

Certified selective interventions and closed-RAF structure in a pooled prokaryotic network

14 September 2026

Abstract

We give a source-specific, computer-assisted analysis of catalytic food-generated reaction structure in a pooled prokaryotic network with 6,039 reaction actions and 9,231 directed reactions. A three-action intervention prevents structural L-valine production while preserving L-methionine production. Ranked restoration witnesses prove inclusion-minimality. An 18-species siphon makes the obstruction independent of catalyst annotations and robust to food additions outside the siphon; with catalyst weakening, the same witnesses establish robust selective minimality. In a ten-reaction valine subsystem, the irreducible RAFs are two singleton seeds, the sole valine-producing RAF is the full subsystem, and there are exactly two closed RAFs. Adding one food-enabled pooling operation collapses this closed-RAF catalogue to one element and identifies the parent closed extensions of the two closed subsets. We also give a constructive lower-bound/RAF upper-bound criterion for parent closed-RAF uniqueness and reproduce equality of the corresponding reaction sets before and after intervention. Conventional proofs are accompanied by Lean certificates for the stated formalized results and a separate audit against the pinned source. The conclusions concern structural generation under an explicit medium, not reactor depletion or concentration equilibria.

1 Introduction

Autocatalytic-set theory describes how a collection of reactions can be sustained by a food supply while its catalysts are supplied or produced within the collection. Its biochemical applications already include metabolic irrRAF sampling and sensitivity to reaction and catalyst removal [6], as well as analysis of prokaryotic networks [8, 7]. More recent work develops algorithms for enumerating irreducible and closed autocatalytic sets [4]. The present contribution is a certified finite case study: a selective intervention, its environmental boundary, and an exact separation between minimal catalytic seeds and target-producing support.

Three distinctions are central. Inclusion-minimal deletion is different from globally minimum deletion cost. A minimal standalone RAF is different from a minimal RAF producing a specified molecule. Finally, closedness is relative to an ambient reaction set. We exhibit each distinction on literal reaction data, preserving conjunctive catalyst requirements and the grouping of reversible reactions. Our proofs use short structural certificates: an inaccessible species set, necessary producers, and constructive activation ranks. They do not require completing the parent irrRAF catalogue.

2 Model, source and intervention semantics

Let X be a finite species set, R a finite set of directed reactions, and $F \subseteq X$ the supplied food. Each $r \in R$ has reactants $I(r)$, products $O(r)$, an action identifier $a(r)$, and a positive catalyst formula ϕ_r . The formula uses atoms in X , conjunction and disjunction; $\phi_r(Y)$ denotes evaluation in $Y \subseteq X$. In particular, $Y \subseteq Z$ implies $\phi_r(Y) \Rightarrow \phi_r(Z)$.

For $S \subseteq R$, $\text{cl}_S(F)$ is the least set containing F and satisfying

$$r \in S, \quad I(r) \subseteq \text{cl}_S(F) \implies O(r) \subseteq \text{cl}_S(F).$$

Catalysts are absent from this closure operation. They are tested subsequently in the final closure, as required by RAF semantics [3].

Definition 1. A nonempty $S \subseteq R$ is a RAF if, for every $r \in S$,

$$I(r) \subseteq \text{cl}_S(F), \quad \phi_r(\text{cl}_S(F)).$$

It is irreducible if it has no proper RAF subset. It is closed relative to R if it is a RAF and every $r \in R$ satisfying these two support conditions already lies in S . The target capability predicate is

$$\text{Cap}_p(U; F) \iff \exists S \subseteq U : \text{RAF}(S; F) \text{ and } p \in \text{cl}_S(F).$$

For a set of action identifiers D , write $R_D = \{r \in R : a(r) \notin D\}$. A successful cut has $\neg \text{Cap}_p(R_D; F)$ and is inclusion-minimal if no proper subset of D is successful. This quantification fixes the food and source semantics.

The source is `prokaryotic-network.crs`, distributed with CatReNet [5, 2], at the commit identified in Supplement S1. We retain all 6,039 source rows and their catalyst formulas. Splitting 3,192 reversible rows yields 9,231 directions. All directions of a source row have the same action identifier. Stoichiometric coefficients remain in the source archive but do not enter the structural presence/absence predicate.

The food set is *exactly* the 68 entries of the pinned file’s food declaration, reproduced in Supplement S1. This includes C00003 and the formal `Pooling` marker, as well as C00019 and C00698. The pooled catalyst symbol `NADs` itself is not supplied. Formal pooling and spontaneous-reaction conventions are retained. This source pools prokaryotic reactions; it is not an organism-specific reconstruction.

Our target is L-valine, C00183, and our preservation control is L-methionine, C00073. Neither belongs to F . The intervention is

$$D = \{\text{R01209}, \text{R01210}, \text{R04441}\}. \tag{1}$$

The first and third actions produce the valine precursor C00141. The reverse direction of R01210 also produces it. R01210 is reversible; hence D removes four directed reactions. None of these actions directly produces valine. Full equations and catalyst formulas are supplied in Supplement S2.

3 A selective inclusion-minimal intervention

Lemma 2 (Food-independent siphon obstruction). *Suppose $B \subseteq X$ satisfies*

$$r \in U, \quad O(r) \cap B \neq \emptyset \implies I(r) \cap B \neq \emptyset. \tag{2}$$

For every F' disjoint from B , $\text{cl}_U(F') \cap B = \emptyset$.

Proof. Induct on a derivation of food generation. A food molecule is outside B . If a generated product belonged to B , condition (2) would supply a reactant in B . That reactant has an earlier generation derivation, contradicting the induction hypothesis. $\square$

Theorem 3 (Selective inclusion-minimal cut). *For the pinned source and food set, the cut (1) eliminates L-valine capability and preserves L-methionine capability. It is inclusion-minimal.*

Table 1: Restoration certificates. The restored action is removed from the deletion set. Witness sizes count directed reactions.

Restored action	Valine witness size	Target verdict
R01209	9	Present
R01210	13	Present
R04441	13	Present
None	—	Absent by the siphon certificate

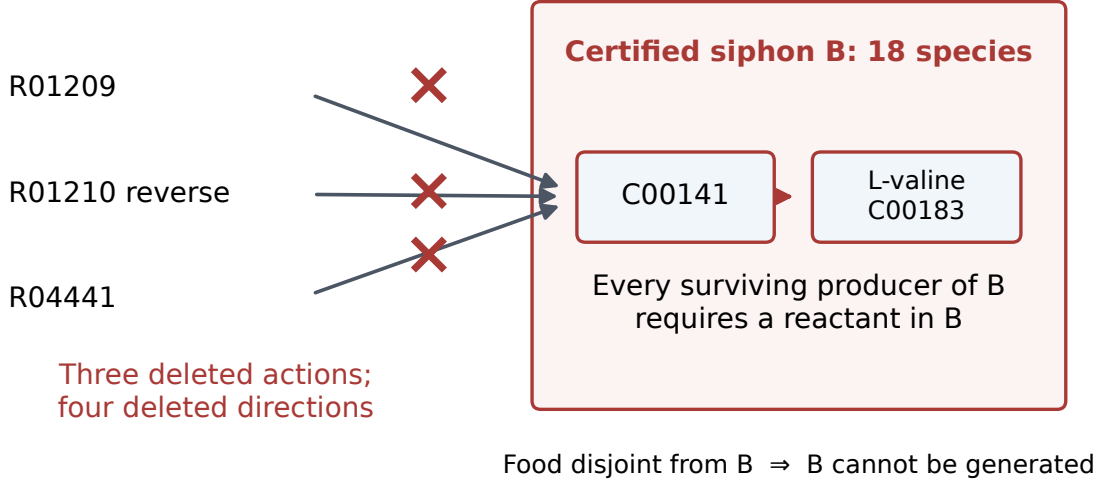
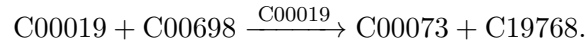


Figure 1: Logic of the intervention certificate. The three action deletions block access to a certified 18-species siphon containing the precursor and valine. The internal arrow indicates the precursor relationship, not an assertion that there is only one biochemical route. The full-source producer condition excludes every surviving bypass.

Proof. The certificate lists an 18-species set B containing C00183, disjoint from F . A check over all 9,231 source directions proves (2) for $U = R_D$. The complete membership of B and its 58 surviving producing directions are provided in the supplement. Lemma 2 excludes valine even from the catalyst-free reactant closure of R_D , and therefore from every RAF contained in it.

For each $a \in D$, a ranked witness in $R_{D \setminus \{a\}}$ is a RAF and generates valine. The three witnesses have 9, 13 and 13 directions (Table 1). Their reaction membership, reactant ranks, positive catalyst formulas in the final closure, and target production are checked. If a successful proper subcut $E \subsetneq D$ existed, choose $a \in D \setminus E$. Then $E \subseteq D \setminus \{a\}$, so $R_{D \setminus \{a\}} \subseteq R_E$. The corresponding witness survives in R_E , a contradiction.

Finally, the forward direction of the reversible source row R09639 is the singleton witness



Its reactants and catalyst are food, and its action is not deleted. Thus methionine remains capable. $\square$

Corollary 4 (Robust selective minimality). *Let $F \subseteq F'$ with $F' \cap B = \emptyset$. Keep reaction reactants, products and action identifiers fixed, and replace each ϕ_r by a positive formula ψ_r satisfying*

$$\phi_r(Y) \implies \psi_r(Y) \quad \text{for every } Y \subseteq X.$$

Then the same cut remains inclusion-minimal for valine, and methionine remains capable.

Table 2: The valine subsystem. Forward and reverse refer to the pinned source equations.

Index	Direction	Index	Direction
0	R00014 forward	5	R01215 reverse
1	R04672 reverse	6	R04441 forward
2	R10916 reverse	7	R00221 reverse
3	R10985 forward	8	R00589 reverse
4	R00585 forward	9	R_NADs_2 forward

Proof. The negative result uses only (2), hence remains true for arbitrary catalyst annotations. Food enlargement preserves every old generation derivation. Formula weakening then preserves the three restoration RAFs and the methionine RAF. Apply the proof of Theorem 3 with these same witnesses. $\square$

The negative obstruction also survives further reaction deletions, food additions outside B , and additions of reactions satisfying (2). These statements do not require catalyst weakening. In contrast, the positive restoration conclusion is not asserted for arbitrary catalyst strengthening, since strengthening may invalidate a witness.

The environmental boundary is concrete. With C00141 supplied as additional food, the four directions R00221 reverse, R00589 reverse, R00585 forward and R01215 reverse form a valine-producing RAF after the cut. This rescue is certified. It does not say that adding any member of B suffices.

4 Seeds, target support and ambient closedness

Let H be the ten-direction subsystem in Table 2, with the original F , and let $L = \{0, 1, 4, 7, 8\}$. Indices in this section refer to that table. The source identities, rather than the indices alone, define the subsystem.

Theorem 5 (Exact subsystem structure). *Relative to the ambient set H and the source food set:*

1. the irreducible RAFs are exactly $\{0\}$ and $\{7\}$;
2. the only RAF producing valine is H itself;
3. the closed RAFs are exactly L and H .

Proof. Every nonempty food-generated RAF contains a reaction whose reactants are food. Otherwise induction on generation gives no molecule beyond food, contradicting the reactant condition for every candidate reaction. A source check shows that only 0 and 7 have food-contained reactants in H , and each singleton is itself a RAF. Every irreducible RAF must therefore be one of these two singletons.

For target support, a nonfood generated molecule must have a producer in the chosen reaction set. Exhausting the producers within H gives the following necessities. Valine forces 5. The two nonfood reactants of 5 force 4 and 6. Reaction 6 forces 2; the inputs of 2 force 1 and 3; and 1 forces 0. The serine requirement of 3 forces 8, which forces 7. The NAD catalyst formula of 3 forces 9, the only producer of the pooled NAD catalyst in H . Each forced species is absent from food. These ten requirements exhaust H . Conversely, a ranked witness verifies that H is a valine-producing RAF.

For closedness, the prefix activation order 0, 7, 1, 8, 4 forces every closed RAF to contain L : at each step its reactants and catalyst are available from food and earlier products. Direct support and upper-pool checks certify that L is closed. Suppose a closed RAF contains a reaction

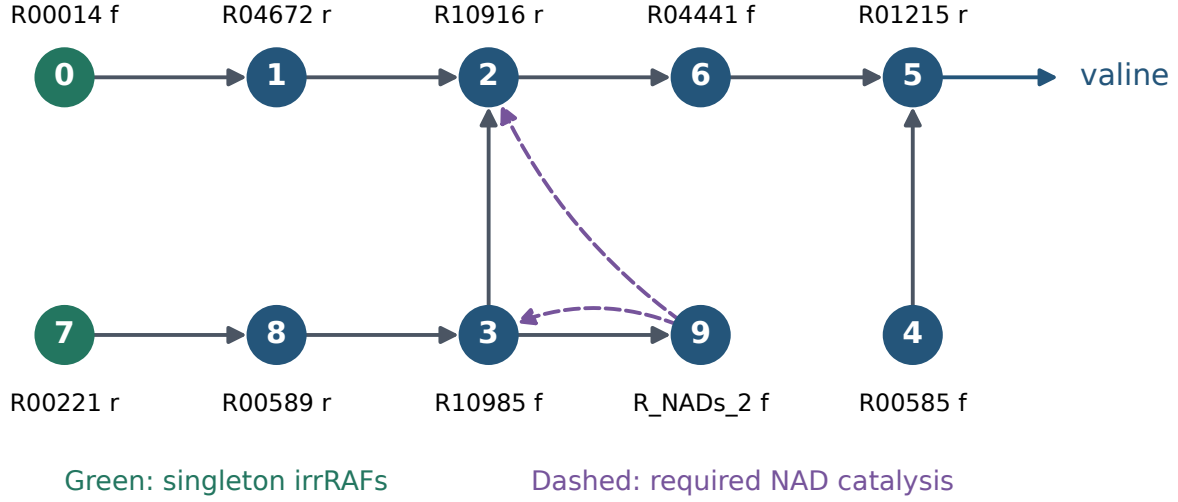


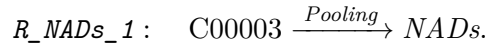
Figure 2: Necessary producer relationships for valine support in H . Solid arrows show relevant reactant dependencies; dashed arrows show required NAD catalysis. The two green reactions are the complete irreducible catalogue, but every reaction is required for the target. This is a dependency diagram, not a complete stoichiometric reaction graph.

outside L . The possible extra indices are 2, 3, 5, 6, 9. Each forces 3, either directly or through $5 \Rightarrow 6 \Rightarrow 2 \Rightarrow 3$; 9 requires the nonfood product of 3. Once 3 is present, closedness successively forces 9, 2, 6, 5. Thus all of H is present. The full H is a RAF and is closed relative to itself. $\square$

Corollary 6 (Information lost by the irreducible catalogue). *The ambient networks L and H have the same food set and the same complete irrRAF catalogue, but only H has valine capability.*

Proof. RAF status and irreducibility of a fixed subset use its own reactions and the common food set. Both singleton RAFs lie in L . Theorem 5 therefore gives the same irreducible catalogue in both ambient networks. Its target-support statement excludes valine capability in the proper subset L . $\square$

Theorem 7 (Pooling changes the closed catalogue). *Let q be the omitted source direction*



The augmented subsystem $H \cup \{q\}$ has exactly one closed RAF, namely all eleven directions. Every parent closed RAF contains H . Consequently the least parent closed extensions of L and H coincide.

Proof. Both the reactant C00003 and the catalyst marker Pooling are food. The order

$$q, 0, 1, 7, 8, 3, 9, 2, 4, 6, 5 \tag{3}$$

activates all eleven directions with catalysts checked against the prefix pool. In any closed RAF of the augmented or parent network, induction through this order forces each direction: the prefix products are generated, and closedness includes the next supported reaction. The augmented set is itself a RAF, so it is its unique closed RAF. The literal embedding into the parent proves the second assertion. Since every parent closed candidate already contains H , the families of parent closed supersets of L and H are identical. Their least elements are therefore identical; existence follows from the parent maximum and the bounds proved in the next section. $\square$

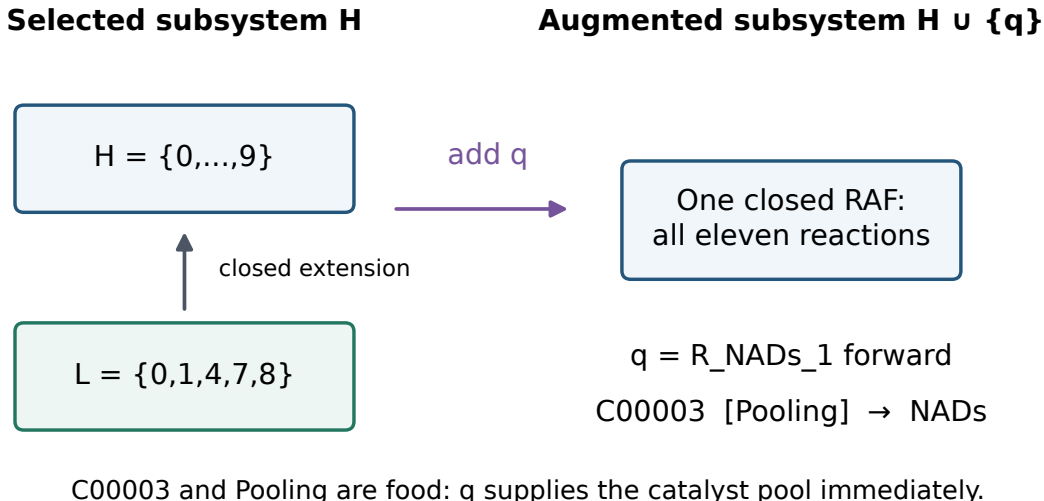


Figure 3: Closedness depends on the ambient reaction set. The food-enabled formal pooling row q supplies NADs without the bootstrap dependency present in H . The two closed subsets of H do not represent two distinct parent closed organizations or two inherited concentration states.

The added row is a formal operation in the source’s representation of catalyst availability. It should not be interpreted as a newly proposed physical conversion of a molecule into an abstract chemical pool.

5 Constructive activation and parent uniqueness

Let A denote the maximal constructively activated set: start from food and repeatedly include any ambient reaction whose reactants and catalyst formula hold in the current pool, adding its products. Finiteness and positivity make the terminal set independent of the activation order. Let M be the maxRAF.

Lemma 8 (Meeting lower and upper bounds). *Every closed RAF C satisfies $A \subseteq C \subseteq M$. If $A = M \neq \emptyset$, then M is the unique closed RAF.*

Proof. For $A \subseteq C$, induct through constructive activation exactly as in (3). Every prefix product lies in $\text{cl}_C(F)$, and positive catalyst formulas remain true in this larger set, so closedness forces the next reaction. The inclusion $C \subseteq M$ is maxRAF maximality. The maxRAF is closed: adding any omitted reaction supported by its closure would give a larger RAF. Equality of the bounds now forces $C = M$. $\square$

The pinned instances give Table 3. Equality here is equality of reaction sets, not merely cardinalities. The constructive certificates have 23 synchronous activation layers in both models. The upper certificates successively bound every RAF using food-containing reactant-closed pools and rejection of unsupported reactions. They reproduce the pruning sequences

$$9231 \longrightarrow 2823 \longrightarrow 2148 \longrightarrow 2148, \quad 9227 \longrightarrow 2584 \longrightarrow 2085 \longrightarrow 2085.$$

The certificates and their formal-verification status are specified in Section 7 and the supplement.

The corresponding parent irrRAF catalogue is a separate problem. We do not claim to have completed that irreducible catalogue. Its incompleteness does not weaken a target-specific intervention certificate or a closed-RAF uniqueness certificate.

Table 3: Full-parent set comparisons. A is the constructively activated set and M the maxRAF.

Ambient model	$ M $	$ A $	Set comparison
Parent R	2,148	2,148	$A = M$
After intervention R_D	2,085	2,085	$A = M$

Table 4: Evidence boundaries.

Evidence	Meaning
Lean theorems	Propositions about the literal encoded finite model.
Source read-back audit	Agreement of rows, catalyst formulas, food and action mapping with the pinned file.
Exact computation	Finite results outside the kernel where explicitly identified.
Biochemical interpretation	Consequences conditional on the source abstraction and medium.

6 Firing and kinetic interpretation

Condition (2) is the usual siphon condition on the directed reactant/product network. Its connection with invariant extinction faces is established prior work; in particular, Angeli, De Leenheer and Sontag give the corresponding equivalence in Proposition 5.5 [1].

Corollary 9 (No startup by surviving firings). *Starting with every species of B absent, any finite sequence of surviving reaction firings leaves every species of B absent, provided external supply introduces no member of B .*

Proof. Induct on the number of firings. A firing producing a member of B would require a reactant in B by (2), contradicting its absence immediately before the firing. Consuming molecules cannot invalidate absence. $\square$

For conventional nonnegative, locally Lipschitz, reactant-respecting kinetics, with no inflow into B , the zero face $x_i = 0$ for $i \in B$ is forward invariant: every rate producing a barrier species vanishes on that face, and every rate consuming it also vanishes. Uniqueness of the local ODE solution preserves the face throughout its existence. This kinetic interpretation is not an additional Lean ODE theorem. Neither it nor the firing result proves depletion when barrier species are initially present. Likewise, a unique closed RAF neither implies a unique concentration equilibrium nor excludes kinetic multistability.

7 Verification and reproducibility

The formal model uses finite reaction lists extensionally, inductive food generation, and positive-formula satisfaction. The proofs distinguish four kinds of evidence (Table 4). No unproved biological assumption is disguised as a Lean certificate hypothesis: the theorem refers to literal source data, and the source interpretation is audited separately.

The original selective intervention, small irreducible catalogue, exact target-support and closed-subsystem classifications, pooling extension, robust selective minimality, precursor rescue and firing corollary have explicit Lean declarations. The proof checker uses kernel reduction of certificates, not `native_decide`; strict compilation rejects warnings and `sorry`. The only logical axioms in the reported proofs are propositional extensionality, classical choice where used, and quotient soundness. The text parser and biochemical annotations are not kernel-verified.

Both concrete parent uniqueness statements are kernel-certified: the parent and postcut modules prove that an ambient subset is a closed RAF if and only if it is extensionally equal to

the respective literal maximum. The cardinalities, pruning totals and layer counts are separately reproduced exact computational data. The common least parent closed extension of the two small closed subsets is the parent maximum.

The supplement supplies the full food and siphon lists, reaction equations, witness identifiers, ranks and certificate paths, toolchain and source fingerprints, verification declarations, and reproduction commands. Compiler diagnostics, AGC history and experimental route changes remain in the research ledger rather than the mathematical argument.

8 Discussion

The intervention theorem is a statement about structural generation under a fixed source-derived medium. Its main strength is that the negative proof does not depend on catalyst annotations, while the positive witnesses make the precise assumptions for selective minimality visible. The precursor-rescue certificate demonstrates an actual environmental boundary rather than a generic sensitivity warning.

The small subsystem shows why a complete irrRAF catalogue cannot serve as a sufficient statistic for biochemical capability. It also shows how an omitted formal pooling operation changes a closed catalogue. Closedness must therefore be reported together with the ambient reaction set and catalyst-pool convention. Stoichiometric self-amplification, growth and persistence are different properties and are not inferred from RAF minimality.

The contribution is the concrete combination of selective intervention, structural robustness, exact target-support separation and ambient closed-RAF analysis, with a traceable boundary between formal proof and source interpretation. A universal enumeration algorithm and a global intervention optimum remain outside its scope.

References

- [1] David Angeli, Patrick De Leenheer, and Eduardo D. Sontag. A Petri net approach to the study of persistence in chemical reaction networks. [arXiv:q-bio/0608019v2](#), 2006. Proposition 5.5.
- [2] CatReNet contributors. Prokaryotic catalytic reaction network, pinned source, 2026. [Source file](#), commit fcb11fb1027f3378075f3a9ef6fc84fae776ff88; accessed 14 September 2026.
- [3] Wim Hordijk. A concise and formal definition of RAF sets and the RAF algorithm. [arXiv:2303.01809](#), 2023.
- [4] Wim Hordijk. Algorithms for enumerating irreducible and closed autocatalytic sets, 2026. [Preprint 202604.1787](#), version 1, 24 April 2026.
- [5] Daniel H. Huson, Joana C. Xavier, and Mike A. Steel. CatReNet: interactive analysis of (auto-) catalytic reaction networks. *Bioinformatics*, 40(8):btae515, 2024. [doi:10.1093/bioinformatics/btae515](#).
- [6] Filipa L. Sousa, Wim Hordijk, Mike Steel, and William F. Martin. Autocatalytic sets in *E. coli* metabolism. *Journal of Systems Chemistry*, 6:4, 2015. [doi:10.1186/s13322-015-0009-7](#).
- [7] Mike Steel, Joana C. Xavier, and Daniel H. Huson. The structure of autocatalytic networks, with application to early biochemistry. *Journal of the Royal Society Interface*, 17:20200488, 2020. [doi:10.1098/rsif.2020.0488](#).
- [8] Joana C. Xavier, Wim Hordijk, Stuart A. Kauffman, Mike Steel, and William F. Martin. Autocatalytic chemical networks at the origin of metabolism. *Proceedings of the Royal Society B*, 287:20192377, 2020. [doi:10.1098/rspb.2019.2377](#).