The biological design variables: kinetic and regulatory constraints

Bernhard Palsson

Hougen Lecture #6

Nov 21th, 2000

INTRODUCTION

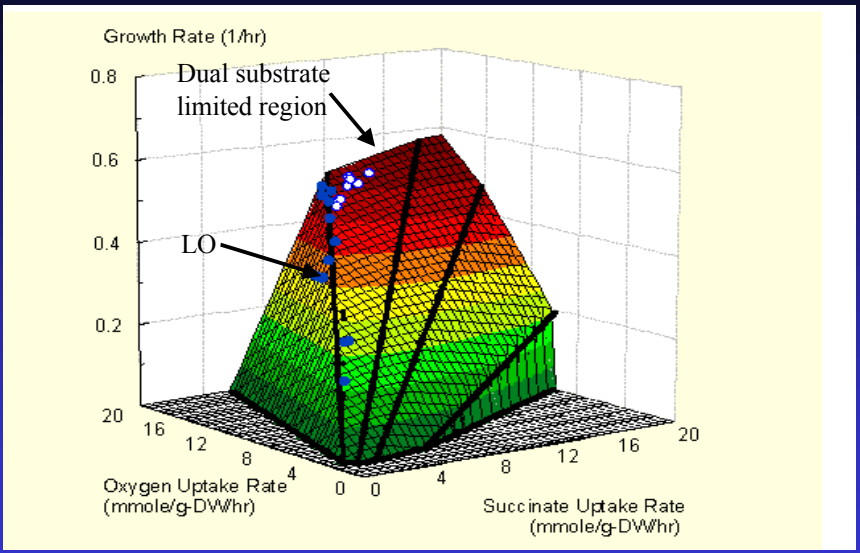
We have up to this point imposed the constraints that arise from basic physico-chemical considerations. Now we look at biological, "self-imposed" constraints.

Lecture #6: Outline

- Brief recap of Lecture #5
- E. coli as an optimizer
- Engineering vs. biological design procedures
- Accounting for regulation of gene expression:
  - Logistical and flux balance representation
  - Examples: multiple substrates
- Dealing with kinetics
  - Numerical values of kinetic constants
  - Relative values
  - Temporal decomposition
- Numerics

LECTURE #6

Succinate 3-D Phenotype Phase Plane



This also works for other substrates!

The case of succinate

This figure shows the succinate-oxygen PhPP in three dimensions.

•The formalism is similar to the 3-D acetate PhPP

•Here the effect of the carbon source on the structure of the PhPP can be seen.

•The LO is shown here, and the data points with reduced succinate uptake rates all lie on (or near) the LO,

•However, when the succinate uptake rate was increased, the experimental data followed the LO until the oxygen mass transfer constraint was reached. At this point, the growth rate and the succinate uptake were increased by moving into region 2 of the phase plane.

•How do cells find this optima?

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

Engineering design begins with a statement of an objective; i.e. separating a protein or building a bridge. The constraints on the design are then defined. Cost and time are always important, but so are material properties (strength, elasticity, etc), physical constants (diffusivities, thermal conductivities), and geometric considerations. These constraints then define a design envelope within which the design must fall. Optimization of the design is then carried out within the allowable ranges to produce the 'best' design.

Constraints on biological networks

- Stoichiometry
- Maximum Capacities
- P/C constraints
  - Diffusion, electroneutrality
- Kinetics/Regulation

→ • Non-adjustable
  - Horizontal gene transfer

→ • Upper limit
  - Downwardly adjustable by gene expression

→ • Non-adjustable

→ • Highly adjustable
  - Evolutionary design

Engineering vs. Biological Design

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

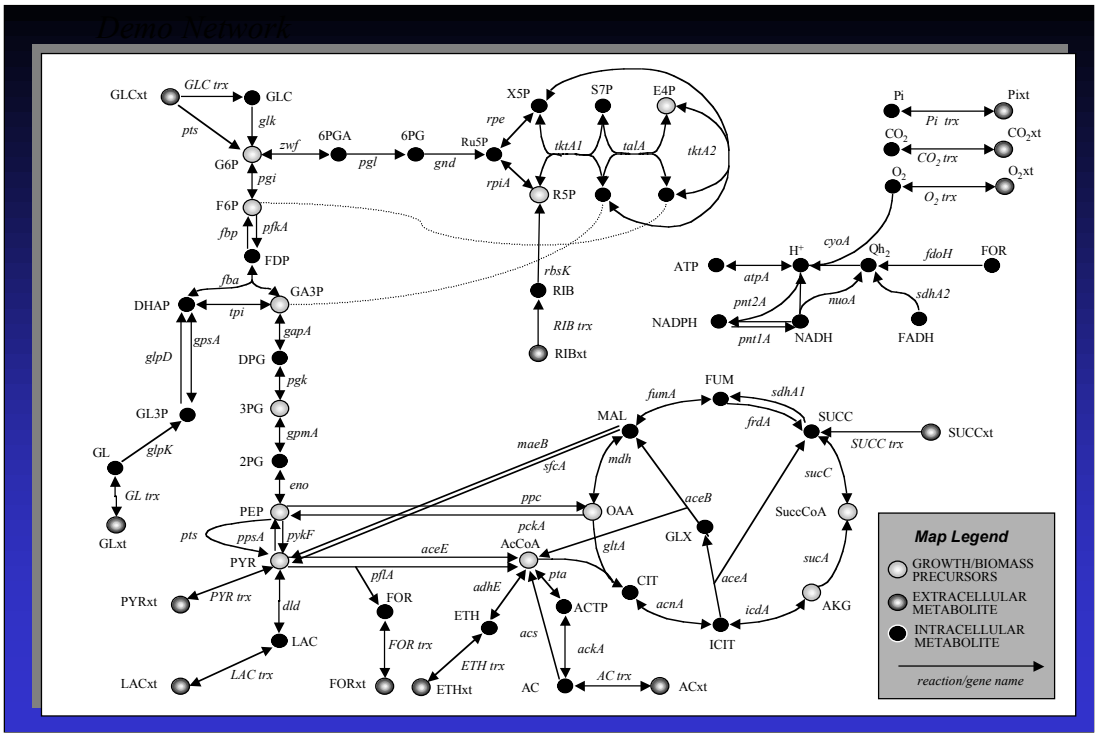
- Objective
  - Survival, growth
- Constraints:
  - Max fluxes
  - Connectivity
  - P/C factors
- Solution space
- Optimize design using kinetic and regulatory variables

There is some uncertainty about how to apply the basic physical laws in the intra-cellular milieu and even if we knew how, we would not have numerical values for the myriad of constants that appear in such equations. The alternative approach relies on the successive imposition of constraints that govern biochemical reaction networks. Such constraints include the maximum flux achievable through a reaction, the connectivity of the network and so forth. The imposition of these constraints defines a solution space, similar to the design envelope discussed above. The 'best' solution in the allowable solution space is then determined based on an optimization procedure. The optimization is based on an assumed objective that the cell is striving to meet. A match has been obtained between measured growth and metabolic by-product secretion of E. coli K-12 for growth on acetate and succinate and the calculated optimal performance based on the constraint-based approach.

Biological Design

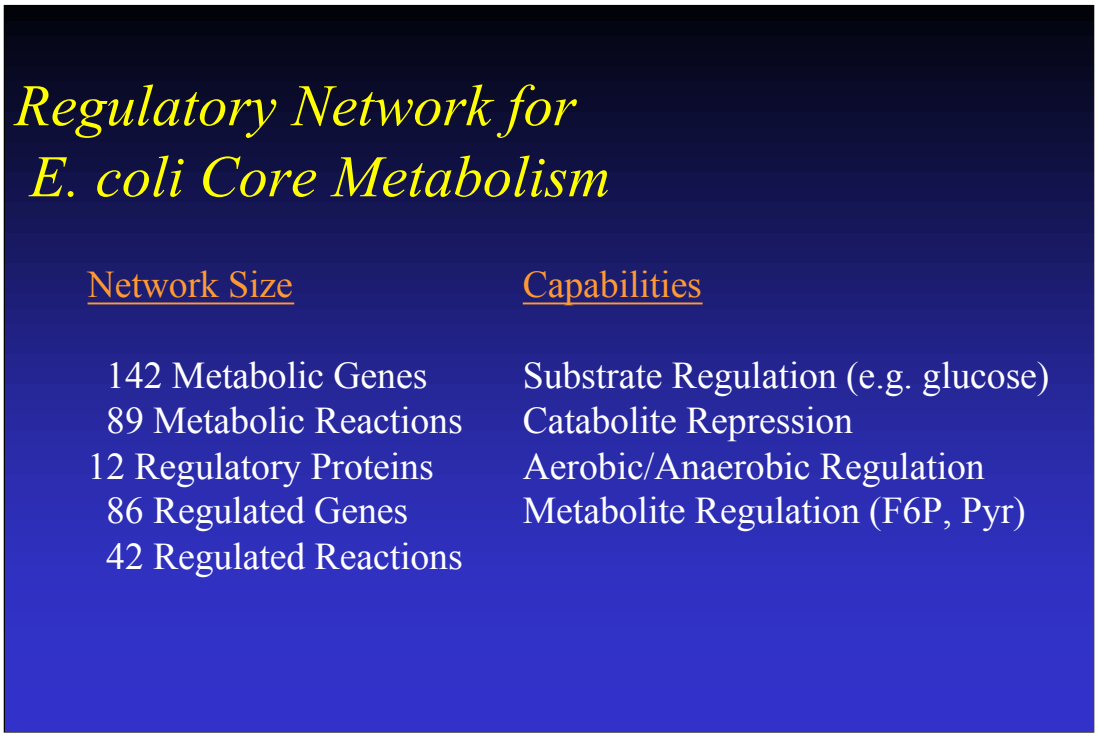
Regulation of expression:
shaping solution spaces
Regulation of activity:
location within a solution space

Given the solution space that is determined in part by hard physicochemical constraints, the exact solution is determined by the kinetic and regulatory parameters that the cell can alter. Thus, we can now view the kinetic and regulatory parameters as 'biological design' variables, based on an analogy with the engineering design procedure. In order for this analogy to hold and to view the kinetics as biological design variables, we must be able to observe the evolutionary motion of a suboptimal design towards an optimal under the given constraints.



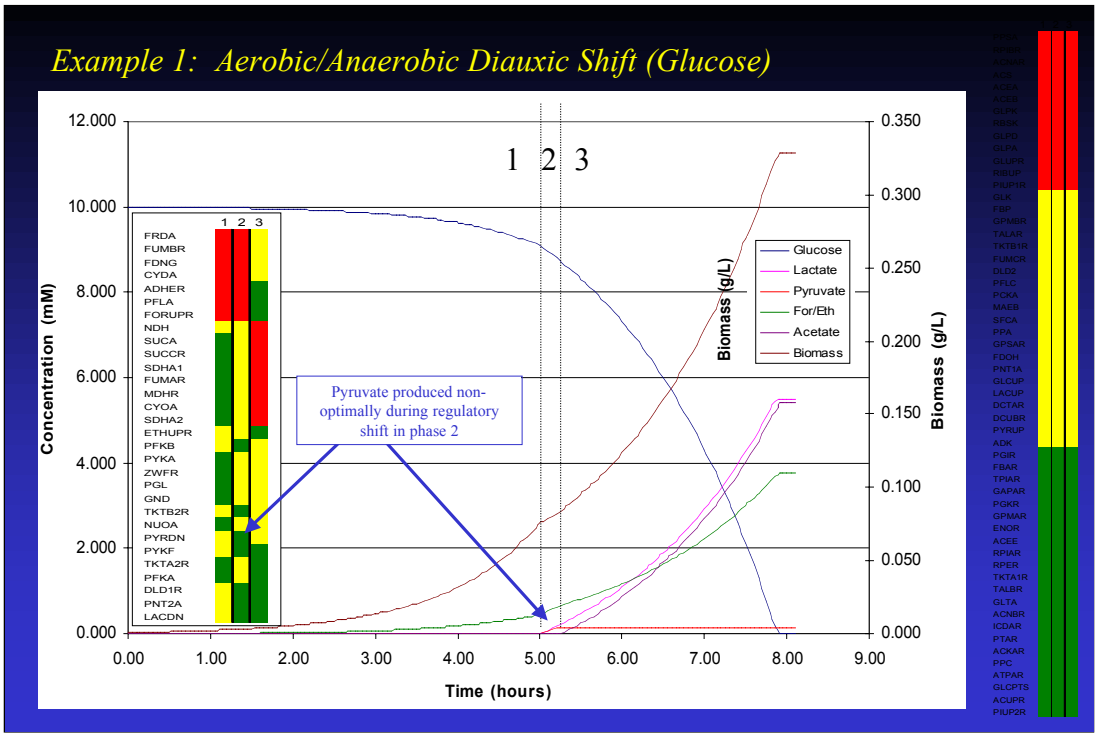
Logistical -FBA Models

Known regulatory effects can be used to close off or open links in the network. The known operon structure for *E. coli* can be used to implement a condition-dependent map available to the cell.

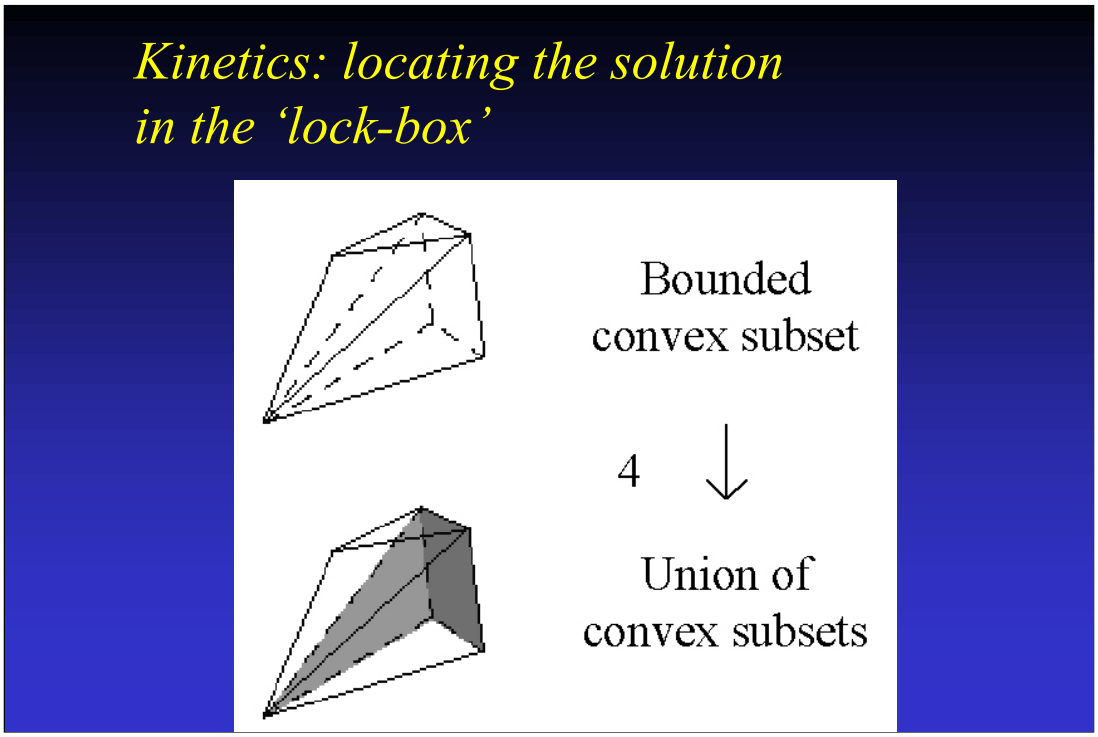


SPECS

These are the specifications on the regulated core *E. coli* metabolic model.



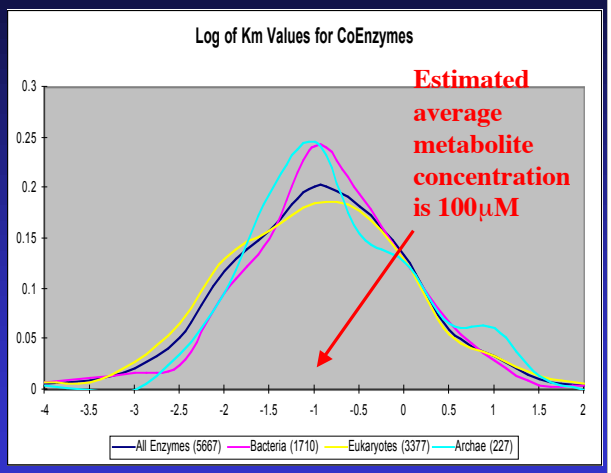
Dynamic simulations of the regulated *E. coli* model. The bar to the left shows changes in gene expression, while the expression of the genes described in the bar on the right does not change.



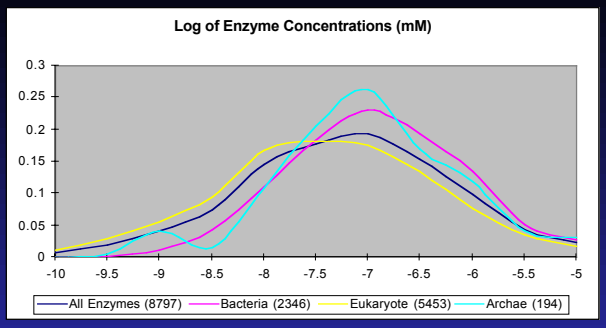
Regulation of gene expression and maximal flux constraints close-off a solution space. The exact location of the solution in the 'lock-box' will be determined by the numerical values of the kinetic constraints.

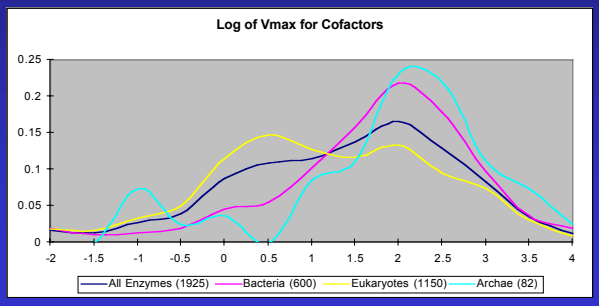
Numerical values of kinetic constants

- Compilations of legacy data
  - i.e. EMP data base
- Determine how well we need to know the kinetic parameters
  - Order-of-magnitude



Enzymes

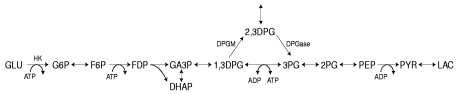

Log of Enzyme Concentrations (mM)
0.3
0.25
0.2
0.15
0.1
0.05
0
-10
-9.5
-9
-8.5
-8
-7.5
-7
-6.5
-6
-5.5
-5
All Enzymes (8797)
Bacteria (2346)
Eukaryote (5453)
Archae (194)


Log of Vmax for Cofactors
0.25
0.2
0.15
0.1
0.05
0
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
3
3.5
4
All Enzymes (1925)
Bacteria (600)
Eukaryotes (1150)
Archae (82)

Orders of Magnitude: Kinetics and edges of solution cones: Use of dimensionless groups

RBC Network:

(a)



Extreme Pathways:

(b)

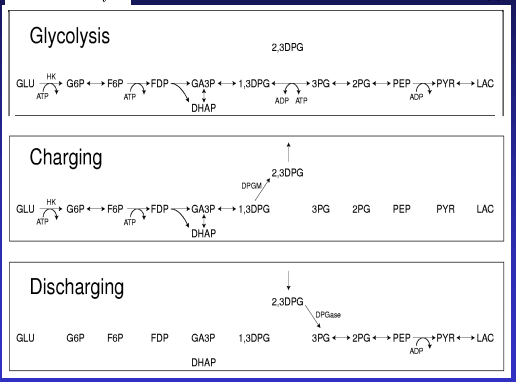


Figure 9. *(A) The simplified red blood cell reation network comprised of only glycolysis and Rapoport-Leubering shunt. (B) The three extreme pathways for this network; glycolysis, charging, discharging.*

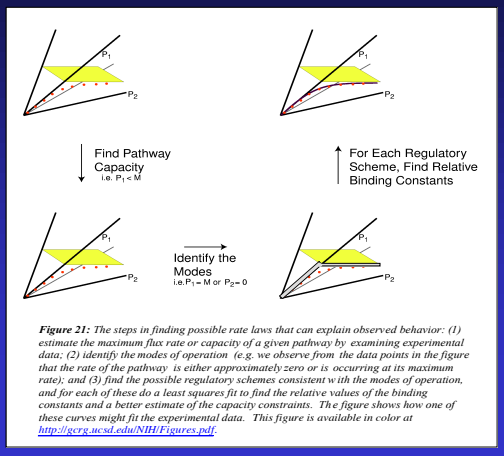
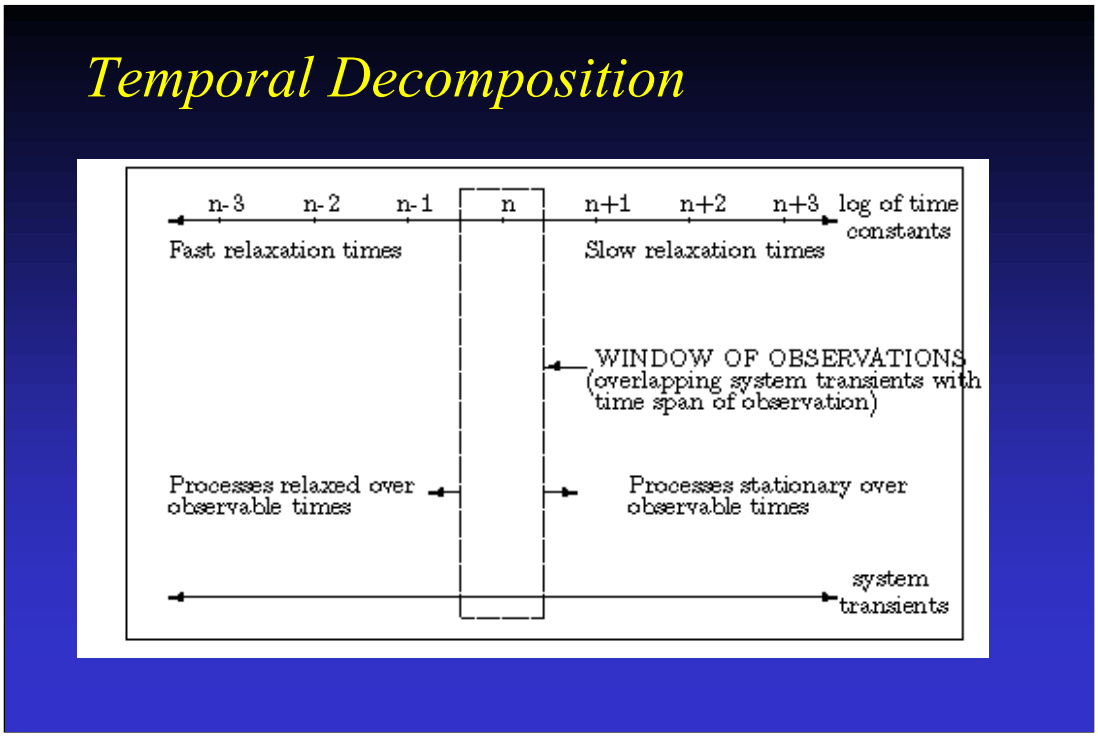


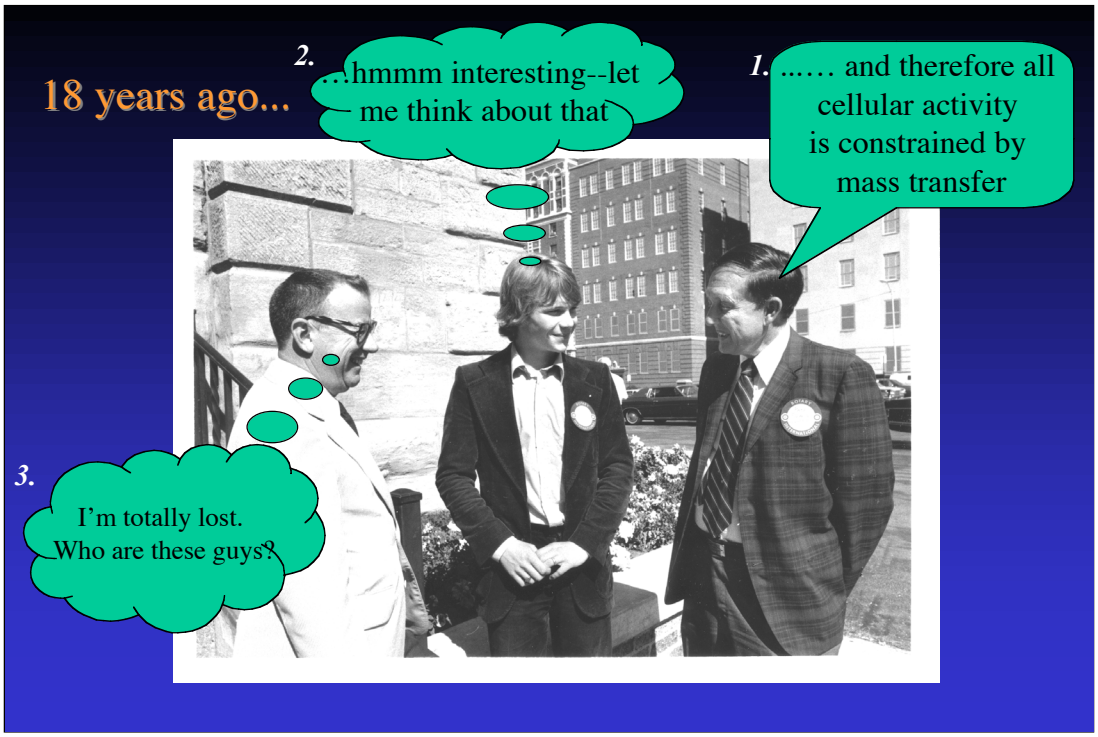
Figure 21: *The steps in finding possible rate laws that can explain observed behavior: (1) estimate the maximum flux rate or capacity of a given pathway by examining experimental data; (2) identify the modes of operation (e.g. we observe from the data points in the figure that the rate of the pathway is either approximately zero or is occurring at its maximum rate); and (3) find the possible regulatory schemes consistent w ith the modes of operation, and for each of these do a least squares fit to find the relative values of the binding constants and a better estimate of the capacity constraints. The figure shows how one of these curves might fit the experimental data. This figure is available in color at* http://gcrg.ucsd.edu/NIH/Figures.pdf.



TEMPORAL DECOMPOSITION

The hierarchy of intrinsic times can be represented by the time axis. Fast transients are characterized by the processes at the extreme left and slow transients at the extreme right. The process time scale, i.e. the time scale of interest, can be represented by a *window of observation* on this time axis. One can conceptualize this readily by looking at a three-dimensional system where one time constant represents the fast motion; the second, the time scale of interest; and the third, a slow motion.

The terms which have time constants faster than the observed window can be eliminated from the dynamic description as these terms are small. However, the mechanisms which have transients slower than the observed time exhibit high "inertia"' and hardly move from their initial state and can be considered constants. One can thus remove slow or fast terms by the appropriate use of the eigenrows and eigenvectors.



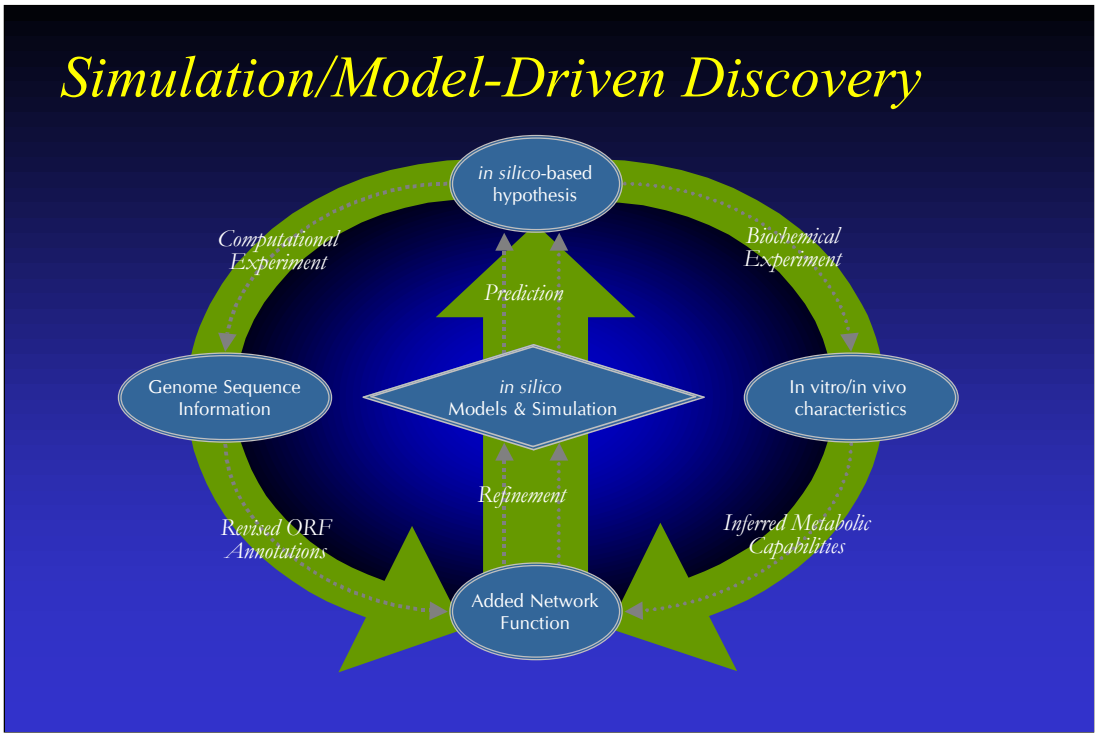
A Personal Reflection

Some Lessons: towards principles

- Importance of Constraints
  - Cells are constrained in their behavior and seem to push close to these constraints ('life on the edge')
  - Extension of the concept of Mass Transfer limitations
    - (E.N. Lightfoot)
- A large number of components (complex genotypes) display relatively few overall types of behaviors (phenotypes)

Simplicity from complexity: the evidence mounts

- Singular value decomposition of genome-scale expression data is in uncovering simple underlying patterns
- Modal analysis of dynamic models of metabolism shows simple dynamic structures
- Robustness analysis of kinetic models of biochemical systems models reveals insensitivity to individual kinetic constants



The model building process is an iterative one. We must learn to embrace failure.

Summary

- Metabolic genotypes can be formulated based on annotated sequence data
- Using the biochemical properties of the gene products and other information, a genome-scale metabolic network can be formulated
- Flux distributions through this network cannot be uniquely calculated, but optimal phenotypes can
- Testable experimental hypotheses can be generated in this way and have been put forth for *E. coli* growth on acetate and succinate
- Further testing is needed to assess the generality of the approach
- It forms the basis for iterative model building within the framework of applying successive constraints

--The End--

Hougen 2000 Lectures