

BRAINSTORMING WORKSHOP ON THE QUESTION OF ORIGIN OF LIFE

Purpose of the Workshop at CERN

The purpose of the Workshop on May 20 is to collect and bounce around ideas on the unsolved question of the Origin of Life. At this stage the aim is to share opinions, to define crucial features of the open problems we would like to tackle, and eventually converge on concrete research plans. Two years ago Nature (Vol. 459, 21 May 2009, pp. 316-319) wrote that ‘Studying the origin of life is not a great way to build a scientific career’, and we must blame ourselves for this public perception on our work. In contrast, the famous organic chemist Albert Eschenmoser once remarked: “the best minds in chemistry should now enter the origins of life field”. Spectacular advances in the field are unlikely to come in a few years from now, but there is a serious risk that the answer to the question ‘how on earth has life appeared on Earth?’ will mainly remain in the realm of philosophy for the years to come unless we can take definitive scientific approaches. We must make a serious effort to clearly separate ‘the essentials’ from ‘the nuances’.

The central purpose of our request to CERN to host this workshop is to ask for help on three fronts:

- 1) Organizational: to build an expanding team of scientists who will collaborate, yet remains adequately diverse to avoid becoming trapped by any single view, and solve the problem.
- 2) Obtain guidance how to access and provide the needed powerful computational resources, such as for an experiment like the ATLAS Project, for very large-scale simulations that are required to drive theory and experiment.
- 3) Lend its high prestige to an Origin of Life international coordinated effort that may make substantial progress in five to ten years and of course over a longer term.

Background

A brief history of the Origin of Life problem begins with Louis Pasteur, who showed using a swan-necked flask with water in the neck and broth in the flask, that the flask remained free of bacterial growth because bacteria from the air could not pass the water in the neck. All life, he said, comes from life. More deeply, his work refused the ‘spontaneous generation principle’ and suggested some essentials life has that are lacked by non-living entities. But there should be a unifying common theme between the living and the non-living to understand the evident problem Pasteur faced: How did life start?

The field lay fallow until Haldane in 1929 and a Russian scientist, Oparin, in the 1930s, proposed key ideas relevant to this question. Haldane suggested that the chemical composition of the primordial Earth was different from what it looks like now, and that without an ozone layer radiation could have provided the energy for the abiogenic build-up of complex organics. Oparin, independently, advocated the view that that jelly like coacervates could absorb small organic molecules and could divide - a possible basis for protolife.

In the 1950s, the famous Miller Urey experiment with small organic molecules simulating the hoped for atmosphere of the early earth, and with electric sparks to mimic lightning and provide energy, all in a flask, yielded a brown “scum”, that contained five of the amino acids of life. The field of abiotic chemistry was born, asking which of the small molecules of life could be assembled under prebiotic conditions. The short answer is that all could, lipids, amino acids, sugars, nucleotides, although under different conditions.

The next step in prebiotic chemistry came with the realization that carbonaceous chondrites, like the Murchison meteorite that fell in Australia in the 1970s, contained organic molecules, including amino acids and lipids. Thus infall from meteors on the early earth may have delivered some of the organic molecules that then dissolved into the “primitive soup”.

There are a number of distinct theories of **how molecular self-reproduction** may have come into existence.

1. **RNA and the RNA world.** Self reproducing single stranded RNA sequences, DNA sequences or PNA sequences. Leslie Orgel, looking at the beautiful symmetry of double stranded DNA or RNA, reasoned that a single strand of RNA, with an ARBITRARY sequence of nucleotides, call it the Watson strand, could line up free nucleotides, A,C,G,U, by base pairing, and, hopefully would be joined in p[roper 3'-5' phosphodiester bonds to create the complementary Crick strand of RNA, then the two would melt and cycle. While conceptually beautiful, the efforts over 45 years have failed. In part this is because nucleotides thermodynamically favor 2'-5' phosphodiester bonds which cannot form the double helix. In part it is because poly G precipitates. The field is gradually abandoning this still hoped for solution to the origin of molecular self reproduction. The RNA World. With the discovery that RNA sequences could act as enzymes, i.e. ribozymes, came an enormous burst of enthusiasm for the RNA World. The fact that ribosome active sites are RNA ribozymes lends support to this possibility. The major strand of work in the RNA world view has been efforts to evolve an RNA ribozyme replicase able to copy any single stranded RNA into its complement, hence able to copy itself. Partial success has been achieved by David Bartel at the Whitehead, with a selected ribozyme from among 10 to the 15th random sequences, able to copy 14 nucleotides. This effort is now stalled for whatever reasons.
2. **Self-reproducing liposomes.** Liposomes are bilipid hollow vesicles easily formed when appropriate lipids, including those from the Murchison meteorite, are dissolved in an aqueous medium. It has been shown that these liposomes can grow by addition of lipids and divide by Luisi, a member of the group. And it has been shown by Deamer, a member of the group, that they can absorb organic molecules, DNA, RNA and small peptides if they go through wet dry cycles. Most workers agree that something like liposomes as containers for further self reproducing molecular systems, are likely to be part of the origin of life on earth.

3. **Metabolism first theories.** Here Harold Morowitz, in our proposed group, has suggested with others that the famous TCA cycle of reactions, run “backwards” is a catalyst as a reaction cycle, producing two copies of one of its molecular constituents. The hope is that from this one can get a primitive metabolism going, which acts as a set of catalytic cycles and may lead to molecular reproduction. Work here is hopeful, although again without decisive evidence.

4. **Collectively autocatalytic sets, ACS,** of polymers, peptides, RNA, both, others. In 1971, Kauffman published a first paper showing that as the diversity of polymers, say linear peptides with two amino acids, alanine and glycine, increased, and if these polymers were candidates to catalyze the ligation and cleavage reactions among themselves, a diversity would be reached at which so many reactions would be catalysed that a collectively autocatalytic set would “crystallize” as a phase transition in the pool of diverse molecules. The theory is analogous to Erdos Renyi random graphs and the phase transition to a Giant Component in the graph as the ratio of arcs to nodes is increased. Here in place of arcs and nodes, we have a bipartite graph with nodes representing each kind of molecule, boxes representing reactions, arrows from substrates of a reaction leading to a reaction box, and arrows from the box to products leading from the box. Of course, the actual direction of reaction depends upon displacement from equilibrium. Under a variety of models of which polymers catalyze which reactions at significant velocity, shown as blue arcs from such catalysts to the reaction box corresponding to the reaction catalyzed, it is now a set of theorems that at some diversity, collectively autocatalytic sets arise as a phase transition to a Giant Autocatalytic Catalyzed Reaction Hypergraph. If so, molecular reproduction is an expected property of chemistry and number, and may be abundant among the Goldilock exoplanets. More, it may arise even were the constants of nature slightly different in “neighboring universes”. Conceptually, a version of this idea was proposed Doron Lancet, a member of our group, has proposed the GARD model of a population of lipids making liposome or micelle structures whose compositions can evolve as these structures divide and vary lipid content.

Autocatalytic chemical (supersystems) as evolvable dynamical systems

Another line of reasoning acknowledges that complex organic *structures* are important, and that replication of structures is an important component process, nevertheless a key question of equal weight arises in the *dynamical domain*: how could chemical supersystems, composed of qualitatively different autocatalytic components, spontaneously assembled and diversified. Eigen in 1971 realized that inaccurate replication poses a severe problem, and that competition of unlinked genome pieces must have led to the first “Intragenomic conflict”, whereas the late Gánti, also in 1971 realized the importance of coupled autocatalytic metabolic cycles and template replication, to be completed with autocatalytic membrane growth in the unified chemoton model in 1974. Kauffman, in 1971 and 1986 proposed an alternative model in which collectively autocatalytic sets of polymers, peptides or RNA for example, arise as a spontaneous phase transition as the diversity of polymers in the reaction system increases.

In the footsteps of these pioneers, if we ask ourselves the question: What is the least common denominator of all theories on the origin of life - regardless of their roots in metabolic, genetic, and compartmentalized reasoning? There is a strikingly simple phenomenon which cannot be omitted in any theory: autocatalysis. Self-sustainment = autocatalysis + transformation of matter. Self-containment = autocatalysis + spatial enclosure. Self-replication = autocatalysis + information transfer.

The COST Action CM0703 (chaired and vice-chaired by group members von Kiedrowski Szathmáry) serves as a handy background both intellectually and organizationally to the present proposal. The main objective of the Action is to investigate autocatalytic reaction systems within supramolecular, prebiotic, and other fields of chemistry and to develop methods for their integration into dynamic supersystems."

In the end, a key issue may be the quest for the **evolution of general evolvibility** - not excluding cases where the space for evolution may be as large as 2 bit: Chiral symmetry breaking and other classes of bifurcation, for example. The logics may be to learn understanding and controlling 2 bit spaces (apparently easy - but only apparently), then n-bit multifurcation, then finally open ended evolution and unlimited heredity of Maynard Smith and Szathmáry.

Eors Szathmáry and colleagues, members of our group, have shown that ACS can evolve, particular if competitive and non-competitive inhibition is included and the system is in a dividing container, like a liposome. Roberto Serra, a member of this group, has shown that ACS and liposome can evolve to divide in synchrony.

Experimental work: Perhaps the most hopeful sign is that collectively autocatalytic sets made of DNA, RNA, and peptides, each, have been made. Gunter von Kiedrowski, a member of our group, made the first autocatalytic DNA hexamer, then two hexamers, A and B, where A catalyzed the ligation of B fragments to form B, while B catalyzed the ligation of A fragments to form A. These are NOT template replication, but simple ligation reactions joining two trimers that bind the hexamer "ligase". Most astonishingly, Reza Ghadiri, and his post doctoral fellow, Gonen Ashkenazy, a member of our group, have made a 9 peptide collectively autocatalytic set from a zinc finger coiled coil 32 amino acids long, able to bind two fragments of itself, 15 and 17 amino acids long to a single autocatalytic molecule. This definitively shows that molecular reproduction need not be based on DNA, or RNA template Watson Crick pairing. The 9 peptide system is collectively autocatalytic.

Meanwhile, Stu, Luigi Luisi and Tetsuya Yomo have made libraries of random peptides. Stu's student, Thom LaBean showed that random peptides fold; this was repeated by Luisi and Yomo. More astonishingly, Yomo has just shown that 140 amino acid long polypeptides each bind to a stable analogue of an esterase transition state, and can each catalyze the esterase activity. Simple selection improves catalysis. Conversely, the probability that a random peptide length 6 - 15 binds a stable analogue of a transition state is about 1 in a million. There is some scaling law about length of random peptides and polypeptides and probability of binding an arbitrary molecular shape and catalyzing an arbitrary reaction. These tie into the theory of the emergence of collectively autocatalytic sets. New theory and experiments, interests of Hordijk, Steel, and

Kaneko, are needed to explore the conditions for the emergence of ACS in peptide or RNA libraries of stochastic sequences.

Finally, Gerold Joyce has selected from among 10 to the 14th RNA sequences and found two collectively autocatalytic ribozyme pairs.

Three major features of protolife that are missing in the above: 1) linking of exergonic and endergonic reactions, an interest of Kauffman, Serra and Szathmáry; 2) work cycles arising from such linking, an interest of Kepa Ruiz Mirazo in our group; and 3) chirality, an interest of Marcelo Gleiser in this group. The molecules of life are chiral, how did this arise? Real cells do all three of these.

Research Aim

The aim of this research group is to create theory and experiments to produce at least one, or several candidate evolving protocells in the next decade. The US has its fine Astrobiology Institute, but it is largely funding work on detecting life on other planets. We wish, of course to expand our initial group to include Astrobiology members, and others on an international basis. The current initial group of 21 includes scientists from Japan, New Zealand, USA and the EU. Time is ripe for an international group to coordinate, tackle and solve the problem of the emergence of protolife. Some members of the group firmly believe we can do so in a decade. From there to the evolution of life with DNA, RNA, coded protein synthesis will require far more work. But protolife that can evolve by Darwinian heritable variations, as Szathmáry and colleagues have shown, is a major first step.

If we do not organize the field it is in danger of drifting with uncoordinated disparate groups making bits of progress. We believe that by coordinating our efforts, without premature lock in to any single view, we can make rapid strides. For this we hope to gain from CERN experience and help.