RNA networks at the origins of life

Jessica A.M. Yeates and Niles Lehman (Portland State University, USA)

The origin of life has often been viewed as the advent of a single self-replicating molecular species, such as RNA. We propose a somewhat different approach in that a network of co-operating molecules could have kick-started life. This view has both theoretical and experimental support. The foundations for life, as we understand it on our planet, began some 4.5 billion years ago with the formation of the Earth and by 4.0 billion years ago evidence for the presence of life existed. Within that timeframe, physical and chemical processes would have produced increasingly more complex interactions, moving from simple inorganic molecules to biopolymers capable of replication and variation. In order to answer the question of how life originated and to even understand what life is, empirical proof-of-concept simple abiotic pathways demonstrating these transitions are needed. In this article, we discuss how networks of molecules, rather than single replicating molecular species, is an emerging view that may unlock some longstanding problems in the origins field.

Prebiotic chemistry

The transition from simple inorganics to biological precursors (i.e. sugars, amino acids, nucleic acids and lipids) has been recorded in the literature for over a century. The most famous of these experiments is the Miller–Urey experiment, in which amino acids were synthesized from ammonia, water, methane and hydrogen in a ‘prebiotic’ environment. Formaldehyde was demonstrated to form sugars in the autocatalytic formose reaction, and hydrogen cyanide can be used to synthesize a pyrimidine nucleobase. A route to abiotic purine nucleobase synthesis had eluded scientists for some time, yet, once produced, a nucleotide in an activated form emerged. Lipids, being another important category of biological molecules, can be synthesized through the reduction of sugars catalyzed by metal or mineral surfaces through a primitive fatty acid synthesis and can be delivered extraterrestrially.

It should be noted that the above processes typically represent low-yield multi-component racemic products, and are far from giving the highly specific building blocks of modern day biology. Yet, many of these biological building blocks have been detected on meteorites indicating that they are more than plausible products. Additional studies have demonstrated pathways to increased specificity and stereoselectivity, for example through the use of borates, lipids and phosphates. Recently, a flow chart of synthesis reactions demonstrated how all major classes of biological precursors could have been interconnected and produced (Figure 1). This diversity and co-operation of molecules and reactions truly invokes the imagination when considering how chemistry can spontaneously build complexity.

Whereas abiotic synthesis of biological precursors and monomers represents one challenge that nascent life overcame, condensation of these units into functional polymerized biomolecules represents a second challenge. Again, we see precursors harnessing the functionality of the environment and/or other molecules. Abiotic oligomerization has been demonstrated using clays, polymer templates (nucleic acids or proteins) and lipids. These processes produce in the laboratory, at most, the synthesis of an oligomer approximately 50 subunits long (i.e. a ‘50-mer’). For RNA, this length is actually sufficient for some small or minimal ribozymes and some have proposed that only 22 nucleotides were needed to lead to the evolution of the ribosome. Yet, a path from functionality to an evolving self-replicating chemical system is still very elusive.

There has been vast progress in making longer oligomers through classic polymerization. A ribozyme has been evolved in a test tube to elongate an RNA strand to over 200 nucleotides. Although this length is impressive and exceeds the length of the ribozyme itself, the product does not represent a self-replicating sequence. An adjacent line of research...
uses the premise that collections of interdependent molecular species – or chemical networks – may have been a more robust pathway to life. A network is a collection of species (nodes) that are connected via interactions such as reactions (edges). Figure 2 shows a simple molecular interaction network.

This concept is not new; as early as the 1920s A. Oparin suggested that a bag of interacting organic chemicals (coacervates) could have led to the first lifeforms. Even the aforementioned work on monomer formation and abiotic polymerization demonstrates the usefulness of multiple molecular species interacting (e.g. template interactions). Theoretical work on chemical networks has been of great interest over the last few decades. Eigen and Schuster proposed the hypercycle and quasispecies, T. Gánti introduced the chemotron, and Kaufmann conceived the collectively autocatalytic set (CAS). These ideas have been very influential in the field.

Yet an empirical approach along this pathway has been much more difficult to implement. This could be due to the traditional reductionist philosophy within chemistry or a lack of technology/methodology for multi-component interactions. Regardless, basic cellular life consists of vast interconnected networks that produce emergent phenomena not easily predicted by the molecular species treated in isolation. An empirical understanding of how simple molecular networks develop and evolve is now needed to add critical insight to the non-life-to-life pathway.
Networks

Molecular networks have recently become a focus towards understanding abiogenesis. In this regard, we imagine a collection of co-operative molecules acting as a unit of origin and evolution rather than a single selfish replicating entity. The very nature of a network’s co-operative and distributed functionality imparts several advantages in the ‘harsh’ prebiotic world. Distributing catalytic activity among members provides robustness against the loss of entities. Co-operative replication among a group of molecules may represent a more facile process than self-replication arising in a single entity and introduces a solution to the error threshold. Lastly, co-operative molecular networks are not just possible and represented in the state of modern life but have demonstrated fitness benefits and selection preferences as compared with selfish entities.

Theoretical work towards understanding a chemical network manifestation of life has looked extensively at autocatalytic sets or collectively autocatalytic sets (CASs) as originally presented by Kauffman. Here the basic concept is that there is catalytic closure of all members of the CAS such that the set can self-replicate. Catalytic function need not reside in any one member, but each member has at least one of the possible last steps of its formation catalysed by some other member of the set. The CAS concept has been extended and the importance of raw materials (i.e. ‘food’) in the environment has been mathematically formalized as a reflexively autocatalytic and food-generated (RAF) set. CAS-like systems were shown further to be relevant in abiogenesis studies with the demonstration that they could arise spontaneously from a simulated pre-life chemical environment. Here, the pre-life scenario was interconnected undefined molecular species able to catalyse the ligation of reactant molecules to produce other catalytic species. The appearance of autocatalytic sets was shown not only to be inevitable, but also to be able to recover from stochastic frequency fluctuations and extinctions.

Although a CAS-like system does embody the self-sustaining aspect of life (with life defined as a self-sustaining chemical system capable of Darwinian evolution), an autocatalytic set itself has not yet been determined to be evolvable. However, an autocatalytic set as part of a chemical network containing multiple autocatalytic sets could be selectable. If these sets are compartmentalized further, competition could exist intra- and inter-compartmentally.

Despite molecular networks potential role in ushering in biotic systems from abiotic systems, concrete chemical and physical dynamics of such networks are still lacking. Empirical work has demonstrated the formation of networks under thermodynamic and kinetic control. Although thermodynamically controlled networks are not discussed here, the experimental manipulations of network constituents at equilibrium using templating and molecular recognition are interesting and may have implications for molecular evolution. Living systems, however, operate far from equilibrium and thus a kinetically controlled network, such as a CAS, is fundamental in understanding the chemical to biological shift.

Empirical manifestations of autocatalytic sets have been demonstrated in a simple two-ligase system by Joyce. Our own laboratory has produced much larger networks based on recombining ribozymes (Figure 3). The system that we used to study these chemical evolutionary dynamics is based on past work with the Azoarcus group I intron. In vivo, this ribozyme splices itself out of a pre-tRNA sequence with catalysis relying on binding of a three-nucleotide internal guide sequence (IGS) with a tag sequence. A decade ago, we determined that this ribozyme can be fragmented into shorter oligomers, which can assemble non-covalently to restore catalytic activity (Figure 3). By placing tag sequences at loop regions that are fragmented, these non-covalent complexes can then recombine with other non-covalent complexes into fully functional covalent ribozymes in a reaction akin to the reverse in vivo function.
Figure 3. Recombinase ribozyme from *Azoarcus* group I intron. (a) Secondary structure illustrating the division of the ribozyme into four fragments: W, X, Y and Z. The IGS (red) is located on the 5’ end of the W fragment, tag sequences (orange) are located on the 3’ end of fragments, and the junctions where catalytic closure occurs are outlined in orange. (b) Reaction scheme of the recombinase ribozyme. W, X, Y and Z first associate via base pair interactions to form a non-covalent WXYZ molecule. The non-covalent WXYZ is then catalytically connected through three separate reactions at the W–X junction, X–Y junction and Y–Z junction by other molecules of non-covalently or covalently linked WXYZ recombinases (colour scheme is the same as in a). (c) A close-up view of the phosphor-ester transfer reaction connecting the Y–Z junction. (d) Small three-membered network. Reproduction of Figure 1b in Vaidya et al. (e) Large 48-membered network. Reproduction of Figure 3d in Vaidya et al.
When the nucleotides of the IGS and tag of the *Azoarcus* ribozyme fragments are varied, different molecular genotypes are created that differentially interact with one another. Thus they can essentially form small networks, depending on these sequences. Importantly, as reaction time goes on, the composition of these networks changes, such that a crude form of evolution can be seen\(^{13}\) (Figure 3). Thus the concept of a prebiotic evolving network has been demonstrated experimentally\(^{14}\), and it has also been mathematically modelled\(^{15}\). Although this work has shown the capability of interacting molecular species to act co-operatively and form networks, the underlying dynamics behind connections, topology and how selective forces could act is largely unknown.

What is known is that network dynamics (from neural, metabolic, social and the World Wide Web) ultimately derive from local structures (i.e. connectivity kinetics between the entities and their links), environmental interactions and long-range topological structures of the network. Thus it would seem that molecular networks could be characterized within the same framework. Recently, we presented a review of six key parameters that could influence and direct the evolvability of networks with the goal of experimental exploration in mind. These parameters include viable cores, connectivity kinetics, scalability, information control, resources, and compartmentalization\(^{16}\).

**Future work**

It is clear from previous work that autocatalytic networks are characterized by cores, where a core is a subset of nodes in which each node is reachable from any other, and *viable cores*, which are cores that are also self-sustaining (providing the food source)\(^{11,16}\). However, it is not known how the connectivity of entities within a core affects growth or how the strengths of connections can affect the ability of a network to add or subtract members (i.e. the connectivity kinetics parameter). A key hurdle to overcome will be to show how small molecular networks possess heritability; in other words, how does the information in a network be transmitted as a function of time? This must be by some sort of process unusual from the standpoint of contemporary biology; it will be inherently ‘non-Darwinian’. But if a simple version of heritability can be understood and demonstrated, then we will have critical insight into how simple abiotic molecular networks could have later given rise to the type of biology – based on discrete individuals – that we are all familiar with today.

---

**References**