Mathematical Representation of Cellular Systems

David Fell

OXFORD BROOKES UNIVERSITY

dfell@brookes.ac.uk

http://mudshark.brookes.ac.uk

OXFORD BROOKES UNIVERSITY

Outline

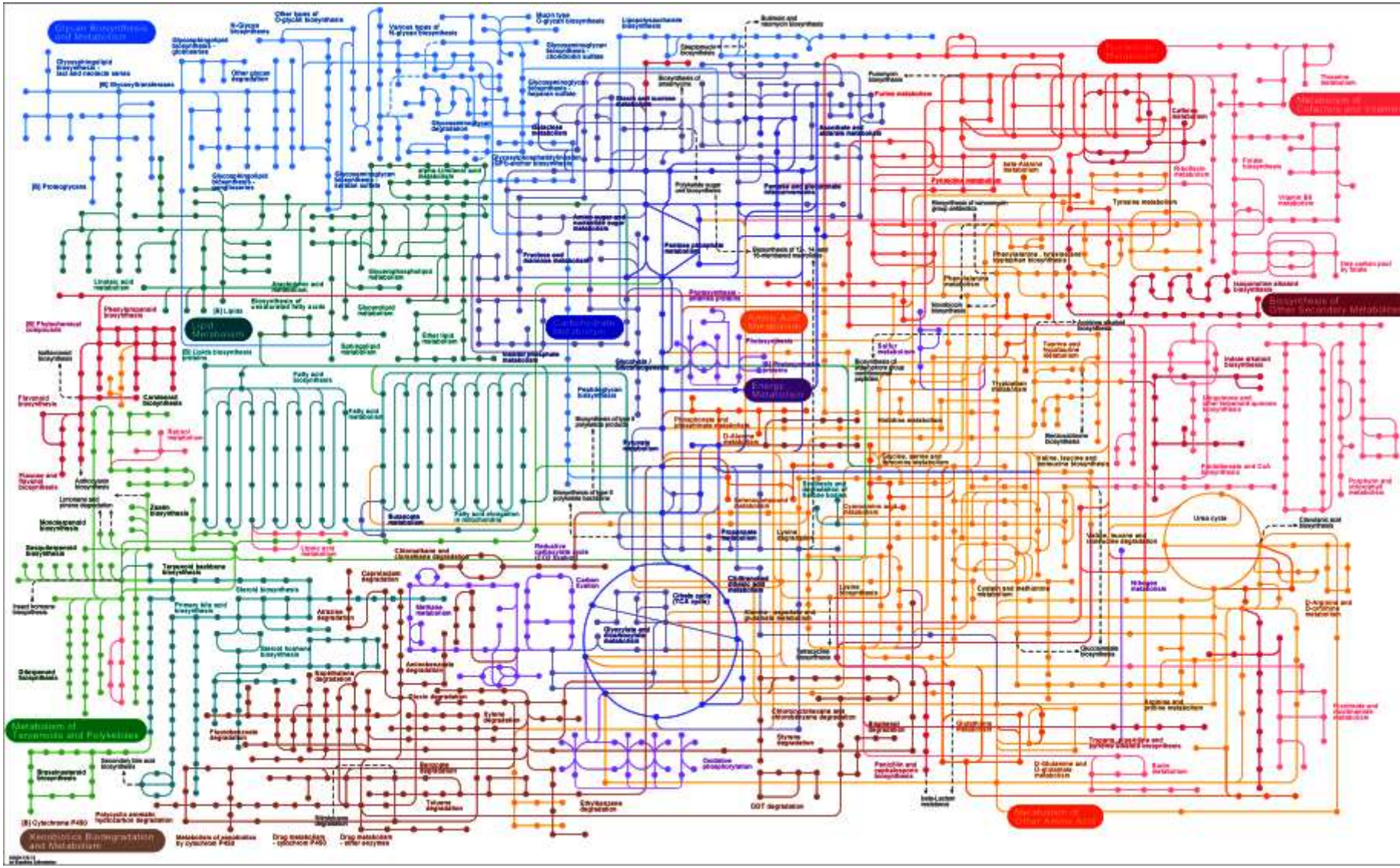
- Formal representation of metabolic networks
- Structural modelling
- Kinetic modelling
- Summary
- Additional information

OXFORD BROOKES UNIVERSITY

Formal Representation of Metabolic Networks

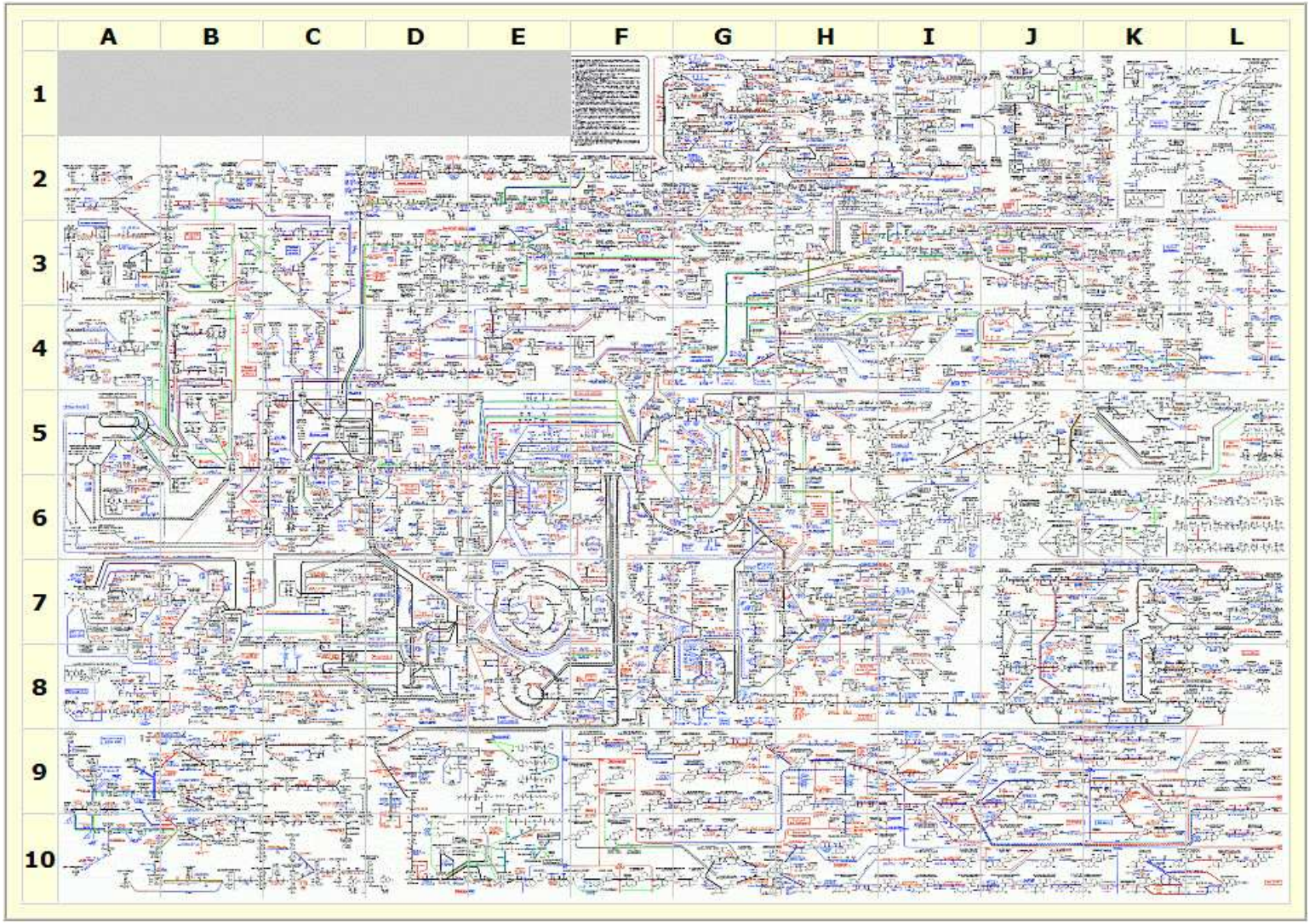
OXFORD BROOKES UNIVERSITY

The Metabolic Network



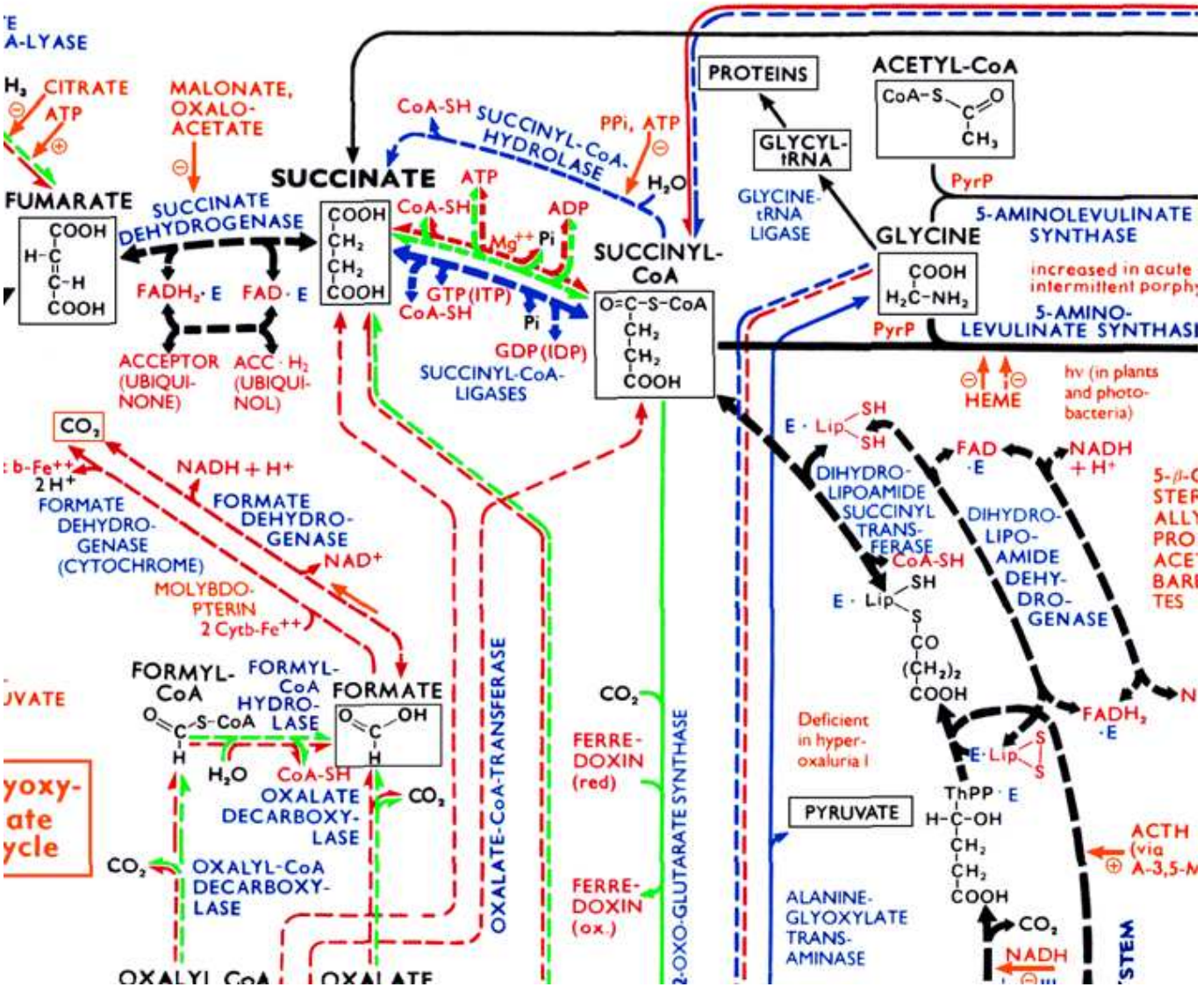
OXFORD BROOKES UNIVERSITY

The Metabolic Network — More Detail



From Expasy Biochemical Pathways: http://www.expasy.ch/cgi-bin/search-biochem-index

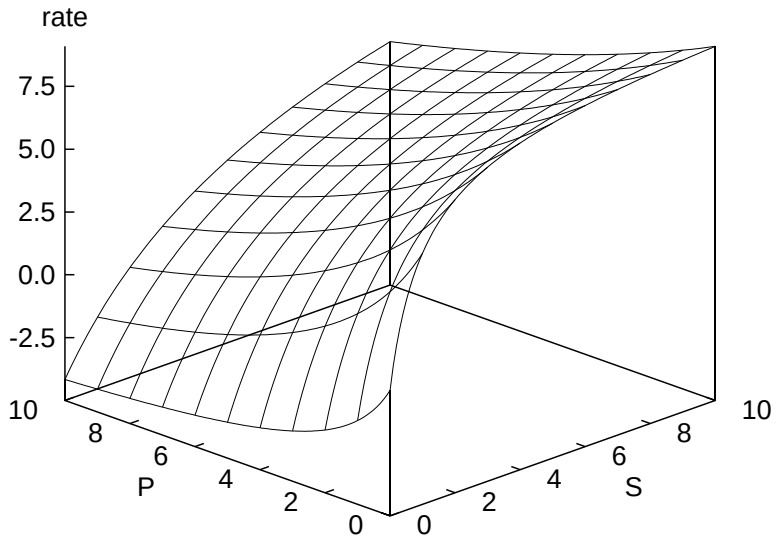
The Metabolic Network — Zoom in



From Expasy Biochemical Pathways: http://www.expasy.ch/cgi-bin/search-biochem-index

The Reversible M–M Eqn.

$$v_{net} = \frac{(V_{\mathrm{f}}/K_{\mathrm{m,S}})\,(S - P/K_{\mathrm{eq}})}{1 + S/K_{\mathrm{m,S}} + P/K_{\mathrm{m,P}}}$$



Simultaneous dependence of enzyme rate on both substrate and product. The parameters have been set to: $K_{\mathrm{m,S}}$ =1; $V_{\mathrm{m,f}}$ = 10; $K_{\mathrm{m,P}} = 2$, and $K_{\mathrm{eq}} = 4$.

The Reaction List

Where does the data for the structure of a metabolic model come from?

- Biochemical literature: books, reviews, journal articles.

The Reaction List

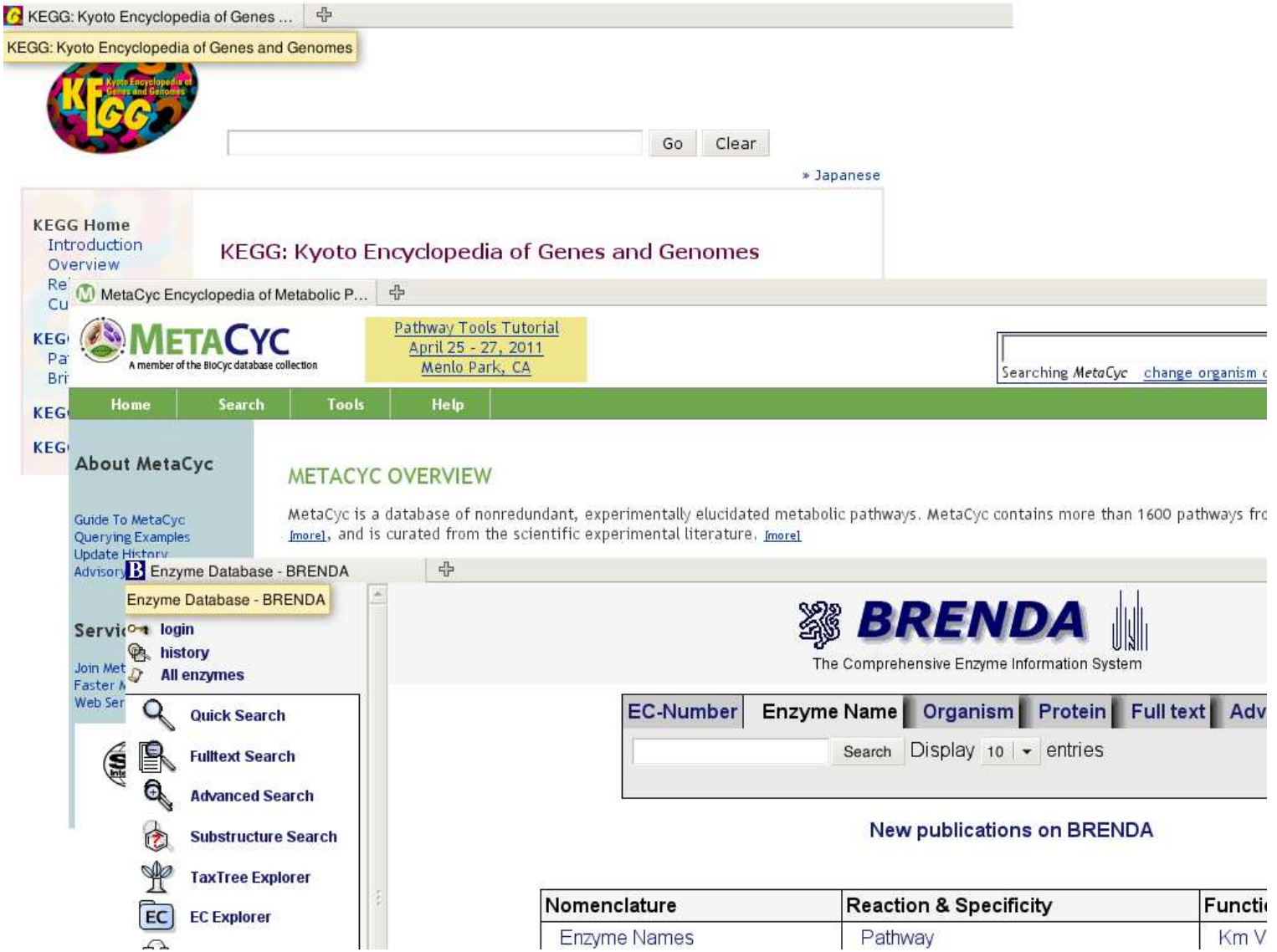
Where does the data for the structure of a metabolic model come from?

- Biochemical literature: books, reviews, journal articles.

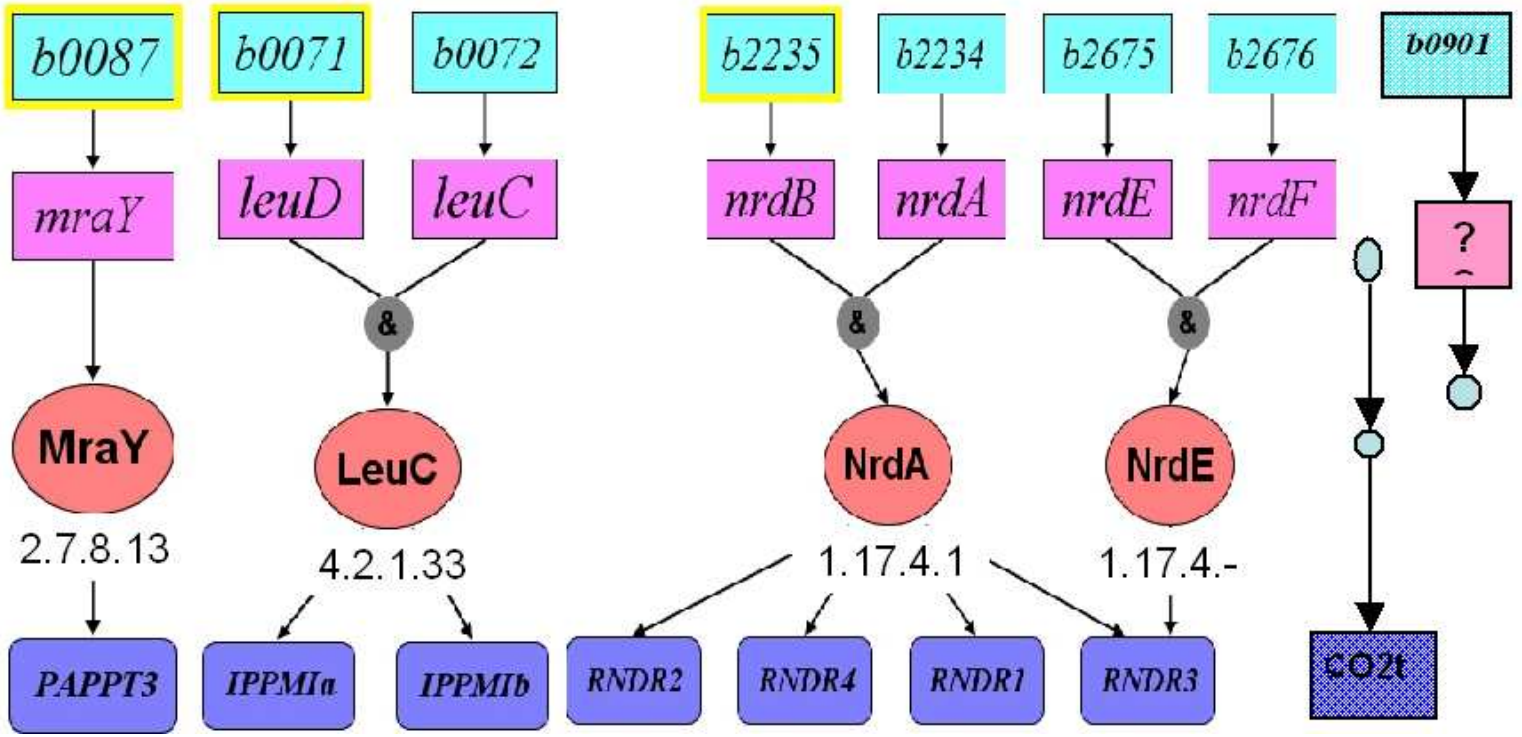
- Genome databases plus annotation plus enzyme database.
  - BioCyc: http://biocyc.org/
  - KEGG: http://www.genome.jp/kegg/
  - IntEnz at http://www.ebi.ac.uk/intenz/ (now the primary source) or
    EXPASY Enzyme: http://www.expasy.org/enzyme/
  - Brenda: http://www.brenda-enzymes.info

OXFORD BROOKES UNIVERSITY

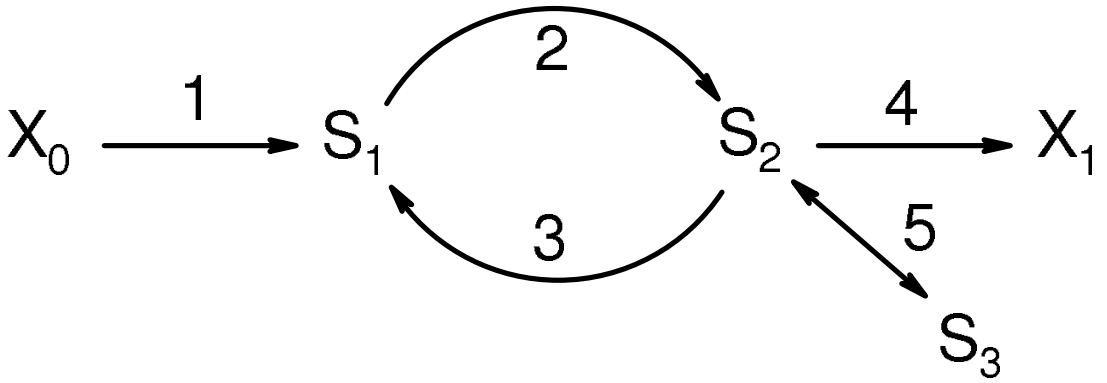
Data Sources



Genes to Reactions



Modelling: a Specimen Network



X_0 is termed the *source*, and X_1 is the *sink*. They are also termed *external metabolites*.

S_1, S_2 and S_3 are the variable, or *internal* metabolites that reach constant levels at steady state, when their rates of formation equal their rates of utilization.

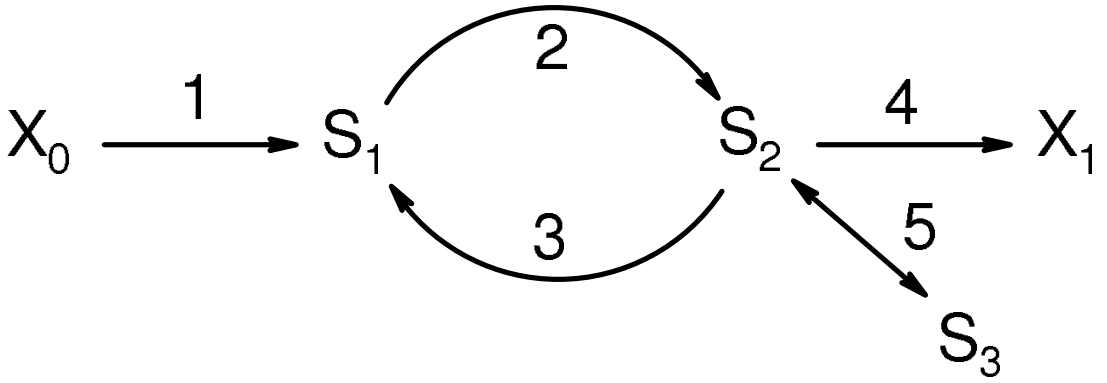
Steady State

In a metabolic network there is a flow of matter from the *source* to the *sink*. At steady state, the concentrations of the intermediates remain constant because their rates of formation exactly equal their rates of degradation. The flow through the pathway also remains constant.
metabolites, or the pathway flux, because of slow changes in the

If there are very slow changes in the concentrations of source or sink, the pathway may be regarded as being in *quasi steady state* provided the time scale of the changes is very much longer than the time taken by the pathway to approach steady state.

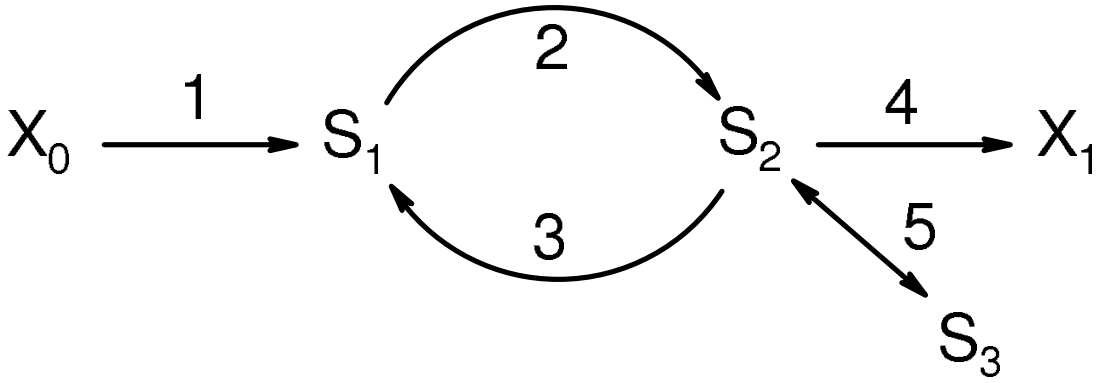
Representation

Consider a simple network, e.g.:



Representation

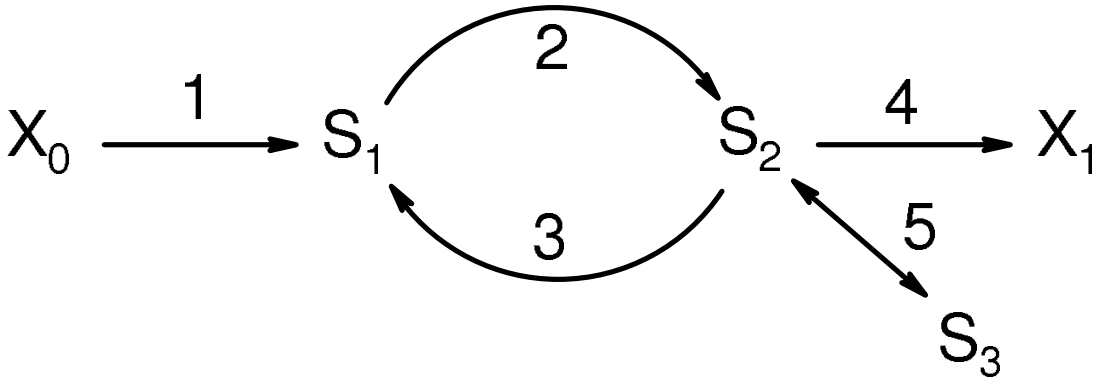
Consider a simple network, e.g.:



$$\begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \end{array} \begin{bmatrix} 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Representation

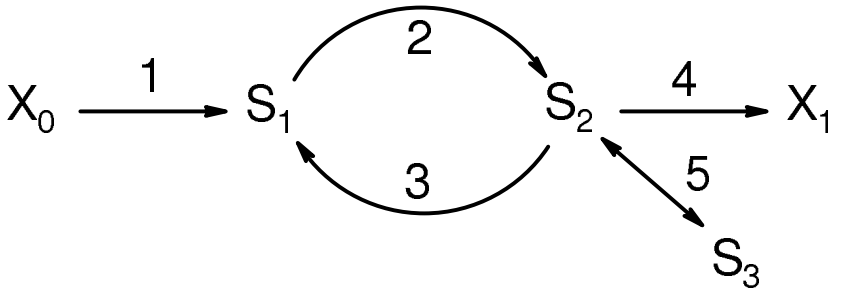
Consider a simple network, e.g.:



$$\begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \end{array} \begin{bmatrix} 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

r1: X_0 -> S_1 ~
r2: S_1 -> S_2 ~
r3: S_2 -> S_1 ~
r4: S_2 -> X_1 ~
r5: S_2 -> S_3 ~

Kinetics of the Metabolites



By inspection of the diagram:

$$\frac{dS_1}{dt} = v_1 - v_2 + v_3$$
$$\frac{dS_2}{dt} = v_2 - v_3 - v_4 - v_5$$
$$\frac{dS_3}{dt} = v_5$$

How can we generalise this?

Separation of Structure and Kinetics

The rate at which the substrate concentrations are changing is given by $\mathbf{N}.\mathbf{v}$, where $\mathbf{N}$ is the stoichiometry matrix, and $\mathbf{v}$ is a vector of enzyme kinetic functions. So for our substrate cycle network:

$$\begin{bmatrix} \frac{dS_1}{dt} \\ \frac{dS_2}{dt} \\ \frac{dS_3}{dt} \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{bmatrix}$$

where each v_i is the rate function for enzyme i, depending on the variable metabolites and the parameters $V_{\mathrm{m,i}}$, $K_{\mathrm{m,i}}$ etc, as $f_i(\mathbf{S})$.

Integrating this set of non–linear differential equations gives a dynamic, or **kinetic** model of our network.

Steady State Solutions

Any metabolic network at steady state satisfies the relationship $\mathbf{N}.\mathbf{v} = \mathbf{0}$, where $\mathbf{N}$ is the stoichiometry matrix, exemplified by our model network:

$$\begin{matrix} S_1 \\ S_2 \\ S_3 \end{matrix} \begin{bmatrix} 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Again, regarding each v_i as a function of the metabolite concentrations, $f_i(\mathbf{S})$ defines a set of non–linear simultaneous equations that can be solved for the steady state values of $\mathbf{S}$. Alternatively, the vector of rates $\mathbf{v}$ can be regarded as the variable, in which case this describes a set of under–determined equations defining constraints on feasible values of the v_i at steady state. This is the basis of **structural modelling** of metabolism.

Metabolic Flux Analysis

If a number of the rates $\mathbf{v}$ are measured (e.g. as input and output fluxes), and arranged as a partition of the rate vector $\mathbf{v_m}$, leaving the unknown (internal) rates in the partition $\mathbf{v_u}$, then:

$$\mathbf{0} = \mathbf{N}.\mathbf{v} = \mathbf{N_m}.\mathbf{v_m} + \mathbf{N_u}.\mathbf{v_u}$$

Hence:

$$\mathbf{N_u}.\mathbf{v_u} = -\mathbf{N_m}.\mathbf{v_m}$$

If this equation can be solved, then (some of) the unknown rates can be expressed in terms of the known rates.
But what can we say about the feasible rates if this equation cannot be solved?

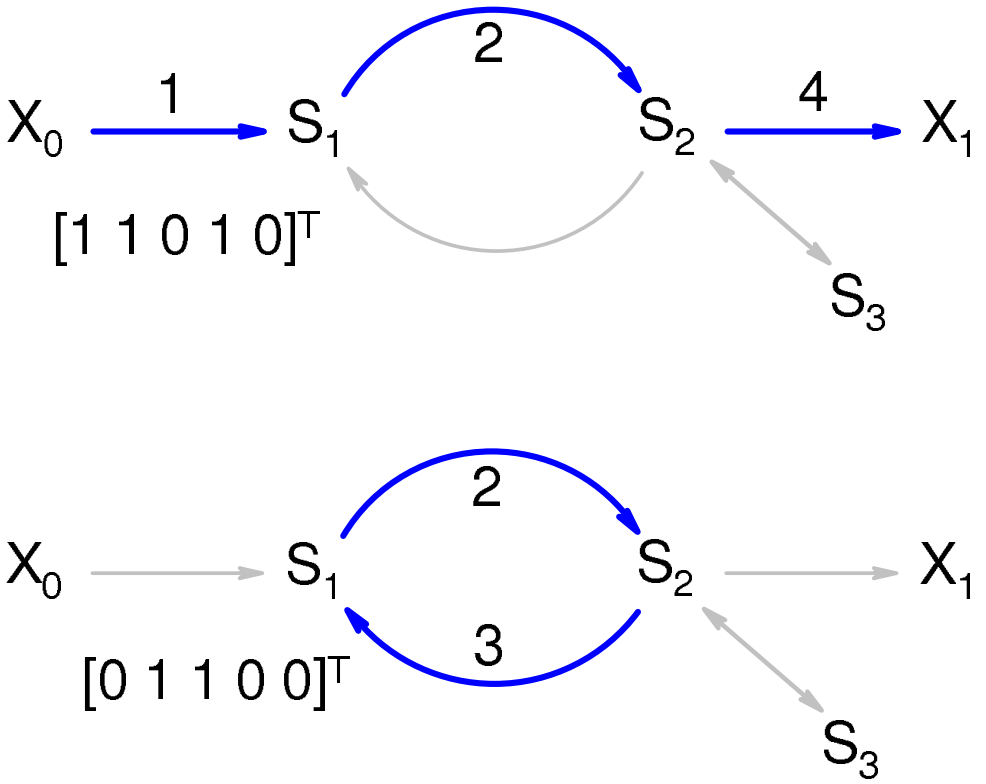
Structural Analysis: Null Space Vectors

Any observed set of velocities at steady state will be a linear combination of a set of vectors $\mathbf{K}$ referred to as a basis for the null space of the stoichiometry matrix. In this case:

$$\mathbf{K} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}$$

The null space can be computed from the stoichiometry matrix using standard algorithms.

Null Space Vectors as Pathways



$[1 \;\; 1 \;\; 0 \;\; 1 \; 0]^T$ and $[0 \;\; 1 \;\; 1 \;\; 0 \; 0]^T$

The Null Space and Pathway Fluxes

Any feasible set of velocities at steady state is a linear combination of these null space vectors, e.g.:

$$\mathbf{K} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}$$

The Null Space and Pathway Fluxes

Any feasible set of velocities at steady state is a linear combination of these null space vectors, e.g.:

$$\mathbf{K} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}$$

and:

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a \\ a+b \\ b \\ a \\ 0 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{bmatrix}$$

The Null Space and Pathway Fluxes

Any feasible set of velocities at steady state is a linear combination of these null space vectors, e.g.:

$$\mathbf{K} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{matrix} \leftarrow \text{subset} \\ \\ \\ \leftarrow \text{subset} \\ \\ \end{matrix}$$

and:

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a \\ a+b \\ b \\ a \\ 0 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{bmatrix}$$

The Null Space and Pathway Fluxes

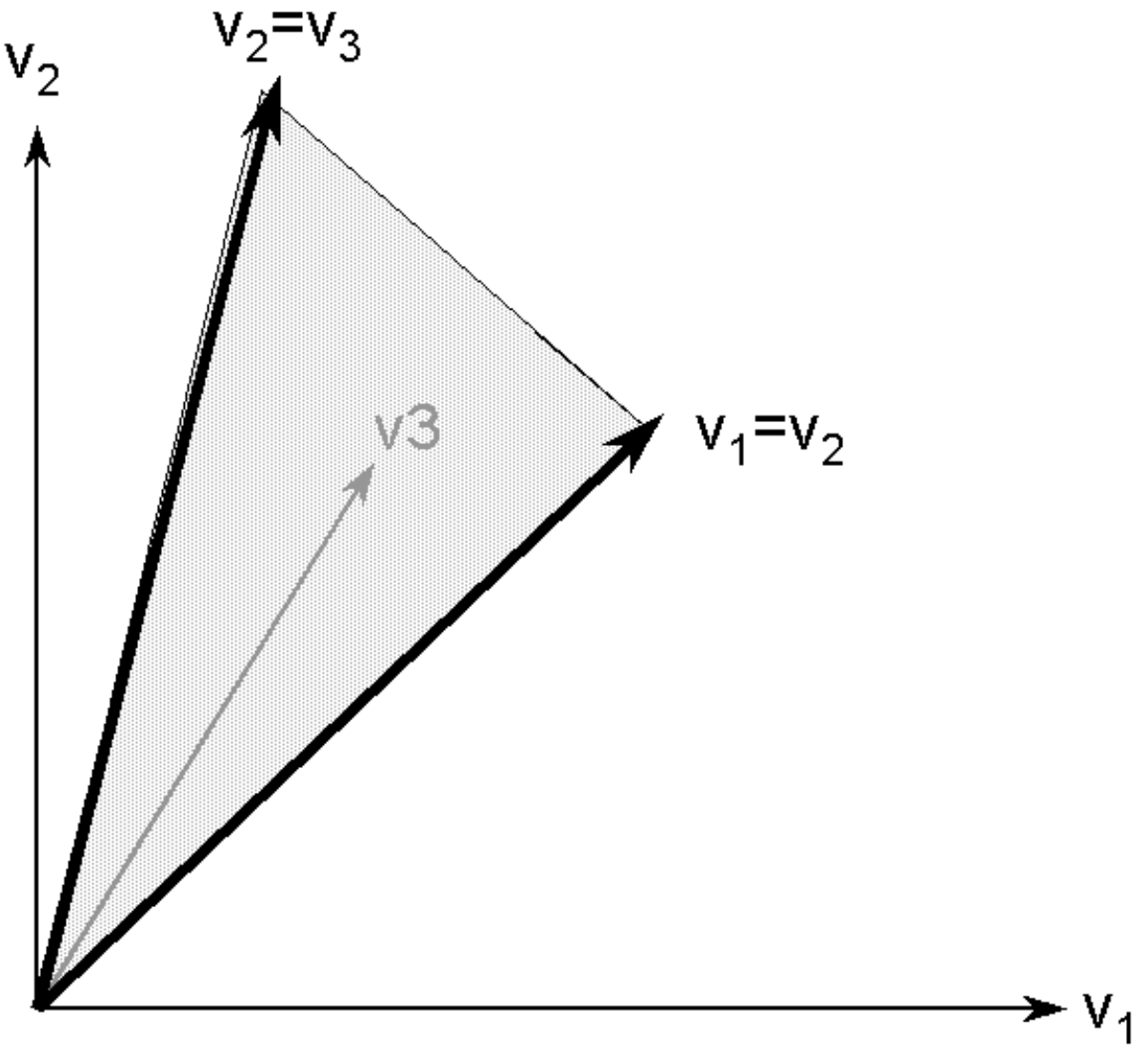
Any feasible set of velocities at steady state is a linear combination of these null space vectors, e.g.:

$$\mathbf{K} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{matrix} \\ \\ \\ \leftarrow \text{subset} \\ \leftarrow \text{dead} \end{matrix}$$

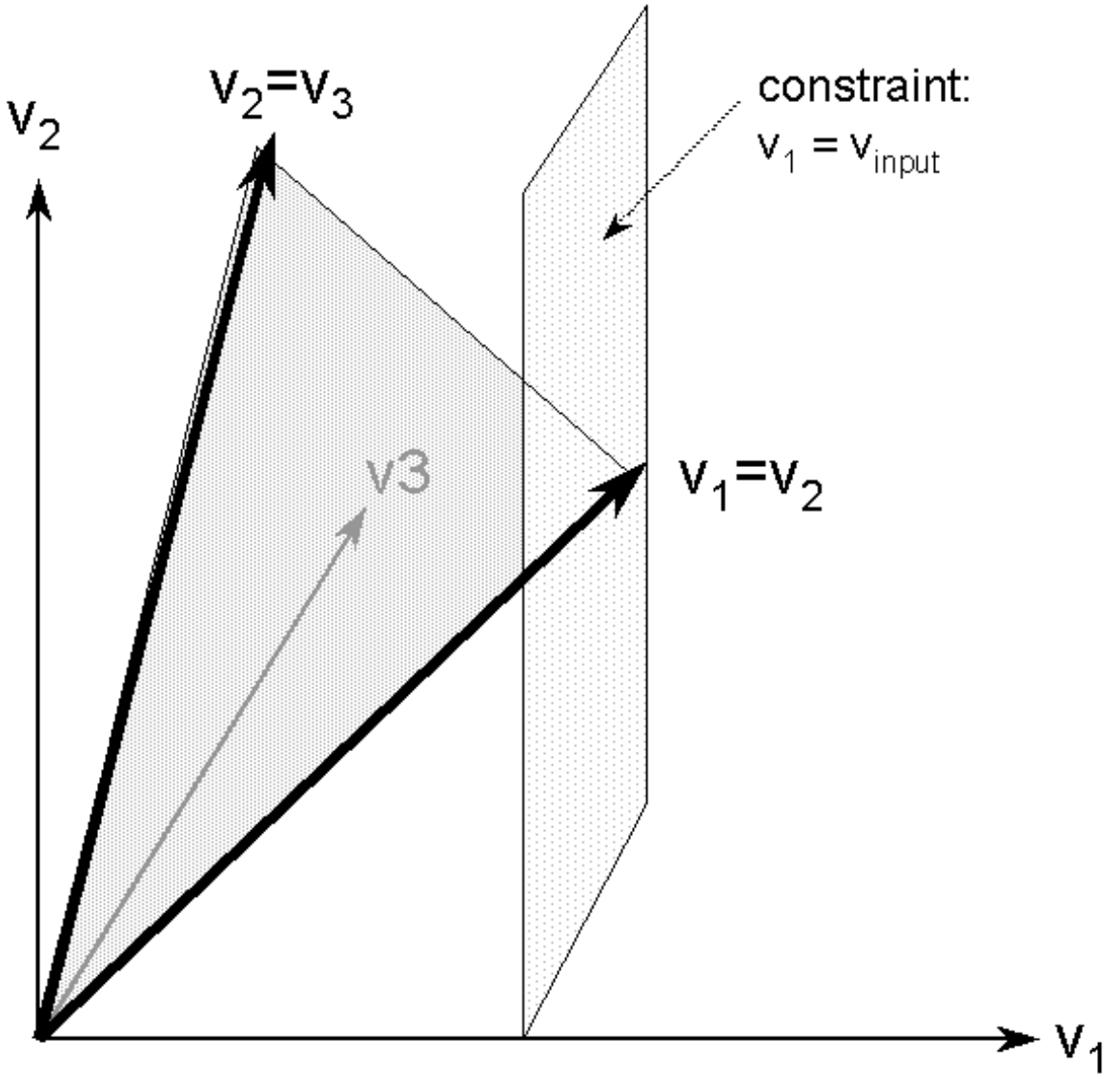
and:

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a \\ a+b \\ b \\ a \\ 0 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{bmatrix}$$

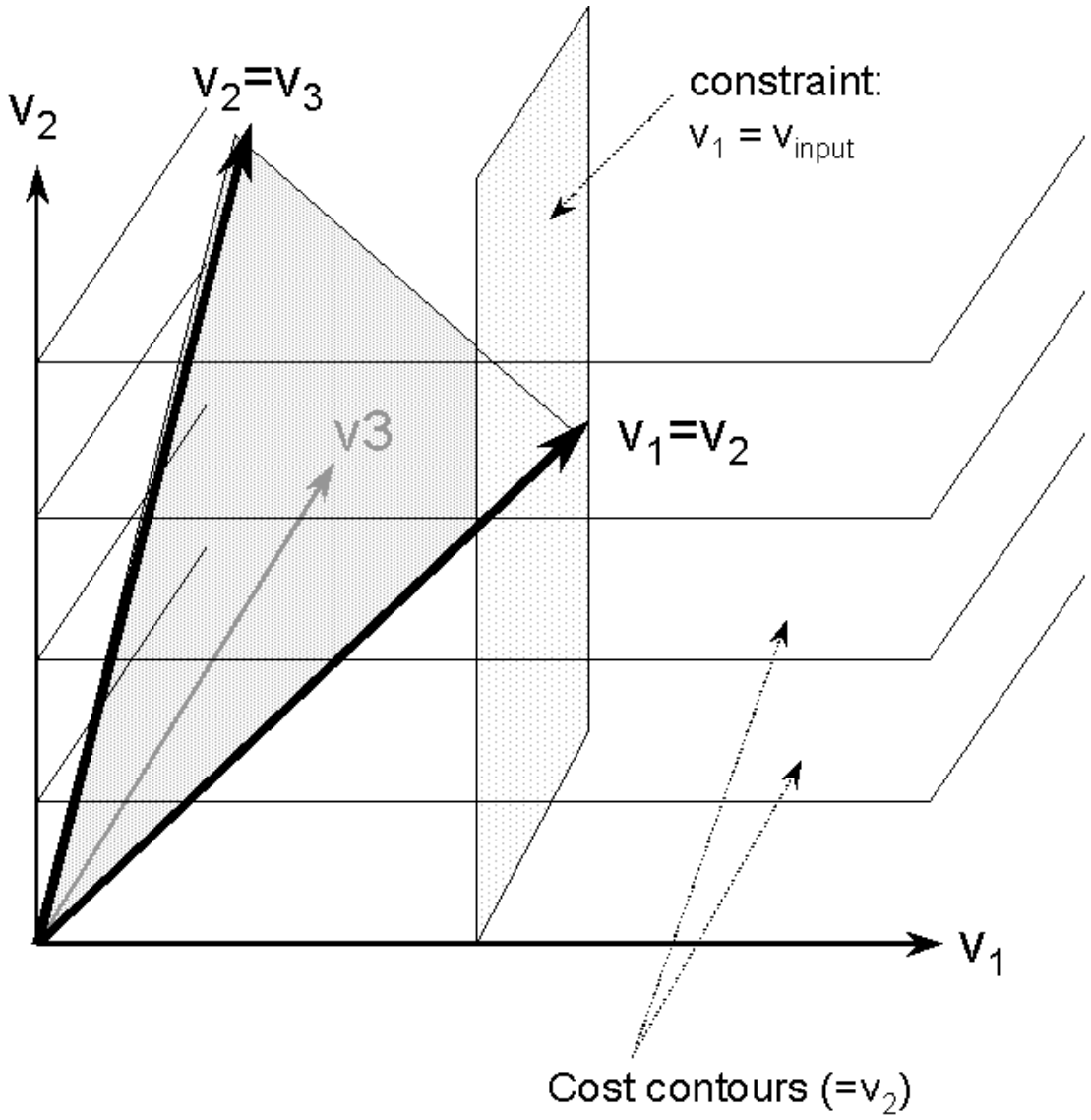
Null Space - Geometrical Interpretation



Linear Programming Solution —2



Linear Programming Solution —3



Limitations of the Null Space

Null space analysis reveals aspects of the network structure, but the set of basis vectors has shortcomings as metabolic routes:

- Is not a unique solution.
- May not respect thermodynamic direction.
- Not necessarily 'simple'.
- Can mislead about the impact of enzyme deletion.

But:

- Computation is rapid, even for genome scale networks.
- Reactions or routes shown to be 'dead' will not be found 'live' by any other approach.

Some types of metabolic model

- Qualitative/pictorial Assembled from successive enzyme reactions along a major metabolic route. Illustrates major metabolic transformations. Often called maps or charts.

Some types of metabolic model

- Qualitative/pictorial Assembled from successive enzyme reactions along a major metabolic route. Illustrates major metabolic transformations. Often called maps or charts.
- Structural — needs reaction list; traces many (all) potential routes through a network; finds optimal conversion stoichiometries, essential reactions and network flux values.

Some types of metabolic model

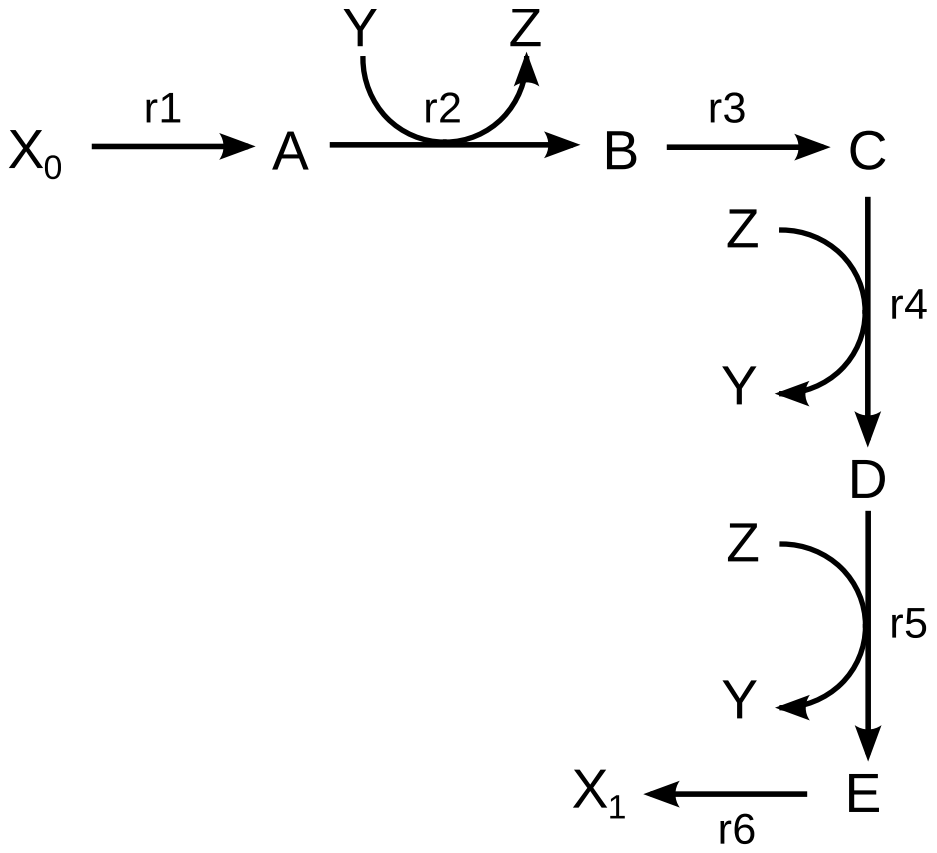
- Qualitative/pictorial Assembled from successive enzyme reactions along a major metabolic route. Illustrates major metabolic transformations. Often called maps or charts.
- Structural — needs reaction list; traces many (all) potential routes through a network; finds optimal conversion stoichiometries, essential reactions and network flux values.
- Kinetic — needs full kinetic description of each enzyme/step; predicts time–courses of rates and concentrations, steady–states, control distribution . . .

Some types of metabolic model

- Qualitative/pictorial Assembled from successive enzyme reactions along a major metabolic route. Illustrates major metabolic transformations. Often called maps or charts.
- Structural — needs reaction list; traces many (all) potential routes through a network; finds optimal conversion stoichiometries, essential reactions and network flux values.
- Kinetic — needs full kinetic description of each enzyme/step; predicts time–courses of rates and concentrations, steady–states, control distribution . . .
- Control analysis — needs effective kinetics near steady–state; predicts control distribution, response of steady state to perturbations.

Problem interlude

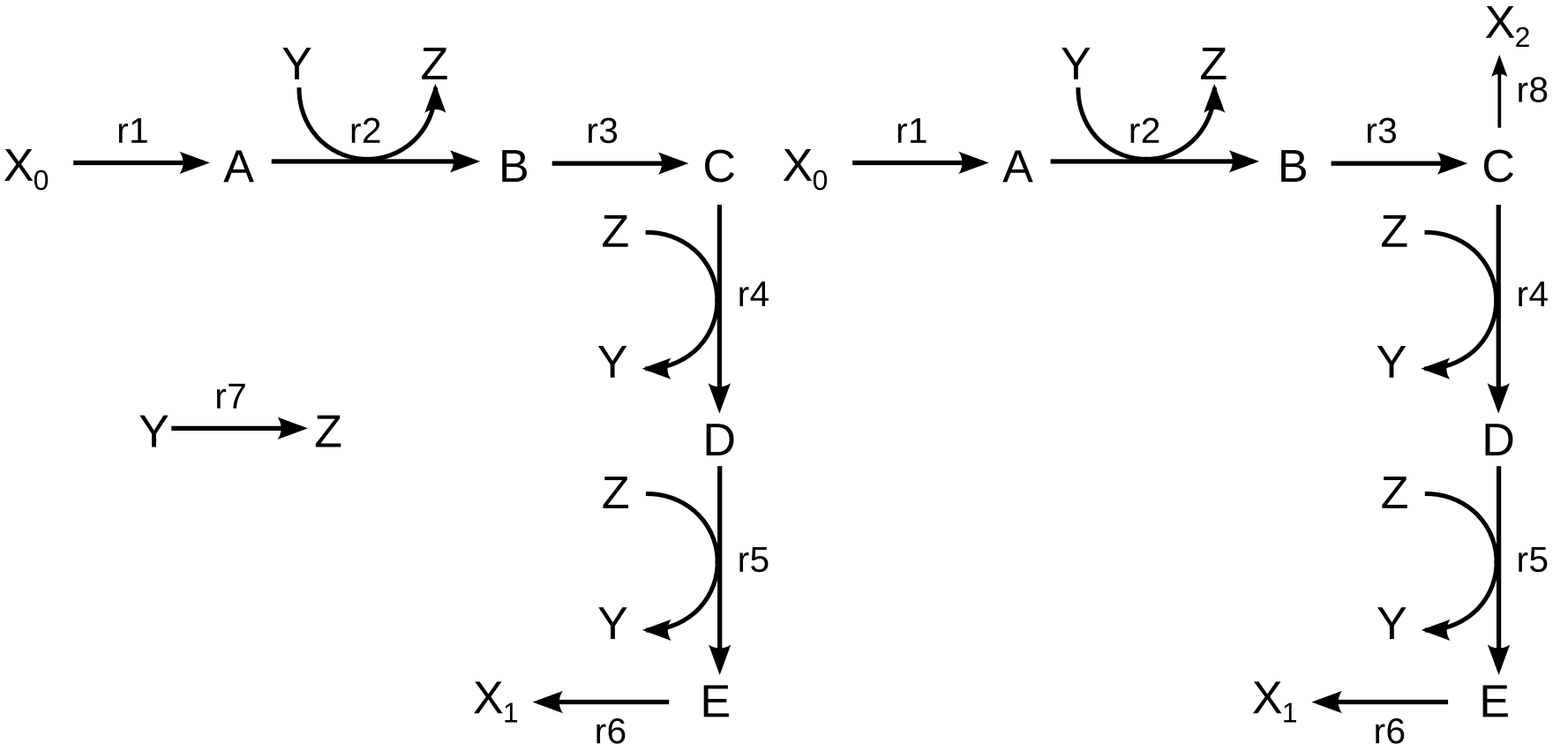
Stoichiometry Matrix of Pathway



- What is the stoichiometry matrix of his pathway?
- Will it have a steady state?

Forward to answer

Would these variants have steady states?



Structural Modelling

Advantages of Structural Analysis

- Knowledge is more complete for network structure than for enzyme kinetics.

Advantages of Structural Analysis

- Knowledge is more complete for network structure than for enzyme kinetics.
- Structural analysis involves simple linear equations; dynamic analysis involves non–linear enzyme kinetic functions.

Advantages of Structural Analysis

- Knowledge is more complete for network structure than for enzyme kinetics.
- Structural analysis involves simple linear equations; dynamic analysis involves non–linear enzyme kinetic functions.
- The network structure places limitations that constrain the network dynamics, irrespective of the kinetics, e.g.:
  - Whether viable routes exist from nutrients to stated metabolic products;
  - Whether some routes remain after deletion of the steps catalysed by a particular enzyme;
  - What the maximum obtainable conversion yield is for formation of any metabolite from a given set of sources, and

Advantages of Structural Analysis

- Knowledge is more complete for network structure than for enzyme kinetics.
- Structural analysis involves simple linear equations; dynamic analysis involves non–linear enzyme kinetic functions.
- The network structure places limitations that constrain the network dynamics, irrespective of the kinetics, e.g.:
  - Whether viable routes exist from nutrients to stated metabolic products;
  - Whether some routes remain after deletion of the steps catalysed by a particular enzyme;
  - What the maximum obtainable conversion yield is for formation of any metabolite from a given set of sources, and
- Structural models underlie kinetic models, and other techniques such as Metabolic Flux Analysis and Metabolic Control Analysis.

Structural Analysis Methods

- Linear programming - single optimal route *Small & Fell, Palsson et al*
- Null space vectors *Fell, Palsson et al*
- Computer construction of transformation routes *Serriotsis & Bailey; Mavrovouniotis et al*
- Graph analysis techniques *various* (but see: de Figueiredo et al (2008) Bioinformatics advance access: doi: 10.1093/bioinformatics/btn500.)
- Elementary modes *Schuster et al*
- Convex basis / Extreme pathways *Palsson et al*
- Enzyme (reaction) subsets

Elementary Modes

- An elementary mode is a minimal set of enzymes that can operate at steady state with all irreversible reactions working in the thermodynamically favoured direction, and enzymes weighted by the *relative* flux they carry.

Elementary Modes

- An elementary mode is a minimal set of enzymes that can operate at steady state with all irreversible reactions working in the thermodynamically favoured direction, and enzymes weighted by the *relative* flux they carry.
- 'Steady state' implies that there is only net production or consumption of external metabolites. Production and consumption of all internal metabolites is balanced.

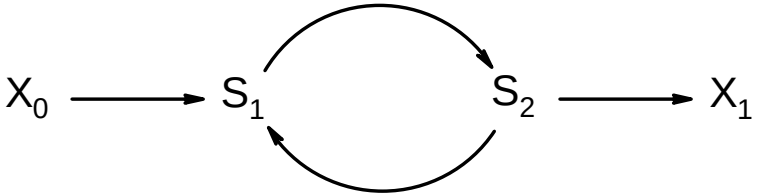
Elementary Modes

- An elementary mode is a minimal set of enzymes that can operate at steady state with all irreversible reactions working in the thermodynamically favoured direction, and enzymes weighted by the *relative* flux they carry.
- 'Steady state' implies that there is only net production or consumption of external metabolites. Production and consumption of all internal metabolites is balanced.
- 'Minimal' means that deleting any enzyme in the set would prevent a steady state. By definition, an elementary mode is not decomposable into component elementary modes.

Elementary Modes

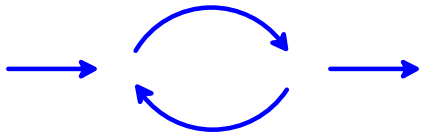
- An elementary mode is a minimal set of enzymes that can operate at steady state with all irreversible reactions working in the thermodynamically favoured direction, and enzymes weighted by the *relative* flux they carry.
- 'Steady state' implies that there is only net production or consumption of external metabolites. Production and consumption of all internal metabolites is balanced.
- 'Minimal' means that deleting any enzyme in the set would prevent a steady state. By definition, an elementary mode is not decomposable into component elementary modes.
- This ensures the set of elementary modes of a reaction network is unique, and each mode is *potentially functional*.

Elementary Modes for a Simple Network



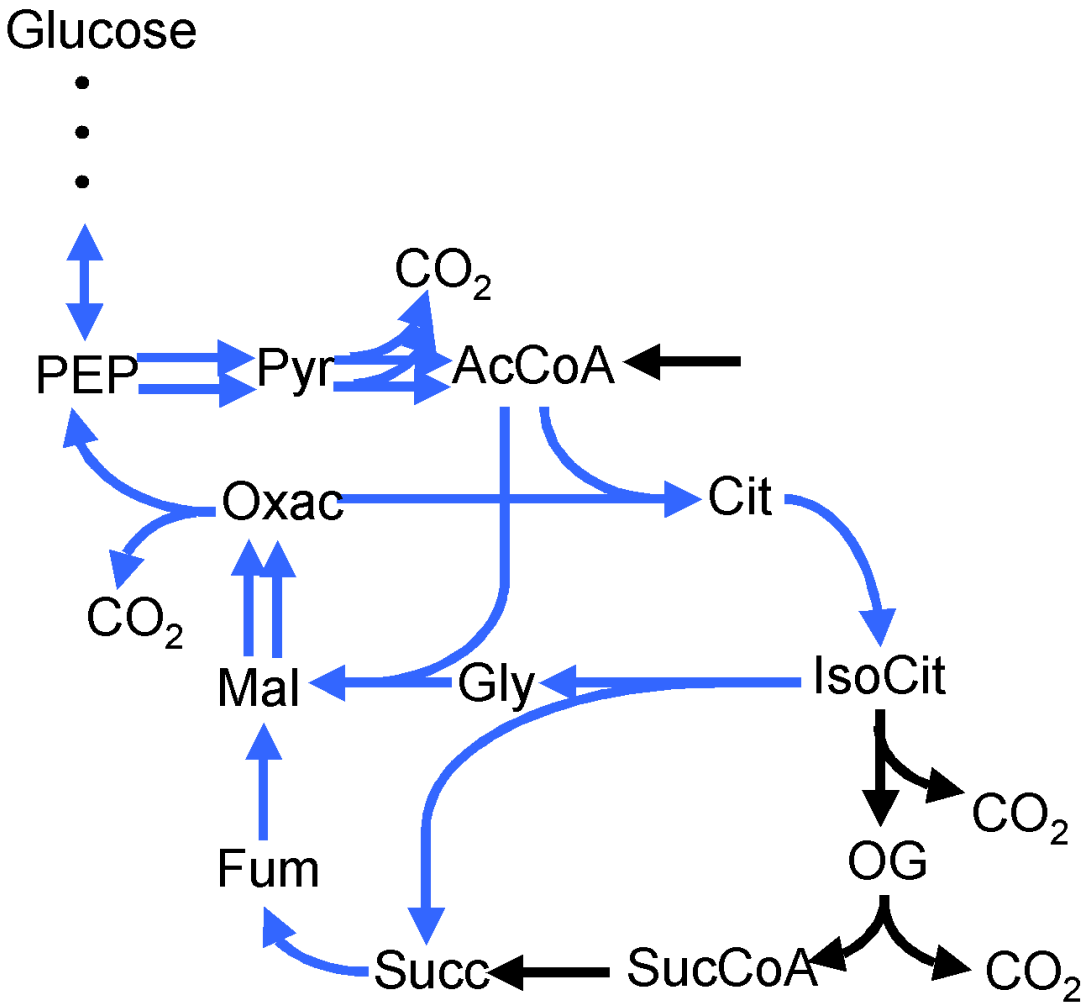
Elementary modes:

Non-elementary modes:



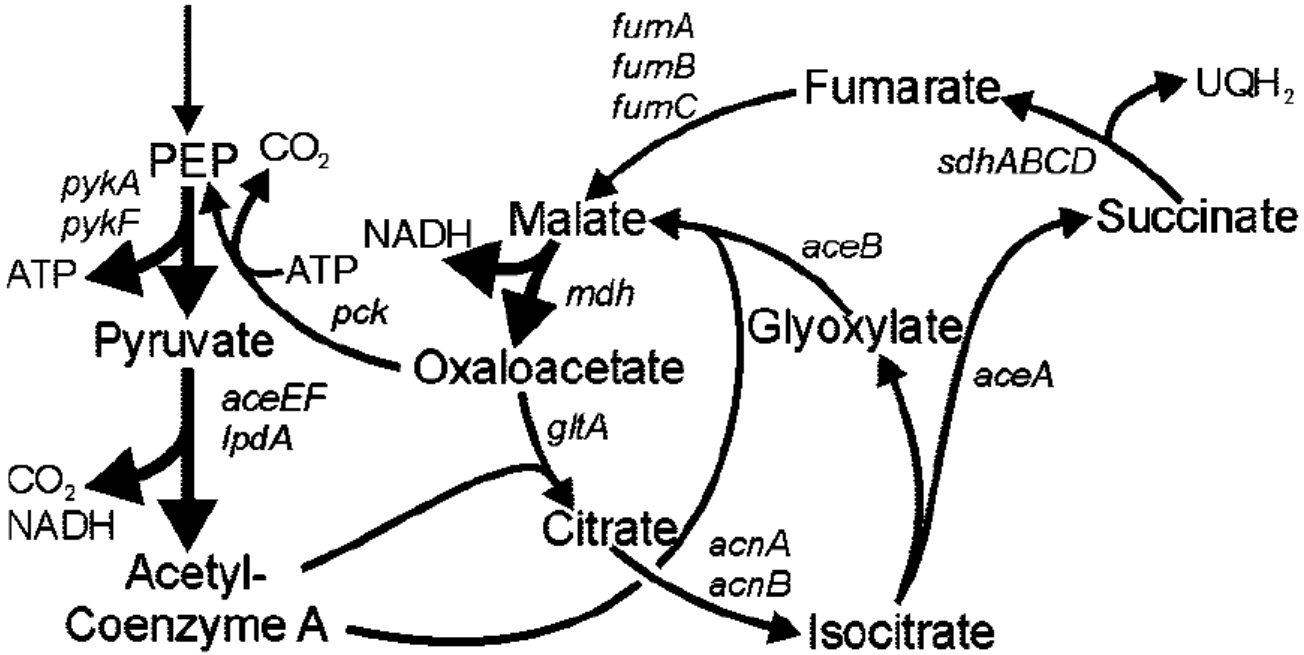
Predicting a New Pathway

In Schuster, Dandekar & Fell, *TIBS* (1999), we described the following:



The Experimental Confirmation

In *J Biol Chem* in 2003, Fischer and Sauer observed this pathway in *E coli* cells growing at low glucose levels:



$$2\ PEP \rightarrow 3\ CO_2 + 4\ NADH + UQH_2 + ATP + PEP$$

Kinetic modelling

Deterministic Dynamic Modelling

- Differential equations — one for each system variable. Ordinary differential equations (ODEs) for spatially homogeneous systems, partial DEs (PDEs) where spatial dependence and diffusion are involved. Species are represented by a continuous, concentration variable.
- Behaviour is (usually) deterministic.
- Equations are integrated numerically by repeatedly calculating new values of the variables after a very small time step.
- Computationally very efficient with appropriate modern algorithms.

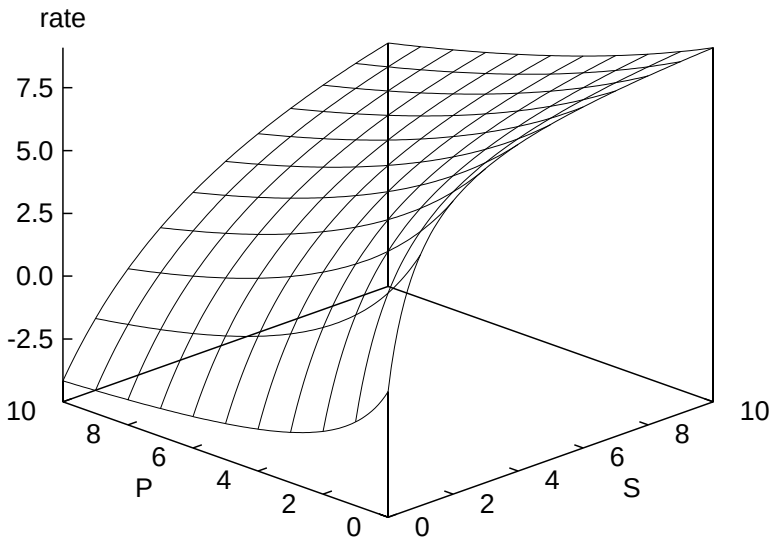
Example: Single Enzyme System

$$\mathrm{S} \leftrightharpoons \mathrm{P}$$

$$-\frac{dS}{dt} = \frac{dP}{dt} = v_{net}$$

where

$$v_{net} = \frac{(V_\mathrm{f}/K_{\mathrm{m,S}})\,(S - P/K_\mathrm{eq})}{1 + S/K_{\mathrm{m,S}} + P/K_{\mathrm{m,P}}}$$



Stochastic Dynamic Modelling

- Individual molecules/particles of the species are represented — again with or without spatial information — and the fate of each particle is followed. Track particle numbers, not concentrations.
- At each small time step, a molecule may move, react or remain unchanged with a probability related to the diffusion and rate constants.
- Computationally demanding; only feasible for modelling small volumes, and outcome is different every time.
- However, represents the intrinsic variability in systems with small numbers (<1000) of reacting particles (e.g. DNA molecules, some transcription factors etc).

Summary

Summary

- There is a range of tools available to model and analyze metabolic networks, some of them applicable to genome scale networks.
- The trick is to choose the appropriate approach for a particular problem.
- The basic principles are similar for modelling other cellular processes, such as signal transduction, cell cycle, apoptosis etc . . .

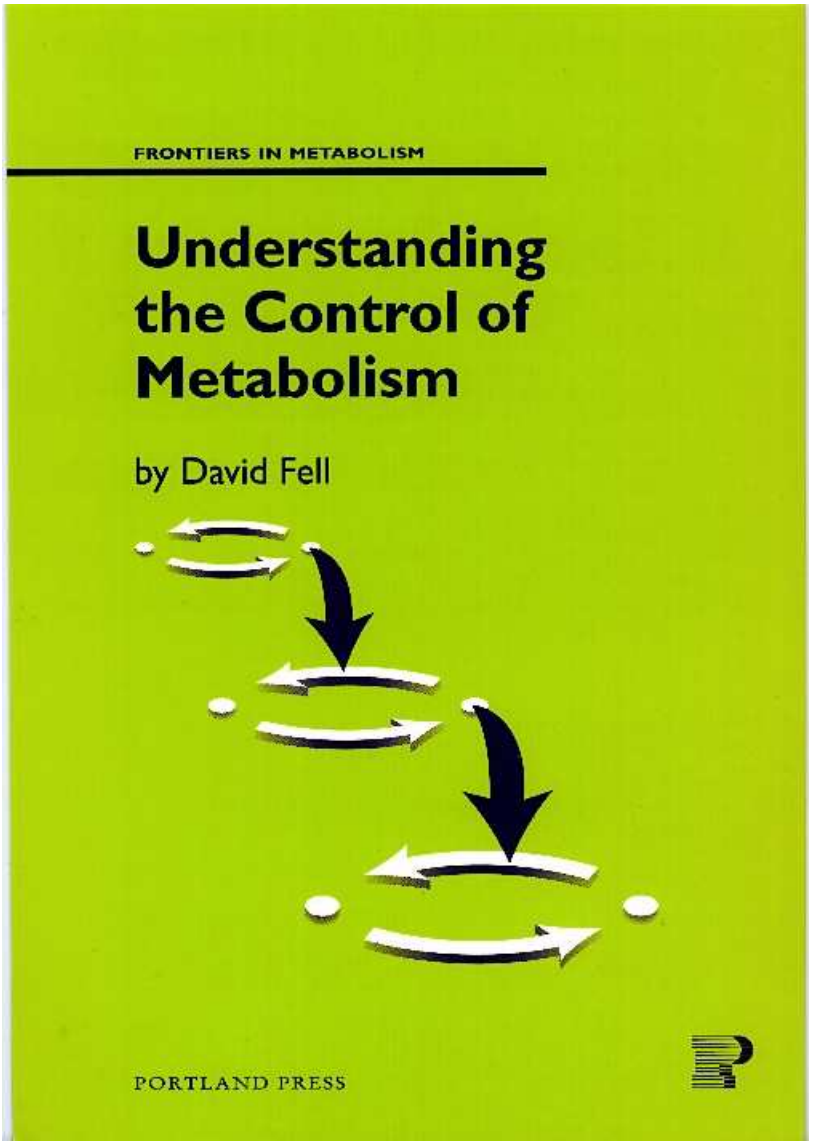
OXFORD BROOKES UNIVERSITY

Reading List

- Schuster, S. and Fell, D. A. (2007) Modelling and simulating metabolic networks. In Lengauer, T. (ed.), *Bioinformatics: From Genomes to Therapies*, vol. 2, pp. 755–806. Wiley–VCH, Weinheim
- Fell, D. A. (2008) Metabolic networks. In Képès, F. (ed.), Biological Networks, vol. 3 of *Complex Systems and Interdisciplinary Science*. World Scientific Publishing Co, Singapore
- Poolman, M. G., Bonde, B. K., Gevorgyan, A., Patel, H. H. and Fell, D. A. (2006) Challenges to be faced in the reconstruction of metabolic networks from public databases. *IEE Proc. Syst. Biol. 153*, 379–384
- de Figueiredo, L.F., Schuster, S., Kaleta, C. and Fell, D. A. (2008) Can sugars be produced from fatty acids? A test case for pathway analysis tools. *Bioinformatics* advance access: doi: 10.1093/bioinformatics/btn500
- Gevorgyan, A., Poolman, M. G., and Fell, D. A. (2008) Detection of stoichiometric inconsistencies in biomolecular models. *Bioinformatics 24*, 2245–2251
- Francke et al (2005) Reconstructing the metabolic network of a bacterium from its genome. *Trends Microbiol. 13*, 550–558
- Palsson, B Ø(2006) *Systems biology : properties of reconstructed networks*, Cambridge University Press, Cambridge. (RSL: 06.B01126)
- Price, N.D., Reed, J.L. and Palsson, B. Ø.(2004)Genome-scale Models of Microbial Cells:Evaluating the consequences of constraints, *Nature Reviews Microbiology, 2*, 886–897
- Borodina et al (2005) Genome scale analysis of *Streptomyces coelicolor* A3(2)metabolism. *Genome Res 15,* 820–829

OXFORD BROOKES UNIVERSITY

Further Details!



Order from Portland Press - printed on demand.