Dynamic Constraints Based Analysis of Cell Free Metabolic Networks

Michael Vilkhovoy*, Mason Minot*, Jeffrey D. Varner*
*Robert Frederick Smith School of Chemical and Biomolecular Engineering, Cornell University, Ithaca, NY 14850 USA

***Abstract*—Mathematical models of biochemical networks are useful tools to understand and ultimately predict how cells utilize nutrients to produce valuable products. Hybrid cybernetic models in combination with elementary modes (HCM) is a tool to model cellular metabolism. However, HCM is limited to reduced metabolic networks because of the computational burden of calculating elementary modes. In this study, we developed the hybrid cybernetic modeling with flux balance analysis or HCM-FBA technique which uses flux balance solutions instead of elementary modes to dynamically model metabolism. We show HCM-FBA has comparable performance to HCM for a proof of concept metabolic network and for a reduced anaerobic *E. coli* network. Next, HCM-FBA was applied to a larger metabolic network of aerobic *E. coli* metabolism which was infeasible for HCM (29 FBA modes versus more than 153,000 elementary modes). Global sensitivity analysis further reduced the number of FBA modes required to describe the aerobic *E. coli* data, while maintaining model fit. Thus, HCM-FBA is a promising alternative to HCM for large networks where the generation of elementary modes is infeasible.**

***Index Terms*—Metabolic models, flux balance analysis, cybernetic models**

I. 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 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.

Corresponding author: J. Varner (email:jdv27@cornell.edu).

In the post genomics world, genome-scale stoichiometric reconstructions of microbial metabolism popularized by static, constraint-based modeling techniques such as flux balance analysis (FBA) have become standard tools [6]. Since the first genome-scale stoichiometric model of *E. coli*, developed by Edwards and Palsson [7], well over 100 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. 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, constraint-based approaches have proven useful in the discovery of metabolic engineering strategies and represent the state of the art in metabolic modeling [14, 15].

II. Results

III. Discussion

IV. Materials and Methods

The balance equation governing species i (x_i) for a cell free metabolic network consisting of $\mathcal{M}$ metabolites and enzymes is given by:

$$\frac{dx_i}{dt} = \sum_{j=1}^{\mathcal{C}} \sigma_{ij} r_j + \sum_{d=1}^{\mathcal{K}} \tau_{id} q_d(\mathbf{x}) \qquad i = 1, \ldots, \mathcal{M} \qquad (1)$$

where $\mathcal{C}$ denotes the number calculated fluxes, and $\mathcal{K}$ denotes the number of kinetic fluxes. The terms σ_{ij} and τ_{ik} denote stoichiometric coefficients which connect the calculated and kinetic rates r_j and $q_k(\mathbf{x})$ to the ith species. If $\sigma_{ij} > 0$ or $\tau_{ik} > 0$, then species i is produced, if $\sigma_{ij} < 0$ or $\tau_{ik} < 0$ species i is consumed. If $\sigma_{ij} = 0$ or $\tau_{ik} = 0$ then species is not connected with either rate process. The balances equations are subject to the initial conditions $x_i(0)$.

The $\mathcal{K}$ kinetic rates take typical saturation forms. The $\mathcal{C}$ calculated rates are estimated by solving a convex optimization

subproblem at each time step k. In this study, we explored both linear and non-linear implementations of the flux estimation subproblem. The convex subproblem was solved at each time step k, using either the GNU Linear Programming Kit (GLPK) plugin for the linear problem [REF], or the Convex plugin for the non-linear problem [REF]. The model code for all examples, encoded in the Julia programming language [REF], can be downloaded from http://www.varnerlab.org under an MIT license.

The convex subproblem consists of an objective function, material balance and flux bounds constraints. The linear objective function:

$$\min_{r_1,\ldots,r_{\mathcal{C}}} \left(\sum_{j=1}^{\mathcal{C}} c_j r_j(k) \right) \tag{2}$$

is minimized subject to the material balance constraints:

$$\sum_{j=1}^{\mathcal{C}} \sigma_{ij} r_j(k) \geq -\left(\frac{\theta_i}{h} + \sum_{d=1}^{\mathcal{K}} \tau_{id} q_d(\mathbf{x}) \right) \;\; i = 1, \ldots, \mathcal{M}$$

and flux bounds constraints:

$$\mathcal{L}_j(\mathbf{x}) \leq r_j \leq \mathcal{U}_j(\mathbf{x}) \qquad j = 1, \ldots, \mathcal{C}$$

at each time step k. The user configurable parameter θ_i, defined by $x_{i,k+1} = (1 - \theta_i)\, x_{i,k}$, governs the maximum rate of decrease of the ith species. The $\mathcal{L}_j(\mathbf{x})$ and $\mathcal{U}_j(\mathbf{x})$ functions encode the lower and upper bound for the jth flux, while h denotes the time step size.

REFERENCES

1. Jewett, M. C., Calhoun, K. A., Voloshin, A., Wuu, J. J. & Swartz, J. R. An integrated cell-free metabolic platform for protein production and synthetic biology. *Mol Syst Biol* **4**, 220 (2008).
2. Matthaei, J. H. & Nirenberg, M. W. Characteristics and stabilization of dnaase-sensitive protein synthesis in e. coli extracts. *Proc Natl Acad Sci U S A* **47**, 1580–8 (1961).
3. Nirenberg, M. W. & Matthaei, J. H. 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 (1961).
4. Lu, Y., Welsh, J. P. & Swartz, J. R. 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 (2014).
5. Hodgman, C. E. & Jewett, M. C. Cell-free synthetic biology: thinking outside the cell. *Metab Eng* **14**, 261–9 (2012).
6. Lewis, N. E., Nagarajan, H. & Palsson, B. O. Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods. *Nat Rev Microbiol* **10**, 291–305 (2012).
7. Edwards, J. S. & Palsson, B. O. The escherichia coli mg1655 in silico metabolic genotype: its definition, characteristics, and capabilities. *Proc Natl Acad Sci U S A* **97**, 5528–33 (2000).
8. Feist, A. M. *et al.* A genome-scale metabolic reconstruction for escherichia coli k-12 mg1655 that accounts for 1260 orfs and thermodynamic information. *Mol Syst Biol* **3**, 121 (2007).
9. Oh, Y.-K., Palsson, B. O., Park, S. M., Schilling, C. H. & Mahadevan, R. Genome-scale reconstruction of metabolic network in bacillus subtilis based on high-throughput phenotyping and gene essentiality data. *J Biol Chem* **282**, 28791–9 (2007).
10. Feist, A. M., Herrgård, M. J., Thiele, I., Reed, J. L. & Palsson, B. Ø. Reconstruction of biochemical networks in microorganisms. *Nat Rev Microbiol* **7**, 129–43 (2009).
11. Ibarra, R. U., Edwards, J. S. & Palsson, B. O. Escherichia coli k-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature* **420**, 186–9 (2002).
12. Schuetz, R., Kuepfer, L. & Sauer, U. Systematic evaluation of objective functions for predicting intracellular fluxes in escherichia coli. *Mol Syst Biol* **3**, 119 (2007).
13. Hyduke, D. R., Lewis, N. E. & Palsson, B. Ø. Analysis of omics data with genome-scale models of metabolism. *Mol Biosyst* **9**, 167–74 (2013).
14. McCloskey, D., Palsson, B. Ø. & Feist, A. M. Basic and applied uses of genome-scale metabolic network reconstructions of escherichia coli. *Mol Syst Biol* **9**, 661 (2013).
15. Zomorrodi, A. R., Suthers, P. F., Ranganathan, S. & Maranas, C. D. Mathematical optimization applications in metabolic networks. *Metab Eng* **14**, 672–86 (2012).