Lecture 10: Genomes

Goal: Model to obtain insights into how the same DNA sequence can result in a diversity of gene expression profiles

https://youtu.be/nQQJNIJbzd8?si=1QLyWGcKHZ0pvAi1

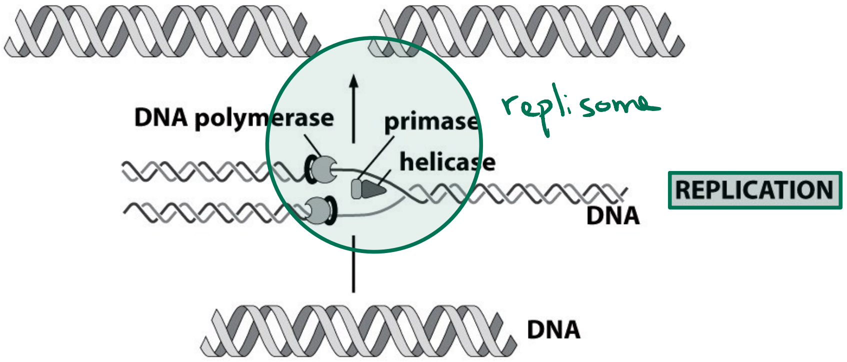
- Central dogma of molecular biology
- Genome regulation
- Genetic networks

PBOC Chapter 3.2.1, 6.1.2 (refresh Lecture 06), 19.2 (except 19.2.5)

Central dogma of molecular biology

Recall (Lecture 1, 6)

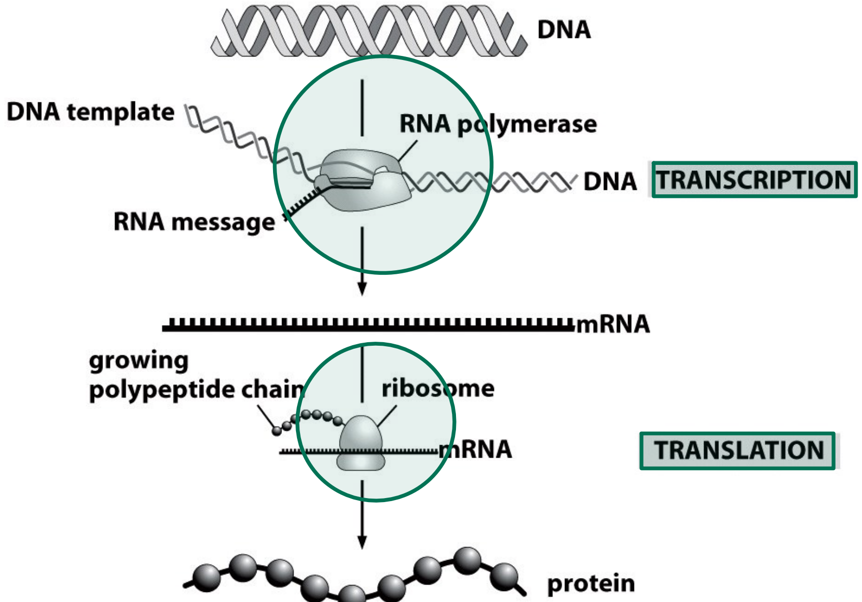
Processes of the central dogma



DNA → DNA

replisome: DNA polymerase + primase + helicase

https://youtu.be/TNKWgcFPHqw?si=lGHHDFNOwNdEYNT7

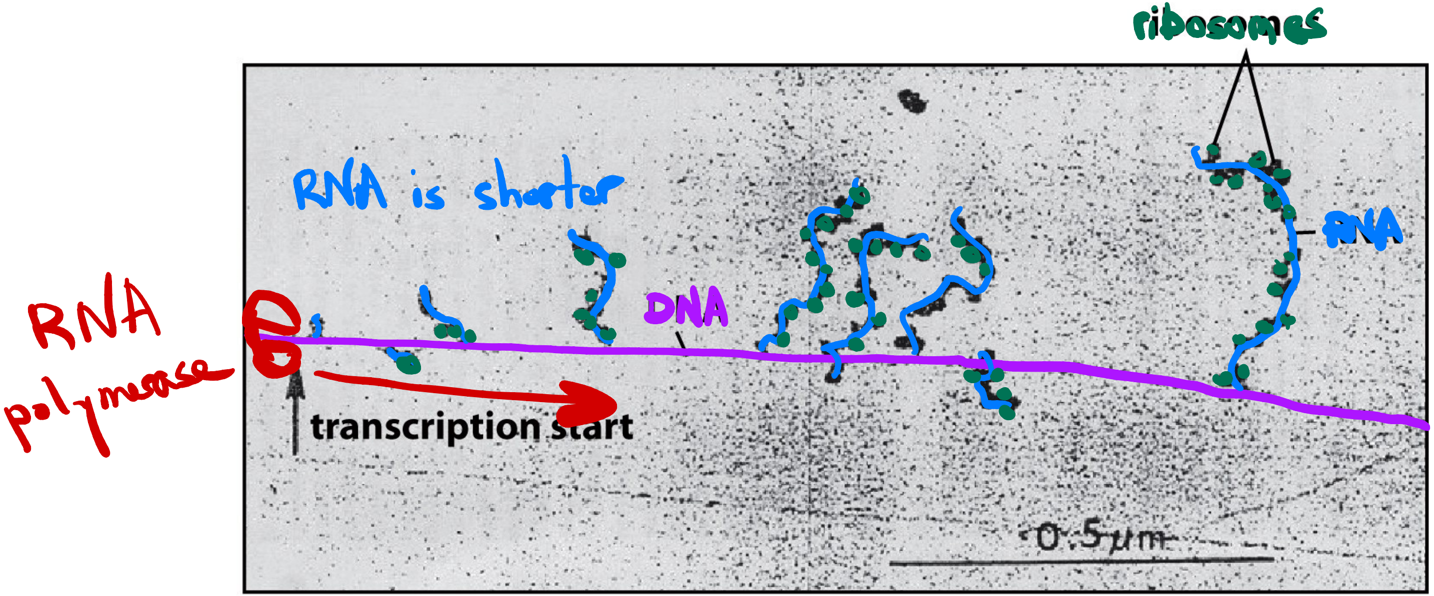


DNA → mRNA → protein

https://youtu.be/gG7uCskUOrA?si=c6mmpHITsrmIUtaF

Central dogma of molecular biology

Processes of the central dogma



In bacteria, ribosomes can bind to RNA even before it has finished being transcribed

Central dogma of molecular biology

Processes of the central dogma

What are the rates of processes involved in the central dogma?

DNA replication

Given replication rate per replisome of 250-1000 bp/s, and *E. coli* genome size of 5×10^6 bp, what range of times should it take to replicate the genome?

transcription

Given an average protein molecular weight of 46 kDa, an average amino acid molecular weight of 110 Da, and a transcription rate of 40 nucleotides/s, how long should it take on average to make an mRNA transcript?

translation

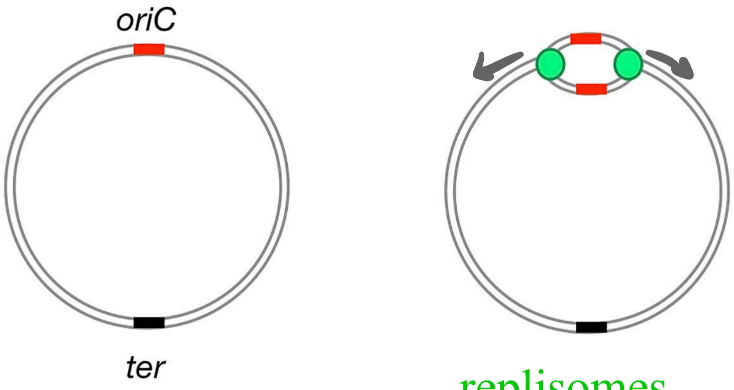
Given an E. coli cell cycle time of 3,000 seconds, and the facts that the number of proteins must double during this time, and that there are ~20,000 ribosomes per cell, what is the rate of translation per ribosome (units aa/second)? If needed, look back at Lecture 2 where we estimated the number of proteins per cell.

Central dogma of molecular biology

Processes of the central dogma: Replication

Given replication rate per replisome of 250-1000 bp/s, and *E. coli* genome size of 5×10^6 bp, what range of times should it take to replicate the genome?

bacteria contain a single circular genome



replisomes

(2 of them work simultaneously)

$$\frac{5 \times 10^6 \text{ bp}}{2 \cdot 250 \text{ bp/s}} = 10^4 \text{ s}$$

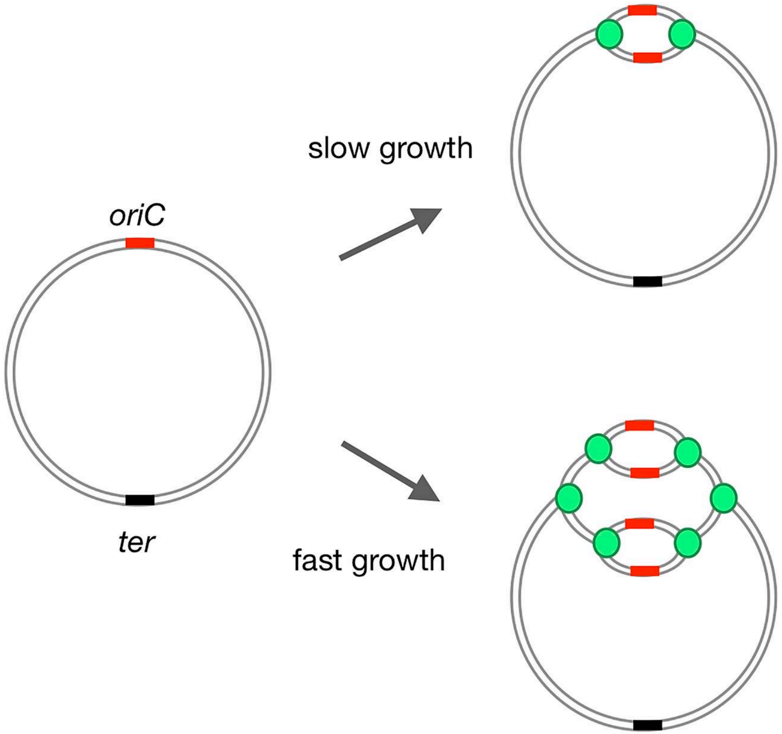
$$\frac{5 \times 10^6 \text{ bp}}{2 \cdot 1000 \text{ bp/s}} = 2.5 \times 10^3 \text{ s}$$

one to a few hours.

Central dogma of molecular biology

Processes of the central dogma: Replication

E. coli can divide in much less than 3000 seconds, in fact, as little as 1000 seconds

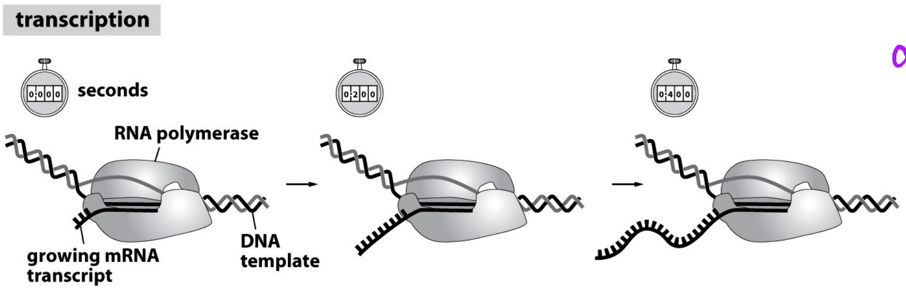


multiple replication forks
⇒ more than 2 replisomes

Central dogma of molecular biology

Processes of the central dogma: Transcription

Given an average protein molecular weight of 46 kDa, an average amino acid molecular weight of 110 Da, and a transcription rate of 40 nucleotides/s, how long should it take on average to make an mRNA transcript?



average protein: $\frac{4.6 \times 10^4 \text{ Da/protein}}{110 \text{ Da/aa}} = 420 \text{ aa/protein}$

Each aa is encoded by three nucleotides (a "codon")

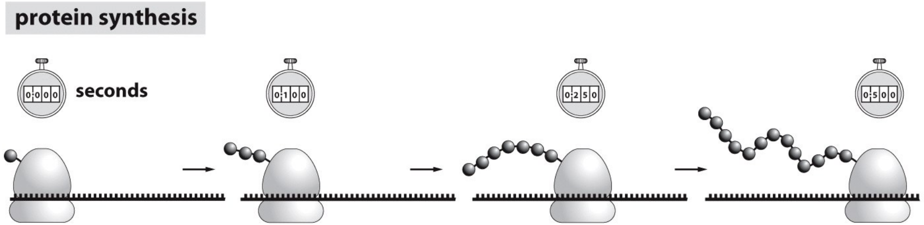
$\left(1260 \text{ nucleotides/protein}\right) / \left(40 \text{ nucleotides/s}\right)$

≈ 31 seconds

Central dogma of molecular biology

Processes of the central dogma: Translation

Given an E. coli cell cycle time of 3,000 seconds, and the facts that the number of proteins must double during this time, and that there are ~20,000 ribosomes per cell, what is the rate of translation per ribosome (units aa/second)? If needed, look back at Lecture 2 where we estimated the number of proteins per cell. 3×10^6 proteins/cell



$$\frac{3\times10^6 \text{ proteins}}{2\times10^4 \text{ ribosome}} = 150 \frac{\text{proteins}}{\text{ribosome}} \text{ (per cell cycle)}$$

$$150 \frac{\text{proteins}}{\text{ribosome}} \cdot 1260 \frac{\text{nucleotides}}{\text{protein}} \Big/ 3000 \text{ s} \approx 63 \text{ nucleotides/ribosome}\cdot\text{s} \approx 21 \text{ aa/ribosome}\cdot\text{s}$$

(last question)

or 420 aa/protein

Around 20 s to make a protein

About same time as to make mRNA.

Genome regulation

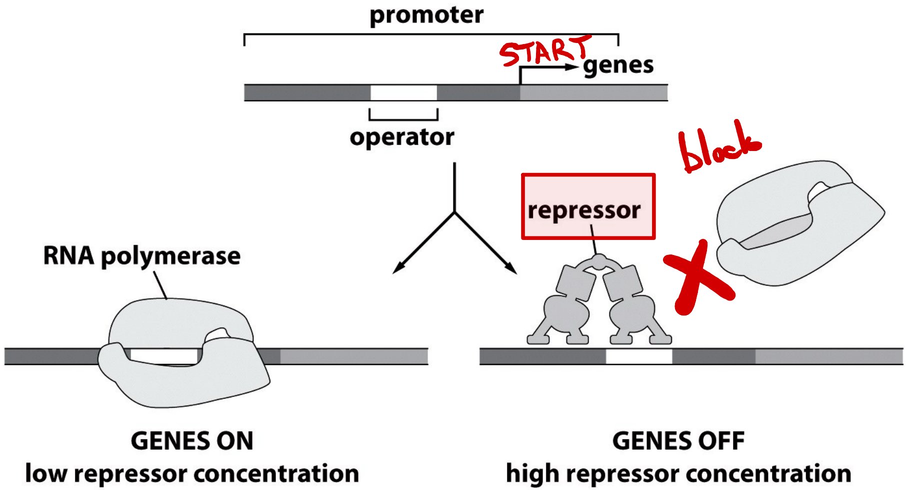
Statistical mechanics of gene expression

Review Lecture 6 annotated slides 13-20, along with PBoC 6.1.2

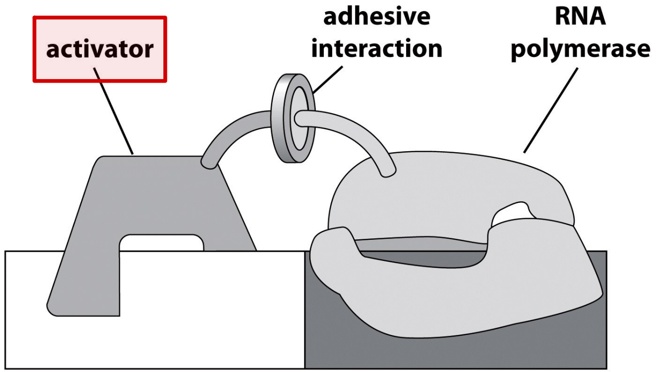
Genome regulation

Genetic networks: Molecules

negative regulation



positive regulation

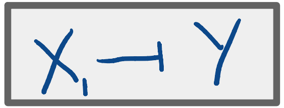


Repressors and activators change the probability of RNA polymerase binding to the promoter of a gene

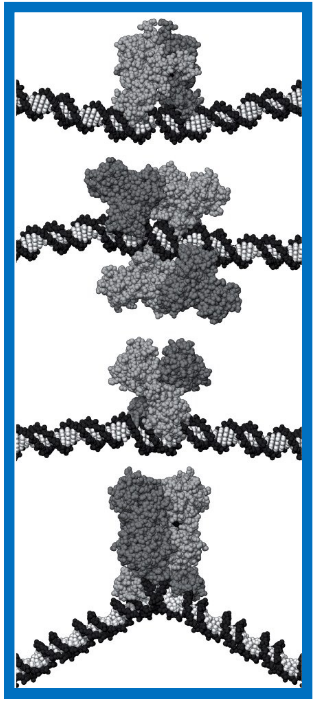
Genome regulation

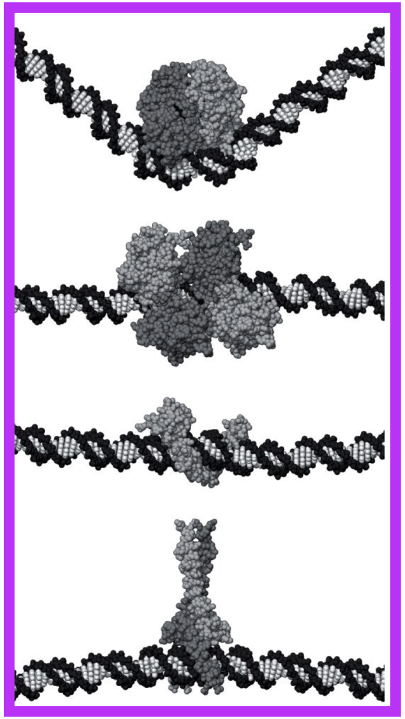
Genetic networks: Molecules

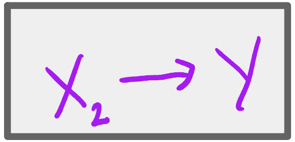
Repressors and activators can both bind to DNA and deform it in the promoter region of a gene.



Repressors bind to the promoter site, to block RNA polymerase from binding







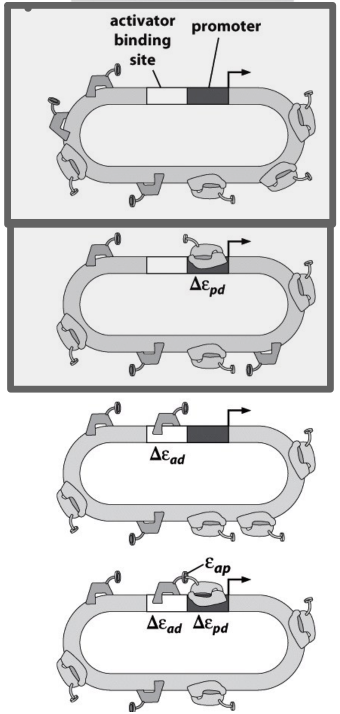
Activators bind upstream of the promoter site, to enhance RNA polymerase binding

Genome regulation

Genetic networks: Models (activation)

P # of polymerases
A # of activators
R # of repressors
N_{NS} # of boxes

renormalize: divide each weight by unoccupied state

STATE	RENORMALIZED WEIGHT	
	1	activator & promoter unoccupied
	$\frac{P}{N_{NS}} e^{-\Delta\varepsilon_{pd}/k_BT}$	promoter occupied
	$\frac{A}{N_{NS}} e^{-\Delta\varepsilon_{ad}/k_BT}$	activator occupied
	$\frac{P}{N_{NS}} \frac{A}{N_{NS}} e^{-(\Delta\varepsilon_{pd}+\Delta\varepsilon_{ad}+\varepsilon_{ap})/k_BT}$	activator & promoter occupied

Energies:

$$\Delta\varepsilon_{pd} = \varepsilon_{pd}^{S} - \varepsilon_{pd}^{NS}$$

$$\Delta\varepsilon_{ad} = \varepsilon_{ad}^{S} - \varepsilon_{ad}^{NS}$$

$$\varepsilon_{ap}$$

NS non-specific

pd polymerase-DNA

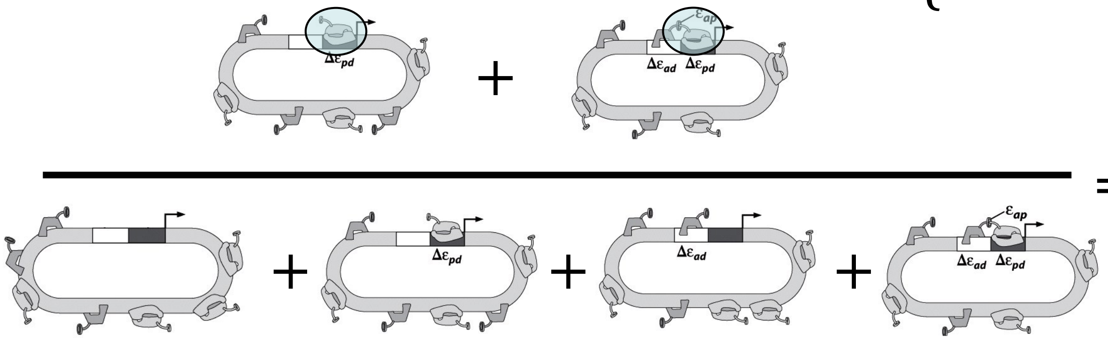
ad activator-DNA

ap activator-polymerase

Genome regulation

Genetic networks: Models (activation)

$P_{bound} =$



$= \dfrac{\text{states with RNAp bound}}{\text{all states}}$

PBOC 19.1-19.6

$$\ldots = \frac{1}{1 + \left[\frac{N_{NS}}{P F^{A}_{reg}(A)}\right] e^{\beta\Delta\varepsilon_{pd}}}$$

like changing [P]

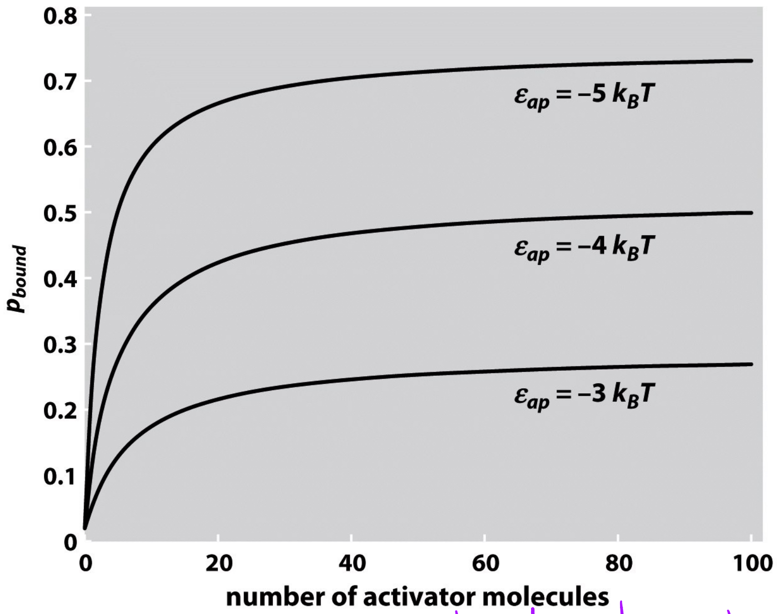
where

$$F^{A}_{reg}(A) = \frac{1 + \frac{A}{N_{NS}} e^{-\beta\Delta\varepsilon_{ad} - \beta\varepsilon_{ap}}}{1 + \frac{A}{N_{NS}} e^{-\beta\Delta\varepsilon_{ad}}}$$

takes on values > 1 for $\varepsilon_{ap} < 0$

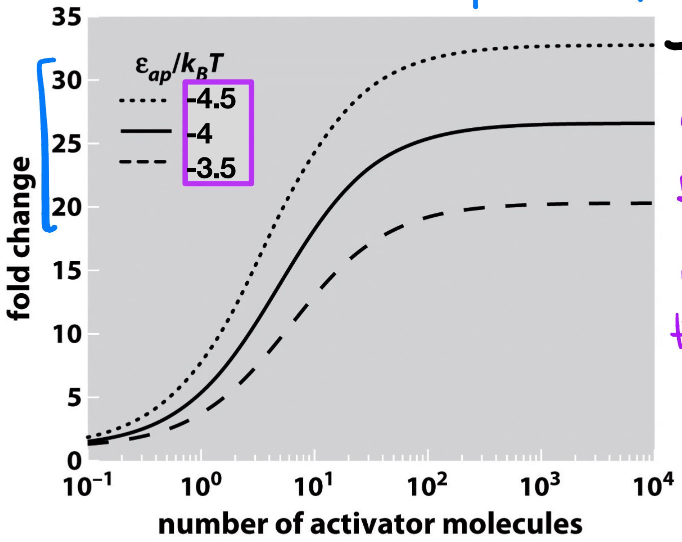
Genome regulation

Genetic networks: Models (activation)



p_{bound} increases with $|\varepsilon_{ap}|$ and [A] concentration of activator molecules

Can have much more expression w/activator



Typo: energies in Fig. 19.12 were positive, that was an error

normalized to display fold-change $p_{bound}(A)/p_{bound}(A=0)$

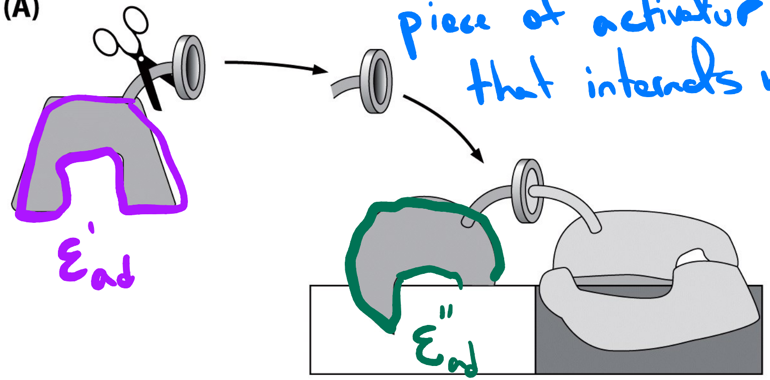
$$= \frac{1+[N_{NS}/P]e^{\beta\Delta\varepsilon_{pd}}}{1+[N_{NS}/PF_{reg}]e^{\beta\Delta\varepsilon_{pd}}} \approx \frac{[N_{NS}/P]e^{\beta\Delta\varepsilon_{pd}}}{[N_{NS}/PF_{reg}]e^{\beta\Delta\varepsilon_{pd}}} = F_{reg}$$

Genome regulation

Genetic networks: Models (activation)

How do we know? mix-and-match activator elements

(A)



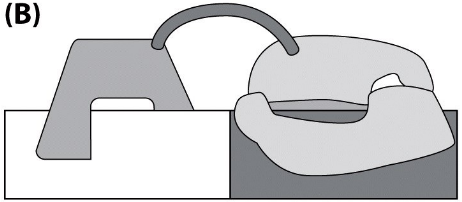
change $\Delta\varepsilon_{ad}$

Experiment:

measure p_{bound}

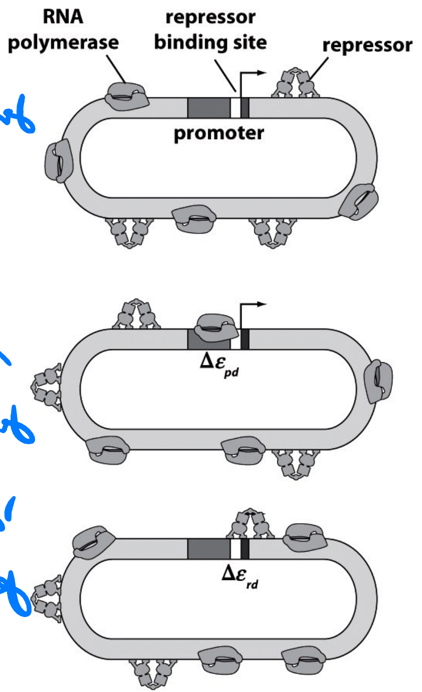
for (A) & (B)

(B)



$\varepsilon_{ap} \to \infty$, two states (1 & 4)

Genetic networks: Models (repression)



STATE	RENORMALIZED WEIGHT
no binding	1
RNAP binding	$\frac{P}{N_{NS}} e^{-\Delta\varepsilon_{pd}/k_BT}$
repressor binding	$\frac{R}{N_{NS}} e^{-\Delta\varepsilon_{rd}/k_BT}$

$$P_{bound} = \frac{1}{1 + \left[\frac{N_{NS}}{P F^{R}_{reg}(R)}\right] e^{\beta\Delta\varepsilon_{pd}}}$$

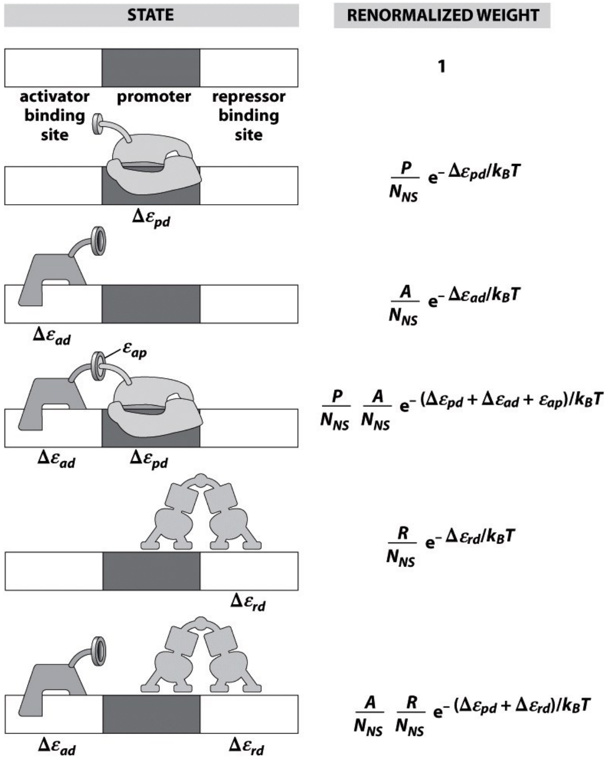
like changing [P]

where $$F^{R}_{reg}(R) = \frac{1}{1 + \frac{R}{N_{NS}} e^{-\Delta\varepsilon_{rd}}}$$

takes on values < 1

Genome regulation

Genetic networks: Models (activation + repression)



PBoC 19.2.4

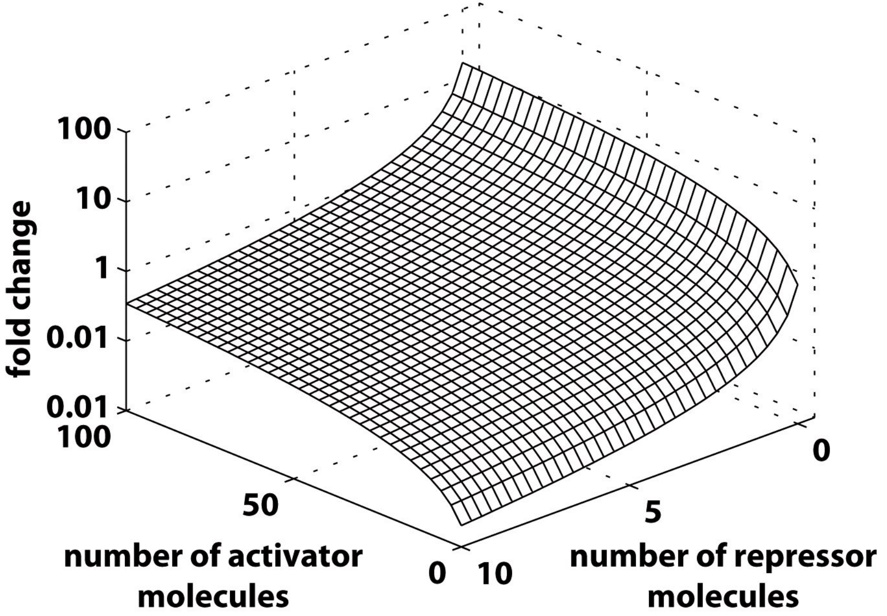
$$P_{bound} = \frac{1}{1 + \left[\frac{N_{NS}}{PF^{AR}_{reg}(A,R)}\right] e^{\beta\Delta\varepsilon_{pd}}}$$

like changing [P]

Note: $F^{A}_{reg}(A) \neq F^{R}_{reg}(R) \neq F^{AR}_{reg}(A,R)$

Genome regulation

Genetic networks: Models (activation + repression)



Lecture 10: Genomes

Summary:

- Processes related to the "central dogma of molecular biology" include transcription, translation, and replication.
- The rates of such processes are governed by enzymatic activity, and can limit achievable rates of cell division.
- Thermodynamic models of gene expression are based on estimating the probability of RNA polymerase to bind to a gene promoter.
- Activators and repressors modify the probability of binding in distinct ways, and their effects can be generalized through a regulation factor, F_{reg}.

Lecture 11: Genomes

Goal: Model interactions between gene regulatory elements, discuss common motifs in genetic circuits

Watch Prof. Uri Alon video before class
Answer guiding questions
In-class small group discussions

Reading : Alon Network Motifs