How did LUCA make a living? Chemiosmosis in the origin of life

Nick Lane,1 John F. Allen,2 and William Martin3*

1Department of Genetics, Evolution and Environment, University College London, London, UK
2School of Biological and Chemical Sciences, Queen Mary University of London, London, UK
3Institut für Botanik III Heinrich-Heine-Universität, Düsseldorf, Germany

Despite thermodynamic, bioenergetic and phylogenetic failings, the 81-year-old concept of primordial soup remains central to mainstream thinking on the origin of life. But soup is homogeneous in pH and redox potential, and so has no capacity for energy coupling by chemiosmosis. Thermodynamic constraints make chemiosmosis strictly necessary for carbon and energy metabolism in all free-living chemotrophs, and presumably the first free-living cells too. Proton gradients form naturally at alkaline hydrothermal vents and are viewed as central to the origin of life. Here we consider how the earliest cells might have harnessed a geochemically created proton-motive force and then learned to make their own, a transition that was necessary for their escape from the vents. Synthesis of ATP by chemiosmosis today involves generation of an ion gradient by means of vectorial electron transfer from a donor to an acceptor. We argue that the first donor was hydrogen and the first acceptor CO2.

Keywords: alkaline hydrothermal vents; ATPase; chemiosmosis; LUCA; proton gradients

Introduction: primordial soup at 81, well past its sell-by date

In 1929, J. B. S. Haldane published an influential essay on the origin of life.(1) In it, he set forth the concept of a prebiotic broth, or primordial soup. Haldane proposed that UV radiation provided the energy to convert methane, ammonia and water into the first organic compounds in the oceans of the early Earth. In the absence of life forms to consume them, these molecules accumulated to the concentration of a ‘hot dilute soup’, in which they reacted further to produce larger molecules, then macromolecules, later virus-sized particles and finally the first fully fledged cells. These, Haldane surmised, were heterotrophic fermenters that consumed the remaining organic soup, stalling evolution until the origin of photosynthesis, at which point life was off and rolling towards present forms. Backed up by Stanley Miller’s (1953) inorganic synthesis of organic molecules in the laboratory,(2) it seemed to generations of scientists that Haldane’s narrative was basically right, and all that was left was to sort out the details.(3)

The concept of organic soup is nowadays closely allied to the idea that the origin of life and the origin of replication are the same thing. Natural selection remains the only mechanism known in which more complex forms can evolve, and natural selection requires a replicator. Regarding the nature of that replicator, there is currently no viable alternative to the idea that some kind of ‘RNA world’ existed, that is, there was a time before proteins and DNA, when RNA was the molecular basis of both catalysis and replication. Some elements of the RNA world concept are almost certainly correct. However, there is a strong version of this theory which states that RNA was once the only catalyst as well as the only replicator and so all the basic chemistry of life was invented by RNA.(4) This ‘RNA first’ theory is difficult to accept from the standpoint of the biochemistry of modern cells. For example, many essential enzymes are metalloproteins that contain mineral centres, such as iron-sulphur clusters, at their heart.(5) There is every reason to believe that such clusters, with the structures of inorganic minerals like greigite,(6,7) have more ancient roots even than those of RNA.

The recent abiotic synthesis of nucleotides using UV radiation and phosphate to purify intermediates(8) seems to lend support to the idea that the primordial oceans became a warm broth filled with nucleotides, which spontaneously polymerised into RNA, able to catalyse its own replication as well as organic transformations that ultimately yielded cells with lipid membranes, proteins and DNA – a purely Haldanian distillation, eight decades after Haldane’s essay. But the fact that nucleotides can be synthesised in ‘warm pond’ conditions hardly makes an oceanic RNA world more likely; by the same measure, the circumstance that amino acids can be present in some meteorites does not mean that life must have arisen in outer space. Setting aside the absence of geochemical evidence that a primordial soup ever existed, there are grave difficulties with the soup theory. To give a single example,
polymerisation into RNA requires both energy and high concentrations of ribonucleotides. There is no obvious source of energy in a primordial soup. Ionizing UV radiation inherently destroys as much as it creates. If UV was the primordial source of energy, why does no life today synthesise ATP from UV radiation? Worse, every time an RNA molecule replicates itself, the nucleotide concentration falls, unless nucleotides are replenished at an equal rate. UV radiation is an unlikely source for rapid polymerisation and replication, and an unpromising initiator of natural selection.

The reason that nucleotides and other organic molecules are reluctant to react further in a soup is that they are at thermodynamic equilibrium. They have already reacted, and the homogeneous soup has no internal free energy that would allow them to react further. Life is not just about replication; it is also a coupling of chemical reactions – exergonic ones that release energy and endergonic ones that utilise it, preventing the dissipation of energy as heat. It is commonplace to say that life requires energy, but the conception of a primordial soup fails to recognise or incorporate the importance of energy flux. On the congruence principle, what life needed was not some harsh and problematic source of energy like UV radiation (or lightning), but a continuous and replenishing source of chemical energy. In all forms of life, hence in the primordial ancestor too, energy flux means making and breaking ‘high-energy’ chemical bonds such as the thioester bond in acetyl CoA or the anhydride bonds in acetyl phosphate and ATP. Release of the energy stored in these bonds fosters the synthesis of organic molecules, activates nucleotides, amino acids, and so on, prompting them to react and polymerise, pushing a dynamic chemical system further and further from futile equilibrium with its surroundings.

**Fermentation is not ‘life without oxygen’**

Following in the footsteps of Pasteur, Haldane argued that fermentation was the primordial mechanism of energy generation, and that – in Pasteur’s words – fermentation is ‘life without oxygen’. Since virtually all geochemists agree that there was little if any free oxygen on the primordial Earth, fermentation in that sense is still widely seen as the most likely primordial source of energy. De Duve, for example, argues that early life was anaerobic and so most likely to have depended on the kind of fermentations that sustain anaerobic life today.[9,10] If there can be said to be a textbook view, this is it.

But there are profound difficulties – both chemical and biological – in viewing fermentation as primitive rather than derived. Fermentation is chemically a disproportionation – not a simple redox reaction, in which electrons are stripped from a donor and passed onto an acceptor, driven by strong thermodynamics. In contrast with respiration, the amount of energy released by fermentation is tiny, reflecting its lack of thermodynamic driving force. To tap such an insignificant source of energy requires more rather than less sophistication, and indeed about 12 enzymes are needed to catalyse a complex succession of steps in glycolytic-type fermentations based around the Embden-Meyerhoff pathway. These enzymes are proteins encoded by genes, which would have had to evolve as a functional unit without any other source of energy in the primordial oceans – close to an impossibility in an RNA world, let alone the only way to evolve one.

The implication that fermentation is a sophisticated derivation, rather than primordial, garners support from phylogenetics. To the extent that we can trust any evolutionary tree that purports to reflect events that occurred almost 4 billion years ago, no tree of life has ever placed ‘pure’ fermenters on the early branches; these branches invariably comprise prokaryotes that gain their energy from chemiosmotic coupling. Today, all known autotrophs generate energy using redox reactions at their plasma membrane, while the majority of heterotrophs, too, depend on oxidative phosphorylation – a process in which oxygen is just one of many possible terminal electron acceptors, like CO2 or ferric iron. Furthermore, essentially all fermenters retain some additional machinery of chemiosmosis, notably the proton-motive ATPases that they use to generate a proton gradient across the plasma membrane to import and export solutes. While fermentation does not itself require such membrane bioenergetics, active transport across the cell membrane does, and is necessary for homeostasis and uptake of nutrients, as well as flagellar motility.[11]

Perhaps most strikingly of all, bacteria and archaea differ markedly in the gene sequences and crystal structures of enzymes that catalyse the individual steps of fermentation.[12] If we follow the straightforward reasoning that traits common to both bacteria and archaea are inherited from a common ancestor, whereas traits that differ substantially probably evolved independently at a later stage, we can draw simple conclusions about ‘LUCA’ – the Last Universal Common Ancestor of cells. Thus RNA, DNA, the universal genetic code, transcription, translation, ribosomes, a rotor-stator-type ATPase, ATP and the Krebs cycle are inherited from LUCA, while traits like oxygenic photosynthesis are not, and evolved later, with the cyanobacteria. In this context, and given the profound differences between enzymes catalysing fermentation in archaea and bacteria, it is overwhelmingly likely that fermentation evolved at least twice, independently, in the archaea and bacteria. LUCA could not perform fermentations of the glucose disproportionation type – modern glycolysis – that de Duve and others envisioned. Fermentations evolved later, once autotrophs had produced organic compounds for fermenters to eat and to breathe. But if there was no soup, and no energy from UV radiation or fermentation, then where was the energy that powered the emergence of life?
Alkaline hydrothermal vents as the primordial source of energy for life

The discovery of deep submarine hydrothermal vents – 'black smokers' – in the late 1970s seemed to offer an answer. In contrast to organic soup, black smokers are far from equilibrium with surrounding seawater. They are volcanic, sitting directly on top of oceanic spreading zones, and fashioned by the interaction of seawater with magma at ~1,200 °C. Their effluent is hot (~350 °C), acidic (~pH 1–2), rich in hydrogen sulphide and dissolved metals, but very low in hydrogen gas. As pointed out by Miller, Bada and others, several issues relating to black smokers as sites of origin of life are problematic, among them their extreme temperature (more likely to break down organics than form them), their low pH, their short lifetimes (in the order of decades), and their lack of compartmentalisation, with its dismal consequence of irretrievable dilution into the ocean.

But black smokers are not the only type of hydrothermal system and the distinction is significant. A second type of vent, discovered in 2000, is formed by the reaction of seawater with minerals like olivine, which comprise much of the oceanic crust. These hydrothermal systems are not volcanic, but the result of a geochemical process known as serpentinisation, in which olivine is hydroxylated to serpentine. The hydroxylated rocks expand and fracture, allowing the entry of even more seawater into the crust, perpetuating the reaction. Remarkably, more water is bound into the Earth's crust as hydroxylated rock than is present in the oceans, forming clusters of minerals found in hydrothermal deposits, as well as aragonite, and metal sulphides rich in hydrogen. Thermal expansion drives the hydrothermal fluids back through the crust up to the sea floor, where they emerge as a radically different kind of vent – one that is warm (ca. 70 °C), alkaline and full of dissolved hydrogen. The importance of alkaline hydrothermal systems, or simply alkaline vents, in the emergence of life was proposed by Russell et al. in 1993, nearly a decade before their first discovery at 'Lost City', on the ocean floor just off the mid-Atlantic ridge. Lost City fulfils many of Russell's requirements for incubators of proto-metabolism – not only in the alkalinity of the fluids emerging, but also in its delicate porous structure. These systems support towers up to 60 m tall, formed from calcium carbonate and riddled with tiny interconnected pores on the micrometer scale, with feathery aragonite walls through which hydrothermal fluids circulate via thermal currents. The hydrothermal fluids are rich in hydrogen, with a smattering of light hydrocarbons, including methane, formate and acetate, some of which appears to be formed abiotically.

Regardless of whether it is produced abiotically or by methanogens deeper within the crust, the methane in alkaline vents gives a clue to the origin of life: life began as a 'side effect' of the direct hydrogenation of carbon dioxide, to form methane or acetate. All autotrophs today fix carbon dioxide using hydrogen, either directly or indirectly (from water or other electron donors like H₂S), yet there are only five known primary pathways of carbon dioxide fixation. All except one consume energy (ATP) to fix carbon. The exception is the acetyl CoA (Wood-Ljungdahl) pathway of direct reaction of hydrogen with carbon dioxide. The acetyl CoA pathway is found in some of the most ancient prokaryotes, including both methanogens (archaea) and acetogens (bacteria), and so is likely to have been present in LUCA as well. This ancient pathway simultaneously fixes carbon while generating energy as ATP – 'a free lunch that you're paid to eat' in Everest Shock's memorable words. In both methanogens and acetogens, the steps of carbon reduction are catalysed by metalloenzymes with mineral clusters containing iron, sulphur and nickel at their centres. The crystal structures of these clusters are essentially those of minerals found in hydrothermal vents, such as greigite, to the point that one could refer to the acetyl CoA pathway as having 'rocky roots'.

During the early phases of Earth's history, >3.8 Gya ago, when life began and CO₂ concentrations in the oceans were 1,000-fold higher than they are today, alkaline vents were the site of a redox interface between H₂-rich hydrothermal and CO₂-rich marine aqueous phases. The absence of oxygen meant that the oceans were full of dissolved ferrous iron, so the mineral cells in alkaline vents could have been bounded by bubbly inorganic membranes rather than, or in addition to, aragonite. Examples of metal sulphide vesicular mineral deposits are found today in the 360-Mya vents at Tynagh in Ireland. Such transition metal sulphides are rich in catalytic capabilities, and would have given the natural inorganic compartments catalytic walls. Alkaline vents would have been present on a massive scale across the sea floor of the early Earth, and were orders of magnitude longer lived than black smokers. Finally, the abiotic chemistry envisaged for alkaline vents, on both thermodynamic and experimental grounds, provides an indication of how life began, and why chemiosmosis has always been the primary process of energy conversion.

The origin of life in alkaline vents

The energy that runs the motor of life in methanogenic and acetogenic metabolism comes from electron transfer from H₂ to CO₂. The formation of either methane or acetate from H₂ and CO₂ releases energy, and it is this energy that methanogens and acetogens harness to fuel all their biosynthetic pathways. A portion of the energy released in
the acetyl CoA pathway is captured in the form of the high-energy thioester bond of acetyl CoA. Because of the simplicity of the C3 compound chemistry involved, Fuchs suggested that the acetyl CoA pathway is as ancient as it gets in biochemistry, and this view remains current among microbiologists. It seems plausible, then, that should hardly be surprising, because the synthesis of phosphate to generate acetyl phosphate, a slightly more analogous to acetyl CoA itself, can in turn react with inorganic metabolism. Under primordial conditions, acetyl thioesters, thioester bonds are easily converted into the phosphate bonds up to C18 have also been detected in Atlantic vent systems, although it is difficult to prove an abiotic origin. Further, ‘high-energy’ thioester bonds are easily converted into the phosphate bonds of ATP in modern biochemistry, powering intermediary metabolism. Under primordial conditions, acetyl thioesters, analogous to acetyl CoA itself, can in turn react with inorganic phosphate to generate acetyl phosphate, a slightly more energy-rich functional analogue of ATP that is still used in some bacteria alongside ATP today. In terms of primordial sources of energy, then, the common currency of ATP is likely to have been preceded in the alkaline vents by the simpler interchangeable currency of acetyl thioesters and acetyl phosphate.

Thus the acetyl CoA pathway produces both simple organics like pyruvate as well as acetyl phosphate, to drive the reductive citric acid cycle and intermediary metabolism more widely. These pathways are beyond the scope of this paper, but are dealt with in detail elsewhere. Suffice to say that the process of serpentinisation generates other reduced compounds beyond hydrogen itself, and that the reducing power of such hydrothermal systems is sufficient to fix N2 as ammonia, so enabling the synthesis of early intermediates in nitrogen metabolism needed for amino acids and nucleotides.

The synthesis of amino acids under vent conditions has been shown, as has the synthesis of long-chain hydrophobic hydrocarbons (possible precursors of biological membranes) via Fischer-Tropsch reactions. Hydrocarbons up to C18 have also been detected in Atlantic vent systems, although it is difficult to prove an abiotic origin. However, no one has yet demonstrated the abiotic synthesis of nucleotide bases under hydrothermal vent conditions, even though FeS will efficiently catalyse the synthesis of purines and pyrimidines from formamide in the laboratory, which should hardly be surprising, because the synthesis of ribonucleotide bases is thermodynamically favourable under hydrothermal conditions. It seems plausible, then, that nucleotides could have been formed under abiotic conditions in alkaline vents, given the continuous source of hydrogen and carbon dioxide, the catalytic properties of transition metal sulphides, and the ability of the inorganic compartments to concentrate intermediates.

Morowitz and coworkers have pointed out that inorganic catalysts can have a profound effect on the outcome of reactions, not by speeding up the reaction enormously, but rather by favouring certain reactions over others, giving rise to proto-metabolic pathways, while minimising contamination with other products. In this view, reactions under kinetic control (the fastest reaction wins) and thermodynamic control (the most stable product wins) determine the nature of the first biochemical pathways, to no small extent via the nature of the catalysts involved – and these pathways furnished the products needed for selection to begin operating in an RNA world. This is an important concept because before there is genetic material, there can be no natural selection. But if thermodynamics channels chemistry towards genetic material under the right kinds of conditions – a vent for example with a replenishing source of chemical energy – then chemistry’s somewhat deterministic version of natural selection (thermodynamics) would interface well indeed with natural selection as biologists know it.

Polymerisation and replication of primitive nucleotides in an RNA world are much more easily envisaged in vents than as a result of the action of UV in soup. Braun and coworkers have shown in simulation that convection currents and thermal diffusion tend to concentrate nucleotides in the cooler regions of the vent system up to millions of times the starting concentration, making polymerisation into RNA or DNA far more likely, especially in the presence of a continuous source of phosphoanhydride bonds. Nucleic acids up to 100 bases in length show even greater propensity to become concentrated, in theory giving up to a trillion times the starting concentration. The experimental effect in the laboratory is orders of magnitude, but not trillions. The alternate warming and cooling produced by thermal circulation also tend to melt and anneal nucleic acids, simulating their amplification by convective PCR. In other words, RNA formation and replication are positivelyfavoured in alkaline vents, whose interiors are close to an ideal RNA world.

The high partial pressures of carbon dioxide gas (which dissolves to give carbonic acid) would have made an equally profound difference: an acidic ocean. Acidic oceans made the alkaline vents inherently chemiosmotic, which is to say there was a natural proton gradient across the inorganic membranes, probably producing a proton-motive force of about 200 mV positive outside, similar in magnitude to that across biological membranes today, and constantly replenished by alkaline fluids emanating from the vents. Notably, the proton gradient had a polarity identical to that in cells today (Fig. 1). The question is, was this natural proton gradient
purely incidental, or did it have a real bearing on how life emerged from the vents? The answer may be that life could not have escaped any other way, explaining just why chemiosmosis is universal in cells today.

Chemiosmosis is fundamental and universal

There is no more cautionary tale in all of biology than that of Peter Mitchell and the discovery of chemiosmosis (derived from Greek ‘to push’). Mitchell’s theory was described by Orgel(48) as the most counter-intuitive idea in biology since Darwin – and the only one to bear comparison with the ideas of Einstein, Heisenberg and Schrödinger. Such an assertion is all the more surprising coming from Orgel, who was a leading advocate of the RNA world, so signally lacking in the spatial organisation that Mitchell himself prized. Yet despite such accolades, the deeper evolutionary significance of chemiosmosis is rarely explored.

Mitchell(49) proposed the radical idea that the energy derived from oxidation of NADH is used to generate a gradient of proton concentration across a membrane, which in turn drives the synthesis of ATP via proteins initially termed ‘coupling factors’ and today known as ATP synthases (or ATPases). The real significance of this idea is that it overcomes the limitations of scalar chemical reactions, relegating the substrate-level phosphorylations (SLP) characteristic of fermentations to a quantitatively insignificant role in the overall scheme of biological energy conservation. In place of molecules reacting directly with each other, as they do in SLP according to the exact stoichiometric laws of chemistry, Mitchell substituted a proton gradient across a membrane. The idea was discounted as absurd by a generation of biochemists who had diligently searched, to no avail, for the covalent, high-energy chemical intermediate (X/C24P) predicted to link oxidation-reduction with phosphate ester formation. Resistance to the idea of proton gradients, perhaps along with Mitchell’s own rather argumentative personality, fuelled more than 20 years of acrimony, the ‘ox-phos wars’. Since then, virtually all work on chemiosmosis has concerned precise mechanistic details, rather than the broader picture that originally drove Mitchell – how and why does life structure chemistry, vectorially, in three dimensions? While Mitchell rebelled against the absence of spatial dimensions in the language of chemistry,(50) the informational vocabulary of molecular biology remains one-dimensional to this day, and has circumscribed spatial thinking on the origin of life.

Apart from being counter-intuitive, the most remarkable feature of chemiosmosis is how universal, how fundamental,
Proton gradients are across all of life. Mitchell appreciated that proton gradients frequently power active transport in bacteria, along with the rotating motor of the bacterial flagellum; and they are as central to ATP synthesis in photosynthesis as they are to oxidative phosphorylation in mitochondria or bacteria.\(^{(51)}\) The fact that a proton-motive force powers chloroplasts, mitochondria and bacteria is already a startlingly diverse heritage; but we now know that proton gradients are ubiquitous in archaea as well as bacteria. Both archaea and bacteria generate proton gradients across their plasma membranes; both synthesise ATP via oxidative phosphorylation; both share similar membrane-bound rotor-stator-type ATP synthase enzymes, albeit with a few minor differences (the bacterial F-type and archaeal A-type ATPases share a number of unquestionably homologous subunits, as well as others apparently unrelated).\(^{(52)}\) Finally, both archaea and bacteria conduct electrons and pump protons using ferredoxins, quinones and cytochromes that are closely related in chemical structure.\(^{(53)}\) Thus the deep and broad phylogenetic distribution of chemiosmosis is precisely the opposite of that of fermentation. Just as we can say that LUCA had a ribosome, which will not startle anyone, we can say that LUCA obtained energy by chemiosmotic coupling. That may startle some.

Despite this centrality, few theories on the origin of life take chemiosmosis explicitly into consideration; it is usually treated as little more than an addendum, incorporated at an arbitrarily late point, or viewed as so advanced relative to Haldanian fermentative origins as to be irrelevant to the origin of life. Wächtershäuser, for example, views the first cells as powered by surface chemistry on pyrite minerals;\(^{(54)}\) when he adds in chemiosmosis,\(^{(34)}\) it is with the reverse polarity to that existing in cells today, with the inside acidic and the outside alkaline. De Duve sees proton gradients as a possible adaptation to acidic conditions, at least providing an explanation as to why cells should pump protons out.\(^{(9,10)}\) But why a proton-motive force should displace virtually all other mechanisms of energy generation (including his own favoured direct coupling of redox reactions to ATP synthesis via thioesters) in organisms with no need to adapt to acidic conditions is not addressed.

More recently, it has been argued that sodium gradients were the primary driving force, resulting from growing ocean salinity over a 500-million-year period.\(^{(55)}\) It is unclear why all living things depend on chemiosmotic mechanisms if the driving force developed not over the seconds and minutes needed for bacterial bioenergetics, but over the millions of years required for ocean salination. Again, why the sodium-motive force should be so comprehensively replaced by a proton-motive force is not addressed. None of these ideas come close to explaining why the proton-motive force, specifically, should be so universal across life.

So what is the alternative? The alternative, as proposed by Russell et al. in 1993,\(^{(10)}\) is that chemiosmosis is an inherent property of life, one inherited from the very place in time and space where it arose.\(^{(19,20,47)}\) This answer is also hinted at by the broad properties of LUCA. As a short aside, we take the properties of LUCA here to be those shared by the last common ancestor of bacteria and archaea, to the exclusion of eukaryotes, on the grounds that all known eukaryotes either possess or possessed mitochondria.\(^{(56,57)}\) Known eukaryotes, the only ones we need to worry about, therefore arose after the eubacteria had already diversified, hence later than the eubacteria as a group. Likewise, most recent genomic studies reveal that the main archaeabacterial genes of eukaryotes (mostly informational genes) arose from within the archaea, not as sisters to them.\(^{(58,59)}\) LUCA is thus the last common ancestor of eubacteria and archaeabacteria, and her properties are those ancestrally shared by these domains.

We have already noted that archaea and bacteria share many fundamental properties, including the ability to generate ATP by way of chemiosmotic coupling, which requires a membrane. But here there is a curious paradox. LUCA was chemiosmotic, and so required a membrane — but there is little in common between archaeal and bacterial plasma membranes or cell walls. Archaeal membranes are composed of isoprenoid side chains, joined by ether bonds to \(\alpha\)-glycerol; bacteria use fatty acids, joined by ester bonds to \(\beta\)-glycerol. While there are points in common, such as the use of glycerol (albeit different enantiomers), the differences are stark, and lipid synthesis is catalysed by distinct enzymes.\(^{(25)}\) Likewise the bacterial cell wall is composed of peptidoglycans, usually missing altogether in archaea; or if not, archaeal cell walls lack the typical \(\alpha\)-amino acids and \(N\)-acetylumuramic acid of bacteria. In short, the defining boundaries that separate the cytoplasms of bacteria and archaea from their environment are chemically unrelated, and their modern forms presumably arose later, independently. One alternative, namely that a free living cell with a perfectly good plasma membrane would go to the trouble of inventing a chemically new plasma membrane via point mutation and/or \textit{de novo} origin of genes,\(^{(60)}\) seems arbitrary by comparison. Another alternative\(^{(54)}\) that LUCA had both kinds of membranes, which segregated as a result of chemical incompatibilities, overlooks the obvious circumstance that having chemically different kinds of lipids (cholesterol, sterols and lipid rafts as tangible examples) in one’s membrane is normal biology,\(^{(61)}\) not a chemical crisis situation of domain-splitting dimensions.

So LUCA was chemiosmotic, requiring a membrane, but apparently did not have a membrane comparable to that in either modern archaea or bacteria. While this might look like a paradox, it is not. The bubbly mineral cells that riddle alkaline vents have their own inorganic walls, which we envisage were lined in some regions by hydrothermally synthesised hydrophobic substances — lipids — that were eventually replaced by enzymatically derived lipid membranes, independently, in the archaeal and eubacterial stem lineages. We posit that the
primordial lipids coating the inorganic cells were synthesised by Fischer-Tropsch-type reactions in hydrothermal systems. At Lost City, such reactions may give rise to the propane, butane and pentane that can be detected today. From the standpoint of chemical thermodynamics, Shock et al. have calculated that under conditions corresponding to Lost City, dodecanoate (C12) would be the preferred organic acid. From the standpoint of biogenetics, the question is, could these inorganically walled cells lined with hydrophobics really have taken advantage of the geochemically existing proton gradient, and later generated their own?

The primordial proton-motive force

One of the biggest stumbling blocks to the idea of ‘chemiosmosis early’ is the daunting complexity of the ATP synthase – a nanomachine comprising a rotary motor powered by a flow of protons through the membrane stalk, coupled to the rotating head that forms ATP from ADP and phosphate. The two main domains, the stalk and the rotating head, have no obvious homologues among other proteins. ATP synthase is certainly the product of Darwinian evolution at the level of genes and proteins. That does not mean that the ATP synthase did not evolve in the vents; even if LUCA was restricted to the vents, she was still a sophisticated organism with DNA, proteins, ribosomes, the genetic code, energy in the form of acetyl phosphate, thioesters, ATP and so forth. The fact that both archaea and bacteria possess similar ATP synthase enzymes (A- and F-type ATPases, respectively) implies that an ATPase ancestral to both types did indeed evolve in the vents.

Although it has been argued that the earliest ATP synthase enzyme was coupled to sodium gradients, not protons, this seems unlikely. Trees claimed to support an ancestral position of sodium-dependent ATPases do not in fact favour either sodium- or proton-first topology; the sequences are interleaved. The amino acid substitutions common to the membrane rotor subunits predict that Na⁺-dependent ATPases arose in parallel in independent lineages. Critically, if the ancestral state of ATPases was to pump cations out of the cell (at the expense of ATP), where did all the ATP come from? Certainly not from fermentation or UV radiation, for reasons discussed earlier. Our interpretation is that proton dependence is the ancestral state of ATPases; that the ancestral function of ATPases is to generate ATP rather than consume it; and that the first ATPases harnessed a natural proton gradient in alkaline vents, the magnitude (roughly 1000-fold concentration difference) and polarity of that gradient (alkaline inside) being virtually identical to that in modern cells. Later adaptations to hyperthermophilic and hypersaline environments led to the replacement of protons with sodium, as membranes are leakier to protons than sodium under these conditions. Sodium ATPases are then derived in several lineages independently, involving a few parallel amino acid substitutions to proton ATPases under the guiding hand of selection. Similar parallel changes occur in mitochondrial DNA under selective constraints.

We therefore argue that the ancestral ATPase arose in alkaline vents, where it harnessed the natural proton gradients to generate ATP, just as it does today. It is worth noting here that we do not envisage the ancestral ATPase as embedded in the inorganic walls, but rather in organic lipids lining the walls. We picture the labyrinth of interconnected cells as equivalent to a chromatography column, open at the bottom and top, through which alkaline fluids percolate. The ‘sides’, however, are coated internally with OR are lined with an organic hydrophobic layer – an abiotically accumulated ‘membrane’ formed by Fischer-Tropsch reactions, as discussed earlier. An ancestral ATPase embedded in such organic membranes could tap proton gradients across the side walls and hydrothermal effluent flow would maintain the gradient through replenishing alkalinity inside.

In such an arrangement, an ATPase, hence any protocells encoding it, could plainly tap an existing proton gradient; but do we need to posit that a vent-bound LUCA had to come up with a machine as complex as the ATPase in one leap? At present, there is nothing to suggest that any simpler intermediates did actually precede the ATPase, and no need for them to do so. Acetyl thioesters, generated spontaneously in vent conditions, could in theory supply sufficient acetyl phosphate, and later ATP for all stages of evolution up to LUCA by SLP alone. In this world of genes and proteins, supplied by plentiful energy, there is no reason why a rotor-stator type ATPase could not be ‘invented’ de novo; in terms of complexity it requires no more evolutionary innovation than the origin of a primordial ribosome, which everyone would agree did evolve. We are simply explicit on the issue of where, why and in the context of which proton gradient, the first ATPase arose.

Nonetheless, for those who feel uncomfortable without a more gradualistic series of plausible intermediates, we note that other, simpler, if not necessarily older, mechanisms could have tied primordial proton gradients into cell physiology at an early stage. For example, pyrophosphate forms spontaneously from phosphate under a proton gradient in hydrophobic membranes and is synthesised today via simple single-subunit proteins like the H⁺-translocating pyrophosphatase. A strong proton gradient can drive pyrophosphate synthesis far from equilibrium, enabling it to directly phosphorylate acetate or ADP. Alternatively, the reverse formate hydrogenase reaction may have boosted the acetyl CoA pathway. Formate hydrogenase is a metalloprotein with mineral centres, which normally generates a proton-motive force from the oxidation of formate to carbon dioxide.
In principle a natural proton gradient could help to drive carbon dioxide reduction to formate via coupling through a protein.\(^{68}\)

Thus there are a number of plausible ways in which a vent-bound LUCA could have taken advantage of a natural proton-motive force, of which the ATPase is undoubtedly the most effective. Any of these would have given protocols an energetic advantage over protocols that did not take advantage of the natural proton gradient, and so would be selected in straightforward Darwinian terms. However, tapping the natural proton gradient offered more than a proliferative advantage; it was also the only way in which cells could escape the vents.

**Why chemiosmosis was necessary**

On the basis of bioenergetics, life could not have left the vents unless it took advantage of the proton gradients on offer. Specifically, the energetics of the acetyl CoA pathway meant that chemiosmosis was strictly *necessary* to leave the vents.

The reason relates to the energetic barrier that must be overcome to go from carbon dioxide to a methyl group by way of formate, a curious bioenergetic puzzle called ‘the early formyl pterin problem’.\(^{30}\) In the vents, geochemically derived reactive intermediates like methyl sulphide could overcome this ‘energy hump’, without cells needing to expend energy in the form of acetyl CoA or acetyl phosphate. Methyl sulphide is predicted (thermodynamically) to form spontaneously in vents,\(^{69}\) and is thought to be central to microbial consortia in vent systems, apparently playing a central role in connecting methanogenesis to anaerobic methane oxidation.\(^{70}\) What’s more, methyl sulphide is a known substrate for methanogenesis by *Methanosarcina acetivorans* and other methanogens.\(^{71–74}\) In bioenergetic terms this is a critically important point, because it means that in the presence of geochemically supplied methyl sulphide, acetogenesis can proceed by SLP alone in alkaline vents, in principle right through to LUCA.

Move out of the vents, however, and there is a problem. Without a free supply of methyl sulphide, cells need to invest ATP or some other form of energy to overcome the energy hump between carbon dioxide and formate. Unfortunately, when growing by SLP alone, cells must invest one ATP to gain one ATP, making growth by SLP impossible. The problem is pure chemistry. Molecules react stoichiometrically. If a reaction releases enough energy to phosphorylate two molecules of ADP, then cells can grow. As it happens, however, SLP releases enough energy to synthesise only about 1.5 molecules of ATP, a chemical impossibility, so only one molecule of ATP can be formed. Thus the ‘quantum’ nature of SLP, and scalar chemical reactions in general, precludes ‘saving up’ towards a second ATP; the additional energy available from the reaction is necessarily lost in SLP, making growth impossible.\(^{50,75}\)

Chemiosmosis overcomes this difficulty by ‘uncoupling’ stoichiometric bonding reactions from ATP synthesis – rather than there being a single reaction, Mitchell’s delocalised proton gradient is an obligatory intermediate. In principle a reaction can be repeated and repeated just to pump a proton over a membrane. It is limited only by the number of possible conformational changes in a protein.\(^{76}\) Eventually the proton gradient becomes sufficient to permit the synthesis of one molecule of ATP. The net effect is that chemiosmosis allows cells to ‘save up’ towards a second ATP. In the case of the acetyl CoA pathway (and even more so in the other four pathways of carbon fixation, which require ATP to operate), chemiosmosis makes the difference between ‘growth’ and ‘no growth’ – between freedom, or a life spent tied to the vents by a thermodynamic umbilical cord. It is no accident that both methanogens and acetogens rely on chemiosmosis to grow. They could not have left the vents unless they invented chemiosmosis first. Indeed, chemiosmosis overcomes the quantum limitation of SLP so effectively that it is still universal across life – the proton-motive force has not been improved in a meaningful way in nearly 4 billion years of evolution.

Of course chemiosmosis in modern free-living cells requires more than simply tapping a ready-made gradient. Cells must have learned at some point to generate their own proton gradient. Why, and how, would cells learn to do that in a vent world already supplied with natural proton gradients? The answer to the first question is quite simple. Proton gradients vary across vents, being weaker further from the flow of hydrothermal fluids. Protocells off the main axis of the vent would gain from regenerating proton gradients, especially if their metabolism had become dependent on the proton-motive force.

The answer to the second question – how? – is not as complex as de Duve and others imply.\(^{9,10}\) While it is true that oxidative phosphorylation in modern organisms requires complex electron transfer chains, usually incorporating iron-sulphur proteins, quinones and cytochromes, far simpler versions are possible.\(^{77}\) For example, some methanogens and acetogens lack cytochromes and quinones or their functional analogues, depending instead on relatively simple conformational changes in proteins to generate ion gradients.\(^{58,75}\) Quinones are undoubtedly ancient, and Mitchell’s Q cycle does not even require conformational changes in proteins, merely shuffling quinones or simpler analogues across a polarised membrane.\(^{78}\) There are even abiotic ways of generating proton gradients by separating the sites of oxidation and reduction to opposite sides of a membrane: in alkaline vents, for example, the oxidation of iron sulphide to pyrites, coupled to the reduction of carbon dioxide in inorganic cells, could generate a proton gradient with the correct polarity.\(^{79}\)
Conclusions

We have not attempted in this essay to plot out exactly what happened at the origin of life, but rather to focus attention on the importance of chemiosmosis as an early bioenergetic process, specifically in the setting of alkaline hydrothermal vents. Not only do such considerations help explain the profound importance of proton gradients to life today, but they also give an indication of how and where such gradients might first have formed, why they are thermodynamically necessary for the evolution of free-living cellular life, and how the first cells might have harnessed, and later generated, their own gradients of proton concentration.

There are many ways to experimentally test the kinds of ideas set forth here. One approach involves chemistry in simulated vent conditions (reviewed in Ref. [40]). With some exceptions, [35] most of the published chemical work done so far has involved closed systems rather than flow-through reactors. The investigation of modern hydrothermal systems like Lost City (alkaline and not too hot) can also go a long way towards improving our understanding of life’s possible starting conditions.

Far from being too complex to have powered early life, it is actually nearly impossible to see how life could have begun in the absence of proton gradients, provided for ‘free’ as the natural result of a global geochemical process. It is time to cast off the shackles of fermentation in some primordial soup as ‘life without oxygen’ – an idea that dates back to a time before anybody had any understanding of how ATP is made – and to embrace the most revolutionary idea in biology since Darwin as ‘life without oxygen’ – an idea that dates back to a time before anybody had any understanding of how ATP is made – and to embrace the most revolutionary idea in biology since Darwin as ‘life without oxygen’ – an idea that dates back to a time before anybody had any understanding of how ATP is made – and to embrace the most revolutionary idea in biology since Darwin as ‘life without oxygen’ – an idea that dates back to a time before anybody had any understanding of how ATP is made – and to embrace the most revolutionary idea in biology since Darwin as ‘life without oxygen’ – an idea that dates back to a time before anybody had any understanding of how ATP is made – and to embrace the most revolutionary idea in biology since Darwin as.

Acknowledgments: We are grateful to Michael J. Russell, Wolfgang Nitschke and Franklin M. Harold for discussions and comments on the manuscript. NL is grateful for funding from the UCL Provost’s Venture Research Fellowship.

References

Problems and paradigms