

Metabolic Networks : structure

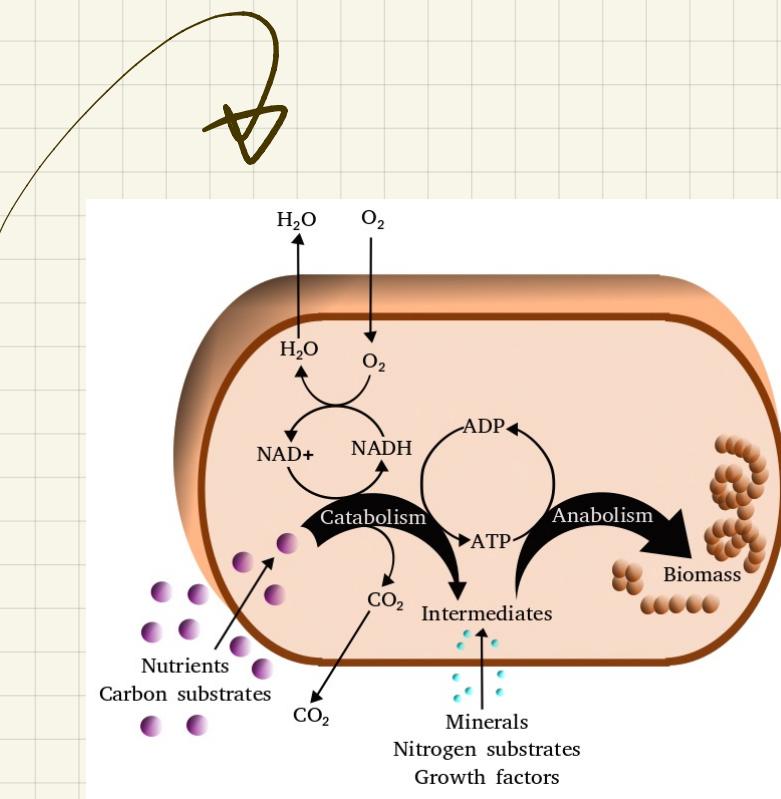
What is metabolism?

- 1) (input) nutrients, resources
- 2) (conversion) molecular transformation into other molecules
- 3) (output) discard waste molecules
new useful stuff

2 kinds of metabolic reactions

① catabolic / catabolism : breaking things down \rightarrow produces "energy"
(currency metabolite)

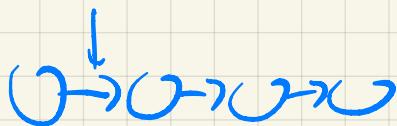
② anabolic / anabolism : building things up \rightarrow requires input of \oplus



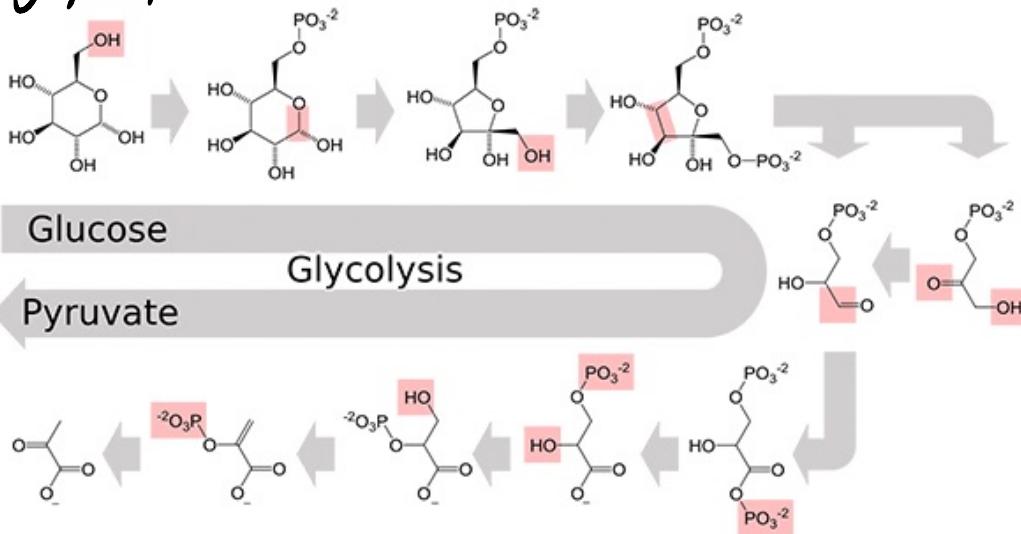
Enzymes and pathways

1. metabolism is not just a collection of passive chemical reactions

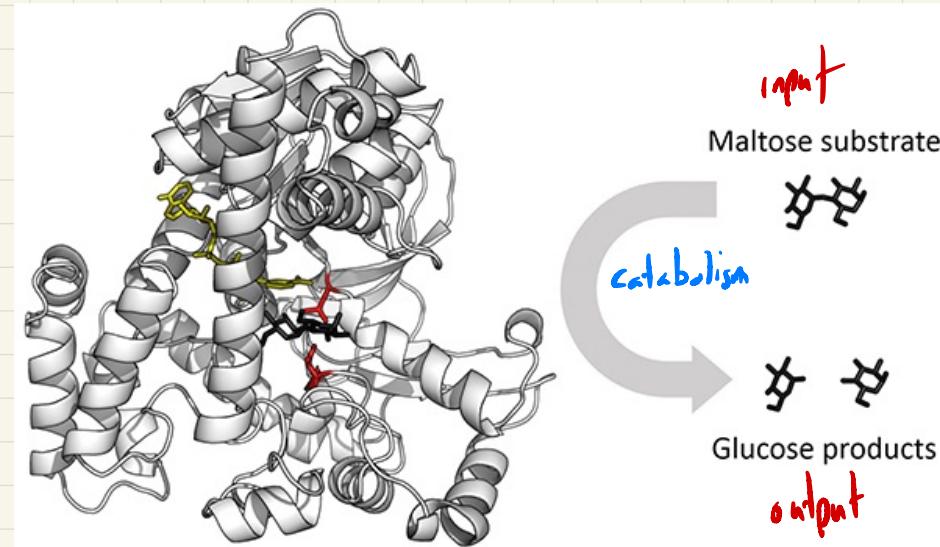
2. cells actively manages ("regulates") metabolism by controlling associated enzymes (proteins)



glycolysis "pathway" \Rightarrow 11 different enzymes



1 such enzyme



glucosidase enzyme \Rightarrow Breaks 1 bond

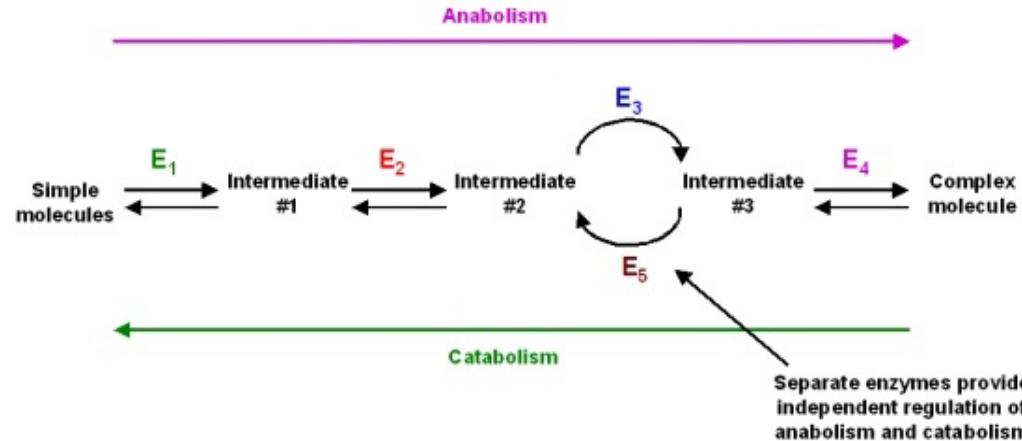
Pathways are overlapping

Metabolic network = \bigcup_i pathway_i $\rightsquigarrow \left\{ \begin{array}{l} \text{all metabolic enzymes} \\ \text{various metabolites} \end{array} \right.$ $\frac{1}{2}$

- Structure of metabolic network
- changes possibilities for regulation (behavior)
 - prohibits / enables some functions

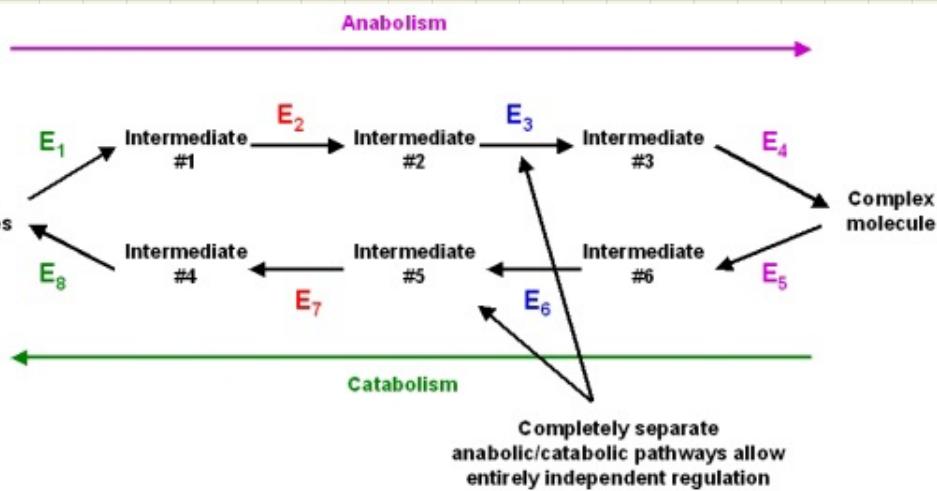
Simple

5 enzymes



Complicated

8 enzymes



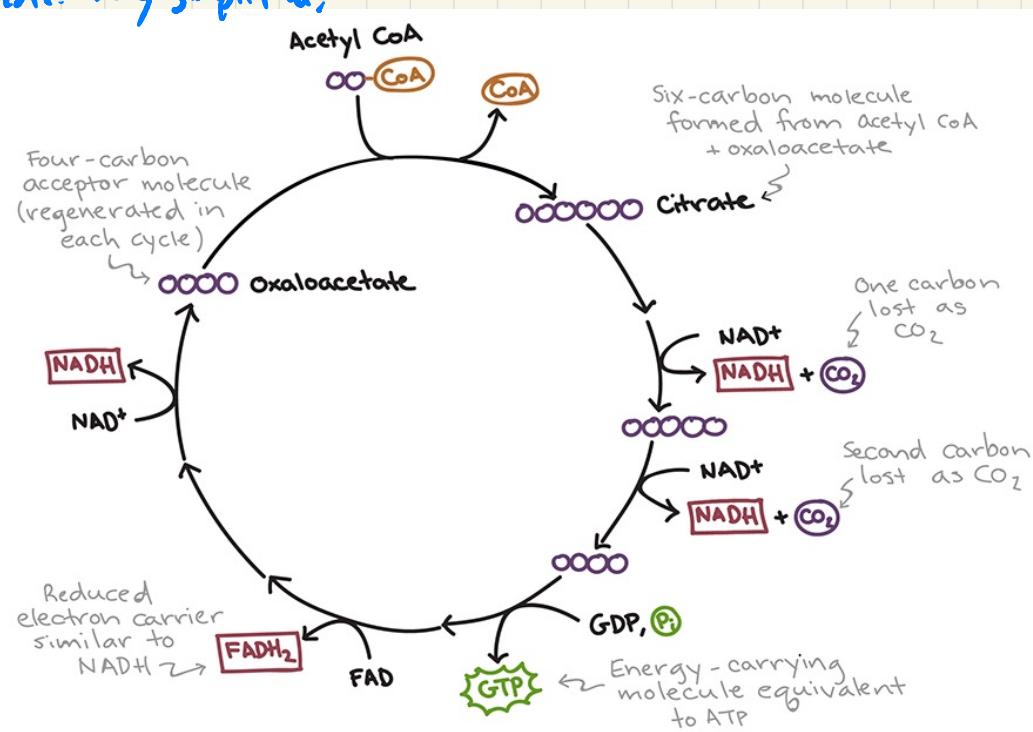
Separate enzymes provide independent regulation of anabolism and catabolism

Completely separate anabolic/catabolic pathways allow entirely independent regulation

Core Metabolism

aerobic organisms \rightarrow have a "core metabolism" \rightarrow Krebs cycle (a cell's "engine")

note: very simplified!

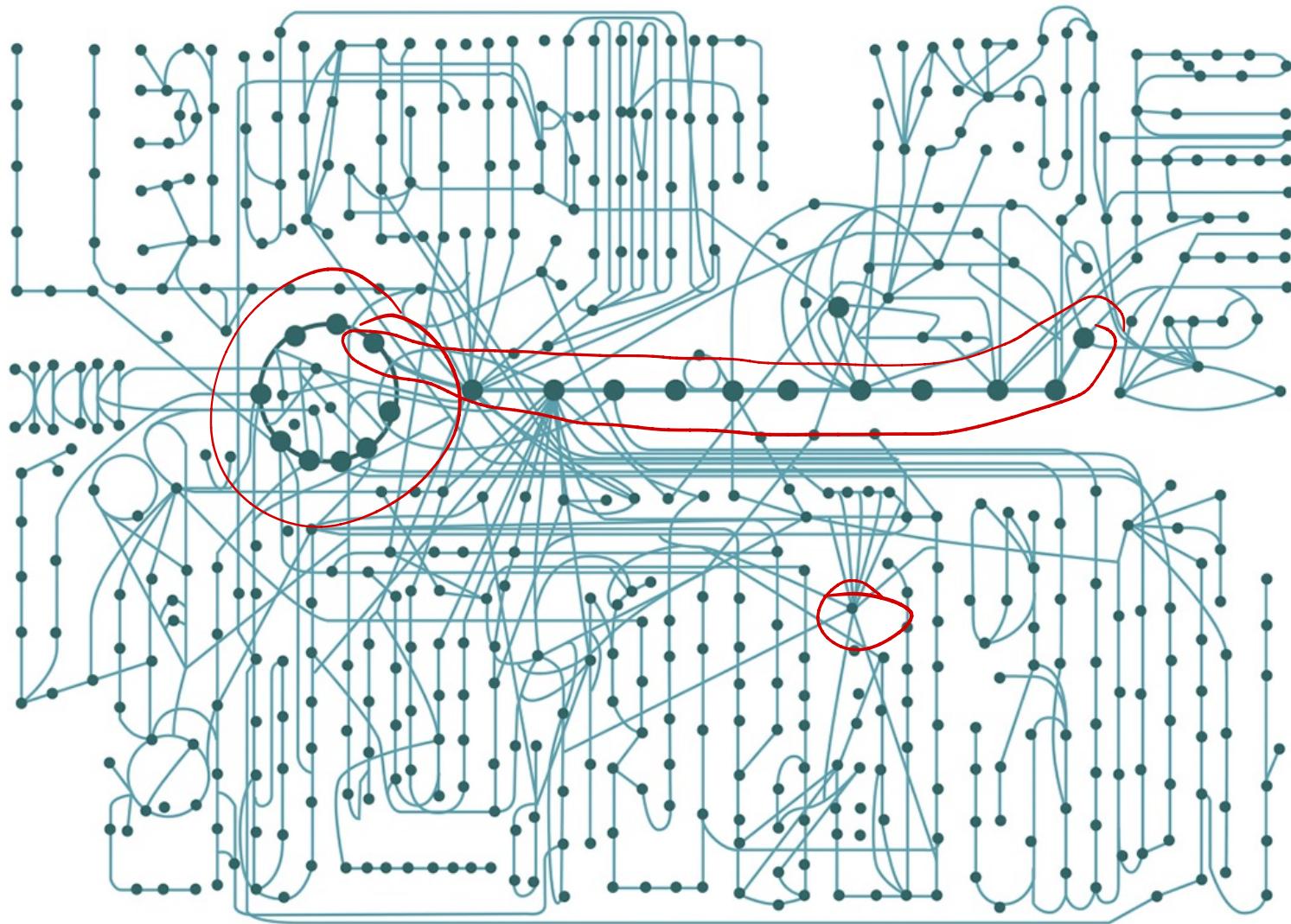


aka: citric acid cycle

- eukaryotes: mitochondria
- prokaryotes: cytosol + cell membrane

but... it's not that simple

less simplified...

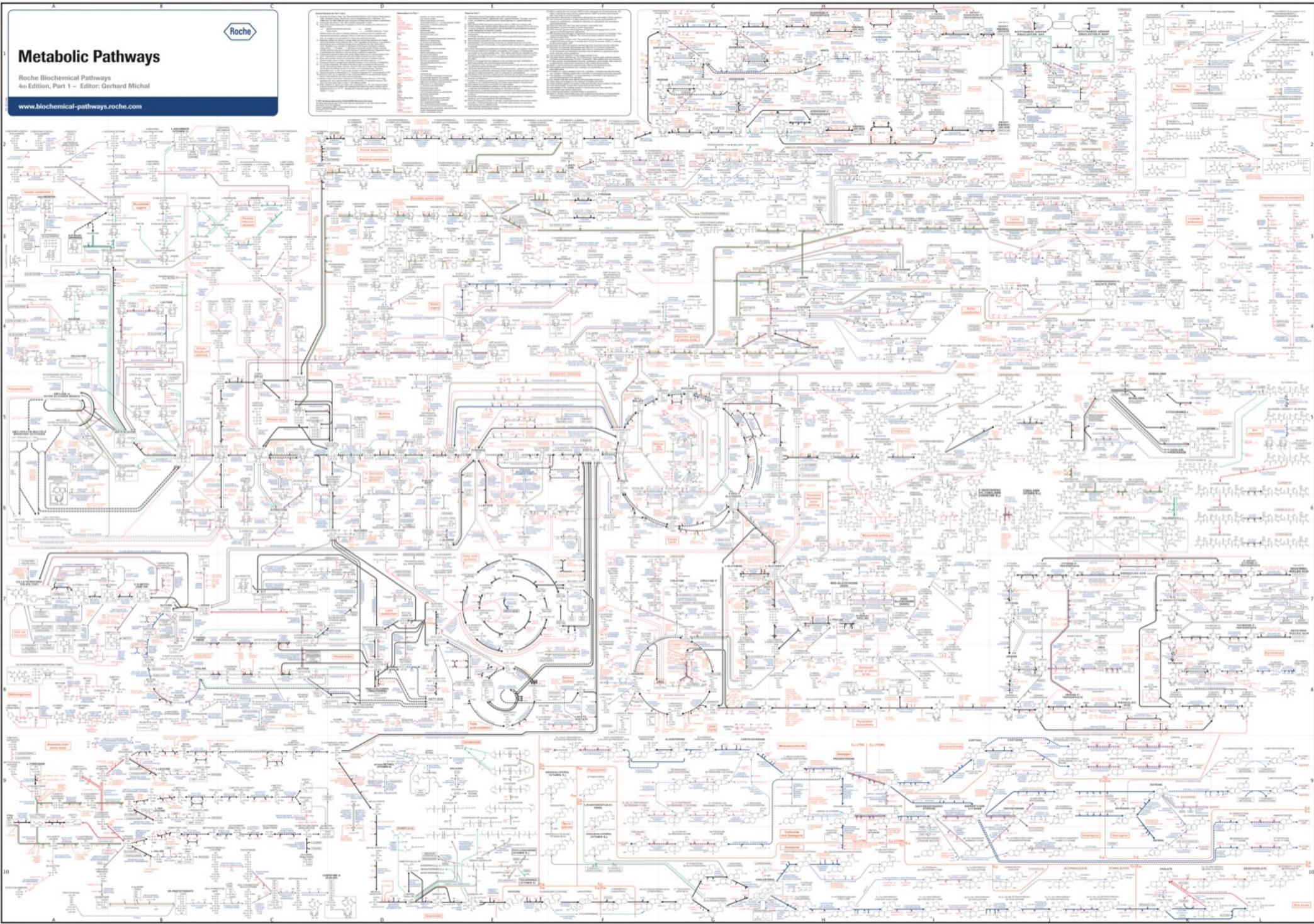


Metabolic Pathways



Roche Biochemical Pathways
4th Edition, Part 1 – Editor: Gerhard Michal

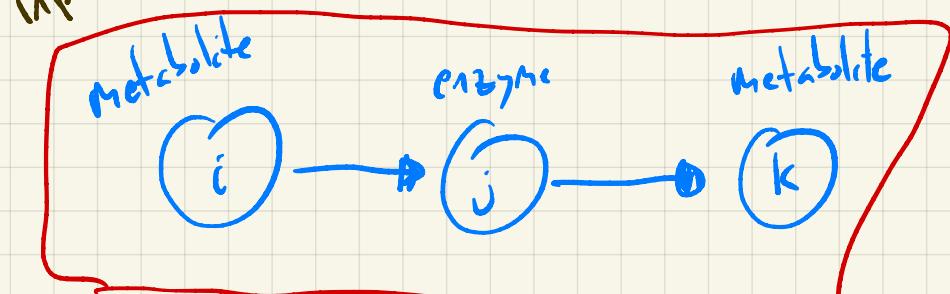
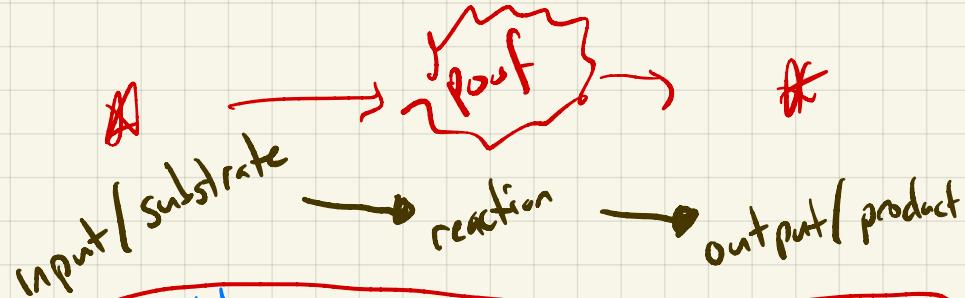
www.biochemical-pathways.roche.com



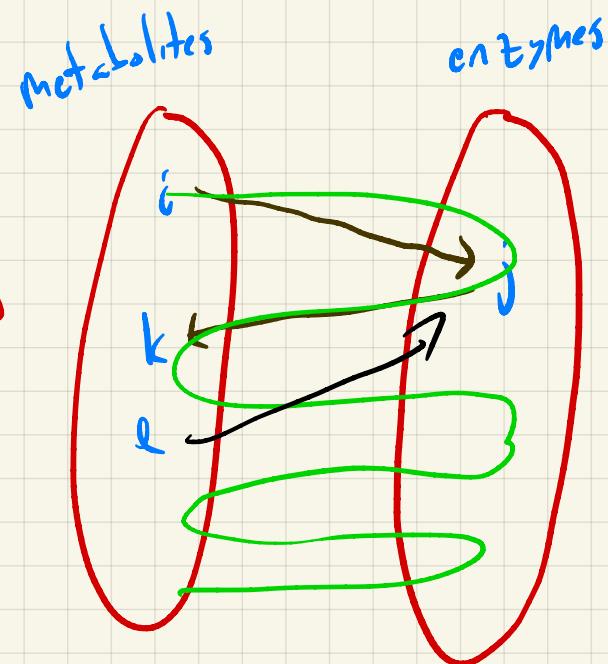
What are the nodes, what are the edges?

↳ multiple representations → most related to a directed bipartite network

→ nodes: { metabolites, enzymes }



$(i \rightarrow j)$ $(j \rightarrow k)$



Multiple ways to derive simpler metabolic networks

↳ "project" onto either reactions or metabolites

two useful ones

1) substrate-product network

* every substrate links to every reaction product

- complete bipartite network b/w inputs + outputs

2) substance network

* edges connect all substances (in or out) in a reaction

- complete network among all inputs + outputs

* no matter what simplification, some information is lost

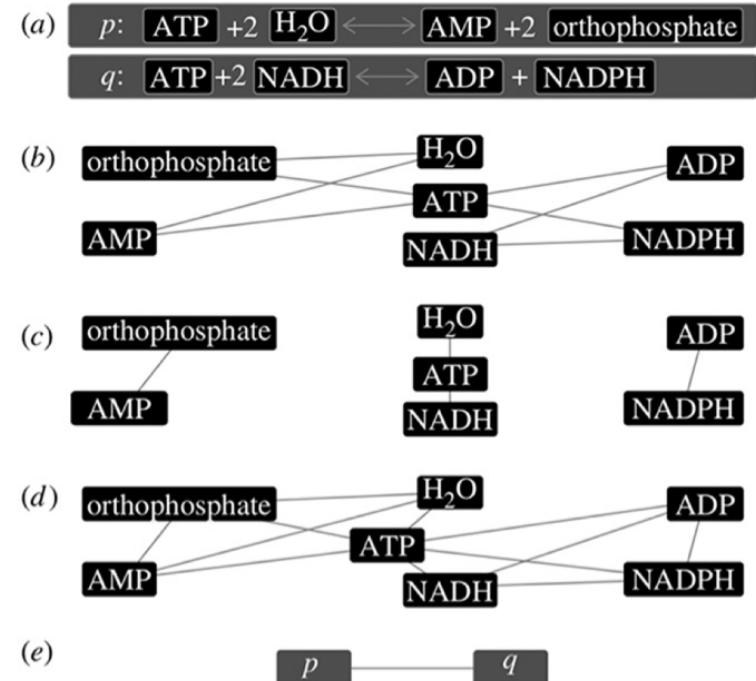
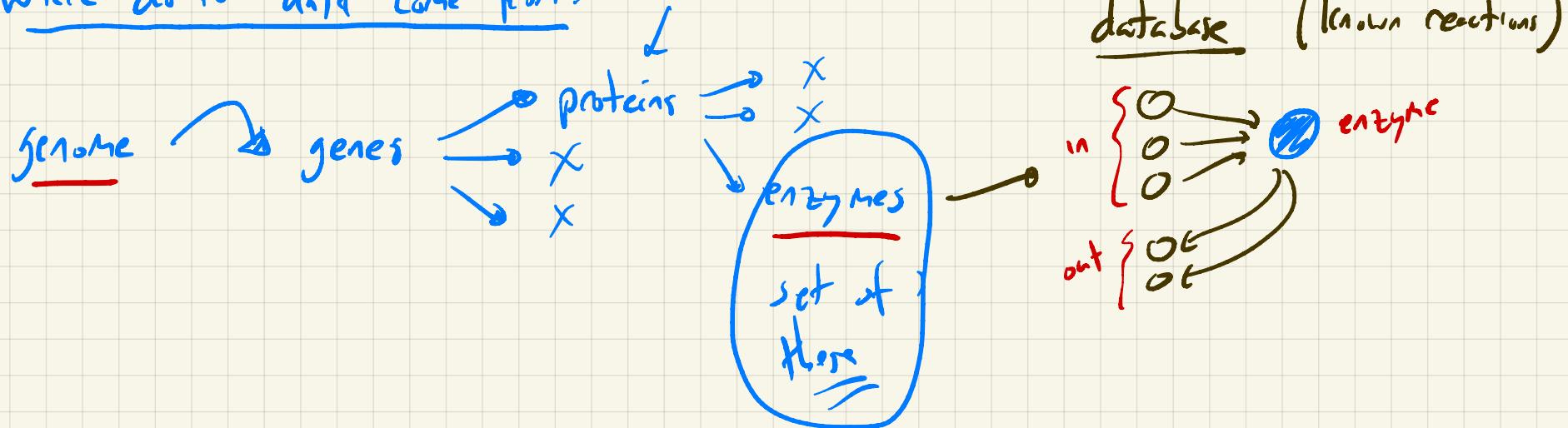


Figure 1. An illustration of different network representations derived from the two hypothetical reactions shown in (a). (b) Substrate-product network, (c) a substrate-substrate network, (d) a substance network (including both the substrate-product and substrate-substrate type edges) and (e) a reaction network where the vertices are reactions connected if they have a substance in common.

Where do the data come from?



Caveats *1) if it's not in the database, we cannot put it into network

*2) most of what's known, is know from model organisms

{ E. coli
D. melanogaster
C. elegans

*3) thus, novel reactions are likely invisible

→ False Pos: low

False Neg: high, but depends on db coverage

(evolutionary distant
to model orgs)

Metadata (sometimes)

- what pathway is rxn in?
- Where in cell? (mitochondria, cytoplasm, cell membrane, etc.)
- what cellular function? (amino-acid metabolism, carbohydrate, lipid, co-factors, etc.)
- What kind of energetics are required?

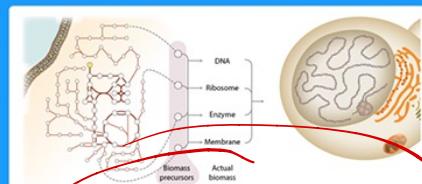
BiGG Models

Search the database by model, reaction, metabolite, or gene 

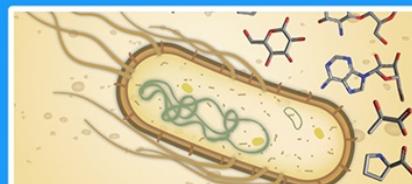
Exclude multistrain models from search

Search

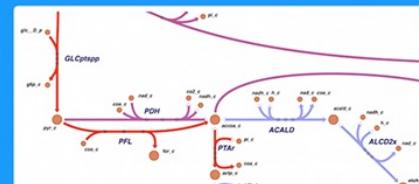
Latest update Version 1.6: Adds 23 new models & more!



View Models



View Metabolites



View Reactions

BiGG Models

[Home](#) [Advanced Search](#) [Data Access](#) [Memote Validator !\[\]\(b52923ac887f6b630066a7f81d758df3_img.jpg\)](#)

Search Database

Search

Search Results

Exclude multistrain models from search

Models

BiGG ID	Organism	Metabolites	Reactions	Genes
e_coli_core	Escherichia coli str. K-12 substr. MG1655	72	95	137
iAB_RBC_283	Homo sapiens	342	469	346
iAF1260	Escherichia coli str. K-12 substr. MG1655	1668	2382	1261
iAF1260b	Escherichia coli str. K-12 substr. MG1655	1668	2388	1261
iAF692	Methanosaerica barkeri str. Fusaro	628	690	692
iAF987	Geobacter metallireducens GS-15	1109	1285	987
iAM_Pb448	Plasmodium berhei	903	1067	448



Model: e_coli_core

Organism:

Escherichia coli str. K-12 substr. MG1655

Genome:

NC_000913.3

Model metrics:

Component	Count
Metabolites	72
Reactions	95
Genes	137

Download COBRA model from the BiGG Database:

SBML ? : e_coli_core.xml.gz (compressed)

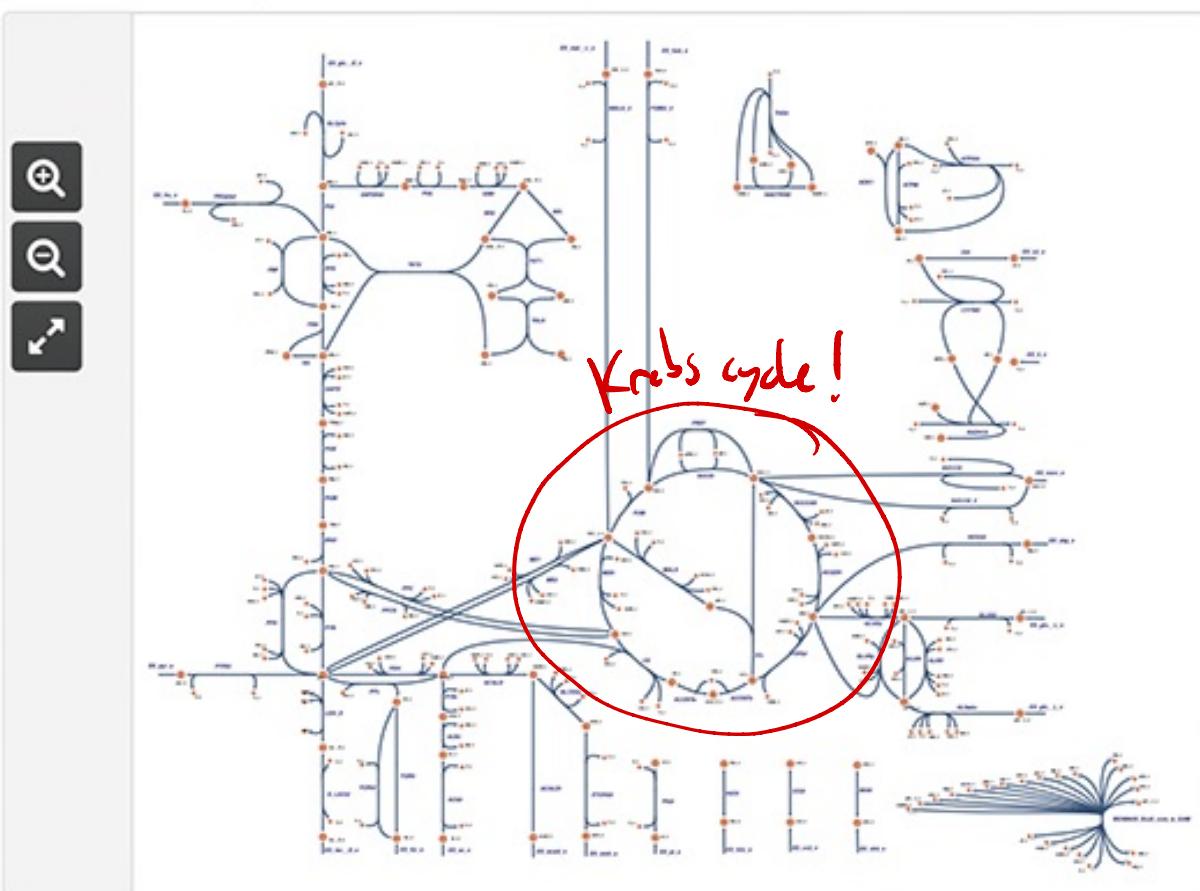
JSON ? : e_coli_core.json

MAT ? : e_coli_core.mat

Downloads last updated Sep 12, 2019 | BiGG License

Publication DOI: 10.1128/ecosalplus.10.2.1 ↗

Escher Map ?



[Open map in a new window](#)

[Escher homepage](#)

What to do with metabolic networks?

1) systems biology

- how cells work? how did they evolve? \Rightarrow look at structure
- are there general principles?

2) flux balance analysis \rightarrow model of mass flow over the network.

- assume conservation of mass $\text{in} \xrightarrow{=} \text{out}$
turn it into linear algebra prob.
- modify the network to produce "desired fluxes"
- add / optimize novel pathways

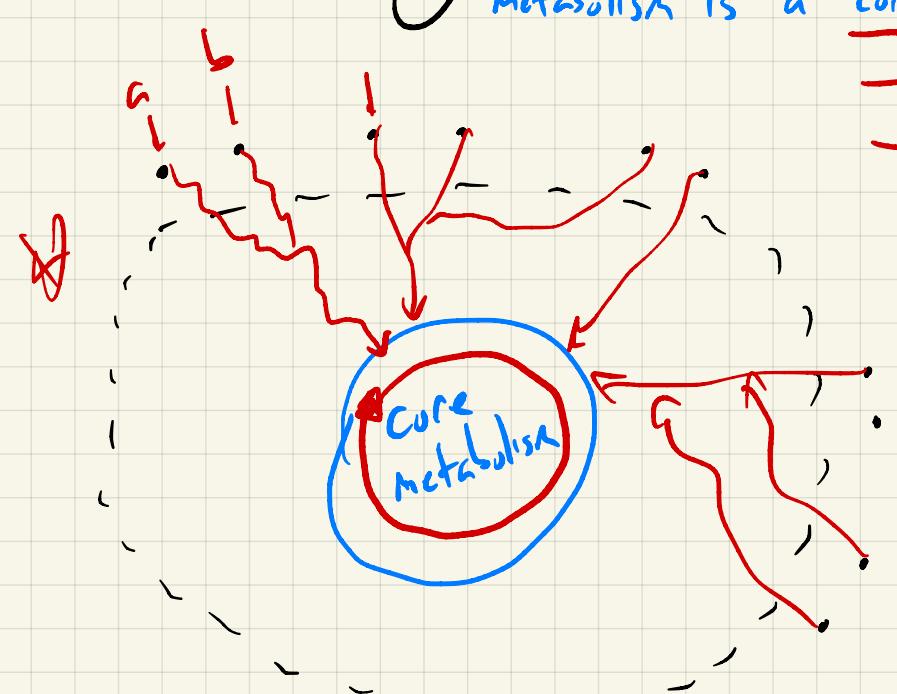
How do metabolic networks set their structure?

* still an active area of research *

* one plausible explanation: Maslov et al. (2009) the toolbox model (of catabolism)

key idea: ① over evolutionary time, a metabolic network must always function
→ no "broken pathways"

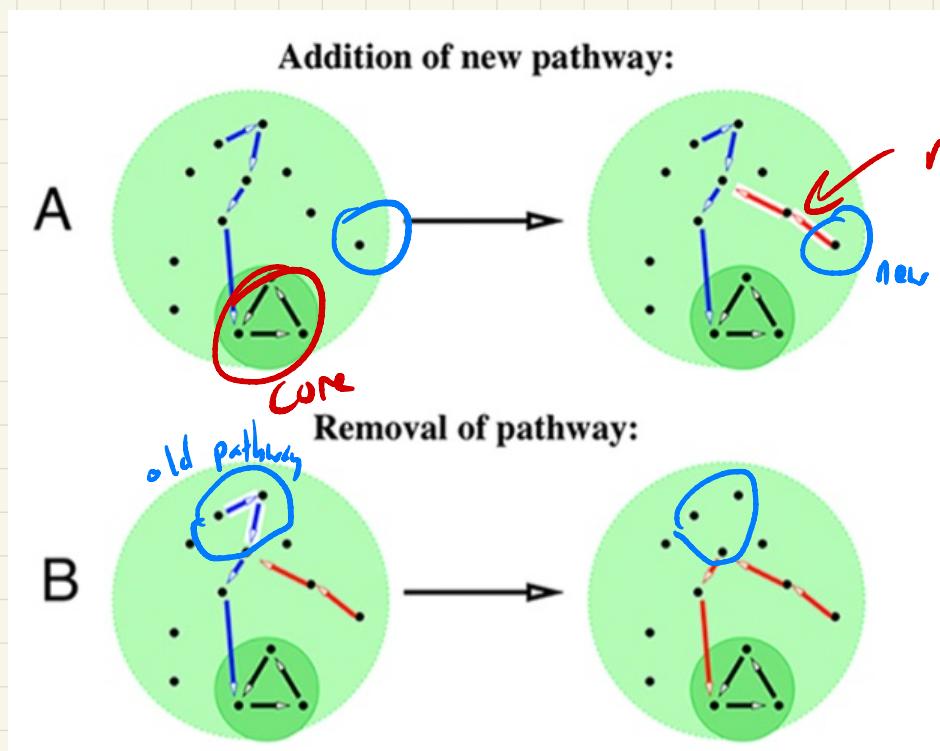
② metabolism is a "core" + "repertoire" of pathways for catalyzing nutrients
→ core cannot change
→ repertoire changes over evolutionary time



③ Separation of time scales
→ additions / deletions happen "quickly"

2 basic steps

(A) pathway addition : new nutrient to use
add to repertoire only necessary steps (new tools)
reuse existing pathways



(B) pathway removal : old nutrient lost
remove from repertoire only nutrient-specific steps (unused tools)
return all other pathways

toolbox simulation

→ all possible metabolic reactions.

* assumes a G_u , the "universal biochemistry network"

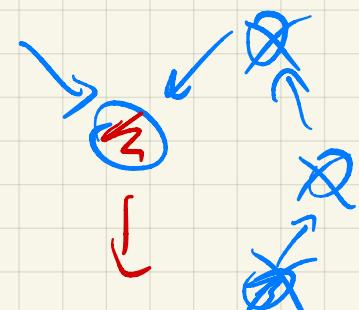
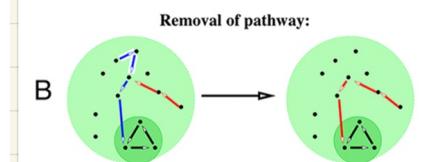
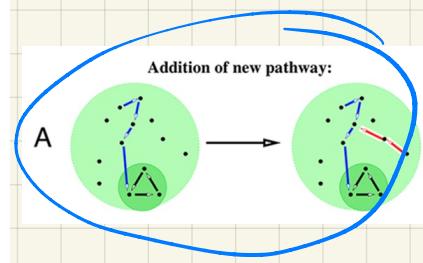
→ each species' metabolic net G is a subset of G_u

→ nodes in G_u are labeled as nutrient and metabolite

→ BigG database

1. Define a universal biochemistry network G_u , which represents the set of all possible metabolic reactions.
2. Initialize a metabolic network G , composed only of core metabolism, which we might say represents the origin of life. Mark all edges in core metabolism so that we never delete them.
3. **A: Pathway addition**
 - (a) uniformly at random, choose a new "nutrient" v , defined as leaf node of G_u that is not in G
 - (b) take a "self-avoiding random walk" (no loops) on the edges of G_u that reaches from v to any node in G
 - (c) call that set of edges σ , and add them all to G
4. **B: Pathway deletion**
 - (a) uniformly at random, choose an existing "nutrient" v from G
 - (b) starting at v , recursively delete the edges in a path toward core metabolism until we reach a node with in-degree $k_{in} > 1$, and then stop.

at each step choose A or B with equal probability.



Comments on the "toolbox model"

- ① pathway-based, which reflects our understanding of metabolism
- ② embodies the "use it or lose it" energetic selection principle of prokaryotic genome evolution
- ③ leverages a "universal biachem. network" of known (possible...) metabolic reactions (taken from a database)

Simulating the model

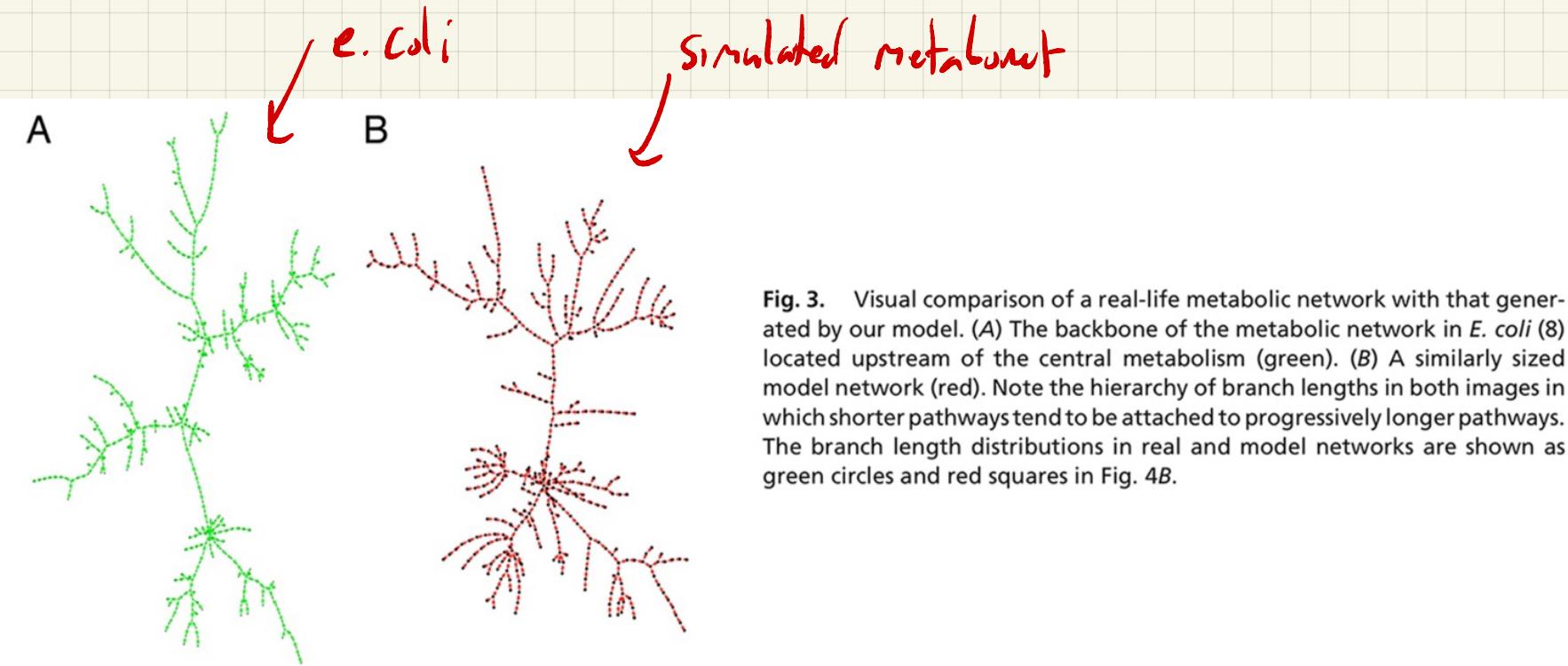


Fig. 3. Visual comparison of a real-life metabolic network with that generated by our model. (A) The backbone of the metabolic network in *E. coli* (8) located upstream of the central metabolism (green). (B) A similarly sized model network (red). Note the hierarchy of branch lengths in both images in which shorter pathways tend to be attached to progressively longer pathways. The branch length distributions in real and model networks are shown as green circles and red squares in Fig. 4B.

only structural!

no flux

no rates

no dynamics

possible variations

- * different rates for addition (A) and deletion (β)
- * non-uniform choice of "new" nutrient (A)
- * non-random walk on G_n
- * non-uniform choice of "old" nutrient (β)