Whole cell modeling

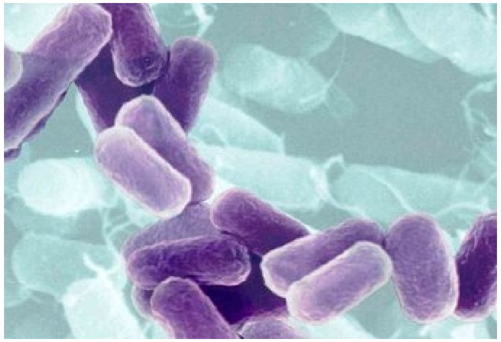
Wolfram Liebermeister, INRA Jouy-en-Josas

wolfram.liebermeister@gmail.com

PSL-ITI – Course "Modelling and engineering of biological systems"

How can a living being emerge from non-living matter?

How can a living cell emerge from sugar, water, and a couple of salts?



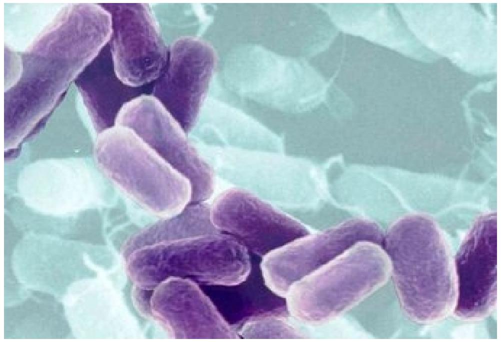
Minimal Medium for *E. coli*

Glucose	5 g/l
Na_2HPO_4	6 g/l
KH_2PO_4	3 g/l
NH_4Cl	1 g/l
NaCl	0.5 g/l
$MgSO_4$	0.12 g/l
$CaCl_2$	0.01 g/l

How can a living cell emerge from sugar, water, and a couple of salts?

How fast can this happen?

How much energy and material will be “wasted”?

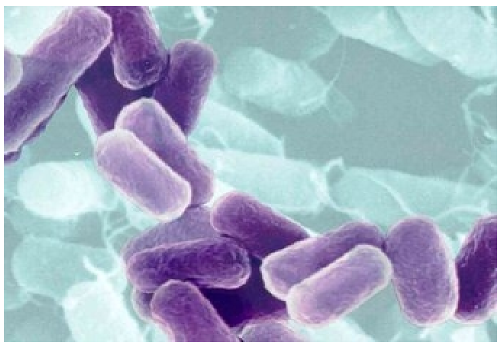


Minimal Medium for *E. coli*

Glucose	5 g/l
Na_2HPO_4	6 g/l
KH_2PO_4	3 g/l
NH_4Cl	1 g/l
NaCl	0.5 g/l
$MgSO_4$	0.12 g/l
$CaCl_2$	0.01 g/l

How can a living cell emerge from sugar, water, and a couple of salts?

How fast can this happen?

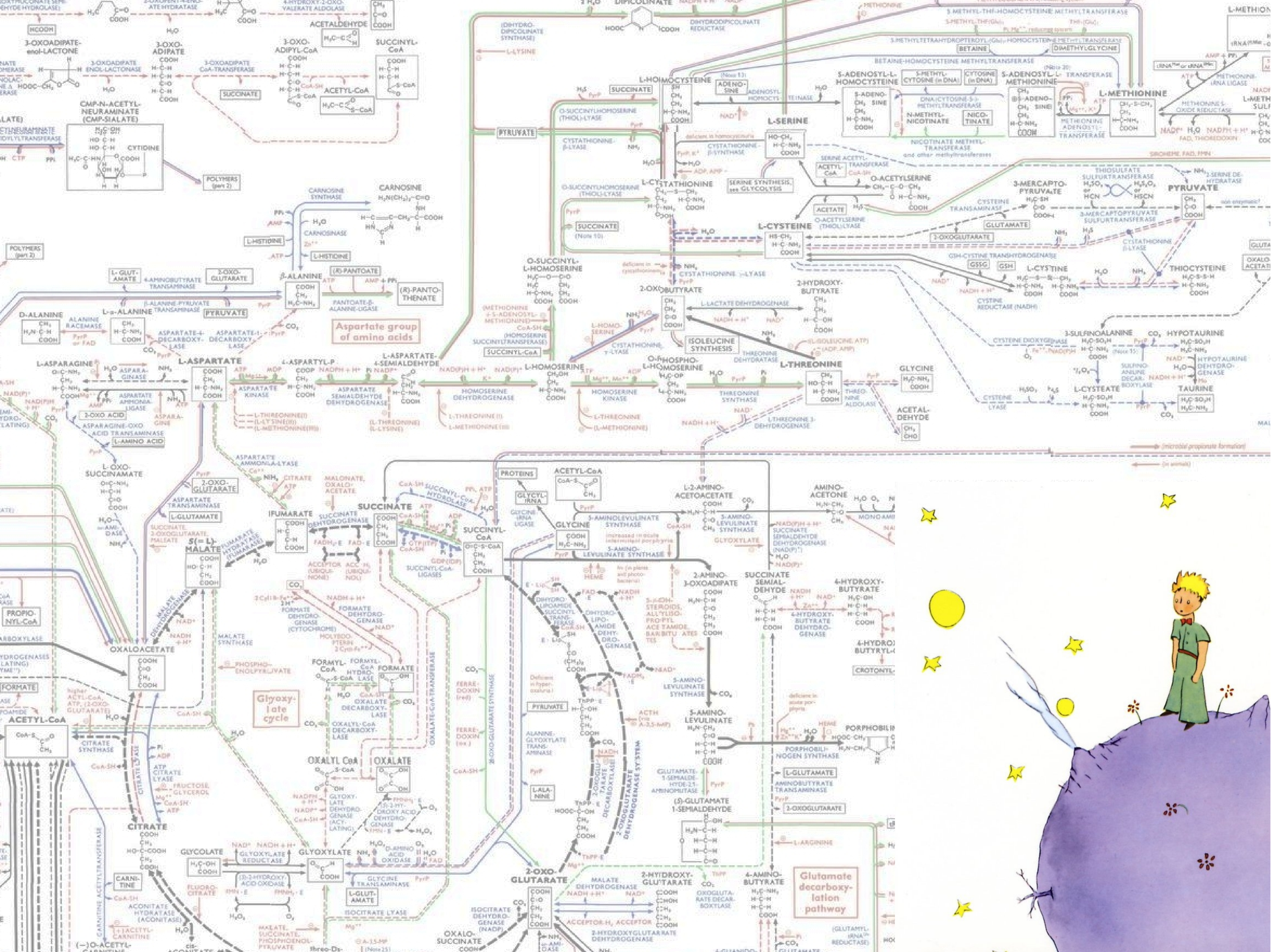


How much energy and material will be "wasted"?

Minimal Medium for *E. coli*

Glucose	5 g/l
Na_2HPO_4	6 g/l
KH_2PO_4	3 g/l
NH_4Cl	1 g/l
NaCl	0.5 g/l
$MgSO_4$	0.12 g/l
$CaCl_2$	0.01 g/l

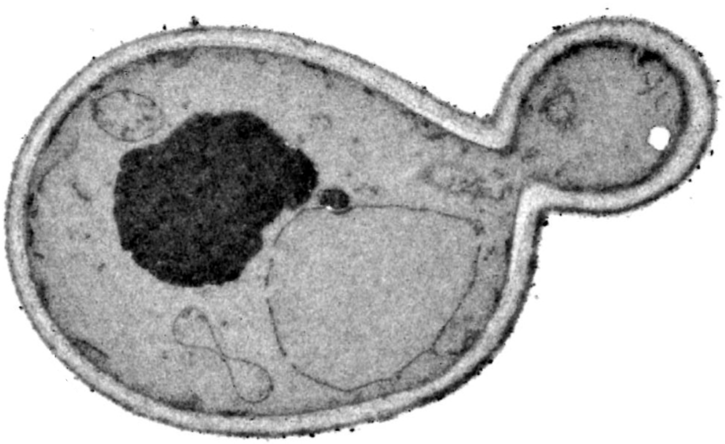
L'essentiel est invisible pour les yeux.



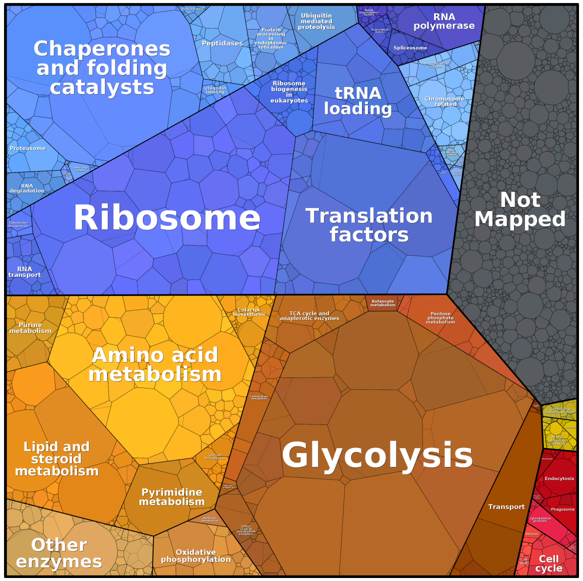
3-OXOADIPATE-enol-LACTONE
3-OXO-ADIPATE
3-OXO-ADIPYL-CoA
SUCCINYL-CoA
ACETALDEHYDE
ACETYL-CoA
3-OXOADIPATE ENOL-LACTONASE
3-OXOADIPATE CoA-TRANSFERASE
SUCCINATE
CMP-N-ACETYL-NEURAMINATE (CMP-SIALATE)
CYTIDINE
POLYMERS (part 2)
DIPICOLINATE
DIHYDRODIPICOLINATE REDUCTASE
L-LYSINE
5-METHYL-THF-HOMOCYSTEINE METHYLTRANSFERASE
BETAINE-HOMOCYSTEINE METHYLTRANSFERASE
BETAINE
DIMETHYLGLYCINE
L-HOMOCYSTEINE
ADENOSINE
S-ADENOSYL-L-HOMOCYSTEINE
S-ADENOSYL-L-METHIONINE
L-METHIONINE
METHIONINE ADENOSYLTRANSFERASE
N-METHYLNICOTINATE
NICOTINATE
NICOTINATE METHYLTRANSFERASE and other methyltransferases
PYRUVATE
CYSTATHIONINE β-LYASE
CYSTATHIONINE β-SYNTHASE
L-SERINE
SERINE ACETYLTRANSFERASE
O-ACETYLSERINE
L-CYSTATHIONINE
O-SUCCINYLHOMOSERINE (THIOL)-LYASE
SERINE SYNTHESIS, see GLYCOLYSIS
3-MERCAPTOPYRUVATE
CYSTEINE TRANSAMINASE
GLUTAMATE
2-OXOGLUTARATE
THIOSULFATE SULFURTRANSFERASE
3-MERCAPTOPYRUVATE SULFURTRANSFERASE
PYRUVATE
L-CYSTEINE
ACETATE
O-ACETYLSERINE (THIOL)-LYASE
GSH-CYSTINE TRANSHYDROGENASE
GSSG
GSH
L-CYSTINE
THIOCYSTEINE
CYSTINE REDUCTASE (NADH)
CYSTATHIONINE γ-LYASE
O-SUCCINYL-L-HOMOSERINE
2-OXOBUTYRATE
2-HYDROXYBUTYRATE
L-LACTATE DEHYDROGENASE
ISOLEUCINE SYNTHESIS
3-SULFINOALANINE
HYPOTAURINE
CYSTEINE DIOXYGENASE
HYPOTAURINE DEHYDROGENASE
L-CYSTEATE
TAURINE
CYSTEINE LYASE
CARNOSINE SYNTHASE
CARNOSINE
CARNOSINASE
L-HISTIDINE
β-ALANINE
(R)-PANTOATE
(R)-PANTOTHENATE
PANTOATE-β-ALANINE LIGASE
4-AMINOBUTYRATE TRANSAMINASE
β-ALANINE-PYRUVATE TRANSAMINASE
L-GLUTAMATE
2-OXOGLUTARATE
D-ALANINE
ALANINE RACEMASE
L-α-ALANINE
Aspartate group of amino acids
ASPARTATE-4-DECARBOXYLASE
ASPARTATE-1-DECARBOXYLASE
L-ASPARAGINE
ASPARAGINASE
ASPARTATE AMMONIA-LIGASE
L-ASPARTATE
4-ASPARTYL-P
ASPARTATE KINASE
L-ASPARTATE-4-SEMIALDEHYDE
ASPARTATE SEMIALDEHYDE DEHYDROGENASE
L-HOMOSERINE
HOMOSERINE DEHYDROGENASE
O-PHOSPHO-L-HOMOSERINE
HOMOSERINE KINASE
THREONINE SYNTHASE
L-THREONINE
THREONINE DEHYDRATASE
THREONINE ALDOLASE
GLYCINE
ACETALDEHYDE
L-THREONINE 3-DEHYDROGENASE
L-2-AMINO-ACETOACETATE
AMINOACETONE
2-OXO ACID
L-AMINO ACID
ASPARAGINE-OXO ACID TRANSAMINASE
ASPARTATE AMMONIA-LYASE
L-OXOSUCCINAMATE
ASPARTATE TRANSAMINASE
FUMARATE
SUCCINATE DEHYDROGENASE
SUCCINATE
SUCCINYL-CoA HYDROLASE
SUCCINYL-CoA
SUCCINYL-CoA LIGASES
PROTEINS
GLYCYL-tRNA
ACETYL-CoA
GLYCINE
5-AMINOLEVULINATE SYNTHASE
FUMARATE HYDRATASE (FUMARASE)
S(−L)-MALATE
MALATE DEHYDROGENASE
MALATE SYNTHASE
OXALOACETATE
PROPIONYL-CoA
FORMATE
FORMATE DEHYDROGENASE (CYTOCHROME)
FORMYL-CoA
FORMYL-CoA HYDROLASE
OXALATE DECARBOXYLASE
OXALYL-CoA DECARBOXYLASE
OXALYL CoA
OXALATE
Glyoxylate cycle
OXALATE-CoA-TRANSFERASE
2-OXOGLUTARATE SYNTHASE
2-OXOGLUTARATE DEHYDROGENASE SYSTEM
DIHYDROLIPOAMIDE SUCCINYLTRANSFERASE
DIHYDROLIPOAMIDE DEHYDROGENASE
PYRUVATE
ALANINE-GLYOXYLATE TRANSAMINASE
L-ALANINE
2-AMINO-3-OXOADIPATE
SUCCINATE SEMIALDEHYDE
SUCCINATE SEMIALDEHYDE DEHYDROGENASE
4-HYDROXYBUTYRATE
4-HYDROXYBUTYRATE DEHYDROGENASE
CROTONYL-
5-AMINOLEVULINATE
PORPHOBILINOGEN SYNTHASE
L-GLUTAMATE
AMINOBUTYRATE TRANSAMINASE
2-OXOGLUTARATE
(S)-GLUTAMATE 1-SEMIALDEHYDE
L-ARGININE
GLUTAMATE-1-SEMIALDEHYDE 2,1-AMINOMUTASE
Glutamate decarboxylation pathway
4-AMINOBUTYRATE
ACETYL-CoA
CITRATE SYNTHASE
CITRATE LYASE
ATP CITRATE LYASE
CITRATE
CARNITINE ACETYLTRANSFERASE
CARNITINE
GLYCOLATE
GLYOXYLATE REDUCTASE
GLYOXYLATE
D-AMINO ACID OXIDASE
GLYCINE TRANSAMINASE
ISOCITRATE LYASE
ACONITATE HYDRATASE (ACONITASE)
2-OXOGLUTARATE
MALATE DEHYDROGENASE
2-HYDROXYGLUTARATE
2-HYDROXYGLUTARATE DEHYDROGENASE
ISOCITRATE DEHYDROGENASE (NADP)
OXALOSUCCINATE

How can a cell reproduce itself?

Budding yeast
(microscope picture of a cell)



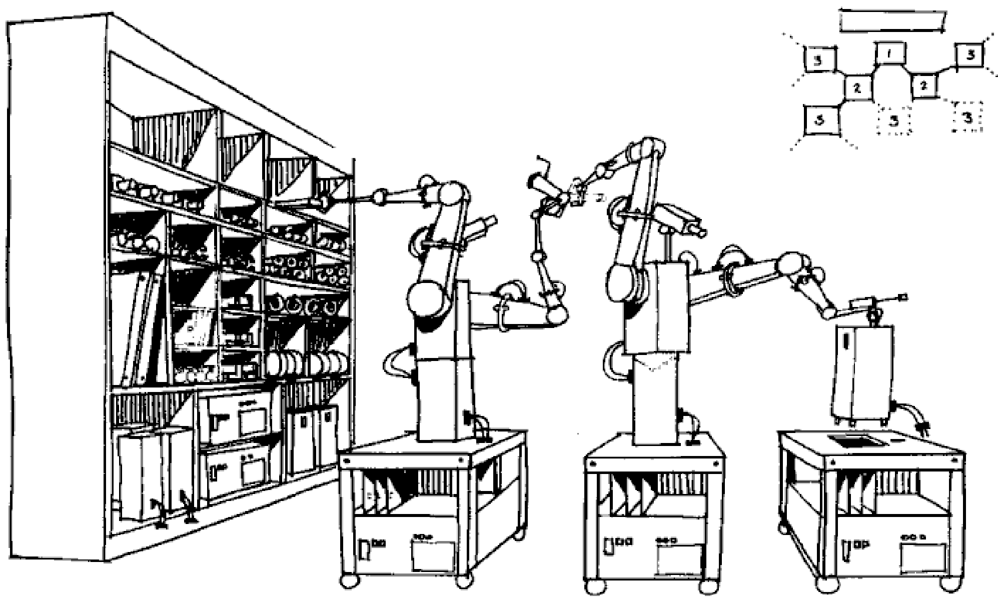
Budding yeast
(relative protein
abundances)



Initial questions for this lecture

- What does a living cell need to do to proliferate, i.e., to reproduce all its components?
- How can it do so in “cost-saving” ways, given physical and biochemical limitations?
- If cells function “optimally”, how will they behave and what will they look like?
- How can we describe all this by mathematical models?

How can a cell reproduce itself?



Analogy: a self-replicating factory

- We consider a factory that produces all the machines it contains, and needs to reproduce itself
- If we know all possible types of machines in this factory, can we predict their optimal numbers?
- Can we predict how these numbers will change depending on energy supply (or energy cost), the need to produce extra products to be sold, the threat of machine failures, etc?

Image source: wikipedia article on self-replicating machines

How can a cell reproduce itself?

Three requirements for living cells to survive

- Cells are alive and fragile – they need to be able to **reproduce (almost) all of their parts**
- Cells are dynamic – they need to **coordinate and regulate all these processes**
- Cells compete – they need to replicate **faster, and be more efficient or resilient** than others

So .. what do we need to study the "economics" of self-replication?

- A **blueprint of the cell** in question
- **Simulation models** for cellular processes (e.g., metabolism and protein production), implementing physical laws and biochemical facts
- Ideas about cells "should" function, formulated as **mathematical optimality problems**

Three levels of description

- "Topics" (in the sense of "components and layout")
- Dynamics
- Economics (flows of production, and allocation of resources)

How can a cell reproduce itself?

Why is fast reproduction a relevant task?

- Fast replication (or withstanding harsh conditions) can be critical for cell survival. Thus, **evolutionary pressures may already have shaped cells** to be "economical"
- In biotechnology, one needs to understand the biological systems that one manipulates (not only how they work physically, but also how they are adjusted to work well).

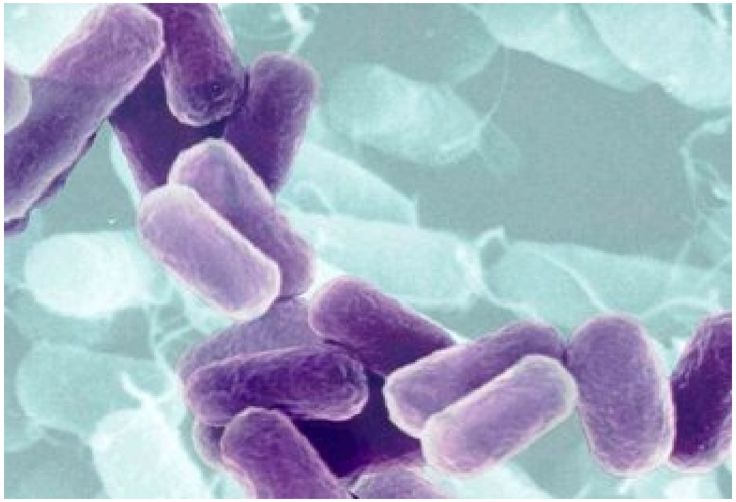
  Genetic modifications that put little burdens on cells are preferable: they increase production or cell growth and reduce the risk to be outcompeted by burden-free mutants.
- Among the many things cells need to do (to survive and proliferate in the long run), self-replication under constant conditions is one important task. it is **easier to understand than many other tasks** (can you think of such other tasks?)
- **Self-replication is just interesting!** A machine that copies itself; a programming language whose compiler is written in the language itself; snowball systems; ...

Part 1: Topics* - A blueprint of the cell

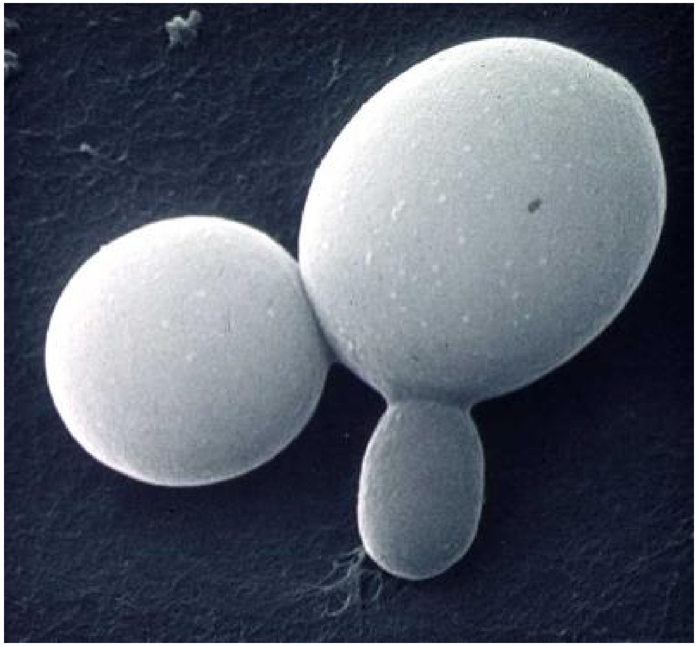
* from topos, place

Microbial cells – external appearance

Escherichia coli
bacteria

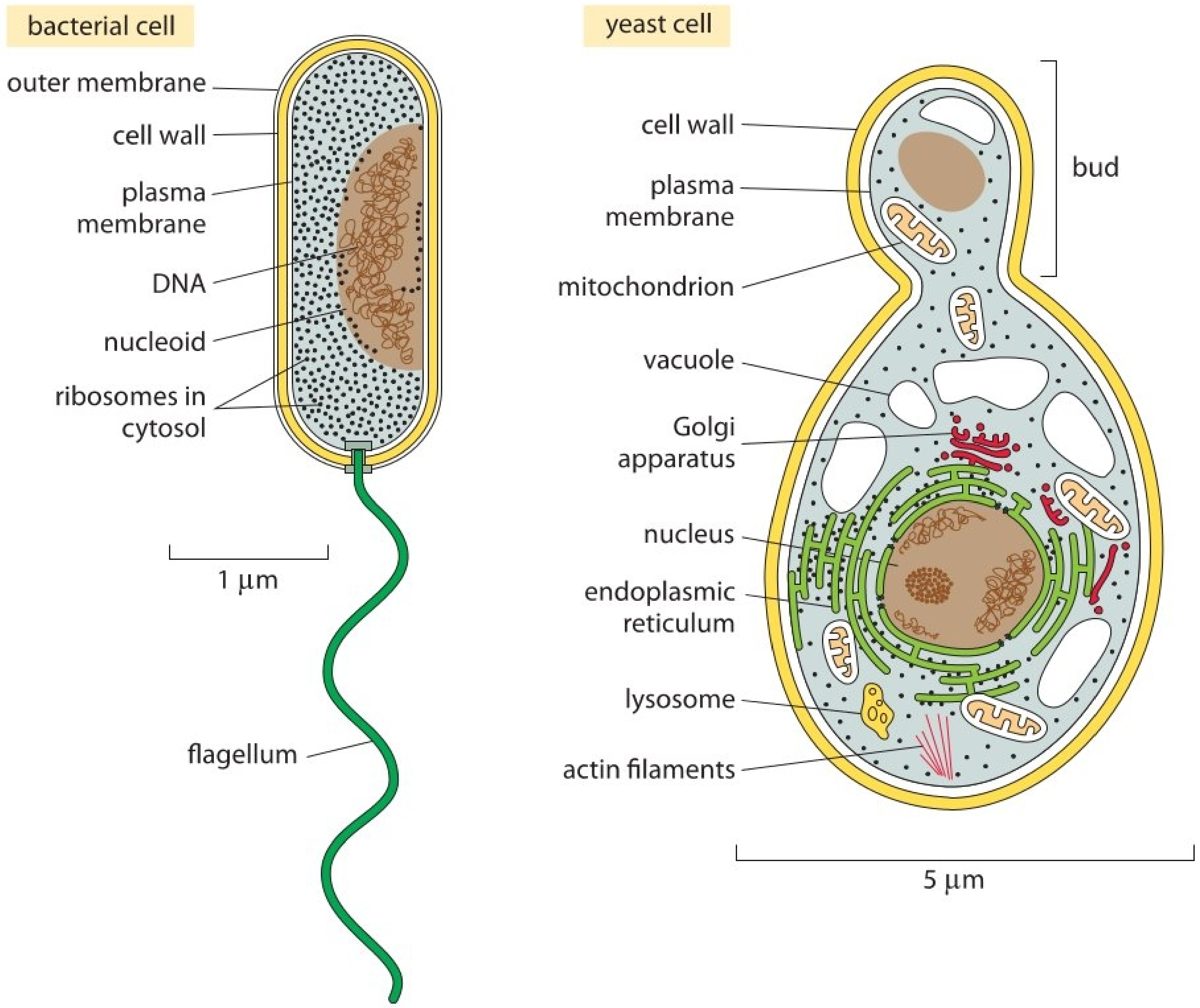


Budding yeast
Saccharomyces cerevisiae



What do you know about cells?

Microbial cells – internal structure



R. Milo and R. Phillips, Cell biology by the numbers

Molecule types and sizes

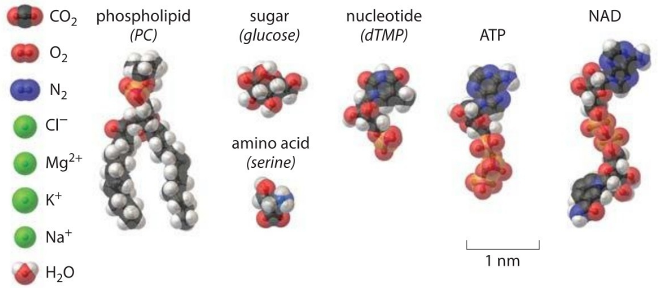
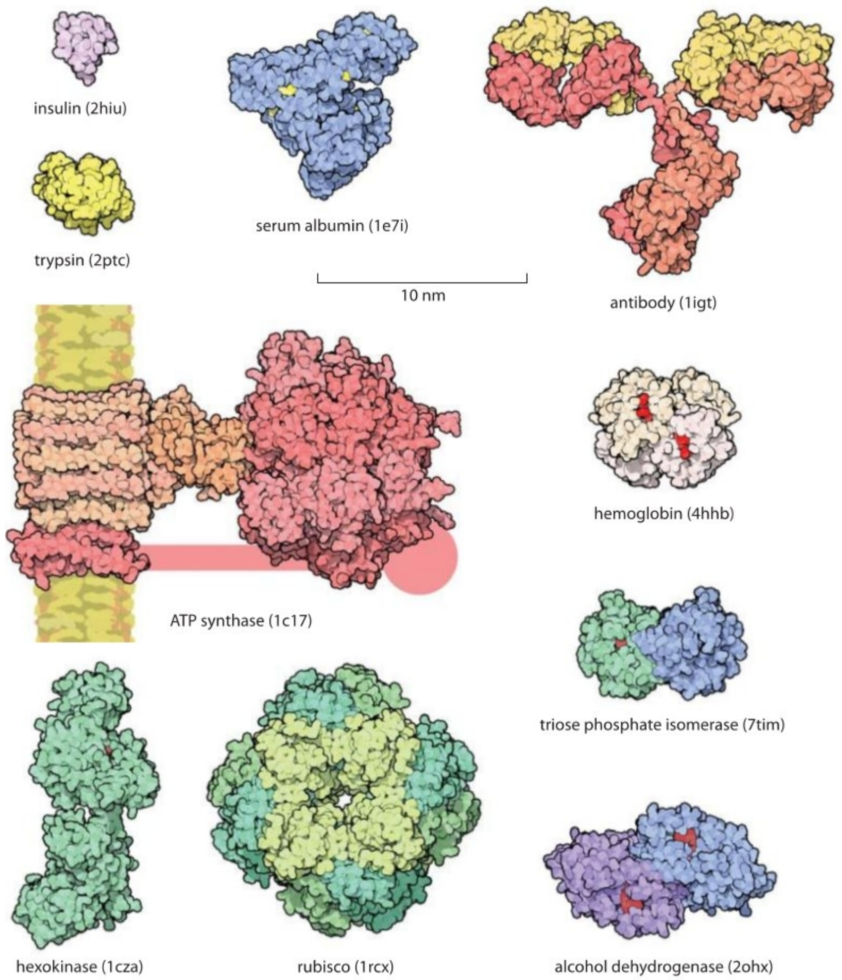


Figure 1: A structural view of some of the basic constituents of a cell.

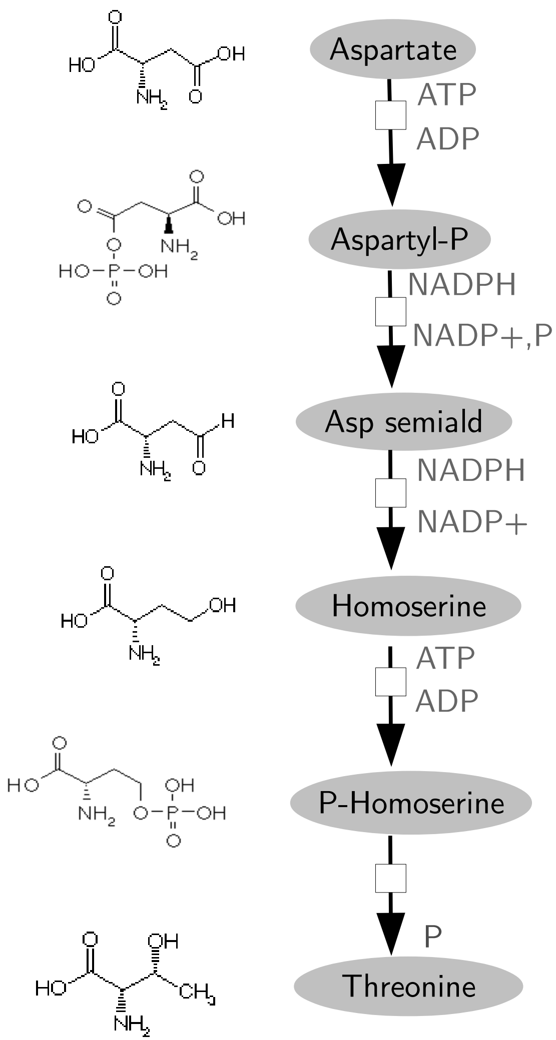


Graphics: R. Milo and R. Phillips, Cell Biology by the Numbers

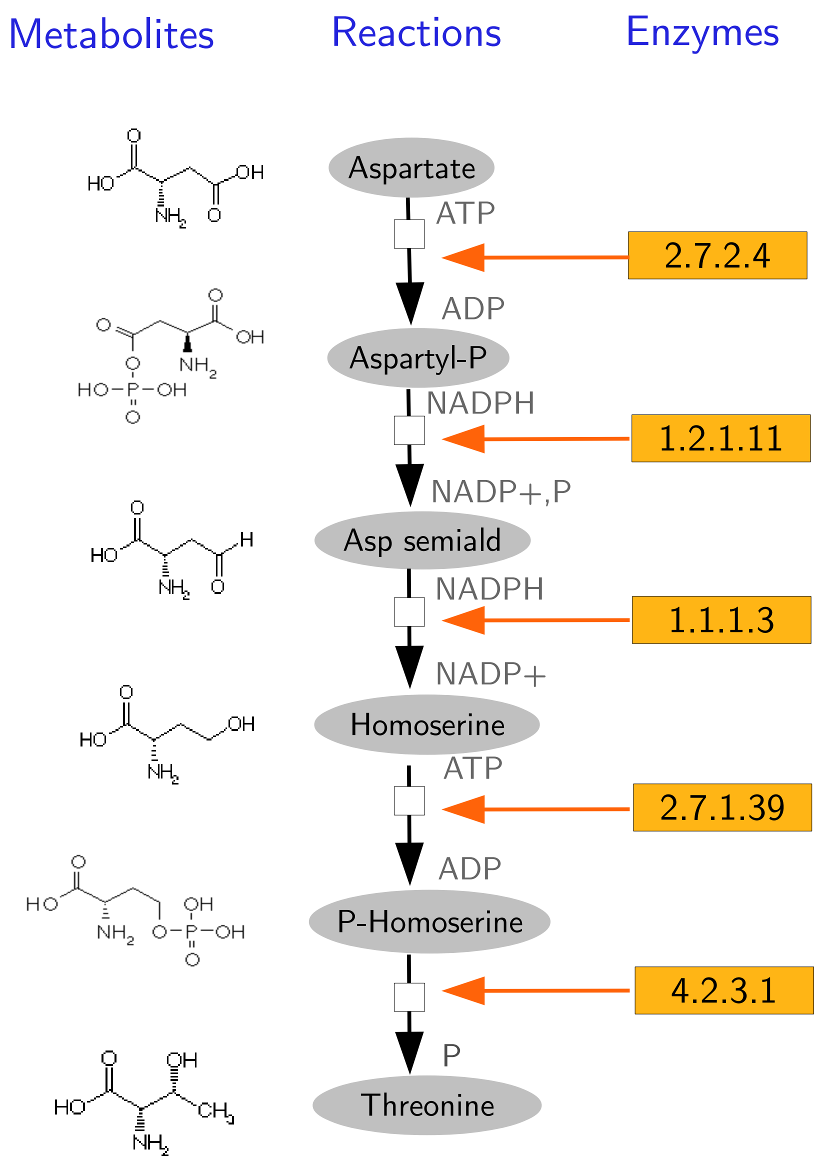
A metabolic pathway and its regulation

Metabolites

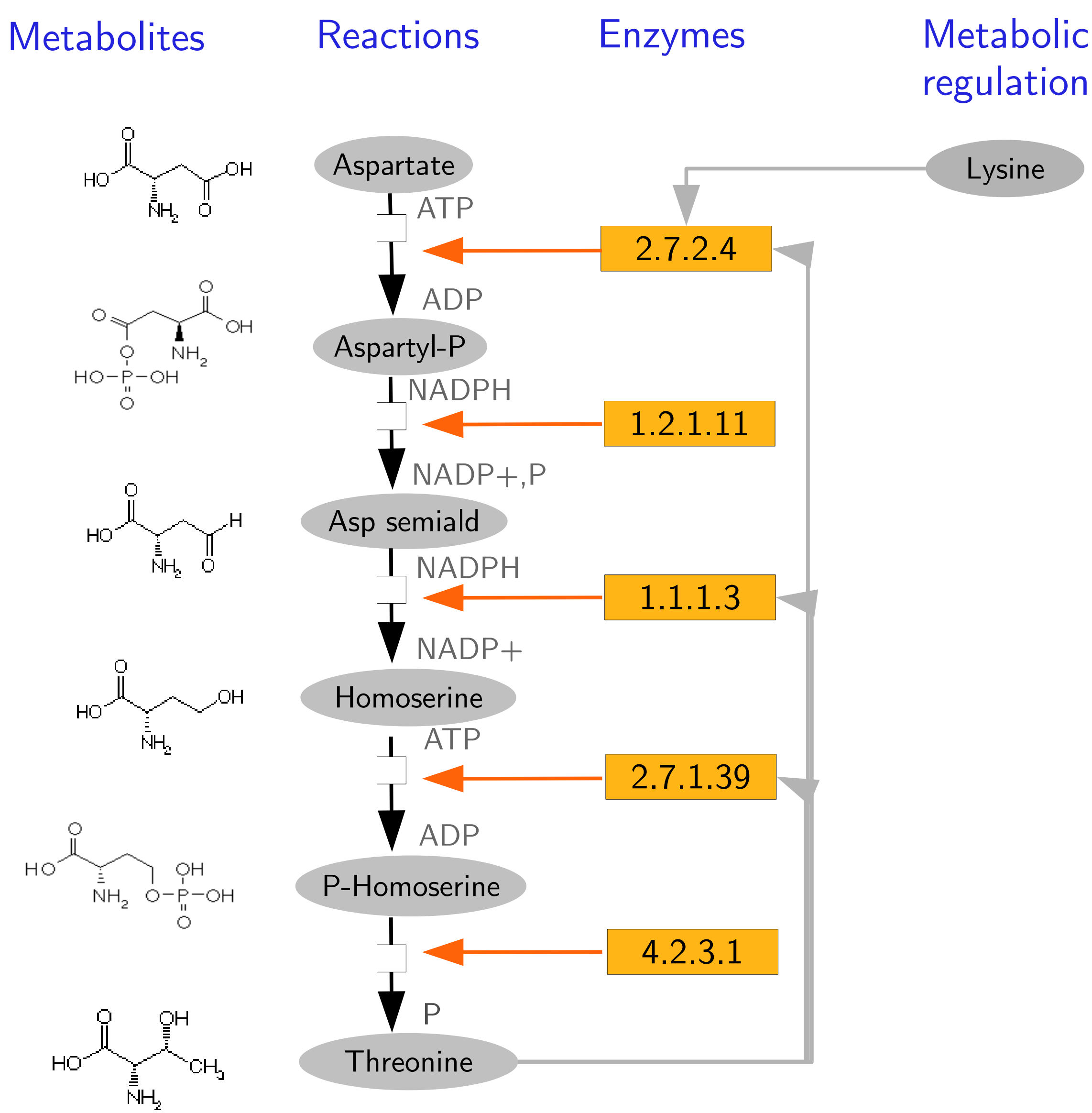
Reactions



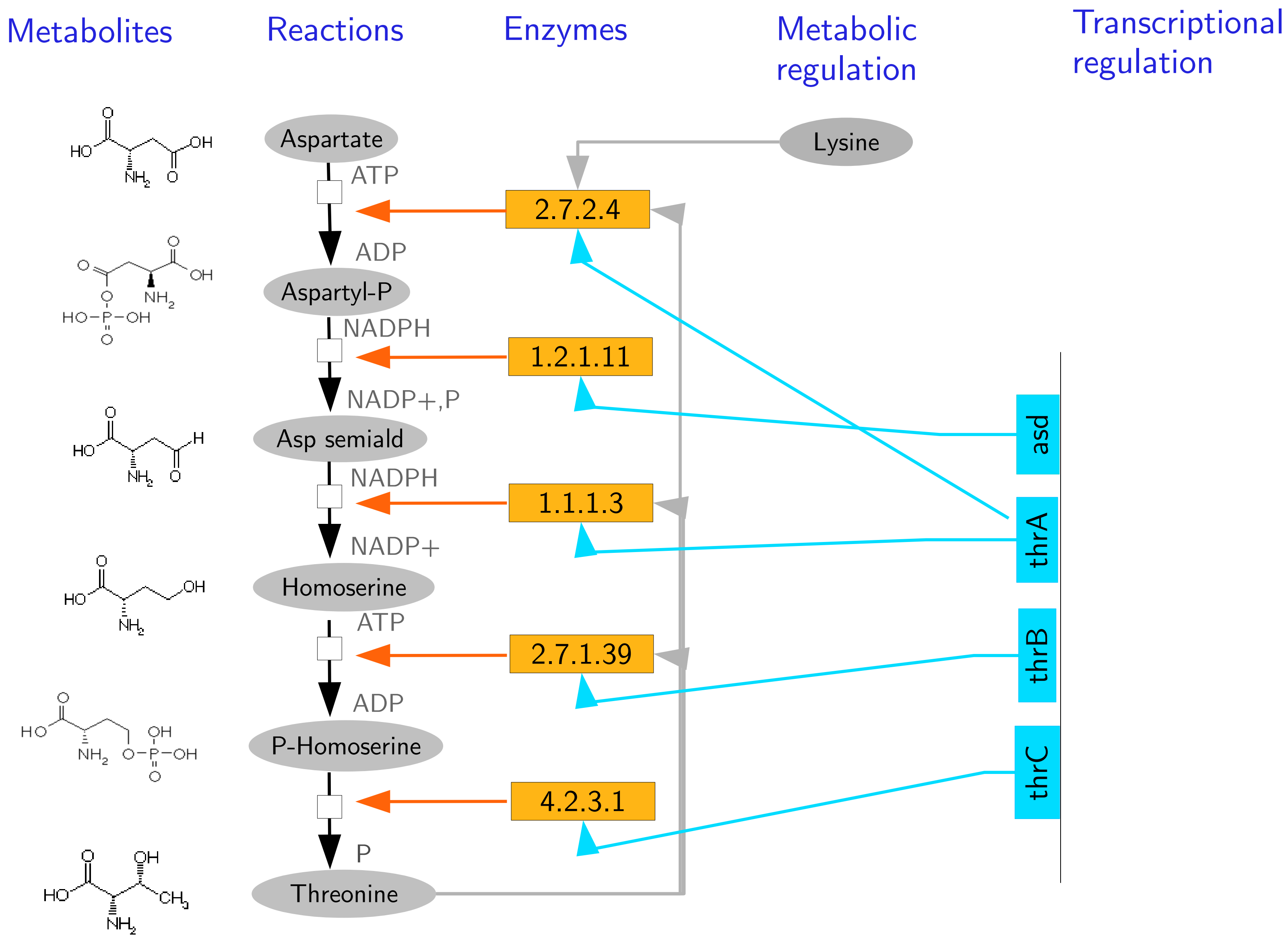
A metabolic pathway and its regulation



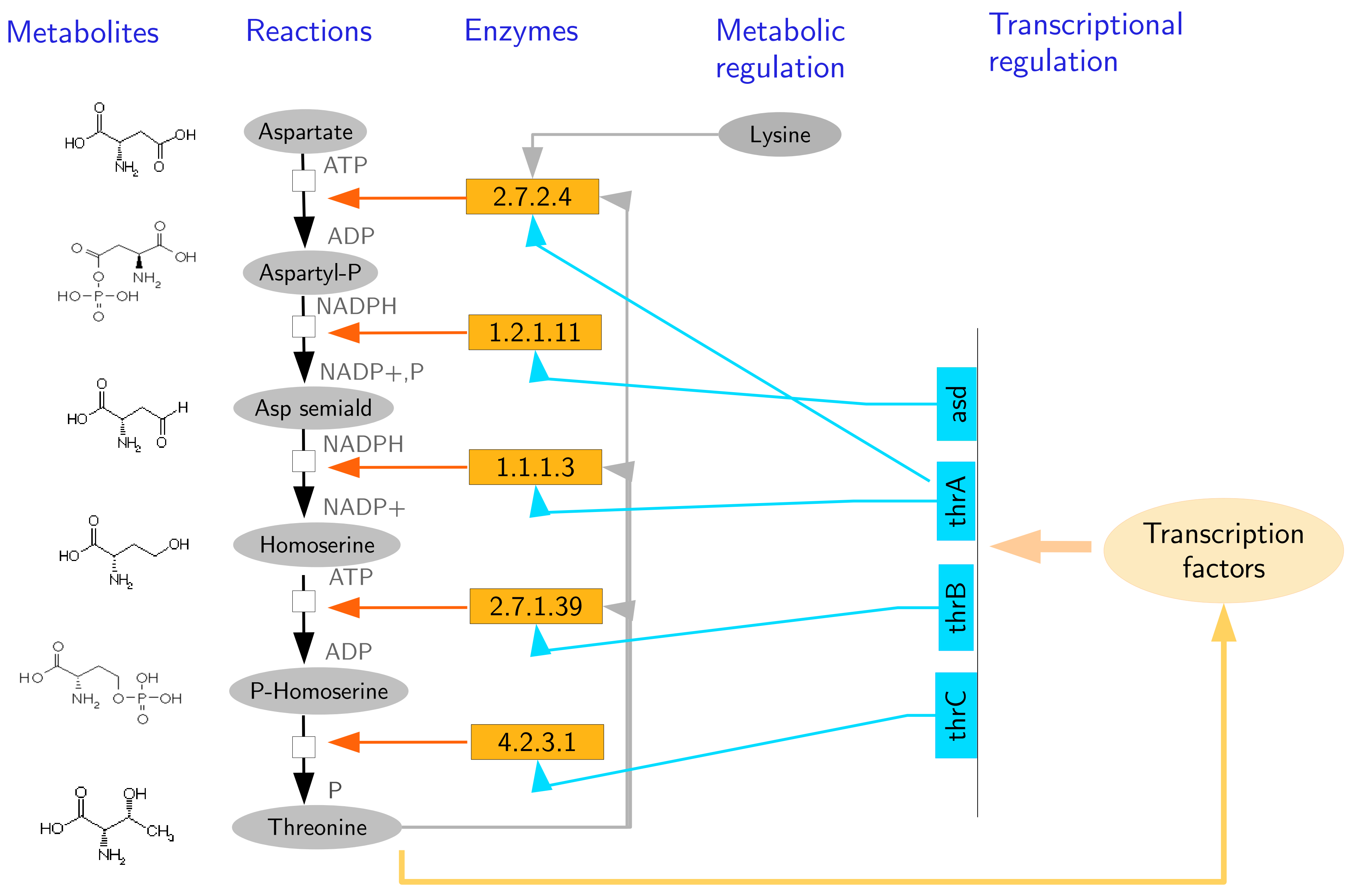
A metabolic pathway and its regulation



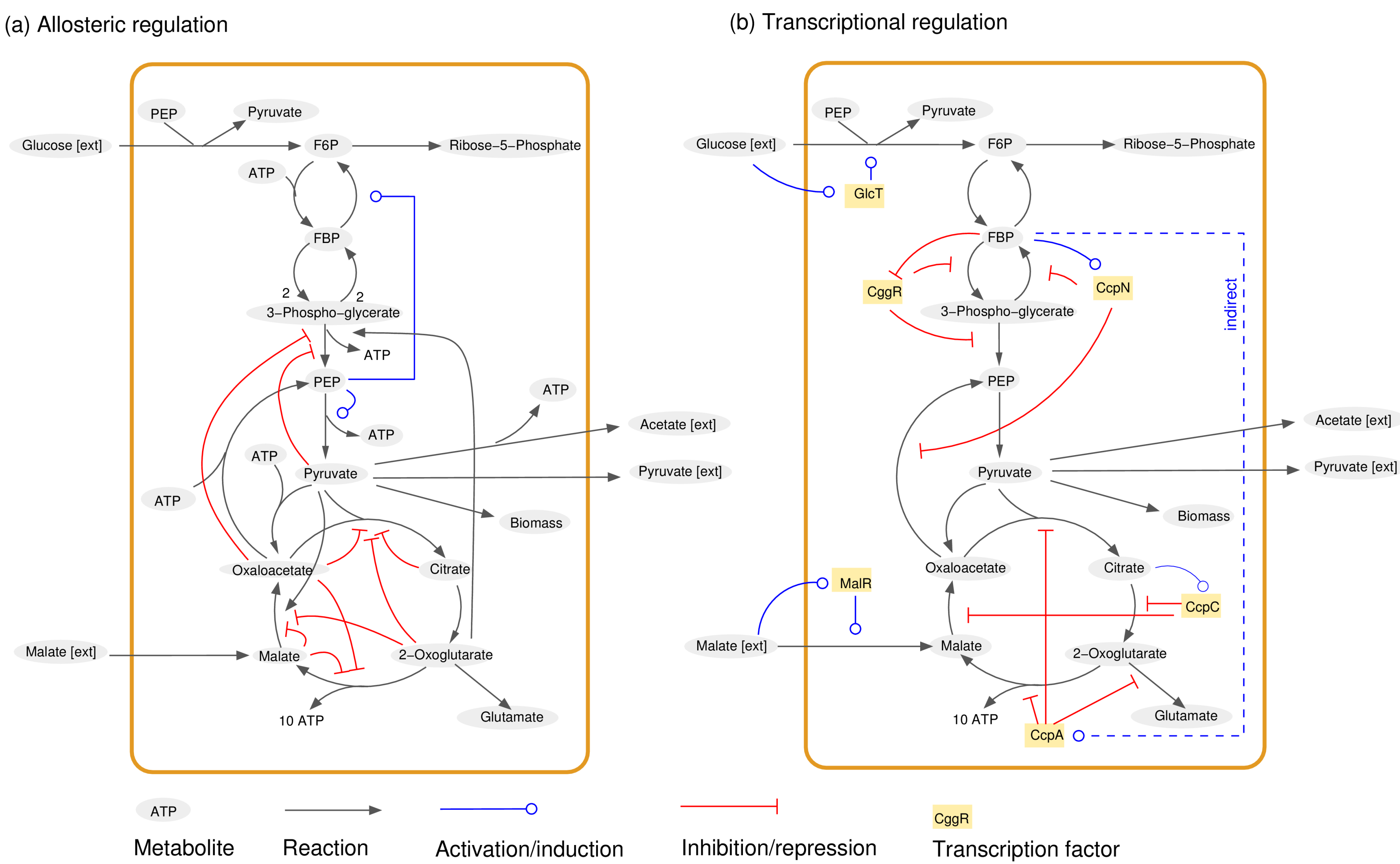
A metabolic pathway and its regulation



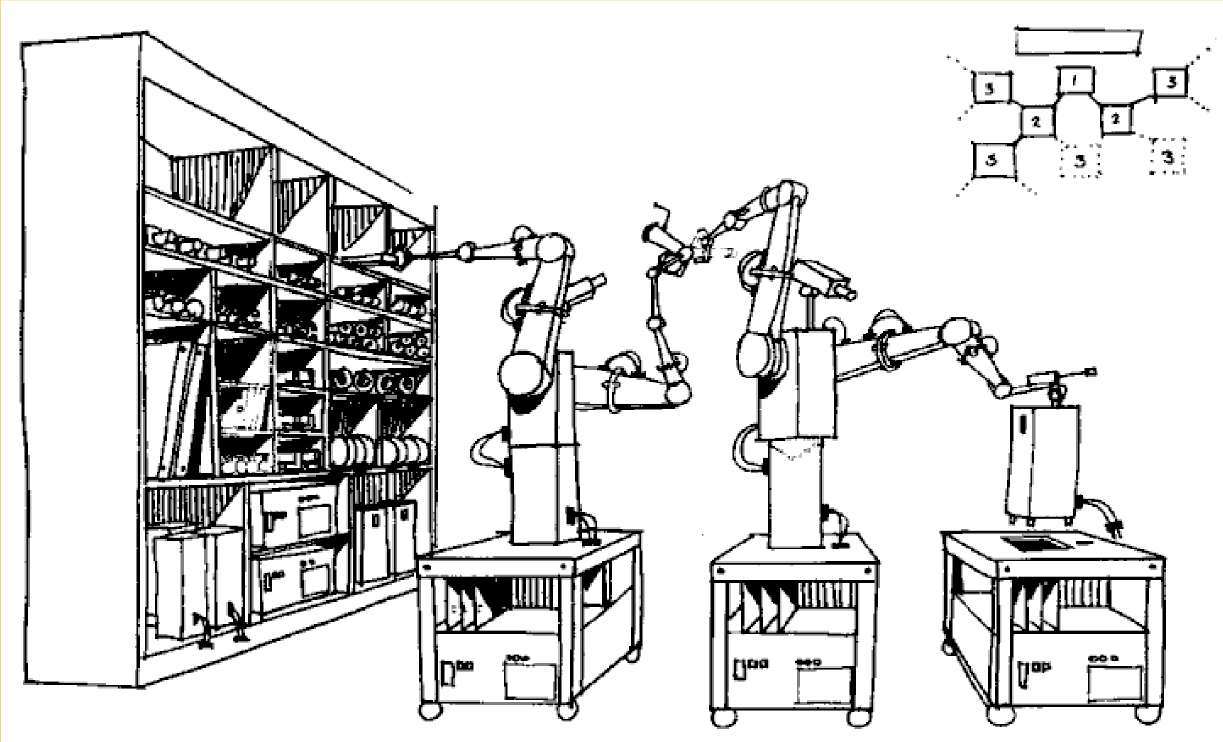
A metabolic pathway and its regulation



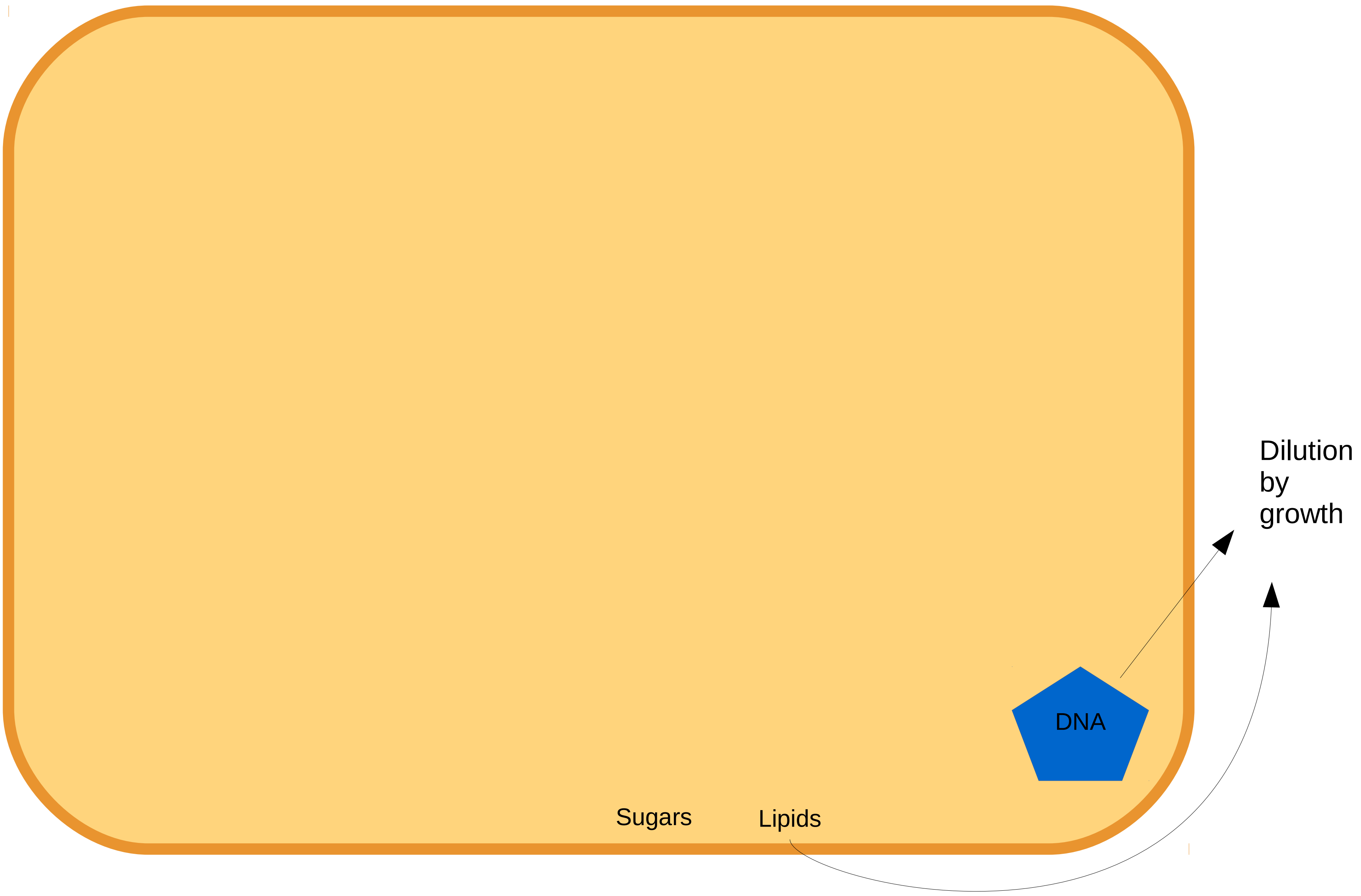
Metabolic networks and their regulation



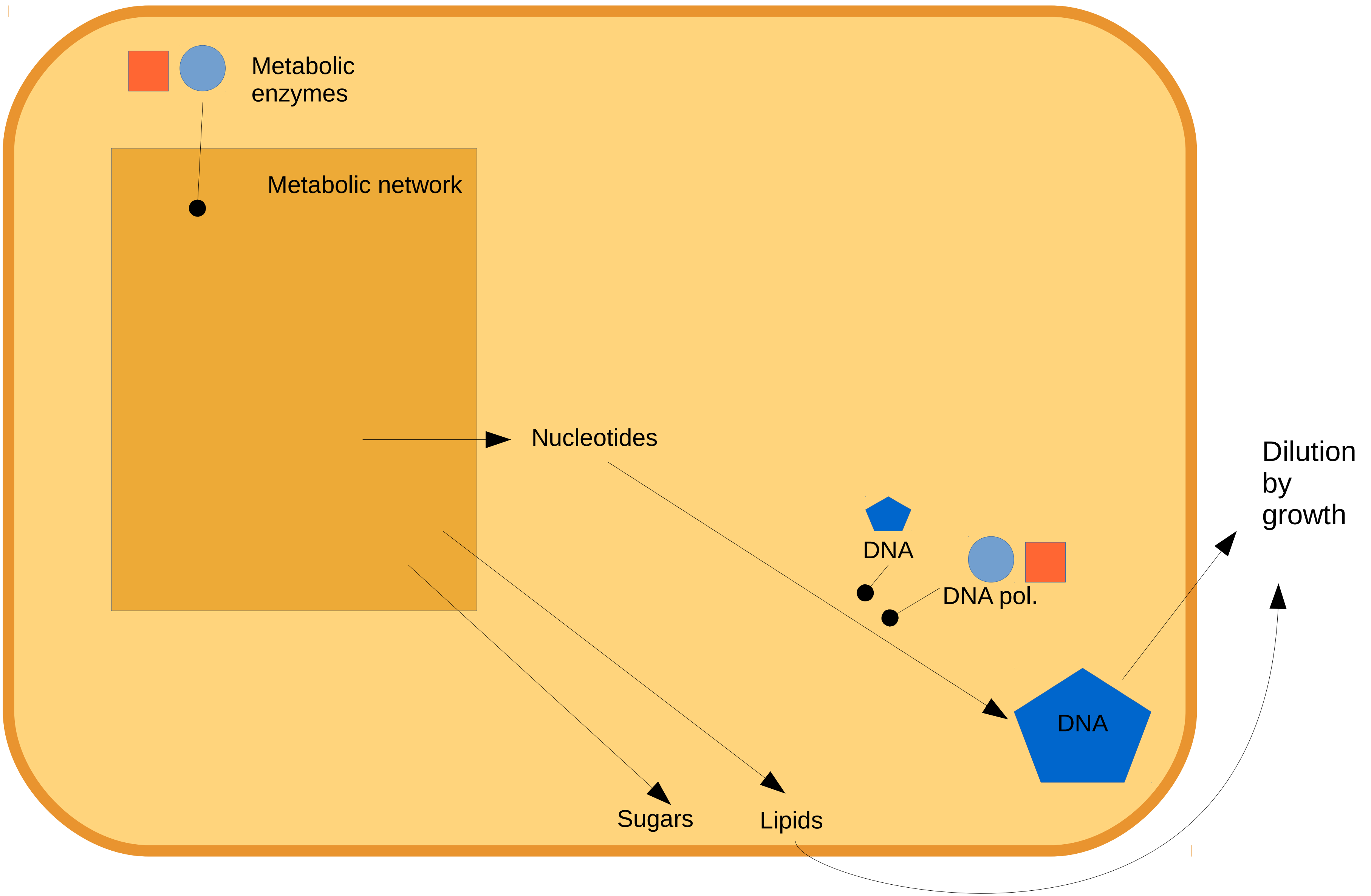
The cell as a self-replicating factory



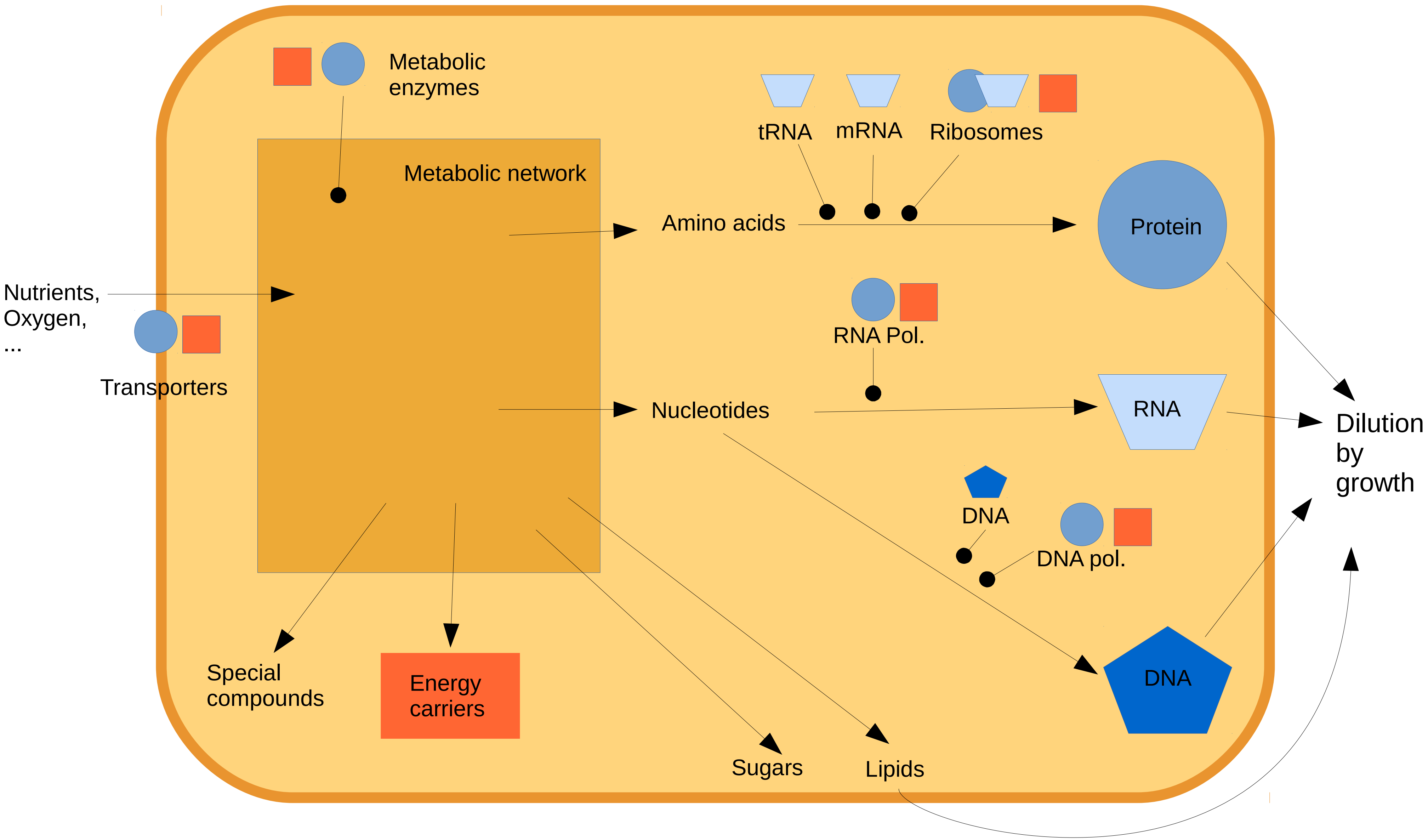
The cell as a self-replicating factory



The cell as a self-replicating factory



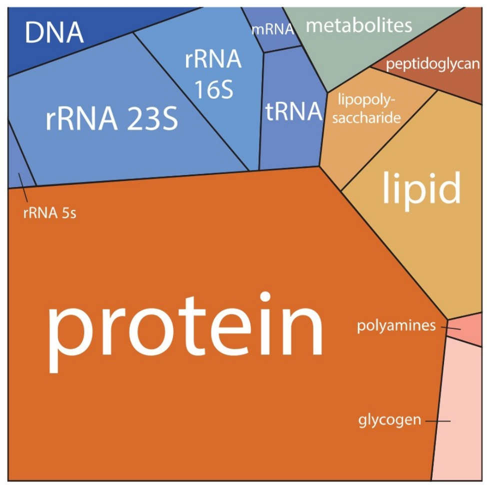
The cell as a self-replicating factory



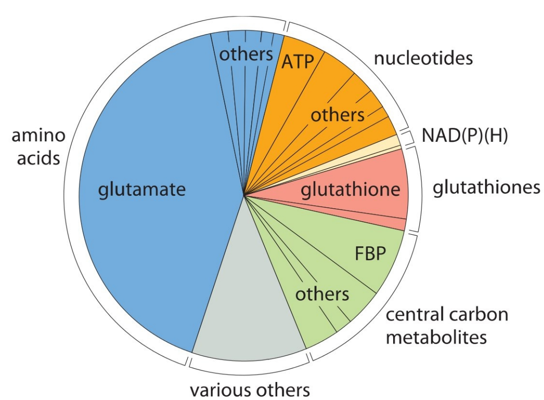
Chemical composition of a cell

Chemical components (mass composition)

- Water
- Protein
- RNA
- DNA
- Lipids
- Small molecules
- ...



Small molecules



sum of concentrations ≈ 200 mM

macromolecule	percentage of total dry weight	weight per cell (fg)	characteristic molecular weight (Da)	number of molecules per cell
protein	55	165	3×10^4	3,000,000
RNA	20	60		
23 S rRNA		32	1×10^6	20,000
16 S rRNA		16	5×10^5	20,000
5 S rRNA		1	4×10^4	20,000
transfer		9	2×10^4	200,000
messenger		2	1×10^6	1,400
DNA	3	9	3×10^9	2
lipid	9	27	800	20,000,000
lipopolysaccharide	3	9	8000	1,000,000
peptidoglycan	3	9	$(1000)_n$	1
glycogen	3	9	1×10^6	4,000
metabolites and cofactors pool	3	9		
inorganic ions	1	3		
total dry weight	100	300		
water (70% of cell)		700		
total cell weight		1000		

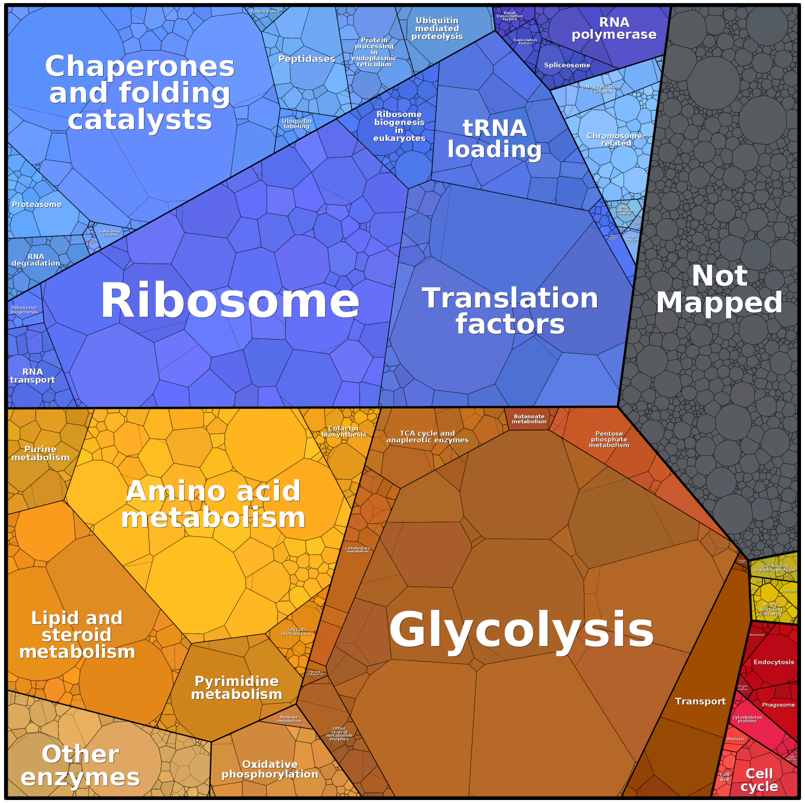
composition rules of thumb

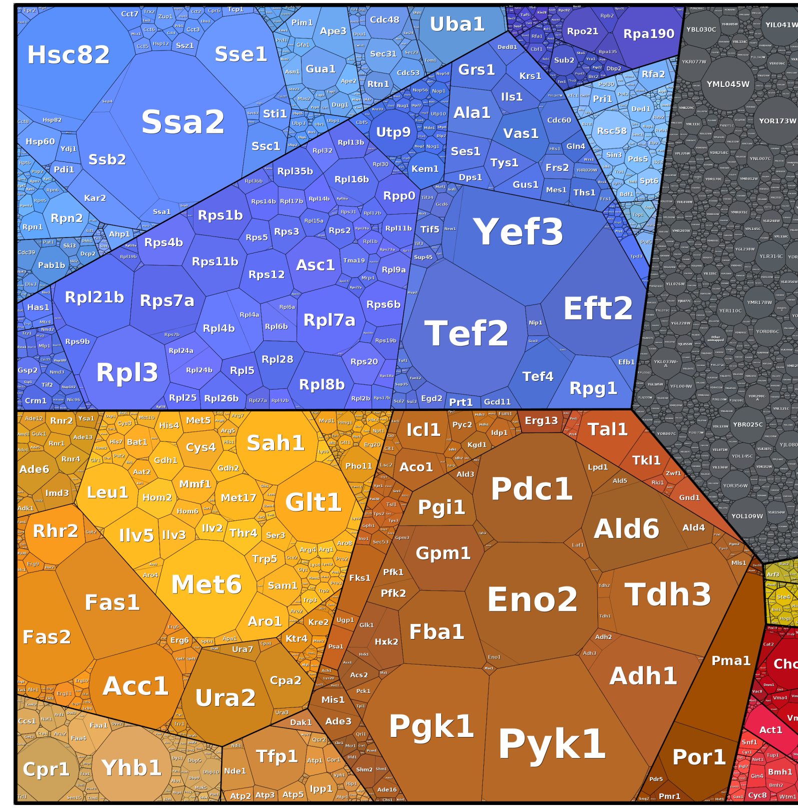
- carbon atoms ~10^{10}
- 1 molecule per cell gives ~1 nM conc.
- ATP required to build and maintain cell over a cell cycle ~10^{10}
- glucose molecules needed per cell cycle ~3×10^9 (2/3 of carbons used for biomass and 1/3 used for ATP)

Molecule numbers in bacterial cells

R. Milo and R. Phillips, Cell biology by the numbers

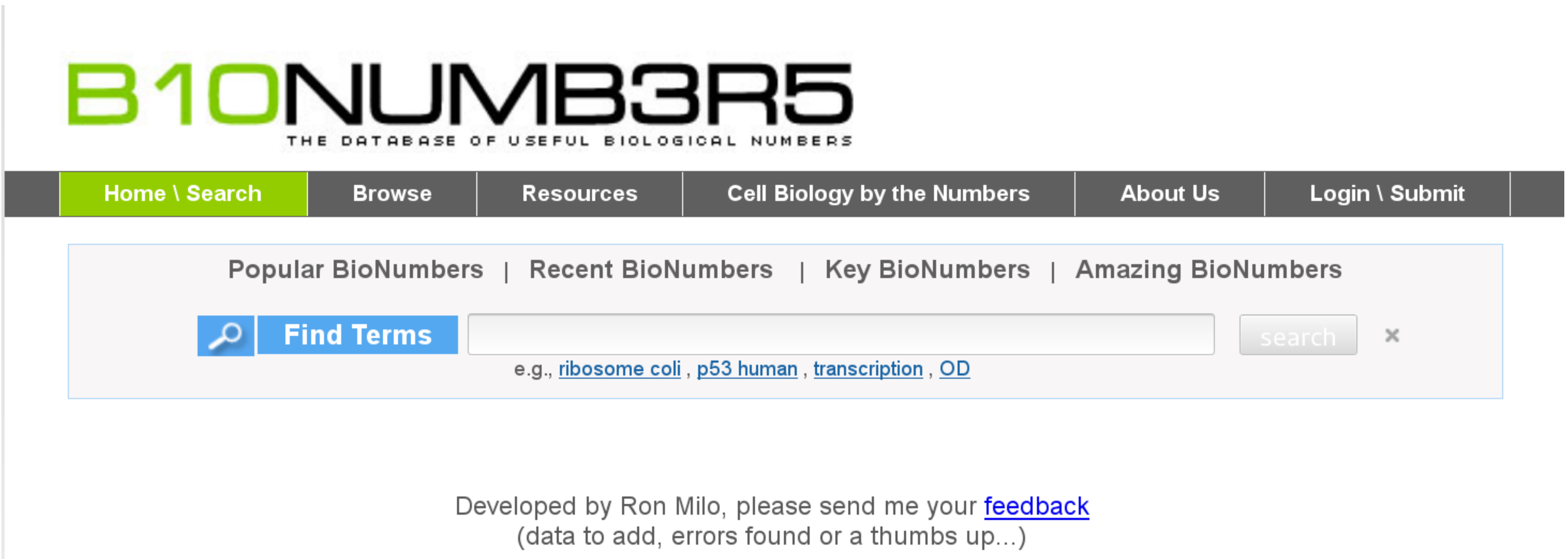
Protein “investment” in functional subsystems





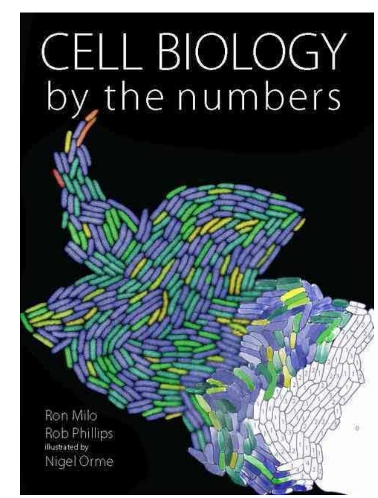
- “Central dogma”: Production of DNA, RNA, and protein
- Metabolism
- Membranes and transport
- Stress reponse, repair
- … and many others

Bionumbers website: relevant numbers for cell biology



http://bionumbers.hms.harvard.edu

Book by R. Milo and R. Phillips,
Cell biology by the numbers



Guess some (typical) numbers

- What is the volume of a cell? ..
- What is the size of a protein?
- How many protein molecules exist in a cell?
- What is the number of genes?
- How long does it take to transcribe a gene?
- How long does it take to produce a protein molecule?
- What is the minimal doubling time of a cell?
- What other questions come to your mind?

Precise values do not matter here – think about orders of magnitude
Work in pairs and check the results at http://bionumbers.hms.harvard.edu

Cells – some typical numbers

property	*E. coli*	budding yeast	mammalian (HeLa line)
cell volume	0.3–3 μm^3	30–100 μm^3	1,000–10,000 μm^3
proteins per μm^3 cell volume		2–4×10^6	
mRNA per cell	10^3-10^4	10^4–10^5	10^5–10^6
proteins per cell	~10^6	~10^8	~10^{10}
mean diameter of protein		4–5 nm	
genome size	4.6 Mbp	12 Mbp	3.2 Gbp
number protein coding genes	4300	6600	21,000
regulator binding site length		10–20 bp	
promoter length	~100 bp	~1000 bp	~10^4–10^5 bp
gene length	~1000 bp	~1000 bp	~10^4–10^6 bp (with introns)
concentration of one protein per cell	~1 nM	~10 pM	~0.1–1 pM
diffusion time of protein across cell (D ≈ 10 μm^2/s)	~0.01 s	~0.2 s	~1–10 s
diffusion time of small molecule across cell (D ≈ 100 μm^2/s)	~0.001 s	~0.03 s	~0.1–1 s
time to transcribe a gene	<1 min (80 nts/s)	~1 min	~30 min (incl. mRNA processing)
time to translate a protein	<1 min (20 aa/s)	~1 min	~30 min (incl. mRNA export)
typical mRNA lifetime	2–5 min	~10 min to over 1 h	5-100 min to over 10 h
typical protein lifetime	1 h	0.3–3 h	10–100 h
minimal doubling time	20 min	1 h	20 h
ribosomes/cell	~10^4	~10^5	~10^6
transitions between protein states (active/inactive)		1–100 μs	
timescale for equilibrium binding of small molecule to protein (diffusion limited)		1–1000 ms (1 μM–1 nM affinity)	
timescale of transcription factor binding to DNA site		~1 s	
mutation rate		10^{-8}–10^{-10}/bp/replication	

Graphics: R. Milo and R. Phillips, Cell Biology by the Numbers