Cambridge University Press
978-0-521-87102-0 - Fitness of the Cosmos for Life: Biochemistry and Fine-Tuning
Edited by John D. Barrow, Simon Conway Morris, Stephen J. Freeland and Charles L. Harper
Frontmatter
More information

Foreword: The improbability of life

George M. Whitesides

How did life begin?

I (and most scientists) would answer, "By accident." But what an absolutely unlikely accident it must have been! The earth on which life first appeared – prebiotic earth – was most inhospitable: a violent place, wracked by storms and volcanoes, wrenched by the pull of a moon that was much closer than the one we know now, still battered by cosmic impacts. On its surface and in its oceans were myriads of organic compounds, some formed in processes occurring on earth, some imported by infalls from space. Out of this universe of tumult and molecules, somehow a small subset of chemical processes emerged and accidentally replicated, thus stumbling toward what became the first cells. How could such a chaotic mixture of molecules have generated cells? Order usually decays toward disorder: *Why do the tracks that led to life point in the opposite direction?*

The origin of life is one of the biggest of the big questions about the nature of existence. *Origin* tends to occur frequently in these big questions: the *origin* of the universe, the *origin* of matter, the *origin* of life, the *origin* of sentience. We, scientists and non-scientists alike, have troubles with such "origins" – we were not there watching when the first events happened, we can never replicate them, and, when those first events happened, there was, in fact, no "we." I believe that one day we will be able to describe life in physical terms – that is, we will rationalize life satisfactorily in molecular detail based on accepted scientific law and scientific theory using the scientific method. But we certainly do not know yet how to do it.

Understanding how organized living cells emerged from disorganized mixtures of molecules is an entrancingly, seductively difficult problem – so difficult, as we now understand it, that science does not even have well-formulated, testable hypotheses about how it might have happened, only guesses and intuitions. This problem deserves our most careful thought. Its solution will tell us about our origins and describe how disorder can spontaneously become order. It will also test the capability of current science to understand systems comprising many interacting parts.

Before trying to answer the question *How did life begin?*, we must first think about what the question really is that we are trying to answer: What is the "life" whose origins we are trying to understand? What are the characteristics of a cell, the simplest embodiment of life, that might allow us to trace back to its origins? How do we recognize an "origin"? When does a set of molecules, and of processes that convert these molecules into one another, cross a line separating "not-alive" from "alive"? And what is the tool – the "scientific method" – that science will use to try to address this problem?

Let us begin with the scientific method, a very useful and quite reliable strategy for doing science. Although it sometimes seems plodding, the scientific method can tease apart astonishingly difficult and complicated problems by careful attention to detail. It starts with rigorously reproducible empirical observations: "Things fall down, not up." "Two objects at different temperatures, when placed in contact, reach the same temperature." "Hydrogen atoms absorb only light that has specific frequencies." The scientific method codifies and quantifies these observations as "physical laws," builds theories (Newtonian mechanics, thermodynamics, quantum mechanics) based on those laws, and then tests new observations or hypotheses for their compatibility with these theories. Based on these theories, science rationalizes the physical world and predicts aspects of it not previously observed. The tools of the scientific method are the millstones and the oven that science uses to grind observations into theory and bake theory into prediction.

The scientific method works most rigorously when it identifies observations that are incompatible with current hypotheses. Faced with a new observation, scientists list all hypotheses that might explain it and then discard those that are incompatible with accepted physical law. Hypotheses that are not discarded as incompatible remain possibilities. If only one remains, it is promoted to theory. If disproving all hypotheses but one is not possible, we may retreat to demonstrating compatibility with theory, recognizing that compatibility is weaker than proof. In science, we use the phrases "I think . . ." and "I believe . . ." as synonyms, both implying ". . . based on known physical law." In other words, "This theory accommodates all the observations that we currently know."

So, what is life? We can describe what it looks like and what it does, but not how it works (most of us are in the same situation even with much simpler systems: computers, electric toothbrushes, refrigerator magnets). I suggest that life has five major physical attributes (other scientists may suggest other lists, but the general principles will usually be the same):

Cambridge University Press
978-0-521-87102-0 - Fitness of the Cosmos for Life: Biochemistry and Fine-Tuning
Edited by John D. Barrow, Simon Conway Morris, Stephen J. Freeland and Charles L. Harper
Frontmatter
More information

1. *Life is compartmentalized.* All life that we know is embodied in cells, and all cells have a continuous, closed membrane that separates "inside" from "outside."
2. *Life is dissipative, or out-of-equilibrium.* Life requires a flow of energy. If the chemical and physical processes in living cells reach equilibrium, and there is no flux of energy through the cell, it is, so far as we know, dead (or, at least, "not-alive").
3. *Life is self-replicating.* The most evident characteristic of the cell is that it was produced by the division of a parent cell, and, in many cases, it too will divide and produce daughter cells.
4. *Life is adaptive.* The cell can adapt its internal environment so that it functions even when the outside environment changes; in some circumstances, it can even modify the outside environment to make its inside more comfortable.
5. *Life occurs in water.* All life, so far as we know, involves molecules and salts dissolved or organized in a medium that is mostly water. We do not know whether water is essential to all life or just to life as we know it. But, at this time, we know no exceptions: life occurs in water.

So, according to this view, life is a spatially distinct, highly organized network of chemical reactions that occur in water and is characterized by a set of remarkable properties that enable it to replicate itself and to adapt to changes in its environment. We can, thus, describe what we are still ignorant about, but not much more.

How remarkable is life? The answer is: *very*. Those of us who deal in networks of chemical reactions know of nothing like it. We understand some – but only some – of the characteristics of the network that make it so remarkable. One key to its behavior is catalysis. The rates of essentially all cellular reactions – the processes that convert one molecule into another – are controlled by other molecules (usually by a class of protein catalysts called enzymes). The catalysts are (in some sense) like valves in a chemical plant (which, in some sense, is what a cell is): they control the rate at which one kind of molecule becomes another in a way loosely analogous to that in which a valve controls the rate at which fluid flows through a pipe. The complexity of the network becomes clear when one realizes that the catalysts – the valves – are themselves controlled by the molecules they produce: the products of one reaction can control the rate at which another reaction takes place.

The catalysts provide plausible connections among the elements of the network. The conversations among catalysts – conversations controlled by the very molecules the catalysts are controlling – allow the components of the network to form a single, coherent, interconnected, albeit very complicated, entity rather than an inchoate collection of independent processes. And how intricate these "conversations" are! The molecules whose production is required for the cell to live and to replicate itself modify the activities of the catalysts that make them. These already very complex interactions are further modulated by additional signals that come from outside the cell and by signals generated by an internal "clock." (This clock – the "cell cycle" – is itself a set of chemical reactions that oscillates spontaneously in time and defines the sequence of stages through which the cell progresses as it replicates.) Many molecules in the cell also have multiple roles: intermediates in one or many synthetic pathways, controllers of the activity of catalysts, signals for generating the catalysts and other molecules, sources of energy, and components of the physical structure of the cell.

Today, we understand many aspects of the behavior of the cell and many fragments of the network, but not how it all fits together. We particularly do not understand the stability of life and of the networks that compose it. Our experience with other very complicated networks (e.g. the global climate, air-traffic-control systems, the stock market) is that they are puzzlingly unstable and idiosyncratic. But unlike these and other such networks, life is stable – it is able to withstand, or adapt to, remarkably severe external jolts and shocks; and its stability is even more puzzling than the instability of the climate. We have a hard enough time understanding even simple sets of coupled chemical reactions. And we have, at this time, no idea how to understand (and certainly not how to construct) the network of reactions that make up the simplest cell.

So, at least for now, the cell is beyond our ability to understand it. The community of people working on the nature of life has, nonetheless, great (and probably warranted) confidence that understanding life in purely physical terms is a tractable, if difficult, problem. This confidence is enormously bolstered by two facts.

First, we are surrounded by uncountable varieties of life, especially by multitudes of different types of living cells; we thus have many examples of different forms of life. We ourselves are communities of cells with the added complexities of hierarchical organization of these cells into tissues, of tissues into organs, and of organs into the organisms that are "we."

Second, the tools of modern molecular biology have given us an astonishing capability to examine, modify, deconstruct, and reconstruct the molecular components of cells to see how they respond to our tinkering. The simplest cells (such as those of the primitive intracellular parasite *Mycoplasma genitalium*) appear to have fewer than a thousand proteins. That number of catalysts is still very complicated, and we have as yet no conceptual tools for understanding a network of reactions of such complexity. But this level of complexity does not, in principle, seem unreachably beyond our understanding. A cellular network of a thousand proteins (catalysts and molecules that sense, signal, and control passage across membranes; act as the structural skeleton; and perform many other functions) talking to one another in groups through the compounds they produce seems to be something that we will be able to disentangle. Certainly, those who call themselves "systems biologists" believe we will. Still, the path that scientists are now following in trying to understand the molecular basis of life will test their creativity and strain their endurance:

Cambridge University Press
978-0-521-87102-0 - Fitness of the Cosmos for Life: Biochemistry and Fine-Tuning
Edited by John D. Barrow, Simon Conway Morris, Stephen J. Freeland and Charles L. Harper
Frontmatter
More information

first, understanding the pieces of the networks as thoroughly as possible; then, perhaps, devising a computer model of a cell; and ultimately, in some distant future, validating the correctness of the principles suggested by this model by designing a set of reactions entirely different from those in the cells we now know.

It is one thing to analyze a Bach fugue; it is quite a different thing to play one, or to write one, or to create the kind of communication between humans that we call "music." We shall, I confidently believe, eventually analyze the fugue of life – the interplay of metabolic processes in the cell – as a network of compartmentalized, adaptive chemical reactions that can, astonishingly, replicate repeatedly into identical, distinct, separate networks. This is a very difficult job, but one that we humans can accomplish. But where did the cell come from? How did this wonderfully, astonishingly complex system come into existence? We do not know. If it is very difficult to understand the operation of cellular life as we observe it today, it is even more difficult to understand how it might have originated in the past.

Thoughtful, deeply creative people from a wide range of backgrounds have been captivated by the question of the origin of life. There is no shortage of ideas about pieces of this puzzle. We know how the surfaces of minerals might have provided elementary, non-biological catalysts to start the process and how heat or sunlight might have contributed other reactions. We can guess why certain types of molecules and reactions tend to occur in metabolism. We understand how any number of plausible natural events occurring in a conceivable prebiotic earth – events that formed complex mixtures of chemicals in geothermal vents, in lightning, on impacts, and under intense solar irradiation – might have contributed relevant bits of chemistry. But we do *not* understand how something as subtle and complicated as the network of reactions that we recognize the cell to be – a network both responsive and robust – might have emerged from these rudimentary processes. How could a chemical sludge spontaneously become a rose, even with billions of years to try?

We can take two approaches in our research directed toward the origin of life: reasoning *backward* and reasoning *forward*. "Backward" starts with life as we know and characterize it now – cells, DNA, RNA, enzymes, membranes, metabolites, membrane receptors, channels, and import/export proteins – and extrapolates back to simpler and simpler systems to try to infer an origin. This approach has been spectacularly successful in "reverse engineering" evolution, at least part of the way; but it has always been guided by examples provided by the types of cells that are now alive. Still, there seems little doubt that evolution could proceed once there was a primitive cell, with RNA or an RNA-like molecule, and reactions that used RNA as a catalyst and also translated RNA into protein or protein-like catalysts that were part of the network of reactions. Several hundreds of millions of tidal pools, together with enormous volumes of lakes and oceans, over several hundreds of millions of years provided many opportunities to produce cellular and organismic

Cambridge University Press
978-0-521-87102-0 - Fitness of the Cosmos for Life: Biochemistry and Fine-Tuning
Edited by John D. Barrow, Simon Conway Morris, Stephen J. Freeland and Charles L. Harper
Frontmatter
More information

complexity. This part of the development of the complexity of life no longer seems to be a serious issue, at least conceptually. And the anatomical and physiological structures that now so enthrall us – the eye, the ear, the kidney, tentacles, muscles – these all seem to me transfixingly interesting products of evolution, but not ones whose origins are incomprehensibly improbable. If we and the squid have the same camera eye, why not? With enough tries, "best" solutions are bound to emerge many times. If some creatures walk on two legs, some on four, some on six or eight – again, why not? Many solutions may work well enough to survive the rigors of evolutionary selection.

Reasoning "forward" is much more problematic. Although we can imagine many possible mangers for the birth of life – deep smokers in the abyssal depths, tidal pools, hot springs, and many others – and although each could plausibly produce primitive precursors to many of the reactions that now constitute cellular metabolism, we have (in my opinion) no idea how these simple reactions might have blundered together to make the first protocell. Monkeys sitting at typewriters pecking out Shakespeare seems child's play by comparison. For example, we still do not know:

- *What were the first catalysts*? Were they protein-analogs or RNA-analogs or minerals or some other species of which there is now no trace?
- *How did the first networks form, and why did they persist?* One can imagine countless catalytic reactions that might have occurred, but how some of these reactions became self-sustaining networks is entirely obscure.
- *How could the process that stores the information that specifies the catalysts – the RNA or precursor of the primitive cells – have evolved?* The connection between RNA (or its younger, more evolved cousin, DNA) and the proteins that are catalysts, the enzymes, is not at all obvious; how the two co-evolved is even less clear.
- *How did the energetic cycles that power every cell emerge?* Why is there potassium ion on the inside of the cell and sodium ion on the outside? What was the origin of chemiosmosis? Given the extraordinary complexity of the ATPases – the complicated aggregates of proteins that generate ATP using the free energy that derives from differences in the concentration of ions across membranes – how could they have evolved? We simply do not know.

Nothing in the cell violates the fundamental laws of physical science. The second law of thermodynamics, the law that describes everything that occurs in the range of sizes relevant to life, can sleep untroubled. The flux of energy – now (although not necessarily originally) produced in nuclear reactions in our sun, transferred to the surface of earth as sunlight, absorbed by plants in photosynthesis, captured as glucose and other compounds, used in the cell to generate the intermediates that make metabolism possible, and ultimately dissipated to space by radiation as heat – can evidently support life. But how life originated is simply not apparent. It seems

Cambridge University Press
978-0-521-87102-0 - Fitness of the Cosmos for Life: Biochemistry and Fine-Tuning
Edited by John D. Barrow, Simon Conway Morris, Stephen J. Freeland and Charles L. Harper
Frontmatter
More information

so improbable! The complexity of the simplest cell eludes our understanding – how could it be that any cell, even one simpler than the simplest that we know, emerged from the tangle of accidental reactions occurring in the molecular sludge that covered the prebiotic earth? We (or, at least, I) do not understand. It is not impossible, but it seems very, very improbable.

This improbability is the crux of the matter. The scientific method can be paralyzed by problems that require understanding the very improbable occurrences that result from very, very large numbers of throws of the dice. Sometimes we can understand the statistics of the problem; sometimes we cannot. How likely is it that a comet will hit the earth? We now have good geological records. How likely is it that a star will explode into a nova? There are many, many observable stars, and we now understand the statistics of nova formation quite well.

But how likely is it that a newly formed planet, with surface conditions that support liquid water, will give rise to life? We have, at this time, no clue, and no convincing way of estimating. From what we do now know, the answer falls somewhere between "impossibly unlikely" and "absolutely inevitable." We cannot calculate the odds of the spontaneous emergence of cellular life on a plausible prebiotic earth in any satisfying and convincing way.

What to do? For all its apparent improbability, life does seem to have happened here (or perhaps on some similar planet that transferred life to here). Rationalizing the origin of life is a problem that chemists are probably best able to solve. Life is a molecular phenomenon. The possibilities of alternative universes and different distributions of the elements are irrelevant from the vantage point of the particular universe and planet – our earth – that we share with so many other forms of life. We understand the chemical elements (we do not need to know about exotic forms of matter or energy in this enterprise), the molecules they form, and their reactivities. We know the players in the game, and we understand the game they play. We can guess (albeit only roughly) the distribution of the elements on the surface of the earth in the epoch in which we believe that life emerged, and we can infer the abundances of the molecules that were probably present. We understand how catalysts function. But we do not see how it all fits together.

Is this a problem in which science can make progress? Yes, and perhaps no. Those researchers who have taken the approach of reasoning "backward" to infer how life might have been born have made rapid progress. They have used the tools of molecular biology to trace the early stages of evolution back to the point where DNA gave way to RNA, which in turn probably gave way to some more primitive molecule whose composition we don't know, but which was probably related to RNA. The paths are fainter and fainter as the trail becomes older and colder and as we move from fact into speculation beyond RNA. We still do not understand the connections between RNA, or its forgotten ancestor, and enzymes, or their also forgotten ancestors, and the metabolic web that supports and constitutes life. Moving "forward" – spinning and weaving the threads that connect "molecules" to "life" – has been technically and conceptually more difficult.

Still, compelling connections are apparent between what might have existed on the prebiotic earth and the molecules of surprising complexity that are now vital to life. We understand, for example, how molecules of astonishing sophistication, such as the porphyrins – the precursors to the "green" of the pigments that serve plants in photosynthesis and the "red" of the hemoglobin that transports oxygen in our blood – could have arisen from aqueous solutions of hydrogen cyanide, one of the simplest of molecules and a possible component of the atmosphere of prebiotic earth. But these demonstrations, marvelous as they are, do not bridge the gap between "forward" pathways from prebiotic molecules to life and "backward" pathways from modern cells to possible progenitors, those emerging from the gray area between "alive" and "not-alive." As yet, no step goes from solutions of molecules to the networks of interconverting molecules that make up living cells. I believe that no one yet knows how to bridge that gap.

How to progress? The best lead to the hardest part of the problem – the "forward" problem – is the hypothesis that life evolved, somehow, from autocatalytic reactions (that is, reactions whose products are themselves catalysts for the reactions that produce them). We know something about autocatalytic reactions: flames are autocatalytic, and so are explosions (and one speaks, sometimes, of the "explosion" of life). We also know other reactions that are autocatalytic, although the subject of "autocatalysis" has not been a particular preoccupation of chemistry or biochemistry. Autocatalysis offers, I believe, a plausible trail into the wilderness.

Here, I suggest, is a *process* that science can use to examine this question. Let us build and understand autocatalytic reactions; extend that understanding to other networks of catalytic reactions; and develop simple, and then more complex, networks of autocatalytic and catalytic reactions. If, in time, we can trace a pathway from "chemical sludge" to "life," we shall have provided an argument based on plausibility, if not on proof, for the origin of life.

If, in time, we cannot trace such a path, what then? In science, until it has been proven that something cannot be done, it is always possible that it can be done. Proving that life did not originate by accident in tidal pools or black smokers will be more difficult than proving that it might have done so. Also, patience may be in order. What is impossible for science today may be trivial for science in the future.

There is still much that we do not understand about nature. As we learn more, I believe that we will ultimately see a path – based on principles of chemistry and physics and geology – that could plausibly have led from disorganized mixtures of inanimate chemicals to the astonishingly ordered, self-replicating networks of reactions that provide the basis for life. The fact that I cannot yet understand how an inconceivably large number of tries at an extraordinarily improbable event might lead to "life" is more a reflection of my limited ability to understand than evidence of a requirement for some new principle. But, having said all of that, I do not know, and in some sense do not care, whether physical science as I now know it ultimately explains the origin of life or whether the explanation will require principles entirely new to me. I do care that science makes every effort to develop the explanation.

Although I believe that science will ultimately be successful in rationalizing the origin of life in terms of physical principles, it should be cautious and claim credit only for the puzzles it has already solved, not those whose solutions still lie in the future. The central conundrum about the origin of life – that, as an accidental event, it seems so very improbable – is not one that science has yet resolved. Claiming credit prematurely – claiming, in effect, that current science holds all the answers – may stunt the growth of the new ideas that a resolution may require.

What, then, do I know? I know that I do not, yet, understand how life originated (and that I may not live long enough to do so). Order from disorder! How could it have happened?

I also know that my father never imagined cloning, and his father would not have believed television. Go far enough back, and the wheel was beyond comprehension. Difficult problems may take time – lots of time – to solve.

And so now, after I wake in the morning – at least on a good morning after I've had my coffee and am not distracted by the countless midges that constitute most of reality-as-we-know-it – my overwhelming response to existence, and to life, remains one of delight in its wonderfully wild improbability.

For now, call it what you will. *L'Chaim!*