Defining Autocatalysis in Chemical Reaction Networks

Jakob L. Andersen[a], Christoph Flamm[d], Daniel Merkle[a], Peter F. Stadler[b,c,d,e,f]

ABSTRACT Autocatalysis is a deceptively simple concept, referring to the situation that a chemical species X catalyzes its own formation. From the perspective of chemical kinetics, autocatalysts show a regime of super-linear growth. Given a chemical reaction network, however, it is not at all straightforward to identify species that are autocatalytic in the sense that there is a sub-network that takes X as input and produces more than one copy of X as output. The difficulty arises from the need to distinguish autocatalysis e.g. from the superposition of a cycle that consumes and produces equal amounts of X and a pathway that produces X. To deal with this issue, a number of competing notions, such as exclusive autocatalysis and autocatalytic cycles, have been introduced. A closer inspection of concepts and their usage by different authors shows, however, that subtle differences in the definitions often makes conceptually matching ideas difficult to bring together formally. In this contribution we make some of the available approaches comparable by translating them into a common formal framework that uses integer hyperflows as a basis to study autocatalysis in large chemical reaction networks. As an application we investigate the prevalence of autocatalysis in metabolic networks.

Keywords: stoichiometry; directed hypergraph; hyperflow; chemical organization; metabolic networks; origins of life

Introduction

The idea of autocatalysis is deceptively simple. A chemical reaction is autocatalytic whenever one of its educts catalyzes its own formation, i.e.,

$$(\mathrm{A}) + \mathrm{X} \longrightarrow 2\,\mathrm{X} + (\mathrm{W})$$

where (A) and (W) denote some sets of extra building material and waste products, respectively. One of the few autocatalytic reactions that is of this simple form is the Soai reaction [1], an alkylation of pyrimidine-5-carbaldehyde with diisopropylzinc. Here, each enantiomer of the product catalyzes only the formation of the same enantiomer.

The concept of autocatalysis goes back to Ostwald [2]. In the context of chemical kinetics, autocatalysis refers to a temporary speed-up of the reaction before it settles down to reach equilibrium, see e.g. [3, 4] for a recent review. In most cases this leads to characteristic sigmoidal time courses. Autocatalysis may also be associated with more complex dynamic behavior, such as oscillations.

Maybe the best-known example of an autocatalytic reaction is the hydrolysis of esters, which is catalyzed by the acid that is one of the reaction products. Even in this simple case, however, we better understand its autocatalytic nature as a generic acid catalysis of the cleavage reaction

$$\mathrm{H^+ + R{-}COOR' + H_2O \longrightarrow R{-}COOH + HO{-}R' + H^+}$$

and the dissociation of the acid

$$\mathrm{R{-}COOH \rightleftharpoons R{-}COO^- + H^+}$$

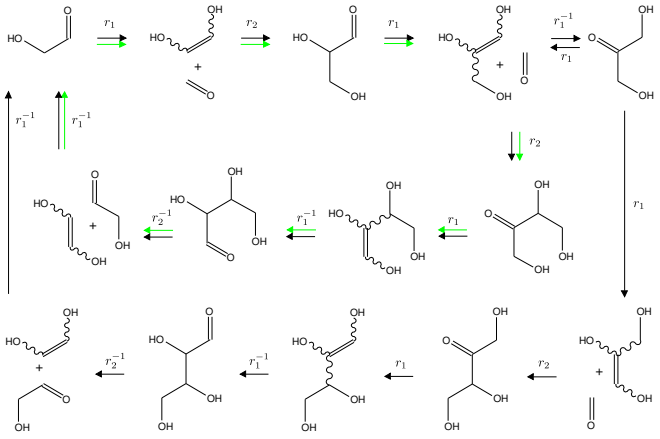


Figure 1: The formose reaction can be understood as a cyclic process involving keto-enol tautomerisation (r_1 and r_1^{-1}), aldol-condensation (r_2) and reverse aldol reaction r_2^{-1}. The inner cycle (green arrows) follows [8]. The outer cycle is the more commonly discussed mechanisms [9]. Figure taken from [10].

Of course, the cleavage reaction itself consists of multiple steps, none of which is overtly catalytic [5]. The mechanisms by which Mn^{2+} catalyzes the oxidation of oxalate by permanganate in this classical example of autocatalysis is much less obvious and can be explained only by an elaborate network of reactions [6, 7].

In fact, one of the earliest autocatalytic reactions reported in the literature, the Formose reaction [11], is a reasonably well-understood example of "network autocatalysis". The simple, autocatalytic overall reaction

$$\mathrm{HOCH_2CHO + 2\,H_2CO \longrightarrow 2\,HOCH_2CHO}$$

has been understood as the net effect of the example re-

Email addresses: jlandersen@imada.sdu.dk (Jakob L. Andersen), xtof@tbi.univie.ac.at (Christoph Flamm), daniel@imada.sdu.dk (Daniel Merkle), studla@bioinf.uni-leipzig.de (Peter F. Stadler)

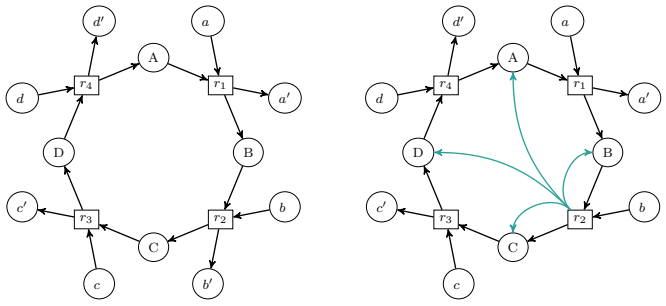


Figure 2: Catalytic and autocatalytic cycle. (a) A catalytic cycle is shown. A set of educts $\{a, b, c, d\}$ is converted into a set of products $\{a', b', c', d'\}$ by a set of species $\{A, B, C, D\}$ which acts as as catalysts. (b) If one of the reactions in the catalytic cycle (here r_2) additionally produces a copy of the species in the cycle (indicated by the green arrows), then the catalytic cycle becomes autocatalytic. Green arrows may also represent multi-step reaction sequences.

action network shown in Fig. 1. Several well-known oscillating reactions, including the Belousov-Zhabotinsky (BZ) reaction [12], are also elaborate examples of network autocatalysis. In fact, the core of the BZ reaction harbors two autocatalytic cycles, one feeding on the other in a *predator-prey* like fashion, resulting in the Lotka-Volterra type oscillatory dynamics [13]. The Whitesides group recently designed an autocatalytic network comprising only a few simple organic compounds that displays oscillatory behavior [14]. A computational study identified coupled autocatalytic cycles in the chemical networks of Eschenmoser's glyoxylate scenario [15]. For the distinction of catalytic and autocatalytic cycles, see Fig. 2.

The concept of autocatalysis plays an important role in metabolic networks. In this context, one frequently speaks of *autocatalytic pathways*, which contain reactions that consume some of the pathway's products. This results in the positive feedback, which in turn explains their characteristic dynamic behavior [16, 17]. A paradigmatic example is glycolysis, which invests two ATP molecules to later produce four.

Autocatalysis plays a key role in most models of the origin of life. Replicating entities — by definition — are autocatalytic. First described theoretically by Manfred Eigen [18], it was soon shown that short nucleic acid templates can be copied, e.g., by ligation of short fragments without the help of enzymes [19]. Alternative models, such as self-replicating peptides [20] or lipid aggregates [21] follow the same logic. Tibor Gánti [22, 23] early-on emphasized the importance of autocatalytic cycles. In order to explain the emergence of replicators, "collectively autocatalytic" networks of interacting molecules have been proposed as precursors of replicating polymers [24]. These chemical reaction networks (CRNs) contain molecules that promote their own synthesis, forming chemical organizations [25, 26]. A distinct concept of "autocatalytic networks" refers to interacting autocatalytic replicators generalizing the hypercycle model of Eigen and Schuster [27, 28, 29]. It describes systems of self-replicating entities rather than chemical reactions of small molecules.

A popular mathematical model of autocatalytic reaction networks are the *Reflexively Autocatalytic Food generated* networks (RAFs) by Steel and Hordijk [30, 31]. Similar to chemical organizations, all chemical species in a RAF $\mathcal{R}$ can be produced from the food or other elements of $\mathcal{R}$ [32]. The model is mathematically much easier to handle than arbitrary CRNs because one considers only reactions of the form

$$\mathrm{C} + \sum_i s_i \mathrm{A}_i \longrightarrow \sum_j s'_j \mathrm{B}_j + \mathrm{C}$$

That is, every reaction is catalyzed by some of the species. A RAF set $\mathcal{R}$ thus also contains a sufficient set of catalysts. While RAF theory is a plausible description, e.g., of ligation networks of simple polymers [33, 34], it does not seem to be a realistic description of reaction networks of small molecules. Here, the assumption that all reactions are catalyzed appears very unrealistic. Unfortunately, the algorithms for recognizing RAFs [35, 36] do not seem to generalize to arbitrary networks composed of non-catalyzed reactions. RAFs are not necessarily meant to model concrete chemical reactions but rather aggregate transformations. The RAF formalism coarse-grains the elementary steps of a catalytic *process* and replaces them by a single influence arrow. This is a valid abstraction if enzymes or other large polymeric entities are the catalysts because they are molecular machines that encapsulate or sequester the individual steps and thus separate the catalytic process from the rest of the system. It is not an appropriate approximation for networks of small (prebiotic) molecules. Here, the intermediates are accessible for alternative reactions. A specific catalytic influence beyond global effects (such as changes in pH or ionic strength) in a CRN is itself a chemical reaction. It remains an open question, therefore, under which conditions a given CRN can be abstracted into the RAF formalism. We suspect that small molecule CRNs are not of this type (recent attempts notwithstanding [37]), limiting RAFs to the realm of macromolecular and supramolecular complexes.

In this contribution, we therefore seek to develop a theory that can be used to identify autocatalytic structures in a given CRN, i.e., a system of chemical reaction equations. To this end we first need to introduce a sound mathematical framework. This is less trivial than it might seem. The notion of autocatalysis in chemical kinetics is difficult to use in a network setting since it strongly depends on the actual choice of rate constants. A natural starting point for a theory of autocatalytic CRNs is to ask for subnetworks for which the rate constants can be chosen such that it shows autocatalytic kinetics. The kinetic criterion, however, is also not entirely unambiguous as we shall see.

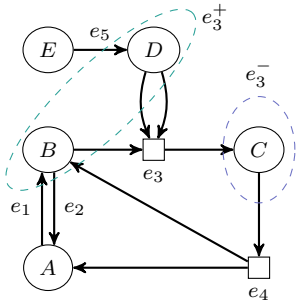


Figure 3: Example of a directed multi-hypergraph $\mathcal{H}$ with vertices $V = \{A, B, C, D, E\}$ and hyperedges $E = \{e_1, e_2, e_3, e_4, e_5\}$, which we will use as a running example to illustrate concepts. Parallel arrows represent the multiplicity of a vertex in a tail/head multiset. For example, the hyperedge e_3 is defined as the pair (e_3^+, e_3^-) with tail $e_3^+ = \{\!\{B, D\}\!\}$ (the cyan ellipse) and head $e_3^- = \{\!\{C\}\!\}$ (the violet ellipse). The multiplicities for the tail/head memberships are $m_B(e_3^+) = 1$, $m_D(e_3^+) = 2$ and $m_C(e_3^-) = 1$. To reduce clutter we often depict a hyperedge with a single tail and head vertex as a lonely arrow without a box (here e_1, e_2, and e_5). The hypergraph represents the reactions $\mathrm{A} \rightleftharpoons \mathrm{B}$, $\mathrm{B} + 2\,\mathrm{D} \longrightarrow \mathrm{C}$, $\mathrm{C} \longrightarrow \mathrm{A} + \mathrm{B}$, and $\mathrm{E} \longrightarrow \mathrm{D}$.

Towards a Structural Theory of Autocatalysis

Directed Multi-Hypergraphs

A CRN consists of a set of molecules V and set of reactions E such that every $e \in E$ is of the form

$$\sum_{x \in V} s_{xe}^+ x \longrightarrow \sum_{x \in V} s_{xe}^- x$$

Note that we regard all reactions as directed. Reversible reactions therefore are represented by a separate forward and backward reaction. This will allows us to use non-negative flows and connects naturally with graph transformations as a means of generating chemical reactions.

Following the notation of [38], a CRN is naturally represented as a directed multi-hypergraph $\mathcal{H} = (V, E)$ where each hyperedge $e \in E$ is a pair of multisets

$$e^+ := \{\!\{x \mid s_{xe}^+ > 0\}\!\} \quad \text{and} \quad e^- := \{\!\{x \mid s_{xe}^- > 0\}\!\}$$

The notation $\{\!\{\dots\}\!\}$ emphasizes that we are dealing with multisets, where an element can be contained more than once. We write $m_x(\,.\,)$ for these *multiplicities*, which in our case are given by the stoichiometric coefficients: $m_x(e^+) = s_{xe}^+$ and $m_x(e^-) = s_{xe}^-$. We call e^+ the *tail* and e^- the *head* of the directed hyperedge. See Fig. 3 for an example.

Every directed multi-hypergraph $\mathcal{H} = (V, E)$ has a faithful representation as a bipartite multi-digraph with vertex set $V' = V \cup E$ and a multiset of edges

$$\begin{aligned} E' = &\{\!\{(v, e) \mid e = (e^+, e^-) \in E, v \in e^+\}\!\} \\ &\cup \{\!\{(e, v) \mid e = (e^+, e^-) \in E, v \in e^-\}\!\} \end{aligned} \tag{1}$$

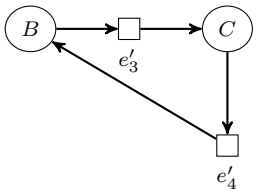


Figure 4: Example of a restriction of the hypergraph shown in Fig. 3. The shown hypergraph is $\mathcal{H}[V', E']$ with $V' = \{B, C\}$ and $E' = \{e_3, e_4\}$. Note that the original e_3 and e_4 have been modified as not all of their tail and head vertices are in V'.

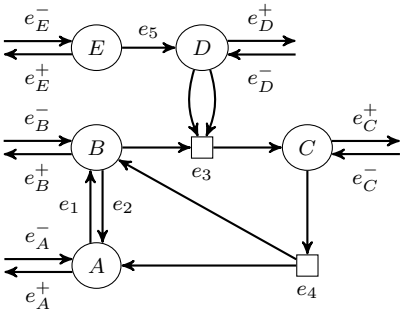


Figure 5: The extended hypergraph $\overline{\mathcal{H}}$ of the one shown in Fig. 3. To reduce visual clutter the boxes of the IO edges are omitted and only an arc are shown for each of them.

with multiplicities of arcs given by the stoichiometric coefficients. We will refer to this as the *König representation* $\mathrm{K}(\mathcal{H})$ of $\mathcal{H}$.

In our discussion we also need restrictions of directed hypergraphs to subsets of vertices and hyperedges in the following way. For $V' \subseteq V$ and $E' \subseteq E$ let $\mathcal{H}[V', E']$ be the directed multi-hypergraph with vertex set V' and the hyperedges $e' = (e^+ \cap V', e^- \cap V')$ for each $e \in E'$. In another view, $\mathcal{H}[V', E']$ is the hypergraph constructed by first taking the König representation $\mathrm{K}(\mathcal{H})$, selecting the subgraph induced by $V' \cup E'$, and then reinterpreting it back into a hypergraph. See Fig. 4 for an example.

The interaction of the CRN $\mathcal{H} = (V, E)$ with its environment is modeled by "exchange reactions" describing the possibility to import or export/accumulate chemical species. In general we add these reactions to every species, and later introduce flow constraints to model specific food and product sets when relevant. The exchange reactions are the *input* edges $E^- = \{e_v^- = (\emptyset, \{\!\{v\}\!\}) \mid v \in S\}$ and the *output* edges $E^+ = \{e_v^+ = (\{\!\{v\}\!\}, \emptyset) \mid v \in T\}$. We therefore define the *extended hypergraph* $\overline{\mathcal{H}} = (V, \overline{E})$ of $\mathcal{H}$ with $\overline{E} = E \cup E^- \cup E^+$. The exchange reactions appear as sources and sinks in the König representation $\mathrm{K}(\overline{\mathcal{H}})$. See Fig. 5 for an example of an extended hypergraph.

Composite Reactions and Formal Autocatalysis

On the set of reactions, i.e., hyperedges of a CRN $\mathcal{H}$ we construct *composite reactions* as integer linear combi-

nations of the form

$$\sum_{e\in E}\left(f_e\sum_{x\in e^+}s^+_{xe}x\right)\longrightarrow\sum_{e\in E}\left(f_e\sum_{x\in e^-}s^-_{xe}x\right)\qquad(2)$$

with $f_e\in\mathbb{N}_0$. A composite reaction often contains one or more species y that appear with the same multiplicity on the both sides, i.e., $\sum_{e\in E}f_e(s^+_{ye}-s^-_{ye})=0$. These are *formal catalysts* for the composite reaction. It is customary to cancel formal catalysts and to retain in the "net reaction" or "overall reaction" only the species for which $\sum_e(s^+_{ye}-s^-_{ye})f_e\neq 0$.

Definition 1. *A composite reaction is* formally autocatalytic for x *if it is of the form*

$$(\mathrm{A})+m\mathrm{x}\longrightarrow n\mathrm{x}+(\mathrm{W})$$

for some integers $n>m>0$.

A CRN is formally autocatalytic if it admits a composite reaction that is formally autocatalytic for one of its constituent compounds x. Def. 1 captures King's notion of autocatalytic sets [39]. It also matches with Gánti's notion that autocatalysis is associated with a cycle that eventually feeds a product back as an educt such that "after a finite number of turns, each constituent multiplies in quantity" [22].

We *conjecture* that it is impossible for a CRN (V,E) to show kinetic autocatalysis (i.e., non-linear acceleration) if, at least, it is not formally autocatalytic, since it seems plausible that the presence of a species x for which the network is formally autocatalytic is a necessary condition for positive feedback of x on its formation. The notion of formal autocatalysis implicitly appears in [33], where such composite reactions are shown to explain superlinear kinetics, i.e., autocatalytic behavior, of certain intermediates in a model of a complex ligation network.

It is easy to see that formal autocatalysis cannot be sufficient. To this end, consider two (possibly composite) reactions

$$(\mathrm{A})+\mathrm{X}\longrightarrow\mathrm{X}+(\mathrm{W})\qquad(\mathrm{B})\longrightarrow\mathrm{X}+(\mathrm{U})\qquad(3)$$

The first one is a net transformation $(\mathrm{A})\longrightarrow(\mathrm{W})$ catalyzed by X, i.e., it does not contribute to the production or degradation of X, while the second one is simply a production reaction for X. Their sum is formally autocatalytic for X with $m=1$ and $n=2$. Assuming that the reactions $(\mathrm{A})+\mathrm{X}\longrightarrow\mathrm{X}+(\mathrm{W})$ and $(\mathrm{B})\longrightarrow\mathrm{X}+(\mathrm{U})$ share only X, there clearly is no feedback between them.

In fact, formal autocatalysis is a very weak condition that includes reaction mechanisms such as the following 2-step decay of A:

$$(\mathrm{A})\longrightarrow 2\,\mathrm{X}\qquad\mathrm{X}\longrightarrow(\mathrm{V})+(\mathrm{W})\qquad(4)$$

This CRN contains the composite reaction $\mathrm{X}+(\mathrm{A})\longrightarrow 2\,\mathrm{X}+(\mathrm{V})+(\mathrm{W})$, making X formally autocatalytic, even though X in no is way involved in its own production or maintenance. This emphasizes that the definition of formal autocatalysis lacks a condition that ties the two "pathways" more closely together. These two simple examples naturally lead to a stricter notion of autocatalysis, "exclusive autocatalysis", where we require that X cannot be produced unless X is already present, e.g. see [16]. However, in order to formalize this idea properly, we first need to consider integer hyperflows as a way to formalize the intuitive notion of a pathway.

Before we proceed, we note that network structure alone is certainly insufficient to imply autocatalysis in the kinetic sense. Even in the setting of simple "autocatalytic cycles", the dynamical behavior depends crucially on the kinetic parameters [40].

Integer Hyperflows

Pathways, understood as systems of reactions with defined input, are naturally described mathematically as integer hyperflows [38]. In this and the following section we introduce some necessary notation and then explain the connection between integer hyperflows and the "algebra" of reactions in a CRN.

For an extended hypergraph $\mathcal{H}=(V,\overline{E})$, we write $\delta^+_A(v)$ as the set of out-edges from a vertex $v\in V$, restricted to the edge set $A\subseteq\overline{E}$, i.e., $\delta^+_A(v)=\{e\in A\mid v\in e^+\}$. Likewise, $\delta^-_A(v)$ denotes the restricted set of in-edges incident v.

Definition 2. *A* hyperflow *on* $\overline{\mathcal{H}}$ *is a function* $f\colon\overline{E}\to\mathbb{R}^+_0$ *satisfying, for each* $v\in V$ *the conservation constraint*

$$\sum_{e\in\delta^+_{\overline{E}}(v)}m_v(e^+)f(e)-\sum_{e\in\delta^-_{\overline{E}}(v)}m_v(e^-)f(e)=0\qquad(5)$$

The sum of flow out of each vertex must be the same as the sum of flow into it. The concept goes back to [41]. It also naturally appears in Metabolic Flux Analysis and Flux Balance Analysis: writing $\mathbf{S}_{ve}:=m_v(e^+)-m_v(e^-)=s^+_{ve}-s^-_{ve}$ indeed allows us to express Eq. (5) in matrix notation as $\mathbf{S}f=0$.

We write $f_1\le f_2$ if $f_1(e)\le f_2(e)$ holds for all $e\in E$. We write $f_1<f_2$ if $f_1\le f_2$ and $f_1\neq f_2$. In contrast we use $f_1\ll f_2$ if $f_1(e)<f_2(e)$ for every hyperedge $e\in\overline{E}$. A key property of flows is that linear combinations of flows are again flows as long as non-negativity is preserved. In particular the difference of two flows f_1 and f_2 is still a flow if and only if $f_1-f_2\ge 0$.

In this contribution we shall be interested mostly in *integer hyperflows*, which for simplicity we will refer to simply as flows unless otherwise specified.

For a flow f on $\overline{\mathcal{H}}$ we denote by $S(f)$ and $T(f)$ the *actual* source and target species in a given flow f, i.e.,

$$S(f)=\{v\mid f(e_v^-)>0\}\text{ and }T(f)=\{v\mid f(e_v^+)>0\}\quad(6)$$

When specifying a model for analysis we may also want to specify *a priori* an *allowed* source set $S\subseteq V$ and target set

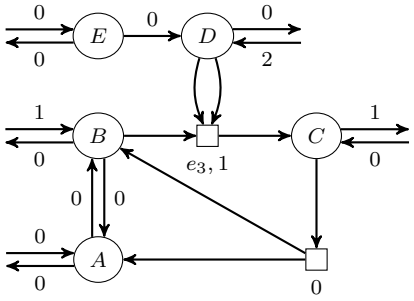


Figure 6: The extended hypergraph $\overline{\mathcal{H}}$ from Fig. 5, with f^{e_3} annotated.

$T \subseteq V$ in $\overline{\mathcal{H}}$. We refer to the triple $(\mathcal{H}, S, T)$ as the *I/O-constrained extended hypergraph*. In this situation we are only interested in flows f satisfying $f(e_v^-) = 0$ for all $v \notin S$ and $f(e_v^+) = 0$ for all $v \notin T$, i.e., $S(f) \subseteq S$ and $T(f) \subseteq T$. In the context of metabolic networks the sources S are usually given by the food set, and the targets T are the products that can be removed or accumulated.

Flows for Composite and Net Reactions

Our next task is to formally connect systems of reactions with flows. Recall that composite reactions are obtained by "adding up" reactions, i.e., hyperedges. The same can be done for flows. To this end, we associate each hyperedge e of $\mathcal{H}$, i.e., each reaction in the CRN, with a flow f^e defined by $f^e(e) = 1$, $f^e(e') = 0$ for $e' \in E \setminus \{e\}$, input-flows $f^e(e_v^-) = s_{xe}^-$ for $x \in e^-$ and output-flows $f^e(e_v^+) = s_{xe}^+$ for $x \in e^+$; all other input- and output-flows are set to zero. We call f^e the *reaction flow* of e. That is, a flow of 1 through reaction e requires an input-flow of its educts and an output-flow of its products in proportions given by the stoichiometric coefficient. The reaction flow f^e thus is simply a representation of a single reaction e in the language of flows. In Fig. 6 a reaction flow is shown.

This mathematical construct is useful because it makes it possible to write the flow f that is associated with a composite reaction (pathway) as a weighted sum of reaction flows. The multiplicity of a reaction e in Eq. (2) is simply the flow $f(e)$ through e and hence we have the formal decomposition

$$f = \sum_{e \in E} f(e) f^e \tag{7}$$

Recall that in constructing a composite reactions we are only allowed to add reactions. Thus every compound v comes with an input-flow $f(e_v^+)$ and an output-flow $f(e_v^-)$ that again matches the stoichiometric coefficients in the composite reactions.

The point of using net reactions, in contrast to using composite reactions, is that we are allowed to *cancel* intermediates, that is, to remove an equal number of copies from both the product and the educt side. This operation can also be formalized in terms of flows. To this end we introduce the *futile flow* f^v for compound v defined as $f^v(e_v^-) = f^v(e_v^+) = 1$ and $f(e) = 0$ for all other reactions $e \in \overline{E}$. Given a flow f, it is easy to see that $\tilde{f} = f - cf^v$ is again a valid flow as long as $c \leq \min\{f(e_v^+), f(e_v^-)\}$. That is, we can reduce in f the input-flow of v and output-flow of v by the same amount as long as we do not attempt to construct a negative input-flow $\tilde{f}(e_v^+)$ and or a negative output-flow $\tilde{f}(e_v^-)$. In terms of net reaction that means we may reduce the stoichiometric coefficients of a compound that appears on both sides by the same amount.

The issue here is that arbitrary canceling of intermediate compounds from a composite reaction does not necessarily leave us with a net reaction that will actually take place because we may have canceled essential catalytic or autocatalytic species. In the flow formalism, however, we can ask which cancellations are allowed and which are not: We only have to ask whether, for a given set S of input species and a given set T of output species there is a flow f with $S(f) \subseteq S$ and $T(f) \subseteq T$, where S and T are subsets of the species on the educt and product side of the composite reaction. If the answer is yes, we can cancel all intermediate species $x \in V \setminus (S(f) \cup T(f))$. Correspondingly, cancellations of $x \in S \cup T$ are not allowed in an I/O constrained networks $(\mathcal{H}, S, T)$.

Let us write $\mathrm{supp}(f) := \{e \in E \mid f(e) > 0\}$ for the set of reactions (not including I/O hyperedges) that are "active". Every flow f can be associated with a composite reaction, namely the one that consists of all reactions $e \in \mathrm{supp}(f)$. The stoichiometric coefficients for each $x \in V$ are given by

$$q_x^- = \sum_{e \in \mathrm{supp}(f)} f(e) m_x(e^-) \text{ and } q_x^+ = \sum_{e \in \mathrm{supp}(f)} f(e) m_x(e^+) \tag{8}$$

Since there is neither an input-flow nor an output-flow for $x \in V \setminus (S(f) \cup T(f))$, we can conclude immediately that stoichiometric coefficients of x as an educt, q_x^-, and as a product, q_x^+, must be the same. We summarize this discussion as

Lemma 3. *There is a flow f on the I/O-constrained network $(\mathcal{H}, S, T)$ if and only if there is a composite reaction $\sum q_x^- x \longrightarrow \sum q_x^+ x$. Moreover, in this case its stoichiometric coefficients satisfy $q_x^- = q_x^+$ for all $x \in V \setminus (S(f) \cup T(f))$.*

In summary, therefore, we can associate a flow with every composite reaction and *vice versa*. An advantage of the flow framework is that it links to a convenient computational paradigm. "Flow queries", i.e., the question whether there exists a flow with prescribed properties, are naturally phrased as (integer) linear programs, and thus can be answered by generic solvers, see e.g. [38] for a more detailed discussion.

Formally Autocatalytic Flows

We next link the flow formalism to the notion of formal autocatalysis introduced in Def. 1. The following statement is a direct consequence of Lemma 3, noting that an

(auto)catalytic species necessarily must be contained in both $S(f)$ and $T(f)$.

Lemma 4. *There is a formally autocatalytic compound reaction for x if and only if there is a flow f on $\overline{\mathcal{H}}$ such that*

$$0 < f(e_x^-) < f(e_x^+) \tag{9}$$

In a practical setting we may additionally I/O-constrain $\mathcal{H}$ with specific source and target sets S and T. The condition matches the definition of "overall autocatalysis" in, e.g., [38]. Naturally, we are interested in minimal formally autocatalytic flows f, i.e., those that do not contain a "smaller" formally autocatalytic flow f_1.

The notion of "smaller" in this context deserves some consideration. It could mean either $\mathrm{supp}(f_1) \subsetneq \mathrm{supp}(f)$ or $f_1 < f$. For not necessarily integer flows, is its well known that the existence of a flow f_1 with $\mathrm{supp}(f_1) \subsetneq \mathrm{supp}(f)$ is equivalent to the existence of a flow f_2 with $f_2 < f$ that is not proportional to f. Analogously, there is an integer flow f_1 with $\mathrm{supp}(f_1) \subsetneq \mathrm{supp}(f)$ if and only if there is an integer flow f_2 that is not proportional to f and an integer $a \geq 1$ such that $f_2 < af$. This suggests to think of "smaller" flows as those that have the smaller support. Support minimality features prominently with Extremal Flux Modes [42, 43] and has been discussed in detail in this context.

Exclusive Autocatalysis

As noted above, the Def. 1 and its counterpart in terms of flows on $\overline{\mathcal{H}}$, Eq. (9), are not satisfactory because parallel reactions such as Eq. (3) and even degradation pathways such Eq. (4) are formally autocatalytic. The most straightforward, but crude way of handling this shortcoming in the definition is to require, in addition, that an autocatalytic species x cannot be produced from within the network unless a minute amount is already present at the outset. In other words, the network under consideration does not contain a pathway that produces x in a non-autocatalytic manner from the same food set. This concept matches the intuition of autocatalysis, e.g., in [16], and was used as a component in [38]. In the language of flows we can formalize it as follows:

Definition 5. *A species x is* exclusively autocatalytic *in an I/O-constrained network $(\mathcal{H}, S, T)$ if there is a flow f such that (i) x is formally autocatalytic in f and (ii) there is no flow f_1 in $(\mathcal{H}, S, V)$ with $f_1(e_x^-) = 0$ and $f_1(e_x^+) > 0$.*

Exclusive autocatalysis is a quite strict requirement: if x in any way can be produced from the sources, without regard to the sinks, it is disqualified from being exclusively autocatalytic. Condition (ii) thus boils down to a simple reachability question in $\mathcal{H}$. In general, for a given set of starting materials ("food set") $F \subseteq V$ and a set $E' \subseteq E$ of reactions the *scope* [44] – or the *closure* in the language of chemical organizations [25, 26] $c(F, E')$ is constructed recursively as $c(F, E') = \bigcup_i Q_i$, where $Q_0 = S$ and, for $i \geq 1$,

$$Q_i = \bigcup \{e^- \mid e \in E' \text{ and } e^+ \subseteq Q_{i-1}\} \tag{10}$$

is the set of a product compounds that can be produced by reactions (in E') whose educts are available in the previous step Q_{i-1}. An equivalent way to define $c(F, E')$ is to require $F \in c(F, E')$, and then for all edges $e \subseteq E'$ if all tail vertices are included, $e^+ \subseteq c(F, E')$, then all head vertices are as well, $e^- \in c(F, E')$. Condition (ii) of Def. 5 can thus be expressed as $x \notin c(S \setminus \{x\}, E)$.

Def. 5 formalizes a very strict interpretation of the idea that x cannot be produced unless it is present to seed to its own production. Condition (ii) is independent of the candidate flow f and pertains to the complete molecule set V as target set. As an object of future study there are several meaningful, less restrictive variations of the definition, for example:

1. f_1 is found in $(\mathcal{H}, S(f), V)$, allowing all edges E,
2. f_1 is found in $(\mathcal{H}, S, T)$, allowing all edges E,
3. f_1 is found in $(\mathcal{H}, S(f), T(f))$, allowing all edges E,
4. f_1 is found in $(\mathcal{H}, S(f), T(f))$, but allowing only edges from $\mathrm{supp}(f)$,

While the first of these variants also can be phrased as a reachability problem, the others are non-trivial hyperflow queries due to the constraint on the output flow to a subset of vertices. The last variant can be interpreted as a question on whether x can be canceled from the educt side of the composite reaction defined by the formally autocatalytic flow f. This condition therefore is in a sense concerned with the connectedness of the formally autocatalytic flow f. All these concepts of exclusive or "obligatory" autocatalysis are very restrictive as far as alternative routes are concerned, while the idea of an underlying autocatalytic cycle is implicit at best.

Autocatalytic Cycles sensu Barenholz et al. (2017)

Several authors have formalized autocatalysis in terms of the algebraic properties of the stoichiometric matrix $\mathbf{S}$. In this and the following section we review two definitions and fit them into the mathematical framework outlined above, and thus translating them into constraints on flows on the extended hypergraph $\overline{\mathcal{H}} = (V, \overline{E})$.

In [40] an autocatalytic cycle is defined as a pair (M, R) of metabolites $M \subseteq V$ and $R \subseteq E$ such that the restriction $\mathbf{S}^*$ to rows M and columns R satisfies the following conditions:

(o) R contains no reversible pair of reactions.

(i) For every $x \in M$ there is $e_1, e_2 \in R$ with $s_{xe_1} > 0$ and $s_{xe_2} < 0$, and
for every $e \in R$ there is $x_1, x_2 \in M$ with $s_{x_1 e} > 0$ and $s_{x_2 e} < 0$.

(ii) There is a strictly positive integer vector $w \in \mathbb{N}^{|R|}$, $w \gg 0$ such that $\mathbf{S}^* w > 0$,

(iii) There is no vector $w' > 0$ with at least one $e \in R$ for which $w'_e = 0$ such that $\mathbf{S}^* w > 0$.

Using the fact that we can express composite reactions as reaction flows we can rewrite condition (ii) in the flow form as

(ii') There is a flow f on $\overline{\mathcal{H}}$ such that $f(e_v^+) \geq f(e_v^-) > 0$ for all $v \in M$ and $f(e) = 0$ for all $e \in E \setminus R$ and $f(e_v^+) > f(e_v^-)$ for at least one $v \in M$.

The first part of condition (i) is equivalent to $x \in M$ appearing on both sides of the composite reaction, and thus $f(e_v^+), f(e_v^-) > 0$ in the corresponding flow. Thus M consists only of species that are catalytic ($f(e_v^+) = f(e_v^-)$) or autocatalytic for v. The second condition constrains M to contain at least one educt and one product of every $e \in \operatorname{supp}(f)$.

Definition 6. *Let f be a flow on $\overline{\mathcal{H}}$. A subset $M \subseteq V$ is a* Milo set *for f if*

(M1) *$v \in M$ implies $f(e_v^+) \geq f(e_v^-) > 0$,*

(M2) *there is $v \in M$ such that $f(e_v^+) > f(e_v^-)$,*

(M3) *for all $v \in M$ there is $e' \in E$ with $f(e') > 0$ and $v \in e'^-$ and $e'' \in E$ with $f(e'') > 0$ and $v \in e''^+$, and*

(M4) *for every $e \in E$ with $f(e) > 0$ holds $M \cap e^- \neq \emptyset$ and $M \cap e^+ \neq \emptyset$.*

A flow f with a Milo set is a Milo flow.

Note that (M2) implies that a Milo set is non-empty. Furthermore, if f is a Milo flow, then the Milo set satisfies $M \subseteq S(f) \cap T(f)$.

Lemma 7. *If f is a Milo flow on $\overline{\mathcal{H}}$ then f is formally autocatalytic for at least one $x \in M$.*

Proof. By (M1), $f(e_v^-) > 0$ for all $v \in M$. Thus (M2) implies that there is $v \in M$ with $f(e_v^+) > f(e_v^-) > 0$, i.e., v is formally autocatalytic according to Lemma 4. □

For a Milo flow f, consider the the König graph $G := \mathrm{K}(\mathcal{H}[M, \operatorname{supp}(f)])$ of its restriction to the Milo set of f. By (M3), G has no source or sink vertices, i.e., every vertex of a Milo set is contained in a cycle of G.

So far, we have not used conditions (o) and (iii).

Definition 8. *A Milo flow f with Milo set M forms an* autocatalytic cycle (f, M) *(*sensu *Barenholz* et al.*, 2007) if there is no flow f_1 with $\operatorname{supp}(f_1) \subsetneq \operatorname{supp}(f)$ that satisfies* (M1) *and* (M2).

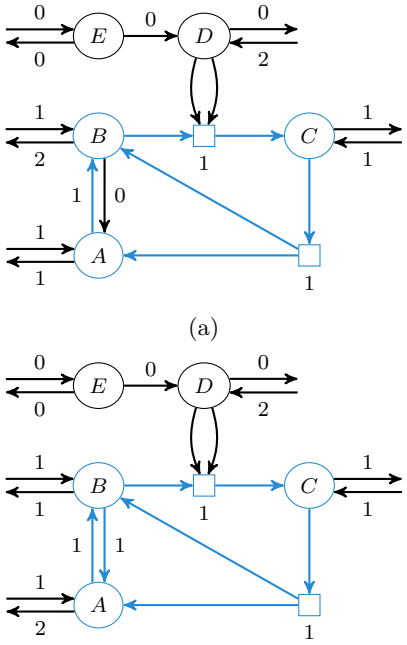


Figure 7: Two examples of Milo flows on the same CRN. (a) The Milo set M and the supporting reactions $\operatorname{supp}(f)$, i.e., restricted network $\mathcal{H}[M, \operatorname{supp}(f)]$ are highlighted in blue. The vertex B is formally autocatalytic. (b) Another Milo flow, containing a pair of reversible reactions. The equivalent flow f' obtained by removing $f(AB) = f(\overline{AB})$ is no longer a Milo flow because A has no outgoing reaction left in $\operatorname{supp}(f')$. Thus, f is a minimal Milo flow. However, (f, M) is not an autocatalytic cycle in the sense of Barenholz [40] since the flow f' is a forbidden flow according to Def. 8. Since f' itself is not a Milo flow, (f', M) is also not an autocatalytic cycle.

The flow f of an autocatalytic cycle in the sense of Def. 8 does not contain a pair of reactions that form a reversible pair e, $\bar{e}$. If f contains such a reaction, consider the flow

$$\begin{aligned} f_1 = & \sum_{e' \in \operatorname{supp} f \setminus \{e, \bar{e}\}} f(e) f^e \\ & + \min(f(e) - f(\bar{e}), 0) f^e \\ & + \min(f(\bar{e}) - f(e), 0) f^{\bar{e}} \end{aligned} \tag{11}$$

By construction, f_1 coincides with f on $\operatorname{supp}(f) \setminus \{e, \bar{e}\}$, has positive input-flow and output-flow, and satisfies that $f_1(e_v^+) - f_1(e_v^-) = f(e_v^+) - f(e_v^-)$ for every $v \in M$. Since $f_1(e) = 0$ or $f_1(\bar{e}) = 0$, it is a forbidden flow according to Def. 8. Thus (o) is in fact a consequence of (iii).

The forbidden flow f_1 in Def. 8 is a very strong condition. In particular [40] states (without proof) that the König graph of every autocatalytic cycle is strongly connected. At this point there is no formal proof for this statement, however.

The class of forbidden flows f_1 in Def. 8 is larger than Milo flows since f_1 is not restricted to flows with inputs

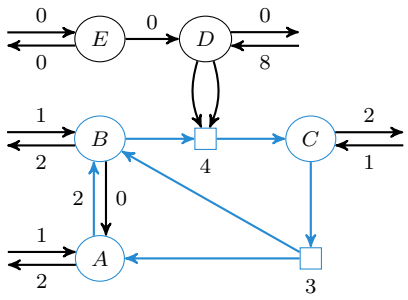


Figure 8: Example of an Nghe flow, with the defining Nghe set and supporting reactions highlighted in blue, i.e., the restricted network $\mathcal{H}[M, \mathrm{supp}(f)]$. All vertices in the Nghe set, A, B, and C are formally autocatalytic.

and output from within the set M. Defining a minimal Milo flow to be one for which there is no Milo flow f' with $\mathrm{supp}(f') \subseteq \mathrm{supp}(f)$, we observe that every autocatalytic cycle is a minimal Milo flow. The converse, however, is not necessarily true, as shown by the example in Fig. 7b. It remains an open question whether all minimal Milo flows are also strongly connected. Fig. 7b also shows that there are strongly connected Milo flows that are not autocatalytic cycles in the sense of Barenholz et al.

Autocatalytic Cores sensu Blokhuis et al. (2020)

The key concept in [45] are submatrices $\mathbf{S}^*$ of the stoichiometric matrix that are autonomous and productive in the following sense:

(i) $\mathbf{S}^*$ is productive if there is a $u \gg 0$ such that $\mathbf{S}^* u \gg 0$

(ii) For every column e of $\mathbf{S}^*$ there are rows v' and v'' such that $\mathbf{S}^*_{ev'} < 0$ and $\mathbf{S}^*_{ev''} > 0$.

A *autocatalytic core* is defined as a minimal submatrix $\mathbf{S}^*$ of $\mathbf{S}$ with these properties. Prop. 1 in [45] shows that in an autocatalytic core, every species x appears both as a substrate and as a product. This concept can be rephrased in terms of flows in a manner that emphasizes its relationship with [40].

Definition 9. *Let f be a flow on $\overline{\mathcal{H}}$. A subset $N \subseteq V$ is a* Nghe set *for f if it satisfies*

(N1) *$v \in M$ implies $f(e_v^+) > f(e_v^-) > 0$*

and conditions (M3) *and* (M4) *of Def. 6. A flow f with a Nghe set $N \neq \emptyset$ is a* Nghe flow.

From Def. 9 we immediately see that every Nghe flow is also a Milo flow with $M = N$ since (N1) obviously implies (M1) and (M2). Thus catalytic cores are Milo flows. In Fig. 8 an example of an Nghe flow is shown.

Condition (N1) appears very restrictive. It will be of immediate interest, therefore, to better understand under which conditions a Milo flow contains a Nghe flow in the sense that for a Milo flow (f, M) there is a Nghe flow (f_1, N) with $N \subseteq M$ and $\mathrm{supp}(f_1, N) \subseteq \mathrm{supp}(f, M)$. The relationships between Milo and Nghe flows deserve attention in future work. Similarly, the connections between autocatalytic cycles *sensu* Barenholz and autocatalytic cores will be of interest.

Proposition 2 of [45] shows that autocatalytic cores f are very restricted structures: it is "square", i.e., $|N| = |\mathrm{supp}(f)|$, every $x \in N$ is "the solitary substrate of a reaction, and is substrate for this reaction only". Proposition 4 of [45], furthermore, states that every autocatalytic core is strongly connected. Thus, strongly connected Nghe flows seem to be interesting objects to study in their own right.

The work of Nghe [45] shows that minimal autocatalytic cores have an essentially geometric characterization that can be expressed largely in terms of the König graph $K := \mathrm{K}(\mathcal{H}[N, \mathrm{supp}(f)])$ of a minimal Nghe flow. In essence they can be understood as "cycles with ears" comprising a simple cycle in K augmented by either "short cut reactions" or a path leading from some starting vertex x in the cycle back to an end-vertex y on the cycle without intersecting the cycle in its interior. In [45], additional algebraic and minimality conditions are required for a complete characterization of minimal autocatalytic cores. This geometric structure suggests to search for hyperflows whose Milo or Nghe sets have cycles or ears as their König graphs.

Mechanistically Simple Flow Solutions

The notion of "autocatalytic cycles" and in particular the idea of "going around a cycle" to produce additional copies of autocatalytic compounds [22] suggests a definite temporal order in which molecules "flow" through the reactions. This matches the chemist's concept of a *mechanism* as a sequence of reactions. Condition (M3) for Milo and Nghe flows addresses this concern to some extent by requiring input-flow and output-flow for every vertex in the distinguished set, thus ensuring that a cycle exists in the König graph of the support. On the other hand, minimal Milo and Nghe flows are rather restrictive in their input/output conditions requiring all vertices in the distinguished set to be a source and a target. The associated concepts of autocatalytic cycles or cores, furthermore, ban pairs of reversible reactions to be used.

The basic flow formulation, and the equivalent formulation based on the stoichiometric matrix, only ensures mass balance, and does not imply any particular ordering of reactions as such. In the following we recap the notion of *expanded flows* from [38], which has constraints that introduce localized temporal order in the flow model. It makes it feasible to keep a predefined source/target specification in terms of a I/O-constrained CRN $(\mathcal{H}, S, T)$, as well as allowing pairs of reversible reactions. The model thus serves as a foundation for finding chemical pathways in general.

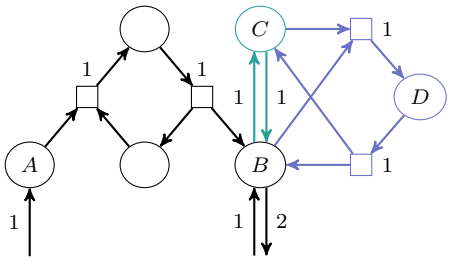


Figure 9: Example for local reasoning of reaction ordering on a seemingly formally autocatalytic flow. In all interpretations of the flow, the violet reactions into and out of D forms a futile two-step sub-pathway. After removing the flow on these reactions we can then apply the same argument to C, and then once more on the I/O flow of B. This leaves the net reaction A $\longrightarrow$ B.

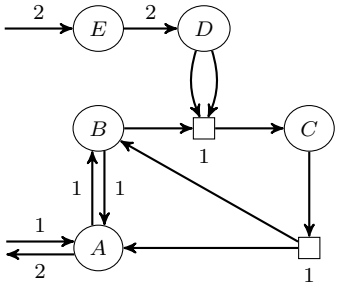


Figure 10: Example of a formally autocatalytic flow using pairs of mutually reverse reactions. Here, the reactions can be ordered such that no two-step futile sub-pathways are present.

In a later section we sketch how the model can be enriched with more computational expensive constraints that ensure the cyclicity required for a comprehensive model of structural autocatalysis.

As a motivating example, consider the expanded network with flow depicted in Fig. 9. The flow, with the net reaction $\mathrm{A}+\mathrm{B} \longrightarrow 2\,\mathrm{B}$, is formally autocatalytic, but due to use of reversible reactions it is equivalent to the simpler reaction A $\longrightarrow$ B. This can be established through a step-wise local reasoning:

1. The only in-flow to D is from $\mathrm{B}+\mathrm{C} \longrightarrow \mathrm{D}$ and the only out-flow is through the reverse reaction $\mathrm{D} \longrightarrow \mathrm{B}+\mathrm{C}$. Any ordering of the reactions in the flow will have this two-step futile part, and the flow on these reactions can thus be removed.
2. Without the violet part of the network, we can apply the same reasoning to vertex C with the reactions B $\longrightarrow$ C and C $\longrightarrow$ B.
3. Without both the violet and cyan parts, we can consider B with its input/output reactions, and decrease the flow by 1.

In this example there is no flow left on reversible reactions, but this is not the case in general. Consider the formally autocatalytic flow shown in Fig. 10, on our running example network. Here there are no vertices where we can apply the local temporal reasoning, and in fact there even exists a partial order for the reactions that have no pairs of reversible reactions in sequence:

1. $\emptyset \longrightarrow$ E, twice
2. E $\longrightarrow$ D, twice
3. $\emptyset \longrightarrow$ A
4. A $\longrightarrow$ B
5. $\mathrm{B}+2\,\mathrm{D} \longrightarrow \mathrm{C}$
6. $\mathrm{C} \longrightarrow \mathrm{A}+\mathrm{B}$
7. B $\longrightarrow$ A
8. A $\longrightarrow \emptyset$, twice

To consider global ordering one must invoke much stronger, and computational expensive, formalisms, such as Petri nets that explicitly "tracks" the paths of molecules through the CRN [46]. Note also that in general a pathway may have cycles even in a fully resolved temporal interpretation.

Fig. 10 also shows the requirement of mechanistic simplicity can enforce topological constraints. The flows obtained by changing the values for the reactions A $\longrightarrow$ B and B $\longrightarrow$ A to 0 or 2 are no longer chemically simply: In the first case it can only be realized by influx of 1 at A that immediately flows out again, and in the second case it required the a flow of 1 reaching B from A is immediately redirected back to A.

Expanded Networks and Flows

To address the need for local routing constraints on flows we introduce the *expanded* hypergraph [38]. Given an extended hypergraph $\overline{\mathcal{H}} = (V, \overline{E})$ we expand each vertex into a complete bipartite graph with vertices corresponding to each in-edge and out-edge. That is, for each $v \in V$:

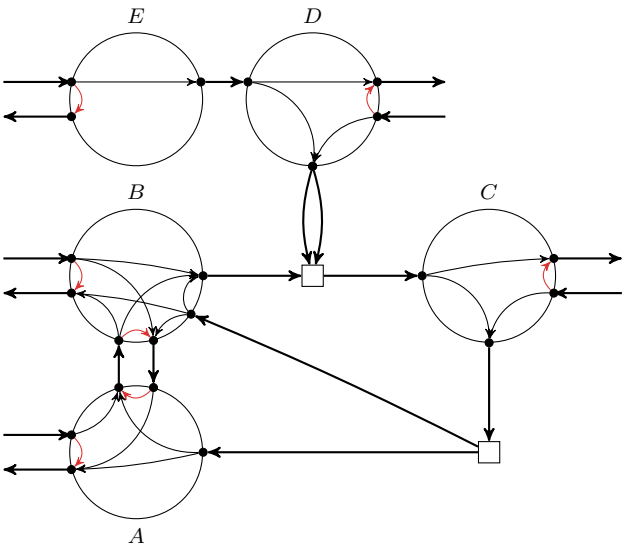
$$
\begin{aligned}
V_v^- &= \{u_{ve}^- \mid \forall e \in \delta_{\overline{E}}^-(v)\} \\
V_v^+ &= \{u_{ve}^+ \mid \forall e \in \delta_{\overline{E}}^+(v)\} \\
E_v &= \left\{ \left(\{\!\{u^-\}\!\}, \{\!\{u^+\}\!\} \right) \mid u^- \in V_v^-, u^+ \in V_v^+ \right\}
\end{aligned}
$$

The hyperedges E_v all have multiplicity 1 for their tail and head vertex, and we call these edges the *transit edges* of v. We then connect the original edges in the natural manner: for each $e = (e^+, e^-) \in \overline{E}$ the reconnected edge is $\widetilde{e} = (\widetilde{e}^+, \widetilde{e}^-)$ with $\widetilde{e}^- = \{\!\{u_{ve}^- \mid v \in e^-\}\!\}$ and $\widetilde{e}^+ = \{\!\{u_{ve}^+ \mid v \in e^+\}\!\}$. The multiplicities of tails and heads correspond to the original multiplicities. We finally define the expanded hypergraph $\widetilde{\mathcal{H}} = (\widetilde{V}, \widetilde{E})$ as

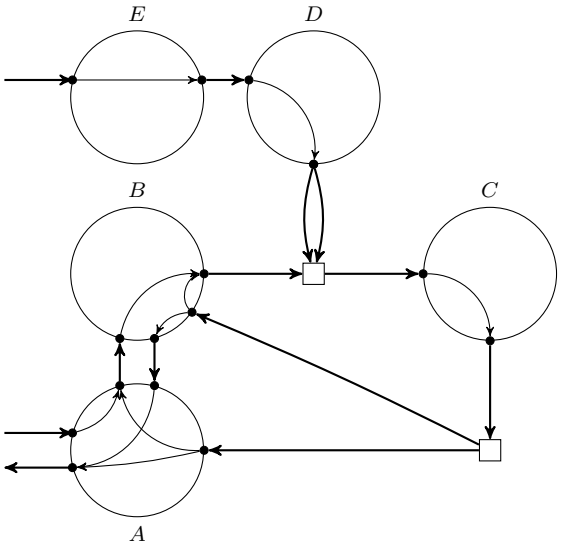
$$
\widetilde{V} = \bigcup_{v \in V} V_v^- \cup \bigcup_{v \in V} V_v^+ \quad \text{and} \quad \widetilde{E} = \bigcup_{v \in V} E_v \cup \{\widetilde{e} \mid e \in \overline{E}\}
$$

This is again a directed multi-hypergraph where (integer) flows are defined as usual. An example of an expanded hypergraph is shown in Fig. 11a.

For each pair of mutually reverse edges $e = (e^+, e^-), \bar{e} = (\bar{e}^+, \bar{e}^-) \in \overline{E}$ and a vertex $v \in e^-$ there is a *futile transit edge* $t = (u_{ve}^-, u_{v\bar{e}}^+)$. These futile transit edges correspond to pushing flow back immediately in the opposite direction of a reversible reaction without first processing the



(a)



(b)

Figure 11: Example of expanded hypergraphs. (a) The expanded hypergraph $\widetilde{\mathcal{H}}$ of the network from Fig. 3. The actual vertices are the small black circles while the large circles only indicate grouping corresponding to the original vertices of $\mathcal{H}$. The red transit edges are those that go between pairs of edges that are mutually reverse of each other in the original network. (b) The effective network that can have non-zero flow, when setting the allowed sources to $\{A, E\}$ and allowed targets $\{A\}$.

products in a different reaction. In Fig. 11 these edges are shown in red. We can now add constraints on flow reversibility by simply requiring that the flow on the red edges vanishes, that is, we enforce the constraint $f(t) = 0$ for all futile transit edges in the expanded hypergraph.

As shown in [38] this expanded network model is computationally not much more difficult to find solutions in than the original network. We can thus simply use the expanded network as a convenient background model for introducing routing constraints. In particular, for each flow $\widetilde{f}$ on the expanded network $\widetilde{\mathcal{H}}$ we can trivially obtain the equivalent flow f on the extended network $\overline{\mathcal{H}}$ simply by contracting the expanded vertices again. This leads us to a type of autocatalysis called *overall autocatalysis* [38].

Definition 10. *A species $x \in V$ is* overall autocatalytic *for a network $\mathcal{H} = (V, E)$ if there exist a flow $\widetilde{f}$ on the expanded network $\widetilde{\mathcal{H}}$ such that $f(t) = 0$ on all futile transit edges and the corresponding contracted flow f on the extended network $\overline{H}$ satisfies $0 < f(e_x^-) < f(e_x^+)$.*

This model of overall autocatalysis has been implemented using Integer Linear Programming as an extension of the software package MØD [47]. It does not constrain solutions to actually contain a cycle, but as outlined in the next section, it is already useful in analyzing chemical systems when coupled with the notion of exclusive autocatalysis described earlier.

Autocatalysis in Metabolic Networks

Metabolism as a whole seems to minimize the generation of waste molecules. Instead, byproducts and waste from one pathway are fed back into the network as a valuable resource for another. The effect of this "molecular recycling", or "metabolic closure", is the emergence of (catalytic) cycles in the reaction network, a necessary precondition for autocatalysis. Autocatalytic cycles can persist under noisy conditions, since they can replace mass loss along the cycle. This feature could be responsible for the inherent robustness of metabolism against fluctuations [48]. The embedding of multiple autocatalytic cycles in a network context results in feedback between cycles, giving rise to a rich repertoire of dynamic behavior and entry-points for regulation and control. Autocatalysis therefore plays an important role in metabolic networks.

Already in 2008, Kun and collaborators [16] published a search for obligatory autocatalytic species in large metabolic network models. Using the RAF framework, autocatalytic sets in the metabolic network of *E. coli* were studied in [49]. In order to illustrate the theoretical considerations in the previous sections we survey overall autocatalytic molecules in the metabolic networks of five very different prokaryotes as retrieved from the BiGG database [50], see Tab. 1. We only give a cursory overview here, a full investigation of autocatalysis using flows in these networks is forthcoming.

The BiGG models contain multiple copies of some molecules representing the compartments cytosol, periplasm, and the external environment. Here, we are only interested in the cytosolic metabolism. We therefore merged the periplasm with the external compartment and removed all reactions without educts or products in the cytosol.

Table 1: Overview of the investigated BiGG models. The original networks were simplified to study the capabilities of the cytosol compartment. The number of molecules that could be detected to be overall autocatalytic, using routing constraints in the expanded network are listed under "#OA". When further applying the strict conditions of exclusive autocatalysis ("EA") we are left with the number of molecules listed in the final column.

		Original		Simplified			
Species	BiGG ID	$\|V\|$	$\|E\|$	$\|V\|$	$\|E\|$	#OA	#(OA + EA)
Escherichia coli	iML1515	1 877	3 005	1 434	2 188	736	580
Helicobacter pylori	iIT341	485	641	485	641	176	143
Methanosarcina barkeri	iAF692	626	809	626	809	154	131
Mycobacterium tuberculosis	iEK1008	969	1 372	969	1 372	459	385
Staphylococcus aureus	iYS854	1 129	1 587	1 126	1 579	506	368

The size of the original and simplified networks are listed in Tab. 1. We then obtained the I/O-constrained hypergraphs interpreting the external molecules as source and target compounds. Furthermore, the explicit exchange pseudo-reactions in the models were converted into source/product specifications.

A molecule can only be (formally or overall) autocatalytic if it appears both as an educt and as a product, thus emulating that it may accumulate in the cell. Fixing a molecule X of interest, we construct an expanded flow model in which we add the condition that X is overall autocatalytic as an additional constraint. In total, this yields 4 640 different flow models of which 2 031 had feasible solutions; see Tab. 1 for a summary. Since many of the solutions in essence conform to Eq. (3) and thus do not represent autocatalysis in a chemically meaningful sense, we restricted ourselves to overall autocatalytic molecules that are also exclusively autocatalytic in the sense of Kun et al. [16]. That is, if a molecule is reachable from the sources (without itself), then it is not considered autocatalytic. This leaves 1 607 solutions.

The intersection of the five models shares 245 cytosolic molecules, of which 87 are overall autocatalytic. Only the 37 molecules listed in Tab. 2 are also exclusively autocatalytic.

This list for the most part comprises the expected "currencies" in the cell, in particular the mono-, di-, and triphosphorylated nucleotides, and the redox cofactors NAD and NADP. This matches the identification of ATP/ADP as ubiquitous "obligatory autocatalysts" in [16] using a very different approach. Furthermore, several tetrahydropholate derivatives, which are essential cofactors in the single carbon metabolism and two prebiotically relevant amino acids aspartate and serine are on the list. Interestingly, also the non-proteinogenic amino acid homoserine, an intermediate in the biosynthesis pathways of the three essential amino acids methionine, threonine, and isoleucine, as well as aspartate-semialdehyde, a building block involved in the biosynthesis of the amino acids lysine and homoserine are present.

Table 2: List of the 37 molecules that are (i) present in all five of the investigated models, (ii) overall autocatalytic, and (iii) exclusively autocatalytic.

BiGG ID	Name
`adp`	ADP
`amp`	AMP
`atp`	ATP
`cdp`	CDP
`cmp`	CMP
`ctp`	CTP
`dudp`	dUDP
`dump`	dUMP
`dutp`	dUTP
`gdp`	GDP
`gmp`	GMP
`gtp`	GTP
`udp`	UDP
`udpg`	UDPglucose
`udpgal`	UDPgalactose
`ump`	UMP
`utp`	UTP
`nad`	NAD
`nadh`	NADH
`nadp`	NADP
`nadph`	NADPH
`10fthf`	10-Formyltetrahydrofolate
`methf`	5,10-Methenyltetrahydrofolate
`mlthf`	5,10-Methylenetetrahydrofolate
`thf`	5,6,7,8-Tetrahydrofolate
`thdp`	2,3,4,5-Tetrahydrodipicolinate
`23dhdp`	2,3-Dihydrodipicolinate
`4pasp`	4-Phospho-L-aspartate
`aspsa`	L-Aspartate 4-semialdehyde
`phom`	O-Phospho-L-homoserine
`pser__L`	O-Phospho-L-serine
`gal1p`	Alpha-D-Galactose 1-phosphate
`13dpg`	3-Phospho-D-glyceroyl phosphate
`prpp`	5-Phospho-alpha-D-ribose 1-diphosphate
`3php`	3-Phosphohydroxypyruvate
`actp`	Acetyl phosphate
`ppi`	Diphosphate

Structural Constraints on Autocatalysis

In the preceding section we have reviewed several ways of formalizing autocatalysis in terms of integer hyperflows. While the comparison of the different approaches provides many open question for future research, it also leaves the

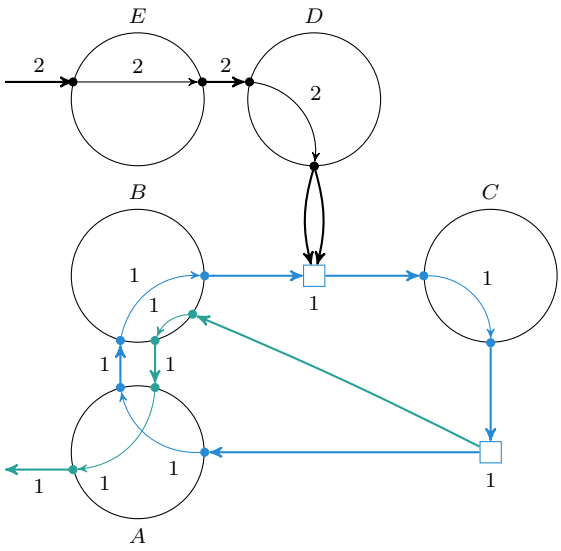


Figure 12: An expanded flow that contains a catalytic cycle (in blue), from which an "ear" (in green), produces an additional copy of the autocatalytic molecule A.

impression that none of them already provides a satisfactory theory. Querying for Milo and Nghe flows, for instance, requires a very loose definition of sources and sinks, and only partially includes structural constraints. Overall autocatalysis provides much more flexibility in the source/sink specification and provides solutions directly interpretable as chemical pathways. However, even with routing constraints on expanded flows the solutions are not guaranteed to have the cyclic motifs we would expect for "true" autocatalysis.

Using the mathematical setup of expanded hypergraphs we can relax the condition of an autocatalytic vertex to have explicit input, and instead require that the flow must induce a cycle that goes through any of the associated vertices in the expanded graph. More formally, for a flow $\widetilde{f}$ on the expanded network $\widetilde{\mathcal{H}} = (\widetilde{V}, \widetilde{E})$, if a vertex $x \in V$ is to be considered autocatalytic then $\mathrm{K}(\widetilde{\mathcal{H}}[\widetilde{V}, \mathrm{supp}(\widetilde{f})])$ must contain a cycle passing through x (indicated in blue in Fig. 12) and an "ear" (indicated in green in Fig. 12) that branches off the cycle before x, rejoins the cycle to pass through x and eventually connects to an outflow, possibly after additional reactions in parallel with the cycle. This "ear" condition may sound deceptively simple, as it directly aligns with the expectation that the cycle must be productive, but providing a formal definition requires careful attention. The cycle condition is mathematically easy to state, but it is a non-local constraint that may require a non-trivial computational effort to handle. We envision that systems such as the one in Fig. 12 will be a paradigmatic example of autocatalytic mechanisms. It is worth noting that at least conceptually this fits with autocatalytic cores of Blokhuis *et al.* [45].

Concluding Remarks

We cannot claim to have a comprehensive mathematical theory of (structural) autocatalysis. However, we have a starting point to develop such a theory and some hints that we can use to guide us into the right direction: integer hyperflows provide a powerful mathematical framework in which *some* of the properties of autocatalytic networks can be expressed very naturally. In addition, it makes the incorporation of stoichiometric balance conditions very easy and natural. On the other hand, flows alone do not seem to be sufficient, since autocatalysis involves pushing material around in a (generalized) cycle, and thus involves a temporal order of reactions that – in general – is not specified completely by a flow, which in essence is just a set of reactions. To this end, we have introduced expanded hypergraphs that encode some of the necessary temporal ordering. Since flows are by construction a description of a steady state, we suspect that flows are an inherently incomplete framework, which need to be complemented by constraints such as the cycle/ear motif sketched in the previous section that imply temporal order of reactions, i.e., *a mechanism*. It remains an interesting mathematical question for future research to what extent routing constraints in extended hypergraphs imply topological orders of reactions for a given flow.

Acknowledgements

We thank Philippe Nghe for stimulating discussions on autocatalysis during a joint visit of CERN in October 2019, and for providing the manuscript "Minimal Autocatalytic Stoichiometries" to us prior to publication. This work was supported in part by the German Federal Ministry of Education and Research (BMBF) within the project Competence Center for Scalable Data Services and Solutions (ScaDS) Dresden/Leipzig (BMBF 01IS14014B). It was also supported by the Independent Research Fund Denmark (DFF-7014-00041.130).

Author Details

[a]Department of Mathematics and Computer Science, University of Southern Denmark, Campusvej 55, Odense M - DK-5230, Denmark
[b]Bioinformatics Group, Department of Computer Science; Interdisciplinary Center for Bioinformatics; German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig; Competence Center for Scalable Data Services and Solutions Dresden-Leipzig; Leipzig Research Center for Civilization Diseases; and Centre for Biotechnology and Biomedicine, University of Leipzig, Härtelstraße 16-18, D-04107 Leipzig, Germany
[c]Max Planck Institute for Mathematics in the Sciences, Inselstraße 22, D-04103 Leipzig, Germany
[d]Institute for Theoretical Chemistry, University of Vienna, Währingerstraße 17, A-1090 Wien, Austria
[e]Facultad de Ciencias, Universidad Nacional de Colombia, Sede Bogotá, Colombia
[f]The Santa Fe Institute, 1399 Hyde Park Rd., Santa Fe, NM 87501, United States

Author Contributions

The authors jointly conceived the study, JLA and PFS developed most of the mathematical framework, JLA performed the computational analysis of the metabolic networks. All authors contributed to the manuscript and approved of its submission.

References

1. Soai, K., Shibata, T., Morioka, H., Choji, K.: Asymmetric autocatalysis and amplification of enantiomeric excess of a chiral molecule. Nature **378**, 767–768 (1995). doi:10.1038/378767a0
2. Ostwald, W.: Über Autokatalyse. Ber Verh Kgl Sächs Ges Wiss Leipzig, Math Phys Class **42**, 189–191 (1890)
3. Bissette, A.J., Fletcher, S.P.: Mechanisms of autocatalysis. Angew. Chem. Int. Ed. **52**, 12800–12826 (2013). doi:10.1002/anie.201303822
4. Schuster, P.: What is special about autocatalysis? Monatsh. Chem. **150**, 763–775 (2019). doi:10.1007/s00706-019-02437-z

5. Bánsági, T., Taylor, A.F.: Ester hydrolysis: Conditions for acid autocatalysis and a kinetic switch. Tetrahedron **73**, 5018–5022 (2017). doi:10.1016/j.tet.2017.05.049
6. Kovács, K.A., Gróf, P., Burai, L., Riedel, M.: Revising the mechanism of the permanganate/oxalate reaction. J. Phys. Chem. A **108**, 11026–11031 (2004). doi:10.1021/jp047061u
7. Kovács, K., Vizvári, B., Riedel, M., Tóth, J.: Decomposition of the permanganate/oxalic acid overall reaction to elementary steps based on integer programming theory. Phys. Chem. Chem. Phys. **6**, 1236–1242 (2004). doi:10.1039/b315211a
8. Benner, S.A., Kim, H.J., Ricardo, A.: Planetary organic chemistry and the origins of biomolecules. Cold Spring Harb Perspect Biol **2**(7), 003467 (2010). doi:10.1101/cshperspect.a003467
9. Breslow, R.: On the mechanism of the formose reaction. Tetrahedron Letters **1**(21), 22–26 (1959). doi:10.1016/S0040-4039(01)99487-0
10. Andersen, J.L., Flamm, C., Merkle, D., Stadler, P.F.: 50 shades of rule composition: From chemical reactions to higher levels of abstraction. In: Fages, F., Piazza, C. (eds.) Formal Methods in Macro-Biology. Lect. Notes Comp. Sci., vol. 8738, pp. 117–135 (2014). doi:10.1007/978-3-319-10398-3_9
11. Butlerow, A.: Bildung einer zuckerartigen Substanz durch Synthese. Justus Liebigs Ann. Chem. **120**, 295–298 (1861). doi:10.1002/jlac.18611200308
12. Treindl, L., Ruoff, P., Kvernberg, P.O.: Influence of oxygen and organic substrate on oscillations and autocatalysis in the Belousov-Zhabotinsky reaction. J. Phys. Chem. A **101**, 4606–4612 (1997). doi:10.1021/jp9705683
13. Gánti, T.: Coupling of autocatalytic cycles as a possible explanation of chemical oscillators. React. Kinet. Catal. Lett. **24**, 197–202 (1984). doi:10.1007/BF02069628
14. Semenov, S.N., Kraft, L.J., Ainla, A., Zhao, M., Baghbanzadeh, M., Campbell, V.E., Kang, K., Fox, J.M., Whitesides, G.M.: Autocatalytic, bistable, oscillatory networks of biologically relevant organic reactions. Nature **537**(7622), 656–660 (2016). doi:10.1038/nature19776
15. Andersen, J.L., Flamm, C., Merkle, D., Stadler, P.F.: *In silico* support for eschenmoser's glyoxylate scenario. Israel J. Chem. **55**(8), 919–933 (2015). doi:10.1002/ijch.201400187
16. Kun, Á., Papp, B., Szathmáry, E.: Computational identification of obligatorily autocatalytic replicators embedded in metabolic networks. Genome Biology **9**, 51 (2008). doi:10.1186/gb-2008-9-3-r51
17. Siami, M., Motee, N., Buzi, G., Bamieh, B., Khammash, M., Doyle, J.C.: Fundamental limits and tradeoffs in autocatalytic pathways. IEEE Trans. Automat. Contr. **65**, 733–740 (2020). doi:10.1109/TAC.2019.2921671
18. Eigen, M.: Selforganization of matter and the evolution of biological macromolecules. Die Naturwissenschaften **10**, 465–523 (1971). doi:10.1007/BF00623322
19. Sievers, D., von Kiedrowski, G.: Self-replication of complementary nucleotide-based oligomers. Nature **369**, 221–224 (1994). doi:10.1038/369221a0
20. Lee, D.H., Granja, J.R., Martinez, J.A., Severin, K., Ghadiri, M.R.: A self-replicating peptide. Nature **382**, 525–588 (1996). doi:10.1038/382525a0
21. Segré, D., Ben-Eli, D., Deamer, D., Lancet, D.: The lipid world. Origins Life Evol. Biosphere **31**, 119–145 (2001). doi:10.1023/A:1006746807104
22. Gánti, T.: Chemical systems and supersystems I: Chemical cycles. Acta Chim Acad Sci Hung **91**, 357–368 (1976)
23. Gánti, T.: Chemical systems and supersystems II: Stoichiometry of self-reproducing chemical systems. Acta Chim Acad Sci Hung **91**, 369–385 (1976)
24. Kauffman, S.A.: Autocatalytic sets of proteins. J. Theor. Biol. **119**, 1–24 (1986). doi:10.1016/S0022-5193(86)80047-9
25. Kaleta, C., Centler, F., Dittrich, P.: Analyzing molecular reaction networks: from pathways to chemical organizations. Mol Biotechnol **34**, 117–123 (2006). doi:10.1385/MB:34:2:117
26. Benkö, G., Centler, F., Dittrich, P., Flamm, C., Stadler, B.M.R., Stadler, P.F.: A topological approach to chemical organizations. Alife **15**, 71–88 (2009). doi:10.1162/artl.2009.15.1.15105
27. Eigen, M., Schuster, P.: The Hypercycle: A Principle of Natural Self-Organization. Springer, Heidelberg (1979). doi:10.1007/978-3-642-67247-7
28. Stadler, P.F., Schuster, P.: Mutation in autocatalytic networks — an analysis based on perturbation theory. J. Math. Biol. **30**, 597–631 (1992). doi:10.1007/BF00948894
29. Stadler, B.M.R., Stadler, P.F., Schuster, P.: Dynamics of autocatalytic replicator networks based on higher order ligation reactions. Bull. Math. Biol. **62**, 1061–1086 (2000). doi:10.1006/bulm.2000.0194
30. Steel, M.: The emergence of a self-catalysing structure in abstract origin-of-life models. Appl. Math. Letters **13**, 91–95 (2000). doi:10.1016/S0893-9659(99)00191-3
31. Hordijk, W., Steel, M.: Detecting autocatalytic, self-sustaining sets in chemical reaction systems. J. Theor. Biol. **227**, 451–461 (2004). doi:10.1016/j.jtbi.2003.11.020
32. Hordijk, W., Steel, M., Dittrich, P.: Autocatalytic sets and chemical organizations: modeling self-sustaining reaction networks at the origin of life. New J Phys. **20**, 015011 (2018). doi:10.1088/1367-2630/aa9fcd
33. Virgo, N., Ikegami, T., McGregor, S.: Complex autocatalysis in simple chemistries. Artif. Life **22**, 138–152 (2016). doi:10.1162/ARTL_a_00195
34. Liu, Y., Sumpter, D.J.T.: Mathematical modeling reveals spontaneous emergence of self-replication in chemical reaction systems. J Biol Chem **293**, 18854–18863 (2018). doi:10.1074/jbc.RA118.003795
35. Hordijk, W., Smith, J.I., Steel, M.: Algorithms for detecting and analysing autocatalytic sets. Algorithms Mol Biol **10**, 15 (2015). doi:10.1186/s13015-015-0042-8
36. Steel, M., Hordijk, W., Xavier, J.C.: Autocatalytic networks in biology: structural theory and algorithms. J. R. Soc. Interface **16**, 20180808 (2019). doi:10.1098/rsif.2018.0808
37. Xavier, J.C., Hordijk, W., Kauffman, S., Steel, M., Martin, W.F.: Autocatalytic chemical networks at the origin of metabolism. Proc. Roy. Soc. B **287**, 20192377 (2020). doi:10.1098/rspb.2019.2377
38. Andersen, J.L., Flamm, C., Merkle, D., Stadler, P.F.: Chemical transformation motifs — Modelling pathways as integer hyperflows. IEEE/ACM Trans. Comp. Biol. **16**, 510–523 (2019). doi:10.1109/TCBB.2017.2781724
39. King, G.A.M.: Autocatalysis. Chem. Soc. Reviews **7**, 297–316 (1978). doi:10.1039/CS9780700297
40. Barenholz, U., Davidi, D., Reznik, E., Bar-On, Y., Antonovsky, N., Noor, E., Milo, R.: Design principles of autocatalytic cycles constrain enzyme kinetics and force low substrate saturation at flux branch points. eLife **6**, 20667. doi:10.7554/eLife.20667
41. Hoffman, A.J.: A generalization of Max Flow-Min Cut. Math. Programming **6**, 352–359 (1974). doi:10.1007/BF01580250
42. Schuster, S., Hilgetag, C.: On elementary flux modes in biochemical reactions systems at steady state. J. Biol. Syst. **2**, 165–182 (1994). doi:10.1142/S0218339094000131
43. Steffen, K., Regensburger, G., Gerstl, M.P., Jungreuthmayer, C., Schuster, S., Mahadevan, R., Zanghellini, J., Müller, S.: From elementary flux modes to elementary flux vectors: Metabolic pathway analysis with arbitrary linear flux constraints. PLoS Comput Biol. **13**, 1005409 (2017). doi:10.1371/journal.pcbi.1005409
44. Handorf, T., Ebenhöh, O., Heinrich, R.: Expanding metabolic networks: Scopes of compounds, robustness, and evolution. J. Mol. Evol. **61**, 498–512 (2005). doi:10.1007/s00239-005-0027-1
45. Blokhuis, A., Lacoste, D., Nghe, P.: Autocatalysis in Chemical Networks: Unifications and Extensions. ChemRxiv (2020). doi:10.26434/chemrxiv.12317273.v1
46. Koch, I.: Petri nets – a mathematical formalism to analyze chemical reaction networks. Molecular Informatics **29**, 838–843 (2010). doi:10.1002/minf.201000086
47. Andersen, J.L., Flamm, C., Merkle, D., Stadler, P.F.: A software package for chemically inspired graph transformation. In: Echahed, R., Minas, M. (eds.) Graph Transformation - 9th International Conference, ICGT 2016, Proceedings. Lect. Notes Comp. Sci., vol. 9761, pp. 73–88. Springer, Heidelberg (2016). doi:10.1007/978-3-319-40530-8_5
48. Piedrafita, G., Montero, F., Morán, F., Cárdenas, M.L., Cornish-Bowden, A.: A simple self-maintaining metabolic system: Robustness, autocatalysis, bistability. PLoS Comp. Biol. **6**, 1000872 (2010). doi:10.1371/journal.pcbi.1000872
49. Sousa, F.L., Hordijk, W., Steel, M., Martin, W.F.: Autocatalytic sets in *E. coli* metabolism. J Syst Chem **6**, 4 (2015). doi:10.1186/s13322-015-0009-7
50. Schellenberger, J., Park, J.O., Conrad, T.M., Palsson, B.T.: BiGG: A biochemical genetic and genomic knowledgebase of large scale metabolic reconstructions. BMC Bioinformatics **11**, 213 (2010). doi:10.1186/1471-2105-11-213