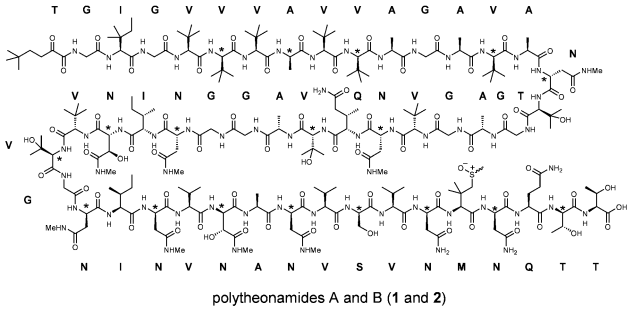
Peptide Biosynthesis

DOI: 10.1002/ange.201400478

Radical *S*-Adenosyl Methionine Epimerases: Regioselective Introduction of Diverse D-Amino Acid Patterns into Peptide Natural Products**

*Brandon I. Morinaka, Anna L. Vagstad, Maximilian J. Helf, Muriel Gugger, Carsten Kegler, Michael F. Freeman, Helge B. Bode, and Jörn Piel**

Abstract: *PoyD is a radical S-adenosyl methionine epimerase that introduces multiple* D*-configured amino acids at alternating positions into the highly complex marine peptides polytheonamide A and B. This novel post-translational modification contributes to the ability of the polytheonamides to form unimolecular minimalistic ion channels and its cytotoxic activity at picomolar levels. Using a genome mining approach we have identified additional PoyD homologues in various bacteria. Three enzymes were expressed in E. coli with their cognate as well as engineered peptide precursors and shown to introduce diverse* D*-amino acid patterns into all-*L *peptides. The data reveal a family of architecturally and functionally distinct enzymes that exhibit high regioselectivity, substrate promiscuity, and irreversible action and thus provide attractive opportunities for peptide engineering.*



Scheme 1. Structure of polytheonamides A and B. Asterisks show D-configured amino acids. Letters denote the precursor amino acids. The two congeners differ in the sulfoxide configuration.

The highly cytotoxic pore-forming polytheonamides A and B (**1** and **2**, Scheme 1) from the marine sponge *Theonella swinhoei* are the most extensively modified peptides known to date.[1] Among their many remarkable structural features are 18 D-configured residues positioned in almost perfect alternation with L-amino acids. We recently identified the polytheonamide (*poy*) biosynthetic pathway and showed that the peptides originate from an uncultivated symbiotic bacterium and are, despite their unusual structure, of ribosomal origin.[2] Among the large diversity of ribosomally synthesized and posttranslationally modified peptides (RiPPs),[3] polytheonamides are the first members of a new natural product family, termed proteusins.[2,4] Although computer analyses of sequenced bacterial genomes revealed a large number of further proteusin biosynthetic gene clusters in diverse bacteria,[4] the identity of the natural products remains obscure.

A unifying feature of proteusin precursor peptides is a large N-terminal leader region with similarity to nitrile hydratases (as for polytheonamides)[2] or Nif11 proteins.[4] In the *poy* pathway, maturation of the precursor is extraordinarily streamlined, presumably involving only 6 modifying enzymes that carry out 48 posttranslational modifications.[2] We previously characterized three of these enzymes, the most intriguing of which is PoyD, a novel epimerase of the radical *S*-adenosylmethionine (rSAM) superfamily.[5] Known rSAM enzymes abstract hydrogen from their substrates to generate radicals that can then undergo diverse follow-up reactions. Initial H abstraction is performed by a 5′-deoxyadenosyl radical, which is in turn generated from SAM by electron transfer from an Fe_4S_4 cluster. PoyD was shown to install numerous D-stereocenters on the polytheonamide backbone as a maturation step,[2] being the first known enzyme that introduces multiple D-amino acids into peptides.

Although PoyD offers considerable applicative potential, further studies on the enzyme have been challenging owing to the lack of an efficient analytical method that would permit

[*] Dr. B. I. Morinaka, Dr. A. L. Vagstad, M. J. Helf, Dr. M. F. Freeman, Prof. Dr. J. Piel
Institute of Microbiology
Eigenössische Technische Hochschule (ETH) Zurich
Vladimir-Prelog-Weg 1-5/10, 8093 Zurich (Switzerland)
E-mail: jpiel@ethz.ch
Dr. M. Gugger
Institut Pasteur, Collection des Cyanobactéries
Département de Microbiologie, 75015 Paris (France)
Dr. C. Kegler, Prof. Dr. H. B. Bode
Merck Stiftungsprofessur für Molekulare Biotechnologie
Fachbereich Biowissenschaften, Goethe Universität Frankfurt
Biozentrum, Campus Riedberg
60590 Frankfurt am Main (Germany)

[**] We thank A. Schneider for assistance with HPLC analysis and T. Coursin for cyanobacterial biomass and genomic DNA. J.P. acknowledges funding from the SNF (31003A_146992/1) and EU (BlueGenics and SYNPEPTIDE); work in the Bode lab was funded by the BMBF within the Biokatalyse 2021 program and an ERC starting grant (agreement no. 311477). B.I.M. is a fellowship recipient of the Alexander von Humboldt foundation. M.J.H. is receiving a fellowship from the German National Academic Foundation. M.F.F. is supported by a Human Frontier Science Program Long-term Fellowship. A.L.V. is the recipient of an ETH Zurich Postdoctoral Fellowship.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201400478.

localization of multiple epimerized residues. It was therefore also unknown whether radical epimerization is limited to products with regular DL-patterns or could be used to generate a broader range of diastereomers.

Herein we report the existence of a larger rSAM enzyme subfamily catalyzing peptide epimerization. rSAM epimerases are architecturally distinct and occur in various culturable and uncultivated bacteria. Characterization of three cyanobacterial representatives from cryptic natural product pathways shows that they irreversibly convert all-L-peptides to single products exhibiting diverse stereo patterns.

In BLAST searches using PoyD, we identified 21 closer homologues, all of which were localized in putative RiPP biosynthetic gene clusters of cyanobacteria, rhizobia, and other bacteria (Table 1; Supporting Information, Table S1, Figure S1). A feature shared by all these proteins but absent in other rSAM members is the presence of a circa 80 aa N-terminal region and a circa 200 aa C-terminal region that lack discernible sequence similarity to characterized enzymes (Supporting Information, Figure S2).

Table 1: Proteusin precursors from cyanobacterial strains analyzed in this study.

Bacterial source (precursor)	Core peptide sequence	Leader peptide type[a]
Oscillatoria sp. PCC 6506 (OspA)	GCWIAGSRGCGFVTRT	NH
A. variabilis ATCC 29413 (AvpA)	IFVSVFVSVFLAVPNQEIA	NH
Pleurocapsa sp. PCC 7319 (PlpA1,A2,A3)	WYFVTNDDEGAIVGSDSN	NH
	VDLSIFELLDEEPLFPIRPLYGLPI	Nif11
	AVAAMYGVVFPWDNEFPWPRWGG	Nif11

[a] NH = nitrile hydratase-like domains

For the present study, homologues from three cyanobacterial strains were selected, termed AvpD (*A. variabilis* ATCC 29413), OspD (*Oscillatoria* sp. PCC 6506),[6] and PlpD (*Pleurocapsa* sp. PCC 7319).[7] These originated from gene clusters encoding a nitrile hydratase-type precursor peptide as well as other presumed modifying enzymes (Figure 1). The *plp* cluster harbored two additional precursor genes of the *nif11* type. The core region of the precursors and the sets of maturation genes found in the three gene clusters were highly dissimilar, indicating significant structural diversity among cyanobacterial proteusins.

We initiated our investigation by cloning genes of the *Oscillatoria* sp. precursor (*ospA*, expression with N-terminal His_6-tag) and *poyD* homologue (*ospD*, untagged) into vectors for expression in *E. coli*. The *ospA* gene was amplified from the genome by PCR and ligated into either pET28b (precursor expression) or the multi-cloning site (MCS) I of pETDuet-1. For coexpression, *ospD* was cloned into the MCSII of pETDuet-1, and the corresponding plasmid was transformed into *E. coli* BL21(DE3). Unlike PoyA, for which all attempts to produce a soluble unmodified peptide had failed, His_6-OspA expressed well and gave soluble protein irrespective of OspD. After affinity chromatography, the core

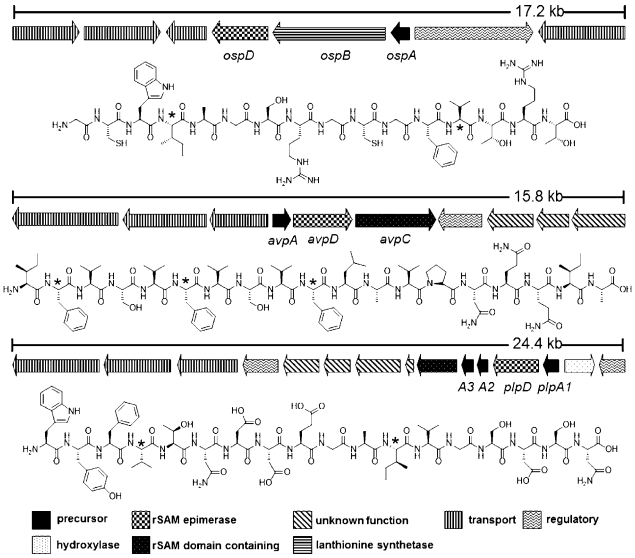


Figure 1. Gene clusters containing rSAM epimerases in cyanobacteria and core peptides before modification. BLAST-based predictions of encoded protein functions are denoted by arrow patterns. a) *osp* cluster, b) *avp* cluster, and c) *plp* cluster. Asterisks denote sites of epimerization by rSAM epimerases OspD, AvpD, and PlpD.

peptide was cleaved in vitro using Factor Xa or GluC protease sites engineered into the leader-core interface.

Reverse-phase HPLC analysis of core peptides cleaved from His_6-OspA-Fx and His_6-OspA-Fx + OspD expression experiments showed remarkably clean conversion into a new product in the coexpression (Figure 2 A,B) with a different retention time and the same molecular formula, indicating irreversible action of the modifying enzyme. To examine whether the newly formed product contained D-residues, it was purified by reverse-phase HPLC and subjected to Marfey-type (L-FDVA) analysis. Both D-Val and D-*allo*-Ile

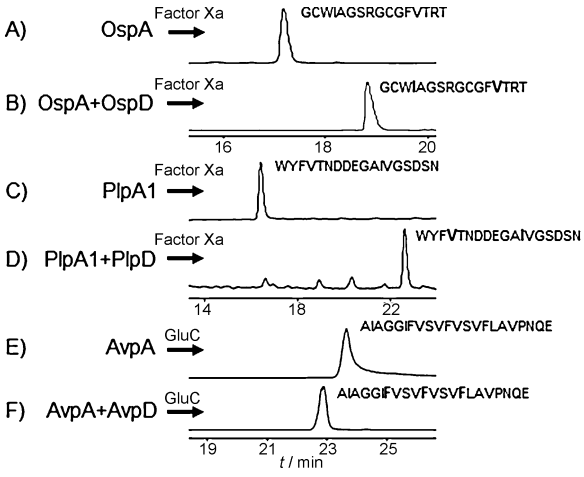


Figure 2. HPLC chromatograms for in vitro-cleaved His-tagged precursor peptides from in vivo (co)expressions. Major peaks correspond to fragments containing core regions. Epimerized residues are shown in bold. A) His_6-OspA-Fx ($\lambda = 280$ nm), B) His_6-OspA-Fx + OspD, C) His_6-PlpA1-Fx ($\lambda = 280$ nm), D) His_6-PlpA1-Fx + PlpD ($\lambda = 280$ nm) E) His_6-AvpA (EIC), F) His_6-AvpA + AvpD (EIC). Samples (E) and (F) were detected by EIC due to peak broadening.

were detected in the derivatized acid hydrolysate, thus showing that OspD indeed acts as an epimerase.

This result also established the regiochemistry of epimerization, as the core peptide contained only single Val and Ile residues. To test whether only these two residues had been epimerized and no other modifications had occurred, the epimerized core was compared by HPLC and NMR to an authentic sample of GCW(D-*allo*-Ile)AGSRGCGF(D-Val)TRT (Supporting Information, Figures S3 and S4), confirming their identity. The fact that only two remote residues were epimerized by OspD in contrast to the highly iterative polytheonamide homologue indicates a broad synthetic range of rSAM epimerases.

The remarkably clean conversion exhibited by OspD encouraged us to probe the substrate specificity for this enzyme. We constructed a series of alanine mutants (Table 2) within the core peptide of OspA to gain site-specific (I86A and V95A point mutations) as well as spatial information for insertions (85A86, 86A87, and 95A96) and deletions (87ADEL and 96TDEL). After coexpression, cleavage with Factor Xa, and HPLC analysis we observed conversion in all cases, most of which showed one major diasteromeric product (see the Supporting Information, Figures S5–S11 for HPLC traces). Marfey-type analysis of the products showed that OspD is not sensitive to the introduced changes at the Ile and Val sites. Furthermore, OspD is able to convert both Val and Ile whether a residue was removed or added before, within, or after epimerized residues.

Table 2: OspA alanine mutants co-expressed with OspD and corresponding products.

Name	Core sequence	Major product[a]
WT	GCWIAGSRGCGFVTRT	GCW**I**AGSRGCGF**V**TRT
85A86	GCWAIAGSRGCGFVTRT	GCWA**I**AGSRGCGF**V**TRT
I86 A	GCWAAGSRGCGFVTRT	GCW**A**AGSRGCGF**V**TRT[b]
86A87	GCWIAAGSRGCGFVTRT	GCW**I**AAGSRGCGF**V**TRT
87ADEL	GCWIGSRGCGFVTRT	GCW**I**GSRGCGF**V**TRT
V95A	GCWIAGSRGCGFATRT	GCW**I**AGSRGCGF**A**TRT[b]
95A96	GCWIAGSRGCGFVATRT	GCW**I**AGSRGCGF**V**ATRT
96TDEL	GCWIAGSRGCGFVRT	GCW**I**AGSRGCGF**V**RT

[a] Bold letters represent epimerized residues. L-Ile is converted to D-allo-Ile. [b] Position of D-alanine inferred from results of all other mutants.

An interesting feature of the *Pleurocapsa* sp. gene cluster is the presence of three putative precursor genes belonging to two different leader types. When the other two precursors (PlpA2 and PlpA3) belonging to the Nif11 family were also coexpressed with the epimerase PlpD, only a small amount of conversion was observed for PlpA3 (Supporting Information, Figure S14), but not for PlpA2 (Supporting Information, Figure S13). This result initially suggested preference for nitrile hydratase over Nif11 leader peptides. We further interrogated the role of leader type and core sequence in recognition by epimerases through constructing a series of chimeric precursors that contained individual leader sequences of PlpA1, PlpA2, and PlpA3 fused to each of the three core peptide sequences. After coexpression with PlpD, analysis revealed that the epimerase accepts remarkably different precursor peptides regarding overall sequence and type of epimerized residue (Table 3; Supporting Information, Figures S12–S14). This also includes precursors containing Nif11 leaders, many of which were epimerized to a significant degree. Surprisingly, a hybrid containing the Nif11 PlpA3 leader and the PlpA1 core was even converted quantitatively. In general, however, PlpD accepted peptides with the PlpA1 leader sequence to a greater extent. For instance, precursors containing the PlpA2 core were only epimerized with the PlpA1 leader sequence, and peptides containing the PlpA3 core showed full conversion with the PlpA1 leader and only partial conversion with the PlpA2 or PlpA3 leader. Of particular note is epimerization of valine, serine, and methionine for different constructs at similar core positions (aa 4-5), suggesting that epimerization might at least be partly guided by regiochemical features.

Table 3: PlpA leader/core peptide exchange coexpressions with PlpD.

Leader	Core	Major Epimerized Product[a]
PlpA1[b]	PlpA1	WYF**V**TNDDEGA**I**VGSDSN
PlpA2	PlpA1	WYFVTNDDEGA**I**VGSDSN
PlpA3	PlpA1	WYF**V**TNDDEGA**I**VGSDSN
PlpA1	PlpA2	VDL**S**IFELLDEEPLFPIRPLYGLPI
PlpA2[b]	PlpA2	NP
PlpA3	PlpA2	NP
PlpA1	PlpA3	AVAA**M**YGVVFPWDNEFPWPRWGG
PlpA2	PlpA3	AVAA**M**YGVVFPWDNEFPWPRWGG
PlpA3[b]	PlpA3	AVAA**M**YGVVFPWDNEFPWPRWGG

[a] Bold letters represent epimerized residues. L-Ile is converted to D-allo-Ile. [b] Wild-type. NP: no product observed.

Two additional epimerase gene candidates, *avpD* from *A. variabilis* and *plpD* from a *Pleurocapsa* sp. were also investigated. Analysis of the core peptides by reversed-phase HPLC again revealed virtually complete conversion when each precursor was expressed with the respective epimerase from its pathway (Figure 2C–F). Marfey-type analysis (Supporting Information) showed that one D-Val and D-*allo*-Ile residue each are present in the epimerized PlpA1 and three residues of D-Phe in epimerized AvpA. The position of D-Val in PlpA1 and of three D-Phe residues in AvpA was established by proteolytic cleavage (Factor Xa, PlpA1 or GluC, AvpA) of His_6-tagged precursor peptides, HPLC purification of peptide fragments containing residues of interest, and Marfey-type analysis (Supporting Information).

The assignment of D-configured amino acids usually requires additional degradative/analytical experiments when multiple residues of the same amino acid are present, but only a number of them are epimerized. The solution to this problem usually relies on partial acid hydrolysis or enzymatic digest followed by isolation of the peptide fragment of interest and amino acid analysis. To work towards a more efficient non-degradative MS-based method for identifying the position of epimerized residues in core peptides resulting from rSAM epimerases, as well as to obtain first insights into the mechanism of epimerization, we turned toward selective labeling strategies. The Osp and Plp heterologous system was

chosen, as soluble protein for the precursors could be obtained without the epimerase in contrast to the polytheonamide system. The strategy was based on the rationale that replacement of the α-H against another hydrogen during epimerization could be detected by a mass shift if the peptide carries α-deuterated residues. This method, which requires feeding of the labeled amino acids into the production strain, was recently developed by one of us (H.B.B.) for the analysis of nonribosomal peptides,[8] which are epimerized by a nonradical mechanism involving deprotonation. In that study, a transaminase-deficient *Photorhabdus* sp. mutant was used for NRPS expression to minimize competing deuterium loss. Using a new *E. coli* mutant defective in five transaminases (the DH10B-derived *E. coli* Δ*penta*)[9] as a host, we tested this method for ribosomally encoded peptides by individually feeding fully deuterated L-Val-d_8 and L-Ile-d_{10} into liquid cultures of OspA + OspD (L-Val-d_8 or L-Ile-d_{10}, individually) and PlpA1 + PlpD (L-Val-d_8 or L-Ile-d_{10}, simultaneously) coexpressions in synthetic medium. LCMS analysis of core peptides showed epimerized residues as detected by shift in retention time. MS and MS^2 analyses confirmed that for each epimerized amino acid in the labeled peptides, selective loss of a single deuterium occurred only after coexpression with the epimerase (Supporting Information, Figures S15–S20). These data suggest that the labeling method provides a convenient platform to identify specific residues epimerized by rSAM enzymes.

The results also provided initial insights into mechanistic aspects of epimerization. Based on general properties of rSAM enzymes,[5] epimerization was suspected to be initiated by α-H abstraction to generate a resonance-stabilized radical. In principle, several alternatives for the source of the backside hydrogen can be envisioned to effect epimerization: 1) The same α-H is removed and transferred to the opposite face; 2) after radical generation at one amino acid, a hydrogen is transferred from another residue within the same peptide; 3) the α-H of the two epimerized residues are switched; or 4) both α-H are replaced by H external from the core. As the label is lost after epimerization, the results disfavor mechanisms 1–3. Unless hydrogens originating from α-positions are transferred to exchangeable sites during epimerization, an interesting consequence would be that the 5′-deoxyadenosyl-radical is not recycled by rSAM epimerases. This mechanism, which is known from several other rSAM homologues,[10] is noteworthy, as it would imply the consumption of 18 SAM equivalents per precursor peptide molecule in polytheonamide biosynthesis.

In conclusion, this study reveals radical epimerization as a more general strategy used by nature to create stereochemically diverse peptides and reports an efficient method for product characterization. Several features of the cryptic proteusins are of interest for natural product discovery and enzymology. The diversity of core regions and modifying enzymes suggest considerable structural heterogeneity beyond epimerization. For example, the *osp* cluster also encodes a LanM homologue predicted to introduce lanthionine bridges,[11] while the *avp* cluster encodes a putative rSAM methyltransferase (AvpC) that likely methylates one or more unactivated carbon centers.[12] The identification of these natural products and their activities are under current investigation.

The presence of D-amino acids in many bioactive natural peptides as well as their use in medicinal synthetic chemistry demonstrate their importance to effect or enhance bioactivity and stability.[13] To our knowledge, proteusin rSAM epimerases are the first enzymes with the capacity to install multiple D-amino acids in genetically encoded peptide chains. The representatives examined so far are readily expressed, act irreversibly, are highly regioselective, accept widely divergent substrates, and exhibit diverse product profiles. These properties will serve as a useful basis to study mechanistic details of these fascinating enzymes and their potential in peptide engineering.

Received: January 16, 2014
Revised: April 30, 2014
Published online: June 18, 2014

Keywords: biosynthesis · D-amino acids · enzymes · natural products · peptides

[1] T. Hamada, S. Matsunaga, G. Yano, N. Fusetani, *J. Am. Chem. Soc.* **2005**, *127*, 110.
[2] M. F. Freeman, C. Gurgui, M. J. Helf, B. I. Morinaka, A. R. Uria, N. J. Oldham, H. G. Sahl, S. Matsunaga, J. Piel, *Science* **2012**, *338*, 387.
[3] P. G. Arnison, M. J. Bibb, G. Bierbaum, A. A. Bowers, T. S. Bugni, G. Bulaj, J. A. Camarero, D. J. Campopiano, G. L. Challis, J. Clardy, P. D. Cotter, D. J. Craik, M. Dawson, E. Dittmann, S. Donadio, P. C. Dorrestein, K.-D. Entian, M. A. Fischbach, J. S. Garavelli, U. Göransson, C. W. Gruber, D. H. Haft, T. K. Hemscheidt, C. Hertweck, C. Hill, A. R. Horswill, M. Jaspars, W. L. Kelly, J. P. Klinman, O. P. Kuipers, A. J. Link, W. Liu, M. A. Marahiel, D. A. Mitchell, G. N. Moll, B. S. Moore, R. Müller, S. K. Nair, I. F. Nes, G. E. Norris, B. M. Olivera, H. Onaka, M. L. Patchett, J. Piel, M. J. T. Reaney, S. Rebuffat, R. P. Ross, H.-G. Sahl, E. W. Schmidt, M. E. Selsted, K. Severinov, B. Shen, K. Sivonen, L. Smith, T. Stein, R. D. Süssmuth, J. R. Tagg, G.-L. Tang, A. W. Truman, J. C. Vederas, C. T. Walsh, J. D. Walton, S. C. Wenzel, J. M. Willey, W. A. van der Donk, *Nat. Prod. Rep.* **2013**, *30*, 108.
[4] D. H. Haft, M. K. Basu, D. A. Mitchell, *BMC Biol.* **2010**, *8*, 70.
[5] a) P. A. Frey, A. D. Hegeman, F. J. Ruzicka, *Crit. Rev. Biochem. Mol. Biol.* **2008**, *43*, 63; b) J. L. Vey, C. L. Drennan, *Chem. Rev.* **2011**, *111*, 2487.
[6] A. Mejean, R. Mazmouz, S. Mann, A. Calteau, C. Medigue, O. Ploux, *J. Bacteriol.* **2010**, *192*, 5264.
[7] P. M. Shih, D. Y. Wu, A. Latifi, S. D. Axen, D. P. Fewer, E. Talla, A. Calteau, F. Cai, N. T. de Marsac, R. Rippka, M. Herdman, K. Sivonen, T. Coursin, T. Laurent, L. Goodwin, M. Nolan, K. W. Davenport, C. S. Han, E. M. Rubin, J. A. Eisen, T. Woyke, M. Gugger, C. A. Kerfeld, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 1053.
[8] H. B. Bode, D. Reimer, S. W. Fuchs, F. Kirchner, C. Dauth, C. Kegler, W. Lorenzen, A. O. Brachmann, P. Grun, *Chem. Eur. J.* **2012**, *18*, 2342.
[9] C. Kegler, F. I. Nollmann, T. Ahrendt, F. Fleischhacker, E. Bode, H. B. Bode, **2014**, *6*, 826 – 828.
[10] E. N. G. Marsh, D. P. Patterson, L. Li, *ChemBioChem* **2010**, *11*, 604.

[11] A. L. McClerren, L. E. Cooper, C. Quan, P. M. Thomas, N. L. Kelleher, W. A. van der Donk, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 17243.
[12] a) J. P. Gomez-Escribano, L. J. Song, M. J. Bibb, G. L. Challis, *Chem. Sci.* **2012**, *3*, 3522; b) W. J. K. Crone, F. J. Leeper, A. W. Truman, *Chem. Sci.* **2012**, *3*, 3516; c) L. J. Huo, S. Rachid, M. Stadler, S. C. Wenzel, R. Müller, *Chem. Biol.* **2012**, *19*, 1278; d) Y. P. Hou, M. D. B. Tianero, J. C. Kwan, T. P. Wyche, C. R. Michel, G. A. Ellis, E. Vazquez-Rivera, D. R. Braun, W. E. Rose, E. W. Schmidt, T. S. Bugni, *Org. Lett.* **2012**, *14*, 5050.
[13] a) H. L. Condurso, S. D. Bruner, *Nat. Prod. Rep.* **2012**, *29*, 1099; b) R. Tugyi, K. Uray, D. Ivan, E. Fellinger, A. Perkins, F. Hudecz, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 413.

