L22 The 'RNA world' hypothesis < emergence of living chemistry >

> In a lecture delivered at Harvard University in 1969, Francis Crick declared that he had become disenchanted with the view that life arose on Earth. Instead, he favored the Theory of Panspermia, which is that life was sent to Earth from elsewhere in the Universe, and suggested that the seed of life would have been an spore of the extraordinarily resistant kind produced by *Bacillus subulis* and related bacteria. Surely, Matthew Meselson rose to point out, having gone to such trouble it send the spore it would have behooved extraterrestrial civilization to put a message in it. If only mankind was able to sequence DNA, he wistfully conjectured, then we could learn the secret of the Universe from the nucleotide sequence of the contemporary genome of the spore-forming bacterium.
>
> —James A. Hoch and Richard Losack, 1997.[1]

> ... the parts of any living cell (proteins, sugars, even nucleic acids) are just molecules–complicated molecules, to be sure, but just inanimate, dead molecules. Put them all together, and they spring into action; they become alive.
>
> —Kenneth R. Miller.[2]

Living systems expand exponentially: one (2^0) DNA (or RNA) molecule begets two (2^1), these two beget four (2^2), which beget eight (2^3), then sixteen (2^4), then thirty two (2^5), and so on. This happens because DNA (and RNA) are made up of long chains of four nucleotides (DNA is a double chain polymer of ACGT, and RNA is a single chain polymer of ACGU) each of which readily pairs off with its complement: purines attach to pyrimidines and vice versa. These connections, James D. Watson and Francis Crick discovered in 1953, structure the DNA molecule as a double-stranded right-coiling helix.[3] In cells, enzymes help split apart each DNA strand during cell division. The newly created single strands then act as templates. Each nucleotide seeks out a new partner, and these partners align to form a complementary strand, thereby creating two new double helices. This is how genes and whole chromosomes, of which the genes are a part, duplicate (replicate) themselves.

But how did DNA (or perhaps RNA) replicate before there were enzymes (***Footnote L22.1***)? In 1994, two experiments have shown that DNA molecules can replicate themselves without the help of any enzymes. Günther von Kiedrowski put nucleotide threesomes into a solution that also contained a six nucleotide strand.[3] The matching threesomes then lined up to make a complementary six-nucleotide strand. This strand, too, began serving as a template for new strands.[4]

Tianhu Li and Kyriacou C. Nicolaou have demonstrated that a palindromic sequence of 24 nucleotides (the order of purines and pyrimidines reads the same from either end of the strand) can assemble into a double-stranded DNA fragment.[5] In a slightly acidic solution, this attracted two shorter, 12-nucleotide fragments, which assembled into a third 24-nucleotide strand upon the addition of a chemical reagent. Making the test-tube solution less acidic or adding more of the 12-nucleotide fragments causes that third strand to separate from the original double strand and to act as a template for a second strand complementary to itself.

These chemical systems increase, in comparison to living systems, at an in-between rate as they add one then two then three copies and so on incrementally.

Amino acids can be made in great variety (some two hundred have been found in plants, mostly present as insect deterrents). Yet the surprise is, proteins of living organisms are assembled typically from only 20 amino acids: 10 "essential" in that these (Phenylalanine, Valine, Tryptophane, Threonine, Isoleucine, Methionine, *Histidine, Arginine*, Lysine, and Leucine) must be obtained from food (although the adult human body can make the two indicated by italics) and 10 "nonessential" in that these (Alanin, Asparagine, Aspartate, Cysteine, Glutamate, Glutamine, Glycine, Proline, Serine, and Tyrosine) are synthesized from the PVT TIM *HA*LL.

Several different codons code for the production of each amino acid and three code for stop production. The reason for 64 codons is because the genetic code alphabet has only four letters (ACGU) and codons are three-letter words (from UUU to GGG). Each of the 4^3 (= 64) codons are active. One codon (UGG) codes for the amino acids Trp and another (AUG) codes for Met. However, for Arg any of 6 codons (UUA, UUG, CUU, CUC, CUA, CUG) can be read for its production. The other amino acids have portioned to them 6, 4, 3, and 2 codons which can be read for their production. Why each codon does not code for something different is an open question.

> Everybody who's considered this problem is tantalized by the question: Is life better with a different set of amino acids? —David R. Liu.[6]

> We don't even know what having 20 amino acids gets you that 16 doesn't. —Olivia P. Judson.[7]

And indeed, the 20 amino acids encoded by the universal genetic code is not slavishly adhered to.

In human cells, a 21st amino acid Selenocystine is needed for a few proteins that cannot function without it. Some codons, that universally are read for amino acid cystine, call for it. In plant cells, mitochondria adhere to the universal genetic code, but mitochondria in other organisms have amino acids read from only one or two codons. This leaves unused some codons.

Life's limited alphabet (ACGU) may well have been set by, and thereby indicate, RNA precursors. RNA is constituted of the same four molecular building blocks. One of these, U (uracil) can bind with formaldehyde to form an amino acid precursor HMU (5-hydroxymethyluracil). Michael P. Robertson and Stanley L. Miller, suggest that formaldehyde existed abundantly on prebiological Earth.[8] In that world, a catalytic role for RNA would have been enabled by HMU "functional groups" (***Figure L22.1***, p. 698). Proteins subsumed RNA's catalytic role and perpetuated the same functional groups.

Life's earliest molecules arose when small RNA fragments came together and served as templates for longer ones. (Lab success has been to a length of 55 nucleotides.)[9] There are many possible natural scenarios. For example, clay may have been the substrate to catalyze life.[10]

A collective of two or more self-replicating molecular strands (species) can become interlinked through a cyclic catalytic (organic catalysts are enzymes) network.[11] This symbiotic group would then be selected for if its competition for resources is more efficient than for any alone member. The shuttling of enzymes, described as a hypercyclic network, is known to be in use by an RNA bacteriophage described by Manfred Eigan in 1991 and two autocatalytic chemical systems may contain vestiges of hypercyclic organizations.[12] The formation of hypercycles that confer group advantage to the symbiotic members could have accelerated the emergence of living chemistry from inanimate chemistry and, predictably, a large number of hypercycles should be embedded in the complex networks of living systems. In 1997, David H. Lee described an autocatalytic (self catalytic) reaction network of two self-replicating coil peptides that could be competitive but are not as they catalyze each other's production.[13]

Given that synthesis mechanisms for simpler precursor molecules, such as sugars and bases are known, the "RNA world" (name coined by Walter Gilbert in his article in *Nature*, 1986) hypothesis is that early-life developed using RNA molecules, rather than proteins, to catalyze biological reactions that promote assemblage of other RNA molecules, the copying of these and the mutations that they accumulate. As Gilberts puts it, "The RNA world idea is an answer to the problem of when the first information began to copy itself and make more information."[14] ❑

Footnote L22.1 In the past biology—with all its details of DNA, proteins, ribosomes and so on—has provided our only example of programmable construction on an atomic scale. But the discoveries in this book [*A New Kind of Science*, 2002] suggest that there are vastly simpler systems that could also be used. —Stephen Wolfram.[15]