# The emergence of a fourth geosphere

Morowitz, Copley, Smith (III)

Complex Systems Summer School, Santa Fe, 2009

# Recap: Two things we have seen so far

- Commitment to universals as the guides to laws
    - Harold: core metabolism is the oldest fossil on earth
- Enriching thermodynamic thinking to pre-biological form / function pairs
    - Shelley: catalysis is the(?) key concept to organize kinetics

# Step back:
# What do we want from a theory of origins?

- Correct statements about historical events
- Ability to handle counterfactuals (can’t happen or can happen elsewhere)
- Insights into organizing principles that apply more generally

# The labyrinth of definitions: What is life? What was the origin?

- Distinctive materials: chemical / organic / DNA-RNA / . . .
- Distinctive processes: replication & evolution / development & regulation / niche construction / succession . . .
- New organizational motifs: individuality , generationality, contingency . . .
- The transition: one event or many? continuous or heterogeneous? accident or necessity?

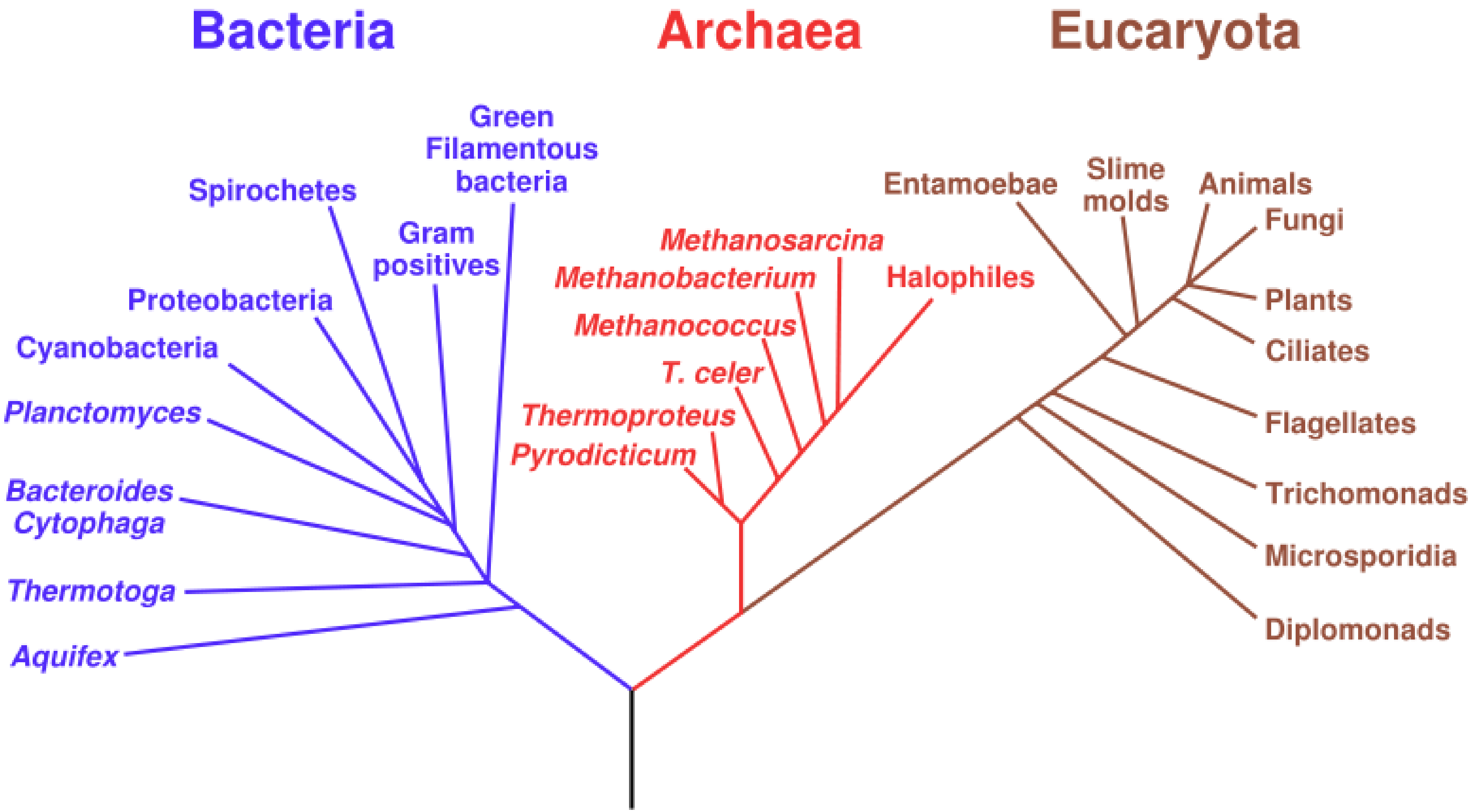
The problem:
choosing what is conceptually central, and what is derived

# Methodological commitments: Respecting physics, chemistry, biology

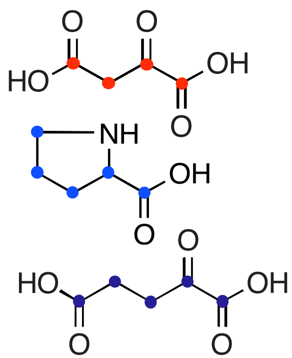
- Chance grew out of necessity in stages
- Universals are clues to the necessities
- Surprising stability provides a way out of the labyrinth

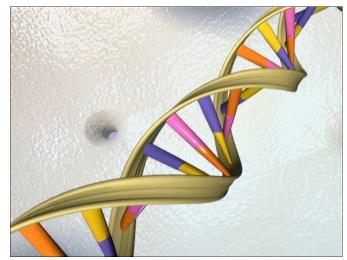
# Chance and necessity:  I know it when I see it . . .

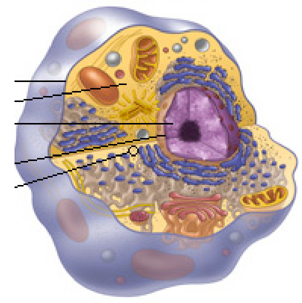
# Evolution has always been about jointly inferring history, process, and causation



# Push a bit more on universality: signatures of chance and necessity





**Necessity**

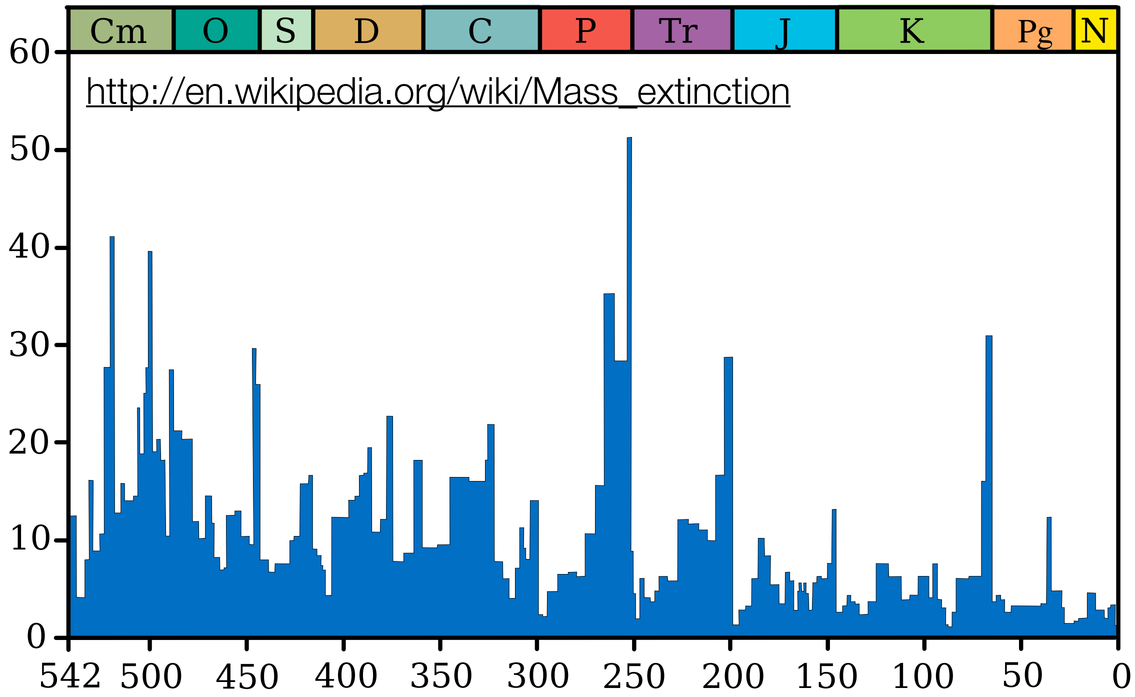
Universal
Steady
Predictable

Ecological order is the natural bridge between geochemistry and life

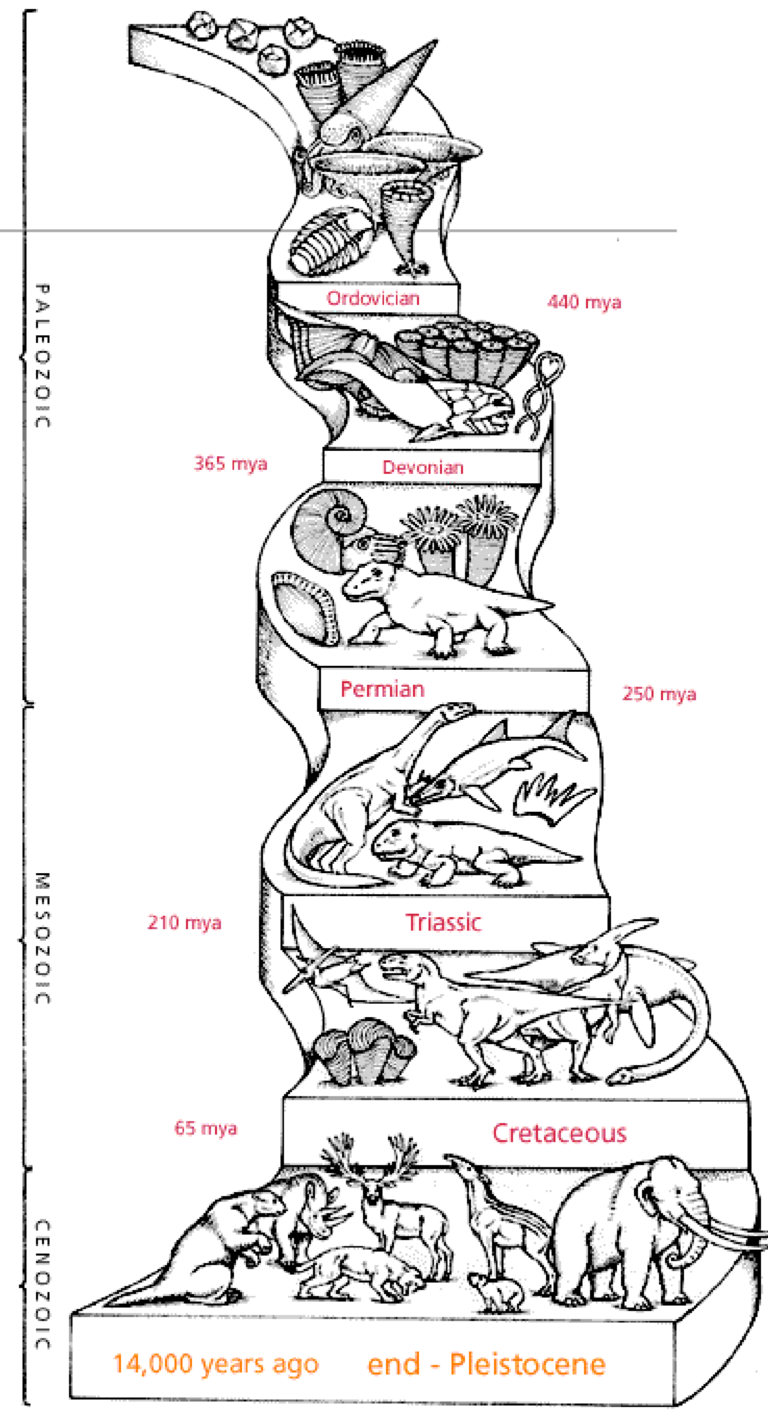
**Chance**

Variable
Fluctuating
Contingent

# Species go extinct; ecosystems restructure

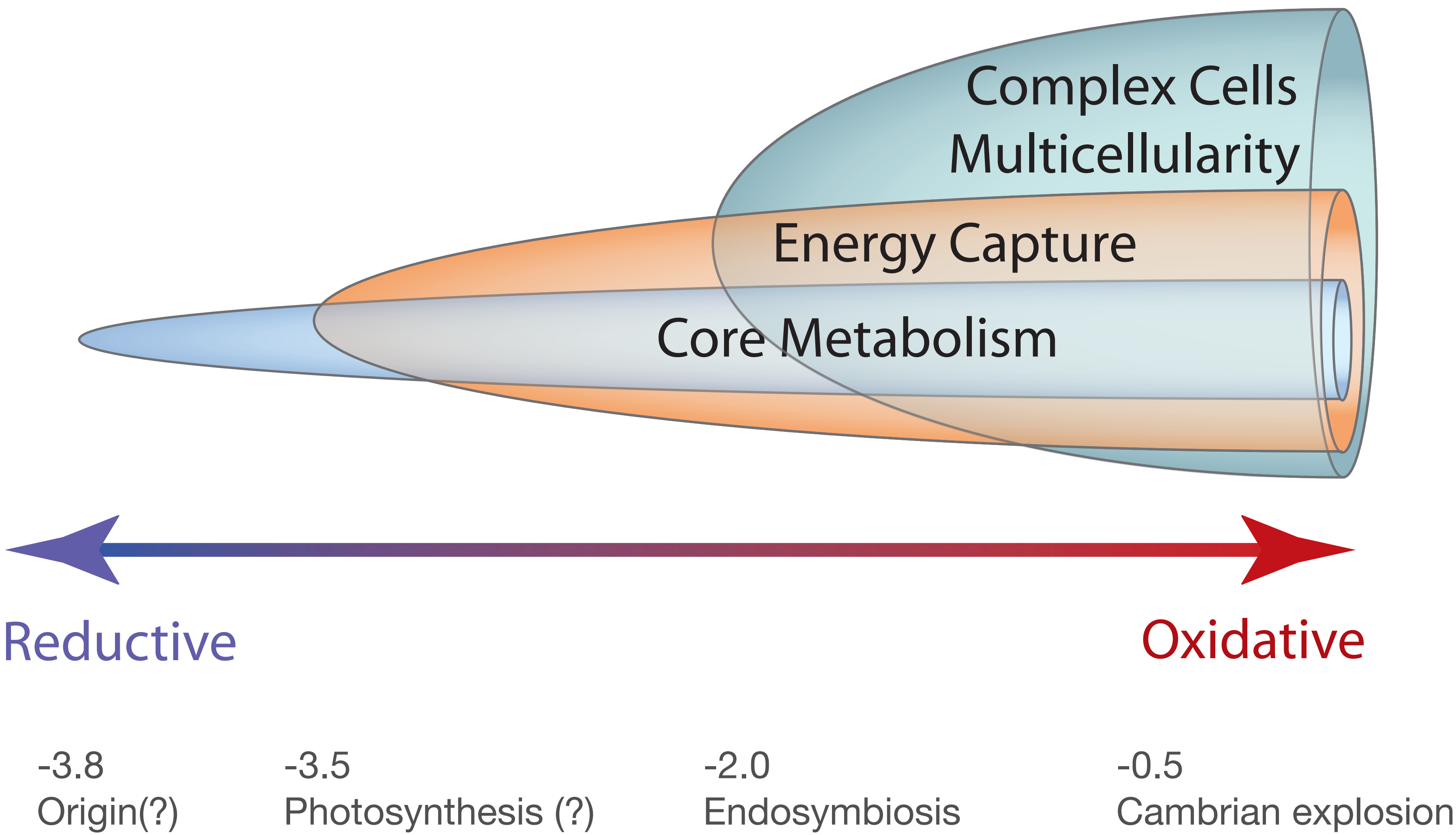



- Extinctions at many levels have happened continually
- Yet core biochemistry has persisted with (we think) little loss, and only occasional major innovations




http://www.amnh.org/science/biodiversity/extinction/Intro/OngoingProcess.html

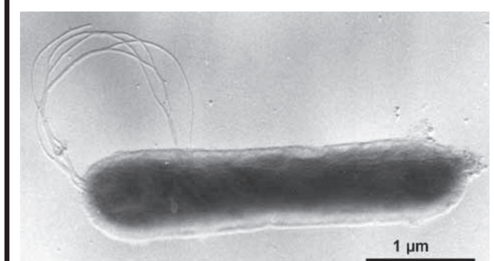
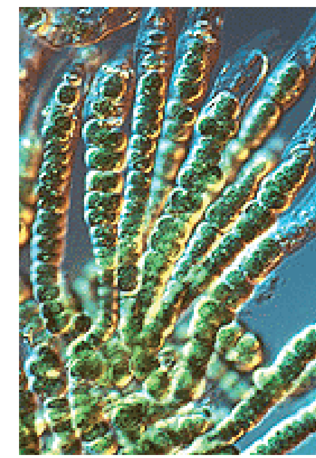
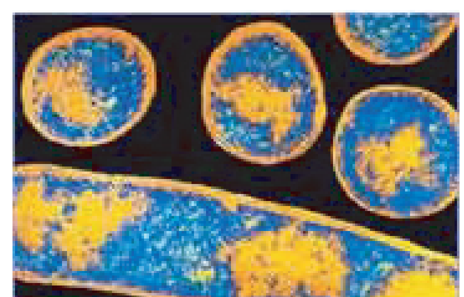
# Chemical versus genotypic major transitions



# Universality does not always respect structures

# A taxonomy based on energetics and ecology

- Physiology is broadly constrained according to directions of electron flow
- Ecological trophic interactions isolate “packaging”, specialization, and interfaces

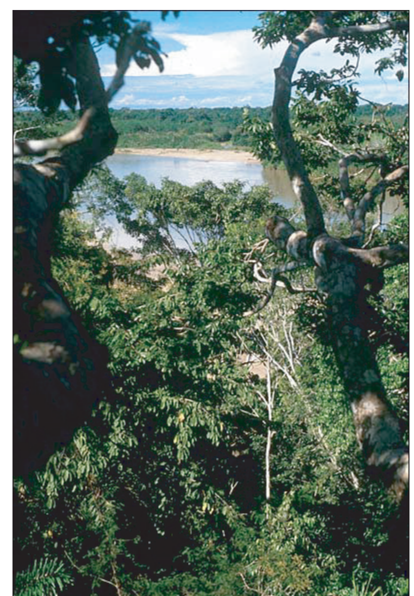
| | Reductive metabolisms | Oxidative metabolisms |
|---|---|---|
| Autotrophs |   |  |
| Heterotrophs |  |  |

# The ecosystem taxonomy: simpler and more *universal*

- All reductive ecosystems are more or less alike (--Tolstoy?)
- Oxidative ecosystems “contain” reductive ecosystems within “energy suits”

Reductive ecologies

Oxidative ecologies



# But what about . . .?

- The cell theory? Viruses, endosymbionts, etc. complicate this story
- Genes and common descent? Less universal than coded reactions
- DNA? Derived from RNA, perhaps twice independently (Koonin and Martin)
- RNA? Easier with metabolism first (our view)

# Chemistry is really surprising

# Why the non-living world doesn't do (much) chemistry

- Energy is related to probability for near-equilibrium systems

$$k_B T \approx 2.6 \times 10^{-2} \text{eV}$$

$$\text{H-bond energy} \approx 0.02 - 0.3 \text{ eV}$$

$$\left\{ \begin{array}{c} \text{C-C} \\ \text{N-N} \\ \text{O-O} \end{array} \right\} \text{bond energy} \approx 1.5 \text{ eV}$$

$$P(\text{state}) \sim e^{-\Delta G / k_B T}$$

$$\text{UV photon energy} \approx 3 - 124 \text{ eV}$$

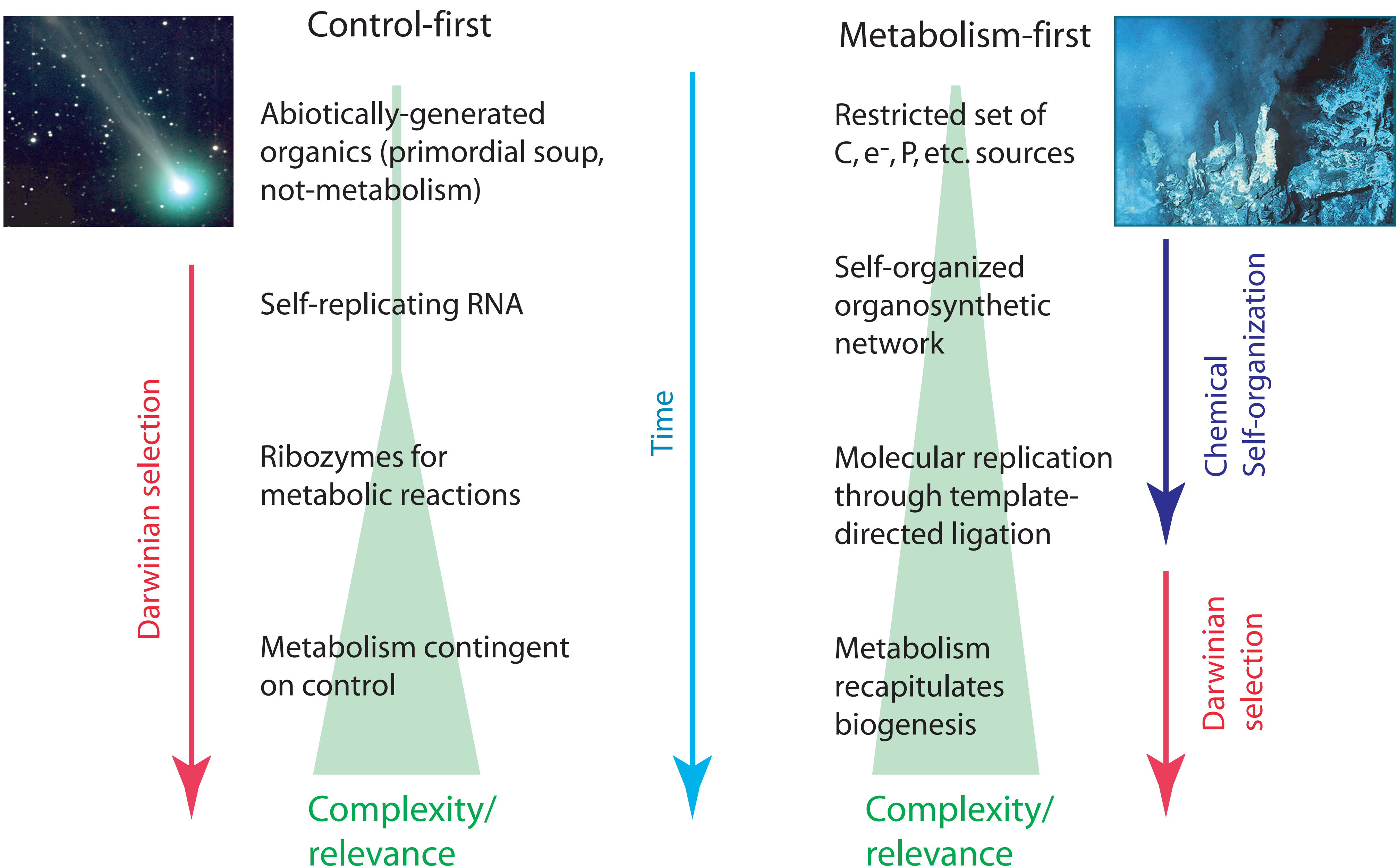
- Morals to the story:
  - We might expect one bonded atom per mole by chance $e^{-50} \approx 2 \times 10^{-22}$
  - (*oh btw. . .* Light capture is hard because it far exceeds chemical bond energies)

# The biosphere as a fourth geosphere

- Lithosphere
  - Solid, radiation heating, metal chemistry
- Hydrosphere
  - Liquid water, absorption albedo, solution chemistry
- Atmosphere
  - Gas, small-light-molecules, ion/radical/electro-chemistry
- Biosphere
  - Steady, high-volume energy transduction through covalent-bond chemistry

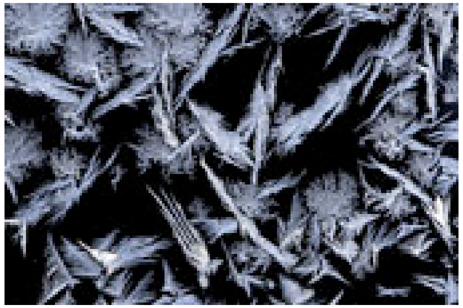
# Reconciling “surprising” order with stability

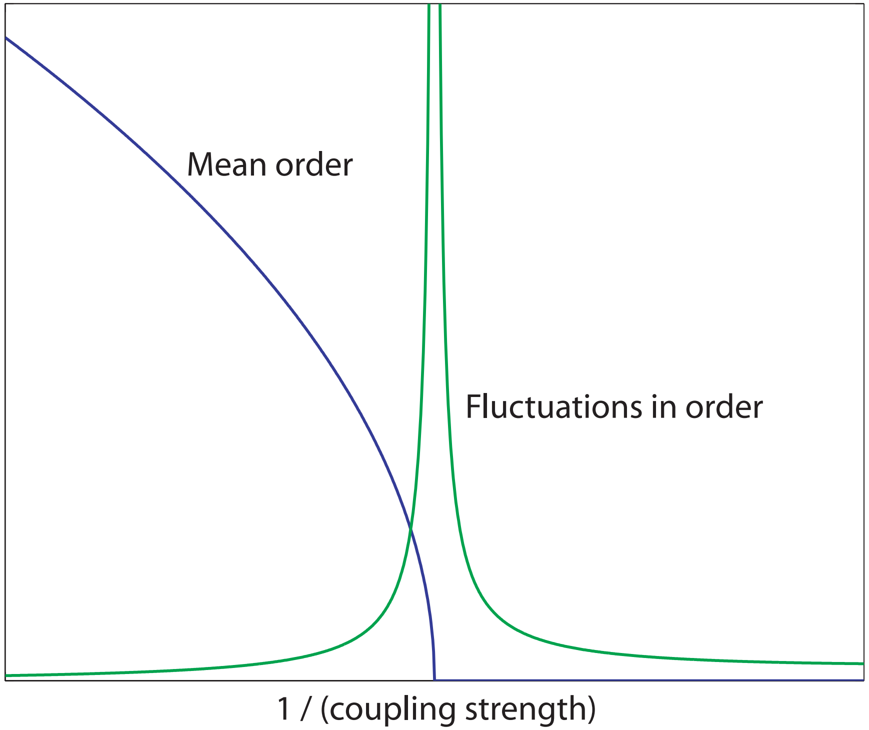
# The choice between hierarchical and flat forces to order



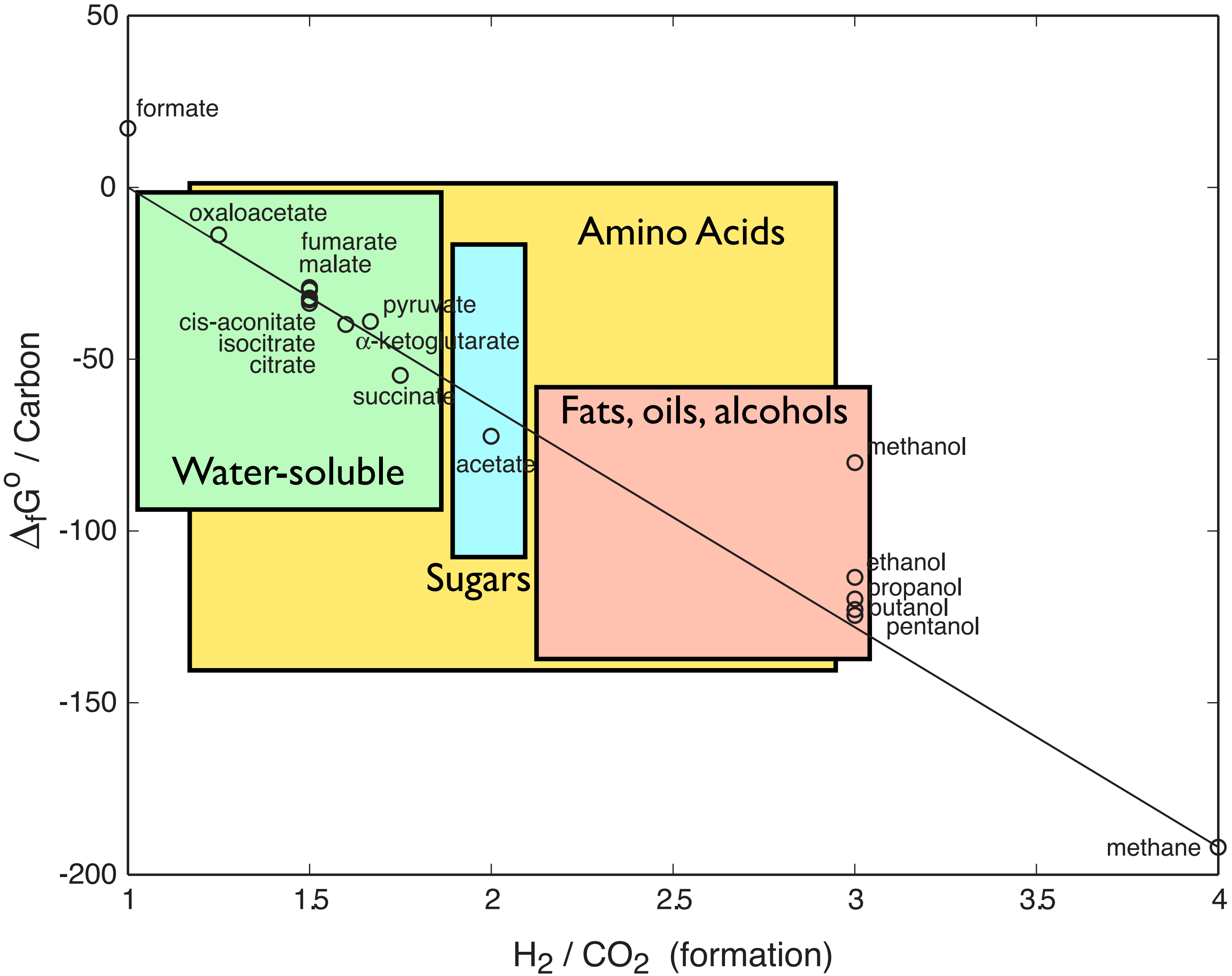
# “Flat” systems and self-organization

- Major lessons from water -> ice
- Order can form without “downward causation” of any nontrivial kind
- Order is *not* easy to form, and cannot be taken for granted scientifically
- Predictable and *inherently unpredictable* components of order depend on one another



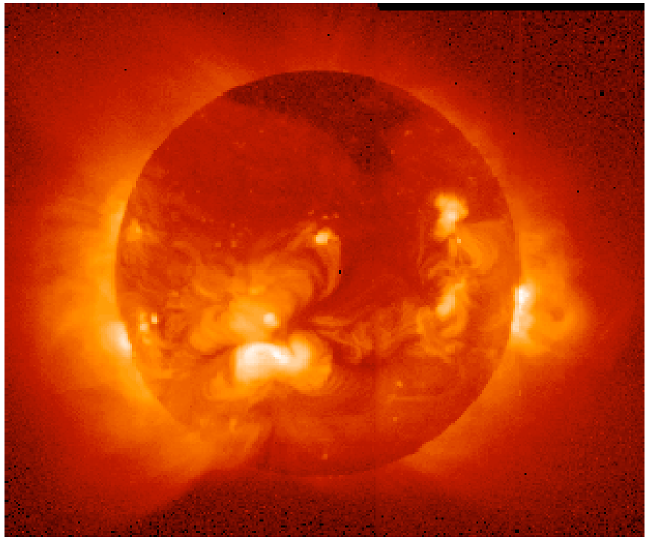


# Biosynthesis is energy-yielding in a reducing world

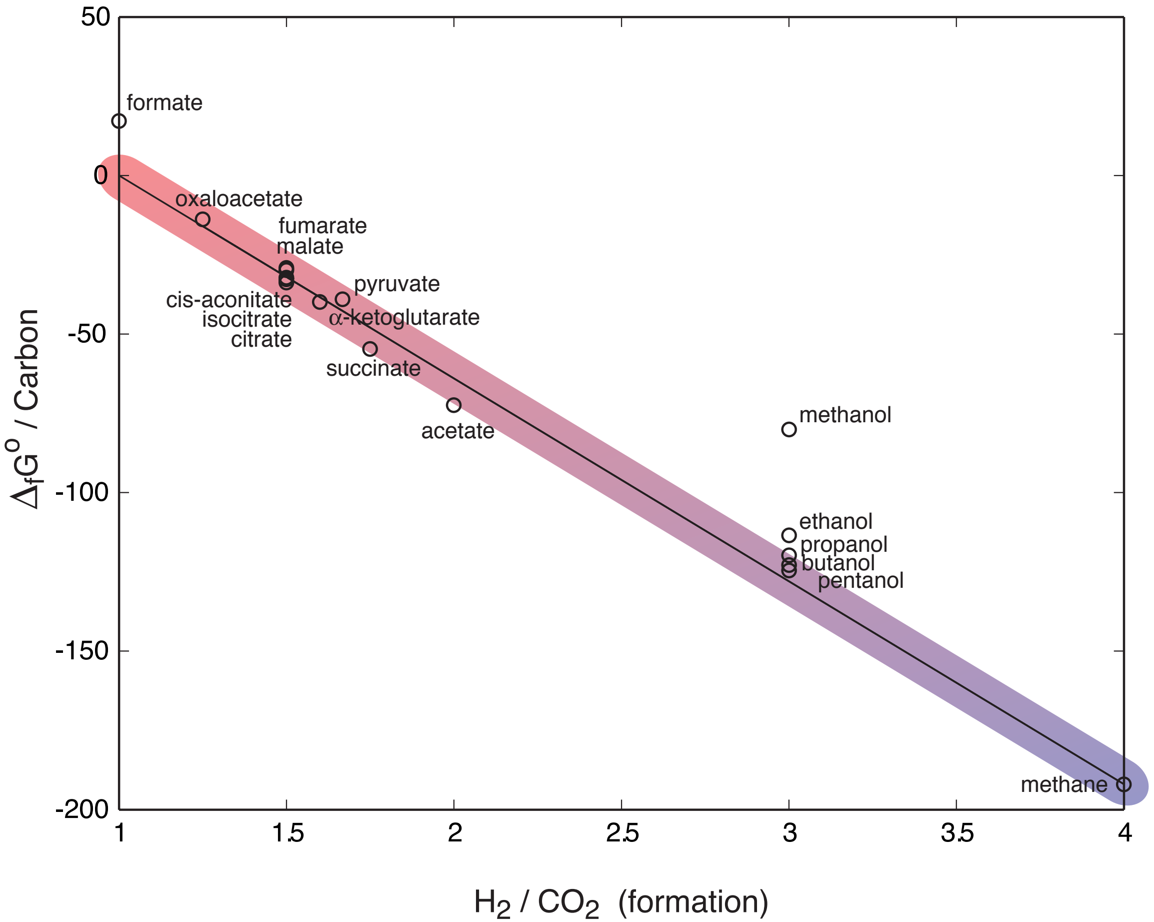


# Biochemistry is deeply tied to its energy sources

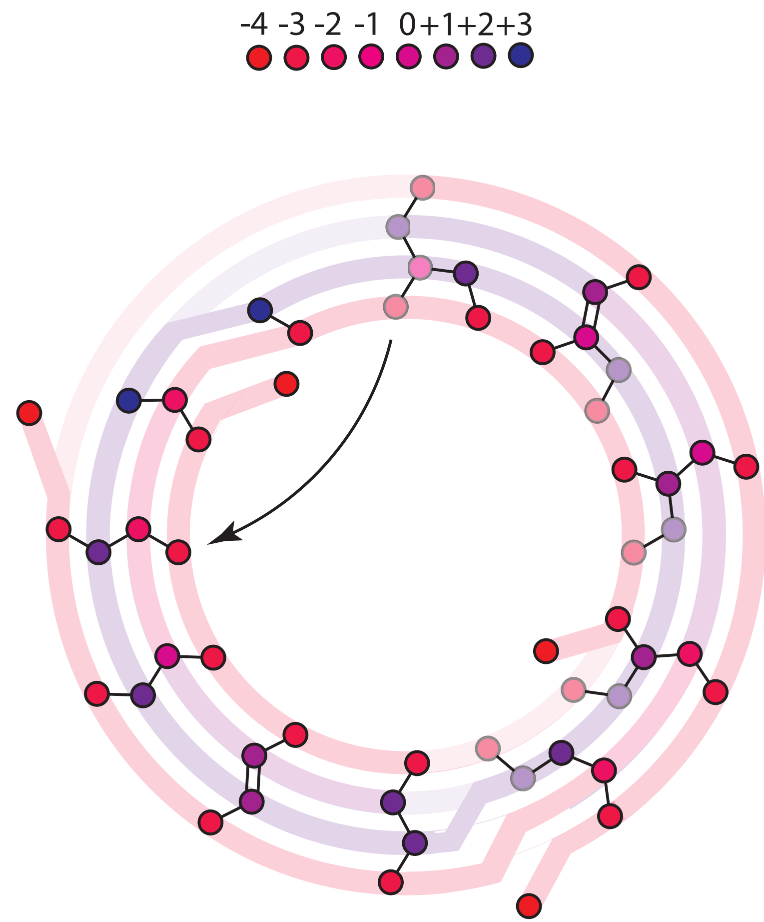
- Major choice is between fission (geothermal heating) and fusion (sunlight)
- Building biomass (anabolism) always requires electrons
- Geochemistry provides electrons directly
- Sunlight must genuinely be “captured” to produce them
- Oxidizing life is (structurally and historically) anabolic life wrapped in an “energy suit”



# Flat organization from non-equilibrium chemistry? the energetics & network topology of carbon fixation

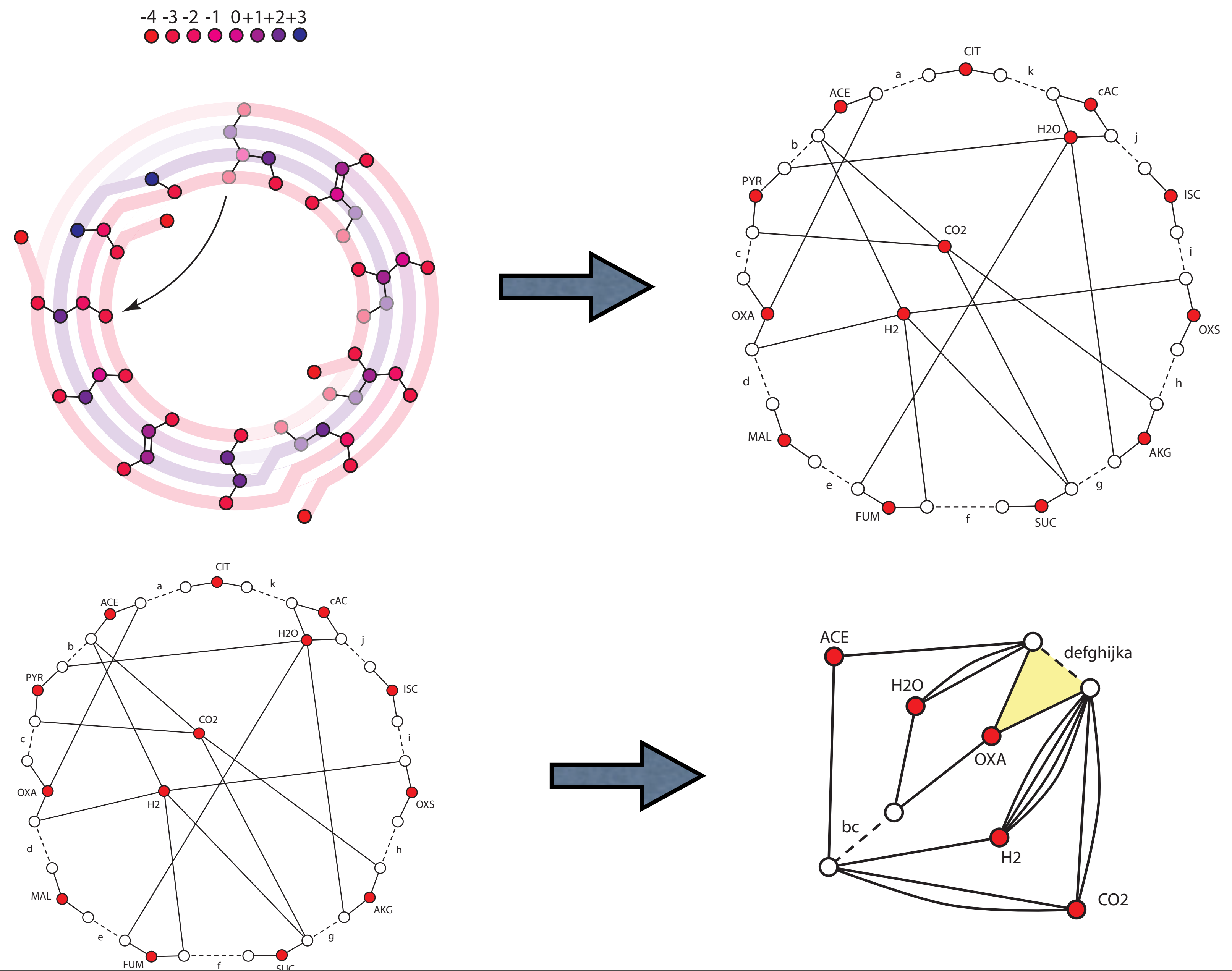



Energetics of the TCA cycle

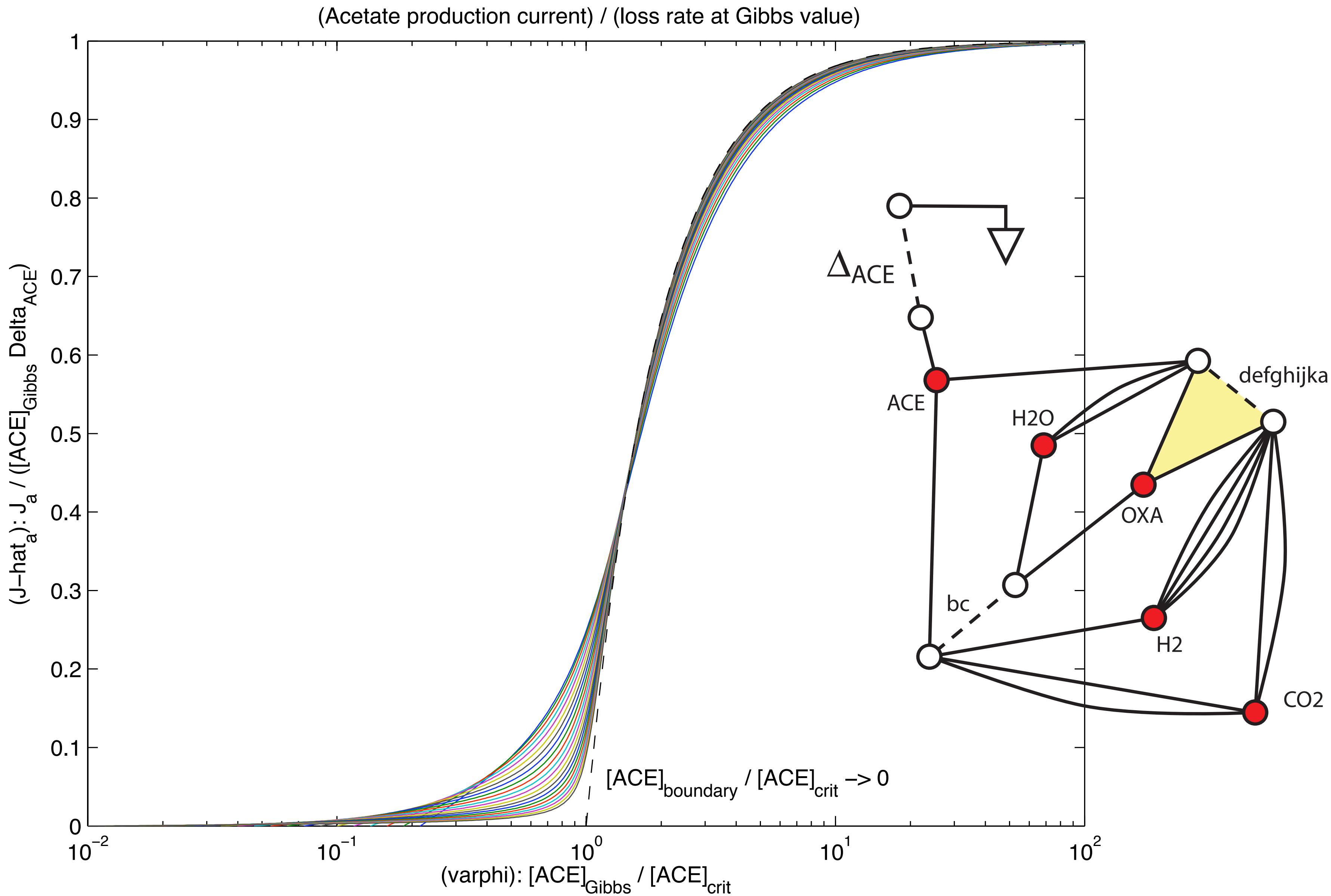



Cycle transport and topology

# Carbon fixation is a microcosm of life

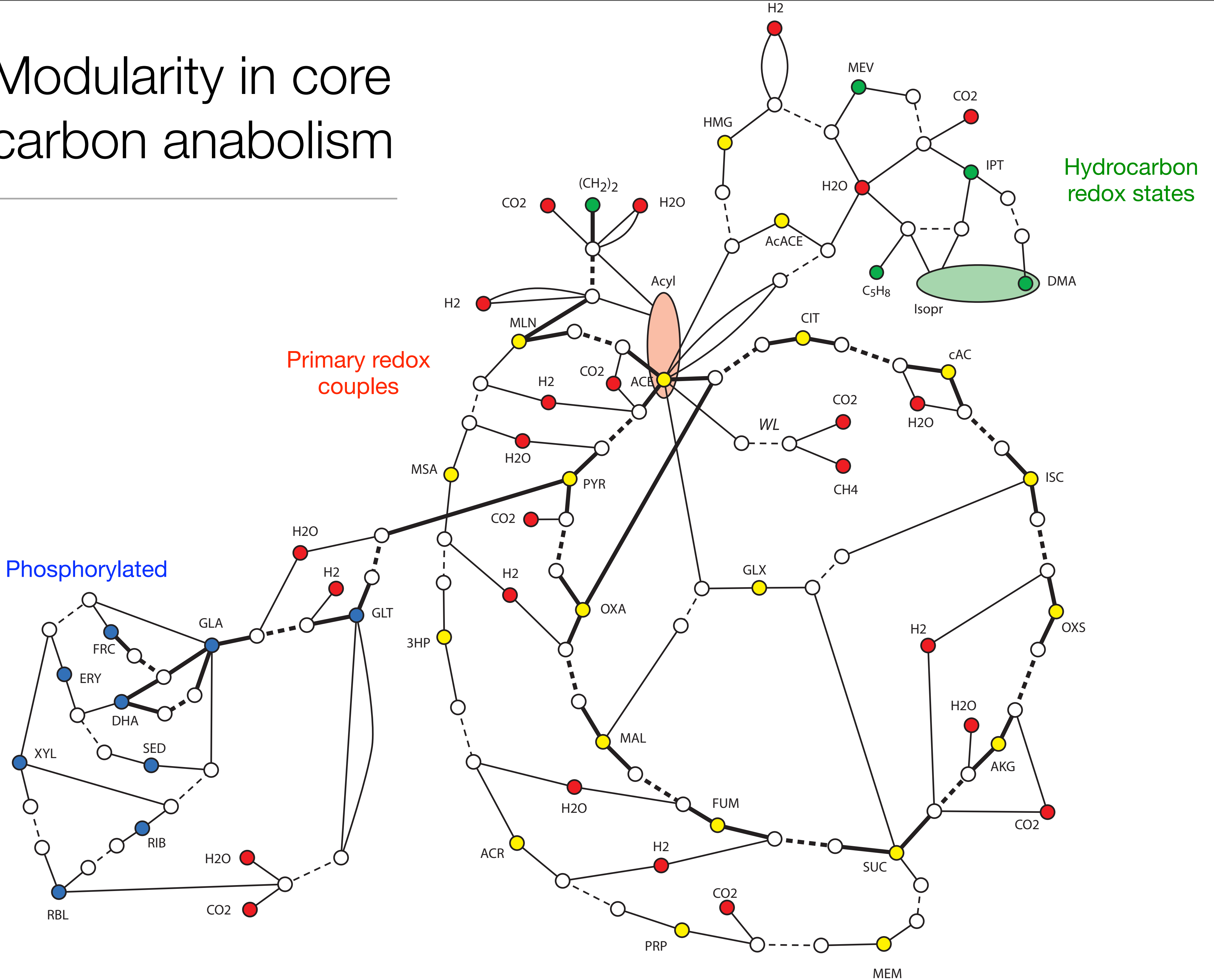


# Autocatalytic cycles can create non-equilibrium phase transitions

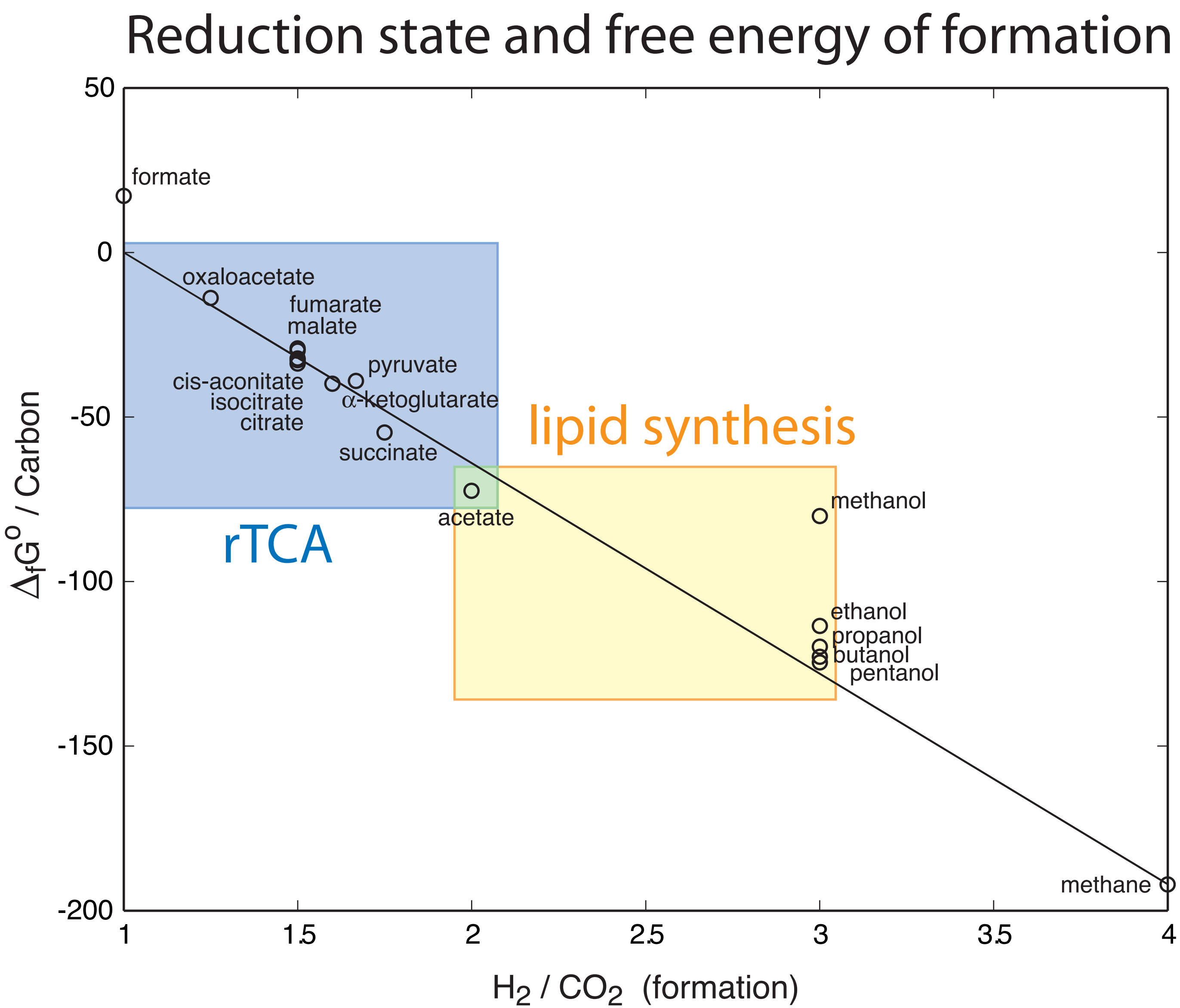


# The structure of metabolism: again

# Modularity in core carbon anabolism

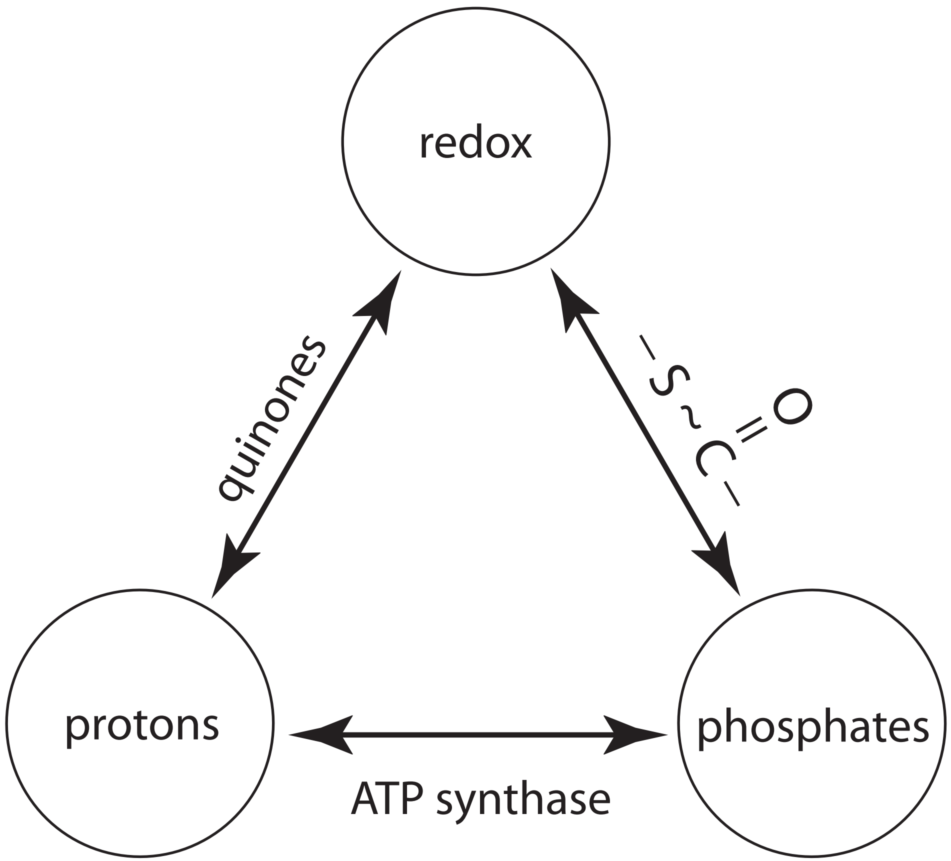


# Modularity in carbon fixation reflects a modularity in chemical energy systems

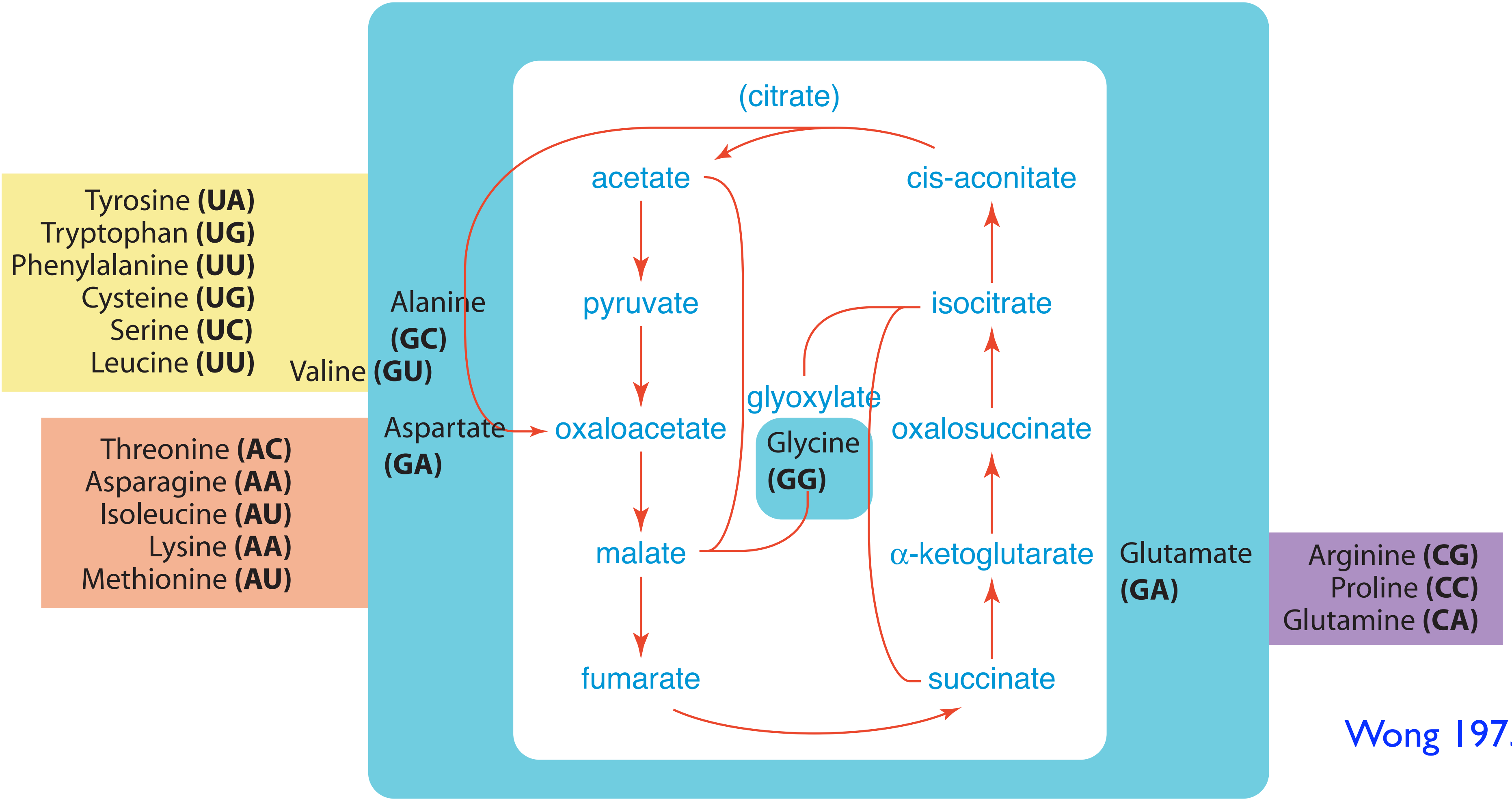


# Modules and micro-environments: implications for contingency of emergence events

- Perhaps best seen in energy metabolism
- Modern cells couple synthetic modules by coupling and buffering energy carriers
- Ox-phos supplements / supercedes substrate-level phosphorylation
- Ox-phos is membrane-mediated, plausibly the last step in emergence (several arguments for this)



# Beyond carbon fixation:
# Synthetic pathways and first-bases in the code




Wong 1975

# Second bases and physical properties in the code

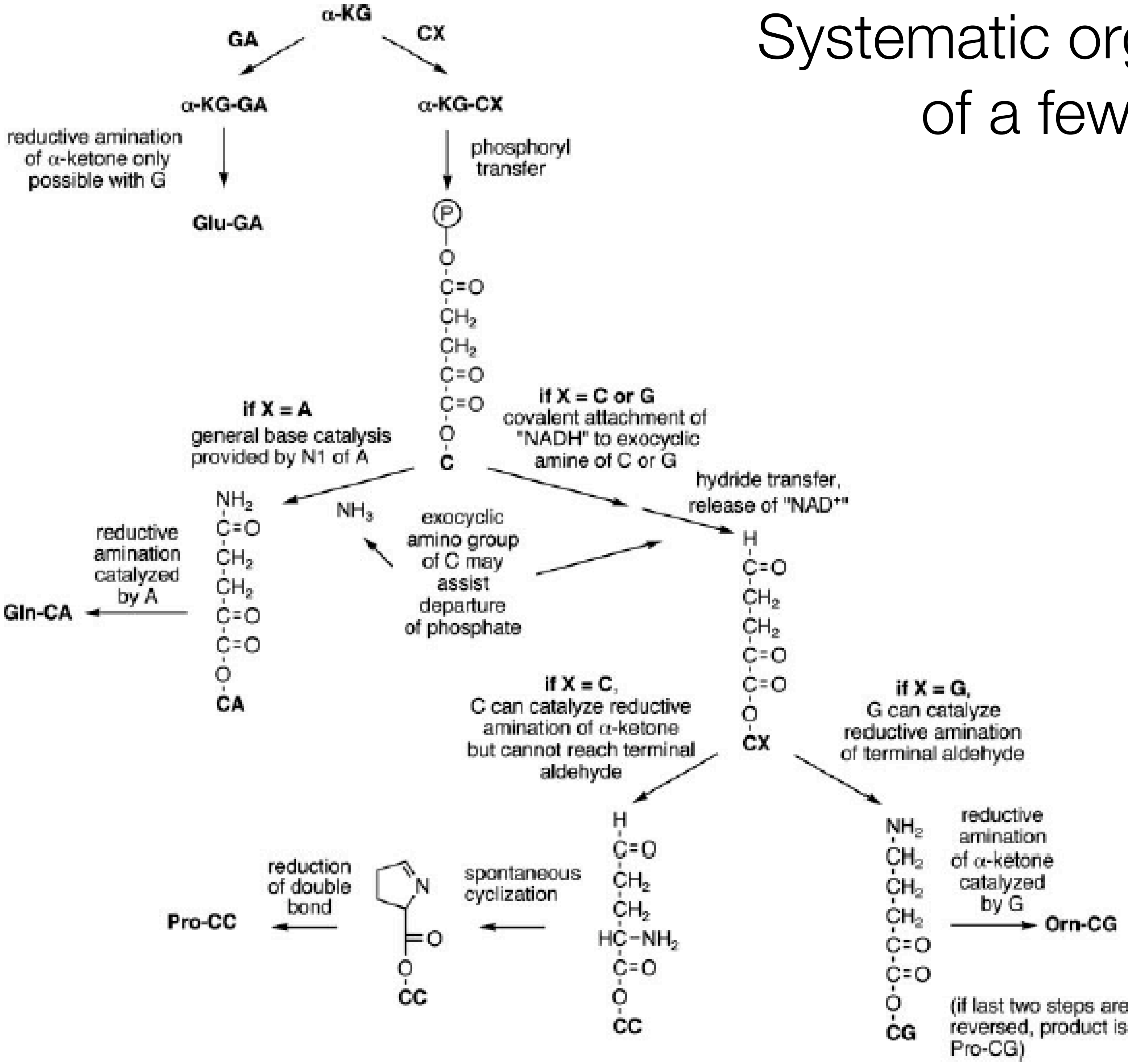
| | G | C | A | U | |
|---|---|---|---|---|---|
| G | Glycine | Alanine | Aspartate | Valine | U |
| | Glycine | Alanine | Aspartate | Valine | C |
| | Glycine | Alanine | Glutamate | Valine | A |
| | Glycine | Alanine | Glutamate | Valine | G |
| C | Arginine | Proline | Histidine | *Leucine* | U |
| | Arginine | Proline | Histidine | *Leucine* | C |
| | Arginine | Proline | Glutamine | *Leucine* | A |
| | Arginine | Proline | Glutamine | *Leucine* | G |
| A | *Serine* | Threonine | Asparagine | Isoleucine | U |
| | *Serine* | Threonine | Asparagine | Isoleucine | C |
| | *Arginine* | Threonine | Lysine | Isoleucine | A |
| | *Arginine* | Threonine | Lysine | Methionine | G |
| U | Cysteine | Serine | Tyrosine | Phenylalanine | U |
| | Cysteine | Serine | Tyrosine | Phenylalanine | C |
| | | Serine | | Leucine | A |
| | Tryptophan | Serine | | Leucine | G |

**Hydrophilic** **Hydrophobic**

Italic: Capture (?)
Font size ~ age

Start
Stop

# Systematic organization of a few reactions



# The “function” of coding enters late (?)
# Signatures mostly in third-base assignments

- Often not used
- Distinguishes simple from complex
- Isolates start and stop positions

| First Codon \ Second Codon | G | C | A | U | Third Codon |
|---|---|---|---|---|---|
| G | Glycine | Alanine | Aspartate | Valine | U |
| G | Glycine | Alanine | Aspartate | Valine | C |
| G | Glycine | Alanine | Glutamate | Valine | A |
| G | Glycine | Alanine | Glutamate | Valine | G |
| C | Arginine | Proline | Histidine | *Leucine* | U |
| C | Arginine | Proline | Histidine | *Leucine* | C |
| C | Arginine | Proline | Glutamine | *Leucine* | A |
| C | Arginine | Proline | Glutamine | *Leucine* | G |
| A | *Serine* | Threonine | Asparagine | Isoleucine | U |
| A | *Serine* | Threonine | Asparagine | Isoleucine | C |
| A | *Arginine* | Threonine | Lysine | Isoleucine | A |
| A | *Arginine* | Threonine | Lysine | Methionine | G |
| U | Cysteine | Serine | Tyrosine | Phenylalanine | U |
| U | Cysteine | Serine | Tyrosine | Phenylalanine | C |
| U | | Serine | | Leucine | A |
| U | Tryptophan | Serine | | Leucine | G |

Start
Stop

# And so on . . .

# So many details; so little time . . .

- The goal remains to explain what is universal and particular (sparse) first
- We look for modularity as evidence of ties to the geosphere, and to separate events of structure emergence (minimize the role assumed for accidents)
- Repeatedly emphasize catalysis as the key to stability and sparseness in the non-equilibrium domain
- These are compatible with being sensitive to the particularity of chemistry

# Making contact with theoretical biology

- Why individuality? What kinds? When and how did they arise?
- Darwinian transitions: (granularity / shared-fate; HGT and speciation . . .)
- How does early self-organization partition into later
  - Population genetics?
  - Development?
  - Niche Construction and ecological dynamics?

# Take-home messages

- The biosphere as a geosphere can incorporate many elements that are not like each other, yet remain a unified and coherent concept
- Continuous steps from geochemistry to biochemistry are possible because part of biochemistry *is* “just” geochemistry
- Some steps in the emergence of life may have been non-equilibrium phase transitions
- Life is complex and Darwinian because chemistry -- in part precisely because of its inaccessibility -- offers complexity in space, structure, and time

# Further reading

Herbert A. Simon, The architecture of complexity, Proc. Am. Phil. Soc. 106:467 -- 482 (1962)

Christian de Duve, Blueprint for a cell (Patterson, Burlington, N.C. 1991) ch.7 (energetics)

Eugene V. Koonin and William Martin, On the origin of genomes and cells within inorganic compartments, Trends Gen. 21:647 -- 654 (2005) (multiple invention of DNA systems)

Shelley D. Copley, Eric Smith, and Harold J. Morowitz, A mechanism for the association oa amino acids with their codons and the origin of the genetic code, Proc. Nat. Acad. Sci. USA 102:4442 -- 4447 (2005) (stereotypical amino acid biosynthetic pathways)

Michael J. Foote, Arnold I. Miller, David M. Raup, and Steven M. Stanley, Principles of paleontology (Freeman, New York, 2006) (concept of “ecospace” and major transitions)

Shelley D. Copley, Eric Smith, and Harold J. Morowitz, The origin of the RNA world: co-evolution of genes and metabolism, Bioorganic chemistry 35:430 -- 443 (2007)

Vijayasarathy Srinivasan and Harold J. Morowitz, The canonical network of autotrophic intermediary metabolism: minimal metabolome of a reductive chemoautotroph, to appear in Biol. Bulletin (2009); Analysis of the intermediary metabolism of a reductive chemoautotroph, submitted (2009)