Kinetic modelling of metabolic pathways: Application to serine biosynthesis

Smallbone, Kieran and Stanford, Natalie J.

2013

MIMS EPrint: **2012.57**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from: http://eprints.maths.manchester.ac.uk/
And by contacting: The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097

Kinetic modelling of metabolic pathways: Application to serine biosynthesis

Kieran Smallbone and Natalie J. Stanford

May 2012

MIMS EPrint: **2012.57**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from: http://www.manchester.ac.uk/mims/eprints
And by contacting: The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097

Kinetic modelling of metabolic pathways: Application to serine biosynthesis

Kieran Smallbone and Natalie J. Stanford

Summary

In this chapter, we describe the steps needed to create a kinetic model of a metabolic pathway using kinetic data from both experimental measurements and literature review. Our methodology is presented by using the example of serine biosynthesis in *E. coli*.

Key words: Mathematical modelling, Metabolism

1. Introduction

Metabolic pathways are collections of enzyme-mediated reactions. Their role in living organisms is to transform cellular inputs, such as glucose, into biomass and energy, so they can grow and function. Nonlinear processes dominate the interactions between enzymes and metabolites, and hence intuitive verbal reasoning approaches are insufficient to describe the complex dynamic behaviour of such biological networks (1-3). Nor can such approaches keep pace with the large increases in 'omics data (such as metabolomics and proteomics) and the accompanying advances in high-throughput experiments and bioinformatics. Rather, kinetic models, continuously revised to incorporate new information, must be used to guide experimental design and interpretation. Moreover, they allow the user to explore a wider variety of biological questions than can be studied using laboratory techniques alone.

In this chapter, we describe the steps needed to create a kinetic model of a metabolic pathway. We use serine biosynthesis in *E. coli* (4) as an example to apply the methodology required for this purpose.

2. Materials

2.1 Data

A mathematical description of a kinetic metabolic model may be given in differential equation form as

```
x' = Nv(x,y,p)
x(0) = x0
```

and this may be used as a guide as to the data that must be collected to create and parameterise a kinetic model.

1. First, *N* is the stoichiometric matrix, which may be derived from the network structure (Table 1). Pathway stoichiometries may be found in generic databases such as Kegg (5), or organism-specific databases such as EcoCyc (6).

2. *x* denotes metabolite concentrations; *y* denotes boundary metabolites, whose concentrations are not allowed to vary, but do affect the reaction rates. Initial concentrations for both *x* and *y* must be defined (Table 2), though note that only concentrations *x* will change over time. For human models, data is available at HMDB (7).

3. Finally, *v* denotes reaction rates; these are dependent on kinetic mechanisms (Table 3), parameters *p* (Table 4), and concentrations *x* and *y*. Databases of kinetic parameters include Brenda (8) and Sabio-RK (9).

2.2 Software

In this chapter, we will demonstrate how to build a model using Copasi (10), but the underlying principles apply to any software. The *SBML Software Guide* (11) lists alternatives that support the systems biology community standards.

3. Methods

3.1 Generating the stoichiometric matrix

1. Identify the important reactions that make up that pathway you wish to study. This information can be obtained from a network database. In this instance, three reactions are of interest: phosphoglycerate dehydrogenase, phosphoserine aminotransferase and phosphoserine phosphatase.

2. Define the metabolites involved in each reaction, including any known allosteric regulation (*see* **Note 1**) or catalysis. This information can be taken from the network database, but should be checked to ensure that the reactions are stoichiometrically consistent. In this instance the reactions are described as follows:

   3-phosphoglycerate ↔ phosphohydroxypyruvate (modifiers: serA, serine)
   phosphohydroxypyruvate ↔ phosphoserine (modifiers: serC)
   phosphoserine ↔ serine (modifiers: serB)

3. Each of the reactions outlined above added to the model. Taking the first reaction as an example, in Copasi we navigate to `Model > Biochemical > Reactions`; enter the following in the chemical equation box:

   ```
   3-phosphoglycerate = phosphohydroxypyruvate; serA serine
   ```

   (*see* **Note 2**) and add the reaction name:

   ```
   phosphoglycerate dehydrogenase
   ```

   (*see* **Note 3**) Return to `Model > Biochemical > Reactions` (*see* **Note 4**). All other reactions should be added in the same way.

4. At `Model > Biochemical > Species` you will notice that seven metabolites are listed. These are the metabolites that participate in the above reactions (*see* **Note 5**). Initially, all metabolites are assumed to be variables that change over time as the system is simulated. In order for a steady state condition to be reached, system inputs and outputs have to be defined through setting some metabolites as boundary conditions. To do this in Copasi, for each metabolite corresponding to the metabolic input or output, change `Simulation Type` from `reactions` to `Fixed`. This indicates that the concentrations of these metabolites should be constant over time. For serine biosynthesis, the following need to be `Fixed`: 3-phosphoglycerate, serine, serA, serC and serB.

5. It is important to ensure that the model is defined in the correct units. Navigate to Model, and ensure that `Time`, `Quantity Unit`, and `Volume Unit` are set accordingly. In this instance `s`, `mmol`, and `l` respectively. In addition ensure that `Rate Law Interpretation` is set to `deterministic` (*see* **Note 6**).

6. The basic reaction network has now been constructed, and the stoichiometric matrix (*N*) is now available by navigating to `Model > Mathematical > Matrices`. It should match Table 5. (*see* **Note 7**)

3.2 Adding in kinetic data

To turn the stoichiometric model into a kinetic model, rate laws have to be defined for each reaction. The rate laws define the kinetic rate (*v*) at which the substrate metabolite/s are converted into product/s.

1. A rate law that mathematically captures reaction behaviour is added to each reaction (*see* **Note 8**). Here we use irreversible Michaelis-Menten kinetics, as can be seen in Table 3.

2. The derived rate laws must be added to corresponding reactions. As an example we take phosphoglycerate dehydrogenase. In Copasi, we navigate to the corresponding reaction in the `Reactions` menu, then add a `New Rate Law` with formula:

```
serA*kcatA*(P3G/KAp3g)/(1+P3G/KAp3g+PHP/KAphp)/(1+SER/KiAser)
```

   Once the rate law has been entered, each separate variable within the math string must be defined as `Substrate` (`P3G`); `Product` (`PHP`); `Modifier` (`serA`, `SER`); or `Parameter` (`KcatA`, `KAp3g`, `KAphp`, `KiAser`).

3. The new rate law needs to be selected for the corresponding reaction back in the `Reaction` menu. At this point the associated parameter values can be added to the reaction, available in Table 4 (*see* **Note 9**).

3.3 Community standards

Describing mathematical models as above is unwieldy and error-prone and naturally leads to difficulties in reproduction of results. Thus standards have been developed to represent models.

1. SBML (12) is an XML-based markup language to unambiguously describe models. When structured in an SBML format, over 200 software tools can be used to study the model; this includes data integration, dynamic simulation and visualisation (11). To turn the Copasi model into SBML, use `File > Export SBML...`

2. SBML may be combined with Miriam (13) to annotate model entities. For example, pointing P3G to the database entry http://identifiers.org/obo.chebi/CHEBI:58272 allows its unambiguous identification and automatically links to many additional sources of information (*see* **Note 10**). The ChEBI ID is available in Table 2 for each metabolite. In Copasi, these may be added to the corresponding metabolite by navigating to the `Species` submenu `Annotation`.

3. BioModels.net (14) is a database that specialises in publishing mathematical models. There are more than 800 available for download in Miriam-annotated SBML format. The model described in this chapter is available at http://identifiers.org/biomodels.db/MODEL1203210000

3.3 Analysing model behaviour

The purpose of a mathematical model is to allow us to study complex cellular interactions more easily. The next important step after model construction is to begin to analyse its behaviour.

.

1. Living organisms can be exposed to varying external conditions, but need to maintain core cellular functions in order to remain viable. On the basis of this it is believed that internal cellular behaviour will attain a specific resting state that allows these functions to be carried out optimally. To calculate the steady state steady state of the model in Copasi, navigate to `Tasks > Steady state` and click `Tasks > Steady state`. The result should match Tables 6 and 7 (*see* **Note 11**).

2. Under different external conditions, the steady state required to maintain core cellular functions might change rapidly. The change in fluxes and metabolite concentrations towards a new steady state is known as transient behaviour. The transient behaviour can be assessed using `Tasks > Time Course` (*see* **Note 12**).

4. Notes

1. Allosteric regulators/modifiers are metabolites that positively or negatively affect the rate of an enzymatic reaction without being used or consumed within the reaction. They can form feedback mechanisms that regulate the flux through the pathway.

2. The metabolite names must be written with no gaps between the letters. For reversible reactions "=" is used, whilst if the reaction is irreversible then "`->`" should be used. The end of the reaction metabolites, and the beginning of the modifiers is signified using "`;`". For reactions that contain no modifiers this should be omitted.

3. The reaction names correspond to the short form available in the below Table (insert table names). You can use either form to build the model. It is important to maintain consistency though.

4. At this point that Copasi automatically assigns a rate law to mathematically describe the reaction. We will address inserting unique rate laws in the next section.

5. This is a good time to check the metabolite list. If there are more or less than the expected number of metabolites present it is likely that there is an error in one of the equations. This should be rectified before proceeding.

6. Deterministic modelling is appropriate when the number of molecules in the system is large. When the number of molecules is small, stochastic modelling may prove more appropriate.

7. Only metabolites that can change over time are included in the stoichiometric matrix, so all `Fixed` metabolites will not appear. The names present in the matrix will also depend on the naming conventions used in Step 3.

8. These can be correctly measured in laboratory experiments. Where this is not possible, there are a range of equations that can be used in order to make an approximation of the reaction rate (v) (15-19). In addition, Copasi also has a collection of equations that can be applied to reactions when the actual kinetic mechanism is not known.

9. When there are multiple substrates, products of modifiers for a given reaction ensure that the parameters correspond to the correct metabolite. This can be altered using the drop down menu next to the parameter.

10. ChEBI (20) is a dictionary of small chemical compounds that unambiguously states the molecular structure and properties of a metabolite. By adding the ChEBI identifier to the compound it allows the compound to always be identified regardless of the name used within the model.

11. Steady states may not closely match the experimental values and thus suggest discrepancies in the model. Refining or fitting the model, or adding new experimental data, can improve the model's utility.

12. You can change the duration of the time course and its granularity, allowing more complex behaviour to be accurately tracked.

Acknowledgement

KS is grateful for the financial support of the EU FP7 (KBBE) grant 289434 "BioPreDyn: New Bioinformatics Methods and Tools for Data-Driven Predictive Dynamic Modelling in Biotechnological Applications". We thank Daniel Jameson for taking the time to comment on the manuscript.

References

1. Lazebnik, Y. (2002) Can a biologist fix a radio? Or, what I learned while studying apoptosis. *Cancer Cell* **2**, 179–182.
2. Mendes, P., and Kell, D. (1998) Non-linear optimization of biochemical pathways: Applications to metabolic engineering and parameter estimation. *Bioinforma* **14**, 869–883.
3. Szallasi, Z., Stelling, J., and Periwal, V. (2006) System Modeling in Cellular Biology: From Concepts to Nuts and Bolts. MIT Press, Boston.
4. Pizer L.I. (1963) The pathway and control of serine biosynthesis in Escherichia coli. *J Biol Chem* **238**, 3934-3944
5. Kanehisa, M., Goto, S., Sato, Y. et al. (2012) KEGG for integration and interpretation of large-scale molecular datasets. *Nucleic Acids Res* **40**, D109-D114 http://www.genome.jp/kegg-bin/show_pathway?org_name=eco&mapno=00260
6. Keseler, I.M., Collado-Vides, J., Santos-Zavaleta, A., et al. (2011) EcoCyc: a comprehensive database of Escherichia coli biology. *Nucleic Acids Res* **39**, D583-590 http://ecocyc.org/ECOLI/NEW-IMAGE?object=SERSYN-PWY
7. Wishart D.S., Tzur D., Knox C., et al. (2007) HMDB: the Human Metabolome Database. *Nucleic Acids Res* **35**, D521-D526
8. Scheer M., Grote A., Chang, I., et al. (2010) BRENDA, the enzyme information system in 2011. *Nucleic Acids Res* **39**, D670-D676
9. Wittig U., Golebiewski, M., Kania, R., et al. (2006) SABIO-RK: integration and curation of reaction kinetics data. *Lect Notes Bioinforma* **4075**, 94-103
10. Hoops S., Sahle S., Gauges R., et al. (2006) COPASI - a COmplex PAthway SImulator. *Bioinforma* **22**, 3067-3074
11. SBML software guide. http://sbml.org/SBML_Software_Guide
12. Hucka M., Finney A., Sauro H.M., et al. (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinforma* **19**, 524-531
13. Le Novère N., Finney A., Hucka M., et al. (2005) Minimum information requested in the annotation of biochemical models (MIRIAM). *Nat Biotechnol* **23**, 1509-1515
14. Li C, Donizelli M, Rodriguez N, et al. (2010) BioModels Database: An enhanced, curated and annotated resource for published quantitative kinetic models. *BMC Syst Biol* **4**:92.
15. Heijnen J.J. (2005) Approximative kinetic formats used in metabolic network modeling. *Biotechnol Bioeng* **91**, 534-545
16. Liebermeister W., Klipp, E. (2006) Bringing metabolic networks to life: convenience rate law and thermodynamic constraints. *Theor Biol Med Model* **3**, 41
17. Savageau, M.A. (1976) Biochemical systems analysis: a study of function and design in molecular biology. Addison-Wesley, Boston
18. Smallbone K., Simeonidis, E., Broomhead, D.S, et al. (2007) Something from nothing: bridging the gap between constraint-based and kinetic modelling. *FEBS J* **274**, 5576-5585
19. Visser D., Heijnen J.J. (2003) Dynamic simulation and metabolic re-design of a branched pathway using linlog kinetics. *Metab Eng* **5**, 164-176

20. Degtyarenko, K., de Matos, P., Ennis, M., et al. (2008) ChEBI: a database and ontology for chemical entities of biological interest. *Nucleic Acids Res* **36**, D344–D350.
21. Ao P., Lee L.W., Lidstrom M.E., et al. (2008) Towards kinetic modeling of global metabolic networks: Methylobacterium extorquens AM1 growth as validation. *Chin J Biotechnol* **24**, 980-994
22. Turnaev I.I., Ibragimova S.S., Usuda Y., et al. (2006) Mathematical modeling of serine and glycine synthesis regulation in Escherichia coli. *Proceedings of the Fifth International Conference on Bioinformatics of Genome Regulation and Structure* **2**, 78-83
23. Zhao G., Winkler M.E. (1996) A novel alpha-ketoglutarate reductase activity of the serA-encoded 3-phosphoglycerate dehydrogenase of Escherichia coli K-12 and its possible implications for human 2-hydroxyglutaric aciduria. *J Bacteriol* **178**, 232-239
24. Drewke C., Klein M., Clade D., et al. (1996) 4-O-phosphoryl-L-threonine, a substrate of the pdxC(serC) gene product involved in vitamin B6 biosynthesis. *FEBS Lett* **390**, 179-182

Table 1: Network stoichiometry. Note that, for clarity, we have ignored the role of cofactors (such as NAD) in this model. Legend: PDH, phosphoglycerate dehydrogenase; PSA, phosphoserine aminotransferase; PSP, phosphoserine phosphatase; P3G, 3-phosphoglycerate; PHP, phosphohydroxypyruvate; PSER, phosphoserine; SER, serine

reaction	**stoichiometry**	**E.C. code**
PDH	P3G ↔ PHP	1.1.1.95
PSA	PHP ↔ PSER	2.6.1.52
PSP	PSER ↔ SER	3.1.3.3

Table 2: Initial metabolite concentrations.

metabolite	**concentration (mM)**	**ChEBI**	**reference**
P3G	2.36	58272	(21)
PHP	0.60	18110	(21)
PSER	0.09	57524	(21)
SER	4.90	17115	(21)

Table 3: Kinetic rate laws. More complex forms are available (22), but are simplified here, for clarity.

reaction	**formula**
PDH	`serA*kcatA*(P3G/KAp3g)/(1+P3G/KAp3g+PHP/KAphp)/(1+SER/KiAser)`
PSA	`serC*kcatC*(PHP/KCphp)/(1+PHP/KCphp+PSER/KCpser)`
PSP	`serB*kcatB*(PSER/KBpser)/(1+PSER/KBpser+SER/KBser)`

Table 4: Kinetic parameters.

parameter	value	units	reference
serA	1.15	mM	(22)
kcatA	0.55	1/s	(23)
KAp3g	1.2	mM	(23)
KAphp	0.0032	mM	(23)
KiAser	0.0038	mM	(23)
serC	0.1	mM	(22)
kcatC	1.75	1/s	(24)
KCphp	0.0015	mM	(24)
KCpser	0.0017	mM	(24)
serB	0.25	mM	(22)
kcatB	1.43	1/s	(22)
KBpser	0.0015	mM	(22)
KBser	0.15	mM	(22)

Table 5: Stoichiometric matrix.

	PDH	PSA	PSP
PHP	1	-1	0
PSER	0	1	-1

Legend: PDH, phosphoglycerate dehydrogenase; PSA, phosphoserine aminotransferase; PSP, phosphoserine phosphatase; P3G, 3-phosphoglycerate; PHP, phosphohydroxypyruvate; PSER, phosphoserine; SER, serine.

Table 6: Steady state concentrations.

metabolite	**concentration (mM)**
PHP	$2.86 \cdot 10^{-06}$
PSER	$4.59 \cdot 10^{-05}$

Table 7: Steady state fluxes.

reaction	**flux (mM/s)**
PDH	$3.24 \cdot 10^{-04}$
PSA	$3.24 \cdot 10^{-04}$
PSP	$3.24 \cdot 10^{-04}$