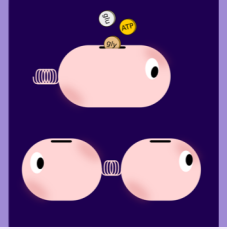
Economic Principles in Cell Biology

Paris, July 8-11, 2024



The Origin of Life

Presented by:
Sanjay Jain

Other authors of chapter:
Christoph Flamm, Ghjuvan M. Grimaud, Shawn E. McGlynn, Marcelo A. Rivas-Astroza

Inria Département MathNum EURIP LEARNING PLANET INSTITUTE

Outline of talk

- The complexity of the simplest life forms
- The simplicity of the prebiotic Earth – hence the puzzle of the origin of life
- Autocatalytic sets – an organizing principle
- Dynamics of autocatalytic sets
- Evolution of autocatalytic sets and protocells
- Other puzzles in the origin of life

Books:

Luisi, P.L., 2016. *The emergence of life: from chemical origins to synthetic biology*. Cambridge University Press.

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

Serra, R. and Villani, M., 2017. *Modelling protocells*. Springer.

Kolb, V.M. ed., 2014. *Astrobiology: an evolutionary approach*. CRC Press.

Kauffman, S.A., 1993. *The Origins of Order*. Oxford University Press

Reviews:

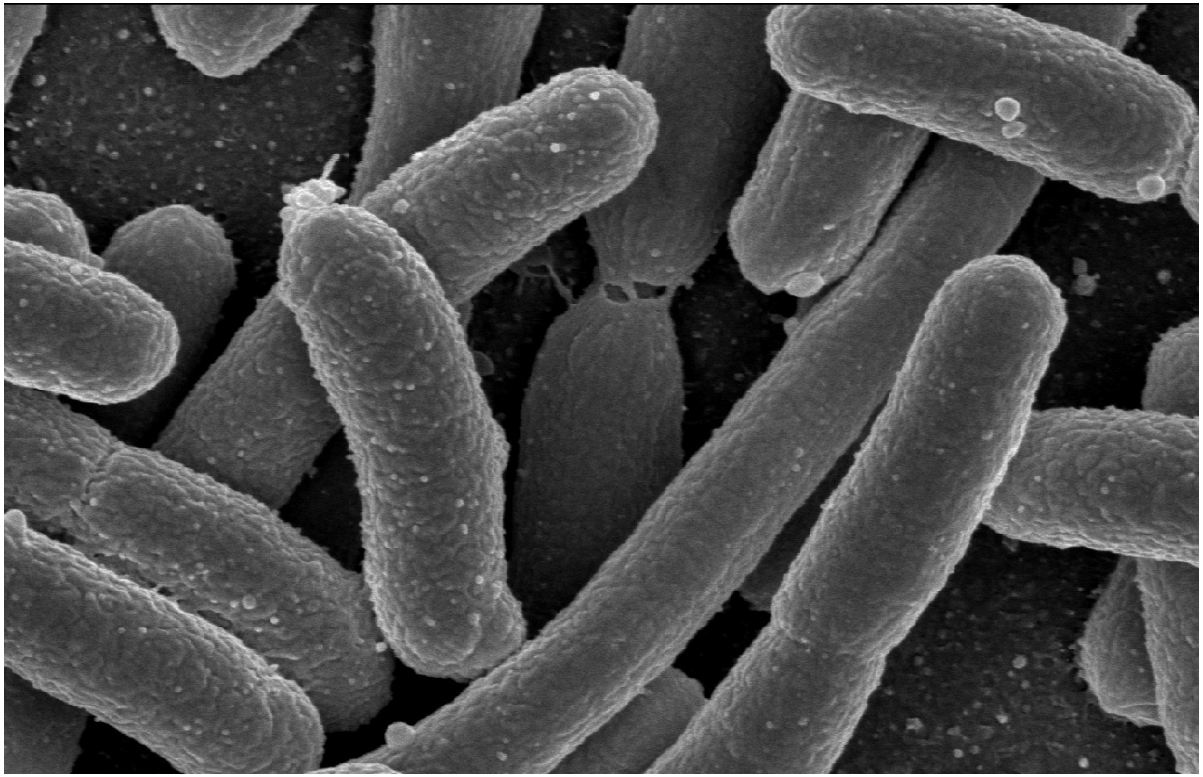
Ameta, S., Matsubara, Y.J., Chakraborty, N., Krishna, S. and Thutupalli, S., 2021. Self-reproduction and Darwinian evolution in autocatalytic chemical reaction systems. *Life*, *11*(4), p.308.

Hordijk, W., 2019. A history of autocatalytic sets. *Biological Theory*, *14*(4), pp.224-246.

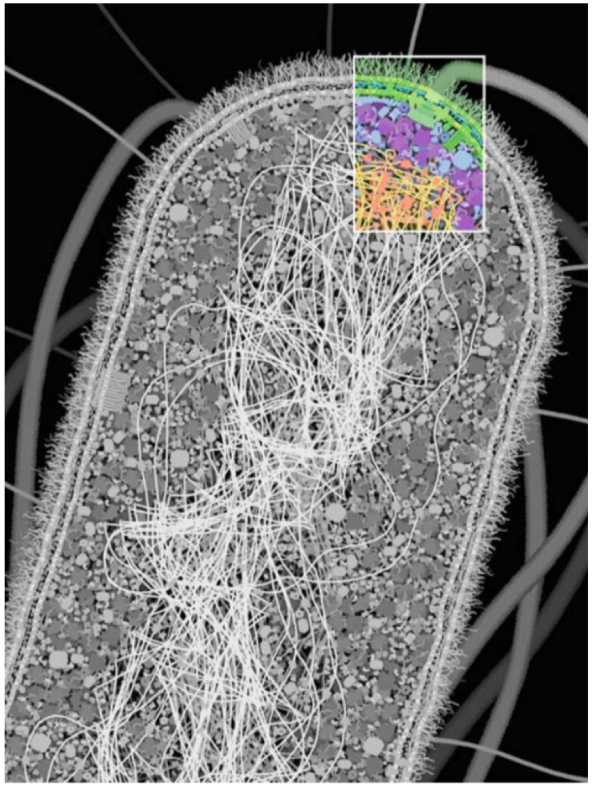
Nghe, P., Hordijk, W., Kauffman, S.A., Walker, S.I., Schmidt, F.J., Kemble, H., Yeates, J.A. and Lehman, N., 2015. Prebiotic network evolution: six key parameters. *Molecular BioSystems*, *11*(12), pp.3206-3217.

Lancet D, Zidovetzki R, Markovitch O. 2018. Systems protobiology: origin of life in lipid catalytic networks. *J. R. Soc. Interface* 15: 20180159.

Bacterial cells - *E. coli*



Source: http://www3.niaid.nih.gov/NR/rdonlyres/49477C30-0513-47BE-88FC-17974CB1F952/0/e_coli.jpg



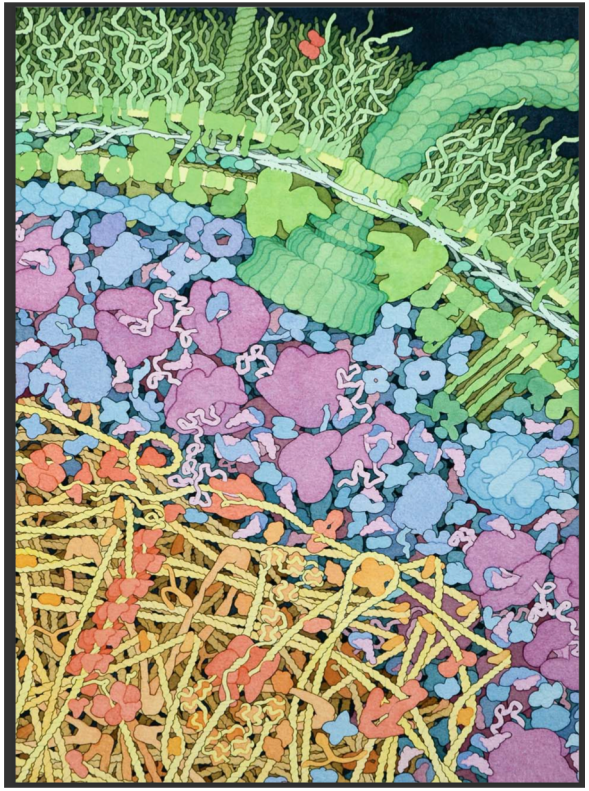
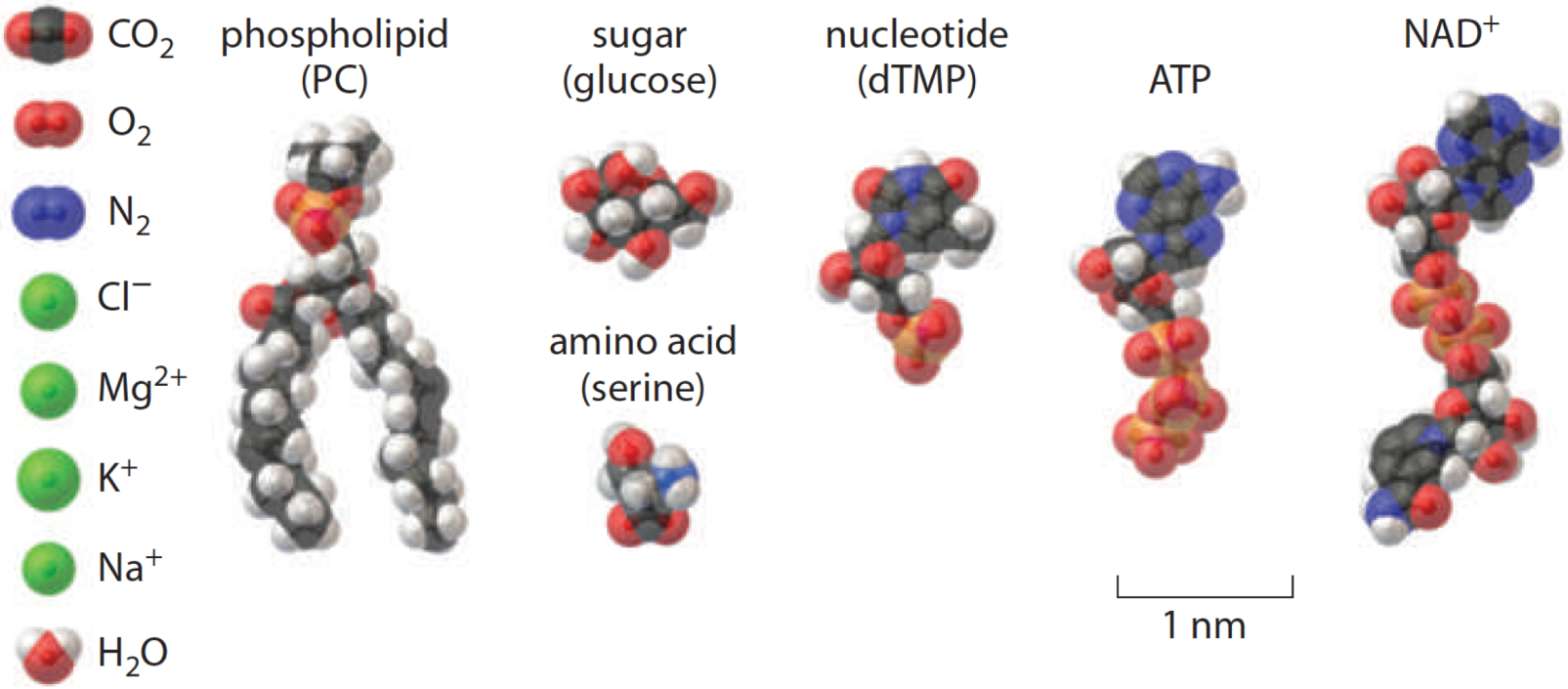
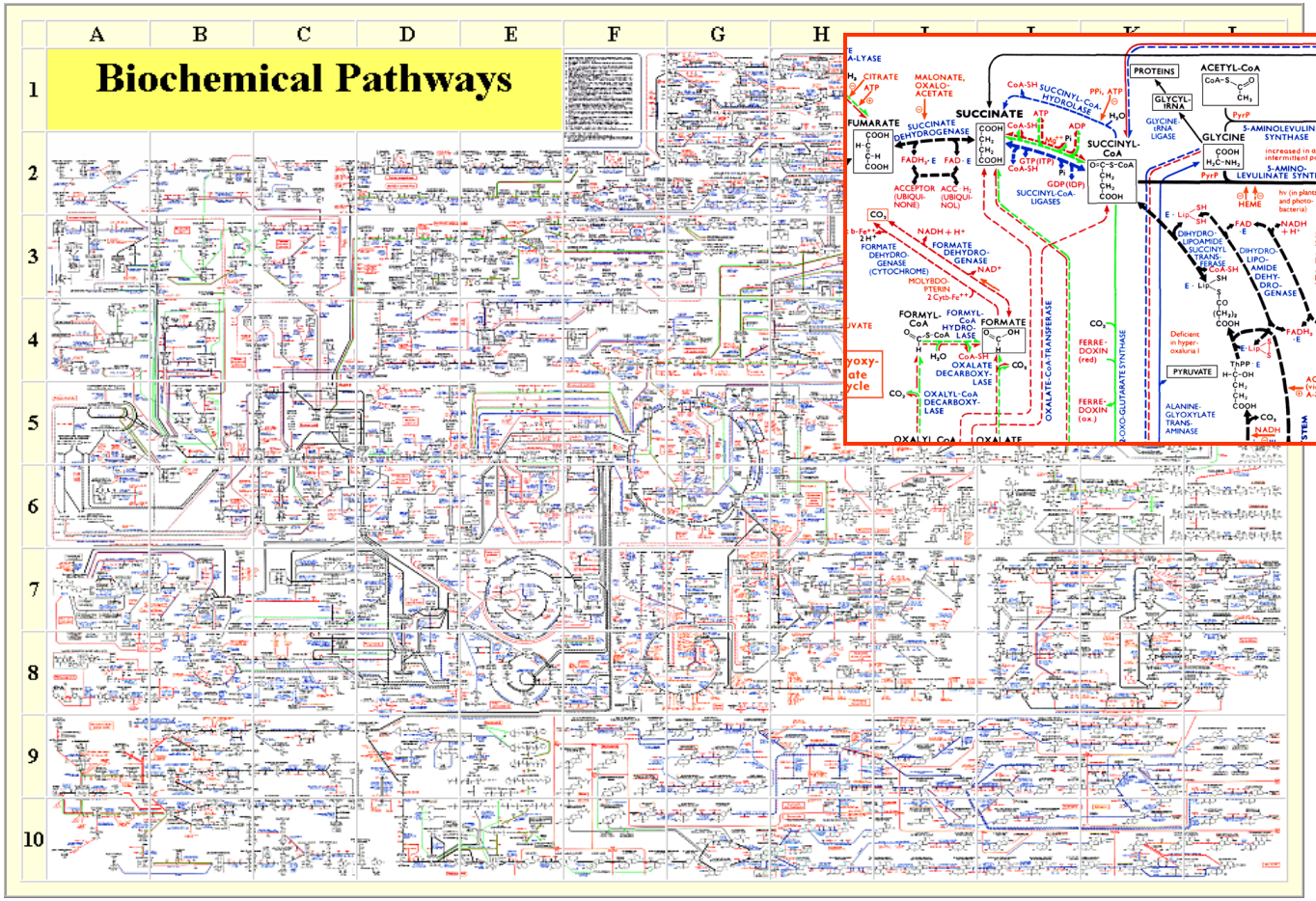


Illustration of Escherichia Coli by Goodsell

The simple molecules of life



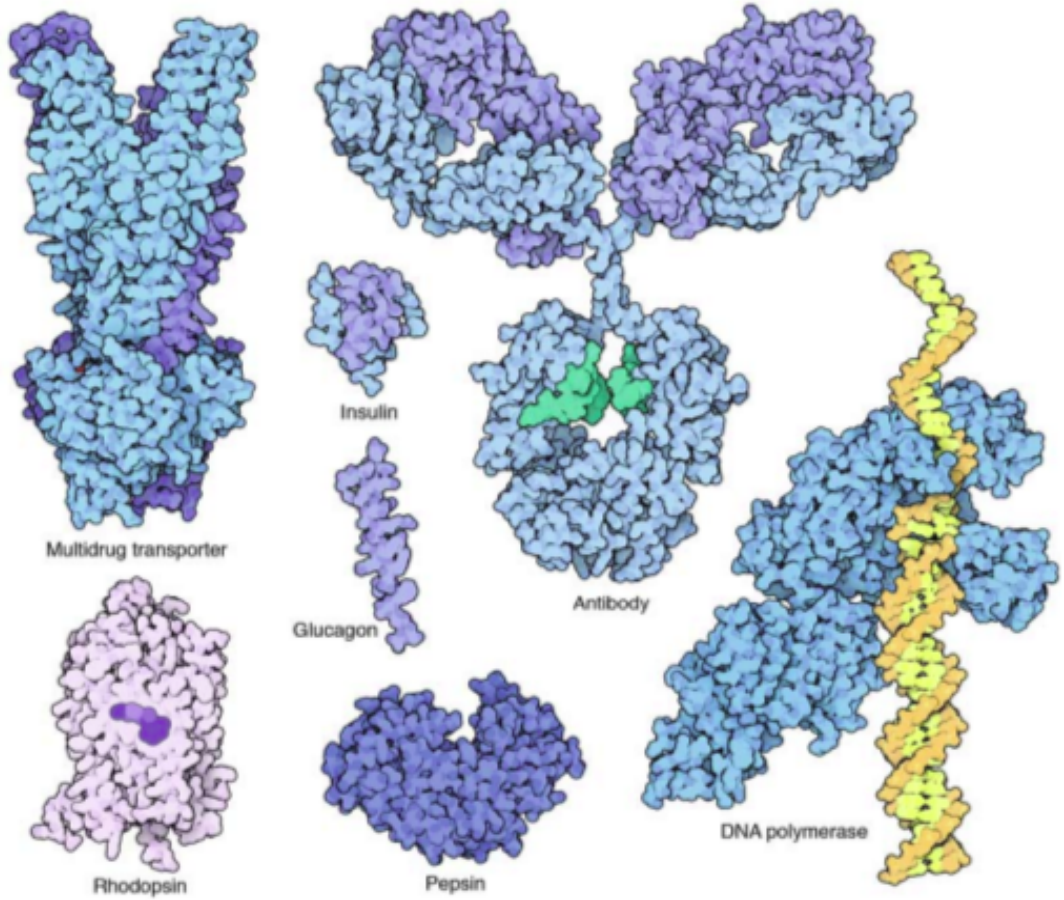
Pictures of molecules from Cell Biology by the Numbers. Authors R. Milo and R. Phillips. Illustrated by N. Orme.



Proteins: Macromolecules with a few hundred to several thousand amino acids.

Produced by a complicated machinery of transcription and translation.

Each playing a specific role in the larger organization



The simplest living cells are very complex
compared to what obtained on the Earth about 4 bya

At the molecular level:

Ingredients (within ~ 1 μm^3): water ~ 10^{10} molecules
small molecules (metabolites) inorganic + organic: ions, sugars, amino acids, nucleobases, lipids, cofactors (~ 1000 species)
macromolecules: protein (~ 500 species, avg length ~ 200 amino acids)
RNA, DNA (~ 1 molecule, ~ 500 genes)
assemblies: cell membrane (made of lipid, transporter protein, receptors)
machines, e.g., Ribosome, RNAP, DNAP (made of protein, RNA)

Molecular Processes: chemical reactions, catalysis, transport by diffusion, lipid assembly, transport by enzymes, signaling, transcription (base pairing), translation (genetic code), DNA replication (with errors), growth, division

The simplest living cells are very complex
compared to what obtained on the Earth about 4 bya

At the level of the whole:

Organization: biochemical networks (metabolic, genetic, signaling)
Role playing: specificity of action, “made for each other” property

System level properties:
(i) The ability, in a suitable environment, to transform raw materials available in the environment into other products needed in the system.
(ii) Reproduction of the whole.
(iii) Capacity to evolve.

NASA definition of life: “Life is a self-sustaining chemical system capable of Darwinian evolution.”

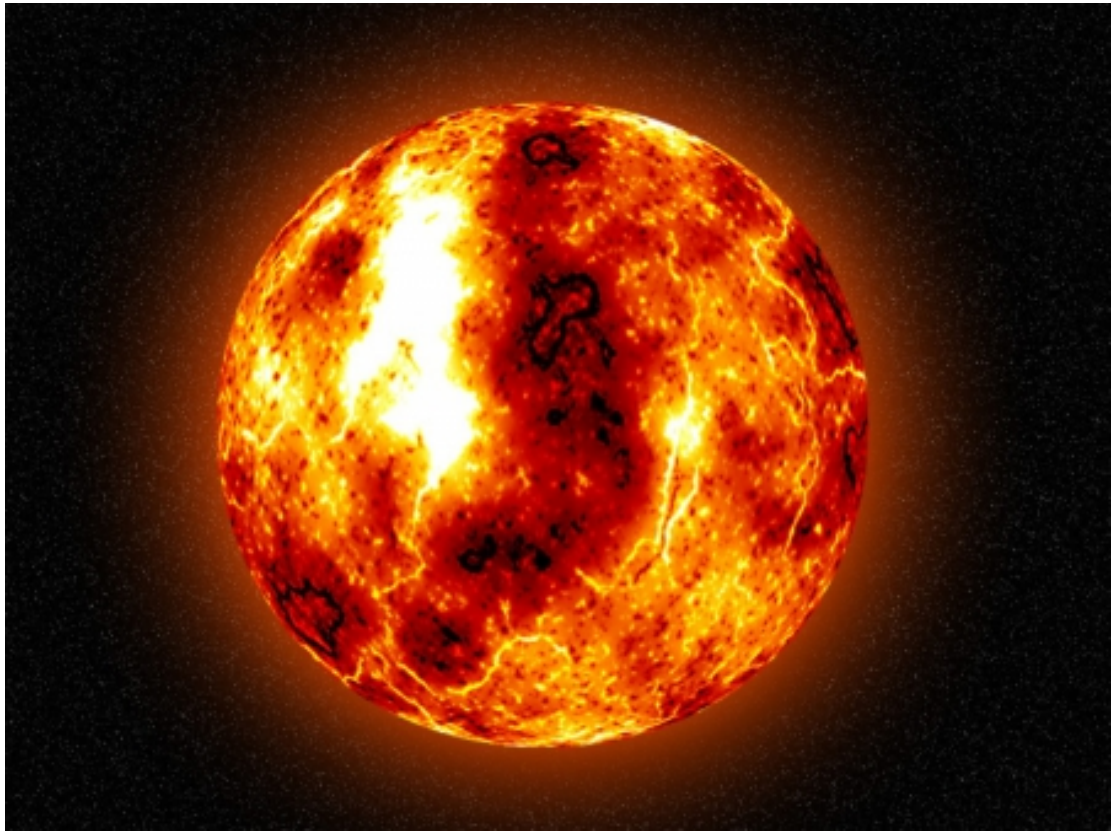
Hence there must have been a period of **chemical evolution** for life to appear. Evolution of chemical complexity.

What came first?

- Organization
- Metabolism
- Self reproduction
- Enclosure
- Evolvability
- Large molecules

There may have been many transitions in the chemical evolution leading to life. However, we do not have evidence of the intermediate stages.

What the early Earth might have looked like – a ball of fire



Source: http://tylkonauka.pl/wiadomosc/poczatki-ziemi

What the prebiotic Earth might have looked like after the oceans condensed

Source: http://tylkonauka.pl/wiadomosc/poczatki-ziemi

What kind of objects and processes existed on the prebiotic Earth?

Chemical molecules undergoing chemical reactions:

$A \rightarrow B$

$A + B \rightarrow C, \qquad C \rightarrow A + B$

Also catalyzed reactions:

$A + C \rightarrow B + C$

$A + B + C \rightarrow P + C$

Catalysis speeds up the rates of reaction and is crucial in cells

Catalytic activity of small molecules

Table 5.1 *Yields of peptides and PNAs, as catalysed by Ser-His and $(Gly)_n$*

Entry #	Catalyst	Conditions	Acyl donor	Free amine	Coupling product	Yield (%)
1	Ser-His	*a*	Ac-Phe-OEt	H-Leu-NH_2	Ac-Phe-Leu-NH_2	15.3
2	Ser-His	*a*	Ac-Phe-OEt	H-Phe-NH_2	Ac-Phe-Phe-NH_2	35.2
3	Ser-His	*a*	Ac-Phe-OEt	H-Leu-Phe-NH_2	Ac-Phe-Leu-Phe-NH_2	25.4
4	Ser-His	*a*	Z-Ala-Phe-OMe	H-Leu-NH_2	Z-Ala-Phe-Leu-NH_2	4.0
5	Ser-His	*a*	Z-Ala-Phe-OMe	H-Phe-NH_2	Z-Ala-Phe-Phe-NH_2	0.5
6	Ser-His	*a*	Z-Ala-Phe-OMe	H-Leu-Phe-NH_2	Z-Ala-Phe-Leu-Phe-NH_2	4.3
7	Ser-His	*b*	PNA monomer	PNA monomer	PNA dimer	4.5[f]
8	Ser-His	*b*	PNA monomer	PNA monomer	PNA trimer	5.6[f]
9	Ser-His	*b*	PNA monomer	PNA monomer	PNA tetramer	9.1[f]
10	Ser-His	*c*	H-Phe-OEt	H-Phe-OEt	H-Phe-Phe-OEt	0.5–3
11	Ser-His	*d*	H-Phe-OEt +H-Trp-OEt	H-Phe-OEt +H-Trp-OEt	H-(Phe)(Trp)-OEt	0.1
12	Gly	*e*	Ac-Phe-OEt	H-Leu-NH_2	Ac-Phe-Leu-NH_2	15.9
13	Gly-Gly	*e*	Ac-Phe-OEt	H-Leu-NH_2	Ac-Phe-Leu-NH_2	15.2
14	Gly-Gly-Gly	*e*	Ac-Phe-OEt	H-Leu-NH_2	Ac-Phe-Leu-NH_2	14.2

Conditions
a: 50 mM reactants (each), 4 mM Ser-His, 24 h, 25 °C;
b: 10 mM PNA monomer; 5.5 mM Ser-His, 35 h, 25 °C;
c: 270 mM H-Phe-OEt; 22 mM Ser-His, 14 days, 4 °C;
d: 50 mM reactants (both), 16 mM Ser-His, 7 days, 60 °C;
e: 50 mM reactants (each), 4 mM Gly (or Gly-Gly; or Gly-Gly-Gly), 15 days, 4 °C.

Note
f: Calculated from the theoretical 100% yield of that product, as the only product in the mixture.

Gorlero et al (2009) FEBS Lett. **583**
Reproduced from Luisi, The Emergence of Life (2016)

Catalysis and Autocatalytic Sets (ACSs)

- Small molecules with catalytic properties are ubiquitous and are readily produced by natural processes on the prebiotic earth

- Consider a class of small molecules which are naturally produced in the prebiotic Earth and are abundant in a particular locale. We will refer to them as the “food set”, and assume that their supply is maintained “for free”.

Autocatalytic set (ACS)

A set of catalyzed reactions with the following properties is referred to as an ACS:

1. The catalyzed reactions form a continuous path from the food set to higher molecules
2. Each catalyst is itself produced in a catalyzed reaction belonging to the set

Eigen, Die Naturwissenschaften (1971); Kauffman, J. Cybernetics (1971); Rossler, Zeitschrift fur Naturforschung B (1971)

Alternative definition of Autocatalytic Set

Let F denote the “food set” (the set of molecules which are naturally produced in the prebiotic Earth and are abundant in a particular locale).

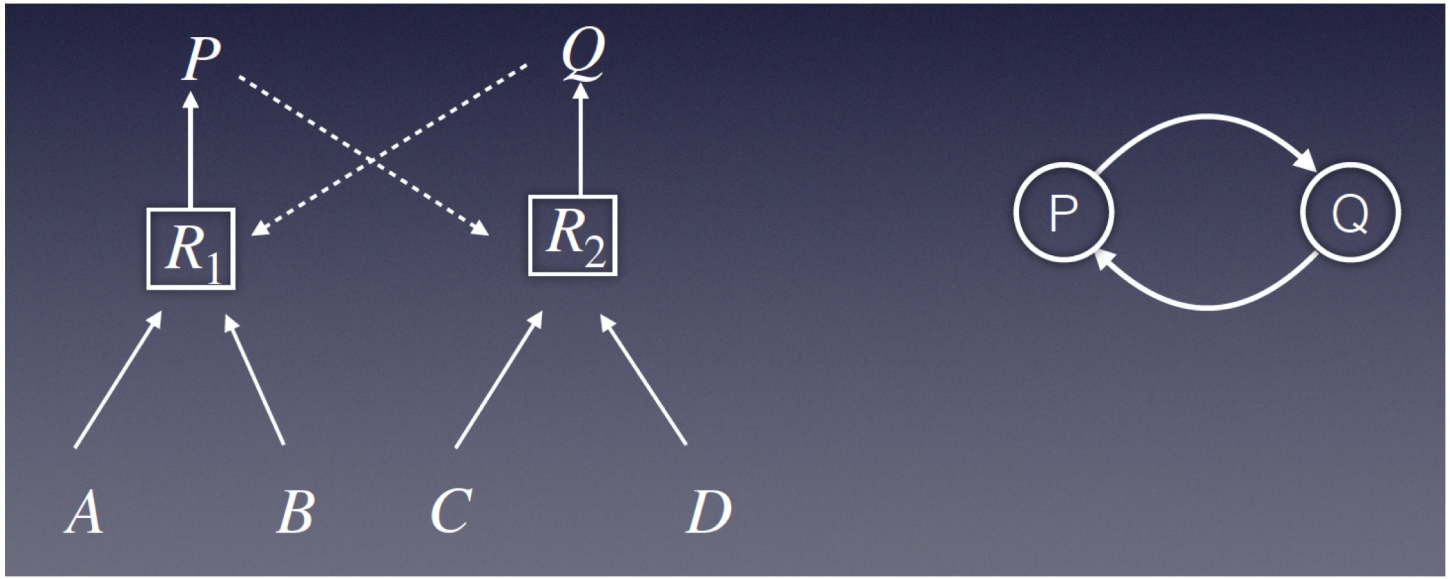
Let S be a set of catalyzed reactions. Every catalyzed reaction has a set of reactants, products, and a catalyst.
Let R, P and C be the unions of reactants, products and catalysts, respectively, of all the reactions in S.

Then S is an ACS if both C and R are subsets of P ∪ F.

In other words, all the catalysts and reactants required to carry out the reactions of S should be either be produced in S itself, or be in the food set.

For a formal definition and an algorithm to find ACSs in a list of reactions, see Hordijk and Steel, J Theor. Biol. (2004)
This is not the only kind of structure that can be called an ACS. For generalizations and a classification, see Blokhuis, Lacoste and Nghe, PNAS (2020)

Example 1



S = {R_1, R_2}

F = {A,B,C,D}, R = {A,B,C,D}, P = {P,Q}, C = {P,Q}

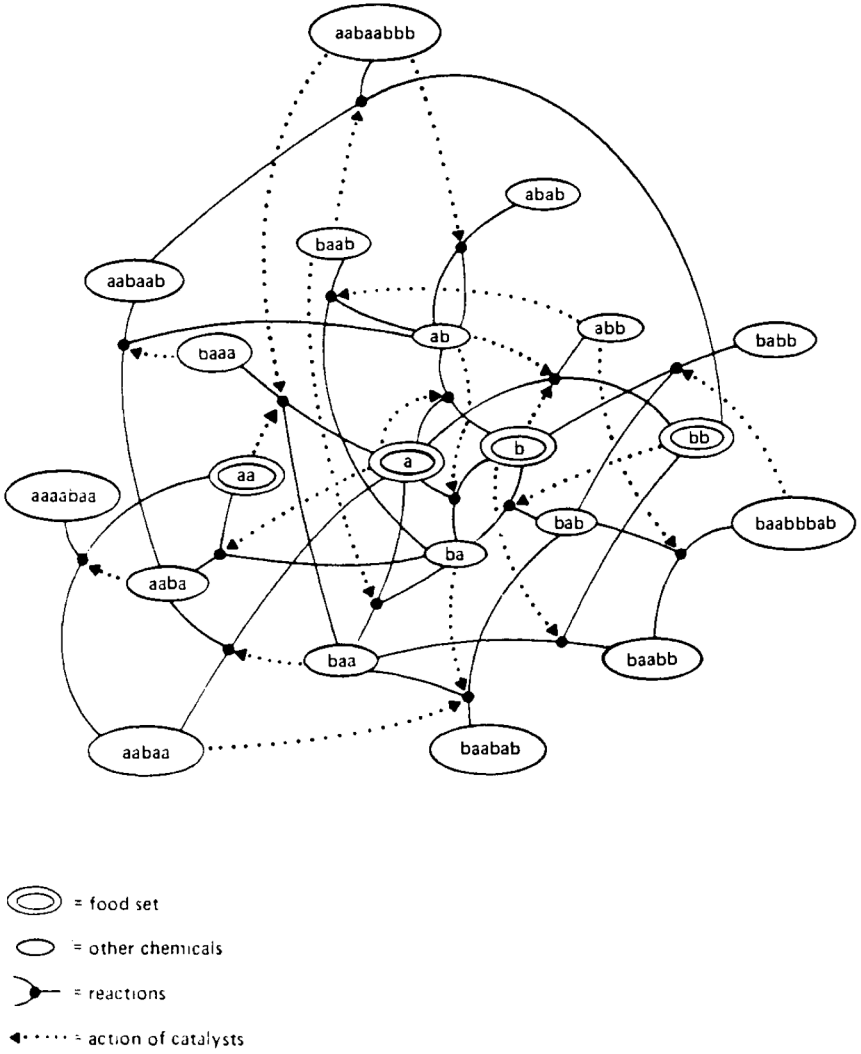
P ∪ F = {A,B,C,D,P,Q}

Catalytic closure;
Positive feedback

Example 2

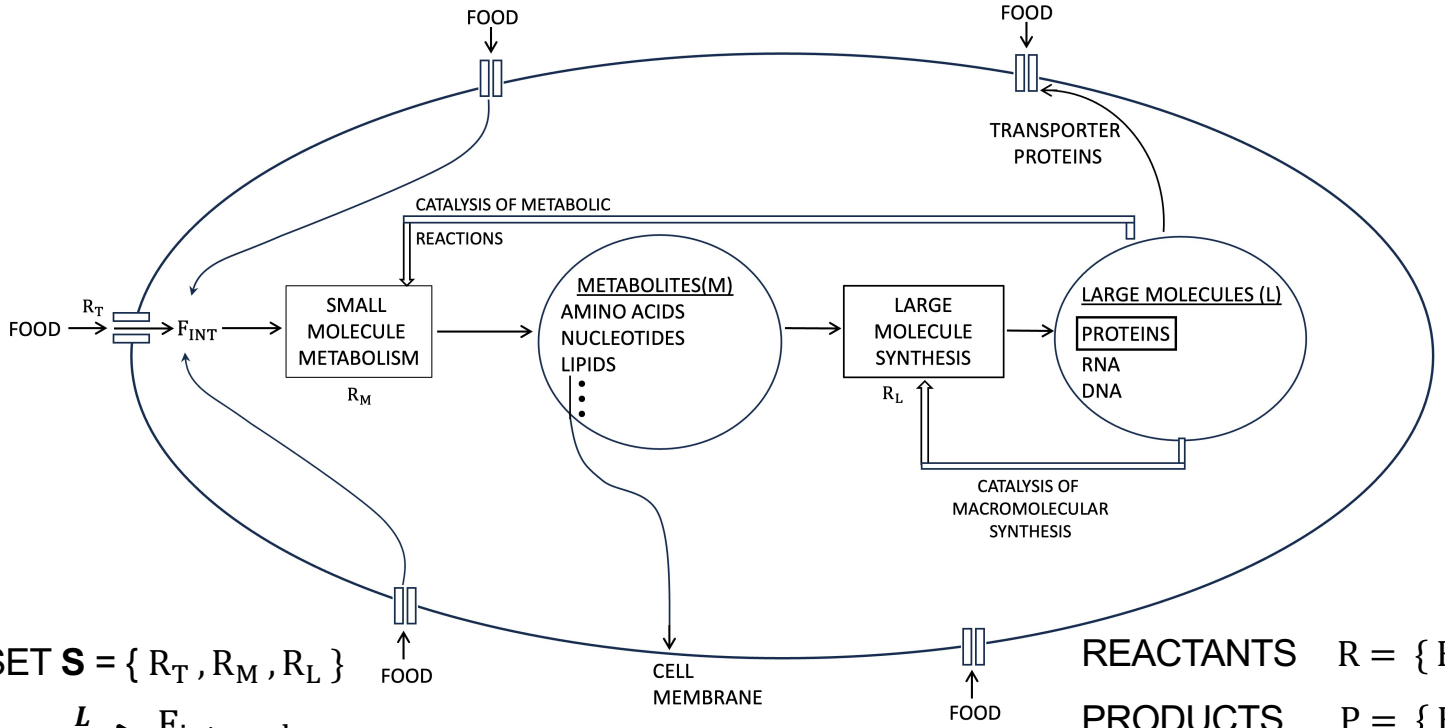
Artificial chemistry:

- All molecules are strings of two monomers a and b
- Reactions are ligations and cleavage of the strings
- Some molecules are catalysts of some reactions



Farmer, Kauffman, Packard (1986) Physica D

Example 3: CELL AS AN AUTOCATALYTIC SET



REACTION SET **S** = { R_T , R_M , R_L }

R_T : F $\xrightarrow{L}$ $F_{internal}$

R_M : $F_{internal}$ $\xrightarrow{L}$ M

R_L. : M $\xrightarrow{L}$ L

REACTANTS $R = \{ F, F_{int}, M \}$

PRODUCTS $P = \{ F_{int}, M, L \}$

CATALYSTS $C = \{ L \}$

$P \cup F = \{ F, F_{int}, M, L \}$

Since R and C are subsets of P ∪ F,

S is an ACS

What came first?

- Organization
- Metabolism
- Self reproduction
- Enclosure
- Evolvability
- Large molecules

The idea of autocatalytic sets naturally embodies the first three. Autocatalytic sets might involve lipid molecules that naturally form enclosures (micelles and vesicles). Theoretical models suggest that autocatalytic sets could be evolvable. Their evolution might be aided by the formation of enclosures. They could also produce large molecules.

Some questions about ACSs

Consider the entire reaction space spanned by organic chemistry.

Clearly, it has subsets of reactions that are autocatalytic (e.g., the set of all reactions in any living cell).

Question 1: Does it also have simpler ACSs in the part of chemical space proximate to what existed on the prebiotic Earth? (Amino acids, their dimers, trimers, other small molecules, minerals)

Question 2: If so, might those have been the primitive chemical organizations that first arose on the Earth?

Question 3: In a large chemistry of catalysed and uncatalyzed reactions, do ACSs stand out? Do the products of the ACS have significantly larger populations than the other molecules? (This is a question that goes from structure (network topology of ACS) to its dynamics (chemical kinetics, population dynamics).

Question 4: Can ACSs evolve to become more complex? Can we imagine evolutionary paths leading to complex ACSs such as those that exist in a living cell?

Q1 is largely an experimental question. Q2-4 can be investigated theoretically for artificial chemistries assuming they have ACSs.

Topology and dynamics of autocatalytic sets

Kauffman, S. A. Autocatalytic sets of proteins, 1986. *J Theor. Biol*, 119(1), 1-24 (1986)

Bagley, R.J. and Farmer, J.D., 1991. Spontaneous emergence of a metabolism. *Artificial life II,* C. G. Langton et al (eds), pp 93-140. Addison Wesley.

Stadler PF, Fontana W, Miller JH, 1993. Random catalytic reaction networks. *Physica* D 63:378–392

Segre D, Lancet D, Kedem O, Pilpel Y, 1998. Graded autocatalysis replication domain (GARD): Kinetic analysis of self-replication in mutually catalytic sets. *Orig Life Evol Biosph* 28:501–514

Ohtsuki, H and Nowak M A, 2009. Prelife catalysts and replicators *Proc. R. Soc.* B.2763783–3790

Giri V and Jain S, 2012. The origin of large molecules in primordial autocatalytic reaction networks, *PLoS One* 7(1): e29546

Matsubara YJ and Kaneko K, 2016. Phys. Rev. E 93:032503

Peng Z, Plum AM, Gagrani P and Baum, DA, 2020. An ecological framework for the analysis of prebiotic chemical networks. *J. Theor. Biol.*, 570:110451

Qualitative summary of ACS dynamics

Answer to Q3:

ACS products do stand out above the background, but that depends on many details:

Non-equilibrium conditions maintained by an influx of food molecules is a must.

ACS dominance depends upon on kinetic rate constants (e.g., sufficiently large catalytic efficiencies, magnitudes of forward and backward reaction rates, dissipation rates).

Depends upon details of ACS network topology (e.g., where the catalysts are located in the reaction network).

Starting from small food molecules, it is difficult to produce large molecules in significant quantities even if they are products of an ACS. A `nested ACS' structure, in which ACSs with small catalyst molecules are embedded inside ACSs with larger catalysts can produce large molecules.

Evolution of autocatalytic sets

Bagley, R.J., Farmer, J.D. and Fontana,W., Evolution of a metabolism. *Artificial life II*, C.G. Langton et al (eds), pp.141-158. Addison Wesley (1991)

Jain S and Krishna S, Autocatalytic sets and the growth of complexity in an evolutionary model. Phys. Rev. Lett. 81, 5684-5687 (1998)

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

Markovitch O, Lancet D, Excess Mutual Catalysis is Required for Effective Evolvability, *Artificial Life* 18: 243–266 (2012)

Hordijk, W., Steel, M. Conditions for Evolvability of Autocatalytic Sets: A Formal Example and Analysis. *Orig Life Evol Biosph* **44,** 111–124 (2014)

Villani M et al, Growth and division in a dynamic protocell model, Life 4:837-864 (2014)

Kahana A, Segov L and Lancet D. Attractor dynamics drives self-reproduction in protobiological catalytic networks, Cell Reports Physical Science 4:101384 (2023)

Autocatalytic sets and enclosures. Protocells.

T. Ganti, Chemoton theory. Theoretical foundations of fluid machineries, theory of living systems, vol. 1, 2 (2003)

Serra R, Villani M, Sustainable growth and synchronization in protocell models. Life 9(3):68 (2019)

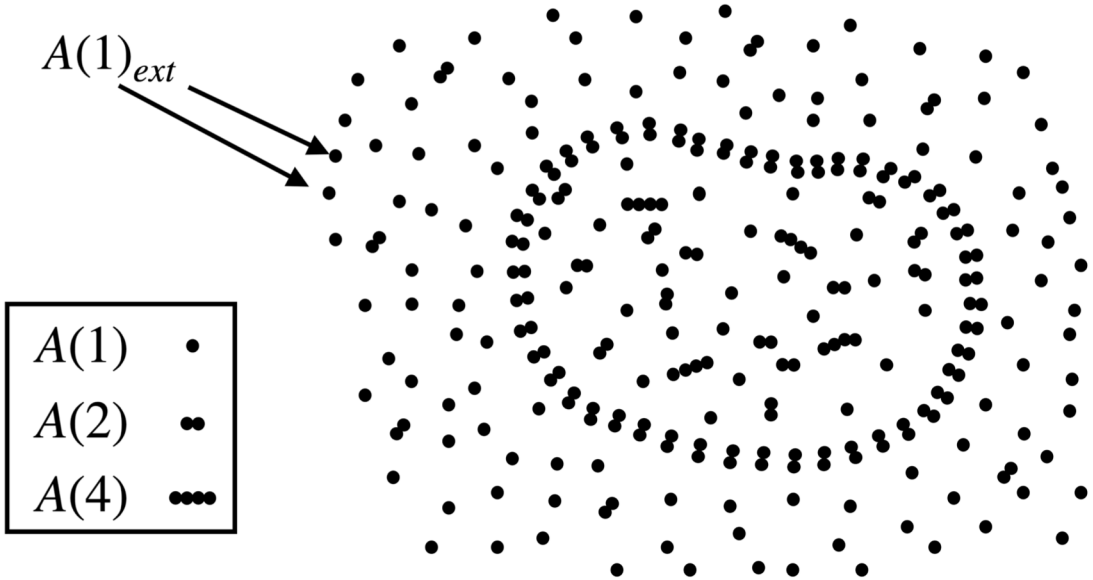
Serra R, Filisetti A, Villani M, Graudenzi A, Damiani C, Panini T, A stochastic model of catalytic reaction networks in protocells, Nat Comput 13(3):367–377 (2014)

Mavelli F and Ruiz-Mirazo K, Theoretical conditions for the stationary reproduction of model protocells, Integr Biol, vol. 5, no. 2, pp. 324–341 (2013)

Kamimura A and Kaneko K, Reproduction of a Protocell by Replication of Minority Molecule in Catalytic Reaction Network. Phys. Rev. Lett. 105: 268103 (2010)

Rasmussen S et al, Bridging nonliving and living matter, Artificial Life 9: 269-316 (2003)

Evolution of an autocatalytic set in a protocell (where the enclosure forming molecule is part of ACS)



Singh, AY and Jain S (2023) *Life* 13, 2327

Reactions

Transport : $A(1)_{ext} \xrightarrow{\alpha X_2} A(1)$

R1 (uncatalyzed) : $2A(1) \underset{k_R}{\overset{k_F}{\rightleftharpoons}} A(2)$

R1 (catalyzed) : $2A(1) + A(4) \underset{\kappa k_R}{\overset{\kappa k_F}{\rightleftharpoons}} A(2) + A(4)$

R2 (uncatalyzed) : $2A(2) \underset{k_R}{\overset{k_F}{\rightleftharpoons}} A(4)$

R2 (catalyzed) : $2A(2) + A(4) \underset{\kappa k_R}{\overset{\kappa k_F}{\rightleftharpoons}} A(4) + A(4)$

Degradation : $A(2) \xrightarrow{\phi} \emptyset, \;\; A(4) \xrightarrow{\phi} \emptyset.$

S = {Transport, R1(catalyzed), R2 (catalyzed)} is an ACS

F = {$A(1)_{ext}$}, R = {$A(1)_{ext}$, A(1), A(2), A(4)}, P = {A(1), A(2), A(4)}, C = {A(2), A(4)}

Equations defining the model

$$\frac{dx_1}{dt} = \alpha x_2 - 2(k'_F x_1^2 - k'_R x_2) - \frac{\dot{V}}{V} x_1,$$

$$\frac{dx_2}{dt} = k'_F x_1^2 - k'_R x_2 - 2(k'_F x_2^2 - k'_R x_4) - (\phi + \frac{\dot{V}}{V}) x_2,$$

$$\frac{dx_4}{dt} = (k'_F x_2^2 - k'_R x_4) - (\phi + \frac{\dot{V}}{V}) x_4,$$

$$k'_F \equiv k_F(1 + \kappa x_4),\ k'_R \equiv k_R(1 + \kappa x_4).$$

$$V(X) = v(X_1 + 2X_2 + 4X_4)$$

$$\mu \equiv \frac{\dot{V}}{V} = \alpha x_2 - \phi(2x_2 + 4x_4)$$

X_i = Population of A(i)

$x_i \equiv X_i / V$

V = Volume of protocell

Rate equations in terms of population variables

$$\frac{dX_1}{dt} = \alpha X_2 - 2\left(\frac{k_F X_1^2}{V} - k_R X_2\right)\left(1 + \kappa \frac{X_4}{V}\right),$$

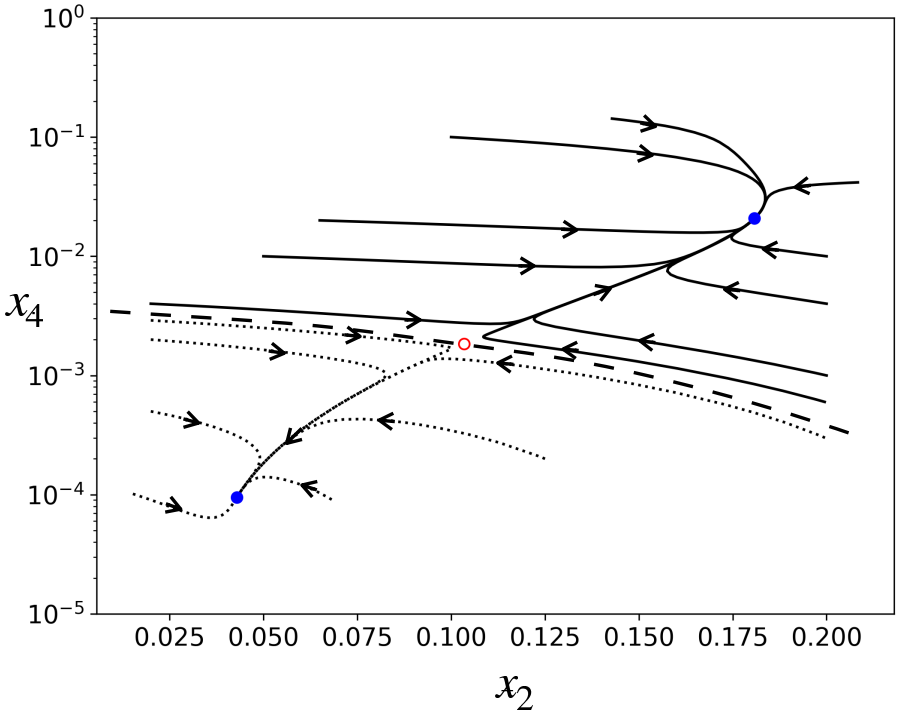
$$\frac{dX_2}{dt} = \left(\frac{k_F X_1^2}{V} - k_R X_2\right)\left(1 + \kappa \frac{X_4}{V}\right) - 2\left(\frac{k_F X_2^2}{V} - k_R X_4\right)\left(1 + \kappa \frac{X_4}{V}\right) - \phi X_2,$$

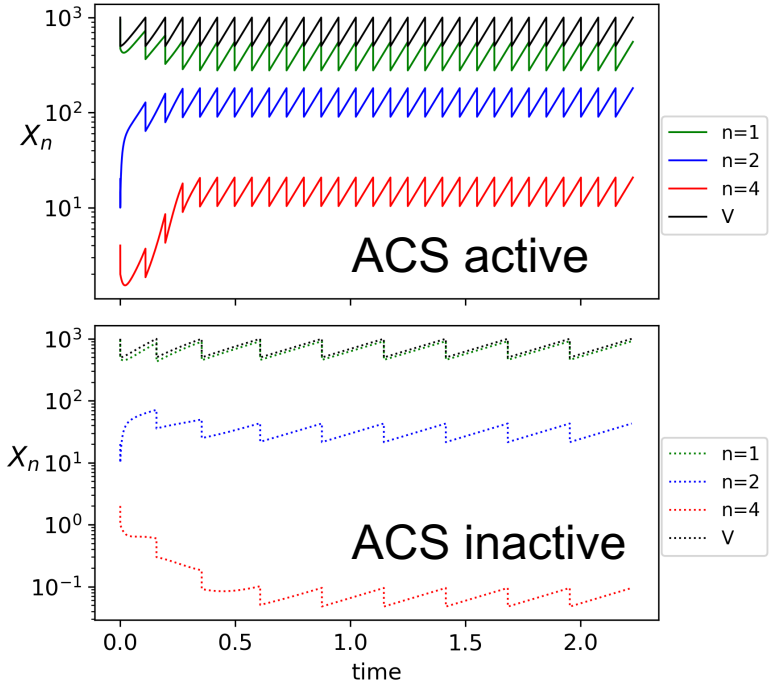
$$\frac{dX_4}{dt} = \left(\frac{k_F X_2^2}{V} - k_R X_4\right)\left(1 + \kappa \frac{X_4}{V}\right) - \phi X_4.$$

$$V(X) = v(X_1 + 2X_2 + 4X_4)$$

Protocell division: If V increases to a critical value V_c, the protocell divides into two identical daughters: $X_i \longrightarrow X_i / 2$

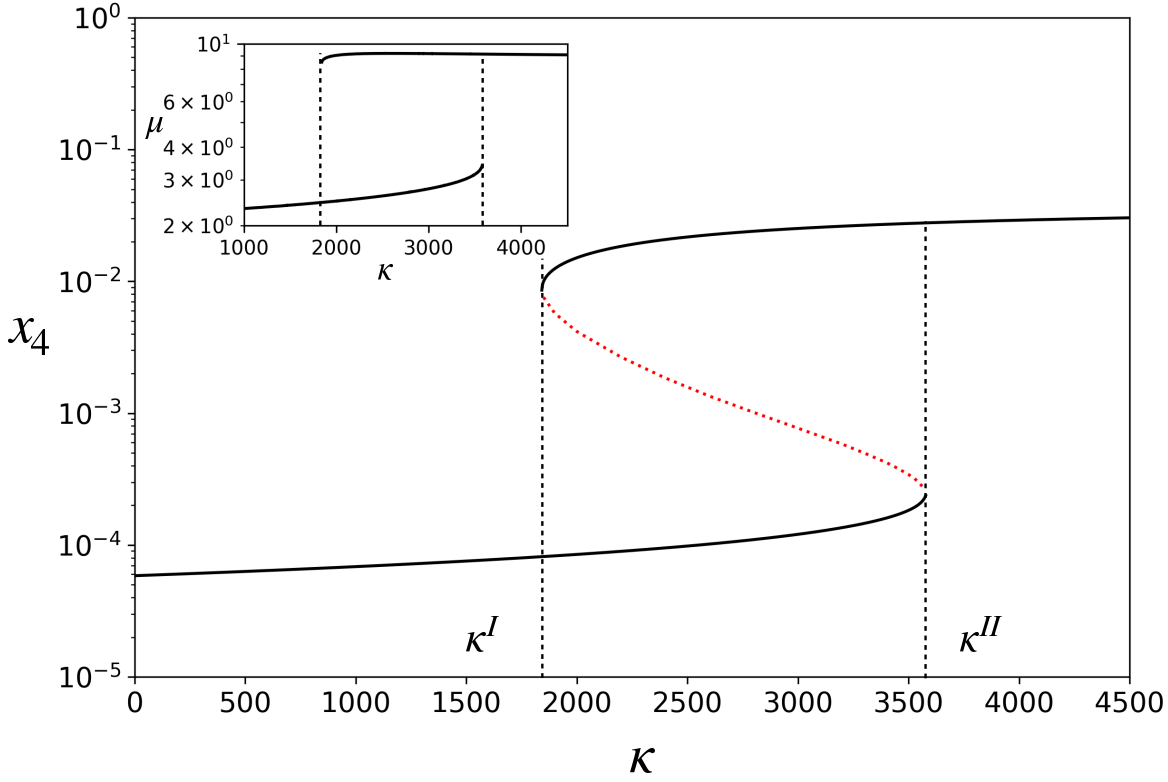
Trajectories



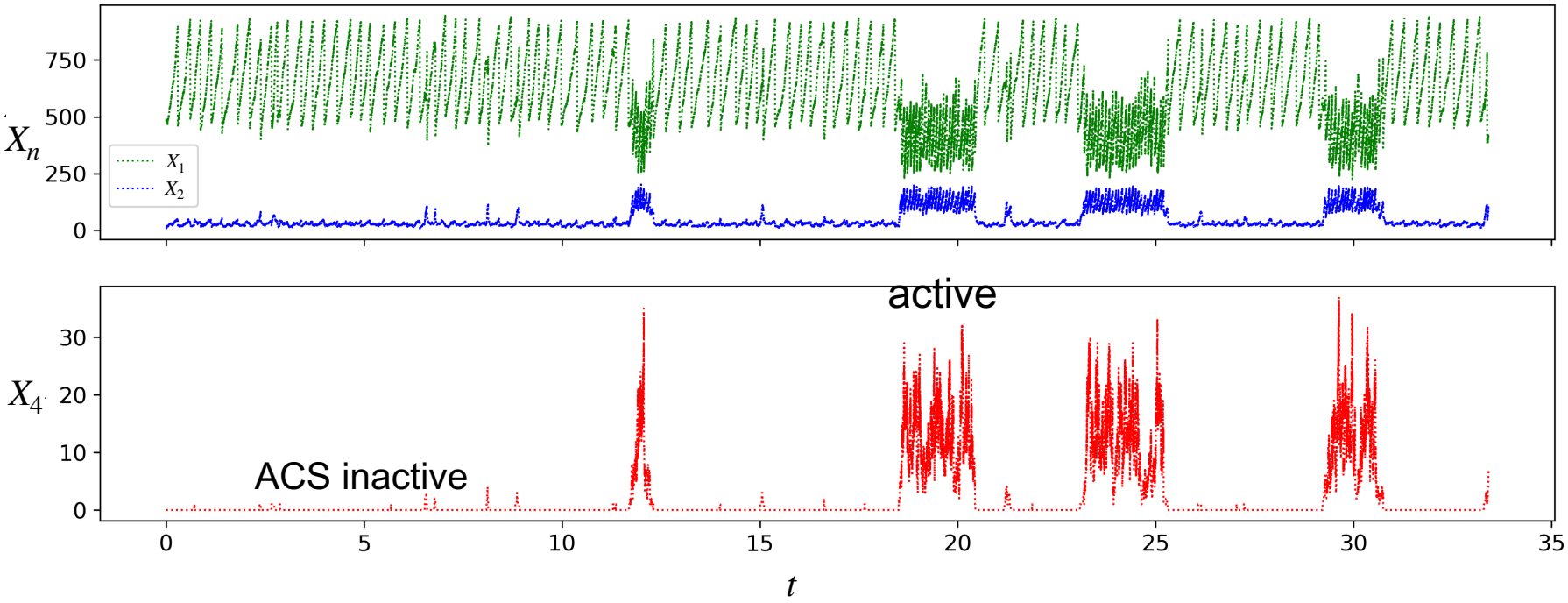


$$k_R = 1,\ \ v = 1,\ k_F = 1,\ \phi = 20,\ \alpha = 100 \quad \kappa = 2400,\ V_c = 1000$$

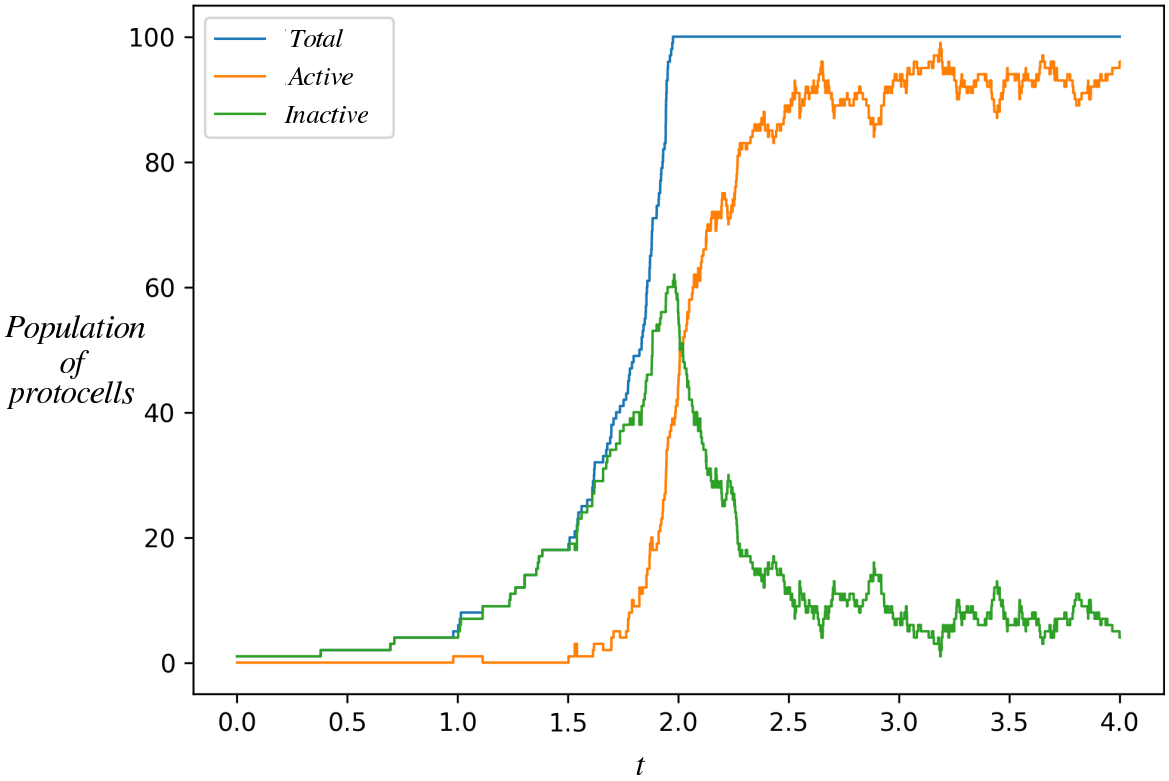
Bifurcation diagram: Bistability in the dynamics



Stochastic dynamics of a single protocell



Evolution of the protocell population: “Natural selection” of the ACS active state



Autocatalytic sets in the lab

S. Ameta, Y. Matsubara, N. Chakraborty, S. Krishna, S. Thuttupalli (2021), Life

Table 1. Various autocatalytic chemistries and their evolutionary properties.

Chemical System	Reproducing Unit	Networks	Boundary Conditions	Variation	Heredity
Lipid-based [94,102]	Chemical composition [1]	N/A	CSTR	Food set-generated [2]	Protocol-mediated [3]
DNA-based [103,105]	Oligonucleotides	1 network, 2 nodes [105]	Equilibrium	Reaction kinetics [4]	N/A
Inorganic-based [113]	Molybdenum clusters	N/A	Stopped-flow	N/A	N/A
Sugar-based [115,116]	Sugars (C2–C5)	N/A	CSTR [116]	Reaction kinetics [4]	N/A
Peptide-based [127,134,138]	Peptides	1 network, >20 nodes [134]	Both [5]	Reaction kinetics [4]	Concentration-mediated [6]
Macrocycle-based [148]	Macrocycle assemblies	N/A	DCL	Reaction kinetics [4]	N/A
RNA-based [155,156,159]	Chemical composition [1]	>20,000 networks [160], >40 nodes [53]	Both [5]	Reaction kinetics [4]	Differential seeding [7]

[1] Chemical composition: abundance (or number) of product molecules in micelles, vesicles [2], or networks of RNA [160]. [2] Food set-generated: variation arises due to the composition of food set flowed under continuous stirred-tank reactor (CSTR) conditions [102]. [3] Protocol-mediated: differential compositional states are obtained due to different experimental protocols, namely equilibrium, phase separation, and CSTR conditions [102]. [4] Reaction kinetics: variation owing to a combination of background reactions and boundary conditions. [5] Both: equilibrium [129,160], and out-of-equilibrium conditions [53,145]. [6] Concentration-mediated: bistable steady states are obtained by starting with different concentrations of reactants and products [129] or by changing the concentration via a chemical fuel [145]. [7] Differential seeding: multiple steady states are obtained by controlled seeding with different RNA catalysts [160].

Lu et al (2023) Nat. Chem.: Small-molecule autocatalysis drives compartment growth, competition and reproduction

Other puzzles in the origin of life

- Origin of enclosure (cell membrane)
- Emergence of macro-molecules (proteins, RNA, DNA)
- Separation of roles (RNA, DNA – Information carriers, Proteins – Functional agents)
- Emergence of the genetic code
- Origin of chirality