Natural Products

DOI: 10.1002/anie.200901808

Thiopeptide Antibiotic Biosynthesis**

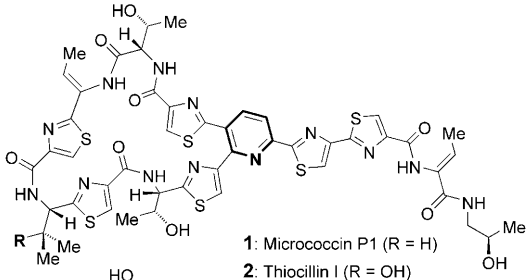
Hans-Dieter Arndt, Sebastian Schoof, and Jin-Yong Lu*

antibiotics · biosynthesis · natural products · ribosomal peptides · thiopeptides

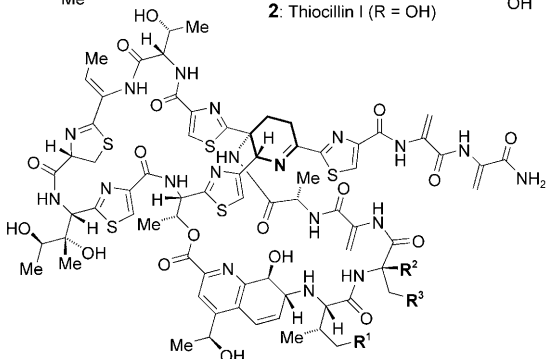
Dedicated to Professor Heinz G. Floss

The family of thiopeptide antibiotics is a large group of highly modified macrocyclic peptides produced by bacteria and comprises more than 80 members.[1] Micrococcin (**1**) was the first thiopeptide to be isolated (in 1948), and the prototypical and easily produced thiostrepton (**3**) has quickly become the most studied compound of this group.[1,2] All thiopeptides are united by their intriguing structures, which always build on a common central pyridine-derived heterocycle. This core forms a macrocyclic array with numerous thiazole and oxazole rings and other specific residues, such as dehydroamino acids. Many thiopeptides show impressive bioactivity,[1] most prominently a very high potency against gram-positive bacteria, including multidrug-resistant *Staphylococcus aureus* strains (MRSA). Their mode of action is based on inhibition of bacterial protein translation by blocking the ribosomal GTPase-associated center[1,3] or by inhibiting the translation factor EF-Tu.[4] Neither of these two cellular targets has yet led to drugs suitable for human therapy, which is currently revitalizing interest in thiopeptides for antibacterial applications.[5]

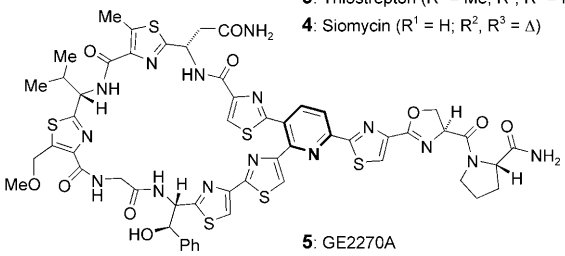
Chemical synthesis has been richly inspired by the intricate molecular architectures of thiopeptides,[1,6] but the biosynthesis of this fascinating compound class remained obscure. Now four studies have been reported almost in parallel,[7–10] substantiating a unifying rationale: The structural complexity of all thiopeptides probably arises from unforeseen posttranslational modifications of genetically encoded and ribosomally translated peptides! These remarkable findings push the limits of current biosynthesis paradigms and will allow us to decipher and utilize the biosynthetic machinery behind this important class of peptide natural products.



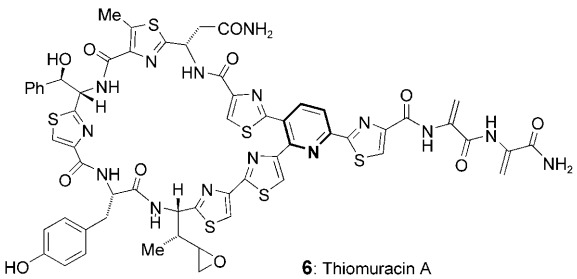
1: Micrococcin P1 (R = H)
2: Thiocillin I (R = OH)



3: Thiostrepton (R^1 = Me; R^2, R^3 = H)
4: Siomycin (R^1 = H; R^2, R^3 = Δ)



5: GE2270A



6: Thiomuracin A

Two main pathways are known for the biosynthesis of peptide natural products.[11] A precursor peptide may be synthesized by ribosomes through translation of genetically encoded mRNA with subsequent posttranslational modification of the linear chain.[11a] Alternatively, nonribosomal

[*] Dr. H.-D. Arndt, S. Schoof, J.-Y. Lu
Fakultät Chemie, Technische Universität Dortmund
Otto-Hahn-Strasse 6, 44221 Dortmund (Germany)
and
Max-Planck-Institut für molekulare Physiologie
Abteilung Chemische Biologie
Otto-Hahn-Strasse 11, 44227 Dortmund (Germany)
Fax: (+49) 231-133-2498
E-mail: hans-dieter.arndt@mpi-dortmund.mpg.de

[**] Our work was supported by the Deutsche Forschungsgemeinschaft (Emmy Noether young investigator award to H.-D. A.) and the Fonds der Chemischen Industrie. S.S. and J.-Y.L. are members of the International Max Planck Research School in Chemical Biology (IMPRS-CB). We thank Prof. Dr. Rolf Müller (Universität Saarbrücken) and Dr. Mark Brönstrup (Sanofi-Aventis, Frankfurt am Main) for discussions.

peptide synthesis is carried out by enzyme complexes (NRPSs: nonribosomal peptide synthetases) that assemble amino acid precursors in a conveyor-belt-like fashion.[11b] Typical products of NRPSs are vastly modified molecules with a high content of nonproteinogenic amino acids, whereas ribosomally synthesized peptides are expected to feature a much lower degree of modification. But with the new results on thiopeptides, these general rules need to be revised.

A visionary hypothesis concerning the key transformation for thiopeptide biosynthesis was drafted by Bycroft and Gowland in 1978, who suggested that the central pyridine ring "*may be derived from the interaction of two dehydroalanine units in a single peptide chain*".[12] This proposed transformation remained speculative for a long time, and it could occur subsequent to either ribosomal or nonribosomal assembly. Detailed investigations of the biosynthesis of thiopeptides were then conducted by Floss and co-workers, who could trace the origin of all structural elements in thiostrepton (**3**) and nosiheptide to standard amino acids from primary metabolism using isotope labeling experiments.[13] However, the enzymatic machinery responsible for the assembly remained elusive, and the identification of typical NRPS gene clusters turned out to be difficult.[13e,14]

Important evidence that complex heterocyclic peptides can be ribosomally synthesized then emerged from the biosynthetic pathway of the patellamides and cyanobactins.[15] Undoubtedly, advancements in whole-genome sequencing and the growth of bioinformatics data facilitated the identification of crucial linear precursor peptides and their processing factors. It could be shown that these processing factors share homology to enzymes implicated in the heterocyclization of Ser/Thr/Cys residues to oxazoles and thiazoles[16a] and to macrolactam-forming enzymes[16b] identified for other natural products.

With this knowledge base and up-to-date sequencing techniques at hand, four groups succeeded at the same time in localizing the biosynthesis genes for the thiocillins (e.g. **2**),[7] thiostrepton and siomycin (**3** and **4**; Δ = double bond),[8,9] GE2270A (**5**),[10] as well as the newly identified thiomuracin (**6**).[10] A direct entry was provided by the fully sequenced genome of *Bacillus cereus*.[7,8] For the producers *Streptomyces lividans*[8,9] and *Nonomurea* sp.[10] genomic libraries were used in combination with partial sequencing. In every case, analogous cluster architectures emerged (Scheme 1 A) that seem to be retained over species boundaries. Sequence homology indicated the structural genes that encode linear precursor peptides and are surrounded by modifying genes. Four of the putative enzymes encoded by these genes were identified as being responsible for the characteristic thiopeptide framework: a cyclodehydratase homologous to PatD of the patellamide biosynthesis and a dehydrogenase that resembles the microcin B17 biosynthesis enzyme McbC; furthermore, two other enzymes with homology to the lantibiotic dehydratase LanB seem to be responsible for dehydroalanine and -butyrine formation.

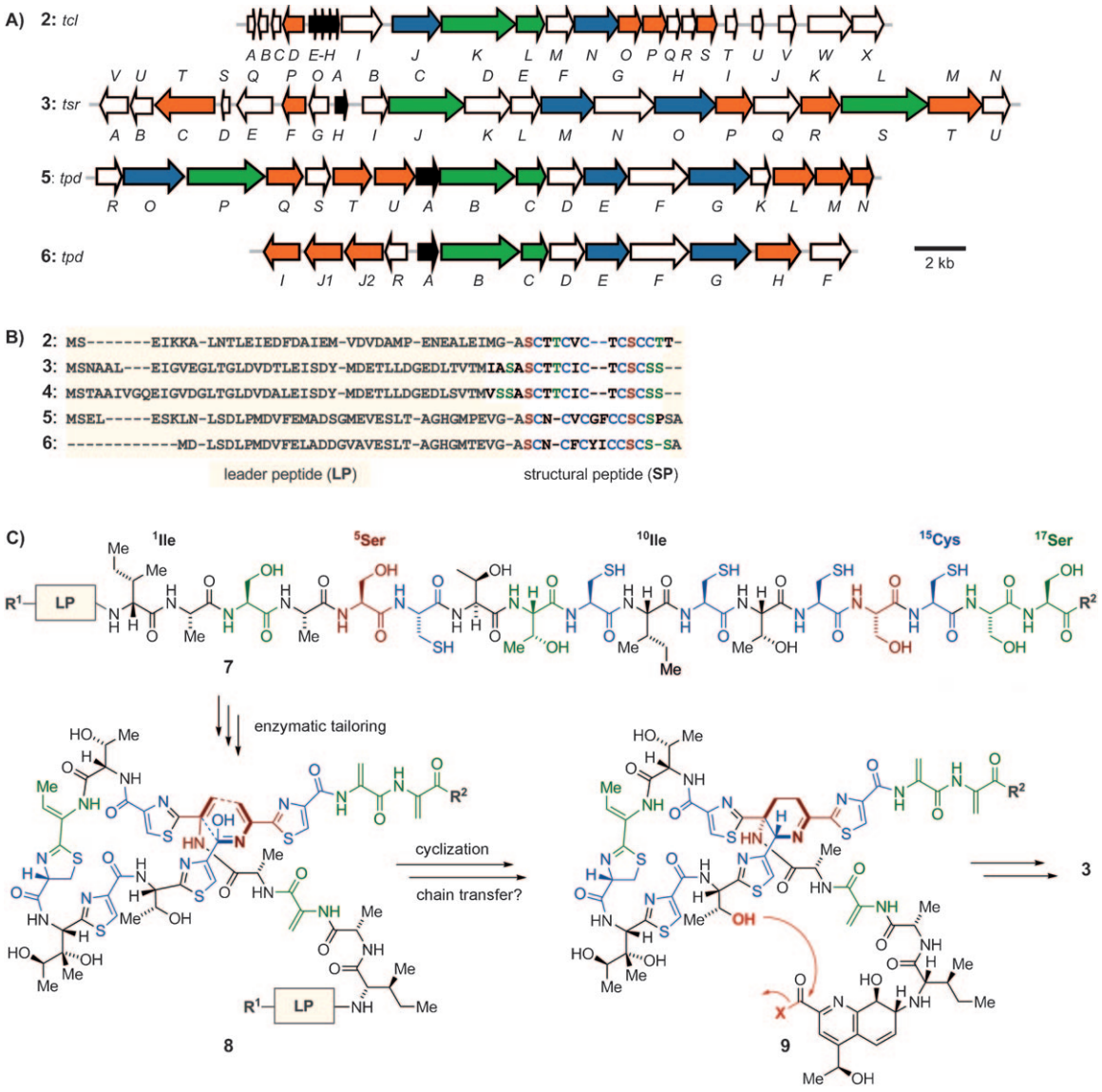
All precursors contain the full peptide sequence of the product and an N-terminal leader peptide (LP) that most likely directs the ensuing modifications before the final product or an advanced intermediate is liberated.[17] Comparison between these structural genes reveals high similarity (Scheme 1 B); especially the highly conserved Ser/Cys placement and the consistent distance between the two key Ser residues is striking. The number of residues incorporated into the final product varies, but it is assumed that for thiopeptides with pyridine nuclei (e.g. **1**, **2**, **5**, **6**) all amino acids N-terminal to the first pyridine-forming Ser residue are lost by elimination upon aromatization of the core.

The anticipated biosynthesis of thiostrepton (**3**) is based on the ribosomal synthesis of the linear peptide precursor **7** encoded by the structural gene *tsrA/tsrH* (Scheme 1 C). All Ser/Thr residues of peptide **7** seem to become dehydrated to dehydroalanines or -butyrines. Likewise, all cysteine residues are converted into thiazolines by cyclodehydration and then mostly oxidized to thiazoles (→**8**). The remarkable central six-membered heterocycle is possibly formed through an intramolecular reaction between two key dehydroalanines and a neighboring carboxyl group (→**9**). This formal hetero-Diels–Alder cycloaddition reaction could be promoted by enzymatically activating a back-folded peptide chain and may either be concerted or an asynchronous sequence of stepwise 1,2/1,4-additions.[9] Nonetheless, these data fully confirm the Bycroft–Gowland hypothesis for thiopeptide biosynthesis.[12]

That such a transformation may be feasible was also suggested by chemical synthesis (Scheme 2). Nicolaou et al. reported the formation of *endo*-dehydropiperidine **13** via **12** by dimerization of 2-azadiene **11** generated in situ from **10.**[18] This method was applied to the total synthesis of thiostrepton (**3**)[19] and GE2270A (**5**).[20] Moody et al. realized a hetero-Diels–Alder reaction of the thiopeptide dienophile **14** with the 2-azadiene **15** to furnish pyridine **16** in one step, which may be close to the anticipated biological process.[21] Whether the putative biosynthesis enzymes really promote a similar transformation of two crossing peptide strands remains to be elucidated.

In addition to these framework-forming enzymes, several other prospective tailoring and decorating enzymes were identified by sequence similarity, for instance SAM-dependent methyltransferases (TclO: thiocillin) or P450 monooxygenases (TpdJ: thiomuracin). However, assignment of the protein factors responsible for the formation of the central six-membered heterocycle will require additional research. Further challenges remain, for example the incorporation of thiostrepton's second macrocycle (Scheme 1 C). Apparently, the precursor amino acid Trp for the quinaldic acid fragment is not gene-encoded, which indicates that the intermediate product of the initial ribosomal peptide biosynthesis must be completed by further manipulations for the bicyclic thiopeptides thiostrepton and siomycin as well as nosiheptide and nocathiacin.[1,22] Some evidence makes the incorporation of the unique quinaldic acid into thiostrepton (**3**) by coenzyme A activation[8] or NRPS-type adenylation domains[9,13e] likely (→**9**). Furthermore, cloning and overexpression of the *tsrV* gene by Kelly et al. gave a protein which transformed 2-methyltryptophan (**17**) to α-keto acid **18** (Scheme 3),[9] suggesting its involvement in the formation of quinaldic acid **20** via the putative diketone **19**.[13d]

The discovery and analysis of thiopeptide gene clusters significantly expands the range of known posttranslational

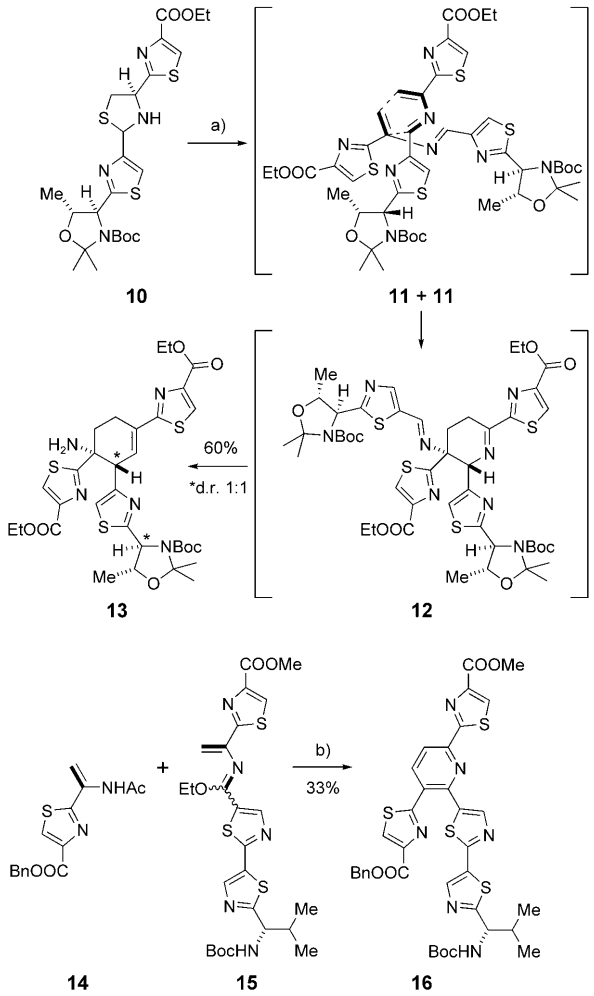


Scheme 1. A) Architectures of the thiopeptide biosynthesis gene clusters of compounds **2**,[7] **3**,[8,9] **5**,[10] and **6**.[10] Note the different numbering schemes and slightly different assignments for the identical *tsr* genes suggested by Liu et al. (bottom)[8] and Kelly et al. (top).[9] Black: structural gene; green: dehydratase; blue: cyclodehydratase or dehydrogenase; orange: tailoring enzymes (monoxygenase, methyl transferase, protease, deaminase, amidotransferase); colorless: other/unknown open reading frame; *tcl* = thiocillin genes, *tsr* = thiostrepton genes, *tpd* = thiopeptide genes; bar: 2 kb. B) Partial alignment of thiopeptide structural gene sequences (ClustalW 2.0) for compounds **2**–**6**. Leader peptides are highlighted in gray, structural peptides are color-coded (see below). C) Emerging picture of thiostrepton structural peptide maturation (simplified). Green: dehydratase-mediated dehydroalanine/-butyrine formation; blue: cyclodehydratase-initiated heterocycle formation; black: structurally unmodified/peripherally decorated residues; red: dehydroamino acids involved in the formation of the aza-heterocyclic nucleus. R^1 = flanking sequence, R^2 = OH, X = leaving group, potentially CoA or adenylate.

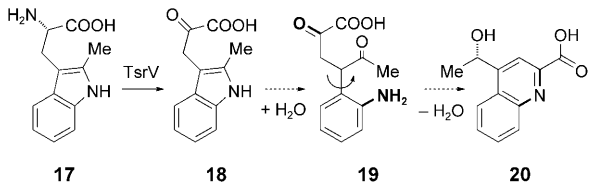
modifications for ribosomal peptides. Moreover, numerous thiopeptides have been described to contain further modifications, such as oxidations of alkyl side chains (berninamycin, thiostrepton), O- and S-alkylations, oxidative transannular bridgings (nocathiacin),[22] and glycosylations (nocathiacin, philipimycin).[23] Unbiased genome searches showed that many more biosynthesis gene clusters of thiopeptide-like metabolites can be identified.[7,8,10] It thus appears that yet unknown thiopeptides still await discovery.

Various steps in the biosynthesis of thiopeptides as well as the order of the modification reactions remain to be more clearly defined, but the principal findings[7–10] that such complex molecular architectures are tailored from ribosomal peptides redefines the paradigms of natural-product biosynthesis. These findings hold high promise for biotechnology and combinatorial biosynthesis of diverse compound collections. Thiopeptides will hence surely continue to stimulate exciting discoveries in chemical biology and natural-products research.

Received: April 3, 2009
Published online: June 17, 2009



Scheme 2. Selected biomimetic chemical syntheses of thiopeptide aza-heterocyclic cores. a) Ag_2CO_3, DBU, pyridine, $BnNH_2$, −12→25 °C; b) toluene, microwave irradiation, 120 °C, 12 h. DBU = 1,8-diazabicyclo-[5.4.0]undec-7-ene, Bn = benzyl, Boc = *tert*-butoxycarbonyl.



Scheme 3. Involvement of TsrV in the biosynthesis of **20** from Trp.

[1] a) M. C. Bagley, J. W. Dale, E. A. Merritt, X. Xiong, *Chem. Rev.* **2005**, *105*, 685–714; b) R. A. Hughes, C. J. Moody, *Angew. Chem.* **2007**, *119*, 8076–8101; *Angew. Chem. Int. Ed.* **2007**, *46*, 7930–7954.

[2] Landmark structure determination of thiostrepton: B. Anderson, D. Crowfoot-Hodgkin, M. A. Viswamitra, *Nature* **1970**, *225*, 233–235.

[3] a) H. R. A. Jonker, S. Ilin, S. K. Grimm, J. Wöhnert, H. Schwalbe, *Nucleic Acids Res.* **2007**, *35*, 441–454; b) J. M. Harms, D. N. Wilson, F. Schlünzen, S. R. Connell, T. Stachelhaus, Z. Zaborowska, C. M. T. Spahn, P. Fucini, *Mol. Cell* **2008**, *30*, 26–38; c) S. Baumann, S. Schoof, S. D. Harkal, H.-D. Arndt, *J. Am. Chem. Soc.* **2008**, *130*, 5664–5666; d) S. Schoof, S. Baumann, B. Ellinger, H.-D. Arndt, *ChemBioChem* **2009**, *10*, 242–245.

[4] a) S. E. Heffron, F. Jurnak, *Biochemistry* **2000**, *39*, 37–45; b) A. Parmeggiani, I. M. Krab, S. Okamura, R. C. Nielsen, J. Nyborg, P. Nissen, *Biochemistry* **2006**, *45*, 6846–6857.

[5] A recent review: F. von Nussbaum, M. Brands, B. Hinzen, S. Weigand, D. Häbich, *Angew. Chem.* **2006**, *118*, 5194–5254; *Angew. Chem. Int. Ed.* **2006**, *45*, 5072–5129.

[6] Review: K. C. Nicolaou, J. S. Chen, D. J. Edmonds, A. A. Estrada, *Angew. Chem.* **2009**, *121*, 670–732; *Angew. Chem. Int. Ed.* **2009**, *48*, 660–719.

[7] L. C. Wieland Brown, M. G. Acker, J. Clardy, C. T. Walsh, M. A. Fischbach, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 2549–2553.

[8] R. Liao, L. Duan, C. Lei, H. Pan, Y. Ding, Q. Zhang, D. Chen, B. Shen, Y. Yu, W. Liu, *Chem. Biol.* **2009**, *16*, 141–147.

[9] W. L. Kelly, L. Pan, C. Li, *J. Am. Chem. Soc.* **2009**, *131*, 4327–4334.

[10] R. P. Morris, J. A. Leeds, H.-U. Nägeli, L. Oberer, K. Memmert, E. Weber, M. J. LaMarche, C. N. Parker, N. Burrer, S. Esterow, A. E. Hein, E. K. Schmitt, P. Krastel, *J. Am. Chem. Soc.* **2009**, *131*, 5946–5955.

[11] Reviews: a) E. M. Nolan, C. T. Walsh, *ChemBioChem* **2009**, *10*, 34–53; b) S. A. Sieber, M. A. Marahiel, *Chem. Rev.* **2005**, *105*, 715–738.

[12] B. W. Bycroft, M. S. Gowland, *J. Chem. Soc. Chem. Commun.* **1978**, 256–258.

[13] a) P. Zhou, D. O'Hagan, U. Mocek, Z. Zeng, L.-D. Yuen, T. Frenzel, C. J. Unkefer, J. M. Beale, H. G. Floss, *J. Am. Chem. Soc.* **1989**, *111*, 7274–7276; b) T. Frenzel, P. Zhou, H. G. Floss, *Arch. Biochem. Biophys.* **1990**, *278*, 35–40; c) U. Mocek, A. R. Knaggs, R. Tsuchiya, T. Nguyen, J. M. Beale, H. G. Floss, *J. Am. Chem. Soc.* **1993**, *115*, 7557–7568; d) U. Mocek, Z. Zeng, D. O'Hagan, P. Zhou, L.-D. G. Fan, J. M. Beale, H. G. Floss, *J. Am. Chem. Soc.* **1993**, *115*, 7992–8001; e) N. D. Priestley, T. M. Smith, P. R. Shipley, H. G. Floss, *Bioorg. Med. Chem.* **1996**, *4*, 1135–1147.

[14] M. C. Carnio, T. Stachelhaus, K. P. Francis, S. Scherer, *Eur. J. Biochem.* **2001**, *268*, 6390–6400.

[15] a) E. W. Schmidt, J. T. Nelson, D. A. Rasko, S. Sudek, J. A. Eisen, M. G. Haygood, J. Ravel, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 7315–7320; b) P. F. Long, W. C. Dunlap, C. N. Battershill, M. Jaspars, *ChemBioChem* **2005**, *6*, 1760–1765.

[16] a) Y.-M. Li, J. C. Milne, L. L. Madison, R. Kolter, C. T. Walsh, *Science* **1996**, *274*, 1188–1193; b) J. O. Solibati, M. Ciaccio, R. N. Farías, J. E. González-Pastor, F. Moreno, R. A. Salomón, *J. Bacteriol.* **1999**, *181*, 2659–2662.

[17] M. R. Levengood, C. C. Kerwood, C. Chatterjee, W. A. van der Donk, *ChemBioChem* **2009**, *10*, 911–919.

[18] K. C. Nicolaou, M. Nevalainen, B. S. Safina, M. Zak, S. Bulat, *Angew. Chem.* **2002**, *114*, 2021–2025; *Angew. Chem. Int. Ed.* **2002**, *41*, 1941–1945.

[19] a) K. C. Nicolaou, B. S. Safina, M. Zak, A. A. Estrada, S. H. Lee, *Angew. Chem.* **2004**, *116*, 5197–5202; *Angew. Chem. Int. Ed.* **2004**, *43*, 5087–5092; b) K. C. Nicolaou, M. Zak, B. S. Safina, S. H. Lee, A. A. Estrada, *Angew. Chem.* **2004**, *116*, 5202–5207; *Angew. Chem. Int. Ed.* **2004**, *43*, 5092–5097.

[20] K. C. Nicolaou, B. Zou, D. H. Dethe, D. B. Li, D. Y.-K. Chen, *Angew. Chem.* **2006**, *118*, 7950–7956; *Angew. Chem. Int. Ed.* **2006**, *45*, 7786–7792.

[21] C. J. Moody, R. A. Hughes, S. P. Thompson, L. Alcaraz, *Chem. Commun.* **2002**, 1760–1761.

[22] K. L. Constantine, L. Mueller, S. Huang, S. Abid, K. S. Lam, W. Li, J. E. Leet, *J. Am. Chem. Soc.* **2002**, *124*, 7284–7285.

[23] C. Zhang, J. Occi, P. Masurekar, J. F. Barrett, D. L. Zink, S. Smith, R. Onishi, S. Ha, O. Salazar, O. Genilloud, A. Basilio, F. Vicente, C. Gill, E. J. Hickey, K. Dorso, M. Motyl, S. B. Singh, *J. Am. Chem. Soc.* **2008**, *130*, 12102–12110.

