BIOCHEMICAL REACTION NETWORKS

Quantifying Metabolic Flux Analysis and Flux Balance Analysis

Dr. Neil Priharto, M.T.

Neil Priharto

+62 877 3039 9192 neil@sith.itb.ac.id https://www.linkedin.com/in/neil-priharto-30b20853/ https://orcid.org/0000-0003-3123-0688

My research interest is in bio-refinery of renewable chemicals and next generation biofuels from biomass

> EDUCATION

UNDERGRADUATE
Microbiology, ITB— 2005 - 2009

GRADUATE
Chemical Engineering, ITB— 2005 – 2009

DOCTORATE
Department of Green Chemistry, Ghent University — 2014 - 2021

Chemicals and fuels from lignin-rich digested stillage and microalgae via thermochemical conversion processes

> PROJECTS




> PUBLICATIONS

Priharto, Neil., Ghysels, Stef., Pala, Mehmet., Opsomer, Wim., Ronsse, Frederik., Yildiz, Guray., Heeres, Hero Jan., Deuss, Peter J., and Prins, Wolter. 2020. ***"Ex-Situ Catalytic Fast Pyrolysis of Lignin-Rich Digested Stillage over Na/ZSM-5, H/ZSM-5, and Fe/ZSM-5".*** Energy Fuels, 34, 10, 12710–12723 https://doi.org/10.1021/acs.energyfuels.0c02390.

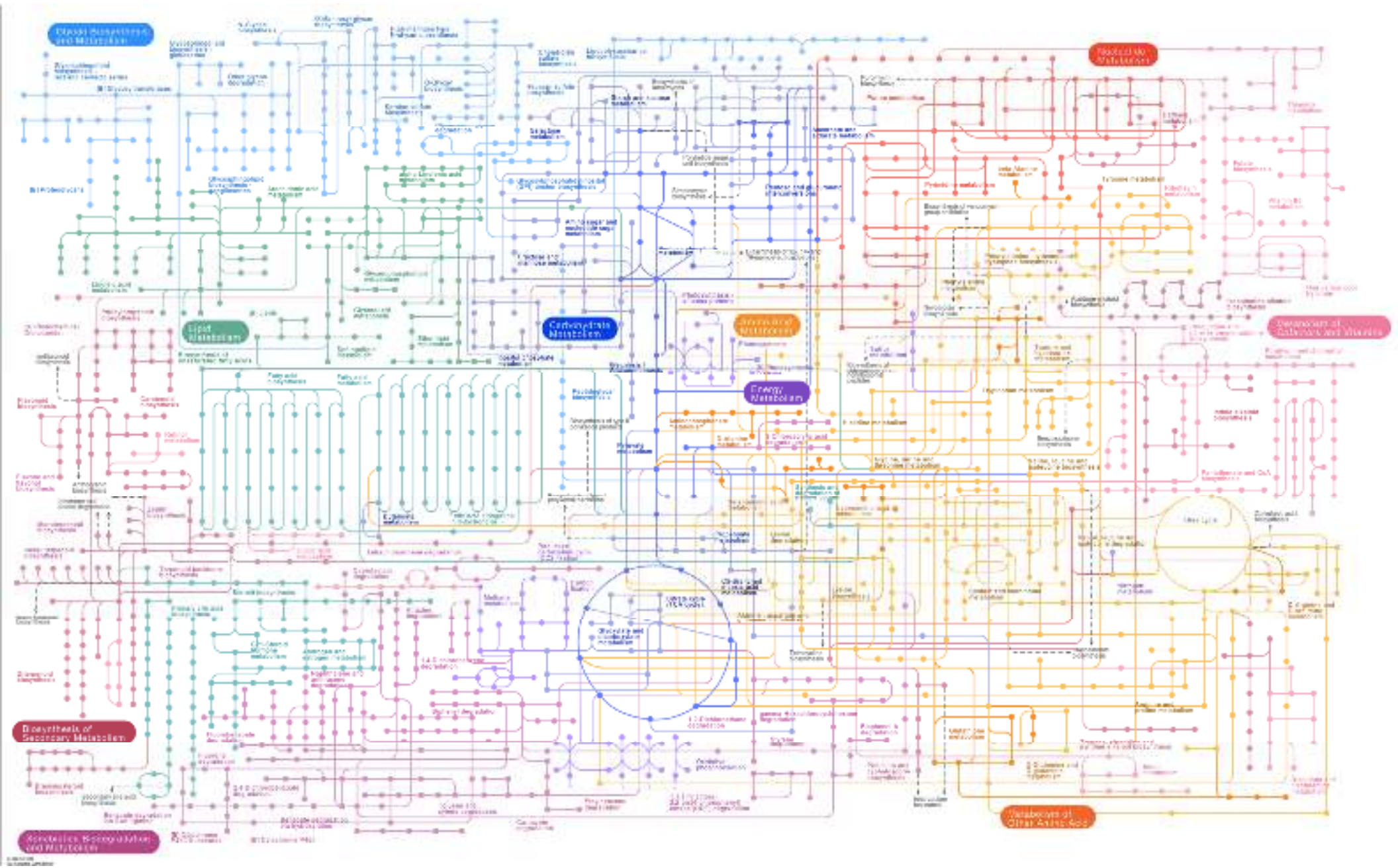
Priharto, Neil, Ronsse, Frederik., Prins, Wolter., Hita, Idoia., Deuss, Peter J., and Heeres, Hero Jan. 2019. ***"Hydrotreatment of Pyrolysis Liquids Derived from Second-Generation Bioethanol Production Residues over NiMo and CoMo Catalysts."*** Biomass and Bioenergy 126 (March): 84–93. https://doi.org/10.1016/j.biombioe.2019.05.005.

Priharto, Neil, Ronsse, Frederik., Prins, Wolter., Carleer, Robert., and Heeres, Hero Jan. 2020. ***"Experimental Studies on a Two-Step Fast Pyrolysis-Catalytic Hydrotreatment Process for Hydrocarbons from Microalgae (Nannochloropsis Gaditana and Scenedesmus Almeriensis)."*** Fuel Processing Technology 206 (February): 106466. https://doi.org/10.1016/j.fuproc.2020.106466.

DEFINITION

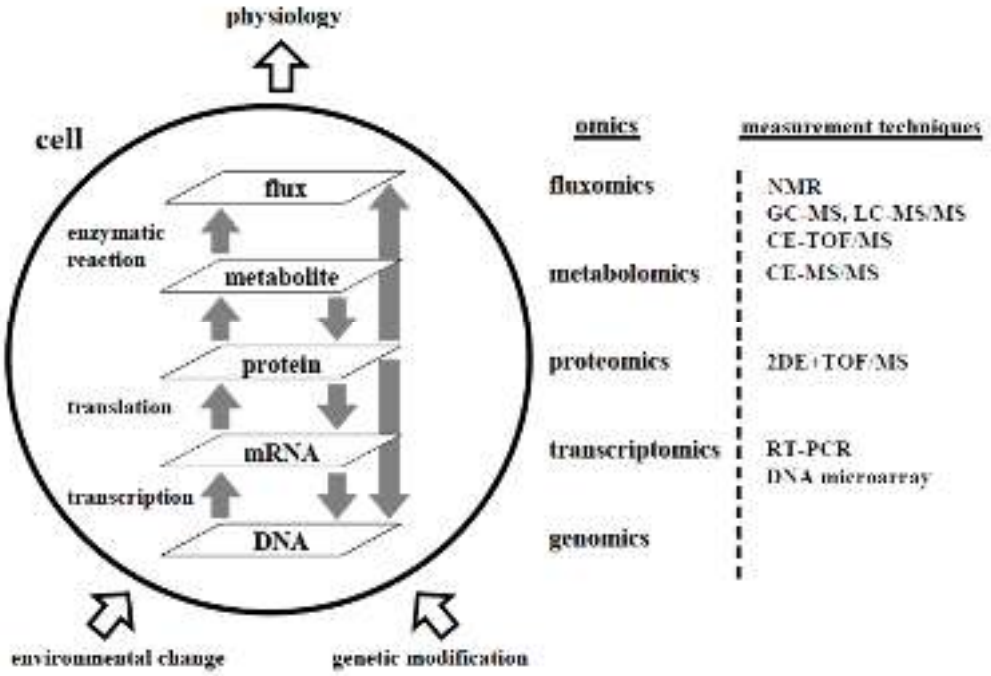
Metabolic Pathway

- In biochemistry, metabolic pathways are series of chemical reactions occurring within a cell.
- In each pathway, a principal chemical is modified by a series of chemical reactions.
- Enzymes catalyze these reactions
- Because of the many chemicals (a.k.a. metabolites) that may be involved, metabolic pathways can be quite elaborate.
- This collection of pathways is called the metabolic network.
- Pathways are important to the maintenance of homeostasis within an organism. Catabolic (break-down) and Anabolic (synthesis) pathways often work interdependently to create new biomolecules as the final end products.


Glycan Biosynthesis and Metabolism
Lipid Metabolism
Carbohydrate Metabolism
Nucleotide Metabolism
Energy Metabolism
Metabolism of Cofactors and Vitamins
Biosynthesis of Secondary Metabolites
Xenobiotics Biodegradation and Metabolism
Metabolism of Other Amino Acids

DEFINITION

The most important information in understanding the complex metabolic control mechanism of the whole cell may be the metabolic flux distribution (Sauer, 2006), as this is the manifestation of gene and protein expressions and the concentrations of intracellular metabolites (Matsuoka & Shimizu, 2010a). The information of the metabolic flux distribution is quite useful for metabolic engineering (Stephanopoulos, 1999).



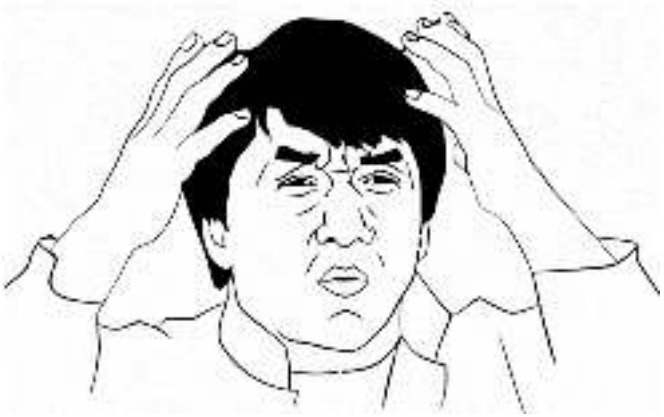
DEFINITION

- Accurate quantification of the magnitude of pathway fluxes in vivo is, therefore, an important goal of cell physiology and metabolic engineering, especially in the context of metabolite production, where the aim is to convert as much substrate as possible to useful products.
- A powerful methodology for the determination of metabolic pathway fluxes is Metabolic Flux Analysis (MFA), whereby intracellular fluxes are calculated by using a stoichiometric model for the major intracellular reactions and applying mass balance around intracellular metabolites

DEFINITION

- The final outcome of flux calculation is a **metabolic flux map** showing a diagram of the biochemical reactions included in the calculations along with an estimate of the steady state rate at which each reaction in the diagram occurs

- The real value of such metabolic flux maps lies in the **flux differences** that are observed when flux maps obtained with different strains or under different conditions are compared with one another



DEFINITION

What is flux?

❖In (Chemical) engineering and physic term :

mass flux (kg m^{-2} s^{-1}) is the rate of mass flow per unit area

The common symbols are *j*, *J*, *q*, *Q*, φ, or Φ

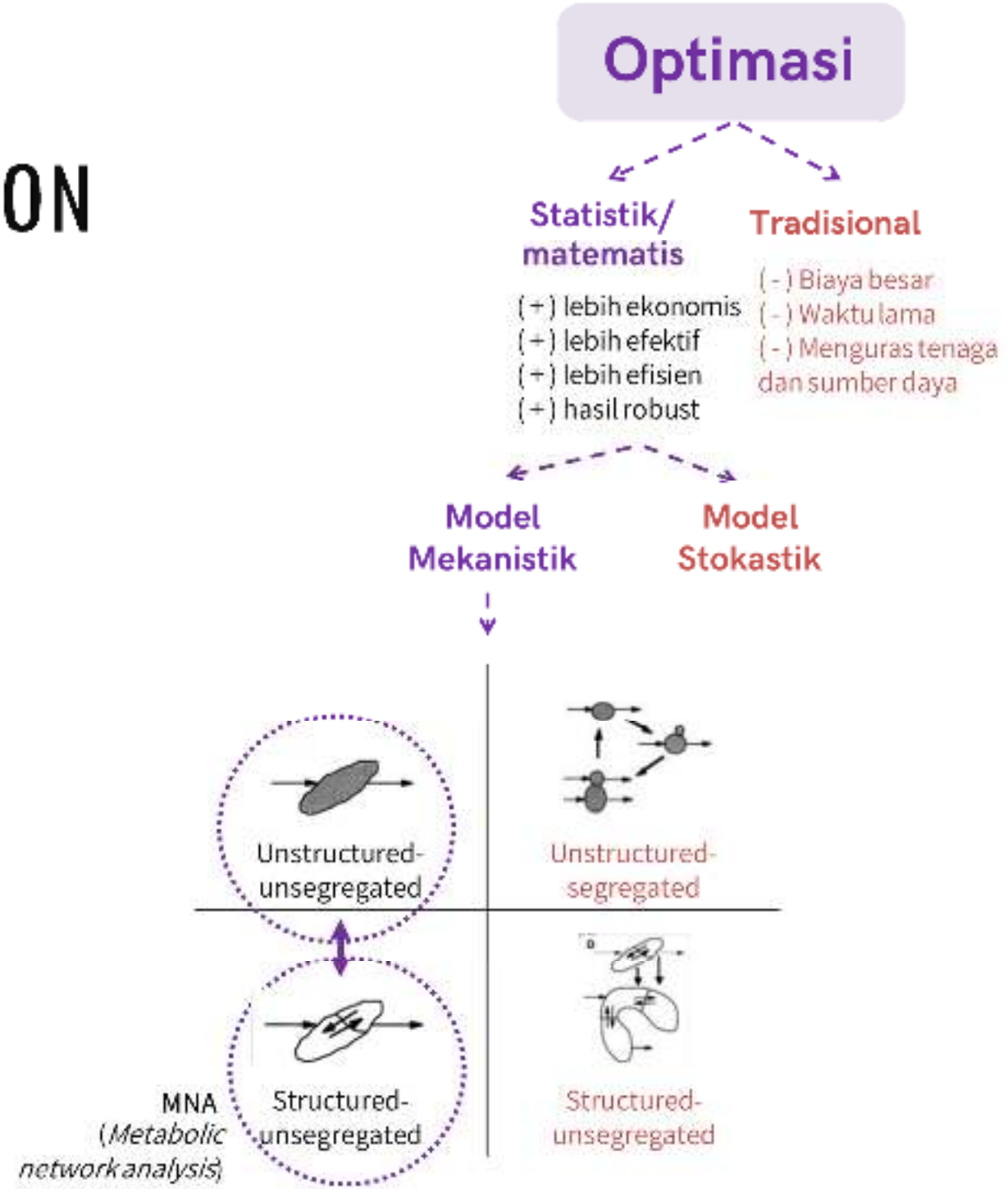
❖**Metabolic flux** is the rate of turnover of molecules through a metabolic pathway

The flux of the metabolites through each reaction (J) is the rate of the forward reaction (V_f), less that of the reverse reaction (V_r)

$$J = v_f - v_r$$

The *in vivo* enzymatic reaction rates (flux) cannot be directly measured.

DEFINITION


Optimasi
Statistik/ matematis
(+) lebih ekonomis
(+) lebih efektif
(+) lebih efisien
(+) hasil robust
Tradisional
(-) Biaya besar
(-) Waktu lama
(-) Menguras tenaga dan sumber daya
Model Mekanistik
Model Stokastik
Unstructured-unsegregated
Unstructured-segregated
Structured-unsegregated
Structured-unsegregated
MNA (Metabolic network analysis)

COMPARE IT TO BLACK BOX STOICHIOMETRIC MODEL

$$-\mathrm{CH_2O} + 0.510\mathrm{CH_3O_{1/2}} + 0.275\mathrm{CO_2} + 0.137X + 0.077\mathrm{CH_{8/3}O} = 0.$$

COMPARE IT TO UNSTRUCTURED-UNSEGREGATED MODEL

1

$$\frac{dX}{dt} = \mu_{max}\left(\frac{S}{K_S + S}\right)X$$

Pertumbuhan sel

μ_{max} : laju pertumbuhan spesifik maksimum (1/jam)

S: konsentrasi substrat

K_S: konstanta afinitas substrat

X: kepadatan sel (x 10^6 sel/mL)

2

$$\frac{dS}{dt} = -\frac{dX}{dt}\frac{1}{Y_{x/s}}$$

Konsumsi substrat

Y_{xs}: Perolehan sel-glukosa (% b/b)

3

$$\frac{dP}{dt} = \alpha\frac{dX}{dt} + \beta X$$

Pembentukan produk
Luedeking-Piret

α: *Growth-related constant* (*dimensionless*)

β: *Non-growth-related constant* (1/jam)

WHAT IS THE BENEFIT OF BIOCHEMICAL REACTION NETWORK VS BLACK BOX STOICHIOMETRIC MODEL?


THERE IS ALWAYS
A CATCH. ALWAYS.
quickmeme.com

EXAMPLE

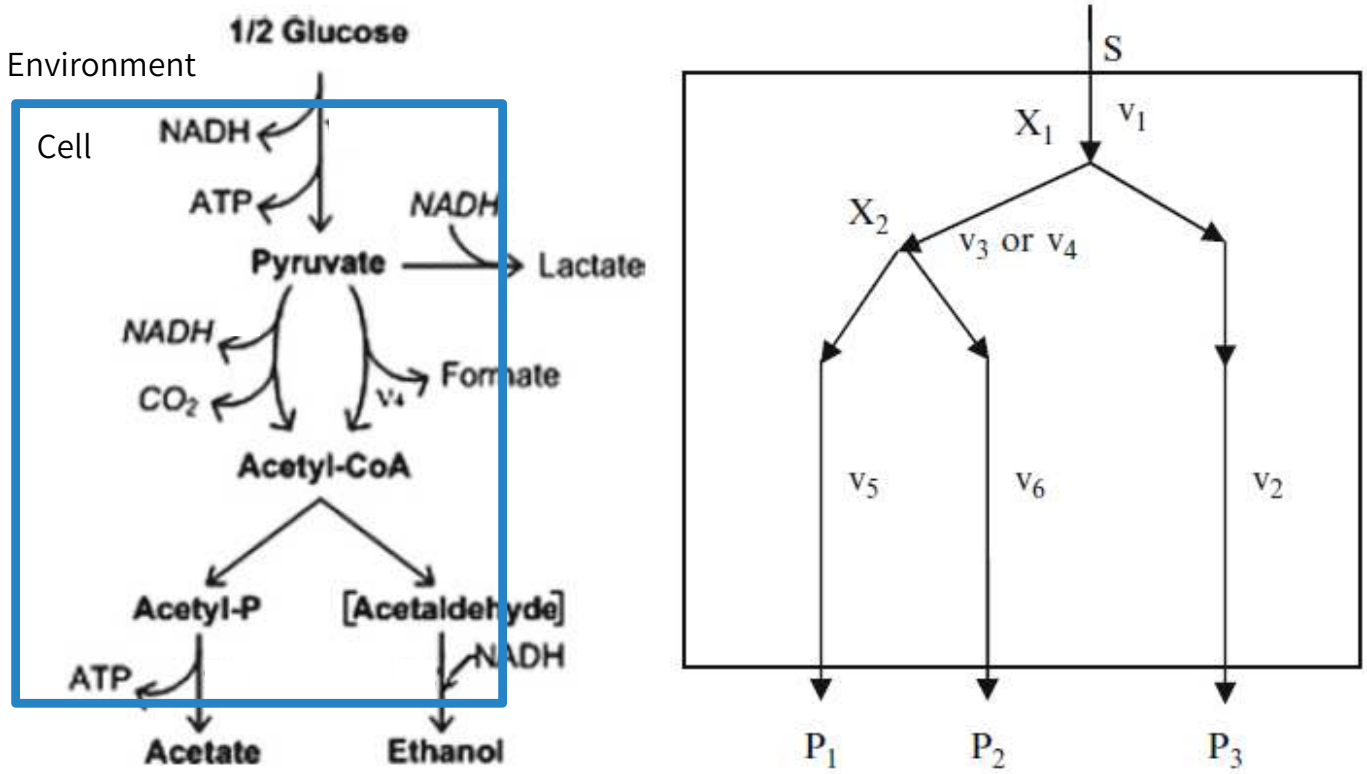


Fig. 5.3 Heterofermentative lactic acid fermentation via pyruvate dehydrogenase (v_3) or pyruvate formate lyase (v_4) to the final products, HAc (P_1), ethanol (P_2), and HLac (P_3). CO_2 is a byproduct of v_3, while formic acid (HCOOH) is a byproduct in v_4. X_1 = pyruvate (PYR) and X_2 = acetyl coenzyme A (AcCoA). S = glucose

Let’s recall basic kinetics rate law, ☺

SIMPLE REACTION KINETICS

- Unimolecular reaction $A \xrightarrow{K} B$

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

- Bimolecular reaction $A + B \xrightarrow{K} P + Q$

$$v = k[A][B]$$

STOICHIOMETRY FBA

For a metabolic network that contains m metabolites and n metabolic fluxes, all the transient material balances can be represented by a single matrix equation

$$\frac{dX}{dt} = S.v - b$$

where X is an m dimensional vector of metabolite amounts per cell, v is the vector of n metabolic fluxes, S is the stoichiometric m × n matrix, and b is the vector of known metabolic demands.

The element S_{ij} is the stoichiometric coefficient that indicates the amount of the i^{th} compound produced per unit of flux of the j^{th} reaction.

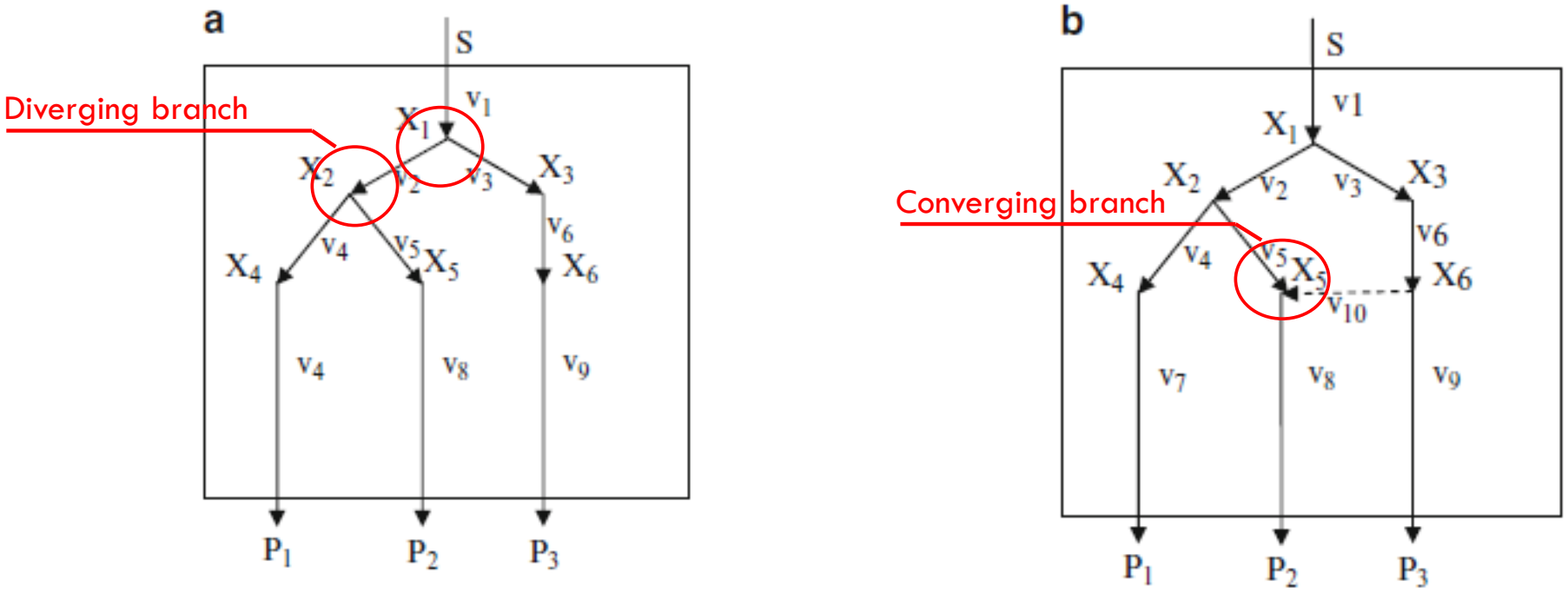
STOICHIOMETRY FBA

The time constants characterizing metabolic transients are typically very rapid compared to the time constants of cell growth, and the transient mass balances can be simplified to only consider the steady-state behaviour (dX/dt=0).

$$S.v = b$$

This equation simply states that over long periods of time, the formation of fluxes of a metabolite must be balanced by the degradation fluxes.

BUILDING A NETWORK MODEL



One carbon substrate S is fed to the network at a rate (rs), which is **equal** to the internal rate, or **flux**, v1, at which carbon is directed toward the first intracellular metabolite pool X1. The carbon flows from pool X1 toward the two pools X2 and X3. The two fluxes v2 and v3 determine the distribution of carbon toward the metabolic products, P1 + P2 and P3, respectively. At X2 there is a further distribution of carbon toward P1 and P2, respectively

BUILDING A NETWORK MODEL

"The objective of Metabolic Flux Analysis is to find the distribution of carbon in the product streams, and next to calculate those production rates P_i that are not given experimentally."

The values of v_i are very relevant when the objective is to learn how the cell responds to different perturbations, either genetic or in the environment.

The general assumption that allows us to calculate fluxes in metabolic networks is:

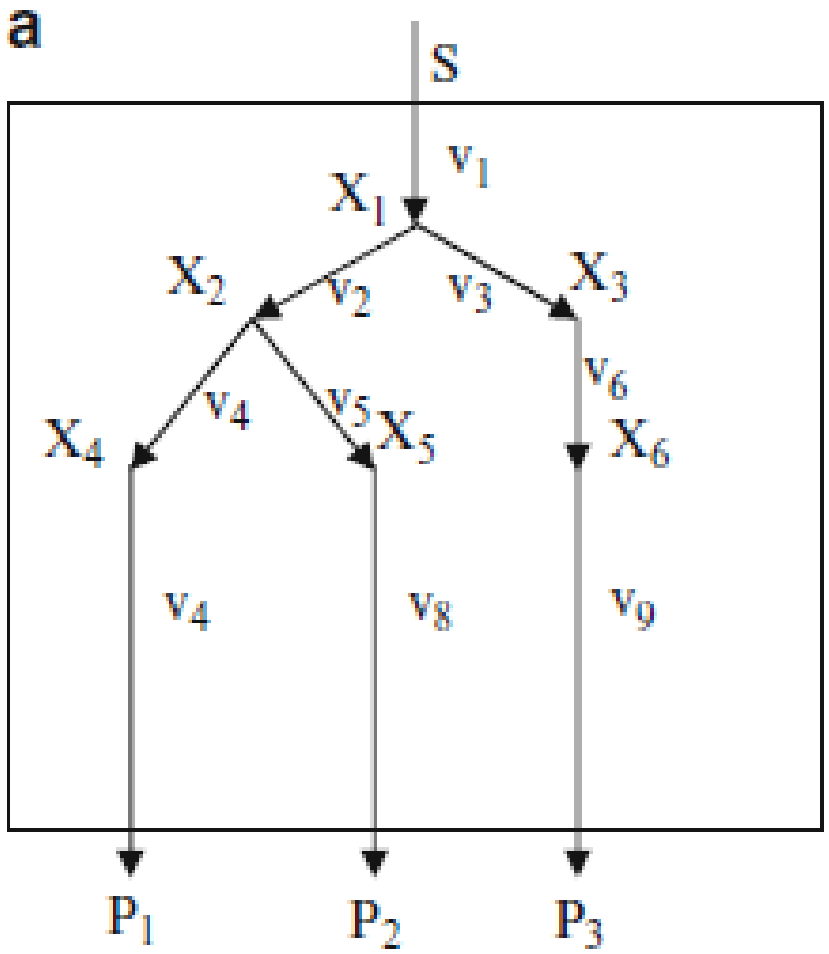
"For all intracellular metabolites the fluxes leading to a given metabolite are balanced with the fluxes leading away from the metabolite. This ensures that in a steady state situation there is no net accumulation of metabolites Xi"

BUILDING A NETWORK MODEL

Besides the constraints obtained from the relationships between the internal fluxes vi, more constraints can be imposed on the network:

1. A total carbon balance shows that the carbon input used to form the key products P_i must equal the carbon fed to the network in S.
2. A redox balance which specifies that the redox generated in one part of the network which is isolated from the rest of the metabolic network, must equal that consumed in the remainder of the network. Thus redox generated in V_1 must be consumed in V_2 and V_3.
3. Similarly an energy balance for an isolated network must stipulate that ATP generated in one part of the network must be consumed in other parts.

BUILDING A NETWORK MODEL



If the values of $[v_1, v_2, v_4]$ are known then the value of all the other fluxes can be calculated as:

$v_3=v_1+v_2$,

$v_5=v_2-v_4$

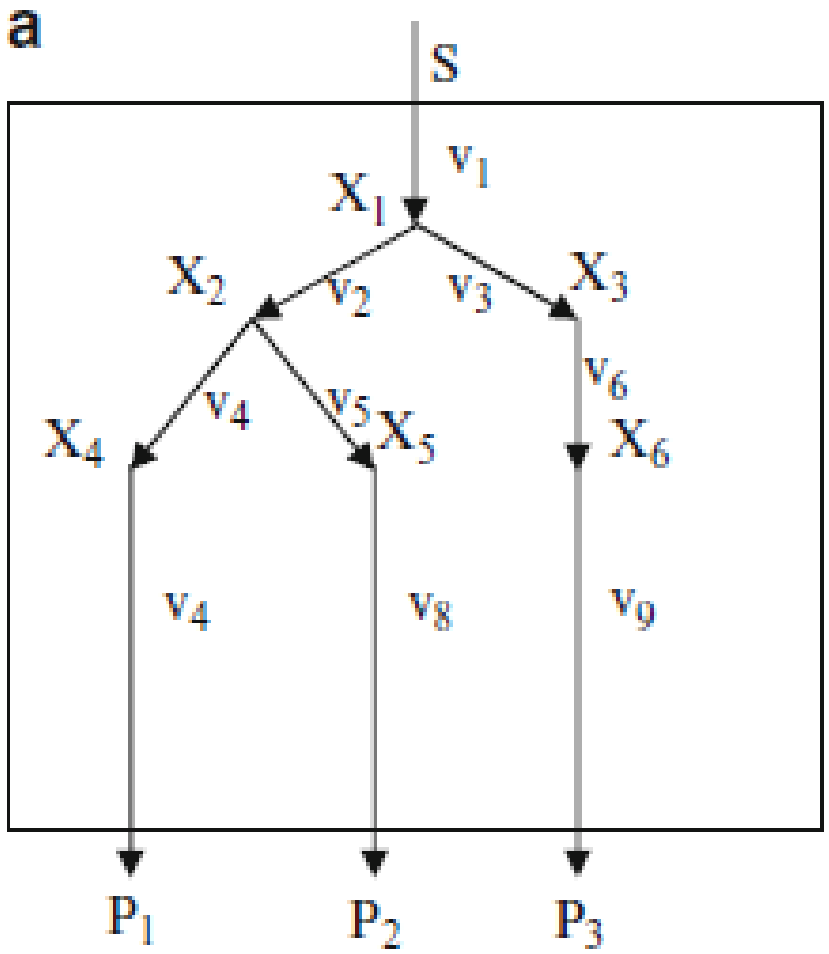
$v_6= v_9=v_3$,

$V_{7=}v_4$

$V_{8=}v_5$.

Thus, the vector $v=[v_1, v_2, v_4]$ is a key to the solution of the flux distribution problem.

BUILDING A NETWORK MODEL

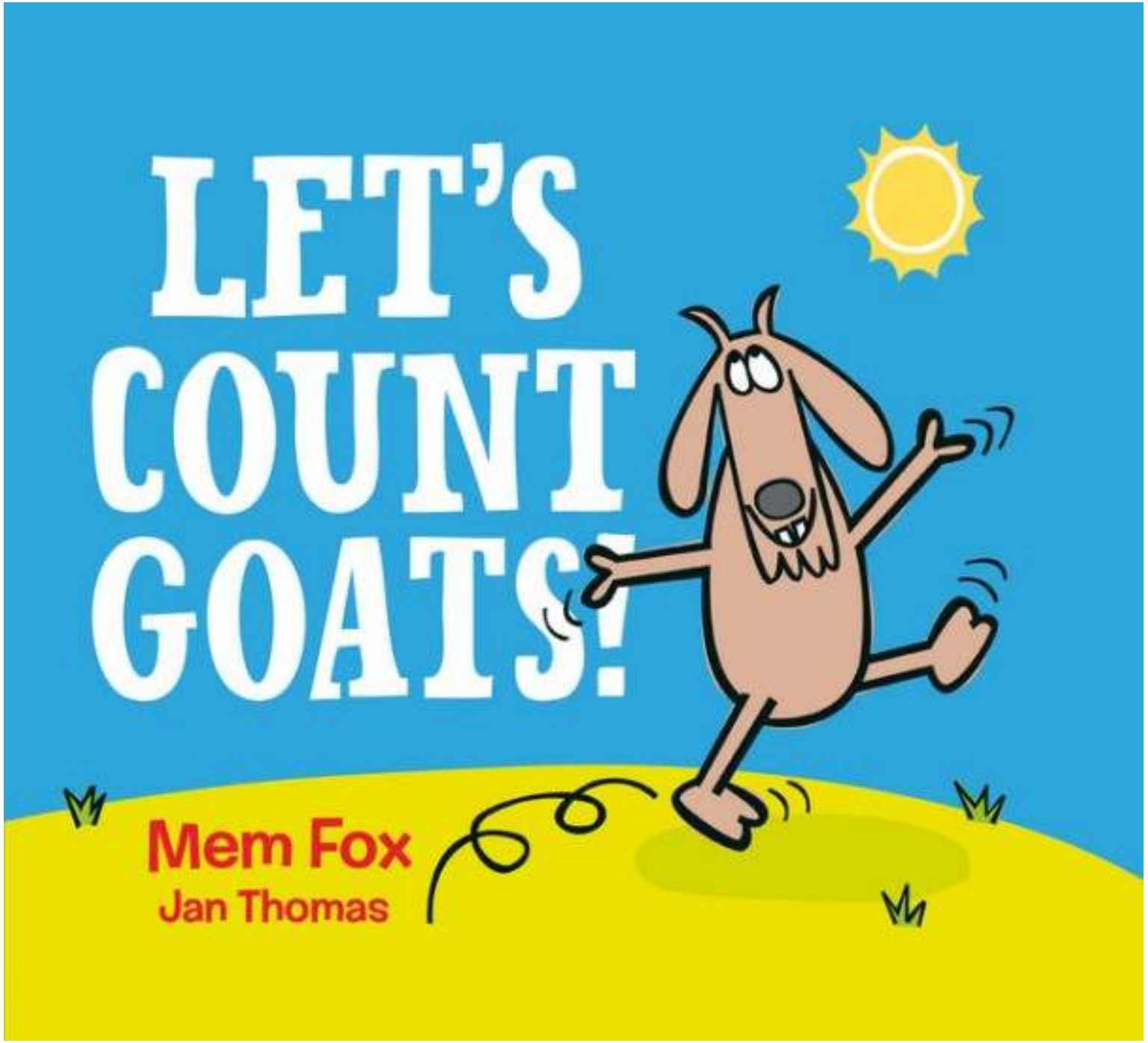


The elements of the key rate vector can be internal fluxes v_i, product formation rates r_{pi}, or the substrate consumption rate r_s.

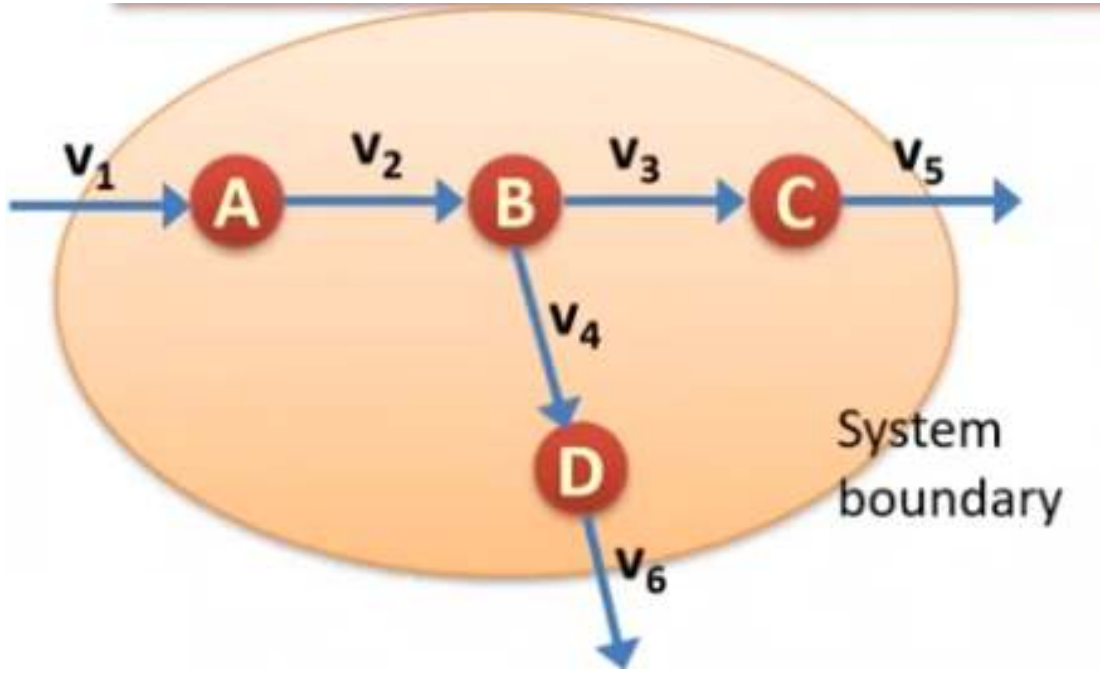
r_{pi} and r_s is the most useful choice, since these are easily available experimentally.

The left figure can be simplified into a network with only three independent paths: S to P_1, S to P_2, and S to P_3.

The vector $V=[V_1, V_2, V_3]$ symbolizes this reduced network, and product P_i is only produced in pathway V_i.

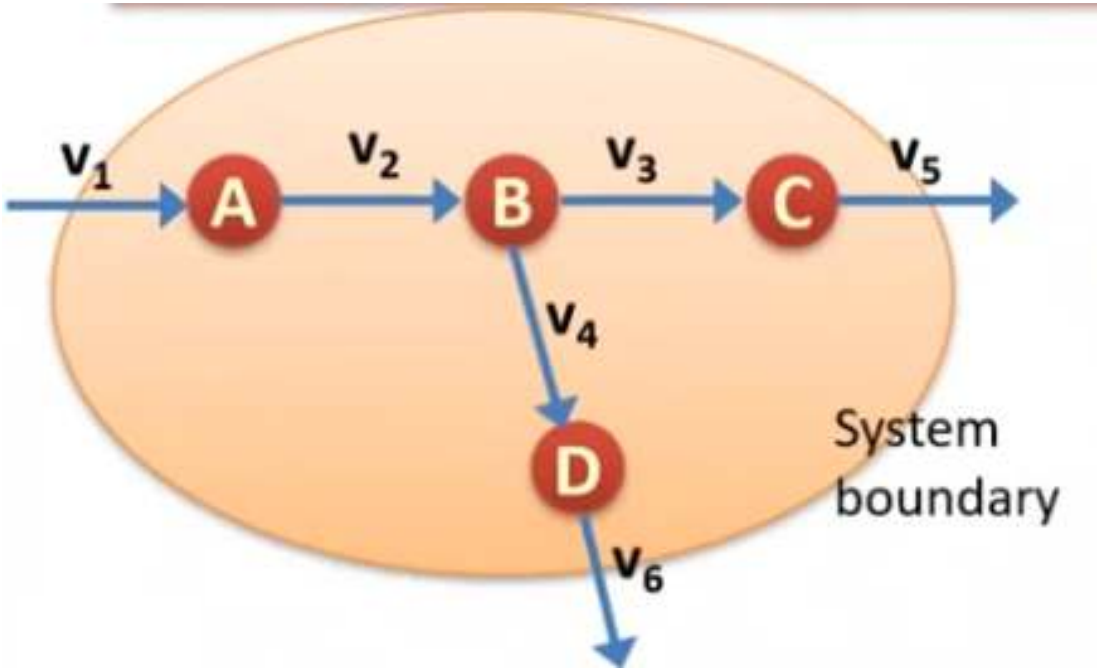

LET'S
COUNT
GOATS!
Mem Fox
Jan Thomas

QUANTIFYING A NETWORK MODEL : STEP I – SYSTEM DEFINITION



A model system comprising four metabolites (A, B, C, and D) with six reactions (three internal fluxes (v_2, v_3, and v_4) and three exchange fluxes (v_1, v_5, v_6).

QUANTIFYING A NETWORK MODEL : STEP I – SYSTEM DEFINITION



$$\frac{dA}{dt} = v_1 - v_2$$

$$\frac{dB}{dt} = v_2 - v_3 - v_4$$

$$\frac{dC}{dt} = v_3 - v_5$$

$$\frac{dD}{dt} = v_4 - v_6$$

QUANTIFYING A NETWORK MODEL : STEP I – SYSTEM DEFINITION

$$\frac{dA}{dt} = v_1 - v_2 \qquad 0 = v_1 - v_2$$

$$\frac{dB}{dt} = v_2 - v_3 - v_4 \qquad 0 = v_2 - v_3 - v_4$$

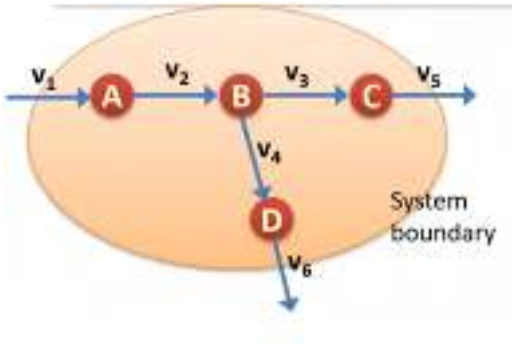
$$\frac{dC}{dt} = v_3 - v_5 \qquad 0 = v_3 - v_5$$

quasi steady state

$$\frac{dD}{dt} = v_4 - v_6 \qquad 0 = v_4 - v_6$$

QUANTIFYING A NETWORK MODEL : STEP I – SYSTEM DEFINITION

$$0 = v_1 - v_2$$
$$0 = v_2 - v_3 - v_4$$
$$0 = v_3 - v_5$$
$$0 = v_4 - v_6$$



Matrix Representation

$$\underbrace{\begin{bmatrix} & v1 & v2 & v3 & v4 & v5 & v6 \\ A & 1 & -1 & 0 & 0 & 0 & 0 \\ B & 0 & 1 & -1 & -1 & 0 & 0 \\ C & 0 & 0 & 1 & 0 & -1 & 0 \\ D & 0 & 0 & 0 & 1 & 0 & -1 \end{bmatrix}}_{\text{Stoichiometric Matrix}} \times \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$S \cdot \vec{v} = 0$$

$S_{m \times n}$

n = number of rows (mass balances)
*m = number of fluxes **
some literature use an interchangeable notation

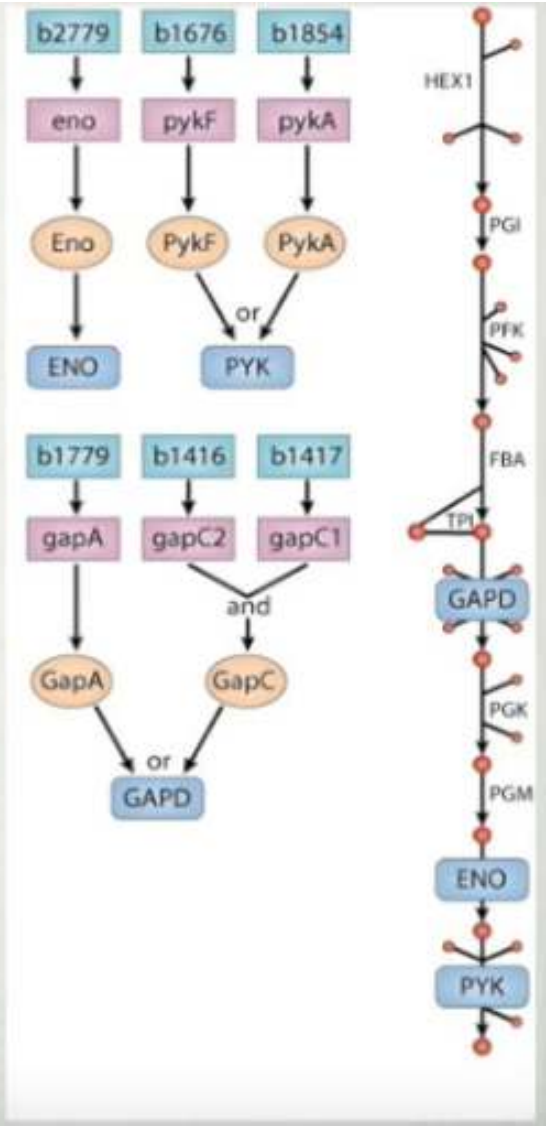
Inherently an underdetermined systems

QUANTIFYING A NETWORK MODEL : STEP I – SYSTEM DEFINITION

stoichiometric matrix is sometimes a bit arbitrary

Abbreviation	Glycolytic reactions	Genes
HEX1	[c]GLC + ATP → G6P + ADP + H	glk
PGI	[c]G6P ↔ F6P	pgi
PFK	[c]ATP + F6P → ADP + FDP + H	*pfkA, pfkB*
FBA	[c]FDP ↔ DHAP + G3P	*fbaA, fbaB*
TPI	[c]DHAP ↔ G3P	tpiA
GAPD	[c]G3P + NAD + PI ↔ 13DPG + H + NADH	gapA, gapC1, gapC2
PGK	[c]13DPG + ADP ↔ 3PG + ATP	pgk
PGM	[c]3PG ↔ 2PG	gpmA, gpmB
ENO	[c]2PG ↔ H_2O + PEP	eno
PYK	[c]ADP + H + PEP → ATP + PYR	pykA, pykF

	HEX1	PGI	PFK	FBA	TPI	GAPD	PGK	PGM	ENO	PYK
ATP	−1	0	−1	0	0	0	1	0	0	1
GLC	−1	0	0	0	0	0	0	0	0	0
ADP	1	0	1	0	0	0	−1	0	0	−1
G6P	1	−1	0	0	0	0	0	0	0	0
H	1	0	1	0	0	1	0	0	0	−1
F6P	0	1	−1	0	0	0	0	0	0	0
FDP	0	0	1	−1	0	0	0	0	0	0
DHAP	0	0	0	1	−1	0	0	0	0	0
G3P	0	0	0	1	1	−1	0	0	0	0
NAD	0	0	0	0	0	−1	0	0	0	0
PI	0	0	0	0	0	−1	0	0	0	0
13DPG	0	0	0	0	0	1	−1	0	0	0
NADH	0	0	0	0	0	1	0	0	0	0
3PG	0	0	0	0	0	0	1	−1	0	0
2PG	0	0	0	0	0	0	0	1	−1	0
PEP	0	0	0	0	0	0	0	0	1	−1
H_2O	0	0	0	0	0	0	0	0	1	0
PYR	0	0	0	0	0	0	0	0	0	1



QUANTIFYING A NETWORK MODEL: STEP II – MASS BALANCE

Metabolic Flux Analysis **(MFA):** Determined system

- Either DOF = 0 or a set of reaction fluxes equal to DOF is measured
- Solution is possible by Gauss-Jordan Elimination or matrix inversion

Flux Balance Analysis **(FBA)** : Under-determined system

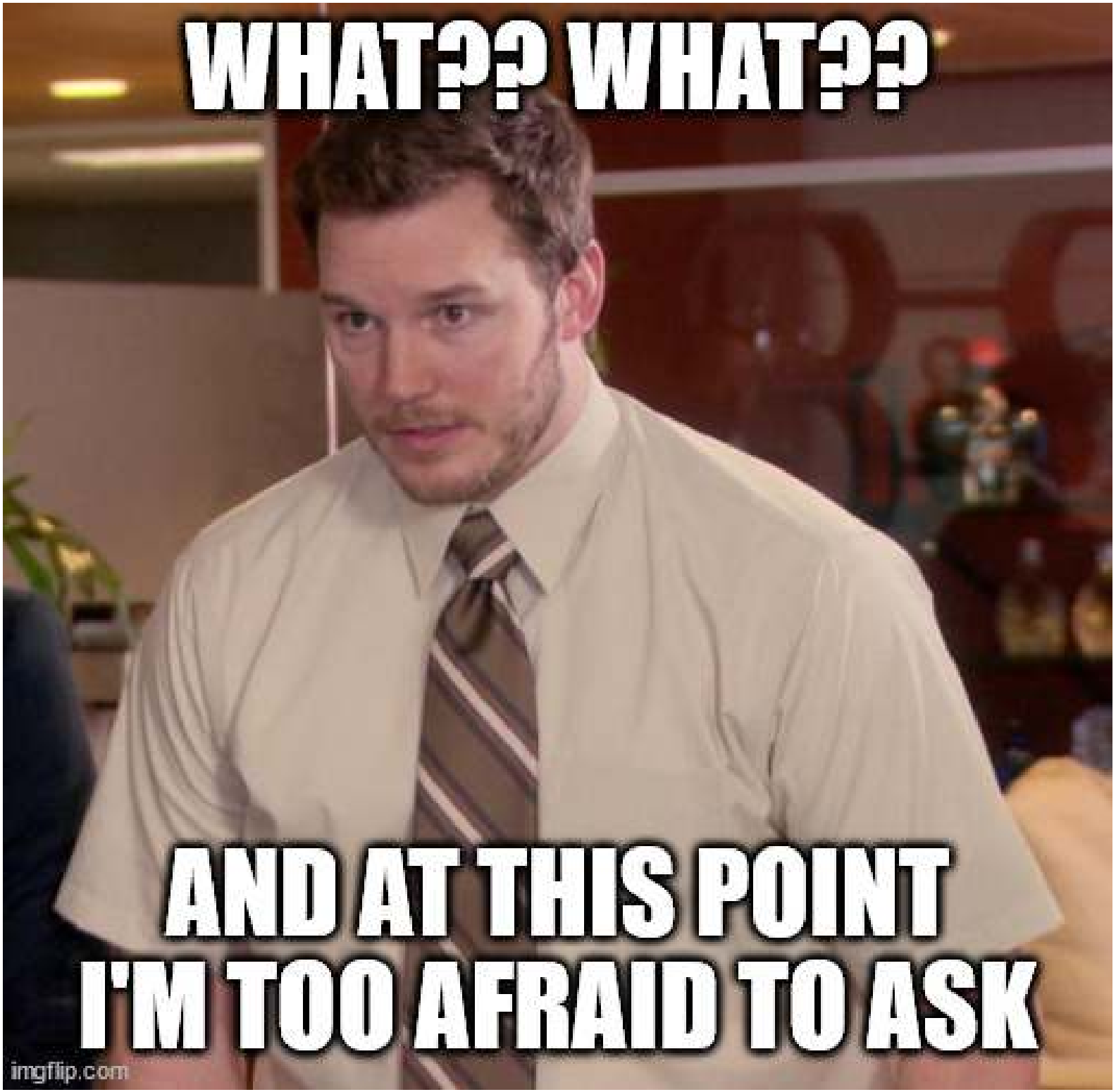
- DOF > 0
- An optimization problem
- Maximizing/minimizing one of the unknows
- Solution is possible by Linier/Quadratic programming

QUANTIFYING A NETWORK MODEL: STEP II – MASS BALANCE

Additional information is needed to solve for all the metabolic fluxes:

Additional assumptions are required, such as neglecting certain reactions occurring within the cell.

However, the measurement of internal fluxes is not always practical, and these measurements can only allow for the determination of the metabolic fluxes in subsystems of the metabolic network.


WHAT?? WHAT??
AND AT THIS POINT
I'M TOO AFRAID TO ASK
imgflip.com

QUANTIFYING A NETWORK MODEL: STEP IIIA – CALCULATING MFA

- For determined systems
  - Partition S into two : i.e., $S^{calculated}$, and $S^{measured}$
  - Partition $\vec{v}$ into two : i.e., $v^{calculated}$, and $v^{measured}$

$$[S^{calculated}\ S^{measured}]\begin{pmatrix} v^{calculated} \\ v^{measured} \end{pmatrix} = 0$$

$$S^{calculated}\ v^{calculated} + S^{measured}\ v^{measured} = 0$$

$$S^{calculated}\ v^{calculated} = -\ S^{measured}\ v^{measured}$$

Multiply both sides by the inverse of $S^{calculated}$

$$v^{calculated} = -\ (S^{calculated})^{-1}\ (S^{measured}\ v^{measured})$$

QUANTIFYING A NETWORK MODEL: STEP IIIA – CALCULATING MFA

- But remember…

A square matrix is only invertible if it does not have any dependent rows

QUANTIFYING A NETWORK MODEL: STEP III – DEFINING MEASURABLE FLUXES & CONSTRAINTS IN FBA

FBA investigates metabolism by involving constraints in the stoichiometric analysis:

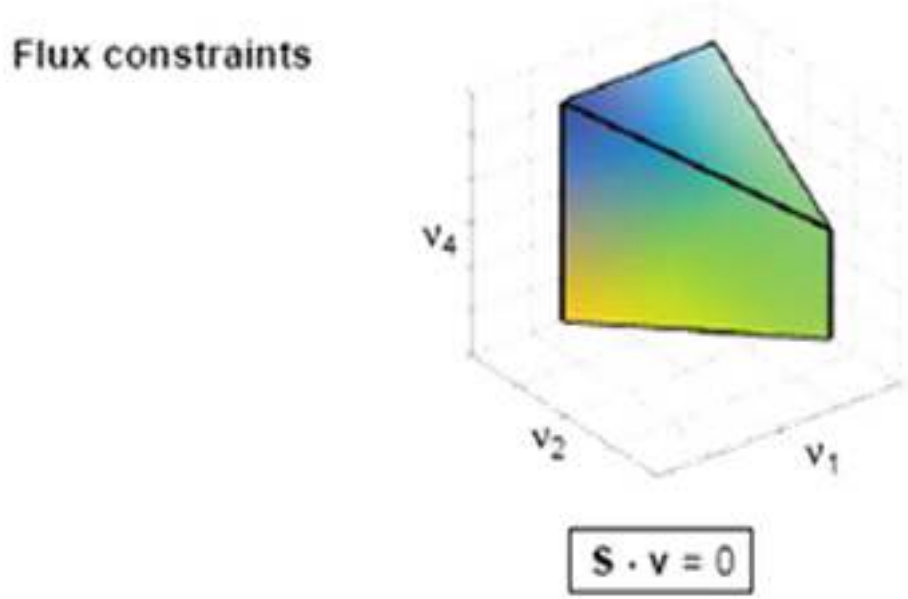
– The first constraint is set by the assumption of a steady state.

– The second constraint is of a thermodynamic nature, respecting the irreversibility of reactions.

– The third constraint may result from the limited capacity of enzymes for metabolite conversion.

– Further constraints may be imposed by biomass composition or other external conditions.

QUANTIFYING A NETWORK MODEL: STEP III – DEFINING MEASURABLE FLUXES & CONSTRAINTS



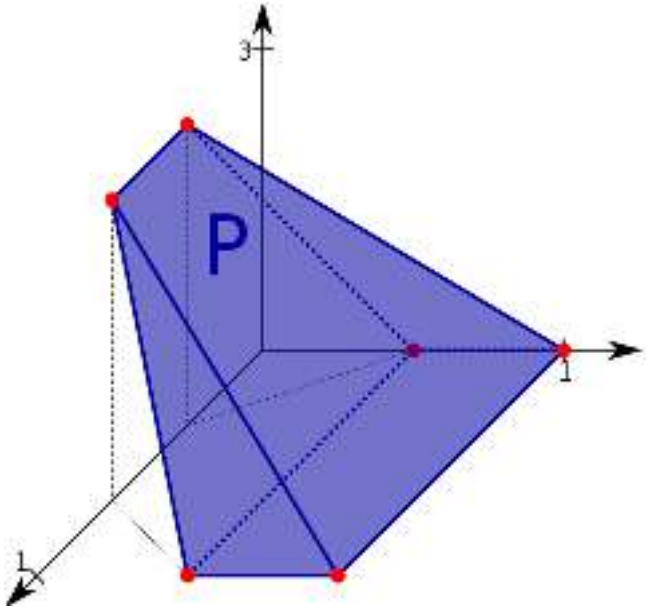
The fluxes of the system are constrained on the basis of thermodynamics and experimental insights. This creates a flux cone corresponding to the metabolic capacity of the organism.

QUANTIFYING A NETWORK MODEL: STEP IV – OPTIMIZATION

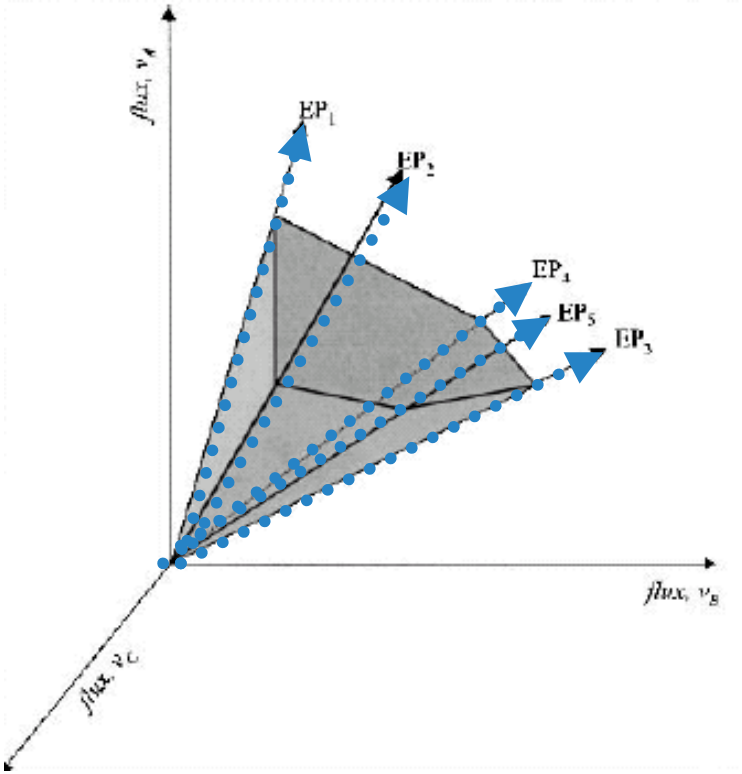
- Define of objective function Z
  - E.g., biomass production in defined proportion.
- → A well-formulated optimization problem P:
  - Maximize (or min.) Z, subject to,
    - $S \cdot v = 0$;
    - $C = 0$;
- If Z is linear, then P could be solved through LP techniques.

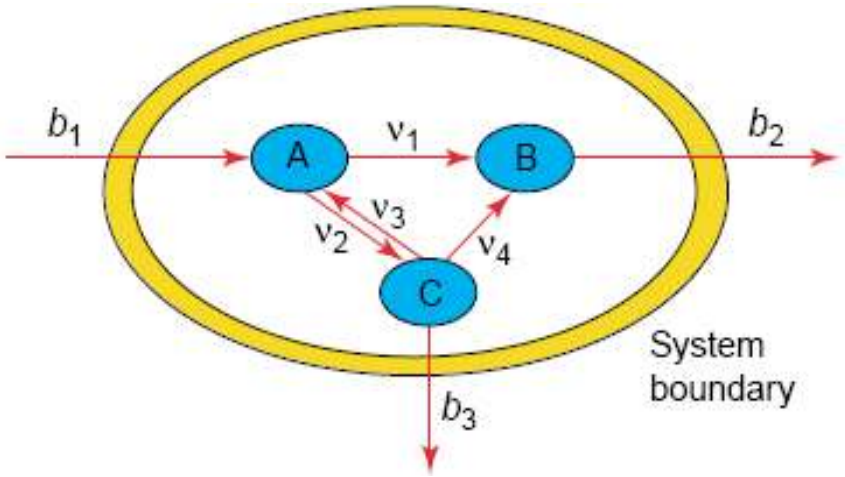
The determination of a particular metabolic flux distribution has been formulated as a **linear programming problem**. The idea is to maximize an objective function Z that is subject to the stoichiometric and capacity constraints

Linear programming (**LP**, also called **linear optimization**) is a method to achieve the best outcome (such as maximum profit or lowest cost) in a mathematical model whose requirements are represented by linear relationships. Linear programming is a special case of mathematical programming (mathematical optimization)



What is the biological interpretation of any point in the flux cone ?





(I) NARROWING THE STEADY STATE FLUX CONE

The steady state flux cone contains infinite flux distributions!

Only a small portion of them is physiologically feasible.

- More constraints on the external fluxes.

  These depend on factors as:
  - Organism
  - Environment and accessibility substrates
  - maximum rates of diffusion mediated transport
  - Etc…


EXAMPLE
EXAMPLE
EXAMPLE
www.shutterstock.com · 424524877

EXAMPLE 1 – SIMPLE MFA

Stoichiometric matrix

$$\begin{bmatrix} B \\ C \\ D \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}$$

$v_4 = 10$ mmol g^{-1} (dw) h^{-1}.

$$\begin{bmatrix} B \\ C \\ D \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

Determine the remaining flux rate!

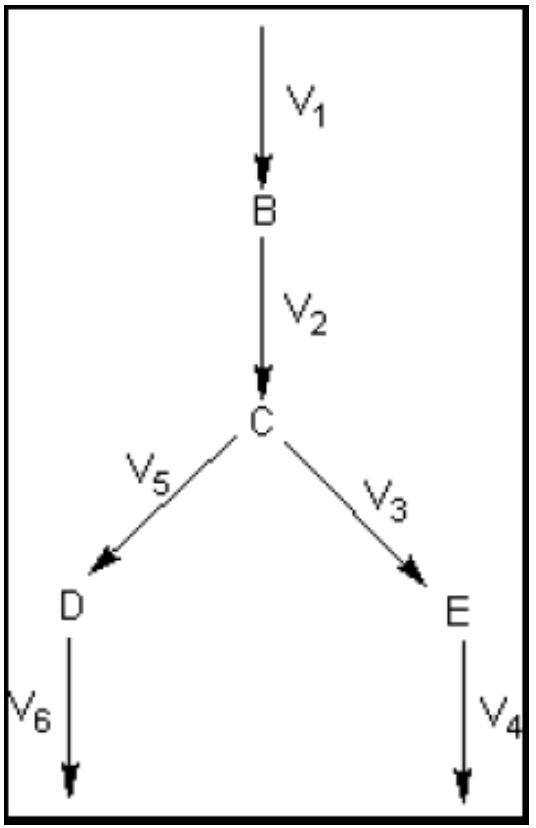
EXAMPLE 2 – SIMPLE BRANCHED MFA

Stoichiometric matrix

$$\begin{bmatrix} B \\ C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{bmatrix}$$

$v_4 = 6$ mmol g^{-1} (dw) h^{-1}
$v_6 = 4$ mmol g^{-1} (dw) h^{-1}

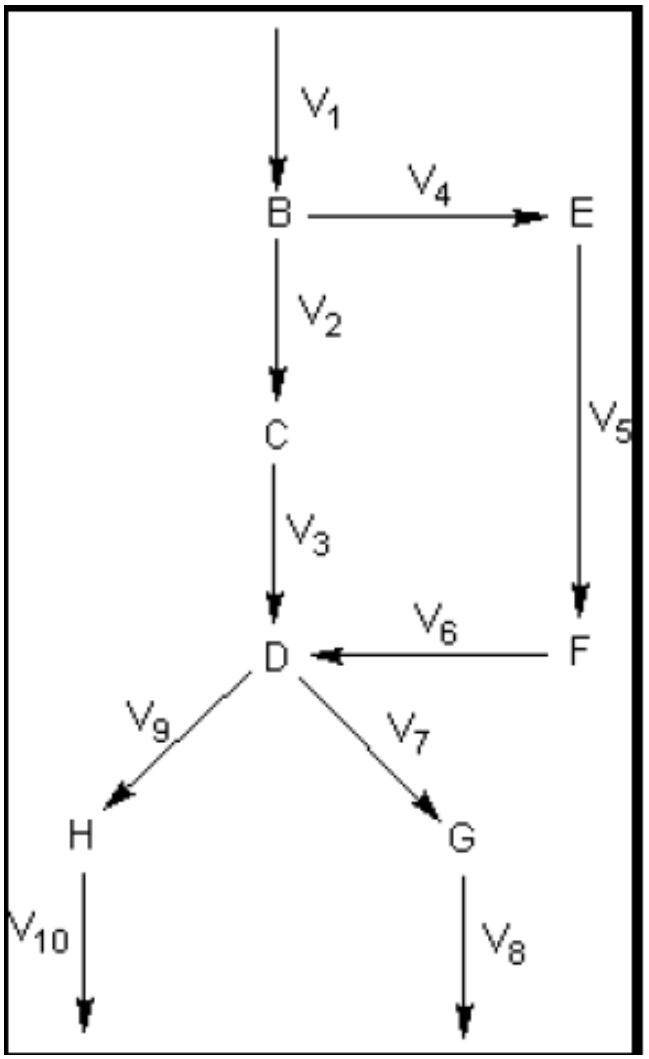
$$\begin{bmatrix} B \\ C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_5 \end{bmatrix}$$



Determine the remaining flux rate!

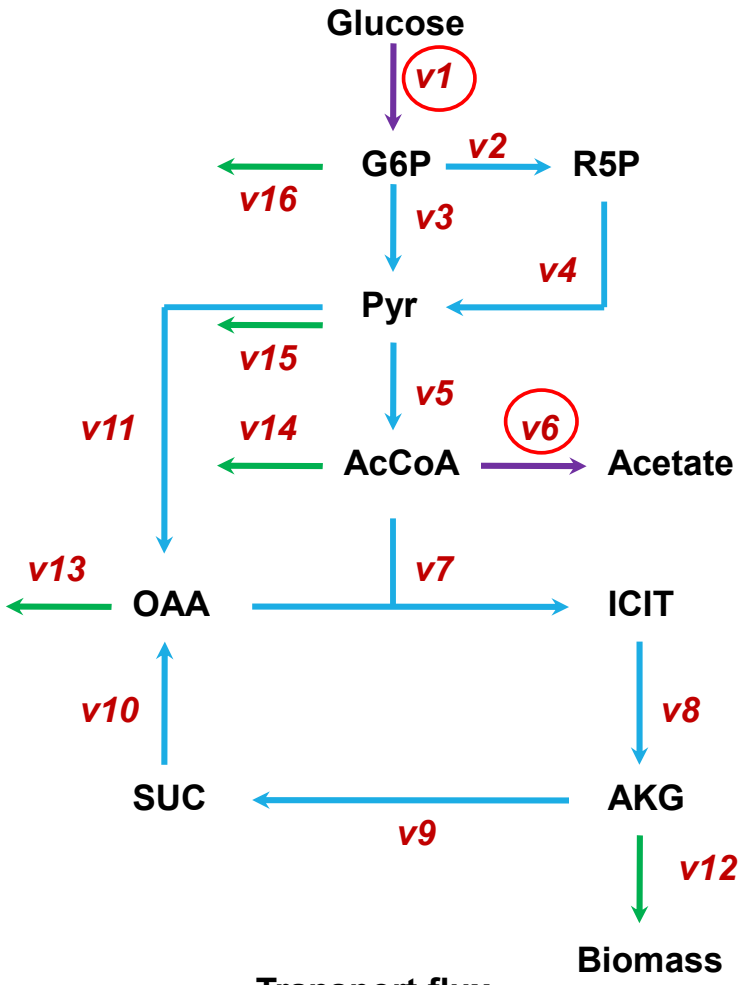
EXAMPLE 3 – BRANCHED MFA

V_1 = 10 mmol g^{-1} (dw) h^{-1}
V_4 = 3 mmol g^{-1} (dw) h^{-1}
V_{10} = 6 mmol g^{-1} (dw) h^{-1}



Determine the remaining flux rate!

EXAMPLE 4 : UNDERDETERMINED SYSTEM (FBA)



(i.e., carbohydrate, protein, lipids, etc.)

16 fluxes, 8 intracellular metabolites

G6P : v1 = v2+v3+v16
R5P : v2 = v4
Pyr : v3+v4 = v5+v11+v15
AcCoA: v5 = v6+v7+v14
ICIT : v7 = v8
AKG : v8 = v9+v12
SUC : v9 = v10
OAA : v10+v11 = v7+v13

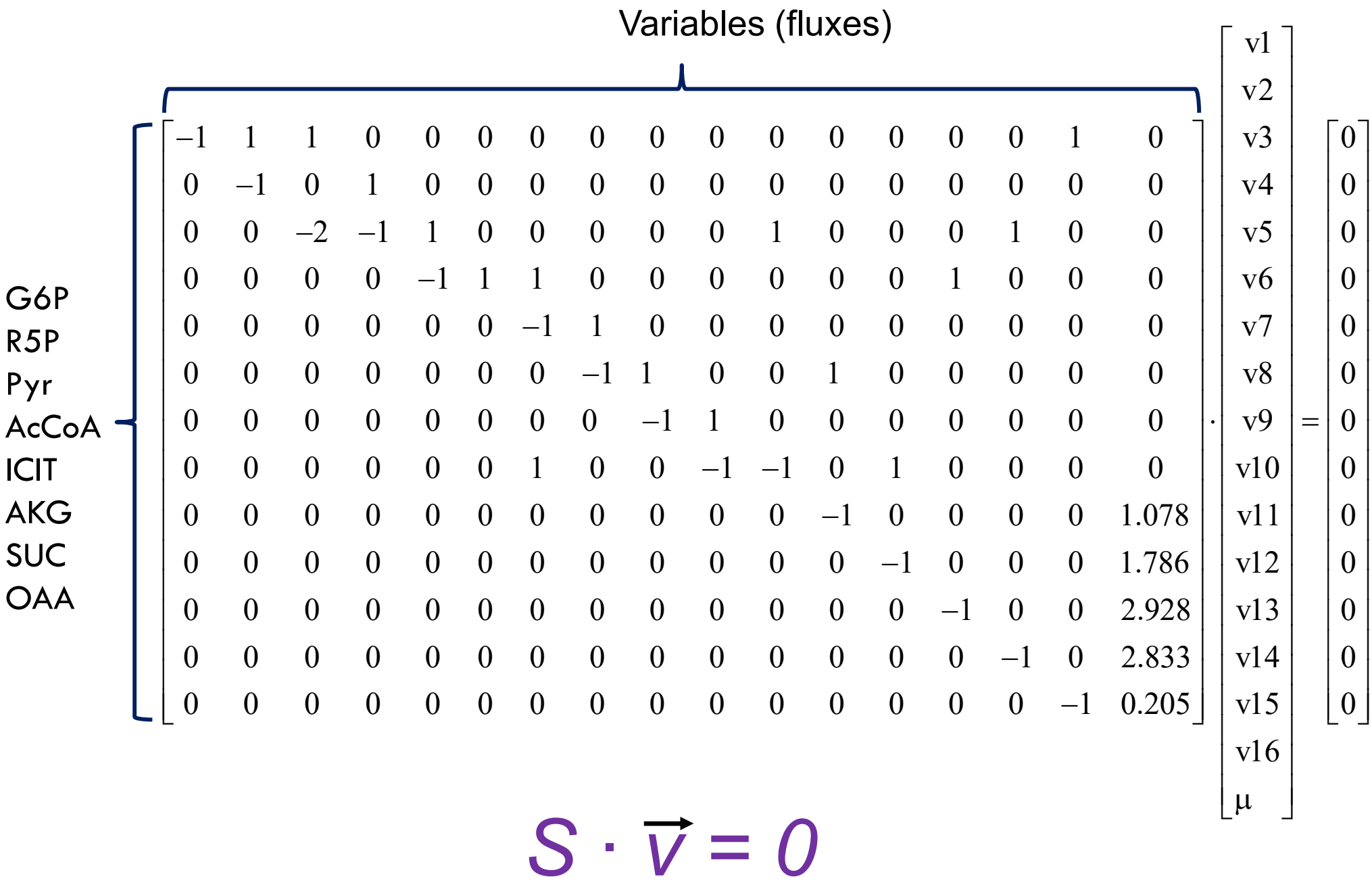
The transport fluxes were measured:

v1=11.0 mmol/g DCW/h

v6=6.4 mmol/g DCW/h

The building block fluxes can be assumed from biomass composition:

v12 : 1.078 μ (in terms of fluxes)
v13 : 1.786 μ
v14 : 2.928 μ
v15 : 2.833 μ
v16 : 0.205 μ


Variables (fluxes)
G6P
R5P
Pyr
AcCoA
ICIT
AKG
SUC
OAA
v1
v2
v3
v4
v5
v6
v7
v8
v9
v10
v11
v12
v13
v14
v15
v16
μ
$S \cdot \vec{v} = 0$

maximize μ

s.t. $S \cdot v = 0$

$0 < v < 20$ mmol/g DCW/h

$$\text{obj} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T$$

$$\text{lb} = \begin{bmatrix} 11.0 & 0 & 0 & 0 & 0 & 6.4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T$$

$$\text{ub} = \begin{bmatrix} 11.0 & 20 & 20 & 20 & 20 & 6.4 & 20 & 20 & 20 & 20 & 20 & 20 & 20 & 20 & 20 & 20 & 20 \end{bmatrix}^T$$