Dynamic Modeling of Cell Free Metabolic Networks using Effective Kinetic Models

Joseph A. Wayman and Jeffrey D. Varner*

School of Chemical and Biomolecular Engineering

Cornell University, Ithaca NY 14853

Running Title: Modeling cell free metabolism

To be submitted: *Processes*

*Corresponding author:

Jeffrey D. Varner,

Associate Professor, School of Chemical and Biomolecular Engineering,

244 Olin Hall, Cornell University, Ithaca NY, 14853

Email: jdv27@cornell.edu

Phone: (607) 255 - 4258

Fax: (607) 255 - 9166

Abstract

Keywords: Cell free metabolism, Mathematical modeling

Introduction

Whole-cell bacterial processes are widely used in biotechnology to produce an array of products including high-value protein therapeutics. However, whole-cell processes share the central limitation of requiring cell growth, which redirects resources away from product synthesis, and cell walls, which complicate interrogation and control of intracellular metabolic processes. On the other hand, cell-free metabolic systems offer many advantages over traditional in vivo production methods. For example, cell-free systems can direct scarce metabolic resources exclusively towards a single product of interest. Moreover, with no cell wall, cell free systems can more easily be interrogated, and substrates of the metabolite processes directly controlled. Cell free production offers the unique opportunity to study metabolism without the complication of cell growth and gene expression processes. For modeling, this implies that we need only consider allosteric regulation of enzyme activity when building and testing cell free metabolic models. Of course, modeling allosteric mechanisms is itself a difficult problem when the model is at a whole genome scale. To address this problem, we have developed a an approach based upon the constrained fuzzy logic framework of Morris and Lauffenburger [REFHERE].

In this study, we present an effective cell free metabolic modeling framework, and test this framework using two proof of concept metabolic networks. [FINISH].

Results

We developed two proof of concept metabolic networks to investigate the features of our effective cell free modeling approach (Fig. 1). In both examples, substrate S was converted to the end-products P_1 and P_2 through a series of enzymatically catalyzed reactions, including a branch point at hypothetical metabolite M_2. Several of these reactions involved cofactor dependence (AH or A), and various allosteric regulation mechanisms. Network A included feedback inhibition of the initial pathway enzyme (E_1) by pathway end products P_1 and P_2 (Fig. 1A). On the other hand, network B involved feedback inhibition of E_1 by P_2 and E_6 by P_1 (Fig. 1B). In both networks, branch point enzymes E_3 and E_6 were subject to feed-forward activation by cofactor AH. Lastly, in both cases pathway enzyme activity was assumed to decay according to a first-order rate law. Allosteric regulation of enzyme activity was represented using a novel rule-based strategy, similar in spirit to the Constrained Fuzzy Logic (cFL) approach of Lauffenberger and coworkers [1]. In this formulation, Hill-like transfer functions were used to calculate the influence of metabolite abundance upon target enzyme activity. When an enzyme was potentially sensitive to more than one regulatory influence, logical rules were used to select which transfer function regulated enzyme activity at any given time (Fig. 2). Thus, our test networks involved important features such as cofactor recycling, enzyme activity and metabolite dynamics, as well as multiple overlapping allosteric regulatory mechanisms. As such, developing our effective modeling approach using these problems gave us valuable insight into how to develop larger network models, without the complication of network size.

The rule based regulatory strategy captured the kinetic behavior of classic allosteric activation and inhibition mechanisms (Fig. 3). We first explored feed-forward substrate activation of enzyme activity (for both positive and negative cooperativity). Consistent with classical data, the rule based strategy predicted a sigmoidal relationship between substrate abundance and reaction rate as a function of the cooperativity parameter (Fig. 3A).

For cooperativity parameters less than unity, increased substrate abundance *decreased* the reaction rate. This was consistent with the idea that substrate binding *decreases* at regulatory sites negatively impacts the ability of the enzyme to bind substrate at the active site. On the other hand, as the cooperativity parameter increased past unity, the rate of conversion of substrate S to product P by enzyme E approached a step function. In the presence of an inhibitor, the rule based strategy predicted non-competitive like behavior as a function of the cooperativity parameter (Fig. 3B). When the control gain parameter, κ_{ij} in Eqn. (10), was greater than unity, the inhibitory force was directly proportional to the cooperativity parameter, η in Eqn. (10). Thus, as the cooperativity parameter increased, the maximum reaction rate decreased (Fig. 3B, orange). However, when the gain parameter was less than unity, enzyme inhibition increased with *decreasing* cooperativity, i.e., smaller η yielded increased inhibition (Fig. 3B). Interestingly, our rule based approach was not able to directly simulate competitive inhibition of enzyme activity. Taken together, the rule based strategy captured classical regulatory patterns for both enzyme activation and inhibition.

End product yield was controlled by feedback inhibition, while selectivity was controlled by branch point enzyme inhibition (Fig. 4). A critical test of our modeling approach was to simulate networks with known behavior. If we cannot reproduce the expected behavior of simple networks, then our effect modeling strategy, and particularly the rule-based approximation of allosteric regulation, will not be feasible for large scale problems. We considered two cases, control on/off, for each network configuration. Each of these cases had identical kinetic parameters and initial conditions; the *only* differences between the cases was the allosteric regulation rules, and the control parameters associated with these rules. As expected, end product accumulation was larger for network A when the control was off (no feedback inhibition of E_1 by P_1 and P_2), as compared to the on case (Fig. 4A). We found this behavior was robust to the choice of underlying kinetic parameters, as we observed that same qualitative response across an ensemble of randomized parameter sets (N = 100). The control on/off response of network B was more subtle. In the off case, the behavior was qualitatively similar to network A. However, for the on case, flux was diverted away from P_2 formation by feedback inhibition of E_6 activity at the M_2 branch point by P_1 (Fig. 4B). Lower E_6 activity at the M_2 branch point allowed more flux toward P_1 formation, hence the yield of P_1 also increased (Fig. 4C). Again, the control on/off behavior was robust to the values of the kinetic parameters, as the same qualitative trend was conserved across an ensemble of possible randomized kinetic parameters (N = 100). Taken together, these simulations suggested that the rule based allosteric control concept could robustly capture expected feedback behavior.

The estimation of kinetic parameters was sensitive to the choice of control structure (Fig. 5). A critical challenge for any dynamic model effort is the estimation of kinetic parameters. However, beyond this challenge is the identification of model structure, i.e., the biological connectivity of the system you are exploring. Of course these two issues are not independent as any description of the control of enzyme activity will be a function of the system state, which in turn is a function of the kinetic parameters.

Discrimination between competing model formulations can be achieved by optimal experimental design (Fig. ZZ). [FINISH]

Discussion

In this study, we proposed a effective modeling strategy to simulate cell free metabolic networks. We tested this strategy using two proof of concept metabolic networks with the same enzymatic connectivity, but differing control structures. [FINISH]

Cybernetic models, other dynamic models of metabolism.

While the results of this study were encouraging, there are several critical next steps that must be accomplished before we can model genome scale cell free metabolic networks. [FINISH]

Materials and Methods

Formulation and solution of the model equations. We used ordinary differential equations (ODEs) to model the time evolution of metabolite (x_i) and scaled enzyme (ϵ_i) abundance in hypothetical cell free metabolic networks:

$$\frac{dx_i}{dt} = \sum_{j=1}^{\mathcal{R}} \sigma_{ij} r_j\left(\mathbf{x}, \epsilon, \mathbf{k}\right) \qquad i = 1, 2, \ldots, \mathcal{M} \tag{1}$$

$$\frac{d\epsilon_i}{dt} = -\lambda_i \epsilon_i \qquad i = 1, 2, \ldots, \mathcal{E} \tag{2}$$

where $\mathcal{R}$ denotes the number of reactions, $\mathcal{M}$ denotes the number of metabolites and $\mathcal{E}$ denotes the number of enzymes in the model. The quantity $r_j\left(\mathbf{x}, \epsilon, \mathbf{k}\right)$ denotes the rate of reaction j. Typically, reaction j is a non-linear function of metabolite and enzyme abundance, as well as unknown kinetic parameters $\mathbf{k}$ ($\mathcal{K} \times 1$). The quantity σ_{ij} denotes the stoichiometric coefficient for species i in reaction j. If $\sigma_{ij} > 0$, metabolite i is produced by reaction j. Conversely, if $\sigma_{ij} > 0$, metabolite i is consumed by reaction j, while $\sigma_{ij} = 0$ indicates metabolite i is not connected with reaction j. Each reaction rate was written as the product of a reaction term ($\bar{r}_j$) and a regulatory term (v_j):

$$r_j\left(\mathbf{x},\epsilon,\mathbf{k}\right)=\bar{r}_j v_j \tag{3}$$

We used multiple saturation kinetics to model the reaction term $\bar{r}_j$:

$$\bar{r}_j=k_j^{max}\epsilon_i\left(\prod_{s\in m_j^-}\frac{x_s}{K_{js}+x_s}\right) \tag{4}$$

where k_j^{max} denotes the maximum rate for reaction j, ϵ_i denotes the scaled enzyme activity which catalyzes reaction j, and K_{js} denotes the saturation constant for species s in reaction j. The product in Eqn. (4) was carried out over the set of *reactants* for reaction j (denoted as m_j^-). The allosteric regulation term v_j depended upon the combination of factors which influenced the activity of enzyme i. For each enzyme, we used a rule based approach to select from competing control factors (Fig. 2). If an enzyme was activated by m metabolites, we modeled this activation as:

$$v_j=\max\left(f_{1j}\left(x\right),\ldots,f_{mj}\left(x\right)\right) \tag{5}$$

Conversely, if enzyme activity was inhibited by a m metabolites, we modeling this inhibition as:

$$v_j=1-\max\left(f_{1j}\left(x\right),\ldots,f_{mj}\left(x\right)\right) \tag{6}$$

Lastly, if an enzyme had both m activating and n inhibitory factors, we modeled the regulatory term as:

$$v_j=\min\left(u_j,d_j\right) \tag{7}$$

where:

$$
\begin{aligned}
u_j &= \max_{j^+}\left(f_{1j}\left(x\right),\ldots,f_{mj}\left(x\right)\right) &(8)\\
d_j &= 1-\max_{j^-}\left(f_{1j}\left(x\right),\ldots,f_{nj}\left(x\right)\right) &(9)
\end{aligned}
$$

where j^+ and j^- denote the sets of activating, and inhibitory factors for enzyme j. If an enzyme had no allosteric factors, we set $v_j = 1$. In this study, each individual factor had the form:

$$f_i\left(\mathbf{x}\right) = \frac{\kappa_{ij}^{\eta}x_j^{\eta}}{1+\kappa_{ij}^{\eta}x_j^{\eta}} \tag{10}$$

where x_j denotes the abundance of metabolite j, and κ_{ij} and η are control parameters. The κ_{ij} parameter was species gain parameter, while η was a cooperativity parameter (similar to a Hill coefficient). The model equations were encoded using the Octave programming language, and solved using the LSODE routine in Octave [2].

Estimation of model parameters from synthetic experimental data Model parameters were estimated by minimizing the difference between simulations and synthetic experimental data using particle swarm optimization [3]. In this study, the parameter estimation problem took the form:

$$\min_{\mathbf{k}}\sum_{\tau=1}^{\mathcal{T}}\sum_{j=1}^{\mathcal{S}}\left(\frac{\hat{x}_j\left(\tau\right)-x_j\left(\tau,\mathbf{k}\right)}{\omega_j\left(\tau\right)}\right)^2 \tag{11}$$

where $\hat{x}_j\left(\tau\right)$ denotes the measured value of species j at time τ, $x_j\left(\tau,\mathbf{k}\right)$ denotes the simulated value for species j at time τ, and $\omega_j\left(\tau\right)$ denotes the experimental measurement variance for species j at time τ. The outer summation is respect to time, while the inner summation is with respect to state. We wanted to approximate a realistic model identification scenario, thus we assumed noisy experimental data, limited sampling resolution (approximately 20 minutes per sample) and that only a subset of metabolites were measurable.

We minimized Eqn. (11) using Particle swarm optimization (PSO). PSO is global optimization procedure which uses a swarming metaheuristic to explore parameter spaces. A strength of PSO is its ability to find the global minimum, even in the presence of potentially many local minima, by communicating the local error landscape experienced by each particle collectively to the swarm using the update rules:

$$\mathbf{\Delta}_i = \theta_1\mathbf{\Delta}_i + \theta_2\mathbf{r}_1\left(\mathcal{L}_i - \mathbf{k}_i\right) + \theta_3\mathbf{r}_2\left(\mathcal{G} - \mathbf{k}_i\right) \tag{12}$$

$$\mathbf{k}_i = \mathbf{k}_i + \mathbf{\Delta}_i \tag{13}$$

where $(\theta_1, \theta_2, \theta_3) = (1.0, 0.05564, 0.02886)$ are adjustable parameters, $\mathcal{L}_i$ denotes local best solution found by particle i, and $\mathcal{G}$ denotes the best solution found over the entire population of particles. The quantities r_1 and r_2 denote uniform random vectors with the same dimension as the number of unknown model parameters ($\mathcal{K} \times 1$). The PSO algorithm, and the objective function were encoded and solved in the Octave programming language [2].

Model discrimination

Acknowledgements

This study was supported by the National Science Foundation GK12 award (DGE-1045513) and by the National Science Foundation CAREER award (FILLMEIN).

References

1. Morris MK, Saez-Rodriguez J, Clarke DC, Sorger PK, Lauffenburger DA (2011) Training signaling pathway maps to biochemical data with constrained fuzzy logic: quantitative analysis of liver cell responses to inflammatory stimuli. PLoS Comput Biol 7: e1001099.

2. Octave community (2014). GNU Octave 3.8.1. URL www.gnu.org/software/octave/.

3. Kennedy J, Eberhart R (1995) Particle swarm optimization. In: Proceedings of the International Conference on Neural Networks. pp. 1942 - 1948.

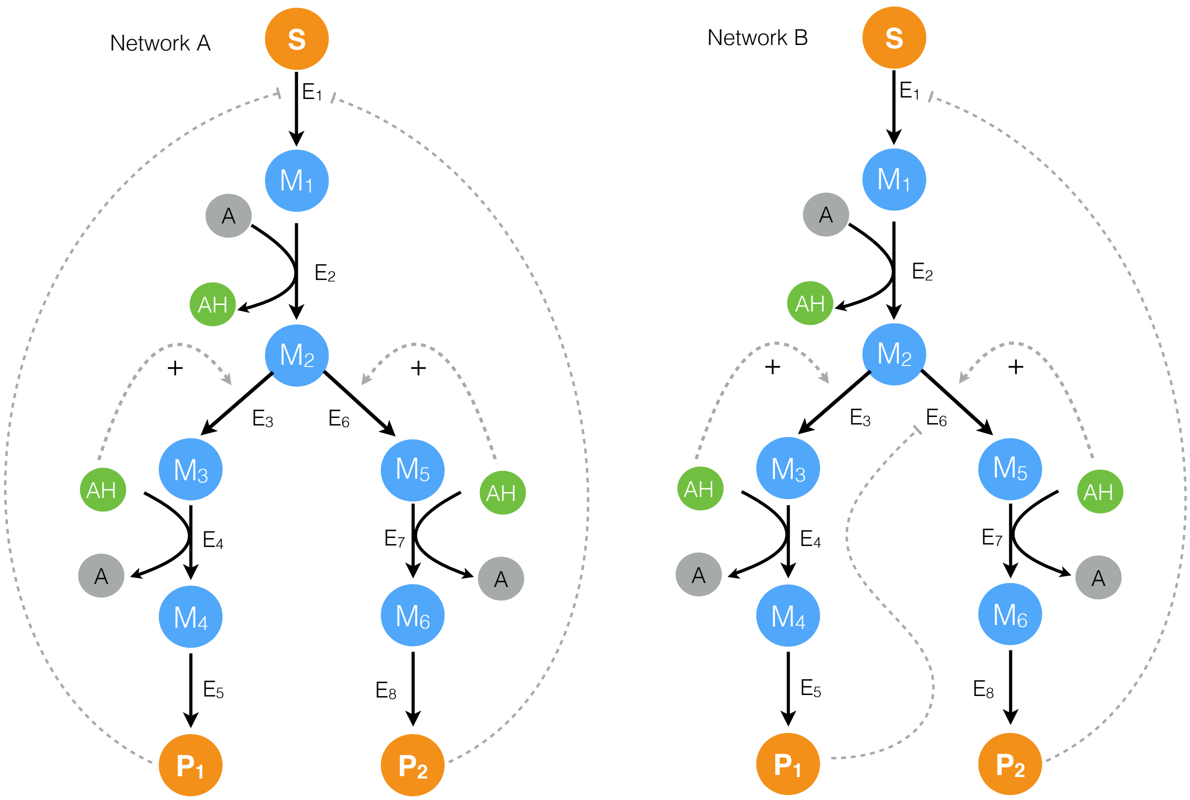

Fig. 1: Proof of concept cell-free metabolic networks considered in this study. Substrate S is converted to products P_1 and P_2 through a series of chemical conversions catalyzed by enzyme(s) E_j. The activity of the pathway enzymes is subject to both positive and negative allosteric regulation.

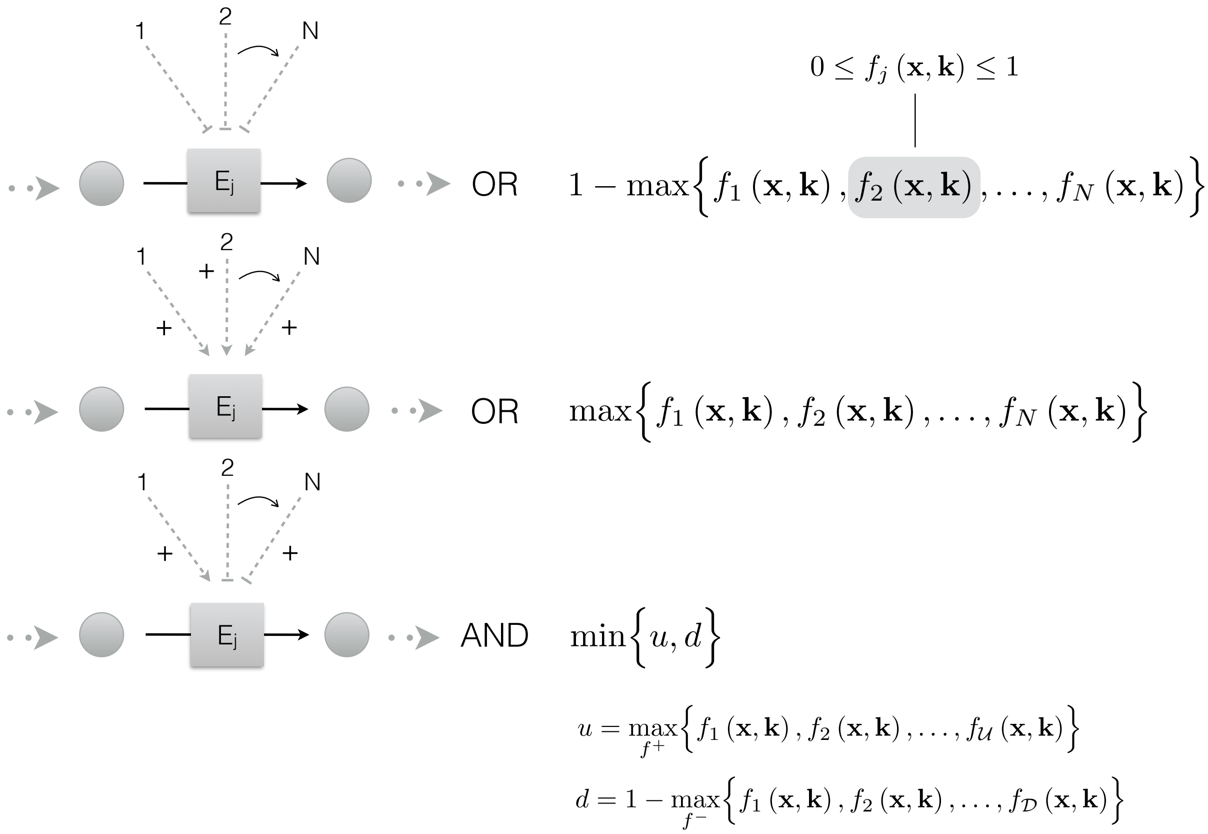

Fig. 2: Schematic of the rule based allosteric enzyme activity control laws.

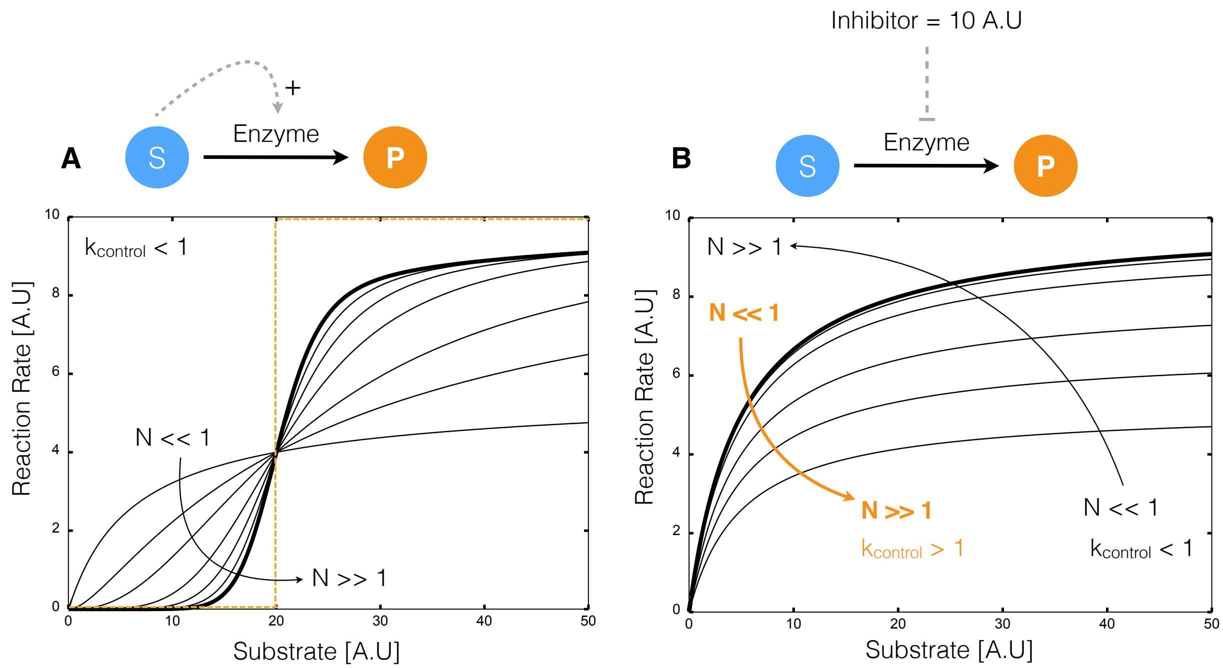

Fig. 3: Kinetics of simple transformations in the presence of activation and inhibition. A:The conversion of substrate S to product P by enzyme E was activated by S. B:The conversion of substrate S to product P by enzyme E was inhibited by inhibitor I.

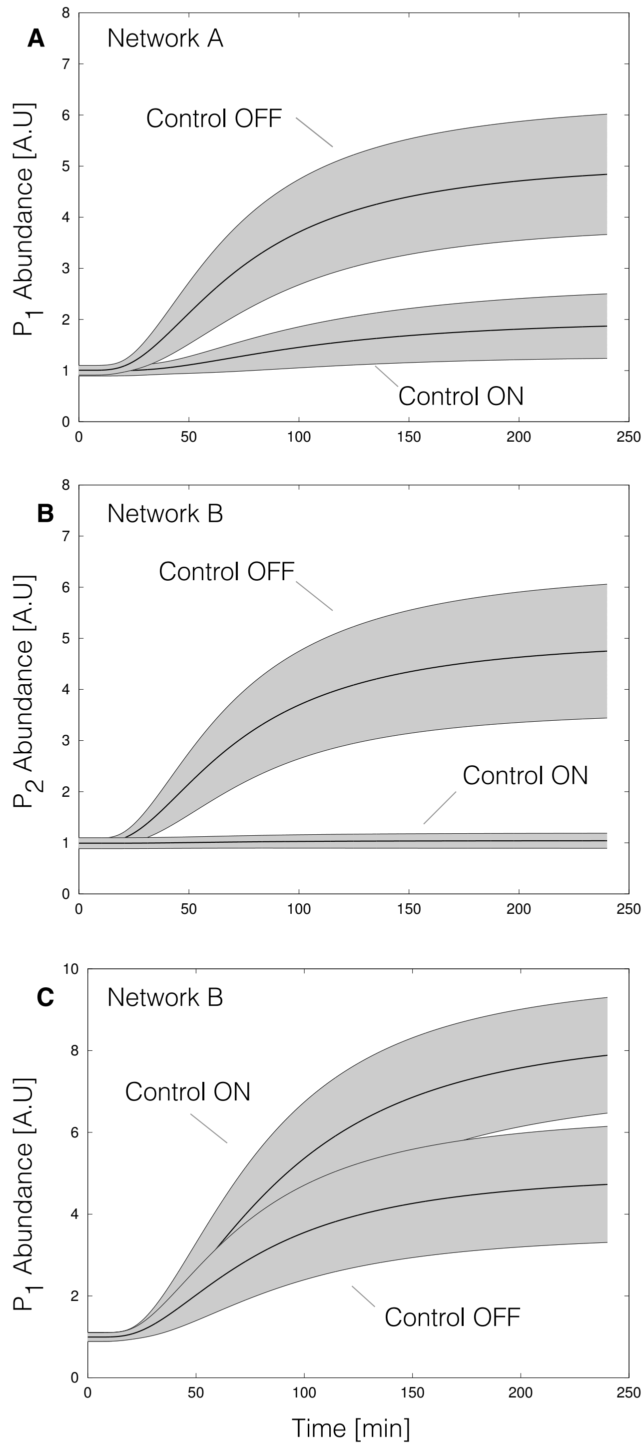

Fig. 4: On/off control simulations for network A and network B for an ensemble of kinetic parameter sets versus time. For each case, N = 100 simulations were conducted using kinetic and initial conditions randomly generated from a hypothetical true parameter set. The gray area represents $\pm$ one standard deviation surrounding the mean. Control parameters were fixed during the ensemble calculations.

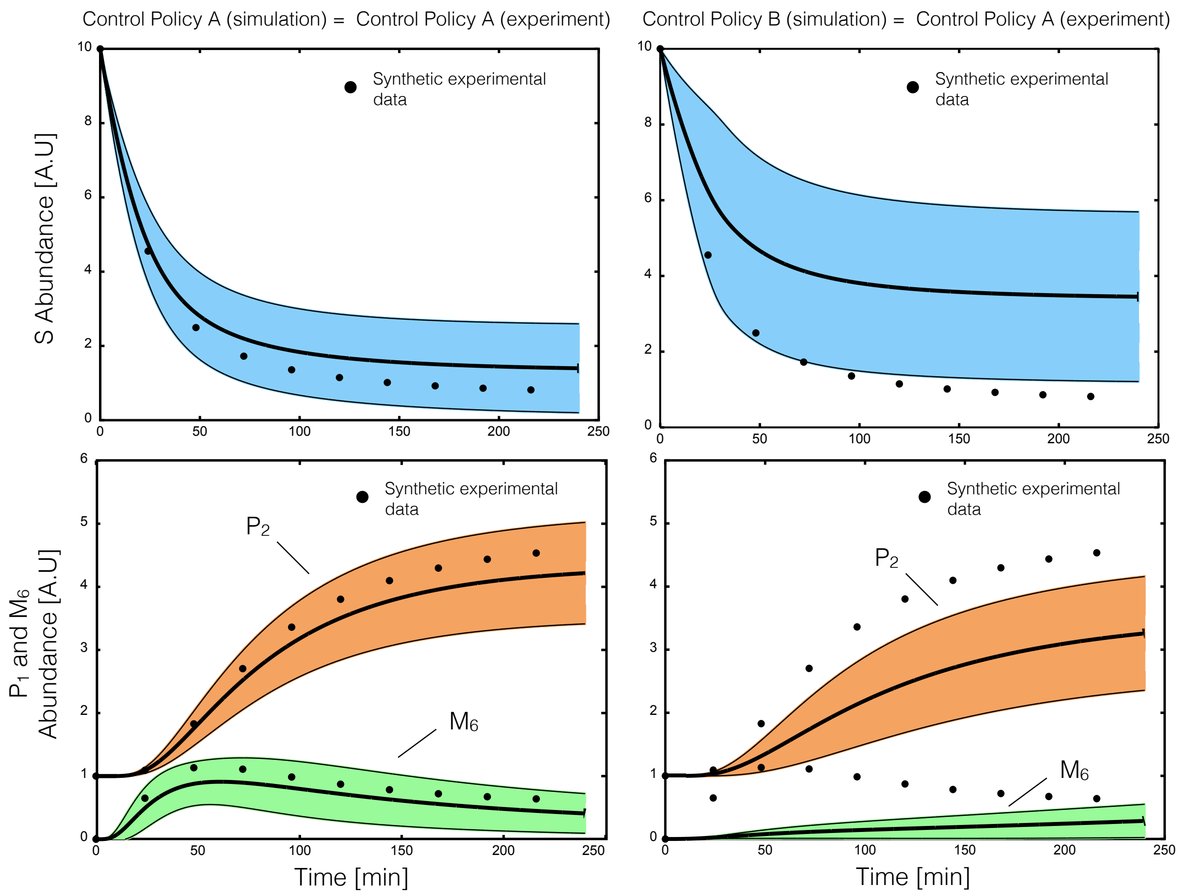

Fig. 5: Parameter estimation from synthetic data for the same and mismatched allosteric control logic.