Syllabus

Principles and Applications of Systems Biology

Vassily Hatzimanikatis
September 2025

Assistants

Omar Keshk

David Liaskos

Luca Milazzo

Syllabus (approximate)

Date	Lecture	Exercise
11.09	Introduction to Systems Biology, FBA, Objective functions, Growth trade off (FBA1)	Q/A - Environment Setup
18.09	Phenotypic phase plane, alternative flux distributions (FBA2) Practical (Software; Simulation exercises) + dFBA	Q/A - Excercises
25.09	Data integration of Stoichiometric models; Practical exercises	Q/A - Excercises
02.10	Thermodynamics flux balance analysis, and GCM (TFA2)	Q/A - Excercises
09.10	Enzyme kinetics/Large-scale modelling + Kinetic modeling of cancer metabolism	Q/A
16.10	Signal transduction systems – Network reconstruction	Q/A
30.10	Signal transduction systems - Dynamics	Q/A
06.11	ME-models	Q/A
13.11	Microbial communities	Q/A
20.11	Other topics (Whole-cell and whole-body models)	Q/A
27.11	Work on projects	Q/A
04.12	*Final Presentations*	
11.12	*Final Presentations*	
18.12		

Brief course description

The course introduces and develops the key concepts from **systems biology** and **systems engineering** in the context of **complex biological networks**.

The lectures elaborate on **techniques** and **methods** to **model** and **analyze** complex biological problems.

(*Focus on Metabolism – and some signaling!*)

Examples and projects apply the model-based and systems engineering integration and analysis of big data from biological systems.

TOPICS

- Mathematical and computational analysis of metabolic reaction networks
- Analysis of metabolomics and bioenergetics data in the context metabolic networks
- Mathematical and computational analysis of protein expression
- Methods and technologies for the analysis of signaling networks
- Computational models of microbial communities
- Interpretation and analysis of single cell data
- Mathematical modeling of spatial effects in biological systems

METHODS

- Metabolic Flux balance analysis (FBA)
  (**Linear programming**)
- Thermodynamics based flux balance analysis (TFA) of metabolic networks
  (**Mixed-integer linear programing).**
- Kinetic models
  (**Ordinary differential equations**)
- Metabolic control analysis
  (**Local and global sensitivity analysis**)
- Stochastic simulation algorithm (SSA) and Particle based simulation methods
  (**Stochastic simulation**)
- Parameter estimation for biological systems
  (**System identification methods**)

RECOMMENDATIONS

The building of working groups will make it possible for people with partial knowledge in these fields to contribute depending on their formation.

COURSES

SV courses:

- Dynamical systems in biology BIO-341
- Numerical analysis MATH-251

ChemE courses:

- Dynamics and kinetics CH-310
- Biochemical engineering ChE-311
- Bioreactor modelling and simulation ChE-320
- Numerical methods ChE-312

RECOMMENDATIONS

The building of working groups will make it possible for people with partial knowledge in these fields to contribute depending on their formation.

COURSES

SV courses:

- Dynamical systems in biology BIO-341
- Numerical analysis MATH-251

ChemE courses:

- Dynamics and kinetics CH-310
- Biochemical engineering ChE-311
- Bioreactor modelling and simulation ChE-320
- Numerical methods ChE-312

Important concepts to start the course

- For the computational exercises, MATLAB® and PYTHON will be the platforms of choice.
- An introductory session on the platforms and software used is part of the course.

Learning Outcomes

By the end of the course, the student must be able to:

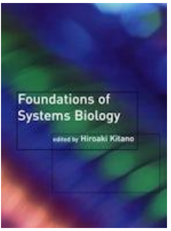
- Formulate mass balances of reaction networks
- Solve mass balance equations using linear programing solvers
- Construct kinetic models of biological reactions
- Create and analyze stochastic models of biological reactions
- Analyze papers on modeling and analysis of biological networks
- Assess / Evaluate alternative methods for the study of biological networks

Transversal skills

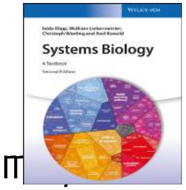
- Plan and carry out activities in a way which makes optimal use of available time and other resources.
- Access and evaluate appropriate sources of information.
- Summarize an article or a technical report.
- Demonstrate the capacity for critical thinking
- Negotiate effectively within the group.

Literature

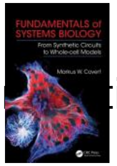
- Foundations of System Biology, Edited by Hiraoki Kitano. MIT Press 2001 (classic reading)



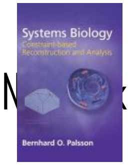
- Systems Biology, a textbook, Edda Klipp



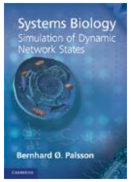
- Fundamentals of Systems Biology: From Circuits to Whole-Cell Models, by Markus Covert



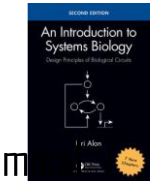
- Systems Biology : Constraint-Based Reco tion and Analysis



- Systems Biology: Simulation of Dynamic N k States

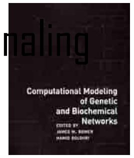


- An Introduction to Systems Biology: Design Principles of Biological Circuits, by Uri Alon. Chapman and Hall/CRC 2006

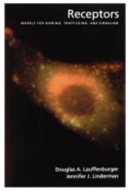


- Computational Modeling of Genetic and Biochem Networks, by James M. Bower and Hamid Bolouri. Bradford 2004

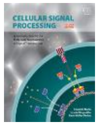
- Receptors: Models for Binding, Trafficking, and Sig aling by Douglas A. Lauffenburger, Jennifer Linderman

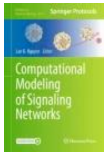


- Cellular Signal Processing: An Introduction to the Molecular Mechanisms of Signal Transduction



- Computational Modeling of Signaling Networks





Grading

- Project based evaluation:
  - Analyze and reproduce the results of a paper (A paper will be assigned to each group)
- TAs will be available to solve the questions during exercise sessions
- One grade per group
- Final presentation 100%

Groups and Project presentations

- Groups:
  - 2-3 People
  - Send one e-mail per group to david.liaskos@epfl.ch

    with the full names of all group members until the 27th Sep
  - If you are not part of a group by that time, we will assign you to a group

- Project presentation:
  - After forming a group, a paper will be assigned to you
  - Goal: Replicate the key results from the article
  - Propose extensions to the work presented

Introduction into Systems Biology

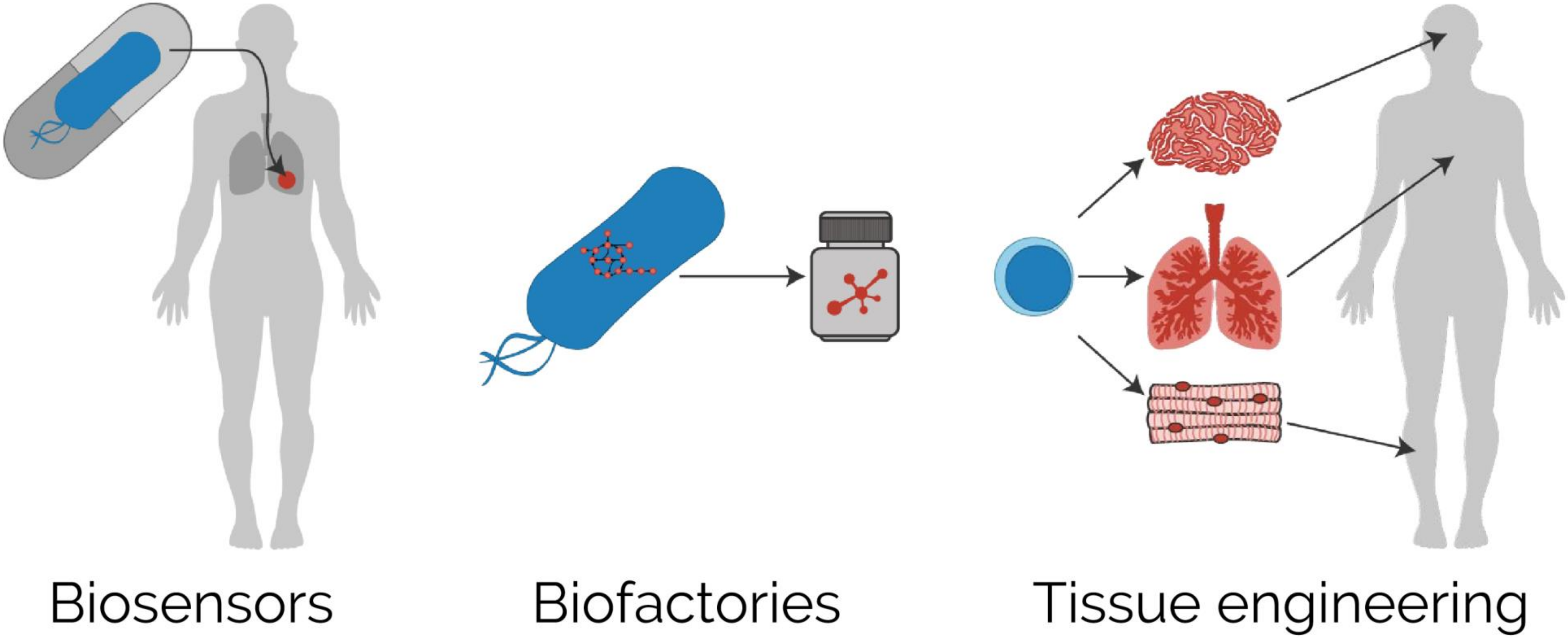
Principles and Applications of Systems Biology

Vassily Hatzimanikatis
September 2025

Systems Biology

Understanding biology on a systems level:

Systems Biology - Why?



Complex Systems

Imagine a system consists of N **agents** that make **individual decisions** and **interact** with each other.

If the behavior of the system **cannot be deduced from the behavior of the individual agents** but can only deduced from **all the participating agents and their interactions** this system is complex.

Let A,B,C and D be departments in a company that perform a special task and interact with each other.

The overall performance of this company will not be determined by how fast the individual departments perform their task with a given input resources but how these departments interchange resources and information.

The cell a complex system

Similar departments exist in the cell:

The cell a complex system

The cell a complex system

External stimuli

Structural building blocks

Genes

Secretion

Signaling

Transcription regulations

Protein supply

Translation regulation

Metabolism

Transcription

Supply building blocks

Supply AA’s

Nutrient uptake

mRNA supply

Protein synthesis

Enzyme supply

Complex subsystems

Structural building blocks

Secretion

Protein sup

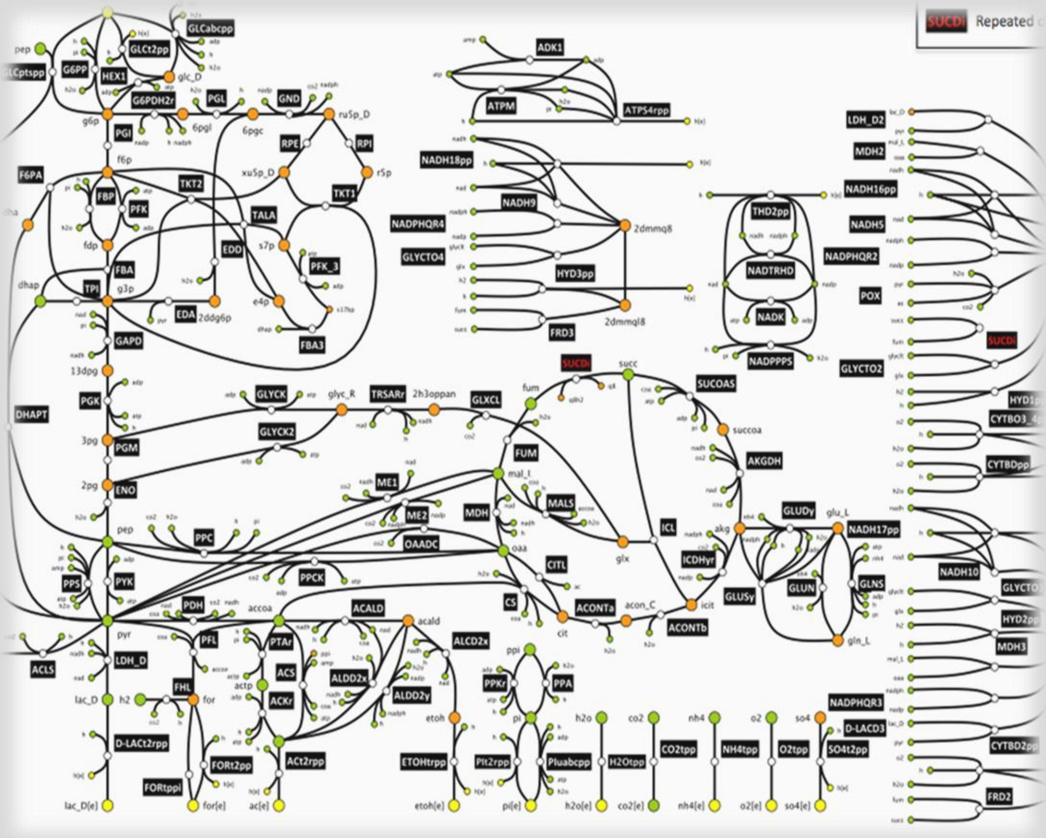
Metabolism

Supply AA's

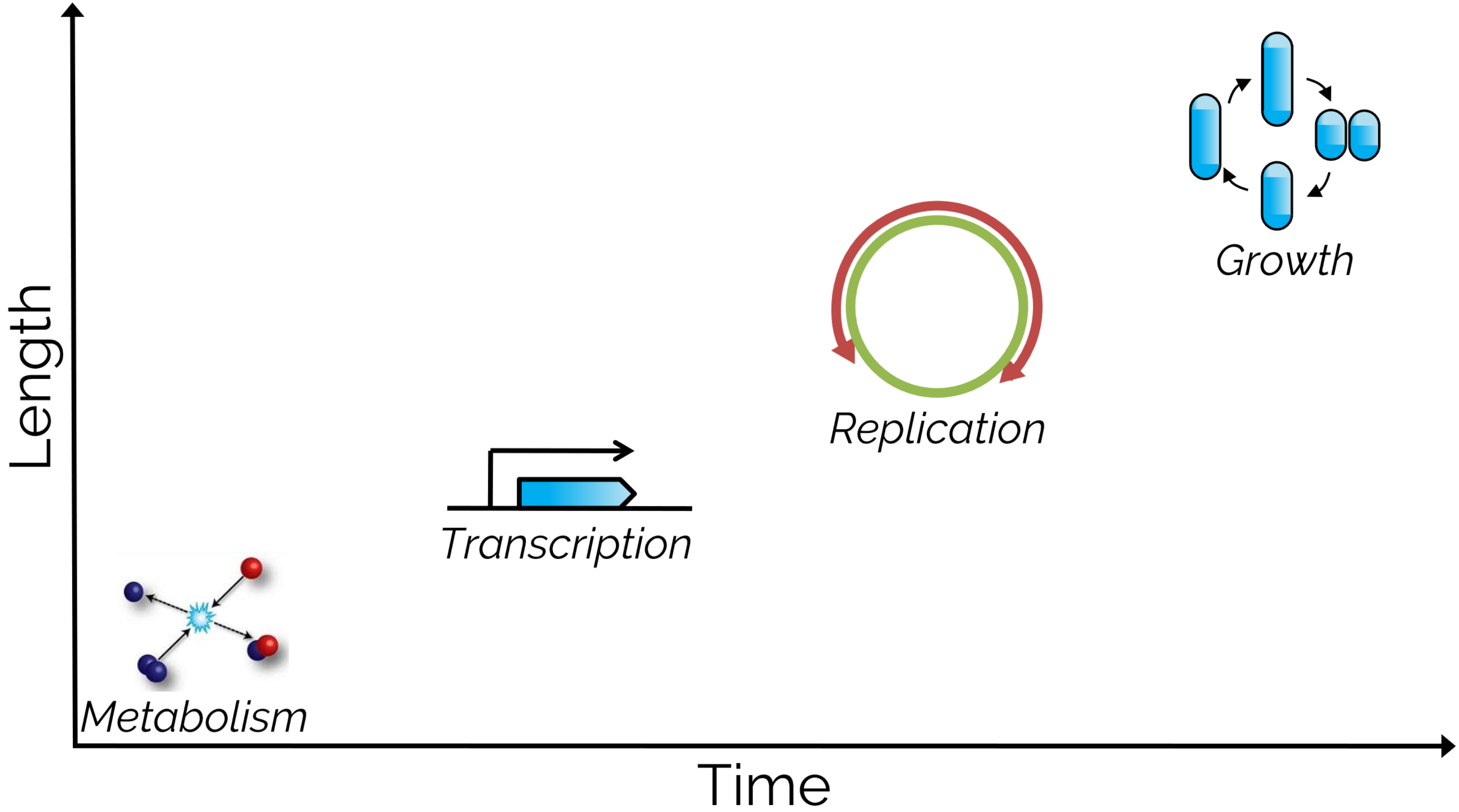
Nutrient uptake

Enzyme supply

These “departments” themselves have then a complex substructure. Consisting of many interacting components to from their behaviour



Complexity in time



Length

Growth

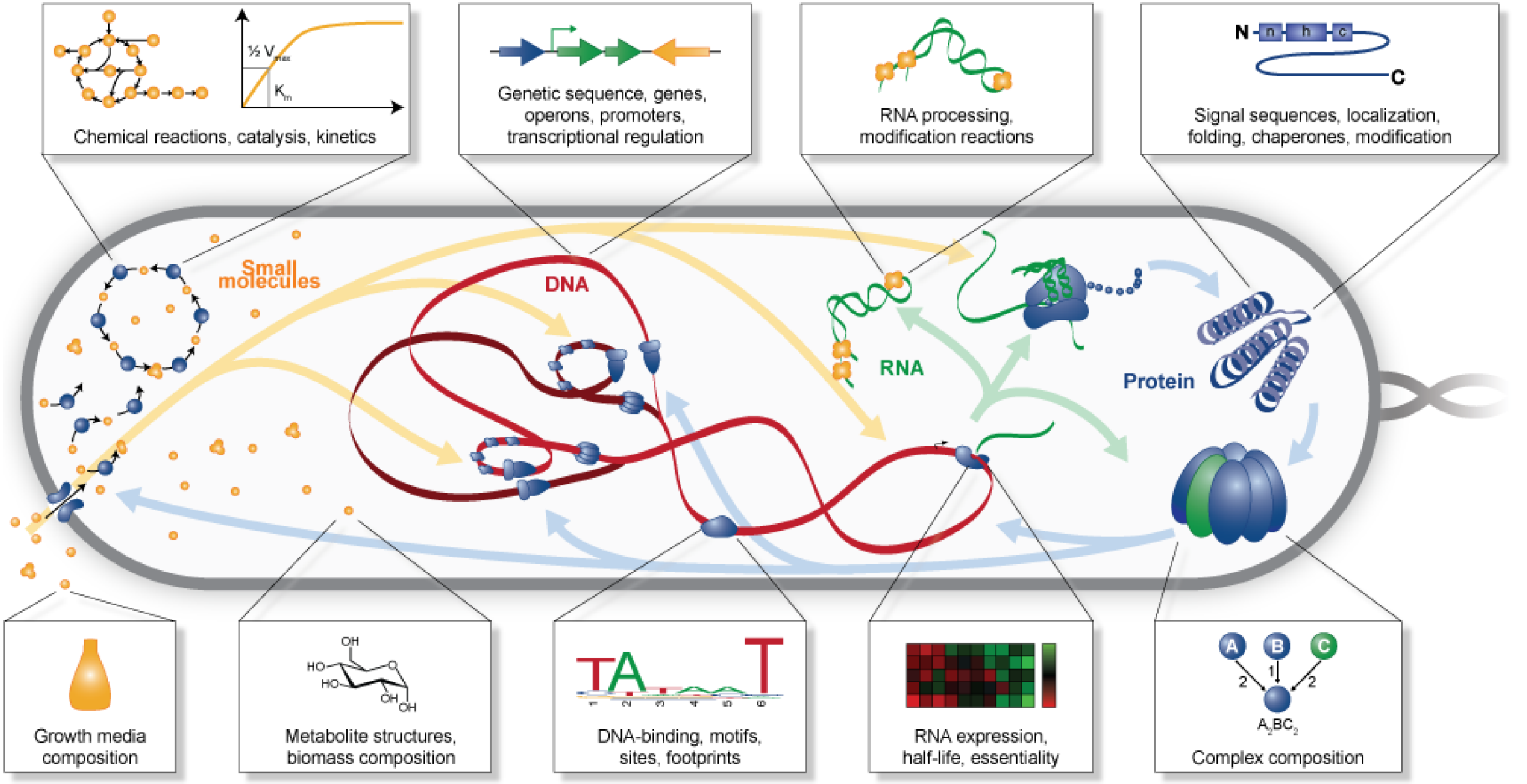
Replication

Transcription

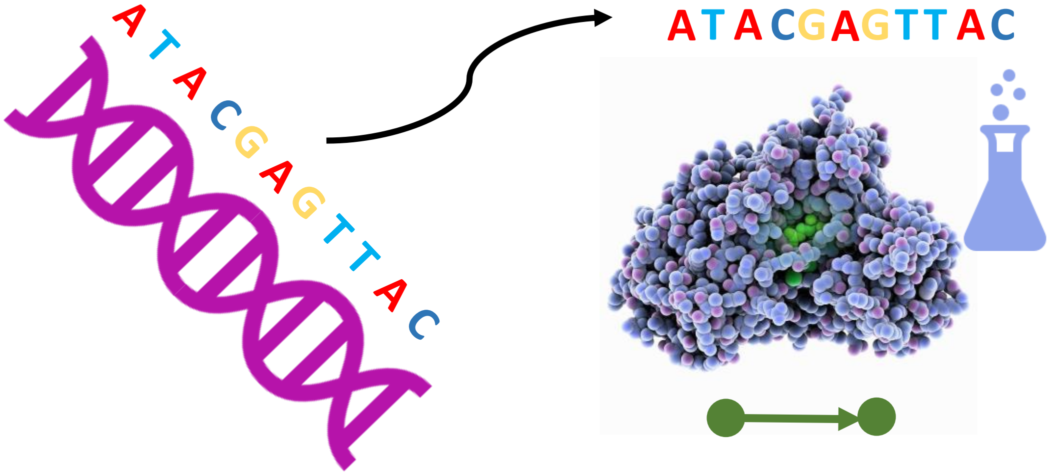
Metabolism

Time

Integration of data and systems



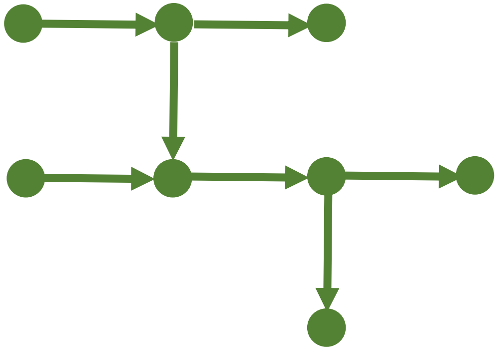
Karr, J. R. et al (2013): Cell, 150(2), 389–40

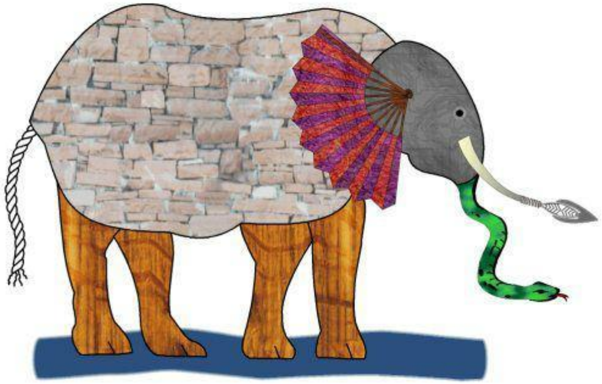


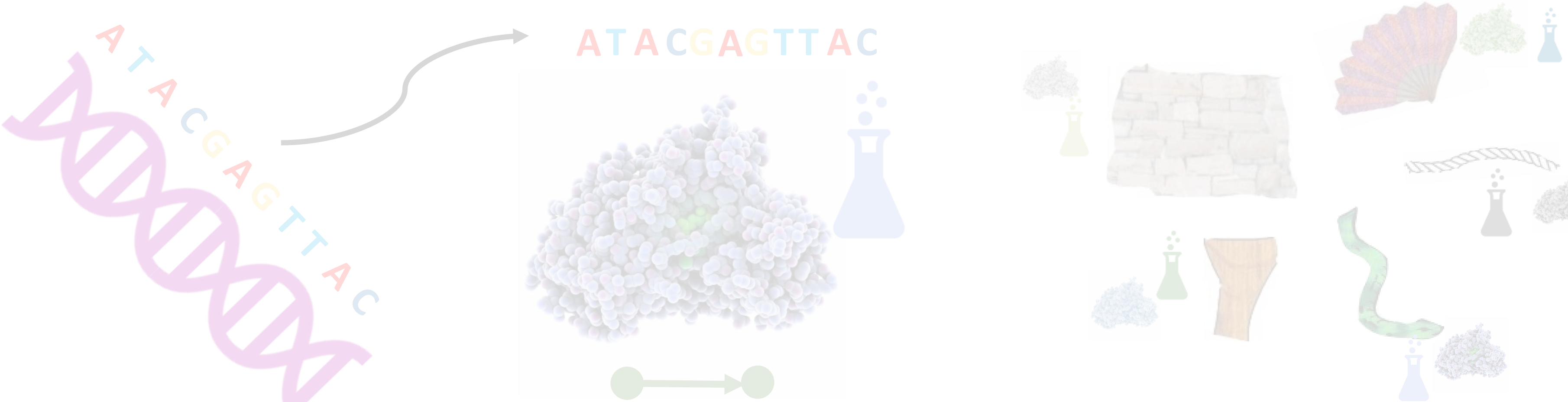
ATACGAGTTAC


ATACGAGTTAC
ATACGAGTTAC

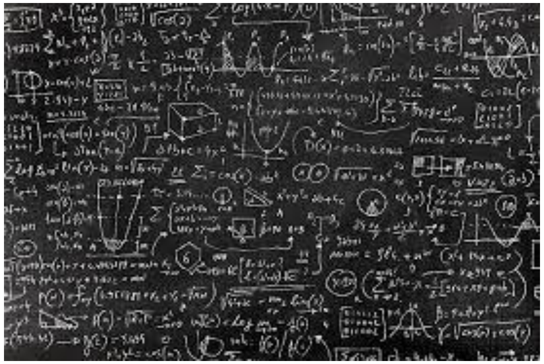
The whole system is different from the individual parts and includes emergent properties

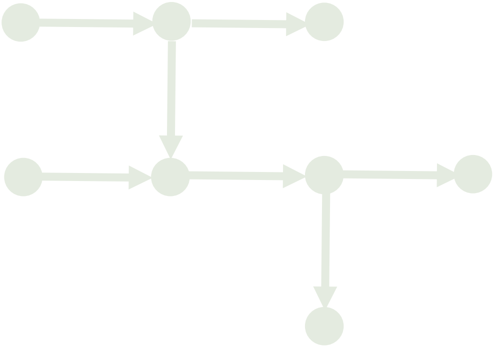


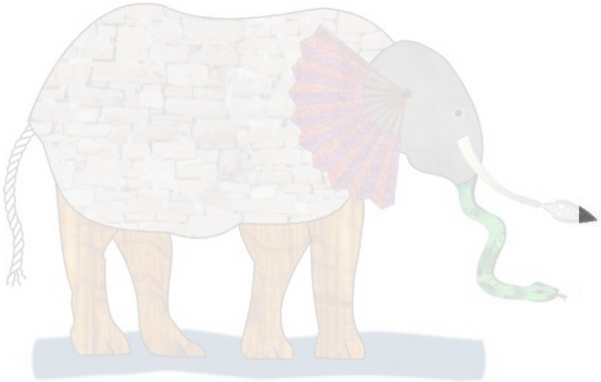




A mathematical model can represent the parts and interactions in the system

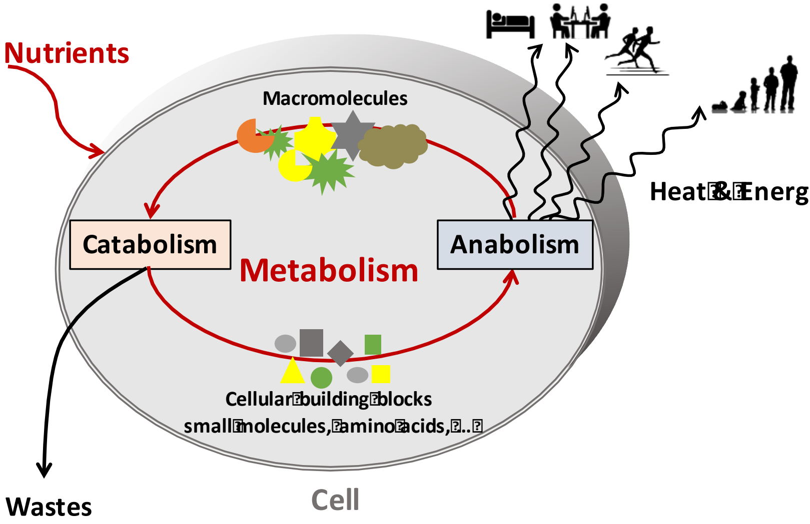






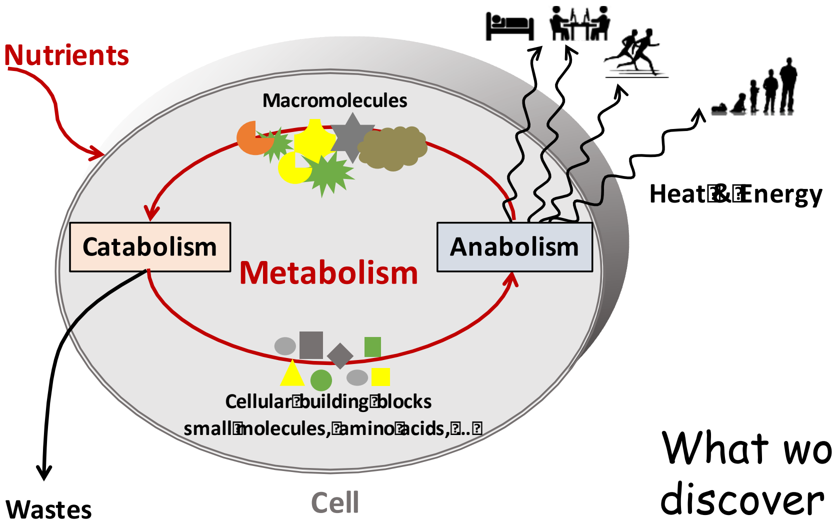
Systems Biology of Metabolism

Ensemble of biochemical reactions that occur in living organisms in order to maintain life



Metabolism

Understanding metabolism



What can we currently measure?

What does omics data mean individually?

Genomics: what *may* happen in cell
Transcriptomics: what *can* happen in cell
Proteomics: What *should* happen in cell
Metabolomics: What *is happening* in cell

How do we integrate the data to understand cell as a system?

What would you need to measure if you wanted to discover the causes of disease?

What are the mechanisms of drugs?

Why pathogens have large metabolic versatility and flexibility?

Design new therapies?

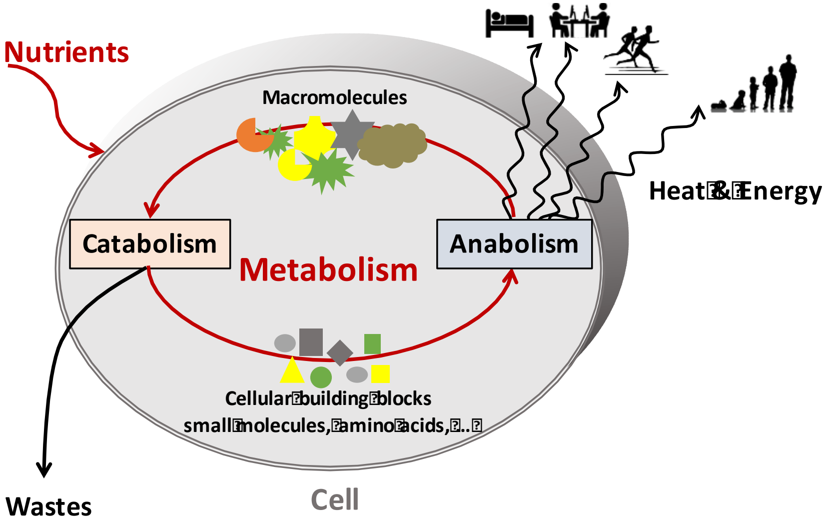
How metabolism is re-programmed in critical condition? Starvation?

How are the metabolic pathways different in micro-organisms?

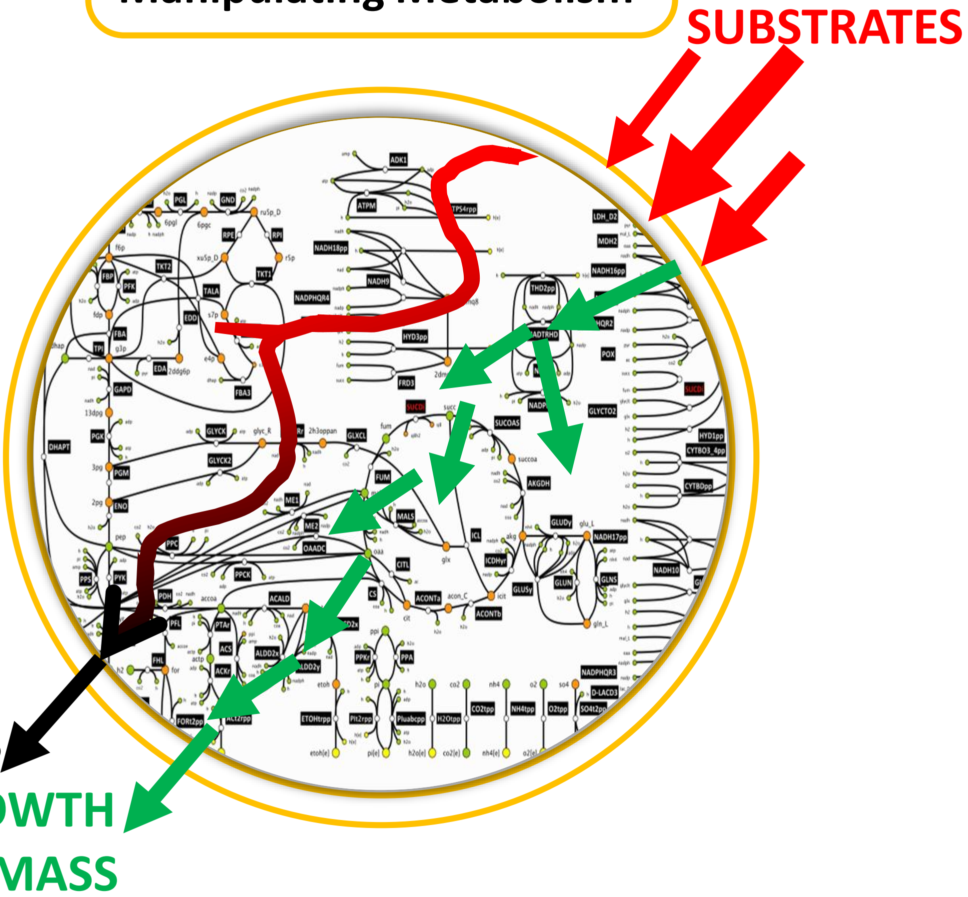
re-engineer organisms for new purposes? Cell factories?

Metabolism

Understanding metabolism



Manipulating Metabolism



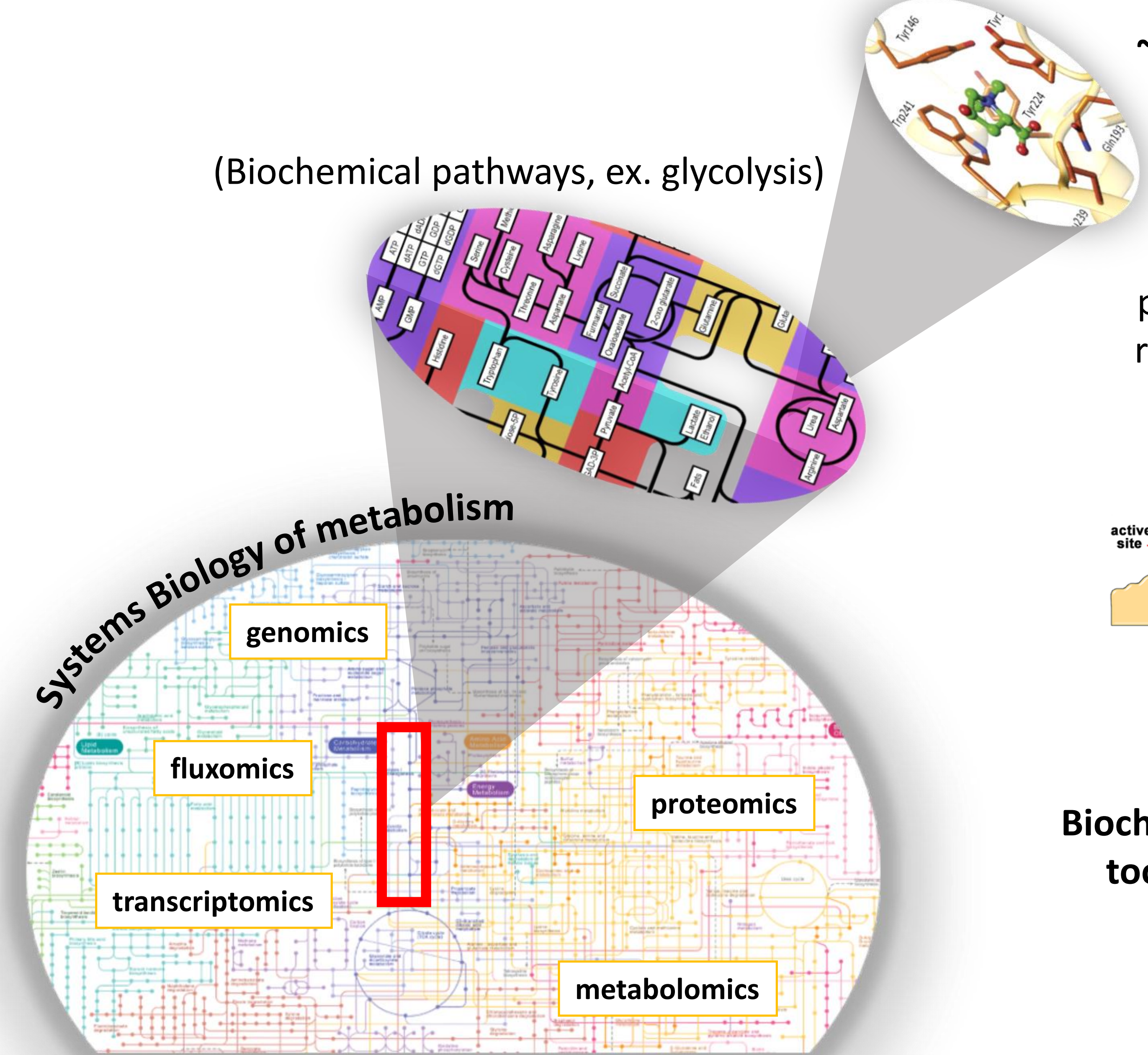
SUBSTRATES

PRODUCTS

GROWTH
BIOMASS

System components

(Biochemical pathways, ex. glycolysis)



Systems Biology of metabolism

genomics

fluxomics

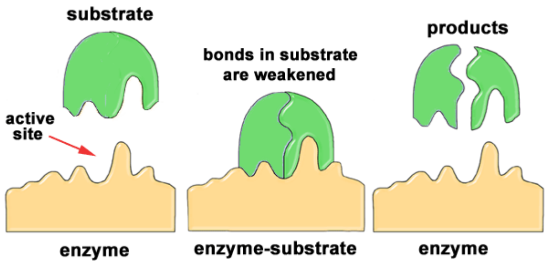
proteomics

transcriptomics

metabolomics

~5800 enzymatic reactions

proteins that accelerate the rate of biochemical reaction

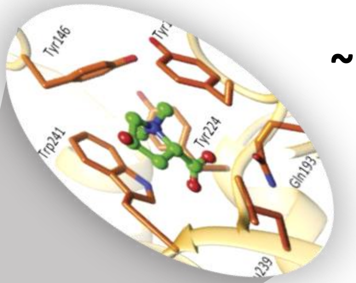


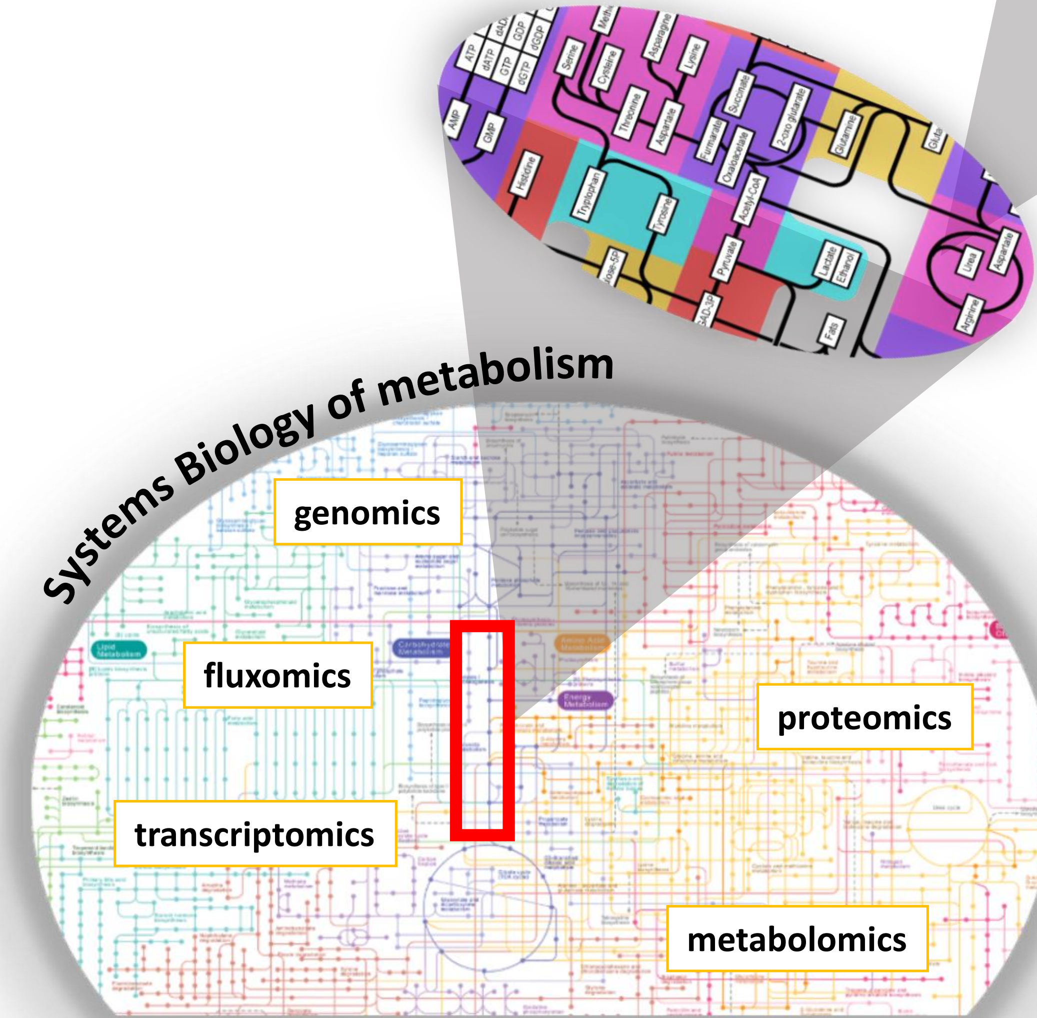
Biochemistry toolbox

Image: https://biochemianzunited.wordpress.com/

System components

(Biochemical pathways, ex. glycolysis)





Systems Biology of metabolism

genomics

fluxomics

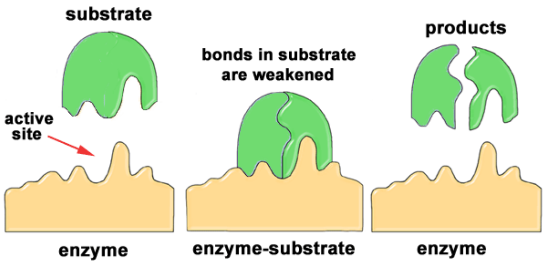
proteomics

transcriptomics

metabolomics

~5800 enzymatic reactions

proteins that accelerate the rate of biochemical reaction



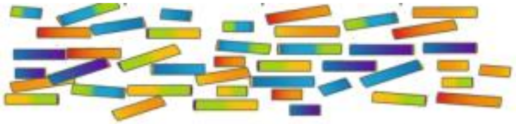
Biochemistry toolbox

Genome sequencing

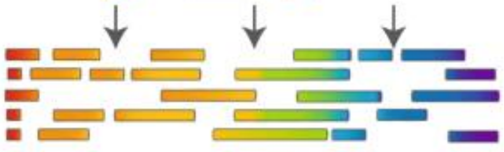
Image: https://biochemianzunited.wordpress.com/

From sequenced genome to enzymatic functions

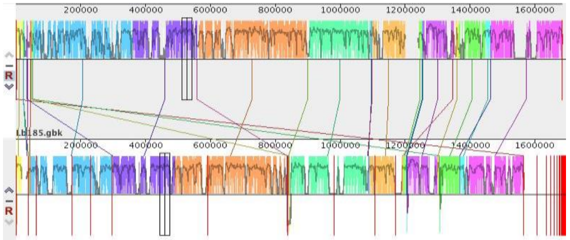
Sequence quality and trimming/filtering
(FastQC, Trimmomatic)



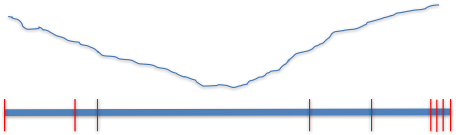
Genome assembly
(SPAdes)

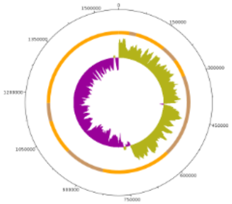


Contigs re-ordering and trimming
(Artemis, Blast, Mauve)

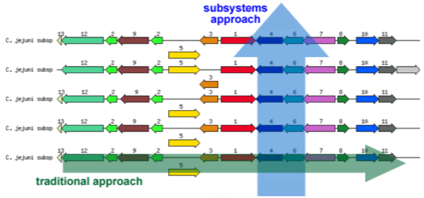


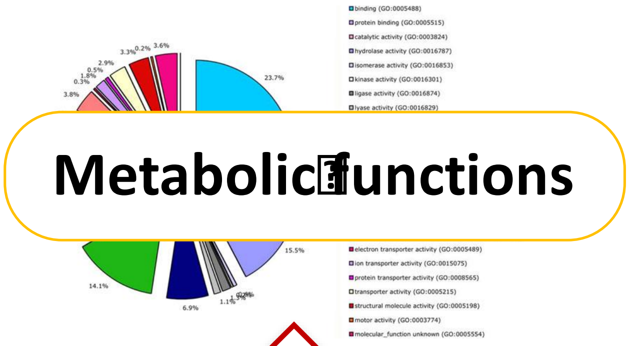
Assembly Validation
(Read coverage, GC-skew/content)





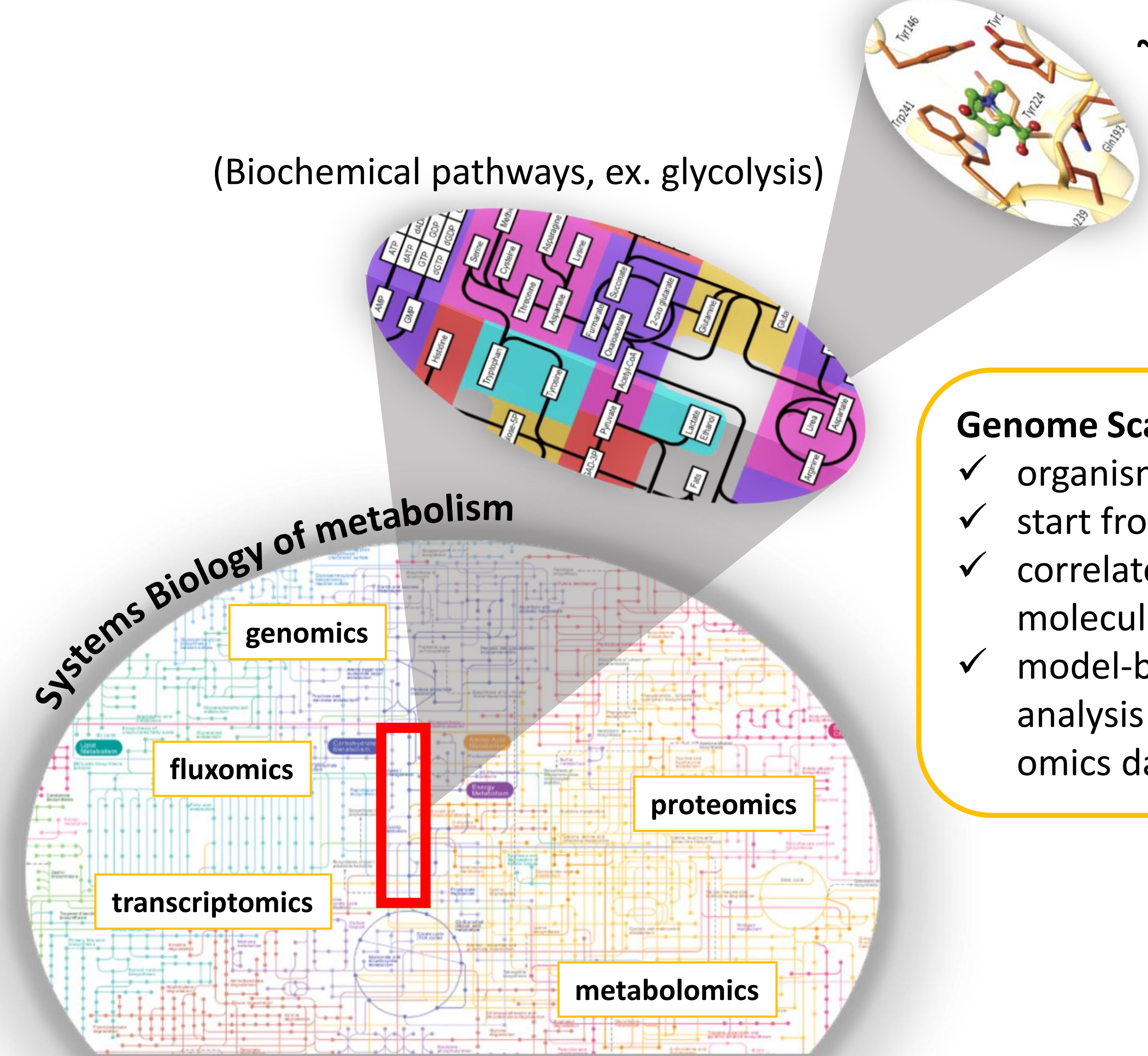
Annotation
(RAST, subsystem approach)





Metabolic functions

System components



~5800 enzymatic reactions

(Biochemical pathways, ex. glycolysis)

Systems Biology of metabolism

genomics

fluxomics

transcriptomics

proteomics

metabolomics

Genome Scale Metabolic Modeling

- ✓ organism-specific models
- ✓ start from sequenced genome
- ✓ correlate genome with molecular physiology
- ✓ model-based integration and analysis of large- and multi-scale omics data

Enzyme Classification

Enzyme Commission number (EC number)

Back in 1955, in the International Congress of Biochemistry in Brussels, the enzyme nomenclature scheme was developed and named: Enzyme Commission (EC) number

EC numbers do not specify enzymes (protein sequences) but enzyme-catalyzed reactions (functions).

If different enzymes (for instance from different organisms) catalyze the same reaction, they receive the same EC number.

http://www.sbcs.qmul.ac.uk/iubmb/enzyme/

EC number (i.j.k.l)

- **i.** general type of chemistry

EC 1 : Oxidoreductases	To catalyze oxidation/reduction reactions
EC 2 :Transferases	Transfer of a functional group from one substance to another
EC 3: Hydrolases	Formation of two products from a substrate by hydrolysis
EC 4: Lyases	Non-hydrolytic addition or removal of groups from substrates
EC 5: Isomerases	Intramolecule rearrangement
EC 6: Ligases	Join together two molecules by synthesis of new C-O, C-S, C-N or C-C bonds

EC number (i.j.k.l)

- **i.** general type of chemistry
- **j.** functional group being acted upon

EC 1 : Oxidoreductases	To catalyze oxidation/reduction reactions

1.1 act on the CH-OH group of donors
1.2 act on the aldehyde or oxo group of donors
1.3 act on the CH-CH group of donors
1.4 act on the CH-NH2 group of donors
1.5 act on CH-NH group of donors
1.6 act on NADH or NADPH
1.7 act on other nitrogenous compounds as donors
1.8 act on a sulfur group of donors
1.9 act on a heme group of donors
1.10 act on diphenols and related substances as donors
1.11 act on peroxide as an acceptor
1.12 act on hydrogen as donors
1.13 act on single donors with incorporation of molular oxygen
1.14 act on paired donors with incorporation of molular oxygen
1.15 act on superoxide radicals as acceptors
1.16 oxidize metal ions
1.17 act on CH or CH2 groups
1.18 act on iron-sulfur proteins as donors
1.19 act on reduced flavodoxin as a donor
1.20 act on phosphorus or arsenic in donors
1.21 act on X-H and Y-H to form an X-Y bond
1.97 includes other oxidoreductases

EC number (i.j.k.l)

- **i.** general type of chemistry
- **j.** functional group being acted upon
- **k.** cofactors the enzyme uses or specific information on the chemistry of the enzyme action

EC 1 : Oxidoreductases	To catalyze oxidation/reduction reactions

1.1 act on the CH-OH group of donors

1.1.1	With NAD or NADP as acceptor
1.1.2	With a cytochrome as acceptor
1.1.3	With oxygen as acceptor
1.1.4	With a disulfide as acceptor
1.1.5	With a quinone or similar compound as acceptor
1.1.9	With a copper protein as acceptor
1.1.98	With other, known, acceptors
1.1.99	With other acceptors

EC number (i.j.k.l)

- **i.** general type of chemistry
- **j.** functional group being acted upon
- **k.** cofactors the enzyme uses or specific information on the chemistry of the enzyme action
- **l.** specific substrates participating in the reaction

EC 1 : Oxidoreductases	To catalyze oxidation/reduction reactions

1.1 act on the CH-OH group of donors

1.1.1 With NAD or NADP as acceptor

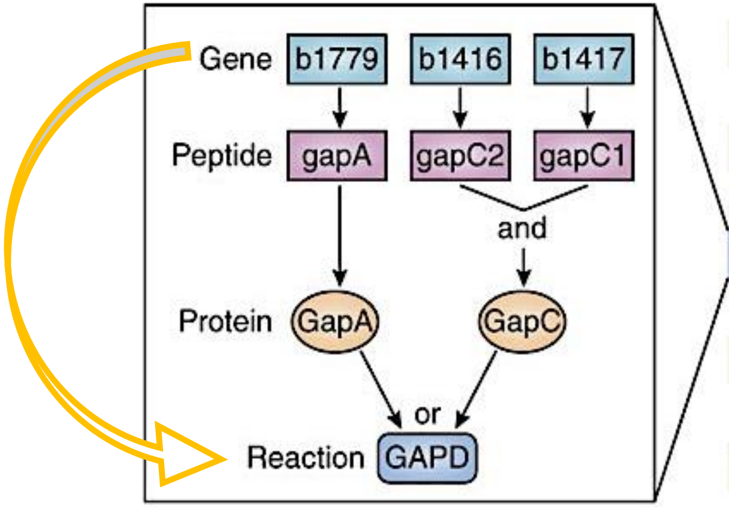
1.1.1.1: alcohol dehydrogenase
1.1.1.2: alcohol dehydrogenase (NADP+)
1.1.1.3: homoserine dehydrogenase
1.1.1.4: (R,R)-butanediol dehydrogenase
1.1.1.5: Transferred to * 1.1.1.303 and * 1.1.1.304
1.1.1.6: glycerol dehydrogenase
1.1.1.7: propanediol-phosphate dehydrogenase
1.1.1.8: glycerol-3-phosphate dehydrogenase (NAD+)
1.1.1.9: D-xylulose reductase
1.1.1.10: L-xylulose reductase
1.1.1.11: D-arabinitol 4-dehydrogenase
1.1.1.12: L-arabinitol 4-dehydrogenase
1.1.1.13: L-arabinitol 2-dehydrogenase
1.1.1.14: L-iditol 2-dehydrogenase
1.1.1.15: D-iditol 2-dehydrogenase
1.1.1.16: galactitol 2-dehydrogenase
1.1.1.17: mannitol-1-phosphate 5-dehydrogenase
1.1.1.18: inositol 2-dehydrogenase

Genome Scale Metabolic Models

✓ Systems Analysis of Metabolism

Sequenced and Annotated Genome

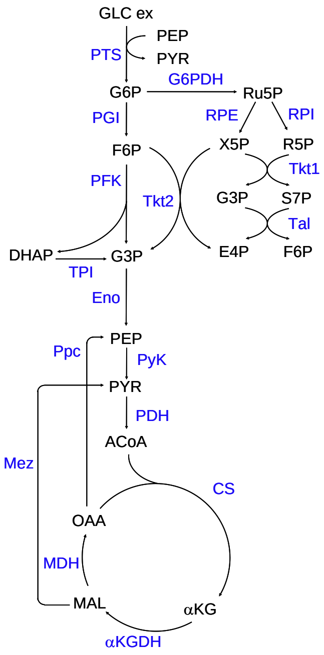
Functional Annotation



Metabolic Reactions -> Glycolysis pathway

Abbreviation	Glycolytic reactions
HEX1	[c]GLC + ATP → G6P + ADP+ H
PGI	[c]G6P ↔ F6P
PFK	[c]ATP + F6P → ADP + FDP + H
FBA	[c]FDP ↔ DHAP + G3P
TPI	[c]DHAP ↔ G3P
GAPD	[c]G3P + NAD + PI ↔ 13DPG + H + NADH
PGK	[c]13DPG + ADP ↔ 3PG + ATP
PGM	[c]3PG ↔ 2PG
ENO	[c]2PG ↔ H_2O + PEP
PYK	[c]ADP + H + PEP → ATP + PYR

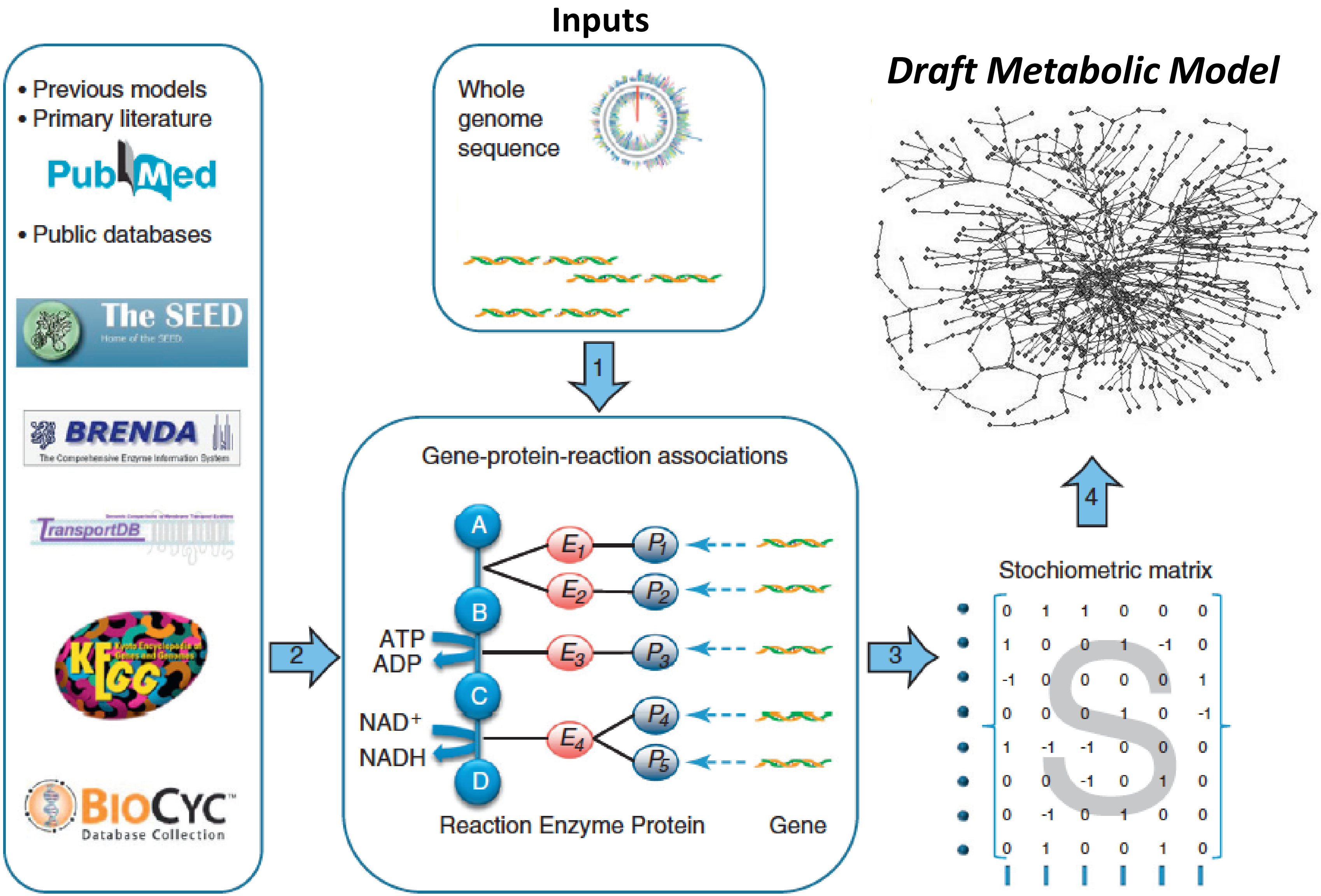
⋮



1000s of reactions for a single species

- Computers can keep this information "in mind" and analyze it in various ways.
- How to digitalize this information and show the interactions between metabolites and reactions?
- ***Interaction Matrices*** summarize any kind of interaction between elements of a system.

Genome Scale Model Reconstruction



Inputs

- Previous models
- Primary literature

- Public databases

Whole genome sequence

1

2

Gene-protein-reaction associations

Reaction Enzyme Protein Gene

3

Stochiometric matrix

$$\begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & -1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

4

Draft Metabolic Model

THE RAVEN TOOLBOX

Agren et al., 2013, The RAVEN Toolbox and Its Use for Generating a Genome-scale Metabolic Model for Penicillium chrysogenum. PLoS Comput Biol

Genome Scale Model Reconstruction

Data Needed

End Product

Automated Reconstruction

RAVEN

Mathematical Representation

Draft Model

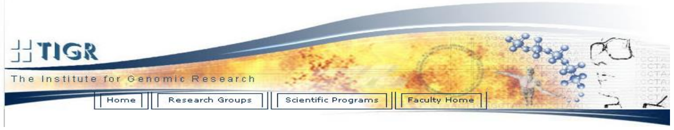
Image adopted from Agren et al., 2013, PLoS Comput Biol 9(3): e1002980

Genome Online Databases

Genome Databases

NCBI Entrez Gene

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=gene


TIGR The Institute for Genomic Research

http://pathema.tigr.org/tigr-scripts/CMR/CmrHomePage.cgi

GOLD

http://genomesonline.org/index2.htm

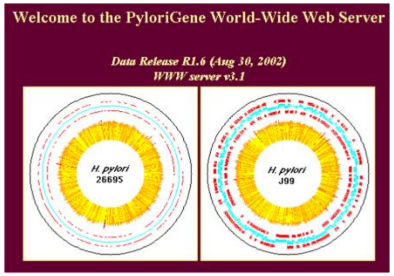
Vega

http://vega.sanger.ac.uk/index.html

TIGR CMR Comprehensive Microbial Resource

http://cmr.tigr.org/tigr-scripts/CMR/CmrHomePage.cgi

Organism-specific databases


Welcome to the PyloriGene World-Wide Web Server

Data Release R1.6 (Aug 30, 2002)
WWW server v3.1

H. pylori: http://genolist.pasteur.fr/PyloriGene/

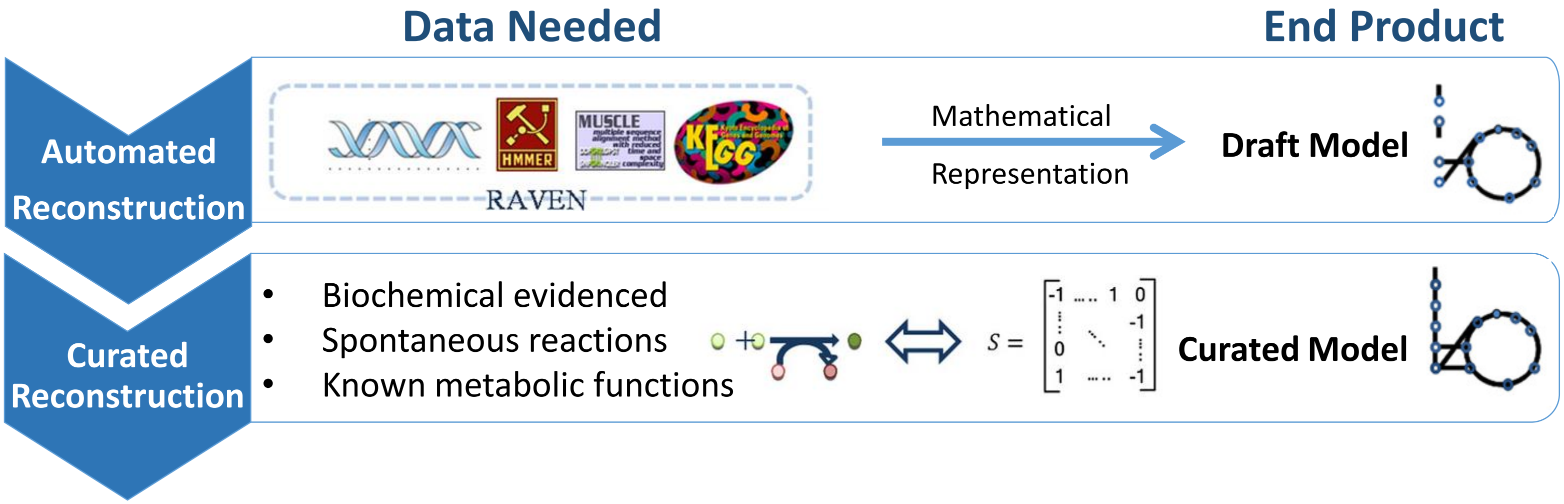
EcoCyc™ Encyclopedia of *Escherichia coli* K-12 Genes and Metabolism

http://ecocyc.org/

SGD *Saccharomyces* Genome Database

http://www.yeastgenome.org/

Genome Scale Model Reconstruction



Data Needed

End Product

Automated Reconstruction

RAVEN

Mathematical Representation

Draft Model

Curated Reconstruction

- Biochemical evidenced
- Spontaneous reactions
- Known metabolic functions

Curated Model

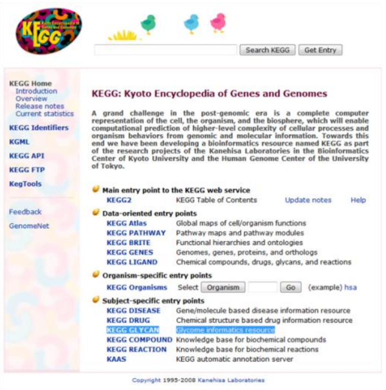
Image adopted from Agren et al., 2013, PLoS Comput Biol 9(3): e1002980

Biochemical databases

- **Enzyme databases:**
  - KEGG:

  http://www.genome.jp/kegg/
  - BRENDA:

  http://www.brenda-enzymes.info/



KEGG: Kyoto Encyclopedia of Genes and Genomes

A grand challenge in the post-genomic era is a complete computer representation of the cell, the organism, and the biosphere, which will enable computational prediction of higher-level complexity of cellular processes and organism behaviors from genomic and molecular information. Towards this end we have been developing a bioinformatics resource named KEGG as part of the research projects of the Kanehisa Laboratories in the Bioinformatics Center of Kyoto University and the Human Genome Center of the University of Tokyo.

Main entry point to the KEGG web service

KEGG2	KEGG Table of Contents	Update notes	Help

Data-oriented entry points

KEGG Atlas	Global maps of cell/organism functions
KEGG PATHWAY	Pathway maps and pathway modules
KEGG BRITE	Functional hierarchies and ontologies
KEGG GENES	Genomes, genes, proteins, and orthologs
KEGG LIGAND	Chemical compounds, drugs, glycans, and reactions

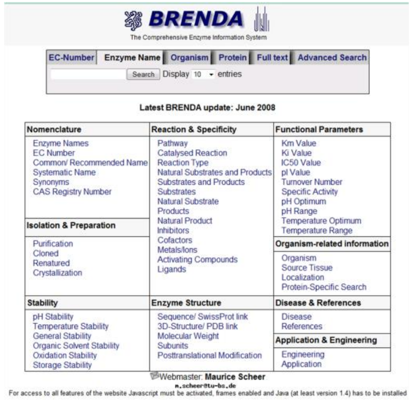
Organism-specific entry points

KEGG Organisms Select Organism Go (example) hsa

Subject-specific entry points

KEGG DISEASE	Gene/molecule based disease information resource
KEGG DRUG	Chemical structure based drug information resource
KEGG GLYCAN	Glycome informatics resource
KEGG COMPOUND	Knowledge base for biochemical compounds
KEGG REACTION	Knowledge base for biochemical reactions
KAAS	KEGG automatic annotation server

Copyright 1995-2008 Kanehisa Laboratories



Latest BRENDA update: June 2008

Nomenclature	Reaction & Specificity	Functional Parameters
Enzyme Names EC Number Common/ Recommended Name Systematic Name Synonyms CAS Registry Number	Pathway Catalysed Reaction Reaction Type Natural Substrates and Products Substrates and Products Substrates Natural Substrate Products Natural Product Inhibitors Cofactors Metals/Ions Activating Compounds Ligands	Km Value Ki Value IC50 Value pI Value Turnover Number Specific Activity pH Optimum pH Range Temperature Optimum Temperature Range
Isolation & Preparation		**Organism-related information**
Purification Cloned Renatured Crystallization		Organism Source Tissue Localization Protein-Specific Search
Stability	**Enzyme Structure**	**Disease & References**
pH Stability Temperature Stability General Stability Organic Solvent Stability Oxidation Stability Storage Stability	Sequence/ SwissProt link 3D-Structure/ PDB link Molecular Weight Subunits Posttranslational Modification	Disease References **Application & Engineering** Engineering Application

Webmaster: **Maurice Scheer**
m.scheer@tu-bs.de

For access to all features of the website Javascript must be activated, frames enabled and Java (at least version 1.4) has to be installed

  - Both databases are great resources for biochemical reactions, but there information are organism-unspecific!

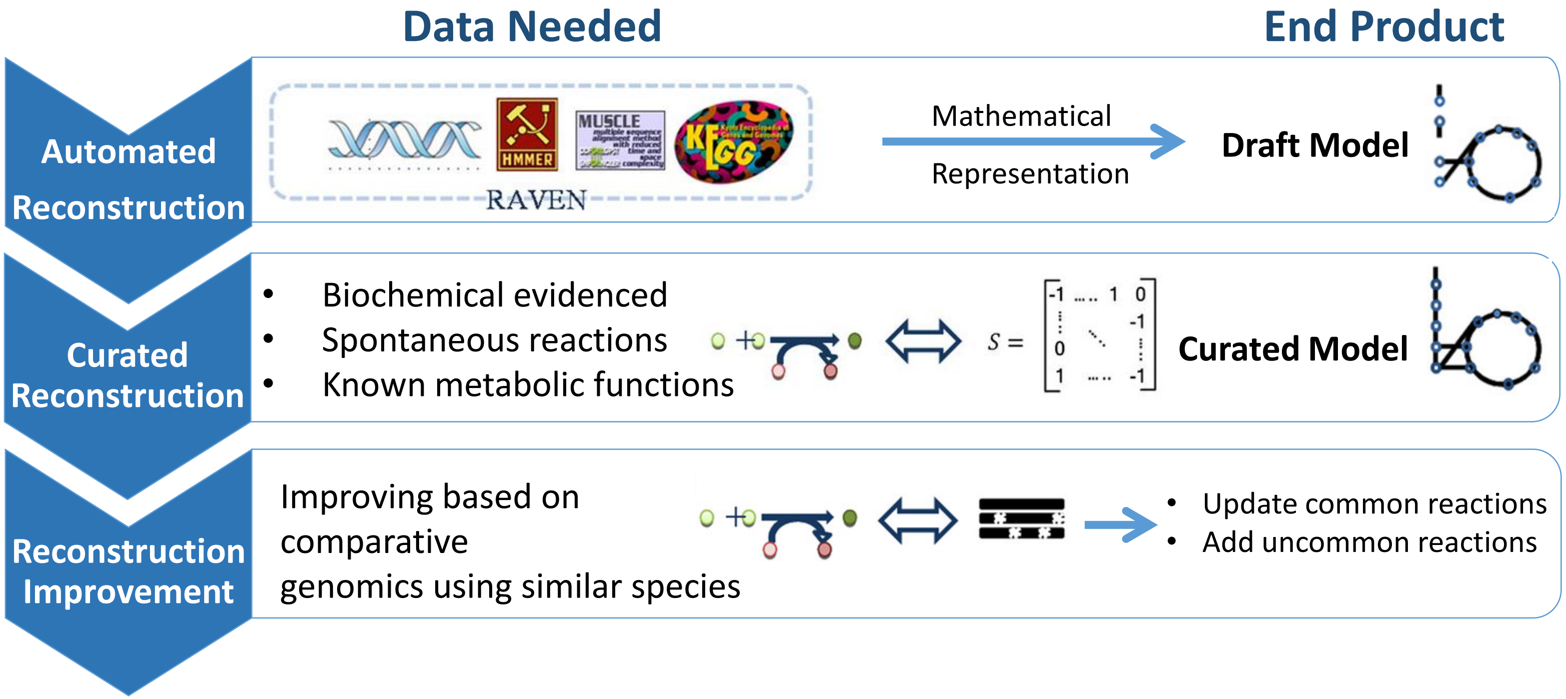
- **Transport database:**
  - Transport DB:

  http://www.membranetransport.org/

  - Transport Classification Database

http://www.tcdb.org/

Genome Scale Model Reconstruction



Data Needed

End Product

Automated Reconstruction

Mathematical Representation

Draft Model

Curated Reconstruction

- Biochemical evidenced
- Spontaneous reactions
- Known metabolic functions

$$S = \begin{bmatrix} -1 & \cdots & 1 & 0 \\ \vdots & & & -1 \\ 0 & & \ddots & \vdots \\ 1 & \cdots & & -1 \end{bmatrix}$$

Curated Model

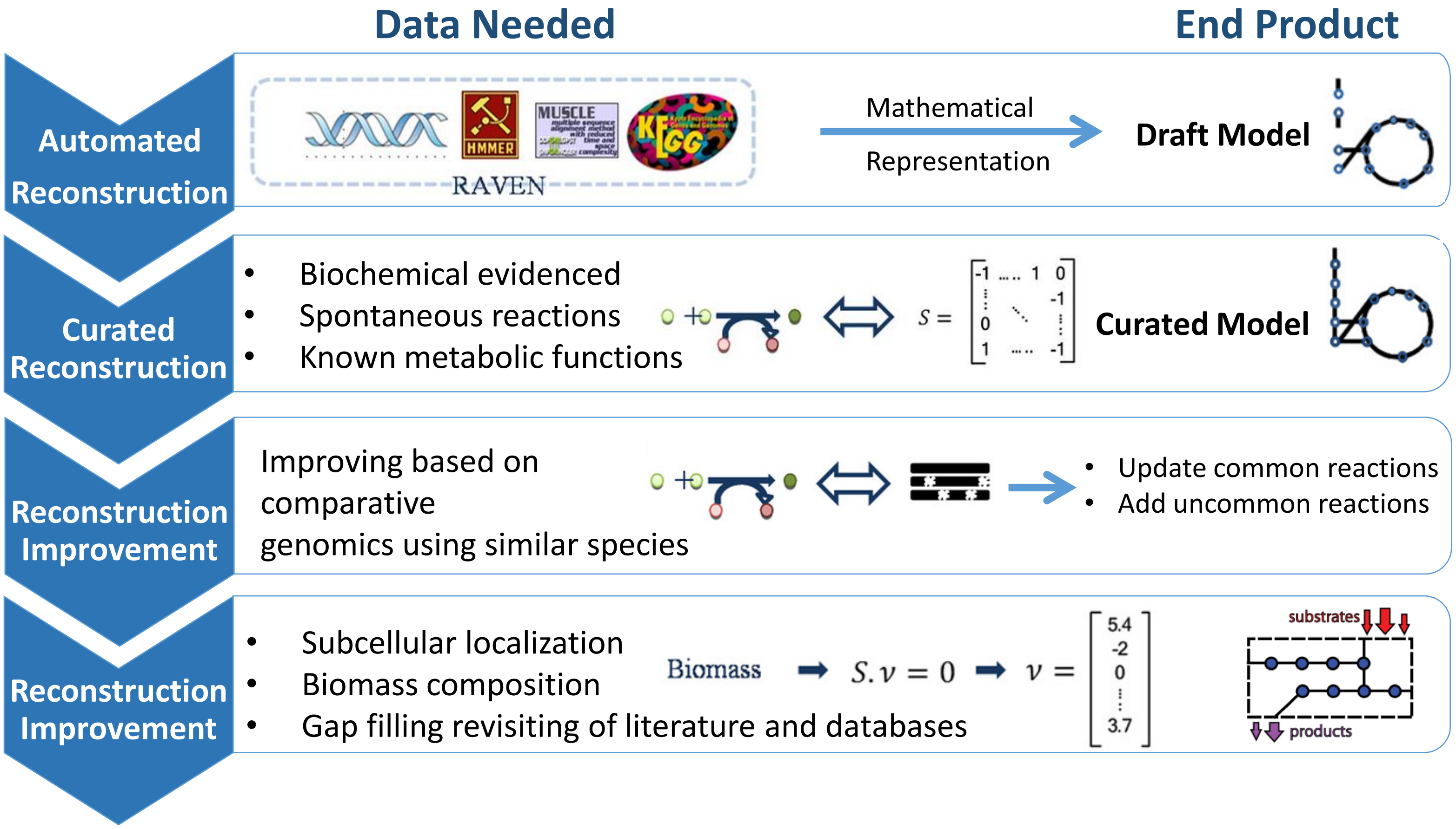
Reconstruction Improvement

Improving based on comparative genomics using similar species

- Update common reactions
- Add uncommon reactions

Image adopted from Agren et al., 2013, PLoS Comput Biol 9(3): e1002980

Genome Scale Model Reconstruction



Data Needed

End Product

Automated Reconstruction

Mathematical Representation

Draft Model

Curated Reconstruction

- Biochemical evidenced
- Spontaneous reactions
- Known metabolic functions

$$S = \begin{bmatrix} -1 & \dots & 1 & 0 \\ \vdots & \ddots & & -1 \\ 0 & & \ddots & \vdots \\ 1 & \dots & & -1 \end{bmatrix}$$

Curated Model

Reconstruction Improvement

Improving based on comparative genomics using similar species

- Update common reactions
- Add uncommon reactions

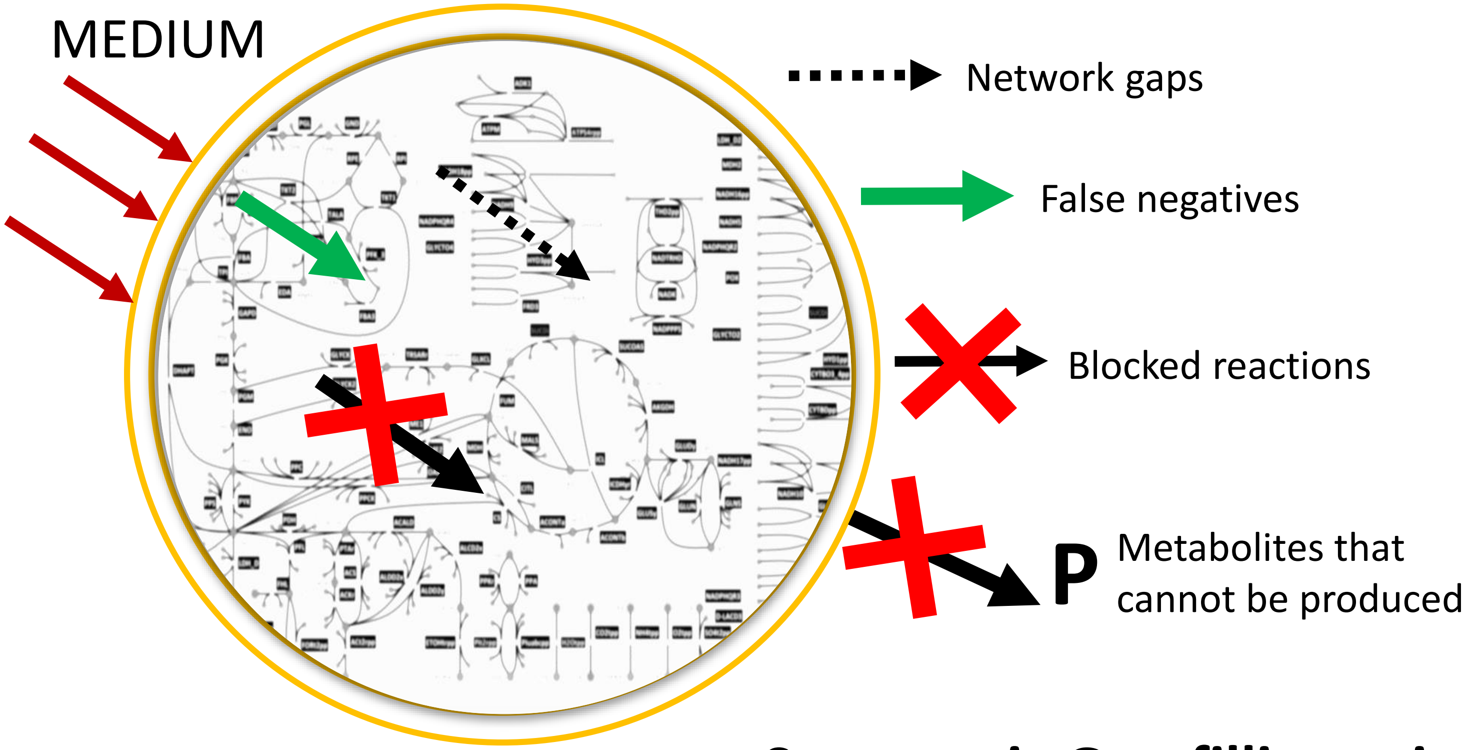
Reconstruction Improvement

- Subcellular localization
- Biomass composition
- Gap filling revisiting of literature and databases

$$\text{Biomass} \rightarrow S.v = 0 \rightarrow v = \begin{bmatrix} 5.4 \\ -2 \\ 0 \\ \vdots \\ 3.7 \end{bmatrix}$$

Image adopted from Agren et al., 2013, PLoS Comput Biol 9(3): e1002980

Knowledge Gaps in metabolic networks



MEDIUM

Network gaps

False negatives

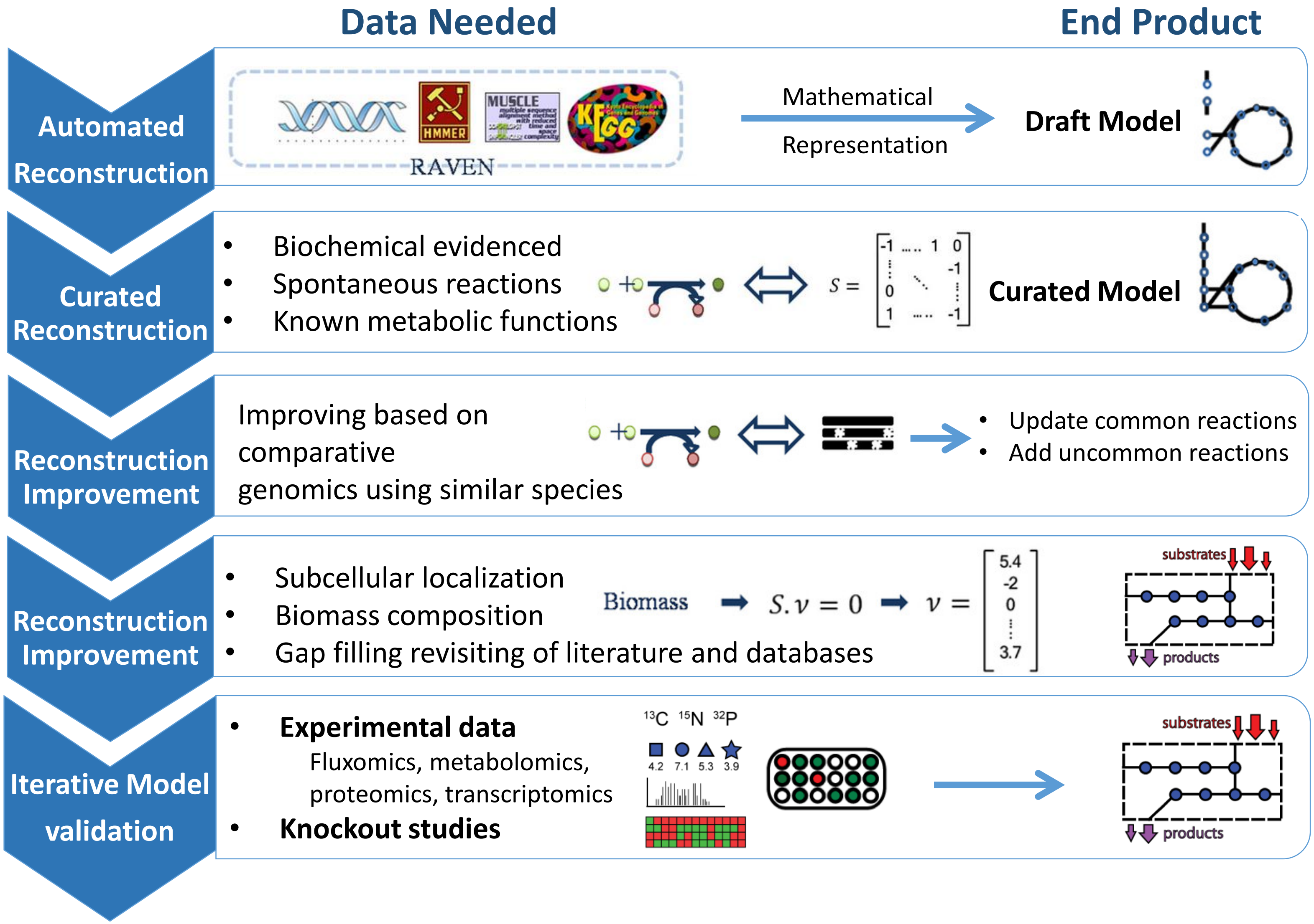
Blocked reactions

P Metabolites that cannot be produced

Systematic Gap-filling using:

- ✓ Other annotation platforms
- ✓ *Close organism*
- ✓ KEGG database
- ✓ ATLAS of Biochemistry

Genome Scale Model Reconstruction



Data Needed **End Product**

Automated Reconstruction

RAVEN

Mathematical Representation → **Draft Model**

Curated Reconstruction

- Biochemical evidenced
- Spontaneous reactions
- Known metabolic functions

$$\Leftrightarrow \quad S = \begin{bmatrix} -1 & \cdots & 1 & 0 \\ \vdots & & & -1 \\ 0 & & \ddots & \vdots \\ 1 & \cdots & & -1 \end{bmatrix}$$

Curated Model

Reconstruction Improvement

Improving based on comparative genomics using similar species

- Update common reactions
- Add uncommon reactions

Reconstruction Improvement

- Subcellular localization
- Biomass composition
- Gap filling revisiting of literature and databases

$$\text{Biomass} \rightarrow S.v = 0 \rightarrow v = \begin{bmatrix} 5.4 \\ -2 \\ 0 \\ \vdots \\ 3.7 \end{bmatrix}$$

Iterative Model validation

- **Experimental data**
  Fluxomics, metabolomics, proteomics, transcriptomics
- **Knockout studies**

Image adopted from Agren et al., 2013, PLoS Comput Biol 9(3): e1002980

Genome Scale Metabolic Models

Genome Scale Models are driven from **sequenced genome**

- They started with the reconstruction of metabolic network of microbes
- Now they exist for several mice strains and human

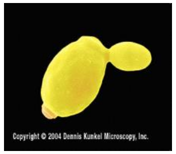
Industrially relevant organisms

E. coli

- 2712 Reactions
- 1516 Genes

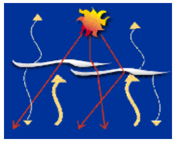
S. cerevisiae

- 1402 Reactions
- 910 Genes



M. barkeri

- 619 Reactions
- 692 Genes



G. sulfurreducens

- 608 Reactions
- 588 Genes

B. subtilis

- 1020 Reactions
- 844 Genes

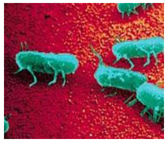
Pathogens

S. aureus

- 640 Reactions
- 619 Genes

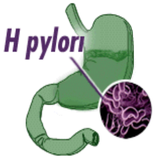
S. typhimurium

- 2545 Reactions
- 1271 Genes



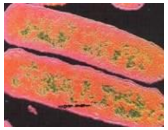
H. pylori

- 558 Reactions
- 341 Genes



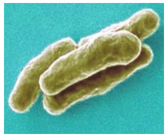
H. influenzae

- 472 Reactions
- 376 Genes



M. tuberculosis

- 939 Reactions
- 661 Genes



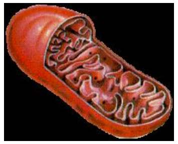
Mammalian cells

H. sapiens

- 10600 Reactions
- 2248 Genes

Human Mitochondria

- 218 Reactions



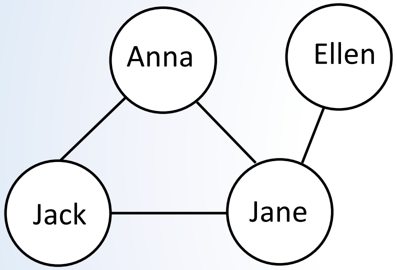
Red blood cell

- 39 Reactions

Interaction Matrices

Represents the **relationship** between all the **elements** (constituent) of a system.

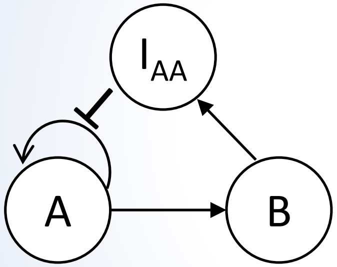
Example 1: social networks



⇔

	Anna	Ellen	Jack	Jane
Anna	0	0	1	1
Ellen	0	0	0	1
Jack	1	0	0	1
Jane	1	1	1	0

Example 2: gene Regulatory Network



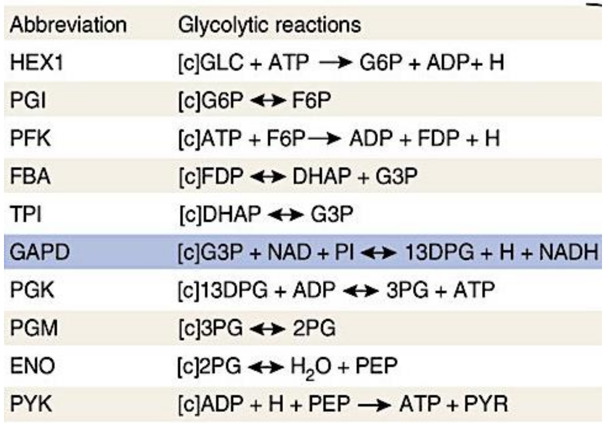
⇔

	A	B	I_{AA}
A	1	1	0
B	0	0	1
I_{AA}	-1	0	0

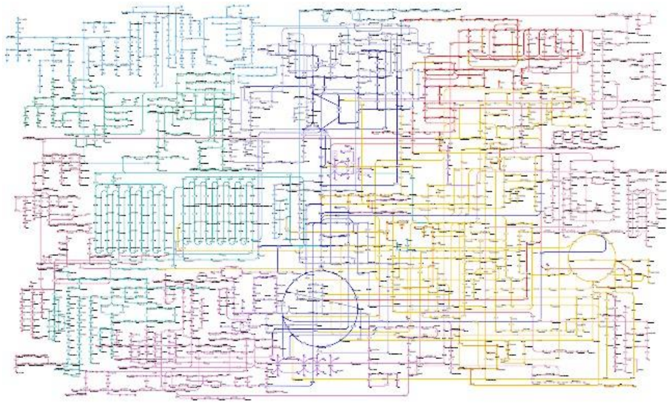
Stoichiometric Matrix

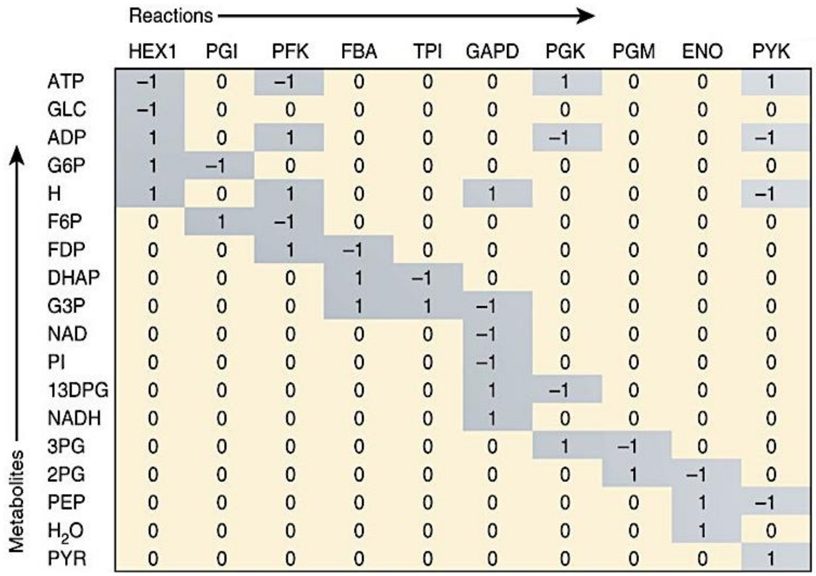
Represents the relationship between *all the* ***metabolites*** in *all the* ***reactions*** in a metabolic network

Metabolic Reactions



Abbreviation	Glycolytic reactions
HEX1	[c]GLC + ATP → G6P + ADP+ H
PGI	[c]G6P ↔ F6P
PFK	[c]ATP + F6P → ADP + FDP + H
FBA	[c]FDP ↔ DHAP + G3P
TPI	[c]DHAP ↔ G3P
GAPD	[c]G3P + NAD + PI ↔ 13DPG + H + NADH
PGK	[c]13DPG + ADP ↔ 3PG + ATP
PGM	[c]3PG ↔ 2PG
ENO	[c]2PG ↔ H_2O + PEP
PYK	[c]ADP + H + PEP → ATP + PYR





Reactions →

↑ Metabolites

	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

$$\begin{array}{cc} & \text{Reactions} \\ \text{Metabolites} & \begin{bmatrix} 1 & 0 & \cdots & 1 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & -1 \end{bmatrix} \end{array}$$

Closing Remarks

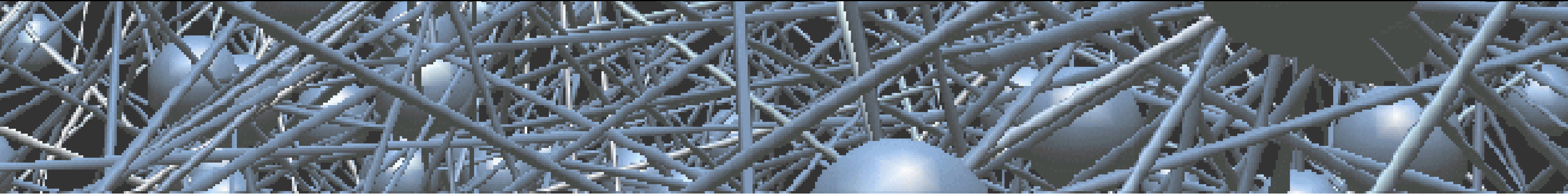
Genome-scale metabolic model is a platform that

agrees with experimentally observed data

allows testing hypotheses and answer metabolically relevant questions

allows generating new hypothesis for experimental validation

RECONSTRUCTION OF METABOLIC NETWORKS
&
FLUX BALANCE ANALYSIS



Principles and Applications of Systems Biology

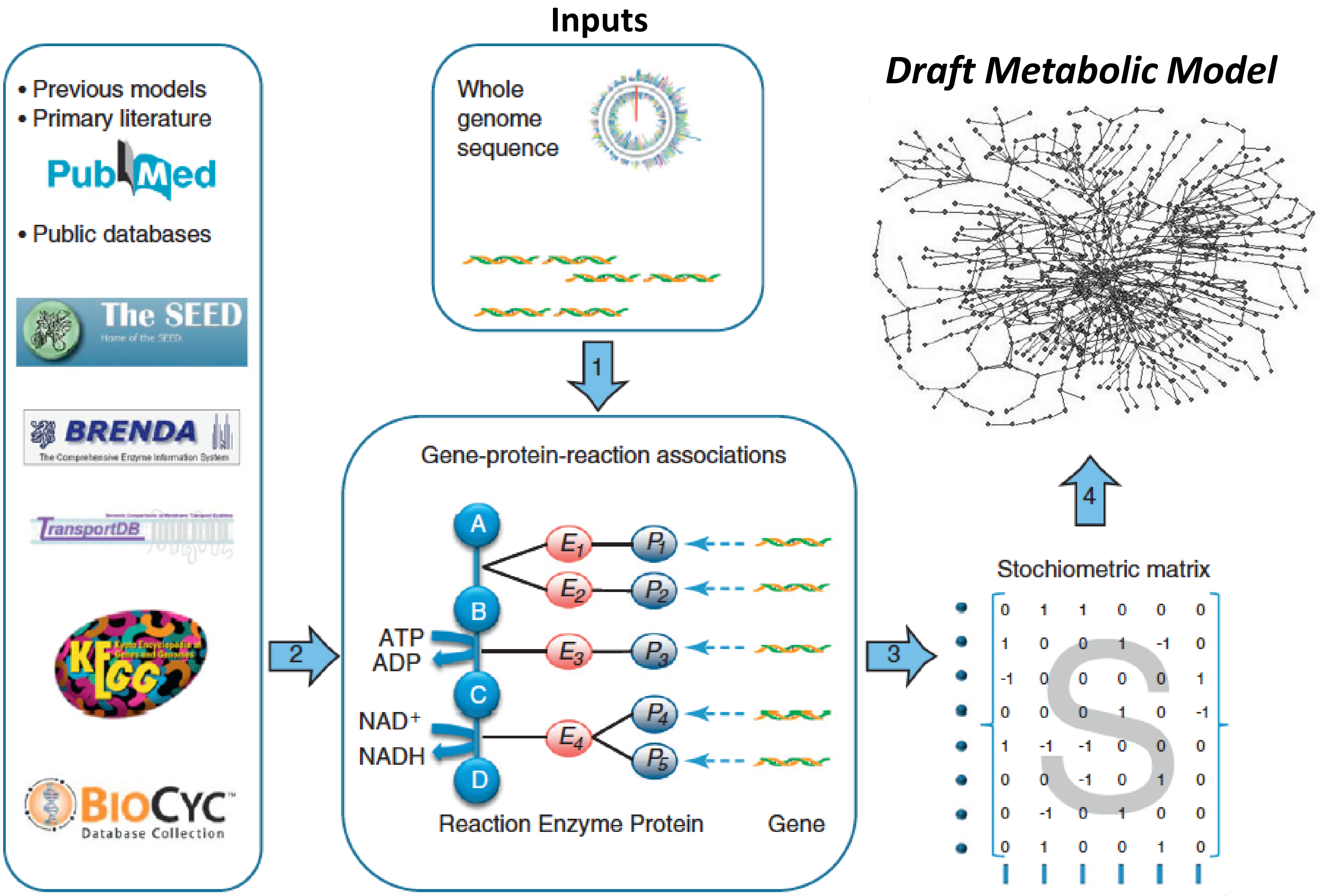
September 2025
Vassily Hatzimanikatis

LECTURE OBJECTIVES

- GENERAL CONCEPTS OF MASS BALANCES IN METABOLIC NETWORKS
- METABOLIC NETWORK RECONSTRUCTION:
  - GENERAL WORKFLOW
  - BASIC CONCEPTS
- INTRODUCTION TO FLUX BALANCE ANALYSIS
  - LINEAR PROGRAMMING & OPTIMIZATION IN METABOLIC NETWORKS

- *WHAT IS THE OPTIMAL GROWTH OF BACTERIUM?*

Genome Scale Model Reconstruction



Inputs

- Previous models
- Primary literature
- Public databases

Whole genome sequence

1

Gene-protein-reaction associations

Reaction Enzyme Protein Gene

2

3

Stochiometric matrix

$$\begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & -1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

4

Draft Metabolic Model

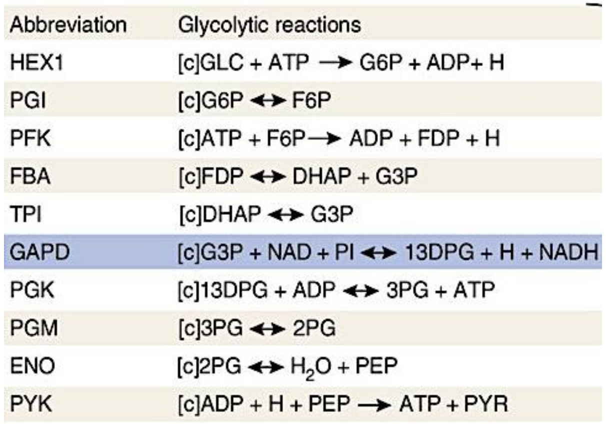
Agren et al., 2013, The RAVEN Toolbox and Its Use for Generating a Genome-scale Metabolic Model for Penicillium chrysogenum. PLoS Comput Biol

THE RAVEN TOOLBOX

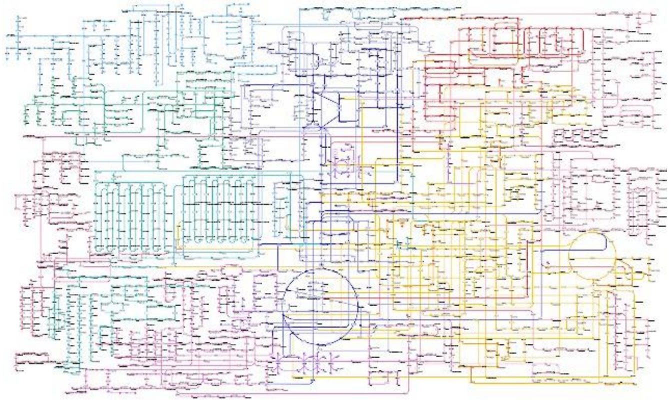
Stoichiometric Matrix

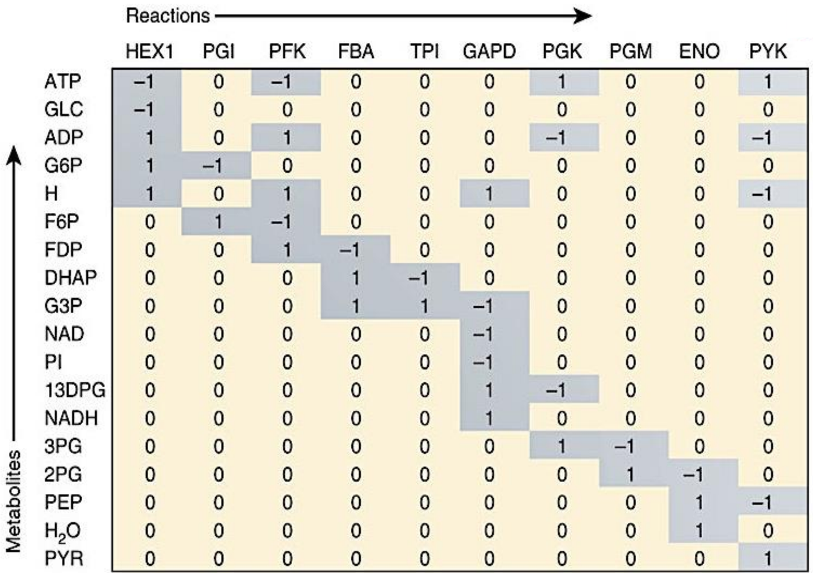
Represents the relationship between *all the* ***metabolites*** in *all the* ***reactions*** in a metabolic network

Metabolic Reactions



Abbreviation	Glycolytic reactions
HEX1	[c]GLC + ATP → G6P + ADP+ H
PGI	[c]G6P ↔ F6P
PFK	[c]ATP + F6P → ADP + FDP + H
FBA	[c]FDP ↔ DHAP + G3P
TPI	[c]DHAP ↔ G3P
GAPD	[c]G3P + NAD + PI ↔ 13DPG + H + NADH
PGK	[c]13DPG + ADP ↔ 3PG + ATP
PGM	[c]3PG ↔ 2PG
ENO	[c]2PG ↔ H_2O + PEP
PYK	[c]ADP + H + PEP → ATP + PYR





Reactions →

Metabolites ↑

	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

$$\begin{array}{cc} & \text{Reactions} \\ \text{Metabolites} & \begin{bmatrix} 1 & 0 & \cdots & 1 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & -1 \end{bmatrix} \end{array}$$

Mass Balances

Metabolic Reactions

$R1: \ Substrate \rightarrow A$

$R2: \ A \rightarrow B$

$R3: \ A \rightarrow C$

$R4: \ B \rightarrow product1$

$R5: \ C \rightarrow product2$

S matrix

Reactions

Metabolites	**R1**	**R2**	**R3**	**R4**	**R5**
Substrate	-1	0	0	0	0
A	1	-1	-1	0	0
B	0	1	0	-1	0
C	0	0	1	0	-1
Product 1	0	0	0	1	0
Product 2	0	0	0	0	1

How much **substrates ($d_{susbstrtae}$/dt)** for how much **products (d_{P1}/dt)**?

Fluxes

Metabolic Reactions

$R1: \ Substrate \rightarrow A$

$R2: \ A \rightarrow B$

$R3: \ A \rightarrow C$

$R4: \ B \rightarrow product1$

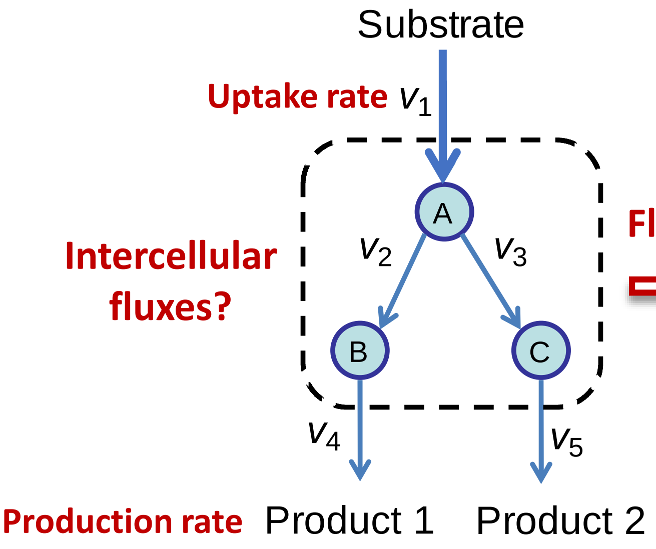
$R5: \ C \rightarrow product2$

S matrix

Reactions

Metabolites	**R1**	**R2**	**R3**	**R4**	**R5**
Substrate	-1	0	0	0	0
A	1	-1	-1	0	0
B	0	1	0	-1	0
C	0	0	1	0	-1
Product 1	0	0	0	1	0
Product 2	0	0	0	0	1

How much **substrates ($d_{susbstrtae}$/dt)** for how much **products (d_{P1}/dt)**?



Flux values ?

✓ v_1 $R1: \ Substrate \rightarrow A$

? v_2 $R2: \ A \rightarrow B$

? v_3 $R3: \ A \rightarrow C$

✓ v_4 $R4: \ B \rightarrow product1$

✓ v_5 $R5: \ C \rightarrow product2$

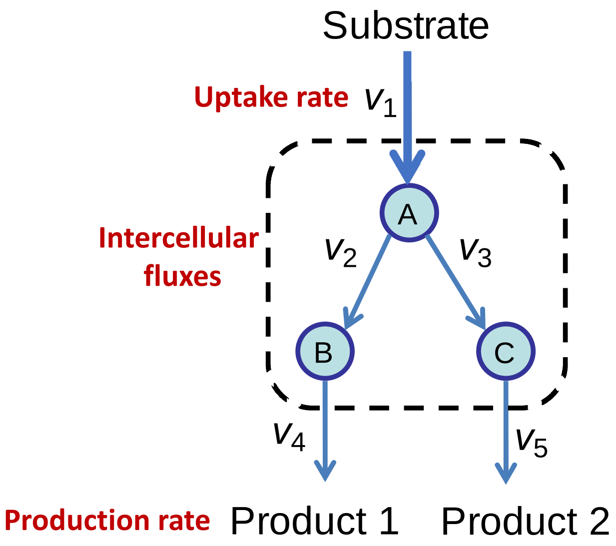
Why calculating the intercellular fluxes?

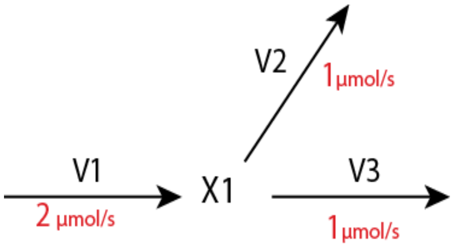
How to calculate the intercellular fluxes?

Metabolic Fluxes Analysis

quantitatively analyze metabolic pathways (calculating/estimating all the missing fluxes).

- ✓ Simulation of the effect of environmental or genetic changes
- ✓ Identification of important/critical/bottleneck reactions or pathways
- ✓ Understanding/controlling the pathway branching points





Mass balances

equations

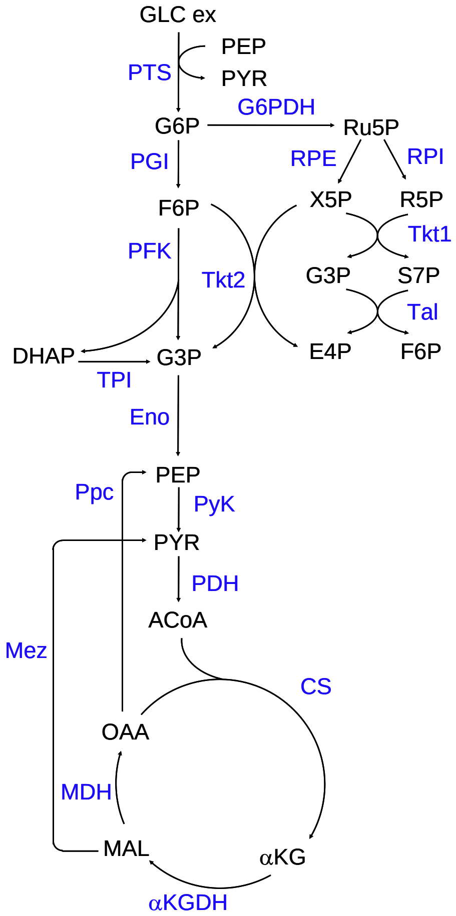
$$\frac{d[A]}{dt} = v_1 - v_2 - v_3$$

$$\frac{d[B]}{dt} = v_2 - v_4$$

$$\frac{d[C]}{dt} = v_3 - v_5$$

Central carbon metabolism of *E. coli*

“Fluxes in many biological networks cannot be uniquely determined”



Measurable flux

$$
\begin{array}{l}
\text{G6P} \\ \text{F6P} \\ \text{DHAP} \\ \text{G3P} \\ \text{PEP} \\ \text{PYR} \\ \text{Ru5P} \\ \text{X5P} \\ \text{R5P} \\ \text{S7P} \\ \text{E4P} \\ \text{ACoA} \\ \alpha\text{KG} \\ \text{MAL} \\ \text{OAA}
\end{array}
\begin{bmatrix}
1 & -1 & & & & & -1 & & & & & & & & & & & \\
 & 1 & -1 & & & & & & & & 1 & 1 & & & & & & \\
 & & 1 & -1 & & & & & & & & & & & & & & \\
 & & 1 & 1 & -1 & & & & & 1 & -1 & 1 & & & & & & \\
-1 & & & & 1 & -1 & & & & & & & & & & & -1 & \\
1 & & & & & 1 & & & & & & & -1 & & & & & 1 \\
 & & & & & & 1 & -1 & -1 & & & & & & & & & \\
 & & & & & & & 1 & & -1 & & -1 & & & & & & \\
 & & & & & & & & 1 & -1 & & & & & & & & \\
 & & & & & & & & & 1 & -1 & & & & & & & \\
 & & & & & & & & & & 1 & -1 & & & & & & \\
 & & & & & & & & & & & & 1 & -1 & & & & \\
 & & & & & & & & & & & & & 1 & -1 & & & \\
 & & & & & & & & & & & & & & 1 & -1 & & -1 \\
 & & & & & & & & & & & & & -1 & & 1 & 1 &
\end{bmatrix}
\cdot
\begin{bmatrix}
v_{PTS} \\ v_{PGI} \\ v_{PFK} \\ v_{TPI} \\ v_{Eno} \\ v_{Pyk} \\ v_{G6PDH} \\ v_{RPE} \\ v_{RPI} \\ v_{TK1} \\ v_{Tal} \\ v_{TK2} \\ v_{PDH} \\ v_{CS} \\ v_{\alpha KGDH} \\ v_{MDH} \\ v_{Ppc} \\ v_{Mez}
\end{bmatrix}
= 0
$$

$$\frac{d[F6P]}{dt} = v_{PGI} - v_{PFK} + v_{Tal} + v_{TK2} = 0$$

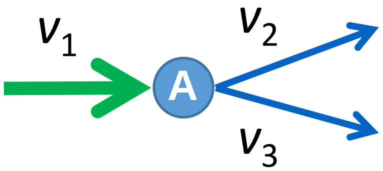
⋮

18 mass balance equations

(Quasi-)Steady state assumption

- Metabolism does not change with respect to time
  - All intermediates do not accumulate
  - The sum of influxes equals the sum of effluxes

$$v_1 = v_2 + v_3$$



Under the ***steady state assumption:***

$$\frac{d[A]}{dt} = v_1 - v_2 - v_3$$
$$\frac{d[B]}{dt} = v_2 - v_4$$
$$\frac{d[C]}{dt} = v_3 - v_5$$

$$\Rightarrow \frac{d[A]}{dt} = \frac{d[B]}{dt} = \frac{d[C]}{dt} = 0 \Rightarrow$$

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

(Quasi-)Steady state assumption

The mass balance equations can be described as follows:

$$\frac{d[A]}{dt} = \boxed{v_1 - v_2 - v_3} = 0$$

$$\frac{d[B]}{dt} = v_2 - v_4 = 0$$

$$\frac{d[C]}{dt} = v_3 - v_5 = 0$$

matrix multiplication

Reactions

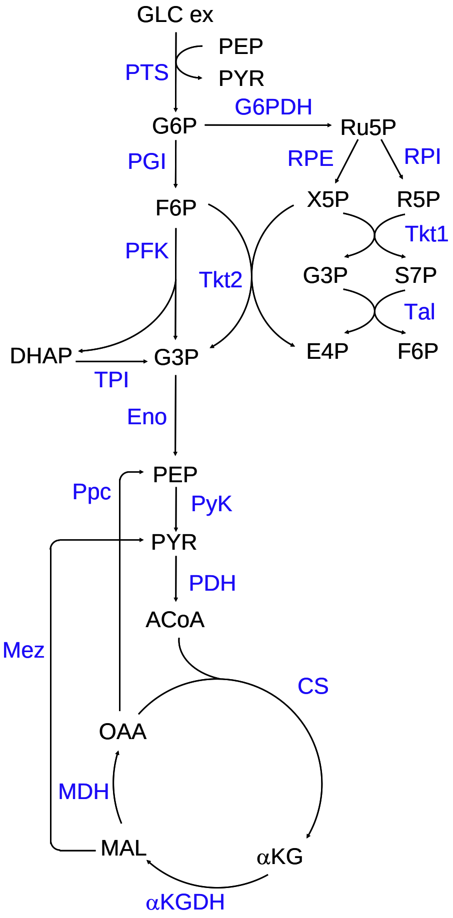
Metabolites

	R1	**R2**	**R3**	**R4**	**R5**
Substrat	-1	0	0	0	0
A	1	-1	-1	0	0
B	0	1	0	-1	0
C	0	0	1	0	-1
Product 1	0	0	0	1	0
Product 2	0	0	0	0	1

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{bmatrix} = 0 \Rightarrow Sv = 0$$

Central carbon metabolism of *E. coli*

“Fluxes in many biological networks cannot be uniquely determined”



Measurable flux

$$
\begin{array}{l}
\text{G6P}\\ \text{F6P}\\ \text{DHAP}\\ \text{G3P}\\ \text{PEP}\\ \text{PYR}\\ \text{Ru5P}\\ \text{X5P}\\ \text{R5P}\\ \text{S7P}\\ \text{E4P}\\ \text{ACoA}\\ \alpha\text{KG}\\ \text{MAL}\\ \text{OAA}
\end{array}
\begin{bmatrix}
1 & -1 & & & & & -1 & & & & & & & & & & & \\
 & 1 & -1 & & & & & & & & 1 & 1 & & & & & & \\
 & & 1 & -1 & & & & & & & & & & & & & & \\
 & & 1 & 1 & -1 & & & & & 1 & -1 & 1 & & & & & & \\
-1 & & & & 1 & -1 & & & & & & & & & & & -1 & \\
1 & & & & & 1 & & & & & & & -1 & & & & & 1 \\
 & & & & & & 1 & -1 & -1 & & & & & & & & & \\
 & & & & & & & 1 & & -1 & & -1 & & & & & & \\
 & & & & & & & & 1 & -1 & & & & & & & & \\
 & & & & & & & & & 1 & -1 & & & & & & & \\
 & & & & & & & & & & 1 & -1 & & & & & & \\
 & & & & & & & & & & & & 1 & -1 & & & & \\
 & & & & & & & & & & & & & 1 & -1 & & & \\
 & & & & & & & & & & & & & & 1 & -1 & & -1 \\
 & & & & & & & & & & & & & -1 & & 1 & 1 &
\end{bmatrix}
\cdot
\begin{bmatrix}
v_{PTS}\\ v_{PGI}\\ v_{PFK}\\ v_{TPI}\\ v_{Eno}\\ v_{Pyk}\\ v_{G6PDH}\\ v_{RPE}\\ v_{RPI}\\ v_{TK1}\\ v_{Tal}\\ v_{TK2}\\ v_{PDH}\\ v_{CS}\\ v_{\alpha KGDH}\\ v_{MDH}\\ v_{Ppc}\\ v_{Mez}
\end{bmatrix}
= 0
$$

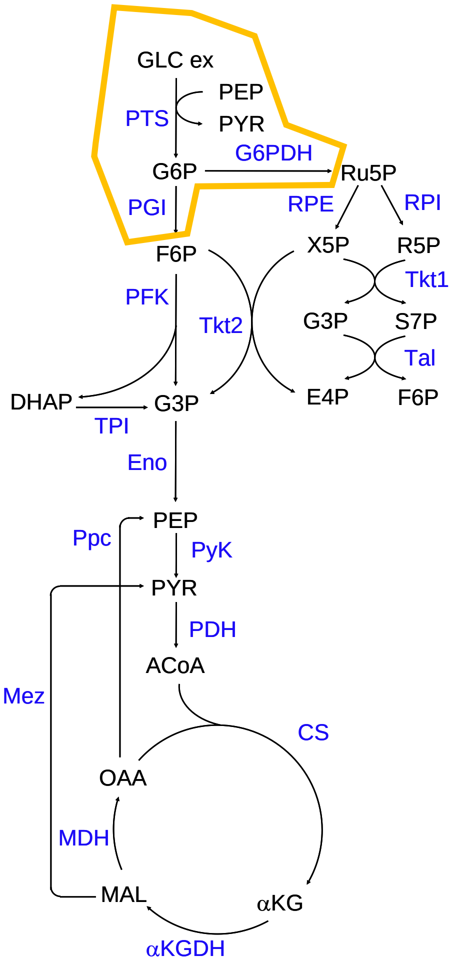
$$\frac{d[F6P]}{dt} = v_{PGI} - v_{PFK} + v_{Tal} + v_{TK2} = 0$$

.
.
.

18 mass balance equations

Underdetermined Systems

“there are fewer equations than unknowns”



$$V^{\checkmark}_{PTS} = V_{PGI} + V_{G6PDH}$$

$$10 = V_{PGI} + V_{G6PDH}$$

2 UNKNOWN
1 EQUATION

10 & 0
1 & 9
9 & 1
2 & 8
3 & 7
4 & 6
6 & 4
5 & 5
-100 & 110
1200 & -1190
.
.
.

space of hundreds of feasible solutions

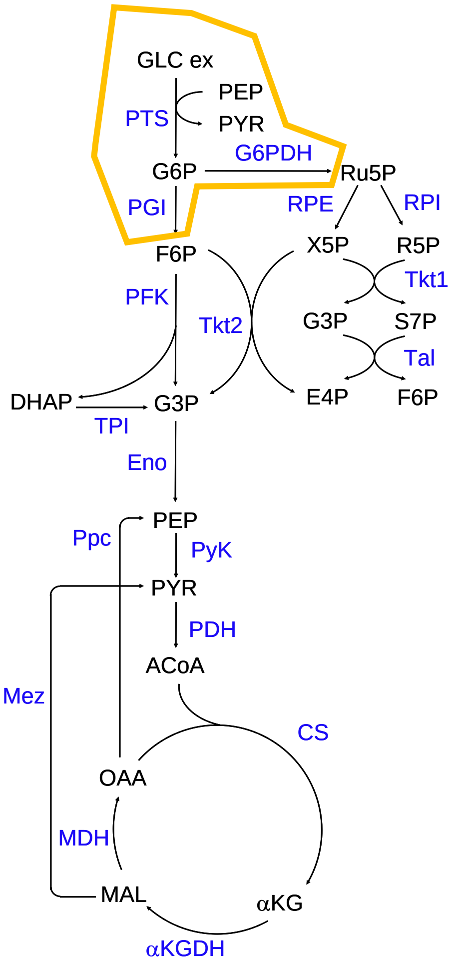
Constraint-based Modeling

Achieving a certain **objective**
in a system
that is shaped by the **defined constraints.**

*There are many different ways from Lausanne to Zurich, but they will all result in reaching Zurich (**objective**) and if I say*
*I should definitely pass through Fribourg & Bern (**defined constraints**)*
some results will be eliminated.

Undetermined Systems

“there are fewer equations than unknowns”



$$V_{PTS} = V_{PGI} + V_{G6PDH}$$

$$10 = V_{PGI} + V_{G6PDH}$$

2 UNKNOWN- 1 EQUATION

~~10 & 0~~

~~1 & 9~~

Trascriptomics ~~9 & 1~~

2 & 8

3 & 7

~~-20 & 30~~ Thermodynamics

6 & 4

Metabolomics 5 & 5

~~-100 & 110~~ Fluxomics

~~1200 & -1190~~

.

.

.

space of hundreds of feasible solutions

Flux Balance Analysis

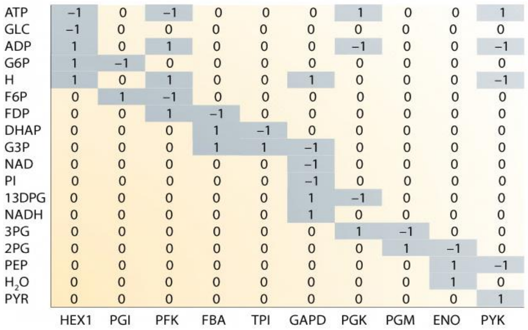
from math to growth

Back to metabolic models ….

Equality constrains:

Stoichiometry:

$$Sv = 0$$



	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

Inequality constraints:

Which substrates are available?

Which of them are less abundant?

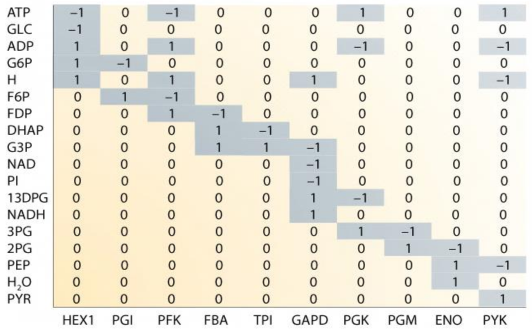
What can be secreted?

Constraints of metabolism

Equality constrains:

Stoichiometry:

$$Sv = 0$$



	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

Inequality constraints:

Uptakes:

$$-10 \leq v_{EX\_glu\_e} \leq 0$$

$$-20 \leq v_{EX\_O2\_e} \leq 0$$

Secretions:

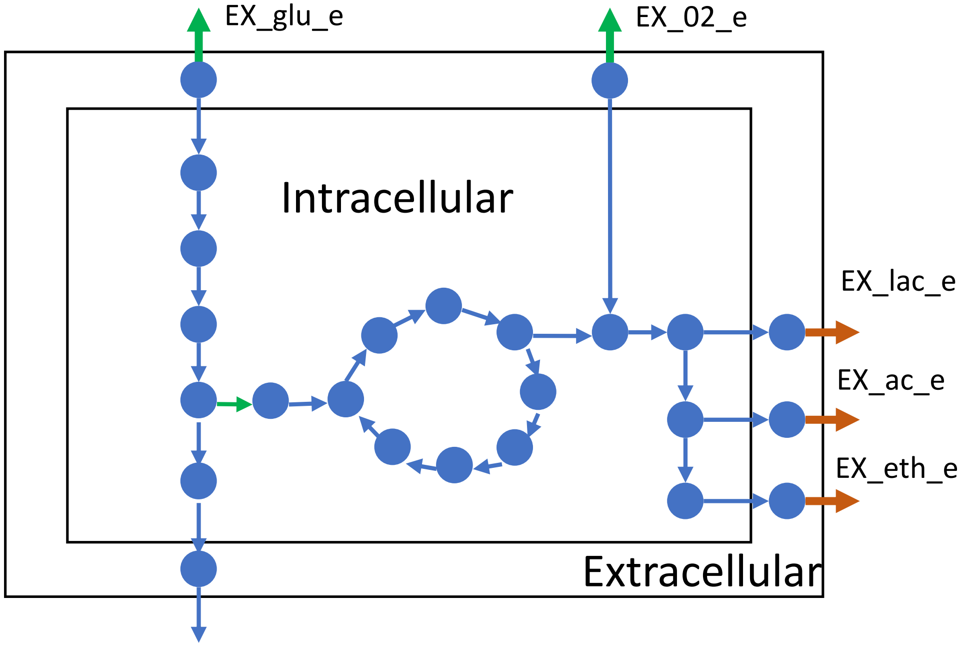
$$0 \leq v_{EX\_lac\_e} \leq 1000$$

$$0 \leq v_{EX\_ac\_e} \leq 1000$$

$$0 \leq v_{EX\_eth\_e} \leq 1000$$

Internal fluxes:

$$10 \leq v_{PYK} \leq 20$$

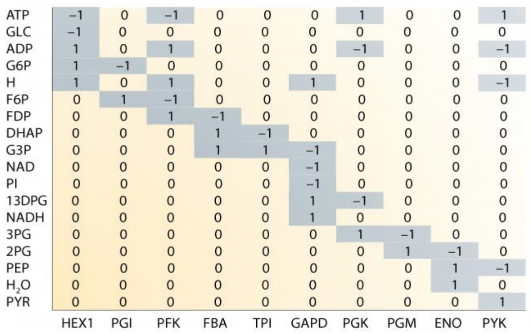


Constraints of metabolism

Equality constrains:

Stoichiometry:

$$Sv = 0$$



	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

Inequality constraints:

Uptakes:

$$-10 \leq v_{EX\_glu\_e} \leq 0$$

$$-20 \leq v_{EX\_O2\_e} \leq 0$$

Secretions:

$$0 \leq v_{EX\_lac\_e} \leq 1000$$

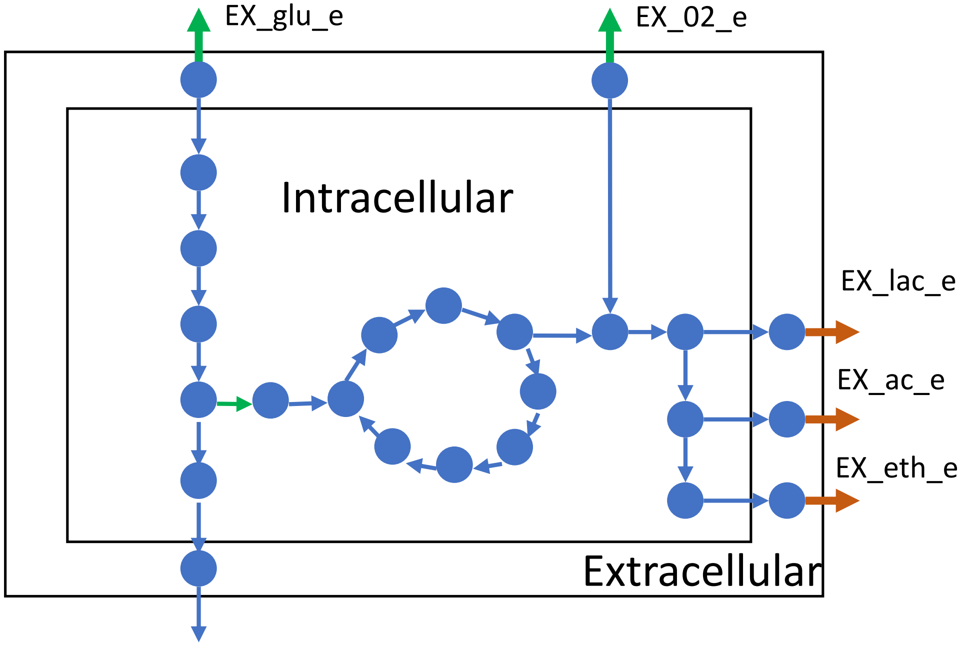
$$0 \leq v_{EX\_ac\_e} \leq 1000$$

$$0 \leq v_{EX\_eth\_e} \leq 1000$$

Internal fluxes:

$$10 \leq v_{PYK} \leq 20$$

If available from data….



Does metabolism have an “*Objective*” ?

Stoichiometry:

$$Sv = 0$$

Uptakes:

$$-10 \leq v_{EX\_glu\_e} \leq 0$$

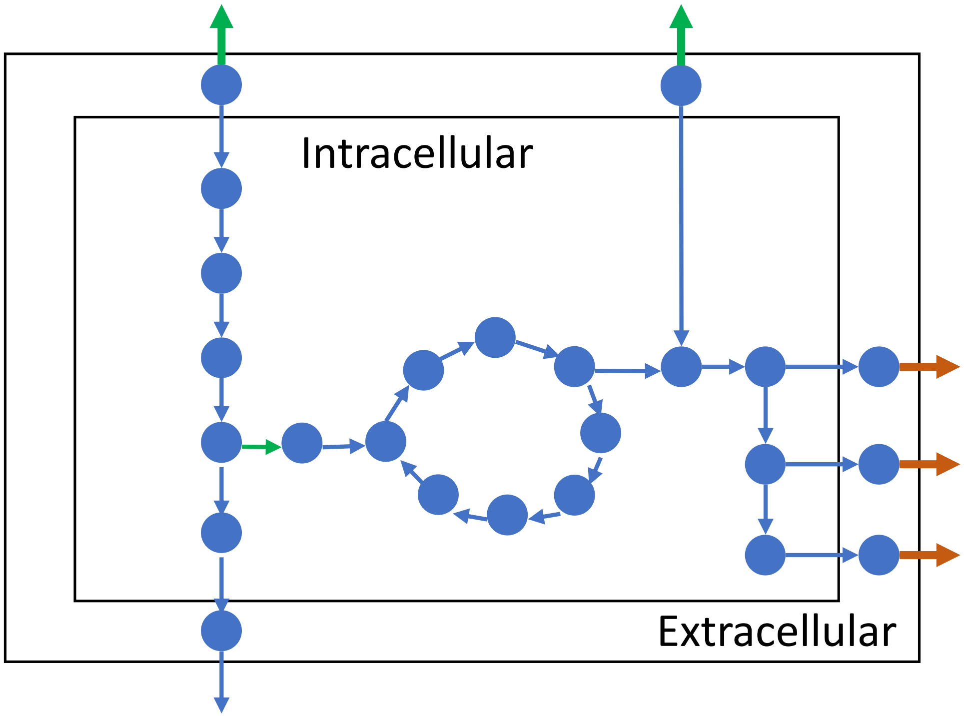
$$-20 \leq v_{EX\_O2\_e} \leq 0$$

Secretions:

$$0 \leq v_{EX\_lac\_e} \leq 1000$$

$$0 \leq v_{EX\_ac\_e} \leq 1000$$

$$0 \leq v_{EX\_eth\_e} \leq 1000$$



What is the objective function?

Does metabolism have an “*Objective*” ?

Stoichiometry:

$$Sv = 0$$

Uptakes:

$$-10 \leq v_{EX\_glu\_e} \leq 0$$

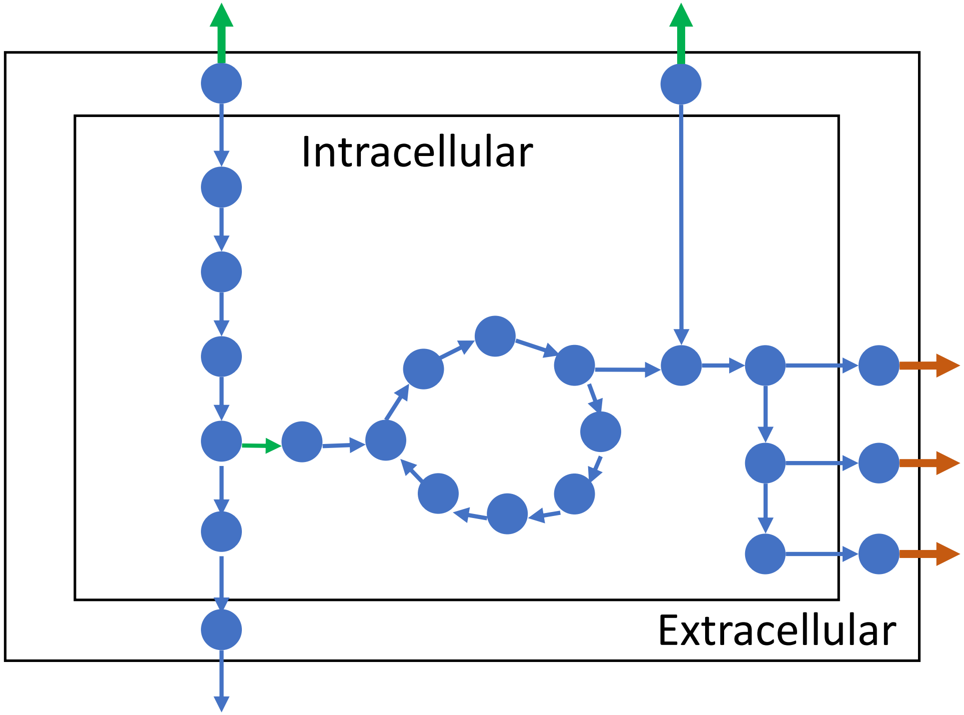
$$-20 \leq v_{EX\_O2\_e} \leq 0$$

Secretions:

$$0 \leq v_{EX\_lac\_e} \leq 1000$$

$$0 \leq v_{EX\_ac\_e} \leq 1000$$

$$0 \leq v_{EX\_eth\_e} \leq 1000$$



What is the objective function?

Evolutionary objective: Utilize **minimal** necessary **resources** for **maximum “growth”**

Does metabolism have an “*Objective*” ?

What is growth?

Does metabolism have an “*Objective*” ?

What is growth?

… doubling time

… maximum theoretical growth rate given
a maximum theoretical uptake rate

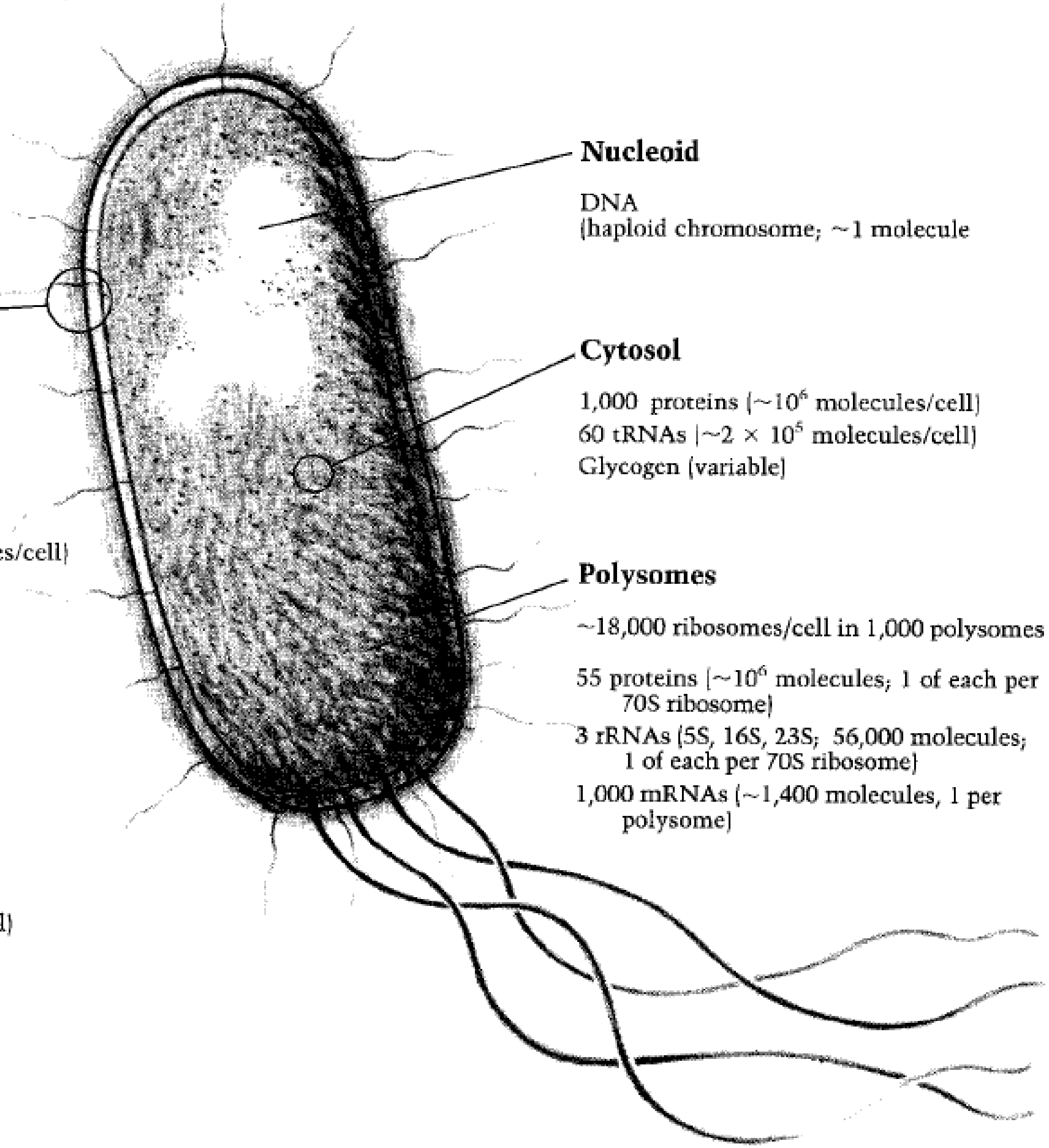
… specific growth rate, yield

Phase diagram … slope of log growth

Is maximization of growth an objective of the cell?
(Darwin…)

How to define growth in the model?

Biomass composition



Envelope

Flagella
6 proteins (~2 × 10^4 molecules/cell)

Pili
1 protein (~2 × 10^4 molecules/cell)

Outer membrane
50 proteins (4 abundant, 10^6 molecules/cell)
5 p-lipids (~5 × 10^6 molecules/cell)
1 LPS (9 × 10^6 molecules/cell)

Capsule
1 complex polysaccharide

Wall
Peptidoglycan (1 molecule/cell)

Periplasm
50 proteins (~10^4 molecules/cell)

Cell membrane
200 proteins (~2 × 10^5 molecules/cell)
7 p-lipids (~15 × 10^6 molecules/cell)

Nucleoid

DNA
(haploid chromosome; ~1 molecule

Cytosol

1,000 proteins (~10^6 molecules/cell)
60 tRNAs (~2 × 10^5 molecules/cell)
Glycogen (variable)

Polysomes

~18,000 ribosomes/cell in 1,000 polysomes

55 proteins (~10^6 molecules; 1 of each per 70S ribosome)
3 rRNAs (5S, 16S, 23S; 56,000 molecules; 1 of each per 70S ribosome)
1,000 mRNAs (~1,400 molecules, 1 per polysome)

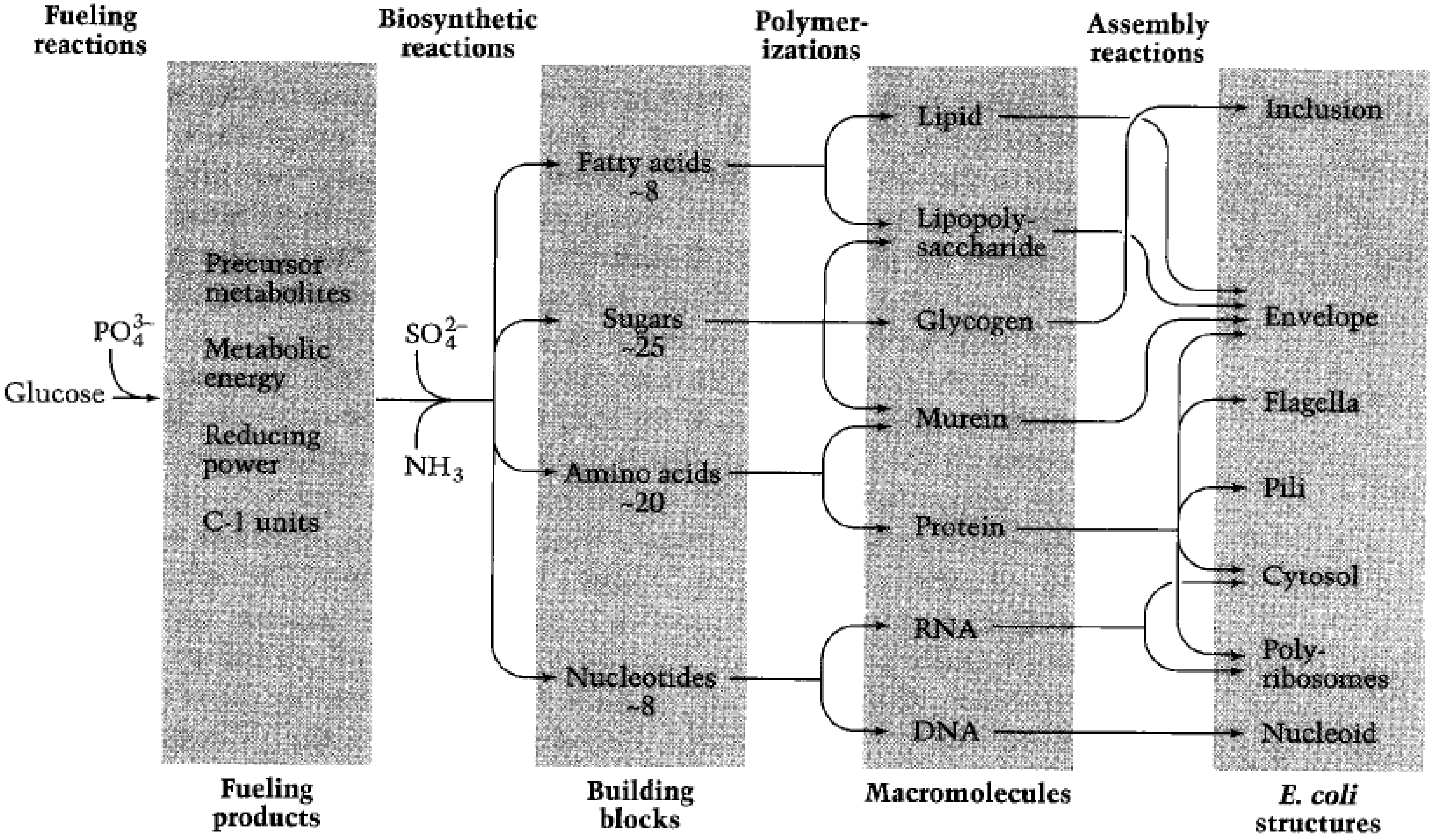


Figure 1

Overview of metabolism leading to the chemical synthesis from glucose of a chemoheterotroph like *E. coli*.

Table 3. Costs of biosynthesis of cellular components from the precursor metabolites

Cellular component	Energy cost[a] (µmole ~P/g cells)	Reducing power cost (µmole NADPH/g cells)
Protein	7,287	11,523
RNA	6,540	427
DNA	1,090	200
Lipid	2,578	5,270
LPS	470	564
Murein	248	193
Glycogen	154	0
1-Carbon	0	48
Polyamines	118	0
Total	18,485	18,225

[a] Each nucleoside triphosphate is assumed to be made by consecutive reactions with ATP that consume 3 ~P per NTP produced. Formation of sugar–nucleotide derivatives are assumed to occur by direct reaction with the appropriate nucleoside triphosphate.

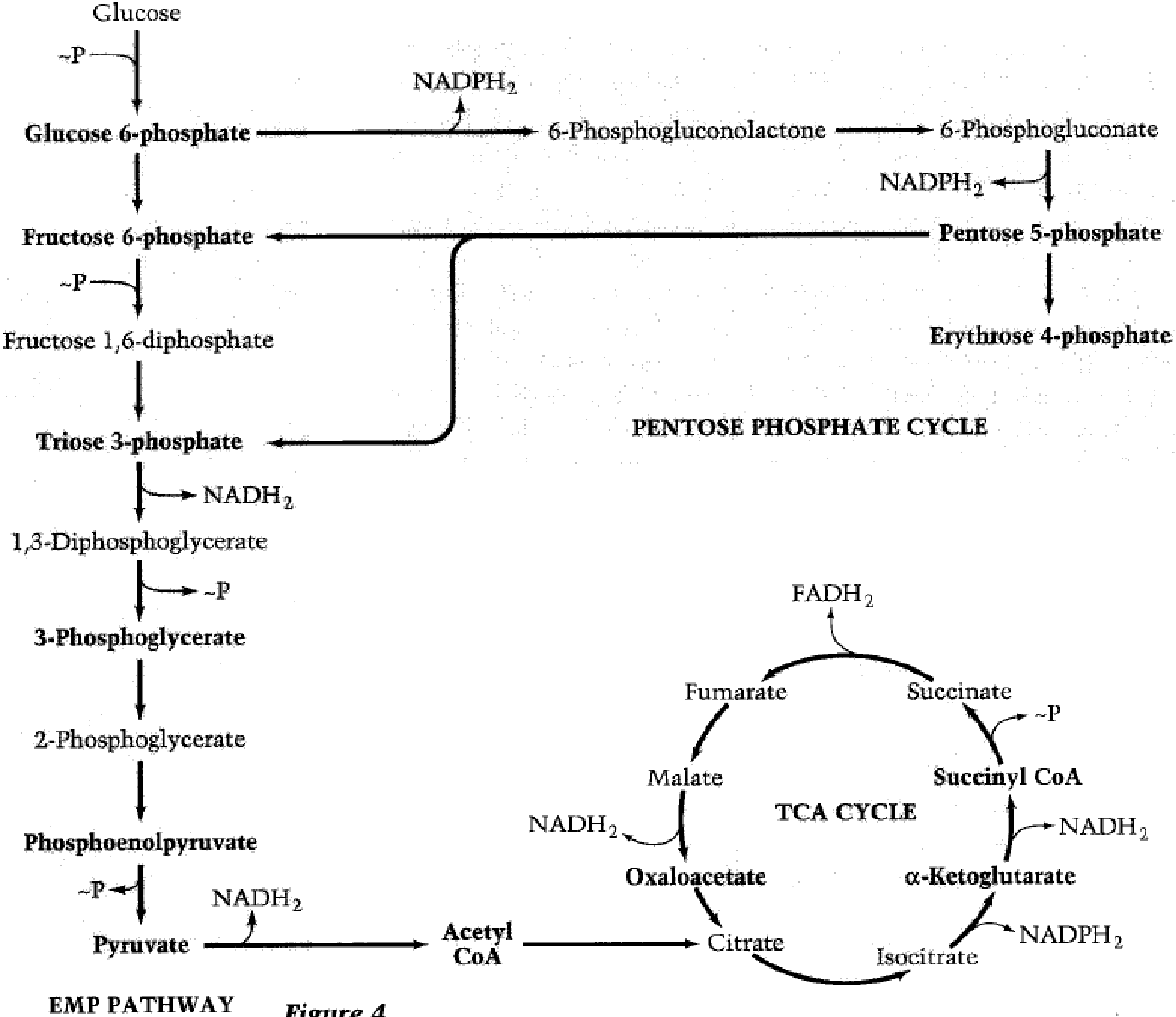
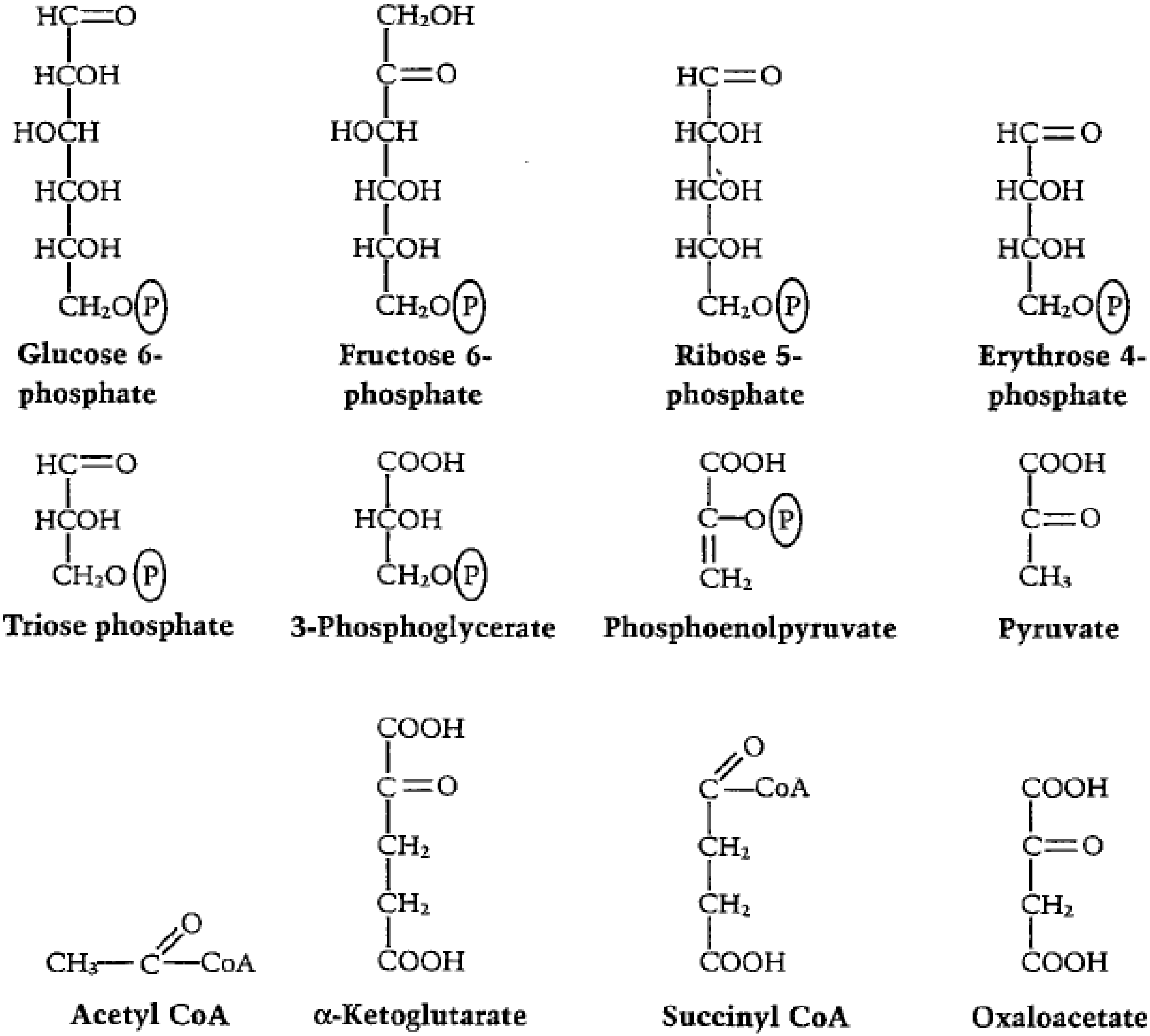

Glucose
~P
Glucose 6-phosphate
NADPH$_2$
6-Phosphogluconolactone
6-Phosphogluconate
NADPH$_2$
Fructose 6-phosphate
Pentose 5-phosphate
~P
Fructose 1,6-diphosphate
Erythrose 4-phosphate
Triose 3-phosphate
PENTOSE PHOSPHATE CYCLE
NADH$_2$
1,3-Diphosphoglycerate
~P
FADH$_2$
3-Phosphoglycerate
Fumarate
Succinate
~P
2-Phosphoglycerate
Malate
Succinyl CoA
TCA CYCLE
NADH$_2$
NADH$_2$
Phosphoenolpyruvate
Oxaloacetate
α-Ketoglutarate
~P
NADH$_2$
NADPH$_2$
Pyruvate
Acetyl CoA
Citrate
Isocitrate
EMP PATHWAY

Figure 4



Glucose 6-phosphate
Fructose 6-phosphate
Ribose 5-phosphate
Erythrose 4-phosphate
Triose phosphate
3-Phosphoglycerate
Phosphoenolpyruvate
Pyruvate
Acetyl CoA
α-Ketoglutarate
Succinyl CoA
Oxaloacetate

Figure 1

Structures of the 12 precursor metabolites.

Biosynthetic pathways

Biosynthetic pathways differ markedly in complexity—some are linear, others branch or are interconnected. A simple pathway, consisting of three sequential enzymatic reactions, might be represented as

$$\underset{(\text{PM})}{\text{Precursor metabolite}} \xrightarrow[(E_1)]{\text{Enzyme 1}} \text{A} \xrightarrow[(E_2)]{\text{Enzyme 2}} \text{B} \xrightarrow[(E_3)]{\text{Enzyme 3}} \underset{(\text{BB})}{\text{Building block}}$$

where A and B are intermediate products in the pathway. Recognizing

$$\mathrm{PM} \xrightarrow{3\mathrm{E}} \mathrm{BB}$$

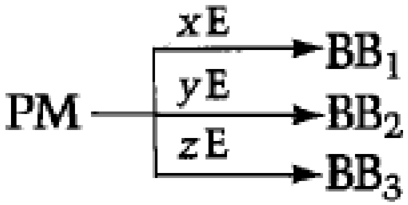
Some pathways produce a building block that, in turn, is converted by a second pathway into another building block:

$$\mathrm{PM} \xrightarrow{x\mathrm{E}} \mathrm{BB}_1 \xrightarrow{y\mathrm{E}} \mathrm{BB}_2$$

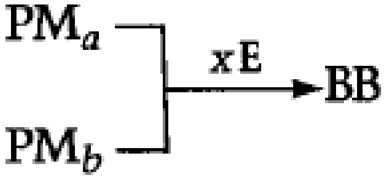
where *x* and *y* represent the number of enzymes in the two pathways. In some cases the pathway is branched:

$$\begin{tikzcd}
 & & \mathrm{BB}_2 \\
\mathrm{PM} \arrow[r, "x\mathrm{E}"] & \mathrm{BB}_1 \arrow[ur, "y\mathrm{E}"] \arrow[dr, "z\mathrm{E}"'] & \\
 & & \mathrm{BB}_3
\end{tikzcd}$$

Most of the 12 precursor metabolites actually serve as the starting point for several pathways:



Finally, in many cases more than one precursor molecule is involved in the biosynthesis of a building block:



Branching and interlocking of this sort are common among biosynthetic pathways. Building blocks that are produced from a common precursor are called a FAMILY. The ASPARTATE FAMILY, for example, consists of seven amino acids (asparagine, aspartate, diaminopimelate, isoleucine, lysine, methionine, and threonine) that are synthesized from the common precursor metabolite oxaloacetate (Figure 2).

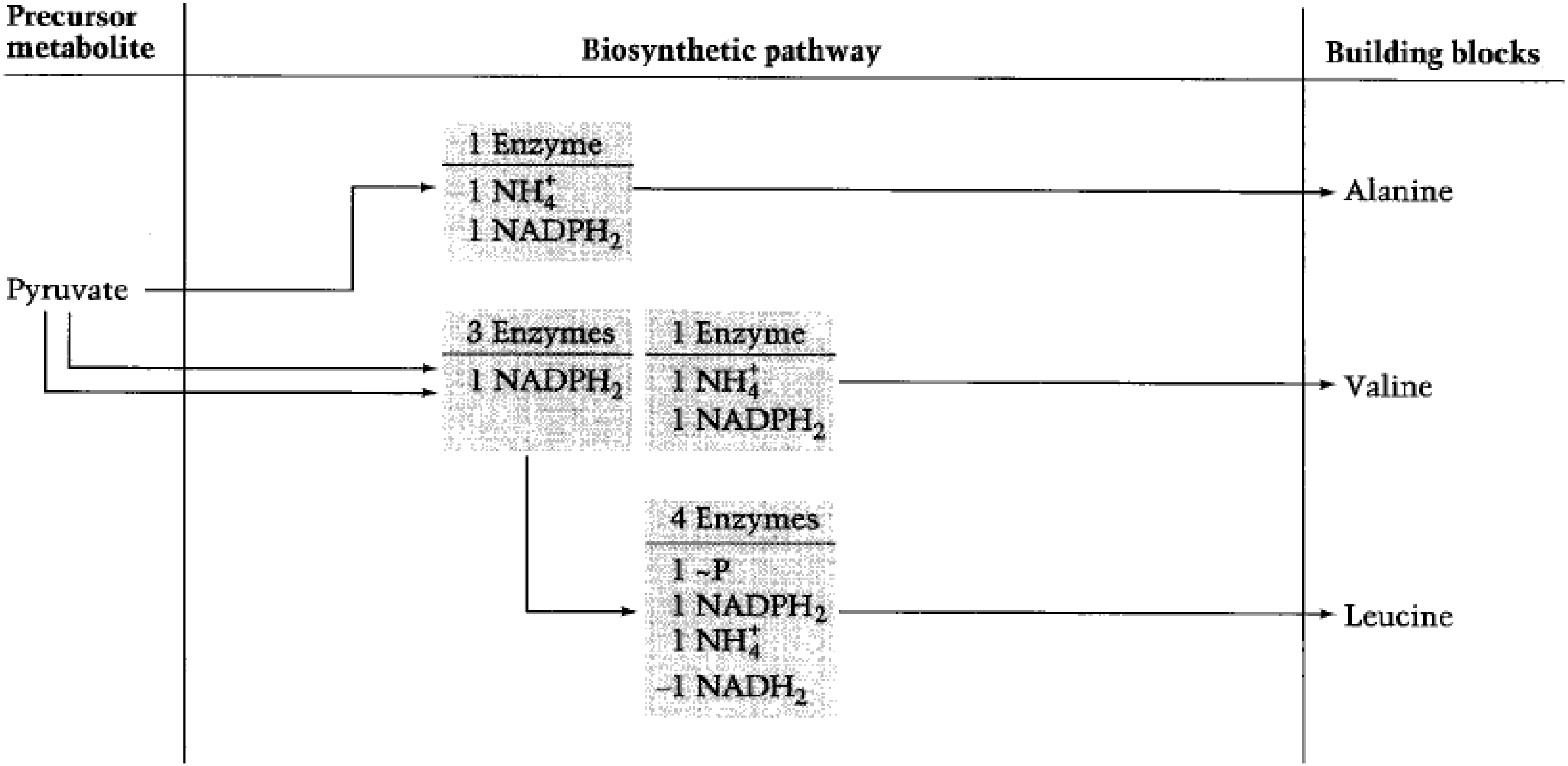
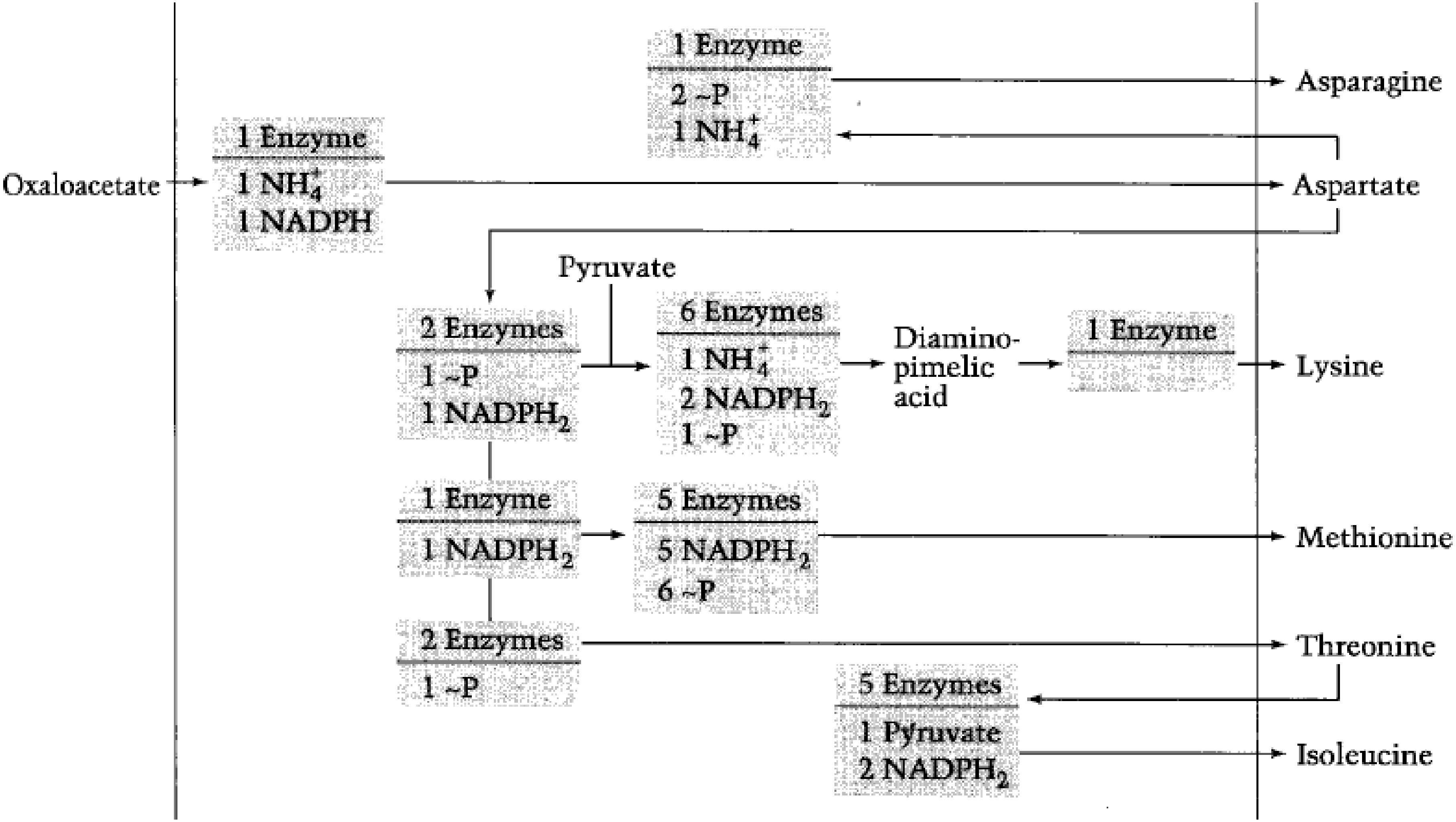
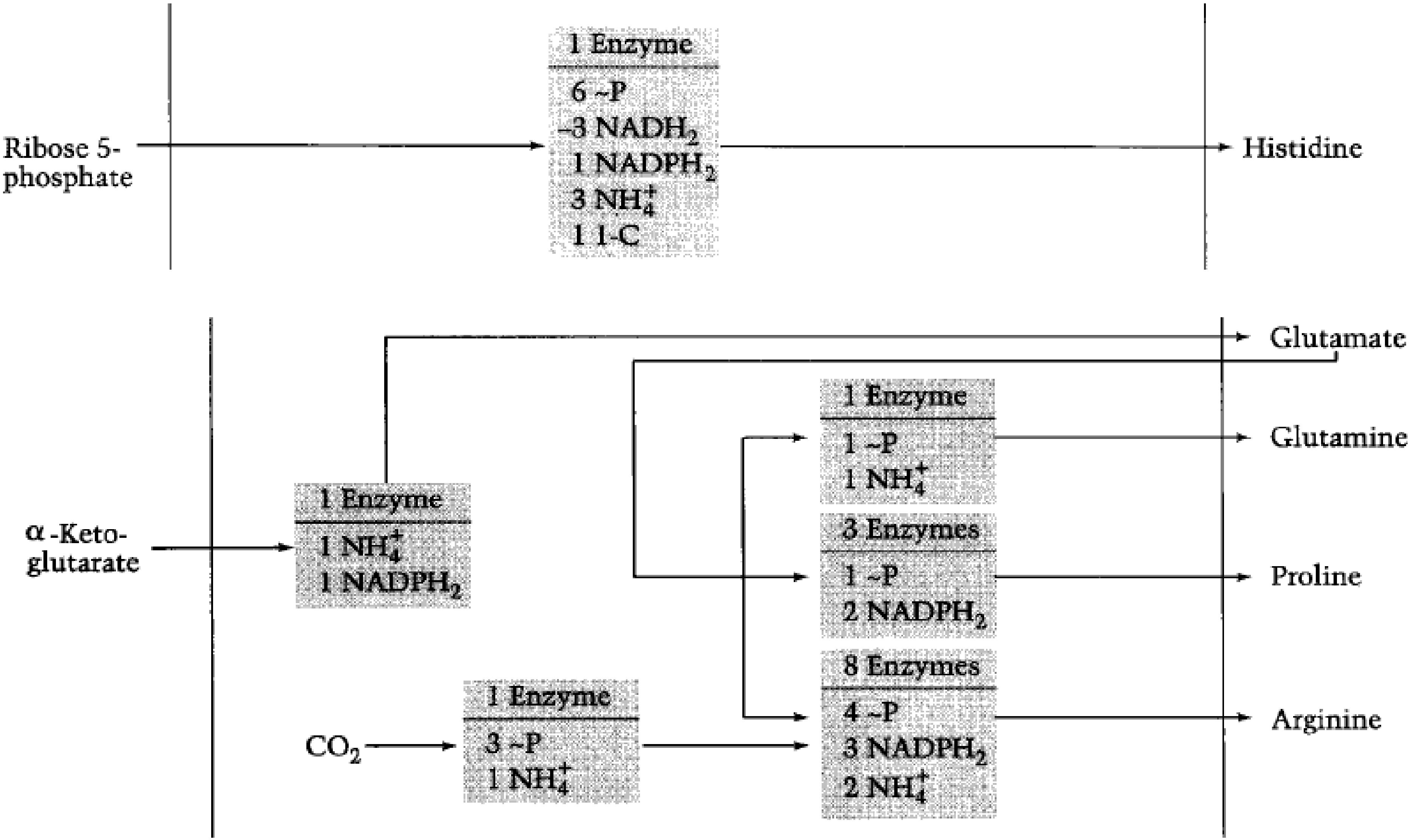
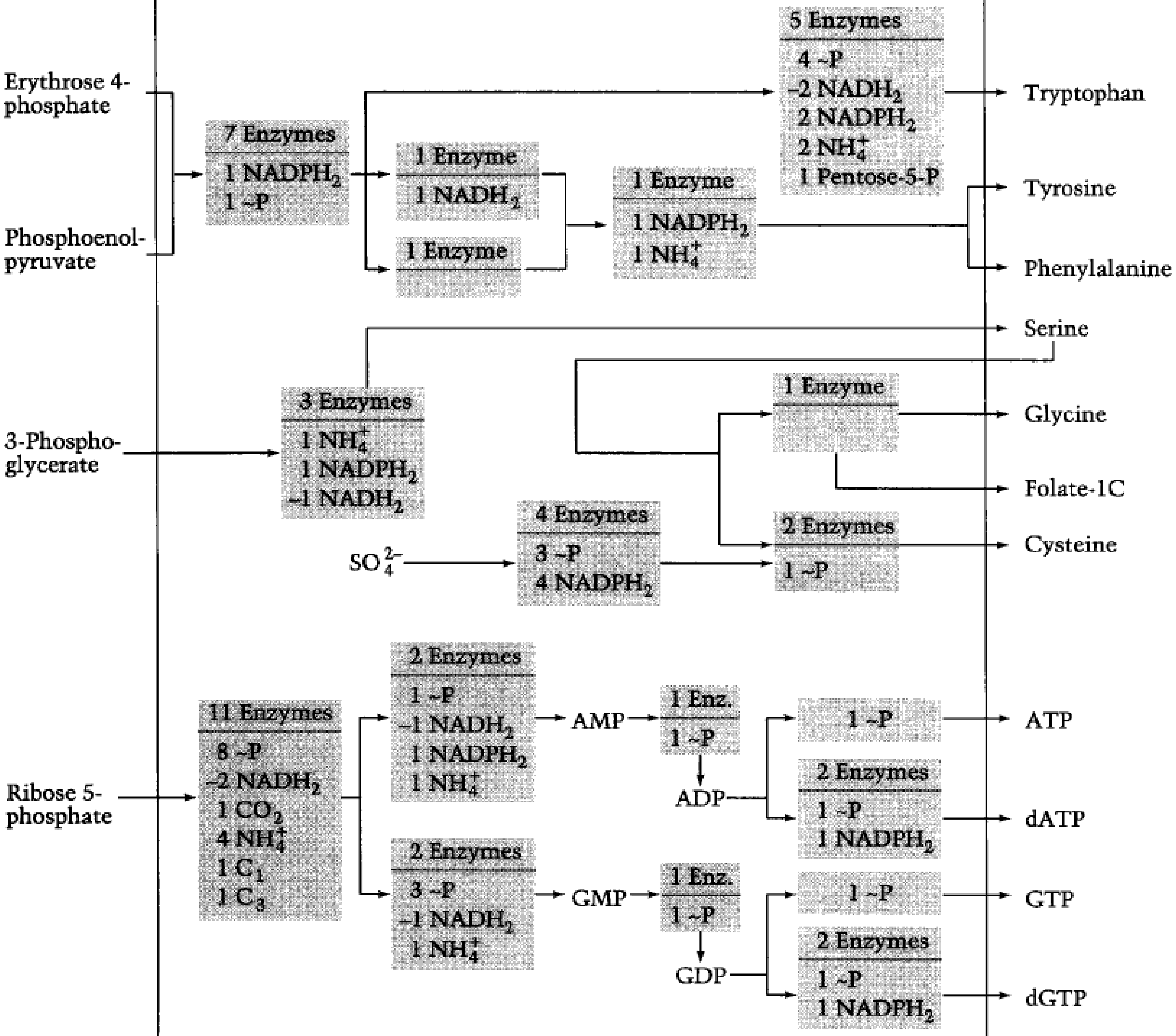

Precursor metabolite
Biosynthetic pathway
Building blocks
Pyruvate
1 Enzyme
1 NH$_4^+$
1 NADPH$_2$
Alanine
3 Enzymes
1 NADPH$_2$
1 Enzyme
1 NH$_4^+$
1 NADPH$_2$
Valine
4 Enzymes
1 ~P
1 NADPH$_2$
1 NH$_4^+$
–1 NADH$_2$
Leucine

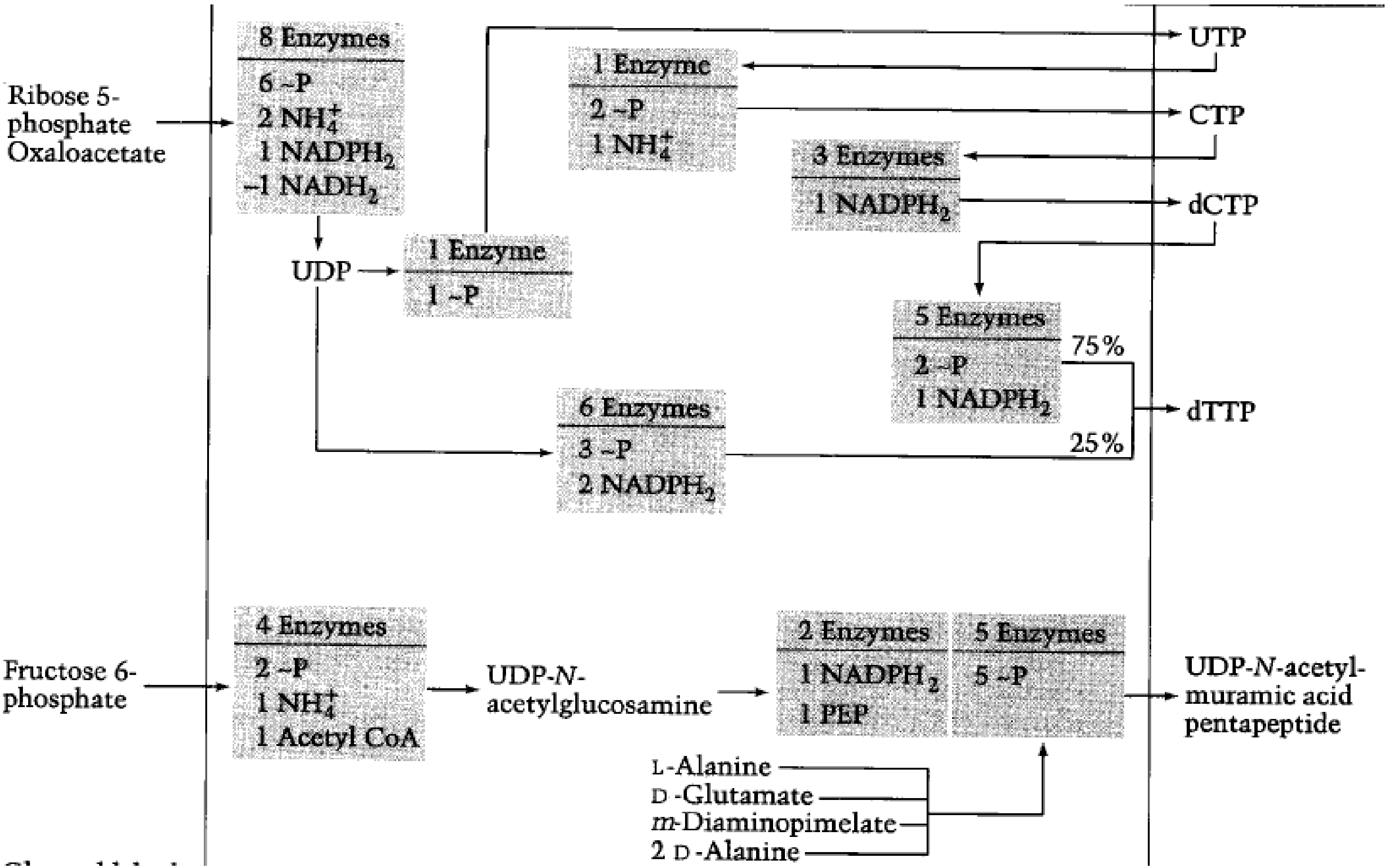
Table 1. Building blocks needed to produce 1 g of *E. coli* protoplasm

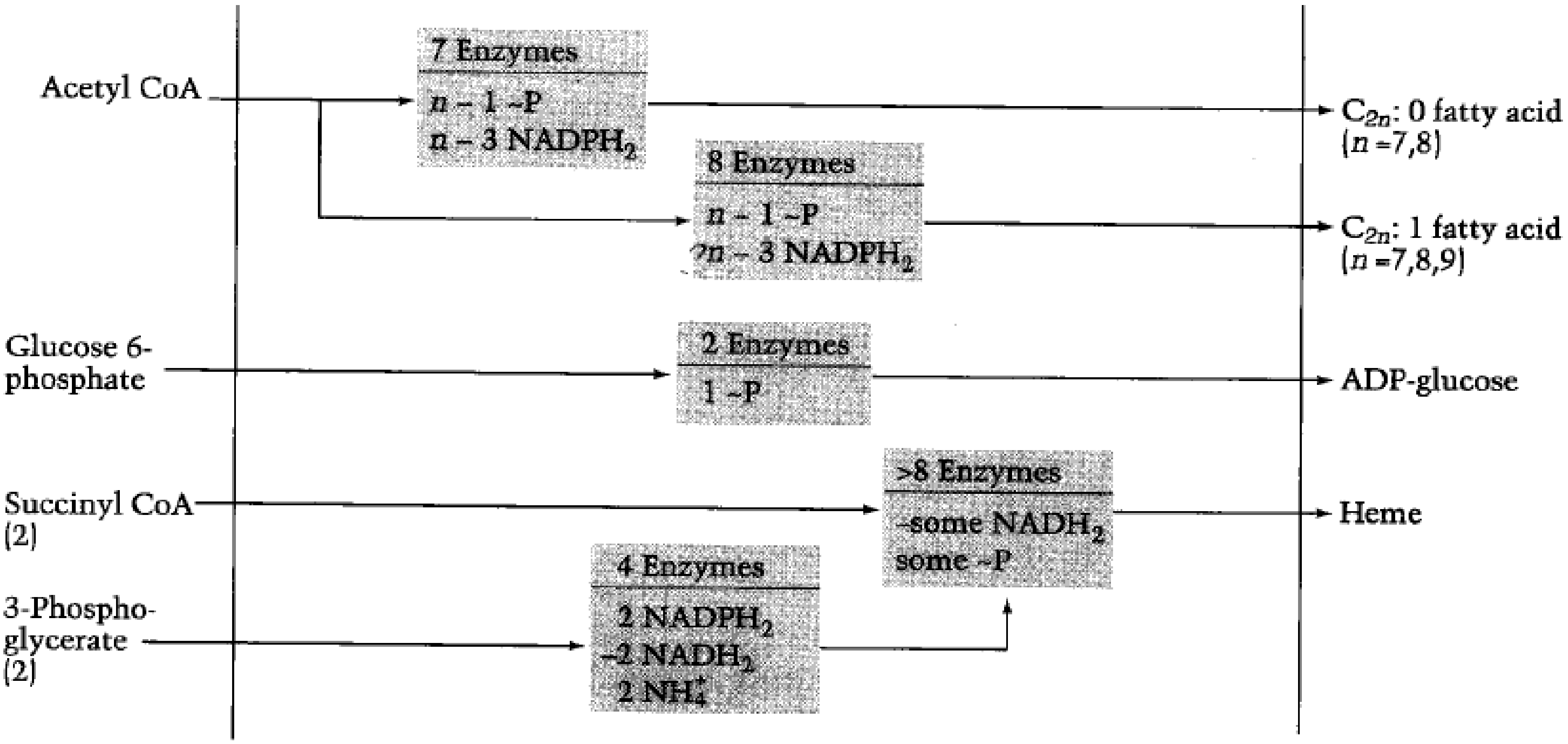
Building block	Amount present in *E. coli* B/r (µmol/g dried cells)	Cost of making 1 µmol of each of these building blocks (µmol/µmol)						
		Metabolites[a]	ATP	NADH	NADPH	1-C	NH_4^+	S
Protein amino acids								
Alanine	488	1 pyr	0	0	1	0	1	0
Arginine	281	1 αkg	7	−1	4	0	4	0
Asparagine	229	1 oaa	3	0	1	0	2	0
Aspartate	229	1 oaa	0	0	1	0	1	0
Cysteine	87	1 pga	4	−1	5	0	1	1
Glutamate	250	1 αkg	0	0	1	0	1	0
Glutamine	250	1 αkg	1	0	1	0	2	0
Glycine	582	1 pga	0	−1	1	−1	1	0
Histidine	90	1 penP	6	−3	1	1	3	0
Isoleucine	276	1 oaa, 1 pyr	2	0	5	0	1	0
Leucine	428	2 pyr, 1 acCoA	0	−1	2	0	1	0
Lysine	326	1 oaa, 1 pyr	2	0	4	0	2	0
Methionine	146	1 oaa	7	0	8	1	1	1
Phenylalanine	176	1 eryP, 2 pep	1	0	2	0	1	0
Proline	210	1 αkg	1	0	3	0	1	0
Serine	205	1 pga	0	−1	1	0	1	0
Threonine	241	1 oaa	2	0	3	0	1	0
Tryptophan	54	1 penP, 1 eryP, 1 pep	5	−2	3	0	2	0
Tyrosine	131	1 eryP, 2 pep	1	−1	2	0	1	0
Valine	402	2 pyr	0	0	2	0	1	0


Oxaloacetate
1 Enzyme
$1 NH_4^+$
1 NADPH
1 Enzyme
2 ~P
$1 NH_4^+$
Asparagine
Aspartate
Pyruvate
2 Enzymes
1 ~P
$1 NADPH_2$
6 Enzymes
$1 NH_4^+$
$2 NADPH_2$
1 ~P
Diamino-pimelic acid
1 Enzyme
Lysine
1 Enzyme
$1 NADPH_2$
5 Enzymes
$5 NADPH_2$
6 ~P
Methionine
2 Enzymes
1 ~P
Threonine
5 Enzymes
1 Pyruvate
$2 NADPH_2$
Isoleucine


Ribose 5-phosphate
1 Enzyme
6 ~P
–3 NADH$_2$
1 NADPH$_2$
3 NH$_4^+$
1 1-C
Histidine
α-Keto-glutarate
1 Enzyme
1 NH$_4^+$
1 NADPH$_2$
Glutamate
1 Enzyme
1 ~P
1 NH$_4^+$
Glutamine
3 Enzymes
1 ~P
2 NADPH$_2$
Proline
CO$_2$
1 Enzyme
3 ~P
1 NH$_4^+$
8 Enzymes
4 ~P
3 NADPH$_2$
2 NH$_4^+$
Arginine


Erythrose 4-phosphate
Phosphoenolpyruvate
7 Enzymes
1 $NADPH_2$
1 ~P
1 Enzyme
1 $NADH_2$
1 Enzyme
1 Enzyme
1 $NADPH_2$
1 NH_4^+
5 Enzymes
4 ~P
−2 $NADH_2$
2 $NADPH_2$
2 NH_4^+
1 Pentose-5-P
Tryptophan
Tyrosine
Phenylalanine
3-Phosphoglycerate
3 Enzymes
1 NH_4^+
1 $NADPH_2$
−1 $NADH_2$
Serine
1 Enzyme
Glycine
Folate-1C
SO_4^{2-}
4 Enzymes
3 ~P
4 $NADPH_2$
2 Enzymes
1 ~P
Cysteine
Ribose 5-phosphate
11 Enzymes
8 ~P
−2 $NADH_2$
1 CO_2
4 NH_4^+
1 C_1
1 C_3
2 Enzymes
1 ~P
−1 $NADH_2$
1 $NADPH_2$
1 NH_4^+
AMP
1 Enz.
1 ~P
ADP
1 ~P
ATP
2 Enzymes
1 ~P
1 $NADPH_2$
dATP
2 Enzymes
3 ~P
−1 $NADH_2$
1 NH_4^+
GMP
1 Enz.
1 ~P
GDP
1 ~P
GTP
2 Enzymes
1 ~P
1 $NADPH_2$
dGTP

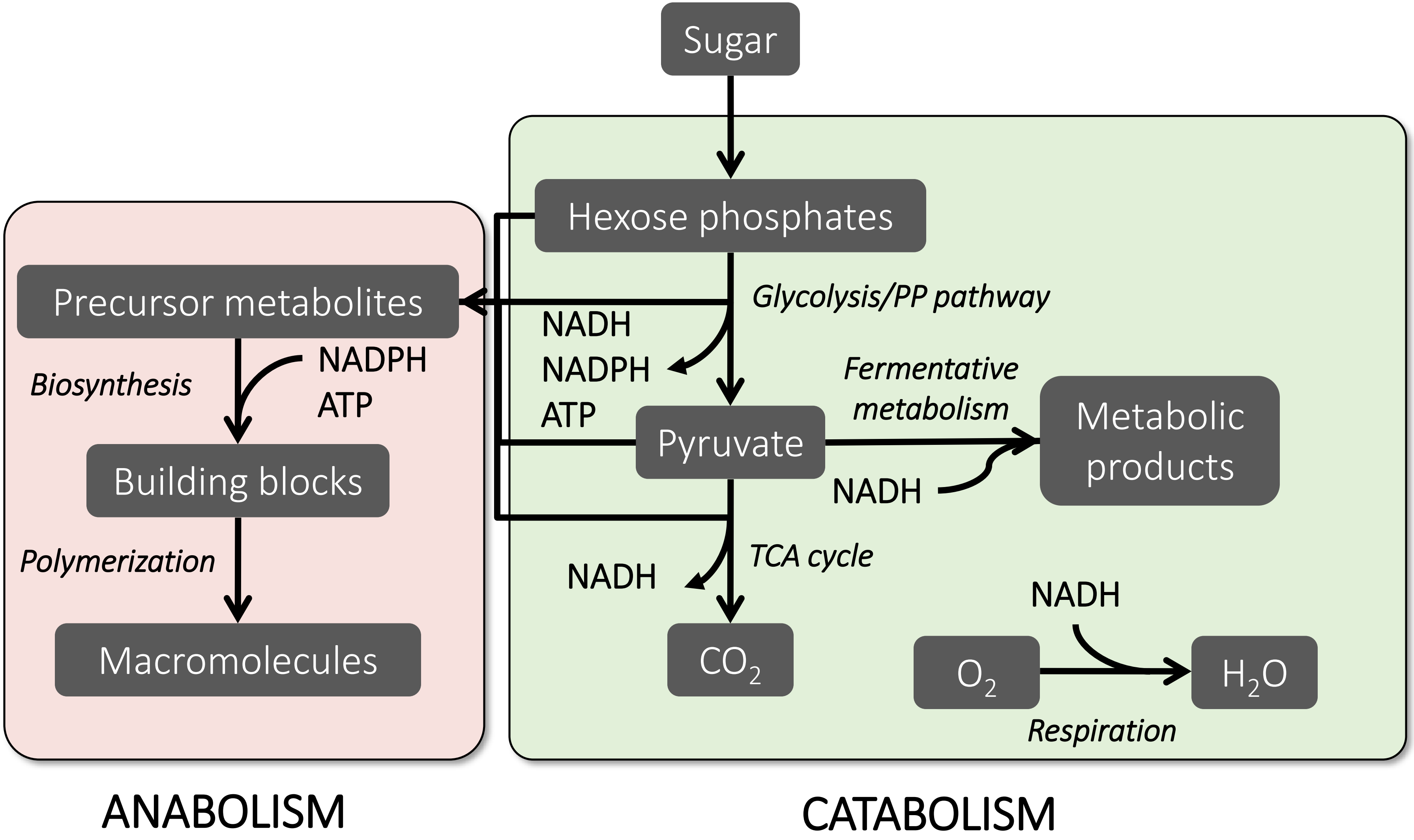

Ribose 5-phosphate
Oxaloacetate
8 Enzymes
6 ~P
2 NH_4^+
1 $NADPH_2$
–1 $NADH_2$
UDP
1 Enzyme
1 ~P
UTP
1 Enzyme
2 ~P
1 NH_4^+
CTP
3 Enzymes
1 $NADPH_2$
dCTP
5 Enzymes
2 ~P
1 $NADPH_2$
75%
dTTP
6 Enzymes
3 ~P
2 $NADPH_2$
25%
Fructose 6-phosphate
4 Enzymes
2 ~P
1 NH_4^+
1 Acetyl CoA
UDP-*N*-acetylglucosamine
2 Enzymes
1 $NADPH_2$
1 PEP
5 Enzymes
5 ~P
UDP-*N*-acetylmuramic acid pentapeptide
L-Alanine
D-Glutamate
m-Diaminopimelate
2 D-Alanine


Acetyl CoA
7 Enzymes
$n - 1 \sim P$
$n - 3$ NADPH$_2$
8 Enzymes
$n - 1 \sim P$
$2n - 3$ NADPH$_2$
C_{2n}: 0 fatty acid
(n = 7,8)
C_{2n}: 1 fatty acid
(n = 7,8,9)
Glucose 6-phosphate
2 Enzymes
$1 \sim P$
ADP-glucose
Succinyl CoA
(2)
>8 Enzymes
–some NADH$_2$
some ~P
Heme
3-Phosphoglycerate
(2)
4 Enzymes
2 NADPH$_2$
–2 NADH$_2$
2 NH$_4^+$

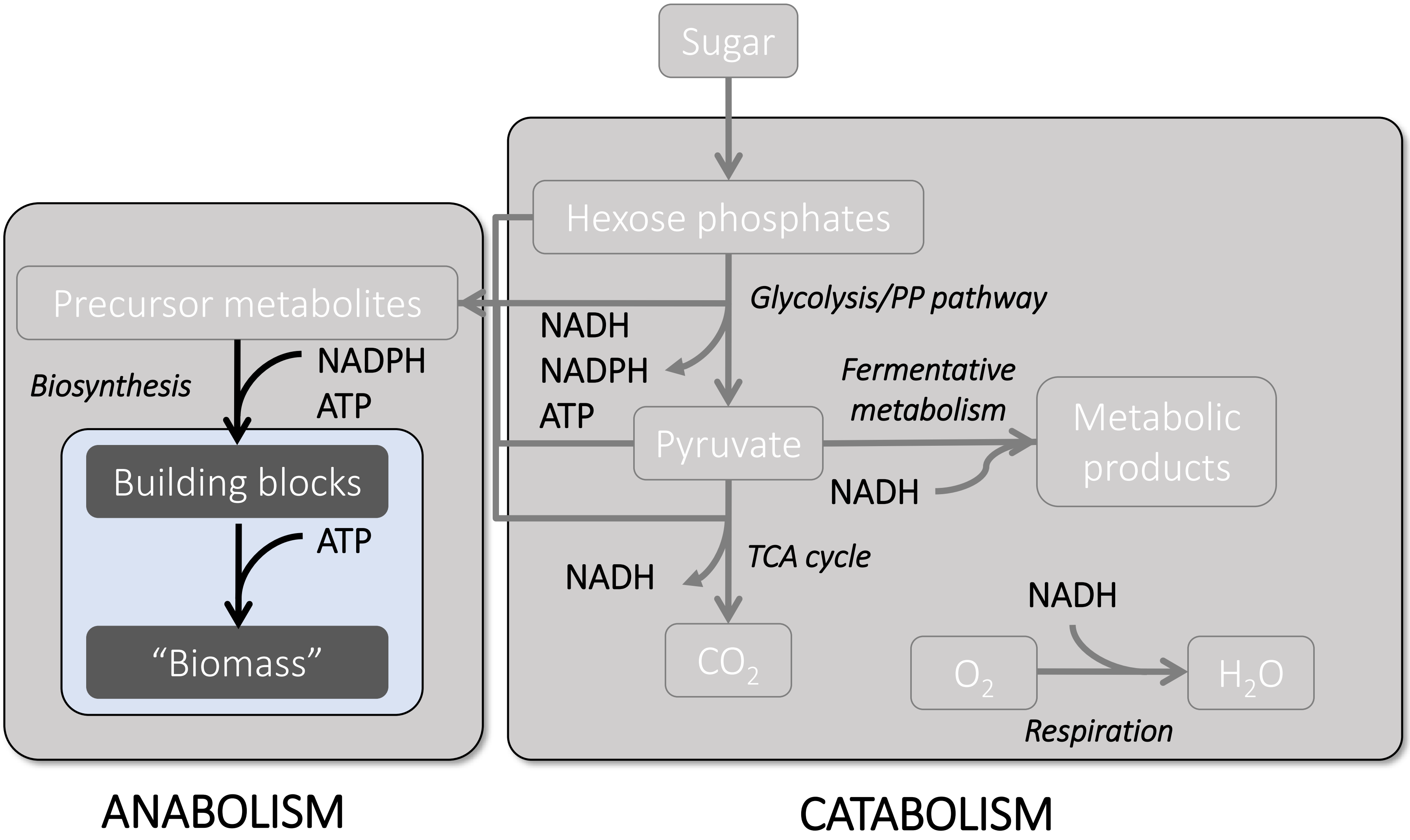
RNA nucleotides								
ATP	165	1 penP, 1 pga	11	−3	1	1	5	0
GTP	203	1 penP, 1 pga	13	−3	0 ?	1	5	0
CTP	126	1 penP, 1 oaa	9	0	1	0	3	0
UTP	136	1 penP, 1 oaa	7	0	1	0	2	0
DNA nucleotides								
dATP	24.7	1 penP, 1 pga	11	−3	2	1	5	0
dGTP	25.4	1 penP, 1 pga	13	−3	1	1	5	0
dCTP	25.4	1 penP, 1 oaa	9	0	2	0	3	0
dTTP	24.7	1 penP, 1 oaa	10.5	0	3	1	2	0
Lipid components								
Glycerol phosphate	129	1 triosP	0	0	1	0	0	0
Serine	129	1 pga	0	−1	1	0	1	0
$C_{16:0}$ fatty acid (43%)		8 acCoA	7	0	14	0	0	0
$C_{16:1}$ fatty acid (33%)		8 acCoA	7	0	13	0	0	0
$C_{18:1}$ fatty acid (24%)		9 acCoA	8	0	15	0	0	0
Average fatty acid	258	8.2 acCoA	7.2	0	14	0	0	0

Building block	Amount present in *E. coli* B/r (μmol/g dried cells)	Cost of making 1 μmol of each of these building blocks (μmol/μmol)						
		Metabolites[a]	ATP	NADH	NADPH	1-C	NH_4^+	S
LPS components								
UDP-glucose	15.7	1 gluP	1	0	0	0	0	0
(CDP) ethanolamine	23.5	1 pga	3	−1	1	0	1	0
OH-myristic acid	23.5	7 acCoA	6	0	11	0	0	0
$C_{14:0}$ fatty acid	23.5	7 acCoA	6	0	12	0	0	0
(CMP) KDO	23.5	1 penP, 1 pep	2	0	0	0	0	0
(NDP) heptose	23.5	1.5 gluP	1	0	−4	0	0	0
(TDP) glucosamine	15.7	1 fruP	2	0	0	0	1	0
Peptidoglycan monomers								
UDP-*N*-acetylglucosamine	27.6	1 fruP, 1 acCoA	3	0	0	0	1	0
UDP-*N*-acetylmuramic acid	27.6	1 fruP, 1 pep, 1 acCoA	4	0	1	0	1	0
Alanine	55.2	1 pyr	0	0	1	0	1	0
Diaminopimelate	27.6	1 oaa, 1 pyr	2	0	3	0	2	0
Glutamate	27.6	1 αkg	0	0	1	0	1	0
Glycogen monomers								
Glucose	154	1 gluP	1	0	0	0	0	0
1-Carbon requirement								
Serine	48.5	1 pga	0	−1	1	0	0	0
Polyamines								
Ornithine equivalents	59.3	1 αkg	2	0	3	0	2	0

The biomass reactions



The biomass reactions



The biomass reactions

1g Biomass

Macromolecules

DNA	6.7 %
RNA	5.9 %
Protein	48.0 %
Lipids	14.1 %
Carbohydrates a.o.	25.3 %

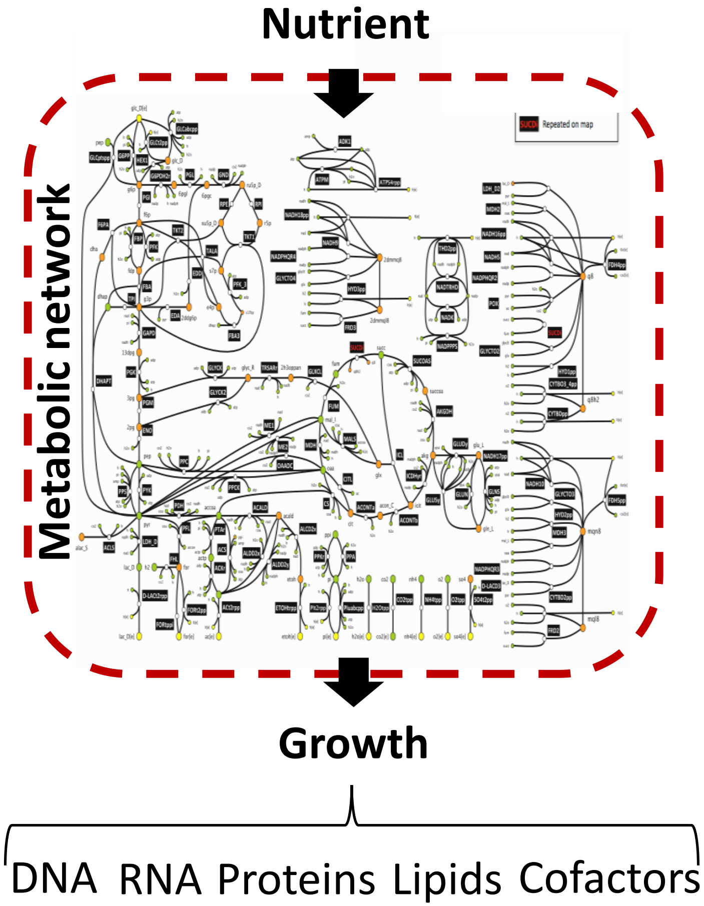
Biomass building blocks

AMP	38.7
UMP	38.7
GMP	11.3
CMP	11.3

etc.

Data for toxoplasma :Thomas Forth PhD thesis, Sept 2012

The biomass reactions



Nutrient

Metabolic network

Growth

DNA RNA Proteins Lipids Cofactors

- Organism-specific models
- Start from sequenced genome
- Correlate genome with molecular physiology

THE RAVEN TOOLBOX

KEGG PlasmoDB GeneDB Malaria Parasite Metabolic Pathways

Defining the growth problem

Equality constrains:

$$Sv = 0$$

Inequality constrains:

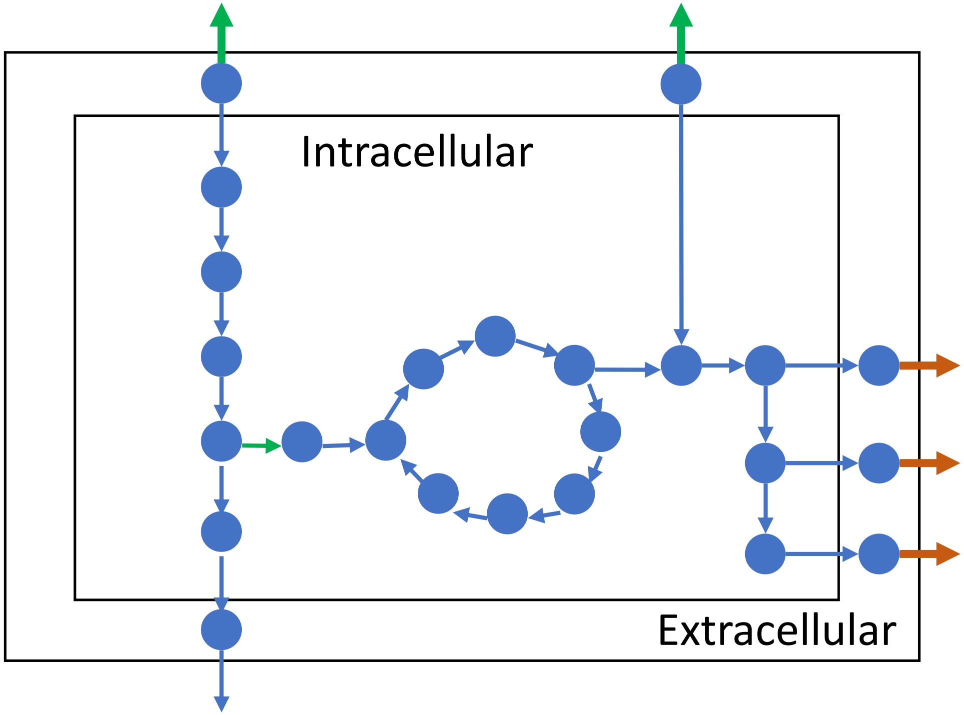
$$-10 \leq v_{EX\_glu\_e} \leq 0$$

$$-20 \leq v_{EX\_O2\_e} \leq 0$$

$$0 \leq v_{EX\_lac\_e} \leq 1000$$

$$0 \leq v_{EX\_ac\_e} \leq 1000$$

$$0 \leq v_{EX\_eth\_e} \leq 1000$$



Objective function:

max(**biomass reaction**):

Linear programing

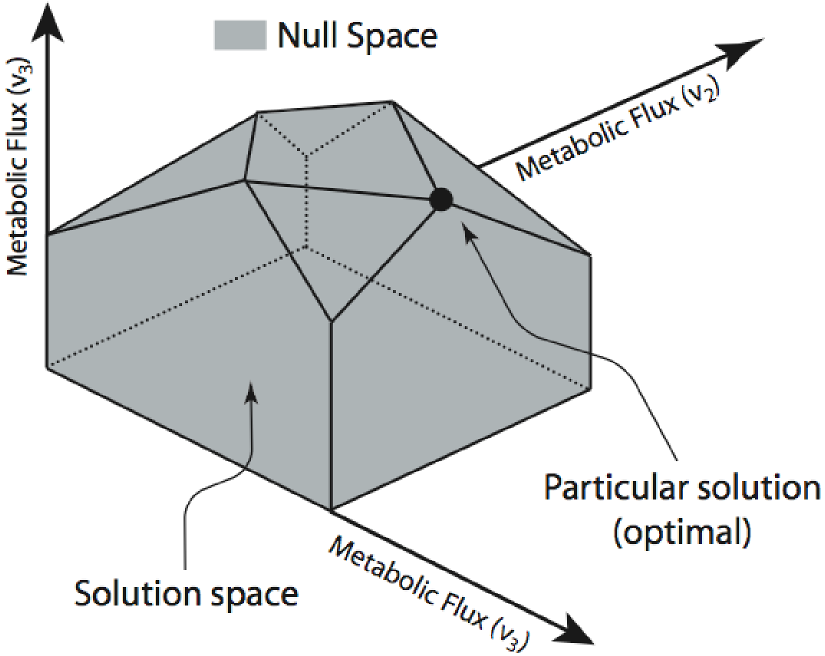
1. Forming the Solution Space

$$\mathbf{Sv} = 0 \quad , \qquad 0 \le v_i \le v_{i,m}$$ internal reactions

$$min \le b_i \le max$$ inputs & outputs

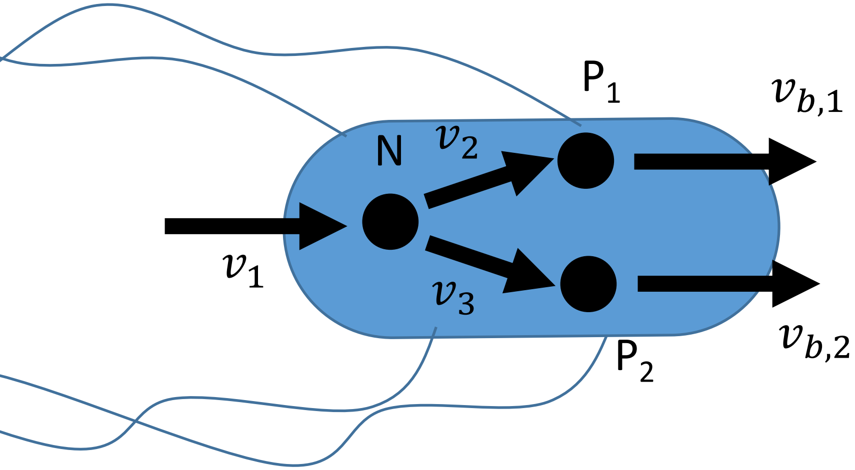
2. Objective Function

$$Z = \sum_{i=1}^{n} c_i v_i$$



Originally from **economics** to **optimize** production processes subject to **linear constraints**.

How Does LP work: 2D example



Biomass building block b$_1$ and b$_2$

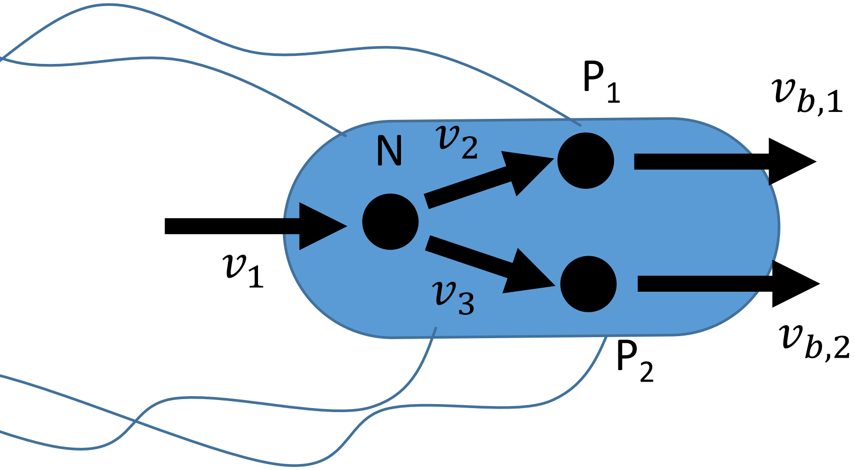
Boundary conditions: $v_1 = 10$

Mass balances:

$$v_1 - v_2 - v_3 = 0$$
$$v_2 - v_{b,1} = 0$$
$$v_3 - v_{b,2} = 0$$

Biomass function: $v_{b,1} + 2v_{b,2} = v_{bio}$

How Does LP work: 2D example



Boundary conditions: $v_1 = 10$

Mass balances:

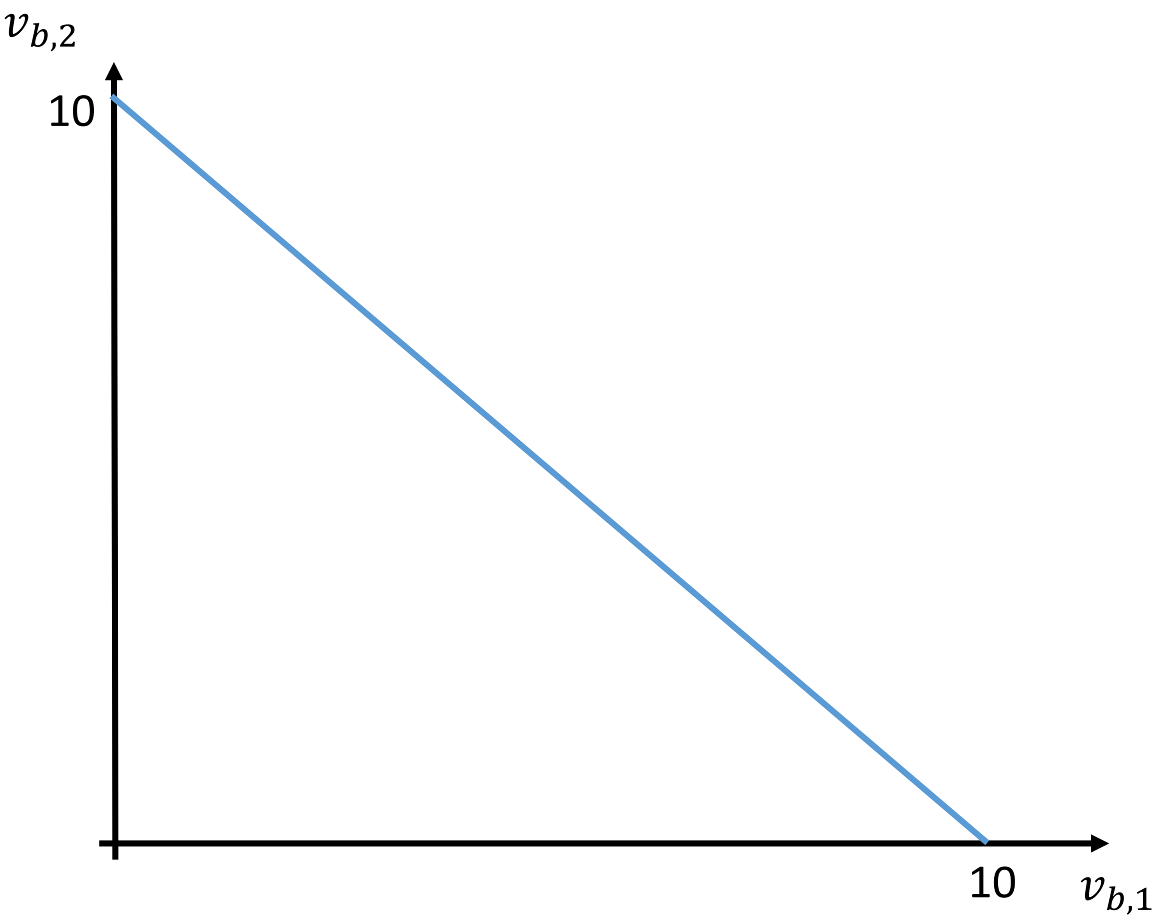
$$\left.\begin{aligned} v_1 &= 10 \\ v_1 - v_2 - v_3 &= 0 \\ v_2 - v_{b,1} &= 0 \\ v_3 - v_{b,2} &= 0 \end{aligned}\right\} \quad 10 - v_{b,1} - v_{b,2} = 0$$

Biomass function: $v_{b,1} + 2v_{b,2} = v_{bio}$

How Does LP work: 2D example

Equality constraints:

$$v_{b,1} = 10 - v_{b,2}$$



How Does LP work: 2D example

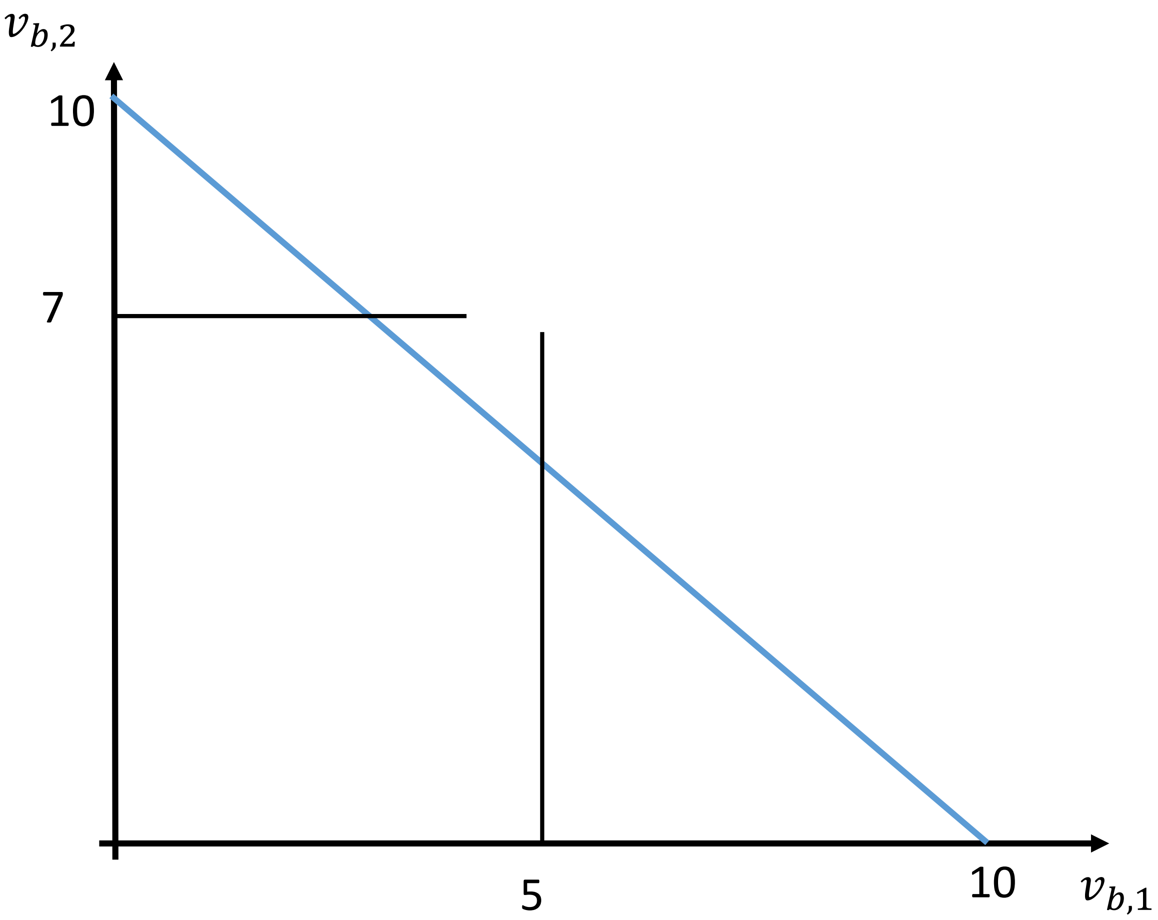
Equality constraints:

$$v_{b,1} = 10 - v_{b,2}$$

Inequality constraints:

$$0 \leq v_{b,2} \leq 7$$

$$0 \leq v_{b,1} \leq 5$$



How Does LP work: 2D example

Equality constraints:

$$v_{b,1} = 10 - v_{b,2}$$

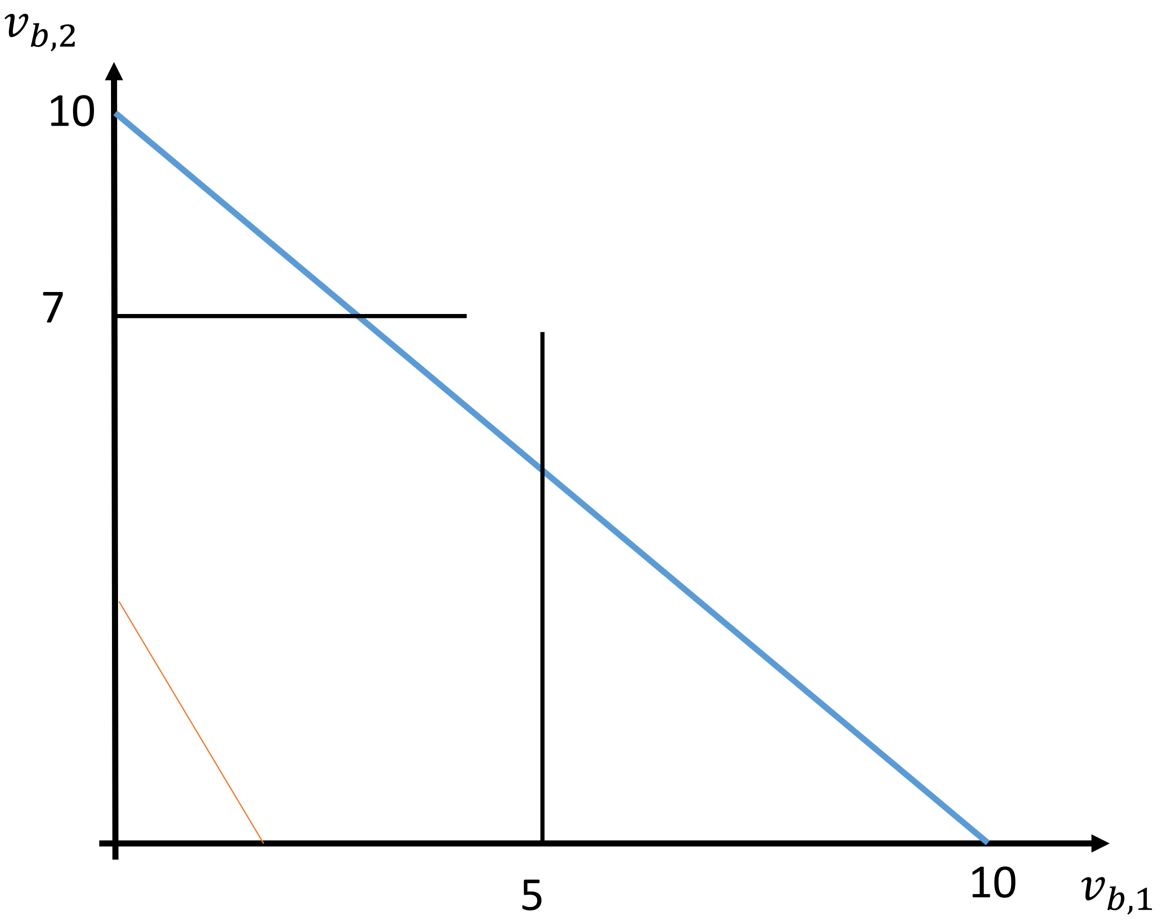
Inequality constraints:

$$0 \leq v_{b,2} \leq 7$$

$$0 \leq v_{b,1} \leq 5$$

Objective function:

$$v_{b,1} + 2v_{b,2} = v_{bio}$$



How Does LP work: 2D example

Equality constraints:

$$v_{b,1} = 10 - v_{b,2}$$

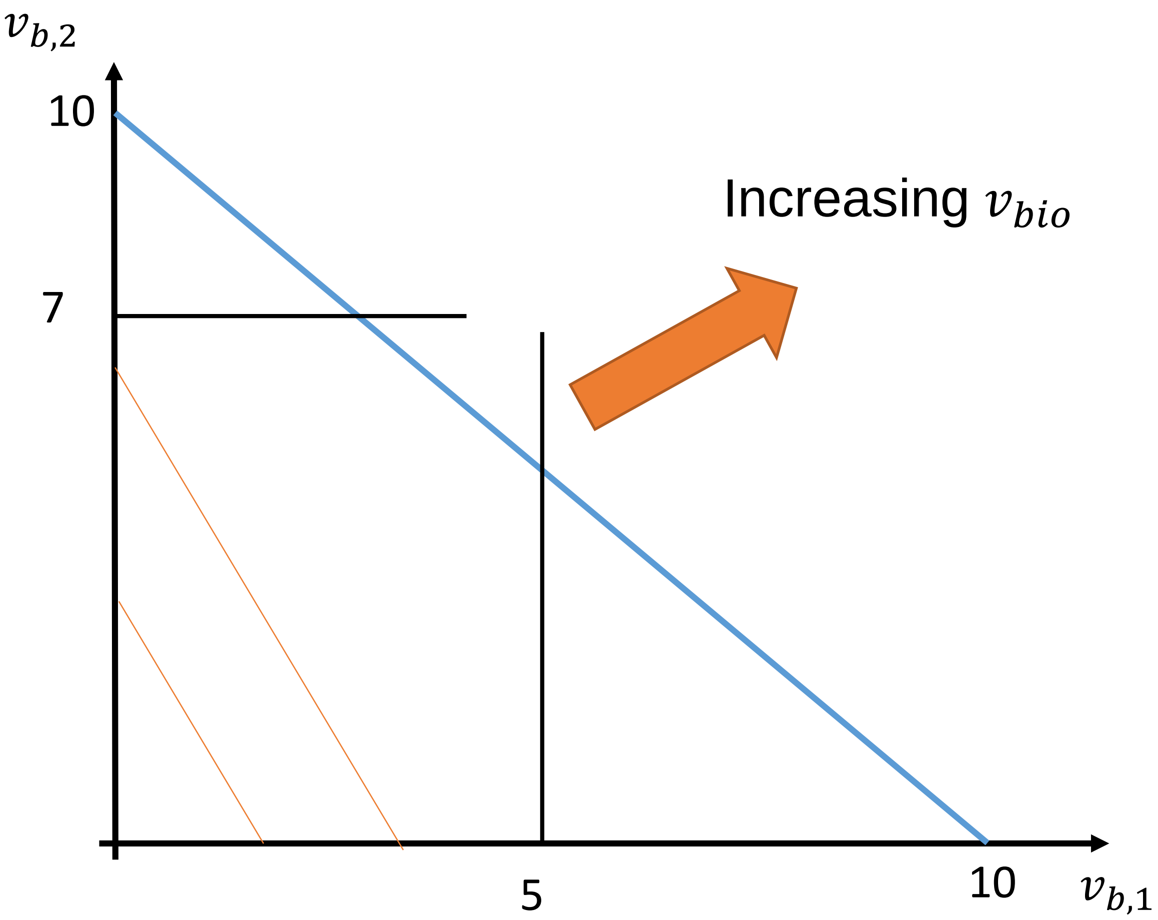
Inequality constraints:

$$0 \leq v_{b,2} \leq 7$$

$$0 \leq v_{b,1} \leq 5$$

Objective function:

$$v_{b,1} + 2v_{b,2} = v_{bio}$$



How Does LP work: 2D example

Equality constraints:

$$v_{b,1} = 10 \; - v_{b,2}$$

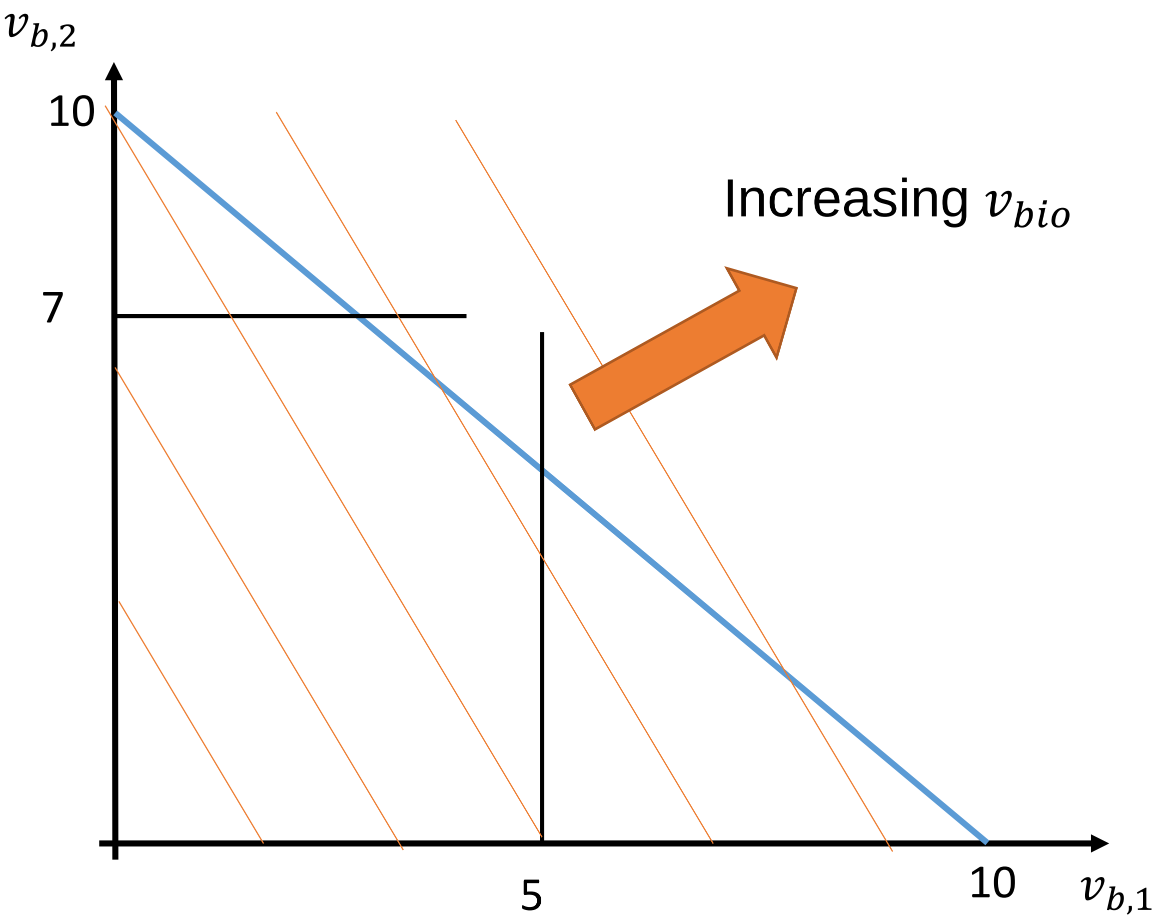
Inequality constraints:

$$0 \leq v_{b,2} \leq 7$$

$$0 \leq v_{b,1} \leq 5$$

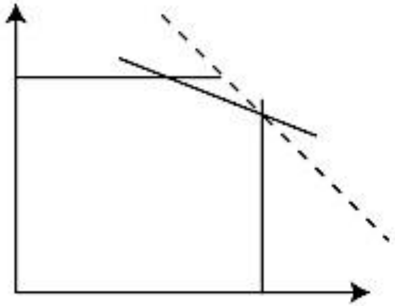
Objective function:

$$v_{b,1} + 2v_{b,2} = v_{bio}$$



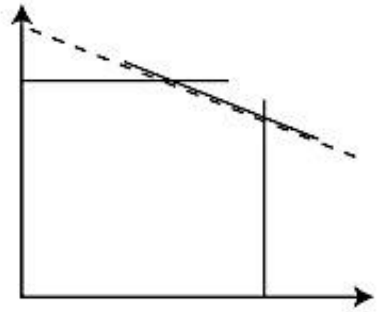
Types of Feasible Solutions found by LP

Unique solution



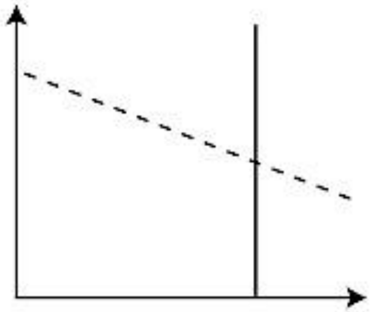
Optimal solution
in a corner

Degenerate solution



Optimal solution
along an edge

Unbounded solution

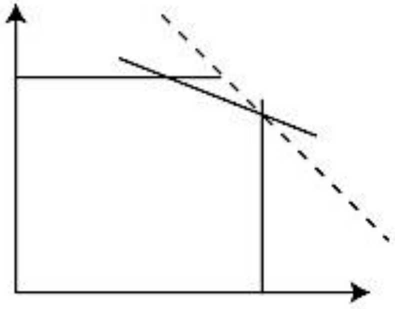


Optimal solution not
found--region unbounded

- - - - - - - - - Lines of constant Z

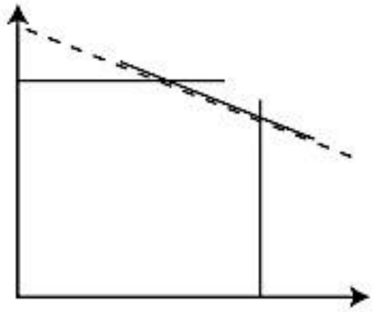
Types of Feasible Solutions found by LP

Unique solution



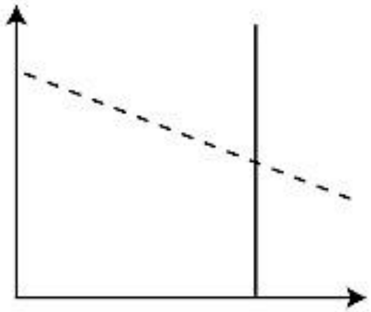
Optimal solution in a corner

Degenerate solution



Optimal solution along an edge

Unbounded solution



Optimal solution not found--region unbounded

--------- Lines of constant Z

Equivalent Optimal Solutions

Equivalent optimal solutions occur frequently in genome-scale networks.
Since, Genome-scale networks are typically able to achieve the same overall functional network state in many different ways.

Flux Balance Analysis

Flux balance analysis (FBA) applies defined **constraints**
on an **objective function**
and find the optimum solutions (flux profiles)

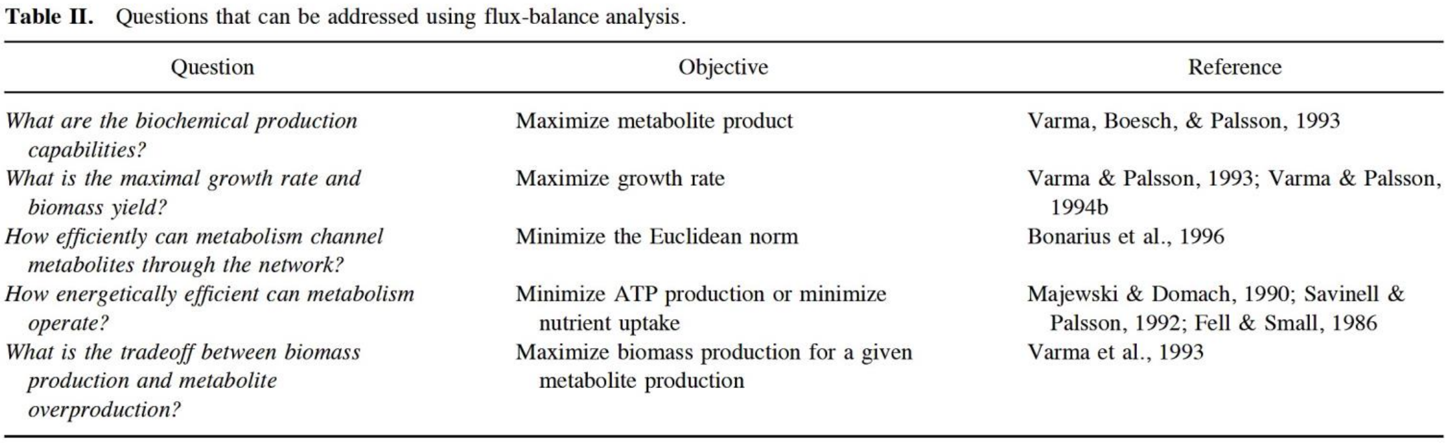


Table II. Questions that can be addressed using flux-balance analysis.

Question	Objective	Reference
What are the biochemical production capabilities?	Maximize metabolite product	Varma, Boesch, & Palsson, 1993
What is the maximal growth rate and biomass yield?	Maximize growth rate	Varma & Palsson, 1993; Varma & Palsson, 1994b
How efficiently can metabolism channel metabolites through the network?	Minimize the Euclidean norm	Bonarius et al., 1996
How energetically efficient can metabolism operate?	Minimize ATP production or minimize nutrient uptake	Majewski & Domach, 1990; Savinell & Palsson, 1992; Fell & Small, 1986
What is the tradeoff between biomass production and metabolite overproduction?	Maximize biomass production for a given metabolite production	Varma et al., 1993

For a given growth condition (e.g. known input nutrients), considering:

- metabolic system operates in a **quasi-steady state**.
- certain **constraints** on system (flux limitations, stoichiometric and reversibility constraints).
- an “objective” that is expected to be maximized (e.g. **biomass** production).

FBA ***predicts*** reaction fluxes and essential enzymes
under a given growth condition

Table from: Edwards, J. S., Palsson, B. O., How will bioinformatics influence metabolic engineering? *Biotechnol Bioeng* 1998

Other biological objectives

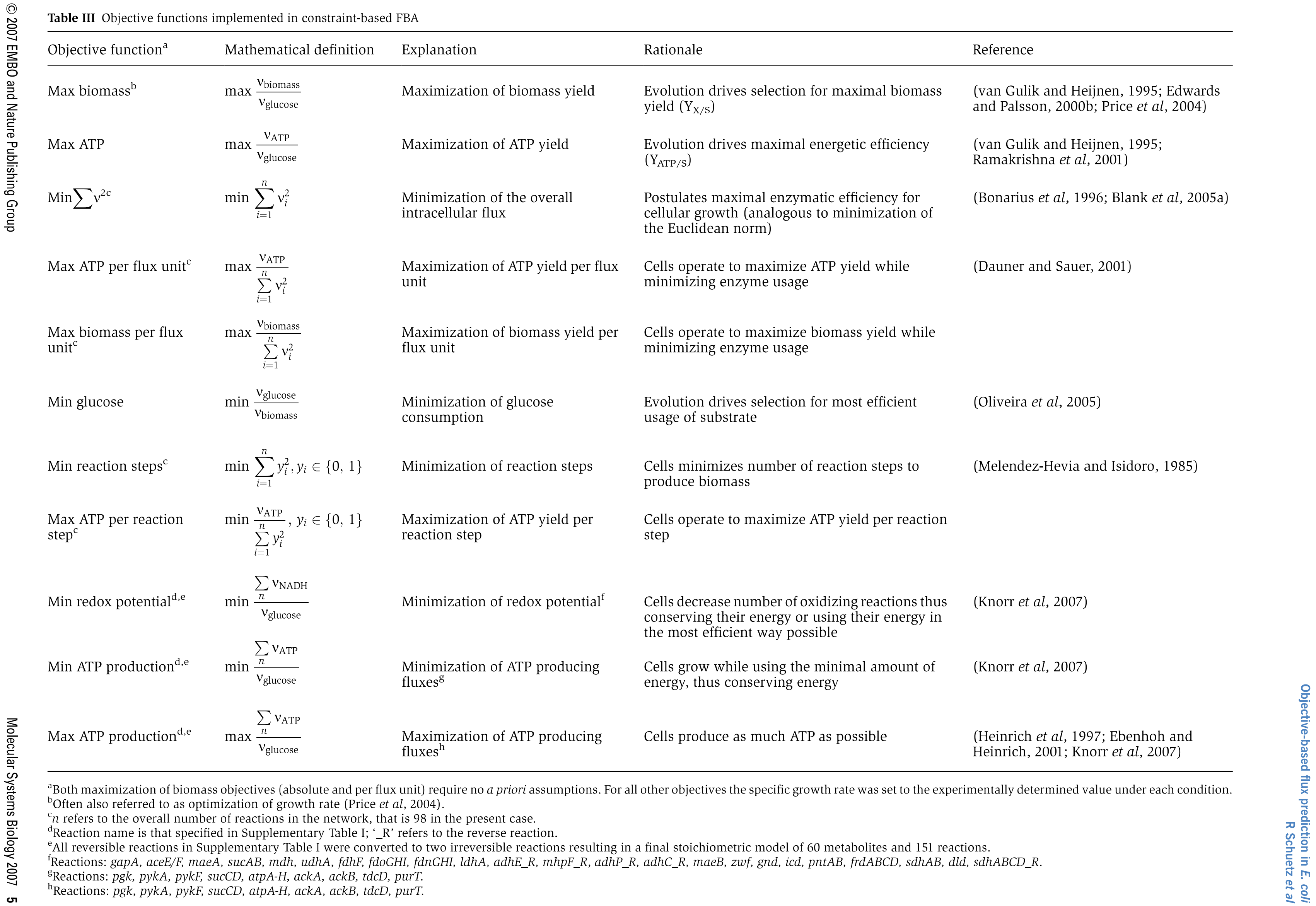
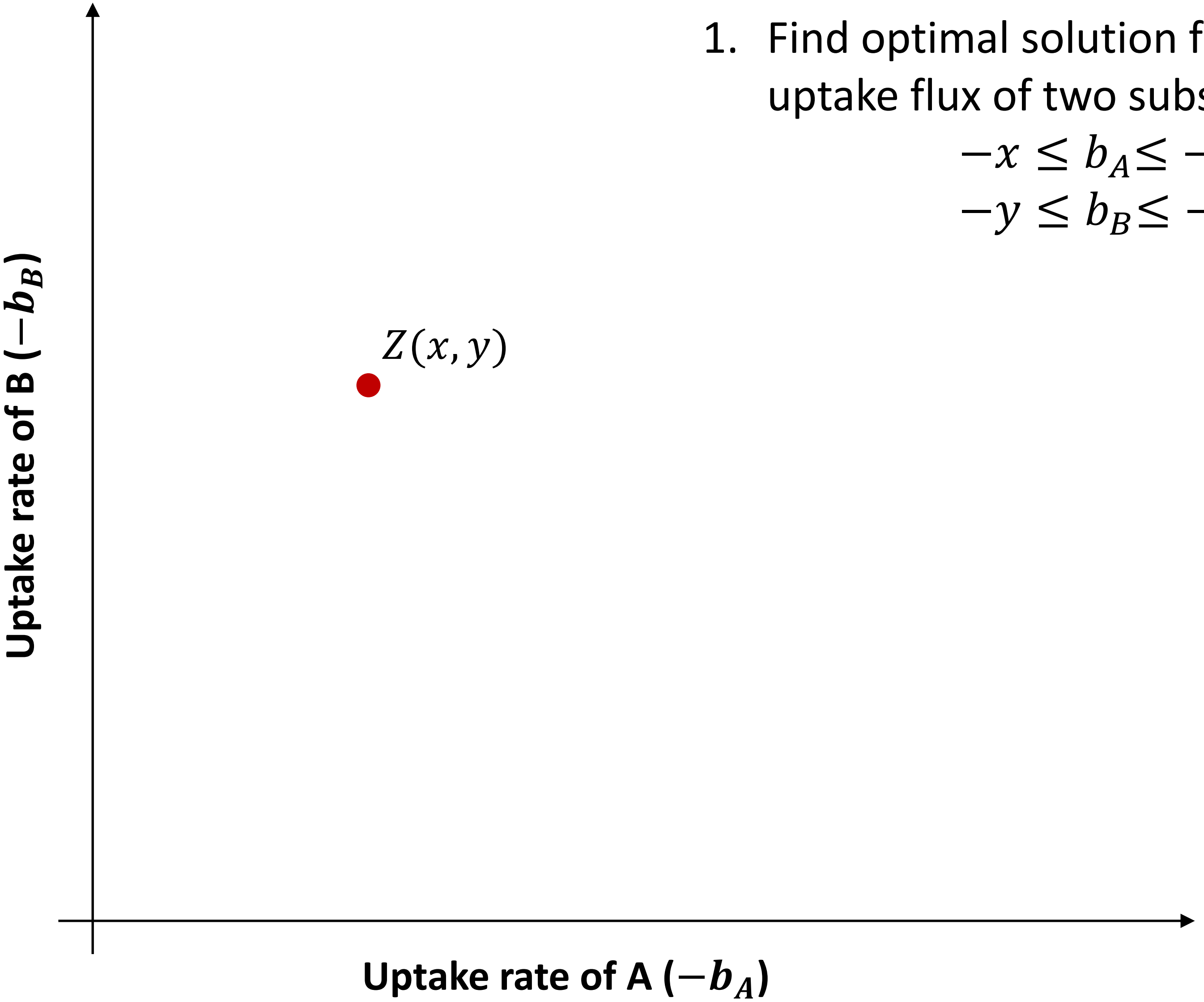


Table III Objective functions implemented in constraint-based FBA

Objective function[a]	Mathematical definition	Explanation	Rationale	Reference
Max biomass[b]	$\max \frac{\nu_{biomass}}{\nu_{glucose}}$	Maximization of biomass yield	Evolution drives selection for maximal biomass yield ($Y_{X/S}$)	(van Gulik and Heijnen, 1995; Edwards and Palsson, 2000b; Price *et al*, 2004)
Max ATP	$\max \frac{\nu_{ATP}}{\nu_{glucose}}$	Maximization of ATP yield	Evolution drives maximal energetic efficiency ($Y_{ATP/S}$)	(van Gulik and Heijnen, 1995; Ramakrishna *et al*, 2001)
Min$\sum \nu^{2}$[c]	$\min \sum_{i=1}^{n} \nu_i^2$	Minimization of the overall intracellular flux	Postulates maximal enzymatic efficiency for cellular growth (analogous to minimization of the Euclidean norm)	(Bonarius *et al*, 1996; Blank *et al*, 2005a)
Max ATP per flux unit[c]	$\max \frac{\nu_{ATP}}{\sum_{i=1}^{n} \nu_i^2}$	Maximization of ATP yield per flux unit	Cells operate to maximize ATP yield while minimizing enzyme usage	(Dauner and Sauer, 2001)
Max biomass per flux unit[c]	$\max \frac{\nu_{biomass}}{\sum_{i=1}^{n} \nu_i^2}$	Maximization of biomass yield per flux unit	Cells operate to maximize biomass yield while minimizing enzyme usage	
Min glucose	$\min \frac{\nu_{glucose}}{\nu_{biomass}}$	Minimization of glucose consumption	Evolution drives selection for most efficient usage of substrate	(Oliveira *et al*, 2005)
Min reaction steps[c]	$\min \sum_{i=1}^{n} y_i^2, y_i \in \{0, 1\}$	Minimization of reaction steps	Cells minimizes number of reaction steps to produce biomass	(Melendez-Hevia and Isidoro, 1985)
Max ATP per reaction step[c]	$\min \frac{\nu_{ATP}}{\sum_{i=1}^{n} y_i^2}, y_i \in \{0, 1\}$	Maximization of ATP yield per reaction step	Cells operate to maximize ATP yield per reaction step	
Min redox potential[d,e]	$\min \frac{\sum_{n} \nu_{NADH}}{\nu_{glucose}}$	Minimization of redox potential[f]	Cells decrease number of oxidizing reactions thus conserving their energy or using their energy in the most efficient way possible	(Knorr *et al*, 2007)
Min ATP production[d,e]	$\min \frac{\sum_{n} \nu_{ATP}}{\nu_{glucose}}$	Minimization of ATP producing fluxes[g]	Cells grow while using the minimal amount of energy, thus conserving energy	(Knorr *et al*, 2007)
Max ATP production[d,e]	$\max \frac{\sum_{n} \nu_{ATP}}{\nu_{glucose}}$	Maximization of ATP producing fluxes[h]	Cells produce as much ATP as possible	(Heinrich *et al*, 1997; Ebenhoh and Heinrich, 2001; Knorr *et al*, 2007)

[a] Both maximization of biomass objectives (absolute and per flux unit) require no *a priori* assumptions. For all other objectives the specific growth rate was set to the experimentally determined value under each condition.
[b] Often also referred to as optimization of growth rate (Price *et al*, 2004).
[c] n refers to the overall number of reactions in the network, that is 98 in the present case.
[d] Reaction name is that specified in Supplementary Table I; '_R' refers to the reverse reaction.
[e] All reversible reactions in Supplementary Table I were converted to two irreversible reactions resulting in a final stoichiometric model of 60 metabolites and 151 reactions.
[f] Reactions: *gapA, aceE/F, maeA, sucAB, mdh, udhA, fdhF, fdoGHI, fdnGHI, ldhA, adhE_R, mhpF_R, adhP_R, adhC_R, maeB, zwf, gnd, icd, pntAB, frdABCD, sdhAB, dld, sdhABCD_R*.
[g] Reactions: *pgk, pykA, pykF, sucCD, atpA-H, ackA, ackB, tdcD, purT*.
[h] Reactions: *pgk, pykA, pykF, sucCD, atpA-H, ackA, ackB, tdcD, purT*.

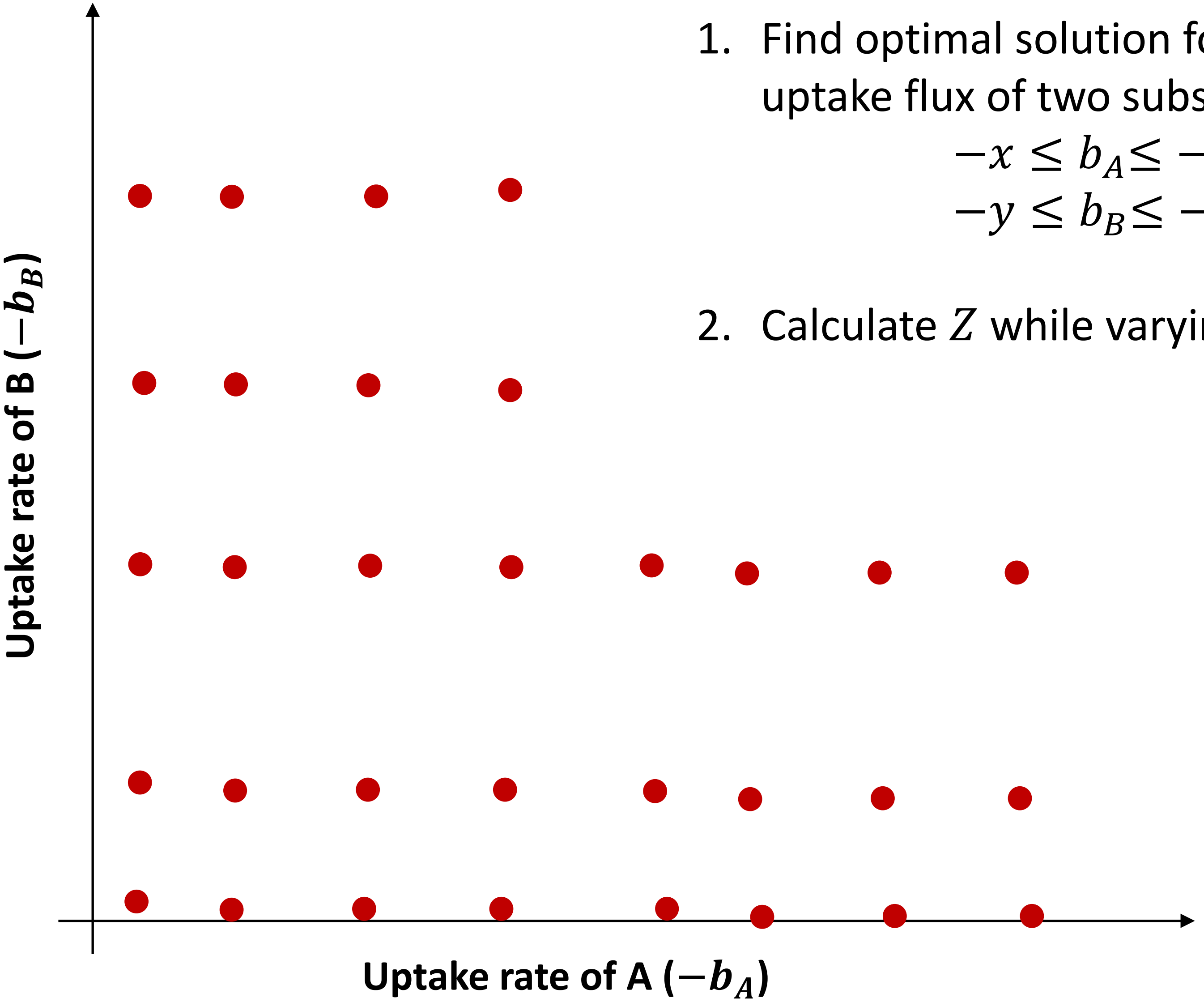
Phenotypic phase planes



1. Find optimal solution for a fixed uptake flux of two substrates (b_A and b_B)

$$-x \leq b_A \leq -x$$
$$-y \leq b_B \leq -y$$

Phenotypic phase planes

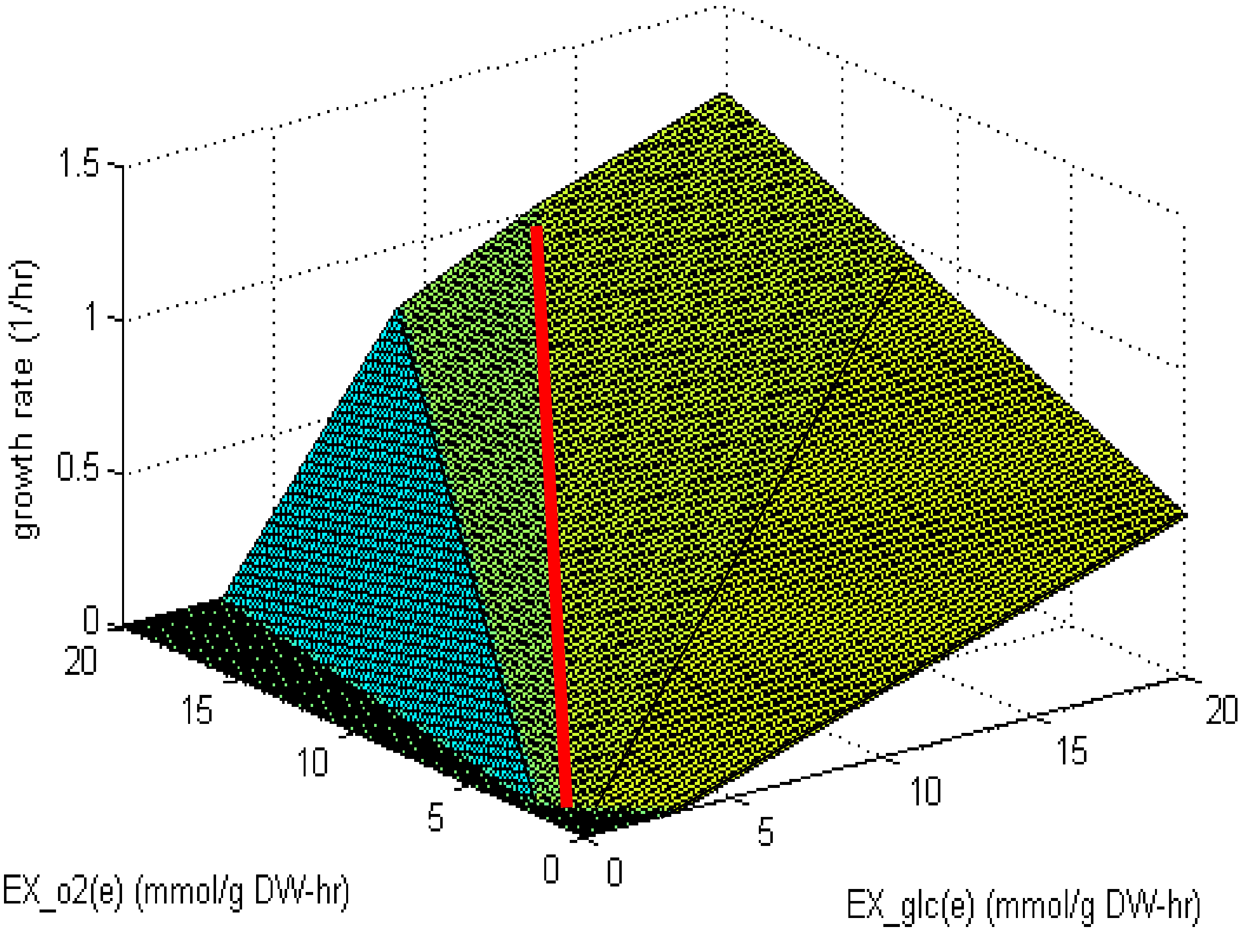


1. Find optimal solution for a fixed uptake flux of two substrates (b_A and b_B)

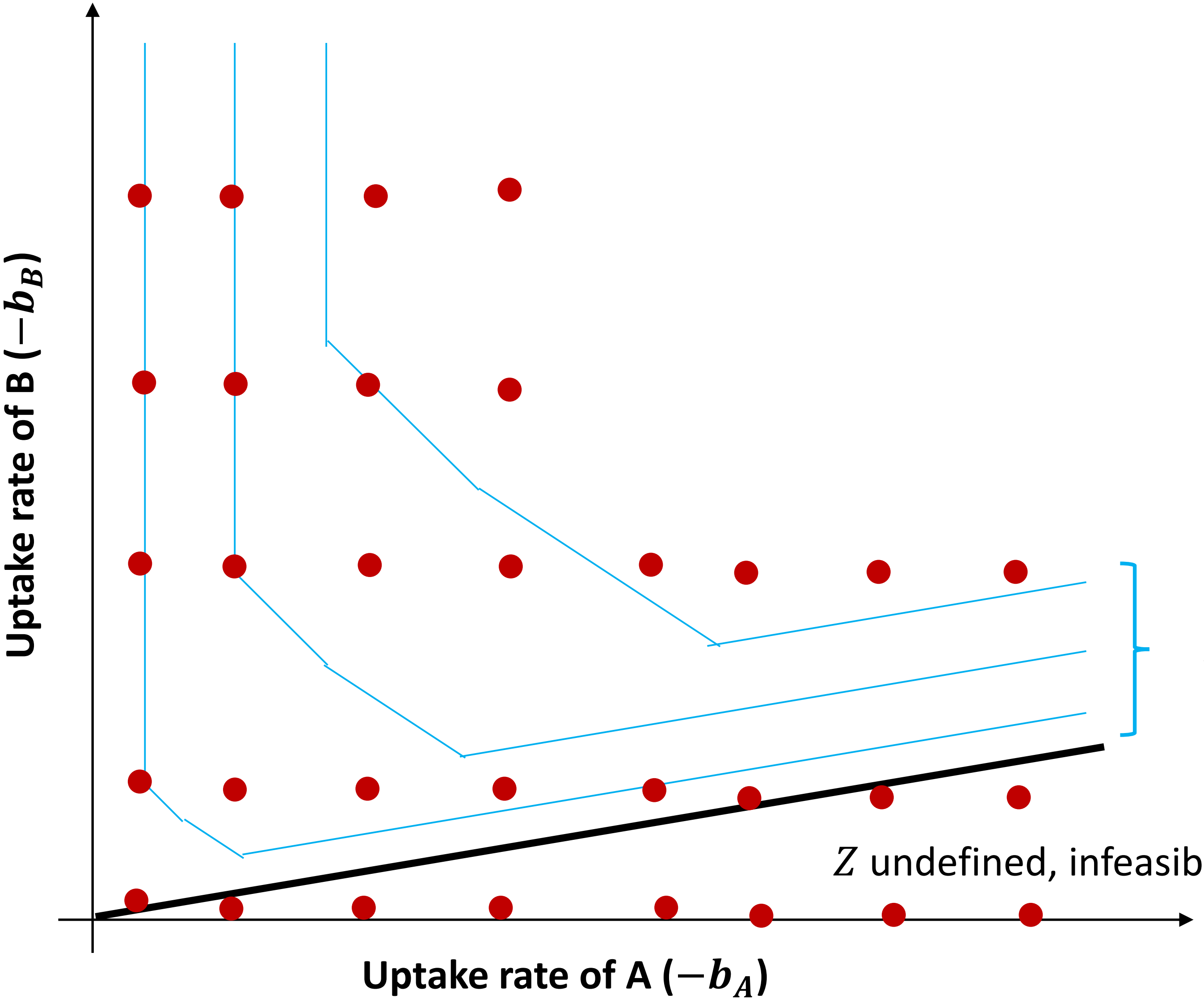
$$-x \leq b_A \leq -x$$
$$-y \leq b_B \leq -y$$

2. Calculate Z while varying x and y

Phenotypic phase planes



Phenotypic phase planes

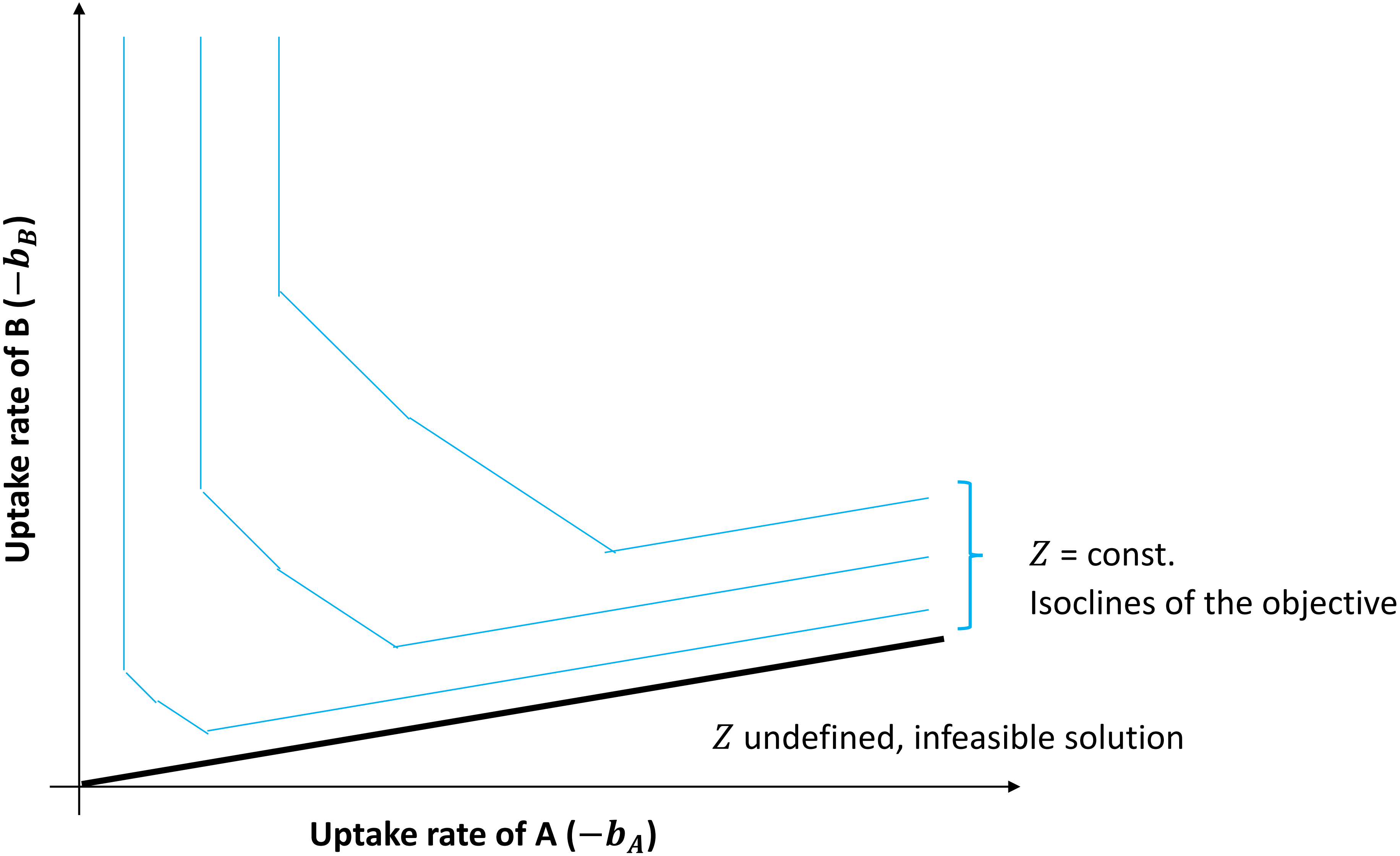


Z = const.

Isoclines of the objective

Z undefined, infeasible solution

Phenotypic phase planes



Z = const.

Isoclines of the objective

Z undefined, infeasible solution

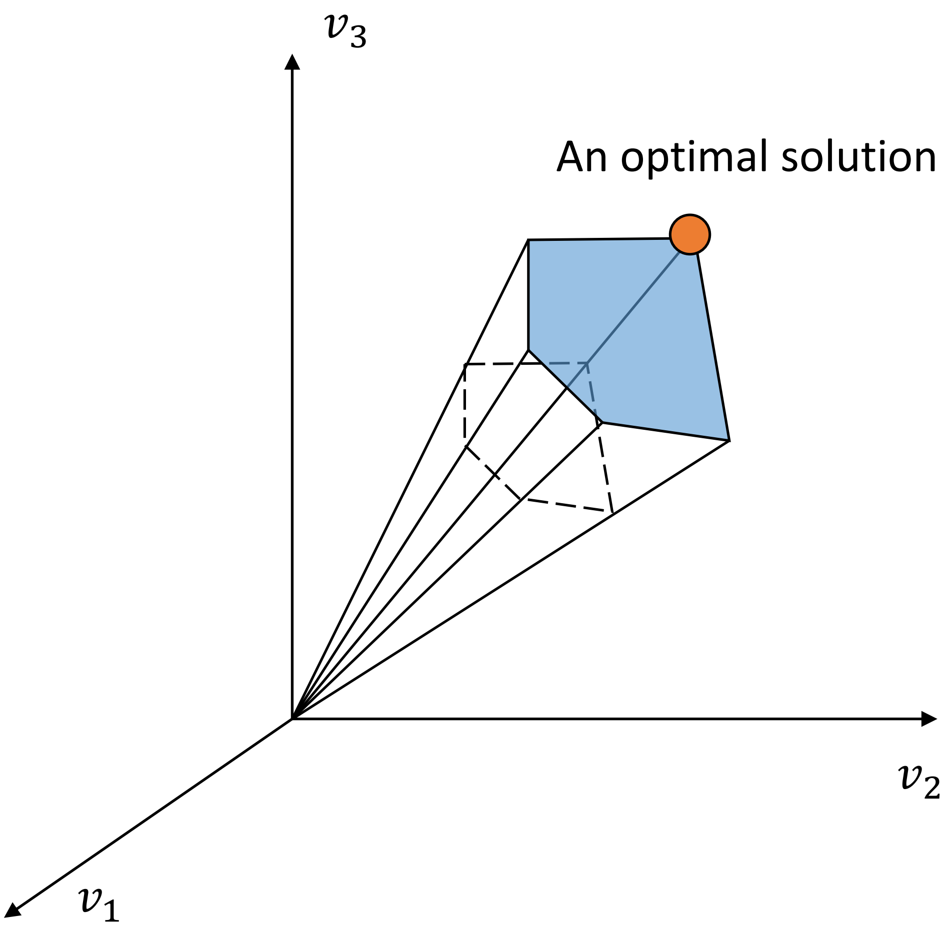
Flux Variability Analysis

Flux Variability Analysis

1) Perform an **FBA** with a given objective e.g:

$$S\vec{v} = \vec{0} \text{ with } v_{j,lb} \leq v_j \leq v_{j,ub}$$

$$b_{max} = max(biomass)$$



Flux Variability Analysis

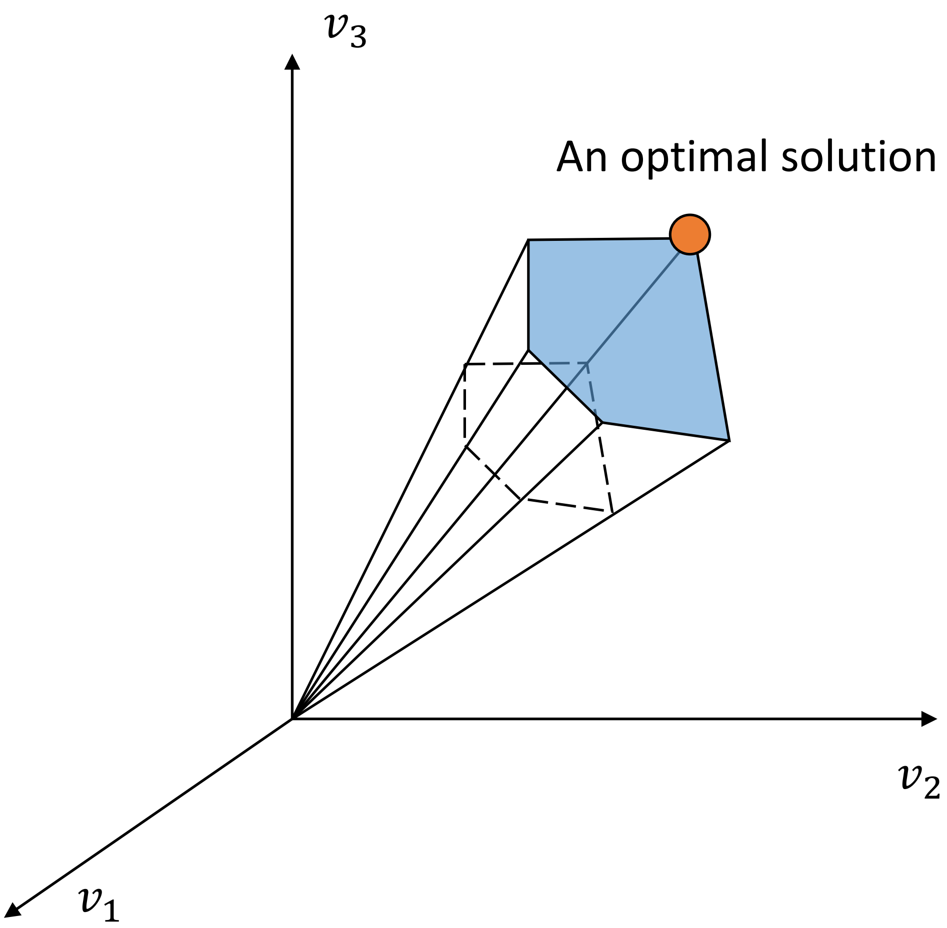
1) Perform an **FBA** with a given objective e.g:

$$S\vec{v} = \vec{0} \text{ with } v_{j,lb} \leq v_j \leq v_{j,ub}$$

$$b_{max} = \ max(biomass)$$

2) **Constraint** the **objective** e.g:

$$b_{max} \leq biomass$$



Flux Variability Analysis

1) Perform an **FBA** with a given objective e.g:

$$S\vec{v} = \vec{0} \text{ with } v_{j,lb} \leq v_j \leq v_{j,ub}$$

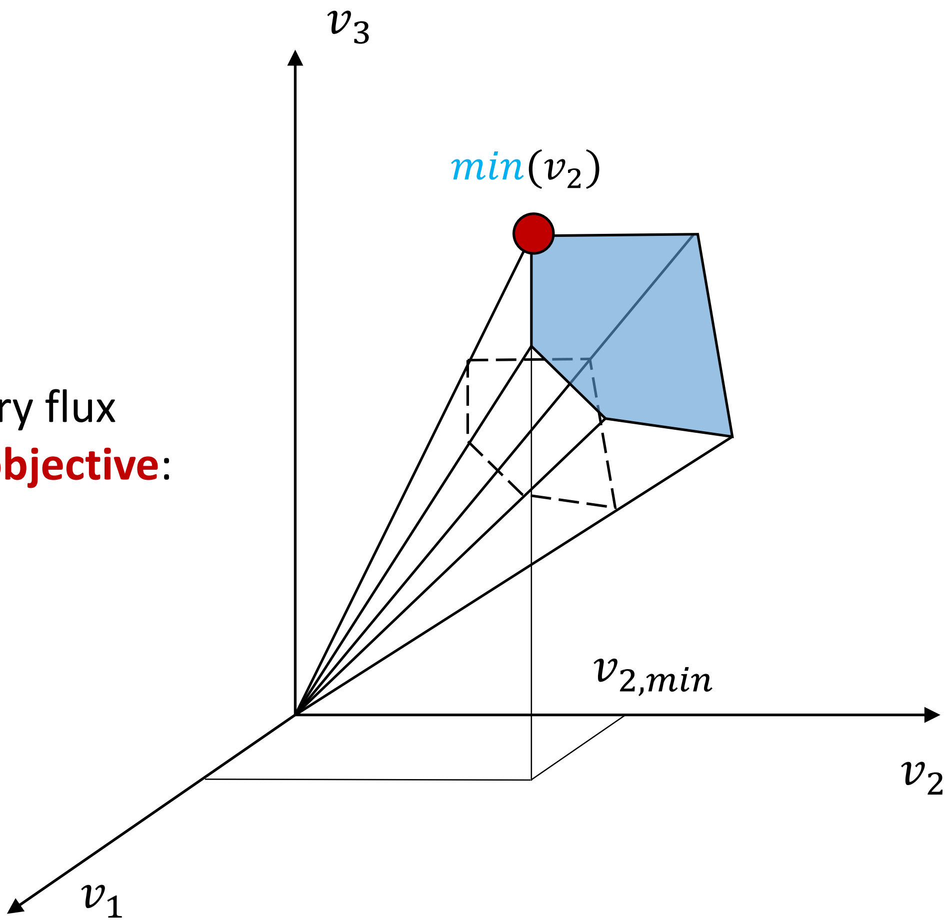
$$b_{max} = max(biomass)$$

2) **Constraint** the **objective** e.g:

$$b_{max} \leq biomass$$

3) Find **minimum** and **maximum** of every flux given the **constraint** of the **optimal objective**:

```
for $v_i$ in $\vec{v}$:
    $S\vec{v} = \vec{0}$ with:
            $v_{j,lb} \leq v_j \leq v_{j,ub}$ and
            $v_{max} \leq biomass$
    $v_{i,min} = min(v_i)$
    $v_{i,max} = max(v_i)$
end
```



Flux Variability Analysis

1) Perform an **FBA** with a given objective e.g:

$$S\vec{v} = \vec{0} \text{ with } v_{j,lb} \leq v_j \leq v_{j,ub}$$

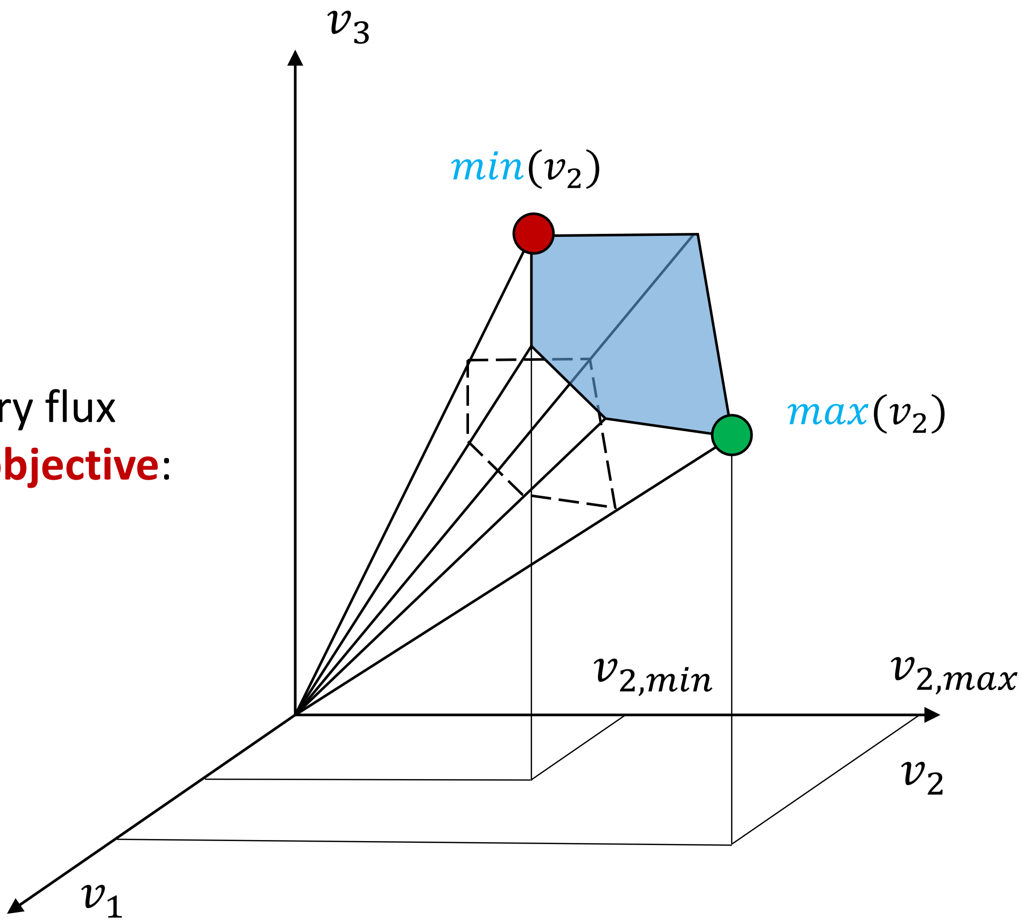
$$b_{max} = max(biomass)$$

2) **Constraint** the **objective** e.g:

$$b_{max} \leq biomass$$

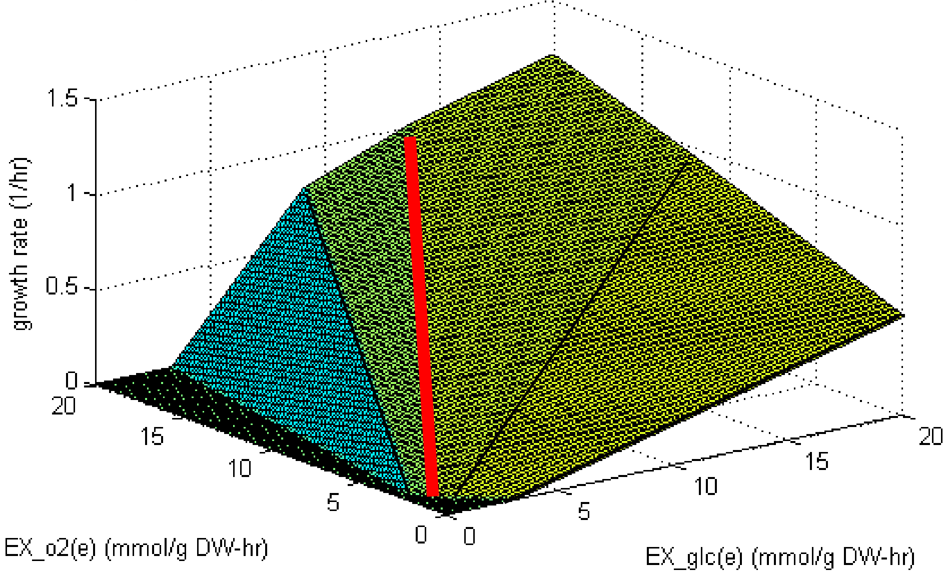
3) Find **minimum** and **maximum** of every flux given the **constraint** of the **optimal objective**:

```
for $v_i$ in $\vec{v}$:
    $S\vec{v} = \vec{0}$ with:
            $v_{j,lb} \leq v_j \leq v_{j,ub}$ and
            $v_{max} \leq biomass$
    $v_{i,min} = min(v_i)$
    $v_{i,max} = max(v_i)$
end
```



Phenotypic phase planes

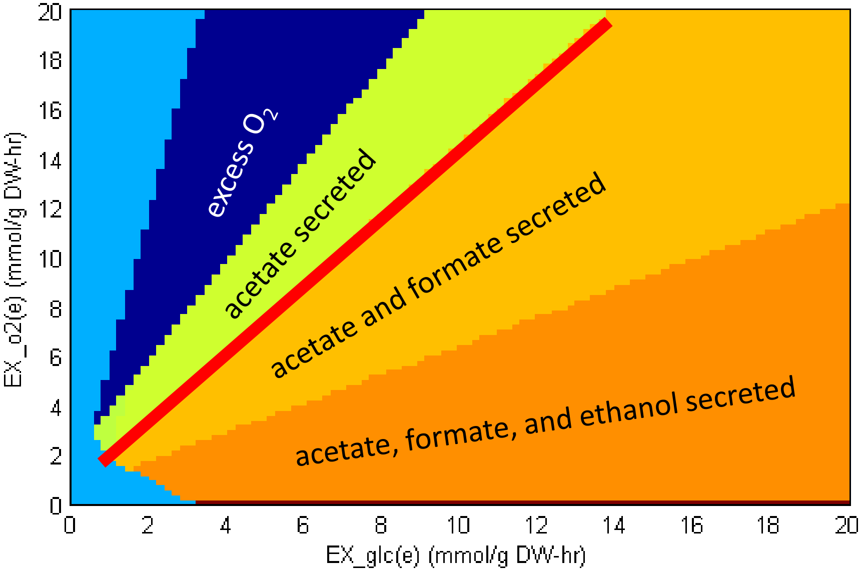
Line of optimality



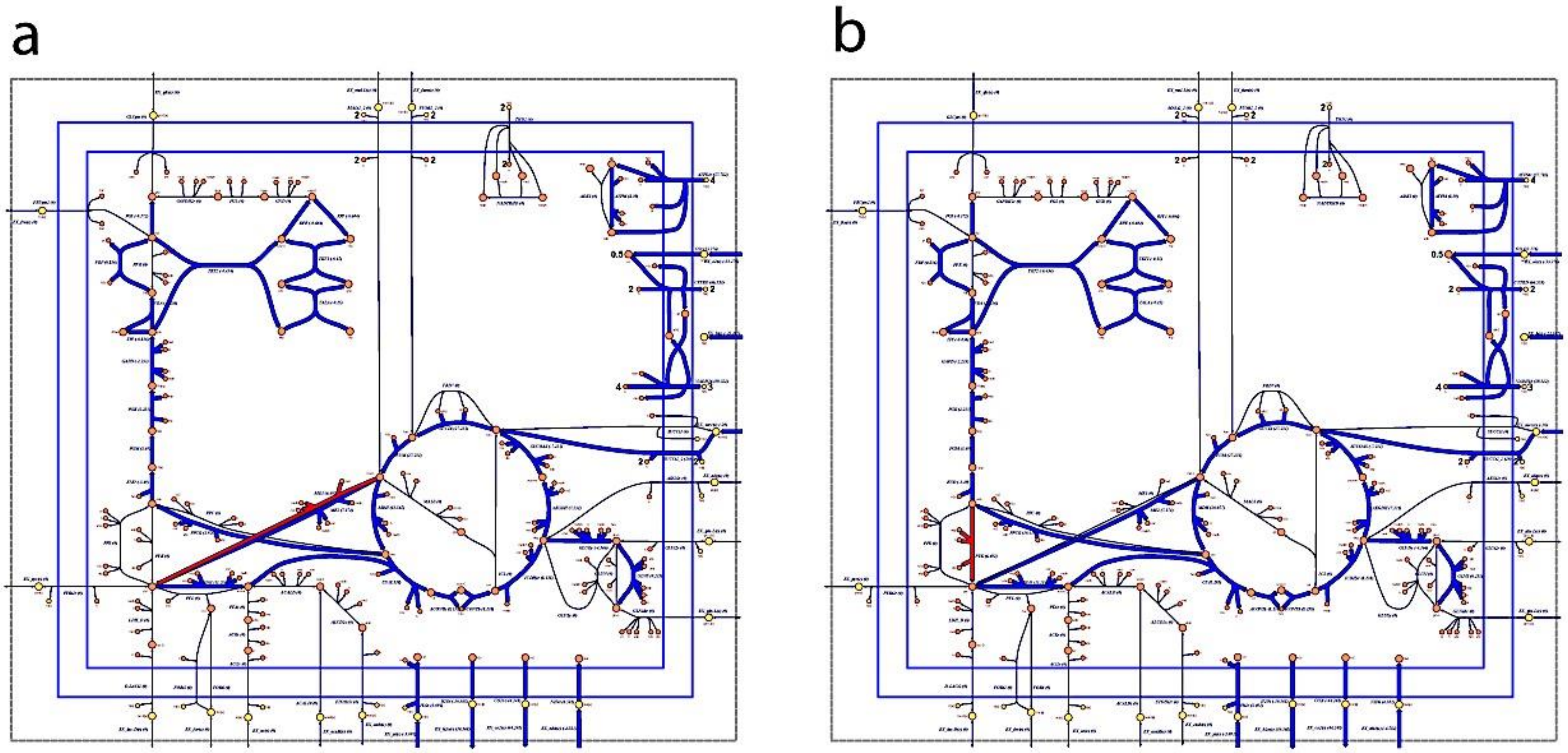
Line of optimality:
Maximal biomass yield with respect to a carbon source (no oxygen limitation!)

Line of optimality

Line of optimality



Alternative flux profiles



Reaction	Minimum Flux (mmol gDW^{-1} hr^{-1})	Maximum Flux (mmol gDW^{-1} hr^{-1})
FRD7	0	972.77
MDH	13.56	20.06
ME1	0	6.49
ME2	7.17	13.67
NADTRHD	0	6.49
PPCK	3.93	10.42
PYK	0	6.49
SUCDi	27.23	1000

Two alternate solutions for maximum aerobic growth on succinate.
a) Reaction ME1 is used to convert L-malate to pyruvate

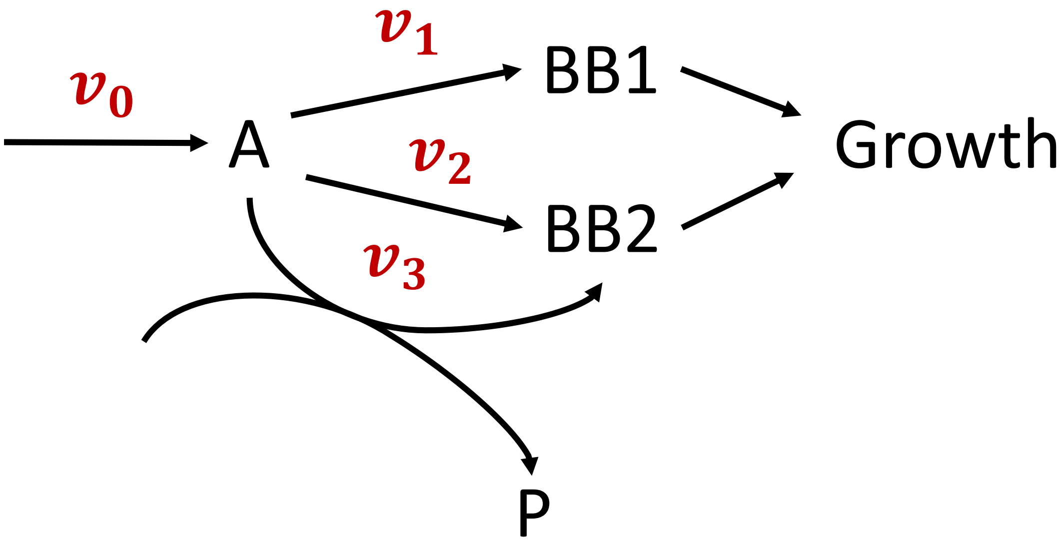
b)The reaction PYK is used to perform this function.

The two alternative reactions are highlighted in red.

Alternative flux profiles

Where do they come from?

$$v_0 = 3\,\frac{\text{mmol}}{gDW\ h}$$

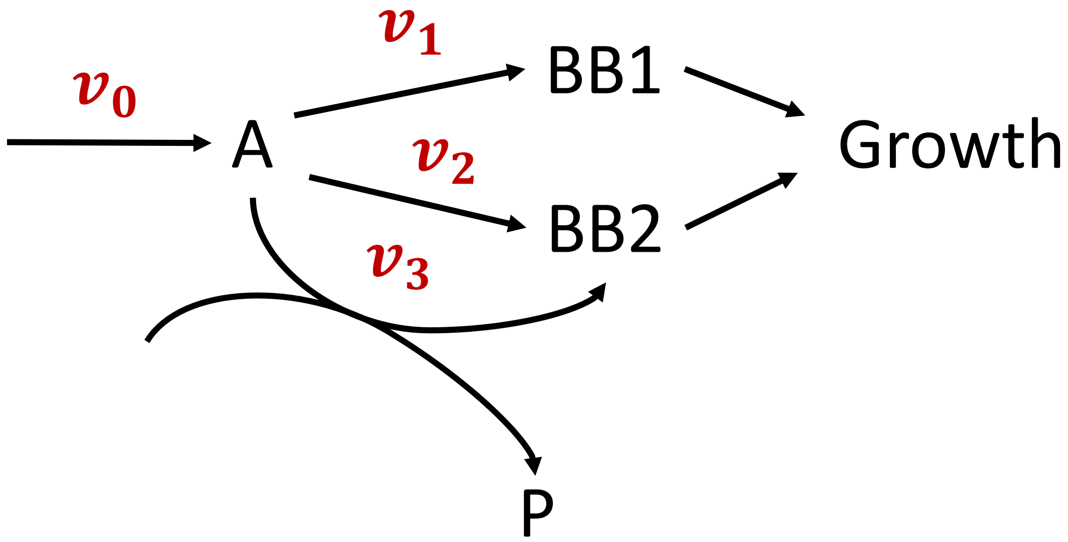


Flux Variability		
Reaction	Min Flux	Max Flux
Growth		

Alternative flux profiles

Where do they come from?

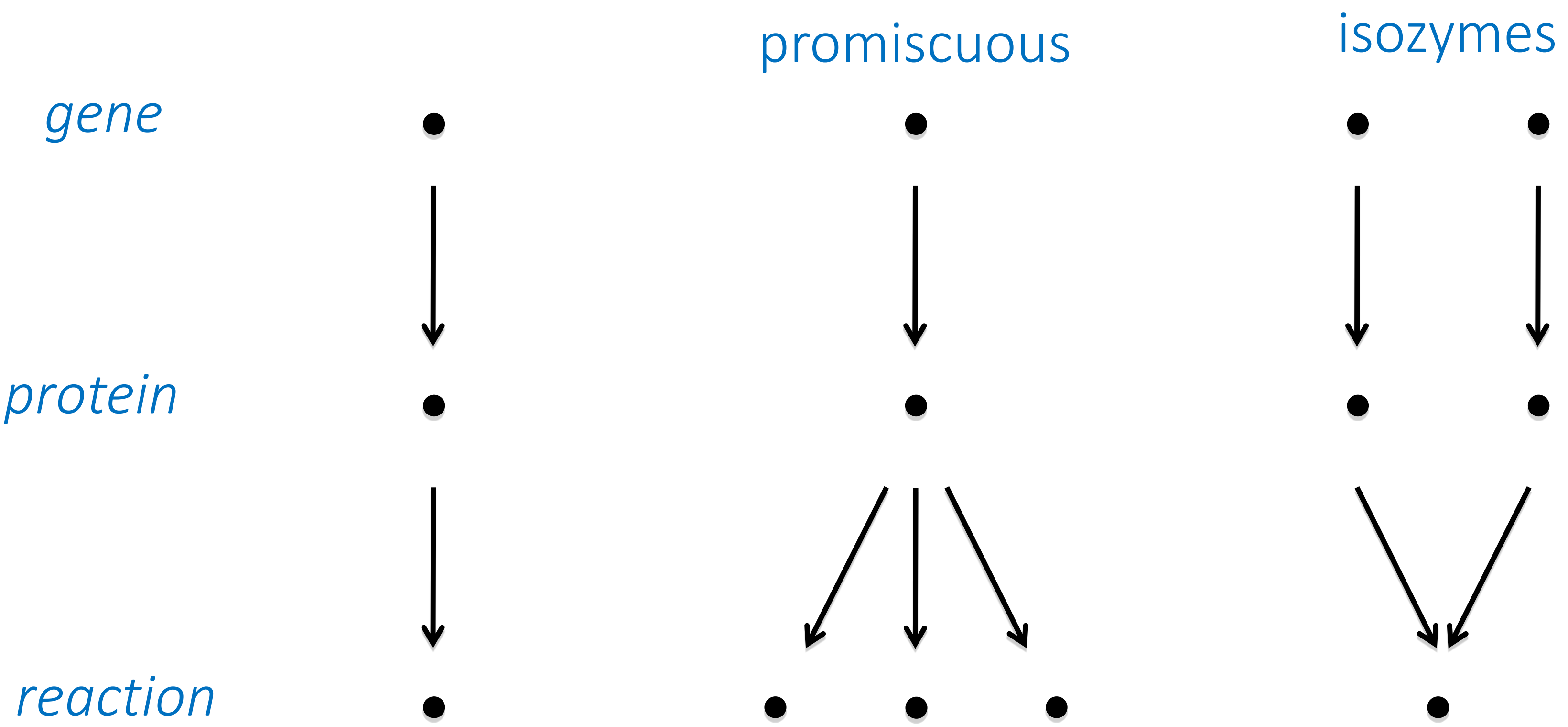
$$v_0 = 3\,\frac{\text{mmol}}{gDW\ h}$$



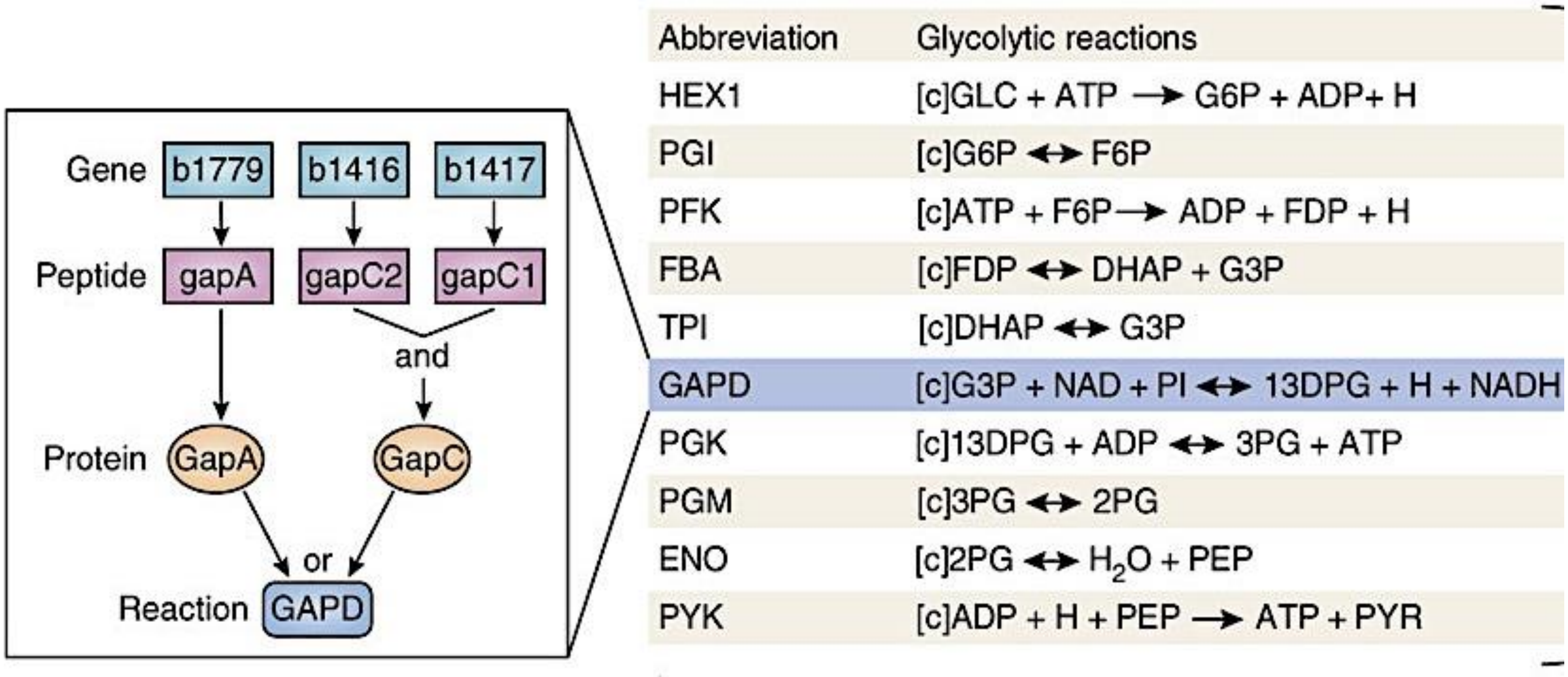
Flux Variability		
Reaction	Min Flux	Max Flux
	1.5	1.5
	0	1.5
	0	1.5
Growth	1.5	1.5

Gene Essentiality

From genes to function



From genes to function

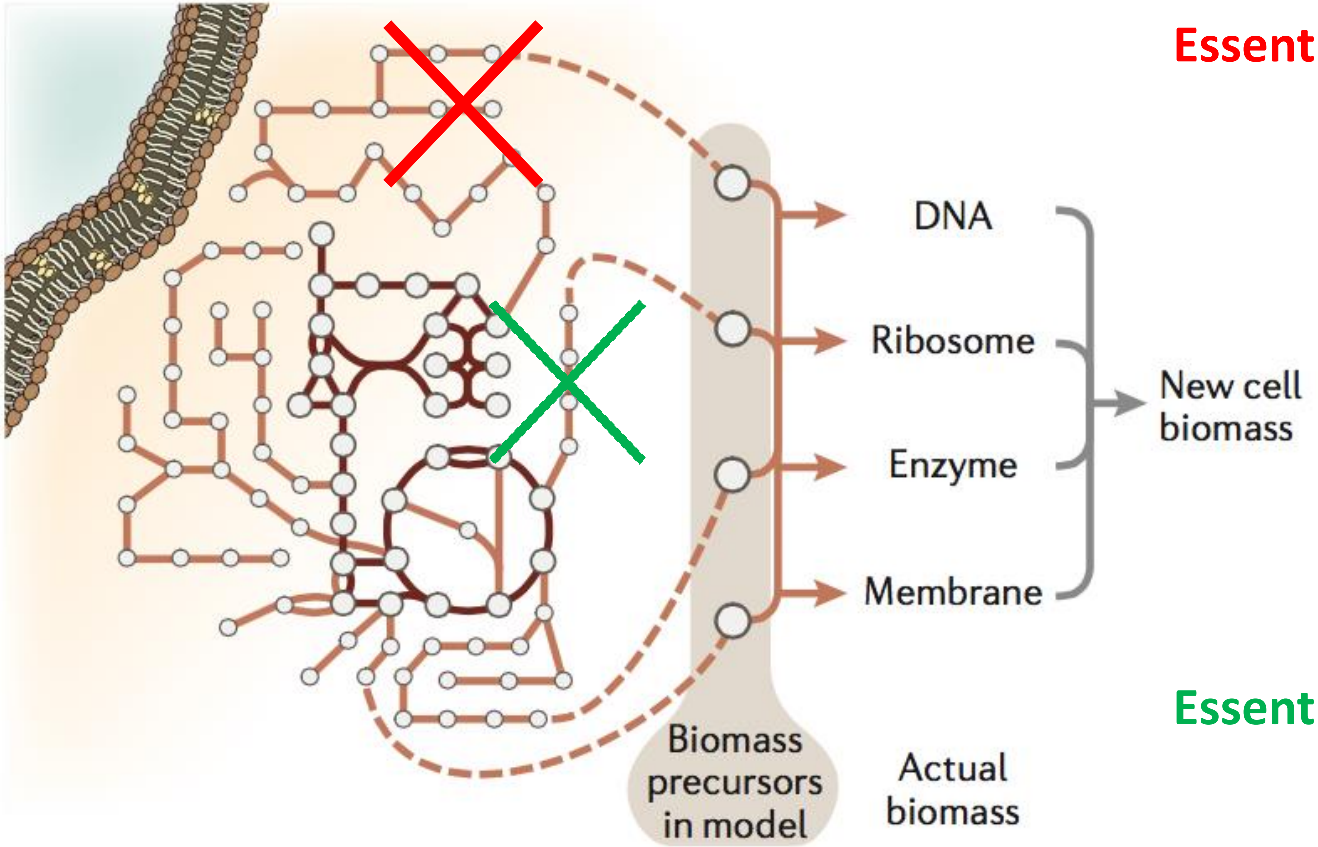


Abbreviation	Glycolytic reactions
HEX1	[c]GLC + ATP → G6P + ADP+ H
PGI	[c]G6P ↔ F6P
PFK	[c]ATP + F6P→ ADP + FDP + H
FBA	[c]FDP ↔ DHAP + G3P
TPI	[c]DHAP ↔ G3P
GAPD	[c]G3P + NAD + PI ↔ 13DPG + H + NADH
PGK	[c]13DPG + ADP ↔ 3PG + ATP
PGM	[c]3PG ↔ 2PG
ENO	[c]2PG ↔ H_2O + PEP
PYK	[c]ADP + H + PEP → ATP + PYR

Knockout Studies -> Gene Essentiality

What makes a gene essential ?

- Biomass precursors required for growth,…



Essential for growth?

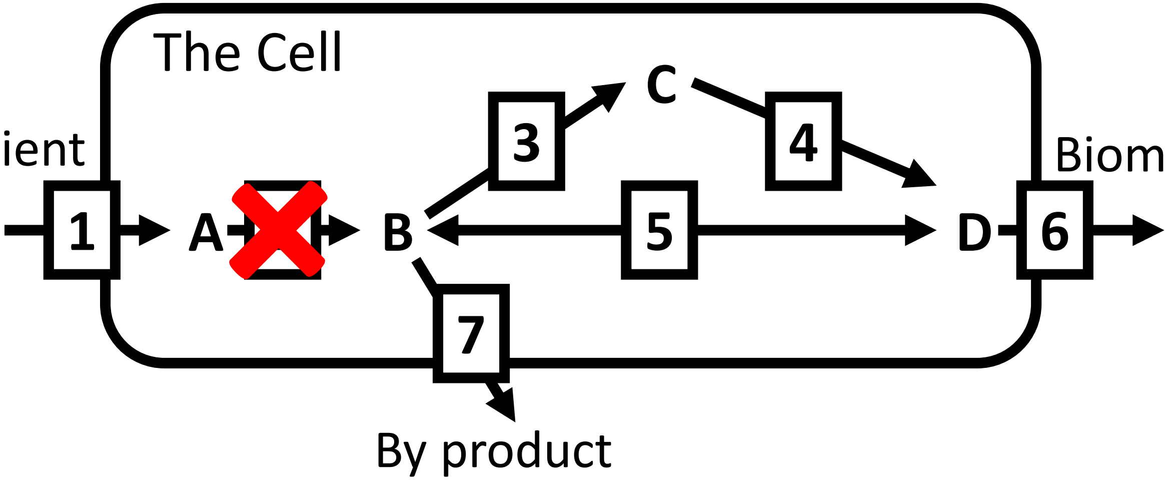
NO!

Essential for growth?

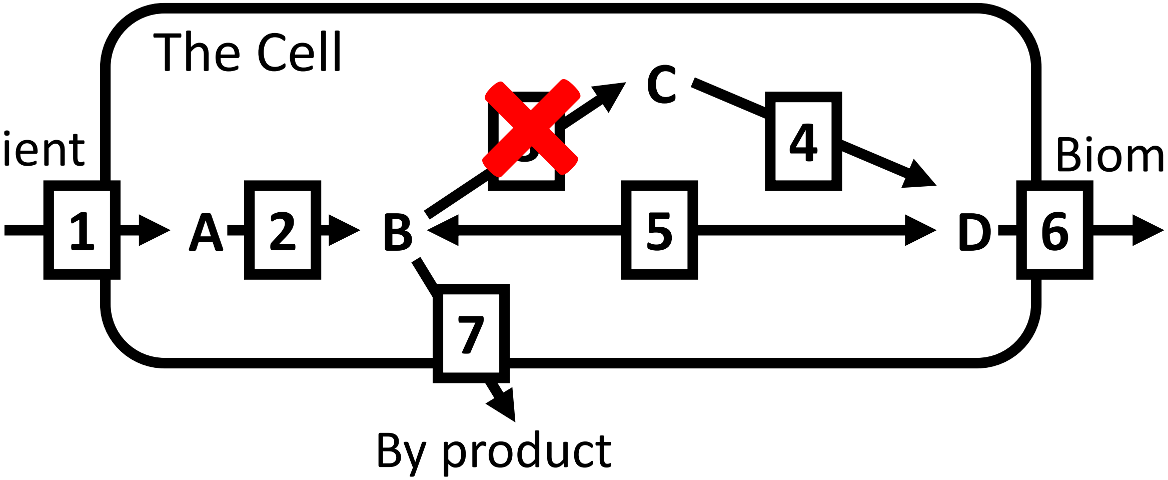
YES!

Lewis, N. E., Nagarajan, H., Palsson, B. O., Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods. *Nature reviews. Microbiology* 2012

Example



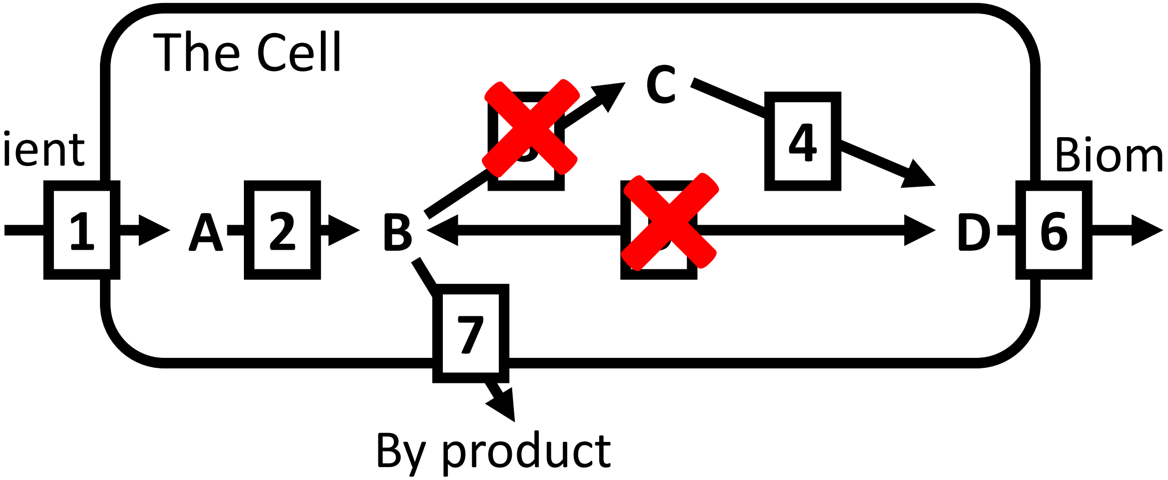
Example



- **Localization:** gene 3 is localized in the cytosol
- **Function:** gene 3 encodes for enzyme 3 that catalyzes reaction from B to C
- **Lethal effect ?:** No the presence of gene 5 enables the production of molecule D that is required for growth

Example

Synthetic lethals



- **Localization:** genes 3 & 5 are localized in the cytosol
- **Function:** both encode for enzymes/pathways that enable production of D
- **Lethal effect:** The knockout of genes 3 & 5 impedes the production of molecule D that is required for growth