Lecture: Computational Systems Biology
Universität des Saarlandes, SS 2012

02 Models of biochemical systems [part 1]

Dr. Jürgen Pahle

8.5.2012

Recap

- In Systems Biology we, *inter alia*, represent processes in living systems using models
- Models can be of various types and use different mathematical formalisms
- Modelling is an iterative process including literature and database search, experimental measurements, simulations, model refinement, etc.
- Models can be used to predict and also to study and understand biological processes (exploratory/explanatory models)

Chemical reactions

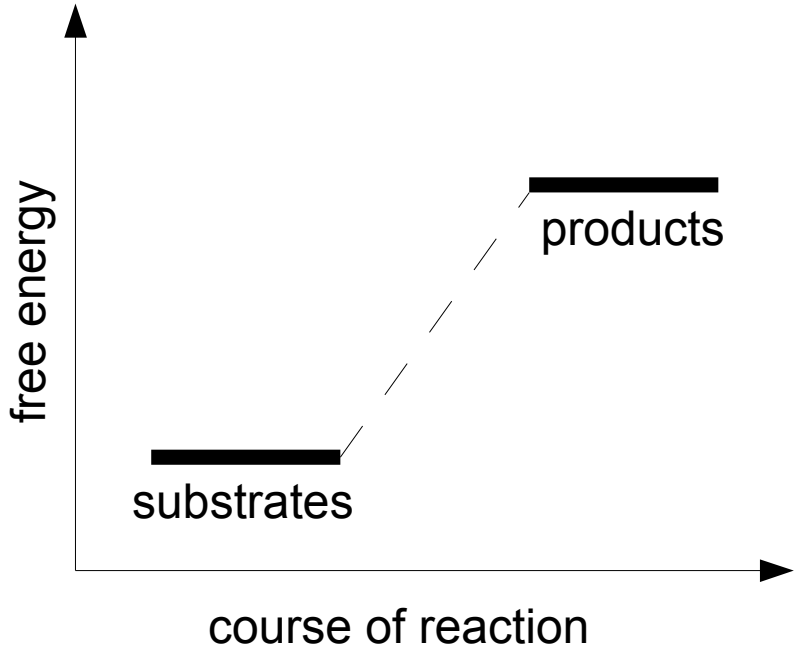
A process that leads to the transformation of one set of chemical substances to another
(IUPAC definition)

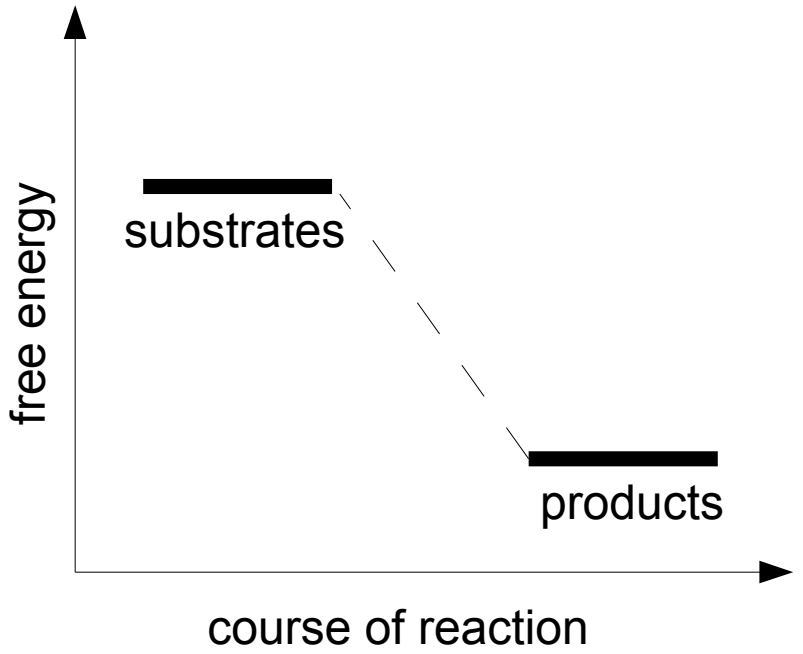
Spontaneous ↔ Non-spontaneous (needs energy input)

Why do some reactions occur spontaneously and others not?

Chemical reactions (cont.)

- Each chemical system has chemical potential energy ("Gibbs free energy" energy capable of carrying out work under isotherm-isobar conditions)
- Low potential energy → high potential energy (endergonic process)
- High potential energy → low potential energy (exergonic process)

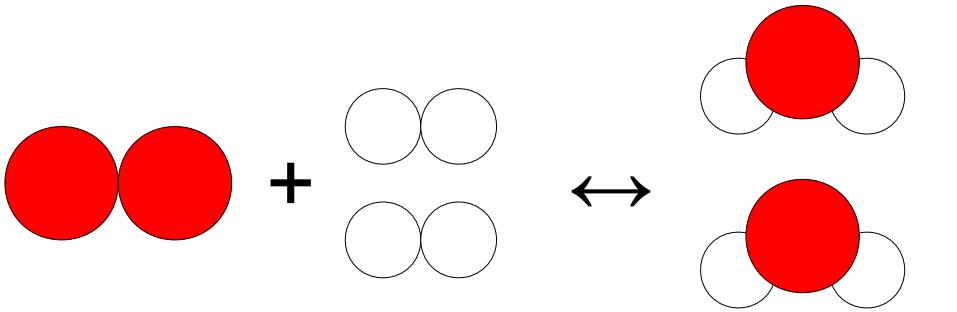




Chemical reactions (cont.)

- Reactions run until an equilibrium state is reached (still conversions going on)
- Equilibrium depends on the energy difference between the states
  - Large difference: in the equilibrium the lower energy state is strongly preferred
  - Small difference: both states can occur in roughly similar amounts

Example: Oxyhydrogen reaction



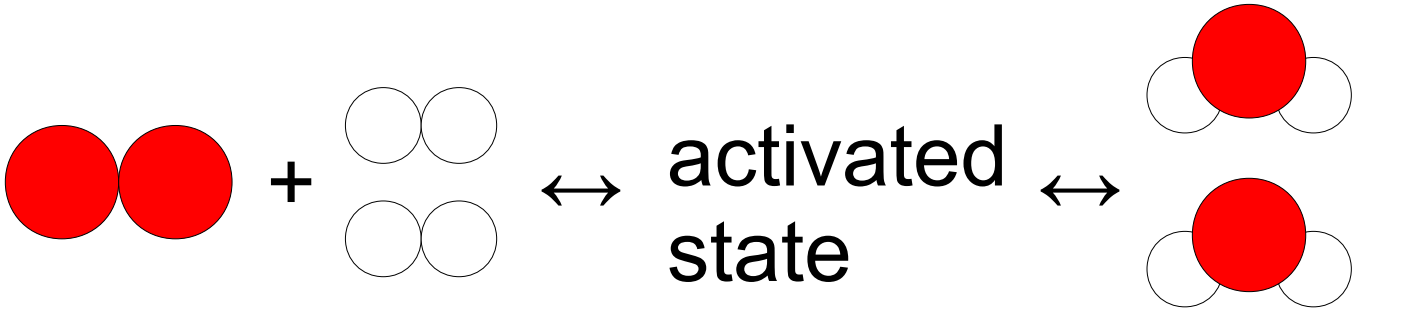
O_2 $2\ H_2$ $2\ H_2O$

Reaction is strongly exergonic & exothermic

$\Delta H = -571.6$ kJ/mol

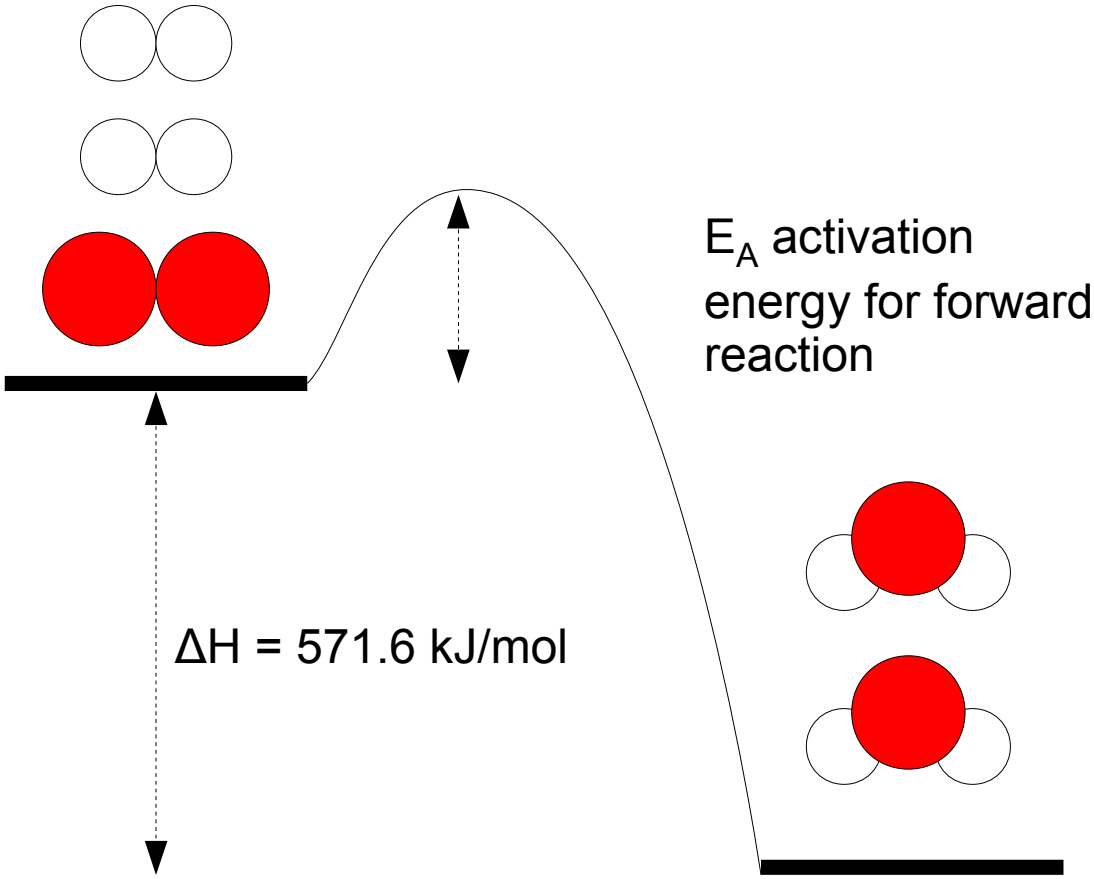
Example: Oxyhydrogen reaction

- Substrates are usually not converted directly to products
- Process often involves one or more elementary steps
- In the simplest case substrates first react to an intermediate (activated) state



Activation energy

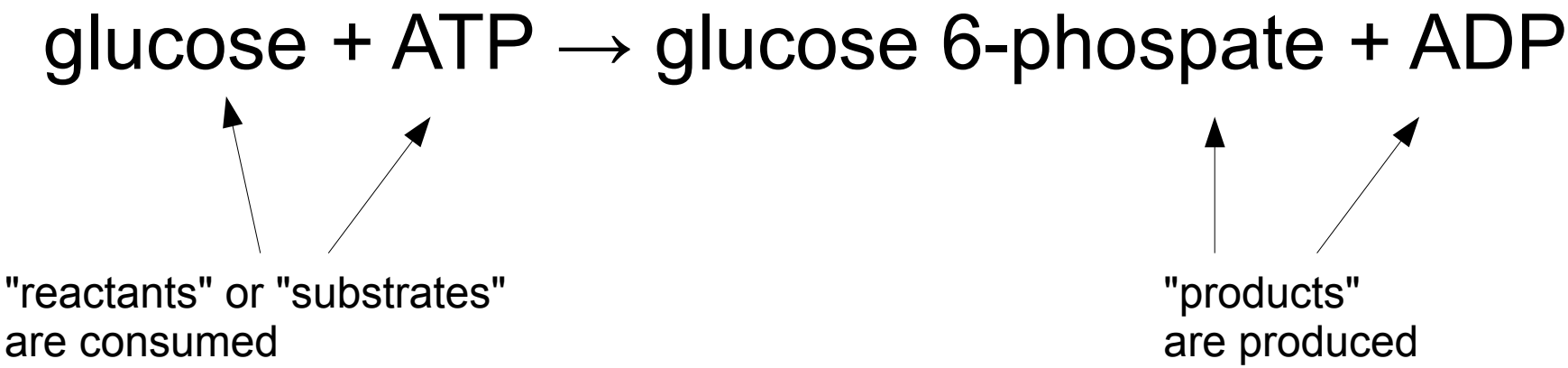
Reaction needs energy to overcome the activation energy barrier



Biochemical reactions

Cells are active compartments where simultaneously thousands of chemical **reactions** take place converting different types of **chemical species** into each other

Chemical species: molecules, ions, radicals, e.g. sugars, amino acids, peroxide, calcium, etc.



Biochemical pathways

Reactions are linked by species:
this leads to **pathways** of reactions

glucose + ATP → glucose 6-phospate + ADP

glucose 6-phosphate ↔ fructose 6-phospate

fructose 6-phosphate + ATP →
fructose 1,6-bisphosphate + ADP

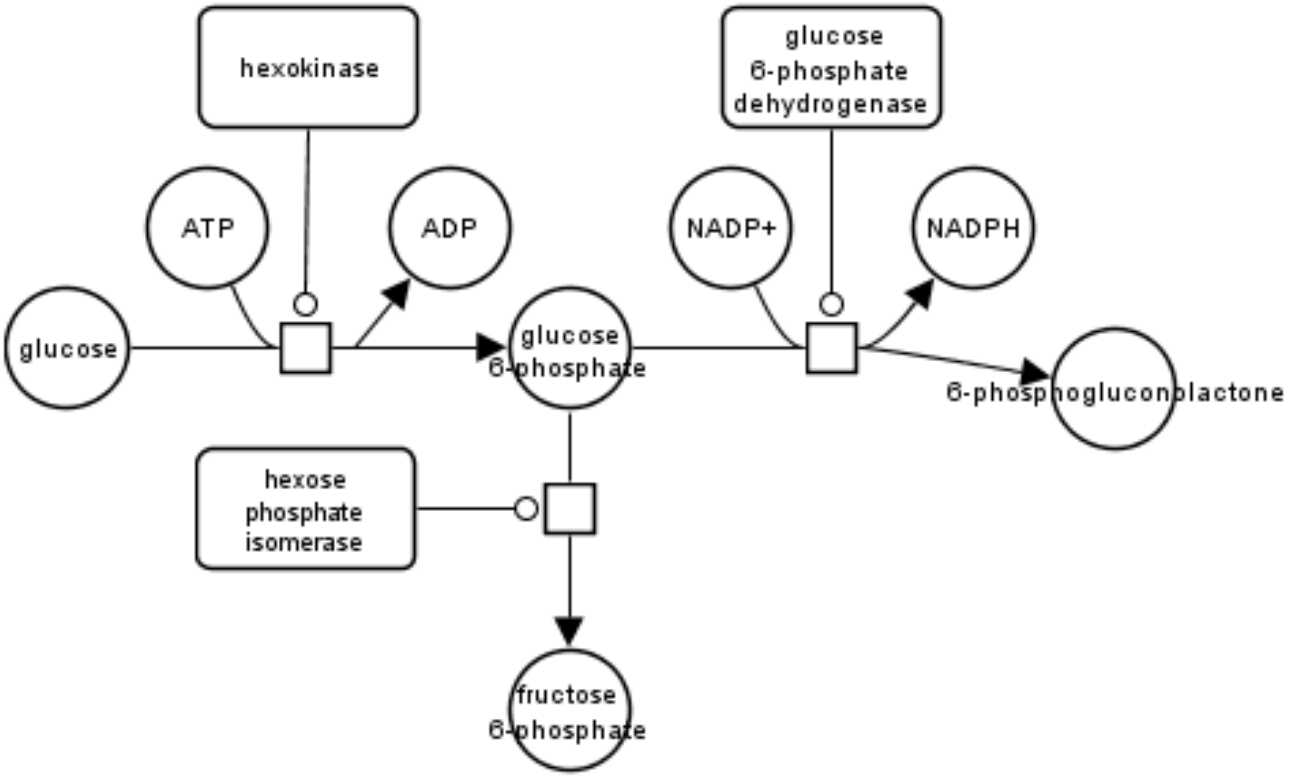
Biochemical networks

Branching points:

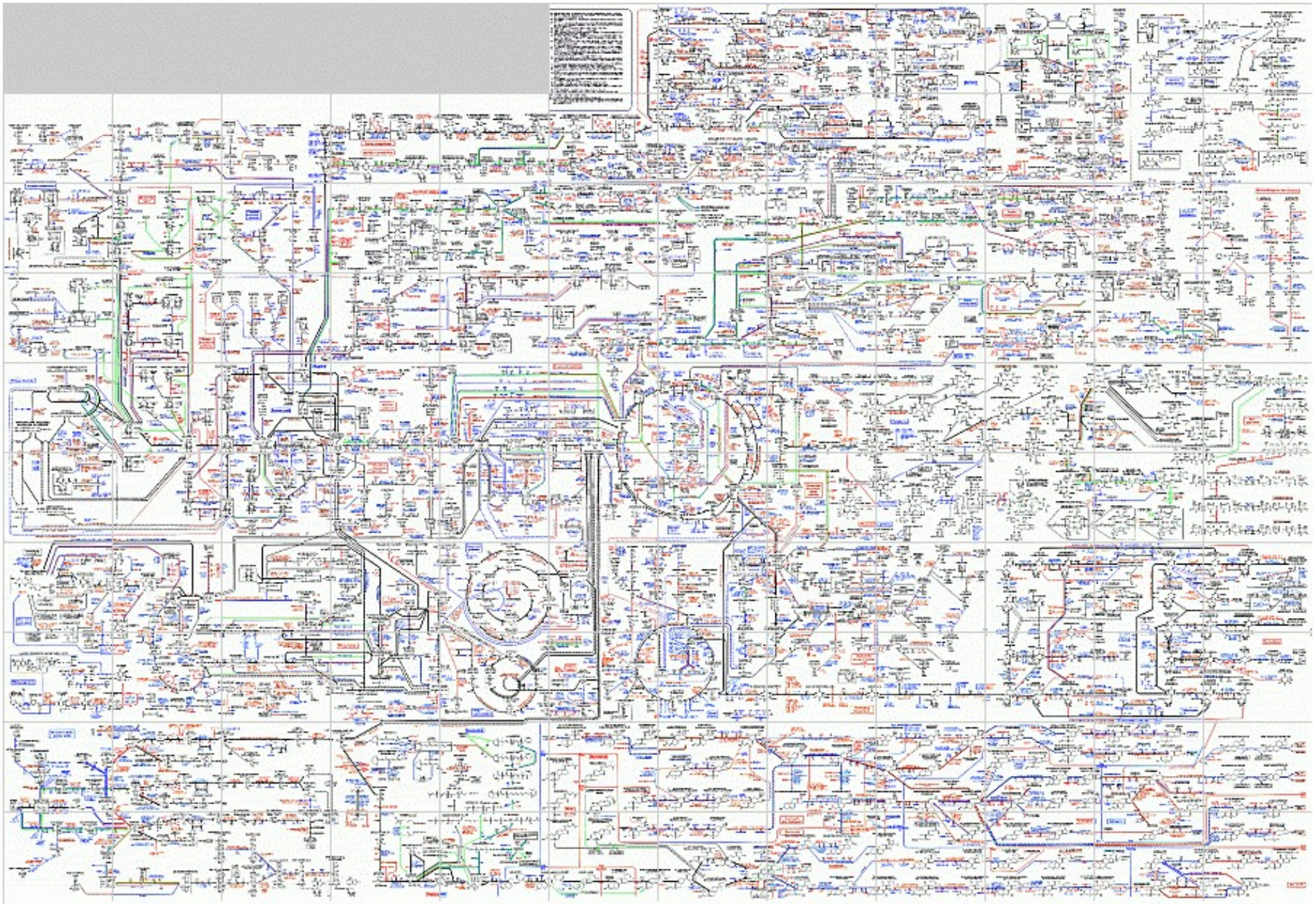
glucose + ATP → glucose 6-phospate + ADP

glucose 6-phosphate ↔ fructose-6-phospate

glucose 6-phosphate + $NADP^+$ → 6-phosphogluconolactone + NADPH

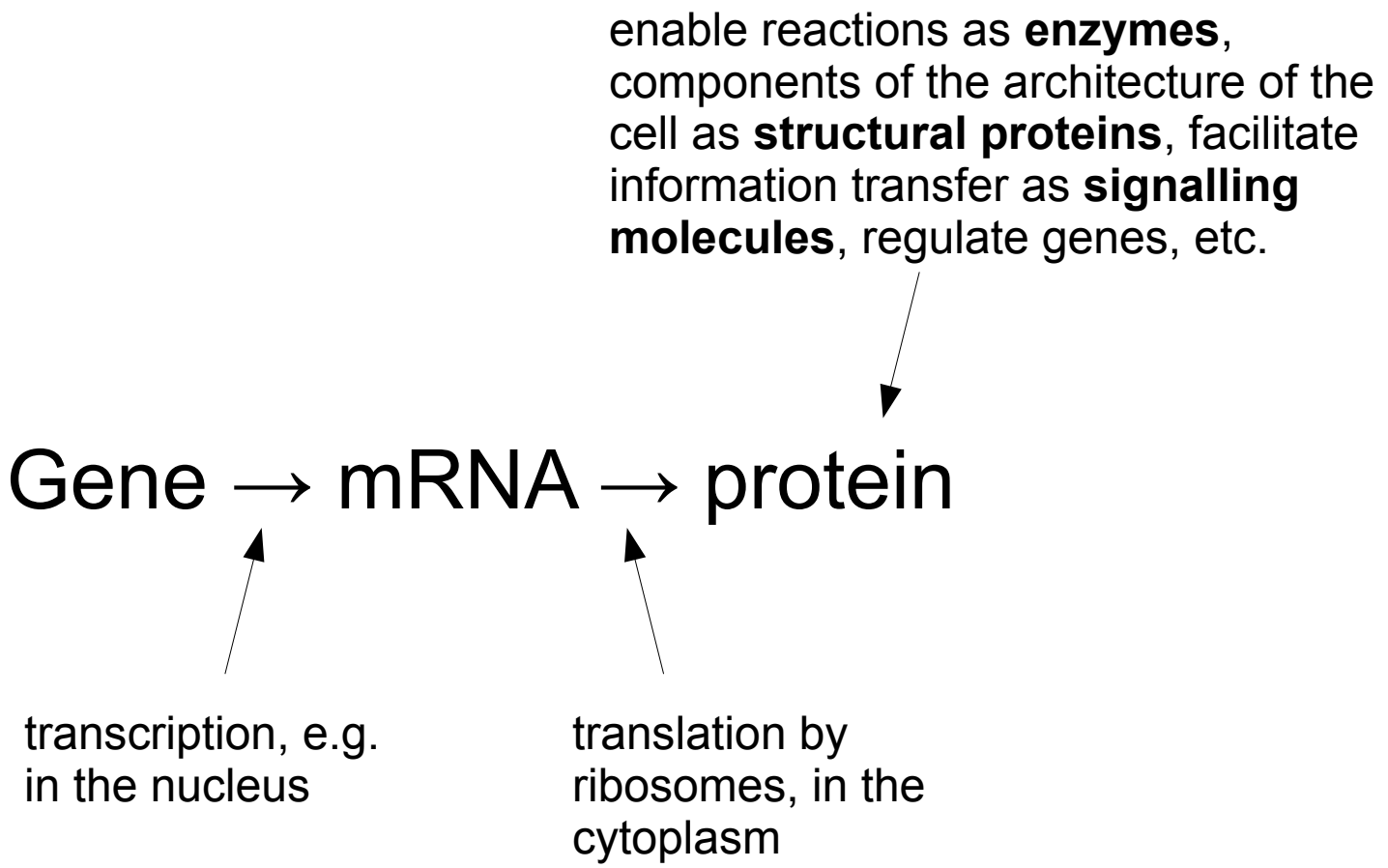


Biochemical maps



taken from http://www.expasy.ch

Central dogma of molecular biology



Biochemical networks

- **Metabolic networks**: transformation of mass of organic molecules
  - catabolic: breaking down of larger molecules to smaller ones. Release of energy
  - anabolic: building larger molecules from smaller components
- **Signalling networks**: transfer of information from signalling molecules (e.g. hormones) to targets in the cell
- **Protein-protein interaction networks**: interactions (e.g. binding) between proteins. Often detailed mechanisms or functions are not known.
- **Gene regulation networks**: high-level conceptual descriptions of interactions between genes

Need for mathematical modelling

- Chemical species are constantly transformed by many different reactions building a complex network
- Therefore, it is difficult to intuitively figure out how the concentrations of the different species behave over time

$\rightarrow$ we need support by mathematical modelling and software tools to do that

Model types (reminder)

- **spatially homogeneous** / spatially explicit
- structural / **quantitative**
- **macroscopic** / mesoscopic / microscopic
- **deterministic** (e.g. ordinary differential equations) / stochastic models

Quantitative models

- Usually derived from qualitative / structural models by specifying equations that describe the interactions of the components ("edges") in a mathematical way, e.g. ordinary differential equation systems
- requires also quantification of initial state of the system, e.g. concentrations of chemical species at time 0

Stoichiometry

- Stoichiometric coefficients indicate the proportions in which substrates are consumed and products are produced in a reaction
- They can be represented as stoichiometric vectors per reaction or a stoichiometric matrix *N* for a whole system (rows: species, columns: reactions, exclude external species with concentrations kept constant)
- Examples...

Kinetic functions

Give the speed of reactions, e.g. rate of change of product concentration, depending on quantities such as substrate concentrations, modifier concentrations, pH, temperature, etc.

"Modifiers" are chemical species that influence the speed of a reaction without being converted in that reaction, e.g. enzymes or regulators of enzymes.

Mass action

Guldberg and Waage (nineteenth century):

Reaction rate is proportional to the probability of a collision of the reactants. This probability is in turn proportional to the concentration of reactants to the power of their molecularity (stoichiometry)

$$v = k \cdot \prod S_i^{n_i}$$

Strictly valid only in one-step gas-phase reactions but, in practise, very often used for other cases as well

Order of a reaction

- The order of a reaction with respect to a certain reactant is defined as the power with which the reactants concentration is used in the rate law (kinetic function)
- Example $2\,A + B \rightarrow C$ with $v = k * [A]^2 * [B]$

  Here the order of the reaction is:

  - 2 with respect to A
  - 1 with respect to B
  - the overall order of the reaction is the sum of the orders of all individual reactants: 3
- The order of a reaction does not necessarily correspond directly to the stoichiometry of the reactants!

Order of a reaction (cont.)

In biochemical systems most often:

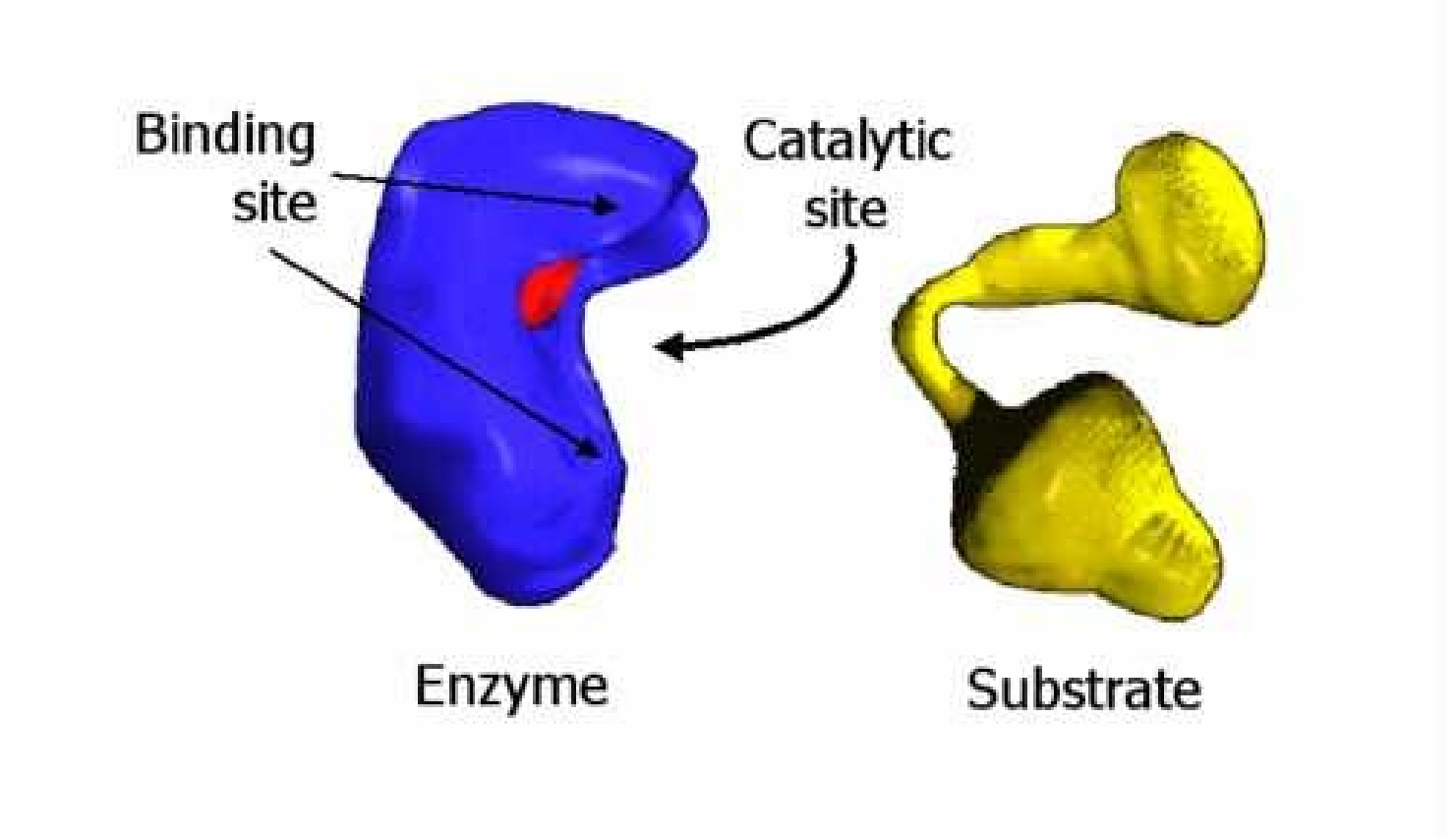
- 0th order, e.g. constant flux
- 1st order, e.g. mass action with one substrate
- 2nd order, e.g. mass action with two substrates

Rate laws of higher orders are rather rare.

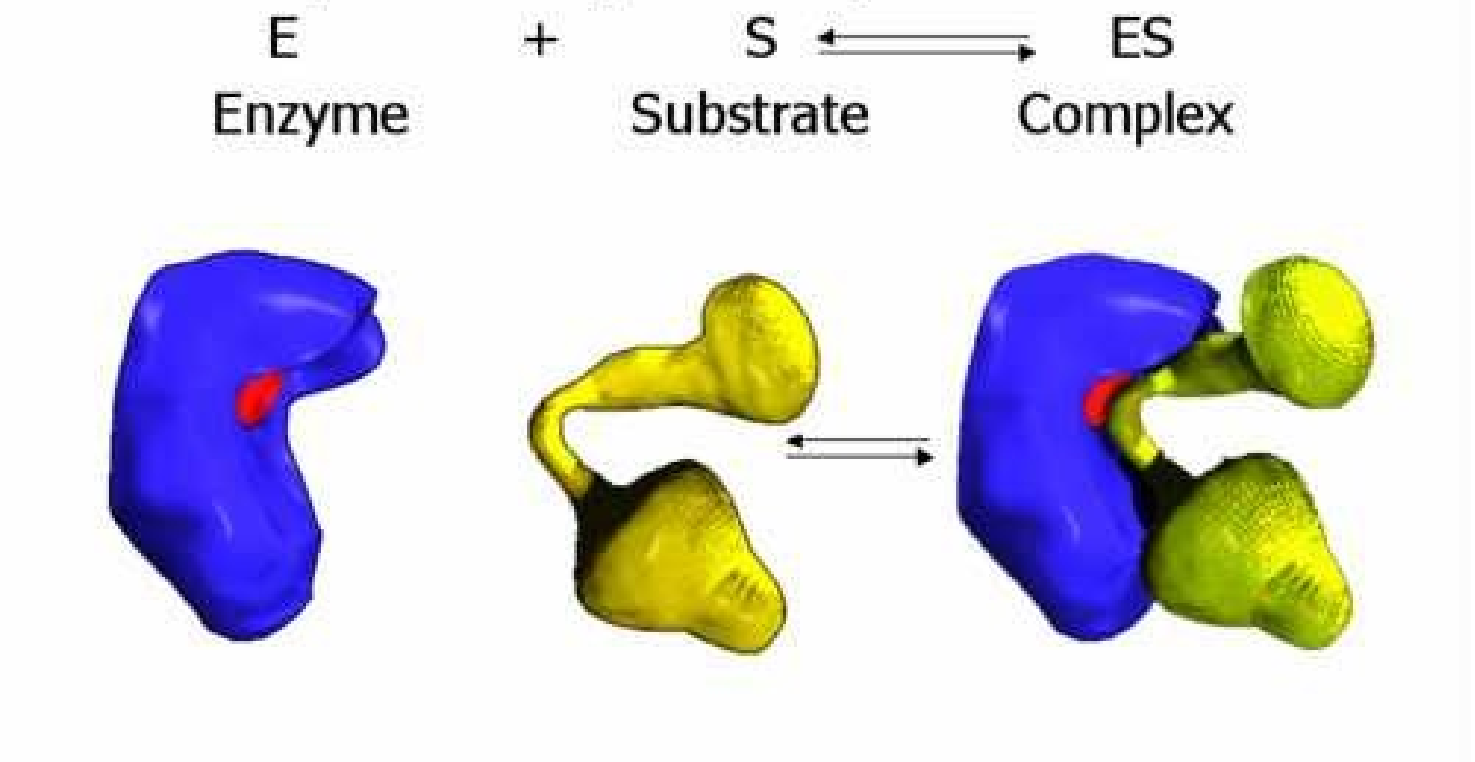
Biochemical reactions

- Biochemical reactions happen in living systems and they usually transform organic substances; (hydro-)carbon based
- Importantly, most often they are catalyzed by **enzymes**. Enzymes reduce the activation energy and, thereby, make reactions possible that would not happen at body temperature with considerable rates
- This also enables the regulation of the speed of reactions by adjusting the concentration of enzymes or modifying the state of enzymes

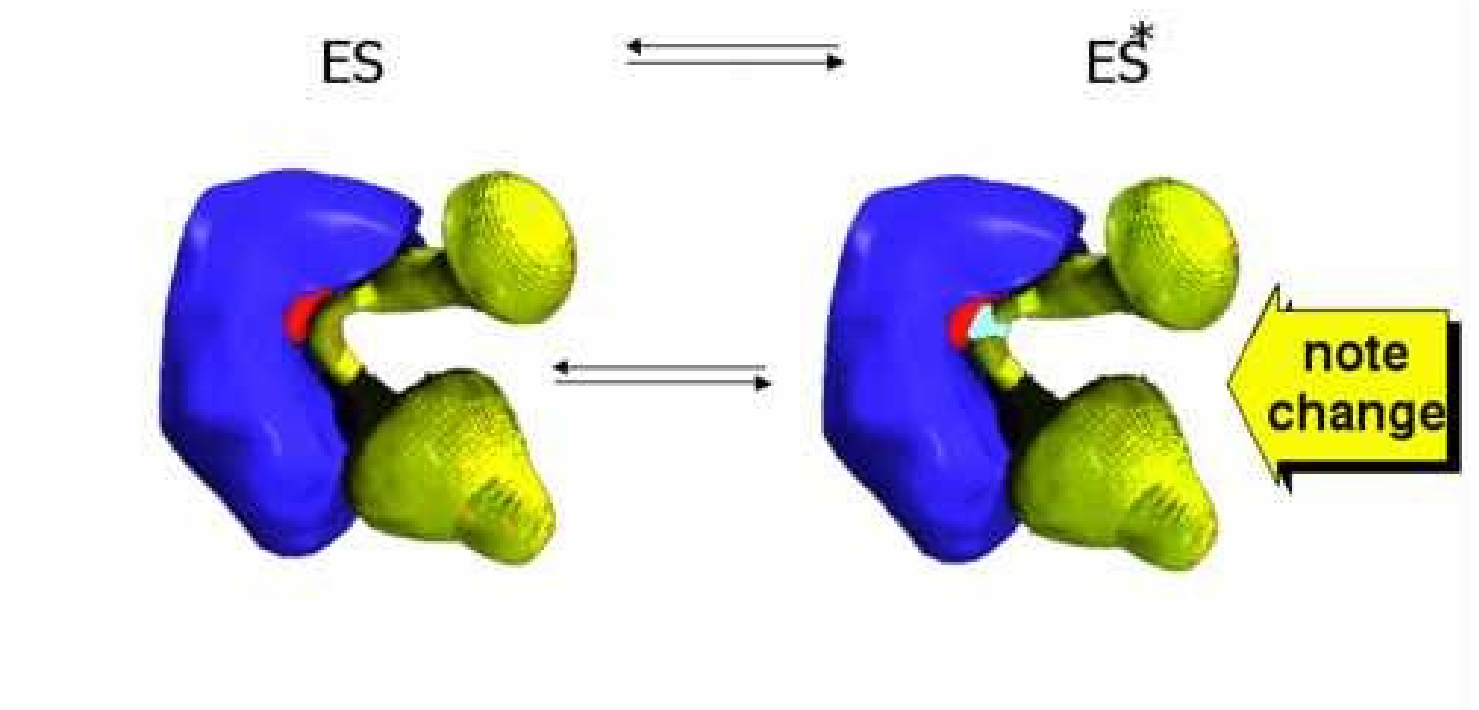
Enzyme kinetics 1/5



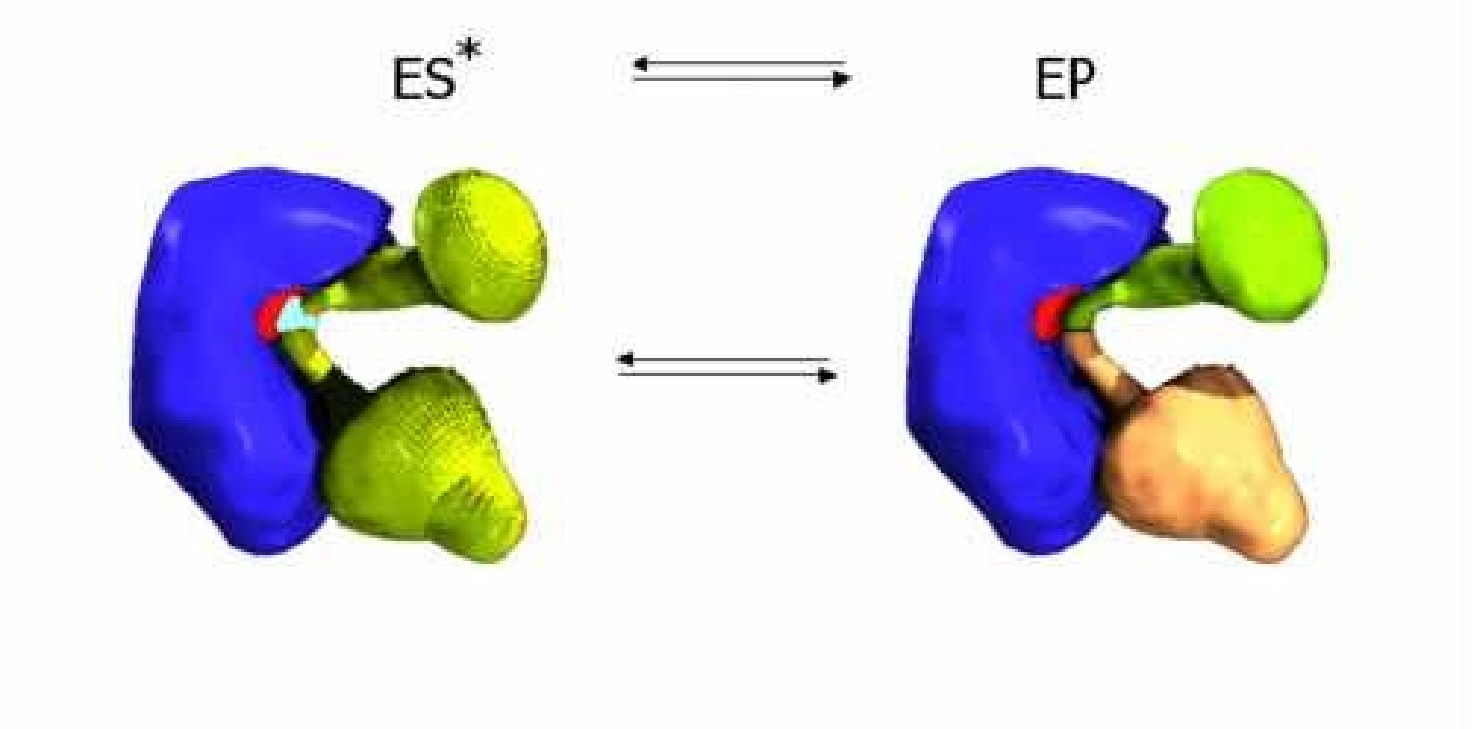
Enzyme kinetics 2/5



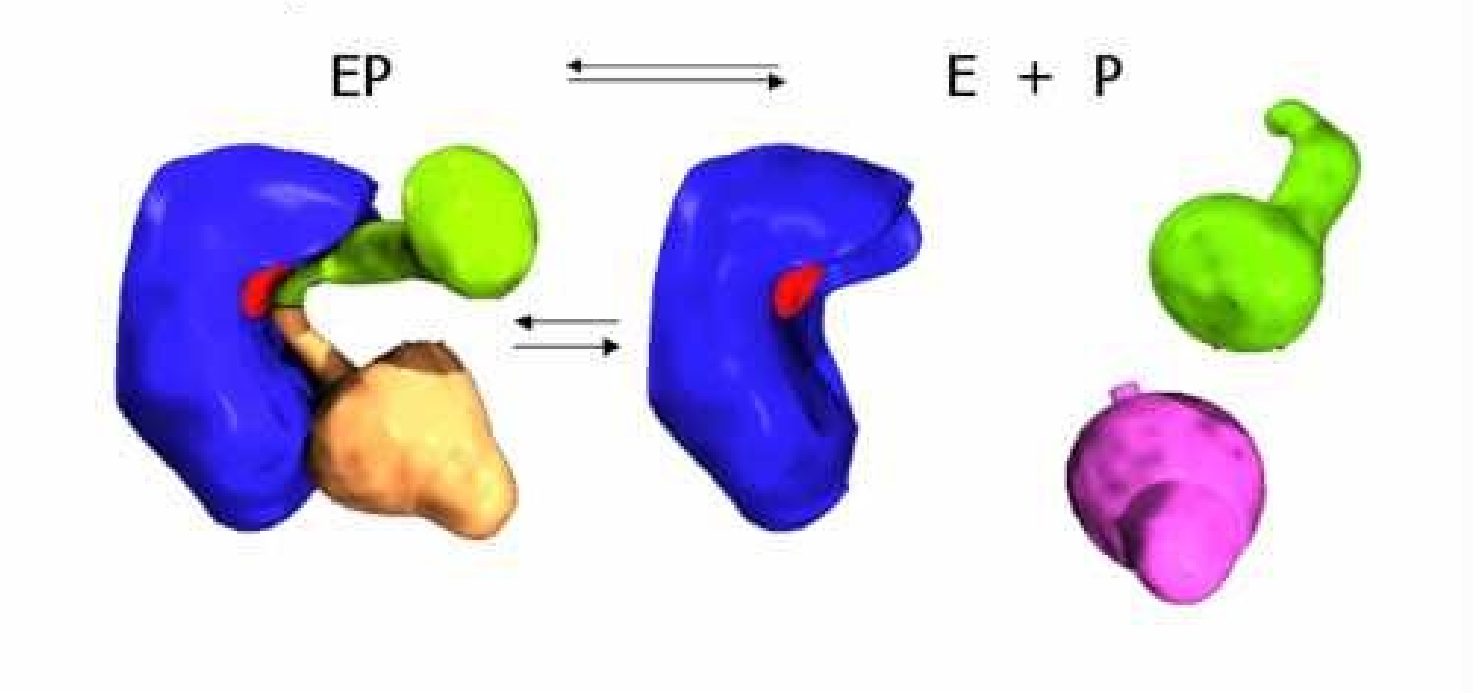
Enzyme kinetics 3/5



Enzyme kinetics 4/5



Enzyme kinetics 5/5

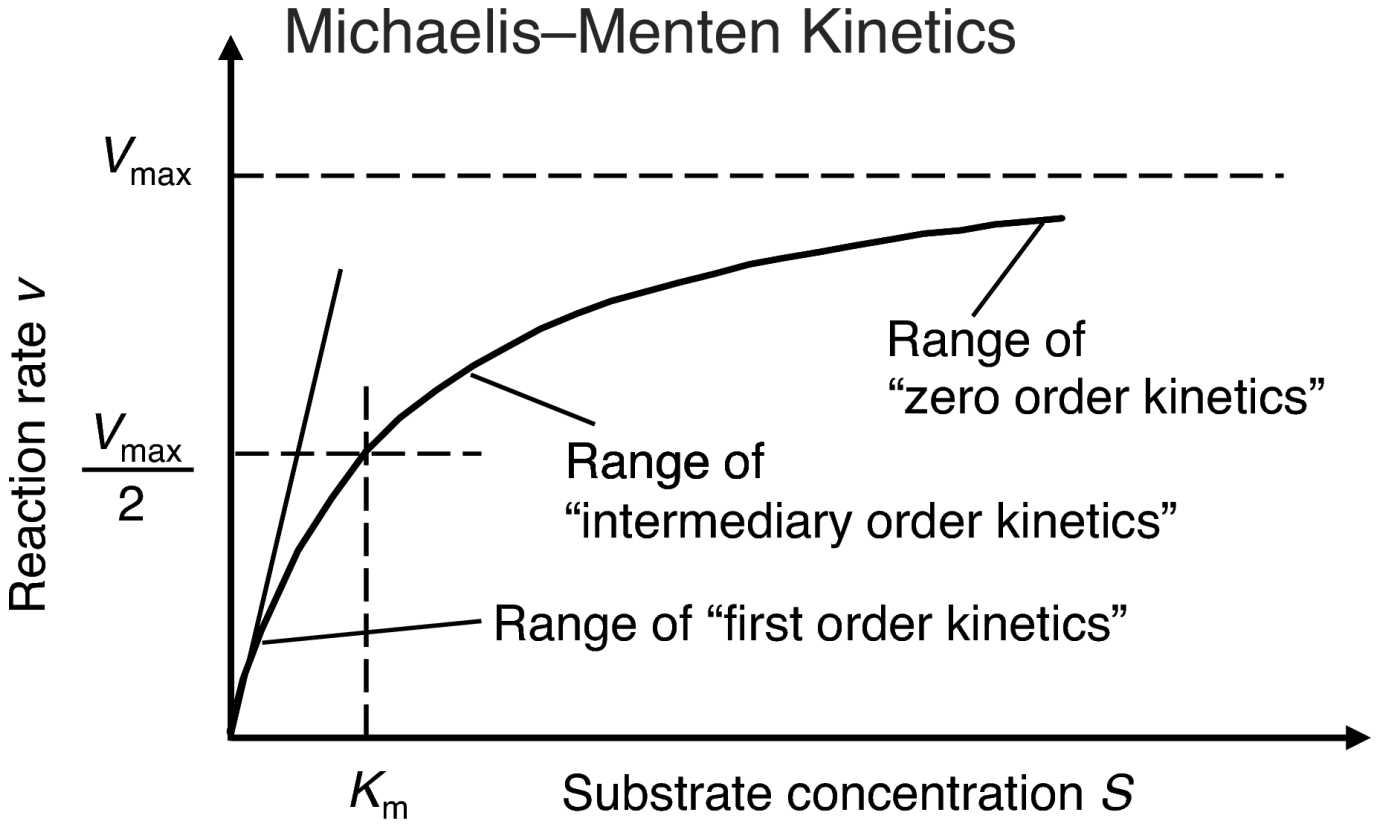


Michaelis-Menten mechanism

$$S + E \Leftrightarrow ES \rightarrow P + E$$

Michaelis-Menten kinetics

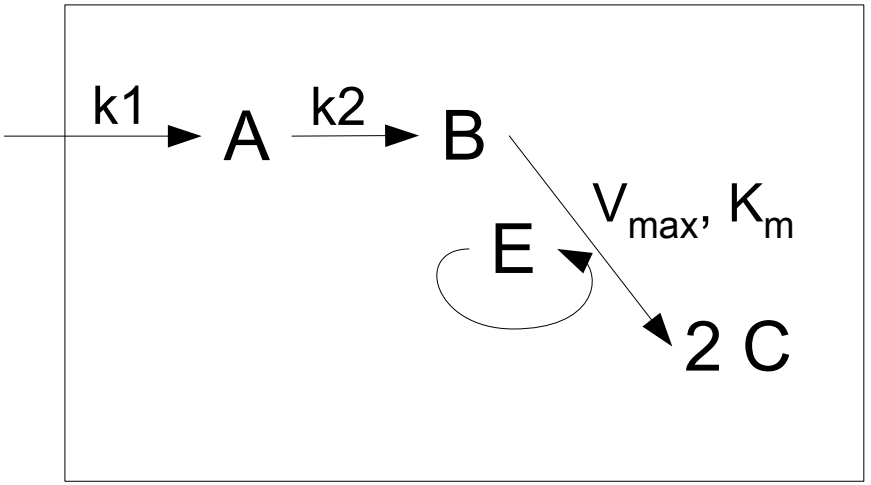
$$v = \frac{V_{max} \cdot [S]}{K_M + [S]}$$



© 2010 Wiley-VCH, Weinheim
Klipp - Systems Biology
ISBN: 978-3-527-31874-2 fig-02-02

Deterministic Approach (ODEs)

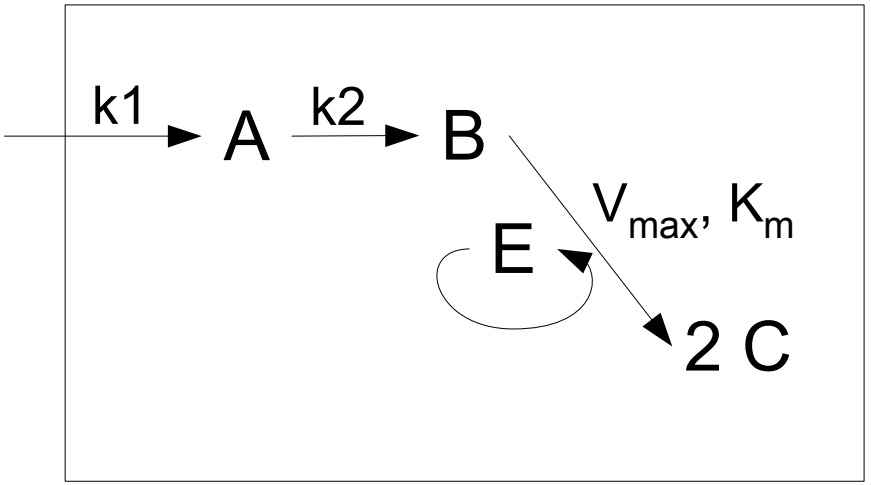
Transformation of a biochemical reaction network into a system of ordinary differential equations (ODEs)



- Constant influx of A from outside
- Simple decay of A to B,
  Mass Action Kinetics
- Reaction of B to 2 C catalyzed by enzyme E,
  Michaelis-Menten Kinetics

Deterministic Approach (ODEs)

Transformation of a biochemical reaction network into a system of ordinary differential equations (ODEs)



- Constant influx of A from outside
- Simple decay of A to B, Mass Action Kinetics
- Reaction of B to 2 C catalyzed by enzyme E, Michaelis-Menten Kinetics

$$d[A]/dt = k1 - k2 * [A]$$

$$d[B]/dt = + k2 * [A] - (V_{max} * [B])/(K_m + [B])$$

$$d[C]/dt = + 2(V_{max} * [B])/(K_m + [B])$$

$$d[E]/dt = 0$$

[E] constant

Formulation using stoichiometry matrix

$$\frac{dS_i}{dt} = \sum_{j=1}^{r} n_{ij} v_j$$

or simply

$$\frac{d\vec{S}}{dt} = N\vec{v}$$

What is needed?

- Structure of the reactions (stoichiometry)
  (what reacts to what? substrates → products)
- Velocity of reactions
  (kinetic functions)
- Initial state $x(t_0)$

Solving the ODE

Simple decay: $S \rightarrow$

$$v = k \cdot [S] \qquad \frac{d[S]}{dt} = -k \cdot [S]$$

Integration from time $t = 0$ with $[S]_0$ leads to exponential behaviour:

$$\int_{[S]_0}^{[S]} dS / S = -\int_{t=0}^{t} k \cdot dt$$

$$[S](t) = [S]_0 \cdot \mathrm{e}^{-kt} \qquad k = \frac{-\ln(1/2)}{t_{1/2}}$$

Deterministic simulation / numerical integration

Example: ODE of a model for calcium oscillations

$$\frac{\mathrm{d}G_\alpha}{\mathrm{d}t} = k_1 + k_2 \cdot G_\alpha - \frac{k_3 \cdot PLC \cdot G_\alpha}{(K_4 + G_\alpha)} - \frac{k_5 \cdot [Ca^{2+}] \cdot G_\alpha}{(K_6 + G_\alpha)} \qquad G_\alpha(t_0) = 0.01\ nmol$$

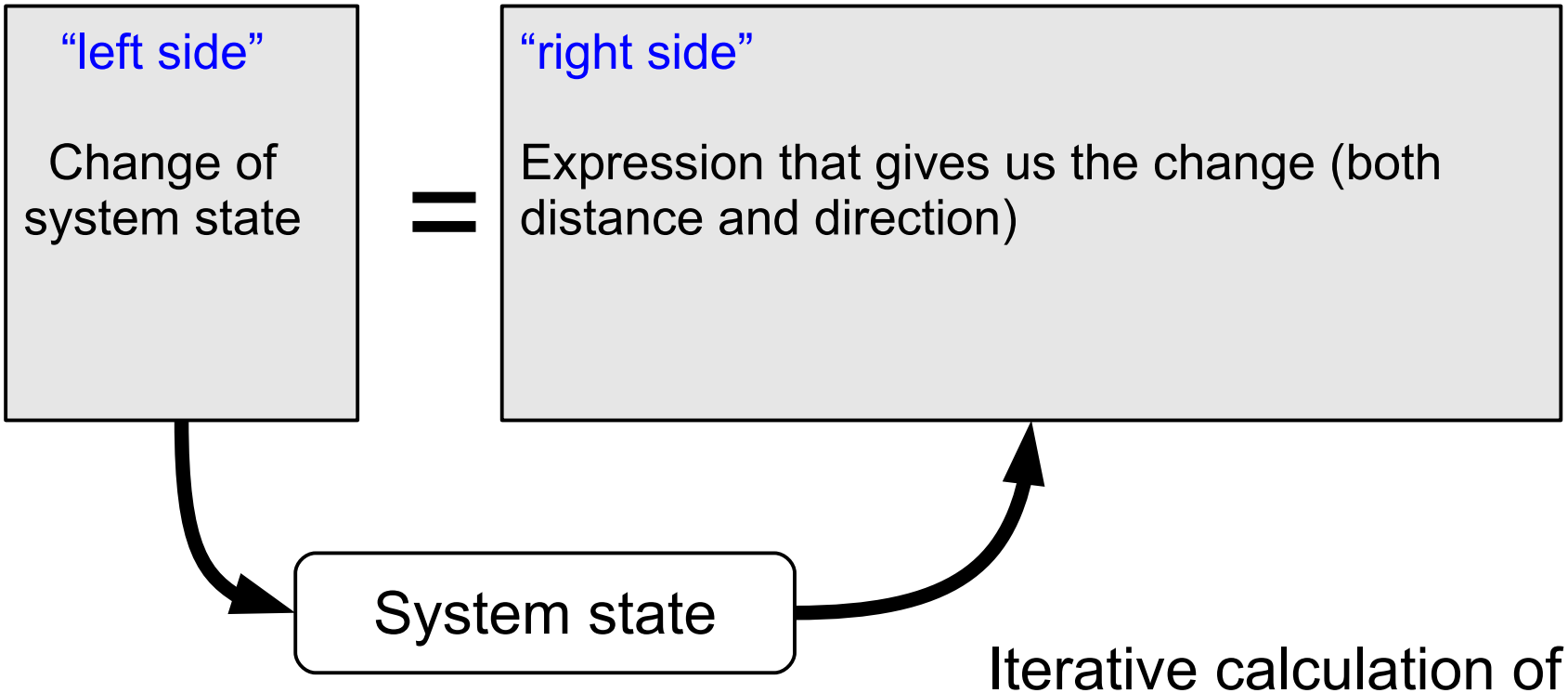
$$\frac{\mathrm{d}PLC}{\mathrm{d}t} = k_7 \cdot G_\alpha - \frac{k_8 \cdot PLC}{(K_9 + PLC)} \qquad PLC(t_0) = 0.01\ nmol$$

$$\frac{\mathrm{d}[Ca^{2+}]}{\mathrm{d}t} = k_{10} \cdot G_\alpha - \frac{k_{11} \cdot [Ca^{2+}]}{(K_{12} + [Ca^{2+}])} \qquad [Ca^{2+}](t_0) = 0.01\ nmol$$

Biochemical ODEs usually can not be solved analytically → numerical integration

Numerical integration

- Differential equation:

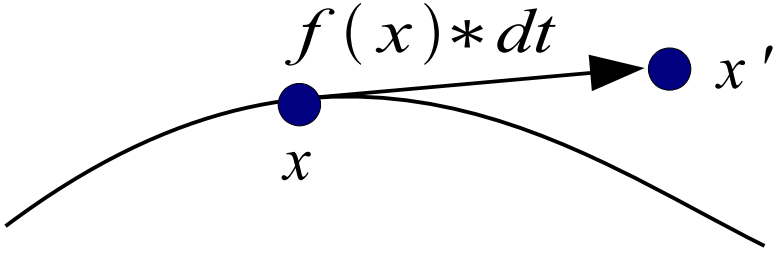


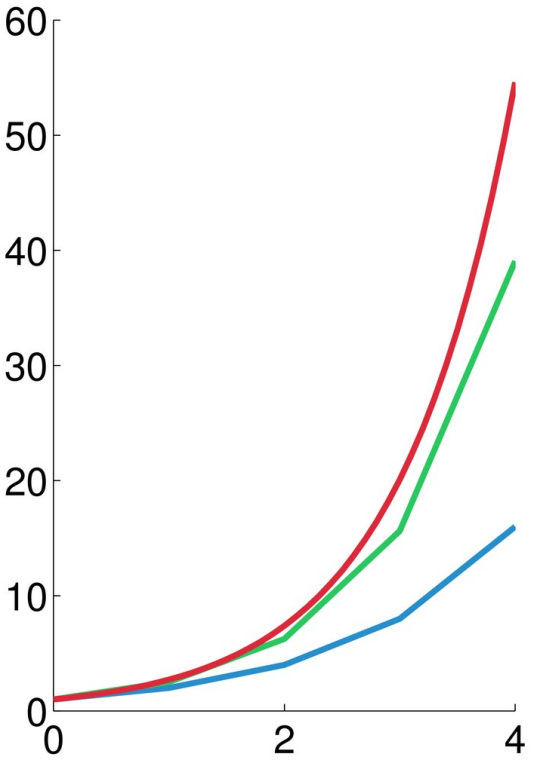
Deterministic Simulation

Euler method (truncated Taylor series expansion)

$$\frac{dx}{dt} = f(x) \Leftrightarrow x' - x = f(x) * dt \Leftrightarrow x' = x + f(x) * dt$$

Explicit first-order method (error is proportional to step size *dt*). Higher order methods include Runge Kutta schemes and others.
The Euler method is NOT recommended for most biochemical systems because of "stiffness"!





Simulation / Numerical integration of ODE

Biochemical systems usually involve processes on different time scales → stiff systems

Step size of, e.g. the first order Euler method, cannot be decreased arbitrarily because of computational demands and resulting numerical problems

Special "stiff solvers" are needed, e.g.

LSODA (Hindmarsh & Petzold), part of ODEPACK

- Non-stiff: Adams multistep method with variable order (1-12)
- Stiff: BDF (backward differentiation formula) with variable order (1-5)

References

- A.C. Hindmarsh (1983) ODEPACK, A Systematized Collection of ODE Solvers. In: R.S. Stepleman and M. Carver and R. Peskin and W.F. Ames and R. Vichnevetsky (eds.) *Scientific Computing, North-Holland, Amsterdam (vol. 1 of IMACS Transactions on Scientific Computation),* pp. 55-64
- L. Petzold (1983) Automatic Selection of Methods for Solving Stiff and Nonstiff Systems of Ordinary Differential Equations. *SIAM J. Sci. Stat. Comput.*, **4**:136-148

Software for numerical integration

- Mathematics packages
  - Maple, Mathematica, Matlab, GNU Octave etc.
- General purpose ODE solvers
  - Berkeley Madonna (commercial but free demo version)
- Specialised systems biology software
  - DBSolve
  - COPASI
  - RoadRunner
  - ...

→ They all provide different interfaces and algorithms

Parameters and variables

- **Parameters** are items that are independent of the system, *i.e.* are set by outside agents (causes)
- **Variables** are items of the system whose values are determined exclusively by the parameters (effects)

- What is a variable and what is a parameter is a modelling question! E.g. the concentration of an enzyme can be a parameter or variable in a specific model.

Important interactions in biochemical models

- **chemical reactions** (particularly in metabolic models)
- **transport reactions** (compartmental modelling)
- **physico-chemical interactions** (e.g. regulatory effects, e.g. complex formation in translation regulation)

Compartment modelling

If transport reactions across compartment boundaries are modelled in a biochemical ODE for species concentrations, different volumes have to be considered by including volume correction factors (concentration is not conserved, but mass is).

Or, reformulate the ODE in terms of particle numbers (see COPASI, example Hendriks 2008 p38 MAPK model)

http://www.copasi.org

Mendes group

Kummer group

- COPASI (COmplex PAthway SImulator)
- Software for the simulation and analysis of biochemical networks
- “Tool kit” with a variety of different methods:
  - Deterministic, stochastic and hybrid simulation methods
  - Metabolic Control Analysis, Elementary Flux Mode Analysis, Sensitivity Analysis
  - Parameter Scanning, Optimization, Parameter Fitting
  - User-friendly GUI, runs under Mac, Linux, Windows and Solaris and command line version
  - Artistic license/open-source
  - reads and writes SBML, etc.