



Turnstiles and bifurcators: The disequilibrium converting engines that put metabolism on the road[☆]

Elbert Branscomb^{a,*}, Michael J. Russell^b

^a Institute for Genomic Biology, UIUC, Champaign-Urbana, IL, USA

^b Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

ARTICLE INFO

Article history:

Received 4 July 2012

Received in revised form 2 October 2012

Accepted 3 October 2012

Available online 10 October 2012

Keywords:

Alkaline hydrothermal vent

Origin of life

Chemiosmosis

Serpentinization

Free energy conversion

Non-equilibrium thermodynamics

ABSTRACT

The Submarine Hydrothermal Alkaline Spring Theory for the emergence of life holds that it is the ordered delivery of hydrogen and methane in alkaline hydrothermal solutions at a spontaneously precipitated inorganic osmotic and catalytic membrane to the carbon dioxide and other electron acceptors in the earliest acidulous cool ocean that, through these gradients, drove life into being. That such interactions between hydrothermal fuels and potential oxidants have so far not been accomplished in the lab is because some steps along the necessary metabolic pathways are endergonic and must therefore be driven by being coupled to thermodynamically larger exergonic processes. But coupling of this kind is far from automatic and it is not enough to merely sum the ΔG s of two supposedly coupled reactions and show their combined thermodynamic viability. An exergonic reaction will not drive an endergonic one unless 'forced' to do so by being tied to it mechanistically via an organized "engine" of "Free Energy Conversion" (FEC). Here we discuss the thermodynamics of FEC and advance proposals regarding the nature and roles of the FEC devices that could, in principle, have arisen spontaneously in the alkaline hydrothermal context and have forced the onset of a protometabolism. The key challenge is to divine what these initial engines of life were in physicochemical terms and as part of that, what structures provided the first "turnstile-like" mechanisms needed to couple the partner processes in free energy conversion; in particular to couple the dissipation of geochemically given gradients to, say, the reduction of CO_2 to formate and the generation of a pyrophosphate disequilibrium. This article is part of a Special Issue entitled: The evolutionary aspects of bioenergetic systems.

© 2012 Elsevier B.V. All rights reserved.

"Synthetic biology is ... the study of the physical forces and conditions which can produce cavities surrounded by osmotic membranes, ... associate and group such cavities, and differentiate and specialize their functions" [115].

1. Introduction

1.1. Life as thermodynamics: it's disequilibria all the way down

In their thermodynamic dimension living systems are staggering and wondrous Taj Mahals of disequilibria—writ in both structure and process; of intense and dynamic disequilibria in fact, precariously lofted high above the quiet repose of the equilibrium state. This is their essential

and distinguishing property. But, of course, the 2nd law dictates that a particular disequilibrium can only arise by a process in which it is "driven" into place—against probability—by being mechanistically tied to the dissipation of a 'greater' disequilibrium; tied through a mechanism that fuses the two processes, driver and driven, into a single—and spontaneous—thermodynamic process; a thermodynamic "Atwood machine". In consequence, life's multitudes of disequilibria are linked together—and to the external world—in cascading networks of driver-driven pairs; each linkage embodying an instance of disequilibria conversion; or, as it is more conventionally termed, "free energy conversion" ("FEC") or "free energy transduction" [1,2]. And "bioenergetics", we here argue, is entirely the study of FEC.

In this introductory section we review the thermodynamic concepts and approaches that are brought into play in the analysis of such disequilibria-converting, FEC processes. Our motivation for doing so in the present context is based in part on an argument that the enabling initial step on the path from the inanimate to the animate was the creation and maintenance, in appropriately localized contexts, and through spontaneously arising, abiotic FEC mechanisms, of two critical 'founding' chemical disequilibria. It is our view, in other words, that at its emergence, life was as much, and as essentially, an economy of disequilibria conversion as it has clearly been since it became recognizably "life like".

[☆] This article is part of a Special Issue entitled: The evolutionary aspects of bioenergetic systems.

* Corresponding author. Tel.: +1 217 333 0515; fax: +1 217 333 0508.

E-mail address: brnscmb@illinois.edu (E. Branscomb).

But on this view, the challenge is to divine what, in this birthing transition, were the driving and driven disequilibria, and perhaps most problematically, what were the likely abiotic mechanisms that, in the right physical context, pulled off the profoundly non-trivial trick of coupling one to the other. The subsequent sections of this paper will discuss this aspect of the problem.

1.1.1. What precisely is a “disequilibrium”?

When a quantity of matter and/or energy (“M/E”) can exist within a spatially bounded system in two alternative states (necessarily equal in all relevant conserved quantities) those states will form a disequilibrium (a “gradient”) if and only if they differ in their entropy content; entropy in Boltzmann’s atomistic terms is the log measure of the state’s microstate “multiplicity”: $S/k = \ln W$ where W is the number of “accessible” microstates available to the M/E [3]; a disequilibrium is therefore fundamentally defined by the condition that the two alternative states differ in entropy content, i.e. in microstate multiplicity: $W_{\text{state1}} \neq W_{\text{state2}}$.

A flux between these two states, if not physically prevented, will arise spontaneously, driving the matter and/or energy from the state of low entropy to that of higher entropy. That is, if there is a flux, i.e. if something (macroscopic) happens, then the 2nd law demands that $W_{\text{after}} > W_{\text{before}}$. Furthermore, this increase in microstate count, in the entropy of the M/E involved in the process, is arguably not merely a restriction on macroscopic processes, but is instead their driving cause and designing hand.

But note that, as presaged above, central to the discussion of FEC is the possibility that the flow induced by a disequilibrium may be a ‘compound’ process—i.e. composed of (at least) two processes—comprising two gradients and two fluxes—mechanistically linked so they function as a single process thermodynamically as we discuss below.

1.1.2. What happened to energy?

Discussions of bioenergetics almost always have “energy” center stage and running the show—albeit energy suitably discounted (in “its ability to perform work”) by the 2nd law tax. But this, in our view, is misleading. Energy doesn’t drive processes—or become “consumed” or “used” in their occurrence. Disequilibria (via their dissipation) alone are to blame for why—and how—things happen (and not just in the processes of life); although it is true that much of the entropy difference that defines these disequilibria is due to differences in how the energy involved is ‘packaged’ or distributed; that is, while energy is typically a dominant contributor to the “medium”, the “message” is disequilibrium dissipation and creation. To better see the point at issue here consider the case of life’s “main fuel” disequilibrium—that of ATP versus its hydrolysis products ADP + Pi.

In actively metabolizing cells, typical ATP free energy potentials are 50–60 kJ/mol—maintained by a furious industry of ATP reconstitution and use; even in sessile ‘sapient’ the daily turnover of ATP is roughly equal to body weight (100–150 mol; each ATP is regenerated 1 to 1.5 K times per day; that is, approximately once per minute).

Using a commonly adopted approximate expression for the Gibbs free energy of a chemical reaction in a state displaced from equilibrium (applicable when the reaction makes a negligible change in the reagent’s chemical potentials [4]), we can express the above ATP free energy in terms of the concentrations involved as:

$$\Delta G = -kT \ln \left[\frac{[ATP]/[ADP][Pi]}{[ATP]_{\text{eq}}/[ADP]_{\text{eq}}[Pi]_{\text{eq}}} \right]$$

where the superscript “eq” labels the equilibrium values. Therefore, at typical temperatures, an ATP “free energy” of ≈ 50 kJ/mol implies that $\Delta G/kT \approx 20$, or that the ratio of [ATP] to the product of the concentrations of its hydrolysis products [ADP][Pi] is approximately 5×10^8 fold greater than it would be at equilibrium. It is this astronomical disequilibrium which the cell is constantly dissipating (in FEC processes in which it is the “driving” disequilibrium for the creation of other disequilibria) and put to the considerable pains of constantly replenishing—and which,

therefore, ‘carries’ the free energy associated with ATP. In other words, ATP (and its bond energy) is by itself useless to the cell as a source of driving force; ATP at equilibrium with its hydrolysis products would not drive a single endergonic reaction no matter what the concentration of ATP. Note also that ΔG never occurs in the expression for a measurable quantity except in the dimensionless ratio $\Delta G/kT$, showing that the only thing that matters physically in the above expression is the dimensionless numerical expression involving the log of a (dimensionless) ratio of concentrations, and that “free energy” is therefore not physically an “energy” (of any stripe), but is just a dimensionless numerical quantity measuring how far from equilibrium the system currently is. Necessarily, of course, the energy exchanged in this reaction is exactly conserved (in the physicist’s literal-minded sense of meaning “not changed”, not the biologist’s sense of “not all wasted as heat”), and is in no sense “consumed”. In this connection we note that in the terms used above to characterize disequilibria, the Gibbs free energy inherent in a chemical reaction displaced from equilibrium can be expressed in general as $\Delta G = -kT \ln(W_{\text{state2}}/W_{\text{state1}})^{\text{mole-reaction-progress}}$ which is negative if $W_{\text{state2}} > W_{\text{state1}}$ —indicating that a spontaneous reaction would take place from state1 to state2.

However, to give ATP’s bond energies their due, the hydrolysis, for example, of the terminal phosphoanhydridic bond dumps into the thermal ‘bath’ a packet of energy that is (typically) about 12 fold larger than the bath’s own energy packets ($\approx kT$)—into which sea the bond energy packet becomes quickly dispersed. It is for the most part this large energy dispersal factor (“one-to-many repackaging”) that makes ATP hydrolysis (from a far-from-equilibrium state) able to drive so many other reactions through relatively simple FEC mechanisms.

1.1.3. The need for non-equilibrium thermodynamics (NET)

For several reasons the thermodynamic analysis of FEC mechanisms is inherently outside the embrace of classical equilibrium thermodynamics (ET) and requires non-equilibrium thermodynamic (NET) concepts and methods [1,2]. For starters, classical ET is limited to the consideration of equilibrium states or to states only incrementally or “quasi-statically” displaced from them. It is inherently non-dynamic (thermo-dynamics it emphatically isn’t). It can characterize a bar fight only in terms of the differences one can measure in the state of the bar between the completely quiet “end” states before and after the fight took place.

NET’s focus is not “states” at or (very) near equilibrium, but fully dynamic “processes” in which states are in general far from equilibrium (FFE) and changing at velocities which are under no obligation to be approximately zero. It undertakes, therefore, to characterize the bar fight during its entire time course. And of course FEC processes are just such FFE, dynamic, finite velocity, systems. Notably, an essential feature of FEC systems and FEC-sustained systems, is that the FEC process can maintain a system far from equilibrium in steady state; a circumstance that ET cannot countenance. In particular, accounting for the key feature of FEC, namely the locking of two thermodynamic flows into one process by a FFE macroscopically organized, dynamic mechanism, especially when that mechanism is driven into existence by the driving gradient itself, is in every regard qualitatively incompatible with the assumptions of classical ET.

To help clarify these points and lay a conceptual groundwork for the subsequent section’s discussion of the problem of finding the FEC processes which we hypothesize got life started, we next undertake a quick tour of NET and of how FEC is analyzed within that framework.

1.1.4. Blitz summary of non-equilibrium thermodynamics and free energy conversion

The defining focus of non-equilibrium thermodynamics is entropy—its rate of change in particular (it is a truly dynamical thermodynamics); NET sees thermodynamics as a (dynamic) play about entropy, not a (static) tableau about energy (as does classical, equilibrium thermodynamics). The enabling assumption allowing the theory to embrace finite velocity

processes (and yet permit feasible calculation of thermodynamic quantities) is that “local equilibrium” is an acceptably good approximation for all points within the physical system under consideration [5]; i.e. that the Boltzmann distribution adequately describes the probability distribution of energy states in a suitably small volume defining the neighborhood of any point—so that a temperature “at each point” is well enough defined.

Many of the concepts and relationships of ET extend to the NET regime, if not always with quite the same meaning; the primary bridge being a re-interpretation of the “fundamental energy equation” of ET which asserts merely that a “simple” system’s energy is a function of three independent variables, its entropy S , volume V , and the vector of the counts of each of the kinds of particles present in the system N : $U = U(S, V, N)$. NET reinterprets this relation, by inverting it, to obtain the “fundamental entropy equation” according to which the system’s entropy is a function of the three independent variables U , V , N : $S = S(U, V, N)$ and where now all of these independent variables can be functions of time (see the discussion in Dill and Bromberg, pp107 [3] in which the authors characterize the fact that thermodynamics and its original definitions were founded on the “fundamental energy equation”, not the “fundamental entropy equation”, as an “inconvenient ... quirk of history”; because “...the microscopic driving forces are better understood in terms of the entropy equation $S = S(U, V, N)$ ”; to which we add that this latter fact is because thermodynamic driving forces are just, and only, entropy differences i.e. disequilibria). In any case, such basic quantities in ET as enthalpy: $H = U + pV$ and Gibbs free energy: $G = H - TS$ can be extended directly to NET where they become dynamic functions changing at finite rates during the course of a reaction.

Consider a spatially bounded system, potentially open to the exchange of matter and energy with its exterior, within which finite velocity (i.e. inherently ‘irreversible’) processes are taking place. A process here can be a chemical reaction, a diffusion, a conduction of heat or electricity, a convective flow, etc. Each such process is the expression of (i.e. is ‘caused by’) a state of disequilibrium (equivalently the existence of a “gradient”) with respect to the energy/material mediating that process. Thus to each such process is associated a “flux” (e.g. heat flow, the progress of a chemical reaction, or an electrical current), and also a “force” (a measure of the strength of the gradient driving the flux). All such processes are “irreversible” because they change (at a finite rate) the entropy of the material or energy involved in the process.

The change in entropy as a function of time of the bounded system is thus due to two contributions: entropy carried by the flow of material and/or energy across the system’s boundary (an incremental amount of which is conventionally labeled $d_e S$), and the changes in entropy of the material/energy within the bounded system due to the irreversible processes taking place within it (labeled $d_i S$). That is, as is drawn in Fig. 1, the total incremental change in the entropy of the system is:

$$dS = d_e S + d_i S$$

where $d_e S$ and dS can be of either sign but by the 2nd law we must have $d_i S \geq 0$ (explanation: $d_e S$ has no effect on the entropy of the universe

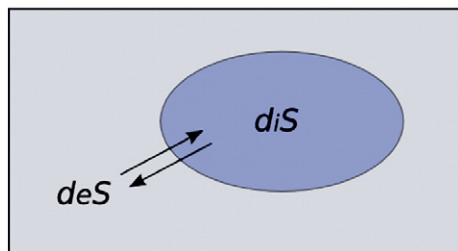


Fig. 1. The dynamics of entropy change in a bounded region open to the exchange of both matter and energy.

since it is just due to moving energy/material from one place to another; therefore it is only the irreversible processes taking place within the system that effect the entropy of the universe; that is: $dS^{\text{universe}} = d_e S$. But by the 2nd law, dS^{universe} must be non-negative) [1,6].

A fundamental relationship of NET expresses the rate of entropy production due to the system’s irreversible processes in terms of a sum of the rates at which each of the processes going on in the system contributes to that entropy production—and further expresses each of these individual contributions as the product of the “flux” (the flow of E/M defining that process) and the “force” for that individual process:

$$\frac{d_i S}{dt} = \sum_j \frac{d_j S}{dt} = \sum_j J_j X_j$$

where J_j and X_j are, respectively, the flux and the force associated with the j th process. Note that the force terms X_j are simply the amount of entropy change that each unit of the j th flux experiences (and that this is, therefore, ‘all there is’ to the ‘forces’ that make macroscopic things happen; they are just—and only—entropy changes). In the case of chemical reactions, and following common units conventions, $J_j = d\xi_j/dt$ and $X_j = A_j/T$ where ξ_j is the “extent” of the j th reaction and $A_j = A_j(\xi_j)$ is the so-called “affinity” of the j th reaction at its current extent ξ_j [8]. Note that $A_j > 0$ (resp. $A_j < 0$) implies that the j th process enjoys an increase—is exergonic (resp. decrease—is endergonic) in entropy, and that $A_j = 0$ when the reaction is at equilibrium (no entropy change to be gained).

Then the ‘big’ points:

- first of all, even though we must have $d_e S \geq 0$ and therefore $\sum_j J_j X_j \geq 0$ this does not imply that all of the terms in this sum must be positive; some fluxes can be reducing the system’s entropy—because their forces are negative ($X_j < 0$ —an endergonic process).
- second, the fluxes can depend on all of the forces acting within the bounded region, not merely on their ‘cognate’ force. In this way, two (or conceivable more) processes act thermodynamically as one; and the dissipation of one gradient, say $X_1 > 0$, can be obligatorily coupled to the creation of another, say $X_2 < 0$. In fact, only in this way can a process proceed in which $X_2 < 0$ —that is, in which a disequilibrium is created.

In a system with just two processes, for example, we could have $J_1 = J_1(X_1, X_2)$ and $J_2 = J_2(X_1, X_2)$ and also $X_1 > 0$ (a ‘driving’ process) and $X_2 < 0$ (a ‘driven’ process), so that we can have $dS/dt = J_1 X_1 + J_2 X_2 > 0$ even though $J_2 X_2 < 0$; this can only arise, of course, if each of the fluxes does in fact depend in an appropriate fashion on its non-cognate force. In this circumstance we have “free energy conversion”; the first disequilibrium is dissipated, the second created - and potentially maintained. All situations in which a non-equilibrium state of matter or energy is created are embraced within this conceptual scheme. No other options exist consistent with the 2nd law; a disequilibrium must be dissipated to create another and the two processes must be reciprocally conditional on each other (through a mechanism that forces inter-process locking) [1,2]. The term “engine” has been apply applied by Alan Cottrell [7] to all systems that carry out free energy conversion since they are all “engines” as physics classically construed the term, and all engines are FECs [2].

To emphasize the critical mechanistic point here, the functional essence of the coupling that achieves FEC is that the driving flux is made conditional on (is ‘gated’ by) the coincident occurrence of the other (driven) flux—which flow, being inherently improbable (i.e. anti-entropic), would, of course, never proceed (‘upstream’) on its own. However, the coupling of two processes as above envisaged is under no stretch ‘automatic’ or trivial; and is in fact a quite special state of physical affairs. In essentially all situations of interest this linking of the two processes into one, requires, and is mediated by, a macroscopically ordered and dynamic ‘structure’ which acts functionally as a “double turnstile”. The turnstile permits a token of the driving flux J_1 to proceed downhill if and only if there is the coincident occurrence of some fixed ‘token’ of the driven

flux J_2 moving “uphill” by chance (albeit as an inherently improbable event) in the same movement of the turnstile. Embodying such conditional, turnstile-like gating mechanisms is what is universally being managed by such evolutionary marvels as the redox-driven proton pumps we will consider in detail later and indeed all other biological devices that carry out what is conventionally termed “energy conservation” (which name, we however argue, misleads in both of its terms). This situation is drawn in Fig. 2.

In many naturally arising FEC engines, i.e. the so-called “emergent” or “self organizing dissipative structures” (“SODS”) [6,8], the driving gradient creates its own coupling “gear” (in Fig. 2) from nothing: the macroscopically organized state of matter that mediates the turnstile coupling between the driving and driven process is produced—forced into existence—by the driving gradient itself. In these cases the FEC-mediating organized structure is itself part of the “disequilibrium” that the FEC is creating through its driven flux. It is due to this positive, ‘autophagic’, feedback behavior that the engine is “autocatalytic” and thereby self-generating or emergent. The engine quite literally consumes (or ‘feeds on’) its own work output to achieve stability and growth (see the remarkable discussion by Cottrell [7]); hydrothermal convection cells are beautifully illustrative examples [9]. Importantly, such entirely self-generating FEC systems come into existence only in response to very far-from-equilibrium disequilibria (as Prigogine and his collaborators were the first to study in detail [8]). Fig. 3 presents a diagram meant to illustrate the ‘autophagic’, self-amplifying behavior of the SODS class of FEC systems.

But as Cottrell [7] has further pointed out, in other cases, due to the prior action of other, unrelated FEC processes, the ordered structures needed to mediate a FEC process can already be present ‘for the taking’—either entirely or in part. In these circumstances, FEC systems can arise in response to relatively weak driving gradients. Note that even in these systems, autocatalytic mechanisms, some quite indirectly mediated (as in the case of living systems) can also be critically in play.

We note that NET had its essential beginning with the work of Onsager [10] who considered the so called “Linear Non-equilibrium Thermodynamics” (“LNET”) approximation in which the fluxes are assumed

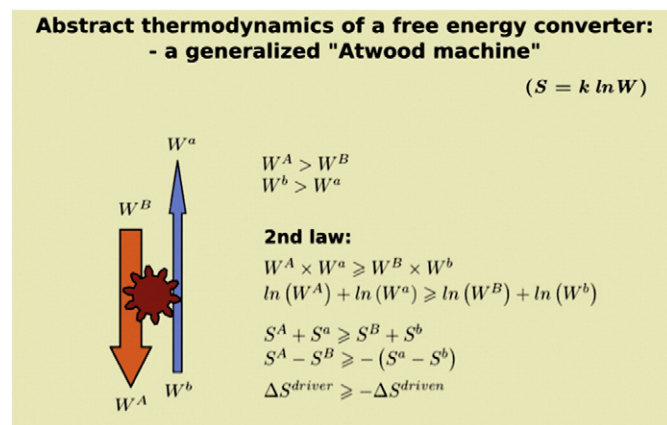


Fig. 2. The driving flux is indicated in this figure by a downward-directed, orange arrow, the driven flux by an upward-directed (and thinner) blue arrow. In a fixed increment of time the driving flux undergoes an increase in microstate count of $W^B \rightarrow W^A$ where $W^A > W^B$ while in the same increment that of the driven flux undergoes a decrease $W^b \rightarrow W^a$ where $W^a < W^b$. The 2nd law simply requires that the total microstate count before: $W^B \times W^b$ be not less than the total after $W^A \times W^a$ (so that the total process is exergonic—and spontaneous). The red gear symbol in the center represents the turnstile coupling mechanism that must exist to tie the two flows together into a single, unitary, thermodynamic process. Note that for engines (e.g. classical “heat engines”) producing mechanical work the output, being motion in one dimension, has one degree of freedom; therefore $W^a = 1$ and $S^a/k = \ln(1) = 0$.

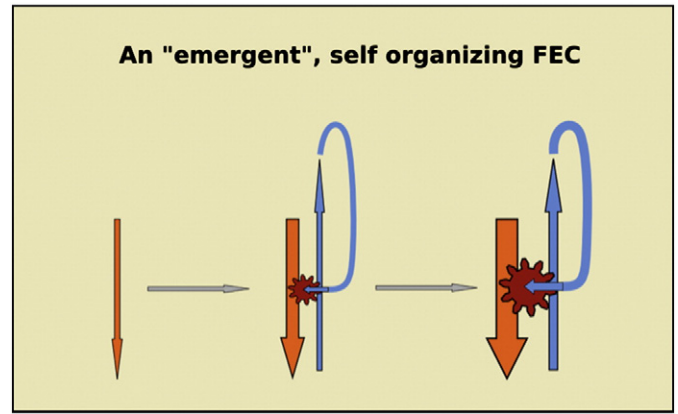


Fig. 3. The left hand downward-directed orange arrow represents a passive “thermodynamic branch” (in Prigogine’s terminology) dissipative flow (such as heat conduction or geochemical reaction in aqueous solution). As the strength of the imposed gradient (the thermodynamic “force”) is increased (left hand horizontal blue arrow) a point is reached at which the passive dissipative modality becomes unstable. The system can then transition to the “organized branch” by creating a macroscopically organized, and dynamic dissipative “structure” (middle figure with the red gear) that: (a) is then stable (to perturbation), and (b) is an FEC driving an anti-entropic process (upward blue arrow)—with the additional magical property of being autocatalytic (such as a convective cell; plastic deformation via linear-discontinuity motion and multiplication is another ‘direct’ example). This property arises because the anti-entropic flow (creating free energy) acts on (provides a free energy input that is consumed by) the engine itself (indicated by the looping blue arrow). The effect of this autophagic consumption of the engine’s free energy output is to stabilize and grow the engine (the right hand horizontal arrow indicating engine growth)—increasing its processing (i.e. dissipative) capacity. When such SODS emerge they invariably accelerate the rate at which the driving gradient is being dissipated.

to depend on the forces through linear relationships (an approximation that is certainly accurate “near equilibrium”): $J_j = \sum L_{jk} X_k$ from which assumption he derived the extremely important reciprocal relations involving the coefficients that specify the coupling between the two processes ($L_{jk} = L_{kj}$) using an argument based on micro-reversibility (“detailed balance”). Most interestingly, linear flux–force relationships and the reciprocal relations appear to be observed by a host of FEC systems (in biology perhaps especially) in which the linear approximation should by no stretch apply (oxidative phosphorylation is one notable example discussed in this connection) [1,4,11].

1.2. Aside on the meaning of the Gibbs’s free energy relation $\Delta G = \Delta H - T\Delta S$ from the viewpoint of NET

Re-arranging this as $\Delta S = \Delta H/T - \Delta G/T$ and, assuming that the chemical reaction implicitly referred to is the only process taking place in NET’s bounded region, integrate the above NET entropy budget relationship $dS = d_e S + d_i S$ over time for the full extent of the reaction to give: $\Delta S = \Delta_e S + \Delta_i S$. Comparing this with the ‘inverted’ Gibbs equation (recognizing that this assumes a closed system with p and T held constant) suggests what is in fact a correct inference, namely that in the circumstance where the Gibbs’s free energy is defined, $\Delta H = T\Delta_e S$ and $\Delta G = -T\Delta_i S$.

That is:

- the “true” (i.e. physical) meaning of the change in enthalpy is that it is just the temperature times the entropy gained (or lost) by the transfer of energy and/or material across the system’s boundary, and that
- the physical meaning of the change in the “Gibbs free energy” is that it is just the negative of the temperature times the entropy produced by the reaction.

Although multiplying the quantities $\Delta_e S$ and $\Delta_i S$ by temperature—thus clothing them in the units of energy—recovers the classical Gibbs’s

free energy equation and mollifies both history and convention, it arguably obscures the physics. In particular, the above discussion makes it clear that in the Gibbs relationship there are not three different types of physical quantities: free energy, enthalpy, and entropy; but just one: entropy, and the relationship is, in physical content, 'really' just the simple entropy budget equation of NET given above. Note however that the NET entropy budget relationship is more general in two fundamental respects; in applying to open systems (where both energy and matter can flow between the system and its environment), and in applying to ongoing processes taking place at finite velocities, not just to difference between the (equilibrium) end states of those processes (or to processes that are obliged to proceed "quasi statically").

1.3. Bridging the concepts and methods of ET into the world of NET and FEC

As was noted above, much of the machinery of ET has a valid generalization in NET. Perhaps most usefully, the classical thermodynamic potential functions, such as the Gibbs free energy $G = U + pV - TS$, carry over under the same conditions (for G , a system in which p and T are held constant by a bath) except that now $G = G(\xi)$ is defined throughout the course of a chemical reaction, however rapidly (within limits) that reaction takes place (noting that the extent of reaction will in general be time dependent $\xi = \xi(t)$). The formal connection between a chemical reaction's affinity in NET and (where it is defined) the Gibbs free energy (in its NET rendering) is given by:

$$A(\xi) = - \left(\frac{\partial G}{\partial \xi} \right)_{p,T}$$

i.e. the affinity is (minus) the rate of change of the Gibbs free energy with respect to the reaction's "extent of progress" [1,6]. Furthermore, if the bath is large enough so that the reagent concentrations can be taken as constant, the change in the system's Gibbs free energy (per mole of reactants) involved in displacing the reaction from equilibrium is numerically equal to minus the affinity: $\Delta G = -A$ (see the discussion of why this should be regarded as only a "numerical equivalence" in Kondepudi and Prigogine p111 [6]).

Thus, correct discussions of FFE processes can often be cast in what are, at least in appearances, the familiar terms of classical thermodynamics. This is particularly the case, common in bioenergetics, in which the disequilibria can be considered to be approximately constant—"in steady state". And it is also particularly the case if the point of the discussion is not to understand how the dissipation of one disequilibrium state has physically led to the creation of another.

However, it's worth emphasizing that in chemical free energy conversion, it is not a requirement that the driver reaction's Gibbs free energy be greater in magnitude than that of the driven reaction; i.e. that $-\Delta G^{\text{driver}} > \Delta G^{\text{driven}}$ (where $\Delta G^{\text{driver}} < 0$ and $\Delta G^{\text{driven}} > 0$). What is required is that $-\Delta G^{\text{driver}} \times \text{moles}^{\text{driver}} > \Delta G^{\text{driven}} \times \text{moles}^{\text{driven}}$; this follows from the analysis presented in the legend to Fig. 2 and the point noted above that: $\Delta G = -kT \ln(W_{\text{after}}/W_{\text{before}})^{\text{mole-reaction-progress}}$.

We are ready now to turn to the place and conditions we consider likely to have acted as life's hatchery on the early earth.

2. The origins of bioenergetics; searching for the founding mothers of biological FEC in the "lost cities" of the Hadean sea

Our central task, as we noted in the opening discussion, is to discover the founding FECs of proto-life; identifying their driving and driven disequilibria, and the naturally arising abiotic turnstile mechanisms that would have coupled them appropriately into single thermodynamic processes. We turn to these issues in the remaining segments of this paper and base our discussion on the "alkaline hydrothermal model" of

the genesis of life originally advanced by Russell et al. [12] and since developed extensively by those and other authors [13–18].

2.1. Initial conditions; the initial driving disequilibria

To see and understand how bioenergetics first emerged on Earth we must transport ourselves back a third of the age of the Universe, ridding ourselves of some of the preconceptions derived from our experience of the world as it is today. The manner by which the planet began to lose its gravitational, and radiogenic heat after the collision with the Mars-sized planet Theia (which produced our moon) and the formation of the iron core, was through many deeply rooted volcanic plumes that dispensed water vapor, carbon dioxide, nitrogen, sulfur and nitric oxides to the atmosphere [19–21]. By 4.4 billion years ago cooling had progressed so far that torrential rains flooded the world, taking around 10,000 years or so to produce an ocean about 10 km deep [20,22,23]. Convective processes continued to transport heat to the surface through a series of stacked and compartmentalized convection cells—one driving another—from the entirely liquid core, via the mantle, crust, and eventually to the surface where it was lost to the cold sink of expanding space [19]. These convection cells are examples of "directly self-generating free energy converters". Convective transfer could eventually involve the all-enveloping ocean as it leaked down faults in the stressed ocean crust, was heated partly through exothermic hydroxylations and returned buoyantly to the ocean bottom in a process known as serpentinization [18]. A large number of convection cells driven by silicic magmatic intrusions and involving ocean water produced a myriad of ephemeral acidic springs thermostated at around 400 °C. They were responsible for introducing transition metals (Fe >> Zn > Ni > Co ~ W) dissolved from the crust in supersaturated solution to the Hadean Ocean. Others, operating at ≤ 150 °C were alkaline, driven merely by heat flow and exothermic hydration of the crust (serpentinization); these supplied the materials (e.g., Ca > Mg >> Mo + OH + NH₃ ~ HS[−]) and fuels (H₂ > CH₄) to what was to become the hatchery of life. This hatchery was a spontaneously precipitated hydrothermal mound which provided the free energy capturing and converting devices to drive incipient metabolism while exporting "waste" heat and unreactive materials in the exhaling hydrothermal effluent back to the ocean [24–26] (Fig. 4). Thus, convection and metabolism were intimately coupled at life's emergence, a coupling that holds, if more loosely, to this day.

From a proto-biological perspective we can view the entire alkaline convective system and hydrothermal mound initially as a single complex proto-metabolizing entity measuring cubic kilometers; where the conduits guiding the carbonic ocean down and through the crust acted as catalyzing redox imports, where water was heated to around 150 °C and partially reduced to hydrogen, where carbon dioxide was in turn partially reduced to methane and minor formate, where protons were spent releasing calcium from silicates to leave an alkaline solution with pH ~ 11, where bisulfide ions were freed from metal sulfides, where the exhaling solutions reacted with the ambient carbonic ocean to generate a catalytic compartmentalized mound, where protons from this ocean leaked into the mound to drive pyrophosphate production and the carbon dioxide leaked through to be hydrogenated to further formate, where oxidants on the outside could oxidize methane on the inside to various intermediates, and where all the uncooperative products could be exhausted with the effluent [16,17,25–29] (Fig. 5). As the mound evolved some of the import paths could be entrained to its base, beginning the miniaturizing tendency to cubic meters, and eventually involving the millimetric to micrometric compartments comprising the outer margins of the mound, a prelude to the development of the first biofilm as organic molecules took over the main structural roles of the cell while leaving the inorganic molecules to their catalytic roles—phosphate for free energy storage and transfer, transition metals, particularly as sulfides, in redox catalysis, and magnesium as a stabilizing influence over negatively charged biomolecules.

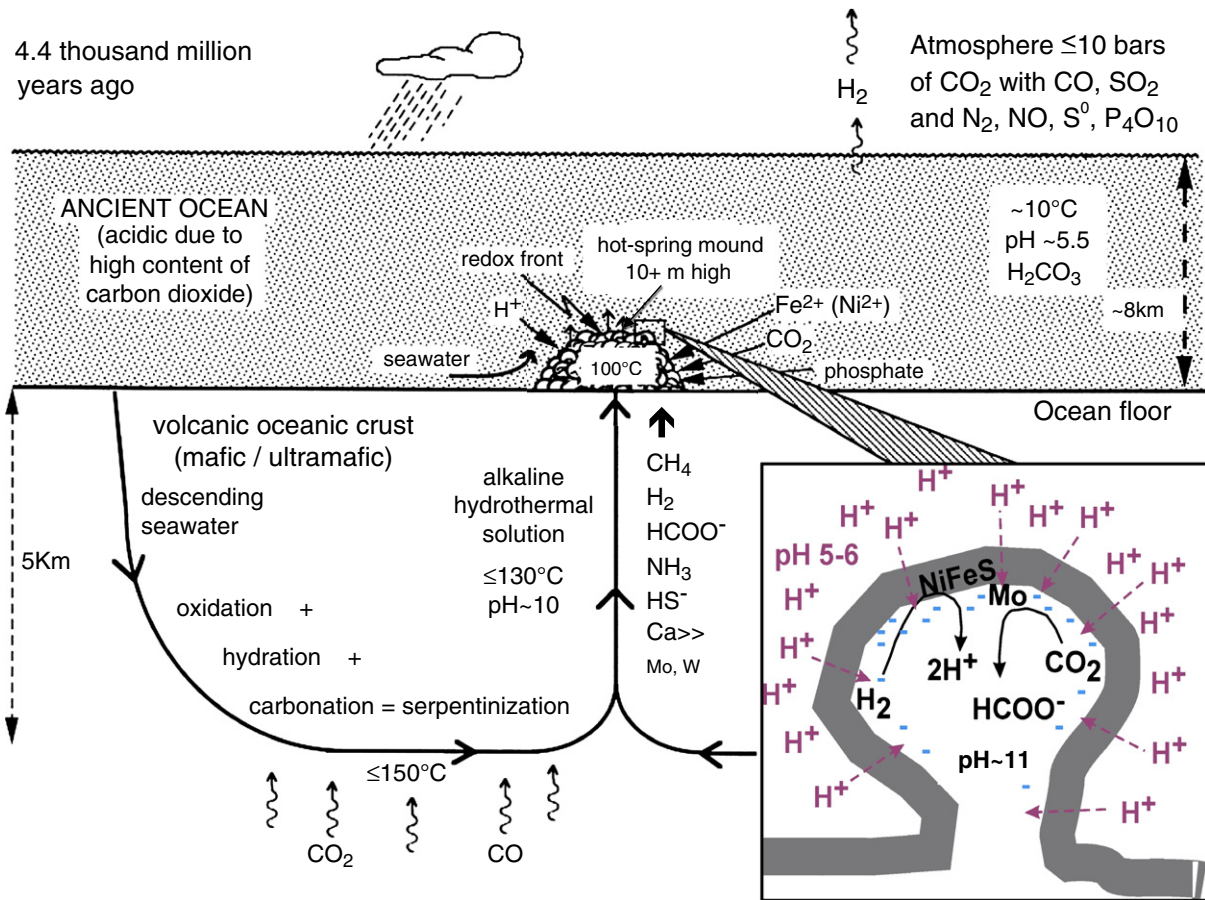


Fig. 4. The type of open system submarine and alkaline (pH ~ 11) hydrothermal convection cell thought to have fed emergent life at a submarine mound 4.4 Ga. Such a system serpentinizes the oceanic crust and lasts for at least 100,000 years ($> 10^{17}$ μm). The all-enveloping ocean is carbonic with a pH of between 5 and 6. Note the provision for the export of waste (entropy). The inset box suggests how the CO₂ may have been reduced to formate [16,44]. Based on [18,32,33,109].

Thus the chemical disequilibria arising within this now miniaturized compartmentalized system that would have been available to power the FEC mechanisms of the first metabolic pathways are the proton gradient amounting to around 5 units (pH 5.5 outside to pH 10.5 inside) or ~300 mV, and the redox gradient (~1 V) involving interior H₂ and a variety of potential exterior electron acceptors [16,17,25] (Fig. 6). How these candidate driving gradients may have been put to work in principle is considered next.

2.2. Putting the driving disequilibria to work on the cusp between aqueous geochemistry and metabolism

Given the conditions in and around the hydrothermal hatchery as re-reviewed above, we now ask what were the initial, “prerequisite” chemical disequilibria that had to be driven into existence and maintained, where were they produced, how was the dissipation of the candidate driver disequilibria we’ve identified above made to be conditional on the production of these driven disequilibria, what specifically were the turnstile mechanisms that managed these FEC processes, and how did they arise, and function, abiotically?

But we note first that, of course, the view that life is essentially an entropy economy driven by free energy converting processes enjoys a long and storied history—beginning not later than with arguments advanced by Boltzmann himself in the 1880s [30] and quite famously promoted by Schrödinger [31]. Moreover, several previous publications have specifically dealt with how such “entropy trapping” devices may have arisen in the context of the alkaline hydrothermal model itself [9,32–34] and it is on these earlier arguments that the present discussion is erected and which we therefore review in summary next.

2.3. Chemical reaction systems and their potential FEC enablers in the “lost cities” of the Hadean; the first “driven” disequilibria of proto-life

The differentiating stack of entropy-producing FEC mechanisms, arising from within the earth’s core and finally involving surface waters, moving, that is, from convection to metabolism, are at base almost entirely physical in character, though they become more chemically complex as the surface of the Earth is approached and finally reached. The sharpest transition in this passage from physical to chemical gradients appears to be that which takes place in the serpentinization compartment itself wherein, in the main, mechanical and thermal disequilibria are converted, as we have noted, to a host of critical chemical and structural disequilibria. Furthermore, the transition to chemical disequilibria brings in an essential new element for the emergence of life, namely that it is the resultant chemical constructs, initially peptidic we argue, that conferred the beginnings of permanence and memory to free energy converters in the hydrothermal hatchery. Now the machinery and products of FEC can persist for periods in which the driving forces have waned, and the basic prerequisite for a system capable of evolution would have arisen.

Since then, of course, cellular life has developed a whole heterarchical set of complex physical processes, the nature and operations of some of which are the concern of this contribution. But all the requirements of chemosynthetic life—free energy (thermal, chemical, electrochemical), materials (C,H,O,N,P,S), catalytic elements (Fe, Ni, Mo/W, Co, Zn, S and Se)—were focused at the Hadean alkaline submarine hydrothermal vent located away from oceanic spreading centers and their accompanying magma chambers (Figs. 4, 5) [16,35]. The rates of delivery of materials from both the alkaline hydrothermal, and the acidulous ocean

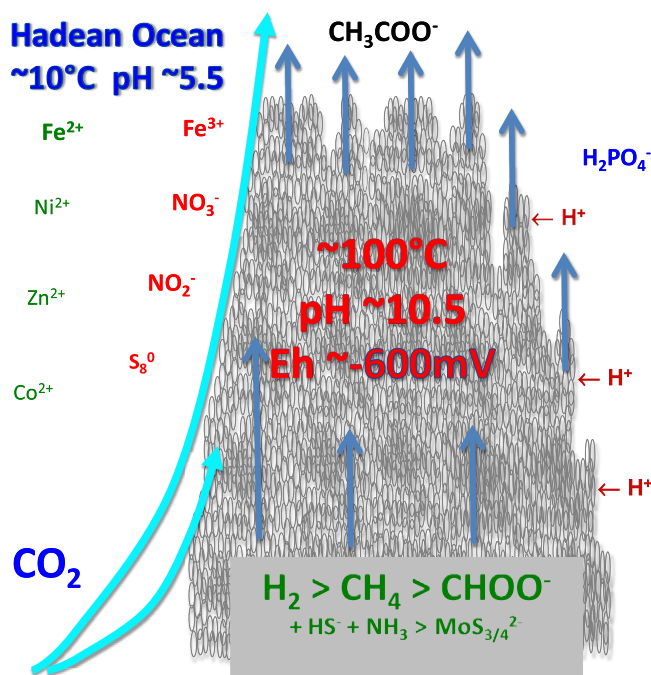


Fig. 5. The hydrothermal mound as a natural carburetor, metacell and hatchery of life. This FEC produces, as its “product” output, extremely complexly organized (low entropy) structures (organized both as literal structures and as processes) which coincidentally provide catalytic centers, electron conductive paths, semi-permeable membranes, compartmentalization at optimal dimensions, chemically active surfaces, which provide such key services as reducing the local water activity and a host of other biasing constraints—geometrical, orientational and electrostatic/polar forces that favor specific reaction outcomes. It also provides for “waste” clearance and disposal. Based on [9,25,33,110,111].

fluids were approximately synchronized partly through the entrainment of the latter by the hot effluent and partly through secondary convection driven by the heat in the mound, both convective and from exothermic mineral precipitation, i.e., life emerged from a low entropy system.

In theory this well-ordered juxtaposition of hydrothermal hydrogen with carbon dioxide dissolved in the Hadean Ocean provided both thermodynamic energy and substrate to emergent life [14,15,29,36–39] and

it has been assumed that metabolism became more and more efficient as geochemical reactions were “quickened” [18]. However, as we’ve argued above, “emergence” requires first and foremost the production—and maintenance—of certain key chemical disequilibria—in useful locations and amounts, and at adequate rates. The central task, therefore, is not just accounting for the right chemistry, it’s accounting for the right “far from equilibrium” and inherently dynamic, chemical FEC processes. This involves, as we earlier stated, identifying the “driven disequilibria” that the earliest possible stirrings of pre-life would have required, the naturally occurring driver disequilibria that could have been exploited for their generation, and the abiotic molecular mechanisms that could have provided the turnstile coupling between them. We have argued above that the available driving disequilibria are primarily two: the redox gradient (≤ 1 V) involving interior H_2 and CH_4 with a variety of potential exterior electron acceptors [16,17,25] and the proton gradient amounting to around 5 units (pH 5.5 outside to pH 10.5 inside) or ~ 300 mV [24].

Given these candidate driver disequilibria, what then were the “driven disequilibria” that would have to have been produced within the mounds micro-compartments for life to have gotten started. First of all, the need for carbon dioxide to be reduced at least as far as formate and carbon monoxide, and for methane to be oxidized to various intermediates such as a methyl and a formyl group—these intermediates interacting to form acetyl-CoA [17]. But life could never have lived on the bread of fixed carbon alone; so it must also have enjoyed the services of a diffusible free-energy carrier system analogous to the ATP disequilibria of contemporary living systems. And as we have discussed, the highly attractive candidate for this role, arguably uniquely so, is pyrophosphate-driven into disequilibrium most credibly by the ambient proton gradient [40]. So provisionally, we suggest that the primary

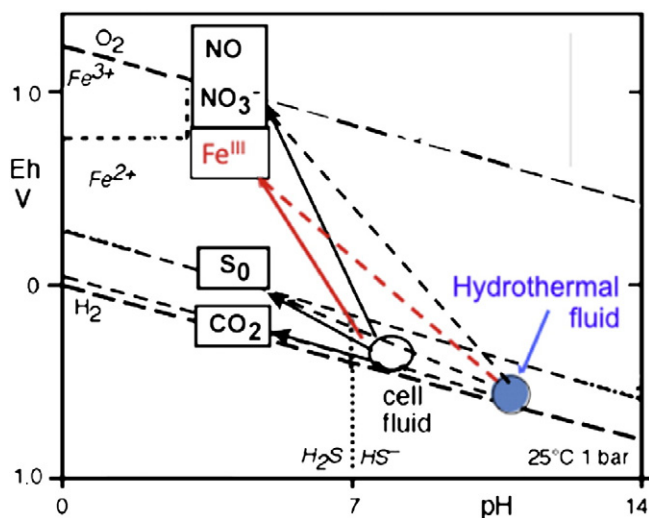


Fig. 6. Eh–pH diagram illustrating the redox potential between hydrothermal hydrogen (and latterly hydrogenotrophs) and redox acceptors available and likely exploitable by an emerging metabolism. After reference [16].

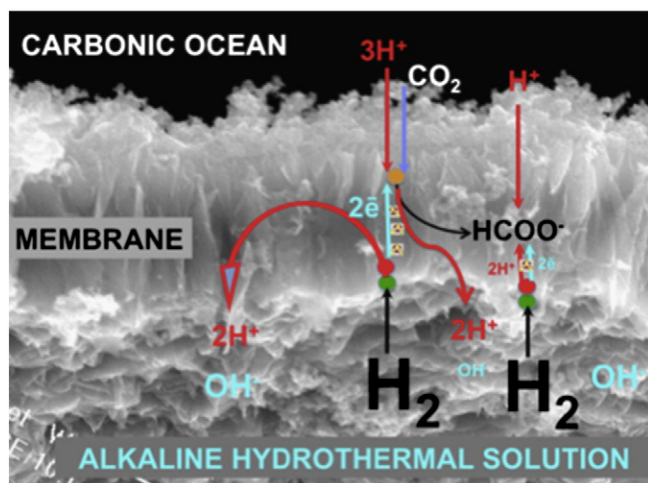


Fig. 7. One mechanistic model for the emergence of the first steps along the reversed acetyl coenzyme-A pathway superimposed on an FeS-bearing low entropy precipitate membrane [53,72–74] (cf. Fig. 5). At $\sim 70^\circ\text{C}$ the sulfides comprise mackinawite $[\text{Fe}_2\text{S}_2]_n$ containing some FeII and subordinate greigite $(\text{SFeS}[\text{Fe}_4\text{S}_4]\text{SFeS})$ [52]. Nickel (green circles) can substitute for iron (red) in both minerals [55,75,76]. As the structural precursors of hydrogenase, carbon monoxide dehydrogenase and acetyl coenzyme A synthase, these mineral nanocrysts may exert some of their catalytic activities, albeit with much lower efficiencies [14,25,32]. The membrane separates a simulacrum of the carbonic Hadean Ocean from the reduced H_2 -bearing alkaline hydrothermal fluid (Fig. 2). The reduction of the carbon dioxide is imagined [16] as a reversal of hydrogenlyase activity as described by Andrews et al. [44]. Carbon dioxide diffuses into the membrane at the same time as the hydrogen diffuses out and meets a nickel-bearing $[\text{NiSFeS}]$ mackinawite that, acting as a protohydrogenase, splits it into two electrons and two protons. The protons are attracted back to the alkaline interior while the electrons are conducted along sulfide nanocrystals toward an adsorbed molybdenum $[\text{Mo}^{\text{IV}}]$ (brown circle) sulfide or selenide entity. Here the CO_2 and one proton are reduced to formate as the molybdenum is oxidized to Mo^{VI} while the two transient protons migrating down gradient stabilize redox cycling. Water activity is kept low by pyrophosphate, recharged through acetyl phosphate thermally and/or by protons, catalyzed by iron sulfides [77]. An alternative model could have involved bifurcation of two electrons downhill to a high potential electron acceptor and two electrons uphill to carbon dioxide [45,46].

initiating event in the emergence of life was the establishment, in the chambers and mineral membranes of inorganic precipitate mounds, of FEC processes that produced and maintained these two “founding” disequilibria, CO_2 versus H_2 and that of methane vs. NO_3^- , NO_2^- and/or

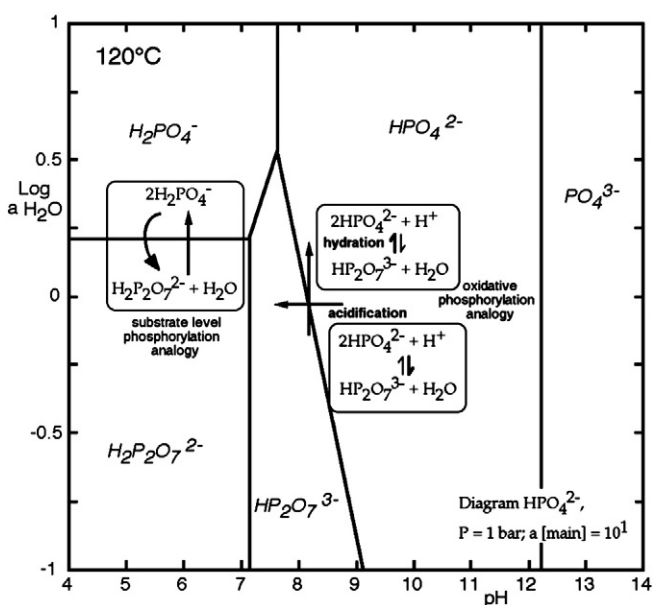


Fig. 8. Diagram drawn with Geochemist's Workbench [114] showing how phosphoanhydride (PPI) can form from orthophosphate as pH and water activity are lowered, as to be expected in the inorganic membrane (cf. Fig. 7) [32].

$\text{Fe}^{\text{III}/3+}$, and $\text{Mn}^{\text{IV}/4+/3+}$, and that of pyrophosphate vs. orthophosphate plus acetyl phosphate [14,17,32,41–43]. Our general reasons for believing that these two classes of disequilibria could plausibly have been driven into existence within the micro-compartments of the serpentine hydrothermal precipitate mounds by the candidate driving disequilibria earlier identified are reviewed next.

As has been argued elsewhere [16,18,24] the ambient proton gradient acting through the molybdenum-bearing inorganic membrane could theoretically drive the hydrogenation of carbon dioxide to produce formate in a reversal of the molybdenum-dependent, proton-translocating formate hydrogenlyase system [44] (Fig. 7). This is a process that should have improved on the merely geochemical/hydrothermal production of the $\sim 150\ \mu\text{mol}$ of formate per kilogram of hydrothermal solution issuing from the vent [27]. An indication of an alternative or additional way that formate may have been produced is given by the operations of *Acetobacterium woodii* [45]. This bacterium overcomes the energetic barrier of CO_2 reduction to formate through the use of soluble electron-bifurcating enzymes of which more later [45–47]. The formate formed within the membrane could have diffused into the alkaline interior of the compartments, possibly facilitated by the ambient proton flow, where, at the high pH, it would convert to CO [24].

A further, and perhaps predominant use of the proton gradient could have been to drive the synthesis of pyrophosphate from ortho- and acetyl-phosphate [15,42,48] (given that at the very origin of life the proton motive force was an ambient condition, i.e., the coenzyme Q pool and cytochrome C were not required inventions before protonic energy could be put to work [49]) (Fig. 8). Various phosphorylations could ensue on mineral surfaces in the confined spaces in the inorganic membrane such as the pyrophosphate-driven phosphorylation of formate to formyl phosphate [15,50]—the activated state reactive with intermediates produced through methane oxidation [17]. Methane oxidation to various C1 organic intermediates, including a methyl group and formaldehyde, may have been effected through the action of external electron acceptors such as NO_3^- , NO_2^- , $\text{Fe}^{3+/\text{II}}$ and $\text{Mn}^{\text{IV}/4+/3+}$ [17,32,41,43]. Iron sulfides such as mackinawite (FeS) in the membrane could have acted as the conducting wire [17,51–53] and molybdenum clusters could have acted to catalyze these reactions in the membrane [17]. Thiolated, we can imagine a methyl group reacting with the CO derived from formate on a nickel iron sulfide cluster with a comparable stoichiometry with greigite (Fe_3NiS_8) [54,55]—a cluster affine with the active center of CODH/ACS [56,57]—thereby producing activated acetate as implied by the Huber–Wächtershäuser experiment [58]. We have argued that such nanocrysts likely acted in many catalytic roles as the precursor components of the redox protein construction kit [24,59–61]. Further hydrogenations of acetate through pyruvate to the higher keto and carboxylic acids catalyzed by the $4\text{Fe}_4\text{S}$ cubanes of the greigite cluster should ensue in a comparable path to the incomplete reduced TCA cycle [62]. Some of these products might then be aminated to amino acids [63]. In turn these amino acids could be phosphorylated within the membrane where water activity was low and there, again driven by the ambient proton motive force, be condensed to peptides [48,64].

Once these amino acids were present within the compartments' mineral walls they could condense to short peptides on the mineral surfaces [65], a process driven by the hydrolysis of pyrophosphates. In doing so they would strengthen the inorganic membranes and perhaps eventually take over their role while retaining the inorganic materials as catalysts as we shall see below [53,66,67]. In lieu of proteins, peptides are likely to have been the first workhorses of metabolism. As such they could have nested the inorganic clusters that otherwise were mineral constituents of the compartment walls; for example, single metal ions such as of Ni and Co, thiolated metal sulfides [48,68,69]. A further example is a peptide nest for phosphate and pyrophosphate [48,67,70]. Two aspects of significance to the emergence of metabolism spring from this: i) one can imagine a feedback mechanism or virtuous circle in which pyrophosphate condenses short peptides that in turn sequester

the phosphates, the better to be condensed by protons coming through the membrane and shuttled into the nest via the carboxylate terminus in what is now a site of even lower water activity, the first inklings of a H^+ -pyrophosphatase [40]. The adsorption to mineral surfaces of the C and N termini of short peptides depends on the pH, and variation in pH causes conformational changes as well as loss and gains of certain ions. These two aspects could both operate at the same time. For example, while the main chain amine groups could sequester the phosphates, any side chain carboxylates could anchor the peptide on a mineral surface [71]. Such a configuration may leave a part of the heterochiral and flexible peptide subject to the charges on proximate ions—changes in conformation that might be useful for shuttling charged entities from one site to another and also acting as gates, levers and even turnstiles. We consider here the potential of such conformational variations to act, as they do in many contemporary examples, as one type of free energy converter mechanism as considered in the following section.

Most of the reactions discussed so far are exergonic at the particular pH values obtaining across the membrane spanning from around pH 6 to pH 11 [38,39] but these only take us so far along the path toward metabolism with its autocatalytic feedback loops and free energy conversions.

2.4. The hard step

The preceding discussion outlines what we believe can currently be said about the possible FEC mechanisms that must have arisen, to initiate the emergence of life, in the alkaline hydrothermal mounds of the Hadean. With some confidence we can propose what chemical disequilibria needed first to be generated, and what disequilibria were available as potential drivers for them. And we can see that within the geochemistry of the serpentinization-driven mounds there are many inviting possibilities as to how the needed FEC turnstile coupling devices could have arisen abiotically. The big leap remains before us, however: what exactly were the pairings of driver to driven, and what exactly were the turnstile mechanisms that made these primal FEC processes work? In the final section we present some general considerations about FEC processes which we hope will be of some help in guiding future efforts to address these deep questions. What kinds of beasts exactly are we looking for and where might we hope to find them?

3. What types of physical mechanisms mediate chemical FEC—and how do they arise in nature?

We begin this section by considering in some detail a small number of specific examples of real-world chemical free energy converting devices. From these we attempt to extract guiding considerations in the search for the mothers of all biological free energy converters. But we emphasize at the outset of this discussion two points. First, the systems discussed below are all examples of FECs that use “pre-existing” or “templated” macroscopically ordered structures to couple driver and driven disequilibria and therefore do not need to create and maintain their own—as do, in contrast, the directly “self organizing dissipative structures” discussed in the opening section (e.g. convection cells). The first example is biological and exemplifies that whereas living systems are certainly autocatalytic “self organizing dissipative systems” they are so “organismically”, and as replicating organisms, each one of which is a vast clockwork of linked FEC processes. The individual FEC devices within are each produced by the indirect processes of biological encoding and cellular reproduction and are only indirectly autocatalytic—through the contribution each makes to the organism's overall function as a replicating FEC. The second is that the essential feature of FEC is that the driving disequilibrium must be natively “frustrated” in its natural drive to dissipate itself—not necessarily completely but “sufficiently” (a river not “frustrated” behind a dam is useless to would-be power producers) [25]. In that situation the FEC mechanism offers to help, but, of course, at a price. In any case, one

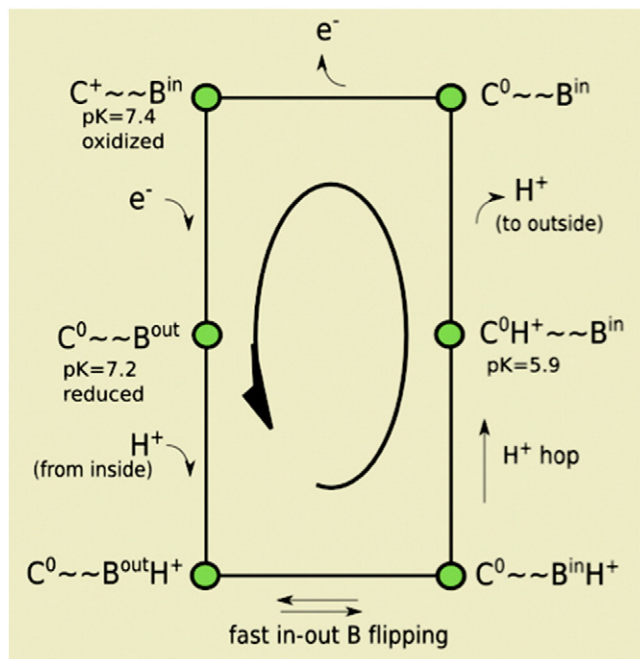


Fig. 9. The symbols $C^{0/+}$ stand, respectively, for the reduced and oxidized forms of the protein's [3Fe4S] cluster, and the symbols $B^{in/out}$ stand, respectively, for the inward and outward facing orientations of the carboxylate O δ atom of the nearby Asp15 (D15) residue [79]. In the “forward” direction, in which the electron (redox) gradient is driving proton translocation (counter-entropically) the reaction proceeds counter-clockwise around the diagram as indicated. In the presumed ground/starting state, the D15 carboxylate moiety is unprotonated (and thus carries an effective negative charge) while the adjacent FeS cluster is oxidized (and is positively charged). In that state the electrostatic attraction between the two draws the flexible carboxylate (its O δ atom) into the protein interior toward the “buried” FeS cluster and away from the surrounding interior medium (the “in” configuration for B). That is, when unprotonated the carboxylate moiety faces ‘in’ toward the adjacent FeS cluster when the latter is oxidized (positively charged), but is turned out into the medium when the FeS cluster is reduced by an incoming electron (delivered by a redox donor “half-reaction”). In that latter orientation the D15 carboxylate group could bind a proton. In the resulting charges state it “oscillates” rapidly between its in-and-out orientations. In this motion the carboxylate acts as a shuttle, able to ferry the proton over the ‘non-conducting’ barrier which sits in the channel between the interior and the FeS cluster. However, in its ‘interior’ orientation, the carboxylate can pass the proton to the then nearby FeS cluster from which, given that the electron driving flow has given it permission, free to dissociate into the “exterior” medium (see the mechanistic model below for a suggestion as to how this comes about). Finally the $C^0 \sim B^{in}$ form can, after it has given the proton leave to pass from the FeS cluster to the “exterior”, lose its electron to the redox reaction's electron acceptor and return to its oxidized state with the Asp15 carboxylate facing inward. Hydrothermal hydrogen is the electron source, freed on the nickel-bearing mackinawite acting as a proto-hydrogenase while the released protons react with the hydroxyls within the mineral compartments.

element of our hunt is to find how the potential driving gradients were naturally frustrated, and then how the production of desired “product” disequilibria could have become mechanistically linked to the “gated” and controlled relief of that frustration.

3.1. The Ferredoxin I protein

The first example considered here is the Ferredoxin I (FdI) protein, a representative biological FEC device which mediates proton translocation anti-entropically (as the disequilibrium-creating driven reaction) using a redox reaction as the driving disequilibrium; in this system, that is, an electron flow and a proton flow are reciprocally coupled [78–80]. An obvious reason for focusing attention on this FEC enzyme, apart from its relative simplicity so far as biological FEC-mediating devices go, is that Fe_4S_4 and Fe_3S_4 clusters were probably readily available in the hydrothermal mound where, in the absence of peptides, they contribute to the mineral greigite— Fe_3NiS_8 [52,55]. Furthermore, short peptides generated in the inorganic membrane would be expected to nest such thiolated iron–sulfur clusters [48,67,81,82] and a C terminal carboxylate could stand in for asp15 to deliver protons from the inner

margins of the membrane as electrons derived from hydrogen in the interior are supplied to the iron sulfur clusters comprising a nickel-bearing mackinawite $(2\text{Fe}_x\text{Ni}_{1-x}\text{S}_2)_n$ which may have acted as an Fe–Ni proto-hydrogenase [16]. And although for our present purposes we are clearly not seeking a means of pumping protons at redox expense, we note that the Fdl redox-driven proton pump would operate in reverse (as do, e.g., ATP synthases [83]), if the thermodynamic strengths of the two opposing gradients were reversed in their relative magnitudes. That is, the same mechanism could allow a proton gradient to drive a reduction reaction endergonically, and it is in this direction of use that primitive abiotic versions of such devices as the Fdl protein might have been employed within the mound.

But to follow up on the possibilities suggested by such examples of chemical FEC mechanisms as the Fdl protein we need, for example, to concretely explicate how the Fe_4S_4 and Fe_3S_4 clusters in the inorganic membranes of the “Lost City” mounds could have functioned as primitive, entirely abiotic versions of the extremely sophisticated FEC mechanisms in modern enzymes in which we find like clusters being used.

To better confront this question we present next a detailed mechanistic model of the Fdl FEC which, to the extent possible, is based directly on the findings of Chen et al. [79]—though to permit a model of the complete FEC process to be framed, some modest extrapolations beyond that paper are required.

3.2. A Hill diagram for the Fdl FEC

The model is presented in two stages. In the first, a “Hill-type” reaction diagram [2] of the model is shown and discussed (Fig. 9). For simplicity we have assumed in this diagram that the process takes place with the two flows “maximally coupled” [1], with no “slippage” or back reactions considered so that the flows are stoichiometrically locked; that is with one proton translocated for each electron driven through the mechanism by the driving redox reaction. In other words, as here imagined, the double turnstile mechanism implicit in the reaction scheme enforces perfect discipline on the two flows it links: an electron is never allowed to sneak downhill without paying its proton pumping ‘probability’ tax, nor are the protons allowed to flow ‘back’ downhill without the coincident event of an electron moving ‘uphill’ against its redox gradient. The real world, of course, is never so tidy. However, in the maximally coupled approximation the reaction scheme is a simple cycle with, e.g., one electron moving from its redox donor to its redox acceptor, and one proton ion being translocated from the ‘inside’ to the ‘outside’ for each traverse of the cycle.

There are two important functional features of this cyclic reaction diagram. First and most essentially, that the reactions occur sequentially in the order indicated. That is, in the model, proton uptake can take place only after electron uptake, the “in–out” flipping of “B”, the carboxylate moiety, can only occur after the uptake of a proton from the membrane-defined “inside”, etc. It is these sequencing requirements that enforce the conditionality that provides the essential coupling feature of an engine (or FEC). Basically, the mechanism says to each individual electron champing at the bit to be allowed to flow downhill from the low voltage redox donor to the high voltage redox acceptor: “I will give you passage if and only if coincident with the event of your passing ‘through’ the FeS cluster, a single proton is translocated (by sheer Brownian chance) against its entropic will from the interior to the exterior”.

Of course, the reaction diagram implies more than this; in particular that the cycle could be reversed, whereby a proton would make its way from the outside to the inside only if an electron has, by improbable chance, reduced the deprotonated FeS cluster by coming from the redox acceptor site when the center was positioned so that it could also take up a proton from the outside; then followed by that proton “hopping” to the unprotonated carboxylate from which it could be ferried across the non-conducting zone from which it could be released into the interior; but that step could take place, presumably, only if the electron on the cluster would have first made the unlikely (anti-entropic) transition

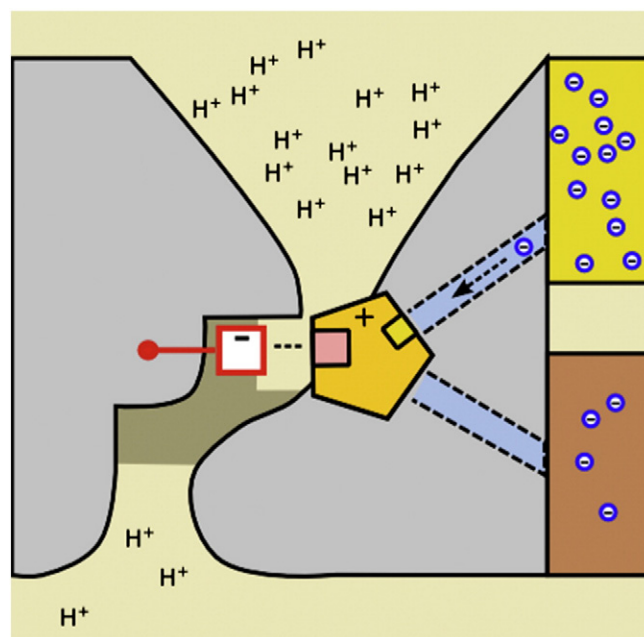


Fig. 10. Mechanism schematic in state 1.

from the cluster to the redox donor—reducing it. On the other hand, if a net entropy gain were to be had by thus driving the cycle “in reverse” (with the proton flowing “in” being the driver of the flow of electrons uphill from redox acceptor to donor, then that is the direction in which the cycle would in fact run—and the engine would be using a proton gradient to drive a chemical reduction. Boltzmann rules!). Implicit in this diagram is that none of the other reactions that could occur between the represented nodes actually do occur. This simplifying approximation is adopted here just to focus on the essential details of the machine; in practice they contribute ‘slippage’ and reverse flow paths that reduce the efficiency of free energy conversion [2].

The next question to address is what specific mechanism details of the protein’s construction and operation would be necessary to accomplish the above cyclic reaction sequence. For this the Chen et al. contribution [79] offers most, but not all, of the answer, so a bit of guessing is required. Of necessity, this Hill-type reaction diagram summarizes a version of ‘what happens’ in the coupling of electron and proton transport but leaves unaddressed how it happens. To address this question, we next present a graphical “mechanistic”, schematic model purporting to represent the functional essentials of the design and operation of the converter. As indicated, this requires making some assumptions that go beyond the explicit content of the current literature [78–80]—in particular concerning the transport and gating, “turnstile” mechanisms effecting the electron transport through the converter, though we believe, these added assumptions are entirely implicit in the cited papers. In this schematic model, the “double turnstile” operation of the converter is not merely evident, but is mechanistically literal.

3.3. A mechanistic model for the Fdl FEC

In Figs. 10 and 11 we present a mechanistic model of Fdl which forms, we propose, a plausible realization of devices that would achieve the above coupling reaction scheme. Most of this model is taken essentially directly from Chen et al. [79], but a few details, having in particular to do with the gating of the electron flow, and by that flow the reciprocal gating of the proton flow, are to one degree or another ‘extrapolations’. However, we think that the real device must be essentially equivalent to this model in functional terms. After showing the model schematic in a presumed ‘start’ configuration, and giving some explanation of its machinery (Fig. 11), we then present at small scale (Fig. 12) a

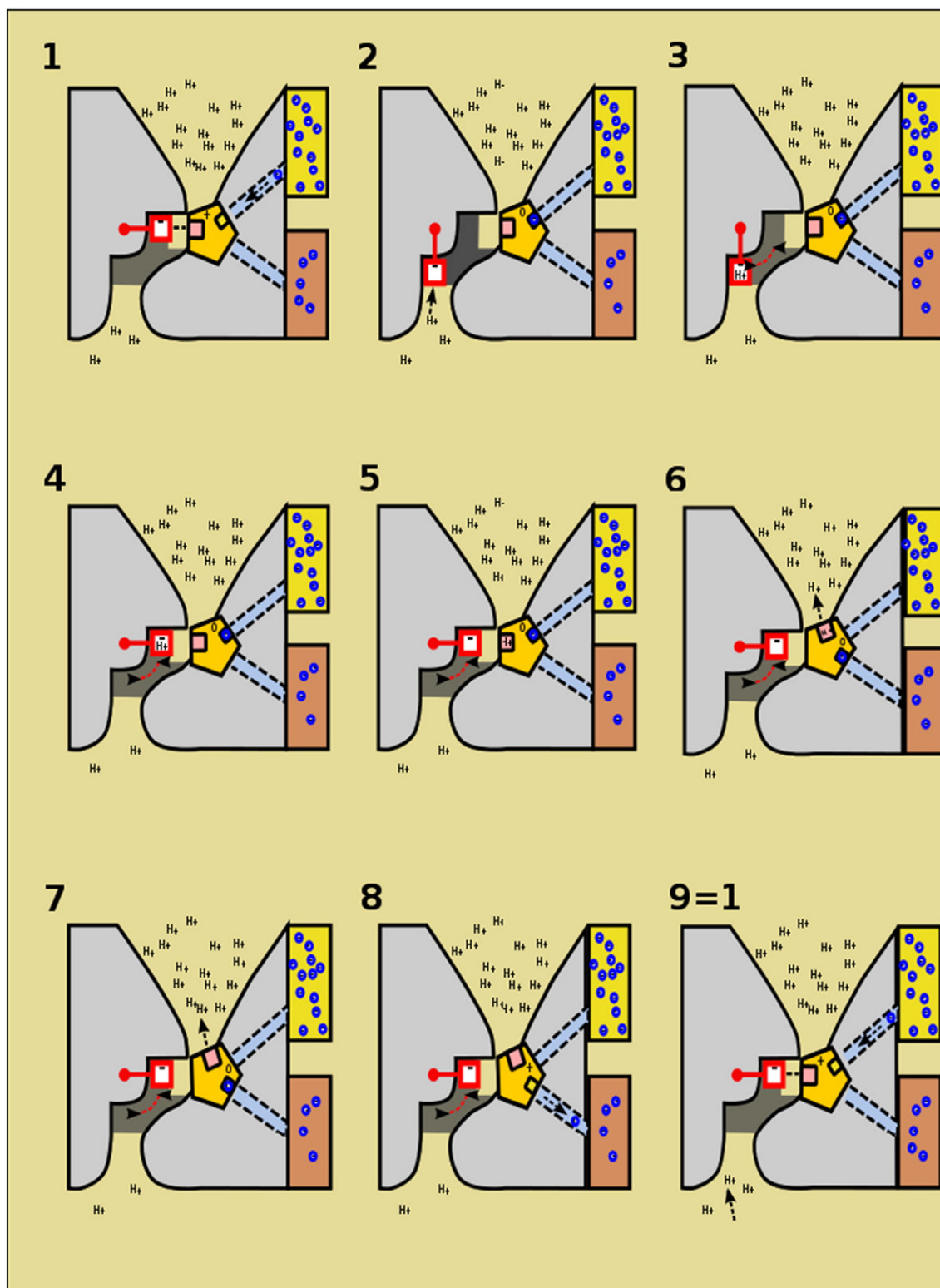


Fig. 11. Mechanism schematic; the FdI reaction cycle involving the gated transfer of protons is represented in 9 sequential diagrams labeled 1 through '9 = 1', of which the first 8 are different and the 9th is identical to the first—indicating the completion of the cycle (cf. references [79] and [91]).

series of 8 different 'state' images that steps the system through 1 cycle (in this representation two of the transitions in the Hill diagram are each resolved into two distinct sub-steps).

There are two 'moving parts' in this schematic diagram, the red box and the orange pentagon (both of which move under Brownian motion and together form a coupled pair of "Brownian ratchets"). The red box attached to the lever arm represents the flexible carboxylate moiety of

the D15 residue; it can rotate between the "buried" position shown and a position in which it is facing "downward" into the interior medium. The orange pentagon represents the $3\text{Fe}4\text{S}$ cluster with its two binding sites indicated: the left (West)-facing pink box binding H^+ , the NE-facing smaller yellow box binding electrons. The cluster can, it is posited, 'rotationally flex' between (1) the orientation shown in which it can accept an electron from the redox donor and then, after

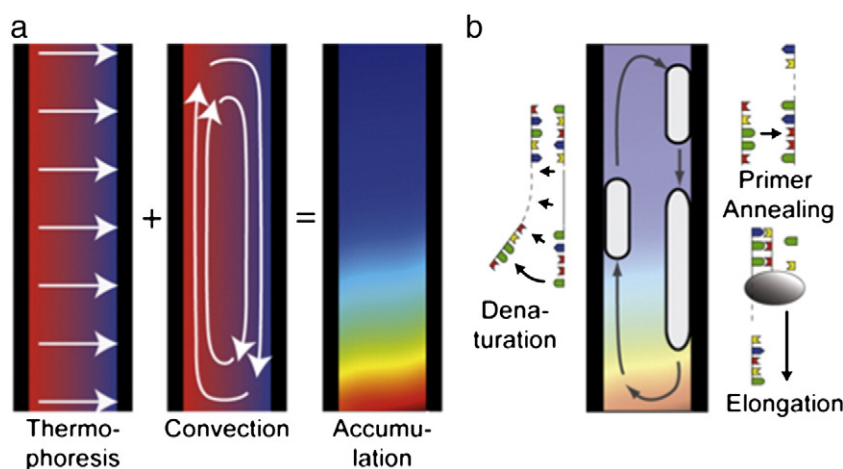


Fig. 12. Experiment modeling the behavior of DNA in a thermal gradient acting across a compartment at the left hand surface of the hydrothermal mound depicted in Fig. 5: (a) the DNA molecules are driven to the right by thermal diffusion and then driven downward by the convection to concentrate at the cool base; (b) illustration of how a natural convective polymerase reaction would operate in the same compartment. This “found art” Free Energy Converter could only come into play after metabolism had reached the stage of generating nucleotides as well as peptides. From reference [85].

it is reduced and electrically neutral, can accept a proton from the facing carboxylate when the latter bears an H^+ —and (2) an orientation in which it can release a bound H^+ to the exterior medium—becoming again neutral, and, following that, release its electron into the channel leading to the redox acceptor (becoming again oxidized and positively charged). The channel connecting the interior and exterior spaces is blocked (is non-conducting) to H^+ translocation in its central portion indicated by dark gray shading. To cross this barrier, protons must be ‘ferried’ over it by being attached to the carboxylate moiety. Electron flow (electrons are represented by small blue circles with interior “-” signs) proceeds from the redox half reaction donor site (upper right box in yellow) through a channel indicated in blue to the FeS cluster to initiate the cycle, and, at the end of the cycle, from the cluster via the lower blue channel to the redox half reaction acceptor site (lower right box in brown). An electron cannot navigate this path unless gated through the FeS cluster by the movements of the FeS cluster—which are controlled by the passage of protons into and out of the cluster.

In the following discussion of the operation of this hypothetical mechanism, we key the steps discussed to the individual numbered “step-by-step” diagrams in Fig. 11 using a curly bracket “{n}” notation. In the configuration shown {1}, the carboxylate is unprotonated and hence bears a (polar) negative charge while the FeS cluster is oxidized and hence bears a positive charge. The electrostatic attraction produced by this charge configuration draws the flexible carboxylate moiety into the protein’s occluded channel “interior” and holds it in close proximity to the H^+ -accepting site of the FeS cluster—and also draws electrons into the cluster from the donor channel. When the FeS cluster is reduced {2} it is (approximately) electrically neutral leaving the carboxylate moiety free to oscillate between its buried and medium-exposed positions. In the latter it can attract (electrostatically) a proton from the proton-conducting channel leading from the “interior” and become charged with H^+ {3}—and have then an approximately neutral charge. In this state it can freely (without charge interference) rotate into the interior, and when there {4} has a significant probability of transferring its proton, via QM tunneling, to the FeS cluster {5}. When that happens the carboxylate is again negatively charged and the FeS cluster has an effective positive charge. To this point, the picture is, we consider, essentially just a summary of the conclusions of Chen et al. [79].

However, for the machine to actually function as a proton pump driven by the redox electron gradient, it must be possible for the proton associated with the reduced FeS cluster to dissociate into a proton-

conductive channel exposed to the exterior, and after that, but only then, for the electron bound to the cluster to dissociate into an electron conductive channel leading to the redox acceptor. These two steps appear to require, either literally, or in effect, that when the FeS cluster is both reduced and charged with a proton (so that its effective charge is positive and the now deprotonated carboxylate is negative) this charge configuration permits the FeS cluster to “flex” (partially rotate ‘clockwise’ in the diagram) to an orientation that exposes its H^+ -site to the exterior medium and its electron site to the channel connected to the redox acceptor {6}. If the proton then dissociates from the cluster into the exterior {7} the bound electron is, in that configuration, freed to dissociate into the electron conductive path leading to the redox acceptor {8}. In this state both the carboxylate and the FeS cluster are uncharged and the cluster is oxidized—so, we presume, the cluster is free to (and, due to the charge attraction, biased to) return to its starting orientation with the deprotonated carboxylate attracted into its “buried” position opposite the FeS cluster—and with that return the entire machine to its original state {9=1}.

Notwithstanding this long-winded mechanistic tale, the FdI protein is an example of a relatively simple class of biological FEC mechanisms, due mostly to its simple operational stoichiometry of coupling the passage of one proton to the passage of one electron. Yet clearly it is at the same time far too complicated and “evolved” to be in any direct way a model of what could have arisen abiotically. On the other hand, it is hard to overlook the suggestions implicit in the similarities noted above between the chemistry of the alkaline hydrothermal mounds and the central molecular devices of FdI and, indeed, of essentially all other biological FECs whose mechanisms have been elucidated. Are modern protein-based FEC devices just the evolutionarily tricked-up direct descendants of vastly simpler “ancestral”, entirely abiotic, and at the root purely geochemical, machines? If this is plausible, how exactly might such progenitors have arisen? What would they have looked like? Are there suggestive examples in the present world of similar simple FECs to be found outside of life’s sphere (or indeed within it)?

To pursue this question a step further we first recognize that even if we could see how to abiotically couple specific redox and concentration disequilibria in a driver-driven FEC pair, that would not meet all of our requirements. We are therefore interested in any mechanism that might couple either of the two types of driver disequilibria we have identified (the geochemical proton gradient and, for example, the $H_2 + CO_2$ redox gradient), to either of the two candidate driven disequilibria (generating formate and pyrophosphate). And it would certainly be helpful, if the

mound provided not only the needed reagents, driving disequilibria, semi-permeable membrane-based compartmentalization, redox active centers, etc., needed, but also, in abiotic form, and as “found art” (so far as our coupling needs are concerned), the turnstile devices required to drive the specific FEC processes we have identified as essential. We explore this possibility briefly in the following section.

3.4. Is it possible that it all began with “found art” FECs?

All biologically ‘encoded’ FEC mechanisms, such as the FdI protein, straddle the pivotal distinction made by Cottrell [7], and discussed in the opening section, between FECs that are directly “self emergent” and those that are “templated” or “facilitated” (by pre-existing structures); they are the former at the organism/evolutionary level and the latter at the molecular machine level. But this magically potent straddling trick would not have been available to the first bio-FECs, so on which chair would they have sat? It is possible to suppose that the first bio-FECs arose as fully “self organizing dissipative structures”—i.e. as a chemical analog to tornadoes and hydrothermal convection cells, generating their own turnstile machines “out of thin air” under the goad of an extremely far-from-equilibrium gradient—and being fully autocatalytic as individual machines from the outset.

But this seems a hard stretch and also to unwisely overlook the possibility of a “facilitated/templated” beginning to bioenergetics afforded by the true Valhalla of macroscopic organization that the alkaline hydrothermal vents generate. As we have seen, these systems are organized and structured in seemingly every sense and manner one might desire: *physically* (compartments and membranes in vast and regular arrays at the ~10 micron level; the formation of a semi-permeable boundary between ~100 °C vent effluent and the cool Hadean sea), *chemically* (the great disequilibria embodied in highly concentrated and spatially differentiated chemical compositions; hydrogen, methane ammonia and bisulfide in the hydrothermal solution presented in ordered manner to the carbon dioxide and phosphate in the ocean); and *geochemically* (e.g. the FeS clusters of mackinawite (FeS) acting as a hydrogenase and Rieske cluster, greigite (Fe₃NiS₈) as a precursor to the low and high potential ferredoxin, as well as ACS/CODH, a molybdenum cluster Mo₃^{IV/VI}Fe₅S₁₄^{0/2-} acting as an electron bifurcating enzyme)—all intercalated within and comprising the well-ordered mineral structures of the inorganic precipitate membranes [16,84].

To see what possibilities might lie in this direction we next briefly examine the nature of “found art” FEC systems. Could the serpentine mounds have “unwittingly” provided not just the driving gradients, the compartments, and the reactants for the required driven gradients, but also the specific organized structures needed to tie driver and driven processes together into just the functioning FEC devices required by proto-biology? If that is indeed what happened, then the alkaline hydrothermal vents and their precipitate mound are rightly to be indicted for having themselves provided everything needed to set in motion, and direct, the greatest qualitative transition in the behavior of matter ever to have occurred.

3.5. FEC for the taking

As noted in the introductory section, many naturally occurring FEC systems exist which exploit preexisting “macroscopically ordered structures” that have arisen through unrelated FEC processes [7]. These structures are used, more-or-less as found, to mediate the linking of driven to driver processes essential to FEC—which therefore do not need to be generated by the driving gradient itself; they are simply “there for the taking” so far as the subject FEC process is concerned. For this reason, as also noted earlier, strong, far-from-equilibrium gradients are not required to get engines in this class started (even those that have, once started, strongly amplifying autocatalytic properties), and quite modest gradients often suffice [7].

Natural examples in this class are numerous and include, for example, a variety of so-called electrokinetic phenomena such as thermophoresis, i.e. the Soret effect and its inverse, the Dufur effect. See also the examples from materials physics described by Cottrell [7] involving structural collapse, plastic deformation, and crack propagation; mechanisms that are likely to be relevant to the action of the serpentinization process itself as a FEC converting mechanical stress into the physical and chemical disequilibria that, we conjecture, make the serpentine mounds the incubators of life.

A striking and relevant example of “found art” FEC involves a thermophoretic flow interacting with a convective engine—and functioning as a highly “compound” system in which no less than six coupled flows are involved. The superimposed thermal gradient drives both the convection cell and thermal diffusion flows; these in turn interact to drive the production of a concentration gradient along the system’s long axis—leading in turn to polymerization and a molecular species disequilibrium [35,85] (Fig. 12). The convective polymerase chain reaction generates long charged polymers of DNA (and predicted for RNA) which are then driven toward the coolest reaches of a compartment or series of interlinked compartments. This might set the scene for the development of variation and selection of what would have been effectively retroviruses living off metabolic energies [38,87,88].

But perhaps the most interesting possibility in regard to the potential role of “found art” FEC devices in the origin of life—“art” generated specifically by the serpentinization process—arises in connection with the fact that one of the key driving disequilibria we have identified involves the oxidation of H₂. Here the point is that H₂ oxidation offers a remarkable thermodynamic gift in that it can yield up pairs of electrons in a single FEC ‘step’. In this two-electrons-at-a-time circumstance specific turnstile mechanisms can arise which permit one electron to undergo an entropy-gaining transition—and thus be a “driver” process—only if its partner electron undergoes an entropy-losing transition—and is thus a “driven” process; in consequence the free energy gain of the driven electron can be much higher than would be possible in an otherwise comparable single electron redox reaction. This remarkable trick is called “electron bifurcation” and its possible role in the ‘abiotic’ emergence of bioenergetics will be discussed next.

3.6. Electron bifurcation

The bifurcation of electrons in the Q cycle was first deduced by Mitchell as a mechanism to explain the operation of the proton motive force and chemiosmosis [89,90]. And since, Brandt [91] has described a proton-gating charge transfer mechanism that drives a two-phase electron bifurcation electron flux in the Q-cycle. It is, in effect, a “turnstile gating” mechanism—that couples the two flows and makes the whole machine a Free Energy Converter. Since these breakthroughs, flavoproteins have recently been added to those organic ring complexes capable of accepting two electrons and two protons from a hydrogenase and dispensing them to electron acceptors of contrasting potentials [86,92,93]. Electron bifurcation also promotes the reduction of electron acceptors that otherwise have too low a potential themselves to attract an electron. In the so-called crossed-over mode a pair of electrons can be almost simultaneously drawn toward different electron acceptors—the first a high potential entity and the second at a low potential [16,93,94]. The significance of the process is that a low potential acceptor can accept the high energy electron left in the same orbit if the first electron was lost to an irresistible high potential acceptor. In biochemistry the two contrasting acceptors are often a low potential ferredoxin and a high potential organic heterodisulfide. However, at the emergence of life the two acceptors might be on the one hand photolytically generated ferric iron and Mn^{IV/3+} or volcanic nitric oxide/nitrate/nitrite [21,41,43,95] and on the other the recalcitrant CO₂ molecule (Fig. 6). The first electron could not be abstracted until a high potential acceptor was contacted. As the oxidized nitrogen anions and/or iron (III) or manganese (IV) cations that had the potential to attract one electron would

have been sparse relative to the high concentrations of carbon dioxide in the Hadean Ocean (~200 mM at a 10 bar pressure) the second “hot” electron could now discharge to the almost all-enveloping CO₂.

But what could have acted in the prebiotic world in the place of the organic complexes involved in the present day cofactors? Molybdenum (and tungsten at higher temperature) in dithiolenes can probably elicit a similar response to electron acceptors of contrasting potential [17]. Molybdenum is known for its catalytic properties in chemistry and as a redox enzyme in an extraordinary number of biochemical roles and across the whole range of electrochemical potentials [96–98]. It is most effective in both roles when ligated with sulfur [99]. Although it is a rare element, especially in cool reduced water, it is relatively soluble in alkaline bisulfide solutions (as is tungsten) while other transition metals are not [84,100,101]. Hence it is likely to have been precipitated in a submarine hydrothermal iron sulfide-rich mound, perhaps as Fe₅Mo₃S₁₄ (or Fe₂MoS₅ or Fe₃Mo₂S₉) [84]. Vorlicek et al. [102] suggest that the precipitation of molybdenum from anoxic sulfidic waters may involve zero-valent sulfur. Notably, polysulfide (S₈⁰) is an expected constituent of the early Earth's atmosphere [103] and, given bidentate sulfur ligands, we might expect molybdenum to be readily oxidized or reduced (Mo^{IV} ⇌ Mo^{VI}) in such an inorganic cluster precipitated in the membrane [102,104]. These considerations suggest that molybdenum, complexed with iron and sulfur, perhaps as Mo₃^{IV/VI}Fe₅S₁₄²⁻ [84] and in contact with the iron sulfide clusters, may have provided the FEC “turnstile” needed to tie proton translocation—or alternatively H₂ oxidation—to CO₂ reduction (Fig. 7).

In any case, one of the external high potential electron acceptors is assumed to have abstracted an electron from a molybdenum dithiolene acting as a two electron source (ultimately derived from H₂ and prior to the invention of quinones and pterins) such that the remaining bifurcating “hot” electron had the free energy to reduce CO₂ to formate [16,17,93]. Thus we speculate that a Fe₅Mo₃S₁₄ cluster precipitated in the inorganic membrane and capable of bifurcating electrons was the precursor to the Complex-Iron-Sulfur-Molybdoenzymes (CISM) prior to the emergence of the pterins [97]. In this form molybdenum may have been involved in the reduction of carbon dioxide to formate in what was to become the so-called “western branch” of the reductive acetyl coA pathway, a pathway common to both the tetrahydromethanopterin and the tetrahydrofolate catalyzed “eastern branch” of the acetyl coA pathway [105].

4. The segue from free proton gradient usage to full blown chemiosmosis

Section 3 outlines examples of how the serpentine mound precipitates (themselves the product of geochemical FEC processes) could have ‘inadvertently’ provided the organized macroscopic turnstile mechanisms needed to effect the transitional FEC processes on the road to life. The critical initial conditions are the “free” provision of thermal, redox and, of course, the proton gradients. How then might we reconcile the use of the ambient proton gradient to get life started with what Mulkidjanian and colleagues have called the “evolutionary primacy of sodium bioenergetics” [106,107]. To avoid misunderstanding we emphasize the need to distinguish the two fundamental FEC aspects of the chemiosmotic hypothesis: its generative, ion-pumping aspect, which drives the production of a trans-membrane ion gradient, and its consumptive aspect through which that gradient is put to useful work in the cell, most universally, to drive the production of the ATP disequilibrium. At its origins in the hydrothermal mound, we posit, a proton gradient was made available “for free” and only the consumptive portion of the chemiosmotic cycle needed at that point to be invented (specifically, we propose the FEC mechanism necessary to allow that gradient to drive the endergonic condensation of pyrophosphate). Thus in this model, the generation of pyrophosphate within the superficial chambers of the mound, driven by the ambient proton gradient is a given of the initial conditions. No pump, neither proton nor sodium,

was required to generate condensed phosphate at this early stage of development. Pumping only became necessary for a metabolizing compartment when it became separated from the natural proton gradient through superimposed precipitates on the mound's margin. Only then did the protocell have to start making its own gradient using whatever free energy resources were still available to energize the pump. While the pH gradient would have been dissipated just beneath the mound's outer surface, the metastable high potential electron acceptors were likely still available so that any abundant small cation could be pumped using electron currency in the manner described in Section 3.3. That this was likely sodium as Mulkidjanian and colleagues have suggested is because there would have been no sodium gradient to resist [112] and anyway the early membrane would have been less permeable to sodium ions than to protons, the better to maintain the imposed gradient and put it to use [106]. Indeed, a sodium ion can replace protons as exemplified in *A. woodii* where a sodium-motive ferredoxin:NAD oxidoreductase (Rnf) is responsible for the Na⁺ translocation in a process that involves electron bifurcation [45]. We can imagine an induced sodium gradient taking over from the ambient proton gradient and while the NAD protein cannot be appealed to at the emergence of life, iron sulfides do have an affinity with sodium as evidenced by erdite (Na⁺Fe³⁺S₂²⁻ · 2H₂O) [113], a mineral cluster produced in hot alkaline solution that could conceivably have taken on a comparable role to the FdI protein.

This adversity having been met left the offspring of such cells much better equipped to face the diaspora where they could use either sodium or proton pumping or both depending on ambient conditions. However, looking back we can see that the consumption side of the chemiosmotic mechanism was there from the beginning and never lost and is the single thread that links the first metabolizing protocells with all modern cellular life.

5. Conclusions

Bioenergetics is the antidote to what the late Bob Shapiro termed ‘molecular vitalism’, a cast of mind held in thrall to the information age and guiding most of the research into the origin of life. Leduc's ‘Mechanism of Life’ reminds us that it was not always thus. Reaching across a century he remarks “All living things consist of closed cavities, which are limited by osmotic membranes, and filled with solutions of crystalloids and colloids. The study of synthetic biology is therefore the study of the physical forces and conditions which can produce cavities surrounded by osmotic membranes, which can associate and group such cavities, and differentiate and specialize their functions. Such forces are precisely those which produce osmotic growths, having the forms and exhibiting many of the functions of living beings. Of all the theories as to the origin of life, that which attributes it to osmosis and looks on the earliest living beings as products of osmotic growths is the most probable and the most satisfying to the reason”.... “During these long ages an exuberant growth of osmotic vegetation must have been produced in these primeval seas. All the substances which were capable of producing osmotic membranes by mutual contact sprang into growth, the soluble salts of calcium, carbonates, phosphates, silicates, albuminoid matter, became organized as osmotic productions, were born, developed, evolved, dissociated, and died. Millions of ephemeral forms must have succeeded one another in the natural evolution of that age, when the living world was represented by matter thus organized by osmosis”.

Bringing these ideas onto a modern stage we develop the view that life sprang into being as the free-energy-converting spawn of a planet itself born far from thermal, geochemical and electrochemical equilibrium. By 4.4 billion years ago Earth was somewhat (photo)oxidized (H₂O > CO₂ > NO ~ S⁰ > Fe^{III/3+} > Mn^{IV/4+/3+}) and cool on the outside and reduced (Fe^{II} >> Fe^{III}) and hot on the inside. Hydrogen emanated from the interior and a mildly acidic carbonic ocean enveloped the planet. The chemical (H₂ vs CO₂) and electrochemical (proton and redox

gradient ≈ 1 V) disequilibria at the surface were relieved, on the lifeless planet, extremely slowly. While life emerged as one of the planet's mechanisms for accelerating the rate of dissipating these disequilibria (equivalently its rate of entropy production), biochemistry should not be thought of as faster or better equilibrium chemistry, as often assumed—it is qualitatively different. It is not enough, in other words, to add ΔG s and be contented when the sum is negative. There is a need to search for specific “turnstile” coupling mechanisms or engines that do the real job of Free Energy Conversion whereby particular exergonic reactions are mechanistically united with—and drive—particular endergonic reactions. And, particularly when searching for the genesis of life, there is likewise a need to place these processes in the context of the entire stack of dissipating mechanisms operating on the planet 4.4 billion years ago. Indeed, we recognize that the primary and secondary convection cells operating from the core to the surface of the planet are themselves FEC “heat engines” accelerating the dissipation of the earth's thermal disequilibrium. At the surface these convective processes also acted as FEC devices which generated, primarily by driving serpentinization, both a new and essential suite of chemical disequilibria, and at the same time new and essential organizational/structural and distributional disequilibria at the sites of alkaline submarine hydrothermal vents fed from serpentinization reactions. The resulting extremely “organized” submarine hydrothermal structure acted as a chemical and electrochemical reactor, affinity column, organic molecule concentrator and condenser and ultimately as the hatchery of life [25,35,108]. It both focused and concentrated on the chemical disequilibria at the “steepest gradient”, i.e., at “membranous” interfaces. These interfaces also provided chemically active, macroscopically organized material structures closely reminiscent of the core devices involved in the “turnstile mechanisms” of the chemical FEC machines of extant life. Indeed, we can see how Mitchell's chemiosmosis—the required addition to Leduc's merely osmotic membranes—may have emerged: first in the form of the “consumptive” side of the chemiosmotic cycle powered by the ambient proton gradient and driving the condensation of pyrophosphate. This stage must have been supplanted by a fundamental transition to self-generated ion-gradients to sustain the, by that time, unalterable dependence of proto-metabolism on the chemiosmotic mechanism. This transition was most likely based on using ambient redox disequilibria to pump sodium and sustain a sodium-based chemiosmotic economy. Subsequently, when fully modern biological membranes evolved and relatively rich sources of free energy were more widely available, there was the predominant return to the use of proton gradients driven by 2-electron redox reactions—most likely exploiting electron bifurcation from the beginning.

It is by means of these and other ‘mineralogical’ and peptidic devices, we propose, that the geochemical disequilibria present within the mound became coupled, as drivers, to the generation of the “founding” bio-chemical disequilibria that began the great, ‘long path’, autocatalytic FEC spiral that is life on this planet; an autocatalytic spiral that has ever since quickened the planet's entropy productivity [26].

Acknowledgements

This contribution is dedicated to Georg Fuchs who cleared the path(s). We appreciate help from Wolfgang Nitschke, Nigel Goldenfeld, Nick Lane, Laurie Barge, Lauren White, Takazo Shibuya, Richard Kidd and Isik Kanik and thank Dr. Carl Pilcher and the members of the NAI-sponsored Thermodynamics Disequilibrium and Evolution Focus Group for discussions. The research described for this publication was carried out at the Institute for Genomic Biology, University of Illinois at Urbana-Champaign and at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration: with support by NASA Exobiology and Evolutionary Biology award (NNH06ZDA001N) and supported by the NASA Astrobiology Institute (Icy Worlds). US Government sponsorship acknowledged.

References

- [1] S.R. Caplan, A. Essig, *Bioenergetics and Linear Nonequilibrium Thermodynamics; The Steady State*, Harvard University, 1999.
- [2] T.L. Hill, *Free Energy Transduction and Biochemical Cycle Kinetics*, Dover edition Springer-Verlag, N.Y., 1989. 2005.
- [3] K.A. Dill, S. Bromberg, *Molecular Driving Forces; Statistical Thermodynamics in Chemistry and Biology*, Garland Science, 2003.
- [4] J.W. Stucki, The optimal efficiency and the economic degrees of coupling of oxidative phosphorylation, *Eur. J. Biochem.* 109 (1980) 269–283.
- [5] S.R. de Groot, P. Mazur, *Non-equilibrium Thermodynamics*, Dover, 1984.
- [6] A. Cottrell, The natural philosophy of engines, *Contemp. Phys.* 20 (1979) 1–10.
- [7] G. Nicolis, I. Prigogine, *Self-organization in Non-equilibrium Systems: From Dissipative Structures to Order Through Fluctuations*, J. Wiley & Sons, New York, 1977.
- [8] D. Kondepudi, I. Prigogine, *Modern Thermodynamics; From Heat Engines to Dissipative Structures*, Wiley & Sons, 2006.
- [9] M.J. Russell, A.J. Hall, A. Mellersh, On the dissipation of thermal and chemical energies on the early Earth: the onsets of hydrothermal convection, chemiosmosis, genetically regulated metabolism and oxygenic photosynthesis, in: R. Ikan (Ed.), *Natural and Laboratory-Simulated Thermal Geochemical Processes*, Kluwer Academic Publishers, Dordrecht, 2003, pp. 325–388.
- [10] L. Onsager, Reciprocal relations in irreversible processes, I, *Phys. Rev.* 37 (1931) 405–426.
- [11] J.W. Stucki, The thermodynamic-buffer enzymes, *Eur. J. Biochem.* 109 (1980) 257–267.
- [12] M.J. Russell, R.M. Daniel, A.J. Hall, On the emergence of life via catalytic iron sulphide membranes, *Terra Nova* 5 (1993) 343–347.
- [13] G. MacLeod, C. McKeown, A.J. Hall, M.J. Russell, Hydrothermal and oceanic pH conditions of possible relevance to the origin of life, *Orig. Life Evol. Biosph.* 24 (1994) 19–41.
- [14] M.J. Russell, W. Martin, The rocky roots of the acetyl coenzyme-a pathway, *TiBS* 24 (2004) 358–363.
- [15] W. Martin, M.J. Russell, On the origin of biochemistry at an alkaline hydrothermal vent, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362 (2007) 1887–1925.
- [16] W. Nitschke, M.J. Russell, Hydrothermal focusing of chemical and chemiosmotic energy, supported by delivery of catalytic Fe, Ni, Mo/W, Co, S and Se, forced life to emerge, *J. Mol. Evol.* 69 (2009) 481–496.
- [17] W. Nitschke, M.J. Russell, Beating the acetyl coenzyme — A pathway to the origin of life, *Philos. Trans. R. Soc.* (in press).
- [18] M.J. Russell, A.J. Hall, W. Martin, Serpentinization and its contribution to the energy for the emergence of life, *Geobiology* 8 (2010) 355–371.
- [19] B.J. Wood, M.J. Walter, J. Wade, Accretion of the Earth and segregation of its core, *Nature* 441 (2006) 825–833.
- [20] K. Zahnle, N. Arndt, C. Cockell, A. Halliday, E. Nisbet, F. Selsis, N.H. Sleep, Emergence of a habitable planet, *Space Sci. Rev.* 129 (2007) 35–78.
- [21] R.S. Martin, T.A. Mather, D.M. Pyle, Volcanic emissions and the early Earth atmosphere, *Geochim. Cosmochim. Acta* 71 (2007) 3673–3685.
- [22] C. Bounama, S. Franck, W. von Bloh, The fate of the Earth's ocean, *Hydrol. Earth Syst. Sci.* 5 (2001) 569–575.
- [23] L.T. Elkins-Tanton, Linked magma ocean solidification and atmospheric growth for Earth and Mars, *Earth Planet. Sci. Lett.* 271 (2008) 181–191.
- [24] M.J. Russell, R.M. Daniel, A.J. Hall, J. Sherrington, A hydrothermally precipitated catalytic iron sulphide membrane as a first step toward life, *J. Mol. Evol.* 39 (1994) 231–243.
- [25] M.J. Russell, A.J. Hall, The emergence of life from iron monosulphide bubbles at a submarine hydrothermal redox and pH front, *J. Geol. Soc. Lond.* 154 (1997) 377–402.
- [26] E. Simioncini, M.J. Russell, A. Kleidon, Modeling free energy availability from Hadean Hydrothermal Systems to the first metabolism, *Orig. Life Evol. Biosph.* 41 (2011) 529–532.
- [27] S.Q. Lang, D.A. Butterfield, M. Schulte, D.S. Kelley, M.D. Lilley, Elevated concentrations of formate, acetate and dissolved organic carbon found at the Lost City hydrothermal field, *Geochim. Cosmochim. Acta* 74 (2010) 941–952.
- [28] W. Martin, J. Baross, D. Kelley, M.J. Russell, Hydrothermal vents and the origin of life, *Nat. Rev. Microbiol.* 6 (2008) 806–814.
- [29] D.S. Kelley, J.A. Karson, G.L. Früh-Green, D.R. Yoerger, T.M. Shank, D.A. Butterfield, J.M. Hayes, M.O. Schrenk, E.J. Olson, G. Proskurowski, M. Jakuba, A. Bradley, B. Larson, K. Ludwig, D. Glickson, K. Buchman, A.S. Bradley, W.J. Brazelton, K. Roe, M.J. Elend, A. Delacour, S.M. Bernasconi, M.D. Lilley, J.A. Baross, R.E. Summons, S.P. Sylva, A serpentine-hosted ecosystem: the Lost City hydrothermal field, *Science* 307 (2005) 1428–1434.
- [30] B. McGuinness, Ludwig Boltzmann, *Theoretical Physics and Philosophical Problems*, Kluwer, 1974.
- [31] E. Schrödinger, *What is Life? The Physical Aspects of the Living Cell*, Cambridge University Press, Cambridge, 1944.
- [32] M.J. Russell, A.J. Hall, The onset and early evolution of life, in: S.E. Kesler, H. Ohmoto (Eds.), *Evolution of Early Earth's Atmosphere, Hydrosphere, and Biosphere—Constraints from Ore Deposits*, Geological Society of America, Memoir, 198, 2006, pp. 1–32.
- [33] M.J. Russell, The alkaline solution to the emergence of life: energy, entropy and early evolution, *Acta Biotheor.* 55 (2007) 133–179.
- [34] W. Nitschke, M.J. Russell, Just like the universe the emergence of life had high enthalpy and low entropy beginnings, *J. Cosmol.* 10 (2010) 3200–3216.
- [35] P. Baaske, F.M. Weinert, S. Duhr, K.H. Lemke, M.J. Russell, D. Braun, Extreme accumulation of nucleotides in simulated hydrothermal pore systems, *Proc. Natl. Acad. Sci. U. S. A.* 104 (2007) 9346–9351.

- [36] E.L. Shock, Chemical environments of submarine hydrothermal systems; marine hydrothermal systems and the origin of life, *Orig. Life Evol. Biosph.* 22 (1992) 67–107.
- [37] M.J. Russell, A.J. Hall, D. Turner, In vitro growth of iron sulphide chimneys: possible culture chambers for origin-of-life experiments, *Terra Nova* 1 (1989) 238–241.
- [38] J.P. Amend, T.M. McCollom, Energetics of biomolecule synthesis on early Earth, *Chemical Evolution II: from origins of life to modern society*, Am. Chem. Soc. (2009) 63–94.
- [39] J.P. Amend, T.M. McCollom, M. Hentscher, W. Bach, Catabolic and anabolic energy for chemolithoautotrophs in deep-sea hydrothermal systems hosted in different rock types, *Geochim. Cosmochim. Acta* 75 (2011) 5736–5748.
- [40] M. Balseffsky, A. Schultz, H. Balseffsky, H⁺-PPases: a tightly membrane-bound family, *FEBS Lett.* 457 (1999) 527–533.
- [41] A.L. Ducluzeau, R. van Lis, S. Duval, B. Schoepp-Cothenet, M.J. Russell, W. Nitschke, Was nitric oxide the first deep electron sink? *Trends Biochem. Sci.* 34 (2009) 9–15.
- [42] L.M. Barge, L.M. White, I. Doloboff, I. Kanik, M.J. Russell, Pyrophosphate generation via an ambient proton-motive force: laboratory tests of an alkaline hydrothermal origin-of-life scenario, in: *Origins 2011: Joint ISSOL and Bioastronomy International Conference*, Montpellier, France, July 2011, 2011.
- [43] A.D. Baner, H.D. Holland, The photochemistry of manganese and the origin of banded iron formations, *Geochim. Cosmochim. Acta* 56 (1992) 2595–2603.
- [44] S.C. Andrews, B.C. Berkst, J. McClay, A. Amblert, M.A. Quail, P. Golby, J.R. Guest, A 12-cistron *Escherichia coli* operon (huf) encoding a putative proton-translocating formate hydrogenlyase system, *Microbiology* 143 (1997) 3633–3647.
- [45] A. Poehlein, S. Schmidt, A.-K. Kaster, M. Goenrich, J. Vollmers, A. Thurmer, J. Bertsch, K. Schuchmann, B. Voigt, M. Hecker, R. Daniel, R.K. Thauer, G. Gottschalk, V. Muller, An ancient pathway combining carbon dioxide fixation with the generation and utilization of a sodium ion gradient for ATP synthesis, *PLoS One* 7 (2012) e33439.
- [46] G.J. Schut, M.W. Adams, The iron-hydrogenase of *Thermotoga maritima* utilizes ferredoxin and NADH synergistically: a new perspective on anaerobic hydrogen production, *J. Bacteriol.* 191 (2009) 4451–4457.
- [47] S. Wang, H. Huang, J. Moll, R.K. Thauer, NADP⁺ reduction with reduced ferredoxin and NADP⁺ reduction with NADH are coupled via an electron bifurcating enzyme complex in *Clostridium kluyveri*, *J. Bacteriol.* 192 (2010) 5115–5123.
- [48] E.J. Milner-White, M.J. Russell, Functional capabilities of the earliest peptides and the emergence of life, *Genes* 2 (2011) 671–688.
- [49] A.R. Crofts, J.T. Holland, D. Victoria, D.R.J. Kolling, S.A. Dikanov, R. Gilbreth, S. Lhee, R. Kuras, M.G. Kuras, The Q-cycle reviewed: how well does a monomeric mechanism of the bc₁ complex account for the function of a dimeric complex? *Biochim. Biophys. Acta* 1777 (2008) 1001–1019.
- [50] A.E. Marolewski, K.M. Mattia, M.S. Warren, S.J. Benkovic, Formyl phosphate: a proposed intermediate in the reaction catalyzed by *Escherichia coli* PurT GAR transformylase, *Biochemistry* 36 (1997) 6709–6716.
- [51] E.J. Beal, C.H. House, V.J. Orphan, Manganese- and iron-dependent marine methane oxidation, *Science* 325 (2009) 184–187.
- [52] R.E. Mielke, K.J. Robinson, L.M. White, S.E. McGlynn, K. McEachern, R. Bharti, I. Kanik, M.J. Russell, Iron-sulfide-bearing chimneys as potential catalytic energy traps at life's emergence, *Astrobiology* 11 (2011) 933–950.
- [53] S.E. McGlynn, I. Kanik, M.J. Russell, Modification of simulated hydrothermal iron sulfide chimneys by RNA and peptides, *Philos. Trans. R. Soc. Lond. A Phys. Sci.* 370 (2012) 1–16.
- [54] B.J. Skinner, R.C. Erd, F.S. Grimaldi, Greigite, the thio-spinel of iron; a new mineral, *Am. Mineral.* 49 (1964) 543–555.
- [55] D.J. Vaughan, J.R. Craig, The crystal chemistry of iron-nickel thiospinels, *Am. Mineral.* 70 (1985) 1036–1043.
- [56] H. Dobbek, V. Svetlitchnyi, L. Gremer, R. Huber, O. Meyer, Crystal structure of a carbon monoxide dehydro-genase reveals a [Ni–4Fe–5S] cluster, *Science* 293 (2001) 1281–1285.
- [57] A. Volbeda, J.C. Fontecilla-Camps, Catalytic nickel-iron-sulfur clusters: from minerals to enzymes, in: G. Simmonaux (Ed.), *Topics in Organometallic Chemistry*, vol. 17, Springer, Berlin, Germany, 2006, pp. 57–82.
- [58] C. Huber, G. Wächtershäuser, Activated acetic acid by carbon fixation on (Fe, Ni) S under primordial conditions, *Science* 276 (1997) 245–247.
- [59] D. Rickard, I.B. Butler, A. Olroyd, A novel iron sulphide switch and its implications for earth and planetary science, *Earth Planet. Sci. Lett.* 189 (2001) 85–91.
- [60] F. Baymann, E. Lebrun, M. Brugna, B. Schoepp-Cothenet, M.T. Giudici-Orticoni, W. Nitschke, The redox protein construction kit: pre-last universal common ancestor evolution of energy-conserving enzymes, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358 (2003) 267–274.
- [61] M.-E. Pandelia, W. Lubitz, W. Nitschke, Evolution and diversification of Group 1 [NiFe] hydrogenases. Is there a phylogenetic marker for O₂-tolerance? *Biochim. Biophys. Acta Bioenerg.* 1817 (2012) 1565–1575.
- [62] C. Furdul, S.W. Ragsdale, The role of pyruvate ferredoxin oxidoreductase in pyruvate synthesis during autotrophic growth by the Wood–Ljungdahl pathway, *J. Biol. Chem.* 275 (2000) 494–499.
- [63] C. Huber, G. Wächtershäuser, Primordial reductive amination revisited, *Tetrahedron Lett.* 44 (2003) 1695–1697.
- [64] E.J. Milner-White, M.J. Russell, Polyphosphate-peptide synergy and the organic takeover at the emergence of life, *J. Cosmol.* 10 (2010) 3217–3229.
- [65] J.F. Lambert, Adsorption and polymerization of amino acids on mineral surfaces: a review, *Orig. Life Evol. Biosph.* 38 (2008) 211–242.
- [66] S. Zhang, T. Holmes, C. Lockshin, A. Rich, Spontaneous assembly of a self-complementary oligopeptide to form a stable macroscopic membrane, *Proc. Natl. Acad. Sci. U. S. A.* 90 (1993) 3334–3338.
- [67] E.J. Milner-White, M.J. Russell, Sites for phosphates and iron-sulfur thiolates in the first membranes: 3 to 6 residue anion-binding motifs (nests), *Orig. Life Evol. Biosph.* 35 (2005) 19–27.
- [68] N.W.H. Ma, D.A. White, R.B. Martin, Metal ion exchange of square-planar nickel(II) tetraglycine with polydentate amines, *Inorg. Chem.* 6 (1967) 1632–1636.
- [69] R.B. Martin, M. Chamberlin, J.T. Edsal, The association of nickel(II) ion with peptides, *J. Am. Chem. Soc.* 82 (1960) 495–498.
- [70] A. Bianchi, C. Giorgi, P. Ruzza, C. Toniolo, E.J. Milner-White, A synthetic hexapeptide designed to resemble a proteinaceous P-loop nest is shown to bind inorganic phosphate, *Proteins Struct. Funct. Bioinform.* 80 (2012) 1418–1424.
- [71] A. Vallee, V. Humblot, C.-M. Pradier, Peptide interactions with metal and oxide surfaces, *Acc. Chem. Res.* 43 (2010) 1297–1306.
- [72] G. Fuchs, Alternative pathways of autotrophic CO₂ fixation, in: H.G. Schlegel, B. Bowen (Eds.), *Autotrophic Bacteria*, Science Technology, Madison, WI, 1989, pp. 365–382.
- [73] G. Fuchs, Alternative pathways of carbon dioxide fixation: insights into the early evolution of life, *Annu. Rev. Microbiol.* 65 (2011) 631–658.
- [74] H. Takami, H. Noguchi, Y. Takaki, I. Uchiyama, A. Toyoda, S. Nishi, G.-J. Chee, W. Arai, T. Nunoura, T. Itoh, M. Hattori, K. Takai, A deeply branching thermophilic bacterium with an ancient acetyl-coA pathway dominates a subsurface ecosystem, *PLoS One* 7 (1) (2012) e30559, <http://dx.doi.org/10.1371/journal.pone.0030559>.
- [75] D.J. Vaughan, Nickelian mackinawite from Vlakfontein, Transvaal, *Am. Mineral.* 54 (1969) 1190–1193.
- [76] J.W. Morse, T. Arakaki, Adsorption and coprecipitation of divalent metals with mackinawite (FeS), *Geochim. Cosmochim. Acta* 57 (1993) 3635–3640.
- [77] I.I. de Zwart, S.J. Meade, A.J. Pratt, Biomimetic phosphoryl transfer catalysed by iron(II)-mineral precipitates, *Geochim. Cosmochim. Acta* 68 (2004) 4093–4098.
- [78] C.G. Schipke, D.B. Goodin, D.E. McRee, C.D. Stout, Oxidized and reduced *Azotobacter vinelandii* Ferredoxin I at 1.4 Å resolution: conformational change of surface residues without significant change in the [3Fe–4S]^{1/0} cluster, *Biochemistry* 38 (1999) 8228–8239.
- [79] K. Chen, J. Hirst, R. Camba, C.A. Bonagura, C.D. Stout, B.K. Burgess, F.A. Armstrong, Atomically defined mechanism for proton transfer to a buried redox centre in a protein, *Nature* 405 (2000) 814–817.
- [80] M. Meuwly, M. Karplus, Theoretical investigations on *Azotobacter vinelandii* ferredoxin I: effects of electron transfer on protein dynamics, *Biophys. J.* 86 (2004) 1987–2007.
- [81] F. Bonomi, M.T. Werth, D.M. Kurtz Assembly, of Fe_nS_n (SR)^{2–} (n=2,4) in aqueous media from iron salts, thiols and sulfur, sulfide, thiosulfide plus rhodanase, *Inorg. Chem.* 24 (1985) 4331–4335.
- [82] J.D. Kim, A. Rodriguez-Granillo, D.A. Case, V. Nanda, P.G. Falkowski, Energetic selection of topology in ferredoxins, *PLoS Comput. Biol.* 8 (4) (2012) e1002463, <http://dx.doi.org/10.1371/journal.pcbi.1002463>.
- [83] H. Wang, G. Oster, Energy transduction in the F1 motor of ATP synthase, *Nature* 396 (1998) 279–282.
- [84] G.R. Helz, E. Bura-Nakić, N. Mikac, I. Ciglenečki, New model for molybdenum behavior in euxinic waters, *Chem. Geol.* 284 (2011) 323–332.
- [85] C.B. Mast, D. Braun, Thermal trap for DNA replication, *Phys. Rev. Lett.* 104 (2010) 188102.
- [86] A.-K. Kaster, J. Moll, K. Parey, R.K. Thauer, Coupling of ferredoxin and heterodisulfide reduction via electron bifurcation in hydrogenotrophic methanogenic archaea, *Proc. Natl. Acad. Sci. U. S. A.* 108 (2011) 2981–2986.
- [87] E. Koonin, W. Martin, On the origin of genomes and cells within inorganic compartments, *TiGs* 21 (2005) 647–654.
- [88] E. Koonin, An RNA-making reactor for the origin of life, *Proc. Natl. Acad. Sci.* 104 (2007) 9105–9106.
- [89] P. Mitchell, Proton-translocation phosphorylation in mitochondria, chloroplasts and bacteria: natural fuel cells and solar cells, *FASEB J.* (1967) 1370–1379.
- [90] P. Mitchell, The protonmotive Q cycle: a general formulation, *FEBS Lett.* 59 (1975) 137–139.
- [91] U. Brandt, Bifurcated ubihydroquinone oxidation in the cytochrome bcl complex by proton-gated charge transfer, *FEBS Lett.* 387 (1996) 1–6.
- [92] G. Herrmann, E. Jayamani, G. Mai, W. Buckel, Energy conservation via electron-transferring flavoprotein in anaerobic bacteria, *J. Bacteriol.* 190 (2008) 784–791.
- [93] H. Huang, S. Wang, J. Moll, R.K. Thauer, Electron bifurcation involved in the energy metabolism of the acetogenic bacterium *Moorella thermoacetica* growing on glucose or H₂ plus CO₂, *J. Bacteriol.* (2012), <http://dx.doi.org/10.1128/JB.00385-12>.
- [94] W. Nitschke, M.J. Russell, Redox bifurcations: mechanisms and importance to life now, and at its origin, *Bioessays* 34 (2012) 106–109.
- [95] A.G. Cairns-Smith, A.J. Hall, M.J. Russell, Mineral theories of the origin of life and an iron sulphide example, *Orig. Life Evol. Biosph.* 22 (1992) 161–180.
- [96] D. Coucouvanis, A. Toupadakis, S.-M. Koo, A. Hadjikyriacou, An inorganic functional group approach to the systematic synthesis and reactivity studies of binuclear Mo/S and Mo/S/O complexes, *Polyhedron* 8 (1989) 1705–1716.
- [97] R.A. Rothery, G.J. Workun, J.H. Weiner, The prokaryotic complex iron-sulfur molybdoenzyme family, *Biochim. Biophys. Acta* 1778 (2008) 1897–1929.
- [98] B. Schoepp-Cothenet, R. van Lis, P. Philippot, A. Magalan, M.J. Russell, W. Nitschke, The ineluctable requirement for the trans-iron elements molybdenum and/or tungsten in the origin of life, *Nat. Sci. Rep.* 2 (2012) 263.
- [99] M.A. Harmer, T.R. Halbert, W.-H. Pan, C.L. Coyle, S.A. Cohen, E.I. Stiefel, Ligand and induced internal redox processes in Mo- and W-S systems, *Polyhedron* 5 (1986) 341–347.
- [100] G.R. Helz, C.V. Miller, J.M. Charnock, J.F.W. Mosselmans, R.A.D. Patrick, C.D. Garner, D.J. Vaughan, Mechanism of molybdenum removal from the sea and its concentration in black shales. EXAFS evidence, *Geochim. Cosmochim. Acta* 60 (1996) 3631–3642.

- [101] B.E. Erickson, G.R. Helz Molybdenum, (VI) speciation in sulfidic waters: stability and lability of thiomolybdates, *Geochim. Cosmochim. Acta* 64 (2000) 1149–1158.
- [102] T.P. Vorlicek, M.D. Kahn, Y. Kasuya, G.R. Helz, Capture of molybdenum in pyrite-forming sediments; role of ligand-induced reduction by polysulfides, *Geochim. Cosmochim. Acta* 68 (2004) 547–556.
- [103] S. Ono, J.L. Eigenbrode, A.A. Pavlov, P. Kharecha, D. Rumble, J.F. Kasting, K.H. Freeman, New insights into Archean sulfur cycle from mass-independent sulfur isotope records from the Hamersley Basin, *Aust. Earth Planet. Sci. Lett.* 213 (2003) 15–30.
- [104] D. Coucouvanis, Synthesis, structures, and reactions of binary and tertiary thiomolybdate complexes containing the (O)Mo(Sx) and (S)Mo(Sx) functional groups (x = 1,2,4), *Adv. Inorg. Chem.* 45 (1998) 1–73.
- [105] S.W. Ragsdale, E. Pierce, Acetogenesis and the Wood–Ljungdahl pathway of CO₂ fixation, *Biochim. Biophys. Acta* 1784 (2008) 1873–1898.
- [106] A.Y. Mulikidjanian, M.Y. Galperin, K.S. Makarova, Y.I. Wolf, E.V. Koonin, Evolutionary primacy of sodium bioenergetics, *Biol. Direct* 3 (2008) 13.
- [107] A.Y. Mulikidjanian, P. Dibrov, M.Y. Galperin, The past and present of sodium energetics: may the sodium-motive force be with you, *Biochim. Biophys. Acta* 1777 (2008) 985–992.
- [108] N. Lane, J.F. Allen, W. Martin, How did LUCA make a living? Chemiosmosis and the origin of life, *Bioessays* 32 (2010) 271–280.
- [109] G. Proskurowski, M.D. Lilley, D.S. Kelley, E.J. Olson, Low temperature volatile production at the Lost City hydrothermal field, evidence from a hydrogen stable isotope geothermometer, *Chem. Geol.* 229 (2006) 331–343.
- [110] G. Proskurowski, M.D. Lilley, J.S. Seewald, G.L. Früh-Green, E.J. Olson, J.E. Lupton, S.P. Sylva, D.S. Kelley, Abiogenic hydrocarbon production at Lost City hydrothermal field, *Science* 319 (2008) 604–607.
- [111] W.J. Brazelton, M.P. Mehta, D.S. Kelley, J.A. Baross, Physiological differentiation within a single-species biofilm fueled by serpentinization, *MBio* 2 (2011) e00127–11.
- [112] D.S. Kelley, J.A. Karson, D.K. Blackman, G.L. Früh-Green, D.A. Butterfield, M.D. Lilley, E.J. Olson, M.O. Schrenk, K.K. Roe, G.T. Leboni, P. Rivizzigno, An off-axis hydrothermal vent field near the Mid-Atlantic Ridge at 30° N, *Nature* 412 (2001) 145–149.
- [113] G.K. Czamanske, B.F. Leonard, J.R. Clark, Erdite, a new hydrated sodium iron sulfide mineral, *Am. Mineral.* 65 (1980) 509–515.
- [114] C. Bethke, *Geochemical Reaction Modeling*, Oxford University Press, 1996.
- [115] S. Leduc, *The mechanism of life*, Rebman, London, UK, 1911.