Analysis of Cell-Free Synthetic Circuits using Sequence Specific Constraints Based Modeling

Michael Vilkhovoy, Nicholas Horvath, and Jeffrey D. Varner*

Robert Frederick Smith School of Chemical and Biomolecular Engineering, Cornell University, Ithaca, NY 14853

E-mail: jdv27@cornell.edu

Phone: +1 607 2554258. Fax: +1 607 2559166

Abstract

In this study, we used sequence specific constraints based modeling to evaluate the performance of synthetic circuits in an *E. coli* TX-TL system. A core *E. coli* metabolic model, consisting of XX metabolites and YY reactions, was developed from literature [REF]. This model, which described glycolysis, pentose phosphate pathway, amino acid biosynthesis and degradation and energy metabolism, was then augmented with sequence specific descriptions of genetic circuits which included mechanistic models of promoter function, transcription and translation. Thus, unlike other synthetic biology modeling efforts, sequence specific constraints based modeling explicitly couples the transcription and translation of circuit components with the availability of metabolic resources. Model parameters were largely taken from literature; our approach had very few adjustable parameters thereby allowing the a first principles prediction of circuit

*To whom correspondence should be addressed

performance. We tested this approach by first simulating σ_{70}-induced deGFP expression and then expanded these studies to more complex multicomponent circuits. First principles predictions of circuit performance were consistent with measurements for a variety of cases. Further, global sensitivity analysis identified the key metabolic processes that controlled circuit performance. Taken together, sequence specific constraints based modeling offers a novel means to *a priori* estimate the performance of cell free synthetic circuits.

Keywords

Synthetic biology, Constraints based modeling, Biochemical modeling

1 Introduction

Cell free systems offer many advantages for the study, manipulation and modeling of metabolism compared to *in vivo* processes. Central amongst these advantages is direct access to metabolites and the microbial biosynthetic machinery without the interference of a cell wall. This allows us to control as well as interrogate the chemical environment while the biosynthetic machinery is operating, potentially at a fine time resolution. Second, cell-free systems also allow us to study biological processes without the complications associated with cell growth. Cell-free protein synthesis (CFPS) systems are arguably the most prominent examples of cell-free systems used today (*1*). However, CFPS is not new; CFPS in crude *E. coli* extracts has been used since the 1960s to explore fundamentally important biological mechanisms (*2*, *3*). Today, cell-free systems are used in a variety of applications ranging from therapeutic protein production (*4*) to synthetic biology (*5*). Interestingly, many of the challenges confronting in-vivo genome-scale kinetic modeling can potentially be overcome in a cell-free system. For example, there is no complex transcriptional regulation to consider, transient metabolic measurements are easier to obtain, and we no longer have to consider cell growth.

Thus, cell-free operation holds several significant advantages for model development, identification and validation. Theoretically, genome-scale cell-free kinetic models may be possible for industrially important organisms, such as *E. coli* or *B. subtilis*, if a simple, tractable framework for integrating allosteric regulation with enzyme kinetics can be formulated.

Stoichiometric reconstructions of microbial metabolism popularized by constraint based modeling techniques such as flux balance analysis (FBA) have become standard tools to interrogate biological networks (*6*). Since the first genome-scale stoichiometric model of *E. coli*, developed by Edwards and Palsson (*7*), stoichiometric reconstructions of hundreds of organisms, including industrially important prokaryotes such as *E. coli* (*8*) or *B. subtilis* (*9*), are now available (*10*). Stoichiometric models rely on a pseudo-steady-state assumption to reduce unidentifiable genome-scale kinetic models to an underdetermined linear algebraic system, which can be solved efficiently even for large systems using linear programming. Traditionally, stoichiometric models have also neglected explicit descriptions of metabolic regulation and control mechanisms, instead opting to describe the choice of pathways by prescribing an objective function on metabolism. Interestingly, similar to early cybernetic models, the most common metabolic objective function has been the optimization of biomass formation (*11*), although other metabolic objectives have also been estimated (*12*). Recent advances in constraint-based modeling have overcome the early shortcomings of the platform, including capturing metabolic regulation and control (*13*). Thus, modern constraint-based approaches are extremely useful for the discovery of metabolic engineering strategies and represent the state of the art in metabolic modeling (*14*, *15*).

In this study, we used sequence specific constraints based modeling to evaluate the performance of synthetic circuits in an *E. coli* TX-TL system. A core *E. coli* cell free metabolic model, consisting of XX metabolites and YY reactions, was developed from literature [REF]. This model, which described glycolysis, pentose phosphate pathway, amino acid biosynthesis and degradation and energy metabolism, was then augmented with sequence specific descriptions of genetic circuits which included mechanistic models of promoter function,

transcription and translation. Thus, sequence specific constraints based modeling explicitly couples the transcription and translation of circuit components with the availability of metabolic resources. Model parameters were largely taken from literature; our approach had very few adjustable parameters thereby allowing the a first principles prediction of circuit performance. We tested this approach by first simulating σ_{70}-induced deGFP expression and then expanded these studies to more complex multicomponent circuits. First principles predictions of circuit performance were consistent with measurements for a variety of cases. Further, global sensitivity analysis identified the key metabolic processes that controlled circuit performance. Taken together, sequence specific constraints based modeling offers a novel means to *a priori* estimate the performance of cell free synthetic circuits.

2 Results and discussion

2.1 Outline

The document layout should follow the style of the journal concerned. Where appropriate, sections and subsections should be added in the normal way. If the class options are set correctly, warnings will be given if these should not be present.

2.2 References

The class makes various changes to the way that references are handled. The class loads natbib, and also the appropriate bibliography style. References can be made using the normal method; the citation should be placed before any punctuation, as the class will move it if using a superscript citation style The use of natbib allows the use of the various citation commands of that package: **?** have shown something, in **?** , or as given by Ref. **?** . Long lists of authors will be automatically truncated in most article formats, but not in supplementary information or reviews. If you encounter problems with the citation macros, please check that your copy of natbib is up to date. The demonstration database file `achemso-demo.bib`

shows how to complete entries correctly. Notice that "et al." is auto-formatted using the `\latin` command.

Multiple citations to be combined into a list can be given as a single citation. This uses the mcitelus package (**?**). Citations other than the first of the list should be indicated with a star. If the mcitelus package is not installed, the standard bibliography tools will still work but starred references will be ignored. Individual references can be referred to using `\mciteSubRef`: "ref. **??**".

The class also handles notes to be added to the bibliography. These should be given in place in the document (*16*). As with citations, the text should be placed before punctuation. A note is also generated if a citation has an optional note. This assumes that the whole work has already been cited: odd numbering will result if this is not the case .

2.3 Floats

New float types are automatically set up by the class file. The means graphics are included as follows (Scheme 1). As illustrated, the float is "here" if possible.

Your scheme graphic would go here: `.eps` format
for LaTeX or `.pdf` (or `.png`) for pdfLaTeX
ChemDraw files are best saved as `.eps` files:
these can be scaled without loss of quality, and can be
converted to `.pdf` files easily using `eps2pdf`.

Scheme 1: An example scheme

As well as the standard float types `table`
and `figure`, the class also recognises
`scheme`, `chart` and `graph`.

Figure 1: An example figure

Charts, figures and schemes do not necessarily have to be labelled or captioned. However, tables should always have a title. It is possible to include a number and label for a graphic without any title, using an empty argument to the `\caption` macro.

The use of the different floating environments is not required, but it is intended to make document preparation easier for authors. In general, you should place your graphics where they make logical sense; the production process will move them if needed.

2.4 Math(s)

The achemso class does not load any particular additional support for mathematics. If packages such as amsmath are required, they should be loaded in the preamble. However, the basic LaTeX math(s) input should work correctly without this. Some inline material $y = mx + c$ or $1 + 1 = 2$ followed by some display.

$$A = \pi r^2$$

It is possible to label equations in the usual way (Eq. 1).

$$\frac{\mathrm{d}}{\mathrm{d}x} r^2 = 2r \tag{1}$$

This can also be used to have equations containing graphical content. To align the equation number with the middle of the graphic, rather than the bottom, a minipage may be used.

$$\begin{array}{c} \text{As illustrated here, the width of} \\ \text{the minipage needs to allow some} \\ \text{space for the number to fit in to.} \end{array} \tag{2}$$

3 Experimental

The usual experimental details should appear here. This could include a table, which can be referenced as Table 1. Notice that the caption is positioned at the top of the table.

Adding notes to tables can be complicated. Perhaps the easiest method is to generate these using the basic `\textsuperscript` and `\emph` macros, as illustrated (Table 2).

Table 1: An example table

Header one	Header two
Entry one	Entry two
Entry three	Entry four
Entry five	Entry five
Entry seven	Entry eight

Table 2: A table with notes

Header one	Header two
Entry one[a]	Entry two
Entry three[b]	Entry four

[a] Some text; [b] Some more text.

The example file also loads the optional mhchem package, so that formulas are easy to input: `\ce{H2SO4}` gives H_2SO_4. See the use in the bibliography file (when using titles in the references section).

The use of new commands should be limited to simple things which will not interfere with the production process. For example, `\mycommand` has been defined in this example, to give italic, mono-spaced text: *some text*.

4 Extra information when writing JACS Communications

When producing communications for *J. Am. Chem. Soc.*, the class will automatically lay the text out in the style of the journal. This gives a guide to the length of text that can be accommodated in such a publication. There are some points to bear in mind when preparing a JACS Communication in this way. The layout produced here is a *model* for the published result, and the outcome should be taken as a *guide* to the final length. The spacing and sizing of graphical content is an area where there is some flexibility in the process. You should not worry about the space before and after graphics, which is set to give a guide to the published size. This is very dependant on the final published layout.

You should be able to use the same source to produce a JACS Communication and a normal article. For example, this demonstration file will work with both `type=article` and `type=communication`. Sections and any abstract are automatically ignored, although you will get warnings to this effect.

Acknowledgement

Please use "The authors thank ..." rather than "The authors would like to thank ...".

The author thanks Mats Dahlgren for version one of achemso, and Donald Arseneau for the code taken from cite to move citations after punctuation. Many users have provided feedback on the class, which is reflected in all of the different demonstrations shown in this document.

Supporting Information Available

A listing of the contents of each file supplied as Supporting Information should be included. For instructions on what should be included in the Supporting Information as well as how to prepare this material for publications, refer to the journal's Instructions for Authors.

The following files are available free of charge.

- Filename: brief description
- Filename: brief description

This material is available free of charge via the Internet at `http://pubs.acs.org/`.

Notes and References

1. Jewett, M. C., Calhoun, K. A., Voloshin, A., Wuu, J. J., and Swartz, J. R. (2008) An integrated cell-free metabolic platform for protein production and synthetic biology. *Mol Syst Biol 4*, 220.

2. Matthaei, J. H., and Nirenberg, M. W. (1961) Characteristics and stabilization of DNAase-sensitive protein synthesis in E. coli extracts. *Proc Natl Acad Sci U S A 47*, 1580–8.

3. Nirenberg, M. W., and Matthaei, J. H. (1961) The dependence of cell-free protein synthesis in E. coli upon naturally occurring or synthetic polyribonucleotides. *Proc Natl Acad Sci U S A 47*, 1588–602.

4. Lu, Y., Welsh, J. P., and Swartz, J. R. (2014) Production and stabilization of the trimeric influenza hemagglutinin stem domain for potentially broadly protective influenza vaccines. *Proc Natl Acad Sci U S A 111*, 125–30.

5. Hodgman, C. E., and Jewett, M. C. (2012) Cell-free synthetic biology: thinking outside the cell. *Metab Eng 14*, 261–9.

6. Lewis, N. E., Nagarajan, H., and Palsson, B. Ø. (2012) Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods. *Nat Rev Microbiol 10*, 291–305.

7. Edwards, J. S., and Palsson, B. Ø. (2000) The Escherichia coli MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. *Proc Natl Acad Sci U S A 97*, 5528–33.

8. Feist, A. M., Henry, C. S., Reed, J. L., Krummenacker, M., Joyce, A. R., Karp, P. D., Broadbelt, L. J., Hatzimanikatis, V., and Palsson, B. Ø. (2007) A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol Syst Biol 3*, 121.

9. Oh, Y.-K., Palsson, B. Ø., Park, S. M., Schilling, C. H., and Mahadevan, R. (2007) Genome-scale reconstruction of metabolic network in Bacillus subtilis based on high-throughput phenotyping and gene essentiality data. *J Biol Chem 282*, 28791–9.

10. Feist, A. M., Herrgård, M. J., Thiele, I., Reed, J. L., and Palsson, B. Ø. (2009) Reconstruction of biochemical networks in microorganisms. *Nat Rev Microbiol 7*, 129–43.

11. Ibarra, R. U., Edwards, J. S., and Palsson, B. Ø. (2002) Escherichia coli K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature 420*, 186–9.

12. Schuetz, R., Kuepfer, L., and Sauer, U. (2007) Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli. *Mol Syst Biol 3*, 119.

13. Hyduke, D. R., Lewis, N. E., and Palsson, B. Ø. (2013) Analysis of omics data with genome-scale models of metabolism. *Mol Biosyst 9*, 167–74.

14. McCloskey, D., Palsson, B. Ø., and Feist, A. M. (2013) Basic and applied uses of genome-scale metabolic network reconstructions of Escherichia coli. *Mol Syst Biol 9*, 661.

15. Zomorrodi, A. R., Suthers, P. F., Ranganathan, S., and Maranas, C. D. (2012) Mathematical optimization applications in metabolic networks. *Metab Eng 14*, 672–86.

16. This is a note. The text will be moved the the references section. The title of the section will change to “Notes and References”.

Graphical TOC Entry

Some journals require a graphical entry for the Table of Contents. This should be laid out "print ready" so that the sizing of the text is correct.
Inside the `tocentry` environment, the font used is Helvetica 8 pt, as required by *Journal of the American Chemical Society*.
The surrounding frame is 9 cm by 3.5 cm, which is the maximum permitted for *Journal of the American Chemical Society* graphical table of content entries. The box will not resize if the content is too big: instead it will overflow the edge of the box.
This box and the associated title will always be printed on a separate page at the end of the document.