Natural Product Purification

DOI: 10.1002/anie.201308363

Use of a Phosphonate Methyltransferase in the Identification of the Fosfazinomycin Biosynthetic Gene Cluster**

Jiangtao Gao, Kou-San Ju, Xiaomin Yu, Juan E. Velásquez, Subha Mukherjee, Jaeheon Lee, Changming Zhao, Bradley S. Evans, James R. Doroghazi, William W. Metcalf, and Wilfred A. van der Donk**

Abstract: *Natural product discovery has been boosted by genome mining approaches, but compound purification is often still challenging. We report an enzymatic strategy for "stable isotope labeling of phosphonates in extract" (SILPE) that facilitates their purification. We used the phosphonate methyltransferase DhpI involved in dehydrophos biosynthesis to methylate a variety of phosphonate natural products in crude spent medium with a mixture of labeled and unlabeled S-adenosyl methionine. Mass-guided fractionation then allowed straightforward purification. We illustrate its utility by purifying a phosphonate that led to the identification of the fosfazinomycin biosynthetic gene cluster. This unusual natural product contains a hydrazide linker between a carboxylic acid and a phosphonic acid. Bioinformatic analysis of the gene cluster provides insights into how such a structure might be assembled.*

Natural products have gained renewed interest with the realization that the genetic capacity of microorganisms to produce these molecules is much larger than anticipated. Indeed, genome mining exercises have resulted in the discovery of a variety of new compounds.[1] Whereas the sequenced bacterial and fungal genomes demonstrate the untapped potential of natural products, connecting the biosynthetic genes in genomes to novel compounds is still a challenge. One aspect of this challenge is the difficulty of purifying these molecules from complex spent media. Recently, chemoselective derivatization approaches have been developed to aid in the purification process.[2] Herein we illustrate the use of an enzymatic approach to rapidly enrich a phosphonate from the spent medium of *Streptomyces* sp. WM6372, a strain previously shown to encode an uncharacterized phosphonate biosynthetic gene cluster. Isolation and structure elucidation identified the molecule as methyl-2-hydroxy-2-phosphonoacetic acid (Me-HPnA). Realization that Me-HPnA is present in fosfazinomycins prompted a targeted search after growth of the strain in various media. We also examined *Streptomyces* XY332, which was previously shown to encode the same phosphonate biosynthetic genes. Both strains produced fosfazinomycins, and heterologous expression of the genes in *S. lividans* led to production of related phosphonic acids, providing strong support for linking the gene cluster to fosfazinomycin production. Analysis of the genes provides insights into the biosynthetic pathway of hydrazine formation in nature.

Phosphonate natural products form a particular challenge in regards to purification, as a consequence of their high polarity and water solubility. These properties may explain why only about 30 phosphonate natural products have been isolated, despite the observation that about 5% of randomly isolated actinomycetes contain the genetic capability to produce phosphonates.[3] The phosphonate *O*-methyltransferase DhpI, from the dehydrophos biosynthetic pathway, is able to non-specifically methylate other phosphonates such as fosfomycin and fosmidomycin under defined reaction conditions.[4] To determine whether DhpI could methylate these compounds in a complex medium, fosfomycin and fosmidomycin were added to international streptomyces project 4 (ISP4) medium. Indeed, when supplied with *S*-adenosyl methionine (SAM), DhpI methylated both compounds as determined by ^{31}P NMR spectroscopy (Figure 1a; see also the Supporting Information, Figures S1 and S2, and Tables S1 and

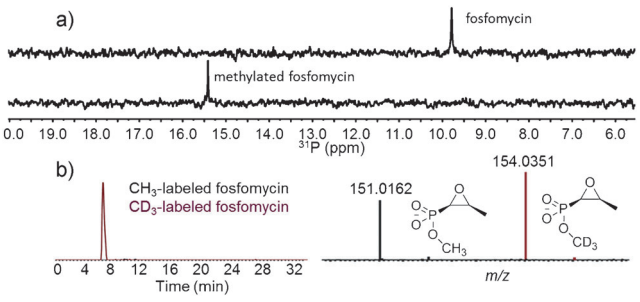

Figure 1. a) ^{31}P NMR spectrum of fosfomycin (top) and after treatment with DhpI and a 1:1 mixture of SAM and [D_3]SAM (bottom). b) LC-MS analysis of the sample in the bottom spectrum of (a).

[*] Dr. J. Gao, Dr. K. Ju, Dr. J. Lee, Dr. C. Zhao, Dr. B. S. Evans, Dr. J. R. Doroghazi
Institute for Genomic Biology, University of Illinois at Urbana-Champaign (USA)
X. Yu, Prof. W. W. Metcalf
University of Illinois at Urbana-Champaign
Department of Microbiology
601 South Goodwin Avenue, Urbana, IL 61801 (USA)
E-mail: metcalf@illinois.edu
Dr. J. E. Velásquez, S. Mukherjee, Prof. W. A. van der Donk
University of Illinois at Urbana-Champaign
Howard Hughes Medical Institute and Department of Chemistry
600 South Mathews Avenue, Urbana, IL 61801 (USA)
E-mail: vddonk@illinois.edu

[**] This work was supported by the National Institutes of Health (GM PO1 GM077596 to W.W.M. and W.A.V.) NMR spectra were recorded on a 600 MHz instrument purchased with support from NIH Grant S10 RR028833.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201308363.

S2). The broad substrate specificity and robust activity of DhpI prompted the development of a method that relies on "stable isotope labeling of phosphonates in extract" (SILPE). The method uses the facilitated identification of pairs of compounds of different stable isotope composition by mass spectrometry, based on their well-defined mass difference, which is similar to the SILAC method in proteomics.[5] Accordingly, ISP4 media spiked with fosfomycin or fosmidomycin were treated with DhpI and a mixture of SAM and CD_3-SAM. The crude material was injected into an LCMS system and the eluent analyzed for compounds giving rise to two ions separated by 3.0188 Da, thus allowing the facile detection of methylated fosfomycin and fosmidomycin (Figure 1b; see also Figures S1–S3).

We then turned our attention to unknown phosphonates produced by *Streptomyces* sp. WM6372. This strain was previously identified as a potential phosphonate producer by screening for the presence of the phosphoenolpyruvate mutase gene (*pepM*), a characteristic marker for phosphonate biosynthesis.[3,6] Spent medium from this strain grown on ISP2 agar plates displayed two resonances in the phosphonate window of the ^{31}P NMR spectrum (Figure 2a). After several unsuccessful attempts to isolate the compounds using traditional purification methods, we attempted the SILPE method. The solid medium was freeze-thawed and compressed to release metabolites, evaporated to dryness, and taken up in 90% MeOH. Insoluble material was removed by centrifugation. The supernatant was concentrated and partially desalted by Sephadex LH20. The resulting crude mixture was treated with DhpI and a mixture of SAM and CD_3-SAM. The two original phosphonate peaks in the ^{31}P NMR spectrum disappeared and two new peaks were generated (Figure 2b; see also Table S1). Analysis by LCMS using the mass difference between compounds containing a CH_3 or a CD_3 group as marker allowed the purification of both molecules and determination of their masses. Characterization by ^{1}H and ^{13}C NMR spectroscopy, as well as ^{1}H-^{31}P-HMBC and mass spectrometry (Figure S4), allowed assignment of the parent compounds before DhpI-treatment as methyl phosphonoacetate (**1**) and methyl 2-hydroxy-2-phosphono-acetate (Me-HPnA (**2**); Figure 2c); these assignments were confirmed by chemical synthesis (Figure S5). A structure search revealed that these molecules are new phosphonate natural products; however, Me-HPnA is also a substructure in fosfazinomycins (Figure 2c). We therefore cultivated *Streptomyces* sp. WM6372 in a variety of media and monitored for fosfazinomycin production by LCMS analysis. In most media, production of fosfazinomycins was not observed; however, the strain produces small amounts of both fosfazinomycin A and B after growth in R2AS medium containing phosphonoacetate. Therefore, the organism has all the genes needed to produce fosfazinomycin, although it appears that unknown regulatory or metabolic limitations prevent its de novo synthesis under the conditions examined.

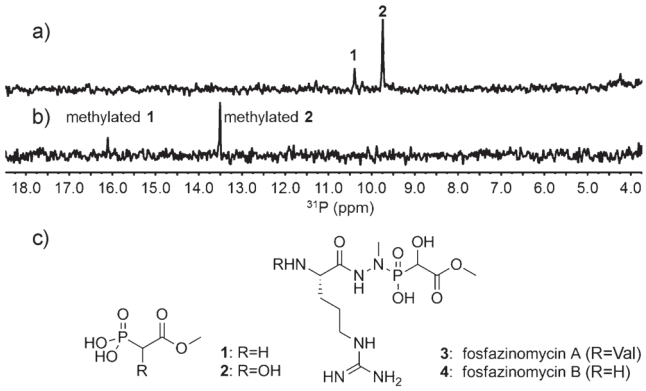

Figure 2. a) ^{31}P NMR spectrum of spent medium of *Streptomyces* sp. WM 6372. Only the chemical shift region of interest with respect to phosphonates is depicted. b) ^{31}P NMR spectrum after treatment with DhpI and SAM. c) Structures of the phosphonates giving rise to the signals in (a) and the structures of fosfazinomycin A and B.

We previously reported that *Streptomyces* sp WM6372 and *Streptomyces* sp. XY332 encode nearly identical phosphonate biosynthetic gene clusters (Figure S6).[6b] Therefore, we tested whether *Streptomyces* sp. XY332 could produce fosfazinomycins. Evaluation of various growth conditions and analysis by LCMS demonstrated that two compounds were produced with masses consistent with fosfazinomycin A and B. A ^{31}P NMR spectrum of the spent medium also revealed two small new signals with close chemical shifts at 13.7 and 13.5 ppm in addition to the peaks for compounds **1** and **2** (Figure 3a). To provide confirmation that the new peaks represented fosfazinomycin A and B, the compounds were partially purified. A ^{1}H-^{31}P HMBC NMR spectrum displayed correlations consistent with the fosfazinomycin core structure, and MS analysis showed molecular ions consistent with fosfazinomycin A and B (Figure S7). Finally, *Streptomyces* XY332 was grown on minimal medium containing $(^{15}NH_4)_2SO_4$ as the sole nitrogen source, and the ^{31}P NMR spectrum was recorded. The two peaks in the ^{31}P spectrum of

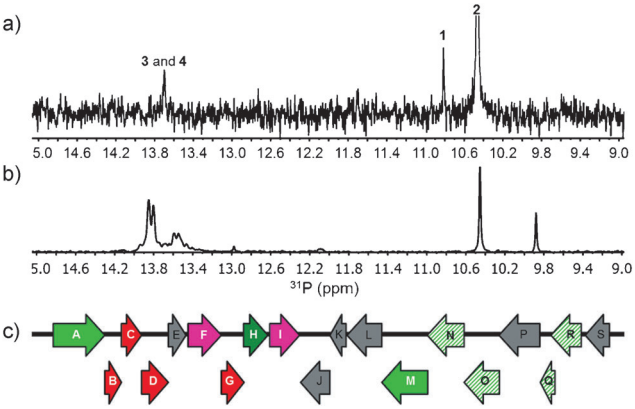

Figure 3. a) ^{31}P NMR spectrum of spent medium of *Streptomyces* sp. XY332. b) ^{31}P NMR spectrum of spent medium of the same strain when grown on $^{15}NH_4SO_4$. The chemical shift variation is the result of small changes in pH, as these compounds have pK_a values near 7. c) DNA fragment of the genome from *Streptomyces* sp. WM 6372 that contains the *pepM* gene; Genes are color-coded according to their putative function: formation of Me-HPnA (red), nitrogen methylation (dark green), amide ligation (pink), installation of the N–N bond (light green; the four genes shared with the clusters for the biosynthesis of diazo-containing natural products are striped in light green).

the spent medium now displayed a doublet splitting pattern with a coupling constant of 11.4 Hz (Figure 3b), and the high resolution mass spectrum indicated the presence of seven ^{15}N atoms, which is consistent with the assignment to fosfazinomycin A (Figure S8).

Two lines of evidence provide strong support to link the gene cluster shown in Figure 3c with the biosynthesis of fosfazinomycins. First, both organisms contain this gene cluster, and both can produce the antibiotic. Moreover, examination of draft genome sequences shows that these clusters are the only PEP mutase-encoding loci in these organisms; PEP mutase is involved in the biosyntheses of all phosphonates that are currently understood. Furthermore, we transferred the cluster to *S. lividans*, a non-phosphonate producer, and showed that the recombinant strain produced **1** and **2**. Thus, the transferred gene cluster clearly directs synthesis of the phosphonate moiety of fosfazinomycin. We suspect that, as with the parental strain, regulatory or metabolic limitations prevent de novo synthesis of the final product in *S. lividans*. Given that all characterized phosphonate biosynthetic pathways require PEP mutase, and that the gene cluster in question directs the synthesis of the phosphonate moiety of fosfazinomycin, we believe that it is highly likely, albeit not definitive in the absence of successful heterologous production, that these loci are responsible for fosfazinomycin synthesis.

Fosfazinomycin A and B were first isolated from *Streptomyces lavendofoliae* 630 as part of a screening program focused on antifungal antibiotics.[7] The two peptide congeners differ at their N termini (Figure 2c). Fosfazinomycins contain a unique hydrazide linkage between the carboxylic acid of Arg and the phosphonate of Me-HPnA. Nitrogen–nitrogen bonds are relatively rare in nature[8] and the mechanism of their formation is generally still poorly understood.[9] Based on bioinformatics analysis of the putative biosynthetic gene clusters of *Streptomyces* sp. WM6372 and *Streptomyces* sp. XY332 (Figure 3c and Table 1; see also Figure S6), a putative biosynthetic pathway for fosfazinomycin biosynthesis can be formulated (Figure 4). FzmC is the PEP mutase that installs the C–P bond in phosphonopyruate

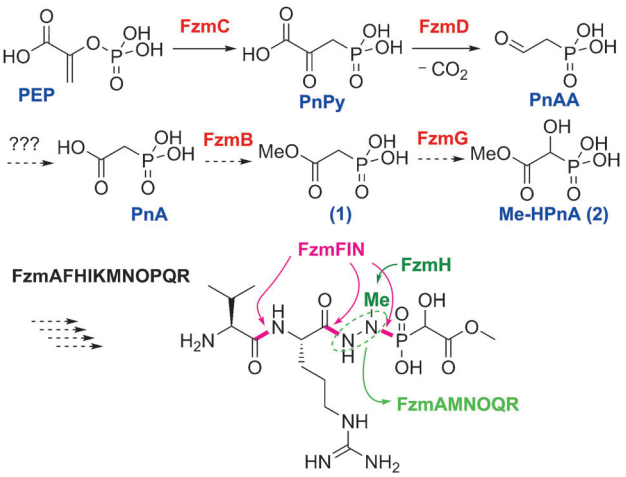

Figure 4. Proposed biosynthetic pathway of fosfazinomycin A. Putative steps are indicated with dashed arrows. FzmFIN are proposed to be involved in formation of the three magenta-colored bonds, whereas the FzmAMNOQR proteins are proposed to be involved in forming and installing the hydrazine core.

(PnPy) and FzmD decarboxylates PnPy to generate phosphonoacetaldehyde (PnAA). This intermediate is then proposed to be oxidized to phosphonoacetate (PnA) and methylated to generate **1**, one of the observed metabolites. The cluster does not contain a canonical aldehyde dehydrogenase, which suggests that a housekeeping enzyme may be used for the conversion of PnAA into PnA, but methylation of the carboxylate is likely catalyzed by the *S*-adenosylme-

Table 1: Summary of open reading frames (ORFs) in the genomic DNA of *Streptomyces* sp. WM 6372 that includes the fosfazinomycin biosynthetic gene cluster.

ORF	No. of residues	Highest identity homologue (as of September 1, 2013)	Amino acid identity/similarity [%]
FzmA	727	*Actinoplanes* sp. SE50/110 asparagine synthase (YP_006269188.1)	48/64
FzmB	233	*Haliangium ochraceum* DSM 14365 type 11 methyltransferase (YP_003269122.1)	48/63
FzmC	284	*Bacillus cereus* VD102 PEP phosphomutase (WP_000571606.1)	52/69
FzmD	378	*Micromonospora* sp. ATCC 39149 PnPy decarboxylase (WP_007071820.1)	52/64
FzmE	253	*Streptomyces ambofaciens* ATCC 23877 phosphodiesterase (CAJ89340.1)	55/66
FzmF	465	*Methanococcoides burtonii* DSM 6242 pyruvate carboxylase (YP_567027.1); ATP-GRASP family.	29/45
FzmG	317	*Streptomyces luridus* dioxygenase (ACZ13452.1)	45/55
FzmH	349	*Nonomuraea* sp. WU8817 N-methyl transferase (ACS83764.1)	35/46
FzmI	426	*Streptomyces rapamycinicus* N-acyltransferase (CAA60475.1)	31/43
FzmJ	411	*Corynebacterium matruchoti* MFS transporter (WP_005523355.1)	16/30
FzmK	231	*Desulfovibrio hydrothermalis* DSM 14728 thymidylate kinase (YP_007326724.1)	18/35
FzmL	492	*Streptomyces bingchenggensis* BCW-1 lyase (YP_004961426.1)	62/70
FzmM	655	*Streptomyces davawensis* JCM 4913 FAD(NAD)-dependent oxidoreductase (YP_007524522.1)	56/65
FzmN	519	uncultured bacterium BAC AB649/1850 glutamine synthetase (AEE65491.1)	56/67
FzmO	500	*Salinispora tropica* CNB-440 amidase (YP_001159034.1)	52/61
FzmP	575	*Staphylococcus epidermidis* hypothetical protein (WP_002477807.1)	24/41
FzmQ	134	*Streptomyces* sp. N-acetyltransferase (WP_008740585.1)	76/85
FzmR	435	*Streptomyces ambofaciens* ATCC 23877 adenylosuccinate lyase (CAI78075.1)	64/76
FzmS	341	*Streptomyces violaceusniger* Tu 4113 transcriptional regulator (YP_004814876.1)	45/57

thionine-dependent *O*-methyltransferase FzmB. FzmG is a likely candidate for the hydroxylation of **1** to produce **2**, because its closest homologue is the α-ketoglutarate-dependent oxygenase that hydroxylates 2-hydroxyethylphosphonate during dehydrophos biosynthesis.[10] To complete the biosynthesis of fosfazinomycin A, three bonds need to be made between nitrogen nucleophiles and acid electrophiles. Interestingly, the cluster encodes proteins from three different amidoligase families:[11] FzmF, a member of the ATP GRASP enzyme family,[11–12] FzmI, a member of the GCN5-related *N*-acetyltransferases (GNATs), which are enzymes that function as aminoacyl tRNA-dependent peptidyl transferases,[11,13] and FzmN, a member of the glutamine synthase family that is known to catalyze amide bond formation.[11,14] Further studies are required to determine which of these three enzymes makes which of the magenta-colored bonds in Figure 4.

The cluster also provides information on the possible construction of the hydrazide core. The methyl group on the hydrazide is probably installed by the *N*-methyltransferase FzmH, whereas one or both of the two nitrogen atoms may originate from asparagine/aspartate by the actions of the Asn synthase FzmA and the adenylosuccinate lyase FzmR; the latter delivers a nitrogen atom during de novo purine biosynthesis.[15] Formation of the N–N bond is likely facilitated by initial oxidation of an amine/amido group to the corresponding hydroxylamine/hydroxamate derivative by the flavin-dependent oxidoreductase FzmM, based on similar activation steps in the N–N bond forming processes in the biosynthesis of valanimycin[9a] and the kutznerides.[9b] Homologues of FzmNOQR (striped green, Figure 3c) are also found in the biosynthetic clusters of kinamycins[16] and lomaiviticin,[17] which contain N–N bonds in diazo groups. The conservation of the Gln synthetase FzmN suggests that the hydrazide group may be assembled on the side chain of glutamate by these enzymes, and that the amidase FzmO may release the hydrazide group. Similar use of glutamate as a molecular scaffold has been previously reported for various processes (Figure S9).[14]

In summary, we used the substrate tolerance of the phosphonate methyltransferase DhpI as a chemoselective tool to purify two phosphonate metabolites that are intermediates in the biosynthesis of fosfazinomycin. Bioinformatics analysis and heterologous production experiments provide support for the involvement of two nearly identical gene clusters in fosfazinomycin biosynthesis, thus providing insights into how the N–N bond might be fashioned and showing the unusual combination of an ATP-GRASP ligase, a tRNA-dependent peptidyl transferase, and a glutamine synthetase in the biosynthesis of a natural product.

Received: September 24, 2013
Revised: November 16, 2013
Published online: December 27, 2013

Keywords: antibiotics · biosynthesis · hydrazine · natural products

[1] a) H. B. Bode, R. Müller, *Angew. Chem.* **2005**, *117*, 6988 – 7007; *Angew. Chem. Int. Ed.* **2005**, *44*, 6828 – 6846; b) G. L. Challis, *J. Med. Chem.* **2008**, *51*, 2618 – 2628; c) J. E. Velásquez, W. A. van der Donk, *Curr. Opin. Chem. Biol.* **2011**, *15*, 11 – 21.
[2] D. J. Trader, E. E. Carlson, *Mol. Biosyst.* **2012**, *8*, 2484 – 2493.
[3] W. W. Metcalf, W. A. van der Donk, *Annu. Rev. Biochem.* **2009**, *78*, 65 – 94.
[4] J. H. Lee, B. Bae, M. Kuemin, B. T. Circello, W. W. Metcalf, S. K. Nair, W. A. van der Donk, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 17557 – 17562.
[5] S. E. Ong, B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey, M. Mann, *Mol. Cell. Proteomics* **2002**, *1*, 376 – 386.
[6] a) S. C. Peck, J. Gao, W. A. van der Donk, *Methods Enzymol.* **2012**, *516*, 101 – 123; b) X. Yu, J. R. Doroghazi, S. C. Janga, J. K. Zhang, B. T. Circello, B. M. Griffin, D. P. Labeda, W. W. Metcalf, *Proc. Natl. Acad. Sci. USA* **2013**, DOI: 10.1073/pnas.1315107110.
[7] a) Y. Kuroda, M. Okuhara, T. Goto, M. Okamoto, H. Terano, M. Kohsaka, H. Aoki, H. Imanaka, *J. Antibiot.* **1980**, *33*, 29 – 35; b) T. Ogita, S. Gunji, Y. Fukuzawa, A. Terahara, T. Kinoshita, H. Nagaki, T. Beppu, *Tetrahedron Lett.* **1983**, *24*, 2283 – 2286.
[8] a) L. M. Blair, J. Sperry, *J. Nat. Prod.* **2013**, *76*, 794 – 812; b) D. G. Fujimori, S. Hrvatin, C. S. Neumann, M. Strieker, M. A. Marahiel, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 16498 – 16503.
[9] a) R. P. Garg, L. B. Alemany, S. Moran, R. J. Parry, *J. Am. Chem. Soc.* **2009**, *131*, 9608 – 9609; b) C. S. Neumann, W. Jiang, J. R. Heemstra, Jr., E. A. Gontang, R. Kolter, C. T. Walsh, *ChemBioChem* **2012**, *13*, 972 – 976.
[10] B. T. Circello, A. C. Eliot, J. H. Lee, W. A. van der Donk, W. W. Metcalf, *Chem. Biol.* **2010**, *17*, 402 – 411.
[11] L. M. Iyer, S. Abhiman, A. M. Burroughs, L. Aravind, *Mol. Biosyst.* **2009**, *5*, 1636 – 1660.
[12] a) M. V. Fawaz, M. E. Topper, S. M. Firestine, *Bioorg. Chem.* **2011**, *40*, 185 – 191; b) M. Y. Galperin, E. V. Koonin, *Protein Sci.* **1997**, *6*, 2639 – 2643.
[13] a) R. P. Garg, J. M. Gonzalez, R. J. Parry, *J. Biol. Chem.* **2006**, *281*, 26785 – 26791; b) R. P. Garg, X. L. Qian, L. B. Alemany, S. Moran, R. J. Parry, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 6543 – 6547; c) W. Zhang, I. Ntai, N. L. Kelleher, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 12249 – 12253; d) D. J. Bougioukou, S. Mukherjee, W. A. van der Donk, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 10952 – 10957.
[14] a) S. I. de Azevedo Wäsch, J. R. van der Ploeg, T. Maire, A. Lebreton, A. Kiener, T. Leisinger, *Appl. Environ. Microbiol.* **2002**, *68*, 2368 – 2375; b) S. Kurihara, S. Oda, K. Kato, H. G. Kim, T. Koyanagi, H. Kumagai, H. Suzuki, *J. Biol. Chem.* **2005**, *280*, 4602 – 4608; c) M. Takeo, A. Ohara, S. Sakae, Y. Okamoto, C. Kitamura, D. I. Kato, S. Negoro, *J. Bacteriol.* **2013**, *195*, 4406 – 4414.
[15] S. Ratner in *The Enzymes*, *Vol. 7* (Ed.: P. Boyer), Academic Press, New York, **1972**, pp. 167 – 197.
[16] a) S. J. Gould, S. T. Hong, J. R. Carney, *J. Antibiot.* **1998**, *51*, 50 – 57; b) R. Bunet, L. Song, M. V. Mendes, C. Corre, L. Hotel, N. Rouhier, X. Framboisier, P. Leblond, G. L. Challis, B. Aigle, *J. Bacteriol.* **2011**, *193*, 1142 – 1153.
[17] R. D. Kersten, A. L. Lane, M. Nett, T. K. Richter, B. M. Duggan, P. C. Dorrestein, B. S. Moore, *ChemBioChem* **2013**, *14*, 955 – 962.

