BioSystems xxx (2014) 38–54

Contents lists available at ScienceDirect

BioSystems

journal homepage: www.elsevier.com/locate/biosystems

Transient and sustained elementary flux mode networks on a catalytic string-based chemical evolution model

José A. Pereira *

ICBAS–Instituto de Ciências Biomédicas de Abel Salazar, Universidade do Porto, Rua de Jorge Viterbo Ferreira no. 228, 4050-313 Porto, Portugal
REQUIMTE/CEQUP – Centro de Química da Universidade do Porto, R. D. Manuel II, Apartado 55142, 4051-401 Porto, Portugal

ARTICLE INFO

Article history:
Received 9 November 2013
Received in revised form 10 May 2014
Accepted 23 June 2014
Available online xxx

Keywords:
Artificial chemistry
Elementary flux modes
Chemical evolution model
Catalytic network model
Metabolism-first
Pattern-matching catalysis

ABSTRACT

Theoretical models designed to test the metabolism-first hypothesis for prebiotic evolution have yield strong indications about the hypothesis validity but could sometimes use a more extensive identification between model objects and real objects towards a more meaningful interpretation of results. In an attempt to go in that direction, the string-based model SSE ("steady state evolution") was developed, where abstract molecules (strings) and catalytic interaction rules are based on some of the most important features of carbon compounds in biological chemistry. The system is open with a random inflow and outflow of strings but also with a permanent string food source. Although specific catalysis is a key aspect of the model, used to define reaction rules, the focus is on energetics rather than kinetics. Standard energy change tables were constructed and used with standard formation reactions to track energy flows through the interpretation of equilibrium constant values. Detection of metabolic networks on the reaction system was done with elementary flux mode (EFM) analysis. The combination of these model design and analysis options enabled obtaining metabolic and catalytic networks showing several central features of biological metabolism, some more clearly than in previous models: metabolic networks with stepwise synthesis, energy coupling, catalysts regulation, SN2 coupling, redox coupling, intermediate cycling, coupled inverse pathways (metabolic cycling), autocatalytic cycles and catalytic cascades. The results strongly suggest that the main biological metabolism features, including the genotype–phenotype interpretation, are caused by the principles of catalytic systems and are prior to modern genetic systems principles. It also gives further theoretical support to the thesis that the basic features of biologic metabolism are a consequence of the time evolution of a random catalyst search working on an open system with a permanent food source. The importance of the food source characteristics and evolutionary possibilities are discussed.

1. Introduction

Our planet begun as a "fire ball", too hot and unstable to sustain life. Since only non-biologic substances were present after the Earth's cool down there had to be a chemical evolution, within the fundamental natural laws, for life to have emerged from those initial molecular conditions. Prebiotic chemical evolution is commonly accepted, almost by default, as the first step in the process by which life originated on Earth but, concerning the path from early molecular pools to the contemporary biological cell, the only fact known for sure is where this path has ended. The biological cell's main molecular and structural features, experimentally established in the 20th century, remain the main source of inspiration for chemical and molecular evolution, theoretical and experimental research.

Considering the prebiotic evolution models proposed until now, a main dichotomy between "genetics-first" and "metabolism-first" models is apparent (Orgel, 1998; Davies, 2001; Mann, 2013). The former are based on the remarkable properties of RNA polymers that can act as catalysts as well as information storage. Although the RNA-world works, and is a base to model further steps of chemical evolution towards a living cell, it remains to be explained how relatively complex molecules like the ribonucleotides were synthesized on the prebiotic world without the help of a previous metabolism that can produce them from the much simpler molecules available (Orgel, 2004).

The "metabolism-first" models, on the other hand, prescribe reaction networks of catalytic polymers that start with the

* ICBAS–Instituto de Ciências Biomédicas de Abel Salazar, Universidade do Porto, Rua de Jorge Viterbo Ferreira no. 228, 4050-313 Porto, Portugal.
Tel.: +351 220428335/6069420.
E-mail address: jpereira@icbas.up.pt (J.A. Pereira).

http://dx.doi.org/10.1016/j.biosystems.2014.06.011
0303-2647/

simplest compounds as the food set to produce a variety of larger molecules. In the absence of genetic support, its pathways would have to endure and evolve without an information storage system for the catalysts. On the experimental front, there is evidence for the prominent prebiotic role of amino acids and peptides since the Miller experiment, following previous hypothesis of Oparin and Urey (Miller, 1953). The plausibility of prebiotic protometabolic peptide cycles, of activated polymerisation and degradation, was concluded in a recent review (Danger et al., 2012). A prebiotic pathway for oligopeptide synthesis from CO_2 and $N_2/NO_3^-/NH_3$, involving metal cations crucial to biological chemistry, was also proposed (Pratt, 2011). Peptides are known to catalyse the condensation of other amino acids, peptides and peptide-nucleic acids hybrids (Gorlero et al., 2009). It was recently found that the peptide Ser-His is able to catalyse the formation of RNA-phosphodiester bonds in water–ice phases (Wieczorek et al., 2013). Non-peptidic "worlds" have also been proposed at the onset of chemical evolution. Metal ions are known for their catalytic properties and were proposed as main catalytic and redox agents for the autocatalytic synthesis of small molecular weight non-hydrolysable carbon compounds from dissolved gases in hydrothermal conditions (Wächtershäuser, 2007).

This and other data suggest a mutating prebiotic polymer world, catalysing the formation of other polymers, like peptides or nucleic acids but also of other kinds of carbon polymers. Metabolism and genetics evolutionary models are complementary because together they span evolution from the simplest to the most complex biological molecules, supporting each other, but it seems more likely that a genetic world is in need of a metabolism than the other way around (Norris et al., 2012; Bernharth, 2012).

Regarding theoretical approaches, the characterization of the living cell as a self-reproducible (autopoietic) autonomous organization is a reference point, expanded by formalizations of the biological organization (Eigen, 1971; Varela et al., 1974; Fontana and Buss, 1994; Rasmussen et al., 2001; Gánti, 2003). One of the main theoretical strategies has been to develop models that achieve sustained agent organizations upon simulation, starting from disorganised systems.

A crucial element of the cell's autopoietic organization are chemical cycles, notably self-catalysed cycles that dissipate external chemical potential by working in non-equilibrium conditions (Eigen and Schuster, 1977). Recently, Kreyssig et al. demonstrated, using chemical organization theory methods, how the existence of cycles is a necessary condition in non-trivial stationary states in reaction networks (Kreyssig et al., 2012). Autocatalysis is the property of a self-catalysed cycle, that is, a cycle that produces its own catalysts, and, naturally, it is also an important aspect of chemical evolution models (Hordijk et al., 2010). Autocatalysis is not a sufficient condition for life but it is nonetheless a necessary condition (Gánti, 2003; Ruiz-Mirazo and Moreno, 2004). It is the main mechanism by which the cell's autopoietic organization is achieved and regulated, nucleic acids included, and it is a property to look for in chemical evolution model simulations because it assures the maintenance and reproduction of the reaction set.

Chemical evolution models aim to simulate the evolution of quantity and quality of complex molecules, like polymers, but the rigorous simulation of a real small polymer on a computational environment is still very difficult, let alone a variety of complex polymers. The traditional workaround to this problem is to use an artificial chemistry (Dittrich et al., 2001; Suzuki et al., 2003). Artificial chemistries are abstractions of real chemistry where only certain key aspects of molecules, interactions, dynamics and space are retained. Their interpretation may be literal or algorithmic. A general way of representing molecules in an artificial chemistry model is to use alphanumerical strings where the properties of the agents are coded in several ways on the character sequence. Catalytic properties, for example, may be algorithmic interpretations of string sequences while reactions or template replication can work orthographically on strings, changing its character sequences (Tominaga et al., 2008).

On many artificial chemistry models, the energy involved in agent transformations is not explicitly calculated, although it may be deducible from kinetic mechanisms and parameters, if present. Some approaches to a formal definition of energy transformations use tools from physics to derive kinetic and potential energies for the artificial chemistry agents (Benkö et al., 2002; Mayer and Rasmussen, 1998; Yoshii and Kakazu, 1997). A different approach is to apply the chemical thermodynamics concept of standard formation energies to assign energetic value to agents and their reactions (Fernando and Rowe, 2007, 2008).

Motivated by the metabolism-first hypothesis of autocatalytic sets of proteins (or polymers in general), Kauffman, Hordijk and Steel (Kauffman et al., 1986; Hordijk and Steel, 2004; Hordijk et al., 2012), Jain (Jain and Krishna, 2000), Bagley and Farmer (Bagley and Farmer, 1991; Farmer et al., 1986) and Vasas et al. (Vasas et al., 2012), among others, developed models of catalysed reactions of string cleavage and condensation using string food sources. The species evolve in variety and concentration on an open well-stirred reactor and the network of catalysed reactions was analysed using graph theory methods. The set of differential equations used for concentration simulation regularly changes to accommodate for new reactions and string species. These non-genetic models revealed sustained metabolisms but, although quite general, their correspondence to biological metabolisms is not easy to establish. In addition, the focus is mainly on kinetics and not on energetics, which makes reasoning about energy flows more difficult. The matter flow would also be easier to follow if some molecular analogies to biological compounds were used. These analogies, if taken coherently, can add a layer of meaning to the strings, the reactions and the energetics, making the results more easy to interpret in chemical or biochemical terms.

In an attempt to develop a chemical evolution model that adheres to these guidelines, the model SSE ("steady state evolution") was developed and is now reported. This model uses an artificial catalytic string-based chemistry, sustained by a food source, but it does not include concentrations or kinetics; the evolution is only on string quality. A detailed approach on string orthography, reaction syntax and energetics was used, expecting some new insights on catalytic systems without the need to make kinetic assumptions like rate constants, reaction mechanisms or initial concentrations, therefore increasing the generality of the model in that respect. Energetics allows to reason about reaction or pathway extension in a time-independent manner by interpreting equilibrium constant values. Although not explicit in the model, time is nevertheless implicit in the SSE algorithm for state evolution, described below. Time can be identified with the number of strings exchanged or the number of updates made, even if the precise relation between these changes and time is unspecified.

Simulations begin with a pool of strings, some permanent, the food set, and some exchangeable. The exchange of strings represents the unspecific and uncatalytic breakdown and random reconstitution of polymers driven by a virtual external energy source. This update mechanism adds random novelty to the pool and is a metaphor of an open, non-conservative, dissipative system with structural instability (Prigogine and Kondepudi, 1999). After each exchange, the specific catalysts are identified and the catalysed reaction set is defined, adding reaction products to the pool. After analysis of the pool state, the algorithm returns to the string exchange step.