SPRINGER BRIEFS IN EARTH SCIENCES

George H. Shaw

Earth's Early Atmosphere and Oceans, and The Origin of Life

Springer

SpringerBriefs in Earth Sciences

SpringerBriefs in Earth Sciences present concise summaries of cutting-edge research and practical applications in all research areas across earth sciences. It publishes peer-reviewed monographs under the editorial supervision of an international advisory board with the aim to publish 8–12 weeks after acceptance. Featuring compact volumes of 50–125 pages (approx. 20,000–70,000 words), the series covers a range of content from professional to academic such as:

- timely reports of state-of-the art analytical techniques
- bridges between new research results
- snapshots of hot and/or emerging topics
- literature reviews
- in-depth case studies

Briefs will be published as part of Springer's eBook collection, with millions of users worldwide. In addition, Briefs will be available for individual print and electronic purchase. Briefs are characterized by fast, global electronic dissemination, standard publishing contracts, easy-to-use manuscript preparation and formatting guidelines, and expedited production schedules.

Both solicited and unsolicited manuscripts are considered for publication in this series.

More information about this series at http://www.springer.com/series/8897

George H. Shaw

Earth's Early Atmosphere and Oceans, and the Origin of Life

Springer

George H. Shaw
Geology Department
Union College
Schenectady
USA

ISSN 2191-5369 ISSN 2191-5377 (electronic)
SpringerBriefs in Earth Sciences
ISBN 978-3-319-21971-4 ISBN 978-3-319-21972-1 (eBook)
DOI 10.1007/978-3-319-21972-1

Library of Congress Control Number: 2015950430

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media
(www.springer.com)

Preface

Surface conditions on the Early Archean (and Hadean) Earth were critical to the origin of life. This is not a recent realization, and origin-of-life considerations have long figured prominently in geological analyses of Earth's earliest history. Several detailed treatments over the past century have discussed prebiotic processes based on assumed conditions, analyzed the (sparse) geologic data from earliest times, quantitatively explored chemical processes on and in the Earth and in its atmosphere, and attempted theoretical reconstructions of processes responsible for the formation and early evolution of the planet. Some of these efforts have also been applied to the other terrestrial planets and/or have been informed by knowledge of conditions throughout the solar system. In many of these efforts the analysis may not have given sufficient weight to one or more factors marginal to the expertise of the analyst, though in general there is a good awareness that the problems are complex and rely on information and constraints from a very broad range of disciplines. A recent spate of books discussing the origin of life suggests both a renewed interest in the topic, and perhaps no need for yet another. That a geologist feels an inclination to add to the discussion may seem misdirected. After all, isn't this a biological problem? On the contrary, I suggest that a careful consideration of the (essentially) geologic conditions leading up to life's origins can lead to a better understanding of where and how life began, and I am not the first (Rutten 1962). This book will cover such considerations in some detail and will also explore some essentially biochemical aspects which I don't believe have been previously discussed.

I will not suggest that my expertise across disciplinary boundaries is necessarily superior to those who have attempted an understanding of early Earth and the origin of life, and especially not as detailed in particular fields. However, as one who has a long-standing interest in astronomy, an abandoned goal of becoming a biochemist, degrees in chemistry and geology, research experience in geophysics and geophysical modeling, and a tendency to approach problems as a generalist, I may be better able to appreciate the range of constraints that help define the problem. A previous attempt at describing the origin and evolution of Earth's atmosphere (Shaw 2008) was somewhat limited by publication constraints, though it contains some of the ideas explored herein. In addition, even in the last few years, there have been developments/discoveries that have affected my thinking, even as I have given more thought to particular aspects of the problem.

This book will not be a critique of past efforts, though it will use and discuss them. There are several aspects of the problem that have been covered in great detail by others, and I make ready reference to those works without repeating previous expositions. I agree with the substance of the vast majority of what has been uncovered and parsed by previous work, and brief expositions of these results will often serve as starting points for my own interpretations. On the other hand, there are many points where I depart from the conventional wisdom, as well as a number of ideas that, if not completely novel, are essential to understanding the origin of conditions conducive to the emergence of life.

The book is divided into three parts. First I discuss the processes and results of the degassing and early processing of Earth's volatile inventory. Next I discuss various aspects of life processes, focusing on the requirements for the earliest emergence of life, including a possible pathway to the generation of the essential complex feedback processes thought to be characteristic of life. The third part examines processes of atmospheric evolution following the emergence of life, especially the development and consequences of photosynthesis, oxygenic photosynthesis in particular. There is a final chapter which examines the implications of the first three parts for understanding the geological and planetological record in terms of atmospheric and surface evolution.

Acknowledgments

My interest in these topics goes back to my school days, but became more focused during a course in economic geology taught by Eric Cheney at the University of Washington in 1969 or 1970. His detailed exposition of banded iron formations included implications for Earth's early atmosphere that I found fascinating. Randall Gresens, who taught a course in geochemistry at UW, was instrumental in teaching me that even apparently outrageous ideas are worth examining, and creative thinking is paramount. A few years later I just missed being a colleague of Preston Cloud, arriving at the University of Minnesota shortly after he had left. But, while teaching a course jointly with V. R. Murthy at UM, I was reminded of Cloud's work and the special features of BIF, and how they could tell us something about the Archean atmosphere. It was a couple of decades later when I had the time to return to thinking about and expanding on these matters at Union College. Brian Cohen in the Biology Department at Union provided invaluable assistance introducing me to the databases and tools of molecular biology. Kurt Hollocher read an earlier version of the manuscript and provided very useful comments. Dean Wendy Sternberg could not have known that providing a little travel support to an emeritus faculty would lead to a book contract, but it was a key step in a lengthy process.

I would like to thank my geologist daughter Katherine F. Shaw for her continued encouragement and for helping with illustrations. My philosopher son, J. Clerk Shaw, provided a nice bit of inspiration by having his own book published last year. And I especially want to thank the light of my life, Paula Fenimore Shaw, with whom I have shared the joys and sorrows of life for nearly 50 years.

Contents

Chapter 1
Introduction

Oparin (1938) provided the first modern, comprehensive treatment relevant to the origin of life. He was not interested in geologic details, many of which were unavailable at that time. His basic premise was that widespread reducing conditions on Earth's early surface were necessary (accurately so, I believe) for the abiotic processing of CHNO compounds in order to yield the building blocks of life, i.e. prebiotic compounds. His work eventually gave rise to a host of experimental studies that reveal the surprising ease with which numerous biologically important compounds can be produced. One of the most striking results of this body of experimental work is the correlation between the compounds resulting from prebiotic chemical processes and those actually used by extant life, including essentially all the monomers (amino acids, nucleotides, simple sugars, etc.) which comprise the complex polymers of all organisms.

Coincidentally, just as these chemical results were being generated in laboratories, Rubey (1951, 1955) carried out the first careful analysis of the origin of Earth's atmosphere using geological reasoning. From sound uniformitarian arguments he concluded that the excess volatiles of Earth's surface, including those now bound in rocks as carbonate minerals and organic-rich sediment, could be accounted for as the result of degassing of Earth's solid interior by magmatic emanation of various gases. He went a bit further by addressing the question of timing of this degassing, concluding that it must have been largely a gradual and continuous process rather than an early, massive discharge. This conclusion was based on both the assumed composition of the volcanic gases (taken to be similar to present volcanic emissions), and the geochemical effects should these be immediately released to the surface in quantities similar to the known excess volatile inventory of sedimentary rocks (which are currently the largest volatile reservoir by a large margin). Rubey's analysis might be faulted because it did not take into account the processes of plate tectonics, unknown at that time, which might have explained the absence of evidence of the effects he described. In addition, much more is now known about ages and amounts of early Archean rocks than was available in the early 1950s. Nevertheless, Rubey's analysis became an important starting point for discussions of Earth's earliest atmosphere and its evolution through the Archean (and later).

G. H. Shaw, *Earth's Early Atmosphere and Oceans, and The Origin of Life,*
SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-21972-1_1

Essentially all subsequent discussions of degassing, the early atmosphere, and atmospheric evolution through time starts with volcanic degassing and the related chemistry of volatiles in magmatic systems, with the major gases thought to be water, carbon dioxide (sometimes monoxide), nitrogen, hydrogen sulfide (later sulfur dioxide), and hydrogen, with the first two dominant. The proportion of hydrogen and relative amounts of carbon dioxide and monoxide have been thought to depend upon whether free iron was present in the source region of the magmas, taken as approximately the upper mantle within a few hundred kilometers of the surface. I will henceforth refer to this as the "standard model".

During the 1970s considerable attention was given to speculation and detailed modeling of planetary accretion. The formation, chemistry, and nebular processing of solids in the pre-planetary solar nebula, including interactions (or lack thereof) with nebular volatiles, was aimed at understanding the compositions and possible histories of the various components of meteorites, usually assumed to be remnants of the earliest accreting material. There is little reason to doubt that the overwhelming bulk of the planets, or at least the non-volatile fraction, was accreted from solid objects resembling meteoritic material in overall composition, perhaps in different proportions depending upon radial distance from the sun. This effort was informed by a large array of elemental and isotopic geochemical data from Earth, meteorites, and eventually other planets. In addition to providing a better understanding of the relationships of various solar system objects and their systematic variations, these efforts provide a theoretical framework for describing, if not understanding in detail, the early history of the solar system. One of the most important, and perhaps well-founded, results was the constraint on the timing of planetary accretion. Establishment of timing constraints on the cratering history of the Moon, and to a somewhat lesser extent Mars and Mercury, helped define the later accretion history, in particular the Late Heavy Bombardment.

Accretion models and the data from planetary missions are largely in agreement that accretion of the terrestrial planets was rather rapid and early, largely completed in the first 100 million years of solar system history (Wetherill and Stewart 1993; Yin et al. 2002; Wood et al. 2006; Halliday 2008). The magnitude of gravitational energy release associated with accreting an Earth-size planet implies substantial heating during accretion, especially later in accretional history. When this is combined with the additional energy that would accompany core formation, it is virtually impossible to escape the conclusion that degassing of Earth must have occurred quite early, and with the release of a large fraction of the volatile inventory from the early accreted material. This conclusion has been largely built into all subsequent treatments of Earth's early atmosphere. Whether accretion was generally chemically homogeneous, that is with the accreting material having a bulk composition similar to that of Earth (or similarly for the other planets with suitable adjustments to match their individual characteristics), or as a sequential non-homogeneous process with refractory components accreting to form a "core", then less refractory material arriving later, and finally with a volatile "veneer" last, can lead to somewhat different results.

Walker's (1977) discussion of Earth's atmosphere added a detailed analysis of the chemical processes, in particular photochemistry, important to both the modern and ancient atmospheres. The addition of this perspective is vital to understanding the short-term dynamics of atmospheric chemistry that can control the properties of the atmosphere important to atmospheric thermal structure and general climate. Walker's quantitative approach to atmospheric chemistry has been continued in essentially all subsequent work, which has gone well beyond just photochemical analysis, to include considerable thermochemistry as well, partly because of the development of models of pre-planetary nebular chemistry involving the condensation sequence of major nebular components. Walker's main model for accretion was explicitly non-homogeneous, with core, mantle and volatiles accreting more or less in that order. This is no longer widely accepted, but it is not certain that a close analysis of likely paths of *late* volatile accretion and degassing would necessarily make a large difference to considerations of early atmosphere composition. The principal reason for rejecting a simple non-homogeneous accretion model lies mainly in consideration of elemental and isotopic analysis of lead and uranium, which indicate a significant (ca. 50–100 MY) delay in core separation and equilibration of core components with mantle material. Recent studies of $^{182}Hf/^{182}W$ indicate core formation was probably no later than 50 MY. In addition, the presence of certain metal-loving components in the upper mantle indicates some late accretion of material with core-like affinities, but which clearly did not completely equilibrate with molten core components during core separation. This late accreting material is often referred to as a "late veneer" (Morgan et al. 2001; Schlichting et al. 2012; but also see Righter et al. 2008).

Walker's conclusions about the early atmosphere follow those of Rubey in that water, nitrogen and carbon dioxide were the major components, not only during degassing but as representative of the earliest atmosphere itself. One major difference in Walker is the realization that accretion dynamics and rates likely resulted in early massive degassing. The disconnect with Rubey's conclusions regarding the lack of geological evidence for early massive degassing was not addressed. One can perhaps assume that this problem was either simply ignored or could be explained as due to the operation of plate tectonics, that is, subduction would have removed most if not all such evidence.

Probably the most exhaustive analysis to date of the chemical state and changes in Earth's atmosphere (and oceans) was provided by Holland (1984). This is a true tour de force and covers essentially all aspects of the problem, starting with the origin of the solar system and the nature of accreting material, including isotopic and elemental analysis of key components (especially the rare gases), accretional history and degassing processes, chemical equilibria expected in the early ocean, early atmospheric chemical processing and dynamics of hydrogen escape, constraints on ocean chemistry through time, evidence for oxidation state of the atmosphere (and oceans) through time, and the major transitions evident from the geological record. As Holland makes clear, it is impossible to undertake such an analysis without making many assumptions, and he is very explicit about the tentative nature of many of the conclusions he draws. In this he seems to be following the views of Walker:

> Where clear evidence is lacking, it has been necessary to guess what might have happened. Much of this account of atmospheric evolution is therefore speculative. Many of my guesses may be wrong, but if my version of atmospheric history stimulates the development of more accurate histories I shall be well pleased. (Walker 1977)

I suggest few would disagree, even 30 years after Holland's comprehensive analysis. Indeed, Holland was careful to describe several unresolved issues. Although in the intervening years there have been exciting discoveries that place stronger constraints on the nature and history of Earth's atmosphere, and while many points enjoy widespread agreement, many questions are still unresolved. Without repeating Holland's reasoning, the following list summarizes many of the major points which would be considered "decided" at present (Shaw 2008):

1) Earth's volatile inventory was originally incorporated in condensed phases in the cooling solar nebula. The volatile inventory of Earth can be largely explained by degassing of a small proportion of the total amount of meteoritic solids accreted to form the planet. The importance of directly accreted volatiles is probably minimal, and homogeneous accretion of Earth was probably dominant. The Earth is strongly depleted in a wide range of volatiles compared to solar abundances, including both light and heavy rare gases.
2) Early degassing may have involved volcanic gases produced by equilibration of magmas in contact with metallic Fe-Ni, but these were probably soon replaced (following sinking of Fe-Ni into the core) by more oxidized gases. Early production of some H_2 and CO was followed by mostly H_2O and CO_2.
3) Primordial gases are still being released from the mantle, but 50 % or more of the volatiles were likely degassed during the first 100 million years of Earth's history. A magma ocean may have been an important factor in developing the early atmosphere, especially in terms of the distribution of magma-soluble volatiles such as water.
4) During the presumably brief interval during which high levels of H_2 (and CO) were initially produced, photochemical and atmospheric processes would have led to the production of reduced carbon compounds, perhaps initially CH_4. The photochemical instability of CH_4, however, would quickly have led to production of more condensable carbon compounds which would end up in the oceans. Reactions between sea water containing organic compounds and volcanic rocks at a few hundred degrees would produce CH_4 (and NH_3), but not in quantities sufficient to maintain the reducing atmosphere envisioned by Miller (1953, 1955), given the sensitivity of both gases to photochemical destruction. "The organic chemistry of the early oceans was probably extraordinarily complex" (Holland 1984). The ultimate quantity of these reduced compounds appears to be unknowable, and unknown.
5) Carbon dioxide soon became the dominant carbon species in the atmosphere, but its concentration during Archean time is unknown and was possibly quite variable. Weathering reactions involving carbonation and hydration have been going on from earliest times. The enhanced greenhouse effect necessary to counteract the faint young sun may have been due to substantially higher levels

of CO_2 in the Archean atmosphere, but very high CO_2 levels would have been required in the absence of other more potent greenhouse gases. Although there is very limited evidence to indicate precise levels of atmospheric CO_2 throughout much of the Archean, analysis of weathering near the end of the Archean suggests that the CO_2 level by that time was lower than required to maintain moderate surface temperatures by itself.

6) Several lines of evidence suggest very low levels of atmospheric oxygen throughout much of the Archean, and low levels of free oxygen by the end of the Archean, but the precise levels are difficult to determine. Oxygen levels remained low during the early Proterozoic. Evidence of oxygenic photosynthesis during the early to mid Archean is minimal, but it probably began by 2.7 Ga (possibly much earlier) and became a significant source of oxygen by the early Proterozoic. A late Archean rise in atmospheric O_2 was likely due to oxygenic photosynthesis. The continued very low level of oxygen following the advent of oxygenic photosynthesis may have been due to nutrient limitations on biologic productivity, in particular phosphorus availability.

These aspects of Earth's early atmosphere are generally accepted at present, although there are a few dissenting voices on some points. Many of these conclusions continue to be tentative, and there are several areas where the available evidence allows, if not encourages, a broader range of interpretations. Since Holland, there have been several discoveries, measurements, and modeling efforts that cast light on these remaining problems, but it is fair to say that several matters have defied the best efforts at resolution. Among these are:

1) Why is there little or no evidence for the early intense weathering and presumed accompanying carbonate deposition that should result from early massive degassing of water and carbon dioxide and a CO_2-rich atmosphere.
2) How can we resolve the Faint Young Sun problem?
3) Is there an explanation for the long-term stability of the carbon isotope signature of carbonate rocks and, to a lesser extent, organic-rich shales?
4) If oxygenic photosynthesis originated by ca 2.7 BYBP, or perhaps as early as 3.5 BYBP, why was oxygenation of the surface environment delayed for at least 500 MY and perhaps more than 1 BY? Are nutrient limitations enough?
5) Does the history of Earth's atmosphere tell us anything about the history of the other terrestrial planets, and especially of Mars?
6) What is the source of the presumably large reservoir of pre-biotic molecules necessary for life to begin and how were significant concentrations of them produced?

My goal with this exposition is not to take issue with much of what has been concluded about the history of Earth's atmosphere and surface environments. Rather, by examining several lines of evidence and providing slightly different interpretations I hope to provide possible answers to some of these remaining questions concerning Earth's Archean history. Moreover, because the origin of life is so closely connected to surface conditions prevailing at the time, and because the requirements for the emergence of life must tell us something about those conditions, I explore certain details relevant to understanding how life actually emerged from non-life.

This attempt will require detailed consideration of several points that have been discussed by previous authors, but with my own perspective. The main areas I will cover include:

1) The nature of accreting material and the history (timing) of accretion consistent with what is now known (or presumed) about the history of the early solar system.
2) The tectonic and thermal state of Earth's earliest surface and its evolution with time.
3) The implications of 1) and 2) for the primordial atmosphere and its early evolution.
4) The carbon cycle through time, and implications for carbon isotopes in rocks.
5) The resolution of the Faint Young Sun problem.
6) How conditions related to 1)–5) can shed light on the emergence of life, and a discussion of how life may have begun.
7) The timing of the emergence of photosynthesis, especially oxygenic photosynthesis, and how/why the oxidation of Earth's surface environment was delayed for so long.
8) A comparison of the histories of the atmospheres of the terrestrial planets, and especially a discussion of the history of Mars atmosphere and surface environment.

I recognize the audacity involved in this effort. It certainly would not have been possible without the major advances and syntheses of Rubey, Walker and Holland. That I depart from one of their major conclusions, namely an early atmosphere dominated by nitrogen and carbon dioxide (and water), is in no way as radical a departure as it seems on the surface. It only requires a slightly, but importantly, different take on the earliest phases of Earth history. The atmospheric conditions they envision do, in fact, come to prevail fairly early, but with vital differences in the resultant chemistry of the early ocean, as well as in the trace constituents of the atmosphere.

References

Halliday AN (2008) A young moon-forming giant impact at 70–110 Ma accompanied by late-staged mixing, core formation and degassing of the Earth. Philos Trans R Soc A 366:4163–4181

Holland HD (1984) The chemical evolution of the atmosphere and oceans. Princeton University Press, Princeton, p 582

Miller SL (1953) Production of amino acids under possible primitive earth conditions. Science 117:528–529

Miller SL (1955) Production of some organic compounds under possible primitive earth conditions. J Am Chem Soc 77:2351–2361

Morgan JKW, Walker RJ, Brandon AD, Horan MF (2001) Siderophile elements in Earth's upper mantle and lunar breccias: data synthesis suggests manifestations of the same late influx, Meteoritics and Planetary. Science 36(9):1257–1275

Oparin AI (1938) The origin of life. Dover, New York, p 495

Righter K, Humayan M, Danielson L (2008) Partitioning of palladium at high pressures and temperatures during core formation. Nat Geosci 1:321–323

Rubey WW (1951) Geologic history of sea water: an attempt to state the problem. Geol Soc Am Bull 62:1111–1147

Rubey WW (1955) Development of the hydrosphere and atmosphere, with special reference to probable composition of the early atmosphere. In: Poldervaart A (ed) Crust of the earth. Geol Soc Am Spec Pap 62 631–664

Schlichting HE, Warren PH, Yin Q-Z (2012) The last stages of terrestrial planet formation: dynamical friction and the Late Veneer. Astrophys J 752(8):1–8

Shaw GH (2008) Earth's atmosphere: Hadean to early Proterozoic. Chem Erde 68:235–264

Walker JCG (1977) Evolution of the atmosphere. Macmillan, NY, p 318

Wetherill GW, Stewart GR (1993) Formation of planetary embryos—effects of fragmentation, low relative velocity, and independent variation of eccentricity and inclination. Icarus 106:190–209

Wood BJ, Walter MJ, Wade J (2006) Accretion of the Earth and segregation of its core. Nature 441:825–833

Yin Q, Jacobsen SB, Yamashita K, Blichert-Toft J, Telouk J, Albarede F (2002) A short timescale for planetary formation from the Hf-W chronometry of meteorites. Nature 418:949–952

Chapter 2
Accretional History of Earth

2.1 The Accreting Material

Before describing the possible scenarios for planetary accretion and their relation to degassing of a primordial atmosphere it is essential to have some idea of just *what* was being accreted. There is general agreement that known meteorites are the best candidate materials based on their ages and compositions. The similarity between the compositions of planets, meteorites, and the refractory element components of the sun argues persuasively for a common origin. Presumably, Earth's volatile elements came from the same or a very similar source. Although extant meteorites may not very precisely match terrestrial compositions, either in terms of elemental abundances or isotopically, the complexities involved in accretional history may be a sufficient explanation. Furthermore, comparisons between meteorites and upper mantle compositions are confounded by the equilibration between metal and silicates during core formation as well as the effect of a possible late veneer. This is probably even more true when considering volatiles, which were undoubtedly affected by various earlier processes in the solar nebula. The idea that most of Earth's volatiles came from late comet impacts has been suggested (Frank et al. 1986; Hartogh et al. 2011), but considerations of elemental and isotopic chemistry apparently limit such contributions to less than 50% (Drake 2005; Altwegg et al. 2015). This could still be a significant amount, but there is little evidence *in favor* of a major cometary component.

Considering the complex processes that would accompany just the latest stages of accretion, it is questionable whether any simple comparison of meteorite types with primitive upper mantle compositions (including considerations of the late veneer) should be expected to produce a close match with particular meteorites or meteorite classes. Earth's upper mantle and Earth's volatile reservoirs must express the results of mixing of a wide variety of late-arriving objects with the pre-existing accreted material, and with complex processing involving impact, magmatic, hydrothermal and atmospheric chemistry. It seems unlikely that any simple model comparisons would be definitive. Aspects of some of these accretional processes relevant to Earth's early atmosphere will be discussed in more detail in the next

G. H. Shaw, *Earth's Early Atmosphere and Oceans, and The Origin of Life*,
SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-21972-1_2

section. Even if volatile components of extant meteorites do not precisely match the chemistry of Earth's atmosphere, the volatiles in that atmosphere must have come from some collection of volatile-bearing accreted materials, whether early or late arriving. Exactly what combination of the various groups of meteorite might apply to any particular (terrestrial) planet is to some extent constrained by the known chemistry of each planet; obviously we require a mix that results in a suitable final average composition. This leaves aside, for now, the question of the order of accretion, or some possible degree of inhomogeneity in the accretion process.

In the case of Earth, for which voluminous compositional data has been available for decades, the so-called "chondritic" model for the bulk silicate composition has long been considered a reference case. When a suitable iron meteorite mix is added we can get a reasonable overall composition for the Earth. The "chondritic" refers, of course, to chondritic meteorites, by far the most abundant of all meteorite classes. Chondrules are spheroidal particles of silicate minerals (mostly olivine and orthopyroxene, which are iron-magnesium silicates) that comprise a large proportion of the meteorites collectively known as stony meteorites. It is thought that chondrites may have been formed by freezing of liquid silicate droplets in the hot solar nebula as it underwent cooling. They are accompanied by a variety of additional silicates, some of which were refractory enough to have remained solid in the nebula, silicates altered by processes of hydration or other reactions in the cooling nebula, and various amounts of iron-nickel metal. They are often put into separate classes depending upon the amount of metal: H for high-iron, L for low-iron and LL for very low-iron. In addition there is a separate class, the carbonaceous chondrites, which have unusually large amounts of volatile components, in particular carbon and carbon compounds. Some rare stony meteorites are chondrite-free and have mineral compositions very different from the chondrites. They may be, in part, the result of early magmatic processes on meteorite parent bodies. Table 2.1 provides an indication of the compositional ranges, with the emphasis on iron content. Metallic meteorites come in a range of types as well, but in general they are largely iron-nickel with a few minor additional components. A final class of meteorite is composed of a roughly 50–50 mix of iron-magnesium silicates and metal.

Table 2.1 Antarctic meteorite summary

Type	Mass in grams	Mass fraction	Metal fraction in class	Overall metal fraction
CC	35254	0.010	0.000	0.000
H	914898	0.255	0.143	0.037
L	1667145	0.466	0.027	0.013
LL	284338	0.079	0.000	0.000
Other stony	89114	0.025	0.000	0.000
Stony-iron	202348	0.057	0.500	0.028
Iron	387799	0.108	1.000	0.108
Total	3580896			0.186

H, L, and LL are the three main categories of chondrites

CC Carbonaceous chondrites of all sub-types

Table 2.2 Volatile components in Meteorites. (after Anders and Ebihara 1982; Huss 1987; Palme and O'Neill 2003; Bradley 2004; Schaefer and Fegley 2007). Orguiel is taken as a typical carbonaceous chondrite for these calculations

	Mass Fraction in each type			Mass Fraction in "Average mix"			Mass fraction in mix as compounds		
	H	C	N	H	C	N	H_2O	CO_2	N_2
CI (Orguiel)	0.020	0.032	0.0032	0.0031	0.0016	0.00006			
H	0.0032	0.0012	0.000034						
L	0.0034	0.0016	0.000034				.028	.0057	0.00006
LL	0.006	0.0024	0.00005						

Even though there are several types of chondritic meteorites they are sufficiently similar to one another and to the known and presumed composition of Earth's mantle to justify the chondritic model of mantle composition, as a first approximation. The basic conclusion is that the Earth was formed by accretion of about 2/3 mass fraction of some kind of average chondrite with about 1/3 iron meteorite (to account for Earth's metallic core), without necessarily specifying a particular accretion sequence or history of accretion rate.

Because we are mostly interested in the volatiles that came to form the atmosphere and oceans it seems reasonable to focus on the components of meteorites which provided those volatiles. The Antarctic Meteorite Program has given us what appears to be a representative sample of meteoritic material arriving on Earth at the present time. Table 2.1 is a summary of those samples, from a compilation made in June 2011 (http://curator.jsc.nasa.gov/antmet/us_clctn.cfm for the most recent compilation).

The mass fraction of iron (actually an iron-rich iron-nickel alloy) in the Earth is close to 0.33. Although this is a factor of two higher than that found in the Antarctic collection summary, the difference is small enough to suggest that the material arriving now is a reasonable approximation to bulk Earth, and suggests homogeneous accretion over time. This does not, of course, prove homogeneous accretion, but it is difficult to imagine how significant inhomogeneity of early accretion could be followed 4.5 BY later by an infalling meteorite mix so similar to bulk Earth.

These values must be considered reasonable (at best) approximations given the rarity of samples and the possibility that even averages of several samples may not be representative. In addition the most volatile-rich carbonaceous chondrites (e. g, Orguiel in Table 2.2) comprise only about half of the mass of Antarctic carbonaceous chondrite samples, the other carbonaceous chondrite classes being typically lower in volatiles. This suggests somewhat smaller values of available volatiles (for degassing) should be taken if using the mass of all carbonaceous chondrites in calculating the percentage of accreting volatiles attributed to them. With that caveat it is possible to calculate the volatiles available from a given mass of "average meteorite" mix. In this analysis no volatile contribution is attributed to the iron meteorite fraction, and thus represents a kind of minimum since iron meteorites actually contain some carbon and nitrogen. For this calculation the following volatile contents (mainly using the average values from Schaefer and Fegley 2007) are combined

Table 2.3 Volatiles on Earth

H_2O in hydrosphere = 1.4×10^{24} g	(Walker 1977)
H_2O in sediments = 1.5×10^{23} g	(Walker 1977)
Total C in crust = 1.0×10^{23} g	(Walker 1977, Wedepohl 1995, Hayes and Waldbauer 2006)
Carbonate rocks in crust (as C) = 8–8.5×10^{22} g	(Walker 1977, Hayes and Waldbauer 2006)
CO_2 in atmosphere = 4.5×10^{18} g	(Walker 1977)
CO_2 production from volcanoes = 1×10^{14} g/yr	(Walker 1977; Hayes and Waldbauer 2006)
Oxygen in atmosphere = 1.2×10^{21} g	(Walker 1977)
Nitrogen in atmosphere = 4×10^{21} g	(Walker 1977)
Nitrogen in crust = 1.4×10^{21} g	(Walker 1977) also see Wedepohl 1995
Total iron resources (mostly BIF) = 8×10^{17} g	(Brobst and Pratt 1973)
C/N in cyanobacteria = 6 (mass), 7 (moles)	(Chorus and Mur 1999)
C/P in cyanobacteria = 42 (mass), 109 (moles)	(Chorus and Mur 1999)
P in Cont. crust = 1.7×10^{22} g ca. 780 ppm	(Wedepohl 1995)
Rate of erosion of continents = 3–5×10^{16} g/year	(Wilkinson 2005; Walker 1977)

with the meteorite percentages from Table 2.1, adjusting to a total chondrite fraction of 81 %.

Each metric ton of meteorite mix provides volatiles equivalent to about 28 kg of water, 6 kg of carbon dioxide and 60 g of nitrogen. Most of the water and carbon dioxide actually come from the ordinary chondrite component, along with about half of the nitrogen. Table 2.2 includes data relevant to volatile budgets for the near-surface parts of the Earth, including the crust.

Using the numbers in Table 2.2 and comparing with Table 2.3 we can see that it takes a rather small amount of meteorite mix (compared to the mass of Earth's mantle) to provide all of the excess volatiles on Earth as shown in Table 2.4. Note that the relative amounts of volatiles in the meteorite mix and Earth are quite similar, which suggests an initial meteorite mix may well have provided Earth's volatiles. Even if there remain residual volatiles in the mantle, this result strongly indicates the necessity of major volatile loss from Earth, or a high degree of early (inhomogeneous) accretion of volatile-poor material. The suggestion that massive volatile loss accompanied the presumed moon-forming impact (Lock et al. 2014) implies relatively late timing for such loss. In any case, whether Earth experienced "dry" inhomogeneous accretion or large-scale volatile loss at a late accretional stage, most of Earth's extant volatiles could have been added late in accretional history, prob-

Table 2.4 Volatiles available from Meteorite mix equal to 1 % of Earth's mass

	Potential from degassing (see Table 2.2)	Earth's near-surface inventory (see Table 2.3)
H_2O	1.7×10^{24} g	1.55×10^{24} g
CO_2	3.4×10^{23} g	3.7×10^{23} g
N_2	3.6×10^{21} g	5.4×10^{21} g
H_2O/CO_2	4.8	4.2
N_2/CO_2	0.010	0.015

ably with other late veneer material (Ballhaus et al. 2013). It is certainly possible that the earlier high temperature conditions of Earth's interior may have left some volatiles in the interior. The fact that they survived that hot phase suggests that they would probably contribute little to subsequent degassing.

2.2 Accretional History

Little detail of accretional history was available until the advent of high speed computing and the formulation of accretion models depending on massive calculations. In fact, earlier ideas about the actual timing and rate of accretion left open the possibility that accretion could have been slow enough to avoid a hot early planet. Such a possibility would have allowed for the formation of a (relatively) cold Earth followed by continuous radiogenic heating and gradual release of gases over geologic time, as originally proposed by Rubey (1951, 1955). Modeling of accretional scenarios in the 1970s indicated fairly rapid formation of terrestrial planets. These results were supported by geochemical studies indicating timing of core formation on Earth at about 50 million years after solar system (meteorite) origin (Wood et al. 2006; Halliday 2008). Such results are also consistent with the suggested moon-forming impact at about the same time. Some recent models (Halliday and Wood 2009) suggest that accretion *ended* with the moon-forming impact (MFI) at about 4.5 BYBP), which seems rather unlikely. Although the Late Heavy Bombardment (LHB) (Koeberl 2003, 2006a, b; Marchi et al. 2012) added only a small amount of additional material, its occurrence suggests a significant amount of material remaining to accrete following the MFI, such accretion taking place between the MFI and LHB, a 600+ million year interval.

Although precise timing for these events is difficult, if not impossible, to determine, the energetics of the moon-forming impact would surely have resulted in core formation if it had not already taken place. The models also indicate continued accretion for tens of millions of years and at least as late as the LHB, as the unaccreted residue impacted. Accretional models in Fig. 2.1 show various results for different average accretion rates. Also shown are a few cases incorporating a moon-forming impact at 50 million years of a Mars-sized object. All except the highest accretion rate (which is probably so high that it would result in core formation earlier than geochemistry suggests) have accretional tails extending out to 100 million years or so, and involving post MFI accretion of a few percent of Earth's total mass (Fig. 2.2).

An interesting aspect of this late accretional era is the Late Heavy Bombardment seen in the cratering history of the Moon. Although this surge in accretion rate produced significant affects on the lunar surface, the quantity of new material added to Earth (or the Moon) was relatively small, in particular in terms of additional volatiles.

These models indicate that the thermal state of Earth's surface was largely established well before the end of accretion. The MFI was a highly energetic event,

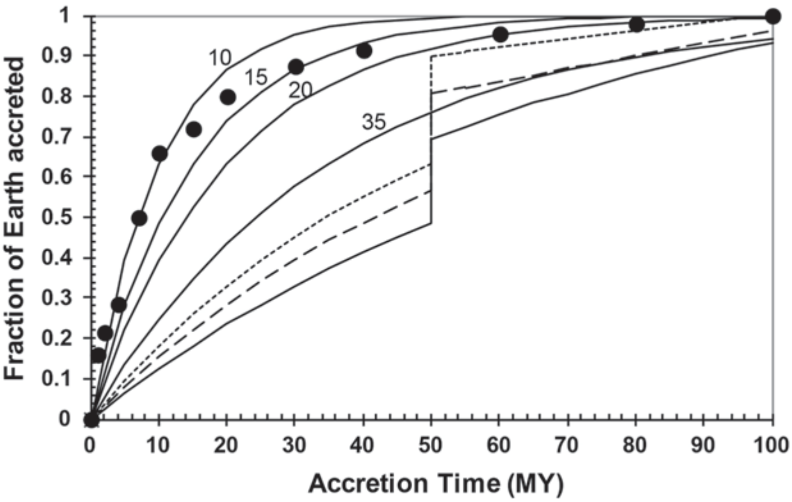
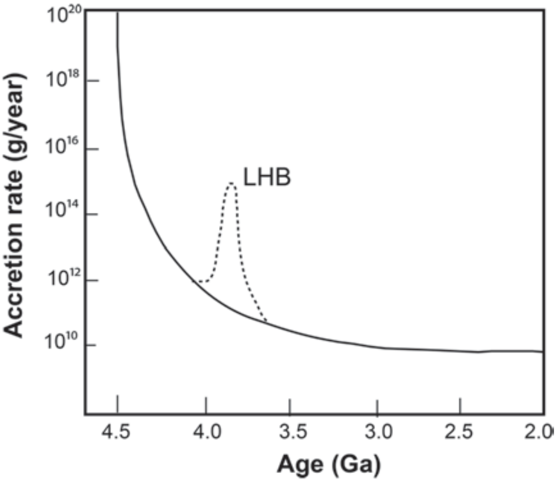


Fig. 2.1 *Accretion models for Earth*. Numbers next to curves are median accretion times in millions of years. Set of curves with "steps" at 50 million years are three cases including a Moon-forming impact at 50 MY (all after Halliday et al. 2000). Black dots are for averages of several models from Wetherill (1976). All but the most rapid of these models supports the idea that more than 1 % of Earth's mass accreted after 50 My, or after a Moon-forming impact at that time

Fig. 2.2 *Accretion rates for the Moon, including a "spike" for the Late Heavy Bombardment*. Accretion rates (mass/year) for Earth would have been at least 20 times higher. The very high, though rapidly falling, accretion rates in the Early Hadean could easily account for accretion of more than 1 % of Earth's total mass following the Moon-forming impact at about 4.5 Ga, but prior to the LHB. After Koeberl 2006a



presumably resulting in a terrestrial magma ocean of unknown, but probably great, depth. The cooling and freezing of this magma ocean seemingly marks the point at which newly accreted volatiles might have accumulated on Earth's surface. Sleep et al. (2001) calculated that, as a best guess, surface temperatures below 100C (and a solid surface, of course) would have been reached in about a million years following the MFI, although they suggest that the continued presence of a dense carbon dioxide atmosphere may have extended this to perhaps 20 million years. Even this outer limit suggests that some liquid water was present on Earth's surface while a few percent of accretion still remained, at least for most accretion models (see Fig. 2.1). This is an important point when considering the conditions under which the latest accretion, and presumably much of the degassing of the permanent atmosphere, took place.

Although large energetic impacts may well have occurred in the 50 million years or so following the MFI, none would have approached its magnitude. These large late impacts may have been important in releasing volatiles during impact and may also have blown off some fraction of the atmosphere, but it is difficult to say what percentage of the volatile inventory was due to impacts of any particular size range, or how much might have been removed. It is important to examine the conditions that would likely have arisen from a range of sizes of impact events in order to assess the nature of the earliest atmosphere. Several aspects of surface conditions may also have changed during the production of the early atmosphere and oceans, as will become clear. An important point is that evidence from early zircons indicates that liquid water (some kind of ocean) was present as early as 4.4 billion years ago, within ca. 150 million years of the beginning of accretion, quite likely even sooner, and certainly as accretion continued (Mojzsis et al. 2001; Wilde et al. 2001; Peck et al. 2001; Valley et al. 2002; Nutman 2006).

During the last stages of accretion, when it appears most volatiles arrived as components of volatile-bearing meteorites, perhaps resembling a mixture similar to the Antarctic collection, Earth's surface would have been recognizable to us. It was probably not the volcanic nightmare so popular with illustrators, but was rather dull, with a large shallow ocean and localized volcanic activity, much of which may have been submarine. Admittedly there were a lot of rocks falling from space. To add 1 % of Earth's mass over a few 100 million years gives an average of about one impact/day of a 1 m object on every square kilometer of Earth's surface, most of which would have been into a shallow global ocean. This is also equivalent to one impact each year by a 2+km diameter object. The presence of the ocean was of great importance to the processes of degassing. The next chapter discusses the nature of this early ocean, the crust beneath the ocean and the tectonic-thermal evolution of the oceans and early continents.

References

Altwegg K, Balsiger H, Bar-Nun A, Berthelier JJ, Bieler A, Bochsler P, Briois C, Calmonte U, Combi M, De Keyser J, Eberhardt P, Fiethe B, Fuselier S, Gasc S, Gombosi TI, Hansen KC, Hässig M, Jäckel A, Kopp E, Korth A, LeRoy L, Mall U, Marty B, Mousis O, Neefs E, Owen T, Rème H, Rubin M, Sémon T, Tzou C-Y, Waite H, Wurz P (2015) 67P/Churyumov-Gerasimenko, a Jupiter family comet with a high D/H ratio. Science 347:1261952-1-1261952-3

Anders E, Ebihara M (1982) Solar-system abundances of the elements. Geochim Cosmochim Acta 46(11):2363–2380

Ballhaus C, Laurenz V, Muenker C, Fonseca ROC, Albarede F, Rohrbach A, Lagos M, Schmidt MW, Jochum KP, Stoll B, Weis U, Helmy HM (2013) The U/Pb ratio of the Earth's mantle-asignature of late volatile addition. Earth Planet Sci Lett 362:237–245

Bradley JP (2004) Interplanetary dust particles. In: Holland HD, Turekian KK (eds) Treatise on geochemistry. Elsevier, Amsterdam, pp 689–711

Brobst DA, Pratt WP (1973) Mineral resources of the United States. Professional Paper P820, U.S. Geological Survey, 722 pp

Chorus I, Mur LR (1999) Preventative measures. In: Chorus I, Bartram J (eds) Toxic cyanobacteria in water: a guide to their public health consequences, monitoring and management. E. and F.N. Spon, London, pp 235–273

Drake MJ (2005) Origin of water in the terrestrial planets. Meteorit Planet Sci 4:519–527

Frank LA, Sigwarth JB, Craven JD (1986) On the influx of small comets into the Earth's upper atmosphere I. Observations. Geophys Res Lett 13(4):303–306

Halliday AN (2008) A young Moon-forming giant impact at 70–110 Ma accompanied by late-staged mixing, core formation and degassing of the Earth. Philos Trans R Soc A 366:4163–4181

Halliday AN, Wood BJ (2009) How did Earth accrete? Science 325:44–45

Halliday AN, Lee D-C, Jacobsen SB (2000) Tungsten isotopes, the timing of metalsilicate fractionation and the origin of the Earth and moon. In: Canup RN, Righter K (eds) Origin of the Earth and Moon. University Arizona Press, Tucson, pp 45–62

Hartogh P, Lis DC, Bockelée-Morvan D, de Val-Borro M, Biver N, Küppers M, Emprechtinger M, Bergin EA, Crovisier J, Rengel M, Moreno R, Szutowicz S, Blake GA (2011) Ocean-like water in the Jupiter-family comet 103P/Hartley 2. Nature 478:218–220

Hayes JM, Waldbauer JR (2006) The carbon cycle and associated redox processes through time. Philos Trans R Soc B 361:931–950

Huss GR (1987) The role of presolar dust in the formation of the solar system, PhD Thesis. University of Minnesota, Minneapolis, 250 pp

Koeberl C (2003) The late heavy bombardment in the inner solar system: is there a connection to Kuiper belt objects? Earth Moon Planet 92:79–87

Koeberl C (2006a) Impact processes on the early Earth. Elements 2:211–216

Koeberl C (2006b) Record of impact processes on the early Earth: a review of the first 2.5 billion years. In: Reimold UW, Gibson RL (eds) Processes on the early Earth. Geological Society of America (Special Paper 405), Denver, pp 1–22

Lock SJ, Stewart ST, Mukhopadhyay S (2014) Was the atmosphere lost during the Moon-forming giant impact? Lunar and Planetary Science Conference, vol 45. Harvard University, Cambridge, p 2843

Marchi S, Bottke WF, Kring DA, Morbidelli A (2012) The onset of the lunar cataclysm as recorded in its ancient crater population. Earth Planet Sci Lett 325:27–38

Mojzsis SJ, Harrison TM, Pidgeon RT (2001) Oxygen isotope evidence from ancient zircons for liquid water at the Earth's surface 4300 Myr ago. Nature 409:178–181

Nutman AP (2006) Antiquity of the oceans and continents. Elements 2:223–227

Palme H, O'Neill HSC (2003) Cosmochemical estimates of mantle composition. In: Heinrich DH, Karl KT (eds) Treatise on geochemistry: the mantle and core. Pergamon, Oxford, pp 1–38

Peck WH, Valley JW, Wilde SA, Graham CM (2001) Oxygen isotope ratios and rare Earth elements in 3.3 to 4.4 Ga zircons: ion microprobe evidence for high $\delta^{18}O$ continental crust and oceans in the early Archean. Geochim Cosmochim Acta 65:4215–4229

Rubey WW (1951) Geologic history of sea water: an attempt to state the problem. Geol Soc Am Bull 62:1111–1147

Rubey WW (1955) Development of the hydrosphere and atmosphere, with special reference to probable composition of the early atmosphere. In: Poldervaart A (ed) Crust of the Earth. Geological Society of America (Special Paper 62), Devver, pp 631–664

Schaefer L, Fegley B Jr (2007) Outgassing of ordinary chondritic material and some of its implications for the chemistry of asteroids, planets and satellites. Icarus 186:462–483

Sleep NH, Zahnle K, Neuhoff PS (2001) Initiation of clement surface conditions on the earliest Earth. Proc Natl Acad Sci U S A 98:3666–3672

Valley JW, Peck WH, King EM, Wilde SA (2002) A cool early Earth. Geology 30:351–354

Walker JCG (1977) Evolution of the atmosphere. Macmillan, New York, p 318

Wetherill GW (1976) The role of large bodies in the formation of the Earth and Moon. Seventh Lunar Science Conference 3:3245–3257

Wedepohl KH (1995) The composition of the continental crust. Geochim Cosmochim Acta 59:1217–1239

Wilde SA, Valley JW, Peck WH, Graham CM (2001) Evidence from detrital zircons for the existence of continental crust and oceans on the Earth 4.4 Gyr ago. Nature 409:175–178

Wilkinson BH (2005) Humans as geologic agents: a deep-time perspective. Geology 33:161–164

Wood BJ, Walter MJ, Wade J (2006) Accretion of the Earth and segregation of its core. Nature 441:825–833

Chapter 3
Ocean Basin and Continental Evolution

A discussion of the evolution of continents may seem out of place in a book about Earth's early atmosphere. However, there is an intimate relationship between the production of ocean crust, continental crust, and volatiles that is essential to understanding the nature of the earliest state of Earth's surface and atmosphere, as well as the changes that took place as new continental material was produced by geological processing. The obvious starting point must be the situation that existed shortly after the moon-forming impact, once the magma ocean's surface finally solidified.

For our purposes we need only a rough idea of the compositional and physical differences between the mantle and the two main types of crust: continental and oceanic. The mantle underlies both oceanic and continental crust, at an average depth of about 10 km beneath the oceans and 40 km beneath the continents. Mantle rock has a density of 3.3 gm/cm^3 and primarily consists of the mineral olivine ($(Mg, Fe)_2SiO_4$), which is also the main mineral found in chondrites, and lesser amounts of the pyroxenes (Mg, Fe)SiO3 and $CaMgSi_2O_6$. There are other minor components that contain elements such as sodium, potassium and aluminum (also found in some pyroxenes). Oceanic crust is largely made of basalt, the rock so prominent in Hawaii and Iceland, and has a density of 2.9–3.0 gm/cm^3. The main minerals in basalt are pyroxenes and plagioclase feldspar (calcium-rich feldspar, $CaAl_2Si_2O_8$, also including a sodium component, $NaAlSi_3O_8$). Again, there are lesser amounts of other chemical elements incorporated in these and other minerals. Continental rocks have average densities of 2.7–2.8 gm/cm^3 and are mainly plagioclase and alkali feldspars ($(K, Na)AlSi_3O_8$) and quartz (SiO_2). Mantle rocks are often referred to as **ultra***maf*ic because they are **ultra**-rich in *ma*gnesium and iron (*f*errous). The basaltic rocks of the ocean crust, which are produced by partly melting mantle rocks, are referred to as mafic for similar reasons. Continental rocks are produced by a bewildering array of processes that keep geologists employed and awake at night, but a critical initial step is partial melting of oceanic and mantle rock in the presence of water. Thus, both oceanic and continental crust are largely (at least originally) igneous rocks formed from silicate melts. The crust, both oceanic and continental, essentially "floats" on the higher density mantle. The thicker continental crust floats higher mostly because of its greater thickness (think of two different thicknesses

G. H. Shaw, *Earth's Early Atmosphere and Oceans, and The Origin of Life*,
SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-21972-1_3

of wood floating on water) but partly because of its lower density (compare oak to balsa wood). Because the oceanic crust floats lower than the continental crust, areas underlain by oceanic crust are at lower elevations and fill with water (hence—oceans).

The freezing of the magma ocean after the MFI may have resulted very briefly in an essentially mantle-like surface, that is, no (or very little) oceanic crust, and certainly no continental crust. This is because the MFI was so energetic that it produced complete melting of essentially meteoritic-type material, i.e. olivine and pyroxenes. Probably, the last fraction of the magma ocean to cool would have reached a basalt-like composition due to fractional crystallization, but continental-type crust was most unlikely. Because of the geological uniformity of the surface there would have been no especially high or low areas to provide dry land and separate ocean basins (possibly excepting impact crater rims). Any water condensed out of the initially hot atmosphere would have spread out across a rather smooth surface as a global ocean.

There can be little doubt that the earliest crustal magmatism was essentially basaltic, though probably of high temperature magnesium-rich (komatiitic) compositions, and produced the earliest oceanic crust. The upper and lower mantles have undergone slow secular cooling as radiogenic heat partly compensated for rapid heat loss from a hot surface. With cooling it is also likely that lower temperature, but still basaltic, magmas continued to be produced by convective processes in the mantle, even as they are the dominant magma in ocean settings to the present day. The nature of this convection may be debatable, especially the question of whether or not it strongly resembled modern plate tectonics. In any case, either large plate-like units were present, or a much larger number of smaller "platelets" were the basic units. There is little likelihood that any geological evidence bearing on this matter survives to the present.

It is certain that convective processes stirred the outer layers of the Earth, mantle-derived magmas rose and froze to make new crust resembling to a high degree modern oceanic crust, and something like subduction (though not necessarily geometrically similar to what we see now) resulted in ocean crust (and upper mantle) being transported back into the interior for recycling. This subduction process almost certainly gave rise to resurgent (secondary) magmatism resembling to some degree modern island are magmatism (as e.g., at Japan, Aleutians and Indonesia), including compositions at least superficially similar to continental crust. The available evidence indicates the presence of a very early ocean, starting about 4.4 billion years ago (Mojzsis et al. 2001; Wilde et al. 2001). This does not mean that there were continents similar to the modern equivalent, rather that crustal rock with comparatively low density, and thus relatively high buoyancy, must have been accumulating from very ancient times. Presumably there was enough stability in these early convecting systems to localize the production of this primeval continental crust and allow the accumulation of thick (relative to oceanic crust), continental-type masses ("croutons" not cratons). Once again, think of Japan, Aleutians, Indonesia.

If we assume that degassing of the first 5 % or so of the late impactors (after surface stabilization following the MFI) resulted in a proto-ocean with a mass approximating 5 % of the present ocean, we may be able to describe what Earth's

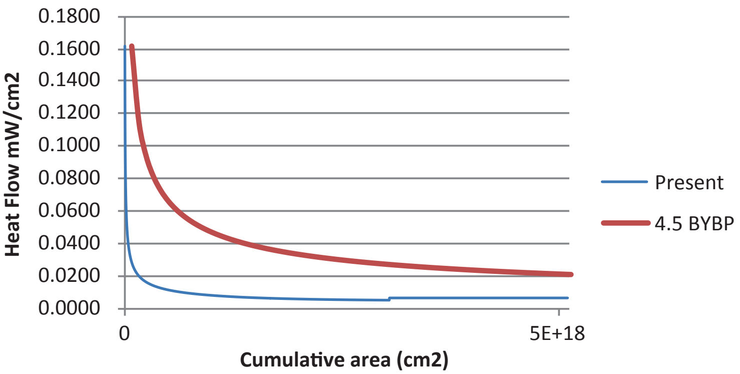


Fig. 3.1 Heat flow versus cumulative area for Earth. The much higher heat flow for 4.5 BYBP is entirely through ocean crust, with no continents yet formed. The high heat flow is mostly due to much younger average crustal ages due to faster rates of convection

surface looked like at that time. Simply put, there were perhaps several small volcanic (island) arcs scattered about a nearly global ocean basin, which would have been covered by a relatively thin layer of water produced by the small amount of degassing (i.e. 5 %) up to that point. It is certainly possible that some atmosphere (and perhaps even more ocean water) may have survived the MFI, but for the moment let us use the 5 % figure for illustrative purposes.

On the present-day Earth thin oceanic crust in isostatic balance (i.e., "floating" on the denser mantle) with thick continental crust defines the basins filled by water, i.e. the oceans. The current global tectonic regime is the result of continuous convective heat loss (and conductive cooling of convectively emplaced ocean crust and upper mantle), which accounts for perhaps 80 % of total heat flow. In Fig. 3.1 the lower (thin blue) line depicts an "average" heat flow distribution at the present time, with very high heat flow at ocean ridges (far left side) and moderate average heat flow from continents (right side), a large fraction of the latter being produced by radioactive decay *within* the continental crust. The slight jump a bit more than half way across the graph is the transition from oceanic to continental crust. The upper (thick red) curve is a schematic approximation to the heat flow situation soon after the establishment of a solid surface following the MFI. It is likely that even higher values existed for an unknown, but brief, interval (perhaps several tens of millions of years) during freezing of any residual molten mantle from the magma ocean phase. The state shown is intended to illustrate the thermal regime in which convective heat emplaced as a thermal boundary layer is responsible for essentially all heat loss, similar to the current plate tectonic regime. There is no implication that the actual tectonic regime closely resembled the present-day situation, only that convective flow continuously produced a conductive boundary layer at a sufficient rate to carry off the additional radiogenic heating that prevailed at 4.5 BYBP (ca. 4 times the present, not including possible short-lived radioactive isotopes). Obviously some additional heat loss would have been necessary for secular cooling of the mantle.

In order to produce the red curve of Fig. 3.1, that is, to get higher total heat flow, it is necessary to invoke either a faster ocean ridge "spreading rate" or a longer length of ocean "ridge crest", or some of each. (The quotes are to indicate the uncertainly about the tectonic details.) For the purposes of this calculation a spreading rate roughly 5 times the current average rate *and* a ridge crest length about 7 times the current length was necessary to yield the requisite heat flow. There is no way of knowing what combination would best describe the actual situation at 4.4–4.5 BYBP, but these values at least suggest something about the likely appearance of Earth's surface at the time. As stated earlier, some of the more fanciful sketches showing widespread volcanism, etc., may be rather overdrawn. The parameters used here imply "platelet" dimensions of 1000 km or so and perhaps 50–100 such platelets. Magmatism would have been locally prominent, but there would have been vast areas of mundane, mostly submarine, lava flows. Note that there were NO continents, as such, present. Continental-scale masses are the result of prolonged magmatic processing at ridges and at subduction zones, and were probably hundreds of millions of years or so in the future.

The solid surface, mostly thinly water-covered, was the "target" for incoming meteoritic material bringing significant amounts of fresh primordial volatiles. Perhaps the very earliest of this material arrived on a relatively dry surface, although it is clearly possible that some volatiles, especially water, were retained at the surface even soon after the violence of the MFI, though it is hard to see how we could ever know. This picture is clearly an approximation but can be used to ascertain the surface conditions likely to accompany subsequent accretion and degassing. Figure 3.2

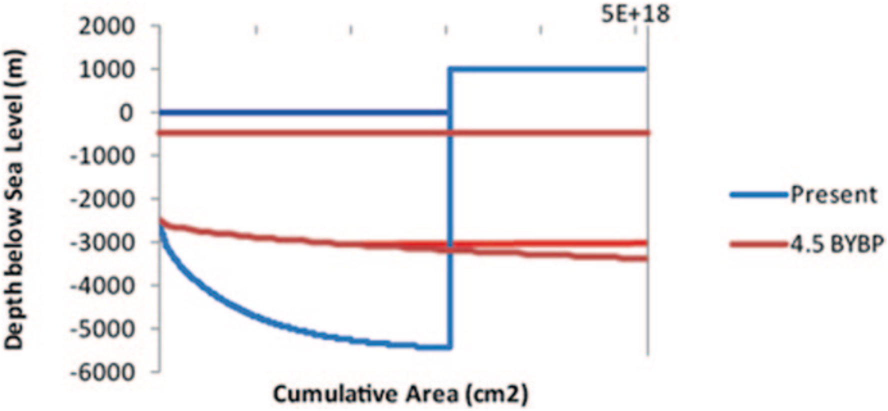


Fig. 3.2 Crust surface topography. The (*red*) curve for 4.5 BYBP is entirely oceanic. The curve for present gives average elevation for continental areas with no attempt to indicate mountains, etc. Sea levels are shown by *horizontal lines*, for oceans with present ocean volume and for 4.5 BYBP two levels, one with the present ocean volume and one with an ocean 5 % of the present volume. Even as little as 5 % of present ocean volume is enough to produce a water surface over more than half the Earth. Two to three times as much water (from earliest degassing of freshly accreted meteoritic material, i.e. post MFI accretion) would produce a nearly global ocean. Note that the offset in sea level at 4.5 BYBP is the result of ocean basin filling by continents that is not compensated for by increased sea floor depth due to cooling

shows the “average” topographic state of the present Earth (blue lines) and the 4.5 BYBP Earth (red lines). This is not a topographic profile, but is simply the areal distribution of elevations (hypsograph) of the crustal surface in both ocean basin and continental areas. At present the ocean basins are filled with the roughly 1.3 billion cubic kilometers of sea water, as shown by the horizontal blue line above the ocean basin. Continental areas are assumed to extend to the edge of the continental shelf.

The 4.5 BYBP curves show a complete absence of continental crust, and sea levels (two horizontal lines) for an ocean as voluminous as at present (upper line) and for an ocean 5 % of the present ocean volume. This shallow proto-ocean would have covered roughly 60 % of the total area, almost the same as the modern oceans, but the dry land in this case would have been the topographic highs associated with the spreading centers. This situation was undoubtedly rather short-lived because continued accretion and degassing of the accreted material, over perhaps 50 million years or so, would have added much more water. I am assuming that a volume equal to the present ocean volume was achieved fairly quickly, that is, degassing of the late accreting meteoritic material was essentially complete during or shortly after accretion. Still later accretion may have added a little, perhaps especially during the Late Heavy Bombardment, but this would not change the picture significantly. Although the last 1 % of accretion would also have added to Earth’s surface area, that would be a minor correction to this picture.

The flatter profile of the early ocean floor is simply a reflection of the more widespread high temperature mantle beneath the generally younger crust (in terms of time since crust formation at upwelling zones). This, of course, is how the necessary additional heat loss is accomplished. The modern ocean floor has a steeper topographic profile away from the ridge crests and modern mean ocean depth is significantly greater. These factors, combined with a greater depth of the modern ridge crest below sea level, are necessary to accommodate the volume of ocean water displaced by continental growth. The question of gain or loss of water from/to the mantle is subsidiary to the general picture, unless very large changes are contemplated, which seems unlikely.

The co-evolution of the ocean basins and continental crust is an interesting topic, but not vital to a discussion of the production and evolution of Earth’s early atmosphere. The important point for the early atmosphere (and degassing) is the clear presence of a large, if rather shallow, proto-ocean during most of the accretion and degassing of the late meteoritic veneer, which is the source of most of Earth’s volatile inventory. The presence of a widespread ocean (which would have expanded to essentially global scale with continued accretion and degassing of an additional 5–10 % of present ocean volume) has a profound effect on the geochemical conditions during late accretion and degassing. It also impacts heat flow by means of hydrothermal circulation. This latter effect is partly included in the picture above (Fig. 3.1), which overestimates “conductive” heat flow near the ridge crest, as a surrogate for substantial hydrothermal convective heat flow near modern ridge crests as well as at 4.5 BYBP. The presence of a global ocean implies hydrothermal alteration/hydration of at least the *upper parts* of the ocean crust as well as cycling of sea water (including dissolved components) through hydrothermal systems. This is perhaps most obvious at “ridge crests”, but is also implied at “subduction zones” where hydrated ocean crust is transported back into the mantle and reheated. This latter process is presumably responsible for the initial (and subsequent) generation of continental-type crust, even as it is in modern plate tectonic settings, and would almost certainly have included shallow hydrothermal release of volatiles at subduction zones even before magmatic temperatures were achieved at greater depth.

The proto-ocean serves another chemical function as a reservoir in which volatiles collect and react with one another. Ultimately the ocean is the site of the origin of life, but pre-biotic chemistry began almost immediately after the proto-ocean formed. Given the high heat flow and consequent hydrothermal activity there was a huge energy source available in addition to solar input. To what extent these various energy resources contributed to pre-biotic chemistry and ultimately to the origin of life is probably unanswerable, but it is hard to argue against the notion that the more (and more varied) energy sources available, the better.

Finally, a large body of surface water was certainly a reagent itself as accretion proceeded to build the Earth's volatile inventory. Impact energy drove numerous chemical reactions involving the impacting object, the solid surface, and the water in the ocean (as well as any dissolved components). Although complex in detail, there are certain aspects of the chemistry that are readily predictable from what we know of each of these chemical reservoirs. Much of this chemistry took place in the primitive atmosphere above the oceans, but included material contributions from the condensed phases in the ocean as well as any pre-existing gaseous phases. The nature of these processes is the topic of the next chapter.

References

Mojzsis SJ, Harrison TM, Pidgeon RT (2001) Oxygen isotope evidence from ancient zircons for liquid water at the Earth's surface 4300 Myr ago. Nature 409:178–181

Wilde SA, Valley JW, Peck WH, Graham CM (2001) Evidence from detrital zircons for the existence of continental crust and oceans on the Earth 4.4 Gyr ago. Nature 409:175–178

Chapter 4
Degassing Processes and Chemistry

There are three main processes that produce degassing of accreted meteoritic material. During impact minor degassing begins with atmospheric ablative heating and dispersal directly into the atmosphere. This is followed by release of kinetic energy on impact, which for a planet the size of the Earth vaporizes a significant fraction of the meteorite (Ahrens et al. 1989). Impact processes *may* release the major fraction of volatiles in accreting meteorites, though this depends on the size of each impactor. Energy release on impact can also have a profound effect on surface conditions immediately following the impact, mostly for larger impactors. Using our assumptions for late-stage accretion, namely one to a few percent of Earth's total mass arriving over ca. 50+ million years after the MFI, we can get some insight into the late accretion environment of Earth's surface. One percent of Earth's mass is about 6×10^{25} g. Spread over 50 million years gives an accretion rate of 1.2×10^{18} g/year or 4×10^{10} g/s. This is roughly equivalent to a 25–30 m diameter object arriving every second and releasing about 5×10^{15} J/s, or about 1 mW/cm^2 averaged over the entire surface. This can be compared to solar input of more than 100 mW/cm^2 and average terrestrial heat flow (at that time) of about 0.3 mW/cm^2. (Note that this calculation is over a shorter period of time (50 MY vs 300 MY) compared to a similar calculation at the end of Chap. 2). Although impact energy would not on average be a major factor in surface energy balance, the local physical effects would be tangible and the atmospheric display spectacular. The images of this era showing widespread volcanism may be overdrawn but the depiction of impacts seems closer to reality. Two additional degassing processes follow the impact. Meteorite material that arrives at the surface intact or rains out from still solid impact generated debris can be processed by heating when it is carried into the interior by convection (subduction). During the shallow phase of subduction, heating takes place under water-rich conditions which can liberate reduced volatiles (e.g. methane and ammonia) at relatively low temperatures and at high water pressure. Deeper convection/subduction leads to magmagenesis and further degassing at high (magmatic) temperatures. Each of these processes implies particular thermochemical environments that affect the chemical composition of the volatiles released.

G. H. Shaw, *Earth's Early Atmosphere and Oceans, and The Origin of Life*,
SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-21972-1_4

Of course, accretion was not uniformly distributed either in terms of rate or size of impactors. In all likelihood rates were somewhat higher early on, while size distributions are speculative. Perhaps we can get some insight into size distributions by examining the impact histories of the other terrestrial planets, though there is no guarantee that the very last phase of impact history (that which we can see now as crater distributions) mimics the earlier phases, which necessarily will have been obliterated. Nevertheless, it is important to look further into the elements of possible impact history. Even though we cannot know the details of size distribution of impactors, there are size-dependent effects we should consider. For example, at the large end of the distribution, the atmosphere (and ocean) will have little direct effect on the incoming object, which will strike the solid surface virtually unimpeded releasing most of its energy on impact. At the small end of the distribution, which may include objects up to several tens of meters or larger, atmospheric effects can be substantial. This is most obvious for objects of dimensions commonly seen in collected meteorites and those which "burn up" on entry. Even meteorites that produce impressive impact structures can be strongly affected by passage through the atmosphere. It has been suggested that the object that produced Meteor Crater, AZ may have initially been 50+ m in diameter, but it partially fragmented on entry to the point that the main impactor that reached the surface and produced the crater was roughly half the original mass (Melosh and Collins 2005). The rest of the object was broken off into smaller pieces, many of which survived to strike the surface and be preserved, *without vaporization*. This latter is important when considering subsequent degassing. Basically, some fraction of the accreting meteoritic material will reach the surface more or less intact, as we know from modern meteorite finds, while the larger meteorites will strike the surface at high speeds, resulting in a high degree of vaporization, and therefore immediate degassing, but also dispersing fragmental material that was not vaporised. That fraction of meteorite reaching the surface intact will subsequently be processed under conditions dictated by prevailing surface (and subsurface) conditions.

We can only guess about the relative total masses in these two "end members", but if we examine a few cases we will probably be covering the full range of conditions (and consequences) associated with degassing of accreted material. At the large mass end it is clear that the maximum impactor size must be much smaller than the MFI, which was probably about 3000 km in diameter. A 500 km diameter object has a mass a bit less than 1 % of the total of the hypothetical late veneer material. Several such objects could provide water for a small proto-ocean, and each impact of an object that size is capable of vaporizing the equivalent of the modern ocean (Zahnle and Sleep 2006). The conditions accompanying any such impact probably approximate the most extreme surface conditions following the MFI. Table 6.1 is an attempt to quantify certain characteristics of these surface conditions for objects covering a range of sizes. There is a major difficulty imposed by the impulsive nature of these events. Peak conditions may be very short-lived making it difficult to estimate their consequences, at least in terms of atmospheric chemistry. Nevertheless, both thermodynamics and kinetics may give us some idea of the chemical consequences of large impacts, at least to a first approximation. Table 6.1

includes results for two orders of magnitude in size range. The largest is probably above the likely high end; sizes less than the smallest would be objects that impact without significant vaporization, where we arrive at the other end-member for accreting material.

The effects of a roughly 500 km impactor have been examined in detail by Zahnle and Sleep (2006) (actually their "test object" was taken to be about 450 km, but the difference is unimportant). They describe in detail the circumstances following such an impact, including the formation of a silicate vapor atmosphere that lasts a few months. The high temperature atmosphere radiates both to the surface and to space. The former process cools the atmosphere by melting surface rock during the highest temperature phase. The latter process ultimately results in atmospheric and surface cooling to the point of water condensation after about 1000 years, i.e. at an atmospheric temperature of about 100 °C. Zahnle and Sleep did not, however, examine details of atmospheric chemistry, particularly of carbon and nitrogen. This is important for an understanding of the oxidation state of the early atmosphere and surface environment.

The direct impact degassing of a large "average composition" meteorite, say 500 km in diameter, results in injection of a considerable volatile inventory into the atmosphere. Using the data of Table 4.1, with the mass of such an object, one can calculate the release from the impactor itself of about 9×10^{21} g of water (equivalent to 10^{21} g of hydrogen), 1.7×10^{21} g of carbon dioxide, and 1.8×10^{19} g of nitrogen. This calculation does not imply that these are either the compounds within the meteorite body or the components that would have remained in the atmosphere after equilibration. In fact meteorites contain little carbonate or carbon dioxide. Most of the carbon and nitrogen in meteorites are in reduced organic compounds which also incorporate an important fraction of the hydrogen, some of which is also certainly in the form of various hydrous compounds. The chemical details surrounding meteorite degassing have been analyzed (Fig. 4.1) (Schaefer and Fegley 2007; Zahnle et al. 2010). Although these calculations are for equilibration between condensed and gaseous phases rather than conditions under impact degassing as such, they are instructive and may reflect overall equilibration of impact generated gases, liquids and solids. They directly apply to processing of meteoritic material that survives impact and is degassed by heating at/near Earth's surface. The proportions of gas phases that would be produced by heating various meteorite compositions as a func-

Table 4.1 Post-large-impact atmosphere scenarios

Diameter (km)	Mass (g)	Possible H_2 produced (g)	H_2 pressure (from Fe) (bar)	T (K)	Time until T drops to 650 K (year)	H_2 loss while T drops (g)
1000	2.4×10^{24}	2.4×10^{22}	5	2500+	5000+/−	10^{17}+
500	3×10^{23}	3×10^{21}	0.6	2500	2500+/−	4×10^{16}
100	2.4×10^{21}	2.4×10^{19}	0.005	ca. 373 after→	1 day	0
50	3×10^{20}	3×10^{18}	0.0006	ca. 373 after→	< 1 day	0
10	2.4×10^{18}	2.4×10^{16}	5×10^{-6}	little lasting effect produced, approx. K-T event		

tion of temperature, including carbonaceous chondrites and H, L, LL and high iron enstatite chondrites (chondrites containing substantial amounts of the pyroxene enstatite), include considerable fractions of reduced gases under readily imaginable conditions. The calculations cover temperatures ranging from 300 to 2500 K and pressures from 10^{-4} to 10^{4} bars. Although the largest impactors might produce somewhat higher temperatures initially, these temperatures would be very short-lived. The effect that a major additional iron-metal component would have on gas compositions were not explicitly included, that is, the calculations are for various possible "pure end member" compositions rather than a hypothetical average meteorite mix. One could argue that smaller impactors might well resemble one or more "pure" meteorite compositions, but it seems likely that large impactors would have more complex compositions and include a metal core or at least a large metal component.

The importance of the metal component can be seen in Table 4.1, which includes a column showing the hydrogen produced from the iron component of a hypothetical average meteorite, amounting to about 18% of the total mass. The argument for the production of this hydrogen is simple. Assuming complete vaporization of the impactor, all of the iron is released as vapor, into an atmosphere containing abundant water from both the meteorite itself and the presumably evaporated small proto-ocean. Reaction between the iron and water liberates hydrogen. It is certainly possible that some iron will survive as particles or condense to settle to the surface, but as a first approximation it is reasonable to assume complete oxidation of the metal. This may even be an underestimate if the impactor had an average composition similar to Earth (i.e. 1/3 metal). Somewhat smaller impactors, which do not produce a molten rock surface layer following impact, as well as lower atmospheric temperatures, could also produce metal "rainout" onto the solid hot surface. This iron would subsequently react with hot rainwater, also yielding hydrogen to the atmosphere. In other words, hydrogen pressure in the early atmosphere is not only the result of the original hydrogen compounds in the meteorites, but is significantly augmented by iron-derived hydrogen, which is, in fact greater (by about a factor of three) than that provided by CHNO compounds in the meteorite.

The hot hydrogen-rich atmosphere would certainly lose hydrogen to space, at rates far higher than the present-day loss of hydrogen from Earth. The high temperature conditions would be relatively short-lived, however, and atmospheric reactions between hydrogen, carbon (especially), oxygen and nitrogen would "fix" hydrogen in CHNO compounds as the atmosphere cooled. The last column in Table 4.1 gives an indication of total hydrogen loss expected (and thus hydrogen retained) during the high temperature phase of the atmosphere, assuming a (probably rather high) hydrogen loss rate a million times the present-day loss rate, and cooling times estimated from Zahnle and Sleep (2006). Obviously this is just a rough guess, but it appears that the post-impact atmosphere will retain a large fraction of hydrogen after temperatures have dropped to relatively cool conditions. This hydrogen abundance strongly impacts atmospheric chemistry, making methane (and possibly ammonia) the stable carbon (and nitrogen) species. Although these conditions are suggested for a meteorite mix, they are consistent with the calculations of Schaefer and Fegley

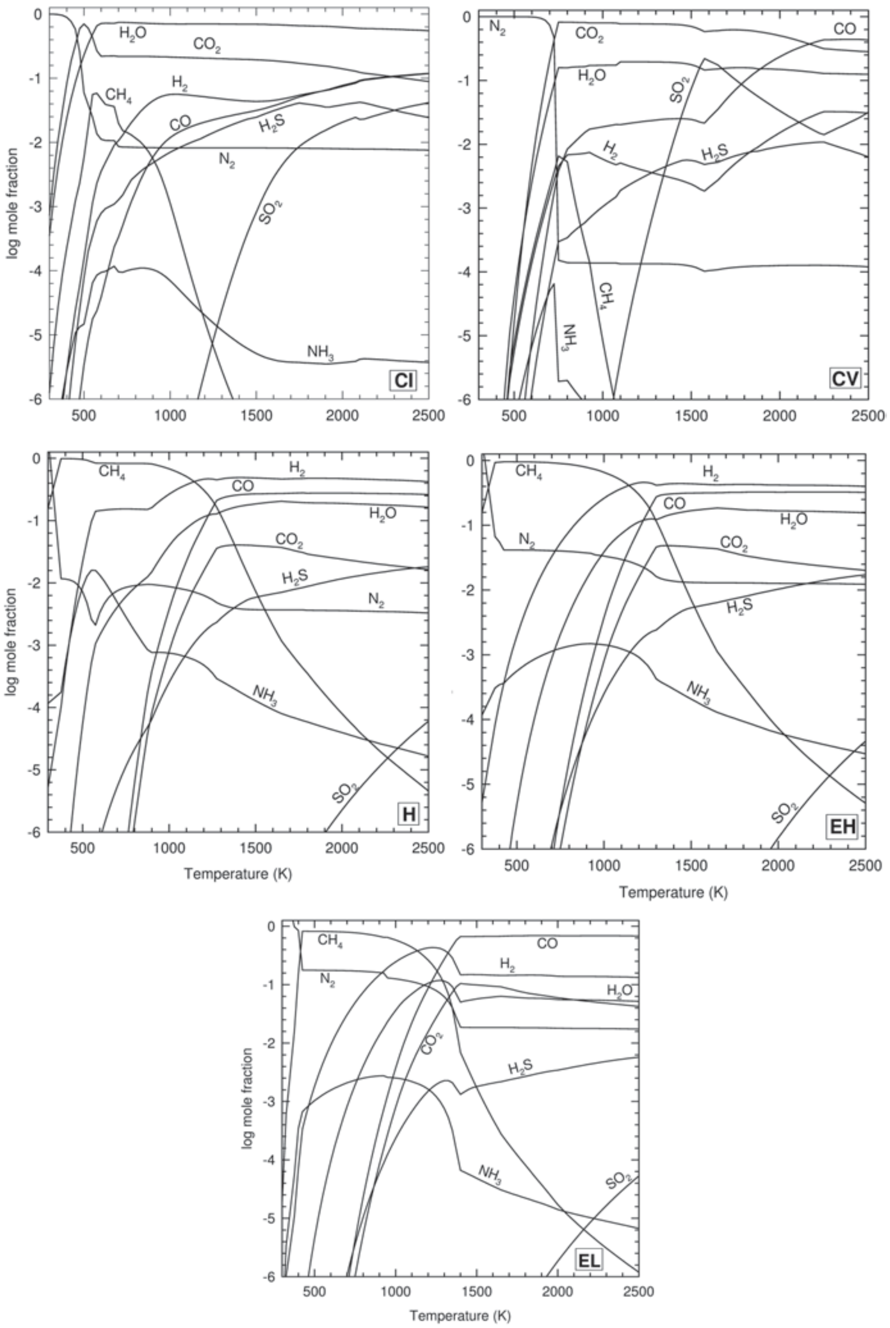


Fig. 4.1 "Atmospheric" compositions for degassing of meteorites. Pressure is set at 100 bars. Meteorite types indicated at *lower right* of each graph. CI = CI carbonaceous chondrite, CV = CV carbonaceous chondrite, H = high iron ordinary chondrite, EH = high-iron enstatite chondrite and EL = low-iron enstatite chondrite. (After Schaefer and Fegley 2010)

(2007) for the heating of various types of ordinary chondrites. A mix containing a percent or so of carbonaceous chondrite would include additional C, H and N from that component, but the overall gas reactions would strongly resemble those for the ordinary chondrites (plus additional hydrogen). Note that smaller meteorite impactors would probably be dominated by "pure" types of individual ordinary chondrites, for which Schaefer and Fegley's calculations would apply more directly.

One question raised by Schaefer and Fegley remains to be addressed. Their equilibrium calculations clearly apply at high temperatures where reaction rates are high, and where the dominant gases are oxidized (mainly CO and CO_2). They point out that reaction kinetics may become an issue as the gases (atmosphere) cool to temperatures at which methane and ammonia become more important (i.e., ca. 1350 K and down to 700 K+). They point out that studies of volcanic gases suggest equilibrium may be maintained down to 700 K (Gerlach 1993; Symonds et al. 1994) and that metal (or maybe silicate?) catalysts may be important at still lower temperatures. In the case of smaller impacts where the impact cloud expands and cools relatively quickly, there would probably be plenty of catalytic material (silicate and metal dust) still entrained in the cloud to maintain equilibrium. In the case of both small and large impacts, the likely accompaniment by smaller meteoroids would add both local energy and catalytic material.

In fact, it may not even be important to consider the equilibrium production of either methane or ammonia in the post-impact atmosphere. The presence of large amounts of hydrogen, by itself, is likely to lead to the production of photochemical products of higher molecular weight. Tian et al. (2005) have proposed an early atmosphere model with a high atmospheric fraction of hydrogen produced by an exceptionally low hydrogen loss rate from the upper atmosphere compared to the modern value coupled with higher mantle hydrogen production, as a means of overcoming the problem of synthesis of pre-biotic organics, a problem long recognized as an accompaniment to the standard model of a CO_2-rich early atmosphere. Regardless of the validity of their assumptions concerning hydrogen loss rate, it is clear that a hydrogen-rich atmosphere, whatever its origin, results in significant abiotic carbon "fixing". Whether the intermediate step of methane (and ammonia) production takes place, the post-impact scenario points in the direction where most of the carbon, and significant amounts of nitrogen, are condensed into higher molecular weight organic (but abiotic) products by atmospheric photochemistry, followed by rainout into the recondensed ocean. This analysis applies to the largest likely impactors. For smaller impactors the situation changes somewhat.

Before discussing the post-impact degassing scenarios it is important to address an even more extreme suggestion about impact degassing as the origin of Earth's atmosphere, namely production of the atmosphere by the MFI itself. A recent discussion of generation of the early atmosphere as it relates to the origin of life closely follows the standard model (Sleep 2010) with the MFI as the starting point. In this example MFI degassing produces a CO_2-rich atmosphere followed by reprocessing of the atmospheric gases in and above a magma ocean, subsequent carbonation/decarbonation of ocean crust, and some form of production of prebiotic compounds through the late addition of reducing capacity, mostly in the form of hydrogen, presumably produced from the mantle or by serpentinisation of ultramafic rocks. The idea of a late veneer is rejected, largely on the basis that the Moon shows little or no evidence of its own late veneer, which would be expected if 1 % of late veneer (after the MFI) were added to Earth as the main source of its volatile inventory. It seems likely that the Moon should exhibit a similar late veneer (Sleep 2010; Zahnle et al. 2007).

The argument against a late veneer on the Moon is based on the absence of any evidence (e.g. iron-rich material from an impactor core) in the lunar highlands, which are presumed essentially as old as the MFI. Obviously there would be some delay in forming the lunar highlands to provide for accretion of the Moon after the MFI, but this is presumed to be short given the ages of the oldest lunar rocks. However, the MFI dates from about 4.5 BYBP, while the oldest lunar rocks are at least 50 MY and perhaps more than 100 MY younger (Halliday 2008; Nemchin et al. 2009, 2012). Figure 2.1 illustrates that this is precisely the time interval during which the late veneer would have been added to Earth. If the Moon's surface remained active for 50–100 MY following the MFI, or if lunar reaccretion from the post-MFI "cloud" were delayed for a similar interval (perhaps due to tidal disruption), any late veneer material added to the Moon would have been simply mixed in before the solidification of the lunar surface, including the lunar highland rocks. Alternatively, the vaporization of impactors on the moon might have resulted in vapor clouds that would simply escape from Moon's lower gravity (Kraus et al. 2015). For Earth the surface would have stabilized rapidly, the presumed post-MFI magma ocean solidifying in about 20 million years or less (Sleep 2010). A solid surface would likely have formed even before the magma ocean had completely solidified, serving to strand late veneer materials at/near Earth's surface. In other words, a late veneer for Earth is quite consistent with lack of evidence for such on the Moon.

An additional point worth noting is the likely nature of the surface volatile inventory even if it was the direct product of the MFI. In fact, the arguments given here for conditions following somewhat smaller impactors (see Table 4.1) would also apply to the MFI, with suitable modifications to adjust for the magnitude of the event. Although MFI scenarios suggest that the core of the impactor is largely retained by Earth (Cameron 1997; Canup 2004), it would be surprising if some fraction of the metallic inventory were not vaporized during the impact (Kraus et al. 2015). If as little as 10 % of the core were vaporized (and added to vaporized ferrous iron from the silicate shell) the amount of hydrogen produced by reaction with water would be at least as much as that produced in the smaller event of a 1000 km impactor (Table 4.1). The delay in surface/atmospheric cooling associated with a 3000 km event would obviously be greater, but not likely by more than an order of magnitude or so, perhaps 50,000–100,000 years. As with the 1000 km event, even with a much higher rate of H loss from the upper atmosphere during this interval, enough hydrogen would be retained in the cooled atmosphere to result in reduced gases, not CO_2.

Another reference point for impact degassing corresponds to the ca. 100 km diameter Imbrium impactor from the Late Heavy Bombardment. The total mass of the LHB striking Earth has been estimated at about 2×10^{23} g (Gomes et al. 2005), roughly equivalent to one hundred objects 100 km in diameter, approximately the size of the Imbrium impactor. Such an impact does not have enough energy to vaporize an ocean, although it would be about enough to vaporize a “proto-ocean” if evenly applied to such a body of water. The energy of such an impact would be more localized and would produce quite high temperatures (2000+ K) in an impact cloud for a very short period of time. Atmospheric temperatures would quickly fall (days or so) to ca. 400 K.

A second degassing scenario involves processing of accreted meteoritic matieral not directly degassed by impact. As suggested above this includes fragments of larger impactors that did not vaporize as well as smaller impactors only partially ablated or slightly heated by impact processes. Smaller impactors imply milder effects, down to sizes that strike the surface without vaporizing. These latter objects simply accumulate at the surface and are incorporated in the motions of the crust, to eventually be subducted (or whatever was the Hadean equivalent).

Subduction would obviously occur in a submarine setting, with water plentiful and probably involving hydrated silicates. Heating at shallow depths would lead to hydrothermal release of gases, mostly methane and ammonia based on theoretical calculations, laboratory experiments, and direct observations of hot spring environments (Seewald 2001; Seewald et al. 1990, 1994, 2006; Cruse and Seewald 2006). Although these are based on non-subduction settings, the conditions, involving heating of organic compounds (from the meteorites) in the presence of water at relatively low temperatures, are chemically very similar. Once released to the atmosphere, photochemical processing of methane and ammonia would return higher molecular weight CNHO compounds to the oceans as rainout.

The third degassing process involves any meteoritic material that survives the shallow, low temperature hydrothermal degassing accompanying subduction. This accreted remnant will be heated in the upper mantle to magmatic temperatures (quite similar to modern subduction volcanism), producing magmas of unknown composition, but perhaps andesitic, that is, essentially like average continental crust. Because the meteorites have (on average) a substantial metal content, both the shallow hydrothermal processing, and importantly the *initial* magmatic processing, will occur with metal present. For this reason, the volatiles released during initial magmatism, should they reach the surface, will probably be strongly reducing, perhaps largely as CO and H_2 (Holland 1984). These components would then be further processed photochemically in the atmosphere, and given the other constituents of the proto-atmosphere, condensed CHNO compounds would be the result.

The essential consequence of all of these pathways for late-accreting meteoritic material is the initial production of a strongly reducing atmosphere that rapidly yields an abundance of high molecular weight CHNO compounds that end up in the ocean. In other words, the earliest proto-atmosphere closely corresponds to the Miller–Urey experiment and with similar results. Perhaps even more important, the main reservoir for carbon on the early Earth is a mix of reduced (prebiotic) compounds dissolved, suspended, and deposited as sediment in a nearly global ocean. Carbon dioxide would have been a minor reservoir for carbon, not being the main carbon product of impact degassing, or high temperature degassing (in the presence of iron, Fig. 4.1), or low temperature (essentially hydrothermal) processing. (Fig 4.2)

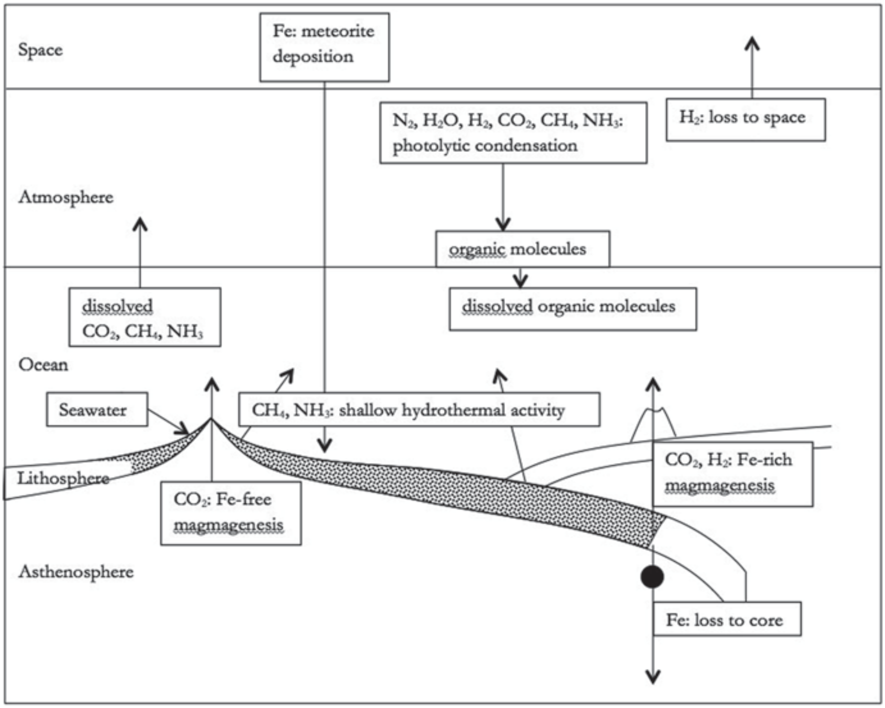


Fig. 4.2 Processing of near-surface materials during late accretion

It would not have been an especially abundant component of the atmosphere, although some atmospheric CO_2 is not unlikely from a fairly early date. This contrasts strikingly with the standard model of degassing of CO_2 (and water) dominated by magmatic processes in the upper mantle to produce a massive atmosphere mostly composed of carbon dioxide and nitrogen. The implications of this are considerable, and they resolve many discrepancies between geological, geochemical and astrophysical information and the conventional view of the early atmosphere. Of particular interest is the impact this conclusion has on the conditions relevant to the origin of life. Before pursuing this it is useful to review the basic mechanisms of life at the molecular level.

References

Ahrens TJ, O'Keefe JD, Lange MA (1989) Formation of atmospheres during accretion of the terrestrial planets. In: Atreya SK, Pollack JB, Matthews MS (eds) Origin and evolution of planetary and satellite atmospheres. University of Arizona Press, Tucson, pp 328–385

Cameron AGW (1997) The origin of the Moon and the single impact hypothesis. Icarus 126:126–137

Canup RM (2004) Dynamics of lunar formation. Annu Rev Astron Astrophys 42:441–475

Cruse AM, Seewald JS (2006) Geochemistry of low molecular weight hydrocarbons in hydrothermal fluids from Middle Valley, northern Juan de Fuca Ridge. Geochim Cosmochim Acta 70:2073–2092

Gerlach TM (1993) Thermodynamic evaluation and restoration of volcanic gas analyses: an example based on modern collection and analytical methods. Geochem J 27:305–322

Gomes R, Levison HF, Tsiganis K, Morbidelli A (2005) Origin of the cataclysmic Late Heavy Bombardment period of the terrestrial planets. Nature 435:466–469

Halliday AN (2008) A young Moon-forming giant impact at 70–110 Ma accompanied by late-staged mixing, core formation and degassing of the Earth. Philos Trans R Soc A 366:4163–4181

Holland HD (1984) The chemical evolution of the atmosphere and oceans. Princeton University Press, Princeton, p 582

Kraus RG, Root S, Lemke RW, Stein ST, Jacobsen SB, Mattsson TR (2015) Impact vaporization of planetesimal cores in the late stages of planet formation. Nature Geosci 8:269–272

Melosh HJ, Collins GS (2005) Planetary science: Meteor Crater formed by low-velocity impact. Nature 434:157

Nemchin AA, Pidgeon RT, Healy D, Grange ML, Whitehouse MJ, Vaughan J (2009) The comparative behavior of apatite-zircon U-Pb systems in Apollo 14 breccias: implications for the thermal history of the Fra Mauro Formation. Meteorit Planet Sci 44:1717–1734

Nemchin AA, Grange ML, Pidgeon RT, Meyer C (2012) Lunar zirconology. Aust J Earth Sci 59:277–290

Schaefer L, Fegley B Jr (2007) Outgassing of ordinary chondritic material and some of its implications for the chemistry of asteroids, planets and satellites. Icarus 186:462–483

Schaefer L, Fegley B Jr (2010) Volatile element chemistry during metamorphism of ordinary chondritic material. Icarus 205:483–496

Seewald JS (2001) Aqueous geochemistry of low molecular weight hydrocarbons at elevated temperatures and pressures: constraints from mineral buffered laboratory experiments. Geochim Cosmochim Acta 65:1641–1664

Seewald JS, Seyfried WE Jr, Thornton EC (1990) Organic-rich sediment alteration: an experimental and theoretical study at elevated temperatures and pressures. Appl Geochem 5:193–209

Seewald JS, Seyfried WE Jr, Shanks WC III (1994) Variations in the chemical and stable isotope composition of carbon and sulfur species during organic-rich sediment alteration: an experimental and theoretical study of hydrothermal activity at Guaymas Basin, Gulf of California. Geochim Cosmochim Acta 58:5065–5082

Seewald JS, Zolotov MY, McCollom T (2006) Experimental investigation of single carbon compounds under hydrothermal conditions. Geochim Cosmochim Acta 70:446–460

Sleep NH (2010) The Hadean–Archean environment. In: Deamer D, Szostak JW (eds) The origins of life. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp 35–48

Symonds RB, Rose WI, Bluth GJS, Gerlach TM (1994) Volcanic-gas studies: methods results and applications. Rev Mineral 30:1–66

Tian F, Toon OB, Pavlov AA, De Steck H (2005) A hydrogen-rich early Earth atmosphere. Science 308:1014–1017

Zahnle K, Sleep NH (2006) Impacts and the early evolution of life. In: Thomas PJ, Hicks RD, Chyba CF, McKay CP (eds) Comets and the origin of life, 2nd edn. Springer, Berlin, pp 175–208

Zahnle K, Arndt N, Cockell C, Halliday A, Nisbet E, Selsis E, Sleep NH (2007) Emergence of a habitable planet. Space Sci Rev 129:35–78

Zahnle K, Schaefer L, Fegley B Jr (2010) Earth's earliest atmospheres. In: Deamer D, Szostak JW (eds) The origins of life. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp 49–66

Chapter 5
The Underlying Chemical Nature of Life

The defining characteristics of life include metabolism and storage and duplication of genetic information. Metabolism involves the processing of "food" molecules to provide energy to carry out synthesis of essential biomolecules in order to grow and maintain the organism (or cell). This may include mechanisms to capture external sources of energy, in particular solar energy for photosynthesis or geochemical energy from minerals for chemolithotrophy. Storage and duplication of information is the role of polynucleotides, mostly DNA but also including RNA in some organisms. Replication of DNA provides the means of reproduction, i.e., duplicating the information to be passed to the next generation of the organism (or additional cells). The information stored in DNA is used by the organism/cell to direct the synthesis of proteins (polypeptides) and various kinds of RNA used in both protein and DNA synthesis. Messenger RNA (mRNA) is produced by "reading" a sequence of nucleotides along the DNA strand and synthesizing (using enzymes) a corresponding RNA polynucleotide. This mRNA is a smaller molecule than DNA and is released into the cytoplasm where it directs the synthesis of a protein. The protein produced will have an amino acid sequence corresponding to the nucleotide sequence "message" encoded in the mRNA (which it received from the DNA). The coding is according to triplets of nucleotides in sequence along the length of the mRNA (and DNA) molecules. Because there are four different kinds of nucleotide in RNA there are 64 possible triplets to code for the 20 amino acids used by most organisms as components of proteins. Some of the triplets are used to control the process of mRNA/protein synthesis, such as defining the "start" and "stop" locations for synthesis. Most amino acids are encoded by more than one triplet.

Translation of the mRNA message into amino acid/protein language is carried out by small (ca. 75 nucleotide) transfer RNA polynucleotides (tRNA) . These small (for RNA) polymers have a complicated geometry (Fig. 5.1), including three loops in a tightly folded pattern resembling a cloverleaf. (It actually folds into a complicated structure not really resembling a cloverleaf, as in Fig. 5.1b)

The "tail end" is where an amino acid, specific to the particular tRNA, is attached. The middle leaf of the clover, opposite the tail end, has a sequence of three nucleotides (the "anti-codon") that matches (defines) the particular amino acid

G. H. Shaw, *Earth's Early Atmosphere and Oceans, and The Origin of Life*,
SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-21972-1_5

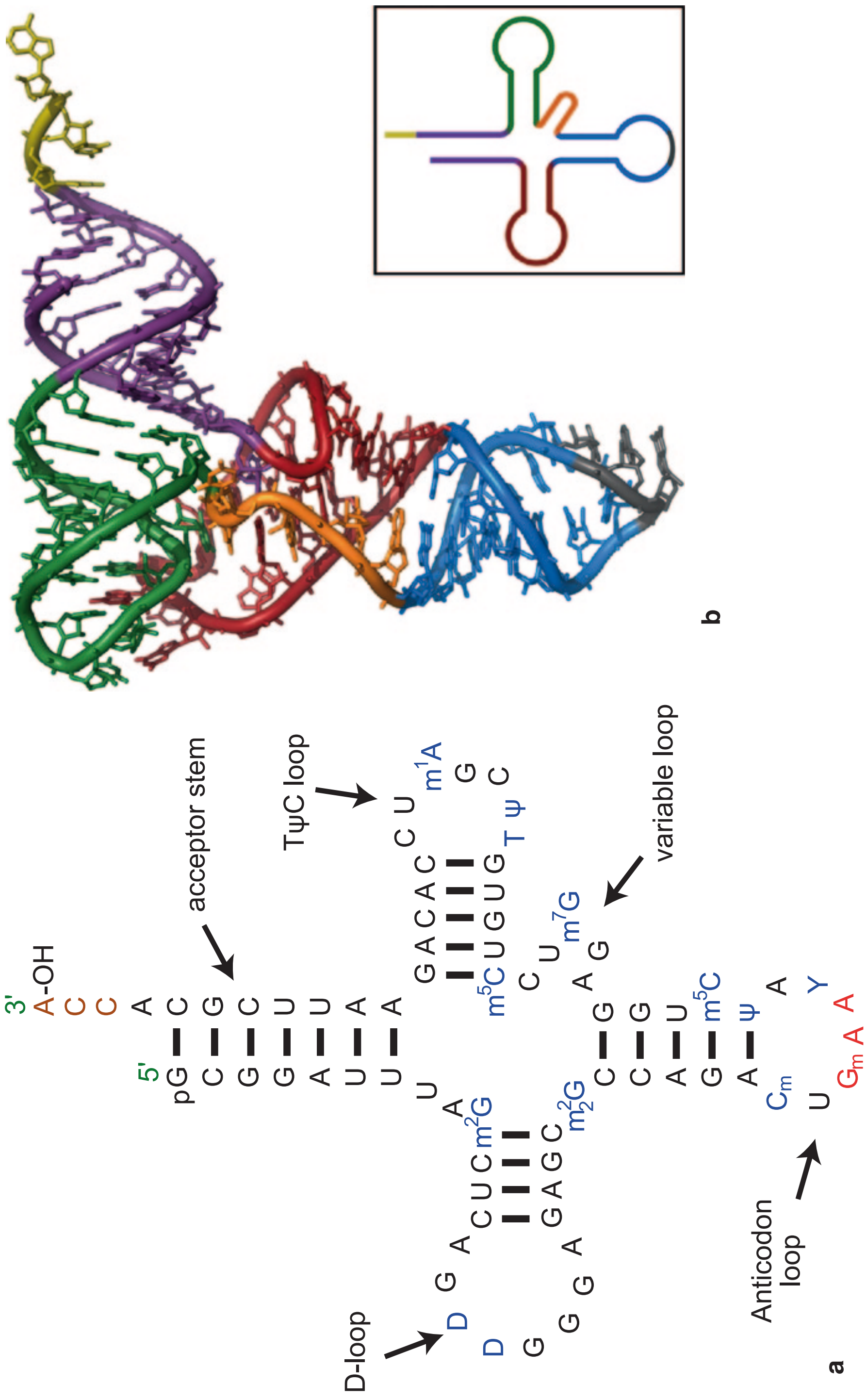


Fig. 5.1 Transfer RNA (tRNA) example. (With permission from Yikrazuul)

attached to that tRNA. These coding nucleotides will match a complementary triplet (codon) on an mRNA molecule. Various tRNA molecules will align sequentially along matching (complementary) codons on the mRNA, carrying their respective amino acids, which will then combine sequentially to make a polypeptide (protein). This process is catalyzed by another relatively small RNA molecular complex (actually an RNA-protein combination enzyme). The process proceeds step by step along the length of the mRNA until the entire amino acid sequence of the protein has been synthesized. There may be additional steps involving processing of the proteins before the final active protein is produced, but this synthesis is the key step in converting the information stored in the DNA to a useful protein. These proteins are then used by the cell to catalyse metabolic energy production and various biological syntheses needed to maintain and grow the cell. Additional RNA enzymes assist the duplication of DNA at the time of cell division/reproduction.

This very abbreviated description of the chemical processes common to essentially all living organisms, from the simplest single-celled organisms to complex multi-cellular organisms including humans, highlights the importance of a small number of key biopolymers to life. The central role of RNA in both protein and DNA synthesis has led to the suggestion that life initially began as some kind of RNA-based complex, and that DNA as a reliable information storage mechanism came later (Woese 1967; Crick 1968; Orgel 1968) . This earlier era is sometimes referred to as the “RNA world”. The role of tRNA is especially striking. It is the sole molecule *directly* connecting polynucleotides to polypeptides. In fact it has been suggested (Eigen 1992) that it is the key molecule for the emergence of life because of this central role.

It should be obvious from this description that it is highly unlikely that life processes were initially as complicated or sophisticated as what we can observe in extant organisms, and that even the simplest organism we can investigate at present has an enormous number of “moving parts”, both “large” and “small” (in molecular terms). There is a huge gap between a prebiotic soup and the simplest bacterium.

What we *can* know is something about the ancestry of all living forms, and the probable characteristics of the “common ancestor” of all life, often called the last universal common ancestor or LUCA. This is possible because the basic processes of DNA, RNA and protein synthesis depend on a small number of enzymatic molecules (or assemblies) that must work with sufficient fidelity to keep life going. This does not mean that they are perfect now, or that they were always as good as they are now, but that they cannot have changed too dramatically for several billion years of biological descent. The central molecules must have maintained a reasonable degree of integrity during the vagaries of evolutionary history in order for the biological machinery to work, and continue to work. At the same time, the variety of extant organisms clearly demonstrates that there are many possibilities for living forms.

Because of the underlying chemical conservatism, it is possible to investigate the nucleotide sequences of certain key molecules in order to relate different organisms to one another. The basic idea is that organisms that are evolutionarily “close” should have a closer correspondence in nucleotide sequences, especially in some

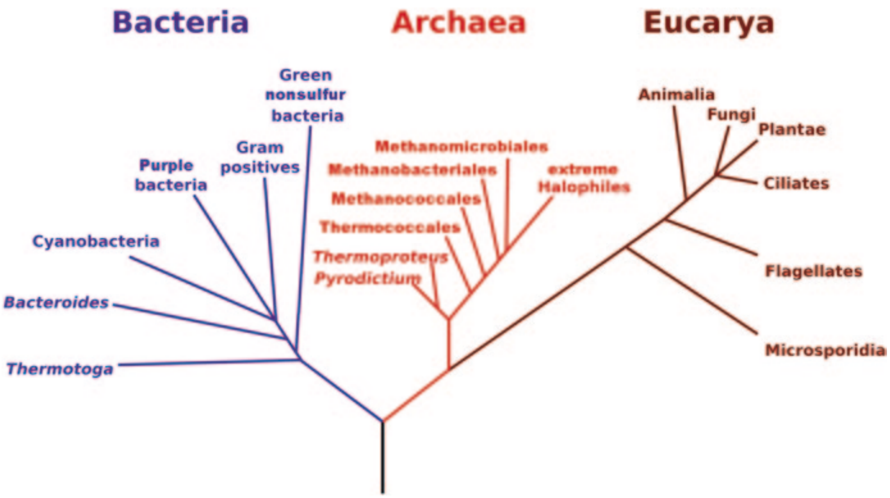


Fig. 5.2 A tree of life. (NASA Astrobiology Institute)

key RNA compounds, compared to less closely related organisms. This is why, for example, chimpanzees share a very large proportion of their DNA (as well as other polynucleotide sequences) with humans.

Detailed sequencing of a particular type of RNA called 16S RNA is one basis for constructing various "trees of life" showing the relationship distances between organisms, from the simplest bacteria to multicellular plants and animals. Figure 5.2 shows one such construction, in which genetic "distance" between groups of organisms is represented by distances along connecting lines, and relatedness is shown by branching.

One of the most interesting features of this diagram is the location of most familiar organisms, such as multicellular plants and animals, in a small corner at the upper right of the diagram. These and other so-called eukaryotic organisms (having large cell sizes, separate nuclei containing the DNA, and several types of complex intracellular structures) are all much more closely related to one another than some bacteria are to one another. More detailed phylogenetic trees focused on, say, animals (i.e. multicellular heterotrophs) , have also been constructed and show a striking correspondence between RNA relatedness and older phylogenetic trees based on macroscopic features which used both modern forms and fossil remains to derive relationships. In fact, this constitutes a powerful test of evolutionary theory, contrary to claims made about the "untestability" of evolution. The much older evolution-based trees existed long before we had any inkling of the mechanisms of genetic information storage and processing. The discovery of molecular phylogenetics, following decades of biochemical research, is a profound confirmation of the previous evolutionary predictions.

References

Crick FHC (1968) The origin of the genetic code. J Mol Biol 38:367–379
Eigen M (1992) Steps towards life—a perspective on evolution. Oxford University Press, Oxford, p 173
Orgel LE (1968) Evolution of the genetic apparatus. J Mol Biol 38:381–393
Woese C (1967) The genetic code: the molecular basis for gene expression. Harper and Row, New York, p 200

Chapter 6
The Miller-Urey Experiment and Prebiotic Chemistry

Before examining the relationship of the previous discussion to the origin of life it is first necessary to review ideas concerning the origin of prebiotic compounds. The literature of prebiotic chemistry is considerable, beginning with Oparin's (1938, 1957) hypothetical ideas and, especially, following Miller's (1953, 1955) groundbreaking experiments (Fig. 6.1, Table 6.1). A brief summary can cover the essentials. Greater detail can be found in several extended, essentially biochemical, treatments (Schopf 1983, 1992; Ponnamperuma 1972; De Duve 1995, 2005; Mason 1991; Calvin 1969; Deamer and Szostak 2010). A few points merit more detailed consideration from a geological perspective, especially as regards specific environments and certain key components. Miller's apparatus is shown in Fig. 6.1. It is comprised of two reservoirs, one containing liquid water (a model ocean) and any constituents that are water soluble, either initially or after chemical reaction. The second reservoir is filled by gases thought to resemble the early atmosphere, in which a spark discharge emulates lightning to energize a variety of gas-phase chemical reactions. Ultraviolet light has been also been used as an energy source in similar experiments. The water in the "ocean" flask is heated to drive water and gases through the system, and water is condensed from the "atmosphere" to rain out, along with water-soluble reaction products, into the "ocean". Further reactions take place between components dissolved in the water, to give a variety of organic compounds containing carbon, hydrogen, oxygen and nitrogen.

The Miller-Urey experiments involved strongly reducing atmospheres. There are many ways to produce these initial conditions in the laboratory, most of which have been tried in subsequent experiments. Whether reducing potential is provided by adding hydrogen, methane or ammonia has little effect on the ultimate results, to a first approximation. More oxidized atmospheres generally produce smaller fractions of interesting prebiotic molecules, or in some experiments essentially none (Abelson 1965; Schlesinger and Miller 1983a, b; Miller and Schlesinger 1983; Stribling and Miller 1987). These latter results (with more oxidizing, that is CO_2-rich, atmospheres) have long been recognized as a barrier to the generation of enough starting materials to make a "prebiotic" soup of significant concentration, and/or with the necessary inventory of precursors. Among the proposed solutions are addition of

G. H. Shaw, *Earth's Early Atmosphere and Oceans, and The Origin of Life*,
SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-21972-1_6

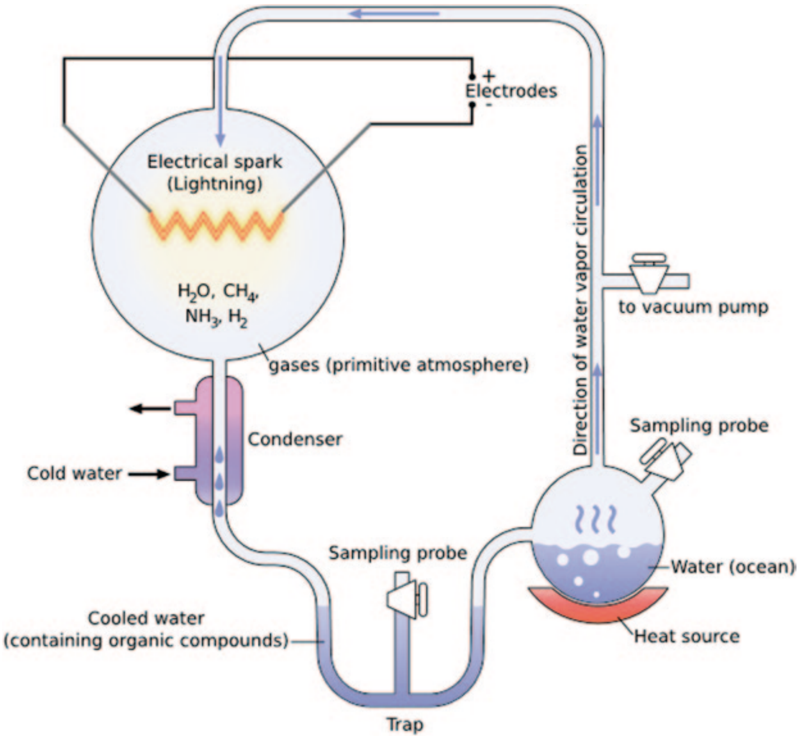


Fig. 6.1 The Miller-Urey experiment. (With permission from YassineMrabet)

Table 6.1 Results of Miller's experiment. Percentage yields of various interesting organic compounds from the original Miller experiment

Compound	% yield
Glycine	2.1
Glycolic acid	1.9
Sarcosine	0.25
Alanine	1.7
Lactic acid	1.6
N-Methylalanine	0.07
α-Amino-*n*-butyric acid	0.34
α-Aminoisobutyric acid	0.007
α-Hydroxybutyric acid	0.34
β-Alanine	0.76
Succinic acid	0.27
Aspartic acid	0.024
Glutamic acid	0.051
Iminodiacetic acid	0.37
Iminioacetic-proprionic acid	0.13
Formic acid	4.0
Acetic acid	0.51
Proprionic acid	0.66
Urea	0.034
N-Methyl urea	0.051

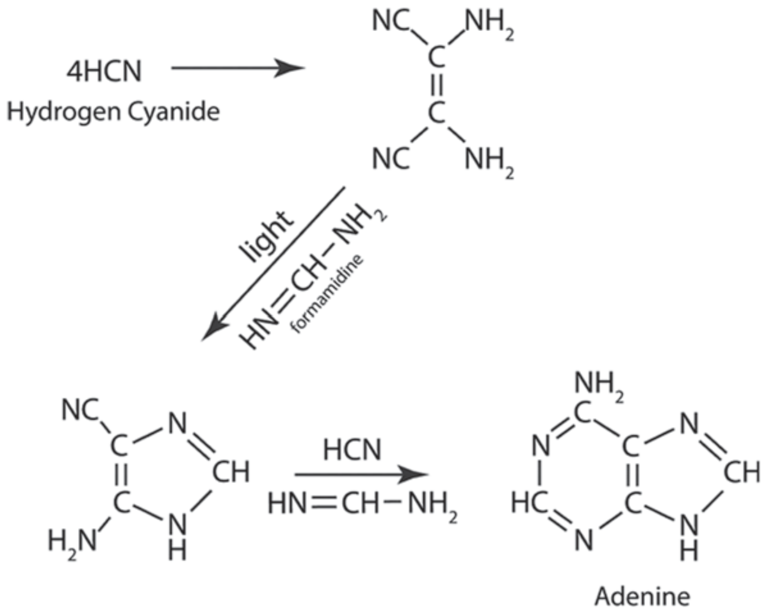


Fig. 6.2 Abiotic pathway to adenine. (After Orgel 2004b)

greater amounts of hydrogen to the atmosphere from enhanced volcanic degassing of mantle volatiles (Holland 1984) or limited hydrogen escape from the early Earth (e.g. Tian et al. 2005) or focusing prebiotic synthesis at submarine hot springs using more concentrated hydrogen in the hydrothermal fluid as a reductant/reactant (Corliss et al. 1981).

Regardless of the exact origin/source of prebiotic biogeochemical conditions, laboratory experiments have demonstrated the synthesis of a wide range of prebiotic compounds, including amino acids, sugars, lipid-like molecules, and even nucleotides (Fig. 6.2). Although not every precursor presumed necessary has been produced by these abiotic syntheses, the results are striking, not only in the broad array of products, but in the surprising presence of specific biomolecules, e.g. many of the essential amino acids. Table 6.1 provides a partial listing of resultant compounds. In few, if any, of these experiments are there significant catalytic surfaces/compounds available to promote reactions; they involve mainly gas reactions followed by hydrolytic reactions in relatively dilute solutions. In addition, there is usually a limited temperature range in these experiments. Seldom are high temperatures, say up to magmatic temperatures (which could obviously facilitate dehydration reactions), or temperatures well below freezing, present in these trials. Some hypothetical reactions have been suggested under these more extreme conditions as possible sources of specific precursors (Orgel 2004a). Considering the potential catalytic activity of various silicates (clays have been proposed as very important (Ferris 2006)), it would not be surprising if "messier" experiments were found to yield many, if not all of the important precursors yet to be synthesized. A recent example of a route to pyrimidine ribonucleotides shows that synthesis of some biological precursors can be accomplished by using more complex mixtures of several components (Powner et al. 2009; Sutherland 2010). Experiments involving clays or other silicates are obviously much more difficult to carry out with the required degree of purification of starting materials necessary to avoid initial bio-contamination of the experiment.

If it is demonstrable that various biological precursors can be produced by simple and reasonably expected natural processes, this is still a long way from the synthesis of life. At the very least there must be a reaction scheme(s) which can yield more complex compounds, especially high molecular weight entities similar to biopolymers, including polypeptides, nucleosides and poly-nucleosides, before we can conceive of a self-sustaining system similar to life. Even with such complex compounds there would remain the problem of a coordinated system of chemical species that can successfully carry out metabolism and reproduction, probably the most essential characteristics of living forms.

Virtually all biopolymerisation reactions involve dehydration. The apparent necessity for aqueous environments to life processes, and the sheer volume of surface water on Earth indicates some difficulty with the generation of biopolymers by dehydration, while at the same time requiring the presence of liquid water as a medium for most biochemical reactions once life is established. Highly dilute solutions, even if they contain all the necessary building blocks, will probably not lead to either "pre-biopolymers" or living material. Several ideas have been proposed to overcome this apparent barrier to the emergence of life.

In the classic example Darwin showed awareness of this problem with his introduction of the "warm pond" model as a site for the origin of life (Darwin 1871, letter to Hooker reproduced in Calvin 1969). He was well aware that duplicating (or looking for) such an environment at present would probably be unsuccessful, because the presence of extant life would result in destruction (consumption, actually) of the necessary nutrients long before de novo life could appear. Indeed, this problem is one of the reasons prebiotic experiments require an extraordinary degree of cleanliness to avoid false positive results. The usual model for the "warm pond" invokes collection and concentration of prebiotic molecules in an isolated, typically small, water body that achieves concentration by continual or periodic drying, with continued addition of sufficient nutrient-laden water to ultimately achieve high organic compound concentrations. This need not occur in any particular type of geologic environment, but a coastal (maybe tide-pool) setting is perhaps the most likely. One could also imagine flood-plain lakes or ponds as additional possibilities. On the Hadean Earth there would have been serious limitations because of the initial absence of continental areas, though it was probably not long before either basaltic plateaus (like Iceland) and/or island-arc settings (like Japan or Philippines) provided considerable lengths of coastal/tidal environments.

Very old zircons (ca. 4.2–4.4 BYBP) are the best evidence for the presence of early oceans and continental lithologies (Wilde et al. 2001). This earliest continental material was probably in the form of island arcs (or something very like them). The much higher rate of mantle thermal convection necessary to dissipate early higher radiogenic heating (and residual accretion energy) implies either faster convective motions or a greater length of ocean ridges and subduction zones, or both (as suggested in Chap. 3). My inclination is toward the latter, which would imply a greater number of island arcs, and therefore a large number of relatively small continental-type crustal fragments. The total length of coastline could easily have been far greater than at present once even a relatively small amount of continental crust had been produced. Although the history of generation of continental crust is somewhat problematical, it seems likely that there was a rapid initial increase in volume (and area) of continental crust (e.g. Armstrong and Harmon 1981; Reymer and Schubert 1984), such that within a few tens of million years after the surface was stabilized and late accretion/degassing of volatiles "filled" the oceans, a multitude of island arcs dotted the globe. "Continental" topography would have resembled a large collection of "Japans, Indonesias and Aleutians" rather than "Americas, Africas or Asias". Average slopes of these island-arc-like land masses would have been greater than at present, which imply faster average areal rates of erosion, especially from the fragmental materials of the arc volcanos. Even with a smaller area of exposed "continent", average global erosion rates might have been comparable to the present. Coastal environments would have included areas of rapid sedimentation, probably with numerous small deltaic settings, and an abundance of tidal zones with isolated pools. Presumably tidal ranges would have been much greater because of the more proximal Moon. Whether these pools were sufficiently isolated and stable to lead to high concentrations of prebiotic organics is speculative at best, but this is probably not an unreasonable assumption, based on what can be observed in modern coastal settings.

In fact, the presence of isolated pools may not be as important for concentrating organics as for providing "test reactors" and "reservoirs" for more highly processed prebiotic materials. If the previous arguments for pervasive reducing conditions during degassing are correct, the production of reduced CHNO compounds from a high degree of degassing of accreted late veneer, and early generation of an ocean approximately the same volume as at present (by the end of accretion and degassing), leads to interesting conclusions about the organic content of that early ocean. As an extreme case one can calculate the amount of organic material that would be found in the early ocean if virtually the entire present inventory of near-surface carbon (most of which is now carbonate) was present in the ocean as reduced organic compounds. Using the data of Table 2.3 and assuming an average organic composition as CH_2O implies a concentration of more than 10 % *in the entire ocean*. Adding nitrogen to the mix would increase this somewhat, but even the 10 % approaches the concentration of organic compounds in protoplasm. This is very simplistic and there are numerous complicating factors one might incorporate to get a "better" number, but it is clear that not much, if any, additional concentration would be necessary to have a "primordial soup".

If the Hadean ocean had such high concentrations of organic compounds it would seem to make the idea of origin of life at submarine hot springs a more tenable idea. Previous suggestions of prebiotic organic synthesis at deep-sea hot springs typically call for such synthesis in hydrogen-rich hydrothermal systems. The observation that chemolithotrophy is very ancient and to a large extent found in hot-spring settings is given as support for this view. There are two major problems with this proposal. First, chemolithotrophy as such obviously cannot occur until life is present. As primitive as they are, chemolithotrophs are highly sophisticated organisms. That they occupy a niche that *may* have been suitable for biogenesis does not necessarily mean they are especially close to the root of biogenesis, or that initial biogenesis would have occurred under conditions closely matching their modern environment (that is, deep-sea). Second, biogenesis requires more than a process for generating precursor compounds, as suggested above. It is also necessary to produce rather high concentrations of these compounds (or at least this seems very likely) *with appropriate conditions for polymerization* (dehydration) reactions. Deep-sea hydrothermal systems may be suitable for prebiotic synthesis, but they are intrinsically dispersive. The heat driven mixing of sea water serves to dilute and distribute any compounds produced throughout the larger ocean. The lack of high concentrations and/or a dehydrating mechanism make polymerizations very difficult. On the other hand, if one accepts the idea expressed above, that virtually all carbon in the Hadean ocean was in the form of reduced organic compounds, it may be that high enough concentrations would be present in the entire ocean, and in particular in deep-sea hydrothermal systems, to accomplish the necessary high degree of synthesis. The intrinsically dispersive nature of such hydrothermal systems and lack of a polymerization/dehydration mechanism probably still constitutes a major barrier to emergence of life in this setting, however. Miller and Bada (1988) examined this issue in some detail not long after the proposal of submarine hot springs as the site of life's origin. Although they were skeptical about several aspects of prebiogenesis at hot springs, and persuasively so to some, the idea of this mode of origin of life continues to receive considerable attention.

It is far more probable that life began in a setting more like that originally envisioned by Darwin. Such "pools" were undoubtedly very abundant in tidal and wave-splash zones during the Hadean. They would have constituted an enormous number of prebiological "test reactors" with high concentrations of prebiotic molecules such as amino acids, nucleotides, etc. In fact, given the high starting concentrations likely in the ocean itself, concentrations produced by periodic drying of these pools should have been even higher. In addition to the drying, the presence of solar radiation (including more intense ultraviolet) was an additional energy source for driving chemical reactions. Add to this the availability of lava flows (think Iceland or Hawaii or Aleutian Islands) for both high temperature reactions and drying, and it is not difficult to imagine an enormous number of prebiotic experiments taking place simultaneously around the active Hadean Earth. These surface environments would also have contained many complex silicate weathering products available for catalysis. Taken together it seems much more likely that the conditions favorable to the initiation of life fit better with (volcanic) coastal zone settings than the deep sea. The presence of primitive chemolithotrophs (and other thermophiles) in modern deep-sea hydrothermal systems is much more likely to be an artifact, in the sense that changed surface conditions, especially oxygenation, drove these organisms into more isolated and protected deep-sea settings where they have survived.

There is yet another reason to suggest coastal pools as the site of the emergence of life. Although drying of an organic-rich solution in a "pond" is one possible way to synthesize "pre-biopolymers", it has also been suggested that polyphosphates may have played a crucial role in these reactions (e.g. Yamagata et al. 1991), by acting as a dehydrating agent. In addition, some of the most critical compounds for life (and presumably necessary before life could emerge) are mono and polyphosphates of nucleotides such as AMP, ADP and ATP, respectively the mono-, di- and triphosphates of adenosine. Synthesis of these compounds presumably requires the presence (especially under pre-biotic conditions) of rather high concentrations of phosphate. Phosphorus is an important trace element in igneous rocks, probably the most abundant rock type present in Hadean crust. Both oceanic mafic (basaltic) and continental felsic (granitic in the broadest sense) rocks contain about 800 ppm phosphorus (Table 2.3). Submarine weathering and/or hydrothermal activity released some phosphorus into sea water, where it would have been diluted. "Continent" (or island arc) weathering would have fed phosphorus into coastal waters where it might well have been concentrated in tidal zones, in a manner similar to concentration of the organic compounds. There would have been some losses into the wider ocean, but continuous delivery by streams from the "island arc" interiors should have resulted in buildup of phosphate along shores. This situation in some ways compares to modern coastal settings in which phosphate delivered by streams helps "feed" biological productivity along ocean shores. In the absence of life, however, the phosphate would more likely accumulate to much higher concentrations (as dissolved phosphate) than are found in modern environments, in which delivered phosphate is rapidly consumed by organisms. A recent alternative view (Jones et al. 2015) suggests a severe shortage of phosphorus in the early oceans due to sorption of phosphorus onto iron oxides. Their reasoning applies to a phosphorus cycle similar to the present, where much oceanic phosphorus is brought to the surface by upwelling of deep ocean water. Biologic production of ferric oxides from the ferrous-iron rich water would likely have the effect they suggest. However, this mechanism would not necessarily apply to phosphorus delivered to coastal waters by erosion of emergent land masses, and in any case would not apply to pre-biotic conditions.

In sum, the combination of high concentrations of biotic precursor compounds and phosphate in myriad coastal pools, accompanied by drying, solar radiation, and heating (both moderate and very high temperatures) is a recipe for production of polymers that could approach the complexity from which life could emerge. The critical step from these "pools" to life seems to require a great leap beyond chemical processing, but this may be less of a leap than it appears.

References

Abelson PH (1965) Abiogenic synthesis in the Martian environment. Proc Natl Acad Sci U S A 54:1490–1494

Armstrong RL, Harmon RS (1981) Radiogenic isotopes: the case for crustal recycling on a near-steady-state no-continental-growth Earth. Philoso Trans R Soc Lond A 301:443–472

Calvin M (1969) Chemical evolution: molecular evolution towards the origin of living systems on the Earth and elsewhere. Oxford University Press, Oxford, 278 pp

Corliss JB, Baross JA, Hoffman SE (1981) A hypothesis concerning the relationship between submarine hot springs and the origin of life on Earth. Oceanologica Acta No Sp, 59–69

De Duve C (1995) Vital dust. Basic Books, New York, 362 pp

De Duve C (2005) Singularities: landmarks on the pathways of life. Cambridge University Press, Cambridge, 258 pp

Deamer D, Shostak JW (eds) (2010) The origins of life. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 318 pp

Ferris JP (2006) Montmorillonite-catalysed formation of RNA oligomers: the possible role of catalysis in the origins of life. Philos Trans R Soc Lond B Biol Sci 361:1777–1786

Holland HD (1984) The chemical evolution of the atmosphere and oceans. Princeton University Press, Princeton, 582 pp

Jones C, Nomosatryo S, Crowe SA, Bjerrum CJ, Canfield DE (2015) Iron oxides, divalent cations, silica, and the early earth phosphorus crisis. Geology 43:135–138

Mason SF (1991) Chemical evolution. Clarendon Press, Oxford, 317 pp

Miller SL (1953) Production of amino acids under possible primitive Earth conditions. Science 117:528–529

Miller SL (1955) Production of some organic compounds under possible primitive Earth conditions. J Am Chem Soc 77:2351–2361

Miller SL, Bada JL (1988) Submarine hot springs and the origin of life. Nature 334:609–611

Oparin AI (1938) The origin of life. Dover, New York, 495 pp

Oparin AL (1957) The origin of life on Earth. Oliver and Boyd, Edinburgh, 495 pp

Orgel LE (2004a) Prebiotic chemistry and the origin of the RNA world. Crit Rev Biochem Mol Biol 39:99–123

Orgel LE (2004b) Prebiotic adenine revisited: eutectics and photochemistry. Orig Life Evol Biosph 34:361–369

Ponnamperuna C (1972) The origins of life. E.P. Dutton, New York, 215 pp

Powner MW, Gerland B, Sutherland JD (2009) Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. Nature 459:239–242

Reymer A, Schubert G (1984) Phanerozoic addition rates to the continental crust and crustal growth. Tectonics 3:63–77

Schlesinger G, Miller SL (1983a) Prebiotic synthesis in atmospheres containing CH_4, CO, and CO_2. I. Amino acids. J Mol Evol 19:376–382

Schlesinger G, Miller SL (1983b) Prebiotic synthesis in atmospheres containing CH_4, CO, and CO_2. II. Hydrogen cyanide, formaldehyde and ammonia. J Mol Evol 19:383–390

Schopf JW (ed) (1983) Earth's earliest biosphere. Princeton University Press, Princeton, 543 pp

Schopf JW (ed) (1992) Major events in the history of life. Jones and Bartlett, Sudbury, 190 pp

Stribling R, Miller SL (1987) Energy yields for hydrogen cyanide and formaldehyde syntheses: the HCN and amino acid concentrations in the primitive ocean. Orig Life Evol Biosph 17:261–273

Sutherland JD (2010) Ribonucleotides. In: Deamer D, Szostak JW (eds) The origins of life. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp 109–122

Tian F, Toon OB, Pavlov AA, De Steck H (2005) A hydrogen-rich early Earth atmosphere. Science 308:1014–1017

Wilde SA, Valley JW, Peck WH, Graham CM (2001) Evidence from detrital zircons for the existence of continental crust and oceans on the Earth 4.4 Gyr ago. Nature 409:175–178

Yamagata Y, Watanabe H, Saitoh M, Namba T (1991) Volcanic production of polyphosphates and its relevance to prebiotic evolution. Nature 352:516–519

Chapter 7

A Pathway to the First Organism

Eigen (1992) pointed out the importance of a "system containing a template-instructed replication cycle over which a catalytic feedback loop is superposed", calling such a system a hypercycle. In a sense all living organisms constitute hypercycles, generally of extraordinary complexity. The simplest of them are probably the RNA viruses (Eigen 1992). A difficult hurdle in the progress from abiotic organochemistry and a living organism is the assembly of the first hypercycle. Current living systems use an integrated polynucleotide-protein system involving several classes of compounds including DNA, tRNA, mRNA, rRNA, and various enzymatic proteins, as described above. The importance of a cellular (or protocellular) structure to keep these units in close enough communication to carry out successful reproduction has been clearly recognized (Eigen 1992; De Duve 1995, 2002). Probably the most impressive approach to simplifying the problem is the postulate of an RNA world, in which RNA serves as both an information carrier and an enzymatic agent (Woese 1967; Crick 1968; Orgel 1968), an idea strongly supported by the proven catalytic activity and/or self-directed assembly of some RNAs (e.g., Orgel 1968).

Even an RNA world presents difficulties, however, especially when considering the subsequent transition to the polynucleotide-protein world of the present: how do we get from self-propagating (and mutating) RNA to the more complex system with two distinct chemical families (polynucleotides and proteins) closely interlinked? It is clear that tRNA plays a critical role, but is there a compelling reason, or probable mechanism for this transition to take place? A path that considers the full potential of tRNA is worthy of consideration. The importance of tRNA as a crucial link between polynucleotides and polypeptides was recognized by Eigen when he proposed the UrGen, or original gene, as being an ancestral tRNA (Eigen 1992). His intent seems to have been to highlight tRNA as the first coherent carrier of genetic information capable of reproduction and evolutionary change.

The earliest proteins may have been very simple, limited in both size and catalytic capacity (De Duve 2002). At the same time there may have been "general" polypeptide catalysts capable of affecting more than one kind of reaction. De Duve has suggested some of these simple polypeptides be called "multimers" to distinguish them from true (modern) enzymatic proteins. He proposed lengths of about

G. H. Shaw, *Earth's Early Atmosphere and Oceans, and The Origin of Life*,
SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-21972-1_7

20 peptide units, which is likely about the smallest length that would exhibit the geometric properties necessary for catalysis. Not coincidentally, De Duve points to the 75 bases of typical tRNA as being capable of coding for just such polypeptides. Surprisingly, he did not further investigate the possible polypeptides that tRNA might encode.

It seems likely that molecular evolution over the last several billion years would conceal, or at least obscure, the coding that may have been carried out by tRNA, both because the tRNA has been subject to some degree of evolutionary change (albeit limited by the necessity to retain critical functionality), and especially by the gradual displacement of primitive proteins by more sophisticated, longer polypeptides. Modern protein and polynucleotide databases provide an abundance of sequencing data for a variety of organisms that allows us to do more than construct phylogenies, it allows us to construct consensus sequences for tRNAs that *might* closely match those of the last universal common ancestor. To be sure, this is complicated by the possibility that early translation systems may have encompassed a smaller amino acid set, with later additions requiring development of tRNA that selectively attach to the "newer" amino acids (e.g., de Duve 2005). On the other hand, it is likely that newer tRNAs were recruited from an existing pool. In any case, it is possible to generate consensus sequences for each of the extant tRNAs. It is, of course necessary to use caution when dealing with the extra loop region of several tRNA (as shown in Fig. 5.1), which can lead to tRNA of more than one hundred nucleotides, but the database construction has largely taken care of that problem.

Sequences from the NCBI database (http://www.ncbi.nlm.nih.gov/nucleotide/) of more than 75,000 tRNA sequences were sorted by coding amino acid and tRNA length, and grouped into sets of up to 2000 for consensus alignment using ClustalW (now ClustalOmega; http://www.ebi.ac.uk/Tools/msa/clustalo/). Because of the program limit of 2000 sequences, the data from most of the tRNAs can be used to produce more than one possible consensus sequence. These batches were grouped in size ranges, excluding some of the very shortest and longest tRNA sequences. The vast majority of consensus alignments were for average lengths between about 70 and 80 nucleotides, except for tRNAs that typically possess a significant "variable loop" (especially leucine, methionine, serine, and tyrosine). Table 7.1 gives several consensus sequences of tRNAs for each amino acid. No attempt was made to sort the database sequences into similar groups, but to simply take the data in batches of 2000 for consensus analysis.

Notice that almost all of the longer sequences have the usual CCA "tail", where the amino acid attaches. Because this common sequence comes at one end of the tRNA it has only a minor effect on any amino acid sequence calculated from the tRNA, and that is at one end of the polypeptide. We should not expect this to have a large effect on subsequent comparisons of the polypeptides "coded" by the tRNA.

Once consensus sequences have been determined it is straightforward to translate the nucleotide string into potential polypeptides. ExPASy (http://www.expasy.org) provides a program for this, which yields results for reading the tRNA string in either the forward or reverse direction, and with each of the three possible offsets (i.e. different start points for translation in each direction) to give 6 possible amino

Table 7.1 Consensus sequences for tRNA

Alanine
GGGGATGTAGCTCAAGTGGTAGAGCGCATGCTTTGCATGTATGAGGTCCTGGGTTCAATCCCCGGCATCTCCA
GGGGAATTAGCTCAGATGGTAGAGCGCTCGCTTAGCATGCGAGAGGTCGGGGGTTCGATGCCCCCATTCTCCA
Arginine
GGCCCGGTGGCCCAATGGATAAGGCGTCTGACTCCGGATCAGAAGATTGTAGGTTCGATTCCTGCCTGGGTCG
GGCCCAGTGGCCTAATGGATAAGGCGTCTGACTACGAATCAGAAGATTGTGGGTTCGAGTCCCACCGGGGTCG
GGCCCCCTTAGCTCAGCTGGATAGAGCACCTGCCTACGAAGCAGGAGGTCGGGGGTTCGAATCCTGCCGGGCGCGCCA
Asparagine
GTCTCTGTGGCGCAATCGGTTAGCGCGTTCGGCTGTTAACTGAAAGGTTGGTGGTTCAAGCCCACCCAGGGACG
GTCTCTGTGGCGCAATTGGTTAGAGCGTTCGGCTGTTAACCGAAAGGTCGGTGGTTCGAGCCCACCCAGGGGCG
Aspartic Acid
TCCTCGGTAGTATAGTGGTGAGTATCCCCGCCTGTCACGCGGGAGATCCCGGGTTCGATTCCCCGCCGGGGAGG
GGAGCGGTAGTTCAGTTGGTTAGAATACCGGCCTGTCACGCCGGGGGTCGCGGGTTCGAGTCCCGTCCGCTCCGCCA
Cystine
TGGCGGTATAGCTCAGTGGTAGAGCATTTGACTGCAGATCAAGAGGTCCCAGGTTCAATTCCTGGTGCCGCCT
GGGGGGGATAGCTAAGTGGTAAAGCATCGGACTGCAAATCCGGAGGACCCCGGTTCGAATCCGGGTGCCGCCTCCA
Glutamine
GGTTCCATGGTGTAATGGTTAGCACTCTGGACTCTGAATCCAGCGATCCGAGGTTCAAATCTCGGTGGGACCT
Glutamic Acid
TCCCTGGTGGTCTAGTGGTTAGGATTCGGGGCTTTCACCGCCGCGGCCCGGGTTCGATTCCCGGTCAGGGAA
TCCCTGTGTGGTCTAGTGGTTAGGACTCGGCGCTTTCACTGCCGCGGCCCGGGGTTCAATCCCTGGTCAGGGAA
GTCCCCTTCGTCTAGAGGCCTAGGACACCGCCCTTTCACGGCGGTGAACAGGGGTTCGAATCCCCTAGGGGACACCA
Glycine
GCATTGGTGGTTCAGTGGTAGAATTCTCGCCTGCCACGCGGGAGACCCGGGTCCGATTCCCGGCCAATGCA
GCATTGGTGGTTCAAGTGGTAGAATTCTCGCCTTCCACTGCGGGAGACCCGGGTTCGATTCCCGGCCAATGCA
TCCCTGGTGGTCCAGTGGTTAGGACTCTGGCCTCCCAAGGCAGGGGGCCCGGGTTCGAATCCCTGCTCAGGCAA
TCCCTGGTGGTCCAGTGGCTAAGACTCTGCGCTCCCAATGCAGGGGGCCCAGGTTCGATCCCTGGTCAGGGAA
TCCCGGGTTGTGCTTAGTTGGTAGAACTCTACCCTGCCAATGCACAAGGTGGGGGTCGCGGGTTCGATTCCCGTTTCCCGCTCCA

Table 7.1 (continued)

Alanine
Histidine
GGCCGTGATCGTATAGTGGTTAGTACTCTGCGTTGTGGCCGCAGCGAACCCCGGTTCGAATCCGGGTCACGGCACCA
Isoleucine
GGCCGGTTAGCTCAGTTGGTTAGAGCGTGGTGCTAATAACGCCAAGGTCGCGGGTTCGATCCCCGCTAGGGGCAA
GGGTCTGTAGCTCAGGTGGTTAGAGCGCACCCCTGATAAGGGTGAGGTCGGTGGTTCAAGTCCACCCAGACCCACCA
Leucine
GTCAAGATGGCCCAAGTGGGTAA[AT] GGCATAAGACTTAAAACCTAAAAGTCAGGGGTTCGAATCCTCTTCGTGGCACCA
GGGTAGCGTGGCCGAGCGGTCTAAGGCGCTGGATTAAGGCTTCCAGTCTCTTCGGGGGCGTGGGTTCGAATCCCACCGCTGCCA
GTGCGGATGTGGGGGAATTGGCTAAGGCGCTGGATTCAGGTTCCAGTGTCCCCAGGGAGGGTGGGGGTTCGAATCCCACTTCTGGCACCA
GCCGGGGTGGCGGAATTGGTAGACGCGCTGGATTTAGGATCCAGTGGCCGTAAAAAGGCCGTGCGGGTTCGAGTCCCGCCCTCGGCACCA
GTCAGGATGGCCGAGTGGTCTAAGGCGCCAGACTCAAGGTGATTCAATCTTCTCAGGTAATGGGGATTCTGGTCTCCGGATGGGAGGCGTGGGTTCGAATCCCACTTCTGACA
Lysine
GCCTGGCTGGCTCAGTCGGTAGAGCATGTGACTCTTGATCTCAGGGTCGTGGGTTCGAGCCCCATGTTGGGTG
GCCCGGCTAGCTCAGTCGGTAGAGCATGAGACTCTTAATCTCAGGGTCGTGGGTTCGAGCCCCACGTTGGGCG
GGGTCGTTAGCTCGCTCAGTTGGTAGAGCAGCTGACTTTTAATCAGTTGGTCGAAGGGTTCGAATCCCTCACGGCCCACCA
Methionine
GAGCAGATTGGCGCAGTGGGAAGCGCGCTGGGCTCATAACCCGGAGGTCGGTGGGTTCGAAACCCACCCCTGGCATA
GGCGGGGTGGCGCAGTTGGGTAGAGCACAGGACTCATAAGCTCGTCGGGCTCATAACCCGAAGGTCGCAGGTTCAAATCCCGCCCCCGCCACCA
Phenylalanine
GCCGAGATAGCTCAGTTGGGAGAGCGTTAGACTGAAGATCTAAAGGTCCCATCCTCGTGTCGCTGGTTCGATCCCGGGTCTCGGCACCA
Proline
GGCTGGTTGGTCTAGTGGTATGATTCTCGCTTTGGGTGCGAGAGGTCCCGGGTTCAAATCCCGGACGAGCCC
CGGGGGGTTGGCGCAGCCTGGTAGCGCACCTGGTTTGGGACCAGGGGGTCGCAGGTTCGAATCCTGTCACCCCGACCA
Serine
GTAGTGGTGGCCGAGTGGTTAAGGCGATGGACTAGAAATCCATTGGGCTCTGCCCGCGCAGGTTCGAATCCTGCCGACGACG
GGAGAGGTGGCCGAGTGGTTGAAGGCGCCGGCCTGGAAAGCCGGTATGGGGGTAAAAACCCCTATCGAGGGTTCGAATCCCTCCCTCTCCGCCA
Threonine
GCCCCCTTAGGTCGAAGTGGTAGAGCGCCTGTCTAGTAAACAGGAGGTCGTGGGTTCGATTCCCGCTGGGGGCT

Table 7.1 (continued)

Alanine
GGCCCTGTTGCTTCAGTTGGTAGAGCACCTGACTTGTAATCAGGAGGTCGGGGGTTCGAATCCCAGCAGCGGCACCA
Tryptophan
GGCTGCTTGGTCCAATGGTAGAGCGTCGGACTCCAAATCAGAAGACCGAAGGTTGCGGGTTCGATTCCCGCCGGGGTCACCA
Tyrosine
CCTTCGATAGCTCAGTTGGTAGAGCGGAGGACTGTAGTAAATCTGTGGGCATCCTTAGCGTCGCTGGTTCGAATCCGGCTCGAAGGACCA
Valine
GTTTCCGTAGTGTAGTGGTTATCACGTTTGCCTAACACGCGAAAGGTCCCCGGTTCGAACCCGGGCGGAAACA
GGGGGTTTTGGTCAAGTTGGTTAGAGCATCTGCCTTACACGGAGGAGGTCGCCGGTTCGAACCCGGTAGCGCCCACCA

acid sequences for each "primal" tRNA. This approach is necessary because there is no certain way of knowing (especially in an ancient RNA world) exactly how the tRNA might have been "read", and this method covers all possibilities. A significant number of these consensus tRNA include a STOP codon (one of three particular nucleotide triads—UAA, UAG or UGA), which in modern organisms terminate protein synthesis at that point. It is impossible to know if this was the case in the most primitive translation system, but since we cannot know what these codons might otherwise have coded for at that time, and as it seems just possible that something about those triads stops attachment of tRNA anticodons, all such translated amino acid sequences are excluded from further analysis. This reduces the total number of polypeptides for analysis from 270 to 101. Only two of the consensus tRNA strings contain STOP codons in all six of the possible sequences. In other words, almost all tRNA produce at least one possible small polypeptide.

Comparisons of the resulting short amino acid sequences with a protein database using BLASTp (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) tests whether extant proteins might contain "residues" in their sequences retained from these potential archaic "multimers" (to use de Duve's term). Interestingly, there are numerous hits, in which the short amino acid sequences match segments of modern proteins to a statistically significant degree. Table 7.2 shows two examples. The E value in the table expresses the likelihood that the match is the result of chance. The smaller the E value the more significant the match: "an E-value of 1E-4 or below seems to be commonly acceptable in the literature for diagnosing a "good" match" (Palumbi Lab, http://sfg.stanford.edu/BLAST.html). There seems to be no particular pattern to the matches, and few striking matches with centrally important enzymes. Although this may be thought somewhat disappointing, it is not really surprising considering the enormous increase in sophistication of proteins that must have taken place over several eons.

The examples in Table 7.2 have been chosen because they clearly show very good "hits" for matching the amino acid sequences defined by proto-tRNA with the

Table 7.2 BLAST characteristics of two selected "matches" between tRNA encoded polypeptides and proteins in the BLAST database

Description	E value
Example 1—from an arginine "proto" tRNA (This is from the third sequence under arginine tRNA in Table 7.1)	
Conserved hypothetical protein [Clostridium ultunense Esp] >emb|CCQ93955.1| conserved hypothetical protein [Clostridium ultunense Esp]	4.00E-08
Arginine-tRNA ligase (EC 6.1.1.19)—Mycobacterium leprae >gb|AAA63099.1| argS [Mycobacterium leprae]	7.00E-08
Predicted protein [Populus trichocarpa]	2.00E-07
Hypothetical protein PI23P_07915 [Polaribacter irgensii 23-P] >gb|EAR12535.1| hypothetical protein PI23P_07915 [Polaribacter irgensii 23-P]	2.00E-07
Hypothetical protein HRM2_17400 [Desulfobacterium autotrophicum HRM2]	4.00E-07
Hypothetical protein RSPPHO_00750 [Rhodospirillum photometricum DSM 122] >emb|CCG07376.1| Putative uncharacterized protein [Rhodospirillum photometricum DSM 122]	5.00E-07
Aldehyde dehydrogenase [Xanthomonas oryzae pv. oryzae KACC 10331]	9.00E-06
Hypothetical protein XAC0392 [Xanthomonas axonopodis pv. citri str. 306] >gb|AAM35283.1| hypothetical protein XAC0392 [Xanthomonas axonopodis pv. citri str. 306]	1.00E-05
Serine/threonine protein kinase [Rhodospirillum photometricum DSM 122] >emb|CCG09408.1| Serine/threonine protein kinase [Rhodospirillum photometricum DSM 122]	4.00E-05
Hypothetical protein SI859A1_00334 [Aurantimonas manganoxydans SI85-9A1] >gb|EAS49676.1| hypothetical protein SI859A1_00334 [Aurantimonas manganoxydans SI85-9A1]	2.00E-04
Example 2—from an aspartic acid "proto" tRNA. (This is from the second sequence under aspartic acid in Table 7.1)	
Hypothetical protein KT99_11413 [Shewanella benthica KT99] >gb|EDP99056.1| hypothetical protein KT99_11413 [Shewanella benthica KT99]	4.00E-17
Hypothetical protein PBPRB1578 [Photobacterium profundum SS9] >emb|CAG23440.1| hypothetical protein PBPRB1578 [Photobacterium profundum SS9]	5.00E-17
Putative membrane protein [Pseudoalteromonas haloplanktis TAC125]	2.00E-14
Hypothetical protein EC55989_0204 [Escherichia coli 55989] >emb|CAU96084.1| conserved hypothetical protein [Escherichia coli 55989]	6.00E-13
Hypothetical protein DAMO_2559 [Candidatus Methylomirabilis oxyfera] >emb|CBE69632.1| protein of unknown function [Candidatus Methylomirabilis oxyfera]	1.00E-08
Hypothetical protein BBFL7_01180 [Flavobacteria bacterium BBFL7] >gb|EAS20298.1| hypothetical protein BBFL7_01180 [Flavobacteria bacterium BBFL7]	1.00E-07
Protease prtH [Bacteroides dorei CL02T12C06] >gb|EIY32994.1| protease prtH [Bacteroides dorei CL02T12C06]	1.00E-05

Table 7.2 (continued)

Description	E value
Hypothetical protein PPOP_0243 [Paenibacillus popilliae ATCC 14706] >dbj|GAC40908.1| hypothetical protein PPOP_0243 [Paenibacillus popilliae ATCC 14706]	3.00E-04
Putative membrane protein [Agromonas oligotrophica S58] >dbj|BAM89399.1| putative membrane protein [Agromonas oligotrophica S58]	3.00E-04
Hypothetical protein DORLON_02056 [Dorea longicatena DSM 13814] >gb|EDM62655.1| hypothetical protein DORLON_02056 [Dorea longicatena DSM 13814]	5.00E-04

protein database. In many cases there was no good match, but that is not surprising considering that some (perhaps most) translations would not be expected to produce amino acid sequences of biological significance. The first example shows an especially interesting match (the second on the list), between a short sequence product from translation of an arginine tRNA and an arginine tRNA ligase. Modern tRNA ligases catalyze the attachment of the particular amino acid to the appropriate (in this case arginine) tRNA. This may be a striking example of the positive feedback one would hope to see in a tRNA world. Whether it has any real significance is perhaps debatable. The second example shows the best match obtained from this set of comparisons, with a very high likelihood that the AA sequence is a match for (some part of) an extant protein. These results (and the total set of results) clearly suggest the real possibility that ancient tRNA may have coded for useful small protein-like molecules. There are two caveats, however. First, many of the matches observed are for “hypothetical”, “predicted”, or “putative” proteins, presumably possible proteins for which specific functionality is not yet known. It is not clear how many of these are really significant in terms of our main question. On the other hand, there are a significant number of good matches for well characterized proteins, which is encouraging. The second point concerns the possibility that matches are simply the result of chance, due to the rather short length of tRNA. Perhaps randomly generated 75 unit polynucleotides would give comparable results.

In order to test this idea 25 random 75-nucleotide “RNAs” were created and processed in the same way as the consensus tRNAs. The results of the BLASTp analysis are shown in Fig. 7.1, compared to those from the tRNAs. For these purposes data from only the first five best matches from the BLASTp analyses were included. The figure shows the negative log of the E value in a histogram. It is clear from the figure that random RNA sequences do not produce significant matches, but form a rather tight group in the vicinity of E values of 1 (+/− a factor of 10). However, although most tRNA results also show a major grouping in the same range of E values, there is a prominent tail to much lower values of E. Keep in mind that this is a log scale.

The inescapable conclusion is that at least some primitive tRNAs probably encoded for some kind of useful small protein-like products, which may well have constituted the necessary positive feedback for self-sustaining regeneration, that

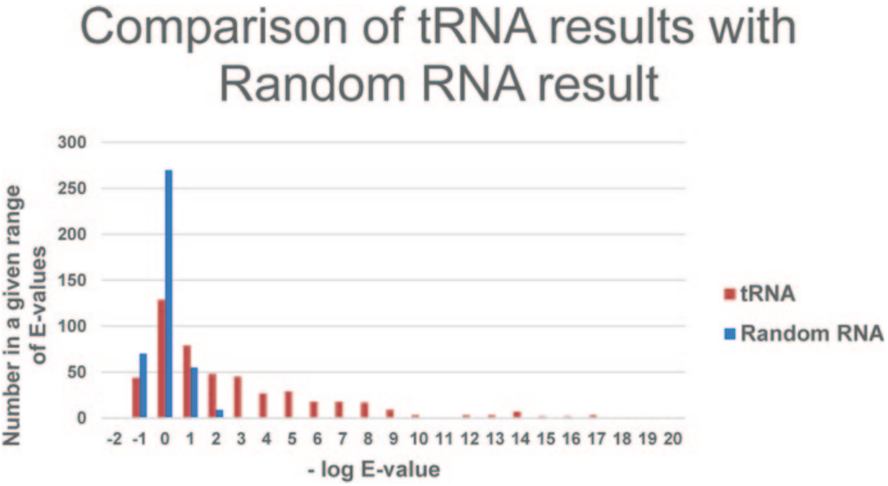


Fig. 7.1 Comparison between consensus tRNA coded and randomly generated "tRNA like" coded amino acid strings

is, life. It is certainly not a wild, outside chance that an Eigen hypercycle based on tRNA(s) and their products formed the first "living entity". The significance of these results lies in an implicit "short-cut" between an abiotic world and the RNA world (presumably an mRNA world), eventually leading to the more complicated system of life as we know it. Initially, no separate mRNA was required to support "life". Rather, tRNA was also a kind of primitive mRNA. This is a significant simplification, one which puts the genetic material and the translation machinery in the same molecules, with production of catalytic molecules occurring at the same site. There are other issues to consider, such as packaging (isolating) such entities in protocellular modules. Consideration of the role that could be played by micellular or protocellular or liposomal assemblies further supports the basic idea of an organic chemical rich early surface environment, as one might expect such structures to form in abundance from such a "soup" (next chapter).

References

Crick FHC (1968) The origin of the genetic code. J Mol Biol 38:367–379
De Duve C (1995) Vital dust. Basic Books, New York, 362 p
De Duve C (2002) Life evolving. Oxford University Press, Oxford, 360 p
De Duve C (2005) Singularities: landmarks on the pathways of life. Cambridge University Press, Cambridge, 258 p
Eigen M (1992) Steps towards life—A perspective on evolution. Oxford University Press, Oxford, 173 p
Orgel LE (1968) Evolution of the genetic apparatus. J Mol Biol 38:381–393
Woese C (1967) The genetic code: the molecular basis for gene expression. Harper and Row, New York, 200 p

Chapter 8
An Organic Carbon Rich Surface and the Source of Primitive tRNA on Earth

Assuming that a collection of tRNA-like polymers emerged as the first living entity (along with various monomers of nucleotides and amino acids and energy-rich compounds in a “primordial soup”) there remains the question of how and why these tRNA-like polymers accumulated and whether we should expect a mix of them that would have the particular nucleotide sequences that lead to life. The answer may lie in the peculiar shape of tRNA. Why is tRNA shaped like that? Of course the current shape clearly relates to tRNA functionality, perhaps especially in the ability of tRNAs to fit together and line up as they attach to mRNA in order to bring their amino acid cargo close enough to polymerize. But this must be a *consequence* rather than a *cause*. The *cause* we are looking for must relate to properties of tRNA that are relevant in a pre-biotic world. Simply stated, tRNA has its structure because the tRNA nucleotide sequences lead to internal base-pairing resulting in the folding pattern we see. This would, of course, be true for any RNA polymer of sufficient size to allow folding. tRNA is one of the smallest polymers for which the complicated folding we observe is possible. The folding helps to “cover up” potential hydration sites, resulting in greater chemical stability in the aqueous environment. One additional factor that might enter into ur-tRNA stability (and maybe account for the CCA tail?) is the possibility that some nucleotide sequences at/near the tail may be especially susceptible to attachment of amino acids, which could conceivably also increase the chemical stability. tRNAs with these chemical attributes would be more likely to accumulate in a prebiotic soup.

This does not quite get us to a “tRNA world”. Even if we assume that chemical conditions were sufficient for producing long RNA polymers (de Duve 2005), the amount of material required to sample the chemical space for even 75-unit RNAs would exhaust the available carbon inventory before all possibilities were tried. For example, if we assume that an average nucleoside (nucleotide + ribose) has, say, a dozen carbon atoms we can easily calculate how much carbon it would take to make all possible 75 unit RNA (similar to de Duve 2005)

> At each nucleoside position in the RNA there are four possible nucleosides, making 4^{75} possible RNAs total. This is equal to about 10^{45} nucleoside monomers to make every possible polymer, which means about 10^{21} moles of nucleoside, equivalent to about 10^{24}–10^{25} grams

G. H. Shaw, *Earth's Early Atmosphere and Oceans, and The Origin of Life*,
SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-21972-1_8

> of carbon. If all of Earth's surface carbon inventory was in the form of organic molecules, that would amount to less than 10^{23} grams of carbon available. It is obvious that much of the available carbon would be taken up by numerous compounds other than our putative RNA, thus it would be most difficult to manufacture every possible 75-unit RNA, even over a long period of time, to say nothing of having just the right mix of tRNA-like polymers in order to start life.

Even worse, if all such possible RNA were adrift in a primordial soup how could we ever expect an *assemblage* of functional tRNA (i.e. capable of mediating polypeptide synthesis) to arise from this mix. A partial answer may lie in the presumed chemical stability of the folded polynucleotides (and polypeptides). The folded structure effectively shields at least some, perhaps most, potential hydrolysis sites from the surrounding chemical medium. A non-folded polynucleotide has many such sites exposed and would probably have a limited lifetime before being hydrolytically broken into smaller pieces, maybe eventually back to monomers. In other words, the RNA polymers that would accumulate over time would be those that resist chemical attack by the hydrous medium. It is certainly essential that polymerization can take place in order to produce polynucleotides of significant size, but it is also clear that depolymerization (hydrolysis) is inherent in the aqueous medium. The "warm, wet pond" would certainly experience periodic dilutions with new water, addition of new chemical species, and drying/concentration as meteorologic and tidal conditions change, even on a daily basis. In such a dynamic regime the polymers most likely to survive will be those with intrinsic chemical stability relative to those changes (de Duve 2005). Those "most fit" will achieve higher concentrations. Moreover, if tRNA is the form in which 75-unit polynucleotides accumulate because of (relative) chemical stability, the previous calculation must be modified.

> In this case, base pairing limits the choices available to the polymer at sites where pairing occurs. One member of the pair could be any of four nucleosides, but the other will be fixed by requisite pairing. This reduces the number of possibilities depending on how many pairs form to give structure/shape to the tRNA. An examination of Fig. 6.1 shows 21 pairs. This reduces the number of possibilities to 4^{54} rather than 4^{75} and the total amount of carbon required to make every possible "tRNA" to somewhat more than 10^{33} nucleoside monomers, or about 10^{10} moles. This requires a much smaller amount of carbon, a bit more than 10^{12} grams, or about 1/100 billionth of the surface carbon inventory.

Although this idea is highly speculative, it may well be testable, at least to a point. Synthesis of random 75-nucleotide RNAs is possible. In a reasonable quantity of material, perhaps 100 mg, there would be 10^{15}–10^{16} different polymers. Subjecting this sample to progressively more hydrolytic conditions, and testing for "survivors" using standard amplification methods, it would theoretically be possible to determine if the surviving polymers resembled tRNA rather than more poorly structured strings. Such an experiment has yet to be performed, but it would be straightforward with current technology. One additional point concerns the necessity for attaching an amino acid to the tRNA tail, without which polypeptide synthesis would not occur. The nature of the selectivity of a tRNA for a specific amino is certainly speculative under prebiotic conditions. In living organisms the attachment process, and specificity, is mediated by enzymes (or ribozymes), which was not likely the case under prebiotic conditions. The prebiotic (and perhaps modern) specificity may have been achieved through details of the nucleotide sequence in the primitive tRNA (de Duve 2005), which assumes additional constraints on the sequences of primitive (but active) tRNA. On the other hand, attachment of an amino acid may further enhance the chemical stability of the tRNA by "tying up" a chemical site on the tRNA tail. This additional stabilization would seem to favor the survival of potentially functional tRNA.

There remains the question of whether we can get localized proto-tRNA assemblages (that is with enough different types of functional tRNA to code for proteinoids) together in a small package at the same time. For this we can imagine yet another thought experiment. The primordial soup would contain many other components, including lipids capable of forming what are known as micelles or liposomes (Fig. 8.1).

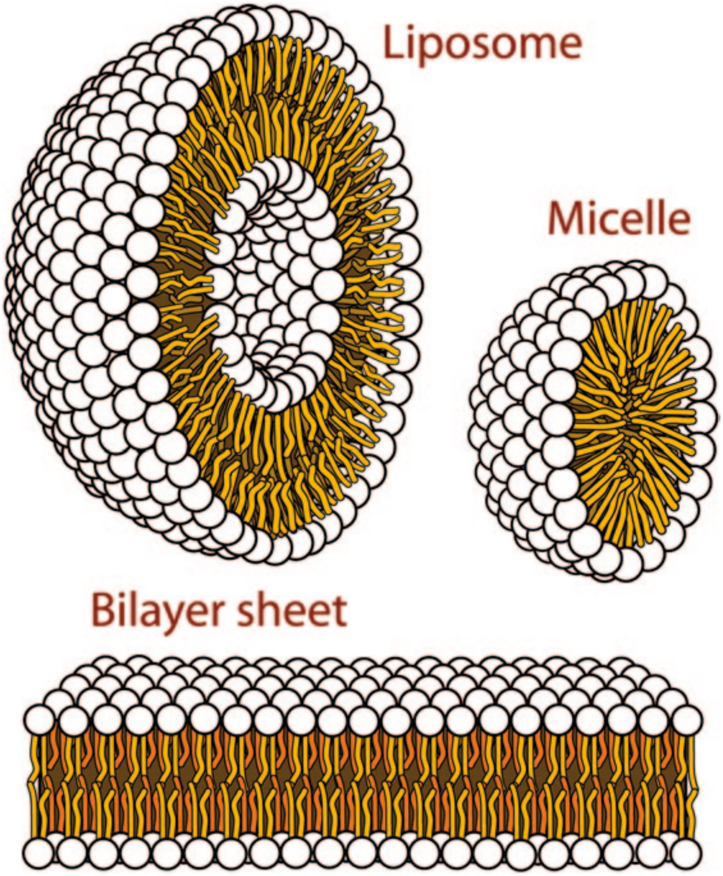


Fig. 8.1 Liposomes, micelles and lipid bilayers. (From Mariana Ruiz 2648 Villarreal)

These are bacterium-sized cell-like blobs formed from a lipid layer (or possibly bi-layer as in liposomes) that enclose a droplet of organic, possibly hydrophobic, material. The liposome and its cargo can float in a hydrous medium (the water of the pond). It is easy to make such structures in the laboratory from materials also likely to be synthesized in a prebiotic environment with organic compounds available (Chen and Walde 2010) If these were the size of a typical coccal bacterium (about 1 μ), there could be as many as 10^{18} in a small pool of water. These small "bags" filled with organic compounds could include a sampling of our putative proto-tRNA. Planet-wide the numbers of such entities could be truly enormous. If half the present continental area was present at about 4 BYBP (see discussion above concerning the evolution of Earth's continents and oceans), which may well have been the case, the length of coastline might have been as much as 2.5×10^{20} m. This is assuming a coastline/area ratio similar to the Philippines or Japan, two classic island-arc terranes. If each hundred meters of coastline harbored just one small tidal pond the total number of "bags" of organic stuff in all of the pools could have approached 10^{36} (though the actual number would probably have been limited by the availability of suitable starting materials). Just as important, the liposomes would have been in a constant state of change—breaking apart and reforming, merging and splitting. The number of "experiments" in which different combinations of internal ingredients get tried out is huge in terms of the space available along early coastlines (assuming again the tide-pool model for the warm, wet pond). Given the rearrangements that take place over time the chance that one combination would arise that had the ability to support a primitive metabolism *and* reproduce its contents (and thus itself) does not strike me as farfetched at all. Remember, it only has to happen once to cross the gap between prebiotic goo and (very) primitive life.

How the organic stuff was distributed between various organic compounds, including interesting polymers like proto-tRNA, is unknowable, but there is certainly scope for having an enormous array of possible "test reactors" and there could have been ample raw material to fill them with reactants, including our proto-tRNA. Modern bacteria contain more than 100,000 tRNA molecules/cell (Jakubowski and Goldman 1984). If we assume a much smaller, but possibly still functional number for a truly primitive organism (say 1000/liposome) we would need 10^{39} t-RNA molecules to "fill" all of the bags. This would be a bit more than 10^{15} mol or about 10^{18} g of carbon equivalent. If this were the case about 10 ppm of the total surface carbon inventory would be in the form of proto-tRNA. Although this seems an unlikely amount, there is no way of evaluating the possibility, as far as I can see. Notice that this calculation implies the presence of multiple copies of each possible folded tRNA polymer. These numbers are only useful in defining some limits on the possible, and are not meant to be descriptive of the actual situation. Indeed there is no reason to believe that every possible tRNA was ever produced, or that such would have been necessary. There are many (thousands, millions?, or more) *possible* functional tRNA sequences for each extant amino acid translator, and the assignment of an original tRNA sequence for each is a matter of trying to identify the most primitive version we can. The specific actual ur-tRNAs were probably the result of the accident in which the first functional assemblage formed, and there can be little doubt that a repetition of the "experiment" would result in quite different "original" sequences.

A very large assumption being made here is that prebiotic processes can form RNA polymers of this length, and those polymers can also be copied abiotically. It is worthwhile noting that polymerization of RNA from monomers and self-replication of small RNA units has received considerable attention from experimentalists (Bowler et al. 2013; Adamala and Szostak 2013; Adamala et al. 2014), including those looking for self-replication of RNA in extant organisms. The presence of an RNA template, say a small RNA polymer such as tRNA, may itself catalyze its own duplication by serving as a template for the assembly of a complementary strand from monomers in the surrounding medium. Once formed, the complementary strand might then serve as a template for the construction of a new strand with the original sequence. Experimental work demonstrates at least the possibility for this to happen, if not a high probability. The initial assumption that polymerization of nucleoside monomers must have been possible also suggests that these template-directed duplications are also likely, and means that generation (and regeneration) of a significant pool of proto-tRNA is reasonable.

Finally, the crucial step for life to emerge may have been the bringing together not only of a sufficient number of different potentially active tRNA to accomplish translation, but of a mix which included one or more tRNAs that coded for an active polypeptide that could act as a catalyst and form the first hypercycle. Once this happened, the positive feedback could have led to the rapid propagation of these packages—the first life.

The precise nature of this early life is highly speculative. That it resembled bacteria in some way is almost certain, but it must have been far less sophisticated than any known bacterium. It is also possible that the first "organism" was actually a close collection of proto-cellular entities that shared proto-enzymatic capabilities with one another. Such a "colonial" assemblage may or may not have been bound together in some way, but simply being in proximity in a small pond would give sufficient advantage that the "group" would proliferate more effectively. At some point the various separate enzymatic units could have coalesced into a single "body" which would more efficiently maintain the necessary biochemical machinery. This would be, in essence, identifiable as the first "bacterium". If the early surface was similar to what I have suggested above the first organism was probably heterotrophic, using the abundant supply of prebiotic compounds already available in the early ocean. This enormous food supply would have allowed nearly unchecked growth for a while, and would have provided the fuel for countless generations to rapidly evolve more sophisticated subsystems in aid of reproduction and metabolism. The available resources would have been far greater than those available to chemotrophs or lithotrophs, which would also have been restricted to limited locations (e.g. hot springs), and such organisms as we know of are, in any case, already rather sophisticated. They probably developed later, perhaps fairly quickly, around those additional energy and "food" sources.

References

Adamala K, Szostak JW (2013) Nonenzymatic template-directed RNA synthesis inside model protocells. Science 342:1098–1100

Adamala K, Anella F, Wieczorek R, Stano P, Chiarabelli C, Luisi PL (2014) Open questions in origin of life: experimental studies on the origin of nucleic acids and proteins with specific and functional sequences by a chemical synthetic biology approach. Comput Struct Biotechnol J 9:e201402004

Bowler FR, Chan CKW, Duffy CD, Gerland B, Islam S, Powner MW, Sutherland JD, Xu J (2013) Prebiotically plausible oligoribonucleotide ligation facilitated by chemoselective acetylation. Nat Chem 5:383–389

Chen IA, Walde P (2010) From self-assembled vesicles to protocells. In: Deamer D, Szostak JW (eds) The origins of life. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp 179–192

De Duve C (2005) Singularities: landmarks on the pathways of life. Cambridge University Press, Cambridge, p 258

Jakubowski H, Goldman E (1984) Quantities of individual aminoacyl-tRNA families and their turnover in Escherichia coli. J Bacteriol 158(3):769–776

Chapter 9
Evolution of the Atmosphere During the Hadean and Archean

Up to this point I have made arguments for an early surface environment in which the vast majority of the carbon is in the form of organic compounds, distributed in several forms: floating as an immiscible layer on the ocean, dissolved in sea water, trapped as an organic component in ocean-bottom sediment, and in terrestrial ponds or along their shorelines, perhaps even coating some emergent dry land. I base these conclusions on an analysis of the processes that most likely gave rise to the earliest atmosphere and how that atmosphere would have responded to surface conditions of sunlight, electrical discharge and various surface reactions. It is also strongly suggested by the apparent necessity of having organic-rich, and probably voluminous, reservoirs as a condition for the emergence of life from abiotic conditions. The modern atmosphere and the modern carbon reservoirs are dramatically different from the scene I have constructed. It is reasonable, and necessary, to ask how that early surface evolved into the present state. Fundamentally this is about oxidation. The presence of free oxygen in our atmosphere is a vital, but actually a quantitatively small part of that problem. The enormous amount of oxidized carbon in carbonate rocks is many orders of magnitude greater than atmospheric oxygen (see Table 2.3), but if it started out as reduced (organic) carbon, virtually all of it must have been produced by oxidative processes. This chapter discusses the geochemical, and eventually biogeochemical, processing of the initial reduced carbon inventory to achieve this oxidation and ultimately the production of free oxygen in the atmosphere (Fig. 9.1).

It is obvious that oxidation of the surface carbon reservoir in essence involves hydrogen loss to space (Zahnle and Catling 2014). Hydrogen production begins immediately by photochemical processing of the methane and ammonia that are generated and regenerated through hydrothermal release from organics in contact with the widespread volcano-hydrothermal systems on the early Earth. Hydrogen is also produced by reaction of subducted organics with water during partial melting in the upper mantle. These systems must have been more extensive and/or more active than the modern ocean-ridge/subduction system in order to accommodate the higher heat production from radioactive decay and from residual impact heating. Radioactive decay alone implies at least a factor of four in heat release. Photolysis

G. H. Shaw, *Earth's Early Atmosphere and Oceans, and The Origin of Life,*
SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-21972-1_9

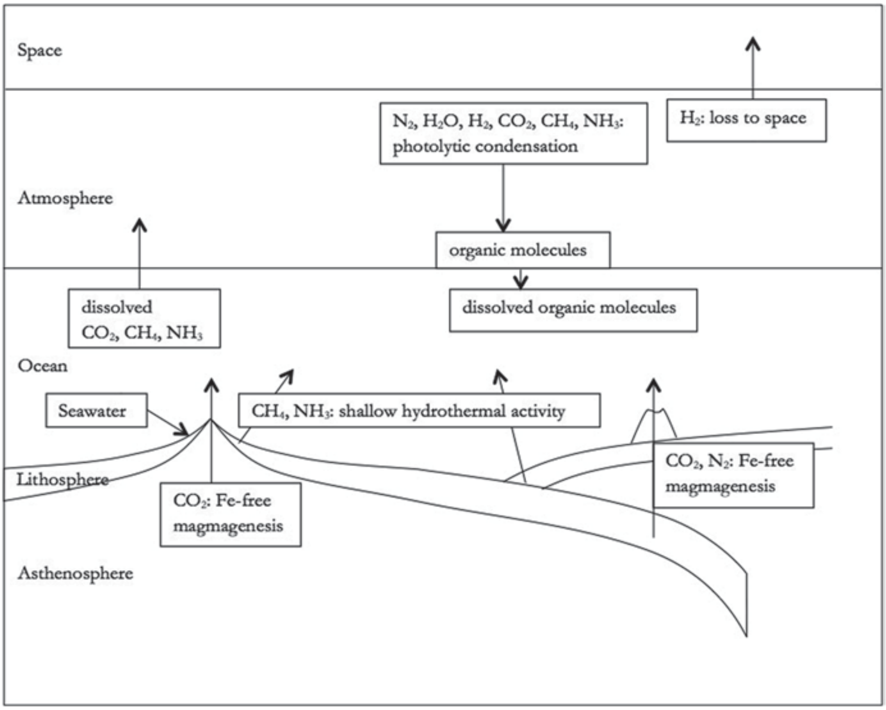


Fig. 9.1 Carbon cycling following establishment of early atmosphere. Note the production by magmatic processes of CO_2 in the mantle at high temperatures. Hydrogen is also produced during this process and is released to the atmosphere and eventually to space

of methane and ammonia would produce hydrogen which is more stable in a reducing atmosphere than it would be in an atmosphere dominated by N_2 and CO_2. In addition, the methane and ammonia are themselves “vehicles” for moving hydrogen to higher levels in the atmosphere where hydrogen produced by photolysis can readily escape Earth’s gravity field. Hydrogen produced hydrothermally and magmatically and released into lower levels of the atmosphere can readily convect to the upper atmosphere for escape. Quantifying this process is rather difficult but hydrogen loss from the upper atmosphere must have been many orders of magnitude greater than the present rate (about 2×10^8 molecules/cm^2-s (Walker 1977)) in order to account for the requisite surface oxidation. If we assume that all carbonate was produced by this hydrogen loss we can calculate a minimum rate necessary to reach the present ratio of carbonate to organic carbon for the entire carbon inventory by the end of the Archean. I choose that timing because that is roughly the point at which significant free atmospheric oxygen is thought to have become established.

The calculation is simple, just provide enough hydrogen loss from water to give the oxygen (two for each carbon atom) to turn reduced carbon to carbon dioxide (or ultimately carbonate). It actually requires about twice this amount of hydrogen loss because you also have to account for the reductants, mostly hydrogen, attached to the original carbon. If we cast the result in terms of rate loss of H_2 in molecules/cm^2-s from the upper atmosphere we get a value of about 5×10^{10} molecules H_2/cm^2-s. The current rate of hydrogen loss from the upper atmosphere is about 2×10^8 molecules/cm^2-s (Walker 1977). This means that the escape rate needed to oxidize the original carbon to carbonate is at least 250 times the current loss rate. In fact, additional hydrogen loss is also necessary to oxidize minerals produced during weathering of rocks to sulfate and to oxidized iron compounds. This is somewhat more difficult to quantify, but is probably not much less than an amount equal to that needed to oxidize the carbon reservoir and probably not ten times as much. If we take the required hydrogen loss as four times that to convert the carbon, we would get a hydrogen loss rate 1000 times the present-day value, or about 2×10^{11} H_2/cm^2-s. There is no compelling reason that this value should necessarily be reduced following the emergence of the earliest life so long as the atmosphere remained generally reducing. After photosynthesis evolved the situation might have changed, but not immediately, as will be discussed in a later chapter.

Even if we consider the estimates of early hydrogen loss to be this high we have yet another question to answer: Was there a mechanism for generating enough hydrogen to maintain high enough atmospheric levels of hydrogen, as (mostly) H_2, H_2O, CH_4 and NH_3, to offset the much higher loss of hydrogen to space? There are three sources of the hydrogen compounds that can provide the hydrogen: release of hydrogen as volcanic gases, generation of methane and ammonia through hydrothermal activity, and serpentinisation of ultramafic rocks. These may not be entirely separable because of the overwhelming importance of sea-floor volcanism, which includes hydrogen release from the mantle, hydrothermal processing of surface material that takes place as sea water circulates through ridge-crest systems, and concurrent serpentisation. Most of the hydrogen was probably volcanic in origin, formed during magmagenesis in the upper mantle. Under the iron-free conditions prevailing after core formation and primary degassing, any carbon present in these magmas would react with water to give CO_2 and H_2 (perhaps some CO). The carbon itself (and water) arrives in the source region of the magma by subduction of surface sediments containing organics deposited on the bottom of the primordial ocean (discussed further in Chap. 11). This recycling of surface reduced carbon provides a continuous source of carbon compounds for reprocessing in the upper mantle, which is in essence where the carbon dioxide is generated. The removal of hydrogen to the surface at volcanic zones, and ultimately to space, is what preserves the carbon dioxide.

Estimates of modern hydrogen production from combined volcanic and hydrothermal systems have been made through field studies, giving a figure of 2×10^{10} H_2/cm^2-s. Note that this number is already far higher than modern hydrogen loss rates from the upper atmosphere. The apparent disparity is due to hydrogen consumption in the modern atmosphere, the vast majority of which is quickly oxidized by oxygen. Such would not have been the case prior to about 2.2 BYBP when there was little or no free oxygen in the atmosphere. The modern hydrogen production is only about 10 % of that necessary to offset the Hadean/Archean loss rate calculated above, but we have to make an adjustment for those earlier conditions. At least in the late Hadean and early Archean volcano-hydrothermal systems must have been considerably more active, at least four to five times as active as at present due to higher radiogenic heat production, and probably significantly more than that when one includes residual impact heat. This means that hydrogen production would probably have been five times higher than at present, or about half the necessary rate.

We must add in the other source of atmospheric hydrogen in the form of other reduced gases, especially methane and ammonia, which deliver hydrogen to the atmosphere. Modern hydrothermal systems are known to generate both methane and ammonia when hydrothermal fluids interact with organic components in sediments or possibly with organic compounds carried into hydrothermal systems as components in sea water (Table 9.1). This process would have been quantitatively more important during the Hadean/Archean because the reduced surface conditions imply significant dissolved organics in sea water at that time, and higher organic content in submarine sediments. We can estimate the potential amounts by assuming a rate of circulation of sea water through submarine hydrothermal systems and likely concentrations of organics in the sea water. Modern circulation through deep-sea systems has been estimated as about 2.5×10^{12} m^3/year (Sleep and Zahnle 2001). If we guess that Hadean/Archean sea water may have contained 1 % dissolved organics (probably a substantial underestimate) and if we further assume that only 1 % of the organics (probably an underestimate) get converted to methane we get a rate of methane production of about 6×10^{10} molecules/cm^2-s, which is equivalent to twice as much H_2, and is nearly half the additional hydrogen equivalent needed to offset the requisite upper atmosphere H_2 loss. An approximately equal amount of ammonia might have also been produced, which brings us up to enough to account for all of the hydrogen loss. Because of the higher heat flow at the time, hydrothermal circulation was probably a factor of 4–5 higher than the modern value, suggesting that there was easily enough generation of CH_4 and NH_3 to make up for H_2 loss to space. This is a very crude calculation, but is almost certainly on the low side.

There is one final point before considering other effects, and the impact of life itself, on the evolution of the atmosphere. During the Archean (and later, of course) radioactive decay is a heat source which wanes through time with the loss of the original radiogenic isotopes. Heat production at the end of the Archean, 2.5 BYBP, had dropped to less than half of the original amount, implying lower rates of tectonism, volcanism and hydrothermal activity. This would lead to a gradual reduction in the production of H_2, CH_4 and NH_3 over time, and would probably lead to a reduction in atmospheric hydrogen (and equivalents) and probably a lower rate of loss of H_2 to space. This suggests that production of these reductants (and H_2 loss to space) was probably greater than the average calculated above during the early Archean to compensate for lower production at the end of the Archean. In addition, as the conversion of reduced carbon to oxidized carbon proceeds there will be less reduced carbon available for processing in the mantle and in hydrothermal systems. There are other factors which complicate this simple picture, as will be evident in the discussions to follow. First, however, it is necessary to address an issue that has been a puzzle for more than 40 years, the problem of how liquid water at Earth's surface was possible in the Hadean/Archean when the level of solar influx to Earth was up to 25 +% less than at present.

Table 9.1 Ammonia and methane production by hydrothermal activity

Laboratory results (see reference for details)				
Run/locality	NH_3 conc. (mmol/kg)	CH_4 conc. in fluid (mmol/kg in fluid)	Time (hours)	T (C)
SED-SW	<2		0	325
	50.4	30	98	325
	52.7	40	241	325
	55.3	46	241	325
(Change T,P)			414	400
	Not given	105	533	400
	88.5	33	1161	400
SED-ESW	<2	<0.1	0	325
	37.5	11	49	325
	40.5	25	288	325
	47.9	29	1751	325
	53.1	46	3672	325
(Change T,P)			3884	375
	60.3	80	4175	375
	69.5	120	6073	375
Field measurements				
Locality	NH_4 (mmol/kg in fluid)	CH_4 (mmol/kg in fluid)		
Dead Dog Field				
H'ken Hollow	3.18	22.6		
Chowder Hill	3.16	20.3		
Insp. Mounds	2.87	21.6		
Puppy Dog	2.99	18.5		
ODP Mound Field				
1035F	1.26	2.99		
Shiner Bock	2.34	7.07		
Spire	2.27	6.71		
1035H	2.1	5.85		

References

Sleep NH, Zahnle K (2001) Carbon dioxide cycling and implications for climate on ancient Earth. J Geophys Res 106:1373–1399

Walker JCG (1977) Evolution of the atmosphere. Macmillan, New York, 318 p

Zahnle K, Catling D (2014) Waiting for O_2. In: Shaw GH (ed) Earth's early atmosphere and surface environment. Geol Soc Am Spec Pap 504:37–48

Chapter 10
The Faint Young Sun Problem

Astrophysical modeling and observations of stars reveal much about the way in which the vast majority of stars evolve through time. The typical star gets its energy from the fusion of hydrogen to helium. Because hydrogen is by far the most abundant element in the universe most stars are composed primarily of hydrogen. They are, in fact, little more than large globs of gas held together by gravity. They do lose mass (mostly hydrogen) into space in the form of the stellar wind and occasional outbursts in flares. But because they are so massive, they are able to retain their gaseous bodies through self-gravitation. They shine because they are large enough that their central pressures and temperatures are high enough to allow hydrogen fusion reactions to take place. It is the energy released by these fusion reactions that provides the heat to offset the effect of gravity and prevent gravitational collapse into a tiny ball of matter. When stars "run out" of fuel they undergo drastic changes driven by gravity that ultimately result in very small and dense objects such as white dwarf stars and even black holes. This only happens late in a star's lifespan, after billions of years for a star the size (mass) of the sun. What happens at earlier stages of the star's history is what concerns us here.

A young star, say the sun 4.5 billion years ago shortly after it formed, has the highest concentration of hydrogen in its core that it will ever have. This means that the fusion reactions required for gravitational stability can happen at a lower temperature because the hydrogen atoms (actually protons) bump into one another and fuse more readily than if they were "diluted" by the presence of other atoms (ions) such as helium. As stars consume hydrogen, converting it into helium, the hydrogen becomes progressively more diluted with the helium. It becomes harder for the hydrogen to fuse because the helium "gets in the way". To compensate for this, the core of the star contracts to a higher density where the fusion reactions can proceed at a high enough rate because at higher density all of the ions, and especially the hydrogen, are closer together. The contraction of the core results in a higher core temperature and a higher temperature gradient between the core and the surface of the star. This higher gradient leads to more rapid heat flow from the center to the surface of the star where it radiates into space. This actually requires a somewhat higher rate of fusion to provide the extra energy. The principle result of interest to us

G. H. Shaw, *Earth's Early Atmosphere and Oceans, and The Origin of Life,*
SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-21972-1_10

is that as it ages the star radiates progressively more energy per unit time into space as it consumes its hydrogen fuel. This process can be readily modeled and over the 4.5 billion years of the sun's life (so far) its luminosity has increased by about 25 %. In other words, the early Archean sun was about 25 % fainter than it is now, the so-called faint young sun.

The faint young sun presents us with a problem unless the early Archean Earth had a strikingly colder climate that at present. If we were to suddenly turn the sun down right now by 25 % all of the oceans would freeze and Earth would be an ice-ball planet. Although there are suggestions that Earth has occasionally experienced so-called "snowball" episodes much more extensive than the recent glacial epochs, the evidence that Earth's climate has generally been similar to the present, with liquid water a prominent surface feature, is overwhelming even as far back as 4.2–4.4 billion years ago (Wilde et al. 2001; Nutman 2006). This presents the puzzle of how these warm conditions were possible with such a faint sun. Most effort has focused on discussions of enhanced levels of greenhouse gases (GHG). Since carbon dioxide is one of the most important GHG for which we can imagine processes for increasing its atmospheric concentration, and because most models of planetary degassing consider CO_2 to be the main carbon species released to the early Earth's atmosphere, it appears relatively easy to compensate for lower solar luminosity by adding CO_2 to the early atmosphere and keeping it there. The concentration required has been calculated to have been as much as 0.2 atmospheres of CO_2, or about 500 times the present concentration (Kasting 1993). There is certainly enough carbon (in the form of carbonate) to accomplish this. The atmosphere contains about 5×10^{18} g of carbon dioxide and carbonate rocks more than 10,000 times that amount, so a fraction of the carbonate rock converted to atmospheric CO_2 would do the trick. The problem with this is similar to the problem raised by Rubey (1951, 1955) when considering whether degassing was rapid very early in Earth's history or gradual over time: an atmosphere so rich in CO_2 would produce noticeable effects in terms of weathering products and subsequent sedimentary deposition. There is also the problem of how to keep the CO_2 in the atmosphere because it reacts so readily with silicates to ultimately make carbonate minerals.

Two attempts to resolve the problem depart slightly from the focus on GHG as such, although they both assume some enhancement of GHG. One suggests a model with a significantly lower albedo for the early Earth due to smaller continental area (Henderson-Sellers 1979; Rossow et al 1982). The model seems compelling because the area of Earth covered by continents has almost certainly increased through time, as discussed above. However, it is also clear that continental area increased rapidly through the early Archean, and surely some emergent land masses existed even during the Hadean. This is important because even a relatively modest land mass in either polar zone would trigger a sharp albedo increase due to formation of polar ice, which could lead to a climate runaway (in the freezing direction) that would be difficult to reverse, even with subsequent (slow) shifts in location of land masses. Obviously the actual magnitude of this effect is highly model dependent, but the sensitivities of climate models during the FYS era certainly fall on the side of cooling (freezing). The uncertainties implicit in climate modeling are so large that incorporating the possibility of polar ice formation only adds to doubts that albedo effects can solve the FYS problem.

Another attempt to solve FYS considers a mechanism to protect a very potent GHG, ammonia, from UV destruction, without creation of an antigreenhouse haze (Pavlov et al 2001; Haqq-Misra et al. 2008). This ingenious approach (Wolf and Toon 2010) considers the detailed optical properties of fractal organic particulates in a reducing atmosphere, with UV selectively absorbed compared to visible wavelengths. This discussion is based on estimates made using measurements on particulates in Titan's atmosphere and it is not clear that these can be extended to early Earth conditions, under which particulate compositions may have been significantly different from Titan, especially with the addition of oxygen-bearing organics to the mix. It remains to be seen if this is really a solution to FYS. Of particular interest in these discussions is the question of the relative importance of CO_2 as a constituent of the earliest atmosphere. A CO_2 dominated atmosphere has long been the conventional wisdom (Kasting 1993, 1997, 2004). In this regard positing addition of significant reducing components to the atmosphere represents a departure from conventional wisdom.

Various efforts to resolve FYS with reduced gases, especially CH_4 and NH_3, reflect recognition that a CO_2 dominated atmosphere cannot fix the problem without unreasonably high levels of atmospheric CO_2 (Sagan and Mullen 1972; Pavlov et al 2000; Kasting 2005; Shaw 2008; Wolf and Toon 2010). Such efforts *may* reflect a shift away from early atmosphere models dominated by CO_2 (and N_2), although skepticism about the importance of reduced gases in the early atmosphere continues. On the other hand, as pointed out above, recent theoretical calculations appear to strongly support the idea that highly reduced gases were abundant during primary degassing (Schaefer and Fegley 2007, 2010; Zahnle et al. 2010).

If any of the previous efforts to solve the FYS paradox were truly convincing we would probably see an end to discussions of the problem. In fact, atmospheric ammonia may well be the answer that has been sought for so long. Rejection of ammonia as the GHG responsible for countering the FYS has largely been based on photolytic destruction rates of ammonia more than sufficient to extinguish present-day natural production (Kuhn and Atreya 1979; Kasting 1982). As noted above, much of the discussion of NH_3 centers on finding a mechanism to shield the gas from UV radiation, without compromising heating of the lower atmosphere and surface by incoming solar radiation. Most of these approaches also add CH_4 to the atmosphere as another component in a rather reduced system. However, shielding is not necessary if rates of ammonia (and methane)production during the Hadean and early Archean were significantly higher than at present. There is evidence to support this idea from studies of hydrothermal systems (Seewald et al. 1990, 1994, 2006; Seewald 2001; Cruse and Seewald 2006).

Ammonia concentrations as low as 1–10 ppm were suggested in the early 1970s as sufficient to counteract the FYS (Sagan and Mullen 1972; Pavlov et al. 2001). Modern biogenic ammonia production is on the order of 10^{10} molecules-cm^{-2}s^{-1} (Pavlov et al. 2001), and the production rates necessary to give 10 ppm ammonia have been estimated to be on the order of 3×10^{12} cm^{-2}s^{-1}, i. e. 2–3 orders of magnitude higher than modern biological production rates. Even with atmospheric shielding it would be difficult for biogenic ammonia to overcome the FYS. And, because biogenic ammonia would have been impossible under pre-biotic or probably even the earliest biotic conditions, ammonia has received little subsequent attention in early Earth climate models, until recently, as indicated above. In any case, *most* recent detailed climate models do not include ammonia as a significant atmospheric component largely because of its presumed rapid destruction.

However, there are abiotic processes that can produce both ammonia and methane, and in significant quantities, as indicated in the previous chapter, in terms of getting enough hydrogen into the atmosphere to support high levels of hydrogen loss to space. Theoretical calculations, laboratory experiments and field observations (Seewald et al. 1990, 1994, 2006; Seewald 2001; Cruse and Seewald 2006) all support the idea that abiotic processes can produce ammonia and methane in quantities sufficient to affect atmospheric chemistry, and thus climate. The relevant processes are not only present now, but were probably quantitatively more important in the Hadean and early Archean. Production of reduced gases by low to moderate temperature hydrothermal processing of organic-rich sediments was probably more intense and more widespread in the early Archean. Based on the likely thermal state of the Earth shortly after formation of a permanent solid surface, this primordial hydrothermal activity was at least 4–5 times the present value and could easily have been more than an order of magnitude greater than at present. This is the combined result of higher primordial concentrations of radiogenic isotopes and residual near-surface impact heating.

Hydrothermal processing of sediments containing 2.5 % organic carbon and.2 % nitrogen yields almost quantitative release of N as NH_3 and a large fraction of the carbon as methane (Table 9.1). If Hadean sea water contained as little as 1 % dissolved organic carbon, and a proportionate amount of organic nitrogen, circulation of such water through hydrothermal systems at rates comparable to present-day rates of 2.5×10^{12} m^3/yr^{31} could have produced approximately 10^{11} molecules/cm^2-s of ammonia, and similar amounts of methane. Considering the presumably higher rates of circulation associated with the higher Hadean heat flow the rate of ammonia production could easily have been close to 10^{12} molecules/cm^2-sec., close to the amount need to counteract the faint young sun. Adding to this the production of ammonia from processing of *sediments* near mid-ocean and subduction-like settings in the ocean suggests even higher levels of atmospheric ammonia. Atmospheric shielding by organics would not even be necessary at these levels, thus avoiding the anti-greenhouse effect of an organic haze. On the other hand, the concomitant hydrothermal release of methane would seem to provide the necessary raw material to produce such a haze, which complicates any detailed modeling.

Volcanogenic production of CO_2 during the Hadean and early Archean is not precluded by these considerations, nor are modest atmospheric CO_2 levels that could suppress the production of an organic haze. In all likelihood the amounts of CO_2 would have been comparable, as a minimum, to present-day volcanic gas emissions. Regardless of the amount of CO_2 generated by volcanism or resident in the atmosphere, the NH_3 and CH_4 present in the atmosphere, *and continually regenerated* *hydrothermally*, would be sufficient to produce a warm climate in spite of the faint young sun.

Obviously the major assumption made here is the presence of large amounts of dissolved organic compounds in the early ocean and/or in ocean sediments. These must have been the result of primordial degassing and early atmospheric and ocean chemistry, chemistry no doubt strikingly similar to that of the Miller-Urey experiment, as has been discussed above. Once present in the amounts suggested, these compounds would certainly be maintained by photochemical processing of NH_3, CH_4, H_2O, etc. in the atmosphere followed by rainout into the ocean, where the hydrothermal systems could regenerate components of a reduced atmosphere as proposed above. This solution to the FYS problem is also, as has been seen in previous chapters, the solution to the problem of generating the large reservoir of prebiotic organic compounds necessary for the emergence of life. The processes by which the surface environment evolved from these strongly reducing conditions to the modern highly oxidized surface depend on the cycling of carbon compounds as discussed in Chap. 9, but especially in the later stages by the development of biological innovations, oxygenic photosynthesis in particular. Without oxygenic photosynthesis it is not only unlikely that Earth could have reached the current oxidation state at the surface, but life as we know it would never have developed.

References

Cruse AM, Seewald JS (2006) Geochemistry of low molecular weight hydrocarbons in hydrothermal fluids from middle valley, northern Juan de Fuca Ridge. Geochim Cosmochim Acta 70:2073–2092

Haqq-Misra JD, Domagal-Goldman SD, Kasting PJ, Kasting JF (2008) A revised, hazy methane greenhouse for the Archean Earth Astrobiology 8:1127–1137

Henderson-Sellers A (1979) Clouds and the long-term stability of the Earth's atmosphere and climate. Nature 279:786

Kasting JF (1982) Stability of ammonia in the primitive terrestrial atmosphere. J Geophys Res 87:3091–3098

Kasting JF (1993) Earth’s early atmosphere. Science 259:920–926

Kasting JF (1997) Warming early Earth and Mars. Science 276:1213

Kasting JF (2004) When methane made climate. Sci Am 291:78–85

Kasting JF (2005) Methane and climate during the precambrian era. Precambrian Res 137:119–129

Kuhn WR, Atreya SK (1979) Ammonia photolysis and the greenhouse effect in the primordial atmosphere of Earth. Icarus 37:207–213

Nutman AP (2006) Antiquity of the oceans and continents. Elements 2:223–227

Pavlov AA, Kasting JF, Brown LL, Rages KA, Freedman R (2000) Greenhouse warming by CH_4 in the atmosphere of early Earth. J Geophys Res 105:11981–11990

Pavlov AA, Brown LL, Kasting JF (2001) Shielding of NH_3 and O_2 by organic hazes in the Archean atmosphere. J Geophys Res 106(E10):23267–23287

Rossow WB, Henderson-Sellers A, Weinreich SK (1982) Cloud feedback: A stabilizing effect for the early Earth? Science 217:1245

Rubey WW (1951) Geologic history of sea water: an attempt to state the problem. Geol Soc Am Bull 62:1111–1147

Rubey WW (1955) Development of the hydrosphere and atmosphere, with special reference to probable composition of the early atmosphere. In: Poldervaart A (ed) Crust of the Earth, Vol. 62. Geological Society of America Special Paper, Denver, pp. 631–664

Sagan C, Mullen G (1972) Earth and Mars: evolution of atmospheres and surface temperatures. Science 177:52–56

Schaefer L, B. Fegley J (2007) Outgassing of ordinary chondritic material and some of its implications for the chemistry of asteroids, planets and satellites. Icarus 186:462–483

Schaefer L, Fegley B Jr (2010) Volatile element chemistry during metamorphism of ordinary chondritic material. Icarus 205:483–496

Seewald JS (2001) Aqueous geochemistry of low molecular weight hydrocarbons at elevated temperatures and pressures: constraints from mineral buffered laboratory experiments. Geochim Cosmochim Acta 65:1641–1664

Seewald JS, Seyfried WE Jr., Thornton EC (1990) Organic-rich sediment alteration: an experimental and theoretical study at elevated temperatures and pressures. Appl Geochem 5:193–209

Seewald JS, Seyfried WE Jr., Shanks WC III (1994) Variations in the chemical and stable isotope composition of carbon and sulfur species during organic-rich sediment alteration: an experimental and theoretical study of hydrothermal activity at Guaymas basin, Gulf of California. Geochim Cosmochim Acta 58:5065–5082

Seewald JS, Zolotov MY, McCollom T (2006) Experimental investigation of single carbon compounds under hydrothermal conditions. Geochim Cosmochim Acta 70:446–460

Shaw GH (2008) Earth's atmosphere: hadean to early proterozoic. Chem Erde 68:235–264

Wilde SA, Valley JW, Peck WH, Graham CM (2001) Evidence from detrital zircons for the existence of continental crust and oceans on the Earth 4.4 Gyr ago. Nature 409:175–178

Wolf ET, Toon OB (2010) Fractal organic hazes provided an ultraviolet shield for early Earth. Science 328:1266

Zahnle K, Schaefer L, Fegley B Jr (2010) Earth's earliest atmospheres. In: Deamer D, Szostak JW (eds) The origins of life. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp. 49–66

Chapter 11
The Origin of Oxygenic Photosynthesis and Its Impact on the Atmosphere

11.1 Earliest Photosynthesis

The evolutionary history of photosynthesis is extraordinarily complicated (Blankenship 2013). In part this is because there are not only a number of specific pathways involved in photosynthetic production of reduced carbon compounds, but there are two major types of photosynthesis: anoxygenic and oxygenic. The former system uses reduced compounds such as molecular hydrogen, sulfides or pre-existing organic compounds as sources of hydrogen (or more precisely, reducing capacity), with no free oxygen produced as a consequence of bio-synthesis. The latter uses water as the source of hydrogen for biosynthesis and releases free oxygen. The only other potential source of free oxygen in the atmosphere is photodissociation of water in the upper atmosphere followed by loss of hydrogen to space (Zahnle and Catling 2014). This process is so slow that it could not provide an atmosphere with much free oxygen, especially starting from generally reducing conditions at the surface. Oxygenic photosynthesis is hundreds or thousands of times more powerful as a source of free oxygen and is certainly the source of our oxygen-rich atmosphere.

Phylogenetic analysis supports the idea that anoxygenic photosynthesis developed first (Blankenship 2013). It is found in bacteria of certain groups and uses one of two distinct pathways starting with trapping of light energy by pigments (bacteriochlorophylls) that pass the energy along a chain of enzymes that use it for biosynthesis. The source of the hydrogen used in these reactions to convert CO_2 into useful compounds is from hydrogen sulfide, pre-existing organic compounds (certainly initially produced by abiotic processes), or directly as molecular hydrogen. Just when life evolved the capability to trap and use solar energy is unknown. No fossil remains are known prior to about 3.5 BYBP, and the oldest may contain no fossil organisms as such, but only the sedimentary depositional signature of their presence. These so-called stromatolites (finely laminated mineral layers) are presumed to be the result of photosynthetic activity that consumed CO_2 and altered the chemistry of the local aqueous environment to the point where minerals precipitated. Stromatolites occur throughout the sedimentary record, from 3.5 BYBP to the present. In modern examples it can be seen that they are formed largely by photosynthetic

G. H. Shaw, *Earth's Early Atmosphere and Oceans, and The Origin of Life,*
SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-21972-1_11

bacteria in thin mats at the top of the stromatolite structure, which is typically in the form of a column perhaps a meter tall and somewhat less than that across, although dimensions can vary considerably. The bacterial community is actually quite complex, but the key element in creating the structure is the photosynthetic mat at the very top. This mat induces precipitation of mineral matter (mostly carbonate, but in earlier environments chert) from the saline water. As mineral collects the bacteria migrate upward to reform the mat, in order to stay exposed to the light and continue photosynthesis. Repetition over an extended period creates the columnar structure. In most older stromatolites we have only the mineral structure to indicate the prior presence of the bacterial mat. The case for old stromatolites as evidence for previous bacterial activity is largely inferential, but by comparison to modern examples and given the geologic settings, which typically include evidence for shallow water and other indications of near coastal environments, it is certainly reasonable to conclude that they are the result of photosynthetic bacteria. Whether the earliest stromatolite-forming bacteria were anoygenic or oxygenic photosynthesizers is a more difficult question to answer; both options have been proposed.

Photosynthesis generates a distinctive chemical signature in the form of the carbon isotopic composition of the organic material produced. During photosynthesis carbon-12 (in CO_2) is selectively taken up by the photosynthetic machinery compared to carbon-13, that is, it reacts faster with the enzymatic system and is therefore selectively incorporated into the organic compounds produced. This means that the carbon in the organic compounds is enriched, typically by about 2–3 % compared to the initial carbon dioxide/carbonate source. This very small difference is easily measured in modern samples and in organic compounds and carbonates deposited in rocks. Carbon samples from minerals as old as 3.8 billion years have been found to have carbon isotopic compositions consistent with photosynthesis, seemingly indicating the presence of photosynthetic bacteria of that age (at least). Similar results are found throughout the geologic record since then. This does not mean that the earliest photosynthesis was oxygenic, and it is probable that it was not (Blankenship 2013). However, another chemical marker has been found in rocks about 2.7 billion years old which is more definitive as an indicator of cyanobacterial (and thus oxygenic) activity. This is a suite of compounds (2-methyl hopanoids) that are known degradation products of compounds thought to be characteristic of cyanobacteria (Brocks et al. 1999).

The timing of various important events in the development of life, and particularly the development of photosynthesis is still rather uncertain. Because oxygenic photosynthesis is so important to the evolution of Earth's atmosphere it would be helpful to have a firm date. Lacking anything better, it seems like a reasonably conservative view to take the 2.7 BYBP figure as a starting point for purposes of discussion of the consequences of this vital event. One thing we can be reasonably certain about: cyanobacteria were the first serious oxygen producers, and probably remained so for hundreds of millions of years, at least until the first eucaryotic algae emerged. This latter event may have been crucial in ultimately producing the diversity of plant and animal life now present, but it was the cyanobacteria that set the scene, and the cyanobacteria that also gave rise to the chloroplasts used by all eukaryotic photosynthetic plants (Blankenship 2013). Indeed, in pelagic ocean settings cyanobacteria remain the main oxygen producers.

The initiation of oxygenic photosynthesis by 2.7 BYBP presents a problem with evolution of the chemistry of Earth's surface and atmosphere. Certain measurements on the isotopes of sulfur indicate that Earth's atmosphere lacked free oxygen as late as 2.3–2.4 BYBP, at least 300–400 million years after the emergence of powerful oxygen producers. In an oxygen-free atmosphere sulfur can readily exist both in oxidized (as in SO_2, etc.) and reduced states (as in H_2S and elemental sulfur). Once there is free oxygen in the atmosphere the reduced forms of sulfur are quickly removed by oxidation. Various reactions of the different oxidized and reduced sulfur species in the atmosphere produce fractionation of the several different isotopes of sulfur, which can be preserved when the reaction products rain out and become incorporated in sediments. If both reduced and oxidized compounds of sulfur are present in the atmosphere a certain kind of fractionation, called mass independent fractionation or MIF, can be produced and preserved. If only the oxidized form of sulfur is present in the atmosphere, as would be the case once free oxygen was present, no MIF will be produced. Detailed and extensive studies of sulfur isotopes in samples going back to nearly 4 BYBP show a marked decrease in MIF after about 2.3 BYBP, to essential zero (Farquhar et al. 2000; Farquhar and Wing 2003). This has been construed as evidence that free oxygen in the atmosphere, at least as a general and widespread condition, began more or less suddenly about 2.3 BYBP. This, of course, does not mean that free oxygen was not present locally before that time, but that the atmosphere as a whole was devoid of free oxygen. What makes this especially interesting is the length of time the atmosphere remained in a highly reduced state following the advent of cyanobacterial oxygenic photosynthesis. This is discussed in the next section, along with considerations of the carbon cycle record preserved in sedimentary rocks.

11.2 The Geologic Record of Carbon Isotopes in Sedimentary Rocks and the Rise of Oxygen

Many sedimentary rocks contain significant amounts of carbon, especially limestones (mostly calcium carbonate), so-called black shales, and the various forms of fossil fuels. The latter include coal and the several types of rock with relatively high concentrations of methane (natural gas), and petroleum. The carbon in these rocks is about 1 % carbon-13 (six protons and seven neutrons in the nucleus) and the rest is carbon-12 (six protons and six neutrons). Carbon-14, so well known from radioactive age determination, is restricted to young rocks less than 100,000 years old because of its short half-life. The two stable isotopes are of interest here because biochemical processes, especially photosynthesis using CO_2 as a carbon source, discriminate between the two isotopes, selectively using the light carbon-12 as discussed above, which is somewhat more concentrated in the products of photosynthesis. The CO_2 left behind, which may be preserved in carbonate rocks, gains a slightly higher concentration of carbon-13. The typical depletion of C-13 in photosynthetic carbon is about 30 parts per 1000 (‰) compared to contemporaneous inorganic (carbonate) carbon, that is, the photosynthetic carbon is "lighter" by about 3 %. Because of the standard used in these measurements, carbonate carbon is often (but not always) near 0 ‰ and photosynthetic carbon near −30 ‰ relative to the standard. These small differences are readily preserved in organic-rich and carbonate rocks and can be measured in the laboratory.

The long-term record of carbon isotopes in rocks, covering more than 3.5 billion years, is one of the most striking results of stable isotope geochemistry. Figure 11.1 is a compilation of results from Early Archean to recent, showing $\delta^{13}C$ for both carbonates and organic-rich sedimentary rocks. The isotopic data strongly support not only a more or less constant ratio between photosynthetic carbon and "residual" carbonate carbon throughout the Precambrian, but values of carbonate carbon stay close to 0 ‰. This near constancy throughout a long span of time, coupled with the established $\delta^{13}C = 30$ ‰ difference (Broecker 1970) between photosynthetic carbon and carbonate carbon, indicates a controlling mechanism for producing and maintaining this isotopic signature over eons. There are some small, but probably significant, variations in the record, such as around 2.5–2.7 billion years before present (BYBP), and toward the end of the Proterozoic, but the overall pattern holds.

Figure 11.2 illustrates why a lot of effort has been expended in trying to analyze and interpret the carbon istotope record. Based on this diagram the carbon isotope record (Fig. 11.1) seems to indicate that throughout much of Precambrian time about 20 % of the active carbon pool has been in the form of reduced (photosynthetically produced) carbon, a ratio that is similar to what we see in existing sedimentary carbon. The underlying importance derives from the notion that the main source of free oxygen, and oxidizing potential generally, is due to photosynthetic production of oxygen. The carbon isotope record may thus provide us with a means of quantifying the production of oxygen through time, assuming we know the composition of the various carbon pools and how they have changed with time. This realization has given rise to a progressively more complicated set of models which try to interpret the carbon isotope record in terms of oxygen (or oxidation capacity) production through time (Hayes and Waldbauer 2006). These efforts are further complicated by the fact that some fraction of oxidation production is certainly due to hydrogen loss from the upper atmosphere, but first we'll examine the implications of the carbon isotope record using the conventional model for Earth's early atmosphere, that is, carbon dioxide was the main carbon species degassed from the early Earth, and continued to be the dominant carbon species at the surface until the present.

The conventional model attributes oxygenation to the more or less gradual photosynthetic conversion of CO_2 to organic compounds, with net oxidation resulting from burial (isolation) of the reduced carbon from the surface environment. This presumably only occurs following the emergence of oxygenic photosynthesis. Anoxygenic photosynthesis may produce a carbon isotope signature but carries no implications in terms of oxidation. For this reason, the time of origin of oxygenic photosynthesis is of considerable importance, and is discussed in essentially all the

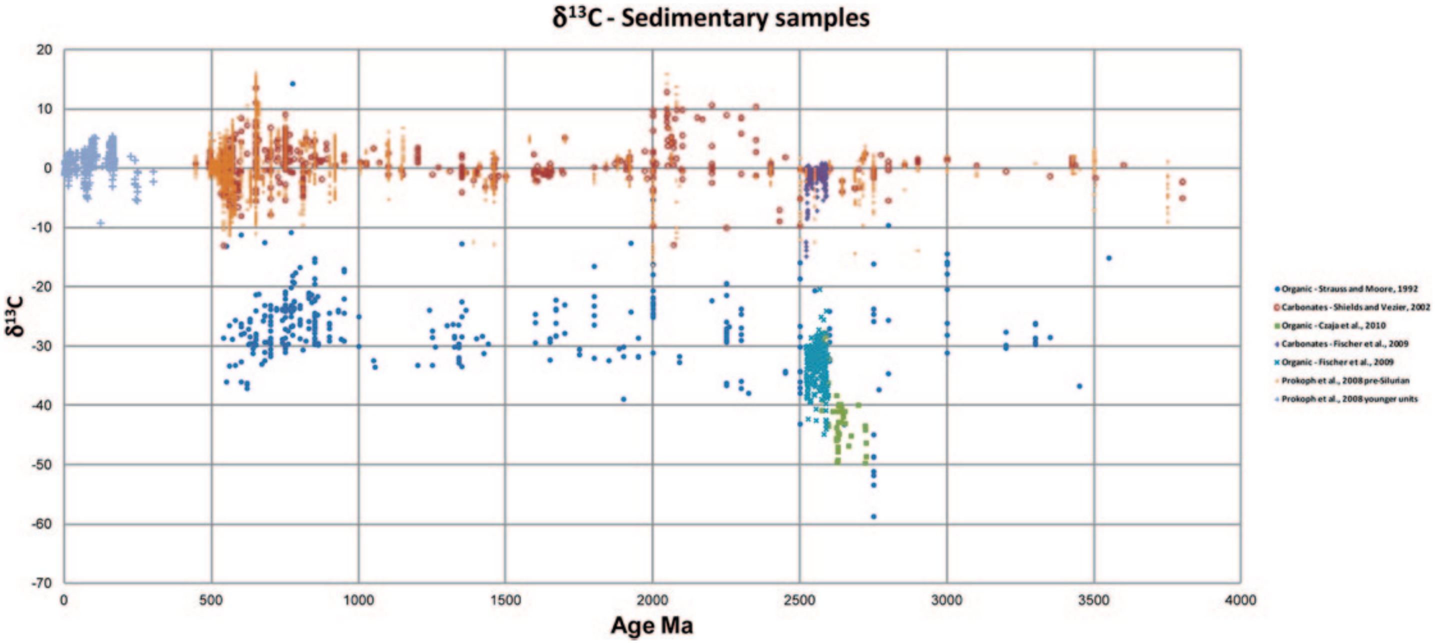


Fig. 11.1 Variations in $\delta^{13}C$ for rocks throughout the geologic record

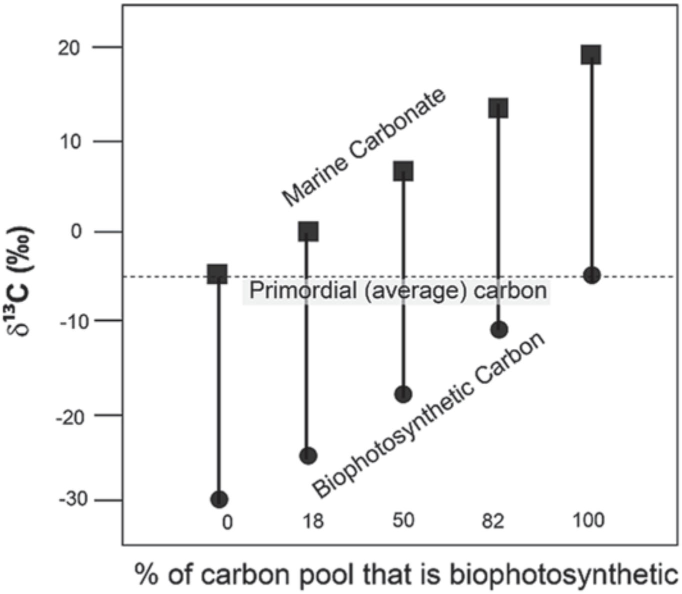


Fig. 11.2 Carbon isotopic signatures for photosynthetic and carbonate carbon assuming photosynthesis operating on an original carbon(ate) reservoir having primordial (e.g. mantle-derived) carbonate ($\delta^{13}C = -4.5$ ‰). (After Broecker 1970)

modeling analyses of this type. The earliest data on carbon isotopes in Fig. 11.2 fall around 3.5 BYBP, with a couple points at about 3.7–3.8 BYBP. These are for carbonates, with rather fewer points for organic material. In fact the earliest points for carbonate give values scarcely different from "primordial" (i.e. unfractionated) carbon, and may not indicate effects of photosynthetic activity at all. It is interesting that the earliest data for which we have both carbonate and organic carbon appear to be affected by photosynthesis, which may or may not be oxygenic. As yet there is no definitive resolution to this issue. As mentioned earlier, rather compelling evidence for oxygenic photosynthesis has been found in samples from about 2.7–2.8 BYBP, although this has not been without challenge, and arguments have been made for an even earlier origin. If we accept the 2.8 BYBP figure as a starting point we can carry out an analysis from there. The state of the carbon pool(s) before this time may or may not be of great importance to oxidation modeling, but the interval of 300–400 MY or so between 2.8 BYBP and the evidence for atmospheric oxygen at 2.3–2.4 BYBP from mass-independent fractionation of sulfur isotopes presents a serious problem.

The essence of this problem is the lability of the carbon pools, that is, the ease and speed with which carbon moves from one reservoir to another, and especially the rate at which it enters and leaves the atmosphere/ocean system at Earth's surface, where the isotopic signature can be altered by photosynthesis. The problem can be illustrated in its simplest form if one assumes that the atmosphere/ocean system contained all the presently known surface carbon (as CO_2 or CO_3^-) at the beginning of oxygenic photosynthesis. In this case the early photosynthesizers were clearly (cyano)bacterial, which means that they could have propagated rapidly, consuming CO_2 and quickly producing equivalent amounts of oxygen. Nutrient limitations are an obvious possible explanation for the 300–400 MY delay in surface oxidation. The three most likely nutrient candidates are iron (needed for the photosynthetic enzymes), and nitrogen and phosphorus, needed for essential biomolecules such as DNA and proteins. Iron was surely present in abundance during the early anoxic phase of the atmosphere, as dissolved ferrous iron in sea water, while being required in only trace amounts.

Some species of cyanobacteria are among the most important and effective nitrogen fixers, including some of the most abundant species in the oceans. Taking current estimates of cyanobacterial nitrogen fixing rates we can estimate that, as a minimum, cyanobacteria could have produced enough fixed nitrogen to yield the current atmospheric inventory of free oxygen in a few million years. In 30 million years the oxygen produced would amount to about half of the oxidation capacity implied by the current pool of fixed photosynthetic carbon stored in organic-rich rock in the crust. Given this timing it is hard to see how surface oxidation could have been delayed for hundreds of millions of years, unless there was a very large sink of oxygen available (more on this later).

Phosphorus then becomes the remaining possible limiting nutrient. Phosphorus is an important trace constituent of igneous rocks, both basaltic ocean crust and granitic (in the more inclusive sense) continental crust. An average concentration is about 800 ppm for either. Weathering of both rock types liberates phosphorus, making it available to organisms. At the point in Earth's history of interest to this discussion all life would have been microbial. There were obviously no large organisms that use phosphorus to make skeletal parts, such as bones. Such organisms in modern oceans can effectively remove phosphorus from the environment, decreasing its availability to others, including photosynthesizers. This was not the case in the Archean because phosphorus was "stored" in readily degraded biological organics. Inorganic precipitation of calcium phosphate is also a possible sink and may have provided a general control on maximum phosphate levels in the oceans. "New" phosphorus from weathering of rock would have been essentially entirely available to microbes, including cyanobacteria. Because the underlying energy resource used by cyanobacteria was abundant solar radiation, they were, as a group, in a very good position to consume phosphorus, essentially as rapidly as it became available. If we use modern rates of erosion of continents and the phosphorus content of rocks we can estimate the potential for oxygen productivity of cyanobacteria. From the data in Table 2.3 we get phosphorus production of about 3×10^{13} g/year from weathering. This converts to about 1.3×10^{15} g/year of carbon, using a C/P mass ratio of 42, giving oxygen production of about 4×10^{15} g/year. Although this assumes modern rates for continental weathering and erosion, it does not include potential phosphorus contributions from alteration of ocean crust, which probably at least partly makes up for smaller continental mass in the Archean. This rate of oxygen production could produce the amount in the modern atmosphere in 300,000 years, less time than that implied by nitrogen limitation. That is, phosphorus limitation on oxygen production is less than for nitrogen, which we have seen is not enough to account for the 300–400 MY delay in oxygenation of the surface following the advent of photosynthesis. Other sinks of phosphorus are possible, and have been suggested, e.g. trapping of phosphorus by ferric hydroxide particles (Jones et al. 2015), but ferric hydroxide was not likely to have been abundant on the anoxic Earth. Their recent discussion of this problem develops the idea that ferric iron was produced in the shallow oceans by photoferrotropic bacteria using aqueous ferrous iron as a source of biosynthetic reductant. They emphasized the importance of the subsequent removal of phosphorus by sorption onto ferric hydroxide particles from upwelling flows of both ferrous iron and phosphorus. This would very likely have been important in limiting upwelling deep water as a source of phosphorus to the near-surface biosphere. Such upwellings are certainly important phosphorus sources in ocean settings at present, where phosphorus in deep water is derived from decay of organisms sinking into the ocean. However, it is likely that the most important source of phosphorus in Archean shoreline environments was from weathering of continental (island-arc) rocks. With virtually no organisms on land to trap phosphorus, or coastal plants and animals to capture it when the phosphorus washed into the ocean, this important nutrient was delivered almost directly to the cyanobacteria. Perhaps some was lost to ferric hydroxide in the near-shore setting, but this would be a relatively small sink, and the cyanobacteria would capture most of it. Interestingly, many cyanobacteria have mechanisms for trapping and storage of phosphorus, an adaptation that would have given them significant advantages in maintaining growth and reproduction, and thus oxygen production (Mur et al. 1999).

This discussion of phosphorus focuses on the quantity of phosphorus available, mainly from weathering of rocks. It does not address the chemical form (oxidation state) of the phosphorus, but assumes it is in the +5 state characteristic of most phosphate minerals, such as apatite. It is instructive to ask about other (more reduced) oxidation states of phosphorus since the model proposed here for the early surface may have resulted in production of such reduced compounds. Interestingly, a considerable amount of recent research looks at the importance of reduced phosphorus compounds as phosphorus nutrient sources in modern ocean settings, in particular in the phosphorus-poor oligotrophic zones of the tropics and sub-tropics (Van Mooy et al. 2015; White et al. 2010; Beversdorf et al. 2010). These studies show that some +3 phosphorus compounds can be both synthesized and used as nutrient sources by certain species of cyanobacteria. One particularly common variety of cyanobacteria, *Trichodesmium* IMS101, is also a major nitrogen-fixer in the ocean (Capone et al. 1997). One point of these studies is to understand the photosynthetic productivity implications for these vast low-nutrient zones of the global ocean. It appears that the ability to produce their own nitrogen removes nitrogen as a limiting nutrient and allows the organisms to deplete phosphorus in the water column, but also requires approaches to phosphorus utilization favoring both production and consumption of reduced phosphorus (+3) compounds. The suggestion is made that certain species have developed biosynthetic adaptations to allow them to thrive in these environments. An additional example is in a sense even more extreme in that a different cyanobacterium (*Prochlorococcus* MIT9301) actually can use phosphite

(P=+1 oxidation state) as a nutrient source (Martinez et al. 2012). This cyanobacterium is the dominant primary producer (numerically) in the oligotrophic zones of the ocean (Martinez et al. 2012), and is a major source of photosynthetic oxygen.

Although it is conceivable that these very ancient cyanobacterial lineages may have evolved their biochemical pathways *relatively* recently, allowing them to thrive in phosphate (+5)-poor environments, it seems more likely that these biochemical pathways may be very ancient, developed at a time when reduced phosphorus compounds may have been a more prominent (and accessible?) part of the surface environment. It is especially striking that these adaptations are found in such ancient oxygen-producing lineages. This suggests that the metabolic pathways involving reduced phosphorus compounds are perhaps "pre-adaptations" allowing the organisms to take advantage of the low nutrient conditions found in the oligotrophic zones of the modern ocean.

The conclusions we are drawn to are either (1) there was a large pool of reducing capacity available to absorb the oxygen as it was produced, such as a large pool of pre-existing abiotic reduced carbon compounds, or (2) carbon dioxide was not readily available in the surface environment to "feed" cyanobacterial growth. The first notion would be a direct consequence of the model for Earth's early atmosphere proposed in part 1 and is discussed further below. Other possible sinks might have been a combination of ferrous iron from weathering and accumulated as a dissolved component of sea water, and sulfur/sulfide from weathering and volcanic emissions. Ferrous iron is readily oxidized to ferric iron by free oxygen and sulfide sulfur may be oxidized to sulfate. Using 6% as the concentration of ferrous oxide in rocks we can calculate the maximum reducing potential of weathering, similar to the approach in calculating the oxidizing potential associated with phosphorus calculated above. The result is oxygen consumption somewhat less than the potential oxygen production as limited by nitrogen availability. This is not too surprising since the presence of oxygen at present clearly indicates that oxygen consumption by weathering has certainly been overcome by oxygen production over time. This calculation seems to show that production of free oxygen might well have been slowed by oxygen consumption due to iron oxidation, and if one adds in something for sulfide oxidation the effect could be even greater. However, these numbers are certainly overestimated by a large amount. This is because a large fraction of sulfur from weathering under anoxic conditions would have been in the form of sulfide minerals, especially pyrite (FeS_2), which would subsequently be deposited unoxidized in sediments (along with some of the ferrous iron), and thus not available to consume oxygen. In addition, the presence of ferrous iron in sea water serves as a trap for hydrogen sulfide as it also leads to precipitation of insoluble iron sulfides. Finally, once carbonate is present in sufficient quantity in sea water (which was probably fairly early), the precipitation of siderite (ferrous carbonate) would also have removed some fraction of the ferrous iron from the active reductant pool.

The second possible explanation for the delay in surface oxidation, which has received most of the attention in discussions of the early Earth, implies storage of carbon dioxide elsewhere, thus isolating it from participation in oxygenic photosynthesis. Given that there is little evidence for major early continental Archean carbonate deposits that could account for the necessary storage, it would have been necessary to store the carbon dioxide within the oceanic crust or put it back into the mantle. It may have been stored by reaction with oceanic crustal rocks (Sleep and Zahnle 2001; Sleep 2010), which can also return the carbon dioxide to the mantle by subduction of carbonates. Long-term reinjection of carbon dioxide into the upper mantle seems unlikely, given that thermal conditions in the upper mantle were presumably, at least in part, responsible for release of carbon dioxide to the atmosphere in the first place. While some have suggested that mantle cooling would allow significant recarbonation of the upper mantle, the amount of cooling, even beyond the end of the Archean, should be relatively small. Heat loss from the mantle is largely accomplished by convective processes. Assuming that modern heat flow is perhaps 1/4 of that during the mid-Archean (due to loss of radioactive heat sources from decay over time, with some additional early heat flow due to residual impact heating), the rate of convection need only fall to 1/4 of the mid-Archean value. Because of the very strong temperature dependence of silicate viscosity, this does not translate into a dramatic decrease in upper mantle temperature, and not likely enough of a decrease to strongly stabilize carbonates in the upper mantle.

This leaves carbonation of oceanic crust as a possible sink and storage site for the "excess" carbon dioxide (Sleep and Zahnle 2001). Such carbonation is, indeed, known from samples of Archean crust (Nakamura and Kato 2004). In fact, carbonation of modern oceanic crust is also well established (e.g. Kerr 2005). How much carbonation are we talking about? For the sake of simplicity let's assume that the current near-surface carbon inventory is 10^{23} g as in Table 2.3 (Hayes and Waldbauer 2006). Furthermore, let's assume that it was all initially released to the atmosphere as CO_2 (or quickly became CO_2 due to early surface conditions). This is approximately the "standard model". If we require the ocean crust to absorb most of this, and continue to do so through most of the Archean, the amount of carbonate mineral (as $CaCO_3$) required would be about 8×10^{23} g. Some small amount or carbonates and CO_2 would, of course, be found in the atmosphere and ocean, but most would necessarily be stored away from immediate contact with the atmosphere. If this amount of carbonate were disseminated throughout the entire thickness of the ocean crust, the concentration of carbonate in (presumably) altered basalt would have been about 10%, assuming that ocean crust covered a larger fraction of the surface, i.e. total continental area and volume were significantly less than at present. Although this seems like a rather large fraction, it is not impossible. In essence this is the model proposed by Sleep and Zahnle for carbon cycling in the Archean (with some, perhaps much, recarbonation of the upper mantle as well, according to their approach.) However, the more fundamental problem is the cycling of carbon through the system once oxygenic photosynthesis begins. If the result of oxygenic photosynthesis is to convert 15–20% of the "active" carbon pool into organic compounds, with the concurrent production of oxygen, we must be concerned with the rate at which carbon dioxide is evolved and processed into organics starting at about 2.8 BYBP. We can probably assume that convective processes at that time were somewhat faster than at present, perhaps by a factor of three, as a consequence of greater radiogenic heat production at that time. In essence this means that the rate of oceanic crust subduction (regardless of the tectonic details) would necessarily have been about three times the present. However, the greater area of oceans would have meant that the average age of ocean crust would not have been 1/3 as much as that at present, but somewhat more. Combining these two effects we get a cycling time for oceanic crust of about 50 million years. In other words, if equilibrium between mantle and surface carbon reservoirs had been established by 2.8 BYBP, which seems likely, whatever storage could be assigned to ocean crustal (and upper mantle) carbonate would be recycled into the atmosphere on a rather short time scale. The crux is that the overall partitioning of the carbon reservoir between depleted photosynthetic fixed carbon (at 15–20 % of the total) and the remaining carbonate carbon would necessarily become established quite soon (ca. 50 My or so) after the emergence of oxygenic photosynthesis, and much sooner than the rise of atmospheric oxygen some 300–400 million years later. The oxygenation problem cannot be solved by carbonation of oceanic crust.

This somewhat simplistic analysis of the problem does not consider the details of decarbonation at subduction zones, subduction of organic carbon (which probably makes things worse since export of carbon to the mantle actually leaves excess oxygen behind), or processing and release of carbon by mantle magmatism. But these are details which would not change the larger conclusions. The bottom line result is that an early atmosphere rich in carbon dioxide is not consistent with what we know of the history of photosynthesis and the oxidation history of Earth's surface as indicated by carbon isotope records. We are left with the proposal with which this book started, namely that primordial degassing of Earth resulted in strongly reducing conditions at the surface and a large pool of organic compounds.

11.3 The Ultimate Source of Oxidation

If the presently proposed model for Earth's early surface is correct it appears we have created another problem, perhaps more serious than those we claim to have resolved. If degassing led to a large reservoir of abiotic reduced organic compounds in the ocean, and an atmosphere with relatively high concentrations of methane and ammonia, along with nitrogen and carbon dioxide, how could the surface and atmosphere ever become oxidized, as we know it did? This looks even more problematic since the abiotic organic carbon must first be oxidized to carbon dioxide before oxygenic photosynthesis can liberate free oxygen. The answer certainly lies in an alternate approach to the net production of oxidation capacity that has been proposed in detailed form by Catling and Claire (2005) and elaborated on by Zahnle and Catling (2014). The essence of this approach is the loss of hydrogen to space referred to previously as a second source of oxidizing capacity (Chap. 9). Indeed, it is the essential, primary source of oxidation capacity and ultimately of atmospheric oxygen. Oxygenic photosynthesis thus becomes the immediate producer of molecular atmospheric oxygen, rather than being the primary or ultimate source of overall oxidation.

None of these treatments of the oxidation problem (and including e.g., Hayes and Waldbauer 2006) start with the highly reduced surface and large reduced compound pool proposed here. They provide net oxidation from a more or less “neutral” atmosphere similar to that of the standard model. Their main point is that the oxidation necessary to provide for free oxygen in the atmosphere, and presumably also to offset the reducing capacity from weathering of rocks containing ferrous iron and sulfide, is ultimately due to hydrogen loss to space, which also applies to a more reduced starting point. Photosynthesis is an important process within the overall system as a modulator of the oxidation state of the surface, but hydrogen escape is the net oxidizer. Because their approach starts from a CO_2-N_2 atmosphere (and surface environment), it requires only relatively modest hydrogen loss over geologic time, at rates within an order of magnitude of current hydrogen loss, and modest levels of atmospheric hydrogen and hydrogen carriers. In order to produce the much greater amount of oxidation implied by the highly reduced early Earth it is necessary to invoke much greater loss of hydrogen to space. Although limited early hydrogen escape has been suggested by Tian et al. (2005) in their effort to explain reducing conditions necessary for prebiotic organic synthesis, there are reasons to doubt that hydrogen escape rates were dramatically lower during earlier eras (Catling 2006).

There is a reasonably straightforward path to oxidizing the large abiotic reduced carbon (ARC) pool and generating the large amounts of oxidized carbon required. While much of the reduced carbon was probably dissolved in sea water, some would have been deposited in sea-floor sediment, especially following the emergence and weathering of early island arc-like mini-continents. Clays from weathering of emergent silicates would serve as useful carriers of organics to form this sediment, some fraction of which would eventually be subducted back into the mantle. In fact it might well have been easier to transport organics into the magmagenic zone of the mantle than to do so for carbonate. Degassing of carbonates can take place at relatively low temperatures if water pressure is high, as would be likely from dehydration of subducted oceanic crust and sediment. Organic compounds in sediment would experience some degree of dehydration during burial and initial subduction, perhaps to the extent of conversion to elemental carbon:

$$CH_2O = C + H_2O$$

This is a well known process and is largely responsible for the formation of carbonaceous fossil matter. The amorphous or graphitic carbon so produced is relatively inert (and thus readily preserved during subduction) until exposed to magmatic temperatures, at which point reaction with water yields CO_2 and H_2:

$$C + 2H_2O = CO_2 + 2H_2$$

To a first approximation it is this equilibrium under magmatic conditions that is often thought to be responsible for the release of carbon dioxide in the standard model, the hydrogen being released at volcanic vents and subsequently lost (or perhaps processed in part into organic compounds at the surface in some variations).

This set of reactions was probably much more active during the Hadean and into the Archean, when abundant organic material was present for the production of carbonaceous sediments. At present it is probably of minor importance given that deep-sea sediment is typically not rich in organics, and carbonate may now well be the most important form of subducted carbon.

If we assume that the current rate of carbon dioxide emission from all volcanic sources is about 1×10^{14} g/year (see Table 2.3) we can calculate the loss rate of hydrogen necessary to remove the hydrogen produced by oxidation of carbon to carbon dioxide by the oxygen in water. The hydrogen equivalent to this amount of carbon dioxide is about 10^{13} moles of H/year, or about 4.5×10^{10} atoms/cm^2-s. Note that this production rate is about the same as the current rate of hydrogen outgassing from volcanoes ($3.6 +/- 2.6 \times 10^{10}$ hydrogen atoms/cm^2-s, Catling and Claire 2005; Holland 1984; but also note the large uncertainty!). However, very little of this hydrogen reaches the upper atmosphere at present, and that mainly in the form of water vapor; it is rapidly oxidized and trapped. This production rate is much greater than the present-day hydrogen escape rate of 10^8 atoms/cm^2-s (Walker 1977), a consequence of the trapping of hydrogen in the lower atmosphere, which keeps stratospheric levels low (ca a mixing ratio of about 4 ppm, (Walker 1977). Overall hydrogen escape depends on delivery of hydrogen to the exosphere, in turn depending on the concentration of hydrogen containing species in the lower atmosphere, which are presently sparse given the oxidizing power of the free oxygen now present. During the Hadean/Archean there was essentially no free oxygen, which could in principle allow reduced hydrogen-containing species the opportunity to transport much more hydrogen to the upper atmosphere. The limiting rate of hydrogen escape from the Earth has been calculated to be:

$$\varphi = 2.5 \times 10^{13} \, f \text{ atoms/cm}^2\text{-s}$$

Where f is the mixing ratio (essentially the concentration) of hydrogen in the atmosphere. F includes components for atomic, molecular and chemically combined hydrogen, including in water, methane, ammonia, etc. The modern escape rate indicates a mixing ratio for Earth's stratosphere of about 4×10^{-6}, about 4 ppm. The model of the early atmosphere proposed here has higher values for reduced gases in the atmosphere than are sustainable at present. The earlier discussion of the resolution of the Faint Young Sun problem suggested that atmospheric abundance of methane, a powerful greenhouse gas, might have been 300–400 ppm, with similar concentrations of CO_2 needed to suppress the generation of an antigreenhouse haze (Zahnle 1986; Haqq-Misra et al. 2008). The discussion of the mechanism for sustaining these high levels in spite of the known photolytic sensitivity of methane was made earlier and need not be repeated. This level of methane, a gas not subject to condensation and removal from the atmosphere, as is water, would allow transport of much more hydrogen into the exosphere than is presently possible. At the mixing ratio for hydrogen due only to methane at 400 ppm the escape rate from the upper atmosphere would have been:

$$\Phi = 2.5 \times 10^{13} \times 4 \times .0004 = 4 \times 10^{10} \text{ atoms/cm}^2\text{-s}$$

The factor of four is to account for the four atoms of hydrogen in each molecule of methane. This loss rate is very close to that necessary to remove the hydrogen produced in mantle magamatism (4.5×10^{10} atoms/cm^2-s.) It is not intended to suggest that methane is the only or even the main atmospheric component responsible for transporting hydrogen to the exosphere. Water, a small amount of ammonia and perhaps considerable molecular hydrogen (following emission from volcanoes) would also have played a role. The loss rate of hydrogen from the top of the atmosphere of about 4.5×10^{10} atoms/cm^2-s could certainly account for removal of all mantle derived hydrogen produced by converting organic carbon to carbon dioxide and hydrogen during magmatism, leading ultimately to a more oxidized surface environment. The details of atmospheric processing of the various gases are beyond the scope of this discussion. Reactions between the reduced and more oxidized gases in the atmosphere, including Miller-Urey type syntheses, would have a great effect on the exact composition of the atmosphere. As suggested earlier, hydrothermal processing of abundant sea water organics (and organic-rich sediments) would have been a powerful source of methane (and some ammonia) necessary to maintain high atmospheric concentrations. The hydrogen added to the atmosphere by mantle magmatism would simply have added more reducing capacity to the atmospheric mix. In fact an even higher outgassing rate of hydrogen could have been accommodated by potential exospheric escape. It seems likely that the mantle production of hydrogen was rather higher than at present, but there is no way of knowing by how much. A greater amount of subducted organics and greater rates of volcanism would both likely have led to more hydrogen production and emission. Using the modern rates sets a lower limit to hydrogen production (and thus CO_2 production).

Assuming that volcanic carbon dioxide was produced during the Hadean/Archean by processing of organic carbon (CH_2O), and at rates similar to modern volcanic CO_2 emissions, how long would it take to produce the modern mix of oxidized and reduced carbon in the surface inventory? If we use 8×10^{22} gm of carbon in modern carbonate as the required product (Hayes and Waldbauer 2006), and a conversion rate of 1×10^{14} g/year, i.e. the modern rate of CO_2 emission, we get 800 MY. This is certainly a short enough time to allow the surface to reach an oxidized state by the end of the Archean. In fact it appears to be somewhat too fast given the 2000 MY+ available between the end of accretion and the rise in oxygen at 2.3–2.4 BYBP. On the other hand we have not taken into account the reducing capacity of weathered materials or the considerable uncertainties in knowing the actual mixing ratios during the Hadean/Archean. It may also be the case that subduction of organics was rather less at earlier times when weathering products were not as available for adsorbing and carrying organics to the sea floor. Perhaps much of the actual conversion/oxidation was taking place in the near-surface environment through hydrothermal production of methane and carbon dioxide:

$$2CH_2O = CH_4 + CO_2$$

There are simply too many possibilities to permit a precise quantitative reckoning of carbon budgets during the Archean. The best we can hope for is to show that the necessary rates are at least reasonable and achievable under the expected geologic conditions.

11.4 A More Realistic (and More Complex) Carbon Budget for the Archean

Highly detailed carbon budgets have attempted to make sense of carbon isotope signatures in terms of oxidation/reduction conditions for Earth (Hayes and Waldbauer 2006), while others have suggested hydrogen loss as the most important element in controlling surface oxidation and its changes through time (Catling and Claire 2005; Zahnle and Catling 2014). These efforts have largely, if implicitly, adopted the standard model for early degassing, that is, an early atmosphere dominated by nitrogen and carbon dioxide. Hayes and Waldbauer include essentially all possible carbon reservoirs, both oxidized and reduced, as well as conversions and transfers between the reservoirs. The success of these models in describing what we can determine about Earth's early history from geologic data is debatable. In particular, there is the problem of the effects expected for a CO_2-rich early atmosphere. Even if we account for much of the CO_2 by reaction with oceanic crust, there should still be an abundance of carbonates in the earliest continental rocks, which is not the case, and this problem actually persists into the mid to late Archean where, although sedimentary carbonates are present (Grotzinger and Kasting 1993), the rate of cycling of carbonate from ocean crust and upper mantle reservoirs should have led to massive carbonate deposition during that time, which does not seem to be the case. One could certainly argue for recycling of this surface carbonate, and for uncertainties about total amounts, but at some point, and presumably by the mid-Archean, there should have been massive surface carbonate deposition.

Even more difficult to overcome is the delay in surface oxidation following the beginning of oxygenic photosynthesis. Hayes and Waldbauer address this issue, at least partly, by including carbonation of oceanic crust in their model. But here again, the rapidity with which carbonate is recycled to the surface via subduction and heating (Fig. 9.1), in combination with the carbon isotope record (Fig. 11.1), suggests that free oxygen should have been present not long after (geologically speaking) the appearance of oxygenic photosynthesis, as previously discussed. Focusing attention on hydrogen loss as the fundamental source of surface oxidation (Catling and Claire 2005; Zahnle and Catling 2014) does not solve this problem, for the simple reason that from an initial oxidized carbon reservoir (carbonate and CO_2), conversion of the requisite (and available) amount of CO_2 to deposited organic carbon as per the carbon isotope record, is in itself such a powerful source of immediate oxidation that rapid surface oxidation should have resulted, regardless of the ultimate source of oxidation capacity. These problems can be solved by a somewhat more complicated carbon budget, involving an initial large reduced carbon pool, as sketched out in Fig. 11.3.

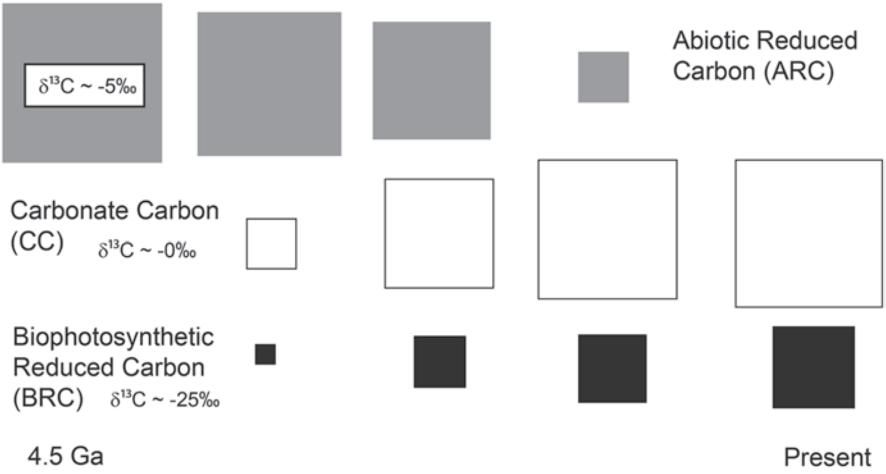


Fig. 11.3 A three component model for carbon reservoir analysis

Figure 11.3 illustrates how a reduced carbon-rich surface can explain the carbon isotope record without encountering problems with the geologic record of surface oxidation. Because the initial carbon pool is in the form of non-photosynthetic reduced carbon (abiotic reduced carbon), its *gradual* oxidation (in the mantle) over geologic time, followed by oxygenic photosynthesis (whenever that may have begun), allows the production of organic and carbonate carbon isotope records without the necessity for rapid production of free oxygen. This suggested modification to carbon cycling and processing should not only open up fertile ground for the carbon cycle modelers, but implies possibilities for modelers of atmospheric chemistry and physics.

The abiotic carbon reservoir dominates early in Earth history, as the main component of the various carbon reservoirs, whether in the atmosphere (certainly a minor amount, but constantly being recycled), dissolved in sea water (initially perhaps the largest quantity of carbonaceous matter), or deposited as sediment on the sea floor (probably initially small, but increasing through time as weathering products become available to carry it to the sea floor.) Cycling of this ARC through hydrothermal and magmatic systems generates carbon dioxide (as it also regenerates methane and ammonia) which gradually accumulates. The rate of this conversion and accumulation is speculative, because it depends upon factors which themselves are speculative, such as rates of sediment deposition and subduction, rates of hydrothermal processing and volcanism, as well as efficiencies of conversion. Nevertheless, early rates were probably higher when mantle heat sources were more abundant. Once continental (initially island-arc-like) areas emerged from beneath the ocean some of this ARC was probably deposited in sediment along their margins, forming an additional reservoir not initially present.

Mantle magmatic production of CO_2 added CO_2 initially to the atmosphere, and dissolved in the ocean. Once a high enough level was reached (probably fairly early) the CO_2 would have enhanced weathering effects, both at/in the ocean crust, perhaps largely at spreading locations, and on the emergent continental fragments. Such weathering would have buffered atmospheric CO_2 (Walker et al. 1981; Gislason et al. 2009), although it is difficult to know at what levels. It is most unlikely that levels of CO_2 were dramatically higher than at present because of the reactivity of silicate rocks with carbonic-rich waters, though much higher levels have been suggested to counter the faint young sun (see earlier discussion). It would not have been long before deposition of carbonates (CC) became possible, both on and within the ocean crust, and eventually in continental areas. This led to additional (oxidized) carbon reservoirs, with carbon isotopic compositions similar to that of ARC, that is, primordial. CC deposited within and on ocean crust would initially have been the main storage site for "excess" CO_2 generated by processing of ARC. In this case "excess" refers to amounts above the level at which weathering becomes rapid. This situation would have continued until the advent of oxygenic photosynthesis. Although it is possible that anoxygenic photosynthesis may have played something of a role a bit earlier, the major change occurs when production of free oxygen becomes possible.

At this point a new component is added to the reservoirs of reduced carbon, namely biophotosynthetic carbon (BRC), which has a distinctly different isotopic signature. Over time the conversion of ARC to CC, and then of CC to BRC reduces the size of the ARC component. However, it takes hundreds of millions of years, perhaps more than a billion years, for earth processes to do this conversion. The BRC produced by photosynthesis would become mixed with unprocessed ARC in each of the reduced carbon reservoirs, which would then be subject to continued reprocessing, both hydrothermal and magmatic. Over time the *average* isotopic signature of the reduced carbon reservoirs would gradually shift from primordial to light. Note, however, that this does not mean we should necessarily expect to see these intermediate organic carbon isotope values reflected in geologic samples. Photosynthesis takes place near the water surface, and especially vigorously where nutrients are available, such as along shorelines. The rapid production of photosynthetic carbon would dominate in just those locations where organic sedimentation is most likely to be preserved in rocks. Note also that photosynthesis changes local water chemistry by consuming carbonic acid, leading to precipitation of calcite, also in a location where preservation is more likely. Open ocean photosynthesis might well have been abundant (though probably less than at present), but the products would largely have been deposited in sea floor sediments destined for reprocessing in a trench setting, and not readily preserved in the geologic record.

The fundamental point in all this discussion is the existence of a surface reservoir for carbon (and the main one initially) that stored the carbon for a long time. The cycling of ARC would probably have been relatively slow, at least compared to a similar sized reservoir of carbonate in ocean crust. ARC was either dissolved in sea water or carried as sediment on the surface of the ocean crust. In these forms most of the ARC would likely remain very near the surface, and not be carried into high temperature zones for reprocessing into CO_2. That would not have been the case for carbonate minerals carried within altered ocean basalt. The model proposed does imply continued processing of ARC, mostly into reduced compounds and largely under hydrothermal conditions at shallow depths. *Some fraction* would certainly have reached the zone of mantle magmatism for conversion to CO_2, as proposed earlier, and some $CO_2$was probably produced during hydrothermal processing, but the main products of processing of ARC were probably the initially reduced gases necessary to maintain atmospheric levels sufficient to counter the Faint Young Sun. I suggest that, as a minimum, modeling of atmospheric/surface evolution involving carbon isotope interpretations should incorporate this more complex picture of carbon components and reservoirs

References

Beversdorf LJ, White AE, Bjorkman KM, Letelier RM, Karl DM (2010) Phosphonate metabolism of *Trichodesmium* IMS101 and the production of greenhouse gases. Limnol Oceanogr 55:1768–1778

Blankenship RE (2013) Molecular mechanisms of photosynthesis. Wiley, New York, p 336

Brocks JJ, Logan GA, Buick R, Summons RE (1999) Archean microfossils and the rise of eucaryotes. Science 285:1033–1036

Broecker WS (1970) A boundary condition on the evolution of atmospheric oxygen. J Geophys Res 75:3553–3557

Capone DG, Zehr JP, Paerl HW, Bergman B, Carpenter EJ (1997) *Trcihodesmium*, a globally significant marine cyanobacterium. Science 276:1221–1229

Catling DC (2006) Comment on “a hydrogen-rich early earth atmosphere”. Science 311:38a

Catling DC, Claire MW (2005) How Earth’s atmosphere evolved to an oxic state: a status report. Earth Planet Sci Lett 237:1–20

Farquhar J, Wing BA (2003) Multiple sulfur isotopes and the evolution of the atmosphere. Earth Planet Sci Lett 213:1–13

Farquhar J, Bao HM, Thiemens M (2000) Atmospheric influence of Earth’s earliest sulfur cycle. Science 289:756–758

Gislason SR, Oelkers EH, Eiriksdottir ES, Kardjilov MI, Gisladottir G, Sigfusson B, Snorrason A, Elefsen S, Hardardottir J, Torssander P, Oskarsson N (2009) Direct evidence of the feedback between climate and weathering. Earth Planet Sci Lett 277:213–222

Grotzinger JP, Kasting JF (1993) New constraints on Precambrian ocean composition. J Geol 101:235–243

Haqq-Misra JD, Domagal-Goldman SD, Kasting PJ, Kasting JF (2008) A revised, hazy methane greenhouse for the Archean Earth. Astrobiology 8:1127–1137

Hayes JM, Waldbauer JR (2006) The carbon cycle and associated redox processes through time. Philos Trans R Soc Lond B Biol Sci 361:931–950

Holland HD (1984) The chemical evolution of the atmosphere and oceans. Princeton University Press, Princeton, p 582

Jones C, Nomosatryo S, Crowe SA, Bjerrum CJ, Canfield DE (2015) Iron oxides, divalent cations, silica, and the early earth phosphorus crisis, Geology, 43:135–138

Kerr AC (2005) La Isla de Gorgona, Colombia: a petrological enigma? Lithos 84:77–101

Martinez A, Osburne MS, Sharma AK, DeLong EF, Chisholm SW (2012) Phosphite utilization by the marine picocyanobacterium *Prochlorococcus* MIT9301. Environ Microbiol 14:1363–1377

Mur LR, Skulberg OM, Utkilen H (1999) Cyanobacteria in the environment. In: Chorus I, Bartram J (eds) Toxic cyanobacteria in water: a guide to their public health consequences, monitoring and management. E. and F. Spon, London, pp 15–40

Nakamura K, Kato Y (2004) Carbonitization of oceanic crust by seafloor hydrothermal activity and its significance as a CO_2 sink in the early Archean. Geochem Cosmochim Acta 68:4595–4618

Sleep NH (2010) The Hadean-Archean environment. In: Deamer D, Szostak JW (eds) The origins of life. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp 35–48

Sleep NH, Zahnle K (2001) Carbon dioxide cycling and implications for climate on ancient Earth. J Geophys Res 106:1373–1399

Van Mooy BAS, Krupke A, Dyhrman ST, Fredricks HF, Frischkorn KR, Ossolinski JE, Repeta DJ, Rauco M, Seewald JD, Sylva SP (2015) Major role of planktonic phosphate reduction in the marine phosphorus redox cycle. Science 348:783–785

Walker JCG (1977) Evolution of the atmosphere. Macmillan, New York, p 318

Walker JGC, Hays PB, Kasting JF (1981) A negative feedback mechanism for the long-term stabilization of Earth's surface temperature. J Geophys Res 86:9776–9782

White AE, Karl DM, Bjorkman KM, Beversdorf LJ, Letelier RM (2010) Production of organic matter by *Trichodesmium* IMS101 as a function of phosphorus source. Limnol Oceanogr 55:1755–1767

Zahnle KF (1986) Photochemistry of methane and the formation of hydrocyanic acid (HCN) in the Earth's early atmosphere. J Geophys Res 91:2819–2834

Zahnle K, Catling D (2014) Waiting for O_2. In: Shaw GH (ed) Earth's early atmosphere and surface environment. Geological Society of America (Special Paper 504), Denver, pp 37–48

Chapter 12
What Can We Learn from Other Planets?

The other terrestrial planets have atmospheres that were also generated by degassing and have undergone evolution over 4.5 BY. Presumably we can learn something about Earth's atmosphere from studying them. At the very least we should expect there to be some consistency in making comparisons. The two planets that have atmospheres most chemically similar to Earth are Mars (surface pressure ca. 0.006 atm) and Venus (surface pressure ca. 90 atm), our near neighbors. (The moon has a very thin atmosphere virtually devoid of nitrogen and carbon compounds, to say nothing of oxygen, a direct consequence of its low mass and thus low escape velocity for light gases.) Although the atmospheres of Venus and Mars are dramatically different, they do share some features. Both have compositions dominated by carbon dioxide and nitrogen, with other minor constituents, such as noble gases. Both have highly oxidized atmospheres, although with little or no free oxygen, and correspondingly oxidized surfaces. This is most apparent for Mars because the iron-stained surface is readily visible through its thin atmosphere. In fact, atomic oxygen, which is highly reactive to reduced carbon species, is a minor constituent of Mars atmosphere. The question that arises immediately is: are the atmospheres of Venus and Mars a direct reflection of primordial degassing or are they the result of whatever processes were active in the evolution of an unknown primordial atmosphere? To a first approximation it appears that their CO_2-rich atmospheres have been taken as support for the standard model for Earth, that is, they are primordial. If, however, the model proposed here for the early Earth is, in fact, correct, it is necessary to consider how primordial reducing atmospheres on both Venus and Mars evolved to the current state. In both cases there is a basic problem (even using the standard model) in that both Venus and Mars have very dry surfaces and atmospheres. Where did the water go? There is certainly no likely place for water on Venus and limited options for Mars, but it is extremely unlikely that there was not a considerable amount of primordial water present.

G. H. Shaw, *Earth's Early Atmosphere and Oceans, and The Origin of Life,*
SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-21972-1_12

12.1 Venus

Walker (1977) certainly recognized the water problem for Venus and resolved it by calculating that hydrogen loss from Venus' atmosphere was sufficient to remove an amount of water (hydrogen, actually) equivalent in mass to Earth's oceans in several 100 million years, and certainly no more than a billion years. He posited that oxidation of rocks was responsible for the uptake of the oxygen so released. Invoking the current model with both water and reduced carbon compounds as primordial actually helps solve the oxygen problem by consuming oxygen as reduced carbon compounds are replaced by carbon dioxide. This implies additional hydrogen loss, but not in quantities that could not be achieved by the previously invoked hydrogen loss from water. The carbon dioxide remains in Venus' present atmosphere, at a level similar to that represented by near-surface carbon compounds on Earth. Nitrogen in Venus' atmosphere is approximately four times the amount in Earth's atmosphere. It is interesting that both carbon and nitrogen are present in approximately equal amounts for the surface regions (atmosphere + oceans + crust) of Earth and Venus. (There is a significant amount of nitrogen incorporated in sedimentary material in Earth's crust, especially in organic-rich deposits.)

12.2 Mars

Mars presents a very different picture from Venus, in some ways perhaps more like Earth. Because we can see Mars' surface we have considerably more information about its history and current state. The Mars landers and orbiters have produced data on composition, topography, surface processes and physical conditions in abundance. Perhaps the most interesting observation is evidence for flowing water. It is not only the enormous Valles Marineris, but there are clear indications of integrated stream networks and examples of likely river channels and deposits (Fig. 12.1). Estimates of ages of Martian surfaces using crater densities indicate that essentially all of this activity ceased prior to 3.8 BYBP. Since that time little or no water flow has occurred. Both higher temperatures and atmospheric pressures than at present would have been required for water flows of this magnitude and duration. The change from an early warm, wet Mars to a cold, dry Mars implies a climate transition early in Martian history; a change to colder conditions occurred even as solar luminosity increased. The early era of abundant liquid water suggests the possibility of life beginning on Mars as it did on Earth, and this is a powerful incentive behind the efforts to search for evidence of life using the surface rovers. As yet this has proven unsuccessful, but the search continues.

Leaving the question of ancient water flows for the moment, it is instructive to consider the inventories of atmospheric volatiles and condensed water on the terrestrial planets (excluding Mercury). The composition of Mars' atmosphere is surprisingly similar to both Earth and Venus (Table 12.1). The CO_2/N_2 ratio is essentially

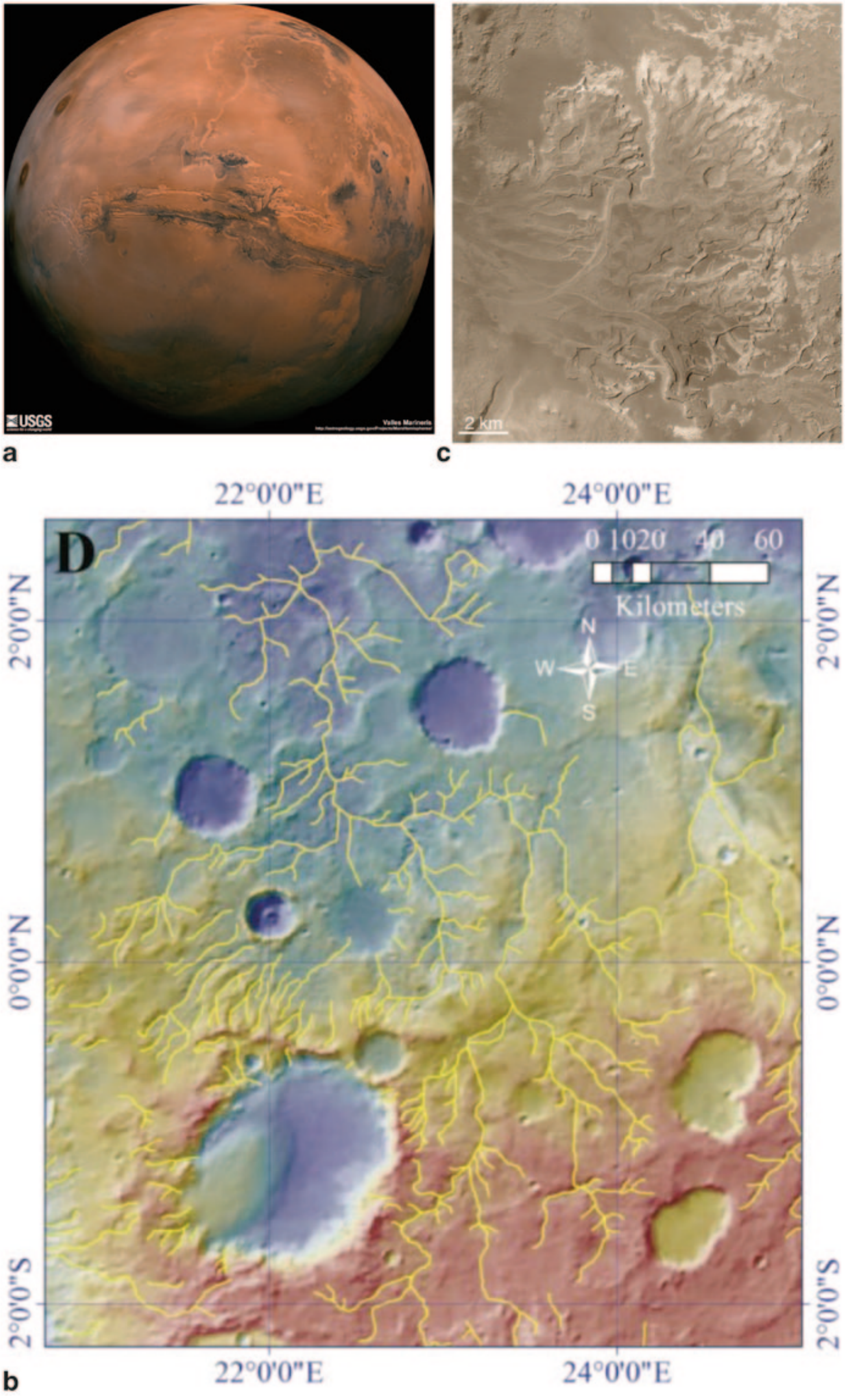


Fig. 12.1 Evidence for flowing water on Mars. **a** Valles Marineris (USGS). **b** Integrated stream networks (with permission, Hynek and Phillips, 4063 Geology, 31, 2003). **c** Deltaic, and stream deposits (MGS MOC Release No. MOC2-1225, 20 4069 September 2005, NASA/JPL/Malin Space Science Systems)

Table 12.1 Comparison of the major component compositions of atmospheres (surface volatiles) of Venus, Earth, and Mars. In the case of Earth the volatiles stored in rocks are included in the ([a]) column. Comparison of surface water on Venus, Earth and Mars

	Venus	Earth[a]	Mars	Earth (atmosphere only)
Mass (gm) (C as CO_2)	4.8×10^{23}	3.7×10^{23}	2.5×10^{19}	5.1×10^{21}
CO_2	96.5 %	98.7 %	95.3 %	300 ppm
N_2	3.5 %	1.3 %	2.7 %	78 %
Other	traces		2.0 %	22 %
Water (mass in gm)	minimal	1.55×10^{24}	5.0×10^{21}	
Water/total volatiles	~0	4.2	~200	

[a] Including volatiles in crustal rocks, which dominate

the same for each, with Earth having a somewhat lower proportion of N_2. Perhaps this is due to an underestimate of nitrogen contents of organic sediments on Earth, but in any case the parallels are striking. Comparing these figures with the availability of these volatiles from a meteorite mix is also striking (Table 2.4). Can it simply be chance that all three terrestrial planets with significant atmospheres have CO_2/N_2 ratios comparable to "average" meteorite? More likely is the notion that the source of volatiles in all three cases was similar to a common late veneer strongly resembling an average meteorite mix.

Looking at surface water suggests both similarities and one significant difference. Venus doesn't enter into this comparison because of the (probable) profound loss of water (hydrogen) from Venus' hot atmosphere. Table 12.1 shows that $CO_2 + N_2$ amounts to about 1/5 of the water for Earth, but 1/200 for Mars. It appears that there is either an excess of water on Mars' surface or a deficiency of CO_2 and N_2. This discrepancy might actually be somewhat greater considering the relative ease with which hydrogen can escape from the lower gravity of Mars. Three possibilities might explain this observation.

1. Mars atmosphere may have suffered a higher loss of CO_2 compared to water, accounting for both the present comparative excess water on Mars' surface as well as the climate transition mentioned above. This differential loss might have resulted from freezing out of atmospheric water, while leaving CO_2 in the atmosphere, resulting in relative mixing ratios of CO_2 and H favoring CO_2 loss. This seems like rather a stretch and in view of the measured CO_2 loss from Mars at present (Barabash et al. 2007), clearly not high enough to account for the current deficiency of CO_2.
2. Another possibility is that the Martian degassing process may have favored water release compared to CO_2 and N_2. Because Mars has about 1/10 the mass of Earth, we might expect as a first approximation that the total amount of chemically bound volatiles available for degassing would, overall, be less for Mars. There is no compelling reason to believe this, assuming that most volatiles came from degassing of a late veneer, but perhaps Mars had a smaller late veneer. Table 12.1

does suggest that volatile evolution from Mars' interior has, in fact, been less effective (or from a more limited source reservoir) than for Earth. Mars' surface inventory of water is about 1/300 that of Earth, although the planetary mass ratio is 1/10. This suggests a deficiency of surface water on Mars compared to Earth, by a factor of 30 or so. The planetary mass difference may affect degassing in a different way. Impact heating for Mars would have been substantially less than for Earth. The lower initial heating also provides less energy to drive degassing, and may well have been low enough to avoid initiation of "plate tectonics" or similar large-scale convective heat loss. A larger surface/volume ratio would also favor heat loss from Mars by conduction rather than convection, at least as a planet-wide mechanism. These factors would limit the volume of material subjected to high temperatures required for degassing. The Martian volcanoes certainly indicate enough internal heat production to drive convective flow (which is what volcanism is, to a first approximation), and such volcanoes would certainly have emitted volatiles to the surface. However, the mass of material affected by volcanism and related hydrothermal activity would have been a fraction of the total of the late veneer. This may explain the overall lower surface volatiles for Mars, but what about the CO_2—water differential? Perhaps this may be explained by a greater ease of removing water by dehydration reactions from accreted silicates compared to oxidation or other chemical processing of carbon compounds in the original mix undergoing degassing during volcanism. A lack of subduction on Mars would also limit the amount of oxidative reprocessing of reduced carbon compounds under magmatic (i.e. CO_2-producing) conditions. This combination seems somewhat contrived, but cannot be excluded.

3. Finally, there may be a near-surface reservoir containing the missing C-N compounds. Such a reservoir could easily be hidden beneath a thin sediment/dust cover on the Northern Plains. An approach to understanding this possibility brings us back to a look at the early Earth. Suppose degassing of Mars was similar to that proposed for the early Earth, with water and reduced gases such as methane and ammonia, as well as nitrogen and some carbon dioxide (in gradually increasing amounts) over time.

If accretion of Mars was similar to what we surmise for Earth, the accreting material would have included mostly reduced carbon compounds and considerable water as hydrous compounds. Given its smaller size, the energetics of accretion imply lower energy impacts and thus less impact heating. Such conditions favor preservation of reduced compounds. This also implies lower internal temperatures from impact heating and core formation. However, internal temperatures were obviously enough to generate basaltic magmatism, which would have driven degassing. The high temperature magmatic degassing may have produced volcanic emissions of water, nitrogen, and oxidized carbon compounds such as CO_2 and CO, but it would also have provided an energy source for near-surface hydrothermal activity and thus low temperature processing of late accreting meteoritic material, which would yield reduced carbon compounds as well as water. On this basis, the early Martian surface could have had an atmosphere and a small ocean containing reduced carbon compounds with some amount of carbon dioxide. The amount of nitrogen in the atmosphere may have been enough to produce surface pressures high enough to stabilize liquid water (more on this later). The volcanic centers we observe today as inactive were certainly active early in Martian history, producing both volcanic emissions and driving reprocessing of surface organics through associated shallow and more widespread hydrothermal systems. This reprocessing was responsible for maintaining atmospheric levels of methane and ammonia, offsetting their photolytic destruction, sufficient to counteract the FYS. This is essentially the same picture as presented above for Earth.

A more or less direct parallel, quantitatively, would suggest an early Martian ocean with about 10–20 % organic compounds dissolved in it, under an atmosphere with small but significant amounts of methane and ammonia regenerated hydrothermally around the margins of the large volcanic centers. Atmospheric pressure could well have been high enough to stabilize liquid water, while the powerful greenhouse gases warmed the surface to produce a warm, wet Mars. Loss of hydrogen to space would convert methane to CO_2 and NH_3 to N_2 over time, but continued hydrothermal processing of the organic-rich ocean would maintain levels of the reduced greenhouse gases sufficient for climate moderation.

This situation would have continued as long as hydrothermal activity was able to maintain enough reduced gases in the atmosphere to yield the necessary warming. However, heat loss from Mars' interior would have been much more rapid than for Earth because of Mar's higher surface/volume ratio, and at some point replenishment of near-surface heat by the volcanic centers would be insufficient to support enough hydrothermal activity to prevent the transition to cold conditions. This apparently happened about 3.8 BYBP. By that time, some fraction of the greenhouse effect warming Mars was being produced by CO_2 and water vapor, as it still is for Earth. A decrease in global temperatures on Mars due to decrease in atmospheric methane and ammonia would have been felt especially in circumpolar regions. Because Mars' ocean at that time would have been located where the Northern Plains are currently, this global ocean would have experienced steady cooling. This implies a feedback accelerating the cooling process, driven by condensation of water vapor from the cooling atmosphere (globally), and a loss of important greenhouse gases from the atmosphere as they dissolved in the cooling ocean. This latter effect would have been particularly great for CO_2 and ammonia, both highly soluble in water, and increasingly so at lower temperatures.

When the Northern Ocean began to freeze at the surface the dissolved compounds became more concentrated, and some may have precipitated on the ocean floor. Eventually the ocean would freeze entirely, not only trapping organic sediment on the ocean floor, but trapping the previously dissolved gases beneath the steadily thickening ice. During the earlier phases of the freezing process, milder conditions away from the pole would have allowed continued rainfall and sediment transport to the southern edges of the ocean. Gradually the sediment would extend northward to bury previously deposited organic sediment under mineral sediment. It is difficult to know how long this may have continued before climate deterioration largely "turned off" temperate zone precipitation and the Northern Ocean was left as a frozen ice plain. This was the beginning of the Northern Plains.

Airborne dust and other fine sediment continued to be added to the surface of the ice plains. At the same time, evaporation of water, especially from the southern margins of the frozen terrane, would snow out near the pole, gradually producing the modern ice cap. As this continued some of the trapped gases were released, but they would not likely have been enough to reverse the cooling trend. Organic deposits may have been locally and partially exposed, but oxidation of most of the buried sediment by the atmosphere would have been limited, and would have more or less ceased with burial beneath the encroaching "dust" cover. The final result is what we see on Mars at present: a sediment/dust covered Northern Plain surrounding a thick ice cap seasonally mantled with a thin layer of CO_2. An ice cap would also have formed (and currently exists) at the South Pole, but without the surrounding low-lying plain. If this scenario is correct most of the organic compounds (comprised of CHNO in various combinations) remain buried beneath the Northern Plains, including under the north polar ice cap. If the analogy with Earth is approximately correct, the original ocean may have contained the majority of the originally degassed carbon and nitrogen, which is still trapped beneath the surface cover, isolated from oxidation. The amount of such a deposit can be readily estimated. Mars surface water inventory is about 5×10^{21} gm, compared to 1.55×10^{24} gm for Earth (Table 12.1). For Earth, the volatile inventory of C and N is about 1×10^{23} gm, with carbon as pure C (Table 12.1 shows C as CO_2, which gives a higher number). The equivalent for Mars would be about 3×10^{20} gm. This compares to 2.5×10^{19} gm of $CO_2 + N_2$ at Mars' surface at present, and implies that most (i.e., about ten times the known inventory) of Mars' original inventory of these volatiles is hidden. For simplicity, assume that this amounts to about 5×10^{20} gm as CHNO compounds, and these compounds are about the same density as water. They would then be about 1/10 the volume of Mars current total water inventory, which itself approximately amounts to a 35 m layer over of water the entire surface of Mars. Limited to just the Northern Plains, this would be a water layer of about 100 m, and the organics would be equivalent to a layer of about 10 m. The thickness of sediments covering the Northern Plains has been estimated at about 100 m (Head et al. 2002), which would easily conceal a deeper 10 m layer of organics as well as 100 m of frozen ocean. It may also be that the organic sediments are thicker and more concentrated under the center of the basin, but that is not essential to their preservation. It is also possible that they are closer to the surface (not as deeply buried) under parts of the basin, perhaps especially toward the southern margins. The occurrence of organic deposits, though probably thinner, in several smaller basins (and larger craters) across Mars is not precluded. It appears that the discrepancy in the last line of Table 12.1, i.e. the excess of water over other volatiles, can be accounted for using a model for Mars' early surface environment similar to that proposed for Earth.

In the process of addressing the volatile balance of Mars we have also seen the solution for the problem of climate change on Mars. The decrease in volcanic/hydrothermal heating necessary to drive regeneration of reduced atmospheric greenhouse gases eventually reduced their resupply below the point of maintaining warm, wet conditions. This was in spite of increasing solar luminosity, a process slow compared to changes in levels of internal heat available to drive volcanism (convective heat flow).

An additional factor in considering the climate transition on Mars is the current paucity of atmospheric nitrogen, probably initially necessary in order to generate atmospheric pressures high enough to stabilize liquid water at the surface. Considering the relative inertness of nitrogen and the difficulty inherent in removing significant nitrogen from the atmosphere by exogenic processes, such as the solar wind, how would it have been possible for Mars to have had enough atmospheric nitrogen to raise atmospheric pressure into the stability field of liquid water? It would, of course be possible to do this by postulating much higher amounts of carbon dioxide necessary to overcome the FYS, but this leads right back to geochemical problems inherent in the lack of abundant carbonates (Bibring et al. 2006; Ehlmann et al. 2008) implicit in a warm, wet, CO_2-rich early Martian surface, similar to the problem with a lack of ancient carbonates on Earth. Other greenhouse gas candidates, such as methane and especially ammonia, as per the model proposed here to account for a mild early Mars, if present in sufficient amounts to overcome the FYS, would have been at levels very marginal to providing atmospheric pressure high enough to stabilize liquid water, at best. The only candidate gas that could solve this problem is, in fact, nitrogen. But we are left with the question: Where did the nitrogen go? The answer may lie in further consideration of the evolution of the early reducing atmosphere of Mars into the strongly oxidizing atmosphere of the present.

Consider again the processes by which the early reducing atmosphere and mild climate was initially maintained for ca 1 BY, and then transformed into a mostly CO_2 atmosphere and a cold, oxidized surface. Remember that this happened in spite of a 20–25 % increase in solar luminosity from 4.5 BYBP to the present. The early carbon cycle on Mars has been described above, with some attention given to the production of ammonia as part of the cycling of carbon compounds. If we further examine the nitrogen cycle on early Mars it becomes apparent what may well have happened to the atmospheric nitrogen. The necessary greenhouse gases according to this model are primarily methane and ammonia, but with some carbon dioxide as well. The amount of carbon dioxide, while much less than necessary to produce the requisite greenhouse effect on its own, probably provided a mix of atmospheric gases sufficient to prevent an anti-greenhouse effect through the photochemical generation of an organic-rich haze (Zahnle 1986; Haqq-Misra et al. 2008). Both of the major greenhouse gases on the early Mars (CH_4 and NH_3) were photochemically unstable, as we know. They were only maintained at sufficient levels by processes that could regenerate them, in particular hydrothermal processing of organic residues rained out of the warm wet, organic-rich atmosphere into an early Martian ocean. This hydrothermal regeneration is presumed to have been sufficient during Mars early more volcanically active era to keep atmospheric levels high enough for climate moderation. The carbon cycle looked much like that in Fig. 9.1 (less the subduction), basically conversion of atmospheric methane to heavier organics ("haze"), rainout into surface waters of these compounds, hydrothermal processing of organic-rich "sea water" and/or sediments at/near volcanic centers, and reinjection of the hydrothermal methane back into the atmosphere. Carbon dioxide in the atmosphere was an additional carbon reservoir in the atmosphere and some carbonate was presumably also present in solution (or precipitated) in the oceans, though it must have been minor given the lack of significant preserved carbonate minerals at present. A similar, and intimately connected, cycle exists for nitrogen, with the added factor of an N_2 reservoir in the atmosphere. This was probably the major atmospheric gas at the time, providing enough atmospheric mass to stabilize liquid water at the surface. It was not, however, an inert component of the atmosphere, but participated in the photochemical processing which produced heavier organic compounds, including nitrogen containing compounds.

There is no good way to know which reservoir contained the majority of the nitrogen, or exactly how much was present, but it seems likely that the oceans contained significant quantities of photochemically produced nitrogenous organics. The climate transition on Mars occurred when hydrothermal activity declined to the point that insufficient methane and ammonia were generated and released into the atmosphere to compensate for low solar output, in spite of the slow (too slow) increase in solar luminosity. The transition may have involved some warm/wet-cold/dry alternation, but was probably fairly sudden on a geological time scale. Methane, however, because of its lower solubility in water, and its low reactivity in aqueous solution, would still have been released to the atmosphere (though more slowly than previously). The drastically decreased atmospheric CO_2 (it having dissolved in the cooling Martian ocean) and the continued presence of atmospheric methane would have led to an antigreenhouse effect, yet further adding to cooling, but even more significantly to photochemical processing of atmospheric nitrogen into nitrogenous organics (haze). Over an extended period the rainout of organics into the (freezing) ocean would extract nitrogen from the atmosphere. Because hydrothermal ammonia was trapped in the cold ocean, there would have been no replenishment of free nitrogen from released ammonia, and atmospheric nitrogen would have fallen over time, eventually to the point where liquid water was no longer stable at the surface. Naturally, the cooling of the planet made liquid water even more problematical, except in areas where hydrothermal activity raised local temperatures. The final state is the present atmosphere and surface. Continued hydrogen loss from the breakdown of water vapor leads to effective oxidation of the dwindling supply of methane, with production of some CO_2, and eventually the oxidation of the surface. At present the atmosphere consists of mostly CO_2 and a small amount of residual nitrogen.

Life on Mars?

If Mars first billion or so years was warm and wet because of a large reservoir of organic compounds in the ocean, recycled continuously into the atmosphere as CH_4 and ammonia, the question naturally arises of why life did not emerge there, as it did on Earth. (Of course this is considered an open question by many, but for the sake of argument I assume that life is NOT present on Mars, and never has been.) If one assumes that the presence of life on Earth is evidence for a reasonably high probability of prebiotic chemistry transitioning into a self-sustaining system (e.g. life), especially considering how quickly life seems to have emerged on Earth (some estimates place the origin of life as little as a few million years or less after the establishment of a stable “climate”(de Duve 2005)), it would be most surprising if the conditions envisioned for the warm, wet early Mars did not also give rise to something recognizable as life in Mars first billion years or so.

A possible answer to this conundrum may tell us something about the early Earth and the conditions necessary for life to arise. There are two differences between Earth and Mars, either or both of which could be factors working against the emergence of life there. First, Mars probably has not experienced plate tectonics, limiting the conditions under which an ocean basin would form. On Earth this is related to plate generation and consumption, and leads to production of extensive complex coastal environments via plate convergence and island arc generation. The huge lengths of coastline that were almost certainly present even on the early Earth when mini-continents (basically numerous island arcs) first emerged, provided opportunities for isolation (and concentration) of organic-rich sea water in tide-pool, lagoonal, etc, settings. These individual pools would have provided a vast number of small test-beds for trying various combinations of precursor "soups". On Mars, the main ocean basin, the Northern Plains, is probably the result of a large late impact (Marinova et al. 2008). As such, it is relatively simple in structure, with a limited length of coast, which is not particularly convoluted. As a consequence, the number of potential "test reactors" would have been much smaller than on Earth. This effect is further accentuated by the much smaller surface (and ocean surface) area on Mars. An estimate of the number of "test pools" that might have existed on Earth and Mars may make this clearer. If we assume that by about 4.1 BYBP (about 10 % into Earth's history) 20 % of Earth's current continental crustal volume (and area) had been produced, there might have been the equivalent of about 80 + Japan's, with total coastline of nearly 2.4 million km. By contrast, the approximate shoreline of Mars' Northern Plains would probably have been not much more than 5–10 Martian circumferences or so, that is about 100–200 thousand km. It is difficult to know where the boundary of a putative early Martian Ocean may have been, but would certainly have been no more than the edge of the Northern Plains. It is possible that lakes in smaller craters may have contributed shore length, but this additional length may have been less useful for reasons discussed next.

The other major difference is the drastically smaller tidal range of the Martian ocean. This is not only due to the lack of a significant satellite, but also to Mars' greater distance from the sun. The solar tide is sometimes ignored in discussions of the importance of the Moon to Earth, but it certainly does factor significantly into Earth's ocean tides (about 1/3 of the total). However, because of the $1/R^3$ dependence of tidal forces, the solar tide on Mars would have been much smaller than that on Earth. As a result, the expected tidal range on Mars should be about 1/7 that of tides on Earth, or on average about 13 cm. Considering the location of the Martian ocean in a circumpolar region, which implies even lower tidal range, the importance of tides in producing and maintaining "test reactors" would have been minimal. In other words, not only would Mars have fewer locations where "warm, wet pools" might exist, the pools themselves would be less likely to develop conditions (concentration of organics, temperature, volume) conducive to the emergence of life. Tidal ranges in even the largest equatorial crater lakes would have been trivial.

In addition to these features, we should consider the lower solar insolation available at Mars for driving chemical reactions both within pools and on drying surfaces adjacent to them (again, compared to Earth) as well as the generally lower availability of thermal energy in either hot spring or lava flow settings. Taking these factors into account, one would expect it would take life much longer to arise on Mars than on Earth, and perhaps (probably, I think) Mars turned cold and dry before such an event occurred, some 1 BY or less into Mars' history.

References

Barabash S, Fedorov A, Lundin R, Sauvaud J-A (2007) Martian atmosphere erosion rates. Science 315:501–503

Bibring J-P, Langevin Y, Mustard JF, Poulet F, Arvidson R, Gendrin A, Gondet B, Mangold N, Pinet P, Forget F, The OMEGA Team (2006) Global mineralogical and aqueous Mars history derived from OMEGA/Mars express data. Science 312(5772):400–404

De Duve C (2005) Singularities: landmarks on the pathways of life. Cambridge University Press, Cambridge, p 258

Ehlmann BL, Mustard JF, Murchie SL, Poulet F, Bishop JL, Brown AJ, Calvin WM, Clark RN, Des Marais DJ, Milliken RE, Roach LH, Roush TL, Swayze GA, Wray JJ (2008) Orbital identification of carbonate-bearing rocks on Mars. Science 322:1828–1832

Haqq-Misra JD, Domagal-Goldman SD, Kasting PJ, Kasting JF (2008) A revised, hazy methane greenhouse for the Archean Earth. Astrobiology 8:1127–1137

Head JW III, Kreslavsky MA, Pratt S (2002) Northern lowlands of Mars: evidence for widespread volcanic flooding and tectonic deformation inthe Hesperian period. J Geophys Res 107(E1):5003. doi:10.1029/2000JE001445

Marinova MM, Aharonson O, Asphaug E (2008) Mega-impact formation of the Mars hemispheric dichotomy. Nature 453:1216–1219

Walker JCG (1977) Evolution of the atmosphere. Macmillan, NY, p 318

Zahnle KF (1986) Photochemistry of methane and the formation of hydrocyanic acid (HCN) in the Earth's early atmosphere. J Geophys Res 91:2819–2834

Chapter 13
Summary

> Where clear evidence is lacking, it has been necessary to guess what might have happened. Much of this account of atmospheric evolution is therefore speculative. Many of my guesses may be wrong, but if my version of atmospheric history stimulates the development of more accurate histories I shall be well pleased. (Walker 1977)

I can only agree once again with this statement by Walker. I have embraced a model for the production and evolution of Earth's atmosphere that is quite divergent from a long established view. It is not original, but hearkens back to an earlier era when there was much more limited information available to constrain the possibilities. The changes that took place during the 1950s led to a new standard model that has held sway for more than six decades. We now have even more data to direct and limit our efforts toward understanding a vital part of Earth's history, and the nature of the surface conditions that governed the emergence of life and all subsequent surface geohistory. I believe what I have presented in this little book provides a more coherent picture of the early Earth (and Mars) than the standard model. I have addressed several pieces of geological and planetological evidence that are problematic when trying to apply the standard model. As in any model, it has been necessary to make some assumptions, some with more justification than others, but I believe my assumptions are no worse (and often better) than many that have been used in applying the standard model. There is little in this approach that lends itself to detailed calculations or the sophisticated computation so often a part of modern model-building. On the other hand, I would be the first to agree with a previous commentary (there focused strictly on computer modeling) and apply it to my own somewhat different efforts here as well (Shaw 1989). We should always be willing to take our results with an appropriate amount of salt, at the very least in the name of properly regulating blood pressure. Believing what I have written here does not reduce my obligation to scientific skepticism, especially of my own ideas.

G. H. Shaw, *Earth's Early Atmosphere and Oceans, and The Origin of Life*,
SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-21972-1_13

References

Shaw GH (1989) Danger warrants rescuing data. EOS 70:195
Walker JCG (1977) Evolution of the atmosphere. Macmillan, New York, pp. 318

Index

G. H. Shaw, *Earth's Early Atmosphere and Oceans, and The Origin of Life,*
SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-21972-1