SYNTHETIC CONNECTIVITY, EMERGENCE AND SELF-REGENERATION IN THE NETWORK OF PREBIOTIC CHEMISTRY

* We still have only a fragmentary understanding of how these individual organic reactions steps combined into complete synthetic pathways to generate life’s building blocks, which other abiotic molecules might have also formed, how independent reactions gave rise to chemical systems, and how membranes encapsulating these systems came into being.
* In recent years we systematized the knowledge about reactions that can be performed under consensus prebiotic conditions, as well as the plausible synthetic routes leading to life’s key molecules.
* In this study the researcher uses large-scale silico network analyses to map the space of molecules synthesizable from basic prebiotic feedstocks, quantifying the structure of this space as well as the abundances and thermodynamic properties of its members.
* After that the researchers demonstrated three notable forms of chemical emergence:
  + The molecules created within the network can themselves enable new types of probiotic reactions, which include multicomponent transformations that lead to complex and useful organic scaffolds.
  + Within a few synthetic generations, simple chemical systems begin to emerge;
  + And last that the network contains prebiotic routes to surfactant species thus outlining a path to biological compartmentalization.
* The researchers demonstrate a self-regenerating cycle of iminodiacetic acid (IDA) that complements prebiotic autocatalysis on the basis of the formose cycle.
* Allchemy’s “Life” module uses 614 reaction rules (“transforming”) involving C, O, N, S, and P elements, grouped within 72 broader reaction classes.
* All of these transforms generalize, are broader than the underlying literature precedents, and are coded to take into account the underlying reaction mechanisms.
* This approach has been recently validated experimentally and is less prone to yield chemically problematic predictions than either machine rule extraction or ab initio methods.
* The researchers transform account for reaction byproducts and specify the scope of admissible substituents, structural motifs the scope of admissible substitutes, structural motifs incompatible with a given reaction, typical conditions accepted in prebiotic chemistry, solvents, temperatures, and more.
* Yields for each type of reaction are approximated on the basis of statistics collected from relevant publications and are categorized as trace (<=3%), low (>3% to <=10%), moderate (>10% to <80%), and high (>=80%).
* The analyses are for H2O, N2, HCN, CH4, and H2S substrates and the corresponding C-, O-, N-, and S- based transforms.
* These substrates were chosen because they are the starting points of many OL studies, and are thought to be the components of Earth’s early atmosphere, and are sufficient to build many common molecules of life.
* The researchers do not include Allchempy-supported P-based chemistries because in this study resultant searchers generate large numbers of chemically distinct scaffolds.
* After each iteration, the newly created molecules were combined with the products of preceding iterations and with the original substrates, and the cycle was repeated until a user-defined limit of synthetic generations.
* The researchers began their quantifying the molecular composition and synthetic structure of the C-, O-, N- and S- based network up to the seventh synthetic generation (G7), with molecular mass limited 300 g/mol.
* In synthetic space which was generated on a simple/standard computer within 2 hours, there are 82 biotic molecules and 36,603 abiotic molecules.
* They find that the biotic compounds are more thermodynamically stable than the synthesizable abiotic compounds with similar masses, are, on average, less hydrophobic than those in the abiotic pool, and are more balanced in terms of hydrogen bond donors and acceptors.
* The biotic molecules also contain fewer reactive groups and fewer distinct functional groups per molecule.
* Pathways leading to biotic molecules entail fewer condition changes than those leading to abiotic ones.
* The network is robust in the sense that removing as many as 34 of 63 C-, O-, N-, and S- based reaction classes still allow for the synthesis of all biotic molecules via bypass routes.
* It has a high degree of robustness which reminiscent of metabolic networks and could indicate that the researchers’ network has similar scale-free architecture and characteristics like internet, airline networks, or some other published organic reactions.
* They trivially expect the network to contain the known prebiotic pathways leading to all of these biotic compounds.
* Analysis of the network in comparison with known literature reveals that most routes to these biotic molecules are a patchwork of steps reported in different publications, some of which are not concerned with OL issues, but all performed under prebiotic conditions.
* The reaction of hydrolysis of fumaronitrile is relatively unimportant, as it exchanges the order of hydrogenation and hydrolysis steps from that in a known method of producing succinic acid.
* The experiment yield depended on time and acid concentration and ranged from 8% in 0.1 M HCI (9 days) to 54% in prebiotically less likely 5 M HCI (1 day).
* Another reaction of hydrolysis of nitrile groups which are more consequential to establishes a route to pyruvic acid that does not require sulfur. It is also shorter than sulfur contain route (that is sulfur contain route take 8 steps where as it takes only six steps).
* Another reaction involves tandom hydrolysis and decarboxylation of cyanoacetaldehyde to produce acetaldehyde. It is the reaction of interest because it enables a sulfur-free synthesis of lactic acid in just five steps.
* The researchers consider the synthesis of citric acid (CA), in this their network analysis suggested that CA could emerge under prebiotic conditions in water from two equivalents of oxaloacetic acid via a tandom aldol self-boxylation sequence followed by a second decarboxylation.
* The researchers observed that at their milder conditions, they obtained CA in ~5% yield, whereas under harsher condition, the citroylformic acid substrate gradually decomposed, reducing the yield to ~2%.
* After that the researchers validate a computer-predicted synthesis of diglycine from *N*-carboxyanhydride (NCA) in a sulfur-rich environment, it is under the different condition describe by Bartlett and Jones.
* In this the first step - thiolysis of glycine *N*-carboxyanhydride, was performed at room temperature in the presence of H2S and potassium carbonate and proceeded in 40% yield.
* The resulting thioacid was then oxidized with K3[Fe(CN)6] to a disulfide, which spontaneously underwent an intramolecular nucleophilic acyl substitution rearrangement followed by hydrolysis. The yield for the two thioacid-to-diglycine steps was 29%.