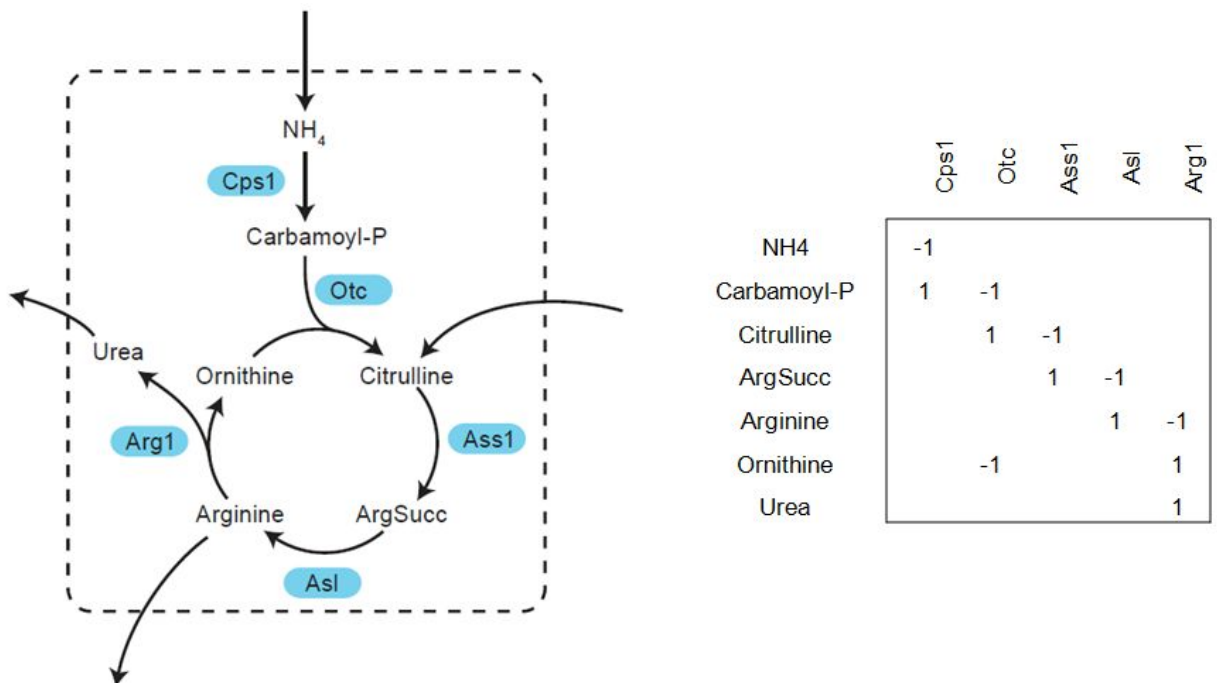# Flux balance analysis

Flux balance analysis (FBA) is a computational method for studying fluxes through metabolic networks at *steady state*. Here, steady state means that metabolic fluxes (reaction rates) and intracellular metabolite concentration are both constant over time. Therefore, the total rates of production and consumption of any *intracellular* metabolite must be equal (otherwise the metabolite concentration would change). We will see how this "balancing" of fluxes help us analyze metabolic networks.

For each reaction acting on metabolite in a network, we assign a negative stoichiometry coefficient $s < 0$ if the metabolite is a substrate, and a positive $s > 0$ if the metabolic is a product of the reaction. For example, for the adenylate kinase reaction

$$AMP + ATP \longleftrightarrow 2\ ADP$$

the stoichiometry coefficient is -1 for AMP, -1 for ATP, and +2 for ADP. To deal with metabolic networks consisting of many reactions and metabolites, we collect all these stoichiometry coefficients in a *stoichiometry matrix S* where metabolites are rows and reactions are columns, and write S_{mr} for the coefficient of metabolite m, reaction r. Below is an example network and corresponding stoichiometry matrix for the urea cycle. For example, the Ass1 has coefficients $S_{33} = -1$ for Citrulline (substrate) and $S_{43} = +1$ for ArgSucc (product). Note that uptake and release fluxes (arrows in/out of the "cell") are not considered as "reactions" here and are not part of the stoichiometry matrix.

	Cps1	Otc	Ass1	Asl	Arg1
NH4	-1				
Carbamoyl-P	1	-1			
Citrulline		1	-1		
ArgSucc			1	-1	
Arginine				1	-1
Ornithine		-1			1
Urea					1

Now let v_r denote the flux through reaction r, in the direction specified by the metabolic network. For example, $v_1 = 100$ in the urea cycle network would denote a flux of 100 (some kind of molar amounts per unit time) through the Cps1 reaction in the forward direction, while negative values mean flux in the reverse direction. Then, the consumption or production of metabolite m due to reaction r is $S_{mr} * v_r$, again positive for production and negative for consumption. Note that since $S_{mr} < 0$ when metabolite m is a substrate, the rate $S_{mr} * v_r$ is negative for substrates, as expected. For example, in the urea cycle network NH_4 is consumed by Cps1 at a rate -100 if $v_1 = 100$, because $S_{11} = -1$. Also, stoichiometry coefficients > 1 will "scale" the production of a metabolite as expected: for example, in the adenylate kinase reaction above, 2 units of ADP is produced per unit time when the flux through the reaction is 1, because the stoichiometry coefficient for ADP is +2.

Now, the *total* production or consumption of a metabolite m in a network is the sum across all n reactions,

$$S_{m1} * v_1 + S_{m2} * v_2 + \ldots + S_{mn} * v_n \qquad (1)$$

(Most of the terms in this sum be zero, since most coefficients $S_{mr} = 0$, as seen in the urea cycle example; only a few reactions in a network will produce or consume any given metabolite.) The steady-state assumption says that this sum should be zero for every *intracellular* metabolite m. For metabolites that are considered to be consumed by the cell (nutrients) the sum is negative, while for byproducts released from the cell, the sum is positive. In the urea cycle model, NH_4 and Citrulline are consumed, while arginine and urea are released; only Carbomyl-P, Ornithine and ArgSucc are considered "intracellular". (Often, we only model a small part of the cellular metabolic network, so "uptake" and "release" and "intracellular" actually refers to the boundaries of the *model*, not of the actual cell.) So in the urea cycle model, we obtain three equations from the "intracellular" metabolites in rows 2, 4 and 6 of the stoichiometry matrix.

$$v_1 - v_2 = 0 \qquad \text{(Carbomyl-P)}$$

$$v_3 - v_4 = 0 \qquad \text{(ArgSucc)}$$

$$-v_2 + v_5 = 0 \qquad \text{(Ornithine)}$$

These equations is the mathematical statement of the fact that fluxes producing and consuming fluxes must balance each other at steady state. *Flux balance analysis (*FBA) refers to computational methods for solving these equations to explore the possible fluxes of a given metabolite network at steady state.

There are many versions of FBA, combining these balancing constraints with experimental data on metabolite uptake or release, with expression data for enzymes; or analyzing effects of loss of an enzyme on metabolic fluxes; or determining fluxes that would be optimal for a certain "task" like maximal ATP production or biomass production. In this computer lab we will explore a simple version of FBA to investigate the capabilities of the pentose phosphate shunt.

Flux balance analysis computer lab

In this exercise we will analyze a model of the pentose phosphate pathway (PPP). This is a good example of a model where FBA can be helpful, as it is fairly complicated with several loops and branches, and its function is not obvious.

The model

First look at the network figure and make sure you recognize the various reactions and metabolites in the pentose phosphate pathway. (The representation may be a bit unusual!) The abbreviated names for reactions and metabolites in this network are derived from a large published model for the human metabolic network [Thiele, 2013].

In Excel (2007 or later), make sure the optimization solver add-on is activated: go to excel options -- add-ins, click "Go … " next to "manage add-ins", and make sure the "Solver" add-in is checked. Open the provided excel file.

The largest part of the sheet is the matrix of stoichiometry coefficients that describe the network model. As with the urea cycle model above, a negative value indicates that a metabolite (in rows) is consumed by a reaction (in columns), while a positive value indicates the metabolite is produced. Check some of the reactions against the network model and make sure you understand the representation.

At the bottom is (row 25) are the values for fluxes through each reaction (the v_r in the above), initially all zero. Positive values will mean net flux in the forward direction; negative values means in reverse. To the right (column P) is the total consumption (negative) or production (positive) of each metabolite, calculated from the stoichiometry and fluxes as in equation (1) above.

1. Try setting some fluxes to specific values (enter into row 25) to see how the totals per metabolite changes. Compare with equation (1) above, and make sure you understand how the total is calculated.

Columns Q, R hold the constraints. (Column Q is only for readability, not used in calculations). Setting a metabolite = 0 means the metabolite is required to balance (is considered intracellular). Glucose (glc-D) is set = -100 here to model uptake of this metabolite at a specified rate. Metabolites set to be >= 0 may be produced (at any rate) but not taken up (co2, for example), while those set to <= 0 may be consumed but not produced (atp, for example). Finally, h2o is not constrained at all in this model, denoted "free".

2. Looking at the pathway diagram, set fluxes manually (by entering values in row 25) so that the model takes up 1 glucose (total for glc-D = -1) and produces glyceraldehyde-2-phosphate

(g3p). This in effect bypasses the PPP. Are all internal metabolites accounted for (all = 0 constraints really zero) ? What byproducts do you get?
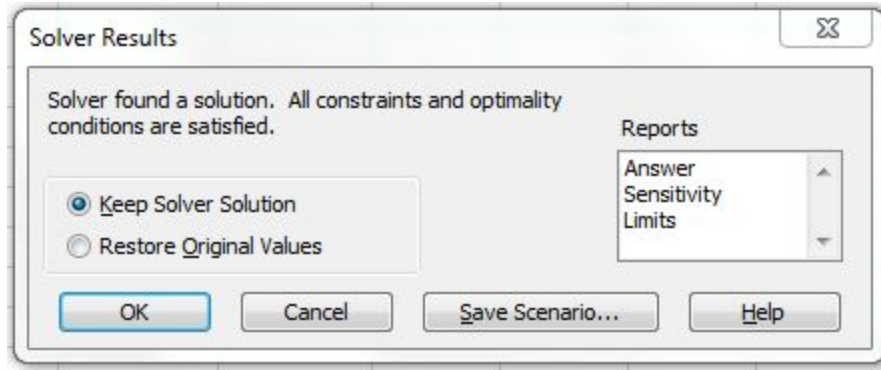
Ribose production

We will next use the Solver add-in to solve the flux balance equations and investigate how much ribose-5-phosphate (r5p) can (theoretically) be generated by the PPP from glucose. But before doing this, study the PPP pathway and try to figure out how r5p should be synthesized.

3. Intuitively, what reactions are used to synthesize r5p? How many r5p do you think can be obtained from 100 units of glucose?

Now use the solver to find a maximal r5p production solution by choosing the total for r5p as the target, set to maximize. Go to Data -- Solver to see the solver configuration; it should look like this:
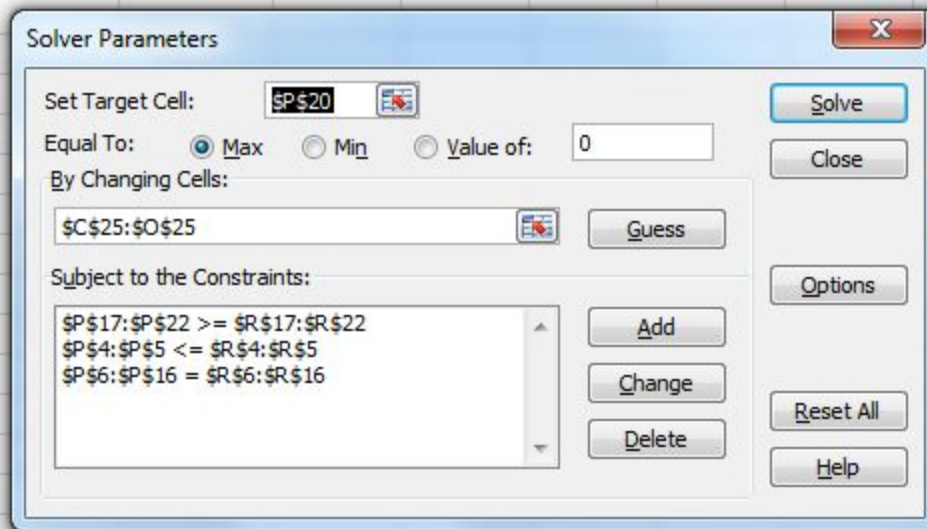
The range under "By Changing Cells" is the cells holding the fluxes (variables). Below this are cells holding the flux balance constraints, defining what metabolites must be produced (>=), consumed (<=) or balanced (=). Set the target cell to P20 to maximize ribose-5-phosphate and click solve. If everything is correct the solver should reply like this:

Click OK and investigate the total r5p production and the values of fluxes through the various reactions.

Q: Is the maximal amount of r5p what you expected? Does it make sense considering the number of carbons in ribose and glucose?

Q: Which reactions produce r5p in this flux state? (Look at the stoichiometry coefficients for r5p and the values of the corresponding fluxes.)

Q: Is this an oxidative process? Does it require energy?

Q: The product r5p contains not only the carbon from glucose, but also a phosphate group. Where does this phosphate derive from in this model?

Q: Write down the net reaction carried out by this network at this flux state.

If you feel the result is counterintuitive, explore the fluxes found for other reactions! Does the result "add up" ? How are the interconnected reactions carried out by the transketolases (TKT1, TKT2) and transaldolase (TALA) exploited in this flux state? It may help to fill in flux values on top of the pathway schematic to visualize the results.

Glucose oxidation to CO2

The PPP can also oxidize glucose to CO2, and use the energy obtain to produce the reducing agent NADPH. Before using the solver to investigate this, answer the following by only looking at the pathway:

**Q:** Which reactions / pathway in the network oxidizes glucose to CO2?

Q: How much CO2 would you think the PPP can obtain from 100 units of glucose? How much NADPH?

As before, use the solver to find the maximal co2 production solution, now choosing total CO2 as target, set to maximize.

Q: Is the maximal amount of co2 what you expected? Does it make sense considering the number of carbons in glucose?

Q: Which reactions are involved in this flux state? Are the values reasonable?

In this model we have not taken into account that reactions may be irreversible in practise, for example phosphofructokinase (PFK) which is driven by ATP. This is reflected by the fact that negative flux values are allowed in row 25. To prevent this, add a new constraint to model, setting the flux through PFK to be positive (cell E25 >=0 in the below).
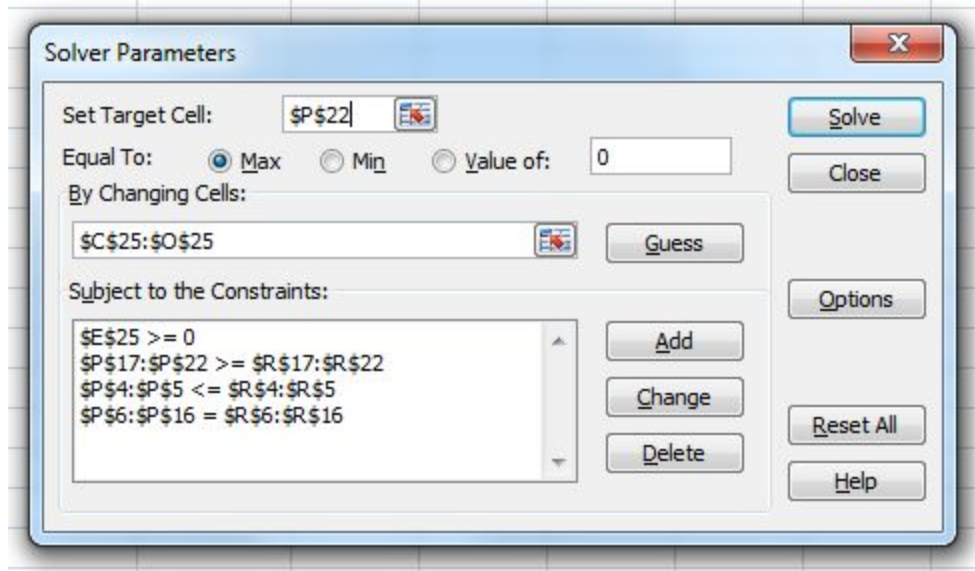
Solve the model again with this change.

Q: How does the result change? How much co2 and nadph is produced, compared to before?

Q: Is the result what you expected by simply looking at the pathway diagram?

Q: What are the fates of carbon from glucose in this flux state, besides co2?

Q: Is this mode of metabolism acidifying?

You can also try producing maximizing g3p (glycolysis), or maximizing flux through a specific reaction (use the flux as target), or modifying the model (what happens if ATP is limited, for example?), or building your own model of a pathway of interest.

Discussion points

The pentose phosphate shunt can operate in more than one "mode". The flux states you obtain from maximizing production of metabolites as above are "extreme" examples, using one "mode" exclusively (ribose or nadph production, respectively). What would a flux distribution look like that produces *both* ribose and nadph?

Results from computational methods like flux balance analysis is only as good as the underlying model. What assumptions are implied in the model we have worked with?

Do you find the results from flux balance analysis counterintuitive? Could the flux solutions you have obtained actually occur in cells? Why / why not?

The steady state assumption is of course an approximation of reality. Can you give examples of conditions where the assumption is reasonable or not? How can the assumption be checked?

References

1. Thiele I, Swainston N, Fleming RMT, Hoppe A, Sahoo S, Aurich MK, et al. A community-driven global reconstruction of human metabolism. Nat Biotechnol 31(5):419–25, 2013.

Pentose phosphate pathway model
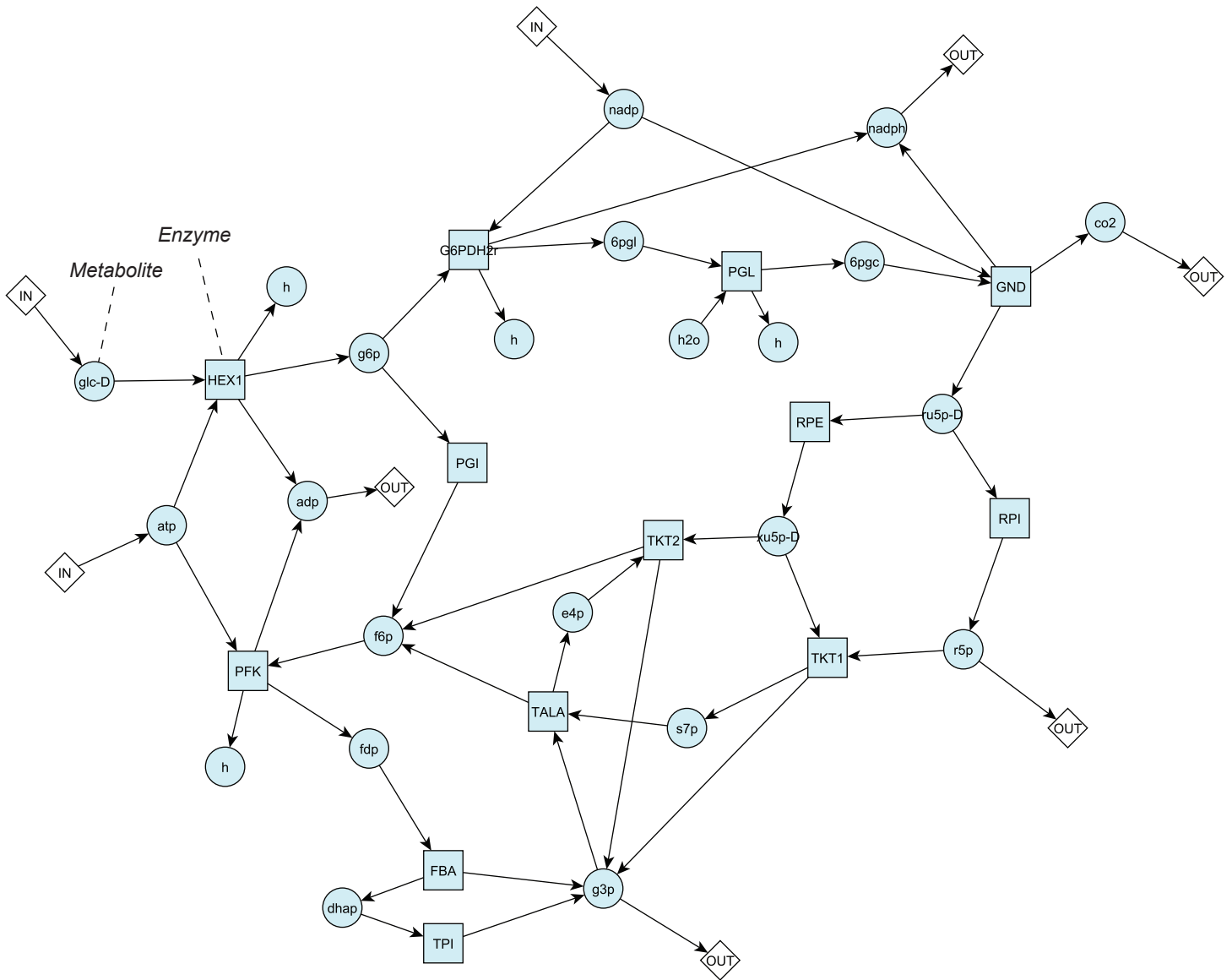
Abbreviations

Reactions

FBA	fructose biphosphate aldolase
G6PDH2r	glucose-6-phosphate dehydrogenase
GND	6-phosphogluconate dehydrogenase
HEX1	hexokinase
PFK	phosphofructokinase
PGI	phosphoglucose isomeras
PGL	6-phosphogluconolactonase
RPE	ribulose epimerase
RPI	ribulose isomerase
TALA	transaldolase
TKT1	transketolase 1
TKT2	transketolase 2
TPI	triose phosphate isomerase

Metabolites

6pgc	6-phosphogluconate
6pgl	6-phosphogluconolactone
adp	adenosine diphosphate
atp	adenosine triphosphate
co2	carbon dioxide
dhap	dihydroxyacetone-phosphate
e4p	erythrose-4-phosphate
f6p	fructose-6-phosphate
fdp	fructose-diphosphate
glc-D	D-glucose
g3p	glyceraldehyde-3-phosphate
g6p	glucose-6-phosphate
h	H+
h2o	water
nadp	NADP
nadph	NADPH
r5p	ribose-5-phosphate
ru5p-D	D-ribulose-5-phosphate
s7p	seduheptulose-7-phosphate
xu5p-D	D-xylulose-5-phosphate