An introduction to modeling of biological systems

Overview

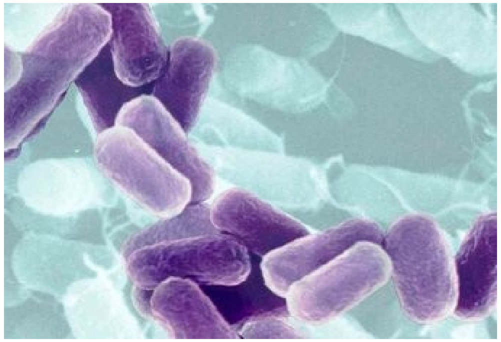
Part I - Kinetic modeling
- What is modelling about?
- Kinetic models of biochemical pathways
- Simulation and dynamic behaviour
- Model fitting

Part II - Constraint-based modeling
- Network reconstruction
- Flux Balance Analysis (FBA)

Part III - Other dynamical cell models
- Whole-cell models
- Gene expression models
- Stochastic simulation
- Spatial simulation models
- Model formats and tools

Part IV - Data analysis and regression
- Principal Component Analysis
- Clustering
- Linear regression

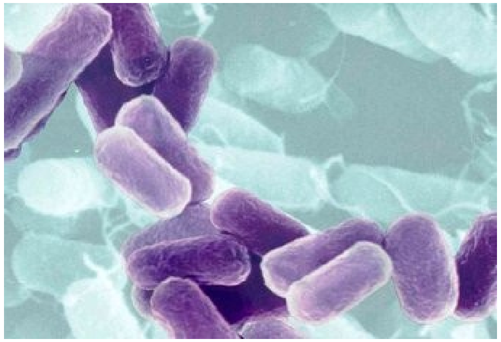
Blackboard session (Wednesday / Thursday)
- Advanced kinetic modeling and enzyme costs



How can a living cell emerge from sugar, water, and a couple of salts?

Minimal Medium for *E. coli*

Glucose	5 g/l
Na_2HPO_4	6 g/l
KH_2PO_4	3 g/l
NH_4Cl	1 g/l
NaCl	0.5 g/l
$MgSO_4$	0.12 g/l
$CaCl_2$	0.01 g/l

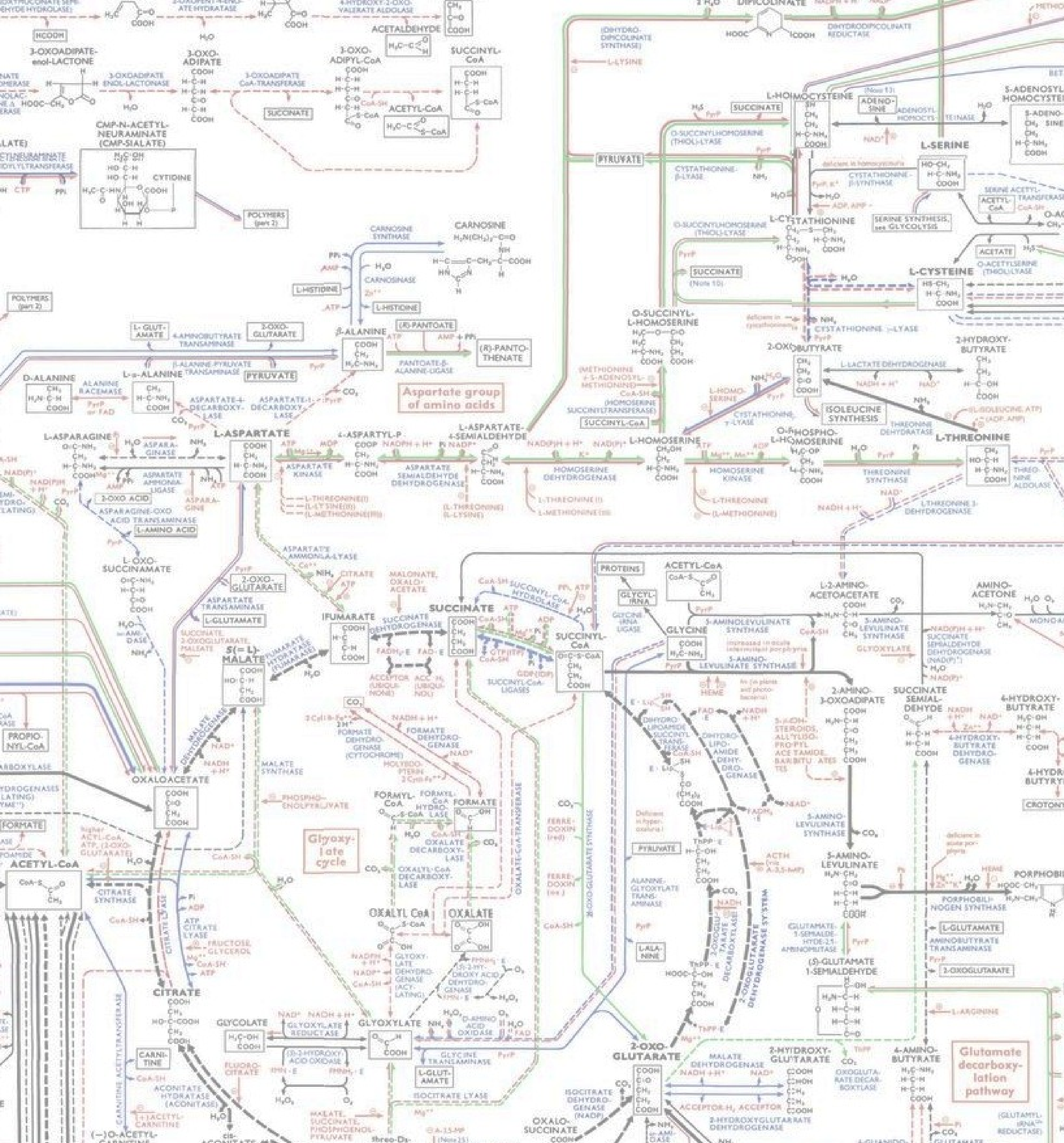


How can a living cell emerge from sugar, water, and a couple of salts?

Minimal Medium for *E. coli*

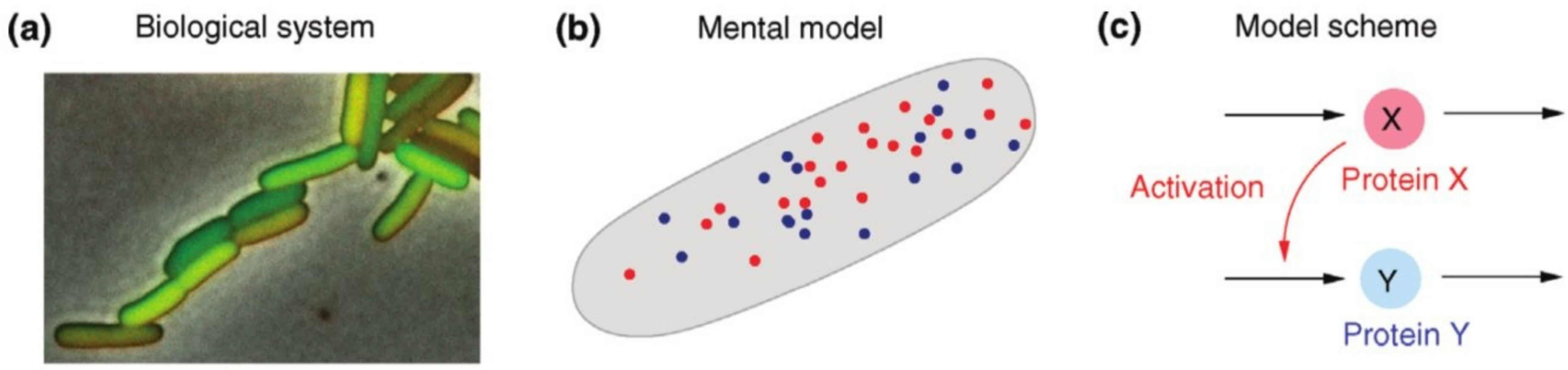
Glucose	5 g/l
Na_2HPO_4	6 g/l
KH_2PO_4	3 g/l
NH_4Cl	1 g/l
NaCl	0.5 g/l
$MgSO_4$	0.12 g/l
$CaCl_2$	0.01 g/l

L'essentiel est invisible pour les yeux.


3-OXOADIPATE-enol-LACTONE
3-OXO-ADIPATE
3-OXO-ADIPYL-CoA
SUCCINYL-CoA
ACETALDEHYDE
ACETYL-CoA
SUCCINATE
3-OXOADIPATE ENOL-LACTONASE
3-OXOADIPATE CoA-TRANSFERASE
CMP-N-ACETYL-NEURAMINATE (CMP-SIALATE)
CYTIDINE
POLYMERS (part 2)
DIPICOLINATE
DIHYDRODIPICOLINATE REDUCTASE
L-LYSINE
L-HOMOCYSTEINE
S-ADENOSYL-L-HOMOCYSTEINE
S-ADENOSYL-L-METHIONINE
L-METHIONINE
BETAINE-HOMOCYSTEINE METHYLTRANSFERASE
BETAINE
DIMETHYLGLYCINE
NICOTINATE METHYLTRANSFERASE
PYRUVATE
CYSTATHIONINE β-LYASE
CYSTATHIONINE β-SYNTHASE
L-SERINE
L-CYSTATHIONINE
O-ACETYLSERINE
3-MERCAPTOPYRUVATE
PYRUVATE
L-CYSTEINE
GLUTAMATE
2-OXOGLUTARATE
L-CYSTINE
THIOCYSTEINE
CARNOSINE
CARNOSINE SYNTHASE
L-HISTIDINE
β-ALANINE
(R)-PANTOATE
(R)-PANTOTHENATE
PANTOATE-β-ALANINE LIGASE
O-SUCCINYL-L-HOMOSERINE
2-OXOBUTYRATE
2-HYDROXYBUTYRATE
L-LACTATE DEHYDROGENASE
ISOLEUCINE SYNTHESIS
L-THREONINE
GLYCINE
D-ALANINE
L-α-ALANINE
ALANINE RACEMASE
L-ASPARAGINE
L-ASPARTATE
4-ASPARTYL-P
L-ASPARTATE-4-SEMIALDEHYDE
L-HOMOSERINE
O-PHOSPHO-L-HOMOSERINE
Aspartate group of amino acids
ASPARTATE KINASE
ASPARTATE SEMIALDEHYDE DEHYDROGENASE
HOMOSERINE DEHYDROGENASE
HOMOSERINE KINASE
THREONINE SYNTHASE
THREONINE ALDOLASE
ACETALDEHYDE
3-SULFINOALANINE
HYPOTAURINE
L-CYSTEATE
TAURINE
L-OXOSUCCINAMATE
2-OXOGLUTARATE
L-GLUTAMATE
ASPARTATE TRANSAMINASE
FUMARATE
SUCCINATE
SUCCINATE DEHYDROGENASE
SUCCINYL-CoA
ACETYL-CoA
PROTEINS
GLYCINE
5-AMINOLEVULINATE SYNTHASE
L-2-AMINO-ACETOACETATE
AMINOACETONE
S(−L)-MALATE
OXALOACETATE
MALATE DEHYDROGENASE
MALATE SYNTHASE
PHOSPHOENOLPYRUVATE
FORMYL-CoA
FORMATE
Glyoxylate cycle
OXALYL CoA
OXALATE
CITRATE
GLYCOLATE
GLYOXYLATE
2-OXOGLUTARATE
2-AMINO-3-OXOADIPATE
5-AMINOLEVULINATE
PORPHOBILINOGEN SYNTHASE
SUCCINATE SEMIALDEHYDE
4-HYDROXYBUTYRATE
(S)-GLUTAMATE 1-SEMIALDEHYDE
4-AMINOBUTYRATE
Glutamate decarboxylation pathway
2-HYDROXYGLUTARATE
OXALOSUCCINATE
ACETYL-CoA
CITRATE SYNTHASE
ISOCITRATE LYASE
ISOCITRATE DEHYDROGENASE (NADP)
ACONITATE HYDRATASE (ACONITASE)
2-OXOGLUTARATE DEHYDROGENASE SYSTEM
PROPIONYL-CoA
FORMATE
CARNITINE


From pictures of cells to mathematical models

Simulation models are simple pictures of cells, in a mathematical form

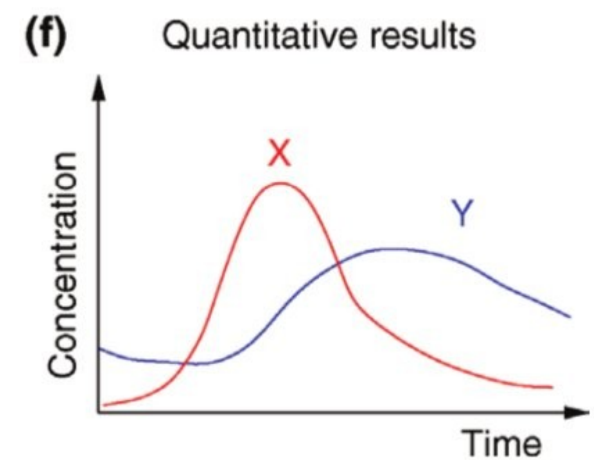


(d) Process model

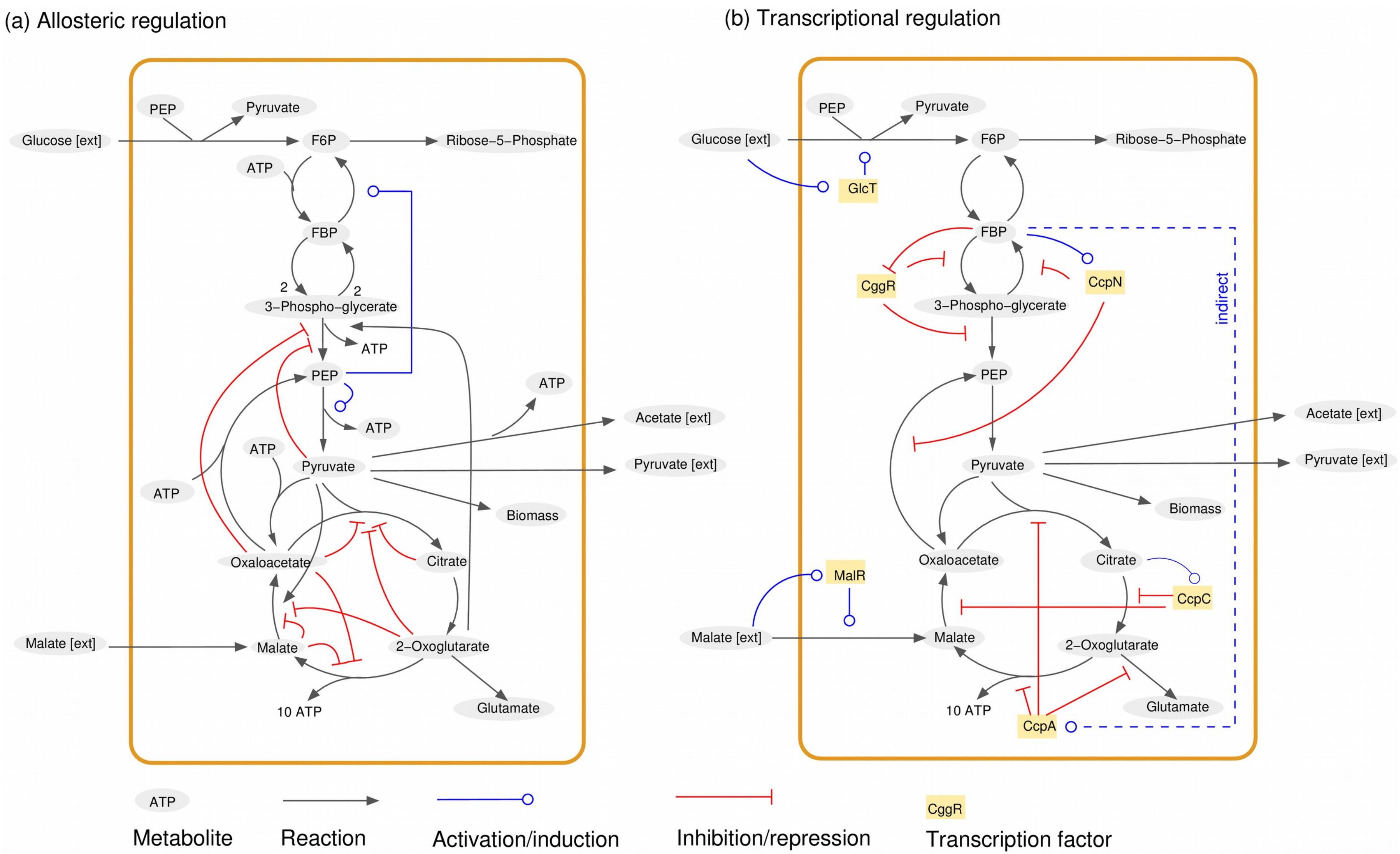
$$
\begin{aligned}
0 \rightarrow \mathrm{X} &\;:\; a_1 = v_0 \text{ (const.)} \\
\mathrm{X} \rightarrow 0 &\;:\; a_2 = \kappa_{\mathrm{x}} \cdot [\mathrm{X}] \\
0 \rightarrow \mathrm{Y} &\;:\; a_3 = f_{\mathrm{y}} \cdot [\mathrm{X}] \\
\mathrm{Y} \rightarrow 0 &\;:\; a_4 = \kappa_{\mathrm{y}} \cdot [\mathrm{Y}]
\end{aligned}
$$

(e) Dynamical model

$$
\begin{aligned}
\mathrm{d}x/\mathrm{d}t &= v_0 - \kappa_{\mathrm{x}}\, x \\
\mathrm{d}y/\mathrm{d}t &= f_{\mathrm{y}}\, x - \kappa_{\mathrm{y}}\, y \\
x(0) &= x_0 \\
y(0) &= y_0
\end{aligned}
$$

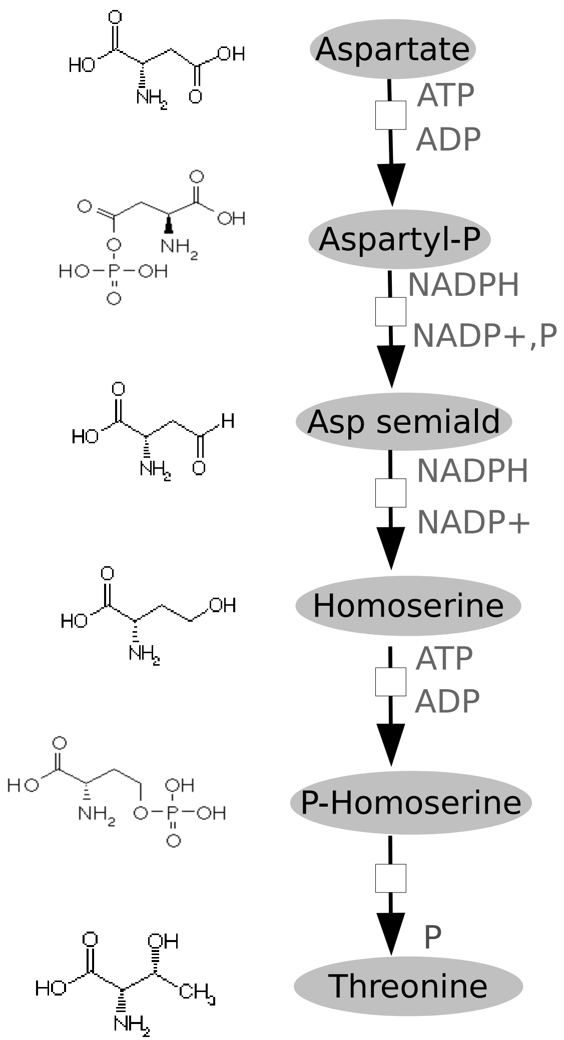


How can we translate network schemes into simulation models?

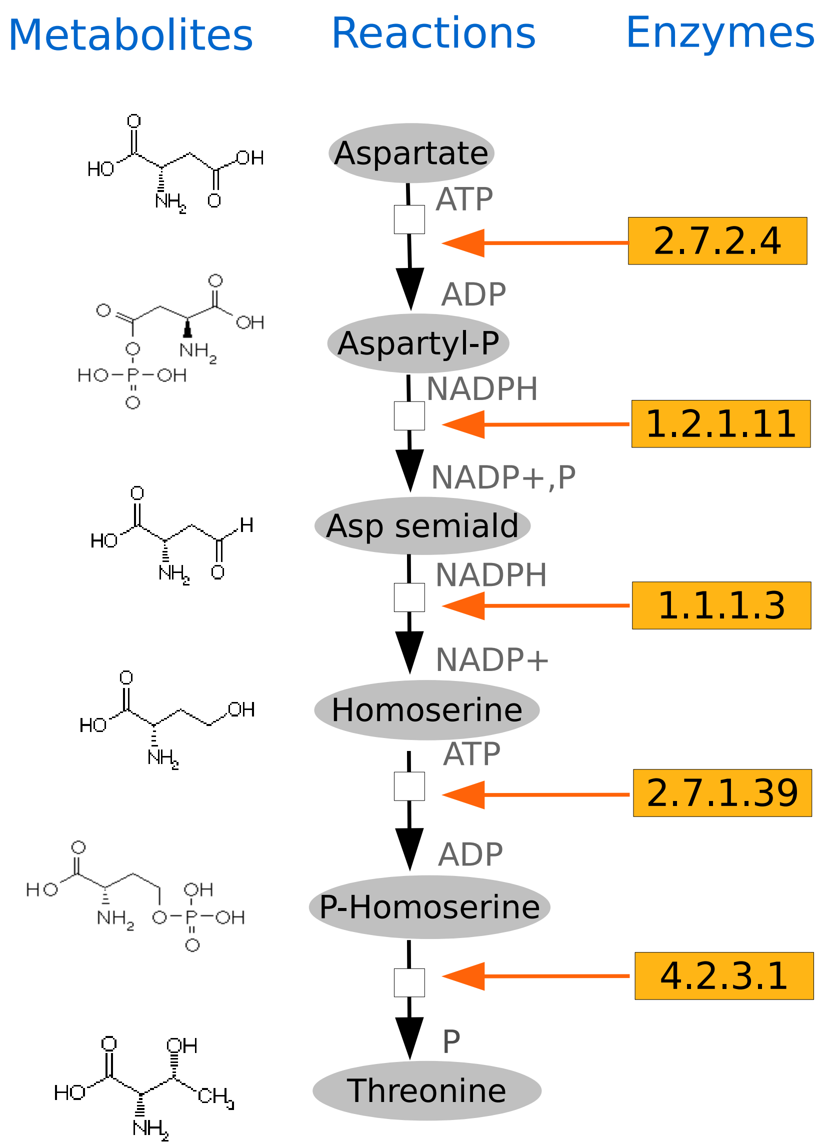


What aspects of a pathway do we want to model?

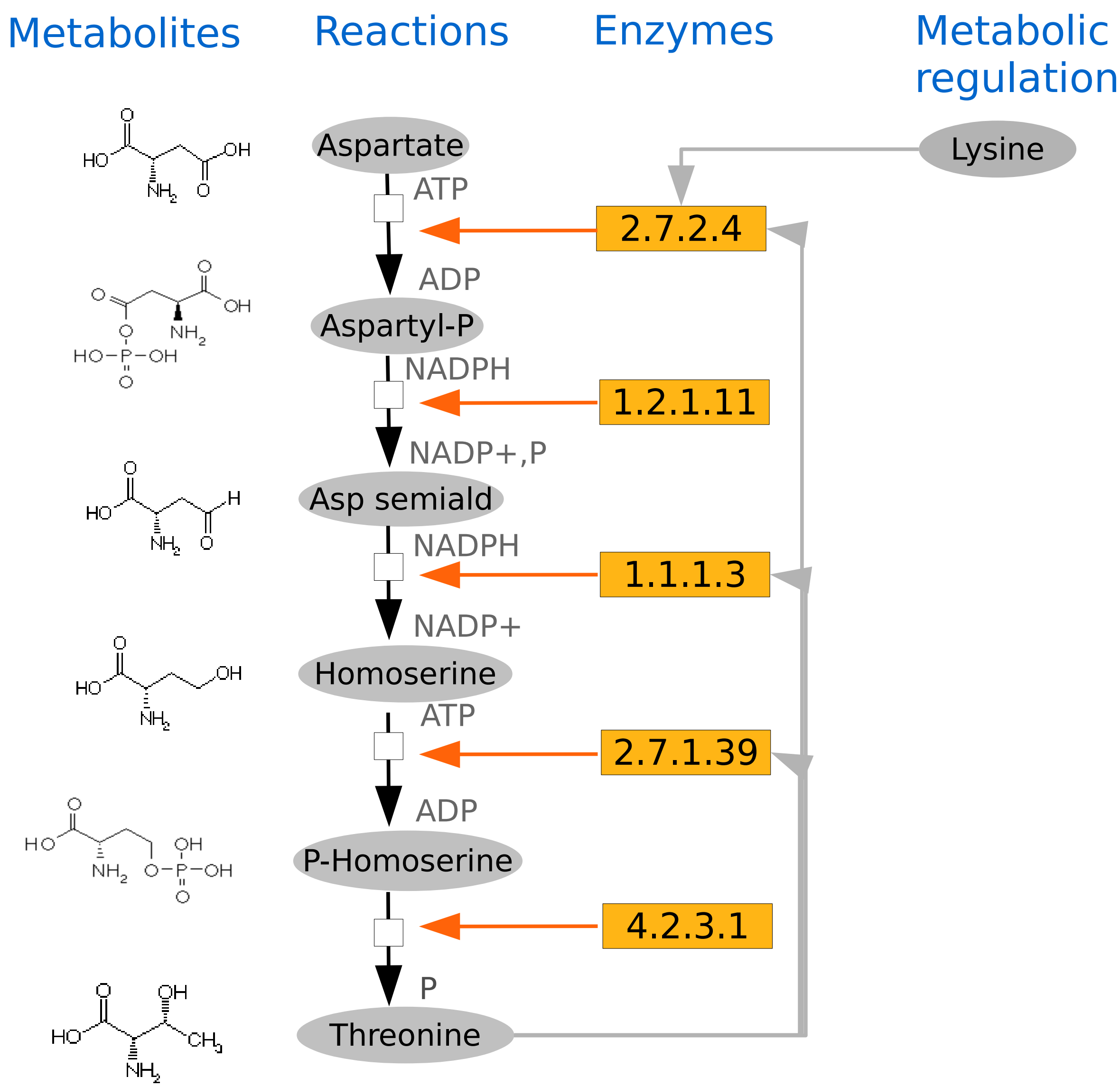
Metabolites Reactions



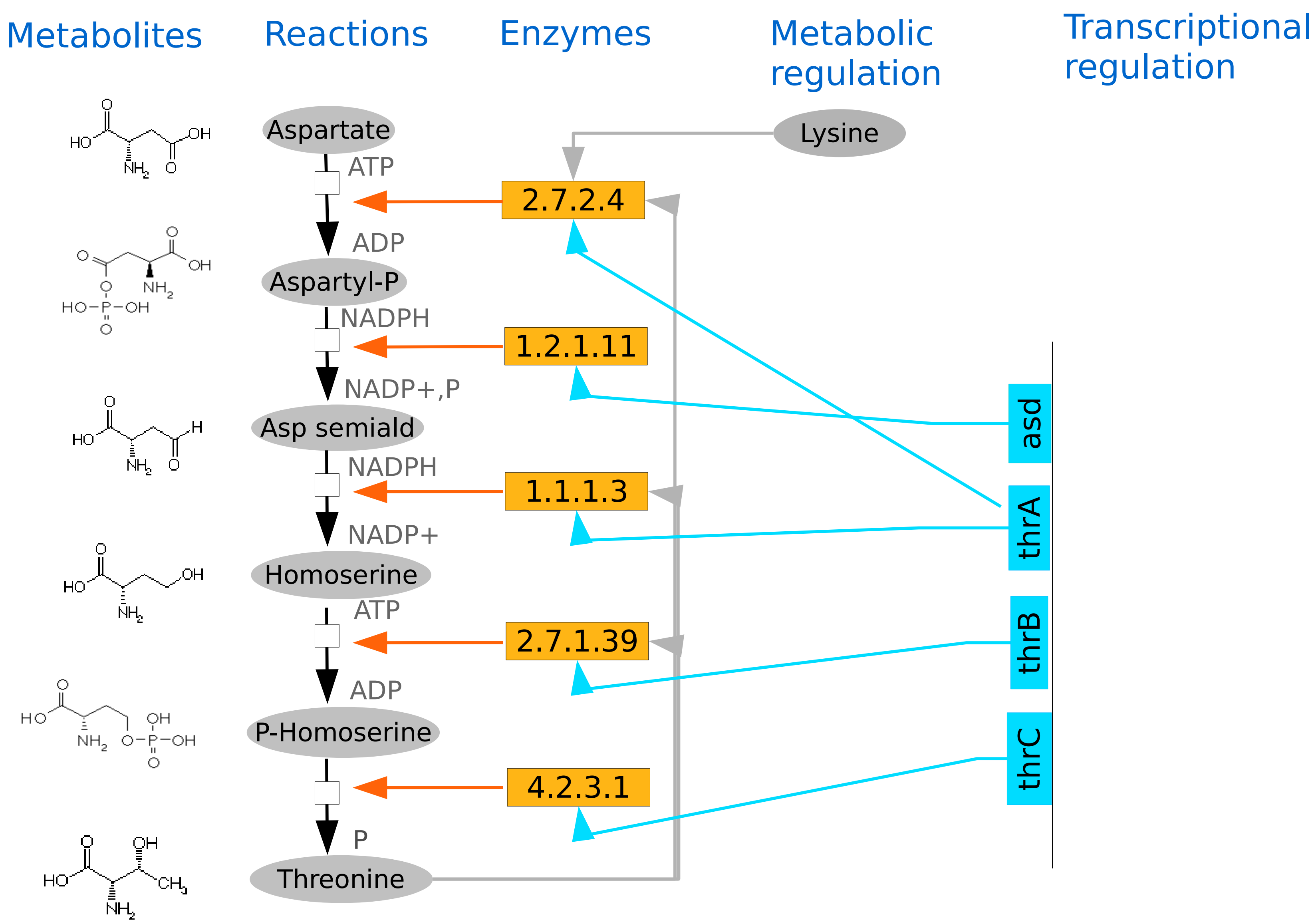
What aspects of a pathway do we want to model?



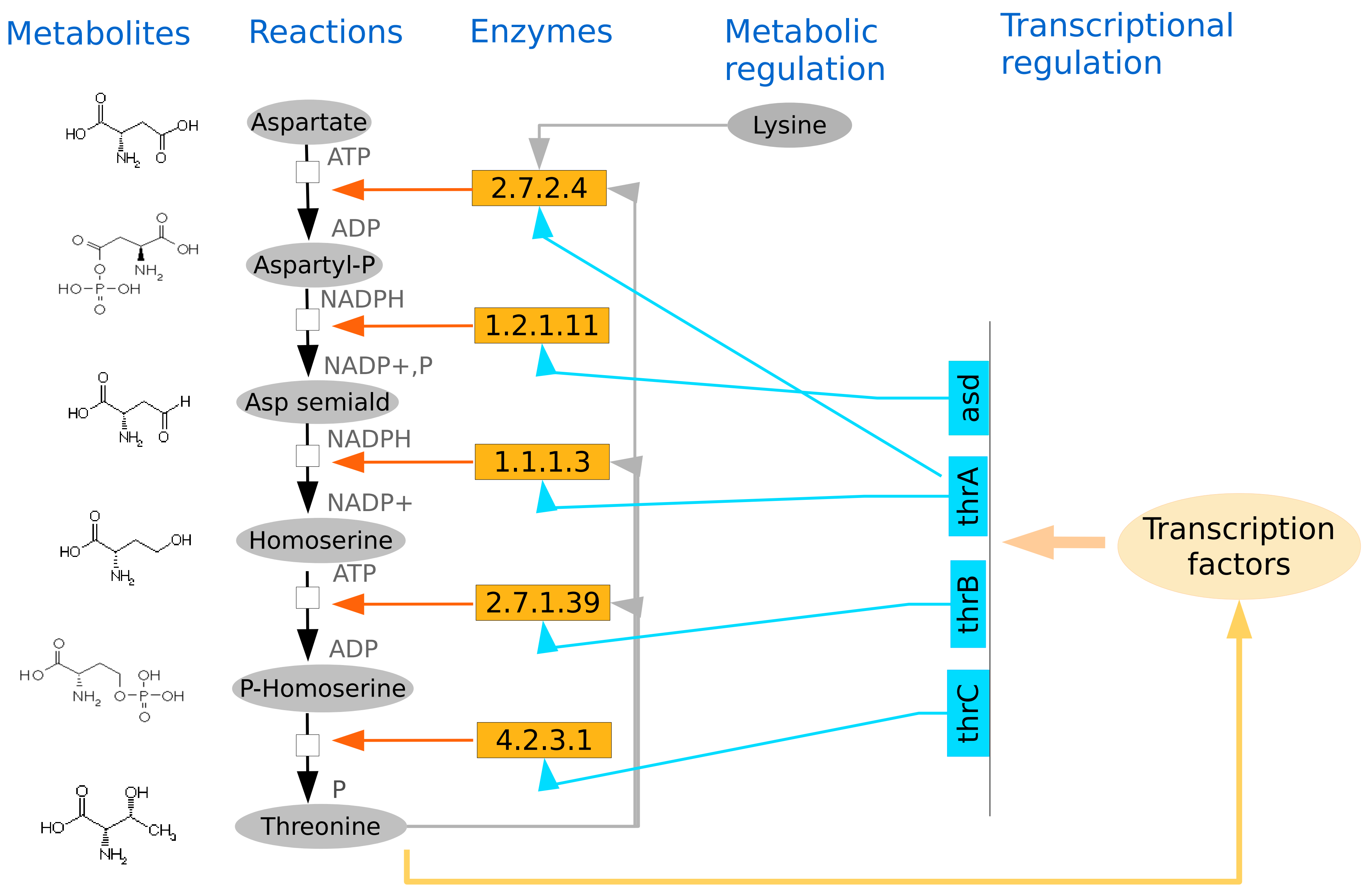
What aspects of a pathway do we want to model?



What aspects of a pathway do we want to model?

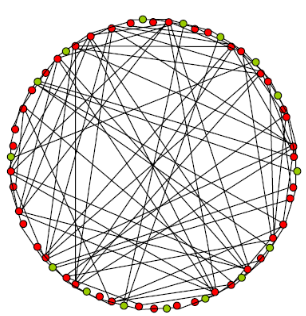


What aspects of a pathway do we want to model?

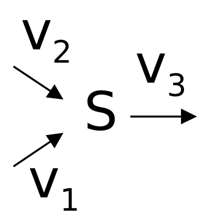


Modelling approaches cover different levels of complexity

Topological Analysis



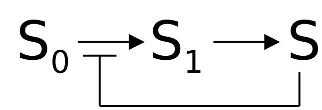
Flux Balance Analysis



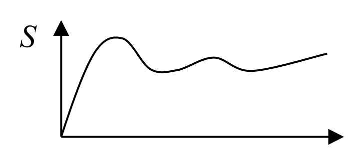
$$\frac{d\boldsymbol{S}}{dt} = \boldsymbol{N} \cdot \boldsymbol{v} = \boldsymbol{0}$$

$$v_1 + v_2 = v_3$$

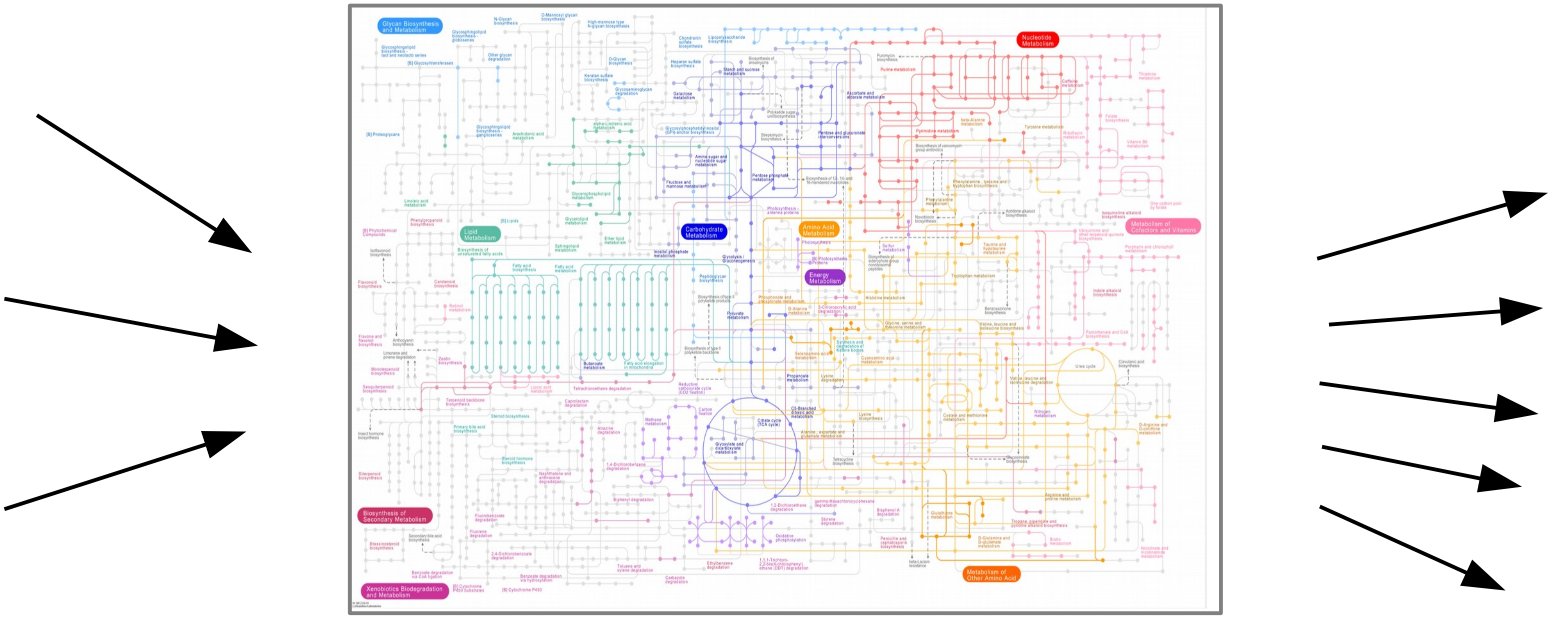
Kinetic modeling



$$\frac{d\boldsymbol{S}}{dt} = \boldsymbol{N} \cdot \boldsymbol{v}(\boldsymbol{S}, \boldsymbol{p})$$



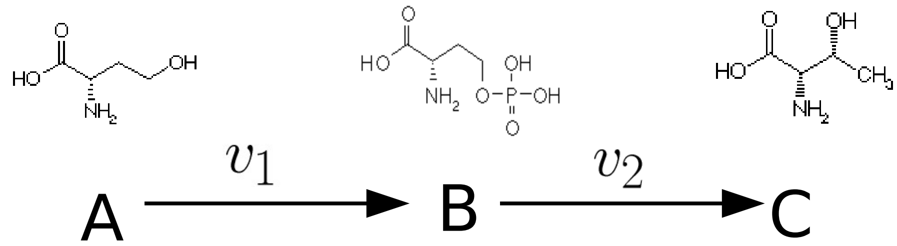
What kinds of questions do we want to answer?



- What compounds can the cell produce, and on what media can it survive?
- What do the metabolic fluxes look like?
- How do fluxes and metabolite levels respond to varying conditions?
- How would a mutation change the cell state?
- How big are the differences between individual cells?
- ...
- How can we answer all these questions with limited data?

Kinetic models of metabolic pathways

Kinetic models describe the dynamics of biochemical reactions

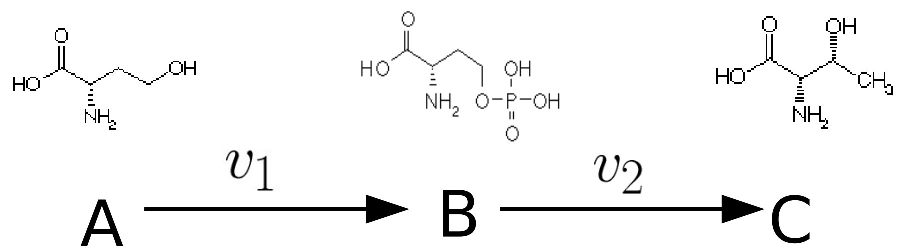


Homoserine

Phospho-
homoserine

Threonine

Kinetic models describe the dynamics of biochemical reactions



Homoserine

Phospho-homoserine

Threonine

Kinetic rate law: "mass-action kinetics"

How often does the reaction occur per time ?

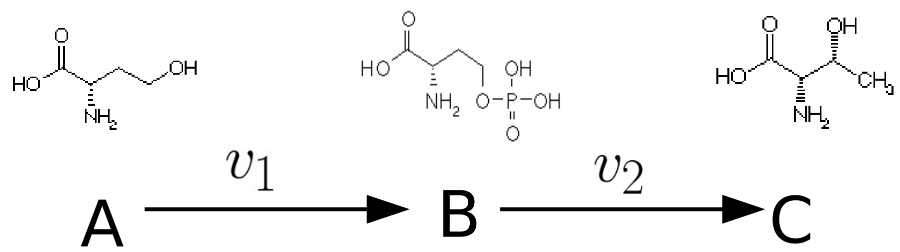
$$v_1(a,b) = k_{1+}\,a - k_{1-}\,b$$
$$v_2(b,c) = k_{2+}\,b - k_{2-}\,c$$

reaction rate

kinetic constant

metabolite concentration

Kinetic models describe the dynamics of biochemical reactions



Homoserine

Phospho-
homoserine

Threonine

Kinetic rate law: “mass-action kinetics”
How often does the reaction occur per time ?

$$v_1(a,b) = k_{1+}\,a - k_{1-}\,b$$
$$v_2(b,c) = k_{2+}\,b - k_{2-}\,c$$

reaction rate

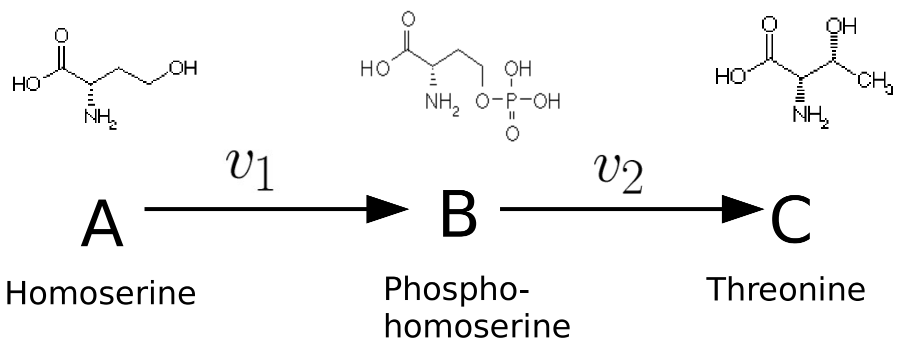
kinetic
constant

metabolite
concentration

System equations
How do the concentrations change over time?

$$\mathrm{d}a/\mathrm{d}t = -v_1$$
$$\mathrm{d}b/\mathrm{d}t = v_1 - v_2$$
$$\mathrm{d}c/\mathrm{d}t = v_2$$

Kinetic models describe the dynamics of biochemical reactions



Kinetic rate law: “mass-action kinetics”
How often does the reaction occur per time ?

$$v_1(a,b) = k_{1+}\,a - k_{1-}\,b$$
$$v_2(b,c) = k_{2+}\,b - k_{2-}\,c$$

reaction rate

kinetic constant

metabolite concentration

concentration change

stoichiometric coefficient

kinetic rate law (reaction rate)

$$\frac{\mathrm{d}s_i}{\mathrm{d}t} = \sum_l n_{il}\, v_l(s,k)$$

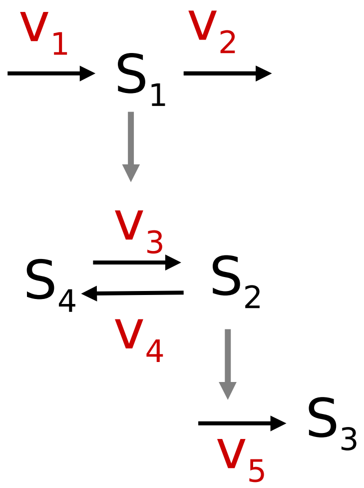
metabolite concentrations

kinetic parameters, enzyme concentrations

System equations
How do the concentrations change over time?

$$\mathrm{d}a/\mathrm{d}t = -v_1$$
$$\mathrm{d}b/\mathrm{d}t = v_1 - v_2$$
$$\mathrm{d}c/\mathrm{d}t = v_2$$

System equations – a more complicated example



Differential equations (ODEs)

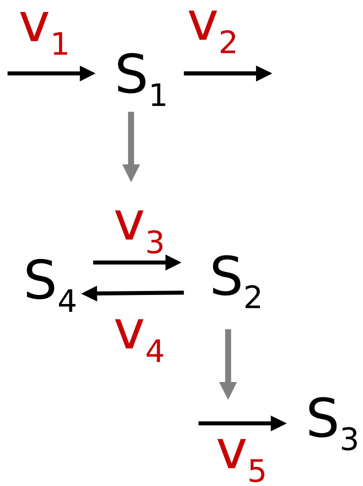
$d[S_1]/dt = v_1 - v_2$

$d[S_2]/dt = v_3 - v_4$

$d[S_3]/dt = v_5$

$d[S_4]/dt = - v_3 + v_4$

System equations – a more complicated example



Metabolite Concentrations

$$\vec{\boldsymbol{S}} = \begin{pmatrix} S_1 \\ S_2 \\ S_3 \\ S_4 \end{pmatrix}$$

Reaction rates

$$\vec{\boldsymbol{v}} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix}$$

Stoichiometric Matrix

$$\boldsymbol{N} = \begin{array}{cc} \begin{array}{ccccc} v_1 & v_2 & v_3 & v_4 & v_5 \end{array} & \\ \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 & 0 \end{pmatrix} & \begin{array}{c} S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \end{array}$$

Differential equations (ODEs)

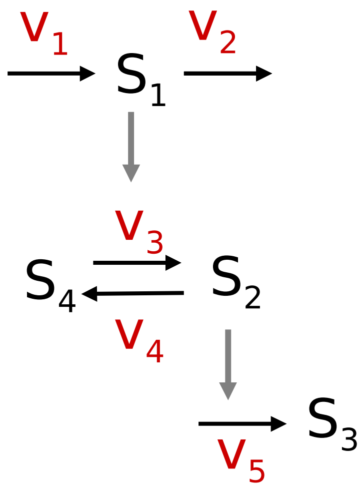
$d[S_1]/dt = v_1 - v_2$

$d[S_2]/dt = v_3 - v_4$

$d[S_3]/dt = v_5$

$d[S_4]/dt = - v_3 + v_4$

System equations – a more complicated example



Metabolite Concentrations

$$\vec{\boldsymbol{S}} = \begin{pmatrix} S_1 \\ S_2 \\ S_3 \\ S_4 \end{pmatrix}$$

Reaction rates

$$\vec{\boldsymbol{v}} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix}$$

Stoichiometric Matrix

$$\boldsymbol{N} = \begin{array}{cc} \begin{array}{ccccc} v_1 & v_2 & v_3 & v_4 & v_5 \end{array} & \\ \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 & 0 \end{pmatrix} & \begin{array}{c} S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \end{array}$$

Differential equations (ODEs)

$d[S_1]/dt = v_1 - v_2$

$d[S_2]/dt = v_3 - v_4$

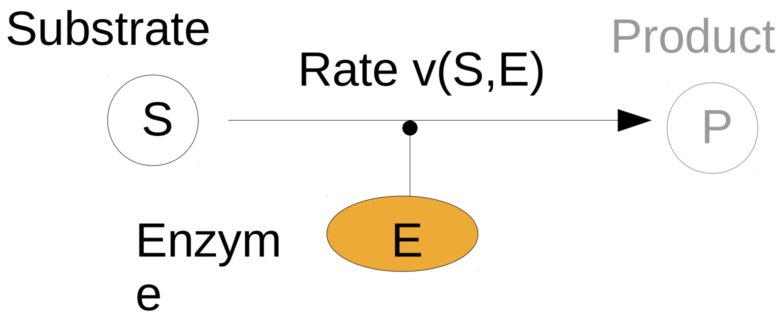
$d[S_3]/dt = v_5$

$d[S_4]/dt = - v_3 + v_4$

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 & 0 \end{pmatrix} \text{x} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix} = \begin{pmatrix} v_1 & -v_2 & +0 & +0 & +0 \\ 0 & +0 & +v_3 & -v_4 & +0 \\ 0 & +0 & +0 & +0 & v_5 \\ 0 & +0 & -v_3 & +v_4 & +0 \end{pmatrix}$$

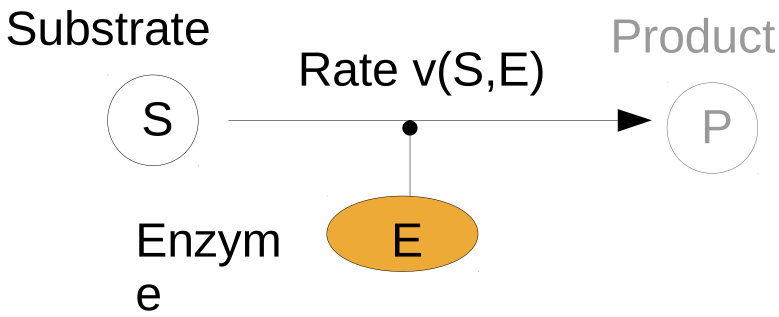
$$\boldsymbol{N} \quad \text{x} \quad \vec{\boldsymbol{v}} = \boldsymbol{d[\vec{S}]/dt}$$

The irreversible Michaelis-Menten rate law

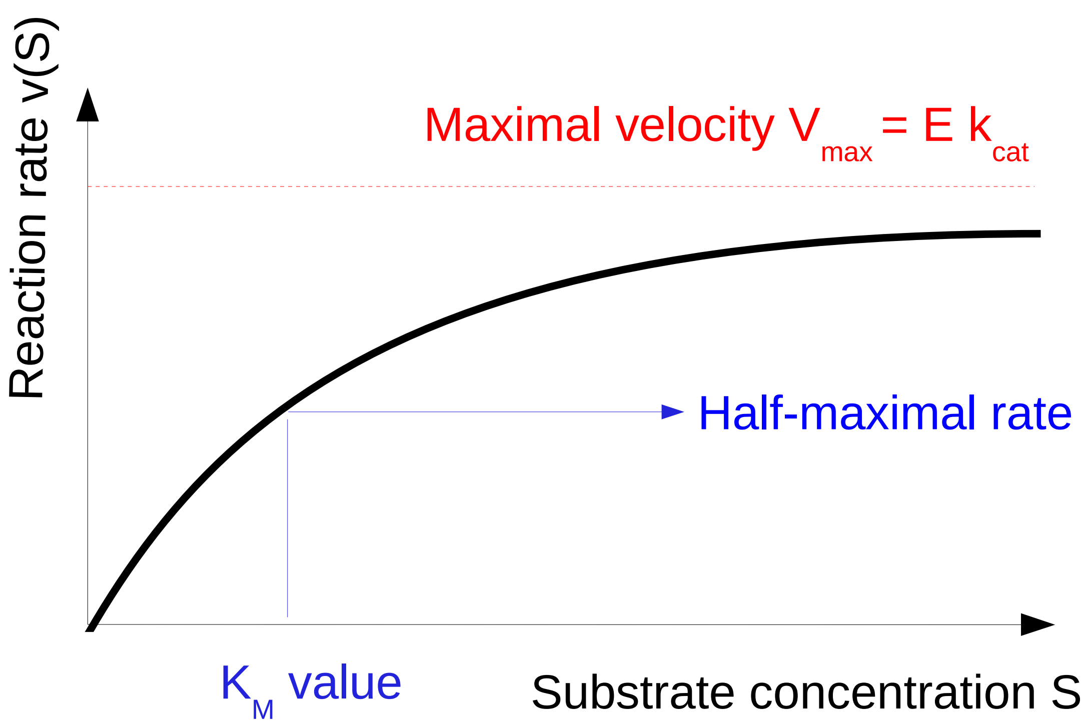


$$v(S,E) = \underbrace{E\,k_{\mathrm{cat}}}_{V_{\mathrm{max}}}\;\frac{S}{S+K_M}$$

The irreversible Michaelis-Menten rate law



$$v(S,E) = \underbrace{E\,k_{\mathrm{cat}}}_{V_{\mathrm{max}}} \frac{S}{S + K_M}$$



Variables:

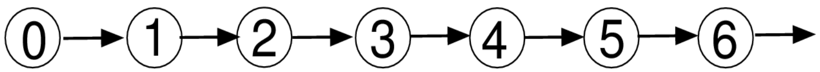
- Substrate concentration s
- Enzyme concentration E

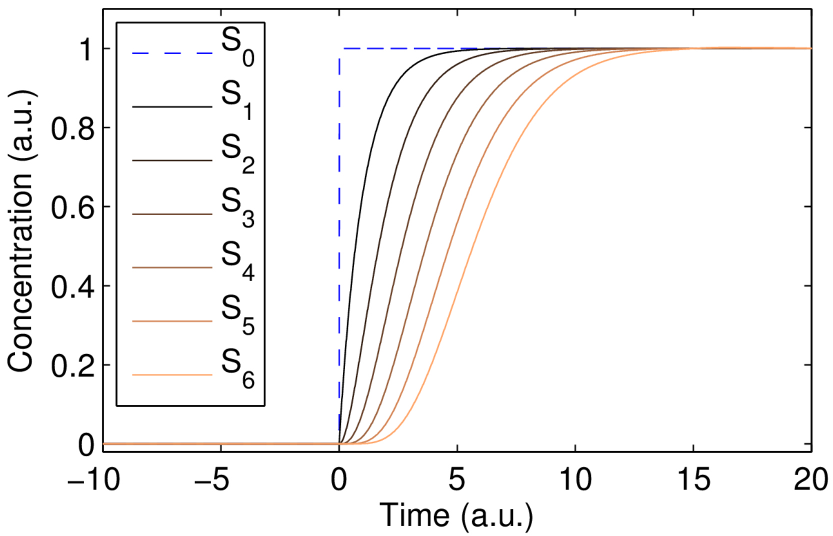
Parameters:

- K_M value (in mM): inverse binding affinity
- Catalytic constant k_{cat} (in 1/s)
  Maximal number of conversions per time and enzyme molecule

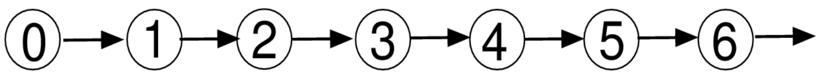
Dynamic behaviour and steady states

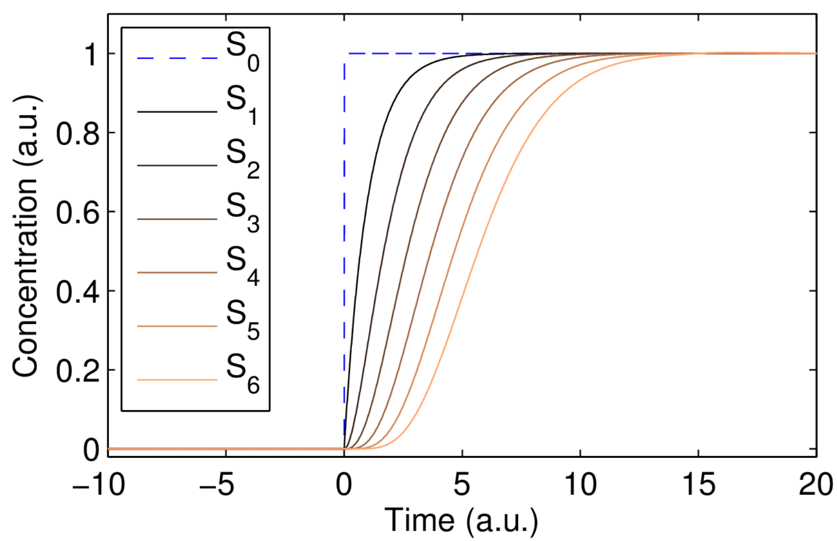
Differential equations describe the change in a moment - numerical integration yields the overall behaviour in time





Differential equations describe the change in a moment - numerical integration yields the overall behaviour in time





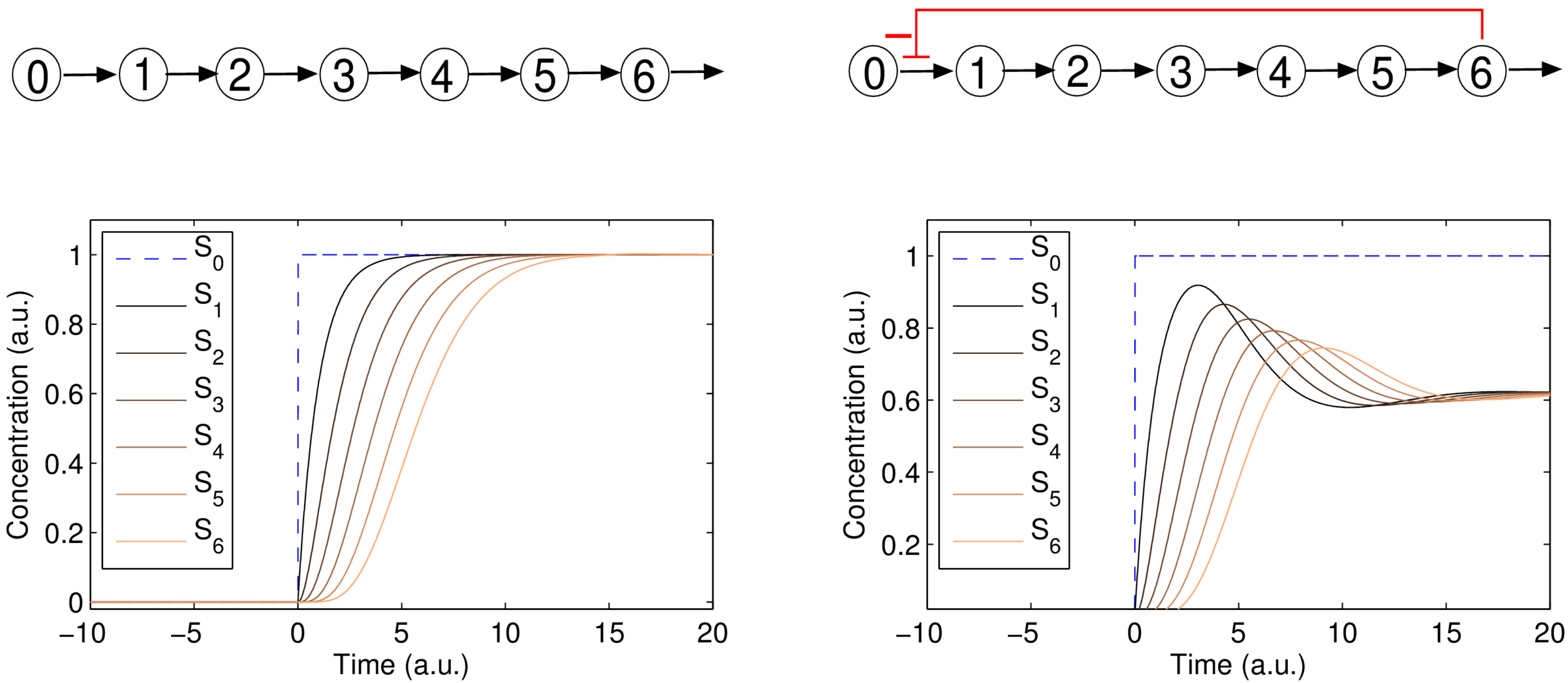
A simple way to solve differential equations numerically ("Euler method")

- Consider fixed, small time step!
- Start with initial values s(t=0)
- Use the updating rule:

$$\mathrm{s}(\mathrm{t} + \Delta t) = s(t) + \frac{\mathrm{d}s}{\mathrm{d}t}\,\Delta t$$

- Repeat the last step many times

Dynamic behaviour depends on small details of a model



In steady states, all substance levels remain constant in time

Stationarity condition in kinetic models

$$\frac{\mathrm{d}c}{\mathrm{d}t} = N\,v = 0$$

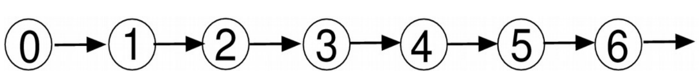
Condition on the flux vector
Kinetic rate laws do not play a role!

External metabolites (e.g. extracellular or buffered)
→ Treated as fixed parameters

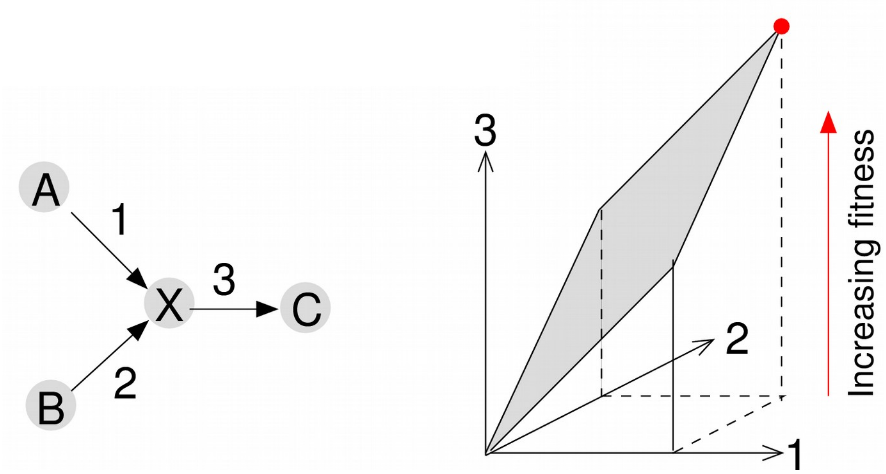
Intracellular metabolites (dynamic)
→ Concentration varies due to chemical reactions

Stationary (=steady) state
A state in which all variables remain constant in time

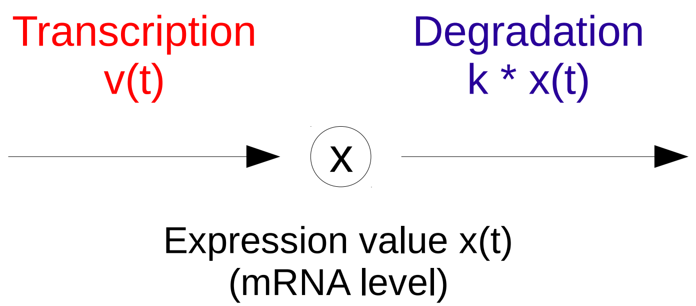
Linear pathway

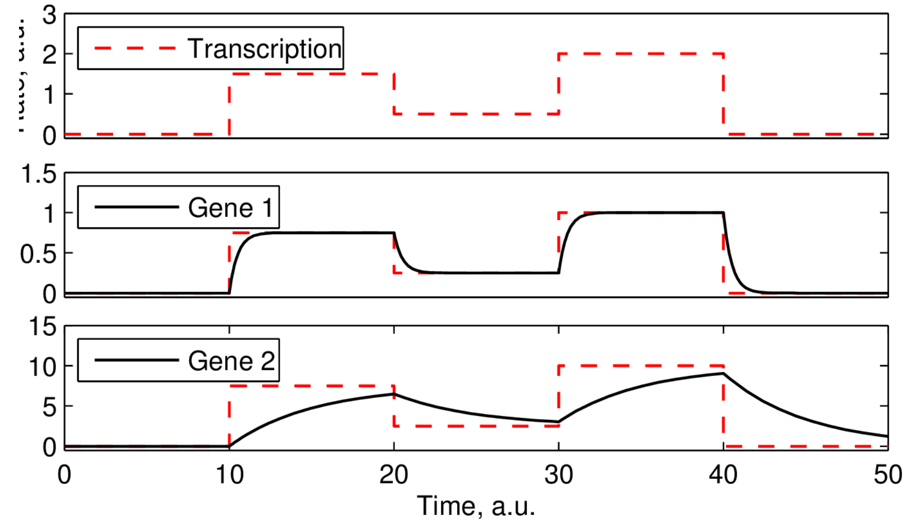


Branch point



An example: transcription rate and mRNA expression level



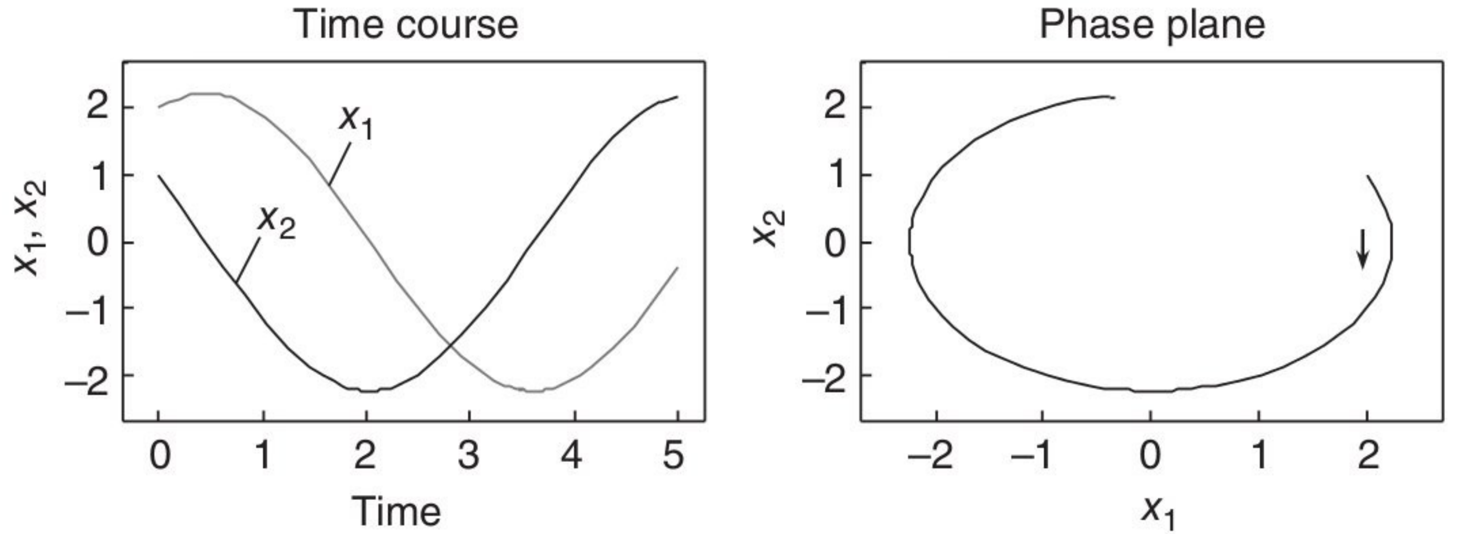


Exercise 1:
Write down the differential equation for x

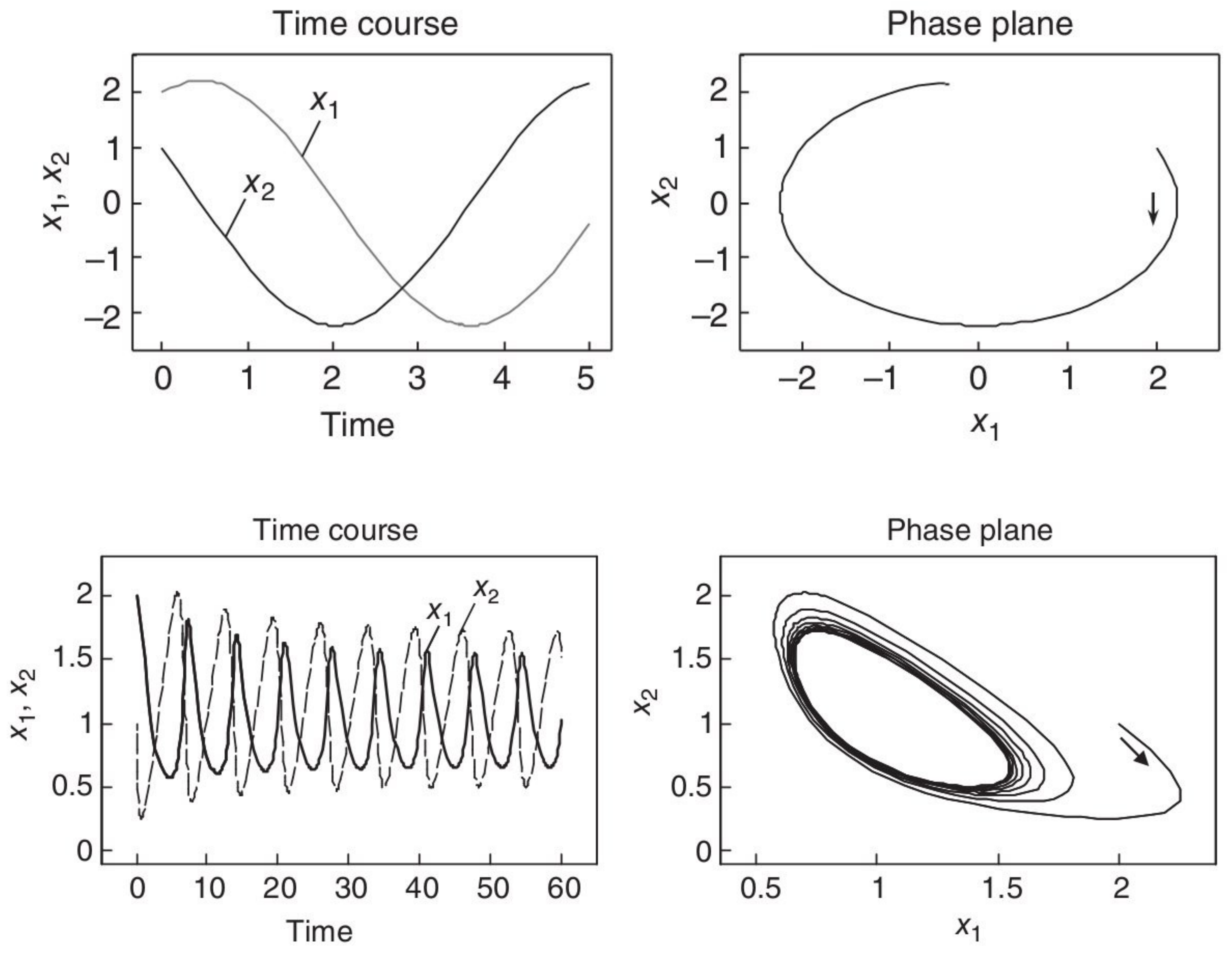
Exercise 2:
Solve the equation. Assume that x(0) = 10 nM, k = 1 /min, and v(t) = 0.

Exercise 3:
Assume a constant v(t) = 20 nM/s, k=0.1 / min, and determine the steady-state value of x.

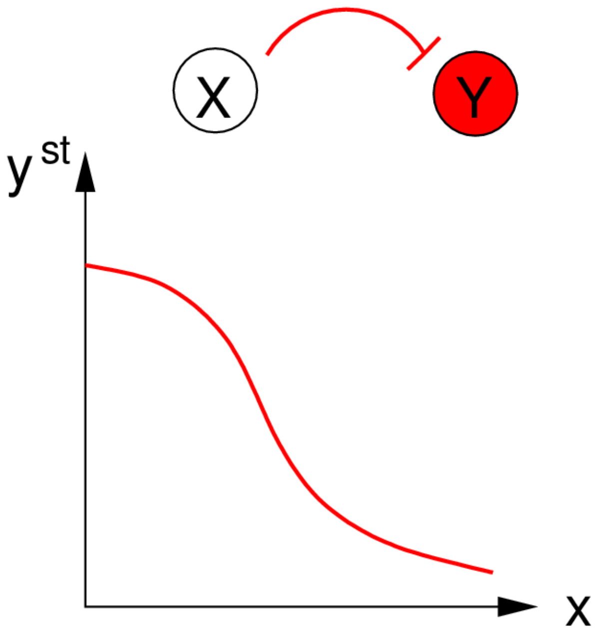
Dynamic behaviour in time and in phase space



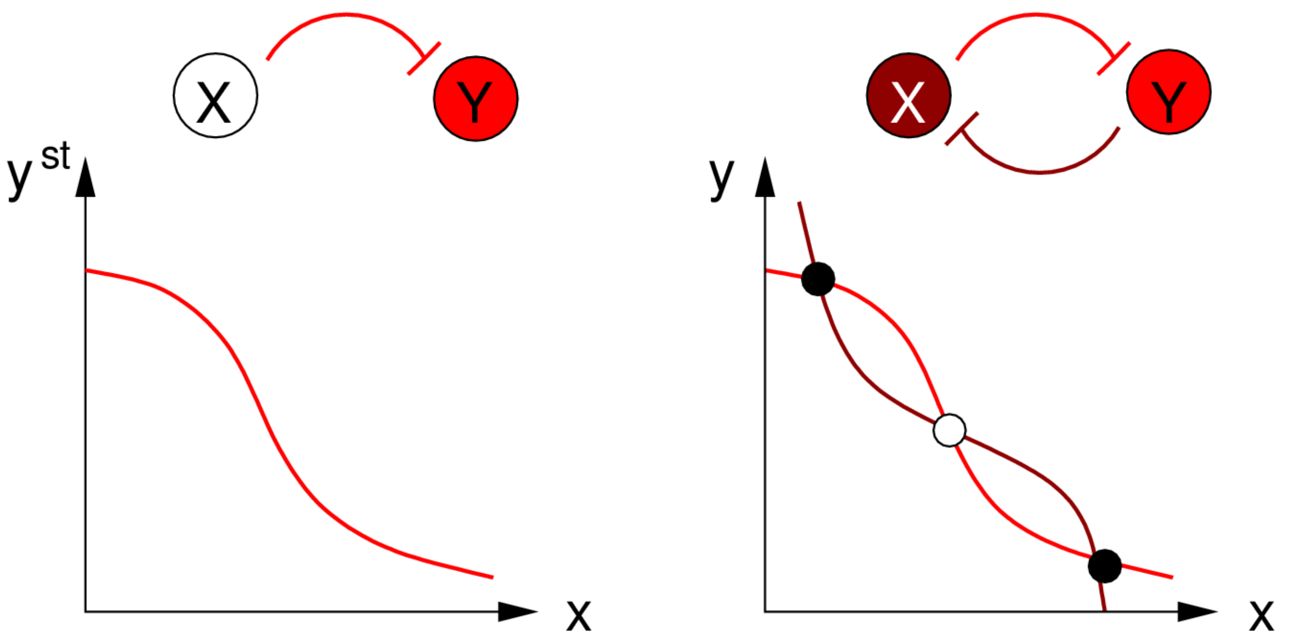
Dynamic behaviour in time and in phase space



Mutual inhibition can lead to bistability as a systemic behaviour



Mutual inhibition can lead to bistability as a systemic behaviour

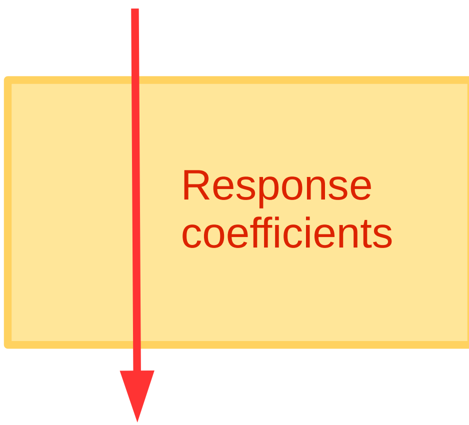


Metabolic control:
quantifying the effects of parameter changes

Metabolic control analysis studies the systemic effects of local parameter perturbations

Parameter change
higher substrate supply?

$$\Delta p_m$$



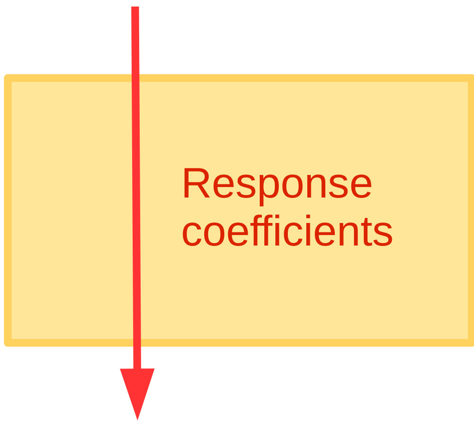
Metabolic change
altered concentrations?
redirected fluxes?

$$\Delta s_i \approx R^{s_i}_{p_m} \, \Delta p_m$$

Metabolic control analysis studies the systemic effects of local parameter perturbations

Parameter change
higher substrate supply?

$$\Delta p_m$$



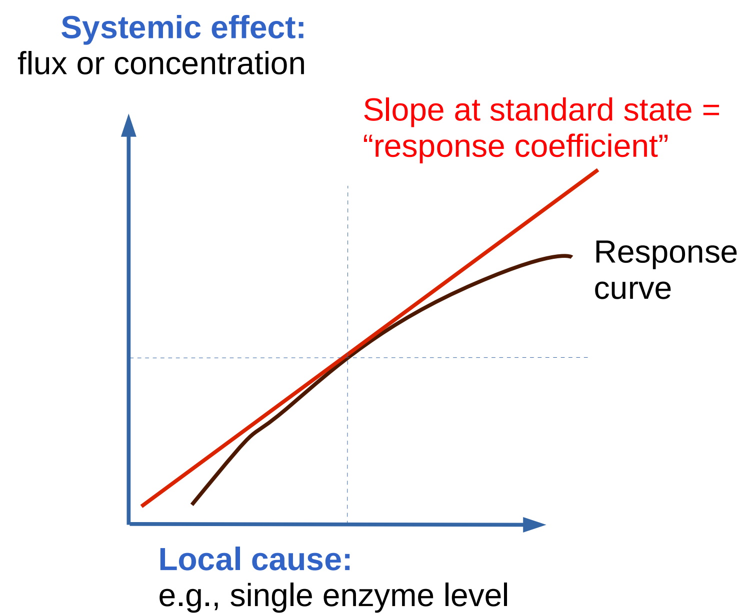
Metabolic change
altered concentrations?
redirected fluxes?

$$\Delta s_i \approx R^{s_i}_{p_m}\, \Delta p_m$$

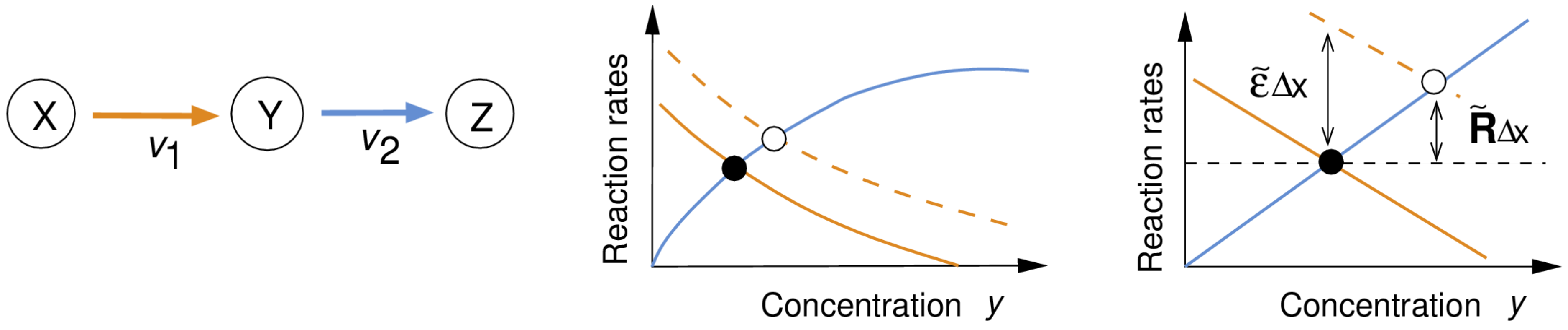
1. Stationary concentrations *s*(*p*)

Solution of $0 = N\, v(s(p), p)$

2. Response coefficients



Local perturbations, in the long run, change the entire metabolic state



Two types of sensitivities in metabolic control analysis:

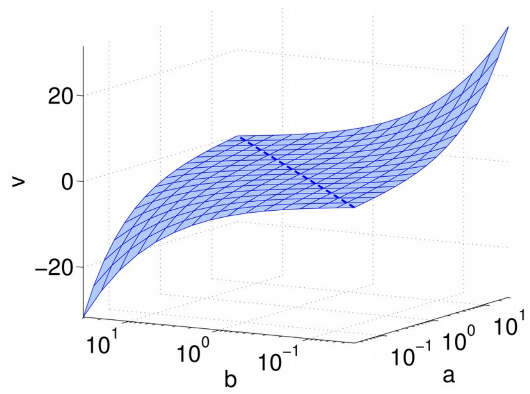
- Reaction elasticities
- Response (or control) coefficients

Model parameters, variability, and model structure

A problem in kinetic modelling: each enzyme is different !!

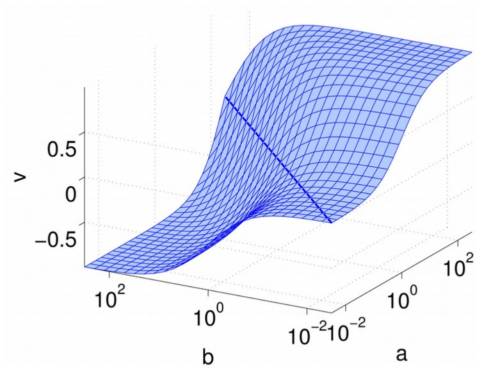
Reversible mass-action kinetics (non-enzymatic)

$$v = k_+ a - k_- b$$



Reversible Michaelis-Menten kinetics

$$v = \frac{v_+^{\max}(a/k_{\mathrm{A}}^{\mathrm{M}}) - v_-^{\max}(b/k_{\mathrm{B}}^{\mathrm{M}})}{1 + (a/k_{\mathrm{A}}^{\mathrm{M}}) + (b/k_{\mathrm{B}}^{\mathrm{M}})}$$

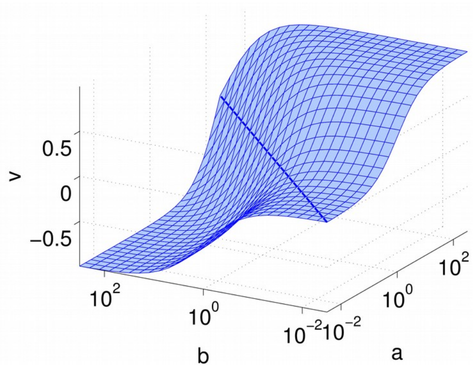


How can we obtain all the necessary parameters ??

Another problem: parameters may depend on each other!

Reversible Michaelis-Menten kinetics

$$v = \frac{v_+^{\max}(a/k_A^M) - v_-^{\max}(b/k_B^M)}{1 + (a/k_A^M) + (b/k_B^M)}$$



Thermodynamic constraints

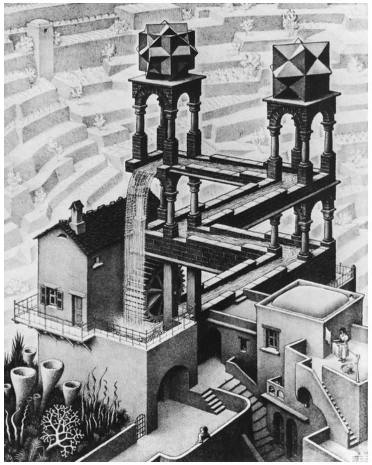
Thermodynamic laws lead to dependencies between kinetic parameters

Chemical equilibrium

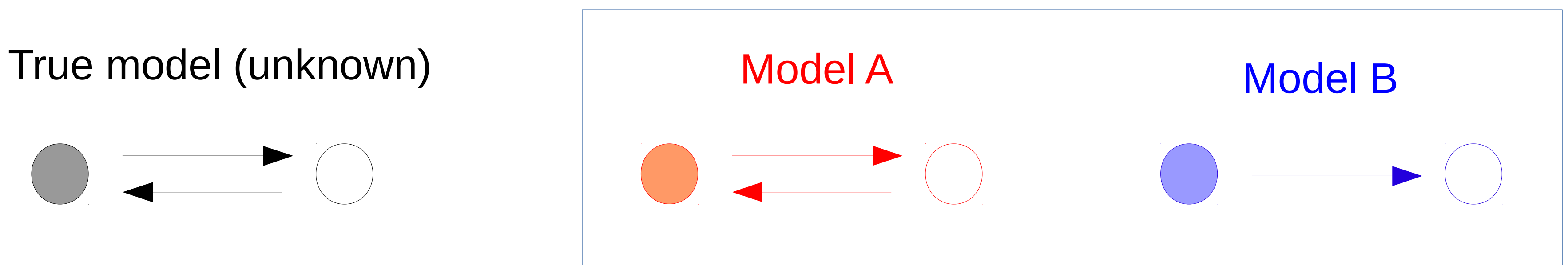
$$0 = v(a^{eq}, b^{eq}) = v_+^{\max}\frac{a^{eq}}{k_A^M} - v_-^{\max}\frac{b^{eq}}{k_B^M}$$

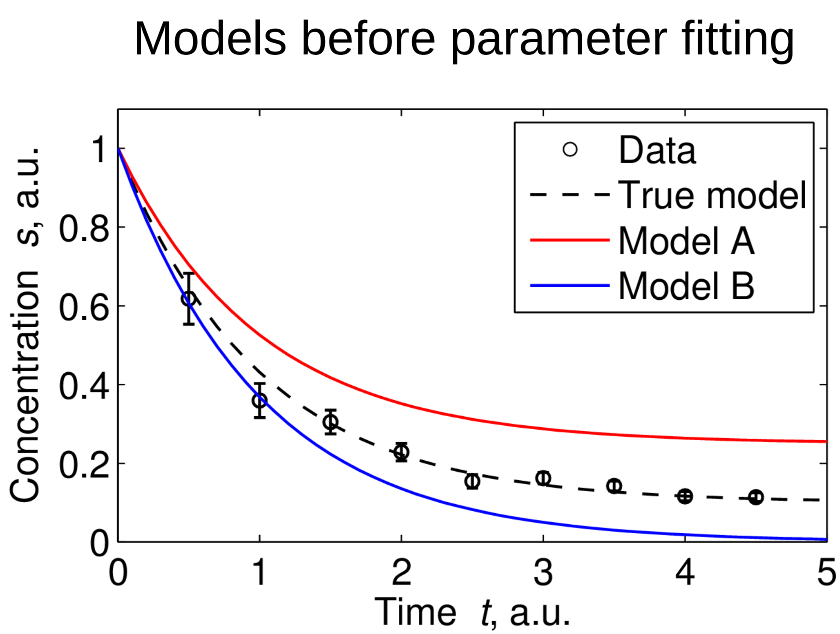
Haldane relationship

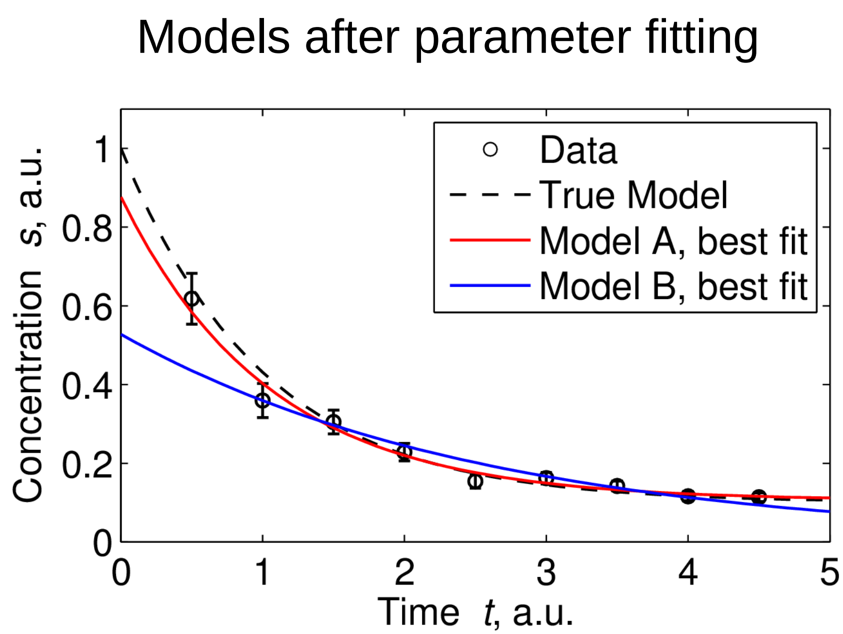
$$k^{eq} = \frac{b^{eq}}{a^{eq}} = \frac{v_+^{\max}k_B^M}{v_-^{\max}k_A^M}$$



How can we choose between two models?

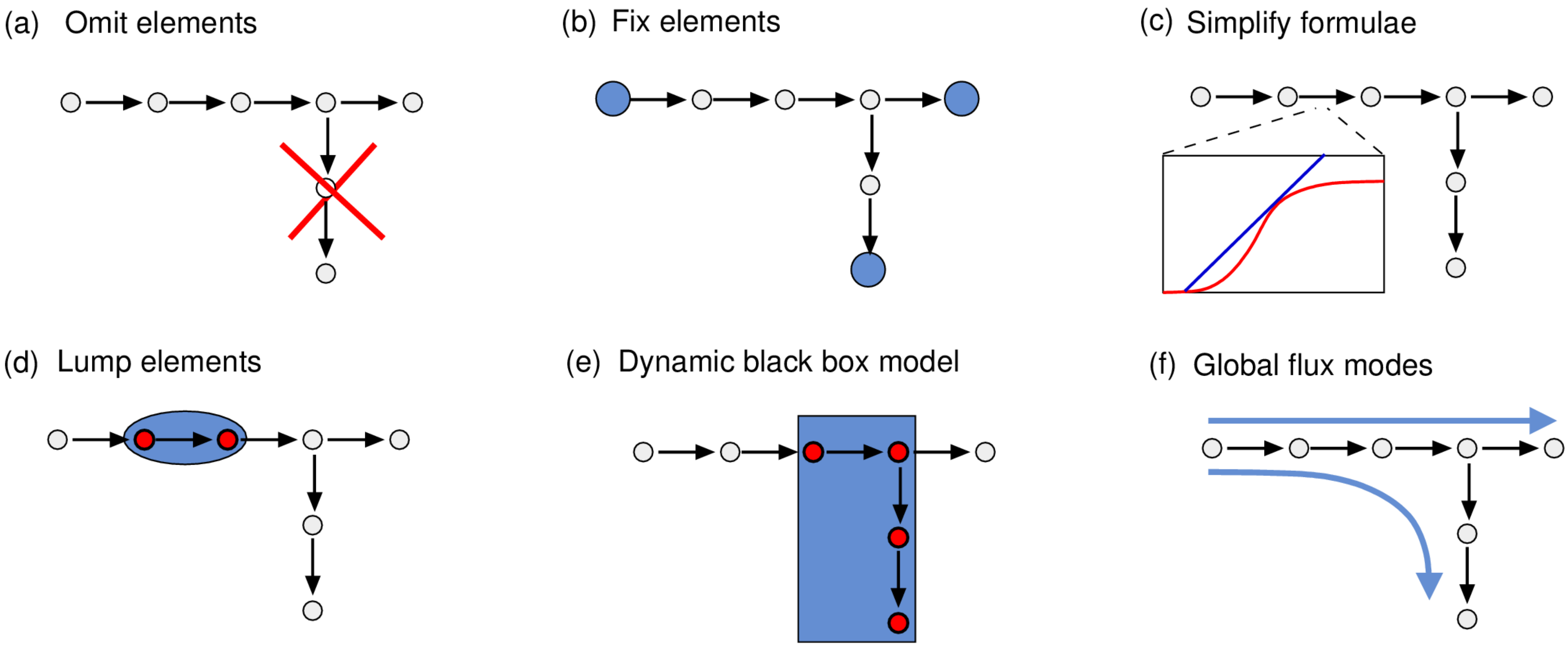




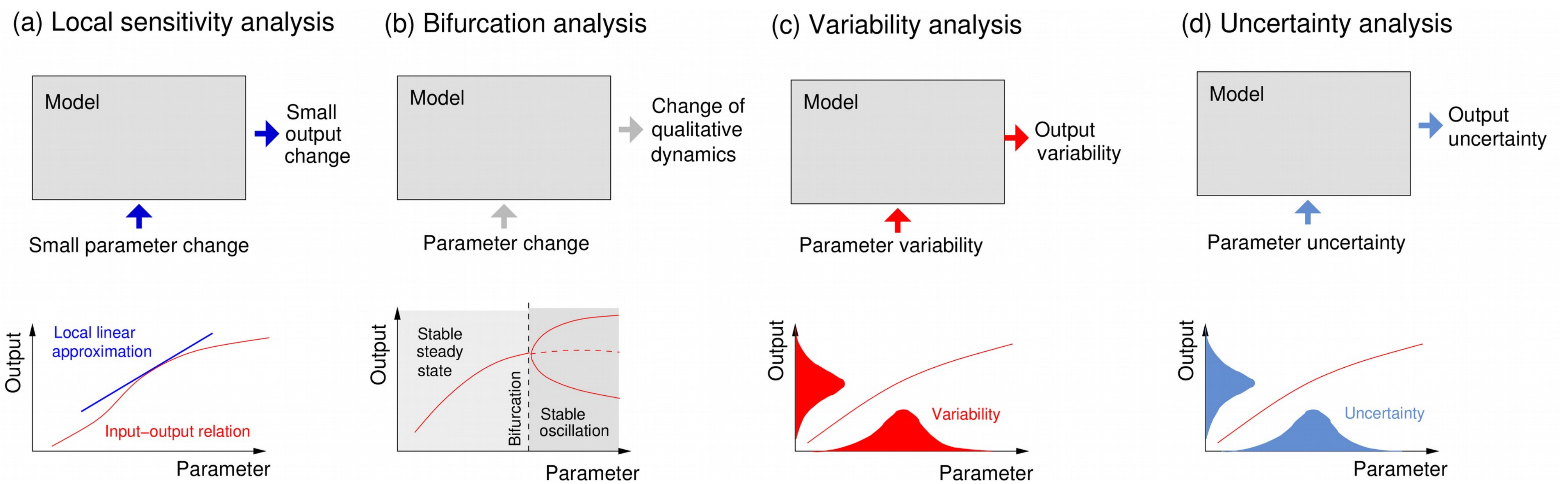


Some methods for model selection: Crossvalidation – "Selection criteria" – Baysian model selection

How models can be simplified
(hopefully, without losing too much accuracy)



Variability and uncertainty of parameters can be mathematically described



Some questions we might care about:

- What parameters have a strong effect on model behaviour?
- What model outputs are strongly affected?
- Under what parameter changes does the qualitative behaviour change, and how?
- If a parameter varies between cells, how much variation do we expect in the model output?
- If we are uncertain about a parameter, how uncertain will we be about model outputs?

Thank you !