Supporting Information

Chubukov et al. 10.1073/pnas.1114235109

SI Text

The following is a list of terms appearing in the differential equation model and explanations for each term.

Model Species.

E_1–E_N: The concentrations of the enzymes of the pathway.

S_0–S_N: The concentrations of the metabolites in the pathway. We assume that $S_0 = \infty$ (i.e., the upstream source is not limiting).

Renamed Species or Variables Calculated Directly from Model Species.

P: The final product of the pathway (S_N).

S_{act}: One of S_1-S_N; the metabolite that interacts with the transcription factor to influence its activity. When $S_{act} = S_N$ (i.e., the end product inhibition architecture) the model is identical except that we change the transcription factor to a repressor.

TF: The fraction of active transcription factor (bound by the metabolic intermediate). We assume that the transcription factor is constitutively bound to the DNA, as is known to be the case for a number of real TFs in the pathways we worked on (i.e., Leu3 and Lys14).

Fixed Parameters.

v_1–v_N: The catalytic constants (rate of enzyme conversion) associated with the enzymes. Changing these affects only the units of E_1-E_N so they are set to be equal to 1.

k_1–k_N: The Michaelis-Menten constants associated with the enzymes. Changing these affects only the units of S_0-S_N so they are set to be equal to 1.

δ: The sum of the rates of dilution and degradation of the enzymes and metabolites. We assume that dilution due to cell division is the major component and as such, use the same δ for all species in the model. Changing this sets the units of the time axis; we set $\delta = 1$ and measure time in (cell cycles/log(2)).

δ_P: The only exception to the above, we assume the product is actively used. Arbitrarily set to 10.

F_{ext}: The external product flux. At $t < 0$ this is positive, set arbitrarily to 10 δ_P, at $t > 0$ it is set to 0.

Parameters Related to the Cost Function.

γ: The growth penalty associated with producing one extra unit of enzyme.

η: The amount of time spent at prestarvation steady state.

T: The amount of time spent poststarvation. This includes both the transient regime and the steady state reached. Typically steady state is reached at approximately $t = 2$ so the major component of C2 is the new steady-state enzyme level. Equal to 10 for all optimizations shown in the main text; we show in Fig. S2 that the results are not sensitive to the exact value.

P_{goal}: The concentration of product necessary for 0 growth penalty. Equal to 1 for all optimizations shown in the main text; we show in Fig. S2 that the results are not sensitive to the exact value. The multiplicative weight on C3 is implicitly set to 1 because changing it would only scale the total sum of the cost components.

Parameters Determined by Optimization.

c_1–c_N: The basal rate of expression of the enzymes.

b_1–b_N: The maximal rate of expression induced by the transcription factor.

a_1–a_N: The fraction of active transcription factor necessary for half-maximal activation of expression.

k_f: The concentration of S_{act} necessary for half-maximal transcription factor activity.

k_{inh}: The concentration of P necessary for half-maximal inhibition of the first pathway step. Our analysis of in vitro data of end product inhibition of Leu4, Lys20/21, Ade4, and Arg2 shows that the activity curves tend to be best fit by Hill coefficients in the range of 1.5–2.5. Thus, we use a Hill coefficient of 2 for this inhibition.

Other Notes. We use a linear model for the growth cost as a function of enzyme production. While it has been shown that at high levels of expression, the growth cost in *Escherichia coli* becomes nonlinear with LacZ protein expression level (1), our calculations show that a typical metabolic protein in yeast is expressed at a level that is several orders of magnitude lower, relative to total cellular protein. At this level, the linear approximation is quite reasonable.

1. Dekel E, Alon U (2005) Optimality and evolutionary tuning of the expression level of a protein. *Nature* 436:588–592.

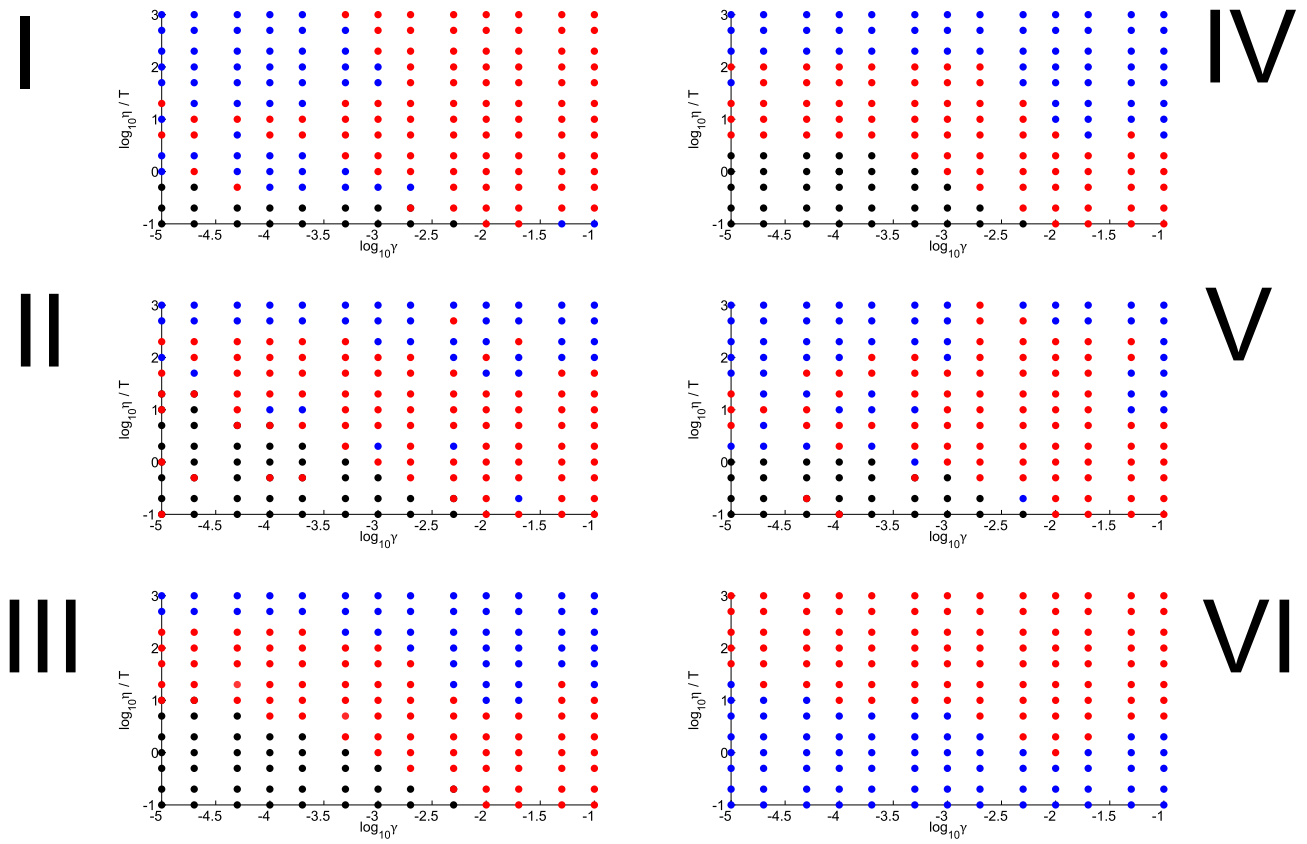

Fig. S1. Phase diagrams for all networks. Networks I–VI correspond to the same networks in Fig. 7 in the main text. For networks I–V, red points correspond to solutions with separation of regulation of downstream and upstream enzymes. For network VI, red points correspond to solutions with almost identical induction levels for all enzymes. Black points correspond to solutions with no appreciable induction of any enzymes. Blue points are solutions with no clear pattern.

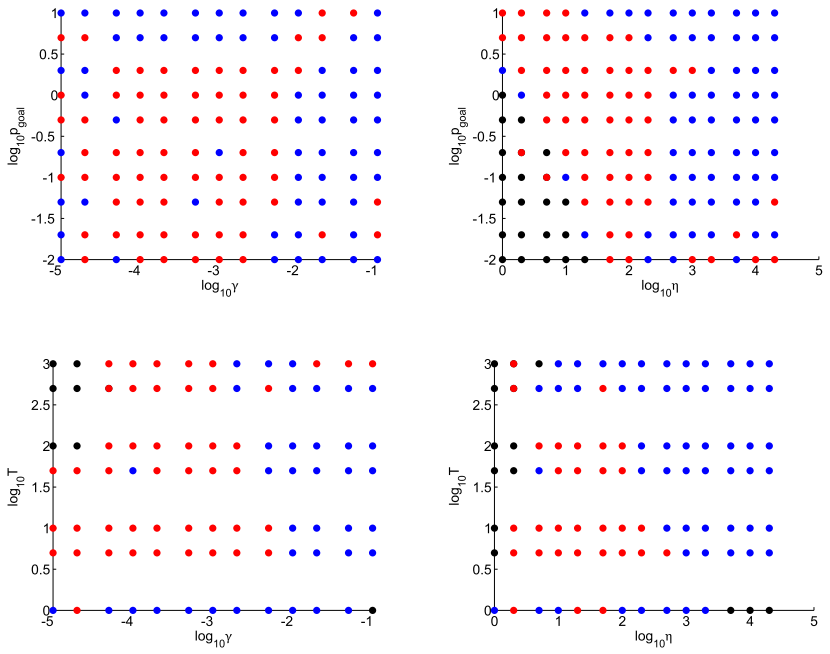

Fig. S2. Phase diagrams for other evolutionary paramers for network III. Red points correspond to solutions with separation of regulation of downstream and upstream enzymes. Black points correspond to solutions with no appreciable induction of any enzymes. Blue points are solutions with no clear pattern. Within a reasonable range, neither P_{goal} nor T has a qualitative affect on the model results. The default values for the parameters were $\gamma = 0.005$, $\eta = 200$, $T = 10$, $P_{goal} = 1$.

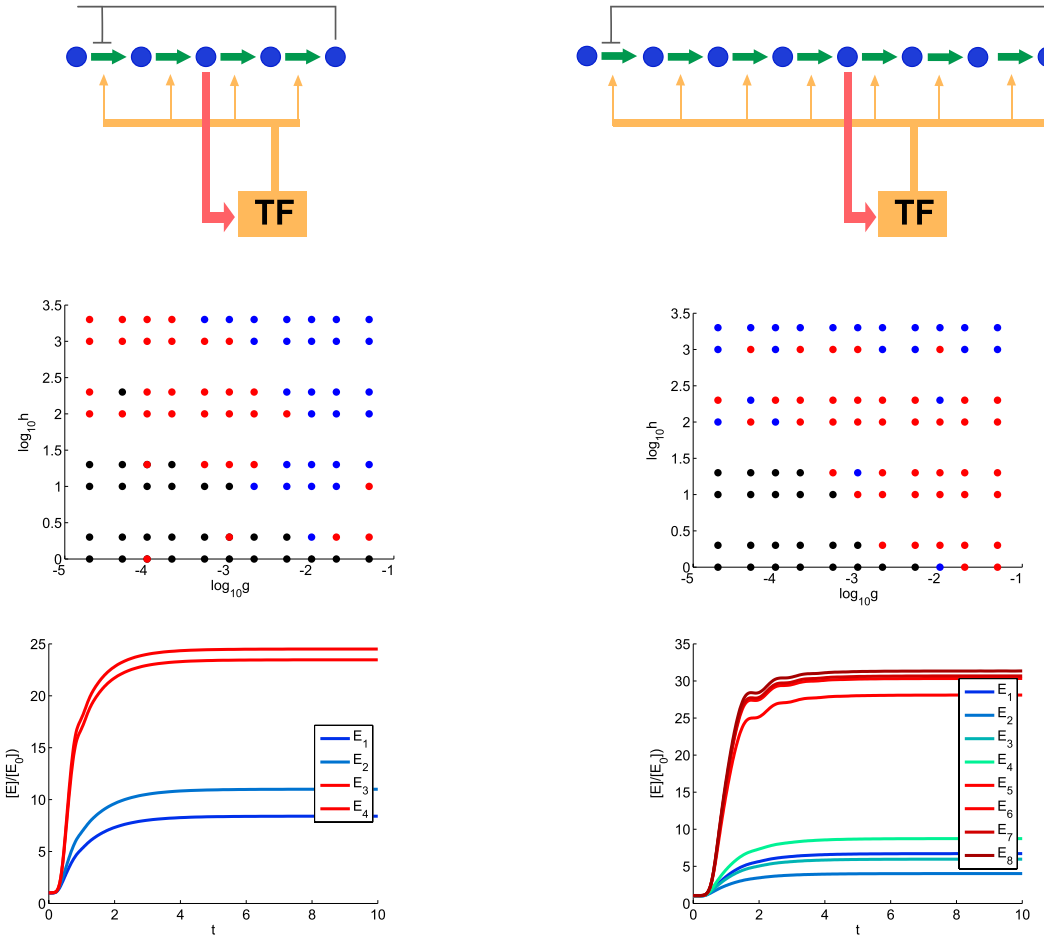

Fig. S3. Phase diagrams and representative curves for the model with four enzymes and eight enzymes. To ensure that the phenomenon we were seeing was not particular to the number of enzymes in the pathway, we also analyzed pathways of four enzymes (*Left*) and eight enzymes (*Right*). Shown are the γ-η phase diagram as well as an example of a case where we clearly see differential regulation. In the phase diagram, red points correspond to solutions with separation of regulation of downstream and upstream enzymes, black points correspond to solutions with no appreciable induction of any enzymes, and blue points are solutions with no clear pattern.

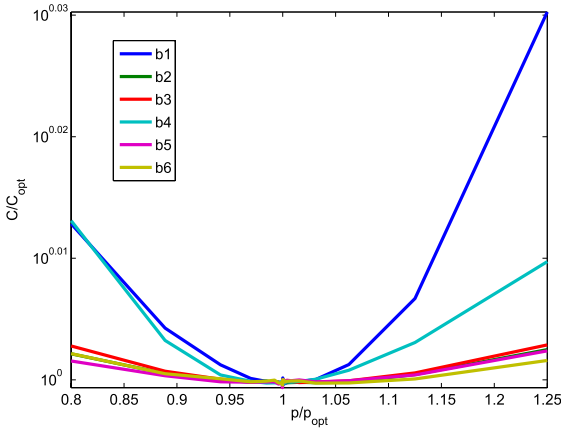

Fig. S4. Sensitivity of the cost function to small changes in parameter values away from the found optimum. Parameters b_1–b_6 correspond to promoter dependence on transcription factor activity for enzymes 1–6 in the pathway. Enzyme 4 is the one immediately downstream of the regulatory intermediate. The cost function is most sensitive to b_1 because enzyme 1 controls basal pathway flux, but also highly sensitive to b_4 because enzyme 4 exerts the most control over the level of regulatory intermediate.

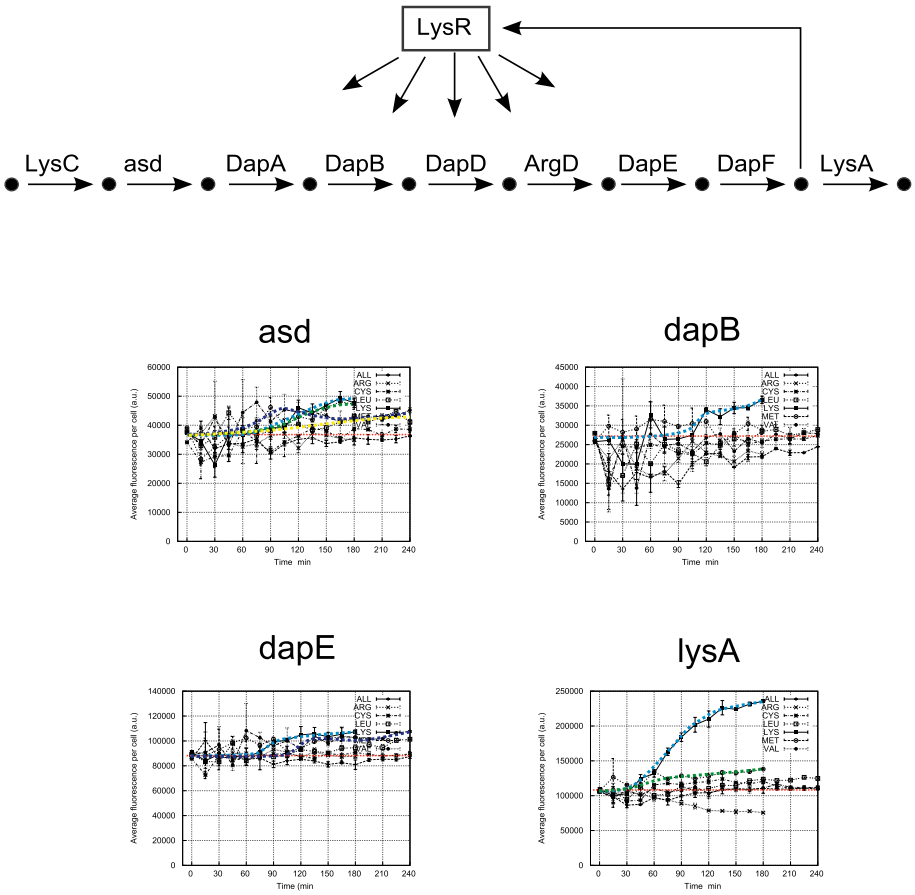

Fig. S5. Profiles of gene expression in *E. coli* in response to lysine starvation, reproduced from ref. 1. Light blue curve corresponds to expression in –Lys media, whereas other curves correspond to starvation for other amino acids. Expression could be measured for four genes in the pathway. LysA, the enzyme downstream of the regulatory intermediate, has a higher level of induction than other pathway enzymes. Fluorescence of wild type (background) is about 30,000 units.

1. Yamada T, et al. (2010) Relationship between noise characteristics in protein expressions and regulatory structures of amino acid biosynthesis pathways. *IET Syst Biol* 4:82–89.

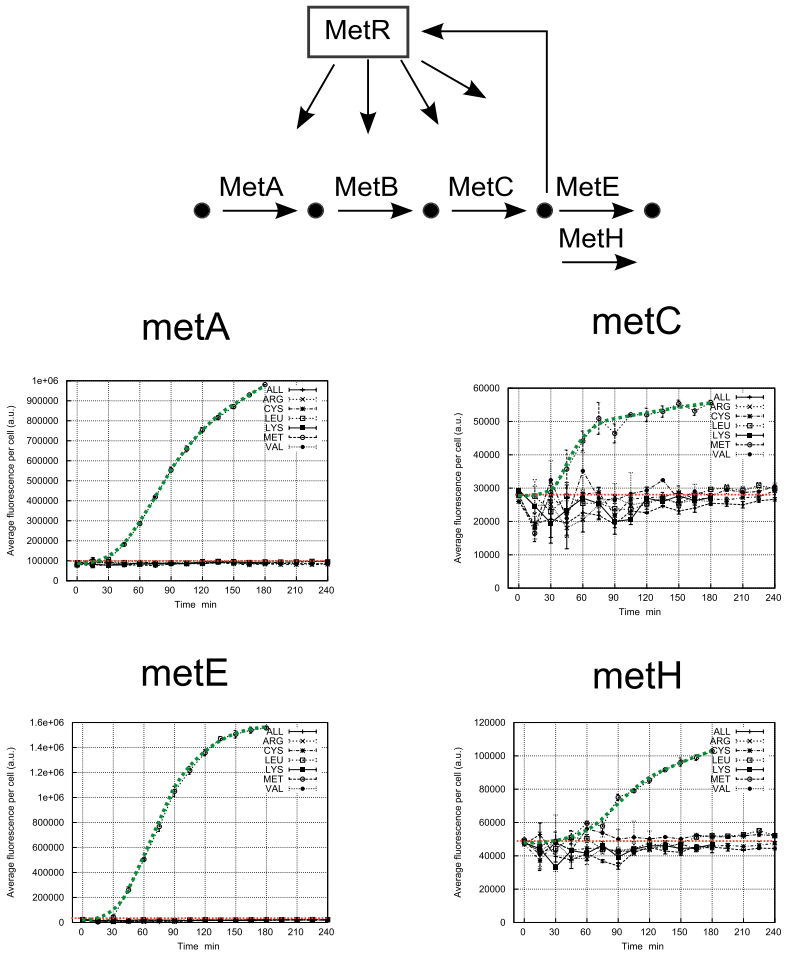

Fig. S6. Profiles of gene expression in *E. coli* in response to methionine starvation, reproduced from ref. 1. Green curve corresponds to expression in −Met media. Expression could be measured for four genes in the pathway. Whereas metA and metE are both expressed at high level, MetE, the enzyme downstream of the regulatory intermediate, shows the highest fold change in response to methionine depletion. Of the two isozymes metE and metH, metE was previously shown to be the one chiefly induced by methionine depletion (2, 3). Fluorescence of wild type (background) is about 30,000 units.

1. Yamada T, et al. (2010) Relationship between noise characteristics in protein expressions and regulatory structures of amino acid biosynthesis pathways. *IET Syst Biol* 4:82–89.
2. Urbanowski ML, Stauffer GV (1989) Role of homocysteine in metR-mediated activation of the metE and metH genes in Salmonella typhimurium and Escherichia coli. *J Bacteriol* 171: 3277–3281.
3. Cai XY, et al. (1989) Methionine synthesis in *Escherichia coli*: effect of the MetR protein on metE and metH expression. *Proc Natl Acad Sci USA* 86:4407–4411.

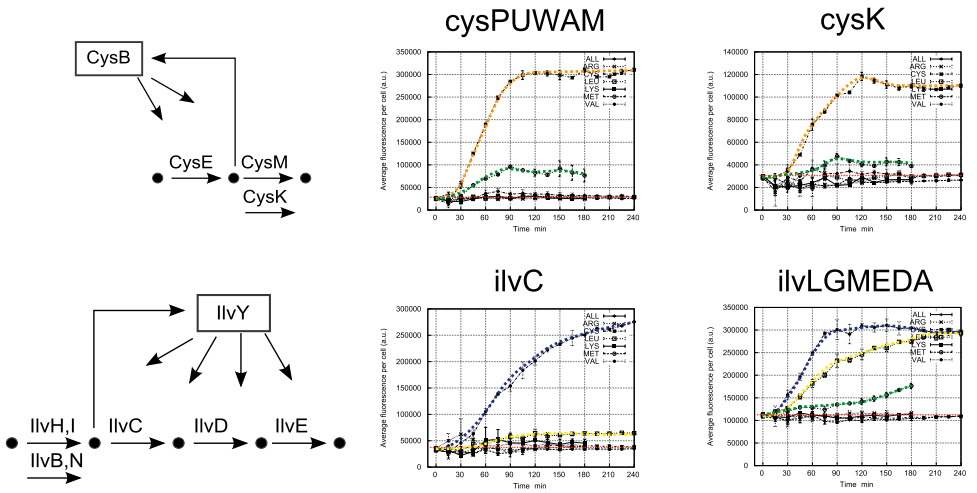

Fig. S7. Profiles of gene expression in *E. coli* in response to cysteine and valine starvation, reproduced from ref. 1. Orange curve corresponds to expression in −Cys media, blue curve corresponds to expression in −Val media. Expression could only be measured for two genes in each pathway because other enzymes had very low levels of expression. Whereas it appears that relative to typical fold changes in other pathways, the enzymes downstream of the regulatory intermediates have high induction, this is only weakly suggestive without more complete data. Fluorescence of wild type (background) is about 30,000 units.

1. Yamada T, et al. (2010) Relationship between noise characteristics in protein expressions and regulatory structures of amino acid biosynthesis pathways. *IET Syst Biol* 4:82–89.

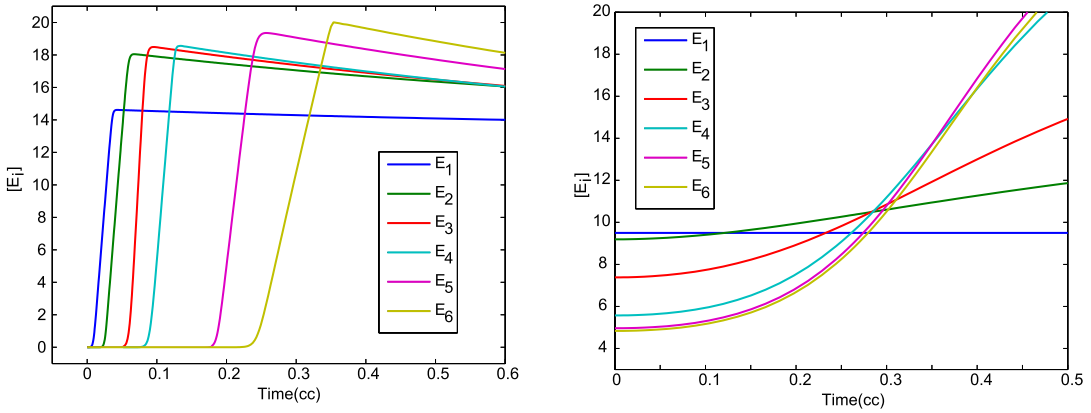

Fig. S8. Examples of just-in-time behavior produced by our model. The *Left* graph shows the profiles when the cost function is minimized with the similar metaparameters to the rest of our analysis, but no regulatory network exists, and enzyme profiles are allowed to be a sigmoid function of three parameters. The *Right* graph shows optimal profiles under network VI (end product inhibition) but with $T = 0.5$, so steady-state levels are largely irrelevant.