Beyond the Human Genome

Metabolic Pathways in the Post-Genomic Era

David Fell

dfell@brookes.ac.uk
Motivation

We understand how the genome encodes the metabolic phenotype. Can we exploit that understanding to make correct predictions about metabolic responses?

One application area: metabolic engineering

Another: design of effective drug therapies

Related cell processes should yield to a similar approach: signal transduction; cell cycle; apoptosis.
Outline

- Formal representation of metabolic pathways
- Methods of structural analysis
  - Elementary modes
  - Enzyme subsets
- Kinetic models of metabolism
- Summary
Formal representation of metabolic pathways
The metabolic network
The Michaelis–Menten equation

\[ v = \frac{SV}{S + K_m} \]

The \( K_m \) and \( V \) have arbitrarily been set to 1, where \( V \) is the *limiting rate* (or maximum velocity, \( V_m \)) and \( K_m \) is the *Michaelis constant*. 

The graph illustrates the relationship between the rate of reaction and substrate concentration. The curve reaches a limiting value when the substrate concentration is equal to the Michaelis constant, \( S = K_m \).
The reversible M–M eqn.

\[
\nu_{\text{net}} = \frac{(V_f/K_{m,S})(S - P/K_{eq})}{1 + S/K_{m,S} + P/K_{m,P}}
\]

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

\[ xase \quad \rightarrow \quad ydh \quad \rightarrow \quad zase \]

\[ X_0 \quad \rightarrow \quad Y \quad \rightarrow \quad Z \quad \rightarrow \quad X_1 \]

\(X_0\) is termed the source, and \(X_1\) is the sink. They are also termed external metabolites.

\(Y\) and \(Z\) 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 pathway 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.
Quasi steady state

If there are very slow changes in the concentrations of metabolites, or the pathway flux, because of slow changes in the 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.
Consider a simple pathway, e.g.:

\[ X_0 \xrightarrow{1} S_1 \xrightarrow{2} S_2 \xrightarrow{3} S_1 \xrightarrow{4} X_1 \]

\[
S_1 = \begin{bmatrix}
1 & -1 & 1 & 0
\end{bmatrix}
\]

\[
S_2 = \begin{bmatrix}
0 & 1 & -1 & -1
\end{bmatrix}
\]

r1: \( X_0 \rightarrow S_1 \sim \)
r2: \( S_1 \rightarrow S_2 \sim \)
r3: \( S_2 \rightarrow S_1 \sim \)
r4: \( S_2 \rightarrow X_1 \sim \)
Rates of change of $S_1$ and $S_2$

By inspection of the diagram:

$$\frac{dS_1}{dt} = v_1 - v_2 + v_3$$

$$\frac{dS_2}{dt} = v_2 - v_3 - v_4$$

At steady state $\frac{dS_1}{dt} = 0$ and $\frac{dS_2}{dt} = 0$. In addition to the above two equations, this implies $v_1 = v_4$.

How can we generalize this?
Separating structure and kinetics

The rate at which the substrate concentrations are changing is given by $N \cdot v$, where $N$ is the stoichiometry matrix, and $v$ are the enzyme kinetic functions. So for our substrate cycle pathway:

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

where each $v_i$ is the rate function for enzyme $i$, depending on the metabolites, $V_m$, $K_m$ etc.
Methods of structural analysis
Steady state solutions

Any metabolic pathway at steady state satisfies the relationship $\mathbf{N} \cdot \mathbf{v} = 0$, where $\mathbf{N}$ is the stoichiometry matrix, exemplified by the substrate cycle pathway:

\[
\begin{align*}
S_1 & : \begin{bmatrix} 1 & -1 & 1 & 0 \end{bmatrix} \\
S_2 & : \begin{bmatrix} 0 & 1 & -1 & -1 \end{bmatrix}
\end{align*}
\]

\[
\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}
\]
Steady state solutions – 2

Any observed set of velocities at steady state will be a linear combination of a set of vectors $K$ referred to as the *null space* of the stoichiometry matrix. In this case,

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

The number of null space vectors tells us the number of independent fluxes that can exist in the pathway — in this case, a linear flux and a cyclic flux.
Null space vectors as pathways

\[ X_0 \rightarrow S_1 \rightarrow S_2 \rightarrow X_1 \]

\[ [1 1 0 1]^T \]

\[ [0 1 1 0]^T \]

\[ [1 1 0 1]^T \text{ and } [0 1 1 0]^T \]
Steady state solutions – 3

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

\[ K = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \]

and:

\[
\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a \\ a+b \\ b \\ a \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}
\]
Aims of structural analysis

If we can prepare a list of the reactions coded by the genome of an organism, can we decide:

- what nutrients it can utilize and what products it can produce?
- is there a route from a particular nutrient to a product?
- which route to a product has the highest yield?
- what are the consequences of deleting an enzyme?
- do the genome annotations generate a connected and self-consistent metabolism?
The reaction list

Where does the data for this metabolic model come from?

- Biochemical literature: books, reviews, journal articles.

- Genome databases plus annotation plus enzyme database.
Database issues

A structural model requires a reaction list in which:

- Each reactant appears with a single unambiguous name.

- Each relevant substrate/product conversion catalysed by a single enzyme is individually identified.

- Incomplete EC number designations in genome annotations have been resolved.
Unambiguous metabolite names

In glycolysis, the route from pyruvate to ethanol is recorded as (in part):

1.2.4.1 Pyruvate dehydrogenase (lipoamide)
Pyruvate + lipoamide <=> S-acetyldihydrolipoamide + CO(2) etc, or

4.1.1.1 Pyruvate decarboxylase
A 2-oxo acid = an aldehyde + CO(2)

1.1.1.1 Alcohol dehydrogenase
An alcohol + NAD(+) <=> an aldehyde or ketone + NADH
Each conversion identified

Transketolase (2.2.1.1) appears as:

Sedoheptulose 7-phosphate + D-glyceraldehyde 3-phosphate = D-ribose 5-phosphate + D-xylulose 5-phosphate.
Wide specificity for both reactants, e.g. converts hydroxypyruvate and R-CHO into CO(2) and R-CHOH-CO-CH(2)OH.

which doesn’t help to find the Calvin cycle reaction:

fructose 6-P + glyceraldehyde 3-P = erythrose 4-P + xylulose 5-P.
Incomplete EC numbers

From E. coli, 4 years ago and today:

1. hisFH: 3.5.1.- imidazole glycerol phosphate synthase
2. pyrH: 2.7.4.- uridylate kinase
3. pabABC: 4.1.3.- (+ 2 others) p-aminobenzoate synthase multi-enzyme complex
4. cls: 2.7.8.- cardiolipin synthase
5. murG: 2.4.1.- N-acetylglucosaminyl transferase
6. dxr: 1.1.1.- now 1.1.1.267
7. ispB: 2.5.1.- octaprenyl diphosphate synthase
Elementary modes analysis
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.

Hence the set of elementary modes of a reaction network is unique.
## Pentose phosphate reactions

<table>
<thead>
<tr>
<th>Gene</th>
<th>Reaction</th>
<th>E. coli</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pgi</td>
<td>G6P = F6P</td>
<td>EC4025</td>
</tr>
<tr>
<td>Pfk</td>
<td>F6P + ATP – ADP + FP2</td>
<td>EC3916</td>
</tr>
<tr>
<td>Fbp</td>
<td>FP2 - F6P + Pi</td>
<td>EC4232</td>
</tr>
<tr>
<td>Fba</td>
<td>FP2 = GAP + DHAP</td>
<td>EC2925</td>
</tr>
<tr>
<td>TpiA</td>
<td>GAP = DHAP</td>
<td>EC3919</td>
</tr>
<tr>
<td>Gap</td>
<td>GAP + Pi + NAD = NADH + BPG</td>
<td>EC1779</td>
</tr>
<tr>
<td>Pgk</td>
<td>BPG + ADP = ATP + P3G</td>
<td>EC2926</td>
</tr>
<tr>
<td>Gpm</td>
<td>P3G = P2G</td>
<td>EC0755</td>
</tr>
<tr>
<td>Eno</td>
<td>P2G = PEP</td>
<td>EC2779</td>
</tr>
<tr>
<td>Pyk</td>
<td>PEP + ADP – ATP + Pyr</td>
<td>EC1854</td>
</tr>
<tr>
<td>Zwf</td>
<td>G6P + NADP = GO6P + NADPH</td>
<td>EC1852</td>
</tr>
<tr>
<td>Pgl</td>
<td>GO6P – P6G</td>
<td>EC0766</td>
</tr>
<tr>
<td>Gnd</td>
<td>P6G + NADP = NADPH + CO2 + Ru5P</td>
<td>EC2029</td>
</tr>
<tr>
<td>Rpi</td>
<td>Ru5P = R5P</td>
<td>EC2914</td>
</tr>
<tr>
<td>Rpe</td>
<td>Ru5P = Xyl5P</td>
<td>EC3386</td>
</tr>
<tr>
<td>TkI</td>
<td>R5P + Xyl5P = GAP + Sed7P</td>
<td>EC2935</td>
</tr>
<tr>
<td>Tal</td>
<td>GAP + Sed7P = Ery4P + F6P</td>
<td>EC2465</td>
</tr>
<tr>
<td>TkI1</td>
<td>Xyl5P + Ery4P = F6P + GAP</td>
<td>EC2935</td>
</tr>
<tr>
<td>Prs</td>
<td>R5P – R5Pex</td>
<td>EC4383</td>
</tr>
</tbody>
</table>
Pentose phosphate pathway: input
Mode 1: glycolysis

- G6P $\rightarrow$ F6P $\rightarrow$ FP2 $\rightarrow$ GAP $\rightarrow$ DHAP $\rightarrow$ 1.3BPG
- ATP $\rightarrow$ ADP $\rightarrow$ Pyr
- PEP $\rightarrow$ 2PG
- 3PG

1. ATP
2. ADP
3. Pyr
Mode 2: G6P to pyruvate and CO$_2$. 

b) 

[Diagram showing the metabolic pathways and reactions involving G6P, F6P, Ru5P, Xyl5P, R5P, Sed7P, Ery4P, GAP, F6P, G6P, Ru5P, CO$_2$, NADPH, NADP, GO6P, ADP, ATP, 1.3BPG, 1.3BPG, ADP, ATP, Pyr, PEP, GAP, 2PG, 3PG, NADPH, NADP, and NADH. The diagram includes numbered reactions and arrows indicating the flow of metabolites and reactions.]
Mode 3: G6P to pyruvate and CO$_2$. 

The diagram illustrates the metabolic pathway from G6P to pyruvate and CO$_2$ through various intermediate metabolites such as 6PG, Ru5P, Xyl5P, R5P, Sed7P, Ery4P, F6P, ATP, ADP, NADP, NADPH, and CO$_2$. The reactions are labeled with numbers indicating the enzymatic steps or chemical reactions involved.

- CO$_2$ is produced from 6PG and Ru5P.
- NADPH and NADP are interconverted through reactions involving 6PG and Ru5P.
- ATP and ADP are involved in energy transfer reactions.
- NAD and NADH are also crucial for redox reactions in the pathway.

The diagram is complex and requires a detailed understanding of metabolic pathways to fully appreciate its significance.
Mode 4: G6P to ribose 5–P and CO$_2$

d)

\[
\begin{align*}
\text{CO}_2 & \rightarrow \text{Ru5P} \\
\text{NADPH} & \rightarrow \text{6PG} \\
\text{NADP} & \rightarrow \text{GO6P} \\
\text{NADPH} & \rightarrow \text{G6P} \\
1 & \rightarrow \text{R5P} \\
1 & \rightarrow \text{R5P}_{\text{ex}} \\
1 & \rightarrow \text{Ru5P} \\
1 & \rightarrow \text{6PG} \\
\text{NADP} & \rightarrow \text{G6P}
\end{align*}
\]
Mode 5: G6P to ribose 5–P, no CO$_2$
Mode 6: cyclic PPP mode
Gene ontologies and annotation

How relevant are they? e.g. the Gene Ontology Consortium:

pentose-phosphate shunt, non-oxidative branch
  * Accession: GO:0009052
  * Aspect: process
  * Synonyms: None
  * Definition:
    o The branch of the pentose-phosphate shunt which does not involve oxidation reactions. It comprises a series of sugar phosphate interconversions, starting with ribulose 5-P and producing fructose 6-P and glyceraldehyde 3-P.
Predicting new pathways

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

**Table 1. Elementary modes of the combined TCA cycle and glyoxylate shunt system . . .**

\[6 \text{ADP} + \text{FAD} + 4 \text{NAD} + \text{PG} \rightarrow \text{ATP} + \text{FADH}2 + 4 \text{NADH} + 3 \text{CO2} (\text{Eno} \ 2\text{Pyk} \ 2\text{AceEF} \ \text{GltA} \ \text{Acn} \ \text{Sdh} \ \text{Fum} \ 2\text{Mdh} \ \text{Icl} \ \text{Mas} \ \text{Pck})\]
The experimental confirmation

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

\[
2 \text{PEP} \rightarrow 3 \text{CO}_2 + 4 \text{NADH} + \text{UQH}_2 + \text{ATP} + \text{PEP}
\]
The experimental confirmation -2

Flux measurements showed it combined with a background level of the TCA cycle:
Gene ontologies and annotation - 2

How relevant are they? e.g. the Gene Ontology Consortium:

glyoxylate cycle
  * Accession: GO:0006097
  * Aspect: process
  * Synonyms: glyoxylate bypass
  * Definition:
    - A modification of the TCA cycle occurring in some plants and microorganisms, in which isocitrate is cleaved to glyoxylate and succinate. Glyoxylate can then react with acetyl-CoA to form malate.
Enzyme knockouts: G6PDH

Glucose-6-phosphate dehydrogenase deficiency (favism) is a common enzymopathy in humans.

Targeted G6PDH knockout in mouse cells produces clones that can grow (i.e. make ribose for nucleic acid synthesis and some NADPH), but that are sensitive to oxidative stress (increased NADPH demand).

Note the mode generating ribose without using G6PDH.
Polyhydroxybutyrate synthesis in yeast

Glycolysis

Pentose-P Pathway

Ethanol

Acetate

PHB Pathway

Succinate

Glycerol

Extracellular Fluid

Mitochondrion

TCA Cycle

Ox. Phos

Oxford BHG–04 – p. 41
Optimal yields of PHB synthesis

Wild-type yeast + PHB pathway

1. 2 Acetate + EtOH $\rightarrow$ PHB + 2 CO$_2$ 0.67
2. 65 Ac. + 31 EtOH $\rightarrow$ 30 PHB + 72 CO$_2$ 0.63

Wild-type yeast + ATP–citrate lyase + PHB pathway

3. 12 EtOH $\rightarrow$ 5 PHB + 4 CO$_2$ 0.83
4. 77 EtOH + 31 Glycerol $\rightarrow$
   48 PHB + 4 Ac. + 47 CO$_2$ 0.78

(Number following each mode is the fractional carbon conversion.)
Enzyme subset analysis
Enzyme subsets

An enzyme subset is defined as a group of enzymes that carry flux in a fixed proportion in any steady state where they are active.

For a simple linear system, all enzymes are in a single enzyme subset.
Subsets in a branched system

Advantages of enzyme subsets:

- Model simplification: each subset can be replaced by a single overall reaction.
- Can aid genome annotation by suggesting missing enzymes from ‘broken’ subsets.
- May give clues about co–regulation of genes.
Gene annotation: *Treponema pallidum*.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Reaction</th>
<th><em>T. pal.</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pgi</td>
<td>G6P = F6P</td>
<td>TP0475</td>
</tr>
<tr>
<td>Pfk</td>
<td>F6P + ATP – ADP + FP2</td>
<td>TP0108</td>
</tr>
<tr>
<td>Fbp</td>
<td>FP2 – F6P + Pi</td>
<td>—</td>
</tr>
<tr>
<td>Fba</td>
<td>FP2 = GAP + DHAP</td>
<td>TP0662</td>
</tr>
<tr>
<td>TpiA</td>
<td>GAP = DHAP</td>
<td>TP0537</td>
</tr>
<tr>
<td>Gap</td>
<td>GAP + Pi + NAD = NADH + BPG</td>
<td>TP0844</td>
</tr>
<tr>
<td>Pgk</td>
<td>BPG + ADP = ATP + P3G</td>
<td>TP0538</td>
</tr>
<tr>
<td>Gpm</td>
<td>P3G = P2G</td>
<td>TP0168</td>
</tr>
<tr>
<td>Eno</td>
<td>P2G = PEP</td>
<td>TP0817</td>
</tr>
<tr>
<td>Pyk</td>
<td>PEP + ADP – ATP + Pyr</td>
<td>—</td>
</tr>
<tr>
<td>Zwf</td>
<td>G6P + NADP = GO6P + NADPH</td>
<td>TP0478</td>
</tr>
<tr>
<td>Pgl</td>
<td>GO6P – P6G</td>
<td>—</td>
</tr>
<tr>
<td>Gnd</td>
<td>P6G + NADP = NADPH + CO2 + Ru5P</td>
<td>TP03319</td>
</tr>
<tr>
<td>Rpi</td>
<td>Ru5P = R5P</td>
<td>TP0616</td>
</tr>
<tr>
<td>Rpe</td>
<td>Ru5P = Xyl5P</td>
<td>TP0945</td>
</tr>
<tr>
<td>TktI</td>
<td>R5P + Xyl5P = GAP + Sed7P</td>
<td>TP0650</td>
</tr>
<tr>
<td>Tal</td>
<td>GAP + Sed7P = Ery4P + F6P</td>
<td>—</td>
</tr>
<tr>
<td>TktII</td>
<td>Xyl5P + Ery4P = F6P + GAP</td>
<td>TP0560</td>
</tr>
<tr>
<td>Prs</td>
<td>R5P - R5Pex</td>
<td>—</td>
</tr>
<tr>
<td>Ppd</td>
<td>Pyr + ATP + Pi = PEP + AMP + Ppi</td>
<td>TP0746</td>
</tr>
</tbody>
</table>
Elementary mode of the *T. pallidum* set

Pgi, Pfk, Fba, TpiA, 2Gap, 2Pgk, 2Gpm, 2Eno, -2Ppd.
(G6P + 2AMP + ADP + 2PPi + 2NAD → 3ATP + 2NADH + 2 Pyr)
‘Unused’ enzymes include Zwf, Gnd, Rpi, Rpe, Tkt. But, in the full pentose phosphate pathway model:
- Zwf Gnd Pgl, and
- Rpe Tkt Tal

are ‘enzyme subsets’, which, if they appear in an elementary mode, always do so in constant proportions.
Analysis of an *E. coli* model

- No. of reactions: 678
- No. of metabolites: 761
- No. of external metabolites: 63
Distribution of enzyme subsets

Size of Enzyme Subset Vs Frequency

Frequency

Size of Enzyme Subset
Enzyme–operon correlation

glgX \rightarrow glgB \rightarrow glgC \rightarrow glgA \rightarrow glgP

glgX: Operon protein

Glu6P \rightleftharpoons \text{Glu1P} \rightleftharpoons \text{ADPG} \rightleftharpoons \text{SGly} \rightleftharpoons \text{Glycogen} \rightleftharpoons X_{Glycogen}

R1, ATP, R2, ADP, R3, Ppi, R4, R5, Pi
Summary - structural models

Structural analysis can offer:

- Valid routes through metabolic networks;
- Analysis of metabolic capabilities;
- A tool for aiding metabolic engineering;
- A functional interpretation of flux patterns;
- Identification of metabolic subsystems;
- Prediction of possible co–regulation/co–expression patterns.
Acknowledgements

Berlin/Jena:  Stefan Schuster, Claus Hilgetag, Thomas Pfieffer & Ferdinand Moldenhauer

Oxford:  Mark Poolman, John Woods, Bhushan Bonde

Heidelberg/Freiburg:  Thomas Dandekar

Minnesota:  Friedrich Srienc, Ross Carlson
Further details

References:


See also the Palsson group work: http://gcrg.ucsd.edu/
Kinetic models
Kinetic models

Needed: a single equation for each enzyme containing all effects from

- substrates
- product inhibition
- reverse reaction
- effectors

using parameters determined at

- in vivo pH
- temperature
- ion concentrations

in the organism, cell type and compartment under consideration.
A kinetic model of *E. coli* threonine biosynthesis
The threonine pathway

Aspartate kinase III (AK)
Aspartate kinase I (AK)
Aspartyl-P
Lysine
Aspartate-semialdehyde dehydrogenase (ASD)
Homoserine dehydrogenase (HDH)
Homoserine kinase (HK)
D-phosphohomoserine
Protein Synthesis
Threonine Synthase (TS)
Threonine
Isoleucine
Methionine
Components of the modelling project

1. Kinetics of the pathway enzymes
2. Dynamics of threonine synthesis in cell-free extracts
3. Building a computer model of the pathway based on kinetics
4. Validation of model with cell-free experiments
5. Extrapolating model to intracellular conditions
1. Kinetic data

Why not use published data?

- pH values.

- Reaction direction.

- Lack of information on product inhibition.

- Analysis didn’t produce a single overall equation in terms of all substrates, products and effectors.
Example: aspartate kinase I

\[
\begin{align*}
Num &= V_f \left( \text{asp}.ATP - \frac{\text{aspp}.ADP}{K_{eq}} \right) \\
D1 &= \left( K_{\text{asp}} \frac{1 + \left( \frac{\text{thr}}{K_{\text{ithr}}} \right)^{n_h}}{1 + \left( \frac{\text{thr}}{\alpha K_{\text{ithr}}} \right)^{n_h}} + \text{aspp} \frac{K_{\text{asp}}}{K_{\text{aspp}}} + \text{asp} \right) \\
D2 &= \left( K_{\text{ATP}} \left( 1 + \frac{ADP}{K_{ADP}} \right) + \text{ATP} \right) \\
v &= \frac{Num}{D1 \times D2}
\end{align*}
\]
Product inhibition of AK I
2. Generating pathway data

Threonine synthesis by cell-free extract:
initial aspartate = 0.5 mM

![Graph showing threonine synthesis over time with various markers and lines representing different compounds and conditions.](image)
3. Simulator (SCAMP) input — 1

```
title threk9_4.cmd: kinetics;

dec asp, aspp, asa, hs, hsp, nadph, nadp, thr, atp, adp, Pi, endonadph;

simulate;

options auto_conserve, integrator=LSODA;

reactions

[ak] asp + atp = aspp + adp;
```
F1*(vm11*(asp*atp - aspp*adp/keqak)/((k11*(1+(thr/k1thr)ˆnak1)/(1+(thr/(alpha*k1thr))ˆnak1)+(k11*aspp/k1aspp) + asp)*(k1atp*(1+adp/k1adp)+atp)) + vm13*(asp*atp - aspp*adp/keqak)/((1 + (lys/k1lys)ˆnak3)*(k13*(1 + aspp/k13aspp) + asp)*(k13atp*(1+adp/k13adp)+atp))));

# F1 is a factor to allow modulation of enzyme group;

[asd] aspp + nadph = asa + nadp + Pi; (vm2f*(aspp*nadph - asa*nadp*Pi/k2eq))/ ((k2aspp*(1 + asa/k2asa)*(1 + Pi/k2p) + aspp)* (k2nadph*(1 + nadp/k2nadp) + nadph));
The differential equations

Automatically derived by SCAMP:

Differential equations (reordered):

1: \( aspp' = V[ak] - V[asd] \)

2: \( asa' = V[asd] - V[hdh] \)

3: \( hs' = V[hdh] - V[hk] \)

4: \( hsp' = -V[ts] + V[hk] \)
4. Simulating dynamics

Threonine synthesis by cell-free extract:
initial aspartate = 0.5mM
5. Extrapolating to *in vivo* behaviour

\[
\text{Asp} \xrightarrow{\text{AK}} \text{aspP} \xrightarrow{\text{ASA DH}} \text{asa} \xrightarrow{\text{HDH}} \text{hs} \xrightarrow{\text{HK}} \text{hsP} \xrightarrow{\text{TS}} \text{Thr}
\]

**Pathway**

- **Aspartate (Asp)**
- **Aspartate semialdehyde (aspP)**
- **Aspartate aminotransferase (ASA)**
- **Dehydrogenase (DH)**
- **Homoserine (hs)**
- **Homoserine kinase (HK)**
- **Homoserine phosphate (hsP)**
- **Threonine (Thr)**

**Metabolites**

- **ATP**
- **NADPH**
- **NADPH**
- **ATP**
“External” concentrations

Measured on cells:

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Content, nmol.(g dry wt)$^{-1}$</th>
<th>Concentration, mM</th>
</tr>
</thead>
<tbody>
<tr>
<td>asp</td>
<td>2854</td>
<td>1.34</td>
</tr>
<tr>
<td>thr</td>
<td>7444</td>
<td>3.49</td>
</tr>
<tr>
<td>lys</td>
<td>984</td>
<td>0.46</td>
</tr>
<tr>
<td>ATP</td>
<td>2792</td>
<td>1.31</td>
</tr>
<tr>
<td>ADP</td>
<td>352</td>
<td>0.17</td>
</tr>
<tr>
<td>NADP</td>
<td>1341</td>
<td>0.63</td>
</tr>
<tr>
<td>NADPH</td>
<td>1197</td>
<td>0.56</td>
</tr>
<tr>
<td>Pi</td>
<td>ND</td>
<td>5</td>
</tr>
</tbody>
</table>
Simulating enzyme over-expression
Definition of the flux control coefficient

Suppose a small change, $\delta E_{xase}$, is made in the amount of enzyme $E_{xase}$, and that this produces a small change in the flux through the step catalysed by $ydh$.

The flux control coefficient $C_{xase}^{J_{ydh}}$ is approximately the % change in $J_{ydh}$ produced by a 1% change in $E_{xase}$. 
Definition of the flux control coefficient

\[ C_{xase} = \frac{\partial J_{ydh}}{\partial E_{xase}} \cdot \frac{e}{J} \]
Control of the pathway

At measured enzyme contents in cells:

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Flux control coeff</th>
<th>Flux control coeff</th>
</tr>
</thead>
<tbody>
<tr>
<td>aspartate kinase</td>
<td>0.280</td>
<td>0.249</td>
</tr>
<tr>
<td>aspartate semialdehyde</td>
<td>0.250</td>
<td>0.570</td>
</tr>
<tr>
<td>dehydrogenase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>homoserine dehydrogenase</td>
<td>0.465</td>
<td>0.180</td>
</tr>
<tr>
<td>homoserine kinase</td>
<td>0.005</td>
<td>0.002</td>
</tr>
<tr>
<td>threonine synthase</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
Threonine synthesis in the cell

Asp → aspP → asa → hs → hsP → Thr<sub>in</sub> → Thr<sub>out</sub>

AK  ASA  DH  HK

ATP  NADPH  NADPH  ATP

Thr<sub>in</sub> → Leak → Thr<sub>out</sub>

Thr<sub>in</sub> → TTS → out1

Thr<sub>in</sub> → TD → Ile

Ile → out2

Protein
Control at different growth rates.

The graph shows the relative growth rate on the x-axis and the flux control coefficient on the y-axis. Different processes, such as Protein synthesis, HDH, AK, ASADH, and Leak, are plotted against the relative growth rate. The graph illustrates how these processes change as the growth rate increases.
Summary
Conclusions

- Modelling and making predictions about the metabolic phenotype is feasible.

- A major issue is the quantity, quality and accessibility of information about gene products.

- Issues remain about scaling modelling approaches to the whole genome or cell.
Perspectives

Reconstruction of the metabolic phenotype of microbial cells by structural modelling is advancing, e.g. Palsson group.

Initiatives are under way to model whole cells:
- The *E coli* alliance: http://www.ieca2004.ca/
- The yeast cell: http://www.siliconcell.net/sica/
- The hepatocyte: http://www.bmbf.de/en/1140.php though these have a long way to go.

Rational approaches to metabolic engineering are beginning to work in an academic setting, but need to show value industrially.

Rational design of drug therapies is being pioneered - the Oxford heart model is an early success - but are not yet embedded in the pharmaceutical industry.
Acknowledgements - Kinetic models

**Oxford:** Mark Poolman

**Bordeaux:** Jean-Pierre Mazat, Christophe Chassagnole
Further details

References:


Programs:

ScrumPy: http://mudshark.brookes.ac.uk/ScrumPy/

Scamp:
http://www.cds.caltech.edu/~hsauro/Scamp/scamp.htm