



Flux Balance Analysis Overview



Learning Objectives

Each student should be able to:

- Explain flux balance analysis (FBA).
- Explain reactions, metabolites, & pathways.
- Explain mass balanced linear equations.
- Explain the biomass reaction.
- Explain how to create a stoichiometric matrix from reactions and metabolites.
- Explain gene-protein-reaction associations.
- Explain the constraint-based modeling.



Flux Balance Analysis Overview

- ➡ • Flux Balance Analysis Overview
- Reactions, Metabolites, & Pathways
- Mathematical Representation of Reactions & Constraints
- Mass Balanced Linear Equations
- Biomass Reaction
- Calculating Fluxes
- COBRA Toolbox v.3.0



Flux Balance Analysis

Orth, J. D., I. Thiele, et al. (2010). "What is flux balance analysis?" Nature biotechnology 28(3): 245-248.

- Through the use of genome-scale metabolic network reconstructions, Flux Balance Analysis (FBA) can be used to calculate the flow of metabolites (flux) through a metabolic network. This capability makes it possible to predict the growth rate of an organism and/or the rate of production of a given metabolite.
- FBA has limitations! It does not use kinetic parameters, thus it cannot predict metabolite concentrations. It is also only capable of determining fluxes at steady state. Typically, FBA does not account for regulatory effects such as activation of enzymes by protein kinases or regulation of gene expression. Therefore, its predictions may not always be accurate.



Flux Balance Analysis Overview

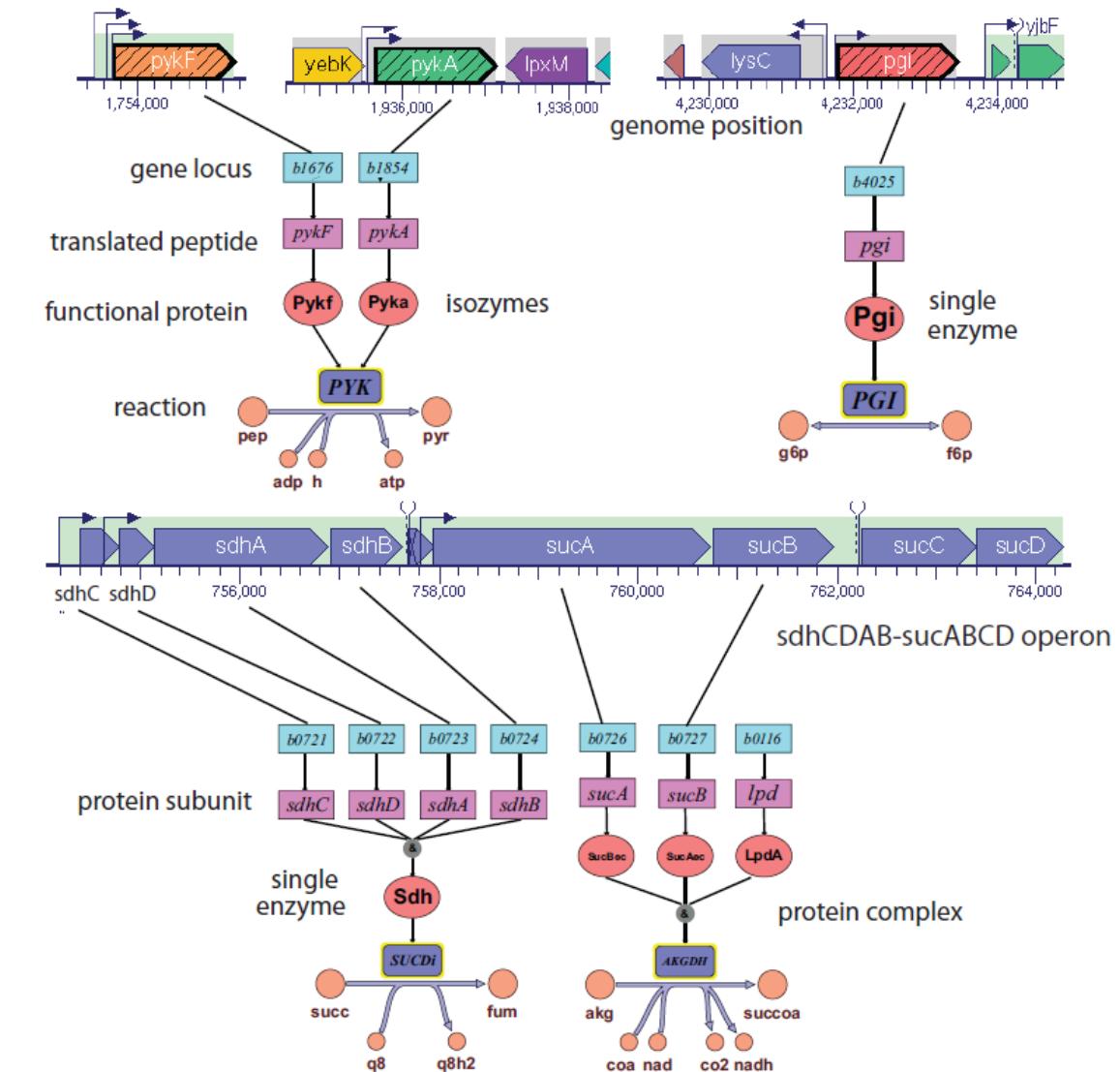
- Flux Balance Analysis Overview
- Reactions, Metabolites, & Pathways
- Mathematical Representation of Reactions & Constraints
- Mass Balanced Linear Equations
- Biomass Reaction
- Calculating Fluxes
- COBRA Toolbox v.3.0



Identifying Metabolic Reactions and Metabolites (Gene-Protein-Reactions)

Objective:

Create A biochemically, genetically and
genomically (BiGG) structured knowledge base.

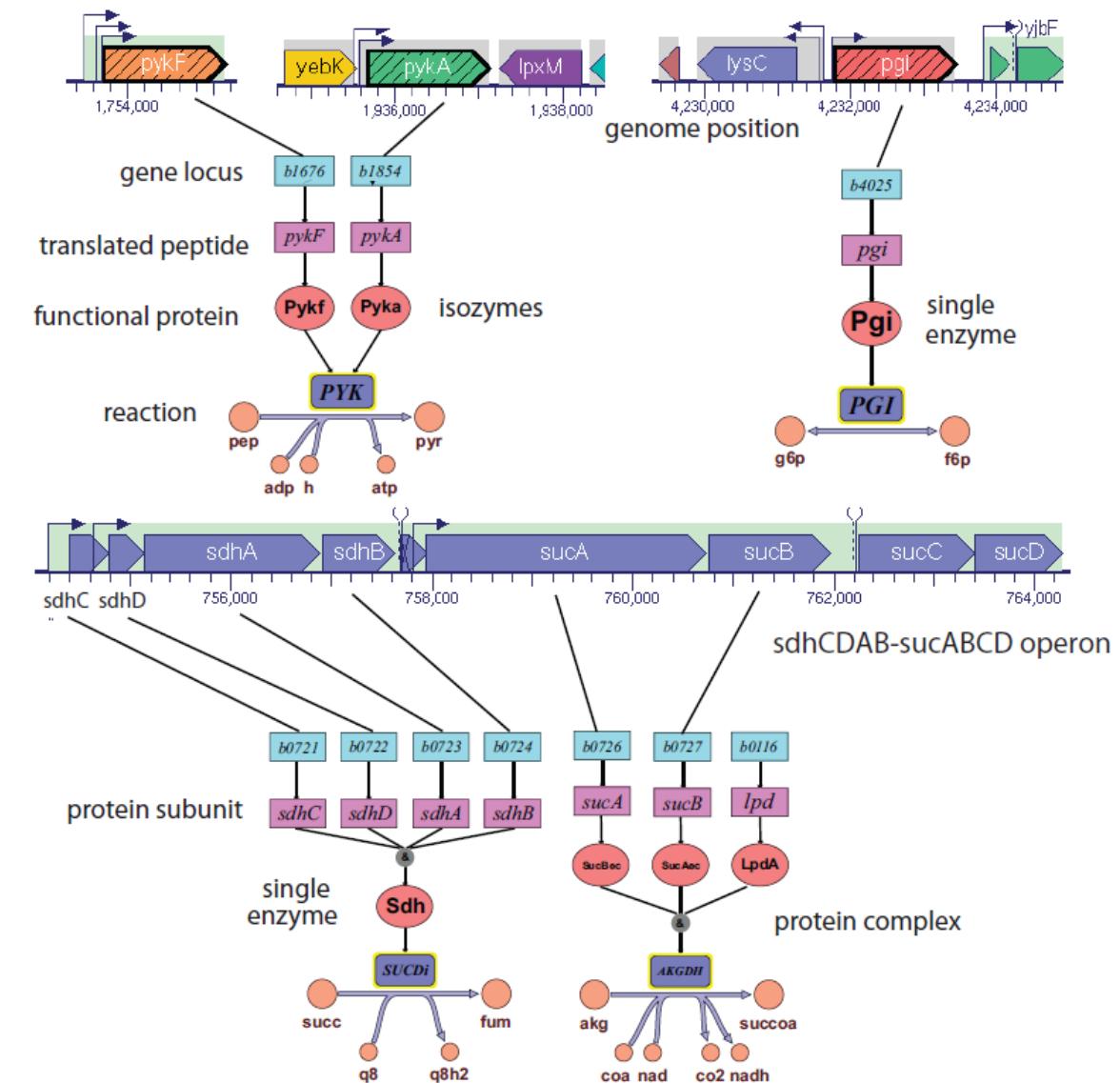




Desired Reaction Information

1. Reaction Name*
2. Reaction Description*
3. Reaction Formula*
4. Gene-reaction Association*
5. Genes (Gene Locus) *
6. Proteins
7. Cellular Subsystem *
(e.g. Glycolysis)
8. Reaction Direction*
9. Flux Lower Bound*
10. Flux Upper Bound*
11. Confidence Score (1-5)
12. EC Number
13. Notes
14. References

* Required



Reconstruction and Use of Microbial Metabolic Networks: the Core Escherichia coli Metabolic Model as an Educational Guide by Orth, Fleming, and Palsson (2010)



Genome-scale Reconstruction Reactions

| abbreviation | officialName | equation | subSystem | proteinClass | delta G (pH 7.2) | Keq | v |
|----------------|---|--|----------------------------|--------------|------------------|------------|---|
| ACALD | acetaldehyde dehydrogenase (acetylating) | [c] : acald + coa + nad <=> accoa + h + nadh | Pyruvate Metabolism | 1.2.1.10 | -4.4 | 1.66E+03 | |
| ACALDt | acetaldehyde reversible transport | acald[e] <=> acald[c] | Transport, Extracellular | | 0.0 | 1.00E+00 | |
| ACKr | acetate kinase | [c] : ac + atp <=> actp + adp | Pyruvate Metabolism | 2.7.2.1 | 4.3 | 7.1204E-04 | |
| ACONTa | aconitase (half-reaction A, Citrate hydro-lyase) | [c] : cit <=> acon-C + h2o | Citric Acid Cycle | 4.2.1.3 | 1.5 | 7.98E-02 | |
| ACONTb | aconitase (half-reaction B, Isocitrate hydro-lyase) | [c] : acon-C + h2o <=> icit | Citric Acid Cycle | 4.2.1.3 | -0.2 | 1.40E+00 | |
| ACt2r | acetate reversible transport via proton symport | ac[e] + h[e] <=> ac[c] + h[c] | Transport, Extracellular | | 0.0 | 1.0000E+00 | |
| ADK1 | adenylate kinase | [c] : amp + atp <=> (2) adp | Oxidative Phosphorylation | 2.7.4.3 | -0.1 | 1.1836E+00 | |
| AKGDH | 2-Oxoglutarate dehydrogenase | [c] : akg + coa + nad --> co2 + nadh + succoa | Citric Acid Cycle | | -8.3 | 1.1896E+06 | |
| AKGt2r | 2-oxoglutarate reversible transport via symport | akg[e] + h[e] <=> akg[c] + h[c] | Transport, Extracellular | | 0.0 | 1.0000E+00 | |
| ALCD2x | alcohol dehydrogenase (ethanol) | [c] : etoh + nad <=> acald + h + nadh | Pyruvate Metabolism | 1.1.1.1 | 6.0 | 4.06E-05 | |
| ATPM | ATP maintenance requirement | [c] : atp + h2o --> adp + h + pi | Oxidative Phosphorylation | | -6.6 | 6.7770E+04 | |
| ATPS4r | ATP synthase (four protons for one ATP) | adp[c] + (4) h[e] + pi[c] <=> atp[c] + (3) h[c] + h2o[c] | Oxidative Phosphorylation | 3.6.3.14 | 6.6 | 1.4756E-05 | |
| Biomass_Ecoli_ | Biomass Objective Function with GAM | [c] : (1.496) 3pg + (3.7478) accoa + (59.8100) atp + (0.3610) e4p + (0.0709) f6p + (0.1290) g3p + (0.2050) g6p + (0.2557) gln-L + (4.9414) glu-L | | | | | |
| CO2t | CO2 transporter via diffusion | co2[e] <=> co2[c] | Transport, Extracellular | | 0.0 | 1.0000E+00 | |
| CS | citrate synthase | [c] : accoa + h2o + oaa --> cit + coa + h | Citric Acid Cycle | | -8.6 | 1.9724E+06 | |
| CYTBD | cytochrome oxidase bd (ubiquinol-8: 2 protons) | (2) h[c] + (0.5) o2[c] + q8h2[c] --> (2) h[e] + h2o[c] + q8 | Oxidative Phosphorylation | | -37.2 | 1.6962E+27 | |
| D_LAct2 | D-lactate transport via proton symport | h[e] + lac-D[e] <=> h[c] + lac-D[c] | Transport, Extracellular | | 0.0 | 1.0000E+00 | |
| ENO | enolase | [c] : 2pg <=> h2o + pep | Glycolysis/Gluconeogenesis | 4.2.1.11 | -0.9 | 4.5580E+00 | |
| ETOHt2r | ethanol reversible transport via proton symport | etoh[e] + h[e] <=> etoh[c] + h[c] | Transport, Extracellular | | 0.0 | 1.0000E+00 | |
| EX_ac(e) | Acetate exchange | [e] : ac <=> | Exchange | | | | |
| EX_acald(e) | Acetaldehyde exchange | [e] : acald <=> | Exchange | | | | |
| EX_akg(e) | 2-Oxoglutarate exchange | [e] : akg <=> | Exchange | | | | |
| EX_co2(e) | CO2 exchange | [e] : co2 <=> | Exchange | | | | |
| EX_etoh(e) | Ethanol exchange | [e] : etoh <=> | Exchange | | | | |
| EX_for(e) | Formate exchange | [e] : for <=> | Exchange | | | | |
| EX_fru(e) | D-Fructose exchange | [e] : fru <=> | Exchange | | | | |
| EX_fum(e) | Fumarate exchange | [e] : fum <=> | Exchange | | | | |
| EX_glc(e) | D-Glucose exchange | [e] : glc-D <=> | Exchange | | | | |
| EX_gln_L(e) | L-Glutamine exchange | [e] : gln-L <=> | Exchange | | | | |
| EX_glu_L(e) | L-Glutamate exchange | [e] : glu-L <=> | Exchange | | | | |

ecoli_textbook.xls



Desired Metabolite Information

- 1. Metabolite Name*
- 2. Metabolite Description*
- 3. Metabolite Neutral Formula
- 4. Metabolite Charged Formula*
- 5. Metabolite Charge*
- 6. Metabolite Compartment*
- 7. Metabolite KEGGID
- 8. Metabolite PubChemID
- 9. Metabolite CheBI ID
- 10. Metabolite Inchi String
- 11. Metabolite Smile

* Required

| | | | | |
|----------------|--|---|--|---|
| Gene | b2388 | Locus | Genomics ORF annotation | |
| Peptide | glk | Gene | Transcriptomics mRNA levels | |
| Protein | Glk | Protein | Proteomics protein levels | |
| | | Reaction | "Fluxomics" flux measurements | |
| Substrates | glc | atp | g6p | adp |
| Neutral | C ₆ H ₁₂ O ₆ ⁰ | C ₁₀ H ₁₆ N ₅ O ₁₃ P ₃ ⁰ | C ₆ H ₁₃ O ₉ P ⁰ | C ₁₀ H ₁₅ N ₅ O ₁₀ P ₂ ⁰ |
| Charged | C ₆ H ₁₂ O ₆ ⁰ | C ₁₀ H ₁₂ N ₅ O ₁₃ P ₃ ⁻⁴ | C ₆ H ₁₁ O ₉ P ²⁻ | C ₁₀ H ₁₂ N ₅ O ₁₀ P ₂ ⁻³ |
| | | | | |
| Stoichiometry | C ₁₆ H ₂₄ O ₁₉ P ₃ , 4e ⁻ | = | C ₁₆ H ₂₃ O ₁₈ P ₃ , 5e ⁻ | |
| | 1 glc + 1 atp | | 1 g6p + 1 adp + 1 h ⁺ | |
| Directionality | 1 glc + 1 atp | → | 1 g6p + 1 adp + 1 h ⁺ | |
| Location | cytosol: 1 glc + 1 atp | → | 1 g6p + 1 adp + 1 h ⁺ | |

Thiele, I. and B. O. Palsson (2010). "A protocol for generating a high-quality genome-scale metabolic reconstruction." *Nature protocols* 5(1): 93-121.



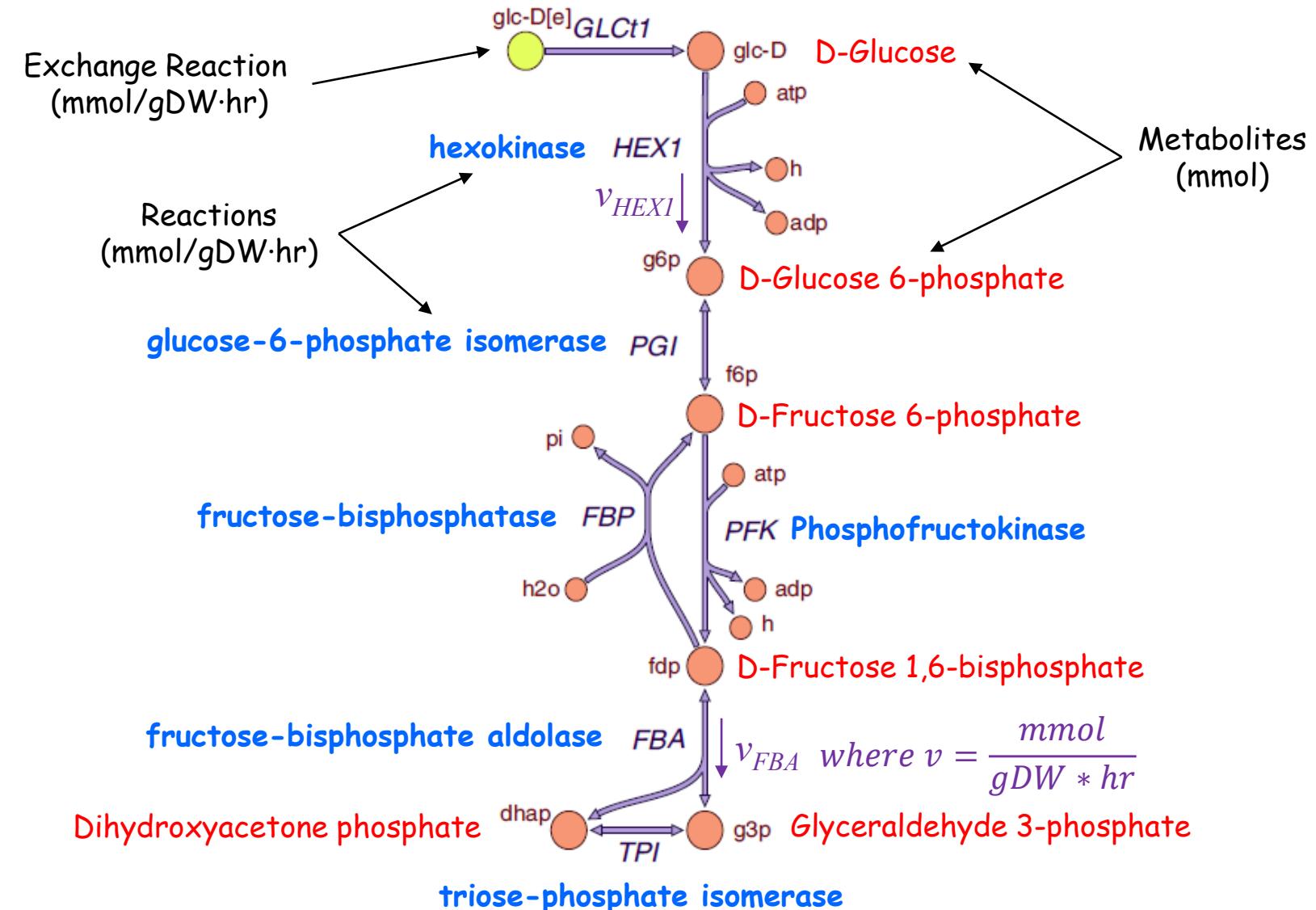
Genome-scale Reconstruction Metabolites

| abbreviation | officialName | formula | charge | casNumber | formulaNeutral | CompoundNames | KeggID |
|--------------|---------------------------------|---------------|--------|------------|----------------|--|--------|
| 13dpG | 3-Phospho-D-glyceroyl phosphate | C3H4O10P2 | -4 | 38168-82-0 | C3H8O10P2 | 1,3-bis-phosphoglycerate/ 3-Phospho-D-glyceroyl phosphate | C00236 |
| 2pg | D-Glycerate 2-phosphate | C3H4O7P | -3 | None | C3H7O7P | 2-phosphoglyceric acid/ 2-Phospho-D-glycerate | C00631 |
| 3pg | 3-Phospho-D-glycerate | C3H4O7P | -3 | None | C3H7O7P | D-Glycerate 3-phosphate/ 3-Phosphoglycerate/ 3-phosphogly | C00197 |
| 6pgc | 6-Phospho-D-gluconate | C6H10O10P | -3 | None | C6H13O10P | 6-phosphogluconic acid/ D-gluconate 6-phosphate | C00345 |
| 6pgl | 6-phospho-D-glucono-1,5-lactone | C6H9O9P | -2 | None | C6H11O9P | D-Glucono-1,5-lactone 6-phosphate | C01236 |
| ac | Acetate | C2H3O2 | -1 | 71-50-1 | C2H4O2 | vinegar/ Ethylic acid/ Vinegar acid/ Methanecarboxylic acid/ A | C00033 |
| ac[e] | Acetate (extracellular) | C2H3O2 | -1 | 71-50-1 | C2H4O2 | vinegar/ Ethylic acid/ Vinegar acid/ Methanecarboxylic acid/ A | C00033 |
| acald | Acetaldehyde | C2H4O | 0 | 75-07-0 | | Ethanal/ Aldehyde C(2)/ acetylaldehyde/ Aceteldehyde/ Acet | C00084 |
| acald[e] | Acetaldehyde (extracellular) | C2H4O | 0 | 75-07-0 | | Ethanal/ Aldehyde C(2)/ acetylaldehyde/ Aceteldehyde/ Acet | C00084 |
| accoa | Acetyl-CoA | C23H34N7O17P3 | -4 | 72-89-9 | C23H38N7O17P3S | Acetyl coenzyme A | C00024 |
| acon-C | cis-Aconitate | C6H3O6 | -3 | 585-84-2 | C6H6O6 | cis-1,2,3-Propenetricarboxylic acid/ (Z)-1-Propene-1,2,3-tricar | C00417 |
| actp | Acetyl phosphate | C2H3O5P | -2 | 19926-71-7 | C2H5O5P | | C00227 |
| adp | ADP | C10H12N5O10P2 | -3 | 58-64-0 | C10H15N5O10P2 | Adenosine 5'-diphosphate | C00008 |
| akg | 2-Oxoglutarate | C5H4O5 | -2 | 328-50-7 | C5H6O5 | Oxoglutaric acid/ 2-Ketoglutaric acid/ alpha-Ketoglutarate/ alp | C00026 |
| akg[e] | 2-Oxoglutarate (extracellular) | C5H4O5 | -2 | 328-50-7 | C5H6O5 | Oxoglutaric acid/ 2-Ketoglutaric acid/ alpha-Ketoglutarate/ alp | C00026 |
| amp | AMP | C10H12N5O7P | -2 | 61-19-8 | C10H14N5O7P | Adenosine 5'-monophosphate/ Adenylic acid/ Adenylate/ 5'-A | C00020 |
| atp | ATP | C10H12N5O13P3 | -4 | 56-65-5 | C10H16N5O13P3 | Adenosine 5'-triphosphate/ 5'-adenylate triphosphate | C00002 |
| cit | Citrate | C6H5O7 | -3 | 77-92-9 | C6H8O7 | Citric acid/ 2-Hydroxytricarballylic acid/ 2/ 2-Hydroxy-1/ 3-pro | C00158 |
| co2 | CO2 | CO2 | 0 | 124-38-9 | | Carbonic anhydride/ Carbon dioxide | C00011 |
| co2[e] | CO2 (extracellular) | CO2 | 0 | 124-38-9 | | Carbonic anhydride/ Carbon dioxide | C00011 |
| coa | Coenzyme A | C21H32N7O16P3 | -4 | 85-61-0 | C21H36N7O16P3S | CoA/ CoA-SH/ CoASH | C00010 |
| dhap | Dihydroxyacetone phosphate | C3H5O6P | -2 | 57-04-5 | C3H7O6P | Glycerone phosphate | C00111 |
| e4p | D-Erythrose 4-phosphate | C4H7O7P | -2 | 585-18-2 | C4H9O7P | 4-phospho D-erythrose | C00279 |
| etoh | Ethanol | C2H6O | 0 | 64-17-5 | | Ethyl alcohol/ Methylcarbinol | C00469 |
| etoh[e] | Ethanol (extracellular) | C2H6O | 0 | 64-17-5 | | Ethyl alcohol/ Methylcarbinol | C00469 |
| f6p | D-Fructose 6-phosphate | C6H11O9P | -2 | 643-13-0 | C6H13O9P | D-Fructose 6-phosphoric acid/ Neuberg ester/ beta-D-Fructos | C00085 |
| fdp | D-Fructose 1,6-bisphosphate | C6H10O12P2 | -4 | 488-69-7 | C6H14O12P2 | fructose-1,6-bisphosphate/ fructose diphosphate/ beta-D-fruc | C00354 |
| for | Formate | CH1O2 | -1 | 64-18-6 | CH2O2 | Methanoic acid/ Formic acid/ Hydrogencarboxylic acid/ aminic | C00058 |

coli_textbook.xls



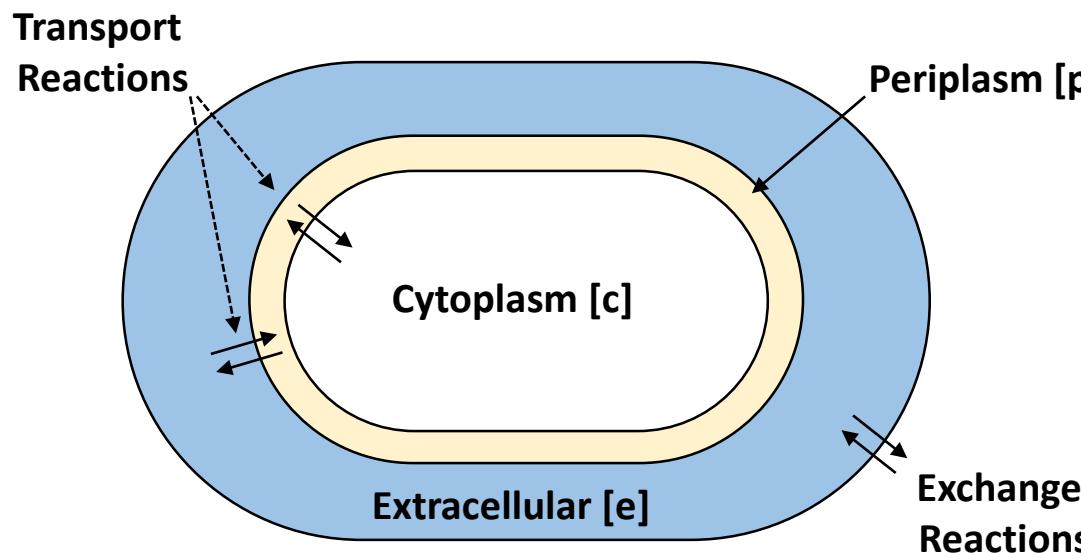
Metabolic Pathway



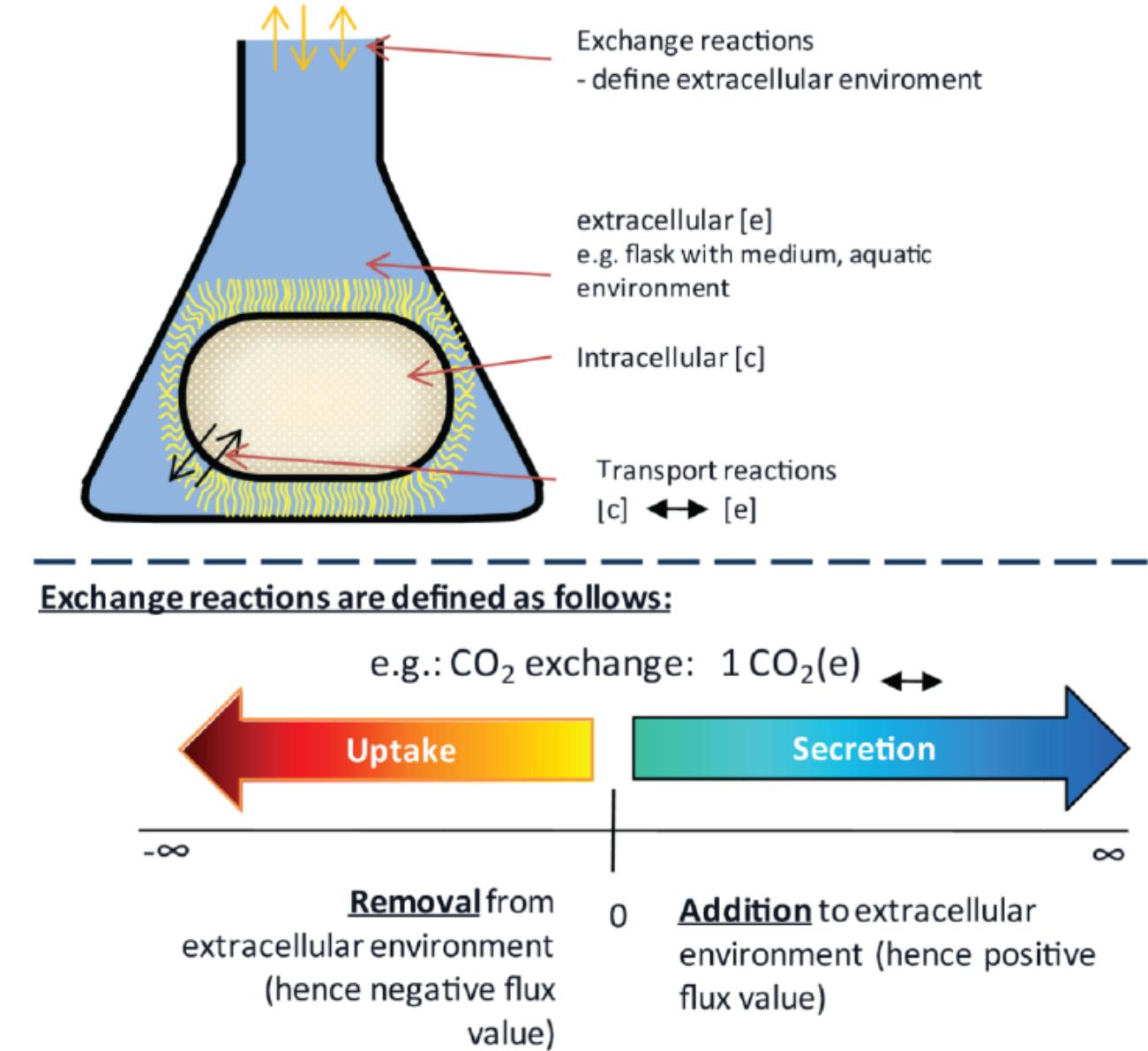
Becker, S. A., et al. (2007). "Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox." *Nature protocols* 2(3): 727-738.



System Boundaries: Exchange & Transport Reactions



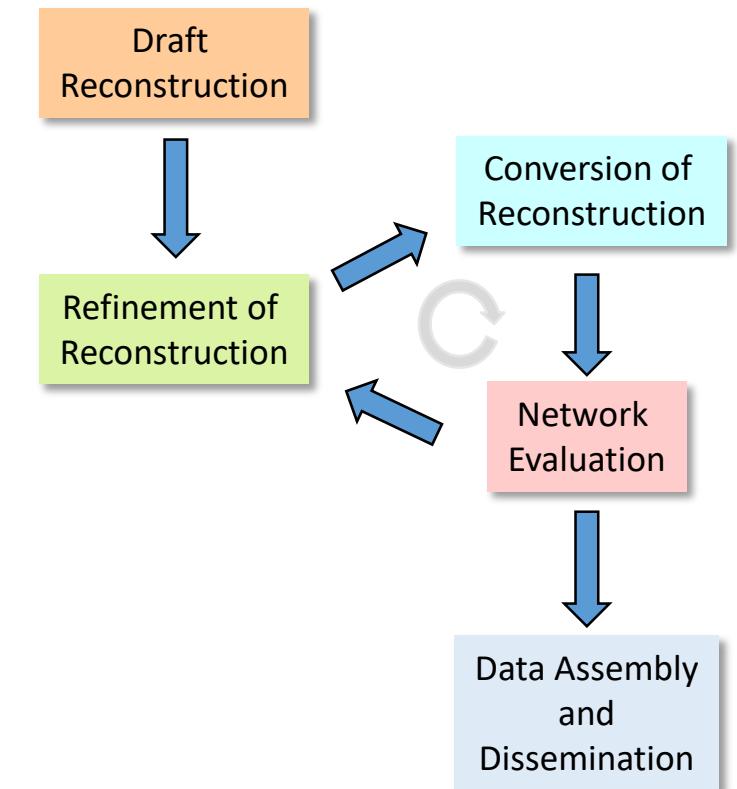
Thiele, I. and B. O. Palsson (2010). "A protocol for generating a high-quality genome-scale metabolic reconstruction." *Nature protocols* 5(1): 93-121.





Genome-scale Metabolic Reconstructions

- Overview
- Draft Reconstruction
- Refinement of Reconstruction
- Conversion of Reconstruction into Computable Format
- Network Evaluation
- Data Assembly and Dissemination

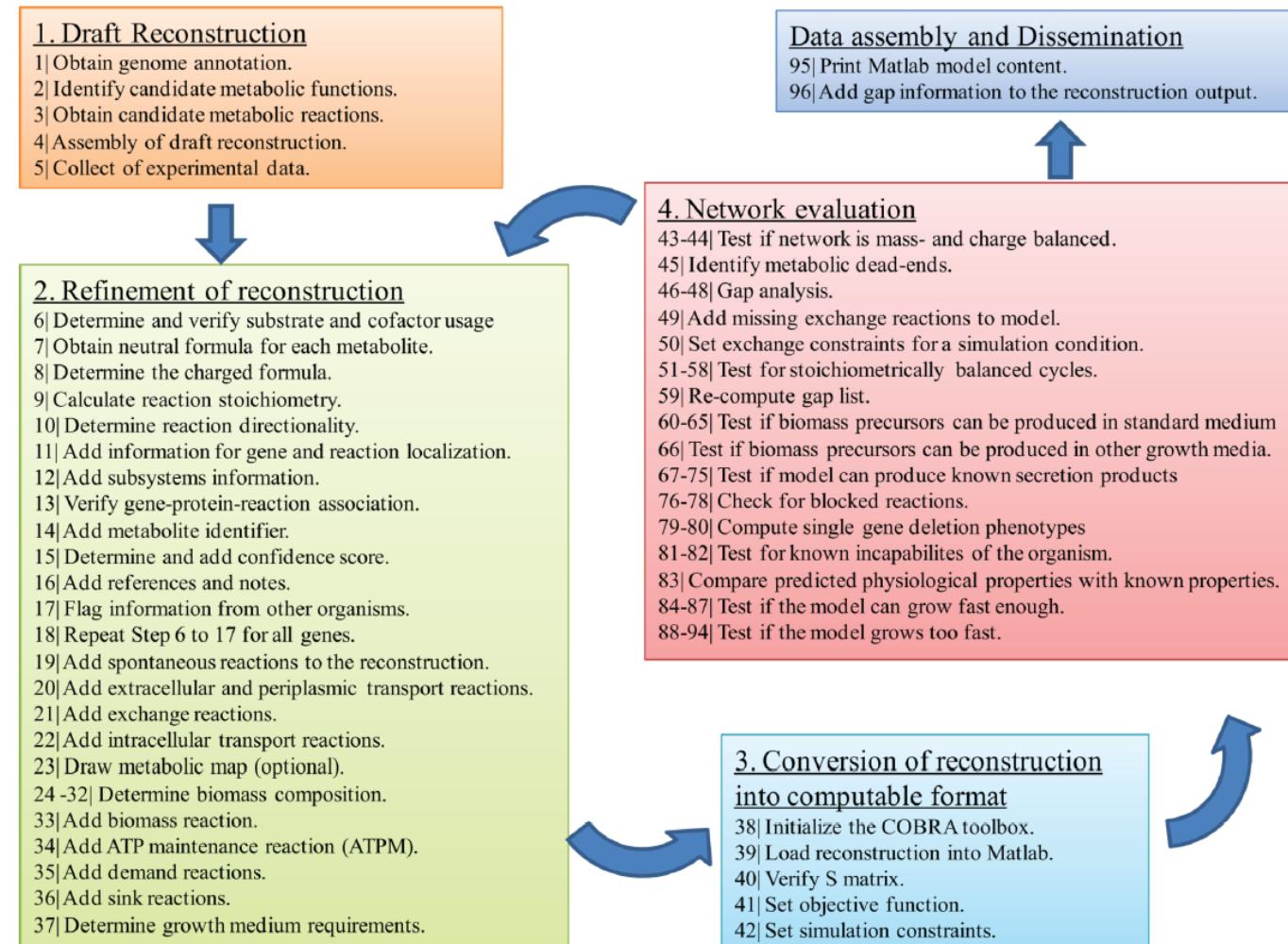


Thiele, I. and B. O. Palsson (2010). "A protocol for generating a high-quality genome-scale metabolic reconstruction." Nature protocols 5(1): 93-121.



Reconstruction Process: 96 Step Protocol

Thiele, I. and B. O. Palsson (2010). "A protocol for generating a high-quality genome-scale metabolic reconstruction." *Nature protocols* 5(1): 93-121.



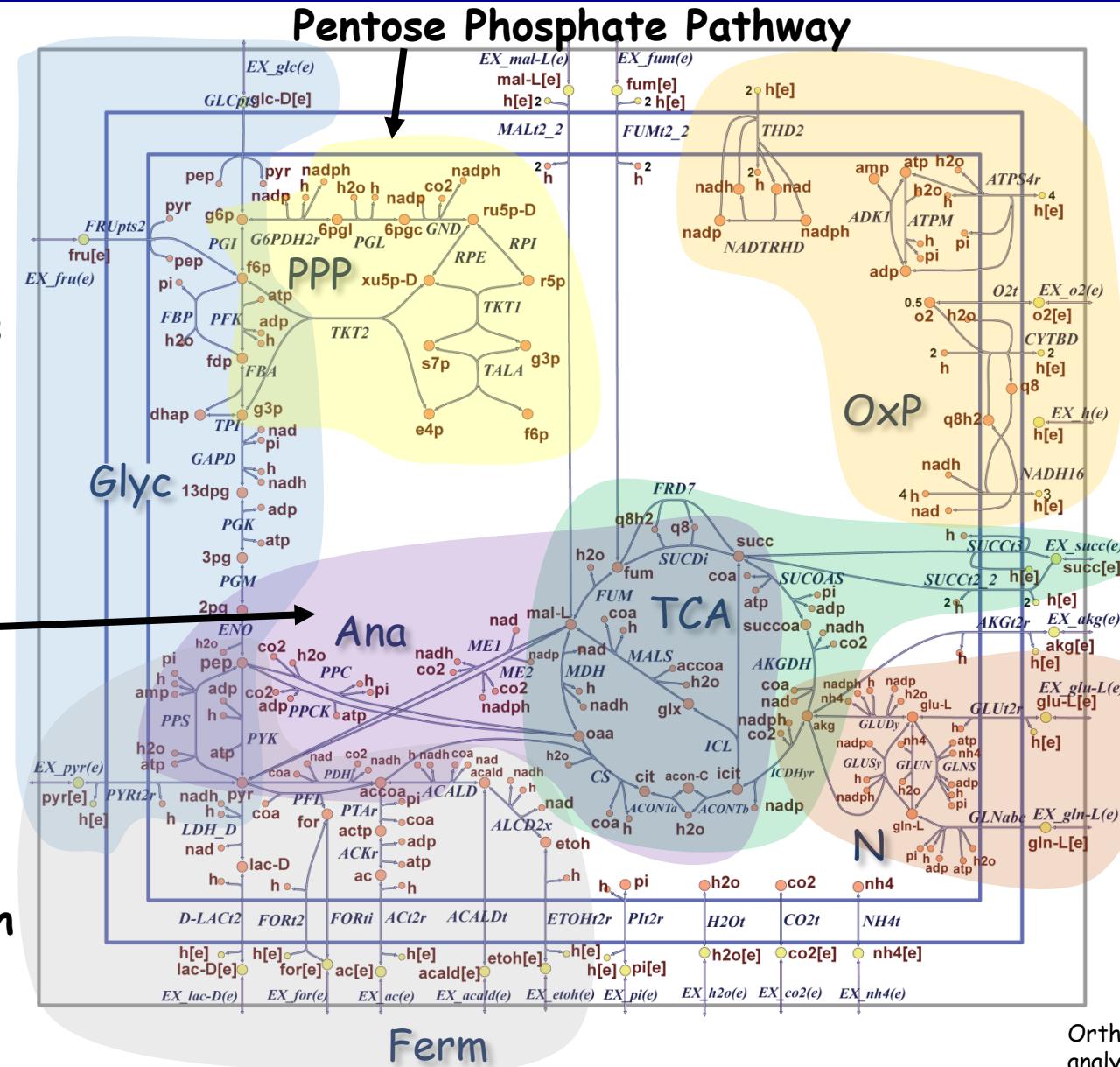


E. coli Core Model

Glycolysis

Glyoxylate Cycle,
Gluconeogenesis, and
Anapleurotic Reactions

Fermentation



Oxidative
Phosphorylation and
Transfer of Reducing
Equivalents

Tricarboxylic Acid
Cycle (TCA)

Nitrogen
Metabolism

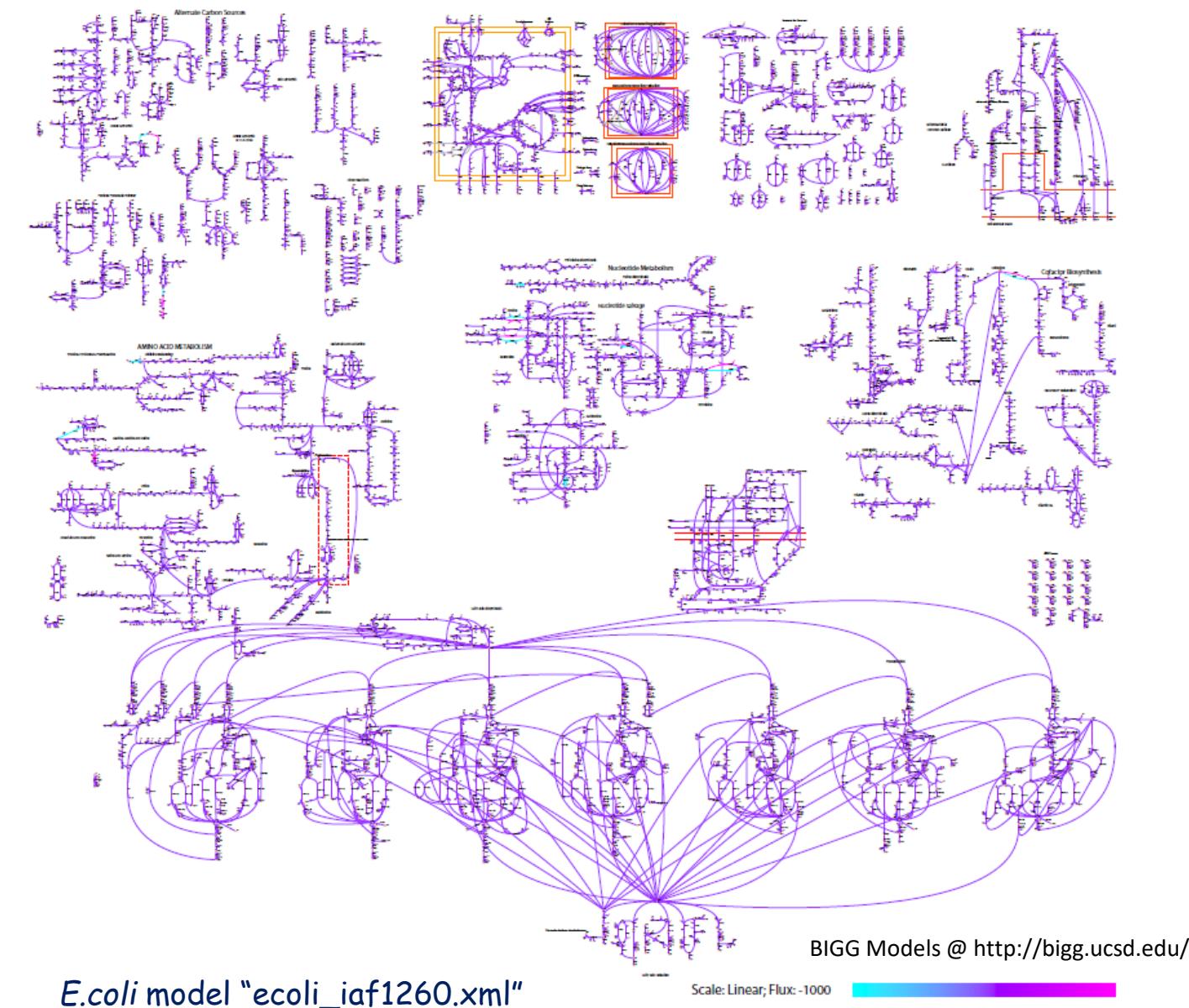
Orth, J. D., I. Thiele, et al. (2010). "What is flux balance analysis?" *Nature biotechnology* 28(3): 245-248.



E. coli K-12 MG1655 Genome-Scale Reconstructions

iAF1260 - Feist, A. M., C. S. Henry, et al. (2007).
"A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information." Molecular Systems Biology 3: 121.

iJO1366 - Orth, J. D. and B. O. Palsson (2012).
"Gap-filling analysis of the iJO1366 Escherichia coli metabolic network reconstruction for discovery of metabolic functions." BMC systems biology 6(1): 30.





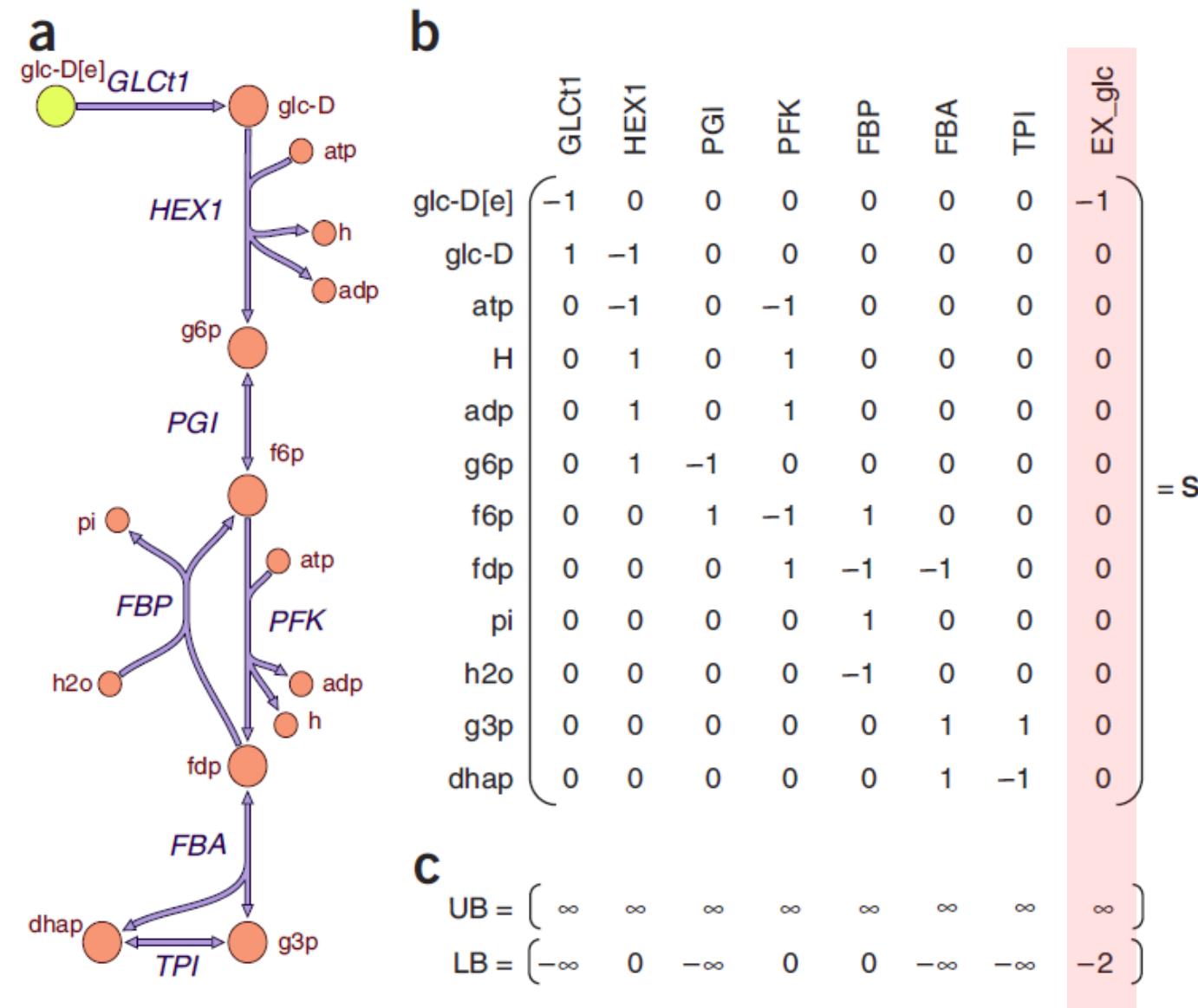
Flux Balance Analysis Overview

- Flux Balance Analysis Overview
- Reactions, Metabolites, & Pathways
- • Mathematical Representation of Reactions & Constraints
- Mass Balanced Linear Equations
- Biomass Reaction
- Calculating Fluxes
- COBRA Toolbox v.3.0



Creating A Stoichiometric Matrix

The stoichiometric matrix, S , is the centerpiece of a mathematical representation of genome-scale metabolic networks. This matrix represents each reaction as a column and each metabolite as a row, where each numerical element is the corresponding stoichiometric coefficient.



Becker, S. A., A. M. Feist, et al. (2007). "Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox." *Nature protocols* 2(3): 727-738.



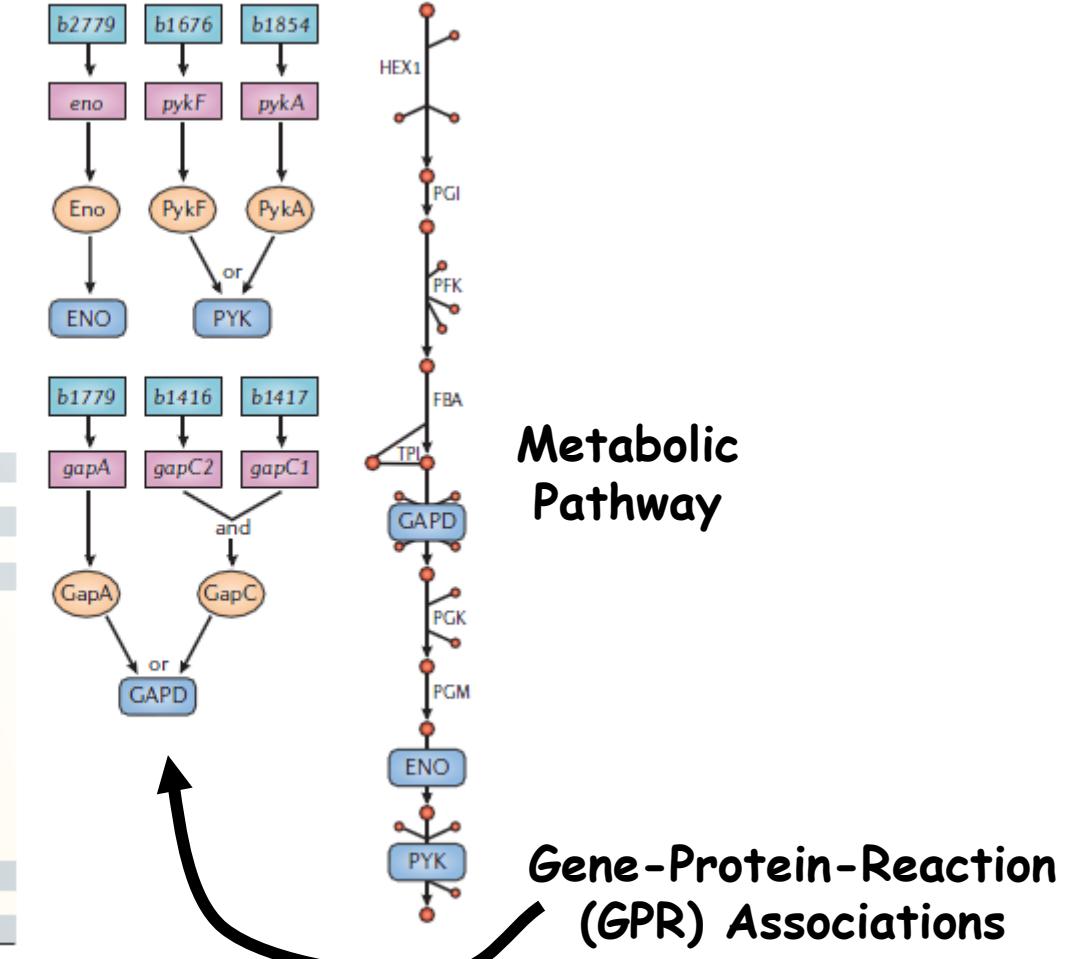
Genome-scale Metabolic Reconstruction

**BIGG
Database**

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

**Stoichiometric
Matrix**

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



Reed, J. L., I. Famili, et al. (2006). "Towards multidimensional genome annotation." *Nature reviews. Genetics* 7(2): 130-141.



Flux Balance Analysis Overview

- Flux Balance Analysis Overview
- Reactions, Metabolites, & Pathways
- Mathematical Representation of Reactions & Constraints
- • Mass Balanced Linear Equations
- Biomass Reaction
- Calculating Fluxes
- COBRA Toolbox v.3.0



How can we use the Stoichiometric Matrix?

- The stoichiometric matrix, S , is a linear transformation of the flux vector, v to a vector of time derivatives of the concentration vector, x .

$$\frac{dx}{dt} = S \cdot v$$

- The concentration vector, x , represents the concentration of each of the metabolites.
- If we assume that a cell will be in a particular phenotype for a time much larger than the changing time of metabolites then we can also assume that the concentration pools for the metabolites will be non-changing thus setting $dx/dt = 0$. This is the steady state assumption of flux balance analysis.

$$\frac{dx}{dt} = 0 = S \cdot v$$

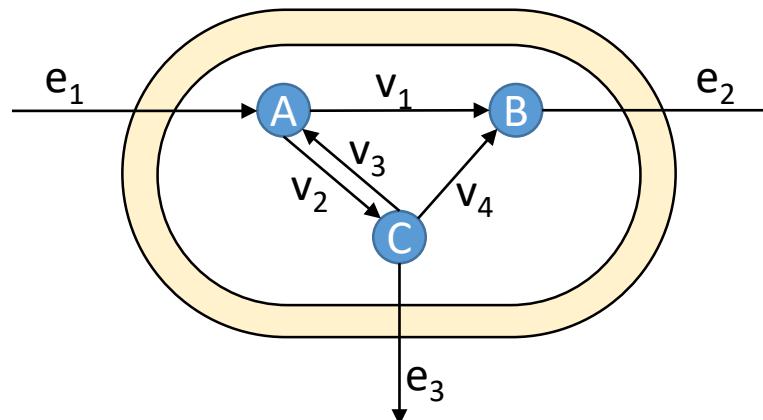
- Since there are normally many more reactions (columns) than metabolites (rows), more unknown variables than equations, then there is no unique solutions (could be a large number of solutions).
- Need to find a way to constrain the solution space!

| Metabolites | Reactions | | | | | | | |
|-------------|-----------|----|----|----|----|----|----|----|
| | -1 | 0 | 0 | 0 | 0 | 0 | 0 | -1 |
| | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 0 | -1 | 0 | -1 | 0 | 0 | 0 | 0 |
| | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 |
| | 0 | 0 | 1 | -1 | 1 | 0 | 0 | 0 |
| | 0 | 0 | 0 | 1 | -1 | -1 | 0 | 0 |
| | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0 |
| | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| | 0 | 0 | 0 | 0 | 0 | 1 | -1 | 0 |



Dynamic Mass Balance

A simple network



Linear Transformation

$$\frac{d\mathbf{x}}{dt} = \mathbf{S} \cdot \mathbf{v}$$

$$\begin{bmatrix} \frac{dA}{dt} \\ \frac{dB}{dt} \\ \frac{dC}{dt} \end{bmatrix} = \begin{bmatrix} v_1 & v_2 & v_3 & v_4 & e_1 & e_2 & e_3 \\ -1 & -1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & -1 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ e_1 \\ e_2 \\ e_3 \end{bmatrix}$$

Stoichiometric Matrix

Linear Differential Equations

$$\frac{dA}{dt} = -v_1 - v_2 + v_3 + e_1$$

$$\frac{dB}{dt} = v_1 + v_4 - e_2$$

$$\frac{dA}{dt} = v_2 - v_3 - v_4 - e_3$$

Dynamic Mass Balance (Steady State)

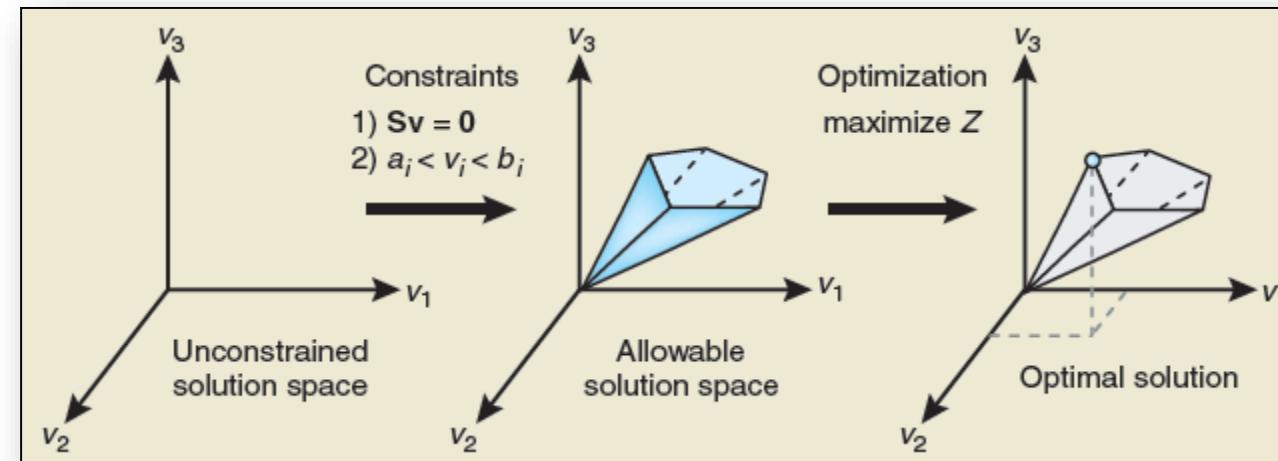
$$0 = \mathbf{S} \cdot \mathbf{v}$$

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 & -1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & -1 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ e_1 \\ e_2 \\ e_3 \end{bmatrix}$$

Note: More unknown variables than equations, thus no unique solutions! Need constraints!



The Conceptual Basis of Constraint-based Modeling



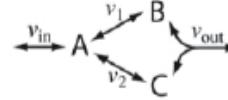
With no constraints, the flux distribution of a biological network may lie at any point in a solution space. When mass balance constraints imposed by the stoichiometric matrix S (label 1) and capacity constraints imposed by the lower and upper bounds (a_i and b_i) (label 2) are applied to a network, it defines an allowable solution space. The network may acquire any flux distribution within this space, but points outside this space are denied by the constraints. Through optimization of an objective function using linear programming, FBA can identify a single optimal flux distribution that lies on the edge of the allowable solution space.

Orth, J. D., I. Thiele, et al. (2010). "What is flux balance analysis?" *Nature biotechnology* 28(3): 245-248.

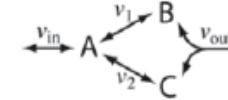


Role of Constraints

Metabolic Network

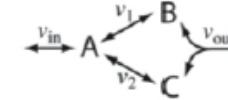


Thermodynamics (reversibility)



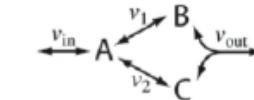
$$v_{out}, v_{in} \geq 0$$

Maximum enzyme capacity



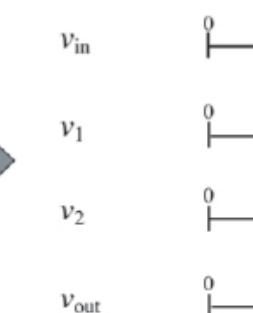
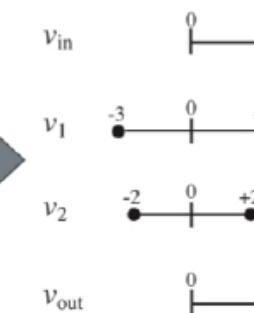
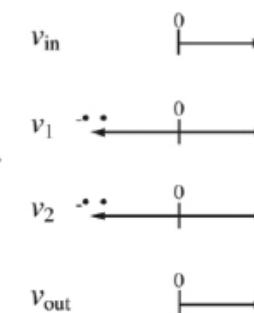
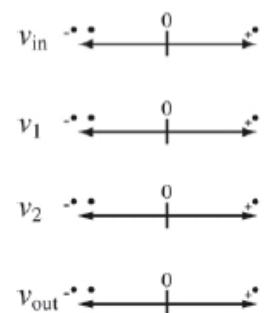
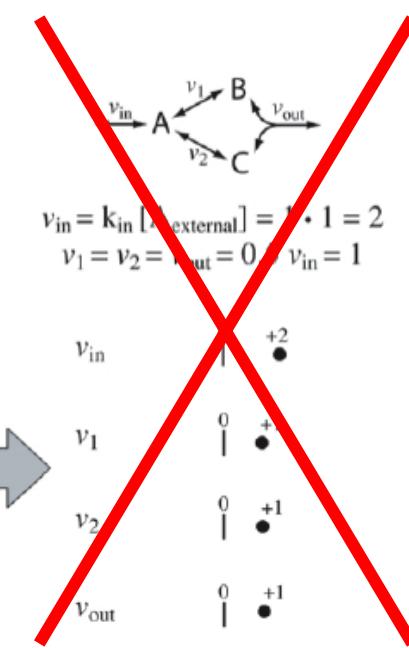
$$|v_1| \leq 3, |v_2| \leq 2$$

Mass balance of metabolites



$$\begin{aligned} v_{in} - v_1 - v_2 &= 0 \\ v_{out} - v_1 &= 0 = v_{out} - v_2 \end{aligned}$$

Kinetics





Flux Optimization

(Linear Programming or Linear Optimization Problem)

Maximize the objective function

$$Z = \sum_i c_i v_i = \mathbf{c} \cdot \mathbf{v}$$

with the following constraints

$$\frac{d\mathbf{x}}{dt} = \mathbf{S} \cdot \mathbf{v} = \mathbf{0}$$

$$\alpha_j \leq v_j \leq \beta_j$$

where

\mathbf{x} = concentration vector

\mathbf{v} = flux vector

\mathbf{c} = objective function weights

\mathbf{S} = Stoichiometric matrix

α_j = Lower bound of flux

β_j = upper bound of flux

The goal is to create an objective function that is biologically meaningful. These could include:

1. Cellular growth (maximization)
2. Particular metabolite engineering (maximization)
3. Energy consumption (minimization)

For the case of cellular growth as the objective function (Biomass Function)

1. "It has been shown that under rich growth conditions (i.e. no lack of phosphate and nitrogen), *E. coli* grows in a stoichiometrically optimal manner." (Schilling 2001, Edwards 1994)
2. "It is reasonable to hypothesize that unicellular organisms have evolved toward maximal growth performance." (Segre, 2002.)



Flux Balance Analysis Overview

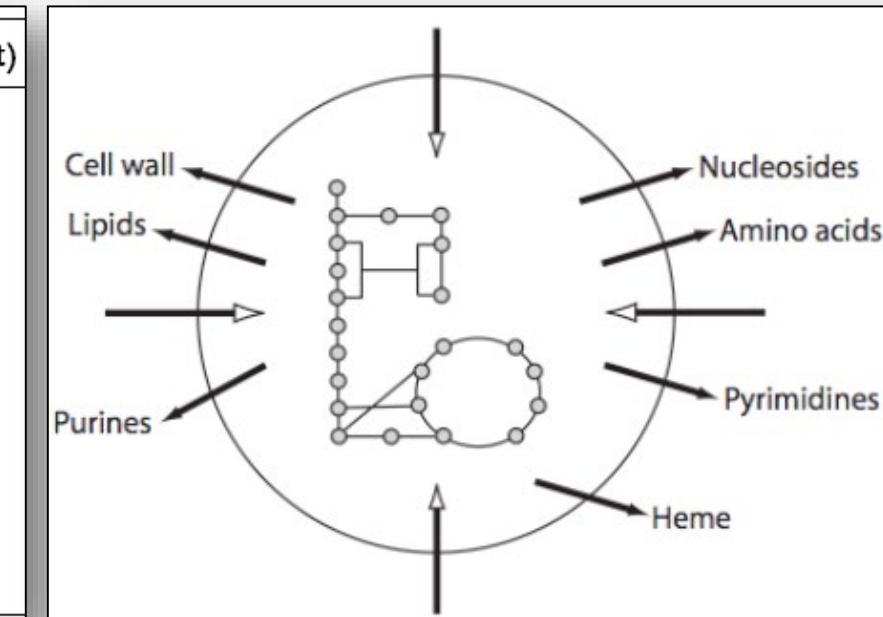
- Flux Balance Analysis Overview
- Reactions, Metabolites, & Pathways
- Mathematical Representation of Reactions & Constraints
- Mass Balanced Linear Equations
- Biomass Reaction
- Calculating Fluxes
- COBRA Toolbox v.3.0



Biomass Precursors

- The biomass reaction accounts for all the fractional contributions from biosynthetic precursors and key cofactors to create 1g of biomass.
- These fractional contributions need to be determined experimentally for cells growing in **log phase**.
- It may not be possible to obtain a detailed biomass composition for the target organism. In this case, one can estimate the relative fraction of each precursor from existing databases.

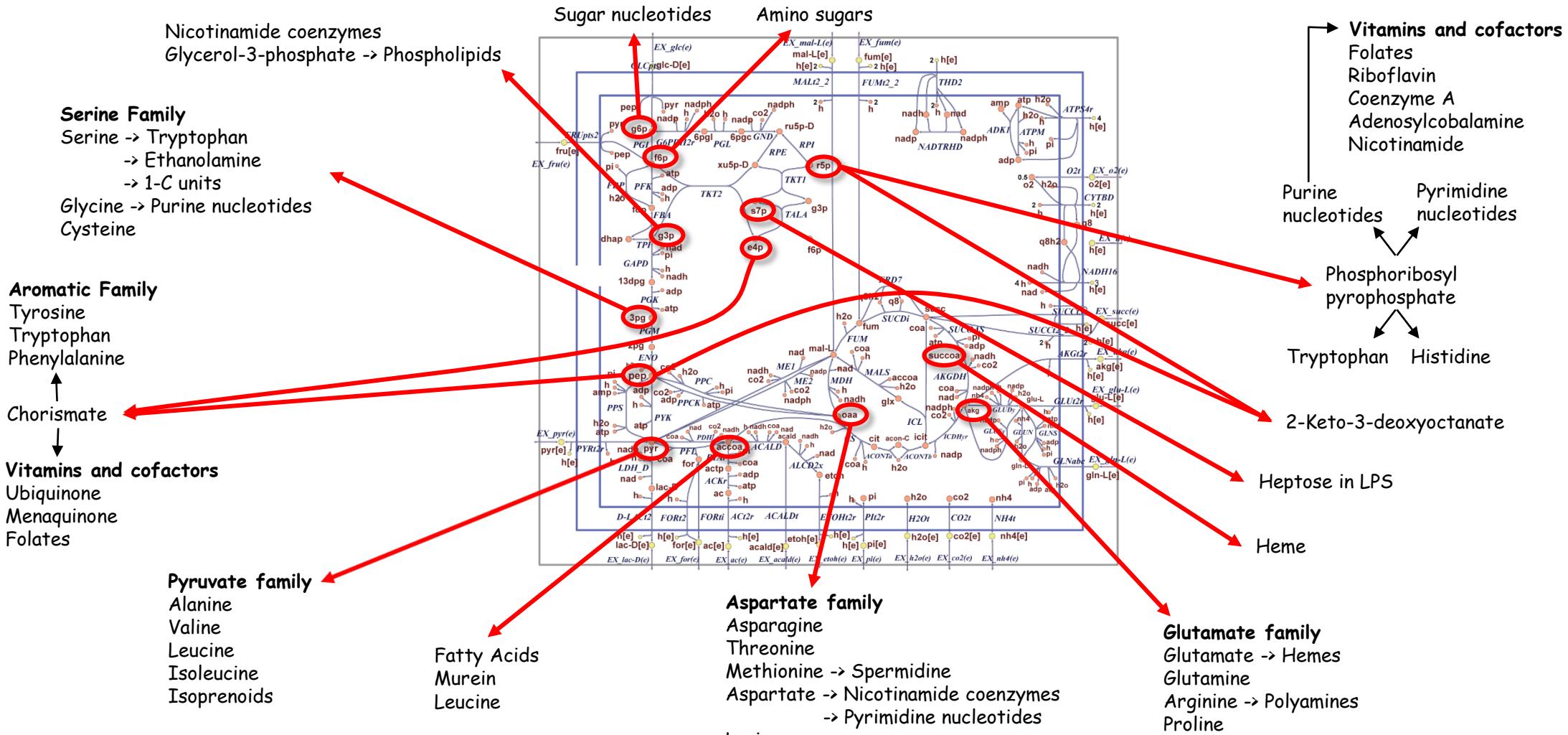
| Cellular component | Cellular content % (wt/wt) |
|--------------------|----------------------------|
| Protein | 55 |
| RNA | 20.5 |
| DNA | 3.1 |
| Lipids | 9.1 |
| Lipopolsaccharides | 3.4 |
| Peptidoglycan | 2.5 |
| Glycogen | 2.5 |
| Polyamines | 0.4 |
| Other | 3.5 |
| Total | 100.00 |



Thiele, I. and B. O. Palsson (2010). "A protocol for generating a high-quality genome-scale metabolic reconstruction." *Nature protocols* 5(1): 93-121.



E.coli Precursor Metabolites





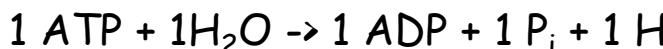
Maintenance Energy Requirements

- To simulate growth, the energy required to maintain the cell growth must be accounted for.
- Two forms of energy are required; growth associated maintenance (GAM) energy and non-growth associated maintenance (NGAM) energy (e.g. turgor pressure).
- GAM reaction accounts for the energy (ATP) necessary to replicate a cell. It is represented in the model by

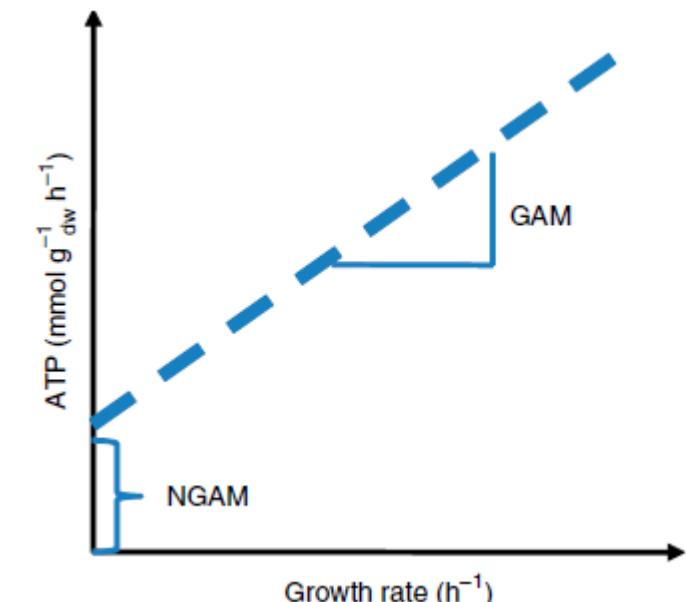


Where x is the number of required phosphate bonds (59.81 in core model). This will be included in the biomass reaction

- The NGAM reaction (ATPM) is given by



where the flux through this reaction is constrained by experimental data to $8.39 \text{ mmol g}_{\text{DW}}^{-1}\text{h}^{-1}$



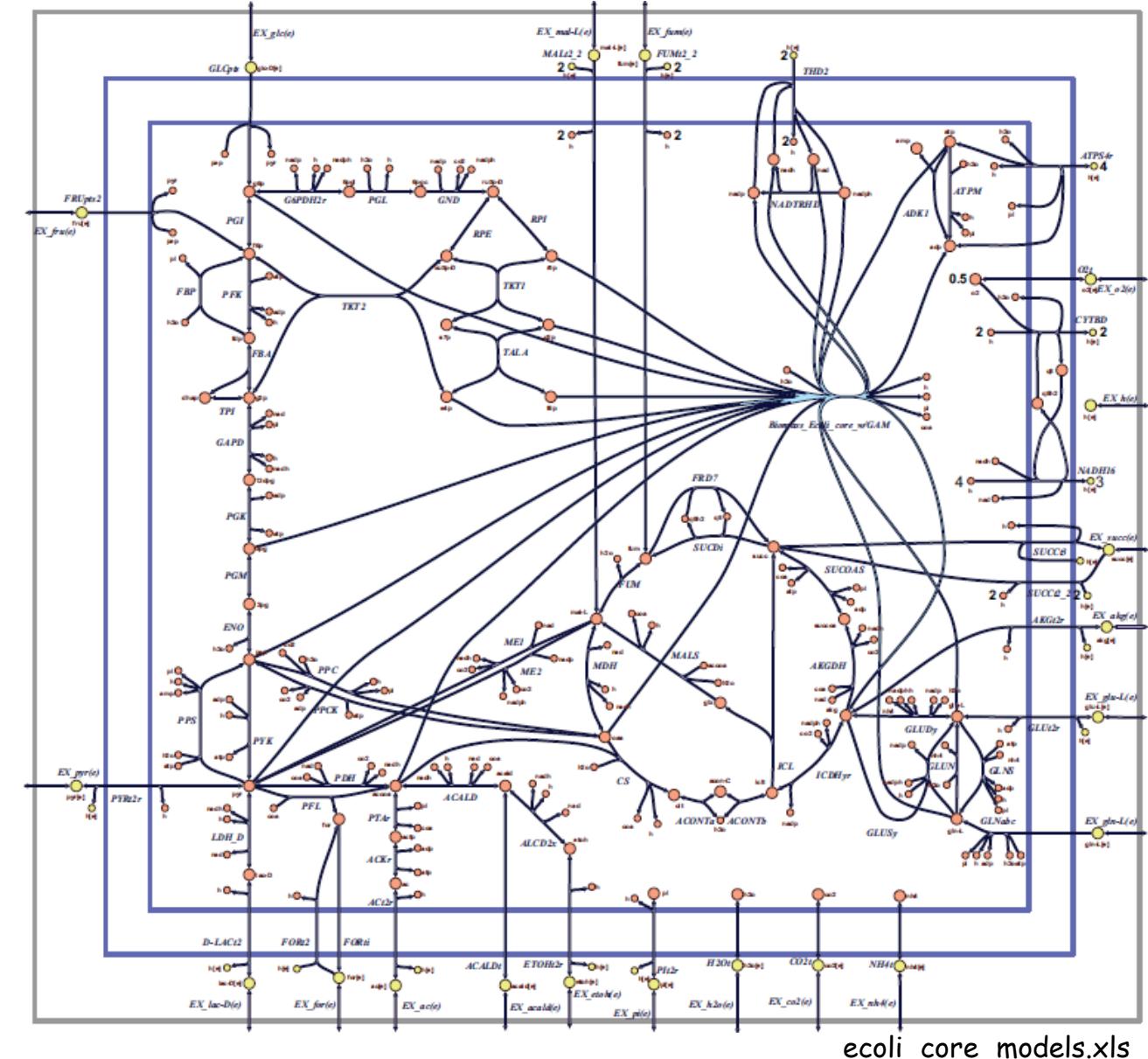
Thiele, I. and B. O. Palsson (2010). "A protocol for generating a high-quality genome-scale metabolic reconstruction." *Nature protocols* 5(1): 93-121.



Biomass Reaction For *E.coli* Core Model

(1.496) 3pg + (3.7478) accoa + (59.8100) atp + (0.3610) e4p + (0.0709) f6p + (0.1290) g3p + (0.2050) g6p + (0.2557) gln-L + (4.9414) glu-L + (59.8100) h2o + (3.5470) nad + (13.0279) nadph + (1.7867) oaa + (0.5191) pep + (2.8328) pyr + (0.8977) r5p --> (59.8100) adp + (4.1182) akg + (3.7478) coa + (59.8100) h + (3.5470) nadh + (13.0279) nadp + (59.8100) pi

* Key Cofactors



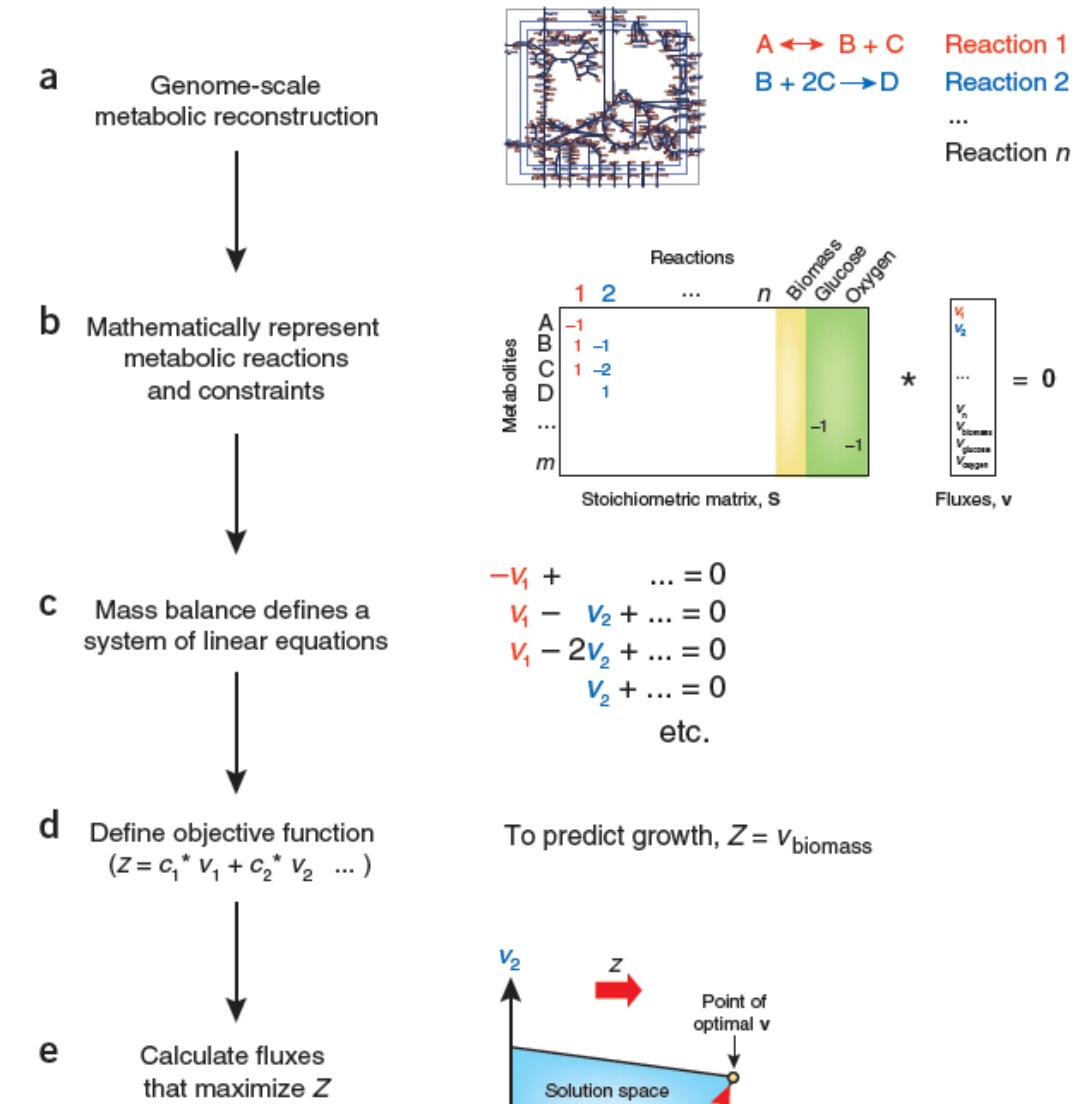


iaf1260 Biomass Objective Function (Ec_biomass_iAF1260_core_59p81M)

$Z = 0.000223 \text{ 10fthf}[c] + 0.000223 \text{ 2ohph}[c] + 0.5137 \text{ ala-L}[c] + 0.000223 \text{ amet}[c] + 0.2958 \text{ arg-L}[c] + 0.2411 \text{ asn-L}[c] + 0.2411 \text{ asp-L}[c] + 59.984 \text{ atp}[c] + 0.004737 \text{ ca2}[c] + 0.004737 \text{ cl}[c] + 0.000576 \text{ coa}[c] + 0.003158 \text{ cobalt2}[c] + 0.1335 \text{ ctp}[c] + 0.003158 \text{ cu2}[c] + 0.09158 \text{ cys-L}[c] + 0.02617 \text{ datp}[c] + 0.02702 \text{ dctp}[c] + 0.02702 \text{ dgtp}[c] + 0.02617 \text{ dtpp}[c] + 0.000223 \text{ fad}[c] + 0.007106 \text{ fe2}[c] + 0.007106 \text{ fe3}[c] + 0.2632 \text{ gln-L}[c] + 0.2632 \text{ glu-L}[c] + 0.6126 \text{ gly}[c] + 0.2151 \text{ gtp}[c] + 54.462 \text{ h2o}[c] + 0.09474 \text{ his-L}[c] + 0.2905 \text{ ile-L}[c] + 0.1776 \text{ k}[c] + 0.01945 \text{ kdo2lipid4}[e] + 0.4505 \text{ leu-L}[c] + 0.3432 \text{ lys-L}[c] + 0.1537 \text{ met-L}[c] + 0.007895 \text{ mg2}[c] + 0.000223 \text{ mlthf}[c] + 0.003158 \text{ mn2}[c] + 0.003158 \text{ mobd}[c] + 0.01389 \text{ murein5px4p}[p] + 0.001831 \text{ nad}[c] + 0.000447 \text{ nadp}[c] + 0.011843 \text{ nh4}[c] + 0.02233 \text{ pe160}[c] + 0.04148 \text{ pe160}[p] + 0.02632 \text{ pe161}[c] + 0.04889 \text{ pe161}[p] + 0.1759 \text{ phe-L}[c] + 0.000223 \text{ pheme}[c] + 0.2211 \text{ pro-L}[c] + 0.000223 \text{ pydx5p}[c] + 0.000223 \text{ ribflv}[c] + 0.2158 \text{ ser-L}[c] + 0.000223 \text{ sheme}[c] + 0.003948 \text{ so4}[c] + 0.000223 \text{ thf}[c] + 0.000223 \text{ thmpp}[c] + 0.2537 \text{ thr-L}[c] + 0.05684 \text{ trp-L}[c] + 0.1379 \text{ tyr-L}[c] + 5.5e-005 \text{ udcpdp}[c] + 0.1441 \text{ utp}[c] + 0.4232 \text{ val-L}[c] + 0.003158 \text{ zn2}[c] \rightarrow 59.81 \text{ adp}[c] + 59.81 \text{ h}[c] + 59.806 \text{ pi}[c] + 0.7739 \text{ ppi}[c]$



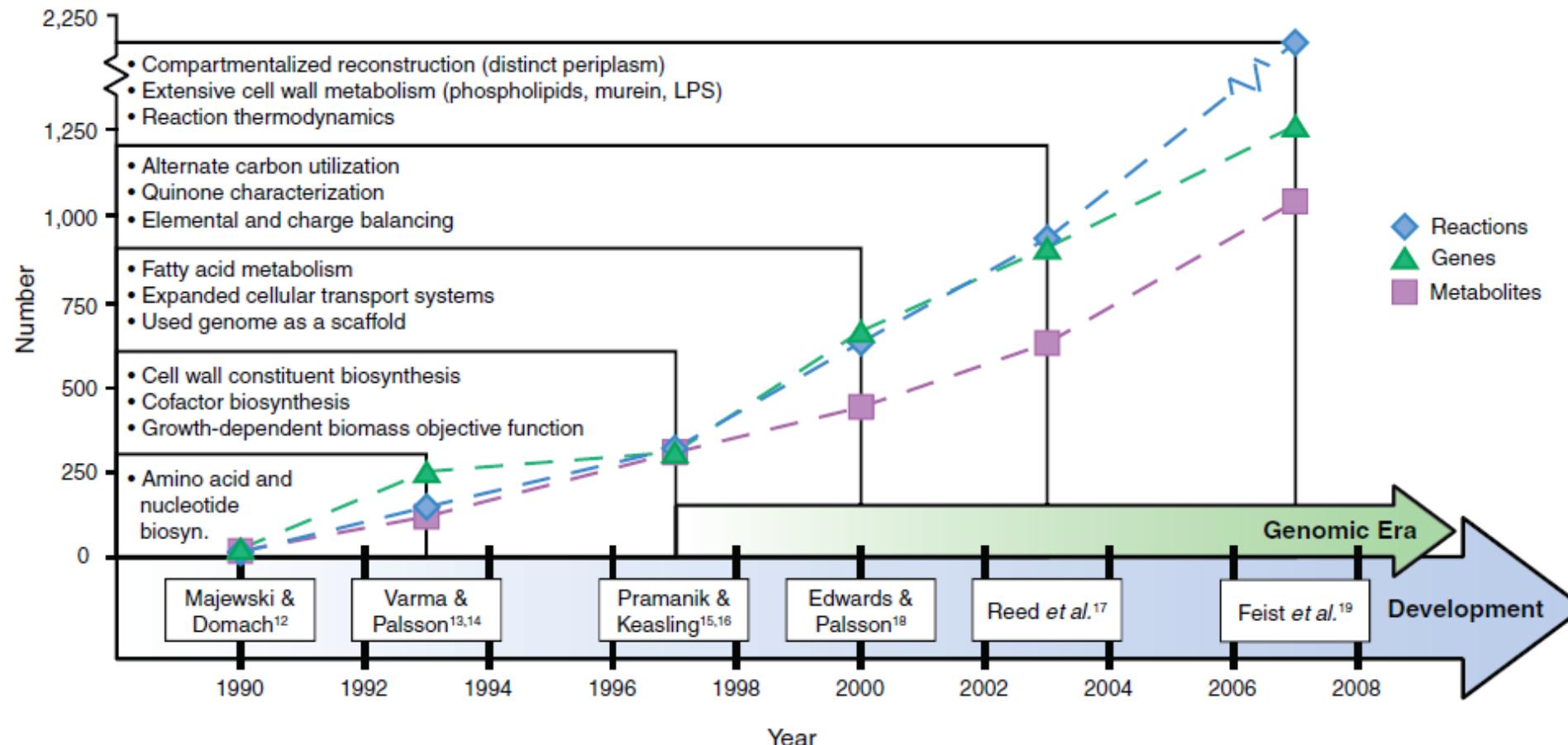
Formulation of Flux Balance Analysis



Orth, J. D., I. Thiele, et al. (2010). "What is flux balance analysis?" *Nature biotechnology* 28(3): 245-248.



The Iterative Reconstruction and History of the *E. Coli* Metabolic Network



Feist, A. M. and B. O. Palsson (2008). "The growing scope of applications of genome-scale metabolic reconstructions using Escherichia coli." *Nature biotechnology* 26(6): 659-667.



E. coli Genome-scale Reconstructions

- Escherichia coli 042
- Escherichia coli 536
- Escherichia coli 55989
- Escherichia coli ABU 83972
- Escherichia coli APEC O1
- Escherichia coli ATCC 8739
- Escherichia coli B str. REL606
- Escherichia coli BL21(DE3) AM946981
- Escherichia coli BL21(DE3) BL21-Gold(DE3)pLySS AG
- Escherichia coli BL21(DE3) CP001509
- Escherichia coli BW2952
- Escherichia coli CFT073
- Escherichia coli DH1
- Escherichia coli DH1 ME8569
- Escherichia coli E24377A
- Escherichia coli ED1a
- Escherichia coli ETEC H10407
- Escherichia coli HS
- Escherichia coli IAI1
- Escherichia coli IAI39
- Escherichia coli IHE3034
- Escherichia coli KO11FL
- Escherichia coli LF82
- Escherichia coli NA114
- Escherichia coli O103:H2 str. 12009
- Escherichia coli O111:H- str. 11128
- Escherichia coli O127:H6 str. E2348/69
- Escherichia coli O157:H7 EDL933
- Escherichia coli O157:H7 str. EC4115
- Escherichia coli O157:H7 str. Sakai
- Escherichia coli O157:H7 str. TW14359
- Escherichia coli O26:H11 str. 11368
- Escherichia coli O55:H7 str. CB9615
- Escherichia coli O83:H1 str. NRG 857C
- Escherichia coli S88
- Escherichia coli SE11
- Escherichia coli SE15
- Escherichia coli SMS-3-5
- Escherichia coli str. K-12 substr. DH10B
- Escherichia coli str. K-12 substr. MG1655
- Escherichia coli str. K-12 substr. W3110
- Escherichia coli UM146
- Escherichia coli UMN026
- Escherichia coli UMNK88
- Escherichia coli UTI89
- Escherichia coli W
- Escherichia coli W CP002185
- Escherichia coli K-12 MG1655

Monk, J. M., P. Charusanti, et al. (2013). Proceedings of the National Academy of Sciences of the United States of America 110(50): 20338-20343.

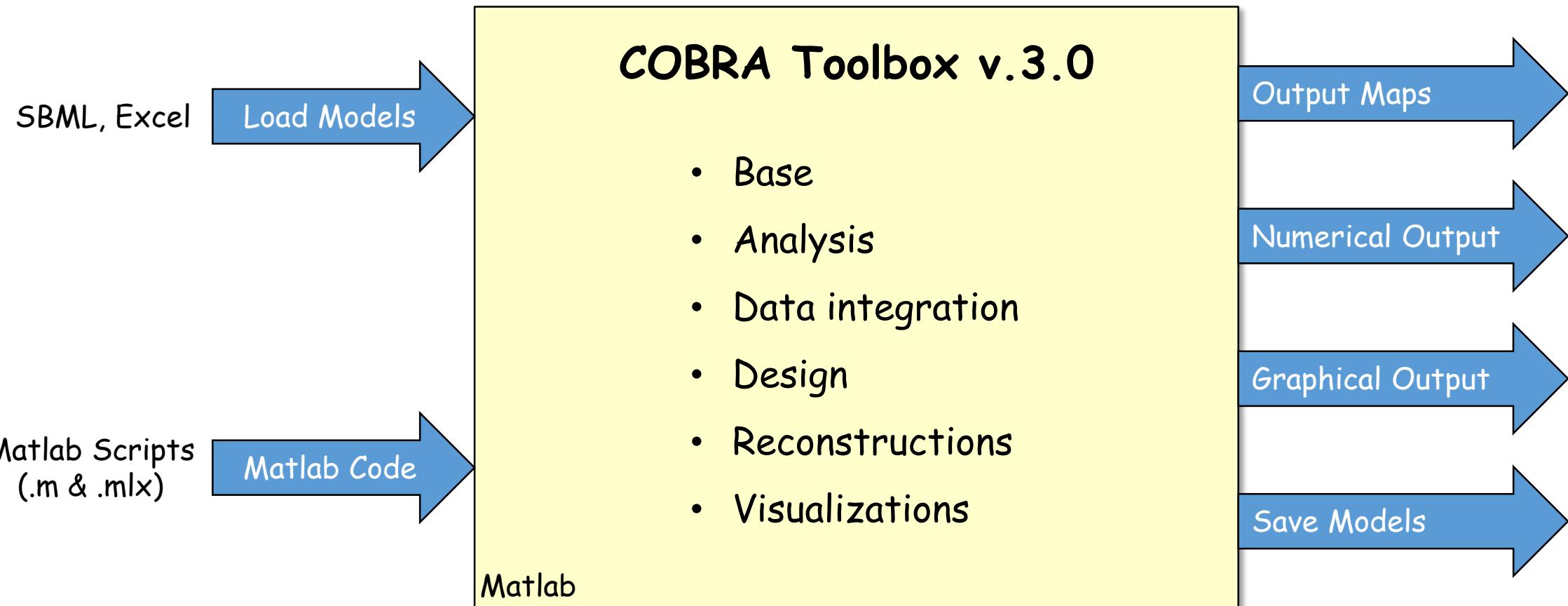


Flux Balance Analysis Overview

- Flux Balance Analysis Overview
- Reactions, Metabolites, & Pathways
- Mathematical Representation of Reactions & Constraints
- Mass Balanced Linear Equations
- Biomass Reaction
- • Calculating Fluxes
- COBRA Toolbox v.3.0



COBRA Toolbox v.3.0 Overview



Laurent Heirendt et al, Creation and analysis of biochemical constraint-based models: the COBRA Toolbox v3.0, Nature Protocols, volume 14, pages 639-702, 2019



Matlab Interface

The screenshot shows the MATLAB 7.12.0 (R2011a) interface with the following windows open:

- Variable Editor - solution**: Displays a table of variables from a struct named "solution".

| Field | Value | Min | Max |
|----------|---------------|-----------|---------|
| x | <95x1 double> | -29.17... | 45.5140 |
| f | 0.8739 | 0.8739 | 0.8739 |
| y | <1x1 struct> | | |
| w | [] | | |
| stat | 1 | 1 | 1 |
| origStat | 2 | 2 | 2 |
| ab | 'gurobi' | | |
| solver | 'gurobi' | | |
| time | 0.1180 | 0.1180 | 0.1180 |
- Workspace**: Shows the current workspace variables.
- Command History**: Shows the command history for the session.
- Current Folder**: Shows the contents of the "Tutorial" folder.
- Script Editor**: Displays the script "AerobicGlucoseBioMass.m" which reads a constraint-based model and performs optimization.



Print Flux Values

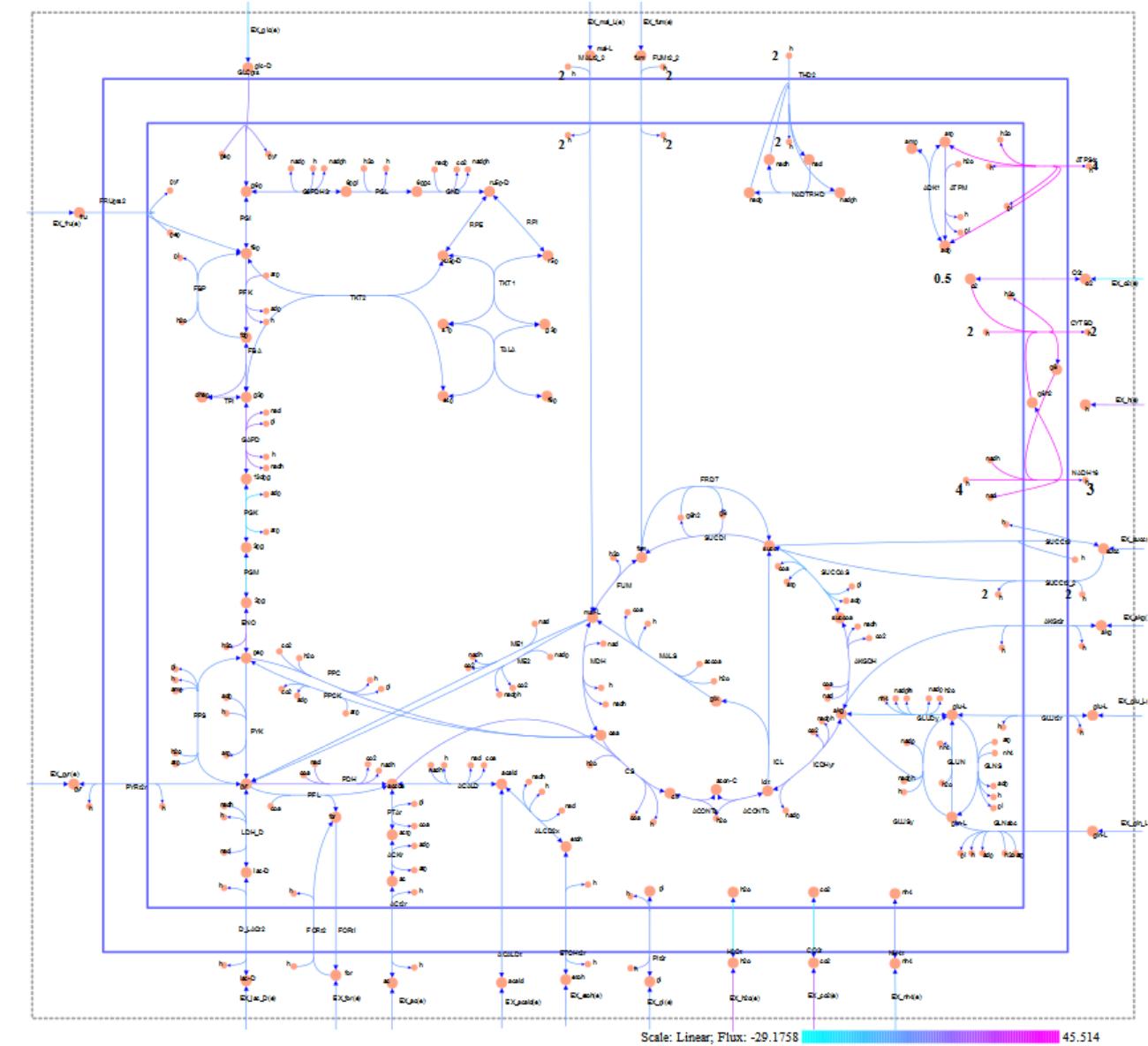
| | | | | | |
|-----------|-------------|---------|----------|--------|----------|
| ACONTa | 6.00725 | FBA | 7.47738 | PGK | -16.0235 |
| ACONTb | 6.00725 | FUM | 5.06438 | PGL | 4.95998 |
| AKGDH | 5.06438 | G6PDH2r | 4.95998 | PGM | -14.7161 |
| ATPM | 8.39 | GAPD | 16.0235 | PIt2r | 3.2149 |
| ATPS4r | 45.514 | GLCpts | 10 | PPC | 2.50431 |
| Biomass | ...0.873922 | GLNS | 0.223462 | PYK | 1.75818 |
| CO2t | -22.8098 | GLUDy | -4.54186 | RPE | 2.67848 |
| CS | 6.00725 | GND | 4.95998 | RPI | -2.2815 |
| CYTBD | 43.599 | H2Ot | -29.1758 | SUCDi | 5.06438 |
| ENO | 14.7161 | ICDHyr | 6.00725 | SUCOAS | -5.06438 |
| EX_co2(e) | 22.8098 | MDH | 5.06438 | TALA | 1.49698 |
| EX_glc(e) | -10 | NADH16 | 38.5346 | TKT1 | 1.49698 |
| EX_h2o(e) | 29.1758 | NH4t | 4.76532 | TKT2 | 1.1815 |
| EX_h(e) | 17.5309 | O2t | 21.7995 | TPI | 7.47738 |
| EX_nh4(e) | -4.76532 | PDH | 9.28253 | | |
| EX_o2(e) | -21.7995 | PFK | 7.47738 | | |
| EX_pi(e) | -3.2149 | PGI | 4.86086 | | |

Growth Rate

Inputs & Outputs
(Exchange Reactions)



Drawing Flux Values on a Map

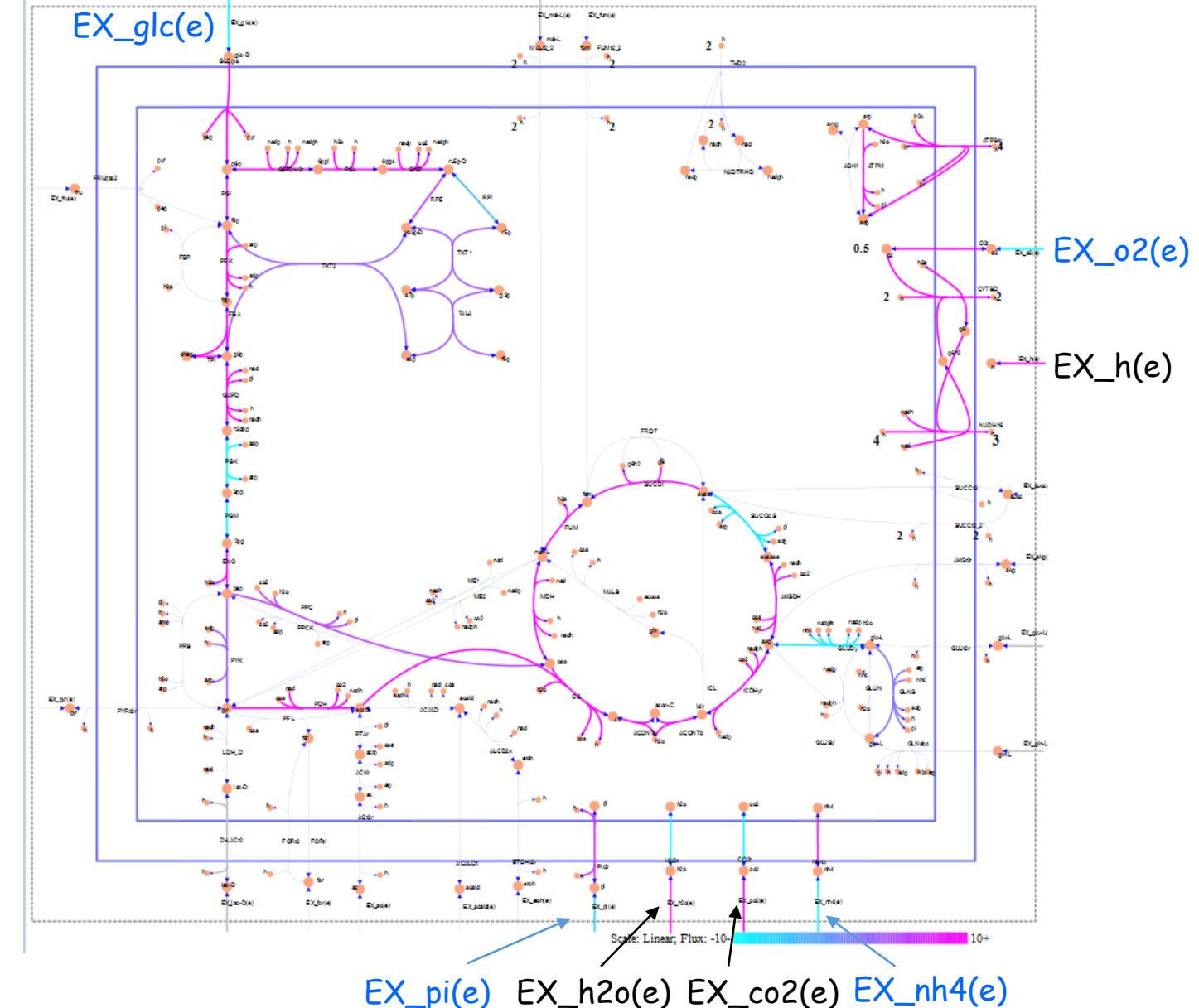




Aerobic Growth on Glucose

Exchange Reactions

| | |
|-----------|----------|
| EX_co2(e) | 40.6527 |
| EX_glc(e) | -18.5 |
| EX_h2o(e) | 52.6943 |
| EX_h(e) | 33.1606 |
| EX_nh4(e) | -9.01387 |
| EX_o2(e) | -38.7416 |
| EX_pi(e) | -6.08116 |

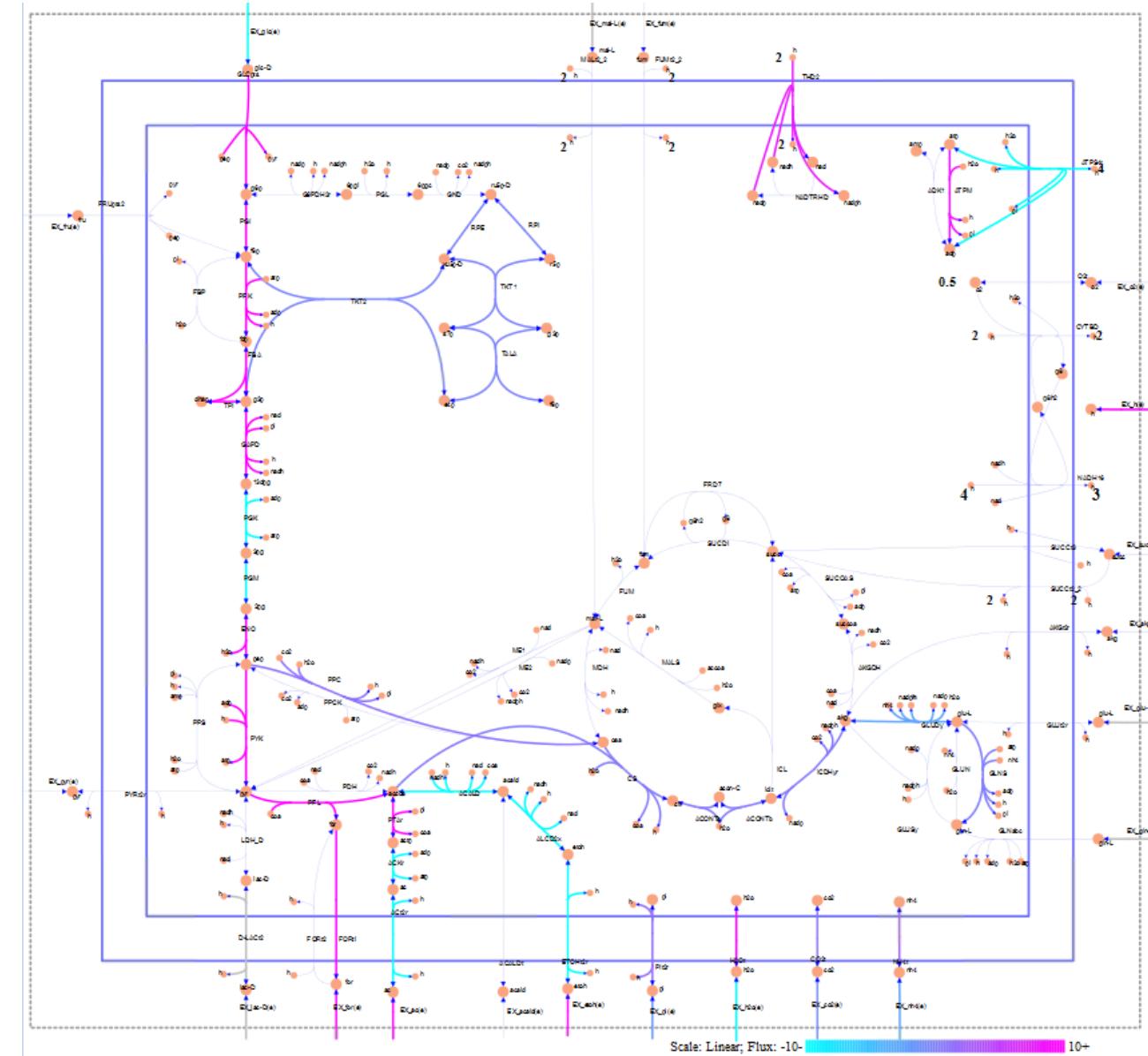




Anaerobic Growth on Glucose

Exchange Reactions

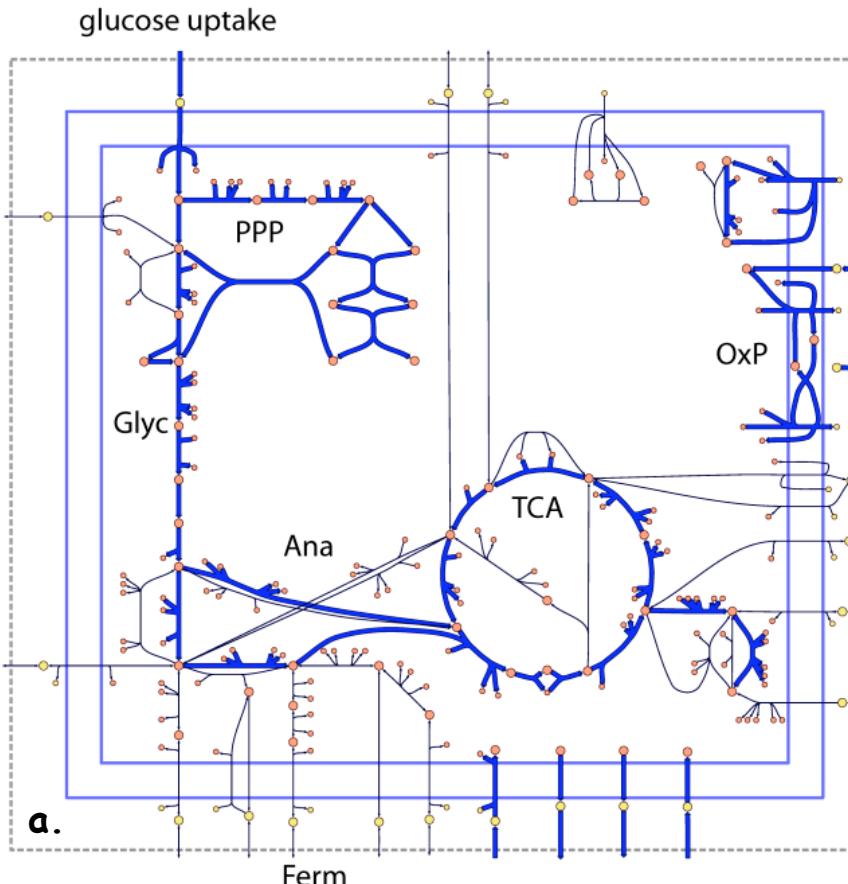
| | |
|-----------|-----------|
| Biomass | 0.470565 |
| EX_ac(e) | 15.1732 |
| EX_co2(e) | -0.840759 |
| EX_eto(h) | 14.6749 |
| EX_for(e) | 32.1194 |
| EX_glc(e) | -18.5 |
| EX_h2o(e) | -12.0879 |
| EX_h(e) | 56.7321 |
| EX_nh4(e) | -2.5659 |
| EX_pi(e) | -1.73107 |



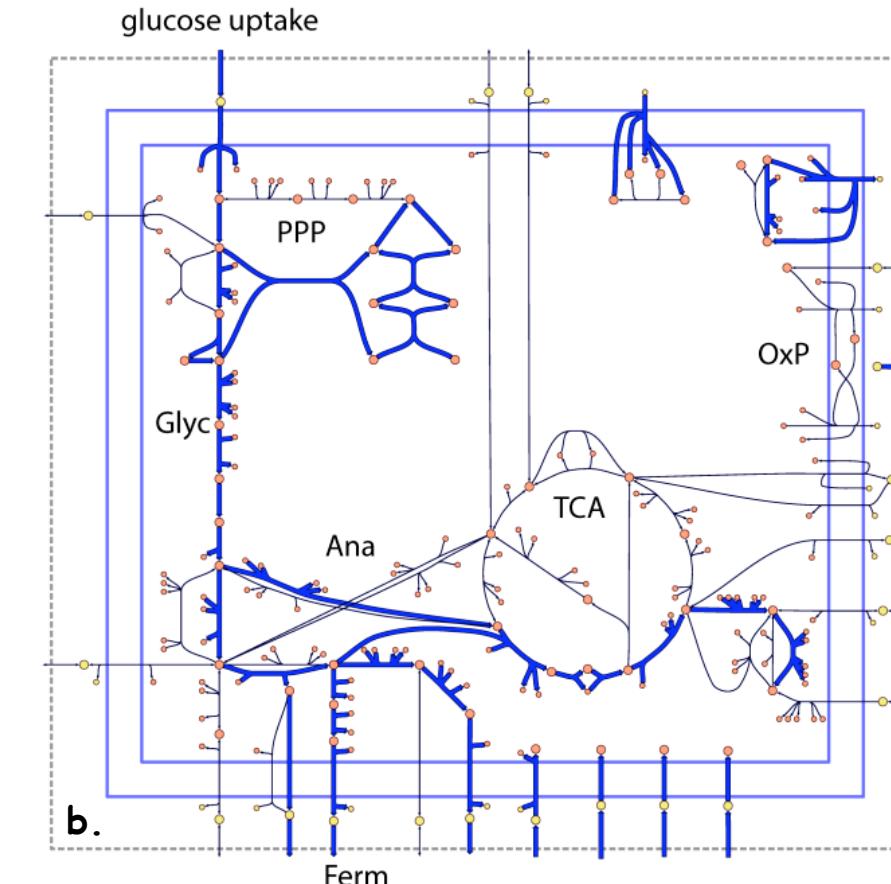


Aerobic vs. Anaerobic Growth

Orth, J. D., I. Thiele, et al. (2010). "What is flux balance analysis?" Nature biotechnology 28(3): 245-248.



Aerobic Growth



Anaerobic Growth



Substrate Maximum Growth Rate

| Substrate | Aerobic (hr ⁻¹) | Anaerobic (hr ⁻¹) |
|----------------|-----------------------------|-------------------------------|
| acetate | 0.3893 | 0 |
| acetaldehyde | 0.6073 | 0 |
| 2-oxoglutarate | 1.0982 | 0 |
| ethanol | 0.6996 | 0 |
| D-fructose | 1.7906 | 0.5163 |
| fumarate | 0.7865 | 0 |
| D-glucose | 1.7906 | 0.5163 |
| L-glutamine | 1.1636 | 0 |
| L-glutamate | 1.2425 | 0 |
| D-lactate | 0.7403 | 0 |
| L-malate | 0.7865 | 0 |
| pyruvate | 0.6221 | 0.0655 |
| succinate | 0.8401 | 0 |

The core *E. coli* model contains exchange reactions for 13 different organic compounds, each of which can be used as the sole carbon source under aerobic or anaerobic conditions.

("What is flux balance analysis? - Supplementary tutorial")



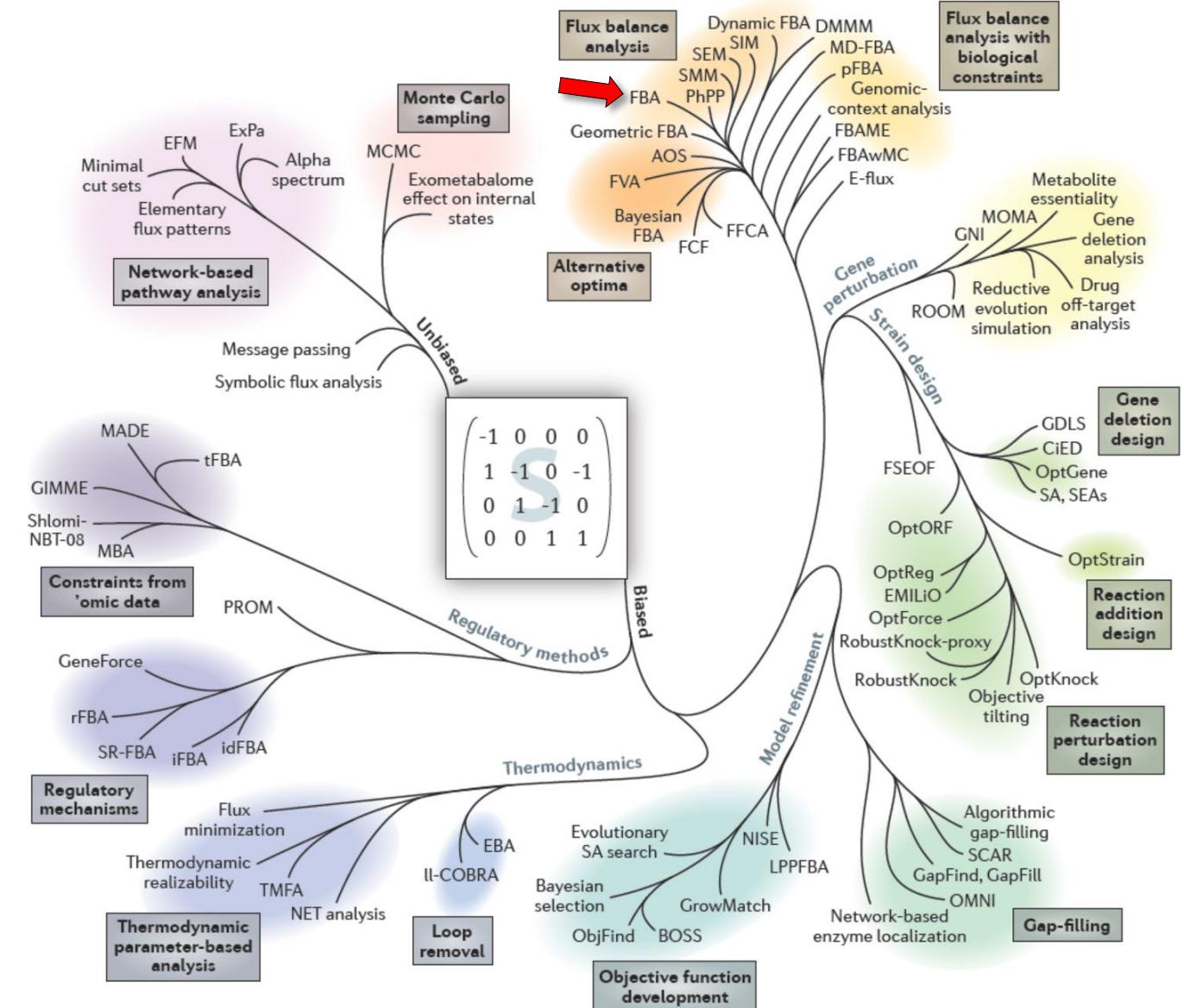
Flux Balance Analysis Overview

- Flux Balance Analysis Overview
- Reactions, Metabolites, & Pathways
- Mathematical Representation of Reactions & Constraints
- Mass Balanced Linear Equations
- Biomass Reaction
- Calculating Fluxes
- • COBRA Toolbox v.3.0



Methods in Constraint-based Reconstruction and Analysis

Lewis, N. E., H. Nagarajan, et al. (2012). "Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods." *Nature reviews Microbiology* 10(4): 291-305.





Flux Balance Analysis Overview

- Flux Balance Analysis Overview
- Reactions, Metabolites, & Pathways
- Mathematical Representation of Reactions & Constraints
- Mass Balanced Linear Equations
- Biomass Reaction
- Calculating Fluxes
- COBRA Toolbox v.3.0



Learning Objectives

Each student should be able to:

- Explain flux balance analysis (FBA).
- Explain reactions, metabolites, & pathways.
- Explain mass balanced linear equations.
- Explain the biomass reaction.
- Explain how to create a stoichiometric matrix from reactions and metabolites.
- Explain gene-protein-reaction associations.
- Explain the constraint-based modeling.



Reflective Questions

1. What is flux balance analysis?
2. What does steady state mean in flux balance analysis (FBA)?
3. What is the difference between a reaction, a metabolite, and a pathway?
4. What is a gene-protein-reaction (GPR)?
5. What is a gene locus?
6. What is the difference between a single enzyme, an isozymes, and a protein complex?
7. What is a reaction formula?
8. What is the difference between a metabolite neutral formula and a metabolite charged formula?
9. What is a metabolite compartment?
10. What is an exchange reaction?
11. What are the units of flux?
12. What is the mathematical sign for uptake and secretion?
13. What is the difference between extracellular environment and intracellular space?
14. What are transport reactions?
15. Are the rows of a stoichiometric matrix metabolites or reactions?
16. Why is the product of the stoichiometric matrix and the flux set to zero?
17. What is the purpose of linear programming?
18. What is the purpose on an objective function?
19. What are biomass precursors?
20. What is the purpose of growth associated maintenance (GAM)?
21. What is the purpose of non-growth associated maintenance energy (NGAM)?
22. What is the purpose of the biomass reaction?
23. What are genome-scale metabolic network reconstruction?
24. What phase of growth does FBA assume (lag, exponential, stationary, death)?
25. What are the limits of FBA?
26. What role does the stoichiometric matrix play in FBA?
27. Why are visualization tools needed?



References

1. Monk, J., J. Nogales, et al. (2014). "Optimizing genome-scale network reconstructions." *Nature biotechnology* 32(5): 447-452.
2. Monk, J. and B. O. Palsson (2014). "Genetics. Predicting microbial growth." *Science* 344(6191): 1448-1449.
3. McCloskey, D., B. O. Palsson, et al. (2013). "Basic and applied uses of genome-scale metabolic network reconstructions of *Escherichia coli*." *Molecular Systems Biology* 9: 661.
4. Orth, J. D., I. Thiele, et al. (2010). "What is flux balance analysis?" *Nature biotechnology* 28(3): 245-248.
5. Reconstruction and Use of Microbial Metabolic Networks: the Core *Escherichia coli* Metabolic Model as an Educational Guide by Orth, Fleming, and Palsson (2010)
6. Oberhardt, M. A., B. O. Palsson, et al. (2009). "Applications of genome-scale metabolic reconstructions." *Molecular Systems Biology* 5: 320.
7. Palsson, B. (2009). "Metabolic systems biology." *FEBS letters* 583(24): 3900-3904.
8. Adam M. Feist, Ines Thiele, and Bernhard Ø. Palsson, Chapter 9: Genome-Scale Reconstruction, Modeling, and Simulation of *E. coli*'s Metabolic Network, S.Y. Lee (ed.), *Systems Biology and Biotechnology of Escherichia coli*, DOI 10.1007/978-1-4020-9394-4_9, Springer Science+Business Media B.V. 2009.
9. Feist, A. M. and B. O. Palsson (2008). "The growing scope of applications of genome-scale metabolic reconstructions using *Escherichia coli*." *Nature biotechnology* 26(6): 659-667.
10. Reed, J. L., I. Famili, et al. (2006). "Towards multidimensional genome annotation." *Nature reviews. Genetics* 7(2): 130-141.
11. Price, N. D., J. L. Reed, et al. (2004). "Genome-scale models of microbial cells: evaluating the consequences of constraints." *Nature reviews. Microbiology* 2(11): 886-897.
12. Price, N. D., J. A. Papin, et al. (2003). "Genome-scale microbial in silico models: the constraint-based approach." *Trends in biotechnology* 21(4): 162-169
13. Reed, J. L. and B. O. Palsson (2003). "Thirteen years of building constraint-based in silico models of *Escherichia coli*." *Journal of Bacteriology* 185(9): 2692-2699.
14. Covert, M. W., I. Famili, et al. (2003). "Identifying constraints that govern cell behavior: a key to converting conceptual to computational models in biology?" *Biotechnology and bioengineering* 84(7): 763-772.