

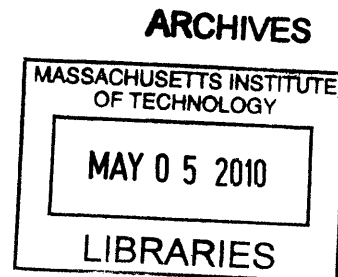
Diversity of polycyclic triterpenoids in *Rhodospirillum rubrum*

by

Katherine Harris

B.S. Biology (2005)
University of North Carolina at Chapel Hill

M.Phil. History of Science (2008)
University of Oxford



SUBMITTED TO THE DEPARTMENT OF
EARTH, ATMOSPHERIC AND PLANETARY SCIENCES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE IN EARTH AND PLANETARY SCIENCES

AT THE

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2010

© 2010 Massachusetts Institute of Technology

All rights reserved

Signature of Author: _____
Department of Earth, Atmospheric and Planetary Sciences
January 15, 2010

Certified by: _____
Tanja Bosak
Cecil and Ida Green Assistant Professor of Geobiology
Thesis Supervisor

Accepted by: _____
Maria T. Zuber
E.A. Griswold Professor of Geophysics
Head, Department of Earth, Atmospheric and Planetary Sciences

Diversity of polycyclic triterpenoids in *Rhodospirillum rubrum*

by

Katherine Harris

Submitted to the Department of Earth, Atmospheric and Planetary Sciences
on January 15, 2010 in Partial Fulfillment of the
Requirements for the Degree of
Master of Science in Earth and Planetary Sciences

ABSTRACT

Sedimentary rocks of all ages abound with geostable lipids of microbial origin, but many biomarkers lack known organismal sources and clear environmental contexts. Here we used *Rhodospirillum rubrum*, a metabolically versatile, genetically tractable α -Proteobacterium, to explore the diversity of its non-polar terpenoids as a function of growth condition and growth phase. We analyzed the non-polar fraction of lipids extracted from *R. rubrum* grown under aerobic, anaerobic, heterotrophic and phototrophic conditions and detected a variety of bicyclic, tricyclic, tetracyclic and pentacyclic triterpenoids, derived from the enzymatic cyclization of squalene and produced in amounts comparable to diploptene. Identified compounds included bicyclic polypodatetraenes, malabaricatriene, euphadiene, adianane, and fernene. Prior to this work, malabaricatriene was an “orphan” biomarker suspected to have a microbial origin, yet it lacked a proven source. We observed similar patterns of polycyclic terpenoids in other hopanoid-producing α -proteobacteria, including *Zymomonas mobilis*, *Rhodopseudomonas palustris*, and *Rhodomicrobium vanniellii*. The presence and relative abundance of polycyclic triterpenoids in *R. rubrum* varied with the growth stage (exponential versus early stationary phase) and growth condition (photoheterotrophic versus photoautotrophic growth). Since *R. rubrum*'s genome contains a single squalene-hopene cyclase gene, the array of triterpenoids produced by it and other α -proteobacteria likely evolves from this enzyme performing low-fidelity cyclization. The observed diversity of sedimentary triterpenoids might therefore result from a select few squalene-hopene cyclase enzymes operating with varying specificity under a range of physiological and environmental conditions, rather than reflecting a great diversity of squalene-hopene cyclases.

Thesis supervisor: Tanja Bosak

Title: Cecil and Ida Green Assistant Professor of Geobiology

INTRODUCTION

The ancient history of life on earth is archived in molecular fossils stored in the sedimentary rocks. Certain molecular fossils are geostable derivatives of ringed lipids called polycyclic terpenoids, a class of compounds that includes tetracyclic sterols, produced mainly by eukaryotes, and pentacyclic hopanoids, produced mainly by bacteria. Because the hydrocarbon skeletons of polycyclic terpenoids resist degradation during the physical, biological, and geochemical processes of diagenesis associated with sedimentary burial up to temperatures of 150°C (Mackenzie, 1984), a vast reservoir of hopanoids has accumulated in rocks and oil deposits. Some estimates of total mass of hopanoids ($\sim 10^{12}$ tons) approach the total mass of carbon contained in all modern living organisms (Ourisson and Albrecht, 1992).

Assuming that the same lipid biosynthetic pathways have operated consistently over time, lipid biomarkers can serve as tools for exploring the coevolution of life and earth, from tracing the evolution of organic matter in sediments (Moldowan et al., 1994) to timing the oxygenation of the early earth (Brocks et al., 2003; Brocks et al., 1999). However, the diagnostic potential of lipid biomarkers depends on understanding the physiological and environmental factors responsible for their synthesis, a task further complicated by the fact that many sedimentary polycyclic terpenoids lack established microbial sources. Cheilanthanes and malabaricatrienes are examples of “orphan” biomarkers abundant in sediments, aquatic environments, and the rock record (Behrens et al., 2000; Behrens et al., 1999; Nytoft and Larsen, 2001; Wang and Simoneit, 1995). Fernene is an example of a biomarker first detected in ferns, and only later discovered in *Rhodomicrobium vanniellii* and *Zymomonas mobilis* (Douka et al., 2001; Howard et al., 1984), as well as lake and marine sediments of varying ages (Brassell and Eglinton, 1983; Volkman et al., 1986).

Identifying microbial parents for orphan biomarker lipids relies as much on fortuity as on systematic analysis, but genetic and genomic approaches can inform the search for candidate microorganisms. The presence of squalene-hopene cyclases (SHC), enzymes that cyclize the linear polyprene squalene into polycyclic terpenoids, within sequenced microbial genomes informs biochemical quests for unique microbial polycyclic terpenoids [e.g., (Blumenberg et al., 2006; Bosak et al., 2008; Fischer et al., 2005; Kontnik et al., 2008)]. Culture-independent

approaches using PCR probes that target conserved SHC domains have demonstrated the presence and diversity of these enzymes in freshwater lakes and the open ocean (Pearson et al., 2007). Phylogenetic analysis reveals a trend toward α -proteobacteria among SHC clones, particularly in marine environments (Pearson et al., 2009). However, the vast majority microbes with SHC genes in the natural world remain uncultured, and so the polycyclic terpenoids associated with them remain uncharacterized.

Even in cultured microbes, little is known about the diversity and function of polycyclic terpenoids. Hopanoids in certain prokaryotes play vital structural roles in membranes, analogous to the roles of sterols in eukaryotic membranes (Horbach et al., 1991; Poralla et al., 1984). Yet, the sporadic phylogenetic distribution and condition-specific expression of squalene-hopene cyclase (SHC) genes in bacteria hints at more functionalized roles for these compounds. In *Rhodopseudomonas palustris*, for example, hopanoids play a role in regulating membrane integrity and pH homeostasis (Welander et al., 2009). The ratio of methylated bacteriohopanepolyols in this same microorganism varies by growth condition, with minimal methylation observed during log-phase heterotrophic growth (Rashby et al., 2007). In *Bacillus subtilis*, regular polycyclic terpenoids are synthesized specifically during sporulation, and are demonstrated to bolster spores against oxidative stress (Bosak et al., 2008). Expanding the range of conditions and stages of growth studied can therefore potentially reveal a hidden diversity and more specialized functions of polycyclic terpenoids.

The biosynthetic mechanism responsible for polycyclic triterpenoid diversity, however, remains unresolved. Targeted mutagenesis of catalytic cavities on the SHC of *Alicyclobacillus acidocaldarius* resulted in altered triterpene product patterns. Wild-type *A. acidocaldarius* SHC produces 95% hopene, with the remaining 5% in diplopterol and minor tetracyclic compounds, but SHC mutants for catalytic cavity residues yield significant amounts of α - and γ -polypodatetraenes (up to 15.9% and 72.3%), as well as minor amounts of malabaricatrienes (Full, 2001; Pale-Grosdemange et al., 1999). Similar suites of polycyclic triterpenoids have been observed in two hopene-producing bacteria, *Rhodomicrobium vannielli* and *Zymomonas mobilis*. Wild-type SHCs in these organism synthesize a variety of bi-, tri-, and tetracyclic triterpenes in quantities comparable to that of diploptene (Douka et al., 2001; Howard et al., 1984). Low-

fidelity squalene cyclization is therefore a mechanism for generating polycyclic triterpenoid diversity from a single or limited number of cyclase enzymes.

To further explore polycyclic triterpenoid diversity, we examined non-pigment lipid biomarkers in *Rhodospirillum rubrum* under a variety of growth conditions and phases. *R. rubrum* ATCC 11170 is a purple non-sulfur α -proteobacterium that produces hopanoids (Rohmer et al., 1984); was reported to produce C-20 compounds with a cheilanthane skeleton (Chuck and Barrow, 1995); and has a fully sequenced genome. *R. rubrum* is capable of photoheterotrophic or photoautotrophic growth under anaerobic conditions in the light, and microaerophilic or aerobic growth in the dark. The metabolic versatility of *R. rubrum*, coupled with its genetic tractability, makes it an ideal model organism for exploring the diversity, and eventually the physiological functions, of its polycyclic terpenoids. Because SHC gene surveys suggest that α -proteobacteria appear to dominate hopanoid production in natural environments (Bosak et al., 2008), *R. rubrum*'s SHC is a likely representative of abundant natural enzymes.

For comparison with *R. rubrum*, we expanded our study to survey the diversity of polycyclic terpenoids in other hopanoid-producing α -proteobacteria, namely *Zymomonas mobilis*, *Rhodopseudomonas palustris*, and *Rhodomicrobium vanniellii*. By surveying the dependence of biomarker lipid diversity on growth conditions and phases in these select cultured α -proteobacteria, we can better understand the origin and diversity of polycyclic terpenoids in natural environments, and learn about the diagnostic potential of fossilized lipid derivatives in the sedimentary rock record.

METHODS

Bacterial strains and growth conditions: *Rhodospirillum rubrum* ATCC11170 (NCBI GI# 83574254) was cultured in minimal medium (Wahlund et al., 1991) under the following carbon and growth conditions: fructose/O₂, fructose/N₂, succinate/O₂, Succinate/N₂, thiosulfate/O₂, thiosulfate/N₂. *Rhodopseudomonas palustris* CGA009 (NCBI GI# 39650627) was cultured in the same minimal media on thiosulfate/N₂. *Zymomonas mobilis* ZM4 (NCBI GI# 405607) was grown semi-anaerobically on minimal medium (Galani et al., 1985). *Rhodomicrobium vanniellii* ATCC 17100 was grown anaerobically on malate salts medium (Whittenbury and Dow, 1977). All cultures were grown at 30°C. Optical density of cultures was monitored on wavelengths 660nm and 880nm. Cells were harvested in both exponential and early stationary phase (*R. rubrum*) or only stationary phase (*R. vannielli*, *R. palustris*, *Z. mobilis*), centrifuged, and preserved at -80°C until further analysis.

Lipid Analysis: Frozen cells were thawed at room temperature, resuspended in 2:1/DCM:MeOH and sonicated for 2 hrs. Total lipid extracts were separated by liquid chromatography using a Merck 60H silica gel column equilibrated with hexane. The first, non-polar fraction was analyzed using a HP 6890 GC with a 30 mls Agilent DB-1ms column using He as carrier gas at constant pressure of 23.2 PSI. Oven temperature increased from 65°C to 100°C at 10 min.

SHC alignments: We obtained protein sequences of squalene-hopene cyclases from the Integrated Microbial Genomes database of the Joint Genomes Institute (<http://img.jgi.doe.gov/pub/main.cgi>) and from the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). Amino acid sequences were aligned using CLUSTALW (Thompson et al., 1994). Multiple alignment parameters were: gap open penalty, 13.0; gap extension penalty 0.05; and BLOSUM weight matrix for proteins. The alignments containing conserved residues that are critical for squalene-hopene cyclization catalysis were identified (Pearson et al., 2007).

FIGURES

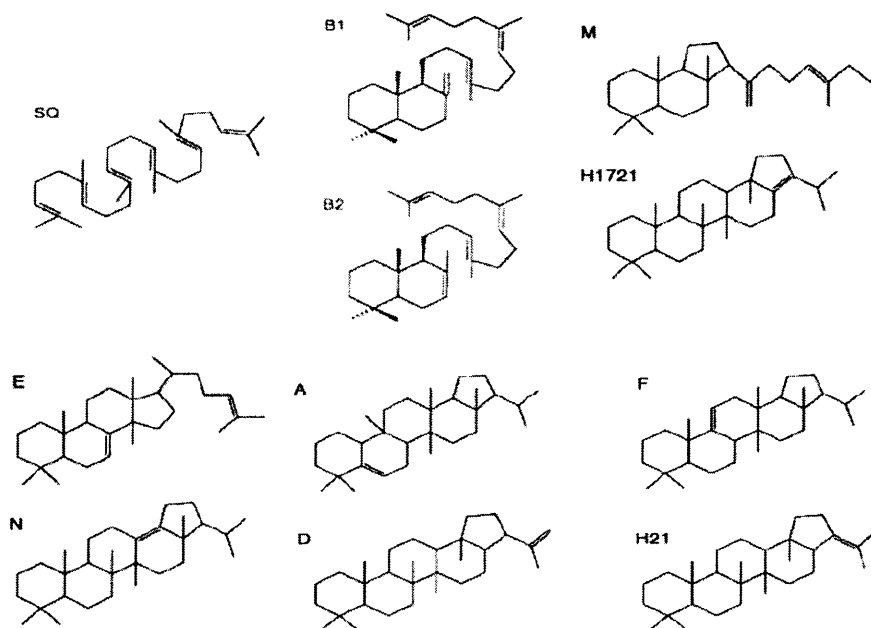


Figure 1. Non-polar triterpenoid compounds identified in the alpha-proteobacteria studied: squalene, bicyclic polypodatetraenes (B1, B2), malabaricatriene (M), eupha-7,24-diene (E), hop-17,21-ene (H1721), neohop-13(18)-ene (N), adianane (A), fernene (F), diploptene (D), and hop-21-ene (H21).

SAMPLE	ABUNDANCE OF NON-POLAR TRITERPENOIDS (% DIPLOPTENE)										
	SQ	B1	B2	M	E	H1721	N	U	A	F	H21
<i>R. rub</i> fructose, exp	43	2.2	3.9	1.8	5	1.6	1.7	3.2	0.8	0.8	nd
<i>R. rub</i> fructose	330	11	26	8.8	13	4.3	5.1	9.7	1.9	2.0	3.1
<i>R. rub</i> fructose, O ₂	416	49	20	22	32	13	12	31	6	7	45
<i>R. rub</i> succinate, exp	225	9.7	17	14	21	9	8.2	15	6.8	8	6.9
<i>R. rub</i> succinate	309	nd	nd	nd	nd	5	nd	nd	nd	nd	11
<i>R. rub</i> succinate, O ₂	164	nd	nd	nd	nd	16	nd	nd	nd	nd	30
<i>R. rub</i> S ₂ O ₃ , exp	159	16	16	20	47	14	14	22	5	8.5	6.2
<i>R. rub</i> S ₂ O ₃	112	52	109	45	71	32	31	34	62	37	50
<i>R. rub</i> S ₂ O ₃ , O ₂	1179	69	120	88	76	32	36	79	23	22	42
<i>R. pal</i> S ₂ O ₃	32	24	26	12	29	17	16	10	nd	20	6.6
<i>Z. mob</i> , O ₂	56	nd	nd	nd	58	45	197	114	nd	1.5	nd
<i>R. van</i>	nd	nd	nd	1.4	3.3	5.2	0.9	4.2	0.6	3.3	17
<i>R. van</i> (Howard, 1984)	nd	nd	nd	nd	nd	12	nd	nd	nd	5	nd

Table 1. Presence/absence table. Amounts (%) of non-polar products of squalene cyclization in *Rhodospirillum rubrum*, *R. palustris*, *R. vannielii*, and *Zymomonas mobilis* relative to diploptene. nd is “not detected”. Cells were harvested at the onset of stationary phase unless indicated for harvest during exponential phase (exp). Cultures were grown under anaerobic, phototrophic conditions unless indicated (ie. O₂ implies aerobic, heterotrophic conditions). Conditions were reproduced once for all *R. rubrum* cultures; other cultures represent preliminary data.

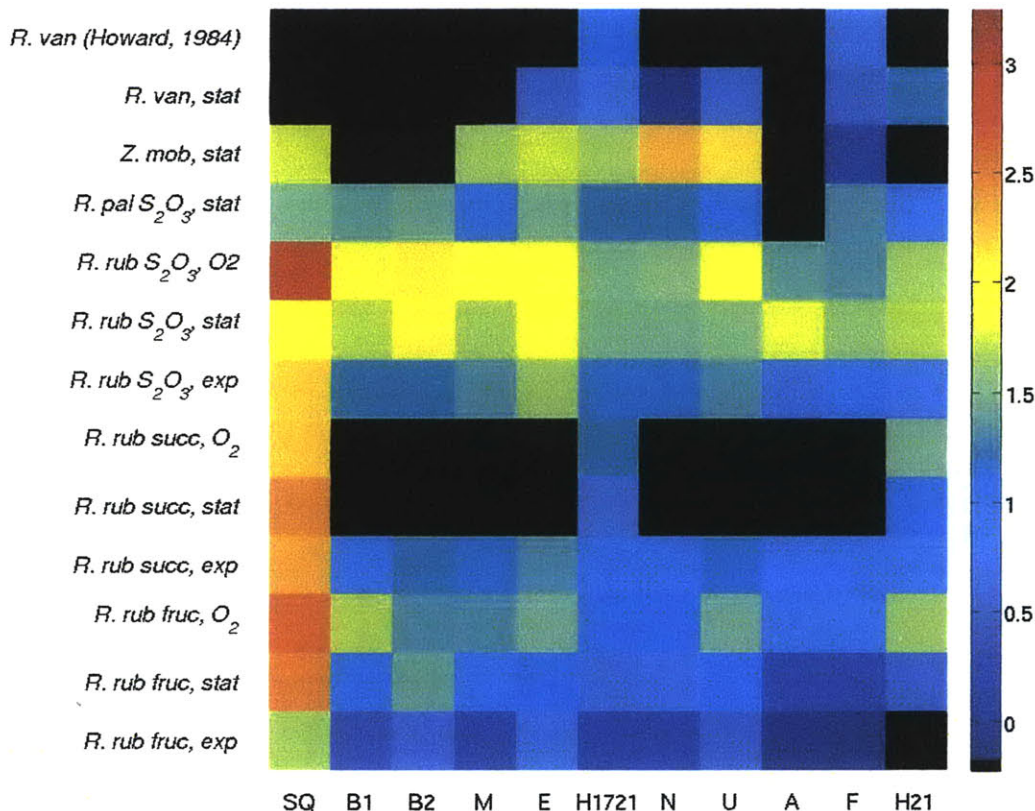


Figure 2. Presence/absence graph. Amount (% diploptene) of non-polar products of squalene cyclization in *Rhodospirillum rubrum* (*R. rub*), *R. palustris*, *R. vannielii*, and *Zymomonas mobilis* (*Z. mob*). Scale is logarithmic with base 10. Refer to table 1 for details.

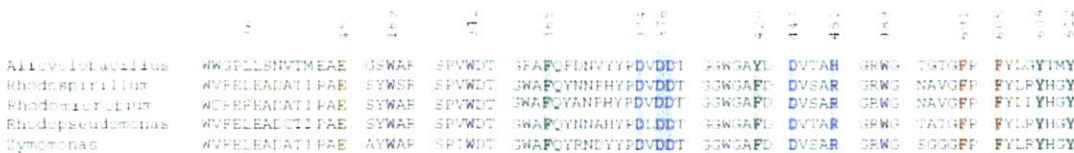


Figure 3. Alignments of amino acids for critical functional domains of SHC. Yellow, highly conserved DxDD motif; pink, substrate-binding residues; green, residues that propagate A- and B-ring carbocations (A and B are the first two rings in the cyclization of squalene to pentacyclic hopene; C and D are subsequent rings); blue, residues for catalytic protonation and reaction initiation; orange; supports the C-ring carbocation; brown; supports D-ring carbocation. Organisms used for the alignment are *Alicyclobacillus acidocaldarius* DSM446 (NCBI GI# 2851526); *Rhodospirillum rubrum* ATCC11170 (NCBI GI# 83574254); *Rhodopseudomonas palustris* CGA009 (NCBI GI# 39650627); *Zymomonas mobilis* ZM4 (NCBI GI# 405607); *Rhodospirillum vannielii* ATCC 17100 (NCBI GI# 10020).

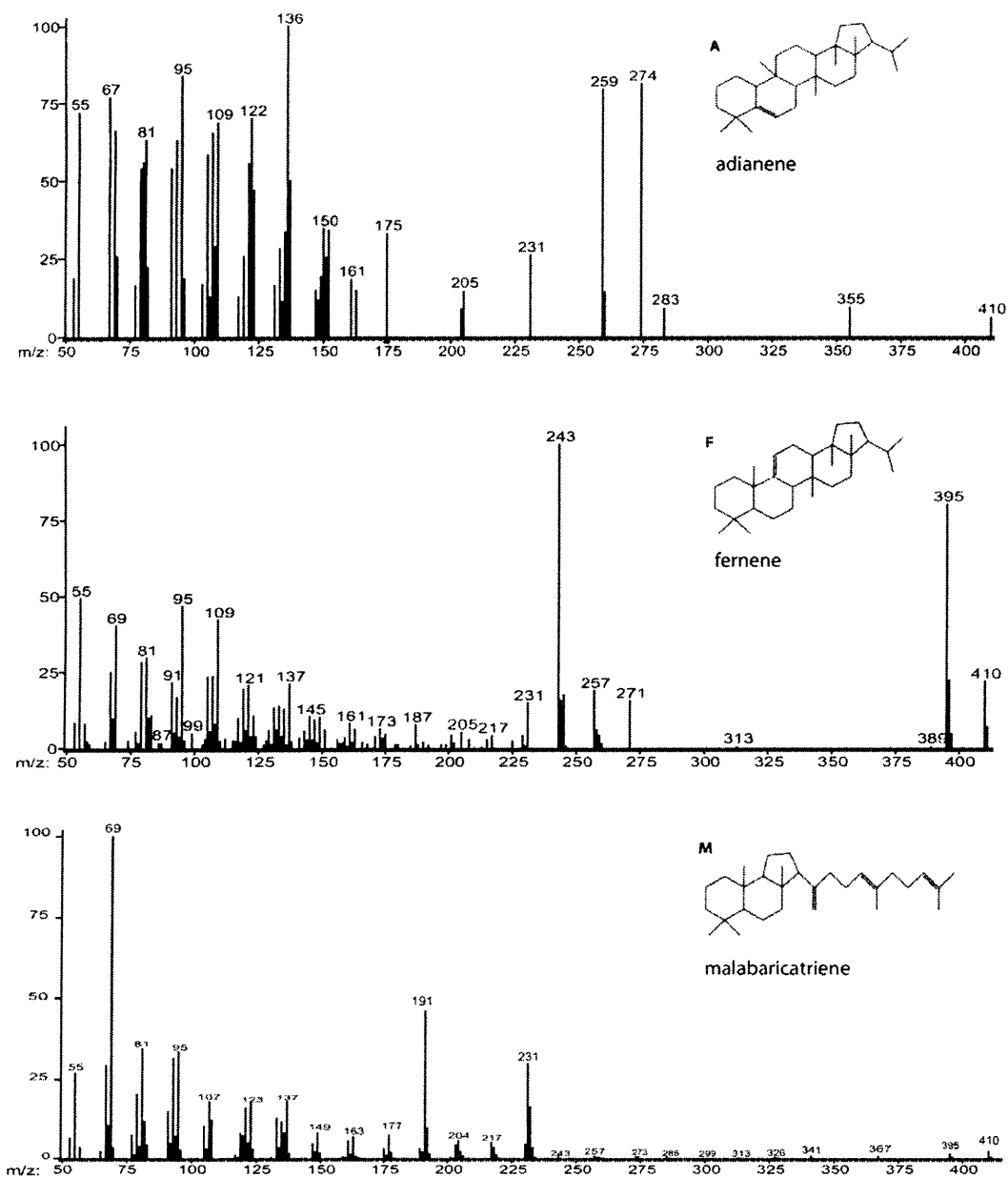


Figure 4. *M/z* spectra of unusual non-polar polycyclic triterpenoids. These spectra were extracted from the GCMS chromatogram for *R. rubrum* grown on fructose under anaerobic, phototrophic conditions, and compounds were identified based on published spectra: A is adianene (Shiojima et al., 1992); F is fernene (Shiojima et al., 1992); M is malabaricatriene (Behrens et al., 1999).

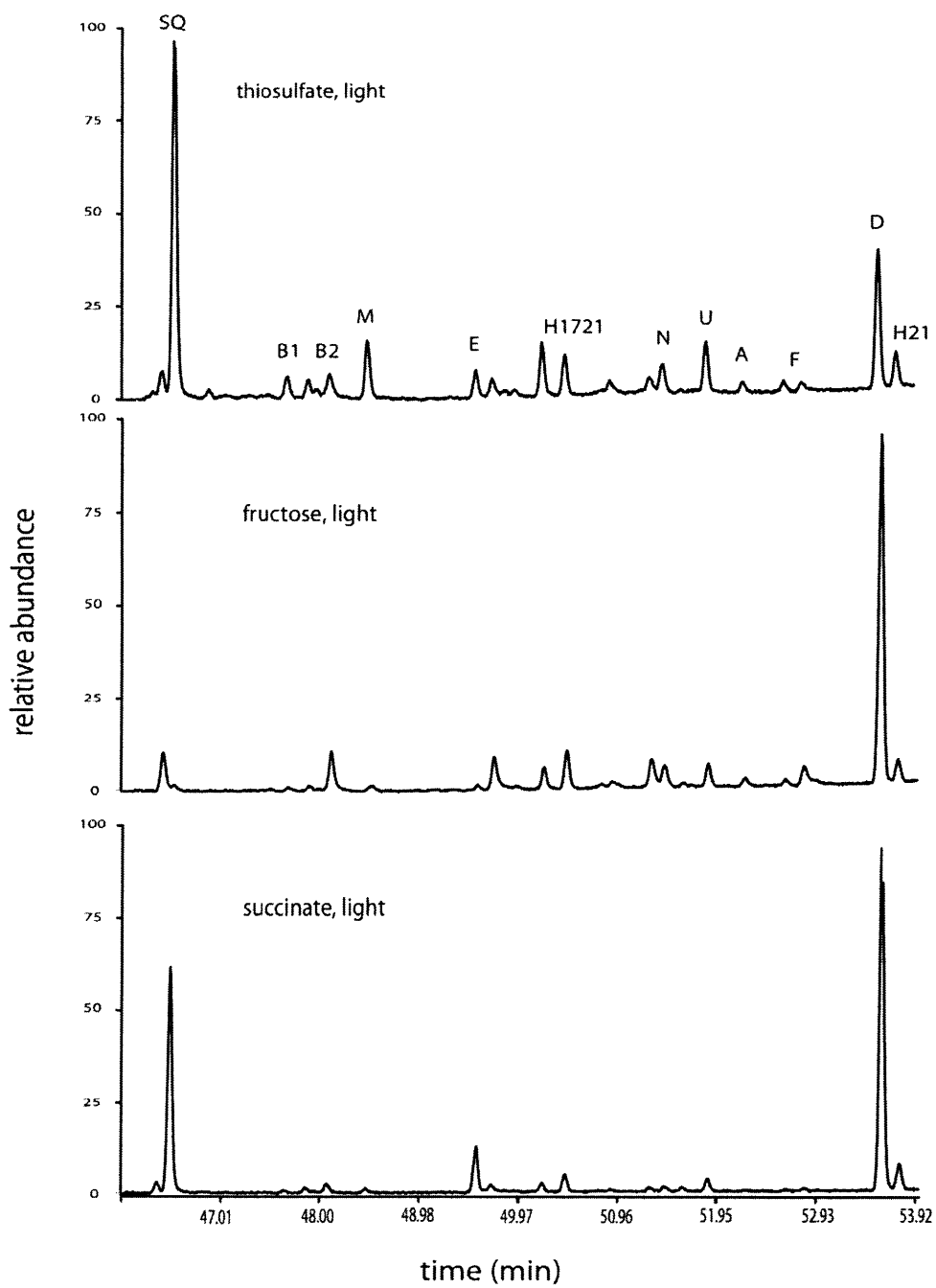


Figure 5. Mass chromatograms (total ion count) of non-polar lipids from *R. rubrum*. Spectra is from cultures grown photosynthetically on thiosulfate (top), fructose (middle), and succinate (bottom). Marked compounds were identified based on comparison of mass spectra with those described in the literature. Cultures were harvested once for each condition, and the presence, but not relative abundance, of these compounds was confirmed in replicate cultures.

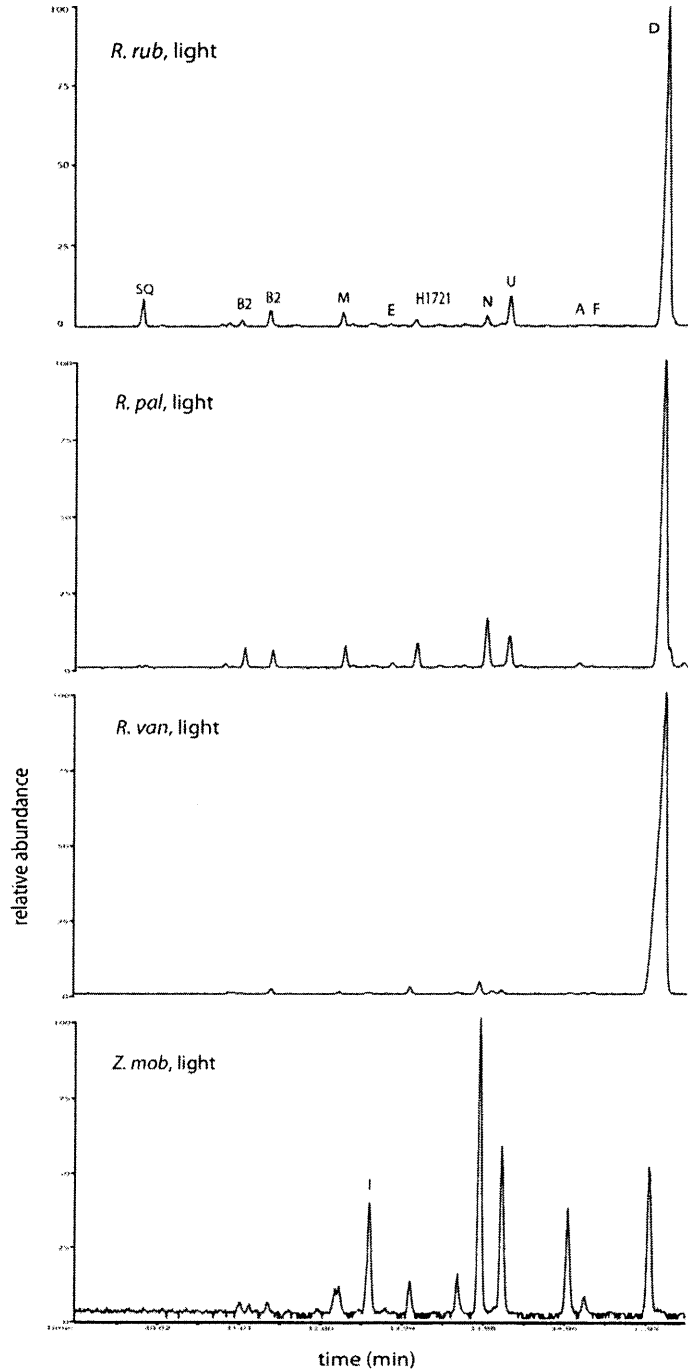


Figure 6. Mass chromatograms (m/z 191) of non-polar lipids from various α -proteobacteria: *R. rubrum*, *R. palustris*, *R. vannielli*, and *Z. mobilis*. All cultures were harvested in early stationary phase. Marked compounds were identified based on comparison of mass spectra with those described in the literature. All compounds identified are listed in Fig. 1, except for “I”, which is 17-isodammara-12,24-diene (Douka et al., 2001). The chromatograms above are each derived from a single culturing experiment.

RESULTS

Diversity of polycyclic terpenoids as a function of growth condition and phase

The GC-MS analysis of the non-polar hydrocarbon fraction of *R. rubrum* revealed the presence of diverse triterpenoids with retention times between those of squalene and diploptene (Fig. 2). An identical suite of compounds was observed in *R. rubrum* under all conditions except for growth on succinate, but the abundance of these polycyclic triterpenoids relative to diploptene varied under different growth conditions (Fig. 5). The maximum relative abundances of compounds were: squalene (1179%) and polypodatetrenes (69% and 120%), malabaricatriene (88%), euphadiene (76%), hop-17,21-ene (32%), neohopene (36%), and the unidentified triterpenoid (79%) in *R. rubrum* grown aerobically on thiosulfate; and hop-21-ene (50%), adianene (62%) and fernene (37%) in *R. rubrum* grown anaerobically on thiosulfate. Minimum abundances for all compounds were observed in *R. rubrum* during heterotrophic growth on succinate, during which only hopenes were detected. Relative abundances of the non-hopene triterpenoids were consistently higher during stationary versus exponential phase for cells grown on thiosulfate and fructose. The exception again was cultures grown on succinate, which showed decreased diversity during stationary phase and under oxic growth conditions. However, to be certain that this pattern holds, growth stage experiments should be repeated.

In the other α -proteobacteria studied, namely *R. palustris*, *Z. mobilis*, and *R. vannielii*, we observed similar but not identical suites of polycyclic triterpenoids. Adianene was below the detection threshold in all these cultures (Table 1). *R. palustris* showed the diversity and relative abundance of compounds most comparable to *R. rubrum* grown phototrophically on thiosulfate. *R. vannielii* had low polycyclic triterpenoid diversity relative to the rest, but greater diversity

than previously reported for this microorganism, with the detection of small amounts of euphadiene (3.3%), and neohopene (0.9%) (Howard et al., 1984). In *Z. mobilis*, we observed triterpenes present in relative amounts comparable with those previously reported (Douka et al., 2001), but we were unable to detect certain reported compounds, namely polypodatraenes.

Identification of malabaricatriene and other unusual polycyclic triterpenoids

A number of bicyclic, tricyclic, tetracyclic, and pentacyclic triterpenoids were detected by GC-MS in *R. rubrum*, *R. palustris*, *R. vanniellii*, and *Z. mobilis*. Compounds were tentatively identified on the basis of published mass spectra in the literature. Polycyclic triterpenoids detected included: bicyclic polypodatetraenes (Pale-Grosdemange et al., 1999); 17(E)-13a(H)-Malabarica-14(27),17,21-triene (Behrens et al., 1999); eupa-7,24-diene (Hoshino et al., 2000); and hop-17,21-ene, neohop-13(18)-ene, adian-5-ene (Hoshino et al., 2000); fernene, diploptene, and hop-21-ene (Shiojima et al., 1992). Malabaricatriene was further identified on the basis of identical elution time and spectra to those of a reference sample from the sediments of Lake Cadagno (Fig. 3; (Behrens et al., 1999). Some of these compounds, such as malabaricatriene (Behrens et al., 1999; Schouten et al., 2000; Werne et al., 2000) and adianene (Ageta and Arai, 1983), previously lacked microbial sources, while ferenes have only been reported in *Z. mobilis* and *R. vanniellii* (Douka et al., 2001; Howard et al., 1984). Tetracyclic triterpenes of the dammarane and euphane series have been previously detected as trace compounds in cells of *A. acidocaldarius* as well as its purified and cloned squalene cyclase (Pale-Grosdemange et al., 1999). Closer scrutiny of GCMS spectra for non-polar fractions extracted from even more microorganisms might therefore reveal an even greater diversity of polycyclic triterpenoids.

DISCUSSION

Low-fidelity squalene cyclization by SHC

The diversity of triterpenoids we observed in α -proteobacteria likely results from low-fidelity squalene cyclization by SHCs. This type of non-specific squalene cyclization was previously observed in *Z. mobilis*, whose SHC yields a range of unusual hydrocarbons lacking clear physiological functions (Douka et al., 2001). Defective squalene cyclization was also induced by the targeted mutagenesis of conserved residues in SHC from *A. acidocaldarius* (Pale-Grosdemange et al., 1999). The hydrocarbon fraction of wild-type SHC from *A. acidocaldarius* contained mainly diploptene, with trace amounts of other hydrocarbons, but the mutated product pattern consisted of a suite of polycyclic triterpenoids, including bicyclic polypodatetraenes, tricyclic malabaricatriene, and tetracyclic dammaradienes (Pale-Grosdemange et al., 1999).

In the case of *R. rubrum*, the fidelity of squalene cyclization in this organism appears to alter in response to both metabolic growth conditions and culture age (Fig. 2). The specificity of this response might indicate unique physiological functions for certain polycyclic terpenoids, such as observed for sporulene during sporulation in *B. subtilis* (Bosak et al., 2008). Alternatively, SHC enzymes might operate with varying fidelity under different conditions in the cell, where polycyclic triterpenoid diversity would be a random offshoot of more or less specific squalene cyclization rather than serving a physiological role. The fact that squalene quantities are extremely high in some cultures with significant polycyclic triterpenoid diversity, such as *R. rubrum* grown aerobically and anaerobically with thiosulfate, suggests that squalene cyclization has been stalled at the enzymatic level. This could result from insufficient expression of the SHC enzyme or the premature quenching of cyclization due to unusual operating conditions, leading to the production of various non-hopene compounds and the accumulation of non-cyclized terpenes.

Acid-induced processes are capable of rearranging hopenes into fernenes and adianenes, among other compounds, although the protocol we employed to extract and separate polycyclic triterpenoids was too benign to induce the abiotic isomerization of diploptene (Ageta and Arai, 1983). But perhaps different growth conditions correspond to locally acidic conditions where SHC operates within cells, leading to increased amounts of fernenes and adianenes, such as observed in *R. rubrum* grown photoautotrophically on thiosulfate. Ultimately, the degree to which condition-dependent polycyclic triterpenoid diversity in *R. rubrum* is a specific physiological response, or simply a side effect of inefficient squalene cyclization, demands further investigation (see Future Work).

The genetic discrepancies between amino acids in critical functional domains of SHC genes for *A. acidocaldarius* compared to the α -proteobacteria studied may contribute to the observed discrepancies in fidelity cyclization among these organisms. In the amino acid sequence alignments for SHC, residue 420, which is responsible for stabilizing the first and second cyclohexyl rings (Merkofer et al., 1999; Pale-Grosdemange et al., 1999), is different in the various organisms studied (Figure 3). In the *A. acidocaldarius* mutagenesis experiments, tyrosine, a polar residue at 420, was substituted for alanine, a non-polar residue. This resulted in the production of bicyclic and tricyclic triterpenes. Similarly, the α -proteobacteria we studied have phenylalanine as residue 420, a non-polar amino acid. Like the mutagenesis experiments, this could lead to subtle variations in substrate positioning for each organism's SHC, resulting in the premature quenching of the carbocation and the production of bicyclic, tricyclic and tetracyclic byproducts. Detailed biochemical and genetic investigation, through enzymatic modeling or targeted mutagenesis, could help reveal the precise mechanism behind the observed

phenomenon of polycyclic triterpenoid diversity generated from a single class of SHC operating under a range of conditions (see Future Work).

Environmental significance of polycyclic terpenoids of α -proteobacteria

While the physiological function, if any, of these unusual polycyclic triterpenoids remains unknown, fernenes and malabaricatriene are environmentally relevant compounds. Fernane skeletons postulated to be microbial in origin were detected in lacustrine sediments ranging from Permian to recent in age [e.g. (Borrego et al., 1997; Brassell and Eglinton, 1983; Hauke et al., 1995; Hauke et al., 1992; Jaffe and Hausmann, 1995)], including anoxic depositional environments in Antarctica (Volkman et al., 1986). Malabaricatriene was first isolated from the Indian tree *Ailanthus malabarica* (Chawla and Dev, 1967), and has since been detected in higher plants (Paton et al., 1979), ferns (Arai et al., 1983), and even a marine sponge *Japsis stellifera* (Ravi et al., 1981). Malabaricatrienes are also abundant in sulfur-rich anaerobic environments where ferns are not found, including anoxic lake sediments from a high-altitude meromictic lake in the Swiss Alps (Behrens et al., 1999), anoxic marine sediments from the Cariaco Basin in Venezuela (Werne et al., 2000), and Arabian Sea sediments (Schouten et al., 2000).

These findings hinted strongly at microbes capable of thriving in anoxic and sulfidic ecosystems as sources for malabaricatriene (Behrens et al., 1999). Now we have detected malabaricatriene in a variety of alpha-proteobacteria, cultivated under both aerobic and anaerobic conditions, with the growth of *R. rubrum* and *R. palustris* on thiosulfate particularly conducive to the synthesis of abundant malabaricatrienes (Fig. 5). Whether environmental strains related to *Rhodospirillaceae*, or other α -Proteobacteria, are responsible for generating the malabaricatrienes and other unusual terpenoids detected in in anoxic, sulfidic natural settings, remains to be

determined, but in our work one of the highest yields of malabaricatriene relative to diploptene was under anoxic conditions (Fig. 2).

Steroid and hopanoid terpene cyclases are hypothesized to have evolved from a common ancestral enzyme that cyclized regular terpenoids, since such an enzyme would contain the minimum set of conserved residues for all known polycyclic terpenoid cyclases. (Ourisson et al., 1987). Pearson et al. (2007) proposed the presence of extant organisms harboring malabaricanoid cyclases based on evidence of malabaricanes in modern environmental settings. However, our work demonstrated that malabaricanes can simply be the byproduct of low-fidelity cyclization of squalene by SHC, obviating the need for a distinct, more primitive cyclase. When malabaricatrienes are detected in natural environments, they are among the most dominant polycyclic terpenoids (Behrens et al., 1999; Schouten et al., 2000; Werne et al., 2000), whereas in our experimental cultures they are not. This suggests that different conditions or even different organisms are responsible for abundant malabaricatrienes in natural settings. Still, *Rhodospirillum rubrum* and other alpha-proteobacteria therefore serve as promising model organisms for exploring the diversity and diagnostic significance of polycyclic triterpenoids, both in the laboratory, in modern environmental settings, and in the sedimentary rock record.

FUTURE WORK

This work characterized the production and diversity of polycyclic terpenoids in *Rhodospirillum rubrum* as a function of different growth conditions and growth phases, and compared this diversity to that observed among various α -proteobacteria. We have shown that low-fidelity squalene cyclization is found in number of α -proteobacteria, including *R. rubrum*, *R. palustris*, *R. vanniellii*, and *Z. mobilis*. Our findings confirm that “orphan” biomarker compounds such as malabaricatriene are synthesized by some of the known hopanoid producers under growth conditions that were not explored in previous studies. Future work should focus on three areas: 1) expanding this study of polycyclic terpenoid diversity in α -proteobacteria to other SHC-bearing microorganisms; 2) exploring the physiological function of polycyclic triterpenoids in *R. rubrum*; and 3) investigating the environmental significance of these compounds.

1) Polycyclic triterpenoid diversity in other SHC-bearing microorganisms

In this work, our investigations of the fidelity of squalene cyclization as the source of polycyclic terpenoid diversity were restricted to α -proteobacteria. To determine whether incomplete squalene cyclization is a phenomenon limited to select α -proteobacteria, or also widespread among other SHC-bearing microorganisms, we can follow the same growth, extraction, and analysis protocols with hopanoid-producing cyanobacteria, and beta- and gamma-proteobacteria. The cyanobacteria would serve as a control organism to demonstrate that the extraction protocol is not responsible for generating the observed polycyclic terpenoid diversity, since these organisms have not been reported to produce a diversity of polycyclic triterpenoids. The investigation of the diversity of polycyclic triterpenoids in beta- and gamma-proteobacteria

would help test whether incomplete squalene cyclization is a signature unique to α -proteobacterial SHCs, or more widespread.

2) Functional roles of polycyclic terpenoids in *R. rubrum*

Using the genetically tractable, metabolically versatile, and environmentally relevant α -proteobacterium *Rhodospirillum rubrum* as a model organism, it should be possible to characterize its suite of polycyclic terpenoids, resolve the genes from which they evolve, determine the conditions that activate these genes, and relate this lab-derived knowledge to modern natural settings and the sedimentary rock record. This involves generating strains of *R. rubrum* mutant for genes related to hopanoid biosynthesis. Mutant strains can be obtained through cloning and site-specific mutagenesis in candidate squalene cyclase genes (Garcia-Contreras et al., 2004; Hessner et al., 1991). If the mutants prove viable, we can monitor the relative fitness of mutant strains under a range of growth conditions to determine the physiological relevance, if any, of polycyclic triterpenoids in *R. rubrum*. To magnify the effects of weak phenotypes, we could conduct competition experiments in mixed cultures of mutants and wild-type *R. rubrum* (Bosak et al., 2008). Additional clues to the physiology of terpenoids could come from screening *R. rubrum*'s genome for genes in the cyclase operons, a search strategy that helped identify the protective role of sporulenes against oxidating stress in *B. subtilis* (Bosak et al., 2008).

3) SHC gene diversity in anoxic, terrestrial environments

Here we can use culture-independent molecular surveys to investigate the diversity and abundance of SHC genes in marine and terrestrial settings where unique polycyclic terpenoids,

such as malabaricatriene, have been detected. In a study of the distribution of microbial terpenoid cyclases in the global ocean metagenome, α -proteobacteria generally and *R. rubrum* specifically had both the greatest number and greatest identity of recruited SHC sequences in all ocean basins sampled (Pearson et al., 2009). The environments sampled in this study encompass diverse ocean climates and depositional settings, supporting a major role for α -proteobacteria as sources of hopanoids in the marine sediments. However, the locales analyzed in this study did not include anaerobic, terrestrial settings. Whether environmental strains of α -proteobacteria related to *Rhodospirillaceae*, or other organisms are responsible for generating the malabaricatrienes detected in natural, anoxic settings remains to be determined. The isotopic compositions of malabaricatriene isomers in Arabian sea sediments (-21.8‰) were relatively enriched compared to algal sterenes, suggesting a distinct and unknown origin for these compounds in this setting (Schouten et al., 2000). Fernenes have been detected in sediments in Ace Lake, a saline, meromictic lake in the Vestfold Hills of Antarctica (Volkman et al., 1986). Surveying SHC-encoding genes in anoxic environments, such as Lake Cadagno (Behrens et al., 1999), could help evaluate the potential for such communities to generate polycyclic terpenoids (Pearson et al., 2009; Pearson et al., 2007).

The production of significant amounts of malabaricatriene and other unusual polycyclic triterpenoids, however, is not simply a function of possessing the SHC gene, but rather a function of the conditions conducive to low fidelity cyclization of squalene. Our work has shown that *R. rubrum* produces the greatest relative abundance of diverse polycyclic triterpenoids during both aerobic or anaerobic growth on thiosulfate. Future work should therefore elaborate on why, when, how, and where polycyclic terpenoids are produced by *R. rubrum* and other proteobacteria.

Is low fidelity squalene cyclization common to all SHCs to greater or lesser degrees, or is it characteristic for α -proteobacteria? What mechanism drives the condition-dependent biosynthesis of diverse polycyclic triterpenoids? What can the physiological functions, if any, of modern polycyclic terpenoids in α -proteobacteria reveal about more ancient metabolisms and environments? By correlating growth conditions with the production of certain polycyclic terpenoids, linking them to microbial communities in modern environments, and relating them to fossilized derivatives in the sedimentary rock record, we can make informed extrapolations into the history and evolution of life on Earth.

BIBLIOGRAPHY

- Ageta, H. and Arai, Y., 1983. Fern constituents: pentacyclic triterpenoids isolated from *Polypodium niponicum* and *Polypodium formosanum*. *Phytochemistry*, 22(8): 1801-1808.
- Arai, Y., Kusumoto, Y., Nagao, M., Shiojima, K. and Ageta, H., 1983. Composite constituents: aliphatics and triterpenoids isolated from the whole plants of *Ixeris debilis* and *Ixeris dentata*. *Journal of the Pharmaceutical Society of Japan*, 103(3): 356-359.
- Behrens, A., Schaeffer, P. and Albrecht, P., 2000. Occurrence of precursors of regular tricyclopolyterpenoids in recent sediments. *Organic Letters*, 2(3): 361-363.
- Behrens, A., Schaeffer, P., Bernasconi, S. and Albrecht, P., 1999. 17(E)-13 alpha(H)-Malabarica-14(27), 17,21-triene, an unexpected tricyclic hydrocarbon in sediments. *Organic Geochemistry*, 30(5): 379-383.
- Blumenberg, M. et al., 2006. Biosynthesis of hopanoids by sulfate-reducing bacteria (genus *Desulfovibrio*). *Environmental Microbiology*, 8(7): 1220-1227.
- Borrego, C.M. et al., 1997. Distribution of bacteriochlorophyll homologs in natural populations of brown-colored phototrophic sulfur bacteria. *FEMS Microbiology Ecology*, 24(4): 301-309.
- Bosak, T., Losick, R.M. and Pearson, A., 2008. A polycyclic terpenoid that alleviates oxidative stress. *Proceedings of the National Academy of Sciences of the United States of America*, 105(18): 6725-6729.
- Brassell, S.C. and Eglinton, G., 1983. Steroids and triterpenoids as environmental and diagenetic indicators. In: M. Bjoroy, C. Albrecht and C. Cornford (Editors), *Advances in Organic Geochemistry*. John Wiley & Sons, New York, pp. 684-697.
- Brocks, J.J., Buick, R., Summons, R.E. and Logan, G.A., 2003. A reconstruction of Archean biological diversity based on molecular fossils from the 2.78 to 2.45 billion-year-old Mount Bruce Supergroup, Hamersley Basin, Western Australia. *Geochimica Et Cosmochimica Acta*, 67(22): 4321-4335.
- Brocks, J.J., Logan, G.A., Buick, R. and Summons, R.E., 1999. Archean molecular fossils and the early rise of eukaryotes. *Science*, 285(5430): 1033-1036.
- Chawla, A. and Dev, S., 1967. A new class of triterpenoids from *Ailanthus malabarica*: derivatives of malabaricane. *Tetrahedron Letters*(48): 4837-&.
- Chuck, J.A. and Barrow, K.D., 1995. The isolation of isoagathenediol: A new tricyclic diterpene from the lipids of *Rhodospirillum rubrum*. *Microbiology-Uk*, 141: 2659-2663.
- Douka, E., Koukkou, A.I., Drainas, C., Grosdemange-Billiard, C. and Rohmer, M., 2001. Structural diversity of the triterpenic hydrocarbons from the bacterium *Zymomonas mobilis*: the signature of defective squalene cyclization by the squalene/hopene cyclase. *FEMS Microbiology Letters*, 199(2): 247-251.
- Fischer, W.W., Summons, R.E. and Pearson, A., 2005. Targeted genomic detection of biosynthetic pathways: anaerobic production of hopanoid biomarkers by a common sedimentary microbe. *Geobiology*, 3(1): 33-40.
- Full, C., 2001. Bicyclic triterpenes as new main products of squalene-hopene cyclase by mutation at conserved tyrosine residues. *FEBS Letters*, 509(3): 361-364.
- Galani, I., Drainas, C. and Typas, M.A., 1985. Growth requirements and the establishment of a chemically defined minimal medium in *Zymomonas mobilis*. *Biotechnology Letters*, 7(9): 673-678.

- Garcia-Contreras, R., Celis, H. and Romero, I., 2004. Importance of *Rhodospirillum rubrum* H⁺-pyrophosphatase under low-energy conditions. *Journal of Bacteriology*, 186(19): 6651-6655.
- Hauke, V. et al., 1995. Isoarborinol through geological times: Evidence for its presence in the Permian and Triassic. *Organic Geochemistry*, 23(1): 91-93.
- Hauke, V. et al., 1992. Novel triterpene-derived hydrocarbons of arborane/fernane series in sediments. *Tetrahedron*, 48(19): 3915-3924.
- Hessner, M.J., Wejksnora, P.J. and Collins, M.L.P., 1991. Construction, characterization, and complementation of *Rhodospirillum rubrum* Puf region mutants. *Journal of Bacteriology*, 173(18): 5712-5722.
- Horbach, S., Neuss, B. and Sahn, H., 1991. Effect of azasqualene on hopanoid biosynthesis and ethanol tolerance of *Zymomonas mobilis*. *FEMS Microbiology Letters*, 79(2-3): 347-350.
- Hoshino, T., Kouda, M., Abe, T. and Sato, T., 2000. Functional analysis of Phe605, a conserved aromatic amino acid in squalene-hopene cyclases. *Chemical Communications*(16): 1485-1486.
- Howard, D.L., Simoneit, B.R.T. and Chapman, D.J., 1984. Triterpenoids from lipids of *Rhodomicrobium vannielii*. *Archives of Microbiology*, 137(3): 200-204.
- Jaffe, R. and Hausmann, K.B., 1995. Origin and early diagenesis of arborinone isoarborinol in sediments of a highly productive freshwater lake. *Organic Geochemistry*, 22(1): 231-235.
- Kontnik, R. et al., 2008. Sporulenes, heptaprenyl metabolites from *Bacillus subtilis* spores. *Organic Letters*, 10(16): 3551-3554.
- Mackenzie, A.S., 1984. Application of biological markers in petroleum geochemistry. In: B. J. and D.H. Welte (Editors), *Advances in Petroleum Organic Geochemistry*. Academic Press, London, pp. 115-214.
- Merkofer, T., Pale-Grosdemange, C., Wendt, K.U., Rohmer, M. and Poralla, K., 1999. Altered product pattern of a squalene-hopene cyclase by mutagenesis of active site residues. *Tetrahedron Letters*, 40(11): 2121-2124.
- Moldowan, J.M. et al., 1994. The molecular fossil record of oleanane and its relation to angiosperms. *Science*, 265(5173): 768-771.
- Nytoft, H.P. and Larsen, B., 2001. Triterpenoids and other organic compounds as markers of depositional conditions in the Baltic Sea deep basins during the Holocene. *Baltica*, 14: 95-107.
- Ourisson, G. and Albrecht, P., 1992. Geohopanoids: the most abundant natural products on Earth. *Accounts of Chemical Research*, 25(9): 398-402.
- Ourisson, G., Rohmer, M. and Poralla, K., 1987. Prokaryotic hopanoids and other polyterpenoid sterol surrogates. *Annual Reviews of Microbiology*, 41: 301-333.
- Pale-Grosdemange, C., Merkofer, T., Rohmer, M. and Poralla, K., 1999. Production of bicyclic and tricyclic triterpenes by mutated squalene-hopene cyclase. *Tetrahedron Letters*, 40(33): 6009-6012.
- Paton, W.F., Paul, I.C., Bajaj, A.G. and Dev, S., 1979. Structure of malabaricol. *Tetrahedron Letters*, (43): 4153-4154.
- Pearson, A. et al., 2009. Diversity of hopanoids and squalene-hopene cyclases across a tropical land-sea gradient. *Environmental Microbiology*, 11(5): 1208-1223.
- Pearson, A., Page, S.R.F., Jorgenson, T.L., Fischer, W.W. and Higgins, M.B., 2007. Novel hopanoid cyclases from the environment. *Environmental Microbiology*, 9(9): 2175-2188.

- Poralla, K., Hartner, T. and Kannenberg, E., 1984. Effect of temperature and Ph on the hopanoid content of *Bacillus acidocaldarius*. FEMS Microbiology Letters, 23(2-3): 253-256.
- Rashby, S.E., Sessions, A.L., Summons, R.E. and Newman, D.K., 2007. Biosynthesis of 2-methylbacteriohopanepolyols by an anoxygenic phototroph. Proceedings of the National Academy of Sciences of the United States of America, 104(38): 15099-15104.
- Ravi, B.N., Wells, R.J. and Croft, K.D., 1981. Malabaricane triterpenes from a Fijian collection of the sponge *Jaspis stellifera*. Journal of Organic Chemistry, 46(10): 1998-2001.
- Rohmer, M., Bouviernave, P. and Ourisson, G., 1984. Distribution of hopanoid triterpenes in prokaryotes. Journal of General Microbiology, 130(May): 1137-1150.
- Schouten, S., Hoefs, M.J.L. and Damste, J.S.S., 2000. A molecular and stable carbon isotopic study of lipids in late Quaternary sediments from the Arabian Sea. Organic Geochemistry, 31(6): 509-521.
- Shiojima, K. et al., 1992. Mass-spectra of pentacyclic triterpenoids. Chemical & Pharmaceutical Bulletin, 40(7): 1683-1690.
- Thompson, J.D., Higgins, D.G. and Gibson, T.J., 1994. ClustalW: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research, 22(22): 4673-4680.
- Volkman, J.K., Allen, D.I., Stevenson, P.L. and Burton, H.R., 1986. Bacterial and algal hydrocarbons in sediments from a saline Antarctic lake, Ace Lake. Organic Geochemistry, 10(4-6): 671-681.
- Wahlund, T.M., Woese, C.R., Castenholz, R.W. and Madigan, M.T., 1991. A thermophilic green sulfur bacterium from New Zealand hot springs, *Chlorobium tepidum* Sp. nov. Archives of Microbiology, 156(2): 81-90.
- Wang, T.G. and Simoneit, B.R.T., 1995. Tricyclic terpanes in Precambrian bituminous sandstone from the Eastern Yanshan Region, North China. Chemical Geology, 120(1-2): 155-170.
- Welander, P.V. et al., 2009. Hopanoids play a role in membrane integrity and pH homeostasis in *Rhodopseudomonas palustris* TIE-1. Journal of Bacteriology, 191(19): 6145-6156.
- Werne, J.P. et al., 2000. Timing of early diagenetic sulfurization of organic matter: A precursor-product relationship in Holocene sediments of the anoxic Cariaco Basin, Venezuela. Geochimica Et Cosmochimica Acta, 64(10): 1741-1751.
- Whittenbury, R. and Dow, C.S., 1977. Morphogenesis and differentiation in *Rhodomicrobium vannielii* and other budding and prosthecate bacteria. Bacteriological Reviews, 41(3): 754-808.