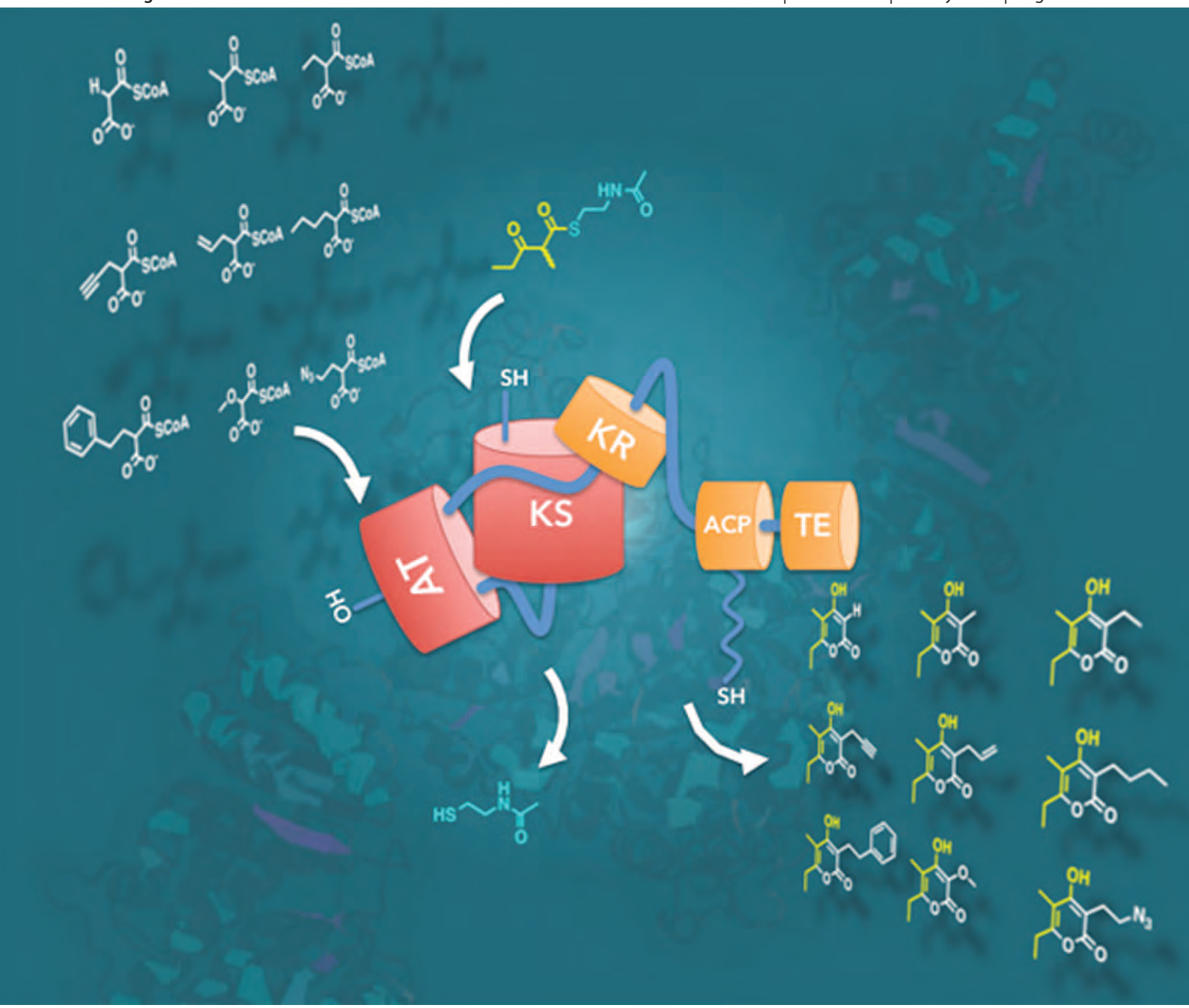
Organic & Biomolecular Chemistry

www.rsc.org/obc

Volume 11 | Number 27 | 21 July 2013 | Pages 4421–4560



ISSN 1477-0520

RSC Publishing

PAPER

Gavin J. Williams *et al.*

Promiscuity of a modular polyketide synthase towards natural and non-natural extender units

Organic & Biomolecular Chemistry

RSC Publishing

PAPER

Promiscuity of a modular polyketide synthase towards natural and non-natural extender units†

Cite this: *Org. Biomol. Chem.*, 2013, **11**, 4449

Irina Koryakina, John B. McArthur, Matthew M. Draelos and Gavin J. Williams*

Received 29th March 2013,
Accepted 8th May 2013

DOI: 10.1039/c3ob40633d

www.rsc.org/obc

Combinatorial biosynthesis approaches that involve modular type I polyketide synthases (PKSs) are proven strategies for the synthesis of polyketides. In general however, such strategies are usually limited in scope and utility due to the restricted substrate specificity of polyketide biosynthetic machinery. Herein, a panel of chemo-enzymatically synthesized acyl-CoA's was used to probe the promiscuity of a polyketide synthase. Promiscuity determinants were dissected, revealing that the KS is remarkably tolerant to a diverse array of extender units, while the AT likely discriminates between extender units that are native to the producing organism. Our data provides a clear blueprint for future enzyme engineering efforts, and sets the stage for harnessing extender unit promiscuity by employing various *in vivo* polyketide diversification strategies.

Introduction

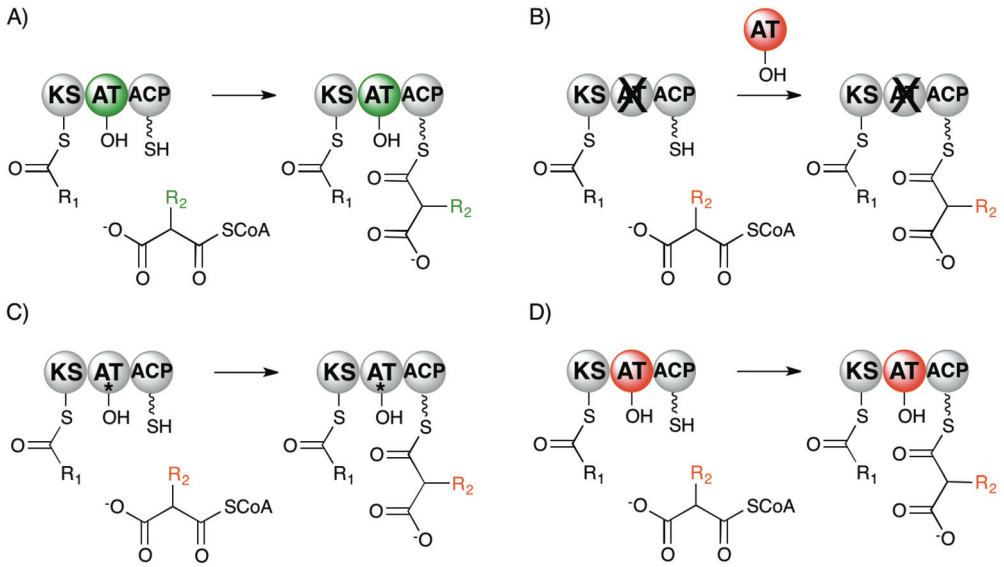
Polyketides are a large class of secondary metabolites that display a broad range of potent biological activities. There is significant interest in developing synthetic strategies that could be used to improve the pharmacological properties of polyketide based drugs or to synthesize polyketide analogs to probe, interrogate and manipulate biological processes. However, notable exceptions aside,[1–4] total synthesis of polyketides remains challenging. In addition, semi-synthesis strategies that rely on diversifying the structures of polyketide scaffolds, for example by 'click' chemistry or other chemoselective reactions,[5–8] are limited in scope and utility, due to the necessity to selectively introduce suitable chemical 'handles' into polyketides that are otherwise not available in naturally produced polyketide structures.[9,10]

Alternatively, biosynthetic strategies for polyketide diversification offer the potential for regioselective modification and combinatorial exploration of polyketide chemical space.[11,12] For example, inactivated acyltransferase (AT) domains can be supplemented with *trans*-ATs that have orthogonal specificity to the target AT[13] (Scheme 1). Site-directed mutagenesis has also been used to relax extender unit specificity of a PKS AT domain,[14–16] while substitution of entire AT domains has also led to production of the corresponding polyketide analogs (Scheme 1).[12,17] Yet, the scope and utility of these biosynthetic strategies is ultimately determined by the substrate specificity of polyketide synthases (PKSs), which catalyze iterative Claisen condensations between activated malonate-derived extender units and other acyl thioesters.[18] Cumulatively, only a modest variety of extender units are available to polyketide biosynthetic pathways,[19,20] and polyketide producing organisms usually only provide biosynthetic routes to a small number of extender units. For example, the most common PKS extender units include malonyl-Coenzyme A (CoA), methylmalonyl-CoA, and ethylmalonyl-CoA (**2a–c**, respectively, Scheme 2).[19] Less prevalent are extender units which are directly functionalized on standalone acyl carrier proteins (ACPs).[21] Several extender units, including chloroethyl-, propyl-, and hexylmalonyl-CoA are available to a subset of PKSs *via* the reductive carboxylation of α,β-unsaturated acyl-CoA precursors by crotonyl-CoA carboxylase/reductase (CCR) homologs.[22–24] Thus, the range of chemical diversity derived from extender unit selection is modest and includes little opportunity for further diversification by semi-synthesis.

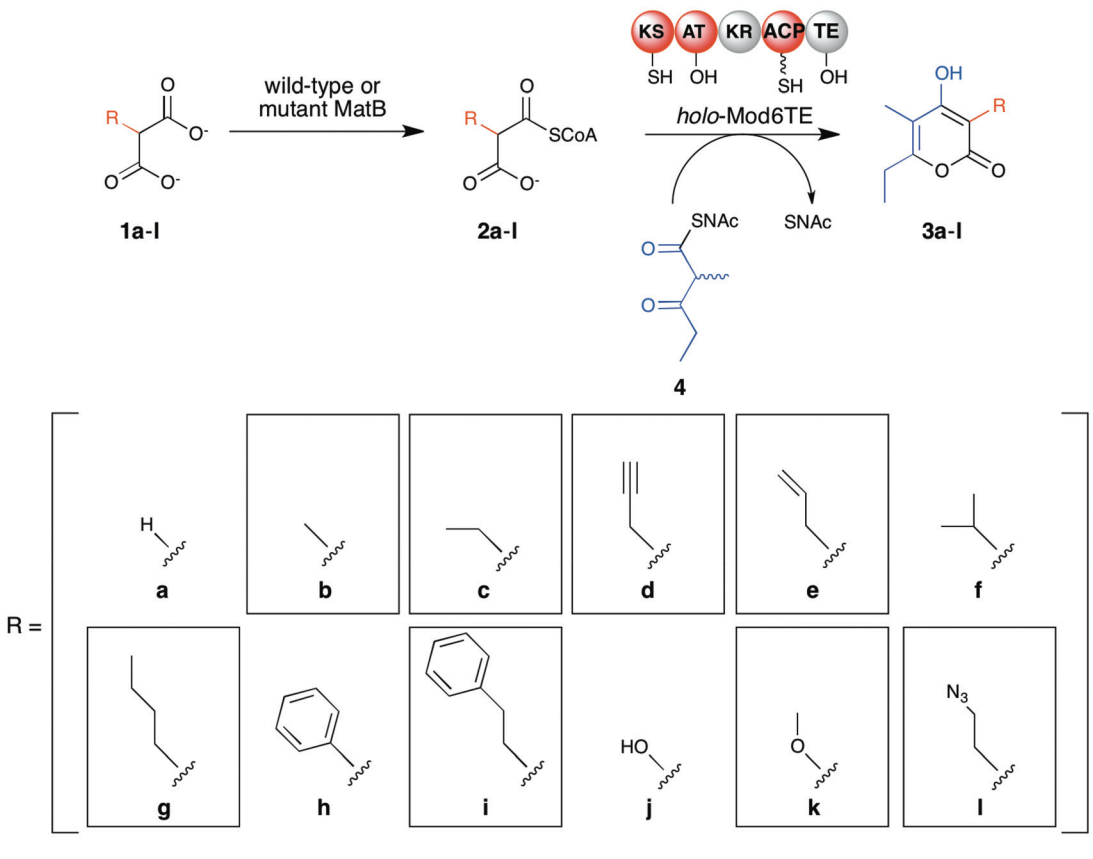
The vast majority of PKS specificity studies have been limited almost entirely to **2a–c**,[17,25–28] likely due in part to the lack of easily accessible biosynthetic routes to other acyl-CoA's or *N*-acetylcysteamine (SNAc) thioesters. Nevertheless, *in vitro* assays using modules from erythromycin[29,30] and pikromycin[27] biosynthesis have revealed tolerance towards one or two non-native extender units (*e.g.* **2c**) that are not available to the natural producer host strain. Accordingly, the synthetic potential of PKSs remains largely unexplored, and we set out to determine whether the inherent promiscuity of AT and ketosynthase (KS) domains of PKSs was sufficient to support the selection and installation of a wide variety of natural and non-natural extender units into polyketide synthase products. We reasoned that the discovery of new extender units could (*i*)

Department of Chemistry, North Carolina State University, Raleigh, NC 27695-8204, USA. E-mail: gavin_williams@ncsu.edu

†Electronic supplementary information (ESI) available: HPLC data, MS data, and NMR characterization of **3b**/**3e**. See DOI: 10.1039/c3ob40633d



Scheme 1 Strategies for regioselective modification of polyketides. (A) AT domains of wild-type PKSs select a cognate extender unit for loading onto the ACP; (B) Selected AT-domains are inactivated by site-directed mutagenesis and supplemented with exogenous *trans*-ATs; (C) Site-directed mutagenesis affords mutant ATs with relaxed substrate specificity; and (D) Entire AT domains can be substituted with those of different acyl-CoA specificity. The majority of these strategies have been employed with a very limited set of extender units (*e.g.* where R_2 = H, Me, or Et). See main text for references.



Scheme 2 Probing the extender unit specificity of the erythronolide B synthase terminal module, Mod6TE. Shown boxed are those acyl-CoA's successfully utilized by *holo*-Mod6TE. SNAc = *N*-acetylcysteamine. Domains in red are likely required for the indicated transformation. Wavy line on Mod6TE represents the phosphopantetheine prosthetic arm.

expand the potential scope and utility of polyketide biosynthesis strategies, (*ii*) provide a blueprint for future enzyme engineering, and (*iii*) potentially lead to development of high-throughput screens and selections for altering PKS substrate specificity.

Results and discussion

Extender unit generation

To provide a panel of potential PKS extender units that offer chemical functionality beyond that provided by natural biosynthetic systems, we recently created mutants of the *Rhizobium trifolii* malonyl-CoA synthetase MatB which are able to utilize commercially available and synthesized malonic acid analogues substituted at the C2 position (Scheme 2) to furnish the desired mono-thioesters in a single step to ~90% yield.[31,32] Conveniently, MatB-catalyzed synthesis of acyl-CoA's can be coupled to downstream biosynthetic systems without purification. For this study, 2-phenylethyl malonic acid (**1i**) was identified as an additional substrate for a previously described MatB variant, T207G/M306I,[31] which was then used for the synthesis of **2i** (see Experimental).

Acyl-CoA extender unit promiscuity of Mod6TE

Next, we set out to use this panel of acyl-CoAs (**2a–l**) to probe the extender unit specificity of the terminal module and thioesterase (TE) from the 6-deoxyerythronolide B synthase (DEBS), Mod6TE.[33] This system was chosen because it represents one of the best-studied PKSs, and a wealth of data[29,30] is available to benchmark any promiscuity revealed in this study. Moreover, this minimal system could be easily dissected and probed in order to identify specificity determinants, and much is known regarding the ability to use DEBS modules in a combinatorial fashion.[11,12,34–36] Thus, we first set out to probe the ability of *holo*-Mod6TE to catalyze the formation of triketide lactones from condensation of the diketide-SNAc **4** (Scheme 2) and each potential extender unit, **2a–l**. In order to generate the expected lactone **3a–l** using this simplified PKS system, the AT domain presumably must select and charge the ACP with a suitable extender unit, and the KS is required to catalyze the Claisen condensation with electrophile **4**. Subsequently, purified *holo*-Mod6TE was incubated with **4** and each acyl-CoA (**2a–l**), and the extender unit specificity was then determined using an HPLC-based end-point assay (Fig. 1 and Table 1) (see Experimental). As expected, no product was detected when **2a** was used as substrate, which is a known poor extender for DEBS and Mod6TE.[37] In contrast, and also as expected,[30] acyl-CoAs **2b** and **2c** each resulted in production of the corresponding triketide lactone product peak (Fig. 1 and Table 1; Fig. S1 and Table S1†). This result confirms that acyl-CoA's synthesized by the engineered MatB variant are catalytically competent with the erythronolide PKS machinery, which accepts only the (2*S*) stereoisomer of **2b**.[38] Remarkably, putative product peaks were also identified when propargyl **2d**, allyl **2e**, methoxy **2k**, and azidoethyl **2l** were included in the reaction, with synthetic conversions ranging from 9–44%, compared to that with **2b**. In addition to these robust substrates, butyl **2g**, phenyl **2h**, and phenylethyl **2i** resulted in very low (<2%) or trace conversions (Fig. 1 and Table 1; Fig. S1 and Table S1†). Analysis of control experiments that lacked either acyl-CoA, *holo*-Mod6TE, or **4** demonstrated that inclusion of each component was required for product formation in every case except when native extender **2b** was used (see Experimental, Fig. S2 and Table S2†), which provided the triketide

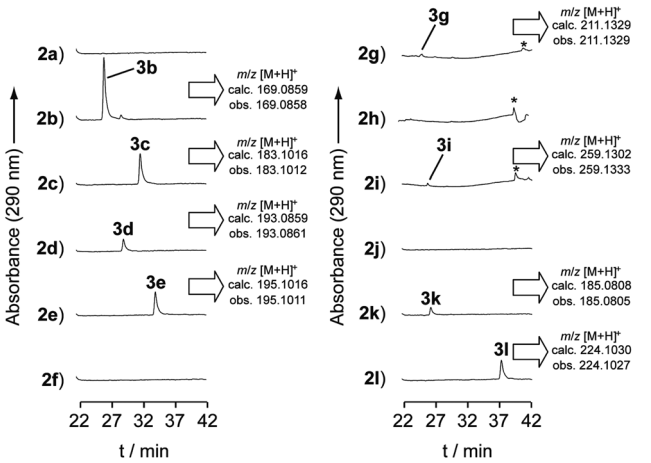


Fig. 1 HPLC and High Res LC-MS analysis of *holo*-Mod6TE-catalyzed triketide lactone formation with natural and non-natural extender units. Compound numbers for the acyl-CoA used are shown to the left of each chromatogram. See Scheme 2 for reaction. See Fig. S1/S2 and Tables S1/S2 in the ESI† for full series of negative controls and LC-MS data. *indicates contaminants present in the **2g–i** acyl-CoA preparations which were detected under the specialized conditions required for HPLC analysis of the corresponding lactones.

Table 1 Extender unit specificity of Mod6TE variants

Acyl-CoA/ lactone[a]	*holo*-Mod6TE[b]	*holo*-AT°-Mod6TE[c]	Sfp-acylated AT°-Mod6TE[d]
2a/**3a**	N.D.[e]	N.D.[e]	58
2b/**3b**	100	N.D.[e]	100
2c/**3c**	44	N.D.[e]	103
2d/**3d**	15	N.D.[e]	45
2e/**3e**	36	N.D.[e]	82
2f/**3f**	N.D.[e]	N.D.[e]	N.D.[e]
2g/**3g**	1	Trace[f]	58
2h/**3h**	Trace[f]	Trace[f]	Trace[f]
2i/**3i**	2	Trace[f]	33
2j/**3j**	N.D.[e]	N.D.[e]	N.D.[e]
2k/**3k**	9	N.D.[e]	21
2l/**3l**	31	N.D.[e]	64

[a] See Scheme 2 for structures of acyl-CoAs and lactones. [b] Reaction included *holo*-Mod6TE, **4**, and each acyl-CoA (Scheme 2). Numbers indicate relative product yield with activity toward **2b** set to 100. Standard error of assay is ±23% of the conversion value (determined with **2b**, $n = 3$). [c] Reaction included *holo*-AT°-Mod6TE, **4**, and each acyl-CoA. [d] Reaction included Sfp, *apo*-AT°-Mod6TE, **4**, and each acyl-CoA (Scheme 3). Numbers indicate relative product yield with activity toward **2b** set to 100. Standard error of assay is ±28% of the conversion value (determined with **2g**, $n = 3$). [e] N.D., non-detected. [f] Product could not be confidently quantified by HPLC (minimal detection limit 1.6%) but was detected by MS analysis.

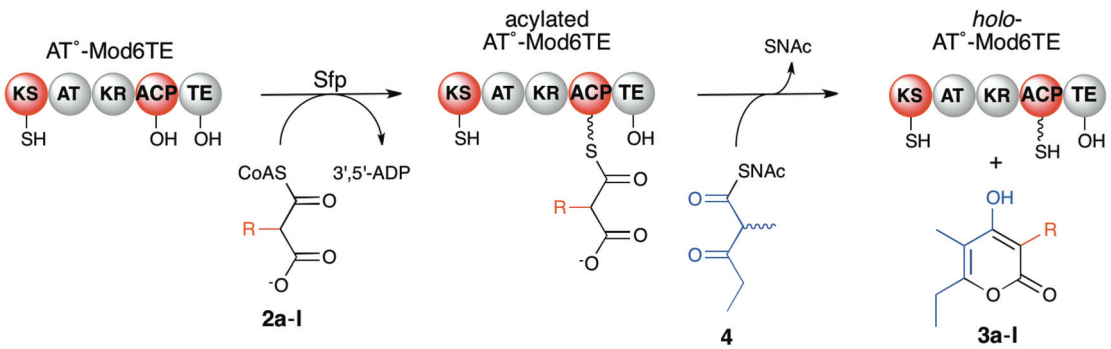
lactone **3b** in the absence of **4** due to 'stuttering'.[39,40] In every successful case, product identity was confirmed by high resolution LC-MS analysis of the product mixtures (Fig. 1; Fig. S1 and Table S1†). Conversion of **2b** and **2e** were selected as representative examples, scaled up, and the lactones **3b** and **3e** each purified by semi-preparative HPLC (see Experimental). Subsequent structural analysis by ^{1}H NMR (Fig. S3 and S4†) was totally consistent with the expected structures. To the best of our knowledge, direct utilization of acyl-CoAs **2e**, **2g–i**, **2k**, and **2l** by a PKS is unprecedented and extends the known substrate tolerance of Mod6TE to include substrates with heteroatoms (*e.g.* **2k**, **2l**), large bulky substituents (**2i**), and sp/sp^2 hybridized carbon atoms (*e.g.* **2d**, **2e**).

Role of AT domain in lactone formation

Intriguingly, self-acylation has been reported in some PKSs, and has been attributed to activity catalyzed by the ACP itself.[41,42] Accordingly, in order to determine whether the AT domain of Mod6TE was absolutely required for the observed Mod6TE promiscuity, the AT active site mutation Ser672Ala was introduced into the Mod6TE gene affording the AT-null mutant AT°-Mod6TE (see Experimental and Scheme 3). The AT-null mutant was then subjected to *in vivo* phosphopantetheinylation, purified, and incubated with each extender unit (**2a–l**) and **4** (see Experimental). Subsequent analysis of the reaction mixtures by HPLC and LC-MS revealed that as expected,[13] introduction of the AT active site mutation failed to support triketide lactone formation when extender units **2a–f** and **2j–l** were used (Table 1 and Fig. S6†). However, low abundance molecular ions corresponding to the expected lactone were observed when **2g–i** were used in this assay (Table 1 and Fig. S6†). While the presence of a contaminant *trans*-AT could in principle explain these results, the specificity of known *Escherichia coli* (*E. coli*) *trans*-ATs are not consistent with the activities observed here.[43] To further probe the source of **2g–i** formation *via* the AT-null mutant, AT°-Mod6TE was expressed and purified from an *E. coli* host that lacked Sfp and therefore is not expected to be phosphopantetheinylated (see Experimental). Interestingly, incubation of the *apo* form of AT°-Mod6TE with each extender unit **2a–l** and **4** led to trace quantities of lactone when extenders **2g–i** (Fig. S7†) were included in the reaction mixture. In addition, crude extract prepared from the host *E. coli* strain used to express Mod6TE failed to support formation of **3a–l** from extender units **2a–l**, (data not shown). Taken together, this set of experiments demonstrates that in the case of **2g–i**, neither a catalytically competent AT or ACP are absolutely required for lactone formation. Presumably, the KS of Mod6TE is able to directly utilize **2g–i**, whereby the CoA moiety mimics the ACP-displayed substrate, albeit inefficiently.

Directly probing promiscuity of the KS domain

Next, it was established whether the KS domain of *holo*-Mod6TE limited utilization of the acyl-CoA's that were not detectable substrates for Mod6TE (**2a**, **2f**, **2j**). It was envisioned that Sfp[44] could be used to transfer each extender unit onto the ACP of the *apo*-form of AT°-Mod6TE (Scheme 3). In this way, AT-specificity is bypassed, and the KS specificity directly probed.[45–47] The acyl-CoA specificity of Sfp was first determined using an end-point MS conversion assay using **2a–l** and *apo*-ACP6 from DEBS (see Experimental). In each case, an increase in mass of the ACP was detected that corresponds to the covalent attachment of the acylphosphopantetheine moiety (Table S3 and Fig. S8†). Next, Sfp and the *apo*-form of AT°-Mod6TE were incubated with each extender unit and **4** and the product mixtures analyzed by HPLC and LC-MS (Table 1, Fig. S10 and Table S4†). Conversion of *apo*-AT°-Mod6TE to the corresponding lactone (Scheme 3) was easily identified with the established Mod6TE substrates **2b–e** and **2k–l** (as expected from the wild-type *holo*-Mod6TE data in Fig. 1), while **2h** produced the phenyl-substituted lactone in trace quantities. Notably, three extender units, **2a** (in agreement with an earlier study)[45] and **2g**/**2i**, (used very poorly by the wild-type *holo*-Mod6TE), were revealed as robust substrates (33–64% conversion, relative to that of **3b**) for the AT-null mutant/Sfp system, illustrating that the KS domain is able to utilize fully 9 out of 12 extender units. Although we cannot completely rule out failure of Sfp to load *apo*-AT°-Mod6TE in



Scheme 3 Bypassing AT specificity of Mod6TE using the broad specificity phosphopantetheinyl transferase Sfp to load extender units onto the inactivated AT-null (AT°) mutant of Mod6TE. Following successful condensation, the subsequent *holo*-AT°-Mod6TE cannot be recharged by Sfp, resulting in single turnover to the lactone. SNAc = *N*-acetylcysteamine. Domains in red are those likely required for the indicated transformation. Wavy line on Mod6TE represents the phosphopantetheine prosthetic arm.

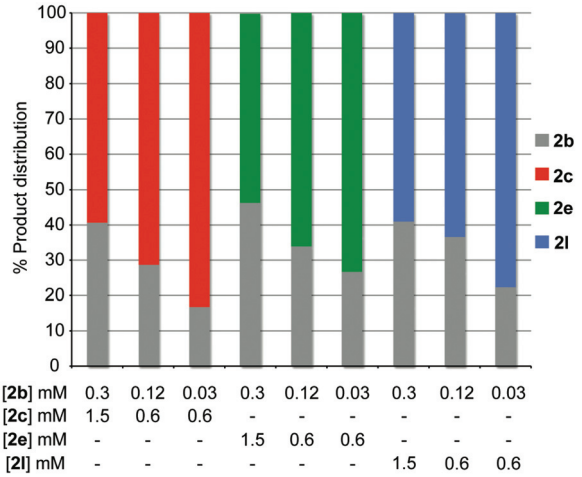


Fig. 2 Mod6TE-catalyzed triketide lactone formation in the presence of **2b** and competing extender unit. The yield of each of the two possible lactones in each reaction is displayed as a percentage of the total product. Each yield was determined in triplicate (standard deviation <10% of the mean). See Experimental for assay details and detection methods.

the case of **2f** and **2j**, it seems likely that the KS of Mod6TE cannot tolerate β-branched extender units or those possessing a free hydroxyl.

Competition experiments

In an effort to describe the relative level of discrimination between the natural extender unit for Mod6TE (**2b**) and the successfully utilized non-native and non-natural extenders, Mod6TE-catalyzed triketide lactone formation was assayed in the presence of acyl-CoA mixtures that included the native extender **2b** and either **2c**, **2e**, or **2l** as competing acyl-CoA's. Several concentrations of **2b**, **2c**, **2e** and **2l** were used in this analysis and are likely below the reported K_M for the **2b**. In this way, the resulting product distribution is likely to at least approximately reflect relative catalytic efficiencies (k_{cat}/K_M) for each substrate. Gratifyingly, when **2b** was present at 0.3 mM and each non-native extender was included at just 5-fold higher concentration, the resulting product mixture contained the lactone derived from the non-native extender unit as the major product, as judged by HPLC analysis (Fig. 2). This fraction was further increased by lowering the concentration of both **2b** and each non-native extender to 0.12 mM and 0.6 mM, respectively, or by using a 20-fold excess of **2c**, **2e**, or **2l**, compared to **2b** (Fig. 2).

Clearly, modulating the extender unit concentrations *in vitro* has some capacity to alter the product distribution to favor production of triketide lactones derived from non-native extender units when the natural extender **2b** is also present.

Discussion

Several strategies are available for the regioselective modification of polyketides, yet the vast majority of these methods are restricted to only a very small number of unique extender units. Interestingly, although some evidence is emerging that PKSs can tolerate at least one or two non-native extender units,[27] the full synthetic scope of PKSs remains poorly described, and AT domains are usually described as being highly selective, at least in terms of discrimination between endogenous extender units. Promiscuous polyketide biosynthetic machinery that can tolerate diverse extender units could enable expansion of diversification strategies to include a broad range of chemical functionality. Here, the promiscuity of the terminal module and thioesterase domain from DEBS was probed using a panel of 12 diverse extender unit acyl-CoA's. Results described in this study demonstrate PKS module turnover of acyl-CoA extender units that (*i*) are non-native for DEBS (**2c–e**, **2g–i**, **2k**, **2l**), (*ii*) are not found in natural biosynthetic systems (**2d**, **2i**, **2k**, **2l**) and (*iii*) contain functional handles (**2d**, **2e**, **2l**) that could enable downstream diversification of polyketide structure *via* chemoselective ligation. Cumulatively, these results considerably expand the known extender unit promiscuity of PKSs.

In an effort to determine whether the substrate tolerance of the KS domain of Mod6TE was limiting utilization of the non-substrate extender unit acyl-CoA's, the AT catalytic residue Ser-672 was mutated to alanine by site-directed mutagenesis, and the AT-null Mod6TE loaded directly using the broad specificity phosphopantetheinyl transferase, Sfp. If the KS domain is broadly tolerant to these extenders, then lactones should be detected in this assay system given the AT domain is inactivated and this is the usual site of extender unit hydrolysis. Our data indicated that the KS domain could not process **2f**/**2j**, but triketide lactone was detected when **2a** was used. Thus, the KS domain will likely require engineering in the case of extender units that resemble **2f**/**2j**. In addition, this result reveals at least some role of the AT-domain in limiting the promiscuity of Mod6TE, likely through hydrolysis of non-native extender units.[27] Presumably, the Ser672Ala mutation reduces or eliminates any editing function normally displayed the AT domain of Mod6TE, and allows loading of **2a** (and perhaps **2g**/**2i**) *via* Sfp and subsequent condensation by the KS. Increasing the effective active site concentration of **2a**/**2g**/**2i** *via* the Sfp-based strategy (Scheme 3) could also play a role in 'rescuing' activity of the AT-null mutant in comparison to the wild-type *holo*-Mod6TE. Although we have yet to directly quantify AT-catalyzed loading and/or subsequent hydrolysis, clearly the intact *holo*-Mod6TE system is sufficiently robust to generate the expected lactone products using most of the extender unit panel. Our results concur with a previous proposal that combinatorial biosynthesis efforts could begin to focus on non-native extenders that are not hydrolyzed by the AT,[27] and now further suggest that this strategy could be expanded to include non-natural extender units. It is intriguing to speculate that given Mod6TE utilizes acyl-CoA extender units with the largest C2 substituent (Me, **2b**) available to the erythromycin producing host, this enzyme does not require discrimination against extender units with larger side-chains (*e.g.* **2c–i**, and **2k–l**), and thereby lacks the hydrolytic editing mechanism that would normally remove the smaller side-chain of **2a**. Consequently, other PKSs that

display promiscuity might be revealed by probing the extender unit specificity of PKS modules that transfer malonyl-CoA derivatives with C2 substituents that are the largest provided by the host organism. The conservation of AT and KS active site amino acid sequences among PKSs[48–50] indicates that such extender unit promiscuity might not be a feature unique to the erythronolide PKS.

Synthetic conversions of the Mod6TE-catalyzed reactions using non-native and non-natural extenders are comparable to previously reported synthesis of triketide lactones using DEBS modules and native extenders, and could be scaled even further to yield ~100 mg lactone.[51] Nevertheless, *in vitro* biosynthesis of complete polyketide scaffolds using type I PKSs[52,53] has yet to match the scale and efficiency of those that involve type II and type III PKSs.[54,55] Clearly, extender unit promiscuity is likely to be better harnessed *via in vivo* polyketide diversification strategies. This study therefore provides a platform for expanding the scope and utility of such strategies. In particular, KS promiscuity could be coupled with inherent[31] or engineered acyl-CoA promiscuity of *trans*-ATs to affect regioselective polyketide modification. Moreover, emerging methods[15] to shift extender unit specificity of a given AT domain towards non-native or non-natural extenders could afford PKS modules tailored towards specific extender units, including those described here. In lieu of detailed kinetic analysis of extender unit specificity, we demonstrated that when Mod6TE was supplied with a mixture of **2b** and non-native extender unit, the major triketide lactone produced was that derived from the non-native substrate when the concentration of the non-native substrate was present at just five-fold excess over the native substrate, **2b** (Fig. 2). This result indicates that specificity between **2b** and other successful extender units might not be too high for alteration by enzyme engineering. Additionally, the modulation of extender unit concentration should prove very helpful for strategies that harness stringent AT/KS domains that have been substituted with promiscuous AT/KS domains, such as those described here. The scope and utility of these *in vivo* strategies could further be expanded by generation of polyketide analogs modified with non-natural handles for chemoselective ligation chemistry (*e.g.* from **2d**, **2e**, **2l**). Incorporation of such handles could enable rapid downstream diversification of polyketides *via* semi-synthesis. Interestingly, promiscuous activities such as those described here often provide suitable starting points for successful directed evolution campaigns or rational redesign. Successful utilization of extender units that include azido and alkynyl handles by Mod6TE or other PKSs suggest various strategies[56–58] for developing high-throughput screens and selections that could be used to identify PKS variants with altered substrate specificities.

Conclusions

The results presented here demonstrate that the AT and KS domains of Mod6TE, and likely those of other PKSs, are highly tolerant of a broad range of extender unit acyl-CoA's. By dissecting and probing Mod6TE *in vitro* by site-directed mutagenesis and complementation *via* Sfp, we determined that the KS is remarkably promiscuous towards diverse extender units, while the AT domain may only play a role in substrate discrimination when native extender units are employed. The vast majority of polyketide biosynthetic diversification strategies have focused on only a very small number of extender units that include limited chemical diversity. The remarkable promiscuity described here sets the stage for significantly expanding the potential scope and utility of such strategies, particularly given the ease with which non-native and non-natural acyl-CoA's can be generated using engineered MatB variants. Future efforts will now focus on harnessing extender unit promiscuity using *in vitro* and *in vivo* methods. In particular, the KS promiscuity discovered here could be harnessed by various precursor directed approaches, and by coupling with *trans*-ATs that display inherent or engineered acyl-CoA promiscuity. Further, a complete description of PKS extender unit promiscuity now provides a guide for future engineering efforts which could include rational redesign of selected AT domain specificity and directed evolution that could for example utilize 'click' handles for high-throughput screens and selections.

Experimental

General

Unless otherwise stated, all materials and reagents were of the highest grade possible and purchased from Sigma (St. Louis, MO). Isopropyl β-D-thiogalactoside (IPTG) was from Calbiochem (Gibbstown, NJ). Bacterial strain *Escherichia coli* BL21 (DE3) pLysS competent cells was from Promega (Madison, WI). Bacterial strain *E. coli* K207-3[59] was a gift from Prof. Keatinge-Clay. Primers were ordered from Integrated DNA Technologies (Coralville, IA). Plasmid pET28a-MatB was as previously described. Analytical HPLC was performed on a Varian ProStar system. Nuclear magnetic resonance spectra were acquired on a Varian Mercury-VX NMR instrument operating at 300 MHz. Chemical shifts (δ) in ^{1}H NMR spectra are expressed in ppm downfield of tetramethylsilane and were referenced to the residual solvent peak.

Synthesis of diketide-SNAc 4

Synthesis of **4** was largely as previously described (Scheme S1†).[60] Meldrum's acid (2.88 g, 20 mmol), recrystallized from toluene, and dry pyridine (3.25 mL, 40 mmol) were added to 40 mL distilled DCM and the solution was stirred at 4 °C. A solution of propionyl chloride (1.75 mL, 20 mmol) in 10 mL distilled DCM was added to the reaction over 10 minutes. The reaction was stirred at 0 °C for 1 hour, then allowed to warm to room temperature and stirred for an additional 5 hours. The reaction was washed with 0.1 M HCl (3 × 50 mL), and the aqueous fractions were extracted with DCM. The combined organic extracts were washed with brine,

dried over $MgSO_4$, and filtered. The solvent was removed under vacuum to yield a red solid, which was recrystallized from petroleum ether to yield the acylated meldrum's acid (**4b**) as a yellow crystalline solid (1.65 g, 41%). ^{1}H-NMR (300 MHz, $CDCl_3$): 3.1 (q, J = 7.5 Hz, 2H), 1.7 (s, 6H), 1.25 (t, J = 7.5 Hz, 3H).

4b (1.0 g, 5 mmol) was dissolved in 5 mL dry toluene. *N*-Acetylcysteamine (0.60 g, 5 mmol) was added, and the reaction was stirred at 80 °C under N_2 for 5 hours. The solvent was removed under vacuum to yield S-(2-acetamidoethyl) 3-oxopentanethioate (**4a**) as a yellow crystalline solid (1.02 g, 94%). ^{1}H-NMR (300 MHz, $CDCl_3$): 6.0 (s, broad, 1H), 3.7 (s, 2H), 3.45 (q, J = 6.0 Hz, 2H), 3.1 (t, J = 6.0 Hz, 2H), 2.56 (q, J = 7.2 Hz, 2H), 1.96 (s, 3H), 1.1 (t, J = 7.2 Hz, 3H).

4a (0.500 g, 2.3 mmol) and potassium *tert*-butoxide (0.30 g, 2.67 mmol) were added to 10 mL distilled THF at 0 °C. Iodomethane (0.85 mL, 13.65 mmol) was added, and the solution was stirred overnight at 0 °C to room temperature. The reaction was quenched with 0.1 M HCl (50 mL) and extracted with ethyl acetate (3 × 50 mL). The combined organic extracts were washed with brine, dried over $MgSO_4$, and filtered. The solvent was removed under vacuum to yield a light yellow oil, which was purified by flash column chromatography (EtOAc) to give the diketide-SNAc (**4**) as a light yellow oil (0.36 g, 68%). ^{1}H-NMR (300 MHz, $CDCl_3$): 6.0 (s, broad, 1H), 3.7 (q, 1H), 3.4 (m, 2H), 3.0 (m, 2H), 2.5 (m, 2H), 1.9 (s, 3H), 1.4 (d, J = 7.2 Hz, 3H), 1.1 (t, J = 7.2 Hz, 3H).

Expression and purification of wild-type and mutant MatB

E. coli BL21(DE3) pLysS competent cells were transformed with the suitable plasmid and positive transformants were selected on LB agar supplemented with 30 μg mL^{-1} kanamycin. A single colony was transferred to LB (3 mL) supplemented with kanamycin (30 μg mL^{-1}) and grown at 37 °C and 250 rpm overnight. The culture was used to inoculate LB media (1 L) supplemented with kanamycin (30 μg mL^{-1}). One liter culture was incubated at 37 °C and 250 rpm to an OD_{600} of 0.6, at which time protein synthesis was induced by the addition of IPTG to a final concentration of 1 mM. After incubation at 18 °C and 200 rpm for 18 h, cells were collected by centrifugation at 5000*g* for 20 min, and resuspended in 100 mM Tris-HCl pH 8.0 (20 mL) containing NaCl (300 mM) and then lysed by sonication. Following centrifugation at 10 000*g*, the soluble extract was loaded onto a 1 mL HisTrap HP column (GE Healthcare, Piscataway, NJ) and purified by fast protein liquid chromatography using the following buffers: wash buffer [20 mM phosphate (pH 7.4) containing 0.5 M NaCl and 20 mM imidazole] and elution buffer [20 mM phosphate (pH 7.4) containing 0.5 M NaCl and 200 mM imidazole]. The purified protein was concentrated using an Amicon Ultra 10000 MWCO centrifugal filter (Millipore Corp., Billerica, MA) and stored as 10% glycerol stocks at −80 °C. Protein purity was verified by SDS-PAGE. Protein quantification was carried out using the Bradford Protein Assay Kit from Bio-Rad.

MatB-catalyzed synthesis of extender unit acyl-CoA's 2a–l

Extender units **2a–h** and **2j–l** were chemo-enzymatically synthesized as previously described.[31,32] Briefly, reactions were performed in 50 μL reaction mixture containing 100 mM sodium phosphate (pH 7), $MgCl_2$ (2 mM), ATP (4 mM), coenzyme A (8 mM), malonate or analog **1a–h**, **1j–l** (16 mM) and wild-type or mutant MatB (10 μg) at 25 °C. The optimal MatB mutant for each conversion was chosen on the basis of steady state kinetic data: WT MatB (**1a**, **1b**, **1j**), T207S/M306I (**1c**, **1e**), T207A/M306I (**1k**), T207G/M306I (**1d**, **1f**, **1h**), T207A (**1g**), and T207G/M306V (**1l**). To ensure >90% conversion to the corresponding acyl-CoA, aliquots were removed after overnight incubation, and quenched with an equal volume of ice-cold methanol, centrifuged at 10 000*g* for 10 min, and cleared supernatants used for HPLC analysis on a Varian ProStar HPLC system. A series of linear gradients was developed from 0.1% TFA in water (A) to methanol (HPLC grade, B) using the following protocol: 0–32 min, 80% B; 32–35 min, 100% A. The flow rate was 1 mL min^{-1}, and the absorbance was monitored at 254 nm using Pursuit XRs C18 column (250 × 4.6 mm, Varian Inc.). Product identity was confirmed by LC-MS as described below and as reported earlier. Synthesis of the 2-phenylethyl analog **2i** proceeded with the commercially available **1i**. MatB mutant T207G/M306I was used to convert **1i** to **2i** and product identity confirmed by LC-MS (Calculated mass 958.1855, observed 958.186 ($[M + H]^{1+}$)).

Mass spectrometry analysis of MatB-synthesized acyl-CoAs

Samples were subjected to negative-ESI LC/MS on a Thermo TSQ Quantum Discovery MAX connected to a UV/Vis diode array detector with a Waters BEH C18, 2.1 × 50 mm, 1.7 μm particle column. A series of linear gradients was developed from water/1 mM ammonium formate (pH 5.3) (A) to methanol (B) using the following protocol: 0–10 min, 3–80% B; 10–11 min, 80–95% B; 11–13 min, 95% B; 13–14 min, 95%–5% B; 14–17.5 min, 5% B.

DEBS *holo*-Mod6TE reactions

DEBS *holo*-Mod6TE reactions were set up by adding 90 μg DEBS *holo*-Mod6TE to 35 μL of 50 mM Tris-HCI (pH 7) containing 2 mM $MgCl_2$, 5 mM **4**, and 4 mM each acyl-CoA (provided by suitable mutant MatB enzymes, see above). Reactions were incubated overnight at room temperature and analyzed by RP-HPLC and mass spectrometry as described below. A series of negative controls that lacked either Mod6TE, each acyl-CoA or diketide-SNAc **4** were also set up and analyzed in the same way (Fig. S2 and Table S2†).

RP-HPLC analysis of DEBS *holo*-Mod6TE reactions

Each reaction sample was quenched with an equal volume of methanol, centrifuged at 10 000*g* for 10 min, and 25 μL used for HPLC analysis. HPLC analysis was performed on a Varian ProStar HPLC system. A series of linear gradients was developed from 0.1% TFA in water (A) to 0.1% TFA in acetonitrile (HPLC grade, B) using the following protocol: 0–40 min,

10–30% B; 40–42 min, 100% B; 42–47 min, 10% B. The flow rate was 1 mL min^{-1}, and the absorbance was monitored at 290 nm using Pursuit XRs C18 column (250 × 4.6 mm, Varian Inc.). For detection of **3g–3i**, the following HPLC protocol was used: 0–40 min, 0–100% B; 40–42 min, 100% B; 42–45 min, 100% A.

Mass spectrometry analysis of *holo*-Mod6TE reactions

For High Res LC-MS analysis of *holo*-Mod6TE reaction products, reaction mixtures were analyzed by positive-ESI LC/MS on a Thermo TSQ Quantum Discovery MAX connected to a UV/Vis diode array detector with a 2.1 mm × 50 mm Agilent XDB C-18 1.8 μm column (Agilent, Santa Clara CA), using a gradient of 25–95% MeOH in 0.1% formic acid/H_2O for 8 min at 1 mL min^{-1}, with detection at 235 nm (thioester bond) and 290 nm (triketide pyrone). For Low Res LC/MS analysis of the Mod6TE reaction products, reaction mixtures were analyzed by positive-ESI LC/MS on a Shimadzu Prominence LC-20 connected to a UV/Vis diode array detector with a 2.1 mm × 50 mm Kinetex C-18 2.6 μm column (Phenomenex, Torrance CA), using the following gradient: 0–4.4 min, 5–99% B; 4–4.9 min, 99% B; 4.9–6 min, 5% B (A – 0.1% formic acid/H_2O, B – 0.1% formic acid/acetonitrile).

Scale up and purification of 3b and 3e

DEBS *holo*-Mod6TE reactions were set up by adding 60 mg DEBS *holo*-Mod6TE to 14 mL of 50 mM Tris-HCI (pH 7) containing 2 mM $MgCl_2$, 10 mM **4**, and 4 mM each acyl-CoA (provided by suitable mutant MatB enzymes, see above). Reactions were incubated for 2 days at room temperature and analyzed by RP-HPLC. The reactions were quenched with an equal volume of ice-cold methanol, centrifuged, and decanted from the precipitated protein. The reactions were concentrated by lyophilization and then extracted 3 times with dichloromethane. The solvent was removed under vacuum, and the residue was dissolved in 2 mL 50% methanol and HPLC purified using solvent A (0.1% TFA in H_2O) and solvent B (0.1% TFA in acetonitrile) and the following protocol: 0–40 min, 10–30% B; 40–42 min, 100% B; 42–47 min, 10% B. The collected fractions were pooled, concentrated by lyophilization, and extracted 3 times with dichloromethane. The solvent was removed under vacuum. The remaining residue was dissolved in $CDCl_3$, and the ^{1}H-NMR spectrum was taken (Fig. S3 and S4†). ^{1}H-NMR **3b** (300 MHz, $CDCl_3$): 2.5 (q, J = 7.2 Hz, 2H), 1.97 (s, 3H), 1.95 (s, 3H), 1.2 (t, J = 7.2 Hz, 3H). ^{1}H-NMR **3e** (300 MHz, $CDCl_3$): 5.6 (m, 1H), 5.2 (m, 2H), 3.1 (m, 2H), 2.5 (q, J = 7.2 Hz, 2H), 1.8 (2, 3H). With purified **3b** in hand, a HPLC calibration curve was constructed (Fig. S5†) and used to determine the % conversion efficiency of the scaled-up Mod6TE reactions, using the same RP-HPLC conditions as described above for analysis of the Mod6TE-catalyzed reactions. Subsequently, the % conversion (from **4**) of the large scale *holo*-Mod6TE catalyzed synthesis of **3b** and **3e** was 4.1 and 2.3, respectively, and is similar to previously reported optimized yields for Mod6TE-catalyzed syntheses.[51]

Preparation of the Mod6TE S672A mutant

Mod6TE S672A plasmid was constructed by using the Stratagene QuikChange II Site-Directed Mutagenesis Kit, as described by the manufacturer using the pET28b/Mod6TE as template and the oligonucleotides Mod6TE-S672A-FOR (5′-TCAGCCGTTATCGGTCATGCTCAGGGCGAAATTGC-3′) and Mod6TE-S672A-REV (5′-GCAATTTCGCCCTGAGCATGACCGATAACGGCTGA-3′) (altered codons underlined). Construct was confirmed to carry the correct mutation by DNA sequencing.

Expression and purification of DEBS *holo*-AT°-Mod6TE

DEBS *holo*-AT°-Mod6TE was over-expressed in *E. coli* K207-3 as an N-terminally His_6-tagged fusion protein as previously described.[29]

DEBS *holo*-AT°-Mod6TE reactions

DEBS *holo*-AT°-Mod6TE reactions were set up by adding 90 μg DEBS *holo*-Mod6TE to 35 μL of 50 mM Tris-HCI (pH 7) containing 2 mM $MgCI_2$, 5 mM **4**, and 4 mM each acyl-CoA (provided by suitable mutant MatB enzymes, see above). Reactions were incubated overnight at room temperature and analyzed by RP-HPLC and mass spectrometry as described for the *holo*-Mod6TE reactions (Fig. S6†).

Expression and purification of DEBS *apo*-AT°-Mod6TE

DEBS *apo*-AT°-Mod6TE was over-expressed from vector pET28b/Mod6TE in *E. coli* BL21(DE3) as an N-terminally His_6-tagged fusion protein as previously described.[29]

DEBS *apo*-AT°-Mod6TE reactions

DEBS *apo*-AT°-Mod6TE reactions were set up by adding 1.3 mg DEBS *apo*-AT°-Mod6TE to 40 μL of 50 mM Tris-HCI (pH 8.8) containing 10 mM $MgCI_2$, 5 mM **4**, and 1.6 mM each acyl-CoA (provided by suitable mutant MatB enzymes, see above). Reactions were incubated for 3 h at room temperature and analyzed by RP-HPLC and mass spectrometry as described for the *holo*-Mod6TE reactions (Fig. S7†).

Cloning, expression and purification of DEBS *apo*-ACP6

DEBS *apo*-ACP6 from the *Saccharopolyspora erythraea* erythromycin biosynthetic gene cluster was cloned in pET28a as described[61] and over-expressed as the *apo*-ACP in *E. coli* BL21 (DE3) as an N-terminally His_6-tagged fusion protein as previously described.[43] The *apo*-ACP6 was purified as described for MatB.

Expression and purification of Sfp

The phosphopantetheinyl transferase Sfp for *in vitro* acyl-CoA specificity studies was over-expressed in *E. coli* BL21(DE3) as an N-terminally His_6-tagged fusion protein as previously described.[62] Sfp was purified as described for MatB.

Sfp reactions with *apo*-ACP6 from DEBS

Reactions containing 100 mM sodium phosphate (pH 7), $MgCl_2$ (2 mM), ATP (4 mM), coenzyme A (4 mM), **1a–l** (16 mM)

and suitable MatB mutant (10 μg) at a final volume of 50 μL were incubated at 25 °C for 24 h, or until conversion to the acyl-CoA was complete (as judged by HPLC, see above). The optimal MatB mutant used was as described for the MatB-catalyzed acyl-CoA syntheses (see above). The reaction product mixture was added directly to 100 μL of 50 mM Tris-HCI (pH 8.8) containing 5 mM DTT, 10 mM $MgCl_2$, 100 μg *apo*-ACP6, and 50 μg Sfp, and incubated at 25 °C for 5 h. For LC-MS analysis of acylated ACPs, reaction mixtures were analyzed by positive-ESI LC/MS on a Thermo TSQ Quantum Discovery MAX connected to a UV/Vis diode array detector with a 2.1 mm × 75 mm Poroshell 300SB-C18 5 μM column (Agilent, Santa Clara CA), using a gradient of 25–100% MeOH in 0.1% formic acid/H_2O for 5 min at 1 mL min^{-1} (Fig. S8/S9 and Table S3†).

Sfp-catalyzed acylation of DEBS *apo*-AT°-Mod6TE and triketide lactone formation assay

Reactions containing 100 mM sodium phosphate (pH 7), $MgCl_2$ (2 mM), ATP (4 mM), coenzyme A (4 mM), **1a–l** (16 mM) and suitable MatB mutant (10 μg) at a final volume of 50 μL were incubated at 25 °C for 24 h, or until conversion to the acyl-CoA was complete (as judged by HPLC, see above). The optimal MatB mutant used was as described for the MatB-catalyzed acyl-CoA syntheses (see above). 20 μL of the reaction product mixture was added directly to 80 μL of 50 mM Tris-HCI (pH 8.8) containing, 10 mM $MgCl_2$, 1.3 mg *apo*-AT °-Mod6TE (see above for expression and purification), and 20 μg Sfp, and incubated at 25 °C for 3 h. Reactions were analyzed by RP-HPLC and LC-MS analysis as described for the *holo*-Mod6TE reactions (Fig. S10 and Table S4†), except 50 μL of the reaction mixture was used for HPLC analysis instead of 25 μL.

Competition experiments

Triketide lactone formation catalyzed by *holo*-Mod6TE was assayed in the presence of varying concentrations of **2b** and each **2c**, **2e**, and **2l**. Concentrations of **2b**/non-native extender unit were 0.3/1.5 mM, 0.12/0.6 mM, and 0.03/0.6 mM, respectively. Aside from extender unit concentrations, assay conditions were the same as that for "DEBS *holo*-Mod6TE reactions", described above, while detection was the same as that for "Sfp-catalyzed acylation of DEBS *apo*-AT°-Mod6TE and triketide lactone formation assay", described above.

Acknowledgements

This study was supported in part by an NSF CAREER Award to G.J.W. (CHE-1151299). The authors would like to thank the Mass Spectrometry Facility at NC State, and Daniel Santi and Dr Adrian Keatinge-Clay for the kind gift of Mod6TE.

Notes and references

1 B. M. Trost and G. B. Dong, *Nature*, 2008, **456**, 485–488.
2 E. M. Stang and M. C. White, *Nat. Chem.*, 2009, **1**, 547–551.
3 B. A. DeChristopher, B. A. Loy, M. D. Marsden, A. J. Schrier, J. A. Zack and P. A. Wender, *Nat. Chem.*, 2012, **4**, 705–710.
4 P. A. Wender and A. J. Schrier, *J. Am. Chem. Soc.*, 2011, **133**, 9228–9231.
5 J. M. Langenhan, N. R. Peters, I. A. Guzei, F. M. Hoffmann and J. S. Thorson, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 12305–12310.
6 X. Fu, C. Albermann, J. Jiang, J. Liao, C. Zhang and J. S. Thorson, *Nat. Biotechnol.*, 2003, **21**, 1467–1469.
7 S. N. Dupuis, A. W. Robertson, T. Veinot, S. M. A. Monro, S. E. Douglas, R. T. Syvitski, K. B. Goralski, S. A. McFarland and D. L. Jakeman, *Chem. Sci.*, 2012, **3**, 1640–1644.
8 B. Ruan, K. Pong, F. Jow, M. Bowlby, R. A. Crozier, D. Liu, S. Liang, Y. Chen, M. L. Mercado, X. Feng, F. Bennett, D. von Schack, L. McDonald, M. M. Zaleska, A. Wood, P. H. Reinhart, R. L. Magolda, J. Skotnicki, M. N. Pangalos, F. E. Koehn, G. T. Carter, M. Abou-Gharbia and E. I. Graziani, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 33–38.
9 J. Kennedy, *Nat. Prod. Rep.*, 2008, **25**, 25–34.
10 A. Kirschning and F. Hahn, *Angew. Chem., Int. Ed.*, 2012, **51**, 4012–4022.
11 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.
12 R. McDaniel, A. Thamchaipenet, C. Gustafsson, H. Fu, M. Betlach and G. Ashley, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 1846–1851.
13 P. Kumar, A. T. Koppisch, D. E. Cane and C. Khosla, *J. Am. Chem. Soc.*, 2003, **125**, 14307–14312.
14 C. D. Reeves, S. Murli, G. W. Ashley, M. Piagentini, C. R. Hutchinson and R. McDaniel, *Biochemistry*, 2001, **40**, 15464–15470.
15 U. Sundermann, K. Bravo-Rodriguez, S. Klopries, S. Kushnir, H. Gomez, E. Sanchez-Garcia and F. Schulz, *ACS Chem. Biol.*, 2013, **8**, 443–450.
16 F. Del Vecchio, H. Petkovic, S. G. Kendrew, L. Low, B. Wilkinson, R. Lill, J. Cortes, B. A. M. Rudd, J. Staunton and P. F. Leadlay, *J. Ind. Microbiol. Biotechnol.*, 2003, **30**, 489–494.
17 D. L. Stassi, S. J. Kakavas, K. A. Reynolds, G. Gunawardana, S. Swanson, D. Zeidner, M. Jackson, H. Liu, A. Buko and L. Katz, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 7305–7309.
18 C. Hertweck, *Angew. Chem., Int. Ed.*, 2009, **48**, 4688–4716.
19 Y. A. Chan, A. M. Podevels, B. M. Kevany and M. G. Thomas, *Nat. Prod. Rep.*, 2009, **26**, 90–114.
20 M. C. Wilson and B. S. Moore, *Nat. Prod. Rep.*, 2011, **29**, 72–86.
21 Y. A. Chan and M. G. Thomas, *Methods Enzymol.*, 2009, **459**, 143–163.
22 A. S. Eustaquio, 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, **106**, 12295–12300.
23 Y. Liu, C. Hazzard, A. S. Eustaquio, K. A. Reynolds and B. S. Moore, *J. Am. Chem. Soc.*, 2009, **131**, 10376–10377.

24 M. C. Wilson, S. J. Nam, T. A. Gulder, C. A. Kauffman, P. R. Jensen, W. Fenical and B. S. Moore, *J. Am. Chem. Soc.*, 2011, **133**, 1971–1977.

25 C. D. Reeves, L. M. Chung, Y. Liu, Q. Xue, J. R. Carney, W. P. Revill and L. Katz, *J. Biol. Chem.*, 2002, **277**, 9155–9159.

26 X. Ruan, A. Pereda, D. L. Stassi, D. Zeidner, R. G. Summers, M. Jackson, A. Shivakumar, S. Kakavas, M. J. Staver, S. Donadio and L. Katz, *J. Bacteriol.*, 1997, **179**, 6416–6425.

27 S. A. Bonnett, C. M. Rath, A. R. Shareef, J. R. Joels, J. A. Chemler, K. Hakansson, K. Reynolds and D. H. Sherman, *Chem. Biol.*, 2011, **18**, 1075–1081.

28 A. Y. Chen, N. A. Schnarr, C. Y. Kim, D. E. Cane and C. Khosla, *J. Am. Chem. Soc.*, 2006, **128**, 3067–3074.

29 A. J. Hughes and A. Keatinge-Clay, *Chem. Biol.*, 2011, **18**, 165–176.

30 N. L. Pohl, M. Hans, H. Y. Lee, Y. S. Kim, D. E. Cane and C. Khosla, *J. Am. Chem. Soc.*, 2001, **123**, 5822–5823.

31 I. Koryakina, J. McArthur, S. Randall, M. M. Draelos, E. M. Musiol, D. C. Muddiman, T. Weber and G. J. Williams, *ACS Chem. Biol.*, 2013, **8**, 200–208.

32 I. Koryakina and G. J. Williams, *ChemBioChem*, 2011, **12**, 2289–2293.

33 R. S. Gokhale, S. Y. Tsuji, D. E. Cane and C. Khosla, *Science*, 1999, **284**, 482–485.

34 A. Ranganathan, M. Timoney, M. Bycroft, J. Cortes, I. P. Thomas, B. Wilkinson, L. Kellenberger, U. Hanefeld, I. S. Galloway, J. Staunton and P. F. Leadlay, *Chem. Biol.*, 1999, **6**, 731–741.

35 B. S. Kim, T. A. Cropp, G. Florova, Y. Lindsay, D. H. Sherman and K. A. Reynolds, *Biochemistry*, 2002, **41**, 10827–10833.

36 K. Watanabe, C. C. C. Wang, C. N. Boddy, D. E. Cane and C. Khosla, *J. Biol. Chem.*, 2003, **278**, 42020–42026.

37 G. F. Liou, J. Lau, D. E. Cane and C. Khosla, *Biochemistry*, 2003, **42**, 200–207.

38 A. F. Marsden, P. Caffrey, J. F. Aparicio, M. S. Loughran, J. Staunton and P. F. Leadlay, *Science*, 1994, **263**, 378–380.

39 S. J. Moss, C. J. Martin and B. Wilkinson, *Nat. Prod. Rep.*, 2004, **21**, 575–593.

40 B. Wilkinson, G. Foster, B. A. Rudd, N. L. Taylor, A. P. Blackaby, P. J. Sidebottom, D. J. Cooper, M. J. Dawson, A. D. Buss, S. Gaisser, I. U. Bohm, C. J. Rowe, J. Cortes, P. F. Leadlay and J. Staunton, *Chem. Biol.*, 2000, **7**, 111–117.

41 C. J. Arthur, A. E. Szafranska, J. Long, J. Mills, R. J. Cox, S. C. Findlow, T. J. Simpson, M. P. Crump and J. Crosby, *Chem. Biol.*, 2006, **13**, 587–596.

42 A. Misra, S. K. Sharma, N. Surolia and A. Surolia, *Chem. Biol.*, 2007, **14**, 775–783.

43 A. T. Koppisch and C. Khosla, *Biochemistry*, 2003, **42**, 11057–11064.

44 L. E. Quadri, P. H. Weinreb, M. Lei, M. M. Nakano, P. Zuber and C. T. Walsh, *Biochemistry*, 1998, **37**, 1585–1595.

45 M. Hans, A. Hornung, A. Dziarnowski, D. E. Cane and C. Khosla, *J. Am. Chem. Soc.*, 2003, **125**, 5366–5374.

46 P. Kumar, C. Khosla and Y. Tang, *Methods Enzymol.*, 2004, **388**, 269–293.

47 E. M. Musiol, T. Hartner, A. Kulik, J. Moldenhauer, J. Piel, W. Wohlleben and T. Weber, *Chem. Biol.*, 2011, **18**, 438–444.

48 G. Yadav, R. S. Gokhale and D. Mohanty, *J. Mol. Biol.*, 2003, **328**, 335–363.

49 S. Donadio and L. Katz, *Gene*, 1992, **111**, 51–60.

50 J. F. Aparicio, I. Molnar, T. Schwecke, A. Konig, S. F. Haydock, L. E. Khaw, J. Staunton and P. F. Leadlay, *Gene*, 1996, **169**, 9–16.

51 A. D. Harper, C. B. Bailey, A. D. Edwards, J. F. Detelich and A. T. Keatinge-Clay, *ChemBioChem*, 2012, **13**, 2200–2203.

52 R. Pieper, G. Luo, D. E. Cane and C. Khosla, *Nature*, 1995, **378**, 263–266.

53 J. D. Mortison, J. D. Kittendorf and D. H. Sherman, *J. Am. Chem. Soc.*, 2009, **131**, 15784–15793.

54 M. I. Kim, S. J. Kwon and J. S. Dordick, *Org. Lett.*, 2009, **11**, 3806–3809.

55 Q. Cheng, L. Xiang, M. Izumikawa, D. Meluzzi and B. S. Moore, *Nat. Chem. Biol.*, 2007, **3**, 557–558.

56 A. J. Link, M. K. Vink, N. J. Agard, J. A. Prescher, C. R. Bertozzi and D. A. Tirrell, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 10180–10185.

57 I. Chen, B. M. Dorr and D. R. Liu, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 11399–11404.

58 Y. Zou and J. Yin, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 5664–5667.

59 S. Murli, J. Kennedy, L. C. Dayem, J. R. Carney and J. T. Kealey, *J. Ind. Microbiol. Biotechnol.*, 2003, **30**, 500–509.

60 S. K. Piasecki, C. A. Taylor, J. F. Detelich, J. Liu, J. Zheng, A. Komsoukaniants, D. R. Siegel and A. T. Keatinge-Clay, *Chem. Biol.*, 2011, **18**, 1331–1340.

61 Z. Ye, M. Bair, H. Desai and G. J. Williams, *Mol. BioSyst.*, 2011, **7**, 3152–3156.

62 J. Yin, A. J. Lin, D. E. Golan and C. T. Walsh, *Nat. Protoc.*, 2006, **1**, 280–285.

