Geobiology (2004), **2**, 199–203

Biosynthetic pathways, gene replacement and the antiquity of life

JASON RAYMOND[1] AND ROBERT E. BLANKENSHIP[2]

[1]*Biology and Biotechnology Research Program, Lawrence Livermore National Laboratory Livermore, CA 94551, USA*
[2]*Department of Chemistry and Biochemistry, Arizona State University Tempe, AZ 85287–1604, USA*

ABSTRACT

The appearance of oxygen in the Earth's atmosphere, a by-product of oxygenic photosynthesis invented by primitive cyanobacteria, stands as one of the major events in the history of life on Earth. While independent lines of geological data suggest that oxygen first began to accumulate in the atmosphere *c*. 2.2 billion years ago, a growing body of biomarker data purports to push this date back fully 500 million years, based on the presumption that an oxygen-dependent biochemistry was functional at this time. Here, we present a cautionary tale in the extension of modern biochemistry into Archean biota, identifying a suite of examples of evolutionary convergence where an enzyme catalysing a highly specific, O_2-requiring reaction has an oxygen-independent counterpart, able to carry out the same reaction under anoxic conditions. The anaerobic enzyme has almost certainly been replaced in many reactions by the more efficient and irreversible aerobic version that uses O_2. We suggest that the unambiguous interpretation of Archean biomarkers demands a rigorous understanding of modern biochemistry and its extensibility into ancient organisms.

Received 16 December 2004; accepted 03 January 2005

Corresponding author: R. Blankenship. Tel.: (480)965 4430; fax: (480)965 2747; e-mail: RBlankenship@asu.edu

Our understanding of the evolution and diversity of life during the Archean, 3.8–2.5 billion years ago (Ga), has been profoundly influenced by a handful of geological signatures that remain recognizable despite eons of metamorphism and diagenesis. This evidence, which ranges from microfossils to stromatolites to fractionated isotopes (Rye & Holland, 1998; Farquhar *et al.*, 2002; Schopf *et al.*, 2002), has been extensively debated in recent literature on the grounds of competing or poorly characterized nonbiological processes that can potentially produce indistinguishable signatures (Grotzinger & Rothman, 1996; Brasier *et al.*, 2002). Importantly though, skepticism gives some ground to consensus when considering derivatized hopanoid and sterol biomarkers, a diverse class of cholesterol-like molecules classified as pentacyclic triterpenes that have been recovered from shales during the Archean eon 2700 million years ago (Mya) (Brocks *et al.*, 1999). It has been argued that these biomarkers are only consistent with a biochemistry in which O_2 is produced by oxygenic phototrophs and subsequently utilized in the biosynthesis of sterols (Brocks *et al.*, 2003). Here, we question that conclusion with examples from modern biochemistry.

In all modern organisms, the biosynthesis of pentacyclic triterpenes proceeds through an unsaturated linear molecule known as squalene. It is at this point that the sterol and hopanoid biosynthetic pathways diverge. In organisms that synthesize sterols or any of their extensive derivatives, a single oxygen atom is enzymatically incorporated into squalene from molecular oxygen, forming the epoxide 2,3-oxidosqualene. In both pathways, squalene and 2,3-oxidosqualene are then cyclized into their familiar pentacyclic ring structures in what has been called one of the most complex single-step reactions in biochemistry. Intriguingly, the two enzymes catalysing this step in both pathways are evolutionary homologues, exemplified by the fact that, in some organisms, the squalene-cyclizing enzyme will also cyclize 2,3-oxidosqualene (Rohmer *et al.*, 1979). Thus the major difference at this critical juncture in the two pathways is the incorporation of an oxygen atom from O_2, specific to the sterol biosynthetic pathway (see, e.g. Wendt *et al.* (2000) for a comprehensive review). This requirement for molecular oxygen and their preservation in Archean rocks suggest that derivatized sterol biomarkers may serve as molecular proxies for the oxidation of Earth's atmosphere. In modern organisms, the presence of specific hopanoid and sterol hydrocarbons has also been argued as diagnostic for cyanobacteria and eukaryotes, respectively (Brocks *et al.*, 1999; Summons *et al.*, 1999). Taken at face value, biomarkers derived from these compounds provide perhaps our earliest unobstructed glimpse into the diversity of

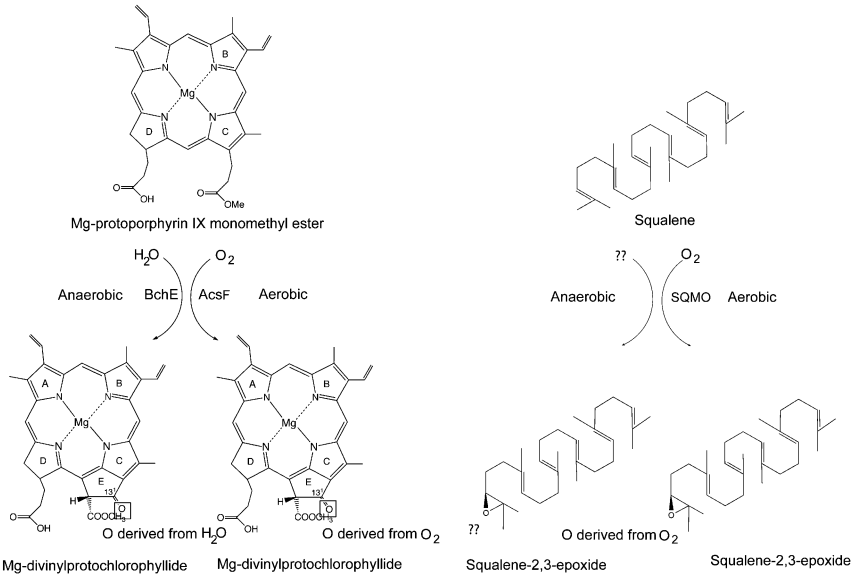

Fig. 1 Gene replacement in the evolution of chlorophyll biosynthesis (left), illustrating the source of the oxygen added to the protoporphyrin precursor in both the aerobic (AcsF, O from O_2) and analogous anaerobic (BchE, O from H_2O) enzymes. At the right is the enzymatic epoxidation of squalene, which is known to incorporate an oxygen atom from O_2 but for which an analogous anaerobic reaction is not known. SQMO: squalene monooxygenase, BchE/AcsF: O_2-independent/O_2-dependent Mg-protoporphyrin IX monomethyl ester cyclases.

Archean life, arguing that 2.7 billion years ago, cyanobacteria – the 'inventors' of oxygenic photosynthesis – had appeared, heralding the impending oxidation of Earth's atmosphere. Importantly, this date sharply contrasts with multiple lines of geological evidence that places the appearance of oxygen at about 2200 Mya (Rye & Holland, 1998), 500 million years later than is suggested by biomarker evidence. Reconciling this discrepancy has generated considerable debate among palaeogeologists and evolutionary biologists alike.

The transition from an anoxic to an oxic world had unfathomable consequences for early life; molecular oxygen and its derivatives went from being poisonous in an obligately anaerobic world, to assuming a central role in the subsequent development of complex, macroscopic life, almost all of which requires O_2 for cellular respiration. Though the aftermath was global in scale, this transition to oxygen tolerance, and ultimately dependence, was first accomplished at the molecular level, which included the invention of a wealth of new enzymes, the extensive modification of many existing ones, and the 'rewiring' of central biochemical pathways to optimize function in the presence of O_2. Intriguingly, many biochemical reactions that were once catalysed in the absence of molecular oxygen, using biosynthetic oxidants such as nicotinamide adenine dinucleotide (NAD^+) or one of various cytochromes, were superseded by a new class of enzymes that utilized O_2 as an electron acceptor. Perhaps selected at first for their ability to reduce O_2 to innocuous H_2O, many of these new enzymes became indispensable to early aerobic organisms, in many cases completely supplanting the function and then displacing the presence of their anaerobic analogues. We argue that, based on remarkable corollaries found throughout enzymology, the oxygenation of squalene during the Archean could plausibly have been carried out by an anaerobic enzyme that either remains to be characterized or that may have been totally lost as organisms, especially early eukaryotes, switched from anaerobic to almost exclusively aerobic lifestyles.

One notable example of an O_2-induced enzyme replacement evidently occurred in cyanobacteria, algae, and plants, the organisms responsible for the oxygen in the Earth's atmosphere. AcsF, genetic shorthand for the so-called aerobic cyclase enzyme, catalyses the O_2-dependent synthesis of the chlorophyll and bacteriochlorophyll precursor Mg-divinyl chlorophyllide, an absolutely required intermediate step in the biosynthesis of photosynthetic pigments (Fig. 1). By requiring molecular oxygen to synthesize chlorophyll, which in turn is required for the organism to produce O_2, AcsF raises the apparent paradox of an enzyme that requires its ultimate product as a substrate. This chicken-egg dilemma is especially problematic because evolutionary analyses suggest that the first photosynthetic organisms on the Earth were incapable of producing oxygen (Xiong *et al.*, 2000). How could these primitive phototrophs synthesize their required pigments in an anaerobic world? In a testament to the innovative ability of microorganisms, it has recently been discovered that anaerobic phototrophs catalyse this cyclization reaction via a completely different and presumably more primitive, vitamin B12-dependent enzyme (known as BchE) that incorporates an oxygen atom from a ubiquitous source – water – rather than from O_2 (Fig. 1) (Pinta *et al.*, 2002). In fact, the presence of the gene for one of these enzymes in a genome is sufficient to predict a photosynthetic organism's preferred lifestyle: O_2-requiring AcsF is found in aerobic organisms, whereas B12-utilizing BchE is found in anaerobes. Moreover, a number of facultative phototrophs capable of growth under either aerobic or anaerobic conditions have both types of enzymes, enabling these organisms to adapt this step of pigment synthesis to the situation at hand (Ouchane *et al.*, 2004).

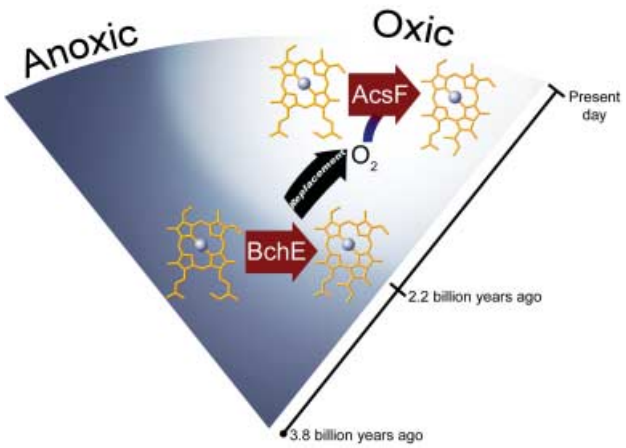

Fig. 2 On the early anoxic Earth (blue), oxygen-independent BchE was required for photosynthetic pigment synthesis and is still used in modern anaerobic phototrophs. Phototrophs living in oxic environments (white) have replaced BchE with an enzyme that catalyses the same steps but utilizes molecular oxygen. This represents just one of many remarkable examples of the independent evolution of enzymes that postdate the oxidation of Earth's atmosphere.

This BchE/AcsF replacement, conceptualized against an evolving biosphere in Fig. 2, is not the only anaerobic-to-aerobic switch that has occurred in the biosynthesis of diverse tetrapyrroles (of which chlorophyll, vitamin B12, and heme are all examples). Further examples include two successive oxidative steps *en route* to synthesizing heme with O_2-dependent and independent enzymes (Ouchane *et al.*, 2004). There are striking cofactor, structural, and possible mechanistic similarities between the BchE enzyme discussed above, and the anaerobic (class III) ribonucleotide reductase essential for synthesizing deoxyribose that is incorporated into DNA. Each is an anaerobic enzyme whose role has been displaced in aerobic organisms by an O_2-requiring or tolerant analogue. Vitamin B_{12} biosynthesis offers yet another exceptional example where a complete rearrangement of biosynthetic steps has taken place in the evolution of the aerobic pathway, centred around enzymatic replacement of cobalt insertion into the B_{12} ring (Raux *et al.*, 2000).

The BioCyc database, a curated, literature-derived repository of nearly 5000 biochemical reactions from 160 species (Krieger *et al.*, 2004), can be used to extend these arguments beyond exemplary anecdotes. We determined that O_2 appears as a reactant in 473 BioCyc pathways, in a diverse range of roles ranging from biosynthesis to electron transfer to detoxification. As a first-order estimate of analogous O_2-independent reactions, we parsed the database for reactions with an identical substrate and product to one of the O_2-dependent reactions, but that did not utilize O_2. The resultant list contained 85 different reaction sets comprised of at least 19 biochemically distinct reactions (Table 1), with redundancy because of the plasticity displayed by some enzymes, such as cytochromes, which can participate in donating electrons in a multitude of different reactions. Many of these displacement examples can be grouped into oxidase/dehydrogenase pairs, which are able

Table 1 Anaerobic/aerobic gene replacements

O_2-independent	O_2-dependent	Conserved reaction
glycerol-3-phosphate dehydrogenase (1.1.1.8)	glycerol-3-phosphate oxidase (.1.3.21)	SN-glycerol-3-phosphate < = > glycerone phosphate
N-acetylhexosamine-1-dehydrogenase (1.1.1.240)	N-acylhexosamine oxidase (1.1.3.29)	N-acetyl-D-glucosamine < = > N-acetyl-D-glucosaminate
Choline dehydrogenase (1.1.99.1)	Choline oxidase (1.1.3.17)	choline < = > betaine aldehyde
L-sorbose dehydrogenase (1.1.99.12)	L-sorbose oxidase (1.1.3.11)	L-sorbose < = > 5-dehydro-D-fructose
glycolate dehydrogenase/reductase (1.1.99.14/1.1.1.26)	glycolate oxidase (1.1.3.15)	glycolate < = > glyoxylate
glucose dehydrogenase (1.1.99.17)	glucose oxidase (1.1.3.4)	D-glucose < = > D-glucono-1,5-lactone
cellobiose dehydrogenase (1.1.99.18)	cellobiose oxidase (1.1.3.25)	cellobiose < = > cellobiose-1,5-lactone
dihydrouracil dehydrogenase (1.3.1.1)	dihydrouracil oxidase (1.3.3.7)	5,6-dihydrouracil < = > uracil
dihydroorotate dehydrogenase (1.3.99.11)	dihydroorotate oxidase (1.3.3.1)	(S)-dihydroorotoate < = > orotate
L-glutamate dehydrogenase (1.4.1.3)	L-glutamate oxidase (1.4.3.11)	L-glutamate < = > 2-oxoglutarate + NH3
L-amino acid dehydrogenase (1.4.3.5)	L-amino acid oxidase (1.4.3.2)	L-amino acid < = > 2-oxo acid + NH3
sarcosine dehydrogenase (1.5.99.1)	sarcosine oxidase (1.5.3.1)	sarcosine < = > glycine + formaldehyde
dimethylglycine dehydrogenase (1.5.99.2)	dimethylglycine oxidase (1.5.3.10)	N,N-dimethylglycine < = > sarcosine +formaldehyde
glutathione peroxidase/dehydrogenase (1.11.1.9/1.8.5.1)	glutathione oxidase (1.8.3.3)	2 glutathione < = > oxidized glutathione
O_2-independent	O_2-dependent	coproporphyrinogen III < = >
coproporphyrinogen oxidase (1.–.–.–)	coproporphyrinogen oxidase (1.3.3.3)	protoporphyrinogen IX
L-aspartate dehydrogenase (Yang *et al*., 2003)	L-aspartate oxidase (1.4.3.16)	L-aspartate < = > oxaloacetate + NH3
O_2-independent oxidative cyclase(BchE)	*O_2-dependent oxidative cyclase(AcsF)*	*Mg-protoporphyrin < = > Mg-protochlorophyllide*
class II/III ribonucleotide reductase	*O_2-dependent (class I) ribonucleotide reductase*	*NTP < = > dNTP*
anaerobic cobalt chelatase	*aerobic cobalt chelatase*	*cobalt insertion into corrin precursor (complete pathway rearrangement)*

Single-step reactions catalysed by both oxygen-dependent and -independent enzymes were identified using the BioCyc database. Reactions not included in BioCyc but discussed in the text are italicized.

to catalyse a broad range of identical substrate oxidations coupled either with reduction of O_2 or $NAD(P)^+$. While diverse, this family of NAD- or nucleotide cofactor-utilizing enzymes catalyses very similar chemical reactions involving electron and proton shuffling (oxidation-reduction) between compounds. Mechanistically, this is a very different class of reactions than would be envisaged for oxygenating squalene anaerobically, which requires incorporating an oxygen atom most likely from water as opposed to molecular oxygen. Importantly, this H_2O- vs. O_2-utilizing oxygenation is exactly the chemical substitution that Ouchane *et al.* (2004) have detailed for the BchE/AcsF enzyme replacement that took place during the evolution of photosynthetic pigment biosynthesis, identifying the oxygen-independent BchE enzyme as a modern analogue to such chemistry on the anoxic Earth. As illustrated in Fig. 1, interpreting squalene epoxide-derived biomarkers as corollaries for the appearance of oxygen in Earth's early atmosphere requires that the oxygen atom added to squalene is and always has been derived from O_2, and not from an alternative source such as H_2O. A mechanistic and evolutionary understanding of O_2-dependent enzymes and their O_2-independent counterparts should give insight into how and why unrelated enzymes sometimes evolve to catalyse identical reactions, as well as into the plausibility of anaerobic squalene epoxidation.

Lastly, it is important to note that this cursory analysis parsed only single-step enzyme displacements. Pathway or network-level rearrangements, such as those illustrated in the aforementioned cobalt insertion into the B_{12} ring, are thereby not recovered by this first-order glimpse at the diverse biochemistry impelled by the transition to an aerobic atmosphere. It therefore seems likely that many examples of the 'new' biochemistry impelled by the evolution of oxygenic photosynthesis remain to be discovered.

Though important new studies have recently begun to unravel the diversity of sterol biosynthesis among prokaryotes and primitive eukaryotes (Brocks *et al.*, 2003; Pearson *et al.*, 2003), the salient implications of biomarker synthesis on the Archean Earth demand a thorough understanding before making a compelling case for the appearance of oxygenic photosynthesis 2.7 billion years ago, a paramount event in the history of life. This is especially true in light of the incredible enzymatic plasticity Nature has demonstrated in the reactions discussed here.

Numerous metabolic pathways have clearly been 'updated' over time by replacement of anaerobic enzymes with more efficient oxygen-dependent versions, so that the modern pathway, while carrying out the same overall chemistry, utilizes very different enzymes than the ancient pathway. Oxygen-dependent enzymes are preferred for reasons of thermodynamic efficiency and irreversibility. Over time, they will have gradually replaced the anaerobic versions, so that the present list of anaerobic enzymes may be only a small fraction of the number that once existed on an anaerobic Earth. This progressive gene replacement process significantly complicates the interpretation of ancient biomarker data to determine the availability of oxygen and date the origin of cyanobacteria and eukaryotes.

ACKNOWLEDGEMENTS

We would like to thank Joe Kirschvink and Jim Kasting for their discussions, insights, and encouragement in developing this manuscript. J.R. acknowledges funding through an E.O. Lawrence Fellowship at Lawrence Livermore National Laboratory. This work has been supported by the NASA Exobiology Program.

REFERENCES

Brasier MD, Green OR, Jephcoat AP, Kleppe AK, Van Kranendonk MJ, Lindsay JF, Steele A, Grassineau NV (2002) Questioning the evidence for Earth's oldest fossils. *Nature* **416**, 76–81.

Brocks JJ, Buick R, Summons RE, Logan GA (2003) A reconstruction of Archean biological diversity based on molecular fossils from the 2.78–2.45 billion-year-old Mount Bruce Supergroup, Hamersley Basin, Western Australia. *Geochimica et Cosmochimica Acta* **67**, 4321–4335.

Brocks JJ, Logan GA, Buick R, Summons RE (1999) Archean molecular fossils and the early rise of eukaryotes. *Science* **285**, 1033–1036.

Farquhar J, Wing BA, McKeegan KD, Harris JW, Cartigny P, Thiemens MH (2002) Mass-independent sulfur of inclusions in diamond and sulfur recycling on early Earth. *Science* **298**, 2369–2372.

Grotzinger JP, Rothman DH (1996) An abiotic model for stromatolite morphogenesis. *Nature* **383**, 423–425.

Krieger CJ, Zhang P, Mueller LA, Wang A, Paley S, Arnaud M, Pick J, Rhee SY, Karp PD (2004) MetaCyc: a multiorganism database of metabolic pathways and enzymes. *Nucleic Acids Research*, 32 Database issue, D438–442.

Ouchane S, Steunou AS, Picaud M, Astier C (2004) Aerobic and anaerobic Mg-protoporphyrin monomethyl ester cyclases in purple bacteria: a strategy adopted to bypass the repressive oxygen control system. *Journal of Biological Chemistry* **279**, 6385–6394.

Pearson A, Budin M, Brocks JJ (2003) Phylogenetic and biochemical evidence for sterol synthesis in the bacterium *Gemmata obscuriglobus*. *Proceedings of the National Academy of Sciences of the USA* **100**, 15352–15357.

Pinta V, Picaud M, Reiss-Husson F, Astier C (2002) *Rubrivivax gelatinosus* acsF (previously orf358) codes for a conserved, putative binuclear-iron-cluster-containing protein involved in aerobic oxidative cyclization of Mg-protoporphyrin IX monomethylester. *Journal of Bacteriology* **184**, 746–753.

Raux E, Schubert HL, Warren MJ (2000) Biosynthesis of cobalamin (vitamin B12): a bacterial conundrum. *Cellular and Molecular Life Sciences* **57**, 1880–1893.

Rohmer M, Bouvier P, Ourisson G (1979) Molecular evolution of biomembranes: structural equivalents and phylogenetic precursors of sterols. *Proceedings of the National Academy of Sciences of the USA* **76**, 847–851.

Rye R, Holland HD (1998) Paleosols and the evolution of atmospheric oxygen: a critical review. *American Journal of Science* **298**, 621–672.

Schopf JW, Kudryavtsev AB, Agresti DG, Wdowiak TJ, Czaja AD (2002) Laser – Raman imagery of Earth's earliest fossils. *Nature* **416**, 73–76.

Summons RE, Jahnke LL, Hope JM, Logan GA (1999) 2-Methylhopanoids as biomarkers for cyanobacterial oxygenic photosynthesis. *Nature* **400**, 554–557.

Wendt KU, Schulz GE, Corey EJ, Liu DR (2000) Enzyme mechanisms for polycyclic triterpene formation. *Angewandte Chemie*. International Ed. In English, **39**, 2812–2833.

Xiong J, Fischer WM, Inoue K, Nakahara M, Bauer CE (2000) Molecular evidence for the early evolution of photosynthesis. *Science* **289**, 1724–1730.

Yang ZR, Savchenko A, Yakunin A, Zhang RG, Edwards A, Arrowsmith C, Tong L (2003) Aspartate dehydrogenase, a novel enzyme identified from structural and functional studies of TM1643. *Journal of Biological Chemistry* **278**, 8804–8808.