The rise of continents—An essay on the geologic consequences of photosynthesis

Minik T. Rosing a,⁎, Dennis K. Bird b, Norman H. Sleep c, William Glassley d, Francis Albarede e

a Geological Museum and Nordic Center for Earth Evolution, University of Copenhagen, Øster Voldgade 5-7, DK-1350 København K, Denmark
b Department of Geological and Environmental Sciences, Stanford University, Stanford, CA 94305, USA
c Department of Geophysics, Stanford University, Stanford, CA 94305, USA
d Lawrence Livermore National Laboratory, Livermore, CA 94550, USA
e Ecole Normale Superieure de Lyon, Lyon, France

Received 26 April 2005; received in revised form 31 December 2005; accepted 7 January 2006

Abstract

Earth accreted 4567 Myr ago from largely homogeneous material. From this initial capital of matter, differentiation formed the chemical and physical compartments of core, mantle, continents, ocean and atmosphere, that characterize Earth today. Differentiation was, and still is, driven by energy from various sources including radioactive heat and relic heat from accretion. With evolution of photosynthesis, living organisms acquired the ability to harvest Solar energy and channel it into geochemical cycles. On our present Earth, the primary production from life contributes 3 times more energy to these cycles than Earth’s internal heat engine. We hypothesize that the emergence of this energy resource modified Earth’s geochemical cycles and ultimately stimulated the production of granite during the earliest Archaean, which led to the first stabilization of continents on Earth. Such biological forcing may explain the unique presence of granite on Earth, and why stable continents did not form during the first half billion years of Earth’s history.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Photosynthesis; Chlorophyll; Continents; Archaean; Granite

1. Introduction

This essay explores geologic consequences of interaction between early planet Earth and its emerging life. We develop the hypothesis that the rise of the continents is a consequence of biological forcing of weathering and diagenetic rock alteration, and that the power needed to drive this process is obtained via photosynthetic energy fixation. Our argument in support of this hypothesis is outlined in Fig. 1. It is based on the argument that biological activity ultimately performs work on the mineral substrate, the consequences of which are weathering and diagenetic alteration. Since granitic melts are generated via dehydration reactions and partial melting of subducted altered basalt crust and sediments, biological forcing of the near surface rock alteration cycle is likely to accelerate granite production, which in turn would lead to the stabilization of continents. This mechanism provides an explanation...
Fig. 1. Cartoon showing some important developments in Earth evolution.

4567 Myr ago Earth accreted from material similar to chondrite meteorites. The metal core segregated shortly after.

During the following 850 Myr (1 1/2 Phanerozoic time span) Earth did not produce any lasting geologic constructions.

Which allows the possibility that life may have emerged 4400 Myr ago.

Oceans most probably formed on Earth’s surface 4400 Myr ago.

3800 Myr ago life flourished in the oceans and photosynthesis was probably invented.

3800 Myr ago continents became stable enough to preserve bits and pieces to our time.

The stabilization of the continents marks the beginning of the geologic record. Some familiar environments are recorded right from the beginning.

87 mW/m²

Earth’s plate tectonic machinery is driven by the internal heat production which provides 87 mW for every m² of Earth’s surface. In comparison sunlight contributes 340 W/m² to Earth surface.

268 mW/m²

Chlorophylls capture 268 mW/m² of the Solar energy and convert it into chemical energy, which help fuel Earth’s geochemical cycles.

340 mW/m²

The forcing of the geochemical cycles by photosynthetic life could have accelerated granite production and led to the stabilization of the continents.
for the apparent lapse of 600–800 Myr between the accretion of Earth and the onset of stable continent formation (Nutman et al., 1996; Bowring and Williams, 1999; Boyet et al., 2003; Rino et al., 2004), and the coincidence that the oldest vestiges of Earth’s crust carry signatures of advanced and probably photosynthesizing life (Buick, 1992; Rosing and Frei, 2004). Continents thus provide a biomarker of this metabolic strategy on silicate planets.

Although difficult to test in detail, basic thermodynamic considerations suggest that this hypothesis is sufficiently consistent with available observations and data to warrant further examination. The analysis presented places first order thermodynamic constraints on the extent and metabolic activity of a biosphere where primary production is purely chemoautotrophic, noting that the energy budget of our planet places strict constraints on fluxes of basic components required for chemoautotrophic life. These relationships indicate that life is unable to influence Earth’s carbon cycle in any significant way in the absence of photosynthesis. We therefore suggest that evidence for extensive bioactivity and management of the carbon cycle by life through the 3800 Myr of geologic record (Schidlowski, 1988; Veizer et al., 1992; Rosing, 1999; Tice and Lowe, 2004) is also evidence for the operation of photosynthesis as far back in time as we are able to see through the rock record.

We develop the argument by first considering differences between physical and biological energy balances and fluxes in the Earth and their respective consequences for earth’s chemical structure. We then consider evidence regarding the rise of continents, and granite genesis. Within that context, the energy and mass constraints on chemautotrophy and photosynthesis are contrasted, and implications for the energetics of granite genesis are considered. Finally, the implications of the record of life for this hypothesis are explored.

2. Energy flow and the compartmentalization of Earth

Chemical differentiation of Earth into compartments is a consequence of the transport of thermal energy. The sequestration of chemical components from a once homogeneous system into separate domains results in a loss of entropy, which is possible because differentiation occurs in conjunction with a greater production of entropy associated with the decay of thermal gradients (heat flow). Chemical differentiation of Earth can be seen as a way of reducing the rate of global entropy production during dissipation of heat, in accord with the principles for thermodynamics of irreversible processes (Prigogine, 1967).

Heat and mass fluxes are controlled by potential differences modified by impeding factors, such as friction, viscosity, permeability or diffusivity. The effect (equivalent to the amount of work done) is proportional to the flux. When we wish to evaluate the global effect of geologic processes over geological timescales, our foremost objective is to quantify fluxes. When first quantified, information can be extracted regarding driving forces and the effectiveness of the impeding factors. Because heat and mass reservoirs are finite, the gradients that drive fluxes decay as a result of these fluxes, and thus the size of the reservoirs ultimately control the integrated flux of any component. In practice, Earth is characterized by a very large number of interlinked systems, in which the abundance of any component in a given reservoir is controlled by the rate of replenishing fluxes into the reservoir and the impeding factors limiting fluxes out of the reservoir.

The main driving force for geological processes in the deep Earth is the flux of thermal energy from its interior to space. The heat flux is controlled by physical properties of Earth and by internal heat production from radioactive decay and fossil heat left over from accretion. Most heat transport within Earth is performed by advection—that is matter moving with its heat content in response to buoyancy forces, which are set up by thermal expansion and phase reactions. The geologic expression of this advective heat transfer is plate tectonics in its broadest sense, including magmatism, sea floor spreading and orogenesis. Lateral forces in tectonism are derived by diversion of vertical forces associated with buoyancy. Erosion and sedimentation, which occur in response to topographic relief developed through tectonism are thus also an expression of Earth’s thermal gradient. Hence, plate tectonism is the main operator in driving mass fluxes on Earth.

In the same way that the flux of thermal energy results in mechanical forces and mechanical work, thermal gradients can establish chemical potential contrasts. Such contrasts arise because the chemical potential of a component of a phase assemblage is temperature dependent. Chemical potential gradients result in mass transfer that is often selective, leading to fractionation of elements and isotopes and differentiation of Earth into reservoirs with distinct chemical compositions, such as the atmosphere, oceans, continents and depleted mantle. The rate of geologic construction and destruction are determined by the heat flux, which can be relatively well characterized at present, and modeled in the past through simple
algorithms for heat production. In Earth’s near surface environments, an additional source of energy to drive mass flux is solar energy. Mechanical work is performed by convection of the atmosphere and hydrosphere, and chemical energy may be added via biological activity. Energy dissipation in the near surface environment is enhanced by the contribution of Gibbs energy from photosynthesis.

An aspect of mass transport important for consideration of continent formation and stabilization is weathering and low-grade alteration of basalt. Both are consequences of disequilibria between the atmosphere and hydrosphere, and the mineralogy of basaltic crust. Mass transfer reactions will tend to approach a state of chemical equilibrium between these reservoirs, decreasing the affinity that drives the reactions responsible for weathering and alteration. In the absence of a persistent and effective energy source, weathering and alteration rates would tend toward zero as the oceans approach saturation with respect to the minerals of the crust. Photosynthesis, on the other hand, maintains a chemical state on the surface of the planet that is far from equilibrium. A photosynthesizing biosphere is thus able to maintain an ocean–atmosphere system that is persistently out of equilibrium with the crust. The result is a forcing of the weathering cycle that achieves a higher rate of production of weathering products, and thus chemical fractionation of the crust, than would otherwise be possible. Particularly important is the production of smectite and illite clays during weathering of basalt, both of which may play a role in granite formation, because they act as alkali exchange media promoting a strong fractionation of K from seawater into the sediment reservoir.

3. The rise of the continents

Earth’s crust is composed of two distinct compartments: the continents and the ocean floor. Neither of these dates back to the origin of the planet. The oldest tiny vestiges of crust on Earth are ca 4000Myr old and granitic in composition (Bowring and Williams, 1999; Nutman et al., 2000). This is in contrast to the Moon and Mars, which both preserve extensive areas of >4400Myr protocrust (Carlson and Lugmair, 1988; Zuber, 2001). On these heavenly bodies, the ancient crust is extensively cratered by meteorite impacts. It has been speculated that Earth’s protocrust was destroyed by meteorite impacts (Priem, 1998; Wilde et al., 2001), and that the antiquity of the oldest preserved crust is controlled by the waning of the Late Lunar Heavy Bombardment (LLHB), which may have affected the entire inner Solar System during the 4000–3450Myr time span (Dalrymple and Ryder, 1990; Cohen et al., 2000; Hartmann, 2002; Ryder, 2002). However, there is no trace of this hypothetical event on Earth or Mars (Arrhenius and Lepland, 2000) and the widespread preservation of protocrust on Mars and Moon suggest that a putative LLHB was not detrimental to planetary crusts. The timing of the onset of stable crust formation must thus be controlled by the dynamics of Earth itself.

Continents have been built gradually over time (Gastil, 1960; Taylor and McLennan, 1995; Abbott et al., 2000; Rogers and Santosh, 2004). The timing of the earliest nucleation of the present continents is unknown, but the oldest known vestiges are ca 4000Myr old, and continental masses with ages around 3500Myr form significant components of the exposed continental crust (Bowring and Housh, 1995; Nutman et al., 1997, 2000; Moser et al., 2001; Rino et al., 2004). The lack of evidence for continental crust older than 4000Myr cannot be taken as evidence against the production of granitic crust prior to that time. However, it is not likely that substantially older granitic materials form significant proportions of the unexposed or unstudied part of the continents. We can assume that stable continent formation started around 4000Myr ago, and that continental growth rates were high from ca 3800Myr. This is supported by the presence of detrital zircons with ages exceeding 3.5Ga in many sedimentary rocks and by the existence of cratonic components in this age range in all the present continents (Rino et al., 2004; Rogers and Santosh, 2004). With the exception of some late Archaean metasedimentary rocks from the Narrier gneiss Complex in Australia (Compston et al., 1985; Wilde et al., 2001) zircons older than 4000Myr have so far never been encountered in modern or Archaean sediments (Nutman, 2001). This evidence precludes the widespread exposure of Hadean rocks at the Earth’s surface through the geologic record, but does not preclude the existence of hidden reservoirs in the deep crust.

Trace element ratios and radiogenic isotope systematics provide further quantitative evidence regarding the extent of differentiation of Earth’s mantle, since they are sensitive even to hidden reservoirs. Based on Nb/U ratios in Archaean basalts and komatiites, (Campbell, 2003) suggested that rapid continental growth began about 4000Myr ago and persisted for 1000Myr during which period the continents grew to near their present size. Although Hf-isotopic studies of Jack Hills zircons indicate the existence of a continent-like, low Lu/Hf reservoir (Amelin et al., 1999; Harrison et al., 2005), the
absence of >4000 Ga old zircons in early Archean rocks suggests that, prior to this date, continental expanses were very limited and probably more short-lived than their modern equivalents. Additional evidence for the lack of extensive stable continental crust prior to 4000 Myr is provided by extinct radio nuclides, particularly \(^{146}\text{Sm}\). During formation of the crust, \(^{146}\text{Sm}\) is fractionated from \(^{147}\text{Nd}\), which is the daughter element for \(\beta\)-decay of radioactive \(^{147}\text{Sm}\) isotopes. Long time sequestration of reservoirs with \(^{147}\text{Sm}/^{146}\text{Nd}\) ratios different from the average silicate Earth will be visible in the Nd isotopic composition of the mantle and any mantle derived rock. If continental crust had stabilized and not re-homogenized back into the mantle by plate tectonics, mantle derived rocks would reveal that the mantle had been depleted in crustal components. There are two radioactive \(^{147}\text{Sm}\) isotopes; \(^{147}\text{Sm}\) which forms the basis for the standard \(^{147}\text{Sm}/^{144}\text{Nd}\) age determination system and has a half-life of 106 Ga, and \(^{146}\text{Sm}\) which has a half-life of only 103 Myr. The short half life means that this nuclide was effectively extinct on Earth ca 4100 Myr ago. The \(^{146}\text{Sm}/^{142}\text{Nd}\) decay couple is thus only sensitive to crust forming events that took place prior to 4.1 Ga. With the exception of a few very ancient rocks that carry a memory of a magma ocean event just after accretion of Earth (Boyet et al., 2003; Caro et al., 2003) there has been observed no \(^{142}\text{Nd}\) variability in rocks throughout the geologic record, which indicates that the mantle and any secreted crust was effectively re-homogenized up to about 4.1 Ga.

Another isotopic tracer relevant to arguments regarding crustal evolution is osmium. During separation of crust forming components from the mantle, Re and Os are strongly fractionated, and as \(^{187}\text{Re}\) decays to \(^{187}\text{Os}\), the rate of crust separation from a volume of mantle can be constrained by Re–Os isotopic systematics. Xenoliths derived from the lithospheric mantle commonly provide Re depletion ages around 3600 Myr or younger (Bernstein et al., 1998). The oldest stabilized lithosphere known is thus broadly similar in age to the oldest exposed crust, i.e. the age distribution of exposed continental surface may express the full history of stabilization of continental crust (Shirey et al., 2001).

The early Archean sedimentary record does not provide any isotopic evidence for exposure of extensive pre-3600 Myr crust. If present during the Archean, ancient Hadean continental crust would have been characterized by distinctive and diagnostic Nd, Hf, Pb and Sr isotopic signatures that would be inherited by Archean sediments, just as the geochemical fingerprint of the average exposed continental crust today can be identified globally in oceanic sedimentary rocks (Galer and Goldstein, 1991; McLennan and Taylor, 1991). The existence of extensive Hadean crustal exposure during the early Archean has not been detected in Archean sedimentary rocks. Late Archean sedimentary rocks from west Australia carry a small fraction of refractory Hadean zircons derived from 4.4 to 4.1 Ga felsic lithologies (Wilde et al., 2001), but similar ancient crustal components have not been identified in any other known rock units. The lack of evidence for the presence of long-lived Hadean continental crust by the beginning of the Archean suggests that the rate of destruction of continental crust was comparable to the rate of production prior to ca 4000 Ga. The end of the Hadean is, almost by definition, the time when the production of continental crust exceeded the rate of its destruction.

4. Granite production

As a means to simplify discussion, we use the term ‘granite’ to mean those rocks that are the primary constructional igneous components of continents, that is rocks rich in plagioclase and/or alkali feldspars, and with significant proportions of quartz. We intend to include a broad range of lithologies with this designation (e.g., tonalites, monzonites, granodiorites, granites sensu stricto, quartz diorites, etc.) whose mineralologies reflect both the silica–aluminum- and alkali metal-rich chemical composition of the magma and the conditions under which the magma crystallized. Such rocks are commonly referred to as “granitoid” or granite sensu lato in the petrologic literature, but for readability and because “granite” is a household term used for such rocks we opt to use it here.

Granite forms the continents and strikes many people as the most common rock type on Earth because of its ubiquity in our surroundings. However, on a cosmic scale, granite is probably a very rare and unusual rock type; it has never been observed elsewhere in our Solar System. In contrast, basalt, which forms Earth’s oceanic crust, is a very common rock type. The crusts of Mars, Venus and the Moon all have significant basaltic components, and even asteroids have basaltic components. This is easy to understand on the basis of our detailed knowledge of the process of basalt formation, which has shown that basalt is the inevitable product of partial melting of the silicate fraction of any terrestrial planetary body. Basalt composition is strongly buffered by a three-phase coticetic phase relationship over a wide range of melt conditions and melt fraction. Thus, the melt will be a rather uniform product, broadly referred to as basalt, a term which in reality covers a suite of rock types with subtle chemical and mineralogical variations.
When the basalt melt solidifies, mechanical sorting of the new-formed minerals may take place, because the different minerals have different densities, and thus also different buoyancies. The lunar highlands formed by the flotation of feldspar from a magma ocean shortly after formation of the Moon (Hartmann, 1989). However, even such a differentiated felsic fraction is still closely related to the basalt family of rocks, and is not related to granites. If dry basalt is reheated again, it will begin to melt along the buffered cotectic, and the melt that forms will also belong to the basalt family. In the absence of processes other than partial melting and fractional crystallization, all differentiated rocks will belong to the extended basalt family.

If, on the other hand, chemical interaction with aqueous fluids takes place, basalt is transformed into various metamorphic or sedimentary rock types depending upon the pressure and temperature during interaction. These alteration products will include hydrated and alkali-rich silicate minerals. Such basalt alteration products will partially melt at a much lower temperatures than the dry basalt solidus, and the partial melt products will be very much different. These melts are strongly enriched in silica, aluminium and alkali metals relative to the original basalt, and constitute the suite of rocks composing the granite family of rocks. The granitic rocks are light in color and have low densities (2.5–2.7 g/cm³ compared to 2.9 g/cm³ for basalt and higher that 3 g/cm³ for the upper mantle). This means that granitic rocks, if they are accumulated in large masses, have enough positive buoyancy to separate by flotation while basaltic crust and lithospheric mantle are subducted. This property is the cause for the gradual build up of the continental crust, while the basaltic crust is continuously destroyed. The rise of stable continents may thus be defined as the beginning of extensive formation of granitic rocks.

We are aware that an important aspect in the stabilization of the continental lithosphere might be the depletion of the lithospheric roots of continents by extensive melting (Abbott et al., 2000). It is possible that continental crust is stable largely because it is perched on top of rafts of lithospheric mantle that are buoyant because FeO has been removed by anomalously high degrees of melt extraction. There may be a causative relationship between the granitic crust and the extremely depleted lithospheric mantle, in that U-, Th- and K-rich granitic crust provides a thermal blanketing leading to accentuated melting of the underlying mantle. Granitic rocks are thermodynamically incompatible with the ultramafic mineral assemblages of the mantle and therefore granites cannot form as melts directly from the mantle, regardless of how hydrous it may be. We endorse the view, that granites are formed through a chain of processes involving two or more stages, including re-melting of oceanic crust processed by low-temperature alteration on the seafloor. The original material subject to fusion can be either normal oceanic crust, in which case granite formation is a continuous process that keeps pace with subduction, or large igneous provinces associated with the head of thermal instabilities (e.g. mantle plumes), in which case granite formation is episodic.

Several independent lines of evidence including rare gas isotopes (Allegre et al., 1983) and oxygen isotopes of ancient zircons (Mojzsis et al., 2001; Wilde et al., 2001; Valley et al., 2002) suggest that Earth oceans had formed by 4400 Myr. There should therefore not have been any shortage of the components necessary for granite formation in the very young Earth. Several inhibiting factors may have impeded formation of continental crust. With significantly higher internal heat production of Earth at 4400 Myr, convective dissipation of heat within the mantle must have been more vigorous. Although it could be suggested that perhaps subduction was more efficient and able to destroy continental crust, we consider the evidence to be against this proposition. Our present level of understanding suggests that simple gravitational settling of cold oceanic lithosphere with a higher mean density than the underlying hot and ductile asthenospheric mantle is the cause for subduction and lithosphere plate motion in general (Kellogg et al., 1999). Since the plate motions are not caused by viscous drag from the convecting asthenosphere, but mainly by the density of the lithosphere, the higher rate of convective overturn in the mantle should not to a first approximation affect the selective preservation of light continental crust during the subduction process. If the effect of the higher internal heat production of the young Earth was higher mantle temperatures, as is suggested by the common presence in the Archaean of komatiite, a basalt type that forms at higher mantle temperatures than those observed today (Nisbet et al., 1993), this should not diminish the selectivity either, since granite begins to melt at temperatures about 400 K lower than the basalt solidus, and the melting process has a positive volume of reaction. A higher mantle temperature should thus increase the density contrast between granite and basalt.

At least two causes can be argued for a lower rate of granite production in the infant Earth. Either the hydration and alteration of the weathering cycle was less effective, or reheating of weathered basalt was prevented. Because the heating of cold subducted...
lithosphere is one of the main cooling mechanisms for the asthenosphere, the latter is highly unlikely, in a hotter Earth. We therefore suggest that a marked increase in the rate of basalt alteration led to the onset of the stabilization of continental crust about 4000 Myr ago. We further suggest that this increased rate of alteration was made possible by the novel ability of the biosphere to convert solar energy to chemical free energy through photosynthesis.

5. The thermodynamics of life

Under all conditions where life is feasible, generic organic matter is thermodynamically unstable relative to isochemical combinations of C, H₂O, CO₂ and CH₄. Life is thus dependent on exploitable energy sources in order to persist and proliferate. Living organisms can harvest energy by catalyzing a plethora of reactions among organic and inorganic compounds in their environment. The energetic consequences of a large number of such reactions have been described and quantified by Amend and Shock (2001) and Amend and Plyasunov (2001). Of these reactions, all the chemoautotrophic ones are minor perturbations in the overall global transfer of thermal energy from the interior of Earth to space. We postulate that the rate of biological activity must be limited by the global energy budget for Earth. Contrary to the implicit assumption in many studies, Earth is not an inexhaustible source of energy and redox potential, and the first law of thermodynamics places tight limits on the extent of biological activity, simply because energy is a rather rare commodity on our planet. The scarcity of energy from chemical sources provides evolutionary advantages to any organism that develops photosynthesis. The theoretical upper bound for chemical energy production on Earth is the heat flow, which has a present day average value of 87 mW/m² (Stein, 1995). However, as most of this energy is transferred to space as heat without being captured via chemical processes, the realistic chemical energy source for life in Earth’s near surface environment is probably several orders of magnitude less than 87 mW/m². In comparison, Earth receives 340 W/m² of solar energy (Wells, 1997). Photosynthesis has allowed life to tap into this source of energy. Probably, the earliest photosynthetic organisms performed anoxygenic photosynthesis, and were dependent on mineral sources as electron donors, but relieved of the energy constraints (e.g. Olson and Blankenship, 2004). With the evolution of the more advanced oxygen producing photosynthetic pathway, life became independent of both energy and reducing power derived from the mineral substrate. The only limiting factors from the environment became the availability of chemical nutrients, such as phosphorous, nitrogen or iron. However, as ample energy became photosynthetically available, these nutrients could be reprocessed and life established biologic communities with high metabolic rates where chemical components were recycled, and energy for keeping the cycle spinning was harvested from light.

6. How heat flow is converted to a biologically available energy source

The internal energy production of Earth is mainly derived from the decay of radioactive nuclides. Heat from the radioactive processes adds to a reservoir of thermal energy still remaining from the accretion of Earth 4567 Myr ago (Amelin et al., 2002). A large part of the remnant heat from accretion is stored as latent heat in the liquid outer metal core, and is gradually released through crystallization and growth of the solid inner core. The heat of crystallization keeps the core at a nearly constant temperature as long as liquid metal remains, and contributes about 20% of the heat production (Stein, 1995; Kellogg et al., 1999). The large temperature difference between Earth’s interior and space drives the transport of heat towards Earth surface where it is lost through radiation. The temperature difference between the center of Earth and the surface is ca 6000 K (Boehler, 1996) or an average of less than 1 K/km. The temperature gradient is not constant with depth. Throughout most of the mantle an adiabatic temperature gradient of 0.3 K/km applies (Boehler, 1996), whereas steeper gradients exist in the lithosphere and at the core mantle boundary. Chemical potential contrasts are linked to temperature gradients, and it is therefore important to acknowledge that the thermal gradients of Earth are generally very small compared to the gradients we experience in our surface environments which we use as a reference frame for our chemical intuition.

Thermal energy cannot be directly used as metabolic fuel for living organisms. In order to be exploitable the heat must be converted to chemical energy, also known as free energy or Gibbs energy, which is the fraction of the total energy of a system that is not related to heat or work, but purely related to chemical interactions. Heat can be converted to chemical energy when materials are heated, either when a thermal gradient is perturbed in a stagnant material, or when materials are carried by advection into new thermal regimes. The basis for conversion of thermal energy to chemical energy is the endothermic reactions that components of a material
undergo when they form new phase assemblages with higher net Gibbs energy as they consume heat. The chemical energy can be released and exploited by life when materials with high net Gibbs energy assemblages are brought to lower temperatures, where lower net Gibbs energy assemblages are stable. Throughout most of Earth interior temperatures are high and reaction kinetics fast, which means that local thermodynamic equilibrium prevail. This means that Gibbs energy is at or near its minimum value for the local pressure and temperature conditions, and as a consequence, chemical energy cannot be harvested. Only when thermal structures are perturbed, the phase assemblages may possess exploitable Gibbs energies. Therefore chemoaotrophy is possible only in dynamic environments with steep thermal gradients or where advection transports matter into new thermal regimes. This will typically be the geothermal systems associated with magmatism in the near surface environment.

The process of formation of high Gibbs energy assemblages during heating at high temperatures and the formation of lower Gibbs energy assemblages during release of heat at lower temperatures is an integral part of the heat transport scheme for Earth. It should be viewed as a side effect of the transport of heat from Earth’s interior towards space, and not as an independent source of energy in Earth. Thus, at steady state, we can regard heat flow as a measure of the energy flux in Earth, and the geological and geochemical phenomena involved in transporting the heat as processes that store energy for a period of time in order to facilitate the transport most efficiently. Earth heat flow is thus an absolute theoretical maximum for the energy dissipation in Earth, in the absence of photosynthesis.

7. Chemoautotrophy and the carbon cycle

A comparison of the energy source for chemoaotrophy from Earth’s heat flow with the flux of CO$_2$ from the mantle to the surface environments reveals the ability of purely chemoaotrophic communities to influence the carbon cycle. Below we estimate the order of magnitude of various aspects of the carbon cycle averaged over the past ~4000Myr of Earth’s history, assuming steady state over geologic timescales. Such models abound with numerous simplifications and sources of error. However, there seems to be convergence on the orders of magnitude of the carbon fluxes estimated by various methods, and a generalized time averaged model provides insight into the basic influence of life on the carbon cycle. The present model is internally consistent, and probably accurate to the level of one order of magnitude.

There is some controversy over the mass of the carbon inventory of the solid Earth, whereas the OAS (Ocean–Atmosphere System) is well characterized. Here we use the reservoir sizes suggested by (Zhang and Zindler, 1993), which are summarized in Table 1, component fluxes which are summarized in Table 2, and energy budgets summarized in Table 3. The rate of CO$_2$ outgassing from the mantle is an important limiting factor in Earth’s carbon cycle. During geologic timescales, the rate of carbonate alteration cannot exceed the CO$_2$ flux from the mantle. Contrary to common speculation, high CO$_2$ concentrations in the atmosphere would indicate inhibited carbonate alteration, such that the rate could only consume the input from the mantle until elevated CO$_2$ activity forced the alteration rate. Conversely, low atmospheric concentration would characterize efficient carbonate weathering for constant mantle flux. The present day rate of CO$_2$ outgassing from the mantle is probably about 10$^{12}$mol/year (e.g. Zhang and Zindler, 1993; Saal et al., 2002). Zhang and Zindler propose a present day outgassing rate of 2.3 ±0.6 × 10$^{12}$mol/year, and suggests that 4000Myr ago it could have been ca 10$^{13}$mol/year. We adopt the average value of 5 × 10$^{12}$mol/year. Following Zhang and Zindler we assume that the present CO$_2$ inventory in the crust + ocean + atmosphere is ca 7 × 10$^{21}$mol and that the mantle originally had 380 ppm CO$_2$ as compared to ca 270ppm in the present depleted mantle. These mantle concentrations indicate that 70% of the CO$_2$ remains in the mantle, suggesting a large rate of recycling from the crust–ocean–atmosphere system. The carbon flux from the OAS to the crust is divided into a flux of organic carbon from primary production, and a flux of CO$_2$ forming sedimentary carbonates. It has been estimated that the present day primary production is 9 × 10$^{15}$mol C/year (Des Marais, 2000). Des Marais further suggested that about 0.1% of the primary production is sequestered in the crust, which would imply that ca 10$^{13}$mol organic carbon is sequestered in the crust each year. This number is greater than the total C flux from the mantle, which indicates that a substantial proportion

<p>| Table 1 | Earth carbon inventories |</p>
<table>
<thead>
<tr>
<th>Reservoir</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mantle (early Earth)</td>
<td>380ppm</td>
</tr>
<tr>
<td>Mantle (current)</td>
<td>270ppm</td>
</tr>
<tr>
<td>Crust–ocean–atmosphere (current)</td>
<td>7.0 × 10$^{11}$mol</td>
</tr>
<tr>
<td>Oceans (current)</td>
<td>3.3 × 10$^{19}$mol</td>
</tr>
<tr>
<td>Atmosphere (current)</td>
<td>6.2 × 10$^{16}$mol</td>
</tr>
</tbody>
</table>
of the C processed through the biosphere must be derived from the crust itself. Mass balance would require that 10% of the carbon sequestered in the crust is subducted back into the mantle. We will therefore assume that the organic carbon sequestration into the mantle is ca 10^{-12} mol/year (Table 2).

The lever rule expressing the mass balance constraint on the distribution of the two stable carbon isotopes 12C and 13C between biomass and carbonate sediments indicates that the flux of carbonate carbon is four times the flux of organic carbon to the sedimentary shell (Broecker, 1970; Schidlowski et al., 1979; Schidlowski, 1988). The total carbon recycling to the mantle (organic plus carbonate) must then be 5 \times 10^{12} mol/year and a total of 5 \times 10^{13} mol carbon is sequestered to the crust. These carbon fluxes would suggest that the crustal reservoir was built during 1400 Myr, if there was no reflux to the mantle or OAS. If, however, we assume that the crustal carbon reservoir of 7 \times 10^{21} mol is built steadily during the 4000 Myr history of crust formation, there would be a return flux from the crust to the mantle of 3.2 \times 10^{12} mol C/year. This is in agreement with the rate of recycling necessary to maintain the present C concentration of the depleted mantle. The estimated annual carbon flux from the mantle is consistent with a model of 10% melting of ca 300 km^3 mantle carrying 270 ppm C to derive the annual basalt production of 30 km^3 (Crisp, 1984). The model allows for the transfer of carbon from crust to mantle, allowing some fraction of the flux out of the crust to enter the OAS without passing through the mantle, for instance by metamorphic decarbonation (Ingebritsen and Manning, 2002).

The present atmosphere holds 6.2 \times 10^{16} mol and the ocean 3.3 \times 10^{18} mol CO_2 and its equivalent in bicarbonate. The OAS thus contains CO_2 equivalent to the amount of C outgassed from the mantle in less than 1 Myr. This short residence time for CO_2 in the OAS indicates that the OAS has little buffering capacity for CO_2 on geological timescales, and is mainly responsible for regulating the relative fluxes of organic relative to carbonate carbon. We can now adopt the carbon fluxes into a global thermodynamic model.

The flux of chemical components that can provide the energy and reducing power for chemosynthetic biosynthesis is coupled to the flux of carbon from Earth’s mantle. Both are supplied via plate motion, which is the main transport agent for the heat flow, and both are released from the mantle source during partial fusion in ascending mantle material. At steady state, we can assume as a first approximation, that the energy available for chemosynthetic carbon fixation per mol carbon flux is independent of the heat flow, and constant through time. The steady state includes the assumption that global magma generation is primarily tapping a MORB type mantle source. Geochemical evidence indicates that this has been the case throughout the known geological record (Gill and Bridgewater, 1979; Hamilton et al., 1983; Gruau et al., 1996; Blichert-Toft et al., 2000). As noted above, the annual photosynthetic carbon fixation is ca 9 \times 10^{15} mol carbon. One mol of carbon transferred to hexose consumes 477 kJ (the energy transformation during photosynthesis was first determined by Bass-Becking and Parks (1927) and literature values vary closely around the 477 kJ (mol CO_2)^{-1} value we have chosen, depending on the choice of thermodynamic input parameters). The primary production is thus equivalent to a conversion of 268 mW/m^2 of solar energy to chemical energy, as a global average (calculation is: 9 \times 10^{15} mol C/year \times 477 kJ/mol/(3.15 \times 10^7 s/year \times 5.1 \times 10^{14} m^2)).

The Solar energy contribution to Earth is 340 W/m^2. Of this ca 30% is scattered by the atmosphere and

Table 2

<table>
<thead>
<tr>
<th>Fluxes related to Earth’s mass transfer and chemical processes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluxes</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>Carbon</td>
</tr>
<tr>
<td>Rate of CO_2 outgassing from the mantle</td>
</tr>
<tr>
<td>4 Gya</td>
</tr>
<tr>
<td>Present</td>
</tr>
<tr>
<td>Average</td>
</tr>
<tr>
<td>Total C_{organic} primary productivity (current)</td>
</tr>
<tr>
<td>C_{organic} sequestered in crust</td>
</tr>
<tr>
<td>C_{organic} sequestered in mantle</td>
</tr>
<tr>
<td>C_{total} sequestered in crust</td>
</tr>
<tr>
<td>C_{total} sequestered in mantle</td>
</tr>
<tr>
<td>Chemoautotrophic-derived C_{organic} burial rate</td>
</tr>
<tr>
<td>Oxygen</td>
</tr>
<tr>
<td>O_2 released to the environment (=C_{organic} sequestered in crust)</td>
</tr>
<tr>
<td>Iron</td>
</tr>
<tr>
<td>Fe added to crust through basalt genesis</td>
</tr>
<tr>
<td>Rate of Fe oxidation as a consequence of sequestering C_{organic} in the crust</td>
</tr>
</tbody>
</table>

Table 3

<table>
<thead>
<tr>
<th>Energy fluxes on Earth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
</tr>
<tr>
<td>Geothermal</td>
</tr>
<tr>
<td>Solar (4 Gya)</td>
</tr>
<tr>
<td>Solar (current)</td>
</tr>
<tr>
<td>Energy captured through primary productivity</td>
</tr>
<tr>
<td>(global ave.)</td>
</tr>
<tr>
<td>Energy sequestered in crust through primary productivity</td>
</tr>
<tr>
<td>Energy sequestered in crust through chemoautotrophy</td>
</tr>
<tr>
<td>(maximum)</td>
</tr>
</tbody>
</table>
Earth surface, and 99.9% of the remaining is converted to heat and back radiated to space as long wave radiation. The remaining 0.07% is captured by phototrophic organisms. An annual sequestration of organic carbon of \( \text{ca} \ 10^{13} \text{mol} \) equals an energy transfer to the crust of 0.27 mW/m\(^2\). This is about 6 orders of magnitude less than the solar energy flux, but is comparable to 0.3% of Earth internal heat production as expressed by the heat flow. If we assume, for the sake of modeling, that the total heat flow is converted to chemical energy and spent solely on chemoautotrophic carbon fixation and we in addition assume that the fraction of the organic carbon produced that is sequestered to the crust is equal to the fraction of the present primary production that is sequestered, the flux rate between organic and inorganic carbon would be about 1:16 indicating that only 6% of the carbon flux to the crust would be organic. This value is unrealistic, even as a maximum value, since it is inconceivable that a purely chemoautotrophic community would be able to exploit the heat flow to this extent, because only an unknown, but miniscule fraction of the heat flow is available as chemical free energy in Earth surface environments. If we make the very optimistic assumption that a chemoautotrophic biosphere could harvest and sequester \( 10^{-6} \) of the energy flux like the extant photoautotrophic biosphere does with its solar energy source, the rate of organic carbon burial would be \( 2.5 \times 10^{8} \text{mol/year} \), which is \( \text{ca} \ 0.005\% \) of the net carbon flux through the OAS. In a world with a purely chemoautotrophic primary production, the organic sequestration would leave no resolvable isotopic imprint on sedimentary carbonates, and there would be no significant biological effect on the global carbon cycle in the absence of photosynthesis. Des Marias (2000) reached a similar conclusion based on the availability of sulfur and iron in the early ocean.

### 8. Photosynthesis

Photosynthesis covers a range of biological processes through which organisms force chemical reactions to proceed in the direction of higher Gibbs energy, rather than towards lower Gibbs energy which would be the spontaneously occurring reaction. The energy for this forcing is provided by involving the high energy component ATP (adenosine triphosphate) on the reactant side of the total reaction, and the lower energy assemblage of phosphate and ADP (adenosine diphosphate) on the product side, such that the change in Gibbs energy of the total reaction is negative, and thermodynamically viable. In phototrophic organisms ATP is manufactured in separate chemical reactions where chlorophylls absorb photon energy and build electrochemical gradients that fuel the ATP synthesis from phosphate and ADP.

The antiquity of photosynthesis is a matter of much debate. Arguments for a large diversity of claims are based on geological, paleontological, and biological data (Des Marias, 2000; Dismukes et al., 2001; Nisbet and Sleep, 2001; Towe, 2001; Line, 2002; Sleep, 2005). However, no combination of observations provide unambiguous evidence for any of the claims, except that it can be safely assumed that the rise of oxygen concentration in the atmosphere around 2.2 Ga (Canfield et al., 2000; Kasting and Siefert, 2002; Towe, 2002) must be the latest possible date for the emergence of oxygenic photosynthesis.

Biologic evolution in a microbial world is difficult to document through the geological record, because microbial organisms by definition are small and fragile with poor preservation potential as fossils. In addition, major evolutionary advances are usually expressed in the metabolic strategies of the organisms rather than in their morphologies, wherefore the taxonomic identity even of rare well preserved fossil microorganisms can be difficult or impossible to establish. The key to describing the geological record of microbial evolution relies on geochemical imprints of metabolic interaction between microbial communities and their environment. These range from supposed microbially constructed mineralized structures such as stromatolites to various isotopic chemofossils linked to metabolic isotope fractionations known from extant organisms and to global scale environmental effects such as the oxidation of Earth’s atmosphere (Rye and Holland, 1998; Canfield et al., 2000; Kasting, 2001; Shen and Buick, 2004).

For the purpose of harvesting solar energy, oxygenic photosynthesis is by far the most efficient strategy, but any type of photosynthesis holds the potential for accelerating biological activity by orders of magnitude relative to a situation with a purely chemoautotrophic primary production. We will here consider the genesis of photosynthesis as a major advance in metabolic strategy, and having irreversible consequences for Earth surface environments whether it is oxygenic or anoxygenic.

The greatest energy source in the surface environment of Earth is sunlight. Today the average solar energy flux to Earth surface is 340 W/m\(^2\) (Wells, 1997). 4000 Myr ago, Solar luminosity was probably ca 25% less (Sagan and Chyba, 1997), or ca 250 W/m\(^2\). The light energy is converted to heat at the surface, and
largely radiated back to space as long wavelength radiation. On a lifeless planet, the solar energy is converted to heat because the energy of individual photons is too small to break chemical bonds in the planetary surface materials. Blue light, the most energetic part of the visible spectrum, possesses 298kJ/mol photons. In comparison the breaking of the hydrogen–oxygen bond in the water molecule requires 492kJ/mol. A single blue light photon thus possesses far too little energy to dissociate water molecules. For this reason, solar energy is not converted to chemical free energy. With the evolution of chlorophylls in living organisms, this situation was dramatically changed. Chlorophylls have the ability to absorb energy from several consecutive photons and accumulate this energy for focused use. This allows organisms that possess chlorophyll to save up energy and use it for the basic CO₂ fixation reaction. Various photosynthetic carbon fixation schemes have been explored by life. Some involve mineral electron donors (reductants) such as ferrous iron or sulfide (e.g. Widdel et al., 1993; Nisbet et al., 1995; Amend and Plyasunov, 2001; Shen and Buick, 2004) while the most versatile and biochemically advanced pathways also produce the reductants in a separate reaction using photon energy to cleave water. The bulk reaction is then:

\[ \text{H}_2\text{O} + \text{CO}_2 + h\nu = \text{CH}_2\text{O} + \text{O}_2 \]

Through such reactions photoautotrophs acquired the ability to build up gradients in chemical potential, rather than just exploiting existing gradients, as was the fate of their chemosynthetic predecessors. The biosphere became able to convert solar radiation into chemical free energy. The energy conversion is 477kJ per mol of C converted into hexose.

It is impossible to measure Archaean bio-productivity, but if the metabolic strategies were available it is conceivable that productivity should not have been dramatically different from that of the present Earth. Today, about half of the primary production is carried out by oceanic plankton (Falkowski et al., 2004), so even though the existence of Archaean land biota has not been confirmed in the geologic record (Buick, 2001), Archaean primary production might have been on a scale similar to the present. The rate limiting factor for primary production was probably the availability of mineral nutrient (Canfield et al., 2000; Bjerrum and Canfield, 2002a,b), which could have been comparable to the present day supply. The present photosynthetic primary production is \(9 \times 10^{15}\) mol y⁻¹ (Des Marias, 2000). If this is converted to its chemical free energy equivalent (477kJ/mol) and evenly distributed over the Earth’s surface, it amounts to 268mW/m². This number should be compared to Earth’s heat flux at the surface, which is 87mW/m² at present (Stein, 1995). An unknown small fraction of this energy is converted to chemical free energy through melting and metamorphic reactions. The advent of oxygendic photosynthesis thus multiplied the available free energy in Earth surface environments. Most of the solar energy taken up in the biologic system, however, is transferred and dissipated within the system by respiration and heterotrophy leaving only a small fraction of the 268mW/m² contribution available for work in Earth’s crust. However, the effect is significant. The basic photosynthetic reaction above indicates that the release of oxygen to the environment associated with the burial of 0.1% of the primary production (Des Marias, 2000), which is ca \(10^{13}\) mol reduced carbon y⁻¹ must also be \(10^{13}\) mol y⁻¹. This would oxidize \(2 \times 10^{13}\) mol FeO. The annual basalts production contributes \(10^{14}\) mol Fe to the crust per year. Burial of 1% of the photosynthetic primary production is thus equivalent to oxidation of 20% of the magmatic Fe flux from the mantle to the crust, and could build the mantle reservoir of ferric iron during ca 1000Myr, if it was irreversibly sequestered in the mantle.

A photosynthetic biosphere could maintain an atmosphere that is not in chemical equilibrium with the rock substrate leading to more efficient rock weathering. One effect of this new photosynthetic energy source is to enhance weathering—the key link in granite formation. It is difficult to envisage a situation where a multiplication of the chemical free energy budget in Earth surface environment would not have left some profound geochemical traces in the rock record. The striking temporal overlap between the possible advent of photosynthesis (Schidlowski, 1988; Buick, 1992; Line, 2002; Dauphas et al., 2004; Rosing and Frei, 2004) and the beginning stabilization of continents, coupled with the correlation that granite only occurs on life bearing planets (from the admittedly limited statistics we have), we hypothesize that granitic continents on Earth are a consequence of the life induced early Archaean energy bonanza, and that granitic continents are likely biomarkers for photosynthetic life on silicate planets in general.

9. Life and the continents

Life was present on Earth 3800Myr ago (Rosing, 1999). At this time life was already rather advanced, and had colonized the open oceans. Geochemical evidence suggests that life had developed the ability to perform
photosynthesis (Dauphas et al., 2004; Olson and Blankenship, 2004; Tice and Lowe, 2004), and some evidence may suggest that oxygenic photosynthesis was active during the early Archaean (Buick, 1992; Nisbet et al., 1995; Line, 2002; Rosing and Frei, 2004). However, it is not the intent of this paper to discuss the evolutionary biology of photosynthesis, but rather to assess its energetic and geochemical consequences. We note that there is evidence in favor of emergence of photosynthetic metabolic strategies on Earth during or before the earliest Archaean, but also that there are arguments that suggest the earliest emergence of oxygenic photosynthesis during the Paleoproterozoic. Evidence for biologic management of Earth’s carbon cycle throughout the 3800 million years of rock record (Schidlowski, 1988; Veizer et al., 1992; Rosing, 1999; Shields and Veizer, 2002) coupled with our observation that Earth’s energy budget is insufficient to sustain biological activities that can affect earth’s carbon cycle in any significant way in the absence of photosynthesis, indicates that Earth’s biosphere has been able to harvest solar energy through some form of photosynthesis for at least 3800 million years. The level of complication of the Early Archaean geologic record is such that biochemical and taxonomic details cannot be reconstructed. Our thermodynamic analysis of the Earth system thus uses a generic carbon fixation equation: 

\[ \text{CO}_2 + \text{H}_2\text{O} + h\nu = \text{CH}_3\text{O} + \text{O}_2 \]

In our energy budget, we have chosen to ignore contributions from abiotic photolysis. We acknowledge that UV radiation does pay a contribution to the Gibbs energy budget in Earth’s surface environment, but we regard it as trivial relative to the contribution from photosynthesis. Photolytic cleaving of water molecules and photolytic oxidation of Fe(OH)$_2$ (Braterman et al., 1983) are viable processes at Earth’s surface, but both processes are associated with loss of hydrogen to space. Earth has a limited H$_2$O budget, and if photolysis had been a significant process in the past, any conceivable water budget for Earth would have been exhausted over a geologically short time span (Sleep, 2005). In contrast to photolysis, oxygenic photosynthesis generates oxygen while fixing hydrogen in organic molecules, without direct loss of hydrogen to space. Thus, it forms a link in a closed oxygen/hydrogen cycle.

The formation and stabilization of granitic continental masses on Earth is governed by fundamental properties of Earth’s geochemical reservoirs and involves a vast range of processes. We do not claim to understand in detail all of these parameters or their mutual interactions. It is impossible to provide any unambiguous evidence that biological energy fixation or any other singular process played a determinant role in continent formation. We are, however, convinced that the scale of photosynthetic energy fixation relative to the energy dissipation in Earth’s internal heat engine is so immense that the emergence of photosynthesis must have marked an important transition in Earth’s geochemical evolution. We find it striking that stable continents did not form during the first 600–800 million years of Earth’s history, and that there exists a temporal correlation between the emergence of photosynthesis and the rise of the continents. Although we do not understand the geochemical processes in detail, we claim that the environmental impact of a high rate of metabolic activity fueled by photosynthetic life caused a forcing of the weathering cycle, which eventually translated into an increased rate of granite production through the plate tectonic cycling of weathered primary crust.

When photosynthetic life forms produce organic matter they also release oxidized chemical components to the environment, be it free oxygen, ferric oxide, sulfate or other electron acceptors. The presence of these in the environment will enhance weathering processes partly by inorganic redox chemistry involving reduced basaltic glass or ferromagnesian minerals, partly by providing electron donors for biological metabolism of these materials (Furnes and Staudigel, 1999; Furnes et al., 2004), and partly by acidifying the hydrosphere leading to dissolution of primary igneous minerals, and precipitation of hydrous alteration minerals. A temporal correlation between the extent of stabilization of continental crust and the rise of oxygen in the atmosphere can be observed. This might in part be caused by a decline in oxygen consumption by crustal weathering as the exposed crust was changing composition from basaltic rocks rich in ferrous iron towards iron-poor granitic lithologies (Holland et al., 1986; Sleep, 2005) coupled with enhanced sequestration of organic matter in epicontinental depositional environments which also reduced the consumption of oxygen by back reaction of organic matter with atmospheric oxygen (Bjerrum and Canfield, 2002a,b). However, we suggest that the increased rate of continental crust construction happened in response to a spur in biologic activity associated with biological evolution.

On the larger scale, there exist some basic relationships inherited from the processes that formed the protosolar nebula that governs the abundances of carbon, hydrogen, oxygen and the long-lived radioactive elements that fuel most of the intra-planetary processes. The relative element abundances are modified during differentiation of the accretionary disc from which the
planets formed, but since the terrestrial planets are all accreted from comparable material within a limited region of the Solar nebula, they share similar element distribution characteristics. The abundances of heat producing elements relative to carbon thus provide insufficient energy to sustain a biosphere that can significantly influence the carbon cycle on any terrestrial planet that does not have photosynthetic organisms. This should be borne in mind when searching for traces of a subsurface biosphere on Mars. Such a biosphere would most likely not have left global geochemical imprints, but would only be encountered in the lucky instance of intersecting members of a very disperse biosphere in a given sample.

Acknowledgments

We thank the Danish Natural Sciences Research Council, The Carlsberg Foundation, Danish National Research Fund and The Greenland Mineral Resources Administration for support for this project. G. Arrhenius, E. Mathez, D. E. Canfield, K. Finster, L. Øvrevås, D. Harper, E. Nisbet, M. Russel are gratefully acknowledged for comments and discussions.

References


