Natural Products

DOI: 10.1002/anie.201200327

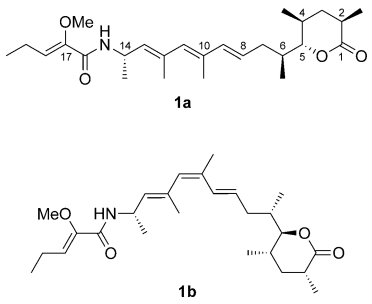
Pellasoren: Structure Elucidation, Biosynthesis, and Total Synthesis of a Cytotoxic Secondary Metabolite from *Sorangium cellulosum***

Christine Jahns, Thomas Hoffmann, Stefan Müller, Klaus Gerth, Peter Washausen, Gerhard Höfle, Hans Reichenbach, Markus Kalesse, and Rolf Müller**

Myxobacteria are efficient producers of numerous secondary metabolites, and the genus *Sorangium* is frequently described as an proficient source for new, biologically active natural products.[1–3] We report here the discovery and complete structure elucidation of pellasoren from the myxobacterium *Sorangium cellulosum*. Identification of the corresponding *pel* gene cluster from *S. cellulosum* So ce38 allowed us to establish a model for pellasoren biosynthesis, providing evidence for an unusual route to glycolate extender unit generation. Moreover, we present the first total synthesis of pellasoren and thereby confirm the absolute configuration of this natural product.

Pellasoren (**1**; Scheme 1) was initially isolated from *S. cellulosum* So ce35 in the course of an activity-guided discovery program.[4] Additionally, we recently identified **1** in relatively high amounts in extracts from the related strain *S. cellulosum* So ce38 by using LC-MS analysis. Here we determine its cytotoxicity against HCT-116 human colon cancer cells at a concentration of 155 nM (IC50). Full structural elucidation by NMR and ESI-MS analysis was performed and confirmed the identity of pellasoren from both sources (Figures S1–S3 and Table S1 in the Supporting Information).

The pellasoren scaffold features an unusual enol ether moiety, also known from a small number of other natural products, which was corroborated by specific HMBC correlations between a methoxy signal and sp^2-hybridized carbon atoms (Figure S1 in the Supporting Information).[5–7] The lactone moiety was identified through HMBC correlation between C1 and C5 and characteristic shifts for H-5 and C5 of $\delta = 4.02$ and 90.3 ppm, respectively. The position of the amide bond was assigned on the basis of indicative fragments in CID spectra. Efforts were made to establish the molecule's relative configuration by using NOE spectroscopy: ROESY interactions together with molecular modeling suggest an *anti* configuration of the substituents at C4 and C5, as well as a *syn* configuration of the methyl groups at C2 and C4 (Figure S2 in the Supporting Information). The stereocenter at C14 maintains the configuration derived from the incorporation of an L-alanine building block during biosynthesis. Stereochemical assignments that could initially not be validated by NOE analysis, such as the configuration at C6, were later established following total synthesis. Additionally, two isomeric pellasorens differing in the configuration of the C10–C11 double bond were isolated from *S. cellulosum* extracts. Pellasoren A (**1a**) represents the all-*E* configuration while pellasoren B (**1b**) has a *Z*-configured double bond at C10–C11 which can be rationalized by photochemical isomerization.



Scheme 1. Structure of pellasoren A (**1a**) and B (**1b**).

Following the structural elucidation of pellasoren, we set out to identify the underlying biosynthetic machinery using fragmentary whole-genome sequence information for strain *S. cellulosum* So ce38. Assuming that pellasoren is most likely the product of a hybrid PKS/NRPS biosynthetic pathway,[8] the full complement of putative PKR/NRPS-related domains encoded in the So ce38 genome was annotated using bioinformatic tools.[9] Using the presumed incorporation of alanine into pellasoren as a guide, we identified a genomic region roughly 57 500 bp in size, containing seven characteristic PKS modules organized as an apparent operon which also encodes one adenylation (A) domain exhibiting the

[*] C. Jahns,[+] Prof. Dr. M. Kalesse
Institut für Organische Chemie, Leibniz Universität Hannover
Schneiderberg 1B, 30167 Hannover
and
Helmholtz-Zentrum für Infektionsforschung (HZI)
Inhoffenstrasse 7, Braunschweig (Germany)
E-mail: markus.kalesse@oci.uni-hannover.de

T. Hoffmann,[+] S. Müller, Prof. Dr. R. Müller
Helmholtz-Institut für Pharmazeutische Forschung Saarland (HIPS), Helmholtz-Zentrum für Infektionsforschung und Pharmazeutische Biotechnologie, Universität des Saarlandes
Campus, Gebäude C2.3, 66123 Saarbrücken (Germany)
E-mail: rom@helmholtz-hzi.de

Dr. K. Gerth, Dr. P. Washausen, Prof. Dr. G. Höfle, Prof. Dr. H. Reichenbach
Helmholtz-Zentrum für Infektionsforschung (HZI)
Inhoffenstrasse 7, Braunschweig (Germany)

[+] These authors contributed equally to this work.

[**] We thank Daniel Krug for proofreading the manuscript, Jennifer Herrmann for performing the bioactivity assays, and Susanne Schneiker-Bekel for genome sequencing. Research in R.M's laboratory was funded by the Bundesministerium für Bildung und Forschung and the Deutsche Forschungsgemeinschaft.

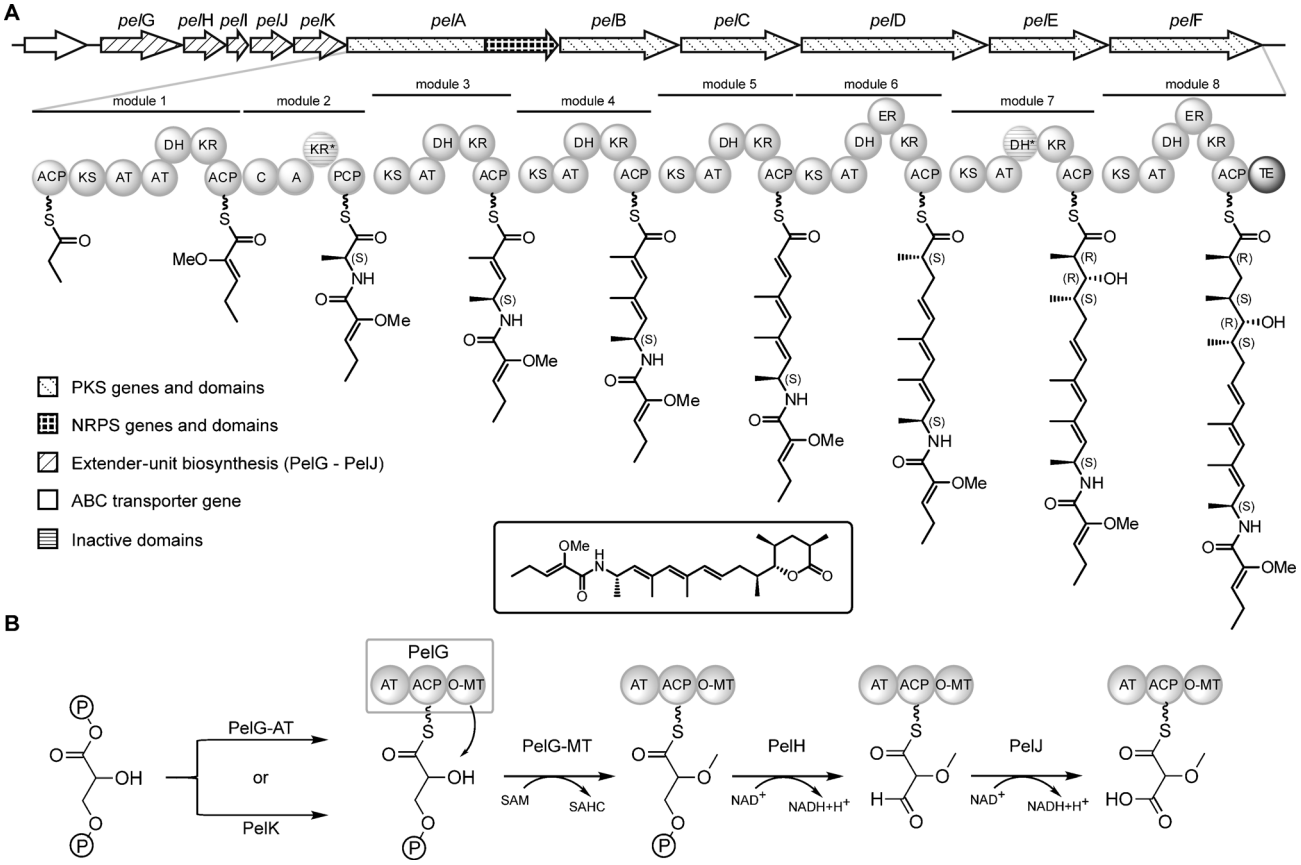
Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201200327.

required predicted substrate specificity for alanine (Scheme 2a). Insertion by single-crossover homologous recombination of a plasmid conferring hygromycin resistance into module 7 led to the abolishment of pellasoren production and thus ultimately confirmed involvement of the *pel* locus in pellasoren biosynthesis (Figure S5 in the Supporting Information).

In line with standard biosynthesis mechanisms for PKS- and NRPS-derived natural products, we reasoned that pellasoren formation starts with the condensation between an ACP-bound propionate and a rare glycolate-derived extender unit catalyzed by module 1, followed by extension with alanine in module 2. Biosynthesis then continues by means of elongation with two methylmalonyl-CoA units (mM-CoA, modules 3 and 4), one malonyl-CoA (M-CoA, module 5), and another three mM-CoA extender units (modules 6 to 8). Finally the full-length intermediate undergoes lactone ring formation and the ACP-bound product is released by the terminal TE domain in module 8 (Scheme 2a). The order and function of domains encoded by genes *pel*A to *pel*F corresponds to the proposed biosynthetic route.

Furthermore, both the predicted substrate specificities of the AT and A domains (Table S5 and Figure S4 in the Supporting Information) and the presence of KR, DH, and ER domains necessary to achieve the appropriate reduction of extender units agree well with the pellasoren scaffold. Modules 2 and 7 encode seemingly superfluous ketoreductase (PelA-KR_2) and dehydratase domains (PelE-DH), respectively; however, these domains are likely inactive as judged on the basis of sequence analysis, which revealed deviations inside otherwise highly conserved amino acid motifs (Tables S4 and S6 in the Supporting Information).

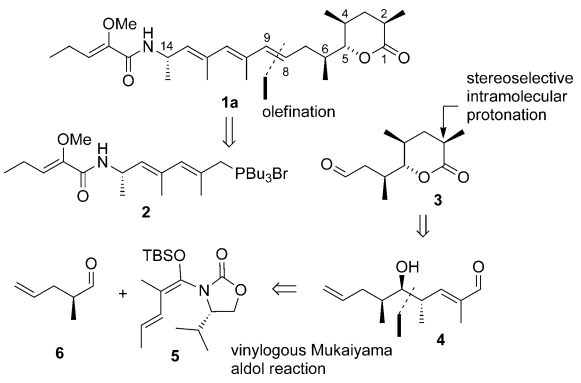
Pellasoren (**1**) features several chiral centers (C2, C4, C5, C6, C14), the configuration of which is set during the biosynthesis. The absence of an epimerization domain within the NRPS module 2 in PelA indicates an *S* configuration at C14 for the L-alanine-derived building block. The configuration at C4 and C5 results from extender-unit reduction by PelE-KR in module 7, whereas stereocenters at C2 and C6 are generated by the ER domains PelD-ER (module 6) and PelF-ER (module 8, Scheme 2a). All active KR domains are of B-type[10] which is in agreement with the 4*R*,5*R* configuration of pellasoren. Kwan et al. highlighted several key amino acid residues in ER domains which appear to exert a crucial influence on the stereochemical outcome of the reduction resulting in either 2*R* or 2*S* configuration of the intermediary building block.[11] A tryptophan residue at position 52 has been shown to direct the conformation to 2*S*, whereas absence of Trp52 favors the 2*R* conformation. In pellasoren biosynthesis, both ER domains should produce 2*R*



Scheme 2. A) The *pel* biosynthetic gene cluster in *Sorangium cellulosum* So ce38 and the proposed biosynthetic route to pellasoren. B) Putative biosynthetic pathway to the glycolate extender unit, allowing for alternative loading of phosphoglycerate onto the PelG-ACP by PelK or PelG-AT. A: adenylation domain, ACP: acyl carrier protein, AT: acyltransferase domain, C: condensation domain, DH: dehydratase domain, ER: enoylreductase domain, KR: ketoreductase domain, KS: ketosynthase domain, TE: thioesterase domain.

centers since they lack Trp52. Nevertheless, they appear to establish converse stereogenic centers during biosynthesis: 2*S* (by PelD-ER) and 2*R* (by PelF-ER). The conflict between the bioinformatics prediction and the proposed NMR-based configuration of pellasoren (**1**) could only be solved by total synthesis of the NMR-based structure. We therefore decided to approach the synthesis of pellasoren A (**1a**).

Our retrosynthetic analysis divides pellasoren into two key building blocks (**2** and **3**) which are supposed to be connected by a Wittig olefination reaction in order to install the sensitive triene at a late stage during the synthesis (Scheme 3). In the synthetic direction, the western fragment **2**

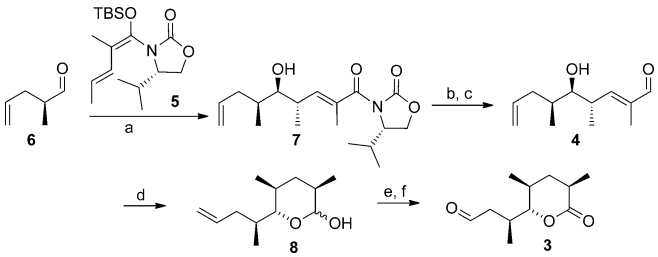


Scheme 3. Retrosynthetic analysis of pellasoren A.

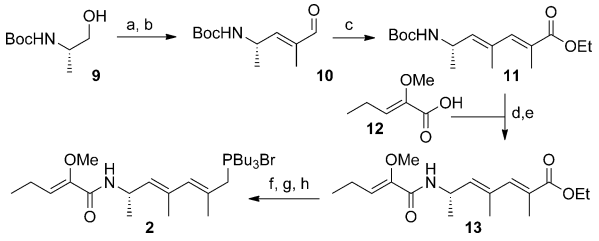
is accessible by a sequence of eight steps, starting from L-alanine (**9**). The eastern fragment **3** contains the lactone subunit and it was planned to be generated by an intramolecular stereoselective protonation of an aldehyde-derived enolate as the key step. This particular transformation takes advantage of conjugate reduction of an α,β-unsaturated aldehyde by Stryker's reagent[12] and the subsequent protonation of the resulting enolate through the secondary alcohol.[13] The required enal **4** is synthesized by a vinylogous Mukaiyama aldol[14] reaction (VMAR) of the chiral vinylketene silyl *N*,*O*-acetal **5** and aldehyde **6** to afford the *anti* aldol product **7** (Scheme 4).

The western fragment **2** could be synthesized starting from Boc-protected amino alcohol **9** (Scheme 5). Swern[16] oxidation established the aldehyde functionality required for the following sequence of two olefination reactions. The first olefination yielded aldehyde **10** which was used in the next transformation to generate ester **11**. The aldehydes were used immediately in the next steps to avoid racemization, yielding ester **11** in 80% *ee*. Liberating the amine with trifluoroacetic acid(TFA) in CH_2Cl_2 at 0 °C and peptide coupling with acid **12** (generated from 2-oxovaleric acid in two steps)[17] gave access to compound **13**. The synthesis continued by reduction of the ester functionality of **13** to the alcohol which was submitted to an Appel reaction to provide corresponding bromide.[18] Reaction with tributylphosphine provided phosphonium bromide **2** and completed the synthesis of this key building block.

With both building blocks in hand we were able to complete the synthesis of pellasoren A (**1a**) by connecting the

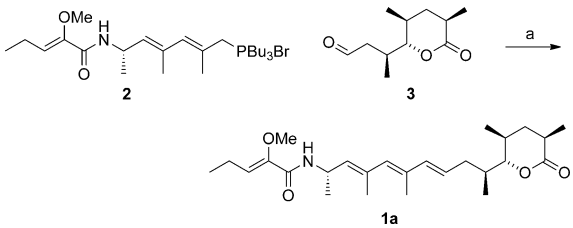


Scheme 4. Synthesis of aldehyde **3** by stereoselective intramolecular protonation. a) $TiCl_4$, CH_2Cl_2, −78 °C to −30 °C, 18 h, 61% (two steps), d.r. 17:1; b) $LiBH_4$, THF, MeOH, 0 °C, 3 h, 90%; c) MnO_2, CH_2Cl_2, RT, 2 h, quant.; d) Stryker's reagent, benzene, RT, 15 h, 65%, d.r. 20:1; e) TPAP, NMO, MS 4 Å, CH_2Cl_2, RT, 14 h, 90%; f) O_3, CH_2Cl_2, −78 °C, 10 min, then PPh_3, RT, 1 h, quant. Stryker's reagent = $[\{(PPh_3)CuH\}_6]$, TPAP[15] = tetrapropylammonium perruthenate, NMO = *N*-methylmorpholine-*N*-oxide, MS = molecular sieves.



Scheme 5. Synthesis of fragment **2**. Reagents and conditions: a) $(COCl)_2$, DMSO, Et_3N, CH_2Cl_2, −78 °C to 0 °C, 2 h; b) $Ph_3P=C(CH_3)CHO$, toluene, 90 °C, 14 h, 65% (two steps); c) $Ph_3P=C(CH_3)CO_2Et$, CH_2Cl_2, RT, 16 h, 85%, 80% *ee* (determined by GC on a chiral stationary phase); d) TFA, CH_2Cl_2, 0 °C, 30 min; e) **12**, EDC, HOBt, CH_2Cl_2, 0 °C, RT, 14 h, 82% (two steps); f) DiBAlH, CH_2Cl_2, −78 °C, 2 h, 87%; g) CBr_4, PPh_3, CH_2Cl_2, RT, 5 min, 76%; h) PBu_3, MeCN, RT, 3 h, quant. TFA = trifluoroacetic acid, EDC = 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride, HOBt = 1-hydroxybenzotriazole, DiBAlH = Diisobutylaluminum hydride.

two segments through an *E*-selective Wittig olefination between C8 and C9 (Scheme 6). Comparison of the ^{1}H NMR data of synthetic and authentic pellasoren confirmed the relative configuration of pellasoren as shown in structure **1a** (see the Supporting Information). Additionally, since the two chromophores are joined at carbon C14, it should be possible to establish the absolute configuration by analyzing the CD data. In order to compare the two isomers, the unnatural diastereomer (14*R*)-pellasoren was also syn-



Scheme 6. Completion of the total synthesis of pellasoren (**1a**). Reagents and conditions: a) KO*t*Bu, THF, −78 °C to RT, 15 h, 46%.

thesized following the same strategy as that described for **1a**. Gratifyingly, comparison of the NMR and CD spectra of pellasoren (**1a**), (14*R*)-pellasoren, and the authentic sample verified the absolute and relative configuration of this natural product as shown for **1a** (Scheme 1).

Verification of the pellasoren configuration by total synthesis confirms the above-described deviations from bioinformatic analysis, and thus underpins the current limits of ER sequence-based stereochemistry prediction. When seeking to improve bioinformatic tools, the pellasoren ER domains may provide a useful starting point for future biochemical and structural studies, as they show 92.4% identity spanning 315 aa residues where Arg43 from PelF-ER is replaced by Leu43 in PelD-ER. We hypothesize here that Arg43 could be another important amino acid to govern the stereochemistry of the reduced extender unit.

The genes involved in biosynthesis of the rarely observed "glycolate" extender unit are located upstream of *pel*A-F. Specifically, we propose that PelG to PelK generate the unusual glycolate extender unit believed to be incorporated by module 2, because these proteins show high similarity to proteins responsible for synthesis of glycolate precursors in the soraphen, ansamitocin, and FK-520 pathways (Table 1). Unfortunately and despite serious efforts, our attempts to prove the hypothesis based on gene cluster analysis by incorporation studies using isotopically labeled precursors and employing MS/MS analysis were not conclusive owing to thorough metabolism of the fed compounds (Figure S7 in the Supporting Information).[7] However, the above-mentioned similarity to known glycolate pathways strongly supports the biosynthesis as shown in Scheme 2.

Table 1: Comparison of proteins involved in biosynthesis of glycolate extender units.

Non-NRPS/PKS proteins of the So ce38 pellasoren biosynthetic machinery

Protein	aa[a]	Proposed function of the homologous protein	Identity/similarity to: soraphen biosynthesis genes	FK-520 biosynthesis genes
PelG	794	acyl transferase (28–355)	57/65% (SorC, AT part)	18/26% (FkbH)
		Ppant attachment site (427–480)	22/28% (SorC, ACP part)	21/29% (FkbJ)
		FkbM methyltransferase (589–744)	39/48% (SorC, OMT part)	8/15% (FkbG)
PelH	283	hydroxyacyl-CoA dehydrogenase	74/82% (SorD)	36/51% (FkbK)
PelI	93	acyl carrier protein	57/67% (SorX)	27/44% (FkbJ)
PelJ	384	acyl-CoA dehydrogenase	49/60% (SorE)	33/48% (FkbI)
PelK	376	HAD superfamily phosphatase	not present	41/57% (FkbH)

[a] Amino acid sequence length.

It is assumed that 1,3-bisphosphoglycerate is loaded to an ACP, and according to soraphen biosynthesis, PelG-ACP binds 1,3-phosphoglycerate loaded by the PelG-AT domain. Methylation could then be catalyzed by the internal SAM-dependent PelG-MT domain. These functions are usually facilitated by discrete proteins rather than a multifunctional protein. Indeed, it appears that a trifunctional enzyme is present (Scheme 2b). Moreover, the presence of the *pelK* gene is fairly unexpected. PelK shows homology to FkbH, a protein that is assumed to load 1,3-bisphosphoglycerate to FkbJ in the FK-520 pathway.[5] The soraphen biosynthesis does not have a FkbH analogue but uses a PelG homologue instead. Since PelG-AT is supposed to catalyze this loading step during pellasoren biosynthesis, PelK seems to be redundant. The *pel* gene cluster actually represents the first finding with this specific gene arrangement. However, when speculating about a hypothetical function for PelK in pellasoren biosynthesis it has to be pointed out that the precise mechanism for loading of phosphoglycerate has to date not been experimentally proven for any of the known biosynthetic pathways which employ glycolate extender units. Thus, the presence of two proteins potentially able to catalyze the same loading reaction in the pellasoren pathway raises more questions concerning the biosynthesis of this extender unit. The involvement of *pelK* in the latter using targeted inactivation is not straightforward because of the *pel* operon organization, but will be the subject of future studies.

Pellasoren A is cytotoxic against human colon cancer cells of the HCT-116 cell line with an IC50 of 155 nM. Interestingly, pellasoren B is around one magnitude less active against HCT-116 with an IC50 of 2.35 μM, emphasizing the importance of the linear an all-(*E*) configuration. Furthermore, both pellasorens show a strong effect on lysosomes. Upon incubation with osteosarcoma cells (U-2 OS) using an acridine orange assay, the pH of the lysosomes changed from acidic to neutral values (Figure S10 in the Supporting Information). This is related to an apoptotic mechanism that remains elusive at the current state of research.[19]

This report features a new cytotoxic compound from *Sorangium cellulosum*. We present the first total synthesis of pellasoren and thereby validated the molecule's stereochemistry. A Wittig reaction in the last step of the synthesis was used to install the sensitive polyene system. Further key steps include a vinylogous Mukaiyama aldol reaction of a chiral vinylketene silyl *N,O*-acetal and a stereoselective intramolecular protonation, which was shown to be an efficient and practical protocol in natural products total syntheses for the first time. Identification of the biosynthetic machinery nicely complements this work by enabling bioinformatics and chemical analyses including an unusual pathway to the glycolate extender unit.

Experimental Section

The pellasoren gene cluster sequence was deposited in the EMBL database with the accession number HE616533. Full experimental details, including stucture elucidation procedures and bioinformatic analysis, can be found in the Supporting Information.

Received: January 12, 2012
Revised: February 16, 2012
Published online: April 5, 2012

Keywords: biosynthesis · myxobacteria · natural products · *Sorangium cellulosum* · total synthesis

[1] K. Gerth, H. Steinmetz, G. Höfle, R. Jansen, *Angew. Chem.* **2008**, *120*, 610 – 613; *Angew. Chem. Int. Ed.* **2008**, *47*, 600 – 602.
[2] S. C. Wenzel, R. Müller, *Curr. Opin. Drug Discovery Dev.* **2009**, *12*, 220 – 230.
[3] H. Irschik, H. Reichenbach, G. Höfle, R. Jansen, *J. Antibiot.* **2007**, *60*, 733 – 738.
[4] H. Reichenbach, G. Höfle in *Drug Discovery from Nature* (Eds.: S. Grabley, R. Thiericke), Springer, Berlin, **1999**, p. 173.
[5] B. J. Carroll, S. J. Moss, L. Q. Bai, Y. Kato, S. Toelzer, T. W. Yu, H. G. Floss, *J. Am. Chem. Soc.* **2002**, *124*, 4176 – 4177.
[6] K. Wu, L. Chung, W. P. Revill, L. Katz, C. D. Reeves, *Gene* **2000**, *251*, 81 – 90.
[7] S. C. Wenzel, R. M. Williamson, C. Grünanger, J. Xu, K. Gerth, R. A. Martinez, S. J. Moss, B. J. Carroll, S. Grond, C. J. Unkefer, R. Müller, H. G. Floss, *J. Am. Chem. Soc.* **2006**, *128*, 14325 – 14336.
[8] S. C. Wenzel, R. Müller, *Nat. Prod. Rep.* **2009**, *26*, 1385 – 1407.
[9] B. O. Bachmann, J. Ravel, *Methods Enzymol.* **2009**, *458*, 181 – 217.
[10] a) P. Caffrey, *ChemBioChem* **2003**, *4*, 654 – 657; b) R. Reid, M. Piagentini, E. Rodriguez, G. Ashley, N. Viswanathan, J. Carney, D. V. Santi, C. R. Hutchinson, R. McDaniel, *Biochemistry* **2003**, *42*, 72 – 79.
[11] D. H. Kwan, Y. Sun, F. Schulz, H. Hong, B. Popovic, J. C. Sim-Stark, S. F. Haydock, P. F. Leadlay, *Chem. Biol.* **2008**, *15*, 1231 – 1240; Ref. [1].
[12] a) W. S. Mahoney, D. M. Brestensky, J. M. Stryker, *J. Am. Chem. Soc.* **1988**, *110*, 291 – 293; b) B. H. Lipshutz, J. M. Serevesko, B. R. Taft, *J. Am. Chem. Soc.* **2004**, *126*, 8352 – 8353.
[13] A. Kena Diba, C. Noll, M. Richter, M. T. Gieseler, M. Kalesse, *Angew. Chem.* **2010**, *122*, 8545 – 8547; *Angew. Chem. Int. Ed.* **2010**, *49*, 8367 – 8369.
[14] a) S. Shirokawa, M. Kamiyama, T. Nakamura, M. Okada, A. Nakazaki, S. Hosokawa, S. Kobayashi, *J. Am. Chem. Soc.* **2004**, *126*, 13604 – 13605; b) for a general review see: G. Casiraghi, L. Battistini, C. Curti, G. Rassu, F. Zanardi, *Chem. Rev.* **2011**, *111*, 3076 – 3154; c) S. Shirokawa, M. Shinoyama, I. Ooi, S. Hosokawa, A. Nakazaki, S. Kobayashi, *Org. Lett.* **2007**, *9*, 849 – 852; d) S. Hosokawa, K. Matsushita, S. Tokimatsu, T. Toriumi, Y. Suzuki, K. Tatsuta, *Tetrahedron Lett.* **2010**, *51*, 5532 – 5536.
[15] S. V. Ley, J. Norman, W. P. Griffith, S. P. Marsden, *Synthesis* **1994**, 639 – 666.
[16] K. Omura, D. Swern, *Tetrahedron* **1978**, *34*, 1651 – 1660.
[17] A. Fürstner, M. M. Domostoj, B. Scheiper, *J. Am. Chem. Soc.* **2005**, *127*, 11620 – 11621.
[18] R. Appel, *Angew. Chem.* **1975**, *87*, 863 – 874; *Angew. Chem. Int. Ed. Engl.* **1975**, *14*, 801 – 811.
[19] a) K. Kågedal, M. Zhao, I. Svensson, U. T. Brunk, *Biochem. J.* **2001**, *359*, 335 – 343; b) A. C. Johansson, H. Appelqvist, C. Nilsson, K. Kagedal, K. Roberg, K. Ollinger, *Apoptosis* **2010**, *15*, 527 – 540.

