Cell PRESS

Special Focus: DNA

Oxygen, epigenetic signaling, and the evolution of early life*

Albert Jeltsch

Institute of Biochemistry, Stuttgart University, Pfaffenwaldring 55, D-70569 Stuttgart, Germany

After approximately 3 billion years of unicellular life on Earth, multicellular animals appeared some 600 million years ago, followed by the rapid emergence of most animal phyla during the Cambrian radiation. This evolutionary jump was paralleled by an increase in atmospheric oxygen, which I propose allowed the generation of epigenetic signaling systems that are essential for cellular differentiation in animals. Epigenetic signaling is based on the reversible deposition of chemically stable marks in DNA and histone proteins, with methylation of cytosine and lysine residues, respectively, playing a central role. Recent evidence indicates that the removal of such methyl groups critically depends on oxygenases. Hence, reversible epigenetic systems could only appear after accumulation of oxygen in the atmosphere.

Life was unicellular for the majority of its existence

After the origin of Earth approximately 4.5 billion years (Ga) ago, it took some time to reach the physical surface conditions that were necessary for life to develop. The first evidence of life in microfossils of stromatolites, which were colonies of photosynthetic cyanobacteria, dates back 3.45 Ga [1–3]. Modern stromatolites are still found in marine lagoons and hypersaline lakes, where high saline levels provide extreme conditions that protect against animal grazing. In addition, carbon isotope enrichment provided evidence of even earlier active metabolism, because 3.75-Ga-old sedimentary rocks in Greenland contain a $^{12}C/^{13}C$ isotope ratio that suggests an organic origin for the rock carbon [4–6]. Although the interpretation of some of these observations is still under dispute, the combined evidence suggests that life appeared relatively shortly after the end of the heavy meteorite bombardment and consolidation of the surface of the Earth (3.8–3.9 Ga). At this point, early life already must have had an established basic cellular metabolism, DNA-based genetic systems, and key structural elements.

Surprisingly, for the following 1.6 Ga these simple cells remained the only kind of living organism until the arrival of eukaryotes, which are single cells with differentiated nuclei and cell organelles. Although representing a large increase in complexity, the eukaryotes were still only single cells or cell aggregates without much cellular differentiation. It took another 1.4 Ga before complex multicellular life appeared. The first larger multicellular animals were observed during the Ediacaran period at approximately 580 Ma ago; most of these organisms were just a few centimeters long but some reached several meters in length [7]. The Ediacaran period was followed by the Cambrian period, starting 540 Ma ago, when most major animal phyla appeared within a relatively short time span, as demonstrated in the fossil record (Figure 1) [8]. This massive radiation, often called the Cambrian explosion, was already discovered in the 1840s [9]. Different explanations have been put forward for the massive and abrupt radiation in the Cambrian explosion, including climate changes, an increase in atmospheric oxygen levels, volcanic activity, or ecological changes such as the appearance of predators. In addition, the abruptness of the change itself is under dispute. These debated details are not the focus of this contribution; rather, we focus on a seemingly simpler question. After the initial appearance of life, it took 3 Ga before complex multicellular life appeared, meaning that life was exclusively single-celled or undifferentiated multicellular for approximately 80% of its history on Earth. Why did the development of differentiated multicellular species take so long?

The Cambrian radiation was accompanied by an increase in atmospheric oxygen content

The earliest atmosphere on Earth contained no free oxygen; the oxygen present today, both in the air and dissolved in water, has been produced by photosynthesis. In general, the concentration of oxygen in the atmosphere increased gradually over about the past 2.5 Ga after the proliferation of photosynthetic Cyanobacteria due to ongoing photosynthesis (Figure 1) [10]. Given the central role of oxygen in modern biochemistry, the lack of oxygen is a straightforward and widely accepted factor limiting the development of life in the Precambrian time. One central need for oxygen lies in the oxidation of food molecules to extract energy. Although there is no doubt that oxidative catabolism is much more efficient than fermentative catabolism, there are clear examples of extant larger animals that live, at least temporarily, under anaerobic conditions [11]. In addition, in light of the obvious ecological advantages of multicellular organisms and the fact that fermentative catabolism is less efficient to a similar degree for both

* This manuscript is dedicated to Prof. Dr Alfred Pingoud (Justus-Liebig University, Giessen) on the occasion of his retirement.

Corresponding author: Jeltsch, A. (albert.jeltsch@ibc.uni-stuttgart.de).

Keywords: Cambrian radiation; origin of life; oxygen; epigenetics; oxygenase; demethylation.

0968-0004/$ – see front matter

 http://dx.doi.org/10.1016/j.tibs.2013.02.001

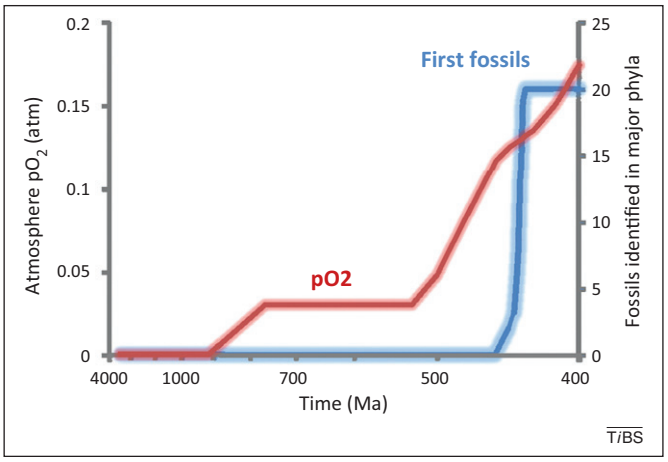

Figure 1. Cambrian radiation and atmospheric oxygen content. The red curve shows atmospheric oxygen pressure based on averages of the values reported by Holland [10]. The blue curve indicates the cumulative first appearance of fossil records in 20 major phyla as compiled by Conway Morris [8]. The *x*-axis shows time in million years (Ma) ago.

unicellular and multicellular organisms, it appears questionable whether the absence of oxygen could preclude the evolution of multicellular species for energetic reasons. Other more specialized metabolic functions that are inhibited by a lack of oxygen have also been connected to the Cambrian radiation, because the appearance of oxygen allowed a large number of new enzymatic reactions catalyzed by several oxygen-dependent enzymes called monooxygenases and dioxygenases [12,13]. These enzymes have a wide range of substrates and are involved in the metabolism of lipids (including membrane lipids), amino acids, aromatic compounds and other secondary metabolites, and nucleotides. Phylogenetic analysis indicated that the appearance of oxygen indeed allowed the generation of several novel metabolites including steroids, alkaloids, and isoflavonoids [14], and the crosslinking of collagen, which has an important role in the stability of connective tissues, is also dependent on oxygen [15]. Still, the question of what central oxygen-dependent reaction was the limiting factor in the development of multicellular animals remains open.

Cellular differentiation is based on epigenetic signaling

One important evolutionary advantage of multicellular organisms is that they can develop various cell types that are specialized for different tasks. Cellular differentiation depends on the ability of an organism to make flexible use of its genome, which enables different cells to follow different genetic programs to generate phenotypically distinct cell types with characteristic morphology and metabolism. This process represents a novel dimension in gene regulation that goes far beyond most regulatory processes in bacteria, which typically transiently change gene expression in response to immediate environmental cues. To generate different phenotypes from one genotype, a complex chromatin-based epigenetic regulatory system has been developed in multicellular eukaryotes that was probably derived from ancient gene regulatory systems. In this epigenetic system, parts of the genome are stably activated whereas others are silenced to induce a heritable lineage commitment among developing cells. Such epigenetic signaling is based on the deposition of stable chemical marks in the form of covalent modifications on histone proteins and on DNA that are propagated during cell division and determine cellular fate. These marks include acetylation, phosphorylation, ubiquitylation, and methylation of histone proteins at Lys, Ser, and Arg residues, as well as methylation of DNA at the cytosine C5 position [16–20].

Epigenetic marks are stable and reversible

The stability and half-life of these modifications vary widely depending of their chemical nature. Histone acetylation, phosphorylation, and ubiquitylation, in which the modification is connected to the protein by an ester or amide bond, can be removed by hydrolytic enzymes and are relatively short-lived. By contrast, the C–N and C–C bonds formed by methylation of lysine and cytosine residues are chemically stable and methyl groups are inert, so their removal is difficult. Consequently, lysine and cytosine methylation have a long half-life and their high chemical stability makes them an ideal modification for introducing long-lived epigenetic marks. In fact, histone methylation was considered a static modification for many years until experimental evidence showed active removal of lysine methylation [21,22]. Similarly, DNA methylation was long considered stable until experimental evidence accumulated indicating that DNA methylation can be lost not only passively through DNA replication but also by active processes [18,23,24]. However, the potential active DNA demethylation pathways were controversial and remained poorly defined for decades. In plants, DNA demethylation is initiated by removal of the modified cytosine by DNA glycosylases and its replacement by an unmodified cytosine during DNA repair [25]. For animals, this and additional pathways, including deamination of the methylcytosine followed by removal of the thymine base and repair of the resulting abasic site, have been discussed [23].

Epigenetic marks must be reversible to allow ongoing differentiation and initiation of new developmental cycles during embryogenesis. Thus, in addition to the evolution of stable chemical marks that could be used for gene regulation and signaling, enzymatic methods to remove them were also necessary to allow cellular differentiation. However, as mentioned above, removal of methyl groups from cytosine and lysine is not chemically trivial, and it appears that Nature had to play a special trick to achieve it. Molecular oxygen is one of the most reactive compounds used in biochemistry, so the attention of researchers investigating the demethylation of cytosine and lysine shifted to oxidative processes.

Demethylation of lysine and cytosine uses oxidative pathways

Long-standing and systematic efforts by several groups identified three classes of oxygen-dependent enzymes responsible for histone lysine demethylation in all species and DNA cytosine demethylation in animals. The first breakthrough came in 2004, when lysine specific demethylase 1 (LSD1), the first histone lysine demethylase, was discovered [26]. The LSD1 protein is a member of the monoamine oxidase family. It catalyzes oxidation of the methylamine group to the imine state, which allows hydrolysis and release of the former methyl carbon atom as