Naturwissenschaften (2009) 96:1265–1292
DOI 10.1007/s00114-009-0602-1

REVIEW

Darwin's warm little pond revisited: from molecules to the origin of life

Hartmut Follmann · Carol Brownson

Received: 10 June 2009 / Revised: 5 August 2009 / Accepted: 10 August 2009 / Published online: 17 September 2009

Abstract All known cosmic and geological conditions and laws of chemistry and thermodynamics allow that complex organic matter could have formed spontaneously on pristine planet Earth about 4,000 mya. Simple gasses and minerals on the surface and in oceans of the early Earth reacted and were eventually organized in supramolecular aggregates and enveloped cells that evolved into primitive forms of life. Chemical evolution, which preceded all species of extant organisms, is a fact. In this review, we have concentrated on experimental and theoretical research published over the last two decades, which has added a wealth of new details and helped to close gaps in our previous understanding of this multifaceted field. Recent exciting progress in the molecular and genetic analyses of existing life, in particular microorganisms of ancient origin, even supports the possibility that a cellular, self-reproducing common ancestor might be assembled and resurrected in anaerobic cultures at some time in the future. Charles Darwin did not, and indeed, could not, address and specify the earliest phases of life which preceded the *Origin of Species*. However, in a famous letter, he sketched "a warm little pond with all sorts of… (chemicals, in which) …a protein was chemically formed." We try to trace the impact of his charming clear-sighted metaphor up to the present time.

Keywords Charles Darwin · Chemical evolution · DNA and RNA · Iron–sulphur world · Last universal ancestor · Organic soups · Origin of life · Thermodynamics

This contribution is part of the Special Issue "Beyond the Origin: Charles Darwin and modern biology". Guest editor: U. Kutschera (see Kutschera 2009).

H. Follmann (✉)
Institute of Biology, University of Kassel,
34109 Kassel, Germany
e-mail: hfollman@uni-kassel.de

C. Brownson
Westferry Road,
London, UK

Preface

> It is often said that all the conditions for the first production of a living organism are now present which could ever have been present. *But If (and oh what a big if) we could conceive in some warm little pond with all sorts of ammonia and phosphoric salts, light, heat, electricity etc. present, that a protein compound was chemically formed, ready to undergo still more complex changes* at the present such matter would be instantly devoured, which would not have been the case before living creatures were formed.

Charles Darwin (1809–1882) wrote these sentences in a letter to his botanist friend Joseph D. Hooker in 1871. What an amazing insight, decades before phosphorus was recognized as an essential element of cellular energy metabolism and heredity, although it had been discovered at the same time in "nuclein" (i.e., DNA; Miescher 1871). It was also long before amino acids, the nitrogen-containing monomers of proteins, were formed in Stanley Miller's seminal "organic soup" experiments conducted under the influence of heat, electricity, and light (Miller 1953, 1955).

Charles Darwin was an outstanding naturalist, versed in biology, husbandry, geology, and mineralogy. Was he also a secret, visionary chemist? We do not really know. Charles

and his brother Erasmus studied the "Chemical catechism" of Samuel Parkes and performed chemical experiments in a little laboratory in their father's garden during their younger years. The formation of gasses from minerals (for example, carbon dioxide from marble and acids, or ammonia liberated from ammonium chloride with CaO or quick lime) interested Charles greatly, to the extent that at school he was nicknamed "Gas." During his medical studies at Edinburgh University (1825–1827), at the age of 17, Charles enjoyed the modern chemistry lectures based on Dalton's theory of atoms taught by Thomas C. Hope. Indeed, Darwin remembered these as the only interesting lectures at Edinburgh. He did not specifically mention chemistry during his years at Cambridge (1828–1831). Later in his life, Darwin exchanged letters with almost 2,000 people and many of them (contained in the Darwin project database; Darwin Correspondence Project, http://www.darwinproject.ac.uk) addressed properties and reactions of lime, carbonates, ammonia, nitrogen in bones, etc. He knew about Justus Liebig, the famous German chemist with a reputation for analyses of the chemical composition of soils, food, and other biological matter which flourished in the nineteenth century. Thus, it would appear that Charles Darwin had acquired a good knowledge of contemporary chemistry.

One can only speculate whether twentieth century chemists were inspired by Darwin's Warm Pond scenario when they set out to discuss and explore the chemical evolution of life in laboratory flasks (Oparin 1936; Urey 1952; Miller 1953, 1955). The above letter to Hooker is quoted by Oparin, Hardin (1950), and Calvin (1969); but as far as we could ascertain, it was not mentioned in the early experimental contributions from other authors. It is a bit curious that Darwin did not mention a carbon source (non-negotiable for organic compounds), but he was probably not aware of the biological significance of methane (marsh gas) in anaerobic places. On the other hand, our colleagues of 50 years ago did not comment on phosphorus, equally important for life. Oh what inciting questions remain for historians!

In his famous *Origin of Species* (1859), Charles Darwin does not reflect upon the *earliest* origins of life, except for one brief sentence in chapter XIV. In a late letter (1881) to the botanist Nathaniel Wallich, he wrote "You expressed quite correctly … that I had intentionally left the question of the Origin of Life uncanvassed as being altogether *ultra vires* in the present state of our knowledge."

About a century of research later, Sol Spiegelman (Mills et al. 1967) and Manfred Eigen (1981) demonstrated both experimentally and theoretically that the Darwinian principles of replication, variation, and selection constitute a law of nature applicable to populations of molecules such as nucleic acids during self-reproduction, as well as to organisms. With the present knowledge of chemistry and molecular biology, the Darwin Jubilee Anniversary 2009 is worthy of an in-depth review of our molecular precursors in an enveloped, cell-like "species" ready to proceed on the long march towards more complex forms of life.

What is "chemical evolution"? It describes the chemical processes that took place on the prebiotic Earth about 4,500 to 3,500 mya. These events preceded biological evolution, a phase which led to the appearance of first living cells capable of self-reproduction at the expense of some rudimentary metabolism. Scientific concepts and laboratory experiments aimed to link life and its material constituents to the laws of chemistry and physics date back almost 100 years (Löb 1913; Oparin 1924). They gained momentum after Stanley Miller reported his now-classic results in 1953 (made public in Germany in the journal *Angewandte Chemie* in the same year) and thereafter. During more recent years, inspiring new ideas and experiments opened new "worlds" such as the iron–sulphur world of prebiotic chemistry and the RNA world of early genetics and catalysis. Today, molecular genetics and bioinformatics contribute deep-rooted sequence trees of all living organisms from which the characteristics of a potential last universal cellular ancestor (LUCA) have emerged.

Individuals who take an interest in evolution and the origins of life hope and, indeed, expect that modern science will provide a detailed and conclusive description. In a scenario of such complexity, when dealing with one's own descent, gaps and uncertainties are uncomfortable and unwelcome. Those of us in science who have touched on this multifaceted matter in theory or practice know that this expectation cannot be met despite the great progress in our knowledge made in recent decades. Authors of review articles and textbooks are not always immune to streamlining the arduous pathways from molecules to cells and organisms. In order to better explain the underlying principles of these pathways, specialists in cell physiology, chemistry, enzymology, or genetics may be tempted to belittle or disregard the problems with certain facts in scientific fields other than their own. This can be a disadvantage when evolution in general is questioned by religious dogmatists.

In this review, we have attempted to avoid bias and to describe diverse or conflicting results where necessary. It is not possible, nor intended, to trace our current understanding of the genesis of life on Earth in full. There have been many hundreds, if not thousands, of chemical and physicochemical experiments and theoretical analyses which have been performed under "warm little pond conditions" and in other environments. Some of the early work remained incomplete or inconclusive because of analytical and other experimental limitations but can now be assessed with modern methodology. We shall point to crossroads

where proponents and opponents of a specific pathway (reducing or redox-neutral conditions? solution or surface chemistry? protein or RNA world?) have difficulties in defining an outcome.

Information presented here places emphasis on studies published during the past two decades and only gives the essence of earlier work. Numerous books and reviews (e.g., Urey 1952; Oparin 1957; Calvin 1969; Lemmon 1970; Eigen 1971; Kvenvolden 1974; Follmann 1981; Kutschera and Niklas 2004; Kutschera 2009) describe in more detail previous experimental studies and theories of chemical and biological evolution.

Chemistry on a pristine planet Earth

Life, as viewed by chemists and physicists, describes highly organized entities ("cells"), composed of various kinds of organic and inorganic molecules and macromolecules, enclosed and separated from the environment by hydrophobic membranes. Moreover, the definition of life requires cells to acquire both energy (i.e., light, or energy-rich molecules) and new organic matter (i.e., carbon compounds and other nutrients) from outside in order to build and feed a genetic apparatus capable of reproducing the system. Finally, transmissible variations (mutations) of the genetic systems are necessary for the continued evolution of vital activities. All these properties demand very special environmental conditions.

A habitable place

Did "bio"chemistry start on Earth or was life brought in from space? Earth-bound chemical reactions and products, such as metabolites, life-like polymers, and cellular precursors, will be at the centre of this review. Nevertheless, large radio telescopes have detected many simple organic molecules in cosmic clouds, and analyses of carbon-containing meteorites that have fallen to the Earth have identified many organic compounds, in particular amino acids. It is conceivable that the European Rosetta space mission and its "cometary sampling and composition" COSAC instrument may find chirality—a characteristic of biomolecules—on arrival at the small comet Churyumov-Gerasimenko, 450 million kilometers away, in the year 2014 (Goesmann et al. 2007). Such findings from space deserve attention and are described below where pertinent but they do not prove that life originated "out there."

Earth is the only planet within the solar system on which the above-mentioned physical conditions for life exist. It is unique because its mean surface temperature, at a distance of 150×10^6 km from the glaring Sun, allows water to exist in all three physical states (as solid ice, liquid, and vapor). This is not the case on our smaller neighbor planets, hot Venus and cold Mars. Liquid water is a superior polar medium that dissolves, transports, and allows reactions between organic and inorganic compounds inside and outside cells. In a gaseous atmosphere, water vapor and gaseous carbon, nitrogen, and sulphur species and their various products coexist in numerous equilibria. Ice can trap, concentrate, and conserve complex products of primordial syntheses. Additionally water and abiotic hydrophobic lipids constitute biphasic media. When agitated by wind or waves, partition equilibria and compartmentalization processes in these media can separate different classes of newly formed and future biomolecules of different polarities.

Astronomers have, in recent years, detected many "exoplanets" and planet-forming clouds of gasses and dust which circle around stars in faraway galaxies (e.g., Beaulieu and PLANET Collaboration 2006; Eisner 2007). Most such planets are of giant Jupiter size but one with only 5.5 Earth masses size was also described. It is unanswerable at this time whether any of these bodies could support life elsewhere in the universe, but future infrared-based remote sensing technologies and a search for key molecules, including water, may someday provide an answer (Bada 2001). More data are expected from the National Aeronautics and Space Administration's current Kepler spacecraft mission.

Analyses of numerous long-lived radioactive and stable isotopes (of hafnium, tungsten, uranium, thorium, lead, zirconium, and others) in meteorites and terrestrial rocks indicate that the Sun and its planets accreted from gasses and cosmic debris 4,600 mya (Halliday 2001; Jacobsen 2003). It is assumed that the primitive solar nebula contained matter of a much older star which had collapsed and exploded as a "supernova." This event must have been significant for life as some contemporary proteins require catalytic metals heavier than iron (e.g., zinc or molybdenum) for catalytic activity. These were not produced by thermonuclear reactions in the young Sun but they came with the remains of the supernova. The fact that atoms such as essential molybdenum, found in our digestive enzyme xanthine oxidase, are direct descendants of the very origin of the universe is an intriguing consequence of this event.

The early Earth must have been a violent and chaotic body of rocks and gasses, with an initial surface temperature of around 1,000°C, and covered with huge impact craters from collisions with large and small asteroids during the "heavy bombardment" phase. Most of the lightest gasses, hydrogen and helium, escaped rapidly, and it is believed that the planet had solid rocks and an atmosphere (see below) between 4,400 and 4,100 mya. When temper-

atures dropped below its critical point (374°C), water condensed and filled crater lakes and early oceans; this occurred an estimated 4,000 million years before present. Indeed, one of the oldest sedimentary rocks (the Isua cherts of West Greenland) is 3,800 million years old.

Carbon of biogenic origin and presumed cellular microfossils has been reported in these and other 3,500-million-year-old sediments (Pflug and Jaeschke-Boyer 1979; Schidlowski 1988; Schopf 1993; Hayes 1996). While these objects were analyzed by powerful microscopic and microchemical techniques, doubts have remained and artifacts cannot be excluded (McLoughlin et al. 2007). Nevertheless, it appears safe to say that "Darwin's dilemma," i.e., the apparent missing Precambrian fossil record is no longer a problem (Schopf 1999).

It is generally agreed that the evolution of life—aggregation of fragile organic molecules—required a liquid, aqueous medium. (Thales of Miletus told us that at about 600 B.C.!) At the outset, however, "ecosystems" of accumulated organic matter would have been annihilated on a global scale in one of the frequent asteroid bombardments that continued to about 3,900 mya (Sleep et al. 1989). For the establishment of life, the time between such events had to be greater than the time required to evolve living organisms. It has been estimated from analyses of impact craters and energies (Oberbeck and Fogleman 1989) that such conditions would have prevailed between 4,000 and 3,700 mya. The agreement between these dates and those quoted above for the presumed existence of very ancient microfossils is remarkable. New numerical models of the habitability of the early Earth have just been published (Abramov and Mojzsis 2009).

Living matter is typically composed of 15–20 chemical elements, including six indispensable macroelements (C, H, O, N, P and S) plus the more or less essential microelements (Na, Mg, Cl, K, Ca, Mn, Fe, Co, Ni, Cu, Zn, Mo, and W). Although some of these elements are rare or sparingly soluble, their availability and uptake from the atmosphere, water, and minerals has obviously not limited chemical and biotic evolution in general. Elements capable of existing in different valence states (C, N, O, S, iron, and other transition metals) will have predominated on the pristine hydrogen-rich planet Earth in low, reduced valences: C as methane, N as ammonia or nitrogen, O in water, S as sulfide but not sulfate, and iron in bivalent oxides or sulfides (FeO, FeS) or even as a metal. The global redox potential began to change when some oxygen was liberated by UV light-induced radiolysis of water and much more drastically after the appearance of oxygenic photosynthesis in primitive bacteria (the later endosymbionts of green algae). The period during which the Earth's atmosphere transmuted from a reduced to a neutral and an oxidizing state is conspicuous in geology for the precipitation of massive red iron oxides ("banded iron formations") up to 2,000 mya. In biology, it opened the transit from nitric oxide (NO), nitrate, and sulfate as electron sinks (Ducluzeau et al. 2009) to modern oxygen respiration, thereby supplying ever more metabolic energy for hungry Proterozoic cells. We will not elaborate these dramatic biochemical changes further.

It is not widely known that Charles Darwin's "warm little pond" metaphor sparked a stimulating study (Spaargaren 1985) on the origin of the inorganic elements of life: From oceans? rivers? or panspermia from space? The latter author noted that the elemental composition of contemporary organisms, marine or terrestrial, is remarkably conservative from bacteria and fungi to fish and man. This is possibly dictated by the chemistry of all universal biochemical reactions taken together and is, therefore, still similar to the composition of early cells. The proportions of macroelements and microelements in organisms differ from those in seawater but the composition of river water is reasonably comparable with living matter—suggesting that organisms of such elemental composition originated in warm, evaporating little ponds fed by rivers. It can be concluded that the elemental profile of earthly life does *not* support Arrhenius' hypothesis of seeds of life propelled through space by light.

Energies

Thermodynamics and reaction kinetics dictate that chemical reactions, in particular those in which complex molecules are formed from simpler precursors, require energy to start and to proceed. The energies available and effective for de novo formation of organic compounds on early Earth are summarized in Table 1. It is important to note that there was not a single specific source: Physically and chemically different forms of energy were present, the most productive being geothermic heating, electric discharges (lightning), and UV light (Bossard et al. 1982).

Again, this gives us every reason to muse about Charles Darwin. Of the seven energy entries in Table 1, he could not have had any knowledge of natural radioactivity or cosmic radiation as these were discovered much later. He would also have been unaware of the abstract laws of thermodynamics (elaborated in depth by J.W. Gibbs only in 1875). Nevertheless, he does identify all the energies that would potentially drive evolution in a warm little pond, viz. heat (including volcanic environments, not specifically mentioned), light, and electricity!

Earth's early atmosphere during the first phases of molecular evolution, when simple organic, "biogenic" chemicals formed, has been difficult to define due to the many constraints (astrophysical, geological, and thermodynamic) that need to be considered. A comprehensive view was first presented by Urey (1952). It is now agreed that, even after much of the original light gasses (hydrogen and

Table 1 Energies at the surface and in the upper layers of early Earth, cf. Miller and Urey (1959) for quantitative estimates

Energy	Abundance	Significance
Heat	Ubiquitous	Standard energy source for chemistry; allows evaporation and sublimation; destructive if too high
Shock heating	Early large impacts	Substantial in earliest atmosphere
Sunlight	Ubiquitous	Dominant, fainter in the beginning; day-night cycles
Ultraviolet	≈ 1 % of total	High energy, photochemistry
Visible	Abundant	Absorbed by pigments, emerging photosynthesis
Infrared		Heat
Volcanism	Stronger than present	Allows surface reactions and dry-wet cycles
Hot lava, hot springs	Local, variable	High-temperature reactions
Natural radioactivity	Local (uranium, thorium, K-40)	Uncertain, maybe in singular events; may induce chiral discrimination
Electric discharges	Ubiquitous, stronger than present	Excitation of molecules, generation of reactive intermediates (ions, radicals)
Cosmic radiation	Relatively low	High energy; contributions uncertain
Reaction enthalpies	Common in chemistry	Driving other reactions, e.g., in iron–sulphur chemistry, redox reactions, etc.

helium) had escaped into space, the atmosphere was still hydrogen-rich, reducing, and almost without oxygen (Kasting 1993; Wiechert 2002; Tian et al. 2005); the stability of such an atmosphere (the "Early Faint Sun Dilemma") has been discussed in great detail (Sagan and Chyba 1997). Additionally, methane, carbon monoxide and dioxide, hydrogen cyanide, ammonia and nitrogen or nitrous oxides, hydrogen sulfide, and carbonyl sulfide (CH_4, CO, CO_2, HCN, NH_3, N_2, NO, H_2S, and COS) together with hot water vapor (H_2O) will have prevailed on the early planet's atmosphere for some time: not a delightful Garden of Eden, but a gigantic, productive outdoor chemical laboratory. Many of these gasses are still ejected in volcanic exhausts today and support some local abiotic chemistry. A glimpse of such very early, very basic organic chemistry is assembled in Fig. 1.

Towards biogenic molecules

A continuing supply of suitable carbon species plus ammonia was obviously critical for the formation of "organic" matter. Besides abundant methane and carbon monoxide, a requirement for carbon dioxide (CO_2, the molecule predestined for the carboxyl group of organic acids) has been implicated for optimum yields, but also disputed. In a recent study (Janda et al. 2008), the authors show that electrical discharges in a gas atmosphere dissociate CO_2 and produce crucial concentrations of carbon monoxide (CO) for the synthesis of organic molecules. For chemists, familiar with equilibrium constants and their temperature dependence, doubts about the contributions of individual gasses are irrelevant because at high temperature they all (plus elemental carbon and carbides) exist in dynamic equilibria, such as:

$$C + 2\,H_2 \rightleftarrows CH_4,$$

$$C + H_2O \rightleftarrows CO + H_2,$$

$$CO + H_2O \rightleftarrows CO_2 + H_2 \text{ (water gas equilibrium)},$$

$$N_2 + 3\,H_2 \rightleftarrows 2\,NH_3 \text{ (ammonia equilibrium)}$$

that will be replenished from the left as products are removed from the right in follow-up reactions, absorption to minerals, etc.

Ammonia, important for abiotic amino acid formation, is considered a component of the earliest atmosphere and/or was dissolved in the oceans. If it was depleted at some time or in a particular location, the catalytic photoreduction of nitrogen gas by UV light-driven water dissociation, in the presence of titanium dioxide or iron minerals, was a rich alternative source (Schrauzer et al. 1983; Brandes et al. 1998). Another reactive component, hydrogen cyanide

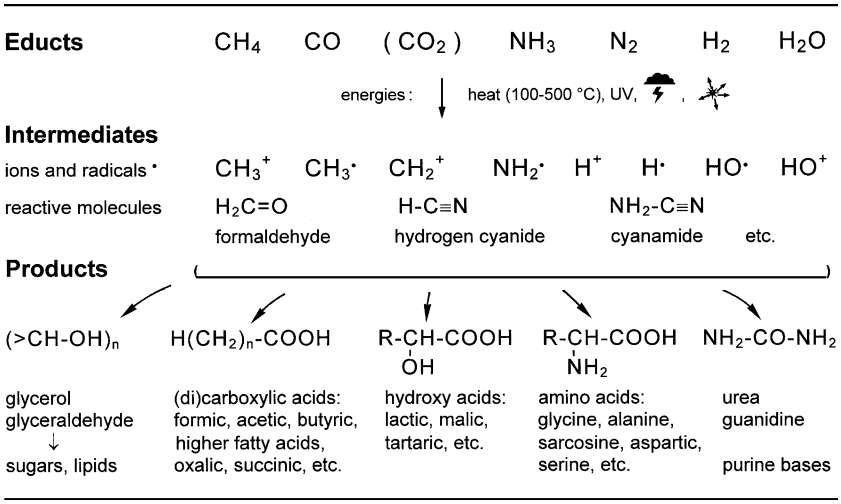


Fig. 1 A selection of educts, intermediates, and products of prebiotic chemistry in the gas phase, in solution, and on hot surfaces. See Fig. 3 for details of the chemistry of hydrogen cyanide and purines and Table 1 for the energies prevailing on the early Earth

(H–C≡N, prussic acid), is formed from methane and reduced nitrogen species in the gas phase and is abundant in volcanic exhalations. Huge amounts (>10 million tons per year) of it may have been produced during impact–shock events (Fegley et al. 1986; Chyba and Sagan 1992). Below 0°C, HCN and water exist in an HCN-rich eutectic state which appears ideally suited for condensation reactions and freezing-out of heterocyclic products.

"Fluctuating environments" are indeed a recurring theme in biogenesis research both in the laboratory and in theory. These include day and night, summer and winter, faint or bright Sun, hot and cold, wet and dry, and dissolved or mineral-adsorbed. Primitive chemistry on Earth differed in the atmosphere, flooded ocean basins, warm little freshwater ponds, submarine hot springs and on hot volcanic rocks, and in many other locations depending on the source of energy (Table 1), temperature, reaction times, presence of inorganic catalysts, or precipitation of products from solution. It is often wrongly presumed that educt concentrations, temperatures, and other variables of a particular prebiotic synthesis were the same and equally productive worldwide. It is not difficult to imagine that however varied the scenarios and the inventory of chemicals, progress towards ever more complex chemical structures, supramolecular aggregates, and "natural selection" cycles became inevitable during the hundreds millions of years.

It has recently been possible (Gaucher et al. 2008) to derive the trend of temperatures in Precambrian oceans for the time after life was established by measuring and comparing the thermostability of ancient ribosomal proteins. Obviously, it would not be possible for any cell to exist under conditions where its ribosomes, delicate nucleoprotein particles, would denature and disintegrate. From these studies, it was concluded that the aqueous environment supporting ancient life changed from a hot start (≥80°C), with thermophiles or even hyperthermophiles, to lukewarm (30°C), which was suitable for mesophiles in the Precambrian period between 3,500 and 500 mya.

Ponds of organic soup and rocky shores with organic scum

In 1953, Stanley Miller published the first results of his Ph. D. thesis, supervised by Harold Urey, in which he demonstrated an abiotic formation of amino acids under primitive Earth conditions. Since these classical experiments (Fig. 2), a wealth of information concerning the conditions and products of prebiotic chemistry has accumulated (Miller 1953, 1955; Miller and Urey 1959). There is no doubt that, under the right conditions, all kinds of "biomolecules" of low molecular weight (compounds identical with or similar to present-day metabolites, coenzymes, nucleotides, lipids, etc.) can form from inorganic elements and their simple hydrides or suboxides. It is more difficult to assess the significance of the still-growing number of studies in further steps of evolution such as the formation of oligopeptides and polypeptides, of sugars, RNA, etc. However, such a complex, multifactorial body of knowledge confers a crucial advantage when it comes to the more final questions such as the assembly of supramolecular structures addressed in the last section of this review.

There have been numerous simulation experiments to produce the main categories of potentially biogenic organic products, and these are summarized in Table 2. Many references to work earlier than 1980 (cited by Follmann 1981) have been omitted here.

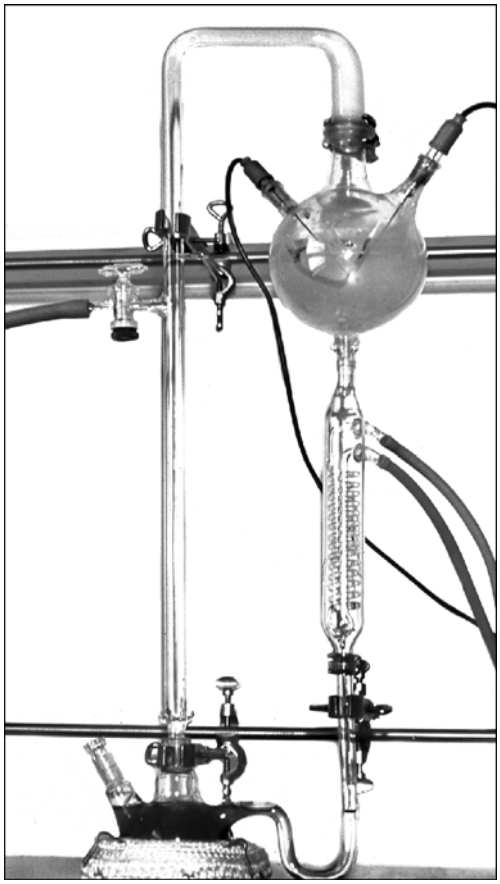


Fig. 2 A spark-discharge apparatus as designed by Miller (1953). Oxygen-free gasses are added at the *upper left side* (usually methane, ammonia, and hydrogen) plus water vapor from the boiling bottom flask (*lower left*), pass through a reaction flask (*upper right side*) in which spark discharges between tungsten electrodes generate slow electrons, UV light, and high temperature, and are cooled down in the condenser at the *right side*. Heat-stable products (cf. Table 2) accumulate in the bottom flask during several days of circulation

Solution chemistry

Amino acids, together with amines, are the most frequent products formed in reaction mixtures of reduced carbon and nitrogen species as in Miller's first prebiotic experiments and in the variants used by other workers (e.g., Bossard et al. 1982; Miyakawa et al. 2002). These observations are in obvious accord with the ease and large yields of classical syntheses of α-amino acids in the laboratory and with the stability of amino acids due to their zwitterionic structure. Amino acids are also found in the organic fraction extractable from carbonaceous meteorites (see below). The number of amino acids which have been characterized over the years is large. Small fractions of additional isomers have been identified in the preserved, dried residues of Miller's early experiments (Johnson et al. 2008), bringing the total number of different products up to at least 30. Unexpectedly, in contrast to hot reaction conditions intense UV irradiation of a CO_2:CO:CH_3OH:H_2O:NH_3 ice matrix frozen in a vacuum at 12 K (−261°C) produced 20 amino acids and other nitrogenous compounds (Munoz Caro et al. 2002). These experiments served to simulate the interstellar medium for comparison with future results expected from the COSAC cometary mission.

Only recently have the circumstances of amino acid synthesis under redox-neutral conditions been explored, based on the assumption that the atmosphere lost its reducing character around 4,000 mya. It was found that product formation also occurred under electric discharges in a neutral mixture of carbon dioxide, nitrogen, and water vapor above liquid water with or without sodium chloride (Plankensteiner et al. 2006) and that the yield of amino acids was greatly increased in the presence of oxidation inhibitors such as ferrous ions (Cleaves et al. 2008). Thus, neutral, hydrogen-depleted atmospheres may have been more important and productive for prebiotic chemistry on Earth than previously thought.

Eight particular amino acids which dominate in abiotic environments, occurred in the genetic ur-code, and are most frequent in contemporary proteins are listed alongside each other in Table 3. The comparison is indeed striking.

Sugars are hydrophilic, conformation-determining essential parts of nucleic acids and many coenzymes and metabolites. (Their important role in extant plant life and energy metabolism is not discussed as it is not relevant here.) Investigations of prebiotic sugar chemistry have had their difficulties. "Useful" C_5 and C_6 species (such as ribose or glucose) form from C_1 and C_2 precursors (formaldehyde CH_2O or glycolaldehyde) in the base-catalyzed "formose reaction" (Pfeil and Ruckert 1961; Follmann 1986) but poor yields, complex isomeric mixtures, and limited chemical stability have not been encouraging. Furthermore, the availability of sufficient prebiotic formaldehyde had been disputed but more recently a rich source has been identified in large atmospheric impacts (Fegley et al. 1986). Another avenue is aldol condensation of glyceraldehyde (C_3) catalyzed by Fe(OH)O that produces fructose and other hexoses (Weber 1992). Finally, a new pathway exploiting the well-known esterification of sugar *cis*-diol structures with boric acid (borate minerals; Ricardo et al. 2004) now permits pentose formation in good stability and yield.

Hydrocarbons, long-chain alcohols, and carboxylic acids Organic molecules with long linear or branched hydrophobic carbon chains are indispensable for the integrity of cells and intracellular compartments in aqueous environments. Present-day biosyntheses are complex and of

Table 2 Selected classes of organic, "biogenic" compounds and structures obtained in prebiotic Earth simulation experiments

Category	Type of chemistry	Products, yields, comments
Amino acids and related amino compounds	Electric discharges[a], heat UV light, HCN, CO Also in ice	Abundant, numerous ways of formation, (reducing or neutral); Simulating interstellar space
Alcohols and carboxylic acids	Electric discharges	Many compounds
Carbohydrates (sugars) Glycerophosphate	Formose reaction (HCHO) + phosphate, cyanamide	Borate minerals stabilize → phospholipids (membranes)
Hydrocarbons (lipids)	Fischer–Tropsch type	More detailed studies desirable
Peptide bond formation	Heat, clay, salt-induced, Iron-sulphur, carbonyl sulfide	Many condensation reactions
(Poly)phosphates	Apatite, carbonyl sulfide, heat	More studies desired for more efficient conditions
Porphyrins	Accompany amino acids and HCN	Low yield, characterized by spectra and metal complexes
Purine bases	HCN chemistry	Easily formed
Pyrimidine bases	Fischer–Tropsch + NH_3, CO–N_2	Numerous heterocycles
Pyrimidine nucleotides	Aldehydes + amino-oxazole	Cytidine and uridine phosphates
Purine and pyrimidine Polyribonucleotides	Evaporation with Mg, Ca, Zn salts	Best yield in presence of complementary strands
Deoxyribonucleotides	Reductive (iron-sulphur?)	More study necessary
Sulphur compounds, thiols	Iron–sulphur; H_2S addition to olefins	Probably abundant
Various metabolites, activated intermediates	Iron–sulphur[b] and other chemical environments, often in low yield	Increasing numbers described

[a] e.g., under conditions introduced by Miller

[b] cf. Huber and Wächtershäuser

little use for prebiotic simulation. Fischer–Tropsch type reactions using iron catalysts (long-known in technical coal chemistry) can produce numerous compounds from $CO/H_2/H_2O$ gas mixtures at around 250°C with product spectra that often resemble extracts from Precambrian cherts (cf. Kvenvolden 1974). Hydrophobic surface layers can also be seen when the product mixtures of Miller-type sparking experiments cool down. Hydrocarbons are photochemically produced from liquid methane even in the very cold atmosphere of Titan, the largest moon of Saturn (Porco et al. 2005).

Heterocyclic bases and nucleosides It is a strange coincidence that one of the strongest poisons, hydrocyanic acid (HCN), has been one of the most versatile and productive chemicals during the origins of life. (The rationale is that the gas, by competing with oxygen binding, is only toxic for higher organisms with *aerobic* energy metabolism.) The abiotic chemistry of HCN is very complex due to the pH-, temperature-, and concentration-dependent presence of dimeric, trimeric, and tetrameric intermediates (Ferris and Hagan 1984). However, this varied chemistry is the very

Table 3 Amino acids dominating under abiotic, early biotic, and contemporary conditions

In a carbon meteorite[a]	In sterile volcanic lava	In abiotic experiments	Early triplet code codons[b]	In average modern proteins (%)
Aspartate	Aspartate	Aspartate	Aspartate	Aspartate[c] (10)
Glutamate	Glutamate	Glutamate	Glutamate	Glutamate[c] (9)
Glycine	Glycine	Glycine	Glycine	Glycine (8)
Alanine	Alanine	Alanine	Alanine	Alanine (8)
Proline	Serine	Serine	Serine	Serine (8)
Leucine	Leucine, valine	Valine	Valine + leucine	Leucine, valine (7 each)
		Phenylalanine	Phe + tyrosine	Phe + Tyr (7)

[a] Engel and Nagy 1982

[b] Wong (1975)

[c] Including asparagine or glutamine, respectively

basis for the formation of biogenic amino acids and *N*-heterocycles. These include the two purine bases adenine and guanine (Fig. 3). Most striking is the smooth abiotic generation of adenine $C_5H_5N_5$, a pentamer of HCN (Voet and Schwartz 1983), considering that more than a dozen adenine-derived nucleotides, coenzymes, and other metabolites in addition to RNA and DNA assume essential functions in present-day biochemistry. An efficient prebiotic synthesis of cytosine and uracil was also described (Robertson and Miller 1995) but the condensation of these bases with ribose in aqueous media has not been possible.

The state of the art changed very recently when an efficient synthesis of cytidine and uridine phosphates was published which bypasses free ribose and pyrimidine bases as precursors (Powner et al. 2009). In an ingenious combination of plausible prebiotic reactions, the authors obtained pyrimidine ribonucleoside 2′,3′-cyclic phosphates from glycolaldehyde and glyceraldehyde (i.e., sugar components) plus cyanamide and cyanoacetylene (forming 2-amino-oxazole as a heterocyclic key intermediate) plus phosphate ions (which serve as general acid/base catalyst and pH buffer). This reaction sequence and the promotion of abiotic sugar formation by borate minerals mentioned above are outstanding examples of how new ideas and synthetic efforts presently expand and will eventually complete our knowledge of prebiotic chemistry.

Phosphate esters and anhydrides If one small molecule was to be named a biochemical symbol of present-day life, adenosine triphosphate (ATP) would be the obvious candidate. It links energy generation and storage with biosyntheses and provides the phosphate ester links in nucleic acids and many coenzymes. (Remember that phosphoric salts were destined to spice the warm little pond!) The problem in early days must have been that mineral orthophosphates are almost insoluble and are unreactive in water and that the energy-rich anhydride bonds of diphosphates and polyphosphates which are

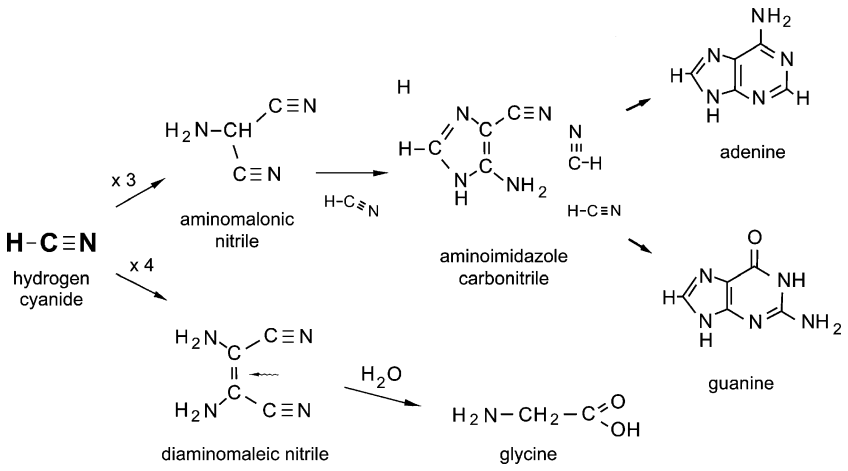


Fig. 3 Formation of *N*-heterocycles (purine bases) and amino acids from hydrogen cyanide. All the intermediates can be isolated. Some resemble intermediates found in present-day purine biosynthesis

needed for organic syntheses undergo slow hydrolysis. Nevertheless, calcium and magnesium diphosphates ($Ca_2P_2O_7$, or $Mg_2P_2O_7$ formed from $MgHPO_4$ under hydrothermal conditions) are good candidates for phosphorylation (Seel and Schinnerling 1978; Seel et al. 1985). Heating calcium phosphates (apatite) in the presence of acidic basaltic rocks to high temperatures (Yamagata et al. 1991) produces volatile phosphorus oxide (P_4O_{10}) which in water yields soluble diphosphates, triphosphates, and polyphosphates capable of phosphorylating alcohols, sugars, etc. The significance of such volcanic heat simulation has been questioned because of low yields and competing hydrolytic reactions (Keefe and Miller 1995); unfortunately, until now, the phosphorylation potential of P_4O_{10} samples generated in this way has not been tested. More recent experiments (Leman et al. 2006) show that inorganic phosphate and amino acids in the presence of carbonyl sulfide (O=C=S), a component of volcanic gas emissions, combine to form aminoacyl phosphates which in turn can serve as phosphorylating agents and promote phosphoryl transfer plus peptide synthesis under mild aqueous conditions.

These results may mark a breakthrough in prebiotic phosphate utilization. Nevertheless, a hypothetical, alternative route via more reduced and more water-soluble phosphites (Pasek 2008) should also be experimentally explored.

Chemistry on surfaces

Another "world" of prebiotic chemistry was proclaimed 20 years ago when Günter Wächtershäuser (1988; Drobner et al. 1990) realized that iron monosulfide (FeS, pyrrhotite) reacts with hydrogen sulfide to yield insoluble FeS_2 (pyrite) in an exergonic reaction. This can drive energetically unfavorable reactions, such as the reduction of carbon dioxide under neutral conditions and at a temperature of 100°C:

$$FeS + H_2S \rightarrow FeS_2 + 2\,H^+ + 2e^- \qquad \Delta G^\circ = -42\,\mathrm{kJ/mol}$$

$$FeS + H_2S + CO_2 \rightarrow FeS_2 + HCOOH \qquad \Delta G^\circ = -11\,\mathrm{kJ/mol}$$

The chemistry is particularly favorable as pyrite surfaces carry positive partial charges and are thus capable of adsorbing acidic products carrying a negatively charged carboxylate group. The thermodynamic potential and geochemical plausibility of iron–sulphur-coupled surface chemistry has since been exploited in an impressive number of ways. Acetic acid, thio-amino acids, hydroxy-amino acid, and amino acids have been produced, albeit in small yield. Amino acids have been activated for peptide formation, too (Huber and Wächtershäuser 1998, 2006). Iron–nickel–sulphur surfaces are active in reactions with carbon monoxide and can, for example, generate activated acetic acid from CO and methyl mercaptan (CH_3SH; Huber and Wächtershäuser 1997). At elevated temperature and pressure, a small amount of the central metabolite pyruvate was also produced from CO (Cody et al. 2000). The search for more organic compounds in the black precipitates incubated in anaerobic warm little bottles continues (Fig. 4).

Many of these laboratory reactions remind us of the metabolic traits and metal–sulphur centres known from chemoautotrophic thermophilic bacteria and archaea of ancient origin. Little doubt remains that the FeS/FeS_2 system once powered a productive surface metabolism on early Earth where abundant iron sulfide membranes existed near hydrothermal vents (Russell et al. 1994). It is evident that this system contributed to primeval carbon fixation and many different life-associated molecules.

Warm organic soup or a hot mineral paste? It has been argued that the organic broth theory and the iron–sulphur world theory of the chemical origins of life are incompatible (Wächtershäuser 2000). We see no reason for such a rigorous interpretation. Both scenarios are supported by robust experimental data and both could have found ample room in different locations on a highly structured, compartmentalized, unstable surface of the early Earth. Initially, their coexistence could even have been an advantage as newly formed anionic compounds (carboxylic acids, organic phosphates) preferred the positive pyrite surfaces whereas cationic products (amines, heterocyclic bases) accumulated on aluminosilicate sediments with negatively polarized surfaces. Frequent local mixing of two different compartments, e.g., in earthquake-rich regions, is easily imaginable. It should also be noted that the aggregation of "true" cellular structures from lipids, macromolecules, etc. most probably required meso-

Fig. 4 Iron–sulphur chemistry can be simulated in small septum bottles kept under strictly anaerobic and near-neutral conditions at around 100°C. Products are recovered from the amorphous black precipitate of iron sulfides (FeS and pyrite FeS_2) after several days. Pyrite crystals ("fool's gold," a common mineral) are shown at the *lower right side*

philic and not (hyper)thermophilic conditions (Miller and Lazcano 1995; Gutfraind and Kempf 2008). Moreover, it is questionable whether iron–sulphur chemistry is able to generate *all* the different kinds of chemicals required for a protocell (Keefe and Miller 1995).

Besides their catalytic action, another general property of mineral surfaces, their capacity to adsorb and activate different molecules next to each other, lowers activation energies and thereby favors bimolecular condensation or other reactions. Abundant surface-rich clay minerals have, therefore, been studied for a possible role in prebiotic syntheses (Paecht-Horowitz 1976; Paecht-Horowitz and Eirich 1988). It has been demonstrated that kaolinite enhances the thermal condensation of glycine to oligopeptides (Lahav et al. 1978). Montmorillonite, a layered aluminosilicate with excess negative surface charge adsorbs amino acid adenylates (the starting substrates of biological, ribosomal peptide synthesis) and enables production of homomeric or heteromeric peptides of up to 40 amino acids chain length at neutral pH and ambient temperature. Other authors have succeeded in generating long oligonucleotides from an activated adenosine phosphate on mineral surfaces (Ferris and Ertem 1993; Ferris et al. 1996). These reactions with complex kinetics deserve further and more intensive study.

Organic molecules in space

Let us finally consider the often-proclaimed arrival on Earth of more or less developed life from space. (Some people, including scientists, believe that life should not and, therefore, could not have originated on humble little planet Earth.) Panspermia or comet nuclei harboring dormant life spores, including DNA, may be charming scenarios but do not deserve much attention. Such "seeds of life" would never have survived the dehydrating vacuum, low temperature, and intense UV and cosmic radiation on their way to Earth (Weber and Greenberg 1985). Nevertheless, extraterrestrial organic matter has always and still does reach Earth, together with silicates and metals, in interplanetary dust particles (IDPs) and in one category of meteorites, the carbonaceous chondrites (see below). A concise analysis of the fluxes of inorganic carbon, nitrogen, and phosphorus from such sources on the early Earth has just been compiled (Pasek and Lauretta 2008). Estimates are up to five million tons of imported "biogenic" elements per year during the time of heavy bombardment which will have mixed with primordial terrestrial matter. As these events happened before and during the earliest phase of complex organic syntheses on the cooling planet, it makes no sense to differentiate between terrestrial and "outside" precursor molecules. After the heavy bombardment had ceased, the influx of IDPs dropped and today only around 100 tons of C, N, and P per year reach the surface of Earth.

One more extraterrestrial, extragalactic source of chemicals should not be forgotten. A multitude of inorganic and organic molecules in the interstellar clouds of our Milky Way galaxy have been identified by radio and infrared astronomy (Zuckerman 1977; Millar 2004). The surveys of these clouds now include more than 100 species ranging from diatomic and triatomic molecules (hydrogen, water, hydrogen cyanide, nitrous oxide, carbonyl sulfide, formic acid, etc.) up to an exotic 13-atom pentaacetylene nitrile, $HC_{11}N$. With ethyl alcohol (C_2H_5OH), it even boasts a molecule of special popularity among us earthlings. It should, however, be noted that their density in space is typically not more than ten atoms per cubic centimeter and that radio telescopes on Earth do not collect and concentrate such matter from a million light-years away but merely record the signatures of rotational and vibrational spectra. There is no opportunity for specific contributions of any such molecule to life.

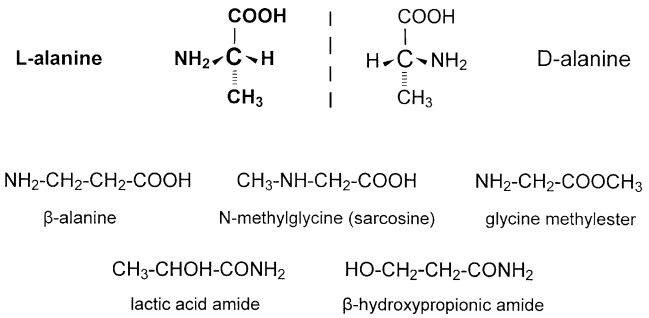


Fig. 5 Isomeric compounds with the elemental formula $C_3H_7O_2N$, which can be thought derived from methane (CH_3, CH_2, and CH), ammonia (NH_2 and NH), and oxidized carbon (CO) plus water. L-Alanine (in *bold*) is one of the 20 universal proteinogenic amino acids and the most dominant among all isomers. D-Alanine (the mirror image of L-alanine), β-alanine (nonchiral), and *N*-methyl glycine (sarcosine) are rarer metabolites of bacteria and animals, and the hydroxy acid amides are not biologically significant. A few D-amino acids occur in antibiotics. The L and D definition of chiral biomolecules as shown for the two alanine stereoisomers is historic; it does *not* necessarily describe the spatial orientation of a substituent towards the *right* or the *left*

Natural selection among molecules

It is most likely that many more organic compounds than those described were formed under all the chemical and local conditions listed above. No doubt more will be identified from new, imaginative laboratory experiments or in space. The reason for such a variety is simple. Carbon atoms, the key element of organic and organismic chemistry, together with a few other light "hetero" elements (hydrogen, nitrogen, oxygen, sulphur, and phosphorus) can exist in an almost

unlimited number of molecular structures held together by stable carbon–carbon and carbon–heteroatom bonds in chains, rings, and heterocycles. Furthermore, due to the steric and electronic properties of the tetravalent C atom, compounds of the same elemental composition can exist in many isomers that differ in their chemical nature and/or geometry (Fig. 5). It might well have been difficult, if not prohibitive, for early forms of life to evolve primitive specificity-enhancing enzymes or genetic elements if *all* these isomeric organics had populated the scene. As there is discrimination between isomers in modern biochemistry (e.g., L-amino acids and D-amino acids), it seems probable that "natural selection" began before there were cells and organisms. Given this scenario, we should consider whether formation, stability, and reactivity of particular molecules were favorable and superior for the increasingly complex molecular interactions that directed the pathway(s) towards living cells. Alternatively, it could be asked whether, in mixtures of abiotic isomers of similar stability and abundance, one particular population become dominant purely by chance and shaped the future in a "frozen accident" scenario?

Chiral molecules

One of the most disputed questions concerning the Origin of Life is the way in which chirality arose in the biogenic amino acids which in turn constitute chiral peptides and proteins. Molecules in which a tetrahedral carbon atom carries four different substituents (an asymmetric or chiral atom) exist in pairs of stereoisomers. D-Alanine and L-alanine (Fig. 5) or D-lactic acid and L-lactic acid (with –OH in place of $–NH_2$ groups) are simple examples. The two "enantiomers" are sterically mirror images of one another, but very similar in chemical behavior, with only tiny differences between them (see below). However, random distribution of both in a polymer (e.g., an enzyme) would produce a set of irregular three-dimensional structures that would be unfavorable for specific functions (e.g., substrate specificity). Such disorderly polypeptides are not found in nature where all the amino acids, incorporated into peptides and proteins, are of the same L form. Chirality is also common among sugar molecules, many of which contain several chiral carbons. A hexose sugar with $n=4$ asymmetric C can occur in $2^n=16$ stereoisomers (and all have been synthesized in the laboratory) but only three of them (D-glucose, D-galactose, and D-mannose) are of biochemical significance. The sugar moiety of nucleotides is always D-ribose or deoxyribose. From all these known coherent properties of life, we can reasonably conclude that natural selection did occur among prebiotic organics from the beginning.

Which subtle chemical and physical forces could establish and augment differences between the stereoisomers of a racemic D,L mixture and produce an enantiomeric excess (ee) of one over the other (Thiemann 1981)? The predominance of chirality or "handedness" of amino acids and other organics remains a difficult problem still today.

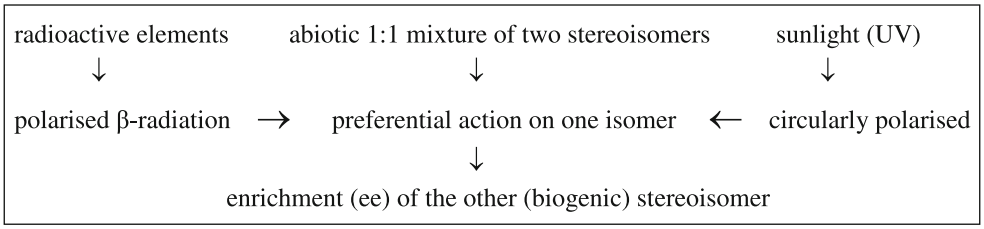
The question "Is the Solar System inherently left-handed?" (Chyba 1997) was prompted by the continuing analyses of the big Murchison meteorite, a carbonaceous chondrite which fell in Australia in 1969. It was quickly recovered and shielded from terrestrial contamination. Its interior was found to contain hydrocarbons, fatty acids and dicarboxylic acids, and pyrimidine heterocycles, but no purines. More than 20 different amino acids (totaling about 50 μg/g stone) were present including some of the common proteinogenic ones. There were also isomers not present in biogenic proteins, such as aminobutyrate, isovaline, and several diamino acids (Meierhenrich et al. 2004). The preponderance of some L-amino acids with 20–30% ee (60% in the case of L-alanine; Engel and Nagy 1982; Engel and Macko 1997) is striking; others exist as racemic mixtures of both D and L forms. Obviously, the inventory of organics in the meteorite does not reflect a biological source but resembles the product spectrum from Miller's and other simulation experiments.

The presence of chiral molecules in extraterrestrial matter has tempted critics of evolution to shift the origin of our biosphere to other cosmic environments. Nobody denies that matter from space reached the planet and mixed with local Earth products which had accumulated in warm little ponds or elsewhere. Thinking in quantitative terms and of the somewhat erratic composition just mentioned, it is unlikely that stone meteorites could have brought in and scattered enantiomeric "seeds" either in bulk amounts or of a defined composition. As the physical and chemical laws of stereoisomer discrimination must have been in force everywhere in the solar system, it should remain our prime interest to understand their effects on Earth.

Fundamental physical discrimination mechanisms

The discrimination between two enantiomers in a D,L racemic mixture by physical forces, which was necessary for life, has been observed in two different setups. One is the phenomenon of a parity-violating energy difference (PVED) for the weak interactions in atomic nuclei (Tranter 1985, 1986). This is responsible for the emission of left-handed polarized electrons (β) from the radioactive isotopes (e.g., potassium-40) widely distributed in nature. β-Radiation will affect the bonding electrons in chiral molecules differentially such that, in an irradiated racemic mixture, one enantiomer could suffer enhanced radiolysis, while the other (e.g., an L-amino acid) could be slowly enriched. Another asymmetry results from

the photochemical action of circularly polarized UV sunlight absorbed by chiral molecules. This again leads to enhanced decomposition of one (the more strongly light-absorbing) enantiomer and enrichment of the other.



The problem with these scenarios is that the asymmetry effects, calculated ab initio or observed in laboratory simulations, are minute for small molecules such as amino acids and sugars. For example, PVED stabilization of L-amino acids in solution is only of the order of -10^{-16} kJ mol^{-1}, which corresponds to an excess of 10^7 molecules of the more stable enantiomers per mole of the racemate (Laerdahl et al. 2000). Nevertheless, it has been estimated that selection towards *homo*-chiral amino acids and peptides could have been possible in an otherwise undisturbed lake of 1 km × 1 km × 4 m volume during 15,000 years (Tranter 1985). The solution to this problem still requires more theoretical and experimental approaches.

Reaction-induced and phase transition-induced enantiomeric enhancement

Much larger PVEDs and powerful mechanisms for chirality amplification are expected where crystals, polymers, and other macromolecules are involved, and indeed, such systems continue to emerge. In one illustrative example (Blair and Bonner 1981), a mixture of L-leucine and D-leucine with 31% initial L-ee was partially polymerized, and the resulting polypeptide then partially hydrolyzed; the L-ee in recovered unhydrolyzed polyleucine had increased to 55%. Repetition of such partial synthetic and degradative phases, probably during fluctuating environmental dry and wet cycles, were likely conditions on the primitive Earth. In a recent study, when amino acids of low (1–5%) initial L-ee were subjected to slow partial sublimation for several hours, an average fourfold enrichment of the L-forms was observed in the sublimates (Fletcher et al. 2007). This effect can be explained in terms of different vapor pressures of a solid D,L-racemate (reduced due to stronger intermolecular interactions) and the pure enantiomers, just as their melting points and solubilities differ. Sublimation could well have provided an important enrichment mechanism on the way from low ee products, generated by any of the previously described mechanisms, towards the homochirality required in amino acid polymers of high activity and specificity.

Adsorption on asymmetric surfaces

Polar organic molecules adsorb readily to polar mineral phases. If the latter are asymmetric, differential interactions (binding, catalytic, etc.) with enantiomeric substrates and/or products should ensue. Indeed, stereospecific binding of natural L-amino acids and D-glucose, but not D-amino acids or L-glucose, to clay silicates has been observed (Bondy and Harrington 1979). Crystals of common calcite ($CaCO_3$) also display differential absorption of D-aspartic acid and L-aspartic acid molecules on mirror-related crystal growth surfaces (Hazen et al. 2001). This effect is thought to serve as a model for the condensation of homochiral polypeptides on such crystals. Many more stereoselective interactions between organic enantiomers and inorganic mineral supports are likely to exist in nature; these may have contributed to the handedness of chemical evolution from the beginning, and further examples should be sought.

Chirality, conformation, and selection of sugars

Sugars are the other major class of chiral natural products in which asymmetric carbon atoms and the resulting three-dimensional structures of monosaccharides, disaccharides, and polysaccharides determine biological activity. D-Glucose, present in sucrose, starch, cellulose, etc., and D-ribose in

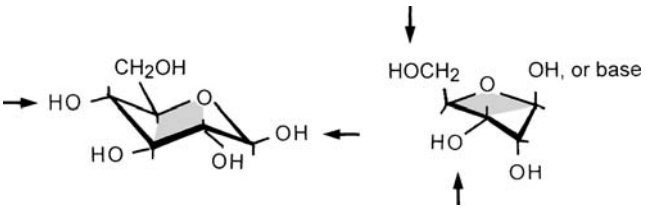


Fig. 6 Sterically relaxed and chemically reactive conformations preferred by biologically important D-hexose and D-pentose sugars. *Left* the "chair" conformation of *hexopyranosyl*-glucose; the *arrows* indicate the sites of 1,4-polysaccharide formation, e.g., cellulose, or bacterial peptidoglycans. *Right* the "envelope" conformation of *pentofuranosyl*-ribose; the *arrows* indicate the sites of 3′–5′-linkage in polynucleotide chains

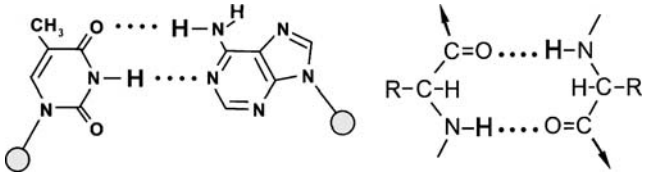


Fig. 7 Hydrogen bonds (symbolized by *dots*) are energetically favorable and exist where an electropositive hydrogen H (in *bold*) is situated close to an electronegative –N= or O= atom. *Left* H-bonded base pairing between uracil and adenine in RNA; the *circles* indicate the glycosidic bonds to ribose–phosphate moieties (not shown). *Right* hydrogen bonds in a unit of two antiparallel polypeptide strands such as found in a β-sheet protein structure; the *arrows* point to the carboxyl termini of the chains

RNA are the most prominent sugar compounds. (The nomenclature and apparent contrast of L-amino acids and D-sugars is due to historic definitions and *not* a basic discrepancy; indeed, both can be derived from the same chiral progenitor D-glyceraldehyde.) The preference of D-sugars has been traced back to parity-violating weak interactions (Tranter 1986), as outlined for enantiomeric amino acids.

An intriguing chemical mechanism of chiral induction was demonstrated (Pizzarello and Weber 2004) when chiral L-amino acids (such as the ones found in meteorites) were used as basic catalysts of the formaldehyde–formose condensation reaction described above and induced 10% ee with "correct" asymmetry in a C_4 sugar product (i.e., D-threose).

Sugars display yet another kind of geometrical constraint in their five-membered or six-membered ring structures (named furanose or pyranose; Fig. 6) which also appear to have directed evolutionary selection. Sugars exert most of their functions by means of the various hydroxyl (–OH) groups. Of all the possible geometric isomers of a C_6 hexose, it is the D-glucopyranose molecule in which *all* OH groups are directed towards the periphery and are freely accessible that is most stable and is preferred in most biochemical reactions. D-Galactose and D-mannose with one OH group pointing upwards (not shown) are part of more specialized metabolic functions; other hexoses do not occur in nature. A structure–function correlation based on selection of the "fittest" molecules can hardly be denied.

In nucleic acids, the pentose (C_5) sugar D-ribose assumes the furanose conformation which permits easy linkage of mononucleotides via their 3′ and 5′ hydroxyl groups to long 3′–5′-phosphodiester-linked chains and base-paired duplex strands. In RNA and DNA, such polynucleotide structures are stable yet allow a certain amount of flexibility and bending during replication, transcription, and translation. In an impressive series of chemical studies, Eschenmoser and coworkers synthesized pyranosyl-RNA (Krishnamurthy et al. 1996) and DNA homologues with D-glucose and hexopyranose components (Eschenmoser and Dobler 1992; Otting et al. 1993). Not surprisingly, all of these polynucleotides can form base-paired double strands. However, they differ from native RNA and DNA in backbone conformations and exhibit stronger and more rigid base pairing that is apparently less favorable for genetic functions than nucleic acid architectures based on furanosylpentoses. Perfection of the genetic material was achieved with the provision of D-2′-deoxyribotides (see below) that resulted in right-handed double-helical B-DNA of optimum structure and stability. None of the other double-helical DNA structures (A-DNA and Z-DNA, formed in nonphysiological media) has been selected under native conditions.

Heterocyclic bases and nucleic acids

The two pyrimidine and two purine bases of nucleic acids (cytosine and uracil or thymine; adenine and guanine) provide the material basis of genetic information by their capability to engage in C≡G and U(or T)=A pairs of "complementary" bases (Fig. 7). Specific pairing through three or two hydrogen bridges between donor and acceptor atoms is spontaneous and well understood. The energy gain for one H bond is small (only $\Delta G = -2.5$ kJ mol^{-1}) but even short stretches of base pairs in small genes include many hundreds of H bonds, not to mention the number in a long nucleic acid. Nevertheless, other isomeric heterocycles, with different substitution patterns, are found in HCN-rich and other abiotic reactions, and other stable base pairs (e.g., purine=purine) exist in vitro. Why, then, are C≡G and U/T=A the universally preferred combinations? Quantum chemistry has dealt extensively with the heterocyclic π-electron systems and conformations of nucleotides (Bergmann and Pullman 1972) and has shown, inter alia, that adenine and guanine are more stable than other isomers. Of all the monomers which were abiotically formed and available, only the common two purine–pyrimidine couples, and not the pyrimidine–pyrimidine or purine–purine pairs (with rare exceptions), enable double-stranded polymers of uniform geometry, stability, and readability in genetic systems. Combined with ribose-3′–5′-internucleotide linkage (explained above), an optimum of structure–function correlations was achieved and biologically exploited in nucleic acids.

Strangely enough, selection of the same four specific bases which allow and determine stable genetic information also introduced the possibility of spontaneous mutations. For example, hydrolysis of the amino group of cytosine to an oxygen function produces uracil, which would lead to a U=A pair in place of the original C≡G pair during RNA or DNA replication. Such deamination of cytosine bases by chance occurs at low rates (i.e., one event per 10^8 cytosine

residues and round of replication), but it is inevitable. Many of these events will be unfavorable or even lethal; however, in other cases, the modification may produce mutated, functionally improved genes and gene products and, thus, benefit molecular evolution.

Polypeptides

Proteins constitute the majority of biological matter and enable a wide variety of cellular functions. The necessity for synthesis of functional polyamino acids from just one kind of stereoisomeric monomer has been established above. A further stipulation is that, not only the genetically defined sequence, but also the macromolecular topography of a polypeptide should be reproducible and predetermined by intramolecular interactions. This requirement of living cells for a stable protein conformation is met by the same universal forces that govern the specificity of nucleic acids. Hydrogen bonds between the H, N, and O atoms of a peptide bond (–CO–NH–; Fig. 7) determine the spontaneous folding of L-amino acid chains into the stable α-helix and β-sheet secondary structures of proteins. The presence of 20 different amino acid side chains attached to a polypeptide backbone allows additional tertiary interactions (not outlined here) that contribute to individual protein stability, folding, and functions. For example, proteins embedded in hydrophobic membranes require specific arrangement of amino acids with hydrophobic residues, while polar side chains are required for catalytic effects, etc. The thermal stability of native structures (inter alia, important for thermophilic organisms) depends on the ratio between the highly ordered and stable and more loosely arranged flexible amino acid sequences. The 20 proteinogenic amino acids fulfill all these chemical demands in an average microbial cell (Weber and Miller 1981), and they were among those available from prebiotic syntheses. Thus, by "natural selection," they became part of the genetic (ur) code (Table 3).

Chance and necessity

Darwinian selection of individual molecules, contained in a large array of congeners and isomers, that are "fit" for specific binding, function, stability, or any other advantageous property is now being practiced on a large scale in high-tech laboratories which employ combinatorial chemistry for rapid, unbiased development of drugs or other desired compounds. Widely differing chemicals are allowed to react in automated evolution machines with thousands of reaction chambers and are analyzed and selected in high-throughput screening systems. One could, in fact, regard this modern kind of molecular design as a rapid-motion picture of the chemical and biological evolution on early Earth, which also selected the "fittest" molecules and cells albeit over thousands millions of years. Such a comparison is not totally foolish. It confirms the principle that numerous diverse reaction sequences, initially giving rise to random products, can and will, after an undetermined period of time, produce unique molecular systems or organisms fit for a particular function.

The logic of such apparently contradictory events and outcomes was substantiated decades ago by Jacques

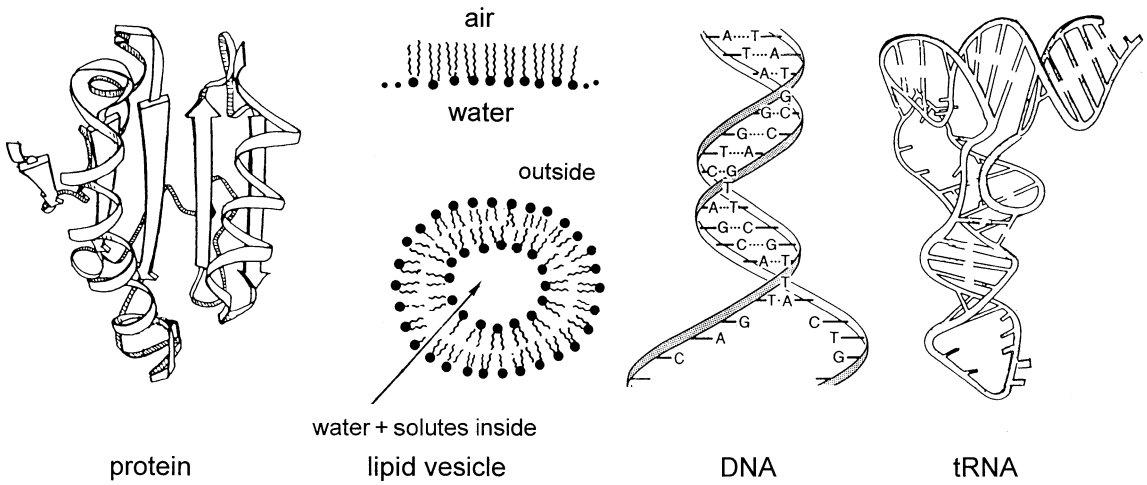


Fig. 8 Spontaneous self-organization of lipids, polypeptides, and nucleic acid strands in aqueous environment leads to ordered structures. Lipid molecules are schematized in the common fashion, consisting of a polar alcohol (*dots*) esterified with hydrophobic long-chain fatty acid residues (*ribbons*). They form vesicles with a double membrane and an aqueous interior when spread on water and agitated. Proteins and nucleic acids assume stable structures by extended hydrogen bonding (cf. Fig. 7). The polypeptide (*left*) is a thioredoxin of 120 amino acids which are arranged in α-helix and β-sheet sections. The polynucleotide at the *right* is an intramolecularly folded transfer-RNA of 80 nucleotides

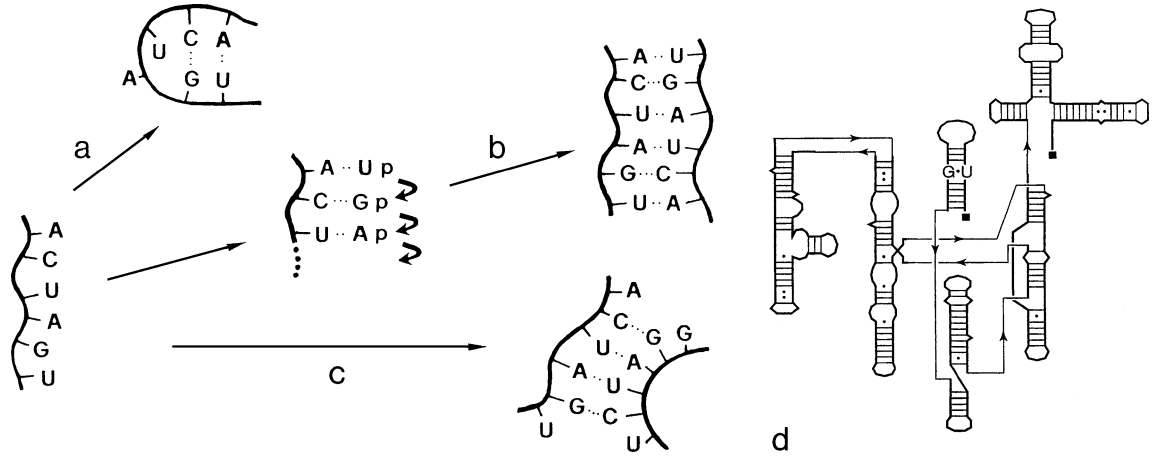


Fig. 9 Intramolecular and intermolecular base pairing of oligonucleotide sequences may permit spontaneous loop formation (**a**), synthesis of a new strand (replication, **b**) by condensation of complementary mononucleotides (*Ap*, *Gp*, *Up*), or recognition by another molecule (**c**, such as in codon–anticodon pairing). **d** Structure of a typical ribozyme sequence composed of numerous short stems, loops, and bulges. Shown is a self-splicing intron (Strobel and Doudna 1997) which serves for cutting out nonrequired RNA stretches

Monod, French biochemist and philosopher in his famous reflections on "Chance and Necessity" (1970). Years of fierce debate followed about the role of pure chance in life. This was disliked by many, especially when related to human life, but it is undoubtedly in action in gene mutations. The matter need not be reopened here; its biological aspects have been addressed at length by Carroll (2001).

Assembly of supramolecular structures

The buildup of cellular life and all its associated functions (energy metabolism, signal processing, motility, etc.) from inanimate small molecules must obey the laws of thermodynamics. It is common knowledge that these laws dictate that such a highly ordered state of chemical matter cannot arise of itself from disordered, chaotic conditions such as those on early Earth. The fundamental equation $\Delta G=\Delta H-T\times\Delta S$ (in which ΔG is the free energy change, ΔH the enthalpy change, ΔS the entropy change, and T the absolute temperature in Kelvin) combines the amounts of energy (usually heat) and ordering (or disordering) which accompany any reaction. A favorable process requires that ΔG be negative, and the second law stipulates an increase of the entropy or *dis*order (ΔS positive). With this in mind, it could appear that the highly organized states of "native" biomolecules such as polypeptides, nucleic acids, and vesicles (Fig. 8) violate a basic law of nature because ΔS of their formation is *negative*, whereby $-T\times\Delta S$ and ΔG become positive terms.

The paradox is easily explained. As biochemical and biological reactions occur in water, the entropy of water molecules, present in large excess around polymer biomolecules, has to be taken into account. The entropy (ΔS) of such a system *as a whole* (i.e., ordered biomolecules with one hydrate shell, $\Delta S<0$, *plus many* disordered water molecules liberated from the monomers, $\Delta S\gg 0$) will indeed *in*crease over an initial situation in which each monomer molecule had its own water shell. The necessity of including the entropy change of water when assessing the possibility that a complex orderly situation such as metabolism, energy generation, macromolecular structure, or integrity of cell membranes arises spontaneously from the individual components is frequently overlooked.

Supramolecular systems such as those schematized in Fig. 8 were all of central significance during the chemical evolution of life. The spontaneous formation of lipid vesicles around an aqueous centre allowed low-molecular-weight hydrophilic compounds to accumulate in elevated concentrations that were sufficient for specific interactions and polymerization (examples are described below). Protein chemistry explains why polypeptide chains with certain amino acids fold into ordered secondary and tertiary structures, in particular α-helix and β-sheet conformations (see above). Likewise, hydrogen-bonded, intramolecularly folded polynucleotides (tRNA-like or ribozyme-like structures; cf. Fig. 9) and later double-helical DNA genes were crucial for early genetics.

Proteins

Polycondensation of amino acids constituted during evolution and still provides today nature's chemical reactor in which new molecules are created for biological functions. Whenever specific sequences and special three-dimensional

structures happen to assemble in random polypeptides (such as synthetic "proteinoids"), these have a certain probability of expressing enzyme-like catalytic activity, inhibitory or regulatory function, metal-binding or other properties.

Problems faced by the chemist when trying to form peptide bonds in aqueous media are that (1) water has to be eliminated from the reactants and (2) amino acids exist in their stable zwitterionic forms $^{+}NH_3$–CH(–R)–COO^{-} (R– are side chains with additional hydrophobic or hydrophilic groups). Expelling water by dry heat usually leads to heterogeneous products and damages chemically sensitive residues R and reduce their biochemical potential. Nevertheless, a variety of thermal polyamino acids has been prepared and studied as "metabolic microsphere" models of protocells (Rohlfing 1976; Fox 1980). Specific amino acid sequences may be enriched in such products (Hartmann et al. 1981). More gentle reaction conditions prevail in the presence of carbonyl sulfide, a simple volcanic gas, which brings about formation of dipeptides and tripeptides by "polymerization on the rock" at moderate temperatures (Leman et al. 2004).

All these attempts to abiotically produce polypeptides suffer from uncertainties such as availability of monomers, suitable reaction conditions on early Earth, and poorly defined product structures. In the early 1990s, Rode and coworkers (Rode and Schwendinger 1990; Schwendinger and Rode 1992) tackled the lingering problem of prebiotic amino acid condensation with a new type of chemistry, salt-induced peptide formation or "SIPF." This has become a most productive approach and now represents the simplest way to oligomerize amino acids in aqueous solution (Rode 1999). Peptide bonds form under the dehydrating action of high (≥3 M) NaCl concentrations in the presence of copper (II) ions, which serve to complex and coordinate educt and product molecules. To mimic early Earth and ocean conditions realistically, it is possible to start from dilute solutions which are then subjected to evaporation cycles. An earlier short study (Okihana and Egami 1979) of amino acid oligomers produced in hot sea water enriched with transition metals may have employed a similar mechanism.

The SIPF reaction produces peptides up to the heptamer stage within a few days but long-term reactions have not been described. When the reaction is combined with adsorption on clay minerals, chain elongation to higher peptides is favored and hydrolytic backreactions are reduced. While the amino acids in general condense to all possible homodipeptides and heterodipeptides, it is an important feature that certain sequences are preferred, e.g., histidine plus hydrophobic residues, and others are under-represented. It is of considerable interest that the dipeptides preferred in SIPF chemistry appear to be reflected in dipeptide frequencies found in membrane proteins of ancient archaebacteria and other prokaryotic cells; this may not be a fortuitous correlation.

Finally, it is overdue and rewarding that attention is now being paid to the rare, yet physiologically important, amino acids histidine, methionine, and tryptophan in peptide formation and chiral discrimination (Reiner et al. 2006). These protein constituents have been neglected for a long time, mostly because they are chemically less stable than other amino acids and also because they were late additions to the genetic code. Histidine deserves interest for its outstanding role in enzyme catalysis and methionine as the N-terminal starter amino acid in ribosomal protein synthesis.

The present knowledge and potential of abiotic amino acid and peptide chemistry have led Rode and coworkers to proclaim an early peptide–protein world that allowed all essential functions of life such as self-organization and growth (Rode 1999). Support for this view has come from the case of a self-replicating α-helical peptide (Lee et al. 1996) which combines several of the preferred dipeptide sequences mentioned above in its 32-amino-acid sequence. A final answer to the sibyllic reflection "Can even peptides do it?" accompanying Lee's communication in Nature is still not known.

Nucleic acids

Both polypeptides and polynucleotides arise by elimination of water between the component monomers. The chemistry of assembling a 3′–5′-linked polyribonucleotide, however, differs from peptide bond formation in that negatively charged phosphodiester bonds –O–PO_2^{-}–O– are more reactive than a neutral, stable peptide bond –CO–NH–. Thus, metal ions that neutralize the negative charge and the adjacent 2′-hydroxyl groups of ribose can enable nucleolytic backreactions. (This is, inter alia, why DNA lacking that group was able to become a second genetic molecule alongside RNA, despite the extra investment in enzyme catalysis.) Moreover, in condensation reactions, the 2′-OH function competes with the formation of the regular 3′–5′ internucleotide bonds. On early Earth, the isomer problems could have been reduced under plausible fluctuating (hot and cold, dry and wet) environmental conditions in which various random copolymers with mixed 2′–5′ and 3′–5′ bonds formed, but the more thermostable double-helical strands enriched in 3′–5′ bonds survived subsequent hydrolysis and accumulated (Usher 1977). As zinc ions also favor 3′–5′-bond formation (Lohrmann et al. 1980), it is probably not fortuitous that modern RNA and DNA polymerases are zinc metalloenzymes.

Nevertheless, generally applicable conditions for abiotic nucleic acid formation are still not readily available. Recent results (Mansy et al. 2008; Rajamani et al. 2008) of "lipid-

assisted" ribonucleotide polymerization were unanticipated previously and may mark a breakthrough. When mixtures of nonactivated mononucleotides (AMP or UMP) and phospholipids were subjected to a series of hydration–dehydration cycles at 90°C, RNA-like polymers of 50 or higher chain length enclosed in lipid vesicles were produced. Other vesicles with "leaky" membranes made of fatty acid and fatty alcohol glycerides (without phospholipids) were able to retain encapsulated oligonucleotides but allowed the passage of nucleotides from the outside and efficient template copying in the interior.

Not all details of abiotic polynucleotide formation can be satisfactorily explained at this time. Therefore, a search continues for chemical alternatives that possibly predated RNA and might have combined simpler backbone structures with the firmly established hydrogen-bonded base-pairing scheme of genetics. Problems inherent in an RNA world (see below) could also have been alleviated by such a precursor. To this aim, ribose-3′–5′-phosphodiester links have been replaced by threose (i.e., a C_4 sugar) or even glycol (C_2) and by peptide links (Schoning et al. 2000; Zhang et al. 2005; Nelson et al. 2000). While interesting synthetic compounds were obtained, it seems unlikely that these "TNA," "GNA," and "PNA" nucleic acid analogs were relevant in chemical evolution. There is no real evidence for their prebiotic occurrence, and the necessary later transitions to present-day molecular biology would have created new, probably insurmountable problems.

Ribosomes and peptide synthesis

Bacterial ribosomes are small particles, composed of three different ribosomal RNAs (rRNA) and more than 50 ribosomal proteins, on which new proteins are synthesized. These are assembled from activated amino acids (aminoacyl adenylates and aminoacyl-tRNAs, –CO–OPO_2–O–AMP, –CO–O–tRNA) in the sequence programmed by messenger RNAs (mRNA), which were transcribed from genes. The various steps of ribosomal protein synthesis, which is universal and highly conserved, are not detailed here. What is important is that the intracellular assembly of ribosomes from their numerous constituent molecules is spontaneous. It is also amazing that, under suitable ionic conditions, reconstitution can proceed in vitro (Nomura and Erdmann 1970; Nissen et al. 2000). In a very complex yet predominantly error-free process, ribosomal proteins bind to each other and to the rRNAs, which serve as scaffold, in an orderly fashion. The particle is held together by numerous hydrogen bond and ionic interactions. The ease and fidelity of ribosome assembly suggests that it was an early invention in the prebiotic era; this new mechanism then overcame the limits of thermal or SIPF. The fact that ribosomes and their peptidyl transferase activity may be considered primordial ribozymes (Nissen et al. 2000) also points to an early origin, dating back to the putative RNA world.

It is appropriate to mention that a seemingly simpler, energy-saving, nonribosomal mechanism of peptide formation exists (Kleinkauf et al. 1992). In this system, amino acids activated by ATP are transferred and linked to an enzyme complex on which a peptide chain grows via condensation of thioester (–CO–S–) functions. Such a transthiolation–thiotemplate pathway could have neatly matched and served an iron–sulphur world. However, the present mechanism is restricted to small peptides of special functions and not suitable to generate larger enzyme or fibrous protein structures. Whether both pathways may have operated in parallel in the beginning remains a matter of speculation.

Lipid membranes

Enclosure of organic matter, be it a watery soup or sedimented particles, inside lipid membranes or vesicles is straightforward. Entropy-driven assembly of monomolecular surface layers and double membrane-coated vesicles or liposomes (Fig. 8), achieved by spreading and agitating lipids upon and in aqueous media, are standard experiments in physical chemistry teaching laboratories. Phospholipids were obtained from glycerol, phosphate, and hydrocarbon derivatives during evaporation on a hot sand or clay matrix (Hargreaves et al. 1977) and indeed formed the expected bilayer membranes. The advantage of lipid vesicles for synthesis of RNA-like polymers (Mansy et al. 2008; Rajamani et al. 2008) has been described above. Entrapment of native nucleic acids in artificial liposomes (Monnard et al. 1997) is also remarkably easy. Lipid membranes themselves integrate hydrophobic proteins (i.e., composed of ancient hydrophobic amino acids such as leucine, valine, phenylalanine; cf. Table 3) as channels or receptors for communication with the aqueous world outside. Current expertise in liposome and protein biochemistry combined has facilitated progress in simulating functional gene expression networks (Ishikawa et al. 2004).

Cell evolution and membrane topology including a reference to warm little pond chemistry have been discussed recently in context (Griffiths 2007). A "living" protocell requires that *many*, if not *all* low-molecular-weight and high-molecular-weight compounds are encapsulated within a membrane-coated vesicle. Experiments to assemble such a body have not yet been tackled and may have to await a future LUCA Revival Project as prophesied below. Meanwhile, the demands for synthetic protocell membranes are obvious. Different lipids such as glycerol esters, fatty acids, and fatty alcohols of various chain lengths and double-bond pattern must be combined to

Table 4 Inventories of genes required for extant minimal bacteria (Mushegian and Koonin 1996) and corrected for the earliest ancestor, LUCA

Gene functions	Minimum number in extant bacteria	Not required in LUCA (cf. text)	Basic genes/gene products[a] required for LUCA (incomplete list)
Translation (ribosomal protein synthesis)	95	Some (<20 amino acids present)	Ribosomal RNAs and proteins
Replication (nucleic acids)	26		Kinases, polymerases, ligases
Transcription (DNA → RNA)	9	Majority (no DNA present)	
Chaperones (protein folding)	13	Uncertain	Uncertain
Nucleotide metabolism	23	Many (no deoxynucleotides)	Kinases, ribosyl transferases
Amino acid metabolism	7	Some (fewer biosyntheses)	Transaminases, proteases
Nonribosomal peptide synthesis			Thiol peptidyl transferase(s)
Lipid metabolism	6		Fatty acid synthetases
Energy production (aerobic)	34	Many (no oxidases, etc.)	–
Energy production (anaerobic)	–		(De)hydrogenases, ferredoxin
			ATP via substrate chain phosphorylation
Coenzyme metabolism	8		(De)carboxylases, thioltransferases
Sugars, polysaccharides	8		Aldolases, glycosyl transferases
Uptake of inorganics (metals)	5		Chelating peptides, thionins, thioredoxins
Secretion	5	Uncertain (leaky membranes?)	Uncertain
Total	256	>50	Estimated ≤200

[a] Many known for ancient origin from sequence trees

permit structural heterogeneity and flexibility (rather than rigidity) and a certain degree of leakiness (rather than impermeability). The art of designing tailor-made membranes for protocells has hardly begun.

Molecular machinery of life: metabolism, energy, catalysis, and a genetic apparatus

Living cells require energy and nutrients to sustain their bare existence and to generate daughter cells by division. Long before chlorophyll-like pigments and phototrophic membranes were present and allowed the harvesting of light as an energy source, early anaerobic protocells could only vegetate by chemolithotrophic forms of metabolism. Ammonia, hydrogen, nitrite, sulfide, and the iron–sulphur system could serve as reductants, and C_1, C_2, and C_3 compounds like methane, methanol, formaldehyde and formate, ethanol and acetate, glyceraldehyde and amino acids taken up from the outside as substrates of an anabolic metabolism. In our opinion, it makes little sense to relate such cells with any kind of extant chemotrophic bacteria (e.g., methanotrophs or methylotrophs) all of which are *much* more specialized today. However, what can be safely assumed, from the great metabolic diversity of modern microbial life, is that *some* chemically and thermodynamically feasible reaction sequence (unknown in detail) was established that was able to produce, inter alia, ATP via substrate chain phosphorylation for intracellular energy. Formation of some group-transferring coenzymes such as ADP-glucose (Mar and Oró 1991) and pantetheine/coenzyme A (Keefe et al. 1995a, b) under primitive conditions has been reported. During a phase of random assembly of both synthetic and lytic pathways, rudimentary catalytic activities would have developed, and these catalysts would have been protected against breakdown when saturated with substrates (de Duve 1987, 1991). An increasingly productive network of metabolic extension and catalytic innovation could thus have been established. Was metabolism first?

Enzyme catalysis

Reactions of small biogenic molecules in aqueous media and at moderate temperature are slow for a variety of structural and kinetic reasons. Life could not have advanced without the substantial increases in reaction rates and specificities enabled by early enzymes, whether proteins or ribozymes (see below). Enzyme proteins require a small number (at least three) of functional amino acid side chains (acidic, basic, nucleophilic, and hydrophobic, as in glutamate, lysine, histidine, serine, and phenylalanine) next to each other in an active site pocket capable of binding and acting on their substrates (Walsh 2001). Such intramolecular domains are highly structured and substrate-specific in present-day enzymes but their assembly in early polypeptides is not as unlikely as it might seem: The energetically favorable folding of many different amino acid backbones (cf. previous chapters) *must* necessarily produce folds or

clefts in which several side chains of catalytic potential are in contact. This is an obvious example of the rules of chance-and-necessity. Catalytic properties were indeed observed in the first synthetic amino acid polycondensates ("proteinoids" and "microspheres"; Fox 1980). These included deaminase and transaminase, decarboxylase, esterase, phosphatase, peroxidase, and even polymerase-like (Jungck and Fox 1973) activities. Photochemical reduction of carbonate to formaldehyde and amino acids was also reported (Folsome and Brittain 1981). We take it for granted that a minimal enzyme set for earliest cellular metabolism and reproduction (cf. Table 4) could have assembled in that way. Unfortunately, systematic advanced studies of primitive enzyme proteins are missing, thereby making it easier for ribozyme aficionados to neglect these fundamental principles of protein-based catalysis.

It is worth noting that the enormous variety of presently known enzymes depends on a much smaller number of principles when it comes to specificities and catalytic mechanisms. The capacity of acidic or basic amino acid side chains (aspartate, glutamate; arginine, histidine, or lysine) to promote acid or base catalyses is obvious. Where substrates and products differ only in the position and number of a few hydrogens, catalysts can produce intramolecularly rearranged isomers and dehydrogenated (=oxidized) or hydrogenated (=reduced) products. Shifting hydrogens plus electrons between carbon or nitrogen atoms in such cases is a common and facile capacity at enzyme active sites as the smallest atoms, hydrogens, are easy to translocate. Another frequent and probably very ancient structure motif is the "nucleotide fold" of dehydrogenases, kinases, and other related proteins which determines their substrate specificity for nucleotides, ATP, and several coenzymes (Rossmann et al. 1974). All these basic principles, optimized in later modification and selection steps up to present-day enzymes, were certainly of great advantage in shaping early metabolism.

The genetic code

The molecular and informational bases of heredity are the nucleotide sequences in DNA genes and RNA transcripts. These carry the genetic code for their translation into the amino acid sequences of enzymes and all other cellular proteins. Initially excluding DNA (see below) but with RNA strands and abiotic amino acids at hand, an ur-code must have been one of the earliest and most important steps in the evolution of life. The rationale and necessity for a degenerate three-letter code of four nucleic bases (A, C, G, U) arranged in $4^3=64$ codons that specify up to 20 (or 21, including selenocysteine) different amino acids plus stop signals, present in all modern organisms, is textbook knowledge. The code is reproduced and discussed in a recent article in this journal (Penzlin 2009). It started with few primeval amino acids (i.e., those listed in Table 3) and subsequently expanded when further metabolic descendants became available, in a fashion that has convincingly been explained by the coevolution theory of Wong (1975). For example, of eight original serine codons UC*x* and UG*x* (*x* = any base), the latter four were used for coding cysteine and tryptophan, respectively, when their biosynthesis pathways from serine became established. In contrast, a possible doublet code with only 16 places would have lacked the potential for such expansions and died out if it had existed as an early competitor.

RNA and ribozymes first?

Ribonucleic acids are complicated and versatile molecules. With an anionic phosphate, hydroxyl groups, hydrophobic heterocyclic bases, basic nitrogen atoms, and hydrogen bonds, they combine different chemical structures and functions, with the notable exception of redox reactions. Their abiotic formation, selection from isomer mixtures, and roles in ribosomal protein biosynthesis and the genetic code have been described above. A conspicuous property of oligonucleotides and polynucleotides is the capacity to engage in *intra*molecular hydrogen bonding and to fold into many specific three-dimensional structures (Fig. 9) that enable different functions. At ambient temperatures, three or four adjacent base pairs (i.e., six to eight hydrogen bonds) suffice for engagement in specific, reversible interactions such as codon–anticodon recognition. Twenty or more base pairs produce thermostable folded structures as in transfer-RNAs and ribozymes. Such structures and, consequently, biological functions arise and exist solely from a defined chemical situation (nucleotide sequence) and the physical conditions.

A boost for evolution research came with the discovery that certain RNA molecules act like enzyme proteins in that they accelerate chemical reactions without themselves being changed in the process (Cech 1987). An "RNA world" of early evolution was proclaimed in which such "ribozymes" serve to catalyze the synthesis of new RNAs and other products (Gilbert 1986). Ribozyme activities have been described in increasing number (Joyce 1998), and the third edition of "The RNA World" (Gesteland et al. 2006) now combines contributions from 66 authors in a volume of 750 pages. Nevertheless, questions like "Did life really start out in an RNA world?" (Waldrop 1989) or "RNA seeks its maker" (Piccirilli 1995) remain and are justified by some serious objections.

Ribozymes such as the self-splicing intron drawn in two dimensions in Fig. 9 are single-stranded RNA molecules, at least 50–60 nucleotides in length, in which short double helices and loops are crowded together in three dimensions

maximizing hydrogen bonding. At an exposed site, phosphodiester bonds can be broken (hydrolyzed) or rearranged (transesterified) between ribozyme and RNA substrates, or ribozyme and amide or peptide substrates. Such reactions have clearly been useful in cutting, joining, and moving pieces of RNA in the course of creating novel polynucleotides and polypeptides with new functions. A reaction sequence in which pyrimidine nucleotides are put together from a ribose phosphate and a base and added onto an RNA substrate (Unrau and Bartel 1998) is of particular interest. In contrast, reactions like ribozyme catalysis of Diels–Alder cycloadditions observed in vitro have no physiological counterpart. Redox reactions like the action of dehydrogenases, which are crucial for cellular energy and anabolic metabolism are hard to imagine in an RNA world lacking enzymes and coenzymes. Finally, the most serious and puzzling question is the chicken-and-egg problem of how first catalytic RNAs came into existence themselves. Claims that RNA was preceded by another class of abiotic self-replicating molecules, probably peptide nucleic acids (PNA; Böhler et al. 1995; Nelson et al. 2000), in which ribose–phosphate links are replaced by amide or peptide bonds do not help in unraveling the problem. The contribution of ribozymes to primordial chemical evolution still requires evaluation.

DNA: the stable genetic information

One major drawback of RNA structures and their biochemistry is that the 2′ hydroxyl group of ribose adjacent to 3′–5′ phosphodiester bonds makes long RNA strands predisposed to hydrolytic cleavage and thus of limited use as a long-time stable genetic material. This problem was overcome with the second, more resistant nucleic acid, DNA. The statement voiced with the proclamation of RNA world, “Finally, DNA appeared on the scene” (Gilbert 1986), however, did not do much to enlighten that important step of evolution. 2-Deoxyribose is a very unstable sugar, claims for its abiotic synthesis were never confirmed, and the matter remained neglected. Formation of deoxyribonucleotide precursors of DNA by reduction of existing ribonucleotides (avoiding free sugars) was suggested as an alternative early on (Follmann 1982, 1986). Present-day ribonucleotide reductases are highly sophisticated enzymes, but their various protein-bound metal sites containing iron, iron–sulphur, cobalt, or manganese centres together with reactive thiols and radical intermediates could conceivably serve as primitive catalysts. Initial experiments under anaerobic iron–sulphur chemistry conditions in which we explored the reaction:

$$\text{ribonucleotide}(-\text{OH}) + \text{FeS} + \text{H}_2\text{S} \rightarrow \text{deoxyribonucleotide}(-\text{H}) + \text{FeS}_2 + \text{H}_2\text{O}$$

(Follmann 2004) give hope that this approach opens a pathway to DNA monomers. While the sequence “RNA came before DNA” is undisputed, it is more than doubtful that any ribozyme could have carried out the complex one-electron mechanism of specific deoxygenation at the ribose 2′ position (Stubbe et al. 2001). Ribonucleotide substrates *plus protein catalysis* is the obvious scenario for the advent of DNA genomes.

Once DNA together with DNA-replicating enzymes had assumed the role as a hydrolysis-resistant carrier of genetic information, the path of evolution opened towards higher complexity and diversity (Carroll 2001). Ribo-organisms like the last universal ancestor LUCA were succeeded by Archaea and Bacteria with their DNA genomes and the nucleated cells of Eukarya with an even higher DNA content. It has been estimated that the potential of DNA genomes is far greater than is actually realized in organisms. Repetitive DNA, nontranscribed introns, “selfish” (presumably useless) DNA in higher organisms (Orgel and Crick 1980), and other peculiarities of extant cells mean a waste of energy and scarce resources when looked at from the very modest beginnings. However, they provided a basis for continuing gene duplications and more complex forms of life. Should the origin of DNA precursors, ribonucleotide reduction, be given yet another “world” status?

Metabolism, protein, RNA, lipid, iron–sulphur world: ur-world prophets unite!

Charles Darwin was an exceptional man, unsurpassed in his ability to bring together divergent facts and ideas. It appears that such skills are needed today to reconcile the various “only” and “first” concepts of molecular evolution. Although our knowledge has become much richer in detail, it is also more puzzling when it comes to integrating the details and incorporating all “worlds” into a coherent and meaningful context. Of course, it is true that, at the outset, Earth had countless physically different microenvironments (Kuhn 1972, 1976; Kuhn and Waser 1981) in which only one or two chemical processes and products flourished, but these could not be called alive. However, over long periods of time, it must have been inevitable that different such localities merged and organic soups, precipitates, and sediments mixed. In this situation, the opportunities for higher complexity and further development would suddenly increase enormously. The significance and superiority of a “first” moment vanishes and becomes irrelevant if we choose a large-scale event of this kind as the impetus to the start of life. Figure 10 presents a minimum scheme of the possible interplay between molecules from an exterior medium and the interior of an early protocell on the brink of integrating metabolism and reproduction. The central

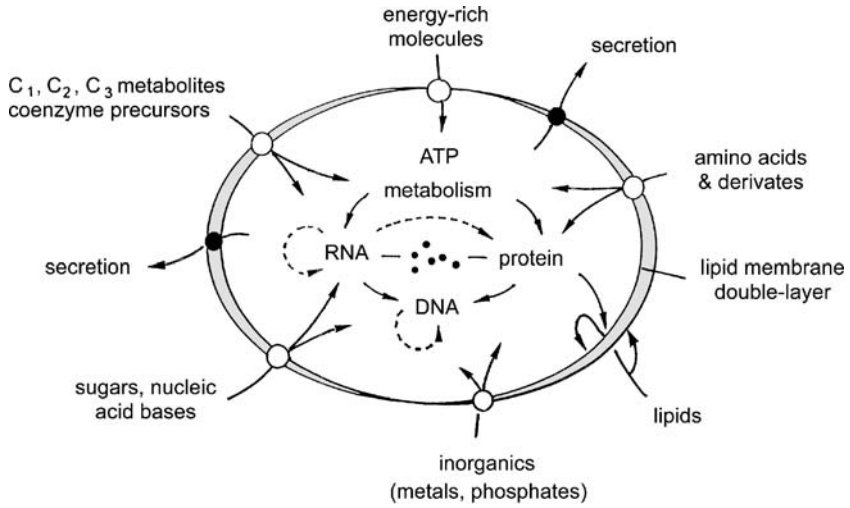


Fig. 10 A scheme of essential material transformations (*arrows*) and information transfer (*dashed arrows*) from an exterior medium and inside a protocell. The RNA–protein particles in the centre represent ribosomes. The lipid membrane must include hydrophobic receptor or channel proteins for the uptake of nutrients (*open circles*) and leaky patches for the secretion of waste (*closed circles*)

role of proteins can hardly be overlooked, but their cooperation with metabolism, RNA, and membranes is also obvious. We conclude that there is no serious reason to rule out a productive coexistence of various classes of abiotically generated organic molecules and that ribozyme and enzyme catalysis may have operated side-by-side. More and innovative experiments to tie them all together would be a welcome addition to our current knowledge.

What was LUCA, and how did it live?

"Probably all of the organic beings which have ever lived on this Earth have descended from some one primordial form" is one of Charles Darwin's rare comments on the earliest form of life in his *Origin of Species* (chapter XIV, 1859). As early as 1837, he had sketched a vision, in one of his notebooks, of a genealogical tree in which different genera (A–D) could have developed from a common origin (①). Its juxtaposition, in Fig. 11, with the generally accepted tree of life (Woese 2000) is amazing and inspiring. The further diversification of the three ur-kingdoms Bacteria, Archaea (archaebacteria), and Eukarya (eukaryotic plants, fungi, and animals) has been addressed in this journal earlier in the year (Penzlin 2009) and is not a subject of our article. However, it is of interest and a challenge to conclude our review of the presumed chemical properties and biochemical capacities of abiotic cellular entities with what might be expected of "live" ancestors at the roots of "big tree."

The last universal common ancestor (LUCA) from which all current living organisms are descended is a hypothetical protocell that arose at an unknown point of time more than 3,500 mya. We are undecided whether to think of LUCA as a genetically defined individual species or a small commune sharing genetic information by frequent horizontal gene transfer (Woese 1998; Whitfield 2004). The division in time and in complexity between LUCA and the most ancient fossil microorganisms, which are more than 3,000 million years of age and abound in Precambrian sediments in several areas of the world (Pflug and Jaeschke-Boyer 1979; Schopf 1993) is also uncertain.

It has been reasoned (Lazcano and Miller 1994) that once cellular life had started the time taken for a transition to autotrophic, photosynthetic cells may have been relatively short in the evolutionary timescale—probably less

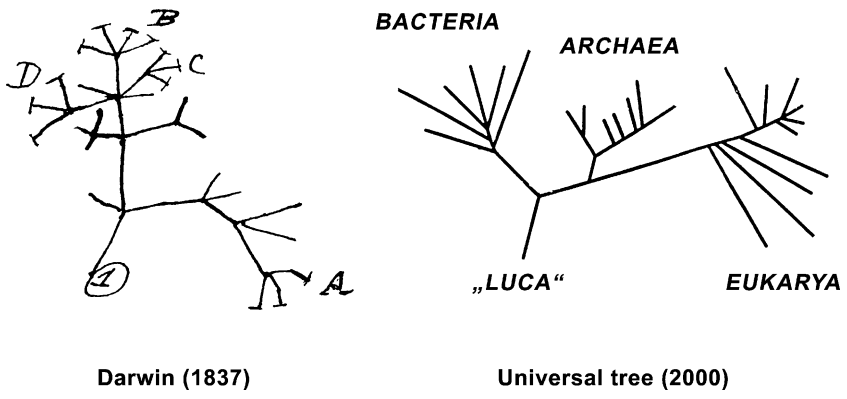


Fig. 11 The imaginary phylogenetic "tree" drawn by Charles Darwin (1837) starts from one origin (marked ①) just like the accepted big tree of life (Woese 2000) springs from one common ancestor (LUCA)

than ten million years. Many metabolic mechanisms and enzymes exhibit considerable analogies and similarity in the different life forms of today; thus, new genes required for photophosphorylation, CO_2 fixation, etc. might have been generated from tried gene sequences of the ancient chemotrophic pathways by straight gene duplication events.

Some characteristics of LUCA and its environment have recently been deduced from the G+C content of ribosomal RNAs (rRNA). This base composition is strongly correlated with the optimal growth temperature of prokaryotes due to the higher stability of G≡C base pairs (Galtier et al. 1999; Boussau et al. 2008). Ribosomal RNAs and protein sequences from hundreds of modern species have been analyzed as "thermometers" for backward extrapolation, and these suggest that LUCA most likely still had an RNA genome. (In possession of proteins, however, it cannot have been "the last ribo-organism" cited by Benner and Ellington 1987). It strived for life in anoxic waters at about 60°C, well *before* thermophilic Bacteria and hyperthermophilic Archaea sprouted from the big tree (Fig. 10). These ancestors of all modern organisms (including ourselves, the Eukarya) may have had to survive a later, hot period in early oceans (≥80°C) after a big impact event to which they responded with the acquisition of more thermostable DNA genes.

LUCA must have possessed a minimum set of genetic elements to code for its metabolic demands and reproduction. Such a set was deduced from the genomes of two very small extant bacteria (*Mycoplasma genitalium* and *Haemophilus influenzae*) and it includes 256 genes (Mushegian and Koonin 1996). LUCA should have managed with even fewer genes if it lacked, inter alia, redox enzymes for aerobic respiration, DNA synthesis, and complex amino acids not yet fixed in the ur-code. Our provisional guess is that no more than 200 genes (Table 4) were sufficient for its existence. This estimate is consistent with earlier figures arrived at in a different context and using different approaches (Gil et al. 2004). Such considerations certainly do *not* imply that LUCA and present-day parasitic microorganisms have anything in common except a very small number of genes.

Can LUCA be reconstructed in the test tube along these lines and revived? Success is not guaranteed but experiments are worth the effort. A promising strategy may be to follow the recent protocols (Gaucher et al. 2003, 2008) to restore just one kind of primeval protein (i.e., ribosomal elongation factor Tu). Retrogression from protein and gene sequences of numerous extant microorganisms near the roots of the big tree towards LUCA's putative genes and gene products would be carried out by computational analyses. From these data, synthetic gene construction, protein expression, and activity measurements might eventually provide the minimum number of catalysts and the genetic elements required for anaerobic, chemotrophic life. Their molecular interactions could be analyzed by modern genomics and proteomics techniques and potentially adverse combinations could be deleted or modified. Inclusion in ready-made lipid vesicles (Fig. 8) spiked with hydrophilic channels might facilitate the project. Incubation in defined artificial media will have to include a wide variety of inorganic and organic potential nutrients, proceed at different (mesophilic? thermophilic?) temperatures, and be monitored by the most sophisticated analytical tools available in chemistry and microbiology.

Oh what a big If : Will, at some point of the LUCA Revival Project, the contents of some culture bottle ("a warm little pond… with all sorts of salts… heat etc. present") sustain a metabolism and even be able to multiply for a little while? Oh what a wonderful Yes! How, then, fellow biologists, do we name the new resurrected species of ancestor or ancestress: *Luca resurrectus* or *resurrecta*?

What a big If : Certainty out of variety

It is often said that all the experiments and facts described above do not and, indeed, cannot prove that cells capable of energy production, anabolic metabolism, and genetic reproduction (i.e., present-day life) could ever have arisen from inanimate chemical matter. Such proof is compounded by the extreme complexity of living systems on one hand, and gaps and inconsistencies that remain in our knowledge of prebiotic chemistry and self-assembly on the other hand. Low yields of sugar formation, poor catalytic power of random polypeptides, and the tiny probabilities calculated for the emergence of large functional polymers are among the purported problems. Such objections, repeated again and again by creationists and other skeptics are fading and will continue to fade as more and more reasonable synthetic pathways, geochemically plausible conditions, and thermodynamic details of molecular interactions become known. Understanding spontaneous structure formation even in large, oligomeric biomolecules such as the ribosome or membrane-integrated electron transport chains have been more recent major advances. Nevertheless, there are still basic problems and misconceptions in our thinking about molecular and cellular evolution that need to be addressed.

Has there been enough time?

Humans are used to thinking about the past in years, centuries, and millennia and possibly even a million years when confronted with geological and paleontological records. Time spans of several thousand million (10^9) years, such as the age of our universe and the solar planets including Earth, are beyond our imagination. Therefore, it is difficult for many to conceive that astrophysical and

chemical processes which are *very* slow on our human timescale have a major impact on atmospheric and terrestrial environments. These can create adverse or, in other situations, favorable conditions over *very long* periods of time (i.e., hundreds of million years) that are sufficient to produce ever more complex molecules stable enough to accumulate and eventually generate cell-like, growing, and multiplying vesicles.

A more precise, statistical argument has been that the age of the Earth was simply not sufficient to allow assembly of oligomeric and polymeric biomolecules from the appropriate monomers *by chance* when the huge number of all possible combinations is considered. Experience among protein chemists and recent progress in prebiotic peptide research (Rode 1999) show that specific monomer alignments are intrinsically preferred, thus reducing the numbers of possible combinations and significantly weakening the argument. Insufficient time for molecular evolution on planet Earth may indeed be a pseudoproblem.

How likely, or unlikely, is the assembly of a cell "from scratch"?

Critics of chemical and cellular evolution sometimes resort to probability calculations. They hope to rule out the aggregation of a self-reproducing cellular body, subcellular fraction, or large molecules without support from unknown, supernatural forces. For example, in a popular German science magazine (1982), the chances for a scenario in which "correct" macromolecules like genes assemble de novo from a collection of atoms or small monomers were figured to be as low as $1{:}1^{1,200,000}$. Not surprisingly, the author concluded that such an origin of life was statistically impossible.

Naïve approaches like this neglect well-established facts. Complex biomolecules, large or small, are built in a modular fashion. Smaller, submolecular structures and functional domains originate separately and then combine, as in the following examples: A heterocyclic base and the sugar phosphate moiety of a nucleotide come together ready-made from different pathways. In proteins, stable subdomains such as fibrous α-helical stretches, globular β-barrel structures, hydrophobic transmembrane regions, or sites specific for binding coenzymes, carbohydrates, or enzyme substrates can be clearly distinguished. Functionally and structurally separate subsections are found in the cloverleaf structure of tRNAs (the amino acid-accepting and the anticodon stem), in ribozymes, and in DNA strands ("genes in pieces"). Many enzymes specifically incorporate magnesium, calcium, iron, zinc, or other metal ions after their synthesis to reach their final activity or stability. Membranes assemble spontaneously due to hydrophobic interactions.

It is obvious that assembling macromolecules from such smaller preformed components in solution or inside the small volume of a lipid vesicle increases the overall probability of establishing favorable new properties in a population of organic molecules by many orders of magnitude. It is also obvious that a sound knowledge of extant biomolecules, including their syntheses, structures, and functions, is required to make reasonable predictions regarding the origin and properties of probable prebiotic progenitors. It seems clear that objections raised by critics who lack or deny such background knowledge can be disregarded.

This is an extensive assertion. Therefore, let us illustrate and quantify the situation with one more specific yet central problem in a statistically proper way. How could cell-like compartments acquire genetic systems made of uniformly configurated and linked monomers if the soup or slime of abiotically formed constituents contained racemic mixtures of ribose and other sugars, heterocyclic bases unsuitable for base pairing, and "wrong" internucleotide bonds (Kuhn 1972)? If at each single position of a growing oligomer chain the probability of incorporation of a "correct" D-ribose, 3′–5′ linkage, and a proper base was only 1/100, then it would be $(1/100)^{10}=10^{-20}$ for a uniform, "useful" decanucleotide and 10^{-100} for a 50-mer (the size of a small tRNA). That a molecule of the latter kind would ever form by chance is indeed unrealistic. However, the first probability is *not* too small to be feasible as Avogadro's number says that 1 mmol of random decanucleotides (about 3 g of abiotic matter) has 6×10^{20} molecules. Thus, the "unlikely" formation of a correctly structured, base-paired decanucleotide fit for replication or another function (Fig. 9) *must* occur and generate new potential. Evolution proceeded in *many small* steps, in parallel or succeeding each other. Students interested in both biological and physical chemistry can simulate and verify this claim on their computers at any time (Försterling et al. 1972; Kuhn et al. 2009). Critics should familiarize themselves with such straightforward ways of thinking, too.

Will we ever and, indeed, do we have to know the historical path of chemical evolution? This may be the most interesting and fascinating question in evolution research. We wish to satisfy our inborn sense of curiosity and our desire for universal knowledge of this world. However, surprisingly, the answer to both questions is "No."

The first "No" is obvious. Progress will be made, but the heritage of 10^9 years of the most complex abiotic, biogenic, and biological chemistry on Earth is not amenable to complete unraveling by us now or in subsequent generations. Even with the very best sets of experimental data available, nobody will be able to define which of innumerable possible steps of synthesis and decay and chaotic and stagnant moments *really* happened. It is also doubtful whether machines can ever do it.

The second "No" is of more interest. It should be possible someday to identify at least *one* model sequence of events,

hypothetical but supported by several lines of solid evidence, which rigorously satisfies all molecular, thermodynamic, and kinetic constraints and requirements. We would still not know the historical path but no one could continue to argue that an evolution of life driven by chance and necessity was *impossible* for objective reasons. Considering the multiple, redundant physical and molecular possibilities now known and described above—not by far comprehensively covered in our review—the authors are certain that we are approaching the stage of enabling transitions from inanimate matter to a replicating entity inside artificial vesicles. Computer studies of intermolecular networks "in silico" have their place, but bright new ideas for intelligent experiments in vitro and in vivo are, at the very least, equally rewarding.

Epilog

In this review, we have listed around a hundred critical steps, reactions, and consequences that must be analyzed and combined in order to proceed from speculation to an increasingly rational understanding of the chemical evolution that occurred very early in the history of our planet. The dictate of physics and chemistry over molecular aggregates and cells did not end after LUCA had started to proliferate, mutate, and "change the world." They continued to exert their influence on energy metabolism and reproduction during the transition from anoxic, reducing conditions to an oxidizing environment when light absorption by chlorophylls and oxygenic photosynthesis (splitting of water to protons, electrons, and molecular oxygen) became established about 2,000 mya. Photosystems and oxygen permitted efficient energy production via electron and proton potentials in specialized intracellular membranes in bacteria, some of which became the energy-providing endosymbionts of eukaryotic cells (Kutschera and Niklas 2004). Multicellular organisms began to step out of the waters and master terrestrial life. These epoch-making events during later eons of the evolution of life on Earth are not the subject of this review. They are in part dealt with in the following papers.

One of the most complex organisms, *Homo sapiens*, is often highly critical of thinking that he or she should be no more than part of an early chemical and long biological evolution that included incredibly numerous, accidental, and random steps. It is, therefore, still believed by some that supernatural principles have to be cited to satisfy the long-cherished yet irrational desire of man to be unique among all other things; the authors do not agree that any are required. As biochemical and biophysical research and reasoning begin to understand even extremely complex physiological phenomena, such as consciousness and memory, sexual differentiation, or aging, to name but a few, there is hope for more informed and tolerant relations between us and our biological descent.

Charles Darwin has been admired, scolded, and ridiculed for 150 years. Next to his *Origin of Species* and *Descent of Man*, the picture of the "warm little pond," probably coined incidentally, has become one of the more popular quotations. The rhetorical question "Goodbye to the Warm Little Pond?" asked by Science in 1990 after the advent of the iron–sulphur surface world idea is not only enjoyable reading but is once again a testament to Darwin's ongoing, amazing intellectual impact on science and culture.

Acknowledgements Vivid discussions with colleagues and friends, in particular Horst-Dieter Försterling, Peter Kaiser, Hans Kuhn, Ulrich Kutschera, Stanley Miller (deceased; La Jolla, CA, USA), Klaus-Heinrich Röhm, and Günter Wächtershäuser are greatly appreciated. The students attending summer academies on evolution at La Villa and Alpbach (Tyrol), sponsored by Studienstiftung, contributed many critical questions and stimulating ideas.

References

Abramov O, Mojzsis SJ (2009) Microbial habitability of the Hadean Earth during the late heavy bombardment. Nature 459:419–422

Bada JL (2001) State-of-art instruments for detecting extraterrestrial life. Proc Natl Acad Sci U S A 98:797–800

Beaulieu JP, PLANET Collaboration (2006) Discovery of a cool planet of 5.5 Earth masses through gravitational microlensing. Nature 439:437–439

Benner SA, Ellington AD (1987) The last ribo-organism. Nature 329:295–296

Bergmann ED, Pullman B (eds) (1972) The purines. Theory and experiment. The Israel Academy of Sciences and Humanities, Jerusalem

Blair NE, Bonner WA (1981) A model for the enantiomeric enrichment of polypeptides on the primitive earth. Orig Life 11:331–335

Böhler C, Nielsen PE, Orgel LE (1995) Template switching between PNA and RNA oligonucleotides. Nature 376:578–581

Bondy SC, Harrington ME (1979) L-Amino acids and D-glucose bind stereospecifically to a colloidal clay. Science 203:1243–1244

Bossard AR, Raulin F, Mourey D, Toupance G (1982) Organic synthesis from reducing models of the atmosphere of the primitive earth with UV light and electric discharges. J Mol Evol 18:173–178

Boussau B, Blanquart S, Necsuela A, Lartillot N, Gouy M (2008) Parallel adaptations to high temperatures in the Archaen eon. Nature 456:942–945

Brandes JA, Boctor NZ, Cody GD, Cooper BA, Hazen RM, Yodor HS (1998) Abiotic nitrogen reduction on the early Earth. Nature 395:365–367

Calvin M (1969) Chemical evolution. Oxford University Press, Oxford

Carroll SB (2001) Chance and necessity: the evolution of morphological complexity and diversity. Nature 409:1102–1109

Cech TR (1987) The chemistry of self-splicing RNA and RNA enzymes. Science 236:1532–1539

Chyba C (1997) A left-handed solar system? Nature 389:234–235

Chyba C, Sagan C (1992) Endogenous production, exogenous delivery and impact-shock synthesis of organic molecules: an inventory for the origins of life. Nature 355:125–132

Cleaves HJ, Chalmers JH, Lazcano A, Miller SL, Bada JL (2008) A reassessment of prebiotic organic synthesis in neutral planetary atmospheres. Orig Life Evol Biosph 38:105–115

Cody GD, Boctor NZ, Filley TR, Hazen RM, Scott JH, Sharma A, Yoder HS (2000) Primordial carbonylated iron–sulfur compounds and the synthesis of pyruvate. Science 289:1337–1340

Darwin C (1859) The origin of species by means of natural selection or the preservation of favoured races in the struggle for life. Murray, London

Drobner E, Huber H, Wächtershäuser G, Rose D, Stetter KO (1990) Pyrite formation linked with hydrogen evolution under anaerobic conditions. Nature 346:742–744

de Duve C (1987) Selection by differential molecular survival: a possible mechanism of early chemical evolution. Proc Natl Acad Sci U S A 84:8253–8256

de Duve C (1991) Blueprint for a cell: the nature and origin of life. Neil Patterson, Burlington

Ducluzeau AL, van Lis R, Duval S, Russell MJ, Nitschke W (2009) Was nitric oxide the first deep electron sink? Trends Biochem Sci 34:9–15

Eigen M (1971) Selforganization of matter and the evolution of biological macromolecules. Naturwissenschaften 58:465–523

Eigen M (1981) Darwin und die Molekularbiologie. Angew Chem 93:221–229

Eisner JA (2007) Water vapour and hydrogen in the terrestrial-planet-forming region of a protoplanetary disk. Nature 447:562–564

Engel MH, Macko SA (1997) Isotopic evidence for extraterrestrial non-racemic amino acids in the Murchison meteorite. Nature 389:265–268

Engel MH, Nagy B (1982) Distribution and enantiomeric composition of amino acids in the Murchison meteorite. Nature 296:837–840

Eschenmoser A, Dobler M (1992) Warum pentose- und nicht hexose-Nucleinsäuren? Teil I. Helv Chim Acta 75:218–259

Fegley B, Prinn RG, Hartman H, Watkins GH (1986) Chemical effects of large impacts on the Earth's primitive atmosphere. Nature 319:305–307

Ferris JP, Ertem G (1993) Montmorillonite catalysis of RNA oligomer formation in aqueous solution. J Am Chem Soc 115:12270–12275

Ferris JP, Hagan WJ (1984) HCN and chemical evolution: the possible role of cyano compounds in prebiotic synthesis. Tetrahedron 40:1093–1120

Ferris JP, Hill AR, Liu R, Orgel LE (1996) Synthesis of long prebiotic oligomers on mineral surfaces. Nature 381:59–61

Fletcher SP, Jagt RBC, Feringa BL (2007) An astrophysically relevant mechanism for amino acid enantiomer enrichment. Chem Commun 2007:2578–2580

Follmann H (1981) Chemie und Biochemie der evolution. Quelle und Meyer, Heidelberg

Follmann H (1982) Deoxyribonucleotides and the emergence of DNA in molecular evolution. Naturwissenschaften 69:75–81

Follmann H (1986) Have deoxyribonucleotides and DNA been among the earliest biomolecules? Adv Space Res 6:33–38

Follmann H (2004) Deoxyribonucleotides: the unusual biochemistry and chemistry of DNA precursors. Chem Soc Rev 33:225–233

Folsome C, Brittain A (1981) Model protocells photochemically reduce carbonate to organic carbon. Nature 291:482–484

Försterling HD, Kuhn H, Tews KH (1972) Computermodell zur Bildung selbstorganisierender Systeme. Angew Chem 84:862–865

Fox SW (1980) Metabolic microspheres. Origins and evolution. Naturwissenschaften 67:378–383

Galtier N, Tourasse N, Gouy M (1999) A non-hyperthermophilic common ancestor to extant life forms. Science 283:220–221

Gaucher EA, Thomson JM, Burgan MF, Benner SA (2003) Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins. Nature 425:285–288

Gaucher EA, Govindarajan S, Ganesh OK (2008) Palaeotemperature trend for Precambrian life inferred from resurrected proteins. Nature 451:704–707

Gesteland RF, Cech TR, Atkins JF (eds) (2006) The RNA world, 3rd edn, chapters 1–3, 7. Cold Spring Harbor Laboratory Press, Cold Spring Harbor

Gil R, Silva FJ, Peretó J, Moya A (2004) Determination of the core of a minimal bacterial gene set. Microbiol Mol Biol Rev 68:518–537

Gilbert W (1986) The RNA world. Nature 319:618

Goesmann F, Rosenbauer H, Roll R, Szopa C, Raulin F, Sternberg R, Israel G, Meierhenrich U, Thiemann W, Munoz-Caro G (2007) COSAC, the cometary sampling and composition experiment. Space Sci Rev 128:257–280

Griffiths G (2007) Cell evolution and the problem of membrane topology. Nat Rev Mol Cell Biol 8:1018–1024

Gutfraind A, Kempf A (2008) Error-reducing structure of the genetic code indicates code origin in non-thermophilic organisms. Orig Life Evol Biosph 38:75–85

Halliday AN (2001) In the beginning. Nature 409:144–145 and references quoted therein

Hargreaves WR, Mulvihill SJ, Deamer DW (1977) Synthesis of phospholipids and membranes in prebiotic conditions. Nature 266:78–80

Hardin G (1950) Darwin and the heterotroph hypothesis. Sci Mon 70:178–179

Hartmann J, Brand MC, Dose K (1981) Formation of specific amino acid sequences during thermal polymerization of amino acids. BioSystems 13:141–147

Hayes JM (1996) The earliest memories of life on Earth. Nature 384:21–22

Hazen RM, Filley TR, Goodfriend GA (2001) Selective adsorption of L- and D-amino acids on calcite: implications for biochemical homochirality. Proc Natl Acad Sci U S A 98:5487–5490

Huber C, Wächtershäuser G (1997) Activated acetic acid by carbon fixation on (Fe, Ni)S under primordial conditions. Science 276:245–247

Huber C, Wächtershäuser G (1998) Peptides by activation of amino acids with CO: implications for the origin of life. Science 281:670–672

Huber C, Wächtershäuser G (2006) Hydroxy and amino acids under possible Hadean, volcanic origin-of-life conditions. Science 314:630–632

Ishikawa K, Sato K, Shima Y, Urabe I, Yomo T (2004) Expression of a cascading genetic network within liposomes. FEBS Lett 576:387–390

Jacobsen SB (2003) How old is planet Earth? Science 300:1513–1514

Janda M, Morvova M, Machala Z, Morva I (2008) Study of plasma induced chemistry by DC discharges in $CO_2/N_2/H_2O$ mixtures above a water surface. Orig Life Evol Biosph 38:23–35

Johnson AP, Cleaves HJ, Dworkin JP, Glavin DP, Lazcano A, Bada JL (2008) The Miller volcanic spark discharge experiment. Science 322:404

Joyce GF (1998) Nucleic acid enzymes: playing with a fuller deck. Proc Natl Acad Sci U S A 95:5845–5847

Jungck JR, Fox SW (1973) Synthesis of oligonucleotides by proteinoid microspheres acting on ATP. Naturwissenschaften 60:425–427

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

Keefe AD, Miller SL (1995) Are polyphosphates or phosphate esters prebiotic reagents? J Mol Evol 41:693–702

Keefe AD, Miller SL, McDonald G, Bada J (1995a) Investigation of the prebiotic synthesis of amino acids and RNA bases from CO_2 using FeS/H_2S as a reducing agent. Proc Natl Acad Sci U S A 92:11904–11906

Keefe AD, Newton GL, Miller SL (1995b) A possible prebiotic synthesis of pantetheine, a precursor to coenzyme A. Nature 373:683–685

Kleinkauf H, van Liempt H, Palissa H, von Döhren H (1992) Biosynthese von Peptiden: Ein nichtribosomales system. Naturwissenschaften 79:153–162

Krishnamurthy R, Pitsch S, Minton M, Miculka C, Windhab N, Eschenmoser A (1996) Pyranosyl-RNA. Angew Chem 108:1619–1623

Kuhn H (1972) Selbstorganisation molekularer Systeme und die evolution des genetischen Apparates. Angew Chem 84:838–862

Kuhn H (1976) Model considerations for the origin of life. Naturwissenschaften 63:68–80

Kuhn H, Waser J (1981) Molekulare Selbstorganisation und Ursprung des Lebens. Angew Chem 93:495–515

Kuhn H, Försterling HD, Waldeck DH (2009) Principles of physical chemistry, 2nd edn, chapter 29: Origin of life - matter carrying information. Wiley-VCH, Weinheim

Kutschera U (2009) Charles Darwin's Origin of Species, directional selection, and the evolutionary sciences today. Naturwissenschaften doi:10.1007/s00114-009-0603-0

Kutschera U, Niklas KJ (2004) The modern theory of biological evolution: an expanded synthesis. Naturwissenschaften 91:255–276

Kvenvolden KA (1974) Geochemistry and the origin of life. Benchmark papers in geology, vol. 14. Dowden, Hutchinson & Ross, Stroudsburg

Laerdahl JK, Wesendrup R, Schwerdtfeger P (2000) D- or L-Alanine: that is the question. ChemPhysChem 2000:60–62

Lahav N, White D, Chang S (1978) Peptide formation in the prebiotic era. Science 201:67–69

Lazcano A, Miller SL (1994) How long did it take for life to begin and evolve to cyanoabacteria? J Mol Evol 39:546–554

Lee DH, Granja JR, Martinez JA, Severin K, Ghadiri MR (1996) A self-replicating peptide. Nature 382:525–528

Leman L, Orgel LE, Ghadiri MR (2004) Carbonyl sulfide-mediated prebiotic formation of peptides. Science 306:283–286

Leman LJ, Orgel LE, Ghadiri MR (2006) Amino acid dependent formation of phosphate anhydrides in water mediated by carbonyl sulfide. J Am Chem Soc 128:20–21

Lemmon RM (1970) Chemical evolution. Chem Rev 70:95–109

Löb W (1913) Über das Verhalten des Formamids unter der Wirkung der stillen Entladung. Ein Beitrag zur Frage der Stickstoff-Assimilation. Chem Ber 46:684–697

Lohrmann R, Bridson PK, Orgel LE (1980) Efficient metal-ion catalysed template-directed oligonucleotide synthesis. Science 208:1464–1465

Mansy SS, Schrum JP, Krishnamurty M, Tobé S, Treco DA, Szostak JW (2008) Template-directed synthesis of a genetic polymer in a model protocell. Nature 454:122–125

Mar A, Oró J (1991) Synthesis of the coenzymes ADP-glucose, GDP-glucose, and CDP-ethanolamine under primitive earth conditions. J Mol Evol 32:201–210

McLoughlin N, Brasier MD, Wacey D, Green OR, Perry R (2007) On biogenecity criteria for endolithic microborings on early Earth and beyond. Astrobiology 7:10–26

Meierhenrich U, Munoz Caro G, Bredehöft JH, Jessberger LK, Thiemann W (2004) Identification of diamino acids in the Murchison meteorite. Proc Natl Acad Sci U S A 101:9182–9186

Miescher F (1871) Über die chemische Zusammensetzung der Eiterzellen. Hoppe-Seylers Medizinische-Chemischen Untersuchungen 4:441–460

Millar TJ (2004) Organic molecules in the interstellar medium. In: Ehrenfreund P et al (eds) Astrobiology: future perspectives. Kluwer Academic, The Netherlands, pp 17–31

Miller SL (1953) A 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, Lazcano A (1995) The origin of life—did it occur at high temperature? J Mol Evol 41:689–692

Miller SL, Urey HC (1959) Organic compound synthesis on the primitive earth. Science 130:245–251

Mills DR, Petersen RL, Spiegelman S (1967) An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. Proc Natl Acad Sci U S A 59:217–224

Miyakawa S, Yamanashi H, Kobayashi K, Cleaves HJ, Miller SL (2002) Prebiotic synthesis from CO atmospheres: iImplications for the origin of life. Proc Natl Acad Sci U S A 99:14628–14631

Monnard PA, Oberholzer T, Luisi PL (1997) Entrapment of nucleic acids in liposomes. Biochim Biophys Acta 1329:39–50

Monod J (1970) Le hasard et la nécessité. Editions du Seuil, Paris; Zufall und Notwendigkeit (1971) Piper-Verlag München; Chance and necessity (1971) Knopf, New York

Munoz Caro GM, Meierhenrich U, Schutte WA, Barbier B, Arcones Segovia A, Rosenbauer H, Thiemann W, Brack A, Greenberg JM (2002) Amino acids from ultraviolet irradiation of interstellar ice analogues. Nature 416:403–406

Mushegian AR, Koonin EV (1996) A minimal gene set for cellular life derived by comparison of complete bacterial genomes. Proc Natl Acad Sci U S A 93:10268–10273

Nelson KE, Levy M, Miller SL (2000) Peptide nucleic acids rather than RNA may have been the first genetic molecule. Proc Natl Acad Sci U S A 97:3868–3871

Nissen P, Hansen J, Ban N, Moore PB, Steitz TA (2000) The structural basis of ribosome activity in peptide bond synthesis. Science 289:920–930

Nomura M, Erdmann V (1970) Reconstitution of ribosomal subunits from dissociated molecular components. Nature 228:744–748

Oberbeck V, Fogleman G (1989) Impacts and the origin of life. Nature 339:434

Okihana H, Egami F (1979) Polymers produced by heating an amino acid mixture in sea water enriched with transition elements. Orig Life 9:171–180

Oparin AI (1924) Proiskhozdenie zhizny (Origin of Life). Izd. Moskovski Rabochii, Moscow

Oparin AI (1936) Origin of life (Moscow, in Russian); (1938, Macmillan, New York, in English)

Oparin AI (1957) Die Entstehung des Lebens auf der Erde. VEB Verlag der Wissenschaften, Berlin

Orgel LE, Crick HC (1980) Selfish DNA: the ultimate parasite. Nature 284:604–607

Otting G, Billeter M, Wüthrich K, Roth HJ, Leumann C, Eschenmoser A (1993) Warum pentose- und nicht hexose-Nucleinsäuren? Teil IV. Helv Chim Acta 76:2701–2755

Paecht-Horowitz M (1976) Clays as possible catalysts for peptide formation in the prebiotic era. Orig Life 7:369–381

Paecht-Horowitz M, Eirich FR (1988) The polymerization of amino acid adenylates on sodium montmorillonite with preadsorbed polypeptides. Orig Life Evol Biosph 18:359–387

Pasek M (2008) Rethinking early Earth phosphorus geochemistry. Proc Natl Acad Sci U S A 105:853–858

Pasek M, Lauretta D (2008) Extraterrestrial flux of potentially prebiotic C, N, and P to the early earth. Orig Life Evol Biosph 38:5–21

Penzlin H (2009) The riddle of "life", a biologist's critical view. Naturwissenschaften 96:1–23

Pfeil E, Ruckert H (1961) Die Bildung von Zuckern aus Formaldehyd unter der Einwirkung von Laugen. Liebigs Ann Chem 641:121–131
Pflug HD, Jaeschke-Boyer H (1979) Combined structural and chemical analysis of 3,800-Myr-old microfossils. Nature 280:483–486
Piccirilli JA (1995) RNA seeks its maker. Nature 376:548–549
Pizzarello S, Weber AL (2004) Prebiotic amino acids as asymmetric catalysts. Science 303:1151
Plankensteiner K, Reiner H, Rode BM (2006) Amino acids on the rampant primordial earth: electric discharges and the hot salty ocean. Mol Divers 10:3–7
Porco CC, Baker E, Barbara J, Beurle K, Brahic A, Burns JA, Charnoz S, Cooper N, Dawson DD, Del Genio AD, Denk T, Dones L, Dyunida U, Evans MW, Fussner S, Giese B, Grazier K, Helfenstein P, Ingersoll AP, Jacobson RA, Johnson TV, McEwen A, Murray CD, Neukum G, Owen WM, Perry J, Roatsch T, Spitale J, Squyres S, Thomas P, Tiscareno M, Turtle E, Vasavada AR, Veverka J, Wagner R, West R (2005) Imaging of Titan from the Cassini spacecraft. Nature 434:159–168
Powner MW, Gerland B, Sutherland JD (2009) Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. Nature 459:239–242
Rajamani S, Vlassov A, Benner S, Coombs A, Olasagasti F, Deamer D (2008) Lipid-assisted synthesis of RNA-like polymers from mononucleotides. Orig Life Evol Biosph 38:57–74
Reiner H, Plankensteiner K, Fitz D, Rode BM (2006) The possible influence of L-histidine on the origin of the first peptides on the primordial earth. Chem Biodivers 3:611–621
Ricardo A, Carrigan MA, Olcott AN, Benner SA (2004) Borate minerals stabilize ribose. Science 303:196
Robertson MP, Miller SL (1995) An efficient prebiotic synthesis of cytosine and uracil. Nature 375:772–774
Rossmann MG, Moras D, Olsen KW (1974) Chemical and biological evolution of a nucleotide-binding protein. Nature 250:194–199
Rode BM (1999) Peptides and the origin of life. Peptides 20:773–786
Rode BM, Schwendinger MG (1990) Copper-catalyzed amino acid condensation in water—a simple way of prebiotic peptide formation. Orig Life Evol Biosph 20:401–410
Rohlfing DL (1976) Thermal polyamino acid synthesis at less than 100°C. Science 193:68–70
Russell MJ, Daniel RM, Hall AJ, Sherringham JA (1994) A hydrothermically precipitated iron sulphide membrane as a first step towards life. J Mol Evol 39:231–243
Sagan C, Chyba C (1997) The early faint sun paradox: organic shielding of ultraviolet-labile greenhouse gases. Science 276:1217–1221
Schidlowski M (1988) A 3,800-million-year isotopic record of life from carbon in sedimentary rocks. Nature 333:313–318
Schoning K, Scholz P, Guntha S, Wu X, Krishnamurthy R, Eschenmoser A (2000) Chemical etiology of nucleic acid structure: the α-threofuranosyl-(3′–2′)oligonucleotide system. Science 290:1347–1351
Schopf JW (1993) Microfossils of the early Archaen apex chert: new evidence of the antiquity of life. Science 260:640–646
Schopf JW (1999) Cradle of life. Earth's earliest fossils. Princeton University Press, Princeton
Schrauzer GN, Strampach N, Hui LN, Palmer MR, Salehi J (1983) Nitrogen photoreduction on desert sands under sterile conditions. Proc Natl Acad Sci U S A 80:3873–3876
Schwendinger MG, Rode BM (1992) Investigations on the mechanism of salt-induced peptide formation. Orig Life Evol Biosph 22:349–359
Seel F, Schinnerling F (1978) Die Cyanat-induzierte Umwandlung von Calciumhydrogen-phosphat in Calciumdiphosphat—eine präbiotische Schlüsselreaktion? Z Naturforsch 33b:373–376
Seel F, Klos KP, Schuh J (1985) Hydrothermale Kondensation von Magnesiumhydrogen-phosphaten zu Magnesiumdiphosphaten. Naturwissenschaften 72:658
Sleep NH, Zahnle KJ, Kasting JF, Morowitz HJ (1989) Annihilation of ecosystems by asteroid impacts on the early Earth. Nature 342:139–142
Spaargaren DH (1985) Origin of life: oceanic genesis, panspermia, or Darwin's warm little pond? Experientia 41:719–727
Strobel SA, Doudna JA (1997) RNA seeing double: close-packing of helices in RNA tertiary structure. Trends Biochem Sci 22:262–265
Stubbe JA, Ge J, Yee CS (2001) The evolution of ribonucleotide reduction revisited. Trends Biochem Sci 26:93–99
Thiemann W (ed) (1981) Generation and amplification of chirality in chemical systems. Origins of life, vol 11. Reidel, Dordrecht
Tian F, Toon OB, Pavlov AA, de Sterck H (2005) A hydrogen-rich early earth atmosphere. Science 308:1014–1017
Tranter GE (1985) Parity-violating differences of chiral minerals and the origin of biomoleculae homochirality. Nature 318:172–173 (1986) Paritätsverletzung: Ursache der biomolekularen Chiralität. Nachr Chem Tech Lab 34:867–870
Tranter GE (1986) Preferential stabilization of the D-sugar series by the parity-violating weak interactions. J Chem Soc Chem Comm 1986:60–61
Unrau PJ, Bartel DP (1998) RNA-catalyzed nucleotide synthesis. Nature 395:260–263
Urey HC (1952) On the early chemical history of the earth and the origin of life. Proc Natl Acad Sci U S A 38:351–363
Usher DA (1977) Early chemical evolution of nucleic acids. Science 196:311–313
Voet AB, Schwartz AW (1983) Prebiotic adenine synthesis from HCN. Evidence for a newly discovered major pathway. Bioorg Chem 12:8–17
Wächtershäuser G (1988) Pyrite formation, the first energy source for life: a hypothesis. Syst Appl Microbiol 10:207–210
Wächtershäuser G (2000) Life as we don't know it. Science 289:1307–1308
Waldrop MM (1989) Did life really start out in an RNA world? Science 246:1248–1249
Walsh C (2001) Enabling the chemistry of life. Nature 409:226–231
Weber AL (1992) Prebiotic sugar synthesis: hexose and hydroxy acid synthesis from glyceraldehyde catalyzed by iron(III) hydroxide oxide. J Mol Evol 35:1–6
Weber AL, Miller SL (1981) Reasons for the occurrence of the twenty coded protein amino acids. J Mol Evol 17:273–284
Weber P, Greenberg JM (1985) Can spores survive in interstellar space? Nature 316:403–407
Whitfield J (2004) Born in a watery commune. Nature 427:674–676
Woese CR (1998) The universal ancestor. Proc Natl Acad Sci U S A 95:6854–6859
Woese CR (2000) Interpreting the universal phylogenetic tree. Proc Natl Acad Sci U S A 97:8392–8396
Wong JT (1975) A co-evolution theory of the genetic code. Proc Natl Acad Sci U S A 72:1909–1912
Wiechert UH (2002) Earth's early atmosphere. Science 298:2341–2342
Yamagata Y, Watanabe H, Saito M, Namba T (1991) Volcanic production of polyphosphates and its relevance to prebiotic evolution. Nature 352:516–519
Zhang L, Peritz A, Meggers E (2005) A simple glycol nucleic acid. J Am Chem Soc 127:4174–4175
Zuckerman B (1977) Interstellar molecules. Nature 268:491–495