Lecture 4: Transcription networks – basic concepts

- Activators and repressors
- Input functions; Logic input functions; Multidimensional input functions
- Dynamics and response time

2.1 Introduction

- The cell is an integrated device made of several thousand types of interacting proteins
- Typical numbers, lengths, timescales: Tables 2.1 & 2.2 / Alon
- A cell continuously monitors its environment and calculates the amount at which each type of protein is needed. This information processing function, which determines the rate of production of each protein, is largely carried out by **transcription networks**

2.2 The cognitive problem of the cell

- Cells respond to environmental signals by producing appropriate proteins that act upon the internal and external environment
- To represent these environmental states, the cell uses special proteins called **transcription factors**
- Transcription factors are usually designed to transit rapidly between *active* and *inactive* molecular states, at a rate that is modulated by a specific environmental signal
- Each active transcription factor can bind the DNA to regulate the rate at which specific target genes are read

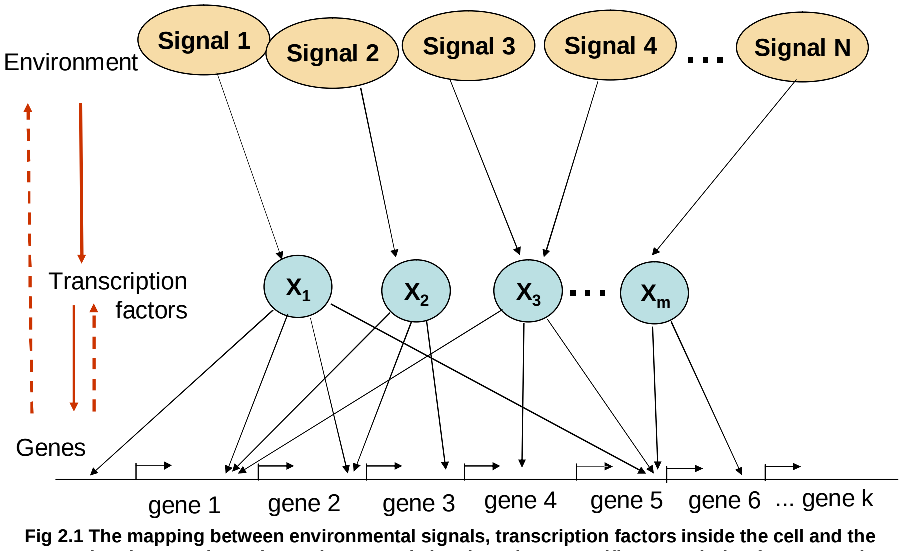

Fig 2.1 The mapping between environmental signals, transcription factors inside the cell and the genes that they regulate. The environmental signals activate specific transcription factor proteins. The transcription factors, when active, bind DNA to change the transcription rate of specific target genes, the rate at which mRNA is produced. The mRNA is then translated into protein. Hence, transcription factors regulate the rate at which the proteins encoded by the genes are produced. These proteins affect the environment (internal and external). Some proteins are themselves transcription factors, that can activate or repress other genes, etc.

2.3 Elements of transcription networks

- ***RNA polymerase*** (RNAp) produces mRNA
- mRNA is translated into a protein (= gene product)
- ***Promoter*:** a regulatory region of DNA that precedes the gene
- Transcription factors can act as ***activators*** or as ***repressors***
- ***Transcription network*** describes all regulatory transcription interactions in a cell. In the network, the *nodes* are genes and *edges* represent transcriptional regulation of one gene by the protein product of another gene
- Directed edge X → Y: product of gene X is a transcription factor protein that binds the promoter of gene Y to control the rate at which gene Y is transcribed
- The inputs to the network are ***signals***. Each signal is a small molecule, protein modification, or molecular partner that directly affects the activity of one of the transcription factors
- Signal S_X can cause protein X to rapidly shift to its active state X*, bind the promoter of gene Y, and increase the rate of transcription, leading to increased production of protein Y

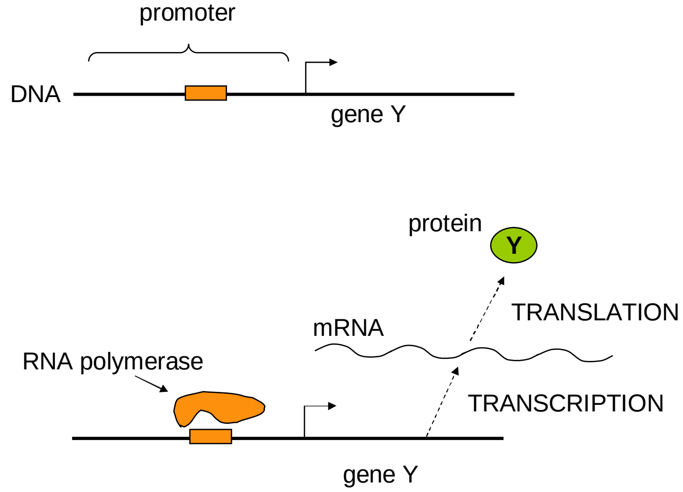

Fig 2.2 (a) GENE TRANSCRIPTIONAL REGULATION, THE BASIC PICTURE. (a) Each gene is usually preceded by a regulatory DNA region called the promoter. The promoter contains a specific site (DNA sequence) that can bind RNA polymerase (RNAp), a complex of several proteins that forms an enzyme that can synthesize mRNA that is complementary to the genes coding sequence. The process of forming the mRNA is called transcription. The mRNA is then translated into protein.

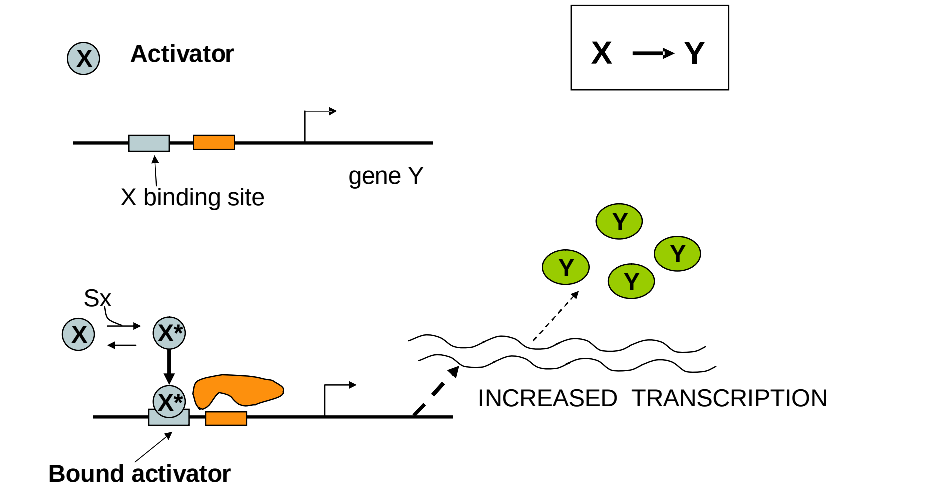

Fig 2.2 (b) An activator X, is a transcription- factor protein that increases the rate of mRNA transcription when it binds the promoter. The activator transits rapidly between active and inactive forms. In its active form, it has a high affinity to a specific site (or sites) on the promoter. The signal Sx increases the probability that X is in its active form X*. Thus, X* binds the promoter of gene Y to increase transcription and production of protein Y. The timescales are typically sub-second for transitions between X and X*, seconds for binding/ unbinding of X to the promoter, minutes for transcription and translation of the protein product, and tens of minutes for the accumulation of the protein,

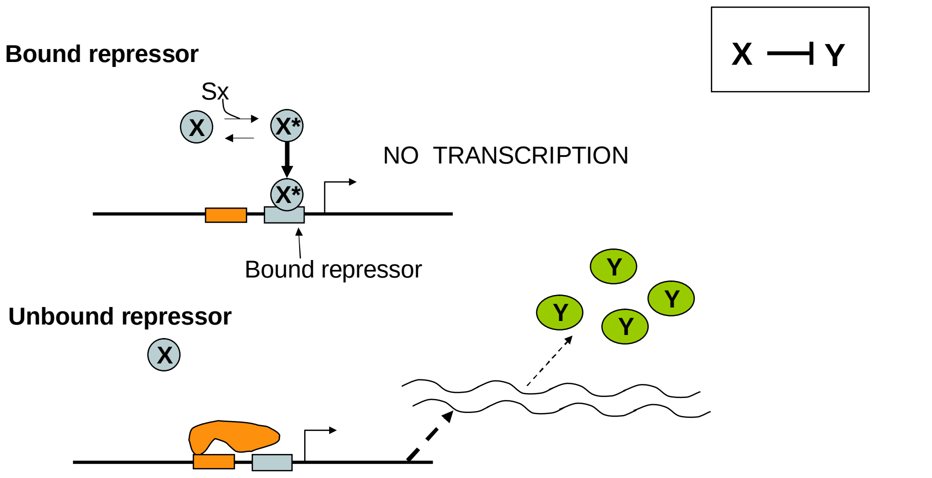

Fig 2.2 (c) A repressor X, is a transcription- factor protein that decreases mRNA transcription when it binds the promoter. The signal Sx increases the probability that X is in its active form X*. X* binds a specific site in the promoter of gene Y to decrease transcription and production of protein Y. Many genes show a weak (basal) transcription when repressor is bound.

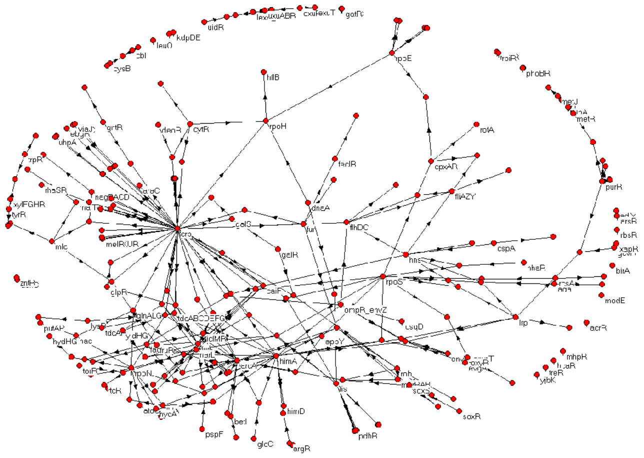

Fig 2.3. A transcription network that represents about 20% of the transcription interactions in the bacterium *E. coli*. Nodes are operons (groups of genes coded on the same mRNA). Some operons encode for transcription factors. Edges directed from node X to node Y indicate that the transcription factor encoded in X regulates operon Y. This network contains data on verified direct transcriptional interactions from many labs, compiled in databases such as regulonDB and Ecocyc.
Source: **Shen-Orr SS, Milo R, Mangan S, Alon U. Network motifs in the transcriptional regulation network of Escherichia coli. Nat Genet. 2002 May;31(1):64-8.**

2.3.1 Separation of timescales

Binding of a small molecule (a signal) to a transcription factor, causing a change in transcription factor activity	~1 msec
Binding of active transcription factor to its DNA site	~1 sec
Transcription + translation of the gene	~5 min
Timescale for 50% change in concentration of the translated protein (stable proteins)	~1 h (one cell generation)

Table 2.2: Timescales for the reactions in the transcription network of the bacterium *E. Coli* (order of magnitude)

2.3.2 The signs of edges: activators and repressors

- Activation or positive control ~ sign of an edge: +
- Repression or negative control ~ sign of an edge: -
- Positive control typically more common
  - 60 – 80% of interactions are positive in *E. Coli* & yeast
- A transcription factor typically acts primarily as either as an activator or as a repressor; but sometimes an activator can act also as a repressor and vice versa

2.3.3 Input function gives weights (numbers) on the edges

- The strength of the interaction associated with an edge of the network is described by an **input function**
- Consider X → Y
- X* = concentration of the active form of X
- Rate of production of Y = f(X*) where f is the input function
- **Hill input function for activators:**

$$f(X^*) = \frac{\beta X^{*n}}{K^n + X^{*n}}$$

  - *Activation coefficient* K = the concentration of active X that is needed to significantly activate expression of Y; half-maximal expression is reached when X* = K
  - *Maximal expression level* β of Y: this is reached when X* >> K
  - *Hill coefficient* n: governs the steepness of the input function; typically n = 1 ... 4

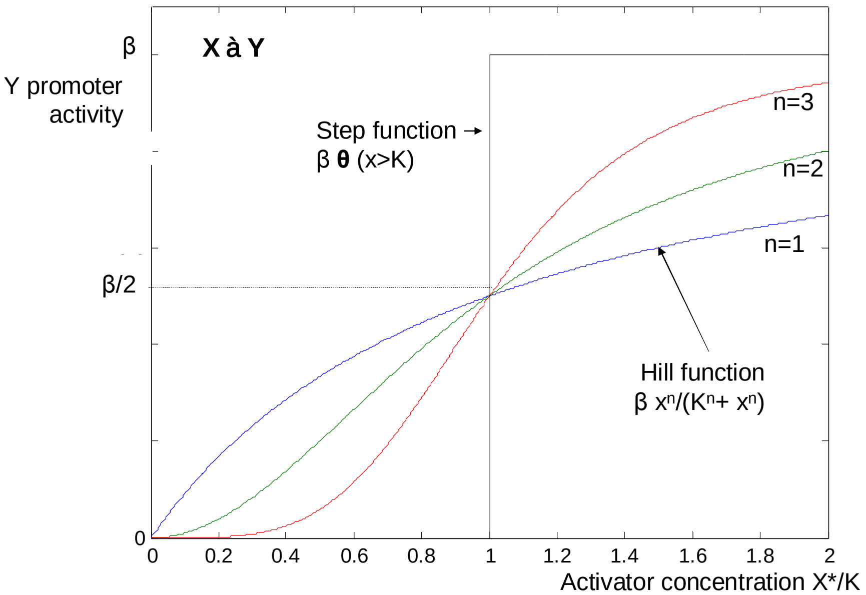

Figure 2.4 (a): Input functions for activator X described by Hill functions with Hill coefficient n=1,2 and 4. Black line: step function, also called logical input function. The maximal promoter activity is β , and K is the threshold for activation of a target gene (the concentration of X* needed for 50% maximal activation).

Input function (cont.)

- **Hill input function for repressors:** $$f(X^*) = \frac{\beta}{1+\left(\frac{X^*}{K}\right)^n}$$
  - Repression coefficient K: half-maximal repression when X* = K
  - Maximal expression level β: reached when X* = 0

- Hill input function can be given by parameters (β, K, n)
- These parameters can readily been tuned during evolution of an organism
  - Lab experiments have shown that when placed in new environment, bacteria can tune these parameters within hundreds of generations to reach otpimal expresion levels

- Many genes have a nonzero minimal expression level, called the **basal expression level**
  - Basal level representation: add term β_0 to the input function

2.3.4 Logic input functions

- Logic approximation:
  a gene is either OFF: f(X*) = 0
  or maximally ON: f(X*) = β
- K = threshold for activation
- Step-function Θ:

$$\Theta(z) = \begin{cases} 0 & if \;\; z = false \\ 1 & if \;\; z = true \end{cases}$$

- => logic approximation for activator:
  f(X*) = βΘ(X* > K)
- => logic approximation for repressor:
  f(X*) = βΘ(X* < K)

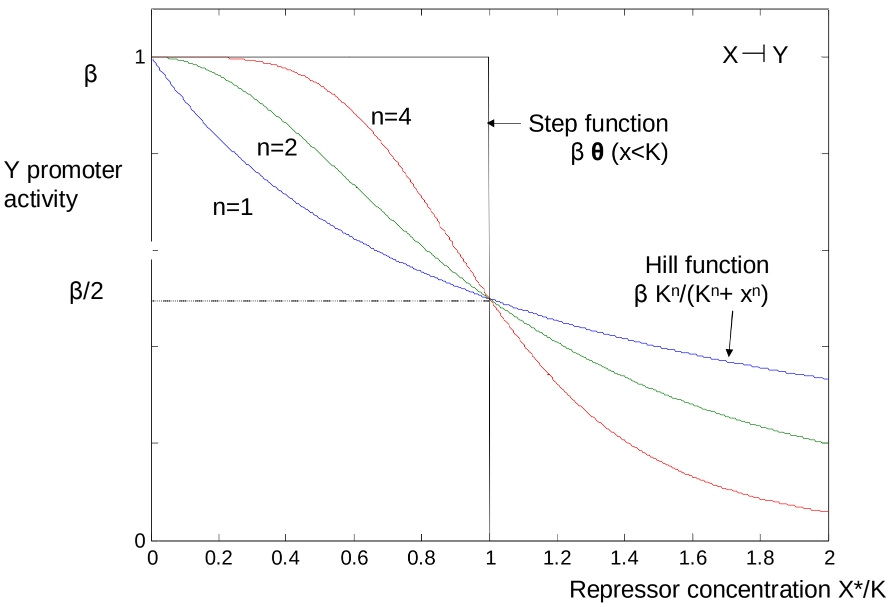

Figure 2.4 (b): Input functions for repressor X described by Hill functions with Hill coefficient n=1,2 and 4. Black line: logical input function (step function). The maximal unrepressed promoter activity is β , and K is the threshold for repression of a target gene (the concentration of X* needed for 50% maximal repression).

2.3.5 Multi-dimensional input functions for genes with several inputs

- AND:
  $f(X^*,Y^*) = \beta\Theta(X^* > K_X)\cdot\Theta(Y^* > K_Y) \sim X^* \text{ AND } Y^*$
- OR:
  $f(X^*,Y^*) = \beta\Theta(X^* > K_X \text{ OR } Y^* > K_Y) \sim X^* \text{ OR } Y^*$
- SUM:
  $f(X^*,Y^*) = \beta_X X^* + \beta_Y Y^*$ (= additive linear inputs)

- Other input functions also possible; see Fig 2.5
- It appears that the precise form of the input function of each gene is under selection pressure during evolution

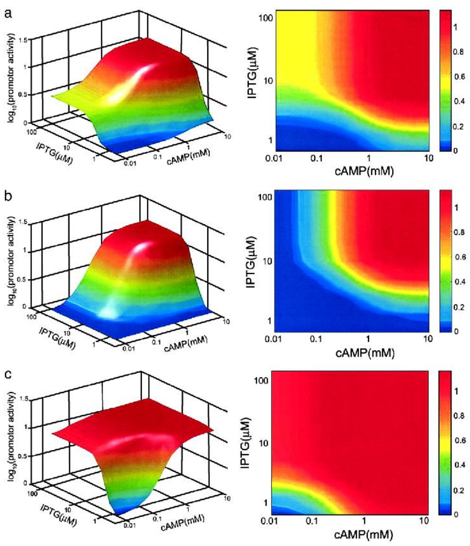

Fig 2.5: Two-dimensional input functions. (a) Input function measured in the lac-promoter of E. coli, as a function of two inducers cAMP and IPTG. (b) an AND-gate like input function, which shows high promoter activity only if both inputs are present. (c) an OR-gate like input function that shows high promoter activity if either input is present. Source: Setty, Y. et al. (2003) Proc. Natl. Acad. Sci. USA 100, 7702-7707

2.4 Dynamics and response time of simple gene regulation

- Consider X → Y (that is: Y is regulated only by X)
- What is the **dynamics** and the **response time** of the concentration of Y? Assume constant input signal S_X.
- A cell produces Y at a constant rate β (in units: concentration / time)
- Concentration of Y is balanced by
  - **Degradation** (= destruction of Y by spezialized proteins in the cell)
  - **Dilution** (= reduction in concentration of Y due to the increase of cell volune during growth)
- **Degradation rate** α_{deg} (in units: 1/time)
- **Dilution rate** α_{dil} (in units: 1/time)

2.4 Dynamics and response time of simple gene regulation

- Total rate: $\alpha = \alpha_{dil} + \alpha_{deg}$
- Change in the concentration of Y is described by a **dynamic equation**

$$dY/dt = \beta - \alpha Y \qquad \text{(DE)}$$

- Steady state = Y reaches a constant concentration Y_{st} that does not change because of dY/dt
- => Solve (DE) for dY/dt = 0. This gives steady state concentration

$$Y_{st} = \beta/\alpha$$

Response time: decay

- What happens if $Y = Y_{st}$ and we take away the input signal S_X and hence $\beta = 0$? The level of Y will decay exponentially in t:
- Solving (DE) for $\beta = 0$ gives

$\boxed{Y(t) = Y_{st}\, e^{-\alpha t}}$ (e is the base of natural log)

- Response time is a measure for the speed at which Y levels change:

- Def.: **Response time** $T_{1/2}$ is the time needed to reach halfway between the initial and final states in a dynamic process

- To find $T_{1/2}$ for the decay from Y_{st}, solve t from $Y(t) = Y_{st}/2$:

  $\Rightarrow Y_{st}\, e^{-\alpha t} = Y_{st}/2 \quad \Rightarrow \quad -\alpha t \ln e = \ln ½$

  $\Rightarrow t = \boxed{T_{1/2} = (\ln 2)/\alpha}$

  Response time depends only on α; Fig 2.6 (a)

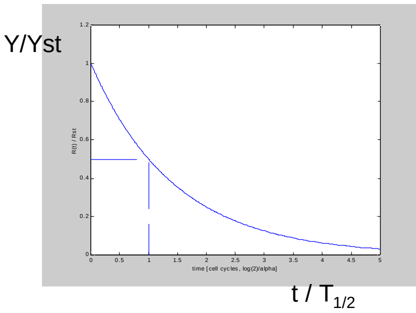

Fig 2.6a: Decay in protein concentration following a sudden drop in production rate. Note the response-time, the time it takes the concentration to reach half of its variation, is $T_{1/2}$=log(2)/ α =1 in these units (dashed lines).

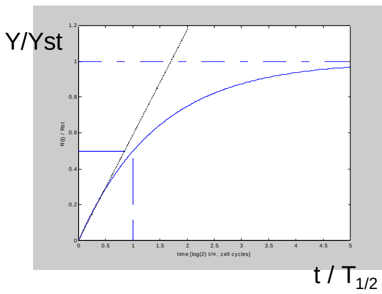

Fig 2.6b: Rise in protein concentration following a sudden increase in production rate. Note the response time, the time it takes the dynamics to reach half of its variation, is $T_{1/2}$=log(2)/ α =1 in these units (dashed lines). At early times, the protein accumulation is approximately linear with time, Y=β t (dotted line).

Response time: accumulation

- What happens if unstimulated gene (i.e., Y = 0) becomes suddenly stimulated by a strong S_X? The level of Y will start to accumulate towards the steady state concentration Y_{st}:
- Solving (DE):
  (DE) is in fact a first order differential equation with constant coefficients
  $dY/dt + \alpha Y = \beta$ whose solution is of the form $Y(t) = \beta/\alpha + Ce^{-\alpha t}$ (Check that this is correct solution by substituting to (DE)!)
- Requiring that Y(0) = 0, gives $C = - \beta/\alpha$
- Hence we obtain

$$\boxed{Y(t) = \beta/\alpha(1 - e^{-\alpha t}) = Y_{st}(1 - e^{-\alpha t})}$$

- At early times when $\alpha t << 1$, we can use the Taylor expansion $e^{-\alpha t} \sim 1 - \alpha t$ (and $Y_{st} = \beta/\alpha$) to obtain

$$Y \sim \beta t$$

  i.e., at early times the accumulation slope equals the production rate; see Fig. 2.6 (b)
- To obtain the **response time** to reach $Y_{st}/2$, solve the time from $Y(t) = Y_{st}/2$:

$$\boxed{T_{1/2} = (\ln 2)/\alpha}$$

- So, response time is the same for both decay and accumulation

Response time of stable proteins

- Stable proteins are such that $\alpha_{deg} = 0$ (stable protein is not actively degraded in growing cells)
- The production of stble proteins is balanced only by dilution due to the increasing volume of the growing cell, that is, $\alpha = \alpha_{dil}$
- *Response time for stable proteins is equal to one cell generation time*:
  - Imagine that a cell produces a stable protein and then suddenly the production stops ($\beta = 0$). The cell grows and, when it doubles its volume, splits into two cells. Thus, after one cell generation time τ, the protein concentration has degreased by 50%, and therefore

$$\boxed{T_{1/2} = (\ln 2)/ \alpha_{dil} = \tau.}$$

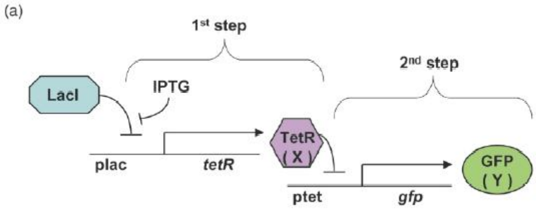

Fig 2.7a A transcriptional cascade made of two repressors in E. coli

The transcription cascade (X à Y à Z) is made of two well-studied repressors, the lacI and tetR . TetR controls the green fluorescent protein gene. Source: Rosenfeld, Alon JMB 2003

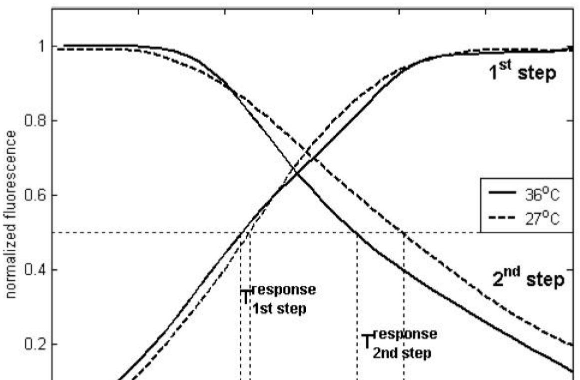

Fig 2.7b Response time is about one cell generation per cascade step.
The first step in the cascade rises in response to the inducer IPTG that inactivates the repressor LacI. This step encodes a repressor TetR that causes a decreases in the activity of the second-step promoter. Both steps were monitored using a green fluorescent protein (GFP) reporter, in two strains, one that reports for each step in the cascade. The experiment was carried out at two temperatures, which show different cell-generation times (about 2-fold slower at 27C compared to 36C). The x-axis shows time in units of the cell generation time in each condition.