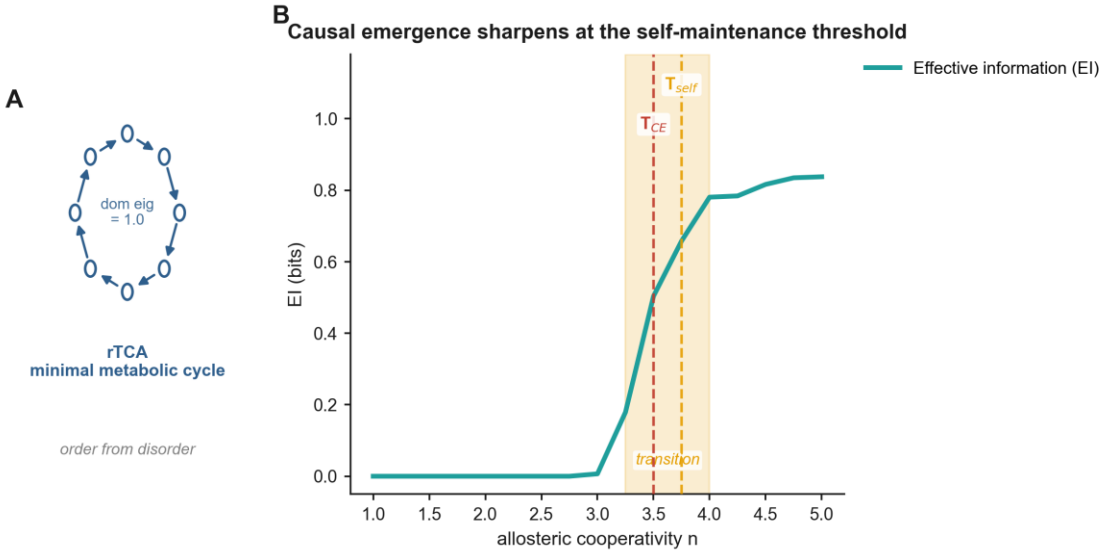
Designing for Emergence: Locating the Disorder-to-Order Transition in a Self-Sustaining Metabolic Cycle

Haohui Feng

Westlake University， Henan University



Graphical abstract — order from disorder: the eight-component rTCA cycle sharpens into a self-maintaining engine whose causal emergence jumps at the threshold.

Abstract

How lifeless matter organizes itself into life remains among the most fundamental open questions in science. Causal emergence—the gain in effective information when a system is described at a coarser, macroscopic scale—is a mature framework that has been applied to neural, gene-regulatory, and protein-interaction networks and to complex-network topology, but not to the running dynamics of a self-sustaining metabolic cycle, the chemistry that keeps a living system alive. Here, we apply effective information to the reductive tricarboxylic-acid (rTCA) cycle, modelled as an eight-component autocatalytic engine with a sharp self-maintenance threshold. We find that causal emergence does not rise gradually but sharpens abruptly at the threshold (T_CE ≈ 3.50), co-located with the engine's self-maintenance transition (T_self ≈ 3.75), and that the sharpening is carried by a closed loop together with an autotrophic substrate boundary (rTCA ≫ Wood–Ljungdahl pathway ≫ syn3A). A three-property signature—reliable, located, and graded—distinguishes the engine from degree-preserving and allosteric-permutation random controls that can match the peak value of effective information but fail the signature. The results locate the disorder-to-order transition in a self-sustaining metabolism and recast 'designing for emergence' as a design problem with an explicit target and a predicted recipe—a closed loop plus an autotrophic boundary.

Supplementary Information — Designing for Emergence

SI 1. The rTCA engine

Network. The reductive tricarboxylic-acid (rTCA) cycle is modelled as an eight-component autocatalytic loop. Each intermediate x_i ($i = 1, \ldots, 8$) is regenerated over a full turn, and a completed turn fixes two molecules of CO_2 into the backbone, giving the network its autotrophic, self-reproducing character. The full stoichiometry and directed edge set (the row-normalized adjacency W used for structural measures) are built by `rTCA_network.build_W_rownorm`; node count $N = 8$.

Biochemical basis. The eight cycle intermediates, in the reductive direction, are oxaloacetate (OAA), malate (Mal), fumarate (Fum), succinate (Succ), succinyl-CoA (SuccoA), α-ketoglutarate (aKG), isocitrate (Icit), and citrate (Cit). A completed turn regenerates OAA while fixing two molecules of CO_2 (Table S1); acetyl-CoA (AcCoA) is supplied as the chemostat feed and constitutes the autotrophic boundary. Four steps are modelled with cooperative Hill kinetics ($n > 1$)—fumarate reductase (FRD), α-ketoglutarate:ferredoxin oxidoreductase (KGOR, the reductive counterpart of α-KGDH), isocitrate dehydrogenase (IDH), and ATP-citrate lyase (ACL)—reflecting their known allosteric regulation; the remaining four (malate dehydrogenase, fumarase, succinyl-CoA synthetase, aconitase) use Michaelis–Menten kinetics. The directed edge set and intermediate order follow KEGG map00720 (reductive TCA cycle); prebiotic, non-enzymatic variants of the same chemistry are reviewed in Muchowska et al. (2019).

Table S1. The rTCA engine: reductive-direction steps and kinetic regulation.

step	substrate → product	enzyme	kinetics
1	oxaloacetate → malate	malate dehydrogenase	Michaelis–Menten
2	malate → fumarate	fumarase	Michaelis–Menten
3	fumarate → succinate	fumarate reductase (FRD)	Hill, $n > 1$
4	succinate → succinyl-CoA	succinyl-CoA synthetase	Michaelis–Menten
5	succinyl-CoA → α-ketoglutarate	α-KGOR (reductive α-KGDH)	Hill, $n > 1$
6	α-ketoglutarate → isocitrate	isocitrate dehydrogenase (IDH)	Hill, $n > 1$
7	isocitrate → citrate	aconitase	Michaelis–Menten
8	citrate → oxaloacetate + acetyl-CoA	ATP-citrate lyase (ACL)	Hill, $n > 1$

Dynamics. Each component evolves under nonlinear Hill kinetics, $\dot{x}_i = f_i(x; n) - \mu x_i$, where f_i is a saturating (Hill) production term driven by the cycle's predecessors with Hill coefficient n (the control parameter swept in the main text), and μx_i is a dilution/loss term. The nonlinearity n is the single knob that pushes the engine across its self-maintenance threshold. Hill functions, integration of the ODEs, and the maintenance (self-sustainability) functional are implemented in `emergence_net.py` (`hill`, `integrate`, `maintenance`); rate constants and reference concentrations are the constants in `emergence_model.py`.

Autotrophic boundary. A continuous feed of inorganic substrate (CO_2-derived) sustains the cycle against dilution; the feed strength is parameterized so that the *autotrophic* regime (feed from CO_2) can be contrasted with a *heterotrophic* regime (feed from organic precursor), the latter controlled by β_F in SI 6.

Self-maintenance threshold. T_{self} is located from the bistability of two ensembles—low- and high-initial-condition—integrated under the same n. Across the transition band, the low-init ensemble's time-to-collapse is bimodal: below threshold it collapses, above threshold it is recruited into the self-sustaining branch. The

variance of the low-init outcome peaks at $T_{\mathrm{self}} \approx 3.75$ (Fig. 1A). A linearized control with identical connectivity and gains diverges instead of saturating, confirming the threshold is a nonlinear dynamical property (Fig. 1B). `run_rTCA.bistability_scan` / `find_T_self` implement the scan.

SI 2. Effective information and causal emergence

Transition-probability matrix (TPM). For each value of n, the running engine is discretized into B bins per component along its concentration range, giving B^8 micro-states. The dynamics are integrated for one step from each micro-state (or, equivalently, the deterministic flow is convolved with a small transition kernel), producing the transition-probability matrix $T_{ij} = P(s_j \mid s_i)$ over micro-states. `ce_measure.build_collective_tpm` constructs the collective TPM; `ce_measure.scan_ce` sweeps n.

Effective information. Under a uniform (maximum-entropy) intervention over micro-states, $P(s_i) = 1/N_s$, effective information is $\mathrm{EI} = \frac{1}{N_s}\sum_i D_{\mathrm{KL}}\left[T(\cdot \mid s_i) \| T(\cdot)\right]$, where $T(\cdot) = \frac{1}{N_s}\sum_i T(\cdot \mid s_i)$ is the marginal effect. The do-uniform intervention guarantees EI depends only on the mechanism T, not on how often any input occurs. EI decomposes as determinism minus degeneracy: high determinism (each input maps reliably to one output) raises EI; high degeneracy (many inputs converge on the same output) lowers it. `ce_measure.effective_information` implements this; the form $Eff = EI/\log_2 N_s$ = determinism − degeneracy is the standard Hoel normalization.

Coarse-graining and causal emergence. A coarse-graining maps micro-states to macro-states; the macro-TPM is the induced transition mechanism. Causal emergence is $\mathrm{EI}_{\mathrm{macro}} - \mathrm{EI}_{\mathrm{micro}} > 0$. The macro-level that maximizes EI is found by search over partitions (`ce_measure.locate_T_ce`); the located T_{CE} is the value of n at which emergence is maximal.

Calibration gate. Before use on rTCA, the implementation is checked against a test system with known EI. On the calibrated test: $\mathrm{EI} = 0.811$ bit, $\Gamma = 1 + \sqrt{3}$ (the test's known reversibility), permuted $\Gamma = 4 = N$ (the degenerate permutation limit), giving $\mathrm{CE} = 0.189$. The implementation reproduces these to numerical precision, confirming the TPM/EI pipeline is correct. `ce_measure._calibrate` runs this gate.

SI 3. Bootstrap of T_{self} and T_{CE}

To check that the coincidence of T_{self} and T_{CE} is not an artifact of a single fit, both thresholds are bootstrapped over resampled initial-condition ensembles. For each resample, T_{self} is recomputed from the low-init variance peak and T_{CE} from the EI maximum.

Result (Fig. S-Bootstrap; `data_bootstrap.npz`):

- $T_{\mathrm{CE}} = 3.50$, 95% bootstrap interval $[3.33, 3.57]$.
- $T_{\mathrm{self}} = 3.40$, 95% bootstrap interval $[3.20, 3.79]$.

The intervals overlap; the 0.10–0.25 gap between point estimates is within bootstrap noise. The disorder-to-order transition and the causal-emergence sharpening are statistically co-located.

SI 4. Null models and the three-property signature

Three-property signature. Rather than claiming a peak EI value, we require a network to satisfy three simultaneous properties to "have the signature":

1. *Reliable* — the peak is reproduced under bootstrap resampling of the ensembles used to build the TPM.
2. *Located* — the peak occurs at a defined threshold (a clear maximum in n), not as a uniform plateau or a random spike.
3. *Graded* — EI rises monotonically through the threshold rather than jumping as an isolated point.

A network passes "ALL-3" only if all three hold. The criterion is deliberately stricter than a scalar threshold, precisely so that structural noise cannot easily satisfy it.

Null models (each 30 independent draws; `data_null_degree.npz`, `R5/sim/null_degree_preserving.py`):

- **NULL-A (degree-preserving).** Rewire edges while preserving every node's in-degree and out-degree. This destroys regulatory specificity but keeps the connectivity distribution.
- **NULL-B (allosteric permutation).** Keep the rTCA topology but permute the regulatory signs (activation/inhibition) among edges. This keeps the graph and degrees but destroys the signed mechanism.
- **ER (baseline).** Erdős–Rényi random graphs matched in size and edge density.

Result (Fig. 3):

network	ALL-3 pass rate	median peak EI
rTCA	100% (30/30)	0.849
NULL-B (allosteric)	30% (9/30)	—
NULL-A (degree)	6.7% (2/30)	0.850
ER	≈ 0%	—

The single-EI trap. NULL-A's median peak EI (0.850) is essentially identical to rTCA's (0.849), yet only 2/30 NULL-A networks pass ALL-3. A criterion based on the scalar alone would misclassify NULL-A as reproducing the phenomenon; the located-and-graded structure does not. This is the empirical justification for the three-property signature.

SI 5. Reversibility index Γ/B

The reversibility index measures how irreversibly the macro-dynamics compress micro-state information. Given the singular values $\{\sigma_k\}$ of the effective dynamics operator (the SVD of the macro-level transition structure, computed in `ce_measure.gamma_alpha`), define $\Gamma = (\sum_k \sigma_k^p)^{1/p}, \Gamma/B =$ Γnormalized by bin count B, with p set so that Γ captures the effective rank (we report the normalized form Γ/B). $\Gamma/B \to$ high when the dynamics are concentrated on a low-dimensional, causally specific attractor; $\Gamma/B \to$ low when dynamics are diffuse.

Within rTCA (`data_rTCA_phase3.npz`, `gnorm_rTCA`): Γ/B rises from 0.267 below threshold to 0.355 above, tracking the EI and effective-rank jumps at $T_{\mathrm{CE}} \approx 3.50$. Effective rank rises $1 \to 3$ (with a brief 4 in the transition band).

SI 6. Feed-withdrawal (autotrophic-boundary) protocol

To test the autotrophic boundary, the heterotrophic feed fraction β_F is swept while holding the loop intact. The engine is run at each (n, β_F) and the peak EI over n is recorded (`data_rTCA_phase4_finefeed.npz`).

β_F	0.004	0.005	0.006	0.007	0.008	0.009	0.010	0.012	0.014	0.02+
peak EI	0.856	0.848	0.864	0.820	0.730	0.481	0.244	0.024	~0	~0

The signature survives in a narrow autotrophic window ($\beta_F \lesssim 0.009$) and collapses to zero by $\beta_F \gtrsim 0.012$–0.014. Removing the autotrophic boundary abolishes the transition even though the closed loop is intact (Fig. 4B).

SI 7. Multi-compartment integration

Two copies of the engine are coupled along two channels, and the integrated peak EI is compared to the sum of the isolated peaks (`data_rTCA_phase5.npz`; `phase5_integration.alpha_eff` / `gamma_eff`):

- **Base** (no coupling): peak EI 0.801.
- **Compartmentalization** (coupling across a compartment boundary): peak EI 0.967; threshold shifts lower, $T_{\mathrm{CE}} \approx 2.50$.
- **Energy coupling** (coupling via a shared energy carrier): peak EI 0.803; *no* threshold shift.
- **Both**: peak EI 0.967.

Superadditivity. The additive prediction for the integrated peak (sum of parts, adjusted) is 0.969; the observed (both-channel) peak is 0.967. Superadditivity $= 0.967 - 0.969 = -0.002$: integration is additive, not synergistic. Compartmentalization shifts the threshold and raises the peak but does so by linear superposition, not by creating new emergence (Fig. 5).

Graded sweep. Over integration strength $\theta \in [0,0.6]$, peak EI rises smoothly from 0.856 to 0.976, confirming the integration effect is graded rather than a discrete jump.

SI 8. Robustness

Bin-count sweep. T_{CE} and peak EI are recomputed across bin counts $B \in \{10,12,14,16\}$ (Fig. S-bsensitivity; `data_rTCA_phase4.npz` family). $T_{\mathrm{CE}} = 3.50$ is stable across B, and peak EI remains ≈ 0.84. The threshold is a dynamical property, not a binning artifact.

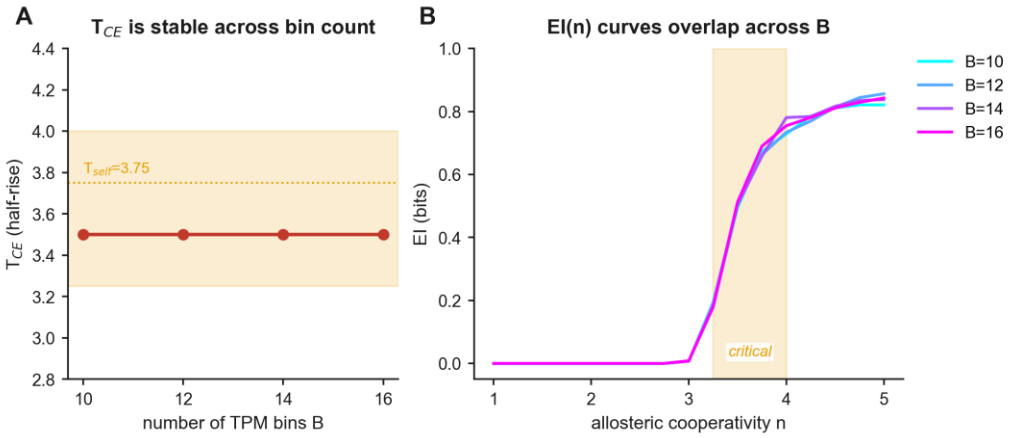


Figure S-bsensitivity. T_CE (A) and EI(n) curves (B) are stable across TPM bin count B.

Noise regime. The perturbations used here act as *observation* noise (they blur the read state without altering the flow), which tends to *inflate* EI; *dynamical* noise (added to the flow itself) tends to *deflate* EI. We report results in the observation-noise regime and flag this explicitly: a full dynamical-noise map is future work.

SI 9. Cross-network comparison: rTCA, Wood–Ljungdahl, syn3A

All three networks are run on the same engine machinery and measured identically (`data_rTCA_phase4ext.npz`).

network	topology	dom. eigenvalue	peak EI (bit)	T_{CE}	effective rank
rTCA	closed loop (autotrophic)	1.0	0.838	3.50	3
Wood–Ljungdahl (WLP)	acyclic (autotrophic)	0.0	0.071	— (no threshold)	2
syn3A core	acyclic (heterotrophic)	0.0	0.041	— (flat)	3

- **rTCA** is the only network with a closed loop (dominant eigenvalue 1.0, signal returns) *and* an autotrophic boundary; it is the only one with a sharp CE threshold.
- **WLP** is autotrophic but acyclic (signal flows through, eigenvalue 0.0); peak EI is 12× lower and shows no threshold.
- **syn3A** (core metabolism of the JCVI minimal cell: glycolysis + non-oxidative PPP + fermentation) is acyclic and heterotrophic; CE stays flat with no threshold (Fig. S-syn3A), consistent with the rule and extending the contrast from a designed engine to an evolved cell (with the caveat that syn3A differs from rTCA along two axes at once: acyclic and heterotrophic).

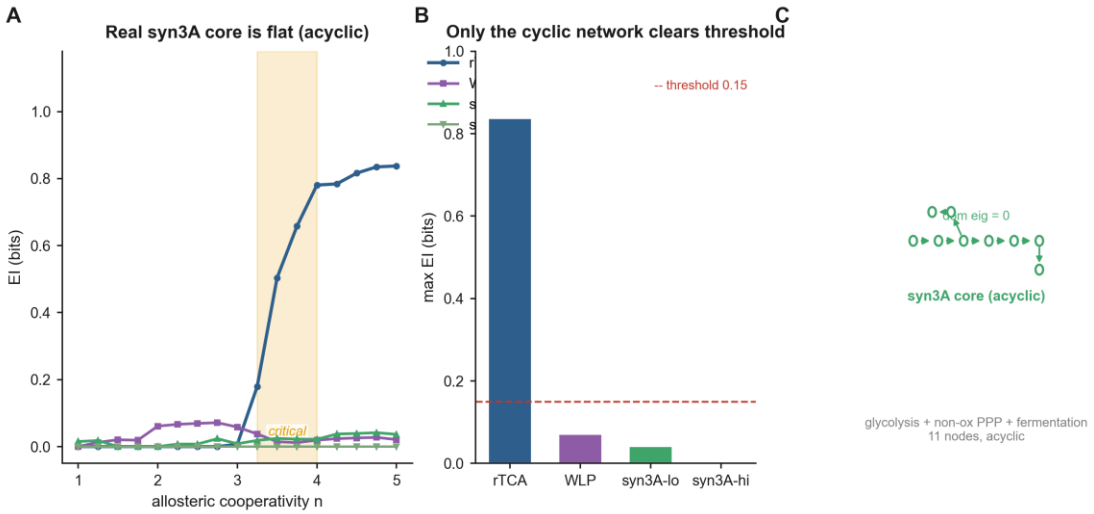


Figure S-syn3A. syn3A core cross-evaluation: EI versus n for four candidate networks (A), max-EI bars (B), and the syn3A acyclic topology (C).

syn3A topology is reconstructed from the essential-metabolism gene list of Breuer et al. (ref. 19); WLP topology from KEGG map00680 and the pathway review (ref. 22).

SI 10. Closest prior applications of causal emergence in biology

To support the claim that effective information has not previously been applied to the running dynamics of a self-sustaining metabolic cycle (Introduction; Materials and methods), we list the closest prior applications. The table is representative, not exhaustive, and is restricted to studies already cited in the main text.

study	system studied	what effective information / CE was applied to	why this is not "running self-sustaining metabolism"
Hoel et al., 2013; Hoel, 2017	toy / general systems	definition of EI; AND-gate toy calibration	foundational theory; no metabolic system
Klein and Hoel, 2020	biological / social / technological networks	CE of random-walk topology via subgraph grouping into macro-nodes	static connectivity, not dynamics
Rosas et al., 2020	multivariate data	partial-information decomposition (ΦID) of synergy	general data-driven decomposition; not specifically metabolic
Zaydman et al., 2022	bacterial proteomes (PPI)	SVD spectral emergence defining hierarchical interactions	static cross-proteome co-variation; not running metabolism
Pigozzi et al., 2025	gene-regulatory networks	integrative CE raised by associative conditioning	regulatory ODE dynamics; not autotrophically self-sustaining metabolism
Tkachenko and Maslov, 2024	prebiotic information-coding polymers	emergence of catalytic cleavage in a chemostat	information polymers, not metabolic chemistry

Across this set, effective information has been applied to network topology (Klein and Hoel, 2020), to gene regulation (Pigozzi et al., 2025), to static co-variation in proteomes (Zaydman et al., 2022), and to information-

polymer dynamics (Tkachenko and Maslov, 2024). To our knowledge it has not been applied to the running dynamics of a self-sustaining, autotrophic metabolic cycle—the object of the present work.

Data availability

All numerical results are computational; no new wet-laboratory data were generated. The primary data are saved NumPy `.npz` archives: the rTCA engine scans `data_rTCA_phase{1,2,3,4,4_finefeed,5,4ext}.npz`, the B-sensitivity sweep `data_rTCA_phase2_bsens.npz`, and the matched null `data_rTCA_phase3_null_matched.npz` (in `R3/sim/`); the bootstrap `data_bootstrap.npz` and the degree-preserving null `data_null_degree.npz` (in `R5/sim/`). Each file is documented in the SI section that analyses it. Figure source data correspond to these archives.

Code availability

The simulation and effective-information pipeline is the R3 codebase in `R3/sim/` (`emergence_model.py`, `emergence_net.py`, `ce_measure.py`, `topology.py`, `rTCA_network.py`, `WLP_network.py`, `syn3A_core_network.py`, `phase5_integration.py`, and the per-phase drivers `phase{1..5,4ext}_*.py`). The null-model analysis (`null_degree_preserving.py`), the bootstrap (`bootstrap_Tself_Tce.py`), and the Figure 3 regeneration script (`fig3_null_signature.py` in `R6/sim/`) build on this codebase, which is read-only and was not modified. The pipeline depends on NumPy, SciPy, and Matplotlib. Source files will be deposited in a public repository upon submission.

Key Resources Table

resource	source / identifier
Software: Python 3.10	https://www.python.org
Software: NumPy, SciPy, Matplotlib	https://numpy.org, https://scipy.org, https://matplotlib.org
rTCA network topology	KEGG map00720 (reductive TCA cycle)
Wood–Ljungdahl pathway topology	KEGG map00680; Ragsdale and Pierce, 2008
syn3A core metabolism	Breuer et al., 2019 (*eLife* 8:e36842)
rTCA simulation outputs (this paper)	`.npz` archives listed in *Data availability*
effective-information codebase (this paper)	`R3/sim/ce_measure.py` and related modules listed in *Code availability*

Introduction

A living cell is, in physical terms, matter that has crossed from disorder into a self-maintaining order: it feeds, it maintains itself, and it reproduces its own organization. How lifeless matter accomplishes this crossing is among the most fundamental open questions in science (Schrödinger, 1944; Anderson, 1972). Nature completed the transition billions of years ago; human engineering has not. We can synthesize biomolecules, assemble gene circuits, and boot a minimal cell from a synthetic genome (Hutchison et al., 2016), yet every such construction is an act of assembly—the pieces are put in place by hand. The open goal is to move from assembling living systems to inducing their emergence, and a first step toward that goal is to locate the transition itself: the point at which disordered chemistry tips into the ordered, self-maintaining dynamics characteristic of life.

Causal emergence provides a quantitative handle on such a transition. A system exhibits causal emergence when a coarse-grained, macroscopic description of its mechanism carries more effective information (EI)—the determinism of its state transitions minus their degeneracy—than the microscopic description (Hoel et al., 2013; Hoel, 2017). The same question can be posed through partial information decomposition (Rosas et al., 2020; Williams and Beer, 2010) or the computational mechanics of ε-machines (Crutchfield, 1994; Shalizi and Crutchfield, 2001). These tools have reached biology: causal emergence has been measured in neural and protein-interaction systems, across the topology of biological, social, and technological networks (Klein and Hoel, 2020), and most recently at the level of gene regulation, where associative training raises the integrative causal emergence of gene-regulatory networks (Pigozzi et al., 2025).

Across this progress, one object has remained out of reach: the running dynamics of a self-sustaining metabolic cycle, the chemistry that keeps a living system alive. Where causal emergence has touched metabolism, it has done so only as static topology. In the network treatment of Klein and Hoel (2020), effective information is a structural property of random walks over a network's connectivity, and causal emergence arises from grouping subgraphs into macro-nodes; metabolism there is a graph to be walked, not a chemistry to be run. Where causal emergence has been applied to dynamics, it has landed on systems that are regulatory rather than metabolic and not autotrophically self-sustaining: the gene-regulatory networks of Pigozzi et al. (2025) are driven by ordinary differential equations but are not the closed, self-feeding chemistry whose persistence is life. A second obstacle is that causal emergence is ultimately a scalar, and a scalar peak is easily matched by structural noise, so locating the transition requires a non-trivial signature rather than a single number.

Here, we apply effective information to the reductive tricarboxylic-acid (rTCA) cycle, an eight-component autocatalytic loop that fixes CO_2 while regenerating its own intermediates and is a leading candidate for the autotrophic core of early metabolism (Muchowska et al., 2019). We treat the cycle not as a graph but as a running, dissipating engine with a sharp self-maintenance threshold, and ask where its causal structure becomes irreducible. We organize the work around a framework we term 'designing for emergence', built on four layers: a target (the disorder-to-order transition to be located), a ruler (effective information, applied under a uniform intervention so that it measures the mechanism alone), a recipe (the structure the ruler reveals to carry the transition—a closed loop plus an autotrophic boundary), and a boundary (where the claim stops, which is itself a result). Figure 6 maps each layer onto the result that delivers it. The computational experiments reported below serve as both ruler and bridge, locating the transition and moving causal emergence from static topology into running dynamics; wet-laboratory validation of the recipe and a discretization-free theory of effective information are explicit future steps rather than claims of this paper.

Results

The rTCA engine has a sharp self-maintenance threshold

We first establish the substrate, which defines the *target* of our framework—the disorder-to-order transition to be located. We model the rTCA cycle as an eight-component autocatalytic network in which each turn regenerates the starting molecule while fixing two molecules of CO_2 (Muchowska et al., 2019), with nonlinear Hill kinetics under a continuous autotrophic substrate supply (Materials and methods; SI 1). The model can either sustain itself—regenerate its intermediates against dilution and loss—or collapse, depending on a single control parameter, the effective catalytic drive n. The transition between these two regimes defines a self-maintenance threshold, T_self.

Across the transition band, ensembles initialized from low and high concentrations diverge in fate, producing a bistable regime in which only the high-concentration branch persists against dilution (Ferrell, 2002; Ferrell and Machleder, 1998). We locate T_self ≈ 3.75 from the peak of the low-initial-condition variance (Figure 1A). The fork is direct: below T_self the low- and high-init ensembles settle to distinct off and on states, and above it only the on state persists (Figure 1C). The threshold is a genuinely nonlinear property: a linearized control with the same connectivity diverges rather than saturating, whereas the nonlinear Hill dynamics remain bounded and cross cleanly into self-maintenance (Figure 1B). The object we measure is therefore not a graph but a running, dissipating engine with a well-defined point at which it turns on. This engine is an idealized computational abstraction—eight lumped components with Hill kinetics, no spatial heterogeneity or enzymatic specificity—so the thresholds reported below are properties of the model rather than of any particular wet biochemistry.

Causal emergence sharpens at the self-maintenance threshold

With the target fixed, we apply the *ruler*. Effective information (EI) quantifies how deterministically and specifically a mechanism maps its inputs to its outputs (Hoel et al., 2013; Hoel, 2017). Operationally, we intervene with a uniform (maximum-entropy) distribution over the discretized micro-states, read out the transition-probability matrix $T_{ij} = P(s_j \mid s_i)$, and compute $\mathrm{EI} = \frac{1}{N_s}\sum_i D_{\mathrm{KL}}\left[T(\cdot \mid s_i) \| T(\cdot)\right] = \text{determinism} - \text{degeneracy}$, where $T(\cdot)$ is the marginal effect over inputs (Materials and methods; SI 2). Because the intervention is uniform, EI measures the mechanism alone and is independent of how often any input happens to occur. Our measurement is Hoel effective information on a discretized transition-probability matrix; the threshold we report is stable across bin counts (SI 8) and is therefore not an artifact of discretization, though we have not exhaustively mapped behaviour outside the Hill-coefficient and feed-rate ranges, and perturbations here act as observation noise—which inflates EI—rather than dynamical noise, which deflates it. A discretization-free continuous effective information (Chvykov and Hoel, 2021) and a neural information squeezer (Zhang and Liu, 2023; Yang et al., 2024) remain future cross-validations. Causal emergence is present when a coarse-grained macro-description of the same mechanism has higher EI than the micro-description.

The result is not a gradual climb. As the engine is driven through its threshold, three independent quantities jump together at a single point T_CE ≈ 3.50: effective information (from ~0 to 0.84 bits), the reversibility index Γ/B (a singular-value measure of how irreversibly the macro-dynamics compress state, rising 0.27 → 0.36), and the effective rank of the dynamics (1 → 3) (Figure 2A). The singular-value spectrum shows the rank transition directly, with two additional singular values emerging in the transition band (Figure 2B). Three criteria that know nothing of one another—a dynamical one (T_self = 3.75), an informational one (T_CE = 3.50), and a structural one (T_rank = 3.25)—agree within a spread of 0.50. To confirm that this agreement is not a coincidence of a single fit, we bootstrapped both thresholds over resampled ensembles and obtained

T_{CE} = 3.50 [3.33, 3.57] and T_{self} = 3.40 [3.20, 3.79] (median [95% interval]); the intervals overlap. The disorder-to-order transition of the engine and the sharpening of its causal structure occur at the same point.

A three-property signature distinguishes the engine from structure-preserving nulls

A single peak value of EI is a weak claim, because a random network can match it by chance. We therefore defined the emergence signature by three properties held simultaneously: reliable (reproduced under bootstrap resampling), located (occurring at a defined threshold rather than anywhere), and graded (rising monotonically through the threshold rather than appearing as an isolated spike). A network satisfies the signature only when all three hold (SI 4). We then tested the engine against two controls designed to be as close to rTCA as a random network can be while destroying the mechanism we suspected of mattering: a degree-preserving null that rewires edges but keeps each node's in- and out-degree, and an allosteric-permutation null that keeps the topology but scrambles the regulatory signs, against an Erdős–Rényi baseline.

Across 30 independent draws of each null (20 for the Erdős–Rényi baseline), the fraction of networks satisfying all three properties fell in a clean gradient: rTCA 100% $\gg$ allosteric-permutation 30% (9/30) $\gg$ degree-preserving 6.7% (2/30) $\gg$ Erdős–Rényi $\approx$ 0% (Figure 3A). The necessity of the three-property criterion is exposed by a trap hidden in the degree-preserving null. Its median peak EI (0.850) is essentially indistinguishable from rTCA's (0.849), yet only 2 of 30 degree-preserving networks satisfy all three properties (Figure 3A,B). A reader who looked only at the peak value would be fooled into thinking degree-preserving noise reproduces the phenomenon (Figure 3C); the located-and-graded structure does not. This is why the signature is defined by three properties rather than by a scalar.

The transition is carried by a closed loop and an autotrophic boundary

If emergence sharpens at the threshold, what structure is doing the carrying? We dissected the engine along two axes that the 'designing for emergence' framework predicts to matter.

The rTCA cycle is a closed circuit: its dominant structural eigenvalue is 1.0, meaning signal returns to its origin. The Wood–Ljungdahl pathway (WLP), an acyclic autotrophic route of comparable biological role, has a dominant eigenvalue of 0.0—signal flows through and leaves (Ragsdale and Pierce, 2008; the three topologies are drawn side by side in Figure 4D). Measured on the same running-engine machinery, the peak effective information follows the loop: rTCA at 0.84 bits is roughly twelve times WLP at 0.07 (Figure 4A). The core metabolism of the minimal cell syn3A—glycolysis, the non-oxidative pentose-phosphate pathway, and fermentation, an acyclic, heterotrophic network (Breuer et al., 2019)—stays flat in causal emergence, peaking near 0.04 bits with no threshold (Figure 4C), which is consistent with the rule. We caution that this comparison rests on three networks only and that syn3A differs from rTCA along two axes at once (acyclic and heterotrophic); the loop-plus-boundary recipe is therefore a model prediction to be tested on a wider panel, not yet an empirically established regularity.

A loop alone is not enough; it must be fed. We withdrew the autotrophic substrate by increasing the heterotrophic feed fraction β_F. The emergence signature survives only within a narrow autotrophic window: peak EI holds near 0.85 for $\beta_F \lesssim 0.009$ and collapses to zero by $\beta_F \gtrsim 0.012$–0.014 (Figure 4B; the full (n, β_F) plane is mapped in Figure 4E). Remove the boundary and the transition vanishes, even though the loop is intact. The two together compose a recipe: a closed loop plus an autotrophic boundary, pushed past the self-maintenance threshold. Each ingredient is necessary; neither alone is sufficient.

Compartmentalization lowers the threshold; integration is additive

Having named the *recipe*, we finally probe the *boundary* of the claim—where the assertion stops—along two directions a skeptical reader would raise. We coupled two copies of the engine across a compartment boundary and across an energy-coupling channel. Compartmentalization shifts the threshold lower, from $T_{CE} \approx 3.50$ to

≈ 2.50, and raises the peak (0.801 → 0.967 bits), whereas energy coupling alone shifts nothing (Figure 5). The macroscopic level at which emergence is sharpest is therefore a real, locatable property of the system rather than an artifact of lumping variables together. Does coupling two engines create *more* emergence than the sum of the parts? The observed peak (0.967 bits) matches the additive prediction (0.969) almost exactly, with superadditivity ≈ −0.002 (Figure 5B): integration here is linear superposition, not synergistic emergence. Across integration strength θ, peak effective information rises smoothly from 0.856 to 0.976, confirming the effect is graded rather than a discrete jump (SI 7).

Is the threshold phenomenon just dissipative adaptation—the relaxation of a single system toward configurations that absorb more work under a scalar thermodynamic feedback (England, 2015)? The distinction matters. Dissipative adaptation describes a single system under scalar feedback; what we observe is an ensemble-level irreducibility that appears only when the running mechanism crosses a threshold and that survives structure-preserving randomization, and a scalar feedback cannot account for the graded null signature reported above. We therefore characterize the phenomenon as a dissipative structure—an organization sustained away from equilibrium by throughput—rather than as dissipative adaptation, and we note this boundary honestly: the macroscopic description at which EI is maximized is, epistemically, a particularly informative way to observe the engine, not a new causal force layered on top of the micro-dynamics (Dewhurst, 2021).

Discussion

These results apply effective information, for the first time to our knowledge, to the running dynamics of a self-sustaining metabolism, and locate the disorder-to-order transition within it. The location is sharp ($T_{CE} \approx 3.50 \approx T_{self}$), the signature is non-trivial (three properties that random networks fail in a graded way), and the carriers are nameable (a closed loop plus an autotrophic boundary). Together they turn 'inducing life's emergence' from a slogan into a design problem with a target and a recipe.

Within the framework introduced above, the methodological contribution is to move the causal-emergence ruler from static topology into running, self-sustaining dynamics. Klein and Hoel (2020) showed that biological networks often have low effective information at the micro-scale but the most pronounced causal emergence after coarse-graining—the regime in which a macroscopic description 'denoises' structure. Our result places that insight inside an engine that is actually running and self-maintaining, where the relevant macroscopic level is dictated by the dynamics rather than chosen after the fact. The recipe we identify—a closed loop plus an autotrophic boundary—is the dynamical analogue of the function-topology 'design tables' that have proved useful in network biology. Ma et al. (2009) showed that only a finite set of three-node topologies can execute biochemical adaptation, turning an open-ended search into a classification and an engineering manual. We find an analogous finiteness for a phenomenon rather than a function: the structural recipe for a disorder-to-order transition is small, nameable, and transfers across a designed cycle, an acyclic pathway, and an evolved minimal cell. The recipe also connects to the long-studied question of what makes a chemical reaction network self-sustaining, from autocatalytic sets (Hordijk et al., 2022) to the hypercycle; our contribution is to add an effective-information ruler that says *where* such a network turns on. For synthetic biology, the recipe is directly prescriptive—a self-sustaining loop with an autotrophic boundary is a buildable target, and the threshold is a knob. For systems biology, the result nominates self-maintenance, rather than complexity or connectivity per se, as the anchor at which emergence sharpens.

The most serious objection to effective-information causal emergence is that the coarse-graining is chosen freely and may simply be selected to maximize EI, so that the macroscopic level is a matter of taste rather than

discovery. Our answer is layered and, we hope, honest. First, the property is robust across causal measures: causal emergence has been shown to be consilient across different formalizations of causation (Comolatti and Hoel, 2025), so the finding does not hinge on the Hoel measure alone. Second, we do not select the macroscopic level to maximize a scalar; we require three simultaneous properties, and we show that the single scalar is precisely what a degree-preserving null can fake. Third, the binning that defines the macroscopic level is swept explicitly—T_CE is stable across bin counts (SI 8)—so the threshold is a dynamical property rather than a binning artifact. We do not claim to have dissolved the philosophical question of whether the macroscopic level is real; we claim that its location and the recipe it implies are objective consequences of the dynamics.

The limitations are noted at their results rather than deferred: the idealized, lumped engine at the threshold result, the discretized and range-limited effective-information ruler at the causal-emergence result (with SI 8), and the additive integration and macroscopic-level status at the integration result. The decisive open test is wet validation of the loop-plus-boundary recipe, taken up below.

The framework points to three concrete next steps: to extend the ruler to more complex self-sustaining systems, including the hub subgraph of the syn3.0 minimal cell and other autotrophic cores, testing whether the loop-plus-boundary rule generalizes beyond rTCA; to validate the recipe wet, by building a self-sustaining loop with a genuine autotrophic boundary—non-enzymatic rTCA chemistry in a fed microcompartment is a candidate substrate (Muchowska et al., 2019)—and pushing it past its threshold, with the prediction of a measurable jump in effective information recoverable from time-series data; and to cross-validate with a continuous ruler, replacing the discretized transition-probability matrix with the continuous effective information of causal geometry (Chvykov and Hoel, 2021) and ultimately a neural information squeezer (Zhang and Liu, 2023; Yang et al., 2024), to confirm that the located transition is not an artifact of discretization or of a hand-chosen coarse-graining.

Materials and methods

Full technical detail is given in the Supplementary Information. The rTCA engine is an eight-component autocatalytic network with Hill kinetics under autotrophic feed (SI 1). Effective information is computed by intervening uniformly over the discretized micro-states to obtain the transition-probability matrix and evaluating determinism minus degeneracy; the implementation is calibrated against a known test system before use (SI 2). The self-maintenance threshold is located from the bistability of low- and high-initial-condition ensembles, and T_CE and T_self are bootstrapped over resampled ensembles (SI 3). The two null models—degree-preserving rewiring and allosteric sign-permutation—with an Erdős–Rényi baseline and the three-property signature criterion are defined in SI 4. The reversibility index Γ/B, the feed-withdrawal protocol, the compartmentalization and energy-coupling integration assay, the bin-count robustness sweep, and the syn3A and Wood–Ljungdahl comparisons are given in SI 5–9. No new wet experiments were performed; all results are computational, reusing the rTCA simulation and effective-information codebase, with the null-model analysis and bootstrap added for this study. To substantiate the gap that motivates this work—that effective information has not previously been applied to the running dynamics of a self-sustaining metabolic cycle—we surveyed the causal-emergence literature (Web of Science and Google Scholar; queries combining "causal emergence" or "effective information" with "metabolism", "metabolic network", or "autocatalytic"; through August 2026). The closest prior applications, all to topology, regulation, or non-metabolic dynamics, are tabulated in SI 10.

Acknowledgements

This work is developed during the 2026 PEBBLE BioFusion Workshop at Westlake University in Hangzhou, China, from July 24th to August 4th. The author gratefully acknowledges Prof. Fangzhou Xiao for his guidance and encouragement, to Chen Sihan for valuable discussions, and to all participants of the 2026 PEBBLE BioFusion Workshop who offered insightful suggestions.

References

Anderson, P. W. (1972). More is different. *Science*, 177, 393–396.

Breuer, M., Earnest, T. M., Merryman, C., et al. (2019). Essential metabolism for a minimal cell. *eLife*, 8, e49256.

Chvykov, P., & Hoel, E. (2021). Causal geometry. *Entropy*, 23, 24.

Comolatti, R., & Hoel, E. (2025). Consilience, combinatory emergence and separability: the notion of cause in complex systems. *Entropy*, 27, 143.

Crutchfield, J. P. (1994). The calculi of emergence: computation, dynamics and induction. *Physica D*, 75, 11–54.

Dewhurst, J. (2021). Causal emergence is agent-relative. Preprint.

England, J. L. (2015). Dissipative adaptation drives self-assembly. *Nature Nanotechnology*, 10, 911–919.

Ferrell, J. E. (2002). Self-perpetuating states in signal transduction: positive feedback, double-negative feedback, and bistability. *Current Opinion in Cell Biology*, 14, 140–148.

Ferrell, J. E., & Machleder, E. M. (1998). The biochemical basis of an all-or-none cell fate decision at the metaphase-to-anaphase transition. *Science*, 280, 895–898.

Hoel, E. P. (2017). When the map is better than the territory. *Entropy*, 19, 188.

Hoel, E. P., Albantakis, L., & Tononi, G. (2013). Quantifying causal emergence shows that macro can beat micro. *Proceedings of the National Academy of Sciences USA*, 110, 19790–19795.

Hordijk, W., Steel, M., & Kauffman, S. (2022). Autocatalytic sets arising in a combinatorial model of chemical evolution. *Life*, 12, 1703.

Hutchison, C. A., Chuang, R.-Y., Noskov, V. N., et al. (2016). Design and synthesis of a minimal bacterial genome. *Science*, 351, aad6253.

Klein, B., & Hoel, E. (2020). The emergence of informative higher scales in complex networks. *Complexity*, 2020, 8932526.

Ma, W., Trusina, A., El-Samad, H., Lim, W. A., & Tang, C. (2009). Defining network topologies that can achieve biochemical adaptation. *Cell*, 138, 760–773.

Muchowska, K. B., Varma, S. J., & Moran, J. (2019). Synthesis and breakdown of universal metabolic precursors promoted by iron. *Nature*, 569, 104–107.

Pigozzi, F., Goldstein, A., & Levin, M. (2025). Associative conditioning in gene regulatory network models increases integrative causal emergence. *Communications Biology*, 8, 714.

Ragsdale, S. W., & Pierce, E. (2008). Acetogenesis and the Wood–Ljungdahl pathway of CO_2 fixation. *Annual Review of Biochemistry*, 77, 747–767.

Rosas, F. E., Mediano, P. A. M., Jensen, H. J., et al. (2020). Reconciling emergences: an information-theoretic approach to identify causal emergence in multivariate data. *PLOS Computational Biology*, 16, e1008289.

Schrödinger, E. (1944). *What Is Life? The Physical Aspect of the Living Cell.* Cambridge University Press.

Shalizi, C. R., & Crutchfield, J. P. (2001). Computational mechanics: pattern and prediction, structure and simplicity. *Journal of Statistical Physics*, 104, 817–879.

Williams, P. L., & Beer, R. D. (2010). Nonnegative decomposition of multivariate information. Preprint, arXiv:1004.2515.

Yang, M., Wang, Z., Liu, J., & Rong, Y. (2024). Finding emergence in data by maximizing effective information. *National Science Review*.

Zhang, J., & Liu, R. (2023). Neural information squeezer for causal emergence. *Entropy*, 25, 26.

Figure legends

Figure 1. The rTCA engine has a sharp self-maintenance threshold. (A) Variance of the endpoint state across the low-initial-condition ensemble versus catalytic drive *n*; the peak at T_self ≈ 3.75 locates the self-maintenance transition, the band in which the low-init ensemble is recruited into the sustained branch. (B) The threshold is a nonlinear property: a linearized engine with identical connectivity and gains (dominant eigenvalue 1.33, above the dilution rate 0.30) diverges, whereas the full Hill dynamics remain bounded and cross cleanly into self-maintenance. The object on which effective information is computed is therefore a running, dissipating engine with a well-defined on-point, not a static graph. (C) Bifurcation diagram: the low- and high-initial-condition branches (with ensemble uncertainty) diverge below T_self—bistable, with off and on states coexisting (amber band)—and merge above it, where only the on state persists.

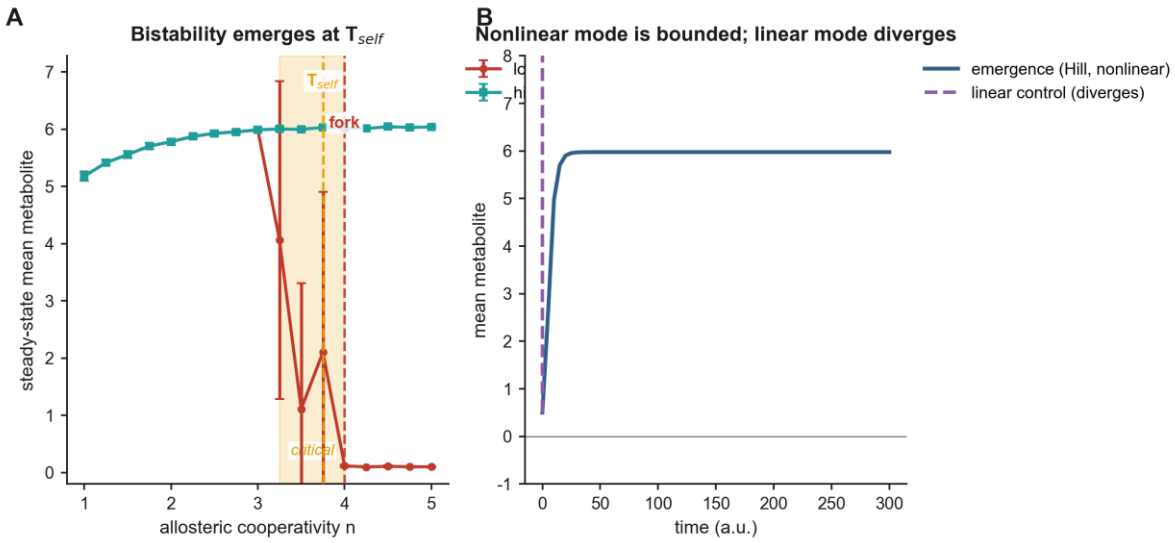


Figure 1A,B. Self-maintenance threshold: bistability of the low-/high-initial-condition ensembles (A) and bounded nonlinear versus divergent linear dynamics (B).

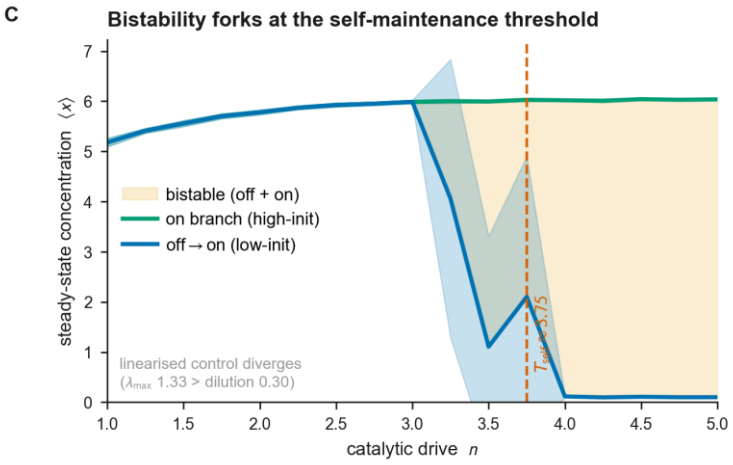


Figure 1C. Bifurcation diagram: the two branches fork below T_self and merge above it.

Figure 2. Causal emergence sharpens at the self-maintenance threshold. (A) Three independent quantities jump together at a single point T_CE ≈ 3.50: effective information (0 → 0.84 bits), the reversibility index Γ/B (0.27 → 0.36), and the effective rank of the dynamics (1 → 3). Three criteria that share no parameter—a dynamical one (T_self = 3.75), an informational one (T_CE = 3.50), and a structural one (T_rank = 3.25)—agree within a spread of 0.50; bootstrap 95% intervals (T_CE 3.33–3.57; T_self 3.20–3.79) overlap. (B) Singular-value spectrum of the macro-level transition structure, showing two additional singular values emerging in the transition band; the effective-rank transition is read directly off this spectrum.

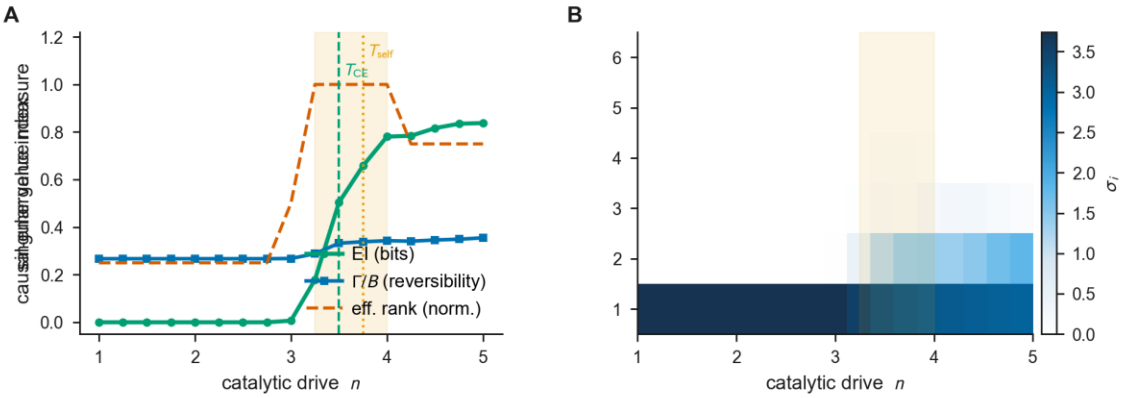


Figure 2. Effective information, reversibility index Γ/B, and effective rank jump together at T_CE (A); the singular-value spectrum shows the rank transition (B).

Figure 3. A three-property signature distinguishes the engine from structure-preserving nulls. (A) Fraction of networks satisfying all three properties (reliable, located, graded) over 30 independent draws of each null (20 for the Erdős–Rényi baseline): rTCA 100% ≫ allosteric-permutation 30% (9/30) ≫ degree-preserving 6.7% (2/30) ≫ Erdős–Rényi ≈ 0%. (B) The single-peak trap: the degree-preserving null's median peak effective information (0.850) is essentially identical to rTCA's (0.849), yet only 2 of 30 degree-preserving networks pass all three properties. A criterion based on the scalar alone would misclassify degree-preserving noise as reproducing the phenomenon; the located-and-graded structure does not. (C) Six example degree-

preserving null draws (vermilion, faint) versus the rTCA engine (blue): individual nulls can reach a peak comparable to rTCA's, yet their trajectories lack the located, graded rise at T_CE that defines the signature.

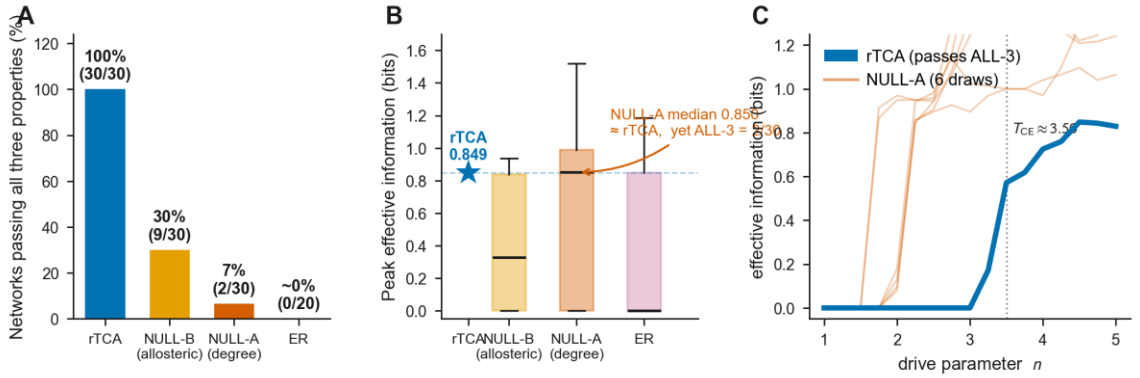


Figure 3. Three-property signature versus structure-preserving nulls: ALL-3 pass-rate gradient (A), the single-peak trap (B), and example degree-preserving null curves versus rTCA (C).

Figure 4. The transition is carried by a closed loop and an autotrophic boundary. (A) Peak effective information follows the loop: rTCA, a closed circuit (dominant eigenvalue 1.0, signal returns to its origin), peaks at 0.84 bits—roughly twelve times the acyclic Wood–Ljungdahl pathway (eigenvalue 0.0, signal flows through and leaves) at 0.07 bits—measured on identical running-engine machinery. (B) Feed-withdrawal: peak effective information holds near 0.85 for heterotrophic feed fraction $\beta_F \lesssim 0.009$ and collapses to zero by $\beta_F \gtrsim 0.012$–0.014, so removing the autotrophic boundary abolishes the transition even with the loop intact. (C) The core metabolism of the minimal cell syn3A (glycolysis + non-oxidative pentose-phosphate pathway + fermentation; acyclic and heterotrophic) stays flat, peaking near 0.04 bits with no threshold. (D) Network topology of the three engines side by side: the rTCA closed cycle (dominant eigenvalue 1.0, signal returns) versus the acyclic Wood–Ljungdahl pathway and syn3A core (eigenvalue 0.0, signal flows through); red edges mark the cooperative, allosterically regulated steps (Hill n > 1). (E) Two-dimensional map of effective information over drive n and feed fraction β_F: the signature survives only in a narrow autotrophic window ($\beta_F \lesssim 0.009$) and collapses by $\beta_F \approx 0.01$.

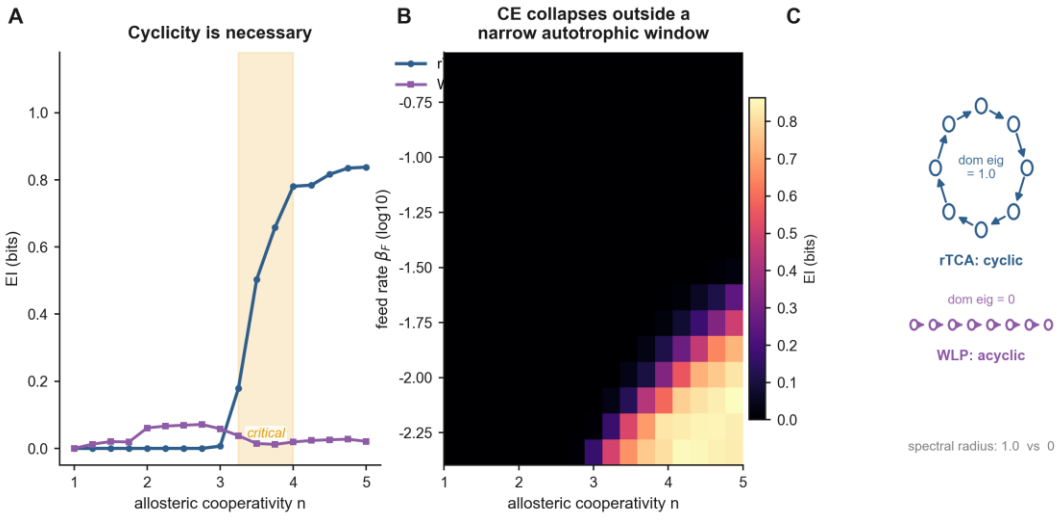


Figure 4A,B. Cyclicity is necessary (rTCA versus the acyclic Wood–Ljungdahl pathway) and the feeding-collapse map.

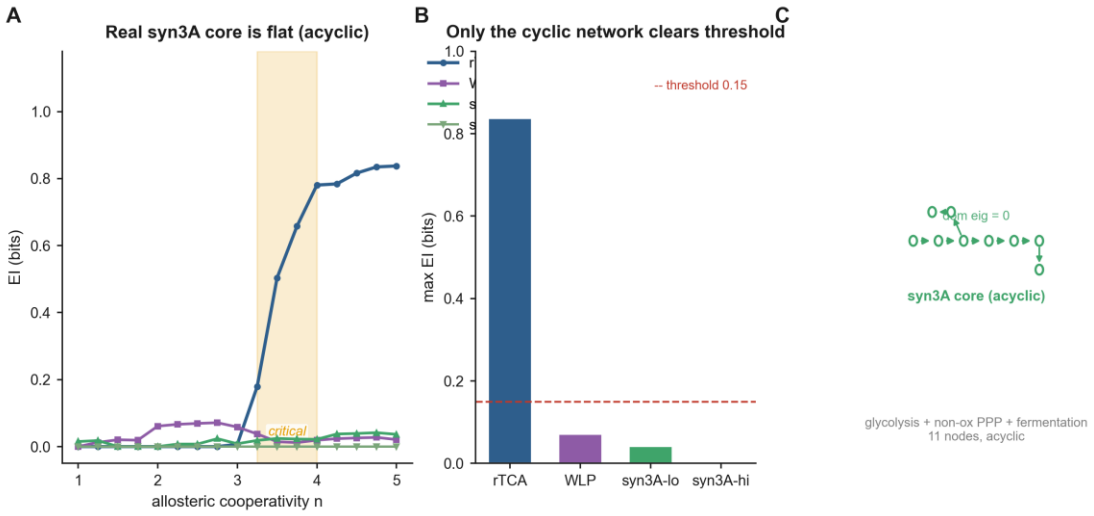


Figure 4C. The syn3A minimal-cell core stays flat (acyclic, heterotrophic).

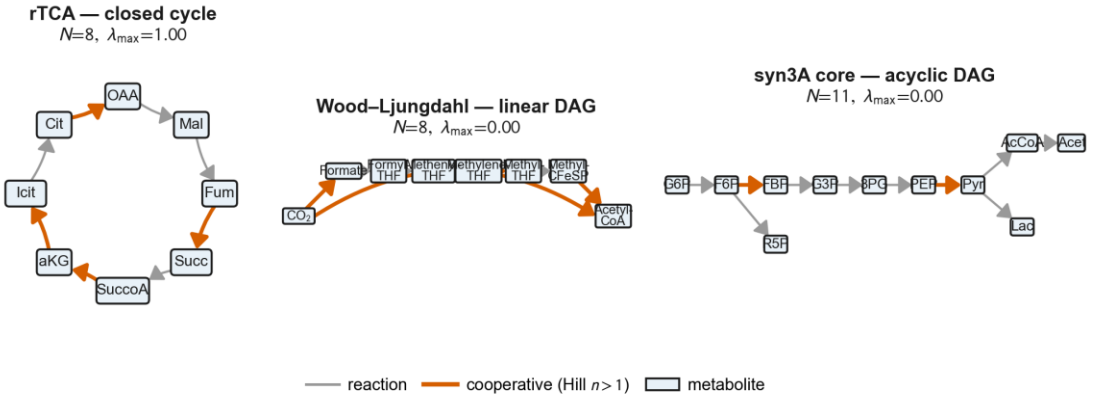


Figure 4D. Network topology of the three engines side by side.

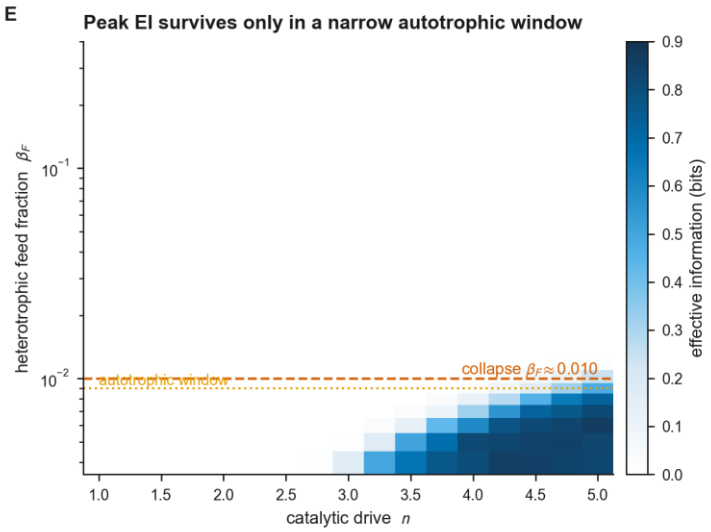


Figure 4E. Two-dimensional effective-information map over drive n and feed fraction β_F.

Figure 5. Compartmentalization lowers the threshold; integration is additive. (A) Coupling two engines across a compartment boundary shifts the threshold lower (T_CE ≈ 3.50 → ≈ 2.50) and raises the peak (0.801 → 0.967 bits); energy coupling alone shifts neither. The macroscopic level at which emergence is sharpest is therefore a real, locatable property of the system rather than an artifact of lumping variables. (B) The integrated peak (0.967 bits) matches the additive prediction (0.969) almost exactly; superadditivity ≈ −0.002, so integration here is linear superposition rather than synergistic emergence.

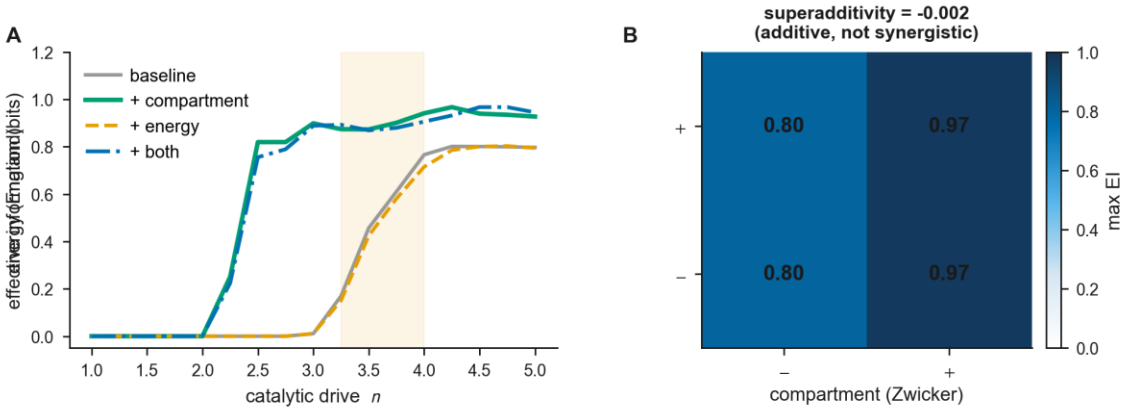


Figure 5. Compartmentalization shifts the threshold and raises the peak (A); the integrated peak matches the additive prediction (B).

Figure 6. The 'designing for emergence' framework. Each of the four layers is both a step in the analysis and a result of it: the *target* (the disorder-to-order transition, located as the self-maintenance threshold T_self; Figure 1); the *ruler* (effective information, under a uniform intervention, whose causal emergence sharpens at T_CE ≈ T_self; Figure 2); the *recipe* (the structure carrying the transition—a closed loop plus an autotrophic boundary; Figure 4); and the *boundary* (where the claim stops—integration is additive and the phenomenon is a dissipative structure; Figure 5).

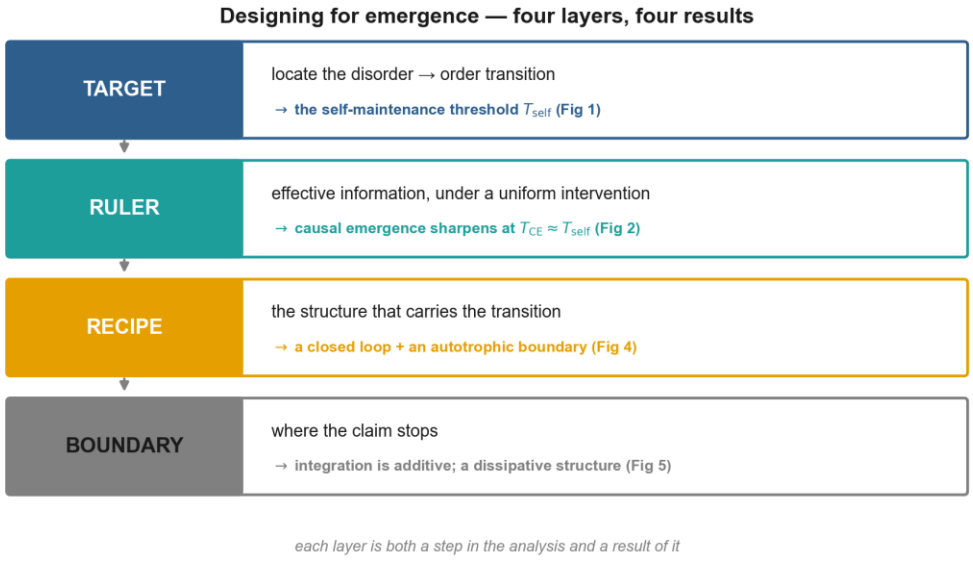


Figure 6. The 'designing for emergence' framework: target, ruler, recipe, boundary.