On the Origins of Life
By: David Berlinski
Commentary
June 14, 2007

Original Article, first published in February, 2006

For those who are studying aspects of the origin of life, the question no longer seems to be whether life could have originated by chemical processes involving non-biological components but, rather, what pathway might have been followed.
—National Academy of Sciences (1996)

It is 1828, a year that encompassed the death of Shaka, the Zulu king, the passage in the United States of the Tariff of Abominations, and the battle of Las Piedras in South America. It is, as well, the year in which the German chemist Friedrich Wöhler announced the synthesis of urea from cyanic acid and ammonia.

Discovered by H.M. Roulle in 1773, urea is the chief constituent of urine. Until 1828, chemists had assumed that urea could be produced only by a living organism. Wöhler provided the most convincing refutation imaginable of this thesis. His synthesis of urea was noteworthy, he observed with some understatement, because "it furnishes an example of the artificial production of an organic, indeed a so-called animal substance, from inorganic materials."

Wöhler's work initiated a revolution in chemistry; but it also initiated a revolution in thought. To the extent that living systems are chemical in their nature, it became possible to imagine that they might be chemical in their origin; and if chemical in their origin, then plainly physical in their nature, and hence a part of the universe that can be explained in terms of "the model for what science should be."*

In a letter written to his friend, Sir Joseph Hooker, several decades after Wöhler's announcement, Charles Darwin allowed himself to speculate. Invoking "a warm little pond" bubbling up in the dim inaccessible past, Darwin imagined that given "ammonia and phosphoric salts, light, heat, electricity, etc. present," the spontaneous generation of a "protein compound" might follow, with this compound "ready to undergo still more complex changes" and so begin Darwinian evolution itself.

Time must now be allowed to pass. Shall we say 60 years or so? Working independently, J.B.S. Haldane in England and A.I. Oparin in the Soviet Union published influential studies concerning the origin of life. Before the era of biological evolution, they conjectured, there must have been an era of *chemical* evolution taking place in something like a pre-biotic soup. A reducing atmosphere prevailed, dominated by methane and ammonia, in which hydrogen atoms, by donating their electrons (and so "reducing" their number), promoted various chemical reactions. Energy was at hand in the form of electrical discharges, and thereafter complex hydrocarbons appeared on the surface of the

sea.

The publication of Stanley Miller's paper, "A Production of Amino Acids Under Possible Primitive Earth Conditions," in the May 1953 issue of *Science* completed the inferential arc initiated by Friedrich Wöhler 125 years earlier. Miller, a graduate student, did his work at the instruction of Harold Urey. Because he did not contribute directly to the experiment, Urey insisted that his name not be listed on the paper itself. But their work is now universally known as the Miller-Urey experiment, providing evidence that a good deed can be its own reward.

By drawing inferences about pre-biotic evolution from ordinary chemistry, Haldane and Oparin had opened an imaginary door. Miller and Urey barged right through. Within the confines of two beakers, they re-created a simple pre-biotic environment. One beaker held water; the other, connected to the first by a closed system of glass tubes, held hydrogen cyanide, water, methane, and ammonia. The two beakers were thus assumed to simulate the pre-biotic ocean and its atmosphere. Water in the first could pass by evaporation to the gases in the second, with vapor returning to the original alembic by means of condensation.

Then Miller and Urey allowed an electrical spark to pass continually through the mixture of gases in the second beaker, the gods of chemistry controlling the reactions that followed with very little or no human help. A week after they had begun their experiment, Miller and Urey discovered that in addition to a tarry residue—its most notable product—their potent little planet had yielded a number of the amino acids found in living systems.

The effect among biologists (and the public) was electrifying—all the more so because of the experiment's methodological genius. Miller and Urey had done nothing. Nature had done everything. The experiment alone had parted the cloud of unknowing.

The Double Helix

In April 1953, just four weeks before Miller and Urey would report their results in *Science*, James Watson and Francis Crick published a short letter in *Nature* entitled "A Structure for Deoxyribose Nucleic Acid." The letter is now famous, if only because the exuberant Crick, at least, was persuaded that he and Watson had discovered the secret of life. In this he was mistaken: the secret of life, along with its meaning, remains hidden. But in deducing the structure of deoxyribose nucleic acid (DNA) from X-ray diffraction patterns and various chemical details, Watson and Crick *had* discovered the way in which life at the molecular level replicates itself.

Formed as a double helix, DNA, Watson and Crick argued, consists of two twisted strings facing each other and bound together by struts. Each string comprises a series of four nitrogenous bases: adenine (A), guanine (G), thymine (T), and cytosine (C). The bases are nitrogenous because their chemical activity is determined by the electrons of the nitrogen atom, and they are bases because they are one of two great chemical clans—

the other being the acids, with which they combine to form salts.

Within each strand of DNA, the nitrogenous bases are bound to a sugar, deoxyribose. Sugar molecules are in turn linked to each other by a phosphate group. When nucleotides (A, G, T, or C) are connected in a sugar-phosphate chain, they form a polynucleotide. In living DNA, two such chains face each other, their bases touching fingers, A matched to T and C to G. The coincidence between bases is known now as Watson-Crick base pairing.

"It has not escaped our notice," Watson and Crick observed, "that the specific pairings we have postulated immediately suggests a possible *copying mechanism* for the genetic material" (emphasis added). Replication proceeds, that is, when a molecule of DNA is unzipped along its internal axis, dividing the hydrogen bonds between the bases. Base pairing then works to prompt both strands of a separated double helix to form a double helix anew.

So Watson and Crick conjectured, and so it has proved.

The Synthesis of Protein

Together with Francis Crick and Maurice Wilkins, James Watson received the Nobel Prize for medicine in 1962. In his acceptance speech in Stockholm before the king of Sweden, Watson had occasion to explain his original research goals. The first was to account for genetic replication. This, he and Crick had done. The second was to describe the "way in which genes control protein synthesis." This, he was in the course of doing.

DNA is a large, long, and stable molecule. As molecules go, it is relatively inert. It is the proteins, rather, that handle the day-to-day affairs of the cell. Acting as enzymes, and so as agents of change, proteins make possible the rapid metabolism characteristic of modern organisms.

Proteins are formed from the alpha-amino acids, of which there are twenty in living systems. The prefix "alpha" designates the position of the crucial carbon atom in the amino acid, indicating that it lies adjacent to (and is bound up with) a carboxyl group comprising carbon, oxygen, again oxygen, and hydrogen. And the proteins are polymers: like DNA, their amino-acid constituents are formed into molecular chains.

But just how does the cell manage to link amino acids to form specific proteins? This was the problem to which Watson alluded as the king of Sweden, lost in a fog of admiration, nodded amiably.

The success of Watson-Crick base pairing had persuaded a number of molecular biologists that DNA undertook protein synthesis by the same process—the formation of symmetrical patterns or "templates"—that governed its replication. After all, molecular replication proceeded by the divinely simple separation-and-recombination of matching (or symmetrical) molecules, with each strand of DNA serving as the template for another.

So it seemed altogether plausible that DNA would likewise serve a template function for the amino acids.

It was Francis Crick who in 1957 first observed that this was most unlikely. In a note circulated privately, Crick wrote that "if one considers the physico-chemical nature of the amino-acid side chains, we do not find complementary features on the nucleic acids. Where are the knobby hydrophobic . . . surfaces to distinguish valine from leucine and isoleucine? Where are the charged groups, in specific positions, to go with acidic and basic amino acids?"

Should anyone have missed his point, Crick made it again: "I don't think that anyone looking at DNA or RNA [ribonucleic acid] would think of them as templates for amino acids."

Had these observations been made by anyone but Francis Crick, they might have been regarded as the work of a lunatic; but in looking at any textbook in molecular biology today, it is clear that Crick was simply noticing what was under his nose. Just where *are* those "knobby hydrophobic surfaces"? To imagine that the nucleic acids form a template or pattern for the amino acids is a little like trying to imagine a glove fitting over a centipede. But if the nucleic acids did not form a template for the amino acids, then the information they contained—all of the ancient wisdom of the species, after all—could only be expressed by an indirect form of transmission: a *code* of some sort.

The idea was hardly new. The physicist Erwin Schrödinger had predicted in 1945 that living systems would contain what he called a "code script"; and his short, elegant book, *What Is Life?*, had exerted a compelling influence on every molecular biologist who read it. Ten years later, the ubiquitous Crick invoked the phrase "sequence hypothesis" to characterize the double idea that DNA sequences spell a message *and* that a code is required to express it. What remained obscure was both the spelling of the message and the mechanism by which it was conveyed.

The mechanism emerged first. During the late 1950's, François Jacob and Jacques Monod advanced the thesis that RNA acts as the first in a chain of intermediates leading from DNA to the amino acids.

Single- rather than double-stranded, RNA is a nucleic acid: a chip from the original DNA block. Instead of thymine (T), it contains the base uracil (U), and the sugar that it employs along its backbone features an atom of oxygen missing from deoxyribose. But RNA, Jacob and Monod argued, was more than a mere molecule: it was a messenger, an instrument of conveyance, "transcribing" in one medium a message first expressed in another. Among the many forms of RNA loitering in the modern cell, the RNA bound for duties of transcription became known, for obvious reasons, as "messenger" RNA.

In transcription, molecular biologists had discovered a second fundamental process, a companion in arms to replication. Almost immediately thereafter, details of the code employed by the messenger appeared. In 1961, Marshall Nirenberg and J. Heinrich

Matthaei announced that they had discovered a specific point of contact between RNA and the amino acids. And then, in short order, the full genetic code emerged. RNA (like DNA) is organized into triplets, so that adjacent sequences of three bases are mapped to a single amino acid. Sixty-four triplets (or codons) govern twenty amino acids. The scheme is universal, or almost so.

The elaboration of the genetic code made possible a remarkably elegant model of the modern cell as a system in which sequences of codons within the nucleic acids act at a distance to determine sequences of amino acids within the proteins: commands issued, responses undertaken. A third fundamental biological process thus acquired molecular incarnation. If replication served to divide and then to duplicate the cell's ancestral message, and transcription to re-express it in messenger RNA, "translation" acted to convey that message from messenger RNA to the amino acids.

For all the boldness and power of this thesis, the details remained on the level of what bookkeepers call general accounting procedures. No one had established a direct—a *physical*—connection between RNA and the amino acids.

Having noted the problem, Crick also indicated the shape of its solution. "I therefore proposed a theory," he would write retrospectively, "in which there were twenty adaptors (one for each amino acid), together with twenty special enzymes. Each enzyme would join one particular amino acid to its own special adaptor."

In early 1969, at roughly the same time that a somber Lyndon Johnson was departing the White House to return to the Pedernales, the adaptors whose existence Crick had predicted came into view. There were twenty, just as he had suggested. They were short in length; they were specific in their action; and they were nucleic acids. Collectively, they are now designated "transfer" RNA (tRNA).

Folded like a cloverleaf, transfer RNA serves physically as a bridge between messenger RNA and an amino acid. One arm of the cloverleaf is called the anti-coding region. The three nucleotide bases that it contains are curved around the arm's bulb-end; they are matched by Watson-Crick base pairing to bases on the messenger RNA. The other end of the cloverleaf is an acceptor region. It is here that an amino acid must go, with the structure of tRNA suggesting a complicated female socket waiting to be charged by an appropriate male amino acid.

The adaptors whose existence Crick had predicted served dramatically to confirm his hypothesis that such adaptors were needed. But although they brought about a physical connection between the nucleic and the amino acids, the fact that they were themselves nucleic acids raised a question: in the unfolding molecular chain, just what acted to adapt the adaptors to the amino acids? And this, too, was a problem Crick both envisaged and solved: his original suggestion mentioned both adaptors (nucleic acids) and their *enzymes* (proteins).

And so again it proved. The act of matching adaptors to amino acids is carried out by a

family of enzymes, and thus by a family of proteins: the aminoacyl-tRNA synthetases. There are as many such enzymes as there are adaptors. The prefix “aminoacyl” indicates a class of chemical reactions, and it is in aminoacylation that the cargo of a carboxyl group is bonded to a molecule of transfer RNA.

Collectively, the enzymes known as synthetases have the power both to recognize specific codons and to select their appropriate amino acid under the universal genetic code. Recognition and selection are ordinarily thought to be cognitive acts. In psychology, they are poorly understood, but within the cell they have been accounted for in chemical terms and so in terms of “the model for what science should be.”

With tRNA appropriately charged, the molecule is conveyed to the ribosome, where the task of assembling sequences of amino acids is then undertaken by still another nucleic acid, ribosomal RNA (rRNA). By these means, the modern cell is at last subordinated to a rich narrative drama. To repeat:

Replication duplicates the genetic message in DNA.

Transcription copies the genetic message from DNA to RNA.

Translation conveys the genetic message from RNA to the amino acids—whereupon, in a fourth and final step, the amino acids are assembled into proteins.

The Central Dogma

It was once again Francis Crick, with his remarkable gift for impressing his authority over an entire discipline, who elaborated these facts into what he called the central dogma of molecular biology. The cell, Crick affirmed, is a divided kingdom. Acting as the cell’s administrators, the nucleic acids embody all of the requisite wisdom—where to go, what to do, how to manage—in the specific sequence of their nucleotide bases. Administration then proceeds by the transmission of information *from* the nucleic acids *to* the proteins.

The central dogma thus depicts an arrow moving one way, from the nucleic acids to the proteins, and never the other way around. But is anything ever routinely returned, arrow-like, from its target? This is not a question that Crick considered, although in one sense the answer is plainly no. Given the modern genetic code, which maps four nucleotides onto twenty amino acids, there can be no inverse code going in the opposite direction; an inverse mapping is mathematically impossible.

But there is another sense in which Crick’s central dogma does engender its own reversal. If the nucleic acids are the cell’s administrators, the proteins are its chemical executives: both the staff and the stuff of life. The molecular arrow goes one way with respect to information, but it goes the other way with respect to chemistry.

Replication, transcription, and translation represent the grand unfolding of the central dogma as it proceeds in one direction. The chemical activities initiated by the enzymes

represent the grand unfolding of the central dogma as it goes in the other. Within the cell, the two halves of the central dogma combine to reveal a *system of coded chemistry*, an exquisitely intricate but remarkably coherent temporal tableau suggesting a great army in action.

From these considerations a familiar figure now emerges: the figure of a chicken and its egg. Replication, transcription, and translation are all under the control of various enzymes. But enzymes are proteins, and these particular proteins are specified by the cell's nucleic acids. DNA requires the enzymes in order to undertake the work of replication, transcription, and translation; the enzymes require DNA in order to initiate it. The nucleic acids and the proteins are thus profoundly coordinated, each depending upon the other. Without amino-acyl-tRNA synthetase, there is no translation from RNA; but without DNA, there is no synthesis of aminoacyl-tRNA synthetase.

If the nucleic acids and their enzymes simply chased each other forever around the same cell, the result would be a vicious circle. But life has elegantly resolved the circle in the form of a spiral. The aminoacyl-tRNA synthetase that is required to complete molecular translation enters a given cell from its progenitor or "maternal" cell, where it is specified by that cell's DNA. The enzymes required to make the maternal cell's DNA do its work enter that cell from *its* maternal line. And so forth.

On the level of intuition and experience, these facts suggest nothing more mysterious than the longstanding truism that life comes only from life. *Omnia viva ex vivo*, as Latin writers said. It is only when they are embedded in various theories about the *origins* of life that the facts engender a paradox, or at least a question: in the receding molecular spiral, which came first—the chicken in the form of DNA, or its egg in the form of various proteins? And if neither came first, how could life have begun?

The RNA World

It is 1967, the year of the Six-Day war in the Middle East, the discovery of the electroweak forces in particle physics, and the completion of a twenty-year research program devoted to the effects of fluoridation on dental caries in Evanston, Illinois. It is also the year in which Carl Woese, Leslie Orgel, and Francis Crick introduced the hypothesis that "evolution based on RNA replication *preceded* the appearance of protein synthesis" (emphasis added).

By this time, it had become abundantly clear that the structure of the modern cell was not only more complex than other physical structures but complex in poorly understood ways. And yet no matter how far back biologists traveled into the tunnel of time, certain features of the modern cell were still there, a message sent into the future by the last universal common ancestor. Summarizing his own perplexity in retrospect, Crick would later observe that "an honest man, armed with all the knowledge available to us now, could only state that, in some sense, the origin of life appears at the moment to be almost a miracle." Very wisely, Crick would thereupon determine never to write another paper on the subject—although he did affirm his commitment to the theory of "directed

panspermia," according to which life originated in some other portion of the universe and, for reasons that Crick could never specify, was simply sent here.

But that was later. In 1967, the argument presented by Woesel, Orgel, and Crick was simple. Given those chickens and their eggs, *something* must have come first. Two possibilities were struck off by a process of elimination. DNA? Too stable and, in some odd sense, too perfect. The proteins? Incapable of dividing themselves, and so, like molecular eunuchs, useful without being fecund. That left RNA. While it was not obviously the right choice for a primordial molecule, it was not obviously the wrong choice, either.

The hypothesis having been advanced—if with no very great sense of intellectual confidence—biologists differed in its interpretation. But they did concur on three general principles. First: that at some time in the distant past, RNA rather than DNA controlled genetic replication. Second: that Watson-Crick base pairing governed ancestral RNA. And third: that RNA once carried on chemical activities of the sort that are now entrusted to the proteins. The paradox of the chicken and the egg was thus resolved by the hypothesis that the chicken *was* the egg.

The independent discovery in 1981 of the ribozyme—a ribonucleic enzyme—by Thomas Cech and Sidney Altman endowed the RNA hypothesis with the force of a scientific conjecture. Studying the ciliated protozoan *Tetrahymena thermophila*, Cech discovered to his astonishment a form of RNA capable of inducing cleavage. Where an enzyme might have been busy pulling a strand of RNA apart, there was a ribozyme doing the work instead. That busy little molecule served not only to give instructions: apparently it took them as well, and in any case it did what biochemists had since the 1920's assumed could only be done by an enzyme and hence by a protein.

In 1986, the biochemist Walter Gilbert was moved to assert the existence of an entire RNA "world," an ancestral state promoted by the magic of this designation to what a great many biologists would affirm as fact. Thus, when the molecular biologist Harry Noller discovered that protein synthesis within the contemporary ribosome is catalyzed by ribosomal RNA (rRNA), and not by any of the familiar, old-fashioned enzymes, it appeared "almost certain" to Leslie Orgel that "there once was an RNA world" (emphasis added).

From Molecular Biology to the Origins of Life

It is perfectly true that every part of the modern cell carries some faint traces of the past. But these molecular traces are only hints. By contrast, to everyone who has studied it, the ribozyme has appeared to be an authentic relic, a solid and palpable souvenir from the pre-biotic past. Its discovery prompted even Francis Crick to the admission that he, too, wished he had been clever enough to look for such relics before they became known.

Thanks to the ribozyme, a great many scientists have become convinced that the "model for what science should be" is achingly close to encompassing the origins of life itself.

“My expectation,” remarks David Liu, professor of chemistry and chemical biology at Harvard, “is that we will be able to reduce this to a very simple series of logical events.” Although often overstated, this optimism is by no means irrational. Looking at the modern cell, biologists propose to reconstruct in time the structures that are now plainly there in space.

Research into the origins of life has thus been subordinated to a rational three-part sequence, beginning in the very distant past. First, the constituents of the cell were formed and assembled. These included the nucleotide bases, the amino acids, and the sugars. There followed next the emergence of the ribozyme, endowed somehow with powers of self-replication. With the stage set, a system of coded chemistry then emerged, making possible what the molecular biologist Paul Schimmel has called “the theater of the proteins.” Thus did matters proceed from the pre-biotic past to the very threshold of the last universal common ancestor, whereupon, with inimitable gusto, life began to diversify itself by means of Darwinian principles.

This account is no longer fantasy. But it is not yet fact. That is one reason why retracing its steps is such an interesting exercise, to which we now turn.

Miller Time

It is perhaps four billion years ago. The first of the great eras in the formation of life has commenced. The laws of chemistry are completely in control of things—what else is there? It is Miller Time, the period marking the transition from inorganic to organic chemistry.

According to the impression generally conveyed in both the popular and the scientific literature, the success of the original Miller-Urey experiment was both absolute and unqualified. This, however, is something of an exaggeration. Shortly after Miller and Urey published their results, a number of experienced geochemists expressed reservations. Miller and Urey had assumed that the pre-biotic atmosphere was one in which hydrogen atoms gave up (reduced) their electrons in order to promote chemical activity. Not so, the geochemists contended. The pre-biotic atmosphere was far more nearly neutral than reductive, with little or no methane and a good deal of carbon dioxide.

Nothing in the intervening years has suggested that these sour geochemists were far wrong. Writing in the 1999 issue of *Peptides*, B.M. Rode observed blandly that “modern geochemistry assumes that the secondary atmosphere of the primitive earth (i.e., after diffusion of hydrogen and helium into space) . . . consisted mainly of carbon dioxide, nitrogen, water, sulfur dioxide, and even small amounts of oxygen.” This is not an environment calculated to induce excitement.

Until recently, the chemically unforthcoming nature of the early atmosphere remained an embarrassing secret among evolutionary biologists, like an uncle known privately to dress in women’s underwear; if biologists were disposed in public to acknowledge the facts, they did so by remarking that every family has one. This has now changed. The

issue has come to seem troubling. A recent paper in *Science* has suggested that previous conjectures about the pre-biotic atmosphere were seriously in error. A few researchers have argued that a reducing atmosphere is not, after all, quite so important to pre-biotic synthesis as previously imagined.

In all this, Miller himself has maintained a far more unyielding and honest perspective. “Either you have a reducing atmosphere,” he has written bluntly, “or you’re not going to have the organic compounds required for life.”

If the composition of the pre-biotic atmosphere remains a matter of controversy, this can hardly be considered surprising: geochemists are attempting to revisit an era that lies four billion years in the past. The synthesis of pre-biotic chemicals is another matter. Questions about them come under the discipline of laboratory experiments.

Among the questions is one concerning the nitrogenous base cytosine (C). Not a trace of the stuff has been found in any meteor. Nothing in comets, either, so far as anyone can tell. It is not buried in the Antarctic. Nor can it be produced by any of the common experiments in pre-biotic chemistry. Beyond the living cell, it has not been found at all.

When, therefore, M.P. Robertson and Stanley Miller announced in *Nature* in 1995 that they had specified a plausible route for the pre-biotic synthesis of cytosine from cyanoacetaldehyde and urea, the feeling of gratification was very considerable. But it has also been short-lived. In a lengthy and influential review published in the 1999 *Proceedings of the National Academy of Science*, the New York University chemist Robert Shapiro observed that the reaction on which Robertson and Miller had pinned their hopes, although active enough, ultimately went nowhere. All too quickly, the cytosine that they had synthesized transformed itself into the RNA base uracil (U) by a chemical reaction known as deamination, which is nothing more mysterious than the process of getting rid of one molecule by sending it somewhere else.

The difficulty, as Shapiro wrote, was that “the formation of cytosine and the subsequent deamination of the product to uracil occur[ed] at about the same rate.” Robertson and Miller had themselves reported that after 120 hours, half of their precious cytosine was gone—and it went faster when their reactions took place in saturated urea. In Shapiro’s words, “It is clear that the yield of cytosine would fall to 0 percent if the reaction were extended.”

If the central chemical reaction favored by Robertson and Miller was self-defeating, it was also contingent on circumstances that were unlikely. *Concentrated* urea was needed to prompt their reaction; an outhouse whiff would not do. For this same reason, however, the pre-biotic sea, where concentrates disappear too quickly, was hardly the place to begin—as anyone who has safely relieved himself in a swimming pool might confirm with guilty satisfaction. Aware of this, Robertson and Miller posited a different set of circumstances: in place of the pre-biotic soup, drying lagoons. In a fine polemical passage, their critic Shapiro stipulated what would thereby be required:

An isolated lagoon or other body of seawater would have to undergo extreme concentration. . . .

It would further be necessary that the residual liquid be held in an impermeable vessel [in order to prevent cross-reactions].

The concentration process would have to be interrupted for some decades . . . to allow the reaction to occur.

At this point, the reaction would require quenching (perhaps by evaporation to dryness) to prevent loss by deamination.

At the end, one would have a batch of urea in solid form, containing some cytosine (and urea).

Such a scenario, Shapiro remarked, "cannot be excluded as a rare event on early earth, but it cannot be termed plausible."

Like cytosine, sugar must also make an appearance in Miller Time, and, like cytosine, it too is difficult to synthesize under plausible pre-biotic conditions.

In 1861, the German chemist Alexander Bulterow created a sugar-like substance from a mixture of formaldehyde and lime. Subsequently refined by a long line of organic chemists, Bulterow's so-called formose reaction has been an inspiration to origins-of-life researchers ever since.

The reaction is today initiated by an alkalizing agent, such as thallium or lead hydroxide. There follows a long induction period, with a number of intermediates bubbling up. The formose reaction is auto-catalytic in the sense that it keeps on going: the carbohydrates that it generates serve to prime the reaction in an exponentially growing feedback loop until the initial stock of formaldehyde is exhausted. With the induction over, the formose reaction yields a number of complex sugars.

Nonetheless, it is not sugars in general that are wanted from Miller Time but a particular form of sugar, namely, ribose—and not simply ribose but dextro ribose. Compounds of carbon are naturally right-handed or left-handed, depending on how they polarize light. The ribose in living systems is right-handed, hence the prefix "dextro." But the sugars exiting the formose reaction are racemic, that is, both left- and right-handed, and the yield of usable ribose is negligible.

While nothing has as yet changed the fundamental fact that it is very hard to get the right kind of sugar from any sort of experiment, in 1990 the Swiss chemist Albert Eschenmoser was able to change substantially the way in which the sugars appeared. Reaching with the hand of a master into the formose reaction itself, Eschenmoser altered two molecules by adding a phosphate group to them. This slight change prevented the formation of the alien sugars that cluttered the classical formose reaction. The products, Eschenmoser reported, included among other things a mixture of ribose-2,4,-diphosphate. Although the mixture was racemic, it did contain a molecule close to the ribose needed by living systems. With a few chemical adjustments, Eschenmoser could plausibly claim, the pre-biotic route to the synthesis of sugar would lie open.

It remained for skeptics to observe that Eschenmoser's ribose reactions were critically contingent on Eschenmoser himself, and at two points: the first when he attached phosphate groups to a number of intermediates in the formose reaction, and the second when he removed them.

What had given the original Miller-Urey experiment its power to excite the imagination was the sense that, having set the stage, Miller and Urey exited the theater. By contrast, Eschenmoser remained at center stage, giving directions and in general proving himself indispensable to the whole scene.

Events occurring in Miller Time would thus appear to depend on the large assumption, still unproved, that the early atmosphere was reductive, while two of the era's chemical triumphs, cytosine and sugar, remain for the moment beyond the powers of contemporary pre-biotic chemistry.

From Miller Time to Self-Replicating RNA

In the grand progression by which life arose from inorganic matter, Miller Time has been concluded. It is now 3.8 billion years ago. The chemical precursors to life have been formed. A limpid pool of nucleotides is somewhere in existence. A new era is about to commence.

The historical task assigned to this era is a double one: forming chains of nucleic acids from nucleotides, and discovering among them those capable of reproducing themselves. Without the first, there is no RNA; and without the second, there is no life.

In living systems, polymerization or chain-formation proceeds by means of the cell's invaluable enzymes. But in the grim inhospitable pre-biotic, no enzymes were available. And so chemists have assigned their task to various inorganic catalysts. J.P. Ferris and G. Ertem, for instance, have reported that activated nucleotides bond covalently when embedded on the surface of montmorillonite, a kind of clay. This example, combining technical complexity with general inconclusiveness, may stand for many others.

In any event, polymerization having been concluded—by whatever means—the result was (in the words of Gerald Joyce and Leslie Orgel) "a random ensemble of polynucleotide sequences": long molecules emerging from short ones, like fronds on the surface of a pond. Among these fronds, nature is said to have discovered a self-replicating molecule. But how?

Darwinian evolution is plainly unavailing in this exercise or that era, since Darwinian evolution *begins* with self-replication, and self-replication is precisely what needs to be explained. But if Darwinian evolution is unavailing, so, too, is chemistry. The fronds comprise "a *random* ensemble of polynucleotide sequences" (emphasis added); but no principle of organic chemistry suggests that aimless encounters among nucleic acids must lead to a chain capable of self-replication.

If chemistry is unavailing and Darwin indisposed, what is left as a mechanism? The evolutionary biologist's finest friend: sheer dumb luck.

Was nature lucky? It depends on the payoff and the odds. The payoff is clear: an ancestral form of RNA capable of replication. Without that payoff, there is no life, and obviously, at some point, the payoff paid off. The question is the odds.

For the moment, no one knows how precisely to compute those odds, if only because within the laboratory, no one has conducted an experiment leading to a self-replicating ribozyme. But the minimum length or "sequence" that is needed for a contemporary ribozyme to undertake what the distinguished geochemist Gustaf Arrhenius calls "demonstrated ligase activity" *is* known. It is roughly 100 nucleotides.

Whereupon, just as one might expect, things blow up very quickly. As Arrhenius notes, there are 4^{100} or roughly 10^{60} nucleotide sequences that are 100 nucleotides in length. This is an unfathomably large number. It exceeds the number of atoms contained in the universe, as well as the age of the universe in seconds. If the odds in favor of self-replication are 1 in 10^{60}, no betting man would take them, no matter how attractive the payoff, and neither presumably would nature.

"Solace from the tyranny of nucleotide combinatorials," Arrhenius remarks in discussing this very point, "is sought in the feeling that strict sequence specificity may not be required through all the domains of a functional oligmer, thus making a large number of library items eligible for participation in the construction of the ultimate functional entity." Allow me to translate: why assume that self-replicating sequences are apt to be rare just because they are long? They might have been quite common.

They might well have been. And yet all experience is against it. Why should self-replicating RNA molecules have been common 3.6 billion years ago when they are impossible to discern under laboratory conditions today? No one, for that matter, has ever seen a ribozyme capable of *any* form of catalytic action that is not very specific in its sequence and thus unlike even closely related sequences. No one has ever seen a ribozyme able to undertake chemical action without a suite of enzymes in attendance. No one has ever seen anything like it.

The odds, then, are daunting; and when considered realistically, they are even worse than this already alarming account might suggest. The discovery of a single molecule with the power to initiate replication would hardly be sufficient to establish replication. What template would it replicate *against*? We need, in other words, at least two, causing the odds of their joint discovery to increase from 1 in 10^{60} to 1 in 10^{120}. Those two sequences would have been needed in roughly the same place. And at the same time. And organized in such a way as to favor base pairing. And somehow held in place. And buffered against competing reactions. And productive enough so that their duplicates would not at once vanish in the soundless sea.

In contemplating the discovery by chance of two RNA sequences a mere 40 nucleotides

in length, Joyce and Orgel concluded that the requisite "library" would require 10^{48} possible sequences. Given the weight of RNA, they observed gloomily, the relevant sample space would exceed the mass of the earth. And this is the same Leslie Orgel, it will be remembered, who observed that "it was almost certain that there once was an RNA world."

To the accumulating agenda of assumptions, then, let us add two more: that without enzymes, nucleotides were somehow formed into chains, and that by means we cannot duplicate in the laboratory, a pre-biotic molecule discovered how to reproduce itself.

From Self-Replicating RNA to Coded Chemistry

A new era is now in prospect, one that begins with a self-replicating form of RNA and ends with the system of coded chemistry characteristic of the modern cell. The *modern* cell—meaning one that divides its labors by assigning to the nucleic acids the management of information and to the proteins the execution of chemical activity. It is 3.6 billion years ago.

It is with the advent of this era that distinctively conceptual problems emerge. The gods of chemistry may now be seen receding into the distance. The cell's system of coded chemistry is determined by two discrete combinatorial objects: the nucleic acids and the amino acids. These objects are discrete because, just as there are no fractional sentences containing three-and-a-half words, there are no fractional nucleotide sequences containing three-and-a-half nucleotides, or fractional proteins containing three-and-a-half amino acids. They are combinatorial because both the nucleic acids and the amino acids are combined by the cell into larger structures.

But if information management and its administration within the modern cell are determined by a discrete combinatorial system, the *work* of the cell is part of a markedly different enterprise. The periodic table notwithstanding, chemical reactions are not combinatorial, and they are not discrete. The chemical bond, as Linus Pauling demonstrated in the 1930's, is based squarely on quantum mechanics. And to the extent that chemistry is explained in terms of physics, it is encompassed not only by "the model for what science should be" but by the system of differential equations that play so conspicuous a role in every one of the great theories of mathematical physics.

What serves to coordinate the cell's two big shots of information management and chemical activity, and so to coordinate two fundamentally different structures, is the universal genetic code. To capture the remarkable nature of the facts in play here, it is useful to stress the word *code*.

By itself, a code is familiar enough: an arbitrary mapping or a system of linkages between two discrete combinatorial objects. The Morse code, to take a familiar example, coordinates dashes and dots with letters of the alphabet. To note that codes are arbitrary is to note the distinction between a code and a purely physical connection between two objects. To note that codes embody mappings is to embed the concept of a code in

mathematical language. To note that codes reflect a linkage of some sort is to return the concept of a code to its human uses.

In every normal circumstance, the linkage comes first and represents a human achievement, something arising from a point beyond the coding system. (The coordination of dot-dot-dot-dash-dash-dash-dot-dot-dot with the distress signal S-O-S is again a familiar example.) Just as no word explains its own meaning, no code establishes its own nature.

The conceptual question now follows. Can the origins of a system of coded chemistry be explained in a way that makes no appeal whatsoever to the kinds of facts that we otherwise invoke to explain codes and languages, systems of communication, the impress of ordinary words on the world of matter?

In this regard, it is worth recalling that, as Hubert Yockey observes in *Information Theory, Evolution, and the Origin of Life* (2005), "there is no trace in physics or chemistry of the control of chemical reactions by a sequence of any sort or of a code between sequences."

Writing in the 2001 issue of the journal RNA, the microbiologist Carl Woese referred ominously to the "dark side of molecular biology." DNA replication, Woese wrote, is the extraordinarily elegant expression of the structural properties of a single molecule: zip down, divide, zip up. The transcription into RNA follows suit: copy and conserve. In each of these two cases, structure leads to function. But where is the coordinating link between the chemical structure of DNA and the third step, namely, translation? When it comes to translation, the apparatus is baroque: it is incredibly elaborate, and it does not reflect the structure of any molecule.

These reflections prompted Woese to a somber conclusion: if "the nucleic acids cannot in any way recognize the amino acids," then there is no "fundamental *physical* principle" at work in translation (emphasis added).

But Woese's diagnosis of disorder is far too partial; the symptoms he regards as singular are in fact widespread. What holds for translation holds as well for replication and transcription. The nucleic acids cannot directly recognize the amino acids (and vice versa), but they cannot *directly* replicate or transcribe themselves, either. Both replication and translation are enzymatically driven, and without those enzymes, a molecule of DNA or RNA would do nothing whatsoever. Contrary to what Woese imagines, no fundamental physical principles appear directly at work *anywhere* in the modern cell.

The most difficult and challenging problem associated with the origins of life is now in view. One half of the modern system of coded chemistry—the genetic code and the sequences it conveys—is, from a chemical perspective, arbitrary. The other half of the system of coded chemistry—the activity of the proteins—is, from a chemical perspective, necessary. In life, the two halves are coordinated. The problem follows: how did *that*—the whole system—get here?

The prevailing opinion among molecular biologists is that questions about molecular-biological systems can only be answered by molecular-biological *experiments*. The distinguished molecular biologist Horoaki Suga has recently demonstrated the strengths and the limitations of the experimental method when confronted by difficult conceptual questions like the one I have just posed.

The goal of Suga's experiment was to show that a set of RNA catalysts (or ribozymes) *could* well have played the role now played in the modern cell by the protein family of aminoacyl synthetases. Until his work, Suga reports, there had been no convincing demonstration that a ribozyme was able to perform the double function of a synthetase—that is, recognizing both a form of transfer RNA and an amino acid. But in Suga's laboratory, just such a molecule made a now-celebrated appearance. With an amino acid attached to its tail, the ribozyme managed to cleave itself and, like a snake, affix its amino-acid cargo onto its head. What is more, it could conduct this exercise backward, shifting the amino acid from its head to its tail again. The chemical reactions involved acylation: precisely the reactions undertaken by synthetases in the modern cell.

Horoaki Suga's experiment was both interesting and ingenious, prompting a reaction perhaps best expressed as, "Well, would you look at that!" It has altered the terms of debate by placing a number of new facts on the table. And yet, as so often happens in experimental pre-biotic chemistry, it is by no means clear what interpretation the facts will sustain.

Do Suga's results really establish the existence of a primitive form of coded chemistry? Although unexpected in context, the coordination he achieved between an amino acid and a form of transfer RNA was never at issue in principle. The question is whether what was accomplished in establishing a chemical connection between these two molecules was anything like establishing the existence of a *code*. If so, then organic chemistry itself could properly be described as the study of codes, thereby erasing the meaning of a code as an arbitrary mapping between discrete combinatorial objects.

Suga, in summarizing the results of his research, captures rhetorically the inconclusiveness of his achievement. "Our demonstration indicates," he writes, "that catalytic precursor tRNA's *could have provided* the foundation of the genetic coding system." But if the association at issue is not a code, however primitive, it could no more be the "foundation" of a code than a feather could be the foundation of a building. And if it is the foundation of a code, then what has been accomplished has been accomplished by the wrong agent.

In Suga's experiment, there was no sign that the execution of chemical routines fell under the control of a molecular administration, and no sign, either, that the missing molecular administration had anything to do with executive chemical routines. The missing molecular administrator was, in fact, Suga himself, as his own account reveals. The relevant features of the experiment, he writes, "allow[ed] *us* to select active RNA molecules with selectivity toward a *desired* amino acid" (emphasis added). Thereafter, it

was Suga and his collaborators who "applied *stringent* conditions" to the experiment, undertook "*selective amplification* of the self-modifying RNA molecules," and "*screened*" vigorously for "self-aminoacylation activity" (emphasis added throughout).

If nothing else, the advent of a system of coded chemistry satisfied the most urgent of imperatives: it was needed and it was found. It was needed because once a system of chemical reactions reaches a certain threshold of complexity, nothing less than a system of coded chemistry can possibly master the ensuing chaos. It was found because, after all, we are here.

Precisely these circumstances have persuaded many molecular biologists that the explanation for the emergence of a system of coded chemistry must in the end lie with Darwin's theory of evolution. As one critic has observed in commenting on Suga's experiments, "If a certain result can be achieved by direction in a laboratory by a Suga, surely it can also be achieved by chance in a vast universe."

A self-replicating ribozyme meets the first condition required for Darwinian evolution to gain purchase. It is by definition capable of replication. And it meets the second condition as well, for, by means of mistakes in replication, it introduces the possibility of variety into the biological world. On the assumption that subsequent changes to the system follow a law of increasing marginal utility, one can then envisage the eventual emergence of a system of coded chemistry—a system that can be explained in terms of "the model for what science should be."

It was no doubt out of considerations like these that, in coming up against what he called the "dark side of molecular biology," Carl Woese was concerned to urge upon the biological community the benefits of "an all-out Darwinian perspective." But the difficulty with "an all-out Darwinian perspective" is that it entails an all-out Darwinian impediment: notably, the assignment of a degree of foresight to a Darwinian process that the process could not possibly possess.

The hypothesis of an RNA world trades brilliantly on the idea that a divided modern system had its roots in some form of molecular symmetry that was then broken by the contingencies of life. At some point in the transition to the modern system, an ancestral form of RNA must have assigned some of its catalytic properties to an emerging family of proteins. This would have taken place at a given historical moment; it is not an artifact of the imagination. Similarly, at some point in the transition to a modern system, an ancestral form of RNA must have acquired the ability to code for the catalytic powers it was discarding. And this, too, must have taken place at a particular historical moment.

The question, of course, is which of the two steps came first. Without life acquiring some degree of foresight, neither step can be plausibly fixed in place by means of any schedule of selective advantages. How could an ancestral form of RNA have acquired the ability to code for various amino acids before coding was useful? But then again, why should "ribozymes in an RNA world," as the molecular biologists Paul Schimmel and Shana O. Kelley ask, "have expedited their own obsolescence?"

Could the two steps have taken place simultaneously? If so, there would appear to be very little difference between a Darwinian explanation and the frank admission that a miracle was at work. If no miracles are at work, we are returned to the place from which we started, with the chicken-and-egg pattern that is visible when life is traced backward now appearing when it is traced forward.

It is thus unsurprising that writings embodying Woese's "all-out Darwinian perspective" are dominated by references to a number of unspecified but mysteriously potent forces and obscure conditional circumstances. I quote without attribution because the citations are almost generic (emphasis added throughout):

• The aminoacylation of RNA initially *must* have provided some selective advantage.

• The products of this reaction *must* have conferred some selective advantage.

• However, the development of a crude mechanism for controlling the diversity of possible peptides *would* have been advantageous.

• [P]rogressive refinement of that mechanism *would* have provided *further* selective advantage.

And so forth—ending, one imagines, in reduction to the all-purpose imperative of Darwinian theory, which is simply that what was must have been.

Now It Is Now

At the conclusion of a long essay, it is customary to summarize what has been learned. In the present case, I suspect it would be more prudent to recall how much has been *assumed*:

First, that the pre-biotic atmosphere was chemically reductive; second, that nature found a way to synthesize cytosine; third, that nature also found a way to synthesize ribose; fourth, that nature found the means to assemble nucleotides into polynucleotides; fifth, that nature discovered a self-replicating molecule; and sixth, that having done all that, nature promoted a self-replicating molecule into a full system of coded chemistry.

These assumptions are not only vexing but progressively so, ending in a serious impediment to thought. That, indeed, may be why a number of biologists have lately reported a weakening of their commitment to the RNA world altogether, and a desire to look elsewhere for an explanation of the emergence of life on earth. "It's part of a quiet paradigm revolution going on in biology," the biophysicist Harold Morowitz put it in an interview in *New Scientist*, "in which the radical randomness of Darwinism is being replaced by a much more scientific law-regulated emergence of life."

Morowitz is not a man inclined to wait for the details to accumulate before reorganizing the vista of modern biology. In a series of articles, he has argued for a global vision based

on the biochemistry of living systems rather than on their molecular biology or on Darwinian adaptations. His vision treats the living system as more fundamental than its particular species, claiming to represent the "universal and deterministic features of *any* system of chemical interactions based on a water-covered but rocky planet such as ours."

This view of things—metabolism first, as it is often called—is not only intriguing in itself but is enhanced by a firm commitment to chemistry and to "the model for what science should be." It has been argued with great vigor by Morowitz and others. It represents an alternative to the RNA world. It is a work in progress, and it may well be right. Nonetheless, it suffers from one outstanding defect. There is as yet no evidence that it is true.

It is now more than 175 years since Friedrich Wöhler announced the synthesis of urea. It would be the height of folly to doubt that our understanding of life's origins has been immeasurably improved. But whether it has been immeasurably improved in a way that vigorously confirms the daring idea that living systems are chemical in their origin and so physical in their nature—that is another question entirely.

In "On the Origins of the Mind," I tried to show that much can be learned by studying the issue from a computational perspective. Analogously, in contemplating the origins of life, much—in fact, more—can be learned by studying the issue from the perspective of coded chemistry. In both cases, however, what seems to lie beyond the reach of "the model for what science should be" is any success beyond the local. All questions about the global origins of these strange and baffling systems seem to demand answers that the model itself cannot by its nature provide.

It goes without saying that this is a tentative judgment, perhaps only a hunch. But let us suppose that questions about the origins of the mind and the origins of life do lie beyond the grasp of "the model for what science should be." In that case, we must either content ourselves with its limitations or revise the model. If a revision also lies beyond our powers, then we may well have to say that the mind and life have appeared in the universe for no very good reason that we can discern.

Worse things have happened. In the end, these are matters that can only be resolved in the way that all such questions are resolved. We must wait and see.

[*]I used this phrase, borrowed from the mathematicians J.H. Hubbard and B.H. West, in "On the Origins of the Mind" (Commentary, November 2004). The idea that science must conform to a certain model of inquiry is familiar. Hubbard and West identify that model with differential equations, the canonical instruments throughout physics and chemistry.

David Berlinski, *the author most recently of* Infinite Ascent: A Short History of Mathematics *(Modern Library), is a senior fellow of the Discovery Institute. The present essay, the second in a three-part series, follows "On the Origins of the Mind" (November 2004), now anthologized in* The Best American Science Writing 2005, *edited by Alan*

Lightman (Harper Perennial). A final installment, on the origins of matter, is scheduled for later this year. Mr. Berlinski would like to thank Ed Peltzer, Gustaf Arrhenius, Jonathan Wells, Arthur Cody, Tyler Hampton, and Morris Salkoff for their helpful comments on early drafts of this essay.