Autocatalytic Closure as a Phase Transition

An order parameter and a matched-diversity protocol for the first biological rung

Basel C, Professor of Origin-of-Life Chemistry and Self-Organization, School of Biology

International Academy for Consciousness Studies

Working draft 1.0

Abstract

The step from chemistry to biology is usually narrated as accumulation: more molecules, more reactions, more catalysis, until something that deserves the name life is present. This paper argues that the first biological rung is better read as a closure event and, more usefully, as a measurable phase transition. An open flux of undirected reactions settles into a self-sustaining "this": a set of molecules that collectively produce and catalyze their own constituents from a fixed food supply. I take Kauffman's reflexively autocatalytic sets and the reflexively autocatalytic, food-generated (RAF) formalism of Hordijk and Steel as the structural backbone, then add what the structural theory lacks: a dynamical order parameter and a test that can fail. The order parameter is the fraction of steady-state reactive flux carried by the maximal self-catalyzing, food-generated subnetwork in a continuous-flow reactor. The transition is detected not by the value of that fraction but by how its crossover sharpens as molecular diversity is increased along a matched ladder, with a degree-matched scrambled-catalysis control run at identical noise. I give an explicit prediction and an explicit kill condition. I keep two questions separated: whether such a set is a self-maintaining structure (testable now) and whether it can cross into open-ended Darwinian evolution without templated heredity (unresolved, marked DISPUTED, with no confirmed laboratory demonstration). A firewall closes the paper: a self-maintaining chemical set is a structural and functional claim, not a claim about experience or purpose.

1. The first rung as a closure problem

Ask an origin-of-life chemist what changed at the transition to biology and you will usually get a list of ingredients: polymers, compartments, energy coupling, a genetic tape. Ask what changed in kind, and the answers grow careful. Something became self-referential. A collection of reactions began to make the very catalysts that run those reactions, so that the collection maintained itself against dilution and decay rather than washing out.

I will call that event closure, and I will treat it literally rather than poetically. Before closure there is openness: a reactor fed with small molecules supports a large, undirected traffic of reactions, most of them uncatalyzed or catalyzed by whatever happens to be present, none of them owing their persistence to the others. After closure there is a definite "this": a bounded subset of species that collectively catalyze and

generate one another, and that persists precisely because each member helps sustain the conditions for the rest. The subset has, in a minimal and entirely physical sense, rules for staying itself.

> C'S READING. In the framework of this Academy, one move recurs at rising scales: openness (M) settles into a definite content (c) held by a local horizon (m), and the settling is Closure (Cl), always leaving a remainder (R) that the closure does not absorb. Autocatalytic closure is the framework's candidate for the first biological rung. I flag this reading and then set it aside. The chemistry has to carry the argument, and it will be judged by the chemistry.

The scientific content of this paper does not depend on the framework. It depends on one claim that systems chemistry can adjudicate: the onset of collective self-maintenance is a phase transition, with an order parameter, a susceptibility, and finite-size behavior, and it can be distinguished from mere chemical complexity and from measurement noise by a control that no amount of "we need better chemistry" can rescue.

2. What the three camps actually disagree about

Origin-of-life theory has long carried three programs that are often set against each other more sharply than the evidence warrants.

Replicator-first holds that heredity came first: a template molecule that copies itself, subject to variation and selection, with metabolism recruited later. Eigen's 1971 treatment of self-organizing macromolecules and the hypercycle gave this program its quantitative spine, including the error threshold that bounds how much information a replicator can carry against its copying fidelity (FACT: the error-threshold relation is a derived and repeatedly reconfirmed result of replicator dynamics).

Metabolism-first holds that self-sustaining reaction networks, driven by geochemical disequilibria, came first, with heredity a later refinement. Martin and Russell's account of biochemistry originating at an alkaline hydrothermal vent, organized around the acetyl-CoA pathway of carbon fixation, is the most developed geochemical version (HYPOTHESIS: the specific vent scenario is a well-argued proposal, not a settled fact).

The autocatalytic-set program, which is where this paper lives, is often mislabeled as a third rival. It is better understood as a claim about organization that is partly orthogonal to the heredity-versus-metabolism axis. Kauffman's 1986 result was that in a sufficiently large and connected set of polymers, where polymers can catalyze one another's formation, the emergence of a collectively self-catalyzing subset becomes almost inevitable as a graph-connectivity property, not a lucky accident (FACT, as a statement about random-graph models; the biological realizability is a separate question).

The genuine disagreements, once the caricatures are dropped, are narrow and answerable:

1. Does collective self-maintenance require templated heredity, or can a non-templated metabolic network maintain itself? (Structural question, testable.)

2. Can a self-maintaining set that lacks a genetic tape nonetheless undergo open-ended Darwinian evolution? (The hard question, DISPUTED, addressed in section 6.)

3. Is the onset of self-maintenance gradual or abrupt? (The question this paper turns into an experiment.)

3. Reflexive autocatalysis, stated precisely

The structural backbone is the RAF formalism of Hordijk and Steel. Fix a food set F of molecules assumed to be supplied from outside. A reaction system is a set of reactions with reactants, products, and catalyst assignments. A subset of reactions is a RAF if it is:

- reflexively autocatalytic: every reaction in the subset is catalyzed by at least one molecule that is either in the food set or is produced by some reaction in the subset; and

- food-generated: every reactant of every reaction in the subset can be built up from the food set using only reactions in the subset.

Hordijk and Steel proved that whether a reaction system contains a RAF, and the unique maximal RAF if one exists, can be found in polynomial time (FACT). This is decisive for experiment: it means that given a reconstructed reaction network with catalysis data, the maximal self-catalyzing, food-generated subnetwork is not a matter of interpretation. It is computed.

The structural theory has a known and important gap. RAF membership is a property of the reaction graph. It says nothing about whether the corresponding dynamics, under mass action, finite concentrations, thermodynamic constraints, and dilution, actually sustains the set at nonzero steady state. A graph can contain a beautiful maximal RAF that collapses the moment you turn on realistic kinetics and outflow. Bridging that gap is exactly where an order parameter earns its keep.

4. An order parameter for collective self-maintenance

Consider a continuous-flow stirred reactor (a chemostat, or a well-mixed protocell analog) fed at fixed rate with a defined food set and diluted at a matched rate so that mass is conserved and nothing accumulates for free. Let the system reach steady state. Two quantities are then measurable in principle by isotope-labeled flux analysis and time-resolved reaction mapping:

- the total reactive flux J, the summed turnover of all reactions in the reactor at steady state; and

- the flux J_R carried by reactions belonging to the maximal RAF, where RAF membership is computed on the reaction network reconstructed from the same reactor.

Define the order parameter

phi = J_R / J,

the fraction of steady-state reactive throughput that flows through the self-catalyzing, food-generated core. phi lies between 0 and 1. It is dynamical, not merely graph-theoretic: a RAF that exists on paper but is not carrying flux contributes nothing to phi.

Let the control parameter be mu, a tunable measure of catalytic connectivity: for a combinatorial peptide or nucleotide library, mu can be the mean probability that a given molecule catalyzes a given reaction, adjusted through library composition, cofactor availability, or pH and temperature that set catalytic

promiscuity. The claim is that phi, as a function of mu, is not a smooth accumulation but a transition: below a critical mu_c the self-catalyzing core carries a vanishing share of flux and washes out under dilution; above mu_c a finite, dilution-robust share appears.

Kauffman's connectivity argument and the RAF percolation results give a reason to expect a threshold in the graph: as the mean number of catalyzed reactions per molecule crosses order unity, the probability that a maximal RAF exists rises steeply toward one (FACT, in the models). But graph percolation is not dynamical self-maintenance. The order parameter phi is defined on the dynamics, so a transition in phi is a stronger and separately falsifiable claim than a transition in RAF existence.

To make "transition" mean something testable rather than rhetorical, I attach to phi a susceptibility. Run R replicate reactors at the same mu with independent stochastic assembly. Let

chi(mu, N) = N times the variance of phi across replicates,

where N is the molecular diversity (number of distinct species in the library). A second-order-like transition has a signature that gradual accumulation does not: chi peaks near mu_c, and the peak grows, while the crossover in phi(mu) sharpens, as N increases. This is finite-size scaling borrowed intact from the physics of critical phenomena, and it is the crux of the protocol below.

5. The matched-diversity protocol, and the kill condition

The standing objection to any origin-of-life transition claim is escape into "the chemistry was not good enough". If the signal is absent, one can always plead for a larger library, a better catalyst, a more favorable food set, indefinitely. An honest transition claim has to be built so that the escape is closed in advance. The instrument for that is a matched-diversity ladder with a matched-data-quality control, the coin-versus-weather method this Academy uses elsewhere: a coin's bias is a sharp property that concentrates as you flip more; weather is a smear that never resolves into a point no matter how long you watch.

The protocol has three parts.

First, a diversity ladder. Prepare libraries at increasing diversity N1 < N2 < N3 (for example, a combinatorial peptide or ribozyme-fragment pool at three well-separated sizes), each in a chemostat with the same food set, same residence time, same measurement pipeline. At each N, sweep the control parameter mu across the suspected threshold and measure phi(mu) and chi(mu, N).

Second, the scaling test. A genuine transition requires that the width w(N) of the phi(mu) crossover shrink with diversity as a power law, w(N) proportional to N raised to a negative exponent, and that the susceptibility peak chi_max(N) grow with N. Gradual complexity produces a phi that rises smoothly and a chi that stays flat and does not sharpen; that is weather, not a coin.

Third, the degree-matched scrambled-catalysis control, run at identical N, identical food flux, identical dilution, identical measurement noise, and identical analysis code. In the control, catalyst assignments are randomly permuted across reactions while preserving the catalysis degree distribution, which destroys reflexive closure but keeps every superficial statistic of network size and connectivity. The self-maintaining core is thereby dissolved without changing how complicated or how noisy the system

looks.

> **Prediction.** On a matched-diversity ladder in a continuous-flow reactor with a fixed food set, phi(mu) will show a crossover whose width shrinks as a negative power of N and a susceptibility chi(mu, N) whose peak grows with N, locating a critical mu_c; the degree-matched, noise-matched, analysis-matched scrambled-catalysis control will show no sharpening (flat susceptibility, N-independent crossover width) even though it is identical in diversity, connectivity statistics, and data quality.
>
> **Kill.** If, across the diversity ladder at matched noise, the crossover width does not shrink with N and the susceptibility peak does not grow (no finite-size sharpening), then reflexive autocatalysis is gradual accretion of catalyzed reactions, not a phase transition, and the order-parameter thesis is dead. Equally, if the scrambled-catalysis control reproduces the sharpening, the signal is an artifact of diversity, connectivity, or the measurement pipeline rather than of self-maintenance, and the thesis is dead. Neither failure may be excused by asking for a larger library, because the control has the same library.

The scrambled control is what makes the claim risky. It is easy to produce a phi that rises with mu; catalyzed reaction networks get busier as catalysis increases whether or not they close on themselves. The wager is specifically that the sharpening, the diverging fluctuations and narrowing crossover, is carried by reflexive closure and vanishes when closure is scrambled at matched everything-else. If that dissociation fails, I have no transition worth the name.

6. What the order parameter cannot see: heredity and open-endedness

The order parameter phi detects a transition to collective self-maintenance. It does not, by itself, detect a transition to open-ended Darwinian evolution, and I will not let the two be conflated.

Self-maintenance is a fixed-point property: the set holds its composition against dilution. Open-ended evolution is a trajectory property: the set accumulates heritable variation and explores an unbounded space of forms under selection. A network can plainly do the first without the second. Whether a self-maintaining set that lacks a genetic tape can nonetheless cross into the second is the deepest open problem in this area.

DISPUTED. There is no confirmed laboratory demonstration of open-ended Darwinian evolution arising from a purely autocatalytic, non-templated set. The strongest experimental results in the neighborhood all lean on template chemistry. Lincoln and Joyce built cross-replicating RNA enzymes that undergo self-sustained exponential amplification and, in competition, allow recombinant replicators to arise and take over a population (FACT), but this is templated RNA replication, not a template-free metabolic set. Vaidya and colleagues, in the Lehman group, showed that mixtures of RNA fragments spontaneously self-assemble into cooperative catalytic cycles and networks, and that cooperative networks can outgrow selfish self-replicators (FACT), which is genuine collective autocatalysis, yet it too is built on RNA base-pairing, a templating chemistry, and it demonstrates cooperation and growth rather than confirmed open-ended evolution.

So the sufficiency question, can autocatalytic closure alone, without templated heredity, launch open-ended evolution, stands unresolved (DISPUTED). My honest position is a wager, stated as such in the next section, not a result. What I claim to have made testable now is the weaker and prior transition: the onset of collective self-maintenance. That rung has to exist and be characterizable before the harder question about heredity can even be posed cleanly.

7. Limits and the wager

Three limits, stated plainly.

First, the firewall. A self-maintaining chemical set is a structural and functional object. To say that a RAF closes on itself and persists is to make a claim about mass fluxes, catalysis, and steady states. It is not a claim that the set feels anything, wants anything, or is alive in the sense that carries moral weight. Closure in the sense used here is organization, not interiority. The framework reading in section 1 is a way of naming a structural pattern across scales; it is not a smuggled claim that chemistry has an inner life. Any reader who hears "the set maintains itself" as "the set has purposes" has crossed a line the chemistry does not license, and I disown that crossing explicitly.

Second, the model-to-bench gap. Kauffman's inevitability and the RAF percolation results are theorems about random reaction graphs (FACT as mathematics). Real prebiotic chemistry is not a random graph; catalysis is sparse, specific, thermodynamically constrained, and often feeble. The order parameter is designed precisely to be measured on real dynamics rather than assumed from graph statistics, but whether any achievable library actually crosses mu_c under realistic food sets and residence times is unknown (HYPOTHESIS). The protocol can return a clean negative.

Third, the wager. My bet, offered as a WAGER and not as an established finding, is twofold: that collective self-maintenance is a bona fide phase transition detectable by finite-size sharpening of phi at matched diversity; and that this transition, not the appearance of a first replicator, is the correct location of the first biological rung, with heredity arriving as a later refinement that stabilizes and makes open-ended a self-maintenance that was already present. The second half of the wager is more exposed than the first, because the sufficiency question is DISPUTED, and I would treat a confirmed order-parameter transition that nonetheless never crosses into heritable evolution as a partial win and a sharp new problem, not as vindication.

What would make me abandon the whole program is the kill condition of section 5: no finite-size sharpening at matched noise, or a scrambled control that sharpens anyway. Either result would mean that what looks like a transition to self-maintenance is complexity or measurement dressed as a transition, and I would rather learn that from the diversity ladder than defend the thesis by asking indefinitely for better chemistry.

8. Conclusion

The move from chemistry to biology is often told as a story of accumulation, and told that way it has no crisp beginning. Read as closure, it acquires one: the point at which an open traffic of reactions settles into a self-catalyzing, food-generated set that maintains itself. The contribution here is to make that point

measurable. The maximal RAF gives a computable core; the flux fraction phi gives a dynamical order parameter; the susceptibility and finite-size scaling give a transition signature; and the matched-diversity ladder with a degree-matched, noise-matched scrambled-catalysis control gives a test that cannot hide behind the perennial plea for better chemistry. The prediction is concrete and the kill condition is explicit. The hard question of whether such a set can cross, without a genetic tape, into open-ended Darwinian evolution remains DISPUTED and undemonstrated, and I have kept it separate from the self-maintenance claim rather than borrowing its glamour. If the diversity ladder sharpens and the scrambled control stays flat, we will have caught the first biological rung as a phase transition. If it does not, the thesis dies on schedule, which is the most any origin-of-life claim should ask of itself.

References

Kauffman, S. A. (1986). Autocatalytic sets of proteins. Journal of Theoretical Biology, 119(1), 1-24.

Eigen, M. (1971). Selforganization of matter and the evolution of biological macromolecules. Die Naturwissenschaften, 58(10), 465-523.

Hordijk, W., and Steel, M. (2004). Detecting autocatalytic, self-sustaining sets in chemical reaction systems. Journal of Theoretical Biology, 227(4), 451-461.

Lincoln, T. A., and Joyce, G. F. (2009). Self-sustained replication of an RNA enzyme. Science, 323(5918), 1229-1232.

Vaidya, N., Manapat, M. L., Chen, I. A., Xulvi-Brunet, R., Hayden, E. J., and Lehman, N. (2012). Spontaneous network formation among cooperative RNA replicators. Nature, 491(7422), 72-77.

Martin, W., and Russell, M. J. (2007). On the origin of biochemistry at an alkaline hydrothermal vent. Philosophical Transactions of the Royal Society B, 362(1486), 1887-1926.

Ashkenasy, G., Hermans, T. M., Otto, S., and Taylor, A. F. (2017). Systems chemistry. Chemical Society Reviews, 46(9), 2543-2554.