# CPIB SUMMER SCHOOL 2011: INTRODUCTION TO BIOLOGICAL MODELLING

## *Lecture 2.1*
## *Building multi-variable differential equation models*

Markus Owen

1

## *Recap...*

- We have introduced simple ordinary differential equation (ODE) models for a single state variable.
- **Steady states** and their **stability** are crucial determinants of system dynamics.
- Changes in number or stability of steady states are called **bifurcations**.
- For 1st order **autonomous** ODEs, the **phase-line diagram** can tell us most of the qualitative information we'd like to know about the system dynamics:
  - if you can sketch the graph, you can sketch the dynamics...
  - steady states, stability AND qualitative solution behaviour (fast, slow, increasing, decreasing, etc), bifurcations.
  - solutions cannot oscillate
- For 1st order **non-autonomous** ODEs (e.g. circadian models with time dependent parameters) solutions can oscillate (driven by e.g. day-night cycle)
- We used CellDesigner to build and simulate simple one variable models.
- **Next, models with more than one state variable:** more complex dynamics possible, analysis more difficult, often resort to computer simulation

2

## *Signalling networks*

- Central dogma of molecular biology:
  - DNA transcribed to RNA (regulated by transcription factors),
  - RNA is translated into Protein.
- Proteins interact, can regulate translation, RNA stability, and transcription.
- RNA can also modulate transcription.
- Signalling networks: interactions between these elements, typically complex and extensive.
- Fundamental approach: decompose into modules that are sufficiently separate from other pathways to be considered on their own.
- Mathematical models: prediction of network behaviour with given topology and interactions.
- Ideas don't just apply to "gene networks", but to many kinds of network: Physiological models, metabolic networks, ecological networks, epidemiology...

3

# The law of Mass Action (1)

- The rate of a chemical reaction is proportional to the product of the concentrations of the reactants.
- The rate of change of a species depends on the rate of reaction and **the net change in the number of molecules** of that species.

- Consider a simple example: $A + B \underset{k_b}{\overset{k_f}{\rightleftharpoons}} C.$
- This is two reactions.
- Forward: reactants $A$ and $B$, and rate $k_f[A][B]$. Consumes one molecule of $A$ and $B$, and produces one molecule of $C$.
- Reverse: single reactant $C$, and rate $k_b[C]$. Consumes one molecule of $C$, and produces one molecule of $A$ and $B$.

$$\frac{d[A]}{dt} = -k_f[A][B] + k_b[C]$$
$$\frac{d[B]}{dt} = -k_f[A][B] + k_b[C]$$
$$\frac{d[C]}{dt} = +k_f[A][B] - k_b[C]$$

Forward Reverse

4

# The law of Mass Action (2)

- Another example, the "Brusselator":

$$A \overset{k_1}{\underset{1}{\rightarrow}} X, \quad B + X \overset{k_2}{\underset{2}{\rightarrow}} Y + D, \quad 2X + Y \overset{k_3}{\underset{3}{\rightarrow}} 3X, \quad X \overset{k_4}{\underset{4}{\rightarrow}} E$$

- We assume the concentrations of substrates $A$ and $B$ are constant.
- $E$ is a product. We are interested in the dynamics of $X$ and $Y$.
- $x = [X]$, $y = [Y]$, the concentrations of $X$ and $Y$.
- 3 has rate $k_3 x^2 y$ which *produces* one molecule of $X$ and consumes one of $Y$.

$$\frac{dx}{dt} = \overset{1}{k_1 A} - \overset{2}{k_2 B} \overset{4}{+ k_4} x + \overset{3}{k_3 x^2 y},$$
$$\frac{dy}{dt} = \overset{2}{k_2 B x} - \overset{3}{k_3 x^2 y}.$$

- This system is a famous example which can have oscillatory solutions.

5

# Michaelis-Menten enzyme kinetics

- $S$, substrate; $E$, enzyme; $P$, product: $S + E \underset{k_2\ 2}{\overset{k_1\ 1}{\rightleftharpoons}} SE \overset{k_3}{\underset{3}{\rightarrow}} P + E$

1 Reactants $S$ and $E$, rate $k_1[S][E]$. Consumes one molecule of $S$ and $E$, and produces one molecule of $SE$.

2 Single reactant $SE$, rate $k_2[SE]$. Consumes one molecule of $SE$, and produces one molecule of $S$ and $E$.

3 Single reactant $SE$, rate $k_3[SE]$. Consumes one molecule of $SE$, and produces one molecule of $P$ and $E$.

$$\frac{d[S]}{dt} = -\overset{1}{k_1[S][E]} + \overset{2}{k_2[SE]}$$
$$\frac{d[E]}{dt} = -\overset{1}{k_1[S][E]} + \overset{2}{k_2[SE]} + \overset{3}{k_3[SE]}$$
$$\frac{d[SE]}{dt} = \overset{1}{k_1[S][E]} - \overset{2}{k_2[SE]} - \overset{3}{k_3[SE]}$$
$$\frac{d[P]}{dt} = \overset{3}{k_3[SE]}$$

- Constant total enzyme
- Substrate assumed in excess
- $[SE]$ at **quasi-steady state**

$$\frac{d[P]}{dt} = \frac{V_{max}[S]}{K_m + [S]} \qquad V_{max} = k_3 E_0, \quad K_m = (k_2 + k_3)/k_1$$

## Michaelis-Menten kinetics: $S + E \underset{k_2}{\overset{k_1}{\rightleftharpoons}} SE \xrightarrow{k_3} P + E$

- Constant total enzyme: $[E] + [SE] = E_0$
- Substrate assumed in excess, $d[S]/dt = 0$
- $[SE]$ assumed to be at **quasi-steady state**

$$0 \cancel{\frac{d[SE]}{dt}} = k_1[S][E] - k_2[SE] - k_3[SE]$$

- A bit more algebra, using the definition: $K = k_1/(k_2 + k_3)$

$$[SE] = K[S][E] = K[S](E_0 - [SE])$$

$$[SE]\,(1 + K[S]) = KE_0[S]$$

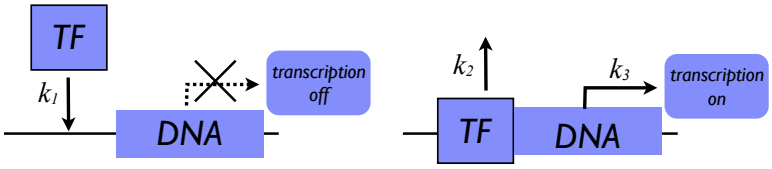
$$[SE] = \frac{KE_0[S]}{1 + K[S]} = \frac{E_0[S]}{(1/K + [S])}$$

Finally: $\dfrac{d[P]}{dt} = \dfrac{V_{max}[S]}{K_m + [S]}$ $\quad V_{max} = k_3 E_0$, $\quad K_m = (k_2 + k_3)/k_1$



## Transcriptional/translational activation

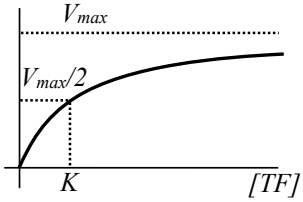
- TF binds to DNA, this complex activates production of protein P.




$$TF + DNA \underset{k_2}{\overset{k_1}{\rightleftharpoons}} TF\text{-}DNA \xrightarrow{k_3} P + TF\text{-}DNA$$

- Assuming TF binding is fast enables use of Michaelis-Menten approach.
- DNA acts as enzyme, $[DNA] + [TF\text{-}DNA] = 1$

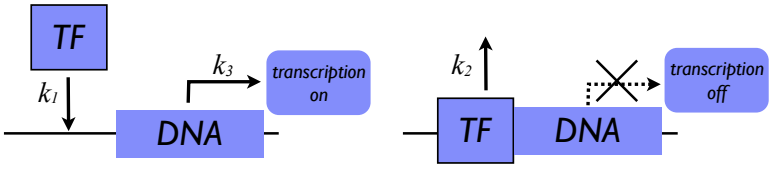
$$\frac{d[P]}{dt} = V_{max}\frac{[TF]}{K + [TF]}$$






## Transcriptional/translational repression

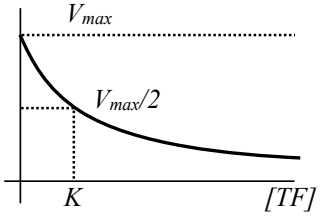
- TF binds to DNA, blocking production of protein P.




$$TF + DNA \underset{k_2}{\overset{k_1}{\rightleftharpoons}} TF\text{-}DNA, \qquad DNA \xrightarrow{k_3} P + DNA$$

- This time, synthesis is a decreasing function of TF concentration:

$$\frac{d[P]}{dt} = V_{max}\frac{K}{K + [TF]}$$



## Hill functions: co-operative activation

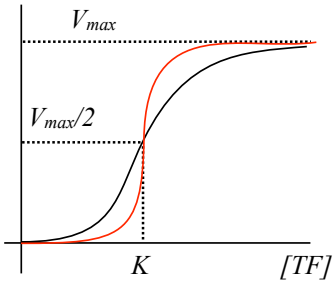
- With the previous expression, the response to changes in substrate (TF) concentration is weak.
- Cooperativity can lead to sharper responses.
- Suppose $n$ molecules of substrate bind to the DNA:

$$nTF + DNA \underset{k_2}{\overset{k_1}{\rightleftharpoons}} TF_n\text{-}DNA \xrightarrow{k_3} P + TF_n\text{-}DNA$$

- After some mass action and some algebra...

$$\frac{d[P]}{dt} = V_{max}\frac{[TF]^n}{K^n + [TF]^n}$$

- Larger $n$ ➔ steeper switch
- Same idea for repression




10

## Hill functions: co-operative repression

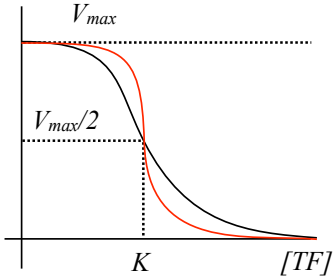
- Suppose $n$ molecules of substrate must bind to the DNA to block transcription:

$$nTF + DNA \underset{k_2}{\overset{k_1}{\rightleftharpoons}} TF_n\text{-}DNA, \qquad DNA \xrightarrow{k_3} P + DNA$$

- After some mass action and some algebra...

$$\frac{d[P]}{dt} = V_{max}\frac{[TF]^n}{K_n + [TF]^n}$$

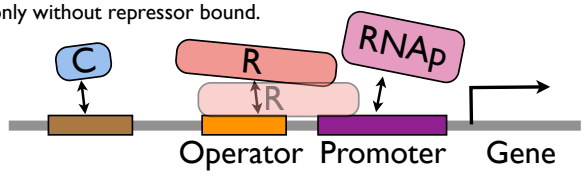
- Larger $n$ ➔ steeper 'off' switch




11

## Multiple TFs?

- What about multiple transcription factors? Some activating and some repressing?
- E.g. **Lac-operon**:
  Catabolite activator protein (CAP), $[C]$, activates (promotes binding of RNAp).
  Lac-repressor, $[R]$, blocks the RNAp binding site.
  Transcription only without repressor bound.




- Catalogue all the relevant states and the contribution of each to transcription rate. Transcription only when RNA-polymerase (RNAp) binds to the promoter.
- Write down ODEs and simplify using Michaelis-Menten approach.

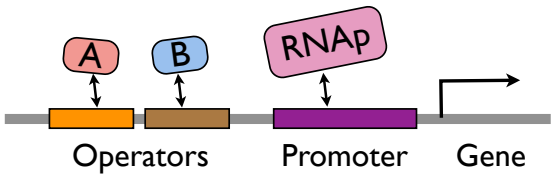
$$\frac{dP}{dt} = \frac{V_{max}(1 + k_1[C])}{1 + k_1[C] + k_2[R] + k_3[R][C]}$$

- or use the Shea-Ackers approach...

12

## Shea-Ackers (1)

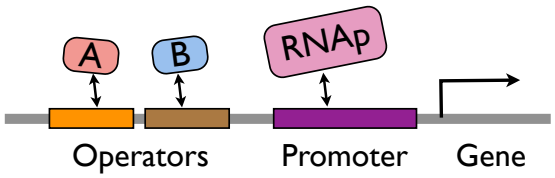
- Method originally developed for the lysis/lysogeny switch in Lambda phage
- Two time scales:
  - Slow: Transcription/Translation/Degradation
  - Fast: Binding/unbinding of TFs to gene – thermal equilibrium
- Possible cases: TF, TF+RNAp, RNAp - probability associated with each
- Enumerate all cases, compute probability of bound RNAp
- Transcription rate is proportional to promoter occupancy (by RNAp)




13

## Shea-Ackers (2)

- Example: two transcription factors, A and B
- Enumerate all possibilities - binding/unbinding of A,B and RNAp
- The "partition function" $Z$ contains $2^3 = 8$ terms




$$Z = \sum_{ijk} [A]^i [B]^j [RNAp]^k \delta_{ijk} = Z_{on} + Z_{off}$$

RNAp bound  RNAp unbound

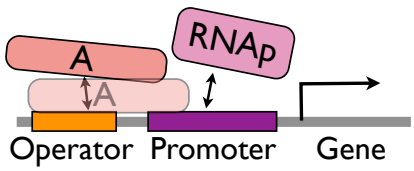
- ***i,j,k = 0*** (unbound) or ***1*** (bound)
- $\delta_{ijk}$ related to binding energy, $\delta_{000} = 1$

Transcription rate proportional to: $\dfrac{Z_{on}}{Z_{on} + Z_{off}}$

14

## Shea-Ackers: simple example

- The *trp* operon of E. coli is regulated by the TrpR repressor protein $A$.
- Tryptophan binds the TrpR repressor enabling TrpR to bind the *trp* operator.
- This prevents transcription: the *trp* operator overlaps the RNAp binding site. $A$ and $R$ cannot be simultaneously bound:




| [A] | [RNAp] | Rate |
|---|---|---|
| 0 | 0 | 1 |
| 1 | 0 | $\delta_{10}A$ |
| 0 | 1 | $\delta_{01}RNAp$ |
| 1 | 1 | - |

$$Z = [A]^0 [RNAp]^0 \delta_{00} + [A]^1 [RNAp]^0 \delta_{10} + [A]^0 [RNAp]^1 \delta_{01}$$

- Only the last term corresponds to a transcriptionally active state, so

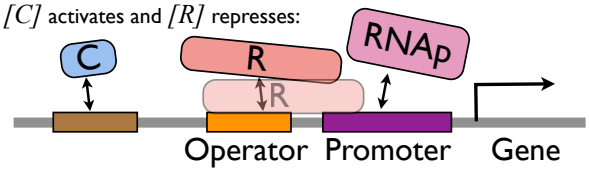
$$T \propto \frac{\delta_{01}[RNAp]}{1 + \delta_{10}[A] + \delta_{01}[RNAp]}$$

- For constant RNAp this is like a decreasing Hill function of order $1$.

15

## *Lac-operon revisited*

E.g. Lac-operon: $[C]$ activates and $[R]$ represses:




Michaelis-Menten approach:

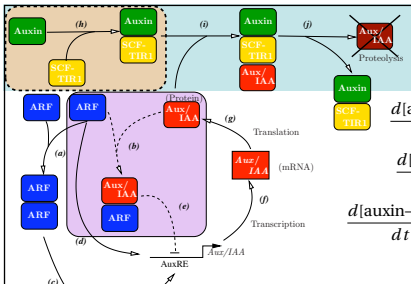
$$\frac{dP}{dt} = \frac{V_{max}(1 + k_1[C])}{1 + k_1[C] + k_2[R] + k_3[R][C]}$$

or use the Shea-Ackers approach... assuming [RNAp] is constant yields the same form as above.

| [C] | [R] | [RNAp] | Rate |
|---|---|---|---|
| 0 | 0 | 0 | 1 |
| 1 | 0 | 0 | $\delta_{100}[C]$ |
| 0 | 1 | 0 | $\delta_{010}[R]$ |
| 0 | 0 | 1 | $\delta_{001}[RNAp]$ |
| 1 | 1 | 0 | $\delta_{110}[C][R]$ |
| 1 | 0 | 1 | $\delta_{101}[C][RNAp]$ |
| 0 | 1 | 1 | - |
| 1 | 1 | 1 | - |

16

## Mass action




$$\frac{d[\text{auxin}]}{dt} = \omega + k_d[\text{auxin–TIR1}] - k_a[\text{auxin}][\text{TIR1}] - \mu_{\text{auxin}}[\text{auxin}],$$

$$\frac{d[\text{TIR1}]}{dt} = -k_a[\text{auxin}][\text{TIR1}] + k_d[\text{auxin–TIR1}],$$

$$\frac{d[\text{auxin–TIR1}]}{dt} = k_a[\text{auxin}][\text{TIR1}] - k_d[\text{auxin–TIR1}] + (l_d + l_m)[\text{auxin–TIR1–IAA}] - l_a[\text{auxin–TIR1}][\text{IAA}_p]$$

$$\frac{d[\text{IAA}_p]}{dt} = \delta[\text{IAA}_m] - l_a[\text{IAA}_p][\text{auxin–TIR1}] + l_d[\text{auxin–TIR1–IAA}] - p_a[\text{IAA}_p][\text{ARF}] + p_d[\text{ARF-IAA}],$$

$$\frac{d[\text{auxin–TIR1–IAA}]}{dt} = l_a[\text{IAA}_p][\text{auxin–TIR1}] - (l_d + l_m)[\text{auxin–TIR1–IAA}],$$

$$\frac{d[\text{IAA}^*]}{dt} = l_m[\text{auxin–TIR1–IAA}] - \mu_{\text{IAA}^*}[\text{IAA}^*],$$

$$\frac{d[\text{ARF}]}{dt} = -2q_a[\text{ARF}]^2 + 2q_d[\text{ARF}_2] - p_a[\text{ARF}][\text{IAA}_p] + p_d[\text{ARF-IAA}],$$

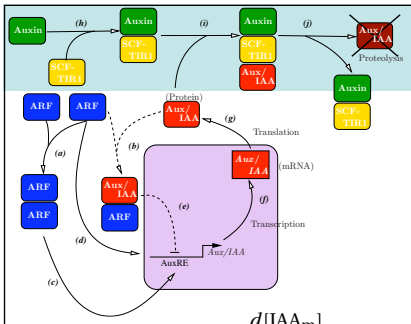
$$\frac{d[\text{ARF-IAA}]}{dt} = p_a[\text{ARF}][\text{IAA}_p] - p_d[\text{ARF-IAA}],$$

$$\frac{d[\text{ARF}_2]}{dt} = q_a[\text{ARF}]^2 - q_d[\text{ARF}_2],$$

$$\frac{d[\text{IAA}_m]}{dt} = F([\text{ARF}], [\text{IAA}_p], [\text{ARF-IAA}], [\text{ARF}_2]) - \mu_{\text{IAA}_m}[\text{IAA}_m],$$

17

## Transcriptional regulation




$$\frac{d[\text{IAA}_m]}{dt} = F([\text{ARF}], [\text{IAA}_p], [\text{ARF-IAA}], [\text{ARF}_2]) - \mu_{\text{IAA}_m}[\text{IAA}_m],$$

$$F([\text{ARF}], [\text{IAA}_p], [\text{ARF-IAA}], [\text{ARF}_2]) = \frac{\lambda_1 \frac{[\text{ARF}]}{\theta_{\text{ARF}}} + \lambda_2 \left( \frac{[\text{ARF}_2]}{\theta_{\text{ARF}_2}} + \frac{[\text{ARF}]^2}{\psi_{\text{ARF}}} \right)}{1 + \frac{[\text{ARF}]}{\theta_{\text{ARF}}} + \frac{[\text{ARF}_2]}{\theta_{\text{ARF}_2}} + \frac{[\text{ARF-IAA}]}{\theta_{\text{ARF-IAA}}} + \frac{[\text{ARF}][\text{IAA}_p]}{\psi_{\text{ARF-IAA}}} + \frac{[\text{ARF}]^2}{\psi_{\text{ARF}}}}$$

Equivalent to Shea-Ackers formulation

18

## Synthetic networks in E. coli

- **Key question:** How do we know that the simple mathematical representations we use are appropriate?
- **One answer:** Test their validity by constructing small networks in cells.
- Measurements on network state can then be compared directly to the behaviour predicted by simple mathematical models.
- Introduce a small number of genes controlled by promoters/repressors.
- Choose regulatory strengths based on a mathematical model of the network.
- A fluorescent reporter gives a read-out of a component of the network.
- The first examples of synthetic networks in *E. coli* were reported in 2000:
  - *Toggle switch* (Gardner, Cantor & Collins)
  - *Oscillator, a.k.a. "Repressilator"* (Elowitz & Leibler)

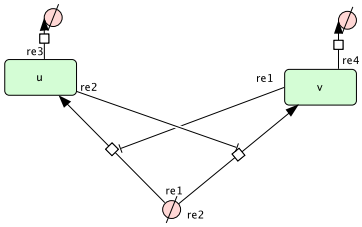


## Toggle switch: two-gene repressor network

Rate of change of ***u*** = production repressed by ***v*** - degradation

Rate of change of ***v*** = production repressed by ***u*** - degradation

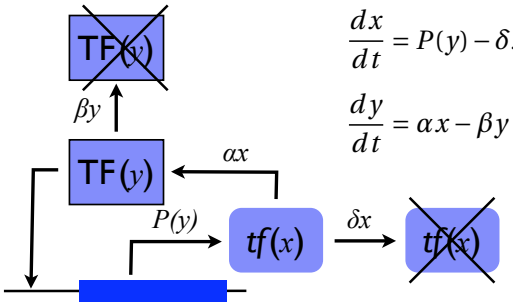
$$\frac{du}{dt} = \frac{\alpha_1}{1+v^\beta} - u, \qquad \frac{dv}{dt} = \frac{\alpha_2}{1+u^\gamma} - v$$




- Can behave as a bistable switch, depending on parameters
- **Phase-plane analysis** very useful - see later (Lecture 2.2)
- Expect high $u$ leads to low $v$ and vice-versa



## Transcriptional regulation and translation




$$\frac{dx}{dt} = P(y) - \delta x$$

$$\frac{dy}{dt} = \alpha x - \beta y$$

- Protein synthesis requires transcription and translation.
- Phase plane analysis quite straightforward (next lecture).

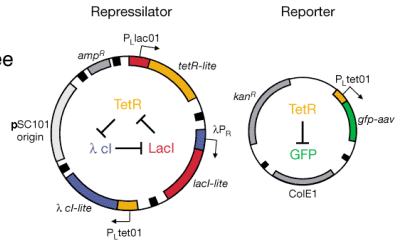
# An engineered negative feedback oscillator

**The repressilator:**

- Transfect *E. coli* with a plasmid containing three repressors:

$$lacI \dashv tetR \dashv cI \dashv lacI$$

- Also transfect with a reporter plasmid (visualise TetR expression)




- Represent the system using six variables: three mRNAs and three proteins.
- Linear degradation.
- "Hill function" transcriptional repression.
- Basal transcription.
- Linear translation.

$$\frac{dm_i}{dt} = \alpha_0 + \frac{\alpha}{1+p_j^n} - m_i \qquad i = lacI,\ tetR,\ cI$$

$$\frac{dp_i}{dt} = \beta(m_i - p_i) \qquad j = cI,\ lacI,\ tetR$$

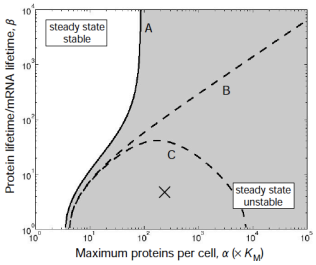
Elowitz, M.B. & Leibler, S. *Nature* **403**, 335 (2000).



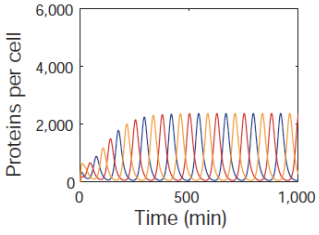
# An engineered negative feedback oscillator

$$\frac{dm_i}{dt} = \alpha_0 + \frac{\alpha}{1+p_j^n} - m_i, \quad \frac{dp_i}{dt} = \beta(m_i - p_i)$$

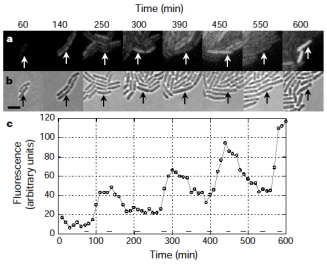
- Use the mathematical model to explore the dynamics as a function of the parameters.
- Engineer the promoters and molecular degradation rates appropriately for oscillations.
- Uses "*LINEAR STABILITY ANALYSIS*"




A: $n = 2.1, \alpha_0 = 0$. B: $n = 2, \alpha_0 = 0$. C: $n = 2, \alpha_0/\alpha = 0.001$.



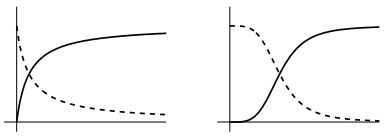

Track bacteria with time lapse over several division cycles (marked with bars in c).






# Gene network modelling

- Variables: mRNAs and proteins.
- ODE models: mass action, sigmoidal transcriptional activation and repression, linear decay and translation.



$$\frac{dx}{dt} = \text{synthesis} - \text{decay} \pm \text{transformation} \pm \text{transport}$$

- Parameters:
  - Thresholds for the sigmoidal functions;
  - effective co-operativities, can be high for indirect pathways;
  - half-lives;
  - relative contributions of multiple transcriptional regulators;
  - transfer rates, e.g. cytosol to cell surface;
  - transformation rates, e.g. cleavage, phosphorylation, binding.
- intracellular species: single equation per cell
- cell-surface: multiple equations per cell (e.g. six if we assume hexagonal cells).

## Epidemiology

**Simplest model: SIR model.**

- Closed population. Individuals do not enter, and leave only by death due to disease.
- Population in 3 compartments: Susceptible, Infective, or Removed (cured and now immune, or dead).
- No spatial effects (uniform mixing), and no heterogeneity in activity (important in, e.g., STDs such as AIDS).
- Negligible incubation time.
- Susceptibles move into Infective class at rate proportional to number of contacts between Susceptibles and Infectives (like law of mass action).
- Infectives removed at some rate into Removed class (*which decouples*).
- An EPIDEMIC if $I(t)>I(0)$ for some $t>0$ (i.e. if the number of infectives goes up)

$$\frac{dS}{dt} = -\beta SI$$

$$\frac{dI}{dt} = \beta SI - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$

- Constant total population

$$S + I + R = N$$

$$S + I \leq N$$

25

## Summary

- Mathematical models can encode our knowledge about signalling networks.
- Gene transcription, mRNA translation, protein interactions, decay, etc, can be described using differential equations.
- There are different approaches to combining multiple transcription factors.
- Mathematical analysis of relatively simple models can be done using **phase-plane** methods - next lecture.
- Complex signalling networks often require a more computational approach.

26