Radiolysis Generates a Complex Organosynthetic Chemical Network

## Zachary R. Adama,b, Albert C. Fahrenbachc, Sofia Marie Jacobsond, Betul Kacard,e and Dmitry Yu. Zubarevf

a Department of Planetary Sciences, University of Arizona, Tucson, AZ 85721

b Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA

c School of Chemistry, University of New South Wales, Sydney, NSW 2052, Australia

d Department of Molecular and Cellular Biology, University of Arizona, Tucson, AZ 85721

e Department of Astronomy, University of Arizona, Tucson, AZ 85721

f IBM Research - Almaden, 650 Harry Rd., San Jose, CA 95120

Correspondence may be directed to: Zachary R. Adam, Department of Planetary Sciences, University of Arizona, Tucson, AZ 85721, zadam@arizona.edu

## ABSTRACT

Origins of life chemistry has progressed from seeking out the production of specific molecules to seeking out conditions in which macromolecular precursors may interact with one another in ways that lead to biological organization. Reported precursor synthesis networks generally lack biological organizational attributes. Radical species are highly reactive, but do their chemical reaction networks resemble living systems? Here we report the results of radiolysis reaction experiments that connect abundant geochemical reservoirs to the production of carboxylic acids, amino acids, and ribonucleotide precursors and study the topological properties of the resulting network. The network exhibits attributes associated with biological systems: it is hierarchically organized, there are families of closed loop cycles, and the species and cycle histograms exhibit heterogeneous (heavy-tailed) distributions. The core cycles of the network are made possible by the high reactivity of radical species such as H and OH. Radiolysis is implicated as a unique prerequisite for driving abiotic organosynthetic self-organization.

## INTRODUCTION

At what point in the chronology of prebiotic chemical compound synthesis did life’s organizational attributes arise? The question strikes to the heart of identifying the circumstances that were likely to have fostered the origins of life. Long after life’s emergence, the circumstances that enable ‘living systems’ with different levels of organization to exist is an intricate amalgamation of abiotic and biotic chemical processes. The network of links between these processes can be informatively studied separately from the synthesis of any particular chemical compound (1). It is in this light that network topology opens a potential investigatory approach that is amenable to describing complex chemical systems across a spectrum of living and non-living possibilities (2-4). By dint of describing the relationships of chemical objects and processes to one another, it is possible to informatively assess whether the organizational attributes of life can arise from a preceding abiotic system (5, 6).

In this paper, we outline the chemical reaction network structure of a radiolytic system of gaseous, aqueous and mineral-derived reactants. Our results describe an enclosed abiotic system driven by powerful photonic and particle irradiation and requires no external exchange beyond the dissipation of waste heat. Regardless of whether it was complicit in the historic process of Earth life’s origins, these radiolytic circumstances outline a candidate for a self-generating kernel of an expansive (and possibly open-ended) organosynthetic system with relevance to the broader field of complex systems (7-9).

### Network Properties of Biological Systems

Complex far-from-equilibrium systems such as biological entities have many interacting and interdependent components and large degrees of freedom, and therefore have access to a multitude of dynamical states as they change over time. Such systems are capable of generating self-organizing phenomena on a macroscopic scale: states of matter capable of ‘evolving’ via the production of state-dependent feedback mechanisms that regulate internal relationships between objects (10-12).

Network theory has been developed as a mathematical tool that can be used to assess the attributes of complex far-from-equilibrium systems. A network is constructed by counting up all unique object types within a given system and graphing each as a node, and then mapping out all of the ways that these nodes are connected to one another by connecting these nodes with lines, which are called edges. The nodes may be molecules, proteins or organisms, and their corresponding edges may consist of a list of chemical reactions, protein-protein interactions or documented instances of ecological interaction. Mapped network interactions can offer clues to how the laws of physics and chemistry can give rise to emergent features (13). By mapping the relationships of chemical objects to one another, it is possible to capture elements of complex hierarchical organization, feedback and order that are otherwise missed (14, 15). Most importantly, the observed topologies of networks give us clues about their evolution and how their organization influences their dynamic responses (16).

Living systems exhibit persistent network attributes, the most obvious being composed of distinct, hierarchically-nested levels (17). Each of these levels is populated with non-identical objects of unique internal architecture that operate at vastly different spatial and temporal scales, but the relationships between objects can nevertheless share general organizational similarities (18). The lowest level consists of interacting metabolic and genetic polymers within cells (19). Intermediate levels consist of populations of cell groups, ecologies of population groups, and biomes of ecological groups, with all of these levels capped at the largest scale by the global biosphere of biome groups that are recognizable over millions of years (20, 21).

Within and across each of these levels, objects interact with each other according to a heterogeneous distribution that appears as a gradually decreasing slope (i.e., a ‘heavy tail’) on a connectivity histogram (22, 23). An observable is considered to be heavy-tailed if the probability of observing extremely large values is more likely than it would be for an exponentially distributed variable (24). Objects at each level are also continually cycled through formation and degradation in closed loops to maintain far-from-equilibrium physicochemical configurations (25-27).

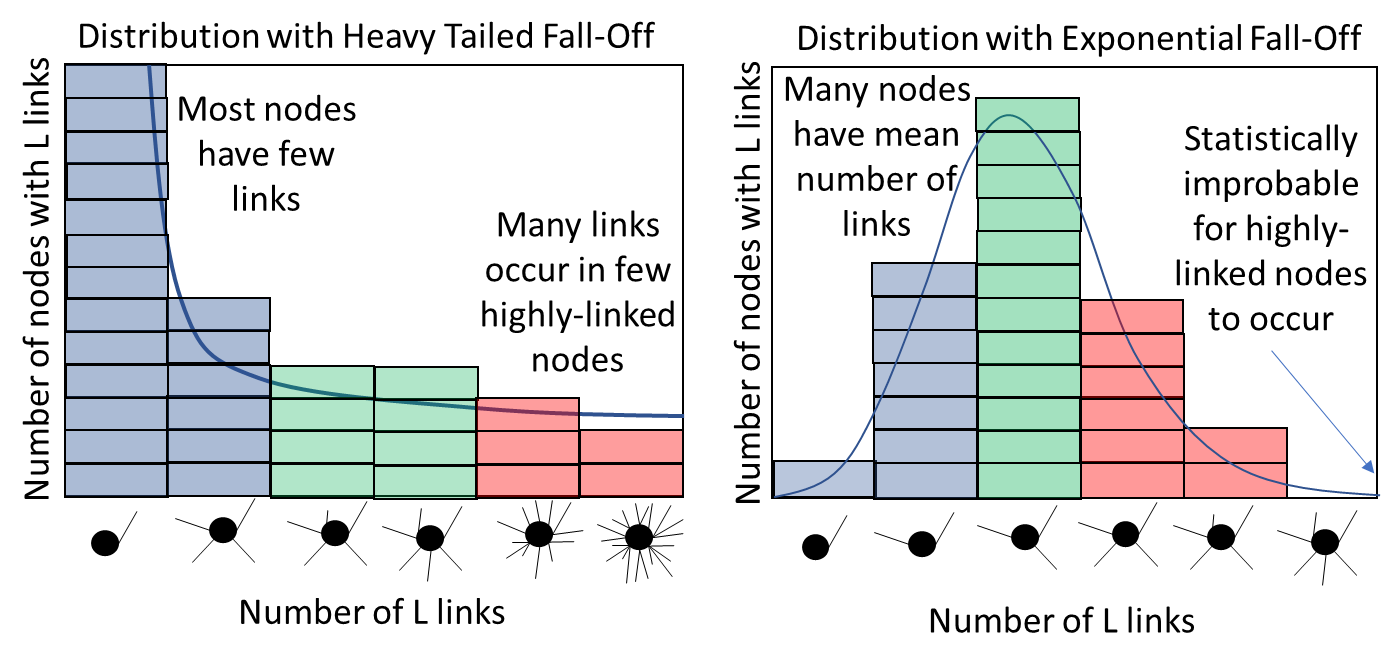
In this view, life’s origins may be recast as a search for biological network attributes within an organosynthetic chemical system (28). The challenge becomes to identify circumstances that yield an organosynthetic system with life’s network properties (e.g., hierarchical organization, closed cycling of objects, and heavy-tailed connectivity distributions) that initially lacks any these relational attributes, and which does not possess these attributes as a result of human intervention.

### Heavy-tailed Distributions in Complex Systems

The underlying generative mechanisms that yield heavy-tailed network topologies, and their implication with capacities for complex behaviors, are active areas of research. Two common modes of heavy-tailed distributions, power law and lognormal scaling, can arise from relatively simple (and under some circumstances, overlapping) generative processes. A generative mechanism known as 'preferential attachment' is frequently invoked as a model for the growth of networks with a power law-scaling degree distribution (29). Preferential attachment refers to a mechanism in which new interactions in a network are more likely to be made with nodes that already have a large number of connections (30). Lognormal distributions can arise as a result of multiplicative products of many independent, identically distributed, positive random variables (31). Under conditions in which the outputs of multiplicative products have bounded minimums, or in which observation times are random, a system generated by multiplicative products can exhibit a power law distribution (32); the generative mechanisms of power law and lognormal distribution models may be viewed as closely related rather than mutually exclusive modes of network growth.

There are many claims of consistent power law scaling of connections between biological entities (13, 19, 23). The uniformity of biological power law scaling relationships has been reconsidered in light of more rigorous statistical treatments, which show that lognormal or other related heavy-tailed distributions are in some cases a better fit (33, 34). In practice, finite or potentially biased sources of information describing real-world networks make it statistically difficult to discern whether a distribution exhibits a power law, lognormal or some other heavy-tailed topology; knowledge of whether or not a distribution is heavy-tailed is far more important than whether it specifically fits a power law (35).

In a network with a heavy-tailed distribution (Figure 1, left), there are few objects connected to many other objects (shaded in red), there are a some objects that are moderately connected to other objects (shaded in green), and the majority of objects are connected to few other objects (shaded in blue). This heterogenous distribution stands in marked contrast to networks in which connections between objects are made at random and with a fixed probability (Figure 1, right). In this case, a statistically homogenous structure arises in which the probability that a node is connected to other nodes is proportional to p(k) = e-λk, where λ describes the rate at which the probability of forming an additional attachment declines with k (29). In statistically homogeneous networks, many nodes possess the mean number of connections, and the probability of a highly connected object occurring is vanishingly small. The generative mechanisms that lead to heterogeneous and homogeneous distribution modes imply markedly different capacities for complex response (36, 37)



**Figure 1. A network with a heavy-tailed distribution (left) has few nodes with a high number of connections (shaded in red), some moderately-connected nodes (green) and a large number of poorly connected nodes (blue). A network with a distribution with an exponential fall-off of connections across nodes (right) clusters around a mean value of links.**

### Why Radiolytic Chemistry?

We hypothesize that conditions that enable organic compound reactivity without leading to thermal degradation can produce a complex organosynthetic system; radiolysis leads to such conditions. Radiolysis is the use of energetic particles to dissociate chemical compounds into fragments called radicals, which can readily react with each other and organic molecules (38). Radicals have unpaired outer valence electrons and can be more reactive than most chemical compounds at standard temperature and pressure (39, 40). High energy atmospheric chemistry is already implicated with heavy-tailed distributions (41) and hierarchical chemical organization (4), but there have been no reported network-level assessments of chemical systems that specifically lead to prebiotic macromolecular precursors.

Chemosynthesis via radiolysis represents an efficient means of connecting geochemistry to macromolecular precursors. Recently, for example, Adam et al. (42) demonstrated that radiolysis of water and nitriles produce the RNA nucleobase precursor formamide, but irradiation does not concomitantly increase its rate of degradation into other compounds. Yi et al. (43) went further, demonstrating that reducing and oxidizing reactions can occur in the same volume without the need for compartmentalization, and that radiolysis of aqueous chlorides can serve a critical role in yielding RNA precursors. These reactions appear to be hierarchically organized (44), with base-level reactions involving highly reactive, short-lived radical species such as H and OH driving the production of successively more stable secondary compounds of increasing molecular mass.

These chemosynthetic properties of radicals (asymmetric productivity, redox promiscuity and incipient hierarchy) should promote heterogeneous connections between compounds in a standalone system. We predict that an assessment of radical equation data will indicate a network in which links between feedstock molecules (CO2, H2O, N2, NaCl, chlorapatite and pyrite) and macromolecular precursors will exhibit network attributes that parallel biological organization. Specifically, the disparity between rapid electron displacement and slow-paced molecule-molecule reactions should facilitate physical prerequisites for self-organized criticality (11, 23, 45-47).

## METHODS

Chemical reaction network distribution data were collected by collating chemical reactions from radiolytic and polar chemical research findings that include all of the key atomic species of living systems. A complete list of these reactions and their source materials is provided as a supplementary data file. Atoms are initially present in the system as their most geochemically abundant molecular species (H2O, CO2, N2, NaCl, a proxy for the apatite mineral group Ca10Cl2(PO4)6, and pyrite FeS2). The electromagnetic spectrum was binned into photonic ‘species’ (gamma rays, X-rays, UV rays, visible and infrared) to account for interactions between photons and chemical species across key atomic and molecular energetic thresholds (i.e., strong nuclear binding force, inner valence electrons, outer valence electrons and ambient system temperature). Hydrated species complexed with water and reactions contingent upon an aqueous medium include water as a bystander species. Amino acid source polymers are chemically uncharacterized, and are binned into polymer placeholder species that can be thermally hydrolyzed to yield amino acids. Reaction rates and kinetic coefficients are omitted from all reactions; this version of the network illuminates network topology only, it is not intended to yield a dynamic simulation.

The energy sources driving physicochemical system response was presumed to be forms of radiation sourced from radiogenic uranium oxide. Uranium is a geochemical proxy for naturally-occurring energy sources such as galactic cosmic radiation, solar flares and planetary radiation belts. The assembled database includes 782 reactions and 386 distinct physical entities (i.e., atomic, molecular and photonic species).

The acquired reaction equations were used to construct a reaction network where nodes are divided in two sets, i.e. partitions, and connections are allowed only between nodes in different partitions. Networks of this type are called bi-partite networks. The first partition in the constructed network includes reaction participants, such as atomic, molecular, and photonic species; the second partition includes reactions themselves. The network is constructed as a directed network where directionality reflects the form of the stoichiometric equation: the entities on the left-hand side of the equation are connected to the node representing the reaction, and the reaction node is connected to the entities on the right-hand side of the equation. This definition of directionality simplifies network analytics, such as cycle search. It is not reflective of the reversibility of the chemical reaction; the information pertinent to the rates of the forward and reverse reactions, equilibrium constants, and such, can be stored as an attribute of the reaction node.

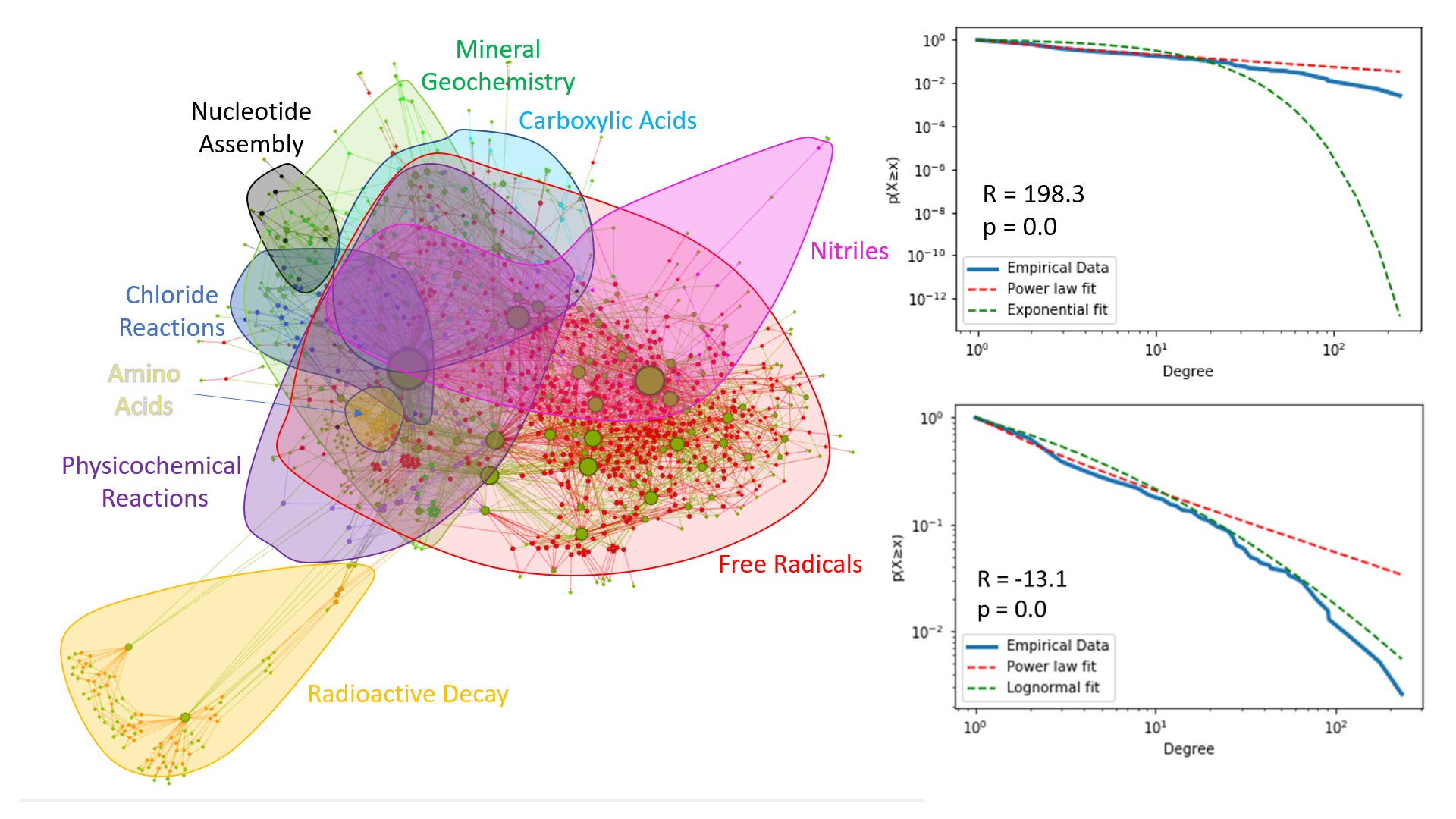
The topological attributes of the resulting network were analyzed and visualized using the Gephi 0.9.2 program (48). A comparison known as a likelihood ratio test (R-value) was employed to evaluate whether the dataset exhibits a heavy-tailed distribution by comparing a power law fit (the simplest approximation of a heavy-tailed distribution on a finite interval) to one with an exponential fall-off. A statistical correlation test (p-Value, with p>0.05 cutoff) was used to make a quantitative judgment about whether the observed value of R is sufficient to assess whether sampling fluctuations from within the dataset would undermine the indicated likelihood ratio result (49, 50). If a heavy-tailed distribution was detected, then the likelihood ratio test and statistical correlation test were also used to ascertain whether a power law, log normal or exponential truncated power law was more likely to be a better fit.

Closed loops of chemical species in the directed bi-partite network were identified through exhaustive enumeration between pairs of nodes by finding the shortest paths from the first node to the second and all the shortest paths from the second node to the first. Pair-wise combinations of the forward/backward paths would constitute distinguishable cycles. Enumerated cycles were expressed as sub-networks formed by these forward-backward paths; each sub-network was labeled with an ordered sequence of the indices of the included nodes to identify unique subnetworks. In this representation, sub-networks included only the nodes along the shortest forward-backward paths; we also produced the expanded sub-networks that included all the components of the reactions along the loop, even if the components were not in the direct paths of the loop.

We recognize that with the total number of reacting entities (~102), the analyzed degree distributions offer extremely limited insight into the scaling behavior associated with the right tails of the degree distributions; we also recognize that scaling fitting is a common practice in consideration of the networks regardless their size. Our goal is, therefore, to bring attention to appropriate techniques of statistical and organizational analyses and to demonstrate the advantages and limitations of each as they relate to putative prebiotic systems.

## NETWORK TOPOLOGY ANALYSIS

A plot of the complete network weighted with a ForceAtlas2 algorithm is depicted in Figure 2. Radical species derived from chemical inputs drive secondary radical and conventional redox reactions that can be grouped by chemical category (i.e., nitrile radical reactions, chloride reactions, etc.; Figure 2, shaded areas). The radiolytic network has a diameter of 28, with an average directed path length of about 5.4 steps. Carboxylic acids are produced directly from radical reactions. Amino acids are produced through a combination of radiolysis, heating and acid-hydrolysis. Nucleotide precursors that lead to cytidine monophosphate (CMP) are all present within the network. Gypsum (CaSO4.2H2O), a moderately soluble phase produced by oxidation of pyrite and subsequent acidic dissolution of apatite, serves as a net sink for Ca, S and O atoms.



**Figure 2. Left: Network visualization with molecular species (green circles) size-weighted according to degree of connectivity to reactions (red circles). Groupings of chemical reaction types are shaded and labeled. Right: A statistical comparison of reaction component connectivity (top) indicates that an exponential fall-off fit is poor compared to a power law fit, indicating a heavy-tailed distribution. Among the two most common classes of heavy-tailed distributions (bottom right) a lognormal distribution fits the data better than a power law distribution.**

A cumulative probability of the node degree for the reaction components found in the system (Figure 2, right) was used to assess whether the node degree distribution within the reaction component partition more closely follows a distribution with an exponential fall-off or a power law scaling relationship, which is the simplest form of a heavy-tailed distribution. A list of the most connected chemical species is provided in Table S1. The topology matched a power law (α ~ -1.5) distribution with a p-value of 0, and the likelihood ratio of a power law to an exponential was much greater than 0, indicating a power law distribution was statistically more likely compared to an exponential fall-off (Figure 2, top right). A similar comparison of heavy-tailed distribution types between power law and lognormal distributions indicated that a lognormal distribution was a better fit than a power law distribution (Figure 2, bottom right; p=0, R<<0). Lognormal and exponentially truncated power law fits were statistically indistinguishable from one another with the available data.

Over 2,600 distinct closed loops of chemical species were identified within the network. Chemical species ranked by frequency of presence within closed cycle subnetworks is provided in Table S2. Comparable analyses of the cycles sizes (number of nodes in each cycle, inclusive of participating chemical species and reactions) also fit a heavy-tailed distribution, but the specific types of heavy-tailed distributions were statistically indistinguishable from one another (Figure S1).

## DISCUSSION

### Heavy-tailed Compound Connectivity Distribution

The topological attributes of this radiolytic network outline chemical conditions that have been sought for plausible prebiotic settings (2). The likelihood ratio test (R-value) indicates that a realized version of this radiolytic chemical system would exhibit heavy-tailed distributions in both compound connectivity and compound cycle sizes. Heavy-tailed distributions are intimately associated with far-from-equilibrium systems and self-organizational properties (23).

The most connected network hub is water (234 total degree connections; Table S1), but the next largest hubs are the low-mass, highly reactive reducing H (165 connections) and oxidizing OH (122 connections) species, which are readily produced by radiolysis of water. The most powerful forms of energy in the system (gamma rays, alpha particles, beta particles, X rays and UV light) are also major hubs within the system.

Two general observations are apparent based on the layout of the network. The first is that the subnetwork of radical interactions (Figure 2, shaded in red) dominates the core of the network and represents an efficient way to connect the other subnetworks to each another. It is not clear whether a network composed of non-radiolytic reactions could connect these subnetworks in a comparable manner.

The second observation is that water can serve as an efficient mediator of far-from-equilibrium energy, in addition to its recognized role as a critical solvent for prebiotic chemistry. Water-derived radical species (H, OH, solvated electron, etc.) dominate the center of the network, yielding some higher mass organic compounds over others, such as in the case of simple hydroxyaldehyde production from hydrogen cyanide via Kiliani-Fischer homologation (43), and the synthesis of 2-aminooxazole and 2-aminoimidazole from the reaction of glycolaldehyde with cyanamide (44).

The most energetic and frequently interacting physical and chemical species possess little or no carbon or nitrogen, but these species in turn drive the production of species with higher numbers of these atoms (Table S1; Figure S2). Under a facile first-order assumption that the relative production of a compound within the system is correlated with the number of connections of that compound, the network topology indicates that heterogeneous concentrations of low-mass chemical species lead to heterogeneous concentrations of higher mass organic species within the system over time.

### Closed-loop Cycling

Repeated interaction of reducing and oxidizing species creates closed loops of molecules (Figure 3). At a basal level, energy-driven loops are composed of low mass radicals directly produced from radiolysis and recombination of feedstock molecules (e.g., N2, CO2, H2O, etc.) such as H, N, O, OH, NO, etc. These cycles occur closer to powerful energetic particles than do higher mass compounds. Closed cycles of these species can form trivial closed systems with no net production of chemical species beyond the basal reactive species themselves and output photons.

The coupling of these basal equations and cycles to one another, and to the initial substrates, undergirds the emergence of more complex cycles with net reactions that do not reduce to other reactions within the network. As energy is inputted into the system, a facile return to the initial state becomes less likely as the numbers of species and possible chemical reaction pathways increase.

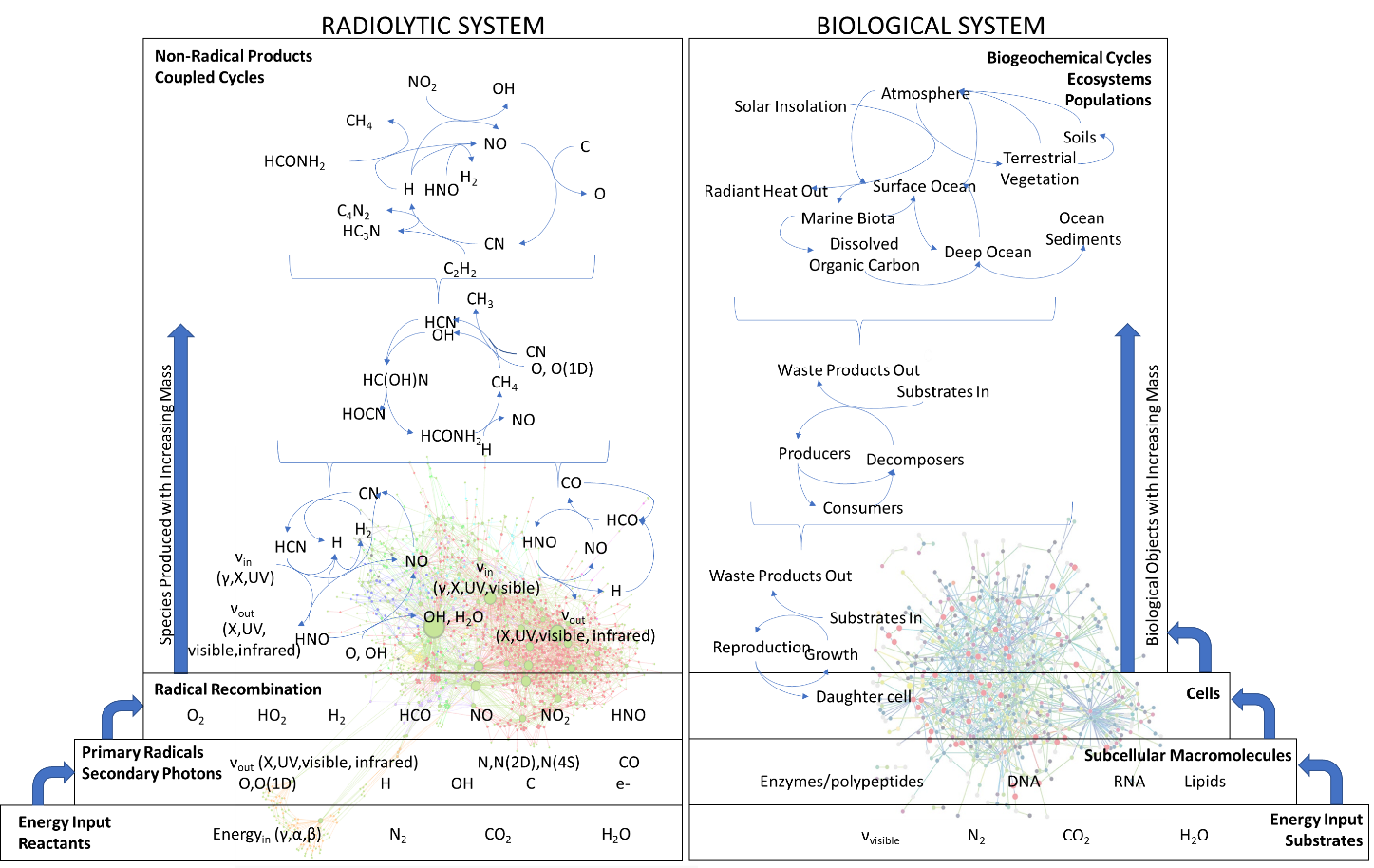
A close up of text on a white background

Description automatically generated

**Figure 3. Four examples of subnetworks of cycles identified within the radiolytic network with net reactions that do not reduce to reactions present elsewhere within the network. Subnetwork images generated using Gephi version 0.9.2.**

### Hierarchical Stratification

Distinct levels can be resolved via inspection of the most prolifically connected species and their placement relative to the feedstock compounds (Figure 4). The link between low-mass species heterogeneity leading to higher-mass species heterogeneity, and the trend of decreasing reactivity and lower G-values associated with higher-mass species (44), implicates an incipient chemical hierarchy in which a larger-scale objects emerge from relationships between and amongst populations of many, smaller-scale objects. This is corroborated by direct observations from individual radiolytic experiments for which quantitation data are available (42) and the atmospheric chemistry of numerous solar system bodies (4).



**Figure 4. Selected examples of network-level organizational features of radiolytic systems (left) and biology (right) including hierarchical stratification, closed-loop cycling and heavy-tailed distributions of object connectivity. Heavy-tailed biological network figure from Titz et al. (51).**

### GENERATIVE MECHANISMS THAT UNDERLIE RADIOLYTIC NETWORK TOPOLOGY

A fit of the degree distribution of the reaction participants to a power law yields an exponent value close to -1.5. Heterogeneous systems with an approximate power law relationship of -1 are implicated with self-organizational behaviors found in the vicinity of a phase transition between stable and chaotic states (11, 45-47, 52). Despite the generality of this attribute, there is no universal prescription for the generative mechanisms that yield such distributions; physical and chemical processes must be distinctly enumerated for each system. For the radiolytic network described here, which specific physical processes are most responsible for the network’s organizational attributes?

One parsimonious explanation for the observed radiolytic network topology is based on the chemical kinetics that describe free radical production, propagation and termination. When a small molecule is split into two radicals, the radicals may recombine in a back-reaction into a molecule in an excited electronic state. The use of far from equilibrium incident particles and photons imparts a bias for the disproportionate production of lower mass radicals (electrons, hydrogen atoms, etc.) that can more easily attain higher translational speeds upon exiting the vicinity of a parent molecule. These species escape back-reaction more readily than larger radicals at a given temperature. Small mass radicals with higher average speeds also have larger collision frequencies with surrounding molecules. Collision frequency is intimately related to chemical reaction rates, so small radicals may be inclined to react more readily than larger molecules during radical propagation steps.

But why should a radiolytic system that contains a disproportionate concentration of small radicals also yield progressively larger molecules over time? One factor is that mon- and di-atomic species such as H, N, O, OH, CN, etc. possess few or no other internal degrees of freedom across which energy released during recombination may be easily distributed. This contrasts with larger molecules, where excess recombination energy can be efficiently redistributed across a higher number of internal degrees of freedom (53). This increases the probability that a recombination reaction that terminates a radical chain is more likely to involve production of a larger molecule than a smaller molecule.

Finally, the specific chemical attributes of the most frequently occurring radicals also contribute to network topology. Hydrogen radicals and solvated electrons have highly negative reduction potentials (-2.3V and -2.87V, respectively), and can be readily produced from both organic and inorganic substrates; they drive a wide array of net-reducing reactions (40). Hydrogen atom transfer reactions are among the most fundamental and common chemical reactions (54, 55). On the other hand, the hydroxyl radical is among the most reactive of oxygen-bearing radical species, with a highly positive reduction potential (+2.65V) and a capacity to react with a broad array of organic molecule types (39, 40). For biological context, the electrochemical potentials of water radicals are much more powerful than the redox potentials associated with metabolic intermediates (56).

Considering the electrochemical attributes of H and OH radicals, along with the chemical kinetics of small radicals described above, it is reasonable to uncover a heterogeneous network with these species as major hubs. It is unclear whether these chemical reactivity attributes meet the formal definition of preferential attachment within the network, but the underlying requirements for growth of scale-free networks (continuous addition of new nodes, and increased likelihood of reactivity of newly added nodes with existing, highly-linked nodes) (22, 57) seem consistent with properties of the radiolytic network.

## INTERPRETATION AND SIGNIFICANCE

The generative mechanisms for this network topology are firmly rooted in radiolysis, which implicates a fundamental interpretive question: are radicals necessary for producing complex abiotic chemical networks? There are no *a priori* reasons to conclude that they are. However, the ease with which interconnections between the various subnetworks can be facilitated by radical reactions (Figure 2, red shaded area) indicates that radiolysis can play a critical role in producing life-like network properties (hierarchical stratification, closed-loop cycles and heavy-tailed distribution) within chemical networks.

A heavy-tailed network distribution driven by radiolysis is, in some respects, not a surprising outcome and entirely consistent with findings from atmospheric models (41). Sole and Munteanu (4), with a dataset at the edge of statistical confidence for discerning a heavy-tailed distribution from one with an exponential fall-off, indicated that Earth’s atmospheric chemistry may fit a power law distribution with an incipient chemical hierarchy. Reactive oxygen species are recognized as limiting organic compound diversity, but previous efforts have not elucidated the underlying structure of the incipient chemical hierarchy or discussed the critical role that reactive nitrogen or chloride radicals can play in selectively producing prebiotic precursors (42-44).

The network’s foundation upon a large number of closed cycles is potentially one of the most significant general attributes of radiolytic systems. Families of radical cycles are known to play key roles in atmospheric composition (58, 59). In one of his last published papers, Orgel (60) wrote: “The demonstration of the existence of a complex, nonenzymatic metabolic cycle, such as the reverse citric acid, would be a major step in research on the origin of life, while demonstration of an evolving family of such cycles would transform the subject…The recognition of sequences of plausible reactions that could close a cycle is an essential first step toward the discovery of new cycles, but experimental proof that such cycles are stable against the challenge of side reactions is even more important.” The selected cycles depicted in Figure 3 illustrate that simple radiolytic cycles, once coupled together, lead to complex multi-reaction entities that do not trivially reduce to any of the network’s constituent reactions. Small-scale radiolytic experiments and large-scale irradiated systems such as upper atmospheres offer examples that systems of closed loops of radical reactions are stable against side reactions. It remains to be seen whether radiolysis-driven cycles meet the formal definition of autocatalytic sets as outlined by previous workers (61), but they do strongly implicate a potential for self-organization through cyclic chemical ‘entities’ prior to the emergence of catalytic macromolecules and full-enclosed cells.

Many of the radiolysis reactions are irreversible, resulting in the generation of new compounds that preclude a facile return of the system to a state that resembles the initial condition. As depicted in Figures 3 and 4, the system attenuates energy input as time progresses via the production and combinatorial interaction of new compounds. The reactive intermediates and energetic circumstances that promote ribonucleotide assembly and amino acid synthesis (i.e., condensing and leaving agents and periodic heating and cooling) could, in theory, simultaneously promote oligomerization of ribonucleotides and peptides or other genetic material candidates (62-66). This opens a tractable investigatory pathway to polymer-facilitated catalysis and information retention within this radiolytic chemical system, which distinguishes this effort from previous atmospheric model assessments. On this basis, the network can serve as a scaffold for developing and testing new hypotheses about prebiotic chronologies such as RNA World (67, 68) or for testing theories about the likelihood of *de novo* emergence of pre-metabolic, autocatalytic cycles (60).

Alternative interpretations of the observed heavy-tailed network topology must include the possibility of systemic bias. The data (chemical equations) that compose this network were gathered from an array of published papers. In some cases, experiments were designed in such a way to search for targeted species (42, 69) or with reference to modeling specific type localities (70, 71). In others, the analytical apparatus, product isolation techniques and compound resolution methods employed were limited to the technology available at the time of publication (72-74) but which have improved in the years that followed. It is therefore possible that the observed network topology represents a subset of reactions that is of particular interest to the scientific community.

On the other hand, there are also reasons to conclude that these sources of bias, though real, exert a minimal impact on the reported topological attributes. The first is that many of the experiments that serve as data sources for this network overlap with one another in terms of design parameters (i.e., total dose, dose rate, initial reactant selections and concentrations, radiolytic energy source, etc.), and that variations in experiment design nevertheless tend to result in similar arrays of products (42, 69). The second is that inferred intermediates are evaluated within an experimental context of relative abundances, not just presence or absence (42, 75). Large portions of the chemical network consist of data collected to build models of the large-scale reservoir of Titan’s atmosphere (70, 71), for which there are independently observable estimates of relative concentration and vertical mixing (76, 77). The combination of small-scale experimentation and large-scale observation indirectly serve to validate the core inter-relationships between the most likely chemical intermediates. A comprehensive array of multi-component experiments composed of the most highly-connected hub species can be used in future efforts to assess whether a realized example of this organosynethic system mimics the ensemble behavior predicted by the assemblage of individual reactions that compose the total network dataset. Preliminary results indicate that the combination of radiolytic products and periodic heating need not to be finely tuned to produce key nucleotide precursors (44).

It remains an open question as to whether life can *only* emerge from non-living settings that already possess life’s network-level organizational attributes prior to the appearance of the first cell, or whether these attributes arise as features at intermediate points throughout the process of life’s emergence. The radiolytic system described here provides one possible example that biological systems, at their deepest roots, may be templated upon a chemical system that already shared key organizational features associated with life. As such, this network may serve as a useful scaffold for testing the plausibility of various prebiotic chronologies. Considered separately from the question of life’s origins, this network may also serve as a basis for investigating whether radiolysis underpins a potential for organosynthetic self-organization.

ACKNOWLEDGMENTS

This work was supported by a grant from the Simons Foundation’s Simons Collaboration on the Origins of Life (494291, Zachary R. Adam). BK and SJ would like to acknowledge funding from the National Science Foundation (#1724090) and the John Templeton Foundation (#58562 and #61239) and ELSI Earth-Life Science Institute at the Tokyo Institute of Technology. We wish to thank H.J. Cleaves and A.H. Knoll for helpful comments on an earlier drafts of this manuscript.

REFERENCES

1. B. Shenhav, A. Solomon, D. Lancet, R. Kafri, "Early systems biology and prebiotic networks" in Transactions on computational systems biology I. (Springer, 2005), pp. 14-27.

2. P. Nghe *et al.*, Prebiotic network evolution: six key parameters. *Molecular BioSystems* **11**, 3206-3217 (2015).

3. C. Jolley, T. Douglas, Topological biosignatures: large-scale structure of chemical networks from biology and astrochemistry. *Astrobiology* **12**, 29-39 (2012).

4. R. V. Sole, A. Munteanu, The large-scale organization of chemical reaction networks in astrophysics. *EPL (Europhysics Letters)* **68**, 170 (2004).

5. S. I. Walker, C. Mathis, "Network Theory in Prebiotic Evolution" in Prebiotic Chemistry and Chemical Evolution of Nucleic Acids. (Springer, 2018), pp. 263-291.

6. W. Hordijk, J. Hein, M. Steel, Autocatalytic sets and the origin of life. *Entropy* **12**, 1733-1742 (2010).

7. M. A. Bedau *et al.*, Open problems in artificial life. *Artificial life* **6**, 363-376 (2000).

8. P. Grassberger, "Information and complexity measures in dynamical systems" in Information dynamics. (Springer, 1991), pp. 15-33.

9. K. Kaneko, Chaos as a source of complexity and diversity in evolution. *Artificial Life* **1**, 163-177 (1993).

10. G. Nicolis, Chemical chaos and self-organization. *Journal of Physics: Condensed Matter* **2**, SA47 (1990).

11. J. Pérez-Mercader, "Scaling phenomena and the emergence of complexity in astrobiology" in Astrobiology. (Springer, 2002), pp. 337-360.

12. W. Li, Expansion-modification systems: a model for spatial 1/f spectra. *Physical Review A* **43**, 5240 (1991).

13. J. H. Brown *et al.*, The fractal nature of nature: power laws, ecological complexity and biodiversity. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **357**, 619-626 (2002).

14. R. Liu, G. Mao, N. Zhang, Research of chemical elements and chemical bonds from the view of complex network. *Foundations of Chemistry* **21**, 193-206 (2019).

15. E. Estrada, The complex networks of earth minerals and chemical elements. *Match* **59**, 605-624 (2008).

16. R. Albert, Scale-free networks in cell biology. *Journal of cell science* **118**, 4947-4957 (2005).

17. J. W. Valentine, C. L. May, Hierarchies in biology and paleontology. *Paleobiology* **22**, 23-33 (1996).

18. D. W. McShea, The hierarchical structure of organisms: a scale and documentation of a trend in the maximum. *Paleobiology* **27**, 405-423 (2001).

19. H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, A.-L. Barabási, The large-scale organization of metabolic networks. *Nature* **407**, 651 (2000).

20. P. A. Marquet *et al.*, Scaling and power-laws in ecological systems. *Journal of Experimental Biology* **208**, 1749-1769 (2005).

21. A. Muscente *et al.*, Quantifying ecological impacts of mass extinctions with network analysis of fossil communities. *Proceedings of the National Academy of Sciences* **115**, 5217-5222 (2018).

22. A.-L. Barabási, R. Albert, Emergence of scaling in random networks. *science* **286**, 509-512 (1999).

23. D. Marković, C. Gros, Power laws and self-organized criticality in theory and nature. *Physics Reports* **536**, 41-74 (2014).

24. R. Adler, R. Feldman, M. Taqqu, *A practical guide to heavy tails: statistical techniques and applications* (Springer Science & Business Media, 1998).

25. A. Kleidon, R. D. Lorenz, *Non-equilibrium thermodynamics and the production of entropy: life, earth, and beyond* (Springer Science & Business Media, 2004).

26. N. Goldenfeld, C. Woese, Life is physics: evolution as a collective phenomenon far from equilibrium. *Annu. Rev. Condens. Matter Phys.* **2**, 375-399 (2011).

27. R. Braakman, E. Smith, The compositional and evolutionary logic of metabolism. *Physical biology* **10**, 011001 (2012).

28. E. Smith, H. J. Morowitz, *The origin and nature of life on earth: the emergence of the fourth geosphere* (Cambridge University Press, 2016).

29. M. Pagel, A. Meade, D. Scott, Assembly rules for protein networks derived from phylogenetic-statistical analysis of whole genomes. *BMC evolutionary biology* **7**, S16 (2007).

30. A.-L. Barabási, R. Albert, H. Jeong, Mean-field theory for scale-free random networks. *Physica A: Statistical Mechanics and its Applications* **272**, 173-187 (1999).

31. E. Limpert, W. A. Stahel, M. Abbt, Log-normal distributions across the sciences: keys and clues: on the charms of statistics, and how mechanical models resembling gambling machines offer a link to a handy way to characterize log-normal distributions, which can provide deeper insight into variability and probability—normal or log-normal: that is the question. *BioScience* **51**, 341-352 (2001).

32. M. Mitzenmacher, A brief history of generative models for power law and lognormal distributions. *Internet mathematics* **1**, 226-251 (2004).

33. G. Lima-Mendez, J. van Helden, The powerful law of the power law and other myths in network biology. *Molecular BioSystems* **5**, 1482-1493 (2009).

34. A. D. Broido, A. Clauset, Scale-free networks are rare. *Nature communications* **10**, 1017 (2019).

35. M. P. Stumpf, M. A. Porter, Critical truths about power laws. *Science* **335**, 665-666 (2012).

36. S. H. Strogatz, Exploring complex networks. *nature* **410**, 268 (2001).

37. S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, Complex networks: Structure and dynamics. *Physics reports* **424**, 175-308 (2006).

38. E. J. Hart, J. Thomas, S. Gordon, A review of the radiation chemistry of single-carbon compounds and some reactions of the hydrated electron in aqueous solution. *Radiation Research Supplement* **4**, 74-88 (1964).

39. M. Gutowski, S. Kowalczyk, A study of free radical chemistry: their role and pathophysiological significance. *Acta Biochimica Polonica* **60** (2013).

40. H. A. Schwarz (1981) Free radicals generated by radiolysis of aqueous solutions. (ACS Publications).

41. P. M. Gleiss, P. F. Stadler, A. Wagner, D. A. Fell, Relevant cycles in chemical reaction networks. *Advances in complex systems* **4**, 207-226 (2001).

42. Z. R. Adam *et al.*, Estimating the capacity for production of formamide by radioactive minerals on the prebiotic Earth. *Scientific reports* **8**, 265 (2018).

43. R. Yi, Y. Hongo, I. Yoda, Z. R. Adam, A. C. Fahrenbach, Radiolytic synthesis of cyanogen chloride, cyanamide and simple sugar precursors. *ChemistrySelect* **3**, 10169-10174 (2018).

44. Y. Ruiqin *et al.*, Continuous Chemical Evolution of RNA Precursors. *Journal of the American Chemical Society* (in prep.).

45. P. Bak, C. Tang, K. Wiesenfeld, Self-organized criticality: An explanation of the 1/f noise. *Physical review letters* **59**, 381 (1987).

46. B. Gaveau, M. Moreau, J. Toth, Scenarios for self-organized criticality in dynamical systems. *Open Systems & Information Dynamics* **7**, 297-308 (2000).

47. C. G. Langton, Computation at the edge of chaos: phase transitions and emergent computation. *Physica D: Nonlinear Phenomena* **42**, 12-37 (1990).

48. M. Bastian, S. Heymann, M. Jacomy (2009) Gephi: an open source software for exploring and manipulating networks. in *Third international AAAI conference on weblogs and social media*.

49. A. Clauset, C. R. Shalizi, M. E. Newman, Power-law distributions in empirical data. *SIAM review* **51**, 661-703 (2009).

50. J. Alstott, E. Bullmore, D. Plenz, powerlaw: a Python package for analysis of heavy-tailed distributions. *PloS one* **9**, e85777 (2014).

51. B. Titz *et al.*, The binary protein interactome of Treponema pallidum–the syphilis spirochete. *PloS one* **3**, e2292 (2008).

52. M. A. Munoz, Colloquium: Criticality and dynamical scaling in living systems. *Reviews of Modern Physics* **90**, 031001 (2018).

53. I. Glassman, R. Yetter, N. Glumac (2015) Chapter 2–Chemical Kinetics. Combustion. (Amsterdam: Elsevier).

54. D. Bím, M. Maldonado-Domínguez, L. Rulíšek, M. Srnec, Beyond the classical thermodynamic contributions to hydrogen atom abstraction reactivity. *Proceedings of the National Academy of Sciences* **115**, E10287-E10294 (2018).

55. J. M. Mayer, Hydrogen atom abstraction by metal− oxo complexes: Understanding the analogy with organic radical reactions. *Accounts of chemical research* **31**, 441-450 (1998).

56. G. Karp, *Cell and molecular biology: concepts and experiments* (John Wiley & Sons, 2009).

57. V. Dančík, A. Basu, P. Clemons, "Properties of Biological Networks" in Systems Biology. (Springer, 2013), pp. 129-178.

58. P. S. Monks, Gas-phase radical chemistry in the troposphere. *Chemical Society Reviews* **34**, 376-395 (2005).

59. U. Platt *et al.*, "Free radicals and fast photochemistry during BERLIOZ" in Tropospheric Chemistry. (Springer, 2002), pp. 359-394.

60. L. E. Orgel, The implausibility of metabolic cycles on the prebiotic Earth. *PLoS biology* **6**, e18 (2008).

61. V. Vasas, C. Fernando, M. Santos, S. Kauffman, E. Szathmáry, Evolution before genes. *Biology Direct* **7**, 1 (2012).

62. A. E. Engelhart, N. V. Hud, Primitive genetic polymers. *Cold Spring Harbor perspectives in biology* **2**, a002196 (2010).

63. P. Karri, V. Punna, K. Kim, R. Krishnamurthy, Base‐Pairing Properties of a Structural Isomer of Glycerol Nucleic Acid. *Angewandte Chemie International Edition* **52**, 5840-5844 (2013).

64. V. B. Pinheiro, P. Holliger, The XNA world: progress towards replication and evolution of synthetic genetic polymers. *Current opinion in chemical biology* **16**, 245-252 (2012).

65. A. Eschenmoser, The TNA-family of nucleic acid systems: properties and prospects. *Origins of Life and Evolution of the Biosphere* **34**, 277-306 (2004).

66. J. G. Forsythe *et al.*, Ester‐mediated amide bond formation driven by wet–dry cycles: A possible path to polypeptides on the prebiotic Earth. *Angewandte Chemie International Edition* **54**, 9871-9875 (2015).

67. M. P. Robertson, G. F. Joyce, The origins of the RNA world. *Cold Spring Harbor perspectives in biology* **4**, a003608 (2012).

68. N. V. Hud, B. J. Cafferty, R. Krishnamurthy, L. D. Williams, The origin of RNA and “my grandfather’s axe”. *Chemistry & biology* **20**, 466-474 (2013).

69. D. Dondi, D. Merli, L. Pretali, A. Buttafava, A. Faucitano, Detailed analytical study of radiolysis products of simple organic compounds as a methodological approach to investigate prebiotic chemistry—Part 1. *Radiation Physics and Chemistry* **80**, 403-407 (2011).

70. L. M. Lara, E. Lellouch, J. López‐Moreno, R. Rodrigo, Vertical distribution of Titan's atmospheric neutral constituents. *Journal of Geophysical Research: Planets* **101**, 23261-23283 (1996).

71. K. J. Zahnle, Photochemistry of methane and the formation of hydrocyanic acid (HCN) in the Earth's early atmosphere. *Journal of Geophysical Research: Atmospheres* **91**, 2819-2834 (1986).

72. I. Draganic, M. Nenadovic, Z. D. Draganic, Radiolysis of HCOOH+ O2 at pH 1.3-13 and the yields of primary products in. gamma. radiolysis of water. *The Journal of Physical Chemistry* **73**, 2564-2571 (1969).

73. L. Josimović, I. Draganić, The radiolysis of acetic acid in aqueous solutions and acetic acid-water mixtures. *International Journal for Radiation Physics and Chemistry* **5**, 505-512 (1973).

74. A. Negrón-Mendoza, Z. Draganić, R. Navarro-González, I. Draganić, Aldehydes, ketones, and carboxylic acids formed radiolytically in aqueous solutions of cyanides and simple nitriles. *Radiation research* **95**, 248-261 (1983).

75. Z. D. Draganić *et al.*, Radiolysis of aqueous solutions of ammonium bicarbonate over a large dose range. *International Journal of Radiation Applications and Instrumentation. Part C. Radiation Physics and Chemistry* **38**, 317-321 (1991).

76. J. Loison *et al.*, The neutral photochemistry of nitriles, amines and imines in the atmosphere of Titan. *Icarus* **247**, 218-247 (2015).

77. M. L. Cable *et al.*, Titan tholins: simulating Titan organic chemistry in the Cassini-Huygens era. *Chem. Rev* **112**, 1882-1909 (2012).