Journal of Theoretical Biology 409 (2016) 1–10

Contents lists available at ScienceDirect

Journal of Theoretical Biology

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

Conditions for duality between fluxes and concentrations in biochemical networks

Ronan M.T. Fleming [a,*], Nikos Vlassis [b], Ines Thiele [a], Michael A. Saunders [c]

[a] Luxembourg Centre for Systems Biomedicine, University of Luxembourg, 7 avenue des Hauts-Fourneaux, Esch-sur-Alzette, Luxembourg
[b] Adobe Research, 345 Park Ave, San Jose, CA, USA
[c] Dept of Management Science and Engineering, Stanford University, Stanford, CA, USA

HIGHLIGHTS

- Flux-concentration duality implies an equivalence between descriptions in terms of concentrations or unidirectional fluxes.
- A novel stoichiometric condition for duality between unidirectional fluxes and concentrations is proposed.
- Flux-concentration duality is a pervasive property of biochemical networks.

ARTICLE INFO

Article history:
Received 8 December 2015
Received in revised form
3 June 2016
Accepted 23 June 2016
Available online 23 June 2016

Keywords:
Biochemical network
Flux
Concentration
Duality
Kinetics

ABSTRACT

Mathematical and computational modelling of biochemical networks is often done in terms of either the concentrations of molecular species or the fluxes of biochemical reactions. When is mathematical modelling from either perspective equivalent to the other? Mathematical duality translates concepts, theorems or mathematical structures into other concepts, theorems or structures, in a one-to-one manner. We present a novel stoichiometric condition that is necessary and sufficient for duality between unidirectional fluxes and concentrations. Our numerical experiments, with computational models derived from a range of genome-scale biochemical networks, suggest that this flux-concentration duality is a pervasive property of biochemical networks. We also provide a combinatorial characterisation that is sufficient to ensure flux-concentration duality.The condition prescribes that, for every two disjoint sets of molecular species, there is at least one reaction complex that involves species from only one of the two sets. When unidirectional fluxes and molecular species concentrations are dual vectors, this implies that the behaviour of the corresponding biochemical network can be described entirely in terms of either concentrations or unidirectional fluxes.

1. Introduction

Systems biochemistry seeks to understand biological function in terms of a network of chemical reactions. Systems biology is a broader field, encompassing systems biochemistry, where understanding is in terms of a network of interactions, some of which may not be immediately identifiable with a particular chemical or biochemical reaction. Mathematical and computational modelling of biochemical reaction network dynamics is a fundamental component of systems biochemistry. Any genome-scale model of a biochemical reaction network will give rise to a system of equations with a high-dimensional state variable, e.g., there are at least 1000 genes in *Pelagibacter ubique* (Giovannoni et al., 2005), the smallest free-living microorganism currently known. In order to ensure that mathematical and computational modelling remains tractable at genome-scale, it is important to focus research effort on the development of robust algorithms with time complexity that scales well with the dimension of the state variable.

Given some assumptions as to the dynamics of a biochemical network, a mathematical model is defined in terms of a system of equations. Characterising the mathematical properties of such a system of equations can lead directly or indirectly to insightful biochemical conclusions. Directly, in the sense that the recognition of the mathematical property has direct biochemical implications, e.g., the correspondence between an extreme ray of the steady state (irreversible) flux cone and the minimal set of reactions that could operate at steady state (Schuster et al., 2000). Or indirectly, in the sense of an algorithm tailored to exploit a recognised property, which is subsequently implemented to derive

* Corresponding author.
E-mail addresses: ronan.m.t.fleming@gmail.com (R.M.T. Fleming), nikos.vlassis@gmail.com (N. Vlassis), ines.thiele@gmail.com (I. Thiele), saunders@stanford.edu (M.A. Saunders).

http://dx.doi.org/10.1016/j.jtbi.2016.06.033

biochemical conclusions from a computational model, e.g., robust flux balance analysis algorithms (Sun et al., 2013) applied to investigate codon usage in an integrated model of metabolism and macromolecular synthesis in *Escherichia coli* (Thiele et al., 2012).

Mathematical duality translates concepts, theorems or mathematical structures into other concepts, theorems or structures in a one-to-one manner. Sometimes, recognition of mathematical duality underlying a biochemical network modelling problem enables the dual problem to be more efficiently solved. An example of this is the problem of computing *minimal cut sets*, i.e., minimal sets of reactions whose deletion will block the operation of a specified objective in a steady state model of a biochemical network (Klamt and Gilles, 2004). Previously, computation of minimal cut sets required enumeration of the extreme rays of part of the steady state (irreversible) flux cone, which is computationally complex in memory and processing time (Haus et al., 2008). By recognising that minimal cut sets in a primal network are dual to extreme rays in a dual network (Ballerstein et al., 2012), one can compute select subsets of extreme rays for the dual network that correspond to minimal cut sets with the certain desired properties in the primal (i.e., original) biochemical network in question (von Kamp and Klamt, 2014). This fundamental work has many experimental biological applications, including metabolic engineering (Mahadevan et al., 2015).

Recognition of mathematical duality in a biochemical network modelling problem can have many theoretical biological applications, in advance of experimental biological applications. For example, in mathematical modelling of biochemical reaction networks, there has long been an interest in the relationship between models expressed in terms of molecular species concentrations and models expressed in terms of reaction fluxes. When concentrations or *net* fluxes are considered as independent variables, a duality between the corresponding Jacobian matrices has been demonstrated (Jamshidi and Palsson, 2009). In this case, the concentration and net flux Jacobian matrices can be used to estimate the dynamics of the same system, with respect to perturbations to concentrations or net fluxes about a given steady state. The primal (concentration) Jacobian and dual (net flux) Jacobian matrices are identical, except that one is the transpose of the other. Matrix transposition is a one-to-one mapping and the aforementioned duality is between the pair of Jacobians. This does not mean that the net flux and concentration vectors are dual variables in the same mathematical sense, and neither are the perturbations to concentrations or net fluxes. This is because the Jacobian duality (Jamshidi and Palsson, 2009), which exists for any stoichiometric matrix, does not enforce a one-to-one mapping between concentrations and net fluxes unless the stoichiometric matrix is invertible, which is never the case for a biochemical network (Heinrich et al., 1978).

Herein we ask and answer the question: what conditions are necessary and sufficient for duality between unidirectional fluxes and molecular species concentrations? We establish a necessary linear algebraic condition on reaction stoichiometry in order for duality to hold. We also combinatorially characterise this stoichiometric condition in a manner amenable to interpretation for biochemical networks in general. In manually curated metabolic network reconstructions, across a wide range of species and biological processes, we confirm satisfaction of this stoichiometric condition for the major subset of molecular species within each reconstruction of a biochemical network. Furthermore, we demonstrate how linear algebra can be applied to test for satisfaction of this stoichiometric condition or to identify the molecular species involved in violation of this condition. We also demonstrate that violation of flux-concentration duality points to discrepancies between a reconstruction and the underlying biochemistry, thereby establishing a new stoichiometric quality control procedure to select a subset of a biochemical network reconstruction for use in computational modelling of steady states.

First, we establish a linear algebraic condition and a combinatorial condition for duality between unidirectional fluxes and concentrations. Subsequently, we introduce a procedure to convert a reconstruction into a computational model in a quality-controlled manner. We then apply this procedure to a range of genome-scale metabolic network reconstructions and test for the linear algebraic condition for flux-concentration duality before and after conversion into a model. We conclude with a broad discussion, with examples illustrating how a recognition of flux-concentration duality could help address questions of biological relevance and improve our understanding of biological phenomena.

2. Theoretical results

2.1. Stoichiometry and reaction kinetics

We consider a biochemical network with m molecular species and n (net) reactions. Without loss of generality with respect to genome-scale biochemical networks, we assume $m \leq n$. We assume that each reaction is *reversible* (Lewis, 1925) and can be represented by a *unidirectional reaction* pair. With respect to the forward direction, in a *forward stoichiometric matrix* $F \in \mathbb{R}^{m \times n}$, let F_{ij} be the *stoichiometry* of molecule i participating as a substrate or catalyst in *forward unidirectional reaction* j. Likewise, with respect to the reverse direction, in a *reverse stoichiometric matrix* $R \in \mathbb{R}^{m \times n}$, let R_{ij} be the *stoichiometry*of molecule i participating as a substrate or catalyst in *reverse unidirectional reaction* j. The set of molecular species that jointly participate as either substrates or products in a single unidirectional reaction is referred to as a *reaction complex*.

One may define the topology of a *hypergraph* of reactions with a *net stoichiometric matrix* $S := R - F$. However, a catalyst, by definition, participates in a reaction with the same stoichiometry as a substrate or product ($F_{ij} = R_{ij}$), so the corresponding row of S is all zeros unless that catalyst is synthesised or consumed elsewhere in the same biochemical network, as is the case for many biochemical catalysts (Thiele et al., 2009). For example, consider the ith molecular species acting as a catalyst in some reactions. If it is synthesised in the jth reaction of a biochemical network, the stoichiometric coefficient in the forward stoichiometric matrix will be less than that of the reverse stoichiometric matrix ($F_{ij} < R_{ij}$), so $S_{ij} := R_{ij} - F_{ij} > 0$. This also encompasses the case of an auto-catalytic reaction.

Before proceeding, some comments on our assumptions are in order. One may derive S from F and R, but the latter pair of matrices cannot, in general, be derived from S because S omits the stoichiometry of catalysis. The orientation of the hypergraph, i.e., the assignment of one direction to be forward (substrates $\rightharpoonup$ products), with the other reverse, is typically made so that net flux is forward (with positive sign) when a reaction is active in its biologically typical direction in a biochemical network. This is an arbitrary convention rather than a constraint, and reversing the orientation of one reaction only exchanges one column of F for the corresponding one in R. Although every chemical reaction is in principle reversible, in a biochemical setting, due to physiological limits on the relative concentrations of reactants and substrates, some reactions are practically irreversible (Noor et al., 2013). Our conclusions also extend to systems of irreversible reactions because the reaction complexes for an irreversible reaction are the same as those for a reversible reaction.

In the following, the exponential or natural logarithm of a vector is meant component-wise, with $\exp(\log(0)) := 0$. Let $v_f \in \mathbb{R}^n_{\geq 0}$ and $v_r \in \mathbb{R}^n_{\geq 0}$ denote forward and reverse unidirectional reaction