

Essence of regulation

The control of a cell's "chemistry"

In simple cases, we can match the “regulation circuit” to the actual molecular processes, and hence have a handle on e.g. the rates through a model that has the “real” physics.

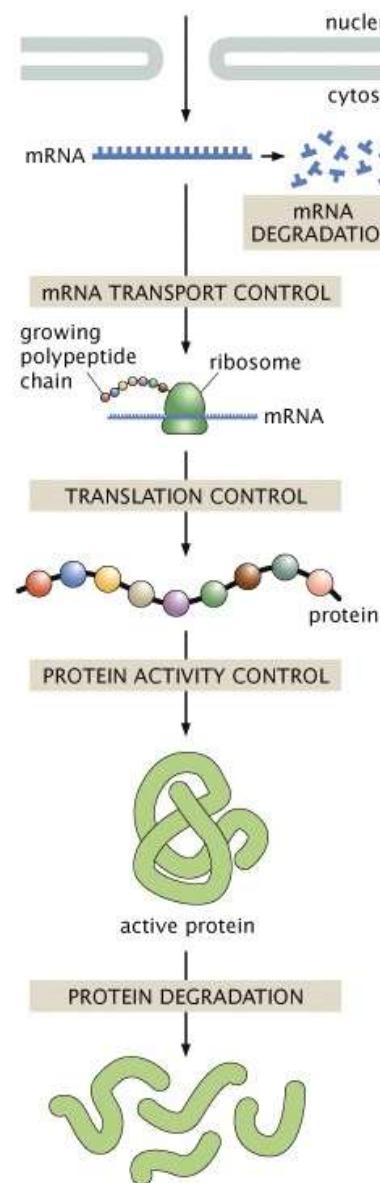
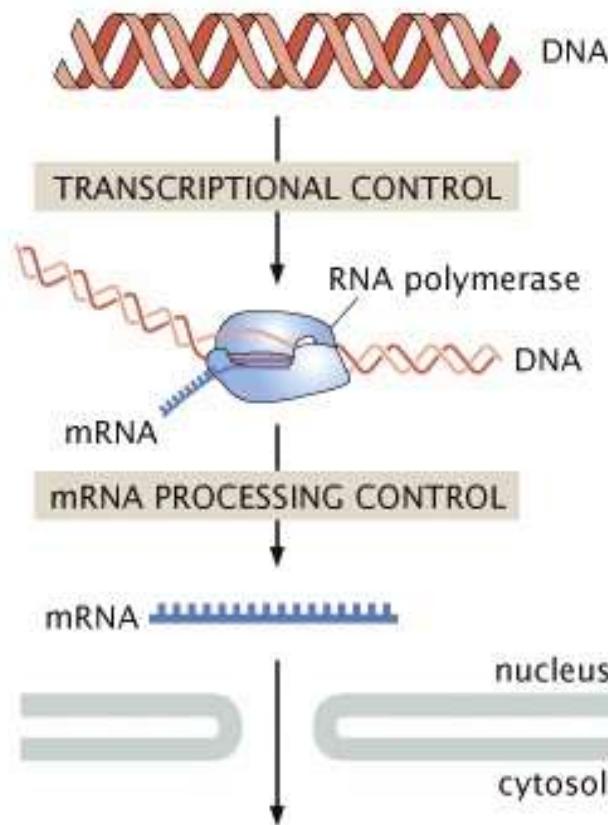


Figure 6.7 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

The λ-phage switch: the study of developmental “decisions” started here

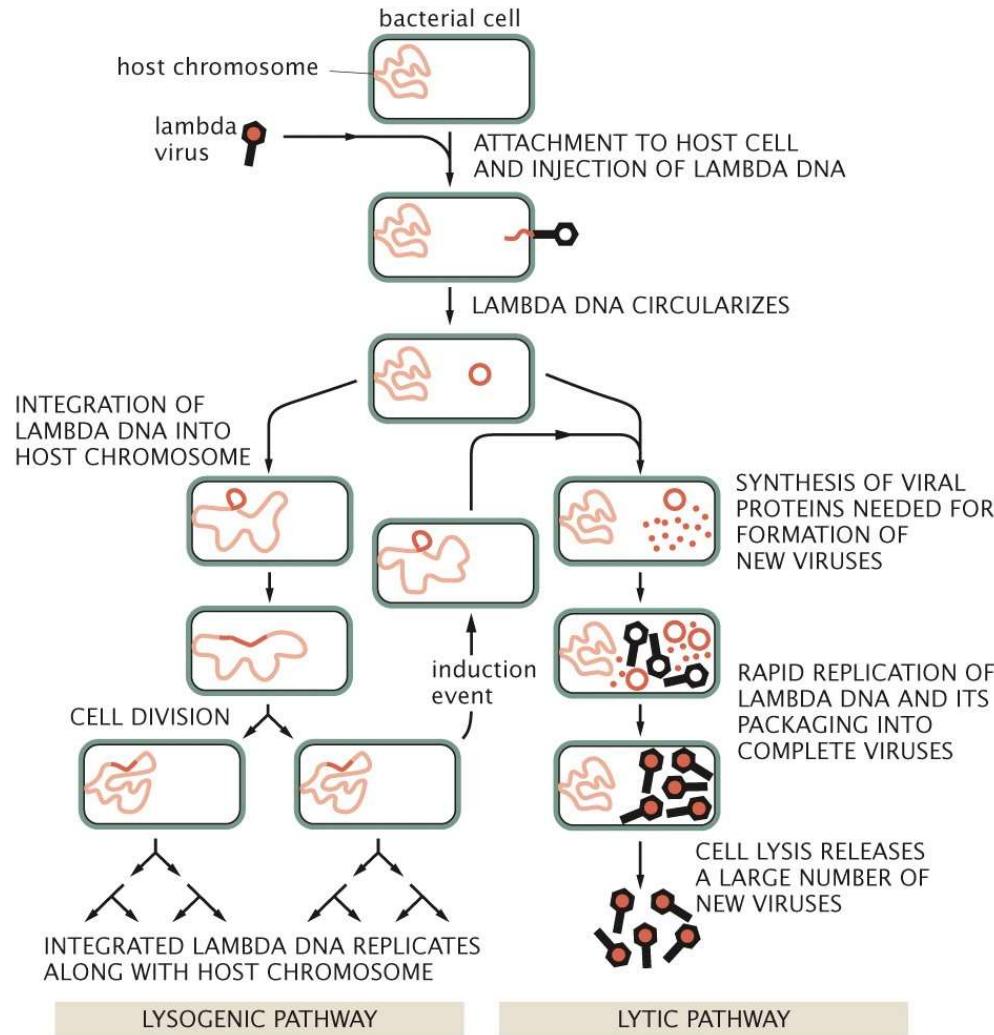


Figure 4.10 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

In simple cases, we can match the “regulation circuit” to the actual molecular processes, and hence have a handle on e.g. the rates.

If proteins have ~ 300 amino acid chains, then 900bp per protein, and around 4500 genes. This is ~ true, in *E.Coli*.

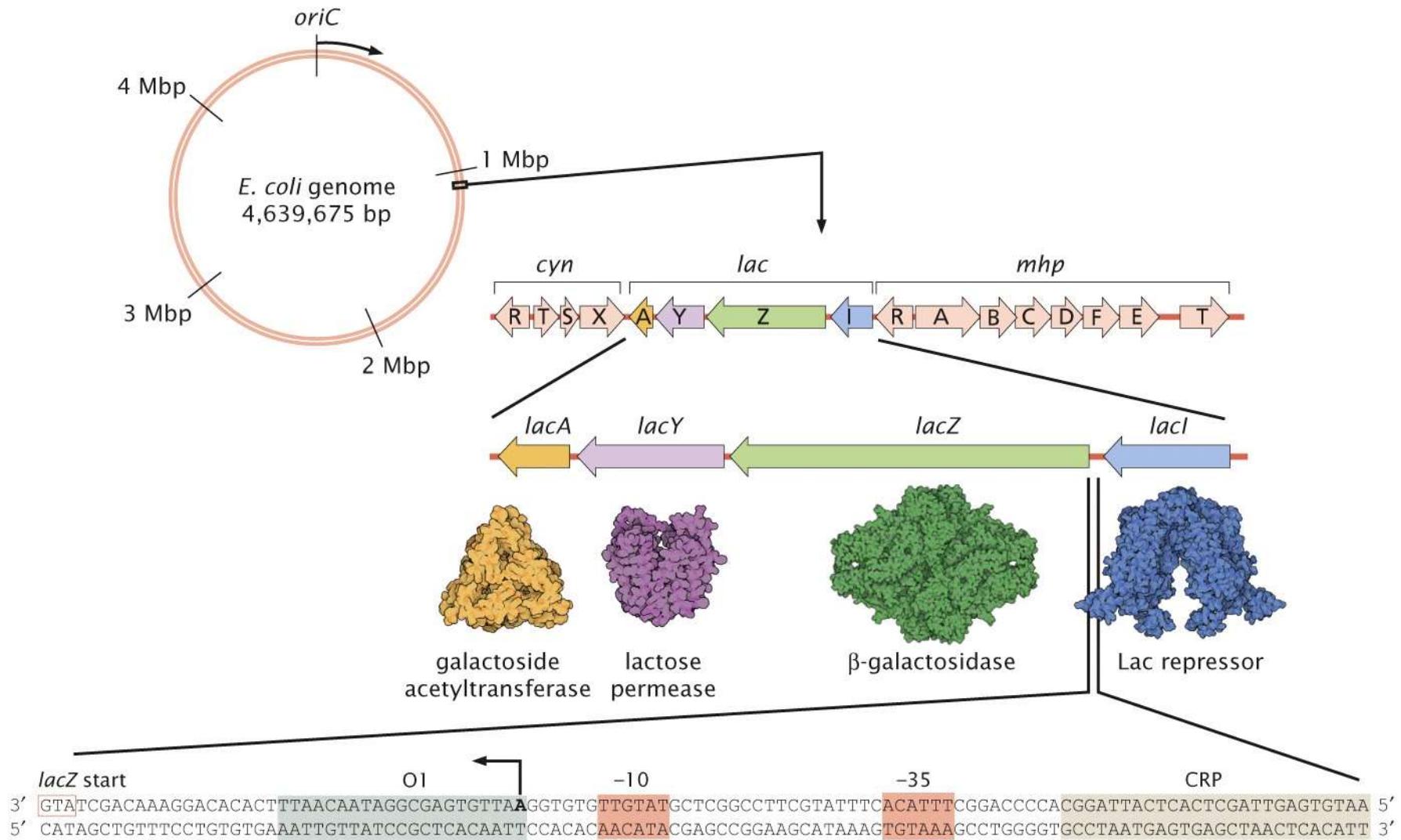


Figure 4.13 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

This is the lac operon: the study of gene regulation started here.

This is the lac operon: the study of gene regulation started here.

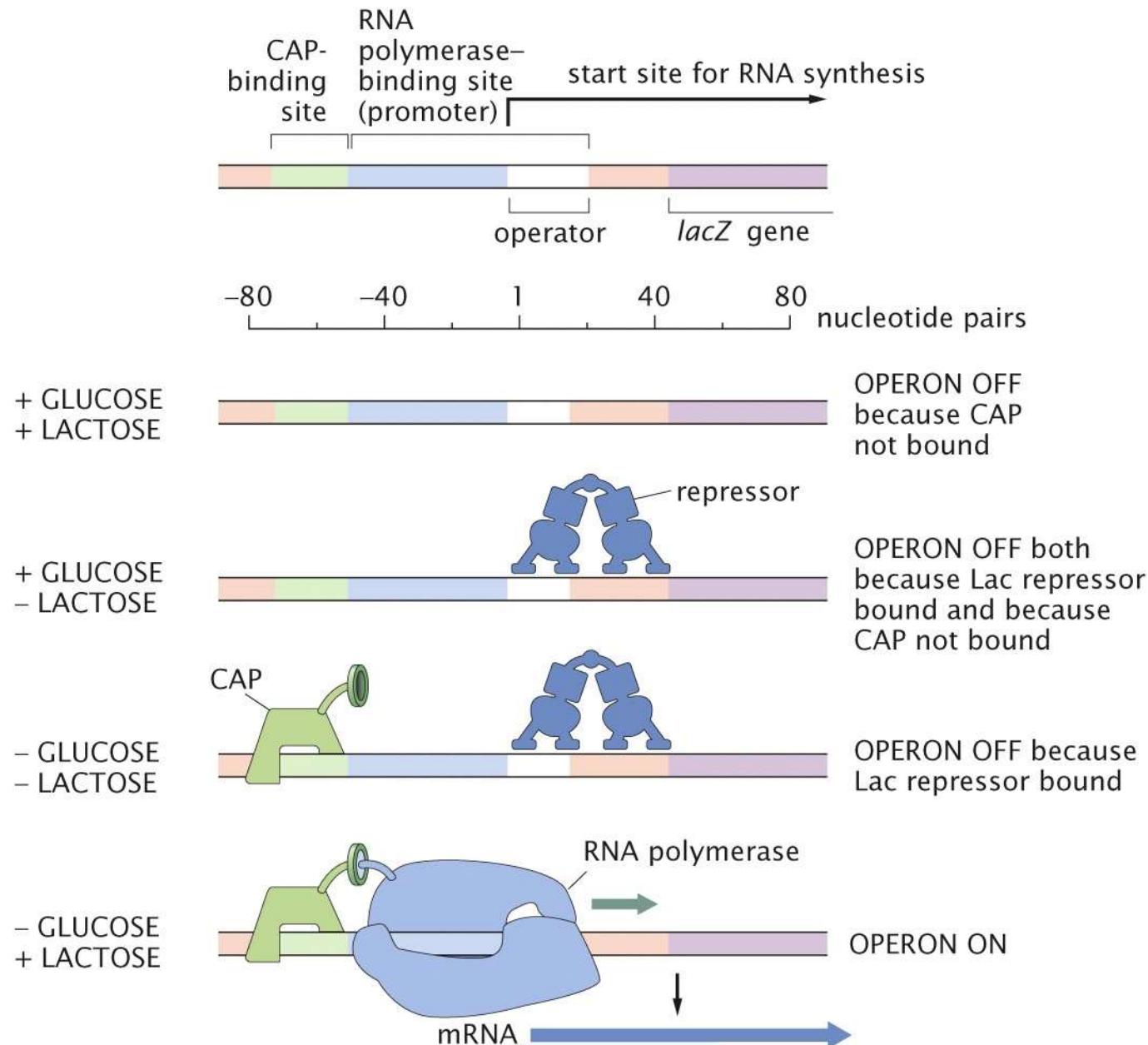
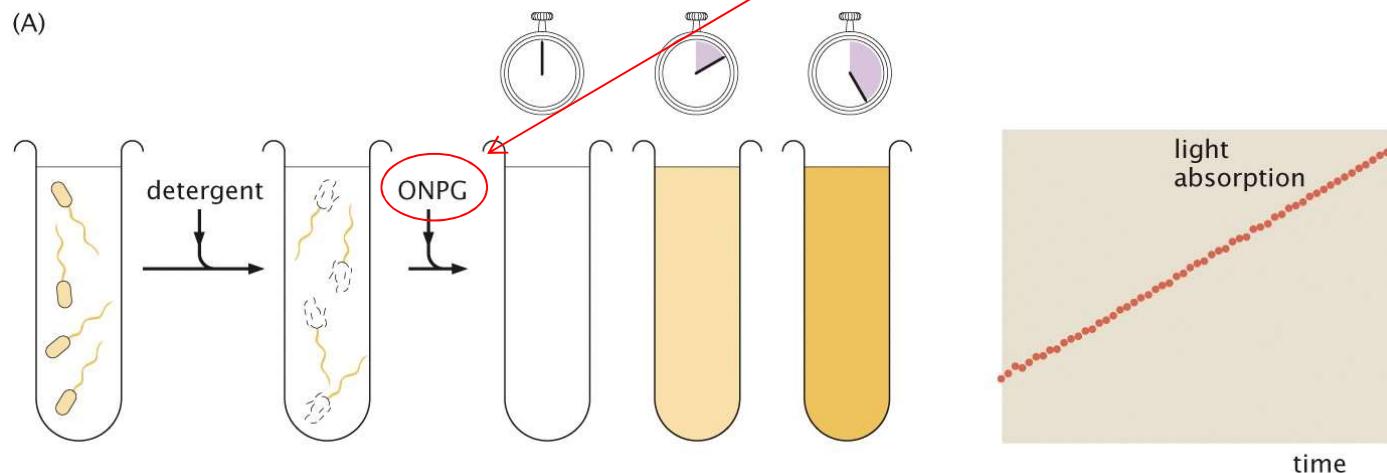


Figure 4.15 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

How do we measure gene expression?

O-nitrophenyl β -galactoside
It's colorless – can be cleaved by
 β -galactosidase to become yellow

(A) Bulk (population): measurement of enzyme activity



(B) Single cell: Make a construct where the promoter drives expression of a fluorescent protein, like GFP. Microscopy.

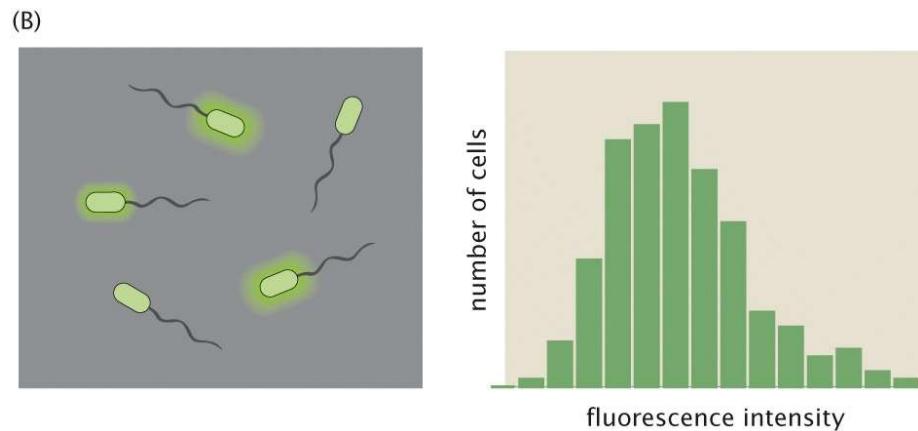


Figure 19.3 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

How do we measure mRNA levels? Microarray or PCR bulk techniques.

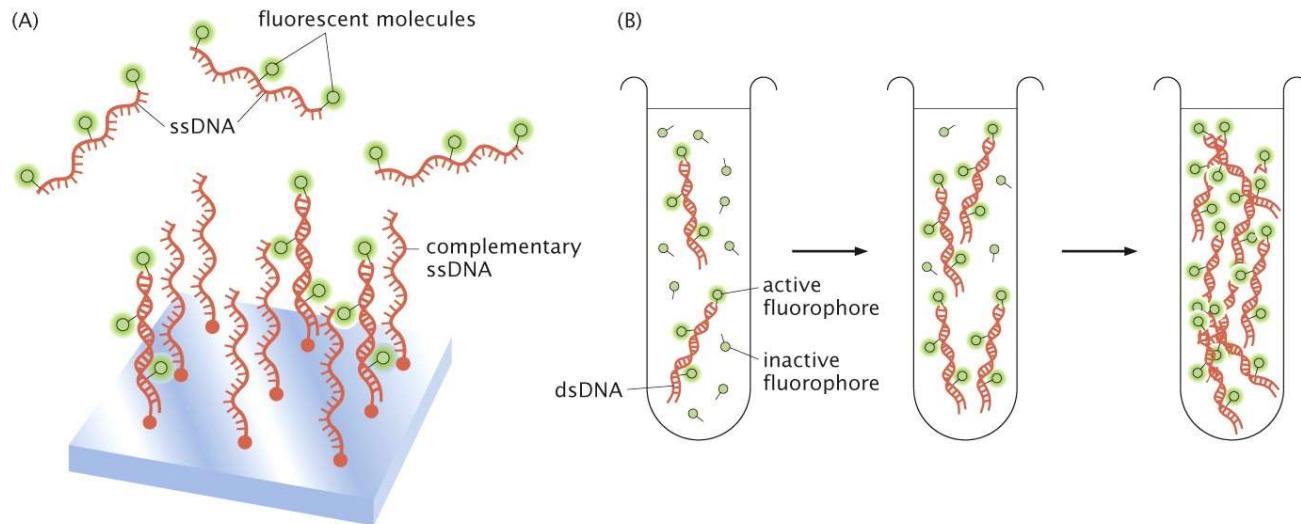
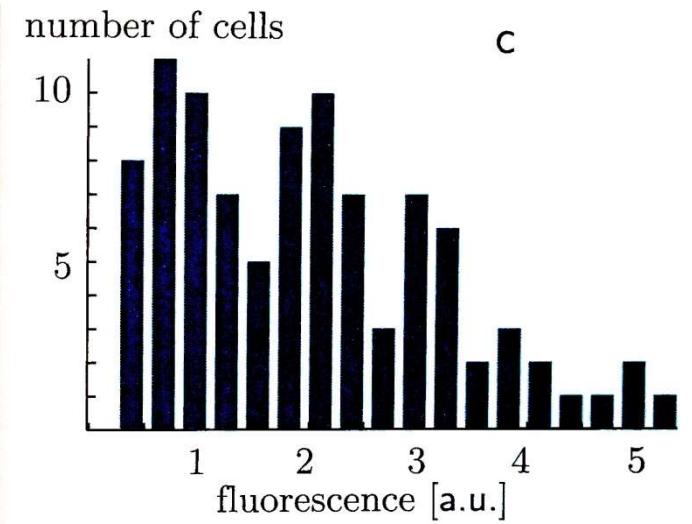
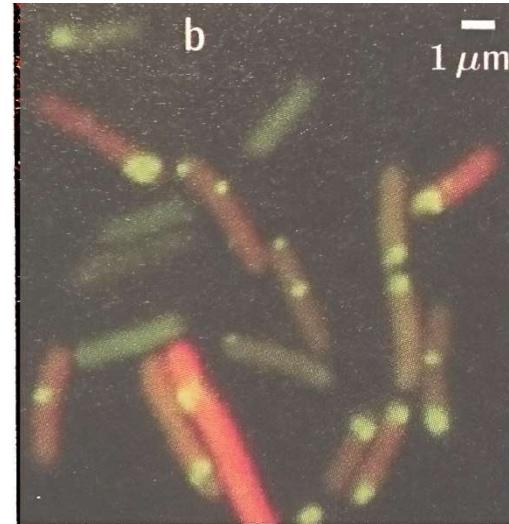
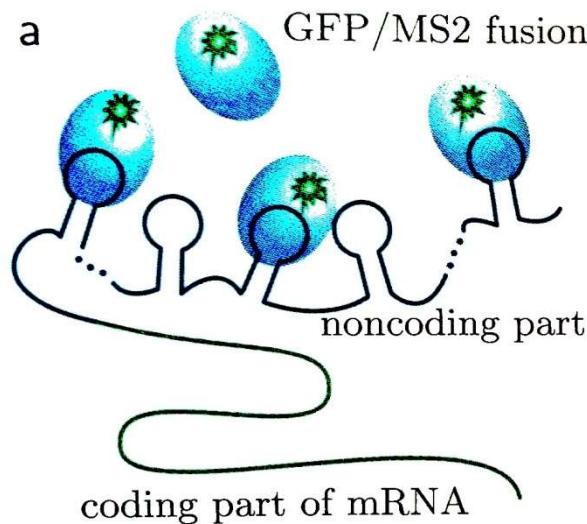


Figure 19.4 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Can also get mRNA levels at single cell, see p188-189 of Nelson



Timescales for various reactions in the transcription network of a bacterium

1ms binding of small molecules (signals) to transcription factors, thus changing the transcription factor activity

1s binding of transcription factor to its DNA site

5min transcription + translation of the gene

~1hr (i.e. 1 cell generation) 50% change in the concentration of the translated protein (assuming no degradation, i.e. for stable proteins)

Repressors hinder the binding of RNA polymerase, or the initiation of transcription.

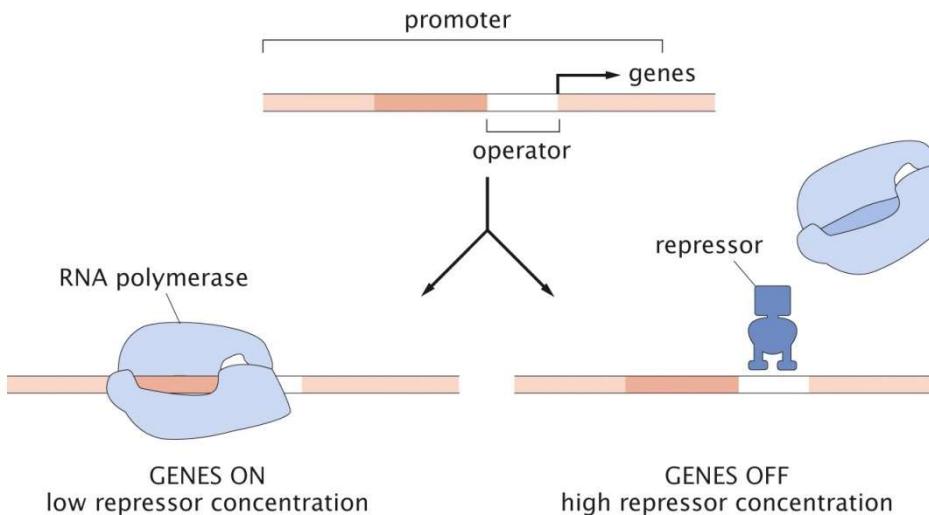


Figure 19.5 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

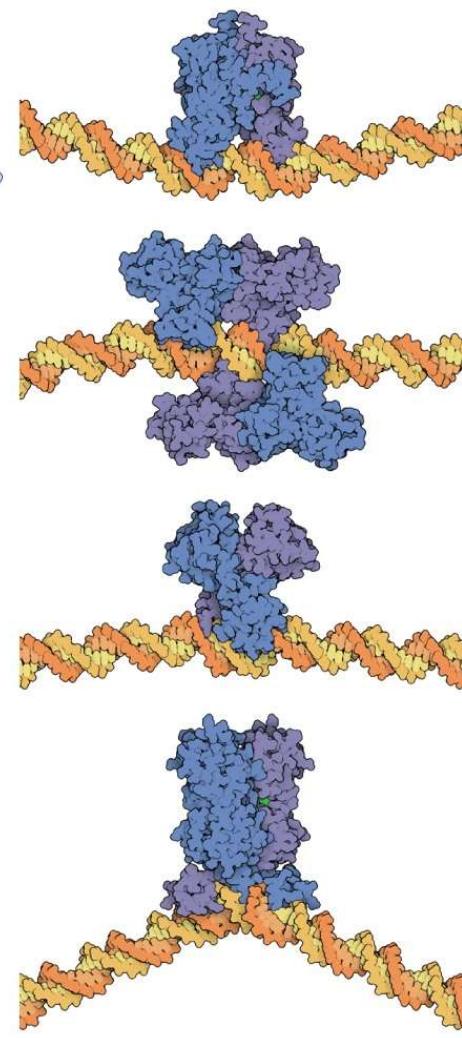


Figure 19.6 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Renditions, to scale with DNA,
4 different
repressors.

Obtained from
X-ray
crystallography.

Many such
structures
discovered @
LMB

Activators promote the binding of RNA polymerase.

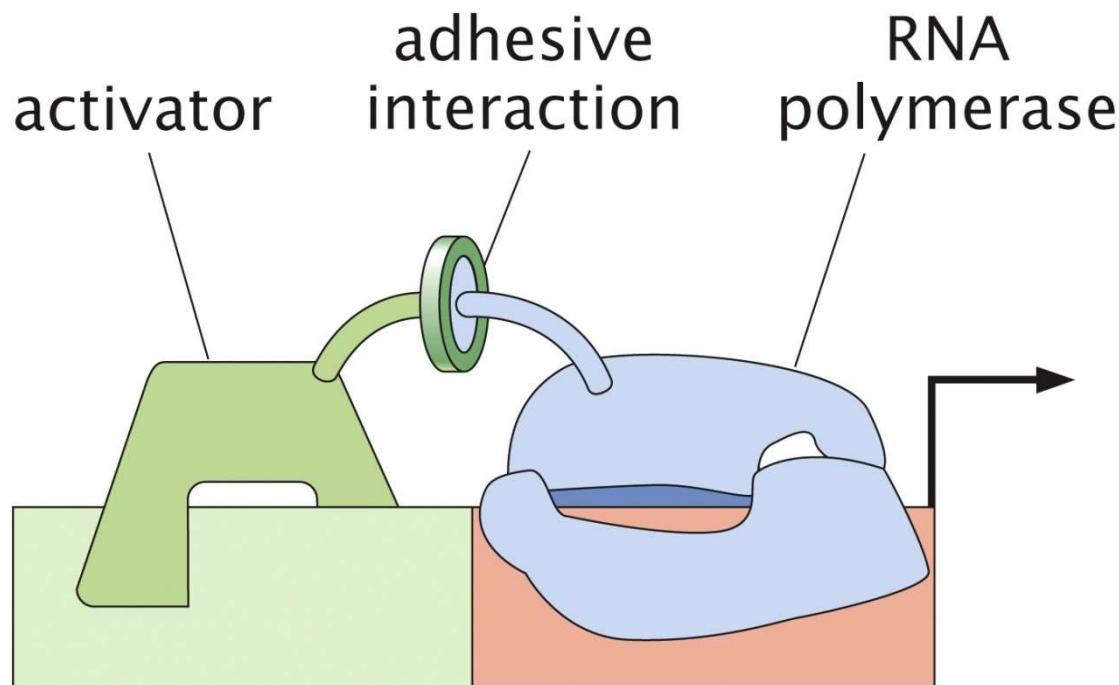


Figure 19.7 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

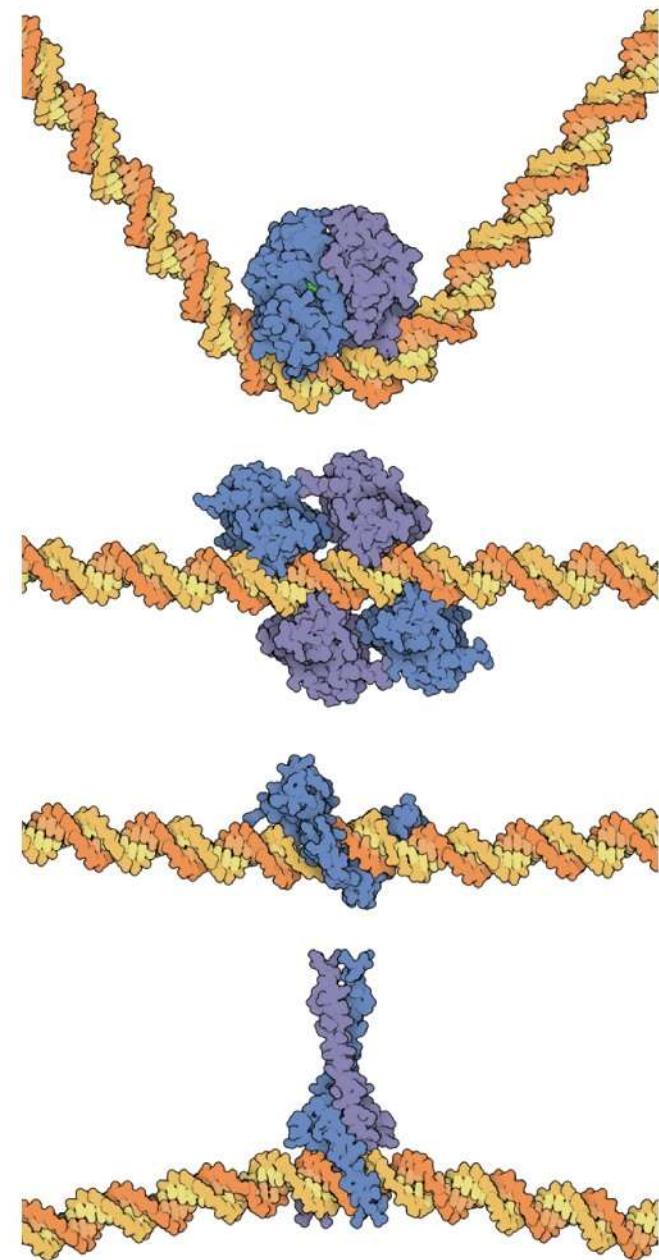


Figure 19.8 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Stat Mech can be used to describe quantitatively the activity of promoters.

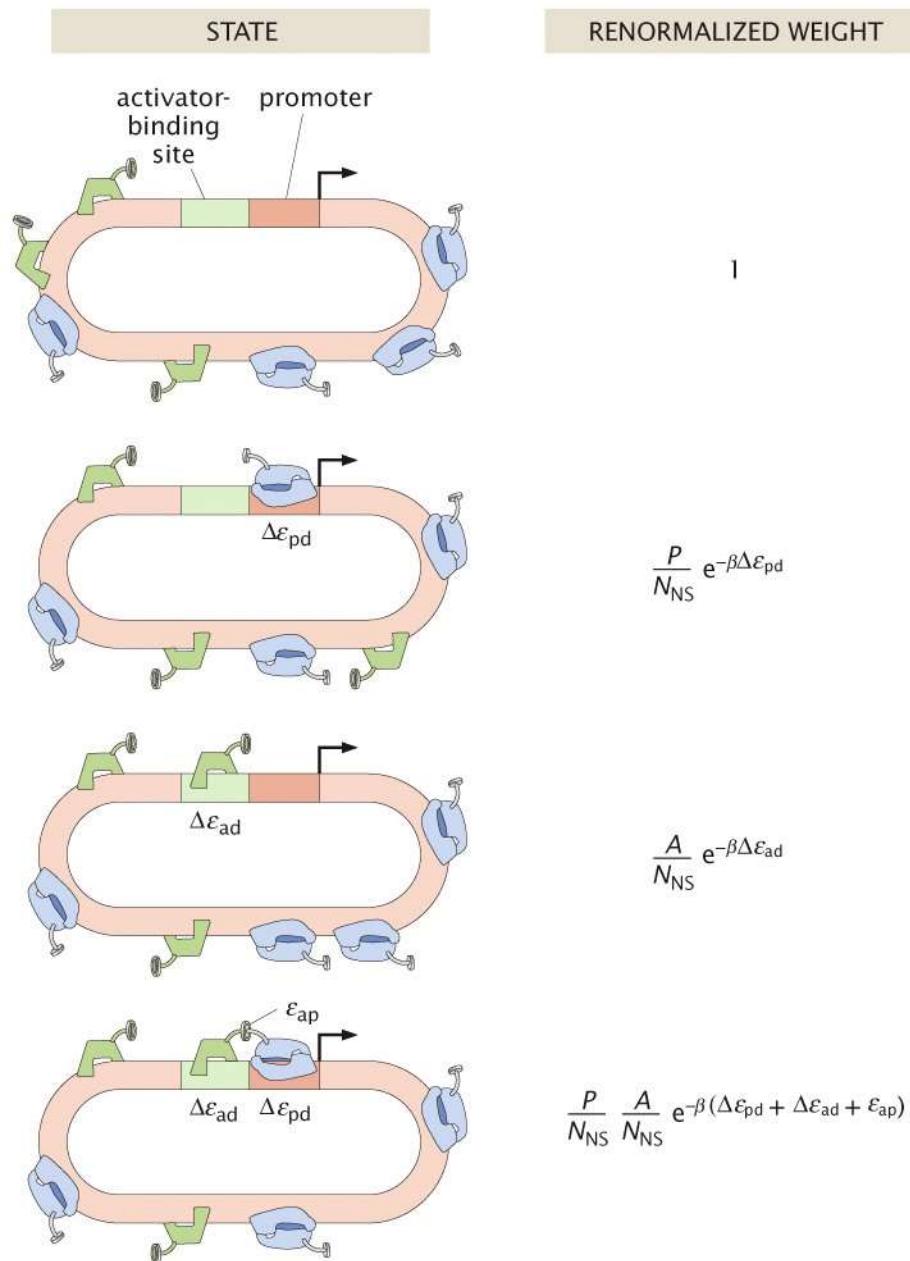


Figure 19.9 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Stat Mech - case of no regulation. [Full details see Phillips or pdf notes.]

RNA polymerase binding to a specific site: $p_{\text{bound}} = \frac{(c/c_0)e^{-\beta\Delta\epsilon}}{1 + (c/c_0)e^{-\beta\Delta\epsilon}}$

RNA polymerase binding: competition between specific and non specific site:

This extends the case above seen in lecture 5. Assume the non-specific sites on the DNA are N_{NS} ‘boxes’. Then the partition function associated with these states is:

$$Z_{NS}(P, N_{NS}) = \frac{N_{NS}!}{P!(N_{NS}-P)!} \times e^{-\beta P \epsilon_{pd}^{NS}}$$

where ϵ_{pd}^{NS} is the energy of binding the polymerase to a non-specific site (and ϵ_{pd}^S will be later the energy of binding the polymerase to the specific site).

Now we can write the total partition function. We need to sum over the states in which the promoter is occupied (hence $P-1$ polymerase molecules in the non-specific sites), and those where it is not:

$$Z(P, N_{NS}) = Z_{NS}(P, N_{NS}) + Z_{NS}(P-1, N_{NS})e^{-\beta \epsilon_{pd}^S}$$

The ratio of configuration weights where promoter is bound, to all weights, is:

$$p_{\text{bound}} = \frac{\frac{N_{NS}!}{(P-1)![N_{NS}-(P-1)]!} e^{-\beta \epsilon_{pd}^S} e^{-\beta(P-1)\epsilon_{pd}^{NS}}}{\frac{N_{NS}!}{P!(N_{NS}-P)!} e^{-\beta P \epsilon_{pd}^{NS}} + \frac{N_{NS}!}{(P-1)![N_{NS}-(P-1)]!} e^{-\beta \epsilon_{pd}^S} e^{-\beta(P-1)\epsilon_{pd}^{NS}}}$$

Stat Mech - case of no regulation (cont).

The factorials can be simplified for $N_{NS} \gg P$, and we can write the result to show that only the energy difference matters:

$$p_{bound} = \frac{1}{1 + \frac{N_{NS}}{P} e^{\beta \Delta \epsilon_{pd}}}$$

Familiar result for two-state models, with the unoccupied state of the promoter having weight =1, and the occupied having weight $P/N_{NS} \exp(-\beta \Delta \epsilon_{pd})$.

The energy differences $\Delta \epsilon_{pd}$ are negative, and can range between minus a few to $\sim -10 k_B T$.

Stat Mech - case of activation.

Activators are proteins that bind to a specific site, and promote the recruitment of RNA polymerase to a nearby promoter site. We now have 4 classes of outcome to sum over to make the total partition function: the activator and promoter site can each be occupied or unoccupied. So:

$$\begin{aligned} Z_{tot}(P, A, N_{NS}) &= Z(P, A, N_{NS}) \text{ (empty)} \\ &+ Z(P-1, A, N_{NS}) e^{-\beta \epsilon_{pd}^S} \text{ (only RNAP on promoter)} \\ &+ Z(P, A-1, N_{NS}) e^{-\beta \epsilon_{ad}^S} \text{ (only activator bound)} \\ &+ Z(P-1, A-1, N_{NS}) e^{-\beta(\epsilon_{pd}^S + \epsilon_{ad}^S + \epsilon_{pa})}. \text{ (both RNAP and activator bound)} \end{aligned}$$

Notation: A, a are the activator, P, p the polymerase, d the DNA). ϵ_{pa} is the energy that favours the activator and the RNA polymerase being close.

Stat Mech - case of activation (cont).

Algebra more lengthy but follows the exact steps as previously.

To get promoter occupancy, we can take the ratios of the weights of the two ‘favorable’ states, against the sum of all weights, and we get:

$$p_{\text{bound}}(P, A, N_{NS}) = \frac{1}{1 + \frac{N_{NS}}{P F_{\text{reg}}(A)} e^{\beta \Delta \epsilon_{pd}}}$$

With $\Delta \epsilon$ the energy differences between specifically and nonspecifically bound conditions and

$$F_{\text{reg}}(A) = \frac{1 + (A/N_{NS})e^{-\beta \Delta \epsilon_{ad}}e^{-\beta \epsilon_{ap}}}{1 + (A/N_{NS})e^{-\beta \Delta \epsilon_{ad}}}$$

Neat result! Shows that activating molecules make an $F > 1$, i.e. have an effect that is mathematically equivalent to increasing the number of polymerases.

Given realistic values of the other energies, a few $-k_B T$ for ϵ_{ap} is enough to significantly change the bound probability.

If the approx $(N_{NS}/P F_{\text{reg}})\exp(\beta \Delta \epsilon_{pd}) \gg 1$ holds, i.e. the promoter is not too strong, then you can obtain (exercise) that the fold increase is approximately $F_{\text{reg}}(A)$ itself.

Details of the promoter (its binding energy) factor out of the problem!

Activation. The stat mech model shows that the presence of activators is equivalent to having a higher concentration of polymerase.

(A)



(B)

$$F_{\text{reg}} = 2$$

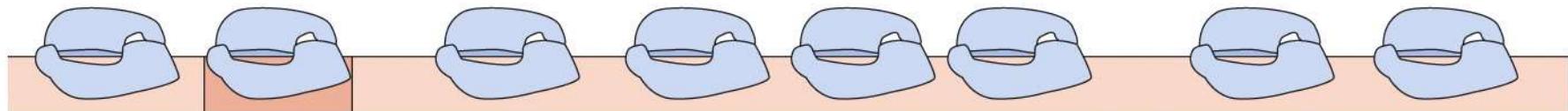


Figure 19.11 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Activation. The probability of binding, which is a good proxy for the rate of expression, depends strongly ('fold increase') on the polymerase-activator binding.

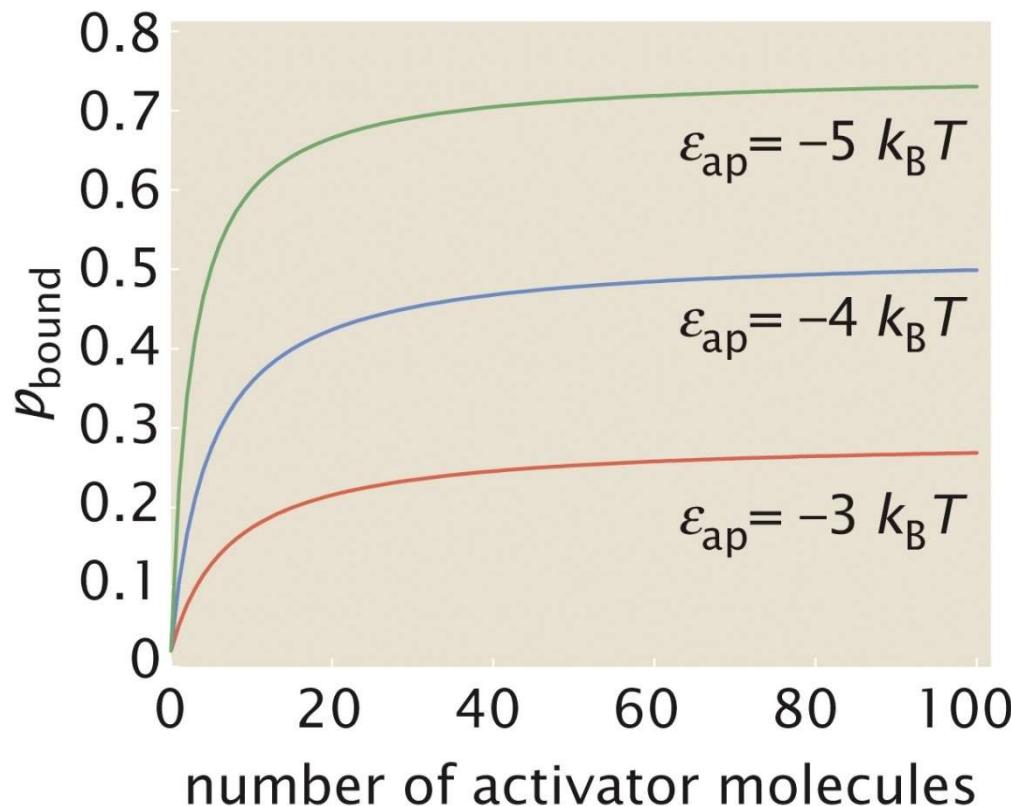


Figure 19.10 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

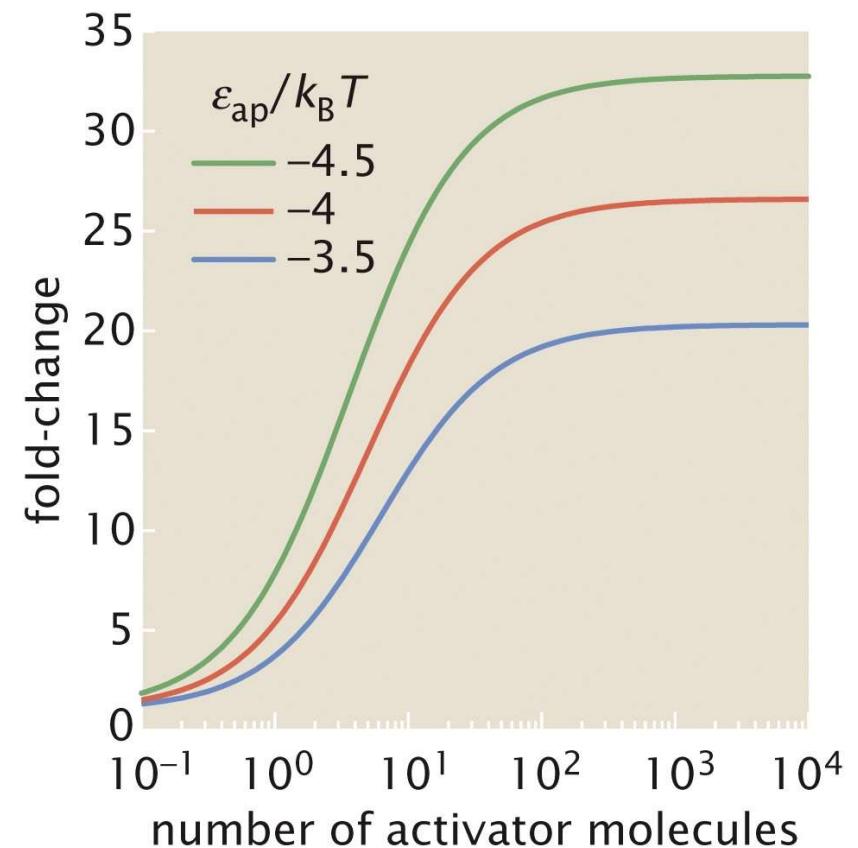


Figure 19.12 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

In plots like this, all other parameters are fixed. Here:

$$P=500, \Delta \epsilon_{pd} = -5.3 k_B T, \Delta \epsilon_{ad} = -13.12 k_B T.$$

Clever experiments have been carried out to specifically target aspects of the activator-polymerase interaction, which correspond directly to how we write the stat mech models.

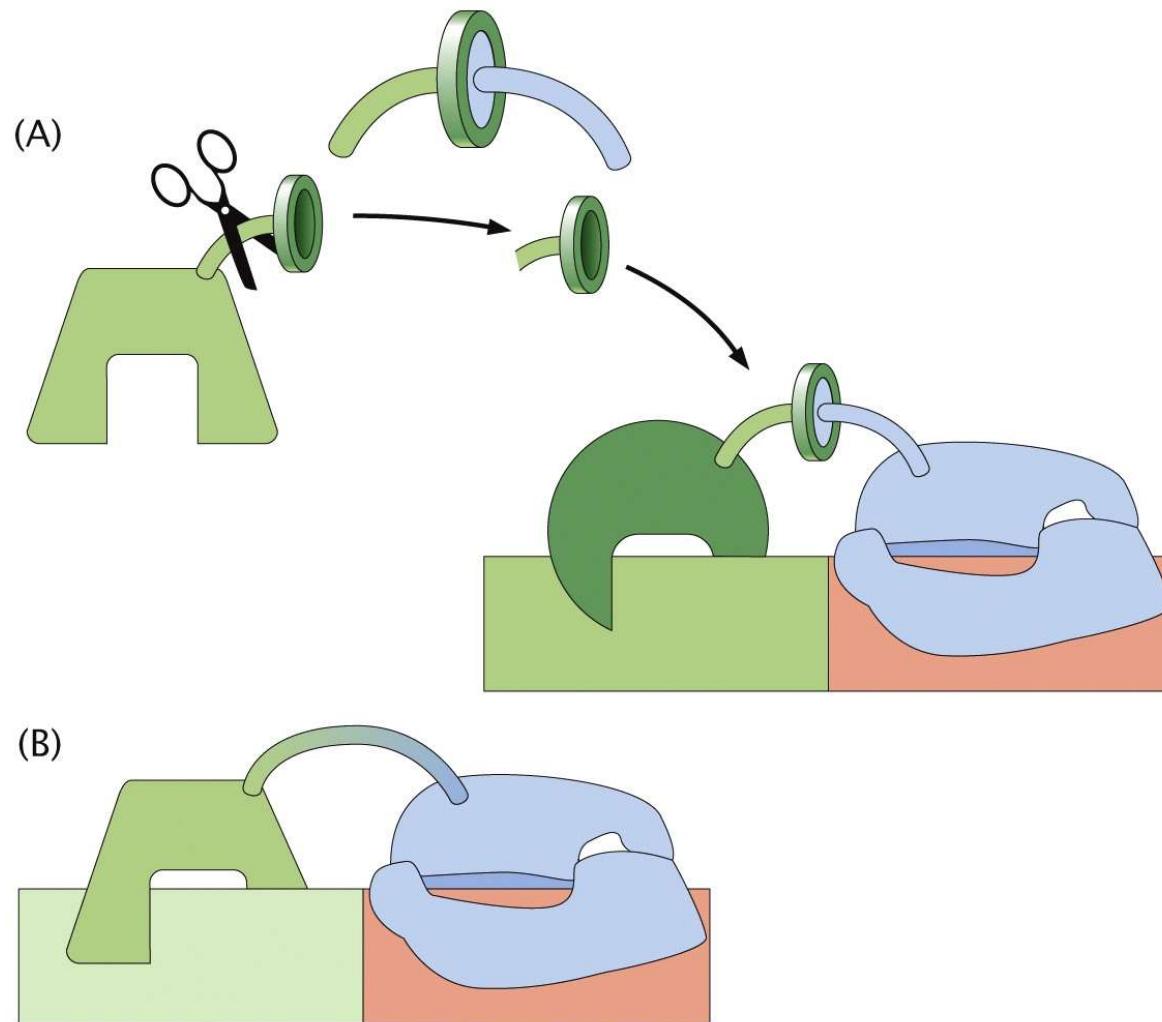


Figure 19.13 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Stat Mech - case of repressor.

Repressor proteins occupy the promoter region, and prevent the RNAP binding there. Physics model is another variant of previous partition function machinery

$$\begin{aligned} Z_{tot}(P, R, N_{NS}) &= Z(P, R, N_{NS}) \text{ (empty promoter)} \\ &+ Z(P - 1, R, N_{NS})e^{-\beta\epsilon_{pd}^S} \text{ (RNAP on promoter)} \\ &+ Z(P, R - 1, N_{NS})e^{-\beta\epsilon_{rd}^S}. \text{ (repressor on promoter)} \end{aligned}$$

We get $p_{bound}(P, R, N_{NS}) = \frac{1}{1 + \frac{N_{NS}}{P}e^{\beta(\epsilon_{pd}^S - \epsilon_{pd}^{NS})}[1 + \frac{R}{N_{NS}}e^{-\beta(\epsilon_{rd}^S - \epsilon_{rd}^{NS})}]}$

And further simplified as done for the activator, but this time

$$F_{reg}(R) = \left(1 + \frac{R}{N_{NS}}e^{-\beta\Delta\epsilon_{rd}}\right)^{-1}$$

with $\Delta\epsilon_{rd} = \epsilon_{rd}^S - \epsilon_{rd}^{NS}$. Here, $F_{reg} < 1$, which means that the system behaves as if fewer polymerases were present.

Repression.

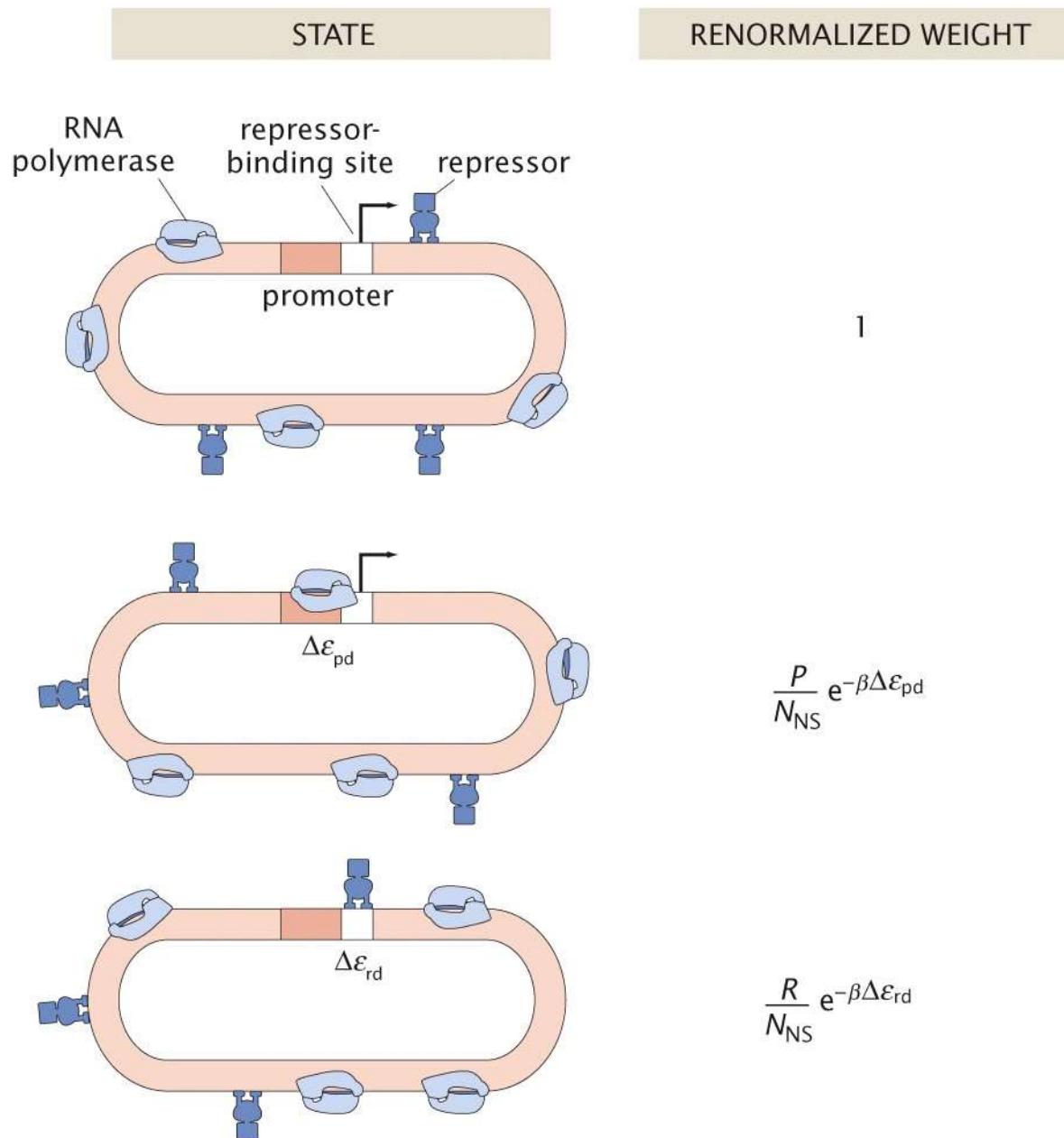


Figure 19.14 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Measurement of fold-change in repression.

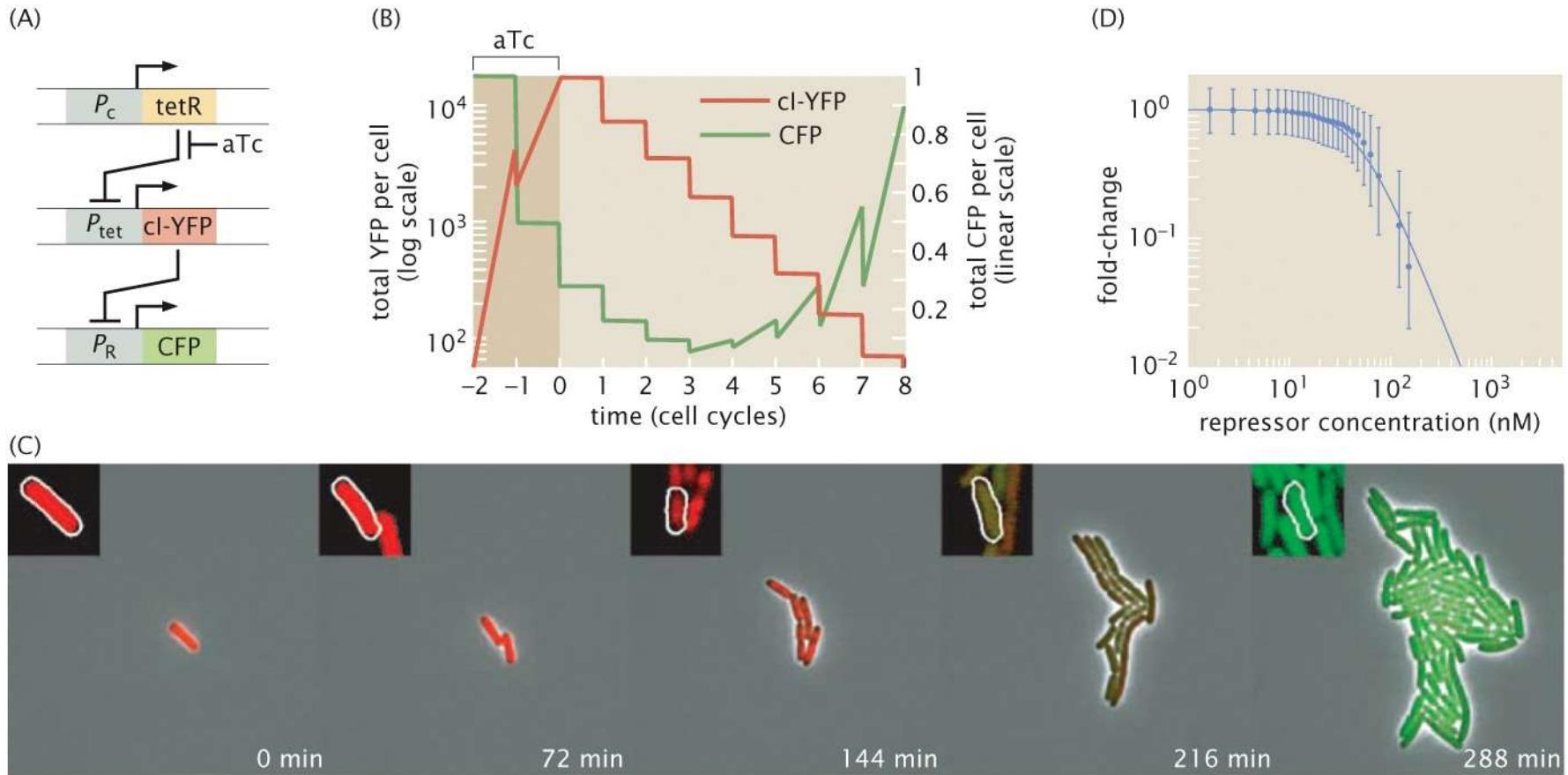


Figure 19.15 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Segmentation and single cell quantification of fluorescence

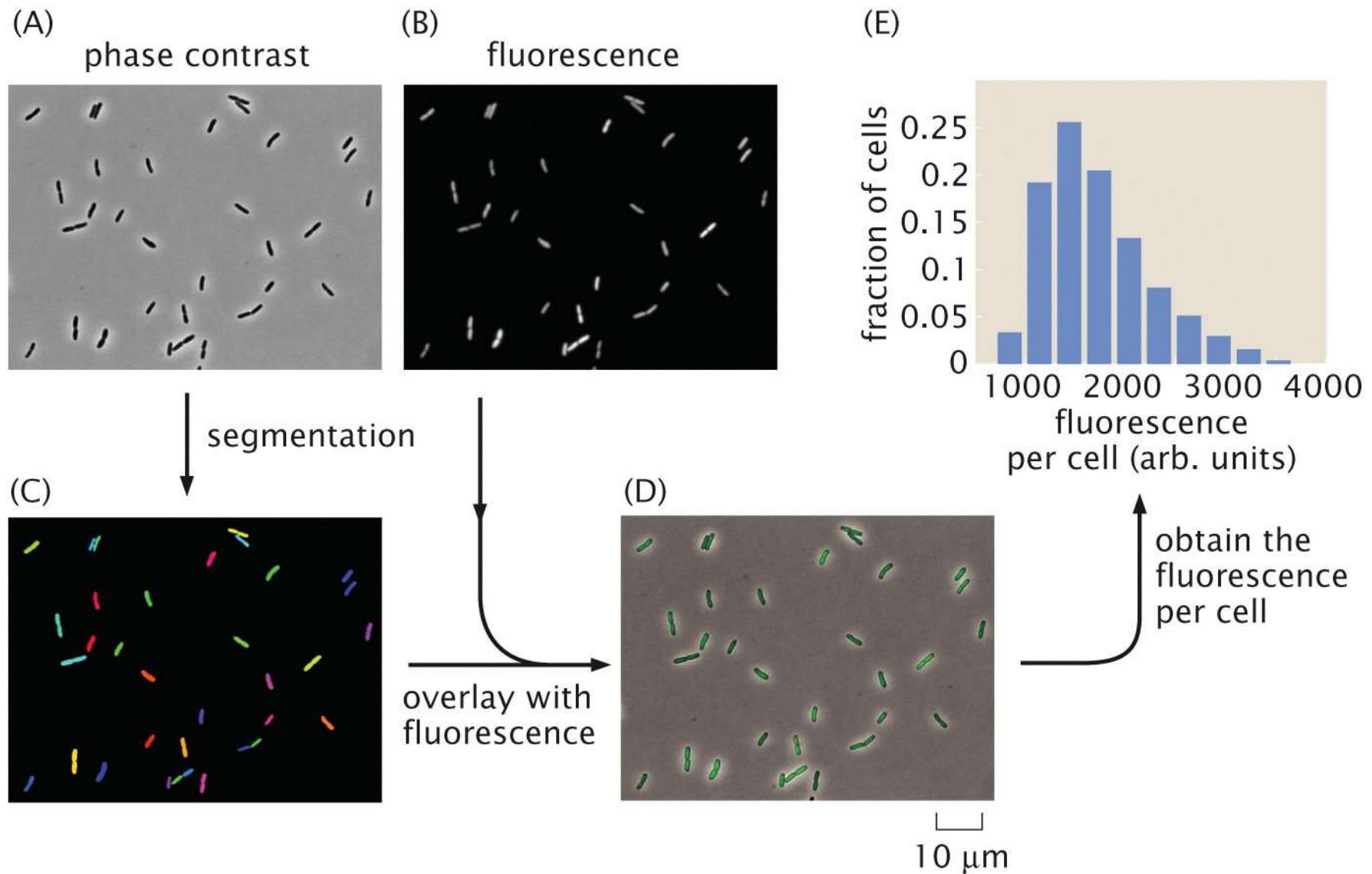


Figure 19.16 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Taking care (same settings, removing autofluorescence background, etc) it is then possible to quantitatively compare across different conditions: a measurement of fold-change.

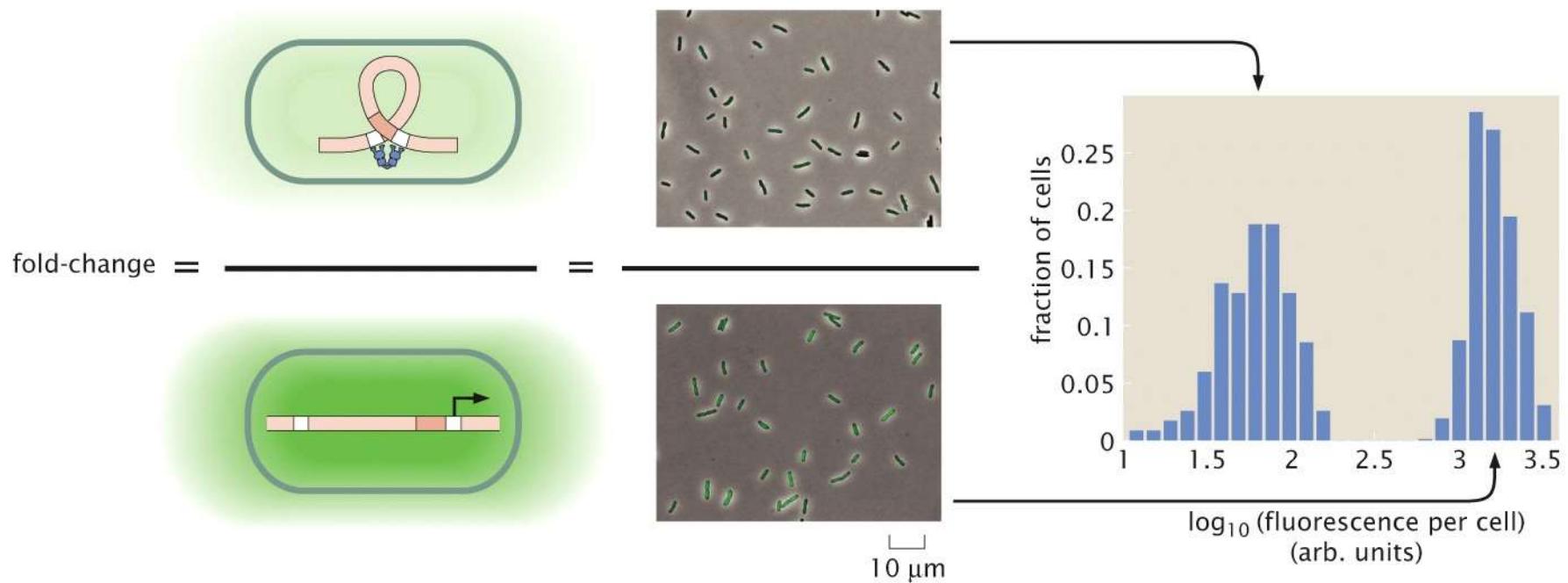


Figure 19.17 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Stat Mech - case of activator and repressor together.

In a real regulatory mechanism there will be both activation and repression, in competition. This tests our algebra skills... there are 6 cases to sum over:

$$\begin{aligned}
 Z_{tot}(P, A, R, N_{NS}) &= Z(P, A, R, N_{NS}) \text{ (empty promoter)} \\
 + Z(P-1, A, R, N_{NS})e^{-\beta\epsilon_{pd}^S} &\text{ (RNAP on promoter)} \\
 + Z(P, A-1, R, N_{NS})e^{-\beta\epsilon_{ad}^S} &\text{ (activator on promoter)} \\
 + Z(P-1, A-1, R, N_{NS})e^{-\beta(\epsilon_{ad}^S + \epsilon_{pd}^S + \epsilon_{pa})} &\text{ (RNAP and activator on)} \\
 + Z(P, A, R-1, N_{NS})e^{-\beta\epsilon_{rd}^S} &\text{ (repressor on promoter)} \\
 + Z(P, A-1, R-1, N_{NS})e^{-\beta(\epsilon_{ad}^S + \epsilon_{rd}^S)} &\text{. (activator and repressor on)}
 \end{aligned}$$

We can follow same steps, with patience, and get:

$$p_{bound}(P, A, R, N_{NS}) = \frac{1}{1 + \frac{N_{NS}}{P F_{reg}(A, R)} e^{\beta(\epsilon_{pd}^S - \epsilon_{pd}^{NS})}}$$

With now a very rich regulation function:

$$\begin{aligned}
 F_{reg}(A, R) = & \left[1 + (A/N_{NS})e^{-\beta(\Delta\epsilon_{ad} + \epsilon_{ap})} \right] / \\
 & \left[1 + (A/N_{NS})e^{-\beta\Delta\epsilon_{ad}} + (R/N_{NS})e^{-\beta\Delta\epsilon_{rd}} + \right. \\
 & \left. (A/N_{NS})(R/N_{NS})e^{-\beta(\Delta\epsilon_{ad} + \Delta\epsilon_{pd})} \right].
 \end{aligned}$$

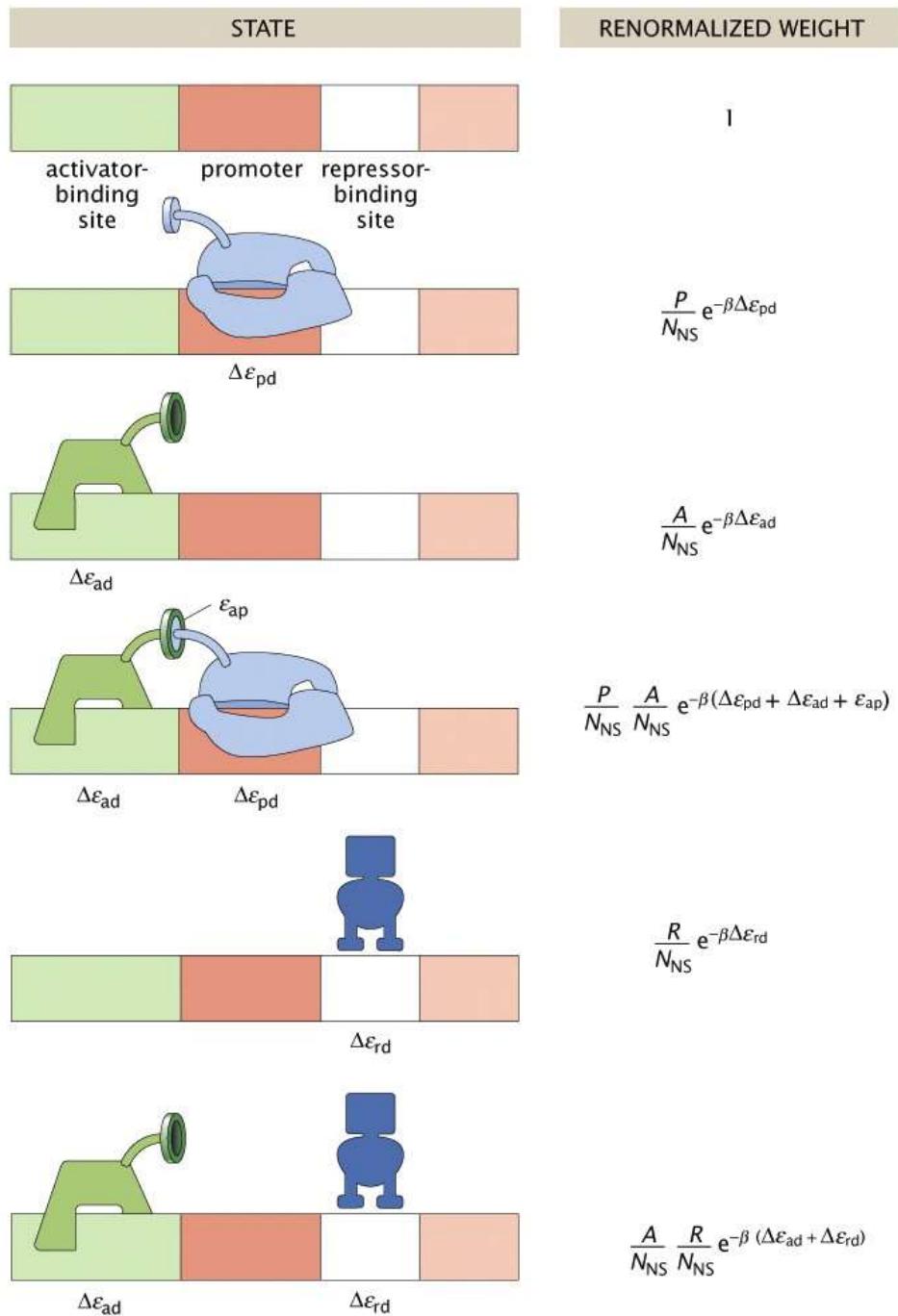
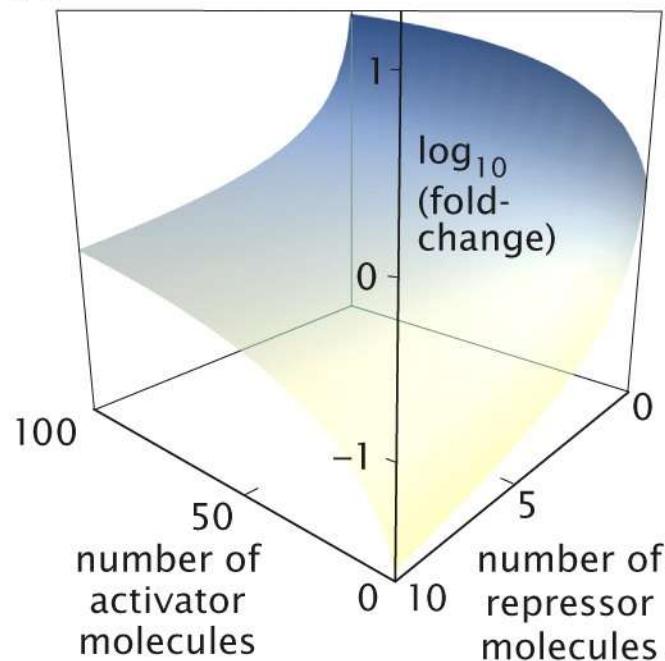


Figure 19.18 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Model and experiment with both positive and negative regulation

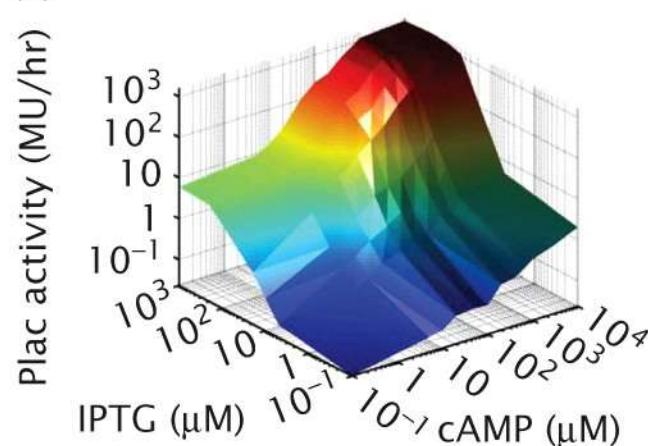
(A)



$$\begin{aligned}\Delta\epsilon_{ad} &= -10k_B T, \\ \Delta\epsilon_{ap} &= -3.9k_B T, \\ \Delta\epsilon_{rd} &= -16.9k_B T.\end{aligned}$$

This is lac operon experiment, from 2007 PNAS paper.

(B)



Rather than change number of activator and repressor molecules in a cell, it is easier to modulate their binding to DNA. IPTG blocks the repressor from binding, cAMP promotes the activator CAP binding.

Figure 19.19 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

What is lac operon ?

The lac Operon is a system that allows bacteria to react to the best sugar present, specifically to ensure that the enzymes to digest lactose are produced only when glucose is **not** present, **and** lactose is present. It seems an apparent simple objective, but selecting reliably for one of four situations requires a mechanism of both activation and repression as outlined here.

This regulatory system has played a key role historically in understanding physical and biological aspects of gene regulation, and is still subject of quantitative investigations and modelling today.

In the lac Operon there is an activator, the protein CAP.

In order to recruit RNAP, CAP has to be bound to a molecule called cyclic AMP (cAMP), whose concentration goes up when amount of glucose decreases.

There is also a repressor, the Lac repressor, which decreases the amount of transcription unless it is abound to allolactose, a byproduct of lactose metabolism. IPTG is a “synthetic inducer”, also binding to the repressor.

Zooming in on *lac* operon

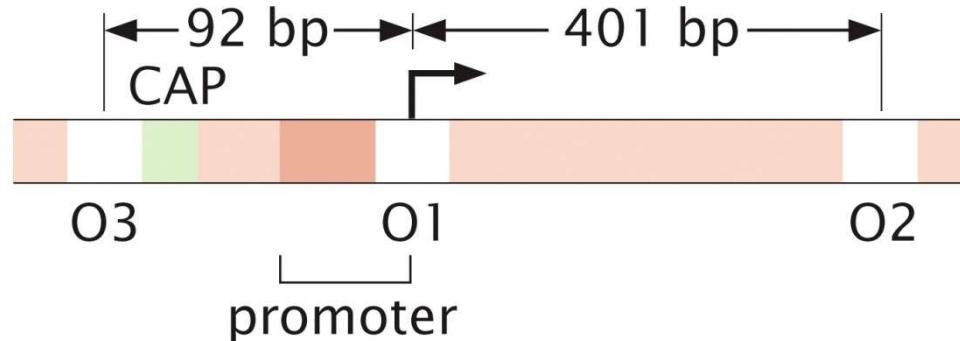


Figure 19.21 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Manipulating the *lac* operon to dissect repression effects due to affinity

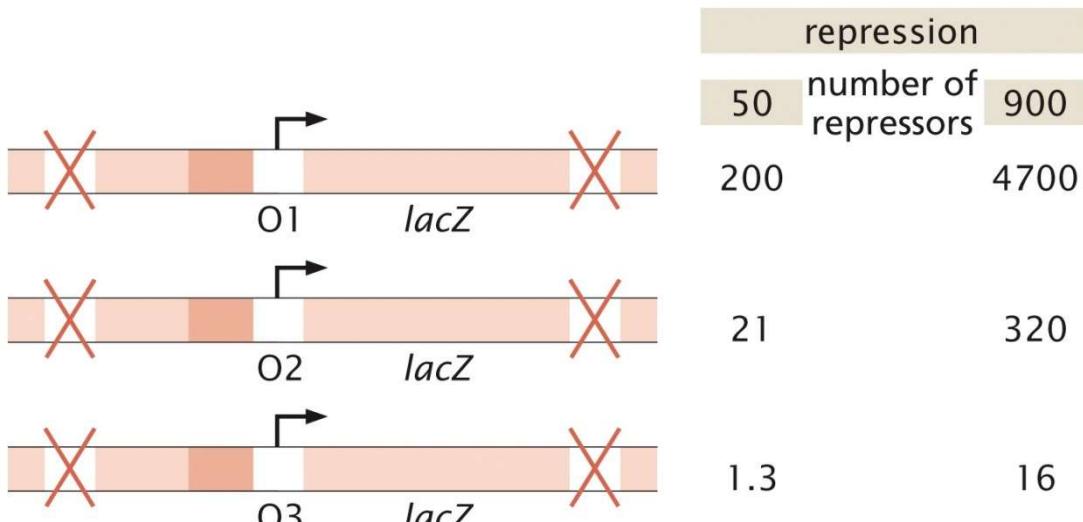


Figure 19.22 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Fitting with the simple repression model, we obtain -16.9, -14.4, -11.2 kT for O1, O2, O3

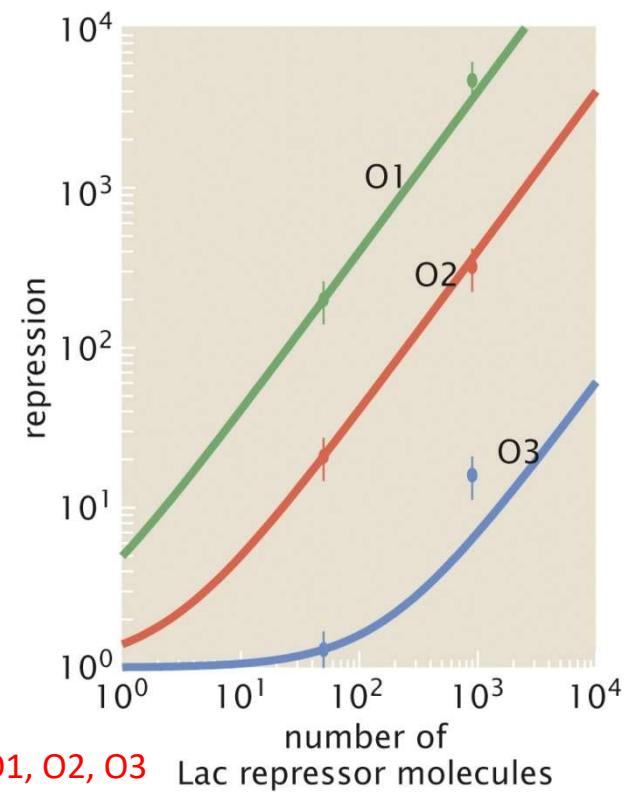


Figure 19.23 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

In the *lac* operon DNA looping also important

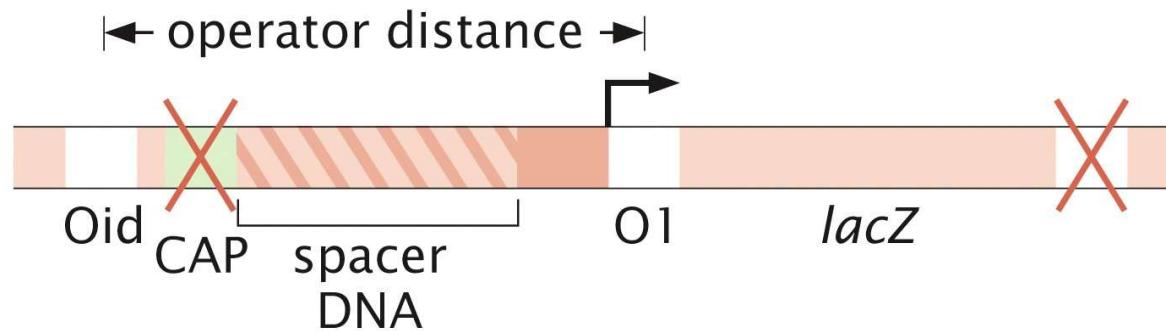


Figure 19.24 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

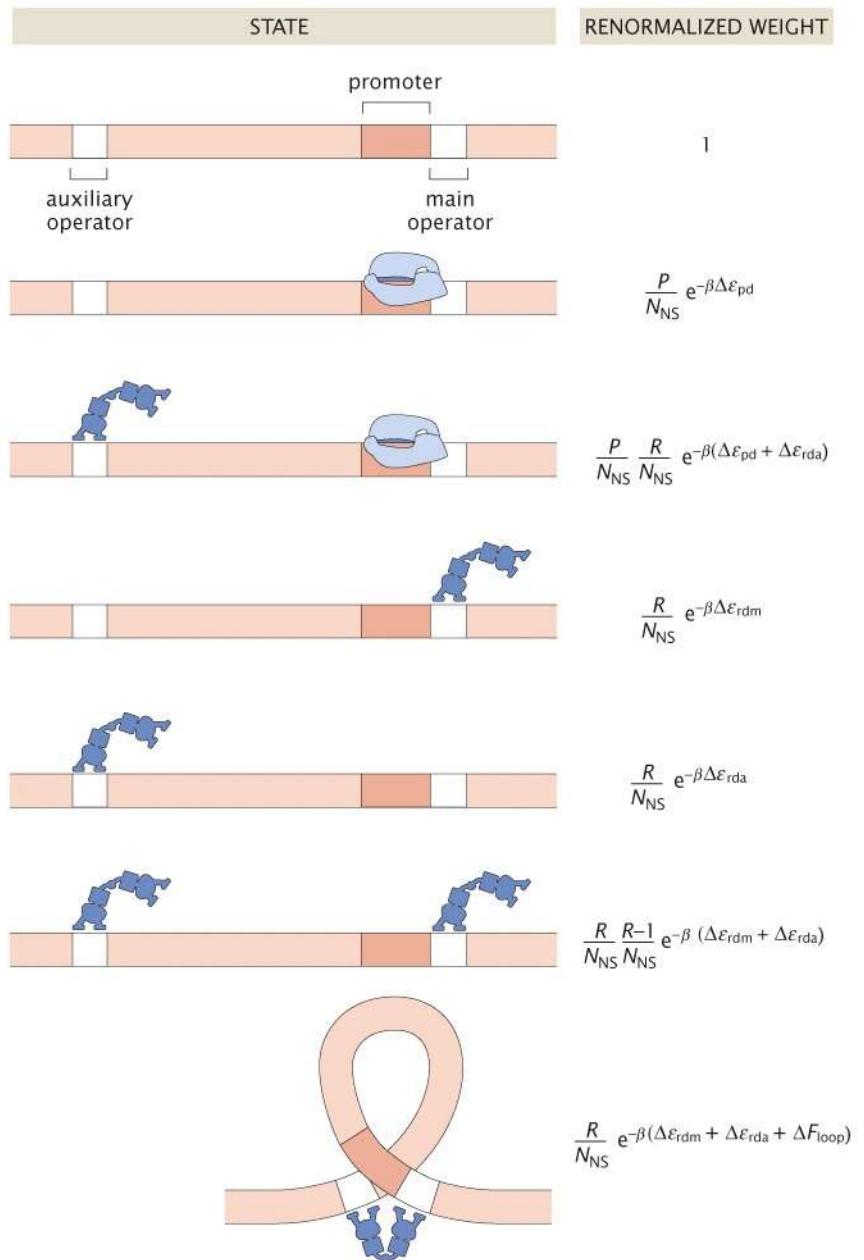


Figure 19.25 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

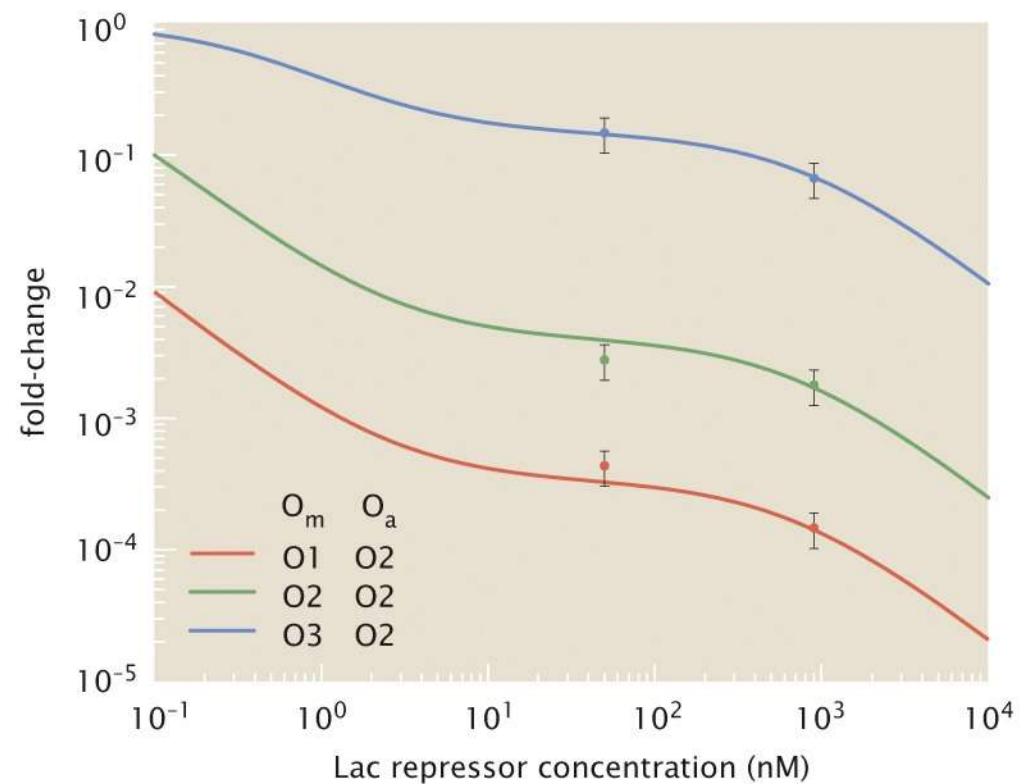
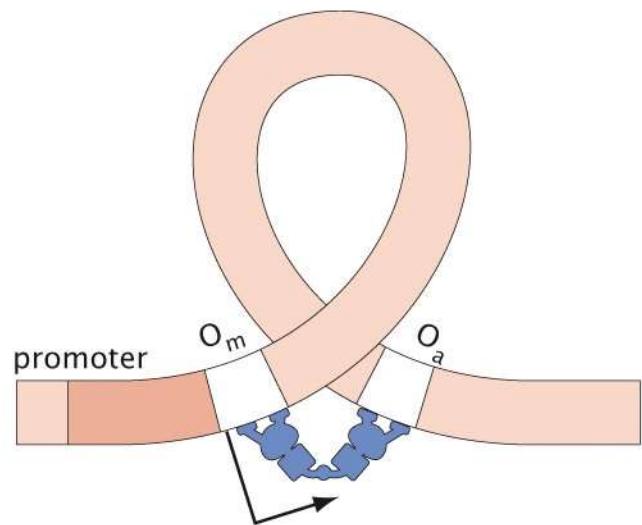


Figure 19.27 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

All parameters are known except the looping free energy which is obtained by these fits.

A summary of regulatory architectures presented so far

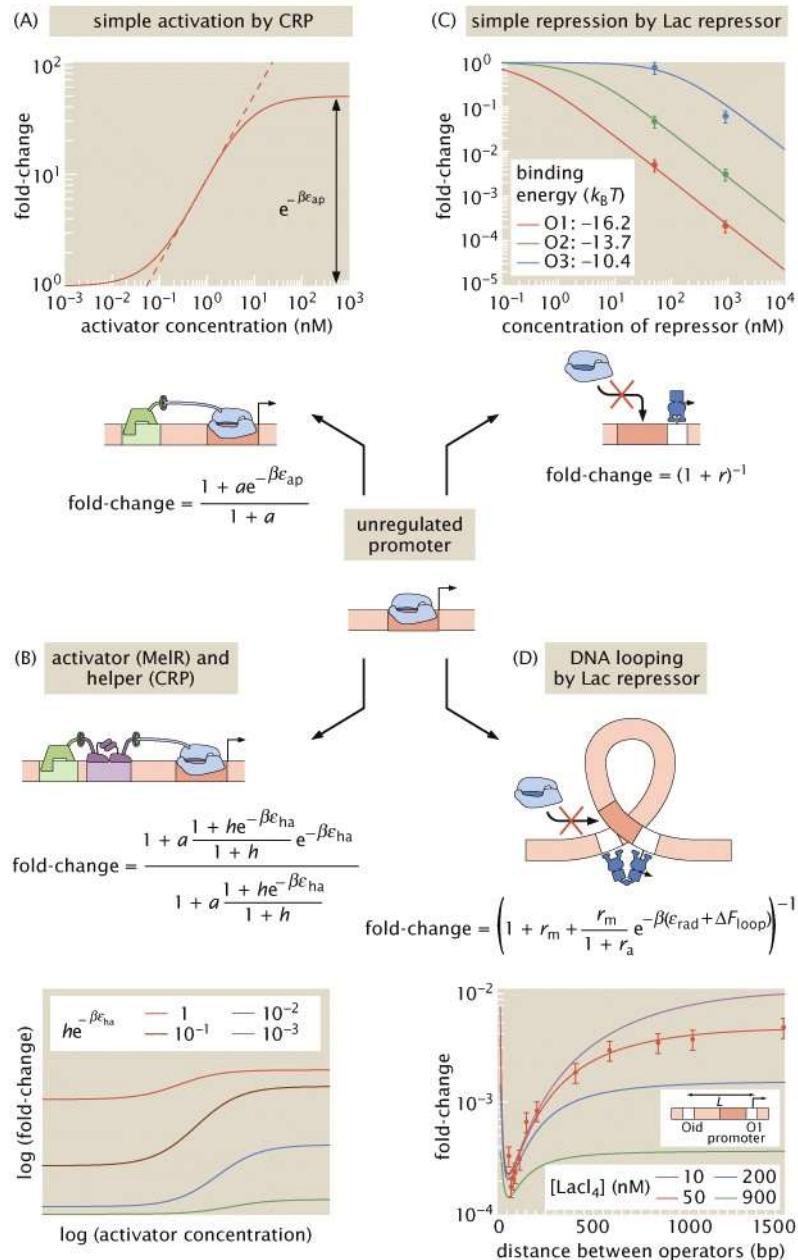


Figure 19.29 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

To regulate or not to regulate?

Consider changing environments

Suppose that gene product Z helps only in the condition C_z

Suppose environment presents condition C_z with probability p . p is called the "demand" for Z

If I have constitutive expression of Z:

have a cost c , and a gain b a fraction p of time

So fitness $F_{\text{const}} = pb - c$

If I have regulated expression of Z, only in proper conditions:

have the cost c and a gain b , a fraction p of time. But regulation cost all the time

So fitness $F_{\text{reg}} = pb - pc - r$

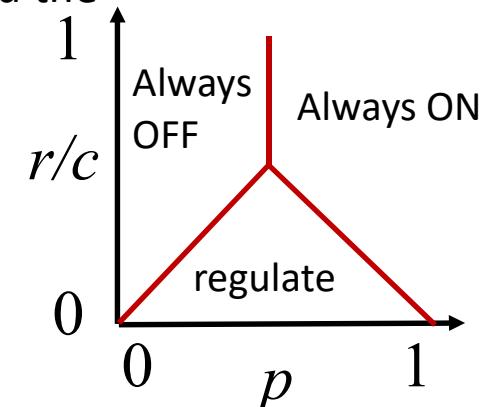
If I have no possible expression of Z, no cost no gain and this is a baseline:

So fitness $F_{\text{no}} = 0$

Regulation will be selected by evolution if its fitness is the highest: $F_{\text{reg}} > F_{\text{const}}, F_{\text{no}}$
which means $p < 1 - r/c$ and $p > r/(b-c)$

No regulation (constitutive expression) will be selected if $p > c/b$ and $p > 1 - r/c$.

Linking a property of the environment (p that C_z occurs) to cost and benefits of the protein Z.



Rich diversity of factors that can regulate rates of gene expression

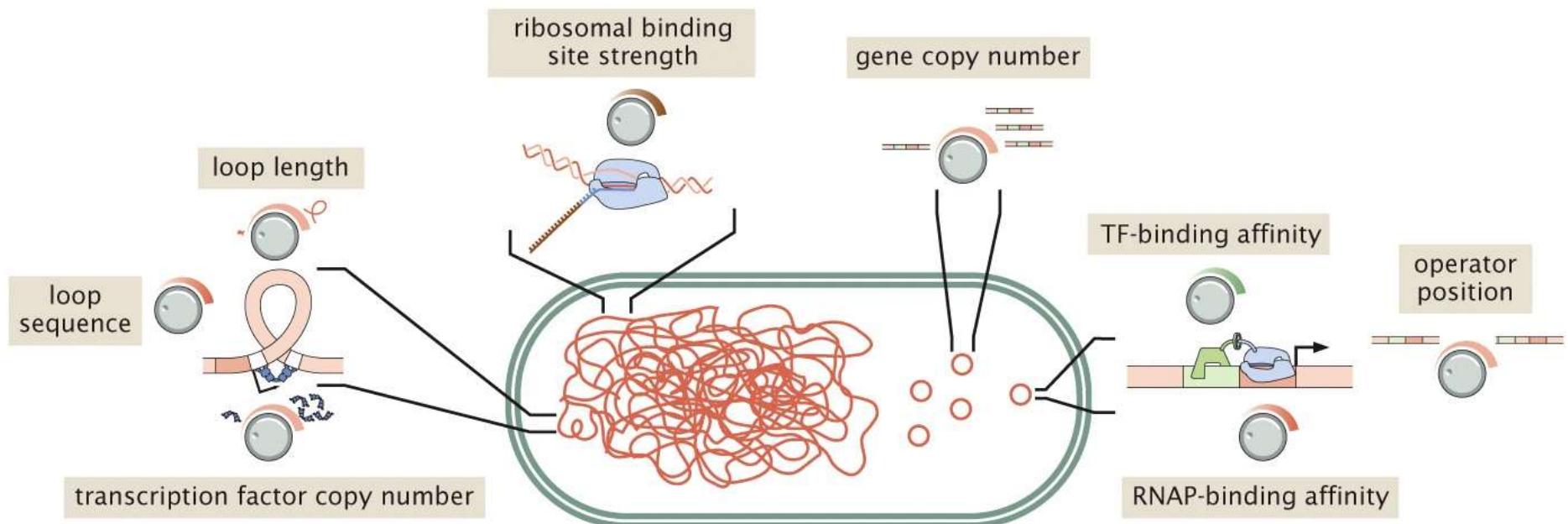


Figure 19.30 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Memo: Absolute numbers of various critical molecules are not large

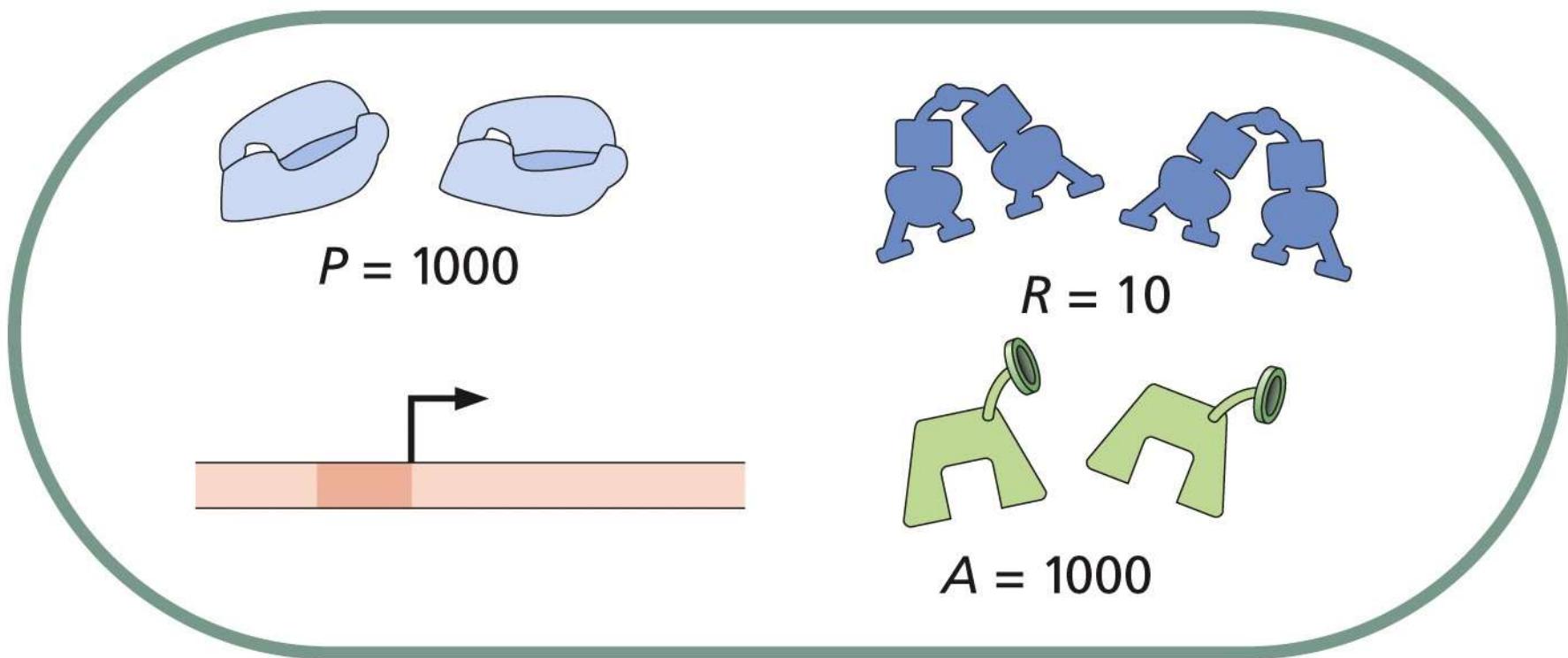


Figure 19.20 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Transcription is not a steady state: it occurs in ‘bursts’

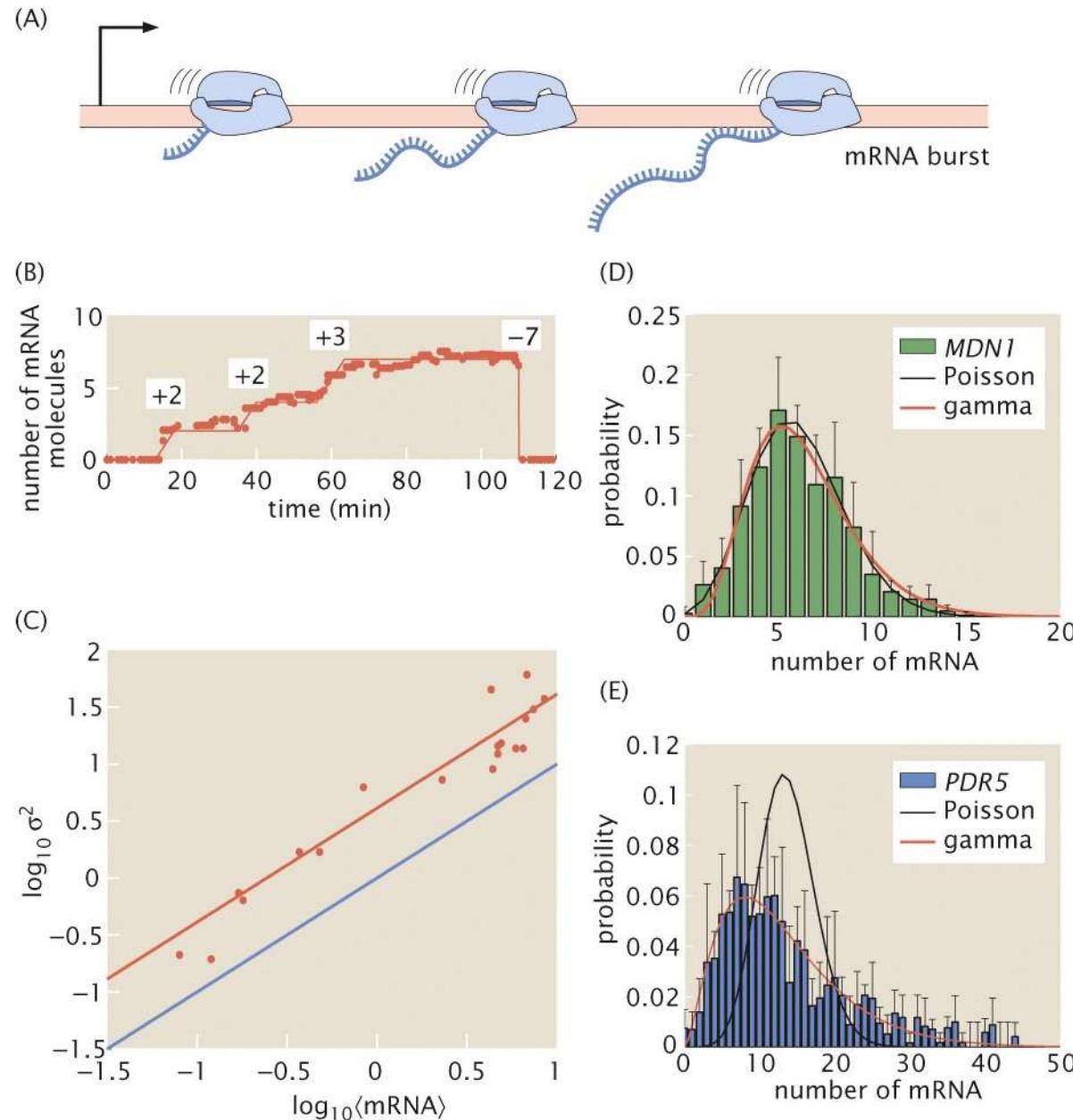


Figure 19.33 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Noise - “Birth-death” model for gene expression

Considering all the (4) ways in which the number of proteins can change, we have:

$$\frac{dp(N,t)}{dt} = kp(N-1) + \mu(N+1)p(N+1) - kp(N) - \mu N p(N)$$

You can check that the steady state solution is:

$$p(N) = \frac{1}{N!} \left(\frac{k}{\mu}\right)^N e^{-\frac{k}{\mu}}$$

... a Poisson distribution!

As (k/μ) increases, the average number of molecules increases. The mean and standard deviation σ_N of the distribution $p(N)$ are given by:

$$\begin{aligned}\langle N \rangle &= \frac{k}{\mu} \\ \sigma_N &= \sqrt{\langle N^2 \rangle - \langle N \rangle^2} = \sqrt{\frac{k}{\mu}}\end{aligned}$$

Can estimate the importance of stochastic effects considering the ratio of the standard deviation to the mean:

$$\frac{\sigma_N}{\langle N \rangle} = \sqrt{\frac{\mu}{k}} = \frac{1}{\sqrt{\langle N \rangle}}$$

We have made models based on molecular numbers. Markov chains. Now we have a molecular/physics picture of the underlying process.

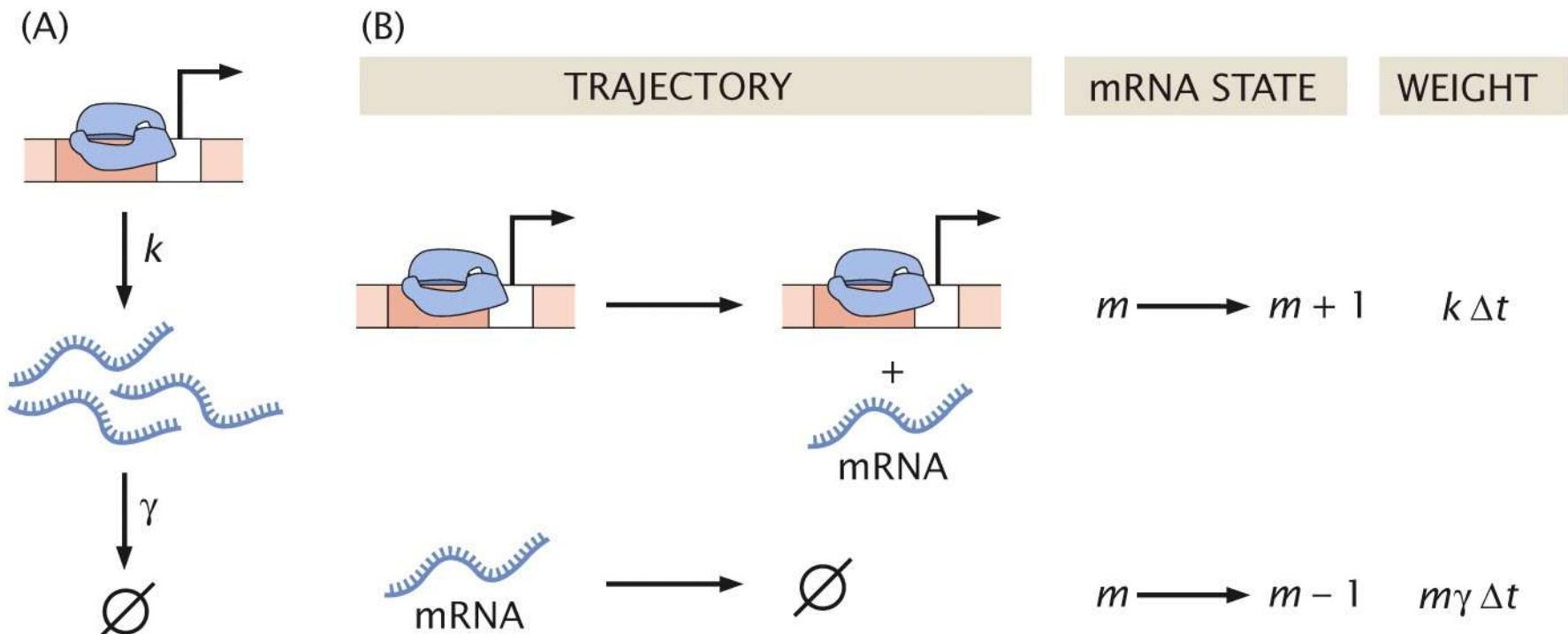


Figure 19.34 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Data of mRNA in *E. coli*

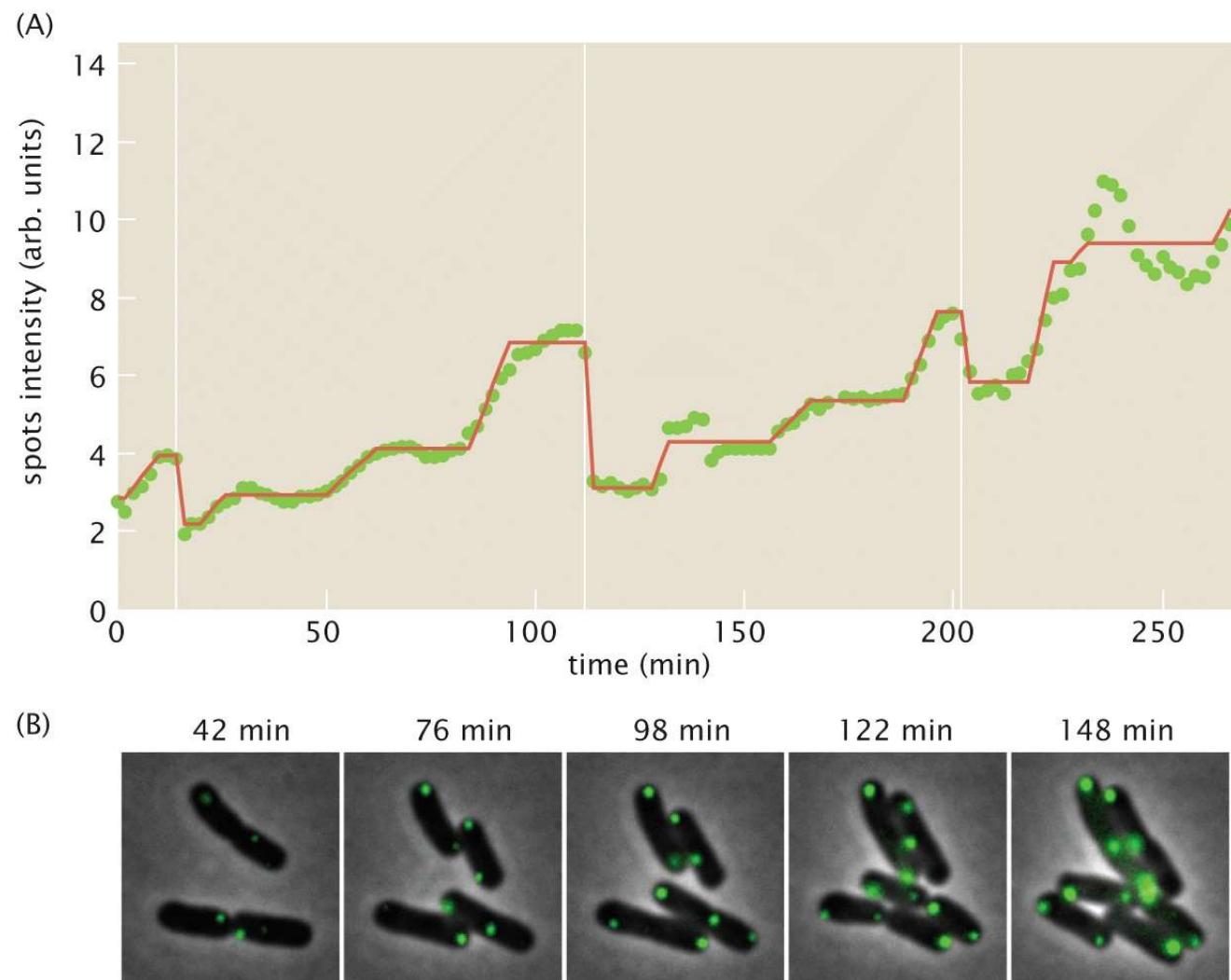


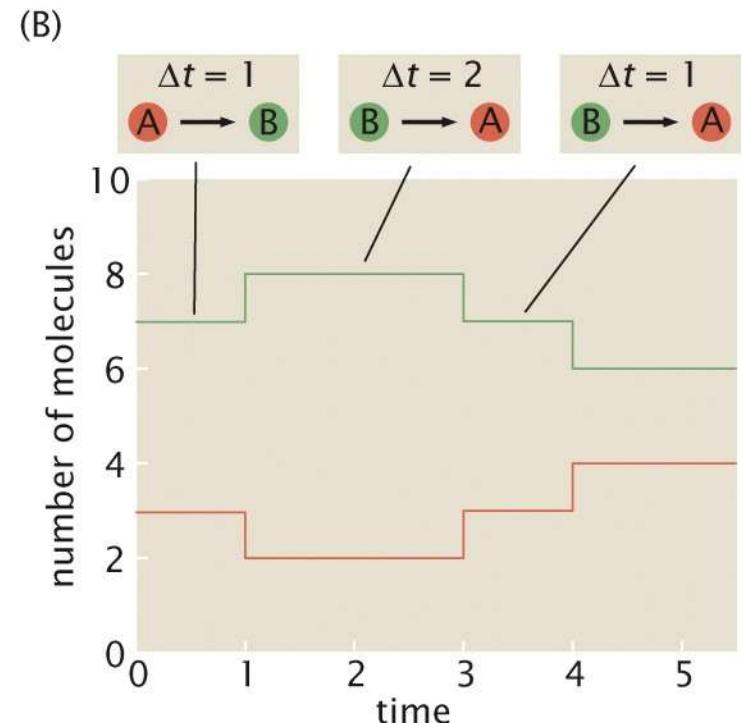
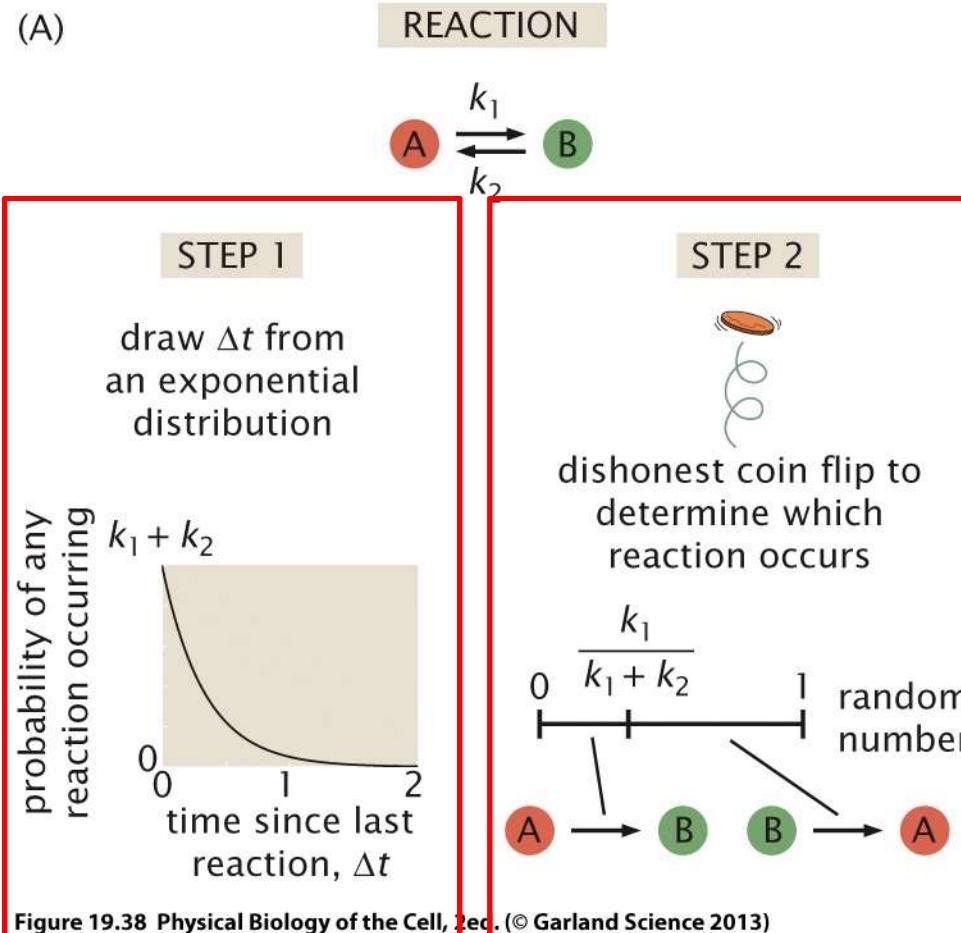
Figure 19.35 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Stochastic modeling

- *The master equation approach is only analytically tractable in very selected cases*
- *A numerical “brute force” approach with fixed discrete time intervals is wasteful because need Δt smaller than inverse of any rates, and any fluctuations that one wants to track. This means *many* Δt typically wasted with the system not changing state, particularly when small number of reactants.*
- *Better approach: have a strategy for adapting Δt*
- *Gillespie approach: randomly assign the time to the next reaction, from a probability distribution. This avoids timesteps where nothing happens.*
- *This creates a realisation of the stochastic time evolution of the system. Running several iterations gives distributions.*
- *Particular example $A \rightleftharpoons B$*

(Daniel Gillespie, Physicist, published milestone paper 1976)

Simulating efficiently networks of reactions: the Gillespie algorithm. Particular example:



2 random numbers:

Step 1- choose Δt to next reaction; Step 2 – biased choice of which reaction

Gillespie algorithm

We want a timestep Δt , and we want to determine $P(i, \Delta t)dt$, the probability that reaction i takes place in the interval $[\Delta t, \Delta t + dt]$.

First, we note that we also want to impose no reaction to have occurred before Δt . We call this probability $P_0(\Delta t)$.

Thus the probability that reaction i takes place in the interval $[\Delta t, \Delta t + dt]$ is:

$$P(i, \Delta t)dt = P_0(\Delta t)k_i dt.$$

Need to calculate $P_0(\Delta t)$: $P_0(\Delta t + dt) = P_0(\Delta t) \left(1 - \sum_i k_i dt\right)$

Taylor expand first term, so get: $\frac{dP_0(\Delta t)}{d\Delta t} = -P_0(\Delta t) \sum_i k_i$

And hence: $P_0(\Delta t) = e^{-\sum_i k_i \Delta t} = e^{-k_0 \Delta t}$ [$P_0(\Delta t = 0) = 1$ and $k_0 = \sum_i k_i$]

Gillespie algorithm (cont)

Substituting back, we get

$$P(i, \Delta t)dt = e^{-k_0 \Delta t} k_i dt.$$

If we sum this over all i , we get the probability that any of the possible reactions happens in the interval $\Delta t, \Delta t + dt$:

$$P(\Delta t)dt = e^{-k_0 \Delta t} k_0 dt.$$

This is the distribution from which one needs to pick Δt .

Now need to work out how to make a distribution from which to pick the random choice of which reaction takes place. The probability that reaction i happens at some time is:

$$P(i) = \int_0^\infty P(i, \Delta t)dt = \frac{k_i}{k_0}.$$

This gives us the criterion to choose (randomly, but with the right bias) which reaction will take place at the simulation timestep.

The unregulated promoter (see also back to first lectures of this set).

There are two reactions ($i=1,2$):

- (1) an mRNA can be produced, with probability k per unit time;
- (2) an mRNA can decay, with probability γ per unit time and per unit molecule.

Call $m(t)$ the number of mRNA molecules at time t . Rates are $k_1 = k$ and $k_2 = m(t) \gamma$.

In algorithm form, the steps in this example are:

1. given $m(t)$, calculate the rates. In this case only k_2 depends on $m(t)$.
2. draw a uniform random number x between $[0, 1]$. Compute k_0 . $\Delta t = (1/k_0) \ln(1/x)$. This last formula is a way (you can check) to turn the uniform random number in a random number from the exponential distribution we want, calculated above. Advance simulation clock by Δt .
3. draw a uniform random number between $[0, 1]$. If the number is between $[0, k_1/k_0]$, increase the mRNA molecule number by one. If it is between $[k_1/k_0, 1]$ then decrease the mRNA molecule number by one.
4. loop back to step (1).

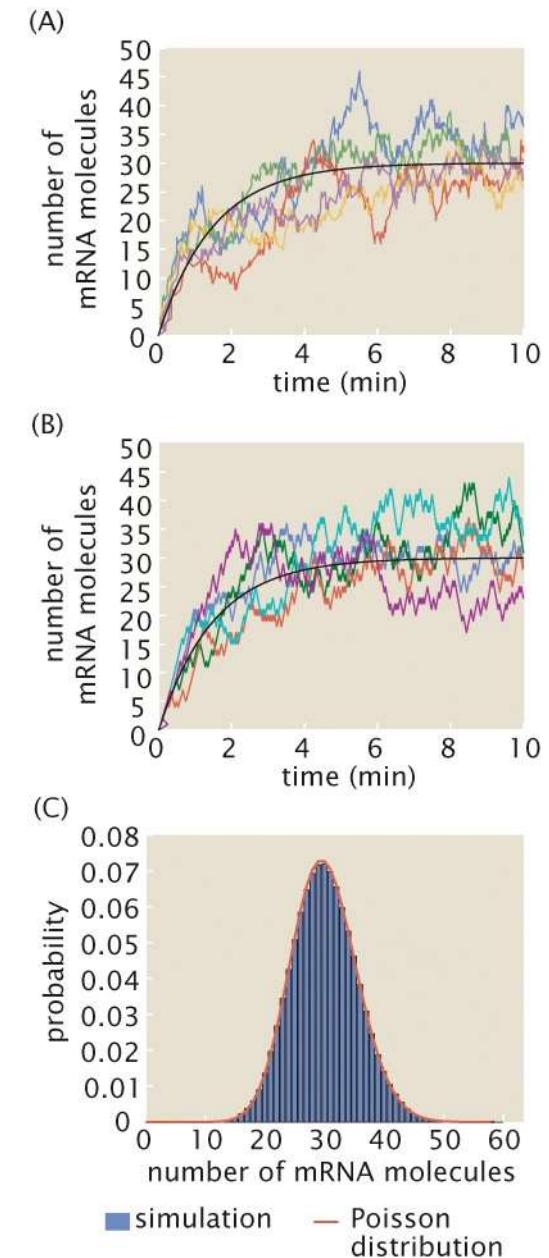
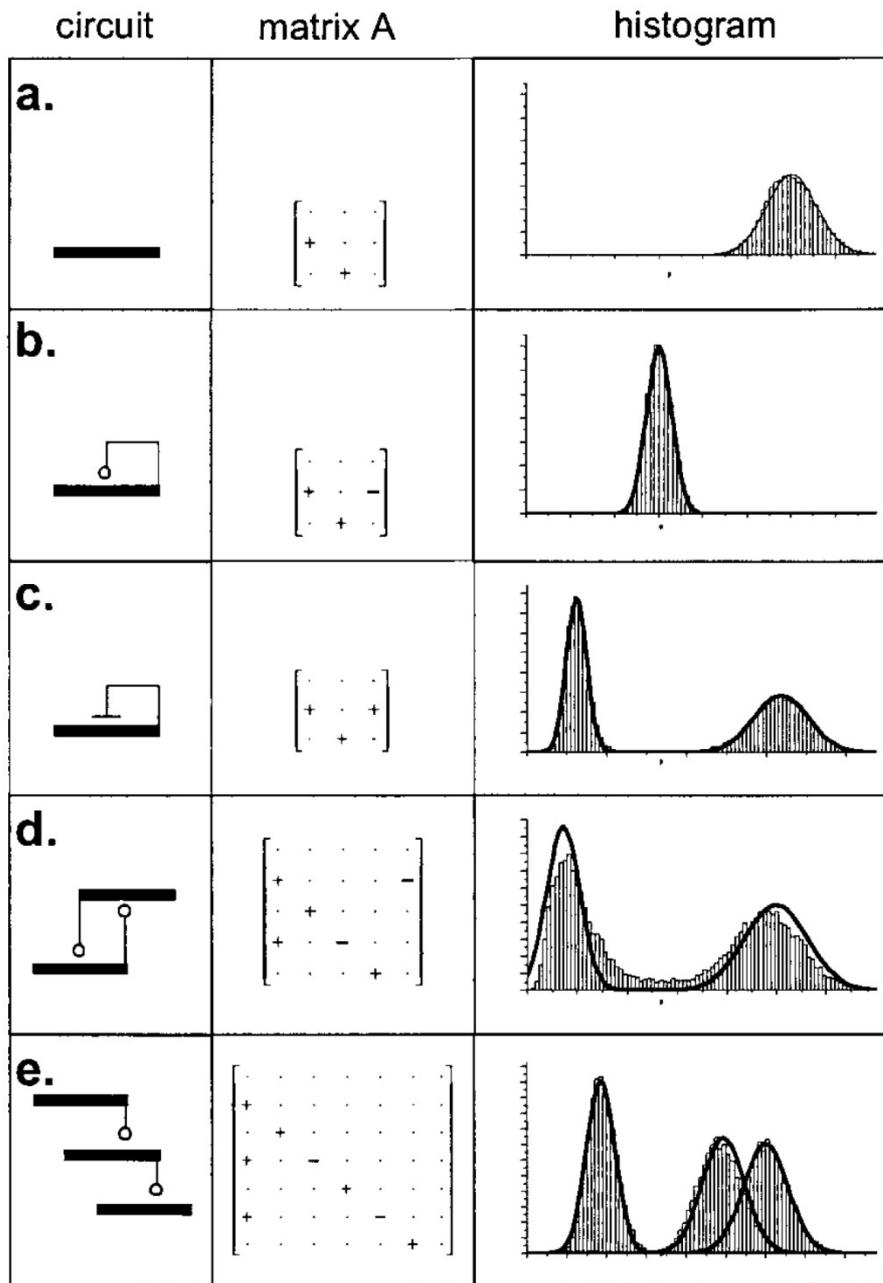


Figure 19.39 Physical Biology of the Cell, 2ed. (© 1

Efficient way to simulate networks.



$$f_{q_i} \xrightarrow{k_i^+(q_j)} f_{q_i+1}$$

$$f_{q_i} \xrightarrow{k_i^-(q_j)} f_{q_i-1}$$

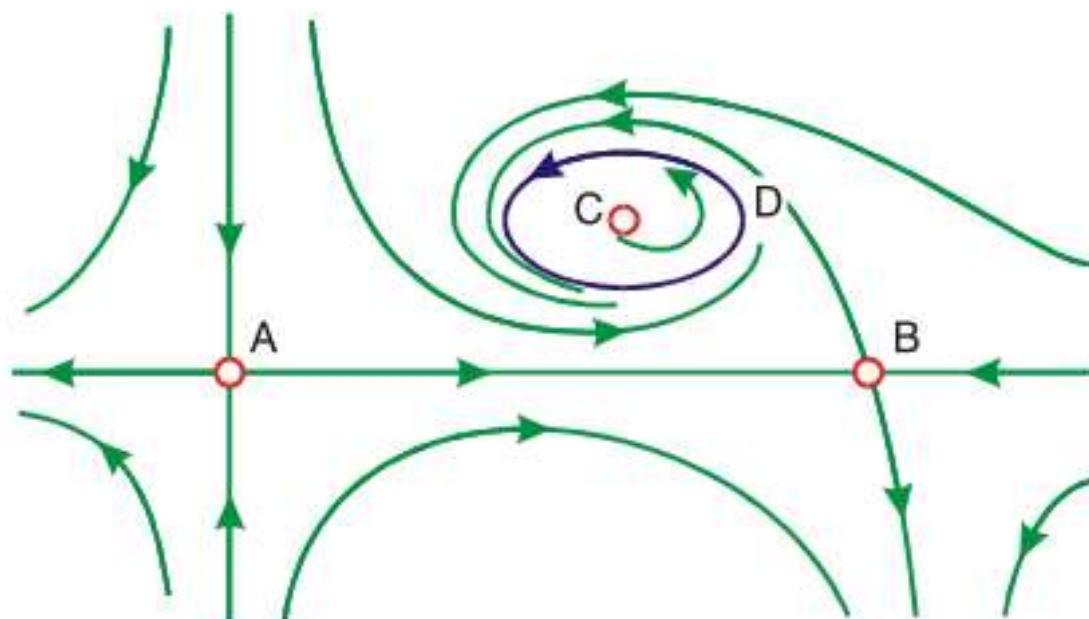
$$\{q_i\} = \{D, r, p\}$$

$$k_i^+(q_j) = \sum_j A_{ij} q_j \quad k_i^-(q_j) = \sum_j \Gamma_{ij} q_j$$

$$A = \begin{matrix} (D) & (r) & (p) \\ (D) & 0 & 0 & 0 \\ (r) & k_R & 0 & 0 \\ (p) & 0 & k_P & 0 \end{matrix} \quad \Gamma = \begin{matrix} (D) & (r) & (p) \\ (D) & 0 & 0 & 0 \\ (r) & 0 & \gamma_R & 0 \\ (p) & 0 & 0 & \gamma_P \end{matrix}$$

These examples from
 “Intrinsic noise in gene
 regulatory networks”
 Thattai and van Oudenaarden,
 PNAS 98, 8614-8619 (2001)

Memo- Example of trajectories in a general two-variable system $\dot{x}=f(x)$



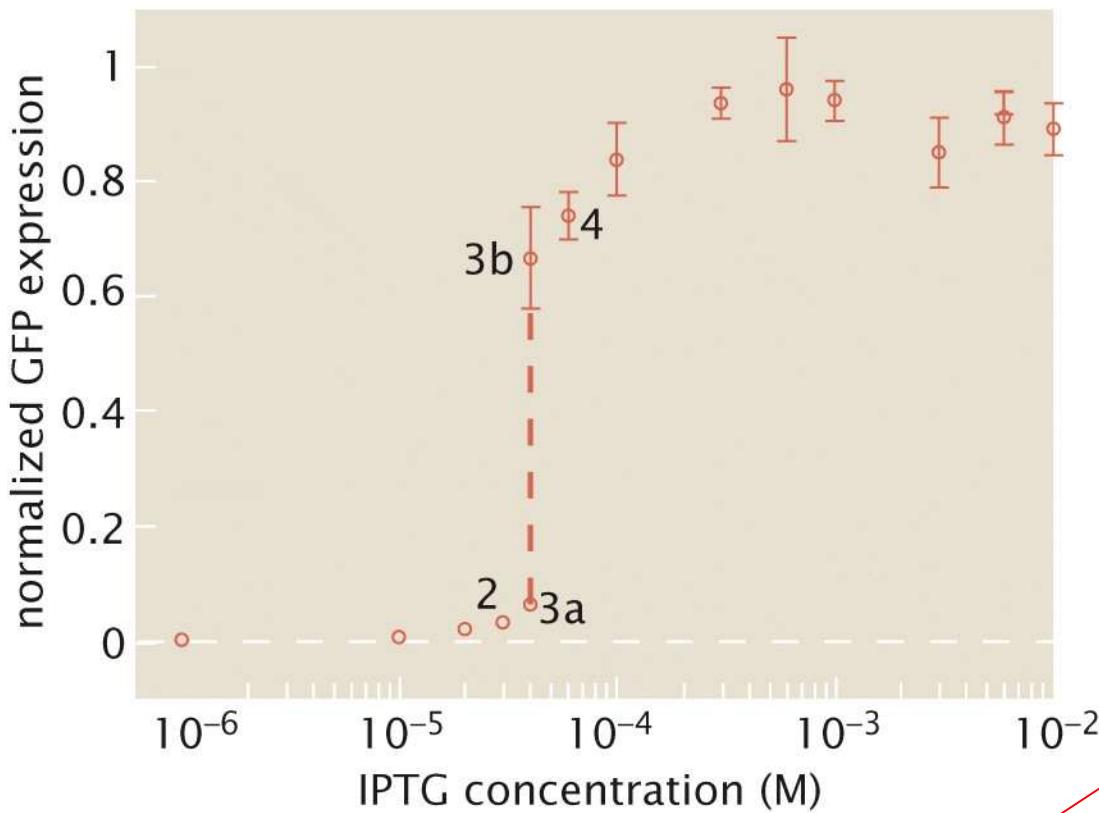
There are a number of fixed points, i.e those for which $\dot{x}=f(x) = 0$.

A and B are saddle nodes;
C is an unstable node (giving rise to the unstable spirals);
D is a limit cycle – that is, an isolated closed trajectory.

Spirals internal and external to D are both attracted towards it.

Switches in genetic circuits

(A)



(B)

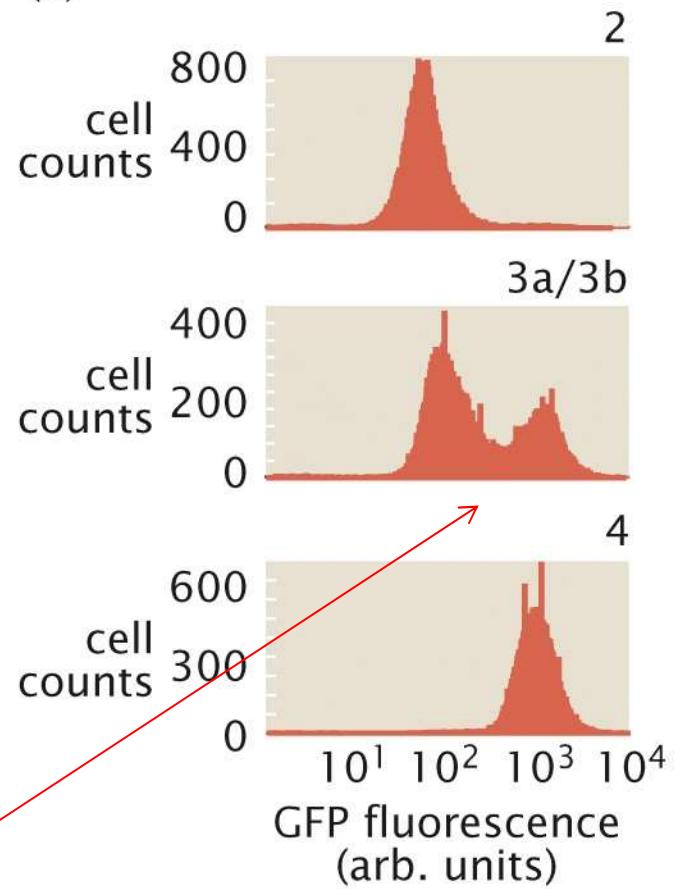


Figure 19.44 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Gardner et al, Nature 2000

IPTG, a lactose analog, is an inducer that flips the “switch” (we saw IPTG binds to the Lac repressor and reduces its affinity).

Note: bistability, not a continuous change. The bimodality occurs due to natural fluctuations in gene expression and the close proximity of the toggle switch to its bifurcation point.

Switches in genetic circuits: 2 repressor proteins mutually regulated

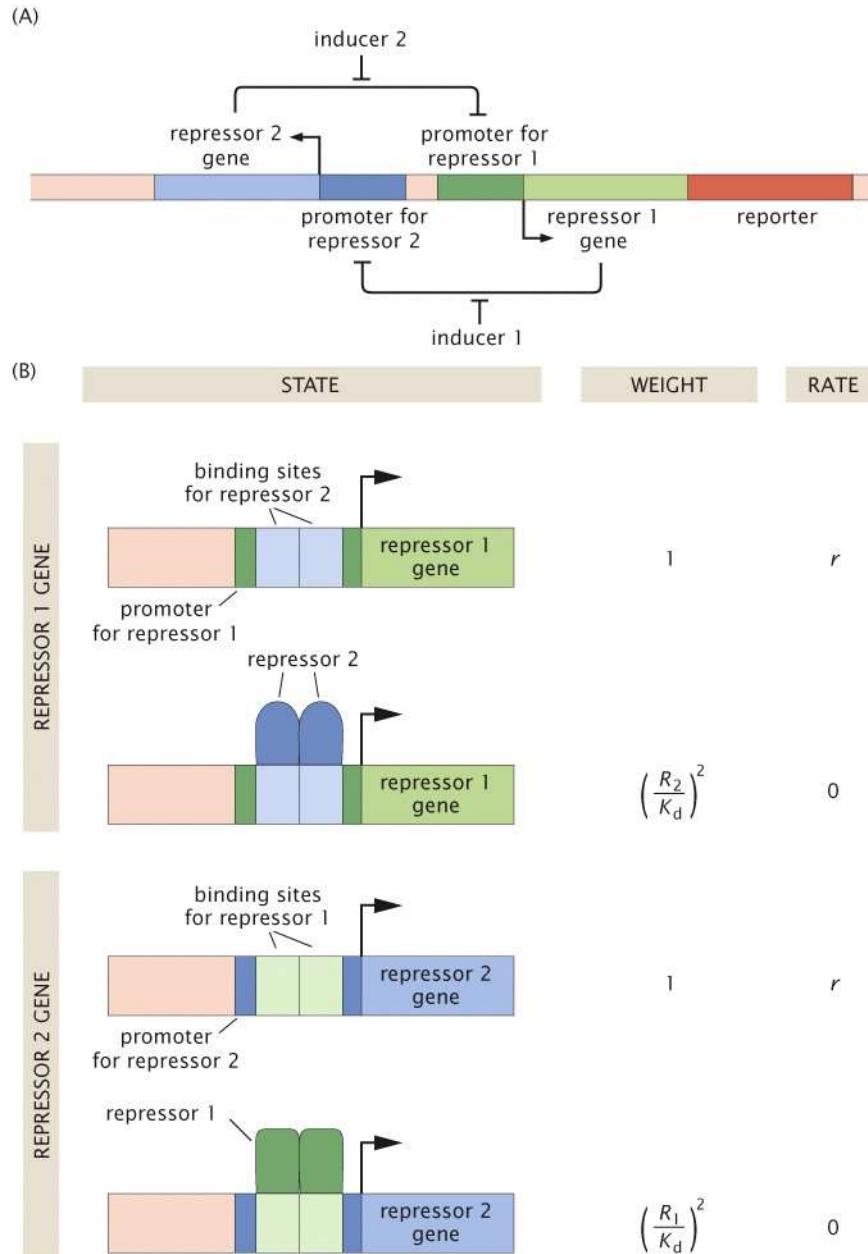


Figure 19.45 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Mathematical model for this genetic switch

c_1, c_2 the two repressor protein concentrations.

Both proteins can degrade, rate γ

To model the repression on "2", consider a modified expression rate of form $r(1-p_{\text{bound}})$, where r is a basal rate.

If binding is a Hill function, $p_{\text{bound}}(c_1) = \frac{K_b c_1^n}{1 + K_b c_1^n}$

$$\begin{aligned}\frac{dc_1}{dt} &= -\gamma c_1 + \frac{r}{1 + K_b c_2^n} \\ \frac{dc_2}{dt} &= -\gamma c_2 + \frac{r}{1 + K_b c_1^n}\end{aligned}$$

Then the reaction system is:

This system has 2 ranges of parameters giving different behaviours.

- A) a range giving a single stable solution, with equal concentrations of both proteins.
- B) a range giving two well distinct protein concentrations.

If the parameters are in the range of (B), then the network exhibits switch like behaviour. The system can be made dimensionless (with $\alpha = r K_b^{1/n} / \gamma$ and time in units of $1/\gamma$):

$$\begin{aligned}\frac{du}{dt} &= -u + \frac{\alpha}{1 + v^n} \\ \frac{dv}{dt} &= -v + \frac{\alpha}{1 + u^n}\end{aligned}$$

Mathematical model for this genetic switch (cont)

One solution is always $u^* = v^* = \frac{\alpha}{1 + v^{*n}}$

Can work out if there are other solutions. Take $n=2$ to proceed simply. For u^* , and similarly v^* :

$$u^* = \frac{\alpha}{1 + \left(\frac{\alpha}{1+u^{*2}}\right)^2}$$

Thus:

$$(u^{*2} - \alpha u^* + 1)(u^{*3} + u^* - \alpha) = 0$$

The cubic polynomial here can be shown to have only one zero, and by some inspection you can see that it is the solution with $u^* = v^*$.

The quadratic however can have 0 (if $\alpha < 2$), 1 (if $\alpha = 2$), or 2 (if $\alpha > 2$) solutions, depending on the value of α . In the 2-solution regime, the concentrations are not the same! The solution with $u^* = v^*$ exists for all α , but it is unstable for $\alpha > 2$.

Can calculate phase portraits of this system:

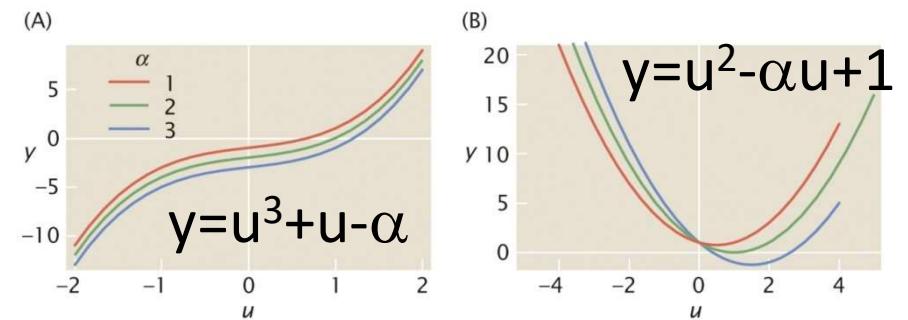


Figure 19.46 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Calculate these in Matlab

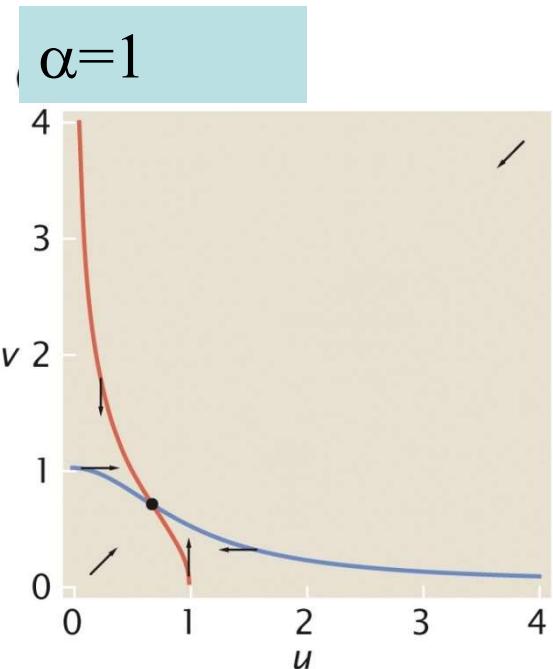
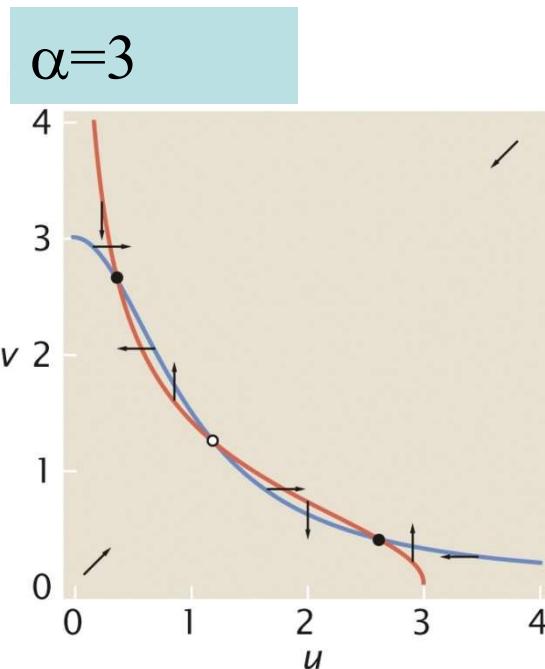


Figure 19.48 Physical Biology of the Cell, 2ed. (© Garland Science 2013)



So, $\alpha=2$ separates region with one single solution of equal concentrations, against having two stable distinct solutions and one unstable.

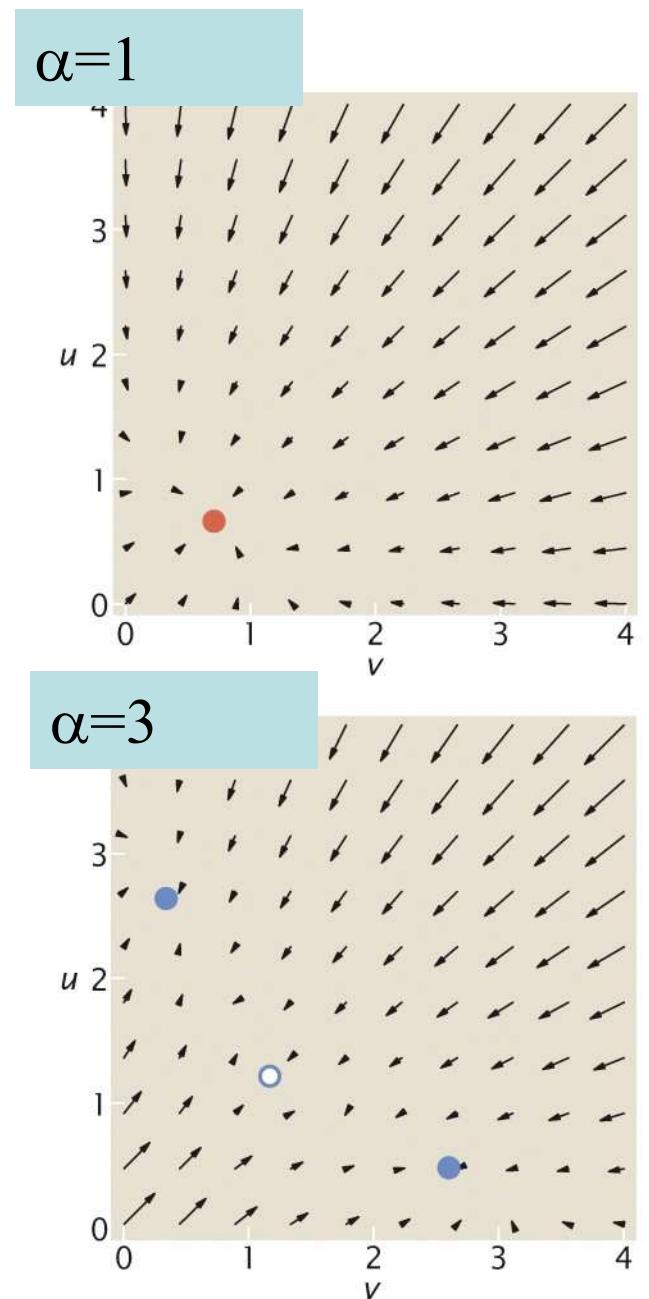


Figure 19.47 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Oscillations in genetic circuits

One simple set of equations that gives rise to oscillations is a gene regulated by both an activator and a repressor:

- the repressor binds as a dimer, and represses production of the activator
- the activator also binds as a dimer, and increases the production of itself, and also of the repressor.

Then the rate equations can be written as:

$$\begin{aligned}\frac{dc_A}{dt} &= -\gamma_A c_A + r_{0A} \frac{1}{1 + (C_A/K_d)^2 + (C_R/K_D)^2} + r_A \frac{(c_A/K_d)^2}{1 + (C_A/K_d)^2 + (C_R/K_D)^2} \\ \frac{dc_R}{dt} &= -\gamma_R c_R + r_{0R} \frac{1}{1 + (C_A/K_d)^2} + r_R \frac{(c_A/K_d)^2}{1 + (C_A/K_d)^2},\end{aligned}$$

Where r_{0A} , r_{0R} are the basal expression rates, and r_A , r_R are the regulated rates in the presence of the activator bound.

Can be made nicely dimensionless:

$$\begin{aligned}\frac{d\tilde{c}_A}{dt} &= -\tilde{\gamma}_A \tilde{c}_A + \frac{\tilde{r}_{0A} + \tilde{r}_A \tilde{c}_A^2}{1 + \tilde{c}_A^2 + \tilde{c}_R^2} \\ \frac{d\tilde{c}_R}{dt} &= -\tilde{c}_R + \frac{\tilde{r}_{0R} + \tilde{r}_R \tilde{c}_A^2}{1 + \tilde{c}_A^2}.\end{aligned}$$

Oscillations in genetic circuits (cont)

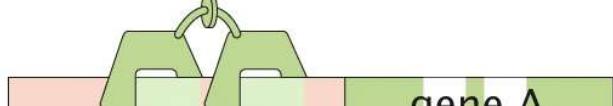
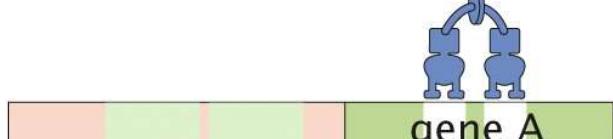
	(A) STATE	WEIGHT	EXPRESSION RATE
Activator gene, with positive feedback on itself		1	r_{0A}
		$\left(\frac{c_A}{K_d}\right)^2$	r_A
		$\left(\frac{c_R}{K_d}\right)^2$	0
Repressor gene, stimulated by the activator		1	r_{0R}
		$\left(\frac{c_A}{K_d}\right)^2$	r_R
			

Figure 19.50 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Oscillations in genetic circuits (cont)

Dimensionless system:

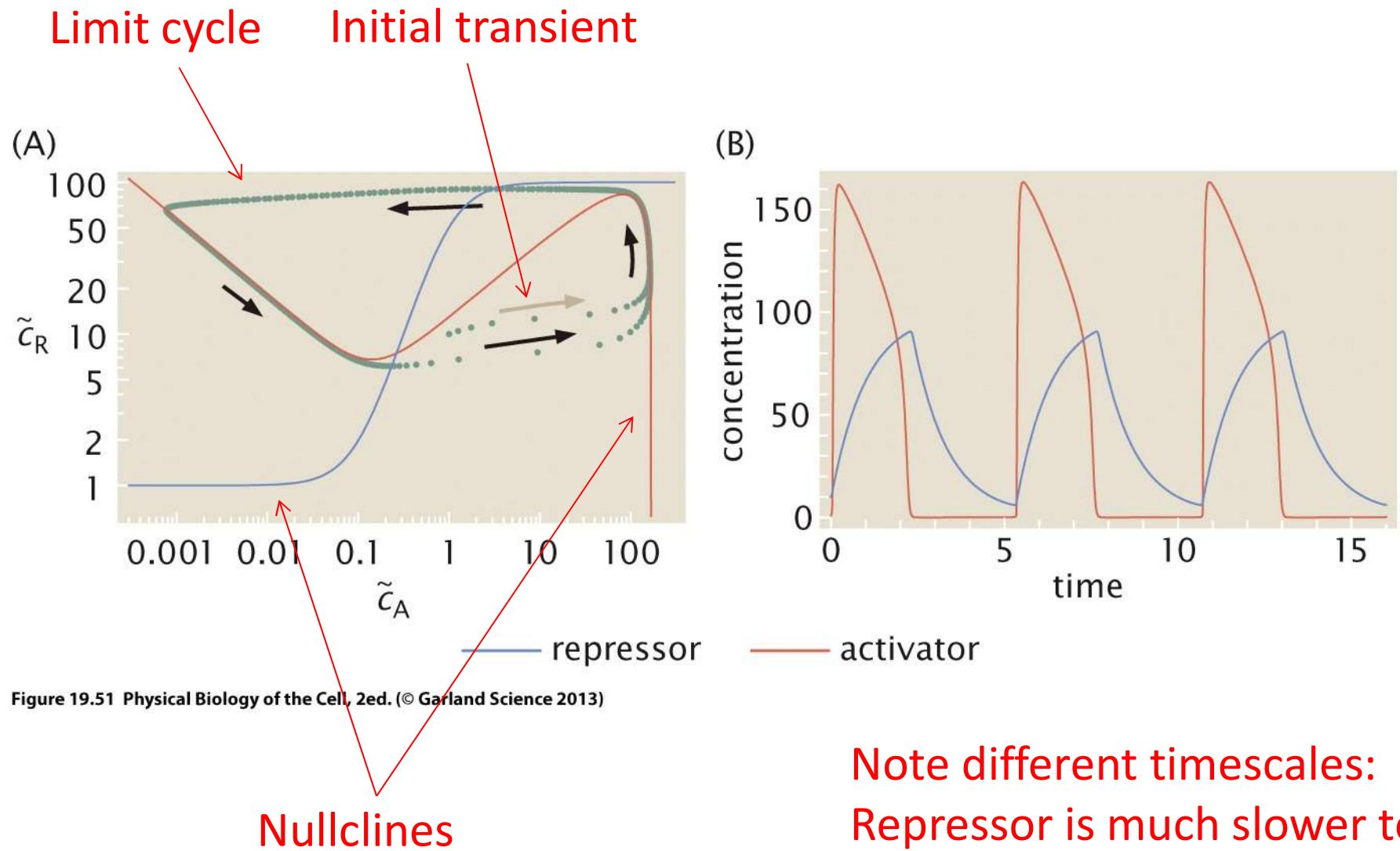
$$\begin{aligned}\frac{d\tilde{c}_A}{dt} &= -\tilde{\gamma}_A \tilde{c}_A + \frac{\tilde{r}_{0A} + \tilde{r}_A \tilde{c}_A^2}{1 + \tilde{c}_A^2 + \tilde{c}_R^2} \\ \frac{d\tilde{c}_R}{dt} &= -\tilde{c}_R + \frac{\tilde{r}_{0R} + \tilde{r}_R \tilde{c}_A^2}{1 + \tilde{c}_A^2}.\end{aligned}$$

Oscillations can arise if there is a separation of timescales between the activator and repressor dynamics.

‘Nullclines’ are the locus of points achieved by the repressor or activator at steady state, given fixed values of activator or repressor, respectively. They are obtained by setting the time derivatives equal to zero, and we have:

$$\begin{aligned}\tilde{c}_R &= \sqrt{-1 - \tilde{c}_A^2 + \frac{\tilde{r}_{0A} + \tilde{r}_A \tilde{c}_A^2}{\tilde{\gamma}_A \tilde{c}_A}} \\ \tilde{c}_R &= \frac{\tilde{r}_{0R} + \tilde{r}_R \tilde{c}_A^2}{1 + \tilde{c}_A^2}.\end{aligned}$$

Oscillations in genetic circuits (cont)



Note different timescales:
Repressor is much slower to equilibrate

This is an example of a “relaxation oscillator”

Natural Oscillators

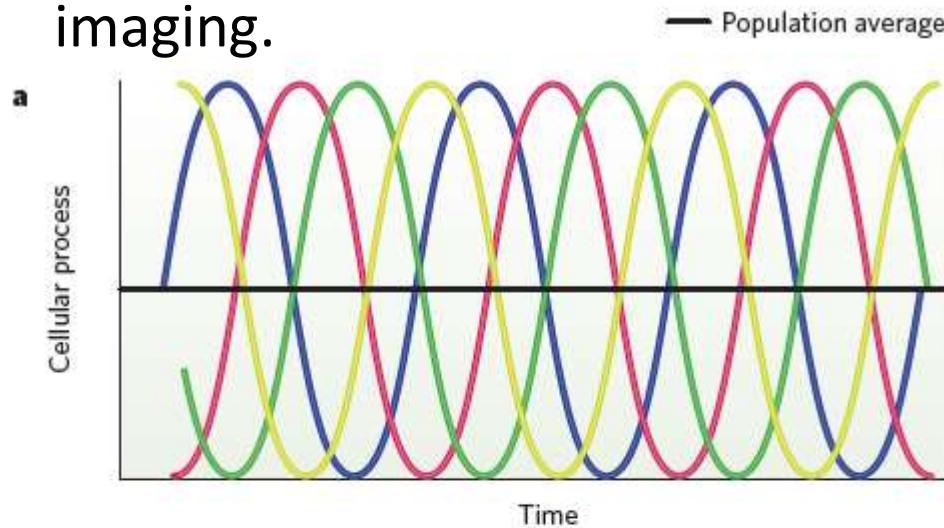
1. Circadian rhythms (eg Drosophila, 24 hour period, feedback oscillator)
2. Ca++ Oscillations
3. Glycolytic Oscillations * (relaxation oscillator)
4. Signalling Pathway Oscillations (P53, ERK, NF-κB)
5. Cell Cycle (relaxation oscillator)
6. Synchronous Rhythmic Flashing Of Fireflies
7. Segmentation during development
8. Many examples of chemical oscillators (mostly relaxation oscillators)
9. Motile Cilia in the airways and in the brain
10.

Buck, John; "Synchronous Rhythmic Flashing of Fireflies. II," *Quarterly Review of Biology*, 63:265, 1988

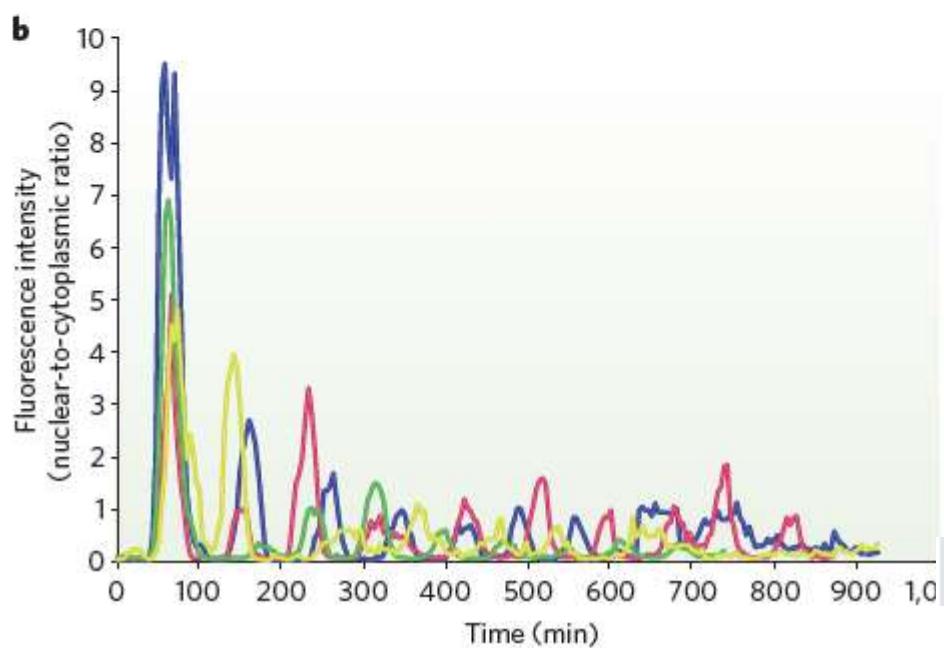
* During growth phase on glucose and ethanol, starve yeast of glucose, add cyanide and glucose, the glycolytic pathway will oscillate (NAD/NADH, ATP/ADP)



Many signals oscillate in cells – experimental importance of single cell imaging.



NATURE|Vol 465|10 June 2010



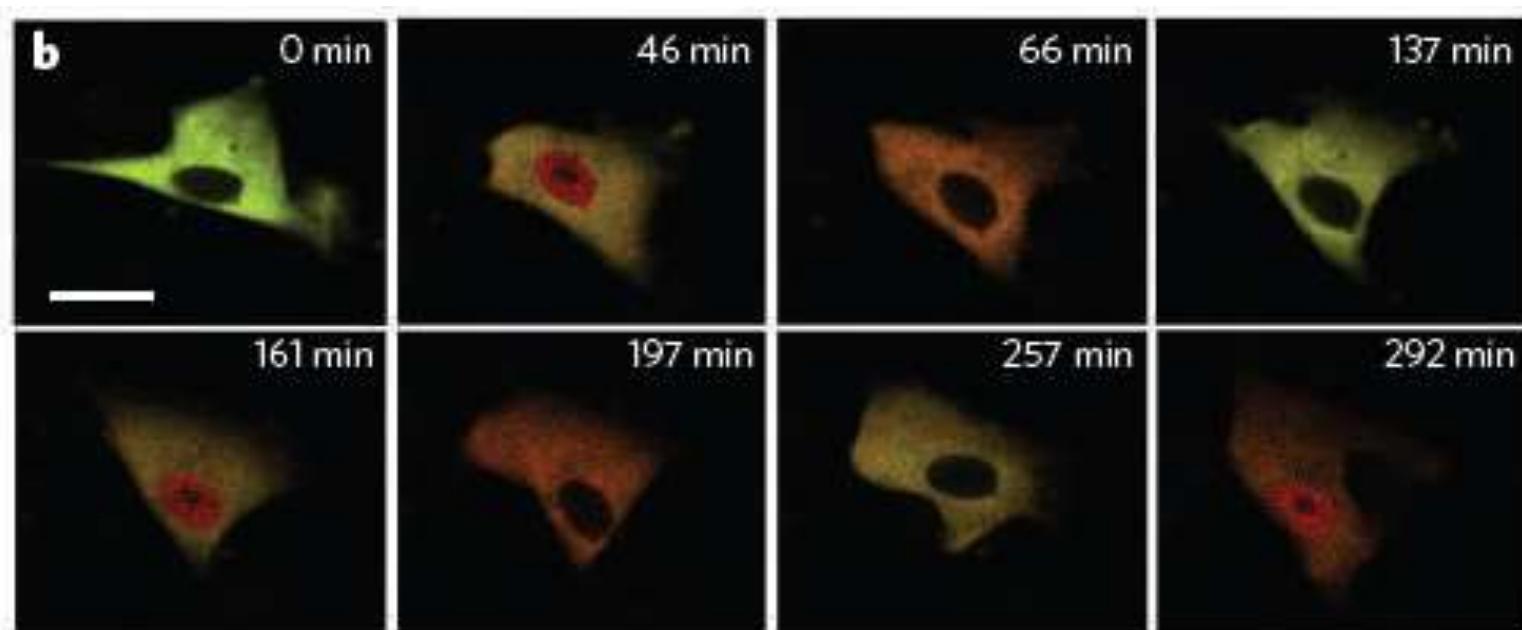
Cellular signal	Period length
Calcium	Seconds to minutes
ERK2	~15 min
Transcription cycles	Tens of minutes
NF-κB	~100 min
Crz1 (similar to NFAT)	Random, pulses occur over minutes
Segmentation (Notch, WNTs, FGFs)	~90 min
STAT signalling	~2 h
p53	5–6 h
Cell cycle	~18 h
Circadian rhythms	24 h

ERK, extracellular signal-regulated kinase; FGF, fibroblast growth factor; NFAT, nuclear factor of activated T cells; STAT, signal transducer and activator of transcription.

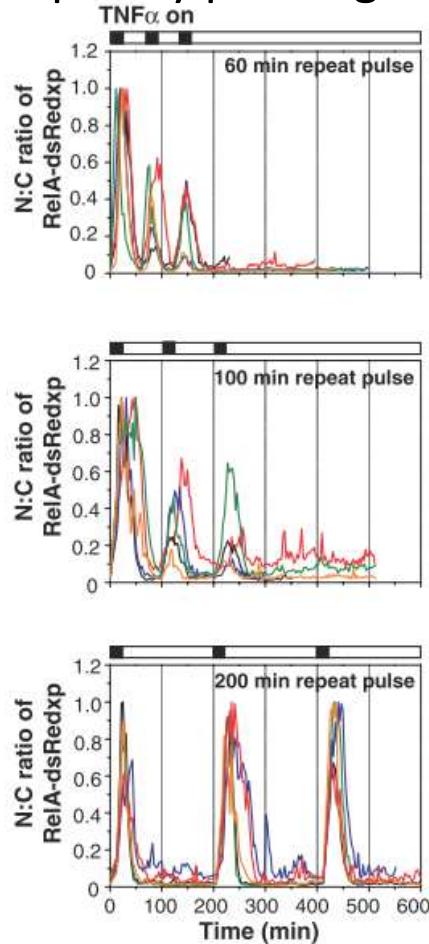
Case of NF-κB

Transcription factor translocates repeatedly between cytoplasm and nucleus of innate immune cells. This is just downstream of the “sensor” for infection, and serves as the trigger for fast innate immune response: “inflammation”.

The nuclear factor κB (NF-κB) transcription factor regulates over 300 genes.



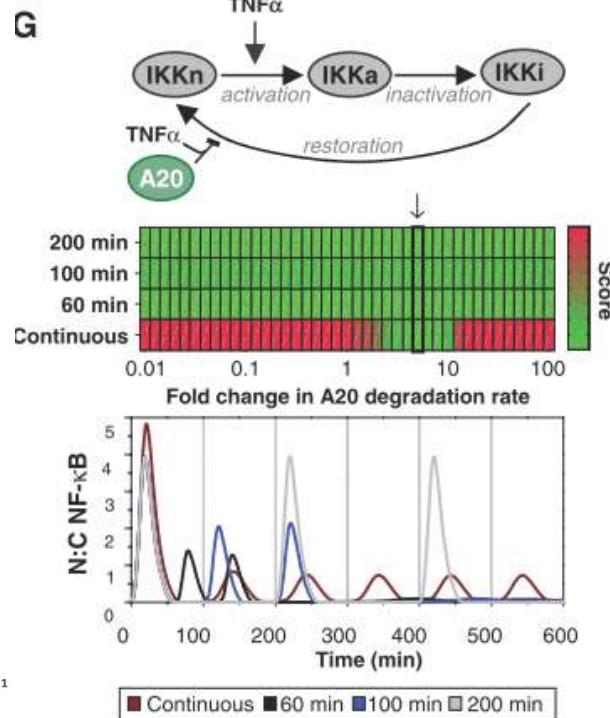
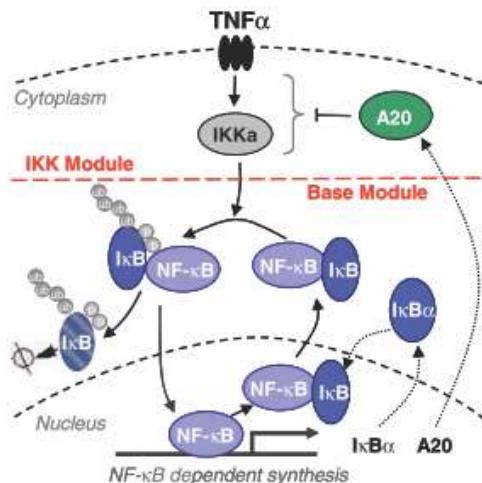
Experiment: cells were subject to pulsatile inflammatory signals. Observed synchronous cycles of NF- κ B nuclear translocation.
 Lower frequency stimulations gave full-amplitude translocations, whereas higher frequency pulses gave reduced translocation, indicating a failure to reset.



Pulsatile Stimulation Determines Timing and Specificity of NF- κ B-Dependent Transcription

Louise Ashall,^{1*} Caroline A. Horton,^{1*} David E. Nelson,^{1*} Paweł Paszek,^{1*} Claire V. Harper,¹ Kate Sillitoe,¹ Sheila Ryan,¹ David G. Spiller,¹ John F. Uniti,² David S. Broomhead,³ Douglas B. Kell,⁴ David A. Rand,⁵ Violaine Séé,¹ Michael R. H. White^{1†}

2009 VOL 324 SCIENCE



Deterministic and stochastic mathematical models predicted how negative feedback loops regulate both the resetting of the system and cellular heterogeneity.

Altering the stimulation intervals gave different patterns of NF- κ B-dependent gene expression, supporting the idea that oscillation frequency has a functional role.

Tying some threads together, and getting closer to research themes

Taming protein production bursts

Noise in dynamical systems

Beyond model experiments in bacteria

Intrinsic noise in gene regulatory networks

Mukund Thattai and Alexander van Oudenaarden*

PNAS | July 17, 2001

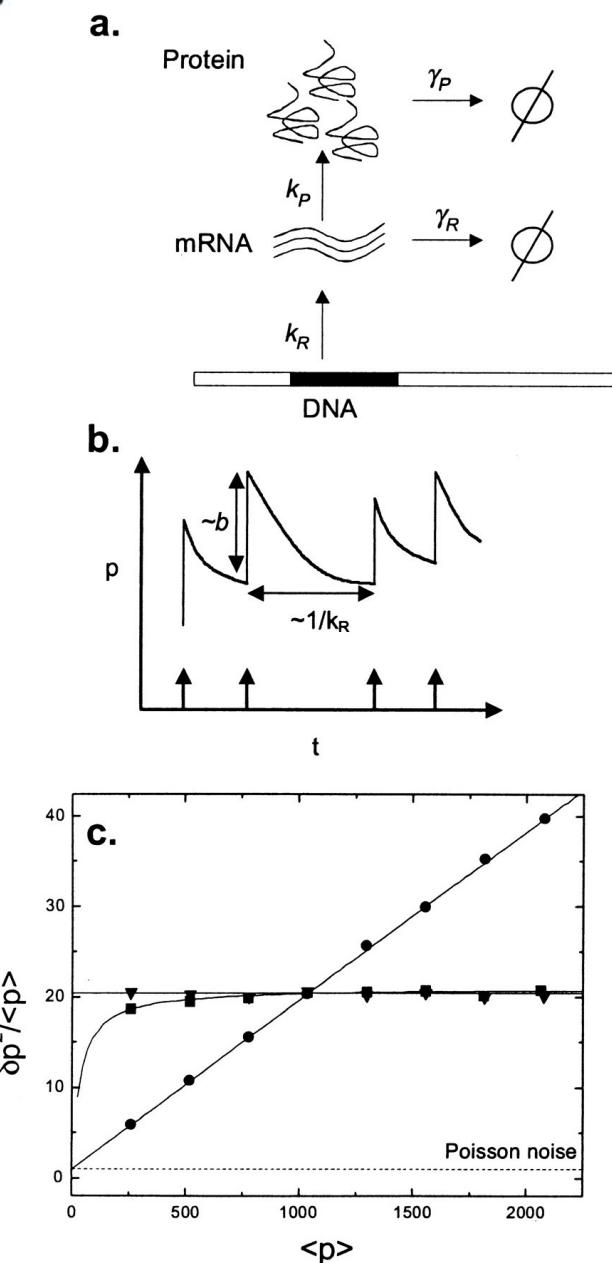
Noise at the protein level can be evaluated analytically. In particular, it can be quantified measuring the relative fluctuations with the coefficient of variation

$$CV_p = \sigma_p / \langle p \rangle$$

(where p represents the protein level), which for a constitutive gene takes the simple form

$$CV_p \approx \sqrt{[(1+b) / \langle p \rangle]},$$

where the mean protein level is simply given by the product of the average size of bursts and their frequency: $\langle p \rangle = b a$.



microRNA (miRNA)

We have assumed so far regulation of protein production by Transcription Factors (TF) proteins regulating promoter region activity.

This is the main regulatory mechanism, but does not account for everything.

“Post-transcriptional” regulation (and post-translational) mechanisms also exist.

Why ?

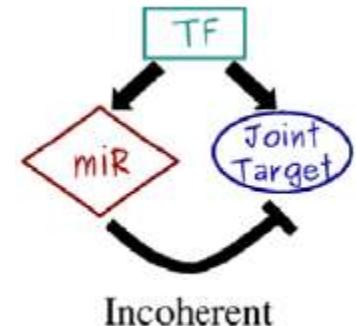
An important mechanism is played by microRNA.

MicroRNAs are non-coding RNAs that negatively regulate the protein production of their mRNA targets in metazoans and plants. They are rather small (about 22 nucleotides long), single stranded RNAs, and are known to target a substantial portion of the human genome. In bacteria, there are similar short single strands of RNA with regulatory function, called sRNA.

The Role of Incoherent MicroRNA-Mediated Feedforward Loops in Noise Buffering

Matteo Osella^{1,2,3*}, Carla Bosia^{1,2,3}, Davide Corá^{2,3}, Michele Caselle^{1,2}

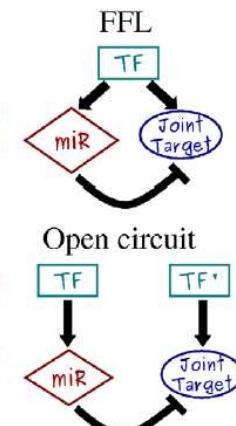
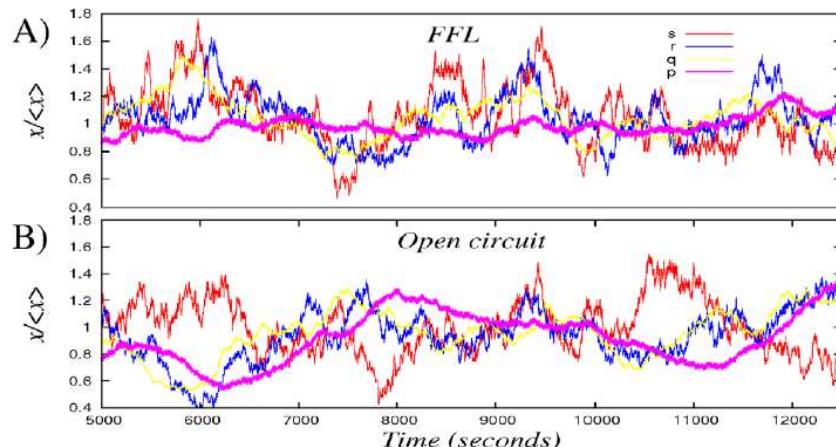
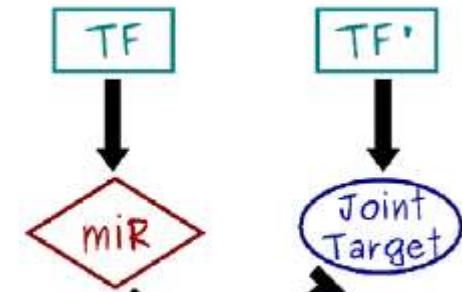
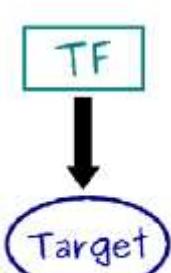
PLoS Computational Biology 2011



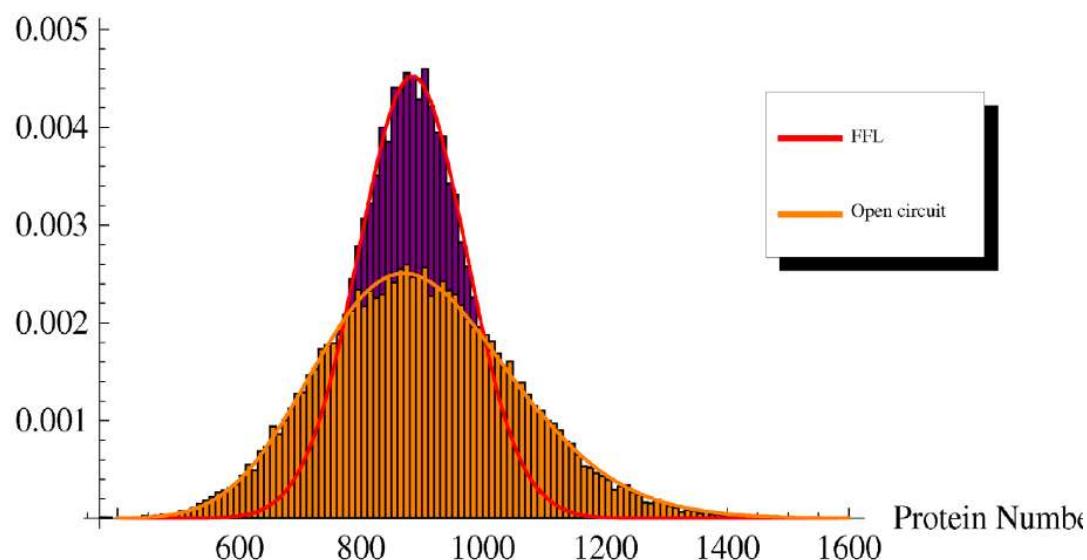
We have seen self-regulation as a vastly overrepresented motif. Another is the microRNA-mediated feedforward loop in which a “master” transcription factor regulates a microRNA and, together with it, a set of target genes.

It can be shown analytically and through simulations that the incoherent version of this motif can couple the fine-tuning of a target protein level with an efficient noise control, thus conferring precision and stability to the overall gene expression program, especially in the presence of fluctuations in upstream regulators.

A nontrivial result is that the optimal attenuation of fluctuations coincides with a modest repression of the target expression. This feature is coherent with the expected fine-tuning function and in agreement with experimental observations of the actual impact of a wide class of microRNAs on the protein output of their targets.



C) Probability



Tying some threads together, and getting closer to research themes

Taming protein production bursts

Noise in dynamical systems – 1st example

Beyond model experiments in bacteria

Noise control with microRNA

– so what can they do

[Osella et al., Front. Genet., 06 October 2014]

Gene expression is inherently a stochastic process , so fluctuations in protein concentration can induce random transitions between the alternative steady states of a bistable genetic circuit like the toggle switch.

A bistable circuit at the basis of cell fate determination is expected to be robust to these stochastic transitions, since they could in principle drive the cell to an undesired phenotype. **miRNA regulation can control the stochastic transitions between the two alternative steady states of a toggle switch.**

A major source of stochasticity in gene expression is due to the burstiness in protein production because (a) during the lifetime of a single mRNA several proteins can be produced, and (b) there are bursts of transcription, probably due to transitions in the promoter state.

A fluctuation at the transcriptional level can be amplified by the translation of a large protein burst stemming from a single mRNA. The average size of these bursts b is given by the product of the rate of translation k_p and the average lifetime of mRNA $1/\gamma_m$ (i.e., $b = k_p/\gamma_m$), while their frequency a is defined by the transcription rate k_m with respect to the timescale set by protein degradation γ_p (i.e., $a = k_m/\gamma_p$).

Regulation at the transcriptional level, e.g. by transcription factors, modulates the transcription rate of the target genes, thus affecting the burst frequency a . On the other hand, miRNAs exert their action by suppressing translation (i.e., decreasing k_p) or promoting mRNA degradation (increasing γ_m). Both regulative modalities affect the target burst size rather than the frequency.

Therefore, the same degree of repression exerted transcriptionally or **post-transcriptionally** via miRNAs will lead to very different levels of noise of the target protein concentration.

In particular, the expression for the coefficient of variation suggests that miRNAs, by reducing the target burst size, are more effective in keeping fluctuations in gene expression under control.

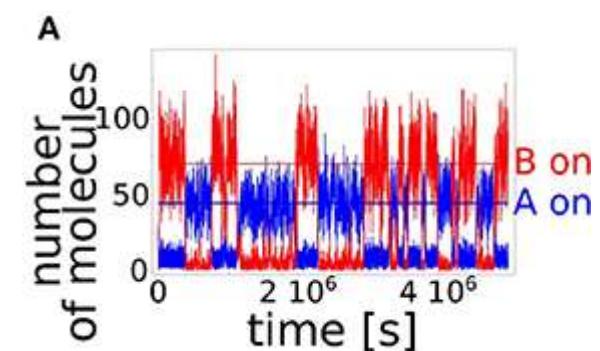
These simple arguments indicate a possible evolutionary reason to prefer miRNA regulation, instead of transcriptional regulation, to build toggle switches involved in cell fate decision: robustness to stochastic fluctuations is probably a crucial feature.

miRNA regulation is an example of **molecular titration** since it requires the direct one-to-one binding of a regulator and its target molecule (different from transcriptional regulation, where even a small amount of transcription factors can influence the target production with no significant consequences on their concentration).

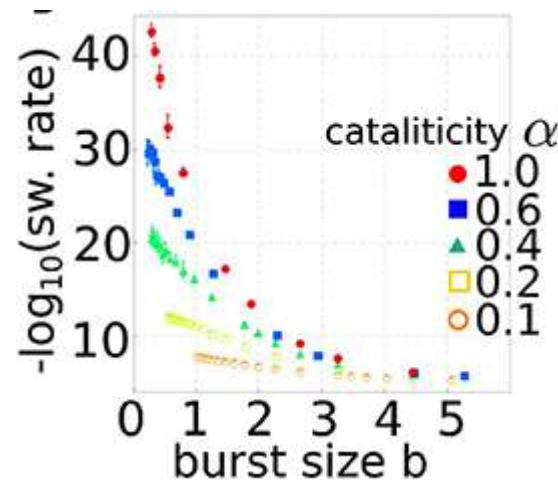
The equations (not empirical - physically justified!) can be written for this situation, terms are slightly different from TF. (- not to worry here.)

Then in order to evaluate the robustness of the circuit to stochastic transitions between the two steady states, the stochastic version of the model has to be considered. The level of stability of a steady state is given by the typical time the system manages to dwell in it before a stochastic transition, and this time can be evaluated using stochastic simulations.

The circuit randomly switches between the equilibrium in which gene A is on while B is shut off and the opposite state. The timing between these transitions defines the switching rate between the stable states.



The switching rate can change by several orders of magnitude depending on the level of miRNA regulation, and thus on the effective target burst size.
(don't worry about "catalicity", it is a parameter in the equations that we are not showing here)



A toggle switch composed only by transcriptional regulators would not be able to reduce the burst size of either of the two genes, and thus could not show this significant reduction in the switching rate.

Tying some threads together, and getting closer to research themes

Taming protein production bursts

Noise in dynamical systems – 2nd example

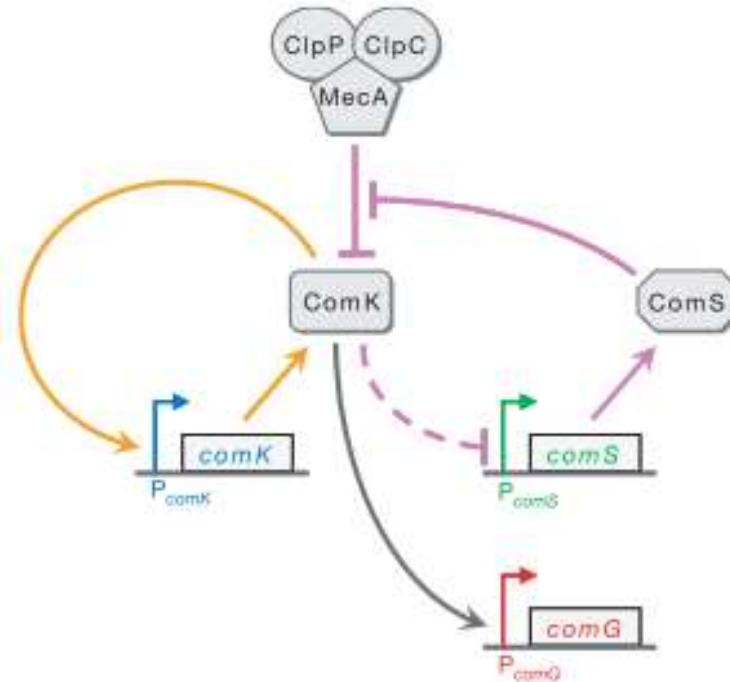
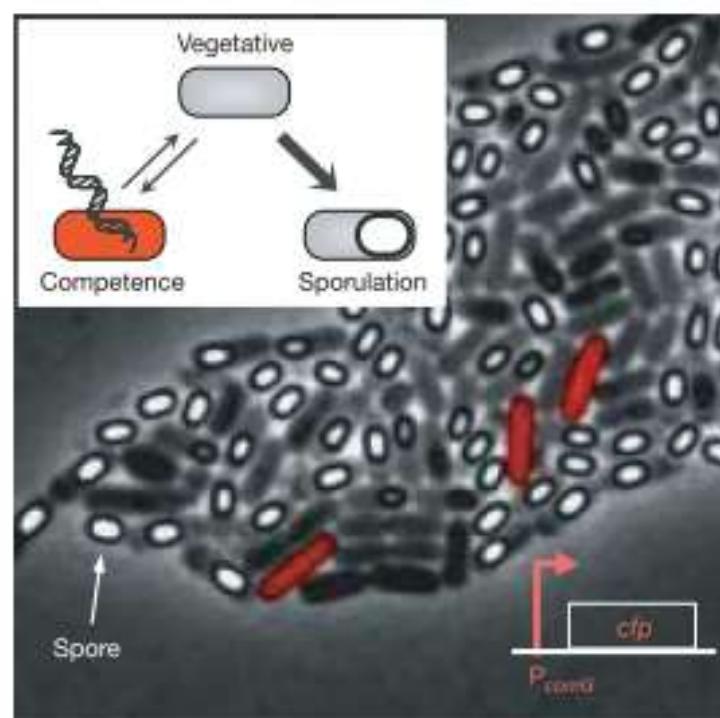
Beyond model experiments in bacteria

An excitable gene regulatory circuit induces transient cellular differentiation

NATURE|Vol 440|23 March 2006

Gürol M. Süel¹, Jordi Garcia-Ojalvo², Louisa M. Liberman¹ & Michael B. Elowitz¹

Certain types of cellular differentiation are probabilistic and transient. In such systems individual cells can switch to an alternative state and, after some time, switch back again. In *Bacillus subtilis*, competence is an example of such a transiently differentiated state: competence is a response to a stress, with the cells becoming able to uptake DNA from the environment.



Encountering nutrient limitation, a minority of *B. subtilis* cells become competent for DNA uptake while most commit irreversibly to sporulation.

Individual genes and proteins underlying differentiation into the competent state have been identified, but it was unclear how these genes interacted dynamically in individual cells to control both spontaneous entry into competence and return to vegetative growth.

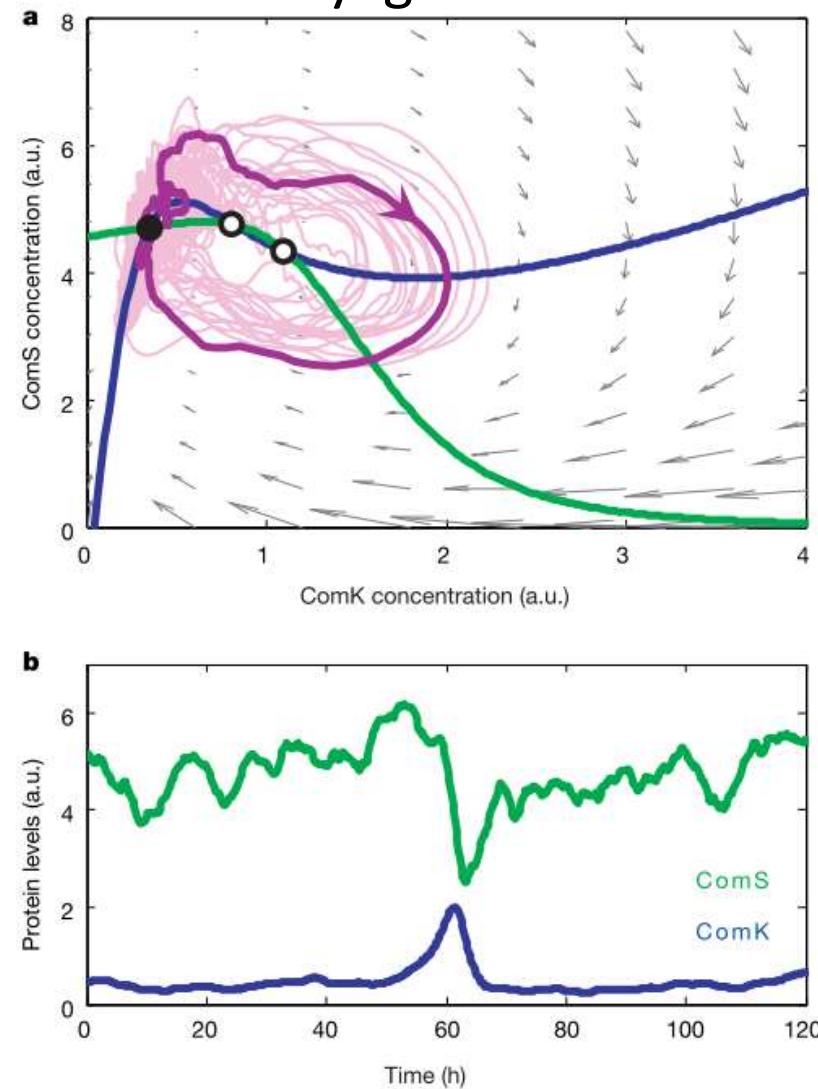
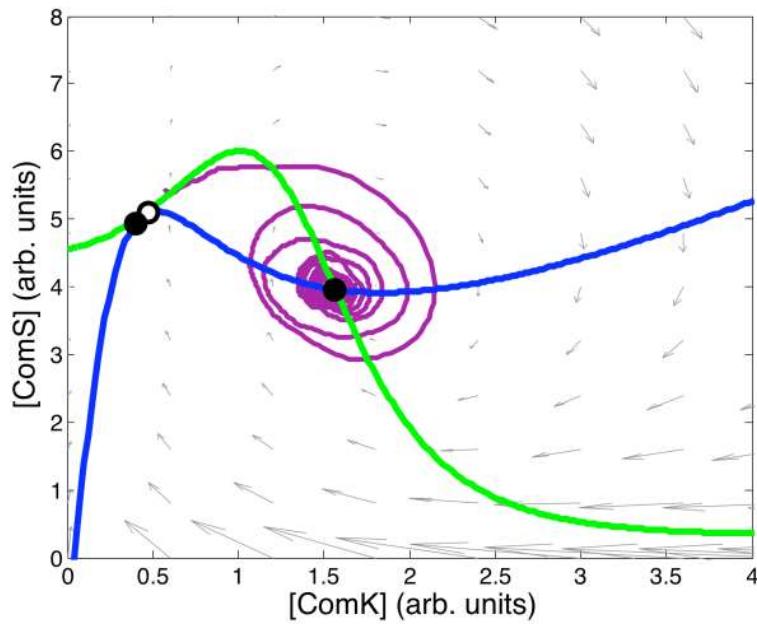
This behaviour can be understood in terms of excitability in the underlying genetic circuit. There is an “excitable” core module containing positive and negative feedback loops.

$$\begin{aligned}\frac{dK}{dt} &= a_k + \frac{b_k K^n}{k_0^n + K^n} - \frac{K}{1+K+S} \\ \frac{dS}{dt} &= \frac{b_s}{1+(K/k_1)^p} - \frac{S}{1+K+S} + \xi(t)\end{aligned}$$

Excitable dynamics driven by noise naturally generate stochastic and transient responses

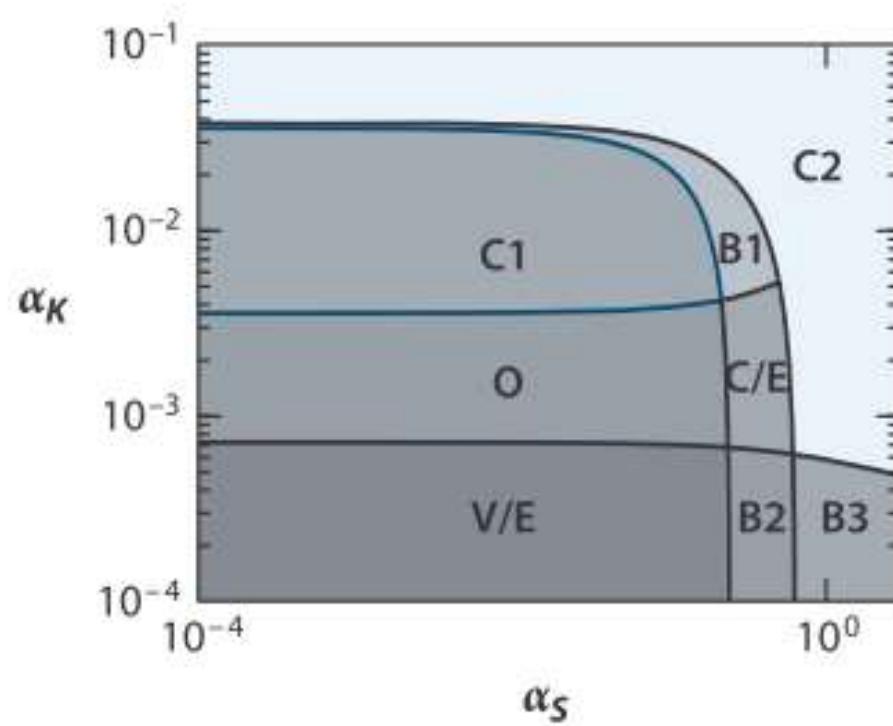
$$\frac{dK}{dt} = a_k + \frac{b_k K^n}{k_0^n + K^n} - \frac{K}{1+K+S}$$

$$\frac{dS}{dt} = \frac{b_s}{1+(K/k_1)^p} - \frac{S}{1+K+S} + \xi(t)$$



An ideal mechanism for competence regulation. A check that this is the correct mechanism was made by interfering with the circuit and showing the bacteria could be “blocked” in the competent state.

Diverse and apparently very complex dynamics can be understood as regimes in a two dimensional parameter space.



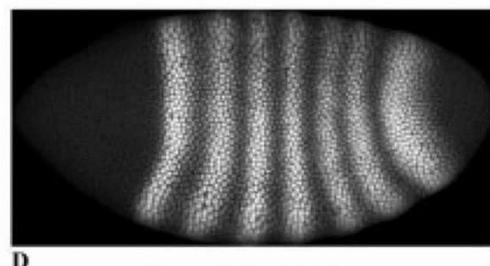
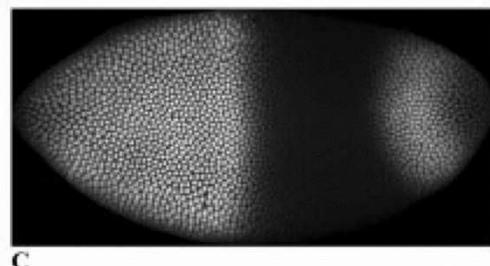
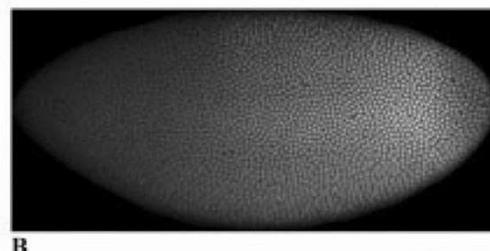
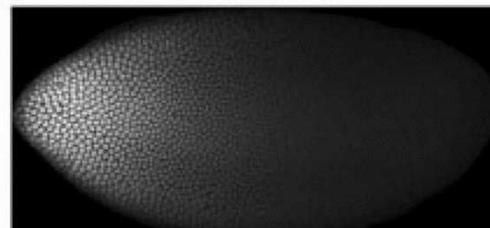
Tying some threads together, and getting closer to research themes

Taming protein production bursts

Noise in dynamical systems

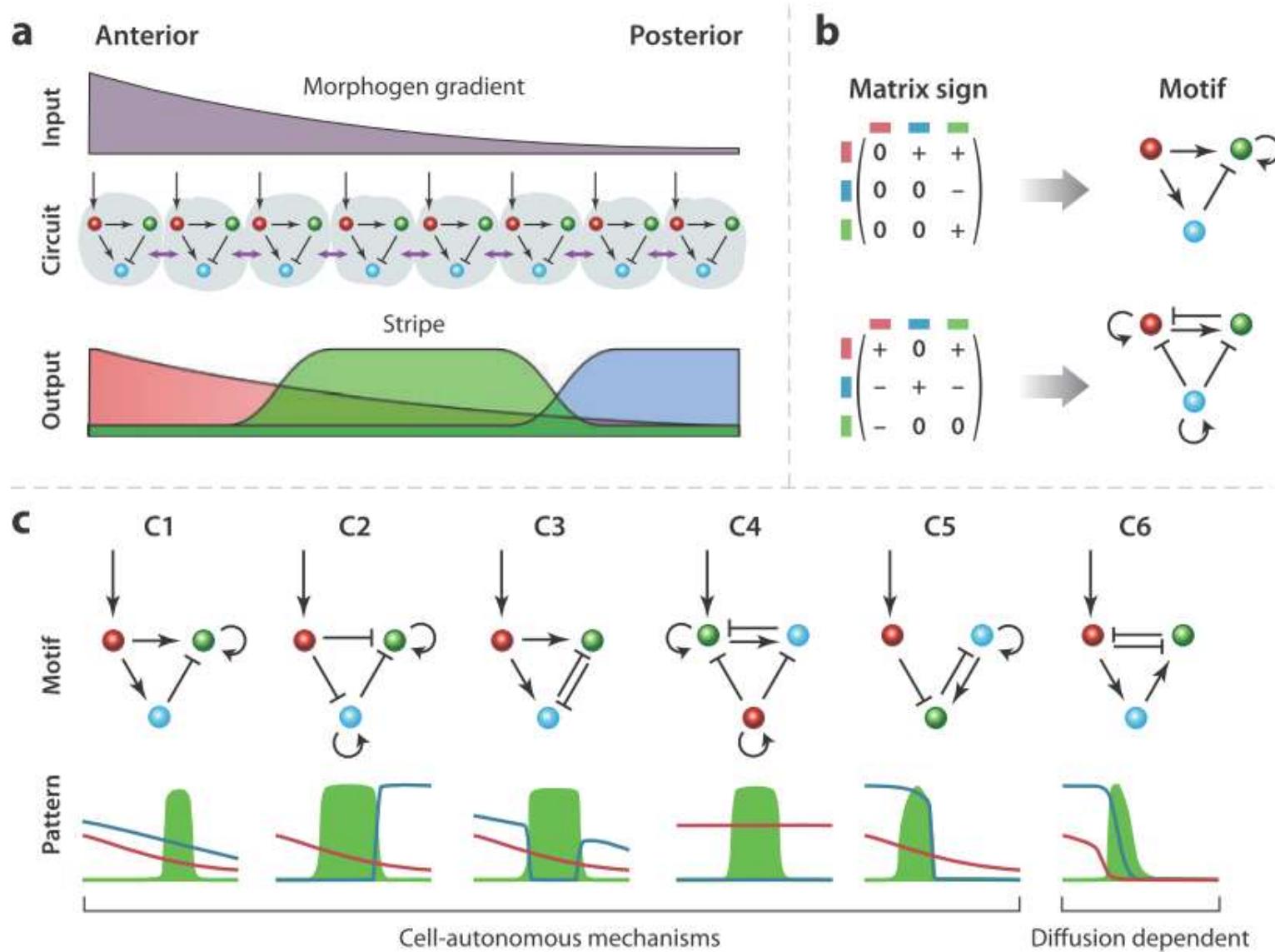
Beyond model experiments in bacteria

What goes on in development ? A classic example of patterning via morphogen gradients is segment determination in the early embryo of *D. melanogaster*.



Patterns of expressions
of different genes.

Morphogen gradients drive the formation of spatial patterns of expression of the gap genes.



Cotterell J, Sharpe J. 2010. An atlas of gene regulatory networks reveals multiple three-gene mechanisms for interpreting morphogen gradients. Mol. Syst. Biol. 6:425