Cite this: *Med. Chem. Commun.*,
2019, **10**, 1256

Diversification of polyketide structures via synthase engineering

Taylor Kornfuehrer and Alessandra S. Eustáquio *

Polyketide natural products possess diverse biological activities including antibiotic, anticancer, and immunosuppressive. Their equally varied and complex structures arise from head-to-tail condensation of simple carboxyacyl monomers. Since the seminal discovery that biosynthesis of polyketides such as the macrolide erythromycin is catalyzed by uncharacteristically large, multifunctional enzymes, termed modular type I polyketide synthases, chemists and biologists alike have been inspired to harness the apparent modularity of the synthases to further diversify polyketide structures. Yet, initial attempts to perform “combinatorial biosynthesis” failed due to challenges associated with maintaining the structural and catalytic integrity of large, chimeric synthases. Fast forward nearly 30 years, and advancements in our understanding of polyketide synthase structure and function have allowed the field to make significant progress toward effecting desired modifications to polyketide scaffolds in addition to engineering small, chiral fragments. This review highlights selected examples of polyketide diversification *via* control of monomer selection, oxidation state, stereochemistry, and cyclization. We conclude with a perspective on the present and future of polyketide structure diversification and hope that the examples presented here will encourage medicinal chemists to embrace polyketide synthetic biology as a means to revitalize polyketide drug discovery.

Received 8th March 2019,
Accepted 9th May 2019

DOI: 10.1039/c9md00141g

rsc.li/medchemcomm

1. Introduction

Polyketides are a structurally diverse group of natural products with equally variable bioactivity. They exhibit intricate structures featuring complex stereochemistry and diverse functionality. Moreover, they provide privileged scaffolds for drug discovery as the structures have evolved to interact with biological systems. As such, many polyketides have been developed into pharmaceuticals such as the antibiotic erythromycin A, the antifungal agent amphotericin B, the immunosuppressant rapamycin, and the antiparasitic avermectin (Fig. 1).

While polyketides have a firmly established role in treating human diseases, polyketide drug discovery and development efforts have waned in recent years, which is on trend with natural product drug discovery as a whole.^{1,2} The interplay of supply issues³ and challenges faced by the synthetic chemist seeking to modify or fully synthesize a polyketide scaffold⁴ have contributed to the deprioritization of natural product research by pharmaceutical companies who opted to redirect their resources to efforts in combinatorial chemistry.⁵ An overall decline in the rate of drug discovery was observed concomitantly with the deprioritization of natural product research. The intrinsic advantage of natural product scaffolds in interacting with biological targets makes it worthwhile to overcome their inherent challenges.⁶

Fortuitously, the biosynthetic logic and machinery that guides the formation of polyketides offers the opportunity to

Department of Medicinal Chemistry and Pharmacognosy and Center for Biomolecular Sciences, College of Pharmacy, University of Illinois at Chicago, Chicago, Illinois 60607, USA. E-mail: ase@uic.edu; Tel: +1 3124137082

Taylor Kornfuehrer received her B.Sc. in Chemistry from The University of Texas at Austin in 2017. She then began her graduate studies in Pharmacognosy at the University of Illinois at Chicago and joined the lab of Dr. Alessandra S. Eustáquio. Taylor is currently working on projects related to enzymology and engineered polyketide biosynthesis.

Alessandra S. Eustáquio is an Assistant Professor of Medicinal Chemistry and Pharmacognosy at the College of Pharmacy, University of Illinois at Chicago. She obtained her B.Sc. in Pharmacy and Biochemistry from the University of São Paulo, Brazil, and Ph.D. in Pharmacognosy from the University of Tübingen, Germany. After postdoctoral training at the Scripps Institution of Oceanography, University of California San Diego, she joined the Natural Products group of Pfizer's Medicinal Chemistry as a Principal Scientist. Her research currently focuses on natural product biosynthesis and the development of synthetic biology tools to facilitate access to natural and engineered compounds.

Fig. 1 Chemical structures of the clinically-relevant polyketides amphotericin B, erythromycin A, rapamycin, and avermectin B1. They are biosynthesized by modular type I polyketide synthases and exhibit many characteristic features of the type I polyketide class such as i) carbon backbones exhibiting an alternating oxygenation pattern, ii) macrocyclic structures, iii) elegant stereochemistry, and iv) diverse functionality, including O-glycosidic bonds to unusual sugar moieties.

revitalize polyketide drug discovery *via* synthetic biology strategies. Polyketides are diverse members of a common biosynthetic class formed by multifunctional polyketide synthase (PKS) enzymes that combine simple carboxyacyl building blocks, or monomers, in a head-to-tail fashion to form complex molecular scaffolds.⁷ Compounds such as those shown in Fig. 1 are biosynthesized by a distinct subclass of PKS enzymes, called modular type I PKSs.⁸ They are distinguished from other subclasses of PKSs – notably, iterative type I PKS and dissociable type II PKS – by their multimodular functional organization and assembly line-style architecture, which are both characteristics that impart utility for synthetic biology.⁹ Two evolutionarily distinct subtypes of modular type I PKSs are known and commonly termed *cis*- and *trans*-acyltransferase PKSs. Only polyketides biosynthesized by the prototypical modular *cis*-acyltransferase PKSs are discussed in

this review, and the term PKS is used to mean this subtype unless otherwise stated.

This review describes how the logic of PKS biosynthesis facilitates the rational engineering of chemical diversity and provides recent examples of polyketide structure modifications achieved *via* synthetic biology. We did not attempt to be exhaustive but have rather decided to present examples that illustrate the types of structure modifications that can be accomplished. With medicinal chemists in mind, our goal was to focus more on the compounds themselves rather than the syntheses. Therefore, we deliberately left out most details regarding the exact modifications done to the synthases to accomplish the desired polyketide structural change. The reader is directed to the original citations for such details, if interested, and to several detailed reviews regarding strategies and challenges for PKS engineering.^{10–13}

We start this review with a section on the rationale for polyketide synthetic biology that includes a brief introduction to PKS architecture and mechanisms of polyketide biosynthesis. After giving specific examples of a) monomer diversification, b) control of the oxidation state and stereochemistry, and c) engineering of cyclized or linear products, we offer our perspective on the present and the future of PKS engineering as it relates to both natural product drug discovery and medicinal chemistry.

2. Rationale for polyketide synthetic biology

The opportunity to utilize PKSs for production of novel chemistry was recognized upon the initial discovery^{14,15} of the catalytic multifunctionality and modular organization of PKS genes involved in the biosynthesis of the macrolide erythromycin (Fig. 2). The PKS enzymatic machinery facilitates access to complex, diverse chemical space from simple monomeric building blocks. Further, the modular organization and assembly line-style architecture of the synthases allow rational prediction of chemical products from an analysis of genetic sequence – known as the principle of collinearity (although exceptions to the collinearity rule do exist¹⁶). Therefore, in principle, it should also be possible to engineer the biosynthetic assembly line to predictably effect a desired chemical change to the final polyketide product.

However, many early attempts to perform “combinatorial biosynthesis” with PKS machinery largely yielded nonfunctional chimeric assembly lines, and those chimeras that were functional suffered from low yields.¹⁷ Developments in PKS structural biology over the last decade – including the first reported structure of an intact PKS module¹⁸ and its rearrangements during a catalytic cycle¹⁹ – have dramatically increased our understanding of PKS structure and function. As we will show in this review, the field is now poised to achieve a number of targeted modifications to a polyketide scaffold in addition to being able to generate small fragments using highly combinatorial PKS chimeras.²⁰

Fig. 2 Organization of the 6-deoxyerythronolide B synthase (DEBS), the PKS responsible for biosynthesis of the erythromycin aglycone. DEBS is encoded by three genes – DEBS1, DEBS2, and DEBS3 – each of which encodes two extension modules. Docking domains located between the three DEBS subunits facilitate self-assembly of the complete PKS assembly line. Note that each module is responsible for incorporation and processing of a ketide monomer in an assembly line fashion. To start, propionate – activated as a coenzyme A (CoA) thioester – is incorporated by the N-terminal loading module. Each extension module then adds a single (2S)-methylmalonate monomer to the growing polyketide intermediate as it proceeds down the assembly line. The C-terminal TE domain catalyzes the release of the polyketide chain *via* macrocyclization. The KR of module 3 is redox-incompetent (KR₀) (AT = acyltransferase, AT_L = loading acyltransferase, ACP = acyl carrier protein, KS = ketosynthase, KR = keto-reductase, DH = dehydratase, ER = enoylreductase, TE = thioesterase).

The following sections will introduce polyketide synthases including the overall architecture and the enzymatic functionalities – or domains – that impart activity for monomer selection, β -carbon processing, α and β stereochemistry, and macrocyclization. For more detailed information, the reader is directed to excellent reviews of polyketide synthase logic/organization,^{7,21} structure and mechanisms,^{22–24} and stereocontrol.^{25,26}

2.1 Polyketide synthase architecture

Representing many of the largest known proteins, PKSs are multifunctional megaenzymes with modular organization.²³ Each module carries out a single round of polyketide chain extension *via* an enzyme-catalyzed decarboxylative Claisen-type condensation of the growing polyketide chain with a new monomeric extension unit. The modules are composed of multiple catalytic domains that each play a distinct role in an extension cycle (Fig. 2).

The N-terminus of the PKS assembly line contains a loading module responsible for selecting the first monomer of the polyketide chain, thereby priming the assembly line for successive rounds of chain extension. The basic catalytic domains required of a module for polyketide chain extension

are acyltransferase (AT) for monomer selection, acyl carrier protein (ACP) for substrate shuttling and activation, and ketosynthase (KS) for decarboxylative condensation. In addition to these essential domains, each module may also contain one or more processing domains – ketoreductase (KR), dehydratase (DH), and enoylreductase (ER) – which act sequentially to either partially or fully reduce the β -carbonyl functionality of the growing polyketide intermediate. The assembly line also contains a C-terminal thioesterase (TE) domain, which catalyzes the release of the polyketide product from the assembly line.

In addition to the catalytic domains, there are critical structural domains associated with the mediation of protein–protein interactions. Because multiple genes are typically involved in encoding the entire PKS, the assembly line may be made of discrete proteins, termed here subunits. For example, 6-deoxyerythronolide B synthase (DEBS), which biosynthesizes the aglycone core of erythromycin, is made of three polypeptide subunits – DEBS1, DEBS2, and DEBS3 (Fig. 2). Regions on the extreme termini of PKS subunits, called docking domains, are responsible for facilitating recognition between the discrete subunits and subsequent self-assembly of the complete PKS assembly line.^{27,28} Additionally, within a given subunit, there are crucial linker regions

that not only connect modules on the same protein but also serve to facilitate dynamic interactions between individual catalytic domains while maintaining the structural integrity of the overall protein.^{18,23}

2.2 Mechanisms of polyketide biosynthesis

2.2.1 Loading. Polyketide biosynthesis is initiated at the N-terminus of the assembly line with selection of a priming monomer, or starter unit, by the loading module. Diverse types of domain organizations exist among PKS loading modules, and their variation in enzymatic functionality allows each type to select monomers of slightly different chemical classes. In the DEBS assembly line, loading is achieved by an AT_L-ACP didomain, though other notable loading module architectures include KS^Q-AT-ACP, A_L-ACP, and GNAT_L-ACP.²¹ While their domain organizations and starter unit selectivity may differ, each type of loading module yields an assembly line that is primed with an ACP-bound thioester (Fig. 3).

AT_L-ACP modules (Fig. 3A) load CoA-bound thioesters. Examples include the erythromycin PKS loading module that primes the assembly line with propionyl-CoA and the avermectin PKS loading module that loads isobutyryl-CoA.^{29,30} Relative to other loading machinery, AT_L-ACP modules can select a broad range of acyl-CoA substrates, with some AT_L domains being promiscuous for multiple starter units.

KS^Q-AT-ACP modules (Fig. 3B) load carboxyacyl-CoA starter units – either malonyl-CoA or (2*S*)-methylmalonyl-CoA. An example of a PKS containing such a loading module is that of amphotericin B.³¹ KS^Q is a condensation-incompetent KS domain capable of decarboxylating the ACP-bound starter

unit to yield either an acetyl group (from malonyl-CoA) or a propionyl group (from (2*S*)-methylmalonyl-CoA) to start the polyketide chain. The “Q” superscript refers to a highly conserved glutamine residue that replaces the active-site cysteine residue found in KS domains, making KS^Q condensation-incompetent.³²

Some assembly lines utilize a loading adenylation (A_L) domain that is equivalent to the adenylation domains of non-ribosomal peptide synthase (NRPS) assembly lines.²¹ A_L-ACP modules (Fig. 3C) can load a free carboxylic acid from the shikimate pathway, as seen in the rapamycin PKS.³³ The biosynthetic nature of the starter units A_L-ACP modules can load – in addition to their accepting free carboxylic acid monomers – makes them noteworthy loading modules to the synthetic biologist.

The least common type of loading module utilizes a GCN5 *N*-acetyltransferase-like (GNAT_L) domain. Similar to KS^Q-AT-ACP modules, GNAT_L-ACP modules load carboxyacyl-CoA starter units to prime the assembly line. However, the GNAT_L domain itself decarboxylates the starter unit. Further, these unusual loading modules sometimes contain *C*-methyltransferases which may install one or more methyl groups on the carboxyacyl starter unit (Fig. 3D). In the gephyronic acid PKS, the malonyl-CoA starter unit is decarboxylated to yield an acetyl-ACP, which is then dimethylated at the α-carbon to yield an assembly line primed with isobutyryl. Alternatively, dimethylation may precede decarboxylation.^{21,34}

2.2.2 Extension. A single round of polyketide chain elongation, or extension, involves translocation of the growing polyketide intermediate to the downstream module where it

Fig. 3 Different domain architectures found in PKS loading modules. (A) AT_L-ACP: AT_L primes the assembly line by selecting an acylthioester-CoA – such as propionyl- or isobutyryl-CoA – and charging ACP with the starter unit. (B) KS^Q-AT-ACP: AT selects a malonyl- or methylmalonyl-CoA monomer and charges ACP with the starter unit. The condensation-incompetent KS^Q domain decarboxylates the ACP-bound starter unit yielding an assembly line primed with acetyl- or propionyl-ACP. (C) A_L-ACP: the adenylation A_L domain prepares a shikimate-derived carboxylic acid monomer for loading in an ATP-dependent manner and loads the activated monomer to ACP. Notably, AT_L and A_L can be embedded in the assembly line or provided as a separate protein (*in trans*). (D) GNAT_L-ACP: the GNAT_L domain selects for a carboxyacyl-CoA starter unit, which it decarboxylates then transfers to ACP. Some GNAT_L loading modules also feature *C*-methyltransferases (CMTs), which may methylate the starter unit one or more times. In the above example, the CMT domain methylates the acetyl-ACP starter unit (from malonyl-CoA) twice at the α-carbon to yield an assembly line primed with isobutyryl-CoA. Such a loading module exists in the gephyronic acid PKS, but GNAT_L modules may contain CMTs that exhibit an alternative methylation pattern or may not contain a CMT at all.

undergoes a KS-catalyzed decarboxylative condensation with a new monomer that is attached to the ACP (Fig. 4). The AT domain of that module acts as a gatekeeper and is responsible for selection of the new monomer, often referred to as an extender unit. Mechanistically, acylation of the AT by an extender unit proceeds in a manner analogous to AT_L acylation (Fig. 3A) by a starter unit. However, the monomer selected by the AT of an extension module is a carboxyacyl-CoA whereas an acyl-CoA starter unit is selected by AT_L . The most commonly used extender unit is malonyl-CoA or (2*S*)-methylmalonyl-CoA, although functionalized extender units such as chloroethylmalonyl-CoA, (2*R*)-hydroxymalonyl-ACP, and (2*S*)-aminomalonyl-ACP also exist.^{35,36}

After the AT selects the appropriate substrate, it loads it onto the downstream ACP, which contains a phosphopantethiene prosthetic group that serves as a flexible tethering point for both the carboxyacyl monomer and the growing polyketide intermediate. The KS domain is responsible for condensing the polyketide intermediate from the upstream module onto the downstream ACP-bound extender unit *via* a decarboxylative Claisen-type condensation.

Following condensation, the new ACP-bound intermediate may undergo full or partial reductive processing. This is mediated by the optional processing domains KR, DH, and ER – collectively termed the reductive loop (even though DH is not an oxidoreductase but rather a lyase). The domains of the reductive loop act in sequence to reduce the β -carbonyl functionality on the ACP-bound intermediate to a β -hydroxyl (KR), to dehydrate the β -hydroxyl product to an α,β -unsaturated alkene (DH), or to reduce the alkene to a saturated alkane (ER), respectively. The reductive loop may contain one of four combinations of these domains – none, KR only, KR–DH, or KR–DH–ER. Once the processing by the reductive loop is complete, the ACP-bound intermediate is translocated to either a downstream KS for another round of extension in the next module or, if it is already in the final module, to the terminal TE domain where it is offloaded from the assembly line (Fig. 4).³⁷

2.2.3 Offloading. The TE domain is responsible for releasing the polyketide chain from the assembly line, either as a macrocyclic lactone or a linear carboxylic acid. Both mechanisms involve nucleophilic attack of the TE-linked polyketide chain, though lactonization is initiated by hydroxyl functionality on the polyketide chain (Fig. 5A), whereas a free water molecule is the nucleophile for linear products (Fig. 5B). The nature of the offloading mechanism employed by a given PKS assembly line is a property of its TE domain,³⁸ and TEs which catalyze lactonization are much more common than TEs that release a linear product *via* hydrolysis.²¹ Atypical, alternative offloading mechanisms include two-electron and four-electron reductions to an aldehyde or alcohol, respectively.³⁹

TEs that catalyze lactonization are usually highly selective in terms of the ring size generated. For example, erythromycin producers generate only 14-membered macrolactones despite the availability of other intramolecular nucleophiles (Fig. 2). An exception to ring size selectivity is known for the pikromycin (Pik) PKS in which 12- and 14-membered macrolactones are produced. However, the mechanism for this promiscuity was shown to be module skipping instead of the Pik TE accepting alternate hydroxyl nucleophiles.⁴⁰ For the DEBS TE, the *D*-configuration of the C-13 hydroxyl group has been shown to be required for macrocyclization – smaller and larger ring sizes can be engineered only if the *D*-configuration of the hydroxyl nucleophile is maintained.³⁸

2.3 Polyketide structure diversification *via* synthase engineering

As evidenced from the above sections, the structural diversity of polyketides can be largely attributed to a few biosynthetic variables.⁷ The length of the polyketide chain is primarily determined by the number of modules present.¹⁶ The selection of monomers to be incorporated into the final scaffold is determined by the gatekeeper AT domain.^{35,36} β -Carbon processing and the degree to which it is reduced are governed by

Fig. 4 Mechanism of polyketide chain extension in a prototypical module *N*. The ACP from the upstream module *N* – 1 transfers the intermediate to KS_n in a process known as transacylation. AT_n selects the appropriate extender unit and charges ACP_n . Then, KS_n catalyzes a decarboxylative Claisen-like condensation of the intermediate onto the ACP_n -bound extender unit. The new ACP_n -bound intermediate can then be processed by the domains of the reductive loop. The depicted module *N* harbors only a KR domain in the reductive loop, but DH or ER domains may also be present. The ACP_n -bound intermediate is then ready to be transferred to either the downstream module *N* + 1 for another round of extension or to the C-terminal TE domain for offloading if module *N* is the final extension module of the assembly line.

Fig. 5 Mechanisms of polyketide offloading. After undergoing extension and processing in the final module of the assembly line, the ACP-bound polyketide intermediate is transferred to the C-terminal TE domain where it can be offloaded from the assembly line. There are two primary mechanisms that may be employed, which are functions of the specific TE present. (A) Intramolecular nucleophilic attack of the ester bond by a hydroxyl group to yield a macrocyclic product. (B) Intermolecular nucleophilic attack of the ester bond by a free water molecule to yield a linear carboxylic acid product.

the domains of the reductive loop, which also hold significant influence on the complex stereochemistry of the final product.^{25,26} Offloading mechanisms mediated by the C-terminal TE domain result in either a linear or cyclized product.^{38,39} Finally, post-PKS tailoring enzymes can install additional chemical functionality to the polyketide product.⁴¹

Given this understanding of how PKS enzymes generate chemical diversity, individual catalytic domains become appealing candidates for modification as a means to alter the structure of the polyketide product formed by the assembly line. As such, domains have been previously described as “the combinatorial units” of PKS combinatorial biosynthesis.⁴² The sections below highlight examples of modifications to polyketide structures that can be achieved by domain engineering.

3. Monomer diversification

The AT domain exerts significant control over the final structure of the polyketide product by selecting which carboxyacyl monomer will be incorporated into the scaffold. By engineering the AT domain, one can achieve a novel polyketide scaffold architecture *via* the incorporation of non-native monomers with variable α -substituent functionality during either the loading or extension steps. The two most common monomers utilized in polyketide biosynthesis are malonyl-CoA and (2*S*)-methylmalonyl-CoA. However, there are a variety of more exotic monomers known to be incorporated by modular PKSs – including analogues exhibiting aromatic functionality, branched-chain alkyl substituents, and halogenation. For details on naturally occurring starter and extender units, we refer the reader to several comprehensive reviews.^{35,36,43} Additionally, alternative monomers may be synthetically produced or engineered, allowing site-specific incorporation of designed functionality.

3.1 AT engineering

To date, several strategies have been employed for AT engineering including domain exchange, site-directed mutagenesis, and AT inactivation with cross-complementation.^{44,45}

3.1.1 Domain exchange. One of the early examples of AT domain exchange was the replacement of the (2*S*)-methylmalonyl-CoA specific AT domain in module 4 of DEBS with a malonyl-CoA specific AT domain from the rapamycin PKS to afford 6-desmethyl erythromycin D, albeit at low yields of $\sim 0.3 \text{ mg L}^{-1}$.⁴⁶ A more recent example involved the production of 25-methyl and 25-ethyl ivermectin analogues at yields of 2 g L^{-1} and 0.95 g L^{-1} , respectively, by exchanging the loading AT_L-ACP didomain of the avermectin PKS with the loading AT_L-ACP from the milbemycin PKS (Fig. 6A).⁴⁷

These examples show that domain exchange may be used to alter monomer incorporation at either loading or extension modules of PKS assembly lines. Yet, domain exchange has proven to be a particularly complicated feat as many attempts have resulted in structurally and/or catalytically compromised assembly lines, leading to low product yields or no product at all. As we will cover in section 6.2, maintaining the integrity of ACP interactions with its cognate enzymes, including AT domains, is crucial in order to generate functional PKS chimeras. Systematic investigations by Yuzawa *et al.* have helped identify the best regions to “cut and paste” in order to enable AT swaps in the context of small chimeras.⁴⁸

3.1.2 Site-directed mutagenesis. Site-directed mutagenesis provides a route for AT engineering that may be less likely to significantly perturb the protein structure and catalytic activity than domain exchange. This technique has been employed to relax the substrate specificity of an AT domain, causing promiscuous incorporation of both native and non-native starter/extender units into the polyketide backbone. Thus, the engineered assembly line may produce multiple analogues in a single fermentation broth. One such early example involved mutation of DEBS AT4, leading to *in vivo* production of a mixture of the natural polyketide, 6-deoxyerythronolide B (6-dEB), and an unnatural analogue, 6-desmethyl-6-dEB, which is formed when DEBS AT4 incorporates malonyl-CoA rather than methylmalonyl-CoA, its native substrate.⁴⁹ Site-directed mutagenesis experiments may be rationally guided by PKS structural information. For instance, Bravo-Rodriguez *et al.* used molecular modeling to inform targeted mutagenesis of DEBS AT6 and produce unnatural derivatives including 2-propargylerythromycin A.⁵⁰ Moreover, recently obtained structural data of a promiscuous AT, that of SpnD-AT from the splenocin PKS, further informed efforts to broaden the substrate scope of DEBS AT6 to include larger functional groups such as benzyl substrates.⁵¹ These studies suggest that broadening the substrate specificity may prove a useful approach for generating suites of analogues, allowing subsequent investigation into structure–activity relationships. However, it may also be desirable to invert the substrate specificity rather than simply relax it.

Inversion of substrate specificity would allow generation of a single non-natural polyketide analogue with site-specific incorporation of a particular monomer. Recently, inversion of substrate specificity was achieved in the last subunit of the erythromycin PKS by Koryakina *et al.* when they modified DEBS AT6 *via* site-directed mutagenesis to preferentially

Fig. 6 Examples of polyketide analogues obtained *via* monomer diversification and the methods employed (A–D).

accept propargylmalonyl-CoA over the native substrate methylmalonyl-CoA.⁵² In a similar *in vitro* system, Kalkreuter *et al.* recently employed site-directed mutagenesis to invert substrate specificity in the last two modules (5 and 6) of the pikromycin (Pik) PKS.⁵³ Both PikAT5 and PikAT6 harbor natural specificity for methylmalonyl-CoA with low native promiscuity. The implemented mutations altered the active site such that the AT was able to accommodate malonyl-CoA derivatives with bulkier α -substituents of propargyl (Fig. 6B), ethyl, allyl or butyl. Importantly, they were able to see single incorporation of non-native extender units by each module as well as double incorporation, where the non-native extender unit was incorporated by both mutated modules in sequence. This is the first time that a full-length polyketide has been generated with the incorporation of two non-natural extender units. They also observed that incorporation of bulky side chains by PikAT5 results in low efficiency of downstream reductive processing by PikKR5. The final round of extension does not occur for intermediates from module 5 that were not reduced by PikKR5, generating a truncated polyketide product with additional keto functionality. The authors propose that this may be the result of the downstream KS harboring specificity for an intermediate exhibiting β -hydroxyl rather than β -carbonyl functionality. If true, this would suggest that alternative monomer incorporation by a non-terminal module may additionally require that downstream domains be engineered to accept the new substrate.

3.1.3 Inactivation and cross-complementation. As briefly mentioned in the Introduction section, modular type I PKSs are categorized into *cis*-AT and *trans*-AT systems. In *cis*-AT systems, such as DEBS and all other assembly lines discussed in this review thus far, each module contains an AT domain that is part of the megaenzyme structure. In contrast, modules of *trans*-AT systems lack their own AT domain and instead utilize a free-standing, separately encoded AT for substrate selection and loading.⁵⁴ The existence of *trans*-AT domains allows for generating chemical diversity *via* inactivation of native *cis*-AT domains and cross-complementation with a heterologous *trans*-AT domain.

This strategy was employed by Walker *et al.* to achieve site-selective incorporation of fluorinated monomers by a

truncated PKS consisting of DEBS modules 2–3 and the TE domain.^{55,56} They tested the ability of each of the two DEBS modules to incorporate the fluoromalonyl-CoA monomer *in vitro*. While the fluorinated product was produced (Fig. 6C), the catalytic efficiency of the system dropped drastically, partially due to substantial hydrolysis of the fluorinated building block by the methylmalonyl-CoA specific AT of DEBS. The authors hypothesized that a malonyl-CoA specific AT may interact more favorably with the fluorinated building block, so they devised a system in which the AT domain of the same DEBS modules was inactivated *via* site-directed mutagenesis, and the activity of the inactivated domain was complemented with a malonyl-CoA specific *trans*-AT. Doing so significantly increased fluoromalonyl-CoA incorporation by the truncated PKS system.⁵⁷ This work shows that it is possible to utilize PKS machinery to incorporate fluorine into a PKS scaffold in a site-specific manner, an application that is particularly attractive to medicinal chemistry.⁵⁸

Another example of cross-complementation used the promiscuous KirCII *trans*-AT to introduce allyl- and propargyl functional groups into kirromycin.⁵⁹ Although kirromycin is encoded by a *trans*-AT PKS, one can envision that KirCII has the potential to be applied with *cis*-AT systems in a manner analogous to what was done for site-specific incorporation of fluorine. Introduction of propargyl groups is significant because alkynes can be further derivatized *via* click chemistry, allowing the integration of synthetic biology and synthetic chemistry for natural product structure diversification.

3.2 Precursor-directed biosynthesis and engineered monomer biosynthesis

In addition to AT engineering, the introduction of modified monomers into polyketide scaffolds can also be achieved by a process called mutasynthesis or precursor-directed biosynthesis. This involves harnessing the promiscuity of certain AT domains coupled with interruption of the biosynthesis of the natural monomer and supplementation of mutant cultures with an alternate monomer.

For instance, fluorosalinosporamide – a hybrid polyketide-nonribosomal peptide (Fig. 6D) – was engineered by

supplementing a fluorinated precursor to cultures of a mutant blocked in the biosynthesis of chloroethylmalonyl-CoA, the natural polyketide monomer.⁶⁰ Alternatively, fluorosalinosporamide can also be obtained *via* total biosynthesis by replacing the chlorinase gene involved in native monomer biosynthesis with a fluorinase gene.⁶¹ The latter was possible because the fluoroacetate pathway from which the fluorinase gene was derived⁶² shares similarities with the chloroethylmalonyl-CoA pathway.^{63,64}

4. Control of the oxidation state and stereochemistry

If present in a given module, the optional reductive loop spans from the C-terminus of the AT domain to the N-terminus of the ACP domain and comprises some combination of the processing domains KR, DH, and ER. In addition to determining the oxidation state of the polyketide scaffold, the reductive loop is also the primary determinant of scaffold stereochemistry (Fig. 7). The KR domain, in particular, is responsible for setting the majority of stereocenters present in a polyketide scaffold.²⁵

KR domains stereoselectively reduce β -ketoacyl intermediates to yield either *L*- β -hydroxyls (A-type KRs) or *D*- β -hydroxyls (B-type KRs). The DH domain introduces alkene functionality into the polyketide scaffold *via* *syn*-coplanar elimination of the β -hydroxyl group and α -proton.⁶⁵ This occurs stereospecifically such that a substrate with an *L*- β -hydroxyl group will yield a *cis*-double bond, while a substrate with a *D*- β -hydroxyl group will yield a *trans*-double bond.^{25,66} Therefore, the stereochemical outcome of a DH-mediated dehydration is tied to the nature of the KR domain that immediately pre-

cedes it. The final activity of the reductive loop is carried out by the ER domain, which reduces the α,β -alkene functionally to a fully saturated alkane (Fig. 7A).

The stereocontrol of the reductive loop becomes even more nuanced when the polyketide intermediate it processes bears an α -substituent, which necessarily creates a chiral center at the α -carbon (Fig. 7B). In addition to their stereoselective reduction of β -carbonyl groups, KR domains also exhibit stereospecificity by reducing substrates with either a *D*- α -substituent (A1- and B1-type KRs) or those with an *L*- α -substituent (A2- and B2-type KRs). A2- and B2-type KRs exhibit additional epimerase activity that allows inversion of the α -substituent orientation from *D* to *L* prior to reduction of the carbonyl. There also exist redox-incompetent KR domains, C2-type KRs (such as DEBS KR3), that can exhibit the same epimerase activity when presented with a *D*- α -substituted intermediate, therefore yielding an *L*- α -substituted- β -ketoacyl intermediate. The final reductive domain, ER, also has the ability to exert stereocontrol over α -substituted intermediates. When presented with a *trans*- α,β -unsaturated intermediate possessing an α -substituent, an *L*-type ER will yield an *L*- α -substituent, and a *D*-type ER will yield a *D*- α -substituent.²⁵

4.1 KR engineering

4.1.1 Site-directed mutagenesis. As with other domains of the reductive loop, the KR can be inactivated *via* site-directed mutagenesis to block reduction, thereby altering the oxidation state of the polyketide backbone.

One such example achieved inactivation of KR domains from the amphotericin B PKS *via* site-directed mutagenesis.⁶⁷

Fig. 7 Control of the oxidation state and stereochemistry effected by domains of the reductive loop when (A) an α -unsubstituted monomer, such as malonate, is used for extension and (B) when an α -substituted monomer, such as (2S)-methylmalonate, is used for extension.

The authors identified the KR domains of modules 12 and 16 as appropriate candidates for modification 1) by eliminating KR domains whose action is known to be essential for downstream biosynthetic steps or pharmacological activity, and 2) because the importance of C-7 (KR16) and C-15 (KR12) to the structure–activity relationship of amphotericin had not yet been investigated.

Each of the two KR domains was successfully inactivated, affording production of several analogues that retained antifungal activity but exhibited additional carbonyl functionality on the polyketide backbone. Notably, inactivation of KR16 generated the analogue 7-oxo-amphotericin, which showed reduced hemolytic activity and has, therefore, the potential to improve the therapeutic window of this clinically relevant drug (Fig. 8). Further, the authors argue that addition of carbonyl functionality serves to prime the structures for further modifications *via* chemical synthesis, such as modifications to increase polarity and solubility.

4.1.2 Domain exchange. Most efforts in KR engineering have focused on understanding how to harness the exceptional stereocontrol exhibited by this domain. One early study on the mechanism of stereocontrol by KR domains identified several active site residues that, when modified, altered the stereochemistry of the product *in vitro*.⁶⁸ However, incorporation of the same mutations did not result in the desired stereochemistry when the KR was part of an intact PKS *in vivo*, indicating gaps in our understanding of the molecular mechanism for KR stereocontrol.⁶⁹ Consequently, domain exchange is the most commonly employed method to change the stereochemical outcome. However, domain exchange is still not a perfectly refined approach, and much of the ongoing research utilizes truncated PKS model systems to aid in

devising methods and principles for predictably altering stereochemical outcomes *via* KR exchange.

One such method involves the use of polylinkers to achieve swaps of an entire reductive loop. Recognizing the common evolutionary origins of the domains in the reductive loop and, consequently, the need to preserve the intimate protein–protein interactions between them, Kellenberger *et al.* devised a polylinker-mediated approach for systematic engineering of reductive loop swaps in the truncated PKS model system DEBS1-TE (the first subunit of DEBS fused to the TE from DEBS3).⁷⁰ By employing this approach *in vitro*, it was possible to alter both the oxidation state and stereochemistry of the intermediate β -hydroxyl group *via* exchange of reductive loop domains in the second module of DEBS.

The same method was used some years later by Annaev *et al.* in the same model system to probe KR stereocontrol over α -substituents.⁷¹ The authors sought to find a donor KR that would afford epimerized α -substituents in the second module of DEBS, which natively harbors an A1-type KR. They tested a donor pool composed of both A2- and B2-type KRs, which are all known to exhibit epimerase activity in their native context. Not all chimeras were active, but all of the active chimeras generated hydroxyl configuration consistent with the donor KR type in a new context. This validated that the direction of keto-reduction is a fundamental property of the KR active site. However, almost half of the donor KRs did not switch the α -methyl configuration. This result indicates that the PKS context of the KR affects the activity and that we still lack a detailed understanding of the molecular basis for epimerization.

In a recent advance towards fully realizing the ability to alter stereochemistry *via* KR domain swap, Eng *et al.* were able to effect a stereochemical change from *anti*- to *syn*- *via* a reductive loop swap in the truncated PKS model system LipPKS1-TE (the first module of the lipomycin PKS fused to the TE from DEBS).⁷² The authors replaced the native A2-type KR with a panel of donor A1-type KRs. In all cases, the stereochemistry was altered from *anti* to *syn*- *via* epimerization of the α -substituent. Further, the native module contains a dimerization element (DE), a structural motif essential for catalytic activity. The authors found that the activity is retained when the native DE is retained but attenuated when the native DE is replaced with a donor DE. This is a significant finding for the advancement of KR domain swapping as this DE is present in approximately half of PKS modules where the KR is the sole domain in the reductive loop.⁷³

4.2 DH engineering

4.2.1 Site-directed mutagenesis. The DH has not been as extensively explored for engineering as other PKS domains, but examples of targeted DH engineering are present in the literature, mostly utilizing site-directed mutagenesis to inactivate its functionality. One such study by Kim *et al.* inactivated the DH domain in the first module of the geldanamycin PKS, leading to retention of hydroxyl functionality at a position which previously had been fully reduced to a saturated alkane

Fig. 8 Diversification of amphotericin *via* KR engineering. Structures of (A) amphotericin B and (B) a novel analogue, 7-oxo-amphotericin, which was generated by inactivating KR16 in the amphotericin PKS.

(Fig. 9).⁷⁴ Additionally, one of the post-PKS tailoring enzymes was disrupted in order to prevent installation of quinone functionality thought to be associated with dose-limiting hepatotoxicity in geldanamycin, which shows anticancer activity.

The combinatorial modifications made to the biosynthetic pathway led to *in vivo* production of a suite of novel analogues, all featuring the desired hydroxylation at C15 but varying slightly in the extent to which they were processed by post-PKS tailoring reactions.⁷⁴ One analogue, DHQ3, was detected at titers comparable to the production of geldanamycin in the wild type strain. Significantly, DHQ3 exhibited 4.6-fold higher biological activity than geldanamycin and is the only C15 hydroxylated analogue to show any significant activity at all. By comparing the structural differences between the analogues, the authors were able to determine the structural features likely responsible for the enhanced activity, showing the ability of PKS engineering to produce novel chemistry that may help tease out structure–activity relationships.

4.2.2 Domain exchange. DH domain swap played a key role in a recent study by Hagen *et al.* that sought to achieve *in vitro* production of the commodity chemical adipic acid *via* engineered PKS machinery.²⁰ Starting with the first extension module of the borrelidin PKS, which contains only a KR processing domain, the authors systematically performed reductive loop swaps using a variety of donor loops containing the full set of processing domains – KR, DH, and ER. However, they found that processing by the reductive loop stalled at the DH domain in each case, presumably as none of the DH domains in any of the donor loops were capable of processing the unnatural substrate – succinyl-ACP – which features a terminal carboxylic acid group that is atypical of polyketide intermediates. As DH domains have been previously reported to display substrate specificity,⁶⁵ the authors sought to identify a DH domain known to process a substrate with a carboxylic acid group and found one in the second extension of the borrelidin PKS. In an exciting combinatorial feat, the

authors replaced the DH domain of the already chimeric modules with the DH domain from the second module of the borrelidin PKS. Finally, the TE from the DEBS assembly line was attached to the end, allowing release of adipic acid from the chimeric PKS.

4.3 ER engineering

4.3.1 Site-directed mutagenesis. Early efforts to engineer the ER domain were focused on altering the stereochemical outcome of the reduction when a methyl branch is present as an α -substituent. One study revealed a point-mutation (Y52V) in the active site of an erythromycin ER that is able to effect the inversion of the methyl branch orientation *in vivo*.⁷⁵ However, the reverse mutation (V52Y) in an ER shown to have the opposite stereospecificity (rapamycin ER of module 13) did not alter the orientation of the methyl branch, indicating an incomplete understanding of the molecular mechanisms of ER-mediated reduction. Recent efforts to engineer ER domains have focused on reductive loop swaps or using site-directed mutagenesis to inactivate the ER domain.

4.3.2 Domain exchange. Two related examples center on affording *in vivo* production of ivermectin, which is a clinically relevant derivative of the antiparasitic agent avermectin that is currently produced semi-synthetically *via* reduction of the natural product with Wilkinson's catalyst.^{76,77} Avermectin has a double bond at C22–C23, but this is reduced to an alkane in ivermectin (Fig. 10). The first successfully engineered *in vivo* production of ivermectin was reported in 2006.⁷⁸ The relevant reductive loop from the second module of the avermectin PKS (AveDH2-KR2) was replaced with a fully-reducing loop from the fourth module of the pikromycin PKS (PikDH4-KR4-ER4). However, the engineered ivermectin producer suffered from low yields (1–3% of parent) that precluded it from commercial viability.

More recently, a separate group attempted to perform a similar reductive loop swap to afford engineered ivermectin production. Zhang *et al.* replaced the AveDH2-KR2 motif with milbemycin reductive loop MilDH2-KR2-ER2.⁴⁷ Excitingly, this led to *in vivo* production of ivermectin B1 in a high yield nearly equivalent to the yield of avermectin in the parent strain.

Milbemycin, also an antiparasitic agent, has a polyketide scaffold that is strikingly similar to that of avermectin (Fig. 10). Furthermore, the donor milbemycin reductive loop shows higher sequence similarity (51.50%) to the avermectin reductive loop than the previously used donor loop from pikromycin (46.48%).⁴⁷ Both sequence similarity of donor and acceptor modules and structural similarity of their native intermediate compounds, respectively, are major factors to consider when exchanging PKS domains. While the original group did take sequence similarity of the donor loop into consideration, they were limited at that time by the available PKS sequence information as the sequence of the characterized milbemycin PKS was not reported until 2010.⁷⁹ This example highlights how increased PKS sequence knowledge can contribute to engineering efforts.

Fig. 9 Diversification of geldanamycin achieved *via* DH engineering and disruption of post-PKS processing. (A) Geldanamycin and (B) novel analogue DHQ3 generated by DH domain inactivation and gene disruption of an oxidative tailoring enzyme responsible for formation of the quinone ring. In addition to retaining the expected hydroxylation at C15 and lacking quinone functionality, DHQ3 also lacks O-methylation of the aromatic ring moiety and an alkene group on the polyketide backbone, both of which are installed during post-PKS tailoring steps.

Fig. 10 Structures of avermectins B1, ivermectins B1, and milbemycins A. The remarkably similar core structures make the milbemycin PKS a promising candidate as a reductive-loop donor to generate the avermectin B1 C22–C23 saturated analogue, ivermectin B1.

5. Engineering cyclized or linear products

Located on the C-terminus of the PKS, the TE catalyzes the offloading of the nascent polyketide product from the assembly line. Therefore, the TE domain is the final biosynthetic determinant originating from the PKS machinery that is capable of altering the primary core structure of the polyketide product.⁷ There are two principal mechanisms for TE-mediated offloading – hydrolysis to generate a linear carboxylic acid product and cyclization to generate a macrocyclic lactone or lactam. Other mechanisms exist, too, such as reductive offloading, which is however more common in nonribosomal peptide synthetase (NRPS) or PKS–NRPS hybrid systems.³⁹

The linear polyketide tautomycin (Fig. 11A) has gained attention for its potent immunosuppressant activity and novel mechanism of action.⁸⁰ Investigation of its biosynthetic pathway resulted in the first high-resolution X-ray crystal structure of a TE domain responsible for offloading of a lin-

ear polyketide product.⁸¹ This structure revealed that the tautomycin TE contains a narrow substrate tunnel, which significantly constricts the substrate and precludes macrocyclization of the polyketide product (Fig. 11B).

In an effort to generate a macrocyclic analogue of tautomycin *in vivo*, Tripathi *et al.* exchanged the native TE for the pikromycin TE, which catalyzes macrocyclization.⁸² In contrast to the tautomycin TE, the pikromycin TE has a large substrate binding tunnel capable of accommodating both 12- and 14-membered rings.⁸¹ The chimeric PKS was able to produce a 14-membered macrocyclic analogue of tautomycin (Fig. 11C), proving the utility of TE domain engineering for the production of novel chemistry. However, only traces of the cyclic analogue were observed, and yields were not quantified, indicating that TE engineering still requires refinement as a synthetic biology strategy in order to become practical.⁸²

In fact, recent studies based on the pikromycin PKS suggest that TE domains may have limited substrate flexibility. Hansen *et al.* showed that the wild-type pikromycin TE has drastically reduced activity with unnatural substrates containing epimerized chiral centers, failing to produce macrocycles.⁸³ Yet, Koch *et al.* demonstrated that TE site-directed mutagenesis can be used to increase catalytic efficiency and broaden substrate scope, enabling effective cyclization.⁸⁴ The authors found that mutation of the catalytic serine to a cysteine improves reaction kinetics by changing the

Fig. 11 Tautomycin and its analogue generated by TE engineering. Structures of (A) tautomycin in its mature, bioactive form following post-PKS tailoring, (B) tautomycin PKS linear polyketide product prior to modification by post-PKS tailoring enzymes, and (C) the macrocyclic tautomycin analogue generated by exchanging the native TE with the pikromycin TE.

reaction mechanism from a step-wise addition–elimination to a concerted acyl substitution as predicted from quantum mechanical calculations. These studies highlight that the nature of the TE should also be considered when attempting AT or reductive loop engineering to modify the polyketide scaffold.

6. The present and the future of polyketide structure diversification

The seminal discovery that polyketide macrolides such as erythromycin are biosynthesized by uncharacteristically large, multifunctional enzymes in the early 1990s^{14,15} has inspired chemists and biologists alike to harness the supposed modularity of polyketide synthases to diversify the polyketide structure. Although initial attempts to perform large-scale, combinatorial biosynthesis by module shuffling failed due to the complete loss or drastic reduction in product yields, a greater understanding of PKS structure and function and of PKS evolution have vitally informed recent efforts that led to the production of polyketide derivatives and small polyketide fragments in acceptable yields.

Further developments are still necessary to make synthetic biology of PKS assembly lines a reliable tool for generating structural diversity in polyketide chemical space. This includes a) continued efforts by natural product researchers to characterize PKS biosynthetic gene clusters; b) additional understanding of the structure and function of PKS enzymes; and c) development of empirically-based engineering principles to better guide PKS engineering attempts.

6.1 The continued role of natural product research

Microbial natural products have long been known to exhibit remarkable bioactivity, and decades of sustained efforts by natural product researchers have led to the discovery of diverse natural product scaffolds. Further, the advances of the genomics era have given us privileged access to the complex synthetic strategies employed by Nature to generate these compounds. The development of bioinformatic tools and databases such as antiSMASH,⁸⁵ MIBiG,⁸⁶ and others,⁸⁷ have created a space for natural product researchers from around the globe to contribute and analyze their sequence data.

Recently, the number of microbial genomes being sequenced has been increasing at an exponential rate, facilitating the discovery of previously unknown natural products *via* genome mining strategies.⁸⁸ It is important for natural product researchers to continue to probe cryptic PKS clusters uncovered by genome sequencing (*i.e.* those PKS clusters not yet connected to a characterized polyketide product). In order for PKS machinery to be useful to the synthetic biologist, it is essential for the PKS genes to be characterized, well-annotated, and associated with their cognate natural product. Recent estimates indicate that only 10% of the known polyketide structures have been associated with their biosynthetic gene cluster.⁸⁹ Therefore, connecting cryptic PKS gene clusters to the compounds they encode and continued efforts to

characterize PKSs will contribute to expanding the PKS toolkit. The ivermectin example presented here (Fig. 10) highlights the importance of continued efforts in PKS and polyketide discovery, since availability of sequence information for the milbemycin PKS was crucial for enabling engineered ivermectin production in commercially viable yields.⁴⁷

6.2 The emergent landscape of PKS synthetic biology

6.2.1 A developing PKS toolkit. Given the supposed modularity of polyketide synthases, one can envision, for instance, starting from a compound of interest and assembling a polyketide synthase from known “parts/devices” (domains/modules) that can catalyze biosynthesis of the target compound.⁹⁰ Such polyketide synthetic biology experiments currently rely on “design–build–test–learn” cycles, as does any other synthetic biology effort. Computational tools, such as the recently developed ClusterCAD,⁹¹ help with design by facilitating the identification of PKS domains and modules that can be used to generate a compound of interest *de novo*,²⁰ or that can be used to modify a given polyketide scaffold. Engineered polyketide synthases are then built and tested for activity, and the more variants are tested, the more we can learn regarding not only which parts to use but also how to connect the parts to produce catalytically active chimeric PKS assembly lines. These kinds of studies are critical for development of reliable principles for PKS engineering.

For example, recent efforts by Yuzawa *et al.* have helped to better identify well-conserved regions of the AT domain and surrounding linker region that serve as “cut–paste” sites for a donor AT domain.⁴⁸ The authors created several variants of a single chimeric module that differed in the boundaries used for the module exchange and then systematically analyzed the variants of this system. This led to identification of a putatively optimized domain boundary that maintained both protein stability and catalytic activity. Importantly, this domain boundary is highly conserved across most PKS modules, meaning that their findings should facilitate AT domain replacements in systems beyond their model. Indeed, tests of this method on a second model system indicate that the identified domain boundaries are likely generalizable.⁴⁸ These results can inform future efforts to perform AT domain exchanges as a method for generating novel polyketide chemistry.

Such insights into AT domain exchanges may provide guidance on how to exchange other catalytic domains, such as those of the reductive loop. While controlling stereochemistry is arguably one of the most exciting engineering outcomes, it has proven to be one of the most challenging tasks. For instance, the KR domain has been described as the “stereochemical workhorse”²⁵ of PKS assembly lines due to the exquisite stereocontrol it can exert over polyketide scaffolds (Fig. 7). However, KR domain swaps to achieve stereochemical changes have been demonstrated only in truncated model systems thus far, generally with mixed results.^{70,71} The ability to control the stereochemical outcome of complex

polyketides in a predictable way – although a welcoming addition to the PKS toolkit – remains outstanding and requires significant continued research.

6.2.2 PKS structure and function. Most early efforts in PKS synthetic biology failed due to an incomplete understanding of the structure and function of PKS megaenzymes. Much work has been done in the last decade to address this gap in our knowledge, and development of techniques such as electron cryo-microscopy has led to significant increases in available structural information, including the first structure of an intact PKS module.¹⁸

This expanded understanding of the PKS structure has also elucidated the importance of protein–protein interactions for not only the structural integrity of the assembly line but also the catalytic integrity.⁹² Additional structural information, including structures of PKS domains/modules at various points in their catalytic cycle, will help to further elucidate which regions of PKS assembly lines are more or less susceptible to perturbation during engineering efforts and how to avoid any issues that may compromise the integrity of the assembly line.

Moreover, the multimodular architecture of PKSs raises an interesting question regarding how synthetic progress along the assembly line is achieved. Mechanisms that prevent a module from being used more than once during biosynthesis of one polyketide molecule must be in place because, although there are exceptions,⁹³ PKS modules are not used iteratively. In other words, a mechanism that prevents the just-generated product from being acted upon again by the same module, and instead to be translocated to the correct downstream module must exist. Understanding how one-directional, or so-called vectorial biosynthesis is controlled is important for PKS engineering purposes as it allows biosynthesis to be directed towards translocation. While the molecular underpinnings of vectorial control remain to be fully elucidated, models have been proposed that include kinetic and structural elements.¹³

Protein–protein interactions are expected to play a crucial role in vectorial control. For instance, the ACP domain has received much attention due to its central role in shuttling the polyketide intermediate not only to downstream modules but also to processing domains within a module.¹⁹ Thus, the ACP domain is known to contain a variety of elements necessary for recognition by its cognate partners. While the ACP domain has not been considered as an engineering target in this review due to its lack of influence on chemical structure, it should be noted that maintaining interactions between an ACP and its cognate partners is of critical importance in generating functional, engineered assembly lines.^{94,95}

6.2.3 PKS evolution. In addition to structure/function studies on PKS machinery and systematic construction of chimeric PKSs, it is essential to study how these assembly lines evolved. Insights into PKS evolution led to recent discoveries that have expanded our understanding of the modular boundaries and various protein–protein interactions present in a PKS. By studying giant, >25-module PKSs encoding

aminopolyols such as mediomycin (Fig. 12A), Zhang *et al.* were able to infer that KS domains move together with the upstream ACP and reductive loop during recombination and evolution of modular PKSs. The authors were also able to show that KSs form substrate-specific, phylogenetic clades and to pinpoint previously unknown sequence motifs presumably associated with KS substrate specificity.⁹⁶ These findings are extremely significant in that they change our understanding of the KS domain by suggesting it to have more substrate specificity than previously thought and by showing that the KS is evolutionarily linked to the domains of the preceding module.

This evolutionary insight led to the proposal that, in order to enable synthetic biology efforts, we ought to update the canonical module boundaries (Fig. 12B) used for *cis*-AT PKS assembly lines such that the KS is grouped with its cognate ACP and reductive loop (Fig. 12C).⁹⁷ In fact, Klaus *et al.* recently demonstrated the recognition of ACP_{*n*} and KS_{*n+1*} to play a major role in turnover,⁹⁸ corroborating the module redefinition proposed by Zhang *et al.* based on phylogenetic analyses.⁹⁶ A similar phylogenetic analysis of ACP and KS domains from *trans*-AT systems suggests that this boundary update applies to these assembly lines as well.⁹⁹

The DEBS assembly line shown in Fig. 2 of this review conforms to the originally proposed boundaries, for it seems appropriate to wait for further consensus on the matter before changing the representation of the canonical PKS assembly line. Further, it is worth noting that this combinatorial definition of a PKS module conflicts with the canonical biosynthetic definition. Until now, a PKS module has been primarily defined as a functional unit capable of catalyzing a single round of polyketide chain extension. However, the combinatorial module (Fig. 12C) no longer fulfils this criterion, as the KS is no longer grouped with the downstream AT and ACP (see Fig. 4). Still, we encourage the reader to consider the module redefinition shown in Fig. 12C when attempting PKS synthetic biology.

Evolution-guided approaches reveal the likelihood of encountering downstream bottlenecks in engineering of larger PKS assembly lines. This prediction seems supported by several examples discussed in this review, such as a) the failure of downstream KR domains to process unnaturally bulky intermediates generated *via* engineered control of monomer incorporation⁵³ described in the “AT engineering” section, and b) the drastically reduced activity of the wild-type pikromycin TE with epimerized, unnatural substrates⁸³ described in the “Engineering cyclized or linear products” section. As the combinatorial abilities of PKS synthetic biology continue to develop, additional downstream bottlenecks are sure to be revealed. This means that we are still not at the point of designing large PKS chimeras with ease, as we lack a holistic understanding of the interdependency between domains of disparate modules. However, certain targeted modifications can be achieved as highlighted in this review. Moreover, PKS synthetic biology has proven to be a useful method for generating small, chiral, functionalized

Fig. 12 Evolutionary studies of large aminopolyl PKS assembly lines leads to proposed redefinition of PKS modular boundaries: (A) structure of the aminopolyl mediomycin 2A. The bimodular second subunit of the mediomycin PKS (*medORF2*) depicted using (B) the canonical PKS module boundaries and (C) the proposed module redefinition based on studies by Zhang *et al.*⁹⁶ The C–C bonds formed by KS₆ and KS₇ of the *med* BGC are indicated on the structure of mediomycin 2A. Gene products of *medORF1* and *medORF3* are not drawn, but their approximate positions in the assembly line are indicated by dashed lines.

fragments,²⁰ which will likely prove to be a useful tool to the medicinal chemist.

6.3 Bridging natural product and medicinal chemistry research

Medicinal chemistry has long picked up where natural product research leaves off, in that it works to take bioactive compounds and develop them into optimized pharmaceutical therapies. In the wake of technological developments surrounding combinatorial chemistry, pharmaceutical companies have opted to deprioritize natural product research. However, a concomitant decline in the overall rate of drug discovery was observed.

Synthetic biology has the potential to help reunite these two fields and revitalize drug discovery by using the machinery of natural product biosynthesis to generate novel chemistry. At present, PKS synthetic biology provides the ability to either generate small fragments, or “oligoketides”, *de novo* from a few disparate PKS parts or to make small, targeted structural modifications to large natural product scaffolds, both of which are valuable tools for the field of medicinal chemistry. The ability of synthetic biology to generate chiral, functionalized synthons for incorporation into synthetic schemes would allow integration of privileged natural product scaffolds with the technological advancements of combinatorial chemistry.¹⁰⁰

Additionally, synthetic biology may be used to make novel analogues of polyketide natural products that help to tease out structure–activity relationships, as highlighted with the geldanamycin example (Fig. 9),⁷⁴ or create sustainable sources of clinically relevant natural product analogues by setting up *in vivo* production of compounds that are currently generated through total or semi-synthesis, as shown with the ivermectin example (Fig. 10).⁴⁷ Medicinal chemists may also meaningfully contribute to the field of synthetic biology by generating functionalized precursors for use in precursor-directed biosynthesis studies as well as by providing guidance for synthetic biologists regarding how and where to

modify polyketide scaffolds to confer an enhanced therapeutic profile to the natural product.

7. Conclusions

Owing to developments in PKS structural biology and enzymology, and the persistence of synthetic biologists to develop methods and principles for PKS engineering, polyketide synthetic biology has emerged as a viable tool for generating chemical diversity. Although further refinement of principles and tools is still necessary, the time is ripe to bridge natural product and medicinal chemical research by incorporating polyketide synthetic biology into medicinal chemistry efforts.

Medicinal chemical knowledge can substantively inform polyketide engineering efforts by directing engineering towards improved biological activity. Conversely, polyketide synthetic biology can be used to generate difficult-to-access, chiral fragments that can then be incorporated into synthetic schemes. Alternatively, targeted modifications to polyketide scaffolds can be made *via* synthetic biology to help tease out structure–activity relationships or address supply issues of natural product analogues. With so much to gain from establishing such a partnership, all that is left is to collaborate.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We thank M. D. Burkart and J. La Clair (University of California at San Diego) for helpful discussions. We also thank anonymous reviewers for their constructive feedback. Financial support for this work was provided by the National Center for Advancing Translational Sciences, National Institutes of Health (NIH), under grant KL2TR002002 (to ASE), the Office of the Director and the National Center for Complementary & Integrative Health, NIH, under grant T32AT007533 (to TK),

and by startup funds from the Department of Medicinal Chemistry and Pharmacognosy and the Center for Biomolecular Sciences, University of Illinois at Chicago (to ASE). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

References

- D. J. Newman and G. M. Cragg, *J. Nat. Prod.*, 2016, **79**, 629–661.
- R. Aminov, *Biochem. Pharmacol.*, 2017, **133**, 4–19.
- D. J. Newman, *Pharmacol. Ther.*, 2016, **162**, 1–9.
- I. Paterson and N. Y. S. Lam, *J. Antibiot.*, 2018, **71**, 215–233.
- A. L. Demain, *J. Ind. Microbiol. Biotechnol.*, 2014, **41**, 185–201.
- A. E. Yñíguez-Gutierrez and B. O. Bachmann, *J. Med. Chem.*, 2019, DOI: 10.1021/acs.jmedchem.9b00246.
- C. Hertweck, *Angew. Chem., Int. Ed.*, 2009, **48**, 4688–4716.
- S. Smith and S. Tsai, *Nat. Prod. Rep.*, 2007, **24**, 1041–1072.
- C. Khosla, S. Kapur and D. E. Cane, *Curr. Opin. Chem. Biol.*, 2009, **13**, 135–143.
- C. L. Bayly and V. G. Yadav, *Molecules*, 2017, **22**, 235.
- E. Kalkreuter and G. J. Williams, *Curr. Opin. Microbiol.*, 2018, **45**, 140–148.
- J. F. Barajas, J. M. Blake-Hedges, C. B. Bailey, S. Curran and J. D. Keasling, *Synth. Syst. Biotechnol.*, 2017, **2**, 147–166.
- M. Klaus and M. Grninger, *Nat. Prod. Rep.*, 2018, **35**, 1070–1081.
- J. Cortes, S. F. Haydock, G. A. Roberts, D. J. Bevirt and P. F. Leadlay, *Nature*, 1990, **348**, 176–178.
- S. Donadio, M. J. Staver, J. B. McAlpine, S. J. Swanson and L. Katz, *Science*, 1991, **252**, 675–679.
- S. J. Moss, C. J. Martin and B. Wilkinson, *Nat. Prod. Rep.*, 2004, **21**, 575–593.
- H. G. Menzella, R. Reid, J. R. Carney, S. S. Chandran, S. J. Reisinger, K. G. Patel, D. A. Hopwood and D. V. Santi, *Nat. Biotechnol.*, 2005, **23**, 1171–1176.
- S. Dutta, J. R. Whicher, D. A. Hansen, W. A. Hale, J. A. Chemler, G. R. Congdon, A. R. H. Narayan, K. Håkansson, D. H. Sherman, J. L. Smith and G. Skiniotis, *Nature*, 2014, **510**, 512–517.
- J. R. Whicher, S. Dutta, D. A. Hansen, W. A. Hale, J. A. Chemler, A. M. Dosey, A. R. H. Narayan, K. Håkansson, D. H. Sherman, J. L. Smith and G. Skiniotis, *Nature*, 2014, **510**, 560–564.
- A. Hagen, S. Poust, T. De Rond, J. L. Fortman, L. Katz, C. J. Petzold and J. D. Keasling, *ACS Synth. Biol.*, 2016, **5**, 21–27.
- A. T. Keatinge-Clay, *Chem. Rev.*, 2017, **117**, 5334–5366.
- A. T. Keatinge-Clay, *Nat. Prod. Rep.*, 2012, **29**, 1050–1073.
- K. J. Weissman, *Nat. Chem. Biol.*, 2015, **11**, 660–670.
- T. Robbins, Y. C. Liu, D. E. Cane and C. Khosla, *Curr. Opin. Struct. Biol.*, 2016, **41**, 10–18.
- A. T. Keatinge-Clay, *Nat. Prod. Rep.*, 2016, **33**, 141–149.
- K. J. Weissman, *Beilstein J. Org. Chem.*, 2017, **13**, 348–371.
- R. W. Broadhurst, D. Nietlispach, M. P. Wheatcroft, P. F. Leadlay and K. J. Weissman, *Chem. Biol.*, 2003, **10**, 723–731.
- C. D. Richter, D. Nietlispach, R. W. Broadhurst and K. J. Weissman, *Nat. Chem. Biol.*, 2008, **4**, 75–81.
- J. Lau, D. E. Cane and C. Khosla, *Biochemistry*, 2000, **39**, 10514–10520.
- H. Ikeda, T. Nonomiya, M. Usami, T. Ohta and S. Omura, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 9509–9514.
- P. Caffrey, S. Lynch, E. Flood, S. Finnan and M. Oliynyk, *Chem. Biol.*, 2001, **8**, 713–723.
- C. Bisang, P. F. Long, J. Cortés, J. Westcott, J. Crosby, A.-L. Matharu, R. J. Cox, T. J. Simpson, J. Staunton and P. F. Leadlay, *Nature*, 1999, **401**, 502–505.
- P. A. S. Lowden, B. Wilkinson, G. A. Böhm, S. Handa, H. G. Floss, P. F. Leadlay and J. Staunton, *Angew. Chem., Int. Ed.*, 2001, **40**, 777–779.
- J. Young, D. C. Stevens, R. Carmichael, J. Tan, S. Rachid, C. N. Boddy, R. Müller and R. E. Taylor, *J. Nat. Prod.*, 2013, **76**, 2269–2276.
- Y. A. Chan, A. M. Podevels, B. M. Kevany and M. G. Thomas, *Nat. Prod. Rep.*, 2009, **26**, 90–114.
- L. Ray and B. S. Moore, *Nat. Prod. Rep.*, 2016, **33**, 150–161.
- B. Lowry, X. Li, T. Robbins, D. E. Cane and C. Khosla, *ACS Cent. Sci.*, 2016, **2**, 14–20.
- M. E. Horsman, T. P. A. Hari and C. N. Boddy, *Nat. Prod. Rep.*, 2016, **33**, 183–202.
- L. Du and L. Lou, *Nat. Prod. Rep.*, 2010, **27**, 255–278.
- B. J. Beck, Y. J. Yoon, K. A. Reynolds and D. H. Sherman, *Chem. Biol.*, 2002, **9**, 575–583.
- C. Olano, C. Méndez and J. A. Salas, *Nat. Prod. Rep.*, 2010, **27**, 571–616.
- K. J. Weissman and P. F. Leadlay, *Nat. Rev. Microbiol.*, 2005, **3**, 925–936.
- B. S. Moore and C. Hertweck, *Nat. Prod. Rep.*, 2002, **19**, 70–99.
- B. J. Dunn and C. Khosla, *J. R. Soc., Interface*, 2013, **10**, 20130297.
- E. M. Musiol-Kroll and W. Wohlleben, *Antibiotics*, 2018, **7**, 62.
- H. Petkovic, R. E. Lill, R. M. Sheridan, B. Wilkinson, E. L. McCormick, H. A. I. McArthur, J. Staunton, P. F. Leadlay and S. G. Kendrew, *J. Antibiot.*, 2003, **56**, 543–551.
- J. Zhang, Y.-J. Yan, J. An, S.-X. Huang, X.-J. Wang and W.-S. Xiang, *Microb. Cell Fact.*, 2015, **14**, 1–13.
- S. Yuzawa, K. Deng, G. Wang, E. E. K. Baidoo, T. R. Northen, P. D. Adams, L. Katz and J. D. Keasling, *ACS Synth. Biol.*, 2017, **6**, 139–147.
- C. D. Reeves, S. Murli, G. W. Ashley, M. Piagentini, C. R. Hutchinson and R. Mcdaniel, *Biochemistry*, 2001, **40**, 15464–15470.
- K. Bravo-Rodriguez, S. Klopries, K. R. M. Koopmans, U. Sundermann, S. Yahiaoui, J. Arens, S. Kushnir, F. Schulz and E. Sanchez-Garcia, *Chem. Biol.*, 2015, **22**, 1425–1430.
- Y. Li, W. Zhang, H. Zhang, W. Tian, L. Wu, S. Wang, M. Zheng, J. Zhang, C. Sun, Z. Deng, Y. Sun, X. Qu and J. Zhou, *Angew. Chem., Int. Ed.*, 2018, **57**, 5823–5827.
- I. Koryakina, C. Kasey, J. B. McArthur, A. N. Lowell, J. A. Chemler, S. Li, D. A. Hansen, D. H. Sherman and G. J. Williams, *ACS Chem. Biol.*, 2017, **12**, 114–123.

- 53 E. Kalkreuter, J. M. Crowetipton, A. N. Lowell, D. H. Sherman and G. J. Williams, *J. Am. Chem. Soc.*, 2019, **141**, 1961–1969.
- 54 E. J. N. Helfrich and J. Piel, *Nat. Prod. Rep.*, 2016, **33**, 231–316.
- 55 M. C. Walker, B. W. Thuronyi, L. K. Charkoudian, B. Lowry, C. Khosla and M. C. Y. Chang, *Science*, 2013, **341**, 1089–1095.
- 56 O. Ad, B. W. Thuronyi and M. C. Y. Chang, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, E660–E668.
- 57 B. W. Thuronyi and M. C. Y. Chang, *Acc. Chem. Res.*, 2015, **48**, 584–592.
- 58 E. P. Gillis, K. J. Eastman, M. D. Hill, D. J. Donnelly and N. A. Meanwell, *J. Med. Chem.*, 2015, **58**, 8315–8359.
- 59 E. M. Musiol-Kroll, F. Zubeil, T. Schafhauser, T. Härtner, A. Kulik, J. Mcarthur, I. Koryakina, W. Wohlleben, S. Grond, G. J. Williams, S. Y. Lee and T. Weber, *ACS Synth. Biol.*, 2017, **6**, 421–427.
- 60 A. S. Eustáquio and B. S. Moore, *Angew. Chem., Int. Ed.*, 2008, **47**, 3936–3938.
- 61 A. S. Eustáquio, D. O. Hagan and B. S. Moore, *J. Nat. Prod.*, 2010, **73**, 378–382.
- 62 K. K. J. Chan and D. O'Hagan, *The Rare Fluorinated Natural Products and Biotechnological Prospects for Fluorine Enzymology*, Elsevier Inc., 1st edn, 2012, vol. 516.
- 63 A. S. Eustáquio, F. Pojer, J. P. Noel and B. S. Moore, *Nat. Chem. Biol.*, 2008, **4**, 69–74.
- 64 A. S. Eustáquio, R. P. McGlinchey, Y. Liu, C. Hazzard, L. L. Beer, G. Florova, M. M. Alhamadsheh, A. Lechner, A. J. Kale, Y. Kobayashi, K. A. Reynolds and B. S. Moore, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, 12295–12300.
- 65 D. Gay, Y.-O. You, A. T. Keatinge-Clay and D. E. Cane, *Biochemistry*, 2013, **52**, 8916–8928.
- 66 J. W. Labonte and C. A. Townsend, *Chem. Rev.*, 2013, **113**, 2182–2204.
- 67 P. Power, T. Dunne, B. Murphy, L. N. Lochlainn, D. Rai, C. Borissow, B. Rawlings and P. Caffrey, *Chem. Biol.*, 2008, 78–86.
- 68 A. Baerga-Ortiz, B. Popovic, A. P. Siskos, H. M. O'Hare, D. Spiteller, M. G. Williams, N. Campillo, J. B. Spencer and P. F. Leadlay, *Chem. Biol.*, 2006, **13**, 277–285.
- 69 D. H. Kwan, M. Tosin, N. Schläger, F. Schulz and P. F. Leadlay, *Org. Biomol. Chem.*, 2011, **9**, 2053–2056.
- 70 L. Kellenberger, I. S. Galloway, G. Sauter, G. Böhm, U. Hanefeld, J. Cortés, J. Staunton and P. F. Leadlay, *ChemBioChem*, 2008, **9**, 2740–2749.
- 71 T. Annaal, C. Paris, P. F. Leadlay, C. Jacob and K. J. Weissman, *ChemBioChem*, 2015, **16**, 1357–1364.
- 72 C. H. Eng, S. Yuzawa, G. Wang, E. E. K. Baidoo, L. Katz and J. D. Keasling, *Biochemistry*, 2016, **55**, 1677–1680.
- 73 J. Zheng, C. D. Fage, B. Demeler, D. W. Ho and A. T. Keatinge-clay, *ACS Chem. Biol.*, 2013, **8**, 1263–1270.
- 74 W. Kim, D. Lee, S. S. Hong, Z. Na, J. C. Shin, S. H. Roh, C. Z. Wu, O. Choi, K. Lee, Y. M. Shen, S. G. Paik, J. J. Lee and Y. S. Hong, *ChemBioChem*, 2009, **10**, 1243–1251.
- 75 D. H. Kwan, Y. Sun, F. Schulz, H. Hong, B. Popovic, J. C. C. Sim-stark, S. F. Haydock and P. F. Leadlay, *Chem. Biol.*, 2008, **15**, 1231–1240.
- 76 J. C. Chabala, H. Mrozik, R. L. Tolman, P. Eskola, A. Lusi, L. H. Peterson, M. F. Woods, M. H. Fisher, W. C. Campbell, J. R. Egerton and D. A. Ostlind, *J. Med. Chem.*, 1980, **23**, 1134–1136.
- 77 S. Ōmura and A. Crump, *Nat. Rev. Microbiol.*, 2004, **2**, 984–989.
- 78 X. Zhang, Z. Chen, M. Li and Y. Wen, *Appl. Microbiol. Biotechnol.*, 2006, **72**, 986–994.
- 79 X. Wang, Y. Yan, B. Zhang, J. An, J. Wang, J. Tian, L. Jiang, Y. Chen, S.-X. Huang, M. Yin, J. Zhang, A. Gao, C.-X. Liu, Z.-X. Zhu and W.-S. Xiang, *J. Bacteriol.*, 2010, **192**, 4526–4527.
- 80 S.-S. Choi, H.-J. Nah, H. Pyeon and E.-S. Kim, *J. Ind. Microbiol. Biotechnol.*, 2017, **44**, 555–561.
- 81 J. B. Scaglione, D. L. Akey, R. Sullivan, J. D. Kittendorf, C. M. Rath, E.-S. Kim, J. L. Smith and D. H. Sherman, *Angew. Chem., Int. Ed.*, 2010, **49**, 5726–5730.
- 82 A. Tripathi, S.-S. Choi, D. H. Sherman and E.-S. Kim, *J. Ind. Microbiol. Biotechnol.*, 2016, **43**, 1189–1193.
- 83 D. A. Hansen, A. A. Koch and D. H. Sherman, *J. Am. Chem. Soc.*, 2017, **139**, 13450–13455.
- 84 A. A. Koch, D. A. Hansen, V. V. Shende, L. R. Furan, K. N. Houk, G. Jiménez-Osés and D. H. Sherman, *J. Am. Chem. Soc.*, 2017, **139**, 13456–13465.
- 85 M. H. Medema, K. Blin, P. Cimermancic, V. De Jager, P. Zakrzewski, M. A. Fischbach, T. Weber, E. Takano and R. Breitling, *Nucleic Acids Res.*, 2011, **39**, 339–346.
- 86 M. H. Medema, *et al.*, *Nat. Chem. Biol.*, 2015, **11**, 625–631.
- 87 M. H. Medema, *mSystems*, 2018, **3**, e00182, DOI: 10.1128/mSystems.00182-17.
- 88 N. Ziemert, M. Alanjary and T. Weber, *Nat. Prod. Rep.*, 2016, **33**, 988–1005.
- 89 C. A. Dejong, G. M. Chen, H. Li, C. W. Johnston, M. R. Edwards, P. N. Rees, M. A. Skinnider, A. L. H. Webster and N. A. Magarvey, *Nat. Chem. Biol.*, 2016, **12**, 1007–1014.
- 90 S. Yuzawa, T. W. H. Backman, J. D. Keasling and L. Katz, *J. Ind. Microbiol. Biotechnol.*, 2018, **45**, 621–633.
- 91 C. H. Eng, T. W. H. Backman, C. B. Bailey, C. Magnan, H. García Martín, L. Katz, P. Baldi and J. D. Keasling, *Nucleic Acids Res.*, 2018, **46**, D509–D515.
- 92 G. J. Dodge, F. P. Maloney and J. L. Smith, *Nat. Prod. Rep.*, 2018, **35**, 1082–1096.
- 93 Y. Sugimoto, K. Ishida, N. Traitcheva, B. Busch, H.-M. Dahse and C. Hertweck, *Chem. Biol.*, 2015, **22**, 229–240.
- 94 S. S. Chandran, H. G. Menzella, J. R. Carney and D. V. Santi, *Chem. Biol.*, 2006, **13**, 469–474.
- 95 S. Kapur, A. Y. Chen, D. E. Cane and C. Khosla, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 22066–22071.
- 96 L. Zhang, T. Hashimoto, B. Qin, J. Hashimoto, I. Kozono, T. Kawahara, M. Okada, T. Awakawa, T. Ito, Y. Asakawa, M. Ueki, S. Takahashi, H. Osada, T. Wakimoto, H. Ikeda, K. Shin-ya and I. Abe, *Angew. Chem., Int. Ed.*, 2017, **56**, 1740–1745.

- 97 A. T. Keatinge-Clay, *Angew. Chem., Int. Ed.*, 2017, **56**, 4658–4660.
- 98 M. Klaus, M. P. Ostrowski, J. Austerjost, T. Robbins, B. Lowry, D. E. Cane and C. Khosla, *J. Biol. Chem.*, 2016, **291**, 16404–16415.
- 99 D. A. Vander Wood and A. T. Keatinge-Clay, *Proteins*, 2018, **86**, 664–675.
- 100 R. Liu, X. Li and K. S. Lam, *Curr. Opin. Chem. Biol.*, 2017, **38**, 117–126.