Published in final edited form as:

Nat Prod Rep. 2018 October 17; 35(10): 1082–1096. doi:10.1039/c8np00058a.

Protein-Protein Interactions in “*cis*-AT” Polyketide Synthases

Greg J. Dodge^{a,b}, Finn P. Maloney^a, and Janet L. Smith^{a,b}

^aLife Sciences Institute

^bDepartment of Biological Chemistry, University of Michigan, Ann Arbor, Michigan, USA 48109

Abstract

Polyketides are a valuable source of bioactive and clinically important molecules. The biosynthesis of these chemically complex molecules has led to the discovery of equally complex polyketide synthase (PKS) pathways. Crystallography has yielded snapshots of individual catalytic domains, di-domains, and multi-domains from a variety of PKS megasynthases, and cryo-EM studies have provided initial views of a PKS module in a series of defined biochemical states. Here, we review the structural and biochemical results that shed light on the protein-protein interactions critical to catalysis by PKS systems with an embedded acyltransferase. Interactions include those that occur both within and between PKS modules, as well as with accessory enzymes.

1. Introduction

Polyketide synthases (PKS) are the source of many chemically diverse and pharmaceutically valuable natural products¹. Among the most versatile of PKS systems, the bacterial type I PKS pathways have long been a source of fascination to (bio)chemists and the targets of engineering efforts². Biosynthesis proceeds in a defined series of steps that include pathway initiation to append an acyl group to an acyl carrier protein (ACP), followed by a series of extensions and modifications. Synthesis involves selection of a building block by an acyltransferase (AT), extension of an intermediate by decarboxylative Claisen condensation in a ketosynthase (KS), and optional modifying reactions by ketoreductase (KR), dehydratase (DH) and enoylreductase (ER) enzymes. In type I PKS pathways, these extension and modification enzymes may exist in a single multi-functional polypeptide that acts iteratively (iPKS), akin to fatty acid biosynthesis, or as a series of individual modules that act in a prescribed order in the biosynthetic scheme (mPKS). The PKS module includes the enzymes for one extension step and any associated modifying reactions as well as the ACP. An mPKS polypeptide may include one or more modules. Following a pathway-specific number of extension and modification cycles, a thioesterase (TE) in the ultimate module offloads the pathway product via hydrolysis or cyclization to a macrolactone.

Bacterial genomes encode a remarkable biosynthetic capacity for both polyketide and non-ribosomal peptide natural products, characterized by gene clusters for PKS, non-ribosomal peptide synthetase (NRPS) and PKS-NRPS hybrid pathways. The common thioester linkage

of a ketide or peptide intermediate to a phosphopantetheine cofactor (Ppant) on the carrier protein facilitates the combination of PKS and NRPS modules within a single pathway. The PKS genes share a common ancestor with the metazoan fatty acid synthase (type I FAS)³, which includes the identical chemical steps of building block selection, intermediate extension, modification, and offloading⁴. However, the multi-module PKS systems afford a major expansion of polyketide chemical diversity at the cost of a requisite expansion of genomic material, as each catalytic unit acts only once in the biosynthetic pathway, and not iteratively as in the FAS or an iPKS. In most cases, genes encoding a PKS pathway are clustered, generally “co-linear” with the order of biosynthesis⁵. Nature further exploited the modular theme by creation of PKS-NRPS hybrids and by insertion of non-PKS/NRPS genes, adding to the wealth of natural product structure and bioactivity⁶. An example is the insertion in many gene clusters of a cassette encoding proteins that construct a carbon branch at the β position of a polyketide intermediate, adapted from the mevalonate pathway in primary metabolism⁷.

The first step of polyketide extension is the selection of a building block by an AT. Modular PKS systems are denoted as “*cis*-AT” or “*trans*-AT” depending on the scheme for building block selection. Each module of a *cis*-AT PKS includes an internal AT that selects a module-specific building block, providing an opportunity for building-block diversity within the pathway. In contrast, in *trans*-AT pathways an external AT provides an identical building block for each extension step. Although the core enzymes are homologs of FAS enzymes, the *cis*-AT and *trans*-AT pathways apparently have separate evolutionary histories, and annotated pathways generally use one AT type or the other. Gene clusters for *trans*-AT pathways more commonly contain inserted genes from outside the PKS realm, and *trans*-AT PKS modules are commonly split between polypeptides⁸. The *cis*-AT pathways adhere more closely to the metazoan type I FAS domain organization (KS-AT-DH-ER-KR-ACP-TE) and are among the most extensively characterized of PKS systems⁴.

Here we review the protein-protein interactions within and between *cis*-AT PKS modules. Arguably the most important of these involve the ACP domain, which must deliver the appropriate intermediate to each enzyme active site within the module and also transfer the module product to the next KS of the pathway for further extension or to the TE for offloading. Much has been learned about PKS protein-protein interactions since an excellent 2008 review by Weissman and Müller⁹ and a 2013 review by Tang and co-workers focused on the structure and function of catalytic domains¹⁰. Here we first review the domain interactions within a module that define the module architecture, including homomeric dimer-forming interactions. This is followed by a discussion of interactions of the ACP with enzyme domains within the module. Protein-protein interactions between modules are reviewed in terms of docking-domain interactions and ACP interactions with the downstream KS. Finally, we review critical ACP interactions in the creation of a β -branch in the polyketide.

2. Domain-domain interactions within a *cis*-AT PKS module

The domain order and composition within *cis*-AT PKS systems have long been understood at the level of primary structure by comparison of amino acid sequences of PKS and FAS

enzymes^{11, 12}, while the domain boundaries were established with the initial crystal structure for each enzyme type^{13–18}. Recent structures of a metazoan FAS, an intact PKS module and a PKS tri-domain have begun to illustrate the protein-protein interactions within and between modules of *cis*-AT PKS systems^{19–22} (Figs. 1 & 2). The structures reveal varying overall architectures, implying that these complex and fascinating machines either have varying architecture or are dynamic. Within an mPKS module or an iPKS, protein-protein interactions include critical contacts of the ACP with each enzyme domain as well as the enzyme-enzyme contacts that define the PKS architecture. Interactions with proteins outside the module involve an ACP, which must recognize the next module in the pathway^{23–25} or interact with other pathway enzymes such as those for β -branching^{7, 26–28}. A full understanding of the protein-protein interactions of PKS modules poses significant technical challenges, ranging from the formidable size of PKS polypeptides to the transience of many interactions.

2.1. PKS module organization

The architecture of a *cis*-AT mPKS extension module or iPKS depends on the domain composition, which varies from the non-reducing KS-AT-ACP to the fully reducing KS-AT-DH-ER-KR-ACP. Despite a plethora of crystal structures of excised PKS domains^{29, 30}, only a handful of structures include multiple domains and thus inform our understanding of module architecture. Within the past decade, structures from crystallography and cryo-electron microscopy (cryo-EM) have provided a few direct views of domain-domain interactions in several natural di-domains, a larger fragment and one full module (Figs. 1 and 2), but the picture is far from complete.

Conjecture about module architecture dates from the discovery that PKS modules are dimeric, leading to competing proposals for head-to-tail and helical head-to-head symmetric structures^{31, 32}. The first breakthrough in visualizing domain-domain interactions within a module came with the near-simultaneous reports of crystal structures of a PKS KS-AT di-domain¹⁴ and a metazoan FAS³³ (Figs. 1A, 2B). The excised PKS KS-AT dimer is extremely similar to the KS-AT dimer within the FAS, reinforcing the widely held assumption that a PKS module with a full set of modifying domains (KR, DH, ER) has a similar or identical architecture to the metazoan FAS. Recently, Maier and co-workers provided highly informative views of the full set of PKS modifying domains (DH-ER-KR) in a crystal structure from “MAS”, a mycocerosic acid synthase-like iPKS²¹ (Fig. 1B). MAS has a domain composition and overall structure similar to the metazoan FAS, but sequence analysis places it firmly in the PKS realm. With few exceptions, subsequent structures have reinforced the assumption of PKS-FAS architectural similarity. The most notable exception, and at present the only direct view of a full PKS module, is a set of cryo-EM reconstructions of pikromycin module 5 (PikAIII), a natural mono-module with domain composition KS-AT-KR-ACP^{20, 22} (Fig. 2A).

The segregation into distinct “condensing” (N_{term}-KS-AT-) and “modifying” (-DH-ER-KR-ACP-C_{term}) regions is a key architectural feature of the *cis*-AT PKS that is shared with the metazoan FAS¹⁹ (Fig. 1A). The condensing-modifying junction occurs immediately after a well conserved “post-AT linker”, which is very clear in sequence alignments. Excisions at

the junction often result in stable condensing and modifying fragments, so the junction is assumed to be flexible, with few contacts between condensing and modifying regions. While the two-region organization is likely a general feature, it is important to remember that we have direct views of several biochemical states of only one full-length PKS module^{20, 22}, which contains only one modifying domain, and only one view of a metazoan FAS, in which the biochemical state is unknown and the ACP and terminal TE domains are not visible due to their flexibility in the crystal lattice¹⁹. A high degree of flexibility in the metazoan FAS, especially between the condensing and modifying regions, was observed directly by cryo-EM³⁴ and atomic force microscopy³⁵, and it is likely that all the megasynthases exhibit varying conformations throughout the catalytic cycle.

2.2. Module architecture: Dimer-forming interactions

The principal domain-domain interactions within a *cis*-AT PKS are the homomeric dimer-forming contacts of some catalytic domains: the KS in the condensing region^{14, 25, 36, 37}, and in the modifying region the DH^{15, 38–40}, the TE^{18, 41} and, in at least some cases, the ER^{16, 21, 42}. Excepting the TE, none of these domains is reliably dimeric when excised from the native polypeptide. Thus, the catalytic domains likely cooperate in forming the module dimer, in some cases assisted by additional small dimerization elements. These other contributors to dimer formation include the coiled-coil docking domain (dd^{KS}) preceding an N-terminal KS^{23–25}, a post-AT dimerization element (post-AT DE)⁴³, and a post-ACP dimerization element (post-ACP DE)²³ (Table 1, Figs. 3 and 4). Interestingly, the size of the dimer interface is largely uncorrelated with the oligomer state of the domain when excised from the module.

KS dimer interaction.—The KS domain has the largest and most conserved interaction surface among enzyme domains in PKS modules (~ 1800 Å²) and thus is likely a major contributor to module dimerization (Fig. 3A). However, excised KS domains, generally in the context of a KS-AT di-domain, are often monomeric in solution despite the large interface. For example, the PikAIII module was monomeric when the small post-ACP dimerization element was removed by truncation²⁰. Remarkably, the excised MAS KS-AT was a monomer in solution and also in crystals where more than 50 amino acids at the usual dimer interface were disordered²¹.

DH dimer interaction.—The PKS DH domain adopts a double hot-dog fold and forms an elongated dimer via the ends of the N-terminal “hot-dogs” (Fig. 3B). Despite the small interface (550–700 Å²), most excised DH domains are dimeric^{15, 38–40}. The multifunctional AmbDH3 DH/cyclase from the ambruticin pathway is a notable exception as it is monomeric in solution and in crystals⁴⁵, but presumably contributes to dimerization of the module. The flat, linear dimer in excised PKS DH domains and in the MAS modifying region²¹ differs from the chevron-shaped dimer in the metazoan FAS¹⁹.

ER dimer interaction.—There is uncertainty about whether the ER domain within a PKS is universally dimeric. In the MAS modifying region²¹ and in the metazoan FAS¹⁹, the ER domain dimerizes by forming a continuous β-sheet between nucleotide-binding sub-domains of the partner subunits (Fig. 3C). This dimerization mode also occurs in two different crystal

structures of an ER nucleotide-binding sub-domain fragment from PpsC, a PKS module closely related to MAS (PDB 1PQW, 4OKI). However, four full-length ER domains excised from modules of the spinosyn, curacin A and jamaicamide pathways are monomeric in solution and in crystals^{16, 42}. Additionally, a short helix is identically positioned in the four monomeric ER structures and is incompatible with the MAS-type dimer. In the dimerization region, the sequences of the MAS and PpsC are more hydrophobic and slightly longer than the corresponding sequences of the monomeric ERs (21 amino acids vs. 18). Each sequence fits its structure and seems incompatible with the alternative structure. It is an open question whether the short helix re-arranges to permit dimer formation in the full module or the corresponding modules have monomeric ERs^{21, 46}.

TE dimer interaction.—TE domains reside in terminal modules of PKS pathways where they offload the final intermediate as either a macrocycle or linear carboxylic acid⁴⁷. In all crystal structures of terminating PKS TEs, a four-helix bundle forms at the dimer interface^{18, 41, 48} and supports a substrate tunnel in each subunit⁴⁴ (Fig. 3D). Although the dimer interface is small (750–850 Å²), terminating TEs are uniformly dimeric when excised and are strong dimerization elements in terminal PKS modules.

Small dimerization elements.—Studies of the non-covalent “docking” interactions of sequential PKS modules revealed two small dimerization elements near the module termini: an N-terminal parallel coiled-coil KS docking domain^{23–25} (Fig. 4A) and a post-ACP dimerization element near the module C-terminus²³ (Fig. 4B). Surprisingly, the 35-residue post-ACP element is required for dimerization of the full-length PikAIII module²⁰. Another dimerization element was discovered in what was previously thought to be an extended linker between the AT and KR in module 2 of the spinosyn pathway⁴³. Once identified, this 50-residue element is readily recognized in sequences of many modules where the modifying regions lack a dimeric enzyme domain (Fig. 4C). Where present, the small dimerization elements play a critical role in module architecture. Additional contributors to module dimerization may yet be discovered.

2.3. Module architecture: Domain-domain interactions

Condensing region (KS-AT).—Domain-domain interactions within the PKS condensing region vary considerably in published structures. Crystal structures of five different excised KS-AT di-domains reveal an elongated shape with a central KS dimer and AT “wings” at the periphery, separated from the KS by a 100-residue linker domain^{14, 25, 36, 37, 49} (Fig. 2B). The linker domain interacts with both KS and AT, which do not contact each other. The orientation of linker-AT relative to KS varies by at least 35° among these structures²¹, but is generally similar to the KS-AT region of the metazoan FAS¹⁹ (Fig. 1A).

The PikAIII cryo-EM structures are a marked contrast to these other PKS and FAS structures, but they are the only structures of a full-length PKS module^{20, 22}. The PikAIII condensing region has a compact, bent shape and the linker-AT domains differ in orientation by ~120° compared to other structures (Fig. 2). The PikAIII AT and KS domains contact one another in an interface that was validated by mutagenesis²⁰. The ACP and KS domains are in defined acylation states in each of the six PikAIII structures, revealing acylation-

dependent positions for the ACP at active site entrances and also for the linker-AT and KR domains. These structures reveal separate entrances to the KS active site for the PikAIII ACP bearing a methylmalonyl building block and for the upstream ACP bearing the module 4 pentaketide product²⁰ (Fig. 5A).

The possibility that other PKS modules may assume a PikAIII-like conformation during catalysis was tested in the erythromycin PKS (DEBS) with module 3, which has the same domain composition as PikAIII (KS-AT-KR-ACP)⁴⁹. An antibody antigen-binding fragment (Fab) with high affinity for the DEBS KS3-AT3 was identified, and the complex visualized in a crystal structure (Fig. 6). The KS3-AT3 was in an extended conformation, as in structures of other excised KS-AT di-domains^{14, 25, 37}. The Fab binding pose was interpreted as specific for the extended KS-AT conformation because the Fab contacted both the KS and AT domains (~4700 Å² interface with KS, ~480 Å² with AT). As Fab binding did not impede catalysis through the full-length module 3, it was concluded that the KS-AT is in an extended conformation throughout the catalytic cycle. However, the crystal structure indicates that the Fab is not conformationally specific. The complementarity-determining regions (CDRs) of the Fab variable domains (V_L and V_H) interact extensively with the KS coiled-coil docking domain, whereas only one amino acid of each AT domain contacts the Fab (Fig. 6). The AT contact with the Fab C_H domain is almost certainly labile, given the famously flexible “elbow angle” between Fab variable (V_L-V_H) and constant (C_L-C_H) regions⁵⁰ and the solvent-exposed and minimal Fab-AT contact. Thus, in solution, Fab binding is not expected to enforce an extended conformation in the KS-AT.

Modifying region (DH-ER-KR-ACP).—Our understanding of domain-domain interactions in the PKS modifying region comes from a crystal structure of the MAS modifying region²¹, which includes all three canonical modifying enzymes (KR, DH, ER) but lacks the ACP (Fig. 1B). Fortuitously, the crystal structure provided nine independent views of the DH-ER-KR dimer, revealing minimal enzyme-enzyme contacts and remarkably flexible connections. Each of the nine dimer views (18 monomer views) is a snapshot in the dynamic landscape of the modifying region. The DH and ER domains contribute to the dimer interface, but they have few contacts with each other and none with the peripheral KR domain. This is in striking contrast to the FAS where both the DH and ER contact the KR (compare Fig. 1A and 1B), but it is consistent with the PikAIII structures where KR has varying orientations dependent on the acyl-ACP state²². Among the nine snapshots of the MAS modifying region, the KR varies in orientation by 40° relative to the DH or the ER²¹. The snapshots also show evidence of a screw motion (14° rotation coupled with 8 Å translation) at the small and variable (350–650 Å²) DH – ER interface, comprising the N-terminal DH β-strands and the ER nucleotide-binding sub-domain.

Like the MAS snapshots, the cryo-EM study of PikAIII provided evidence for motion of the monomeric domains (linker-AT and KR), but the trapped conformations were coupled to the ACP acylation state. The deleterious effects on catalytic throughput from mutagenesis of residues at domain interfaces reinforces their importance to function^{20, 22}. These results suggest an additional complication in understanding domain interfaces within a PKS module – the interfaces themselves may be transient and, at least in PikAIII, are dependent on the

catalytic state of the module. Like the MAS and PikAIII, dynamic domain arrangements were also observed in a PKS loading module using negative-stain EM⁵¹.

Inter-domain linkers and structural domains.—The architecture of *cis*-AT PKS modules depends critically on the interactions of enzyme domains with linker peptides and “structural” or pseudo-catalytic domains. These non-catalytic elements have remarkable parallels with the metazoan FAS, for example the 100-residue KS-AT linker domain in the condensing region of both PKS and FAS. Additionally, an extended post-AT linker crosses the surface of the KS-AT linker domain and the KS (Fig. 1A), makes both hydrophobic and hydrogen bonding contacts with the KS, and is conserved in *cis*-AT PKS and metazoan FAS sequences. Studies of excised KS, AT and ACP domains from the DEBS pathway showed that the post-AT linker appended to the AT was required for the Claisen condensation reaction of acylated KS with methylmalonyl-ACP when the three fragments, KS, AT and ACP, were *in trans*⁵², suggesting a role for the post-AT linker in productive ACP interaction with the KS active site. Consistent with this finding, in the methylmalonyl-PikAIII EM structure, the methylmalonyl-ACP was positioned at the KS active site entrance adjacent to the position of the post-AT linker in other structures (Figs. 2A & 5A), although the post-AT linker was not visible in the cryo-EM maps²⁰.

In the modifying region, all ketoreductases are composed of two tightly associated, homologous domains¹³: a pseudo or structural domain (KR_S) and a catalytic domain (KR_C). The evolutionary history of the KR is apparent by comparison to the bacterial type II FAS, which is composed of discrete enzymes including a KR tetramer (FabG). The associated KR_S-KR_C of PKS and metazoan FAS resembles half of the FabG KR tetramer⁵³ and is a clear product of gene duplication. The duplication of an ancestral KR gene must have preceded the gene fusions leading to the type I PKS and metazoan FAS because the ER domain is inserted between the KR_S and KR_C domains. Furthermore, some *cis*-AT PKS modules include a C-methyltransferase domain (MT) for α-methylation of a β-keto polyketide intermediate^{17, 54}, whereas the metazoan FAS has a pseudo-MT (ΨMT) domain at an identical position in the modifying region: DH-(Ψ)MT-KR_S-ER-KR_C-ACP. This suggests that the last common ancestor of the *cis*-AT PKS module and the metazoan FAS may have been a PKS with an active MT domain. In the FAS context, the KR_S is truncated, the ΨMT lacks catalytic activity, and these two pseudo domains share a large interface. How the MT and KR_S domains interact in PKS modules is unknown.

KR_S and KR_C are bridged by two long, antiparallel β-strands. This “β-zipper” is an integral part of the ketoreductase structure and contributes to the central β-sheets of both KR_S and KR_C¹³. Where present, the MT (ΨMT in FAS) resides between the first β-zipper strand and KR_S¹⁷, and the ER between the second β-zipper strand and KR_C (Fig. 7A). The β-zipper has a conserved sequence that became easily recognizable once the boundaries of the inserted ER and MT domains were established^{13, 17}.

While much of the initial structural and biochemical research to characterize PKS protein-protein interactions has focused on experimentally accessible mono- and multi-domains, it has become increasingly clear that the linker regions play an important role both in organization and function of the module^{19–21, 56}. The linkers are generally longer in MAS

than in the metazoan FAS, perhaps accounting for the more open structure of the MAS modifying region (compare Fig. 1A and 1B). For example, the MAS DH-KR_S linker includes a 12-residue α -helix that does not exist in the FAS, and the KR_S-ER linker is eight amino acids longer in the MAS than in the FAS. Moreover, the linkers in MAS, an iterative PKS, also differ from the modular PKS linkers, which are inferred from structures of mono- and di-domains. For example, in MAS and the metazoan FAS, the ER domain is bracketed by linker peptides that form a short two-stranded antiparallel β -sheet. No such β -sheet exists in a crystal structure of the KR_S-ER-KR_C excised from module 2 of the spinosyn pathway¹⁶, and it may not exist in other modular PKS systems where the ER-KR_C linker is considerably shorter than in MAS (8 vs. 21 amino acids)⁴⁶. The picture that emerges from these crystal structures and from the PikAIII cryo-EM structures is one of a highly mobile KR that contacts other enzyme domains only transiently, if at all. The open structure of the MAS modifying region may be a hallmark of the many related PKS pathways, in which a lack of contacts between catalytic domains facilitates adaptation by domain deletion and permits rapid expansion through gene duplication.

Other linkers function as simple tethers. Within an mPKS module or an iPKS, the linker between the condensing and modifying regions is typically short (< 20 amino acids) and flexible. In contrast, the linker preceding the ACP is longer (~ 30 amino acids) and poorly conserved. Even with a 30-residue tether, great flexibility is needed for the ACP to interact with every catalytic domain in the module and with the KS of the downstream module. Similarly, the inter-module linker between the ACP and KS domains of naturally fused modules (ACP_{*i*-1} – KS_{*i*}) also must be flexible. Distinct positions for modules 5 and 6 in the DEBS3 polypeptide, a natural dimodule, were proposed in a small-angle x-ray scattering (SAXS) study⁵⁷, however the extremely low resolution (50 Å) SAXS model was created by enforcing collinearity of the module twofold axes in fitting a single model to the SAXS data. In general, a fixed orientation for naturally fused modules seems unlikely given the dynamics within each module and other constraints on the system.

3. ACP-enzyme interactions during polyketide elongation and modification

The function of a *cis*-AT PKS is critically dependent on the intramolecular contacts of the ACP with each of the catalytic domains according to the reaction sequence. The ACP is present at a high effective concentration, perhaps millimolar⁵⁸, therefore strong protein-protein affinity should not be required for throughput and may be disadvantageous. The natural fusion complicates analysis of the intramolecular ACP-enzyme interactions within a module, but the ability to prime catalysis by acylation of KS domains from activated thioester substrates^{59, 60} facilitates biochemical studies with ACPs *in trans*. Several studies have aimed to identify regions of the ACP with specificity for particular enzyme domains. For example, ACP helix II has been implicated in interaction with phosphopantetheinyl transferases⁶¹, loop I and helix II in chain elongation interactions^{62, 63}, helix I in inter-module transfer^{62, 63}, and the helix II-III surface with the AT domain^{64, 65}. It is not clear whether any of the findings are general.

The PikAIII cryo-EM structures provide direct, albeit low-resolution, views of ACP interaction with catalytic domains^{20, 22}. In these structures, where the module was trapped in each of the biochemical states of its catalytic cycle, the ACP position was dependent on the acylation state, leading to the overall conclusion that the substrate plays an important role in the dynamics of the protein-protein interactions over the steps of catalysis within a PKS module. Due to the resolution of the EM reconstructions, it is unknown whether the observed ACP-enzyme interactions are intra- or inter-subunit. Acyl-dependent ACP positioning is consistent with biochemical studies on a variety of ACP-catalytic domain pairs excised from different PKS modules^{59, 66}.

Ketosynthase-ACP interaction.

In the cryo-EM reconstruction of the PikAIII module with its ACP in the methylmalonyl state and in absence of other substrates, methylmalonyl-ACP is localized at the entrance of a previously unidentified channel to the KS active site²⁰ (Figs. 2A & 5A). The channel entrance is on the opposite side of the KS dimer from the docking domain (dd^{KS}). Tryptophan substitutions designed to block the channel were deleterious to module throughput²⁰. Evidence of a second active site entrance also exists in crystal structures of those PKS KS-AT di-domains where two different ACPs interact with the KS during the catalytic cycle. The crystallographic evidence is structural floppiness at the second entrance: missing amino acids, high temperature factors, and weak electron density for loops. In contrast, KSs that do not interact with two ACPs lack such evidence and have well ordered structures at the analogous site.

A role for protein-protein specificity in the ACP delivery of building blocks to the module KS has been investigated for the DEBS pathway in assays of KS extension reactions using an *in trans* malonyl- or methylmalonyl-ACP^{67, 68}. In these studies, the ACPs of modules 3 and 6 were compatible with the cognate KS domain but were incompatible with the KS domains of modules 6 and 3, respectively. Subsequent studies with ACP3/ACP6 chimeras suggested that selectivity comes from the ACP loop 1 – helix 2 region, which encompasses the serine attachment point for Ppant^{62, 63}. This corresponds with the region of the PikAIII methylmalonyl-ACP that contacts the second KS active site entrance²⁰ (Fig. 5A).

Acyltransferase-ACP interaction.

The AT – ACP interaction is of critical importance for selection and loading of the acyl building block to be incorporated during the extension of an intermediate. Again using the DEBS module 3 and 6 domains, excised ATs exhibited an ~tenfold preference for acylation of their cognate ACPs *in trans*⁶⁹. In the cryo-EM structure of PikAIII in a state with the upstream intermediate transferred to the KS (pentaketide-PikAIII), the holo-ACP contacted the AT lid sub-domain; this region was validated by mutagenesis of two charged amino acids on the AT lid^{20, 22}. Variability in the AT-ACP interaction is seen in crystal structures of crosslinked ACP-ATs from two *trans*-AT pathways, where the standalone ATs and cognate ACPs naturally act *in trans*^{64, 65}. The *in trans* AT-ACP interactions differ not only from each other, but also from the PikAIII AT.

Ketoreductase-ACP interactions.

Ketoreductase domains act on β -keto polyketide intermediates to reduce the keto group or in some cases to epimerize an α -methyl substituent. Recently, Broadhurst and coworkers reported affinities of two KR domains for two ACPs excised from the mycolactone PKS, ranging from $\sim 3 \mu\text{M}$ for apo-ACP to $0.2 \mu\text{M}$ for a β -hydroxyacyl-ACP product mimic⁵⁹. Moreover, each KR showed no preference for its cognate ACP. The affinities are unexpectedly strong, given that ACP binding should be transient. For example, the affinities exceed all reported docking-domain affinities^{24, 25}. In the PikAIII cryo-EM study, the ACP contacted the KR only when it carried a KR substrate (β -keto-hexaketide-ACP), and a charge-reversal substitution in the interface reduced module throughput by tenfold²². Interestingly, the ACP contacted the KR lid that covers the NADPH cofactor during catalysis. In the DEBS pathway, the catalytic effectiveness of excised KR - ACP pairs revealed a mixture of situations where, for the majority of KRs, the cognate ACP was most effective ($\sim 10\text{X}$), but in one case a non-cognate ACP was preferred (ACP5-KR6)⁷⁰.

Dehydratase-ACP interactions.

An ACP - DH interaction in a *cis*-AT PKS has not been viewed directly. The entrance to the active site channel and substrate binding site are well established by analogy to bacterial type II FAS DH, which has a nearly identical active site to PKS homologs⁷¹. The ACP - DH interaction was visualized directly in a crystal structure of the *E. coli* type II FAS DH in which the catalytic histidine was crosslinked to an acyl-ACP substrate mimic⁷². However, the highly electronegative ($\text{pI} = 4.0$) type II FAS ACP interacts with a strongly electropositive surface of the FAS DH adjacent to the active site entrance. The analogous surface of *cis*-AT PKS DHs does not appear to be the ACP interaction zone, as it is blocked by the ER in both MAS and the metazoan FAS^{19, 21}. Moreover, strong electrostatic steering does not appear to guide ACP interaction with *cis*-AT PKS DHs as PKS ACPs are not uniformly electronegative ($\text{pI} \approx 5 - 8.5$)^{28, 64, 65, 73}, and the DHs do not have a strongly electropositive surface near the active site entrance. Thus, the ACP-DH interaction in the type II FAS DH cannot be used as a guide for the *cis*-AT PKS systems.

Thioesterase-ACP interactions.

Available data indicate a lack of protein-protein interactions of ACP and TE domains in *cis*-AT PKS. The crystal structures of several TEs established that the active site is at the center of a substrate tunnel^{18, 41, 48} (Fig. 3D), whose entrance and exit sides were defined in structures of the pikromycin TE inactivated with affinity-label substrate mimics^{44, 74}. Biophysical studies demonstrate a striking lack of ACP-TE protein-protein contacts. For example, the DEBS TE had no detectable affinity for its excised cognate ACP in the apo form, whereas the affinity was $\sim 20 \mu\text{M}$ for holo-ACP or acyl-ACP^{66, 75}. Moreover, the addition of TE caused NMR line broadening for the acyl group of acyl-ACP but not for the ACP protein. These data are consistent with many biochemical studies in which a TE was fused to a non-cognate PKS module in order to offload pathway intermediates⁷⁶⁻⁷⁹. Such TEs retain reactivity with non-cognate ACPs and non-natural substrates. Most of the engineered TE chimeric modules have been used to offload small, early-stage intermediates. The situation differs for late-stage intermediates, where the TE can exhibit strong acyl

selectivity. This has been studied most extensively for the pikromycin TE, which is exceptional in its ability to produce both 12- and 14-member macrolactones in the producing organism⁸⁰. However, the Pik TE also has exquisite selectivity for its native hexaketide and heptaketide substrates, and has been proposed as a “gatekeeper” for pathway fidelity⁸¹.

4. Protein-protein interactions during intermediate transfer between modules

Following extension and modification in a *cis*-AT PKS module, the polyketide intermediate is transferred from the module ACP to the KS catalytic cysteine of the next module. For sequential modules that are fused in a single polypeptide, the *in cis* intermediate transfer is aided by the high effective concentration of ACP and KS and a short inter-module linker (typically ~20 amino acids). This was illustrated by genetic manipulation of the producing strains for pikromycin and erythromycin. Separation of the naturally fused modules 1 and 2 reduced pikromycin production more than 1000-fold⁸². Conversely, fusion of the naturally *in trans* polypeptides DEBS1 and DEBS2, or DEBS2 and DEBS3, maintained or even enhanced erythromycin production relative to the wild type⁸³. Inter-module linkers have poorly conserved sequences and are thought to be flexible. When sequential modules are not fused, intermediate transfer occurs *in trans* through the association of “docking domains” at the C-terminus of the upstream ACP (dd^{ACP}) and at the N-terminus of the downstream KS (dd^{KS})^{84, 85}. Docking domain compatibility is essential to pathway throughput. Some cases of selective ACP-KS interaction have been detected, although productive chain transfer between non-cognate ACP-KS pairs can also occur^{86, 87}.

4.1. Structure and selectivity of docking domain interactions

The C-terminal dd^{ACP} and the N-terminal dd^{KS} are extensions to the respective polypeptides, first identified in the DEBS pathway⁸⁴. Docking domains of cognate ACP-KS pairs have a moderate affinity (4 – 100 μ M based on measurements with purified proteins^{24, 25}). Although relatively weak, this is sufficient to prevent non-cognate ACP-KS pairings, nearly all of which have undetectable affinity^{24, 25}. The docking complexes are dimeric ((dd^{ACP})₂(dd^{KS})₂), and affinities are greater in the presence of other dimeric module elements (KS, TE, post-AT DE, post-ACP DE). Structures have been determined for two docking complexes from actinobacterial PKS systems^{23, 24} and two from a cyanobacterial PKS²⁵. In all cases, the weak docking affinity necessitated fusion of the dd^{ACP} C-terminus to the dd^{KS} N-terminus. The actinobacterial and cyanobacterial docking domains were recognized as belonging to different classes by sequence analysis⁸⁸ and have distinctly different structures (Fig. 8).

The current understanding of the actinobacterial docking complex comes from an NMR structure of the DEBS2-DEBS3 complex²³ (comprising the DEBS module 4 dd^{ACP} and module 5 dd^{KS}, Fig. 8A) and a crystal structure of the PikAIII-PikAIV complex²⁴ (between modules 5 and 6 of the pikromycin pathway, Fig. 8B). In each structure, the dd^{KS} is a 30-residue helix that extends from the KS homodimer and forms a parallel coiled-coil with the dd^{KS} of the partner subunit (Figs. 2B, 4A). The single dd^{ACP} helix binds to the dd^{KS} coiled-coil in a parallel orientation to form a four-helix bundle in which two ACPs are tethered to

the bundle N-terminus and the KS dimer to the C-terminus (Fig. 8A,B). The small post-ACP dimerization element (Fig. 4B) is a feature common to actinobacterial *cis*-AT PKS modules with an ACP at the C-terminus. The dimerization element and dd^{ACP} docking helix are connected by a flexible linker and appear to function independently²³.

In each structure, the docking interaction is formed by a core of complementary van der Waals contacts and peripheral polar or ionic interactions^{23, 24}. This overall framework appears to be conserved⁸⁸, although the sequence identity among these and related docking domains is typically low, leading to poor predictability of specific contacts. For example, single charge-reversal substitutions that disrupted docking of DEBS1 and DEBS2 could be rescued by a complementary charge-reversal substitution in the partner protein, however the sites of substitution were outside the DEBS1-DEBS2 docking interface in a model based on the structure of the DEBS2-DEBS3 dock⁸⁹. The size of the docking interface also differs in the docking complex structures from DEBS2-DEBS3 (15-residue dd^{ACP} helix) and PikAIII-PikAIV (9-residue dd^{ACP} helix). Nevertheless docking interactions are selective; binding could be detected only for the expected cognate partners in a study that measured affinities for each of the two DEBS and three Pik dd^{ACP}s with every DEBS and Pik dd^{KS}²⁴.

The cyanobacterial PKS docking interaction is visualized in crystal structures of two docking complexes from the curacin A biosynthetic pathway (CurG-CurH and CurK-CurL)²⁵. The cyanobacterial docking complex differs substantially from the actinobacterial complex. The dd^{KS} consists of two helices; the second KS-proximal helix forms a coiled-coil extension with its partner helix of the KS dimer, as in the actinobacterial dd^{KS}. The dd^{ACP} also consists of two helices, but there appears to be no post-ACP dimerization element. The two dd^{KS} and two dd^{ACP} helices contribute to the docking interface (Fig. 8C). In contrast to the actinobacterial docking complexes, the cyanobacterial ACP and KS domains are tethered at the same end of the complex (Fig. 8C).

The native cyanobacterial docking affinities (5–20 μ M) were similar to those of the actinobacterial docks²⁵. However, the CurK dd^{ACP} was promiscuous and interacted with two non-cognate partners with reduced affinity (20–50 μ M to CurH and CurM dd^{KS} compared to 5 μ M for the cognate CurL dd^{KS}). A charge-reversal substitution (Glu to Arg) in the CurK dd^{ACP} abolished binding to the natural partner CurL dd^{KS} but enhanced binding to the CurH dd^{KS} (9 μ M). This result highlights the importance of charge-charge interactions in docking selectivity and suggests that docking interactions may be engineered when affinity for a non-cognate interaction is already present.

As expected from the structures, docking domain partners can guide biosynthesis when transplanted in heterologous PKS pathways. Pikromycin production was lost *in vivo* by separation of the naturally fused modules 1 and 2 but was recovered by transplantation of docking domains from another actinobacterial PKS⁸². Even the actinobacterial and cyanobacterial docks are interchangeable despite their dissimilar sequences and structures. The CurG-CurH docking domains enabled catalytic throughput when substituted in a two-module PikAIII-PikAIV assay system²⁵. Maintenance of the module dimer enhanced throughput, which was greatest when the chimera retained the native PikAIII post-ACP dimerization element.

A different situation pertains in hybrid pathways where dimeric PKS modules must dock with ordinarily monomeric NRPS modules. For PKS-to-NRPS docks, a novel N-terminal NRPS docking domain (dd^{NRPS}) was detected in several hybrid pathways⁹⁰. The dimeric dd^{NRPS} has a β - α fold and binds C-terminal peptides from the partner module⁹⁰. This docking mode appears to be unique to hybrid pathways and is distinct from docking interactions in purely NRPS pathways⁹.

4.2. ACP-KS interactions during intermediate transfer

The possibility of pathway engineering through combination of non-natural modules inspired studies of KS-ACP interactions during intermediate transfer between modules, most intensively in the DEBS pathway. In assays of forward and back transfer of di- and triketide intermediates between ACP and KS, these studies detected that some DEBS ACPs were selective for their natural downstream partner KS independent of the docking domains, whereas other combinations required only that the dd^{ACP} and dd^{KS} were matched^{86, 87}. Through the use of ACP chimeras (ACP3/ACP2 and ACP4/ACP2), ACP2 helix 1 was identified as a region that conferred selectivity for KS3^{62, 63}. It is unknown whether this result is relevant to other ACP – KS partners. The result may be compatible with the cryo-EM reconstruction of PikAIII fused to the natural partner ACP from PikAII, where the ACP helix 1 is near but does not contact the KS²⁰ (Fig. 5A).

5. ACP-enzyme interactions during β -branching

Several specialized ACP-enzyme interactions are involved in β -branching, a process by which a carbon branch is incorporated in a polyketide by substitution of the β -keto product of the KS extension reaction⁷. Although β -branching is more common in *trans*-AT pathways⁸, many conclusions about the highly specialized enzymatic machinery are applicable to both *trans*-AT and *cis*-AT pathways. The defining β -branching enzyme is a 3-hydroxy-3-methylglutaryl synthase (HMGS), which forms the initial branch from a 3-keto intermediate⁹¹. In systems where a methyl branch is created, the 3-hydroxy-3-methylglutaryl-like HMGS product is dehydrated, decarboxylated and reduced to yield a 3-methyl-butyryl-like intermediate. The HMGS and the dehydrating enzyme (ECH₁) exist as standalone proteins, but in many cases the decarboxylase (ECH₂) and enoylreductase (ER) are domains within a module. Moreover, one or more post-HMGS enzymes may be absent or have alternative functions, and other enzymes such as halogenases may be involved in tailoring the initial intermediate into the final branch, resulting in a wide variety of potential outcomes^{7, 92}. Additionally, β -branching requires specialized ACPs: a specific ACP (“acceptor” ACP, ACP_A) within a module tethers the β -keto intermediate subject to branching, and a second standalone ACP (“donor” ACP, ACP_D) delivers the branch unit to the HMGS. The HMGS is not fused to a module, and β -branching enzymes lack docking domains, thus highly selective ACP-enzyme interactions must facilitate the β -branching reactions.

Details of the ACP_D-HMGS interaction were observed directly in crystal structures of the HMGS from the curacin A pathway²⁸. ACP_D has an unusual 3₁₀ helix III and an adjacent hydrophobic cleft with exquisite surface complementarity to an HMGS helix, creating a

unique interaction that is not possible for other ACPs (Fig. 5B). As with other ACP-enzyme interactions, electrostatic contacts also play a role. The HMGS-ACP_D affinity (0.5 μ M) exceeds that of docking domains. The characteristic features of ACP_D are presumably also recognized by another β -branching enzyme, the KS-decarboxylase (KS_{DC}) that converts malonyl-ACP_D to the acetyl-ACP_D substrate for the HMGS^{91, 93}. Given the significant structural differences between ACP_D and other ACPs, it is unclear what enzyme acylates the ACP_D of *cis*-AT pathways. While the standalone AT of *trans*-AT pathways can load ACP_D *in vitro*^{91, 93}, an AT from primary metabolism or a PKS enzyme with unexpected activity may be required for ACP_D of *cis*-AT pathways.

In contrast to ACP_D, ACP_A must deliver substrates to enzymes within the PKS module, but also must possess unique features that are recognized by the branching enzymes. A tryptophan six amino acids after the serine Ppant attachment site (DSxxxxxW) is strongly predictive for identifying an ACP_A among the ACPs in a pathway, and by extension for identifying the β -branching substrate⁹⁴. Structures of ACP_A from the curacin A *cis*-AT pathway and the mupirocin *trans*-AT pathway reveal that the characteristic tryptophan is embedded in the ACP core and may influence the position of helix III^{94, 95}. An ACP_A complex with a β -branching enzyme has not been viewed directly, but ACP_A mutagenesis and docking studies have suggested that helix II and helix III are involved in recognition by a halogenase domain in the curacin A pathway⁹⁵ and by the HMGS in the mupirocin pathway⁹⁴. The features of the ACP_D group and the ACP_A group are sufficiently distinctive that homologs from heterologous β -branching systems can support catalysis^{26, 94}. Still, it is not understood how β -branching enzymes exclusively select ACP_A or conversely how ACP_A can promiscuously interact with module and β -branching enzymes.

ACP_A often exists as a tandem repeat^{27, 96}, which, in the case of ACP_A in the curacin A pathway, is followed by a small C-terminal dimerization element²⁷. Inactivation of single ACP_A domains within a repeat did not abolish mupirocin production *in vivo*⁹⁶ or the activity of the curacin A β -branching enzymes *in vitro*²⁷. However, the multiple active ACP_As are synergistic²⁷. The curacin A triple tandem ACP_A dimer exhibited weak synergy when only one ACP_A was acylated, indicating that the repeat may generate a high local ACP_A concentration that can interact effectively with the HMGS dimer or can organize the successive *in trans* interactions that are involved in β -branching. Much greater synergy was observed when all domains within the repeat were active, which could result from processive enzymatic action.

Even greater ACP selectivity exists in the myxovirescin pathway, which includes two β -branching systems⁹⁷. The first (ACP_D TaB, HMGS TaC, ECH₁ TaX) inserts a methyl branch in the C12 intermediate carried by ACP5 (ACP_A), and the other (ACP_D TaE, HMGS TaF, ECH₁ TaX, ECH₂ TaY) inserts an ethyl branch in the C16 intermediate carried by ACP7. The separate ACP_D and HMGS proteins serve the two β -branching systems with exceptional selectivity, whereas the ECH₁ and KS_{DC} enzymes act in both systems and the ECH₂ is specific to the ethyl branch at the C16 intermediate⁹³.

Conclusions

The *cis*-AT PKS systems comprise a complex clockwork of both intramolecular and intermolecular protein-protein interactions. Structures of excised catalytic domains, multi-domains, and an intact module together with biochemical data have begun to illuminate the molecular cogs of these fascinating catalytic machines. Protein-protein interactions control the PKS architecture as well as the passage of polyketide intermediates through the catalytic steps of an mPKS module or an iPKS. A dimeric association is fundamental to PKS architecture, supported by dimeric KS, DH, TE and (at least some) ER domains, as well as additional small dimerization elements. However, it is now apparent that other protein-protein interactions are dynamic and may vary among PKS systems.

The ACP domain is by far the most dynamic PKS element, as it interacts with each of the enzyme domains within the PKS. While it is clear that some enzyme domains are selective for the cognate ACP, PKS pathways must balance selectivity with throughput. Enzyme selectivity for the catalytic region of its substrate acyl group seems most beneficial to throughput. In cases of ACP selectivity, rapid dissociation of the acyl product may overcome protein-protein affinity, but any preferential affinity of an ACP for an enzyme domain would be seem to be deleterious to catalytic throughput.

The few direct views of protein-protein interactions in *cis*-AT PKS systems reveal that, beyond the wanderings of the ACP, the mPKS module and the iPKS are remarkably dynamic, perhaps explaining the failure to crystallize a complete PKS module or iPKS. Structures of both the PikAIII module and the MAS modifying region capture each PKS in a range of conformations. For the future, it is important to broaden our view of PKS architecture to other systems and to determine which features of these two structures can be applied broadly and which are specific to the system. The cryo-EM reconstructions of PikAIII in a series of defined biochemical states indicate that the acylation state of the module restricts a large population of conformers to a smaller subset. Thus, both inherent dynamics and natural variation among modules may be features of PKS architecture. Given the dynamic behavior of PikAIII and MAS and the recent technical advances in application of cryo-EM to module-size macromolecules, EM is currently the most promising route to a richer database of structures.

The flexibility of PikAIII, MAS and PKS di-domains resides in the links connecting PKS enzyme domains and not within the enzymes. Many features of these links are remarkably conserved in the metazoan FAS, which for more than a decade has been the 3D model for the *cis*-AT PKS. However, today more structures are available for multi-domain PKS systems than for the metazoan FAS. Perhaps the PKS systems can now inform research on the FAS.

Acknowledgements

Research was supported by US National Institutes of Health grant DK042303 and the Margaret J. Hunter Professorship to JLS. The authors gratefully acknowledge Timm Maier (University of Basel) for sharing a composite model of the MAS and David Akey (University of Michigan) for critical reading of the manuscript.

References

1. Newman DJ and Cragg GM, *J. Nat. Prod.*, 2016, 79, 629–661. [PubMed: 26852623]
2. Cane DE, Walsh CT and Khosla C, *Science*, 1998, 282, 63–68. [PubMed: 9756477]
3. Jenke-Kodama H, Sandmann A, Muller R and Dittmann E, *Mol. Biol. Evol.*, 2005, 22, 2027–2039. [PubMed: 15958783]
4. Smith S and Tsai SC, *Nat. Prod. Rep.*, 2007, 24, 1041–1072. [PubMed: 17898897]
5. Staunton J and Weissman KJ, *Nat. Prod. Rep.*, 2001, 18, 380–416. [PubMed: 11548049]
6. Du L, Sanchez C and Shen B, *Metab. Eng.*, 2001, 3, 78–95. [PubMed: 11162234]
7. Calderone CT, *Nat. Prod. Rep.*, 2008, 25, 845–853. [PubMed: 18820753]
8. Helfrich EJ and Piel J, *Nat. Prod. Rep.*, 2016, 33, 231–316. [PubMed: 26689670]
9. Weissman KJ and Muller R, *ChemBioChem*, 2008, 9, 826–848. [PubMed: 18357594]
10. Xu W, Qiao K and Tang Y, *Crit. Rev. Biochem. Mol. Biol.*, 2013, 48, 98–122. [PubMed: 23249187]
11. Cortes J, Haydock SF, Roberts GA, Bevirt DJ and Leadlay PF, *Nature*, 1990, 348, 176–178. [PubMed: 2234082]
12. Donadio S, Staver MJ, McAlpine JB, Swanson SJ and Katz L, *Science*, 1991, 252, 675–679. [PubMed: 2024119]
13. Keatinge-Clay AT and Stroud RM, *Structure*, 2006, 14, 737–748. [PubMed: 16564177]
14. Tang Y, Kim C-Y, Mathews II, Cane DE and Khosla C, *Proc. Natl. Acad. Sci. USA*, 2006, 103, 11124–11129. [PubMed: 16844787]
15. Keatinge-Clay A, *J. Mol. Biol.*, 2008, 384, 941–953. [PubMed: 18952099]
16. Zheng J, Gay DC, Demeler B, White MA and Keatinge-Clay AT, *Nat. Chem. Biol.*, 2012, 8, 615–621. [PubMed: 22634636]
17. Skiba MA, Sikkema AP, Fiers WD, Gerwick WH, Sherman DH, Aldrich CC and Smith JL, *ACS Chem. Biol.*, 2016, 11, 3319–3327. [PubMed: 27723289]
18. Tsai S-C, Miercke LJ, Krucinski J, Gokhale R, Chen JC-H, Foster PG, Cane DE, Khosla C and Stroud RM, *Proc. Natl. Acad. Sci. USA*, 2001, 98, 14808–14813. [PubMed: 11752428]
19. Maier T, Leibundgut M and Ban N, *Science*, 2008, 321, 1315–1322. [PubMed: 18772430]
20. Dutta S, Whicher JR, Hansen DA, Hale WA, Chemler JA, Congdon GR, Narayan ARH, Håkansson K, Sherman DH, Smith JL and Skiniotis G, *Nature*, 2014, 510, 512–517. [PubMed: 24965652]
21. Herbst DA, Jakob RP, Zhringer F and Maier T, *Nature*, 2016, 531, 533–537. [PubMed: 26976449]
22. Whicher JR, Dutta S, Hansen DA, Hale WA, Chemler JA, Dosey AM, Narayan ARH, Håkansson K, Sherman DH, Smith JL and Skiniotis G, *Nature*, 2014, 510, 560–564. [PubMed: 24965656]
23. Broadhurst RW, Nietlispach D, Wheatcroft MP, Leadlay PF and Weissman KJ, *Chem. Biol.*, 2003, 10, 723–731. [PubMed: 12954331]
24. Buchholz TJ, Geders TW, Bartley FE, Reynolds KA, Smith JL and Sherman DH, *ACS Chem. Biol.*, 2009, 4, 41–52. [PubMed: 19146481]
25. Whicher JR, Smaga SS, Hansen DA, Brown WC, Gerwick WH, Sherman DH and Smith JL, *Chem. Biol.*, 2013, 20, 1340–1351. [PubMed: 24183970]
26. Buchholz TJ, Rath CM, Lopanik NB, Gardner NP, Håkansson K and Sherman DH, *Chem. Biol.*, 2010, 17, 1092–1100. [PubMed: 21035732]
27. Gu L, Eisman EB, Dutta S, Franzmann TM, Walter S, Gerwick WH, Skiniotis G and Sherman DH, *Angew. Chem.*, 2011, 50, 2795–2798. [PubMed: 21387490]
28. Maloney FP, Gerwick L, Gerwick WH, Sherman DH and Smith JL, *Proc. Natl. Acad. Sci. USA*, 2016, 113, 10316–10321. [PubMed: 27573844]
29. Akey DL, Gehret JJ, Khare D and Smith JL, *Nat. Prod. Rep.*, 2012, 29, 1038–1049. [PubMed: 22498975]
30. Keatinge-Clay AT, *Nat. Prod. Rep.*, 2012, 29, 1050–1073. [PubMed: 22858605]
31. Staunton J, Caffrey P, Aparicio JF, Roberts GA, Bethell SS and Leadlay PF, *Nat. Struct. Mol. Biol.*, 1996, 3, 188–192.

32. Kao CM, Pieper R, Cane DE and Khosla C, *Biochemistry*, 1996, 35, 12363–12368. [PubMed: 8823171]
33. Maier T, Jenni S and Ban N, *Science*, 2006, 311, 1258–1262. [PubMed: 16513975]
34. Brignole EJ, Smith S and Asturias FJ, *Nat. Struct. Mol. Biol.*, 2009, 16, 190. [PubMed: 19151726]
35. Benning FM, Sakiyama Y, Mazur A, Bukhari HS, Lim RY and Maier T, *ACS Nano*, 2017, 11, 10852–10859. [PubMed: 29023094]
36. 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]
37. Tang Y, Chen AY, Kim C-Y, Cane DE and Khosla C, *Chem. Biol.*, 2007, 14, 931–943. [PubMed: 17719492]
38. Akey DL, Razelun JR, Tehranisa J, Sherman DH, Gerwick WH and Smith JL, *Structure*, 2010, 18, 94–105. [PubMed: 20152156]
39. Faille A, Gavalda S, Slama N, Lherbet C, Maveyraud L, Guillet V, Laval F, Quémard A, Mourey L and Pedelacq J-D, *J. Mol. Biol.*, 2017, 429, 1554–1569. [PubMed: 28377293]
40. Gay D, You Y-O, Keatinge-Clay A and Cane DE, *Biochemistry*, 2013, 52, 8916–8928. [PubMed: 24274103]
41. Tsai S-C, Lu H, Cane DE, Khosla C and Stroud RM, *Biochemistry*, 2002, 41, 12598–12606. [PubMed: 12379102]
42. Khare D, Hale WA, Tripathi A, Gu L, Sherman DH, Gerwick WH, Håkansson K and Smith JL, *Structure*, 2015, 23, 2213–2223. [PubMed: 26526850]
43. Zheng J, Fage CD, Demeler B, Hoffman DW and Keatinge-Clay AT, *ACS Chem. Biol.*, 2013, 8, 1263–1270. [PubMed: 23489133]
44. Akey DL, Kittendorf JD, Giraldes JW, Fecik RA, Sherman DH and Smith JL, *Nat. Chem. Biol.*, 2006, 2, 537–542. [PubMed: 16969372]
45. Sung KH, Berkhan G, Hollmann T, Wagner L, Blankenfeldt W and Hahn F, *Angew. Chem. Int. Ed.*, 2018, 57, 343–347.
46. Keatinge-Clay AT, *Cell Chem. Biol.*, 2016, 23, 540–542. [PubMed: 27203374]
47. Du L and Lou L, *Nat. Prod. Rep.*, 2010, 27, 255–278. [PubMed: 20111804]
48. Scaglione JB, Akey DL, Sullivan R, Kittendorf JD, Rath CM, Kim ES, Smith JL and Sherman DH, *Angew. Chem.*, 2010, 49, 5726–5730. [PubMed: 20623733]
49. Li X, Sevillano N, La Greca F, Deis L, Liu Y-C, Deller MC, Mathews II, Matsui T, Cane DE, Craik CS and Khosla C, *J. Am. Chem. Soc.*, 2018, 140, 6518–6521. [PubMed: 29762030]
50. Stanfield RL, Zemla A, Wilson IA and Rupp B, *J. Mol. Biol.*, 2006, 357, 1566–1574. [PubMed: 16497332]
51. Skiba MA, Sikkema AP, Moss NA, Lowell AN, Su M, Sturgis RM, Gerwick L, Gerwick WH, Sherman DH and Smith JL, *ACS Chem. Biol.*, 2018, 13, 1640–1650. [PubMed: 29701944]
52. Chen AY, Cane DE and Khosla C, *Chem. Biol.*, 2007, 14, 784–792. [PubMed: 17656315]
53. Price AC, Zhang YM, Rock CO and White SW, *Structure*, 2004, 12, 417–428. [PubMed: 15016358]
54. Wagner DT, Stevens DC, Mehaffey MR, Manion HR, Taylor RE, Brodbelt JS and Keatinge-Clay AT, *Chem. Commun.*, 2016, 52, 8822–8825.
55. Bonnett SA, Whicher JR, Papireddy K, Florova G, Smith JL and Reynolds KA, *Chem. Biol.*, 2013, 20, 772–783. [PubMed: 23790488]
56. Yuzawa S, Kapur S, Cane DE and Khosla C, *Biochemistry*, 2012, 51, 3708–3710. [PubMed: 22509729]
57. Edwards AL, Matsui T, Weiss TM and Khosla C, *J. Mol. Biol.*, 2014, 426, 2229–2245. [PubMed: 24704088]
58. Kuriyan J and Eisenberg D, *Nature*, 2007, 450, 983–990. [PubMed: 18075577]
59. Moretto L, Vance S, Heames B and Broadhurst RW, *Chem. Commun.*, 2017, 53, 11457–11460.
60. Hansen DA, Rath CM, Eisman EB, Narayan AR, Kittendorf JD, Mortison JD, Yoon YJ and Sherman DH, *J. Am. Chem. Soc.*, 2013, 135, 11232–11238. [PubMed: 23866020]

61. Weissman KJ, Hong H, Popovic B and Meersman F, *Chem. Biol*, 2006, 13, 625–636. [PubMed: 16793520]
62. Kapur S, Lowry B, Yuzawa S, Kenthirapalan S, Chen AY, Cane DE and Khosla C, *Proc. Natl. Acad. Sci. USA*, 2012, 109, 4110–4115. [PubMed: 22371562]
63. Kapur S, Chen AY, Cane DE and Khosla C, *Proc. Natl. Acad. Sci. USA*, 2010, 107, 22066–22071. [PubMed: 21127271]
64. Miyanaga A, Iwasawa S, Shinohara Y, Kudo F and Eguchi T, *Proc. Natl. Acad. Sci. USA*, 2016, 113, 1802–1807. [PubMed: 26831085]
65. Miyanaga A, Ouchi R, Ishikawa F, Goto E, Tanabe G, Kudo F and Eguchi T, *J. Am. Chem. Soc*, 2018, 140, 7970–7978. [PubMed: 29870659]
66. Tran L, Tosin M, Spencer JB, Leadlay PF and Weissman KJ, *ChemBioChem*, 2008, 9, 905–915. [PubMed: 18348128]
67. Kim CY, Alekseyev VY, Chen AY, Tang Y, Cane DE and Khosla C, *Biochemistry*, 2004, 43, 13892–13898. [PubMed: 15518537]
68. Chen AY, Schnarr NA, Kim CY, Cane DE and Khosla C, *J. Am. Chem. Soc*, 2006, 128, 3067–3074. [PubMed: 16506788]
69. Wong FT, Chen AY, Cane DE and Khosla C, *Biochemistry*, 2010, 49, 95–102. [PubMed: 19921859]
70. Ostrowski MP, Cane DE and Khosla C, *J. Antibiot*, 2016, 69, 507–510. [PubMed: 27118242]
71. Leesong M, Henderson BS, Gillig JR, Schwab JM and Smith JL, *Structure*, 1996, 4, 253–264. [PubMed: 8805534]
72. Nguyen C, Haushalter RW, Lee DJ, Markwick PR, Bruegger J, Caldara-Festin G, Finzel K, Jackson DR, Ishikawa F and O'Dowd B, *Nature*, 2014, 505, 427–431. [PubMed: 24362570]
73. Alekseyev VY, Liu CW, Cane DE, Puglisi JD and Khosla C, *Protein Sci*, 2007, 16, 2093–2107. [PubMed: 17893358]
74. Giraldes JW, Akey DL, Kittendorf JD, Sherman DH, Smith JL and Fecik RA, *Nat. Chem. Biol*, 2006, 2, 531–536. [PubMed: 16969373]
75. Tran L, Broadhurst RW, Tosin M, Cavalli A and Weissman KJ, *Chem. Biol*, 2010, 17, 705–716. [PubMed: 20659683]
76. Cortes J, Wiesmann KE, Roberts GA, Brown MJ, Staunton J and Leadlay PF, *Science*, 1995, 268, 1487–1489. [PubMed: 7770773]
77. Pieper R, Luo G, Cane DE and Khosla C, *Nature*, 1995, 378, 263–266. [PubMed: 7477343]
78. Chen S, Xue Y, Sherman DH and Reynolds KA, *Chem. Biol*, 2000, 7, 907–918. [PubMed: 11137814]
79. Menzella HG, Reid R, Carney JR, Chandran SS, Reisinger SJ, Patel KG, Hopwood DA and Santi DV, *Nat. Biotechnol*, 2005, 23, 1171–1176. [PubMed: 16116420]
80. Xue Y, Zhao L, Liu HW and Sherman DH, *Proc. Natl. Acad. Sci. USA*, 1998, 95, 12111–12116. [PubMed: 9770448]
81. Hansen DA, Koch AA and Sherman DH, *J. Am. Chem. Soc*, 2017, 139, 13450–13455. [PubMed: 28836772]
82. Yan J, Gupta S, Sherman DH and Reynolds KA, *ChemBioChem*, 2009, 10, 1537–1543. [PubMed: 19437523]
83. Squire CM, Goss RJ, Hong H, Leadlay PF and Staunton J, *ChemBioChem*, 2003, 4, 1225–1228. [PubMed: 14613116]
84. Gokhale RS, Tsuji SY, Cane DE and Khosla C, *Science*, 1999, 284, 482–485. [PubMed: 10205055]
85. Tsuji SY, Cane DE and Khosla C, *Biochemistry*, 2001, 40, 2326–2331. [PubMed: 11327852]
86. Wu N, Tsuji SY, Cane DE and Khosla C, *J. Am. Chem. Soc*, 2001, 123, 6465–6474. [PubMed: 11439032]
87. Wu N, Cane DE and Khosla C, *Biochemistry*, 2002, 41, 5056–5066. [PubMed: 11939803]
88. Thattai M, Burak Y and Shraiman BI, *PLoS Comp. Biol*, 2007, 3, 1827–1835.
89. Weissman KJ, *ChemBioChem*, 2006, 7, 1334–1342. [PubMed: 16871615]

90. Richter CD, Nietlispach D, Broadhurst RW and Weissman KJ, *Nat. Chem. Biol.*, 2008, 4, 75–81. [PubMed: 18066054]
91. Calderone CT, Kowtoniuk WE, Kelleher NL, Walsh CT and Dorrestein PC, *Proc. Natl. Acad. Sci. USA*, 2006, 103, 8977–8982. [PubMed: 16757561]
92. Gu L, Wang B, Kulkarni A, Geders TW, Grindberg RV, Gerwick L, Hakansson K, Wipf P, Smith JL, Gerwick WH and Sherman DH, *Nature*, 2009, 459, 731–735. [PubMed: 19494914]
93. Calderone CT, Iwig DF, Dorrestein PC, Kelleher NL and Walsh CT, *Chem. Biol.*, 2007, 14, 835–846. [PubMed: 17656320]
94. Haines AS, Dong X, Song Z, Farmer R, Williams C, Hothersall J and Posko, *Nat. Chem. Biol.*, 2013, 9, 685–692. [PubMed: 24056399]
95. Busche A, Gottstein D, Hein C, Ripin N, Pader I, Tufar P, Eisman EB, Gu L, Walsh CT, Sherman DH, Lhr F, Gntert P and Dtsch V, *ACS Chem. Biol.*, 2012, 7, 378–386. [PubMed: 22103656]
96. Rahman AS, Hothersall J, Crosby J, Simpson TJ and Thomas CM, *J. Biol. Chem.*, 2005, 280, 6399–6408. [PubMed: 15583005]
97. Simunovic V, Zapp J, Rachid S, Krug D, Meiser P and Muller R, *ChemBioChem*, 2006, 7, 1206–1220. [PubMed: 16835859]

A. Metazoan FAS

B. Iterative PKS modifying region

Fig. 1.

Comparison of metazoan FAS and a PKS modifying region. (A) Porcine FAS crystal structure (PDB 2VZ8)¹⁹. (B) Modifying region of MAS, an iterative PKS (based on PDB 5BP1)²¹. The KS (blue), DH (yellow) and ER (beige) domains are dimeric, whereas the AT (green), KR (purple) and ΨMT (orange) domains have no homomeric contacts. The PKS and FAS DH dimers have different overall shapes, but the domain N-termini form the dimer interface in both megasynthases. The PKS KR domain has no contacts with either DH or ER whereas many inter-domain contacts exist in the FAS. The post-AT linker is rendered in red, and KS-AT linker domain is in gray. Helices are rendered as cylinders and β-strands as arrows.

A. PikAIII

B. DEBS KS-AT

Fig. 2.

Comparison of PikAIII module to an excised KS-AT di-domain. **(A)** Methylmalonyl-PikAIII (EMDB 5653)²⁰. Methylmalonyl-ACP is docked at the bottom entrance to the active site. The domains are colored as in Fig. 1, with ACPs in red. **(B)** KS-AT from DEBS3 module 5 (PDB 2HG4)¹⁴. The linker-AT (gray-green) is differently oriented relative to the KS dimer (blue) in two structures. The post-AT linker (red in B) was not visible in the cryo-EM reconstruction of PikAIII.

Fig. 3. PKS dimer-forming catalytic domains. **(A)** KS dimer from DEBS2 module 3 (PDB 2QO3)³⁷. The catalytic Cys, His and Glu side chains are shown as spheres in atomic colors with gray C atoms. **(B)** DH dimer from CurK (PDB 3KG9)³⁸. The catalytic His and Asp side chains are shown as spheres in atomic colors with gray C atoms. **(C)** ER dimer from MAS (based on PDB 5BP4)²¹. The NADPH cofactor is shown as spheres with gray C atoms. **(D)** TE dimer from PikaIV (PDB 2HFJ)⁴⁴. The substrate tunnel in each subunit is rendered as a gray surface. All dimers are viewed approximately along the twofold symmetry axis. Each dimer is colored in the same hue as in Fig. 1, with monomers in different brightness.

Fig. 4. Small PKS dimerization elements. (A) dd^{KS} from PikaIV (PDB 3F5H)²⁴. (B) post-ACP dimerization element from DEBS2 (PSB 1PZQ)²³. (C) post-AT dimerization element from SpnC module 3 (PDB 4IMP)⁴³. The dimerization elements are colored in the same hue as an adjacent catalytic domain (Fig. 1), with monomers in different brightness. Adjacent domains are shown as circles. The twofold symmetry axes are vertical in these images.

A. PikAIII KS

B. HMGS from curacin pathway

Fig. 5.

ACP interaction with KS and HMGS. The PKS KS and β -branching HMGS are members of the thiolase family. Here they are depicted in identical orientation with the dimer axis vertical on the page with monomers colored in shades of blue (KS) and cyan (HMGS). **(A)** PikAIII KS dimer with upstream and module ACPs as observed in functional complexes²⁰. The active site cavity is shown as a solid surface inside the KS in this composite image from two structures. In the complex with pentaketide-ACP_{*i-1*}, the upstream ACP_{*i-1*} from the preceding module (dull red) delivers the pentaketide product of the previous module to a site near the active site side entrance (arrow). In the complex with methylmalonyl-ACP_{*i*}, the module ACP_{*i*} (bright red) delivers a methylmalonyl building block to the lower active site entrance. **(B)** HMGS from the curacin A pathway in complex with holo-ACP_D (gold)²⁸. The holo-ACP Ppant (rendered as sticks with gray C atoms) enters the active site through the side entrance, analogous to the KS side entrance.

Fig. 6. Fab-KS-AT complex. In the top view, the dimer axis is vertical. In the bottom view, the structure is rotated 45° towards the viewer to illustrate the separation of the AT and the Fab (arrow). Domains within the DEBS2 module 3 KS-AT dimer (PDB 6C9U)⁴⁹ are colored as in Fig. 2B, and the Fab heterodimer (V_L-C_L, V_H-C_H) is colored with the variable domains (V_L and C_L) in magenta and the constant domains (V_H and C_H) in cyan.

Fig. 7. Insertion of catalytic domains between KR structure elements. **(A)** Composition of canonical *cis*-AT PKS extension modules. Domains are represented by circles colored as in Fig. 1 and Fig. 2, and the strands of the KR β -zipper as arrows. The schematic illustrates how MT and ER domains, when present, are inserted within the KR. The gray circle is the linker domain between KS and AT, and the post-AT linker peptide is represented as a red line. Docking domains and the small dimerization elements are not depicted. **(B)** Structure of KR from Plm1 (PDB 4IMP)⁵⁵. The two β -zipper strands contribute to the core β -sheets of both KR_S (light purple) and KR_C (dark purple). Approximate insertion points for MT and ER in other modules are indicated by arrows.

Fig. 8. PKS docking complexes. (A) DEBS2-DEBS3 dock (PDB 1PZR)²³. (B) PikAIII-PikAIV dock (PDB 3F5H)²⁴. (C) CurG-CurH dock (PDB 4MY Y)²⁵. The dd^{KS} dimers are in blue and the dd^{ACP}s in red and mauve with the corresponding KS and ACP domains in schematic form. Polypeptide N and C-termini are labeled.

Table 1.

Dimer-forming PKS domains

Domain	Abbreviation	Size *	Interface area (Å ²)	References
Ketosynthase	KS	425	1800	14, 25, 36, 37
Dehydratase	DH	280	550–700	15, 38–40
Enoylreductase	ER	325	1400	16, 21, 42
Thioesterase	TE	275	750–850	18, 41
KS docking domain	dd ^{KS}	30	850	23–25
Post-AT dimerization element	post-AT DE	50	990	43
Post-ACP dimerization element	post-ACP DE	35	1100	23

*
number of amino acids