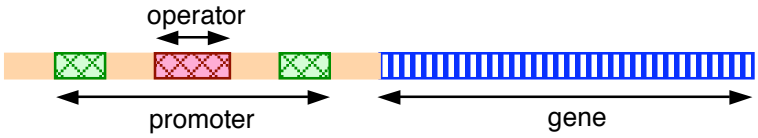
4xx Control 4 - Molecular basis of regulation

The cell cycle involves changes in cell shape and activity, which means that the concentrations of various proteins or other molecular constituents of the cell may change with time, particularly for eukaryotes. Consequently, mechanisms must be available to control the production rate for many of the cell's molecular products, both individually and as members of a group involved in a collective activity.

Transcription control in prokaryotes

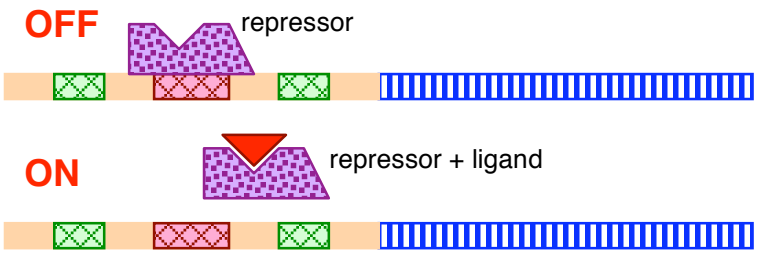
At the molecular level, the first step in the transcription process is the attachment of RNA polymerase to double-stranded DNA at a *promoter* site



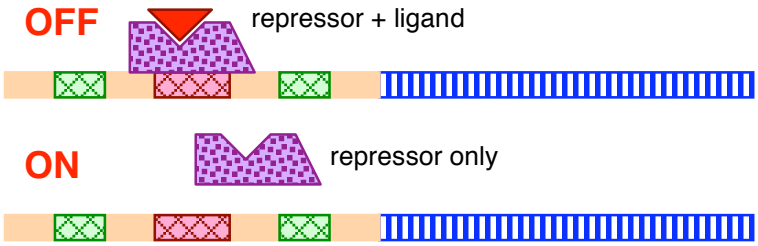
For transcription to occur, RNA polymerase first encompasses the double helix at the promoter, then opens the helix and begins to read along the gene (moving to the right in the diagram). In some situations in bacteria, a single promoter is the start point for a series of related genes, all transcribed sequentially onto the same mRNA; in such instances, the promoter and associated genes are collectively called an *operon*. Transcription can be influenced by the attachment of *regulatory proteins* that serve as sensors to their environments: for example, they may be able to bind a particular chemical compound or *ligand*.

Under *negative* control, a *gene repressor protein* (also called a *transcriptional repressor*) attaches to a site called the *operator* and is able to block the attachment of RNA polymerase. The operator provides a ligand-sensitive switch that can function in one of two ways:

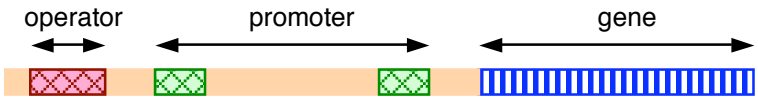
Example 1 - negative control The repressor binds in the absence of the ligand, turning the gene off. Then, when the ligand binds to the repressor, the pair is released from the DNA helix and the gene turns on.



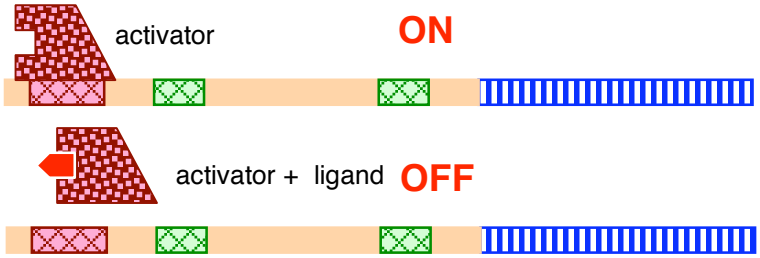
Example 2 - negative control The repressor binds in the presence of the ligand, turning the gene off. Then, when the ligand is absent, the repressor is released from the helix and the gene turns on.



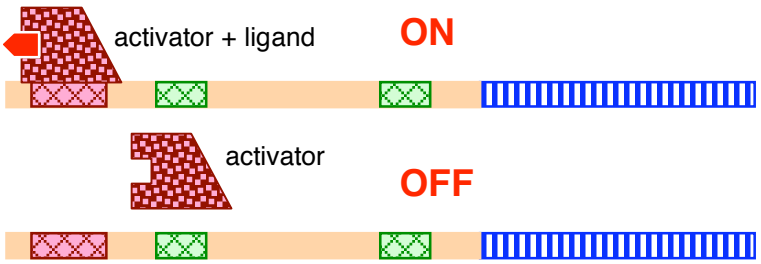
There are situations in which a bound protein can have a positive effect by enhancing the function of RNA polymerase, for example by helping it open the DNA helix. Such proteins are called *gene activator proteins* (or *transcriptional activators*), and they can function as switches if their ability to bind to DNA is affected by the presence/absence of a ligand.



Example 3 - positive control The activator binds to DNA in the absence of the ligand, turning the gene on. Then, when the ligand binds to the activator, the pair is released from the helix and the gene turns off or the frequency of its transcription is dramatically curtailed.



Example 4 - positive control The activator binds to DNA in the presence of the ligand, turning the gene on. When the ligand is not present, the activator is released from the helix and the gene turns off.



Note the reversal of the on-off states here compared to the negative control situations. Even though binding sites may be well separated along the helical contour, they may be spatially close if the helix forms a loop.

We now begin to describe the regulation mechanisms in Examples 1 - 4 in terms of the mathematical formalism developed in lecture Control 2 for coupled equations. Suppose that we can isolate the production and degradation of two particular proteins A and B from the remaining biochemical pathways in the cell. If these proteins are present at some initial concentration in the cell, but are not produced or destroyed during the cell cycle, then their concentration will decline as the cell volume expands. The rate at which the concentration falls is proportional to the concentration itself at any given time, so that $dc_A/dt = -\lambda c_A$. This equation will also apply if the protein is degraded by additional means, but the rate constant λ will be different and may depend on the concentration of the species, protein or otherwise, that removes protein A from the system. Next, let protein A also be produced in the cell at a rate γ, independent of the cell volume. As a result, the rate equation will be modified to read $dc_A/dt = -\lambda c_A + \gamma$, where the plus sign indicates A is being produced, not destroyed.

As a further extension, let another protein B act as a repressor to A, so that the rate of production γ will be reduced according to the probability p_B that B can bind to the operator site of the gene that codes for protein A. To accommodate the repressor, the equation governing dc_A/dt becomes

$$dc_A/dt = -\lambda c_A + \gamma(1 - p_B), \qquad (1a)$$

where the factor $(1 - p_B)$ means that the more likely the binding of repressor B, the smaller the production rate $\gamma(1 - p_B)$. The possibility of switch-like behavior arises if there is another pathway in which A acts as a repressor to the production of B, so that the time evolution of species B obeys

$$dc_B/dt = -\lambda c_B + \gamma(1 - p_A), \qquad (1b)$$

where the same degradation rate λ and production rate γ have been used in both equations for simplicity.

We need to determine the functional form of the binding probabilities p_A and p_B. The probability of a ligand at concentration c_L to bind to a receptor is given by the Hill function [Eq. (9.93) of *Mechanics of the Cell*]

$$p_{bound} = (c_L/c_o)^n \exp(-\beta\Delta\varepsilon) \,/\, \{1 + (c_L/c_o)^n \exp(-\beta\Delta\varepsilon)\}, \qquad (2)$$

where n is the Hill coefficient. As applied to the two-species transcription problem we have formulated, the reference concentration c_o, receptor-ligand binding energy $\Delta\varepsilon$, and inverse temperature β are all fixed, so they can be wrapped together into a single constant K as

$$K = \exp(-\beta\Delta\varepsilon) / c_o^{\,n}, \quad (3)$$

permitting us to rewrite p_{bound} as

$$p_{bound} = Kc_L^{\,n} / (1 + Kc_L^{\,n}). \quad (4)$$

Thus,

$$1 - p_{bound} = 1 / (1 + Kc_L^{\,n}), \quad (5)$$

and Eq. (1) can be expressed as

$$dc_A/dt = -\lambda c_A + \gamma/(1 + Kc_B^{\,n}) \quad (6a)$$

$$dc_B/dt = -\lambda c_B + \gamma/(1 + Kc_A^{\,n}). \quad (6b)$$

Once again, the combination of terms in K is taken to be the same for species A and B to simplify the mathematics.

Leaving aside the Hill coefficient, there are three parameters in Eq. (6), two of which can be absorbed into the definition of the concentration c and time t, leaving only a single combination $\gamma K^{1/n}/\lambda$. Thus, Eq. (6) has the same coupled structure as seen in the system in Control 2, where we showed that under weak coupling (small values of $\gamma K^{1/n}/\lambda$ here), the equations permitted only a single stable solution, whereas at large coupling, there were two stable solutions and the system possessed switch-like behavior. The stable solutions in the switch regime are asymmetric, in which one of the protein concentrations is large and the other is small, depending on the initial conditions of the system.

This genetic switch is not the same as an oscillator; rather, it is a system that can be driven between two different states. However, a simple extension of the two-component switch model to include a third repressor leads to oscillatory behavior in some ranges of its parameter space. The network, dubbed a repressilator, involves three repressor proteins interacting in a loop:

- Protein A represses the expression of protein B
- Protein B represses the expression of protein C
- Protein C represses the expression of protein A.

To describe the concentrations of each protein and its corresponding mRNA requires six coupled rate equations of the generic form:

$$dm_j/dt = -\lambda_m m_j + \gamma/(1 + Kp_{j-1}^{\,n}) + \gamma_o \quad (7a)$$

$$dp_j/dt = -\lambda_p p_j + \zeta m_j. \quad (7b)$$

The index $j = 1 - 3$ refers to species A, B and C periodically (that is, $j = 0$ is species C). Here, all proteins have the same rate parameters, as do all types of mRNA: the rates do not depend on the protein species A, B or C. However, the degradation rates (λ_m and λ_p) and production rates (γ and ζ) are different for mRNA and proteins. Other than the fact that twice as many equations are required for determining the concentrations of

both protein and mRNA, the only new term to appear here is γ_o, which allows for the production of each protein even when its associated repressor concentration is at saturation.

The appearance of these equations can be improved with the usual conversion to dimensionless variables, as done in reducing Eq. (6). The six equations have a large parameter space, but the starting point for their solution is the same as that outlined in Control 2: find the steady state solutions and then evaluate their stability and other properties. Some of the steps in the process are easy, others require numerical evaluation. For example, imposing $dp_j/dt = 0$ on Eq. (7b) immediately yields

$$\lambda_p p_{j,ss} = \zeta m_{j,ss}, \tag{8}$$

for the steady-state concentrations of each of the three protein - mRNA pairs. To find the individual concentrations, we impose dm/dt on Eq. (13.51a) to obtain

$$\lambda_m m_j = \gamma/(1 + K p_{j-1}^{\ n}) + \gamma_o \tag{9}$$

Substituting Eq. (8) gives

$$m_j = (\gamma/\lambda_m)/[1 + K(\zeta m_{j-1}/\lambda_p)^n] + (\gamma_o/\lambda_m) \tag{10}$$

Each m_j is related to its neighbor through an equation like this, so that m satisfies the iterative equation $m_j = H\{H[H(m_j)]\}$ for a function $H(m)$. Hence, $m_j = H(m_j)$ is an allowed solution for monotonically decreasing $H(m)$, and the steady-state solution satisfies

$$m = (\gamma/\lambda_m)/[1 + K(\zeta m/\lambda_p)^n] + (\gamma_o/\lambda_m). \tag{11}$$

The stability analysis of the solution can be performed as in Control 2.

The stability analysis reveals that there are large regions of parameter space where the steady-state solutions become unstable and, as a result, oscillatory (for example, see Chapter 19 of Phillips *et al.* (2008). This is in contrast to the two-protein switch model where at least one stable steady state was present for any value of the parameter α.