Dynamics of Biological Systems

Part I - Biological background and mathematical modelling

Introduction

The recent developments in biology have produced a huge amount of data about the structure of living matter;

- consider as an example the success of the Human Genome Project

Less is known about the versatile functions that cells and their components show.

In the last few years the scientific interest has started to move from structures to functionalities

The complexity of the cellular processes has stimulated the growth of a new paradigm, that moves from the classical reductionist approach to a system level understending of life

- Such a paradigm is called systems biology

Introduction

A better understanding of the funcitoning of cellular processes may allow:

- a better undertanding of diseases
- the development of more effective drugs
- the development of preventive and early diagnosis techniques

Mathematical and computational modelling may contribute to the study of cellular processes with simulation tools that, based on data from laboratory experiments, could be used to:

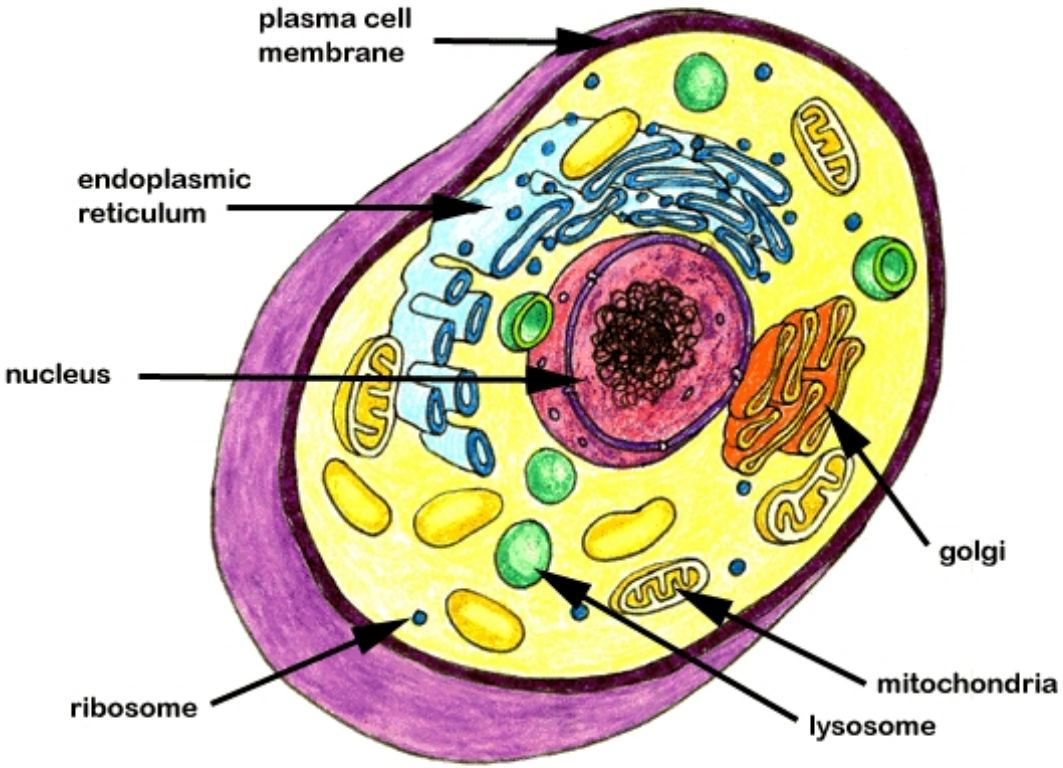
- validate hypotheses
- suggest further experiments
- predict the effect of some treatments

In the future, treatment of diseases will be based on patient-specific therapies

- simulation tools capable to predict the effect of some therapy on a specific patient will be essential

Outline of the talk

Cells: complex systems of interactive components



- Two classifications of cell:
  - prokaryotic
  - eukaryotic
- Main actors:
  - membranes
  - proteins
  - DNA/RNA
  - ions, macromolecules,. . .
- Interaction networks:
  - metabolic pathways
  - signaling pathways
  - gene regulatory networks

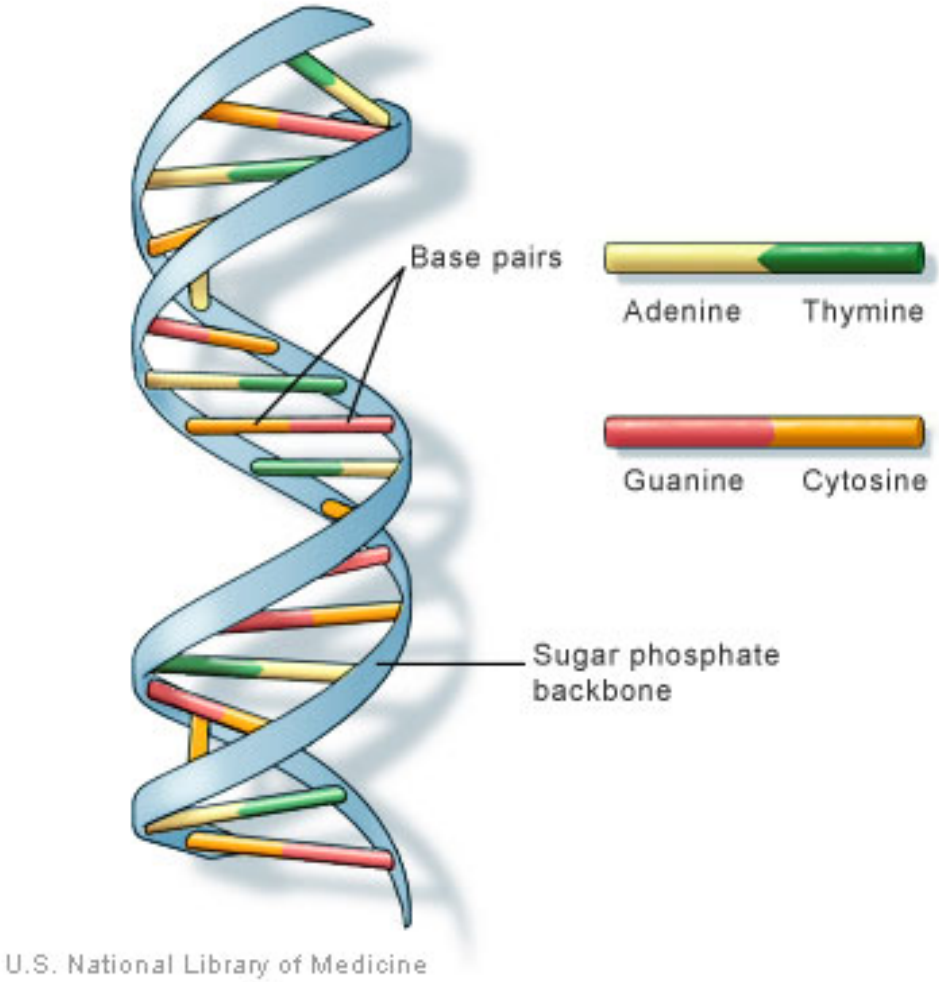
The DNA

The DNA is:

- a molecule
- structured as a string
- over an alphabet of four elements (nucleic acids, bases) denoted A,T,C,G

DNA forms double-stranded helices:

- Base pairing: A–T,C–G
- The *c*omplement of a string is obtained by replacing A with T and C with G, and viceversa
- Two complementary strings form a helic



Proteins

A gene is a substring of the DNA

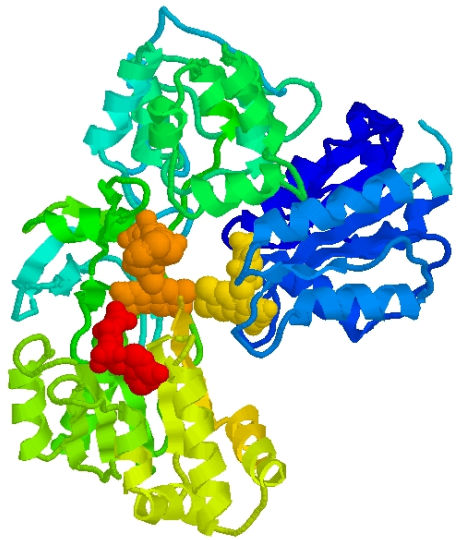
- some genes are the "source code" of proteins

A protein is:

- a molecule
- structured as a string
- over an alphabet of twenty elements (amino acids)

Proteins have complex 3D structures related with their functions:

- Catalysis of chemical reactions (enzymes)
- Transport
- Structure
- .....



The central dogma of Molecular Biology

Schematically, in cells we have this flux of information:

$$DNA \xrightarrow{transcription} RNA \xrightarrow{translation} Protein$$

Where the RNA is a molecule structured as a string over the alphabet A,U,C,G (similar to that of DNA)

- It is essentially a copy of the DNA (this motivates the terminology of transcription)

Both transcription and translation can be regulated in order to synthesize proteins only when necessary

On the importance of membranes

Eukariotic cells have a large number of compartments separated by membranes

- this allows a more efficient cell management (based on "divide et impera")

Membranes have (at least) two fundamental functions:

- as separators of compartments
- as channels of communication between compartments

The main classes of cellular processes

- Gene regulation networks: are collection of genes in a cell which interact with each other (indirectly through their RNA and protein expression products) and with other substances in the cell, thereby governing the rates at which genes in the network are transcribed into RNA.
- Metabolic pathways: are series of chemical reactions occurring within a cell aimed at transforming some initial molecule into another product.
- Cell signalling pathways: (or signal transduction pathways) are series of chemical reactions occurring within a cell aimed at converting some chemical stimulus into a specific cellular response. Signal transduction starts with a signal to a receptor, and ends with a change in cell behavior.

Outline of the talk

Example of gene regulation network: the *lac* operon

E. coli is a bacterium often present in the intestine of many animals. It is one of the most completely studied of all living things.

As most bacteria, E.coli is often exposed to a constantly changing physical and chemical environment, and reacts to changes in its environment through changes in the kinds of enzymes it produces.

In order to save energy, bacteria do not synthesize degradative enzymes unless the substrates (e.g. lactose) for these enzymes are present in the environment.

This result is obtained by controlling the transcription of some genes into the corresponding enzymes.

Example of gene regulation network: the *lac* operon

Two enzymes are mainly involved in lactose degradation:

- the *lactose permease*, which is incorporated in the membrane of the bacterium and actively transports the sugar into the cell,
- and the *beta galactosidase*, which splits lactose into glucose and galactose.

The bacterium produces also the *transacetylase* enzyme, whose role in the lactose degradation is marginal.

The sequence of genes in the DNA of E. coli which produces the described enzymes, is known as the *lactose operon*.

The lactose operon consists of six genes:

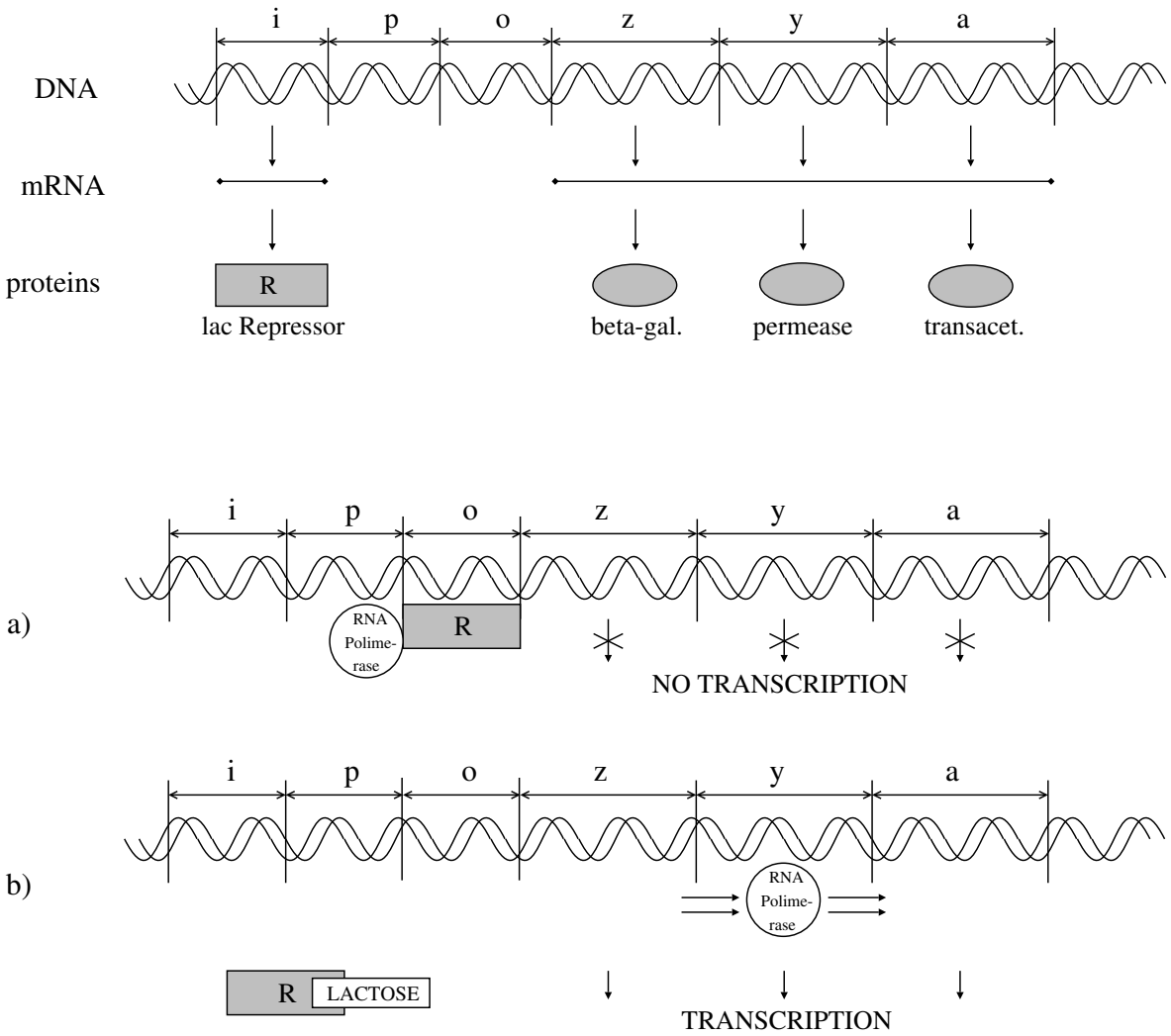
- The first three genes of the operon (i, p and o) regulate the production of the enzymes,
- the last three (z, y and a), called *structural genes*, are transcribed (when allowed) into the mRNA for beta galactosidase, lactose permease and transacetylase, respectively.

Example of gene regulation network: the *lac* operon

The regulation process is as follows:

- Gene i encodes the *lac Repressor*, which, in the absence of lactose, binds to gene o (the *operator*).
- Transcription of structural genes into mRNA is performed by the RNA polymerase enzyme, which usually binds to gene p (the *promoter*) and scans the operon from left to right by transcribing the three structural genes z, y and a into a single mRNA fragment.
- When the lac Repressor is bound to gene o, it becomes an obstacle for the RNA polymerase, and the transcription of the structural genes is not performed.
- On the other hand, when lactose is present inside the bacterium, it binds to the Repressor and this cannot stop anymore the activity of the RNA polymerase. In this case the transcription is performed and the three enzymes for lactose degradation are synthesized.

Example of gene regulation network: the *lac* operon



Example of metabolic pathway: the glycolysis pathway

Glycolysis is the metabolic pathway that converts glucose, C6H12O6, into pyruvate, CH3COCOO- and H+.

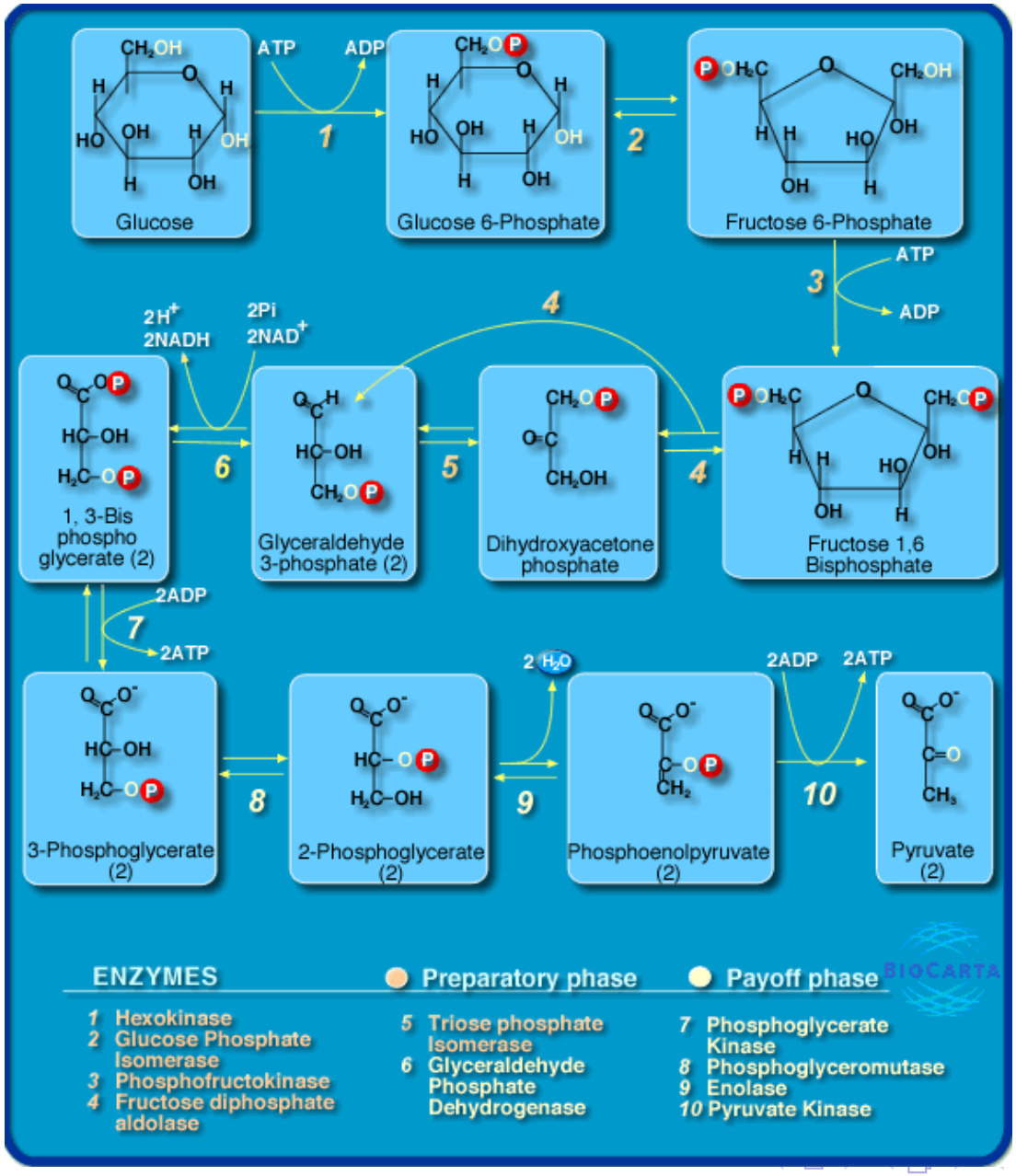
The free energy released in this process is used to form the high energy compounds, ATP (adenosine triphosphate) and NADH (reduced nicotinamide adenine dinucleotide).

Glycolysis is a definite sequence of ten reactions involving intermediate compounds.

The intermediates provide entry points to glycolysis. For example, most monosaccharides, such as fructose, glucose, and galactose, can be converted to one of these intermediates.

The intermediates may also be directly useful. For example, the intermediate dihydroxyacetone phosphate is a source of the glycerol that combines with fatty acids to form fat.

Example of metabolic pathway: the glycolysis pathway



Example of signalling pathway: the EGF pathway

A classical example of biological system is the EGF signal transduction pathway.

If EGF proteins are present in the environment of a cell, they must be interpreted as a signal from the environment meaning that new cells are needed.

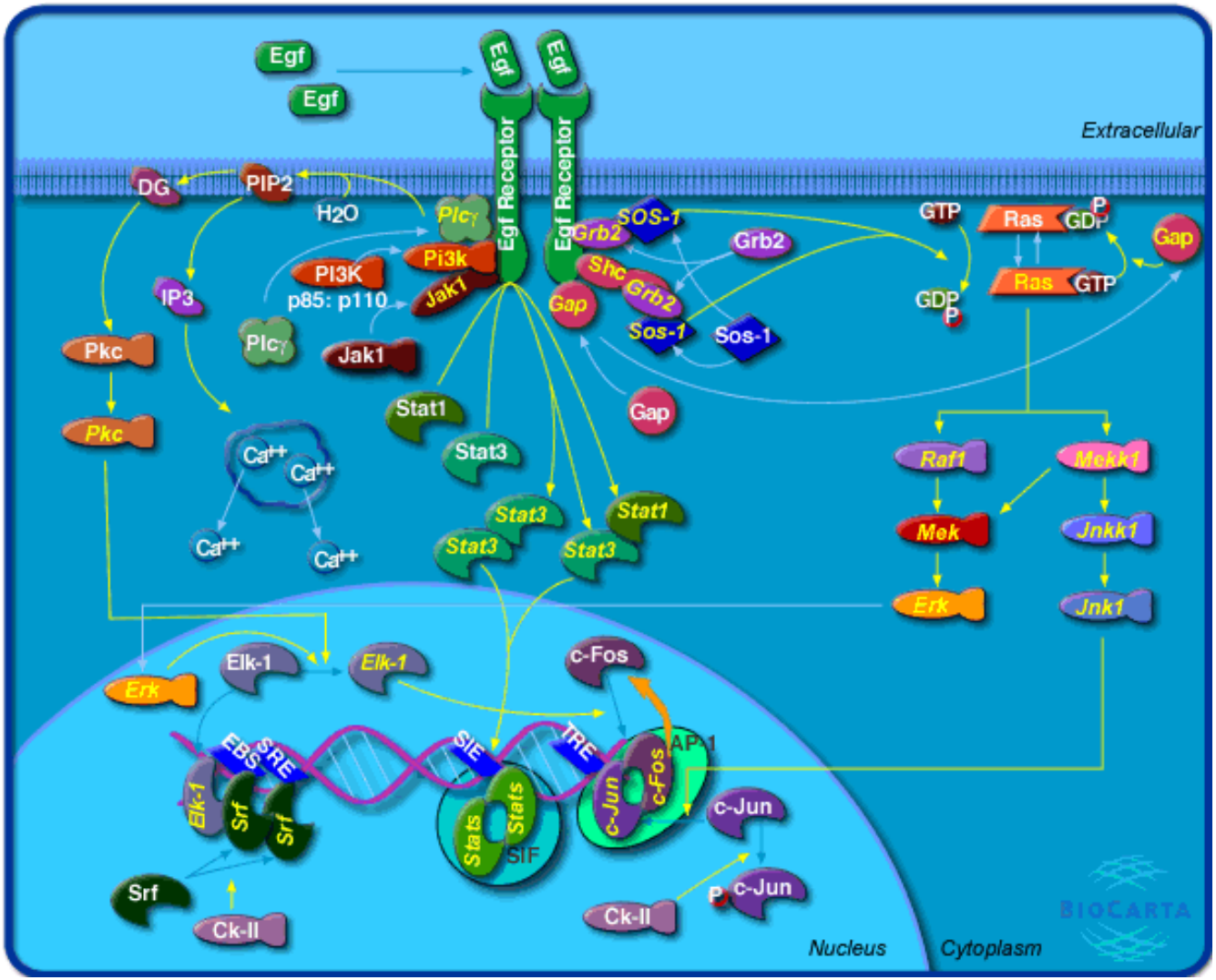
A cell recognizes the EGF signal from the environment because it has on its membrane some EGF receptor proteins (EGFR), which are transmembrane proteins (they have some intra–cellular and some extra–cellular domains).

Example of signalling pathway: the EGF pathway

The signalling pathway is as follows:

- One of the extra–cellular domains binds to one EGF protein in the environment, forming a signal–receptor complex on the membrane.
- This causes a conformational change on the receptor protein that enables it to bind to another one signal–receptor complex.
- The formation of the binding of the two signal–receptor complexes (called dimerization) causes the phosphorylation (addition of some phosphate groups PO_4) of some intra–cellular domains of the dimer.
- This causes the internal domains of the dimer to be recognized by proteins that are in the cytoplasm, which bind to the dimer, enabling a chain of protein–protein interactions inside the cell.
- This chain of interactions finally activate some proteins which bind to the DNA and stimulate synthesis of proteins for cell proliferation.

Example of signalling pathway: the EGF pathway



The cell cycle

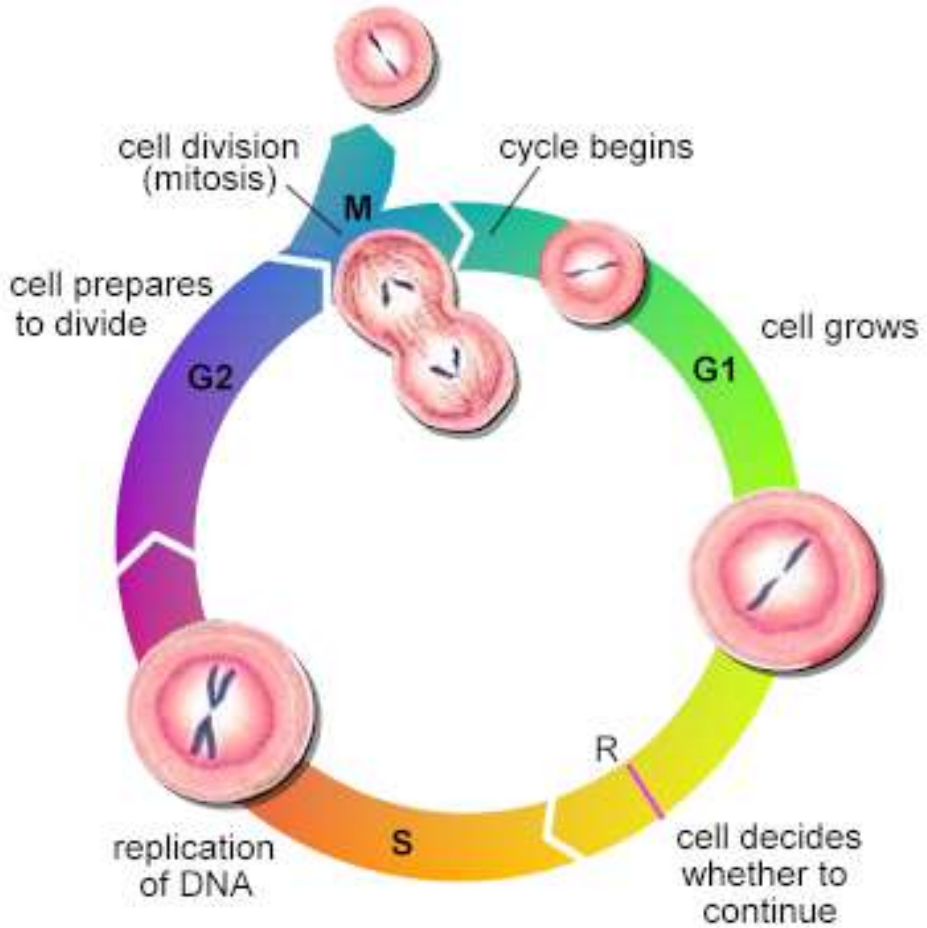
The cell cycle is a series of sequential events leading to cell replication via cell division.

It consists of four phases: G_1, S, G_2 and M.

- G_1 and G_2 are gap phases in which the cell prepares itself to enter phases S and M, respectively
- S is a synthesis phase, in which DNA is replicated
- M is a mitosis phase, in which the cell segregates the duplicated sets of chromosomes between daughter cells and then divides.

The duration of the cell cycle depends on the type of cell (e.g a human normal cell takes approximately 24 hours to perform a cycle).

The cell cycle



Outline of the talk

Background

- The modelling of the dynamics of biological systems is essentially based on the modelling of the dynamics (kinetics) of some (bio)chemical reactions
- The modelling of chemical reactions using deterministic rate laws has proven extremely successful in both chemistry and biochemistry for many years
- This deterministic approach has its core the law of mass action, an emprirical law giving a simple relation between reaction rates and molecular component concentrations
- Given knowledge of inital concentrations, the law of mass action provides a complete picture of the component concentrations at all future time points

Keywords

Some definitions (mostly from Wikipedia...)

- chemical solution: a homogeneous mixture composed of two or more substances (molecules)
- biochemical solution: chemical solution in which substances have an organic nature (proteins, DNA, etc...)
- chemical reactions: a process that leads to the transformation of one set of chemical substances to another
- biochemical reactions: chemical reactions in which substances have an organic nature (proteins, DNA, etc...)
- kinetics: the study of (bio)chemical reaction rates (which govern the dynamics of (bio)chemical solutions)
- rate: frequency of an event (e.g. a reaction) over time
- concentration: average quantity of molecules of a solution in a volume unit
- deterministic model: model which exhibit a single (usually average) behaviour (dynamics)
- stochastic model: model which takes probabilistic/stochastic behaviour (dynamics) into account

The mass action kinetics of chemical reactions

Usual notation for chemical reactions:

$$\ell_1 S_1 + \ldots + \ell_\rho S_\rho \underset{k_{-1}}{\overset{k}{\rightleftharpoons}} \ell'_1 P_1 + \ldots + \ell'_\gamma P_\gamma$$

where:

- S_i, P_i are molecules (reactants)
- ℓ_i, ℓ'_i are stoichiometric coefficients
- k, k_{-1} are the kinetic constants

Examples of chemical reactions:

$$2H_2 + O_2 \underset{k_{-1}}{\overset{k}{\rightleftharpoons}} 2H_2O$$

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E + P$$

The mass action kinetics of chemical reactions

Usual notation for chemical reactions:

$$\ell_1 S_1 + \ldots + \ell_\rho S_\rho \underset{k_{-1}}{\overset{k}{\rightleftharpoons}} \ell'_1 P_1 + \ldots + \ell'_\gamma P_\gamma$$

The kinetics is described by the *law of mass action*:

> *The rate of a reaction is proportional to the products of the concentrations of the participating molecules.*

The rate of a reaction (the left-to-right direction of the reversible reaction above) is defined as follows:

$$k[S_1]^{\ell_1} \cdots [S_\rho]^{\ell_\rho}$$

similarly, the rate of the opposite reaction (the right-to-left direction) is

$$k_{-1}[P_1]^{\ell'_1} \cdots [P_\gamma]^{\ell'_\gamma}$$

The mass action kinetics of chemical reactions

For example, the rates of

$$2H_2 + O_2 \underset{0.5}{\overset{5}{\rightleftharpoons}} 2H_2O$$

are

- $5[H_2]^2[O_2]$ for the left-to-right direction
- $0.5[H_2O]^2$ for the right-to-left direction

Similarly, the rates of

$$E + S \underset{1000}{\overset{0.1}{\rightleftharpoons}} ES \underset{0.01}{\overset{0.3}{\rightleftharpoons}} E + P$$

are

- $0.1[E][S]$ for the formation of the complex ES from E and S
- $1000[ES]$ for the unbinding of ES into E and S
- $0.3[ES]$ for the unbinding of ES into E and P
- $0.01[E][P]$ for the formation of the complex ES from E and P

The mass action kinetics of chemical reactions

Usual notation for chemical reactions:

$$\ell_1 S_1 + \ldots + \ell_\rho S_\rho \underset{k_{-1}}{\overset{k}{\rightleftharpoons}} \ell'_1 P_1 + \ldots + \ell'_\gamma P_\gamma$$

A reversible reaction is said to be in dynamic equilibrium when the rates of its two directions are the same, namely:

$$k[S_1]^{\ell_1} \cdots [S_\rho]^{\ell_\rho} = k_{-1}[P_1]^{\ell'_1} \cdots [P_\gamma]^{\ell'_\gamma}$$

When the dynamic equilibrium is reached, often the chemical solution seems stable (it seems that nothing is happening inside it). Instead, reactions happen continuously, compensating each other.

It is easy to see that at the equilibrium we have:

$$\frac{k}{k_{-1}} = \frac{[P_1]^{\ell'_1} \cdots [P_\gamma]^{\ell'_\gamma}}{[S_1]^{\ell_1} \cdots [S_\rho]^{\ell_\rho}}$$

Ordinary differential equations (ODEs)

Reaction rates can be used to construct a mathematical model (based on ordinary differential equations - ODEs) of the dynamics of a set of chemical reactions

Ordinary differential equations (ODEs) are mathematical equations for un unkonwn function of one or of several variables that relates the values of the function itself and its derivatives.

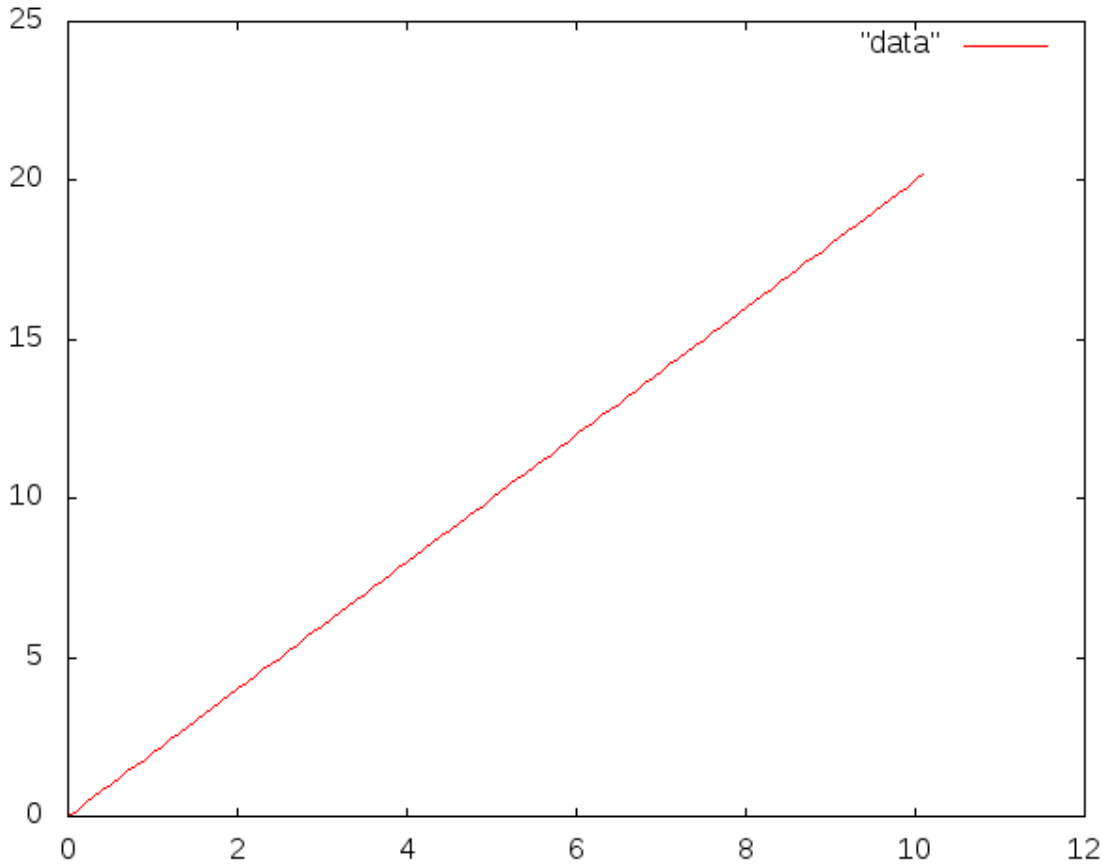
In other words, an ODE describes (by means of derivatives) how a variable of an unknown function changes over time

The unknown function could be obtained by integrating the corresponding ODEs (often not feasible)

Ordinary differential equations (ODEs)

Some examples:

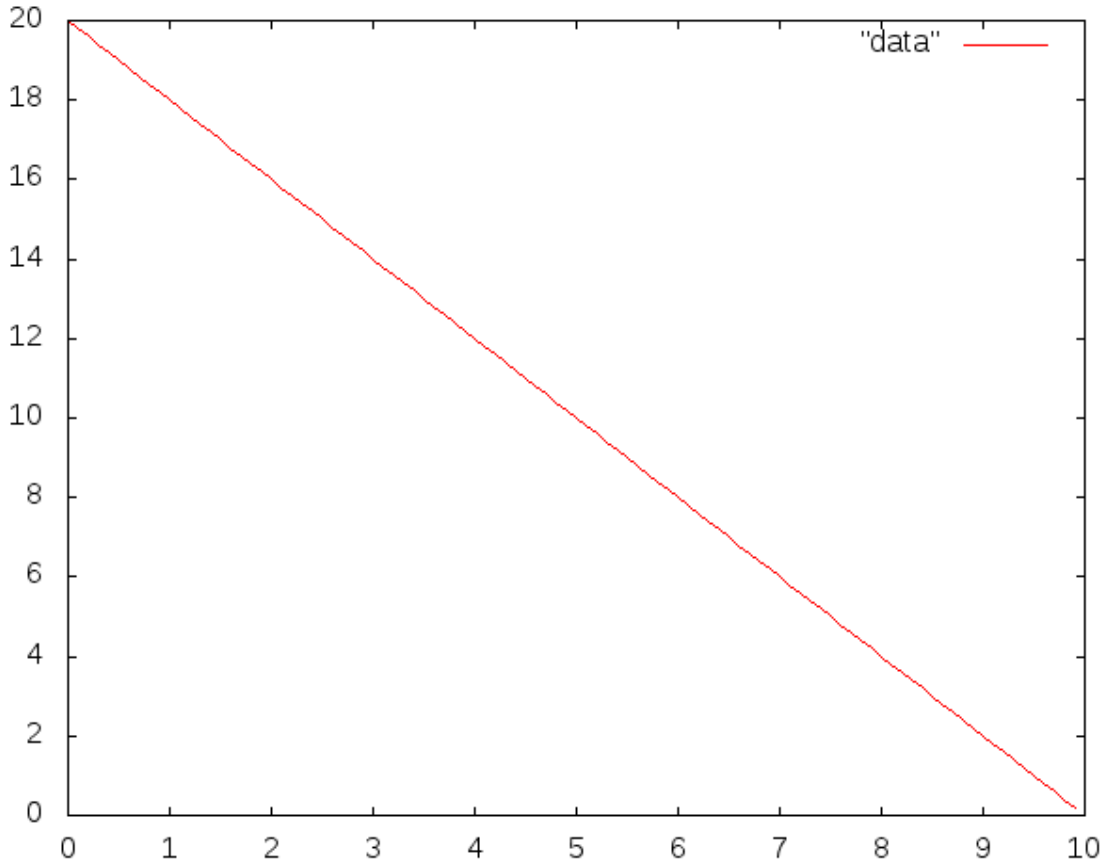
$$\frac{dx}{dt} = 2$$



Ordinary differential equations (ODEs)

Some examples:

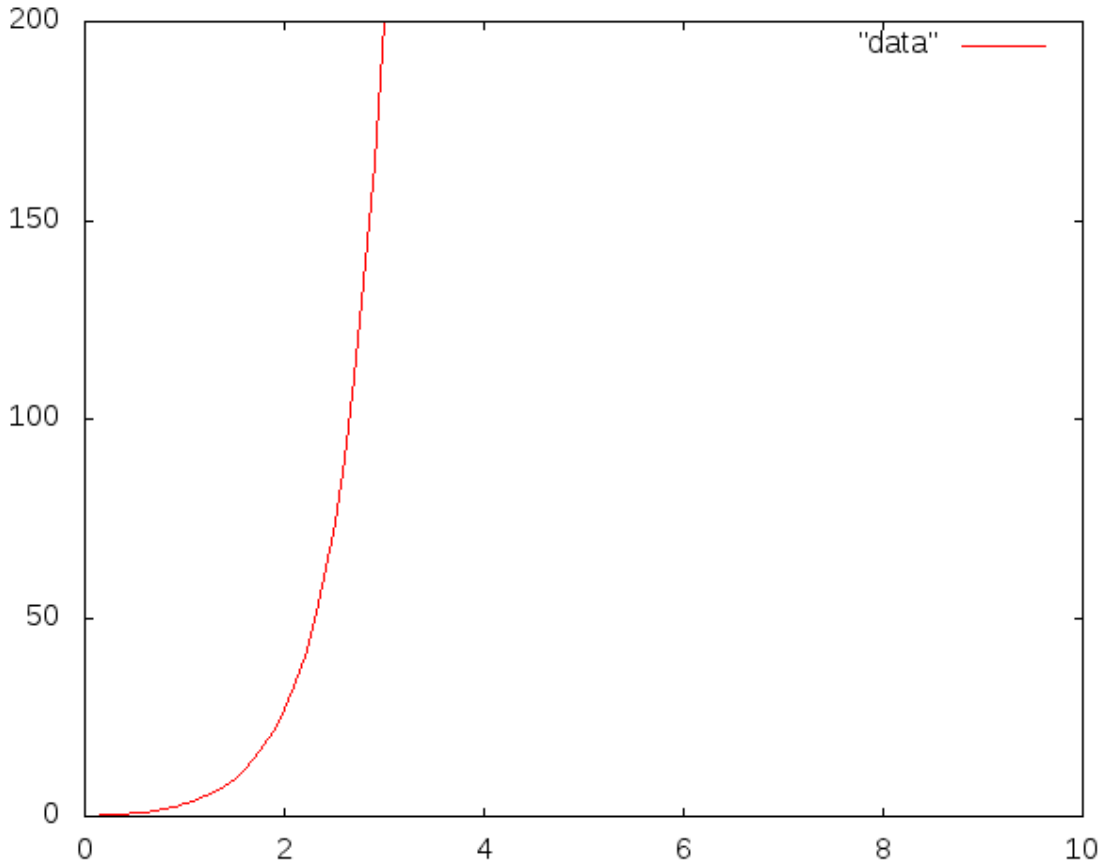
$$\frac{dx}{dt} = -2$$



Ordinary differential equations (ODEs)

Some examples:

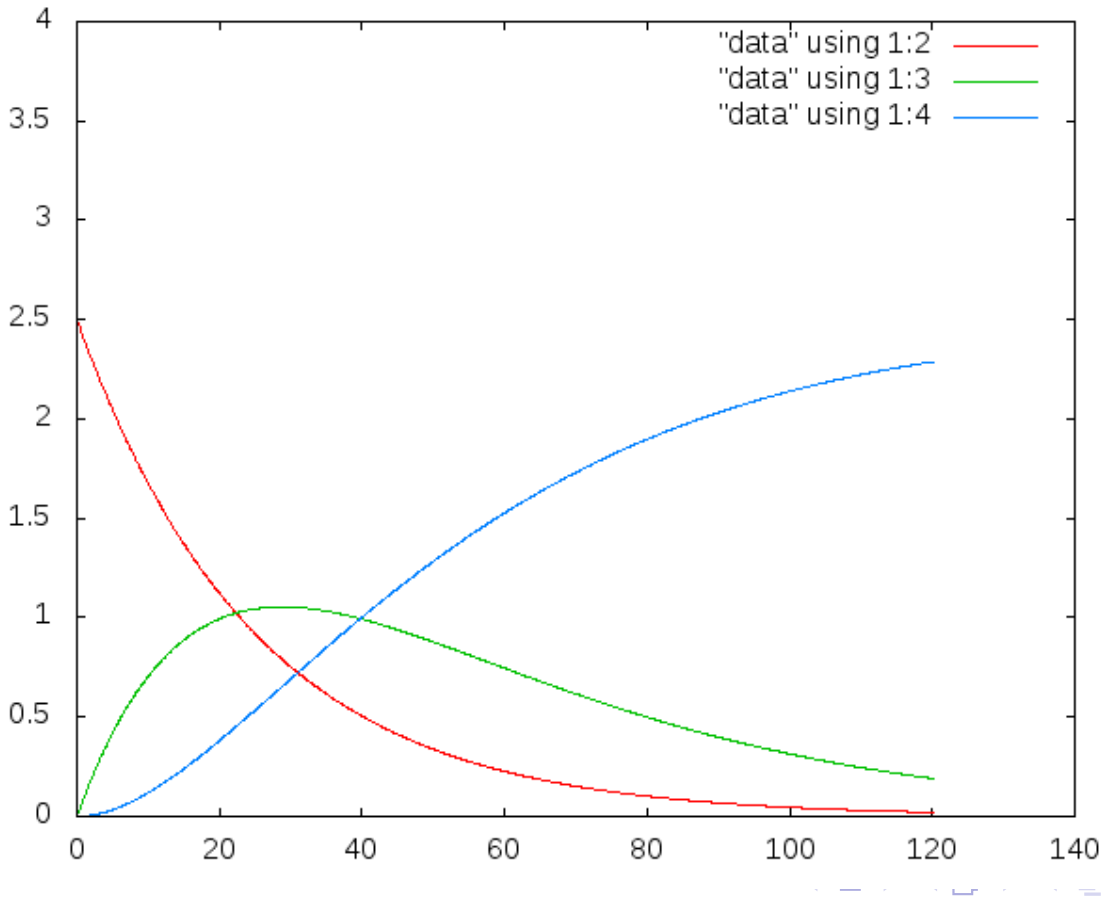
$$\frac{dx}{dt} = 2x + 1$$



Ordinary differential equations (ODEs)

Some examples:

$$\begin{cases} \frac{dx_1}{dt} = -0.04x_1 \\ \frac{dx_2}{dt} = 0.04x_1 - 0.03x_2 \\ \frac{dx_3}{dt} = 0.03x_2 \end{cases}$$



Numerical solution of ODEs

Since ODEs often cannot be solved analytically, numerical integration (or simulation) techniques are usually applied to find approximate solutions.

The simplest numerical integration algorithm is Euler's method.

Euler's method approximates a continuous function $x = f(t)$ on $0 \leq t \leq L$ (where L is a constant representing the simulation time) which is a solution of the differential equation $\frac{dx}{dt} = F(x, t)$ with initial condition $x_0 = \alpha$ (where α is a given constant).

The interval $0 \leq t \leq L$ is then divided into n equal parts, each of length h, with h very small (the smaller it is, the more precise the approximation will be). We obtain the points $t_i = ih$ with $i = 0, 1, \ldots, n$.

Euler's method consist in computing values x_i with $i = 0, 1, \ldots, n$ obtained by the following recurrence relation:

$$x_0 = \alpha$$

$$x_i + 1 = y_i + hF(x_i, t_i) \quad \text{with } i = 0, 1, \ldots, n - 1$$

The mass action kinetics of chemical reactions

The ODE model of a set of chemical reactions contains one differential equation for each molecular species

From these two reactions: $\ell_1 S_1 + \ldots + \ell_\rho S_\rho \underset{k_{-1}}{\overset{k}{\rightleftharpoons}} \ell'_1 P_1 + \ldots + \ell'_\gamma P_\gamma$

we obtain the following mathematical model:

$$\begin{cases} \frac{d[S_1]}{dt} = \ell_1 k_{-1}[P_1]^{\ell'_1} \cdots [P_\gamma]^{\ell'_\gamma} - \ell_1 k[S_1]^{\ell_1} \cdots [S_\rho]^{\ell_\rho} \\ \vdots \\ \frac{d[S_\rho]}{dt} = \ldots \\ \frac{d[P_1]}{dt} = \ell'_1 k[S_1]^{\ell_1} \cdots [S_\rho]^{\ell_\rho} - \ell'_1 k_{-1}[P_1]^{\ell'_1} \cdots [P_\gamma]^{\ell'_\gamma} \\ \vdots \\ \frac{d[P_\gamma]}{dt} = \ldots \end{cases}$$

The mass action kinetics of chemical reactions

$$2H_2 + O_2 \underset{0.5}{\overset{5}{\rightleftharpoons}} 2H_2O$$

ODEs obtained from mass action kinetics:

$$\begin{cases} \frac{d[H_2]}{dt} = 2 \cdot 0.5[H_2O]^2 - 2 \cdot 5[H_2]^2[O_2] \\ \frac{d[O_2]}{dt} = 0.5[H_2O]^2 - 5[H_2]^2[O_2] \\ \frac{d[H_2O]}{dt} = 2 \cdot 5[H_2]^2[O_2] - 2 \cdot 0.5[H_2O]^2 \end{cases}$$

$$E + S \underset{1000}{\overset{0.1}{\rightleftharpoons}} ES \underset{0.01}{\overset{0.3}{\rightleftharpoons}} E + P$$

ODEs obtained from mass action kinetics:

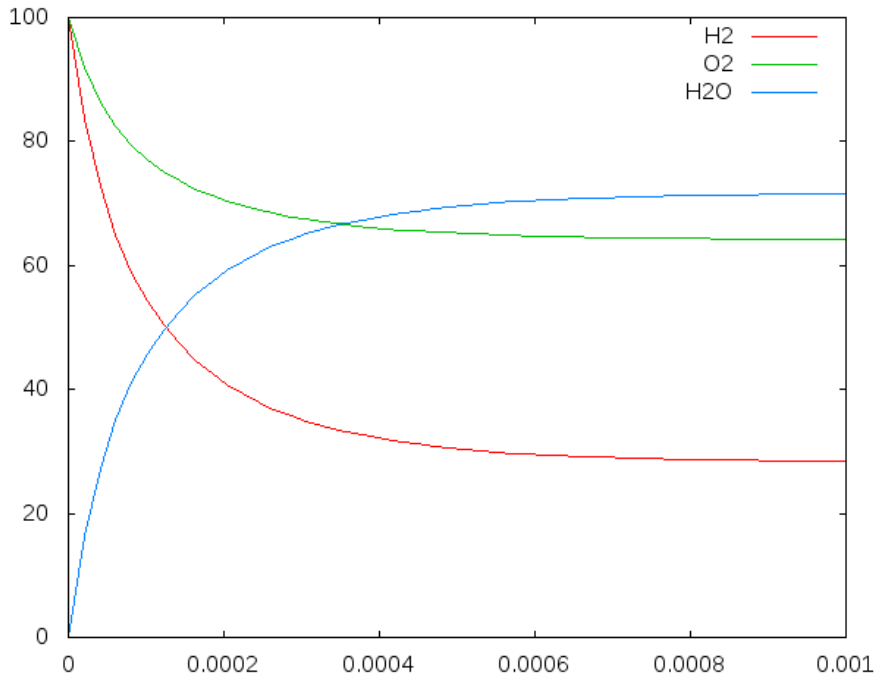
$$\begin{cases} \frac{d[E]}{dt} = 1000[ES] + 0.3[ES] - 0.1[E][S] \\ \frac{d[ES]}{dt} = 0.1[E][S] + 0.01[E][P] - 1000[ES] - 0.3[ES] \\ \frac{d[S]}{dt} = 1000[ES] - 0.1[E][S] \\ \frac{d[P]}{dt} = 0.3[ES] - 0.01[E][P] \end{cases}$$

The mass action kinetics of chemical reactions

We can now simulate the dynamics of the reactions (from given initial concentrations) by using numerical integration

$$2H_2 + O_2 \underset{0.5}{\overset{5}{\rightleftharpoons}} 2H_2O \qquad \begin{cases} \frac{d[H_2]}{dt} = 2 \cdot 0.5[H_2O]^2 - 2 \cdot 5[H_2]^2[O_2] \\ \frac{d[O_2]}{dt} = 0.5[H_2O]^2 - 5[H_2]^2[O_2] \\ \frac{d[H_2O]}{dt} = 2 \cdot 5[H_2]^2[O_2] - 2 \cdot 0.5[H_2O]^2 \end{cases}$$

with initial concentrations $[H_2]_0 = [O_2]_0 = 100$ and $[H_2O]_0 = 0$.



Other kinetic models of chemical reactions

In some cases the mass action kinetics could be considered too "low-level", and introduce unneccesary details in the models of some phenomena.

Two examples of more abstract kinetic models are:

- the Michaelis-Menten kinetics of enzyme-catalysed reactions
- the logistic function for reactions subject to saturation

These kinetic models are approximations of the mass action kinetics that could be considered reasonable under certain conditions and that make models simpler (hence, easier to be simulated numerically)

Outline of the talk

The Michaelis-Menten kinetics

The Michaelis-Menten kinetics describes the dynamics of irreversible enzyme-catalysed reactions such as:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

The assumptions under which the kinetics is an acceptable approximation of the mass action kinetics are the following:

- the whole enzyme-catalysed reaction is irreversible, namely the reaction $E + P \xrightarrow{k_{-2}} ES$ has a very small kinetic constant k_{-2} and can be ignored
- rates of $E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES$ as given by the mass action kinetics are much higher than the rate of $ES \xrightarrow{k_2} E + P$. Hence, the first two reactions can be considered in dynamic equilibrium.

The Michaelis-Menten kinetics

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

- by the dynamic equilibrium assumption we have $k_1[E][S] = (k_{-1} + k_2)[ES]$
- since the enzyme is not consumed we have $[E]_0 = [E] + [ES]$ constant
- hence, $k_1([E]_0 + [ES])[S] = (k_{-1} + k_2)[ES]$
- hence, $(k_{-1} + k_2)[ES] - k_1[ES][S] = k_1[E]_0[S]$
- hence, $[ES] = \frac{k_1[E]_0[S]}{(k_{-1}+k_2-k_1[S]}$ is the concentration of the complex ES at the dynamic equilibrium
- let $K_M = \frac{k_{-1}+k_2}{k_1}$, we obtain $[ES] = \frac{[E]_0[S]}{K_M+[S]}$
- we know that $\frac{d[P]}{dt} = k_2[ES]$, hence, $\frac{d[P]}{dt} = \frac{k_2[E]_0[S]}{K_M+[S]}$
- let $V_{max} = k_2[E]_0$, we finally obtain:

$$\frac{d[P]}{dt} = \frac{V_{max}[S]}{K_M + [S]}$$

that is the Michaelis-Menten equation.

The Michaelis-Menten kinetics

In the Michaelis-Menten equation

$$\frac{d[P]}{dt} = \frac{V_{max}[S]}{K_M + [S]}$$

we have only two parameters:

- K_M (called Michaelis-Menten constant) that corresponds to $\frac{k_{-1}+k_2}{k_1}$
- V_{max} that corresponds to $k_2[E]_0$ and is the max speed of production of P

The concentrations of the enzyme $[E]$ and $[ES]$ are no longer necessary variables in the ODE model of the considered reactions

The Michaelis-Menten kinetics

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

ODEs obtained from mass action kinetics:

$$\begin{cases} \frac{d[E]}{dt} = k_{-1}[ES] + k_2[ES] - k_1[E][S] \\ \frac{d[ES]}{dt} = k_1[E][S] - k_{-1}[ES] - k_2[ES] \\ \frac{d[S]}{dt} = k_{-1}[ES] - k_1[E][S] \\ \frac{d[P]}{dt} = k_2[ES] \end{cases}$$

ODEs obtained from Michaelis-Menten kinetics:

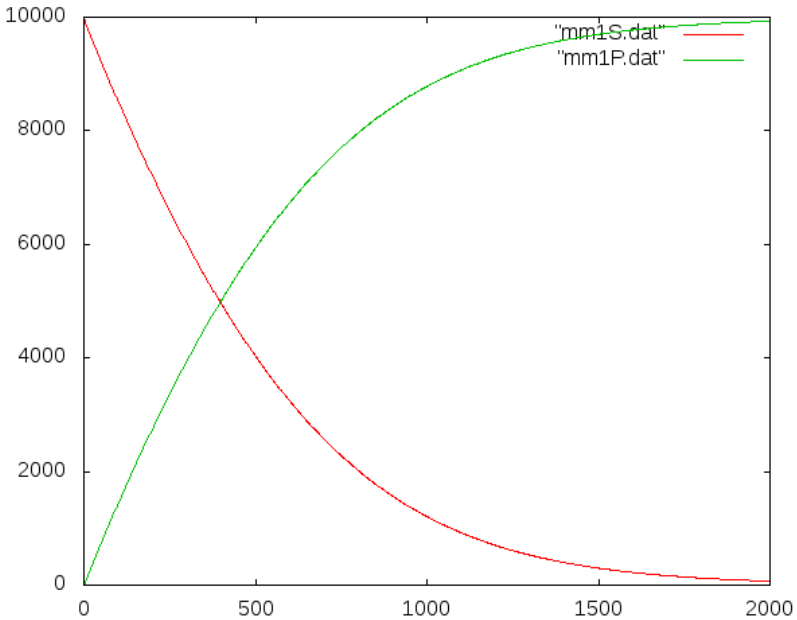
$$\begin{cases} \frac{d[S]}{dt} = -\frac{V_{max}[S]}{K_M+[S]} \\ \frac{d[P]}{dt} = \frac{V_{max}[S]}{K_M+[S]} \end{cases}$$

The Michaelis-Menten kinetics

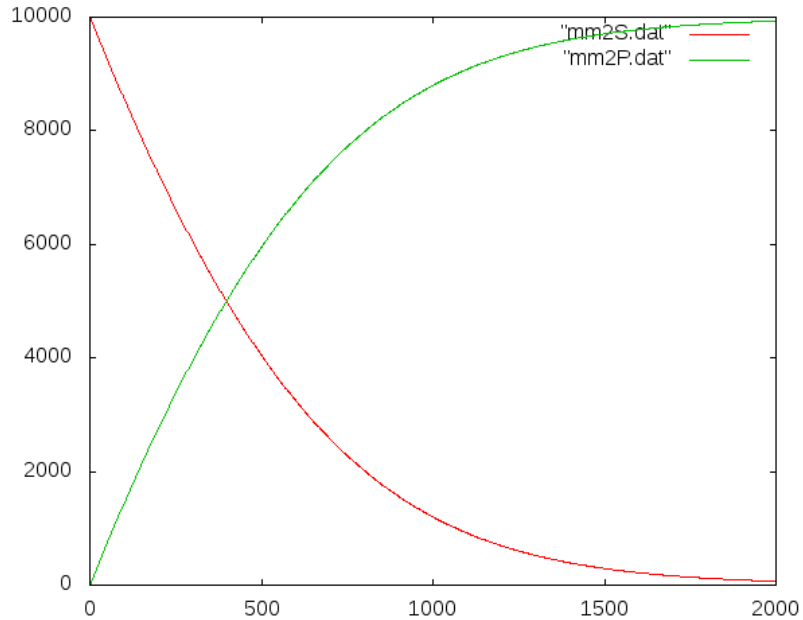
$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

Numerical simulation with $k_1 = 0.1, k_{-1} = 1000, k_2 = 0.3$ and with initial concentrations $[E]_0 = 100, [S]_0 = 10000, [ES]_0 = [P]_0 = 0$:

Mass Action



Michaelis-Menten



Outline of the talk

The logistic function

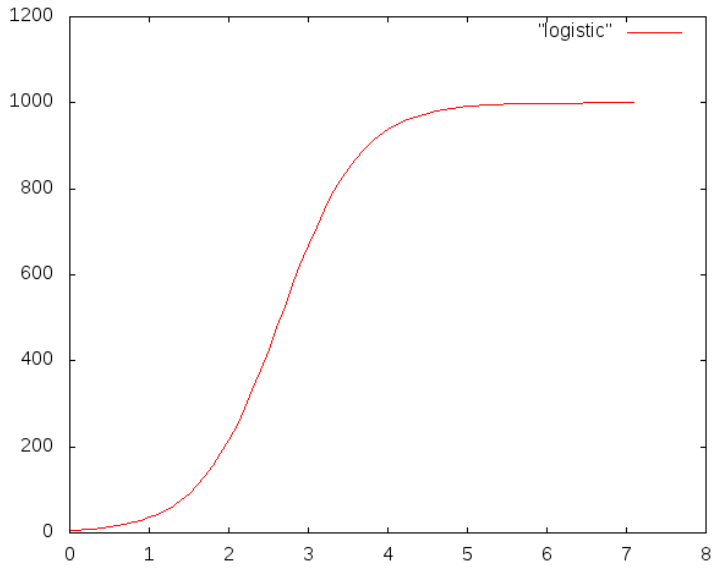
The logistic funciton finds common application in ecology to describe population growth in an enviroment with limited resources

A differential equation in which the logistic function is used is as follows:

$$\frac{dP}{dt} = \alpha P(1 - \frac{P}{K})$$

where α is a constant representing the growth rate and K is the carrying capacity, namely the maximal sustainable value of P

For example, with $\alpha = 2$ and $K = 1000$, starting from $P = 5$ we obtain



The logistic function

In general, the logistic function can be used to descibe the dynamics of something subject to an exponential growth combined with a saturation effect

An example of such kind of dynamics in (bio)chemistry are autocatalytic reactions, namely reactions such as

$$A + B \xrightarrow{k} 2B$$

Another biological example that can often be described with a logistic function is tumour growth

The logistic function

$$A + B \xrightarrow{k} 2B$$

ODEs obtained from mass action kinetics:

$$\begin{cases} \frac{d[A]}{dt} = -k[A][B] \\ \frac{d[B]}{dt} = k[A][B] \end{cases}$$

ODEs obtained by using the logistic function:

$$\left\{ \frac{d[B]}{dt} = k[B][A]_0(1 - \frac{[B]}{[A]_0+[B]_0}) \right.$$

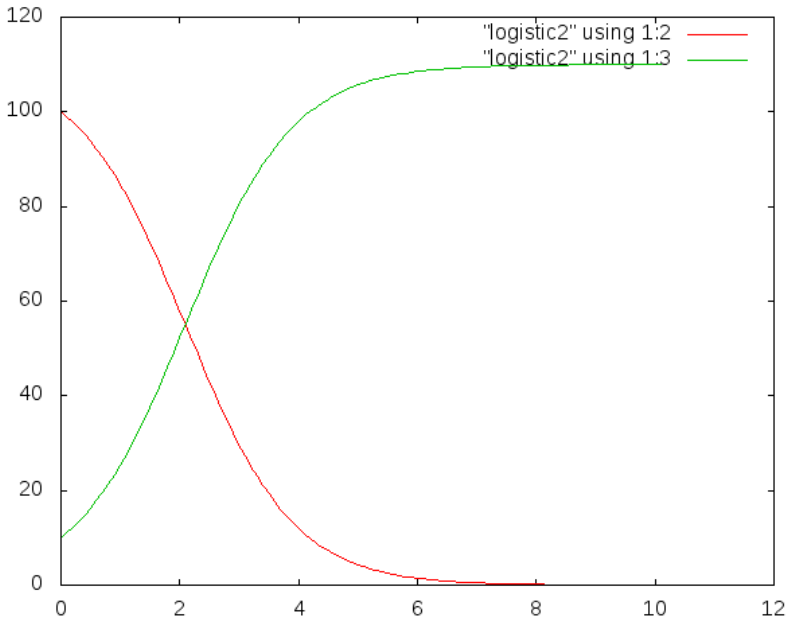
where $[A]_0$ and $[B]_0$ are constants representing the initial concentrations of A and B, respectively

The logistic function

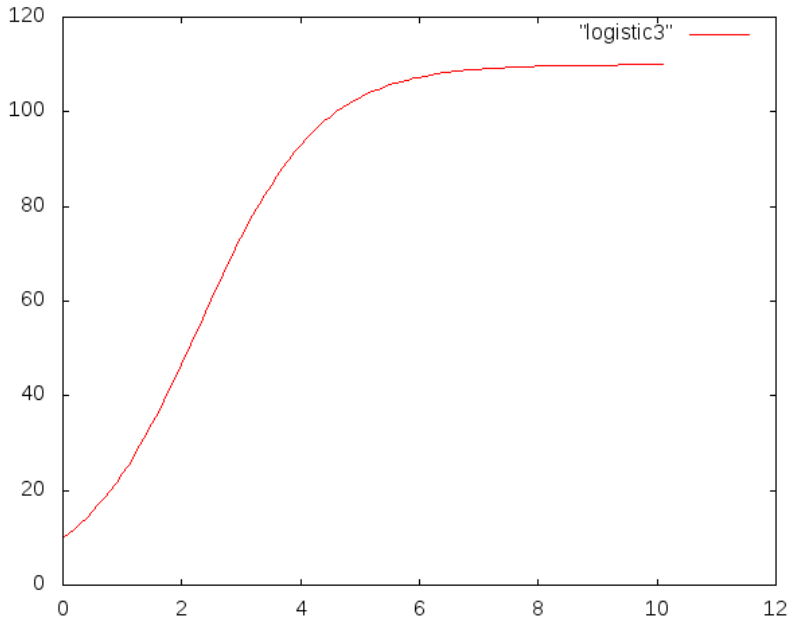
$$A + B \xrightarrow{k} 2B$$

Numerical simulation with $k = 0.01$, $[A]_0 = 100$ and $[B]_0 = 10$:

Mass Action (A and B)



Logistic function (only B)



Outline of the talk

Advantages of ODEs

We have seen that ODEs can be used to describe the dynamics of biological systems.

Advantages:

- ODEs offer tools for reasoning about the functioning of biological systems (validation of hypotheses, suggestions for new laboratory experiments, etc...)
- models based on ODEs can be used to make predictions
- ODEs can be also studied analytically
- the theory of ODEs is well-established (ODEs are the same in all books)
- a lot of rather efficient numeric solvers are available

Disadvantages of ODEs

We have seen that ODEs can be used to describe the dynamics of biological systems.

Disadvantages:

- When the complexity of the modelled system increases, ODE models become difficult to manage
- ODEs are continuous and deterministic. This makes them unsuitable to describe systems in which components occur in small numbers and subject to stochastic events (e.g. systems involving interaction with the DNA)

Next step: stochastic modelling and simulation...