Reachability, Persistence, and Constructive Chemical Reaction Networks (Part III): A Mathematical Formalism for Binary Enzymatic Networks and Application to Persistence

Gilles Gnacadja

Research and Development Information Systems, Amgen, Inc.
One Amgen Center Drive, Thousand Oaks, California 91320-1799, USA
gilles.gnacadja@gmail.com

Published in *Journal of Mathematical Chemistry*
DOI 10.1007/s10910-011-9895-3
Preprint revision E.04 31 August 2011

Abstract

Chemical Reaction Network Theory is concerned with understanding the properties of systems of reactions from their structure. Enzymatic networks receive significant attention in the field because they are crucial in biochemistry and often illustrate the network features that are studied. In this paper we propose a formalism for binary enzymatic networks which can be used to research their mathematical properties. The networks are binary in that every enzyme-substrate complex consists of one enzyme and one substrate. Many connected concepts, e.g. futile enzymatic cycles and enzymatic cascades, are defined rigorously and so as to reflect the corresponding biochemical phenomena. We prove that binary enzymatic networks that are futile and cascaded are vacuously persistent: no species will tend to extinction if all species are implicitly present at initial time. This result extends prior work of Angeli, De Leenheer and Sontag in which a theorem was applied to show that certain individual enzymatic networks are persistent. This paper completes a series of three articles. It applies both the first paper which studies vacuous persistence and the second paper which describes a formalism for species composition.

Keywords. Binary Enzymatic Networks; Futile Enzymatic Cycle; Enzymatic Cascade; Vacuous Persistence; Reachability; Chemical Reaction Network.

Mathematics Subject Classification (2010): 92C42, 92C45, 34D05.

1 Introduction

Enzymes are molecules that catalyze, i.e. accelerate, the conversion of certain molecules, the substrates, into other molecules, the products. They are fundamental to countless biochemical processes. There is abundant work using mathematics to gain insight into the properties of enzymatic networks. Enzyme kinetics in particular is extensively researched and discussed in references such as Cornish-Bowden [2], Salazar and Höfer [8], and numerous others. There is also research concerned primarily with equilibrium and limit states, specifically their numbers, their parameterizations and their asymptotic properties. Examples include but are not limited to Craciun, Tang and Feinberg [3], Angeli, De Leenheer and Sontag [1], Wang and Sontag [11], Thomson and Gunawardena [10], and Pérez Millán, Dickenstein, Shiu and Conradi [7].

In this paper we propose a mathematical definition of binary enzymatic networks with two intended goals. We seek to faithfully represent the biochemical mechanisms in which one enzyme and one substrate bind into an intermediate enzymatic complex which, possibly after isomerization, dissociates into the same enzyme and one product, which may be identical to or different than the substrate. At the same time, we want to facilitate the mathematical deduction of the properties of such networks. We build upon reaction networks as they are classically defined in Chemical Reaction Network Theory. A binary enzymatic network will be a reaction network that satisfies five additional conditions referred to as Conditions (Enz1)-(Enz5). We deliberately omit a number of biochemically important mechanisms so as to focus the discussions. In particular, we do not represent the simultaneous or stepwise binding of several substrates onto an enzyme, and the simultaneous or stepwise dissociation of several products from an enzyme. This specific restriction is what makes binary the enzymatic networks we consider.

We define several related concepts which represent features observed in enzyme chemistry. In particular, we define futility, which is the phenomenon whereby every enzyme performs actions that reverse the actions of some other enzyme. We also define cascades, which are schemes in which products of enzymatic reactions may serve as enzymes in other enzymatic reactions.

This paper is the last in a series of three articles that investigate the persistence of reaction networks. In the first article, Gnacadja [4], we introduce

and prove a structural characterization of vacuous persistence, which is the property that no species tend to extinction if all species are implicitly present at initial time. In the second article, Gnacadja [5], we develop a formalism for species composition and use it to find a class of biochemically relevant networks that are vacuously persistent. Both works are used here to prove the following result.

Theorem 1.1 (Theorem 6.7). *If a binary enzymatic network is futile and cascaded, then it is vacuously persistent.* □

This theorem covers the examples of enzymatic networks that were found to be persistent in Angeli, De Leenheer and Sontag [1, Sections 6.1-6.3]. We think that this theorem and the methods employed to obtain it suggest further potential for the formalism we develop.

Ours is not the first effort to formulate a mathematical definition of enzymatic networks. We note in particular the formalism of Thomson and Gunawardena [10]. There are stylistic differences between the two approaches due in part to the intended applications and the methods used to pursue them. For instance, our formalism allows from the onset a species to be in the two roles of enzyme and substrate/product.

The rest of the paper consists of five sections. Section 2 presents the formal definition of a binary enzymatic network and Section 3 illustrates this with three examples. Section 4 formalizes the notions of futility and cascades. In Section 5, we process binary enzymatic networks through the concepts related to species composition studied in Gnacadja [5]. Finally we establish the persistence result in Section 6. Background material not elaborated on here can be found in the first two articles of this three-part series, Gnacadja [4] and [5]. In particular, Section 3 of Gnacadja [4] provides the basics on reaction networks.

2 Structure of Binary Enzymatic Networks

The simplest enzymatic reaction has the form

$$E + A \to EA \to E + B$$

where E is the enzyme, A is the substrate, B is the product, and EA is the intermediate. The enzyme enables or accelerates the conversion of the substrate into the product through the formation and dissociation of the

intermediate. This process can be more elaborate in a number of ways, such as:

- The dissociation of the intermediate EA may create more than one product, and the substrate A may be one of the products;
- The intermediate EA may convert into other intermediates, which either convert into yet other intermediates or dissociate into the enzyme E and other products;
- There could be enzymes that revert the actions of the enzyme E;
- Some of the products could act as enzymes in other reactions.

Illustrated examples of intricacies that may occur can be seen in Thomson and Gunawardena [10, Figure 1] and Salazar and Höfer [8, Figure 2]. We also have here the enzymatic networks of Section 3 and Figure 6.1.

The definition of binary enzymatic networks we are about to formulate will capture the features discussed above and more. We proceed by formulating a series of conditions which will lead to the definition. While working on this definition and its implications, we found it motivating and rewarding to bear in mind that

> *Half the battle in understanding is having the right representation.*
> Attributed to Pierre-Simon Laplace.

The understanding here is that of enzymatic networks as a class of reaction networks. Research work on specific preselected enzymatic networks may not require this formalism. A reaction network $\mathscr{N} = (\mathscr{S}, \mathscr{C}, \mathscr{R})$ is fixed for this section.

Condition (Enz1). Four proper and nonempty subsets Enz, Sub, Pro, Int of $\mathscr{S}$ are given and we have $\mathsf{Enz} \cup \mathsf{Sub} \cup \mathsf{Pro} = \mathscr{S} \backslash \mathsf{Int}$. □

The species in Enz, Sub, Pro, Int are respectively the *enzymes*, the *substrates*, the *products* and the *intermediates*. We collectively refer to the substrates and the products as the *enzyme partners* or simply the *partners*. We pose

$$\mathsf{Par} := \mathsf{Sub} \cup \mathsf{Pro} \quad \text{and} \quad \mathsf{Enz}_0 := \mathsf{Enz} \backslash \mathsf{Par} \,.$$

These are respectively the set of partners and the set of enzymes that are not partners. We allow species to be both enzyme and partner, so Enz_0 may

be a proper subset of Enz. But intermediates may be neither enzymes nor partners. We have

$$\mathscr{S} = \mathsf{Enz}_0 \sqcup \mathsf{Par} \sqcup \mathsf{Int} \,.$$

Condition (Enz2). A subset $\mathsf{Cat} \subseteq \mathsf{Enz} \times \mathsf{Sub} \times \mathsf{Pro}$ of *catalysis triples* is given. Every enzyme occurs in some catalysis triple, so does every substrate and so does every product. □

A catalysis triple (E, A, B) indicates that the enzyme E catalyzes the conversion of the substrate A into the product B. It does not record the arrangement of intermediates that achieves this conversion.

A *substrate-product pair* is any $(A, B) \in \mathsf{Sub} \times \mathsf{Pro}$ such that (E, A, B) is a catalysis triple for some enzyme E. We introduce the *partner graph* $\mathsf{ParGraph}$, the directed graph for which the set of vertices is the set $\mathsf{Par} = \mathsf{Sub} \cup \mathsf{Pro}$ of enzyme partners and the set of edges is the set of substrate-product pairs in which substrate and product are not the same. We call *undirected partner graph* and denote $\underline{\mathsf{ParGraph}}$ the corresponding undirected graph. We equip the set Par with the equivalence relation whose equivalence classes are the connected components of the graph $\underline{\mathsf{ParGraph}}$. The quotient map is $\mathsf{cl} : \mathsf{Par} \to \overline{\mathsf{Par}}$.

The (directed) partner graph $\mathsf{ParGraph}$ is obtained by simplifying the original network in a process that eliminates not only the various arrangements of intermediates that achieve the conversions, but also the enzymes that catalyze them. This process deletes any pattern of the form

$$E + A \;\to\; (\text{network of intermediates}) \;\to\; E + A$$

and transforms any pattern of the form

$$E + A \;\to\; (\text{network of intermediates}) \;\to\; E + B$$

with $A \neq B$ into the simplified and hypothetical isomerization reaction

$$A \;\to\; B \,.$$

Thus, the partner graph $\mathsf{ParGraph}$ is the reaction network of isomerization reactions which records the core mechanisms, i.e. the substrate-to-product conversions devoid of any information as to how they occur. The undirected partner graph $\underline{\mathsf{ParGraph}}$ erases any distinctions between substrates and products and lets us assemble the partners into equivalence (linkage) classes.

Condition (Enz3). A surjective map $\mathsf{par} : \mathsf{Enz} \to \overline{\mathsf{Par}}$ is given. We have $E \notin \mathsf{par}(E)$ for every enzyme $E \in \mathsf{Enz}$ and $A, B \in \mathsf{par}(E)$ for every catalysis triple $(E, A, B) \in \mathsf{Cat}$. □

For an enzyme E, $\mathsf{par}(E)$ is the equivalence class of partners of E. Each class of partners must be matched with one or several enzymes in this way. The requirement $E \notin \mathsf{par}(E)$ says that an enzyme may not be partner with itself. In particular, there may not be autocatalysis.

For $E \in \mathsf{Enz}$, we set

$$\mathsf{sub}(E) := \mathsf{par}(E) \cap \mathsf{Sub} \quad \text{and} \quad \mathsf{pro}(E) := \mathsf{par}(E) \cap \mathsf{Pro} \; .$$

These are the sets of partners in $\mathsf{par}(E)$ that are substrates and products respectively. Since a partner must be one or the other (or both), we have

$$\mathsf{par}(E) = \mathsf{sub}(E) \cup \mathsf{pro}(E) \; .$$

And since the map $\mathsf{par} : \mathsf{Enz} \to \overline{\mathsf{Par}}$ is surjective and the set $\overline{\mathsf{Par}}$ is a partition of the set Par, we have

$$\mathsf{Sub} = \bigcup_{E \in \mathsf{Enz}} \mathsf{sub}(E) \quad \text{and} \quad \mathsf{Pro} = \bigcup_{E \in \mathsf{Enz}} \mathsf{pro}(E) \; .$$

Condition (Enz4). Given are an equivalence relation on Int with quotient map $\mathsf{cl} : \mathsf{Int} \to \overline{\mathsf{Int}}$ and two mutually inverse bijective maps $\mathsf{int} : \mathsf{Enz} \to \overline{\mathsf{Int}}$ and $\mathsf{enz} : \overline{\mathsf{Int}} \to \mathsf{Enz}$. □

The equivalence relation with quotient map $\mathsf{cl} : \mathsf{Int} \to \overline{\mathsf{Int}}$ partitions the intermediates according to the enzyme that catalyzes the conversions in which they occur. For an enzyme $E \in \mathsf{Enz}$, $\mathsf{int}(E)$ is the class of intermediates in the conversions that are catalyzed by E. For a class of intermediates $\mathscr{Y} \in \overline{\mathsf{Int}}$, $\mathsf{enz}(\mathscr{Y})$ is the enzyme that catalyzes the conversions in which those intermediates occur.

The class map cl has been defined for the partners following the definition of the undirected partner graph $\overline{\mathsf{ParGraph}}$. It has just been defined for the intermediates. For a non-partner enzyme $E \in \mathsf{Enz}_0$, we set $\mathsf{cl}(E) := \{E\}$. Also, we denote $\overline{\mathsf{Enz}_0}$ the set of singletons of elements of Enz_0. The map cl is now defined for all species as the juxtaposition of the three maps $\mathsf{cl} : \mathsf{Enz}_0 \to \overline{\mathsf{Enz}_0}$, $\mathsf{cl} : \mathsf{Par} \to \overline{\mathsf{Par}}$ and $\mathsf{cl} : \mathsf{Int} \to \overline{\mathsf{Int}}$. We explained above what the map cl does to the intermediates. As for the enzymes and the partners, it will be seen in

Section 5 that the map cl partitions them (they are the elementary species) into isomerism classes.

Given a catalysis triple $(E, A, B) \in \mathsf{Cat}$, we shall call an *intermediates path* of (E, A, B) any finite nonempty tuple $(Y_1, \ldots, Y_\ell)$ of intermediates in $\mathsf{int}(E)$ such that the following $\ell + 1$ reactions are in the network.

$$E + A \to Y_1 \to Y_2 \to \cdots \to Y_\ell \to E + B$$

We denote $\mathscr{R}_{(Y_1,\ldots,Y_\ell)}(E, A, B)$ the set consisting of these $\ell + 1$ reactions.

Condition (Enz5). For every catalysis triple $(E, A, B) \in \mathsf{Cat}$, we are given a nonempty set $\mathsf{IntPath}(E, A, B)$ of intermediates paths such that the set $\mathscr{R}$ of all reactions is given by

$$\mathscr{R} \;=\; \bigcup_{(E,A,B)\in\mathsf{Cat}} \;\bigcup_{(Y_1,\ldots,Y_\ell)\in\mathsf{IntPath}(E,A,B)} \mathscr{R}_{(Y_1,\ldots,Y_\ell)}(E, A, B)\,.$$

□

This condition precisely prescribes the reactions. Note that the intermediates paths belonging to the same catalysis triple need not be of the same length and may share nodes.

Definition 2.1. The reaction network $\mathscr{N} = (\mathscr{S}, \mathscr{C}, \mathscr{R})$ is a *binary enzymatic network* provided the five conditions (Enz1)-(Enz5) are satisfied. □

This definition is illustrated with examples in Section 3. The definition is formulated so as to start with a "naive" reaction network and progressively reveal the additional features that together make it a binary enzymatic network.

- Condition (Enz1) assigns the species with the roles of enzyme, substrate, product and intermediate.
- Condition (Enz2) records the substrate-to-product conversions that occur along with the enzymes that catalyze them, but not (yet) the intervening steps.
- Condition (Enz3) constrains how the enzymes relate to substrates and products.
- Condition (Enz4) constrains how intermediates relate to enzymes.
- Condition (Enz5) specifies the reactions.

It could be tempting to think that Conditions (Enz1), (Enz2) and (Enz5) are sufficient: Condition (Enz1) tells us the roles of the species, Condition (Enz2) tells what the catalyzed conversions are, and Condition (Enz5) specifies the reactions that make these conversions happen. The purpose of Conditions (Enz3) and (Enz4) is to ensure that overlaps of various paths in the network occur in accordance with what we observe in actual enzymatic networks. Figure 2.1 presents a global view of some of the maps used in the conditions.

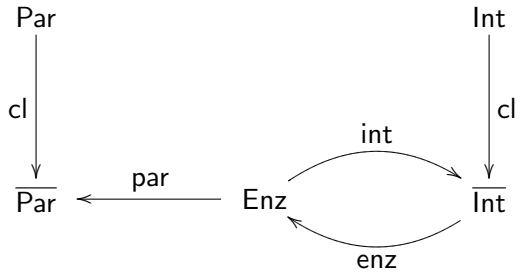


Figure 2.1: Collective diagrammatic view of some of the maps in the Conditions (Enz1)-(Enz5) which define a binary enzymatic network. The two maps between Enz and $\overline{\mathsf{Int}}$ are bijective and the inverse of each other. The other three maps are surjective.

Enzyme-catalyzed mechanisms may be distributive or processive. These concepts are reviewed in Salazar and Höfer [8] with illustrations on Figure 2 therein. See also Gunawardena [6]. Definition 2.1 accommodates both distributive and processive mechanisms. Also, because enzymes are allowed to be partners of other enzymes, Definition 2.1 accommodates cascaded mechanisms, as is seen in the example of Section 3.2 and more generally in Section 4. We expect that for a network of enzyme-catalyzed conversions of substrates into products to not be covered by Definition 2.1, there must be intermediate species that are at least ternary in terms of their enzyme and substrate constituents.

3 Examples of Binary Enzymatic Networks

We present three examples of binary enzymatic networks and make explicit how they are instances of Definition 2.1.

3.1 The Simplest Futile Enzymatic Cycle

$$E + A \rightleftharpoons EA \to E + B$$
$$F + B \rightleftharpoons FB \to F + A$$

This network is the simplest futile enzymatic cycle. It is a futile cycle in that the two enzymes interconvert the two substrates. Section 4 formally defines futility. We now list the attributes that make this network a binary enzymatic network in the sense of Definition 2.1.

$$\mathsf{Enz} = \{E, F\} \qquad \mathsf{Sub} = \{A, B\} \qquad \mathsf{Par} = \{A, B\}$$
$$\mathsf{Enz}_0 = \{E, F\} \qquad \mathsf{Pro} = \{A, B\} \qquad \mathsf{Int} = \{EA, FB\}$$

$$\mathsf{Cat} = \big\{(E, A, A), (E, A, B), (F, B, B), (F, B, A)\big\}$$

$$\mathsf{ParGraph} : \quad A \rightleftarrows B$$

$$\overline{\mathsf{Par}} = \big\{\{A, B\}\big\} \qquad \overline{\mathsf{Int}} = \big\{\{EA\}, \{FB\}\big\}$$

$$\mathsf{cl}(E) = \{E\} \qquad \mathsf{cl}(A) = \{A, B\} \qquad \mathsf{cl}(EA) = \{EA\}$$
$$\mathsf{cl}(F) = \{F\} \qquad \mathsf{cl}(B) = \{A, B\} \qquad \mathsf{cl}(FB) = \{FB\}$$

$$\mathsf{sub}(E) = \{A\} \qquad \mathsf{pro}(E) = \{A, B\}$$
$$\mathsf{sub}(F) = \{B\} \qquad \mathsf{pro}(F) = \{A, B\}$$

$$\mathsf{par}(E) = \{A, B\} \qquad \mathsf{int}(E) = \{EA\} \qquad \mathsf{enz}\big(\{EA\}\big) = E$$
$$\mathsf{par}(F) = \{A, B\} \qquad \mathsf{int}(F) = \{FB\} \qquad \mathsf{enz}\big(\{FB\}\big) = F$$

$$\mathsf{IntPath}(E, A, A) = \mathsf{IntPath}(E, A, B) = \big\{(EA)\big\}$$
$$\mathsf{IntPath}(F, B, B) = \mathsf{IntPath}(F, B, A) = \big\{(FB)\big\}$$

3.2 A Cascade of Three Simple Enzymatic Conversions

$$\begin{array}{ccccc} E_0 + S_0 & \rightleftharpoons & Y_0 & \to & E_0 + E_1 \\ & & & & + \\ & & & & S_1 \\ & & & & \Uparrow\Downarrow \\ & & & & Y_1 \\ & & & & \downarrow \\ & & & & E_1 \\ & & & & + \\ & & & & E_2 + S_2 \quad \rightleftharpoons \quad Y_2 \quad \to \quad E_2 + E_3 \end{array}$$

This network is a cascade of three simple enzymatic conversions. It is a cascade because the product of the first conversion is the enzyme in the second conversion, and the product of the second conversion is the enzyme in the third conversion. Section 4 formally defines cascades. In an example such as this one, each enzymatic conversion is usually accompanied with another so that the pair forms a futile cycle (see Section 3.1). This is omitted here to keep the illustration simple. We list the attributes that make this network a binary enzymatic network in the sense of Definition 2.1.

$$\mathsf{Enz} = \{E_0, E_1, E_2\} \qquad \mathsf{Sub} = \{S_0, S_1, S_2\} \qquad \mathsf{Par} = \mathsf{Pro}$$
$$\mathsf{Enz}_0 = \{E_0\} \qquad \mathsf{Pro} = \{S_0, S_1, S_2, E_1, E_2, E_3\} \qquad \mathsf{Int} = \{Y_0, Y_1, Y_2\}$$

$$\mathsf{Cat} = \bigcup_{i=0,1,2} \big\{(E_i, S_i, S_i), (E_i, S_i, E_{i+1})\big\}$$

$$\mathsf{ParGraph}: \qquad S_0 \longrightarrow E_1 \qquad S_1 \longrightarrow E_2 \qquad S_2 \longrightarrow E_3$$

$$\overline{\mathsf{Par}} = \big\{\{S_0, E_1\}, \{S_1, E_2\}, \{S_2, E_3\}\big\} \qquad \overline{\mathsf{Int}} = \big\{\{Y_0\}, \{Y_1\}, \{Y_2\}\big\}$$

In the following, $i = 0, 1, 2$.

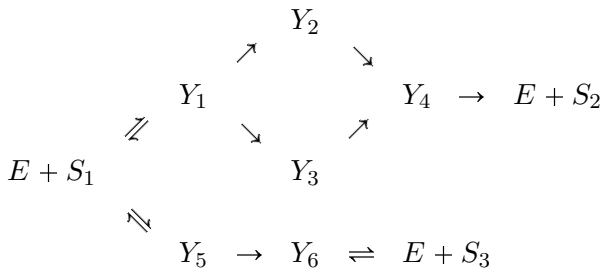
$$\mathsf{cl}(E_0) = \{E_0\} \qquad \mathsf{cl}(S_i) = \mathsf{cl}(E_{i+1}) = \{S_i, E_{i+1}\} \qquad \mathsf{cl}(Y_i) = \{Y_i\}$$

$$\mathsf{sub}(E_i) = \{S_i\} \qquad \mathsf{pro}(E_i) = \mathsf{par}(E_i) = \{S_i, E_{i+1}\}$$

$$\mathsf{int}(E_i) = \{Y_i\} \qquad \mathsf{enz}\big(\{Y_i\}\big) = E_i$$

$$\mathsf{IntPath}(E_i, S_i, S_i) = \mathsf{IntPath}(E_i, S_i, E_{i+1}) = \big\{(Y_i)\big\}$$

3.3 A Hypothetical Enzymatic Network with Parallel Paths



This is a hypothetical enzymatic network obtained by assembling some of the examples of Thomson and Gunawardena [10, Figure 1]. Following are

the attributes that make this network a binary enzymatic network in the sense of Definition 2.1.

$$\begin{array}{lll} \mathsf{Enz} = \{E\} & \mathsf{Sub} = \{S_1, S_3\} & \mathsf{Par} = \mathsf{Pro} \\ \mathsf{Enz}_0 = \{E\} & \mathsf{Pro} = \{S_1, S_2, S_3\} & \mathsf{Int} = \{Y_1, Y_2, Y_3, Y_4, Y_5, Y_6\} \end{array}$$

$$\mathsf{Cat} = \big\{(E, S_1, S_1), (E, S_1, S_2), (E, S_1, S_3), (E, S_3, S_3)\big\}$$

$$\mathsf{ParGraph}: \quad S_1 \to S_2, \quad S_1 \to S_3$$

$$\overline{\mathsf{Par}} = \big\{\mathsf{Par}\big\} \qquad \overline{\mathsf{Int}} = \big\{\mathsf{Int}\big\}$$

$$\begin{array}{lll} \mathsf{sub}(E) = \mathsf{Sub} & \mathsf{int}(E) = \mathsf{Int} & \mathsf{cl}(E) = \{E\} \\ \mathsf{pro}(E) = \mathsf{Pro} & \mathsf{enz}(\mathsf{Int}) = E & \mathsf{cl}(S_i) = \mathsf{Par}\,,\ i = 1, 2, 3 \\ \mathsf{par}(E) = \mathsf{Par} & & \mathsf{cl}(Y_j) = \mathsf{Int}\,,\ j = 1, \ldots, 6 \end{array}$$

$$\begin{aligned} \mathsf{IntPath}(E, S_1, S_1) &= \big\{(Y_1), (Y_5)\big\} \\ \mathsf{IntPath}(E, S_1, S_2) &= \big\{(Y_1, Y_2, Y_4), (Y_1, Y_3, Y_4)\big\} \\ \mathsf{IntPath}(E, S_1, S_3) &= \big\{(Y_5, Y_6)\big\} \\ \mathsf{IntPath}(E, S_3, S_3) &= \big\{(Y_6)\big\} \end{aligned}$$

4 Futile and Cascaded Networks

Futility and cascadedness are two important properties of enzymatic networks which we formalize in this section. We fix a binary enzymatic network $\mathscr{N} = (\mathscr{S}, \mathscr{C}, \mathscr{R})$ and we use the notations of Section 2.

For each enzyme $E \in \mathsf{Enz}$, let

$$\mathscr{C}(E) \;:=\; \mathsf{int}(E) \sqcup \big\{E + A : A \in \mathsf{par}(E)\big\}\,.$$

By Condition (Enz5), the set $\mathscr{C}$ of complexes is given by

$$\mathscr{C} \;=\; \bigsqcup_{E \in \mathsf{Enz}} \mathscr{C}(E)\,.$$

Definition 4.1. Let $E \in \mathsf{Enz}$ be an enzyme.

- The set $\mathsf{isub}(E) \subseteq \mathsf{sub}(E)$ of *initial substrates* of E is defined as follows: for $A \in \mathsf{sub}(E)$, we have $A \in \mathsf{isub}(E)$ if and only if the complex $E + A$ ultimately reacts to every complex in $\mathscr{C}(E)$.

- The set $\mathsf{tpro}(E) \subseteq \mathsf{pro}(E)$ of *terminal products* of E is defined as follows: for $B \in \mathsf{pro}(E)$, we have $B \in \mathsf{tpro}(E)$ if and only if every complex in $\mathscr{C}(E)$ ultimately reacts to the complex $E + B$. □

Definition 4.1 is pertinent because initial substrates and terminal products possess reachability features we use in Section 6. We illustrate these notions for the networks of Sections 3.1, 3.2 and 3.3 respectively:

- $\mathsf{isub}(E) = \mathsf{tpro}(F) = \{A\}$ and $\mathsf{isub}(F) = \mathsf{tpro}(E) = \{B\}$;
- $\mathsf{isub}(E_i) = \{S_i\}$ and $\mathsf{tpro}(E_i) = \{E_{i+1}\}$ for $i = 0, 1, 2$; and
- $\mathsf{isub}(E) = \{S_1\}$ and $\mathsf{tpro}(E) = \varnothing$.

One can readily observe the following from Definition 4.1.

Remark 4.2. Let $E \in \mathsf{Enz}$ be an enzyme.

- If $\mathsf{isub}(E) \neq \varnothing$, then $\mathsf{par}(E) = \mathsf{isub}(E) \cup \mathsf{pro}(E)$.
- If $\mathsf{tpro}(E) \neq \varnothing$, then $\mathsf{par}(E) = \mathsf{sub}(E) \cup \mathsf{tpro}(E)$. □

Definition 4.3. An enzyme F is a *reversing enzyme* for an enzyme E if $\varnothing \neq \mathsf{tpro}(E) = \mathsf{isub}(F)$. The network $\mathscr{N}$ is *futile* if every enzyme is a reversing enzyme. □

Partner classes either are disjoint or coincide, so:

Remark 4.4. Suppose that an enzyme F is a *reversing enzyme* for an enzyme E. Then $\mathsf{par}(E) = \mathsf{par}(F)$. Furthermore, with Remark 4.2, each species in this partner class is both a substrate and a product. So in a futile network, every substrate is a product and every product is a substrate. □

Our definition of a futile network is sufficient for the intended use. But it often also holds that every enzyme has a reversing enzyme. In fact, enzymes often occur in pairs of mutually reversing enzymes, whence the following definition.

Definition 4.5. A *futility involution* of the network $\mathscr{N}$ is a map $\varphi : \mathsf{Enz} \to \mathsf{Enz}$ such that $\varphi^2 = \varphi \circ \varphi = \mathrm{Id}_{\mathsf{Enz}}$ and for every enzyme E, $\varphi(E)$ is a reversing enzyme for E. □

The networks of Section 3.1 and Figure 6.1 are binary enzymatic networks that are futile and each has a futility involution.

We relate futility to the partner graph $\mathsf{ParGraph}$.

Remark 4.6. Recall that the partner graph $\mathsf{ParGraph}$ may be regarded as a reaction network of isomerization reactions. Let E and F be enzymes.

- For any $A \in \mathsf{par}(E)$ and $C \in \mathsf{tpro}(E)$, A ultimately reacts to C in $\mathsf{ParGraph}$.
- For any $B \in \mathsf{par}(F)$ and $C \in \mathsf{isub}(F)$, C ultimately reacts to B in $\mathsf{ParGraph}$.

Therefore:

- If F is a reversing enzyme for E, then $\mathsf{par}(F)$ (which coincides with $\mathsf{par}(E)$) is a strongly connected component of $\mathsf{ParGraph}$.
- If the network $\mathscr{N}$ is futile, then $\mathsf{ParGraph}$ is weakly reversible. □

We now turn our attention to enzymatic cascades. These are networks in which there are species in the dual roles of product and enzyme. Recall that the set Enz_m for $m = 0$ is already defined as

$$\mathsf{Enz}_0 \quad = \quad \mathsf{Enz} \setminus \mathsf{Par} \,.$$

We define the sets Enz_m for $m \in \mathbb{Z}_{\geqslant 1}$ by induction as follows.

$$\mathsf{Enz}_m \quad = \quad \Big(\mathsf{Enz} \setminus \big(\mathsf{Enz}_0 \cup \cdots \cup \mathsf{Enz}_{m-1}\big)\Big) \cap \bigcup_{E \in \mathsf{Enz}_{m-1}} \mathsf{tpro}(E) \,.$$

The sets Enz_m for $m \in \mathbb{Z}_{\geqslant 0}$ are pairwise disjoint and the set Enz is finite, so there exists $m_0 \in \mathbb{Z}_{\geqslant 0}$ such that $\mathsf{Enz}_m = \varnothing$ for $m > m_0$.

Definition 4.7. The network $\mathscr{N}$ is *cascaded* if $\mathsf{Enz} = \bigsqcup_{m=0}^{\infty} \mathsf{Enz}_m$.
An enzyme $E \in \mathsf{Enz}_m$ is said to have *cascade index* $\gamma(E) = m$. □

Note that if $\mathscr{N}$ is cascaded, then $\mathsf{Enz}_0 \neq \varnothing$. This is because $\mathsf{Enz} \neq \varnothing$ and it holds that $\mathsf{Enz}_m = \varnothing \Rightarrow \mathsf{Enz}_{m+1} = \varnothing$.

If $\mathsf{Enz} = \mathsf{Enz}_0$, then the network is cascaded in a trivial way: all enzymes have cascade index zero. This is the case for the networks of Sections 3.1 and 3.3. For the network of Section 3.2, one can verify that it is cascaded with the enzymes E_0, E_1, E_2 having cascade index 0, 1, 2 respectively.

5 Binary Enzymatic Networks are Explicitly Constructive

The material in this section requires familiarity with the formalism for species composition we develop in Gnacadja [5]. Included in that paper is the definition and a study of what it means for a reaction network to be explicitly constructive. Basically, it is a well-formedness condition that says that the network possesses an intrinsic notion of species composition which is comprehensive, explicit and minimal. We fix a binary enzymatic network $\mathscr{N} = (\mathscr{S}, \mathscr{C}, \mathscr{R})$ and use the notations of Section 2.

To show that the network is constructive, we need to present a core composition for it. We will define a map $\mathscr{E}$ and prove that it fulfills that role. We will use the $\mathbb{Z}$- and $\mathbb{R}$-linear spaces $\mathbb{Z}\big(\overline{\mathsf{Enz}_0} \sqcup \overline{\mathsf{Par}}\big)$ and $\mathbb{R}\big(\overline{\mathsf{Enz}_0} \sqcup \overline{\mathsf{Par}}\big)$. Consistently with notations discussed in Section 2.1 of Gnacadja [4] and used throughout this series of three papers, vectors in these spaces are regarded either as formal $\mathbb{Z}$- and $\mathbb{R}$-linear combinations of elements of $\overline{\mathsf{Enz}_0} \sqcup \overline{\mathsf{Par}}$, or as tuples indexed by $\overline{\mathsf{Enz}_0} \sqcup \overline{\mathsf{Par}}$ with entries in $\mathbb{Z}$ and $\mathbb{R}$.

We define the map $\mathscr{E} : \mathscr{S} \to \big(\mathbb{Z}_{\geqslant 0}\big(\overline{\mathsf{Enz}_0} \sqcup \overline{\mathsf{Par}}\big)\big) \setminus \{0\}$ as follows.

$$\text{For } X \in \mathsf{Enz} \cup \mathsf{Sub} \cup \mathsf{Pro} = \mathsf{Enz}_0 \sqcup \mathsf{Par},\ \mathscr{E}(X) := \mathsf{cl}(X)\,.$$

$$\text{For } Y \in \mathsf{Int},\ \mathscr{E}(Y) := \mathsf{cl}(E) + \mathsf{par}(E),\ \text{ where } E = \mathsf{enz}\big(\mathsf{cl}(Y)\big).$$

The map $\mathscr{E}$ is a composition map of the network $\mathscr{N}$ with composition tuples indexed by $\overline{\mathsf{Enz}_0} \sqcup \overline{\mathsf{Par}}$. The $\mathscr{E}$-elementary species are the enzymes, the substrates and the products, and the $\mathscr{E}$-composite species are the intermediates. Let $\tilde{\mathscr{E}}$ be the linear extension $\mathbb{R}\mathscr{S} \to \mathbb{R}\big(\overline{\mathsf{Enz}_0} \sqcup \overline{\mathsf{Par}}\big)$ of $\mathscr{E}$.

To perfectly match the wording of the definition of a composition map, we would have to number the elements of $\overline{\mathsf{Enz}_0} \sqcup \overline{\mathsf{Par}}$. But this is not necessary and there is no natural way to do so. We avoid doing it in the interest of not introducing new and arbitrary notation.

To put the definition of $\mathscr{E}$ in words, for a species X which is an enzyme or a partner, $\mathscr{E}(X)$ is simply the class $\mathsf{cl}(X)$ of X. And for a species Y which is an intermediate, finding $\mathscr{E}(Y)$ is just slightly more elaborate: first we find the class $\mathsf{cl}(Y)$ of the intermediate Y, then we find the enzyme for that class, $E = \mathsf{enz}\big(\mathsf{cl}(Y)\big)$, and then $\mathscr{E}(Y)$ is the sum of $\mathsf{cl}(E)$ and $\mathsf{par}(E)$, where $\mathsf{cl}(E)$ is the class of this enzyme E and $\mathsf{par}(E)$ is the class of the partners of this same enzyme E.

Here is the result of applying this to the network of Section 3.2.

$$\begin{aligned} \mathscr{E}(E_0) &= \{E_0\} \\ \mathscr{E}(S_0) = \mathscr{E}(E_1) &= \{S_0, E_1\} & \mathscr{E}(Y_0) &= \{E_0\} + \{S_0, E_1\} \\ \mathscr{E}(S_1) = \mathscr{E}(E_2) &= \{S_1, E_2\} & \mathscr{E}(Y_1) &= \{S_0, E_1\} + \{S_1, E_2\} \\ \mathscr{E}(S_2) = \mathscr{E}(E_3) &= \{S_2, E_3\} & \mathscr{E}(Y_2) &= \{S_1, E_2\} + \{S_2, E_3\} \end{aligned}$$

This does not plainly suggest the idea of a composition map. The way $\mathscr{E}$ is defined is more suited to mathematical deductions in a general context than to illustrating particular examples. We can have $\mathscr{E}$ expressed with tuples by numbering from one to four the classes of $\mathscr{E}$-elementary species, $\{E_0\}$, $\{S_0, E_1\}$, $\{S_1, E_2\}$, $\{S_2, E_3\}$, say in the order just listed. Here is what $\mathscr{E}$ then becomes.

$$\begin{aligned} \mathscr{E}(E_0) &= (1,0,0,0) \\ \mathscr{E}(S_0) = \mathscr{E}(E_1) &= (0,1,0,0) & \mathscr{E}(Y_0) &= (1,1,0,0) \\ \mathscr{E}(S_1) = \mathscr{E}(E_2) &= (0,0,1,0) & \mathscr{E}(Y_1) &= (0,1,1,0) \\ \mathscr{E}(S_2) = \mathscr{E}(E_3) &= (0,0,0,1) & \mathscr{E}(Y_2) &= (0,0,1,1) \end{aligned}$$

With the general definition and meaning of $\mathscr{E}$ now established, we can state the theorem that describes its purpose.

Theorem 5.1. *The composition map $\mathscr{E}$ is a core composition of the network $\mathscr{N}$. The elementary species are the enzymes, the substrates and the products, while the composite species are the intermediates. The network $\mathscr{N}$ is explicitly constructive.*

Proof. We show that the three conditions of Theorem 4.2 of Gnacadja [5] are realized.

We prove condition (1) of Gnacadja [5, Theorem 4.2], i.e. that $\mathscr{E}$ is a near-core composition of $\mathscr{N}$. We already noted that the $\mathscr{E}$-elementary species are the species $X \in \mathsf{Enz} \cup \mathsf{Sub} \cup \mathsf{Pro}$. Furthermore, every $\mathscr{E}$-elementary composition occurs as $\mathsf{cl}(X)$ for such a species X because the map cl is surjective from $\mathsf{Enz} \cup \mathsf{Sub} \cup \mathsf{Pro} = \mathsf{Enz}_0 \sqcup \mathsf{Par}$ onto $\overline{\mathsf{Enz}_0} \sqcup \overline{\mathsf{Par}}$. So it remains to show that all reactions are $\mathscr{E}$-conservative. Let $(E, A, B) \in \mathsf{Cat}$ and let $(Y_1, \ldots, Y_\ell) \in \mathsf{IntPath}(E, A, B)$. We have $A, B \in \mathsf{par}(E)$, which is equivalent to $\mathsf{cl}(A) = \mathsf{cl}(B) = \mathsf{par}(E)$, and we have $Y_1, \ldots, Y_\ell \in \mathsf{int}(E)$, which is equivalent to $E = \mathsf{enz}(\mathsf{cl}(Y_1)) = \cdots = \mathsf{enz}(\mathsf{cl}(Y_\ell))$. So by the definition of $\mathscr{E}$, we have $\mathscr{E}(Y_1) = \cdots = \mathscr{E}(Y_\ell) = \mathsf{cl}(E) + \mathsf{cl}(A) = \mathsf{cl}(E) + \mathsf{cl}(B)$. On another hand, we have $\tilde{\mathscr{E}}(E + X) = \mathscr{E}(E) + \mathscr{E}(X) = \mathsf{cl}(E) + \mathsf{cl}(X)$ for $X = A$ and

$X = B$. It results that all reactions in $\mathscr{R}_{(Y_1,\ldots,Y_\ell)}(E, A, B)$ are $\mathscr{E}$-conservative. Thus, all reactions are $\mathscr{E}$-conservative. This concludes the proof that $\mathscr{E}$ is a near-core composition of $\mathscr{N}$.

As a preparation for proving condition (2) of Gnacadja [5, Theorem 4.2], we note that if (A, B) is a substrate-product pair, then the species A and B are stoichiometrically isomeric. Indeed, let E be an enzyme such that (E, A, B) is a catalysis triple, and let $(Y_1, \ldots, Y_\ell) \in \mathsf{IntPath}(E, A, B)$. We have $B - A = (E + B - Y_\ell) + \sum_{j=2}^{\ell}(Y_j - Y_{j-1}) + (Y_1 - E - A)$, and therefore $B - A$ lies in the stoichiometric space.

We prove condition (2) of Gnacadja [5, Theorem 4.2]. From the definition of the composition map $\mathscr{E}$, we get that the $\mathscr{E}$-isomerism classes of $\mathscr{E}$-elementary species are the singletons of elements of Enz_0 and the elements of $\overline{\mathsf{Par}}$. So we need to show that if $A, B \in \mathsf{Par}$ and $\mathsf{cl}(A) = \mathsf{cl}(B)$, then A and B are stoichiometrically isomeric. Consider such A and B. Then there exists $C_0, \ldots, C_r \in \mathsf{Par}$ such that $C_0 = A$, $C_r = B$, and for each $j \in [1..r]$, (C_{j-1}, C_j) or (C_j, C_{j-1}) is a substrate-product pair. In either case, C_{j-1} and C_j are stoichiometrically isomeric. Consequently, A and B are stoichiometrically isomeric.

We prove condition (3) of Gnacadja [5, Theorem 4.2]. The $\mathscr{E}$-composite species are the intermediates. Let $Y \in \mathsf{Int}$. Because the species Y participates in at least one reaction, there exists a catalysis triple (E, A, B) and an intermediates path $(Y_1, \ldots, Y_\ell) \in \mathsf{IntPath}(E, A, B)$ such that Y is one of the intermediates $Y_1, \ldots, Y_\ell$. We have $\mathscr{E}(Y) = \mathscr{E}(E) + \mathscr{E}(A)$ and $Y - E - A$ is in the stoichiometric space. □

Remark 5.2. It is apparent from the definition of a binary enzymatic network that:

- Enzymes are both explicitly constructive and explicitly destructive;
- The substrates are the partners that are explicitly constructive;
- The products are the partners that are explicitly destructive; and
- Intermediates are both explicitly constructible and explicitly destructible.

Therefore, with Theorem 5.1, a binary enzymatic network is explicitly-reversibly constructive if and only if all substrates are also products and all products are also substrates. Hence, with Remark 4.4, futile binary enzymatic networks are explicitly-reversibly constructive. □

The two well-established notions of reversibility in Chemical Reaction Network Theory are reversibility and weak reversibility. In our observation however, biochemically valid networks that are weakly reversible are in fact reversible, and these are not the majority. Explicitly-reversibly constructive networks seem more suited to model large classes of biochemically valid reaction networks. Nevertheless, Remark 4.6 shows that weak reversibility has relevance in the biochemical context.

One benefit of an explicitly constructive network is that we have for the conservation space (the orthogonal of the stoichiometric space) a canonical basis consisting of vectors that are linear combinations of species with non-negative integer coefficients. In more direct terms, we have a set of vectors which express the conservativeness of the network in a comprehensive and minimal fashion. This set is commonly found in examples by visual inspection of the network. The fact that it is a basis is usually tacitly taken for granted. As preparation for presenting this basis, we introduce the following sets of species for non-partner enzymes $E \in \mathsf{Enz}_0$ and for isomerism classes of partners $\mathscr{X} \in \overline{\mathsf{Par}}$.

$$\begin{aligned}
\mathscr{S}(E) &:= \{E\} \sqcup \mathsf{int}(E)\,, \\
\mathscr{S}'(\mathscr{X}) &:= \big\{Y \in \mathsf{Int} : \mathsf{cl}\big(\mathsf{enz}\big(\mathsf{cl}(Y)\big)\big) = \mathscr{X}\big\}\,, \\
\mathscr{S}''(\mathscr{X}) &:= \big\{Y \in \mathsf{Int} : \mathsf{par}\big(\mathsf{enz}\big(\mathsf{cl}(Y)\big)\big) = \mathscr{X}\big\}\,, \\
\mathscr{S}(\mathscr{X}) &:= \mathscr{X} \sqcup \mathscr{S}'(\mathscr{X}) \sqcup \mathscr{S}''(\mathscr{X})\,.
\end{aligned}$$

A description of these sets follows.

- The set $\mathscr{S}(E)$ consists of the non-partner enzyme E and the intermediates that contain it.
- The set $\mathscr{S}'(\mathscr{X})$ consists of certain intermediates: an intermediate is in $\mathscr{S}'(\mathscr{X})$ if and only if it contains an element of $\mathscr{X}$ as an enzyme.
- The set $\mathscr{S}''(\mathscr{X})$ consists of certain intermediates: an intermediate is in $\mathscr{S}''(\mathscr{X})$ if and only if it contains an element of $\mathscr{X}$ as a partner.

Here is what this gives for the network of Section 3.2: $\mathscr{S}(E_0) = \{E_0, Y_0\}$,

$$\begin{aligned}
&\mathscr{S}'\big(\{S_0, E_1\}\big) = \{Y_1\}, \quad \mathscr{S}'\big(\{S_1, E_2\}\big) = \{Y_2\}, \quad \mathscr{S}'\big(\{S_2, E_3\}\big) = \varnothing, \\
&\mathscr{S}''\big(\{S_0, E_1\}\big) = \{Y_0\}, \quad \mathscr{S}''\big(\{S_1, E_2\}\big) = \{Y_1\}, \quad \mathscr{S}''\big(\{S_2, E_3\}\big) = \{Y_2\}.
\end{aligned}$$

The disjoint union in the definition of $\mathscr{S}(\mathscr{X})$ is justified because Condition (Enz3) implies that $\mathscr{S}'(\mathscr{X}) \cap \mathscr{S}''(\mathscr{X}) = \varnothing$. With sum denoting the function that sums the elements of a finite subset when such operation makes sense, we set

$$\begin{aligned} T_E &:= \operatorname{sum}\big(\mathscr{S}(E)\big) = E + \operatorname{sum}\big(\mathsf{int}(E)\big)\,, \\ T_{\mathscr{X}} &:= \operatorname{sum}\big(\mathscr{S}(\mathscr{X})\big) = \operatorname{sum}(\mathscr{X}) + \operatorname{sum}\big(\mathscr{S}'(\mathscr{X})\big) + \operatorname{sum}\big(\mathscr{S}''(\mathscr{X})\big)\,. \end{aligned}$$

With Theorem 3.6 of Gnacadja [5], we get:

Theorem 5.3. *The vectors T_E for $E \in \mathsf{Enz}_0$ and $T_{\mathscr{X}}$ for $\mathscr{X} \in \overline{\mathsf{Par}}$ form a basis of the conservation space (the orthogonal of the stoichiometric space).* □

For the network of Section 3.2, the basis of Theorem 5.3 consists of the following four vectors.

$$\begin{aligned} T_{E_0} &= E_0 + Y_0 \\ T_{\{S_0,E_1\}} &= S_0 + E_1 + Y_0 + Y_1 \\ T_{\{S_1,E_2\}} &= S_1 + E_2 + Y_1 + Y_2 \\ T_{\{S_2,E_3\}} &= S_2 + E_3 + Y_2 \end{aligned}$$

These are "the conservation laws" of the network. Again, a visual inspection of the network could yield these vectors, and even an intuition, but not a proof, that they form a basis of the conservation space.

6 Persistence

The main result in this section is Theorem 6.7 which states that a binary enzymatic network that is futile and cascaded is vacuously persistent. We study vacuous persistence for reaction networks in general in Gnacadja [4] and for constructive networks in particular in Gnacadja [5]. Vacuous persistence is the property that no species will tend to extinction if all species are implicitly present at initial time. To obtain Theorem 6.7, we collect a number of interesting results and eventually apply the following one.

Theorem 6.1 (Gnacadja [4, Theorem 5.5]). *Consider a mass-action reaction network for which that all trajectories are bounded. Then the following are equivalent:*

- *The reaction network is vacuously persistent.*

- *Among the subsets of the set of all species, only the full set is both reach-closed and stoichiometrically admissible.* □

The paper cited contains the necessary explanations, including discussions on stoichiometric admissibility and reachability. Let $\mathscr{N} = (\mathscr{S}, \mathscr{C}, \mathscr{R})$ be a binary enzymatic network and let $\mathscr{Z} \subseteq \mathscr{S}$ be a subset of species.

The following result is an immediate application of Proposition 6.2 of Gnacadja [5].

Lemma 6.2. *Suppose that $\mathscr{Z}$ is stoichiometrically admissible. Then $\mathscr{Z} \cap \mathscr{S}(E) \neq \varnothing$ for all $E \in \mathsf{Enz}_0$ and $\mathscr{Z} \cap \mathscr{S}(\mathscr{X}) \neq \varnothing$ for all $\mathscr{X} \in \overline{\mathsf{Par}}$.* □

The following result is an immediate application of Lemma 6.3 of Gnacadja [5].

Lemma 6.3. *Suppose that $\mathscr{Z}$ is reach-closed. Let $\mathscr{X} \in \overline{\mathsf{Par}}$.*
If $\mathscr{Z} \cap \mathscr{S}(\mathscr{X}) \neq \varnothing$, then $\mathscr{Z} \cap \mathscr{X} \neq \varnothing$. □

Following is a trivial but instrumental observation.

Remark 6.4. Let $E \in \mathsf{Enz} \cap \mathscr{Z}$. Suppose that $\mathscr{Z}$ is reach-closed.

- If $\mathscr{Z} \cap \mathsf{isub}(E) \neq \varnothing$, then $\mathsf{par}(E) \subseteq \mathscr{Z}$ and $\mathsf{int}(E) \subseteq \mathscr{Z}$.
- If $\mathscr{Z} \cap \mathsf{par}(E) \neq \varnothing$ or if $\mathscr{Z} \cap \mathsf{int}(E) \neq \varnothing$, then $\mathsf{tpro}(E) \subseteq \mathscr{Z}$. □

The preceeding two lemmas and remark combine nicely to yield the following result.

Theorem 6.5. *Suppose that the network $\mathscr{N}$ is futile. If $\mathscr{Z}$ is stoichiometrically admissible and reach-closed, and if $\mathsf{Enz} \subseteq \mathscr{Z}$, then $\mathscr{Z} = \mathscr{S}$.*

Proof. Let $F \in \mathsf{Enz}$. There exists $E \in \mathsf{Enz}$ such that F is a reversing enzyme for E. By Lemma 6.2, we have $\mathscr{Z} \cap \mathscr{S}\big(\mathsf{par}(E)\big) \neq \varnothing$. Then by Lemma 6.3, we have $\mathscr{Z} \cap \mathsf{par}(E) \neq \varnothing$. Next, the second assertion of Remark 6.4 implies that $\mathsf{tpro}(E) \subseteq \mathscr{Z}$. Therefore, by Definition 4.3, $\varnothing \neq \mathsf{isub}(F) \subseteq \mathscr{Z}$. Then, with the first assertion of Remark 6.4, we have $\mathsf{par}(F) \subseteq \mathscr{Z}$ and $\mathsf{int}(F) \subseteq \mathscr{Z}$. This holds for all $F \in \mathsf{Enz}$, so $\mathscr{Z} = \mathscr{S}$. □

We see next a way to satisfy the condition $\mathsf{Enz} \subseteq \mathscr{Z}$ which is required in Theorem 6.5.

Theorem 6.6. *Suppose that the network $\mathscr{N}$ is cascaded. If $\mathscr{Z}$ is stoichiometrically admissible and reach-closed, then* $\mathsf{Enz} \subseteq \mathscr{Z}$.

Proof. Let $E \in \mathsf{Enz}_0$. By Lemma 6.2, we have $\mathscr{Z} \cap \mathscr{S}(E) \neq \varnothing$, i.e. $E \in \mathscr{Z}$ or $\mathscr{Z} \cap \mathsf{int}(E) \neq \varnothing$. But we have $\mathscr{Z} \cap \mathsf{int}(E) \neq \varnothing \Rightarrow E \in \mathscr{Z}$ because $\mathscr{Z}$ is reach-closed. So $E \in \mathscr{Z}$. Hence, $\mathsf{Enz}_0 \subseteq \mathscr{Z}$. Let $m \in \mathbb{Z}_{\geqslant 1}$ and assume for induction that $\mathsf{Enz}_{m-1} \subseteq \mathscr{Z}$. Then for every $E \in \mathsf{Enz}_{m-1}$, we successively have: $\mathscr{Z} \cap \mathscr{S}\big(\mathsf{par}(E)\big) \neq \varnothing$ by Lemma 6.2; $\mathscr{Z} \cap \mathsf{par}(E) \neq \varnothing$ by Lemma 6.3; and $\mathsf{tpro}(E) \subseteq \mathscr{Z}$ by Remark 6.4. So $\bigcup_{E \in \mathsf{Enz}_{m-1}} \mathsf{tpro}(E) \subseteq \mathscr{Z}$, whence in particular, $\mathsf{Enz}_m \subseteq \mathscr{Z}$. □

By combining Theorems 6.5 and 6.6, and then using Theorem 6.1, we get:

Theorem 6.7. *If the binary enzymatic network $\mathscr{N}$ is futile and cascaded, then it is vacuously persistent.* □

$$\begin{array}{ccccccccc} E+S_0 & \rightleftharpoons & ES_0 & \to & E+S_1 & \rightleftharpoons & ES_1 & \to & E+S_2 \\ F+S_2 & \rightleftharpoons & FS_2 & \to & F+S_1 & \rightleftharpoons & FS_1 & \to & F+S_0 \end{array}$$

(a) A futile cycle of two two-step pathways.

$$\begin{array}{ccccccccc} E+S_0 & \rightleftharpoons & ES_0 & \to & E+S_1 & \rightleftharpoons & ES_1 & \to & E+E^\star \\ F+E^\star & \rightleftharpoons & FS_2 & \to & F+S_1 & \rightleftharpoons & FS_1 & \to & F+S_0 \\ E^\star+S_0^\star & \rightleftharpoons & ES_0^\star & \to & E^\star+S_1^\star & \rightleftharpoons & ES_1^\star & \to & E^\star+S_2^\star \\ F^\star+S_2^\star & \rightleftharpoons & FS_2^\star & \to & F^\star+S_1^\star & \rightleftharpoons & FS_1^\star & \to & F^\star+S_0^\star \end{array}$$

(b) Two futile cycles in a cascade.

$$\begin{array}{ccccccccc} M+E & \rightleftharpoons & ME & \to & M_y+E & \rightleftharpoons & M_yE & \to & M_2+E \\ M+E & \rightleftharpoons & ME^\star & \to & M_t+E & \rightleftharpoons & M_tE & \to & M_2+E \\ M_2+F & \rightleftharpoons & M_2F & \to & M_y+F & \rightleftharpoons & M_yF & \to & M+F \\ M_2+F & \rightleftharpoons & M_2F^\star & \to & M_t+F & \rightleftharpoons & M_tF & \to & M+F \end{array}$$

(c) One futile cycle with two alternate pathways in each direction.

Figure 6.1: Binary enzymatic networks from Angeli, De Leenheer and Sontag [1].

In Angeli, De Leenheer and Sontag [1], a Petri net approach is used to study the persistence of reaction networks and it is shown that the three networks of Figure 6.1 are persistent. One can observe that these networks are binary enzymatic networks as defined here. Furthermore, they are futile and cascaded – the networks of Figures 6.1(a) and 6.1(c) are trivially cascaded, while in the network of Figure 6.1(b), $E^\star$ is an enzyme of cascade index 1. So these three networks are vacuously persistent. Tables 6.1 and 6.2

illustrates for two of these networks some of the concepts we introduced for binary enzymatic networks. We find four other examples of enzymatic mechanisms in Siegel and MacLean [9, Section 4]. The mechanism with no inhibitor and the one with a competitive inhibitor (respectively in Sections 4.1 and 4.3 in the reference) are futile, trivially-cascaded binary enzymatic networks in our terminology. (The competitive inhibitor is simultaneously a substrate and a product.) Hence, consistently with results in the reference (Theorems 4.2 and 4.3 in the no-inhibitor case and Theorems 4.6 and 4.7 in the competitive inhibitor case), these networks are vacuously persistent. The mechanism with a noncompetitive inhibitor and the one with an uncompetitive inhibitor (respectively in Sections 4.2 and 4.4 in the reference) are not binary enzymatic networks because there are ternary species.

$X \in \mathsf{Enz}$	$\mathsf{par}(X)$	$\mathsf{int}(X)$	$\mathsf{isub}(X)$	$\mathsf{tpro}(X)$	$\varphi(X)$	$\gamma(X)$
E	$\{S_0, S_1, E^\star\}$	$\{ES_0, ES_1\}$	$\{S_0\}$	$\{E^\star\}$	F	0
F	$\{S_0, S_1, E^\star\}$	$\{FS_1, FS_2\}$	$\{E^\star\}$	$\{S_0\}$	E	0
$E^\star$	$\{S_0^\star, S_1^\star, S_2^\star\}$	$\{ES_0^\star, ES_1^\star\}$	$\{S_0^\star\}$	$\{S_2^\star\}$	$F^\star$	1
$F^\star$	$\{S_0^\star, S_1^\star, S_2^\star\}$	$\{FS_1^\star, FS_2^\star\}$	$\{S_2^\star\}$	$\{S_0^\star\}$	$E^\star$	0

Table 6.1: Selected concepts illustrated for the network of Figure 6.1(b).

$X \in \mathsf{Enz}$	$\mathsf{par}(X)$	$\mathsf{int}(X)$	$\mathsf{isub}(X)$	$\mathsf{tpro}(X)$	$\varphi(X)$	$\gamma(X)$
E	$\{M, M_y, M_t, M_2\}$	$\{ME, M_yE, M_tE\}$	$\{M\}$	$\{M_2\}$	F	0
F	$\{M, M_y, M_t, M_2\}$	$\{MF, M_yF, M_tF\}$	$\{M_2\}$	$\{M\}$	E	0

Table 6.2: Selected concepts illustrated for the network of Figure 6.1(c).

7 Conclusion

We have come to the end of this series of three papers investigating persistence in reaction networks. We had several motivations for this effort. One was that biochemically relevant persistence should account for trajectories originating at states where all species are present implicitly, but not necessarily explicitly; whence the notion of vacuous persistence in the first paper. Another motivation was that persistence should be in effect when species

are made of building blocks that are conserved and processes are fundamentally reversible; whence the theory of species composition and constructive networks in the second paper. Yet another motivation was that there should be theorems that expressly affirm mathematical properties that interested bioscientists would deem obvious. Indeed, if a biochemist were to look at the futile cascaded binary enzymatic networks in this paper, they would readily conclude that the conservation and self-compensating characteristics of the networks could not allow the depletion of any species. Effectively, they would conclude that the networks are vacuously persistent by conducting (in ways that mathematicians would consider handwavy) the relevant reachability analysis. This is what we do with mathematical generality and rigor in this series of papers, first for reaction networks in general, then for constructive networks, and finally here for binary enzymatic networks. The formalism of binary enzymatic networks could serve in further research. It could also be extended. One useful extention would be to remove the restriction to binary intermediates. Another one could account for enzymatic networks that are genuinely binary but are not accounted for in our formalism because there would be an enzyme that is not selective at catalyzing conversions in only one isomerism class of substrates and products.

Acknowledgments

This paper benefited from the comments the author received from Eduardo Sontag, Martin Feinberg, Anne Shiu and the anonymous reviewers.

References

[1] D. Angeli, P. De Leenheer, and E. D. Sontag, *A Petri Net Approach to the Study of Persistence in Chemical Reaction Networks*, Mathematical Biosciences **210** (2007), no. 2, 598–618, DOI 10.1016/j.mbs.2007.07.003.

[2] A. Cornish-Bowden, *Fundamentals of Enzyme Kinetics*, 3rd ed., Portland Press, 2004, ISBN 1855781581.

[3] G. Craciun, Y. Tang, and M. Feinberg, *Understanding Bistability in Complex Enzyme-Driven Reaction Networks*, Proceedings of the National Academy of Sciences USA **103** (2006), no. 23, 8697–8702, DOI 10.1073/pnas.0602767103.

[4] G. Gnacadja, *Reachability, Persistence, and Constructive Chemical Reaction Networks (Part I): Reachability Approach to the Persistence of Chemical Reaction Networks*, Journal of Mathematical Chemistry, DOI 10.1007/s10910-011-9894-4.

[5] ———, *Reachability, Persistence, and Constructive Chemical Reaction Networks (Part II): A Formalism for Species Composition in Chemical Reaction Network Theory and Application to Persistence*, Journal of Mathematical Chemistry, DOI 10.1007/s10910-011-9896-2.

[6] J. Gunawardena, *Distributivity and Processivity in Multisite Phosphorylation can be Distinguished through Steady-State Invariants*, Biophysical Journal **93** (2007), no. 11, 3828–3834, DOI 10.1529/biophysj.107.110866.

[7] M. Pérez Millán, A. Dickenstein, A. Shiu, and C. Conradi, *Chemical Reaction Systems with Toric Steady States*, http://arxiv.org/abs/1102.1590.

[8] C. Salazar and T. Höfer, *Multisite Protein Phosphorylation – From Molecular Mechanisms to Kinetic Models*, FEBS Journal **276** (2009), no. 12, 3177–3198, DOI 10.1111/j.1742-4658.2009.07027.x.

[9] D. Siegel and D. MacLean, *Global Stability of Complex Balanced Mechanisms*, Journal of Mathematical Chemistry **27** (2000), 89–110, DOI 10.1023/A:1019183206064.

[10] M. Thomson and J. Gunawardena, *The Rational Parameterisation Theorem for Multisite Post-Translational Modification Systems*, Journal of Theoretical Biology **261** (2009), no. 4, 626–636, DOI 10.1016/j.jtbi.2009.09.003.

[11] L. Wang and E. D. Sontag, *On the Number of Steady States in a Multiple Futile Cycle*, Journal of Mathematical Biology **57** (2008), no. 1, 29–52, DOI 10.1007/s00285-007-0145-z.