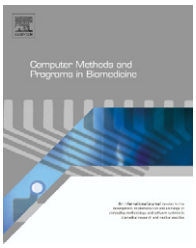
COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE 104 (2011) 70–80

journal homepage: www.intl.elsevierhealth.com/journals/cmpb



Indistinguishability and identifiability of kinetic models for the MurC reaction in peptidoglycan biosynthesis

J.G. Hattersley[a], *J. Pérez-Velázquez*[a], *M.J. Chappell*[a], *D. Bearup*[b], *D. Roper*[b], *C. Dowson*[b], *T. Bugg*[c], *N.D. Evans*[a,*]

[a] *School of Engineering, University of Warwick, Coventry, CV4 7AL, UK*
[b] *School of Biological Sciences, University of Warwick, Coventry, UK*
[c] *School of Chemistry, University of Warwick, Coventry, UK*

ARTICLE INFO

Article history:
Received 14 December 2009
Received in revised form 9 June 2010
Accepted 22 July 2010

Keywords:
Indistinguishability
Identifiability
Experiment design
MurC
Biomedical systems
Parameter identification

ABSTRACT

An important question in Systems Biology is the design of experiments that enable discrimination between two (or more) competing chemical pathway models or biological mechanisms. In this paper analysis is performed between two different models describing the kinetic mechanism of a three-substrate three-product reaction, namely the MurC reaction in the cytoplasmic phase of peptidoglycan biosynthesis. One model involves ordered substrate binding and ordered release of the three products; the competing model also assumes ordered substrate binding, but with fast release of the three products. The two versions are shown to be distinguishable; however, if standard quasi-steady-state assumptions are made distinguishability cannot be determined. Once model structure uniqueness is ensured the experimenter must determine if it is possible to successfully recover rate constant values given the experiment observations, a process known as structural identifiability. Structural identifiability analysis is carried out for both models to determine which of the unknown reaction parameters can be determined uniquely, or otherwise, from the ideal system outputs. This structural analysis forms an integrated step towards the modelling of the full pathway of the cytoplasmic phase of peptidoglycan biosynthesis.

1. Introduction

In the biological sciences it is becoming increasingly common to collect data in high-throughput experiments on several scales: genomic, proteomic, or metabolic [30,27]. These data hold the promise of identifying the mechanisms of interactions that comprise large-scale regulatory biochemical networks. An important step in this work is the design of experiments to allow discrimination between two (or more) competing pathway models or biological mechanisms. Structural indistinguishability provides a formal approach to distinguish between competing model mechanisms [19].

In Systems Biology mathematical models that are generated invariably include large numbers of state variables with numerous model parameters, many of which are unknown, or cannot be directly measured. With such highly complex systems there are often few direct measurements that can be made and limited access for input perturbation to elucidate system dynamics. These limitations cause problems when investigating the existence of hidden pathways or attempting to estimate unknown parameters. Identifiability analysis provides a formal approach to determine what additional inputs and/or measurements are necessary in order to reduce, or remove, these limitations and permit the derivation of models that can be used for practical purposes with greater confidence [30,27,19].

* *Corresponding author*. Tel.: +44 24 765 22062.
E-mail addresses: john@jhattersley.com (J.G. Hattersley), Neil.Evans@warwick.ac.uk (N.D. Evans).
0169-2607/$ – see front matter © 2010 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.cmpb.2010.07.009

Structural indistinguishability for system models is concerned with determining the uniqueness between possible candidates for the model (or mechanism) structure [13]. The analysis is concerned with whether the underlying possibilities for the parameterised mathematical model can be distinguished using the inputs (perturbations or interventions) and observations (or measurements) available for the biological system under investigation.

In chemical kinetics it is key to characterise reaction mechanisms, however there are often several different process models that are consistent with the available data. These mechanisms may be described by the same mathematical representation [9,10] but without formal analysis of the mathematical model a reaction mechanism's validity is only disproved by showing inconsistency with available data. Whilst this problem has been recognised [5,6], structural indistinguishability is not routinely applied to chemical kinetics experiments and model development. For linear systems structural indistinguishability analysis is generally exhaustive with all competing mechanisms generated from a given one (see [16]). For nonlinear systems, approaches are generally only for pairs of candidate models, though in some cases a parameterised family of such candidates can be generated. There has been limited application of structural indistinguishability analysis to the chemical kinetic models. In Schnell et al. [28] the issues of distinguishability with respect to biochemical kinetics was considered via application of structural indistinguishability to classical models for a single-enzyme, single-substrate reaction. In addition, simple kinetic models are incorporated into studies of structural analysis methods, often with Michaelis–Menten type reparameterisation (see [25,3]); however, to the authors' knowledge this is the first time indistinguishability and identifiability have been applied to a full (three-substrate/three-product) enzyme kinetic model.

Structural identifiability arises in the inverse problem of inferring from the known, or assumed, properties of a biological system, estimates for the corresponding rate constants and other parameters; as such it can be considered as a special case of the structural indistinguishability problem. Structural identifiability analysis considers the uniqueness of the unknown model parameters from the input–output structure corresponding to proposed experiments to collect data for parameter estimation. This is an important, but often overlooked, prerequisite to experiment design, system identification and parameter estimation, since estimates for structurally unidentifiable parameters are effectively meaningless. If parameter estimates are to be used to inform intervention or inhibition strategies, or other critical decisions, then it is essential that the parameters be uniquely identifiable. Numerous techniques for performing a structural identifiability analysis on linear parametric models exist (see [17,34]). In comparison, there are relatively few techniques available for nonlinear systems such as the Taylor series approach [23], similarity transformation based approaches [31,32,12] and differential algebra techniques [20,26]. Unfortunately for systems with a complex structure significant computational problems can arise even for relatively low dimensional models. At present there has been relatively little work on techniques for large-scale, highly complex systems, which are typical in Systems Biology. As shall be shown, the analytic identifiability approaches can generate computationally intractable solutions; therefore, an alternative numerical approach that uses the sensitivity of the observation functions to changes in the parameters is implemented to suggest the local identifiability of the parameters and associated model [15,18].

The purpose of this paper is to explore the possible effectiveness of using structural indistinguishability and identifiability techniques in model discrimination within Systems Biology networks, using MurC as a case study; to this end, a structural indistinguishability analysis is performed between two different models describing the kinetic mechanism of the MurC (Ter-ter) reaction in the cytoplasmic phase of bacterial peptidoglycan biosynthesis (BPB). One model assumes the kinetic mechanism proposed by Emanuele et al. [8], which involves step-by-step release of the three products; the competing model assumes that the release of the products is simultaneous. A structural identifiability analysis is also carried out for both models to ensure that the model output uniquely determines the unknown parameters.

2. The models

Prior to describing the mathematical models of the Ter-ter enzyme reaction, it is necessary to review the basic biology and context of the MurC reaction. The cell wall of many bacteria is composed of peptidoglycan, which is made up of a combination of peptide bonds and carbohydrates. Peptidoglycan serves a structural role in the bacterial cell wall, giving rigidity, as well as counteracting the osmotic pressure of the cytoplasm [7]. The BPB pathway (shown in Fig. 1) is a significant target in the development of antibacterial agents [33]. A detailed understanding of the biosynthesis pathway is essential for the development of new strategies for antibacterial action to compensate for the emergence of clinical resistance to penicillin antibiotics in *Staphylococcus aureus*, *Streptococcus pneumoniae* and Gram-negative pathogens such as *Pseudomonas aeruginosa* [1,21], and the emergence of vancomycin resistance in Enterococci, together with the lack of new classes of antibacterial agent. Although this pathway is quite well known, characterisation, especially of the later lipid linked steps, has been hampered by difficulties in making the natural substrates and there remain some uncertainties within the wider reaction network. As such this pathway is ideal for a feasibility study of the effectiveness of structural indistinguishability and identifiability in mechanism discrimination, model formulation and experiment design [30,27].

Within the biosynthetic pathway for bacterial peptidoglycan there is already a reasonable (but not complete) understanding of the cytoplasmic pathways. Whilst a basic pathway scheme is recognised for this region, there are still a number of steps within the pathway where competing reaction schemes may exist (e.g. multiple isoforms of an enzyme within the same cell) and also where feedback inhibition may play an important (but not yet fully understood) role. The study of the cytoplasmic phase of the biosynthetic process comprises a comprehensive understanding of the MurA to MurF reactions of the full pathway (see Fig. 1). In this paper the focus is on MurC.