

Insights into 6-Methylsalicylic Acid Bio-assembly by Using Chemical Probes

James S. Parascandolo, Judith Havemann, Helen K. Potter, Fanglu Huang, Elena Riva, Jack Connolly, Ina Wilkening, Lijiang Song, Peter F. Leadlay, and Manuela Tosin*

Abstract: Chemical probes capable of reacting with KS (ketosynthase)-bound biosynthetic intermediates were utilized for the investigation of the model type I iterative polyketide synthase 6-methylsalicylic acid synthase (6-MSAS) *in vivo* and *in vitro*. From the fermentation of fungal and bacterial 6-MSAS hosts in the presence of chain termination probes, a full range of biosynthetic intermediates was isolated and characterized for the first time. Meanwhile, *in vitro* studies of recombinant 6-MSA synthases with both nonhydrolyzable and hydrolyzable substrate mimics have provided additional insights into substrate recognition, providing the basis for further exploration of the enzyme catalytic activities.

Together with peptidic molecules, polyketide natural products constitute one of the most abundant sources of bio-derived and bio-inspired pharmaceuticals: these include widely utilized antibiotic and antitumor agents (for example, erythromycin and doxorubicin) and cholesterol-lowering (statin) drugs.^[1] Polyketide biosynthesis resembles that of fatty acids in the use of decarboxylative Claisen condensation to generate enzyme-bound β -keto thioester intermediates, which are variably processed and ultimately converted to highly diversified products in structure and function. Polyketide synthases (PKSs) can be distinguished as either

modular or iterative: the former (best known as type I modular) are assembly lines comprising multiple sets of domains (modules), with each catalyzing at least one round of chain extension and downstream translocation for further processing. For these PKSs the module order and composition determine the sequence of biosynthetic events,^[2] which can be predicted and manipulated to generate novel compounds.^[3] Conversely iterative synthases are constituted by single enzymes (type I and III) or clusters of discrete proteins (type II) that harbor a limited set of catalytic domains iteratively utilized for intermediate chain growth and elaboration. The rapid and unpredictable nature of substrate processing makes the investigation of iterative synthases much more challenging compared to their modular counterparts. Type I iterative polyketide synthases (iPKSs) most closely resemble fatty acid synthases in their domain organization and modus operandi. They are typical of fungi,^[4] although an increasing number have been reported in bacteria.^[5] Type I iPKSs utilize a diverse range of acyl building blocks (for example, acetate, hexanoate, and benzoate) as starter units for polyketide chain building, and have been classified as non-reducing (NR), partially reducing (PR), or highly reducing (HR) according to the degree of ketone moiety processing occurring throughout product assembly.^[4] Tailoring modifications can occur during and after assembly to yield the final bioactive molecules.^[8] Established mechanisms of iPKS product release include cyclization and hydrolysis mediated by thioesterase (TE) and Claisen cyclase thioesterase (CLC-TE) domains, thioester reduction,^[4,5] pyrone formation,^[6] and product transfer to a nonribosomal peptide synthetase (NRPS) assembly line.^[7] Important mechanistic insights on iPKSs have been gathered by the *in vitro* reconstitution of enzyme activity with putative synthetic substrates,^[9] by genetic manipulation,^[10] and by analysis of protein site-occupancy using advanced mass spectrometry of PKS-bound precursors.^[11] However, key details of the timing of biosynthetic transformations and the basis for substrate discrimination for iPKSs remain elusive, and new tools are needed to uncover them.

6-methylsalicylic acid (6-MSA, **1**, Scheme 1) was the first polyketide to be biosynthetically investigated.^[12] It is produced in various fungi including *Penicillium patulum*, where it is a precursor to the toxin patulin.^[13] 6-MSA is also a key structural moiety of promising antibiotic and anticancer agents, such as chlorothricin, maduropeptin, and neocarzinostatin.^[14] The iPKS 6-methylsalicylic acid synthase (6-MSAS) was first purified from *P. patulum* and characterized as a 188 kDa tetrameric protein.^[15] Early labeling experiments established that 6-MSAS requires one acetyl-CoA (**2**)

[*] J. S. Parascandolo, Dr. J. Havemann, Dr. E. Riva, J. Connolly, Dr. I. Wilkening, Dr. L. Song, Dr. M. Tosin
Department of Chemistry, University of Warwick
Library Road, Coventry CV4 7AL (UK)
E-mail: m.tosin@warwick.ac.uk

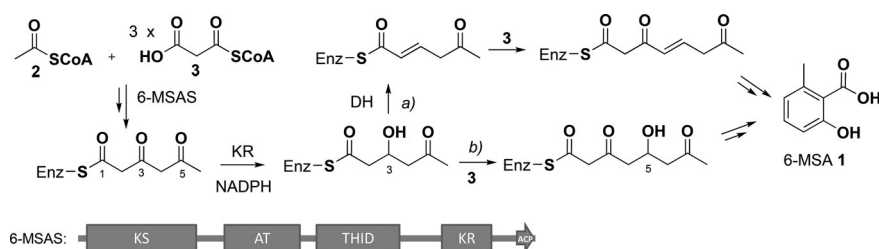
H. K. Potter, Dr. F. Huang
Department of Chemistry, University of Cambridge
Lensfield Road, Cambridge CB2 1EW (UK)

Dr. F. Huang, Prof. P. F. Leadlay
Department of Biochemistry, University of Cambridge
80 Tennis Court Road, Cambridge CB2 1GA (UK)

J. Connolly
Current address:
School of Biosciences, The University of Birmingham
Birmingham B15 2TT (UK)

Supporting information for this article (including synthesis of novel chemical probes **5**, **13**, and **15**; detailed LC-HR-MSⁿ analyses of captured polyketide intermediates *in vivo* and *in vitro*; and the cloning, expression, purification and activity assays of 6-MSAS and 6-MSAS H958A) is available on the WWW under <http://dx.doi.org/10.1002/anie.201509038>.

© 2015 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.



Scheme 1. Overview of biosynthetic hypotheses leading to 6-methylsalicylic acid (6-MSA, **1**): a) enzymatic dehydration of a 3-hydroxytriketide followed by further chain extension leads to a dehydrated enzyme-bound tetraketide, eventually giving **1**; b) formation of a 5-hydroxytetraketide eventually leading to **1** via TH-mediated hydrolysis.^[18c,19b] 6-MSAS comprises ketosynthase (KS), acyltransferase (AT), thioester hydrolase (THID),^[19b] ketoreductase (KR), and acyl carrier protein (ACP) domains.

and three malonyl-CoA molecules (**3**, Figure 1 A) to generate 6-MSA.^[16] Analysis of its gene cluster revealed it encodes one polypeptide chain harboring ketosynthase (KS), acyltransferase (AT), dehydratase (DH), ketoreductase (KR), and acyl carrier protein (ACP) domains^[17] as in a vertebrate fatty acid synthase. Mechanistic studies of purified 6-MSAS using substrate/intermediate analogues and enzyme inhibitors,^[18] as well as enzyme mutagenesis,^[14a,19] have led to two distinct biosynthetic proposals: in the first, DH-catalyzed dehydration of a 3-hydroxytriketide intermediate is followed by a further round of chain extension, *trans* to *cis* isomerization of a double bond, aromatization, and finally thioester hydrolysis (Scheme 1 a).^[18c,19a] In the second, a 3-hydroxytriketide intermediate is directly extended to a 5-hydroxytetraketide, which cyclizes, dehydrates, and aromatizes prior to final product release (Scheme 1 b).^[18c,19b] A recent study of the 6-MSAS-

like enzyme ATX from *Aspergillus terreus* has supported this second route and provided evidence of involvement of a so-called thioester hydrolase (THID) domain in product release.^[19b] The THID domain comprises the previously identified dehydratase (DH) domain together with an adjacent region termed the interdomain (ID) linker, originally identified as a core domain required for subunit-subunit interaction within ATX.^[19a] THID has been shown to catalyze 6-MSA release from a mutant form of ATX (H972A, which would inactivate the DH function); it also catalyzes hydrolysis of the *N*-acetylcysteamine thioester of 6-MSA.^[19b] This suggests that enzyme-catalyzed dehydration of a 3-hydroxytriketide intermediate is not necessary for 6-MSA formation. However, in the absence of direct evidence for the biosynthetic intermediates involved, it remains unclear whether triketide dehydration takes place and whether the sole role of THID is the hydrolytic release of enzyme-bound 6-MSA.

To obtain a complete mechanistic picture of 6-MSA assembly, we have used chain-termination probes for the capture and identification of polyketide intermediates.^[20,21] By competing with ACP-bound malonate extension units for the growing polyketide chain, the chemical probes react with enzyme-bound intermediates and off-load them for LC-MS characterization (Figure 1). The use of these tools has already allowed fresh insights into the timing and the mechanism of modular assembly-line biosynthesis *in vitro*^[20] and *in vivo*,^[21]

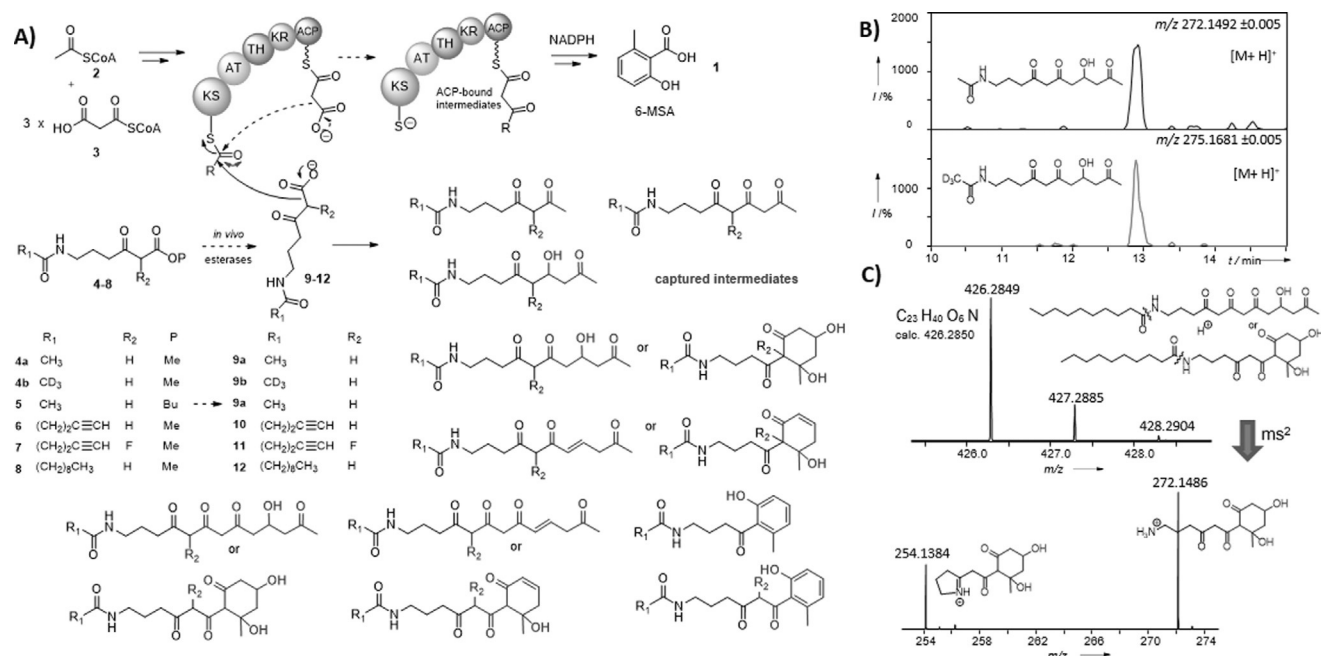


Figure 1. A) Chain termination probes **9–12**, generated *in vivo* from the hydrolysis of the corresponding esters **4–8**,^[21] compete with ACP-bound malonate to off-load 6-MSAS-bound intermediates in *P. patulum*, *E. coli* heterologously expressing 6-MSAS,^[22] and *S. antibioticus* DSM40725 (Supporting Information, Tables 1S–3S); B) LC- HR-MS detection of putative hydroxytetraketide captured from *P. patulum*; C) HR-MSⁿ analyses of a putative hydroxytetraketide resulting from the off-loading of a KS-bound hydroxytetraketide in *P. patulum*.

and, more recently, has unveiled novel opportunities for the generation of unnatural polyketide derivatives.^[21c] We initially used our intermediate-capturing probes for in vivo studies on fungal and bacterial strains harboring 6-MSAS genes, including the natural 6-MSA producer *P. patulum*, an engineered *E. coli* host strain heterologously expressing *P. patulum* 6-MSAS (*E. coli* BAP1 pKOS007-109),^[22] and *S. antibioticus* DSM40725 (producer of chlorothricin).^[14a] Each strain was grown in the presence of substrates **4–8**, which are hydrolyzed in vivo to the corresponding carboxylates **9–12** (Figure 1 A; Supporting Information, Figure 2S).^[20]

The overall outcome of these in vivo experiments is illustrated in Figure 1 (for details, see the Supporting Information, Tables 1S–3S and following figures). In most of the ethyl acetate extracts from both fungal and bacterial hosts, a series of trapped intermediates, including diketides, triketides, reduced triketides, and a range of putative hydroxy, dehydrated, and aromatized tetraketides, were identified by HR-LC-MS: these would directly reflect the nature of ACP-bound substrates in 6-MSA assembly. Besides, putative hydroxy, dehydrated, and aromatized pentaketides arising from the off-loading of 6-MSAS-bound tetraketides were also identified (Figure 1 C and the Supporting Information). All the captured intermediates, absent in control samples, were characterized by MSⁿ analysis, showing diagnostic peaks resulting from the loss of *N*-acyl chains and subsequent cyclic imine formation (Figure 1 C and the Supporting Information). For the majority of the putative tetraketides and pentaketides, multiple peaks were observed: these may arise from isomerization, cyclization and dehydration events which can be spontaneous or enzyme-catalyzed. A distinction between hydroxy, dehydrated, and aromatized advanced species was made on the basis of variable LC retention times as well as of detected accurate masses. On the same basis, distinct species with masses corresponding to dehydrated triketides could not be identified. From the lack of direct evidence for dehydrated triketides and the identification of the putative hydroxy tetraketides and pentaketides, it appears that, whether the PKS is of fungal or of bacterial origin, route b) of Scheme 1 is followed.

Nonetheless, to seek additional confirmation of these insights and further dissect 6-MSAS catalytic activities, we also utilized recombinant *P. patulum* 6-MSAS from heterologous *E. coli* BAP1 host strain,^[23] as well as an additional mutant form of the enzyme (6-MSAS H958A) bearing an alanine in place of a histidine in the THID active site for in vitro assays.^[22] The capture of biosynthetic intermediates in vitro proved much more challenging than in vivo. Using probes **9a–b** (generated from pig liver esterase-assisted hydrolysis of **4a–b**),^[19b] only intermediates from two rounds of chain extension were consistently identified in the ethyl acetate extracts of 6-MSAS assays (Supporting Information, Figure 43S). When recombinant 6-MSAS was primed with acetoacetyl-CoA instead of acetyl-CoA in the attempt to improve advanced intermediate capture, the accumulation of a possibly dehydrated triketide was observed (Supporting Information, Figure 46S).

To investigate whether this species could have been enzymatically formed, a racemic 3-hydroxytriketide substrate

mimic **13** was synthesized in five steps from **4a** (Supporting Information) and tested as a substrate for both 6-MSAS and its H958A mutant. In neither case was **13** enzymatically dehydrated (Figure 2 A; Supporting Information, Figure 47S); rather, we found that purified **13** dehydrated over

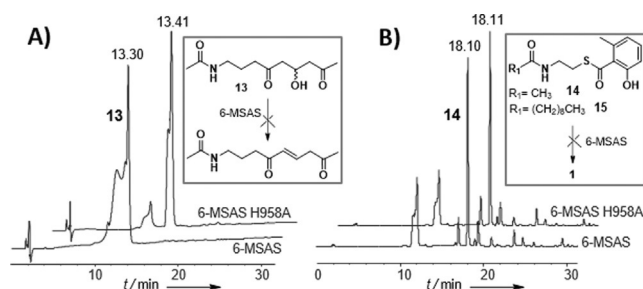


Figure 2. Analytical HPLC analyses showing that: A) the synthetic triketide mimic **13** is not dehydrated by 6-MSAS nor by its mutant 6-MSAS H958A and B) the thioester substrates **14** and **15** are not hydrolyzed by 6-MSAS nor by its mutant 6-MSAS H958A.

long-term storage. Surprisingly, when 6-MSAS and the H958A mutant were incubated with the *N*-acetylcysteamine thioester of 6-MSA **14**, previously utilized to probe the THID function in ATX,^[12] no free 6-MSA was generated (Figure 2 B; Supporting Information, Figure 48S). An *N*-decanoyl thioester analogue **15** was synthesized as an additional substrate, with the idea of utilizing a long acyl chain to mimic the phosphopantetheine cofactor of ACP. However, **15** was also resistant to hydrolysis by either enzyme (Supporting Information, Figure 49S). This unexpected outcome suggests that 6-MSAS differs from ATX in that covalent attachment of tetraketide intermediates to the ACP or coenzyme A might be necessary for their processing. Alternatively, this may indicate that readily aromatized thioesters are not true substrates for THID domains. The results reported herein strongly point towards no enzymatic dehydration taking place at a triketide stage of 6-MSA assembly, so loss of water must occur at the tetraketide stage. If this is enzyme-catalyzed, the configuration of the resulting alkene (from an *R*-alcohol, as recently established for the 6-MSA-like mellein synthase)^[24] would likely be *trans*. On the basis of multiple peaks observed for dehydrated tetraketides (for example, Supporting Information, Figures 24S), it is tempting to speculate that, along with final thioester hydrolysis, the THID might act as a template domain to aid *trans* to *cis* double bond isomerization, and/or cyclization and aromatization; however, this remains undetermined. Amongst iPKSs with significant homology to 6-MSAS and leading to 6-MSA related products, the THID domain is highly conserved (Supporting Information, Figure 50S). Its presence in non-reducing iPKSs such as the orsellinic acid synthase supports its role in product cyclization and release.^[14a] Intriguingly, a THID is not present in MicC, an iPKS responsible for the formation of 6-pentasalicylic acid in micacocidin biosynthesis, whose assembly allegedly proceeds similarly to that of 6-MSA.^[7b,25] Although bioinformatic analysis usefully pinpoints the sim-

ilarities of 6-MSAS to ATX and other iPKSs (Supporting Information, Figure S1S), it cannot be used to infer subtle differences in substrate recognition and processing between these multienzymes. For this the use of chemical probes appears a promising approach towards a fuller understanding and exploitation of iPKS catalysis. Through the use of the chain termination probes 9–12 we have indeed gathered the first direct and comprehensive evidence for the course of substrate processing on a type I iterative PKS. The probes were successfully used in a fungal host and both Gram-positive and Gram-negative bacteria, with the putative intermediates providing preliminary insights into the kinetics of 6-MSA assembly. In general, the first two rounds of chain extension/processing appear relatively slow, whereas more distinct differences can be observed in advanced intermediate accumulation in the different 6-MSAS hosts (Supporting Information, Tables 1S–3S and following analyses): these may depend on probe efficiency and uptake to some extent, as well as on specific kinetic programming of product assembly within a particular host. Significantly 6-MSAS displayed unexpected substrate flexibility, in that it was able to accept malonate surrogates of different chain lengths and bearing various functionalities (including alkyne and fluorine moieties) at different positions and for every round of chain extension, and also to generate novel pentaketide products. This opens the possibility of utilizing iPKSs for the generation of novel unnatural products employing different extender^[21c] as well as starter units.^[26]

Acknowledgements

The authors gratefully acknowledge BBSRC (project grant BB/J007250/1 to M.T.); EPSRC (DTA studentships to J.S.P. and H.K.P.); Institute of Advanced Studies (IAS) at Warwick (Postdoctoral Fellowship to E.R.); FP7 (Marie Curie Intra-european Fellowship to I.W.); Prof. C. Khosla (Stanford University) for providing the plasmids pKOS12-65a and pKOS007-95 for construction of pKOS007-109; Dr Sue Slade and Dr Cleidiane Zampronio (Warwick), and Asha Boodhun and Paul Skelton (Cambridge) for assistance with LC-HRMSⁿ analyses on Orbitrap Fusion and Classic instruments; and Prof Greg Challis for the use of a MaXis Bruker Impact instrument.

Keywords: chemical probes · dehydratase · iterative polyketide catalysis

How to cite: *Angew. Chem. Int. Ed.* **2016**, *55*, 3463–3467
Angew. Chem. **2016**, *128*, 3524–3528

- [1] a) K. J. Weissman, *Methods Enzymol.* **2009**, *459*, 3–16; b) C. Hertweck, *Angew. Chem. Int. Ed.* **2009**, *48*, 4688–4716; *Angew. Chem.* **2009**, *121*, 4782–4811.
[2] a) J. Cortés, S. F. Haydock, G. A. Roberts, D. J. Bevitt, P. F. Leadlay, *Nature* **1990**, *348*, 176–178; b) S. Donadio, M. J. Staver, J. B. McAlpine, S. J. Swanson, L. Katz, *Science* **1991**, *252*, 675–679.

- [3] a) M. Cumming, R. Breitling, E. Takano, *FEMS Microbiol. Lett.* **2014**, *351*, 116–125; b) C. Hertweck, *Trends Biochem. Sci.* **2015**, *40*, 189–199.
[4] R. J. Cox, *Org. Biomol. Chem.* **2007**, *5*, 2010–2026.
[5] Q. Zhang, B. Pang, W. Ding, W. Liu, *ACS Catal.* **2013**, *3*, 1439–1447.
[6] a) J. Kennedy, K. Auclair, S. G. Kendrew, C. Park, J. C. Vederas, C. R. Hutchinson, *Science* **1999**, *284*, 1368–1372; b) K. Kasahara, T. Miyamoto, T. Fujimoto, H. Oguri, T. Tokiwano, H. Oikawa, Y. Ebizuka, I. Fujii, *ChemBioChem* **2010**, *11*, 1245–1252.
[7] a) K. L. Eley, L. M. Halo, Z. Song, H. Powles, R. J. Cox, A. M. Bailey, C. M. Lazarus, T. J. Simpson, *ChemBioChem* **2007**, *8*, 289–297; b) H. Kage, M. F. Kreutzer, B. Wackler, D. Hoffmeister, M. Nett, *Chem. Biol.* **2013**, *20*, 764–771.
[8] J. Huffman, R. Gerber, L. Du, *Biopolymers* **2010**, *93*, 764–776.
[9] a) Z. Gao, J. Wang, A. Norquay, K. Qiao, Y. Tang, J. C. Vederas, *J. Am. Chem. Soc.* **2013**, *135*, 1735–1738; b) S. A. Ma, J. Li, J. W. Choi, H. Zhou, M. Lee, V. Moorthie, X. Xie, J. T. Kealey, N. Da Silva, J. C. Vederas, Y. Tang, *Science* **2009**, *326*, 589–592.
[10] a) K. M. Fisch, W. Bakeer, A. A. Yakasai, Z. Song, J. Pedrick, Z. Wasil, A. M. Bailey, C. M. Lazarus, T. J. Simpson, R. J. Cox, *J. Am. Chem. Soc.* **2011**, *133*, 16635–16641; b) T. Liu, J. F. Sanchez, Y.-M. Chiang, B. R. Oakley, C. C. C. Wang, *Org. Lett.* **2014**, *16*, 1676–1679; c) A. L. Vagstad, A. G. Newman, P. A. Storm, K. Belecki, J. M. Crawford, C. A. Townsend, *Angew. Chem. Int. Ed.* **2013**, *52*, 1718–1721; *Angew. Chem.* **2013**, *125*, 1762–1765.
[11] a) M. J. Meehan, X. Xie, X. Zhao, W. Xu, Y. Tang, P. C. Dorresteijn, *Biochemistry* **2011**, *50*, 287–299; b) J. M. Crawford, P. M. Thomas, J. R. Scheerer, A. L. Vagstad, N. L. Kelleher, C. A. Townsend, *Science* **2008**, *320*, 243–246.
[12] A. J. Birch, P. A. Massy-Westropp, C. J. Moye, *Aust. J. Chem.* **1955**, *8*, 539.
[13] a) J. D. Bu'Lock, A. J. Ryan, *Proc. Chem. Soc.* **1958**, 222–223; b) S. W. Tanenbaum, E. W. Bassett, *J. Biol. Chem.* **1959**, *234*, 1861–1866.
[14] a) W. Ding, C. Lei, Q. He, Q. Zhang, Y. Bi, W. Liu, *Chem. Biol.* **2010**, *17*, 495–503; b) S. G. Van Lanen, T. J. Oh, W. Liu, E. Wendt-Pienkowski, B. Shen, *J. Am. Chem. Soc.* **2007**, *129*, 13082–13094.
[15] J. B. Spencer, P. M. Jordan, *Biochem. J.* **1992**, *288*, 839–846.
[16] a) P. Dimroth, H. Walter, F. Lynen, *Eur. J. Biochem.* **1970**, *13*, 98–110; b) P. Dimroth, E. Ringelmann, F. Lynen, *Eur. J. Biochem.* **1976**, *192*, 487–498.
[17] a) F. Lynen, M. Tada, *Angew. Chem.* **1961**, *73*, 513–519; b) J. Beck, S. Ripka, A. Siegner, E. Schiltz, E. Schweizer, *Eur. J. Biochem.* **1990**, *192*, 487–498.
[18] a) A. I. Scott, L. C. Beading, N. H. Georgopapadalou, C. R. Subbarayan, *Bioorg. Chem.* **1974**, *3*, 238–248; b) J. B. Spencer, P. M. Jordan, *Chem. Soc. Chem. Commun.* **1990**, 1704–1706; c) J. B. Spencer, P. M. Jordan, *Biochemistry* **1992**, *31*, 9107–9116; d) K.-M. Cheung, S. J. Coles, M. B. Hursthouse, N. I. Johnson, P. M. Shoolingin-Jordan, *Angew. Chem. Int. Ed.* **2002**, *41*, 1198–1202; *Angew. Chem.* **2002**, *114*, 1246–1250.
[19] a) T. Moriguchi, Y. Ebizuka, I. Fujii, *ChemBioChem* **2008**, *9*, 1207–1212; b) T. Moriguchi, Y. Kezuha, T. Nonaka, Y. Ebizuka, I. Fujii, *J. Biol. Chem.* **2010**, *285*, 15637–15643.
[20] a) M. Tosin, D. Spittler, J. B. Spencer, *ChemBioChem* **2009**, *10*, 1714–1723; b) M. Tosin, L. Betancor, E. Stephens, W. M. A. Li, J. B. Spencer, P. F. Leadlay, *ChemBioChem* **2010**, *11*, 539–546.
[21] a) M. Tosin, Y. Demydchuk, J. S. Parascandolo, C. Blasco-Per, F. J. Leeper, P. F. Leadlay, *Chem. Commun.* **2011**, *47*, 3460–3462; b) M. Tosin, L. Smith, P. F. Leadlay, *Angew. Chem. Int. Ed.* **2011**, *50*, 11930–11933; *Angew. Chem.* **2011**, *123*, 12136–12139; c) E. Riva, I. Wilkening, S. Gazzola, W. M. A. Li, L. Smith, P. F. Leadlay, *Angew. Chem. Int. Ed.* **2014**, *53*, 11944–11949; *Angew. Chem.* **2014**, *126*, 12138–12143.

- [22] H. K. Potter, *PhD Thesis*, University of Cambridge, **2011**.
- [23] B. A. Pfeifer, S. J. Admiraal, H. Gramajo, D. E. Cane, C. Khosla, *Science* **2001**, *291*, 1790–1792.
- [24] I. Soehano, L. Yang, F. Ding, H. Sun, Z. J. Low, X. Liu, Z. X. Liang, *Org. Biomol. Chem.* **2014**, *12*, 8542–8549.
- [25] H. Kage, E. Riva, J. S. Parascandolo, M. F. Kreuzer, M. Tosin, M. Nett, *Org. Biomol. Chem.* **2015**, *13*, 11414–11417.
- [26] a) C. R. Huit-Roehl, E. A. Hill, M. M. Adams, A. L. Vagstad, J. W. Li, C. A. Townsend, *ACS Chem. Biol.* **2015**, *10*, 1443–1449; b) M. F. Kreuzer, H. Kage, J. Herrmann, J. Pauly, R. Hermenau, R. Müller, D. Hoffmeister, M. Nett, *Org. Biomol. Chem.* **2014**, *12*, 113–118.

Received: September 26, 2015

Revised: November 19, 2015

Published online: February 2, 2016