

GEMs

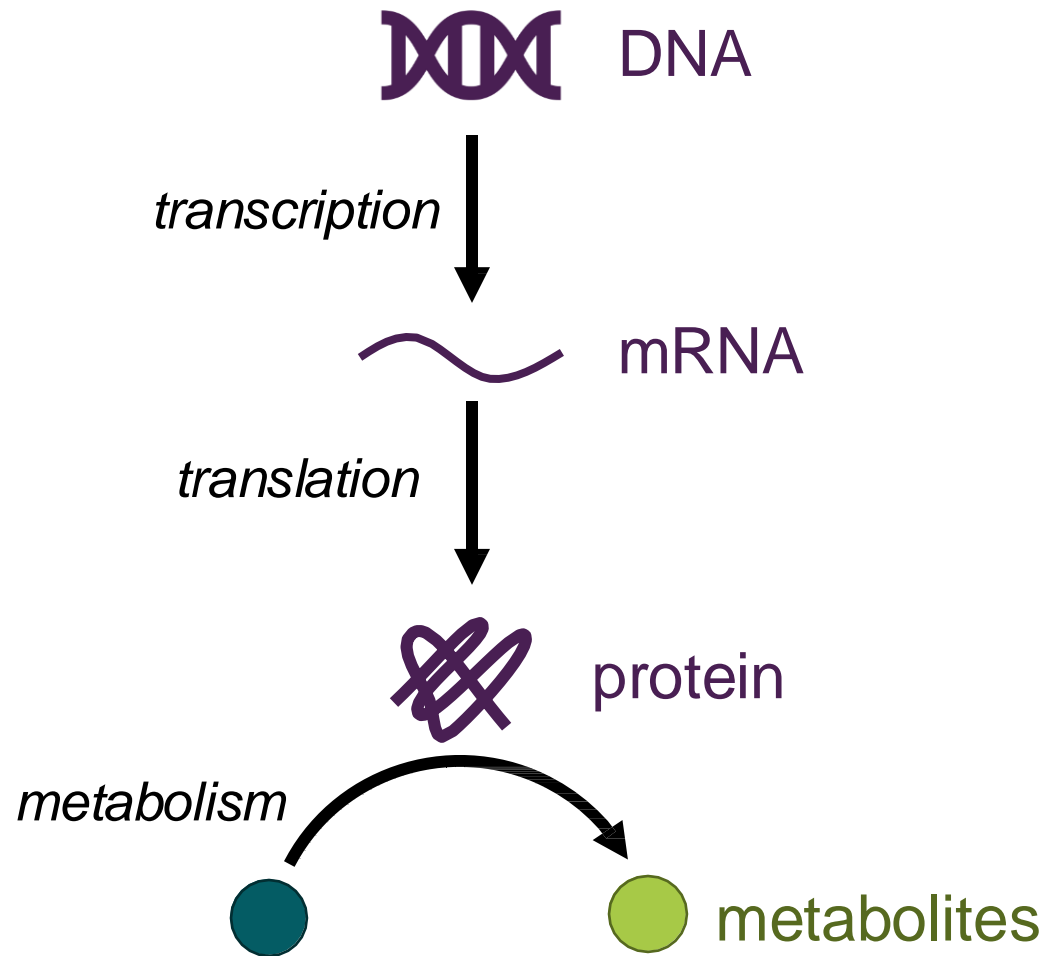
Structure & Simulation

NBIS Omics Integration and Systems Biology workshop
Oct 2024, Lund University

Rasool Saghaleyni

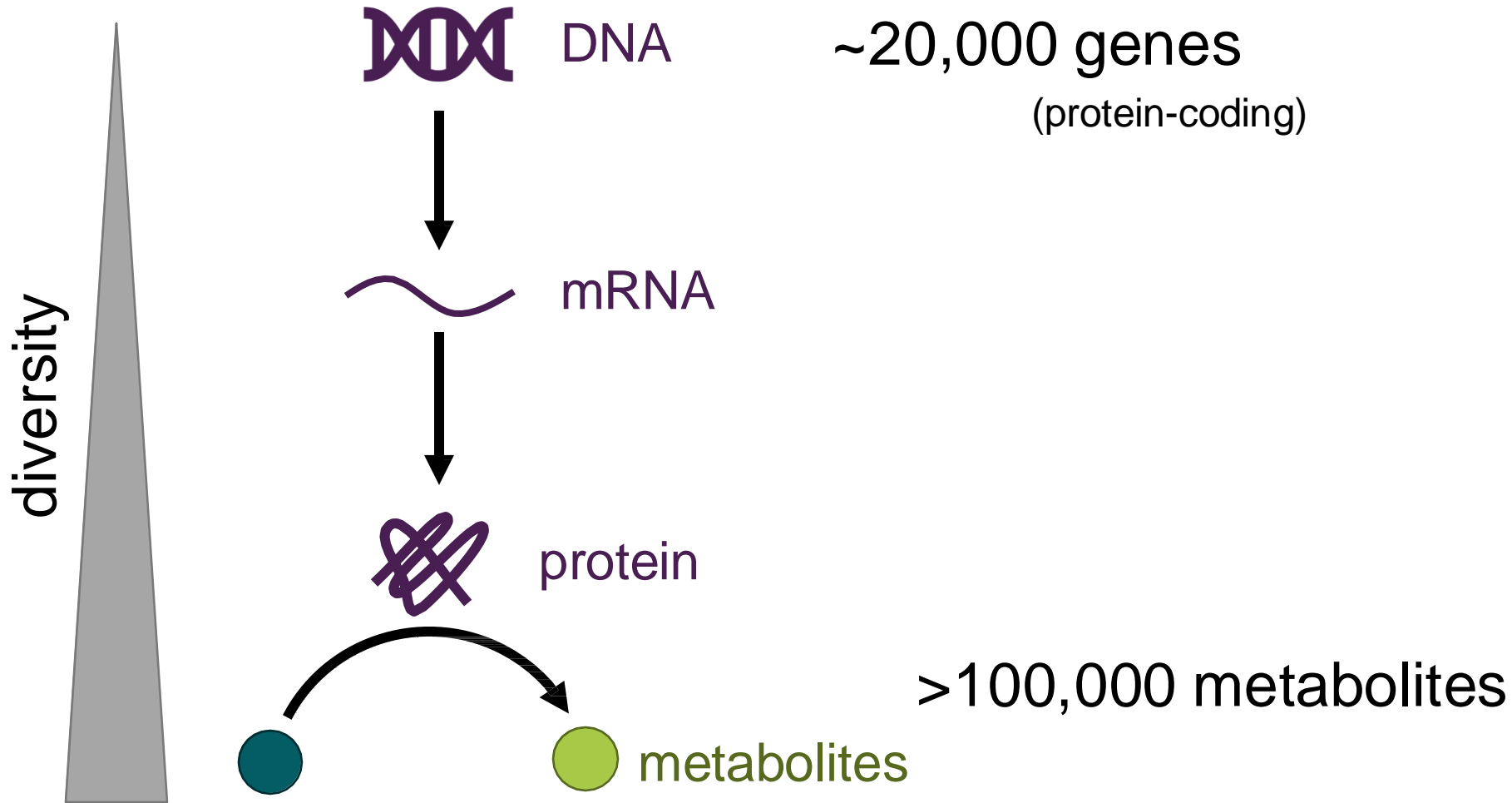
National Bioinformatics Infrastructure Sweden (NBIS)
Science for Life Laboratory (SciLifeLab)
Chalmers University of Technology
rasool.saghaleyni@scilifelab.se

Background

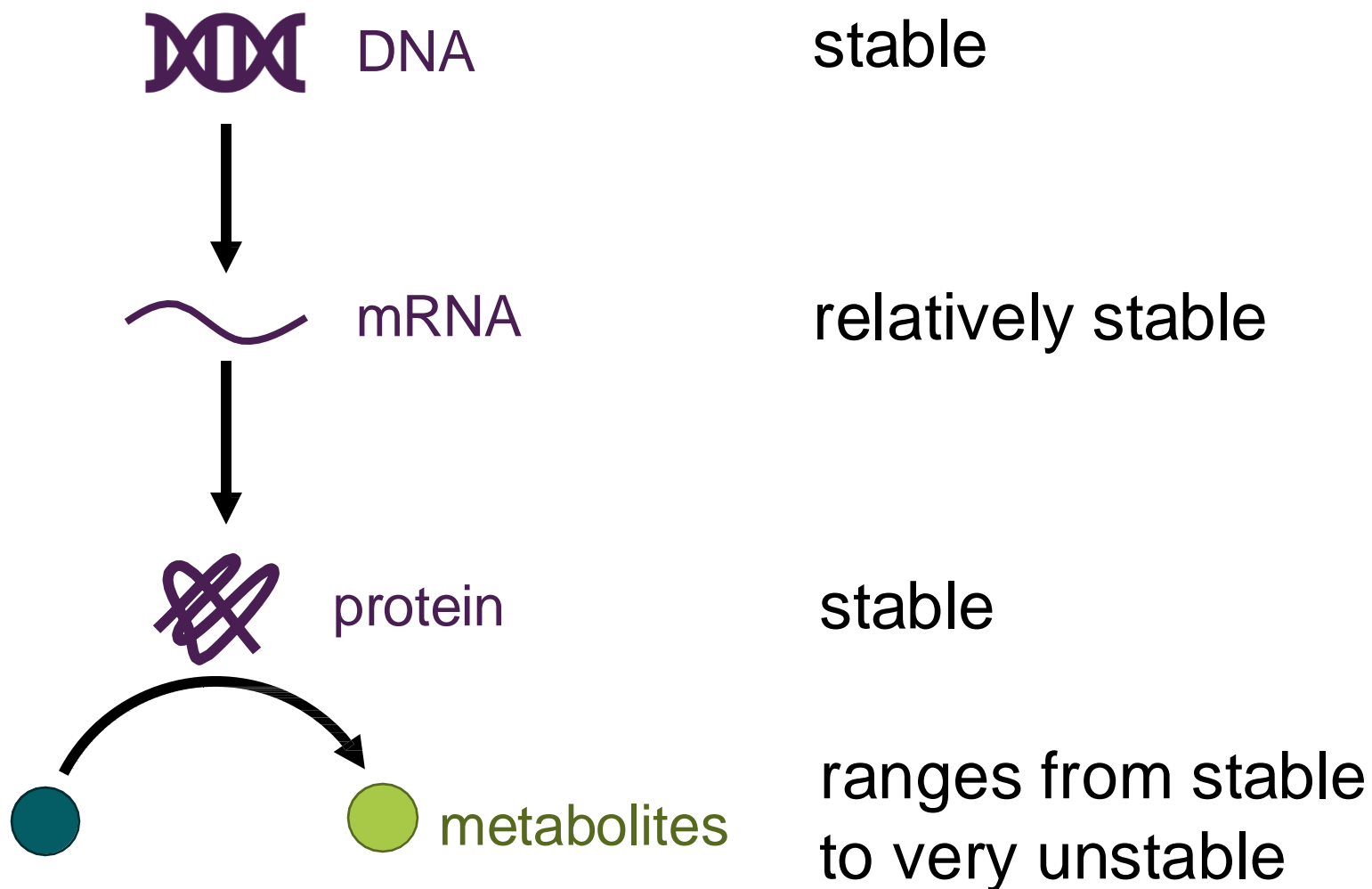


Metabolism provides the **energy** and **building blocks** necessary to sustain life.

Background



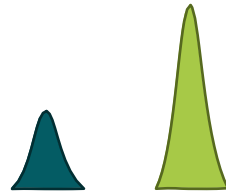
Background



Background



We can generally measure
metabolite concentrations



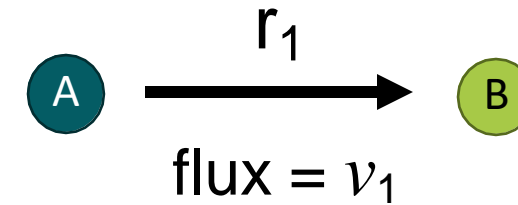
...but what is often important is
the flow or **flux** of metabolites
through the reactions.



Background

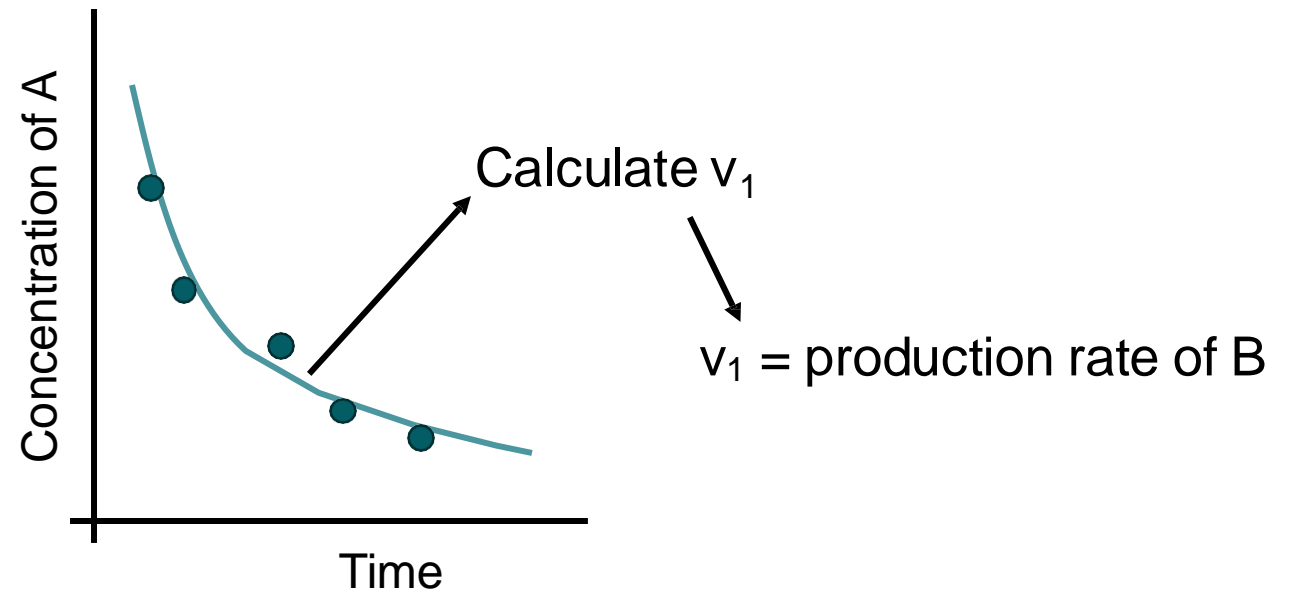


Assume that we want to know the production rate of **B**, but can only measure the concentration of **A**

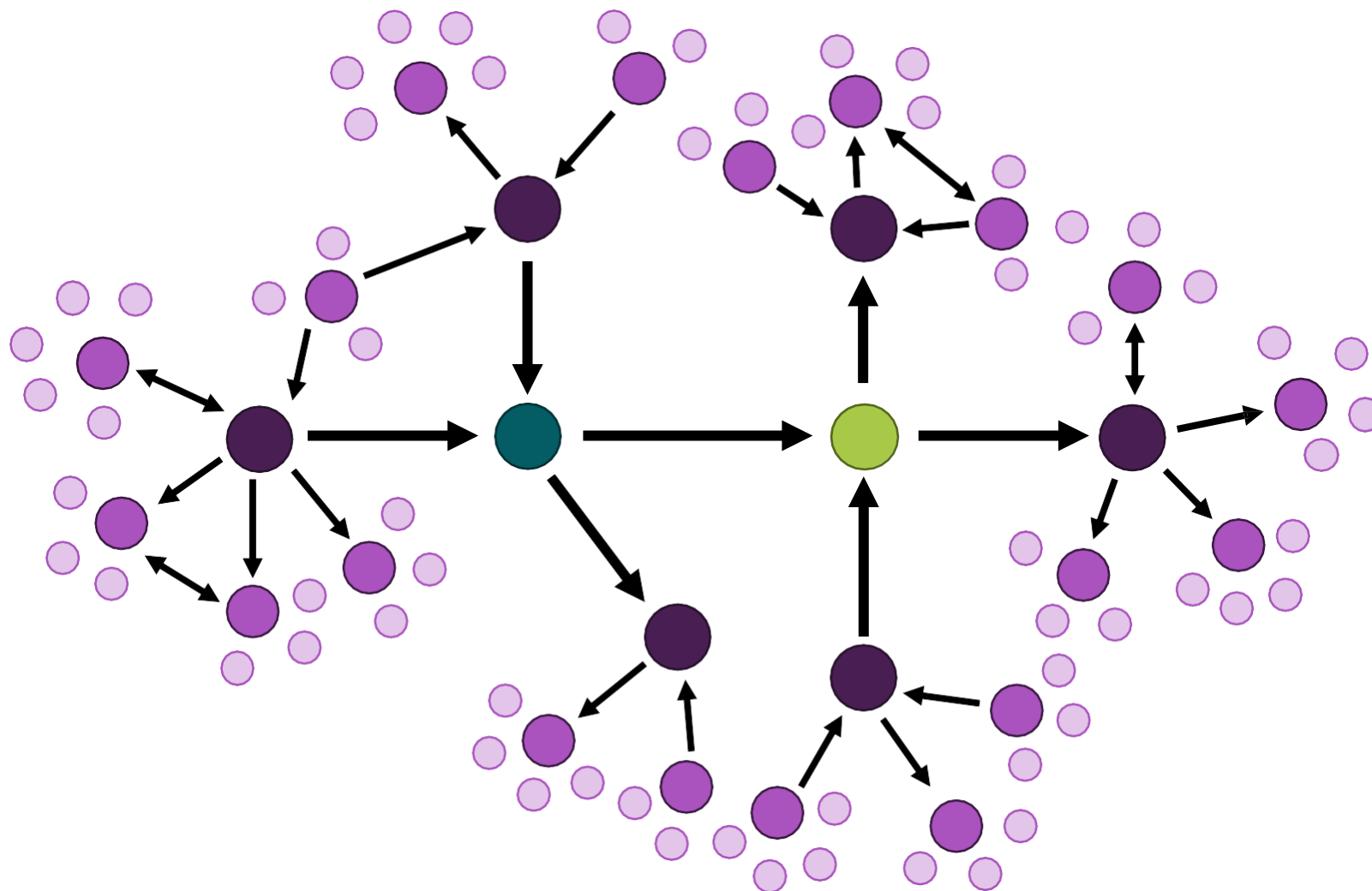


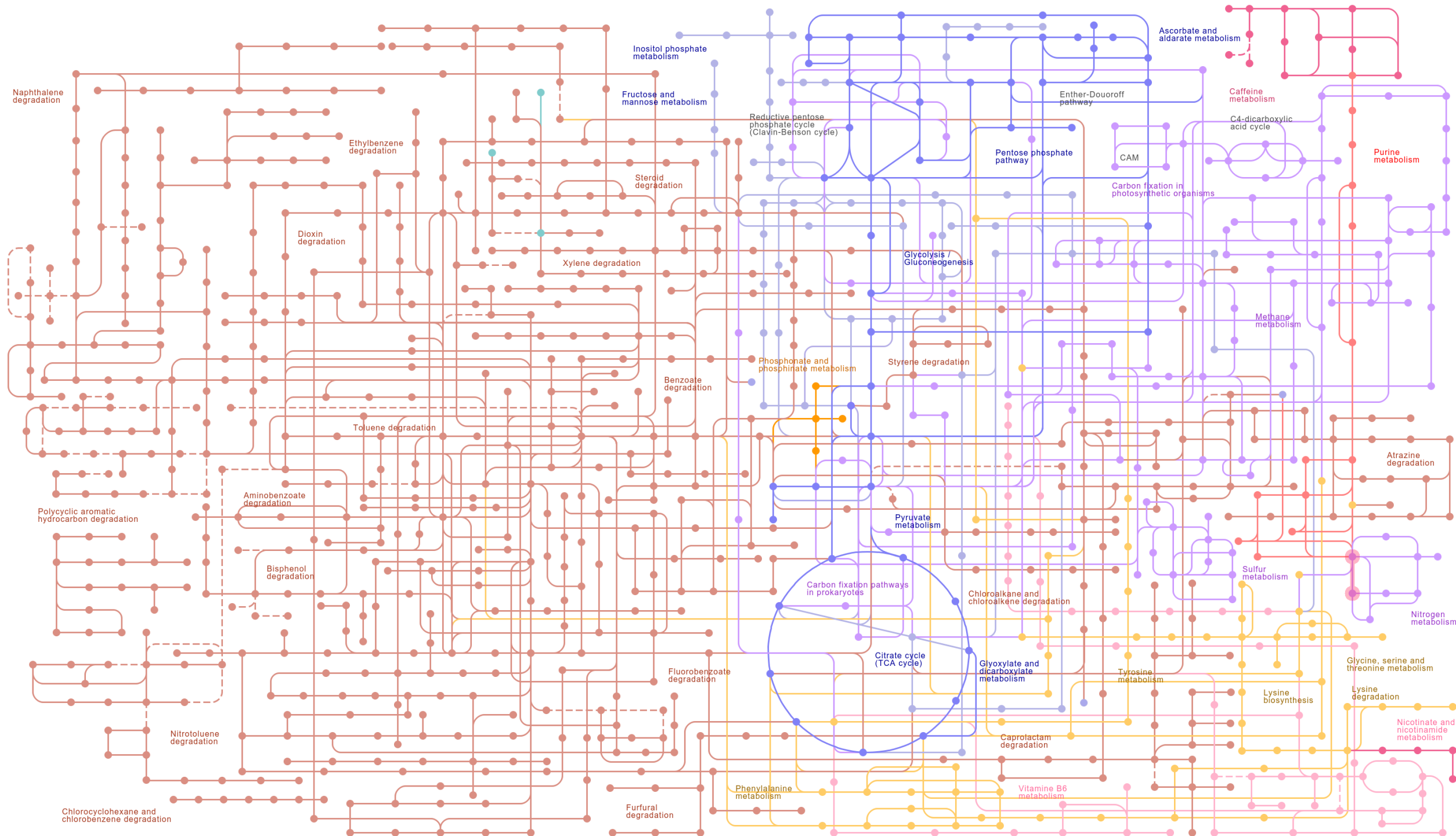
$$\frac{dA}{dt} = -v_1$$

$$\frac{dB}{dt} = v_1$$

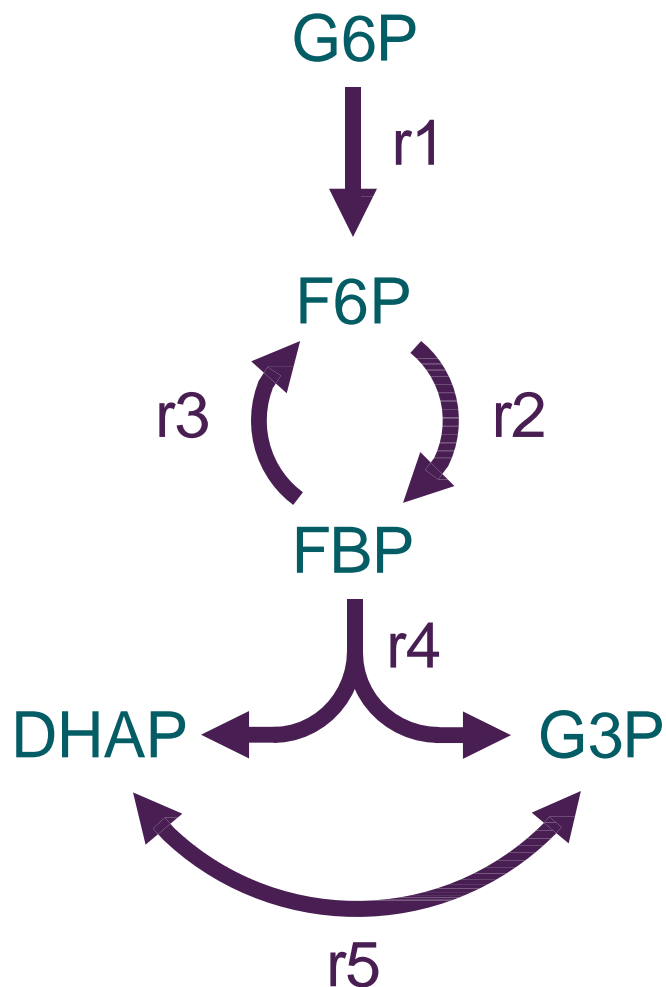


Background





The Stoichiometric Matrix

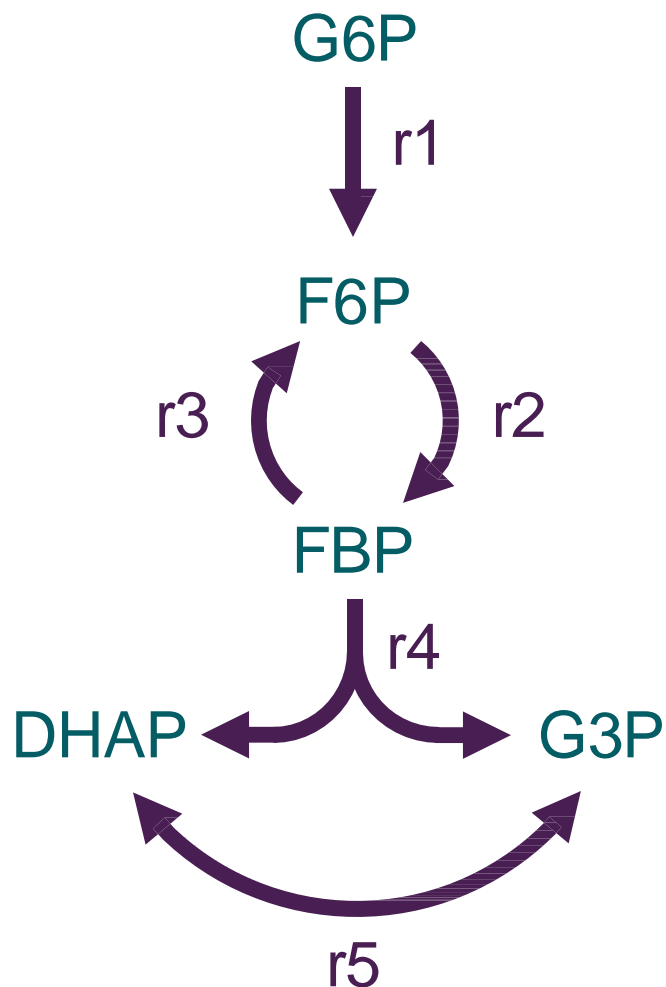


Metabolites

Reactions

	r1
G6P	-1
F6P	1
FBP	0
DHAP	0
G3P	0

The Stoichiometric Matrix

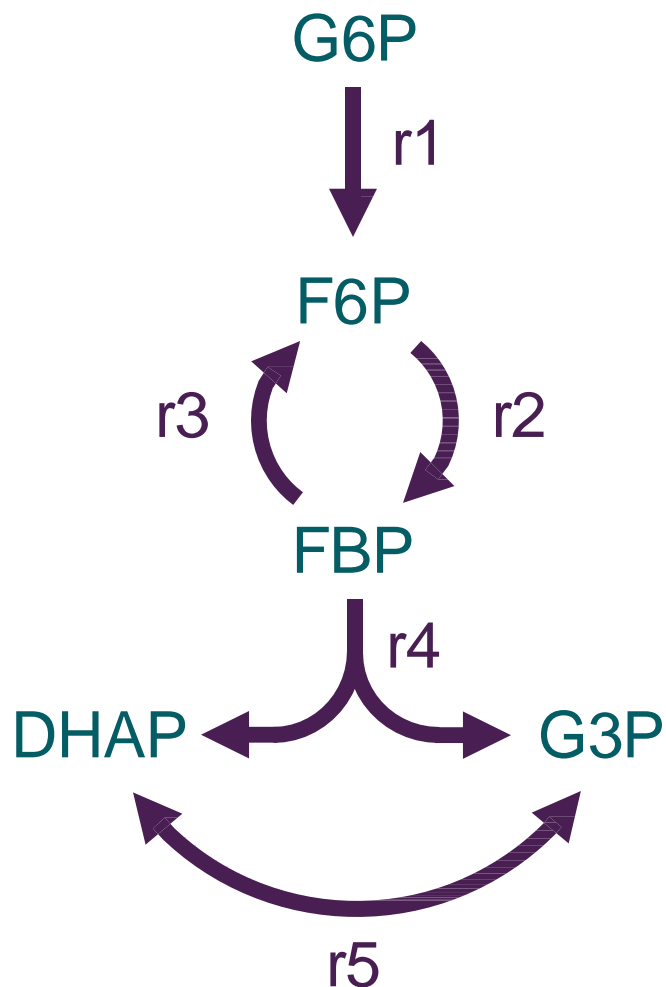


Metabolites

Reactions

	r1	r2
G6P	-1	0
F6P	1	-1
FBP	0	1
DHAP	0	0
G3P	0	0

The Stoichiometric Matrix



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
...

...
Other IDs
Name

KEGG ID	Compartment	Name	Symbol	r1	r2	r3	r4	r5	Symbol
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 ID)	Transcript IDs	GO Terms	Orthologs
					GPI	P06744			
					<i>n/a</i>				
					FBP1, FBP2	P09467, O00757			
					ALDOA, ALDOB, ALDOC	P04075, P05062, P09972			
					TPI1	P60174			

Symbol	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

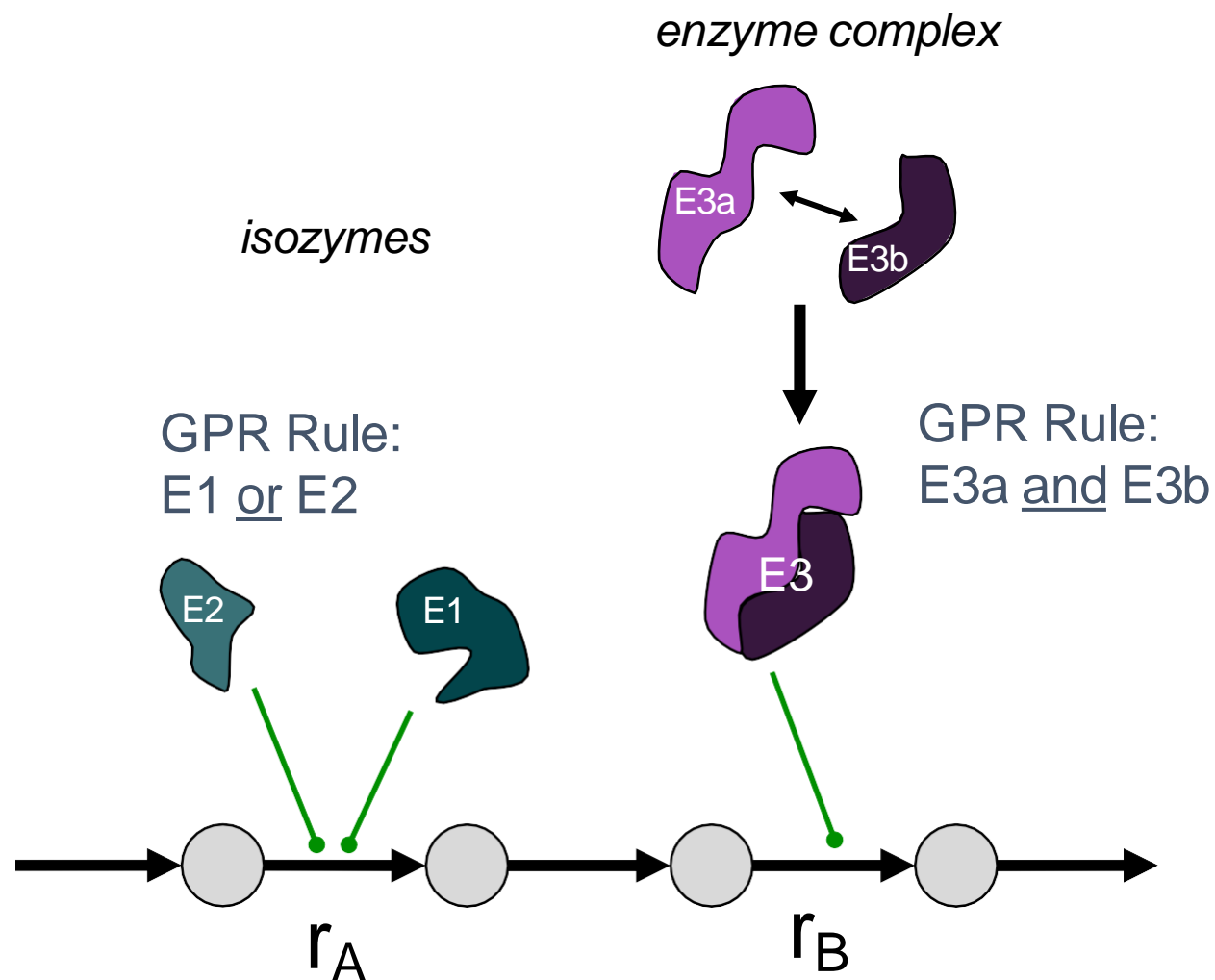
Reactions are linked to genes that encode the enzymes that catalyze the reaction.

These associations are often called “gene-protein rules” (GPR rules)

GPR Rules

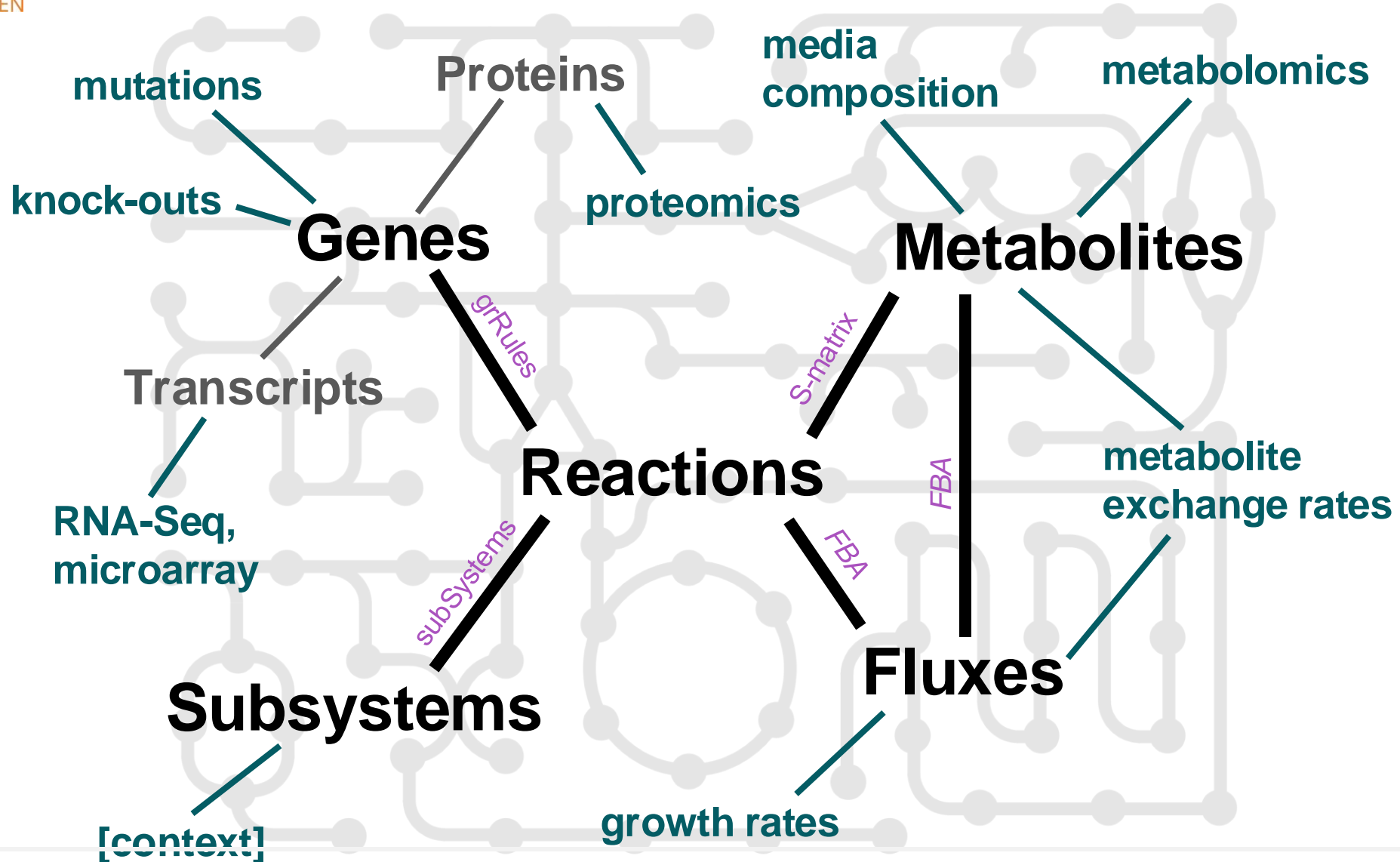


GPR Rules enable more accurate simulation of gene inactivation/knock-out

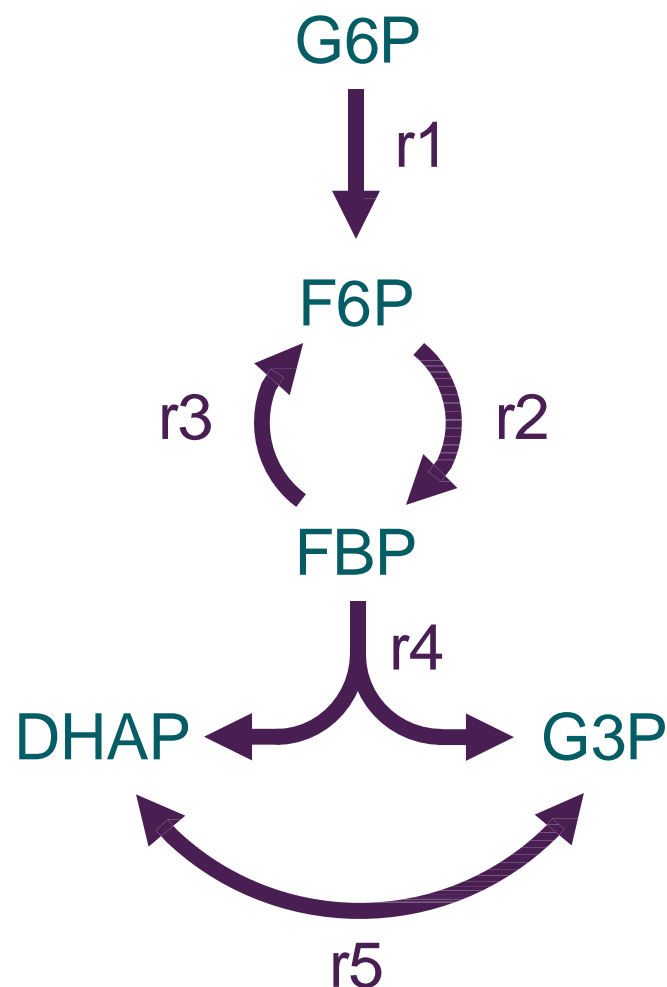


Knockout	Effect
E1	none
E2	none
E1 + E2	rA inactive
E3a	rB inactive
E3b	rB inactive
E3a + E3b	rB inactive

GEMs as an integrative tool



Flux Balance Analysis (FBA)



Metabolites

Reactions

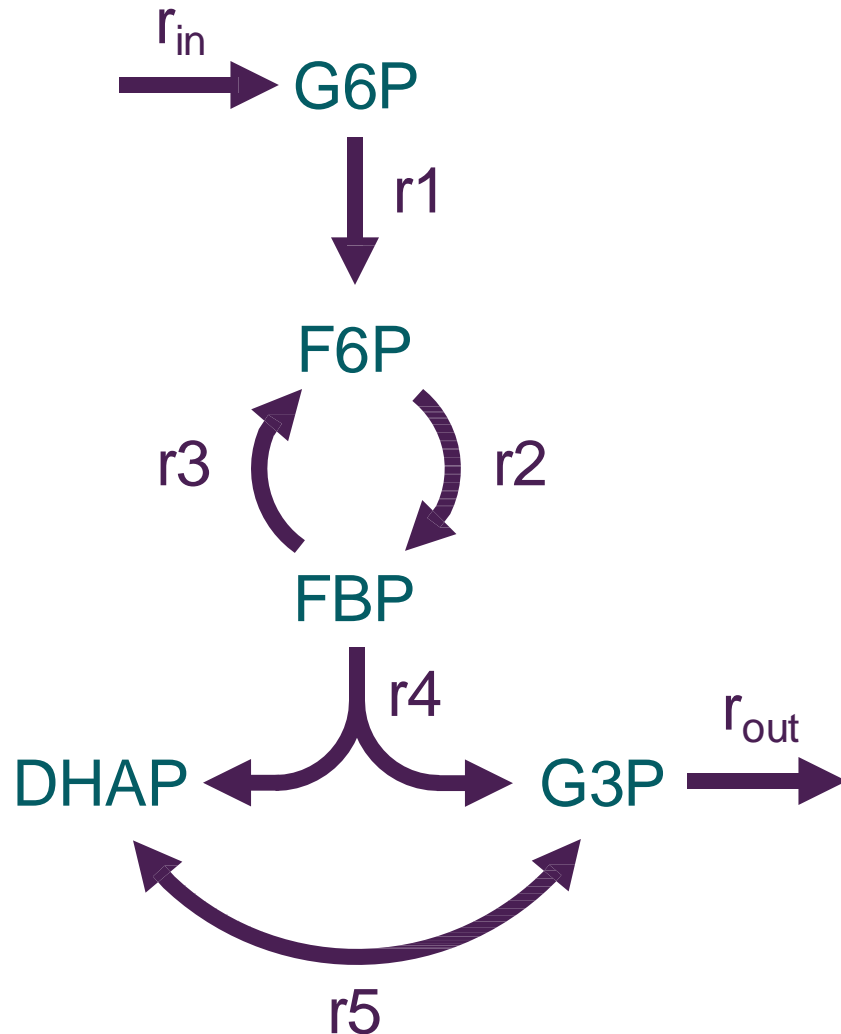
	r1	r2	r3	r4	r5
G6P	-1	0	0	0	0
F6P	1	-1	1	0	0
FBP	0	1	-1	-1	0
DHAP	0	0	0	1	-1
G3P	0	0	0	1	1

Flux Balance Analysis (FBA)



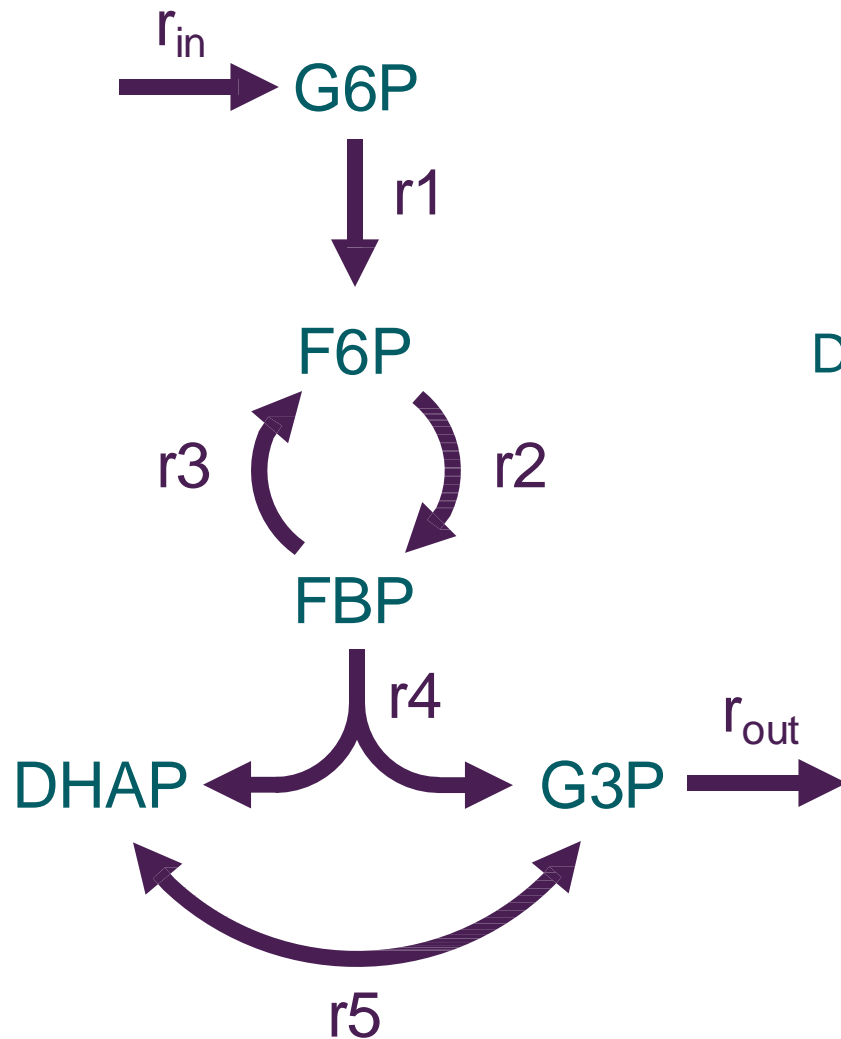
FBA seeks to calculate the reaction **fluxes** (v) of a network

The calculation is based on the **conservation of mass**: it cannot be created or destroyed



$$\frac{dX}{dt} = v_{produce} - v_{consume}$$

Flux Balance Analysis (FBA)



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

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

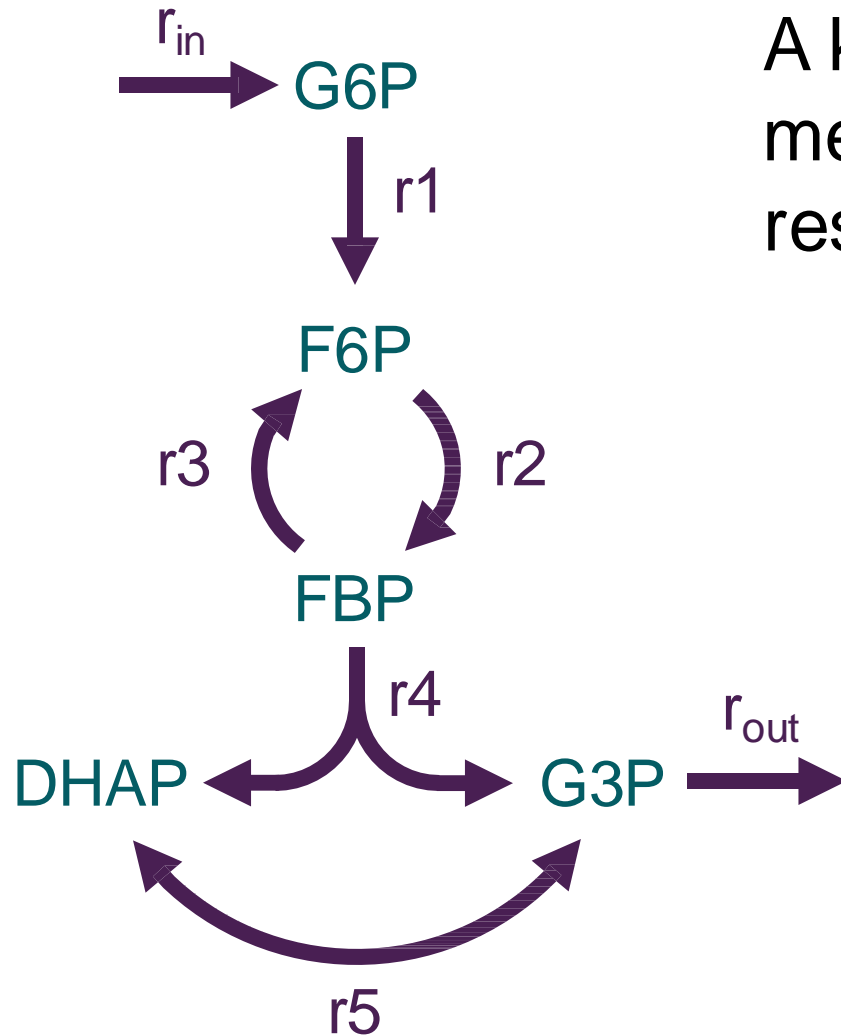
$$\frac{d[\text{G3P}]}{dt} = v_4 + v_5 - v_{out}$$

Flux Balance Analysis (FBA)



A key assumption to FBA is **steady state**: metabolite concentrations are **constant** with respect to time!

$$\frac{dX}{dt} = v_{produce} - v_{consume} = 0$$



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



Flux Balance Analysis (FBA)

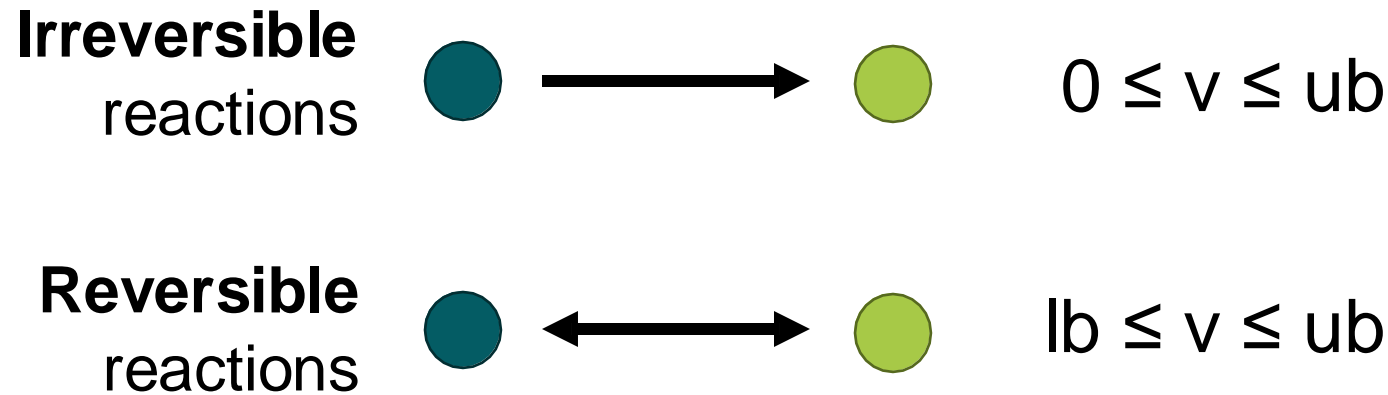
$$\begin{array}{c}
 \text{G6P} \\
 \text{F6P} \\
 \text{FBP} \\
 \text{DHAP} \\
 \text{G3P}
 \end{array}
 \begin{array}{c}
 r_1 \quad r_2 \quad r_3 \quad r_4 \quad r_5 \quad r_{in} \quad r_{out} \\
 \begin{pmatrix}
 -1 & 0 & 0 & 0 & 0 & 1 & 0 \\
 1 & -1 & 1 & 0 & 0 & 0 & 0 \\
 0 & 1 & -1 & -1 & 0 & 0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

$$S \cdot v = 0$$

Flux Balance Analysis (FBA)



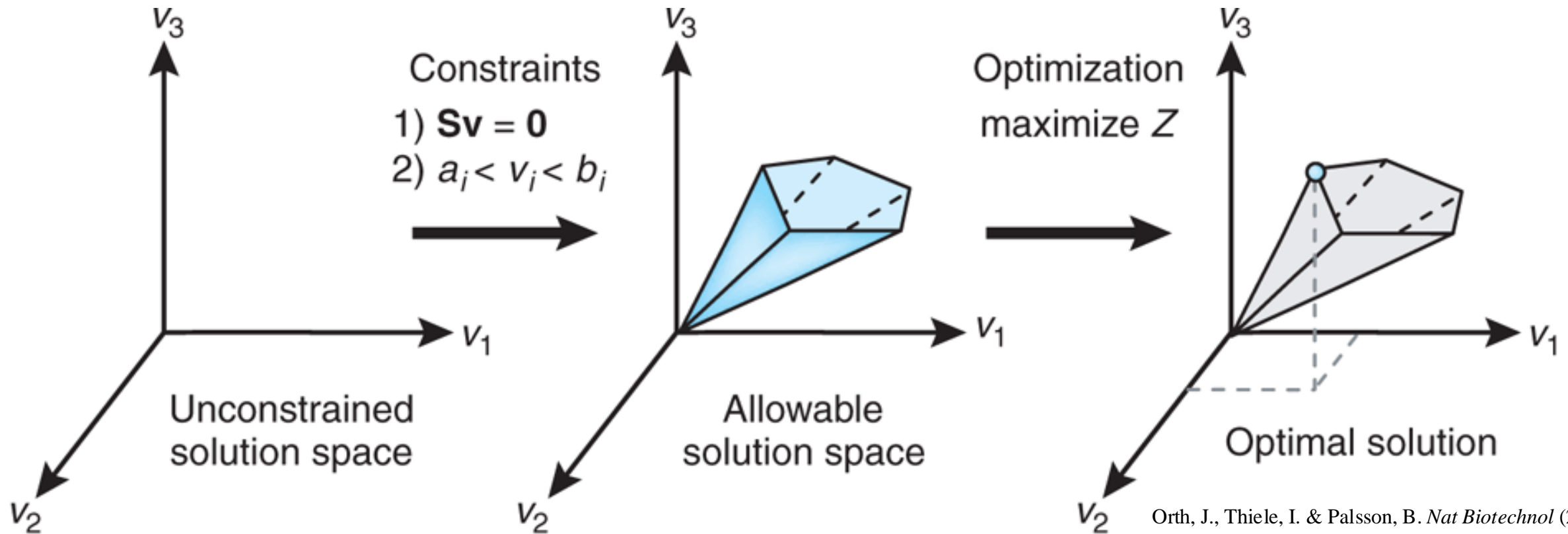
We can further constrain the solution space by limiting reaction fluxes based on their reversibility:



Flux Balance Analysis (FBA)

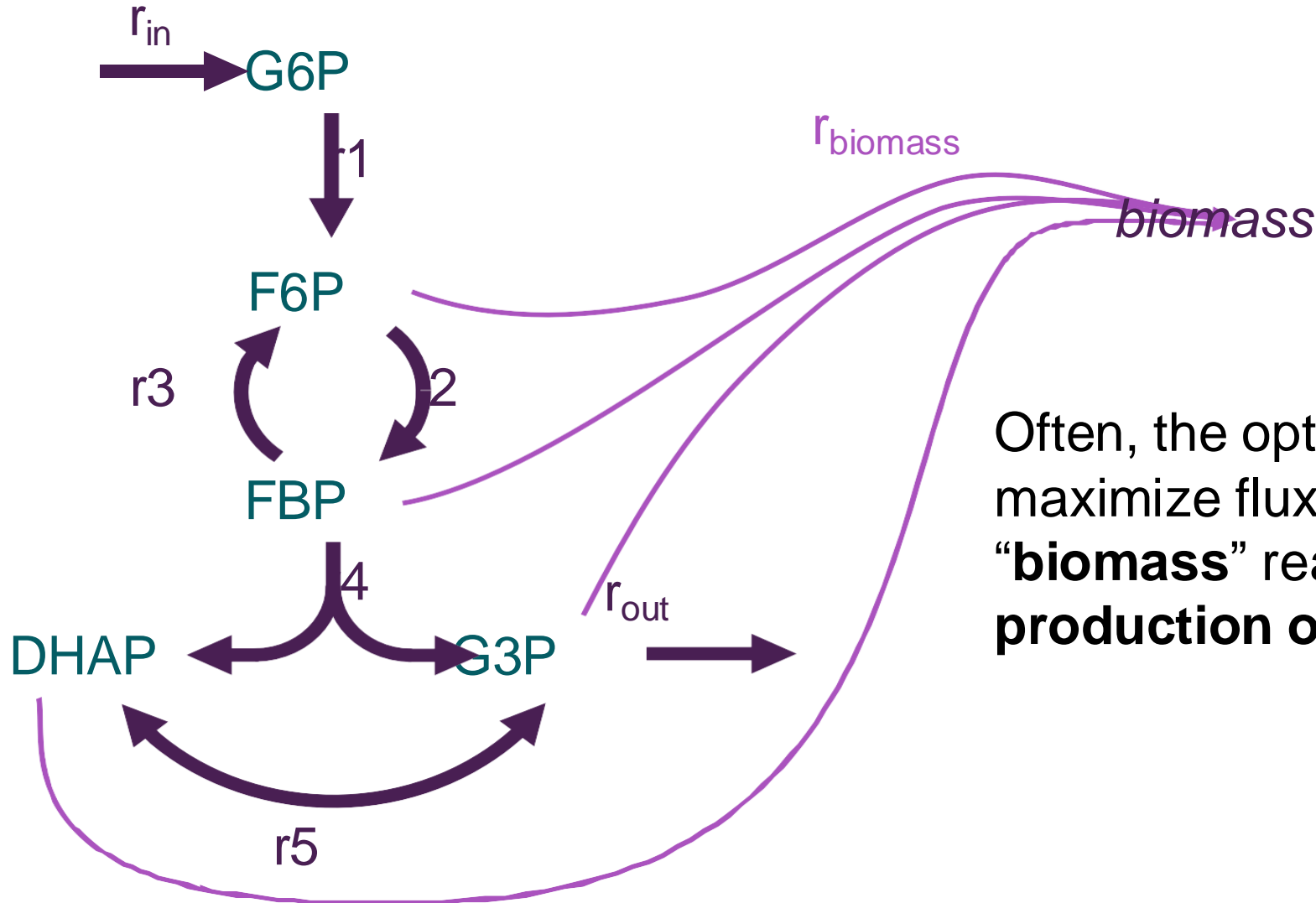


Since the problem is still **under-defined**, FBA uses linear **optimization** to identify a solution that maximizes (or minimizes) some **objective**



Orth, J., Thiele, I. & Palsson, B. *Nat Biotechnol* (2010).

Flux Balance Analysis (FBA)



Often, the optimization objective is to maximize flux through an artificial “**biomass**” reaction, or to maximize production of **ATP**.

Exercise: COBRApy



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

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