Is life a cosmic imperative: How would thermodynamics force life into existence?

Eric Smith
Earth-Life Science Institute: Tokyo Institute of Technology
Santa Fe Institute
Georgia Institute of Technology

Searching for Life across Space and Time
Irvine, CA, 5-6 December 2016
Outline

• What does it mean for emergence of a biosphere to have been a “necessary” stage in planetary maturation?

• What detailed features of biochemistry and higher-level architecture can be invoked as having this interpretation?

• When does this move beyond being “essay” and become a theoretical claim? What conceptual systems do we have to support such questioning?

• Problems, opportunities, needed work, and cautions for exoplanet studies
The mathematical concept at issue: breakdown as a state of order in its own right

Breakdown processes are robust states of dynamical order

Q: Is the emergence of a biosphere a chemical version of this?
System characteristics that lead to breakdown states: self-amplifying positive feedbacks
Earth’s redox battery is the most obviously-coupled disequilibrium to autotrophy and core biochemistry.
Where would this show up? Evidence for biochemistry as an outgrowth of geochemistry

- Continuity of ancient and universal biochemistry with selective organic geochemistry

- Could happen by “lock-in” or as shared reflection of paths of least resistance

- Infer phylogenetically from evolution’s “dogs that didn’t bark” -- Small-molecule chemistry unchanged even with modern enzymes

- “Upward causation”: metabolic patterns imprinted at higher levels where Central Dogma reasoning would argue it doesn’t belong
Universality and recapitulation in autotrophic Carbon Fixation and core anabolic pathways

A universal biosynthetic core exists, organized around TCA intermediates.

All innovation in Carbon Fixation has remained close to a TCA template.

Sugar-phosphates, RNA, Calvin cycle

Pathway innovations concentrate in ~single Carbon-addition steps;

Associated with conserved protein families and metal centers
Upward causation: suggestions that the genetic code was entrained on biosynthetic regularities

- Ideal code should be a “firewall”

- Central Dogma suggests arbitrary assignments (Crick’s “adapter” argument)

- Error buffering (polar requirement etc.) departs from naive “adapters”, but still only applies to properties of the full amino-acids as used

- Yet: the code is most “compressible” in terms of early-biosynthetic pathway regularities
Code First Base: Organized according to backbone starting point in the Citric Acid Cycle

Tyrosine (UA)
Tryptophan (UG)
Phenylalanine (UU)
Cysteine (UG)
Serine (UC)
Leucine (UU)

Valine

Aspartate (GA)

Alanine (GC) (GU)

Glycine (GG)

Glutamate (GA)

Arginine (CG)
Proline (CC)
Glutamine (CA)

Tyrosine (UA)
Tryptophan (UG)
Phenylalanine (UU)
Cysteine (UG)
Serine (UC)
Leucine (UU)

Valine

Aspartate (GA)

Alanine (GC) (GU)

Glycine (GG)

Glutamate (GA)

Arginine (CG)
Proline (CC)
Glutamine (CA)
Code Second Base: physical properties reflect reused modifications to TCA backbones

<table>
<thead>
<tr>
<th>first position</th>
<th>second position</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td>Asp/Glu</td>
</tr>
<tr>
<td></td>
<td>Val</td>
</tr>
<tr>
<td>Gly</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td>Asp/Glu</td>
</tr>
<tr>
<td></td>
<td>Val</td>
</tr>
<tr>
<td>Gly</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td>Asp/Glu</td>
</tr>
<tr>
<td></td>
<td>Val</td>
</tr>
<tr>
<td>C</td>
<td>Arg</td>
</tr>
<tr>
<td></td>
<td>Pro</td>
</tr>
<tr>
<td></td>
<td>Gln</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
</tr>
<tr>
<td>(α-ketoglutarate)</td>
<td>Orn</td>
</tr>
<tr>
<td></td>
<td>Pro</td>
</tr>
<tr>
<td></td>
<td>Gln</td>
</tr>
<tr>
<td></td>
<td>?</td>
</tr>
<tr>
<td>A</td>
<td>Ser / Arg</td>
</tr>
<tr>
<td></td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td>Asn</td>
</tr>
<tr>
<td></td>
<td>Ile</td>
</tr>
<tr>
<td>(oxaloacetate)</td>
<td>Dab</td>
</tr>
<tr>
<td></td>
<td>Hsr</td>
</tr>
<tr>
<td></td>
<td>Asn</td>
</tr>
<tr>
<td></td>
<td>Ile</td>
</tr>
<tr>
<td>U</td>
<td>Cys</td>
</tr>
<tr>
<td></td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td>Tyr/stop</td>
</tr>
<tr>
<td>(Pyruvate)</td>
<td>Cys</td>
</tr>
<tr>
<td></td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
</tr>
</tbody>
</table>

Copley et al. PNAS 201:4442 (2005)
A Decision Tree:

Dependence on early biosynthetic pathway steps is even stronger than properties of the terminal amino acids:

Requires some imprint of history, beyond a simple version of error-buffering

Copley et al. PNAS 201:4442 (2005)
Main premises of the usual “metabolism-first” interpretation

• **Continuity premise:** some network selectively organized around these pathways was already a Hadean geochemical feature

• “The rest” of metabolism was what could be most easily built from these precursors, which were present in excess abundance

• As organic control systems came to be found among this inventory, fitness was pre-disposed by the system contexts already in place

• **Consequence:** evolution refined and re-invented catalysts, but many of the pathways were locked in place

• In some cases they were the best solutions anyway; no fitness advantage for deviating from them, even with refined genes and enzymes (evolutionary constraint and convergence tend to exist along a continuum)
When does this “essay” become a theory?
Rehearsing Boltzmann is a start, but not enough

Second law? No problem!

The problem:
When does repair of things that constantly fall apart become a robust and generic outcome?

This is fundamentally a question about **kinetics** and **trajectories**
Thermodynamics neither must nor should mean equilibrium: need a theory of *process* and *history*.

- Concept behind ordered macro-phases:
  - convergence on exponential families
  - scale separates from structure

- Dynamic context: Equilibrium Free Energy $\rightarrow$ Effective Action

- Dynamic phase transitions $\leftrightarrow$ shifts of central tendencies of macrostates

- **Connection to the concept of “repair”:** Optimal error correction is a *thermodynamic* theory in this sense

- Consequence for evolution:
  3-way trade-off between **robustness**, capacity for **complexity**, and **cost**

\[
P_{\text{fluct}} = e^{-N\hat{S}}
\]


\[
P_{\text{error}} \sim e^{-D(C-R)}
\]
Universal metabolism through the lens of dynamical phase transition theory

- Abiotic Earth becomes a dynamically metastable condition

- Maximum-path-entropy condition has the most ways to fluctuate into it and the fewest ways to fluctuate out: definition of “path of least resistance”

- Contributions from familiar equilibrium thermodynamics and chemical kinetics
  1) “Easy” chemistry is water-based pair/group transfer, duplicated widely
  2) “Hard” chemistry is all metal-center, suggesting mineral or MLC origin (?)

- Contributions from network-level of fluctuation/control statistics
  3) Self-amplifying positive feedbacks concentrate matter and energy flows
  4) Feedbacks on shortest-possible (?) loops; less to diffuse/less to control
Current problems: connecting the biological story to convincing synthetic and geo-organic chemistry

- The pathways that seem **necessary and specific** biologically do not seem either **inevitable or special** geologically

- **Reactivity is self-defeating:** e.g. $\text{C}_1$ reduction

  Corollary: formose reaction is in a qualitatively different class from reductive carboxylation

- **What is the right concept of disequilibrium?**
  “Chasing equilibrium?” *(hot/cold or wet/dry cycling)*

  Biochemistry uses **group-transfer cascades** and **“compartmented quasi-equilibrium”** *(acknowledge Y. Oono)*
Proposed useful next steps

• Rudolf Thauer et al., Everett Shock et al. (et al. et al.) have for decades assembled essential databases of thermodynamic landscapes for biochemically relevant organic reactions

• We need a similar encyclopedia for kinetics:
  -- Mineral metal-center catalytic selectivity and activity
    Edge/vertex effects
    Impurities and mineral/mineral interfaces in complexes
  -- Metal-ligand complex catalysis in solution phase
    Element, redox state, ligand-field geometry and orbital properties

• Much of the progress in my generation (microbiology, geo-energetics) has come from decades of painstaking (non-flashy) highly professional work; How best to organize and support such work in the above topics?
Need to connect Hadean atmosphere/ocean calibration to main controlling reaction classes

- Ueno et al. suggest, from quadruple S-isotope fingerprints, more reducing atmosphere with significant CO
- CO as an addition reagent rather than CO$_2$ is intriguing, but much is left to do to tell a full story

**Fig. 5** Vertical mixing ratio profiles of H$_2$, CH$_4$, and CO in our Case 1 ecosystem for $f_{\text{eq}}$(H$_2$) = (a) 200 ppmv and (b) 800 ppmv. The CO$_2$ mixing ratio was fixed at 2500 ppmv for both cases. In the absence of a biotic sink, CO can quickly accumulate to very high levels in a methanogenic ecosystem.

Closing thoughts: Even if this is somehow right…. Cautions for exoplanet studies

• We need to move beyond the mentality of “getting to a material”
  Many conversations are already shifting well in that direction

  • Not “organics” per se, but organics out of equilibrium with the context

  • Big-and-random (tar) is not “complex” -- complex implies elements of selectivity. How much of the selection of metabolism was geochemical?
    Gell-Mann and Lloyd, Complexity 2:44—52 (1996)

• Not all disequilibria are equivalent, w.r.t. problems of biological origin
  Redox $\neq$ radioactivity $\neq$ thermal activation $\neq$ dehydration
Inevitable does not mean generic: Complex Order is often Context Sensitive

Remember: crystallography only has to solve constraints of matching representations of the rotation group to spatial lattices
Thanks to our sponsors

NSF

wpi

NASA

Tokyo Tech