Protein2: Properties & Quantitation
(Last week)

- Separation of proteins & peptides
- Protein localization & complexes
- Peptide identification (MS/MS)
  - Database searching & sequencing.
- Protein quantitation
  - Absolute & relative
- Protein modifications & crosslinking
- Protein - metabolite quantitation

Net1: Simulation & optimization

- **Macroscopic continuous concentration rates**
  - **Cooperativity & Hill coefficients**
  - **Bistability**
- **Mesoscopic discrete molecular numbers**
  - **Approximate & exact stochastic**
- **Chromosome Copy Number Control**
- **Flux balance optimization**
  - **Universal stoichiometric matrix**
  - **Genomic sequence comparisons**

Networks Why model?

Red blood cell metabolism	Enzyme kinetics (Pro2)
Cell division cycle	**Checkpoints** (RNA2)
Plasmid Copy No. Control	**Single molecules**
Phage λ switch	Stochastic bistability
Comparative metabolism	**Genomic connections**
Circadian rhythm	Long time delays
E. coli chemotaxis	Adaptive, spatial effects

also, all have large genetic & kinetic datasets.

Types of interaction models

Quantum Electrodynamics	subatomic
Quantum mechanics	electron clouds
Molecular mechanics	spherical atoms (101Pro1)
Master equations	stochastic single molecules (Net1)
Phenomenological rates ODE	**Concentration & time (C,t)**
Flux Balance	dC_{ik}/dt optima steady state (Net1)
Thermodynamic models	$dC_{ik}/dt = 0$ k reversible reactions
Steady State	$\Sigma dC_{ik}/dt = 0$ (sum k reactions)
Metabolic Control Analysis	$d(dC_{ik}/dt)/dC_j$ (i = chem.species)
Spatially inhomogenous models	dCi/dx

Increasing scope, decreasing resolution

In vivo & (classical) *in vitro*

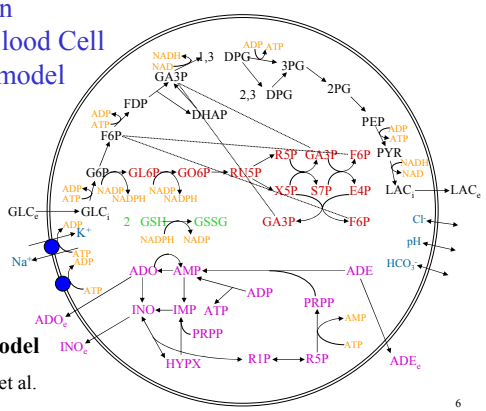
1) "Most measurements in enzyme kinetics are based on initial rate measurements, where only the substrate is present… enzymes in cells operate in the presence of their products" Fell p.54 (Pub)

2) Enzymes & substrates are closer to equimolar than in classical *in vitro* experiments.

3) Proteins close to crystalline densities so some reactions occur faster while some normally spontaneous reactions become undetectably slow.

e.g. Bouffard, et al., Dependence of lactose metabolism upon mutarotase encoded in the gal operon in *E.coli.* J Mol Biol. 1994; 244:269-78. (Pub)

Human Red Blood Cell ODE model



ODE model

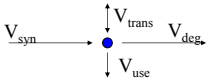
Jamshidi et al.

2000 (Pub)

Factors Constraining Metabolic Function

- Physicochemical factors
  - Mass, energy, and redox balance:
    - Systemic stoichiometry
  - osmotic pressure, electroneutrality, solvent capacity, molecular diffusion, thermodynamics
  - Non-adjustable constraints
- System specific factors
  - Capacity:
    - Maximum fluxes
  - Rates:
    - Enzyme kinetics
  - Gene Regulation
  - Adjustable constraints

Dynamic mass balances on each metabolite



$$\frac{dX_i}{dt} = (V_{syn} - V_{deg} - V_{use}) - V_{trans} = (S_{ij} v_j) - b_i$$

Time derivatives of metabolite concentrations are linear combination of the reaction rates.

The reaction rates are non-linear functions of the metabolite concentrations (typically from in vitro kinetics).

1. v_j is the jth reaction rate, b is the transport rate vector,

S_{ij} is the "Stoichiometric matrix" = moles of metabolite i produced in reaction j

RBC model integration

Reference	Glyc-olysis	PPP	ANM	Na^+/K^+ Pump	Osmot.	Trans-port	Hb-5 ligands	Gpx Hb	Ca	Shape
Rapoport '74-6	+	-	-	-	-	-	-	-	-	-
Heinrich '77	+	-	-	-	-	-	-	-	-	-
Ataullakhanov'81	+	+	-	-	-	-	-	-	-	-
Schauer '81	+	-	+	-	-	-	-	-	-	-
Brumen '84	+	-	-	+	+	-	-	-	-	-
Werner '85	+	-	-	+	+	+	-	-	-	-
Joshi '90	+	+	+	+	+	+	-	-	-	-
Yoshida '90	-	-	-	-	-	-	+	-	-	-
Lee '92	+	+	+	+	+	+	(+)	-	-	-
Gimsa '98	-	-	-	-	-	-	-	-	-	+
Destro-Bisol '99	-	-	-	-	-	-	-	(-)	-	-
Jamshidi '00	+	+	+	+	+	+	-	-	-	-

Scopes & Assumptions

- Mechanism of ATP utilization other than nucleotide metabolism and the Na^+/K^+ pump (75%) is not specifically defined
- Ca^{2+} transport not included
- Guanine nucleotide metabolism neglected
  - little information, minor importance
- Cl^-, HCO_3^-, LAC, etc. are in "pseudo" equilibrium
- No intracellular concentration gradients
- Rate constants represent a "typical cell"
- Surface area of the membrane is constant
- Environment is treated as a sink

Glycolysis Dynamic Mass Balances

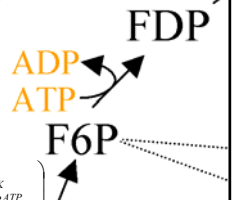
$$\frac{d}{dt}(G6P) = v_{HK} - v_{PGI} - v_{G6PDH}$$
$$\frac{d}{dt}(F6P) = v_{PGI} - v_{PFK} + v_{TA} + v_{TKII}$$
$$\frac{d}{dt}(FDP) = v_{PFK} - v_{ALD}$$
$$\frac{d}{dt}(DHAP) = v_{ALD} - v_{TPI}$$
$$\frac{d}{dt}(GA3P) = v_{ALD} + v_{TPI} - v_{GAPDH} + v_{TKI} + v_{TKII} - v_{TA}$$
$$\frac{d}{dt}(1,3DPG) = v_{GAPDH} - v_{PGK} - v_{DPGM}$$
$$\frac{d}{dt}(2,3DPG) = v_{DPGM} - v_{DPG_{ase}}$$

$$\frac{d}{dt}(3PG) = v_{PGK} - v_{PGM} + v_{DPGase}$$
$$\frac{d}{dt}(2PG) = v_{PGM} - v_{EN}$$
$$\frac{d}{dt}(PEP) = v_{EN} - v_{PK}$$
$$\frac{d}{dt}(PYR) = v_{PK} - v_{PYR_{ex}} - v_{LDH}$$
$$\frac{d}{dt}(LAC) = v_{LDH} - v_{LAC_{ex}}$$
$$\frac{d}{dt}(NADH) = v_{GAPDH} - v_{LDH}$$

$$\frac{dX_i}{dt} = (V_{syn} - V_{deg} - V_{use}) - V_{trans} = (S_{ij} v_j^1) - b_i$$

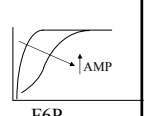
Enzyme Kinetic Expressions

Phosphofructokinase



$$v_{PFK} = \frac{v_{max}^{PFK}}{N_{PFK}} \left(\frac{F6P / K_{F6P}^{PFK}}{1 + F6P / K_{F6P}^{PFK}} \right) \left(\frac{Mg \bullet ATP / K_{Mg \bullet ATP}^{PFK}}{1 + Mg \bullet ATP / K_{Mg \bullet ATP}^{PFK}} \right)$$

$$N_{PFK} = 1 + L_0^{PFK} \frac{\left(1 + ATP_{free} / K_{ATP}^{PFK}\right)^4 \left(1 + Mg / K_{Mg}^{PFK}\right)^4}{\left(1 + AMP / K_{AMP}^{PFK}\right)^4 \left(1 + F6P / K_{F6P}^{PFK}\right)^4}$$



Kinetic Expressions

- All rate expressions are similar to the previously shown rate expression for phosphofructokinase.
- Model has 44 rate expressions with ~ 5 constants each → ~ 200 parameters
- What are the assumptions associated with using these expressions?

Kinetic parameter assumptions

- in vitro values represent the in vivo parameters
  - protein concentration in vitro much lower than in vivo
    - enzyme interactions (enzymes, cytoskeleton, membrane, …)
  - samples used to measure kinetics may contain unknown conc. of effectors (i.e. fructose 2,6-bisphosphate)
  - enzyme catalyzed enzyme modifications
- all possible concentrations of interacting molecules been considered (interpolating)
  - e.g. glutamine synthase (unusually large # of known effectors)
    - 3 substrates, 3 products, 9 significant effectors
    - 4^{15} (~10^9) measurements: 4 different conc. of 15 molecules (Savageau, 1976)
      - in vivo probably even more complex, but approximations are effective.
- have all interacting molecules been discovered?
- and so on …

Additional constraints:
Physicochemical constrains

Osmotic Pressure Equilibrium (interior & exterior, m chem. species)

$$\pi_i = \pi_e$$

$$RT\sum_{j=1}^{m}\phi_{ij}C_{ij} = RT\sum_{j=1}^{m}\phi_{ej}C_{ej}$$

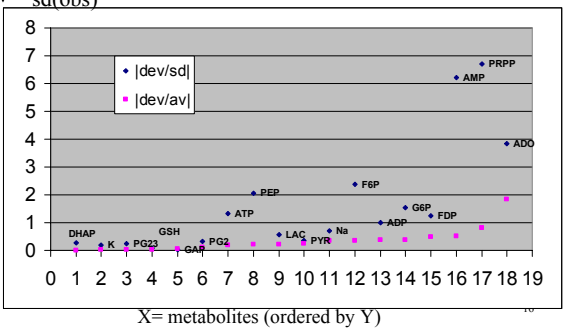
Electroneutrality (z = charge, Concentration)

$$\sum_{j=1}^{m} z_{ij}C_{ij} = 0$$

$$\sum_{j=1}^{m} z_{ej}C_{ej} = 0$$

RBC steady-state in vivo vs calculated

$$\frac{|obs - calc|}{sd(obs)} = Y$$



X= metabolites (ordered by Y)

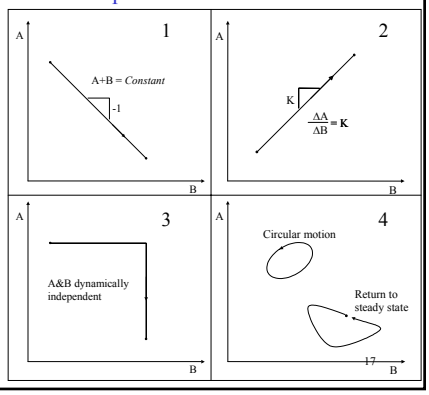
Phase plane diagrams: concentration of metabolite A vs B over a specific time course

1: conservation relationship.

2: a pair of concentrations in equilibrium

3: two dynamically independent metabolites

4: a closed loop trace

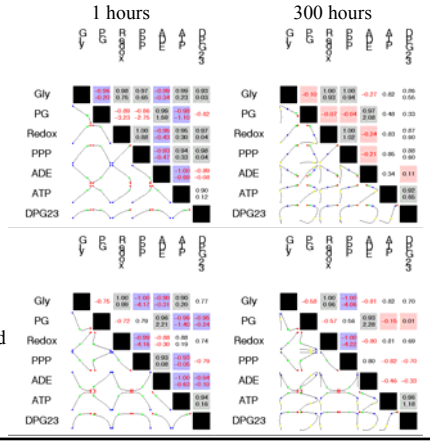


ATP & Redox loads

ATP load

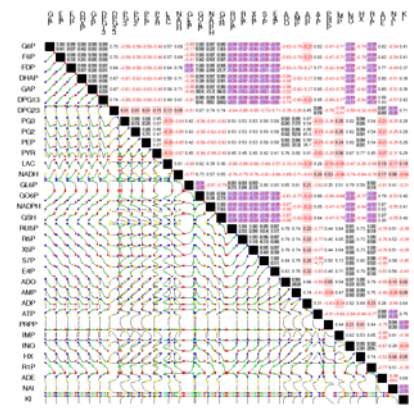
Red 0 hours
Green 0.1
Blue 1.0
Yellow 10
End 300

Redox load



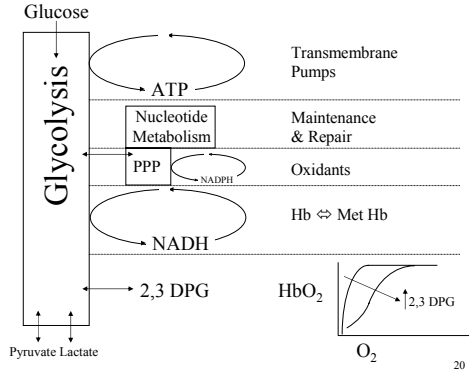
Redox Load

0 to 300 hour dynamics
34 metabolites calculated



ODE model
Jamshidi et al. 2000 (Pub)

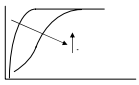
RBC Metabolic "Machinery"

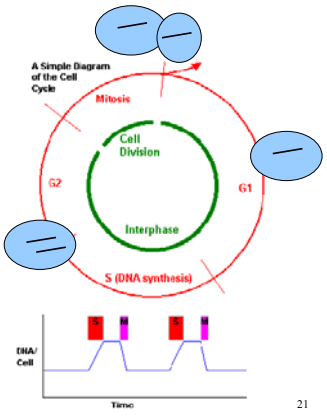


20

Cell Division Cycle

G2 arrest to M arrest switch



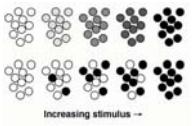


21

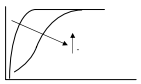
Hill coefficients

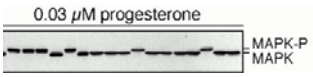
Response

$$R = \frac{1}{1+(K/S)^H}$$



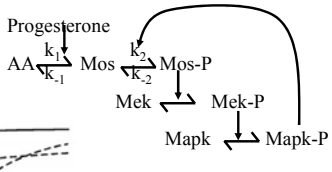
H simple hyperbolic = 1
H (R=HbO_2, S=O_2) sigmoidal = 2.8
H (R=Mapk-P, S=Mos) = 3
H (R=Mapk-P, S=Progesterone in vivo) = 42

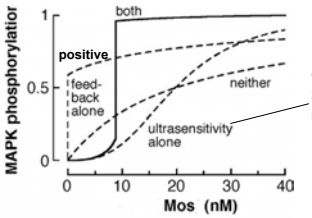




22

"The biochemical basis of an all-or-none cell fate switch in Xenopus oocytes."





(a chain of enzyme modifiers close to saturation generate higher sensitivity to signals than one enzyme can)

Science 1998;280:895-8 Ferrell & Machleder, (Pub)

23

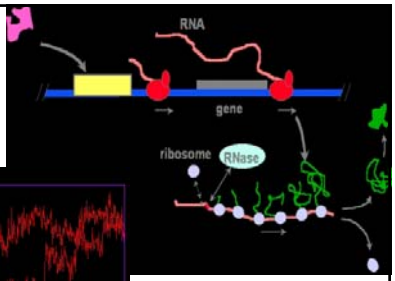
Net1: Simulation & optimization

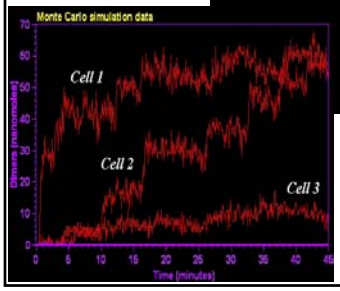
- **Macroscopic continuous concentration rates**
  - **Cooperativity & Hill coefficients**
  - **Bistability**
- **Mesoscopic discrete molecular numbers**
  - **Approximate & exact stochastic**
- **Chromosome Copy Number Control**
- **Flux balance optimization**
  - **Universal stoichiometric matrix**
  - **Genomic sequence comparisons**

24

Arkin A, Ross J, McAdams HH *Genetics* 1998 149(4):1633.

Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *E. coli* cells.





Variation in level, time & whole cell effect

Efficient exact stochastic simulation of chemical systems with many species & many channels

"the Next Reaction Method, an exact algorithm ...time proportional to the logarithm of the number of reactions, not to the number of reactions itself". Gibson & Bruck, 1999; J. Physical Chemistry. (Pub)

Gillespie J.Phys Chem 81:2340-61. 1977. Exact stochastic simulation of coupled chemical reactions

Utilizing Noise

Hasty, et al. PNAS 2000; 97:2075–2080, Noise-based switches and amplifiers for gene expression (Pub)

"Bistability ... arises naturally... Additive external noise [allows] construction of a protein switch... using short noise pulses. In the multiplicative case, ... small deviations in the transcription rate can lead to large fluctuations in the production of protein".

Paulsson, et al. PNAS 2000; 97:7148–53. Stochastic focusing: fluctuation-enhanced sensitivity of intracellular regulation. (Pub) (exact master equations)

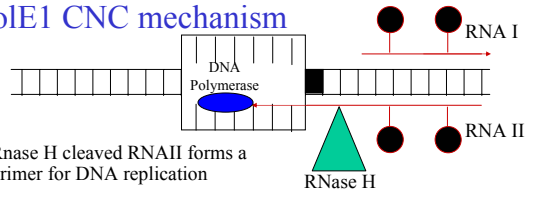
Net1: Simulation & optimization

- **Macroscopic continuous concentration rates**
  - **Cooperativity & Hill coefficients**
  - **Bistability**
- **Mesoscopic discrete molecular numbers**
  - **Approximate & exact stochastic**
- **Chromosome Copy Number Control**
- **Flux balance optimization**
  - **Universal stoichiometric matrix**
  - **Genomic sequence comparisons**

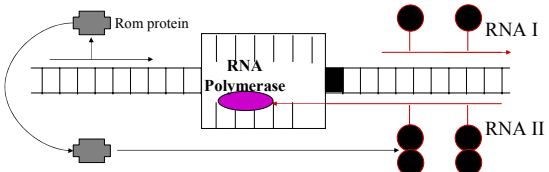
Copy Number Control Models

- Replication of ColE1 & R1 Plasmids
- Determine the factors that govern the plasmid copy number
  - cellular growth rate
  - One way to address this question is via the use of a kinetic analysis of the replication process, and relate copy number to overall cellular growth.
- Why? the copy number can be an important determinant of cloned protein production in recombinant microorganisms

ColE1 CNC mechanism



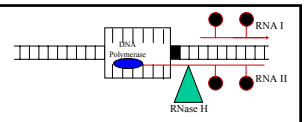
Rnase H cleaved RNAII forms a primer for DNA replication



RNA I binding to RNA II prevents RNaseH from cleaving RNA II

Where do we start?

Dynamic mass balance

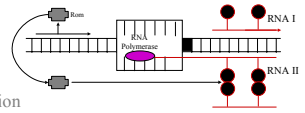


What are the important parameters?

Plasmid, RNA I, RNA II, Rom, μ

All the constants

degradation, initiation, inhibition



Assumptions?

RNaseH rate is very fast ➔ instantaneous

DNA polymerization is very rapid

Simplify by subsuming [RNA II] ➔ model RNA I inhibition

RNA I and RNA II transcription is independent (neglect convergent transcription)

Rom protein effects constant

Consider 2 species: RNA I and plasmid

Many more assumptions...

Dynamic Mass Balance: ColE1 RNAI

[concentration in moles/liter]

Rate of change of [RNA I] = Synthesis of RNA I - Degradation of RNA I - Dilution due to cell growth

R = [RNA I]
k_1 = rate of RNA I initiation
N = [plasmid]
k_d = rate of degradation
μ = growth rate

$$\frac{dR}{dt} = k_1 N - (k_d + \mu)R$$

Keasling,& Palsson (1989) *J theor Biol* **136,** 487-492; **141,** 447-61.

Dynamic Mass Balance: ColE1 Plasmid

Rate of change of [N] = Plasmid Replication - Dilution due to cell growth

R = [RNA I]
k_2 = rate of RNA II initiation
N = [plasmid]
K_1 = RNA I/RNA II binding constant (an inhibition constant)
μ = growth rate

$$\frac{dN}{dt} = k_2\left(\frac{1}{1+K_1 R}\right)N - \mu N$$

Solve for N(t).

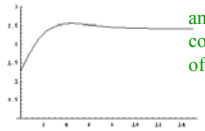
Mathematica ODE program

```
dr = k1 * n - (kd + mu) * r;
dn = k2 * (1 / (1 + kI * r)) * n - mu * n;
sol = Solve[{dr == 0, dn == 0}, {r, n}]
```

Formulae for steady state

```
{{n → -(-k2 + mu) (kd + mu) / (k1 kI mu), r → -(-k2 + mu) / (kI mu)}}
n0 = sol[[1, 1, 2]] /. mu -> 1 /. k1 -> 1.05 /. k2 -> 2 /. kI -> 1 /. kd -> .35;
r0 = sol[[1, 2, 2]] /. mu -> 1 /. k1 -> 1.05 /. k2 -> 2 /. kI -> 1 /. kd -> .35;
dsol = NDSolve[
  {x'[t] == dr /. r -> x[t] /. n -> y[t],
   y'[t] == dn /. r -> x[t] /. n -> y[t],
    x[0] == r0, y[0] == n0}
  /. k1 -> 1.05 /. mu -> .5 /. k2 -> 2 /. kI -> 1 /. kd -> .35, {x, y}, {t, 0, 20}];
Plot[y[t] /. dsol, {t, 0, 15}, PlotRange -> {0, 3}]
```

start at mu=1 shift to mu=.5



and then solve for plasmid concentration N as a function of time.

```
▪ Graphics ▪
```


Stochastic models for CNC

Paulsson & Ehrenberg, J Mol Biol 1998;279:73-88. Trade-off between segregational stability and metabolic burden: a mathematical model of plasmid ColE1 replication control. (Pub),

J Mol Biol 2000;297:179-92. Molecular clocks reduce plasmid loss rates: the R1 case. (Pub)

While copy number control for ColE1 efficiently corrects for fluctuations that have already occurred, R1 copy number control prevents their emergence in cells that by chance start their cycle with only one plasmid copy. Regular, clock-like, behaviour of single plasmid copies becomes hidden in experiments probing collective properties of a population of plasmid copies ... The model is formulated using master equations, taking a stochastic approach to regulation"

From RBC & CNC to models for whole cell replication?

e.g. *E. coli* ?

What are the difficulties?

- The number of parameters
- Measuring the parameters
- Are parameters measured *in vitro* representative to the parameters *in vivo*

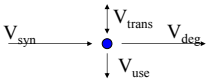
Factors Constraining Metabolic Function

- Physicochemical factors:
  - **Mass, energy, and redox balance:**
    - **Systemic stoichiometry**
  - osmotic pressure, electroneutrality, solvent capacity, molecular diffusion, **thermodynamics**
  - Non-adjustable constraints
- System specific factors:
  - **Capacity:**
    - **Maximum fluxes**
  - Rates:
    - Enzyme kinetics
  - Gene Regulation
  - Adjustable constraints

Net1: Simulation & optimization

- **Macroscopic continuous concentration rates**
  - **Cooperativity & Hill coefficients**
  - **Bistability**
- **Mesoscopic discrete molecular numbers**
  - **Approximate & exact stochastic**
- **Chromosome Copy Number Control**
- **Flux balance optimization**
  - **Universal stoichiometric matrix**
  - **Genomic sequence comparisons**

Dynamic mass balances on each metabolite



$$\frac{dX_i}{dt} = (V_{syn} - V_{deg} - V_{use}) - V_{trans} = (S_{ij} v_j) - b_i$$

Time derivatives of metabolite concentrations are linear combination of the reaction rates. The reaction rates are non-linear functions of the metabolite concentrations (typically from in vitro kinetics).

Where v_j is the jth reaction rate, b is the transport rate vector, S_{ij} is the "Stoichiometric matrix" = moles of metabolite i produced in reaction j

Flux-Balance Analysis

- Make simplifications based on the properties of the system.
  - Time constants for metabolic reactions are very fast (sec- min) compared to cell growth and culture fermentations (hrs)
  - There is not a net accumulation of metabolites in the cell over time.
- One may thus consider the steady-state approximation.

$$\frac{d\mathbf{X}}{dt} = \mathbf{S} \cdot \mathbf{v} - \mathbf{b} = 0$$

Flux-Balance Analysis

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

- Removes the metabolite concentrations as a variable in the equation.
- Time is also not present in the equation.
- We are left with a simple matrix equation that contains:
  - Stoichiometry: *known*
  - Uptake rates, secretion rates, and requirements: *known*
  - Metabolic fluxes: Can be solved for!

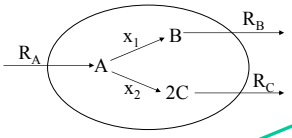
In the ODE cases before we already had fluxes (rate equations, but lacked C(t).

Additional Constraints

- Fluxes >= 0 (reversible = forward- reverse)
- The flux level through certain reactions is known
- Specific measurement – typically for uptake rxns
- maximal values
- uptake limitations due to diffusion constraints
- maximal internal flux

$$\alpha_i \le v_i \le \beta_i$$

Flux Balance Example



Flux Balances:
A: $R_A - x_1 - x_2 = 0$
B: $x_1 - R_B = 0$
C: $2\,x_2 - R_C = 0$

Constraints:
$R_A = 3$
$R_B = 1$

Equations:
A: $x_1 + x_2 = 3$
B: $x_1 = 1$
C: $2\,x_2 - R_C = 0$

$$\begin{array}{c} \\ A \\ B \\ C \end{array}\begin{array}{c} \begin{array}{ccc} x_1 & x_2 & R_C \end{array} \\ \begin{bmatrix} 1 & 1 & \\ 1 & & \\ & 2 & -1 \end{bmatrix} \end{array}\begin{bmatrix} x_1 \\ x_2 \\ R_C \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$$

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

FBA Example

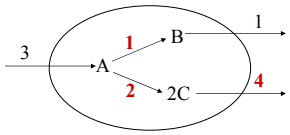
$$\mathbf{S \cdot v = b}$$

$$\mathbf{v = S^{-1} b}$$

$$\mathbf{v = S^{-1} b}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ R_C \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -1 & 0 \\ 2 & -2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ R_C \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

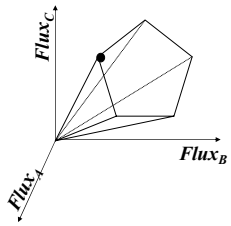


FBA

- Often, enough measurements of the metabolic fluxes cannot be made so that the remaining metabolic fluxes can be calculated.
- Now we have an underdetermined system
  - more fluxes to determine than mass balance constraints on the system
  - what can we do?

Incomplete Set of Metabolic Constraints

- Identify a specific point within the feasible set under any given condition
- Linear programming - Determine the optimal utilization of the metabolic network, subject to the physicochemical constraints, to maximize the growth of the cell



Assumption:
The cell has found the optimal solution by adjusting the system specific constraints (enzyme kinetics and gene regulation) through evolution and natural selection.

Find the optimal solution by linear programming

Under-Determined System

- All real metabolic systems fall into this category, so far.
- Systems are moved into the other categories by measurement of fluxes and additional assumptions.
- Infinite feasible flux distributions, however, they fall into a solution space defined by the **convex polyhedral cone.**
- The actual flux distribution is determined by the cell's regulatory mechanisms.
- It absence of kinetic information, we can estimate the metabolic flux distribution by postulating **objective functions(Z)** that underlie the cell's behavior.
- Within this framework, one can address questions related to the capabilities of metabolic networks to perform functions while constrained by stoichiometry, limited thermodynamic information (reversibility), and physicochemical constraints (ie. uptake rates)

FBA - Linear Program

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

- For growth, define a growth flux where a linear combination of monomer (M) fluxes reflects the known ratios (d) of the monomers in the final cell polymers.

$$\sum_{allM} d_M \cdot M \xrightarrow{v_{growth}} biomass$$

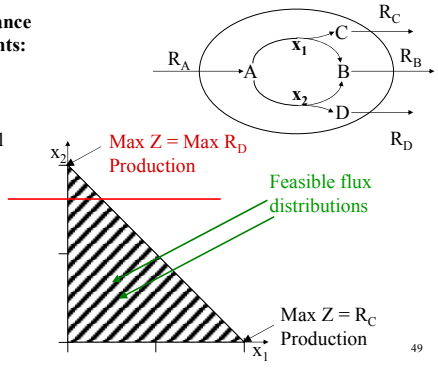
- A linear programming problem is formulated where one finds a solution to the above equations, while minimizing an objective function (Z). Typically $Z = v_{growth}$ (or production of a key compound).
- Constraints to the LP problem:
- i reactions

$$\begin{array}{l} \mathbf{S \cdot v = b} \\ v_i \geq 0 \\ \alpha_i \leq v_i \leq \beta_i \\ v_i = X_i \end{array}$$

Very simple LP solution

Flux Balance Constraints:

$R_A = R_B$
$R_A < 1$
$x_1 + x_2 < 1$
$x_1 > 0$
$x_2 > 0$



Applicability of LP & FBA

- Stoichiometry is well-known
- Limited thermodynamic information is required
  - reversibility vs. irreversibility
- Experimental knowledge can be incorporated in to the problem formulation
- Linear optimization allows the identification of the reaction pathways used to fulfil the goals of the cell if it is operating in an optimal manner.
- The relative value of the metabolites can be determined
- Flux distribution for the production of a commercial metabolite can be identified. Genetic Engineering candidates

Precursors to cell growth

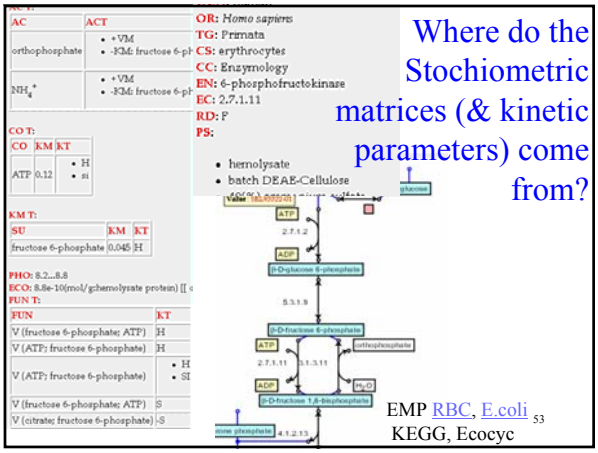
- How to define the growth function.
  - The biomass composition has been determined for several cells, *E. coli* and *B. subtilis*.
    - This can be included in a complete metabolic network
  - When only the catabolic network is modeled, the biomass composition can be described as the 12 biosynthetic precursors and the energy and redox cofactors

in silico cells

	E. coli	*H. influenzae*	*H. pylori*
Genes	**695**	**362**	**268**
Reactions	**720**	**488**	**444**
Metabolites	**436**	**343**	**340**
(of total genes	**4300**	**1700**	**1800)**

Edwards, et al 2002. Genome-scale metabolic model of Helicobacter pylori 26695. J Bacteriol. 184(16):4582-93.

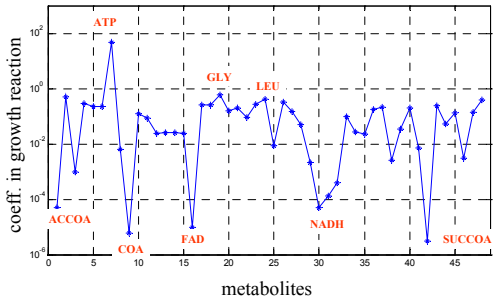
Segre, et al, 2002 Analysis of optimality in natural and perturbed metabolic networks. PNAS 99: 15112-7. (Minimization Of Metabolic Adjustment) http://arep.med.harvard.edu/moma/



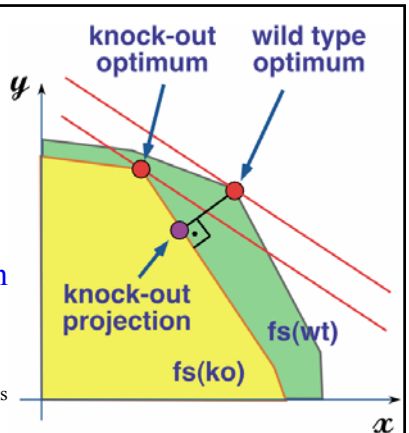
Where do the Stochiometric matrices (& kinetic parameters) come from?

EMP RBC, E.coli
KEGG, Ecocyc

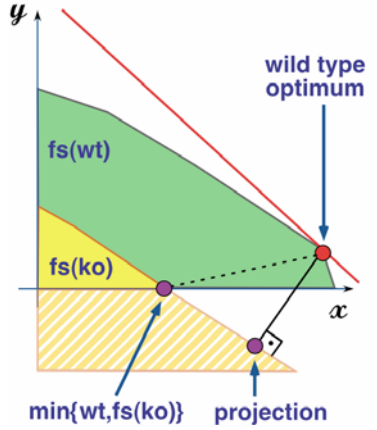
Biomass Composition



Flux ratios at each branch point yields optimal polymer composition for replication

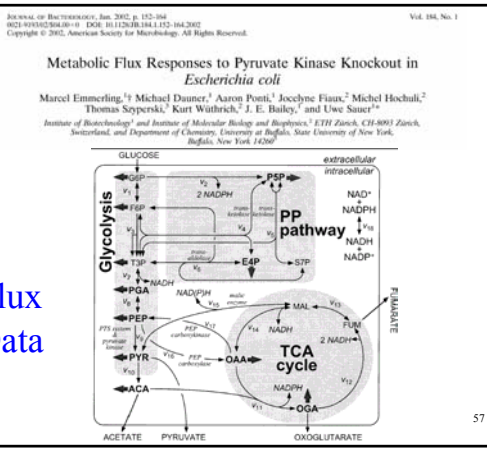


x,y are two of the 100s of flux dimensions



56

Flux Data



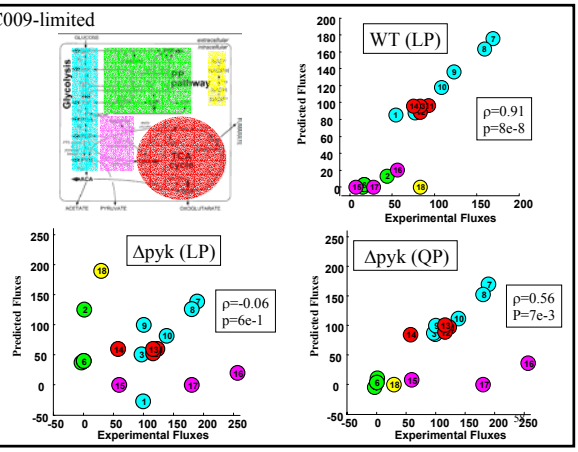
Metabolic Flux Responses to Pyruvate Kinase Knockout in *Escherichia coli*

Marcel Emmerling,[1]† Michael Dauner,[1] Aaron Ponti,[1] Jocelyne Fiaux,[2] Michel Hochuli,[2] Thomas Szyperski,[2] Kurt Wüthrich,[2] J. E. Bailey,[1] and Uwe Sauer[1]*

GLUCOSE
extracellular
intracellular
Glycolysis
PP pathway
TCA cycle
ACETATE
PYRUVATE
OXOGLUTARATE
FUMARATE

57

C009-limited

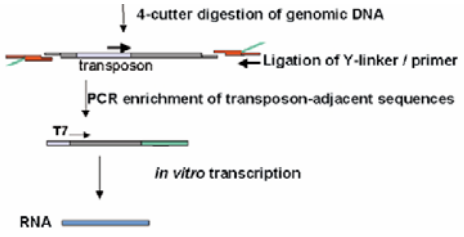


Flux data (MOMA & FBA)

Condition	Method	ρ_1	p-val (a)	p-val (b)	ρ_2	p-val (c)	p-val (d)
C-0.09	wt	0.91	8.2E-08				
	ko (FBA)	-0.064	**6.0E-01**	3.3E-03	-0.36	**9.0E-01**	2.4E-04
	ko MoMA	0.56	7.4E-03		0.48	2.3E-02	
C-0.4	wt	0.97	8.1E-12				
	ko (FBA)	0.77	8.1E-05	2.5E-03	0.36	**7.0E-02**	1.4E-02
	ko MoMA	0.94	2.6E-09		0.74	2.3E-04	
N-0.09	wt	0.78	7.1E-05				
	ko (FBA)	0.86	3.0E-06	**9.0E-02**	0.096	**3.5E-01**	4.6E-02
	ko MoMA	0.73	2.8E-04		0.49	2.0E-02	

59

Replication rate of a whole-genome set of mutants

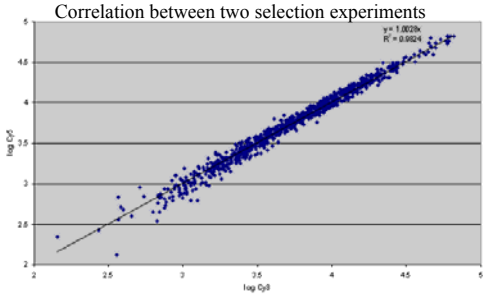


Badarinarayana, et al. (2001) Nature Biotech.19: 1060

60

Reproducible selection?

Correlation between two selection experiments



Badarinarayana, et al. (2001) Nature Biotech.19: 1060

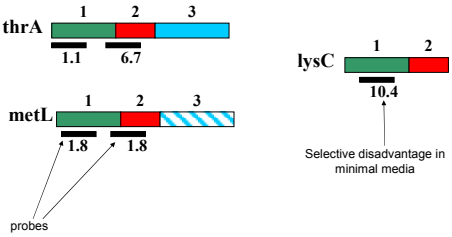
Competitive growth data

		Negative selection	no selection	X^2 p-values
FBA	Essential	142	80	62	
	Reduced growth	46	24	22	
	Non essential	299	119	180	$4x10^{-3}$
MOMA	Essential	162	96	66	
	Reduced growth	44	19	25	
	Non essential	281	108	173	$1x10^{-5}$

Position effects

Novel redundancies

Replication rate challenge met: multiple homologous domains



Net1: Simulation & optimization

- **Macroscopic continuous concentration rates**
  - **Cooperativity & Hill coefficients**
  - **Bistability**
- **Mesoscopic discrete molecular numbers**
  - **Approximate & exact stochastic**
- **Chromosome Copy Number Control**
- **Flux balance optimization**
  - **Universal stoichiometric matrix**
  - **Genomic sequence comparisons**