Prof. Jörg Stelling
D-BSSE
ETH Center UNG H 14
Universitätstrasse 19
CH-8092 Zürich
https://csb.ethz.ch

Computational Systems Biology
636-0007-00 U, Autumn 2025

Assignment 3

(Issue: 10-Oct-2025)

Introduction

A biotech startup company provides the simple network given in figure 1 describing some metabolic pathway of interest. Two substrates (A and B) are converted into the pharmaceutically relevant precursor P and a byproduct E. Your task is to analyze the conversion process in more detail using stoichiometric network analysis methods.

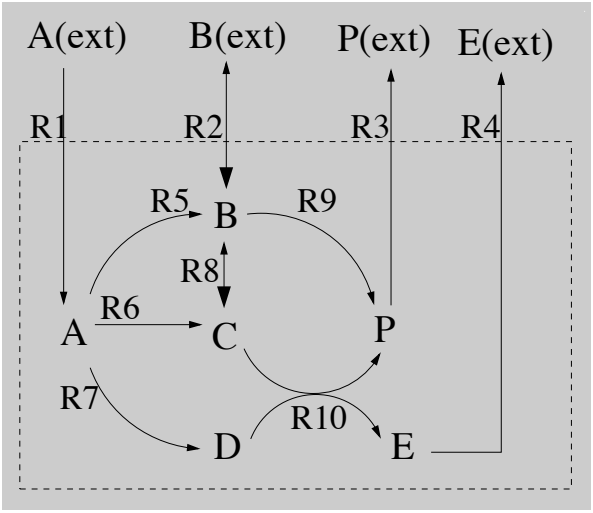

Figure 1: Metabolic network, *Klamt & Stelling, Stoich. and Constraint-based Modeling, MIT Press, Apr. 2006*

1 Stoichiometric Matrix and Nullspace Analysis

a) Derive the stoichiometric matrix $\mathbf{N}^{m\times n}$ from figure 1, where $m = 6$ denotes the number of metabolites (A, B, C, D, E, P), and $n = 10$ the number of reactions.

b) Determine the nullspace (kernel matrix) $\mathbf{K}$ of $\mathbf{N}$ using MATLAB (check null command and argument 'r') or python (rational basis for example from sympy.rref and sympy.nullspace. sicpy.linalg.nullspace is also an option). What is the difference between the orthonormal and the rational basis, and what advantages do they have?

c) Are the computed basis vectors feasible flux distributions? Explain why / why not.

2 Flux Balance Analysis (FBA)

Since substrate A is cheaper and easier to come by on the marked than B, we are especially interested in the maximal yield of P on A. Our job is thus to

$$\begin{array}{lrcll} \text{maximize} & Y_{P/A} & = & r_3/r_1 & \\ \\ \text{subject to} & \mathbf{Nr} & = & \mathbf{0} & \textit{quasy steady state constraint} \\ & \mathbf{r}_{irrev} & \geq & \mathbf{0} & \textit{irreversibility constraints} \end{array}$$

a) Perform a linear programming (LP) optimization using `linprog` from the MATLAB optimization toolbox (available via the ETH campus license, in Python you can use `scipy.optimize.linprog`), assuming a fixed uptake rate of A and no uptake of B, i.e. $r_1 = 1, r_2 = 0$. What is the maximal yield $Y_{P/A}$?

b) What is the associated flux distribution vector $\mathbf{r}_{\max(Y_{P/A})}$?

c) Perform two further optimizations, this time explicitly providing starting vectors $\mathbf{r'}^0$ and $\mathbf{r''}^0$ using

$$\begin{aligned} \mathbf{r'}^0 &= (1 \quad 0 \quad 0.5 \quad 0.5 \quad 0 \quad 0.5 \quad 0.5 \quad 0 \quad 0 \quad 0.5)^T \\ \mathbf{r''}^0 &= (1 \quad 0 \quad 0.7 \quad 0.3 \quad 0.3 \quad 0.4 \quad 0.3 \quad -0.1 \quad 0.4 \quad 0.3)^T \end{aligned}$$

To set starting vectors, use the function `'fmincon'` instead of `'linprog'` in MATLAB. For Python you can keep the same function. What influence does the variation of starting values have on the outcome of the optimization?

d) How does $Y_{P/A}$ change if you simulate the same scenario for a "single-gene knockout" scenario, i.e. by setting $r_i = 0$ $(1 \leq i \leq 10)$ as additional constraint (resulting in 10 optimizations)?

e) On the basis of these results, what modification (gene knockout) would you suggest to construct an optimized production strain?

Submission:

Please address any questions to

alix.moawad@bsse.ethz.ch, UNG H11