Terms of Use

Sporulenes, Heptaprenyl Metabolites from *Bacillus subtilis* Spores

Renee Kontnik,[†] Tanja Bosak,*[,‡,§,||] Rebecca A. Butcher,[†] Jochen J. Brocks,[⊥] Richard Losick,[||] Jon Clardy,*[,†] and Ann Pearson[§]

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts 02138, Department of Molecular and Cell Biology, Harvard University, Cambridge, Massachusetts 02138, and Research School of Earth Sciences, The Australian National University, Canberra ACT 0200, Australia

tbosak@mit.edu; jon_clardy@hms.harvard.edu

Received June 10, 2008

ABSTRACT
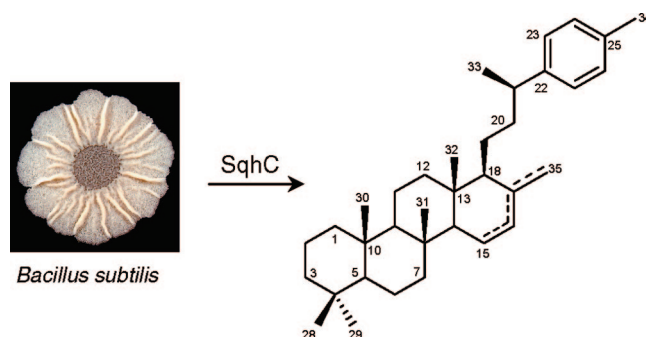
Sporulene, a C_{35}-terpenoid hydrocarbon with an unusual pentacyclic structure, is produced by *Bacillus subtilis* during sporulation.

All bacterial genomes reveal a common lesson: our knowledge of the small molecules made by bacteria, even very well-studied ones, is woefully incomplete, and numerous small molecules that are likely to carry out important biological functions have been overlooked. For example, *Bacillus subtilis* and its relatives are some of the best studied bacteria on the planet, but a recent comparative analysis of their genomes revealed that a giant polyketide synthase cluster produced bacillaene,[1] a complex polyunsaturated antibiotic, that was subsequently isolated and characterized.[2] In a similar vein, a very recent study on *B. subtilis* drew

attention to a putative squalene cyclase gene, *sqhC*, that makes heptaprenyl lipids named sporulenes that increase the resistance of spores to reactive oxygen species.[3] In this letter, we report the characterization of these three pentacyclic terpenoids, sporulenes A, B, and C (**1**–**3**, Scheme 1) from *B. subtilis* spores.

To identify the product of the SqhC cyclase enzyme in *B. subtilis*, lipids were extracted from spores of a wild-type strain (PY79) and a mutant strain lacking *sqhC* (strain TB10). Comparison of the lipid profiles by GC-MS demonstrated that three isoprenoid lipids were present in wild-type spores and absent in *sqhC*-deficient spores (Supporting Information).[3] These lipids were extracted from a large-scale spore
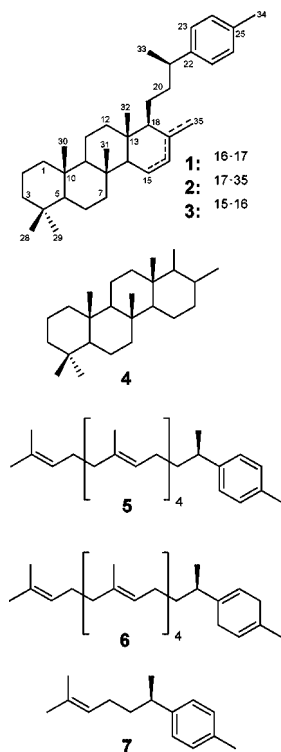
[†] Harvard Medical School.
[‡] Massachusetts Institute of Technology.
[§] Department of Earth and Planetary Sciences, Harvard University.
[||] Department of Molecular and Cell Biology, Harvard University.
[⊥] The Australian National University.
(1) Chen, X. H.; Vater, J.; Piel, J.; Franke, P.; Scholz, R.; Schneider, K.; Koumoutsi, A.; Hitzeroth, G.; Grammel, N.; Strittmatter, A. W.; Gottschalk, G.; Sussmuth, R. D.; Borriss, R. *J. Bacteriol.* **2006**, *188*, 4024.

(2) Butcher, R. A.; Schroeder, F. C.; Fischbach, M. A.; Straight, P. D.; Kolter, R.; Walsh, C. T.; Clardy, J. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1506.

(3) Bosak, T.; Losick, R. M.; Pearson, A. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 6725.

Scheme 1

no.	δ_C, mult.	δ_H	COSY	HMBC
1	40.1, CH_2	0.81, 1.70	2, 30	9
2	22.9, CH_2	1.30, 1.68	1	
3	42.3, CH_2	1.16, 1.38	28	1
4	33.2, qC			
5	56.7, CH	0.81	6	
6	22.7, CH_2	1.32, 1.67	5	5
7	40.0, CH_2	0.81, 1.68	9	6, 8, 31
8	37.6, qC			
9	61.3, CH	0.98	7	10
10	34.6, qC			
11	19.6, CH_2	1.37, 1.52		12, 10
12	40.6, CH_2	1.08, 1.81		11
13	40.5, qC			
14	61.0, CH	1.04		
15	22.9, CH_2	1.30, 1.66	16	14
16	38.6, CH_2	1.93, 2.34	15,35	15, 17, 35
17	149.0, qC			
18	57.7, CH	1.52	35	17
19	22.5, CH_2	1.27		18
20	34.3, CH_2	1.52	21	18, 19, 21
21	40.3, CH	2.62	20,33	23, 27, 33
22	135.1, qC			
23	127.1, CH	7.05		21, 25, 27
24	129.4, CH	7.08	34	26, 34
25	135.5, qC			
26	129.4, CH	7.08	34	24, 34
27	127.1, CH	7.05		21, 23, 25
28	21.2, CH_3	0.81	3	3, 4, 5, 29
29	33.3, CH_3	0.84		4, 5, 28
30	15.5, CH_3	0.78	1	10
31	16.2, CH_3	0.87		8, 9
32	15.2, CH_3	0.68		13, 14, 18
33	22.4, CH_3	1.67	21	
34	20.8, CH_3	2.30	24,26	24, 25, 26
35	105.8, CH_2	4.47, 4.77	16,18	16, 18

preparation derived from a *B. subtilis* strain overexpressing SqhC (strain TB28). All three compounds (**1**−**3**) had identical molecular ion peaks at *m/z* 474. HRMS yielded a common molecular mass of 474.420, indicating molecular formulas of $C_{35}H_{54}$ (calculated *m/z* 474.4226). Complete hydrogenation of sporulenes on Pt/C indicated four double bonds, signifying a pentacyclic structure. MS fragmentation patterns show features characteristic of the tetracyclic scalarane system (**4**), a sesterpene skeleton previously observed in metabolites isolated from marine sponges.[4] Additional fragments with *m/z* 119, 132, and 145 in all compounds are consistent with the presence of ethylene- ($C_9H_{11}^+$), propylene- ($C_{10}H_{12}^+$), and butylene-methylphenyl ($C_{11}H_{13}^+$) fragments, respectively, and indicate an aromatic side-chain moiety (Figure 1a).

The structures were further investigated by 1D- and 2D-NMR studies including homonuclear $^1H−^1H$ (COSY, ROESY, and TOCSY) and heteronuclear $^1H−^{13}C$ (HMQC and HMBC) correlation experiments. Due to consistent coelution of the three sporulenes, it was easier to perform NMR analysis on a mixture. As a result, the density of upfield aliphatic signals in the ^1H NMR spectrum made assignments of peak multiplicity and coupling constants impossible for some of the overlapped signals. However, careful analysis of $^1H−^{13}C$ correlation experiments allowed the complete spectral assignment for all proton and carbon atoms in the sporulenes (Table 1).

HMBC and COSY correlations defined the *para*-oriented methyl group on the aromatic ring, along with the *sec*-butyl

linkage between this phenyl group and C-18 of the scalarane system (Figure 1b). Remaining correlations were consistent with the scalarane backbone predicted by mass spectral fragmentation patterns. Multiple vinylic proton signals in the ^1H NMR spectrum are present in three distinct spin systems, corresponding to the differing locations of the C-ring double bond among the three sporulenes ($\Delta^{16−17}$, $\Delta^{17−35}$, $\Delta^{15−16}$, see **1**−**3**). TOCSY, COSY, and HMBC experiments were used to elucidate the positions of the double bonds in these compounds (Figure 1b and c; Supporting Information).

Relative stereochemistry of the sporulenes was determined by analysis of ROESY correlations. All of the cyclohexane rings have chair conformations, and the methyl substituents at quarternary centers are all axial (Figure 1d). Stereochemistry at the benzylic position (C-21) could not be determined through ROESY correlations alone but was established based on the stereochemistry of linear analogues which were also isolated, as described below.

At least two minor peaks with molecular ions at *m/z* 476 coeluted with compounds **1**−**3**, which suggested a set of reduced analogues of the three major sporulenes (Supporting Information). These minor components are also discernible

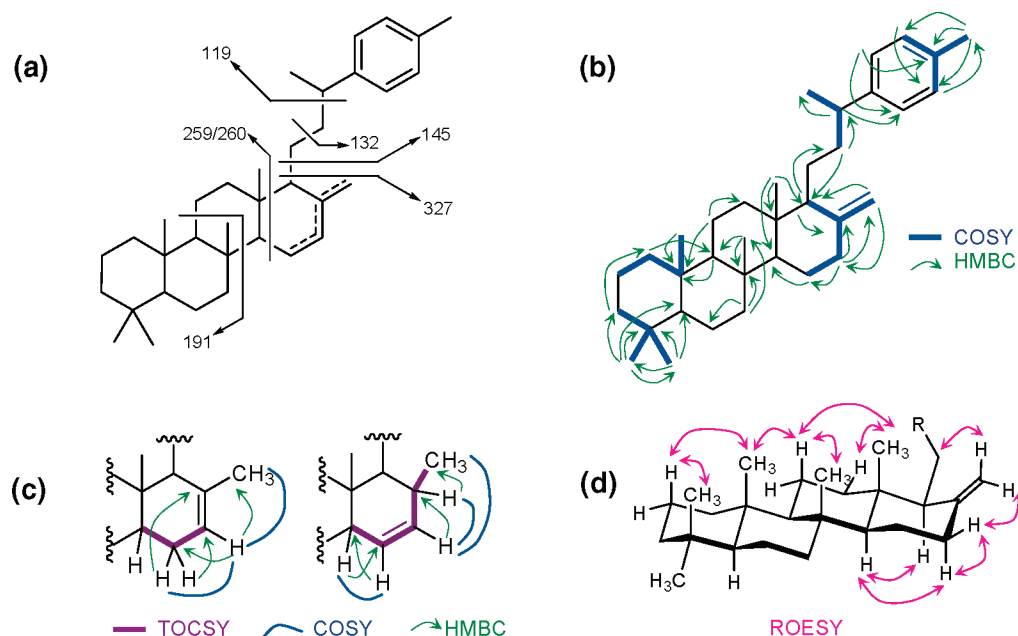(4) Renoux, J. M.; Rohmer, M. *Eur. J. Biochem.* **1986**, *155*, 125.

Figure 1. Structure elucidation of the sporulenes. (a) Fragmentation scheme that accounts for major fragments observed in the mass spectra of the sporulenes. (b) Carbon sequence of sporulene B as established from HMBC and COSY correlations. (c) Key TOCSY, COSY, and HMBC correlations used in determination of double bond locations in sporulenes A and C. (d) Key ROESY correlations used in determination of relative configuration of the sporulenes.

upon close analysis of the vinylic proton signals in the HMBC, COSY, and TOCSY spectra. The 2 Da increase in mass was found to correspond to a reduction of the aromatic ring. While the resulting cyclohexadiene ring can be detected in the NMR spectra, long-range correlations could not determine whether this reduced analogue exists for all three sporulene structures. Analysis of the mass chromatogram, however, reveals at least two reduced forms (Supporting Information).

Two additional lipids with molecular ions m/z 474 and 476 were recognized in the extracts from both wild-type and *sqhC*-lacking spores, and these were identified as the tetraprenyl curcumenes **5** and **6** originally described by Böröczky et al.[5] (Supporting Information). These regular polyprenoids are likely the linear precursors to the polycyclic sporulenes—substrates of SqhC. Stereochemistry at the benzylic carbon of **5** and **6** was previously established by oxidative degradation,[5] and this orientation is assumed to be the same in the *B. subtilis* metabolites. The stereochemistry shown in structures **1**–**3** thus reflects the relative stereochemistry for the tetracyclic core defined by NMR coupled with plausible assumptions about the absolute configurations of the core and C-21.

The sporulene structures raise questions about both the biosynthesis of these unusual C_{35}-pentacyclic hydrocarbons and the relation of these structures to their biological function. The sporulenes most plausibly originate from a linear C_{35} heptaprenyl unit with all head-to-tail isoprene linkages, distinguishing them from the hopanoids and steroids, which are derived from squalene. While heptaprenyl-derived molecules are not common, *B. subtilis* possesses heptaprenyl−diphosphate synthase genes (*hepS* and *hepT*)[6] that could generate a linear heptaprenyl unit. The linear heptaprenyl diphosphate must then undergo two cyclizations: one to form what ultimately becomes the phenyl group and another to form the tetracyclic unit. In light of the products found, formation of the cyclohexadiene fragment probably occurs first, followed by formation of the tetracyclic fragment. While not common, many aromatic terpenes are known, including simple sesquiterpenes such as curcumene (**7**). The biosynthesis of the phenyl ring in curcumene and similar terpenes, however, is not presently understood. The formation of a cyclohexadiene requires only a straightforward cyclization (Figure 2a), but the mechanism of its oxidative conversion to an aromatic compound is unclear. A similar cyclization of the heptaprenyl diphosphate precursor would yield **6**, and oxidation of **6** would give **5**.

The SqhC-catalyzed cyclization of **5** and/or **6** would generate the polycyclic skeleton of the sporulenes if they were folded in the typical all prechair conformation typical of squalene−hopane cyclases (Figure 2b). If the substrate was **5** or **6**, the cyclization cascade would end with a C-17 carbocation after forming the tetracyclic core. The C-17 carbocation could eliminate a proton to give **1** and **2** or undergo rearrangement prior to proton elimination to give **3**.

The biological role(s) of the sporulenes is related to sporulation since *B. subtilis* produces them only during

(5) Böröczky, K.; Laatsch, H.; Wagner-Dobler, I.; Stritzke, K.; Schulz, S. *Chem. Biodivers.* **2006**, *3*, 622.

(6) (a) Zhang, Y. W.; Koyama, T.; Marecak, D. M.; Prestwich, G. D.; Maki, Y.; Ogura, K. *Biochemistry* **1998**, *37*, 13411. (b) Zhang, Y. W.; Li, X. Y.; Sugawara, H.; Koyama, T. *Biochemistry* **1999**, *38*, 14638.
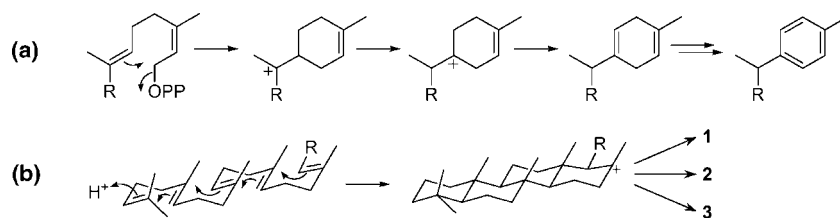
Figure 2. Proposed scheme for the biosynthesis of the sporulenes. (a) Cyclization of a heptaprenyl precursor yields a cyclohexadiene that is further oxidized to an aromatic moiety. (b) Mechanism of SqhC-catalyzed cyclization of **5** and/or **6**, folded in an all prechair conformation, to yield the polycyclic scalarane skeleton of the sporulenes.

sporulation, and their abundance in wild-type *B. subtilis* (0.1 mg/g dry spores) suggests a physical or stoichiometric mechanism.[3] There is a clear association between sporulene production and relief of oxidative stress, as wild type spores of *B. subtilis* resist oxidative stress more successfully than spores of a mutant strain that cannot produce sporulenes. The mutant strain grows and sporulates equally well as the wild type in the absence of aggressive oxidants such as hydrogen peroxide.[3] In a possibly related case, when *Streptomyces coelicolor* sporulates on solid medium, hopanoids which are believed to protect against dehydration are produced.[7]

Sporulenes could protect *B. subtilis* spores against oxidative stress in two ways: one physical and the other chemical. Physical protection could come from a protective barrier made by packing sporulenes into cellular membranes in a manner that functionally resembles the bacteriohopanetetrol phenylacetate monoesters found in the nitrogen-fixing Gram-positive root symbiont *Frankia* that prevent the diffusion of oxygen into the *Frankia*-containing vesicles.[8] The reduced sporulenes could alternatively provide a chemical barrier as oxygen scavengers through the oxidative conversion of their cyclohexadiene to a phenyl moiety. In this scenario, compounds **1**−**3** and **5** would be the oxidized remnants of the biologically relevant compounds.

The identification of SqhC and its hydrocarbon products provides a basis for searching for other cyclases in microbial genomes and for their small molecule products in a variety of milieus—including sediments and sedimentary rocks as clues to the redox conditions of ancient environments. Knowledge of the sporulene structural array coupled with organic synthesis will provide a clearer picture of the biological processes related to these unusual metabolites.

Supporting Information Available: Experimental procedures and full spectroscopic data for all new compounds. This material is available free of charge via the Internet at http://pubs.acs.org.

OL801314K

(7) Poralla, K.; Muth, G.; Hartner, T. *FEMS Microbiol. Lett.* **2000**, *189*, 93.

(8) Berry, A. M.; Harriott, O. T.; Moreau, R. A.; Osman, S. F.; Benson, D. R.; Jones, A. D. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 6091.