Genetic transcription and regulation
Central dogma of biology

DNA codes for DNA
- replication (DNA → DNA)

DNA codes for RNA
- transcription (DNA → RNA)

RNA codes for proteins
- translation (RNA → Protein)

Not surprisingly, many points for regulation of the process
DNA codes for DNA

Okazaki fragments - 100-200 nucleotides long in eukaryotes, 10x longer in bacteria!

“replisome”

error rate is 1 in $10^8 - 10^{10}$ bp!
DNA codes for RNA

Why the need for a messenger?

RNA is more flexible

additional regulation

DNA is sequestered in the nucleus in eukaryotes (protected)

and...

AMPLIFICATION

RNA polymerase

Roger Kornberg received the 2006 Nobel Prize (Chemistry) for structures
DNA codes for RNA

Can get MANY copies simultaneously from the same DNA strand
Transcription starts at a promoter

promoter is a sequence recognized by the polymerase upstream of the transcription initiation site consensus sequence in bacteria: TATAAT @ -10 bps and TTGACA @ -35 bps (always just a fraction present though!).

RNA polymerase is promiscuous, can bind non-specifically but with lower energy. Variation in promoter sequence can lead to strong or weak promoters.

Transcription factors are often required for binding

\[
p_{\text{bound}} = \frac{1}{1 + \frac{N_{NS}}{P} e^{\beta \Delta \epsilon_{pd}}} = \frac{(N_{NS})^P}{P!} e^{-P \epsilon_{NS} / k_{B}T} = \frac{(N_{NS})^{P-1}}{(P-1)!} e^{-(P-2)\epsilon_{NS} / k_{B}T} e^{\epsilon_{pd} / k_{B}T}
\]

\[
\frac{N_{NS}!}{P! (N_{NS}-P)!} = \frac{(N_{NS})^P}{P!}
\]

\[
\frac{N_{NS}!}{(P-1)! [N_{NS}-(P-1)]!} = \frac{(N_{NS})^{P-1}}{(P-1)!}
\]

\[
\frac{(N_{NS})^{P-1}}{(P-1)!} e^{-(P-2)\epsilon_{NS} / k_{B}T} e^{\epsilon_{pd} / k_{B}T}
\]
The *lac* operon

*Operon* is just a cluster of genes with one promoter.

---

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

- **promoter sequences**
- **activator binding site**
- **repressor binding site**
- **no lactose**
- **no glucose**

---

**PBoC 4.4.1**
Gene regulation

- Genes can be regulated by controlling polymerase access to promoter.

**Gene repression**

+ Glucose + Lactose: OPERON OFF because CAP not bound.

+ Glucose - Lactose: OPERON OFF both because Lac repressor bound and because CAP not bound.

- Glucose - Lactose: OPERON OFF because Lac repressor bound.

- Glucose + Lactose: RNA polymerase and mRNA indicate OPERON ON.

**Gene activation**

- About 10% (1500-3000) of our genes code for transcription factors!
Gene activation

instead of two, now have four distinct states
also have three relevant interaction energies (P-DNA, A-DNA, P-A)

What is \( p_{\text{bound}} \) for the promoter?

\[
p_{\text{bound}} = \frac{1}{1 + \left[ \frac{N_{NS}}{P F_{\text{reg}}(A)} \right] e^{\beta \Delta \epsilon_{pd}}} \\
F_{\text{reg}}(A) = \frac{1 + (A/N_{NS}) e^{-\beta \Delta \epsilon_{ad}} e^{-\beta \Delta \epsilon_{ap}}}{1 + (A/N_{NS}) e^{-\beta \Delta \epsilon_{ad}}} 
\]
Genetic switches

A genetic switch is a repressor and the gene it controls. Most interesting are cases where two genes regulate each other’s expression in a feedback loop.
Genetic switches

\[ p_{\text{bound}}(c_1) = \frac{K_b c_1^n}{1 + K_b c_1^n} \]

Hill function - \( n \) is cooperativity, i.e., number required for reaction

\[ r(1 - p_{\text{bound}}(c_1)) = \frac{r}{1 + K_b c_1^n} \]

\( r \) is basal (default) rate in absence of repressor

\[
\begin{align*}
\frac{dc_1}{dt} &= -\gamma c_1 + \frac{r}{1 + K_b c_2^n} \\
\frac{dc_2}{dt} &= -\gamma c_2 + \frac{r}{1 + K_b c_1^n}
\end{align*}
\]

degradation rate
basal rate of production
due to repression

\[
\begin{align*}
\frac{du}{dt} &= -u + \frac{\alpha}{1 + \nu^n} \\
\frac{d\nu}{dt} &= -\nu + \frac{\alpha}{1 + u^n}
\end{align*}
\]

make dimensionless by change of variables

PBoC 19.3.5
Genetic switches

\[
\frac{du}{dt} = -u + \frac{\alpha}{1 + \nu^2}
\]

\[
\frac{d\nu}{dt} = -\nu + \frac{\alpha}{1 + u^2}
\]

**steady state**

\[
(u^2 - \alpha u + 1)(u^3 + u - \alpha) = 0
\]

- \((u^2 - \alpha u + 1)\) two zeroes for \(\alpha > 2\)
- \((u^3 + u - \alpha)\) one real zero

\[\frac{-\nu + \frac{1}{u}}{u} = 0 \rightarrow uv = 1\]

\[u = \nu\]

**not switch-like!** (concentrations are always the same)

assume \(n = 2\)
Stability analysis

vectors denote \((\frac{du}{dt}, \frac{dv}{dt})\)

\(\alpha = 1\)

\(\alpha = 3\)

\[a = rK_b^{1/n} / y\]

filled - stable

unfilled - unstable

Stable equilibrium

for \(\alpha > 2\), observe bistable behavior (two dominant states)
**Genetic circuits and clocks**

*Synthetic circuit* could be used to, e.g., report on concentrations of multiple chemicals in environment.


**Genetic clock** the regulated expression of multiple genes generates ~24h molecular oscillations establishing the circadian rhythm.


2017 Nobel Prize in Physiology or Medicine awarded to Jeff Hall, Michael, Rosbash, & Michael Young.
Summary

Regulation of gene expression at transcription level is important

Controlled by access of promoter sequence to RNA-polymerase by activators are repressors

Ultimately binding energies affect the probability of RNA-polymerase binding to the promoter

Interplay between expression of various genes: genetic switches, genetic circuits, genetic clocks