

Utahmycins A and B, Azaquinones Produced by an Environmental DNA Clone

John D. Bauer,^{†,‡} Ryan W. King,[†] and Sean F. Brady^{*}

Laboratory of Genetically Encoded Small Molecules, Howard Hughes Medical Institute, The Rockefeller University, New York, New York 10065

Received December 4, 2009

Two new azaquinones, utahmycins A (**1**) and B (**2**), were isolated from cultures of *Streptomyces albus* J1704 transformed with the environmental DNA-derived Erd gene cluster. The structures of **1** and **2** were elucidated by spectroscopic analyses. The structure of **1** was confirmed by single-crystal X-ray diffraction analysis. Both metabolites appear to arise from the addition of a nitrogen atom to erdacin biosynthetic intermediates. Utahmycin A (**1**) is the first example of a biologically derived 1,3-dimethyl-2-azaanthraquinone.

Azaanthraquinones are reported to exhibit a wide range of biological activities, including antibacterial activity against drug-resistant *Staphylococcus aureus*, antimalarial activity, anti *Trypanosoma congolense* activity, and potential anticancer activity as a result of their DNA intercalation properties.^{1–9} Azaanthraquinones are rare in nature, having been isolated mainly from lichens and fungi as derivatives of 3-methyl-2-azaanthraquinones (Figure 1).^{2,10–17} As part of an ongoing effort to identify novel bacterial metabolites using culture-independent approaches, we are exploring environmental DNA (eDNA)-derived type II (iterative, aromatic) polyketide synthase (PKS) gene clusters for the ability to confer the production of new small molecules to laboratory-grown heterologous hosts.^{18–20} Here we report the isolation and characterization of utahmycin A (**1**) a 2-azaanthraquinone based on a new 1,3-dimethyl-2-azaanthraquinone skeleton, as well as utahmycin B (**2**), a closely related azaquinone derivative. Both utahmycins A and B are produced by *Streptomyces albus* transformed with the eDNA-derived Erd gene cluster.

DNA extracted from soil collected in Utah was used to construct a large cosmid-based eDNA library in *Escherichia coli*.^{18–20} Using a set of minimal PKS-specific degenerate primers, this library was screened for clones that contain type II PKS biosynthetic machinery. Minimal PKS-containing clones were then shuttled from *E. coli* into *Str. albus*, and the resulting recombinants were screened for the ability to produce clone-specific small molecules. Clone V167, which contains the eDNA-derived Erd gene cluster, was found to produce a large quantity of the novel metabolite erdacin (Figure 2).²⁰ Erdacin appears to be biosynthesized from the heterodimerization of two 13-carbon intermediates that arise from octaketide precursors.

While erdacin is the major metabolite produced by cultures transformed with the Erd gene cluster, a detailed TLC analysis of V167 culture broth extracts indicated the presence of small quantities of two additional clone-specific metabolites. LCMS analysis of EtOAc extracts obtained from cultures of V167 confirmed the presence of two clone-specific metabolites with odd masses (observed $[M + H]^+$ $m/z = 228$ and 254). The odd masses suggested that, unlike erdacin or any of its predicted biosynthetic intermediates, each of these additional metabolites must contain at least one nitrogen atom.

Ethyl acetate extracts derived from one liter cultures of V167 grown for 7–14 days in R2YE media were initially partitioned by silica gel flash chromatography using a $CHCl_3/MeOH$ step gradient.

^{*} To whom correspondence should be addressed. Tel: (212) 327-8280. Fax: (212) 327-8281. E-mail: sbrady@rockefeller.edu.

[†] These authors contributed equally to this work.

[‡] Current address: College of Pharmacy and Health Sciences, Mercer University.

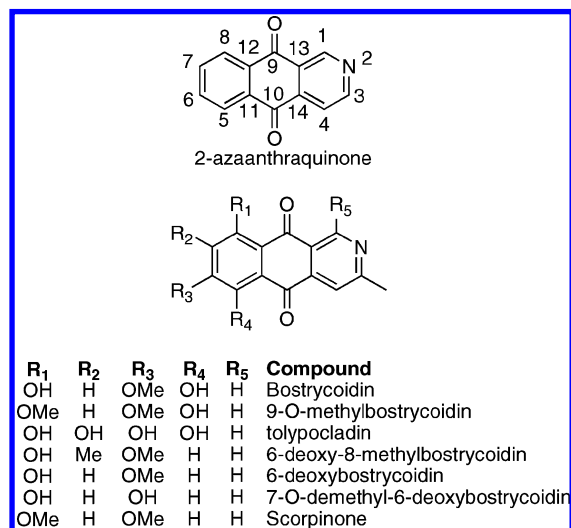


Figure 1. Known 2-azaanthraquinones from fungi and lichens.

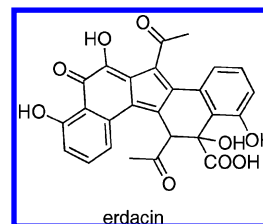


Figure 2. Erdacin is produced by heterologous expression of the eDNA-derived Erd gene cluster in *S. albus*.

The 99:1 $CHCl_3/MeOH$ fraction from this column was found to contain both of the new metabolites. Utahmycins A (observed $m/z = 254$) and B (observed $m/z = 228$) were purified from this column fraction by preparative reversed-phase HPLC.

The structure of each utahmycin was established using 1- and 2-D NMR and HRMS data. ¹³C chemical shift, ¹H–¹³C HMQC, and ¹H–¹³C HMBC data from utahmycin A confirmed the presence of two methyls, two carbonyls, and 11 olefinic carbons. Four of the olefin carbons are attached to protons, and the deshielding of three others suggested they were directly attached to heteroatoms.

The presence of the disubstituted phenol-like moiety in utahmycin A was inferred from the three-proton spin system seen in ¹H–¹H COSY experiments and the collection of ¹H–¹³C HMBC correlations from this spin system to six olefin carbons, including one that is highly deshielded (C-8, 163.1 ppm) (Figure 3a). An additional HMBC correlation from the proton doublet at one end of the three-

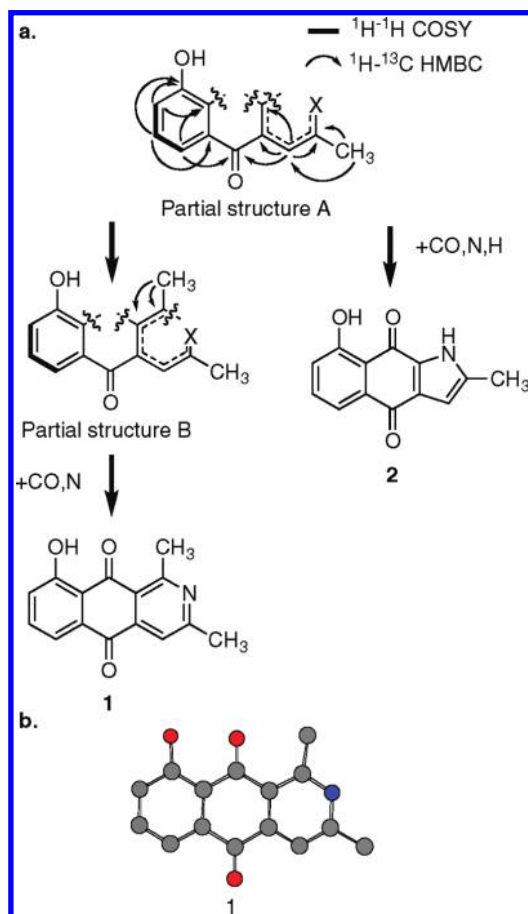


Figure 3. (a) $^1\text{H}-^1\text{H}$ COSY correlations and key long-range $^1\text{H}-^{13}\text{C}$ HMBC correlations used to establish the structures of utahmycins A (**1**) and B (**2**) are shown. (b) The structure of **1** was confirmed by X-ray crystallography (CCDC 734992). A computer-generated perspective drawing of **1** is shown (gray = carbon, red = oxygen, blue = nitrogen). Hydrogen atoms are not shown.

proton spin system to a carbonyl carbon (C-10, 183.1 ppm) allowed us to position one of the two carbonyls directly adjacent to this aromatic ring. The C-10 carbonyl is connected by an HMBC correlation from H-4 to the right side of partial structure A. The right side of partial structure A is defined by HMBC correlations from the H-16 methyl singlet to C-3 and C-4 and by HMBC correlations from the H-4 singlet to C-3, C-10, C-13, and C-14.

In the case of utahmycin A, partial structure A (Figure 3) can be expanded to partial structure B on the basis of the collection of HMBC correlations observed from the second methyl singlet (C-15). The H-15 methyl singlet shows HMBC correlations to both C-13 and C-1, the only remaining heteroatom-substituted olefin that had not yet been positioned in the structure. On the basis of the predicted molecular formula, the unsaturation index, and the ^{13}C chemical shift data, the remaining carbon, oxygen, and nitrogen that must still be incorporated into utahmycin A can be satisfactorily introduced into partial structure B only by closing the second and third rings with a carbonyl (C-10) and the nitrogen (N-2), respectively (Figure 3).

The final predicted structure for utahmycin A indicates that it is a 1,3-dimethyl-substituted 2-azaanthraquinone. 2-Azaanthraquinones are rarely seen in nature. Utahmycin A differs from other naturally occurring 2-azaanthraquinones by the presence of the C-1 methyl substituent and the absence of the C-7 oxidation (Figure 1). With no natural precedent for this type of 2-azaanthraquinone, we chose to verify the NMR structure prediction by X-ray crystallography. Utahmycin A was crystallized by slow evaporation from CH_2Cl_2 ,

and the predicted NMR structure was confirmed by single-crystal X-ray diffraction analysis (Figure 3).

^{13}C and $^1\text{H}-^{13}\text{C}$ HMBC data from utahmycin B indicate that all of the structural arguments used to generate partial structure A for utahmycin A also hold true for utahmycin B. However, the ^1H and ^{13}C spectra of utahmycin B contain neither the methyl singlet nor the additional olefin that was used to generate partial structure B. The only way to satisfy the predicted molecular formula, the unsaturation index, and the carbon chemical shift data for utahmycin B is to close the second and third rings with a carbonyl (C-8) and an amine (N-1), respectively. The resulting tricyclic structure differs from utahmycin A only by the presence of the five-membered nitrogen-containing heterocycle. While utahmycin B has been reported as a synthetic compound, to the best of our knowledge it has not been previously isolated from a biological system.²¹ Utahmycins A and B were assayed in agar disk diffusion assays against a panel of Gram-positive and Gram-negative bacteria. At the highest concentration tested (100 $\mu\text{g}/\text{disk}$) they did not exhibit antibiotic activity against any of the bacteria tested.

Utahmycin A is the first 1,3-dimethyl-2-azaanthraquinone structure to be characterized from nature. The origin of structural differences between known fungal 2-azaanthraquinones and the utahmycins can be rationalized on the basis of differences in fungal and *Streptomyces* polyketide folding paradigms as well as differences in the size of the predicted polyketide precursors.²² Known 2-azaanthraquinones are thought to arise from the fungal-specific ("F-type") folding of a heptaketide (Figure 4b),^{22,23} while the utahmycins are predicted to arise from the *Streptomyces*-specific ("S-type") folding of a decarboxylated octaketide intermediate (Figure 4a).²⁰

The incorporation of nitrogen into predicted erdacin biosynthetic intermediates would give rise to both utahmycins A and B (Figure 4a). In both this study and in fungal studies, 2-azaanthraquinones appear alongside biosynthetically related polyketides that do not contain nitrogens. In fungal studies, 2-azaanthraquinones are often found with fusarbuins or harbarin, structurally related metabolites that contain an oxygen atom in place of the nitrogen atom that appears in 2-azaanthraquinones. The ratio of 2-azaanthraquinones to fusarbuins differs depending on the culture conditions and is known to be sensitive to the nitrogen content of the media.^{4,15,24-27} For neither the fungal-derived 2-azaanthraquinones nor the utahmycins is the biosynthetic origin of the nitrogen clear. In the case of the utahmycins we know that these compounds appear only in cultures transformed with the Erd gene cluster. The absence of a predicted amino transferase in the Erd gene cluster suggests that the utahmycins likely arise from erdacin polyketide intermediates by a mechanism that is independent of the Erd gene cluster. Whether 2-azaanthraquinones are unintended shunt products of polyketide biosynthetic pathways or the result of evolutionarily relevant strategies for the generation of structurally diverse, and potentially biologically relevant, metabolites remains to be seen.

Our studies on the Erd gene cluster suggest that the heterologous expression of eDNA-derived gene clusters will likely provide access to both naturally occurring metabolites encoded by these previously inaccessible gene clusters and an orthogonal collection of compounds that arises from the interaction of these gene clusters with a host's native metabolic processes.

Experimental Section

General Experimental Procedures. Melting points were measured on a Stanford Research Systems MPA100 OptiMelt melting point apparatus. UV data were obtained using a Varian Cary 100-Bio UV-vis spectrophotometer. IR data were obtained on a Bruker Tensor-27 spectrometer. ^1H NMR and 2D NMR spectra were obtained on a Bruker Avance-600 spectrometer. ^{13}C NMR data were obtained on a Bruker Avance-800 spectrometer. Preparative HPLC was performed using an XBridge C18 (10 \times 150 mm, 5 μm) column with a 7 mL/min flow rate.

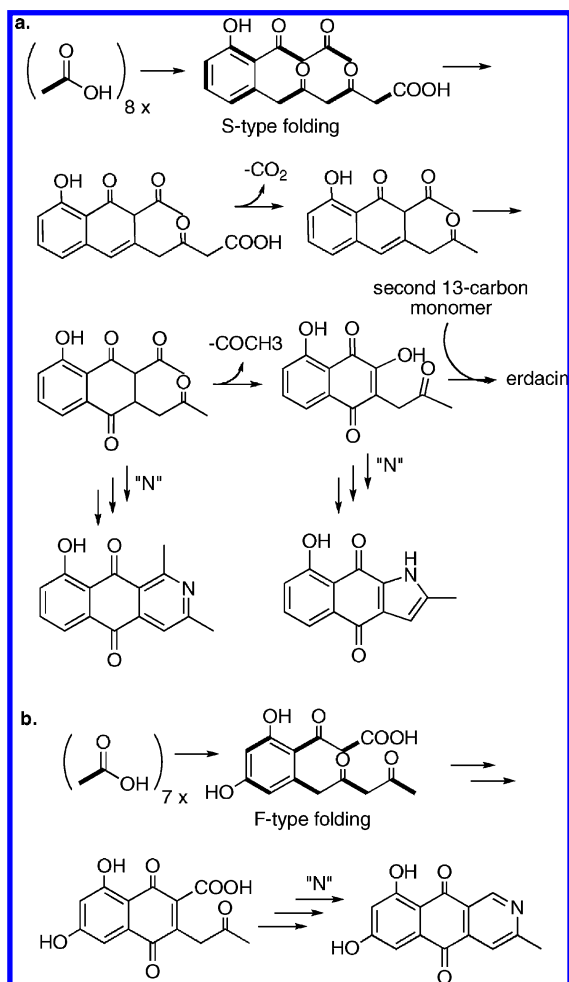


Figure 4. (a) The two halves of erdacin are predicted to arise from a common octaketide precursor that loses three carbons. Utahmycins A (**1**) and B (**2**) likely arise from the incorporation of nitrogen into two different erdacin intermediates. (b) While previous 2-azaanthraquinones appear to arise from the fungal (F)-type folding of a septaketide, utahmycins A and B appear to arise from a *Streptomyces* (S)-type folding of an octaketide.

Strains. *Streptomyces albus* J1074, a gift from Prof. Jose A. Salas Universidad de Oviedo, Spain, was used for heterologous expression studies.

Extraction and Isolation. A spore stock (75 μ L) of *S. albus* transformed with the retrofitted V167 clone²⁰ was inoculated into 1 L flasks of modified R2YE medium. Modified R2YE medium was prepared by adding 10 mL of KH_2PO_4 (0.5%), 80 mL of $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ (3.68%), 15 mL of L-proline (20%), 25 mL of TES buffer (5.73%, pH 7.2), 2 mL of R2 trace element solution, 5 mL of NaOH (1 M), and 50 mL of yeast extract (10%) to an 800 mL solution containing 103 g of sucrose, 250 mg of K_2SO_4 , 10.12 g of $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$, 10 g of glucose, and 100 mg of casamino acids. Following incubation at 30 °C with shaking (250 or 300 rpm) for 7–14 days, the pH of the culture broth was adjusted to 3 with concentrated HCl and then extracted with two volumes of EtOAc. The dried extracts were partitioned by silica gel flash chromatography using a $\text{CH}_3\text{Cl}/\text{MeOH}$ gradient in which 100–150 mL fractions were collected for each of the following $\text{CH}_3\text{Cl}/\text{MeOH}$ ratios: 100:0, 99:1, 97:3, 95:5, 90:10, 80:20, and 0:100. The fraction containing **1** and **2** (99:1, $\text{CH}_3\text{Cl}/\text{MeOH}$) was determined by analytical TLC (silica gel F₂₅₄) and further fractionated by reversed-phase HPLC. Preparative HPLC, with a steady flow of 20% CH_3CN for 2 min, followed by a 30 min gradient of 20–100% CH_3CN in H_2O (0.1% TFA), was used to isolate **1** and **2**. Fractions containing **1** (16.5 min, 1 mg/L) and **2** (12.0 min, 2.5 mg/L) were dried *in vacuo* and analyzed by HRMS and NMR.

Utahmycin A (1): yellow needles (CH_2Cl_2); mp 199–201 °C; UV (CH_2Cl_2) λ_{max} (log ϵ) 312 (0.72), 412 (1.32) nm; IR (thin film) ν_{max}

3358, 2918, 2850, 1737, 1675, 1633, 1581, 1465, 1380, 1322, 1265, 1231, 795 cm^{-1} ; ^1H NMR (CD_2Cl_2 , 600 MHz) δ 12.78 (1H, s, OH), 7.85 (1H, s, H-4), 7.77 (1H, dd, $J = 7.5, 1.1$ Hz, H-5), 7.69 (1H, t, $J = 7.9$ Hz, H-6), 7.35 (1H, dd, $J = 8.4, 1.1$ Hz, H-7), 3.05 (3H, s, H-15a-c), 2.70 (3H, s, H-16a-c); ^{13}C NMR (CD_2Cl_2 , 200 MHz) δ 190.3 (C-9), 183.1 (C-10), 165.2 (C-3), 163.1 (C-8), 162.1 (C-1), 136.9 (C-6), 133.1 (C-11), 125.8 (C-7), 122.8 (C-13), 119.5 (C-5), 117.9 (C-14), 117.8, (C-4), 117.2 (C-12), 26.8 (C-15), 25.2 (C-16); HRMS (ESI) m/z 254.0829 [$\text{M} + \text{H}$]⁺ (calcd for $\text{C}_{15}\text{H}_{12}\text{NO}_3$, 254.0817).

Utahmycin B (2): amorphous, yellow solid; UV (EtOH) λ_{max} (log ϵ) 321 (0.79), 438 (1.53) nm; IR (thin film) ν_{max} 3584, 3264, 2924, 1675, 1624, 1556, 1483, 1451, 1350, 1329, 1263, 1239, 753 cm^{-1} ; ^1H NMR (acetone- d_6 , 600 MHz) δ 12.45 (1H, s, OH), 11.77 (1H, br s, NH), 7.62 (2H, m, H-4 and H-5), 7.18 (1H, dd, $J = 6.6, 2.8$ Hz, H-6), 6.49 (1H, s, H-3), 2.44 (3H, s, H-14a-c); ^{13}C NMR (acetone- d_6 , 200 MHz) δ 180.8 (C-8), 180.5 (C-9), 163.0 (C-7), 140.9 (C-2), 136.5 (C-5), 135.5 (C-10), 131.9, 130.0 (C-12, 13), 124.5 (C-6), 119.7 (C-4), 116.5 (C-11), 108.2 (C-3), 13.1 (C-14); HRMS (ESI) m/z 228.0672 [$\text{M} + \text{H}$]⁺ (calcd for $\text{C}_{13}\text{H}_{10}\text{NO}_3$, 228.0661).

X-ray Crystallographic Analysis of Utahmycin A (1). A $0.8 \times 0.1 \times 0.025$ mm³ crystal of **1** was transferred from the crystallization vessel into a drop of inert oil, then mounted on a Bruker X8 APEX II diffractometer using a nylon loop, and cooled to -100 °C. X-ray data were collected using Cu K α radiation and processed, and a structure was obtained with Bruker APEX2 and SAINT+ software packages. Crystals are triclinic, space group *P*1, with $a = 7.268(2)$ Å, $b = 7.996(2)$ Å, $c = 10.8965(16)$ Å, $\alpha = 72.201(2)^\circ$, $\beta = 72.244(11)^\circ$, $\gamma = 76.278(15)^\circ$, $V = 566.9(2)$ Å³. Overall 5459 reflections were collected; 3090 of these were symmetry independent and 1529 were considered “strong” ($F_o > 4\sigma F_o$). An empirical absorption correction was applied with SADABS. The structure was solved by direct methods and refined on F^2 by full matrix least-squares techniques using the SHELXTL software package (final R1 = 4.15%). All non-hydrogen atoms were refined anisotropically. Methyl group hydrogen atom positions were calculated geometrically; all other hydrogen atoms were found in a difference Fourier map and refined isotropically. Data have been deposited with the Cambridge Crystallographic Data Centre under the organic and organometallic compounds as entry CCDC 734992. Copies of the data can be obtained, free of charge, on application to the Director, CCDC, 12 Union Road, Cambridge CB2 1EZ, UK (fax: +44-(0)1223-336033 or e-mail: deposit@ccdc.cam.ac.uk).

Acknowledgment. This work was supported by NIH GM077516, Howard Hughes Medical Institute Early Career Scientist, The Beckman Foundation, The Searle Foundation, and The Hershel Foundation. We thank E. Lobkovsky at Cornell University for his assistance with X-ray crystallography.

Supporting Information Available: ^1H , ^{13}C , HMQC, and HMBC NMR spectra for all compounds. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References and Notes

- (1) Koyama, J.; Morita, I.; Kobayashi, N.; Osakai, T.; Usuki, Y.; Taniguchi, M. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 1079–1082.
- (2) Nok, A. J. *Cell Biochem. Function* **2002**, *20*, 205–212.
- (3) Baker, R. A.; Tatum, J. H.; Nemeč, S., Jr. *Mycopathologia* **1990**, *111*, 9–15.
- (4) Medentsev, A.; Akimenko, V. *Biokhimiya* **1988**, *52*, 289–296.
- (5) Medentsev, A. G.; Maslov, A. N.; Akimenko, V. K. *Biokhimiya* **1990**, *55*, 1766–1772.
- (6) Grafe, U.; Romer, W. *Pharmazie* **1991**, *46*, 297–298.
- (7) Okunade, A. L.; Clark, A. M.; Hufford, C. D.; Oguntimein, B. O. *Planta Med.* **1999**, *65*, 447–448.
- (8) Krachko, A. P.; Waterhouse, D. J. *Heterocycles* **1999**, *51*, 737–749.
- (9) Khanapure, S. P.; Biehl, E. R. *Heterocycles* **1988**, *27*, 2643–2650.
- (10) Arsenault, G. P. *Tetrahedron Lett.* **1965**, *45*, 4033–4037.
- (11) Hamilton, M. A.; Knorr, M. S.; Cajori, F. A. *Antibiot. Chemother.* **1953**, *3*, 853–856.
- (12) Miljkovic, A.; Mantle, P. G.; Williams, D. J.; Rassing, B. J. *Nat. Prod.* **2001**, *64*, 1251–1253.
- (13) Moriyasu, Y.; Miyagawa, H.; Hamada, N.; Miyawaki, H.; Ueno, T. *Phytochemistry* **2001**, *58*, 239–241.
- (14) Parisot, D.; Devys, M.; Barbier, M. Z. *Naturforsch., B: Chem. Sci.* **1989**, *44*, 1473–1474.
- (15) Parisot, D.; Devys, M.; Barbier, M. *Microbios* **1990**, *64*, 31–47.

- (16) Gräfe, U.; Ihn, W.; Tresselt, D.; Miosga, N.; Kaden, U.; Schlegel, B.; Bormann, E. J.; Sedmera, P.; Novak, J. *Biol. Met.* **1990**, *3*, 39–44.
- (17) Zhang, W.; Watanabe, K.; Wang, C. C. C.; Tang, Y. *J. Nat. Prod.* **2006**, *69*, 1633–1636.
- (18) Brady, S. F. *Nat. Protocol.* **2007**, *2*, 1297–1305.
- (19) Banik, J. J.; Brady, S. F. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 17273–17277.
- (20) King, R. W.; Bauer, J. D.; Brady, S. F. *Angew. Chem., Int. Ed.* **2009**, *48*, 6257–6261.
- (21) Sakai, M.; Abe, M.; Takagaki, H.; Ogawa, S.; Yamazaki, K.; Ikegawa, T. *Jpn. Kokai Tokkyo Koho JP 04-211650*, 1992.
- (22) Thomas, R. *ChemBioChem* **2001**, *2*, 612–627.
- (23) Van Wagoner, R. M.; Mantle, P. G.; Wright, J. L. *J. Nat. Prod.* **2008**, *71*, 426–430.
- (24) Medentsev, A. G.; Arinbasarova, A. Y.; Akimenko, V. K. *Appl. Biochem. Microbiol.* **2005**, *41*, 503–507.
- (25) Kurobane, I.; Vining, L. C.; McInnes, A. G.; Gerber, N. N. *J. Antibiot.* **1980**, *33*, 1376–1379.
- (26) Kurobane, I.; Zaita, N.; Fukuda, A. *J. Antibiot.* **1986**, *39*, 205–214.
- (27) Medentsev, A. G.; Akimenko, V. K. *Mikrobiologiya* **1992**, *61*, 824–829.

NP900786S