1
From Chemical Invariance to Genetic Variability

Günter Wächtershäuser

The archaeologist of nature is at liberty to go back to the traces that remain of nature's earliest revolutions, and, appealing to all he knows or can conjecture about its mechanism, to trace the origin of that great family of creatures … down even to mosses and lichens, and finally down to the lowest perceivable stage of nature, to crude matter. From this and from the forces within, by mechanical laws, like those that are at work in the formation of crystals, seems to be derived the whole technique of nature.

Immanuel Kant [1]

1.1
Heuristic of Biochemical Retrodiction

Darwin (1863) wrote in a letter to Hooker [2]: "It is mere rubbish, thinking at present of the origin of life; one might as well think of the origin of matter." Studies of nucleosynthesis are now quite advanced, but research into the origin of life is still an immature science. The problem of early evolution of life is unique and requires its own heuristic. A commonly used heuristic consists of one-to-one back-extrapolations of individual biochemical features (Figure 1.1a), for which Lipmann [3] coined the term *backward projection*. More and more backward projections add evermore ingredients to the recipe. Inevitably, this way of thinking leads to the notion of a "primordial broth." No one has ever spelled out all that what would or would not have been in the broth and how precisely the organization of life could have come about within such a chaotic situation.

This conceptual hodgepodge is overcome by a heuristic of convergent back-extrapolation, termed *biochemical retrodiction* (Figure 1.1b) [4]. Some extant features are still projected all the way back to the origin. Typically, however, several extant biochemical features are related to one simpler common functional precursor feature and several precursor features are related to a still deeper common precursor. This pattern is applied over and over again, drawing in more and more extant features, progressing to ever deeper, fewer, and simpler precursor features, and generating an overall pattern of backward convergence. Ultimately,
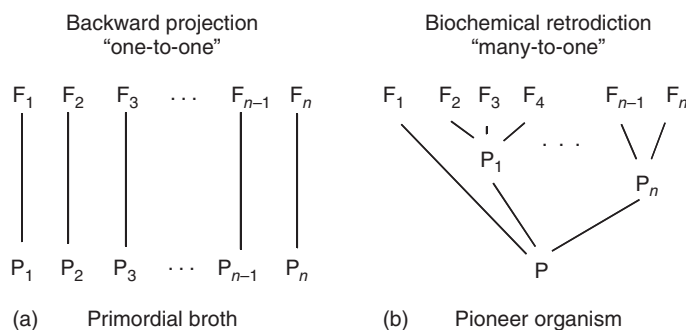
Backward projection "one-to-one" Biochemical retrodiction "many-to-one"

F_1 F_2 F_3 $\cdots$ F_{n-1} F_n F_1 F_2 F_3 F_4 F_{n-1} F_n

P_1

$\cdots$

P_n

P_1 P_2 P_3 $\cdots$ P_{n-1} P_n P

(a) Primordial broth   (b) Pioneer organism

Figure 1.1 Heuristics of (a) parallel backward projection and (b) convergent biochemical retrodiction (F1, F2, etc., extant features; P, P1, P2,etc., precursor features). (From Ref [47] © (2010), Springer.)

the heuristic of biochemical retrodiction aims at a restricted set of chemical compounds and processes, which cooperate to form a distinct chemical entity with the ability to reproduce and evolve: the "pioneer organism." Its chemistry is intrinsically synthetic, thereby imposing from the start directionality from simple to complex on the overall process of evolution. In this sense, the pioneer organism paves the way for all future evolution, hence its name.

Specifically, we include in our platform for retrodiction extant biochemical features, which combine aspects of evolutionary change with aspects that have been largely invariant over time by the universal laws of chemistry. It is precisely the aspect of chemical invariance within evolved biochemical features that provides directionality and allows us to unravel evolutionary history backward to the very pioneer organism of life. For the biochemical retrodiction of multistep biosynthetic pathways we employ in addition the Florkin–Granick rule [5, 6] that earlier steps in a pathway have greater evolutionary antiquity than later steps. This rule is based on the assumption that biosynthetic pathways evolve by terminal extensions. We should apply caution, however, because pathway evolution comprises also lateral branchings, recruitments, reversals, and eliminations [7].

The results of biochemical retrodiction are evaluated empirically by chemical experiments and theoretically by quantum-chemical calculations [8], with the perceived geochemical scenario determining the parameters. Biochemical retrodiction suggests chemical experiments and experimental results inform revised retrodictions, and such iterative procedure promises a progressive exploration of the pioneer organism. Finally, when our procedure leads to competing hypotheses, we prefer the one with the greater explanatory power, that is, the ability to explain a greater number of extant biochemical facts with fewer evolutionary assumptions [9]. We shall now use this methodology for a step-by-step reconstruction of the pioneer organism, beginning at the simplest level: the elements of life.

1.2
Retrodicting the Elements of Life

The elements of central biochemistry [10] fall into two distinct subsets. (i) The main group nonmetal bioelements (H, C, N, O, P, S, Se) make up the bulk of the biomass with mostly structural roles. They originate deep in the mantle of the Earth and form volcanic gases (H_2, N_2, CO_2, CO, CH_4, NH_3, H_2O, SO_2, H_2S, H_2Se, COS, HCN, CH_3SH, P_4O_{10}). (ii) The transition metal bioelements (Fe, Co, Ni, V, Mo, W, Mn, Cu, Zn) occur in organisms only in trace amounts, with mostly catalytic functions. Together with the main group biometals Mg, Ca they form essentially stationary crustal minerals. These two classes of bioelements come into close encounter at volcanic-hydrothermal flow sites in the presence of liquid water [11], cf. [12].

Four main group nonmetal bioelements (H, C, N, O) form the structural basis for biochemistry. They are involved in all biochemical reactions, each one with its unique roles, indispensable for life, down to the pioneer organism. Sulfur and its companion selenium have diverse biocatalytic and bioenergetic functions. They are projected into the pioneer organism. Phosphorus is indispensible in genetics and bioenergetics, but limited to phosphate group chemistry. It must have been acquired after the origin of life. Therefore, the pioneer organism is defined *prima facie* by main group system $H-C-O-N-S-Se$.

Among the transition metal bioelements we find some of the most crucial, indispensible catalysts of central anaerobic biochemistry. Iron in the form of moderately soft ferrous ions has diverse biochemical functions. It is the most abundant transition metal in aqueous, anaerobic, volcanic-hydrothermal settings. In the same settings, hydrogen sulfide (H_2S), the source for soft thio ligands is a ubiquitous volcanic gas. These two locally coinciding components have a high bonding affinity for each other, as evident from the abundance of iron–sulfur clusters in extant metalloenzymes [13] and of iron–sulfur minerals in extant volcanic-hydrothermal flow sites. Therefore, the world of the pioneer organism has been dubbed "*iron–sulfur world*" [7]. The Fe–S bonding strength under anaerobic conditions serves as a gauge for the suitability of other transition metal bioelements, notably Co and Ni. These three iron group metals form the catalytic core of the pioneer organism.

Cu and Mo were unavailable for the anaerobic pioneer organism, because they form extremely insoluble sulfides (Cu_2S and MoS_2). They could have entered the biosphere only after oxygenation of the oceans. Zn is also discounted. Under sulfidic conditions, it forms highly insoluble ZnS and it is not redox-active. Mn (as Mn^{2+}) is too hard for the iron–sulfur world. Cr, the group companion of Mo, exists under volcanic-hydrothermal conditions as hard cation (Cr^{3+}) and had no chance to enter early biochemistry. Compared to its extremely low overall abundance in the Solar System, W is highly enriched in the walls of hydrothermal flow ducts and it does not form an extremely insoluble sulfide [14]. It would have been available for the pioneer organism. Vanadium has a remarkable chemical similarity to its diagonal neighbor Mo [15] and a high crustal abundance without being trapped

as an insoluble sulfide. We conclude that the pioneer organism was catalytically defined *prima facie* by the transition metal system $Fe-Co-Ni-W-(V)$.

1.3
Retrodicting Pioneer Catalysis

Extant biocatalysis is dominated by enzymes. These classify into metalloenzymes and nonmetalloenzymes. Nonmetalloenzymes require a large number of weak group interactions to fold and to stabilize transition states. This means high sequence accuracy, that is, late evolutionary arrival. The accuracy need of metalloenzymes is more relaxed. They typically exhibit a few strong coordination bonds to transition metals, reacting molecules, or transition states in addition to weak group interactions. Effectiveness of such coordination bonds is relatively insensitive to sequence variations. Therefore, a few coordination bonds may have been sufficient for protein folding and enzyme catalysis at the beginning of translation.

Next we note that evolutionary variability is not uniform throughout the structure of a metalloenzyme. Variability increases and invariance decreases from the (innermost) transition metal through the (inner) ligators to the (outer) ligand moieties. Hence, in the course of evolution central transition metals and ligators, once established remained invariant, with rare replacements ($Fe \rightarrow Mo$, V in nitrogenases; $W \rightarrow Mo$ in tungsto/molybdopterins; $Fe \rightarrow Co$, Ni in tetrapyrroles; $S \rightarrow Se$ or $S \rightarrow O$ in $Fe-S$ clusters). The outer protein ligands evolved to modulate the catalytic properties of otherwise invariant transition metals and ligators.

We now apply the heuristic of biochemical retrodiction. We replace the protein ligands and any highly evolved organic ligands by simpler small-molecule ligands with the same ligator (e.g., $CysS \rightarrow HS$). We do so formally without much concern for proof of chemical stabilities of retrodicted intermediates. We arrive ultimately at transition metal complexes/clusters with inorganic ligands that may derive directly from volcanic gases. These small structures are characterized by a low ratio of bridging to nonbridging ligands and by a total lack of periodicity so that the use of the term *nanocrystal* for such a small structure would be utterly misleading. How do these small structures relate to large transition metal minerals with their high connectivity and short-range to long-range periodicity (crystallinity)? In order to find an answer to this question, we should appreciate that volcanic-hydrothermal flow settings are characterized by mineral transformations: thermal dehydration, ligand exchange, hydrolysis, and mobilization by leaching, carbonylation, and cyanidation. These transformations are governed by the Ostwald–Volmer step rule, whereby the withdrawal of energy from a chemical system that can exist in several states of density will proceed stepwise from low to high density (typically = stability). Mineral transformations that obey the principle of minimal structural (topological) change may occur by solid-state reactions. All others occur via mobilization and recrystallization. Aqueous recrystallization begins with endergonic nucleation by transient, fluctuational

aggregation of small, aperiodic clusters. When a critical size is reached, exergonic growth to (nano)crystals sets in. We now are ready to answer our question. We simply view ligand evolution as an adaptation to the intermediate structures *en route* to mineral nucleation. Nucleation intermediates are transient, yet invariant in type. They cannot be directly observed, but they become fixed by protein ligands and thereby become observable. Protein ligands are variable in sequence and they adapt to catalytically competent complexes/clusters. Mineral catalysis is heterogeneous surface catalysis. Metalloenzyme catalysis, from the point of view of the tiny catalytic centers, is homogeneous catalysis. And the evolutionary transition from mineral catalysis to metalloenzyme catalysis must have proceeded from versatile (multifunctional) heterogeneous catalysis on large mineral surfaces to specialized (oligofunctional) complex/cluster catalysis. This evolutionary transition backtraces mineral formation from endergonic nucleation to exergonic crystal growth.

In another vein, we see the evolution of the multistep cyclical process of extant metalloenzyme catalysis as the result of a progression from linear to cyclical transformations. Primordial metallocatalysis is seen as a linear transformation with few reaction steps and as evolving later into a catalytic cycle with an increased number of reaction steps (much like industrial development of cyclical catalysis that also begins typically with the discovery of a simpler, linear transformation). Let us briefly look at some examples.

- *[Fe−S] proteins* contain a family of simple, structurally related [Fe−S] complexes [13] that are biosynthetically connected as follows: $[2Fe-2S]CysS_3L \rightarrow [4Fe\text{-}4S]CysS_3L$ (L = CysS or RO); $[4Fe\text{-}4S]CysS_3L' \rightarrow [Fe3\text{-}S4]CysS_3$ (L' = RO) (R = SerO, AspO, or the like). The retrodiction of analogous inorganic complexes (e.g., CysS $\Leftarrow$ SH; RO $\Leftarrow$ OH) has experimental support [16, 17].
- *Nitrogenases* evolved in the direction [FeFe] $\Rightarrow$ {[FeMo] or [FeV]} nitrogenase. Their catalytic Fe−S−C clusters derive biosynthetically from an ancestral [FeFe] cluster, which in turn derives from two [4Fe-4S] clusters and from AdoMet as C source [18]. They all have the same redox-active iron−sulfur P cluster. We project all nitrogenases back to an inorganic [Fe−S] system of volcanic-hydrothermal provenance.
- The *[Fe−Ni−S] enzymes* carbon monoxide dehydrogenase (anaerobic) (CODH) and acetyl-CoA synthase (ACS) [19] may be traced back to an inorganic [Fe−Ni−S] system, with ligands derived from H_2S, H_2O, NH_3 [20].
- *Molybdopterin* and *tungstopterin* enzymes of extant biochemistry trace back to an ancestral family of common precursor enzymes that were exclusively endowed with oxygen-sensitive tungstopterin cofactors [14]. With oxygenation of the surface of the Earth the fastidiously anaerobic tungstopterin enzymes receded into anaerobic niches and the aerobic world became molybdenized [14]. Among the tungstoenzymes that are holdovers to this day, we find extraordinary catalytic abilities, for example, direct reduction of carboxylic acids [21] or hydration of acetylene [22]. Tungstopterins are traced back to inorganic tetrathiotungstate dianions $[WS_4]^{2-}$ [4] with coexistent thiovanadate anions ($VO_xS_{4-x}^{3-}$, $x = 0-3$).

Such retrodictions are specific enough to benefit from, and find support in the study of (aqueous or nonaqueous) model compounds [23, 24]. In a broader context, biocatalyst retrodictions should be correlated with global geochemical developments: for example, cooling of the mantle and crust; depletion of nickel in the ocean; oxygenation of the ocean and atmosphere [25]; onset of aerobiosis; and onset of molybdenization of the biosphere [14]. In another broad sense, we see two causes for complications: (i) habitats became dislocated away from volcanic-hydrothermal sites and (ii) formerly nutritious volcanic-hydrothermal components became toxic (e.g., H_2S). Both causes led to the emergence of strategies for biosynthesis of nutrients from metabolites (e.g., from cysteine by desulfurases [26]) or from other nutrients; or strategies for the conversion of one ligand into another ligand – all under highly controlled scaffolding conditions. Let us look at the biochemistry of H_2.

- The three *hydrogenases* have deep dissimilarities, but also startling commonalities in terms of active site $Fe-(CO)-(CN)$ coordination. It is frequently concluded that their commonalities are "convergent" and not due to a common history. This conclusion is obscure and, strictly speaking, not supported by facts. Surely, the protein structures do not reveal common ancestry at the level of the last universal common ancestor (LUCA). Their catalytic clusters differ in terms of structure and biosynthesis. A deeper, pretranslational common ancestry, however, may well be a realistic assumption (Figure 1.2). We shall attempt to trace the evolution of the three hydrogenases back into invariant inorganic chemistry.
- [*Fe*] *hydrogenase* (anaerobic) converts H_2 in *one* step into H^+ and a H^- ligand that is transferred directly and reversibly to the target substrate (methenyl-H_4MPT) [28]. The redox-inactive Fe center has two CO ligands, which are derived from CO_2 via an internal CO pool [29]. The bidentate organic guanylylpyridinol ligand may be a replacement for inorganic ligands: pyridine-*N* for cyano [30] and acyl-CO for CO [28]. This leads us to a retrodicted ancestral $Fe(CO)_3(CN)(SH)$ complex that may be formally related to the carbonyl $Fe(CO)_5$ (Figure 1.2).
- [*FeFe*] *hydrogenase* (strictly anaerobic) has a redox-active $[Fe-Fe]$ center and converts H_2 into *two* e^- (via intermediate H^-). These are transferred efficiently to a $[4Fe-4S]$ cluster that is covalently attached via its CysS ligand as the bridge [31]. The set of inorganic ligands comprises CN^-, CO, and $-S-CH_2-NH-CH_2-S-$. They all derive from dehydroglycine (formed from tyrosine by an AdoMet-dependent radical mechanism) [32, 33]. Retrodiction of an inorganic ancestral cluster (Figure 1.2) is based on chemical synthesis (Rauchfuss reaction) [34]. Incidentally, dehydroglycine is related to the formation of glycine from a CN/CO ligand system [27].
- [*NiFe*] *hydrogenases* have a redox-active $[Ni-Fe]$ center. It converts H_2 into *two* e^- (via intermediate H^-) that are transferred via [4Fe-4S] clusters. Its catalytic $Fe(CN)_2(CO)$ moiety is preassembled on a scaffold [35], with CN^- (from carbamoyl phosphate [36]) added before CO (from a metabolite-derived internal
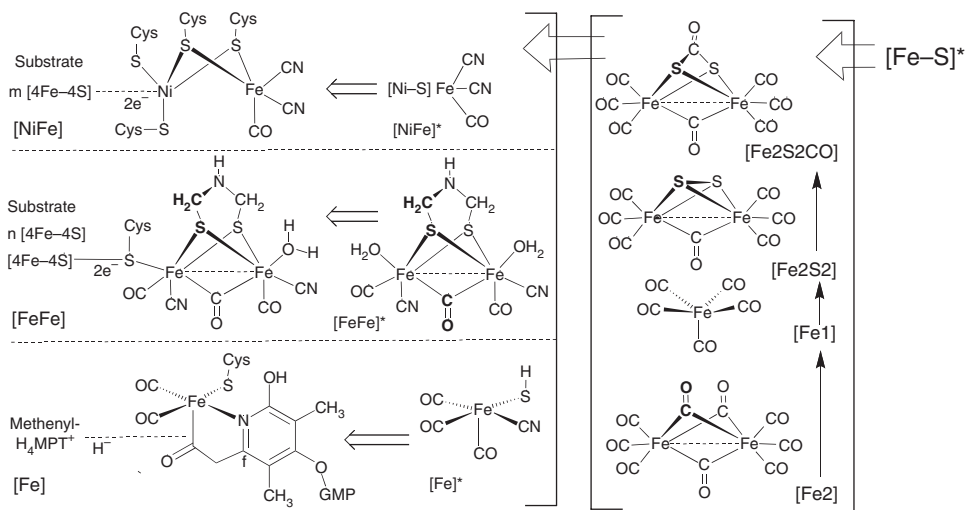
Figure 1.2 Common origin of hydrogenases in a set of carbonylated precursors ([Fe1] = Fe(CO)$_5$) in pseudorotation; [Fe2S2CO] forms from [Fe2] under primordial conditions [27].
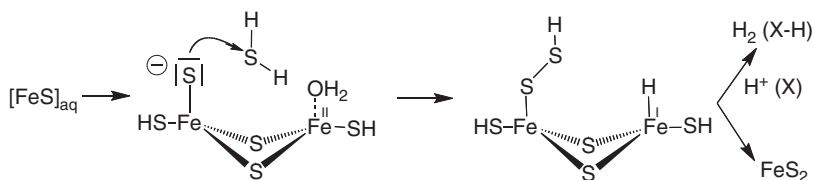
Figure 1.3 Notional mechanism of pyrite formation from FeS/H$_2$S (X = substrate).

CO pool [37]). This leads us to the ancestral structure [Ni−S]Fe(CN)$_2$(CO) (Figure 1.2).

All three hydrogenases may ultimately go back to the conversion of FeS/H$_2$S to FeS$_2$/H$_2$ [38, 39], for which a mechanism is postulated (Figure 1.3) that starts from aqueous [FeS]$_{aq}$ [40] and assumes a nucleophilic attack of a sulfide ligand on the sulfur atom of H$_2$S to form a persulfide ligand (as source for FeS$_2$) [40] in combination with either formation of H$_2$ by reaction of a hydride ligand with H$^+$ or reduction of a substrate X by hydride or electron transfer. From this ur-hydrogenase/reductase chemistry, the evolution of hydrogenases began by abandonment of irreversible pyrite formation in favor of catalytic cycling, reversal of H$_2$-formation into H$_2$-uptake, and substrate reduction (e.g., persulfide cleavage).

It proceeded to carbonylation with volcanic CO (cf. carbonyl transformations shown in brackets [41, 42]). The next evolutionary step would have been cyanidation by an exogenous cyanide source or by an unknown endogenous ligand conversion, with the effect of stabilizing the CO/SR ligand sphere [43, 44].

1.4
Retrodicting Metabolic Reproduction and Evolution

Extant organisms reproduce by sequential phases of metabolic reproduction, genetic reproduction, and cellular reproduction. We retrodict an early stage of life when all reproduction was metabolic reproduction. For conceptualizing the evolution of metabolic reproduction, we subdivide the metabolism into pathways and catalysts (Figure 1.4). Pathway evolution is broken down into elementary transformations: terminal extension, lateral branching, recruitment of nutrients or intermediates, pathway cyclization, pathway reversal, and eliminations [6]. Further, we project all enzymes into metalloenzymes, these into metallocatalysts with inorganic and simple organic ligands, and these finally into an inorganic starter catalyst system, which converts volcanic gases into organic compounds.

Let us assume one of the produced organic compounds turns into a ligand of a transition metal center of the starter catalyst system to augment its catalytic activity, cf. [45]. Consider terminal pathway extension and a proximal ligand effect of an organic product B (Figure 1.5). There are two possibilities. (i) The organic product B feeds back as ligand to enhance catalyst K_m for producing the same organic product B (metabolic reproduction). (ii) The organic product B feeds forward as ligand to enhance a catalyst K'_m for converting the organic product B into another product C. It has the effect of boosting the concentration of C to a level that is sufficient for eliciting a new ligand function (metabolic evolution) [46, 47]. Similar feed-forward effects concern lateral branching, recruitment, or cyclization. In addition, ligand effects may recruit new catalytic transition metals or colonize new chemical habitats. Feed-forward effects may also work over a distance.

Pathway evolution

Terminal extension (+elimination?)

Lateral branching

Recruitment (+elimination?)

Cyclization

Pathway reversal
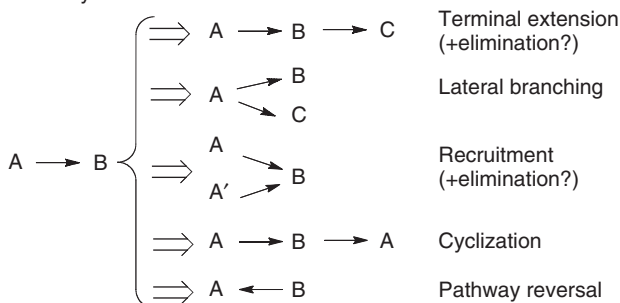
Metallocatalyst evolution

$K_m \implies K_m\text{-}L_1 \implies K_m\text{-}L_2$ Ligand acceleration

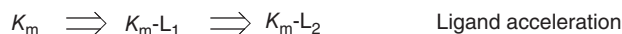**Figure 1.4** Metabolic evolution (A, A′: nutrients; B, C, organic products; ⇒: evolutionary transformation; K_m: catalytic transition metal center; and L_1, L_2: organic ligands), with permission from Springer.

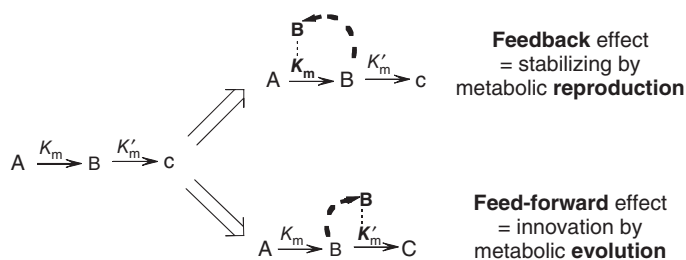**Figure 1.5** Product ligand effects (dashed arrow: ligand effect and ⇒: evolutionary change).

Metabolic expansions by feed-forward effects weaken the preexistent reaction network and require adequate stabilization by feedback effects. Metabolic stability requires multiple product effects, one product enhancing at least two transition metals, or one transition metal that catalyzes at least two reaction steps. With this realization we arrive at the notion of a metabolic avalanche breakthrough. Moreover, by bonding to transition metal centers, organic products may be stabilized against decomposition or hydrolysis (self-selection).

With every feedback effect the metabolism deepens its autonomy from the environment. Eventually, the sum total of all feedback effects will cause the metabolism to run even under conditions that would no longer permit its *de novo* initiation. This marks the beginning of the precariousness of life and the potential for death. It may be said that the emerging organisms are "alive" to the extent that they are "mortal." Chemistry, by acquiring historicity, turns into biology. Each catalytic product effect, as ligand or otherwise, constitutes an instance of memory, or dynamic inheritance. All subsequent evolution is a concatenation of memory effects, progressing from dynamic inheritance of "analog" feedback loop information to genetic inheritance of "digital" sequence information.

1.5
Retrodicting Pioneer-Metabolic Reactions

We now turn to the viability of retrodicted starter catalysts and reducing agents for reactions at or upstream from the location of the pioneer organism. With regard to reaction *temperature*, experiments have proved that the catalytic burden of biochemical reactions suffers a steep increase with decreasing temperature [48]. Therefore, primitive primordial catalysts could generate required reaction rates only at sufficiently high reaction temperatures (e.g., $130-160\,^{\circ}$C), determined by the slowest step in a multistep pathway. Subsequent evolution proceeded irreversibly down the temperature scale to the extent that catalytic activities could increase sufficiently by ligand modification to keep up with the increasing catalytic burden (Wolfenden theorem). The *pH* of the pioneer

organism is determined by the minimum value (lower at higher temperatures) for aqueous H_2S, CH_3SH, or NH_3 to exist as reactive free bases. Primordial rocks were ultramafic (<45% SiO_2), that is, orthosilicates, generating alkaline conditions (pH 9–12). Acidic volcanic gases caused a neutralization front to move in flow direction. Thus, alkalinity decreased from pH 12 to 9 and beyond by leaching first $Ca(OH)_2$ and then $Mg(OH)_2$. Later approach of neutrality permitted condensation reactions and RNA. Thus, the homestead of life was characterized as a "place with liquid water having a nearly neutral pH … where hot volcanic exhalations clash with a circulating hydrothermal water flow" [7, p. 480].

Reductants are needed for carbon/nitrogen fixation. Volcanic-hydrothermal H_2 is available, but quite inert. It requires a transition metal ur-hydrogenase. As to the reductant CO, water gas equilibration between CO/H_2O and CO_2/H_2 replaces CO by H_2, increasingly so with decreasing temperature. In hot liquid water, this occurs without a catalyst, but not below about 200 °C [49]. Minerals also serve as reductants, notably $Fe(OH)_2$ (from hydrolysis of orthosilicates) [50] and FeS/H_2S (Figure 1.3) for driving metabolic reactions.

Inorganic conversions are foundational for the pioneer organism. The system FeS/H_2S has been revealed as ur-nitrogenase ($^{15}N_2 \rightarrow {}^{15}NH_3$) (Weigand reaction) [51], as reductant for forming NH_3 from NO_3^- (due to hydrolysis of NO_x that forms in a $CO_2/N_2/H_2O$ atmosphere by lightning) [52], and as reductant for converting CO_2 to CH_3SH via COS [53]. The aqueous system $H_2S/CO/(NiFe)S$ (~neutral pH) generates COS, CH_3SH, and $CH_3CO(SCH_3)$ [20]. The water-free system $FeS/HCOOH/C_9H_{19}SH$ converts at 250 °C (2000 bar) to system $FeS/H_2O/CO/C_9H_{19}SH$ to generate $C_9H_{19}COOH$, and pyruvate [54]. Iterative carbon fixation (up to C5) was demonstrated for $CO/H_2S/NiS$: $C_n\text{-SH} \rightarrow C_n\text{-COSH} \rightarrow C_n\text{-CHO} \rightarrow C_{n+1}\text{-SH}$ [55]. Aqueous double carbonylation was demonstrated in the presence of $Ca(OH)_2$ (H^+ scavenger) for RSH (R = CH_3, $PhCH_2$) and $CO/Co_2(CO)_8/NiS$ to form RCO-COOH and RCHOH-COOH [27]. Double carbonylation of $PhCH_2$-SH was also shown for $CO/(Co,Ni)(OH,S)$, but dramatically higher productivity was found for $CO/(Co,Ni)(CN)_2$ or $CO/(Co,Ni)(OH)(CN)$ along two pathways: (i) double CO fixation and (ii) CO fixation for CHOH and cyanidation for COOH [27].

Pairs of *α-amino/α-hydroxy acids* and derivatives of formula R-CHA-COY (A = OH, NH_2, $NHCH_3$, $NHCOOH$; Y = OH, NH_2, $NHCH_3$; R = H, CH_3, CH_2OH, C_2H_5) were formed in good yield by direct, one-pot reaction in the aqueous system (CO or H_2)/Ni(OH)(CN) at 100–280 °C, optimally 130–180 °C (70 bar) with (Ca, Mg)(OH)$_2$ as H^+ scavenger (Figure 1.6). Some higher homologs and α-keto acids were also found. CN ligands were the main C source and CO was e^- source and minor C source. The reaction system comprises layered $Ni(OH)_2$ and $[Ni(CN)_4]^{2-}$. When glycine or alanine as products of this reaction were added to the chemical system (CO/Ni(OH)(CN)) for double carbonylation of $PhCH_2$-SH, the yield of $PhCH_2$-CHOH-COOH was significantly increased – a model for the postulated feed-forward effect [27].

α-Amino acids are activated as *N*-carboxy-aminoacyl anhydrides (L) with COS/NiS (derived from $CO/H_2S/NiS$) at pH ~ 9, which are drawn into a *peptide*
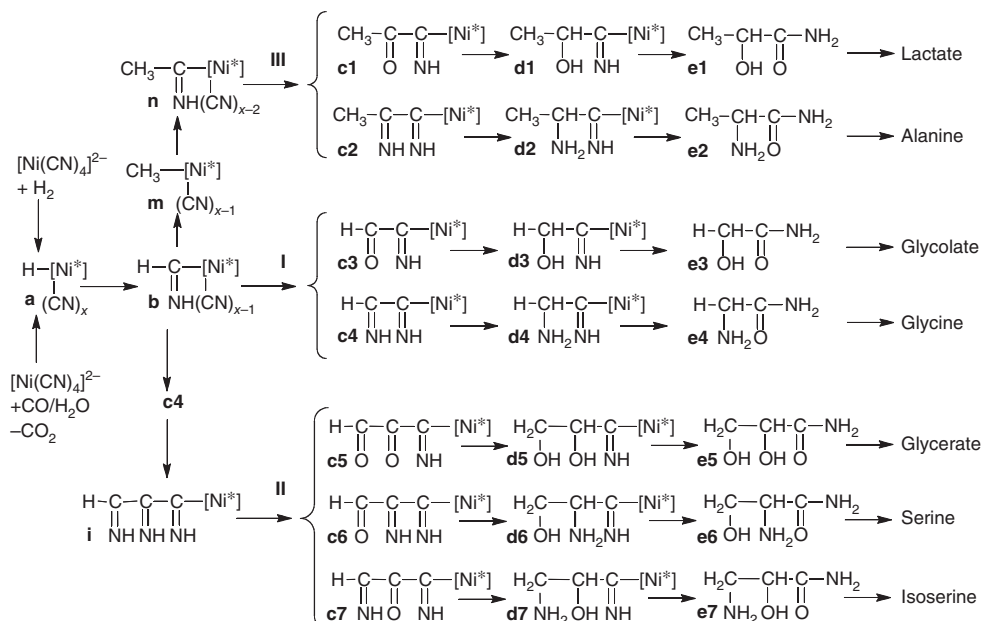
Figure 1.6 Notional mechanism of formation of α-amino/α-hydroxy acids by serial H–N=C insertion [56] ([Ni*], Ni center of unknown nuclearity, oxidation state, and ligand sphere; CO insertion not shown). (From Ref [27] © (2012), Wiley.)

cycle [57, 58] (Figure 1.7). They react with the free amino group of another α-amino acid to form a dipeptide, or with the free amino group of the dipeptide to form a tripeptide, and so forth. The peptides also react with the system COS/NiS to acquire an N-terminal hydantoin ring that hydrolyzed via an N-terminal urea moiety to a free amino acid and a peptide that is shortened N-terminally by one amino acid. By this peptide cycle, the sequence space is scanned for the most stable peptide structures (self-selecting peptide library).

1.6
Early Evolution in a Spatiotemporal Flow Context

An ocean may have existed 4.4 billion years ago, as evidenced by the composition of particles of $ZrSiO_4$ that were formed at that time and preserved in much younger sedimentary rocks [59]. Therefore, the pioneer organism may well date back deep into the Hadean eon, when the Earth had just emerged from its globally molten state. The crust was thin and frequently fragmented by large impactors, generating a great variety of chemical conditions and a bed of debris with a maze of flow ducts for hydrothermal venting and volcanic outgassing [60]. The styles of volcanic-hydrothermal typology differed drastically from those of the present day,
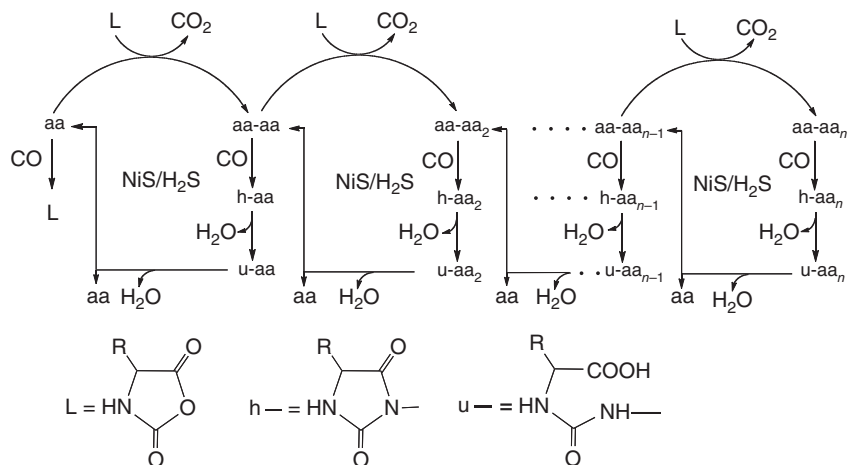
Figure 1.7 Peptide cycle driven by COS derived from CO/H_2S (aa, amino acid), with permission from Wiley.

and the ocean experienced continuous cleansing by intense sedimentation of volcanic ashes and hydrothermal percolation. In short, uniformitarianism does not apply.

The Hadean mantle must have had temperatures well above the melting point of komatiite ($1700\,°C$). Under conditions of graphite saturation, the molar equilibrium ratio $CO:CO_2$ increases with increasing temperature and decreasing pressure. Hence, it must have been higher than the ratio $1:1$, which exists at $1200\,°C$ ($2\,kbar$) [61]. By mixing with cooler hydrothermal water the hot volcanic fluids were quenched, preserving a high activity of CO.

A Hadean volcanic flow path is subdivided into three regions. (i) A deep flow region with a temperature above the critical temperature of water operates as a gas/solid reactor. (ii) A medium flow region with a temperature below the critical temperature of water operates as a gas/liquid/solid trickle-bed reactor. (iii) An uppermost flow region operates as a chromatographic reactor with a stationary packed bed of catalytic particles and a mobile phase of hydrothermal water and dissolved (volcanic) reactants. It exhibits gradients of temperature, pressure, and pH and chemical zoning, as well as chemical reactions in sync with chromatographic separation of organic products (by differential coordination bonding) in flow direction. The strongest surface bonders are the slowest travelers. They concentrate in early "ligand zones" to engage in a surface metabolism. In such a location, the pioneer organism starts by ligand-based feedback and feed-forward effects. It renders the pioneer organism a spatially and temporally coherent entity.

While reactive interactions occur on a molecular scale, larger dimensions are also of interest. Kuhn [62] considered rocks with an open pore structure for size exclusion separation of RNA molecules in an (RNA world) origin of life. Generally speaking, porosity and surface irregularities increase catalytic activity and cause molecular sieve effects. An ingenious experiment-based "chemical garden"

proposal sees early evolution inside a mound of open FeS-bounded "cells." A FeS membrane precipitated at the top of the mound, where hot (e.g., 150 °C), alkaline, sulfidic (e.g., 0.5 M HS$^-$) vent fluid contacts warm (e.g., 90 °C), acidic, carbonated ($\sim$10 bar CO_2) ocean water with dissolved Fe^{2+} (e.g., 0.5 M). The mound grew by breaking and remaking of the FeS membrane. A pH gradient across the top FeS membrane fueled the synthetic reactions inside the mound by chemiosmotically driven ATP synthesis [63]. As the mound grew in height (up to 10 m), the metabolism inside the open-cell structure evolved from FeS-catalyzed carbon fixation at the bottom via the RNA world in the center to LUCA higher up – with full-fledged cells of bacteria and archaea breaking loose at the top [64]. Unfortunately, however, the experimentally detected cell structure is a freeze-drying artifact [46]; RNA cannot exist under hot alkaline conditions; and an experimental FeS membrane "behaved more like a permeable reactive barrier than a membrane" [40].

A volcanic-hydrothermal flow duct is a local affair, which means that the origin of life must have occurred in many locations. Owing to the invariant laws of chemistry, the essential origin chemistry must have been the same everywhere, slight parameter-dependent differences of composition notwithstanding. With the beginning of evolution as a historic process, locally different varieties with different catalytic endowments would have arisen. There was, however, a grand uniformization process at work: large impactors ejected chunks of crustal material high into the atmosphere or even into orbit. Inside these ejected (orbiting) chunks, cellular organisms are believed to have survived until later reentry [60]. By the same argument (sets of) catalysts composed of transition metal centers and evolved organic ligands would have survived *a fortiori* – a truly global lateral transfer of catalytic endowment. Global lateral transfer by this and other mechanisms continued unabated until the advent of LUCA. Lateral transfer of catalytic endowment by physical scattering of ejected crustal material has another important consequence. Fluid flow through volcanic-hydrothermal flow ducts from deep hot locations into the ocean is of too short a duration to accommodate a long period of evolution. Intense scattering of crustal materials, however, and their settlement in renewed flow beds would restart the process time and again with evermore advanced catalytic material. With the formation of lipids by carbon fixation and subsequent cellularization, the metabolism became an internal affair, segregated from the chemical environment – the beginning of an increasing autonomy from the mineral base of life.

The start of the genetic machinery is marked by the advent of organophosphate groups for hard ligation. The components of the genetic machinery (phospho)-sugars, bases, (oligo)nucleos(t)ides, and an ever-expanding set of oligopeptides are all ligand formers that earned their keep with ligand effects. Up to this point, evolution is a *direct* process. New metallocatalysts lead to new organic products that turn *directly* into ligands for more advanced metallocatalysts. With the beginning of ribosomal translation, evolution turns into an *indirect* affair. Mutated polynucleotide sequences are translated into new peptides and these then become ligands for more advanced metallocatalysts. The ribosome is to this day a metallocatalyst with a seemingly inorganic magnesium-phosphate

core, controlled by invariant laws of coordination and surrounded by a gigantic ligand sphere that expanded layer by variable layer over time in radial direction [65]. With an increasingly dominant role of the genetic machinery, the inorganic origins of life in coordination/organometal chemistry became more and more overshadowed. But to this day, the genetic machinery remains what it was at the beginning: the great randomizer of life – invented by life for adaptation for a continued existence in regions evermore remote from the original mineral homestead in a volcanic-hydrothermal flow setting. The genetic machinery works by stochastic variations, which give us the impression of evolution as a directionless, random walk. But in its core and from the beginning the evolution of life is chemically unique, predetermined by universal laws and progressing along the arrow of time.

Acknowledgments

References

1. Kant, I. (1790) Kritik der Urteilskraft. §81, Translation by Meredith, J.C. (1952) *The Critique of Judgment*, Clarendon Press, Oxford.
2. Darwin, C. (1863) Letter to J.D. Hooker, 29 March, 1863.
3. Lipmann, F.A. (1965) in *The Origin of Prebiological Systems and of Their Molecular Matrices* (ed, S.W. Fox), Academic Press, New York, pp. 259–280.
4. Wächtershäuser, G. (1997) The origin of life and its methodological challenge. *J. Theor. Biol.*, **187**, 483–494.
5. Florkin, M. (1949) *Biochemical Evolution*, Academic Press, New York.
6. Granick, S. (1957) Speculations on the origin and evolution of photosynthesis. *Ann. N. Y. Acad. Sci.*, **69**, 292–308.
7. Wächtershäuser, G. (1992) Groundworks for an evolutionary biochemistry: the iron-sulphur world. *Prog. Biophys. Mol. Biol.*, **58**, 85–201.
8. Marx, D. and Hutter, J. (2009) *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*, Cambridge University Press, Cambridge.
9. Popper, K.R. (1963) *Conjectures and Refutations: The Growth of Scientific Knowledge*, Routledge & Kegan Paul, London.
10. Fraústo da Silva, J.J.R. and Williams, R.J.P. (1991) *The Biological Chemistry of the Elements*, Clartendon Press, Oxford.
11. Wächtershäuser, G. (1988) Before enzymes and templates: theory of surface metabolism. *Microbiol. Rev.*, **52**, 452–484.
12. Corliss, J.B., Baross, J.A., and Hoffman, S.E. (1981) An hypothesis concerning the relationship between submarine hot springs and the origin of life on Earth. *Oceanol. Acta* (Proceedings of the 26th International Geological Conference), **4**, 59–69.
13. Johnson, D.C., Dean, D.R., Smith, A.D., and Johnson, M.K. (2005) Structure, function, and formation of biological iron-sulfur clusters. *Annu. Rev. Biochem.*, **74**, 247–281.
14. Kletzin, A. and Adams, M.W.W. (1996) Tungsten in biological systems. *FEMS Microbiol. Rev.*, **18**, 5–63.
15. Crans, D.C., Smee, J.J., Gaidamauskas, E., and Yang, L. (2004) The chemistry and biochemistry of vanadium and the

biological activities exerted by vanadium compounds. *Chem. Rev.*, **104**, 849–902.

16. Malkin, R. and Rabinowitz, J.C. (1966) The reconstitution of clostridial ferredoxin. *Biochem. Biophys. Res. Commun.*, **23**, 822–827.

17. Müller, A. and Schladerbeck, N.H. (1986) Einfache aerobe Bildung eines [Fe$_4$S$_4$]$^{2+}$-Clusterzentrums. *Naturwissenschaften*, **73**, 669–670.

18. Wiig, J.A., Hu, Y., Lee, C.C., and Ribbe, M.W. (2012) Radical SAM-dependent carbon insertion into the nitrogenase M-cluster. *Science*, **337**, 1672–1675.

19. Ragsdale, S.W. (2009) Nickel-based enzyme systems. *J. Biol. Chem.*, **284**, 18571–18575.

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

21. White, H., Strobl, G., Feicht, R., and Simon, H. (1989) Carboxylic acid reductase: a new tungsten enzyme which catalyzes the reduction of non-activated carboxylic acids to aldehydes. *Eur. J. Biochem.*, **184**, 89–96.

22. Seiffert, G.B., Ullmann, G.M., Messerschmidt, A., Schink, B., Kroneck, P.M.H., and Einsle, O. (2007) Structure of the non-redox-active tungsten/[4Fe:4S] enzyme acetylene hydratase. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 3073–3077.

23. Groysman, S. and Holm, R.H. (2009) Biomimetic chemistry of iron, nickel, molybdenum and tungsten in sulfur-ligated protein sites. *Biochemistry*, **48**, 2310–2320.

24. Hsieh, C.-H., Erdem, Ö.F., Harman, S.D., Singleton, M.L., Reijerse, E., Lubitz, W., Popescu, C.V., Reibenspies, J.H., Brothers, S.M., Hall, M.B., and Darensbourg, M.Y. (2012) Structural and spectroscopic features of mixed valent Fe^{II}Fe^{I} complexes and factors related to the rotated configuration of diiron hydrogenase. *J. Am. Chem. Soc.*, **134**, 13089–13102.

25. Kornhauser, K.O., Pecoits, E., Lalonde, S.V., Papineau, D., Nisbet, E.G., Barley, M.E., Arndt, N.T., Zahnle, K., and Kamber, B.S. (2009) Oceanic nickel depletion and a methanogen famine before the Great Oxidation Event. *Nature*, **458**, 750–753.

26. Py, B. and Barras, F. (2010) Building Fe–S proteins: bacterial strategies. *Nat. Rev. Microbiol.*, **8**, 436–446.

27. Huber, C., Kraus, F., Hanzlik, M., Eisenreich, W., and Wächtershäuser, G. (2012) Elements of metabolic evolution. *Chem. Eur. J.*, **18**, 2063–2080.

28. Dey, S., Das, P.K., and Dey, A. (2013) Mononuclear iron hydrogenase. *Coord. Chem. Rev.*, **257**, 42–63.

29. Schick, M., Xie, X., Ataka, K., Kahnt, J., Linne, U., and Shima, S. (2012) Biosynthesis of the iron-guanylylpyridinol cofactor of [Fe]-hydrogenase in methanogenic archaea as elucidated by stable-isotope labeling. *J. Am. Chem. Soc.*, **134**, 3271–3280.

30. Shima, S., Pilak, O., Vogt, S., Schick, M., Stagni, M.S., Meyer-Klaucke, W., Warkentin, E., Thauer, R.K., and Ermler, U. (2008) The crystal structure of [Fe]-hydrogenase reveals the geometry of the active site. *Science*, **321**, 572–575.

31. Rauchfuss, T.B. (2010) Unraveling the biosynthesis of nature's fastest hydrogenase. *Angew. Chem. Int. Ed.*, **49**, 4166–4168.

32. Pilet, E., Nicolet, Y., Mathevon, C., Douki, T., Fontecilla-Camps, J.C., and Fontecave, M. (2009) The role of the maturase HydG in [FeFe]-hydrogenase active site synthesis and assembly. *FEBS Lett.*, **583**, 506–511.

33. Kuchenreuther, J.M., George, S.J., Grady-Smith, C.S., Cramer, S.P., and Swartz, J.R. (2011) Cell-free H-cluster synthesis and [FeFe] hydrogenase activation: all five CO and CN$^-$ ligands derive from tyrosine. *PLoS One*, **6**, e20346.

34. Li, H. and Rauchfuss, T.B. (2002) Iron carbonyl sulfides, formaldehyde, and amines condense to give the proposed azadithiolate cofactor of the Fe-only hydrogenases. *J. Am. Chem. Soc.*, **124**, 726–727.

35. Bürstel, I., Siebert, E., Winter, G., Hummel, P., Zebger, I., Friedrich, B., and Lenz, O. (2012) A universal scaffold for synthesis of the Fe(CN)$_2$(CO) moiety of [NiFe] hydrogenase. *J. Biol. Chem.*, **287**, 38845–38853.

36. Böck, A., King, P.W., Blokesch, M., and Posewitz, M.C. (2006) Maturation of hydrogenases. *Adv. Microb. Physiol.*, **51**, 1–71.

37. Bürstel, I., Hummel, P., Siebert, E., Wisitruangsakul, N., Zebger, I., Friedrich, B., and Lenz, O. (2011) Probing the origin of the metabolic precursor of the CO ligand in the catalytic center of [NiFe] hydrogenase. *J. Biol. Chem.*, **286**, 44937–44944.

38. Taylor, P., Rummery, T.E., and Owen, D.G. (1979) Reactions of iron monosulfide solids with aqueous hydrogen sulfide up to 160 °C. *J. Inorg. Nucl. Chem.*, **41**, 1683–1687.

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

40. Rickard, D. and Luther, G.W. III, (2007) Chemistry of iron sulfides. *Chem. Rev.*, **107**, 514–562.

41. Hieber, W., Beck, W., and Braun, G. (1960) Anionische Kohlenoxidkomplexe. *Angew. Chem.*, **72**, 795–874.

42. Nametkin, N.S., Kolobkov, B.I., Tyurin, V.D., Muratov, A.N., Nekhaev, A.I., Mavlonov, M., Sideridu, A.Y., Aleksandrov, G.G., Lebedev, A.V., Tashev, M.T., and Dustov, H.B. (1984) Synthesis and structure of dithiodiiron heptacarbonyl. *J. Organomet. Chem.*, **276**, 393–397.

43. Rauchfuss, T.B., Contakes, S.M., Hsu, S.C.N., Reynolds, M.A., and Wilson, S.R. (2001) The influence of cyanide on the carbonylation of iron(II): synthesis of Fe—SR—CN—CO centers related to the hydrogenase active sites. *J. Am. Chem. Soc.*, **123**, 6933–6934.

44. Lyon, E.J., Georgakaki, I.P., Reibenspies, J.H., and Darensbourg, M.Y. (2001) Coordination sphere flexibility of active-site models for Fe-only hydrogenase: studies in intra- and intermolecular diatomic ligand exchange. *J. Am. Chem. Soc.*, **123**, 3268–3278.

45. Berrisford, D.J., Bolm, C., and Sharpless, K.B. (1995) Ligand accelerated catalysis. *Angew. Chem. Int. Ed. Engl.*, **34**, 1059–1070.

46. Wächtershäuser, G. (2006) From volcanic origins of chemoautotrophic origin of life to bacteria, archaea and eukarya. *Philos. Trans. R. Soc. London, Ser. B*, **361**, 1787–1808.

47. Wächtershäuser G. 2010. In: (Eds.) L. L. Barton, M. Mandl, A. Loy *Geomicrobiology: Molecular and Environmental Perspective*. Springer, Dordrecht, pp. 1-35.

48. Stockbridge, R.B., Lewis, C.A. Jr., and Wolfenden, R. (2010) Impact of temperature on the time required for the establishment of primordial biochemistry, and for the evolution of enzymes. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 22102–22105.

49. Seewald, J.S., Zolotov, M.Y., and McCollom, T. (2006) Investigation of single carbon compounds under hydrothermal conditions. *Geochim. Cosmochim. Acta*, **70**, 446–460.

50. Huber, C. and Wächtershäuser, G. (2003) Primordial reductive amination revisited. *Tetrahedron Lett.*, **44**, 1695–1697.

51. Dörr, M., Käßbohrer, J., Grunert, R., Kreisel, G., Brand, W.A., Werner, R.A., Geilmann, H., Apfel, C., Robl, C., and Weigand, W. (2003) A possible prebiotic formation of ammonia from dinitrogen on iron-sulfide surfaces. *Angew. Chem. Int. Ed.*, **42**, 1540–1543.

52. Blöchl, E., Keller, M., Wächtershäuser, G., and Stetter, K.O. (1992) Reactions depending on iron sulfide and linking geochemistry with biochemistry. *Proc. Natl. Acad. Sci. U.S.A.*, **89**, 8117–8120.

53. Heinen, W. and Lauwers, A.M. (1996) Organic sulfur compounds resulting from the interaction of iron sulfide, hydrogen sulfide and carbon dioxide in an anaerobic aqueous environment. *Orig. Life Evol. Biosph.*, **26**, 131–150.

54. Cody, G.D., Boctor, N.Z., Filley, T.R., Hazen, R.M., Scott, J.H., Sharma, A., and Yoder, H.S. Jr., (2000) Primordial carbonylated iron-sulfur compounds and the synthesis of pyruvate. *Science*, **289**, 1337–1340.

55. Loison, A., Dubant, S., Adam, P., and Albrecht, P. (2010) Elucidation of an iterative process of carbon-carbon bond

formation of prebiotic significance. *Astrobiology*, **10**, 973–988.

56. Fehlhammer, W.P. and Fritz, M. (1993) Emergence of CNH and cyano complex based organometallic chemistry. *Chem. Rev.*, **93**, 1243–1280.

57. Huber, C. and Wächtershäuser, G. (1998) Peptides by activation of amino acids on (Fe,Ni)S surfaces: implications for the origin of life. *Science*, **281**, 670–672.

58. Huber, C., Eisenreich, W., Hecht, S., and Wächtershäuser, G. (2003) A possible primordial peptide cycle. *Science*, **301**, 938–940.

59. Harrison, T.M. (2009) The Hadean crust: evidence from >4 Ga zircons. *Annu. Rev. Earth Planet. Sci.*, **37**, 479–505.

60. Cockell, C.S. (2006) The origin and emergence of life under impact bombardment. *Philos. Trans. R. Soc. London, Ser. B*, **361**, 1845–1875.

61. Holloway, J.R. and Blank, J.G. (1994) Application of experimental results to C-O-H species in natural melts. *Rev. Mineral.*, **30**, 187–230.

62. Kuhn, H. and Waser, J. (1981) Molecular self-organization and the origin of life. *Angew. Chem., Int. Ed. Engl.*, **20**, 500–520.

63. Russell, M.J. and Hall, A.J. (1997) The emergence of life from iron monosulphide bubbles at a submarine hydrothermal redox and pH front. *J. Geol. Soc.*, **154**, 377–402.

64. Martin, W. and Russell, M.J. (2003) On the origin of cells: an hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. *Philos. Trans. R. Soc. London, Ser. B*, **358**, 27–85.

65. Hsiao, C., Mohan, S., Kalahar, B.K., and Williams, L.D. (2009) Peeling the onion: ribosomes are ancient molecular fossils. *Mol. Biol. Evol.*, **26**, 2415–2425.