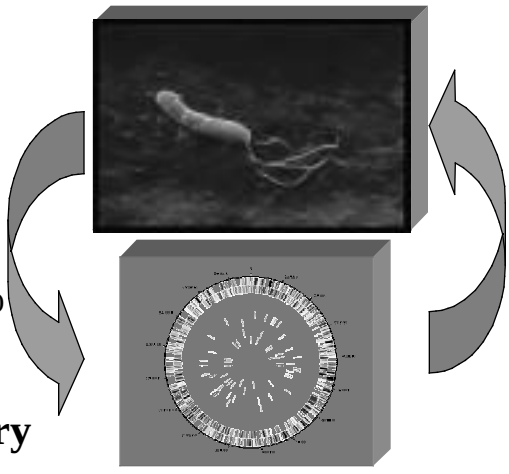
Data-driven Constraints-based models of Cellular Metabolism

Bernhard Palsson
BioEngineering, UCSD & Genomatica

Reconstructing Cellular Functions



Nature Biotech 18:1147 (2000)

Reductionistic Approach
(Genome Sequencing, DNA arrays, proteomics)

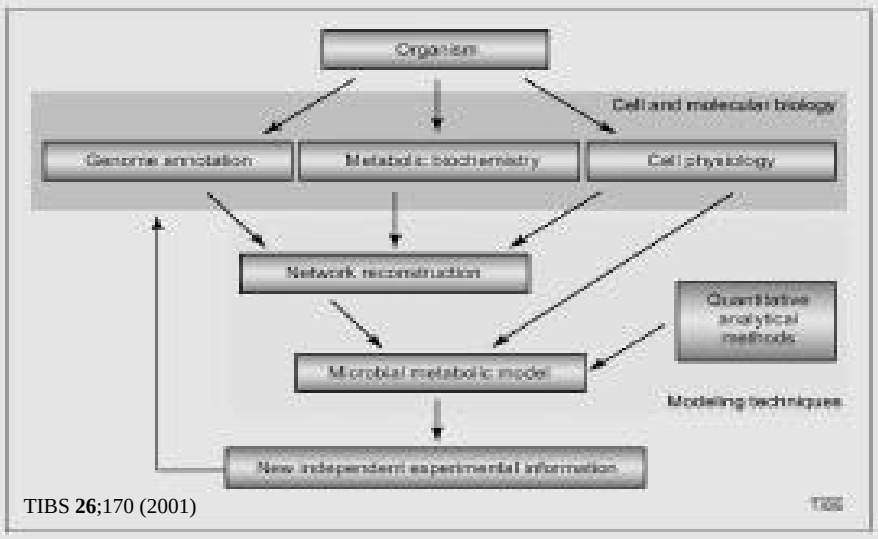
Integrative Approach
(Bioinformatics, Systems Science, modeling & simulation)

20th Century Biology

21th Century Biology

Genetic Circuits
Research Group

1. Reconstructing Metabolic Networks



TIBS **26**;170 (2001)

The Size of Reconstructed Networks

(S is metabolites x reactions)

	E. coli	*H. influ.*	*H. pylori*	*Yeast*
Sequenced	***1997***	***1995***	***1997***	***1998***
Model built	***1999***	***1998***	***2000***	***2001***
Publication	***PNAS 5/00***	***JBC 6/99***	***submitted***	***submitted***
Reactions	***720***	***488***	***444***	***1294***
Metabolites	***436***	***343***	***340***	***801***
Genes	***695***	***362***	***268***	***932***

2. Coping with incomplete constraints: solution spaces vs. single points

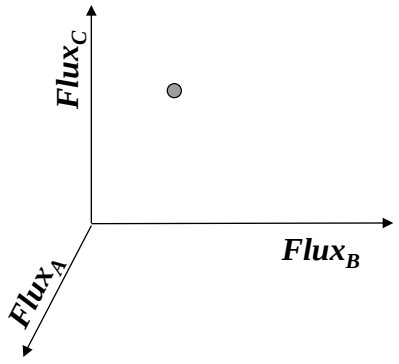
--Cannot describe cellular networks in the same detail as we are used to in the P/C sciences

1) do the P/C laws apply?
2) will never get numerical values for parameters
3) even if you do, they change with time and differ from one individual to another

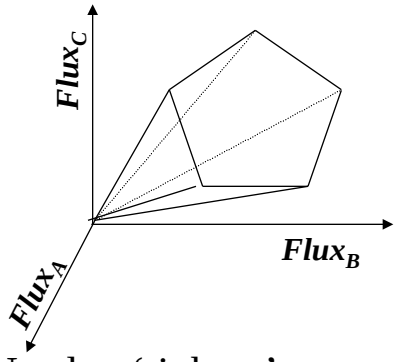
--However, we can subject the networks to known constraints and analyze them given these constraints

Incomplete Set of Metabolic Constraints

- Complete knowledge
- Solution a single point

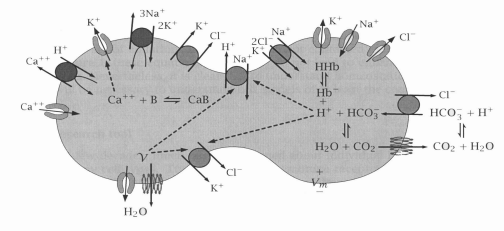


- Incomplete constraints
- Solution space



Need to 'tighten' constraints

Factors Constraining Metabolic Function



Hard

- Capacity:
  - Maximum fluxes
- Connectivity:
  - Systemic stoichiometry
- P/C factors:
  - osmotic pressure, electro-neutrality, solvent capacity, molecular diffusion

Adjustable

- Rates:
  - Mass action, Enzyme kinetics, Regulation

3. 1st Generation of Constraints-Based models

A. Genome-scale analysis of metabolic pathways

PNAS **95**: 4193-4198, (1998)
Biotechnology Progress, **15**: 296, (1999)
J. theor. Biol., **203**: 229 &249, (2000)

Polyhedral Cones and Pathways

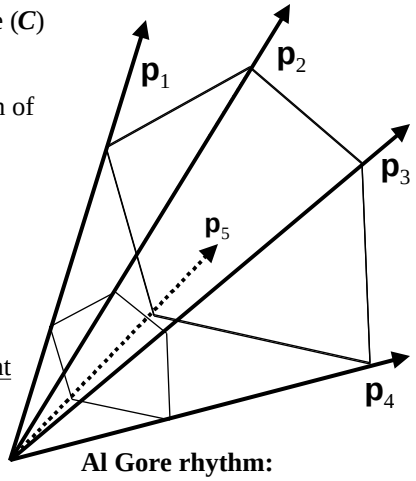
• Region determined by a linear homogeneous equation/inequality system is a convex polyhedral cone (***C***)

• Every point in the cone is a non-negative combination of the generating vectors (Extreme Pathways) of the cone

• The number of generating vectors can exceed the dimensions of the cone

• Generating vectors represent systemically independent pathways which can theoretically be "switched" on or off

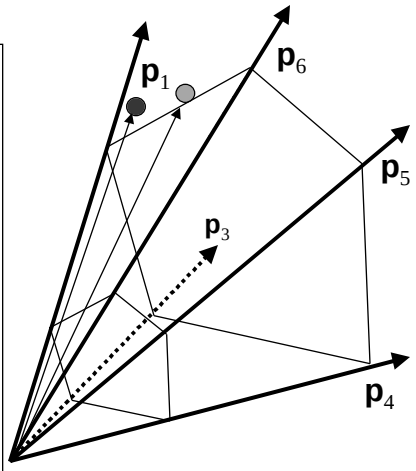
• Generating vectors are unique for a system



Al Gore rhythm:
J. theoret. Biol, 203:229 2000

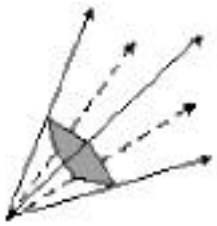
Pathway Utilization in the Red Cell:
Geometric representation

Basis Pathway	Net reaction Equation	Primary Functional Attribute
p1	glucose + 2 Pi +2 ADP -> 2 lactate +2 ATP	ATP energy production for metabolic energy
P2	glucose + 2 NAD+ -> 2 pyruvate + 2 NADH + 2 H+	NADH energy production for methemoglobin reduction
P3	glucose + 2 Pi + 2 ADP + 2 NAD+ -> pyruvate +2 ATP 2 NADH+ + 2 H+	ATP production for metabolic energy and NADH production for methemoglobin reduction
P4	glucose + 2 ATP +2 NAD+ -> 2 2,3DPG +2 Pi +2 ADP +2 NADH +2 H+	2,3DPG production for oxyhemoglobin modulation
P5	glucose + ATP + 2 NADP+ -> R5P +CO2 + ADP + 2 NADPH + 2 H+	R5P production for adenosine salvaging
P6	glucose +12 NADP+ -> 6 CO2 + 12 NADPH + 12 H+	NADPH energy production for glutathione reduction and subsequent antioxidant activity
p7-p22	(no net reaction)	



3.B. Finding Optimal Phenotypes:
LP & Flux Balance Analysis

Optimizing cellular growth **(=max likelihood of survival?)**



Convex cone

3 ↓

Bounded
convex subset

Mathematics

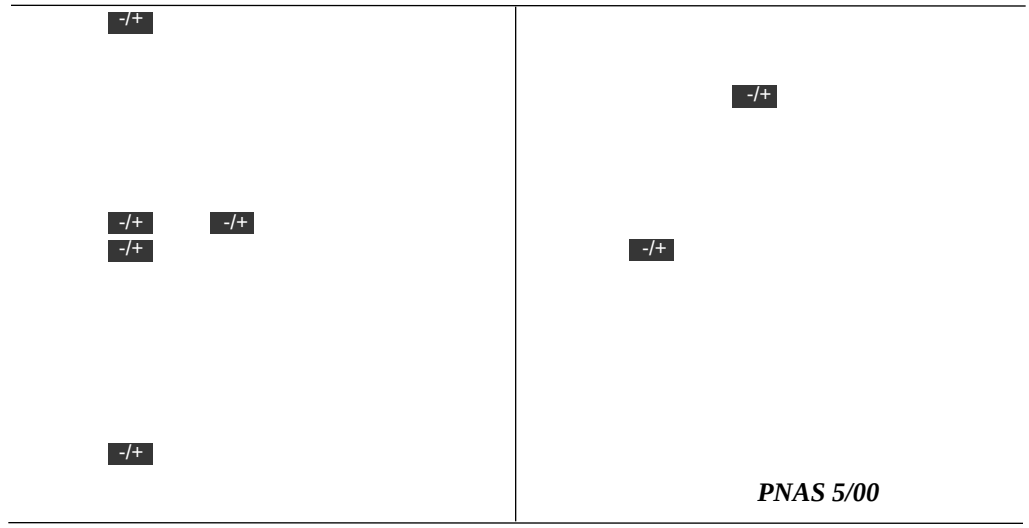
Maximize

Subject to

Genetic Circuits
Research Group

E. coli in silico vs. in vivo

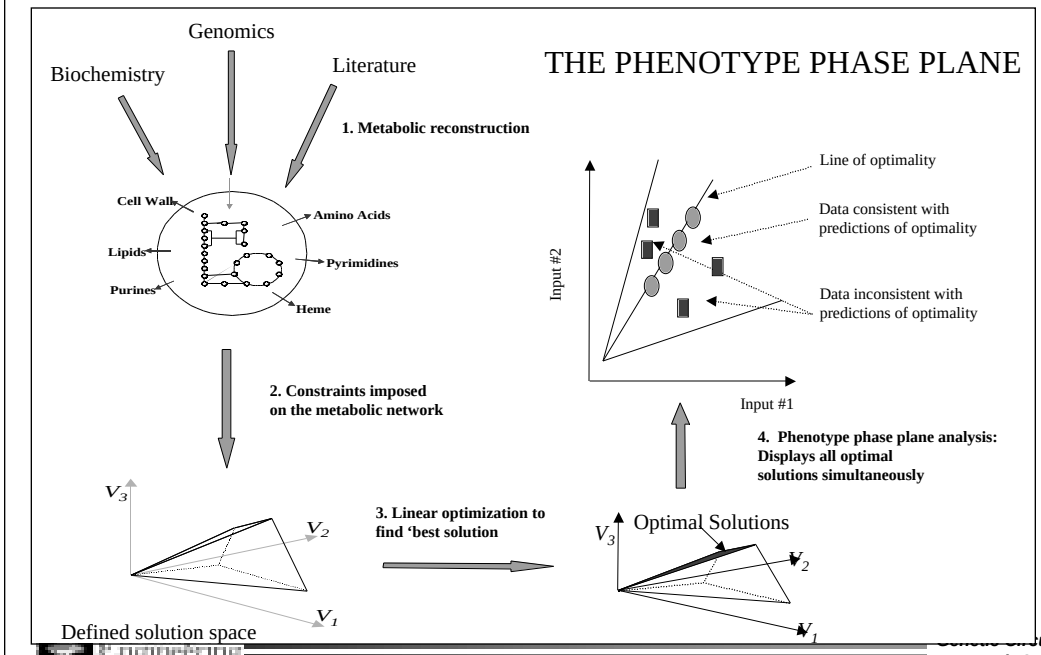
Experimental/*in silico*



PNAS 5/00

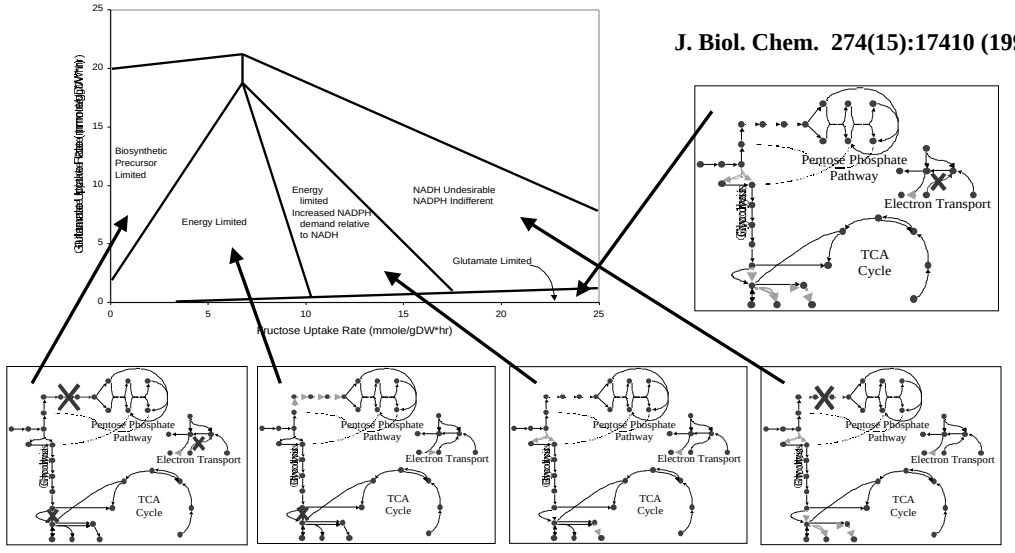
Phase Plane Analysis: phenotypic behavior under a wide range of growth conditions

Quantitative predictions



The H. influenzae Metabolic Phase Plane

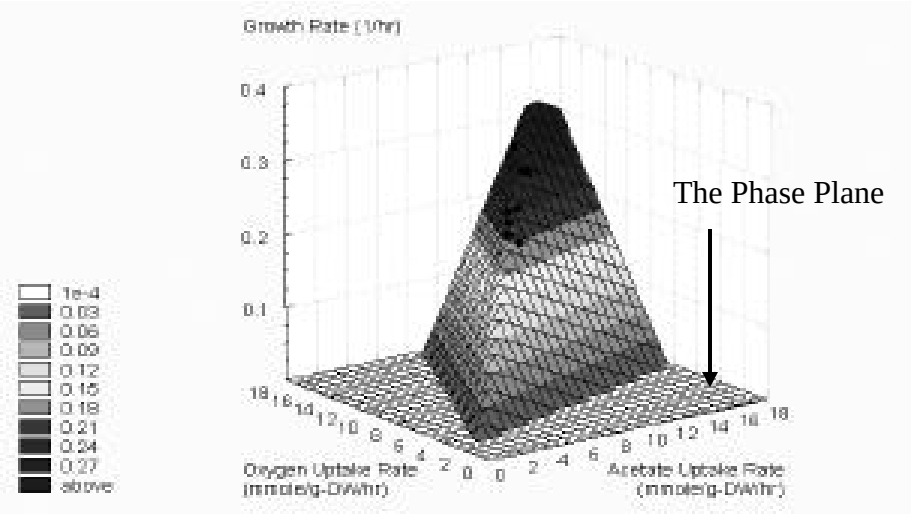
J. Biol. Chem. 274(15):17410 (1999)



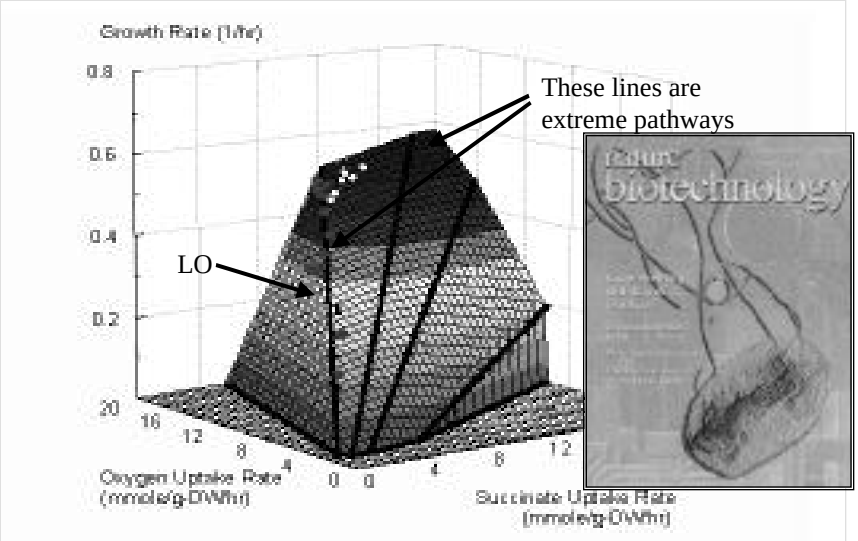
3.C. 'Living on the edge'

Formulating experimentally testable hypotheses

Growth on Acetate Graphed above Phase Plane:



Growth on Succinate



The Yeast model and predictive in silico biology

- Can predict the P/O ratio
- Can predict the ATP maintenance cost
- Can predict growth and by-product secretion patterns
- Can predict expression profiles during metabolic shifts
- Can account for the Crabtree effect
- Gene deletions (87/113 now about 480/520)

4. 2nd Generation Constraints-Based Models

Kinetic and Regulatory Constraints

Engineering Design

- Objective
  - separation of protein, building a bridge, designing a car, etc
- Constraints:
  - geometry, materials, diffusion constants, cost, time
- Design envelope
- Optimize design using free design variables
  - optimal engineering designs do evolve
  - see Detroit's industrial history museum

Engineering vs. Biological Design

- Objective
  - separation of protein
- Constraints:
  - Geometry
  - Materials
  - Diffusion constants
- Design envelope
- Optimize design using free design variables

- Objective
  - Survival, growth
- Constraints:
  - Max fluxes
  - Connectivity
  - P/C factors
- Solution space
- Optimize design using kinetic and regulatory variables

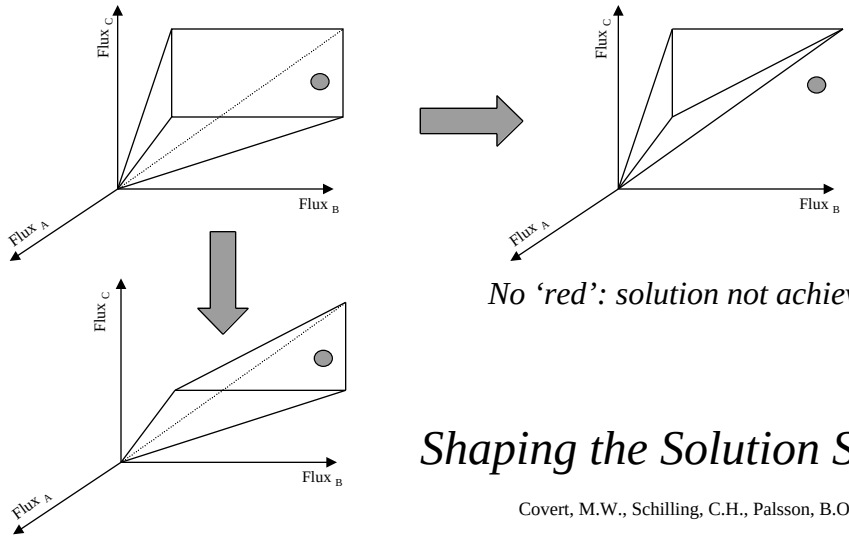
Biological Design

Regulation of expression:

shaping solution spaces

Regulation of activity:

location within a solution space



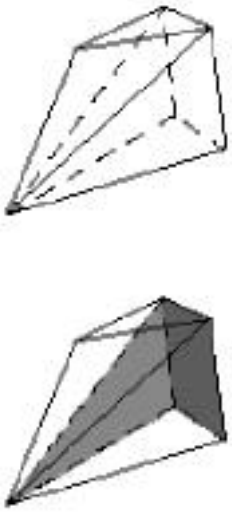
No 'red': solution not achievable

No 'blue': solution achievable but solution space smaller

Shaping the Solution Space

Covert, M.W., Schilling, C.H., Palsson, B.O.

Kinetics: locating the solution within the 'lock-box'



Bounded convex subset

4 ↓

Union of convex subsets

Some Lessons: towards principles

- The Importance of Constraints
  - Cells are constrained in their behavior and seem to push close to these constraints ('life on the edge')
    - Mass transfer limitations (P. Weisz, E.N. Lightfoot)
    - Solvent capacity (D. Atkinson)
- A large number of components (complex genotypes) display relatively few overall types of behaviors (phenotypes)

Summary:

- Active support NIH R01, NSF/KDI, NSF
- This particular program supports two Ph.D. students for *in silico* work only
- Students funded: Markus Covert, Sharon Smith, Jennifer Marciniac, Tim Allen
- Helped support the genome-scale construction of several organisms
- Helped support the synthesis of analysis methods such as the phase plane, and 'cone shrinking'