

# Chance and necessity do not explain the origin of life

J.T. Trevors<sup>a,\*</sup>, D.L. Abel<sup>b</sup>

<sup>a</sup>Laboratory of Microbial Technology, Department of Environmental Biology, Room 3220, Bovey Building,  
University of Guelph, Guelph, Ontario, Canada, N1G 2W1

<sup>b</sup>The Gene Emergence Project, The Origin-of-Life Foundation Inc., 113 Hedgewood Dr., Greenbelt, MD 20770-1610, USA

Received 8 April 2004; revised 19 May 2004; accepted 24 June 2004

## Abstract

Where and how did the complex genetic instruction set programmed into DNA come into existence? The genetic set may have arisen elsewhere and was transported to the Earth. If not, it arose on the Earth, and became the genetic code in a previous lifeless, physical–chemical world. Even if RNA or DNA were inserted into a lifeless world, they would not contain any genetic instructions unless each nucleotide selection in the sequence was programmed for function. Even then, a predetermined communication system would have had to be in place for any message to be understood at the destination. Transcription and translation would not necessarily have been needed in an RNA world. Ribozymes could have accomplished some of the simpler functions of current protein enzymes. Templating of single RNA strands followed by retemplating back to a sense strand could have occurred. But this process does not explain the derivation of “sense” in any strand. “Sense” means algorithmic function achieved through sequences of certain decision-node switch-settings. These particular primary structures determine secondary and tertiary structures. Each sequence determines minimum-free-energy folding propensities, binding site specificity, and function. Minimal metabolism would be needed for cells to be capable of growth and division. All known metabolism is cybernetic – that is, it is programmatically and algorithmically organized and controlled.

© 2004 International Federation for Cell Biology. Published by Elsevier Ltd. All rights reserved.

**Keywords:** Cellular communication; Chance; Necessity; Genetic control; DNA; RNA; Evolution; Information theory; Life origin; Astrobiology; Panspermia

## 1. Introduction

Genetic information consists of linear digital algorithmic programs. These programs are recorded into the sequencing of DNA's primary structure. Algorithms consist of sequences of decision-node switch-settings. Each nucleotide selection in the sequence constitutes a quaternary decision-node selection. “Quaternary” is not used here in a chemical sense, but to distinguish four-branch decision nodes from binary decision nodes with only two branches (a fork in the road). Each

selection of a nucleotide is from among the four real options.

Not all segments of the sequence are critical. But in those sections that are, each switch-setting is highly determinative of minimum-free-energy folding and binding success (Fontana and Schuster, 1998; Reidys et al., 2001; Schuster, 1995; Wuchty et al., 1999).

Peer-reviewed life-origin literature presupposes that, given enough time, genetic instructions arose via natural events. Thus far, no paper has provided a plausible mechanism for natural-process algorithm-writing. Only 200 million years separated the end of Earth's bombardment (Davies, 2001; Whittet, 1997) from the presumed first appearance of life on Earth 3.8 billion years ago (Martin and Russell, 2003; Mojzsis et al., 1996; Parsons et al., 1998; Schopf, 1993; Van Zuilen

\* Corresponding author. Tel.: +1 519 824 4120x53367; fax: +1 519 837 0442.

E-mail address: jtrevors@uoguelph.ca (J.T. Trevors).

et al., 2002). Following cooling, it is difficult to understand how natural processes could have generated the following aspects of life in such a short time:

- (1) a genetic operating system with which to record programming instructions,
- (2) the programs themselves for production or assembly of every individual building block, biochemical pathway, and metabolic cycle needed for even the simplest protometabolism to develop, and
- (3) a coding system with which to translate triplet codon “language” into polyamino acid language.

## 2. The derivation of genetic instruction in nature

There is an immense gap from prebiotic chemistry and the lifeless Earth to a complex DNA instruction set, code encryption into codonic sequences, and decryption (translation) into amino acid sequences. Sound reductionistic science should keep breaking down the immensity of the life-origin problem into its component problems. Life-origin specialists should appreciate that the generation of instructions is a separate and distinct problem from that of devising a language system with which to record those instructions. And these two problems are separate and distinct from the problem of encryption/decryption (coding) associated with translation to proteins. We tend to naively refer to genomes as genetic code. Molecular biological cybernetics is not that simple. The phenomenon of instructions must be constantly delineated from the phenomenon of an overall language, and also from the phenomenon of code encryption/decryption. Translating one operating system into another is not the same problem as writing algorithmic programs. Genes and emergent gene networks represent programming. These algorithms are written in a pre-existent operating system environment. As in computer science, this language is used by the programmer. We must not only find models for specific genetic programming, but for the genetic operating system.

Natural processes have never been observed to write conceptual instructions, to symbolize such algorithmic meaning, or to translate it using code bijection (a one-to-one correspondence of “meaning” between alphanumeric symbols) from one language system into another. In addition, inanimate nature seems to possess no attributes capable of encryption/decryption of coded messages. Genes are algorithmic programs “instantiated” into the physical medium of nucleotide sequence. What do we mean by “instantiated”? Algorithmic programs are abstract. They are integrated strings of fundamentally nonphysical choice commitments. But of necessity any program must be recorded into a physical medium

(e.g., a computer disc, a typed letter, an email, a telephone or TV signal, DNA). This recordation into physicality of choices is called “instantiation”. Successive ribonucleotide selections instantiate the abstract “instructions” for ribozyme secondary and tertiary structures. Successive triplet codon selections incorporate the instructions for polyamino acid sequence. This primary structure ultimately determines protein conformation and contextual function. The genetic operating system uses a bijective coding system whereby a certain triplet codon represents a certain amino acid. Thus, we are left with three separate missing mechanisms of molecular evolution theory which remain to be explained:

How did inanimate nature write

- (1) the conceptual instructions needed to organize metabolism?
- (2) a language/operating system needed to symbolically represent, record and replicate those instructions?
- (3) a bijective coding scheme (a one-to-one correspondence of symbol meaning) with planned redundancy so as to reduce noise pollution between triplet codon “block code” symbols (“bytes”) and amino acid symbols?

We could even add a fourth question. How did inanimate nature design and engineer

- (4) a cell [Turing machine? (Turing, 1936)] capable of implementing those coded instructions?

## 3. The role of natural selection

Natural selection could select the fittest already-programmed phenotypes. Evolution works through differential survival and reproduction of the superior members of each species. Phenotypes are the finished products of nucleic acid (genetic) algorithms. Natural selection could not have programmed nucleic acid algorithms at the covalently-bound primary structure (sequence) level. The environment does not select nucleotide or codonic sequences. The environment favors only the fittest phenotypes. It knows nothing of genotypic programming directly. Nature has no ability to optimize a conceptual cybernetic system at the decision node (covalently-bound sequence) level. Nature cannot organize conceptual, holistic operating systems and instructions from “necessary” (Monod, 1972) mass/energy relationships. Freedom of selection is necessary at each decision node. Gene regulation and coordination are programmed algorithmically. No known hypothetical mechanism has even been suggested for the generation of nucleic acid algorithms.

The standard genetic code table is published in a multitude of scientific papers and books. Numerous papers have been published in an attempt to explain the origin of the genetic code (Barbieri, 2003; Beebe et al.,

2003b; Crick, 1968; Di Giulio, 1997, 1999a,b, 2000, 2001a,b, 2002, 2003; Di Giulio and Medugno, 1998, 1999, 2000, 2001; Freeland et al., 2003; Guimaraes, 2004; Jukes, 1993; Jukes and Osawa, 1993; Knight et al., 1999; Osawa et al., 1992; Ribas de Pouplana et al., 1998; Ribas de Pouplana and Schimmel, 2000, 2001c; Ronneberg et al., 2000, 2001; Schimmel and Ribas de Pouplana, 1999; Schimmel and Wang, 1999; Seligmann and Amzallag, 2002; shCherbak, 2003; Skouloubris et al., 2003; Stevenson, 2002; Szymanski and Barciszewski, 2000; Szymanski et al., 2000; Trevors, 2003; Wang and Schultz, 2002; Woese et al., 2000; Wong, 1975, 1988; Xue et al., 2003; Yarian et al., 2002; Yarus and Christian, 1989; Yarus, 2000). All code origin models have problems and lack detail. In addition, no physical mechanism has been suggested for the source of abstract genetic instructions themselves. New formats and approaches are needed to investigate the origin of instructions, coding, biochemical pathways, cycles, metabolism, and life itself.

#### 4. The role of long periods of time

The argument has been repeatedly made that given sufficient time, a genetic instruction set and language system could have arisen. All that would be needed would be diversification, environmental selection, and continuing optimization. But extended time does not provide an explanatory mechanism for spontaneously generated genetic instruction. What is needed is a plausible mechanism for natural-process-generation of functional algorithms. We need empirical evidence of prescriptive genetic information arising spontaneously, without artificial investigator selection and amplification. A fulfilled prediction of the latter would be ideal. So far, none has occurred.

No amount of time proposed thus far, can explain this type of conceptual communication system. It is not just complex. It is conceptually complex. First, the ribosome/tRNA/aminoacyl tRNA synthetase/amino acid holistic translative *system* would have had to pre-exist any messages. Then we have to explain how the DNA and mRNA sequence provided the codon-encrypted instructions for the correct proteins to be synthesized. Only then could the receiver and destination have known what those instructions meant.

The appearance of genetic control does not seem possible unless the transmitted message and the decoded outcome were pre-arranged. It is an immense challenge to envision how this coding/decoding could have occurred on the early natural-process Earth, especially under harsh conditions for life. One possibility is that the first coding/decoding system was very simple and the first code produced only a few useful proteins. The code was then optimized to produce larger and larger

genomes and more complex organisms. If this did occur, then the first small genetic instruction set would have needed the capacity to become more complex and diverse. As new more complex instructions were sent, the translative system would have had to maintain relative constancy for needed proteins to be synthesized. This could only be achieved if the three-base codonic code for one amino acid was already established. A change in the number of nucleotides per codon at some time during early evolution would have resulted in reading frame shifts. This would have resulted in a catastrophic loss of everything gained up to that point. Life would have had to start over again.

#### 5. Code bijection is a separate problem from programming

Code bijection is a one-to-one correspondence of “meaning” between alphanumeric symbols in different languages or operating systems. The phrase “genetic code” should be reserved for describing this one-to-one correspondence between each triplet codon and its corresponding amino acid (Yockey, 1992, 2000, 2002a,b,c). Proper use of the term genetic code applies only to the redundant (usually erroneously called degenerative in the literature), noise-reducing table of codon assignments. These assignments are widely regarded as an amazingly optimal coding system (Bradley, 2002; Freeland and Hurst, 1998; Gilis et al., 2001; Labouygues and Figureau, 1984). Such optimization makes it all the more difficult to explain the molecular evolution of such exquisite non-human genetic algorithms. Random walks (Markov Processes) (Holland, 2003) will never provide an adequate explanation for the generation of such a highly refined translative coding system. Nor will it provide an explanation for the codonic operating system and specific programs generated through selection of each nucleotide in a strand.

Similarly, it is difficult to envision how the laws of physics and chemistry could explain encryption/decryption. No physicochemical link exists between programmed nucleotide sequence and amino acid functional sequence. How did each tRNA acquire the correct anticodon? In addition, how would each tRNA link up with the correct amino acid on its opposite end? Each aminoacyl tRNA synthetase would then have to establish a link with the correct amino acid. Paul Schimmel’s group and others have done some fine work in this area (Beebe et al., 2003a,b; Hendrickson et al., 2002; Nangle et al., 2002; Nomanbhoy and Schimmel, 2001; Ribas de Pouplana et al., 2001; Ribas de Pouplana and Schimmel, 2001a,b,c; Schimmel and Ribas de Pouplana, 2001; Tamura and Schimmel, 2001, 2002). But much of this research involves human engineering.

Cause-and-effect physicality has no ability to anticipate or devise a conceptual system that employs symbolic representationalism. Both the semantics and syntax of codonic language must translate into appropriate semantics and syntax of protein language. That symbolization must then translate into the “language” of three-dimensional conformation via minimum-free-energy folding. No combination of the four known forces of physics can account for such conceptual relationships. Symbolism and encryption/decryption are employed. Codons represent functional meaning only when the individual amino acids they prescribe are linked together in a certain order using a different language. Yet the individual amino acids do not directly react physicochemically with each triplet codon. Even after a linear digital sequence is created in a new language, “meaning” is realized at the destination only upon folding and lock-and-key binding.

How did 20 specific tRNA, aminoacyl tRNA synthetases, and amino acids self-organize into a holistic translative operating system? The origin of translation defies natural-process modeling as a holistic system. The ribosome is only one aspect of translation. Yet we do not even know how ribosomes formed to provide such sophisticated translation machinery.

The four known forces of physics know nothing of the phenomenon of linguistic translation. The laws of physics and chemistry cannot explain why each tRNA just happens to have the correct anticodon and links up with the correct amino acid and the correct aminoacyl tRNA synthetase. Cause-and-effect physicality has no ability to anticipate or devise a conceptual system that employs symbolic representationalism. Both the semantics and syntax of codonic language must translate into appropriate semantics and syntax of protein language.

The “chicken and egg” paradox, therefore, remains a stubborn problem. Little progress has been made on the origin of the genetic code on the Earth or elsewhere. Its improbable delivery to the Earth by bombardment or impact events billions of years ago has its own problems.

## 6. Astrobiological considerations

Exploring other planets and moons will provide new scientific knowledge. But astrobiology seems unlikely to explain the origin of the phenomenon of genetic instruction. Most investigators agree that the origin of life on the Earth was a very improbable event. Thus far, no evidence exists of life elsewhere in the universe. Life may have arisen on another planet or moon. But some regard it extremely unlikely that spores could have survived prolonged space travel and rapid entry into our atmosphere (Dose, 1986; Weber and Mayo Greenberg, 1985). Others disagree (Davies, 2001; Secker et al., 1996), citing spore coatings of silicon or carbon

(lithopanspermia) as a mechanism of shielding even interstellar organisms from UV radiation, excessive speed, shock, and heat (Melosh, 1993; Weiss et al., 2000). Even if panspermia were true (Arrhenius, 1908; Hoyle and Wickramasinghe, 1981, 1986; Parsons, 1996), the origin of an operating system language and specific prescriptive genetic information (source code) would remain unexplained. No fixed laws or formulae can program metabolic computation. Instruction is abstract and conceptual. Yet instruction is exactly what genomes do. In addition to instructing, they actually perform metabolic function in a hardware-implementation sense.

Much attention has been given to recent media announcements by NASA suggesting that liquid water once existed on Mars (Associated Press, USA Today, April 1 2004 “NASA rover finds more signs of past water on Mars”). Water is central to life as we know it (Ball, 2004; Good, 1973; Papagiannis, 1992). Water is the biological solvent. Its high polarity and high dielectric constant contribute to protein and ribosomal folding through hydrogen-bond and van der Waals like forces. Hydrophobic and hydrophilic groups within the same polypeptide help determine protein conformation (e.g., thermostable beta-sheets) (Brack, 1993; Trevors, 2002). Some proteins’ conformation depends upon trapped water molecules within their innermost folds. The lipid bilayers of membranes are permeable to water and ionic solutes (Deamer and Bramhall, 1986). Clay surface adsorption thought to assist early RNA polymerization depends upon one or two molecular layers of water molecules (Anderson and Banin, 1975). Water is a powerful hydrolytic agent. Hydrolysis is used to break down biomolecules for recycling into different primary structures (sequences). Organic solvents could not do this (Brack, 1993).

The presence of water suggests that a planet has (1) enough mass to retain an atmosphere, (2) enough rotation to aid cooling, and (3) the right distance from its sun to provide a habitable zone or ecosphere (Papagiannis, 1992). But the mere presence of liquid water on a planet tells us little about life origin. Astrophysicist Paul Davies calls this “the ingredients fallacy” (Davies, 2003). He points out that even the presence of carbon, hydrogen, nitrogen, oxygen, sulfur, and water on a planet no more guarantees life than the presence of silicon guarantees the presence of computers. The algorithmic complexity of life puts our finest computers to shame. Water has been present in abundance on Earth for approximately 4 billion years. Yet many investigators have abandoned earth-based life-origin models for more likely astrobiological sources. Water’s presence does not create or explain life.

The role of carbon polymers is also important to life. Cairns-Smith proposed clay-life models of life (Cairns-Smith, 1966, 1977, 1990; Cairns-Smith et al., 1972; Cairns-Smith and Walker, 1974). Clay-life models have

not been pursued for decades for multiple reasons. Information retention would be limited to crystal irregularities in the otherwise regular, highly-ordered crystal matrix. But the number of such irregularities would be insufficient to retain the amount of information required by life. In addition, no satisfactory mechanism of genetic takeover was ever proposed. The information would have to be translated from clay crystal irregularities into nucleotide sequences. No theoretical means of code bijection (one-to-one correspondence between languages) exists. Access to the non-surface and deeper clay layers to read clay information also remained unexplained by this model. Finally, clay crystal irregularities are randomly distributed. As such, they may meet the Shannon definition of “information” (which Shannon himself eschewed). But such randomly distributed irregularities would never have programmed sophisticated metabolic functions.

The only other options discussed for non-carbon life have been nonlayered silicon (Mann and Perry, 1986; Trevors, 1997a; Williams, 1986) and boron polymers. Silicon polymers cannot gain sufficient length for adequate information retention. Silicon forms bonds with other elements that would interfere with silicon–silicon chain formation. Silicon lacks the relatively easily-broken-and-rejoined covalent-like bonds enjoyed by carbon, hydrogen, and oxygen in organic compounds. Silicon bonds are too rigid and irreversible for cellular metabolic recycling of structural, enzymatic, regulatory, and informational biopolymers. Silicon is too insoluble in an aqueous environment. Sand, a typical silicon compound, is a good example. No organisms could have been produced except in an aqueous environment. Carbon, unlike silicon, is amenable with the help of catalysts to dehydration synthesis even in an aqueous environment. Yet carbon-based organisms do not dissolve in ponds, rivers, and oceans. Carbon chains are unique. Finally, silicon chains lack the ability of carbon chains to establish a lipoprotein-like connection between different kinds of biomolecules. Lipids have a different solubility and serve different functions from proteins. Both are needed for life as we know it. Carbon–carbon bonds provide both kinds of branching using the same basic building blocks. Lipoprotein molecules can cooperate to contribute to cellular survival through such functions as membrane formation. Silicon oxide can form layers, but lacks the unique properties of lipoprotein needed for semi-permeable membranes, active transport, secretion, and excretion.

The main interest in silicon has not been as the backbone molecule of a life system, but its role as a surface adsorbent and catalyst for proper alignment and polymerization of polyadenines and polyuracils (Burton et al., 1974; Ding et al., 1996; Ertem and Ferris, 1996, 1998, 2000; Ferris et al., 1996, 1988, 1989; Ferris and Ertem, 1992; Friebele et al., 1980, 1981; Huang and

Ferris, 2003; Kawamura and Ferris, 1999; Miyakawa and Ferris, 2003; Paecht-Horowitz and Eirich, 1988; Trevors, 1997a). But polyadenines and polyuracils, like the monotonous clay crystals to which they adsorb, contain almost no Shannon uncertainty (often misnamed “information”). Clay surface adsorption could not possibly be the source of highly informational genetic instructions.

The notion of boron life has never received serious attention. Based on current astronomical knowledge, there appears to be insufficient boron in the cosmos to support life on any planet. Any evolutionary scenario would require large quantities of boron compounds to provide enough diversity from which to select happenstantial algorithmic metabolic function. The replicative potential of carbon biopolymers is lacking in the case of both boron and silicon polymers. In addition, metabolic function depends largely upon folding. Neither silicon nor boron offers the peculiar secondary and tertiary folding-versatility needed to catalyze and support life. Unique lock-and-key binding fits are not afforded by boron or silicon molecules. No empirical evidence exists for any form of life other than carbon-chemistry life. Silicon and boron life models present so many challenges that it is difficult to imagine either using even the most elementary definition of life.

One could postulate that the extraterrestrial origin of life would have provided a longer time period for life to have originated and evolved elsewhere before it was delivered to the ancient Earth by one or more impact events. If we ignore the need for cosmic cooling following a hot Big Bang, we only extend the available time from 4 billion to 14 billion years. This does little to overcome the statistical prohibitiveness of algorithmic self-organization.

Organic compounds were found in the Murchison meteorite (Deamer and Pashley, 1989; Kvenvolden et al., 1970; Ponnampereuma, 1972). The transport of life in rock fragments from Mars has been suggested. This possibility need not be limited to a nearby planet. The maximum amount of additional time available for evolution to arise on Mars over that available on Earth is only about 100 million years (Line, 2002). A maximum extension of 600 million additional years anywhere in the Solar System exists over that available on habitable Earth (Line, 2002). This suggests that the origin of life on Mars, other planets and moons, and its transport to the Earth becomes highly improbable.

Many biochemical questions remain as to how exclusively right-handed ribonucleotides could have been activated and polymerized with 3'5' bonds in an aqueous environment (Shapiro, 1984, 1987, 1988, 1999, 2000, 2002). Ribose and ribonucleotides are hard to make and are not stable. In addition, when delivery of life from one planet to another is proposed, it focuses on the delivery of spores and/or bacterial cells. Possibly, the

spores were simply the carrier for the genetic instruction set that was inserted into the primitive lifeless Earth. But neither DNA nor its instantiated instructions are themselves alive. All of the genetic instructions in DNA are usually still present milliseconds after cell death. Yet the cell is dead. Clearly, genetic programs are not synonymous with life itself. If viable spore-seeding of Earth did occur, an intervening time period of about 3.8 billion years (Mojzsis et al., 1996; Parsons et al., 1998; Schopf, 1993; Van Zuilen et al., 2002) passed, to bring us to our present state of evolution.

## 7. The nature of prescriptive information

All of the above problems pale in comparison to the difficulty of explaining the origin of (1) an operating system, (2) genetic programming, and (3) encryption/decryption coding. Natural processes, mechanisms, and chemical catalyses do not explain any of these emergent conceptual phenomena. There is an immense paucity of information/knowledge that we do not have. For example, transcription and translation can only function if there is a predetermined shared meaning as to what required proteins are synthesized at translation. How was this shared meaning between source and destination pre-established? Each specific genetic message from DNA to RNA to protein can only be decoded if the coding/decoding apparatus and operating system pre-exist the message. The message received by the ribosome must be decrypted and translated into particular proteins needed for certain tasks. These proteins in turn must be transported to the correct binding site, the true destination of the source's original message. Even "meaningful" RNA or DNA inserted into a lifeless physical world such as the ancient Earth, would not be "readable". It could not communicate its coded message for protein synthesis unless a language (operating system) context already existed. Programs must be executable. This requires the equivalent of a hardware and an operating system context. All necessary structures/functions for protein synthesis would have had to be in place, and a predetermined specific correspondence between codon sequence and amino acid sequence had to have predated translation.

## 8. The RNA world

It is not surprising that the RNA world model is so appealing despite its many biochemical problems (Joyce and Orgel, 1999; Shapiro, 1984, 1987, 1988, 1999, 2000, 2002). The RNA world conveniently bypasses translative coding issues. The same molecule acts as a catalyst and physical information matrix. Ribonucleotides do not have to be grouped into triplet "block codes" such

that each symbolize a different amino acid letter in a protein sentence. No plausible natural-process mechanism for development of such an ingenious noise-reducing, redundancy-based coding scheme is needed for RNA world theory.

RNA may have been copied to DNA (as by current retroviruses). This would have allowed DNA to take over as the more stable, double-stranded, genetic instruction set. DNA has many other advantages such as the capability of exact replication and partitioning between offspring cells, and transcription to mRNA (Line, 2002). But retroviruses depend upon reverse transcriptase, a complex protein enzyme not available in the RNA world. Most metabolic functions require highly tailored protein enzymes that not only act to regulate DNA, but do most of the work of the cell. The "Which came first, DNA or proteins?" question, otherwise known as the "chicken and the egg problem", is a real enigma and not easily solved despite a century of life-origin research. In addition, the conversion from RNA to DNA worlds does not explain the origin of initial RNA genetic programming. How did covalently-bound nucleotide sequencing anticipate what amino acid sequences would be needed? Moreover, the instruction code for the enzymes needed to make all this function are contained in the genetic instruction set. The instruction set needs protein synthesis to replicate the instruction set and regulate cell division (Trevors, 2004).

## 9. Did genetic instruction arise by "necessity"?

We keep intuiting, "there must be some natural mechanism that we are overlooking which will explain the origin of genetic instruction". By this we mean some cause-and-effect "necessity" rather than "chance" (Monod, 1972). But is such a natural mechanism a plausible scenario for the origin of genetic instructions?

Natural mechanisms are all highly self-ordering. Reams of data can be reduced to very simple compression algorithms called the laws of physics and chemistry. No natural mechanism of nature reducible to law can explain the high information content of genomes. This is a mathematical truism, not a matter subject to overturning by future empirical data. The cause-and-effect necessity described by natural law manifests a probability approaching 1.0. Shannon uncertainty is a probability function ( $-\log_2 p$ ). When the probability of natural law events approaches 1.0, the Shannon uncertainty content becomes miniscule ( $-\log_2 p = -\log_2 1.0 = 0$  uncertainty). There is simply not enough Shannon uncertainty in cause-and-effect determinism and its reductionistic laws to retain instructions for life. Prescriptive information (instruction) can only be explained

by algorithmic programming. Such DNA programming requires extraordinary bit measurements often extending into megabytes and even gigabytes. That kind of uncertainty reflects freedom from law-like constraints. This is exactly what we find at each decision-node selection of an additional untemplated nucleotide.

Living cells are capable of replication, error correction, and sufficient mutation for diversity; yet sufficient accuracy of replication is maintained to preserve the relative constancy of any species. The DNA template can be transcribed to mRNA which is then translated into predetermined, useful, necessary, proteins. This is an immense instructional complexity and prone to errors at several steps and at every base. However, within the confines of a cell, the entire process functions at a high level of fidelity. Exceptions include the effects of mutagen(s), cell stress, injury, and death. Under appropriate environmental conditions, minimal errors are made. This means that the genetic communication system is functionally optimized (Bradley, 2002; Freeland and Hurst, 1998; Freeland et al., 2000; Gilis et al., 2001; Labouygues and Figureau, 1984).

The sequence of nucleotides in DNA determines the coded instruction set. A predetermined knowledge of the decryption cipher and the cellular enzymes and organelles is needed to translate the coded information. Specific sequences of deoxyribonucleotides are essential to communicate each biomessage. The correct protein must be synthesized in the correct amounts and at the correct time. Without a coding/decoding system, message sequences in the first mRNA and DNA molecules would have been meaningless (nonfunctional metabolically). Communication within the protocell could not have been established. The nucleotide sequence is also meaningless without a conceptual translative scheme and physical “hardware” capabilities. Ribosomes, tRNAs, aminoacyl tRNA synthetases, and amino acids are all hardware components of the Shannon message “receiver”. But the instructions for this machinery is itself coded in DNA and executed by protein “workers” produced by that machinery. Without the machinery and protein workers, the message cannot be received and understood. And without genetic instruction, the machinery cannot be assembled.

## 10. Did the genetic code arise by “chance”?

It is not reasonable to expect hundreds to thousands of random sequence polymers to all cooperatively self-organize into an amazingly efficient holistic metabolic network. The spontaneous generation of long sequences of DNA out of sequence space ( $\Omega$ ) does have the potential to include the same sequences as genetic information. But there is no reason to suspect that any instructive biopolymer would isolate itself out of  $\Omega$  and

present itself at the right place and time. Eigen and Schuster, along with others, have pointed out repeatedly that a competition for resources would have existed in any prebiotic environment. This would have greatly limited both sequence space and hypercyclic advance (Eigen, 1971a,b, 1983, 1987, 1992; Eigen et al., 1980a,b; Schuster, 1984; Smith, 1979). The latter is especially true of a theoretical RNA world where the number and length of RNA strands is greatly limited. In aqueous solution, a maximum of 8–10 RNA mers can polymerize (Ferris et al., 1996; Joyce and Orgel, 1999). Up to 55 mers can polymerize on montmorillonite (Ferris et al., 1996), but only at the expense of information content. The polyadenines and polyuracils contain essentially no Shannon uncertainty. They could not have contributed to algorithmic programming of genes.

Even if all the right primary structures (digital messages) mysteriously emerged at the same time from  $\Omega$ , “a cell is not a bag of enzymes”. And, as we have pointed out several times, there would be no operating system to read these messages.

Without selection of functional base sequencing at the covalent level, no biopolymer would be expected to meet the needs of an organizing metabolic network. There is no prescriptive information in random sequence nucleic acid. Even if there were, unless a system for interpreting and translating those messages existed, the digital sequence would be unintelligible at the receiver and destination. The letters of any alphabet used in words have no prescriptive function unless the destination reading those words first knows the language convention.

## 11. The role of biological editing functions

No new information can be inserted into existing DNA without sophisticated restriction and ligase enzymes. The editing function of these enzymes, including that of the far less sophisticated ribozymes, must be itself algorithmically instructed. All in vitro ribozymic editing requires extensive artificial selection by humans (e.g., SELEX) (Ellington and Szostak, 1990; Robertson and Joyce, 1990; Tuerk and Gold, 1990). Nucleotide sequence is deliberately manipulated and steered through many iterations to achieve the experimenter’s goal. SELEX experiments demonstrate the extraordinary inventive prowess of some excellent RNA chemists. But the fine work of these biopolymer engineers has little or nothing to do with *natural* selection. In addition, “directed evolution” is a self-contradictory nonsense phrase that has no place in the literature. If an experiment is directed, it is not evolutionary. Evolution has no goal.

Whatever prescriptive information DNA contains has to be instantiated into its physical matrix as the

strand forms with covalent bonds. Conformation and function are ultimately determined by primary structure. The linear digital sequence of nucleotide selections constitutes the message of “messenger molecules”. A genetic information system or convention must have been devised prior to ribonucleotide sequencing. Only then could message source and destination (binding sites) “be on the same page”.

Random sequences are the antithesis of prescribed genetic information. There is no empirical or rational justification for theorizing that the random shuffling of nucleotides could generate instructions for a metabolic network. Progress has been made, however, on the evolution of already existing genetic instructions (Altman, 2000; Altreuter and Clark, 1999; Alves et al., 2002; Baltscheffsky, 1997; Baltscheffsky et al., 1999; Benner et al., 1987; Benner and Ellington, 1987, 1988; Blankenship and Hartman, 1998; Castresana and Saraste, 1995; Cech 1993, 2000; Cunchillos and Lecointre, 2002, 2003; Goossens et al., 2003; Hartman, 1975; Morowitz et al., 2000; Trevors, 1997b; Wachtershauser, 1990, 1992). But none of these papers provide mechanisms whereby stochastic ensembles in prebiotic environments acquire algorithmic programming prowess. Even the earliest protometabolism would have needed integrative management.

## 12. Conclusions

New approaches to investigating the origin of the genetic code are required. The constraints of historical science are such that the origin of life may never be understood. Selection pressure cannot select nucleotides at the digital programming level where primary structures form. Genomes predetermine the phenotypes which natural selection only secondarily favors. Contentions that offer nothing more than long periods of time offer no mechanism of explanation for the derivation of genetic programming. No new information is provided by such tautologies. The argument simply says it happened. As such, it is nothing more than blind belief. Science must provide rational theoretical mechanism, empirical support, prediction fulfillment, or some combination of these three. If none of these three are available, science should reconsider that molecular evolution of genetic cybernetics is a proven fact and press forward with new research approaches which are not obvious at this time.

## Acknowledgements

JTT is supported by an NSERC (Canada) Discovery grant; DLA is supported by grants from The Origin-of-Life Foundation, Inc. a 501-c-3 science foundation.

## References

- Altman S. The road to RNase P. *Nat Struct Biol* 2000;7:827–8.
- Altreuter DH, Clark DS. Combinatorial biocatalysis: taking the lead from nature. *Curr Opin Biotechnol* 1999;10:130–6.
- Alves R, Chaleil RAG, Sternberg MJE. Evolution of enzymes in metabolism: a network perspective. *J Mol Biol* 2002;320:751–70.
- Anderson DM, Banin A. Soil and water and its relationship to the origin of life. *Orig Life* 1975;6:23–36.
- Arrhenius S. *Worlds in the making*. New York: Harper and Row; 1908.
- Ball P. Water, water, everywhere? *Nature* 2004;427:19–20.
- Baltscheffsky H. Major “anastrophes” in the origin and early evolution of biological energy conversion. *J Theor Biol* 1997;187:495–501.
- Baltscheffsky M, Schultz A, Baltscheffsky H. H<sup>+</sup>-proton-pumping inorganic pyrophosphatase: a tightly membrane-bound family. *FEBS Lett* 1999;452:121–7.
- Barbieri M. *The organic codes: an introduction to semantic biology*. Cambridge: Cambridge University Press; 2003.
- Beebe K, Merriman E, Schimmel P. Structure-specific tRNA determinants for editing a mischarged amino acid. *J Biol Chem* 2003a;278:45056–61.
- Beebe K, Ribas de Pouplana L, Schimmel P. Elucidation of tRNA-dependent editing by a class II tRNA synthetase and significance for cell viability. *EMBO J* 2003b;22:668–75.
- Benner SA, Allemann RK, Ellington AD, Ge L, Glasfeld A, Leanz GF, et al. Natural selection, protein engineering, and the last riboorganism: rational model building in biochemistry. *Cold Spring Harb Symp Quant Biol* 1987;52:53–63.
- Benner SA, Ellington AD. The last ribo-organism. *Nature* 1987;329:295–6.
- Benner SA, Ellington AD. Return of the ‘last ribo-organism’. *Nature* 1988;332:688–9.
- Blankenship RE, Hartman H. The origin and evolution of oxygenic photosynthesis. *Trends Biochem Sci* 1998;23:94–7.
- Brack A. Liquid water and the origin of life. *Orig Life Evol Biosph* 1993;23:3–10.
- Bradley D. Informatics. The genome chose its alphabet with care. *Science* 2002;297:1789–91.
- Burton FG, Lohrmann R, Orgel LE. On the possible role of crystals in the origins of life. VII. The adsorption and polymerization of phosphoramidates by montmorillonite clay. *J Mol Evol* 1974;3:141–50.
- Cairns-Smith AG. The origin of life and the nature of the primitive gene. *J Theor Biol* 1966;10:53–88.
- Cairns-Smith AG, Ingram P, Walker GL. Formose production by minerals: possible relevance to the origin of life. *J Theor Biol* 1972;35:601–4.
- Cairns-Smith AG, Walker GL. Primitive metabolism. *Curr Mod Biol* 1974;5:173–86.
- Cairns-Smith AG. Takeover mechanisms and early biochemical evolution. *Biosystems* 1977;9:105–9.
- Cairns-Smith AG. *Seven clues to the origin of life*. Cambridge: Cambridge University Press; 1990.
- Castresana J, Saraste M. Evolution of energetic metabolism: the respiration – early hypothesis. *Trends Biochem Sci* 1995;20:443–8.
- Cech TR. Catalytic RNA: structure and mechanism. *Biochem Soc Trans* 1993;21:229–34.
- Cech TR. Structural biology. The ribosome is a ribozyme. *Science* 2000;289:878–9.
- Crick FHC. The origin of the genetic code. *J Mol Biol* 1968;38:367–79.
- Cunchillos C, Lecointre G. Early steps of metabolism evolution inferred by cladistic analysis of amino acid catabolic pathways. *C R Biol* 2002;325:119–29.

- Cunchillos C, Lecointre G. Evolution of amino acid metabolism inferred through cladistic analysis. *J Biol Chem* 2003;278:47960–70.
- Davies P. The origin of life. I: When and where did it begin? *Sci Prog* 2001;84:1–16.
- Davies PCW. How biofriendly is the universe? *Int J Astrobiol* 2003;2:115–20.
- Deamer DW, Bramhall J. Permeability of lipid bilayers to water and ionic solutes. *Chem Phys Lipids* 1986;40:167–88.
- Deamer DW, Pashley RM. Amphiphilic components of the Murchison carbonaceous chondrite: surface properties and membrane formation. *Orig Life Evol Biosph* 1989;19:21–38.
- Di Giulio M. On the origin of the genetic code. *J Theor Biol* 1997;187:573–81.
- Di Giulio M, Medugno M. The historical factor: the biosynthetic relationships between amino acids and their physicochemical properties in the origin of the genetic code. *J Mol Evol* 1998;46:615–21.
- Di Giulio M. The coevolution theory of the origin of the genetic code. *J Mol Evol* 1999a;48:253–5.
- Di Giulio M. The non-monophyletic origin of the tRNA molecule. *J Theor Biol* 1999b;197:403–14.
- Di Giulio M, Medugno M. Physicochemical optimization in the genetic code origin as the number of codified amino acids increases. *J Mol Evol* 1999;49:1–10.
- Di Giulio M. Genetic code origin and the strength of natural selection. *J Theor Biol* 2000;205:659–61.
- Di Giulio M, Medugno M. The robust statistical bases of the coevolution theory of genetic code origin. *J Mol Evol* 2000;50:258–63.
- Di Giulio M. The non-universality of the genetic code: the universal ancestor was a progenote. *J Theor Biol* 2001;209:345–9.
- Di Giulio M. A blind empiricism against the coevolution theory of the origin of the genetic code. *J Mol Evol* 2001;53:724–32.
- Di Giulio M, Medugno M. The level and landscape of optimization in the origin of the genetic code. *J Mol Evol* 2001;52:372–82.
- Di Giulio M. Genetic code origin: are the pathways of type Glu-tRNA(Gln) → Gln-tRNA(Gln) molecular fossils or not? *J Mol Evol* 2002;55:616–22.
- Di Giulio M. The early phases of genetic code origin: conjectures on the evolution of coded catalysis. *Orig Life Evol Biosph* 2003;33:479–89.
- Ding PZ, Kawamura K, Ferris JP. Oligomerization of uridine phosphorimidazolides on montmorillonite: a model for the prebiotic synthesis of RNA on minerals. *Orig Life Evol Biosph* 1996;26:151–71.
- Dose K. Hypotheses on the appearance of life on Earth (review). *Adv Space Res* 1986;6:181–6.
- Eigen M. Self-organization of matter and the evolution of biological macromolecules. *Naturwissenschaften* 1971a;58:465–523.
- Eigen M. Molecular self-organization and the early stages of evolution. *Experientia* 1971b;27:149–212.
- Eigen M, Gardiner Jr WC, Schuster P. Hypercycles and compartments. Compartments assists – but do not replace – hypercyclic organization of early genetic information. *J Theor Biol* 1980;85:407–11.
- Eigen M, Schuster P, Sigmund K, Wolff R. Elementary step dynamics of catalytic hypercycles. *Biosystems* 1980;13:1–22.
- Eigen M. Life from the test tube? *MMW Munch Med Wochenschr* 1983;(Suppl 1)S125–35.
- Eigen M. New concepts for dealing with the evolution of nucleic acids. *Cold Spring Harb Symp Quant Biol* 1987;52:307–20.
- Eigen M. Steps toward life. Oxford, UK: Oxford University Press; 1992.
- Ellington AD, Szostak JW. In vitro selection of RNA molecules that bind specific ligands. *Nature* 1990;346:818–22.
- Ertem G, Ferris JP. Synthesis of RNA oligomers on heterogeneous templates. *Nature* 1996;379:238–40.
- Ertem G, Ferris JP. Formation of RNA oligomers on montmorillonite: site of catalysis. *Orig Life Evol Biosph* 1998;28:485–99.
- Ertem G, Ferris JP. Sequence- and regio-selectivity in the montmorillonite-catalyzed synthesis of RNA. *Orig Life Evol Biosph* 2000;30:411–22.
- Ferris JP, Huang CH, Hagan Jr WJ. Montmorillonite: a multifunctional mineral catalyst for the prebiological formation of phosphate esters. *Orig Life Evol Biosph* 1988;18:121–33.
- Ferris JP, Ertem G, Agarwal V. Mineral catalysis of the formation of dimers of 5'-AMP in aqueous solution: the possible role of montmorillonite clays in the prebiotic synthesis of RNA. *Orig Life Evol Biosph* 1989;19:165–78.
- Ferris JP, Ertem G. Oligomerization of ribonucleotides on montmorillonite: reaction of the 5'-phosphorimidazolide of adenosine. *Science* 1992;257:1387–9.
- Ferris JP, Hill Jr AR, Liu R, Orgel LE. Synthesis of long prebiotic oligomers on mineral surfaces. *Nature* 1996;381:59–61.
- Fontana W, Schuster P. Shaping space: the possible and the attainable in RNA genotype–phenotype mapping. *J Theor Biol* 1998;194:491–515.
- Freeland SJ, Hurst LD. The genetic code is one in a million. *J Mol Evol* 1998;47:238–48.
- Freeland SJ, Knight RD, Landweber LF, Hurst LD. Early fixation of an optimal genetic code. *Mol Biol Evol* 2000;17:511–8.
- Freeland SJ, Wu T, Keulmann N. The case for an error minimizing standard genetic code. *Orig Life Evol Biosph* 2003;33:457–77.
- Friebele E, Shimoyama A, Ponnampertuma C. Adsorption of protein and non-protein amino acids on a clay mineral: a possible role of selection in chemical evolution. *J Mol Evol* 1980;16:269–78.
- Friebele E, Shimoyama A, Hare PE, Ponnampertuma C. Adsorption of amino acid enantiomers by Na-montmorillonite. *Orig Life* 1981;11:173–84.
- Gillis D, Massar S, Cerf NJ, Rooman M. Optimality of the genetic code with respect to protein stability and amino-acid frequencies. *Genome Biol* 2001;2:R49.1–R49.12.
- Good W. The role of water in the origin of life and its function in the primitive gene. *J Theor Biol* 1973;39:249–76.
- Goossens A, Hakkinen ST, Laakso I, Seppanen-Laakso T, Biondi S, De Sutter V, et al. A functional genomics approach toward the understanding of secondary metabolism in plant cells. *PNAS* 2003;100:8595–600.
- Guimaraes RC. The Genetic Code as a Self-Referential and Functional System. *International Conference on Computation, Communications and Control Technologies* 2004;7:160–5.
- Hartman H. Speculations on the origin and evolution of metabolism. *J Mol Evol* 1975;4:359–70.
- Hendrickson TL, et al. Mutational separation of two pathways for editing by a class I tRNA synthetase. *Mol Cell* 2002;9:353–62.
- Holland JH. A derived Markov process for modeling reaction networks. *Evol Comput* 2003;11:339–62.
- Hoyle F, Wickramasinghe NC. Space travelers. Cardiff: University College Cardiff Press; 1981.
- Hoyle F, Wickramasinghe NC. The case for life as a cosmic phenomenon. *Nature* 1986;322:509–11.
- Huang W, Ferris JP. Synthesis of 35–40 mers of RNA oligomers from unblocked monomers. A simple approach to the RNA world. *Chem Commun (Camb)* 2003;12:1458–9.
- Joyce GF, Orgel LE. Prospects for understanding the origin of the RNA world. In: Gesteland RF, Cech TR, Atkins JF, editors. *The RNA world*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 1999. p. 49–78.
- Jukes TH. The genetic code – function and evolution. *Cell Mol Biol Res* 1993;39:685–8.
- Jukes TH, Osawa S. Evolutionary changes in the genetic code. *Comp Biochem Physiol B* 1993;106:489–94.

- Kawamura K, Ferris JP. Clay catalysis of oligonucleotide formation: kinetics of the reaction of the 5'-phosphorimidazolides of nucleotides with the non-basic heterocycles uracil and hypoxanthine. *Orig Life Evol Biosph* 1999;29:563–91.
- Knight RD, Freeland SJ, Landweber LF. Selection, history and chemistry: the three faces of the genetic code. *Trends Biochem Sci* 1999;24:241–7.
- Kvenvolden K, Lawless J, Pering K, Peterson E, Flores J, Ponnampereuma C, et al. Evidence for extraterrestrial amino-acids and hydrocarbons in the Murchison meteorite. *Nature* 1970;228:923–6.
- Labouygues JM, Figureau A. The logic of the genetic code: synonyms and optimality against effects of mutations. *Orig Life* 1984;14:685–92.
- Line MA. The enigma of the origin of life and its timing. *Microbiology* 2002;148(Pt 1):21–7.
- Mann S, Perry CC. Structural aspects of biogenic silica. In: *Silicon Biochemistry*. CIBA Foundation Symposium 121. Sussex, UK: John Wiley & Sons; 1986. p. 40–58.
- Martin W, Russell MJ. On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. *Philos Trans R Soc Lond B Biol Sci* 2003;358:59–83 [discussion 83–5].
- Melosh HJ. Blasting rocks off planets. *Nature* 1993;363:498–9.
- Miyakawa S, Ferris JP. Sequence- and regioselectivity in the montmorillonite-catalyzed synthesis of RNA. *J Am Chem Soc* 2003;125:8202–8.
- Mojzsis SJ, Arrhenius G, McKeegan KD, Harrison TM, Nutman AP, Friend GRL. Evidence for life on Earth before 3,800 million years ago. *Nature* 1996;384:55–9.
- Monod J. *Chance and necessity*. New York: Knopf; 1972.
- Morowitz HJ, Kostelnik JD, Yang J, Cody GD. From the cover: the origin of intermediary metabolism [see comments]. *Proc Natl Acad Sci U S A* 2000;97:7704–8.
- Nangle LA, De Crecy Lagard V, Doring V, Schimmel P. Genetic code ambiguity. Cell viability related to the severity of editing defects in mutant tRNA synthetases. *J Biol Chem* 2002;277:45729–33.
- Nomanbhoy TK, Schimmel P. Active site of an aminoacyl-tRNA synthetase dissected by energy-transfer-dependent fluorescence. *Bioorg Med Chem Lett* 2001;11:1485–91.
- Osawa S, Jukes TH, Watanabe K, Muto A. Recent evidence for evolution of the genetic code. *Microbiol Rev* 1992;56:229–64.
- Paecht-Horowitz M, Eirich FR. The polymerization of amino acid adenylates on sodium-montmorillonite with preadsorbed polypeptides. *Orig Life Evol Biosph* 1988;18:359–87.
- Papagiannis MD. What makes a planet habitable, and how to search for habitable planets in other solar systems. *J Br Interplanet Soc* 1992;45:227–30.
- Parsons I, Lee MR, Smith JV. Biochemical evolution. II: origin of life in tubular microstructures on weathered feldspar surfaces. *Proc Natl Acad Sci U S A* 1998;95:15173–6.
- Parsons P. Exobiology: dusting off panspermia. *Nature* 1996;383:221–2.
- Ponnampereuma C. Organic compounds in the Murchison meteorite. *Ann N Y Acad Sci* 1972;194:56–70.
- Reidys C, Forst CV, Schuster P. Replication and mutation on neutral networks. *Bull Math Biol* 2001;63:57–94.
- Ribas de Pouplana L, Turner RJ, Steer BA, Schimmel P. Genetic code origins: tRNAs older than their synthetases? *Proc Natl Acad Sci U S A* 1998;95:11295–300.
- Ribas de Pouplana L, Schimmel P. A view into the origin of life: aminoacyl-tRNA synthetases. *Cell Mol Life Sci* 2000;57:865–870.
- Ribas de Pouplana L, Brown JR, Schimmel P. Structure-based phylogeny of class IIa tRNA synthetases in relation to an unusual biochemistry. *J Mol Evol* 2001;53:261–8.
- Ribas de Pouplana L, Schimmel P. Operational RNA code for amino acids in relation to genetic code in evolution. *J Biol Chem* 2001a;276:6881–4.
- Ribas de Pouplana L, Schimmel P. Two classes of tRNA synthetases suggested by sterically compatible dockings on tRNA acceptor stem. *Cell* 2001b;104:191–3.
- Ribas de Pouplana L, Schimmel P. Aminoacyl-tRNA synthetases: potential markers of genetic code development. *Trends Biochem Sci* 2001c;26:591–6.
- Robertson DL, Joyce GF. Selection in vitro of an RNA enzyme that specifically cleaves single-stranded DNA. *Nature* 1990;344:467–8.
- Ronneberg TA, Landweber LF, Freeland SJ. Testing a biosynthetic theory of the genetic code: fact or artifact? *Proc Natl Acad Sci U S A* 2000;97:13690–5.
- Ronneberg TA, Freeland SJ, Landweber LF. Genview and gencode: a pair of programs to test theories of genetic code evolution. *Bioinformatics* 2001;17:280–1.
- Schimmel P, Ribas de Pouplana L. Genetic code origins: experiments confirm phylogenetic predictions and may explain a puzzle [published erratum appears in *Proc Natl Acad Sci U S A* 1999 May 11;96(10):5890]. *Proc Natl Acad Sci U S A* 1999;96:327–8.
- Schimmel P, Wang CC. Getting tRNA synthetases into the nucleus. *Trends Biochem Sci* 1999;24:127–8.
- Schimmel P, Ribas de Pouplana L. Formation of two classes of tRNA synthetases in relation to editing functions and genetic code. *Cold Spring Harb Symp Quant Biol* 2001;66:161–6.
- Schopf JW. Microfossils of the early fossil archaean apex chert: new evidence of the antiquity of life. *Science* 1993;260:640–6.
- Schuster P. Polynucleotide evolution, hypercycles and the origin of the genetic code. *Adv Space Res* 1984;4:143–51.
- Schuster P. How to search for RNA structures. Theoretical concepts in evolutionary biotechnology. *J Biotechnol* 1995;41:239–57.
- Secker J, Wesson PS, Lepock JR. Astrophysical and biological constraints on radiopanspermia. *J R Astro Soc Can* 1996;90:184–92.
- Seligmann H, Amzallag GN. Chemical interactions between amino acid and RNA: multiplicity of the levels of specificity explains origin of the genetic code. *Naturwissenschaften* 2002;89:542–51.
- Shapiro R. The improbability of prebiotic nucleic acid synthesis. *Orig Life* 1984;14:565–70.
- Shapiro R. *Origins: a skeptic's guide to the creation of life on earth*. New York: Bantam; 1987.
- Shapiro R. Prebiotic ribose synthesis: a critical analysis. *Orig Life Evol Biosph* 1988;18:71–85.
- Shapiro R. Prebiotic cytosine synthesis: a critical analysis and implications for the origin of life. *Proc Natl Acad Sci U S A* 1999;96:4396–401.
- Shapiro R. A replicator was not involved in the origin of life. *IUBMB Life* 2000;49:173–6.
- Shapiro R. Comments on 'concentration by evaporation and the prebiotic synthesis of cytosine'. *Orig Life Evol Biosph* 2002;32:275–8.
- shCherbak VI. Arithmetic inside the universal genetic code. *Bio-systems* 2003;70:187–209.
- Skouloubris S, de Pouplana LR, de Reuse H, Hendrickson TL. A noncognate aminoacyl-tRNA synthetase that may resolve a missing link in protein evolution. *Proc Natl Acad Sci U S A* 2003;100:11297–302.
- Smith JM. Hypercycles and the origin of life. *Nature* 1979;280:445–6.
- Stevenson DS. Co-evolution of the genetic code and ribozyme replication. *J Theor Biol* 2002;217:235–53.
- Szymanski M, Barciszewski J. Aminoacyl-tRNA synthetases database Y2K. *Nucleic Acids Res* 2000;28:326–8.
- Szymanski M, Deniziak M, Barciszewski J. The new aspects of aminoacyl-tRNA synthetases. *Acta Biochim Pol* 2000;47:821–34.

- Tamura K, Schimmel P. Oligonucleotide-directed peptide synthesis in a ribosome- and ribozyme-free system. *Proc Natl Acad Sci U S A* 13 February 2001;98:1393–7.
- Tamura K, Schimmel P. Ribozyme programming extends the protein code. *Nat Biotechnol* July 2002;20:669–70.
- Trevors JT. Bacterial evolution and silicon. *Antonie Van Leeuwenhoek* 1997a;71:271–6.
- Trevors JT. Bacterial evolution and metabolism. *Antonie Van Leeuwenhoek* 1997b;71:257–63.
- Trevors JT. Hydrophobic Medium (HM) water interface, cell division and the self-assembly of life. *Theory Biosci* 2002;121:163–74.
- Trevors JT. Genetic material in the early evolution of bacteria. *Microbiol Res* 2003;158:1–6.
- Trevors JT. Evolution of cell division in bacteria. *Theory Biosci* 2004;123:3–15.
- Tuerk C, Gold L. Systematic evolution of ligands by exponential enrichment – RNA ligands to bacteriophage-T4 DNA-polymerase. *Science* 1990;249:505–10.
- Turing AM. On computable numbers, with an application to the *entscheidungs problem*. *Proc R Soc Lond Math Soc* 1936;42:230–65 [correction in 43, 544–46].
- Van Zuilen MA, Lepland A, Arrhenius G. Reassessing the evidence for the earliest traces of life. *Nature* 2002;418:627–30.
- Wächtershäuser G. Evolution of the first metabolic cycles. *Proc Natl Acad Sci U S A* 1990;87:200–4.
- Wächtershäuser G. Groundworks for an evolutionary biochemistry: the iron–sulphur world. *Prog Biophys Mol Biol* 1992;58:85–201.
- Wang L, Schultz PG. Expanding the genetic code. *Chem Commun (Camb)* 2002;1–11.
- Weber P, Mayo Greenberg J. Can spores survive in interstellar space? *Nature* 1985;316:403–7.
- Weiss BP, et al. A low temperature transfer of ALH84001 from Mars to Earth. *Science* 2000;290:791–5.
- Whittet DCB. Is extraterrestrial organic matter relevant to the origin of life on Earth? *Orig Life Evol Biosph* 1997;27:249–62.
- Williams RJP. Introduction to silicon chemistry and biochemistry. In: *Silicon Biochemistry*. CIBA Foundation Symposium 121. Sussex, UK: John Wiley & Sons; 1986. p. 24–39.
- Woese CR, Olsen GJ, Ibba M, Soll D. Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiol Mol Biol Rev* 2000;64:202–36.
- Wong JT. A co-evolution theory of the genetic code. *Proc Natl Acad Sci U S A* 1975;72:1909–12.
- Wong JT. Evolution of the genetic code. *Microbiol Sci* 1988;5:174–81.
- Wuchty S, Fontana W, Hofacker IL, Schuster P. Complete suboptimal folding of RNA and the stability of secondary structures. *Biopolymers* 1999;49:145–65.
- Xue H, Tong KL, Marck C, Grosjean H, Wong JT. Transfer RNA paralogs: evidence for genetic code-amino acid biosynthesis coevolution and an archaeal root of life. *Gene* 2003;310:59–66.
- Yarian C, Townsend H, Czestkowski W, Sochacka E, Malkiewicz AJ, Guenther R, et al. Accurate translation of the genetic code depends on tRNA modified nucleosides. *J Biol Chem* 2002;277:16391–5.
- Yarus M, Christian EL. Genetic Code origins. *Nature* 1989;342:349–50.
- Yarus M. RNA-ligand chemistry: a testable source for the genetic code. *RNA* 2000;6:475–84.
- Yockey HP. Information theory and molecular biology. Cambridge: Cambridge University Press; 1992.
- Yockey HP. Origin of life on earth and Shannon's theory of communication. *Comput Chem* 2000;24:105–23.
- Yockey HP. Information theory, evolution, and the origin of life. In: Palyi G, Zucchi C, Caglioti L, editors. *Fundamentals of life*. Paris: Elsevier; 2002a. p. 335–48.
- Yockey HP. Informatics, information theory, and the origin of life. Fourth International Conference on Computational Biology and Genome Informatics, Research Triangle Park: Duke University; 2002b.
- Yockey HP. Information theory, evolution and the origin of life. *Inf Sci* 2002c;141:219–25.