



Structure Elucidation and Activity of Kolossin A, the D-/L-Pentadecapeptide Product of a Giant Nonribosomal Peptide Synthetase**

Helge B. Bode,* Alexander O. Brachmann, Kirtikumar B. Jadhav, Lydia Seyfarth, Christina Dauth, Sebastian W. Fuchs, Marcel Kaiser, Nick R. Waterfield, Holger Sack, Stefan H. Heinemann, and Hans-Dieter Arndt*

Dedicated to Professor Peter B. Dervan on the occasion of his 70th birthday

Abstract: The largest continuous bacterial nonribosomal peptide synthetase discovered so far is described. It consists of 15 consecutive modules arising from an uninterrupted, fully functional gene in the entomopathogenic bacterium *Photobacterium luminescens*. The identification of its cryptic biosynthesis product was achieved by using a combination of genome analysis, promoter exchange, isotopic labeling experiments, and total synthesis of a focused collection of peptide candidates. Although it belongs to the growing class of D-/L-peptide natural products, the encoded metabolite kolossin A was found to be largely devoid of antibiotic activity and is likely involved in interspecies communication. A stereoisomer of this peculiar natural product displayed high activity against *Trypanosoma brucei rhodesiense*, a recalcitrant parasite that causes the deadly disease African sleeping sickness.

Advanced sequencing methods and bioinformatics-based analysis have enabled the investigation of underexplored or even unknown compounds in natural products research.^[1–4] The determination of their constitution has greatly profited from improvements in MS-based methods.^[5–8] Furthermore, advances in genomic data acquisition and correlation has enabled prediction of the product stereochemistry of polyketides and nonribosomal peptide synthetases (NRPSs) with increasing fidelity.^[9–11] However, in order to make putative

products of cryptic biosynthesis accessible for further research, material from fermentation must be isolated on a preparative scale. With only analytical amounts of material accessible, chemical synthesis becomes increasingly important with respect to determining the identity and stereochemistry of the products of cryptic gene clusters, as well as for assigning their bioactivity and function.

Particularly challenging in this regard are very large gene clusters that cannot easily be cloned or expressed in heterologous hosts, as well as producer organisms with a complex life style, such as symbionts or species-selective pathogens. Herein, we describe the computational prediction of the cryptic biosynthesis product kolossin A (**1**) from *Photobacterium luminescens*, the induced expression of its single 1.8 MDa NRPS Kol, and the stringent confirmation of the formylated D-/L-pentadecapeptide product structure (**1**) through labelling experiments and chemical synthesis, which also enabled bioactivity profiling.

In the genome of the entomopathogenic model bacterium *Photobacterium luminescens* TTO1, several biosynthesis gene clusters have been characterized that are involved in the biosynthesis of diverse natural products.^[12,13] Among them we identified the *kol* gene (*plu2670*, 49.1 kbp), which encodes the (to our knowledge) largest single NRPS of bacterial origin (Kol) with a predicted molecular weight of 1.81 MDa, a size

[*] Prof. Dr. H. B. Bode, Dr. A. O. Brachmann, Dipl.-Chem. C. Dauth, Dr. S. W. Fuchs
Merck-Stiftungsprofessur für Molekulare Biotechnologie
Goethe Universität Frankfurt, Biozentrum, Campus Riedberg 250
Max-von-Laue-Str. 9, 60438 Frankfurt/Main (Germany)
E-mail: h.bode@bio.uni-frankfurt.de

Dr. K. B. Jadhav, Dr. L. Seyfarth, Prof. Dr. H.-D. Arndt
Friedrich-Schiller-Universität
Institut für Organische Chemie und Makromolekulare Chemie
Humboldtstr. 10, 07743 Jena (Germany)
E-mail: hd.arndt@uni-jena.de

Prof. Dr. H. B. Bode
Buchmann Institut für Molekulare Lebenswissenschaften (BMLS)
Goethe Universität Frankfurt
Max-von-Laue-Str. 15, 60438 Frankfurt/Main (Germany)

Dr. M. Kaiser
Schweizerisches Tropen- und Public-Health-Institut
Parasiten-Chemotherapie, Socinstr. 57, 4051 Basel (Switzerland)

Dr. N. R. Waterfield
University of Warwick, Warwick Medical School
Coventry CV4 7AL (UK)

Dipl.-Phys. H. Sack, Prof. Dr. S. H. Heinemann
Zentrum für Molekulare Biomedizin/Biophysik
Friedrich-Schiller-Universität Jena und Universitätsklinikum Jena
Hans-Knöll-Str. 2, 07745 Jena (Germany)

[**] Support by the DAAD (predoctoral fellowship to K.B.J.) is gratefully acknowledged. This work profited in part from support by the TMBWK (grant No. 43-5572-321-12040-12 to H.-D.A.), the DFG (SFB 1127, to H.-D.A.) and the European commission (ERC starting grant agreement No. 311477 to H.B.B.). We thank Dr. O. Makarewicz (UKJ Jena) for conducting antibacterial activity tests.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201502835>.

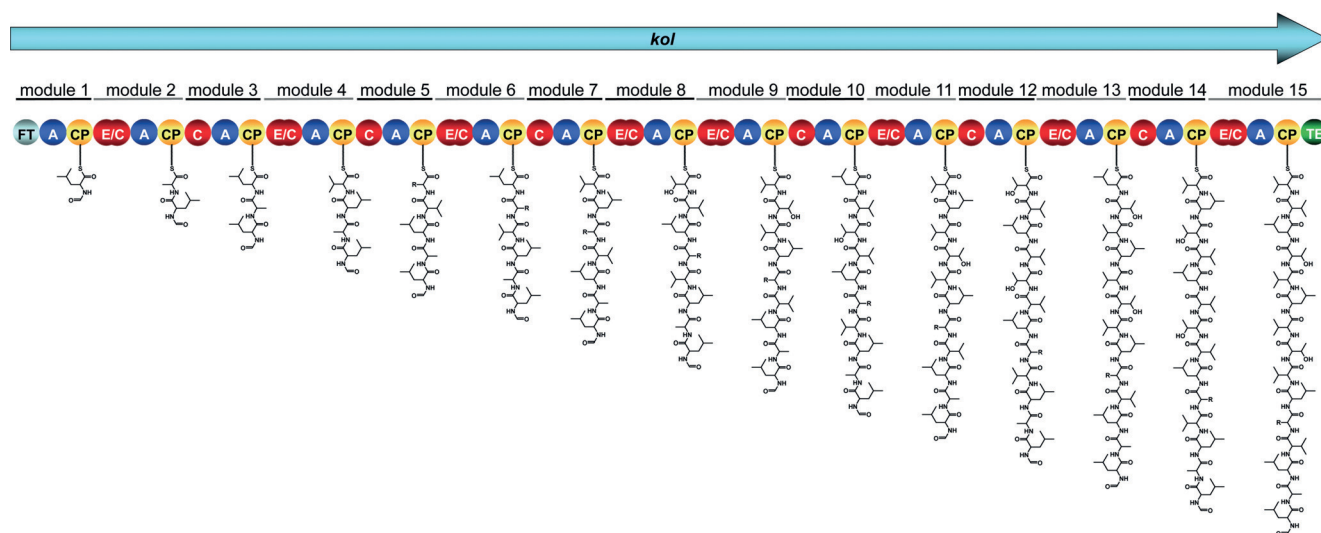


Figure 1. Model of the NRPS Kol, which is responsible for the biosynthesis of kolossin A (**1**) and contains 15 consecutive modules. Domains: C: condensation, E/C: dual epimerase/condensation, A: adenylation, T: thiolation, TE: thioesterase, FT: formyltransferase, CP: peptidyl carrier protein.

comparable to that of the bacterial ribosome. Detailed analysis indicated 15 NRPS modules with 45 domains (Figure 1) and Kol was thus expected to produce a pentadecapeptide.^[14] However, we were unable to detect such peptides under a variety of fermentation conditions, including cultivation of *P. luminescens* in the larvae of the greater waxmoth *Galleria mellonella* to imitate the natural life cycle of *Photorhabdus* (data not shown). Similarly, transcription analysis by RNAseq did not show significant activation of the *kol* gene, either in the exponential or stationary growth phases or in the presence of insect hemolymph (Figure S1 in the Supporting Information). Fortunately, artificial activation of this silent gene cluster could be achieved by applying a promoter exchange approach as described earlier^[15] to the *kol* gene, thus resulting in the detection of three novel biosynthesis products by HPLC–MS monitoring (Figure 2).

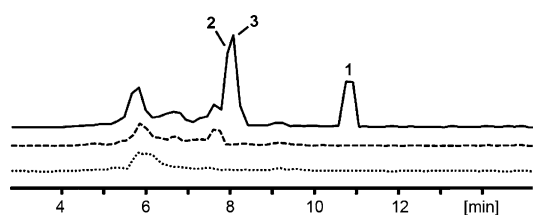


Figure 2. HPLC–MS chromatograms of XAD extracts from metabolites produced by wild type *P. luminescens* (.....) and *P. luminescens* pCEP-kol without (----) and with (—) arabinose induction. Extracted ion chromatograms (EICs) of the *m/z* for 1–3 (Table S3) are shown.

Unfortunately, the amount of material produced did not allow any detailed chemical analysis or structure determination. However, molecular formulae, as well as amino acid sequences, were delineated from MSⁿ experiments (Figure 1, Tables S3 and S4, and Figure S2). These data suggested the hydrophobic compound **1** to be the full-length kolossin A,

with kolossin B (**2**) and kolossin C (**3**) as C-terminally truncated fragments. The configurations of the stereogenic centers were predicted based on the presence of dual condensation/epimerization (C/E) domains^[16,17] in Kol, which indicate fully alternating D-/L-configuration with the exception of two consecutive D-configurations at Val⁷ and Thr⁸ (Table S5). This places kolossin A among rare linear D-/L-peptide natural products such as gramicidin A,^[18] feglymycin,^[19] or the polytheonamides,^[20] all of which feature a propensity for β -helical structures and often show potent bioactivity. Interestingly, similar hydrophobic D-/L-depsipeptides, the xenoamcins, have been recently identified in *Xenorhabdus*, but with no linear derivatives.^[21]

To test the activity of the C/E domains in the Kol NRPS and thus the highly relevant configuration of the biosynthesis product,^[18–20] labelling experiments with fully deuterated alanine, valine, leucine, tyrosine, and threonine were performed in an $\Delta ilvE$ mutant of *P. luminescens*.^[8] The loss of one deuterium atom for tyrosine and threonine (Figures S2 and S4) strongly suggest that at least the C/E domains following the Tyr- and Thr-incorporating modules 5, 8, and 12 were functional, and thus that both Thr residues should feature the D-configuration. The $\Delta ilvE$ mutant is defective in the amino-transferase reaction for Leu and Ile and therefore was ideal for analyzing the epimerization of these amino acids. Indeed, incorporation of one L- and one D-leucine was detected for the truncated products **2** and **3**, thus further supporting the proposed configuration of these amino acids.

The configuration at the β -substituted Thr residues was not directly accessible by isotopic labelling. In order to confirm the likely structure and provide material for bioactivity tests, access to synthetic kolossin A was necessary. Furthermore, to unequivocally validate the compound identity by LC–MS and to complement biological testing, chemical synthesis of stereoisomers was pursued. Accordingly, by employing permutations of D-allo-Thr, D-Thr, and L-

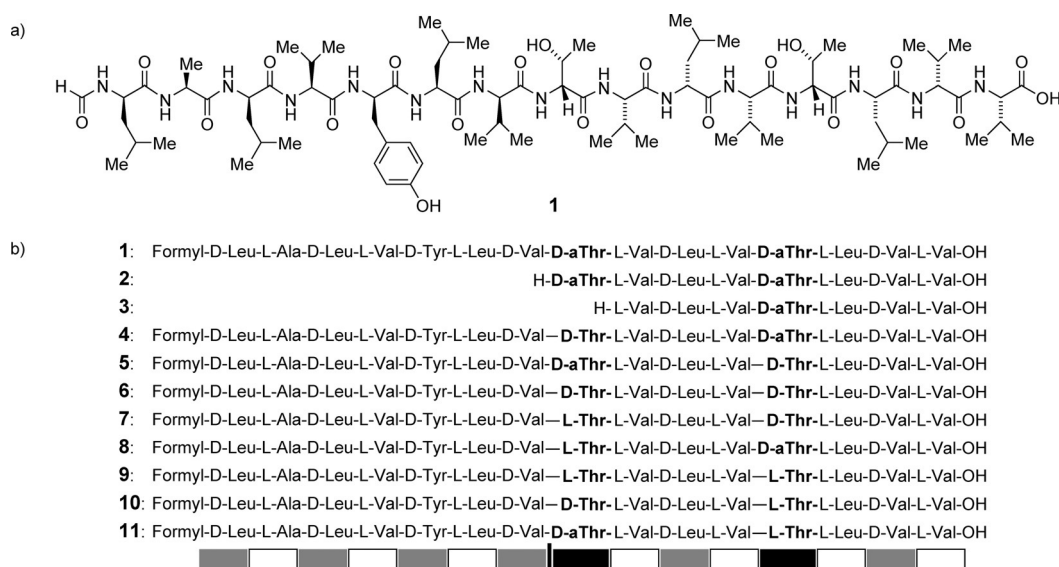
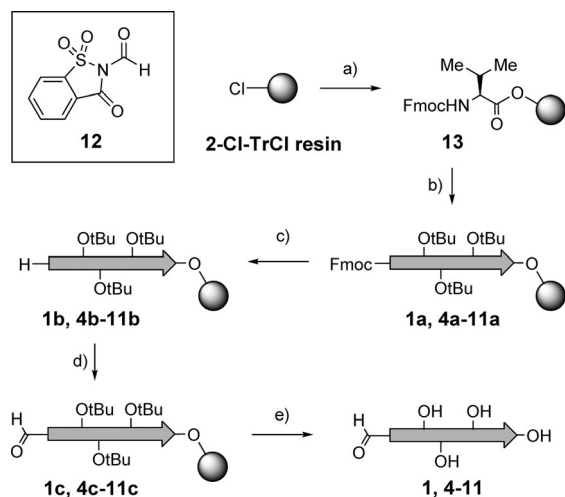


Figure 3. a) Structural formula of kolossin A (**1**). b) Sequences of the investigated natural products **1–3** and isomers **4–11**. Positions of variation (residue 8 and 12) and the backbone stereochemistry change between D (gray) and L (white) are highlighted, as well as the irregular continuation between residues 7 and 8 (aThr denotes *allo*-Threonine).

Thr at positions 8 and 12, a nine-membered library of kolossin A analogues was designed (Figure 3).

All stereoisomers were independently synthesized by using Fmoc-based solid phase peptide synthesis (SPPS)^[22] on 2-chlorotrityl chloride resin.^[23] Peptides were sequentially assembled from C terminus to N terminus using an automated peptide synthesizer (Scheme 1). Considerable exper-



Scheme 1. Optimized solid-phase synthesis of kolossin A library members (**1**, **4–11**). Reagents and conditions: a) Fmoc-L-Val-OH, EtN(*i*Pr)₂, CH₂Cl₂, 16 h; b) piperidine/DMF 1:3 + 2% DBU; c) amino acid *n* (0.40 M in DMF), HBTU, HOBT (0.42 M in DMF), EtN(*i*Pr)₂ (2 M in DMF), 1 h; d) Ac₂O/pyridine/CH₂Cl₂ = 1:1:4, repeat from b) for all *n* = 14 → *n* = 1; e) **12** (saturated in THF), THF/EtN(*i*Pr)₂ 9:1, 2 × 15 min; f) TFA/Et₃SiH/H₂O/CH₂Cl₂ (70:7.5:7.5:15), 20 °C, 1.5 h. DBU = 1,8-diazabicyclo[5.4.0]undec-7-en; DMF = dimethylformamide; HBTU = O-(benzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate; HOBT = 1-hydroxybenzotriazole; HFIP = 1,1,1,3,3,3-hexafluoroisopropanol; TFA = trifluoroacetic acid.

imentation based on precedence with difficult lipophilic D-/L-peptides^[24,25] led to an optimized synthetic method featuring a coupling cycle consisting of accelerated Fmoc deprotections, HBTU couplings in DMF as the solvent, and systematic capping (Scheme 1). Initial loading of resin **13** was 0.4 mmol g⁻¹. Conventional N-terminal formylation of fully deprotected D-/L-peptides in solution or by using standard reagents^[26] on resin-bound peptide led to impure prod-

ucts, apparently owing to lack of chemoselectivity and difficult removal of side products. Fortunately, formylation of the resin-bound, side-chain-protected peptide through treatment with saccharin formate (**12**)^[27] in THF was successful and selective for the terminal amino group. The crude peptides were released from the solid support and deprotected under carefully optimized conditions that left the N-terminal formyl group intact to give the corresponding kolossin A analogues **1** and **4–11** in more than 90% crude yields. The low solubility and the aggregation associated with these peptides were overcome by dissolving crude peptides in HFIP immediately prior to RP-HPLC purification (C4 column, MeOH/H₂O, pH 3.5). On average, 20–23 mg of the kolossin A analogues were obtained from 100 mg of resin in high purity with excellent overall yields ranging from 31–34%. All compounds were fully characterized and distinguished by NMR spectroscopy and by their chromatographic properties despite their being diastereoisomeric in one amino acid side chain only. By using a combination of HPLC-MS and fragmentation analysis, the chemical structure and configuration of the scarce (μg) natural product was clearly correlated to **1**, thus confirming the structure of kolossin A and the configuration of residues 8 and 12 as D-*allo*-Thr (Figure S5).

Upon biological screening, neither kolossin A nor its isomers showed cytotoxicity against rat myoblast L6 cells or antibacterial activity against Gram-positive (*Staphylococcus aureus*, *S. epidermidis*, *Enterococcus faecalis*, *E. faecium*) or Gram-negative (*Escherichia coli* wt, *Klebsiella pneumonia*, *Proteus mirabilis*, *Pseudomonas aeruginosa*) bacteria. To test for pore-forming properties in membranes, as often associated with D-/L-peptides,^[18–20] all of the candidates were investigated in a black lipid bilayer test system. None of the compounds was found to induce measurable conductivity (< 1 μS) at a bath concentration of 1 μM (100 mM KCl, pH 8.0

TRIS/HCl), whereas gramicidin A (gA) as a control induced conductivities up to 100 μ S at 0.5 nM (Figure S6). Hence, the general pore-forming properties of the kolossins are 10^6 – 10^7 -fold less pronounced than for gA and likely absent at biological concentrations. Further tests of compounds **1** and **4–11** were carried out with the eukaryotic parasites *Trypanosoma brucei rhodesiense*, *T. cruci*, *Leishmania donovani*, and *Plasmodium falciparum*.^[21,28] Notable activity (IC_{50} = 2.7 μ M) was observed for compound **8** against *T. brucei rhodesiense*, the causative agent of African sleeping sickness, and for compound **11** against *T. brucei r.* (IC_{50} = 8.9 μ M) and *P. falciparum* (IC_{50} = 16.1 μ M), the causative agent of malaria (Table S6). All other isomers and the natural product itself were inactive.

In summary, bioinformatic prediction of a NRPS-derived peptide, stimulated overexpression, MS analysis and labelling experiments were combined to unequivocally prove the chemical structure and configuration of the giant silent NRPS gene cluster product kolossin A. The mere size of the *kol* gene and Kol NRPS, the amount of precursors, and energy and time required for biosynthesis suggests an important biological role for the kolossins. Enabled by total synthesis, no toxic or antibiotic bioactivities were found to be associated with the kolossins. Pore-forming properties were not detected, which in line with the interruption of the regular D-/L-stereochemistry of the backbone. Furthermore, its own producer *P. luminescens* was found to be insensitive to the compound. In line with the observation that kolossin production could not be effectively induced under culture conditions in the laboratory, our data suggest that the production of kolossin might require a specific signaling factor, potentially one connected to the insect host or a specific stage of its life cycle. The non-natural isomer **8** was found to be specifically active against *T. brucei rhodesiense*, a consistently fatal infection if untreated and a serious health issue in sub-Saharan Africa. Further studies will concentrate on pursuing this promising activity, as well as potentially linking it to the elusive endogenous function of kolossin in its producing bacterium *P. luminescens*.

Keywords: biological activity · biosynthesis · peptides · natural products · total synthesis

How to cite: *Angew. Chem. Int. Ed.* **2015**, *54*, 10352–10355
Angew. Chem. **2015**, *127*, 10493–10496

- [1] A. O. Brachmann, F. Kirchner, C. Kegler, S. C. Kinski, I. Schmitt, H. B. Bode, *J. Biotechnol.* **2012**, *157*, 96–99.
- [2] J. B. Biggins, X. Liu, Z. Feng, S. F. Brady, *J. Am. Chem. Soc.* **2011**, *133*, 1638–1641.
- [3] J. Piel, *Annu. Rev. Microbiol.* **2011**, *65*, 431–453.
- [4] K. Yamanaka, K. A. Reynolds, R. D. Kersten, K. S. Ryan, D. J. Gonzalez, V. Nizet, P. C. Dorrestein, B. S. Moore, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 1957–1962.
- [5] R. D. Kersten, Y.-L. Yang, Y. Xu, P. Cimermancic, S.-J. Nam, W. Fenical, M. A. Fischbach, B. S. Moore, P. C. Dorrestein, *Nat. Chem. Biol.* **2011**, *7*, 794–802.
- [6] R. D. Kersten, P. C. Dorrestein, *ACS Chem. Biol.* **2009**, *4*, 599–601.
- [7] S. W. Fuchs, C. C. Sachs, C. Kegler, F. I. Nollmann, M. Karas, H. B. Bode, *Anal. Chem.* **2012**, *84*, 6948–6955.
- [8] C. Cregler, F. I. Nollmann, T. Ahrendt, F. Fleischhacker, E. Bode, H. B. Bode, *ChemBioChem* **2014**, *15*, 826–828.
- [9] M. H. Medema, R. van Raaphorst, E. Takano, R. Breitling, *Nat. Rev. Microbiol.* **2012**, *10*, 191–202.
- [10] K. Blin, M. H. Medema, D. Kazempour, M. A. Fischbach, R. Breitling, E. Takano, T. Weber, *Nucleic Acids Res.* **2013**, *41*, W204–12.
- [11] A. O. Brachmann, H. B. Bode, *Adv. Biochem. Eng. Biotechnol.* **2013**, *135*, 123–155.
- [12] E. Duchaud, C. Rusniok, L. Frangeul, C. Buchrieser, A. Givaudan, S. Taourit, S. Bocs, C. Boursaux-Eude, M. Chandler, J.-F. Charles, et al., *Nat. Biotechnol.* **2003**, *21*, 1307–1313.
- [13] a) H. B. Bode, *Curr. Opin. Chem. Biol.* **2009**, *13*, 224–230; b) J. Yin, H. Zhu, L. Xia, X. Ding, T. Hoffmann, M. Hoffmann, X. Bian, R. Müller, J. Fu, A. F. Stewart, Y. Zhang, *Nucleic Acids Res.* **2015**, *43*, e36.
- [14] The biosynthesis of the D-/L-pentadecapeptide gramicidin A features four genetically independent NRPS modules: N. Kessler, H. Schumann, S. Morneweg, U. Linne, M. A. Marahiel, *J. Biol. Chem.* **2004**, *279*, 7413–7419.
- [15] E. Bode, A. O. Brachmann, C. Kegler, R. Simsek, C. Dauth, Q. Zhou, M. Kaiser, P. Klemmt, H. B. Bode, *ChemBioChem* **2015**, *16*, 1115–1119.
- [16] H. B. Bode, D. Reimer, S. W. Fuchs, F. Kirchner, C. Dauth, C. Kegler, W. Lorenzen, A. O. Brachmann, P. Grün, *Chem. Eur. J.* **2012**, *18*, 2342–2348.
- [17] T. Stachelhaus, C. T. Walsh, *Biochemistry* **2000**, *39*, 5775–5787.
- [18] a) R. Sarges, B. Witkop, *J. Am. Chem. Soc.* **1965**, *87*, 2020–2027; b) D. W. Urry, *Proc. Natl. Acad. Sci. USA* **1971**, *68*, 672–676; c) R. R. Ketchum, W. Hu, T. A. Cross, *Science* **1993**, *261*, 1457–1460.
- [19] a) G. Bunkóczi, L. Vértessy, G. M. Sheldrick, *Angew. Chem. Int. Ed.* **2005**, *44*, 1340–1342; *Angew. Chem.* **2005**, *117*, 1364–1366; b) F. Dettner, A. Hänchen, D. Schols, L. Toti, A. Nusser, R. D. Süßmuth, *Angew. Chem. Int. Ed.* **2009**, *48*, 1856–1861; *Angew. Chem.* **2009**, *121*, 1888–1893.
- [20] a) T. Hamada, S. Matsunaga, G. Yano, N. Fusetani, *J. Am. Chem. Soc.* **2005**, *127*, 110–118; b) M. Inoue, N. Shinohara, S. Tanabe, T. Takahashi, K. Okura, H. Itoh, Y. Mizoguchi, M. Iida, N. Lee, S. Matsuoka, *Nat. Chem.* **2010**, *2*, 280–285.
- [21] Q. Zhou, F. Grundmann, M. Kaiser, M. Schiell, S. Gaudriault, A. Batzer, M. Kurz, H. B. Bode, *Chem. Eur. J.* **2013**, *19*, 16772–16779.
- [22] I. Coin, M. Beyermann, M. Bienert, *Nat. Protoc.* **2007**, *2*, 3247–3256.
- [23] K. Barlos, D. Gatos, J. Kallitsis, G. Papaphotiu, P. Sotiriou, Y. Wenqing, W. Schäfer, *Tetrahedron Lett.* **1989**, *30*, 3943–3946.
- [24] a) H.-D. Arndt, A. Knoll, U. Koert, *Angew. Chem. Int. Ed.* **2001**, *40*, 2076–2078; *Angew. Chem.* **2001**, *113*, 2137–2140; b) H.-D. Arndt, D. Bockelmann, A. Knoll, S. Lamberth, C. Griesinger, U. Koert, *Angew. Chem. Int. Ed.* **2002**, *41*, 4062–4065; *Angew. Chem.* **2002**, *114*, 4234–4238.
- [25] a) R. J. Lichtenecker, B. Ellinger, H.-M. Han, K. B. Jadhav, S. Baumann, O. Makarewicz, M. Grabenbauer, H.-D. Arndt, *ChemBioChem* **2013**, *14*, 2492–2499; b) K. B. Jadhav, R. J. Lichtenecker, A. Bullach, B. Mandal, H.-D. Arndt, *Chem. Eur. J.* **2015**, *21*, 5898–5908.
- [26] a) L. I. Krimen, *Org. Synth.* **1970**, *50*, 1–2; b) G. H. Zayia, *Org. Lett.* **1999**, *1*, 989–991; c) A. R. Katritzky, H.-X. Chang, B. Yang, *Synthesis* **1995**, 503–505.
- [27] T. Cochet, V. Bellosta, A. Greiner, D. Roche, J. Cossy, *Synlett* **2011**, 1920–1922.
- [28] F. Grundmann, M. Kaiser, M. Kurz, M. Schiell, A. Batzer, H. B. Bode, *RSC Adv.* **2013**, *3*, 22072–22077.

Received: March 27, 2015

Published online: June 26, 2015