

# Defining Life

Steven A. Benner

## Abstract

Any definition is intricately connected to a theory that gives it meaning. Accordingly, this article discusses various definitions of life held in the astrobiology community by considering their connected “theories of life.” These include certain “list” definitions and a popular definition that holds that life is a “self-sustaining chemical system capable of Darwinian evolution.” We then act as “anthropologists,” studying what scientists do to determine which definition-theories of life they constructively hold as they design missions to seek non-terran life. We also look at how constructive beliefs about biosignatures change as observational data accumulate. And we consider how a definition centered on Darwinian evolution might itself be forced to change as supra-Darwinian species emerge, including in our descendents, and consider the chances of our encountering supra-Darwinian species in our exploration of the Cosmos. Last, we ask what chemical structures might support Darwinian evolution universally; these structures might be universal biosignatures. Key Words: Evolution—Life—Life detection—Biosignatures. *Astrobiology* 10, 1021–1030.

## 1. Introduction

**A**STROBIOLOGISTS ARE COMMITTED to studying life in the Cosmos, the terran life we know as well as the extra-terran life we do not know but hope to encounter. But what exactly do we seek?

The question is hardly new, nor is the recognition of its difficulty. Also not new is a certain imprecision in the language used to address this question and therefore an imprecision in the consequent ideas.

Daniel Koshland recently provided an anecdote that illustrates this imprecision (Koshland, 2002). As president of the American Association for the Advancement of Science (which publishes the prestigious journal *Science*), Koshland recounted his own experience with a committee that was charged to generate a definition of life:

What is the definition of life? I remember a conference of the scientific elite that sought to answer that question. Is an enzyme alive? Is a virus alive? Is a cell alive? After many hours of launching promising balloons that defined life in a sentence, followed by equally conclusive punctures of these balloons, a solution seemed at hand: “The ability to reproduce—that is the essential characteristic of life” said one statesman of science. Everyone nodded in agreement that the essentials of a life was the ability to reproduce, until one small voice was heard. “Then one rabbit is dead. Two rabbits—a male and female—are alive but either one alone is dead.” At that point, we all became convinced that although everyone knows what life is, there is no simple definition of life.

The imprecise use of language is manifest. The “elite” have confused the concept of “being alive” with the concept of “life.” This is not simply the mistaking an adjective for a noun. Rather, it represents the conflation of a *part* of a system with its *whole*. Parts of a living system might themselves be alive (a cell in our finger may be “alive,” as might a fertilized ovum *in utero*). But those living parts need not be coextensive with a living system and need not represent life. Using language precisely, one rabbit may be alive even though he or she is not life.

Koshland, distinguished as an experimental biochemist, dug the hole a bit deeper as he tried to manage the problem that his committee had created. He set out an alternative definition of “life” that listed items that he thought were needed to give life the features that he valued. He called this his PICERAS definition of life, where PICERAS was an acronym for “program, improvisation, compartmentalization, energy, regeneration, adaptability, and seclusion.” Unfortunately, a single rabbit remains dead if the listed elements must all be present, while the mated pair remains alive. Also, while the list captures the thermodynamic, genetic, physiological, metabolic, and cellular features of the terran life that we know, it offers little by way of theory to argue that these features must be universal in the life that we do not know.

Other committees have come up with different definitions for life that attempt to offer such a theory. For example, Joyce summarized the discussion of a committee assembled in

1994 by NASA to consider the possibility of life in the Cosmos (Joyce, 1994). Following a suggestion by Carl Sagan, the committee proposed that life is a “self-sustaining chemical system capable of Darwinian evolution.”

Through their use of the word “system,” the committee intended to recognize that entities can be alive (a cell, virus, or a single rabbit) without themselves individually exemplifying life. They used the phrase “self-sustaining” to imply that a living system should not need continuous intervention by a higher entity (a graduate student or a god, for example) to continue as “life.” They exploited the phrase “Darwinian evolution” as a shorthand for a process, elaborated over the past 150 years, that involves a molecular genetic system (DNA in terran life) that can be replicated imperfectly, where mistakes arising from imperfect replication can themselves be replicated, and where various replicates have different “fitnesses.”

## 2. A Definition Embodies a Theory

The “NASA definition” of life avoids some of the simplest counterexamples that often defeat “list definitions” for life. The requirement for reproduction with errors, where the errors are themselves reproducible, excludes a variety of non-living chemical systems that can reproduce.

For example, a crystal of sodium chlorate ( $\text{NaClO}_3$ ) can be powdered and used to seed the growth of other sodium chlorate crystals (Kondepudi *et al.*, 1990). Therefore, we might say that the original crystal can reproduce through this process. Further, features of the original crystal, such as its chirality (sodium chlorate crystals can either be “right handed” or “left handed”) can be passed to its descendants via this process.

The replication is imperfect, as a real sodium chlorate crystal contains many defects. Indeed, to specify *all* the defects in any real crystal of sodium chlorate would require an enormous amount of information, easily exceeding the 10 billion bits of information contained in a human genome.

But the information in these defects is *not* itself inheritable via a powdering-then-seeding process. The defects in the parent crystal of sodium chlorate cannot be passed to the descendent crystals via this process. Therefore, the sodium chlorate system cannot support Darwinian evolution, even if various descendent crystals with different sets of defects might be differentially “fit.” Thus, the system cannot deliver improved descendents via Darwinian processes. As such processes are viewed under the theory behind this definition as the only way matter can become organized to give the properties that we value in living systems, a system of sodium chlorate crystals is not life.

The NASA definition also avoids other popular counterexamples of list definitions. One from the popular series *Star Trek* (life “absorbs compounds from its environment,” “excretes waste,” and “grows”) was challenged by the android Data, who pointed out that fire consumes “food,” excretes “waste,” metabolizes, moves, and grows, but is not life. But fire is not capable of Darwinian evolution. Its growth may be imperfect, but those imperfections are not heritable.

Indeed, the NASA definition avoids problems encountered with most other definitions of life. Those familiar with thermodynamics often attempt to distinguish “fire” from “life” by observing that fire simply dissipates available free energy. In this view, living systems are different because

they use free energy to produce order as part of their dissipative process.

But what is order, and within what bounds? After all, life on any planet almost certainly cannot violate the second law of thermodynamics; it is also dissipative if its “within” and its “without” are considered together. And those who attempt to distinguish “life” from “fire” by pointing to the structure within living systems are promptly refuted by those who note that fire (whirlwinds) and hurricanes *also* generate order within as they dissipate free energy without.

## 3. The “NASA Definition” Also Captures a Theory of Life

The NASA definition of life has another virtue: it conveys information about what its framers believed to be possible ways to generate features we value in life. By doing so, the NASA definition captures a *theory* of life.

As discussed by Cleland and Chyba (2002), definitions are inseparable from the theories that give them meaning. Conversely, a definition-theory gives motive and context to experimental work to “test” the definition by testing the theory. In this way, studies of life done in terran laboratories can be relevant to non-terran life we cannot access (Benner, 2009).

The value of a theory associated with a definition is illustrated by examples of life that are conceivable but are excluded by the definition, where the exclusion is justified by a constructive belief that such life-forms are not possible. Many of these can be extracted from popular culture.

For example, the crew of *Star Trek (The Next Generation)* has encountered conceptual aliens that do not fit the NASA definition. The nanites that infected the computer of the next-generation Enterprise in Episode 50 (“Evolution”) are informational (or perhaps electromechanical, but in any case not chemical); their evolution is not tied to an informational molecule like DNA (although they require a chemical matrix to survive). The Crystalline Entity of Episode 18 (“Home Soil”) appears to be chemical but not obviously Darwinian; it seems to have no children. The Calamarain (Episode 51: “Déjà Q”) are made of pure energy, not chemicals. And the sentient being known as Q (Episode 1: “Encounter at Farpoint,” and others) appears to be neither matter nor energy, flitting instead in and out of the Continuum without the apparent need of either.

Off screen, other fictional forms of life appear to defy the NASA definition. For example, Fred Hoyle published a story *The Black Cloud*, a fictional entity that floats into our Solar System and blocks our sunlight, placing Earth in distress. After the black cloud realizes that the Earth holds self-aware forms of life, it politely moves out of the way and apologizes.

If we were to encounter Q, the Calamarain, or any of these other conjectural entities during a real, not conceptual, trek through the stars, we would be forced to concede that they *do* represent living systems, because they have the attributes that we value in living systems. We would also be forced to agree that they do not fall within the NASA definition of life. We would therefore be forced to agree that we need a new definition for life. If a black cloud were to float into our Solar System and begin to apologize to us, we would certainly reject any definition that does not include it as life, especially

if we want to encourage particularly polite life-forms (Benner, 2009).

We do not *now* change our definition of life so as to accommodate Q, the black cloud, or other examples of “weird” life from Hollywood because *we do not constructively believe that this kind of life is possible*.

A brief comment on the word “constructive” is in order. A constructive belief is one that we actually act upon. For example, many people say that they believe global warming will raise sea levels in a few years. Those who constructively believe this do not buy homes in Miami Beach. Those who do buy homes in Miami Beach do not constructively believe that the sea level will rise, even if they say they do.

With respect to Q or the black cloud, no one is proposing an early warning system to detect such entities, as disruptive as they would be if they were to appear. We do not devote resources to protect ourselves from transdimensional beings or black clouds because we constructively do not believe that they can exist.

Likewise, a group of beings who have achieved perfect immortality may no longer need to reproduce and may therefore lose their ability to reproduce. These, according to our definition-theory, have lost their status as living systems. We do not now change our definition of life to include such beings because we do not constructively believe that perfect immortality is possible.

Biosignatures can also be incorporated into this view. According to our definition-theory of life, nanites and androids (Data of *Star Trek*, Marvin of *Hitchhiker's Guide to the Galaxy*) are examples of artificial life. We do not doubt that Darwinian evolution can be simulated in a computer. We do not doubt that androids can be created, including androids (such as Data) that (who?) wish to be human. Our definition-theory of life, however, excludes the possibility that computers, their viruses, or androids could have arisen without a creator that had already emerged by Darwinian process.

Instead, our definition-theory regards these as *bio-signatures* (evidence that life exists or existed), not life itself. Likewise, no matter how intelligent these android-robots are, our definition-theory of life requires us to regard their intelligence as *artificial*.

Following similar reasoning under our definition-theory, the computer in which nanites reside is not life but is evidence of a life-form that created it. In this view, the computer is a biosignature and the nanites are an “artificial” life-form. Any intelligence that either displays would be “artificial.” They both are derived from a self-sustaining chemical system capable of Darwinian evolution, which must have created them.

#### 4. Is Humankind Life?

Certainly, Darwinian theory holds that humankind attained its present forms via events where natural selection was superimposed upon random variation. By “random,” Darwinian theory requires that the variation must not be biased with respect to future outcome. The genetic mutations that we pass to our children in their DNA cannot have arisen in anticipation of their ability to make our children fitter, even if they actually end up doing so.

For all their power to create life in the world that we know, Darwinian processes have some well-understood

disadvantages. For example, they condemn some of our children to die of genetic diseases in order to “allow” others among our children to adapt. For every mutation that allows some children to be bigger, better, and smarter, Darwinian processes require dozens of other mutations that make some children sick. Death from genetic maladaptation inherently goes with adaptation.

However, according to at least one model, *Homo habilis* learned, a few million years ago, to make tools (Huang *et al.*, 1995). These tools conferred fitness that could, in a sense, be inherited, but in a new way. While the details are still sketchy, it appeared that *Homo habilis* had acquired a genetic change that allowed its children to be taught. *Homo habilis* moms and dads could point, unlike chimpanzee parents. Their children looked where their parents pointed and understood that it was important to understand what their parents were pointing at (at least until the age of 16). This converted an older trait of simple imitation (monkey see, monkey do) into the ability to educate and be educated. This, in turn, allowed extragenetic fitness to be passed from parents to children not by DNA but through education.

Is this process Darwinian? Is the variation in the structure of manufactured tools (not inherited genes) random with respect to future outcome? Perhaps in human history some tool development may have been random, allowing the emergence of better tools to be broadly called Darwinian. In this model, our ancestors chipped rocks randomly. Those rocks that accidentally killed mastodons helped the ancestors survive and thus were perpetuated through teaching. Chipped rocks that did not kill mastodons did not, and therefore were not.

But at least some tool development in *Homo* is prospective. Some tool variants *are* constructed with foreknowledge of a desirable outcome. Certainly today, modern engineers think that they do not generate solutions to problems in a Darwinian fashion. They tell us that they perceive a problem and then build tools prospectively to solve the problem. They then teach those tools to their students.

Thus improvements in fitness that come via designed variation are also not Darwinian. Novel variation is introduced prospectively with foresight of its potential value.

Therefore, *Homo sapiens* as a species is already stretching our definition-theory of life. We will stretch and confound that definition still more when technology emerges that allows the inheritance of designed fitness in ways other than education.

For example, technology may soon be available to identify DNA sequences that prospectively help our children survive better, marry better, and have better children. We may soon gain the technology that allows our pediatrician to place those DNA sequences into our eggs and sperm, creating mutant children that are fitter by design. If this happens, then our species will escape Darwinian mechanisms for improving our genes. Our species will have become supra-Darwinian.

A good news–bad news dialectic relates to this. The good news is that supra-Darwinian species need not see children die of genetic disease as a price to pay for its continued existence. The bad news is that we do not know today how to beneficially change the sequence of our DNA and may not be smart enough to learn before technology enables us to drive ourselves to extinction.

## 5. Supra-Darwinian Life

Setting aside the desirability of goal-driven evolution, what should we do with a definition-theory of life that features “Darwinian evolution” so prominently, given the manifest fact that our species is on the verge of improving itself using quite different processes? Must we now modify that definition-theory to read “A self-sustaining chemical system capable of Darwinian or supra-Darwinian evolution”?

Here, we cannot escape the dilemma by claiming that we constructively believe that supra-Darwinian evolution is impossible. On the contrary, many laboratories are working today to develop gene therapy technology to make it so. The elected representatives of the taxpayers are funding this work. We constructively believe that that supra-Darwinian evolution *is* possible

To save our definition-theory, we might notice that even as we are happily becoming cerebral beings by prospectively altering our personal DNA by design, we still are *capable of Darwinian evolution*. Last, we might argue that, like an intelligent android, we could not have come into being had our ancestors not first had access to Darwinian evolution.

We are crossing into uncharted philosophical territory here. Philosophers like definitions of natural kinds to consider only the object being defined, not its history. We will have a hard time persuading serious philosophers to accept a system as “life” if, some time in the past (but no longer), it had access to Darwinian evolution.

For astrobiology, we might minimize the dilemma by arguing that supra-Darwinian life is rare. After all, stars like our Sun live for only about 10 billion years. Biology on Earth consumed about half that time to get from the origin of Earth to a form of life that is considering surpassing the constraints of Darwinian evolution. Thus, only around a relatively small proportion of the stars in the Cosmos might orbit supra-Darwinian beings.

Why did it take so long for supra-Darwinian life to emerge on Earth? One hypothesis is that educable toolmakers are hard to get via Darwinian processes. According to this hypothesis, 4.5 billion years were needed to have enough random mutations for genes to arise that allow parents to point at tools and children learn as a consequence. Generalizing this to the Cosmos as a whole, supra-Darwinian systems might be expected to be rare.

Alternatively, supra-Darwinian life may have been slow to emerge on Earth because environments where tool-making was useful generally come along only infrequently. After all, in a static environment, Darwinian evolution prefers specialists, species that do just one thing and do it well. A tool-making species is a generalist, and generalists have advantages over specialists only when the environment is diverse and, especially, changing. Tools and education are necessary in a rapidly changing environment, as DNA cannot keep up without having so many mutations that create too much genetic disease.

At least this is so with *Homo sapiens*. Because we make tools, we can survive in a huge variety of environments, more than any other species. Further, we can survive when our environment changes rapidly, far better than if we needed to rely on Darwinian mechanisms to adapt.

Dramatically changing environments are rare in the history of Earth. Yet, recently, after the onset of the ice ages,

they became the rule. Under this model, the ice ages selected for human toolmakers by creating an environment where a generalist lifestyle based on toolmaking was fitter than a specialist lifestyle based on conventional (but slow) optimization of DNA sequences. In this view, without the ice ages, humans and the type of intelligence valued by humans would likely not have emerged, even on Earth. And if the exceptionality of ice age-like climate change on a planet is universal, then supra-Darwinian life should also be exceptional universally.

With these thoughts in mind, we can ask whether the life that we are most likely to encounter in the Cosmos has evolved supra-Darwinian approaches to fitness? This would seem to depend on whether we find that life or whether that life finds us.

Consider an *a fortiori* argument. *Homo sapiens* is now much closer to doing directed genetic therapy than it is to doing interstellar travel. Extracting a universal law from this, we might propose that, in the natural progression of biological evolution, species learn how to prevent their children from dying of genetic diseases (and how to drive their own evolution) before they learn to be interstellar space travelers.

This *a fortiori* argument has some practical consequences. If you have been abducted by an alien from another star system, your abductor was probably not “life” according to the NASA definition-theory. The technology that he, she, or it used to come to Earth to abduct was, *a fortiori*, developed after technology that allowed his, her, or its DNA to be prospectively engineered by his, her, or its parents. That engineering might even have been done specifically to allow him, her, or it to travel between stars.

However, if *we* encounter life in our Solar System, it is not likely to be as intelligent as we are (as it has not yet encountered us). Thus, the life that *we* find will not yet have developed either space travel or gene therapy. Thus, the NASA definition-theory of life is likely to be useful.

## 6. The Utility of the NASA Definition-Theory of Life

The practical purpose of a definition-theory of life is to guide our design of missions to the Cosmos to find life. For this purpose, the definition-theory drives us to adopt two propositions: (a) Every system that we encounter that produces behaviors we value from life will prove to be a chemical system capable of Darwinian evolution. (b) Any chemical system that is capable of Darwinian evolution is capable of producing the behaviors we value from life.

These propositions reflect the possibility that, at some time in the future, we may find life that occupies places in the Cosmos where the terran life that we know could not possibly survive, even in highly mutated form. For example, we constructively do not believe that terran biochemistry, even after an arbitrarily large amount of adaptation, could support life at 300°C (572 K) in sulfuric acid clouds above Venus or in liquid methane at –179°C (94 K) in the oceans of Titan. Chemistry will not allow it.

But the NASA definition-theory of life says: never mind. If something exists in those environments that has the attributes that we value in life, then it will be a chemical system capable of Darwinian evolution. We must simply ask *what kinds* of chemistry can be performed in those environments to support Darwinian evolution.

A commitment to a clearly stated definition-theory focuses subsequent efforts to answer universal questions. This is scientific pragmatism. Let us illustrate this in the context of recent experience searching for life by way of NASA and ESA missions.

In 1976, the Viking mission delivered three explicit life-detection tests duplicated in two separate landers to the surface of Mars (Levin and Straat, 1977, 1981). One experiment added martian soil to a solution in water of seven organic compounds labeled with carbon-14. Life in the martian soil was expected to oxidize these compounds to release radioactive carbon dioxide.

A second life-detection test also placed a sample of martian soil into a nutrient broth, but with a different detection scheme. If life were present and capable of doing photosynthesis, molecular oxygen should emerge from the broth. Production of oxygen was to be interpreted as a positive sign of martian life.

A third life-detection test presented radioactive carbon dioxide and carbon monoxide to the martian surface in the presence of sunlight. If life were present on the martian surface having the capability to “fix” carbon from the atmosphere (like terran plants), radiolabel from these gasses should be fixed into organic compounds on the soil. Observation of carbon fixation was to be interpreted as a positive sign of martian life. Further, since both gasses are present naturally in the martian atmosphere, this was viewed as an especially relevant test for life.

The Viking scientists did not clearly state their definition-theory of life. But the design of these life-detection tests speaks clearly to what definition-theory they constructively held. As “anthropologists” of science (Fig. 1), we can infer from their actions what beliefs those scientists constructively held. As all their life-detection tests looked for products from

metabolism, the Viking mission designers evidently placed “metabolism” high on their list of criteria for life.

The life-detection tests also corresponded closely to metabolism known in terran life. If you were to consume the seven radioactive organic molecules in the second life-detection tool and then exhale into the Viking detector, that detector would conclude that you are alive because it would detect radioactive CO<sub>2</sub> in your breath.

The other Viking tests for life were reminiscent of terran photosynthesis. On Earth, if you put water and nutrients on terran soil and expose that soil to sunlight, photosynthetic organisms in the soil will emit oxygen. If you expose a terran plant to radioactive carbon dioxide and give it sunlight, the plant will fix the CO<sub>2</sub> into organic compounds in the plant, which will become radioactive.

So what happened on the surface of Mars when the three life-detection experiments were run? Without dwelling on the details, all the life-detection experiments gave positive results. Radiolabeled CO<sub>2</sub> was released when the seven radiolabeled organic compounds were added in water to the martian soil. Release was not observed if the martian soil had been heated, as expected if the heat had killed something living in the soil. Molecular oxygen was released when water was added to martian soil, just as if the soil contained photosynthetic organisms. Radioactive carbon was fixed in the martian soil when exposed to light and radiolabeled carbon monoxide and dioxide, consistent with the presence of photosynthetic organisms.

The results could have been more definitive. The presence of life would have been more clearly indicated if more radioactive carbon had been fixed from the atmosphere, not the small amounts that were observed. The pattern of oxygen release was a bit perplexing upon heating and cooling. Nevertheless, Gilbert Levin, who designed the labeled release instrument, still argues that his test detected life on Mars. And why not? The results observed were the results that the designers of the experiment intended to interpret as positive signs for life.

Nevertheless, scientists concluded that the martian surface held no life. Why? Briefly, the positive indicators of life were set aside because of results from another pair of instruments delivered to Mars, a gas chromatograph (GC) and a mass spectrometer (MS). These were used in tandem. First, a sample of martian soil was delivered to a cup. The cup was sealed and heated. Vapors emerging from the heated martian soil were blown by a stream of dihydrogen gas into the GC column, where they were separated. Then, the separated compounds were injected into the MS to determine their presence and masses (Biemann *et al.*, 1977).

Surprisingly, given the results of the life-detection test, the GC-MS instruments did not detect any organic molecules at all other than what the instruments had brought with them from Earth. This result drove the community to conclude that the martian surface contained no life at all, even though all three experiments designed under a definition-theory of life valuing metabolism had produced positive signs of life.

This says something about the definition-theory of life that these scientists constructively had: a *reduced-carbon* definition-theory of life. That definition-theory placed a higher value on organic composition of a sample than on whatever metabolism it displayed. As anthropologists, we can say that the scientists interpreting the Viking results constructively



FIG. 1. Directly studying as “anthropologists of science” what scientists do allows us to understand what scientists constructively believe. By permission from Benner (2009). Color images available online at [www.liebertonline.com/ast](http://www.liebertonline.com/ast).

felt that the reduced-carbon theory of life trumped the metabolism theory of life.

### 7. The Cell-Based Definition-Theory of Life Trumps the Carbon Theory of Life

The Viking experiment provides a fascinating look into the minds of practicing scientists who use definitions set within the context of theory to guide their actions. In one sense, they were operating with “list” definitions of life, where the first set of experiments had a list consisting of one item: “metabolism.”

Their view was, of course, tightly tied to a theory that life could not possibly exist without metabolism. Leaving aside the fact that the *particular* metabolisms sought were terran-inspired, the theory of life behind these assays is almost certainly universal. One feature that we value from living systems is their ability to exploit free energy to transform matter “without” into matter “within.” Since metabolism is simply the collection of those transformations, our view of life would need to change dramatically were we to include life without metabolism.

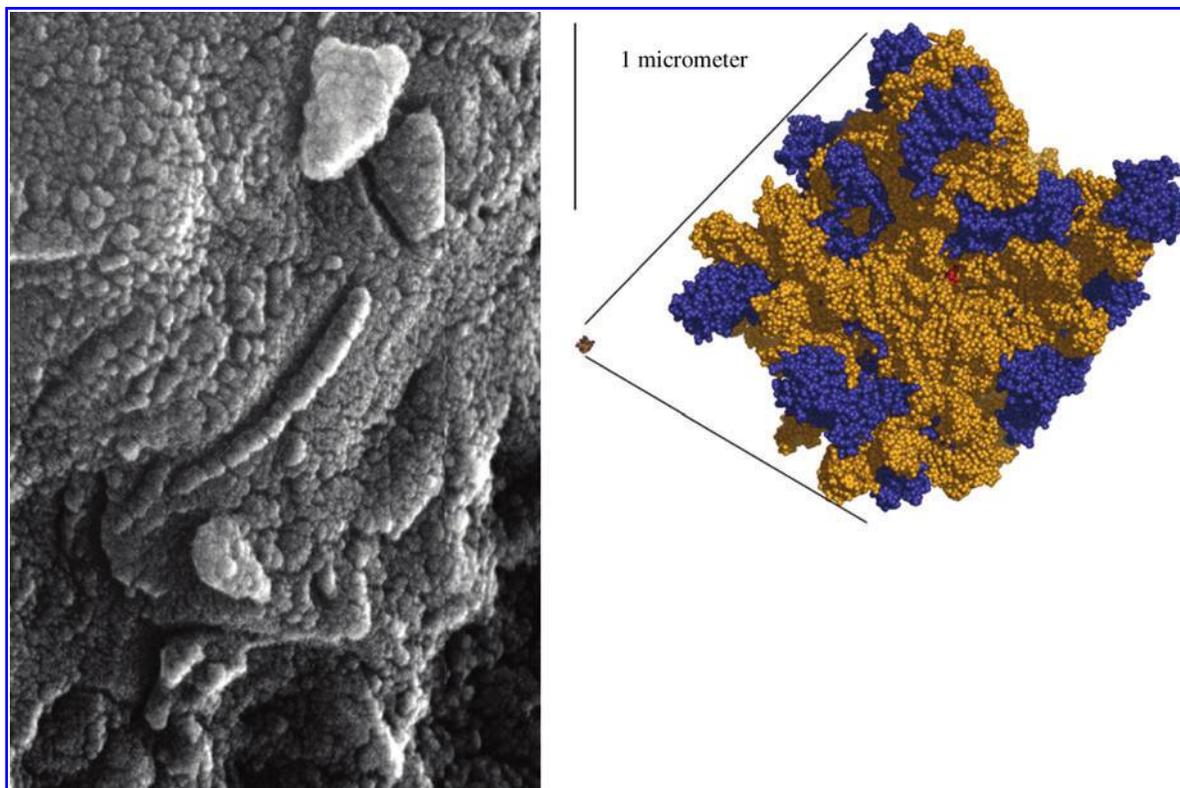
This view, which is rooted deeply in thermodynamic theory, is stronger than the view that this metabolism must be based on reduced carbon species. We can certainly conceive of metabolism without organic carbon. Indeed, it is not entirely correct to call even terran life “carbon based.” If terran biomolecules were simply alkanes (that is, made of nothing but carbon and hydrogen), they would not support life on Earth. The interesting parts of terran organic bio-

molecules, the parts that permit metabolism at terran temperatures, are the parts that include non-carbon atoms, including oxygen, nitrogen, sulfur, and phosphorus.

Notwithstanding, we as anthropologists can infer that the “reduced-carbon” theory of life constructively trumped the “metabolism” theory of life in the minds of the scientists who interpreted the Viking results. In part, their subordination of the metabolism theory of life was rational, as non-living systems may also interconvert compounds. Indeed, as soon as the failure of the GC-MS to detect reduced carbon was announced, proposals were presented to explain how carbon dioxide would emerge from martian soils, how dioxygen would be released from martian soils, and how carbon would be fixed on martian soils, all without a chemical system capable of Darwinian evolution. These explanations focused on the possibility that ultraviolet light near the surface might create conditions for reactivity, including inorganic oxidizing power (Fenton and Jones, 1900).

As a consequence of this collection of constructive views, “non-life” explanations for the Viking results came to be dominant within the community. For the following 20 years, most scientists viewed the martian surface as highly oxidizing and devoid of life.

This changed when David McKay and his colleagues at the Johnson Space Center focused on a meteorite that had been collected from the top of an ice field in Antarctica near the Allan Hills, ALH 84001. McKay *et al.* published images of small, cell-like structures in the Allan Hills 84001 meteorite (Fig. 2). These, he suggested, could be the remnants of microbial life on Mars (McKay *et al.*, 1996).



**FIG. 2.** A scaled comparison of the structure of the ribosome with the structures observed in the Allan Hills meteorite from Mars suggested that the structures were too small to be life. This suggestion assumed that any martian life must make proteins by using terran-sized ribosomes. Color images available online at [www.liebertonline.com/ast](http://www.liebertonline.com/ast).

McKay's suggestion drew upon an old theory of life, "cell theory." Cell theory finds its origin in observations by Robert Hooke (1635–1703), who used a microscope to observe cells in slices of cork. Later studies showed that animal tissues were also made from cells. In 1847, Theodor Schwann and M.J. Schleiden suggested that animal and vegetable biology could be unified under a theory that all living systems are built from cells.

Because of McKay's observations, Mars immediately came alive (at least constructively in the minds of the community). To us as anthropologists of science, this shows that the cell definition-theory of life must trump the reduced carbon definition-theory. Consistent with this view, many list definitions of life capture this feature in multiple ways. For example, both the C (compartmentalization) and the S (seclusion) components of the PICERAS definition of life proposed by Daniel Koshland are closely related to cell theory.

You can decide for yourself whether the structures discovered in the Allan Hills meteorite look like cells from living systems (Fig. 2). But as the decade progressed, some in the community began to argue against those structures being biosignatures based on their small size. The structures are only 100 nanometers across. This was "too small," according to some commentaries, to be the signature of life.

But too small for what? The most frequently cited "too small" argument compared the size of the ALH "cells" to the ribosome, the molecular machine used by terran life to make proteins (Fig. 2). The ribosome is approximately 25 nanometers across. This means that the "cells" in Allan Hills 84001 can hold only four ribosomes across. Many argued that this was too few for a viable cell.

Again, anthropologists of science can ask what this argument implies about the definition-theories for life held by those who advance it. Those who advanced this argument evidently had a different theory of life: a protein theory. The syllogism is clear. Proteins are necessary for life. Ribosomes are necessary to make proteins. Martian cells are too small to hold ribosomes. Therefore, martian cells are too small to make proteins. Therefore, martian cells are too small to be life.

But what is the argument that proteins are universally necessary components of life? Could martian life not have had proteins?

Curiously, in another corner of modern astrobiology lies a model for an ancient form of life on Earth that did *not* use proteins. This model, advanced by Alex Rich a half century ago (Rich, 1962), suggested that the first forms of life used RNA to do both genetics and catalysis. The hypothesis that life on Earth began as an RNA world solves, at least to some, the "chicken-or-egg" problem associated with the origin of life. In modern life, proteins are needed to make DNA, while DNA is needed to make proteins. DNA and proteins certainly could not have emerged spontaneously from inanimate matter, it is thought. But if RNA, now the intermediate in modern terran life between DNA and proteins, originally played the roles of both DNA and proteins, life could have originated by the (still perplexing) abiotic emergence of RNA.

Under the RNA world model for life on early Earth, life is possible without proteins. Indeed, this model claims that we are all descendants of a form of life that lived without proteins and sustained Darwinian evolution.

This makes it possible to dismiss the "small size" argument. What if the structures observed in the Allan Hills

meteorite were remnants of life from an RNA world on Mars, the same RNA world that was proposed to have existed on early Earth? Such a form of life would not need ribosomes. It therefore need not be large enough to hold ribosomes (Benner, 1999).

## 8. What Kind of Polymers Will Support Darwinian Evolution?

These examples of practical attempts to detect evidence for life on Mars illustrate the importance of the question: What is our definition-theory of life? It also opens a way to investigate the definition-theory that holds that life is a self-sustaining chemical system capable of Darwinian evolution. Here, we start with a model that places a biopolymer at the center of the Darwinian processes. We can ask: what kind of biopolymer can be copied, with mutation, where the mutations can themselves be copied?

Not many, it turns out. In fact, the demands of Darwinian evolution place some unusual chemical constraints on any biopolymeric system that might support it. That biopolymer must be capable of changing its structure and, therefore, changing its encoded information without changing its bulk properties, in particular, properties that are essential for its replication.

Very few organic molecular systems meet this constraint, a fact well known to synthetic organic chemists. Synthetic efforts on natural products are, for example, often preceded by exploratory work with a "model system," a simpler molecule that is more accessible than the natural product itself but that is presumed to represent the kinds of problems that will be encountered when the natural product itself is synthesized. As often as not, reactivities of the model and the real natural product are sufficiently different that chemistry developed on the first is defeated by the second. Changing a single methyl group in a molecular system can be sufficient to alter its physical properties dramatically.

The same sensitive connection between structure and behavior is seen with proteins. A single amino acid replacement, for example, converts normal hemoglobin into sickle cell hemoglobin, a mutant protein with very different physical properties and reactivities than the native protein.

In this respect, DNA (and, to a lesser extent, RNA) is unusual. Changing the sequence of a DNA molecule generally does not change the overall physical properties of the molecule or its general reactivity. This feature of DNA allows DNA to support Darwinian evolution.

This feature arises because of a particular structural feature of DNA: the phosphates that link the nucleoside building blocks together in the DNA backbone each carry a negative charge (Benner and Hutter, 2002). This makes DNA a *polyelectrolyte*, a molecule with multiple charges. The repeating backbone charge dominates the properties of the DNA molecule so much that changing one of the uncharged nucleobases (and thereby changing the information encoded by the DNA molecule) scarcely alters the physical behavior of the DNA molecule.

The repeating backbone charge helps DNA support Darwinian evolution in other ways (Fig. 3). For example:

- (a) The repeating backbone charge keeps RNA and DNA dissolved in water.
- (b) The repeating backbone charge forces interaction between strands to occur as far from the backbone as

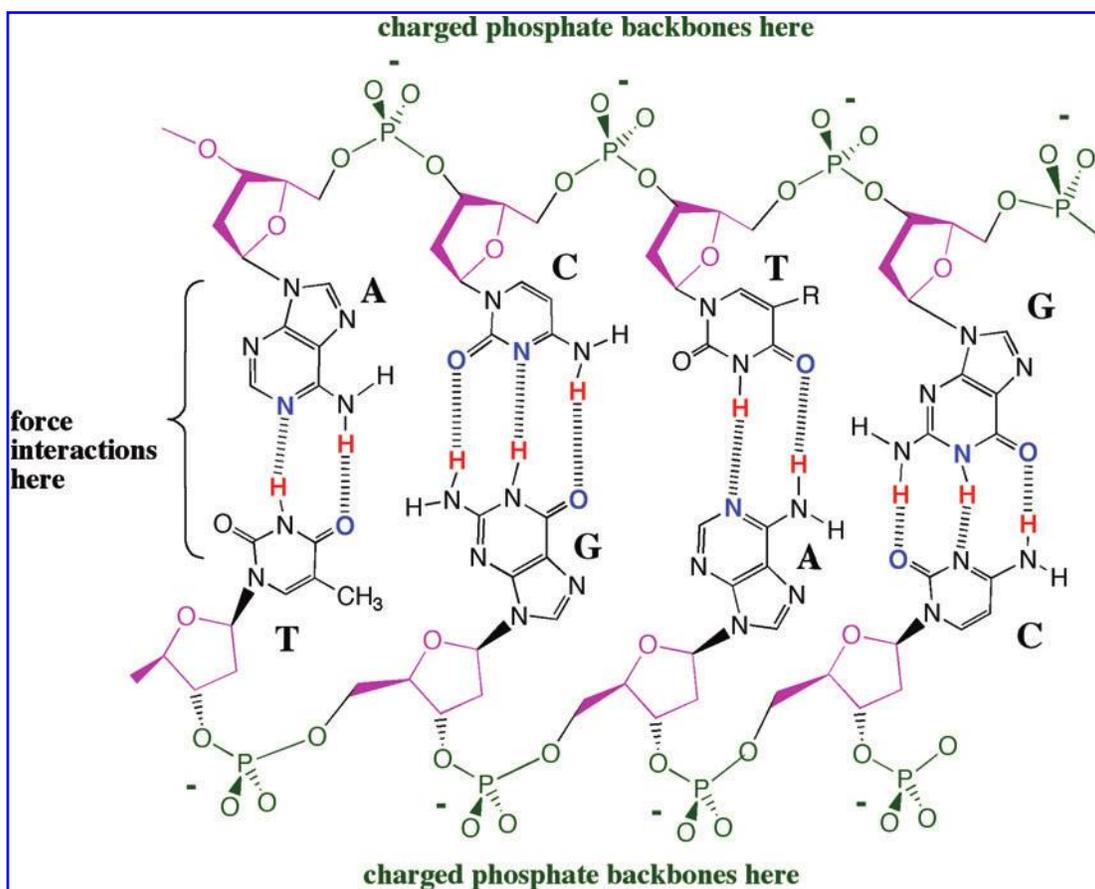


FIG. 3. The repeating charge in the backbones of nucleic acids is hypothesized to be key to allowing DNA and RNA to support Darwinian evolution. It helps force strand-strand interactions as far from the backbone as possible (the origin of Watson-Crick pairing rules). Further, the repeating charge so dominates the biophysical properties of the molecular system that changing a nucleobase (a mutation) does not have any significant impact on the behavior of the molecule, a feature essential to allow the system to evolve. Color images available online at [www.liebertonline.com/ast](http://www.liebertonline.com/ast).

possible, as the backbone charges from one strand repel the backbone charges from the other (Fig. 3). The Watson-Crick interactions essential to the ability of DNA to replicate arise because of interstrand interactions far from the backbone.

- (c) The repeating backbone charge keeps the RNA and DNA molecules from folding, allowing them to act as templates.

Based on these arguments, we hypothesized some time ago that this particular structural feature (a repeating charge) was general to all life in water, universally, given a definition-theory of life that requires Darwinian evolution (Benner and Hutter, 2002). This is called the “polyelectrolyte theory of the gene.”

### 9. Operational Tests for Life

Some have objected to the NASA definition-theory of life because it does not seem usefully “operational” (Cleland and Chyba, 2002). After all, defining life based on its ability to evolve in a Darwinian sense would seem to imply that a useful life-detection instrument must perturb the alien environment and then wait long enough to observe something actually evolving.

While this objection is well taken, it ignores other ways that the NASA definition-theory of life can guide our search for non-terran life. Here, we need not construct and fly an instrument hoping to actually observe Darwinian evolution. Rather, we can construct and fly an instrument designed to observe *molecular structures necessary* for Darwinian evolution.

The polyelectrolyte theory of the gene illustrates how one might do this. The polyelectrolyte theory identifies features in the chemical structure of genetic molecules that are necessary to support Darwinian evolution universally. Thus, rather than waiting around for an alien glob to actually evolve, we can look within it for molecules having those features. As chemistry is (we presume) universal, and as the feature that we have hypothesized to be universal in a gene (a polyelectrolyte, in water) is based on chemical theory, that feature should also be universal.

When that feature is a repeating charge, instruments can be easily designed to detect it. A genetic polyanion with a repeating negative charge in a sample from Mars, for example, will bind specifically and tightly to a polycation in an instrument that we send to Mars. Conversely, a genetic polycation with a repeating positive charge in a sample from Mars will bind specifically and tightly to a polyanion in an instrument that we send to Mars. Binding is a first step to

ward detection. We can even bring the bound martian polyanion or polycation back to Earth for further study, where structural regularities in its building blocks might be dispersive of biology.

Thus, with the polyelectrolyte theory of the gene, we can write a recipe to search for Darwinian chemical systems in the Cosmos. First, we identify places where liquid water is possible, organic materials are available, and free energy is accessible. Then, we design an instrument that carries polycations (to detect polyanionic genetic material) and polyanions (to detect polycationic genetic material). Then we raise the funding to build and fly the instrument.

### 10. Even the Polyelectrolyte Theory of the Gene is Rather Terra-centric

The NASA definition-theory of life can be used to ask what kinds of biopolymers might support Darwinian evolution universally. But who says that Darwinian evolution needs a biopolymer? For example, physicist Freeman Dyson suggested that a form of life might be possible that reproduces without replication (Dyson, 1985). His hypothetical life-form has a “compositional genome,” a collection of molecules that adapts without an encoding biopolymer. The chemist Graham Cairns-Smith has suggested that Darwinian evolution might be supported by minerals (Cairns-Smith, 1982).

The NASA definition-theory for life has scarcely been explored in the context of these alternative “weird” concepts for life. But it does constrain experimental work with them. If a definition-theory of life insists upon Darwinian evolution, then we must ask in the laboratory what kinds of metabolic networks might support Darwinian evolution? Or what kinds of mineral assemblages might do so? And if metabolic cycles or minerals are found in the laboratory to create features in their respective systems that we value in life *without* having a process that falls within our concept of Darwinian processes as it has been developed over the past 150 years, then we will need to alter our definition-theory of life or, perhaps, our concept of Darwinian processes.

### 11. Synthesis as a Tool to Expand our Definition-Theory of Life

This approach to developing a definition-theory of life captures the pragmatism of recent discussions of scientific epistemology (Benner, 2009). This pragmatism suggests yet another approach to developing our understanding of life: synthetic biology.

The goal of synthetic biology is to get, in the laboratory, a chemical system that can support Darwinian evolution. Our activities as synthetic biologists need not be constrained by any particular model for how life might have emerged on Earth. Any system will do, including one based on a biopolymer, or a collection of metabolic processes, or a mineral assembly.

Further, the system need not be self-sustaining; we would be happy if it were able to evolve and adapt, even if it needed continuous attention from a sentient being. Any synthetic molecular system that reproduces with error, if those errors are themselves reproducible, should be able to adapt to environmental changes, at least to the degree that its fundamental molecular capabilities allow.

Various efforts are underway to obtain such systems. These include work with DNA-like molecules that are built from six different nucleotide “letters” (Yang *et al.*, 2010), artificial genetic systems that can be copied, with errors, where those errors are replicable.

Should artificial Darwinian chemical systems be obtained, they present a direct test to the definition-theory of life. They should be able to produce, *in vitro*, features that we value from living systems. Should they fail to do so, they will be analyzed to learn why they fail. This might lead to the identification of chemical features *other* than a polyelectrolyte that are needed to support Darwinian evolution. Or they may challenge the centrality of Darwinian evolution in any theory-definition of life.

### 12. Summary

The dynamic between theory, observation, and definition has been the focus of those who study science and philosophize over the structure of its theories (Suppe, 1977). Many (and perhaps most) practicing scientists are unaware of the “philosophy of science,” which has come to include many who study scientists and their behavior as anthropologists.

Many (and perhaps most) of those scientists who are aware of philosophers on their periphery find their approach not particularly useful. In part, this is undoubtedly because philosophers too often deliver complex, abstruse, and perhaps nihilistic answers to questions that scientists view as concrete.

For example, in preparing this chapter, I consulted the local philosopher of science who sent back the message that “philosophers are weary of satisfactory definitions of any non-trivial term these days.” He continued:

According to the classical philosophical understanding of “definition,” a definition must give both necessary and sufficient conditions, and must do that as a matter of the meaning of the term. For instance, the claim that water equals H<sub>2</sub>O arguably specifies both necessary and sufficient conditions, but it doesn’t do that as a matter of the meaning of the word “water.” The claim is *a posteriori*. A definition, on this classical understanding, must be *a priori*—at least its justification must be *a priori* (because it is supposed to be an analytic claim—true solely in virtue of the meaning of the terms involved). It turns out that, when understood this way, [a definition] is almost impossible to find.

Little wonder that scientists attempting to create life in the laboratory or launch an instrument to Mars find little useful in these comments. We do what we generally do when a reality is too complex to meet our constructive needs: we ignore it and continue with a simpler, if arguably false, view.

For astrobiologists, a need remains for some pragmatic philosophies of science, if only in the training of our youth. This may best come from those who are practicing astrobiologists (Grinspoon, 1998; Benner, 2009), or philosophers closely connected with them (Cleland and Chyba, 2002). I suspect that an understanding dynamic between theory, observation, and definition will be important to these.

### Abbreviations

GC, gas chromatograph; MS, mass spectrometer.

## References

- Benner, S.A. (1999) How small can a microorganism be? In *Size Limits of Very Small Microorganisms: Proceedings of a Workshop*, Steering Group for the Workshop on Size Limits of Very Small Microorganisms, National Research Council, The National Academies Press, Washington, DC, pp 126–135.
- Benner, S.A. (2009) *The Life, the Universe and the Scientific Method*, FfAME Press, Gainesville, FL.
- Benner, S.A. and Hutter, D. (2002) Phosphates, DNA, and the search for nonterrestrial life. A second generation model for genetic molecules. *Bioorg Chem* 30:62–80.
- Biemann, K., Oro, J., Toulmin, P., III, Orgel, L.E., Nier, A.O., Anderson, D.M., Flory, D., Diaz, A.V., Ruchneck, D.R., and Simmonds, P.G. (1977) The search for organic substances and inorganic volatile compounds in the surface of Mars. *J Geophys Res* 82:4641–4658.
- Cairns-Smith, A.G. (1982) *Genetic Takeover and the Mineral Origins of Life*, Cambridge University Press, Cambridge.
- Cleland, C.E. and Chyba, C.F. (2002) Defining “life.” *Orig Life Evol Biosph* 32:387–393.
- Dyson, F. (1985) *Origins of Life*, Cambridge University Press, Cambridge.
- Fenton, H.J.H. and Jones, H.O. (1900) The oxidation of organic acids in the presence of ferrous iron. *J Chem Soc* 77:69–76.
- Grinspoon, D. (1998) *Lonely Planets: The Natural Philosophy of Alien Life*, Harper Collins, New York.
- Huang, W., Ciochon, R., Gu, Y., Larick, R., Fang, Q., Schwarcz, H., Yonge, C., Devos, J., and Rink, W. (1995) Early *Homo* and associated artefacts from Asia. *Nature* 378:275–278.
- Joyce, G.F. (1994) Foreword. In *Origins of Life: The Central Concepts*, D.W. Deamer and G. Fleischaker, Jones and Bartlett, Boston.
- Kondepudi, D.K., Kaufman, R.J., and Singh, N. (1990) Chiral symmetry breaking in sodium chlorate crystallization. *Science* 250:975–976.
- Koshland, D.E. (2002). The seven pillars of life. *Science* 295:2215–2216.
- Levin, G.V. and Straat, P.A. (1977) Recent results from the Viking labeled release experiment on Mars. *J Geophys Res* 82:4663–4667.
- Levin, G.V. and Straat, P.A. (1981) A search for a nonbiological explanation of the Viking labeled release life detection experiment. *Icarus* 45:494–516.
- McKay, D.S., Gibson, E.K., Thomas-Keprta, K.L., Vali, H., Romanek, C.S., Clemett, S.J., Chlilier, X.D.F., Maechling, C.R., and Zare, R.N. (1996) Search for past life on Mars: possible relic biogenic activity in martian meteorite ALH84001. *Science* 273:924–930.
- Rich, A. (1962) On the problems of evolution and biochemical information transfer. In *Horizons In Biochemistry*, edited by M. Kasha and B. Pullmann, Academic Press, New York, pp 103–126.
- Suppe, F., editor. (1977) *The Structure of Scientific Theories*, Illini Books, Champaign-Urbana, IL.
- Yang, Z., Chen, F., Chamberlin, S.G., and Benner, S.A. (2010) Expanded genetic alphabets in the polymerase chain reaction. *Angew Chem Int Ed Engl* 49:177–180.

Address correspondence to:

Steven A. Benner  
 Foundation for Applied Molecular Evolution  
 720 S.W. 2<sup>nd</sup> Avenue  
 Gainesville, FL 32601

E-mail: sbenner@ffame.org

Submitted 20 July 2010  
 Accepted 6 November 2010