Review

Electrons, life and the evolution of Earth’s oxygen cycle

Paul G. Falkowski1,2,* and Linda V. Godfrey1

1Institute of Marine and Coastal Sciences and 2Department of Earth and Planetary Sciences, Rutgers University, New Brunswick, NJ 08901, USA

The biogeochemical cycles of H, C, N, O and S are coupled via biologically catalysed electron transfer (redox) reactions. The metabolic processes responsible for maintaining these cycles evolved over the first ca 2.3 Ga of Earth’s history in prokaryotes and, through a sequence of events, led to the production of oxygen via the photobiologically catalysed oxidation of water. However, geochemical evidence suggests that there was a delay of several hundred million years before oxygen accumulated in Earth’s atmosphere related to changes in the burial efficiency of organic matter and fundamental alterations in the nitrogen cycle. In the latter case, the presence of free molecular oxygen allowed ammonium to be oxidized to nitrate and subsequently denitrified. The interaction between the oxygen and nitrogen cycles in particular led to a negative feedback, in which increased production of oxygen led to decreased fixed inorganic nitrogen in the oceans. This feedback, which is supported by isotopic analyses of fixed nitrogen in sedimentary rocks from the Late Archaean, continues to the present. However, once sufficient oxygen accumulated in Earth's atmosphere to allow nitrification to out-compete denitrification, a new stable electron ‘market’ emerged in which oxygenic photosynthesis and aerobic respiration ultimately spread via endosymbiotic events and massive lateral gene transfer to eukaryotic host cells, allowing the evolution of complex (i.e. animal) life forms. The resulting network of electron transfers led a gas composition of Earth’s atmosphere that is far from thermodynamic equilibrium (i.e. it is an emergent property), yet is relatively stable on geological time scales. The early coevolution of the C, N and O cycles, and the resulting non-equilibrium gaseous by-products can be used as a guide to search for the presence of life on terrestrial planets outside of our Solar System.

Keywords: biogeochemical cycles; nitrogen isotopes; Archaean ocean; Wilson cycle; oxygenic photosynthesis

1. INTRODUCTION

Fundamentally, life on Earth comprises six major elements, namely H, C, N, O, S and P (Williams & Frausto da Silva 1996). These elements are the building blocks of all the major biological macromolecules including proteins, nucleic acids, lipids and carbohydrates. The production of macromolecules requires an input of energy. The hallmark of biological energy transduction is non-equilibrium redox chemistry. Indeed, the evolution of life during the first half of Earth’s history led to a network of electron transfer reactions that have driven fluxes of the first five of the ‘big six’ elements (Falkowski 2001). The vast majority of the fluxes are the result of microbially catalysed, energy transducing, redox reactions (Nealson & Conrad 1999). Because all redox reactions are paired,

the resulting network is a linked system of elemental cycles in which a forward reaction in one biogeochemical guild is complemented by a reverse reaction, usually, but not always, in another biogeochemical guild. The input of energy is almost exclusively derived from the Sun. For example, the reduction of inorganic carbon to organic molecules requires the addition of hydrogen and is endergonic. This reaction is catalysed solely by chemoautotrophs and photoautotrophs, which use either the energy of preformed chemical bonds or light (the Sun) to drive the reduction reaction. The oxidation of organic molecules to inorganic carbon is catalysed by heterotrophs, with the resulting liberation of energy. In some cases, the forward and reverse reactions can be catalysed within the same biogeochemical guilds by altering either the mass balance or conditions of the reaction pathway (Falkowski et al. in press). For example, under reducing conditions, purple photosynthetic bacteria are photoautotrophs, while under oxidizing conditions the same organisms lose their photosynthetic capability, operate their electron transport pathways in reverse, and become facultative heterotrophs. Indeed, this type of ‘dual fuel’ metabolic flexibility is relatively common in microbes,
but became rare with increasing evolutionary complexity (Maynard Smith & Szathmáry 1997).

Since all organisms exchange gases with their environment, the network of fluxes is maintained on a planetary body by movement of substrates and products through the atmosphere and oceans of Earth (Falkowski 2006). The combinatorial metabolic outcomes of the reactions of the five major elements connected by electron transfers ultimately are driven by thermodynamics, but selected by biologically derived catalysts. In this paper, we consider how the linked evolution of the five elements irreversibly altered the geochemistry of Earth and allowed oxygen to accumulate in the atmosphere.

A naive chemist examining the atmosphere on Earth may be completely surprised that the two most abundant gases are N\textsubscript{2} and O\textsubscript{2}. N\textsubscript{2} behaves as a noble gas and it is virtually non-reactive. Geochemists assume that the amount of N\textsubscript{2} in the atmosphere has remained constant since the planet was first formed ca 4.6 Ga ago. Indeed, the turnover time for N\textsubscript{2} in the atmosphere is estimated to be ca 1 Ga (Berner 2006). By contrast, O\textsubscript{2}, the second most abundant gas in Earth’s atmosphere, is highly reactive, and without a continuous source would become rapidly depleted (Keeling et al. 1993). This gas exists far from thermodynamic equilibrium with a virtually infinite source of reductant in Earth’s mantle. Indeed, high concentrations of gaseous diatomic oxygen are unique to this planet in our Solar System and this feature of our planetary atmosphere has not yet been found on any other planet within approximately 20 parsecs of our Solar System (Kasting 1993). The presence of high concentrations of the gas in a planetary atmosphere is presently understood to be a virtually irrefutable indication of life on other terrestrial planets. Why is the gas so abundant on Earth yet so scarce on other planets in our Solar System and beyond?

2. CHEMICAL ORIGIN OF OXYGEN

All oxygen in the Universe is formed along the so-called ‘main line’ sequence from the high-temperature fusion of four \(^{4}\text{He}\) atoms in hot stars. The concentration of the resulting element is approximately equal to or slightly higher than that of carbon in the solar atmospheres in this region of our galaxy (figure 1). Molecular orbital calculations reveal that the atom has six valence electrons, a valence of two and naturally forms a diradical molecule with one \(\sigma\) and one \(\pi\) bond and two unpaired electrons in degenerate lower (anti-bonding) orbitals; hence the groundstate of molecular O\textsubscript{2} is a triplet. This unusual electron configuration prevents O\textsubscript{2} from reacting readily with atoms or molecules in a singlet configuration without forming radicals (Valentine et al. 1995); however, although reactions catalysed by metals or photochemical processes often lead to oxides of group I, II, III, IV, V and even VI elements spanning H\textsubscript{2}O, MgO and CaO, AlO, CO\textsubscript{2}, SiO\textsubscript{2}, NO\textsubscript{3}, PO\textsubscript{4} and SO\textsubscript{3}, oxygen also reacts with many trace elements, especially Mn and Fe\textsubscript{2}, which in aqueous phase form insoluble oxyhydroxides. The reactivity of oxygen is driven by electron transfer (redox) reactions, leading to highly stable products, such as H\textsubscript{2}O, CO\textsubscript{2}, HNO\textsubscript{3}, H\textsubscript{2}SO\textsubscript{4} and H\textsubscript{3}PO\textsubscript{4}. The abiotic reactions of oxygen often involve unstable reactive intermediates such as H\textsubscript{2}O\textsubscript{2}, NO, NO\textsubscript{2}, CO and SO\textsubscript{2}. The reactions of oxygen with the other abundant light elements are almost always exergonic, meaning that, in contrast to N\textsubscript{2}, without a continuous source, free molecular oxygen would be depleted from Earth’s atmosphere within a few million years.

The presence of free diatomic oxygen in Earth’s early atmosphere had multiple effects and feedbacks, many of which are not yet fully understood. Free molecular oxygen allowed for a much greater thermodynamic efficiency in the oxidation of organic matter via aerobic respiration and, simultaneously, the presence of the gas in the atmosphere allowed Earth to form a stratospheric ozone layer that absorbed UV radiation (Farquhar et al. 2000). However, oxygen also greatly influenced the evolution of lipids, secondary metabolites and the selection of trace elements that are critical for maintaining electron traffic within organisms (Anbar & Knoll 2002; Quigg et al. 2003; Raymond & Segre 2006).

Given the reactivity of free molecular oxygen, the question arises: how did this gas become so abundant in Earth’s atmosphere?

3. THE WILSON CYCLE AND CARBON BURIAL

It remains unclear how the evolution of the two photosystems that gave rise to oxygenic photosynthesis within a single clade of bacteria came about (Blankenship 2001; Falkowski & Knoll 2007). The prevailing hypothesis, based largely on structural studies of the reaction centres, is that oxygenic photosynthesis resulted from lateral gene transfer.
between a purple non-sulphur bacterium with a quinone-based reaction centre and a green sulphur bacterium with an iron–sulphur-based reaction centre, giving rise to a chimeric organism. However, while both the electron transfer processes and amino acid sequences of the two reaction centre complexes are significantly different, the structural homology between them is strikingly similar, suggesting that they may have evolved from a single ancestor in one organism via gene duplication events, followed by divergent evolution (Blankenship et al. 2007). Whatever process led to oxygenic photosynthesis, this energy transduction machine is undoubtedly the most complex in nature. In extant cyanobacteria, well over 100 genes are required for the construction of the protein scaffolds as well as the enzymes required for biosynthesis of the prosthetic groups (Shi et al. 2005). Consequently, oxygenic photosynthesis is, unlike any other core prokaryotic metabolic pathway, completely isolated to cyanobacteria. Owing to extensive operon splitting, only subsections of the photosynthetic apparatus appear to be laterally transferred, and such piecemeal transfers do not appear to have been successful in establishing the pathway in any other clade of prokaryotes (Shi & Falkowski 2008). Rather, transfer of the machinery was achieved only by wholesale engulfment and massive horizontal gene transfers in a series of endosymbiotic events, leading to the formation of plastids, the second genetically identifiable organelle to be so appropriated in the evolution of eukaryotes (Dellwicke 2000). Even in the most ancient photosynthetic eukaryotes, plastids still maintain a complement of key genes required for photosynthetic function. It is not entirely clear if the initial oxidation of Earth’s atmosphere in the Late Archaean–Early Proterozoic aeons was mediated solely by cyanobacteria, or whether photosynthetic eukaryotes may also have contributed.

Regardless of the pathway(s) that led to the evolution of oxygenic photosynthesis, the biological process was a necessary, but not sufficient, condition to allow for a net accumulation of oxygen in Earth’s atmosphere. The key reaction in oxygenic photosynthesis is the water-splitting oxygen-evolving complex that contains a quartet of Mn atoms and an essential Ca atom. The origins of this metal cluster are unclear; however, the quartet of Mn atoms is clearly essential to the key reaction (Allen & Martin 2007; Barber 2008). Through sequential single photon-driven electron transfers, the Mn atoms remove electrons (without protons) one at a time from two water molecules, ultimately liberating O₂. Ca appears to stabilize the intermediate O₂ until the second atom is released. The resulting electrons (and protons) are used to reduce inorganic carbon (as well as sulphate and nitrate). The production and consumption of oxygen are almost always very closely balanced on local scales. Indeed, on time scales of millions of years, there is virtually no change in the oxygen concentration of the atmosphere or in its isotopic composition (Bender & Sowers 1994; Falkowski et al. 2005). These results strongly suggest that photosynthesis and respiration are extremely tightly coupled (Berry 1992).

On longer geological time scales, however, there are net changes in atmospheric oxygen (Berner et al. 2000). A net accumulation of oxygen in the atmosphere requires a net sink for reductants; that is, there must be an imbalance between oxygen production and its biological and abiological consumption (Holland 2006). Indeed, the very presence of oxygen in the atmosphere implies a permanent sink for reductants (hydrogen atoms). The primary biological sink for the reductants generated by oxygenic photoautotrophs is carbon dioxide, leading to the formation of reduced (organic) carbon (Berner 2004). Indeed, the counter-part of the story of O is, in large measure, the formation and sequestration of C–H bonds. Mixing of these two reservoirs would, purely from a thermodynamic perspective, consume both pools, leading to the formation of water and inorganic carbon. Hence, a permanent reservoir for organic matter is imperative if oxygen is to accumulate from water splitting coupled to carbon fixation.

By far, the largest sink for organic matter is the lithosphere (Falkowski & Raven 2007). In the contemporary ocean, the vast majority (more than 99%) of organic matter produced by photosynthetic organisms is respired in the water column or in the sediments. However, a very small fraction is buried in sediments, especially along continental margins and in shallow seas (Hedges & Keil 1995; Aller 1998). Similar processes must have operated in the Late Archaean and Early Proterozoic oceans ca 2.5 Ga ago, but some key factors differed.

It is currently thought that during the Late Archaean and Early Proterozoic, primary production in general was exclusively conducted by prokaryotes, and oxygenic photosynthesis specifically was carried out exclusively by cyanobacteria (Knoll & Baud 1989). Grazing pressure on these organisms was almost certainly nil; there were no metazoan grazers. However, there may have been eukaryotic heterotrophic or anaerobic photosynthetic protists (Fenchel & Finlay 1995). A small fraction of the organic matter sank out of the water column (and it probably was a much smaller fraction than in the contemporary ocean) to become incorporated into sediments. The rate of delivery of sediments was also probably much slower than in the contemporary oceans; there were no soils on the continents and no terrestrial plants to help catalyse weathering reactions (Berner 1997). Abiological authigenic carbonate precipitation on biological particles may have been a major source of sedimentary materials that could have facilitated sinking of organic matter.

The burial of organic matter in the lithosphere is critically dependent on the Wilson cycle (Wilson 1965; Fisher 1983). Basically, in this cycle, radiogenic heat and mantle convection leads to the fracture of supercontinents, followed by the production of oceanic basalts along mid-ocean ridges and the formation of new ocean basins (Worsley et al. 1986). As more oceanic crust forms, the newly formed basins continue to expand with passive non-subducting margins. As this cycle matures, the oceanic crust becomes cooler and denser and begins to subduct under the continental crust forming a destructive margin. Eventually the ocean is subsumed and the continents collide.
This cycle takes ca 400 Myr to complete; Earth is presently near the middle of a cycle with the opening of the Atlantic basin. The storage of organic matter along passive margins helps to reduce atmospheric CO₂ while simultaneously allowing for a production of O₂ (Katz et al. 2004; Falkowski et al. 2005).

Regardless of the initial fate for organic matter buried in marine sediments, unless it is removed from the Wilson cycle (Wilson 1965), it will be tectonically reprocessed by arc volcanism and the stored reducing equivalents will potentially be re-exposed to the atmosphere (Canfield 2005; Hayes & Waldbauer 2006). While this process may not be 100% efficient, it would greatly constrain the amount of oxygen that could accumulate in the atmosphere. Ultimately to trap the reducing equivalents more permanently requires that organic-rich sediments be uplifted onto continents. This process occurs along subducting margins leading to the formation of orogenic sedimentary units along the edges of cratons (e.g. shales along coastal mountain ridges), and in the Archaean, through the accretion of continental crust. Indeed, the largest reservoir of organic carbon on Earth’s surface are sedimentary rocks.

4. THE ROLE OF NITROGEN

The production and burial of organic matter in the ocean are not simply dependent on carbon dioxide being reduced and buried. Marine microbes, be they prokaryotes or eukaryotes, are, from a biogeochemical perspective, primarily porous bags containing enzymes that allow exchange of gases with the environment. Enzymes, being proteins, require a source of fixed nitrogen in the form of NH₄⁺ for their synthesis. Ultimately, the source of NH₄⁺ is biological reduction of N₂. The enzyme system responsible for this reaction, nitrogenase, is widely dispersed among bacteria and archaea (Postgate 1998) and contains 19 iron–sulphur clusters, which upon exposure to molecular oxygen become oxidized, thereby rendering the enzyme complex catalytically inert.

The evolution of nitrogenase, like that of photosynthesis, is not well constrained (Raymond et al. 2004). The metabolic pathway is distributed in several clades of bacteria and archaea, but has never been stably incorporated into any known extant eukaryotic genome. The energetics of the pathway are rather remarkable; approximately 16 ATPs are required to fix a single atom of N to NH₃, yet there is no known phosphorylated intermediate in the reaction pathway. The enzyme can use free H₂ to fix N₂; however, free H₂ is very rare on Earth at present and probably has not been a major gas in Earth’s atmosphere for at least 3 Ga. In vitro, the enzyme can be made to behave as an H₂ requiring ATPase (Mortenson 1964), but in vivo, the source of the reductant ultimately is organic carbon. Hence, the reduction of N₂, a reaction that is arguably as critical to the evolution and perpetuation of life on Earth as photosynthesis, is metabolically coupled to the oxidation of organic carbon. Further, given the extremely high-energy requirements of nitrogenase (a requirement that is not well understood from a mechanistic perspective), in the absence of a high ATP flux, the enzyme would be extremely inefficient. Under strictly anaerobic conditions, heterotrophic ATP supply is coupled to substrate phosphorylation, and only approximately two ATPs can be generated per glucose oxidized. The best source of ATP is proton-coupled phosphorylation; the best supply of the energy is photosynthesis. Indeed, nitrogenase is contained in some, but not all cyanobacteria. Did the ancestral cyanobacterium contain nitrogenase and then one or more subgroups subsequently lost the capability of fixing nitrogen? Or, was the capability of nitrogen fixation acquired following the evolution of oxygenic photosynthesis? Perhaps surprisingly, consensus trees of cyanobacteria, based on 16S rDNA sequences, concatenated protein sequences, and three, independent analyses of ‘core gene sets’, all suggest that nitrogenase was not present in the last common thermophilic photosynthetic ancestor of modern cyanobacteria, but rather was acquired later via one or more lateral gene transfers from another clade (Shi & Falkowski 2008). The question then is: what were the selection pressures that led to the stable incorporation of nitrogenase into this clade and how has it remained stably incorporated to present time in the presence of a machine that generates O₂ within microns of the enzyme complex that fixes N₂?

5. FEEDBACKS BETWEEN THE C, N AND O CYCLES

Let us now consider the nitrogen cycle in the context of the H, C and O cycles and examine some of the feedbacks imposed by metabolic linkages of the four elements. Before the evolution of nitrogenase, primary production in the oceans (if, indeed there was such) was almost certainly limited by the rate of supply of fixed inorganic nitrogen (Falkowski 1997). Assuming that the planetary atmosphere was not so reducing as to allow NH₄⁺ to be a major component of the atmosphere, the only known abiotic sources of fixed inorganic nitrogen would have included lightning, bolide impacts and possibly hydrothermal reactions in a reduced environment. At most, these reactions would have delivered approximately 1 Tmol yr⁻¹; a flux that would have constrained CO₂ fixation to less than 0.1% of modern values (Henderson-Sellers & Schwartz 1980; Summers & Chang 1993; Behrenfeld & Falkowski 1997; Brandes et al. 1998). While there is evidence that processes that strongly fractionated the isotopes of carbon (such as contemporary Rubiscos) were present by at least 3.8–3.4 Ga (Rosing 1999; Rosing & Frei 2004; Tice & Lowe 2004), the first appearance of cyanobacteria is not well constrained, but present biogeochemical markers place their occurrence later, approximately 2.8 Ga (Brocks et al. 1999; Knoll et al. 2007). The coincidence of the timing of the first appearance of cyanobacterial biomarkers in the sedimentary record and a global superplume event associated with a peak in continental crust growth (Condie 2004; Eriksson et al. 2006b, 2007) suggests that the influx of critical chemicals to the surface ocean may have provided critical trace elements, such as P, which potentially help accelerate primary production in the ocean via oxyogenic photosynthesis.

Phil. Trans. R. Soc. B (2008)
The accumulation of N fixed by these organisms could have increased well above the abiotic flux of fixed N until it reached the amount that would match the inventory of phosphate multiplied by the sedimenting N : P ratio (16 : 1 or greater; Falkowski 1997; Fennel et al. 2005).

There appears, however, to be a gap of some 300–500 Myr (or possibly even more) from the evolution of cyanobacteria and the accumulation of significant enough amounts of oxygen in the surface ocean and atmosphere to initiate the surface oxidation of the Earth (Summons et al. 1999). Models suggest that during this critical period a negative feedback between the oxygen and nitrogen cycles would have developed, which would have drastically constrained the accumulation of O$_2$ in Earth’s atmosphere (Fennel et al. 2005).

In the presence of oxygen, ammonium can be oxidized by chemooautotrophic bacteria to nitrate (via nitrite; figure 2). Both oxidation processes are strictly aerobic; the thermodynamic gradient strongly favours the reactions. The electrons and protons extracted from ammonium are used by the bacteria to reduce inorganic carbon without photon-derived energy; indeed, throughout most of the interior of the contemporary ocean, nitrate rather than ammonium vastly predominates as the stable form of dissolved inorganic nitrogen. However, nitrate itself is a substrate for a large number of facultative anaerobic bacteria: the denitrifiers. In the absence of oxygen, these organisms use nitrate as a respiratory electron sink, reducing nitrate back to N$_2$. Hence, at its core, the nitrogen cycle is mediated by three groups of microbes: nitrogen fixers, nitrifiers and denitrifiers, which sequentially couple the cycle of N to the reciprocal reaction of carbon via a reduction, oxidation and reduction step. That is, the initial reduction of N$_2$ to NH$_3$ is at the expense of the oxidation of organic carbon to inorganic carbon, while the oxidation of NH$_3$ to NO$_3$ is coupled to the reduction of inorganic carbon to organic matter and the reduction of NO$_3$ to N$_2$ is coupled to the oxidation of organic matter to inorganic carbon. All the three pathways are related to oxygen; the two reduction reactions are inhibited by O$_2$, while nitrification is strictly dependent upon the gas.

In modelling, the N cycle as a function of O$_2$, Fennel et al. (2005) suggest that there was a ‘node’; that is, as free molecular oxygen began to accumulate, nitrate would have initially formed from the abundant pool of ammonium, but as the oxidation state was low, denitrifiers would have efficiently removed the nitrate to form N$_2$. The model indicates that there is a critical concentration of O$_2$ above which nitrifiers successfully out-compete denitrifiers, allowing nitrate to remain as a stable intermediate in the dissolved phase of the oceans; based on experimental studies with extant microbes, that threshold concentration of O$_2$ is between approximately 15 and 30 μM. Hence, until sufficient O$_2$ accumulated in the upper ocean and atmosphere, nitrate would be rapidly depleted, causing a global ‘nitrogen crisis’, in which virtually all fixed inorganic nitrogen would be removed by denitrifiers. The amount of oxygen that could be produced under the most ideal conditions would have been constrained by the phosphate inventory of the ocean, while the rate of accumulation of O$_2$ in the atmosphere would have been constrained by the rate of burial of organic matter. In practical terms, the latter is dependent on the areal extent of shallow continental shelves, which are the main repository of organic matter in ocean basins.

6. EVIDENCE FOR THE NITROGEN CRISIS

Did this modelled nitrogen crisis really happen? The effects of oxygen production on the N cycle can be followed by changes in the isotope composition of the fixed N that was assimilated by organisms and incorporated into the sedimentary record (figure 3). Reactions within the N cycle are often accompanied by isotopic fractionations between the two stable isotopes of the element, $^{15}$N and $^{14}$N. The strongest isotope effects are associated with denitrification and nitrification, which can enrich the substrate by 22–30‰ (Cline & Kaplan 1975; Mariotti et al. 1981; Yoshida 1988; Brandes et al. 1998; Altabet et al. 1999; Voss et al. 2001; Casciotti et al. 2003). Indeed, in the modern ocean, nitrate is enriched by approximately 5‰ relative to N$_2$ primarily as a result of denitrification in the water column (Brandes & Devol 2002; Altabet 2006). Under the very low oxygen conditions of the Late Archaean, it is possible that nitrification could have been of greater importance in imparting a $\delta^{15}$N signature to fixed N than in the contemporary ocean if enhanced nitrification in modern, low O$_2$ environments (Goreau et al. 1980; Jiang & Bakken 1999) is analogous to the processes in deep time.

The isotope composition of fixed N delivered to the ocean by abiotic processes or by biological N fixation is shifted to slightly more negative $\delta^{15}$N values relative to the atmosphere, by far the largest surface reservoir of N$_2$, that has (by definition) a $\delta^{15}$N value of 0. We implicitly assume that, given the turnover time of N$_2$ of...
ca $10^9$ years, the isotopic pool of the gas has remained constant. During the Archaean, as fixed N was added to the ocean by microbes, there would have been an increase in N inventory (assuming efficient N recycling) but no expected change in $\delta^{15}$N. Ammonium adsorption into clays and its substitution for K$^+$ could have produced a weak $^{15}$N enrichment of dissolved ammonium during weathering reactions between young crust and seawater, but the production of oxygen by cyanobacteria must have had the greatest impact on $\delta^{15}$N$_{\text{fixed}}$. Early in the evolution of oxygenic photosynthesis, nitrification would have had greater impact on $\delta^{15}$N$_{\text{ammonium}}$ than denitrification since the instability of nitrate must have led to its total loss, either through denitrification or by reaction with reduced forms of S, Fe, Mn, etc., reactions that presumably would not impart an isotope signature on fixed nitrogen. While large-scale complete denitrification would have been invisible in isotopic records, its effect on lowering the ocean inventory of fixed nitrogen must have been profound. As production rates of oxygen increased in the upper ocean, it is probable that strong vertical redox gradients would have developed. Under such conditions, the inventory of fixed inorganic nitrogen in the upper ocean would reflect the balance between the flux of ammonium from the ocean interior and nitrification/denitrification rates in the surface. In such an environment most, but not all nitrates would have been lost. As these conditions became more widespread, the isotopic fractionation associated primarily with partial denitrification would imprint its signature on the sedimentary record of $\delta^{15}$N in organic matter.

Analyses of the N isotopic composition of Late Archaean and Early Proterozoic sedimentary rocks reveal a clear enrichment in $^{15}$N. These results imply that there was sufficient O$_2$ to permit nitrification/denitrification reactions to proceed. However, the S isotopic data do not show strong deviation from mass independent fractionation (MIF), assumed to arise from UV-based photochemical reactions. This apparent discrepancy can be reconciled by assuming that in the Late Archaean biologically driven O$_2$ fluxes were sufficient to allow a small accumulation (or ‘whiff’) of O$_2$, but not sufficient to allow the accumulation of stratmospheric ozone, the presence of which would have obliterated the MIF sulphur isotope signal (Kaufman et al. 2007). Indeed, the apparent contradiction between N and S isotopic signatures suggests an extremely unstable condition, where small changes in net carbon burial ultimately determine whether the atmosphere became oxidized. These signals, which persisted for several hundred million years before the MIF signal was finally obliterated, strongly suggest that fixed inorganic N was a very scarce commodity in the upper ocean; that is, the isotopic record appears to support the nitrogen crisis hypothesis generated from the coupled biogeochemical model of Fennel et al. (2005).

The superplume event that approximately coincided with the appearance of cyanobacteria may have also played a significant role in tipping the oxidation state of Earth. This event appears to have initiated the onset of ‘modern style’ plate tectonics, the growth of cratons that led to the creation of wide shelves and carbonate platforms (Eriksson et al. 2006), allowing for massive burial of organic C, a key step in the accumulation of oxygen in Earth’s atmosphere (Fennel et al. 2005). From 2.4 to 2.6 Ga, large epeiric seas occupied the

---

![Figure 3](http://rstb.royalsocietypublishing.org/)

Figure 3. (a,b) Variations in $\Delta^{33}$S and $\delta^{15}$N for the first 2.5 Ga of sedimentary isotope records that include the Archaean and the Palaeoproterozoic. The Archaean saw the start of life and the initiation of modern plate tectonics; it ended at the great oxidation event. Deviations from zero of $\Delta^{33}$S indicate mass-independent fraction of S isotopes and the absence of UV shielding ozone in the atmosphere and disappear after 2.5 Ga (see Kaufman et al. 2007). The increase in the minimum value of $\delta^{15}$N through the Early and Middle Archaean, which probably reflects the least altered values of $\delta^{15}$N, indicates changes in seawater fixed N–$\delta^{15}$N arising from seafloor weathering and initiation of oxygenic photosynthesis and nitrification–denitrification processes. Data from Hayes et al. (1983), Sano & Pilleri (1990), Pinti & Hashizume (2001), Hu et al. (2003), Mosjzsis et al. (2003), Ono et al. (2003, 2006), Jia & Kerrich (2004), Bekker et al. (2004), Johnston et al. (2005), Papineau et al. (2005), Ohmoto et al. (2006) and Kaufman et al. (2007).
passive margins of the newly forming supercontinent Kenorland and its fragments that hosted thick successions of clastic and chemical sediments. As organic matter sank from the surface ocean and was sequestered in shelf and shallow sea sediments, two processes occurred: (i) the primary driver of denitrification, organic C, was removed and isolated from available nitrate and (ii) faster remineralization of N relative to C returned fixed N to the ocean. Increased shelf slope upwelling and recycling of nutrients within these shallow seas would have increased the production of O₂ and nitrate, but rates of denitrification decreased owing to the sequestration and burial of C org. Thus, absence of a significant O₂ concentration in the upper ocean. Given an adequate supply of this critical transition metal, microbes possessing a nitrogenase (nif) operon would have a selective advantage for acquiring fixed inorganic N, albeit with a heavy energetic cost. Indeed, horizontal gene transfers of the nif operon appear to have occurred across the two major prokaryotic domains of life and ultimately were acquired by a clade of cyanobacteria. However, an alternative solution for conserving fixed inorganic nitrogen also arose: endosymbiosis.

Consensus evolutionary trees strongly suggest that mitochondria were acquired as an organelle by a (presumably) heterotrophic host cell through an endosymbiotic appropriation of an α proteobacterium very closely related to extant purple non-sulphur photosynthetic bacteria (Yang et al. 1985; Pace 1997; Embley & Martin 2006). While the primary role of extant mitochondria is to ‘supercharge’ the host cell by catalysing the aerobic oxidation of organic carbon to CO₂, this role would not have been possible in the absence of a significant concentration of O₂ in the upper ocean and the atmosphere. Under anaerobic conditions, the acquired protomitochondrion would have probably operated as an intracellular photosynthetic organelle, providing a source of carbon to the host cell, but more importantly, recycling fixed inorganic nitrogen (and possibly phosphate), which otherwise would have been excreted from the heterotroph. That is, the protomitochondrion became an intracellular septic tank, assimilating catabolically derived ammonium from the heterotroph and resynthesizing amino acids and nucleic acids, a portion of which would be returned to the host cell. Indeed, this recycling system is the very basis for a wide variety of extant, plastid-based symbiotic associations, including those observed in zooxanthellate cnidarians, symbiotic foraminifera and hundreds of facultatively photosynthetic protists (Falkowski et al. 1993). The nascent symbiont may also have fixed nitrogen, although that hypothesis is difficult to constrain. Regardless, the acquisition of an anaerobic photosynthetic organelle appears to have been the first process leading to the last major branch in the tree of life: the eukaryotes. Indeed, many extant protists are facultative anaerobes; however, the mitochondrion permanently lost its photosynthetic machinery. We suggest that the primary selective pressure for such an event was not energy production or even organic carbon acquisition, but rather conservation of fixed inorganic nitrogen during a transition period when oxygen concentrations were high enough to allow the completion of the nitrogen cycle and hence significantly denitrify the ocean, but not so high as to allow aerobic respiration to be efficient or to allow stratospheric ozone to form.

With the rise of oxygenic photosynthesis, new opportunities for endosymbiotic associations evolved; one of which ultimately led to the acquisition of a second organelle. Indeed, the second, and only other known true capture of an organelle occurred via the appropriation of a cyanobacterium by a eukaryotic cell that already contained a photosynthetic proteobacterium (Palmer & Delwiche 1998). Initially, this triumvirate of metabolic pathways was almost certainly deadly; the generation of oxygen within a host cell would have killed the anaerobic proteobacterium, upon which the host cell was now so dependent. However, repeated attempts at this type of macro-metabolic merger appear to have led to an interesting outcome. When extant purple photosynthetic bacteria are exposed to oxygen, they operate their photosynthetic electron chain in ‘reverse’; that is, the cytochrome b/c complex, which under anaerobic conditions ferries electrons from a photochemically catalysed donor to a low-potential electron acceptor that is used to reduce CO₂, under aerobic conditions is used to facilitate the oxidation of organic matter and reduce O₂ (Blankenship et al. 1995). In an aerobic setting, the organisms suppress the synthesis of their photosynthetic apparatus and become facultative aerobic heterotrophs! The efficiency of oxidation of organic carbon under aerobic conditions ultimately became extraordinary. The proton gradient (which did not have to be ‘reinvented’; it can be produced under both anaerobic, photosynthetic conditions and under aerobic heterotrophic conditions) allowed for an 18-fold increase in free energy from the oxidation of glucose. Indeed, the protomitochondrion and the proplastid ultimately emerged as the major energy production machine—the intracellular production of oxygen by the proplastid allowed the protomitochondrion to generate enormous amounts of energy from the reduced carbon produced through oxygenic photosynthesis. Simultaneously, the catabolic reactions of the protomitochondrion and the host cell, both of which would generate ammonium and phosphate, would supply the newly acquired proplastid with essential elements for its own survival.

Taming and coordinating the supplies and demands of each of these complementary metabolic pathways

7. THE EFFECT OF THE NITROGEN CRISIS ON MICROBIAL EVOLUTION: THE RISE OF EUKARYOTES

The hypothesized nitrogen crisis of the Late Archaean/Early Proterozoic aeons would have exerted a profound selective influence on the evolutionary trajectory of microbial evolution. That banded iron formations persist throughout this period, while red beds are not yet a significant geological feature, strongly suggests that Fe(II) was probably still significantly abundant in the upper ocean. Given an adequate supply of this critical transition metal, microbes possessing a nitrogenase (nif) operon would have a selective advantage for acquiring fixed inorganic N, albeit with a heavy energetic cost. Indeed, horizontal gene transfers of the nif operon appear to have occurred across the two major prokaryotic domains of life and ultimately were acquired by a clade of cyanobacteria. However, an alternative solution for conserving fixed inorganic nitrogen also arose: endosymbiosis.

Consensus evolutionary trees strongly suggest that mitochondria were acquired as an organelle by a (presumably) heterotrophic host cell through an endosymbiotic appropriation of an α proteobacterium very closely related to extant purple non-sulphur photosynthetic bacteria (Yang et al. 1985; Pace 1997; Embley & Martin 2006). While the primary role of extant mitochondria is to ‘supercharge’ the host cell by catalysing the aerobic oxidation of organic carbon to CO₂, this role would not have been possible in the absence of a significant concentration of O₂ in the upper ocean and the atmosphere. Under anaerobic conditions, the acquired protomitochondrion would have probably operated as an intracellular photosynthetic organelle, providing a source of carbon to the host cell, but more importantly, recycling fixed inorganic nitrogen (and possibly phosphate), which otherwise would have been excreted from the heterotroph. That is, the protomitochondrion became an intracellular septic tank, assimilating catabolically derived ammonium from the heterotroph and resynthesizing amino acids and nucleic acids, a portion of which would be returned to the host cell. Indeed, this recycling system is the very basis for a wide variety of extant, plastid-based symbiotic associations, including those observed in zooxanthellate cnidarians, symbiotic foraminifera and hundreds of facultatively photosynthetic protists (Falkowski et al. 1993). The nascent symbiont may also have fixed nitrogen, although that hypothesis is difficult to constrain. Regardless, the acquisition of an anaerobic photosynthetic organelle appears to have been the first process leading to the last major branch in the tree of life: the eukaryotes. Indeed, many extant protists are facultative anaerobes; however, the mitochondrion permanently lost its photosynthetic machinery. We suggest that the primary selective pressure for such an event was not energy production or even organic carbon acquisition, but rather conservation of fixed inorganic nitrogen during a transition period when oxygen concentrations were high enough to allow the completion of the nitrogen cycle and hence significantly denitrify the ocean, but not so high as to allow aerobic respiration to be efficient or to allow stratospheric ozone to form.

With the rise of oxygenic photosynthesis, new opportunities for endosymbiotic associations evolved; one of which ultimately led to the acquisition of a second organelle. Indeed, the second, and only other known true capture of an organelle occurred via the appropriation of a cyanobacterium by a eukaryotic cell that already contained a photosynthetic proteobacterium (Palmer & Delwiche 1998). Initially, this triumvirate of metabolic pathways was almost certainly deadly; the generation of oxygen within a host cell would have killed the anaerobic proteobacterium, upon which the host cell was now so dependent. However, repeated attempts at this type of macro-metabolic merger appear to have led to an interesting outcome. When extant purple photosynthetic bacteria are exposed to oxygen, they operate their photosynthetic electron chain in ‘reverse’; that is, the cytochrome b/c complex, which under anaerobic conditions ferries electrons from a photochemically catalysed donor to a low-potential electron acceptor that is used to reduce CO₂, under aerobic conditions is used to facilitate the oxidation of organic matter and reduce O₂ (Blankenship et al. 1995). In an aerobic setting, the organisms suppress the synthesis of their photosynthetic apparatus and become facultative aerobic heterotrophs! The efficiency of oxidation of organic carbon under aerobic conditions ultimately became extraordinary. The proton gradient (which did not have to be ‘reinvented’; it can be produced under both anaerobic, photosynthetic conditions and under aerobic heterotrophic conditions) allowed for an 18-fold increase in free energy from the oxidation of glucose. Indeed, the protomitochondrion and the proplastid ultimately emerged as the major energy production machine—the intracellular production of oxygen by the proplastid allowed the protomitochondrion to generate enormous amounts of energy from the reduced carbon produced through oxygenic photosynthesis. Simultaneously, the catabolic reactions of the protomitochondrion and the host cell, both of which would generate ammonium and phosphate, would supply the newly acquired proplastid with essential elements for its own survival. Taming and coordinating the supplies and demands of each of these complementary metabolic pathways
would require some serious management; the ‘command and control’ system ultimately became the nucleus of the host cell through massive transfer of genes from each of the protoorganelles (Raven & Allen 2003). Thus, although both organelles retain the capability of protein synthesis, neither organelle is capable of reproduction outside of the host cell (Palmer 1997; Martin et al. 1998).

As oxygen increasingly accumulated in the atmosphere, anaerobic photosynthetic bacteria were relegated to increasingly smaller environmental niches, but several proteobacteria retained the functional photosynthetic apparatus acquired from their anaerobic ancestors and operate as obligate aerobic photosynthotrophs in the contemporary ocean. Indeed, these aerobic anoxygenic photosynthotrophs comprise up to approximately 10% of all the prokaryotes in the upper ocean. They possess reaction centres and light-harvesting complexes that are virtually identical to those of purple sulphur bacteria; however, their photochemical activity is not productively coupled to a carbon fixation capacity and, lacking Rubisco and/or photochemical activity is not productively coupled to harvesting complexes that are virtually identical to upper ocean. They possess reaction centres and light-harvesting complexes that are virtually identical to those of purple sulphur bacteria; however, their photochemical activity is not productively coupled to a carbon fixation capacity and, lacking Rubisco and/or a reverse tricarboxylic acid cycle, cannot generate sufficient fixed inorganic carbon to grow without a source of organic carbon (Kolber et al. 2000; Kobližek et al. 2003). However, these obligate aerobic photosynthotrophs use the photochemical production of reductant and ATP to offset heterotrophic respiratory costs, thereby giving them a competitive advantage over obligate heterotrophs, especially under conditions where dissolved organic carbon is plentiful (e.g. coastal oceans; Kolber et al. 2001).

8. The effect of oxygen on metabolic processes: Feedbacks on the oxygen cycle

In general, the biological evolution of Earth can be divided into two major evolutionary aeons (Falkowski 2006). The first half, up until the oxidation of the atmosphere, we can classify as an ‘aeon of biological innovation’, when the major metabolic pathways of life that facilitate all the electron transfers, which would become biogeochemical cycles, evolved. The second half of Earth’s history is one of ‘biological adaptation’, where organisms appropriated metabolic pathways that evolved early on, but ‘repackaged’ them in new body plans which allowed them to be carried forward. Before discussing how these ‘insurance policies’ that preserved core metabolic pathways through a wide variety of planetary traumas evolved, let us consider the problem of how the oxygen cycle interacted with the metabolic pathways which evolved, for the most part, under anaerobic conditions.

The major metabolic pathways consist of multimeric nanomachines, containing numerous proteins, prosthetic groups and, for several, lipids. For example, oxygenic photosynthesis is itself the most complex metabolic machine known in nature. The minimal structure consists of two reaction centres. The electrons generated by the first water-splitting unit (PSII) are transferred to the second reductant generating unit (PSI) via a cytochrome complex. The core of the water-splitting complex is a heterodimer of two closely related proteins, namely D1 and D2. D1 is among the most highly conserved proteins in nature (Shi et al. 2005), yet the protein turns over in vivo approximately every 30 min (Prasil et al. 1992); it works for approximately 10 000 electron transfers and then becomes irreversibly damaged. The damage leads to a loss of photosynthetic capacity, especially at high light under nutrient-limiting conditions, yet literally billions of years of evolution have not led to a ‘better’ reaction centre. The damage is clearly due to the evolution of oxygen (Ohad et al. 1990). In anaerobic photosynthetic bacteria, the evolutionary orthologue of D1 does not turnover nearly as rapidly and when oxygenic photosynthetic electron transfer is blocked with DCMU, an incredibly specific inhibitor that docks to on the plastoquinone-binding site on the stromal side of the D1 protein, the protein is also rescued from degradation even at high light.

There have been literally thousands of papers discussing the mechanism of the degradation of D1 in vivo and while the exact mechanism for the process remains controversial and the repair mechanism remains almost completely unknown, it is clear that the production of a reactive oxygen species, an ‘unintended’ consequence of the production and consumption of O₂, is almost certainly the proximal cause. So why did natural selection not produce a better D1 protein rather than selecting for an elaborate repair mechanism to take the damaged protein out of the reaction centre complex and insert a new protein? This process almost certainly would be circumvented if a better D1 could evolve.

Clearly, the employment of the D1 orthologue from a PSII ancestor to function in a new task, transferring electrons from a new proximal donor (a tyrosine in the protein’s own complex) to an ‘old friend’, a quinone, via a phaeophytin intermediate (a lot of electron hopping) was a difficult task. It was a biological contract that asked for on-the-job suicide, but a contract that ensured that every time the protein died it would be recycled. The problem in breaking the deal is that D1 became irreversibly associated with a close cousin, D2. These two proteins are ‘frozen metabolic accidents’, protein structures that conduct critical reactions, yet the core machinery is maladapted to oxygen, even after 2.4 Ga of evolution (Shi et al. 2005).

9. Frozen metabolic accidents

So why has nature developed a better oxygen-evolving machine?

We suggest that the answer is: ‘Because it can’t’. Over 2.4 Ga of evolution, mutations of every amino acid in every combination of positions possible surely occurred, yet D1 is one of the most conserved proteins in the photosynthetic apparatus; the sequences are more than 80% identical between cyanobacteria and higher plants, and approximately 90% similar. The extraordinary conservation of this protein, which is irreversibly damaged by a product of its own catalysis, suggests that the restrictions imposed by protein–protein, protein–prosthetic group and protein–lipid interactions cannot be overcome.

Phil. Trans. R. Soc. B (2008)
Are there other examples of the effects of oxygen in the electron market of life on Earth? Indeed, there are several.

Rubisco is an enzyme that evolved under anoxic conditions; in the presence of oxygen, rather than carboxylating ribulose 1,5-bisphosphate to form two molecules of 3-phosphoglycerate, the enzyme can catalyse a non-productive oxygenase reaction leading to the formation of glyceraldehyde and glycolate (Tabita et al. 2008). These two products are respired, leading to an unproductive consumption of photosynthetically generated reducing power. Indeed, in the contemporary atmosphere, upwards of 40% of Rubisco activity is oxygenase in a typical C3 terrestrial plant. Many algae have developed carbon concentrating mechanisms (Raven et al. 2003), which allow the cell to increase the concentration of CO₂ in the vicinity of Rubisco (Coleman 1991); this process suppresses but does not eliminate the oxygenase activity of the enzyme. The primary mechanism for reducing the impact of the oxygenase activity is to synthesize large amounts of Rubisco (Tabita et al. 2008); indeed, this is probably the most abundant protein in nature. Such a solution requires a significant investment in fixed inorganic nitrogen and as nitrogen is frequently itself a limiting element, this investment comes at the expense of growth.

Another example is the nitrogen-fixing enzyme, nitrogenase, a highly conserved heterodimer, which invariably contains Fe–S motifs (Peters et al. 1995). Fixation of nitrogen from the atmospheric N₂ reservoir is critical for life on Earth, yet the enzyme is relatively rare in nature (Capone & Carpenter 1982). When exposed to O₂, nitrogenase is irreversibly inactivated due to the oxidation of Fe(II) in the Fe–S clusters. Indeed, at 21% oxygen, approximately 30% of the nitrogenase activity is lost in cyanobacteria (Berman-Frank et al. 2003). To compensate for this loss, the cells must overproduce the enzyme, at a great expense to their own nitrogen economy. Even after 2.4 Ga of evolution, the solutions allowing nitrogenases to function in an aerobic environment are invariably to reduce the partial pressure of oxygen in the immediate vicinity of the enzyme; this is achieved primarily through metabolic processes such as the Mehler reaction (a photosynthetic reaction that reduces O₂ back to H₂O, usually via an H₂O₂ intermediate), or by aerobic respiration (Milligan et al. 2007).

Thus, while the evolution of the oxygen cycle on Earth ultimately facilitated a much more efficient pathway for the oxidation of organic matter, it also posed feedbacks and constraints at a molecular level on key metabolic pathways including oxygenic photosynthesis itself, carbon fixation and nitrogen fixation. In every case, the ‘fixes’ for these pathways circumvented major evolutionary changes in the core pathways and, instead, built mechanisms to repair or overproduce the ‘inefficient’ pathways, leaving their defects permanently imprinted in life on Earth.

(i) Where did we come from? The question of how the electron market evolved and life came to be sustained on Earth, and
(ii) Are we alone? The question of how Earth came to be a unique planet in our Solar System, and is it a unique planet in the Universe?

The evolution of the network of non-equilibrium redox reactions that came to be the primary energy-transforming pathways of life on Earth remains largely unknown. It was not inevitable that oxygen became the ultimate electron sink. The biological chemistry of oxygen is still a mystery; we do not fully understand how oxygen is produced in the photosynthetic reactions, or its feedbacks in metabolic and biogeochemical cycles. The first of the two fundamental questions posed above reflects our ignorance of basic chemistry of the electron transfers that bring the ensemble of molecules in cells to ‘life’. Despite 150 years since the publication of ‘The Origin of Species’, the origin of life remains a black box.

Whether we are alone is also still an unanswered question. However, our search strategy relies on detecting the changes in planetary atmospheres that arise from non-equilibrium metabolic processes. Detecting O₂ on a terrestrial planet would certainly indicate the potential presence of life; however, other trace gases that cannot exist without non-equilibrium processes, such as CH₄ and N₂O, would also be strong indications of life, even in the absence of O₂ or ozone on the planet. These gases almost certainly marked a long period of Earth’s atmospheric spectrum prior to the ‘great oxidation event’, and technology for detecting them in planets far removed from our Solar System is feasible. Within the next several decades, we will almost certainly know if there is life on terrestrial planets within approximately 15 parsecs of our Solar System.

We thank NASA and the Agouron Foundation for its financial support, and Euan Nesbit and Derek Bendall for inviting us to present our ideas.

ENDNOTES

1 While some microbes can reverse a metabolic pathway and operate forward and reverse reactions, this process cannot be sustained without the influx of energy in one of the reaction pathways or it would violate the second law of thermodynamics. In other words, a cycle of elements cannot be sustained by one organism that spends part of its lifecycle in the direction and then reversing the process unless there is an input of free energy in one reaction pathway.

2 Indeed, nitrogen fixation can be thought of as a type of respiratory pathway; however, unlike other respiratory pathways, it consumes rather than produces energy.

3 By ‘significant’, we are purposefully vague as the exact concentration of O₂ required, for this process is not well constrained; however, based on experiments with extant purple sulphur bacteria, the value was probably approximately 1% by volume.

REFERENCES


