Universality in intermediary metabolism

Eric Smith[*][†] and Harold J. Morowitz[‡]

[*]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501; and [‡]Krasnow Institute for Advanced Study, George Mason University, Fairfax, VA 22030

We analyze the stoichiometry, energetics, and reaction concentration dependence of the reductive tricarboxylic acid (rTCA) cycle as a universal and possibly primordial metabolic core. The rTCA reaction sequence is a network-autocatalytic cycle along the relaxation pathway for redox couples in nonequilibrium reducing environments, which provides starting organic compounds for the synthesis of all major classes of biomolecules. The concentration dependence of its reactions suggests it as a precellular bulk process. We propose that rTCA is statistically favored among competing redox relaxation pathways under early-earth conditions and that this feature drove its emergence and also accounts for its evolutionary robustness and universality. The ability to enhance the rate of core reactions creates an energetic basis for selection of subsequent layers of biological complexity.

Widespread or universal structures and processes in cellular biochemistry are central to a coherent understanding of life, much as universality in physics has become central to understanding order in condensed-matter systems (1). Yet the default interpretation of universality is different in biology than in condensed-matter physics because genetic inheritance controls faithful reproduction of all known organisms. Modern biochemical reactions, and by extension all those reconstructed phylogenetically, are regulated by sophisticated enzymes, and the enzymes are encoded in highly conserved DNA or RNA templates. Where statistical or dynamical extremization principles would be sought to explain universal phenomena in abiotic finite-temperature systems, even the chemical universals in biology are generally presumed to result from common ancestry.

Universals: Inherited or Caused?

Genetics-first origin of life scenarios (ref. 2; see also ref. 3) can be viewed as a reification of this interpretation, an attempt to impose Darwinian selection as the sole dynamical principle governing all stages of life. Self-reproducing genetic templates are proposed as the first emergent biological structures, which then exclusively determine through catalysis what biochemistry is supported. Compartments are needed to enhance the second-order rate kinetics of catalyzed reactions and to associate specific biochemistry with each template to form a unit of selection (4). The emergence of life generally must be seeded by relatively complex molecules generated either astrophysically (5) or from reactions of gas-phase free radicals (6), and only if an opportune genome is found that enables a metabolism does the system become self-sustaining (7, 8). Common ancestry becomes the only plausible explanation for universal processes of any complexity, but what forms emerge are determined by frozen accidents of genetic assembly and as such are not predictable.

The irreducible complexity of genetics-first origin scenarios is high, requiring joint emergence of catalysis, compartmentation, and heritability to make the minimal self-perpetuating structures. The concentration dependence of their synthesis also has been criticized as geophysically unrealistic (9, 10). Metabolism-first scenarios are therefore gaining acceptance as both more plausible and potentially more predictive of observed forms. However, with a few important exceptions, like those of Wächtershäuser (11–13) and Corliss (14), these share with genetics-first theories some version of the Oparin–Haldane conjecture (15, 16), that life emerged in an organically rich environment by

catabolic reactions very different from those that sustain it today (3). As a result, theories founded on Oparin–Haldane imply few specific constraints on primordial metabolic reactions from universals among modern forms.

We argue that the chart of intermediary metabolism (17) has a universal anabolic core, which should not be understood as merely a result of common ancestry but rather as a solution imposed on life already within the energetically structured environment of the early earth, by details of carbon chemistry and certain transport and transformation functions performed only by biomass.[§] This part of biochemistry was selected by statistical and kinetic factors independent of genetics, catalysis, or even cellular compartmentation, factors that governed the emergence of life and its evolution to complex forms in the period before the earliest reconstructible genetic ancestor. By relating the universality of modern organisms to the geochemistry in which cellular life emerged, we are able to propose a specific metabolism for the first organisms and introduce statistical optimization principles under which it may be unique.

Our interpretation of universality in metabolism replaces a pure paradigm of inheritance with a mixed paradigm, where energetics, transport and transformation properties, and ergodic sampling select biological processes sufficiently close to bulk physical chemistry (18), many of which precede the genome structurally and phylogenetically. Genetic inheritance with its element of frozen accident still may determine the possibilities for regulatory structures at higher levels of complexity and contingency, but there remains a unique role for metabolism in biasing selection on those structures that can affect the net anabolic rate. Here we do not pursue a complete scenario for the emergence of life but rather a quantitative and predictive foundation for the treatment of metabolism, on which theories of both origination and the selection of modern forms can be based. We find support for our more causal interpretation of some universals in the large degree of lateral gene transmission inferred among early prokaryotes (19) and expect that similar interpretations will be motivated at higher levels of structure, possibly in protein folding (20), or will unify the growing inventory of evolutionary convergences of phenotypes (21).

Our argument begins in *Metabolic Universality and Geochemical History* with three empirical observations: (*i*) The 11 carboxylic acids of the tricarboxylic acid (TCA) cycle are a unique anabolic core for all of life; (*ii*) the sequence of reactions in the cycle is run oxidatively in modern photoautotrophs and oxidizing heterotrophs but reductively in several chemolithoautotrophs; and (*iii*) the reducing chemistry likely to have characterized at least some environments on the early earth makes the reductive cycle a candidate primordial form. We then look in *Energetic Embedding of Biomass in Its Environment* at the transport functions unique to biomass in a reducing environment and at the free energies of formation and reactivities of the TCA
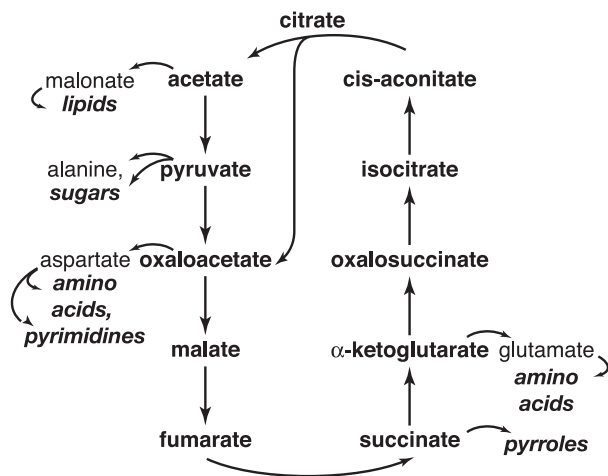

Fig. 1. The reductive TCA cycle (boldface) as an engine of synthesis of the major classes of biomolecules (bold italic). Synthesis of categories usually begins with a specific molecule (lightface).

intermediates as mediators of that transport. In addition to forming an anabolic core, the cycle intermediates form a natural relaxation pathway for the free energy of redox couples created from the earth's ordinary volcanism. In *Concentration Dependence and Network Topology of rTCA* we study the specific reaction network of the reductive TCA (rTCA) cycle, noting that it is topologically network-autocatalytic over a short loop; that its reactions arise from the projection onto a cycle of an indefinitely repeated synthesis of homologous acetate moieties, further reducing possible dependence on multiple prebiotic catalytic innovations; and that the reaction concentration dependence is compatible with a precellular bulk process in appropriate inorganic environments. From these observations we propose that autotrophic rTCA cycling was the metabolism of the first cellular life and that it [or some close variant (12, 13)] may even have preceded cellularity as a bulk relaxation phenomenon.

The Statistical Chemistry of Metabolic Networks discusses the role of ergodic sampling in metabolism-first scenarios like the one we propose and briefly describes extensions of equilibrium maximum-entropy methods for predicting the emergence of autocatalytic cycles like rTCA as the order parameters of dynamical phase transitions. We note which laboratory investigations will be essential to provide kinetic factors as inputs to such calculations. *Feed-Down of Regulation Systems onto Metabolism* then introduces a relation we call "feed-down," through which statistical-mechanical selection pressures on core metabolism induce Darwinian selection biases on higher-level regulatory structures, from catalysts to ecological niches.

Metabolic Universality and Geochemical History

Individual modern species vary widely in their inventories of anabolic and catabolic reactions, but collectively all ecosystems are autotrophic from small C, H, and O molecules, ammonia, and simple inorganic salts and acids. At this level of aggregation, they also share a common chart of intermediary metabolism (17). All anabolic pathways in this chart originate from TCA cycle intermediates, as shown in Fig. 1, making the TCA cycle a unique core for biosynthesis. All lipids come from acetate, all sugars come from pyruvate or 3-phosphoglycerate, all amino acids come from keto acids, and other synthetic pathways are based on these starting materials (17, 22). Every synthetic pathway is used by some autotroph, although among modern organisms there are alternative pathways for synthesis of isolated compounds, such as photosynthesis of 3-phosphoglycerate.

All of biomass is itself composed of <500 small molecules and polymers thereof, most of which participate in many reactions in the metabolic chart. This fact alone constitutes a form of universality, although less restrictive than that of the anabolic core. The most tractable and to us the most natural interpretation of this empirical regularity is that the metabolites were selected through a process of ergodic sampling of common reactions among small C, H, and O molecules (23), rather than by infinitesimally sparse sampling of nucleic acid polymers or proteins catalyzing those reactions. The existence of a small subset forming an anabolic core further suggests that their reactions somehow dominate the sampling process, but it is hard to see how from the role of the oxidizing TCA cycle (Krebs cycle) in those organisms where it has been most studied (24). This cycle is a catabolic cycle extracting energy from phosphenolpyruvate, derived from photosynthetic pathways whose smallest self-regenerating reaction network is the whole metabolic chart of photoautotrophs.

We know, although, that cellular life emerged between 3.5 and 4 billion years ago (22), whereas banded iron formations show that the earth's atmosphere was reducing or at least neutral until ≈2 billion years ago (25). Because photosynthesis is believed to be responsible for loading the atmosphere with molecular oxygen (26) and, even today, is a capability limited to a subset of the ancient lineages (27), life must have achieved nearly modern levels of catalytic sophistication and complexity over 2 billion years without significant use of the Krebs cycle. Further, we will argue below that photosynthesis as a primordial energy source is both unnecessary and suffers from the same problems of discoverability as genes or proteins in an abiotic milieu.

Variations on the Oparin–Haldane conjecture in both gene-first and protein-first origin scenarios separate the capture of energy from the synthesis of biomass, drawing carbon from such fully reduced abiotic precursors as methane, from which hydrogen must be driven endergonically to make organics. Biosynthesis from these is then catabolic and constitutes a preecological extreme form of heterotrophy. In contrast, nonequilibrium-reducing environments provide other abiotic sources of carbon, such as CO_2, and reductant such as H_2, and species exist that are chemoautotrophic from these inputs (28). The discovery of living organisms at the efflux of submarine hot springs (29) suggests that the earliest life may have been autotrophic from magmatic redox couples, making the Oparin–Haldane conjecture and its variants unnecessary. The ability of the nonequilibrium steady-state of vents to provide such redox couples does not depend on whether the surrounding atmosphere is strongly reducing or neutral, and these environments are likely to be the nearest modern equivalents to their counterparts on the early earth.

A reducing pathway involving the TCA cycle intermediates was proposed as a core metabolism for such organisms (30) and later confirmed as the rTCA cycle with the discovery of citrate lyase (31). Our proposal that a preenzymatic rTCA cycle or some variant on it was the first metabolism most nearly resembles the views of Wächtershäuser (12, 13), who has studied detailed mechanisms by which surface chemistry may be able to overcome some of the difficulties of sustaining the necessary reactions in liquid phase (11). As described in the next section, energy capture and anabolism coincide within the rTCA cycle, and many of the side-reactions that potentially remove cycle intermediates are themselves the beginnings of the biosynthetic pathways of Fig. 1.

Energetic Embedding of Biomass in Its Environment

Investigations of how it is possible to generate the complex structures of life tend to have a teleological character that they deny to the process itself. They underutilize a necessary premise for any statistical interpretation of either emergence or universality: A chemical origin of life must have been due to thermo-
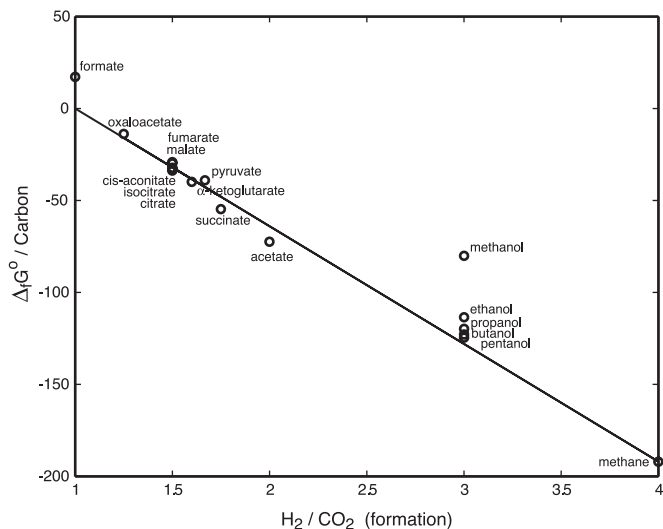
Fig. 2. $\Delta_f G^0$ in kJ/mol per carbon atom of TCA cycle intermediates from $CO_2(aq)$ and $H_2(aq)$, together with reference molecules on the reduction pathway to $CH_4(aq)$. For the reaction $xCO_2 + yH_2 \leftrightarrows X + zH_2O$, reducing potential per carbon taken from the environment to form species X is plotted as y/x on the abscissa. $\Delta_f G^0$ values for cycle intermediates from ref. 35 and for other organic compounds from ref. 36.

dynamic forces away from the abiotic state. As in breakdown of a dielectric under stress, the abiotic state must be unstable against collapse into some more ordered form that incrementally relieves the stress. The consequence, that the conducting form becomes stable, has been likened by Eschenmoser (32) to a dynamic Le Chatelier principle.

The stress relieved by modern deep-ocean chemolithoautotrophic bacteria is the free energy density of those pyrolitically generated redox couples that survive passage from the magma–water interface to the cooler environment of hydrothermal vents (14). With respect to the more unstable redox couples that do degrade, the cooling schedule of the flow is annealing, whereas for the more stable species it is a quench. A corollary is that the kinetic barriers to reaction of the surviving sources of free energy create a bottleneck to its spontaneous degradation, unrelieved by all interactions with abiotic reagents. The existence of persistent reservoirs of free energy requires such bottlenecks; thus, only where these bottlenecks are present can life emerge and persist. Conversely, metastable redox couples introduce a heterogeneous chemical boundary condition that no longer excludes the more ordered living configurations from maximum-entropy ensembles, as equilibrium boundary conditions would (33, 34).

An attempt to relate the universality of the metabolic chart and within it the generative rTCA core to primordial metabolism must concern the energetic and probabilistic accessibility of these molecules, and the transport channel they create for electron pairs from high energy to low energy bond types in a medium stressed by redox free energy. The geochemically primitive sources of redox energy are CO_2 and reductants such as H_2 (12, 13). As shown in Fig. 2, all TCA intermediates have free energies of formation between those and the fully reduced terminal molecules CH_4 and H_2O. They lie in a narrow range of free energy of formation per carbon, and degree of reduction defined as the ratio $[H_2]/[CO_2]$ in the formation reaction, at a local maximum in reactivity driven by carbonyl groups from incompletely reacted CO_2. Smaller molecules along the stoichiometric pathway from CO_2 to CH_4, such as formaldehyde, have higher free energies of formation per carbon and are unstable against collapse to cycle intermediates. A roughly linear relation

between reducing potential from the environment and $\Delta_f G^0$ per carbon, characterizes the most stable molecules, as shown in Fig. 2.

Thus, even non-rTCA relaxation pathways would be expected to generate the more stable of these acids (acetate and succinate) in some excess over equilibrium concentrations. Once present, their enhanced reactivity relative to CO_2 and H_2 makes them more likely participants in further relaxation reactions, as well as providing accessible pathways for anabolic side-reactions. In a nonequilibrium reducing environment, anabolic and energy-transducing reactions can coincide, whereas in the oxidizing environment they are opposed.

Although we do not pursue it here, the energetic role of photosynthesis may be similarly assessed. The stress is free energy density from visible light, which remains unequilibrated with the thermal microwave background in passage both through space and through the atmosphere. The abiotic environment is an imperfect spectral cross-band conductor, because of quantum selection rules in small gas molecules and to a lesser extent the tendency toward photodissociation in larger surface molecules. Photosynthesis enables rapid, repeatable high-energy photon absorption, with the captured energy transduced to microwaves by the combination of biosynthesis and subsequent chemical degradation.

G. Wald (37) has interpreted the absorbance mismatch of chlorophylls and rhodopsins with the terrestrial spectrum as evidence for the difficulty of this innovation, even given the full machinery of modern cells. Together with its noncentrality in the metabolic chart and the fact that it is not an anabolic core, this evidence suggests to us that photosynthesis was discoverable only within the context of a sophisticated synthetic metabolism as an alternative mechanism to supply reductant to an established rTCA core.

The anoxygenic green sulfur bacterium *Chlorobium thiosulfatophilum*, in which rTCA was discovered (30), is a suggestive model for organisms at the first stage of this transition. Unlike most photosynthesizers, it does not use the Calvin–Benson cycle for carbon fixation but drives rTCA directly with reductant produced photosynthetically from sulfides. Reduced ferredoxin drives carboxylation of acetyl-CoA and succinyl-CoA, overcoming the two energy barriers to the reductive cycle, which below we argue should be regarded as a single evolutionary innovation.

The rTCA intermediates were not abandoned in the subsequent large-scale adoption of oxygenic photosynthesis and the resulting oxygen catastrophe, but rather were adapted to the oxidation of acetate (24), releasing CO_2 and reducing NAD^+ into $NADH + H^+$. Only at this stage did the cycle become a mechanism for energy transduction. Anaplerotic reactions added reagents to the TCA cycle, and it remains the synthetic core of everything in the photoautotrophs except sugars, which may come directly from 3-phosphoglycerate. It seems unlikely that this ensemble robustness of TCA is due to genetic inheritance, because different enzymes catalyze parts of the oxidative and reductive cycles, even when both are used by an individual organism such as *Desulfobacter hydrogenophilus* in response to alternative environments (38). It may, however, reflect a genetically induced stability of the whole metabolic chart, within which TCA reactions provided the natural decomposition route for chemicals they had formerly produced (12, 13).

Concentration Dependence and Network Topology of rTCA

The rTCA cycle has three additional features favoring chemical self-organization. It is topologically network-autocatalytic over a single short loop, its synthesis steps are redundant, and all reactions are first-order in cycle intermediates. The first feature makes the cycle self-enhancing at a low level of complexity, the second further reduces the number of innovations for its discovery in a preenzymatic world, and the third frees it from
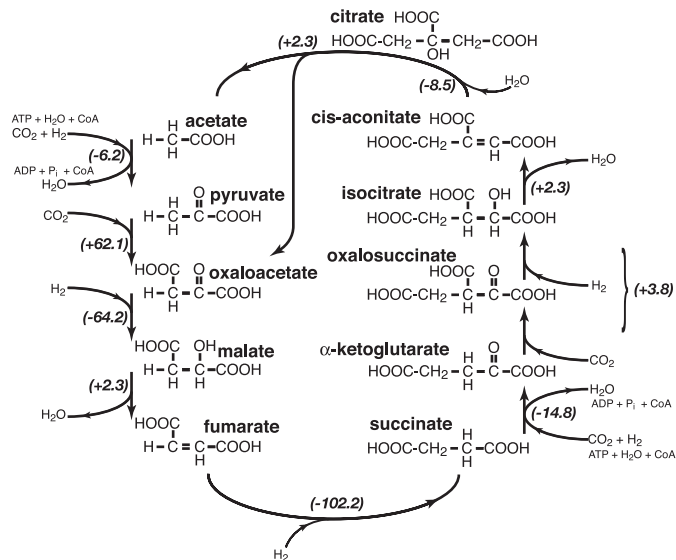
Fig. 3. The rTCA cycle reactions. Synthesis from acetate → fumarate is repeated from succinate → *cis*-aconitate, with CH₂COOH replacing H as end group. $\Delta_r G^0$ values (kJ/mol, bold italic) are for reactions from CO_2(aq) and H_2(aq) in equilibrium with gases at 1atm partial pressure, and H_2O(l). Computed from $\Delta_f G^0$ values in ref. 35, using $\Delta_f G^0 = -386$ kJ/Mol for CO_2(aq), +17.7 kJ/Mol for H_2(aq), and −237.2 kJ/Mol for H_2O(l). $\Delta_r G^0 = -34$ kJ/Mol for ATP hydrolysis corresponds to a local stationary region ≈pH 6 and pMg 1.5 in ref. 39.

requiring diffusion barriers to enhance reaction rates. Fig. 3 shows the rTCA reaction sequence and the environmental molecules, which form essential intermediates to those reactions in modern organisms, ATP and CoA (12, 13).

Without the regeneration of acetate → oxaloacetate, the rTCA cycle would be a synthetic pathway for 2 CO_2 + 4 H_2 + oxaloacetate → acetate + 2 H_2O + oxaloacetate, requiring only first-order reactions in environmental CO_2 and H_2 by using oxaloacetate as a "network catalyst." Acetate → oxaloacetate synthesis introduces the possibility for positive feedback 4 CO_2 + 5 H_2 + oxaloacetate → 3 H_2O + 2 oxaloacetate, or equivalently 6 CO_2 + 9 H_2 + citrate → 5 H_2O + 2 citrate, creating the property of network autocatalysis above a finite threshold for preservation of intermediate states against removal by parasitic reactions.

The synthesis acetate → fumarate, followed by the strongly exergonic saturation of the fumarate C═C bond, generates succinate, whose symmetric halves are copies of the starting acetate. In rTCA, the synthetic sequence is applied again with the same prosthetic groups to one of these legs to go from succinate → *cis*-aconitate, with nearly the same free energies, indicating that the steps are chemically or energetically linked but that the different end-attached groups H and CH₂COOH in the two executions of the sequence are essentially irrelevant.

Repeated application of this minimal synthetic pathway to individual acetate moieties produced by saturation of its ending C═C bond defines an indefinite synthesis with the topology of a line. Projection of the acetate moieties at the end of each synthetic sequence onto the original acetate projects the line to a minimal loop of unique synthetic reaction types, in the manner of a topological covering space. Because of the acetate-repeat structure of the endpoints of synthesis, many different fragmentations lead to two molecules in or near earlier points in the synthesis pathway. At some stage the size of the molecule must favor fragmentation, and in rTCA this action is done in a regular way by nearly reversible rehydration of *cis*-aconitate to citrate,

which then fragments by a retro-aldol reaction to acetate and oxaloacetate. Both an indefinite synthesis creating a cycle by unregulated fragmentation and the actual rTCA cycle deserve consideration as primordial relaxation pathways by the criteria we have proposed.

Although there are 11 rTCA intermediates, the acetate-projection onto the minimal cycle gives carbon incorporation through only 2 types of reactions and hydrogen through only 3. First reverse aldol-condensation leads to the insertion of a carbonyl group, consuming an energetic phosphodiester bond and apparently requiring a thioester intermediate state. Reduction of this C═O subsequent to carboxylation of the adjacent C stabilizes the added carboxyl group.

The coincidence between increasing molecular complexity and decreasing redox free energy per carbon enables an autocatalytic network for capturing redox energy to be far simpler than any autocatalytic network for capturing photon energy suggested by modern organisms. Because all known chromophores use porphyrins in electron transfer (another metabolic universal), the smallest networks that synthesize these contain succinate, and hence potentially rTCA, in a reducing environment.

Finally, all reactions in Fig. 3 involve the concentrations of the cycle intermediates linearly, if usable sulfhydryl and pyrophosphate groups are provided by the environment. If the rTCA reactions can be driven above their autocatalytic threshold with inorganic pyrophosphate and thiols or their equivalent, it will follow that the cycle does not require compartmentation in environments that provide these.

If the acetate pathway for lipid synthesis also can be driven as a bulk process starting from malonate, then the raw materials for vesicles are generated by a parasitic reaction from rTCA, and the first biological need for membranes can be postulated elsewhere. We note that pyrophosphate is the only non-C,H,O intermediate that appears energetically essential to rTCA, and that membrane transduction from redox couples to phosphodiester bonds is another empirical universal of all cellular life (22). A simple speculation is that rTCA may be viable as a bulk process under geochemical conditions that supply pyrophosphate and that the first essential role for the membranes it produced was to generate pyrophosphate in compartments from general redox sources, enabling the enclosed metabolism to expand to a much wider range of environments.

The Statistical Chemistry of Metabolic Networks

We have identified two first-principle approaches to test for the statistical extremality of rTCA in a reducing environment. The more laborious but more straightforward approach is to explicitly account for the chemical potentials of the environmental species and identify the carbon currents that arise when the potentials are not at equilibrium. A less direct but simpler program is to take the characteristic degree of reduction {[H₂]/[CO₂] of formation} and free energy of formation of biomass as given and ask whether the rTCA compounds are in some sense maximal-entropy configurations given these constraints. We sketch the two methods here and treat them fully in separate publications.

Omitting the acetate → oxaloacetate pathway and other side reactions, the remainder of the rTCA pathway is a loop coupled to reservoirs of CO_2, H_2, H_2O, and CH_3COOH. Because there are only three atomic species (C, H, and O), independently specified chemical potentials for the four molecular species are generally incompatible with equilibrium. However, the loop reactions couple to different environmental species on different nodes, creating the chemical equivalent of a spatially extended system in condensed-matter thermodynamics. It is known that such extended systems can be coupled consistently to heterogeneous thermodynamic potentials and that the excess constraints

result in steady-state currents. The current within the network enabling net flow from 2 CO_2 + 4 H_2 → acetate + 2 H_2O is cycling of oxaloacetate. Rate-kinetic evaluation of such transport relations leads to the cycling theorem (40), but following Onsager and Machlup (41), many such near-equilibrium reactions can be evaluated in an effective-potential framework, making the extremization principle explicit.

Adding back the acetate → oxaloacetate pathway and side reactions that result in removal of TCA intermediates, one obtains the full statistical chemistry of redox relaxation in a C, H, and O world. The opposition between positive feedback from autocatalysis and removal by parasitic reactions must produce a phase transition across an autocatalytic threshold, as a function of chemical potential differences from equilibrium. If the network as a whole is sufficiently dominated by the few pathways we have shown, the transition should relate to the simpler loop as a standard nonlinear system with fixed points relates to its linear-response limit.

The difficult inputs to both calculations are of course activation energies and complexes, which must be determined from laboratory synthesis together with geophysically provided context. The abiotic cleavage of citrate has been studied at high temperatures and pressures (42), but energy extraction from inorganic pyrophosphate and the role of thiols remain open problems.

One can obtain a calculation that is well defined without knowledge of kinetic factors by considering all possible bond types in C, H, and O molecules as repositories for electron pairs and asking what is the maximum-entropy distribution of pairs into bonds, constrained by chemical potentials for C, H, and O, and free energy of formation. This calculation is an equilibrium representation of the statistically typical bond distribution if kinetic factors act uniformly on species at common degrees of reduction along the redox relaxation pathway. Preliminary results suggest that the distribution at the reduction typical of biomass is indeed dominated by C—O and C=O bonds, predicting that carboxylic acids are typical. Whether the individual rTCA intermediates are distinct enough to be resolved by this distribution or are more likely constrained by cycling of oxaloacetate has not been determined.

Feed-Down of Regulation Systems onto Metabolism
We have argued that the reactions within the rTCA cycle are part of the natural relaxation pathway for nonequilibrium reducing environments, characterized by positive feedback as a result of their network topology. The population of cycle intermediates relative to all other small molecules at comparable redox potential is determined by chemical competitive exclusion among reaction pathways driven by shared sources of excess free energy. The rate at which rTCA as an anabolic core can deliver organic molecules and free energy to biosynthetic pathways is set by the gross rate of consumption of CO_2 and reductant, the ratio of reactions feeding into cycle intermediates to those draining them away, and the number and inefficiency of steps both within the cycle and leading to other biomolecules. The centrality of rTCA in the chart of intermediary metabolism expresses the fact that biological use is made of a large number of molecules requiring few additional synthetic steps (22).

Many of these properties determining the output of a metabolic core result from the set of possible synthetic processes and the free energies of formation of the species involved and are not changed by catalytic rate enhancement. Others, like reactions within the core and the synthetic spokes radiating from it, can be enhanced catalytically relative to decay, whereas some parasitic reactions that do not lead to biomolecules can be eliminated by selecting the intracellular environment. However, harmful side reactions that cannot be eliminated cost additional metabolic energy to handle. Thus, a metabolic core with high intrinsic

efficiency and statistically favored reactions will in general leave more free energy for the synthesis of higher-level regulatory structures than less intrinsically efficient alternatives. Among alternative pathways exploiting fixed resources, those regulatory systems that augment the statistically favored networks can be expected to exclude alternatives by outgrowing them. This basic premise motivates our presumption that the modern universality of core metabolism reflects the same stabilizing forces that drove prebiotic emergence.

A different form of chemical or even Darwinian competitive exclusion affects alternative catalytic, compartmental, or trophic schemes that influence the bulk rate of a common core like rTCA. Those leading to higher net self-synthesis, through higher core-metabolic rate or more efficient exploitation of side reactions, outgrow those with lower rates. We say that the regulatory structures feed down onto core metabolism and observe that competitive exclusion by primary production is an energetic foundation for the reproductive fitness of any regulatory structure, prebiotic or Darwinian. The feed-down relation between structures that both augment and are generated by metabolic reactions is reciprocal like feedback but explicitly involves a hierarchical relation in which the regulatory structure inherits construction and an arrow of time from the underlying metabolism.

Feed-down defines a quantitative fitness difference for internal structures such as catalysts in competition among autotrophs. More generally it defines a differential growth rate in comparisons of ecosystems that are collectively autotrophic. Because ecosystems are not in Darwinian competition, the interpretation of this growth rate must be in terms of ecological succession or resilience under perturbations. The projection of whole-ecology alternatives onto the interactions of the individual then determines its Darwinian fitness as well as the long-range consequences of variation.

Discussion
The interpretation of metabolic universals and the emergence of metabolism that we are trying to construct differs in several important respects from views generally implicit in discussions of the origin of life (43). First, we emphasize that under preenzymatic conditions, all reactions that relax the free energy of the more stable, small inorganic molecules are impeded. The idea of a "reasonable rate" for a reaction network (43) must be defined relative to the rates of competing reactions with the same inputs and outputs and need not be competitive with the rates in a modern, life-rich world. We look for the emergence of life in a dynamical generalization of the statistical principles of energy distribution, although not necessarily a thermally near-equilibrium state. Just as the equilibrium Gibbs distribution at given temperature and chemical potentials describes a spectrum of molecular species, so we anticipate that an energetically stressed environment should be populated by a distribution of reaction currents (33, 34, 40), of which rTCA is a plausible core for the reasons given above.

In such networks, topology, rate kinetics, and the free-energetic stability of molecules together determine the favored pathways. Whereas the Oparin–Haldane conjecture makes it natural to focus on simple pathways to combine potentially complex inputs, in an autotrophic origin the reaction network *is* the entire synthetic pathway. We have shown (for a few molecules) that on the reduction sequence from CO_2 to CH_4, the smallest molecules are generally less stable than somewhat larger species. The moderate complexity of the rTCA compounds thus does not imply that their decomposition in side-reactions is energetically favored. An exhaustive empirical analysis of the network of side reactions is an important area for future work. When decomposition is not favored, the fact that rTCA is network-autocatalytic from any of its species implies that op-

portune condensations into the cycle from smaller molecules are amplified, as are fragmentations of longer molecules that create cycle intermediates, a relatively easy occurrence because of the repeated-carboxyl structure of the acids. These observations remain valid in the presence of mineral or surface catalysis (11), while making it clear that catalytic enhancement of all reactions in this particular cycle may not be required for the cycle to be a most-favored pathway. Even without specific mineral catalysis, several stages relevant to the rTCA cycle have been demon-strated under plausible conditions (42). Perhaps the best way to view our proposal for a primordial metabolism is as an autotrophic foundation for organic synthesis, on which more complex stages may arise, by something like proposed mechanisms or otherwise.

1. Wilson, K. G. & Kogut, J. (1974) *Phys. Rep., Phys. Lett.* **12C,** 75–200.
2. Miller, S. L. & Orgel, L. E. (1974) *The Origins of Life on the Earth* (Prentice-Hall, Engelwood Cliffs, NJ).
3. Fry, I. (2000) *The Emergence of Life on Earth: A Historical and Scientific Overview*, (Rutgers Univ. Press, New Brunswick, NJ).
4. Hanczyk, M. M., Fujikawa, S. M. & Szostak, J. W. (2003) *Science* **302,** 618–622.
5. Whittet, D. C. B. (1997) *Origins Life Evol. Biosphere* **27,** 249–262.
6. Miller, S. L. (1953) *Science* **117,** 528–529.
7. Crick, F. H. C. (1968) *J. Mol. Biol.* **38,** 367–379.
8. Crick, F. H. C. (1982) *Life Itself, Its Origin and Nature* (Simon and Schuster, New York).
9. Shapiro, R. (1987) *Origins: A Skeptic's Guide to the Creation of Life on Earth* (Bantam, New York).
10. Horgan, J. (1991) *Sci. Am.* **264,** 100–109.
11. Wächtershäuser, G. (1988) *Microbiol. Rev.* **52,** 452–484.
12. Wächtershäuser, G. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 200–204.
13. Wächtershäuser, G. (1992) *Prog. Biophys. Mol. Biol.* **58,** 85–201.
14. Corliss, J. B. (1990) *Nature* **347,** 624 (lett.).
15. Oparin, A. I. (1967) in *The Origin of Life*, ed. Bernal, J. D. (Weidenfeld and Nicolson, London), pp. 199–234.
16. Haldane, J. B. S. (1967) in *The Origin of Life*, ed. Bernal, J. D. (Weidenfeld and Nicolson, London), pp. 242–249.
17. Metabolic Pathways (2003), ed. Nicholson, D. E. (Sigma, St. Louis), product M3907, chart, 22nd Ed.
18. Williams, R. J. P. (2002) *J. Inorg. Biochem.* **88,** 241–250.
19. Morowitz, H. J. (2003) *Complexity* **8,** 12–13.
20. Denton, M. (2002) *J. Theor. Biol.* **219,** 325–342.
21. Conway-Morris, S. (2003) *Life's Solution: Inevitable Humans in a Lonely Universe* (Cambridge Univ. Press, New York).
22. Morowitz, H. J. (1992) *Beginnings of Cellular Life* (Yale Univ. Press, New Haven, CT).
23. Morowitz, H. J., Kostelnik, J. D., Yang, H. & Cody, G. D. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 7704–7708.
24. Melendez-Hevia, E., Waddell, T. G. & Cascante, M. (1996) *J. Mol. Evol.* **43,** 293–303.
25. Greenwood, N. N. & Earnshaw, A. (1984) *Chemistry of the Elements* (Pergamon, New York), pp. 700–701.
26. Blankenship, R. E. (2002) *Molecular Mechanisms of Photosynthesis* (Blackwell, Oxford).
27. Margulis, L. & Schwartz, K. V. (1998) *Five Kingdoms: An Illustrated Guide to the Phyla of Life on Earth* (Freeman, New York).
28. Lengeler, J. W., Drews, G. & Schlegel, H. G. (1999) *Biology of the Prokaryotes* (Thieme, Stuttgart), p. 234.
29. Corliss, J. B., Dymond, J., Gordon, L. I., Edmond, J. M., Von Herzen, R. P., Ballard, R. D., Green, K., Williams, D., Bainbridge, A., Crane, K., *et al.* (1979) *Science* **203,** 1073–1083.
30. Evans, M. C. W., Buchanan, B. B. & Arnon, D. I. (1966) *Proc. Natl. Acad. Sci. USA* **55,** 928–934.
31. Ivanovsky, R. N., Sintov, N. V. & Kondratieva, E. N. (1980) *Arch. Microbiol.* **128,** 239–241.
32. Eschenmoser, A. (1994) *Origins Life Evol. Biosphere* **24,** 389–423.
33. Smith, E. (2003) *Phys. Rev. E* **68,** 046114.
34. Jaynes, E. T. (1983) in *Papers on Probability, Statistics, and Statistical Physics*, ed. Rosenkrantz, R. D. (Kluwer, Boston).
35. Miller, S. L. & Smith-Magowan, D. (1990) *J. Phys. Chem. Ref. Data* **19,** 1049–1073.
36. Plyasunov, A. V. & Shock, E. L. (2000) *Geochim. Cosmochim. Acta* **64,** 439–468.
37. Wald, G. (1974) *Origins Life Evol. Biosphere* **5,** 7–27.
38. Fuchs, G. (1989) in *Autotrophic Bacteria*, eds. Schlegel, H. G. & Bowien, B., (Springer, New York), p. 370.
39. Alberty, R. A. (1969) *J. Biol. Chem.* **244,** 3290–3302.
40. Morowitz, H. J. (1979) *Energy Flow in Biology* (Ox Bow Press, New York).
41. Onsager, L. & Machlup, S. (1953) *Phys. Rev.* **91,** 1505–1512.
42. Cody, G. D., Boctor, N. Z., Hazen, R. M., Brandeis, J. A., Morowitz, H. J. & Yoder, H. S., Jr. (2001) *Geochim. Cosmochim. Acta* **65,** 3557–3576.
43. Orgel, L. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 12503–12507.

BIOCHEMISTRY