Long-term behaviours of Autocatalytic Sets

Alessandro Ravoni

Department of Mathematics and Physics, University of Roma Tre, Via della Vasca Navale 84, 00146 Rome, Italy

ARTICLE INFO

ABSTRACT

Autocatalytic Sets are reaction networks theorised as networks at the basis of life. Their main feature is the ability of spontaneously emerging and self-reproducing. The Reflexively and Food-generated theory provides a formal definition of Autocatalytic Sets in terms of graphs with peculiar topological properties. This formalisation has been proved to be a powerful tool for the study of the chemical networks underlying life, and it was able to identify autocatalytic structures in real metabolic networks. However, the dynamical behaviour of such networks has not been yet complitely clarified. In this work, I present a first attempt to connect the topology of an Autocatalytic Set with its dynamics. For this purpose, I represent Autocatalytic Sets in terms of Chemical Reaction Networks, and I use the Chemical Reaction Network theory to detect motifs in the networks'structure, that allow to determine the long-term behaviour of the system.

© 2021 Elsevier Ltd. All rights reserved.

1. Introduction

Autocatalytic Sets (ASs) are sets of chemical species that mutually catalyse each other's production through chemical reactions, starting from a basic food source. Introduced by Kauffman (1971, 1986, 1993), the notion of AS captures the idea of a system capable of spontaneously emerging and self-reproducing, using the resources provided by the environment. These features are commonly accepted as basic qualities of early life forms (Luisi, 2016; Kauffman, 1993; Higgs and Lehman, 2015; Nghe et al., 2015; Rasmussen et al., 2016), and many authors have suggested that the appearance of ASs can have been a fundamental step in the transition from non-living to living matter, possibly preceding RNA-based systems (Kauffman, 2011; Hordijk et al., 2010; Hordijk and Steel, 2018; Dyson, 1985; Vasas et al., 2012; Xavier et al., 2020). However, some properties of ASs are still under debate, such as their effective ability to form spontaneously (Lifson, 1997; Szathmáry, 2000) and undergo adaptive evolution (Vasas et al., 2010, 2012), a feature that is generally referred to as *evolvability*.

A formal definition of ASs has been proposed in the context of graph theory (see Section 2.2 for details) by the *Reflexively Autocatalytic and Food-generated* (RAF) theory (Hordijk and Steel, 2004; Hordijk et al., 2011). This allowed to overcome some limits associated with ASs and to make evident their relevance in the context of the origin of life. In particular, RAF theory has successfully proved that ASs (or *RAF sets*, using the nomenclature of RAF theory) are

highly likely to exist in a catalytic reaction system (Hordijk and Steel, 2004; Mossel and Steel, 2005; Hordijk et al., 2011; Hordijk and Steel, 2012; Vasas et al., 2012), and RAF sets have been detected in the metabolic networks of Escherichia Coli (Sousa et al., 2015) and ancient anaerobic autotrophs (Xavier et al., 2020). Moreover, RAF theory shows that RAF sets are often composed by RAF subsets, (Hordijk and Steel, 2018; Hordijk et al., 2012; Hordijk and Steel, 2014), and it has been argued that (some of) these subsets could be the elementary units on which natural selection may act (Hordijk and Steel, 2014; Hordijk et al., 2012; Vasas et al., 2012; Hordijk and Steel, 2018; Hordijk et al., 2018). First numerical results support this scenario, at least for compartmentalised RAF sets (Vasas et al., 2012; Hordijk et al., 2018; Serra and Villani, 2019; Ravoni, 2020). However, in the same works it has been observed that the conditions a graph must satisfy to be a RAF set (see Section 2.2) are not sufficient to strictly constrain its dynamics, and the system may not exhibit the features necessary to be evolvable. More generally, the connection between the topology of a RAF set and its dynamics is not yet fully understood (Filisetti et al., 2011; Serra and Villani, 2019; Hordijk et al., 2018; Hordijk and Steel, 2012),

In this work I make a first attempt to solve this issue, by studying RAF sets through the formalism of the *Chemical Reaction Network* (CRN) theory (Feinberg, 2019). The CRN theory offers powerful tools that allow to connect the topology of a network with its dynamical behaviour, for a large class of network dynamics (Feinberg, 2019, 1987, 1995; Ellison, 1998; Ellison et al., 2011). Note that the CRN theory usually deals with the study of real-world chemical systems. Starting from the results of the CRN

theory, I identify which conditions on the structure of RAF sets are useful for predicting their long-term behaviours. In particular, I investigate which conditions determine a dynamics with desired properties for systems at the origin of life (e.g., self-maintenance, concentrations growth, homeostasis), adding new elements in the scenario previously proposed for the evolvability of RAF sets (Hordijk et al., 2018; Vasas et al., 2012).

The paper is organised as follows. In Section 2 I recall notions of the CRN theory useful for our study, and the definition of RAF sets. In Section 3 I show my representation of RAF sets in terms of CNRs. In Section 4 I introduce motifs in the structure of RAF sets that are connected with long-term behaviours of the networks, and I provide an illustrative example of evolvability of RAF sets by numerically simulate the dynamics. Finally, in Section 5 I discuss the conclusions.

2. Background

2.1. Chemical reaction networks and the deficiency zero theorem

In this Section I recall some useful definitions and introduce the *Deficiency Zero Theorem* (DZT) (Feinberg, 1987; Feinberg, 1995). An interested reader can find a detailed discussion of the CRN theory in Feinberg (2019). Let $\mathcal{S}$ denote the set of chemical *species* and $\mathcal{N} = \{v_i, v'_i : i = 1, \cdots, r, \ v_j \neq v'_i\}$ the set of *complexes*, that is, the members of the vector space $\mathbb{R}^{\mathcal{S}}_{>0}$ generated by the species that provide the inputs and the outputs of a reaction. I denote with $v(s)$ the *s*-th component of the vector v, that is, the *stoichiometric coefficient* of species s in the complex v. Let $\mathcal{R} = \{v_i \to v'_i : i = 1, \cdots, r\}$ be the set of *reactions*, that is, the set of relations among complexes. For each reaction $v_i \to v'_i \in \mathcal{R}$, the component $v_i(s)$ indicates the number of molecules of the *s*-th chemical species consumed by the *i*-th reaction, while the component $v'_i(s)$ indicates the number of molecules of the *s*-th species produced by the *i*-th reaction. A CRN is a triple $(\mathcal{S}, \mathcal{N}, \mathcal{R})$. Note that it is possible to consider *open* CRNs introducing pseudo reactions such as $0 \to v$ and $v \to 0$, where 0 is called the *zero complex* and it is the zero vector of $\mathbb{R}^{\mathcal{S}}$.

I denote with $\mathcal{G}$ the directed graph with complexes as nodes and reactions $v_i \to v'_i$ as directed edges. The connected components of $\mathcal{G}$ are called the *linkage classes* of the CRN. A *strong-linkage class* is a strongly connected linkage class. A *terminal strong-linkage class* is a strong-linkage class containing no complex that is a source for an edge pointing to a different strong-linkage class. Note that a complex not involved in any reaction is a terminal strong-linkage class. A CRN is said to be *weakly reversible* if each linkage class in $\mathcal{G}$ is a strong-linkage class, or, equivalently, if each linkage class is a terminal strong-linkage class. The *reaction vectors* for a CRN are the members of the set $\{v'_i - v_i \in \mathbb{R}^{\mathcal{S}} | v_i \to v'_i \in \mathcal{R}\}$. The span generated by the reaction vectors is called the *stoichiometric subspace* of the network, and its dimension is denote by d. The *deficiency* of a CRN is:

$$\delta = n - l - d, \qquad (1)$$

where n and l are the number of complexes and linkage classes, respectively. Note that the dimension d of the span of two networks with the same complexes and the same linkage classes is the same. This implies that the two networks have the same deficiency δ, and that it results $\delta \geq 0$ (Feinberg, 1995).

In this work I refer to *mass-action systems*, that are CRNs taken together with an element $\kappa \in \mathbb{R}^{\mathcal{R}}_{>0}$. The number $\kappa_{v \to v'}$ is the rate constant for the reaction $v_i \to v'_i \in \mathcal{R}$. Hereafter, I use $\mathcal{G}$ to indicate both a CRN, its associated directed graph and the resulting mass-action system. The *species-formation rate function* for a mass-action system is the function $f(\cdot, \kappa) : \mathbb{R}^{\mathcal{S}}_{\geq 0} \to \mathbb{R}^{\mathcal{S}}$ defined by:

$$f(c, \kappa) = \sum_{v \to v'} \kappa_{v \to v'} c^v (v' - v), \qquad (2)$$

where c is the (time-dependent) vector of species concentration and c^v is defined as:

$$c^v := \prod_{s \in \mathcal{S}} c(s)^{v(s)}. \qquad (3)$$

Here $c(s)$ is the concentration of species s. The dynamics of $\mathcal{G}$ is described by the equation:

$$f(c, \kappa) = \dot{c}, \qquad (4)$$

where the point indicates the time derivative. Eq. (4) correspond to a set of ordinary differential equations (ODEs) that describe the time evolution of concentration of each species in S. Indeed, it is natural to define a *positive equilibrium* as an element $c \in \mathbb{R}^{\mathcal{S}}_{>0}$ such that $f(c, \kappa) = 0$. A positive equilibrium is then a steady state of the mass-action system characterized by a positive concentration and a zero net production rate for every species of the CRN. Note that a necessary condition for the existence of a positive equilibrium for a mass-action system is the existence of a *positive vector* $\alpha \in \mathbb{R}^{\mathcal{R}}_{>0}$ such that (Feinberg, 1995):

$$\sum_{v \to v'} \alpha_{v \to v'} (v' - v) = 0. \qquad (5)$$

If exists α such that Eq. (5) is satisfied, I say that the network's reaction vectors are positively dependent. Note that weak reversibility is a sufficient (but not necessary) condition for the existence of a positive equilibrium (Boros, 2019; Feinberg, 1995).

The DZT provided a connection between the topology of a network, expressed in terms of its deficiency, and the existence and the uniqueness of steady states for the associated mass-action system (Feinberg, 1987, 1995). In fact, the DZT does much more, and its results can be applied to a broad class of kinetics systems. However, in this work I limit my attention on a specific result (again, see Feinberg (2019) for a comprehensive illustration of the theorem and its application). In particular, I will use the following statement:

Theorem 1 (*The deficiency zero theorem*). *Let* $(\mathcal{S}, \mathcal{N}, \mathcal{R})$ *be a reaction network of deficiency zero.*

> i *If the network is weakly reversible, the resulting mass-action system admits precisely one positive equilibrium for each initial condition, no matter what (positive) rate constants are assigned to the reactions.*
> ii *If the network is not weakly reversible, there can be no assignment of (positive) rate constants such that the resulting mass-action system admits the existence of a positive equilibrium (the network's reaction vectors are not positively dependent).*

2.2. Reflexively autocatalytic and food-generated sets

Within the framework of RAF theory (Hordijk and Steel, 2004; Hordijk et al., 2011), a network of interacting species is represented by a *Catalytic Reaction Systems* (CRS) (Lohn et al., 1998; Hordijk and Steel, 2004; Hordijk and Steel, 2017), that is a tuple (S, R, C, F) such that:

– S is the set of chemical species;
– R is the set of reactions, $\rho \to \pi$, where $\rho, \pi \subset S$ are the *reactants* and *products* of a reaction, respectively;
– C is the *catalysis* set, that is, a set of pairs $\{(s, r) : s \in S, r \in R\}$ indicating the species s as the *catalyst* of reaction r;

– $F \subset S$ is the *food* set, that is a distinguished subset of S such that each species $s \in F$ is assumed to be available from the environment.

Let R' be a subset of R. The *closure* $cl_{R'}(F)$ is defined to be the (unique) minimal subset of S that contains F together with all species that can be produced from F by repeated applications of reactions in R'. Note that $cl_{R'}(F)$ is well defined and finite[1] (Hordijk and Steel, 2004). In RAF theory, the notion of ASs in represented by the RAF sets. Given a CRS (S, R, C, F), a RAF set is a nonempty set of reactions $R' \subseteq R$ (and associated chemical species) that satisfies the following properties:

– Reflexively autocatalytic: for each reaction $r \in R'$ there exists at least one species $s \in cl_{R'}(F)$ such that $(s, r) \in C$;
– F-generated: for each reaction $r : \rho \to \pi, r \in R'$ and for each species $s \in S$ such that $s \in \rho$, it is $s \in cl_{R'}(F)$.

Thus, a RAF set is a set of reactions able to catalytically produce all its reactants starting from a suitable food set. It is also possible to define the closure of a set of reactions, introducing the notion of *closed RAF set* (Smith et al., 2014). Given a CRS (S, R, C, F), a subset R' of R is said to be a closed RAF set if:

– R' is a RAF set;
– for each r such that all its reactants and at least one catalyst are either part of the set F or are produced by a reaction from R', it is $r \in R'$.

Fig. 1 shows an example of a RAF set and its constituent closed RAF sets.

In Smith et al. (2014), authors claimed that "informally, a closed RAF captures the idea that any reaction that can occur, will occur". It has been shown that closed RAF sets have interesting dynamical properties and they may be the relevant units for the potential evolvability of ASs (Hordijk et al., 2018). However, the structure of a closed RAF set seems unable to guarantee its dynamical stability, and monotonic growth of all its associated species (Serra and Villani, 2019; Ravoni, 2020; Dittrich et al., 2007). In Section 4 I introduce further constraints on the topology of closed RAF sets that allow to predict their dynamics and to identify networks with peculiar dynamical properties.

3. Reflexively autocatalytic and food-generated chemical reaction networks

In this Section I study RAF sets in the context of CRN theory. Our first step is to represent F-generated sets in terms of CRNs (Section 3.1). In particular, I focus my attention on zero deficiency CRNs (this restriction is dropped when I introduce catalysis, see Section 3.2). I open the system to the inflow of matter by introducing pseudo reactions that produce species of the food set. I consider both reversible and irreversible reactions and allow only elementary reactions (i.e., reactions involving one or two chemical species as reactants) to occur[2].

---

[1] Note that the finiteness of $cl_{R'}(F)$ follows from the assumption of finiteness of the sets R and S. This assumption, although not binding in the definition of RAF sets, is necessary for finding a RAF set in a generic CRS in a finite time (Hordijk and Steel, 2004).

[2] Reactions involving three or more molecules as reactants are rare; they are often introduced as an approximation of sequences of elementary reactions (Gillespie, 2007). By excluding non-elementary reactions, the model I propose is suitable for being studied through stochastic simulations (Gillespie, 1976, 1977), without the need to introduce further assumptions.
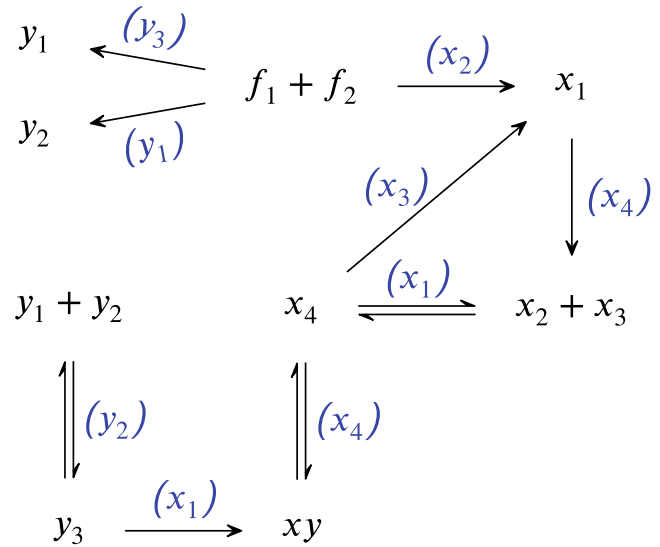

Fig. 1. RAF set example. Complexes are represented by black letters (the letter f indicates food species), reactions by arrows. Blue letters in brackets indicate catalysts. The entire reaction network is a RAF set. Two closed RAF sets are present within the RAF set, namely: $R^{(1)} = \{f_1 + f_2 \to^{(x_2)} x_1, x_1 \to^{(x_4)} x_2 + x_3, x_2 + x_3 \overset{(x_1)}{\rightleftharpoons} x_4, x_4 \to^{(x_3)} x_1, x_4 \overset{(x_4)}{\rightleftharpoons} xy\}$ and $R^{(2)} = \{f_1 + f_2 \to^{(y_3)} y_1, f_1 + f_2 \to^{(y_1)} y_2, y_2 + y_3 \overset{(y_2)}{\rightleftharpoons} y_3\}$.

Finally, in Section 3.2 I introduce the notion of catalysis within the framework of the CRN theory. In RAF theory, a catalyst is a species able, with its presence, to facilitate (i.e. accelerate) the proceeding of a reaction (Hordijk and Steel, 2004). However, in real systems, catalysed reactions usually involve one or more intermediate chemical complexes and can exhibit various dynamical behaviours (Michaelis and Menten, 1913; Ricard and Cornish-Bowden, 1987).

I define catalysis within CRNs by introducing an enzymatic mechanism such that a catalysed reactions consumes a catalyst x and a substrate s to form an intermediate species e, which in turn releases product species and the original catalyst x:

$$s + x \overset{\kappa_1}{\underset{\kappa_{-1}}{\rightleftharpoons}} e \overset{\kappa_2}{\underset{\kappa_{-2}}{\rightleftharpoons}} p + x \tag{6}$$

where $\kappa_{-2} = 0$ for irreversible reactions. Note that this modeling is the common *Michaelis–Menten scheme* (Michaelis and Menten, 1913), extensively used to describe enzymatic processes, and that this modelisation has already been used to simulate RAF sets dynamics (Serra and Villani, 2019). Furthermore, I assume that, due to its highly reactivity, the intermediate species e reaches a quasi-equilibrium state very quickly, compared to the characteristic evolutionary time scale of other species. Therefore, the derivative in Eq. (4) relating to the intermediate species can be set approximately equal to zero:

$$\dot{c}(e) \approx 0. \tag{7}$$

I refer to Eq. (7) as the *quasi steady-state* (QSS) assumption. It is important to underline that the QSS assumption is at the basis of the *quasi steady-state approximations* (QSSAs) (Briggs and Haldane, 1925; Borghans et al., 1996; Tzafriri, 2003; Tzafriri and Edelman, 2004), that are theoretical tools widely used to study enzymatic reactions. In particular, the QSSAs start from the QSS assumption in order to derive algebraic relations among intermediate species and the other reactants, reducing the number of ODEs that describe the evolution of the system. Interested readers can consult, for example, Bersani et al. (2015) to learn more about these approaches.

The applicability of the QSSAs depends on the initial conditions and the parameters describing the system. If their conditions of

validity are not satisfied, the use of these approximations leads to a wrong description of the dynamics of interest (see, for instance, Flach and Schnell (2006, 2007)). However, as shown in Borghans et al. (1996, 2015, 2003, 2004), the so-called *total quasi steady-state approximation* (tQSSA) is able to accurately reproduce the dynamics of enzymatic reactions for a wide range of parameters. In particular, it holds whenever it is (Tzafriri, 2003; Tzafriri and Edelman, 2004):

$$\frac{\kappa_{-1} + \kappa_2}{\kappa_1} \gg 1, \tag{8}$$

$$\frac{\kappa_{-1} + \kappa_2}{\kappa_1} \gg \frac{\kappa_2}{\kappa_1}, \tag{9}$$

$$\frac{\kappa_{-2}}{\kappa_1} \ll 1, \tag{10}$$

Therefore, for these values of the rate constants one can be reasonably confident that the QSS assumption (Eq. (7)) is valid. This assumption allows to provide a relation between the topology of the CRN without catalysis and the corresponding dynamics including enzymatic processes (Appendix A).

3.1. Food-generated chemical reaction networks

Consider a flow reactor kept at constant pressure and temperature, containing a homogeneous and ideal dilute solution of $\mathcal{S}$ interacting chemical species. The system of chemical interactions that occur within the reactor is described by the CRN $\mathcal{G} = (\mathcal{S}, \mathcal{N}, \mathcal{R})$. The reactor allows inflow and outflow of species in the set $F \subset \mathcal{S}$. I consider F to be the food set of $\mathcal{G}$. As in RAF theory, I say that $\mathcal{G} = (\mathcal{S}, \mathcal{N}, \mathcal{R})$ is a *F-generated CRN* if, for a certain set $F \subset S$ and for each $s \in S$, there exists a sequence of reactions $(v'_0 - v_0), (v'_1 - v_1), \ldots, (v'_i - v_i)$ such that:

- $v'_i(s) \neq 0$;
- $v_0(s') \neq 0 \Longleftrightarrow s' \in F$;
- $\forall s' \in S$ such that $v_j(s') \neq 0, j = 1, \ldots, i$ exists $k < j$ such that $v'_k(s') \neq 0$.

Many CRNs describing real systems have zero deficiency. However, constructing "artificial" zero deficiency CRNs is not always an easy task. An example of zero deficiency CRNs are the *elementary* CRNs: given a F-generated CRN $\mathcal{G} = (\mathcal{S}, \mathcal{N}, \mathcal{R})$, I say that $\mathcal{G}$ is elementary if the following condition holds:

$$\forall (v, \mu) \in \mathcal{N} \exists s \in \mathcal{S} | \mu \neq v \Rightarrow v(s) \neq 0, \mu(s) = 0, \tag{11}$$

that is, a F-generated CRN $\mathcal{G}$ is elementary if all its complexes contain at least one species that does not appear in any other complex of the network. In order to show that an elementary F-generated CRN is a zero deficiency network, let $\mathcal{G}'$ be a elementary F-generated CRN having just one linkage class that connects all its n complexes through exactly $n-1$ reactions. From Eq. (11) it follows that the reaction vectors $(v' - v)$ are linear independent. It follows that $d_{\mathcal{G}'} = n - 1$, and it is $\delta_{\mathcal{G}'} = 0$. Let $\mathcal{G}$ be a CRN obtained from $\mathcal{G}'$ by removing exactly m reactions. Note that $\mathcal{G}$ is again elementary. It follows that $l_{\mathcal{G}} = 1 + m, d_{\mathcal{G}} = n - 1 - m$, and $\delta_{\mathcal{G}} = 0$. From the arbitrary choice of m and the equivalence of the deficiency for CRNs having the same complexes and the same linkage classes (Feinberg, 1995), it follows that each CRN satisfying Eq. (11) is a deficiency zero network.

In order to model the flow of food species, I introduce pseudo reactions for each species $f \in F$ such that:

$$0 \underset{o_f}{\overset{\iota_f}{\rightleftharpoons}} f \tag{12}$$

where ι_f and o_f are the constant rates for reactions $0 \to f$ and $f \to 0$, respectively. I denote with $\mathcal{G}_F = (\mathcal{S}_F, \mathcal{N}_F, \mathcal{R}_F)$ the CRN such that

$\mathcal{S}_F = F, \mathcal{N}_F = F \cup 0$ (where 0 is the zero complex) and $\mathcal{R}_F$ is the set of the pseudo reactions. Note that for each pair of reaction vectors $(v'_i - 0), (v'_j - 0) \in \mathcal{R}_F$, for $i \neq j$, we have $(v'_i - 0) \cdot (v'_j - 0) = 0$, where $\cdot$ is the scalar product. Thus, it results $d_{\mathcal{G}_F} = |F|$ and $\delta_{\mathcal{G}_F} = 0$, that is, $\mathcal{G}_F$ is a (weakly) reversible zero deficiency CRN. The DZT states that $\mathcal{G}_F$ admits exactly one steady state. In particular, for each $f \in F$, the positive equilibrium is given by $c(f) = \iota_f / o_f$. Therefore, $c(f)$ is the concentration that the species f would reach if only pseudo reactions are included.

Given a CRN $\mathcal{G}' = (\mathcal{S}', \mathcal{N}', \mathcal{R}')$ and a food set $F \subset \mathcal{S}'$, the equivalent *open CRN* $\mathcal{G}$ of $\mathcal{G}'$ is obtained from $\mathcal{G}'$ by adding $\mathcal{G}_F$, such that it is $\mathcal{G} = (\mathcal{S}, \mathcal{N}, \mathcal{R})$, where $\mathcal{S} = \mathcal{S}', \mathcal{N} = \mathcal{N}' \cup \mathcal{N}_F$ and $\mathcal{R} = \mathcal{R}' \cup \mathcal{R}_F$. Note that, in general, the addition of $\mathcal{G}_F$ to $\mathcal{G}'$ changes the topology of $\mathcal{G}'$; for instance, $\mathcal{G}_F$ can connect separated linkage classes of $\mathcal{G}'$. Moreover, the introduction of various food species affects the dynamical behaviour of the network, both by breaking conservation laws and with the emergence of irreversible cycles (Polettini and Esposito, 2014). In this work, I focus my attention on CRNs having only one complex that includes all the food species used by the F-generated CRN to built itself, that is, given an F-generated CRN $\mathcal{G}' = (\mathcal{S}', \mathcal{N}', \mathcal{R}')$ there exist exactly one complex $v_0 \in \mathcal{N}$ such that, for each $v \in \mathcal{N}$, it is:

$$v(f) \neq 0, f \in F \Rightarrow v = v_0. \tag{13}$$

Eq. (13) implies that, if the F-generated CRN $\mathcal{G}'$ is a deficiency zero network, then the equivalent open CRN $\mathcal{G}$ still has $\delta = 0$. Indeed, let $\delta' = n' - l' - d'$ and $\delta = n - l - d$ be the deficiencies of $\mathcal{G}'$ and $\mathcal{G}$, respectively. It is:

$$n = n' + 1 + |F| - \alpha, \tag{14}$$

$$l = l' + (1 - \alpha), \tag{15}$$

$$d = d' + |F|, \tag{16}$$

where the number 1 in Eq. (14) derives from the introduction of the zero complex, $|F|$ is the number of food species and $\alpha = 1$ if $|F| = 1, \alpha = 0$ if $|F| > 1$. Therefore, it is $\delta = \delta'$. Note that the constrain expressed by Eq. (13) is quite restrictive. The author is currently investigating the impact of its relaxation. Moreover, in the present study I do not allow food species to be catalysts[3], that is, for each $s \in \mathcal{S}$ it is:

$$(s, v \to v') \in C \Rightarrow s \notin F. \tag{17}$$

The constraint expressed by Eq. (17) is not a necessary condition for the results I will show in Section 4 and in Appendix A. However, allowing food elements to act as catalysts can still play an important role in the emerging dynamics. For the clarity of the results of this paper, Eq. (17) is assumed to be valid in all of the CRNs studied in this paper. I will deepen the current study by relaxing this condition in a future work.

3.2. Addition of catalysis

In this Section I introduce a procedure to include catalysis within a CRN. In RAF theory, a catalysis is a pair (x, r) indicating that species x catalyses reaction r. As previously indicated, in the model, I add a catalysis as a substrate-catalyst interaction that produces a *substrate-catalyst species*. In particular, for a monomolecular reaction $s_1 \to p$ catalysed by a molecule x, I assume that the species s_1 reacts with the catalyst x to form the substrate-

---

[3] Note that, by imposing Eq. (17), we are excluding the constructively autocatalytic and food generated (CAF) set from the current study. CAF sets are particular RAF sets for which reactions can take place on the condition that their catalysts are already made available by the occurrence of catalysed reactions, starting from the food set (Mossel and Steel, 2005).
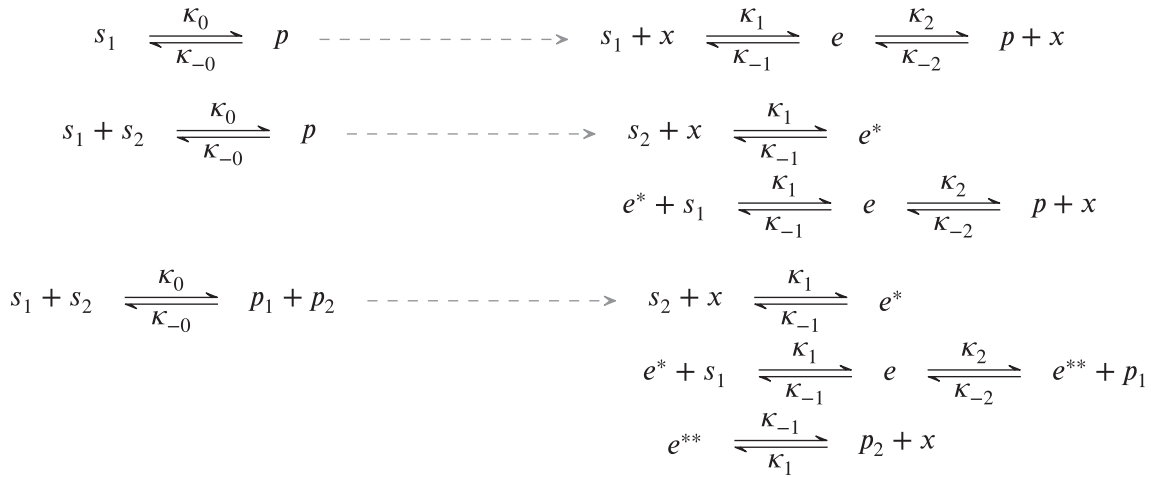
Fig. 2. Catalysis scheme. The dashed gray arrows associate spontaneous reactions (terms on the left) with the corresponding catalysed reactions (terms on the right). Catalysis associated with irreversible reactions can be obtained from the first two catalysis of the scheme by setting κ_{-0} and κ_{-2} equal to zero. Note that for irreversible reactions, p is in turn a complex, i.e., it can be constituted by one or more chemical species. The associated ODEs are given by Eq. (4), where the associated function $f(c, \kappa)$ is described in Appendix A.

catalyst species e, which can perform the inverse reaction or release the product p and the catalyst x. Similarly, for a bimolecular reaction $s_1 + s_2 \rightarrow p$ catalysed by x, I assume that one reactant, says s_2, is the substrate that reacts with x to produce e^*, while the other reactant, s_2, reacts with e^* producing e, that eventually releases x and products p. Fig. 2 sketched the procedure.

Note that the substrate-catalyst species e, e^*, e^{**} (Fig. 2) are considered as different complexes, that is:

$$(e_i \cdot e_j), (e_i^* \cdot e_j^*), (e_i^{**} \cdot e_j^{**}) \neq 0 \Longleftrightarrow i = j \tag{18}$$

$$e_i \cdot e_j^* = e_i \cdot e_k^{**} = e_j^* \cdot e_k^{**} = 0 \ \forall i, j, k. \tag{19}$$

Moreover, it is important to underline that original spontaneous reactions are still present in the network with addition of catalysis. Given a CRN $\mathcal{G}$ and a set of catalysis C, I call $\mathcal{G}_C$ the CRN obtained from $\mathcal{G}$ by adding reactions according to the scheme of Fig. 2 for each catalysis in C. Therefore, $\mathcal{G}_C$ represents the "real" mass-action system underlying a CRN to which a set of catalysis is associated.

As in RAF theory, given a F-generated CRN $\mathcal{G} = (\mathcal{S}, \mathcal{N}, \mathcal{R})$ and a set of catalysis $C = \{(s, v \rightarrow v') : s \in \mathcal{S}, v \rightarrow v' \in \mathcal{R}\}$, I say that $\mathcal{G}$ is a *RAF CRN* if, for each reaction $v \rightarrow v' \in \mathcal{R}$, there exists at least one species $x \in \mathcal{S}$ such that $(x, v \rightarrow v') \in C$; I say that a RAF CRN is a *closed RAF CRN* if, for each reaction $v \rightarrow v'$ such that all its reactants and at least one catalyst are either part of the set F or are produced by a reaction from $\mathcal{R}$, it $v \rightarrow v' \in \mathcal{R}$; finally, I say that a closed RAF CRN is a *minimal closed RAF CRN* if it does not contain any other closed RAF CRNs.

4. Long-term behaviour

In this Section, I identify motifs in the toplogy of a minimal closed RAF CRN $\mathcal{G}$ that allow to predict the dynamics of its associated network $\mathcal{G}_C$ (note that, although not specified, I am considering open CRNs). In general, adding catalysis increases the deficiency of a CRN (Polettini et al., 2015), and higher deficiency theories are necessary in order to study the steady states of the network (Feinberg, 1995; Ellison, 1998).

However, the topology of the zero deficiency CRN $\mathcal{G}$ provides some information on the dynamics of $\mathcal{G}_C$, at least if Eq. (7) holds (see Appendix A for more details). In fact, the relation among complexes provided by $\mathcal{G}$ does not change in $\mathcal{G}_C$: the terms $\alpha_{v \rightarrow v'}(v - v')$ in the sum of Eq. (4) for the two networks, differ only in the values of $\alpha_{v \rightarrow v'}$, with $\alpha_{\mathcal{G}_C} > \alpha_{\mathcal{G}}$ (Appendix A). This implies, for instance, that

if $\mathcal{G}$ is not weakly reversible, thanks to the DZT it can be say that the reaction vectors of $\mathcal{G}_C$ are not positively dependent; on the other hand, if $\mathcal{G}$ is weakly reversible, $\mathcal{G}_C$ admits the existence of a positive equilibrium. Nevertheless, it is worth to underline that $\mathcal{G}$ and $\mathcal{G}_C$ are not equivalent. For instance, consider the CRN shown in Fig. 3a. The network without catalysis $\mathcal{G}$ is a weakly reversible zero deficiency network, and it admits exactly one positive equilibrium for each choice of initial condition. Conversely, by analysing the associated CRN $\mathcal{G}_C$ through the *CRN Toolbox* (Ellison et al., 2011), I find that $\mathcal{G}_C$ admits the existence of multiple stationary solutions starting from the same initial conditions, for an appropriate choice of the constant rates κ.

I use these results in order to predict the dynamics of the mass-action system $\mathcal{G}_C$ associated with a RAF CRN $\mathcal{G}$. In particular, let $\mathcal{G}$ be a minimal closed RAF CRN, and let $\mathcal{G}_-$ be the CRN obtained from $\mathcal{G}$ by removing all the reactions $v_0 \rightarrow v'$, where the complex v_0 contains (only) food species. I define:

i. *Fully connected* (FC) motif: $\mathcal{G}$ is a FC motif if it is a weakly reversible CRN;
ii. *Non connected* (NC) motif: $\mathcal{G}$ is a NC motif if both $\mathcal{G}$ and $\mathcal{G}_-$ are not weakly reversible CRNs;
iii. *Core connected* (CC) motif: $\mathcal{G}$ is a CC motif if it is not a weakly reversible CRN and $\mathcal{G}_-$ is a weakly reversible CRN.

Fig. 3 shows examples of the three motifs. They exhibit dynamics with different characteristics (Fig. 4a). In particular, I focus my attention on three dynamical properties of the concentrations of species that constitute a motif:

1. *homeostasis*: the ratio among concentrations of all the species is constant over time;
2. *continuous growth*: the concentrations of all the species increase in time;
3. *self-conservation*: if the inflow of matter towards the CRN is interrupted[4], the concentrations of all the species do not decrease in time.

---

[4] Note that by interrupting the inflow of species in F, the system is no longer a proper flow reactor. However, assumption $\iota_f = 0, \forall f \in F$ is useful to represent, for example, a protocell with a semipermeable membrane that allows the outflow of species from the food set to an environment where the concentration of these species is very low, i.e., in the case where the protocell is in a resource-free environment.
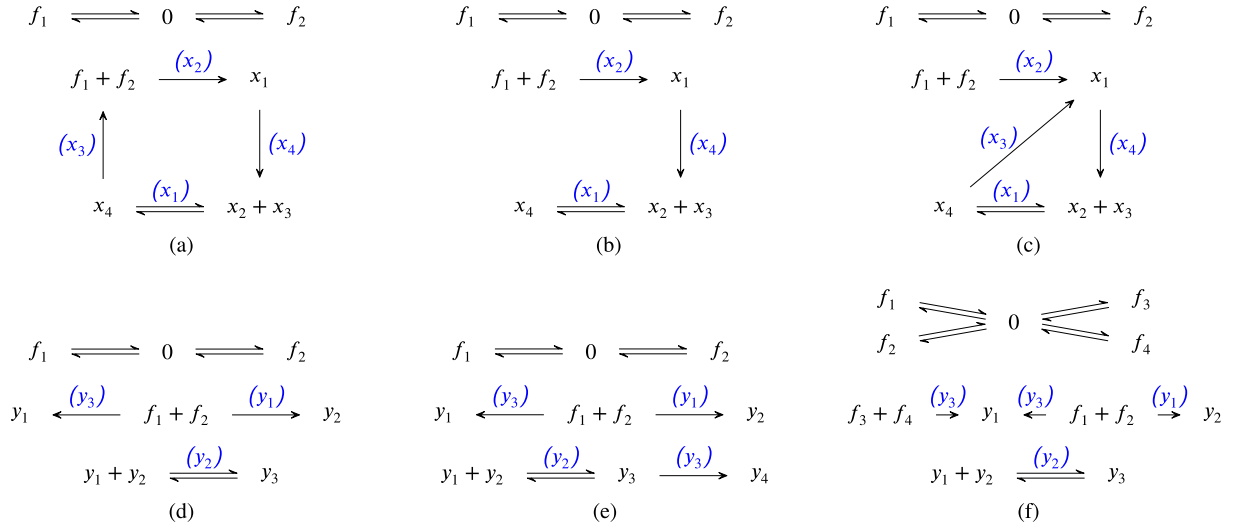
Fig. 3. Examples of motifs. Complexes are represented by black letters (the letter f indicates food species), reactions by arrows. Blue letters in brackets indicate catalysts. (a) FC motif: the CRN is weakly reversible. (b,e) NC motif: the CRN is not weakly reversible, and there is not a weakly reversible (sub) network that includes all the non-food species. (c,d,f) CC motif: the CRN is not weakly reversible, and there is a weakly reversible (sub) network that includes all the non-food species. Note that all the showed motifs are minimal closed RAF CRNs.
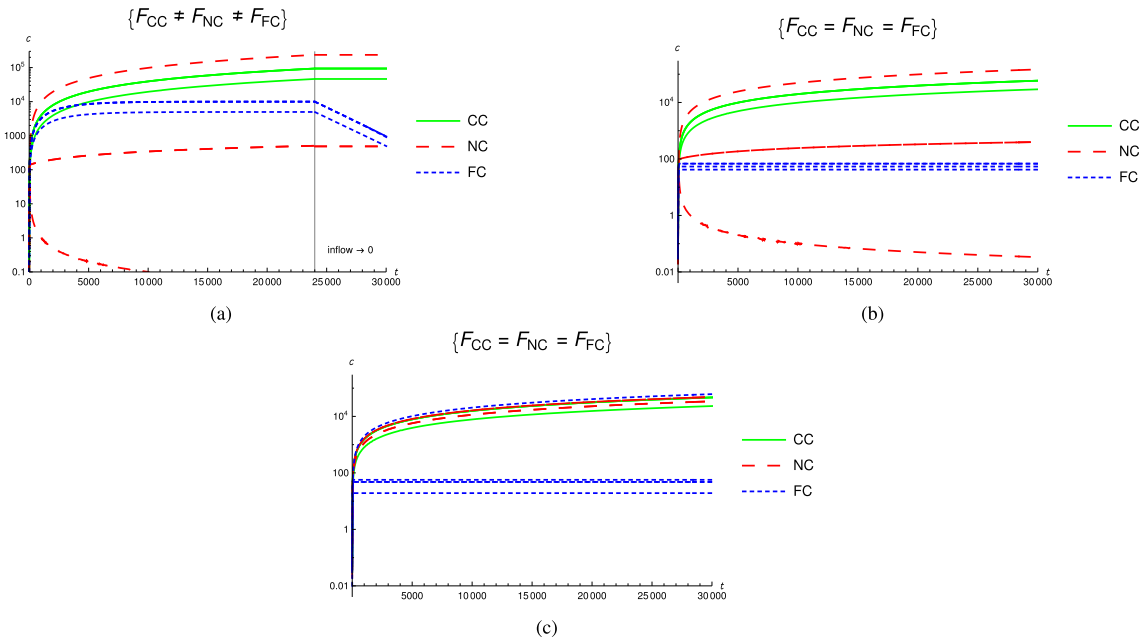

Fig. 4. Dynamics of motifs. On the x-axis: time steps t. On the y-axis: species concentration c. (a) Dynamics of the three isolated motifs introduced in Fig. 3a, Fig. 3b and Fig. 3c. For $t > 24000$ the inflow rate ι_f is set equal to 0 for each pseudo reaction $0 \to^{\iota_f} f$. (b) Dynamics of the three motifs introduced in Fig. 3a, Fig. 3b and Fig. 3c, that draw from the same food set $F = \{f_1, f_2\}$. (c) Resulting dynamics after mutations (see Section 4.1) that change the FC motif of Fig. 3a in a NC motif, and the NC motif of Fig. 3b in a CC motif. Different curves of the same colour correspond to different species belonging to the same motif. The plotted curves are obtained by numerically solving the ODEs associated with the CRNs (Eq. (4)), setting $\iota_f = 10, o_f = 10^{-2}, \kappa_0 = \kappa_{-0} = 10^{-3}, \kappa_1 = 10, \kappa_{-1} = 10^7, \kappa_2 = 10^4, \kappa_{-2} = 10^{-1}$, and an initial concentration of $c_0 = 10^{-2}$ for all the species.

Note that these are desired features for a system at the basis of the origin of life (Luisi, 2016; Varela, 2000). In fact, homeostasis is considered to be a fundamental property connected with self-maintenance of an organism. On the other hand, the growth of species concentration is necessary for self-reproduction (an example is the interphase of the cell cycle that precedes cell division). Finally, self-conservation prevents the decay of the system in case of insufficient resources, in analogy with the dormancy observed in many microorganisms (Rittershaus et al., 2013).

I argue that FC motif exhibits homeostasis, NC motif exhibits none of the properties and CC motif exhibits continuous growth and self-conservation. It is important to underline that I am referring to long-term dynamics of the networks and not to their transient behaviour.

In fact, FC motif admits the existence of a positive equilibrium for the concentration of all its chemical species, meaning that, for $t \to \infty$, the ratio $c(s)/c(s')$ is constant for each pair $s, s' \in \mathcal{S}$ (Fig. 4a). However, if the inflow rate ι_f is set equal to zero for each $f \in F$, the resulting network is no longer weakly-reversible. In par-

ticular, there is a terminal strong-linkage class (i.e., the zero complex) that is an *absorbing linkage class*, meaning that, once the flow of matter falls in the strong-linkage class, it can no longer leave. Therefore, the matter consumed by reactions involving the species of the CRN is not replaced, the species concentrations decrease and the network is not self-conservative (Fig. 4a).

On the other hand, NC motif contains an absorbing linkage class involving a strict subset $\mathcal{S}'$ of $\mathcal{S}$, both in case of $\iota_f > 0$ and $\iota_f = 0$. As a result, species in $\mathcal{S}'$ increase their concentrations at the expense of other species $s \notin \mathcal{S}'$. Thus, even if a subset of $\mathcal{S}$ can experience continuous growth, the whole NC motif does not satisfy any of the properties listed above (Fig. 4a).

In contrast, all the non-food species of a CC motif belong to a weakly reversible (sub) network that is absorbing for the inflowing matter coming from the environment. This implies that, if $\iota_f > 0$, each non-food species is reached by a positive flow of matter (continuous growth); on the other hand, if $\iota_f = 0$, the matter is redistributed among non-food species and does not escape from the CRN (self-conservation). Note that a CC motif could exhibit homeostasis, depending on the fine details of its topology.

4.1. An illustrative example of evolvability of RAF sets

In this Section I present illustrative examples on how the detected motifs (Section 4) impact the long-term dynamics of a (larger) RAF CRN. I start the study by simulating the dynamics of a RAF CRN constituted by three different closed RAF CRNs that share the same food set *F*. The closed RAF CRNs correspond to the different motifs I have introduced in Fig. 3.

Fig. 4b shows the dynamics of this particular network: as expected, the resulting dynamics is substantially equivalent to that shown by the non composed motifs (Fig. 4a, $\iota_f > 0$). In fact, letting the motifs share the food set does not affect their structures, nor does significantly change the ODEs describing the dynamics (apart from a change in the equilibrium constants of FC motif). Therefore,

the overall dynamics of the network is given by the union of the dynamics of the single motifs.

Let us now consider minimal changes in topology of the networks. Given a CRN $\mathcal{G} = (\mathcal{S}, \mathcal{N}, \mathcal{R})$, I refer to the addition of a reaction $r_j \notin \mathcal{R}$ that produces the new network $\mathcal{G}_j = (\mathcal{S}_j, \mathcal{N}_j, \mathcal{R}_j)$ with the term *mutation*. Mutations can have a major role in the resulting dynamics of the network. For instance, adding the reaction $x_4 \rightarrow^{(x_1)} x_5$ to the FC motif of Fig. 3a changes it into a NC motif. At the same time, adding the reaction $x_4 \rightarrow^{(x_4)} x_1$ to the NC motif of Fig. 3b changes it into a CC motif. The dynamics of the overall network, in fact, is the result of the evolution of a FC motif and two CC motifs (Fig. 4c). This is an example of a general case: it is always possible to pass from one motif to another by adding (or removing) one or more reactions.

This last point has important consequences for the evolvability of RAF sets. In fact, the characteristic dynamical behaviour of a motif can be interpreted as its phenotype. Furthermore, it is possible to associate the appearance of a mutation with the result of a certain stochastic process (as the occurrence of a rare spontaneous reaction). In this framework, the examples of Fig. 4b and Fig. 4c show that closed RAF CRNs can acquire (or lose) dynamical properties that can be, potentially, picked by natural selection. As further evidence to support this observation, in Fig. 5 I show different frames of an hypothetical evolution of a RAF CRN in which mutations occur. In particular, the starting network is formed by two closed RAF CRNs that share the same food set $F = \{f_1, f_2\}$. The closed RAF CRNs are indicated with letters *X* (CC motif shown in Fig. 3c) and *Y* (CC motif shown in Fig. 3d), respectively. Initially, both the closed RAF CRNs show a continuous growth of species concentrations, with similar growing rates (Fig. 5a). However, the addition of a new reaction in the network allow *Y* to drawn resources from a supplementary source compared to *X*, represented by the food elements f_3 and f_4 (Fig. 3f). Consequently, this favorable mutation provides an advantage for *Y*, which is able to grow faster than *Y* (Fig. 5b). At the same time, as previously
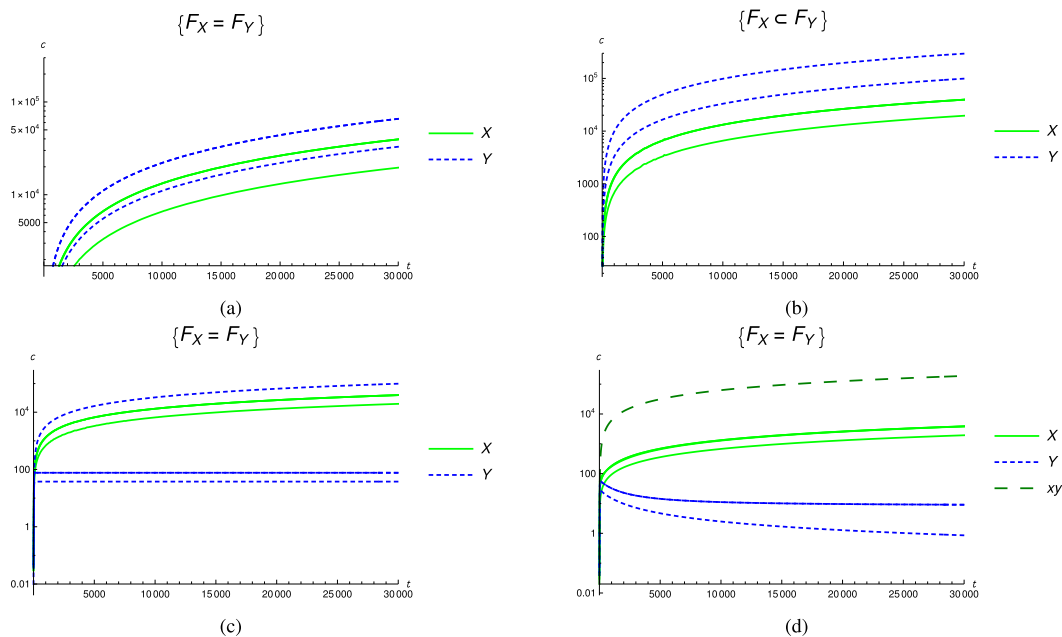

Fig. 5. Evolution of a RAF CRN induced by mutations. On the *x*-axis: time steps *t*. On the *y*-axis: species concentration *c*. (a) Dynamics of a RAF CRN constituted by the CC motif of Fig. 3c (*X*) and the CC motif of Fig. 3d (*Y*), that draw from the same food set $F = \{f_1, f_2\}$. (b) A favorable mutation allows *Y* to draws resources from the expanded set $\{f_1, f_2, f_3, f_4\}$ (Fig. 3f), resulting in a higher self-production rate. (c) A unfavorable mutation changes species *Y* in a NC motif (Fig. 3e), preventing from its continuous growth. (d) A mutation allows *X* to steal matter from *Y*, causing the dynamical disappearance of *Y* (the CRN associated with this dynamics is shown in Fig. 1). Different curves of the same colour correspond to different species belonging to the same motif. The plotted curves are obtained by numerically solving the ODEs associated with the CRNs (Eq. (4)), setting $\iota_f = 10, o_f = 10^{-2}, \kappa_0 = \kappa_{-0} = 10^{-3}, \kappa_1 = 10, \kappa_{-1} = 10^7, \kappa_2 = 10^4, \kappa_{-2} = 10^{-1}$, and an initial concentration of $c_0 = 10^{-2}$ for all the species.

observed, an unfavorable mutation (Fig. 3e) can decrease the efficiency of a network's self-reproduction (Fig. 5c). Finally, note that a mutation introducing a connection between X and Y can have major effects on their dynamics. In particular, consider the RAF CRN shown in Fig. 1: it can be built starting from the networks of Fig. 3c and Fig. 3d, by adding the production of the species xy. This latter expansion of the network can be interpreted, for instance, as the result of a mutation that allows species x_4 to produce species xy in a reversible way, and species x_1 to catalyse the irreversible production of species xy starting from species y_3. Therefore, thanks to a favorable mutation, one motif can steal resources from the other. The dynamics of the resulting RAF CRN is shown in Fig. 5d: due to the mutation, the closed RAF CRN X' (the symbol X' indicates the network X to which the reaction $x_4 \rightarrow^{(x_4)} xy$ is added) is an absorbing linkage class in which the flow of matter of the net Y falls; consequently, network X still shows a continuous growth in species concentrations, while Y dynamically disappears.

In this Section, I have shown how the presence of motifs FC, NC and CC within the structure of a RAF CRN affects the dynamics of the system. Moreover, the examples I have introduced suggest that closed RAF CRNs could be the elementary units that experience adaptive evolution. Regarding this issue, it is important to underline two points. First, note that, in this work, I am dealing with a deterministic model. This approximation does not allow to directly simulate neither the spontaneous appearance of novelty, nor the probabilistic effects, that are characteristics of systems with few elements. However, recent numerical simulations have shown how these stochastic processes can determine both a dependence on different initial conditions (for instance, deriving from an inhomogeneous distribution of resources, or from random events, such as the division of a protocell), and an actual competition among closed RAF CRNs (Hordijk et al., 2018; Serra and Villani, 2019; Ravoni, 2020). Therefore, the addition of stochastic elements to the RAF CRN model introduced in this work should reproduce the scenario proposed in (Vasas et al., 2012; Hordijk et al., 2018; Ravoni, 2020) for the evolvability of a population of compartmentalised RAF sets.

Moreover, it is very important to note that these results show that a RAF CRN does not always fully emerge in the long-term dynamics, and that an actual competition can therefore occur also among subsets of the same RAF network. Regarding this point, in the future it would be interesting to study the dynamics of various RAF sets contained in a larger CRN, within which chemical species of the RAF sets can also participate in reactions that are not members of the RAF sets themselves. In particular, it would be useful to investigate how the dynamics of the introduced motifs respond to the presence of side reactions, which have been shown to be able to prevent the realization of self-replicating chemical systems (Szathmáry, 2000; Liu et al., 2020).

I am currently investigating the impact of some of the assumptions made in this work, such as the QSS assumption or the limited role of the food set in the network reactions, and will discuss all these aspects in an upcoming work.

5. Conclusions

In this work, I study long-term behaviours of RAF sets, particularly focusing on those systems that are considered to be the building blocks from which life emerged.

For this purpose, I represent RAF sets in terms of CRNs, and I use CRN theory to predict the dynamics of the networks starting from their topology. The main result of this paper is the identification of additional topological conditions compared to the one introduced by the RAF theory, which allow to predict the dynamics of the sys-

tem and to identify networks with interesting dynamical properties. To my knowledge, this is the first time that the actual dynamics of RAF sets have been successfully connected with their topology.

Furthermore, starting from their structure, I display that not all the subsets of a RAF network eventually dynamically emerge, contrary to what has been observed in previous works, where overly simple CRNs were considered (Hordijk et al., 2018; Ravoni, 2020). These results show that the increase of complexity of a RAF set significantly affects the emergent dynamics, supporting the potential evolvability of RAF sets and their importance in the transition from inanimate to living matter. A possible next step could be to explore the introduced motifs in real metabolic networks.

Thermodynamic aspects of these systems are going to be investigated in a forthcoming paper, allowing to formalise a consistent theory of autocatalytic networks.

CRediT authorship contribution statement

Alessandro Ravoni: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing.

Appendix A

Let $\mathcal{G} = (\mathcal{S}, \mathcal{N}, \mathcal{R})$ be a CRN, $\mathcal{R}^+$ and $\mathcal{R}^-$ be the sets of reversible and irreversible reactions (such that $\mathcal{R} = \mathcal{R}^+ \cup \mathcal{R}^-$), and i, j denote reactions $v \rightarrow v' \in \mathcal{R}$. The function $f(c, \kappa)$ for $\mathcal{G}$ can be written as:

$$f_{\mathcal{G}}(c, \kappa) = \sum_{i \in \mathcal{R}^+} \kappa_i c^{v_i}(v_i' - v_i) + \sum_{j \in \mathcal{R}^-} \kappa_j c^{v_j}(v_j' - v_j). \tag{20}$$

Similarly, for $\mathcal{G}_c$ it is:

$$f_{\mathcal{G}_c}(c, \kappa) = \sum_{i \in \mathcal{R}^+} [\kappa_i c^{v_i}(v_i' - v_i) + P_i] + \sum_{j \in \mathcal{R}^-} [\kappa_j c^{v_j}(v_j' - v_j) + I_j], \tag{21}$$

where P_i and I_j denote the additional terms introduced by adding catalysis according to the scheme of Fig. 2 for reversible and irreversible reactions, respectively. For each pair (i, i') of reversible reactions such that $i : v \rightarrow v'$ and $i' : v' \rightarrow v$, the term $P_{(i,i')}$ can be written as:

$$
\begin{aligned}
P_{(i,i')} = & |s2|[\kappa_1 c(s2)c(x) - \kappa_{-1}c(e^*)] \times (e^* - s2 - x) \\
& + \{\kappa_1 c^{(v-s2)}[|s2|c(e^*) + (1 - |s2|)c(x)] - \kappa_{-1}c(e)\}[e - v \\
& + s2 - |s2|e^* - (1 - |s2|)x] + \{\kappa_2 c(e) \\
& - \kappa_{-2}c^{(v'-p2)}[|p2|c(e^{**}) + (1 - |p2|)c(x)]\}[v' - p2 \\
& + |p2|e^{**} + (1 - |p2|)x - e] + |p2|[\kappa_1 c(p2)c(x) \\
& - \kappa_{-1}c(e^{**})](e^{**} - p2 - x).
\end{aligned} \tag{22}
$$

Here, $s2$ and $p2$ are the vectors having all components equal to zero, except one (corresponding to species $s2$ and $p2$, respectively). Note that due to the above definition, $P_{(i,i')}$ describes the terms to be added for all the catalysis shown in Fig. 2, depending on whether the relations $s2 = 0$ and $p2 = 0$ are valid or not, where 0 is the zero complex.

Similarly, for each irreversible reaction $j : v \rightarrow v'$ the term I_j can be written as:

$$
\begin{aligned}
I_{(j)} = & |s2|[\kappa_1 c(s2)c(x) - \kappa_{-1}c(e^*)] \times (e^* - s2 - x) \\
& + \{\kappa_1 c^{(v-s2)}[|s2|c(e^*) + (1 - |s2|)c(x)] - \kappa_{-1}c(e)\}[e - v \\
& + s2 - |s2|e^* - (1 - |s2|)x] + [\kappa_2 c(e)](v' + x - e).
\end{aligned} \tag{23}
$$

Note that I_j can be obtained from $P_{(j,j')}$ assuming $\kappa_{-2} = 0$ and $p2 = 0$. From Eq. (22) it results:

$$\dot{c}(e) = \{\kappa_1 c^{(v-s2)}[|s2|c(e^*) + (1 - |s2|)c(x)] - \kappa_{-1}c(e) + \kappa_2 c(e) \\ - \kappa_{-2}c^{(v'-p2)}[|p2|c(e^{**}) + (1 - |p2|)c(x)]\}, \tag{24}$$

$$\dot{c}(e^*) = |s2|[\kappa_1 c(s2)c(x) - \kappa_{-1}c(e^*) - \kappa_1 c^{(v-s2)}c(e^*) \\ + \kappa_{-1}c(e)], \tag{25}$$

$$\dot{c}(e^{**}) = |p2|[\kappa_1 c(p2)c(x) - \kappa_{-1}c(e^{**}) - \kappa_{-2}c^{(v'-p2)}c(e^{**}) \\ + \kappa_2 c(e)]. \tag{26}$$

By imposing $\dot{e} = 0, \dot{e}^* = 0$ and $\dot{e}^{**} = 0$ (quasi steady-state assumption, see Section 3), it follows:

$$\kappa_2 c(e) - \kappa_{-2}c^{(v'-p2)}c(e^{**}) = \kappa_1 c^{(v-s2)}c(e^*) - \kappa_{-1}c(e). \tag{27}$$

Again, equivalent expressions for irreversible reactions are obtained assuming $\kappa_{-2} = 0, p2 = 0$ and (or) $s2 = 0$. Note that the equivalences of Eq. (27) imply that, if for $f_{\mathcal{G}}$ it is:

$$f_{\mathcal{G}} = \sum_{v \to v'} \alpha_{v \to v'}(v' - v), \tag{28}$$

then $f_{\mathcal{G}_C}$ can be written as:

$$f_{\mathcal{G}_C} = \sum_{v \to v'} \alpha'_{v \to v'}(v' - v), \tag{29}$$

where:

$$\alpha'_{v \to v'} = \alpha_{v \to v'} + \Delta_{v \to v'}, \tag{30}$$

with $\Delta_{v \to v'} \geq 0$. For instance, if $\kappa_{-2} \neq 0, p2 \neq 0$ and $s2 \neq 0$ it is:

$$\Delta_{v \to v'} = \kappa_2 c(e) \tag{31}$$

for a forward reaction (that is, a reaction with a "spontaneous" rate constant κ_0) and:

$$\Delta_{v \to v'} = \kappa_{-2}c^{(v'-p2)}c(e^{**}) \tag{32}$$

for a reverse reaction (that is, a reaction with a "spontaneous" rate constant κ_{-0}). This implies, in particular, that $\mathcal{G}_C$ admits a positive equilibrium only if $\mathcal{G}$ is weakly reversible. An equivalent result holds if $\kappa_{-2} = 0, p2 = 0$ and (or) $s2 = 0$.

References

Bersani, Alberto M., Bersani, Enrico, Dell' Acqua, Guido, Pedersen, Morten G., 2015. New trends and perspectives in nonlinear intracellular dynamics: one century from michaelis-menten paper. Continuum Mech. Thermodyn. 27 (4–5), 659–684.

Borghans, José A.M., Boer, De, Rob, J., Segel, Lee A., 1996. Extending the quasi-steady state approximation by changing variables. Bull. Math. Biol. 58 (1), 43–63.

Boros, Balázs, 2019. Existence of positive steady states for weakly reversible mass action systems. SIAM J. Math. Anal. 51 (1), 435–449.

Briggs, George E., Haldane, John B.S., 1925. A note on the kinetics of enzyme action. Biochem. J. 19 (2), 338–339.

Dittrich, Peter, Fenizio, Di, Pietro, S., 2007. Chemical organisation theory. Bull. Math. Biol. 69 (4), 1199–1231.

Dyson, Freeman, 1985. Origins of Life. Cambridge University Press, Cambridge.

Ellison, Philipp R., 1998. The Advanced Deficiency Algorithm and its applications to mechanism discrimination, Ph.D. thesis, Department of Chemical Engineering, University of Rochester..

Feinberg, Martin, 1987. Chemical reaction network structure andthe stability of complex isothermal reactors I: the deficiency zero and deficiency one theorems. Chem. Engr. Sci. 42 (10), 2229–2268.

Feinberg, Martin, 1995. The existence and uniqueness of steady states for a class of chemical reaction networks. Arch. Ration. Mech. Anal. 132 (4), 311–370.

Feinberg, Martin, 2019. Foundations of Chemical Reaction Network Theory. Springer.

Filisetti, Alessandro and Graudenzi, Alex and Serra, Roberto and Villani, Marco and De Lucrezia, Davide and Füchslin, Rudolf M. and Kauffman, Stuart A. and Packard, Norman and Poli, Irene, 2011. A stochastic model of the emergence of autocatalytic cycles. J. Syst. Chem. 2(1): 2..

Flach, Edward H., Schnell, Santiago, 2006. Use and abuse of the quasi-steady-state approximation. IEE Proceedings - Systems Biology 153 (4), 187.

Gillespie, Daniel T., 1976. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J. Comput. Phys. 22 (4), 403–434.

Gillespie, Daniel T., 1977. Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem. 81 (25), 2340–2361.

Gillespie, Daniel T., 2007. Stochastic simulation of chemical kinetics. Annu. Rev. Phys. Chem. 58 (1), 35–55.

Higgs, Paul G., Lehman, Niles, 2015. The RNA World: molecular cooperation at the origins of life. Nat. Rev. Genet. 16 (1), 7–17.

Hordijk, Wim, Steel, Mike, 2004. Detecting autocatalytic, self-sustaining sets in chemical reaction systems. J. Theor. Biol. 227 (4), 451–461.

Hordijk, Wim, Hein, Jotun, Steel, Mike, 2010. Autocatalytic Sets and the Origin of Life. Entropy 12 (17), 1733–1742.

Hordijk, Wim, Kauffman, Stuart A., Steel, Mike, 2011. Required levels of catalysis for emergence of autocatalytic sets in models of chemical reaction systems. Int. J. Mol. Sci. 12 (5), 3085–3101.

Hordijk, Wim, Steel, Mike, 2012. Predicting template-based catalysis rates in a simple catalytic reaction model. J. Theor. Biol. 295, 132–138.

Hordijk, Wim, Steel, Mike, 2012. Autocatalytic sets extended: Dynamics, inhibition, and a generalization. J. Syst. Chem. 3 (1), 5.

Hordijk, Wim, Steel, Mike, Kauffman, Stuart A., 2012. The structure of autocatalytic sets: Evolvability, enablement, and emergence. Acta Biotheor. 60 (4), 379–392.

Hordijk, Wim, Steel, Mike, 2014. Conditions for evolvability of autocatalytic sets: A formal example and analysis. Origins Life Evol. Biospheres 44 (2), 111–124.

Hordijk, Wim, Steel, Mike, 2017. Chasing the tail: The emergence of autocatalytic networks. Biosystems 152, 1–10.

Hordijk, Wim, Steel, Mike, Dittrich, Peter, 2018. Autocatalytic sets and chemical organizations: Modeling self-sustaining reaction networks at the origin of life. New J. Phys. 20, (1) 015011.

Hordijk, Wim, Naylor, Jonathan, Krasnogor, Natalio, Fellermann, Harold, 2018. Population dynamics of autocatalytic sets in a compartmentalized spatial world. Life 8 (3), 33.

Hordijk, Wim, Steel, Mike, 2018. Autocatalytic networks at the basis of life's origin and organization. Life 8 (4), 62.

Ellison, Philipp R. and Feinberg, Martin and Haixia, Ji and Knight, Daniel, 2011. The Chemical Reaction Network Toolbox, Ver. 2.x, http://www.chbmeng.ohio-state.edu/ijfeinberg/crntwin/..

Kauffman, Stuart A., 1971. Cellular homeostasis, epigenesis and replication in randomly aggregated macromolecular systems. J. Cybern. 1 (1), 71–96.

Kauffman, Stuart A., 1986. Autocatalytic sets of proteins. J. Theor. Biol. 119 (1), 1–24.

Kauffman, Stuart A., 1993. The origins of order: Self-organization and selection in evolution. Oxford University Press, USA.

Kauffman, Stuart A., 2011. Approaches to the origin of life on earth. Life 1 (1), 34–48.

Lifson, Shneior, 1997. On the Crucial Stages in the Origin of Animate Matter. J. Mol. Evol. 44 (1), 1–8.

Liu, Yu., Hjerpe, Daniel, Lundh, Torbjörn, 2020. Side Reactions Do Not Completely Disrupt Linear Self-Replicating Chemical Reaction Systems. Artif. Life 26 (3), 327–337.

Lohn, Jason D., Colombano, Silvano P., Scargle, Jeffrey, Stassinopoulos, Dimitris, Haith, Gary L., 1998. Evolving catalytic reaction sets using genetic algorithms. In: Proceedings of the 1998 IEEE InternationalConference on Evolutionary Computation, Anchorage, AK, USA, pp. 487–492.

Luisi, Pier Luigi, 2016. The emergence of life: from chemical origins to synthetic biology. Cambridge University Press.

Michaelis, Leonor, Menten, Maud.L., 1913. Die kinetik der invertinwirkung. Biochem. Z. 49, 333–369.

Mossel, Elchanan, Steel, Mike, 2005. Random biochemical networks: The probability of self-sustaining autocatalysis. J. Theor. Biol. 233 (3), 327–336.

Nghe, Philippe, Hordijk, Wim, Kauffman, Stuart A., Walker, Sara I., Schmidt, Francis J., Kemble, Harry, Yeates, Jessica A.M., Lehman, Niles, 2015. Prebiotic network evolution: six key parameters. Mol. BioSyst. 11 (12), 3206–3217.

Pedersen, Morten G., Bersani, Alberto M., Bersani, Enrico, 2007. Quasi steady-state approximations in complex intracellular signal transduction networks - a word of caution. J. Math. Chem. 43 (4), 1318–1344.

Polettini, Matteo, Esposito, Massimiliano, 2014. Irreversible Thermodynamics of Open Chemical Networks. I. Emergent Cycles and Broken Conservation Laws. J. Chem. Phys. 141, (2) 024117.

Polettini, Matteo, Wachtel, Artur, Esposito, Massimiliano, 2015. Dissipation in Noisy Chemical Networks: The Role of Deficiency. J. Chem. Phys. 143, (18) 184103.

Rasmussen, Steen, Constantinescu, Adi, Svaneborg, Carsten, 2016. Generating minimal living systems from non-living materials and increasing their evolutionary abilities. Phil. Trans. R. Soc. B: Biol. Sci. 371 (1701), 20150440.

Ravoni, Alessandro, 2020. Impact of composition on the dynamics of autocatalytic sets. Biosystems 198, 104250.

Ricard, Jacques, Cornish-Bowden, Athel, 1987. Co-operative and allosteric enzymes: 20 years on. Eur. J. Biochem. 1987 (166), 255–272.

Rittershaus, Emily S.C., Baek, Seung-Hun, Sassetti, Christopher M., 2013. The Normalcy of Dormancy: Common Themes in Microbial Quiescence. Cell Host Microbe 13 (6), 643–651.

Serra, Roberto, Villani, Marco, 2019. Sustainable growth and synchronization in protocell models. Life 9 (3), 68.

Smith, Joshua I., Steel, Mike, Hordijk, Wim, 2014. Autocatalytic sets in a partitioned biochemical network. J. Syst. Chem. 5 (1), 2.

Sousa, Filipa L., Hordijk, Wim, Steel, Mike, Martin, William F., 2015. Autocatalytic sets in E. coli metabolism. J. Syst. Chem. 6 (1), 4.

Szathmáry, Eörs, 2000. The evolution of replicators, Philosophical. Trans. Roy. Soc. Lond, Ser. B: Biol. Sci. 355 (1403), 1669–1676.

Tzafriri, Abraham R., 2003. Michaelis-menten enzyme kinetics at high enzyme concentrations. Bull. Math. Biol. 65 (6), 1111–1129.

Tzafriri, Abraham R., Edelman, Elazer R., 2004. The total quasi-steady-state approximation is valid for reversible enzyme kinetics. J. Theor. Biol. 226 (3), 303–313.

Varela, Francisco J., 2000. El fenómeno de la vida. Dolmen Esayo, Santiago, Chile.

Vasas, Vera, Szathmáry, Eörs, Santos, Mauro, 2010. Lack of evolvability in self-sustaining autocatalytic networks constraints metabolism-first scenarios for the origin of life. Proc. Nat. Acad. Sci. 107 (4), 1470–1475.

Vasas, Vera, Fernando, Chrisantha, Santos, Mauro, Kauffman, Stuart A., Szathmáry, Eörs, 2012. Evolution before genes. Biol. Direct 7 (1), 1.

Xavier, Joana C., Hordijk, Wim, Kauffman, Stuart A., Steel, Mike, Martin, William F., 2020. Autocatalytic chemical networks at the origin of metabolism. Proc. Roy. Soc. B 287 (1922), 20192377.