

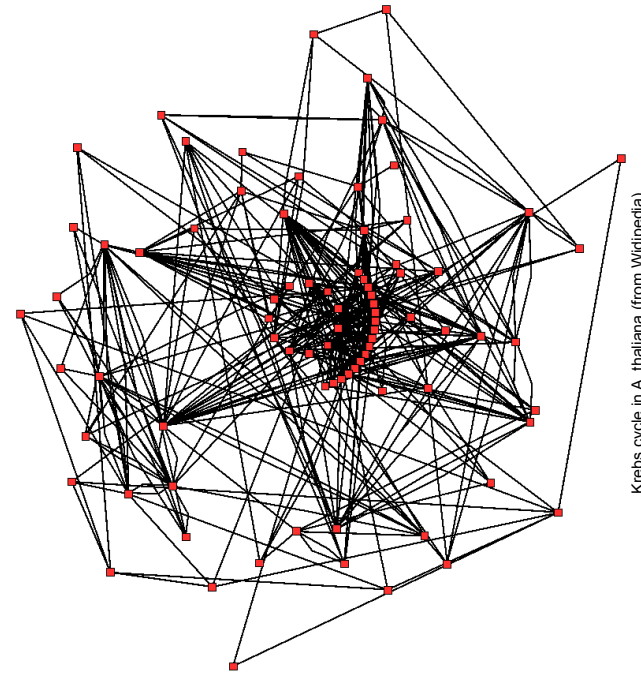
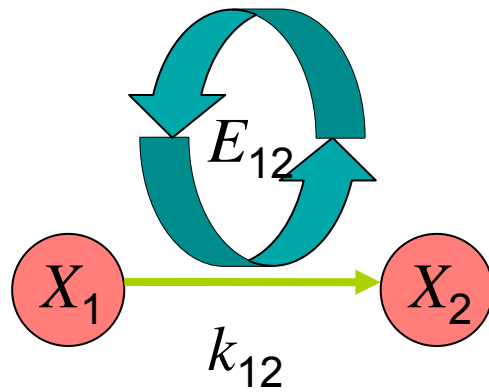
Using Pathway Tools & Matlab for Flux Balance Analysis

Kent Peterson
25 Aug. 2009

Summary

- Flux Balance Analysis (FBA)
- Workflow: from Ptools to Matlab, and back again
- A simple example
- What next?

Let's start at the very beginning ...



Bio Chemical Reaction Network

Systems biology: Stoichiometry and kinetics

Matrix of stoichiometric weights (S)

$$S_{ij} = \frac{\dot{x}_i}{v_j} \Big|_{v_{k \neq j} = 0}$$

Matrix of linearized kinetics ($D_x r$)

$$D_x r_{ij} = \frac{r_i}{x_j} \Big|_{x_{k \neq j} = 0}$$

Chemical reactions

	1	j	n
1	of each pool in the j^{th} reaction		
i			
m	Stoichiometric weightings		

Chemical pools

Rate of chemical reactions

	1	j	m
1	of each pool on the j^{th} reaction rate		
i			
m	Differential effect		

Chemical pools

Systems biology:

Stoichiometric (a.k.a. kinematic, structural) constraints

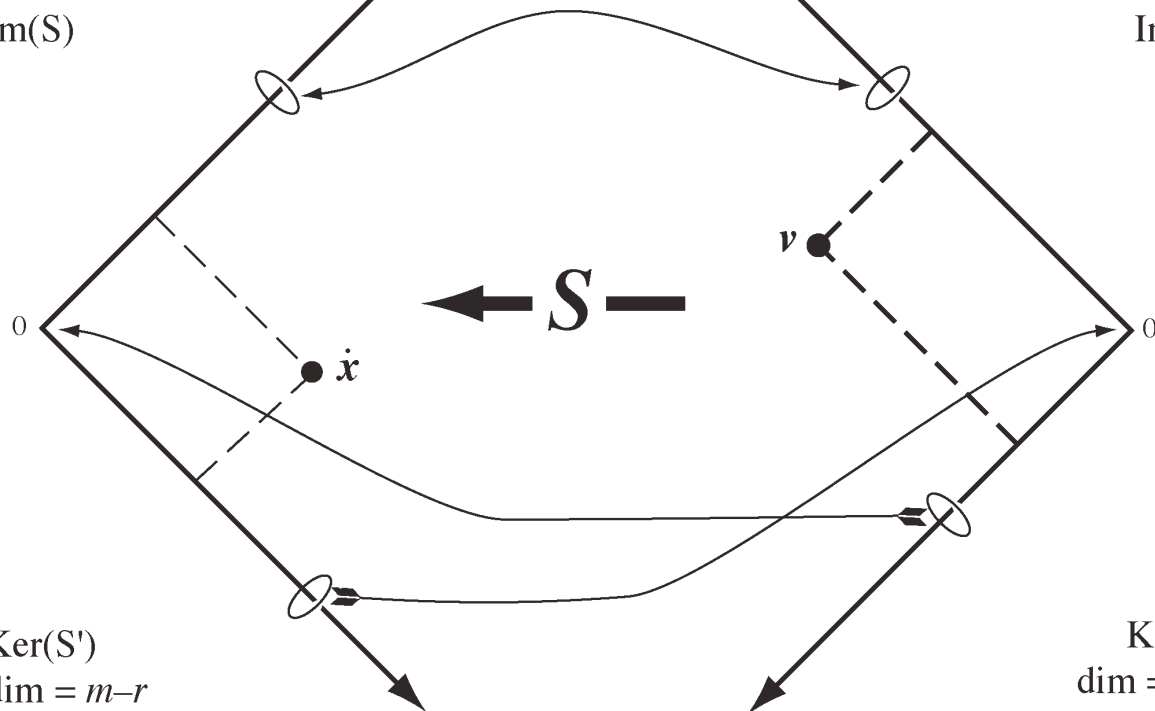
**Chemical pool
growth rates, \dot{x}_i**

$$[\dot{x}_i] = [S_{ij}] [v_j]$$

**Chemical reaction
velocities, v_j**

Modifiable pool "mixtures"
Column space
dim = r
Im(S)

Pool-varying flux "patterns"
Row space
dim = r
Im(S')



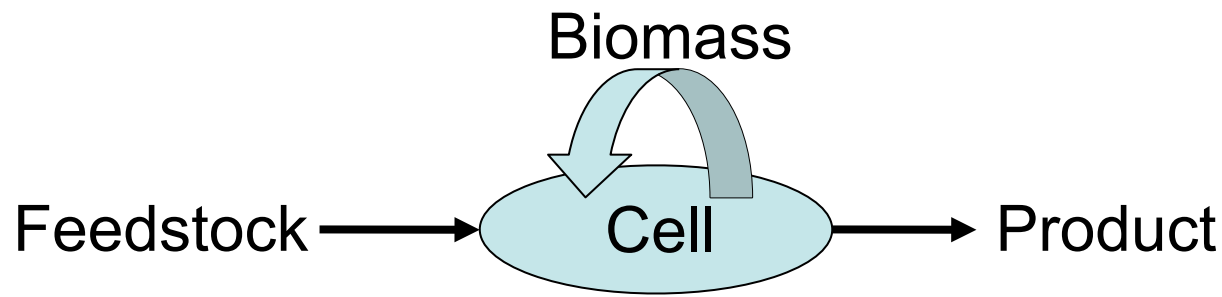
$\text{Ker}(S')$
dim = $m-r$
Left null space
Conserved pool "mixtures"

$\text{Ker}(S)$
dim = $n-r$
Right null space
Steady-state flux "patterns"

Flux Balance Analysis (FBA)

- Flux = “rate at which chemicals are transformed or transported”
- Balanced = “input & output rates are equal” (a.k.a. steady-state)
- Analysis = “optimize (balanced) flux with constraints”

- Example: maximize product synthesis per unit feedstock, while satisfying maintenance (biomass) requirements



Software components

- Pathway Tools¹ (PGDB for organism)
- SBML² Toolbox (export metabolic description files)
- COBRA³ (setup the LP)
- GLPK⁴ (solve the LP)
- Matlab⁵ (math's programming & graphics)

1. <http://bioinformatics.ai.sri.com/ptools/>

2. <http://sbml.org/Software/SBMLToolbox>

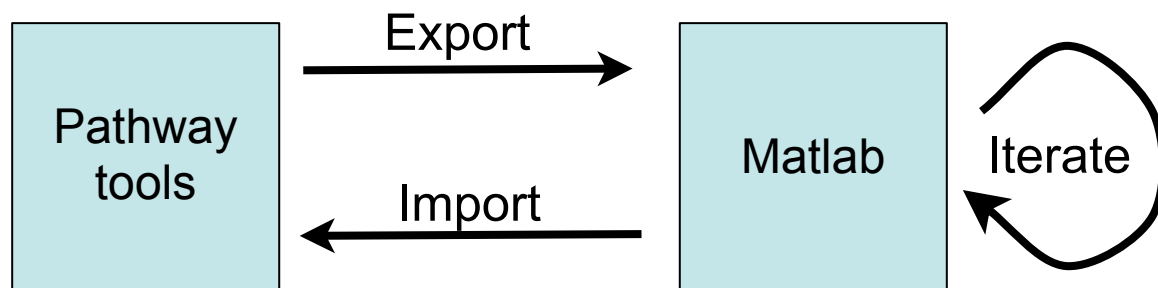
3. Palsson lab @ UCSD, http://gcrq.ucsd.edu/Downloads/Cobra_Toolbox

4. GNU Linear Programming Kit, <http://glpkmex.sourceforge.net/>

5. <http://www.mathworks.com/>

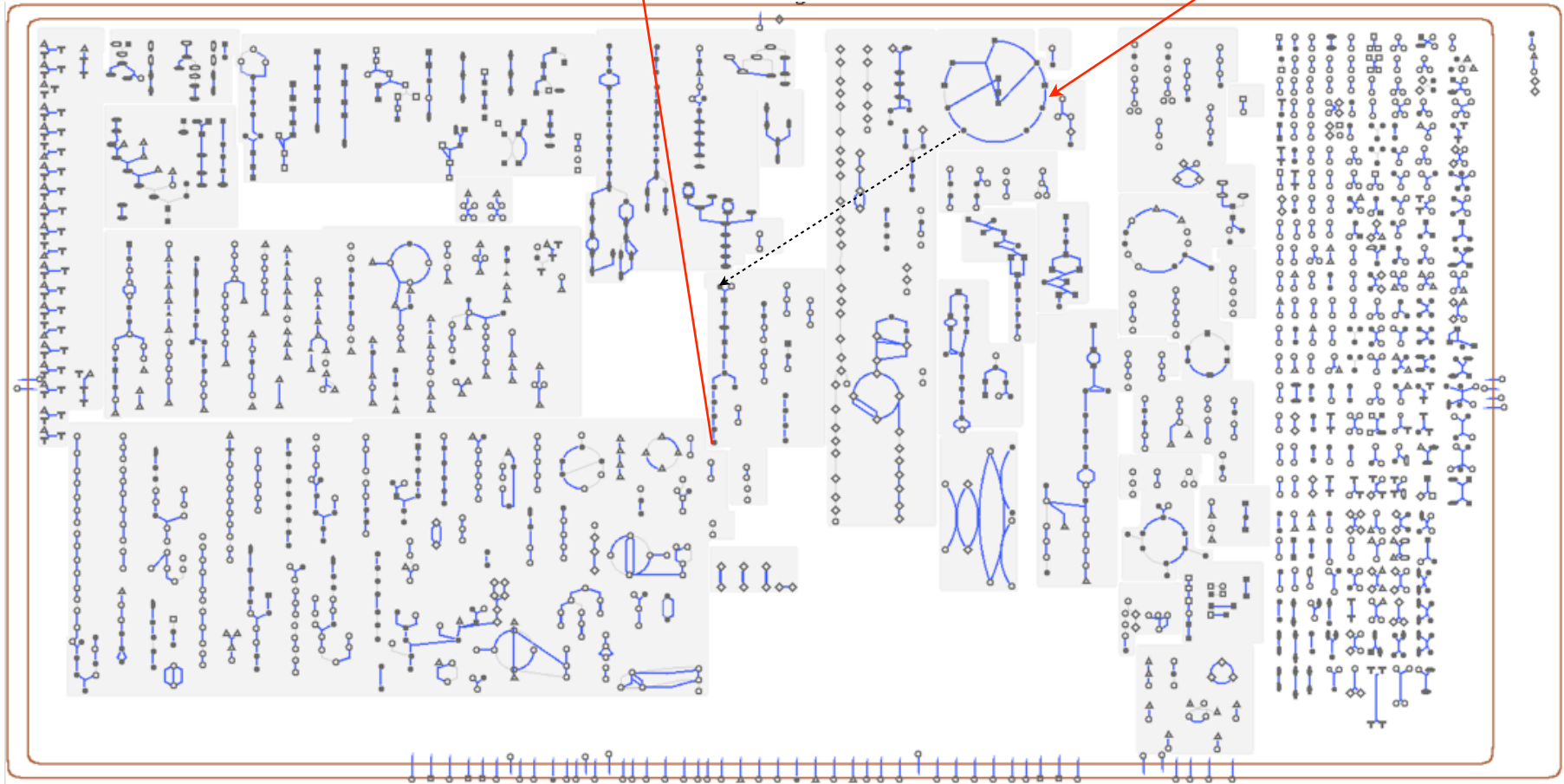
Workflow

- Load PGDB into PathwayTools
- Export metabolic network into SBML file
- Import SBML file into Matlab
- Compute FBA solution in Matlab
- Examine network & flux sol'n, using Matlab GUIs
- Iterate
- Write FBA sol'n to text file
- Import sol'n to Pathway Tools Omics Viewer



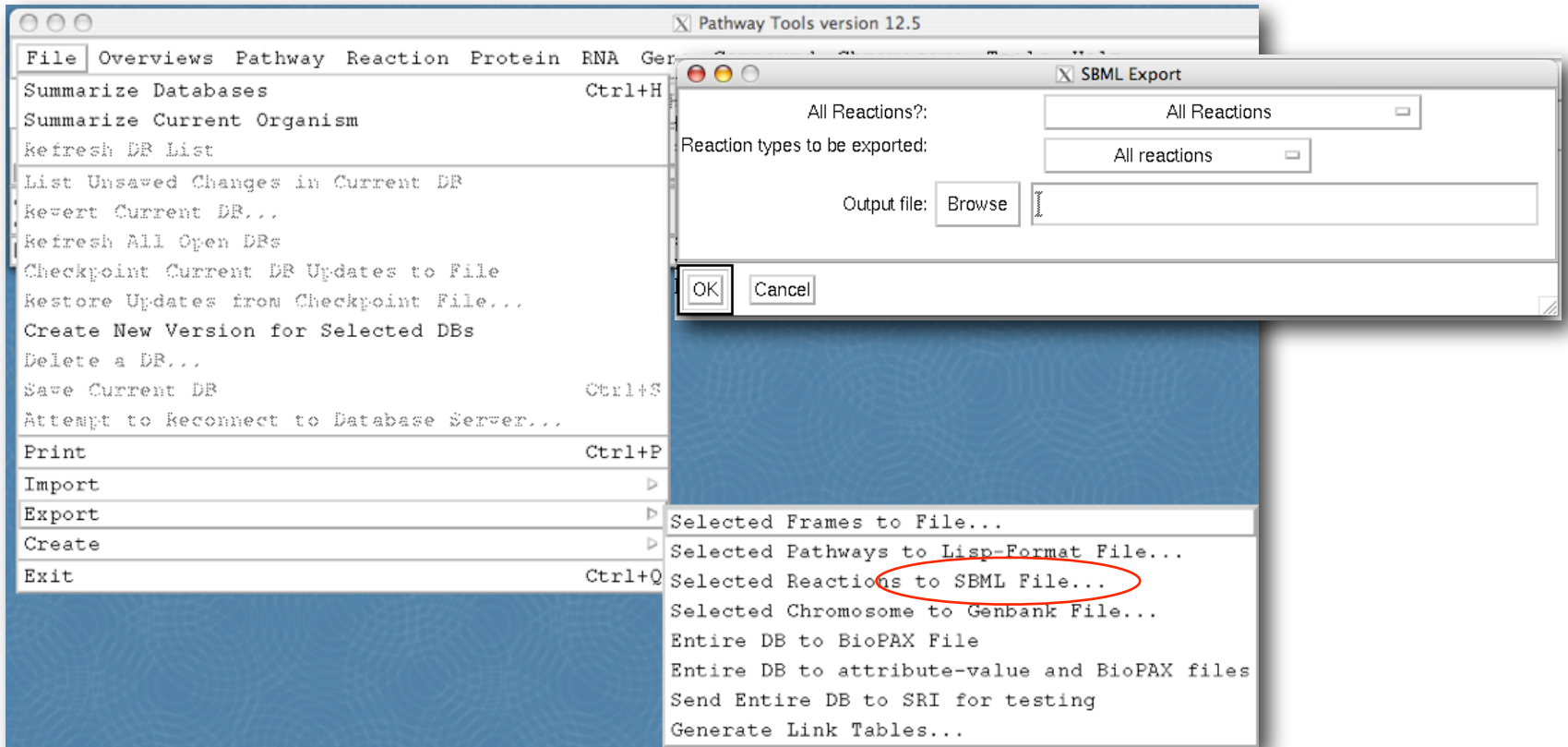
Example

Maximize anabolic “biomass” produced from a “feedstock”



Getting the data from Ptools to Matlab

- SBML export from Pathway Tools

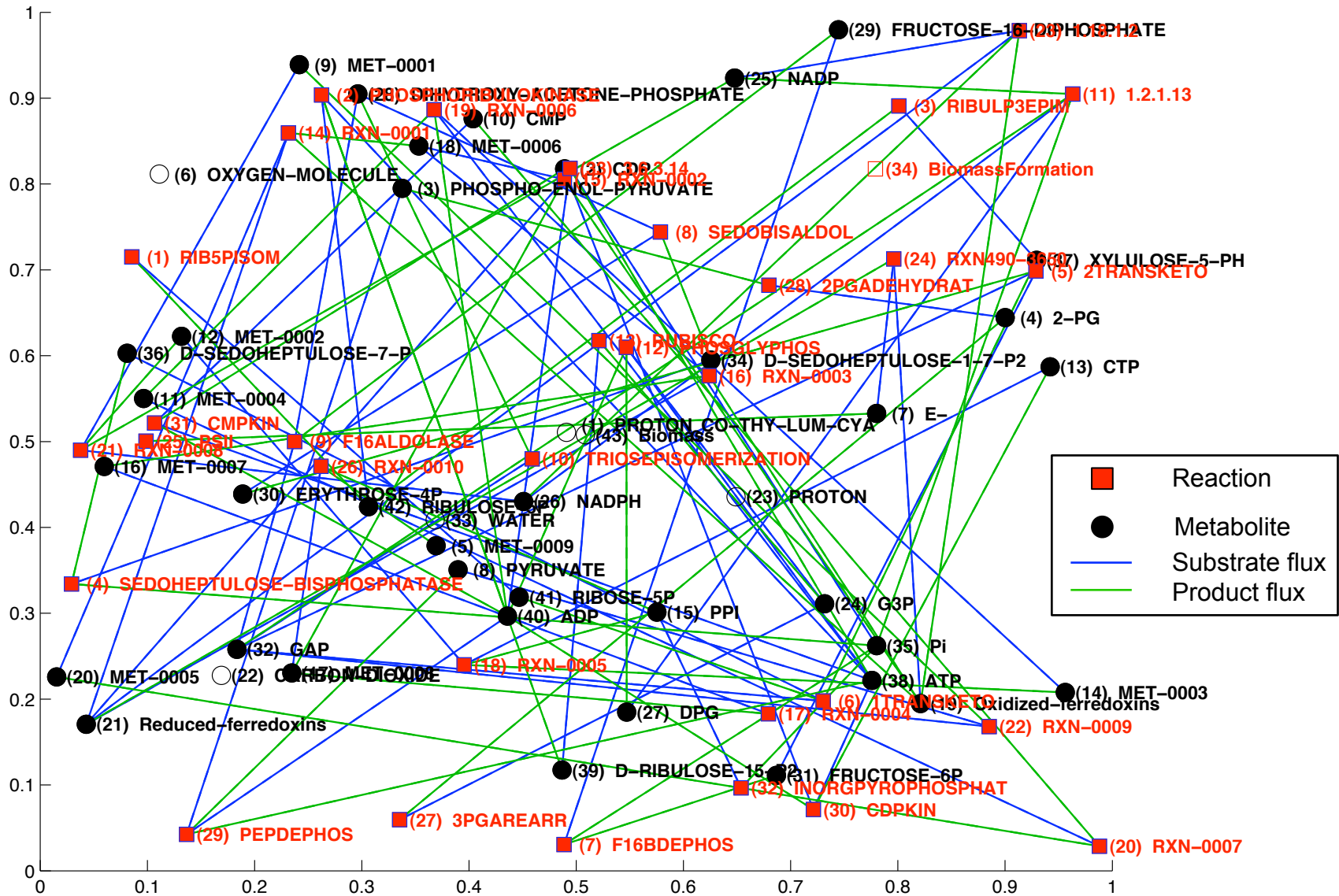


- SBML import into Matlab

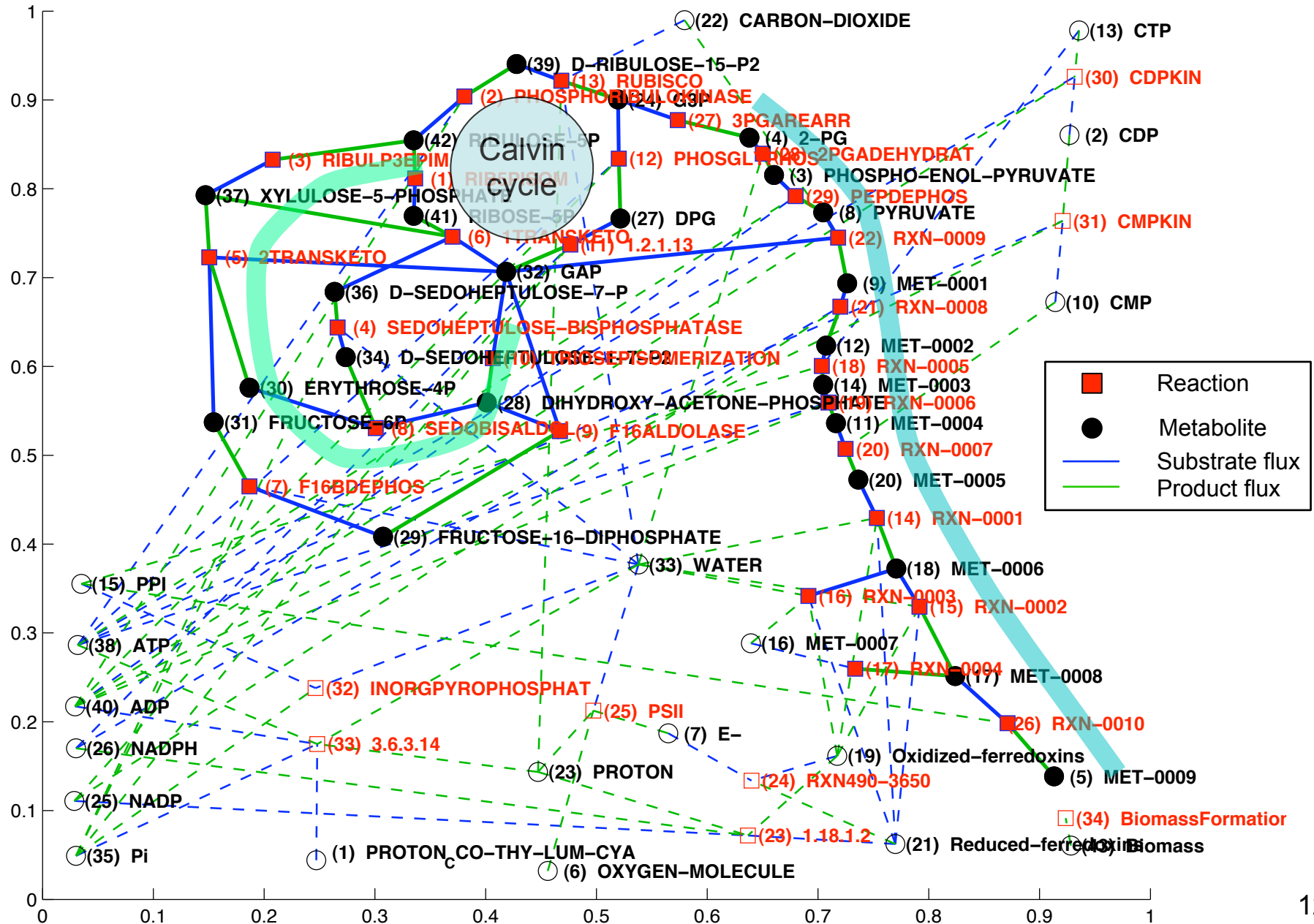
```
>>  
>>  
fx >> model = readCbModel('FBA_example.sbml');
```

Start

What a mess!



Matlab tools (1): Met/Rxn graph



Matlab tools (2): Met/Rxn browser

The screenshot displays the 'Met/Rxn browser' interface. The main window is titled 'fbagui4' and is divided into two main sections: 'Metabolites' and 'Reactions'. A 'Reload/Recompute' button is highlighted in green at the top.

Metabolites Section:

- 41 RIBOSE-5P
- Sort mets by:

-1..22	(CARBON-DIOXIDE)
2..23	(PROTON)
2..24	G3P
-1..33	(WATER)
-1..39	D-RIBULOSE-15-P2

19	Oxidized-ferredoxins
20	MET-0005
21	Reduced-ferredoxins
22	(CARBON-DIOXIDE)
23	(PROTON)
24	G3P
25	NADP
26	NADPH
27	DPG
28	DIHYDROXY-ACETONE-PHOSPHATE
29	FRUCTOSE-16-DIPHOSPHATE
30	ERYTHROSE-4P
31	FRUCTOSE-6P
32	GAP
33	(WATER)
34	D-SEDOHEPTULOSE-1-7-P2
35	Pi
36	D-SEDOHEPTULOSE-7-P
37	XYLULOSE-5-PHOSPHATE
38	ATP
39	D-RIBULOSE-15-P2
40	ADP
41	RIBOSE-5P
42	RIBULOSE-5P
43	(Biomass)

Reactions Section:

- 13 RUBISCO
- Sort rxns by:

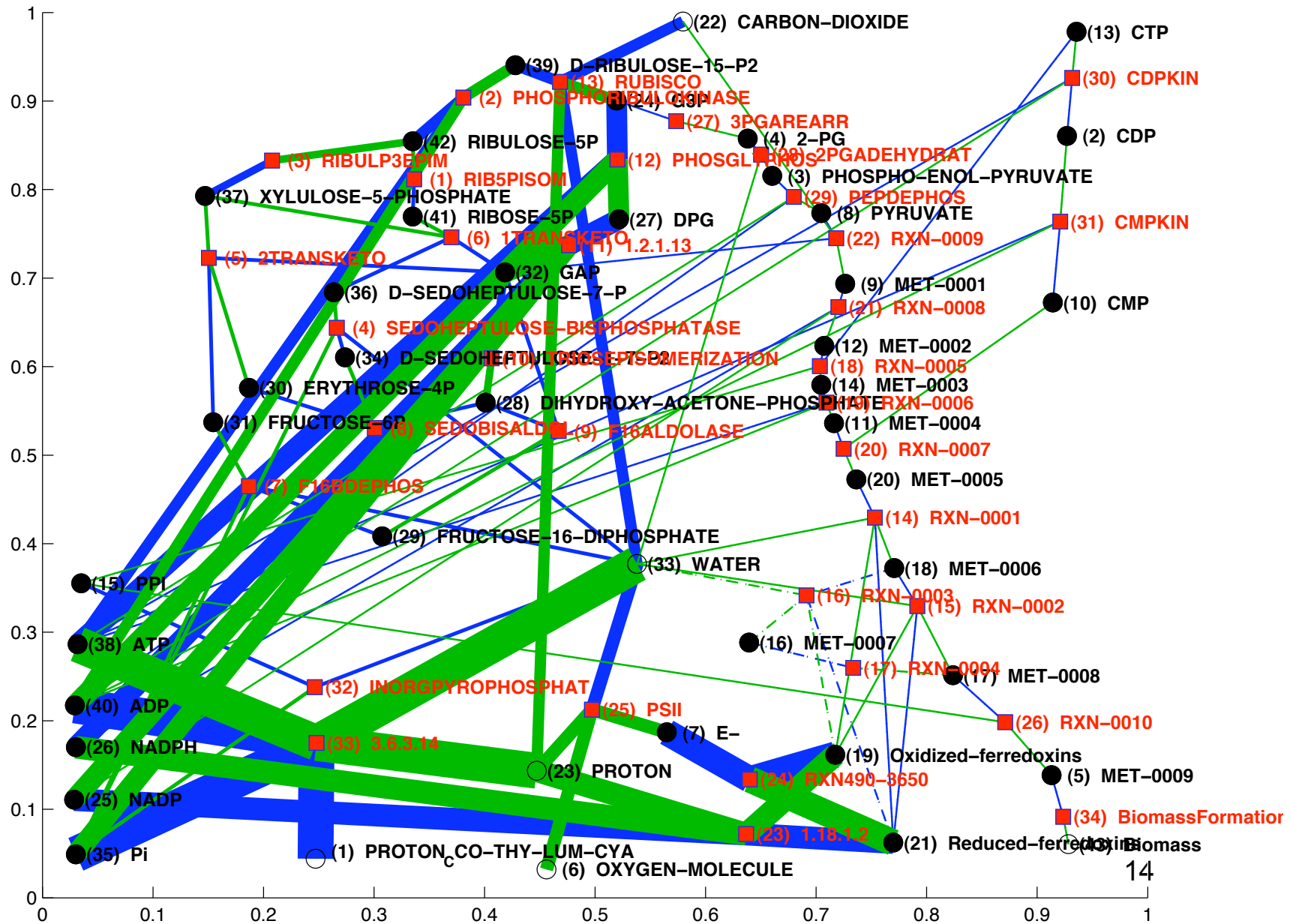
-1..1	0.00e+00	0.00e+00	1.00e+03	0.00e+00	2.00e+00	RIBSPISOM
1..6	1.00e+00	-1.00e+03	1.00e+03	0.00e+00	2.00e+00	1TRANSKETO

1	0.00e+00	0.00e+00	1.00e+03	0.00e+00	2.00e+00	RIBSPISOM
2	0.00e+00	0.00e+00	1.00e+03	0.00e+00	6.00e+00	PHOSPHORIBULOKINASE
3	0.00e+00	0.00e+00	1.00e+03	0.00e+00	4.00e+00	RIBULP3EPIM
4	0.00e+00	0.00e+00	1.00e+03	0.00e+00	2.00e+00	SEDOHEPTULOSE-BISPHOSPHATASE
5	0.00e+00	0.00e+00	1.00e+03	0.00e+00	2.00e+00	2TRANSKETO
6	1.00e+00	-1.00e+03	1.00e+03	0.00e+00	2.00e+00	1TRANSKETO
7	0.00e+00	0.00e+00	1.00e+03	0.00e+00	2.00e+00	F16BDEPHOS
8	1.00e+00	-1.00e+03	1.00e+03	0.00e+00	2.00e+00	SEDOBISALDOL
9	0.00e+00	0.00e+00	1.00e+03	0.00e+00	2.00e+00	F16ALDOLASE
10	1.00e+00	-1.00e+03	1.00e+03	0.00e+00	4.00e+00	TRIOSEPISOMERIZATION
11	0.00e+00	0.00e+00	1.00e+03	0.00e+00	1.10e+01	1.2.1.13
12	0.00e+00	0.00e+00	1.00e+03	0.00e+00	1.10e+01	PHOSGLYPHOS
13	0	0	0	0	6.00e+00	RUBISCO
14	0	0	0	0	1.00e+00	RXN-0001
15	0	0	0	0	1.00e+00	RXN-0002
16	0	0	0	0	0.00e+00	RXN-0003
17	0	0	0	0	0.00e+00	RXN-0004
18	0	0	0	0	1.00e+00	RXN-0005
19	0	0	0	0	1.00e+00	RXN-0006
20	0	0	0	0	1.00e+00	RXN-0007
21	0	0	0	0	1.00e+00	RXN-0008
22	0	0	0	0	1.00e+00	RXN-0009
23	0	0	0	0	1.20e+01	1.18.1.2
24	0	0	0	0	1.40e+01	RXN490-3650
25	0	0	0	0	7.00e+00	PSII

Overlaid Windows:

- Starting Compounds.txt:** PROTON, WATER, PROTON_CCO-THY-LUM-CYA, CARBON-DIOXIDE, OXYGEN-MOLECULE
- Biomass Compounds.txt:** MET-0009
- LimitingRxn.txt:** RUBISCO

Final product: an FBA solution



Import to Omics Viewer

The image displays the Cellular Omics Viewer software interface. The main window shows a complex metabolic pathway map with various nodes and connections, color-coded by expression levels. A legend on the left indicates expression levels from 1 to 19, with a color scale from yellow to red. A histogram below the legend shows the distribution of expression levels. The Omics Viewer dialog box is open, showing options for experiment title, file path, paint data on (Cellular Overview Diagram, Regulatory Overview Diagram, Genome Overview Diagram), type of display (Single Experiment, Animation), type of experimental data (relative or absolute), scale type (Relative, Absolute), centering (0-centered (log scale), 1-centered (linear scale)), column zero contains (Reaction IDs/EC#s), data column, highlighting color scheme (Use default color scheme with the maximum value bin cutoff being either computed from the data values, or specified), and options to save or retrieve color scheme parameters.

Cellular Omics Viewer

Refresh Display
Reload Data
Save as HTML
Print
Close

Expression Levels
19
13
9
6.1
4.4
3
2.1
1.5
1

Histogram
21
01
left side =
s in overview
right side =
s not in overview

Omics Viewer

Experiment Title: []

File: /Users/kent/Fluxes [Help]

Paint data on:
 Cellular Overview Diagram
 Regulatory Overview Diagram
 Genome Overview Diagram

Type of display:
 Single Experiment
 Animation

Type of experimental data:
Data can either be relative, in which there is a fixed center point (either 0 or 1) indicating no change, and values extending in either direction, or it can be absolute, in which the minimum value is 0 and values extend only in the positive direction. If using relative values, you must specify whether the data is centered around 0 (as in a log scale) or 1 (as in a linear scale).
If using either a 1-centered or an absolute scale, then any negative values will be discarded.

Scale type:
 Relative
 Absolute

Centering:
 0-centered (log scale)
 1-centered (linear scale)

Column zero contains: Reaction IDs/EC#s

Data column: [1]

Highlighting Color Scheme:
 Use default color scheme with the maximum value bin cutoff being either
 computed from the data values, or [] or []
Note: Specifying a maximum cutoff value is the simplest way to ensure that the same color scale is used for multiple experiments even if the range of data in the experiments differ. The minimum value bin cutoff will be 0 for absolute data, 1/(max-cutoff) for relative data in normal format, and -(max-cutoff) for relative data in log format.
 Specify color value cutoffs (numbers, one per line): []

Assign colors automatically:
 Assign colors manually:

Save Color Scheme Parameters Retrieve Saved Color Scheme Parameters

OK Cancel

Recap

- **What have we accomplished?**
 - From stoichiometry (structure) of reaction network ...
 - Steady-state flux “subspace”, $v \in \mathcal{N}(S)$
 - Conserved “motifs”, $\dot{x} \in \mathcal{N}(S^T)$
 - From flux constraints ...
 - Unique steady-state operating point (solution of LP)
 - We used “biomass” optimization, but there are many variations on the theme (e.g., MOMA, ROOM)
- **What is not represented?**
 - Quantity (Concentrations)
 - Thermodynamics (chemical potentials)
 - Dynamics (approach to steady-state, stability,...)
 - Spatial inhomogeneities
 - Regulation (neither metabolic nor transcriptional)

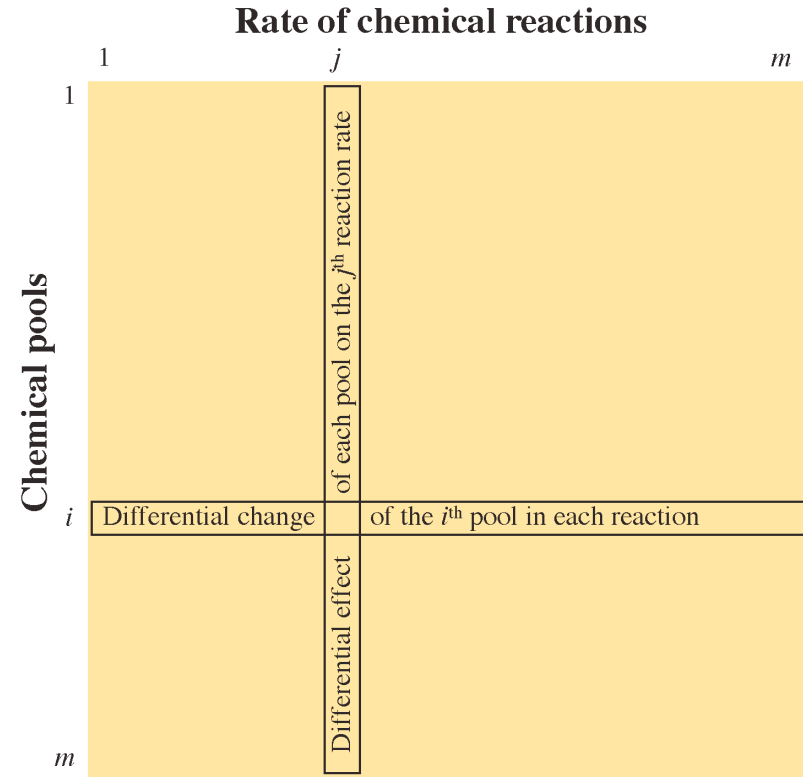
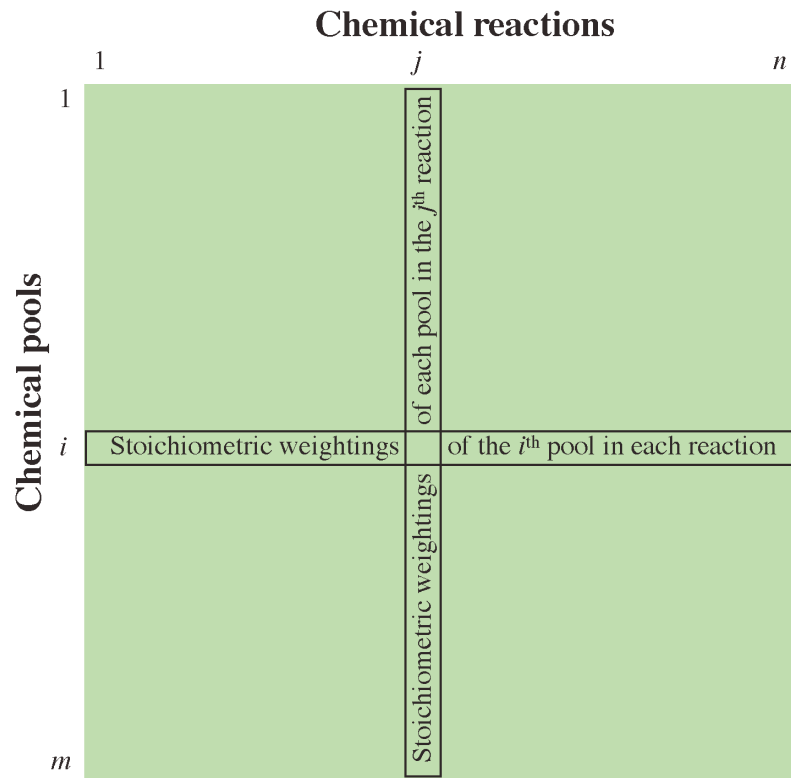
Systems biology: Stoichiometry and kinetics

Matrix of stoichiometric weights (S)

$$S_{ij} = \frac{\dot{x}_i}{v_j} \Big|_{v_{k \neq j} = 0}$$

Matrix of linearized kinetics ($D_x r$)

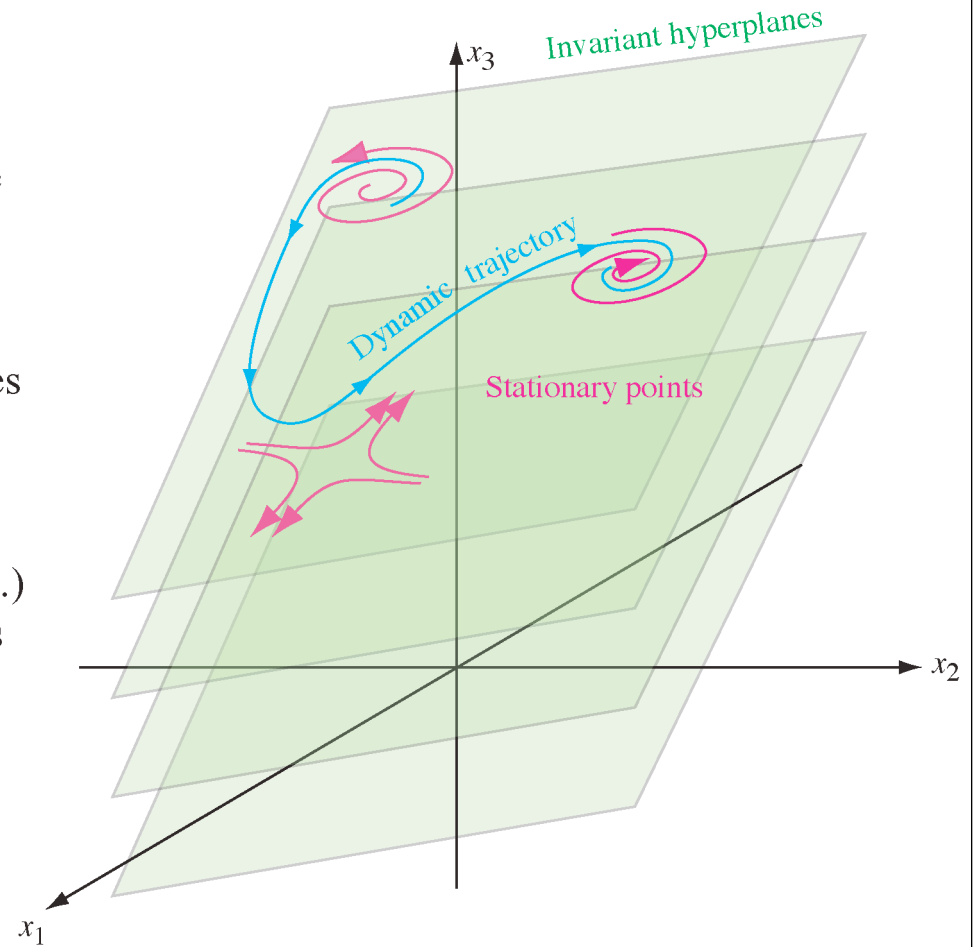
$$D_x r_{ij} = \frac{r_i}{x_j} \Big|_{x_{k \neq j} = 0}$$



Systems biology:

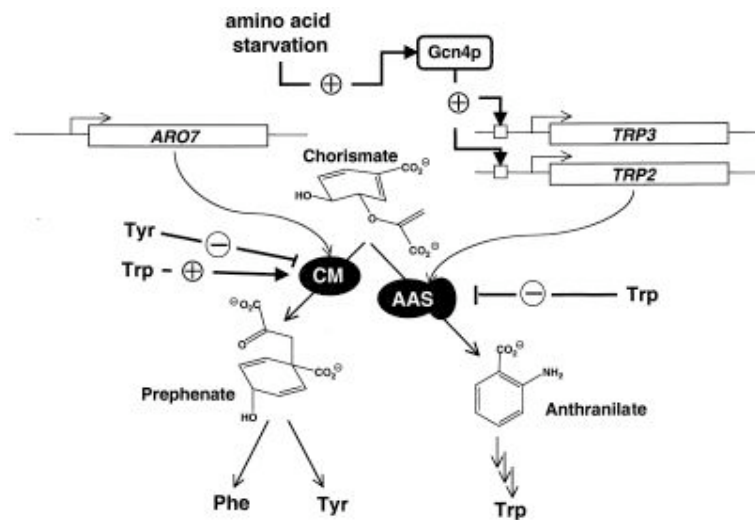
Dynamical behavior of chemical & electrical networks

	Chemical	Electrical
Network:		
State	Chem.pools	Electrical charge
Flow	Accum.rates	Current
Force	Chem.potential	Voltage
Connections	Stoichiometry	Wiring
Init.cond.	Chem.pools	Capacitor charges
Dynamics determined by:		
Constitutive relations	React.kinetics	Device physics (transistors, etc.)
Ext.enviroin.	Exchange pools	I/O signals/loads
Params.	Regulatory modes	Supervisory modes

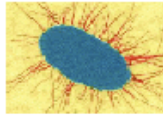


A network model can (should?) include:

- Structure (connections, stoichiometric weights)
- Flux (chemical flow)
- Concentration (chemical potential)
- Kinetics (flux/potential constitutive model)
- Regulatory “feedback”



Example of “Metabolic Control” from EcoCyc



E. coli K12 Pathway: tryptophan biosynthesis

[More Detail](#)
[Less Detail](#)
[Cross-Species Comparison](#)
[Download](#)

chorismate biosynthesis

chorismate

anthranilate

N-(5'-phosphoribosyl)-anthranilate

1-(*o*-carboxyphenylamino)-1'-deoxyribulose-5'-phosphate

indole-3-glycerol-phosphate

Mtr tryptophan ArAAP transporter

indole

SstT DAACS transporter

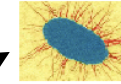
L-serine

L-tryptophan

Summary:

The TrpE protein is also called Component I of the anthranilate synthase enzyme complex. *In vitro* in the absence of the TrpD (Component II) protein, TrpE can use ammonia as the amino donor for the synthesis of anthranilate from chorismate at approximately 20% efficiency [[Ito69](#)]. As a component of the anthranilate synthase complex, TrpD (Component II) provides the glutamine amidotransferase function that allows glutamine to serve as the amino donor in anthranilate formation.

The TrpE subunit contains the tryptophan binding site for feedback inhibition [[Crawford89](#)]. Both the TrpE and TrpA polypeptides in the trp operon lack tryptophan residues [[Nichols81](#), [Nichols81](#)].



E. coli K12 Enzyme: anthranilate synthase

Summary:

The native anthranilate synthase enzyme exists as a tetrameric complex of two subunits each of the TrpD protein. The TrpD protein is bifunctional; it also catalyzes the second reaction in the biosynthesis of tryptophan.

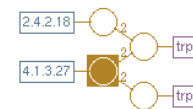
Citations: [[L74](#), [Nichols81](#), [Yanofsky81a](#)]

Subunit composition of anthranilate synthase = [TrpE]₂[TrpD]₂

[anthranilate synthase component I = TrpE](#)
[anthranilate synthase component II = TrpD](#)

Genetic Regulation Schematic:

Gene-Reaction Schematic:



Transcriptional control



MultiFun Terms: [location of gene products -> cytoplasm](#)

[metabolism -> biosynthesis of building blocks -> amino acids -> tryptophan](#)

Enzymatic reaction of: anthranilate synthase

Synonyms: ASase

[chorismate + L-glutamine <=> anthranilate + pyruvate + L-glutamate](#)

The reaction direction shown, that is, A + B <=> C + D versus C + D <=> A + B, is in accordance with the reaction direction shown in the pathway. This reaction is reversible.

Alternative Substrates for L-glutamine: [ammonia](#)

In Pathways: [tryptophan biosynthesis](#)

Summary:

The complex of the TrpE (Component I) and TrpD (Component II) proteins uses either glutamine or ammonia as the amino donor for the synthesis of anthranilate from chorismate. In the absence of TrpD it cannot catalyze the formation of anthranilate using ammonia instead of glutamine as the amino donor [[Ito69](#)]. Anthranilate is responsive to changes in the free tryptophan concentration [[Nichols81](#)].

Citations: [[Nichols81](#), [Crawford89](#), [Ito69a](#)]

Cofactors: [Mg²⁺](#) [[Ito69b](#)]

Alternative Cofactors for Mg²⁺: [Co²⁺](#), [Fe²⁺](#)

Inhibitors (Competitive): [L-tryptophan](#) [[Comment 1](#)]

Inhibitors (Noncompetitive): [L-tryptophan](#) [[Comment 2](#)]

Primary Physiological Regulators of Enzyme Activity: [L-tryptophan](#)

Modulatory control