

Introduction to Genome Scale Metabolic Models

Rui Benfeitas

NBIS - National Bioinformatics Infrastructure Sweden
Science for Life Laboratory, Stockholm
Stockholm University

rui.benfeitas@scilifelab.se



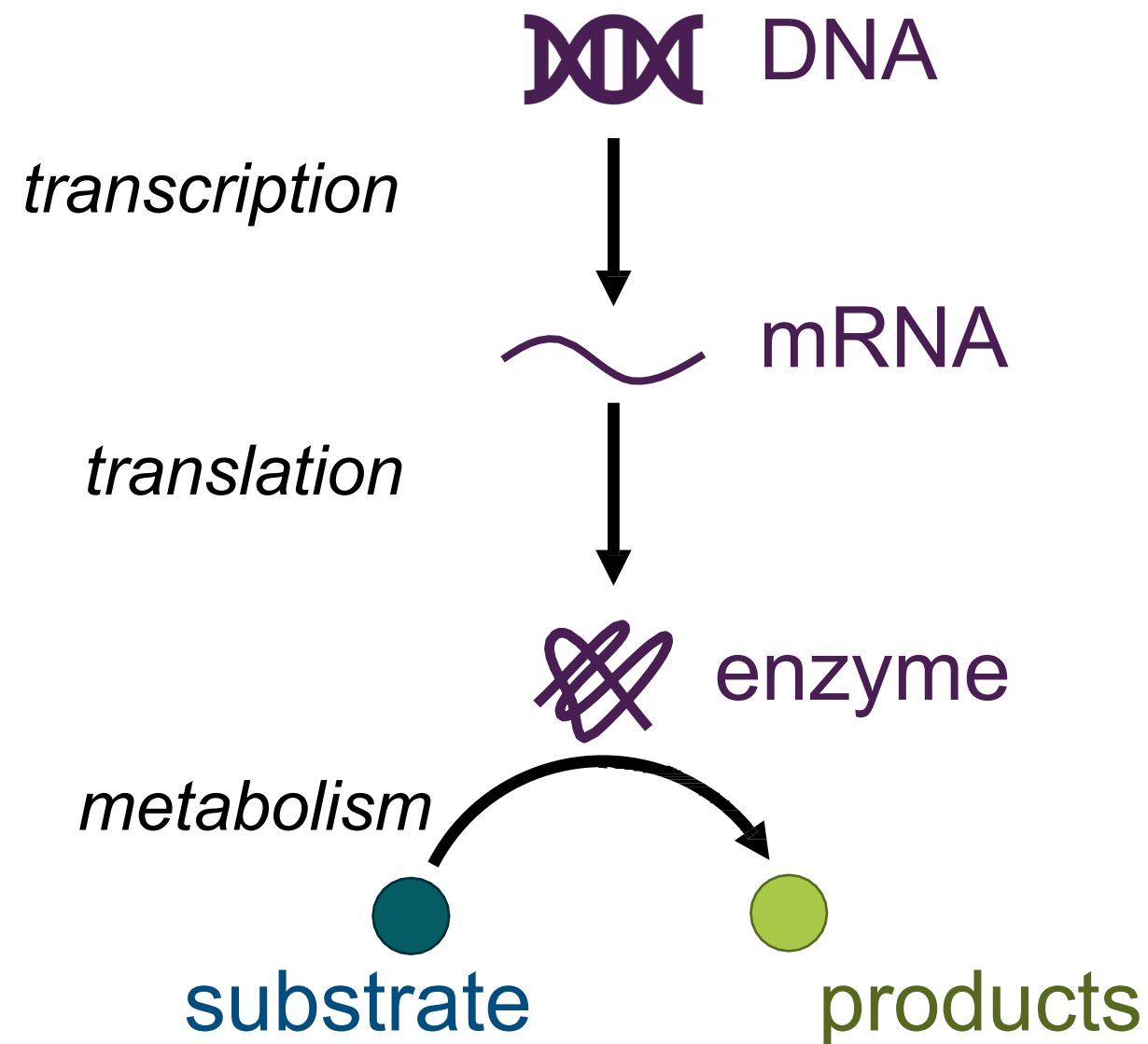
SciLifeLab



Overview

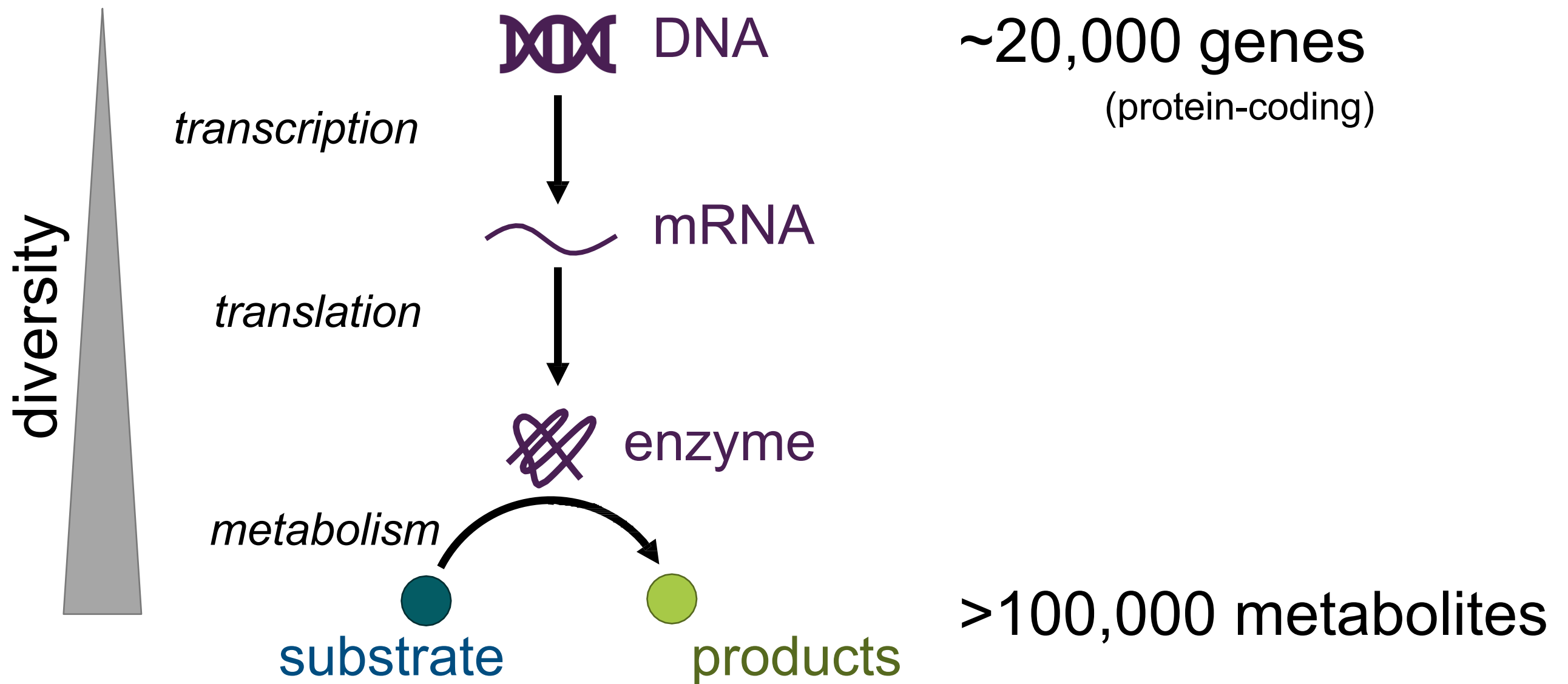
1. The problem in characterising fluxes
2. Rationale behind metabolic modelling
3. Employing GEMs in simulating metabolic fluxes

Central dogma as starting point in modelling metabolism

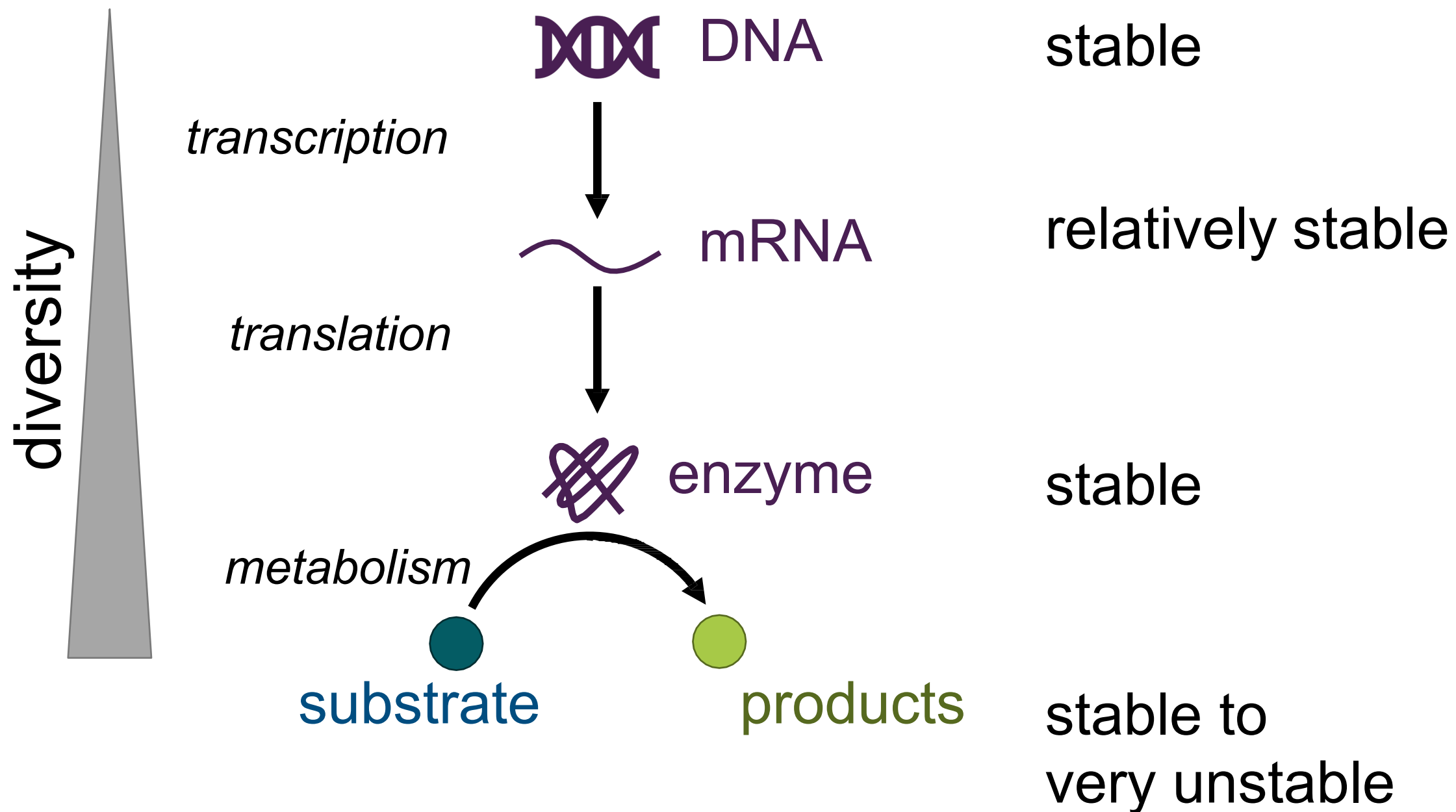


Metabolism provides the **energy** and **building blocks** necessary to sustain life.

Central dogma as starting point in modelling



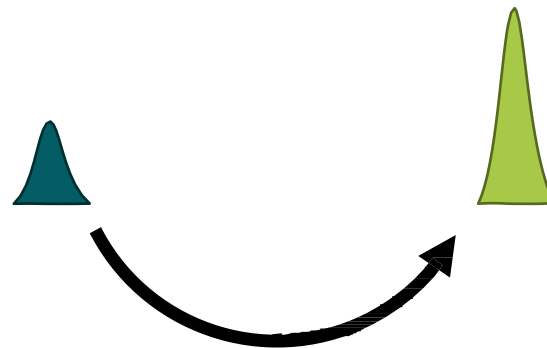
Central dogma as starting point in modelling



Quantifying fluxes



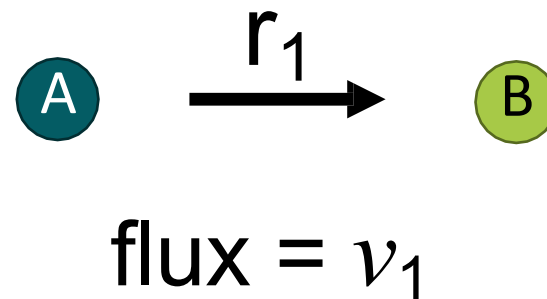
We can generally measure metabolite concentrations



...but what is often important is the flow or **flux** of metabolites through the reactions.

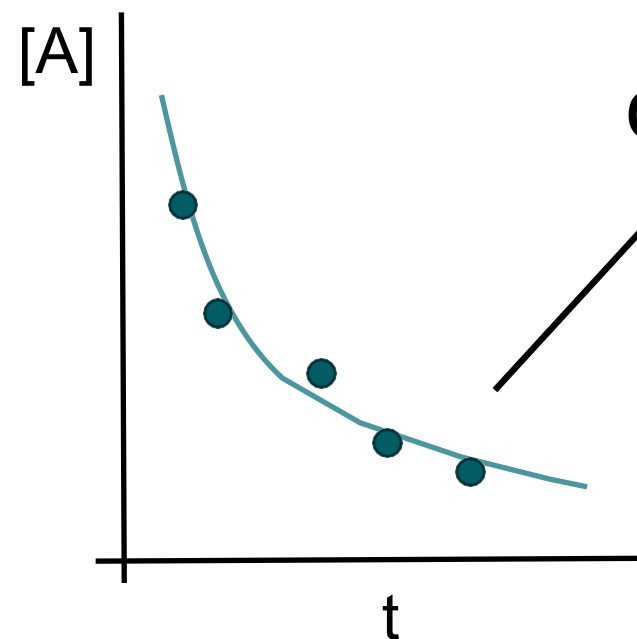


Quantifying fluxes



$$\frac{d[A]}{dt} = -v_1$$

$$\frac{d[B]}{dt} = v_1$$

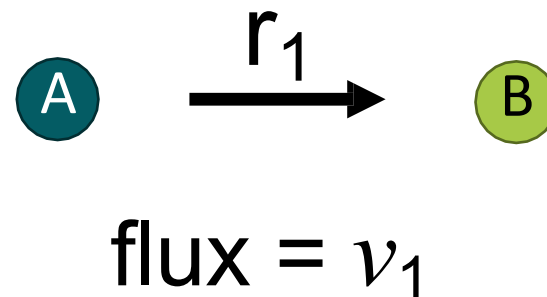


calculate v_1

v_1 = production rate of B

$$v : \frac{mmol}{g_{DCW} \cdot h}$$

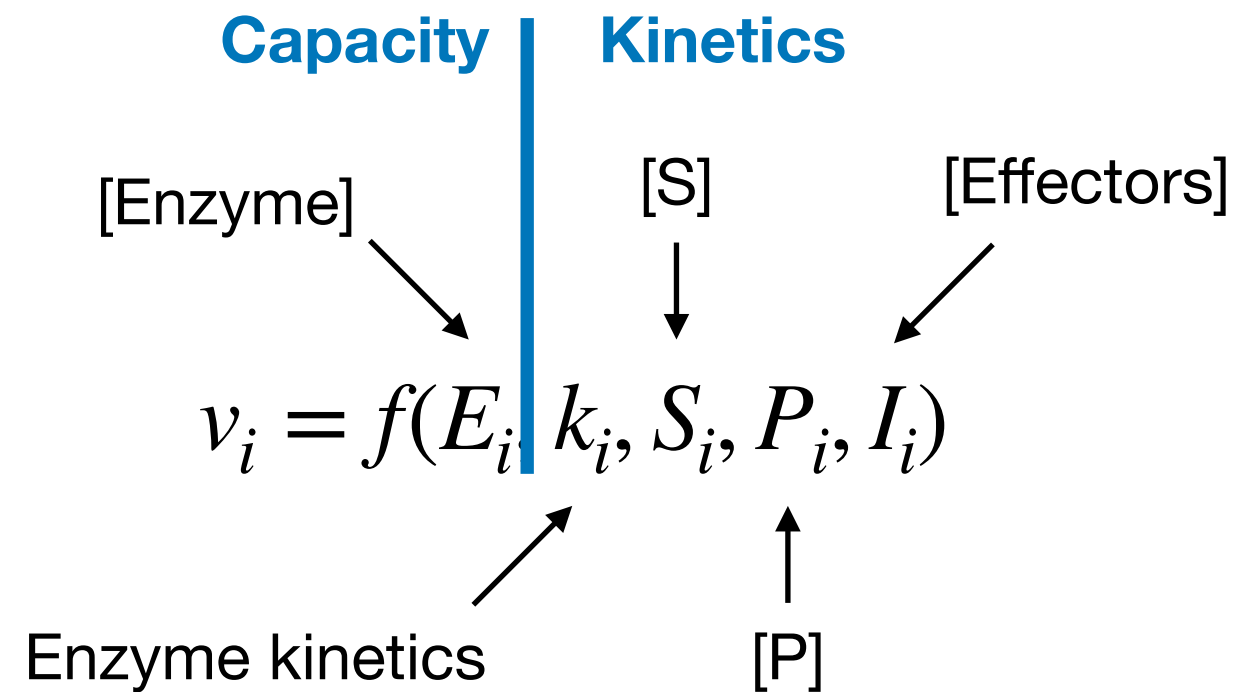
Enzyme kinetics require knowledge of many kinetic parameters



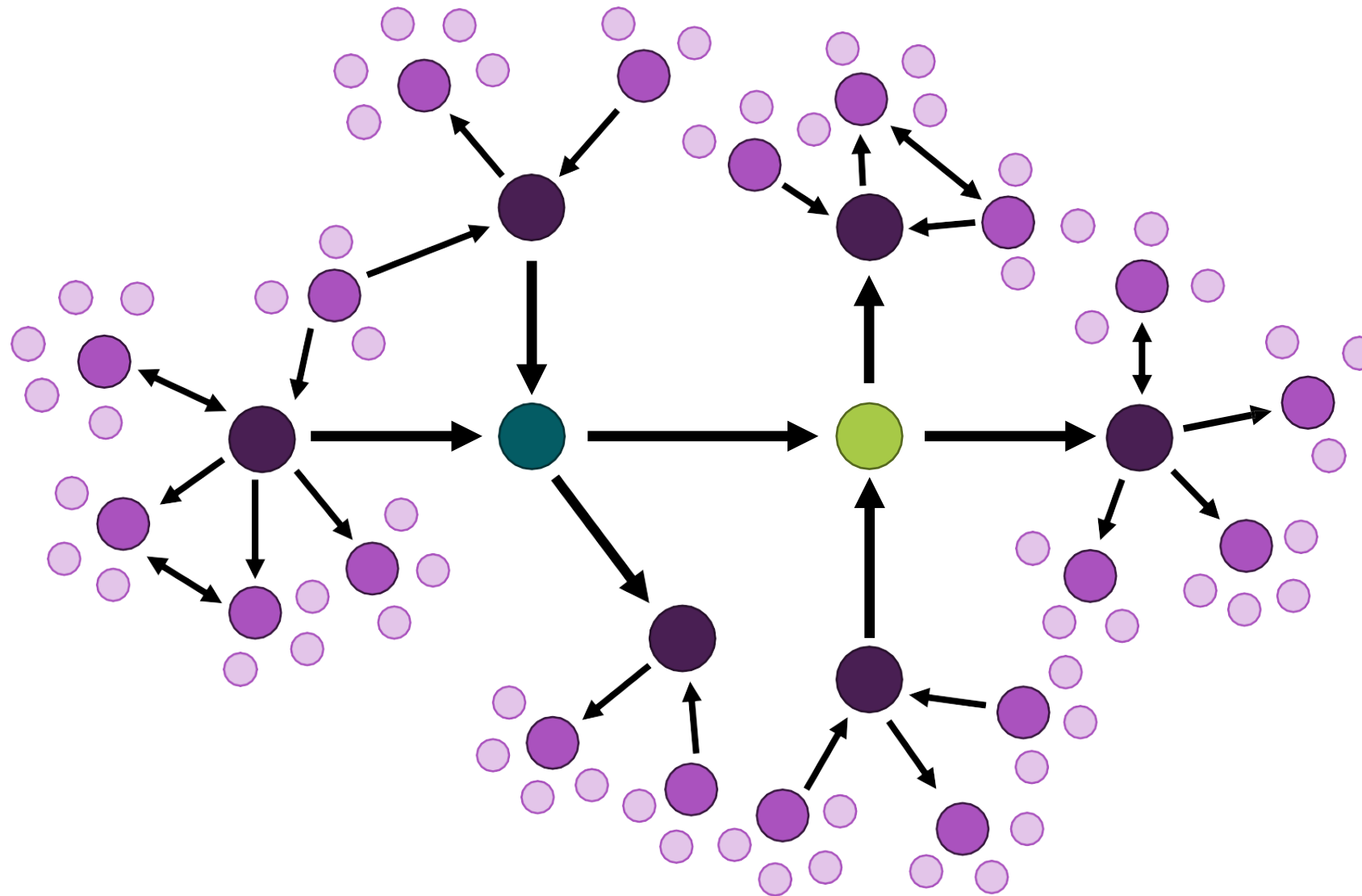
Estimated experimentally

$$\frac{d[A]}{dt} = -v_1 = \underline{k_1} \times [A]$$

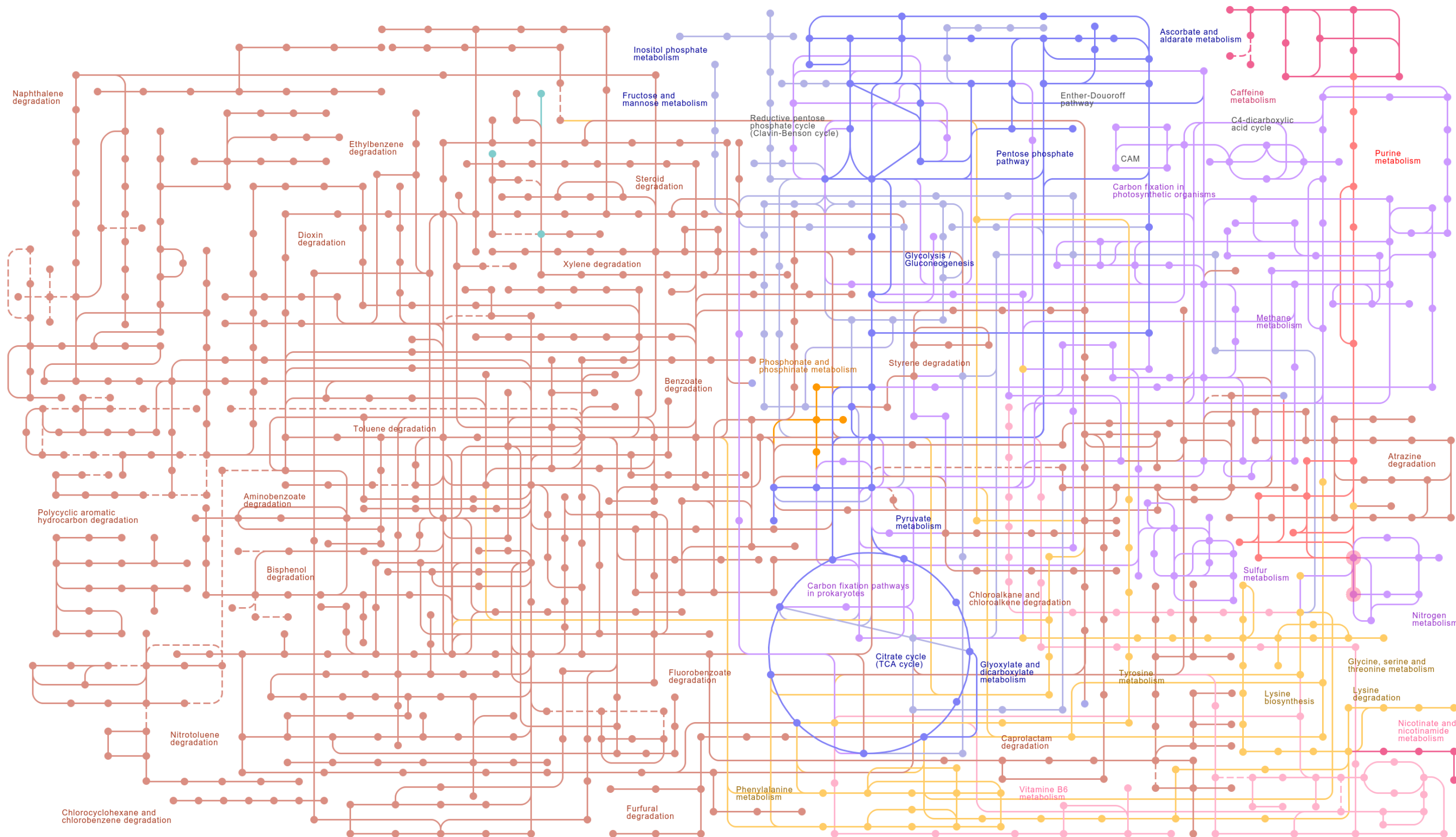
$$\frac{d[A]}{dt} = -v_1 = \frac{V_{max} \times [A]}{\underline{K_M} + [A]}$$



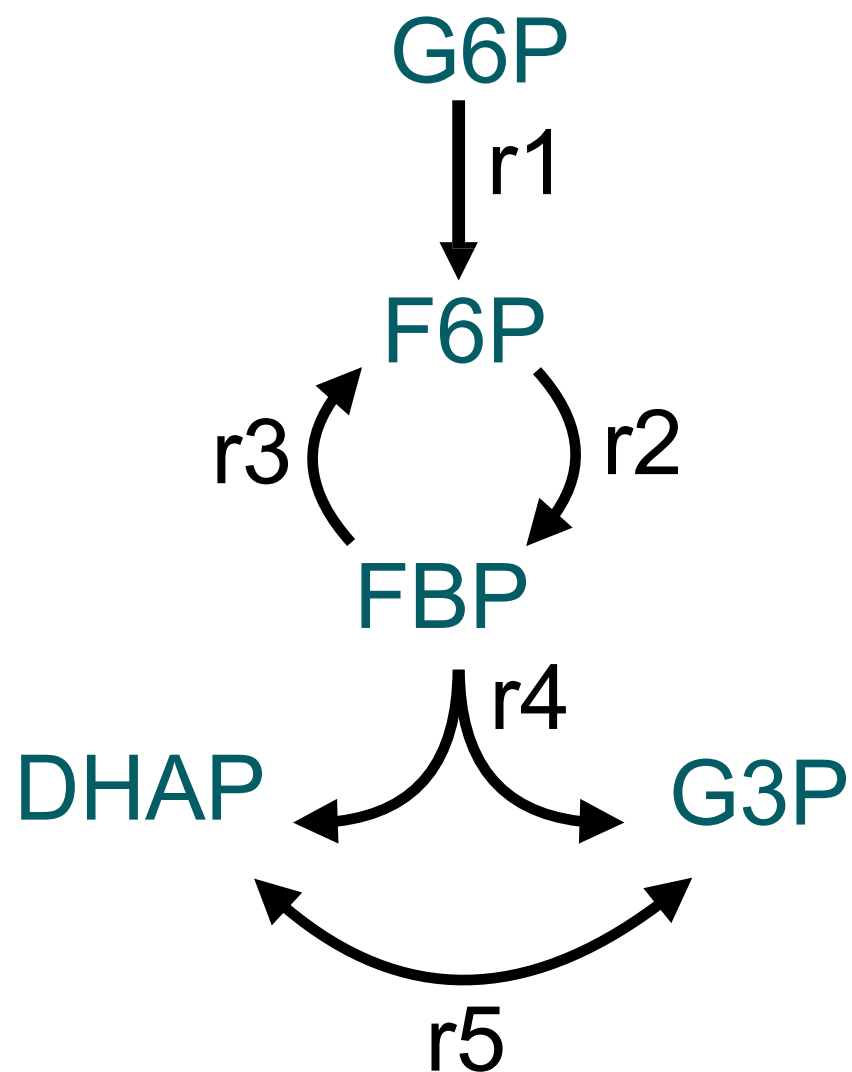
Expanding flux simulations globally



Expanding flux simulations globally



Using reaction stoichiometry to describe metabolism



Metabolites

Reactions

| | r1 | r2 | r3 | r4 | r5 |
|------|----|----|----|----|----|
| G6P | -1 | 0 | 0 | 0 | 0 |
| F6P | 1 | -1 | 1 | 0 | 0 |
| FBP | 0 | 1 | -1 | -1 | 0 |
| DHAP | 0 | 0 | 0 | 1 | -1 |
| G3P | 0 | 0 | 0 | 1 | 1 |

Genome-scale model (GEM)

Chemical formula
Charge
InChI code
Other external IDs
...

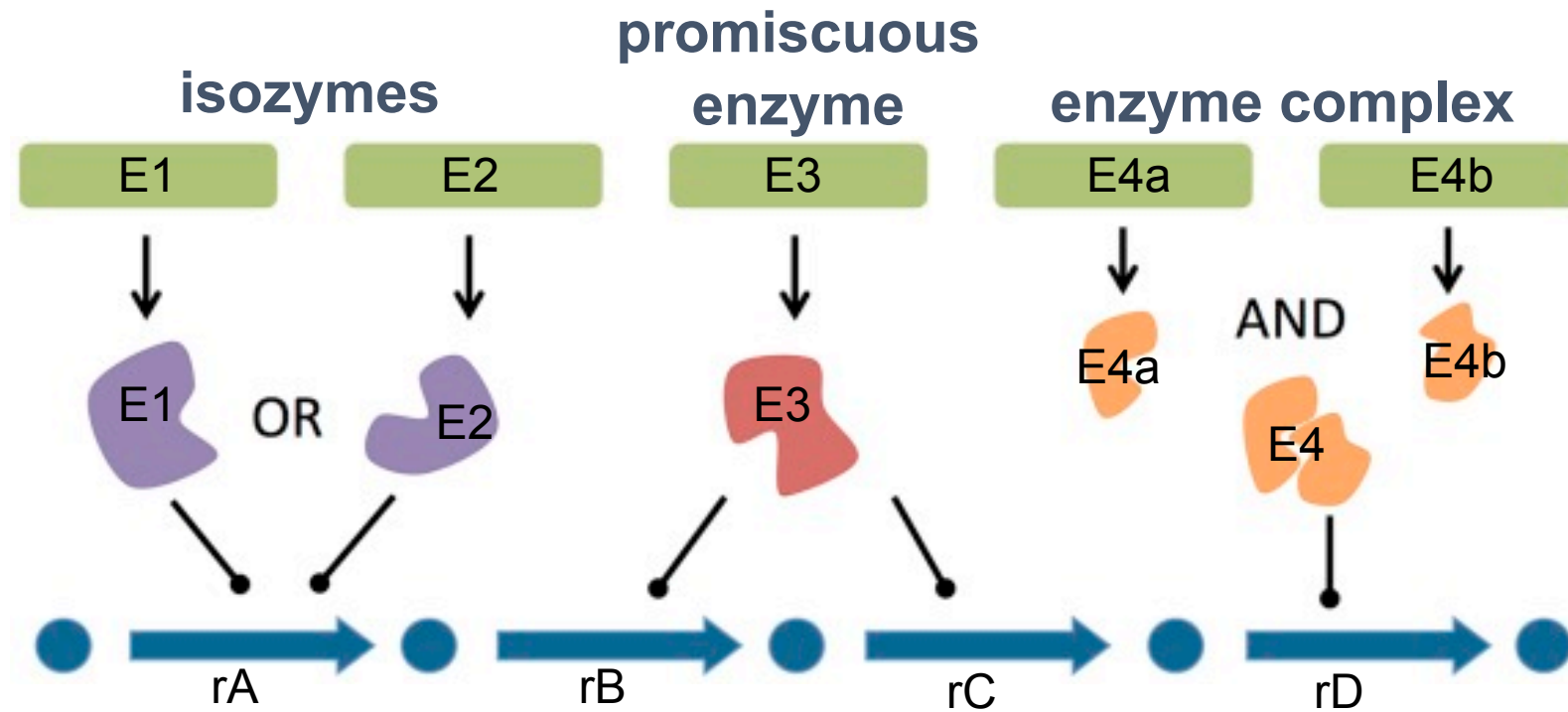
...
...

| KEGG ID | Compartment | Name | Symbol | r1 | r2 | r3 | r4 | r5 |
|---------|-------------|----------------------------|--------|----|----|----|----|----|
| C00668 | cytosol [c] | glucose 6-phosphate | G6P | -1 | 0 | 0 | 0 | 0 |
| C00085 | cytosol [c] | fructose 6-phosphate | F6P | 1 | -1 | 1 | 0 | 0 |
| C00354 | cytosol [c] | fructose-1,6-bisphosphate | FBP | 0 | 1 | -1 | -1 | 0 |
| C00111 | cytosol [c] | dihydroxyacetone phosphate | DHAP | 0 | 0 | 0 | 1 | -1 |
| C00118 | cytosol [c] | glyceraldehyde 3-phosphate | G3P | 0 | 0 | 0 | 1 | 1 |
| | | | ... | | | | | |

Genome-scale model (GEM)

| | | | | | | Genes (symbol) | Proteins (UniProt) | Transcript IDs | GO Terms | Orthologs |
|------|----|----|----|----|----|--|--------------------|----------------|----------|-----------|
| | | | | | | GPI | P06744 | | | |
| | | | | | | <i>n/a</i> | | | | |
| | | | | | | FBP1, FBP2 | P09467, O00757 | ... | | |
| | | | | | | ALDOA, ... | P04075, ... | | | |
| | | | | | | TPI1 | P60174 | | | |
| | | | | | | r1 | r2 | r3 | r4 | r5 |
| G6P | -1 | 0 | 0 | 0 | 0 | | | | | |
| F6P | 1 | -1 | 1 | 0 | 0 | | | | | |
| FBP | 0 | 1 | -1 | -1 | 0 | | | | | |
| DHAP | 0 | 0 | 0 | 1 | -1 | | | | | |
| G3P | 0 | 0 | 0 | 1 | 1 | | | | | |
| | | | | | | Reactions linked to genes that encode the enzymes that catalyze the reaction | | | | |
| | | | | | | “gene-protein rules” (GPR rules) | | | | |

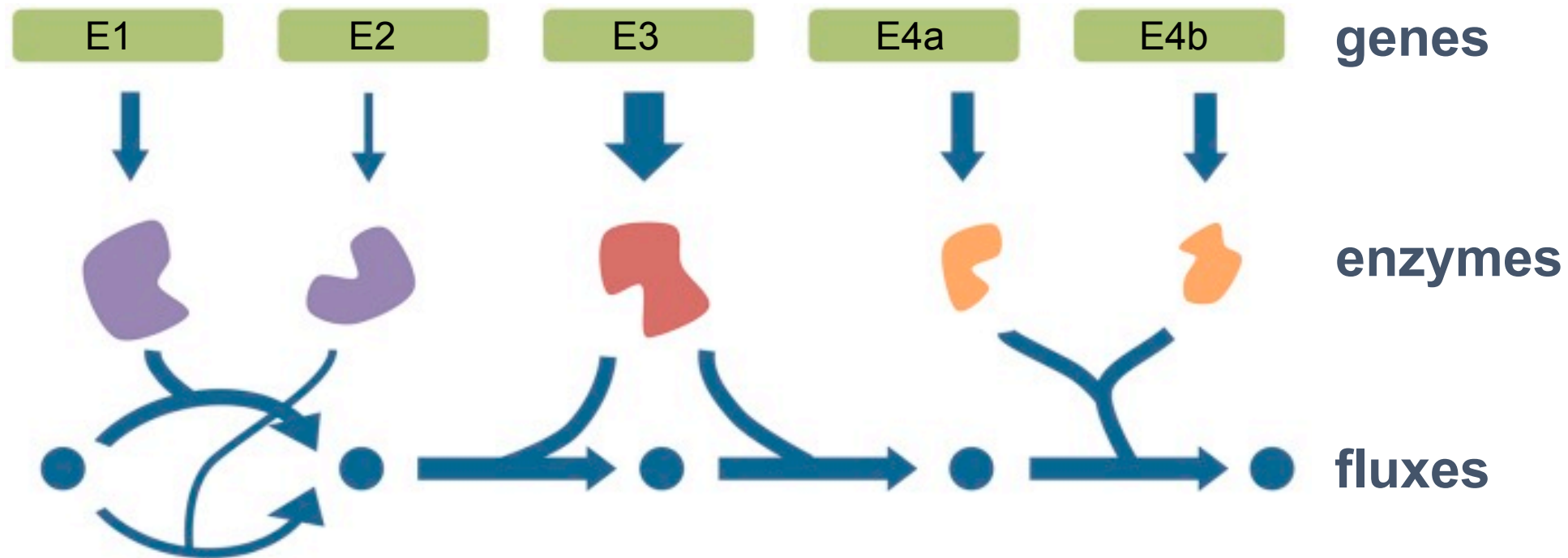
GPR rules



GPR Rules enable more accurate simulation of gene inactivation/knock-out

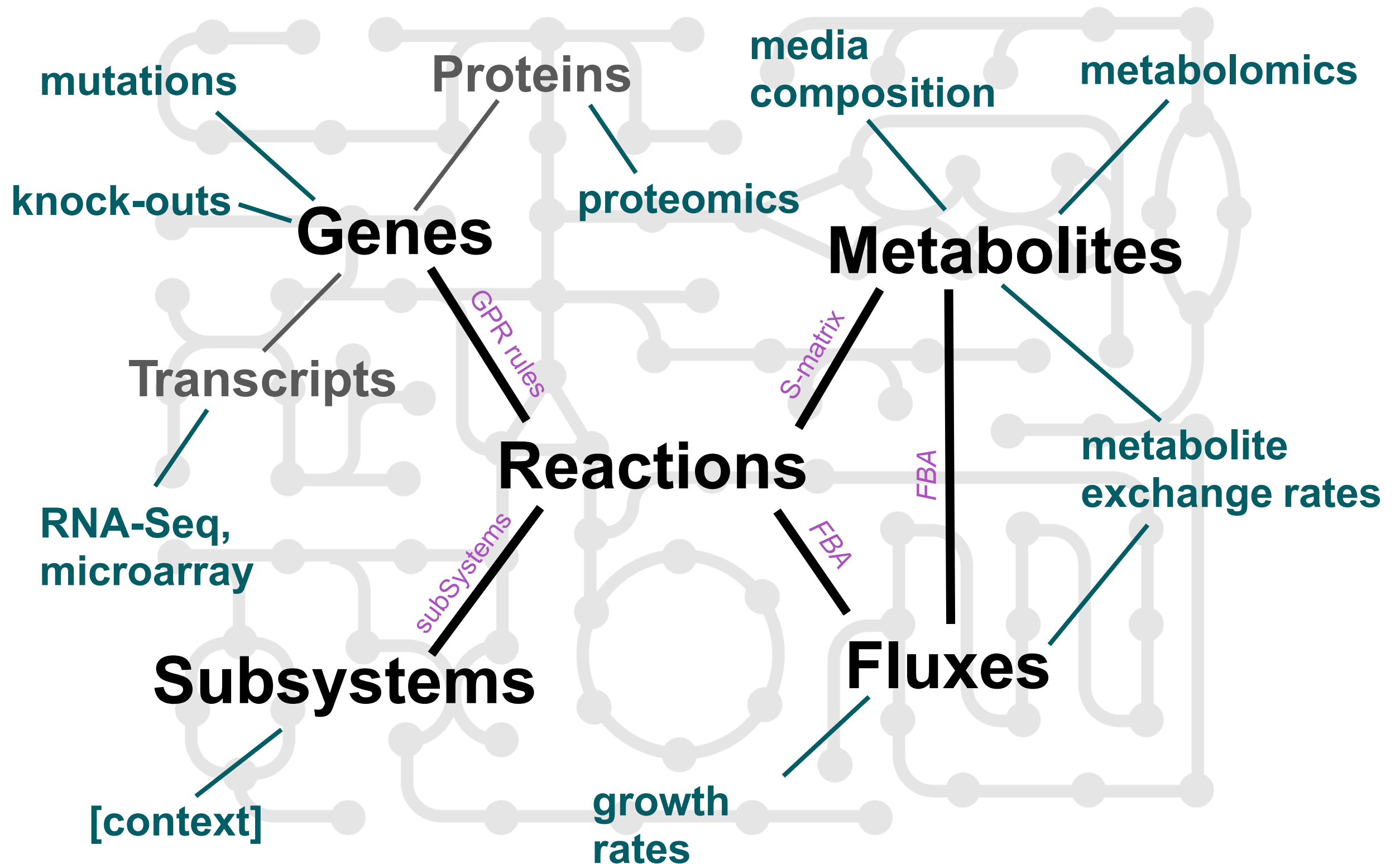
| Knockout | Effect |
|-----------|----------------|
| E1 | none |
| E2 | none |
| E1 + E2 | rA inactive |
| E3 | rB rC inactive |
| E4a | rD inactive |
| E4b | rD inactive |
| E4a + E4b | rD inactive |

GPR rules

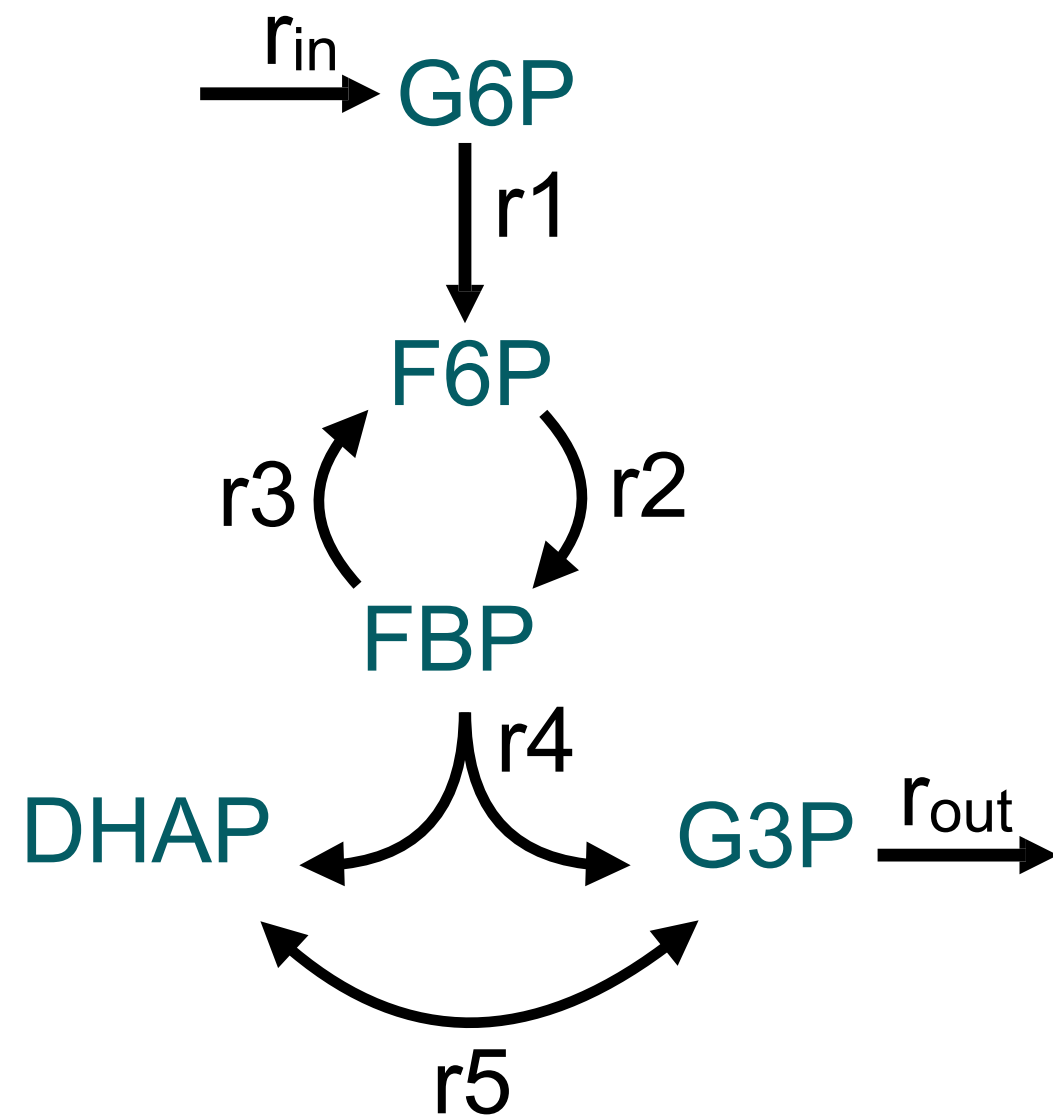


GPR rules can be linked
with gene expression

GEMs as an integrative tool



Flux Balance Analysis (FBA)

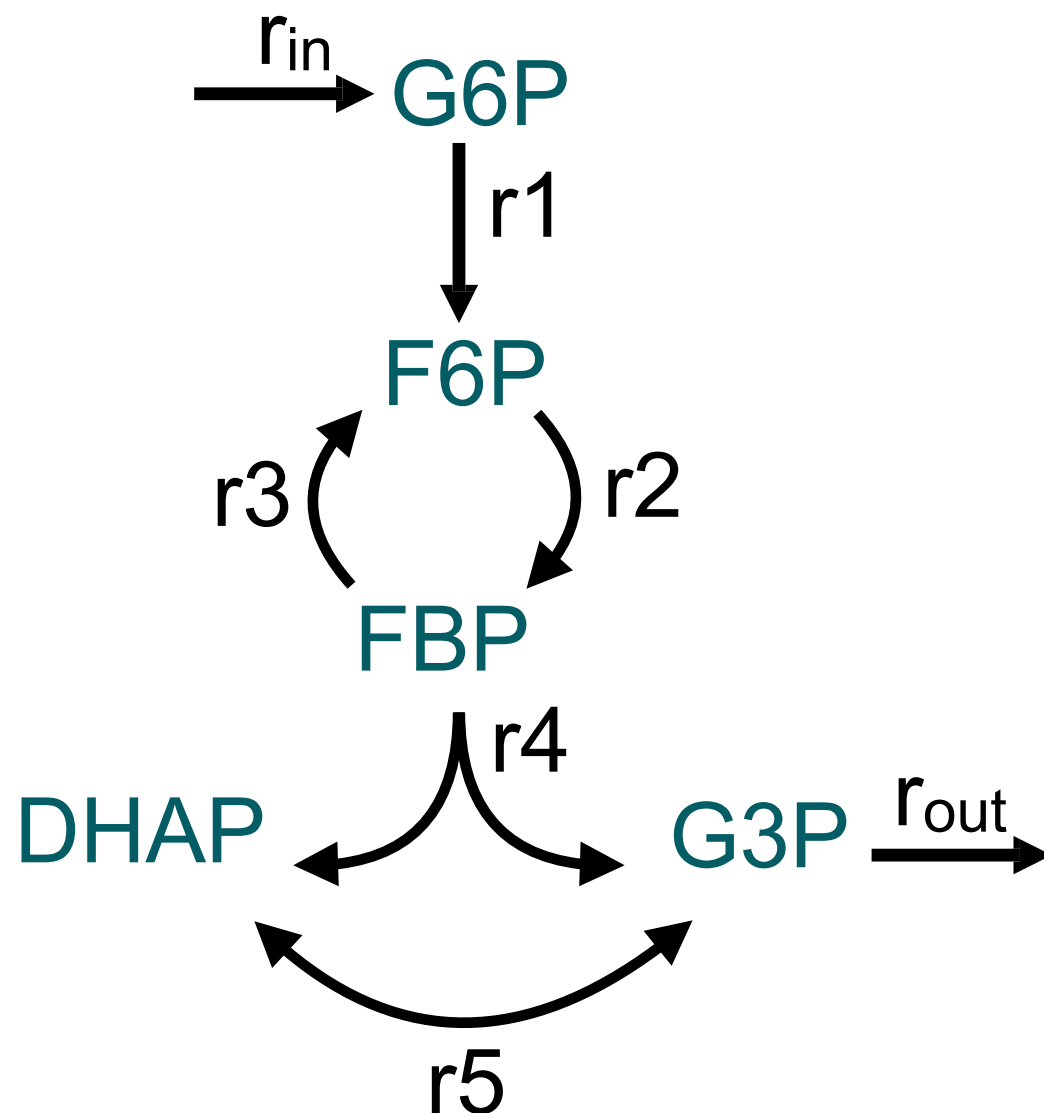


| | r1 | r2 | r3 | r4 | r5 | r _{in} | r _{out} |
|------|----|----|----|----|----|-----------------|------------------|
| G6P | -1 | 0 | 0 | 0 | 0 | 1 | 0 |
| F6P | 1 | -1 | 1 | 0 | 0 | 0 | 0 |
| FBP | 0 | 1 | -1 | -1 | 0 | 0 | 0 |
| DHAP | 0 | 0 | 0 | 1 | -1 | 0 | 0 |
| G3P | 0 | 0 | 0 | 1 | 1 | 0 | -1 |

Flux Balance Analysis (FBA)

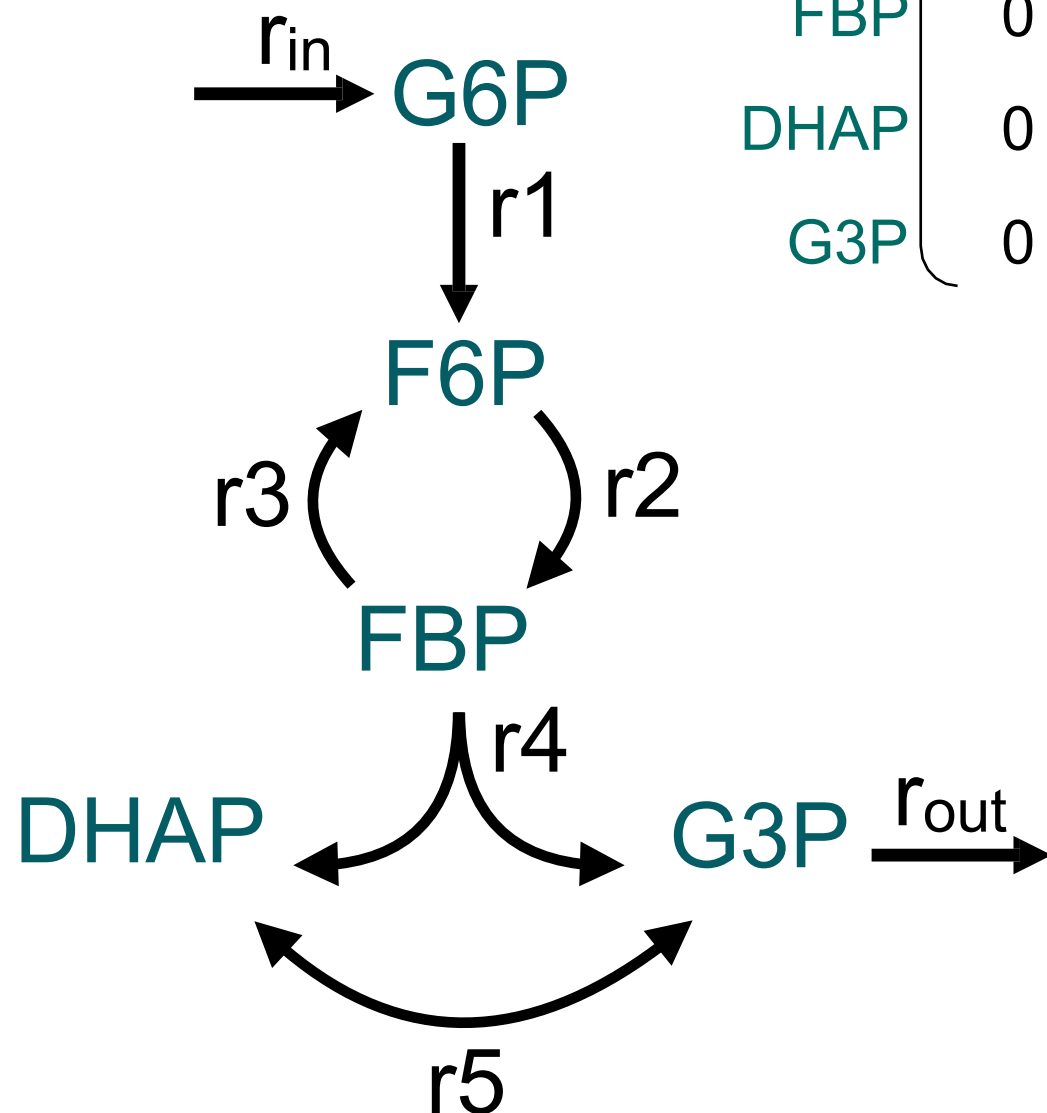
FBA seeks to calculate the reaction **fluxes** (v) of a network

Based on the **conservation of mass**:
it cannot be created or destroyed



$$\begin{aligned}\frac{d[A]}{dt} &= v_{prod} - v_{cons} \\ &= S \times v\end{aligned}$$

Flux Balance Analysis (FBA)



$$\begin{array}{c}
 \text{G6P} \\
 \text{F6P} \\
 \text{FBP} \\
 \text{DHAP} \\
 \text{G3P}
 \end{array}
 \begin{array}{c}
 r_1 \quad r_2 \quad r_3 \quad r_4 \quad r_5 \quad r_{in} \quad r_{in} \\
 \left(\begin{array}{cccccc}
 -1 & 0 & 0 & 0 & 0 & 1 & 0 \\
 1 & -1 & 1 & 0 & 0 & 0 & 0 \\
 0 & 1 & -1 & -1 & 0 & 0 & 0 \\
 0 & 0 & 0 & 1 & -1 & 0 & 0 \\
 0 & 0 & 0 & 1 & 1 & 0 & -1
 \end{array} \right)
 \times
 \begin{array}{c}
 v_1 \\
 v_2 \\
 v_3 \\
 v_4 \\
 v_5 \\
 v_{in} \\
 v_{out}
 \end{array}
 =
 \begin{array}{c}
 d\text{G6P}/dt \\
 d\text{F6P}/dt \\
 d\text{FBP}/dt \\
 d\text{DHAP}/dt \\
 d\text{G3P}/dt
 \end{array}$$

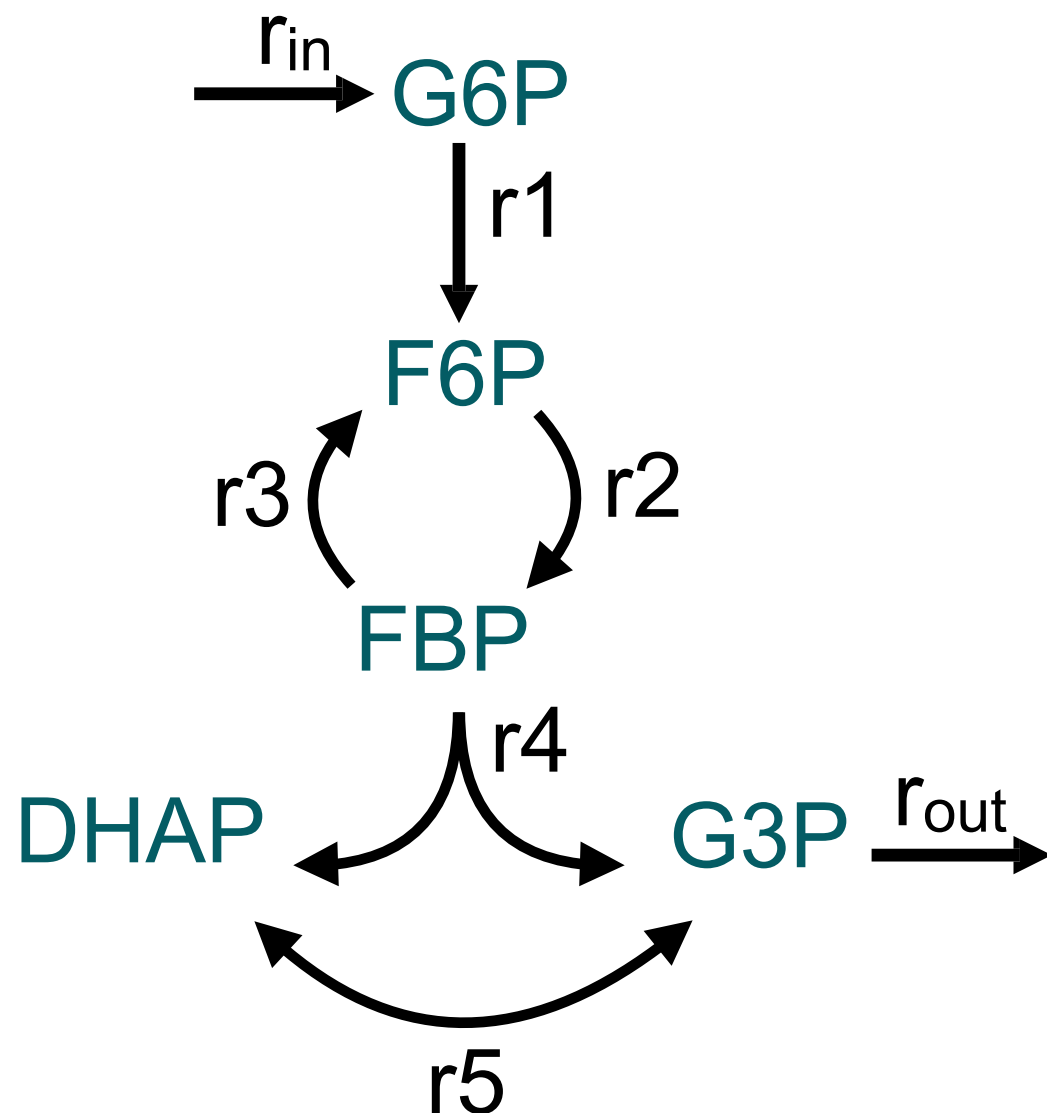
$$\frac{d[\text{G6P}]}{dt} = v_{in} - v_1$$

$$\frac{d[\text{G3P}]}{dt} = v_4 + v_5 - v_{out}$$

$$v_i = f(E_i, k_i, S_i, P_i, I_i)$$

Flux Balance Analysis (FBA)

A key assumption to FBA is **steady state**:
metabolite concentrations are **constant** through time



$$\frac{d[X]}{dt} = v_{prod} - v_{cons} = 0$$

$$\Rightarrow v_{prod} = v_{cons}$$

This assumption allows us to **ignore enzyme kinetics**, thus eliminating **many** unknown parameters

Flux Balance Analysis (FBA)

$$\begin{array}{c}
 \text{G6P} \\
 \text{F6P} \\
 \text{FBP} \\
 \text{DHAP} \\
 \text{G3P}
 \end{array}
 \begin{array}{c}
 r1 \quad r2 \quad r3 \quad r4 \quad r5 \quad r_{in} \quad r_{in} \\
 \left(\begin{array}{cccccc}
 -1 & 0 & 0 & 0 & 0 & 1 & 0 \\
 1 & -1 & 1 & 0 & 0 & 0 & 0 \\
 0 & 1 & -1 & -1 & 0 & 0 & 0 \\
 0 & 0 & 0 & 1 & -1 & 0 & 0 \\
 0 & 0 & 0 & 1 & 1 & 0 & -1
 \end{array} \right)
 \end{array}
 \begin{array}{c}
 v_1 \\
 v_2 \\
 v_3 \\
 v_4 \\
 v_5 \\
 v_{in} \\
 v_{out}
 \end{array}
 =
 \begin{array}{c}
 d\text{G6P}/dt \\
 d\text{F6P}/dt \\
 d\text{FBP}/dt \\
 d\text{DHAP}/dt \\
 d\text{G3P}/dt
 \end{array}
 =
 \begin{array}{c}
 0 \\
 0 \\
 0 \\
 0 \\
 0
 \end{array}$$

$$S \cdot v = 0$$

Flux Balance Analysis (FBA)

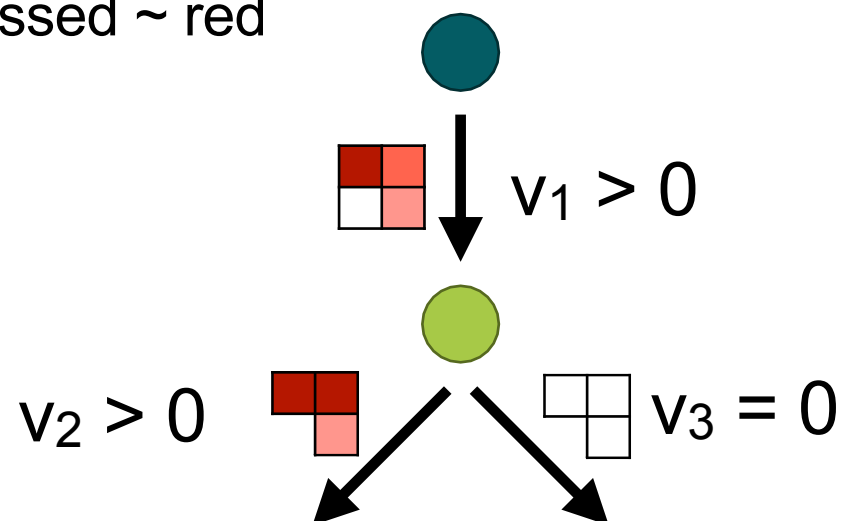
We can further constrain the solution space by limiting reaction fluxes based on their reversibility:

Irreversible reactions  $0 \leq v \leq ub$

Reversible reactions  $lb \leq v \leq ub$

Gene expression:

Expressed ~ red



Others:

Enzyme capacity

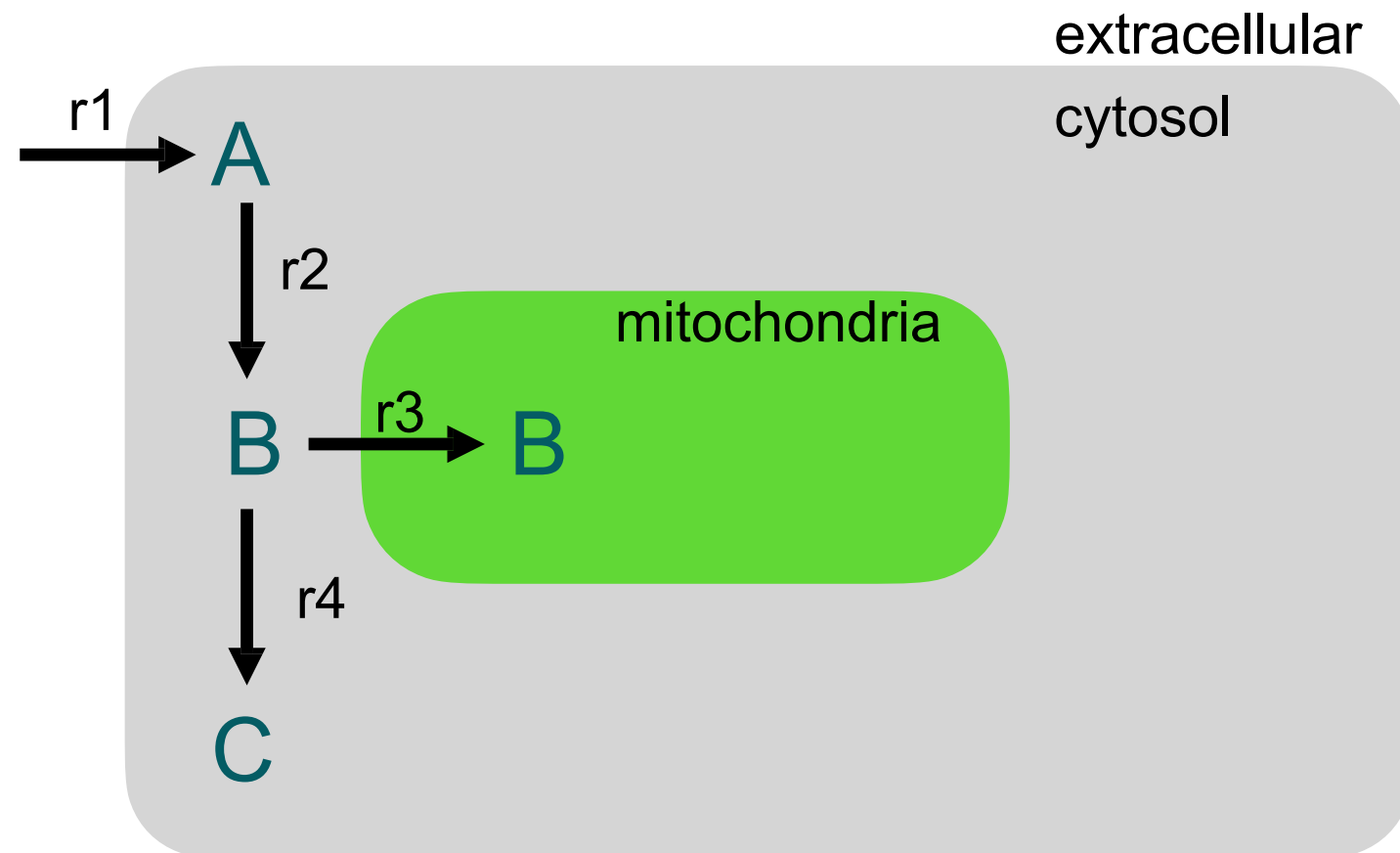
Kinetics

Thermodynamic constraints

Flux Balance Analysis (FBA)

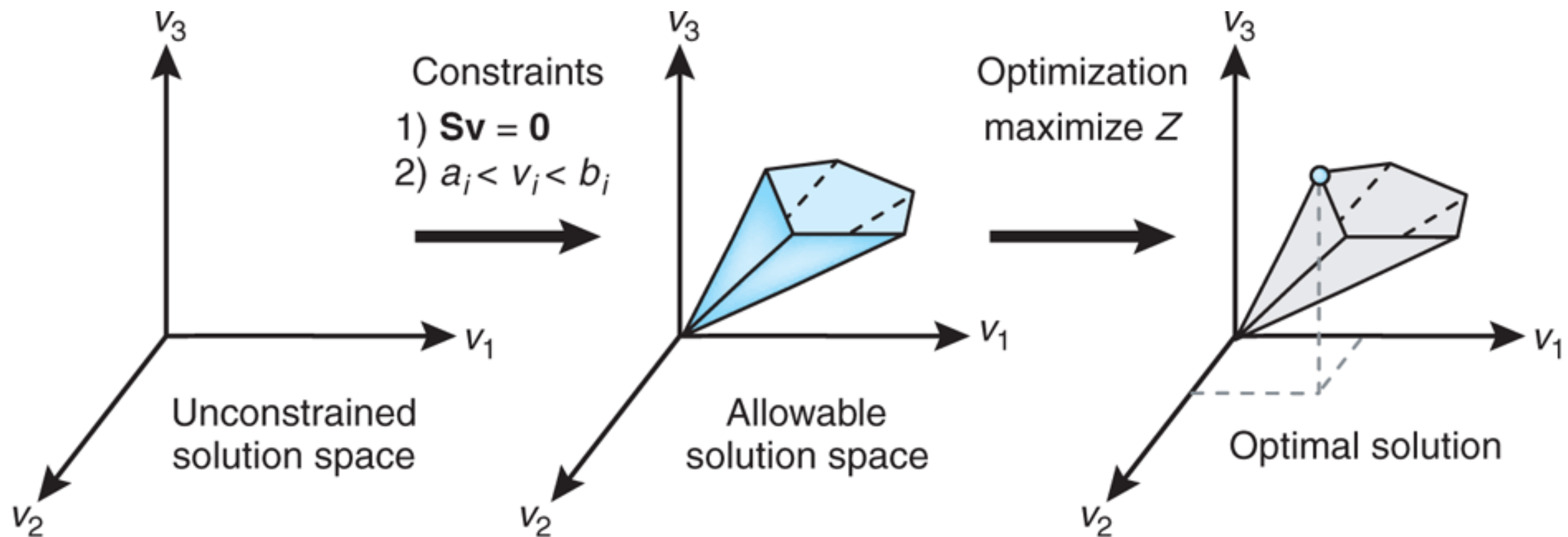
Models account for compartments:

- Exchange reactions
- Intracellular compartments



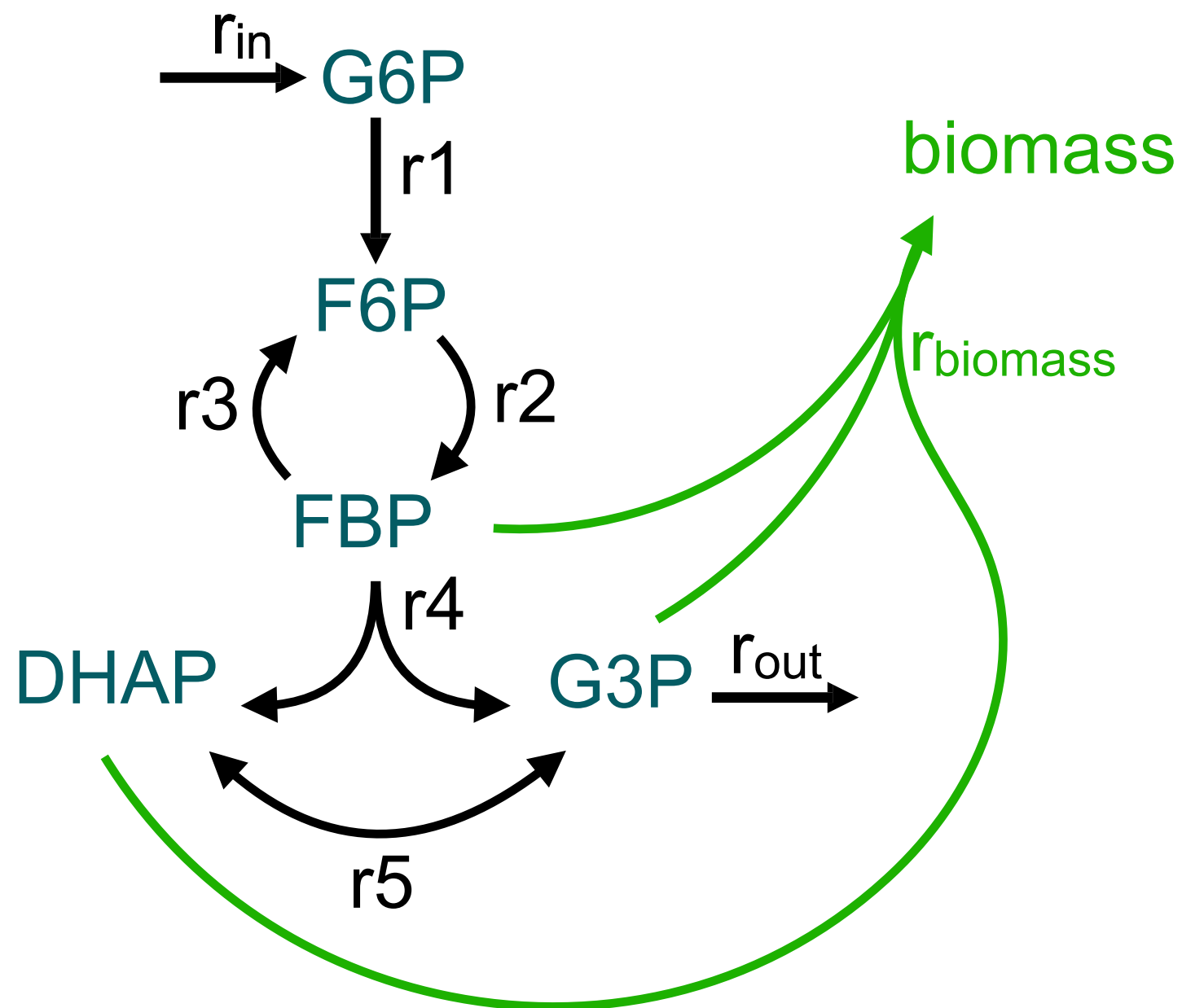
Flux Balance Analysis (FBA)

Since the problem is still **under-defined**, FBA uses linear **optimization** to identify a solution that maximizes (or minimizes) some **objective**

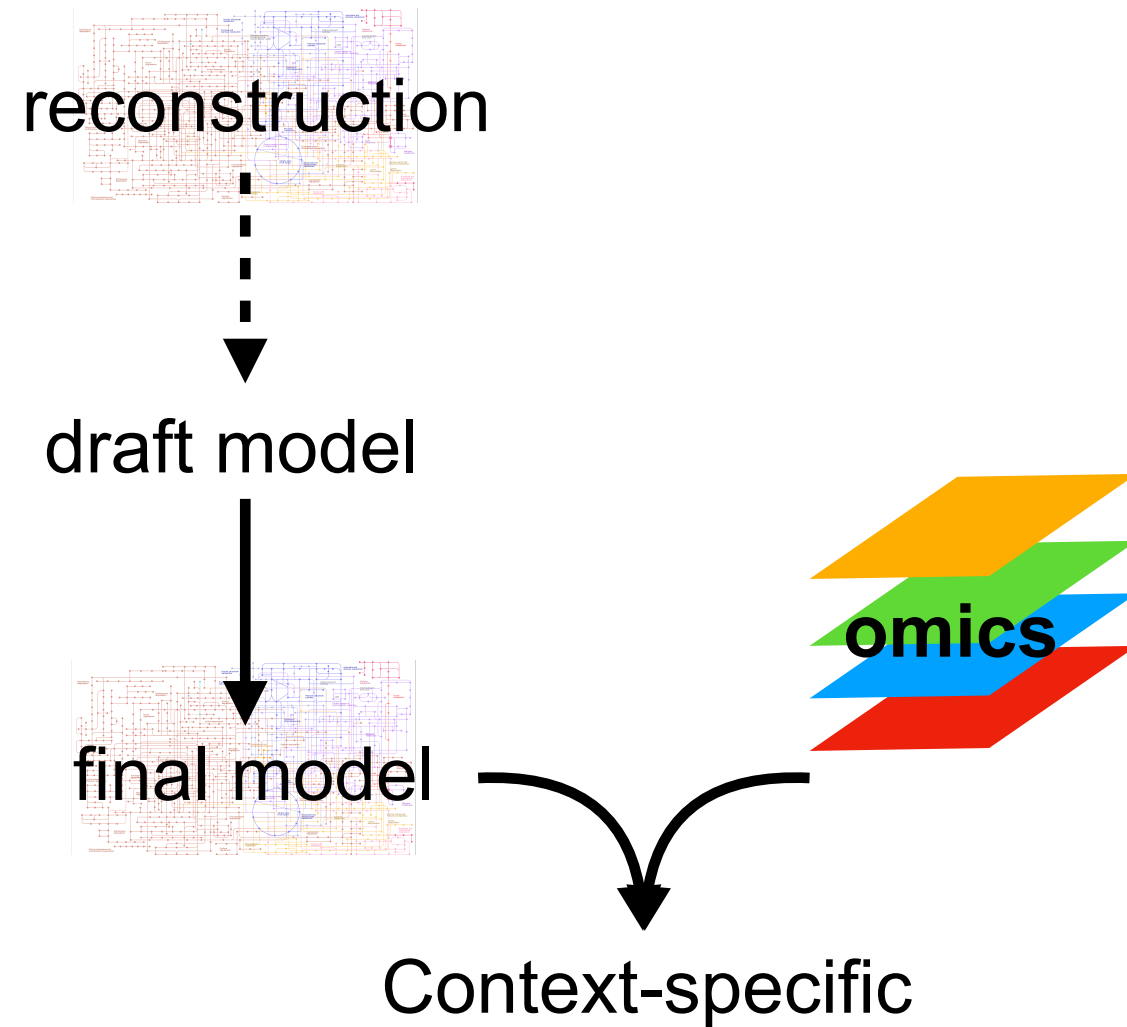


Flux Balance Analysis (FBA)

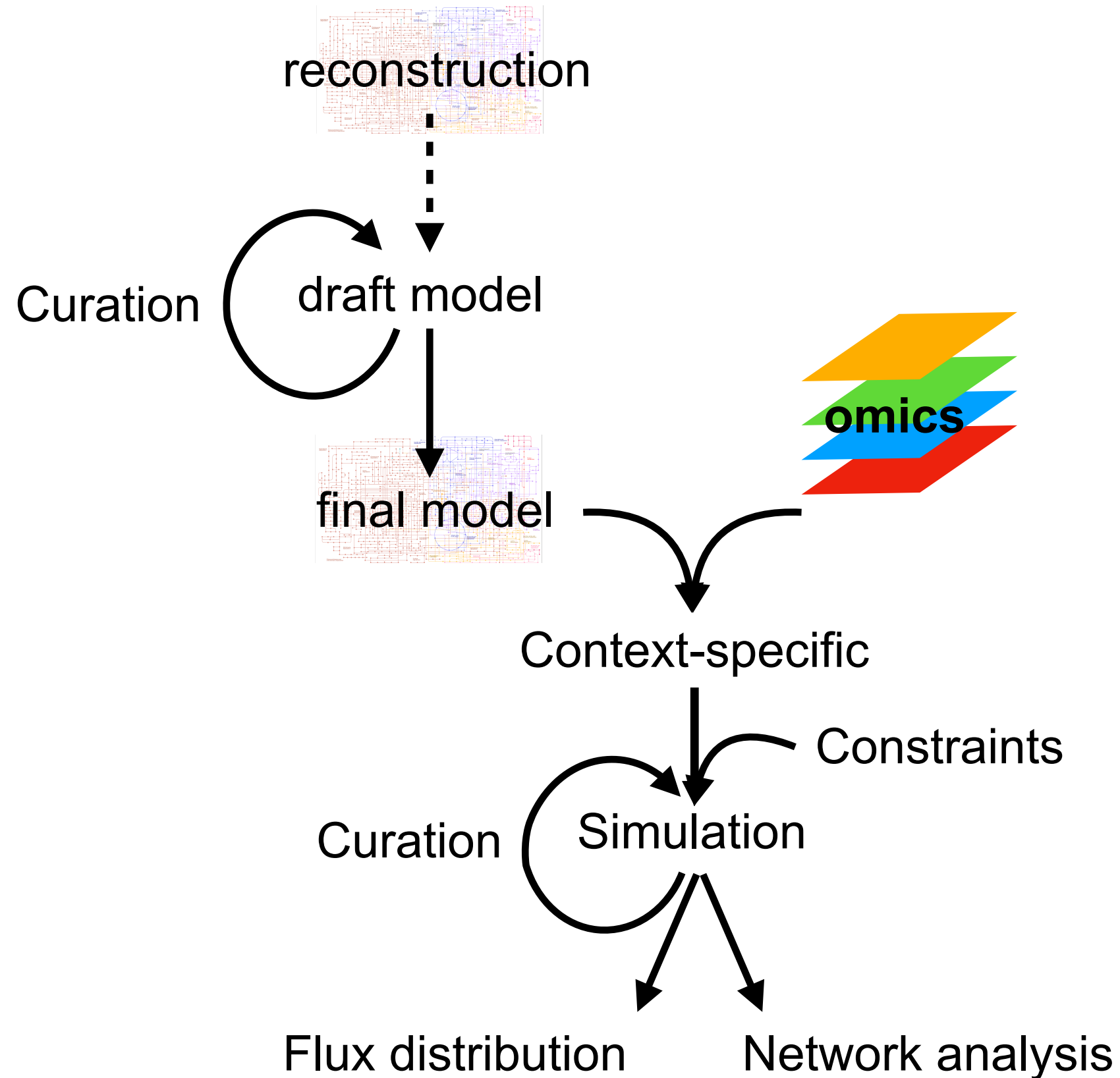
Objective function (i.e. optimisation objective) is often:
maximise an artificial “**biomass**” reaction or **ATP production**



Approach for analysis



Approach for analysis



Exercise: COBRApy



COBRApy (COnstraint-Based Reconstruction and Analysis) facilitates the use of GEMs in python

The exercise will walk through the basics of GEM structure, functionality, FBA, and gene knockouts.

Acknowledgements: Jonathan Robinson