CHAPTER TEN

Ultimate Riddle

Origin of Cellular Life

A riddle wrapped in a mystery inside an enigma — Winston Churchill, speaking of Russia

Prebiotic Chemistry
The Beginnings of Order
Protocells
From Protocells to Progenotes
Enigma

The difference between a puzzle and a mystery is that the former can be solved within the framework of known principles, while the latter cannot. Over the past sixty years, dedicated and skillful scientists have devoted much effort and ink to the origin of life, with remarkably little to show for it. Judging by the volume of the literature, both experimental and theoretical, the inquiry has thrived prodigiously. But unlike more conventional fields of biological research, the study of life's origins has failed to generate a coherent and persuasive framework that gives meaning to the growing heap of data and speculation; and this suggests that we may still be missing some essential insight.

The search for answers is firmly constrained by the insistence that acceptable hypotheses be formulated in naturalistic terms, excluding *a priori* any intervention by supernatural forces. Scientists' refusal to grant some space to the mind and will of God may strike the majority of mankind as arbitrary and narrow-minded, but it is essential if the origin of life is to remain within the domain of science. A nudge from the divine would help us clear some very high hurdles; but once that possibility is admitted

there will be no place to stop, and soon the settled principle of evolution by natural selection would be thrown into doubt. Scientific students of genesis are also inclined to set aside the possibility that life had an extraterrestrial cradle or that life began with some exceedingly rare chance event or fluctuation—effectively, with a miracle. As a practical matter, we assume that life originated here on earth by a natural and probable outgrowth of the chemical and physical circumstances prevailing some four billion years ago. The origin of life can therefore be construed as a problem in geochemistry; but it is also the black hole at the root of biological organization, and there's the rub.

What we have learned supports the conclusion that, once the organization and machinery of contemporary cells had come into existence (genes and their expression, catalytic proteins, phospholipid membranes, metabolic pathways, and so on), evolution was safely in Darwin's hands. Many puzzles remain to be solved, but there is no mystery: cellular life was shaped by the same evolutionary forces still at work today, including the transmission of genes and the creation of new ones, mutation, symbiosis, and cell heredity, all channeled by natural selection towards rising complexity and autonomy. But how could life ever get to that cellular stage? The answer must turn on the spontaneous emergence of molecular systems that first drew matter and energy unto themselves, maintained their integrity, reproduced their own kind, and evolved over time into cells as we know them. What little we have learned of these matters is the subject of the present chapter.

Life's origin has been most ardently pursued by chemists, apparently on the unspoken premise that once the molecular building blocks are on hand, cellular organization will take care of itself. That premise is surely incorrect. Modern cells do not assemble themselves from preformed constituents, and they would not have done so in the past. Instead, if we exclude explanations not grounded in natural law, we must envisage a protracted phase of prebiotic chemical evolution that generated most of the molecular parts *pari passu* with their increasingly purposeful and elaborate organization. How this might have come about remains one of the deep mysteries of biology. It has engaged the attention of numerous capable scientists and generated an enormous literature, including a number of recent book-length treatments.[1] Yet we seem to be little closer to enlightenment than were A. I. Oparin and J. B. S. Haldane, who pioneered this line of inquiry seventy years ago.[2] The timeline sketched in figure 6.2, with its stately progression from a stable earth through prebiotic chemistry to

protocells to the diversification of cellular life may look like a hypothesis, but it is not. Its purpose is merely to frame the questions and help readers keep a handhold on uncertainty.

Prebiotic Chemistry

A pioneering experiment published half a century ago laid the foundation and set the tone for almost all the subsequent research into the origin of life.[3] The late Stanley Miller, then a graduate student in the laboratory of Harold Urey at the University of Chicago, set out to test ideas formulated in the 1930s by J. B. S. Haldane and A. I. Oparin, who argued that abiotic chemical processes on the lifeless earth would generate a large stock of sundry organic molecules. In the absence of living things to consume them, these substances would accumulate in ponds and in the sea and supply all the essential raw materials for the assembly of primitive cells. These earliest protocells would be heterotrophs, reliant on the "primordial soup" for both constituents and energy. Miller constructed an ingenious apparatus in which mixtures of carbon dioxide, methane, ammonia, and hydrogen (thought at the time to resemble the composition of the early earth's atmosphere) were subjected to electric discharges (to simulate lightning), while the products were collected in a water trap representing the ocean. Over a period of days, organic matter accumulated in the trap. Most of it consisted of black tar, but there was also a significant yield of soluble small molecules, including several of biological relevance: glycine, alanine, glutamic, and aspartic acids and traces of other amino acids. There were also products that play no role in contemporary organisms, such as alpha-aminobutyric acid. The door to the origin of life had, it seemed, been cracked open.

From a contemporary perspective, the details of Miller's experiments are less important than their role in seeding a new branch of organic chemistry, sometimes called "prebiotic synthesis." Over the past fifty years, dozens of biological molecules have been produced under conditions that the authors, at least, consider to be consistent with circumstances that prevailed on earth four billion years ago.[4] These include a suite of the simpler amino acids, purines, and pyrimidines, an array of sugars, and simple carboxylic acids. Fatty acids and alcohols are harder to make, but complex mixtures of lipoidal substances can be extracted from stony meteorites; remarkably, these assemble spontaneously into membrane-bound vesicles.

Theoretical arguments suggest that members of the citric acid cycle were likely to have been available, especially under the conditions that prevail in hydrothermal vents. The roster of metabolites is far from complete, but growing. Recent discoveries include a route for the formation of pyrimidine nucleotides and the discovery of small organic molecules of biological interest in certain meteorites[5].

Macromolecules and molecular assemblages have also been generated by abiotic procedures. Mixtures of amino acids, when heated dry, polymerize into "proteinoids," chains linked together by peptide bonds, though not necessarily in the alpha position, as biological proteins are. Activated nucleotides adsorbed onto montmorillonite, a kind of clay, link up into polymers akin to RNA, with a length of up to forty residues. The operative principle here is that the catalyst, like heating and drying, promotes polymerization by driving the removal of water to form the bonds between monomers. Clays are also effective in promoting the self-assembly of alcohols and fatty acids into vesicles.[6]

The holy grail of prebiotic synthesis is the production of a linear macromolecule capable of replicating its own sequence, or that of an external template. In the eyes of many students of the origin of life, self-replicating RNA makes the most plausible starting point, and the quest has generated a devoted following.[7] Nevertheless, efforts to generate such a molecule abiotically, in the absence of some kind of biological catalyst, have not been successful. The most promising approach relies on the selection of a catalytically active RNA, a ribozyme, out of a large pool of random RNAs generated with the aid of enzymes. One such ribozyme proved capable of extending a primer, directed by an external RNA template, by as much as fourteen monomers, with a fidelity of 97 percent. More recently, Wochner and colleagues succeeded in engineering a ribozyme capable of synthesizing catalytically active RNA molecules ninety-five nucleotides in length, while Lincoln and Joyce produced a system in which a pair of ribozymes catalyze each other's synthesis.[8] However, it has thus far not been possible to isolate an efficient, self-replicating species of RNA; the exercise is clearly far more difficult than had been anticipated, and may not be feasible at all. This bleak assessment has sparked a search for other kinds of self-replicating structures that might have preceded RNA, again with partial success, but the relevance of those products to the origin of life remains to be demonstrated.

Even if the next issue of *Nature* heralds the successful isolation of a straightforward self-replicating species of RNA, the notion that the first

protocells assembled themselves spontaneously from a generous menu of precursor molecules conveniently supplied by abiotic chemistry (or imported by way of comets and meteorites) is now widely recognized as simplistic and effectively has been abandoned.[9] Among its most cogent critics are experienced masters of the art of prebiotic synthesis, who are well aware of the shortcomings of many of the proposed routes and of the wide gap between the range of molecules that living things employ and those that can be made in the laboratory. The significance of the "prebiotic" molecules is not that they served as the raw materials for life's assembly but that they illustrate chemical structures that nature could readily discover and put to use. Almost certainly, life's chemical complexity did not spring directly from lifeless precursors, but arose in the context of an evolving cellular organization.

The Beginnings of Order

"The great mystery of life's origin lies in the huge gap between molecules and cells."[10] We do not presently understand how emerging life bridged this gulf. However, if we reject both the spontaneous coalescence of prefabricated components and supernatural intervention, we must envisage a protracted transition that began in geochemistry and led to molecular systems that exhibited at least some of the qualities of life. It's a tall order, and the more you reflect on it the taller it grows. In this section we shall consider not so much *how* this might have happened but *why*. Let us grant that a prebiotic broth of diverse organic molecules had come into existence. What natural causes might have ignited the rise of complex systems, and promoted the emergence of biological functions, autonomy and purposeful action?

THE POWER OF REPLICATION. Of all the qualities that distinguish living things from inanimate objects, none is so spectacular as the capacity of the former to multiply and reproduce their own kind. Sprinkle a few bacterial cells into a flask containing a solution of salts and a pinch of sugar; next morning the sugar has been consumed and the medium teems with bacteria, each one a replica of the few that began it all. How bacteria grow and divide is a study in complicated cell physiology, but embedded deep in its core is a more elemental chemical process, the replication of those cells' genetic material. Like the multiplication of cells, replication of nucleic acids is an autocatalytic process: the product catalyzes its own formation,

and so in principle the reaction proceeds at an ever-accelerating rate until it exhausts its substrates. The idea that life began its long journey with the appearance of a self-replicating molecule of some kind is both powerful and seductive; and it has come to exercise almost a stranglehold on the biological imagination.

The standard version of the hypothesis was formulated some forty years ago by Manfred Eigen, who proposed that life took its first baby step when a self-replicating molecule, probably RNA, arose by chance in the prebiotic soup of organic molecules.[11] Popular from the start, the hypothesis was heavily bolstered in the 1980s by the demonstration that RNA is not only a carrier of sequence information but a chemical catalyst as well. RNA thus embodies the essential qualities of both genes and enzymes, making it a prime candidate for the precursor to all the marvels of biology. That first clumsy replicator evolved thanks to imperfect replication ("mutation") and chemical selection for enhanced replication and stability (a function of the relationship between the rates of production and of degradation). Replicators clothed themselves in accessory molecules, including abiotic peptides, and eventually became protogenes. With the invention of translation and the coding mechanism, naked protogenes turned into assemblies; metabolism, membranes, and cell organization came later.

What driving force would underpin the progression from molecule to system? In a series of lucid and closely reasoned articles, the Israeli chemist Addy Pross argues that it is the kinetic power of replication.[12] All chemical reactions are governed by the laws of thermodynamics; a reaction can proceed only to an extent consistent with the chemical potential of its participants, as expressed in the change of free energy. This holds for replication too; molecular replication entails an input of substrates and energy (in the simplest instance, in the form of nucleoside triphosphates ready to be polymerized), and it can proceed only to the extent allowed by their chemical potential. But thermodynamics does not determine reaction *rates*. The autocatalytic nature of replication means that it will proceed much faster than any competing reaction, even one that would be thermodynamically favored if it were given time to go to completion. "Thus we can consider this kinetic phenomenon as an enormously powerful driving force responsible for the emergence and evolution of life. This force operates at every stage—from the primal replication reaction stage through to the single-cell stage and on to the multi-cell stage."[13]

Beginning with a basic molecular replicator, kinetic drive can support the exploration of complexity space by mechanisms that are wholly chemical yet guided by natural selection. Sequences tend to grow longer by

accident during replication, increasing the amount of information they can potentially encode. For elongation to be favored despite its costs, increasing length must enhance the "effectiveness" of replication, defined by features that favor production of replicas over their decay. Such benefits may be purchased by incorporating accessory molecules into the replication pathway, with a premium on anything that enhances the supply of precursors and energy. So long as these can be found, increasing complexity can confer advantages that exceed its costs. The power of replication does not in itself explain how translation, coding, or membranes came to be, but it ensures that, if and when they are invented, they will be drawn into the service of replication. Pross stands on its head the traditional premise that "replication is a manifestation of life." On the contrary, from the chemical perspective, "life is a manifestation of replication." And the chemical, kinetic drive of replication suffices to explain life's teleonomic character, the undeniably purposeful behavior of living things. Initially, at the stage of plain molecular replication, that teleonomic character is rudimentary, expressed only in the "desire" to replicate. As evolution proceeds matters become more complicated, but "when all is said and done, all entities along that entire route, from replicating molecules through to human beings, in essence share the same 'dream'—to replicate."[14]

This is heady stuff, argued with conviction and style. I find the explosive power of replication very helpful in understanding how life has colonized every habitable nook and cranny, even the most inhospitable (a point also made by Richard Dawkins years ago). But there are at least two reasons to reserve judgment about the case Pross makes. First, the argument hinges on the chance appearance of a primal self-replicating molecule in the prebiotic broth. How likely is that? We know so little about the time, place, and circumstances that nothing can be ruled out. But the fact is that chemists have encountered insuperable difficulties in generating a working replicator, and many have expressed doubts about the project.[15] It is at least incumbent upon proponents of its spontaneous genesis to explain how the "correct" monomers could have been selected from the "prebiotic clutter," how a sufficient concentration of monomers was maintained, where the energy came from, and how the replicator evaded the tendency of polymers to break down by hydrolysis. Second, the gradual accretion of accessories implies that replication continued in a sustained manner over a protracted period, which calls for a very special locale. These misgivings would be somewhat allayed if, instead of a naked replicator, we envisage that replication was from the beginning a property of an organized system, sequestered in a compartment and sustained by a flux of energy.

ORGANIZED BY ENERGY FLOW. Nonconformists who doubt that life began with a molecular replicator and grew complex over time are apt instead to argue that life began in complexity, with self-organized networks of chemical reactions.[16] The basic premise is that the prebiotic environment supplied a rich diversity of organic substances, a variety of energy sources, and also potential catalysts, including clays, transition metals, and abiotic peptides. Chemistry was everywhere rampant, with reactions coalescing into networks of mounting complexity. A few will have attained "catalytic closure," the capacity to reproduce as a collective entity even though no single constituent replicates itself. Prebiotic reaction networks will have made up an ancestral version of metabolism, whose lineaments can still be traced in the metabolic machinery of contemporary cells. At the inception of life, the critical qualities were kinetic stability and error tolerance. Enclosure within a boundary must have been an early feature, but replicators, proteins, and structural order were later accretions.

These are stirring ideas. As Leslie Orgel put it, "The demonstration of the existence of a complex nonenzymatic metabolic cycle . . . would be a major step in research on the origin of life, while demonstration of an evolving family of such cycles would transform the subject."[17] But the proposal that metabolic organization can arise and persist of its own accord in the absence of both enzymes and genes seems to fly in the face of reason. There is no familiar precedent to lean on, and the spontaneous emergence of order out of chaos violates one's intuitive sense of how the world works. Indeed, if chemical networks could arise spontaneously, one would expect them to show up as a cause of spoilage in canned food. The case for self-organization becomes plausible only when the emergence of networks is strongly coupled to a flow of energy.

The cardinal point was made by Harold Morowitz forty years ago: the flow of energy through a system acts to organize that system. At the chemical level, coupled energy flux promotes the emergence of cyclic reaction pathways. At the physical level it causes the appearance of dissipative structures, including such familiar objects as flames, hurricanes, whirlpools, and those remarkable "Bènard cells" that appear when a pan filled with liquid is uniformly heated from below. There is nothing mystical about the emergence of large-scale patterns in systems kept far from equilibrium by the flux of energy; the phenomenon is well documented, and seems certain to play a major role in the inception of chemical and structural organization. I am inclined to go much further, in company with those who claim that the flux of energy is what called life into being.[18]

If this general proposition is to become a testable hypothesis for the

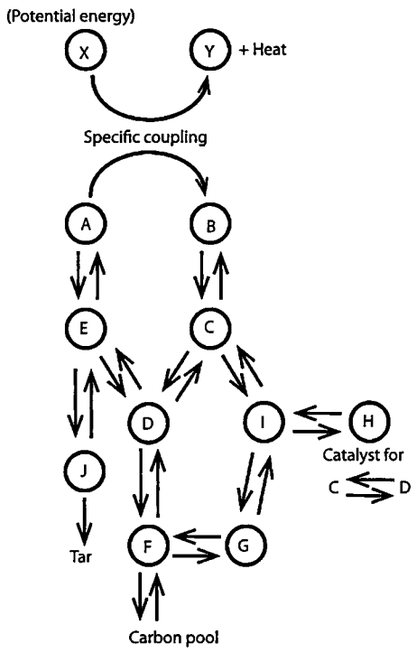


FIGURE 10.1. A metabolic network organized by a "driver" reaction. The exergonic (energy-yielding) transformation of X to Y, external to the cycle, is chemically coupled to the transformation of A to B (the "driver" reaction). This joint reaction moves the cycle represented by the letters A to E in the clockwise direction. The movement of material from side reactions (indicated by the equilibria associated with substances F to J), and from an external carbon source ("carbon pool") into the cycle is favored by the release of the potential energy in the transformation of X to Y. (From Shapiro 2006, with permission of the University of Chicago Press.)

origin of chemical complexity, it is necessary to specify both an energy source and a credible coupling reaction to funnel that energy into a reaction sequence. An abstract example, formulated by the late Robert Shapiro, is sketched in figure 10.1.[19] It postulates an external energy source, the conversion of X to Y, which is chemically coupled to the conversion of A to B, one reaction in a hypothetical metabolic cycle. Thanks to this coupling, the reaction sequence represented by A to E is moved in the clockwise direction, ultimately regenerating A; the cycle can expand by drawing in materials from external sources (reactions F through J) and generate a net yield of some product. Coupled reactions of this sort are a

staple of today's biochemistry. Most cellular metabolites are made by reactions that couple biosynthesis to the hydrolysis of ATP, which is regenerated elsewhere (see chapter 5). For a network of reactions to operate, it must be sequestered in a compartment of some kind, if only to keep the reactants from diffusing away. There was no shortage of energy, with light and geochemical reactions being the most realistic.

All this sounds sensible enough in principle, at least to my mind, but a little reflection suffices to show how thin the ice is.[20] Modern biochemical cycles are catalyzed by specific and dedicated enzymes, manufactured by the cell for that purpose. But here we are thinking about a time before enzymes, when the only catalysts were minerals and metal ions. Is it really likely that all the requisite catalysts would occur together in one place, and would catalyze precisely the steps required to keep the cycle turning, rather than side-reactions that draw off the participants into byproducts (tar, in fig. 10.1)? It would be reassuring if one could point to a single known cycle, complete with energy source and coupling reaction, that operates in the absence of enzymes! Here, if anywhere, is an opportunity for experimentalists to make a contribution.

SELECTED FROM THE OUTSET. There are, Richard Dawkins once observed, only two possible explanations for our existence: God and natural selection. With the former excluded *a priori*, scientists must account for the transition from chemistry to biology by some sort of undirected evolution. The principles that drive biological evolution today are well understood: reproduction, heritable variation, competition, selection for reproductive advantage, and hence adaptation to a changing environment. The central issue is whether abiotic chemical systems, devoid of the apparatus for the expression and transmission of genetic information, could still exhibit some degree of heredity with variation and be subject to natural selection. If so, they might serve as a platform upon which evolution could construct more sophisticated mechanisms. On that premise, the origin of cells has much in common with subsequent evolution, even though the objects of selection, the criteria for success, and the mechanisms responsible have changed as the transition progressed.

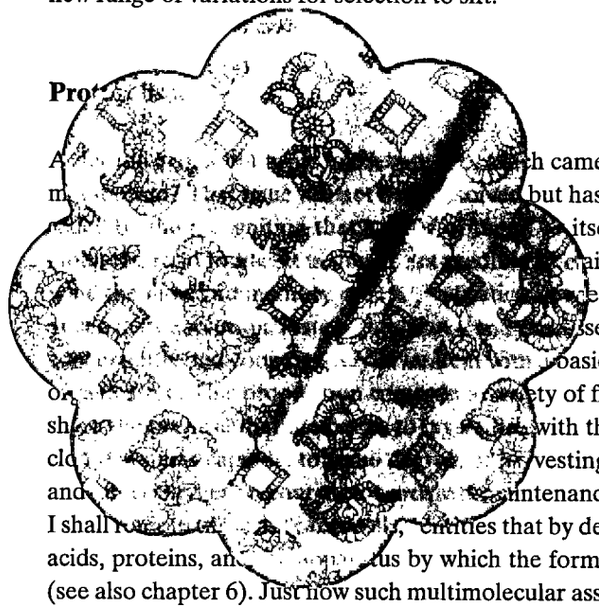
On the premise that life began with a self-replicating molecule, conventionally taken to be RNA, the pattern of prebiotic evolution is comprehensible, at least in outline. In the absence of enzymes, replication was slow and subject to frequent errors that generated a cloud of mutants derived from the master sequence. It was natural selection that steered the kinetic

power of replication on a course that led to more accurate and stable replication and, later on, to complex assemblies culminating in cells.[21] The first product of selection will have been RNA itself, selected out of a mixture of abiotic molecules by virtue of its capacity to replicate. Selection would have favored the acquisition of ancillary factors including amino acids, abiotic peptides, and nucleic acids. At a later stage selection would have favored replicators capable of specifying the structures of these auxiliary factors; only then did replicators become primitive genes. Just how this came about, and what were the criteria for selection, is not altogether clear to anyone. But it seems fair to say that the evolution of replicators cannot have gone very far in the absence of mechanisms to ensure a stable milieu and a supply of energy and precursors. If ever there was an era of free replicators, it must have been short-lived.

It's a grand leap from independent replicators to cells, because in a cell replication is constrained for the benefit of the collective. The persistence of a gene depends on the function it encodes and thus on the contribution it makes to the survival of the community; it has little to do with the replication of that gene itself (there are exceptions, of course—parasitic replicators that make no contribution to the common welfare—but I shall here set those aside). In fact, natural selection today does not see replicators at all. It judges the phenotype of the cell, which may be quite remote from the functions encoded by the genes. Just how those (hypothetical) free replicators would have been herded into communities and then taught to accept the collective discipline is one of the questions under debate.

On the alternative premise, that the first steps to life were represented by self-organized metabolic cycles, evolution must be grounded in the properties of systems. Individual cycles, like eddies, could "compete" for resources, producing winners and losers. In theory, at least, they can undergo enlargement and a kind of propagation, even the transmission of their compositional identity. I quite like this notion, but there was obviously a limit to the range of possible variations until the advent of digital genes made unlimited heredity possible. Besides, a substantial baseline level of functional organization seems to be a prerequisite for natural selection to favor progressively higher levels of adaptive order. The implication is that, even if nucleic acids were not first on the stage, they must have made their entrance very early in the drama. Cellular organization, founded on the collaboration of genes, catalysts, and membranes, was not a late stage in the advance of either nucleic acids or metabolic cycles; it is what the origin of life was all about.

How cellular organization came about is a great mystery, but it is not hard to see why natural selection would favor systems of this kind over unstructured ones. Enclosure in a pod, such as a membrane-bound vesicle, keeps multiple parts together and makes it possible to maintain an ionic environment that suits their operations. Membranes today are crucial to the generation of useful energy, and it is likely that this was true from the start. Free replicators compete for raw materials, but genes cooperate for the good of a larger unit that maintains autonomy from its environment. And with the advent of cells came the rudiments of individuality and a vast new range of variations for selection to sift.



Pro

A h came first, replication or m but has been largely super- itself, can take one far aim that anything be- etabolism could have sses must have evolved basic mode of structural ty of flavors,[22] all of which sh with the emergence of en- cl vesting matter and energy an intenance and reproduction. I shall r entities that by definition lack nucleic acids, proteins, an us by which the former specify the latter (see also chapter 6). Just how such multimolecular assemblies arose is not well understood and is the subject of much research. Here we shall focus on the principles, beginning with the need for a physical boundary.

NO LIFE WITHOUT MEMBRANES. Cells as we know them are defined by their plasma membrane, which segregates the workplace from the environment. It is only thanks to the membrane that cells can maintain a distinctive cytoplasmic milieu, which differs from all the rest of the world in its pH, ionic composition, and its complement of small molecules. The barrier function of the membrane stems from its structure, a phospholipid bilayer that is almost impermeable to ions and polar molecules. This feature in turn imposes a requirement for selective portals and conduits to