8.03 Sedimentary Hydrocarbons, Biomarkers for Early Life

J. J. Brocks
Harvard University, Cambridge, MA, USA
and
R. E. Summons
Massachusetts Institute of Technology, Cambridge, MA, USA

8.03.1 INTRODUCTION

8.03.2 BIOMARKERS AS MOLECULAR FOSSILS

8.03.2.1 The Fate of Dead Biomass: Diagenesis, Catagenesis, and Metagenesis

8.03.2.2 Compound-specific Stable Isotopes

8.03.3 THERMAL STABILITY AND MATURITY OF BIOMARKERS

8.03.3.1 Biomarkers as Maturity Indicators

8.03.3.2 The Survival of Biomarkers with Increasing Temperature and Time

8.03.4 EXPERIMENTAL APPROACHES TO BIOMARKER AND KEROGEN ANALYSIS

8.03.5 DISCUSSION OF BIOMARKERS BY HYDROCARBON CLASS

8.03.5.1 Advantages and Limitations of the Biomarker Approach

8.03.5.2 n-Alkanes, Algoenan, and other Polymethylene Biopolymers

8.03.5.3 Methyl and Ethyl Alkanes

8.03.5.4 Alkyl Cyclohexanes and Cyclopentanes

8.03.5.5 Isoprenoids

8.03.5.6 Carotenoids

8.03.5.6.1 Aromatic carotenoids and arylisoprenoids

8.03.5.6.2 Bacterioruberin

8.03.5.7 Chlorophylls and Maleimides

8.03.5.8 Sesquiterpanes (C_{15}) and Diterpanes (C_{20})

8.03.5.9 Cheilanthanes and other Tricyclic Polyprenoids

8.03.5.10 Hopanoids and other Pentacyclic Triterpanes

8.03.5.11 Steroid Hydrocarbons

8.03.6 RECONSTRUCTION OF ANCIENT BIOSPHERES: BIOMARKERS FOR THE THREE DOMAINS OF LIFE

8.03.6.1 Bacteria

8.03.6.1.1 Hopanoids as biomarkers for bacteria

8.03.6.1.2 Cyanobacteria

8.03.6.1.3 Methanotrophs, methylotrophs, and acetic acid bacteria

8.03.6.1.4 Phototrophic sulfur bacteria

8.03.6.2 Archaea

8.03.6.2.1 Methanogens

8.03.6.2.2 Biomarkers and ecology at marine methane seeps

8.03.6.2.3 Halobacteria

8.03.6.2.4 Marine Crenarchaeota

8.03.6.3 Eukarya

8.03.7 BIOMARKERS AS ENVIRONMENTAL INDICATORS

8.03.7.1 Marine versus Lacustrine Conditions

8.03.7.2 Hypersaline Conditions
8.03.1 INTRODUCTION

Molecular biological markers, or biomarkers, are natural products that can be assigned to a particular biosynthetic origin. For environmental and geological studies, the most useful molecular biomarkers are organic compounds with high taxonomic specificity and potential for preservation. In other words, the most effective biomarkers have a limited number of well-defined sources; they are recalcitrant against geochemical changes and easily analyzable in environmental samples. Accordingly, biomarkers can be proxies in modern environments as well as chemical fossils that afford a geological record of an organism’s activities. One of the first significant outcomes of biomarker research was Treibs’ (1936) recognition of unquestionable biological signatures in sedimentary organic matter. Subsequent research (Eglinton and Calvin, 1967; Eglinton et al., 1964) pursued the concept that biomarkers can provide information about the nature of early life in the absence of recognizable fossils and that petroleum is composed of biological remains (Whitehead, 1973). As of early 2000s, thirty years of accumulated facts about sedimentary organic matter are clearly commensurate with the aforesaid and falsify the hypotheses (e.g., Gold, 2001) about the primordial origins of petroleum and natural gas.

Largely as a result of efforts to understand the detail of the transformation of biogenic organic matter into petroleum (Hunt, 1996; Tissot and Welte, 1984) and individual chemical fossils, geochemists began to appreciate the value of biomarkers as tools for environmental research (e.g., Brassell et al., 1986) and their potential for elucidating biogeochemical processes (e.g., Hinrichs et al., 1999; Kuypers et al., 2003). The structural and isotopic information in biomarkers allows them to be distinguished from abiotic organic compounds that are widely distributed throughout the cosmos (e.g., Cronin and Chang, 1993; Engel and Macko, 1997). Consequently, biomarkers will be an important tool in the search for extraterrestrial life. A thorough review of recent biomarker research is not possible within the limitations of this chapter. Instead, this chapter introduces some of the general principles, provides examples of their use for discerning the identities and physiologies of microbes in contemporary environments and summarizes biomarker research aimed at elucidating aspects of biological and environmental evolution in the Precambrian.

8.03.2 BIOMARKERS AS MOLECULAR FOSSILS

Molecular fossils that are stable under geological conditions mostly originate from biologic lipids. These biomarkers encode information about ancient biodiversity, trophic associations, and environmental conditions. They are recorders of element cycling, sediment and water chemistry, redox conditions, and temperature histories. Most importantly, however, hydrocarbon biomarkers are stable for billions of years if they are enclosed in intact sedimentary rocks that have only suffered a mild thermal history. Therefore, biomarkers offer a powerful means to study life and its interaction with the environment as recorded in rocks of Precambrian age. In sedimentary environments, and under appropriate diagentic conditions, functionalized biolipids are reduced to hydrocarbon skeletons (e.g., \(C_{31}\) to \(C_{32}\)). During this process, much of the biological information is retained and it is thereby possible to assign specific hydrocarbon skeletons to specific taxa (Figure 1) wherever their biosynthetic pathways are exclusive. For example, pentacyclic terpanes of the \(C_{31}\) to \(C_{35}\) extended hopane series (55) are diagnostic biomarkers for the domain Bacteria. The biologic precursors of extended hopanes (55), the bacteriohopanepolyols (56), probably have the physiological function of membrane rigidifiers, a role in Eukarya fulfilled by sterols. Important hydrocarbon fossils of eukaryotic sterols such as cholesterol (65) are steranes (e.g., (66)) and aromatic steroids (e.g., (68)). Although some Bacteria are capable of synthesizing a limited variety of sterols, including lanosterol and 4-methylsterols, the wide structural range of
fossil steranes typically found in oils and bitumens is diagnostic for organisms of the domain Eukarya. Similarly, a range of structurally distinctive acyclic and cyclic isoprenoids found in sedimentary rocks can be assigned exclusively to the domain Archaea (Figure 1). Their precursor lipids are hydrocarbon chains bound to glycerol through ether linkages with varied chain lengths, branching patterns and modes of cyclization (e.g., (12)–(14)). Other biomarkers are evidently diagnostic for taxonomic groups below domain level. These include extended 2α-methylhopanes (57) for cyanobacteria, 24-n-propylcholestanes (66d) for pelagophyte algae, 24-isopropylcholestane (66e) for certain sponges, and a large number of very distinctive polycyclic...
compounds characteristic of various plant taxa (e.g., (53) and (61)–(63)). Botryococcanes (e.g., (20)) are hydrocarbon fossils that appear to be diagnostic for a single taxon, the alga Botryococcus braunii. The study of biomarkers in sedimentary rocks thus allows the existence of a taxonomic group to be established for a given geological period. This capability is especially useful prior to the Cambrian where diagnostic body fossils are mostly absent and the affinities of microfossil are less certain.

8.03.2.1 The Fate of Dead Biomass: Diagenesis, Catagenesis, and Metagenesis

Organic matter from defunct organisms is almost quantitatively remineralized back to carbon dioxide in aquatic environments. However, a small fraction of total biomass, on an average less than 0.1% (Holser et al., 1988), escapes remineralization, and eventually accumulates in sediments. As compounds with rapid biological turnover rates—including carbohydrates, proteins, and nucleic acids—are most prone to recycling, more resistant molecules such as lipids and recalcitrant structural biopolymers become concentrated (Tegelaar et al., 1989). During transport through the water column, and subsequently in the unconsolidated sediment, this organic matter is further altered by a variety of chemical and biological processes commonly referred to as diagenesis (e.g., Hedges and Keil, 1995; Hedges et al., 1997; Rullkötter, 1999). During diagenesis a large fraction of the lipid and other low-molecular weight components react via condensation and sulfur vulcanization reactions (e.g., Sandison et al., 2002) and combine with degradation-resistant macromolecules to form kerogen (e.g., de Leeuw and Largeau, 1993; Derenne et al., 1991). Formally, kerogen is defined as the fraction of large chemical aggregates in sedimentary organic matter that is insoluble in solvents. In contrast, the fraction of organic matter that can be extracted from sediments with organic solvents such as dichloromethane and methanol, is defined as bitumen (pyrobitumen and radiobitumen are residues of migrated petroleum that was cross linked and immobilized by heat and radioactivity, respectively). Bitumen in fresh sediments is predominantly composed of functionalized lipids. During diagenesis, these lipids undergo oxidation, reduction, sulfurization, desulfurization, and rearrangement reactions, generating an array of partly or entirely defunctionalized breakdown products that can have different stereo- and structural isomers. Analysis of these alteration products often yields valuable information about prevailing chemical conditions in the sediment during and after deposition because the extent and relative speed of diagenetic reactions is dependent on environmental conditions such as redox state, pH, and availability of catalytic sites on mineral surfaces. Where reducing conditions prevail in the sediment, biolipids eventually lose all functional groups but remain identifiable as geologically stable hydrocarbon skeletons.

Diagenetic reactions in the presence of reduced sulfur species have a profound effect on the sedimentary fate of lipids and other biological debris (Sinninghe Damsté and de Leeuw, 1990) and the preservation of diagnostic carbon skeletons (e.g., Adam et al., 1993; Kenig et al., 1995a; Kohnen et al., 1992, 1993, 1991a,b; Schaeffer et al., 1995; Wakeham et al., 1995) in complex, sulfur-rich macromolecules. The sequestration and subsequent release of these skeletons upon burial provides one of the most important mechanisms for preserving the structural integrity of organism-specific biomarkers.

With increasing burial over millions of years, geothermal heat will initiate catagenesis, the thermal degradation of kerogen and bitumen. Kerogen is cracked into smaller fragments, releasing increasing volumes of bitumen that might eventually be expelled from its source rock as crude oil. Weaker chemical bonds, such as S—S and S—C, are cleaved at relatively low temperatures with the result that sulfur-rich kerogens might commence oil generation at lower temperatures (e.g., Koopmans et al., 1997; Lewan, 1985). Hydrocarbon chains attached to kerogen via stronger C—O and C—C bonds are sequentially released at higher temperatures. Also, with increasing heat flux, biomarkers and other components in the bitumen undergo thermal rearrangement and cracking reactions. By measuring the relative abundances of these thermal products, it is possible to assess the maturity of an oil or bitumen (Section 8.03.3). With continuing burial, and at temperatures and pressures that initiate low-grade metamorphism of the host rock, most or all of residual bitumen is expelled or cracked to gas and the kerogen becomes progressively depleted in hydrogen to form a partly crystalline, highly aromatic carbon phase (metagenesis). The exact temperature and time constraints of metagenesis and the preservation of hydrocarbon biomarkers are much debated (e.g., Mango, 1991; Price, 1997) and are briefly reviewed later in Section 8.03.3.2.

8.03.2.2 Compound-specific Stable Isotopes

The carbon-isotopic content of organic matter carries information about the immediate
environment of an organism, its primary carbon assimilation pathways and subsequent processing of its metabolic products in the environment. While isotopic measurements of bulk organic materials (e.g., biomass, kerogen, bitumen, petroleum) allow some correlations between precursor and product, measurements at the molecular or intramolecular level reveal stunningly detailed information about the biosynthetic pathways and organismic sources of individual carbon skeletons. The feasibility of developing a tool such as gas chromatography–isotope ratio mass spectrometry for routine natural-abundance isotopic measurements of individual organic compounds was first demonstrated by Matthews and Hayes (1978). Subsequent improvements in sensitivity, precision, calibration, analysis software, and general ease of use then enabled a wide range of biogeochemical applications for compound-specific carbon isotope analysis to be explored and exploited (e.g., Freeman et al., 1990; Hayes et al., 1990; Jahneke et al., 1999). Hydrogen, nitrogen, and oxygen isotopes are also amenable to analysis, and multiple isotope ratios for the same compound offers a precise way to determine its provenance (e.g., Engel and Macko, 1997; Hinrichs et al., 2001). Compound-specific carbon-isotopic patterns also reveal much about fractionation during carbon assimilation and biosynthesis (e.g., Grice et al., 1998a; Hayes, 1993, 2001; Jahneke et al., 1999; Kohnen et al., 1992; Rieley et al., 1993; Schouten et al., 1998a; Summons et al., 1994a; van der Meer et al., 2001). Measurements of the individual radiocarbon ages of different organic compounds add a further dimension to studies of recent sediments (e.g., Eglinton et al., 1997). It is now clear that compound-specific isotope analysis has completely revolutionized biomarker research.

8.03.3 THERMAL STABILITY AND MATURITY OF BIOMARKERS

8.03.3.1 Biomarkers as Maturity Indicators

One of the most widely used applications for biomarkers is for the measurement of thermal maturity of organic matter to estimate the petroleum-generation potential and temperature history of sedimentary basins (e.g., Mackenzie, 1984; Radke et al., 1997). A large number of biomarker parameters that are sensitive to different stages of maturity have been developed and are reviewed in Peters and Moldowan (1993). Two examples are described further below. For the interpretation of maturity parameters in the literature, it is important to note that the thermal evolution of biomarkers, and organic matter in general, might be widely different in rocks of different lithological compositions and from different basins and formations. Clay minerals, for example, provide catalytic sites for degradation and isomerization reactions and strongly influence the type and extent of isomer conversion (e.g., Moldowan et al., 1991a). Moreover, the range of biological inputs, presence of organic sulfur compounds and a host of other factors might cause disparate maturity values from the outset. Therefore, considerable caution has to be applied when comparing maturity parameters across disparate sample sets. Similar caution is necessary for the interpretation of conventional organic geochemical nomenclature for hydrocarbon thermal maturity (Figure 2). The generation and maturation process of petroleum can follow

![Figure 2](https://example.com/figure2.png)

*Figure 2* Terminology for bitumen maturity commonly used in the literature (e.g., Peters and Moldowan, 1993). The “modified temperature scale” pertains to hydrocarbon preservation under ideal conditions and was derived from data in Table 1.
markedly different pathways between different samples and between different components of the same sample (e.g., Radke et al., 1997). Thus, the
description of the preservation state of petroleum and bitumen using such terminology as “peak oil
generation” or “overmature” is quite vague and clearly qualitative. It does not necessarily corre-
late with kerogen maturity data (e.g., vitrinite reflectance) or absolute temperatures unless cali-
ibrated for each sample set. Figure 2 should therefore only be used as a visual indicator of the relationships between bitumen descriptions expressed in words, temperatures, and vitrinite reflectance data.

As an example of a typical biomarker maturity parameter, the ratio of 20S/(20S + 20R) isomers in a sterane measures the relative abundance of the S and R configurations at C-20 of sterane hydrocarbons with 5α, 14α, 17α (H) stereochemistry (Figure 3; for sterane nomenclature see Section 8.03.5.11). In living organisms, sterols exclusively possess the 20R configuration, but during diagenesis and catagenesis steranes are gradually transformed to a mixture of 20R and 20S isomers. The thermal equilibrium value of ~0.55 for the 20S/(20S + 20R) ratio is apparently reached close to the peak of oil generation (Peters and Moldowan, 1993). Ratios based on triaromatic steroids (TA) (68a) and (68b) are an example of parameters sensitive at higher thermal maturities, i.e., the late stage of petroleum generation (Riolo et al., 1985). Triaromatic steroids (68) form apparently predominantly by aromatization of monoaromatic steroid precursors (Mackenzie et al., 1981). Thermal cleavage of the side chain of intact C26 to C28-TAs (68b) (TA-II) leads to the generation of degradation products with 20 to 21 carbon atoms (TA-I (68a)). Consequently, in the transition from immature through mature to overmature petroleum, the ratio TA-II/(TA-I + TA-II) increases from <5% to close to 100% (Figure 4).

Hydrocarbons with typical “overmature” compositions and isomer distributions are character-
istically found in rocks with deep-burial history, e.g., in Archean sequences (Brocks et al., 2003a). Adamantanes and diamantanes, for example, are diagnostic classes of “diamondoid” hydrocarbons that persist and become concentrated at extreme levels of thermal maturity (Chen et al., 1996; Dahl et al., 2002, 1999). In contrast to burial metamorphism, hydrocarbons may also be gener-
ated over short timescales at very high tempera-
tures such as those prevailing at recent hydrothermal vents (e.g., Simonite and Fetzer, 1996; Simonite et al., 1992), in shales proximal to centers of hydrothermal ore formation (e.g., Brooks et al., 2003d; Chen et al., 2003; Gize, 1999; Landais and Gize, 1997; Püttmann et al., 1988) or near-volcanic intrusions (e.g., Farrimond et al., 1999; George, 1992). Bitumens that form in these extreme environments also have very distinctive hydrocarbon distribution patterns.

8.03.3.2 The Survival of Biomarkers with Increasing Temperature and Time

The number of sedimentary units that contain indigenous bitumens drastically decreases with increasing age of the rock. However, time alone is not the driver of degradation of organic molecules. For example, amino acids and other highly sensitive compounds may survive in carbonaceous meteorites for many billion years (Engel and Macko, 1997), and hydrocarbon biomarkers have endured in sedimentary rocks with little alteration for as long as 1.7 Gyr (Jackson et al., 1986). Instead, the main factor driving molecular degradation, next to oxidation and erosion of the host rock, is thermal cleavage of covalent bonds during catagenesis and meta-

morphism. All known sedimentary successions older than ~1.7 Ga have suffered burial meta-

morphism to at least prehnite-pumpellyte facies at temperatures between 175 °C and 280 °C. Unfor-

tunately, it is not entirely clear whether these temperatures, if experienced over geological periods of time, are consistent with preservation of biomarker hydrocarbons (Price, 1997). How-

ever, it is possible to obtain reliable minimum estimates of molecular preservation by observing biomarkers in deep-subsurface petroleum reservoirs (Table 1). Some of these reservoirs produce gas condensate and oil at present-day temperatures up to 200 °C, but still contain intact C15+ biomarkers (Brigaud, 1998; Knott, 1999; McNeil and BeMent, 1996; Pepper and Dodd, 1995). So far, the highest reliable temperature was observed in the rapidly subsiding Los Angeles Basin that contains moderately mature kerogens and bio-

markers at 223 °C (Price, 2000; Price et al., 1999).

Earlier reports of relatively immature bitumens allegedly preserved at even higher temperatures (226–296 °C) lack credible information about syngeneity and require reconsideration (Price, 1982, 1983, 1993, 1997; Price et al., 1981). These observations are consistent with kinetic models of petroleum degradation that predict

![Figure 3 Equilibration between the biological 20R epimer and the geological 20S epimer of cholestane 66a.](image-url)
persistence of aliphatic hydrocarbons over geological periods of time at 250 °C (Burnham et al., 1997; Domine` et al., 2002; Pepper and Dodd, 1995). Therefore, the existence of residual biomarkers in the lowest grade of metasedimentary rocks and, therefore, in units older than 1.7 Ga is at least theoretically possible.

The preservation of commercial quantities of oil at reservoir temperatures of 200 °C and the existence of moderately mature biomarkers in the lowest grade of metasedimentary rocks and, therefore, in units older than 1.7 Ga is at least theoretically possible.

The preservation of commercial quantities of oil at reservoir temperatures of 200 °C and the existence of moderately mature bitumen in source rocks at temperatures of 200–220 °C, though evidently real, are certainly not normal. Such excellent thermal preservation therefore requires exceptional conditions. One favorable condition is rapid heating caused by fast-basin subsidence and/or high geothermal gradients increasing maximum temperatures for hydrocarbon preservation by up to 30 °C (Waples, 2000). A second favorable condition is increased fluid or gas pressure that has the potential to retard petroleum generation, biomarker maturation and hydrocarbon thermal degradation to a significant degree (Fang et al., 1995; Lewan, 1997; Price and Wenger, 1992). The absence of catalysts that promote the degradation of hydrocarbons is a third factor influencing petroleum preservation at high temperatures. Some organometallic complexes and active mineral surfaces have the capacity to induce hydrocarbon cracking at considerably reduced temperatures (Mango, 1990; Mango and Elrod, 1999; Mango and Hightower, 1997; Mango et al., 1994). Mango (1987) has even argued that purely thermally induced, uncatalyzed cracking of higher hydrocarbons to gas is generally an insignificant process. Furthermore, hydrocarbon degradation is strongly retarded and shifted to higher temperatures in the absence of sulfur and organic sulfur compounds that are known to initiate radical chain reactions (Lewan, 1998). Some or all of these exceptional conditions might have prevailed in some Paleoproterozoic and Archean

Figure 4  Distribution of triaromatic steroids (68) in GC-MS m/z = 231 selected ion chromatograms in (a) a Phanerozoic oil of low thermal maturity, (b) a mature Phanerozoic oil, and (c) an overmature bitumen from the late Archaean Fortescue Group in Western Australia. The inset in (c) is a 20× magnification of the elution range of C26 to C28 triaromatic steroids (68b) (Brocks et al., 2003a,b) (reproduced by permission of Elsevier from Geochim. Cosmochim. Acta 2003, in press).
sedimentary basins and might be used as a guide to find preserved biomarkers in low-grade metasedimentary rocks.

8.03.4 EXPERIMENTAL APPROACHES TO BIOMARKER AND KEROGEN ANALYSIS

Analysis of organic matter in sediments generally begins with determination of total organic carbon content (TOC) and an approximate evaluation of thermal maturity and organic matter type using a screening tool such as Rock-Eval pyrolysis (e.g., Espitalié et al., 1977; Peters and Moldowan, 1993). This may be accompanied by organic petrographic analysis, palynology, determination of elemental carbon, hydrogen, oxygen, nitrogen, and sulfur contents and bulk isotope analyses. Bitumen is isolated from the sediment by extraction with solvents such as dichloromethane and methanol and further separated into components of different molecular sizes and polarities by liquid chromatography. Saturated hydrocarbons, aromatic hydrocarbons and a polar fraction with organic oxygen, nitrogen, and sulfur compounds are readily separated from macromolecular material of the asphaltene fraction in this way. The insoluble organic component, kerogen, is then obtained from the rock residue after demineralization with hydrochloric and hydrofluoric acids.

Gas chromatography (GC) and combined gas chromatography-mass spectrometry (GC-MS) are the primary instrumental means for identifying and quantifying biomarkers in the saturated and
aromatic hydrocarbon fractions of bitumen. Mass spectrometers operated in full scan mode provide detailed information of fragmentation pathways and the identity of compounds. When operated to detect selected fragment ions (selected ion monitoring or SIM) there is considerable enhancement of signal to noise and more sensitive detection of trace components. Instruments with tandem mass analyzers (GC-MS-MS) allow compounds to be detected through their most-specific fragmentation reactions (multiple reaction monitoring (MRM)) with further enhancements in detection, provided one knows the structure of the compound being sought and some details of its mass spectrum (e.g., Philp and Ong, 1992; Summons, 1987). Improvements in chromatographic resolution (e.g., multidimensional GC; Reddy et al., 2002) and the high mass-spectrometer scan rates of time-of-flight (TOF) mass analyzers are creating the means to identify more of the thousands of components that occur in fossil-hydrocarbon mixtures. Liquid chromatography and mass spectrometry is opening new windows on the structures and compositions of intact polar lipid mixtures from cultured organisms and environmental samples (e.g., Hopmans et al., 2000; Rutters et al., 2001; Talbot et al., 2001). This new direction has important consequences for improved knowledge of biomarker sources and the identification of high molecular weight compounds that are inaccessible by GC-MS.

Obtaining information about the overall structures and biomarker contents of kerogen and other kinds of macromolecular organic matter is complex and best accomplished with a combination of controlled chemical and pyrolytic degradation techniques and solid-state spectroscopic methods such as Nuclear Magnetic Resonance (NMR) (e.g., Cody et al., 2002; Wilson, 1987; Wilson et al., 1994) and Fourier Transform Infrared Spectroscopy (FT-IR) (e.g., Ganz and Kalkreuth, 1987; Marshall et al., 2001; Solomon and Carangelo, 1987). Successful approaches that target the structures of component biomolecules have been based on analytical pyrolysis techniques (Larter and Horsfield, 1993, and references therein), catalytic hydrolysis (Love et al., 1995) and various types of chemical degradation (e.g., Kohnen et al., 1991b).

8.03.5 DISCUSSION OF BIOMARKERS BY HYDROCARBON CLASS

8.03.5.1 Advantages and Limitations of the Biomarker Approach

In contrast to provenance and authenticity issues faced by paleontologists studying Proterozoic and Archean rocks (e.g., Brasier et al., 2002; Schopf et al., 2002) it is generally straightforward, by virtue of their chemical structures, to recognize when complex hydrocarbon molecules are genuine biogenic remains. However, a more detailed interpretation of these molecular structures is complicated by three factors. The first is the fragmentary knowledge of biomarker distributions across the wide range of extant organisms, second is the presence of compounds that are obviously biogenic but which have no known precursor organism, and third are the uncertainties associated with the extrapolation of the biomarker relationships of extant organisms back in time over hundreds of millions to billions of years. The biological interpretation of molecular fossils is almost exclusively based on the distribution of biomarkers in living organisms. However, the full repertoire of lipid biosynthetic capabilities is only known for a small fraction of microorganisms that have been cultured (Volkman et al., 1993). Therefore, it is probable that some biomarkers have a broader taxonomic distribution and less diagnostic value than is currently accepted. Moreover, pathways for the biosynthesis of particular lipids might have evolved independently in different lineages or could have been acquired by horizontal gene transfer between lineages. Lastly, lipids believed to be diagnostic for specific taxonomic groups might also be representative of unrelated extinct clades.

Some tests are available to verify the accuracy of biomarker assignments. In many cases, unusual biomarkers are associated with organisms known for a particular physiology or biochemical capacity. Examples of this would be an uncommon carbon fixation pathway (e.g., van der Meer et al., 2001), the consumption of methane or capacity to survive anoxia or hypersalinity. These physiologies may be associated with specific isotopic fractionations or with specific geological settings. In such cases, isotopic ratios are particularly valuable in determining whether or not to assign a specific source organism, biochemical process or environmental niche to a particular compound (e.g., Jahnke et al., 1999; Summons and Powell, 1986).

8.03.5.2 n-Alkanes, Algaenans, and other Polymethylenic Biopolymers

n-Alkanes, such as hexadecane (1),

(1) Hexadecane (n-C_{16})

are the most abundant hydrocarbons in all nonbiodegraded oils and mature bitumens. Their potential biological precursors can be
found in virtually all extant organisms. The bulk of n-alkanes in most Phanerozoic bitumens is derived from membrane components such as phospholipids and sphingolipids produced by bacteria and algae, polymeric amphiphilic biopolymers such as algaenans biosynthesized by microalgae (Tegelaar et al., 1989), and waxes introduced by vascular plant debris (Hedberg, 1968). However, despite the ubiquity of straight-chain lipids in the biosphere, some n-alkane profiles can be environmentally and taxonomically diagnostic (Table 2), especially if combined with micro-paleontological and carbon-isotopic analysis (Hoffmann et al., 1987; Rieley et al., 1991). For example, elevated concentrations of n-alkanes with odd carbon numbers between n-C_{15} and n-C_{19} in Ordovician rocks point to the presence of the marine cyanobacterium or alga Gloeocapsomorpha prisca (Fowler, 1992; Hoffmann et al., 1987; and references therein). Long chain n-alkanes with more than ~27 carbon atoms and a predominance of odd-over-even carbon numbers (OEP) are frequently derived from plant waxes indicating a post-Silurian age and organic matter input from terrestrial sources (Hedberg, 1968; Tissot and Welte, 1984).

Also abundant in most crude oils, coals, and bitumens are n-alkanes with more than 40 and up to 110 carbon atoms (del Rio and Philp, 1999; Hsieh et al., 2000; Killops et al., 2000; Mueller and Philp, 1998). In coals and oils from dominantly terrigenous sources, high-molecular weight alkanes are likely to be the diagenetic alteration products of cuticular waxes and plant-derived aliphatic macromolecules such as cutan (McKinney et al., 1996; Nip et al., 1986a,b) and suberan (Tegelaar et al., 1995). However, in sedimentary rocks with little terrigenous organic matter input, algaenans are probably the most important sources for high-molecular weight aliphatic hydrocarbons.

Algaenans are insoluble, nonhydrolyzable, and highly aliphatic macromolecules that serve as a structural component in the cell wall of several marine (Derenne et al., 1992; Gelin et al., 1996) and freshwater (Blokker et al., 1998) green algae (chlorophytes), and marine eustigmatophytes (Gelin et al., 1996, 1999). Chemists are still uncertain about the precise structures of algaenans. Elucidation with various chemical and thermal degradation techniques suggests that the biopolymers comprise mainly linear, long-chain aliphatic building blocks, derived from even-carbon-numbered C_{30} to C_{34} mono- and di-unsaturated α,ω-dialdehydes via an aldolization-dehydration mechanism (Bertheas et al., 1999; Gelin et al., 1994; Metzger et al., 1993). Algaenans likely play an important role in the marine carbon cycle (Derenne and Largeau, 2001; Volkman et al., 1998). Although the full extent to which algaenans are present in photosynthetic microorganisms of marine and lacustrine ecosystems is not known, their resistance against chemical and biological degradation leads to selective preservation during diagenesis (Tegelaar et al., 1989). This recalcitrance also ensures that algaenan is one form of organic matter which may be quantitatively exported from the surface ocean and eventually into sediments where further selectivity in its preservation leads to accumulation in kerogen. Degradation-resistant qualities probably make algaenan and related materials one of the major sedimentary sinks for organic carbon (Derenne and Largeau, 2001; Gelin et al., 1996). With burial of the host rock, and upon cracking of kerogen, algaenans become an important source of crude oil hydrocarbons and are, therefore, relevant to understanding petroleum occurrence (Tegelaar et al., 1989). It is probably not just coincidence that the world’s oldest known commercial deposits of petroleum in Oman and Siberia are from rocks of Late Neoproterozoic to Early Cambrian age (Grantham et al., 1988), which correspond in age to the rising prominence of marine planktonic algae as suggested by a major diversification in acritarchs (Knoll, 1992; Mendelson, 1993; Zang and Walter, 1989). These old oils are geochemically distinctive with high abundances of long-chain methyl-alkanes with chain length and branching patterns consistent with thermal cracking of a specific type of aliphatic biopolymer such as algaenan (Fowler and Douglas, 1987; Höld et al., 1999; Klomph, 1986). An exceptionally high content of 24-ethylcholestanes in the same oils suggests an overwhelming input of green algal (Chlorophycean) biomass to the source kerogen. Resolution of the exact structures of various algaenans might yield a new biomarker tool that could help to recognize major algal groups that contributed to organic matter in sedimentary rocks (Blokker et al., 2000). Different algal groups biosynthesize algaenans with different monomeric units and various modes of linking,
Table 2  Aliphatic and monocyclic saturated hydrocarbons in the molecular fossil record and their paleobiological interpretation.

<table>
<thead>
<tr>
<th>Biomarkers</th>
<th>Biological and environmental interpretation</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>n-Alkanes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outstanding concentrations of ( n-C_{15} ), ( n-C_{17} ), and ( n-C_{19} ) in early Paleozoic rocks</td>
<td>Gloeocapsomorpha prisca, marine phytoplankton of uncertain affinity, probably an alga; identified in Cambrian–Devonian sediments but most prominent in Ordovician. Estonian kukersite is a typical source.</td>
<td>Blokker et al. (2001) and Fowler (1992)</td>
</tr>
<tr>
<td>( &gt; n-C_{27} ) with OEP(^a)</td>
<td>Waxes derived from higher plants; terrestrial input; post-Silurian age.</td>
<td>Hedberg (1968) and Tissot and Welte (1984)</td>
</tr>
<tr>
<td>( &gt; n-C_{40} )</td>
<td>Predominantly degradation products of aliphatic macromolecules such as algaenan (marine, lacustrine), cutan and suberan (terrestrial, plant derived).</td>
<td>Allard et al. (2002) and Killops et al. (2000)</td>
</tr>
<tr>
<td><strong>Branched alkanes and acyclic isoprenoids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monomethylalkanes and dimethylalkanes (MMA and DMA)</td>
<td>Cyanobacteria both cultured and in mat communities from hypersaline and hydrothermal environments.</td>
<td>For example, Dembitsky et al. (2001), Kenig et al. (1995b), Köster et al. (1999), and Shiea et al. (1990)</td>
</tr>
<tr>
<td>5,5-diethylalkanes with OEP(^a) (wrongly reported as 3,7- or 3,6,7-dimethylalkanes)</td>
<td>These structures widely and incorrectly assigned. Chemical synthesis of a 5,5-diethylalkane indicates this is a major series. Often occurs with other alkanes with quaternary carbon centers (BAQC’s). Source organisms not known but commonly found in association with benthic microbial mats.</td>
<td>Arouri et al. (2000a,b), Kenig et al. (2002), Logan et al. (1999), Logan et al. (2001) and Simons et al. (2002)</td>
</tr>
<tr>
<td>Pristane (9) (Pr) and phytane (10) (Ph)</td>
<td>From chlorophylls of cyanobacteria; algae and plants; bacteriochlorophylls ( a ) and ( b ) of phototrophic bacteria; tocopherols; Ph: archaeal membrane lipids.</td>
<td>Peters and Moldowan (1993)</td>
</tr>
<tr>
<td>Regular acyclic isoprenoids (6) ( i_{21} ) to ( i_{30} )</td>
<td>Probable source is halophilic Archaea; abundant in evaporitic environments.</td>
<td>Grice et al. (1998b)</td>
</tr>
</tbody>
</table>

(continued)
Table 2  (continued).

<table>
<thead>
<tr>
<th>Biomarkers</th>
<th>Biological and environmental interpretation</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Squalane (15) (tail–tail C_{30} acyclic isoprenoid)</td>
<td>All organisms produce some squalene; most sedimentary squalane probably from Archaea.</td>
<td>Grice et al. (1998b)</td>
</tr>
<tr>
<td>Crocetane (17)</td>
<td>Archaea (anaerobic methane oxidizers); associated with sub-sea gas, gas hydrate, and mud volcanoes.</td>
<td>Bian et al. (2001) and Thiel et al. (1999)</td>
</tr>
<tr>
<td>PMI (18) (2,6,10,15, 19-pentamethyllicosane)</td>
<td>Methanogenic and methanotrophic archaea.</td>
<td>Elvert et al. (1999), Schouten et al. (1997), and Thiel et al. (1999)</td>
</tr>
<tr>
<td>TMI (2,6,15,19-tetramethyllicosane)</td>
<td>Only reported from a mid-Cretaceous oceanic anoxic event; nonhyperthermophilic marine Crenarchaeota?</td>
<td>Kuypers et al. (2001)</td>
</tr>
<tr>
<td>C_{20}, C_{25}, C_{30} and C_{35} highly branched isoprenoids (19)</td>
<td>Un saturated and polyunsaturated isoprenoid hydrocarbons are prominent biochemicals in some diatom taxa such as <em>Rhizosolenia</em>, <em>Haslea</em>, <em>Pleurosigma</em>, and <em>Navicula</em>.</td>
<td>Sinninghe Damsté et al. (1999a), Volkman et al. (1994), Belt et al. (2000), and Rowland et al. (2001)</td>
</tr>
<tr>
<td>Botryococenes and botryococcanes (20), cyclobotryococenes, polymethylsqualenes</td>
<td>The unsaturated, sometimes cyclic, biogenic hydrocarbons and their saturated fossil counterparts are diagnostic markers of the chlorophyte <em>B. braunii</em> and their preferred habitat of fresh to brackish water.</td>
<td>Huang et al. (1988), Metzger and Largeau (1999), and Summons et al. (2002)</td>
</tr>
</tbody>
</table>

**Monocyclic saturated hydrocarbons**

| C_{12}–C_{16} cyclopentylalkanes (3) with OEP                           | Oils from marine environments; unknown biological source.                                                   | Carlson et al. (1993), and Hsieh and Philp (2001) |
| C_{12}–C_{16} cyclopentylalkanes (3) with no distinct carbon preference | Oils from freshwater lacustrine settings; unknown biological source.                                        |                                                |
| C_{12}–C_{16} cyclopentylalkanes (3) with strong EOP<sup>b</sup>         | Oils from saline lacustrine settings; unknown biological source.                                          |                                                |
| Cyclohexylalkanes (4) without carbon number preference                  | Formed during pyrolysis of biopolymers with long aliphatic carbon chains suggesting an origin from acyclic polymethylenic precursors. | For example, Gelin et al. (1994) |
| Macrocyclic alkanes C_{15}–C_{34}                                      | Bitumens extracted from torbanites containing remains of *B. braunii*; fresh to brackish water.          | Audino et al. (2002)                    |

<sup>a</sup> Odd-over-even carbon number predominance.  <sup>b</sup> Even-over-odd carbon number predominance.
and the resulting structures are resistant and often preserved in sedimentary rocks with minor alterations. Thus, analysis of the structures of sedimentary algaenans and comparison with their counterparts in extant organisms might eventually enable organic matter from different algae to be distinguished, possibly down to the family level (Blokker et al., 2000). Prior to the Neoproterozoic, and the major algal diversifications, aliphatic algaenans probably played a relatively minor role in organic matter accumulation and oil generation (Brocks et al., 2003c).

8.03.5.3 Methyl and Ethyl Alkanes

Acyclic alkanes with one or more sites of branching are notably abundant components of Archean, Proterozoic, and Early Paleozoic bitumens with most reported occurrences being low-molecular weight (C_{14–15}) monomethylalkanes (e.g., Hoering, 1976, 1981; Summons and Walter, 1990). Microbial mat communities, particularly those where cyanobacteria are the predominant organism, are well known for having high abundances and distinctive patterns of short-chain (C_{15–20}) methyl alkanes and are considered to be one of the major sources since these same hydrocarbons have been identified in cyanobacterial cultures (Dembitsky et al., 2001; Köster et al., 1999), as well as modern and ancient sediments with actual or remnant cyanobacterial mat assemblages (Kenig et al., 1995b; Robinson and Eglinton, 1990; Shiea et al., 1995b, 1991; Summons and Walter, 1990). Hydrolysis and decarboxylation of branched fatty acyl bacterial lipids is another possible origin for C_{15–20} methyl alkanes.

As mentioned above (Section 8.03.5.2), a striking feature of some Neoproterozoic to Early Cambrian oils from Oman and Siberia are C_{20+} methylalkanes with the locus of branching located towards the centers of the chains and a marked reduction on their abundances above C_{24} (Höld et al., 1999). These compounds were attributed to C_{24+} precursor lipids with alkyl substituents at C-12 or C-13.

Other Proterozoic sediments contain abundant pseudohomologous series of odd carbon-numbered C_{19–35} branched alkanes that were originally and mistakenly assigned as 3,7-dimethylalkanes on the basis of similar GC and MS data to published literature (Logan et al., 1999, 2001; Mycke et al., 1988). These hydrocarbons have been reported as major components of 1,640 Ma microbial mat sediments in the Barney Creek Formation, Australia (Logan et al., 2001) and the Tanana Formation and correlatives of the Centralian Superbasin, Australia (Arouri et al., 2000a; Arouri et al., 2000b; Logan et al., 1999), while a separate series of branched alkanes, consisting of predominantly even carbon numbers ranging from C_{22} to C_{36} were also found in Barney Creek sediments in association with assemblages of large filamentous microfossils (Logan et al., 2001). Some uncertainties and errors concerning the exact structures of these odd- or even-carbon numbered series of branched alkanes have recently been resolved. Kenig et al. (2001) identified a series of even-numbered carbon monoethylalkanes in a Mesozoic black shale by comparisons with reference mass spectra of ethylalkanes earlier identified by Wharton et al. (1997). Chemical synthesis of a member of another series, namely 5,5-diethylalkanes (Kenig et al., 2002), has led to an appreciation of their widespread occurrence. Mass-spectral and gas chromatographic analysis of related series suggests that branched alkanes with quaternary carbon centers (BAQCs) may be ubiquitous and, although the full range of structures and their biological sources are not established, they appear to be especially abundant in ancient sediments and associated with microbial mats (Kenig et al., 2002; Logan et al., 2001; Simons et al., 2002).

Audino et al. (2001) have reported a unique distribution of branched alkanes ranging from C_{23} to C_{31+} in the extractable organic matter and kerogen of several Peruvian torbanites. Every series begins with the 2-methylalkane. Each member of a particular homologous series has a common alkyl group and each series differs from the next by two carbon atoms. These components were assigned either to an origin from the A-race of Botryococcus braunii based on structural similarities to the botryals biosynthesized by these organisms or by subsequent heterotrophic organisms reworking the Botryococcus braunii biomass.

8.03.5.4 Alkyl Cyclohexanes and Cyclopentanes

Although specific biological sources for alkyl-cyclohexanes (3) are unknown, the distribution
of high-molecular weight homologs in the C41 to C46 range may be a useful tool to obtain information about the depositional environment (Carlson et al., 1993; Hsieh and Philp, 2001) (Table 2). A predominance of odd-over-even (OEP) carbon numbers in the above range seems to indicate petroleum from marine sources, while petroleum hydrocarbons with no distinct carbon number preference or a low even-over-odd (EOP) predominance might have a freshwater origin. A strong EOP of C41 to C46 alkylcyclopentanes may be a useful indicator for oils sourced from saline lake sediments. However, the statistical basis for the above interpretations is still limited and requires a study of a larger set of oils and bitumens from different depositional environments.

\[ R = n\text{-alkyl} \]

(3) Alkylcyclopentanes

\[ R = n\text{-alkyl} \]

(4) Alkylcyclohexanes

\[ R = n\text{-alkyl} \]

(7) Tail-to-tail

n-Alkylcyclohexanes (4), methyl-n-alkylcyclohexanes and related compounds such as alkyl phenols have long been recognized as important components of sedimentary hydrocarbon assemblages. Potential precursors are cyclohexyl fatty acids that are known from some thermophilic and nonthermophilic bacteria (e.g., De Rosa et al., 1971; Suzuki et al., 1981). However, the limited carbon number distributions of these biological lipids compared to the long chain lengths of the cyclohexanes in sediments suggest there are less exotic sources. A wide variety of alkylcyclohexanes has been reported in pyrolysis products of fatty acids, aliphatic polyaldehydes, and algaenans (e.g., Fowler et al., 1986; Gelin et al., 1994; Rubinstein and Strausz, 1979). Moreover, a homologous series of n-alkylcyclohexanes was identified in pyrolysis products of microbial mats (Kenig, 2000) suggesting that they can arise from chemical or thermal alteration of acyclic precursors.

### 8.03.5.5 Isoprenoids

Hydrocarbons formally constructed from repeating C5 isoprene (5) units, are ubiquitous in ancient sediments and petroleum. The most common and abundant of these are the C19 and C20 regularly branched (head-to-tail linking of isoprene units (6)) compounds pristane (9) and phytane (10) which are widely viewed as transformation products of phytol (11), the esterifying alcohol of cyanobacterial and green-plant chlorophylls (e.g., Chlorophyll a (43)) (Didyk et al., 1978). Tocopherols are additional plant and phytoplanktonic sources of pristane (Goosens et al., 1984). Archaeol (12) (diphytanylglycerol)
Discussion of Biomarkers by Hydrocarbon Class

(12) Archaeol

(13) Caldarchaeol

(14) A polycyclic caldarcheol characteristic of nonthermophilic crenarcheota

(15) Squalane

(16) Biphytane

(17) Crocetane

(18) 2,6,10,15,19-Pentamethyllicosane (PMI)
is the most commonly reported core lipid in Archaea, occurring in both major kingdoms, the Euryarchaeota and Crenarchaeota (Kates, 1993; Koga et al., 1993) and is also an important source of sedimentary phytane, especially in samples from extreme environments. Many oils and bitumens, however, also contain varying abundances of C$_{21+}$ regularly branched acyclic isoprenoids which must have originated from C$_{20}$ precursors. Archaea are presumed to be the major source of compounds of this type. Although Albaiges (1980) has reported extended regular isoprenoids with chains as long as C$_{45}$ in oils, there is limited knowledge of their occurrence in cultured organisms. Langworthy et al. (1982) have cited the presence of regular isoprenoid chains as long as C$_{30}$ in the neutral lipid fractions of thermoacidophiles but this does not explain the wider range of carbon numbers in fossil assemblages. The polar ether lipids of extreme halophiles have often been reported to contain the C$_{20}$–C$_{25}$ and C$_{25}$–C$_{25}$ diether analogues of archaeol (12) and are, therefore, a logical source of the C$_{25}$ and lower regular acyclic isoprenoid hydrocarbons that are invariably prominent in bitumens and oils from saline lakes (e.g., McKirdy et al., 1982). Other C$_{21+}$ regular isoprenoids in extant organisms might have remained unnoticed as the majority of lipid profiling studies to date have focused on compounds that are able to be made volatile for analysis by GC-MS. Many of the recently identified archaeal lipids have irregular C$_{40}$ isoprenoid chains and it is quite probable that other, presently unknown, high-molecular weight polar lipid precursors exist but have escaped detection through conventional analytical windows.

Irregularly branched isoprenoids are also prominent sedimentary hydrocarbons. Squalane (15), comprising two tail–tail (7) linked C$_{15}$ isoprenoid moieties, is a very common component of bitumens and oils and, although its logical precursor squalene occurs in most organisms, Archaea are likely to be the predominant sources. Squalane and a variety of unsaturated derivatives are present in the neutral lipid fractions of many Archaea and their abundances are highest in environmental samples with overall elevated acyclic isoprenoid content such as those from saline lakes (McKirdy et al., 1986; ten Haven et al., 1988). The tail-to-tail (7) linked C$_{40}$ isoprenoid lycopane (22) is often detected in lacustrine and marine sediments (e.g., Freeman et al., 1990, 1994; Wakeham et al., 1993) and in particulate organic matter from anoxic water columns (e.g., Wakeham et al., 1993). Feasible precursors include carotenoids of the lycopene (21) family that occur, for example, in anoxogenic phototrophic bacteria (Section 8.03.6.1.4), or the lycopadiene-like precursors produced by algae such as Botryococcus braunii (Derenne et al., 1990; Wakeham et al., 1993).

A major source of biphytane (16), the C$_{40}$ isoprenoid with head-to-head (8) branching, is caldarchaeol (13) (dibiphytanyl—diglycerol–tetraether) which is a prominent core lipid in methanogens (Koga et al., 1993) and members of the kingdom Crenarchaeota (e.g., Kates, 1993). Crenarchaeotes and some methanogens are known to produce polar lipids with variants of the caldarchaeol core with cyclopentane, and, occasionally, cyclohexane rings (e.g., (14)). These complex lipids have been discovered in abundance in filtrates from open ocean waters attesting to the probability that these Archaea are an important component of ocean plankton (DeLong et al., 1998; Sinninghe Damsté et al., 2002a). Biphytane (16) has long been recognized as a prominent sedimentary hydrocarbon (Moldowan and Seifert, 1979) and can have both crenarchaeote and euryarchaeote origins.

Lower molecular weight irregularly branched isoprenoids are also sometimes prominent in sediments and oils. The irregular tail-to-tail (7) linked C$_{20}$ isoprenoid hydrocarbon 2,6,11,15-tetramethylhexadecane (17) (crocetane) and its C$_{25}$ counterpart 2,6,10,15,19-pentamethylicosane (18) (often referred to as PMI, or in older literature PME) are considered diagnostic markers for...
Archaea that are central to the methane cycle. These compounds have been detected in various cultured organisms, microbial communities and sediments comprising methanogenic (e.g., Brassell et al., 1981; Koga et al., 1993; Risatti et al., 1984; Schouten et al., 2001a,b,c, 1997) and methanotrophic Archaea (e.g., Bian et al., 2001; Elvert et al., 1999; Hinrichs et al., 2000; Pancost et al., 2000; Thiel et al., 1999). Crocetane has also recently been reported in crude oils (Barber et al., 2001; Barber et al., 2002; Greenwood and Summons, 2003). PMI appears to be confined to Mesozoic and younger rocks, whereas crocetane probably has a much longer geological record since it has been detected in Triassic, Devonian, and Proterozoic rocks (Greenwood and Summons, 2003).

There are further distinctive classes of isoprenoids which are thought to have quite restricted biological origins. Compounds referred to as “highly branched isoprenoids” or HBIs with C20, C25, and C30 (e.g., (19)) members are biosynthesized by some diatoms (Volkman et al., 1994) and are therefore considered very specific biomarkers for these organisms (Allard et al., 2001; Belt et al., 2000; Robson and Rowland, 1986; Rowland et al., 2001; Sinninghe Damsté et al., 1999a,b; Summons et al., 1993; Volkman et al., 1994). The biomarker botryococcane (20) and related compounds are derived from botryococenes, C30–C37 polymethylated, and polyunsaturated derivatives of an irregularly constructed isomer of squalene, and are only known to be biosynthesized by the green alga Botryococcus braunii (Metzger and Largeau, 1999). Certain strains biosynthesize cyclobotryococenes (e.g., Metzger et al., 1985) and polymethylsqualenes which occur as their saturated hydrocarbon analogs in ancient sediments and oils (Summons et al., 2002).

8.03.5.6 Carotenoids

Carotenoids are usually yellow to red colored lipids formally derived from the C40 isoprenoid lycopene (21) carbon skeleton by varied hydrogenation, dehydrogenation, cyclization and oxidation reactions. In excess of 600 different carotenoid structures have been identified (Britton, 1995). They are biosynthesized de novo by all photosynthetic bacteria, eukaryotes and halophilic archaea, but also occur in a large variety of nonphotosynthetic organisms. Vertebrates and invertebrates have to incorporate carotenoids through their diet, but often have the capacity to generate structurally modified products from ingested precursors (Liaaen-Jensen, 1979). Carotenoids function most commonly as accessory pigments in phototrophs, as pigments for photoprotection, as photoreceptors for phototropism and phototaxis, and as pigments for the coloration of plants and animals (Liaaen-Jensen, 1979). Several hundred natural carotenoids have been described that are distinguished by different cyclic and linear end-groups, and a large variety of functionalities in various positions such as keto, aldehyde, ester, hydroxy, methoxy, and glycoside groups. Many of the functionalized carotenoids extracted from living organisms and recent sediments have been used to obtain information about biological origins, evolution, and ecology (e.g., Britton et al., 1995; Frank et al., 1999; Liaaen-Jensen, 1979; Watts et al., 1977; Xiong et al., 2000). However, the large variety of biological carotenoids, such as (21), (23), and (25) is based on a limited number of different carbon skeletons. Thus, most carotenoids, lose their diagnostic value during diagenesis by reduction of all functional groups and generation of much less specific fossil hydrocarbons such as lycopane (22) and β-carotane (24).
However, some carotenoids retain a taxonomically diagnostic structure during diagenesis and belong to the most important biomarkers for paleoenvironmental reconstructions as discussed below.

8.03.5.6.1 Aromatic carotenoids and arylisoprenoids

The only significant biological source for aromatic carotenoids in aquatic sedimentary environments are phototrophic green (Chlorobiaceae) and purple (Chromatiaceae) sulfur bacteria (Table 3). The growth of most phototrophic sulfur bacteria requires the presence of light and reduced sulfur species in the absence of oxygen. Thus, aromatic carotenoids are often applied as biomarkers for photic-zone euxinia (Koopmans et al., 1996a; Requejo et al., 1992; Summons and Powell, 1986). Okenone (25), the potential precursor of yet undiscovered okenane (26), is exclusively known from planktonic species of Chromatiaceae, while chlorobactane (28), the fossil equivalent of chlorobactene (27) and hydroxychlorobactene, is a biomarker for planktonic as well as benthic mat-forming green pigmented species of Chlorobiaceae. Brown pigmented species of Chlorobiaceae, in contrast, predominantly contain the carotenoids isorenieratene (31) and β-isorenieratene (29), the precursors for sedimentary isorenieratane (32) and β-isorenieratane (30) (Liaaen-Jensen, 1965). As carbon assimilation in Chlorobiaceae follows the reductive or reversed tricarboxylic acid cycle (TCA), their biomass is often distinguished by a strong carbon-isotopic enrichment in $^{13}$C by more than $\sim10\%e$ relative to that of oxygenic photoautotrophs (e.g., Kohnen et al., 1992). The distinctive carbon-isotopic composition of Chlorobiaceae, and the ecology of phototrophic sulfur bacteria are further discussed in Section 8.03.6.1.4.

A second source for aromatic carotenoids are some genera of actinomycetes, such as Mycobacterium and Streptomyces (Krügel et al., 1999). However, the contribution of carotenoids from these organisms to organic matter in aquatic sediments is probably insignificant. A larger variety of aromatic carotenoids also occurs in selected species of marine sponges (Liaaen-Jensen et al., 1982) suggesting the presence of bacterial symbionts. These carotenoids include isorenieratene (31) and β-isorenieratene (29) also found in Chlorobiaceae, but in addition two aromatic structures
Table 3  Aromatic carotenoids and maleimides as indicators for photic zone euxinia.

<table>
<thead>
<tr>
<th>Geological carotenoid</th>
<th>Possible biological precursors</th>
<th>Biological sources</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Okenane (26)</td>
<td>Okenone (25)</td>
<td>Chromatiaceae</td>
<td>Schaeffer et al. (1997)&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Chlorobactane (28)</td>
<td>Chlorobactene (27); hydroxychlorobactene Isorenieratene (32) Isorenieratene (31)</td>
<td>Green pigmented Chlorobiaceae Brown pigmented Chlorobiaceae</td>
<td>Grice et al. (1998c) Bosch et al. (1998), Grice et al. (1996b), Hartgers et al. (1993), Koopmans et al. (1996a), Pancost et al. (1998), Putschew et al. (1998), Simons and Kenig (2001), and Sinninghe Damste´ et al. (2001)</td>
</tr>
<tr>
<td>β-isorenieratane&lt;sup&gt;b&lt;/sup&gt; (30)</td>
<td>β-isorenieratene (29); β-carotene&lt;sup&gt;d&lt;/sup&gt; (23)</td>
<td>Brown pigmented Chlorobiaceae</td>
<td>Grice et al. (1998c)</td>
</tr>
<tr>
<td>Renieratane (34)</td>
<td>Renieratene (33)</td>
<td>Sponges or sponge symbionts? phototrophic sulfur bacteria?</td>
<td>Hartgers et al. (1993), and Schaeﬂé et al. (1977)</td>
</tr>
<tr>
<td>Renierapurpurane&lt;sup&gt;b&lt;/sup&gt; (36)</td>
<td>Renierapurpurin (35)</td>
<td>Sponges or sponge symbionts? phototrophic sulfur bacteria?</td>
<td>Schaeﬂé et al. (1977)</td>
</tr>
<tr>
<td>Palaerenieratane&lt;sup&gt;f&lt;/sup&gt; (37)</td>
<td>Unknown</td>
<td>Chlorobiaceae?</td>
<td>Hartgers et al. (1993), Koopmans et al. (1996a), and Requejo et al. (1992)</td>
</tr>
<tr>
<td>2,3,6-TMAs&lt;sup&gt;g&lt;/sup&gt; (38)</td>
<td>Chlorobactene (27); hydroxychlorobactene; isorenieratene (31); β-isorenieratene (29); β-carotene (23)&lt;sup&gt;d&lt;/sup&gt; and similar structures</td>
<td>Mostly Chlorobiaceae&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Hartgers et al. (1993), Requejo et al. (1992), Summons and Powell (1986), Summons and Powell (1987)</td>
</tr>
<tr>
<td>2,3,4-TMAs&lt;sup&gt;g&lt;/sup&gt; (39)</td>
<td>Okenone (25); renieratene (33) renierapurpurin (35)</td>
<td>Chromatiaceae; Chlorobiaceae?</td>
<td>Summons and Powell (1987)</td>
</tr>
<tr>
<td>3,4,5-TMAs&lt;sup&gt;g&lt;/sup&gt; (40)</td>
<td>Precursor of palaerenieratane (37)</td>
<td>Chlorobiaceae?</td>
<td>Hartgers et al. (1993), Requejo et al. (1992), Summons and Powell (1987)</td>
</tr>
<tr>
<td>Me i-Bu maleimide (49d)</td>
<td>BChl c, d, and e (46)–(48)</td>
<td>Chlorobiaceae, Chloroflexaceae</td>
<td>Grice et al. (1996a, 1997), and Pancost et al. (2002)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Report of okenane as a hydrogenation product of okenone after H2/PtO2 treatment of a polar fraction extracted from a Recent lake sediment. Okenane is unknown as a hydrocarbon biomarker from sedimentary rocks.  
<sup>b</sup> Also including a large variety of di- to pentacyclic early diagenetic cyclization and rearrangement products of isorenieratene (Grice et al., 1996b; Koopmans et al., 1996a).  
<sup>c</sup> Only diagnostic for Chlorobiaceae if the carbon isotopic composition of individual arylisoprenoids shows an enrichment in 13C diagnostic of the reversed tricarboxylic acid cycle.  
<sup>d</sup> According to Koopmans et al. (1996b), β-carotene can undergo aromatization during diagenesis to β-isorenieratane and further degrade to 2,3,6-TMAs.  
<sup>e</sup> New trivial name suggested here (¼ perhydrorenierapurpurin).  
<sup>f</sup> New trivial name suggested here.  
<sup>g</sup> TMA = trimethylarylisoprenoids.
unknown from other organisms: renieratene (33) and renierapurpurin (35). However, sponges are not capable of de novo carotenoid biosynthesis. Therefore, these biomarkers are either generated by the sponge by modification of dietary carotenoids, or they are derived from sponge symbionts, possibly phototrophic sulfur bacteria (Liaaen-Jensen et al., 1982). Although phototrophic sulfur bacteria have not yet been reported as symbionts in sponges, determination of the carbon-isotopic composition of the aromatic carotenoids could confirm their presence. The diagenetic products renieratane (34) and renierapurpurane (36) (= perhydrorenierapurpurin), are rare in the geological record. In the Upper Devonian Duvernay Formation of the Western Canada Basin, renieratane (34) occurs together with isorenieratane (32) (Hartgers et al., 1993). In these particular samples isorenieratane was enriched in $^{13}$C by up to 15‰ relative to aliphatic hydrocarbons and, therefore, clearly derived from Chlorobiaceae. Although the carbon-isotopic composition of renieratane (34) was not reported, its co-occurrence with isorenieratane (32) in the Duvernay Formation indirectly suggests that it is also product of phototrophic sulfur bacteria (Hartgers et al., 1993). The Duvernay Formation also contains a third diaromatic carotenoid (37), here named palaerenieratane, unknown from extant organisms (Hartgers et al., 1993; Requejo et al., 1992). Palaerenieratane (37)
Discussion of Biomarkers by Hydrocarbon Class

(31) Isorenieratene

(32) Isorenieratane

(33) Renieratene

(34) Renieratane

(35) Renierapurpurin

(36) Renierapurpurane

(37) Palaerenieratane
is formally derived from renieratane (34) by dislocation of one methyl group at one of the terminal aromatic rings. In the Duvernay samples, palaerenieratane (37) is strongly enriched in $^{13}$C just as isorenieratane (32). As the enrichment is diagnostic for carbon assimilation via the reversed TCA cycle (Section 8.03.6.1.4), palaerenieratane (37) is also almost certainly derived from either an extinct species, or an as yet undetected species, of extant Chlorobiaceae (Hartgers et al., 1993).

Arylisoprenoids with the 2,3,6-(38), 2,3,4-(39), and 3,4,5-trimethyl (40) substitution patterns are diagenetic and catagenetic cracking products of the above C$_{40}$ aromatic carotenoids (Hartgers et al., 1993; Summons and Powell, 1986, 1987). Therefore, arylisoprenoids are also applied as biomarkers for phototrophic sulfur bacteria. However, 2,3,6-trimethyl aromatic arylisoprenoids (38) are also purported to form by diagenetic aromatization and rearrangement of cyclic, nonaromatic carotenoids (Koopmans et al., 1996b). (38) are therefore more clearly biomarkers for Chlorobiaceae if they also show the carbon-isotopic enrichment typical for the reductive TCA cycle. Aromatic carotenoids might also form a large variety of other rearrangement, cyclization, and degradation products that have diagnostic value for phototrophic sulfur bacteria (Grice et al., 1996b; Koopmans et al., 1996a; Sinninghe Damsté et al., 2001).

8.03.5.6.2 Bacterioruberin

Pigments of the bacterioruberin group (41) are an example for uncommon C$_{50}$ carotenoids. The unique carbon skeleton (42) is biosynthesized by addition of two C$_{5}$ isoprenoid units to the 2 and 2' positions of the C$_{40}$ carotenoid lycopene (21) (Kushwaha and Kates, 1979; Kushwaha et al., 1976). Bacterioruberin is a ubiquitous and abundant, red-orange pigment in moderately (Rønnekleiv and Liaaen-Jensen, 1995) to extremely halophilic archaea (Halobacteria) (Liaaen-Jensen, 1979) (Section 8.03.6.2.3). Located in the membrane of Halobacteria, it plays a role in the photoprotection system (Cockell and Knowland, 1999), but might also be important for the adaptation of membrane fluidity to changing osmotic conditions (D’Souza et al., 1997). Carotenoid pigments with the bacterioruberin skeleton have also been detected in several species of the class Actinobacteria. These include the plant pathogen Curtobacterium flaccumfaciens (Häberli et al., 2000), the psychrotrophic Micrococcaceae Micrococcus roseus (Strand et al., 1997) and Arthrobacter agilis found in Antarctic soil and ice (Fong et al., 2001), and the highly radioresistant Rubrobacter radiotolerans (Saito et al., 1994). In psychrotrophic species, C$_{50}$ carotenoids play an adaptive role in membrane stabilization at low temperature (Fong et al., 2001).

The fossil equivalent of bacterioruberin (41), perhydro bacterioruberin (42), has yet to be discovered in geological samples. However, the abundance and ubiquity of bacterioruberin in Halobacteria, and the occurrence of dense blooms in salt lakes and pools of evaporating seawater, makes (42) a potential, highly diagnostic biomarker for Halobacteria and moderate-to-extreme hypersaline conditions. It is worth bearing in mind that some high-molecular weight biomarkers may have escaped detection because they are difficult to analyze by conventional GC-MS methods.

8.03.5.7 Chlorophylls and Maleimides

The major chlorophyll (Chl) found in all oxygenic photosynthetic organisms, i.e., prochlorophytes, cyanobacteria, and photosynthetic Eukarya, is Chl a (43). However, the partially defunctionalized diagenetic products of Chl a can usually not be distinguished from products of bacteriochlorophyll (BChl) a (44) and b (45), which are mostly derived from anoxygenic phototrophic purple sulfur bacteria. However, BChl c, d, and e (46)–(48) are highly specific. BChl c (46) and d (47) are restricted to green
filamentous bacteria (Chloroflexaceae) and green sulfur bacteria (Chlorobiaceae), while BChl \( e \) is found as a major component only in brown pigmented strains of the Chlorobiaceae.

An elegant methodology to study the input of BChl \( c \), \( d \), or \( e \) into ancient sediments was developed by Grice et al. (1996a). The tetrapyrrole structure of Chl and BChl is only rarely preserved in thermally mature sedimentary rocks. Grice et al. (1996a) observed that Chl and BChl might undergo oxidative degradation to maleimides (49) (1H-pyrrole-2,5-diones), possibly induced by enzymatic activity or light. As the major distinguishing structural characteristics of BChl \( c \), \( d \), and \( e \) in comparison to (46)–(48), in addition to positions C-8, C-12, and C-20; their oxidative degradation will generate a distinctive suite of maleimides. While the major products of Chl \( a \) (43) degradation are indistinct 3,4-dimethyl (49(a)) and 3-ethyl-4-methylmaleimide (49(b)), the oxidation of BChl \( c \), \( d \), and \( e \) (46)–(48) of Chlorobiaceae additionally generates the 3-isobutyl-4-methylmaleimide (49(d)). The diagnostic value of 3-isobutyl-4-methylmaleimide (49(d)) in the Permian Kupferschiefer was confirmed by determination of the carbon-isotopic composition of individual maleimides. 3-methyl-4-propyl (49(c)) and 3-isobutyl-4-methylmaleimide (49(d)) were enriched in \(^{13}\)C by 10–11‰ relative to 3-ethyl-4-methylmaleimide (49(b)) (Grice et al., 1996b). This isotopic enrichment is typical for the reductive TCA cycle, the pathway of \( \text{CO}_2 \) fixation followed by Chlorobiaceae (Section 8.03.6.1.4).

8.03.5.8 Sesquiterpanes (C\(_{15}\)) and Diterpanes (C\(_{20}\))

Bicyclic terpanes are common in oils and bitumens and can have separate origins in bacteria and plants. Compounds of the drimane (50) series, which are ubiquitous and occur in rocks of all ages, are thought to be degradation products of bacteriohopanoids (Alexander et al., 1983). Diterpanes with a far more restricted distribution appear to be derived from vascular plant precursors such as abietic acid. Prominent sedimentary hydrocarbons include bayerane, kaurane, phyllocladane (51), and isopimarane. These compounds and structurally related aromatic hydrocarbons regularly co-occur with resins and other remains of conifers and are therefore considered biomarkers for vascular plants and, more specifically, for gymnosperms (e.g., Noble et al., 1985). Compound-specific isotopic data support the gymnosperm-diterpane relationships (e.g., Murray et al., 1998).

Another important class of diagnostic plant terpenoids is the cadinane group derived from cadinene-based polymers of resinous tropical...
Sedimentary Hydrocarbons, Biomarkers for Early Life

(43) Chl a

(44) BChl a

(45) BChl b

(46) BChl c

(47) BChl d

(48) BChl e

(49) Maleimides

(a) R = Me; (b) R = Et; (c) R = n-Pr; (d) R = isobutyl
angiosperms. Cadinane (52), isomeric bicadinanes (53), and tricadinanes are generated by thermal alteration of polycadinene resins and are putative biomarkers for the Dipterocarpaceae (van Aarssen et al., 1990; Sorsowidjojo et al., 1996).

8.03.5.9 Cheilanthanes and other Tricyclic Polyprenoids

The most common compounds in this class are cheilanthanes (54) (= 13-methyl, 14-alkylpodo-carpanes), tricyclic terpanes that extend from C$_{19}$ to at least C$_{45}$ (Moldowan and Seifert, 1983). These compounds are, on theoretical grounds, derived by cyclization of regular polyprenol precursors (Aquino Neto et al., 1983). The only known natural products that have the cheilanthane skeleton are very unlikely precursors for the ubiquitous hydrocarbon counterparts found in bitumens and petroleum: for example cytotoxins with exotic structures found in sponges (e.g., Gomez Paloma et al., 1997; Manes et al., 1988) and nudibranchs (Miyamoto et al., 1992), and cheilanthatriol, extracted from the fern Cheilanthes, the organism that gave the compound class its name (Khan et al., 1971). Cheilanthanes are “orphan biomarkers” because their actual source remains unknown. While Bacteria have been hypothesized as their precursors, they have been found to occur abundantly in association with tasmanite algae, and are cogenerated with related monoaromatic to triaromatic tricyclic hydrocarbons during pyrolysis of Tasma-nites kerogen (e.g., Aquino Neto et al., 1992; Greenwood et al., 2000; Revill et al., 1994). Therefore, cheilanthanes could be derived from an unusual algal biopolymer. This could explain why feasible precursors have not been found in the extractable lipids of extant organisms using conventional techniques. Tricyclic terpanes occur widely in the geological record but are most abundant in mature shales and their derived oils. More highly cyclized structures based on the same regular polyprenol carbon chain also occur in sediments and have been positively identified by comparison with standards produced by chemical synthesis (Grosjean et al., 2001).

8.03.5.10 Hopenoids and other Pentacyclic Triterpanes

It is often said that hopanoids are “the most abundant natural products on Earth” (Ourisson and Albrecht, 1992) and a major body of work exists on their distributions in sediments, in prokaryotes and in plants. Most commonly, hopanoids are found in select groups of Bacteria, all of which are aerobic (Farrimond et al., 1998; Rohmer et al., 1984). In fact,
Hopanoids were recognized as chemical fossils (e.g., pentakishomohopane (55)) well before their bacterial origins were established. Hopanoids are ubiquitous components in sedimentary organic matter and petroleum of all geological eras. The functional forms of hopanoids in bacteria are the amphiphilic bacteriohopanepolyols (BHP) where a five carbon sugar-derived moiety is C-bound to the C30 pentacyclic hopane skeleton. This C5 unit may have additional sugar, amino acid, or other polar groups attached. It is hypothesized that BHP are the bacterial surrogates of sterols which perform a role as membrane modifiers in eukaryotic cells (Ourisson and Albrecht, 1992; Ourisson et al., 1987).

Although they are known to be synthesized by a wide variety of cultured aerobic bacteria there does not appear to be any obligate requirement for oxygen in their biosynthesis. The biosynthesis and cyclization of squalene to a pentacyclic triterpenoid with a hopane skeleton does not seem to require oxygen and, therefore, hopanoid synthesis might also be possible in anaerobes. For instance, analysis of microbial mats at methane seeps under anoxic Black Sea water revealed the presence of 13C-depleted (δ13C = -78‰) hopanoids with an unusual stereochemistry. This isotopic depletion indicates in situ production and, therefore, suggests that anaerobes are responsible (Thiel et al., 2003).

Besides the apparent paradox of finding BHP only in cultured aerobic bacteria, specific precursor-to-hopane product relationships are very poorly constrained. The major problem in elucidating their sources lies in the huge variety of potential contributing organisms, the low number of these that have been cultured for screening and the relatively low number of individual compounds that have been so far identified in both cultures and in environmental samples. For the vast majority of natural situations, the hopanoid content of a particular sediment- or water-column sample cannot be reliably attributed to any specific source without additional information. Such information might include the amounts of a specific hopanoid known to be contained in different bacteria versus their quantitative importance in a particular setting. Or, it might be the presence of characteristic chemical attributes or stable carbon isotopic compositions as in the case of (57) and (58) (Table 4).

The presence of alkyl substituents on the hopanoid skeleton, for example, A-ring methyl groups, appears to be limited to specific physiological types. For example, methanotrophic bacteria and acetic acid bacteria biosynthesize a range of 3β-hopanoids (Summons and Jahnke, 1992; Zundel and Rohmer, 1985a,b,c). The corresponding 3β-methylhopane hydrocarbons (58) could be derived from either group of bacteria but a profound 13C depletion that has been observed in several of their sedimentary occurrences points to methanotrophic sources being more important (e.g., Burhan et al., 2002; Collister et al., 1992). 2β-Methylhopanoids are produced by many cyanobacteria and have few other demonstrated sources (Bisseret et al., 1985) and, accordingly, it is hypothesized that the corresponding sedimentary 2α-methylhopane (57) hydrocarbons are biomarkers for cyanobacteria (Summons et al., 1999).

It also appears that further clues about hopanoid origins can be drawn from the polar side-chains which carry different numbers and types of substituents. This, in turn, affects their subsequent diagenesis and the types of hopane hydrocarbon, ketone, and other products that are recorded in sediments. In addition to the diagnostic 3β-methyl substituents, hexafunctionalized side-chains are prevalent in Type 1 methanotrophic bacteria (Neunlist and Rohmer, 1985; Zundel and Rohmer, 1985a). The hydroxy substituent at C-31 of these compounds appears to assist oxidative loss of this carbon or the one at C-30, resulting in a predominance of C30 hopane and 30-norhopane products where methanotrophs are prevalent or even dominant (e.g., Burhan et al., 2002; Rohmer et al., 1992). Anomalous 13C-depletion of these hopanoids often observed in sediments and oils is quite consistent with this interpretation (e.g., Summons et al., 2002).

28,30-Dinorhopane (a.k.a. 28,30-bisnorhopane) and 25,28,30-trinorhopane (59) are often very prominent hydrocarbons in sediments from euxinic environments and their derived oils...
<table>
<thead>
<tr>
<th>Biomarkers</th>
<th>Biological and environmental interpretation</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hopanoids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C_{30}-hopanes</td>
<td>Diverse bacterial lineages; few eukaryotic species (e.g., some cryptogams, ferns, mosses, lichens, filamentous fungi, protists).</td>
<td>Rohmer et al. (1984)</td>
</tr>
<tr>
<td>Extended C_{31} to C_{35}</td>
<td>Diagnostic for Bacteria; biosynthesis appears to be restricted to lineages that are not strictly anaerobic (with a possible exception (Thiel et al., 2003)).</td>
<td>Ourisson and Albrecht (1992), Rohmer et al. (1984)</td>
</tr>
<tr>
<td>2\alpha\text{-}methylhopanes (57)</td>
<td>Diagnostic for cyanobacteria and prochlorophytes.</td>
<td>Bisseret et al. (1985), Summons et al. (1999)</td>
</tr>
<tr>
<td>Extended C_{32} to C_{36}</td>
<td>Diagnostic for some microaerophilic proteobacteria (certain methylotrophs, methanotrophs, acetic acid bacteria).</td>
<td>Zundel and Rohmer (1985a,b), (1985c), Summons and Jahnke (1992) Grantham et al. (1980), Peters and Moldowan (1993)</td>
</tr>
<tr>
<td>28,30-Dinorhopane; 25,28,30-trinorhopane TNH (59)</td>
<td>Often prominent in sediments from euxinic environments.</td>
<td></td>
</tr>
<tr>
<td><strong>Steranes and steroids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24-Norcholestane (C_{26})</td>
<td>Possible diatom origin; high concentrations relative to 27-norcholestane indicate Cretaceous or younger crude oil.</td>
<td>Holba et al. (1998a,b)</td>
</tr>
<tr>
<td>Cholestane (66a)</td>
<td>In aquatic sources probably almost exclusively derived from diverse eukaryotes; in organic matter from terrestrial sources (e.g., paleosols) input from soil bacteria of the order Myxococcales conceivable.</td>
<td>Volkman (2003), Bode et al. (2003), Kohl et al. (1983)</td>
</tr>
<tr>
<td>Ergostane (66b), stigmastane (66c)</td>
<td>Exclusively eukaryotic; but usually no distinct sources discernible.</td>
<td>Volkman (2003)</td>
</tr>
<tr>
<td>24-n-propylcholestane (66d)</td>
<td>Pelagophyte algae; a biomarker for marine conditions with few exceptions.</td>
<td>Moldowan et al. (1990)</td>
</tr>
<tr>
<td>4-Methylcholestane (69a); 4,4-dimethylcholestone</td>
<td>Diverse eukaryotic sources; high concentrations likely indicate a dinoflagellate origin.</td>
<td>Volkman (2003)</td>
</tr>
<tr>
<td>4-Methylergostane (69b); 4-methylstigmastane (69c)</td>
<td>Diverse eukaryotic sources; high concentrations likely indicate a dinoflagellate origin.</td>
<td>Volkman (2003)</td>
</tr>
<tr>
<td>Dinosterane (70)</td>
<td>In the Mesozoic and Cenozoic specific for dinoflagellates (with possible minor diatom contribution); in Paleozoic and Neoproterozoic samples probably derived from protodinoflagellates.</td>
<td>Moldowan and Talyzina (1998), Robinson et al. (1984), Volkman et al. (1993)</td>
</tr>
</tbody>
</table>
The association between these compounds and evidence of sulfidic water columns is very strong. Furthermore, it is not easy to rationalize how these compounds could arise by diagenesis of known BHP precursors. Thus, their occurrence is an indicator that some, presently unidentified, bacteria specific to these environments are the ultimate source and we predict they will be found in due course.

The pentacyclic triterpenoid gammacerane (60) occurs in trace amounts in almost all bitumens and oils, but is often abundant in sediments that were deposited under a stratified water column, a condition often observed in lacustrine and hypersaline settings (Sinninghe Damsté et al., 1995). The most likely diagenetic precursor of gammacerane (60) is tetrahymanol (ten Haven et al., 1989). Tetrahymanol has multiple sources. It has been isolated from a fern (Zander et al., 1969), a fungus (Kemp et al., 1984) and the ubiquitous phototrophic purple nonsulfur bacterium Rhodopseudomonas palustris (Kleemann et al., 1990). However, the most likely source for abundant tetrahymanol in sediments is bacterivorous ciliates (e.g., Harvey and McManus, 1991). Predatory ciliates can thrive under oxic and anoxic conditions and are known to graze microorganisms across the oxic–anoxic interface, where they might feed on phototrophic sulfur bacteria or methanotrophs. In some cases it was possible to reconstruct this diverse bacterial diet by measuring the carbon-isotopic composition of ciliate lipids. Gammacerane enriched in $^{13}$C relative to most other lipids in the Miocene Gessos-solfifera Formation suggests that ciliates were partially feeding on green sulfur bacteria (Sinninghe Damsté et al., 1995); and tetrahymanol, strongly depleted in $^{13}$C, extracted from cold-seep sediments from Kazan mud volcano in the eastern Mediterranean Sea indicates that ciliates were probably grazing on methane metabolizing prokaryotes (Werne et al., 2002).

Plant-derived triterpenoids are overtly abundant in sediments from the late Mesozoic onwards. These include oleanane (61), lupane (62), and taraxastane (63). There is a clear relationship between these compounds and triterpenoid precursors such as $\beta$-amyrin (64).
in angiosperms and, consequently, they are considered excellent biomarkers. The appearance of oleanane (61) and its increasing abundance in the Cenozoic is clearly related to the radiation of flowering plants (Moldowan et al., 1994), although there are also numerous diagenetic controls on its occurrence and preservation (e.g., Murray et al., 1997; ten Haven et al., 1992). In fact, the sedimentary distributions of vascular plant triterpanes reflect not only the existence of precursor triterpenoids but the outcome of numerous, kinetically controlled diagenetic reactions. The ultimate preservation of just a few of the most thermodynamically stable isomers probably masks much more diverse contributions from the original biological precursors (Rullkötter et al., 1994; ten Haven et al., 1992).

8.03.5.11 Steroid Hydrocarbons

Sterols, such as cholesterol (65), are essential lipids in all eukaryotic organisms (Table 4). They are often quantitatively important components in membranes where they control membrane permeability and rigidity. Recent sediments contain an extensive variety of different functionalized sterols characterized by the position and number of double bonds, hydroxy groups, alkyl and various other substituents. Many are widespread among eukaryotes, but a considerable number is diagnostic for certain taxonomic groups (e.g., Volkman, 2003). Although double bonds and heteroatomic groups are commonly lost during diagenesis, it is still possible in mature sedimentary rocks to distinguish fossil steroids with different alkyl substituents (e.g., (66a)–(66e)). Furthermore, it has been established that zooplankton feeding on phytoplankton do not alter the stable carbon-isotopic compositions of lipids such as the sterols (Grice et al., 1998b), therefore stable carbon-isotopic composition of steranes in sedimentary material is assumed to be unaltered. There is also evidence that phytoplankton grazing by zooplankton has only a minor impact on the composition of sterols present in fecal pellets (e.g., Mejanelle et al., 2003). Steranes (66), diasteranes (67) and aromatic steroids (e.g., (68)) with 26 to 30 carbon atoms are abundant in most oils and bitumens from the Cenozoic to the Paleoproterozoic (Summons and Walter, 1990) and possibly the Archaean (Section 8.03.9.2).

During diagenesis and catagenesis the biological stereospecificity of sterols, particularly at C-5, C-14, C-17, and C-20, is usually lost (see structure (66) for numbering) and a diverse range of isomers is generated. The nomenclature for the structural and stereoisomers used in the literature, and also in this review, requires a short explanation. The term αββ sterane (sometimes just αβ) is commonly used as short-hand to denote steranes with the 5α(H), 14β(H), 17β(H) configuration, while ααα sterane refers to those with 5α(H), 14α(H), 17α(H) stereochemistry. The notation 14α(H) indicates that the hydrogen is located below the plane of the paper whereas in 14β(H) it is above the plane. In steranes, if no other carbon number is cited, S and R always refer to the stereochemistry at C-20. The prefix “nor”,

\[
\text{Skeleton (63) Taraxastane} \\
\text{Skeleton (64) β-Amyrin} \\
\text{Skeleton (65) Cholesterol} \\
\text{Skeleton (66) Steranes} \\
\text{(a) } R = \text{H (cholesterol)}; \text{ (b) } R = \text{Me (ergostane)}; \\
\text{ (c) } R = \text{Et (stigmastane)}; \text{ (d) } R = n-\text{Pr (24-n-propylcholesterol)}; \\
\text{ (e) } R = i-\text{Pr (24-isopropylcholesterol)}. \\
\]
as for example, in 27-norcholestanate, indicates that the molecule is formally derived from the parent structure by loss of the indicated carbon atom, i.e., in the above example, C-27 is removed from cholestanate (66a). The term “desmethylsteranes” is sometimes used to distinguish steroids that do not possess an additional alkyl group at ring A, i.e., at carbon atoms C-1 to C-4. Diasteranes refers to hydrocarbons with the distinctive structure (67). Diasteranes have no direct biological precursors (Ourisson, 1994) and form by diagenetic rearrangement of sterols or sterenes (Sieskind et al., 1979). The rearrangement is probably catalyzed by clay minerals, regularly leading to elevated concentrations of diasteranes in petroleum derived from clay-rich source rocks (van Kaam-Peters et al., 1998) (Section 8.03.7.4). Finally, monoaromatic and triaromatic steroids (68) form either by diagenetic alteration of unsaturated and polyunsaturated steroids or by dehydrogenation of steranes during catagenesis (de Leeuw and Baas, 1986; Moldowan and Fago, 1986).

Desmethylsteranes with 26–30 carbon atoms have a large number of different sources. C_{26} steranes are ubiquitous in sedimentary rocks, although usually in relatively low concentrations. Moldowan et al. (1991b) have identified three series of C_{26} steranes. The 21- and 27-norcholestanes have apparently no direct biological precursors and are probably degradation products of steroids with higher carbon numbers. In contrast, the third series, 24-norcholestanes, probably has a direct biological source as corresponding sterols are commonly found in recent marine sediments. Circumstantial evidence points to diatoms, or at least to organisms or diagenetic processes associated with diatom blooms (Holba et al., 1998a). A marine algal origin for these sterols is corroborated by the compound-specific radiocarbon ages of C_{26}–C_{29} sterols in shallow marine sediments (Pearson et al., 2000, 2001). The abundance of 24-norcholestanes relative to 27-norcholestanes in crude oils increases considerably from the Jurassic to the Cretaceous and again in the Tertiary, a distribution that appears to coincide with diatom radiation and deposition of major diatomaceous sediments (Holba et al., 1998b). Therefore, the abundances of 24-norcholestanes relative to the more common 27-nor isomers is considered to be an age-diagnostic marker for post-Jurassic oils and bitumens.

Desmethylsteranes with 27–29 carbon atoms are the most abundant steranes and occur in virtually all bitumens and oils that are not overmature. Biological precursors of cholestanate (66a) (C_{27}) are common in animals and red algae (Rhodophyceae), while precursors of ergostane (C_{28}) (66b) are frequently found in yeast and fungi, diatoms (Bacillariophyceae), and several other classes of microalgae (Volkman, 2003). Sterols with the stigmastane skeleton (C_{29}) (66c) typically occur in higher plants (Volkman, 1986), but are also the major sterols in many microalgae, such as several freshwater eustigmatophytes and chrysophytes, and green algae of the class Chlorophyceae. Unfortunately, the C_{27} to C_{29} desmethylsteranes are not characteristic for any specific taxon, because the precursors are widely distributed in the domain Eukarya. Even related species within the same class may contain major sterols with different carbon numbers or even mixtures of all three carbon skeletons (Volkman, 1986, 2003, 1980).

Highly specific, on the other hand, are the C_{30} desmethylsteranes (66d) and (66e). 24-n-propylcholestane (66d) is even regarded as one of the most specific indicators for marine conditions (Moldowan et al., 1985). Its potential biological precursors have only been detected in five marine algae of the class Pelagophyceae. These include the “brown tide” algae Aureoumbra (Giner and Li, 2000; Giner et al., 2001) and Aureococcos (Giner and Boyer, 1998) and three species of the order Sarcinochrysidales (Moldowan et al., 1990; Raederstorff and Rohmer, 1984) (Sarcinochrysidales were previously grouped with the class Chrysophyceae but were reclassified into the new class Pelagophyceae (Saunders et al., 1997)). Sterols with the 24-isopropylcholestane skeleton (66e) are only abundant in extant demosponges. Therefore, 24-isopropylcholestane (66e) in sedimentary rocks is generally attributed to the contribution of sponges (McCaffrey et al., 1994b). The ratio 24-isopropylcholestane/24-n-propylcholestane is high in the terminal Proterozoic to Ordovician but low in all following periods. This distribution might reflect the radiation of early sponges or sponge-related organisms that were the dominant reef builders during this time (McCaffrey et al., 1994b).

Steranes (66a)–(66e)) with alkyl substituents at C-2 or C-3 are ubiquitous in oils and bitumens of all ages (Summons and Capon, 1988). 2- and 3-methylsteranes are usually most abundant, but many oils also contain alkyl substituents at C-3 with up to seven and possibly more than ten carbon atoms (Dahl et al., 1995). Biological steroids with an alkyl substituent in 2- or 3- position have not been observed in extant organisms, and a direct biological source seems unlikely. Instead, 2- and 3-alkylsteranes probably form by addition of a substituent to diagenetically-formed Δ^2-sterenes, possibly mediated by heterotrophic organisms (Summons and Capon, 1991). It is possible that pentose and hexose sugars are important reactants in this process, as substituents with five and six carbon atoms are particularly abundant in some samples (Dahl et al., 1992; Schouten et al., 1998b). Moreover, by desulfurization of the polar fraction of oils it was possible to show that the diagenetic
precursors of these alkyl substituents originally carried multiple functionalities, consistent with a sugar origin (Dahl et al., 1992). The analysis of the sedimentary processes that lead to the formation of steroids functionalized at C-2 or C-3 could, in principle, lead to the discovery of heterotrophic organisms that mediate the reaction. In this case, 2- and 3-alkylsteranes might eventually gain biomarker status.

A third series of alkylsteranes (69) common in bitumens and oils carries a methyl group at C-4. Sterols with the corresponding carbon structure are ubiquitous in eukaryotic organisms because 4-methylsterols and 4,4-dimethylsterols (e.g., lanosterol and cycloartenol) are intermediates in the biosynthesis of all other sterols (Volkman, 2003). However, the usually low concentration of these reaction intermediates suggests that their contribution to sedimentary organic matter is not significant (Volkman et al., 1990). The most important source for sedimentary 4-methylsteranes (69) appear to be dinoflagellates. Dinoflagellates contain relatively high concentrations of sterols with the 4-methyl cholestane (69a), 4-methylergostane (69b) and 4-methylstigmastane (69c) skeletons (e.g., Piretti et al., 1997; Robinson et al., 1984; Volkman et al., 1999). Although dinoflagellates are probably the only significant origin of 4-methylsteranes (69) in the majority of sedimentary rocks, multiple other potential sources are known (Volkman, 2003). Sterols with the 4-methylergostane (69b) and 4-methylstigmastane (69c) skeletons have been isolated from a slime mold (Nes et al., 1990) and further potential precursors for 4-methylstigmastanes (69c) occur in the Pavlovales order of haptophyte algae (Volkman et al., 1990). Methylo trophic bacteria of the family Methylococcales biosynthesize sterols with the (69a) structure (Bird et al., 1971; Schouten et al., 2000), and potential (69)-precursors were also detected in red algae (Beastall et al., 1974), higher plants (Menounos et al., 1986; Yano et al., 1992) and fungi (Me´ janelle et al., 2000). Therefore, low relative concentrations of regular 4-methylsteranes (69) are not specific for any particular taxon, but high concentrations likely indicate biomarker contribution from dinoflagellates.

A distinct group of 4-methylsteranes, the dinosteranes (4α,23,24-trimethylcholestanes (70)), possess a unique side-chain alkylation pattern with an additional methyl group at C-23. Dinosteranes are regarded as very sound biomarkers for dinoflagellates (Robinson et al., 1984; Summons et al., 1987). Their biological source, dinosterol and related compounds, (Robinson et al., 1984) are the most abundant sterols in the majority of dinoflagellates species (Volkman, 2003, and references therein). The only other organism that is known to contain sterols with the dinostan skeleton is a single diatom species (Nichols et al., 1990; Volkman et al., 1993). Although dinosteranes (70) and triaromatic dinosteroids
Myxococcales appear to generate C_{27}-cholester-tin pattern and do not have the capacity to 4,4-dimethyl sterols with an uncommon unsaturation. Methylosphaera (both Methylococcaceae) (Bird, 2003) and Brocks bacteria was critically reviewed by Volkman (2003). The only known bacteria with the unequivocal capacity for de novo sterol biosynthesis appear to be the methylotrophic bacteria Methylococcus and Methylophphaera (both Methylococcaceae) (Bird et al., 1971; Schouten et al., 2000), and several species of soil bacteria of the order Myxococcales, for example, Nannocystis exedens (Bode et al., 2003; Kohl et al., 1983). However, the Methylococcaceae synthesizes exclusively 4-methyl and 4,4-dimethyl sterols with an uncommon unsaturation pattern and do not have the capacity to alkylate the sterol side chain at C-24. Some Myxococcales appear to generate C_{27}-cholester-tinoids, but they also do not have the biosynthetic capacity to alkylate the side-chain. Sterols have also been detected repeatedly in cyanobacterial cultures, although only in trace amounts. It appears now likely that these low quantities were introduced by eukaryotic culture contamination, probably fungi (Summons et al., 2001). Moreover, the complete DNA sequence data of several cyanobacterial lineages are available now, and they do not indicate that cyanobacteria possess the genes required for full sterol biosynthesis (Volkman, 2003). The same criticism (Volkman, 2003) applies to sterols allegedly biosynthesized by mycobacteria.

In conclusion, sterol biosynthesis in Bacteria is probably limited to a small number of taxa that either have an incomplete sterol biosynthetic pathway or lack the capacity to alkylate the side chain. Therefore, steranes ((66b)–(66e) and (69b) and (69c)) in bitumens and oils can be reliably attributed to the activity of eukaryotic organisms (contra Cavalier-Smith, 2002). Additionally, some steranes with a diagnostic alkylation pattern ((66d)–(66e), and (70)) have taxonomic value below domain level.

8.03.6 RECONSTRUCTION OF ANCIENT BIOSPHERES: BIOMARKERS FOR THE THREE DOMAINS OF LIFE

8.03.6.1 Bacteria

8.03.6.1.1 Hopanoids as biomarkers for bacteria

The first extensive survey of biohopanoids in bacteria (Rohmer et al., 1984) indicated that biosynthesis of this important class of biomarkers was the province of aerobic bacteria. Subsequent research, using new screening approaches such as liquid chromatography–mass spectrometry (LC-MS) to measure intact polar lipid structures, has verified their widespread occurrence in cultured aerobic bacteria and environmental samples (e.g., Farrimond et al., 1998; Talbot et al., 2001). It also showed that the isoprenoid building block of the hopanoid skeleton was produced by a biosynthetic pathway, the methyletheral phosphates “MEP” pathway, new to science (Rohmer et al., 1993) as well as other distinctive biochemistries (Rohmer, 1993). These discoveries illustrate the extent to which specific aspects of lipid biosynthetic pathways can also function as a biomarker. Since it would be an impossible task to measure the hopanoid contents of all bacteria growing under the full diversity of natural situations, we have to look to other methods to extend our knowledge of these biomarkers. One of the most promising approaches flows from studies of the DNA that codes for enzymes of key biosynthetic pathways and making use of the genomes of cultured organisms and sequences of DNA cloned from natural environmental samples. Although triter-penoids with a hopane skeleton occur in some plants, hopanoids with an extended side chain (i.e., C_{35} bacteriohopanes (55)) have only ever been found in the Bacteria (Section 8.03.5.10).

8.03.6.1.2 Cyanobacteria

Many, but not all, cyanobacteria biosynthesize bacteriohopanepolyol (56) (BHP) (e.g., Rohmer et al., 1984). As discussed above (Section 8.03.5.10), cyanobacterial hopanoids have
a number of distinctive attributes such as specific polar side-chain groups and sometimes an additional methyl substituent at position 2 of the hopane skeleton (Bisseret et al., 1985) that makes them readily distinguished from other hopanoids (e.g., Summons and Jahnke, 1990). It is these 2-methylhopanoids (57) that are recognized in ancient sediments and oils as being largely of cyanobacterial origin (Summons et al., 1999). A survey of cultured cyanobacteria indicates that biosynthesis of both BHP and 2-Me-BHP is widely and evenly distributed through cyanobacterial phylogeny (L. Jahnke, personal communication). However, compared to their freshwater counterparts, cyanobacteria from saline and hypersaline environments are poorly studied in this respect and this is an obvious target for further research.

Monomethyl and dimethylalkanes in the range C_{16}-C_{20} are prominent in many cultured cyanobacteria as well as most cyanobacterial mat communities that have been studied (Section 8.03.5.3). No specific physiological role has been assigned to these hydrocarbons. Because they have probably multiple origins in ancient sediments and petroleum, these monomethyl and dimethylalkanes alone probably have limited chemotaxonomic specificity. However, they may be very useful in multivariate approaches for linking isotopic and molecular-structure data for a less ambiguous identification of sedimentary cyanobacterial lipids.

8.03.6.1.3 Methanotrophs, methylotrophs, and acetic acid bacteria

These are further groups of aerobic bacteria that produce distinctive hopanoids in abundance. In this case the distinctive features are an additional methyl group at C-3 of the hopane skeleton or the degree of functionality of the polar side chain (Farrimond et al., 2000; Zundel and Rohmer, 1985b) and these hopanoids are also easily distinguished from other series on the basis of their GC-MS or LC-MS behavior (e.g., Summons and Jahnke, 1992; Talbot et al., 2001). In the case of hopanoids from methanotrophic bacteria, an additional signature for their physiology can be a depletion in 13C content compared to co-occurring compounds (Jahnke et al., 1999; Summons et al., 1994a). This isotopic characteristic is preserved along with the diagnostic carbon skeleton in sedimentary hydrocarbons (58) from communities supported by methane oxidation (e.g., Burhan et al., 2002). Some methylotrophic bacteria are also unusual in having the capacity to simultaneously biosynthesize hopanoids along with 4-methyl and 4,4-dimethylsterols (Section 8.03.5.11) with both groups of compounds recording comparable isotopic depletion (Ourisson et al., 1987; Summons and Capon, 1988).

8.03.6.1.4 Phototrophic sulfur bacteria

Anoxygenic phototrophic bacteria are a taxonomically very heterogeneous group. Based on phenotypic criteria they are divided into heliobacteria, purple nonsulfur bacteria, green filamentous bacteria (Chloroflexaceae), green sulfur bacteria (Chlorobiaceae), and purple sulfur bacteria (Chromatiaceae and Ectothiorhodospiraceae; Imhoff, 1995). Among these groups, diagnostic and geologically stable hydrocarbon biomarkers are known for the Chromatiaceae and Chlorobiaceae (Sections 8.03.5.6.1 and 8.03.5.7). As purple and green sulfur bacteria are highly specialized organisms, these biomarkers provide important paleoenvironmental tools. To form blooms, purple and green sulfur bacteria require reduced sulfur species and light. They represent the only known indicators for euxinic conditions in the photic zone of ancient lacustrine and marine environments.

The Chlorobiaceae form a monophyletic group, separated from other phototrophs (Figure 1). Brown pigmented strains contain bacteriochlorophyll e (48) (BChl e) and the major specific carotenoids isorenieratene (31) and β-isorenieratene (29). Green strains obtain their distinctive color from BChl c (46) or d (47) (Section 8.03.5.7) and the diagnostic carotenoids chlorobactene (27) and dihydroxyclochlorobactene (Imhoff, 1995) (Section 8.03.5.6.1). Chlorobiaceae are strictly anaerobic, obligate phototrophs that utilize only photosystem I (PS I). In contrast to cyanobacteria that have the capacity to oxidize water, green sulfur bacteria require sulfide or other reduced sulfur species as the electron donor. CO_2 is the sole carbon source and is assimilated via the reductive or reversed TCA cycle. This mode of carbon fixation gives biomarkers of Chlorobiaceae a diagnostic isotopic fingerprint. Isorenieratane (32), chlorobactane (28), and other biomarkers are often enriched in 13C by ~10‰ relative to organic matter from co-occurring oxygenic phototrophs. According to 16S rRNA (ribosomal RNA) analyses, purple sulfur bacteria form a well separated group in the γ-subgroup of Proteobacteria. Several genera of the family Chromatiaceae contain the taxonomically diagnostic monomaromatic carotenoid okeneone (25). Although okeneone has been extracted from recent sediments (Schaeffer et al., 1997), the equivalent fossil hydrocarbon okeneane (26) has surprisingly not been reported from sedimentary rocks. However, if okeneane is discovered it should be possible to establish its specific biological origin by

Reconstruction of Ancient Biospheres
95
measuring the carbon-isotopic composition. CO₂ fixation in Chromatiaceae, as in cyanobacteria and algae, follows the Calvin-Benson cycle. However, the CO₂ utilized by the Chromatiaceae originates in the anoxic zone of the water column and usually carries the distinct carbon isotopic depletion of remineralized organic matter. Accordingly, the biomass of Chromatiaceae should be depleted in ¹³C relative to organic matter derived from oxygenic phototrophs. Moreover, biomarkers of Chlorobiaceae, using the same source of ¹³C depleted CO₂ but following the reversed TCA cycle, should be strongly enriched in ¹³C relative to biomarkers of Chromatiaceae. Accordingly, it has been observed that okenane from recent sediments is depleted in ¹³C by ~20‰ relative to isorenieratene derived from Chlorobiaceae (Schaeffer et al., 1997). A similar depletion is predicted for okenane (26) relative to isorenieratane (32) extracted from sedimentary rocks.

8.03.6.2 Archaea

The Archaea is often considered Life’s “extremist” domain because of their overwhelming presence in volcanic vent systems, strongly acidic and alkaline springs, evaporitic settings, and in deep-subsurface sediments (e.g., Rothschild and Mancinelli, 2001). However, recent research is showing that archaeans are also quite abundant in the picoplankton of the open ocean (e.g., DeLong, 1992; DeLong et al., 1998). The two broad metabolic themes of archaea, both of which rely on molecular hydrogen as an energy source, make them an important driving force in biogeochemical cycles. In the Euryarchaeota, CO₂ is the predominant electron acceptor and methane the predominant product. In the Crenarchaeota, there is a strong bias toward the oxidation of molecular hydrogen using sulfur compounds as electron acceptors. Given the importance of these processes in biogeochemical cycling it is not surprising, therefore, that biomarkers from archaea are widely and abundantly present in environmental samples (e.g., Sinninghe Damsté et al., 2002a) as well as bitumen and petroleum (e.g., Moldowan and Seifert, 1979).

8.03.6.2.1 Methanogens

Methanogen lipids have been intensively studied and characterized due to their structures being one of the most remarkable features that distinguish the Archaea from all other organisms (Woese et al., 1990). The polar lipids of methanogens comprise both di- and tetra-ethers of glycerol and isoprenoid alcohols with most compounds being based on the core lipids archaeol (12) or caldarchaeol (13). Minor core lipids are sn-2- and sn-3-hydroxyarchaeol and macrocyclic archaeol (Koga et al., 1993). As discussed earlier (Section 8.03.5.5), nonpolar lipids are also distinctive with many methanogens having high contents of hydrocarbons including the characteristic irregularly branched compound PMI (18) and structurally related analogs (e.g., Risatti et al., 1984; Schouten et al., 1997; Tornabene et al., 1979).

8.03.6.2.2 Biomarkers and ecology at marine methane seeps

The advent of compound-specific isotope analysis (CSIA) has forever altered the way geochemists approach the analysis of sediments. One example of this is the capacity to screen environmental samples for the carbon-isotopic signatures of a process as opposed to the traditional mode of analyzing for the diagnostic molecular structures. CSIA of lipids from near-surface sediments of the Kattegat Strait draining the Baltic Sea showed extreme ¹³C-depletion of a chromatographic peak, normally attributed to phytane, within a zone corresponding to high rates of sulfate reduction and concomitant methane oxidation. Closer inspection revealed the localized occurrence of the hydrocarbon crocetane (17), hitherto rarely reported isomer of phytane (Bian et al., 2001) and pointed to a relationship between organisms biosynthesizing crocetane and the oxidation of methane with sulfate as the terminal electron acceptor. This had been hypothesized on the basis of other geochemical indicators (Hoehler et al., 1994). The introduction of gene surveys of small subunit ribosomal RNA (16S rRNA) in concert with data from lipid biomarkers and their individual isotopic compositions further demonstrated that the anaerobic oxidation of methane (AOM) was conducted by archaea in close association with sulfate-reducing bacteria (Hinrichs et al., 1999) and that the process was characterized by distinctive assemblages of lipids such as sn-2 hydroxyarchaeol, crocetane (17), PMI (18) from the Archaea and nonisoprenoid branched fatty acids and ether lipids from the sulfate-reducing partners (Elvert et al., 1999; Hinrichs et al., 1999, 2000; Pancost et al., 2000; Thiel et al., 1999). Recent lipid work suggests thermophilic archaea can mediate anaerobic oxidation of methane in environments with steep geothermal gradients (Schouten et al., 2003).

The above studies established a precedent for the combined use of microbiological, genomic, and isotopic methods to study important biogeochemical processes. Furthering the revolution of culture-independent methods for studying these processes on natural samples has been the visualization of the active microbes through fluorescence in situ hybridization (Boetius et al., 2000),
8.03.6.2.3 Halobacteria

Halophiles are chemo-organotrophic Euryarchaeota that are often the predominant organisms in salt lakes, pools of evaporating seawater, solar salterns and other hypersaline environments with salt concentrations as high as halite saturation (e.g., Oren, 2002). Their lipids closely resemble those of methanogens with the principal difference (e.g., Oren, 2002) with likely implications for rapid climate change (Hinrichs et al., 2003) and with directly associated biomarkers that facilitate studies of its occurrence in the geological past (e.g., Greenwood and Summons, 2003; Thiel et al., 1999).

8.03.6.2.4 Marine Crenarchaeota

An amazing example of a massive occurrence of archaea was reported by Kuypers et al. (2001). They discovered that black shales from the Mid-Cretaceous Oceanic Anoxic Event OAE1b contained an unusual assemblage of cyclic and acyclic isoprenoids including a lipid (14) diagnostic for nonthermophilic Crenarchaeota. Kuypers et al. (2001) calculated that up to 80% of the sedimentary organic matter deposited during this event was derived from nonthermophilic Crenarchaeota. The archaeal biomarkers were enriched in $^{13}\text{C}$ by more than 10‰ relative to algal lipids. This isotopically heavy biomass is not only responsible for the positive carbon-isotopic excursion of organic matter during OEB1b, but also suggests that the marine Crenarchaeota did not live heterotrophically but followed a chemoheterotrophic metabolism (e.g., Høef et al., 1997; Kuypers et al., 2001; Pearson et al., 2001; Sinninghe Damsté et al., 2002a). Further work on elucidating the precise structures of lipids from Crenarchaeota is underway (e.g., Sinninghe Damsté et al., 2002b) and will likely result in profound new insights into the paleobiology and biogeochemistry of this group of Archaea.

8.03.6.3 Eukarya

Eukaryotes are an ancient clade. The oldest acritarchs that are clearly eukaryotic come from shales of the 1.49–1.43 Ga Roper Group, McArthur Basin, Australia (Javaux et al., 2001) and the oldest body fossils believed to have eukaryotic affinity have been found in rocks ~1.8–1.9 Gy old (Figure 5) (Hofmann and Chen, 1981; Zhang, 1986; Han and Runnegar, 1992; see Schneider et al., 2002 for an up-to-date age of the Negaunee Iron-Formation). Biomarker evidence for eukaryotes comes from steranes with diagnostic alkylolation patterns in the side chain ((66b)–(66d)) extracted from ~1.64 Gy old rocks of the Barney Creek Formation (Summons et al., 1988b) and possibly the ~2.7 Ga Fortescue Group (Brocks et al., 1999) (Section 8.03.9.3), both in Australia (Figure 5).

Extant eukaryotes contain thousands of natural products that are only found in members of their domain. Although much information in these molecules is lost in the diagenetic transition to hydrocarbon fossils, many retain a structure based on a specific carbon skeleton (Figure 1). Diagnostic hydrocarbon skeletons include ergostane (66b) and stigmastane (66c) for Eukarya as a whole, 24-n-propylcholestone (66d) for pelagophyte algae, dinosteranes (70) for dinoflagellates and, possibly, a few diatoms (Section 8.03.5.11), and botryococccane (20) for the chlorophyte Botryococcus (Section 8.03.5).

The vast majority of biomarkers that can be traced to distinct branches in the eukaryotic tree belong to higher plants, for example, oleanane (61) (e.g., Moldowan et al., 1994; Murray et al., 1997), taraxastane (63) (e.g., Perkins et al., 1995), and bicadinane (53) (e.g., Cox et al., 1986; van Aarssen et al., 1992). Bicyclic and tricyclic diterpenoid compounds such as abietic acid are major components of conifer resins (e.g., Sinninghe Damsté, 1977). These are the proposed biological precursors of sedimentary diterpane biomarkers retene, simonellite, phyllocladane (51), kaurane, bayerene, and many others (e.g., Alexander et al., 1988, 1992, 1987; Sinninghe Damsté, 1977; Noble et al., 1985, 1986; Otto and Sinninghe Damsté, 2001, 2002). For example, retene has been detected in high relative abundance in Tertiary carbonaceous shales and has been attributed to Podocarpaceae and Araucariaceae conifer resins (Villa et al., 1988). The biomarker cadalene occurs widely in recent and ancient sediments (e.g., Noble et al., 1991; Wang and Sinninghe Damsté, 1990). Cadinenes and cadinols in plants, bryophytes, fungi, and extant and fossils plant resins (e.g., Grantham and Douglas, 1980; van Aarssen et al., 1990) are the proposed precursors for cadalene (Sinninghe Damsté et al., 1986). Surprisingly, there appears to be only one
Figure 5 Geological timescale with important biological events, and observations of well-preserved Precambrian biomarkers (gray) and crude oil (black). (a) Xiao et al. (1998); (b) Knoll (1992); (c) Jackson et al. (1986); (d) Hofmann and Chen (1981); (e) Hofmann (1976); (f) Buick (1992); (g) Hayes (1983); (h) Rasmussen (2000); (i) Shen et al. (2001); (j) Buick et al. (1981), Walter et al. (1980); (k) Rosing (1999); (l) Arouri et al. (2000a,b); Logan et al. (1997); (m) Jiang et al. (1995); (n) e.g., Grantham (1986); Klomp (1986); McCaffrey et al. (1994b); Summons et al. (1999); (o) e.g., Fowler and Douglas (1987); McCaffrey et al. (1994b); Summons et al. (1988b); (p) Peters et al. (1995); (q) Logan et al. (1999, 1997); Summons and Powell (1991); (r) Summons et al. (1988a, 1999); Höld et al. (1997); (s) Wang (1991); Wang and Simoneit (1995); (t) Logan et al. (1997); McCaffrey et al. (1994b); Summons et al. (1988a); Summons et al. (1999); Summons and Powell (1991); (u) Ho et al. (1990); Pratt et al. (1991); (v) Brocks et al. (2003c); Crick et al. (1988); George and Ahmed (2003); George and Jardine (1994); Summons et al. (1999, 1988b, 1994b); Taylor et al. (1994); (w) Summons et al. (1999); (x) Crick et al. (1988); Greenwood and Summons (2003); Logan et al. (2001); McCaffrey et al. (1994b); Summons et al. (1999, 1988b), Jackson et al. (1986); (y) Peng et al. (1998); (z) Brocks et al. (2003a,b,c,d, 1999); Arouri et al. (2000a,b).
hydrocarbon fossil in oils and bitumens that is clearly derived from an animal: 24-isopropylcholestane diagnostic for sponges (Section 8.03.5.11), although cholestane is likely to have significant contributions from the cholesterol of animals.

8.03.7 BIOMARKERS AS ENVIRONMENTAL INDICATORS

Organic matter can provide important clues for paleoenvironmental assessments (Table 5) (de Leeuw et al., 1995). Because some biomarkers point to specific taxa, they can also act as indicators of specific habitats. Paleoenvironmental conditions that are often readily inferred from the presence and distribution patterns of biomarkers are marine (e.g., \((66d)\)), terrestrial (e.g., \((61)\)), and deltaic environments where plant and algal hydrocarbons are mixed or show stratigraphy-related fluctuations in abundance.

8.03.7.1 Marine versus Lacustrine Conditions

As discussed above (Section 8.03.5.11), 24-n-propylcholestane \((66d)\) is considered an unambiguous indicator of marine depositional environments. Additionally, marine conditions can often be inferred from high abundances and the compositions of organo-sulfur compounds as the prevalence of sulfide in euxinic marine environments strongly affects the diagenetic pathways and preservation of many classes of lipids (e.g., Kohnen et al., 1992, 1993, 1991a; Schouten et al., 2001a; Wakeham et al., 1995). Organosulfur compounds are usually less abundant in sediments that were deposited in freshwater. However, freshwater environments are often indicated by the presence of biomarkers of typical freshwater organisms such as \textit{Botryococcus braunii}. Lacustrine conditions are often indicated by preponderances of algal steroids (e.g., Chen and Summons, 2001), biomarkers for aerobic methanotrophs (Collister et al., 1992) and, very often, by the presence of certain \(C_{30}\) tetracyclic polyprenoid hydrocarbons (Holba et al., 2003). The Cenozoic lacustrine basins of China provide numerous examples of biomarker patterns that are characteristic of nonmarine (freshwater and saline) depositional systems (e.g., Chen and Summons, 2001; Chen et al., 1989; Li et al., 2003; Philp et al., 1992; Ping’an et al., 1992).

### Table 5 Biomarkers as environmental indicators.

<table>
<thead>
<tr>
<th>Depositional environment</th>
<th>Typical biomarker patterns</th>
<th>Reference example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marine</td>
<td>24-n-Propylcholestane ((66d))</td>
<td>Moldowan et al. (1985)</td>
</tr>
<tr>
<td></td>
<td>Botryococcane ((20)) and other biomarkers of \textit{Botryococcus} (fresh to brackish water).</td>
<td>Metzger and Largeau (1999)</td>
</tr>
<tr>
<td></td>
<td>Elevated concentrations of (C_{30}) tetracyclic polyprenoids (fresh to brackish water).</td>
<td>Holba et al. (2003)</td>
</tr>
<tr>
<td>Hypersaline</td>
<td>(C_{21}) to (C_{25}) regular isoprenoids (6) enriched in (^{13}C) relative to biomarkers of phytoplanktonic origin.</td>
<td>Grice et al. (1998b)</td>
</tr>
<tr>
<td></td>
<td>High gammacerane (60).</td>
<td>Sinninghe Damsté et al. (1995)</td>
</tr>
<tr>
<td>Terrestrial organic matter input</td>
<td>Diverse biomarkers of higher plants</td>
<td>Section 8.03.6.3</td>
</tr>
<tr>
<td>Strongly anoxic conditions (water column anoxia?)</td>
<td>28,30-Dinorhopane;</td>
<td>Peters and Moldowan (1993)</td>
</tr>
<tr>
<td></td>
<td>25,28,30-trisnorhopane ((59)).</td>
<td>Sinninghe Damsté et al. (1995)</td>
</tr>
<tr>
<td></td>
<td>Gammacerane (60).</td>
<td>Grice et al. (1996a), Hartgers et al. (1993), Koopmans et al. (1996a), Summons and Powell (1987)</td>
</tr>
<tr>
<td>Photic zone euxinia (^{b})</td>
<td>Isorenieratane ((32)); 2,3,6- ((38)) and 2,3,4- ((39)) trimethylarylisoprenoids; chlorobactane ((28)); Me i-Bu maleimide ((49d))</td>
<td>van Kaam-Peters et al. (1998)</td>
</tr>
<tr>
<td>Carbonates and evaporites</td>
<td>Low diasterane (67)/sterane ((66)) ratios(^a)</td>
<td>Summons et al. (1999)</td>
</tr>
<tr>
<td></td>
<td>High 2α-methylhopane ((57)) concentrations(^a)</td>
<td>Subroto et al. (1991)</td>
</tr>
<tr>
<td></td>
<td>High 30-norhopanes(^a)</td>
<td></td>
</tr>
</tbody>
</table>

\(^{a}\) Typical for, but not necessarily restricted to, this depositional environment. \(^{b}\) This might include environments with an anoxic and sulfidic water column that persists into the photic zone, or microbial mats in very shallow water settings that become anoxic within millimeters below the sediment water interface.
8.03.7.2 Hypersaline Conditions

Halophiles are found in all three domains of life with a wide diversity of metabolisms such as aerobic heterotrophy and fermentation, sulfate reduction, denitrification, methanogenesis, and anoxygenic and oxygenc phototrophy. Many sediments deposited under hypersaline conditions contain abundant biomarkers probably derived from archaeal Halobacteria (Section 8.03.6.2.3). For instance, Miocene/Pliocene halite deposits from the Dead Sea Basin in Israel contain pristane, phytane and C_{21} to C_{25} regular isoprenoids as the dominant lipids of the hydrocarbon fraction (Grice et al., 1999). These isoprenoids are enriched in $^{13}$C by up to 7‰ relative to biomarkers of presumed phytoplanktonic origin, consistent with Halobacteria as the dominant source. Hypersaline lakes and ponds often develop anoxic conditions if saline deep water is covered with water of lower density. Sedimentary rocks that were deposited under these conditions often contain high relative concentrations of gammacerane (60), a biomarker generally associated with water column stratification (Section 8.03.5.10) (Sinninghe Damsté et al., 1995). However, as water column stratification occurs under other conditions as well, gammacerane is also often abundant in freshwater sediments (e.g., Grice et al., 1998c).

8.03.7.3 Anoxic and Euxinic Conditions

Biomarker analysis is one of the best paleoenvironmental tools to identify anoxic and euxinic conditions in the water column. Biomarkers of phototrophic sulfur bacteria such as isorenieratane (32) (Section 8.03.5.6.1) and 3-isobutyl-4-methylmaleimide (49d) (Section 8.03.5.7) unambiguously indicate euxinic conditions within the photic zone of the water–sediment system. Other biomarkers that are often associated with sediments deposited beneath anoxic waters are 28,30-dinorhopane, 25,28,30-trisnorhopane (59) and gammacerane (60) (Section 8.03.5.10).

Nitrogen cycling in anoxic water columns and sediments is another prominent biogeochemical process that can have direct molecular and isotopic indicators, Sinninghe Damsté et al. (2002c) discovered that organisms (Planctomycetales) capable of oxidizing ammonia with nitrate (anammox) biosynthesize unprecedented glycerol ester and ether lipids with hydrocarbon chains comprising concatenated cyclobutane rings, or ladderanes (e.g., (71)).

A combination of water column nutrient profiles, fluorescently labeled RNA probes, the vertical distribution of specific “ladderane” membrane lipids, and experiments with $^{15}$N labeled ammonium and nitrate demonstrate that anammox organisms are presently active in anaerobic oxidation of ammonia below the oxic zone of the Black Sea (Kuypers et al., 2003). These observations suggest that the anammox pathway of nitrogen cycling may be widespread in suboxic environments in the modern marine realm. 16S rRNA sequences indicate that the Planctomycetales are a distinct and ancient lineage within the bacterial domain (Brochier and Philippe, 2002). Further, the anammox reaction may have been even more significant in times past when ocean-water columns were largely anaerobic (Anbar and Knoll, 2002; Canfield, 1998). If ladderane-like lipids evolve diagenetically to recognizable chemical fossils, we expect this would be prominently recorded in Proterozoic sediments.

8.03.7.4 Carbonates versus Clay-rich Sediments

Acid-catalyzed rearrangement reactions are promoted during diagenesis of organic matter adsorbed to clay particles (e.g., Rubinstein et al., 1975). Accordingly, rearranged sterenes (diasteranes) are relatively more abundant in clastic sediments than in carbonates (van Kaam-Peters et al., 1998). Hopanoids appear to be similarly affected so that diahopanes and neohopanes are relatively more prominent in bitumens and oils derived from shales as opposed to carbonates (Peters and Moldowan, 1993). However, increasing thermal maturity is also a key factor in the conversion of biomarkers to their rearranged forms.

Carbonate-dominated sediments tend to be deposited in low-latitude environments and, therefore, biomarkers for organisms that preferentially colonize warm waters tend to be important signatures in these sediments. Cyanobacterial 2α-methylhopanes (57) (Summons et al., 1999) and 30-norhopanes (Subroto et al., 1991) are generally elevated in bitumens from carbonates and marls.

8.03.7.5 Paleotemperature and Paleolatitude Biomarkers

Paleotemperature might be reconstructed in ancient sediments through biomarker signals if cold-adapted and warm-adapted organisms produced distinctive lipids. The best-known example of such a signal is the long-chain ketones produced by haptophytes that carry patterns of unsaturation determined by sea-surface temperature (Brassell et al., 1986).
Water temperature is one factor that can influence the concentration of dissolved CO₂ and, thereby, the isotopic fractionation encoded during photosynthetic carbon assimilation. This has been suggested as a means through which paleolatitude could be reconstructed from the carbon-isotopic composition of petroleum hydrocarbons sourced from rocks laid down during time intervals when significant pole to equator temperature gradients prevailed (Andrusevich et al., 2001).

8.03.8 AGE DIAGNOSTIC BIOMARKERS

It is now well known that the hydrocarbon composition of petroleum has evolved over geological time reflecting a corresponding evolution in sedimentary organic matter and, hence, biology. Unusual oils with high abundances of branched alkanes appeared to be exclusively associated with “Infracambrian” source rocks of Siberia and Oman (e.g., Fowler and Douglas, 1987; Grantham et al., 1988) while oils with strong odd carbon number predominances at \( n-C_{15} \), \( n-C_{17} \), and \( n-C_{19} \), and extremely low abundances of acyclic isoprenoids are often found in Ordovician strata (e.g., Hoffmann et al., 1987; Reed et al., 1986). These are features of bulk hydrocarbon composition that denote an overwhelming input from a single organism such as Gloeocapsomorpha prisca, in the case of Ordovician oils and kukersite oil shales (Section 8.03.5.2). Further major changes in petroleum composition accompanied vascular plant radiations in the Late Paleozoic and again during the Cenozoic, with organic matter contributions from leaf waxes, resins, and other terpene-based biopolymers. The isotopic composition of marine organic carbon has changed over geological time (Hayes et al., 1999) and there is a concomitant secular variation in the isotopic compositions of petroleum from marine source rocks (Andrusevich et al., 1998).

More subtle changes occur in the distribution of hydrocarbons that reflect the radiation of specific taxa and their distinctive biochemicals. Algal steroids show particularly strong age-related trends and can be used in a forensic sense to constrain the age of organic sedimentary matter, including petroleum. Prime examples are the heightened occurrences of 24-isopropylcholestanes (66e) in Proterozoic and Early Paleozoic sediments (McCaflrey et al., 1994a), dinosteroids (e.g., 70i) in the Mesozoic and Cenozoic (e.g., Moldowan and Taltyzina, 1998; Summons et al., 1992) and 24-norcholestanes (Holba et al., 1998a) in the Cenozoic. Triterpenoids from angiosperms are another class of compounds that show very strong age-related patterns of occurrence (e.g., Moldowan et al., 1994).

8.03.9 BIOMARKERS IN PRECAMBRIAN ROCKS

8.03.9.1 Biomarkers in the Proterozoic (0.54–2.5 Ga)

There are numerous sedimentary sequences in the Proterozoic that contain abundant and well-preserved organic matter. Characterization of this organic matter, and especially the establishment of its age, has provided a major challenge for geochemists, and much of this progress has been reviewed by Hayes et al. (1983), Summons and Walter (1990), and Brocks et al. (2003a). Organic matter in the form of distinctive, morphologically diverse, organic-walled microfossils abounds in otherwise organic-lean shales and carbonates (e.g., Butterfield et al., 1988) and, as with other paleoflora, carry information about biota and environments. Rocks with the high contents of organic matter tend to have amorphous kerogen which is difficult to study by optical methods but may be amenable to pyrolysis and chemical degradation studies for paleoenvironmental and paleobiological reconstruction (e.g., Aroui et al., 2000a, 1999). For this, identification and selection of sediments with very mild thermal histories and freedom from the damaging effects of ionizing radiation (e.g., Dahl et al., 1988) is essential for making accurate assessments of these issues.

Studies of bitumens have been far more extensive than studies of kerogen due to the relative ease with which extractable and volatile hydrocarbons can be analyzed (e.g., Summons and Walter, 1990). The carbon skeletons found prominently include algal steroids, bacterial hopanoids, and archaeal polyisoprenoids. Distinctive biomarker distribution patterns are common in the Neoproterozoic. Prime examples include organic-rich shales and marls within the Chuar Group, Grand Canyon, USA, (Summons et al., 1988a), Rodda Beds, Bitter Springs and Pertatataka Formations of Central Australia (Hayes et al., 1992), and the Terminal Proterozoic of Oman and the Siberian Platform (Figure 5) (e.g., Fowler and Douglas, 1987; Grantham, 1986; Klomp, 1986; Summons and Powell, 1992). It is in these sediments that one finds unprecedented predominances of a single-sterane homolog, either C_{27} (cholestanes (66a) or C_{29} (stigmastanes (66c)), signals that might be related to the radiation and massive occurrence of specific algal clades. Similarly, 24-isopropylcholestone (66e) shows a unique predominance in the Neoproterozoic to Ordovician, and this is hypothesized to be a consequence of the radiation of sponges and their archaeocyathid or stromatoporoid relatives (McCaflrey et al., 1994b).

Heightened relative abundances of monomethyl, dimethyl, and other branched acyclic alkanes is another distinctive feature of Proterozoic
bitumens (e.g., Höld et al., 1999; Klomp, 1986; Logan et al., 1999). This is most often seen in clastic lithologies and only rarely in carbonates. Some patterns of highly branched alkanes appear to be specific to benthic microbial mats and, on the basis of carbon and sulfur isotope anomalies, are hypothesized to be associated with sulfide-oxidizing microbial communities (Kenig et al., 2002; Logan et al., 1999).

Evidence from the age-distribution of mineral deposits, sedimentary patterns of redox-sensitive trace metals combined with sulfur- and carbon-isotope systematics point to a profound evolution of the ocean redox structure during the Proterozoic eon (e.g., Anbar and Knoll, 2002; Canfield, 1998; Des Marais et al., 1992). In particular, it is hypothesized that the oceans were sulfide-rich and sulfate-poor after the cessation of deposition of banded iron formations (BIF) in the Paleoproterozoic roughly 1.8 Ga and prior to the existence of ventilated oceans, possibly as early as the end of the Mesoproterozoic (1.0 Ga) (Canfield, 1998; Shen et al., 2003) or at the end of the Neoproterozoic (Logan et al., 1995). Analyses of organic matter provide some evidence for unusual diagenetic pathways and support the hypothesis that the biogeochemical carbon cycle in the Proterozoic was fundamentally different from that of the Phanerozoic. Studies of kerogens indicate that Proterozoic sedimentary organic matter, despite having high elemental hydrogen to carbon ratios, tends to be unusually aromatic in nature and yields relatively low amounts of aliphatic hydrocarbons during burial maturation (Summons et al., 1994b) and also during catalytic hydropyrolysis (Brocks et al., 2003c). Carbon-isotopic compositions of kerogens and co-occurring individual hydrocarbons in sediments throughout the Proterozoic show a different order to those observed in the Cambrian (Logan et al., 1997). This was hypothesized to be a hallmark of a major re-organization of the biogeochemical carbon and sulfur cycles at the Proterozoic–Phanerozoic transition (Logan et al., 1995). Rothman et al. (2003) analyzed fluctuations in the isotopic records of sedimentary organic and inorganic carbon through the Neoproterozoic and found evidence for non-steady-state behavior of the carbon cycle at this time. Thus, there are numerous clues pointing to an evolution in carbon cycle and in the type of organic matter that was being buried. While the actual compounds that are found in Proterozoic sediments tend to be the same ones that are encountered in younger rocks, their relative abundances, distribution patterns, and isotopic characteristics can be quite different. It is in this regard that studies of kerogen composition, biomarkers and compound-specific isotope data may prove to be most useful for evaluating environmental and ecological evolution during the Proterozoic and especially across the Proterozoic–Phanerozoic transition.

8.03.9.2 Biomarkers Extracted from Archean Rocks (>2.5 Ga)

An example that illustrates the difficulties that might be associated with establishing the age of solvent extractable organic matter are biomarkers detected in 2.7–2.5 Ga rocks from the Hamersley Basin, Western Australia (Brocks et al., 1999). The host rocks from the Hamersley and Fortescue Group, although exceptionally well preserved by Archean standards, have suffered low-grade metamorphism at temperatures between 175 °C and 300 °C (Brocks et al., 2003a). Yet, solvent extraction of kerogen-rich shales unexpectedly yielded 1 ppm–1,000 ppm n-alkanes, methylalkanes, acyclic isoprenoids, adamantanes, tri- to penta-cyclic terpanes, sterenes, and polycyclic hydrocarbons (PAH). One sample, a black shale from a hydrothermally altered iron mine in the Hamersley Group, exclusively contained adamantanes, parent PAH and minor concentrations of methylated PAH, patterns indicating extremely high thermal maturity and possibly hydrothermal alteration. Moreover, PAH with the same overmature pattern were also released by pyrolytic degradation of isolated kerogens from other iron deposits in the Hamersley Basin (Brocks et al., 2003c). The unusual composition, extreme thermal maturity and covalent bonding to kerogen rank these adamantanes and PAH as the by far oldest known “certainly syngenetic” bitumens in terrestrial rocks.

However, most samples from the Hamersley Basin additionally contain aliphatic hydrocarbons and polycyclic biomarkers in mixture with the certainly syngenetic adamantanes and PAH. The origin and age of these thermally less-stable components is less well constrained. Arguments against their syngeneity are a pronounced carbon-isotopic difference between bitumen and kerogen (Brocks et al., 2003a), the absence of saturated hydrocarbons in kerogen pyrolysates (Brocks et al., 2003c) and, most significantly, a strong inhomogeneous distribution of bitumen in individual drill core samples that is potentially consistent with surficial staining and migration of hydrocarbons into the rock (Brocks et al., 2003a). However, the samples come from eight independent drill cores, drilled by several different companies, stored several hundred kilometers apart, collected by different workers over several years and analyzed in two laboratories with consistent results. All samples contain bitumen with a typical earlier Precambrian composition: absence of plant biomarkers, predominance of C_{27}-steranes, high C_{31}-2α-methyllopane indices
(8–20%), and phytane isotopically depleted relative to n-C₁₈. Moreover, the thermal maturity of the biomarkers is within the wet-gas zone of petroleum generation, younger petroleum source rocks are absent within the basin and were never deposited over the top, and the shales were collected from diamond drill core over an area of several hundred kilometers (Brocks et al., 2003a). Therefore, the biomarkers are characterized “probably syngenetic.” A less ambiguous classification might become available when fresh material is collected in the Hamersley Basin under controlled conditions as part of the Deep Time Drilling Project (Dalton, 2001).

However, if the biomarkers are in fact syngenetic, then they provide new insights into Archean biodiversity and ecology (Brocks et al., 2003b; Brocks et al., 1999). The presence of hopanes confirms the antiquity of the domain bacteria, and biomarkers of the 3β-methylhopane series suggest that microaerophilic Proteobacteria, probably methanotrophs, were active in Late Archean marine environments. High relative abundances of C₃₀ to C₃₆ 2α-methylhopanes indicate that cyanobacteria were important primary producers in the Late Archean. Therefore, oxygenic photosynthesis probably evolved before 2.7 Ga. High relative concentrations of cyanobacterial biomarkers were also detected in thin layers of Late Archean shales interbedded with oxide-facies banded iron formations (BIF) suggesting that, although some Archean BIF might have been formed by anoxygenic phototrophic bacteria or nonbiological photochemical processes, those in the Hamersley Group formed as a direct consequence of biogenic oxygen production. As chlorophyll biosynthesis in cyanobacteria probably succeeded the evolution of bacteriochlorophylls in anoxygenic phototrophic bacteria (Xiong et al., 2000), the 2α-methylhopanes also give indirect evidence that all lineages of anoxygenic phototrophs—heliobacteria, purple bacteria, green sulfur bacteria, and green nonsulfur bacteria—evolved before 2.7 Ga (Des Marais, 2000). Steranes, including 4-methylsteranes (69), desmethylsteranes alkylated at C-24 ((66b)–(66d)), and aromatic steroids (68), occur in relative abundances similar to those from other Precambrian sources, providing evidence that ancestral eukaryotes existed ~900 Ma before the earliest microfossil evidence indicates that the lineage arose (Hofmann and Chen, 1981; Zhang, 1986). Sterol biosynthesis in extant eukaryotes requires dissolved molecular oxygen in concentrations equivalent to ~1% of the present atmospheric level (Jahnke and Klein, 1979; Jahnke and Klein, 1983). Therefore, it is likely that oxygen concentrations in Archean surface waters were high enough to support aerobic respiration to some extent.

8.03.10 OUTLOOK

In recent years, the discovery of new biomarkers and their sources has been greatly aided by the combination of molecular and compound-specific isotopic analysis methods (e.g., Hinrichs et al., 2000), the advent of genomic tools to screen natural samples for the identities of dominant taxa (e.g., Boetius et al., 2000; Hinrichs et al., 1999) and the advent of culture-independent methods for studying important biogeochemical processes (e.g., Orphan et al., 2001). The extensive screening of cultured extant organisms in the past has shown very general connections between taxa and their diagnostic markers. While this will continue, access to important organisms that are difficult or impossible to grow in the laboratory can be accomplished by studies of their genomes. Moreover, these genomes also encode an evolutionary history so that it may eventually be possible to reconstruct genetic information (paleogenomics) about extinct ancestors and their biochemical capacities (Benner, 2001). Accurate timing of evolutionary events can only be accomplished by studies of the rock record, and exploration of the fossil biomarkers will continue to be an important activity in the search for life’s early history on Earth.

ACKNOWLEDGMENTS

The manuscript was greatly improved by helpful comments on various sections of this chapter by Alison Cohen, Jennifer Eigenbrode, David Fike, Kliti Grice, John Hayes, Fabien Kenig, Ann Pearson, and John Volkman. We thank Ann Pearson for interesting discussions about the occurrence of sterols in bacteria. Roger Summons is supported by grants from the NASA Exobiology program and the NASA Astrobiology Institute and Jochen Brocks gratefully acknowledges the Harvard Society of Fellows for providing support during the preparation of this work.

REFERENCES


References


Suzuki K.-I., Saito K., Kawaguchi A., Okuda S., and Komagata

Summons R. E., Jahnke L. L., Hope J. M., and Logan G. A.

Summons R. E., Jahnke L. L., and Roksandic Z. (1994a)

Summons R. E., Taylor D., and Boreham C. (1994b)

Taylor D., Kontorovich A. E., Larichev A. I., and Glikson M.

Talyzina N. M., Moldovan J. M., Johannisson A., and

Summons R. E., Brassell S. C., Eglinton G., Evans E.,


Summons R. E. and Walter M. R. (1990) Molecular fossils and

in lacustrine sediments and oils. Geochim. Cosmo-

B22D-0184.

Tornabene T. G., Langworthy T. A., Holzer G., and Oro J.

Ten Hagen J. T., van Aarssen B. G. K., Cox H. C., Hoogendoorn P., and de

van Aarssen B. G. K., Cox H. C., Hoogendoorn P., and de

and fatty acids of the marine diatom


ten Haven H. L., Rohmer M. Rullkötter J., and Bisseper R. (1989) Tetracyhanomol, the most likely precursor of gamma-

determinations using physical and chemical methods. APEA J. 34, 279–296.


Volkmann J. K. (1986) A review of steroid markers for marine and


Watts C. D., Maxwell J. R., and Kjøsøen H. (1977) The potential recording or otherwise, without prior written permission of the Publisher.

References

115


