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

Results

2.1 Model Derivation

The goal of this work was first to construct a modeling framework to describe cell-free protein synthesis systems and to examine its performance in productivity and yield for a protein of interest. One mathematical framework that has found wide use in modeling metabolism are constraint based models such as flux balance analysis (FBA). FBA can predict how cells utilize nutrients to produce products by using the biochemical stoichiometry and thermodynamical feasibility under pseudo steady-state conditions. Traditionally, FBA is used to model *in vivo* processes, however cell-free systems do not have growth associated reactions or transport through the cell membrane. Thus, to model cell-free metabolism we constructed a cell-free stoichiometric network by removing growth associated reactions from the MG1655 reconstruction (*8*), and incorporating amino acid synthesis and transcription/translation associated reactions (*16*) for a protein of interest to be expressed. The network consisted of 281

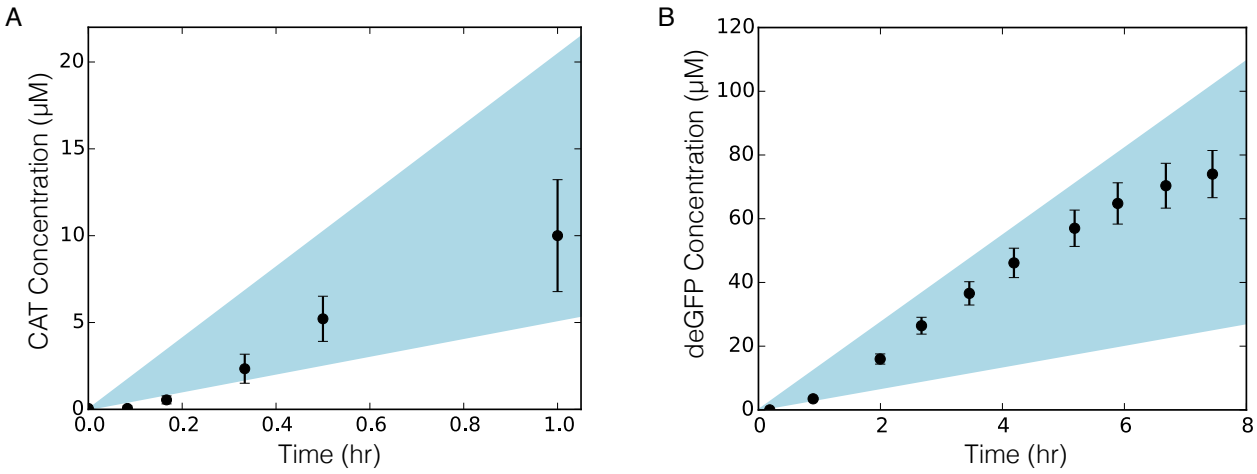

Figure 1: Sequence specific flux balance analysis of protein production vesus time. A. CAT production under a T7 promoter in CFPS *E. coli* extract for 1 hr under glucose consumption. B. deGFP production under a P70 promoter in TXTL 2.0 *E. coli* extract for 8 hr under glucose consumption. 95% CI (blue region) over the ensemble of 100 sets.

reactions and 132 species. We developed a cell-free sequence specific flux balance analysis (ssFBA) with a detailed promoter model (*17*) to examine the performance of CFPS. We first validated the ssFBA approach by comparing simulated and measured concentrations of two proteins from two different cell-free *E. coli* extract systems. The first protein, chloramphenicol acetyltransferase (CAT), was produced under a T7 promoter in a glucose/NMP cell-free system (*18*) for a duration of 1 hour under glucose consumption (Fig 1A). The second protein, dual emission green fluorescent protein (deGFP), was produced under a P70 promoter in TXTL 2.0 *E. coli* extract for a duration of 8 hours under maltose consumption(Fig 1B). The ssFBA simulations predicted the production of both proteins for the duration of both CFPS batch reactions. Uncertainty in experimental factors such as RNA polymerase, ribosome concentrations, elongation rates or the upper bounds for oxygen and glucose consumption rates did not alter the qualitative performance of the model. Thus, the metabolic network and molecular description of trancription and translation were consistent with experimental measurements.

Next, ssFBA predicted deGFP production as a function of plasmid concentration (Fig 2). Concentration of deGFP at each plasmid concentration was calculated by multiplying the flux of deGFP synthesis by the active time of production, approximately 8 hours in TXTL

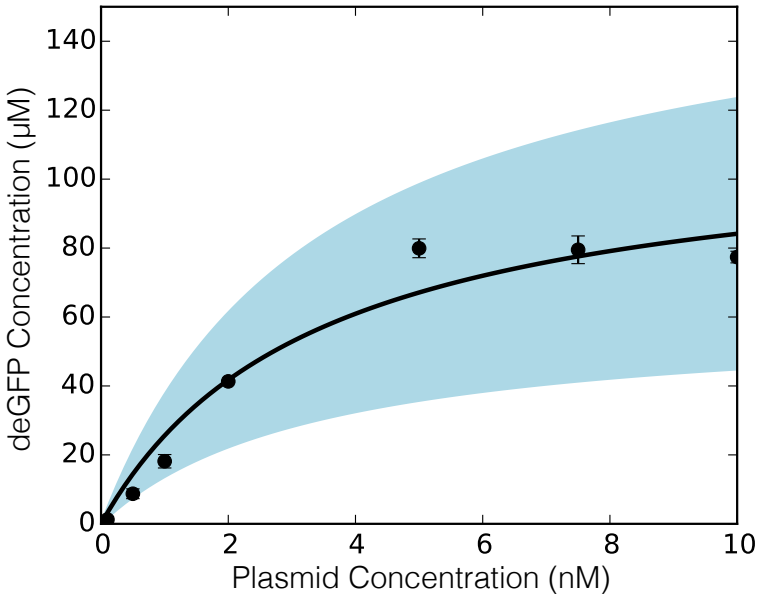

Figure 2: Predicted deGFP concentration (black line) at different plasmid concentrations versus measurements of deGFP (dots) synthesized in TXTL 2.0. 95% CI (blue region) over the ensemble of 100 sets.

2.0 (**?**). The mean of the ensemble shows a good prediction against the measured deGFP levels, even though it under predicted deGFP concentration at the saturating point of 5 nM of plasmid concentration. However, the ensemble and the mean of the ensemble captured the overall saturating dynamics of deGFP production as a function of plasmid concentration.

These results validated our mathematical framework to model CFPS systems and predict the production of two proteins with no adjustable parameters. It also showed that the sequence specific reactions were sufficient to predict the production of two different proteins under different promoters and cell-free systems. Since the model accurately predicted the behavior of protein production, we wanted to use our mathematical framework to help us understand the performance limits of CFPS and how these shortcomings could be addressed.

2.2 Examining CFPS Performance: Productivity

Our next goal was to examine the performance of CFPS productivity for eight different proteins under three different cases. Each of the proteins were produced under a P70 promoter, except for CAT which was produced under a T7 promoter. In all our cases, CFPS

is supplied with glucose. The first case, CFPS is supplied with amino acids and the system can synthesis amino acids (control). In the second case, CFPS is supplied with amino acids, however the system cannot synthesis amino acids (AA uptake w/o synthesis). We turned off these synthesis reactions because during the cell-free extract preparation the cells are often supplied with amino acids, thus the enzymes responsible for amino acid synthesis would not be present. In the third case, CFPS is not supplied with amino acids, but the system can synthesis them (AA synthesis w/o uptakte). We used ssFBA to estimate the productivity of eight proteins for each case (Fig 3A). The second case (without amino acid synthesis) showed the highest productivity for each of the proteins, however the control case had very similar performance. This shows the system has sufficient substrates to power the system and synthesis each protein. The third case had the lowest productivity for each protein, thus the addition of amino acids to CFPS extract is important for maintaining a relatively high productivity. The qualitative trend of productivity for the three different cases was the same, however some proteins had higher productivity than others. For instance, in the second case FGF21 had a productivity of 17 (μM/h) whereas FII had a productivity of 3 (μM/h). To examine this further, the mean productivity was plotted against the carbon number of each protein (Fig 3B). The proteins with the highest productivity had the lowest carbon number whereas proteins with low productivity had higher carbon numbers. This inverse qualitative trend was due to the fact that larger proteins require more amino acids and substrates to assemble them resulting in lower productivity.

2.3 Examining CFPS Performance: Carbon Yield

Following the same outline as in examining the productivity, we calculated the carbon yield for each protein (Fig 4A). The same trends followed, where the case without amino acid synthesis showed the highest carbon yield and the control case had comparable performance. The third case (with no amino acid uptake) had the lowest yields; this is most likely because glucose is now utilized to synthesize the necessary amino acids for each protein as well

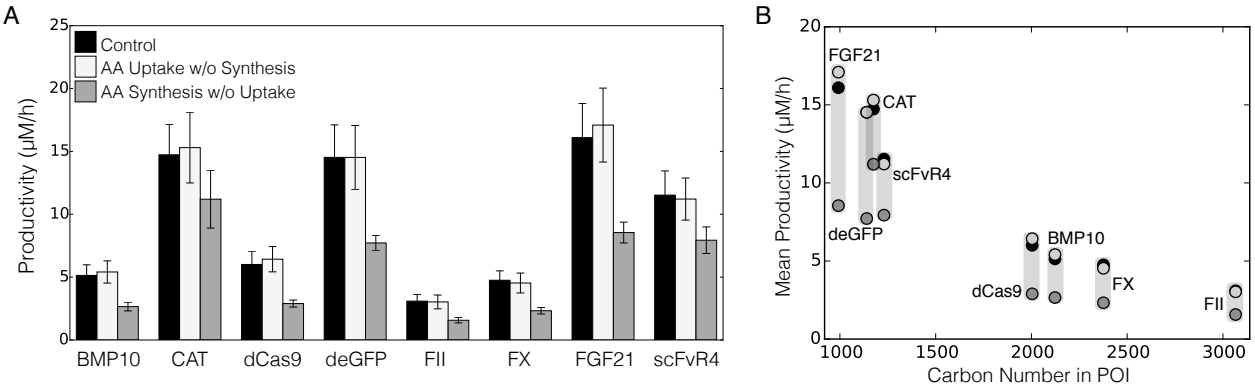

Figure 3: CFPS productivity performace for eight proteins for the control (black), AA uptake w/o synthesis (off white), and AA synthesis w/o uptake (grey). A. Productivity normalized to the specific glucose uptake rate (Error bars represent the 95% CI of the ensemble). B. Mean productivity versus the carbon number for each corresponding protein.

as power the system. Interestingly, CAT carbon yield showed similar performance for all three cases. Each protein has different energy requirements for transcription and translation depending on its sequence. Thus, for the case of CAT, energy requirements for transcription or translation are low enough that it's carbon yield is not hindered in the third case. We next investigated the effect of the carbon number of each protein to the carbon yield (Fig 4B). The same inverse qualitative trend is observed, however it is less significant. Sequence specific flux balance analysis assumes a psuedo steady state, thus intermediate metabolites cannot be accumulated within the cell-free extract. In addition, ssFBA is solved by setting the production of the protein as the objective function. Therefore, carbon flux will travel throughout the network to optimize the flux through the protein synthesis reaction. This leads to slightly similar carbon yields for all the proteins, however, the lower carbon number protein still observe a higher carbon yield.

2.4 Sensitivity analysis on CFPS Performance

To better understand the effect of substrate utilization and transcription/translation parameters on CFPS performance we performed global sensitivity analysis on productivity and carbon yield for deGFP, a representative protein (Fig 5). In examining productivity performance (Fig 5A), the significance of transcription/translation parameters were fairly constant

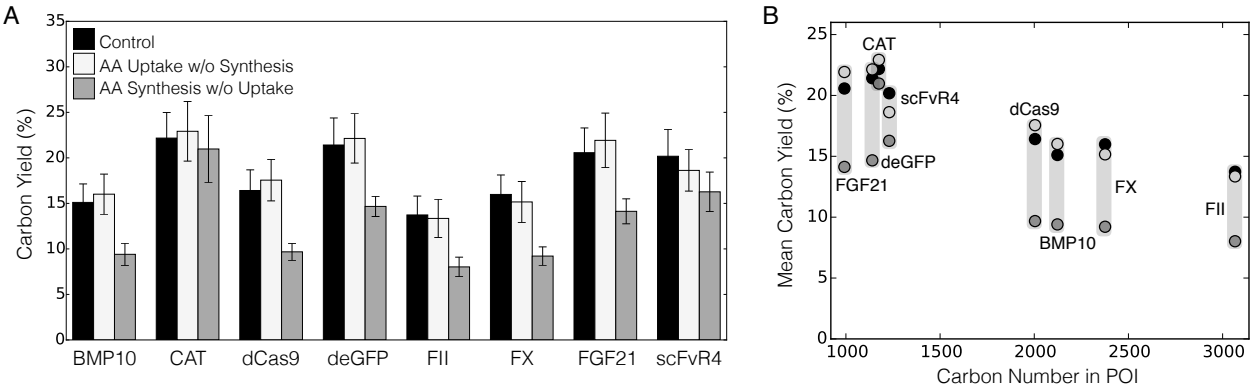

Figure 4: CFPS carbon yield performace for eight proteins for the control (black), AA uptake w/o synthesis (off white), and AA synthesis w/o uptake (grey). A. Carbon yield (Error bars represent the 95% CI of the ensemble). B. Mean carbon yield versus the carbon number for each corresponding protein.

across all three cases, with the elongation rate by ribosomes being the most sensitive. This shows the rate of translation is instrumental in achieving high productivity, and should be the first step to investigate in order to optimize productivity, prior to examining transcription rates, this is consistent with literature findings. Underwood and coworkers have also shown that an increase in ribosome levels does not significantly increase protein yields or rates; however, adding elongation factors increased yields by 23% at 30 minutes (*19*). In addition, Li et al. have increased productivity of firefly lucifease by 5-fold in CFPS systems by adding and adjusting factors that affect transcription and translation such as elongation factors, ribosome recyclcing factor, release factors, chaperones, BSA, and tRNAs (*20*). In examining the substrate utilization, glucose uptake was not very important for productivity in the first two cases, but its significance increased when amino acids was removed from CFPS. This makes sense, as amino acid synthesis can only come from glucose and becomes the only way to power protein synthesis. On the other hand, when amino acid synthesis is removed, amino acid uptake becomes more important, as it becomes the only source of amino acids.

When considering carbon yield performance (Fig 5B), the substrate and oxygen uptake fluxes become much more important while the sensitivity to kinetic parameters decreases slightly. This is likely because glucose and oxygen uptake determine the mechanism of en-

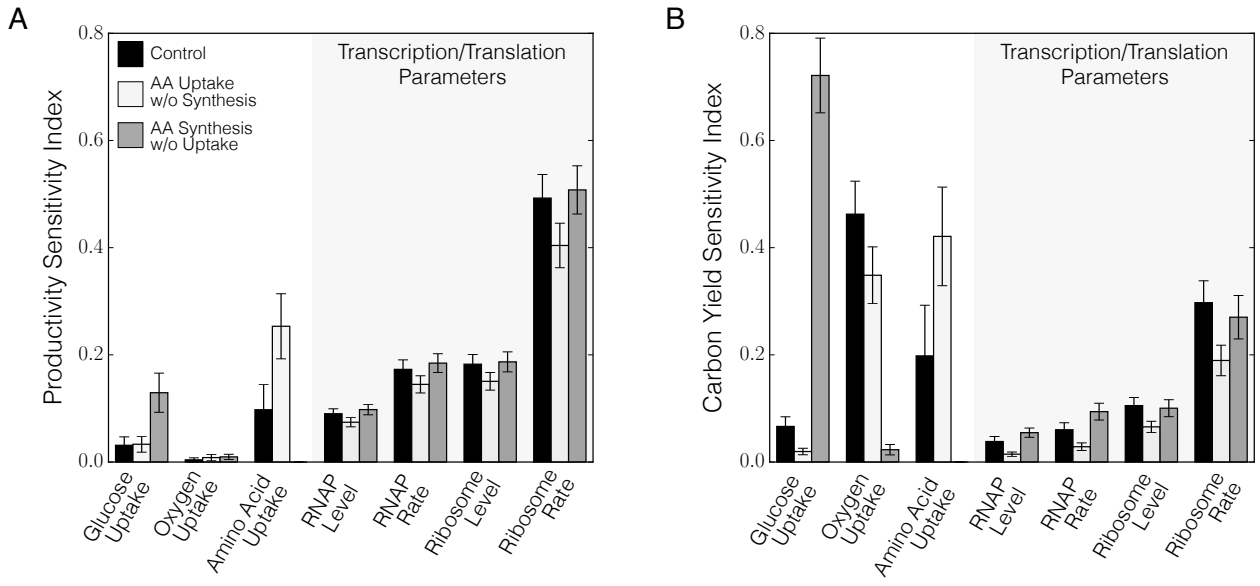

Figure 5: Total order sensitivity of deGFP productivity (A) and carbon yield (B) to specific uptake rates and transcription/translation parameters for three cases: control (black), amino acid uptake without synthesis (off white), and amino acid synthesis without uptake (gray). (Error bars represent 95% CI of the ensemble).

ergy generation, which is critical to efficient protein production. Meanwhile, productivity is determined primarily by the rate of the most downstream steps, transcription and translation. Interestingly, in the first two cases, oxygen uptake has a significant effect on carbon yield. This is most likely due to the importance of oxidative phosphorylation for efficient energy generation. However, in the third case (with no amino acid uptake), oxygen uptake becomes insignificant compared to the other parameters, while glucose uptake becomes the most siginificant. This may be because glucose is required for both amino acid synthesis and efficient energy generation, both of which are important for a high yield. Thus, there is a tradeoff between energy generation to power transcription/translation as well as synthesize amino acids which are required to assemble the protein of interest. Across all three cases, substrate utilization (amino acid uptake and glucose) showed to be the next important as these substrates are contributing to the carbon yield of deGFP. The transcription/translation parameters had the same trend as in the productivity trend, where the ribosome elongation rate was the most sensitive compared to the transcription/translation parameters and showed significant importance across all parameters. Thus, in investigating CFPS performance of

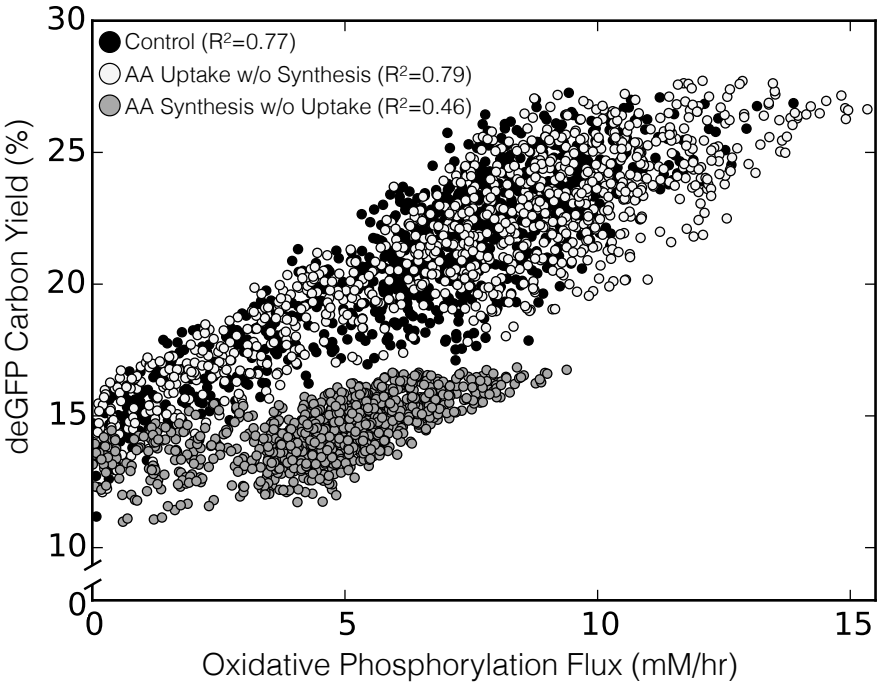

Figure 6: An ensemble of 1000 ssFBA solutions for deGFP carbon yield versus oxidative phosphorylation flux for three cases: control (black), amino acid uptake without synthesis (off white), and amino acid synthesis without uptake (gray).

productivity and yield, the ribosome elongation rate seems to be an important parameter to optimize, as been already shown in literature (*19*, *20*).

We also performed this sensitivity analysis on CAT production, which did not show the same trend in yield across the three cases as the other proteins. For CAT, glucose uptake was as important to productivity as the kinetic parameters. Also, the increase in amino acid uptake importance to both productivity and yield was much greater. Finally, oxygen uptake became the most important factor in yield and maintained this across all three cases. Apart from these differences, the general trend of transcription and translation mattering more for productivity and uptake mattering more for yield was still true.

Table 1: Carbon contribution from glucose and each amino acid toward deGFP production for three cases: control, amino acid uptake without synthesis, and amino acid synthesis without uptake.

Carbon Produced (mM)	Control	AA uptake w/o synthesis	AA synthesis w/o uptake
deGFP	9.8	9.6	9.9
Carbon Consumed (mM)			
GLC	37.3	33.7	66.9
ALA	0.0	0.2	-
ARG	0.3	0.3	-
ASN	0.5	0.4	-
ASP	0.4	0.6	-
CYS	0.2	0.1	-
GLU	2.1	0.6	-
GLN	0.4	0.3	-
GLY	0.3	0.3	-
HIS	0.5	0.5	-
ILE	0.0	0.6	-
LEU	1.0	1.0	-
LYS	1.0	0.9	-
MET	0.0	0.2	-
PHE	1.0	0.9	-
PRO	0.4	0.4	-
SER	0.0	0.2	-
THR	1.0	0.5	-
TRP	0.1	0.1	-
TYR	0.8	0.8	-
VAL	0.6	0.7	-
Sum	47.9	43.3	66.9
Yield	20.5%	22.2%	14.8%
Yield w/o GLC	92.5%	100%	-

Materials and Methods

Formulation and solution of the model equations.

The flux balance analysis problem was formulated as:

$$\max_{\boldsymbol{w}} \left(w_{obj} = \boldsymbol{\theta}^T \boldsymbol{w}\right)$$
$$\text{Subject to : } \mathbf{Sw} = \mathbf{0}$$
$$\alpha_i \leq w_i \leq \beta_i \qquad i = 1, 2, \ldots, \mathcal{R}$$

where $\mathbf{S}$ denotes the stoichiometric matrix, $\mathbf{w}$ denotes the unknown flux vector, $\boldsymbol{\theta}$ denotes the objective selection vector and α_i and β_i denote the lower and upper bounds on flux w_i, respectively. The flux balance analysis problem was solved using the GNU Linear Programming Kit (v4.52) (*21*). In this study, the objective w_{obj} was to maximize the production of circuit output. The specific glucose uptake rate was constrained to allow a maximum flux of 10 mmol/hr (**?**); the amino acids were also bound to allow a maximum flux of 10 mmol/hr, but did not reach this maximum flux.

Transcription and translation template reactions.

The transcription and translation template reactions are based off sequence specific FBA (**?**) involving transcription initiaion, transcription, mRNA degradation, translation initiation, translation, and tRNA charging. The mRNA and protein sequence of each protein was determined from literature. The transcription rate was constrained by the following formulation:

$$w_{tx} = [RNAP]\frac{v_{RNAP}}{l_{mRNA}}\left(\frac{[Gene]}{km + [Gene]}\right)P$$

where $[RNAP]$ is the concentration of RNA polymerase which was determined from literature values based on the number of copies per cell, v_{RNAP} is the elongation rate (nu-

cleotides/hr) of the RNA polymerase, l_{mRNA} is the number of nucleotides in the mRNA for the protein of interest, $[Gene]$ is the gene concentration of the protein of interest, km is the plasmid saturation coefficient, and P is the promoter activity. The promoter activity was formulated following Moon et al. for synthetic circuits by the following:

$$P = \frac{K_1 + K_2 f_{p70}}{1 + K_1 + K_2 f_{p70}}$$

where K_1 represents the state of RNA polymerase binding, K_2 is the state of sigma-70 binding along with RNA polymerase, and f_{p70} is the fraction of the transcription factor, sigma-70, bound to the promoter following Hills kinetics.

The translation rate was constrained by the following formulation:

$$w_{tl} = [Ribo] K_P \frac{v_{Ribo}}{l_{protein}} [mRNA_{ss}]$$

where $[Ribo]$ is the ribosome concentration determined from literature values based on the number of copies per cell, K_P is the polysome amplification constant, v_{Ribo} is the elongation rate (amino acids/hr) of the ribosome, $l_{protein}$ is the number of amino acids in the protein of interest, and $[mRNA_{ss}]$ is the mRNA concentration at steady state determined by the transcription rate divided by the degradation rate of mRNA.

Theoritical carbon yield.

The theortical carbon yield of each protein was formulated as:

$$Yield = \frac{C_{POI} v_{POI}}{\sum_{i=1}^{\mathcal{R}} C_i v_i}$$

where C_{POI} and C_i denote the carbon number of the protein of interest (POI) and substrate i, respectively, v_{POI} and v_i denote the flux of the POI and substrate i, respectively, and $\mathcal{R}$ denotes the number of substrates consumed.

Global sensitivity analysis.

We conducted a global sensitivity analysis, using the variance-based method of Sobol, to estimate which parameters controlled the performance of synthetic circuits (*22*). We computed the total sensitivity index of each parameter relative to two performance objectives, the peak thrombin time and the area under the thrombin curve (thrombin exposure). We established the sampling bounds for each parameter from the minimum and maximum value of that parameter in the parameter set ensemble. We used the sampling method of Saltelli *et al.* (*23*) to compute a family of $N\,(2d+2)$ parameter sets which obeyed our parameter ranges, where N was the number of trials, and d was the number of parameters in the model. In our case, N = 10,000 and d = 22, so the total sensitivity indices were computed from 460,000 model evaluations. The variance-based sensitivity analysis was conducted using the SALib module encoded in the Python programming language (*24*).

2.5 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 (*25*). 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 .

2.6 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

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.

2.7 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 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. Allen, T. E., and Palsson, B. Ø. (2003) Sequence-based analysis of metabolic demands for protein synthesis in prokaryotes. *J Theor Biol 220*, 1–18.

17. Moon TS, T. A. S. B. V. C., Lou C (2012) Genetic programs constructed from layered logic gates in single cells. *Nature 491*.

18. Calhoun, K. A., and Swartz, J. R. (2005) An Economical Method for Cell-Free Protein Synthesis using Glucose and Nucleoside Monophosphates. *Biotechnology Progress 21*, 1146–53.

19. Underwood, K. A., Swartz, J. R., and Puglisi, J. D. (2005) Quantitative polysome analysis identifies limitations in bacterial cell-free protein synthesis. *Biotechnology and Bioengineering 91*, 425–35.

20. Li, J., Gu, L., Aach, J., and Church, G. M. (2014) Improved Cell-Free RNA and Protein Synthesis System. *PLoS ONE 9*, 1–11.

21. GNU Linear Programming Kit, Version 4.52. 2016; `http://www.gnu.org/software/glpk/glpk.html`.

22. Sobol, I. (2001) Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation 55*, 271–80.

23. Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S. (2010) Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications 181*, 259–70.

24. Herman, J. D. http://jdherman.github.io/SALib/.

25. 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.

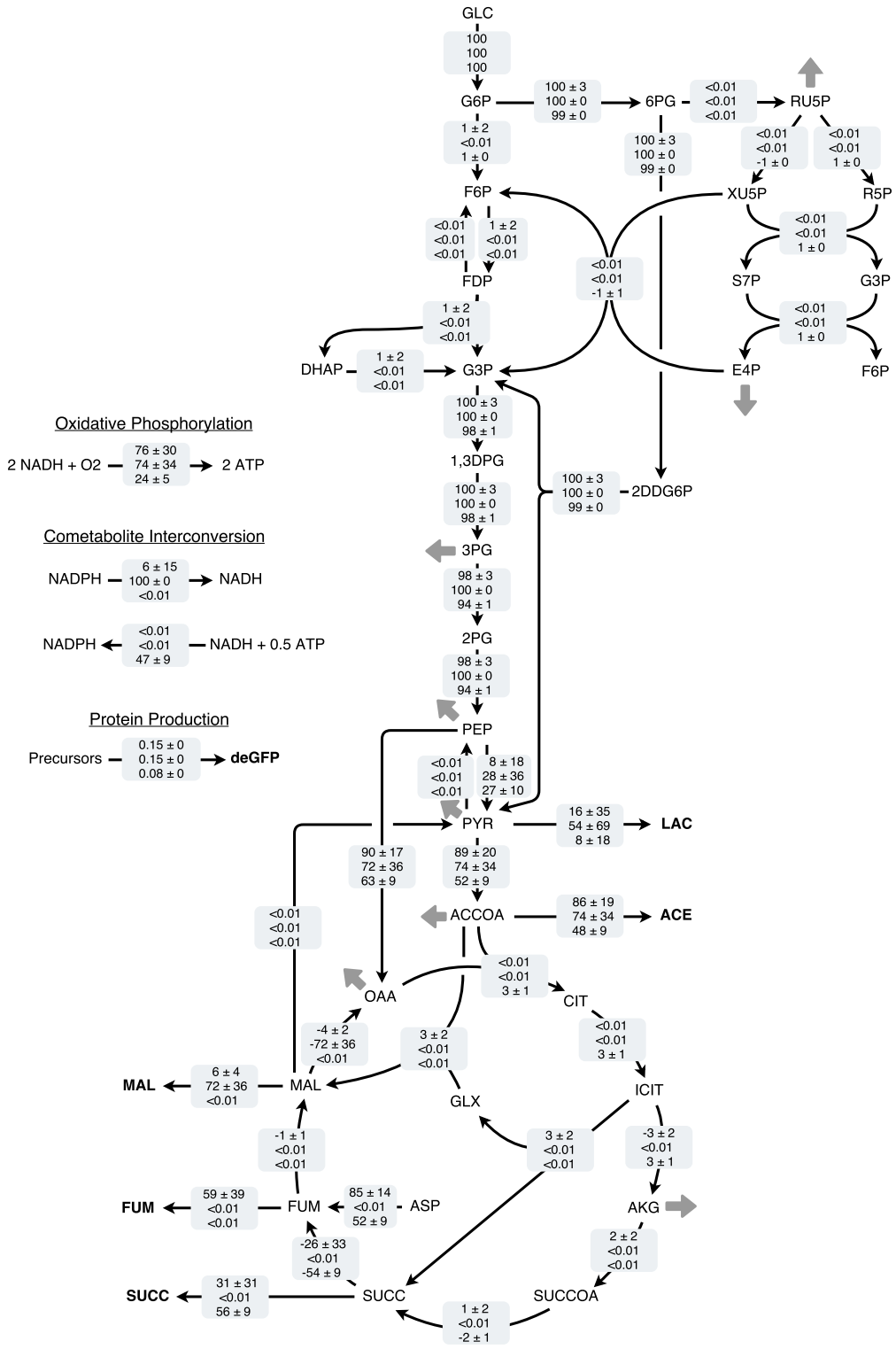

Figure 7: deGFP production flux profile for glycolysis, pentose phosphate pathway, Entner-Doudoroff pathway, TCA cycle, NADPH/NADH transfer, and oxidative phosphorylation. Flux (mean ± standard deviation) across ensemble normalized to glucose uptake flux. Flux distribution for three different cases: control (top row), amino acid uptake without synthesis (second row), and amino acid synthesis without uptake (third row). Bold metabolites are allowed to accumulate; grey arrows lead to amino acid synthesis.

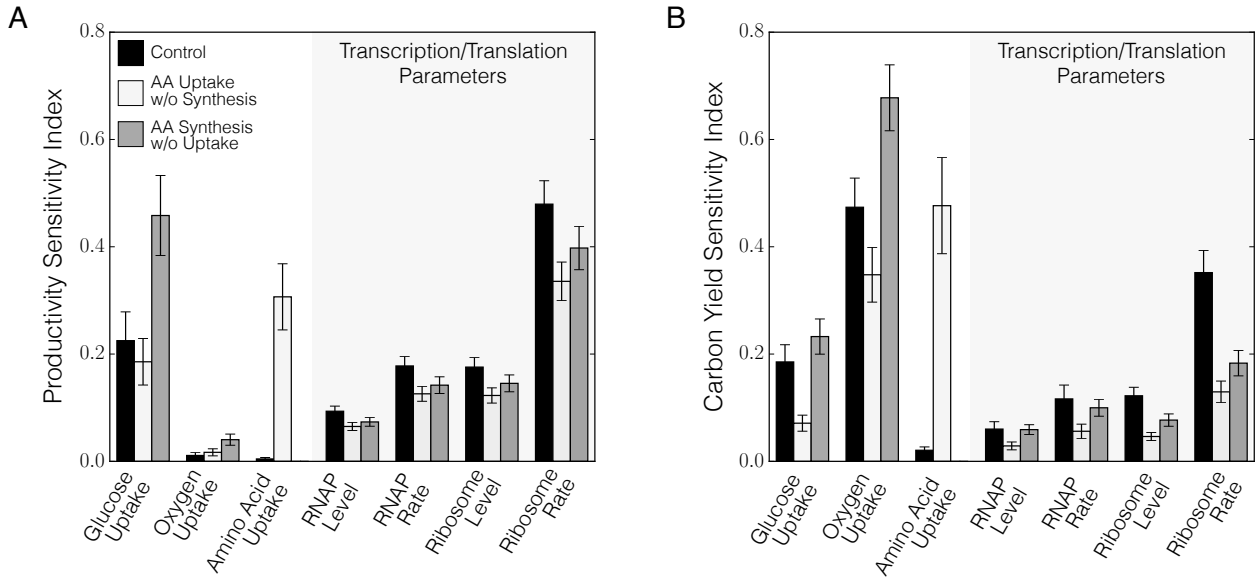

Figure 8: Total order sensitivity of CAT productivity (A) and carbon yield (B) to specific uptake rates and transcription/translation parameters for three cases: control (black), amino acid uptake without synthesis (off white), and amino acid synthesis without uptake (gray). (Error bars represent 95% CI of the ensemble).

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

Figure 9: An example figure