Acyl chain elongation drives ketosynthase substrate selectivity in *trans*-acyl transferase polyketide synthases

Matthew Jenner[a], José Pedro Afonso[a], Hannah R. Bailey[a], Sarah Frank[b], Annette Kampa[b], Jörn Piel*[b] and Neil J. Oldham*[a]

Abstract: Type I modular polyketide synthases (PKSs), responsible for the biosynthesis of many biologically active agents, possess a ketosynthase (KS) domain within each module to catalyze chain elongation. Acylation of the KS active site Cys residue is followed by transfer to malonyl-acyl carrier protein, yielding an extended β-ketoacyl chain. To date, the precise contribution of KS selectivity in controlling product fidelity has been unclear. We submitted six KS domains from the *trans*-acyl transferase PKSs to a mass spectrometry-based elongation assay, and identified higher substrate selectivity in the elongating step than in preceding acylation. A close correspondence between observed KS selectivity and that predicted by phylogenetic analysis was seen. Our findings provide insights into the mechanism of KS selectivity in this important group of PKSs, can serve as guidance for engineering, and show that targeted mutagenesis can be used to expand the repertoire of acceptable substrates.

Polyketide synthases (PKSs) are responsible for the production of a vast array of biologically active compounds[1]. All PKSs build-up polyketide products by a series of chain-elongation steps, and are classified into types based upon their architecture. Bacterial type I modular PKSs are large multi-domain enzyme complexes. Each module is dedicated to one round of chain elongation and subsequent processing of its particular intermediate. The array of modules has been likened to a production line, with intermediates passed along in sequence. Within a PKS module, a single cycle of biosynthetic chain elongation minimally requires the activity of an acyltransferase (AT), to supply a malonyl or related extender unit, a ketosynthase (KS) domain to catalyze a Claisen-type chain elongation between the extender and growing acyl chain, and an acyl carrier protein (ACP) domain to tether the intermediates to the PKS via a thioester link (Scheme 1a)[2]. Chain elongation can be followed by modification of the β-keto product by processing domains such as ketoreductases (KR), dehydratases (DH) and enoylreductases (ER)[3]. These molecular factories therefore offer the prospect of novel bioactive natural products through rational engineering strategies[4], but questions remain as to how 'acceptable' intermediate are recognized by PKS domains.

In 'textbook' PKSs, such as that responsible for erythromycin biosynthesis, a 1:1 correspondence between PKS architecture and product structure usually occurs. This principle of co-linearity has greatly assisted assignment of each product to its synthase[5]. PKSs that fall into this category are termed *cis*-AT PKSs, due to the presence of an integral AT domain within each module. The *trans*-AT PKSs are an architectural variant of type I synthases that utilize free-standing AT enzymes[6].They often incorporate non-canonical modules and novel enzymatic domains, which results in poor biosynthetic assignment of clusters using the co-linearity rules referred to above[7]. *trans*-AT systems constitute approximately 40% of all bacterial multimodular PKSs[8], and represent a major, but poorly characterized, enzyme class of high relevance for drug discovery.

One proposed mechanism for the maintenance of polyketide structural integrity is based on substrate specificity of the KS domains. Previous work on the *cis*-AT PKSs responsible for the biosynthesis of erythromycin and its relatives, has revealed that unnatural substrates were able to acylate a KS domain, but no elongated product was observed to be released by a downstream thioesterase[9]. It was suggested that the intermediate 'stalled' on the KS domain. Clear evidence for a KS 'gatekeeper' role has been found in the iterative *cis*-AT PKS which produces aureothin.[10] In contrast, some chemobiosynthetic studies, where the PKS is fed a non-natural acyl starter unit, have successfully produced polyketide analogs, and attempts to engineer artificial fusions of modules from eight *cis*-AT PKSs resulted in competent enzymatic assemblies, indicating more relaxed KS substrate selectivity in these cases[11]. Recent work on pikromycin module 5 has provided new structural insights into the dynamics of the ACP domain within a module[12]. It showed that the KR domain undergoes a large conformational change upon acylation, and that the position of the ACP domain within the module is dependent upon the degree of processing of its acyl chain. This mechanism may contribute to PKS selectivity by preventing access of the incorrect substrate to the downstream KS. Thus, the precise role of KS domains, from *cis*-AT PKSs, in controlling product fidelity remains unclear.

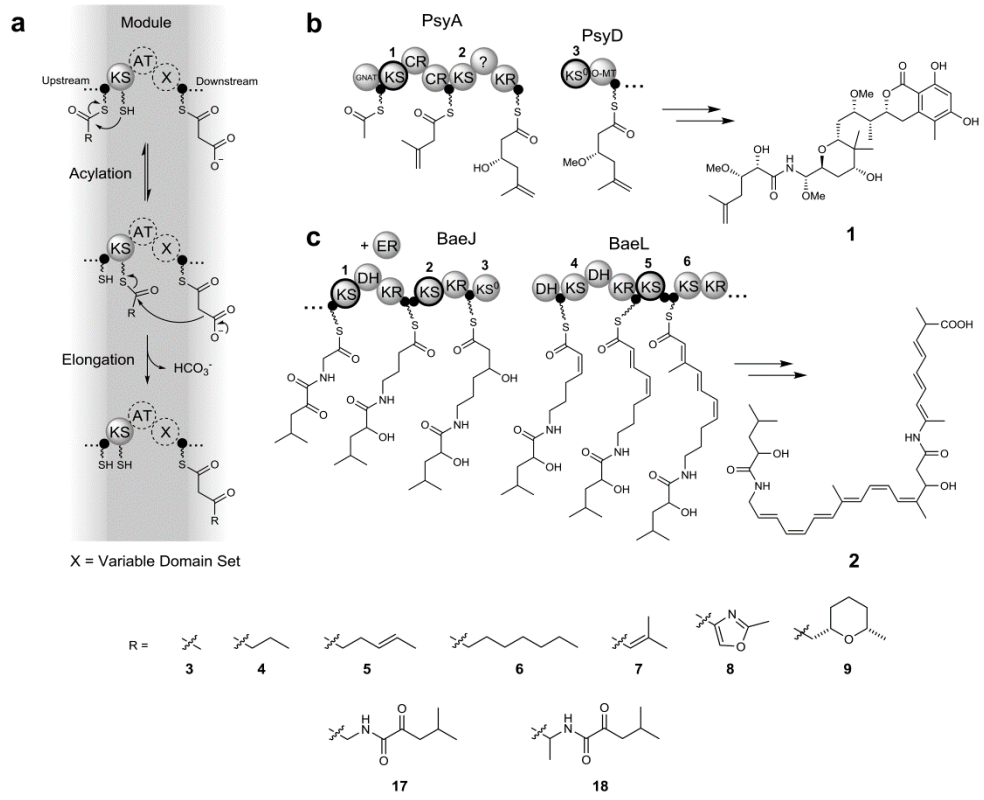
In *trans*-AT PKSs the situation is believed to be rather different. Here KS domains have been shown to be closely correlated, at the sequence level, with their predicted substrates[13]. Phylogenetic analysis gives rise to a series of KS clades, each linked to specific types of intermediate (e.g. β-keto, β-hydroxy, enoyl etc.). No such correlation was found for *cis*-AT PKSs. This approach provides improved functional prediction of *trans*-AT biosynthetic clusters when compared with co-linearity rules, suggesting that KS substrate selectivity is an important property of *trans*-AT PKSs. To test these predictions, previous work by us examined the initial acylation step of *trans*-AT KSs (Scheme 1a). Although some specificity was observed, the tolerance profile was not as well defined as the phylogenetic grouping might suggest[14]. We suspected that further selectivity might be afforded at the elongation step of the KS-catalyzed reaction, where the acyl chain is irreversibly transferred from the KS to ACP-bound malonate to form a new C-C bond. Currently no

[a] M. Jenner, Dr. J. P. Afonso, H. R. Bailey, Dr. N. J. Oldham
School of Chemistry
University of Nottingham
University Park, Nottingham NG7 2RD (UK)
E-mail: neil.oldham@nottingham.ac.uk
[b] S. Frank, Dr. A. Kampa, Prof. Dr. J. Piel
Institute of Microbiology
Eidgenössische Technische Hochschule (ETH) Zurich
Wolfgang-Pauli-Strasse 10, 8083, Zurich (Switzerland)

Supporting information for this article is given via a link at the end of the document.((Please delete this text if not appropriate))

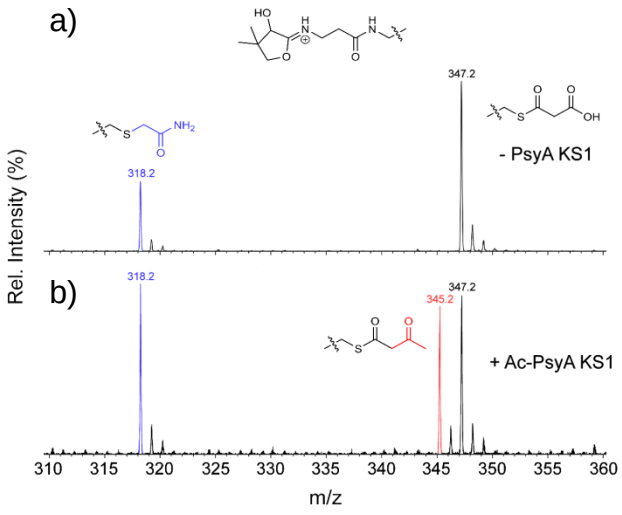
conclusive data on the relative selectivity of these two steps exist for KSs from either *cis*- or *trans*-AT PKSs.

To provide new insights into substrate selectivity exhibited during the elongation step of KS-catalyzed polyketide biosynthesis in *trans*-AT PKSs, a range of KS domains from the psymberin (Psy) and bacillaene (Bae) PKSs [15] were examined.



Scheme 1. KS catalysis and predicted substrates of Psy and Bae KS domains. (a) The catalytic cycle of a KS domain. Acylation by an upstream ACP loads the KS with an acyl intermediate, followed by elongation through a Claisen condensation between the acyl intermediate and a malonyl unit. In *cis*-AT PKSs the AT domain is integral to the PKS, whereas in *trans*-AT systems the activity is supplied by a free-standing AT domain. X represents optional domains. (b) & (c) Partial proposed biosynthetic schemes for psymberin (**1**) and bacillaene (**2**). KS, ketosynthase; KS^0, non-elongating ketosynthase; KR, ketoreductase; DH, dehydratase; CR, crotonase (also known as enoyl-CoA dehydratase, ECH); O-MT, O-methyl transferase; C, NRPS condensation domain; A, NRPS adenylation; GNAT, acetyl-loading AT; and •, ACP domains. KS domains are numbered and those used in this study highlighted in red.

A new mass spectrometry-based assay was developed to quantify ACP-bound elongated β-ketoacyl chains arising from incubation of free-standing malonyl-ACP with a pre-acylated KS domain (see Supporting Information for details). A productive KS-catalyzed reaction resulted in detection of β-ketoacyl-ACP. Quantification was made possible by the addition of an alkyl-ACP standard at the end of the incubation. Using this assay, chain elongation catalyzed by PsyA KS1 was analyzed. The natural substrate for this KS is the acetyl chain, which it is predicted to convert to an acetoacetyl unit (Scheme 1b). Following incubation of acetylated PsyA KS1 with malonyl-ACP, tandem MS activation of the ACP-bound derivatives revealed the presence of elongated acetoacetyl-ACP, with an associated decrease in malonyl-ACP, by detection of the ejected phosphopantetheinyl (PPant) ions (**Figure 1, Figure S5**). PsyA KS1 is located in phylogenetic clade VI with other acetyl-accepting KS domains (**Scheme 1b, Figure S14**). Previous work has shown that functionally, PsyA KS1 is capable of accepting a range of substrates at the acylation stage[14a]. This was tentatively attributed to the lack of evolutionary pressure upon the domain to confer selectivity towards an acetyl unit, due to the presence of a highly specific upstream *N*-acetyl transferase (GNAT) loading domain[16]. A set of seven simple acyl-SNAC thioesters (**3** - **9**) were screened against Psy KS1 for elongation activity by first acylating the KS domain and incubating the product with malonyl-ACP, using the assay described. PsyA KS1 showed a clear preference towards elongation of the acetyl unit (**3**), with the longer carbon chains (**4** - **6**) disfavored (**Figure 2a**, **Figure S7-13**). There was no observable elongation for the β-methyl substrate (**7**), despite initial acylation of the KS. In addition, no elongation was observed with oxazolyl and pyranyl substrates (**8**) and (**9**), which are moieties present in various PKS intermediates[17]. Overall, these results suggest that, not only does PsyA KS1 impose



selectivity at the elongation step of its catalytic action, but that

the selectivity is in excellent agreement with sequence-based phylogenetic predictions (**Figure S14**)[13].

Figure 1. KS catalysed production of β-ketoacyl chains. MS/MS spectra of the activated 7^+ PsyA ACP_3 species following acetyl elongation reactions in the absence (a), and presence (b) of PsyA KS1. Production of the acetoacetyl-ACP (PPant ion *m/z* 345.2) is only seen after incubation with the acetyl-KS domain. Internal standard (PPant ion m/z 318.2) was added after the elongation reaction at a final concentration of 10 μM.

The non-elongating domain, PsyD $KS3^0$ (Scheme 1b), was employed as a control for these experiments since it lacks an essential active-site His residue required for decarboxylation of malonate[7], and was therefore expected to yield no elongated product. Incubation of malonyl-ACP with PsyD $KS3^0$ acylated with a range of acyl groups (**3** - **9**) generated no observable β-ketoacyl products (**10** - 1**6**), despite effective acylation of the KS domain[14a].

In contrast to PsyA KS1, BaeJ KS2 and BaeL KS5 are predicted to accept and elongate much longer acyl chains (**Scheme 1c**). In phylogenetic terms, BaeJ KS2 is located in clade V (**Figure S14**) with other similar KS domains predicted to accept either fully saturated or β,γ-olefinic, unbranched acyl chains. BaeL KS5 is located in clade IX, and is also believed to accept unbranched intermediates (**Scheme 1c, Figure S14**). Acylation studies, carried out on BaeJ KS2 and BaeL KS5 by us previously[14a,17], indicated that unbranched acyl chains were clearly favored. Elongation assays conducted on BaeJ KS2 revealed a remarkably different substrate profile from that seen for PsyA KS1, with the longer acyl chains preferentially elongated to form their β-ketoacyl equivalents (**11 – 13**) (**Figure 2b**). In contrast, the short acetyl chain was elongated only poorly by BaeJ KS2, to yield little of the acetoacetyl-ACP (**10**).

The X-ray structure of BaeJ KS2, acylated with its natural substrate, shows that the entire 10-atom chain is accommodated within a relatively hydrophobic binding pocket[18]. Presumably, favorable interactions between longer carbon-chains and this pocket account for the observed substrate preference of BaeJ KS2. A similar elongation profile was produced by BaeL KS5, with clear preference towards longer unbranched acyl chains (**Figure 2c**). In agreement with previous acylation data, no elongation was observed for a β-Me substrate (**7**), as evidenced by the absence of (**14**). Additionally, the oxazolyl and pyranyl substrates (**8**) and (**9**) yielded no elongated product. Overall, these data demonstrated that both BaeJ KS2 and BaeL KS5 possessed a preference for substrates bearing longer acyl chains, and that both KSs were unable to catalyze elongation of β-branched chains, as predicted by phylogenetic analysis.

Previously we have shown that an active-site mutant of BaeL KS5(M237A) was, in contrast to the wild type (WT), able to accept a β-methyl branched acyl chain at the acylation stage[14a]. Whether this KS variant had the ability to process the bulkier substrate fully, yielding elongated product (**14**), remained an open question. The elongation profile of BaeL KS5(M237A) revealed a similar trend to that of WT BaeL KS5 with regard to the straight chain acyl units, and preferential elongation of longer acyl chains was observed. However, the mutant was also found to be competent to elongate the β-methyl branched chain, as evidenced by the detection of β-ketoacyl-ACP (**14**) in the assay (**Figure 2d**). This suggested that the single Met→Ala point mutation in the active site of KS5 not only permitted acylation by the β-methyl SNAC, but that it also allowed the elongation step to be catalyzed. Pleasingly, the bulky oxazolyl and pyranyl substrates (**8**) and (**9**) were also elongated by this KS variant. This functional modification of BaeL KS5 highlights the attractive possibility of creating engineered *trans*-AT PKSs harboring KS domains with 'non-native' selectivity profiles.

BaeJ KS1 (**Scheme 1c**) was also submitted to the elongation assay. Phylogenetically, BaeJ KS1 is located in clade XVI (**Figure S14**) and must accept and elongate 2-amidoacetyl intermediates. The elongation profile of BaeJ KS1 revealed categorical preference towards 2-amidoacetyl substrates, as evidenced by the elongation of (**17**) and (**18**), with preference shown towards the glycine substrate. There was no observable elongated product for any acyl chains which did not contain the 2-amidoacetyl functionality (**3** - **6**) (**Table S1**).

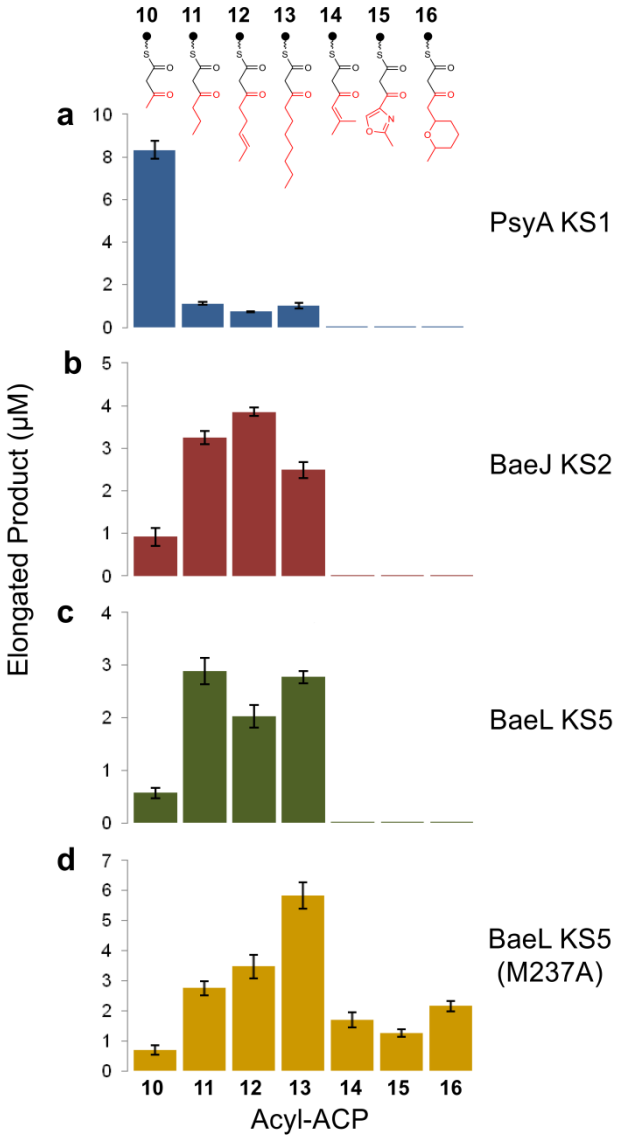


Figure 2. Elongation-based substrate selectivity profiles of KS domains. Elongation selectivity profiles for Psy and Bae KS domains measured by the generation of β-ketoacyl products (**10**-**16**).

The elongation assay performed on PsyA KS1 demonstrated a clear preference for the natural acetyl substrate over longer acyl chains contrasting with the rather broad substrate profile in the acylation step[14a], as described previously. To investigate whether the apparent lack of selectivity during acylation was due to the use of simple SNAC substrate analogues used, in an artificially bimolecular reaction, we cloned the PsyA *apo*-ACP1-KS1 di-domain to provide an accurate representation of intramolecular ACP-KS transfer. When the ACP was loaded with either acetyl- or butyryl-PPant chains, from the corresponding Coenzyme A, this construct showed similar transfer of the two acyl groups from ACP1 to KS1 (see **Figs S16 and S17** and Supporting Information for details of controls). This result shows that the selectivity of KS acylation, even within a native di-domain, is lower than the subsequent elongation step.

Our findings that KS substrate selectivity at the second elongation step exceeds that of acyl loading highlight the potential problem of KS enzymatic stalling. In order to restore enzymatic activity either the acyl chain must hydrolyze off the active site Cys, or the upstream ACP must deacylate the KS in a reversal of the initial acylation step. Once on the upstream ACP, the acyl chain may be removed from the ACP, possibly catalyzed by an acyl hydrolase (AH), or by cleavage at the β–position, as seen in the mupirocin system[19]. In theory, both mechanisms are plausible, but we are unaware of any clear evidence in the literature for reversibility of the KS acylation step in a 'stalled' situation. In order to test the ability of the upstream ACP domain to deacylate the KS active site the ACP1-KS1 didomain was again utilized. The KS1 active site Cys residue was acetylated prior to loading the PPant chain onto the *apo*-ACP1 domain (**Figure S19**)**.** Intramolecular transfer of the acetyl chain from the KS1 domain to the PPant chain of ACP1 was detected by MS/MS analysis as shown in **Figure S21**, demonstrating the inherent reversibility of ACP↔KS acyl transfer (**Scheme 1a**). Control measurements ruled out other sources of PPant acylation (see Supplementary for details).

During polyketide biosynthesis, the maintenance of product fidelity requires a mechanism for intermodular transfer of the correct intermediate and in *trans*-AT PKSs the hypothesis that KS domains fulfil a 'gatekeeper' role has been proposed[17,19]. In principle, KS selectivity could be manifest at the initial acylation step, at the C-C forming elongation step, or both. Our previous work, based on examining substrate tolerance during the KS acylation step, has shown that some selectivity does exist. However, the selectivity is not as strictly defined as phylogenetic analysis would predict and here we show that a high level of selectivity is observed during the KS-catalyzed elongation step, in line with that predicted phylogenetically. Moreover, by using the PsyA ACP1-KS1 didomain we confirm that relatively low selectivity is exhibited at the acylation step, even for a tethered substrate. Thioester bonds are high energy linkages, which form and break rapidly in the presence of a thiol, or thiolate, by a process of transthioesterification[21]. Thus, acyl transfer between an ACP and the immediate downstream KS may be difficult to control, leading to relatively limited enzymatic selectivity, and a potentially reversible process. Indeed, it has been observed that a large excess of SNAC thiol is able to remove an acyl group from the KS active site by transthioesterification[22].

The differential selectivity observed in KS acylation and elongation steps does raise a potential problem: namely non-productive acylation of the enzyme's active site by a substrate incapable of being elongated. Providing that KS acylation is reversible, however, this problem can be overcome, since the acyl chain can be transferred back to the upstream module. The rate constant for back transfer from KS to upstream ACP has been measured for the *cis*-AT DEBS module 2 and found to be very low, leading to the conclusion that acylation is practically irreversible[36]. It should be borne in mind, however, that these observations were made for a productive substrate under steady state conditions, and not for a stalled substrate incapable of elongation. Our findings using PsyA ACP1-KS1 demonstrate that KS acylation in this *trans*-AT PKS didomain is a reversible process (in the absence of a downstream malonyl-ACP), and that this mechanism may allow unloading of non-productive substrates from the KS domain.

In summary, we have shown that five KS domains from the *trans*-AT PKSs responsible for the biosynthesis of psymberin and bacillaene exhibit substrate preference that is in agreement with that predicted by phylogenetic analysis of their sequences. These findings provide functional testing of the proposal that KS domains from *trans*-AT PKSs do possess selectivity for their intermediate. Moreover, we also show that this selectivity can be modulated by rational enzyme modification.

Acknowledgements

We are grateful for funding from the Leverhulme Trust (RPG-2012-578) to J.A. and N.J.O., from the Biotechnology and Biological Science Research Council (BBSRC) for a studentship to M.J, and from the DFG (PI 430/8-1) and the EU (BlueGenics) to J.P.

Keywords: • ketosynthase • mass spectrometry • polyketides • *trans*-AT polyketide synthase

[1] D. J. Newman, G. M. Cragg, *J. Nat. Prod.* **2007**, *70*, 461-477.

[2] a) R. J. Heath, C. O. Rock, *Nat. Prod. Rep.* **2002**, *19*, 581-596; b) Y.–M. Zhang, J. Hurlbert, S. W. White, C. O. Rock, *J. Biol. Chem.* **2006**, *281*, 17390-17399.

[3] a) J. Staunton, J. K. Weissman, *J. Nat. Prod. Rep.* **2001**, *18*, 380-416; b) S.-C. Tsai, B. D. Ames in *Complex Enzymes in Microbial Natural Product Biosynthesis, Part B: Polyketides, Aminocoumarins and Carbohydrates* (Ed. David A. Hopwood), Academic Press, **2009**, pp. 459; c) M. A. Fischbach, C. T. Walsh, *Chem. Rev.* **2006**, *106*, 3468-3496; d) C. Hertweck, *Angew. Chem. Int. Ed.* **2009**, 48, 4688-4716.

[4] M. Till, P. R. Race, *Biotechnol. Lett.* **2014**, *36*, 877-888.

[5] a) S. Kapur *et al.*, *Proc. Natl. Acad. Sci.* **2012**, *109(11)*, 4110-4115; b) C. Khosla, Y. Tang, A. Y. Chen, N. A. Schnarr, D. E. Cane, *Ann. Rev. Biochem.* **2007**, *76*, 195-221.

[6] a) J. Piel, *Nat. Prod. Rep.* **2010**, *27*, 996-1047; b) Y. Q. Cheng, G. L. Tang, B. Shen. *Proc. Natl. Acad. Sci.* **2003**, *100*, 3149-3154.

[7] T. Gulder, M. Freeman, J. Piel in *Topics in Current Chemistry*, Springer Berlin Heidelberg, **2012**, pp.1-53.

[8] R. V. O'Brien, R. W. Davis, C. Khosla, M. E. Hillenmeyer, *J. Antibiot.* **2014**, *67*, 89-97.

[9] K. Watanabe, C. C. C. Wang, C. N. Boddy, D. E. Cane, C. J. Khosla, *J. Biol. Chem.* **2003**, *278*, 42020-42026.

[10] B. Busch *et al. Angew. Chem. Int. Ed.* **2013**, 52, 5285-5289.

[11] a) H. G. Menzella *et al*, *Nat. Biotechnol.* **2005**, *23*, 1171-1176; b) S. Murli *et al. Appl. Environ. Microbiol*. **2005**, 71, 4503-4509; c) J. R. Jacobsen, C. R. Hutchinson, D. E. Cane, C. Khosla, *Science*, **1997**, 277, 367-369.

[12] a) S. Dutta *et al*, *Nature* **2014**, *510(7506)*, 512-517; b) J. R. Whicher *et al*, *Nature* **2014**, *510(7506)*, 560-564.

[13] T. Nguyen *et al*, *Nat. Biotechnol.* **2008**, *26*, 225-233.

[14] a) M. Jenner *et al*, *Angew. Chem. Int. Ed.* **2013**, *52*, 1143-1147; b) C. Kohlhaas *et al*, *Chem. Sci.* **2013**, *4*, 3212-3217.

[15] a) J. Moldenhauer, X. H. Chen, R. Borriss, J. Piel, *Angew. Chem. Int. Ed.* **2007**, *46*, 8195-8197; b) M. Bielitza, J. Pietruszka, *Angew. Chem. Int. Ed.* **2013**, *52*, 10960-10985; c) K. M. Fisch *et al*, *Nat. Chem. Biol.* **2009**, *5*, 494-501.

[16] L. Gu *et al*, *Science* **2007**, *318*, 970-974.

[17] a) B. Kusebauch, N. Brendel, H. Kirchner, H.-M. Dahse, C. Hertweck, *Chembiochem* **2011**, *12*, 2284-2288; b) P. Pöplau, S. Frank, B. I. Morinaka, J. Piel, *Angew. Chem. Int. Ed.* **2013**, *52(50)*, 13215-13218.

[18] D. C. Gay *et al*, Structure **2014**, *22*, 444-451.

[19] a) K. Jensen *et al*, *Chem. Biol.* **2012**, *19(3)*, 329-339; b) J. Wu *et al. ChemBioChem* **2008**, *9*, 1500-1508.

[20] B. Busch *et al*, *Angew. Chem. Int. Ed.* **2013**, *52*, 5285-5289.

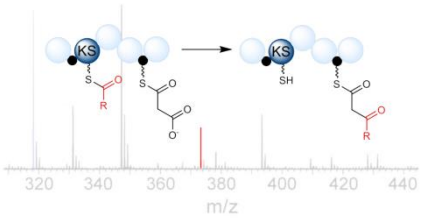
[21] P. J. Bracher, P. W. Snyder, B. R. Bohall, G. M. Whitesides, *Origins Life Evol. B.* **2011**, *41*, 399-412.

[22] J. Q. Wu, K. Kinoshita, C. Khosla, D. E. Cane, *Biochemistry* **2004**, *43*, 16301-16310.

Entry for the Table of Contents (Please choose one layout)

Layout 2:

COMMUNICATION



Extension granted: *In vitro* studies on ketosynthase (KS) domains from *trans*-AT polyketide synthases reveal selectivity within the chain elongation step that mirrors predictions from phylogenetic analysis. Substrate tolerance profiles provide valuable information for bioengineering of polyketide assembly lines.

Matthew Jenner, José Pedro Afonso, Hannah R. Bailey, Sarah Frank, Annette Kampa, Jörn Piel and Neil J. Oldham**

Page No. – Page No.

Acyl chain elongation drives ketosynthase substrate selectivity in *trans*-acyl transferase polyketide synthases