alternative architecture remains unclear. The ΨKR/KR is laterally attached to the DH and ER in the MAS-like PKS structural model, but, in contrast to mFAS, reveals a different tilt angle and has no interface with neighboring domains.

In fact, the entire modifying region architecture in MAS-like PKSs is based on linker-mediated interactions rather than on direct domain-domain interfaces. The DH connects the ΨKR via a linker containing a spacer helix, interactions with the ΨKR/KR surface and the ER-KR linker, and a small disordered segment that is probably used for CMeT integration in homologous PKSs. The ER is connected to the ΨKR via linkers that interact with the ER-KR linker. Thus, the ΨKR/KR is connected via a double tethered hinge with its two partner domains.

Such a linker-mediated architecture provides a framework for evolutionary domain shuffling without a requirement for adjustment of direct domain interfaces; it rationalizes the evolutionary modularity and variability of modifying regions in PKS.

The observation of nine independent dimers in the crystallographic asymmetric unit of the MAS-like Pks5 modifying region revealed conformational variability and coupling of the lateral ΨKR/KR domain hinge-bending (up to 40° rotation) via a screw motion of the ER dimer (translation of 8.5 Å) on the DH platform (Fig. 10D). This coupling creates an entirely dynamic architecture, which affords additional room to compensate for ACP linker constraints and might provide a mechanism for coordinated asymmetric polyketide biosynthesis in the two lateral MAS-like PKS reaction clefts. The condensing region presumably also contributes to the overall dynamics of MAS-like PKSs, because a comparison with homologous structures indicates a hinge in the LD. Conformational coupling has also been observed for mFAS at low resolution<sup>38</sup>, which supports the general relevance of a dynamic architecture, rather than a static scaffold, in ACP-dependent multienzymes. Notably, in MAS-like Pks5, domain flexibility increases laterally with the ΨKR/KR domain being the most flexible catalytic domain and the ACP the most flexible non-catalytic part.

While phylogenetic data indicate that the architecture of mFAS is representative for HR-PKSs, the MAS-like architecture appears to be more related to modPKSs. The SpnB fragment-based prediction of a modPKS architecture with outward facing monomeric ERs<sup>100</sup> was supported by differences in mFAS and modPKS linkers. Also the MAS-like Pks5 reveals a relatively long ER-KR linker of 20 aa, while many, but not all, modPKSs have short linkers (5–10 aa) as well as low sequence conservation in the putative ER dimerization interface; however, the short linkers are tolerated in a MAS-like Pks5-based modPKS model.

In order to experimentally validate a MAS-like PKS based architecture of modPKS, two fully reducing modPKS modifying regions with short ER-KR linkers and low sequence conservation in the ER dimerization interface were analyzed by small angle X-ray scattering (SAXS), which is suited to distinguish the two principal architectures. Both modPKS modifying regions conformed to the MAS-like architecture, supporting a general relevance of this architecture in modPKSs. However, the existence of an alternative architecture in other modPKSs cannot be completely ruled out.

Despite some HR-PKS-NRPS hybrids as well as HR-PKSs that pass their substrates on to NR-PKSs, most iterative PKS lack integration into an assembly line that imposes vectorial transport of intermediates. In modPKSs, the shuttling direction is imposed by ACP–KS recognition<sup>52, 103, 104</sup> as well as kinetic mechanisms<sup>105</sup>. In some cases, modPKS modules, such as BorA5<sup>51</sup> (Fig. 6F), were also found to act/function iteratively (Fig. 8D). Additionally, modPKSs modules such as PikAIII, can be switched to an iterative mode of action when isolated from the assembly line<sup>106</sup> or by engineering ACP recognition<sup>104</sup>.

Altogether, the structural data on MAS-like PKSs and the functional intermixing of iPKSs and modPKSs create close links between members of the two families. MAS-like iPKSs may actually be viewed as an iterative version of a modPKS module.



## VI. Non-Reducing Iterative Polyketide Synthases

Although aromatic metabolites isolated from fungi are among the earliest characterized natural products, the nature and mechanisms of the underlying polyketide synthases remained a mystery and lagged behind the rapid advances that occurred after 1990<sup>2, 107</sup> in understanding modPKSs<sup>108</sup>. In no small measure this comparative ignorance owed not only to technical barriers to experiment but also to the long-held belief based on sequence comparisons to mFASs that the NR-PKSs contained four catalytic domains, a KS, a malonyl-specific AT (MAT; not to be confused with the dual-specific mFAS MAT), ACP and TE. An extended *N*-terminus and internal region were thought to be dimerization interfaces to support  $\alpha_2$ -structures akin to mFASs. A quite different picture emerged, however, when the predictive algorithm UMA was applied to PksA, the NR-PKS that lies at the heart of aflatoxin biosynthesis, and the limited number of other NR-PKS sequences known at the time (Fig. 11)<sup>109</sup>. This bioinformatics tool can be used to analyse a set of related protein sequences to identify linker regions between presumably compactly folded domains. While the KS, MAT, ACP and TE were correctly discerned, the elongated *N*-terminus and the previously undefined internal region both appeared to be well-structured.

Unique to the NR-PKS family, the *N*-terminus was identified as a starter-unit acyltransferase (or transacylase) SAT domain, which loads a primer for polyketide extension and accounts for the classical “starter unit effect”, a phenomenon long known for the anomalous incorporation level of radiolabeled precursors at this site<sup>20, 110</sup>. Strikingly absent are the three reduction/dehydration domains of FASs or HR-PKSs. Highly reactive poly- $\beta$ -ketone chains are synthesized, but controlled in length and cyclization behaviour. The internal product template (PT) domain dictates a defined cyclization pattern, as detailed below, of the polyketide intermediate to particular monocyclic or fused ring systems. The coupled loss of water and resonance stabilization of the aromatic ring(s) generated provide enormous driving force to product.

### VI. 1 The starter-unit acyltransferase (SAT) domain.

Using the simple, original phylogenetic designations of Kroken<sup>111</sup> Clade I and II NR-PKSs have the domain organization depicted in Figure 11A, although two, even three, ACP domains in row can occur<sup>112</sup>. Members of the related Clade III NR-PKSs carry out *C*-methylation(s) coupled to polyketide elongation<sup>59</sup>, and can terminate in either a TE or NADPH-dependent reduction (R) domain to release the product typically as an aldehyde<sup>113</sup>. *C*-methylation does not take place in every cycle of polyketide extension, but at “programmed” sites/intervals determined primarily by the CMeT domain itself<sup>59</sup>.

In most NR-PKSs (Fig. 11A, path **a**) the SAT domain is modestly selective for acetyl-CoA, a fundamental metabolic building block of the cell<sup>114</sup>. There are many examples (Fig. 11B) beginning with molecules as simple as orsellinic acid to the tricyclic structure of *nor*-toralactone in which the formation of the terminal pyrone is catalyzed by the TE domains<sup>115</sup>. This cyclization is an exception among TE-catalyzed reactions, but sets the pathway in motion to the biosynthesis of the widespread phytotoxin cercosporin<sup>115</sup>.

The substrate adaptability of the SAT domain, however, considerably expands the synthetic capabilities of NR-PKSs. For example, PksA accepts a hexanoyl starter unit from a specialized yFAS pair of HexA and HexB, together also referred to as HexS, (Fig. 11A, path **b**) and carries out seven malonyl extensions to give the C<sub>20</sub>-anthraquinone norsolorinic acid, the first committed intermediate in a ~15 step reaction sequence to the environmental toxin aflatoxin B<sub>1</sub><sup>116</sup>. With an eye to redirecting synthesis to non-natural products, the substrate tolerance of the PksA SAT domain has been tested and >20 alternative starter units were successfully processed by the SAT–KS–MAT + PT–ACP to correctly cyclized products<sup>117</sup>.

Finally, a HR-PKS (Fig. 11A, path **c**) can be functionally linked to a NR-PKS to generate, for example, the extended primer shown bolded in the resorcylic acid lactone (*R*)-monocillin II, the precursor of radicicol, a potent inhibitor of heat shock protein 90 (Hsp90)<sup>118</sup>. Similarly, a HR-PKS creates the (*S*, *E*)-4-methyl-hex-2-enoyl triketide starter unit in chaetoviridin A (bolded carbons)<sup>65</sup>.

The SAT domain of the Clade III NR-PKS CazM transfers this initiating component into the rest of the biosynthetic machinery. A crystal structure of the isolated CazM SAT domain has been obtained with a simpler hexanoyl surrogate of the native starter bound to the active site cysteine<sup>65</sup>. The isolated CazM SAT closely resembles the general fold of PKS AT domains (Fig. 11C). These proteins share an  $\alpha/\beta$ -hydrolase fold with a small ferredoxin-like subdomain forming a palm and thumb motif with the active site at the base of a cleft between the two. The active sites are largely hydrophobic corresponding to the generally hydrophobic nature of starter unit substrates. This SAT domain and the related SAT domain of PksA serve as adapter proteins recognized *in trans* by a donor ACP from a HR-PKS or FAS and an *in cis* ACP acceptor through a single-access transfer site<sup>63</sup>.

## VI.2 Product template (PT) domain.

Transfer of the starter unit from the SAT to the KS domain initiates polymerization of successive malonyl units to form a polyketide intermediate whose programmed length is dictated mainly by the KS<sup>119–121</sup>. PT domains are known that specifically catalyze C2–C7, C4–C9 and C6–C11 first-ring cyclizations as shown schematically and exemplified in Figure 12<sup>122</sup>. Guided by UMA, it has been possible to dissect various NR-PKS proteins to give a library of mono-, di- and tridomain fragments. Reconstitution of native components has faithfully restored *wild-type* synthetic activity<sup>20, 114, 119, 123</sup>. Similarly, heterocombinations of these domains have enabled synthesis of non-native products and, in so doing, given a detailed picture of the division of labor shared among the domains and the extent of control each exerts on product outcome<sup>121</sup>.

Analysis of the catalytic cycle of PksA by mass spectrometry revealed that the time-average population of species bound to the ACP was dominated by building blocks (hexanoyl, malonyl and acetyl, presumably arising from decarboxylation of malonyl units) and lower intensity mass signatures in keeping with the fully extended polyketide intermediate and two successive dehydrations (cyclizations) enroute to product release (Fig. 13A). Importantly no intermediate chain lengths were detectable despite the introduction of mutations intended to slow polyketide extension<sup>119, 124</sup>. We attribute these observations to rapid, processive extension and, in contrast to the behavior of Clade III NR-PKSs, the extending chain

probably never leaves the KS, shuttling rapidly back and forth between the active site Cys and transient reactions with malonyl-ACP (Fig. 13A).

For Clade III enzymes polyketides of intermediate length leave the KS to the CMeT in a programmed manner for  $\alpha$ -methylation, which gives fleeting exposure of this reactive ACP-bound species outside of either protein. The opportunity presents in reconstitution experiments for spontaneous intramolecular pyrone formation and the appearance of derailed, immature side products in competition with return to the KS and completion of programmed polyketide elongation<sup>59</sup>. This contrasting behavior is illustrated for PksCT, which lies at the beginning of citrinin biosynthesis (Fig. 13B)<sup>59</sup>.

The X-ray structure of only one PT domain is known, that for PksA. While at the time bioinformatics tools did not recognize the evolutionary relationship of PT and DH domains, the 1.8 Å crystal structure showed the PksA PT as a homodimer of unmistakable double “hot dog” fold monomers<sup>125</sup> (Fig. 14A). Close to the dimer interface is the entrance to a deep active site pocket extending 30 Å into the interior. A fairly linear 14 Å channel binds the Ppant of ACP to deliver the C<sub>20</sub>-polyketide intermediate inward to an 8 × 13.5 Å reaction chamber housing a His–Asp catalytic dyad seen in DH domains (Fig. 14B). The deepest extent of the active site is then reached by a largely hydrophobic tunnel, which was designated the “hexyl binding” region to accommodate the hexyl starter of the polyketide intermediate. In support for this picture was the fortuitous co-crystallization of palmitate with its hydrocarbon tail deep in this pocket and arrayed in a fully extended conformation to place the carboxylate in the reaction chamber proximal to, but not binding with, the His–Asp dyad (Fig. 14C). Using the well-resolved C<sub>16</sub> chain of palmitate as a template, the C<sub>20</sub>-polyketide intermediate was overlaid and docked in the active site such that C4 was constrained to be nearby the catalytic His. The energy minimized structure that emerged revealed dual hydrogen bonds specifically to the C9 carbonyl from crystallographic (static) water molecules on the “wet” side of the reaction chamber. This unanticipated outcome of the computational model gave sudden insight into how PT domains might control the pluripotent reactivity of an octa-carbonyl intermediate and guide it to a *single* cyclization product<sup>125</sup>. Generation of a nucleophilic anion or enol at C4 by the catalytic His can now partner with the C9 carbonyl, which has been specifically activated by dual hydrogen bonds<sup>126</sup> as the most reactive electrophile for intramolecular aldol reaction (Fig. 14D). Dehydration, the primordial DH function, seals first-ring formation and aromatization. The PT cyclization chamber is large enough to accommodate a second ring. One can visualize that after formation of the first ring, C2 is drawn inward to the catalytic His and the C11 carbonyl is correspondingly displaced and activated by the water network. Analogous aldol ring closure takes place to the trihydroxynaphthalene, which had been detected in earlier experiments by mass spectrometry and UV-vis spectroscopic comparison<sup>119</sup>.

In support of this proposal for PT function, a bisisoxazole model of the experimentally inaccessible PksA polyketide intermediate was synthesized with thioether linkers and *N,O*-heteroatoms intended to mimic the alternating pattern of keto/enol sites in the native substrate. A 1.8 Å structure was obtained (Fig. 14E), which gratifyingly revealed the well-resolved mimic in an extended conformation where the Ppant stretched from the terminal phosphate interaction with a PT surface Arg to deliver the critical C4 methylene precisely to

the catalytic His, in keeping with the computational model. The short hydrophobic tail is positioned as predicted in the “hexyl binding” region<sup>127</sup>. The rigid heterocycles and unavoidable 90° C–S–C bond angles of the bridging thioethers give a congested, highly bent presentation of the carbonyl surrogates in the reaction chamber. While the mimetic accuracy of their interactions with the ordered active site water molecules is uncertain, in sum, the structure is a major advance and provides important experimental support for the *in silico* model of PT function proposed above<sup>125</sup>.

### VI.3 Thioesterase (TE) domain

With no further ring formation possible in the PT cyclization chamber, the significantly more stable naphthaloyl aldol product is transferred to the Ser of the catalytic triad of the TE domain, where it undergoes not hydrolysis to the carboxylic acid, but a final Dieckmann cyclization and concomitant release of norsolorinic acid anthrone, which spontaneously oxidizes to the corresponding anthraquinone (Fig. 15A). This interesting variant of TE behavior is commonly seen among fungal NR-PKSs<sup>128</sup>. The overall fold is of an  $\alpha/\beta$ -hydrolase in which the active site is closed at each end, but a large lid region opens from the side to admit the substrate and release the ring-closed product most like the macrocycling NRPS TE domains<sup>129</sup> (Fig. 15B). It has been proposed that as the ACP-Ppant departs the TE active site after transacylation, the thiolate removes the triad histidinium proton (Fig. 15A) to create a base for the final condensation. However, for the large TE lid to close over the substrate, the  $\beta$ -diketone of the substrate rotates into the ACP-Ppant binding channel to be oriented now proximal to the His for final Dieckmann reaction<sup>130</sup>.

### VI.4 Substrate loading architecture of NR-PKSs

Enormous progress has been made in the last dozen years toward understanding the iterative catalytic cycle of NR-PKS mega-enzymes. The discovery and functional characterization of SAT and PT domains, which are unique to this family of PKSs, made it possible to assign individual synthetic tasks to all domains. Structural studies of individual domains have greatly amplified detailed understanding of key mechanistic steps. While some sense of the overall orchestration of iterative synthesis has been gained from these experiments, it has been an underlying goal for more than a decade since the mFAS and yFAS structures were reported<sup>5, 37</sup> to obtain comparable structural information about iterative PKS enzymes. Recently a major step in that direction has been made with a 2.8 Å crystal structure of the dimeric CTB1 loading/condensing region (SAT–KS–MAT) together with trapping of the KS in its substrate-loading state by mechanism-based chemical crosslinking to the ACP and analysis by cryo-EM at 7.1 Å resolution<sup>78</sup>.

The crystal structure reveals a symmetrical, compact dimer of overall rhomboidal shape (Fig. 16). The integrated SAT domain closely resembles the CazM SAT and has *N*- and *C*-termini on opposite sides as oriented along the central dimer axis as in modPKSs. The two CTB1 SAT domains reveal an intertwined integration, where each SAT interfaces the KS–MAT condensing region of the neighboring protomer. The SATs form a weak dimer and are held in place by extensive interactions with the MATs and stabilization via an extensive SAT–KS linker, which binds to the surface of each neighboring SAT (Fig. 16B). As compared to the monomeric character of the excised SAT of CazM this represents another

example for the relevance of avidity in multienzymes for oligomerization (see above). The condensing region (KS-MAT) reveals the same conserved organization as observed in other PKSs and mFASs<sup>4, 30, 39, 67, 75, 77, 131</sup>, except for PikAIII<sup>76</sup>. The cavity formed by the S-shaped condensing region (view along the dimeric axis) is closed at the bottom by the SATs. This arrangement forms a surface channel that connects the active site entrances to the SAT and KS active sites and might guide ACP translocation.

Structural and functional analysis of productive ACP interactions with its target in most cases require a mechanism-based crosslinking approach owing to the transient nature of this interaction. This approach was developed in the Burkart laboratory<sup>132</sup> and has become indispensable for functional and structural studies on carrier protein-based molecular machines. So far, this approach has been employed for structural studies of the bacterial type II FAS DH<sup>133</sup> and for the NR-PKS CTB1 loading/condensing region (SAT-KS-MAT)<sup>78</sup>.

Upon crosslinking of the *holo*-ACP to the KS of the CTB1 loading/condensing region, conformational changes are evident that propagate throughout the dimeric structure. While crosslinking was carried out to >90% at both KS and KS', the ACP is plainly resolved at only one KS but completely disordered at the other (Fig. 17A). Both SAT domains move toward the central KS•KS' and the KS-MAT linker domain (LD) and MAT itself shift upward 4.4 and 8.4 Å, respectively. The high asymmetry of the domain movements gives a dramatic glimpse of an important step in the iterative catalytic cycle where C<sub>2</sub>-extension occurs. It would appear that coordinated conformational changes block reaction at KS' on the companion protomer. Such a view is in accord with the kinetics of analogous tridomain crosslinking experiments using two other NR-PKS systems, but where a 1:2 stoichiometry (ACP:KS) was observed as rapidly as for CTB1. To achieve a full 1:1 ratio required excess ACP and prolonged reaction times<sup>78, 134</sup>. These results indicate a mechanism that mediates asymmetric coupling of polyketide biosynthesis in the two reaction clefts of the dimeric synthase (Fig. 17B, C). Notably, conformational asymmetry and coupling has also been observed during the reaction cycle of mFAS<sup>38</sup> and for the modifying region of the MAS-like Pks5<sup>77</sup>, pointing towards a general biosynthetic strategy.

## VII. Conclusion

For a decade, high-resolution structural data have been available for the two distinct types of FAS multienzymes, yFAS<sup>6</sup> and mFAS<sup>30</sup>. These data provided a new foundation for conceptual understanding of the architecture, evolution and function of the FAS and PKS families of ACP-containing multienzymes.

For yFAS, the structural insights revealed the previously enigmatic route of evolution to yeast FAS from early single-functional ancestors in bacterial primary and secondary metabolism via bacterial yFAS<sup>17, 18</sup>. Structure-guided re-engineering of yFAS has most recently succeeded in creating novel reaction modules for the synthesis of new product families<sup>24, 25</sup>.

For mFAS, the structural model of the core elongation domains has enabled rational dissection into crystallizable fragments<sup>39, 40, 135</sup>. These fragments have successfully been

employed in screening small-molecule FAS inhibitors as potential leads against cancer and metabolic disorders. mFAS continues to serve as a test-bed for functional studies of ACP-containing multienzymes, as exemplified by attempts towards analysing mFAS dynamics at the single molecule level via fluorescence methods<sup>136</sup> or high-speed atomic force microscopy<sup>41</sup>.

Structural insights into PKS architectures massively lagged behind studies on mFAS, despite their related domain organization. Meanwhile, different conceptual models have been discussed based on structures of isolated domains or limited size fragments of PKS. The intermediate resolution visualization of a minimal reducing modPKS module, PikAIII, in different conformational states served as a first example of PKS organization<sup>76, 137</sup> until the hybrid structural model of a MAS-like PKS was obtained with demonstrated relevance also for modPKSs<sup>77</sup>. These structures depict different conformations and states, with additional data supporting the functional relevance of either of the structures.

A key aim of ongoing structural biology work on PKS is to conclusively reveal the principal architecture of different classes of PKSs as well as the ensemble of conformational states contributing to catalytic efficiency. The available data have reinforced the notion that overall conformational dynamics, and linker-mediated interactions are crucial for PKS function.

Obtaining a “blueprint” of the organization of different types of PKSs remains a key aim in natural product research. Site specific crosslinking of ACPs to their catalytic partner domains has proven to provide crucial “snapshots” of catalytic states not seen before. They are of fundamental value to visualize and understand the dynamic, coordinated function of sophisticated PKS molecular machines and will guide their reprogramming to new synthetic objectives.

## Acknowledgements

This work has been supported by the Swiss National Science Foundation grants 159696 and 179323 to TM, DAH has been supported by a Fellowship for Excellence from the Werner - Siemens Foundation. CAT gratefully acknowledges support of the National Institutes of Health grant RO1 ES001670.

## Notes and references

1. Maier T, Leibundgut M, Boehringer D and Ban N, Q. Rev. Biophys, 2010, 43, 373–422. [PubMed: 20731893]
2. Cortes J, Haydock SF, Roberts GA, Bevirt DJ and Leadlay PF, Nature, 1990, 348, 176–178. [PubMed: 2234082]
3. Leibundgut M, Maier T, Jenni S and Ban N, Curr. Opin. Struct. Biol, 2008, 18, 714–725. [PubMed: 18948193]
4. Rittner A, Paithankar KS, Huu KV and Grninger M, ACS Chem. Biol, 2018, 13, 723–732. [PubMed: 29328619]
5. Jenni S, Leibundgut M, Maier T and Ban N, Science, 2006, 311, 1263–1267. [PubMed: 16513976]
6. Jenni S, Leibundgut M, Boehringer D, Frick C, Mikolasek B and Ban N, Science, 2007, 316, 254–261. [PubMed: 17431175]
7. Leibundgut M, Jenni S, Frick C and Ban N, Science, 2007, 316, 288–290. [PubMed: 17431182]
8. Lomakin IB, Xiong Y and Steitz TA, Cell, 2007, 129, 319–332. [PubMed: 17448991]
9. Johansson P, Wiltschi B, Kumari P, Kessler B, Vonrhein C, Vonck J, Oesterhelt D and Grninger M, Proc. Natl. Acad. Sci. U. S. A, 2008, 105, 12803–12808. [PubMed: 18725634]



10. Johansson P, Mulinacci B, Koestler C, Vollrath R, Oesterhelt D and Grininger M, *Structure*, 2009, 17, 1063–1074. [PubMed: 19679086]
11. Gipson P, Mills DJ, Wouts R, Grininger M, Vonck J and Kuhlbrandt W, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, 107, 9164–9169. [PubMed: 20231485]
12. Perez DR, Leibundgut M and Wider G, *Biochemistry*, 2015, 54, 2205–2213. [PubMed: 25774789]
13. Sumper M, Riepertinger C and Lynen F, *FEBS Lett*, 1969, 5, 45–49; [PubMed: 11947235]  
Gajewski J, Buelens F, Serdjukow S, Janssen M, Cortina N, Grubmuller H and Grininger M, *Nat. Chem. Biol*, 2017, 13, 363–365. [PubMed: 28218912]
14. Brindley DN, Matsumur S and Bloch K, *Nature*, 1969, 224, 666–669.
15. Schweizer E and Hofmann J, *Microbiol. Mol. Biol. Rev.*, 2004, 68, 501–517. [PubMed: 15353567]
16. Gago G, Diacovich L, Arabolaza A, Tsai SC and Gramajo H, *FEMS Microbiol. Rev.*, 2011, 35, 475–497. [PubMed: 21204864]
17. Bukhari HST, Jakob RP and Maier T, *Structure*, 2014, 22, 1775–1785. [PubMed: 25456814]
18. Grininger M, *Curr. Opin. Struct. Biol.*, 2014, 25, 49–56. [PubMed: 24457260]
19. Brown DW, Adams TH and Keller NP, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, 93, 14873–14877. [PubMed: 8962148]
20. Crawford JM, Dancy BCR, Hill EA, Udway DW and Townsend CA, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, 103, 16728–16733. [PubMed: 17071746]
21. Boehringer D, Ban N and Leibundgut M, *J. Mol. Biol.*, 2013, 425, 841–849. [PubMed: 23291528]
22. Ciccarelli L, Connell SR, Enderle M, Mills DJ, Vonck J and Grininger M, *Structure*, 2013, 21, 1251–1257. [PubMed: 23746808]
23. Fischer M, Rhinow D, Zhu Z, Mills DJ, Zhao ZK, Vonck J and Grininger M, *Protein Sci*, 2015, 24, 987–995. [PubMed: 25761671]
24. Gajewski J, Pavlovic R, Fischer M, Boles E and Grininger M, *Nat. Commun*, 2017, 8, 14650. [PubMed: 28281527]
25. Zhu Z, Zhou YJ, Krivoruchko A, Grininger M, Zhao ZK and Nielsen J, *Nat. Chem. Biol*, 2017, 13, 360–362. [PubMed: 28218911]
26. Saito J, Yamada M, Watanabe T, Iida M, Kitagawa H, Takahata S, Ozawa T, Takeuchi Y and Ohsawa F, *Protein Sci*, 2008, 17, 691–699. [PubMed: 18305197]
27. Witkowski A, Ghosal A, Joshi AK, Witkowska HE, Asturias FJ and Smith S, *Chem. Biol*, 2004, 11, 1667–1676. [PubMed: 15610851]
28. Smith S, Witkowski A and Joshi AK, *Prog. Lipid Res*, 2003, 42, 289–317. [PubMed: 12689621]
29. Brink J, Ludtke SJ, Yang CY, Gu ZW, Wakil SJ and Chiu W, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, 99, 138–143. [PubMed: 11756679]
30. Maier T, Leibundgut M and Ban N, *Science*, 2008, 321, 1315–1322. [PubMed: 18772430]
31. Ploskon E, Arthur CJ, Evans SE, Williams C, Crosby J, Simpson TJ and Crump MP, *J. Biol. Chem.*, 2008, 283, 518–528. [PubMed: 17971456]
32. Chakravarty B, Gu Z, Chirala SS, Wakil SJ and Quioco FA, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, 101, 15567–15572. [PubMed: 15507492]
33. Pemble C. W. t., Johnson LC, Kridel SJ and Lowther WT, *Nat. Struct. Mol. Biol*, 2007, 14, 704–709. [PubMed: 17618296]
34. Joshi AK, Witkowski A, Berman HA, Zhang L and Smith S, *Biochemistry*, 2005, 44, 4100–4107. [PubMed: 15751987]
35. Ritchie MK, Johnson LC, Clodfelter JE, Pemble C. W. t., Fulp BE, Furdui CM, Kridel SJ and Lowther WT, *J. Biol. Chem.*, 2016, 291, 3520–3530. [PubMed: 26663084]
36. Joshi AK, Rangan VS, Witkowski A and Smith S, *Chem. Biol*, 2003, 10, 169–173. [PubMed: 12618189]
37. Maier T, Jenni S and Ban N, *Science*, 2006, 311, 1258–1262. [PubMed: 16513975]
38. Brignole EJ, Smith S and Asturias FJ, *Nat. Struct. Mol. Biol*, 2009, 16, 190–197. [PubMed: 19151726]
39. Pappenberger G, Benz J, Gsell B, Hennig M, Ruf A, Stihle M, Thoma R and Rudolph MG, *J. Mol. Biol*, 2010, 397, 508–519. [PubMed: 20132826]

40. Hardwicke MA, Rendina AR, Williams SP, Moore ML, Wang L, Krueger JA, Plant RN, Totteritis RD, Zhang G, Briand J, Burkhart WA, Brown KK and Parrish CA, *Nat. Chem. Biol.*, 2014, 10, 774–779. [PubMed: 25086508]
41. Benning FMC, Sakiyama Y, Mazur A, Bukhari HST, Lim RYH and Maier T, *ACS Nano*, 2017, 11, 10852–10859. [PubMed: 29023094]
42. Hertweck C, *Angew. Chem. Int. Ed. Engl.*, 2009, 48, 4688–4716; [PubMed: 19514004] Smith S and Tsai SC, *Nat. Prod. Rep.*, 2007, 24, 1041–1072. [PubMed: 17898897]
43. Weissman KJ, *Nat. Chem. Biol.*, 2015, 11, 660–670; [PubMed: 26284673] Weissman KJ, *Nat. Prod. Rep.*, 2015, 32, 436–453. [PubMed: 25310997]
44. Bretschneider T, Zocher G, Unger M, Scherlach K, Stehle T and Hertweck C, *Nat. Chem. Biol.*, 2011, 8, 154–161. [PubMed: 22179067]
45. Sirakova TD, Thirumala AK, Dubey VS, Sprecher H and Kolattukudy PE, *J. Biol. Chem.*, 2001, 276, 16833–16839. [PubMed: 11278910]
46. Gokhale RS, Saxena P, Chopra T and Mohanty D, *Nat. Prod. Rep.*, 2007, 24, 267–277. [PubMed: 17389997]
47. Weissman KJ, *Chembiochem*, 2010, 11, 485–488. [PubMed: 20127928]
48. Van Lanen SG, Lin S and Shen B, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, 105, 494–499; [PubMed: 18182490] Liu W, Christenson SD, Standage S and Shen B, *Science*, 2002, 297, 1170–1173. [PubMed: 12183628]
49. Blodgett JA, Oh DC, Cao S, Currie CR, Kolter R and Clardy J, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, 107, 11692–11697; [PubMed: 20547882] Li Y, Chen H, Ding Y, Xie Y, Wang H, Cerny RL, Shen Y and Du L, *Angew. Chem. Int. Ed. Engl.*, 2014, 53, 7524–7530. [PubMed: 24890524]
50. Kaulmann U and Hertweck C, *Angew. Chem. Int. Ed. Engl.*, 2002, 41, 1866–1869. [PubMed: 19750617]
51. Olano C, Wilkinson B, Sanchez C, Moss SJ, Sheridan R, Math V, Weston AJ, Brana AF, Martin CJ, Oliynyk M, Mendez C, Leadlay PF and Salas JA, *Chem. Biol.*, 2004, 11, 87–97. [PubMed: 15112998]
52. Moss SJ, Martin CJ and Wilkinson B, *Nat. Prod. Rep.*, 2004, 21, 575–593. [PubMed: 15459756]
53. Proctor RH, Brown DW, Plattner RD and Desjardins AE, *Fungal Genet. Biol.*, 2003, 38, 237–249. [PubMed: 12620260]
54. Hendrickson L, Davis CR, Roach C, Nguyen DK, Aldrich T, McAda PC and Reeves CD, *Chem. Biol.*, 1999, 6, 429–439. [PubMed: 10381407]
55. Campbell CD and Vederas JC, *Biopolymers*, 2010, 93, 755–763. [PubMed: 20577995]
56. Singh SB, Zink DL, Goetz MA, Dombrowski AW, Polishook JD and Hazuda DJ, *Tetrahedron Lett.*, 1998, 39, 2243–2246.
57. Auclair K, Sutherland A, Kennedy J, Witter DJ, Van den Heever JP, Hutchinson CR and Vederas JC, *J. Am. Chem. Soc.*, 2000, 122, 11519–11520; Kelly WL, *Org. Biomol. Chem.*, 2008, 6, 4483–4493. [PubMed: 19039353]
58. Cacho RA, Thuss J, Xu W, Sanichar R, Gao Z, Nguyen A, Vederas JC and Tang Y, *J. Am. Chem. Soc.*, 2015, 137, 15688–15691. [PubMed: 26630357]
59. Storm PA, Herbst DA, Maier T and Townsend CA, *Cell. Chem. Biol.*, 2017, 24, 316–325. [PubMed: 28238725]
60. Fisch KM, Bakeer W, Yakasai AA, Song ZS, Pedrick J, Wasil Z, Bailey AM, Lazarus CM, Simpson TJ and Cox RJ, *J. Am. Chem. Soc.*, 2011, 133, 16635–16641. [PubMed: 21899331]
61. Roberts DM, Bartel C, Scott A, Ivison D, Simpson TJ and Cox RJ, *Chem. Sci.*, 2017, 8, 1116–1126. [PubMed: 28451251]
62. Hitchman TS, Schmidt EW, Trail F, Rarick MD, Linz JE and Townsend CA, *Bioorg. Chem.*, 2001, 29, 293–307. [PubMed: 16256699]
63. Foulke-Abel J and Townsend CA, *Chembiochem*, 2012, 13, 1880–1884. [PubMed: 22807303]
64. Bedard LL and Massey TE, *Cancer Lett.*, 2006, 241, 174–183. [PubMed: 16458422]
65. Winter JM, Cascio D, Dietrich D, Sato M, Watanabe K, Sawaya MR, Vederas JC and Tang Y, *J. Am. Chem. Soc.*, 2015, 137, 9885–9893. [PubMed: 26172141]

66. Broadhurst RW, Nietlispach D, Wheatcroft MP, Leadlay PF and Weissman KJ, *Chem. Biol.*, 2003, 10, 723–731; [PubMed: 12954331] Buchholz TJ, Geders TW, Bartley FE, 3rd, Reynolds KA, Smith JL and Sherman DH, *ACS Chem. Biol.*, 2009, 4, 41–52; [PubMed: 19146481] Dorival J, Annaval T, Risser F, Collin S, Roblin P, Jacob C, Gruez A, Chagot B and Weissman KJ, *J. Am. Chem. Soc.*, 2016, 138, 4155–4167; [PubMed: 26982529] Zeng J, Wagner DT, Zhang Z, Moretto L, Addison JD and Keatinge-Clay AT, *ACS Chem. Biol.*, 2016, 11, 2466–2474; [PubMed: 27362945] Richter CD, Nietlispach D, Broadhurst RW and Weissman KJ, *Nat. Chem. Biol.*, 2008, 4, 75–81. [PubMed: 18066054]
67. Whicher JR, Smaga SS, Hansen DA, Brown WC, Gerwick WH, Sherman DH and Smith JL, *Chem. Biol.*, 2013, 20, 1340–1351. [PubMed: 24183970]
68. Dowling DP, Kung Y, Croft AK, Taghizadeh K, Kelly WL, Walsh CT and Drennan CL, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, 113, 12432–12437. [PubMed: 27791103]
69. Schumann J and Hertweck C, *J. Am. Chem. Soc.*, 2007, 129, 9564–9565. [PubMed: 17636916]
70. Jiao WX, Feng YJ, Blunt JW, Cole ALJ and Munro MHG, *J. Nat. Prod.*, 2004, 67, 1722–1725. [PubMed: 15497948]
71. Kennedy J, Auclair K, Kendrew SG, Park C, Vederas JC and Hutchinson CR, *Science*, 1999, 284, 1368–1372. [PubMed: 10334994]
72. Ma SM, Li JWH, Choi JW, Zhou H, Lee KKM, Moorthie VA, Xie XK, Kealey JT, Da Silva NA, Vederas JC and Tang Y, *Science*, 2009, 326, 589–592. [PubMed: 19900898]
73. Sims JW, Fillmore JP, Warner DD and Schmidt EW, *Chem. Commun. (Camb.)*, 2005, 0, 186–188.
74. Ames BD, Nguyen C, Bruegger J, Smith P, Xu W, Ma S, Wong E, Wong S, Xie X, Li JW, Vederas JC, Tang Y and Tsai SC, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, 109, 11144–11149. [PubMed: 22733743]
75. Tang Y, Chen AY, Kim CY, Cane DE and Khosla C, *Chem. Biol.*, 2007, 14, 931–943; [PubMed: 17719492] Tang Y, Kim CY, Mathews II, Cane DE and Khosla C, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, 103, 11124–11129. [PubMed: 16844787]
76. Dutta S, Whicher JR, Hansen DA, Hale WA, Chemler JA, Congdon GR, Narayan AR, Hakansson K, Sherman DH, Smith JL and Skiniotis G, *Nature*, 2014, 510, 512–517. [PubMed: 24965652]
77. Herbst DA, Jakob RP, Zahringer F and Maier T, *Nature*, 2016, 531, 533–537. [PubMed: 26976449]
78. Herbst DA, Huitt-Roehl CR, Jakob RP, Kravetz JM, Storm PA, Alley JR, Townsend CA and Maier T, *Nat. Chem. Biol.*, 2018, 14, 474–479. [PubMed: 29610486]
79. Skiba MA, Sikkema AP, Fiers WD, Gerwick WH, Sherman DH, Aldrich CC and Smith JL, *ACS Chem. Biol.*, 2016, 11, 3319–3327. [PubMed: 27723289]
80. Chopra T and Gokhale RS, *Methods Enzymol.*, 2009, 459, 259–294; [PubMed: 19362644] Brown L, Wolf JM, Prados-Rosales R and Casadevall A, *Nat. Rev. Microbiol.*, 2015, 13, 620–630. [PubMed: 26324094]
81. Quadri LEN, *Crit. Rev. Biochem. Mol. Biol.*, 2014, 49, 179–211. [PubMed: 24625105]
82. Astarie-Dequeker C, Le Guyader L, Malaga W, Seaphanh FK, Chalut C, Lopez A and Guilhot C, *PLoS Pathog.*, 2009, 5, e1000289; [PubMed: 19197369] Cambier CJ, Takaki KK, Larson RP, Hernandez RE, Tobin DM, Urdahl KB, Cosma CL and Ramakrishnan L, *Nature*, 2014, 505, 218–222; [PubMed: 24336213] Reed MB, Domenech P, Manca C, Su H, Barczak AK, Kreiswirth BN, Kaplan G and Barry CE, 3rd, *Nature*, 2004, 431, 84–87. [PubMed: 15343336]
83. Camacho LR, Constant P, Raynaud C, Laneelle MA, Triccas JA, Gicquel B, Daffe M and Guilhot C, *J. Biol. Chem.*, 2001, 276, 19845–19854. [PubMed: 11279114]
84. Cox JS, Chen B, McNeil M and Jacobs WR, Jr., *Nature*, 1999, 402, 79–83. [PubMed: 10573420]
85. Rombouts Y, Alibaud L, Carrere-Kremer S, Maes E, Tokarski C, Ellass E, Kremer L and Guerardel Y, *J. Biol. Chem.*, 2011, 286, 33678–33688. [PubMed: 21803773]
86. Etienne G, Malaga W, Laval F, Lemassu A, Guilhot C and Daffe M, *J. Bacteriol.*, 2009, 191, 2613–2621. [PubMed: 19181796]
87. Buglino J, Onwueme KC, Ferreras JA, Quadri LE and Lima CD, *J. Biol. Chem.*, 2004, 279, 30634–30642; [PubMed: 15123643] Trivedi OA, Arora P, Vats A, Ansari MZ, Tickoo R, Sridharan V, Mohanty D and Gokhale RS, *Mol. Cell.*, 2005, 17, 631–643. [PubMed: 15749014]

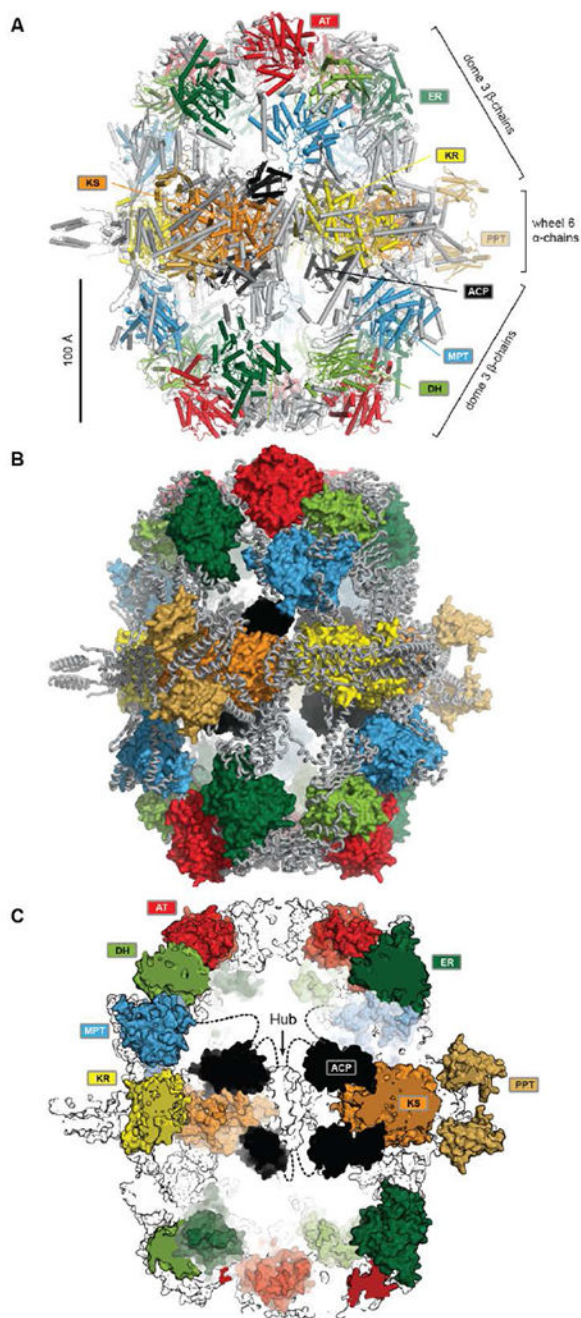
88. Onwueme KC, Ferreras JA, Buglino J, Lima CD and Quadri LE, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, 101, 4608–4613. [PubMed: 15070765]
89. Xie XK, Watanabe K, Wojcicki WA, Wang CCC and Tang Y, *Chem. Biol.*, 2006, 13, 1161–1169. [PubMed: 17113998]
90. Boritsch EC, Frigui W, Cascioferro A, Malaga W, Etienne G, Laval F, Pawlik A, Le Chevalier F, Orgeur M, Ma L, Bouchier C, Stinear TP, Supply P, Majlessi L, Daffe M, Guilhot C and Brosch R, *Nat Microbiol.*, 2016, 1, 15019. [PubMed: 27571976]
91. Edwards AL, Matsui T, Weiss TM and Khosla C, *J. Mol. Biol.*, 2014, 426, 2229–2245. [PubMed: 24704088]
92. Rittner A and Grininger M, *Chembiochem.*, 2014, 15, 2489–2493. [PubMed: 25318851]
93. Keatinge-Clay A, *J. Mol. Biol.*, 2008, 384, 941–953. [PubMed: 18952099]
94. Akey DL, Razelun JR, Tehranisa J, Sherman DH, Gerwick WH and Smith JL, *Structure*, 2010, 18, 94–105. [PubMed: 20152156]
95. Fiers WD, Dodge GJ, Sherman DH, Smith JL and Aldrich CC, *J. Am. Chem. Soc.*, 2016, 138, 16024–16036; [PubMed: 27960309] Faille A, Gavalda S, Slama N, Lherbet C, Maveyraud L, Guillet V, Laval F, Quemard A, Mourey L and Pedelacq JD, *J. Mol. Biol.*, 2017, 429, 1554–1569. [PubMed: 28377293]
96. Gay D, You YO, Keatinge-Clay A and Cane DE, *Biochemistry*, 2013, DOI: 10.1021/bi400988t.
97. Bretschneider T, Heim JB, Heine D, Winkler R, Busch B, Kusebauch B, Stehle T, Zocher G and Hertweck C, *Nature*, 2013, 502, 124–128. [PubMed: 24048471]
98. Bonnett SA, Whicher JR, Papireddy K, Florova G, Smith JL and Reynolds KA, *Chem. Biol.*, 2013, 20, 772–783; [PubMed: 23790488] Schafer M, Stevenson CE, Wilkinson B, Lawson DM and Buttner MJ, *Cell. Chem. Biol.*, 2016, 23, 1091–1097. [PubMed: 27617849]
99. Keatinge-Clay AT, *Chem. Biol.*, 2007, 14, 898–908; [PubMed: 17719489] Keatinge-Clay AT and Stroud RM, *Structure*, 2006, 14, 737–748; [PubMed: 16564177] Zheng J and Keatinge-Clay AT, *J. Mol. Biol.*, 2011, 410, 105–117; [PubMed: 21570406] Zheng J, Piasecki SK and Keatinge-Clay AT, *ACS Chem. Biol.*, 2013, 8, 1964–1971; [PubMed: 23755878] Zheng J, Taylor CA, Piasecki SK and Keatinge-Clay AT, *Structure*, 2010, 18, 913–922. [PubMed: 20696392]
100. Zheng J, Gay DC, Demeler B, White MA and Keatinge-Clay AT, *Nat. Chem. Biol.*, 2012, 8, 615–621. [PubMed: 22634636]
101. Lu TB, Alexander R, Bignan G, Bischoff J, Connolly P, Cummings M, De Breucker S, Esser N, Fraiponts E, Gilissen R, Grasberger B, Janssens B, Ludovici D, Meerpoel L, Meyer C, Parker M, Peeters D, Schubert C, Smans K, Van Nuffel L and Vermeulen P, *Cancer Res.*, 2014, 74.
102. Khare D, Hale WA, Tripathi A, Gu L, Sherman DH, Gerwick WH, Hakansson K and Smith JL, *Structure*, 2015, DOI: 10.1016/j.str.2015.09.013.
103. Busch B, Ueberschaar N, Behnken S, Sugimoto Y, Werneburg M, Traitcheva N, He J and Hertweck C, *Angew. Chem. Int. Ed. Engl.*, 2013, 52, 5285–5289. [PubMed: 23606628]
104. Kapur S, Lowry B, Yuzawa S, Kenthirapalan S, Chen AY, Cane DE and Khosla C, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, 109, 4110–4115. [PubMed: 22371562]
105. Lowry B, Li X, Robbins T, Cane DE and Khosla C, *ACS Cent Sci.*, 2016, 2, 14–20. [PubMed: 26878060]
106. Beck BJ, Aldrich CC, Fecik RA, Reynolds KA and Sherman DH, *J. Am. Chem. Soc.*, 2003, 125, 4682–4683. [PubMed: 12696866]
107. Donadio S, Staver MJ, McAlpine JB, Swanson SJ and Katz L, *Science*, 1991, 252, 675–679. [PubMed: 2024119]
108. Staunton J and Weissman KJ, *Nat. Prod. Rep.*, 2001, 18, 380–416; [PubMed: 11548049] Khosla C, Tang Y, Chen AY, Schnarr NA and Cane DE, *Annu. Rev. Biochem.*, 2007, 76, 195–221. [PubMed: 17328673]
109. Udway DW, Merski M and Townsend CA, *J. Mol. Biol.*, 2002, 323, 585–598. [PubMed: 12381311]
110. Turner WB and Aldridge DC, *Fungal Metabolites II*, Academic Press, London, 1983.
111. Kroken S, Glass NL, Taylor JW, Yoder OC and Turgeon BG, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, 100, 15670–15675. [PubMed: 14676319]

112. Bradshaw RE and Zhang S, *Mycopathologia*, 2006, 162, 201–213. [PubMed: 16944287]
113. Bailey AM, Cox RJ, Harley K, Lazarus CM, Simpson TJ and Skellam E, J.C.S., *Chem. Commun.*, 2007, 4053–4055; Cox RJ, *Org. Biomol. Chem.*, 2007, 5, 2010–2026. [PubMed: 17581644]
114. Crawford JM and Townsend CA, *Nat. Rev. Microbiol.*, 2010, 8, 879–889. [PubMed: 21079635]
115. Newman AG and Townsend CA, *J. Am. Chem. Soc.*, 2016, 138, 4219–4228. [PubMed: 26938470]
116. Minto RE and Townsend CA, *Chem. Rev.*, 1997, 97, 2537–2555. [PubMed: 11851470]
117. Huitt-Roehl CR, Hill EA, Adams MM, Vagstad AL, Li JW and Townsend CA, *ACS Chem. Biol.*, 2015, 10, 1443–1449. [PubMed: 25714897]
118. Zhou H, Qiao KJ, Gao ZZ, Vederas JC and Tang Y, *J. Biol. Chem.*, 2010, 285, 41412–41421. [PubMed: 20961859]
119. Crawford JM, Thomas PM, Scheerer JR, Vagstad AL, Kelleher NL and Townsend CA, *Science*, 2008, 320, 243–246. [PubMed: 18403714]
120. Li YR, Xu W and Tang Y, *J. Biol. Chem.*, 2010, 285, 22762–22771.
121. Newman AG, Vagstad AL, Storm PA and Townsend CA, *J. Am. Chem. Soc.*, 2014, 136, 7348–7362; [PubMed: 24815013] Vagstad AL, Newman AG, Storm PA, Belecki K, Crawford JM and Townsend CA, *Angew. Chem. Int. Ed. Engl.*, 2013, 52, 1718–1721. [PubMed: 23283670]
122. Zhou H, Li Y and Tang Y, *Nat. Prod. Rep.*, 2010, 27, 839–868. [PubMed: 20358042]
123. Ma SM, Watanabe K, Xie X, Zhang W, Wang CCC and Tang Y, *J. Am. Chem. Soc.*, 2007, 129, 10642–10643. [PubMed: 17696354]
124. Vagstad AL, Bumpus SB, Belecki K, Kelleher NL and Townsend CA, *J. Am. Chem. Soc.*, 2012, 134, 6865–6877. [PubMed: 22452347]
125. Crawford JM, Korman TP, Labonte JW, Vagstad AL, Hill EA, Kamari-Bidkorpheh O, Tsai SC and Townsend CA, *Nature*, 2009, 461, 1139–1143. [PubMed: 19847268]
126. Hine J and Ahn K, *J. Org. Chem.*, 1987, 52, 2083–2086; Hibbert F and Emsley J, *Adv. Phys. Org. Chem.*, 1990, 26, 255–379.
127. Barajas JF, Shakya G, Moreno G, Rivera H, Jr., Jackson DR, Topper CL, Vagstad AL, La Clair JJ, Townsend CA, Burkart MD and Tsai SC, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, 114, E4142–E4148. [PubMed: 28484029]
128. Fujii I, Watanabe A, Sankawa U and Ebizuka Y, *Chem. Biol.*, 2001, 8, 189–197. [PubMed: 11251292]
129. Brunner SD, Weber T, Kohli RH, Schwarzer D, Marahiel MA, Walsh CT and Stubbs MT, *Structure*, 2002, 10, 301–310; [PubMed: 12005429] Samel SA, Wagner B, Marahiel MA and Essen L-O, *J. Mol. Biol.*, 2006, 359, 876–889. [PubMed: 16697411]
130. Korman TP, Crawford JM, Labonte JW, Newman AG, Wong J, Townsend CA and Tsai SC, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, 107, 6246–6251. [PubMed: 20332208]
131. Li X, Sevillano N, La Greca F, Deis L, Liu YC, Deller MC, Mathews II, Matsui T, Cane DE, Craik CS and Khosla C, *J. Am. Chem. Soc.*, 2018, 140, 6518–6521. [PubMed: 29762030]
132. Meier JL, Mercer AC, Rivera H, Jr. and Burkart MD, *J. Am. Chem. Soc.*, 2006, 128, 12174–12184; [PubMed: 16967968] Worthington AS and Burkart MD, *Org. Biomol. Chem.*, 2006, 4, 44–46; [PubMed: 16357994] Worthington AS, Rivera H, Torpey JW, Alexander MD and Burkart MD, *ACS Chem. Biol.*, 2006, 1, 687–691. [PubMed: 17184132]
133. Nguyen C, Haushalter RW, Lee DJ, Markwick PR, Bruegger J, Caldara-Festin G, Finzel K, Jackson DR, Ishikawa F, O'Dowd B, McCammon JA, Opella SJ, Tsai SC and Burkart MD, *Nature*, 2014, 505, 427–431. [PubMed: 24362570]
134. Bruegger J, Haushalter RW, Vagstad AL, Shakya G, Mih N, Townsend CA, Burkart MD and Tsai SC, *Chem. Biol.*, 2013, 20, 1135–1146. [PubMed: 23993461]
135. Sippel KH, Vyas NK, Zhang W, Sankaran B and Quijoch FA, *J. Biol. Chem.*, 2014, 289, 33287–33295. [PubMed: 25301948]
136. Heil CS, Rittner A, Goebel B, Beyer D and Grininger M, *bioRxiv*, 2018, DOI: 10.1101/282525.
137. Whicher JR, Dutta S, Hansen DA, Hale WA, Chemler JA, Dosey AM, Narayan AR, Hakansson K, Sherman DH, Smith JL and Skiniotis G, *Nature*, 2014, 510, 560–564. [PubMed: 24965656]
138. Heneghan MN, Yakasai AA, Williams K, Kadir KA, Wasil Z, Bakeer W, Fisch KM, Bailey AM, Simpson TJ, Cox RJ and Lazarus CM, *Chem. Sci.*, 2011, 2, 972–979.

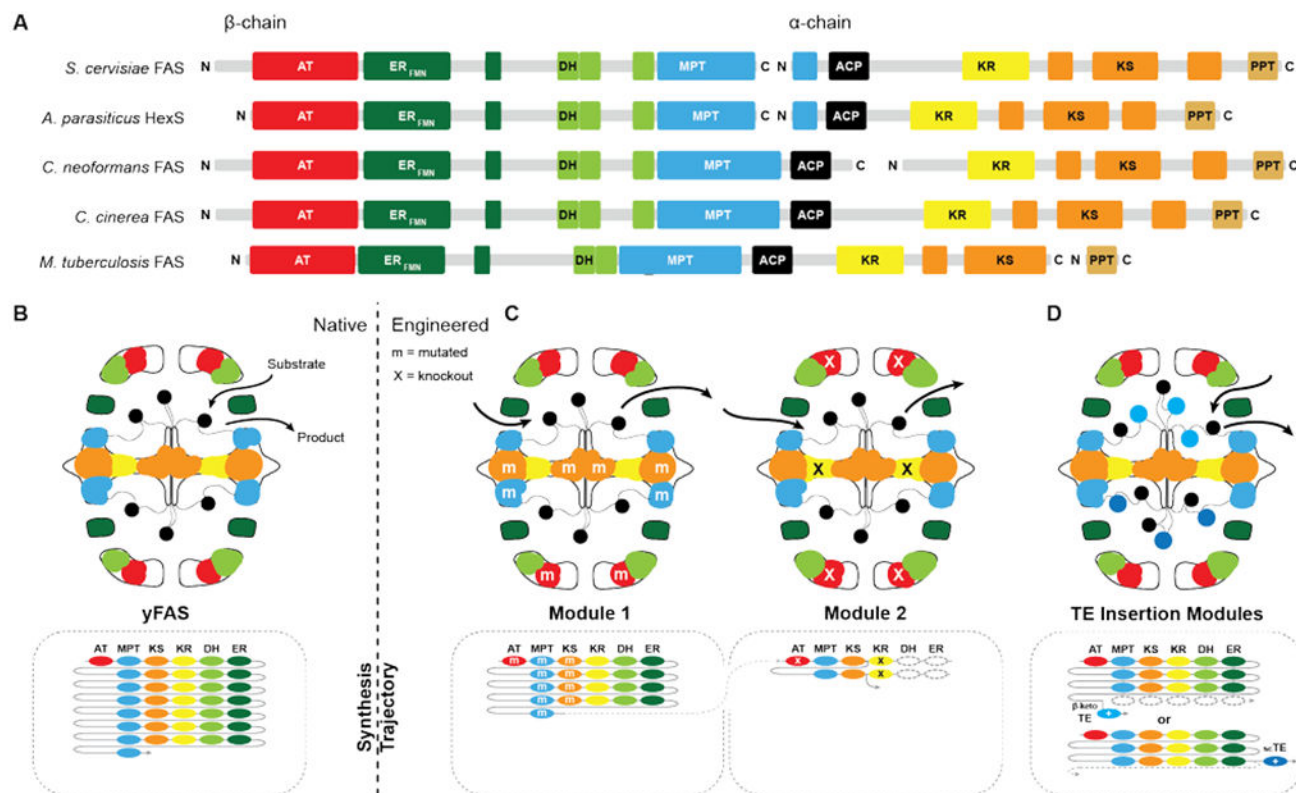
139. McInnes AG, Smith DG, Walter JA, Vining LC and Wright JLC, J.C.S., Chem. Commun., 1974, 0, 282–284; Eley KL, Halo LM, Song ZS, Powles H, Cox RJ, Bailey AM, Lazarus CM and Simpson TJ, Chembiochem, 2007, 8, 289–297. [PubMed: 17216664]
140. Proctor RH, Desjardins AE, Plattner RD and Hohn TM, Fungal Genet. Biol, 1999, 27, 100–112. [PubMed: 10413619]
141. Burmeister HR, Bennett GA, Vesonder RF and Hesseltine CW, Antimicrob. Agents Chemother, 1974, 5, 634–639. [PubMed: 15825417]
142. Cochrane RVK, Sanichar R, Lambkin GR, Reiz B, Xu W, Tang Y and Vederas JC, Angew. Chem. Int. Ed. Engl, 2016, 55, 664–668. [PubMed: 26783060]
143. Reeves CD, Hu ZH, Reid R and Kealey JT, Appl. Environ. Microbiol, 2008, 74, 5121–5129. [PubMed: 18567690]
144. Kim YT, Lee YR, Jin JM, Han KH, Kim H, Kim JC, Lee T, Yun SH and Lee YW, Mol. Microbiol, 2005, 58, 1102–1113; [PubMed: 16262793] I G and Trail F, Appl. Environ. Microbiol, 2006, 72, 1793–1799. [PubMed: 16517624]
145. Xu YQ, Espinosa-Artiles P, Schubert V, Xu YM, Zhang W, Lin M, Gunatilaka AAL, Sussmuth R and Molnar I, Appl. Environ. Microbiol, 2013, 79, 2038–2047. [PubMed: 23335766]
146. Chiang YM, Szweczyk E, Davidson AD, Keller N, Oakley BR and Wang CCC, J. Am. Chem. Soc, 2009, 131, 2965–2970. [PubMed: 19199437]
147. Winter JM, Sato M, Sugimoto S, Chiou G, Garg NK, Tang Y and Watanabe K, J. Am. Chem. Soc, 2012, 134, 17900–17903. [PubMed: 23072467]
148. Kasahara K, Miyamoto T, Fujimoto T, Oguri H, Tokiwano T, Oikawa H, Ebizuka Y and Fujii I, Chembiochem, 2010, 11, 1245–1252. [PubMed: 20486243]
149. Cox RJ, Glod F, Hurley D, Lazarus CM, Nicholson TP, Rudd BAM, Simpson TJ, Wilkinson B and Zhang Y, Chem. Commun. (Camb.), 2004, DOI: 10.1039/b411973h, 2260–2261.
150. Yang G, Rose MS, Turgeon BG and Yoder OC, Plant Cell, 1996, 8, 2139–2150; [PubMed: 8953776] Baker SE, Kroken S, Inderbitzin P, Asvarak T, Li BY, Shi L, Yoder OC and Turgeon BG, Mol. Plant. Microbe Interact, 2006, 19, 139–149. [PubMed: 16529376]
151. Song ZS, Cox RJ, Lazarus CM and Simpson TJ, Chembiochem, 2004, 5, 1196–1203. [PubMed: 15368570]
152. Bergmann S, Schumann J, Scherlach K, Lange C, Brakhage AA and Hertweck C, Nat. Chem. Biol, 2007, 3, 213–217. [PubMed: 17369821]
153. Zabala AO, Chooi YH, Choi MS, Lin HC and Tang Y, ACS Chem. Biol, 2014, 9, 1576–1586. [PubMed: 24845309]
154. Abe Y, Suzuki T, Ono C, Iwamoto K, Hosobuchi M and Yoshikawa H, Mol. Genet. Genomics, 2002, 267, 636–646. [PubMed: 12172803]



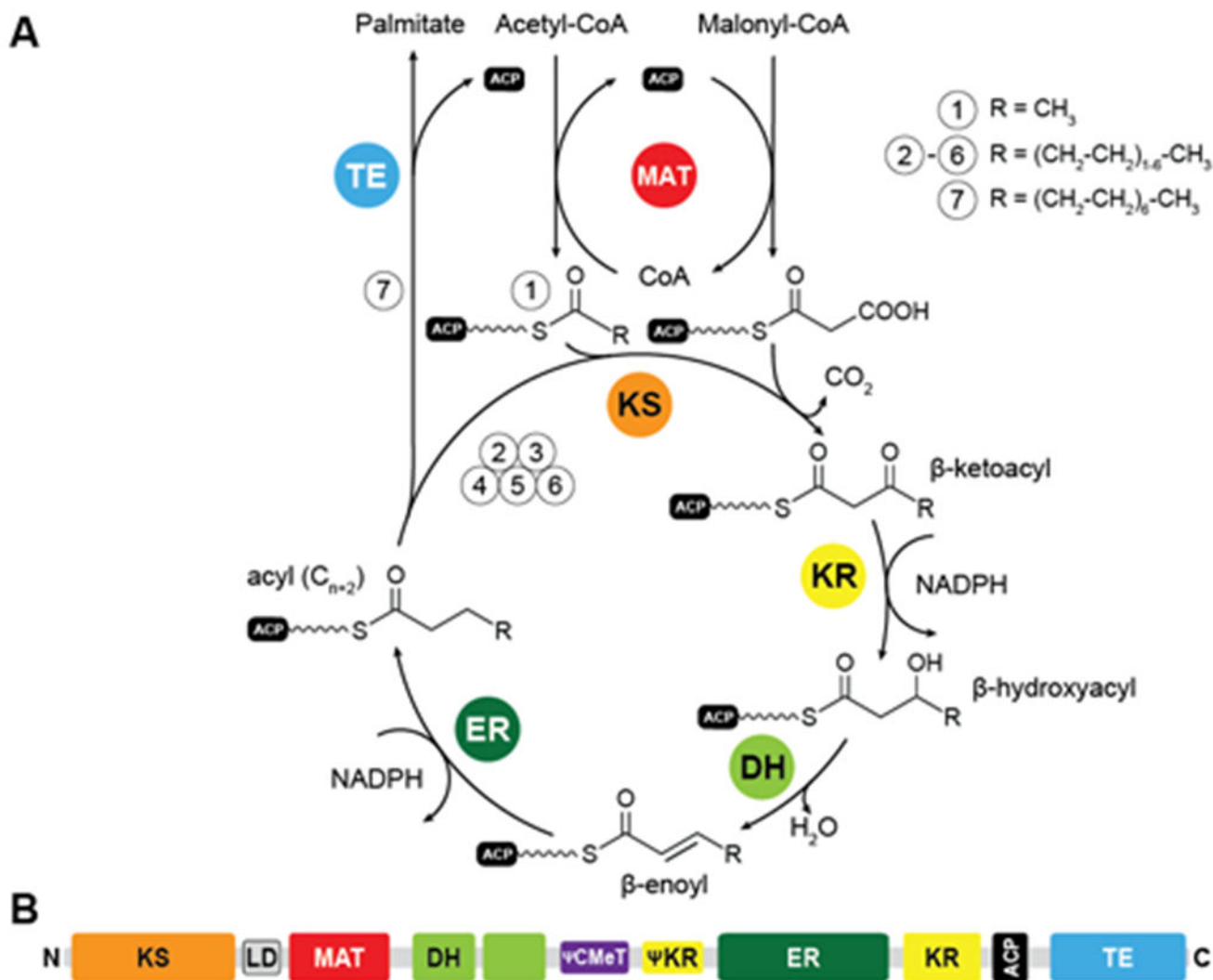


**Figure 2.**

The barrel-like architecture of yeast FAS. (A) Cartoon representation of the heterododecameric yeast FAS shows a central wheel of six  $\alpha$ -chains and two domes of three  $\beta$ -chains on each side. (B) Approx. 50% of the ~2.6 MDa architecture consist of scaffolding linkers (grey loops) that embed 42 catalytic domains and six ACPs (colored surfaces). (C) Cut-open view of the barrel. Each side of the barrel contains two reaction chambers with three double-tethered ACPs, which are organized by a central hub. Domain names correspond to Fig. 1.

**Figure 3.**

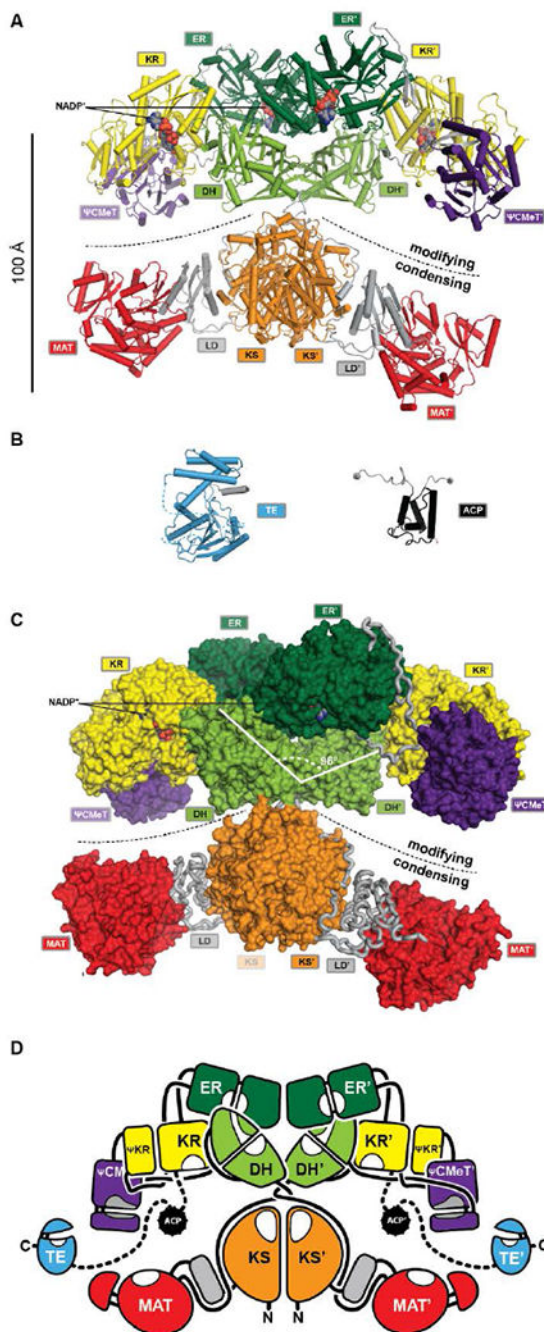
Domain organization and engineering approaches for yFASs. (A) Domain representation in sequence of evolutionarily related yFAS. (B) Schematic representation of yFAS architecture and synthesis trajectory. Domains are colored according to A. The sequential domain utilization of each domain in the reaction cycle is indicated below the barrel. (C) Two engineered yFAS assembly lines according to Gajweski *et al.* Domain mutagenesis modified the biosynthetic trajectory (m: mutation, X: knockout) and yielded novel products. (D) Engineered yFAS according to Zhu *et al.* Two separate yFAS variants are indicated in each reaction chamber of the barrel that harbor a TE insertion for the production of short chain fatty acids. TEs can be inserted either into the C-terminal (indicated for the upper chamber) or N-terminal (indicated for the lower chamber) ACP linker. Domain names correspond to Fig. 1. scTE: Short chain TE.



**Figure 4.**

Catalytic reaction cycle and domain organization of mFAS. (A) Fatty acid biosynthesis is initiated by loading an acetyl-unit from acetyl-CoA onto the ACP, which is subsequently elongated by the KS in iterative cycles of decarboxylative condensations using the extender unit malonyl-CoA. All building blocks are loaded by the MAT. Fully saturated acyl products are generated by sequential reduction and water elimination (KR, DH, ER) at the  $\beta$ -carbon position in each cycle and serve as substrate for further elongation reactions by the KS. (B) Linear domain organization of mFAS. KS: ketosynthase, MAT: malonyl-acetyl transferase, DH: dehydratase, CMeT: C-methyltransferase, ER: enoylreductase, KR: ketoreductase, ACP: acyl-carrier protein, TE: thioesterase. Inactive domains are indicated by leading  $\Psi$  (pseudo) symbols.





**Figure 5.**

The architecture of mFASs. (A) Cartoon representation of the asymmetric porcine FAS crystal structure reveals a structural separation into a modifying and condensing regions with a central hinge. (B) Crystal structure of human FAS TE and NMR structure of rat FAS ACP in cartoon representation. (C) In contrast to yFAS, mFAS invests only 9% of the total polypeptide into scaffolding. In the modifying region all domains (surface representation) are connected by direct domain-domain interfaces. The DH dimer has a V-shaped organization. (D) Schematic architectural drawing of the mFAS domain connectivity. Dotted

lines indicate linkers that have not been experimentally observed. Domain names correspond to Fig. 4.

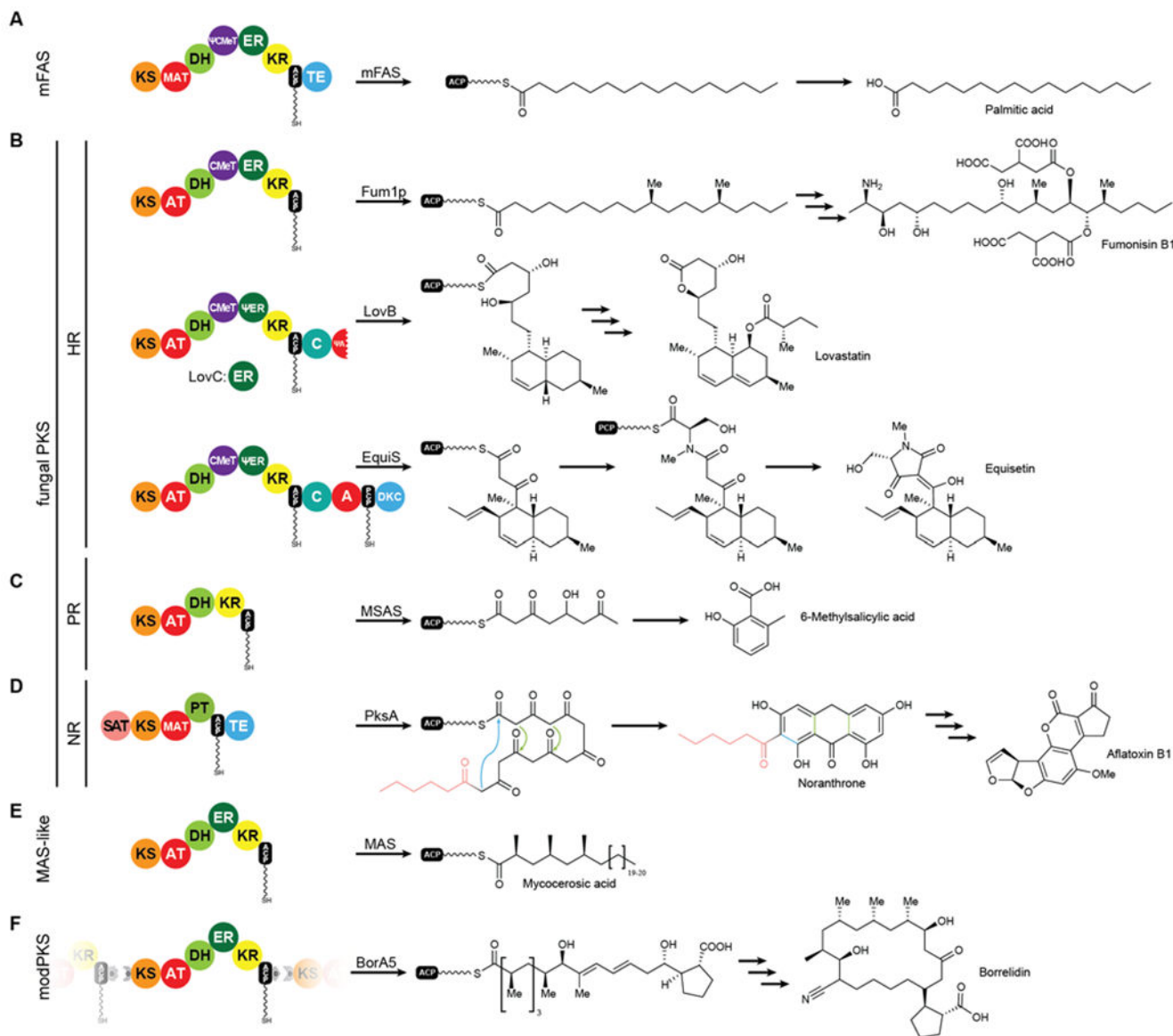
Author Manuscript

Author Manuscript

Author Manuscript

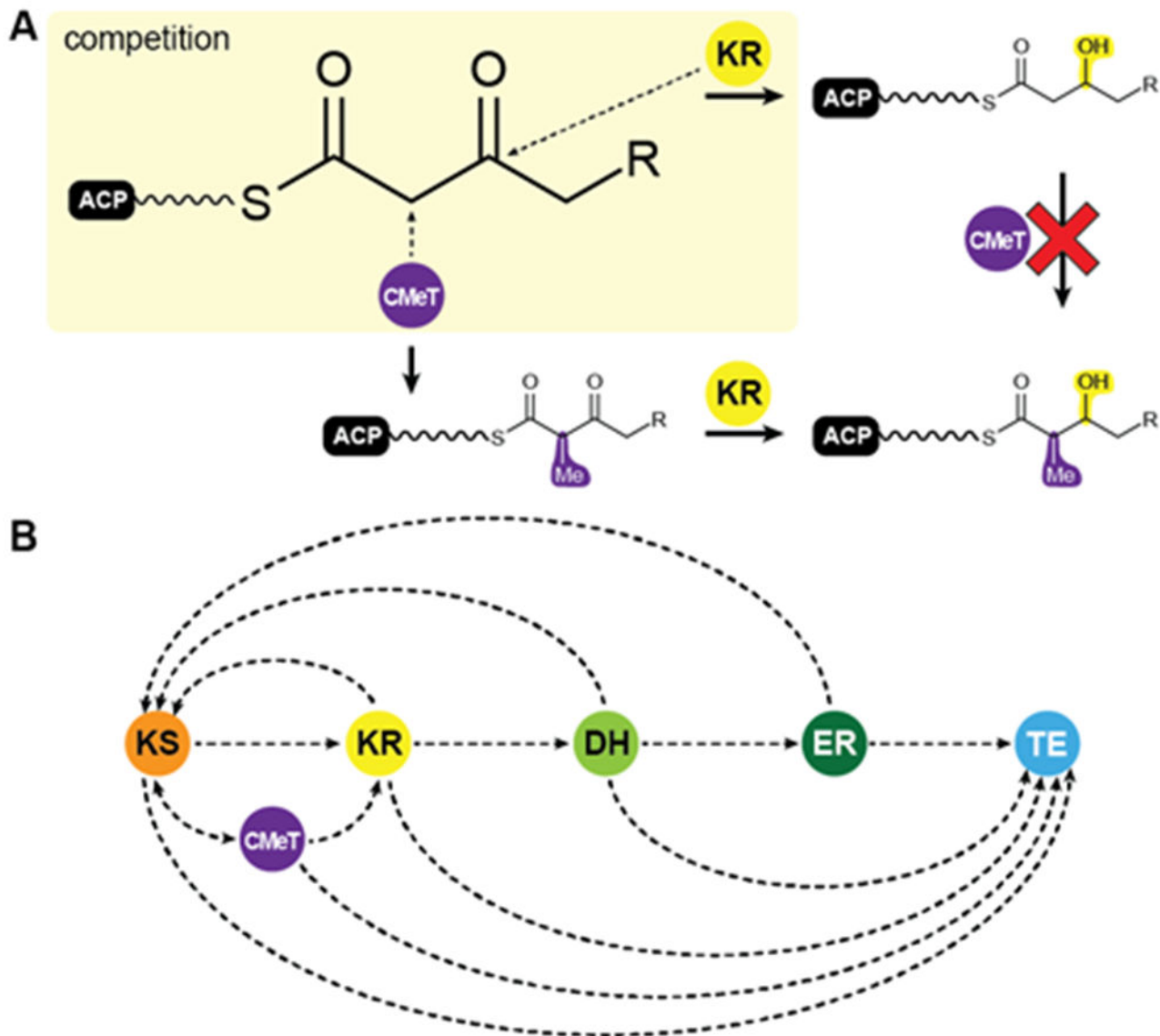
Author Manuscript





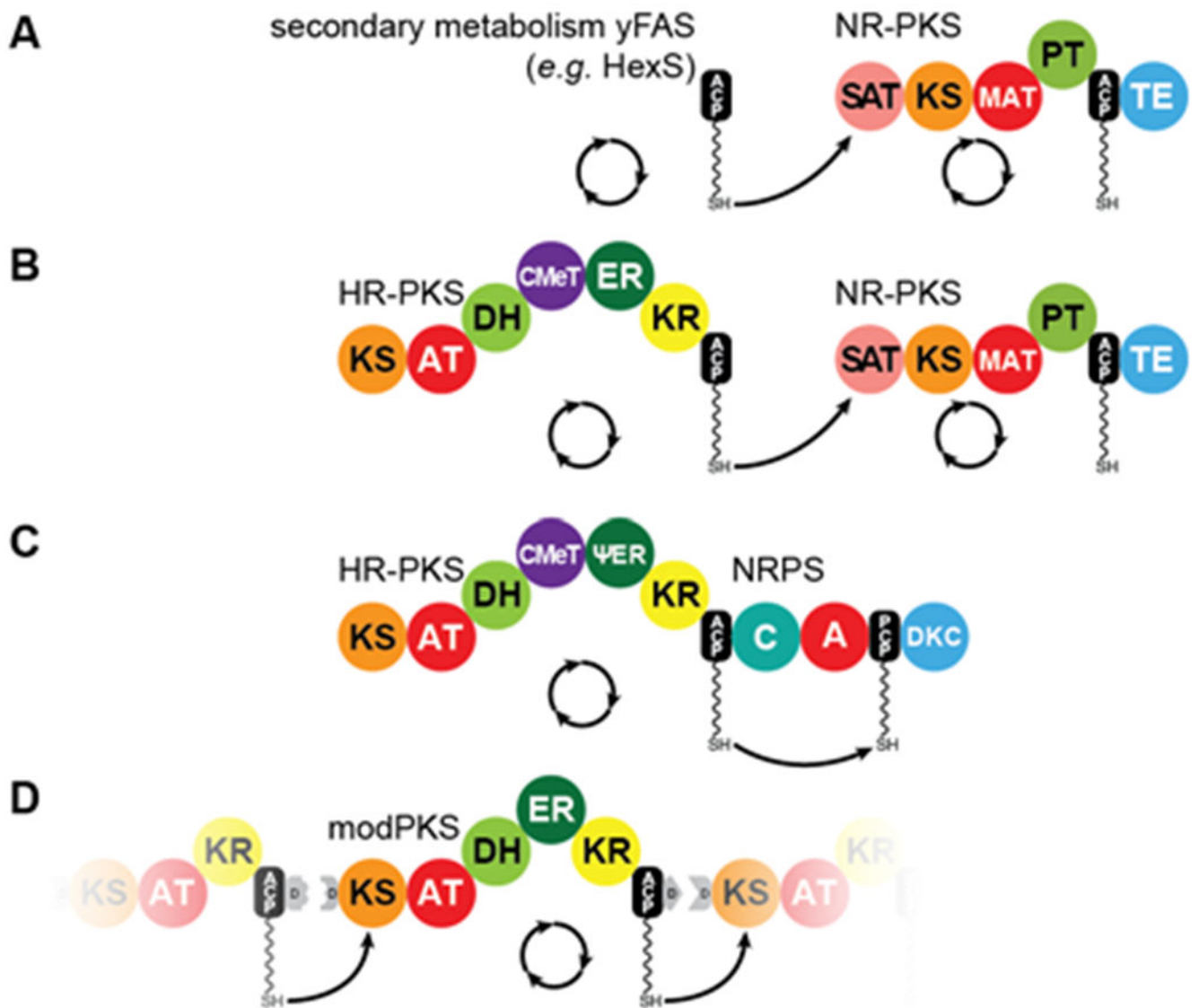
**Figure 6.**

Examples for iteratively acting mFAS and PKS modules. The domain organization and the corresponding biosynthetic products are indicated for mFASs (A), HR-PKSs (B), PR-PKSs (C), NR-PKSs (D), MAS-like PKS (E), and modPKSs (F). KS: ketosynthase, AT: acyltransferase, MAT: malonyl-acetyl transferase (mFAS context) / malonyl-specific acyltransferase (NR-PKS context), DH: dehydratase, CMeT: C-methyltransferase, ER: enoylreductase, KR: ketoreductase, ACP: acyl-carrier protein, TE: thioesterase, C: condensing domain (NRPS), A: adenylation domain (NRPS), PCP: peptidyl-carrier protein (NRPS), DKC: Diekmann cyclase, SAT: starter-unit acyltransferase, PT: product template domain, D: docking domain (modPKS). Inactive domains are indicated by leading Ψ (pseudo) symbols.

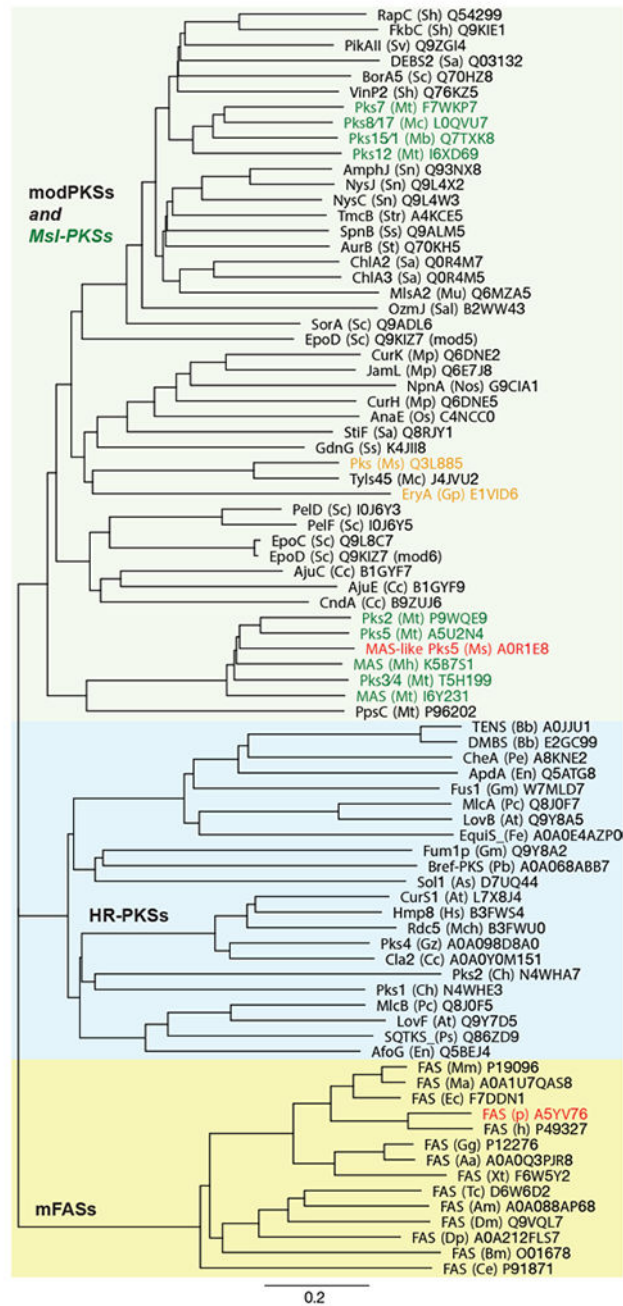


**Figure 7.**

Putative programming trajectory in fungal iPKS. (A) CMeTs and KR compete for their substrates: CMeT domains require  $C_{\alpha}$  acidity for methylation, which is lost after  $\beta$ -keto reduction. (B) Programming trajectory in fungal iPKS. The main programmed path appears to be described by specific ACP-substrate interactions with each target domain in every elongation and modification cycle. Domain names correspond to Fig. 6.

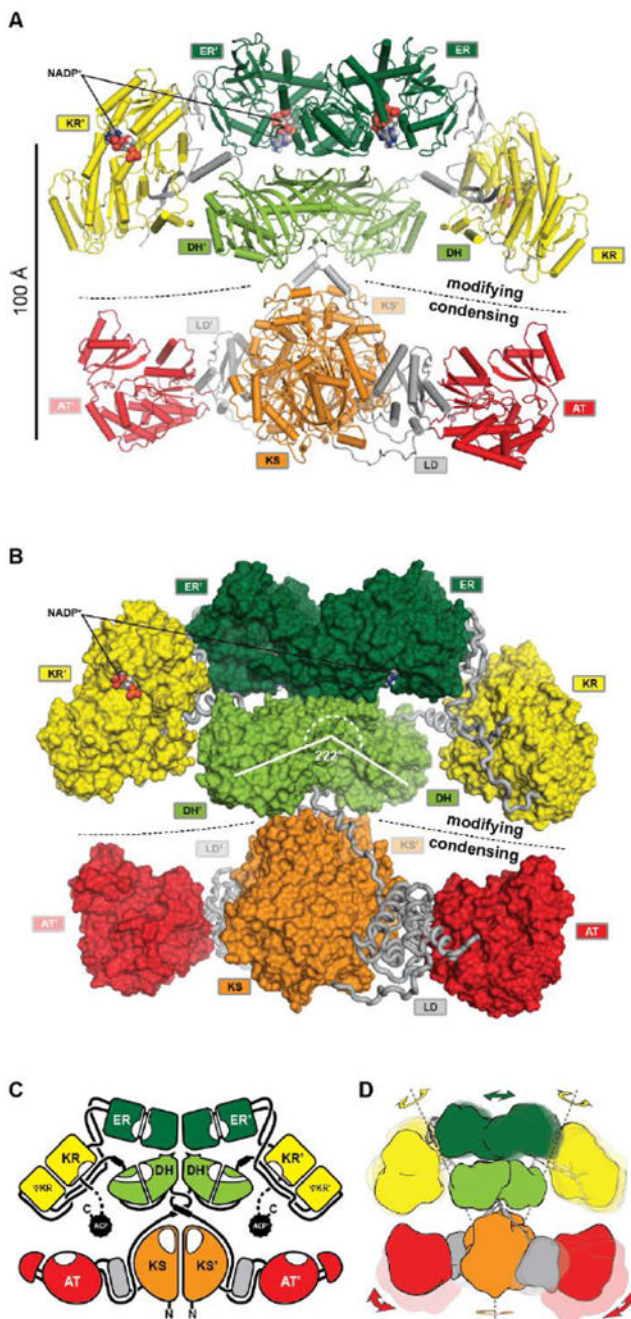


**Figure 8.** Iterative PKS assembly lines. Direct transfer from iteratively acting PKS and secondary metabolic yFAS ACPs to downstream acting modules has been observed for the pairs of yFAS:NR-PKS (A), HR-PKS:NR-PKS (B), HR-PKS:NRPS (C), and modPKS:modPKS (D). Domain names correspond to Fig. 6.



**Figure 9.**

Phylogeny of fully reducing modules of mFASs, HR-PKSs, and modPKSs/MAS-like PKSs. The three groups are indicated in colored boxes. Green: MAS-like PKSs; Red: Known structures of complete modules; Orange: Representative modPKSs with short ER-KR linker, low ER interface conservation and same architecture as MAS-like Pks5. All modules are labeled as: Protein name (organism abbr.) Uniprot number. Units are given as amino acid substitutions per site.



**Figure 10.**

The architecture of MAS-like PKSs. (A) Cartoon representation of the hybrid crystal structure of MAS-like Pks5 showing the canonical separation into a modifying and condensing regions. (B) In contrast to mFAS, the modifying region reveals a linker-based architecture. The KR is double tethered by linkers (loops) to the DH and ER without a direct interface. The DH-ER interface is small and highly variable. Compared to mFAS the DH dimer angle is inverted. (C) Schematic representation of the MAS-like architecture. Dotted lines indicate linkers that have not been experimentally observed. (D) Scheme indicating the



conformational variability in fully reducing PKS modules. Domain movements in the modifying region are conformationally coupled across the dimeric synthase. Domain names correspond to Fig. 6.

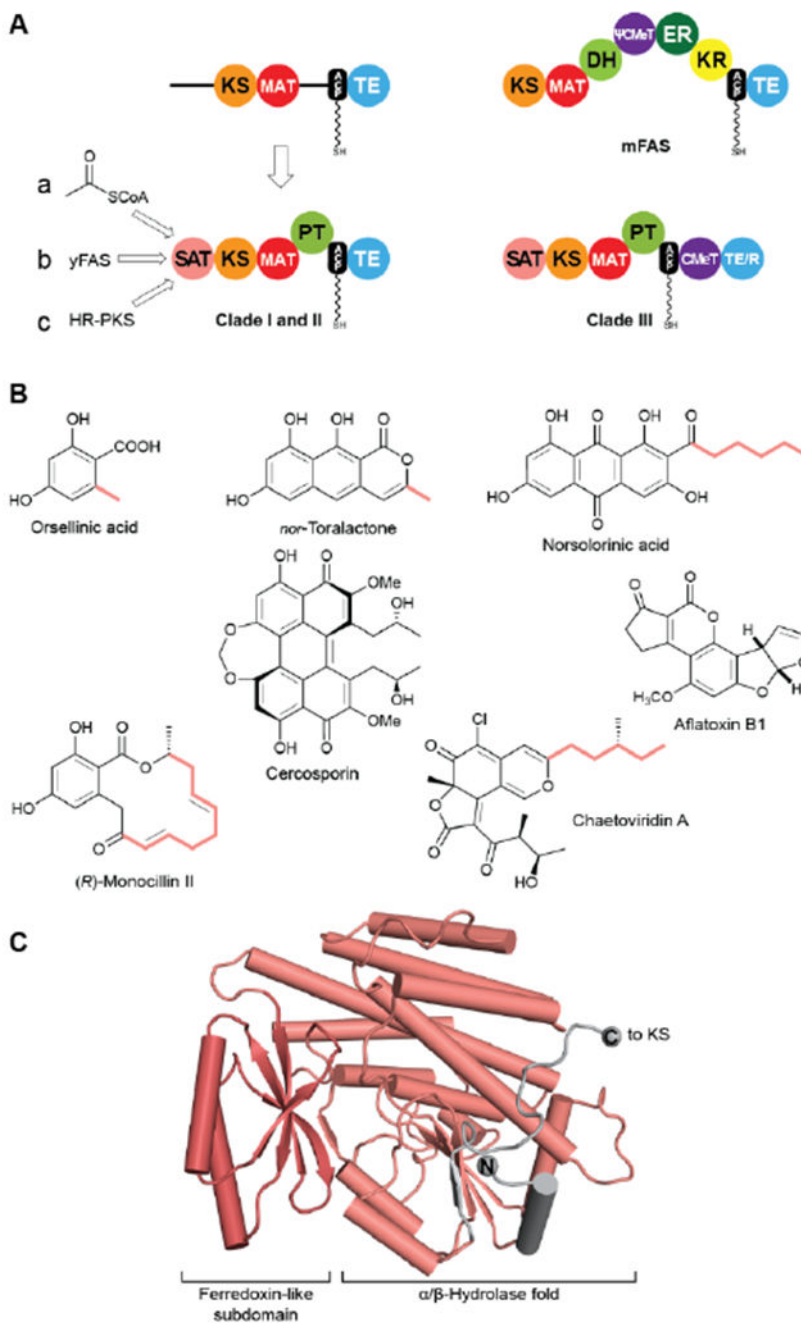
Author Manuscript

Author Manuscript

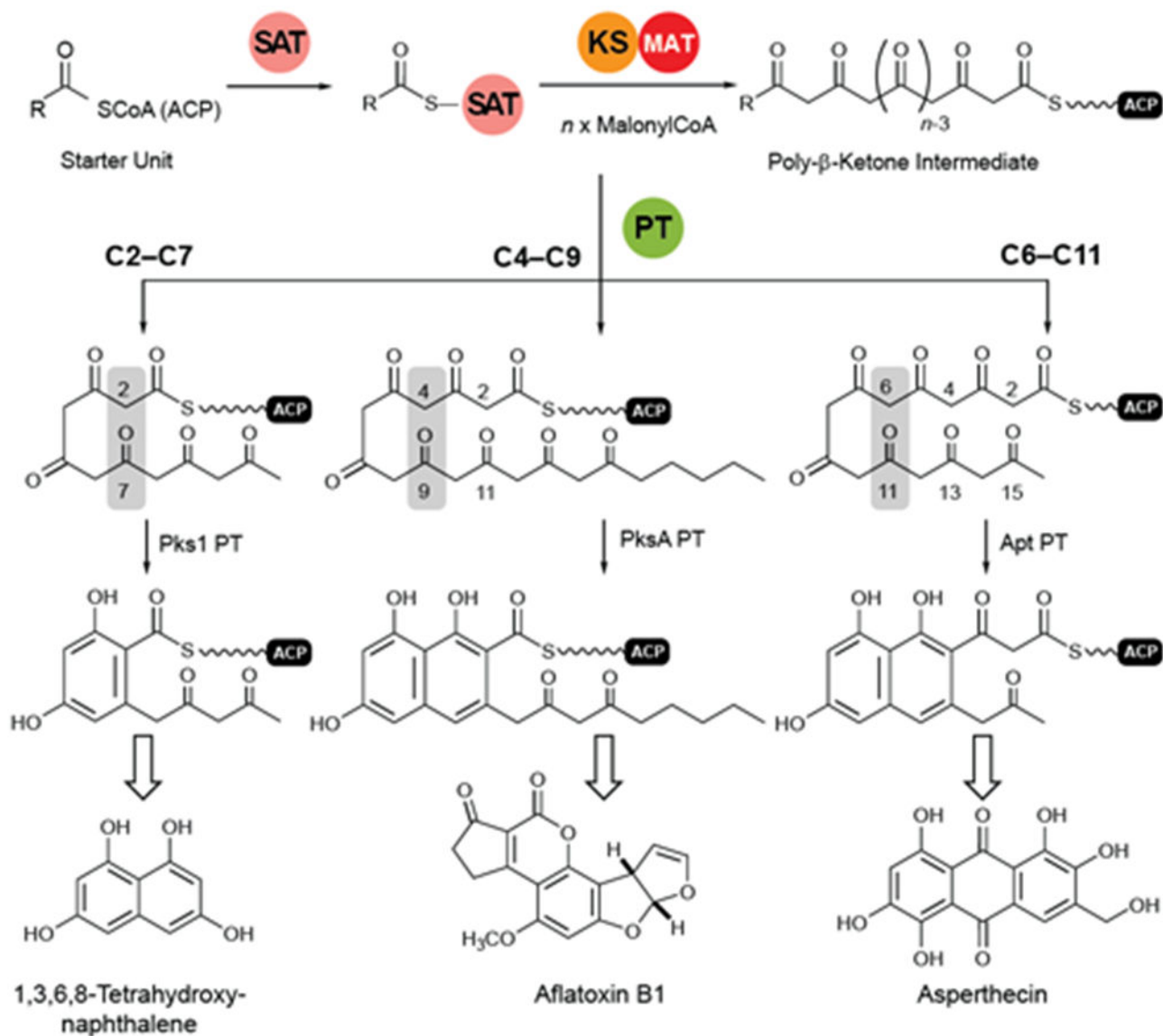
Author Manuscript

Author Manuscript

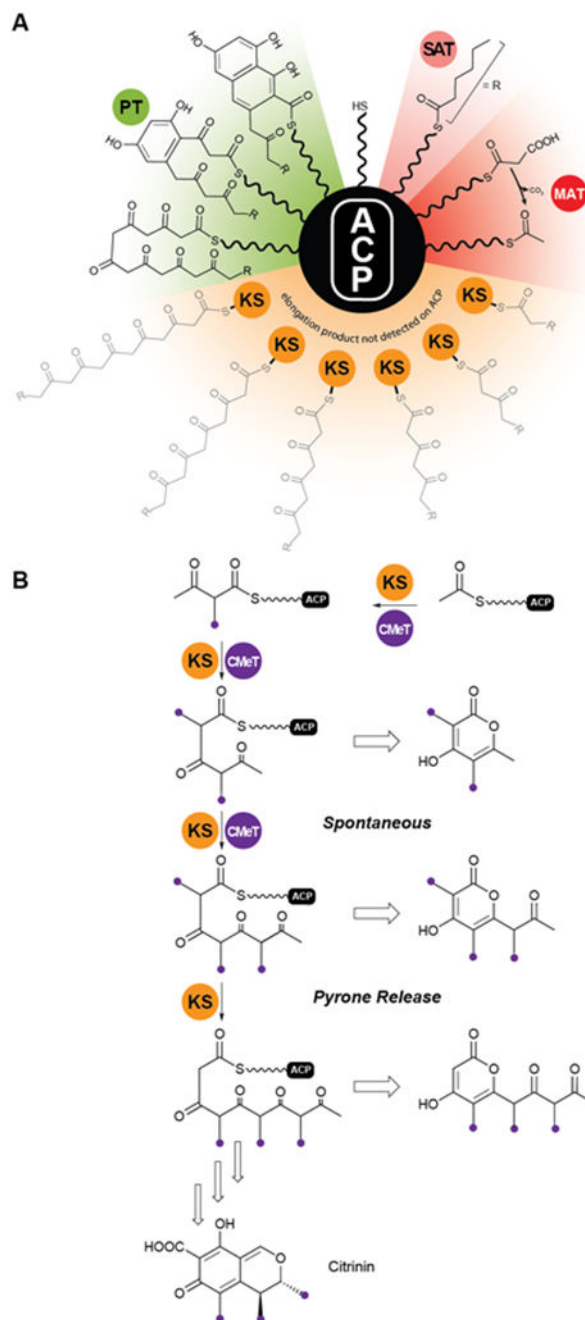


**Figure 11.**

Identification of SAT and PT domains in NR-PKS. (A) SAT and PT domains of NR-PKSs were identified by the UMA algorithm and deconstruction experiments. SATs can accept substrates from three different pathways (a-c). Only Clade III NR-PKS contain CMeT domains. (B) Products of pathways involving NR-PKSs. The identity of some substrates loaded by the SAT are indicated (colored). (C) Crystal structure of the CazM SAT domain. Domain names correspond to Fig. 6. R: reduction domain.

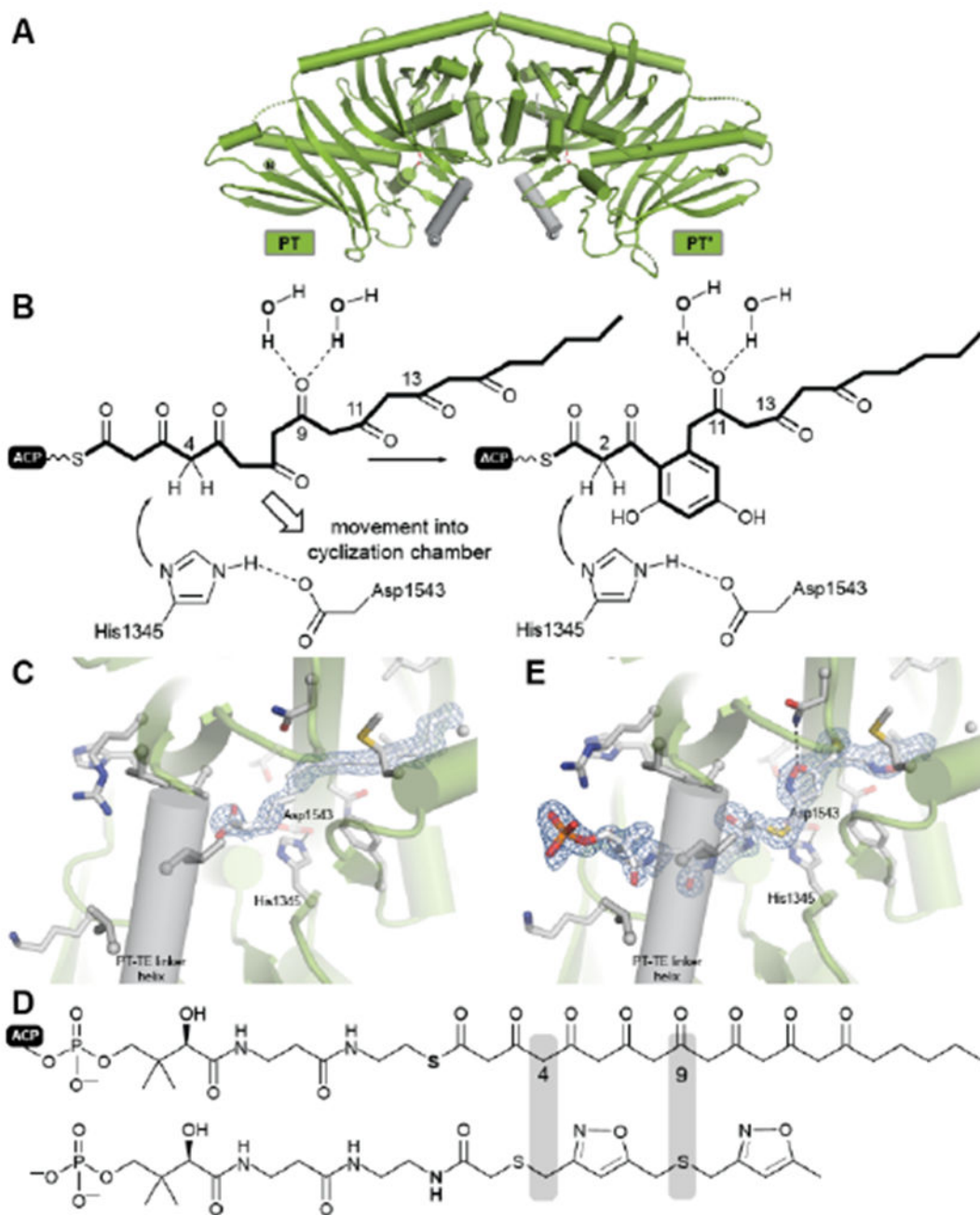


**Figure 12.** Cyclization reactions catalyzed by PT domains. The product of the loading/condensing region (SAT-KS-MAT) can be cyclized via three different routes. Domain names correspond to Fig. 6.



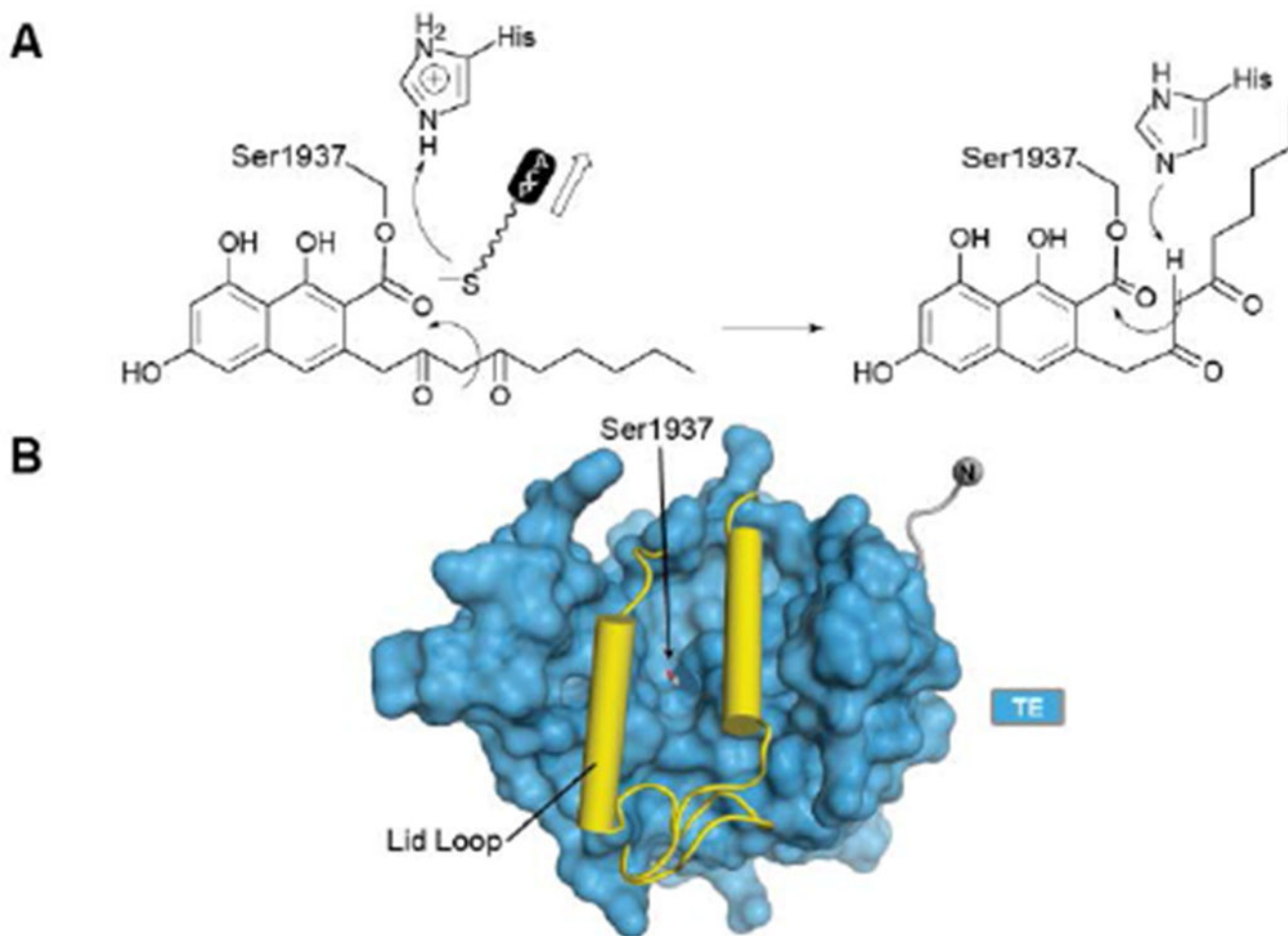
**Figure 13.**

Protection of the nascent polyketide and programming. (A) In clade I and II NR-PKS, intermediates of the nascent polyketide prior to achieving full length (shown in grey) could not be detected on the ACP, indicating that the intermediates stay tethered to the KS during the chain elongation cycles. (B) In Clade III NR-PKS the nascent polyketide chain has to leave the KS active site for C-methylation, which is a programmed event and required for complete chain elongation, but allows derailed pyrone products to be observed. Domain names correspond to Fig. 6.



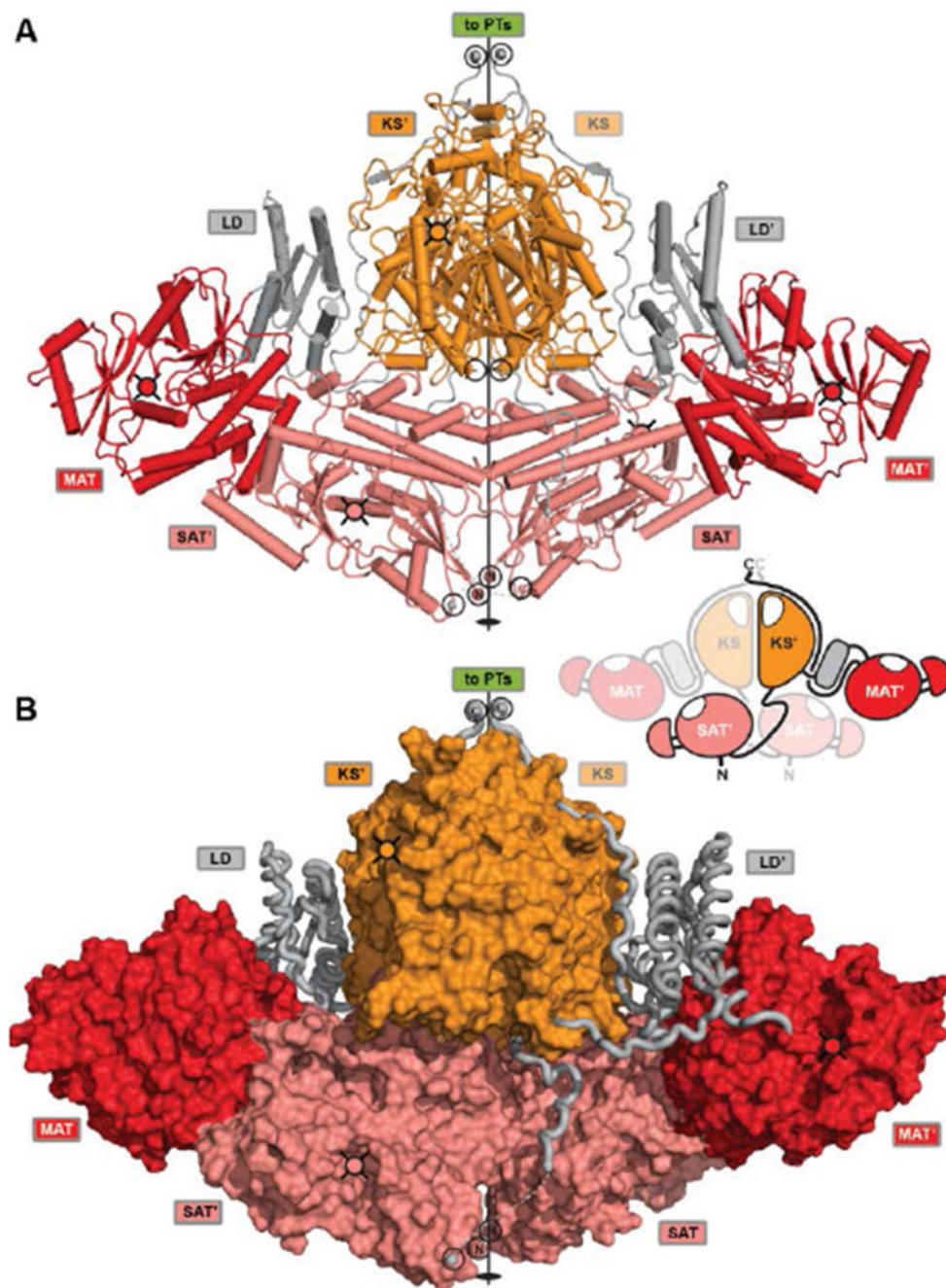
**Figure 14.**

Cyclization by the PksA PT domain. (A) Structure of the dimeric PksA PT domain. (B) The proposed cyclization mechanism involves a movement of the polyketide chain into the PT cyclization chamber. Fo-Fc shaked omit electron density maps at  $3.0\sigma$  of the palmitate (C) and a polyketide substrate analog (E) bound to the PT active site. (D) Derivation of the crystallized substrate analog.



**Figure 15.** The PksA TE domain. (A) Proposed reaction mechanism and product release. (B) TE domains have a lid loop that closes the active site.





**Figure 16.**

Crystal structure of the CTB1 NR-PKS loading/condensing region. (A) The SAT domains are integrated into the canonical condensing region architecture and align the *N*- and *C*-termini along the central two-fold axis (black circles) together with the termini of the KS. The inset representation shows the intertwined organization where the SAT mainly interfaces with the neighboring protomer. (B) Linker representation of (A) showing an extended SAT-linker interaction with its neighboring partner SAT domain. Active site locations and



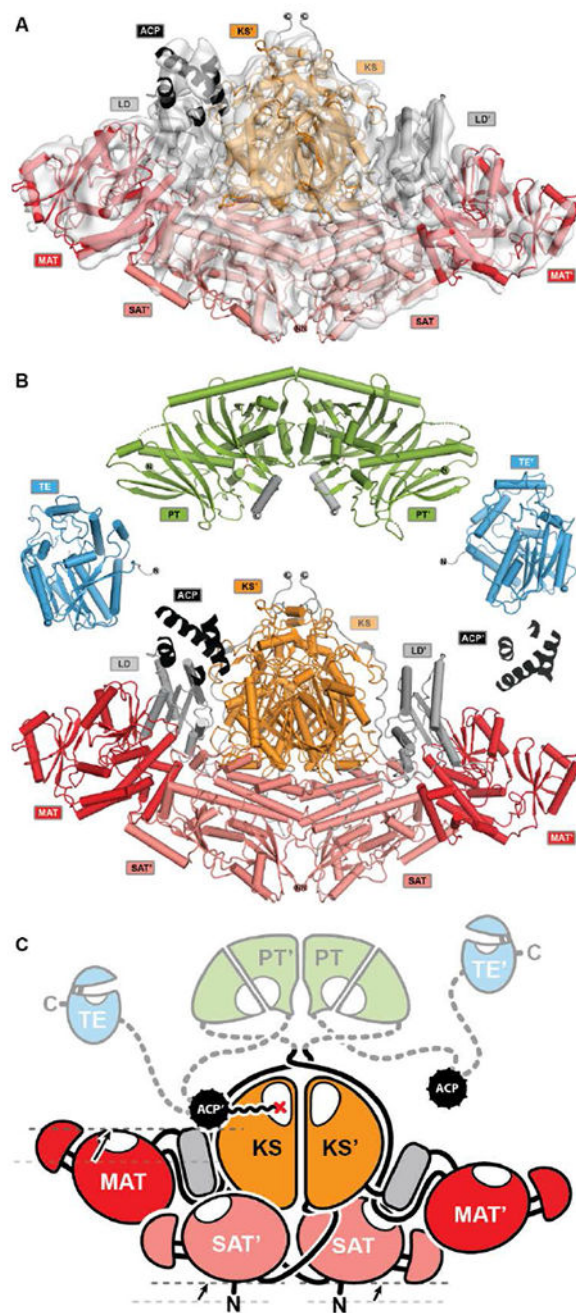
entrances, respectively, are indicated by cross-hatched filled-in circles. Domain names correspond to Fig. 6.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Figure 17.**

Proposed mode of conformational coupling in NR-PKS substrate loading. (A) Asymmetric cryo-EM structure of crosslinked ACP to the KS-active site of the loading/condensing region CTB1. (B) Structures of all NR-PKSs catalytic domains have been elucidated. The structures of CTB1 SAT-KS-MAT=ACP2 (=, crosslinked to KS) and the structures of PksA TE and PT are shown. (C) Schematic representation of the NR-PKS architecture in a catalytically

relevant asymmetric state. Conformational changes upon ACP binding are indicated with arrows. Domain names correspond to Fig. 6.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 1.**

Fungal biosynthetic pathways involving HR-PKSs.

Pathway Product	PKS name	Organism	Reference	Remark
Lovastatin	LovB, LovF	<i>Aspergillus terreus</i>	71	LovF: diketide synthase
Desmethylbassianin	DMBS	<i>Beauveria bassiana 992.05</i>	138	Hybrid HR-PKS-NRPS
Tenellin	TENS	<i>Beauveria bassiana 110.25</i>	139	Hybrid HR-PKS-NRPS
Fumonisin	Fum1p	<i>Gibberella fujikuroi</i>	53, 140	
Equisetin	EquiS	<i>Fusarium equiseti</i>	73, 141	Hybrid HR-PKS-NRPS
Cladosporin	Cla2	<i>Cladosporium cladosporioides</i>	142	Transfer to Cla3 (NR)
Hypothemycin	Hmp8	<i>Hypomyces subiculosus</i>	143	Transfer to Hmp3 (NR)
Radicicol	Rdc5	<i>Pochonia chamydosporia</i>	143	Transfer to Rdc1(NR)
Zearaleone	Pks4	<i>Gibberella zeae</i>	144	Transfer to Pks13 (NR)
10,11-Dehydrocurvularin	CURS1	<i>Aspergillus terreus</i>	145	Transfer to CURS2 (NR)
Asperfuranone	AfoG	<i>Aspergillus nidulans</i>	146	Transfer to AfoE (NR)
Chaetomugilin / Chaetoviridin	CazF	<i>Chaetomium globosum</i>	147	Transfer to CazM (NR)
Cytochalasin	CheA	<i>Penicillium expansum</i>	69	Hybrid HR-PKS-NRPS
Prosolanapyrone	Sol1	<i>Alternaria solani</i>	148	
Squalestatin	SQTKS	<i>Phoma sp. C2932</i>	149	
T-toxin	Pks1, Pks2	<i>Cochliobolus heterostrophus</i>	150	
Fusarin C	Fus1	<i>Fusarium moniliforme</i>	151	Hybrid HR-PKS-NRPS
Aspyridone	AdpA	<i>Aspergillus nidulans</i>	152	Hybrid HR-PKS-NRPS
Brefeldin A	Bref-PKS	<i>Eupenicillium brefeldianum</i>	153	
Compactin	MlcA, MlcB	<i>Penicillium citrinum</i>	154	MlcB: diketide synthase