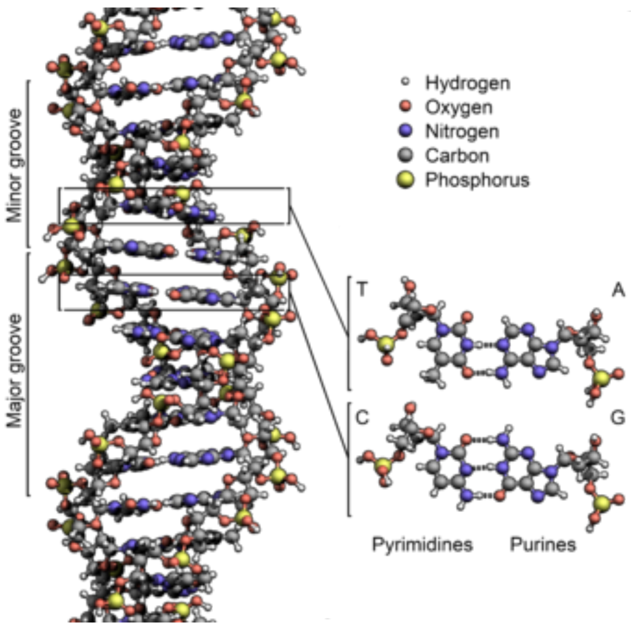
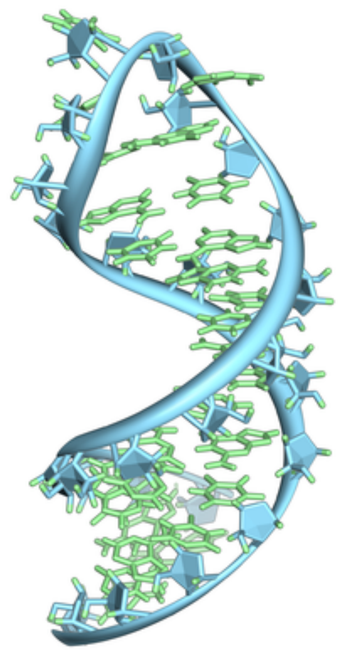
Gene Transcription Networks

Genes and Proteins

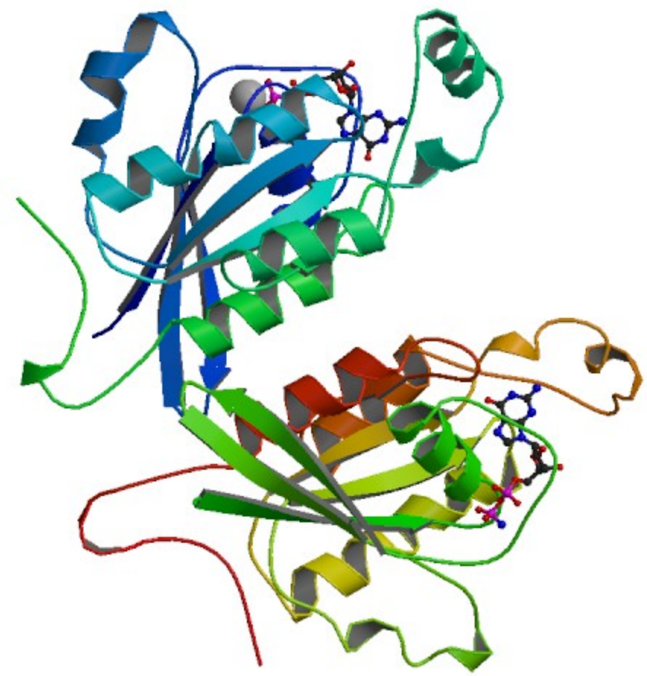
Three Key Biopolymers



Deoxyribonucleic Acid
(DNA)



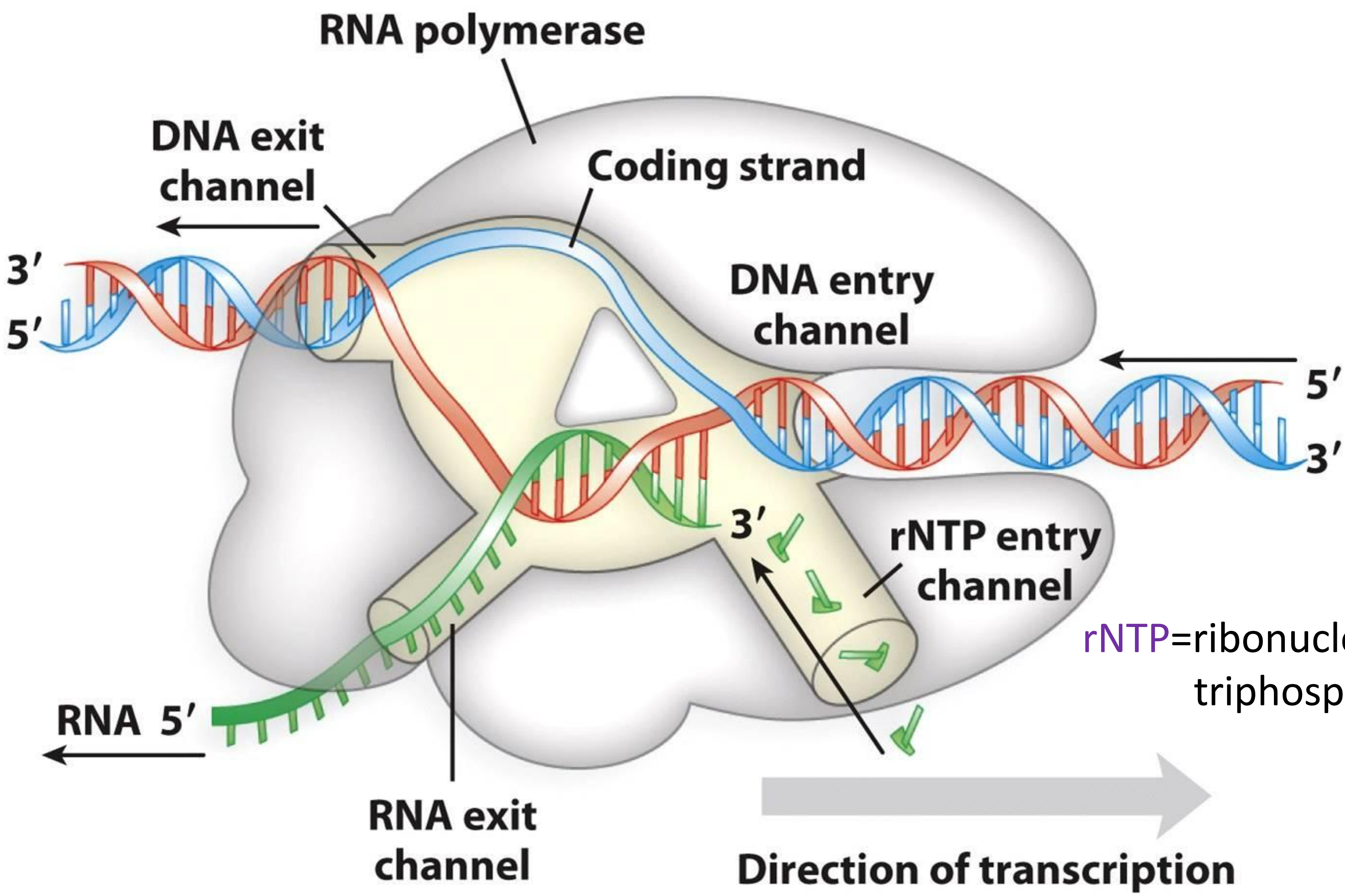
Ribonucleic Acid
(RNA)



Proteins

A polymer is a molecule made up of a sequence of simpler molecules

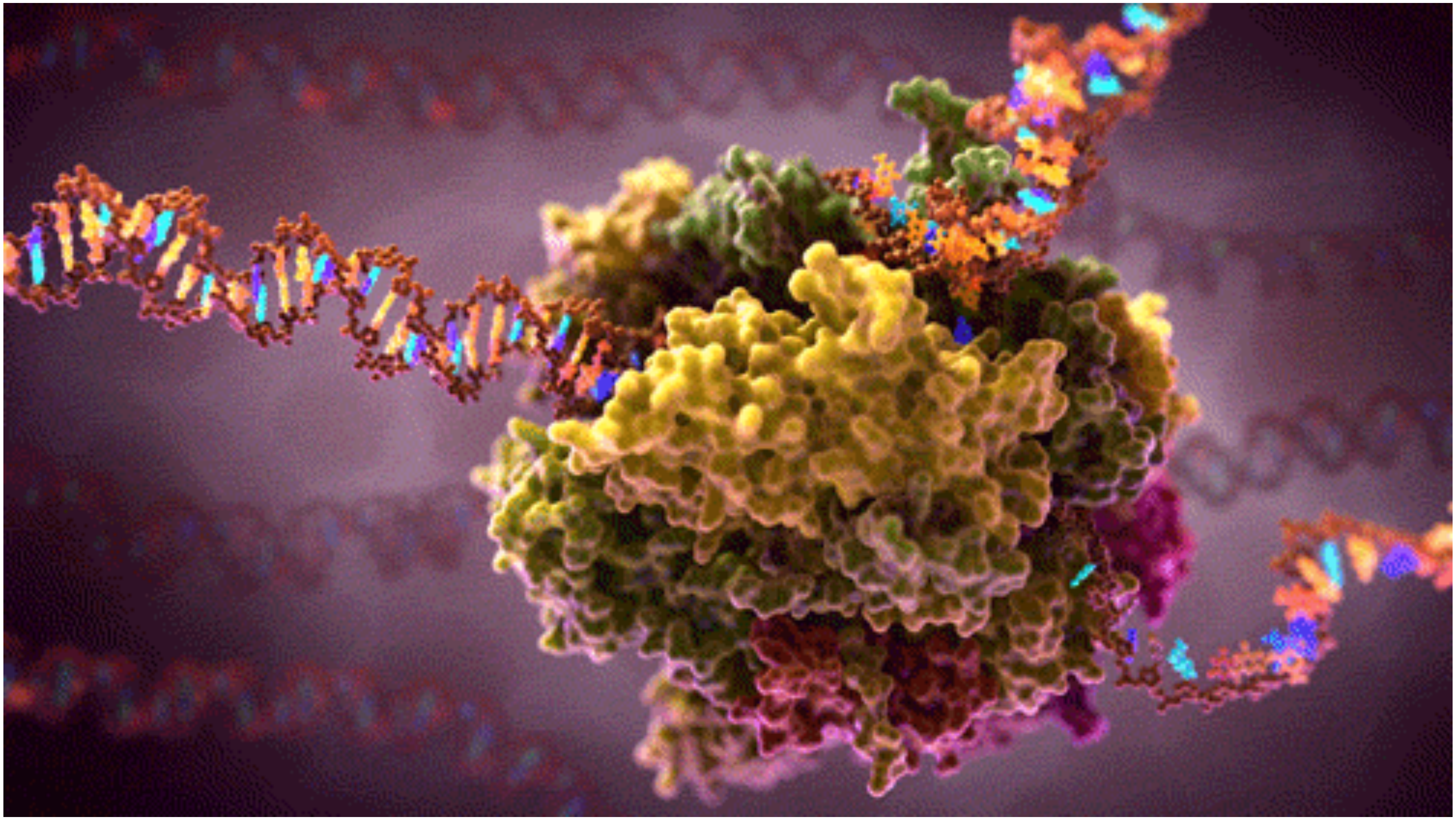
RNA Polymerase Transcribes a DNA Fragment to Messenger RNA (mRNA)



rNTP=ribonucleoside triphosphate

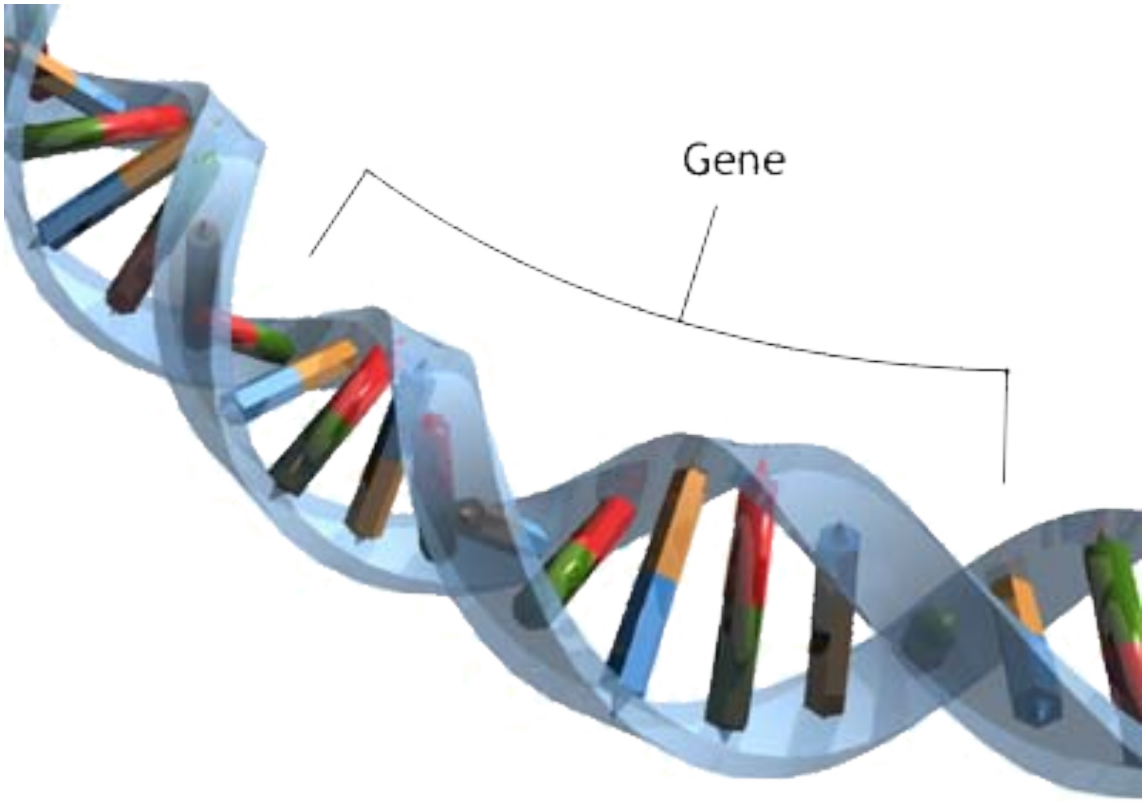
Figure 15-14
Molecular Biology: Principles and Practice

RNA Polymerase Transcribes a DNA Fragment to Messenger RNA (mRNA)



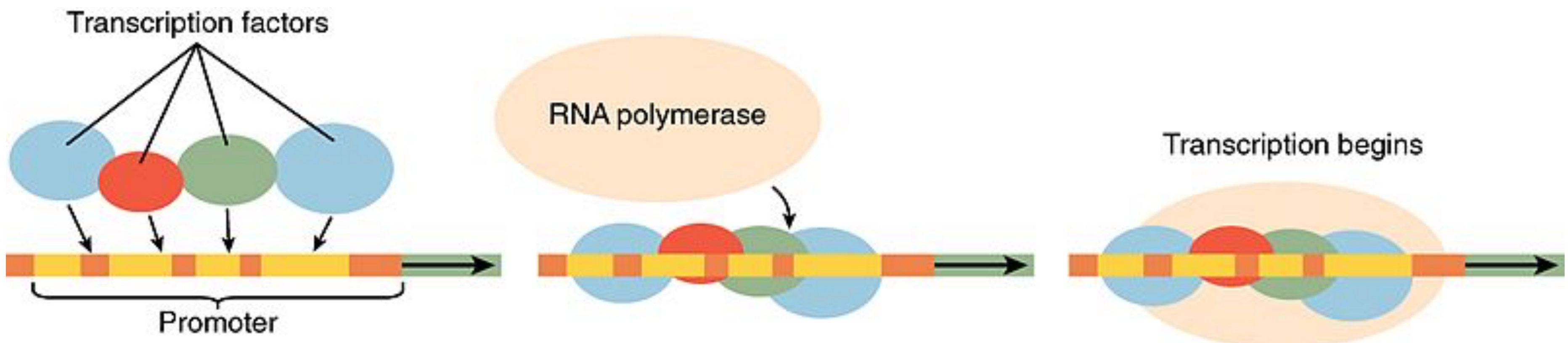
What Are Genes?

A codon is a triplet of three adjacent nucleotides . A gene is a sequence of codons between a start codon and an end codon.



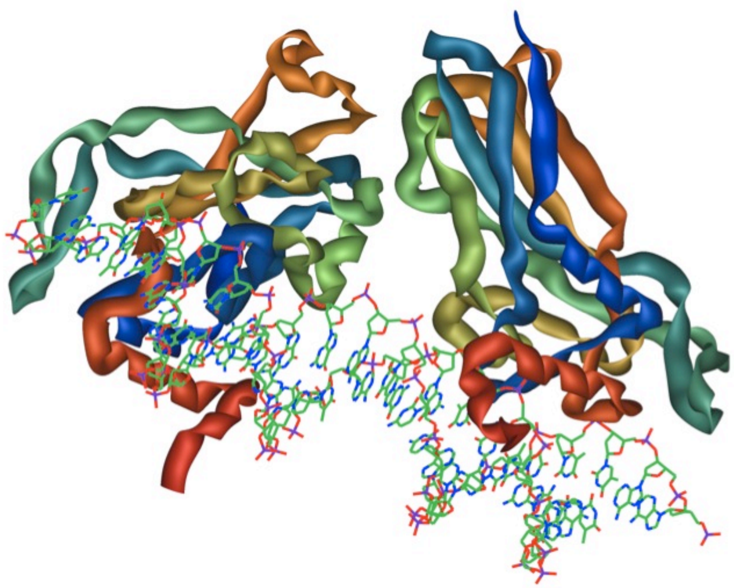
What Causes Genes to be Expressed?

A gene is expressed if it is first transcribed into mRNA and then translated into protein. This happens when a specific transcription factor binds to the promoter region for a gene. This is a segment of DNA near the gene.

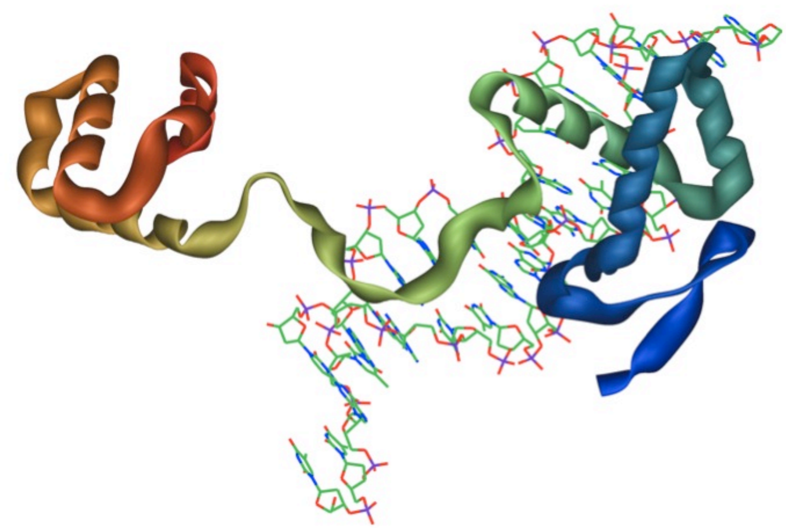


Transcription factors may be activators (promote binding of RNA polymerase) or repressors (prevent the binding of RNA polymerase).

Transcription Factors Determine Which Genes are Transcribed



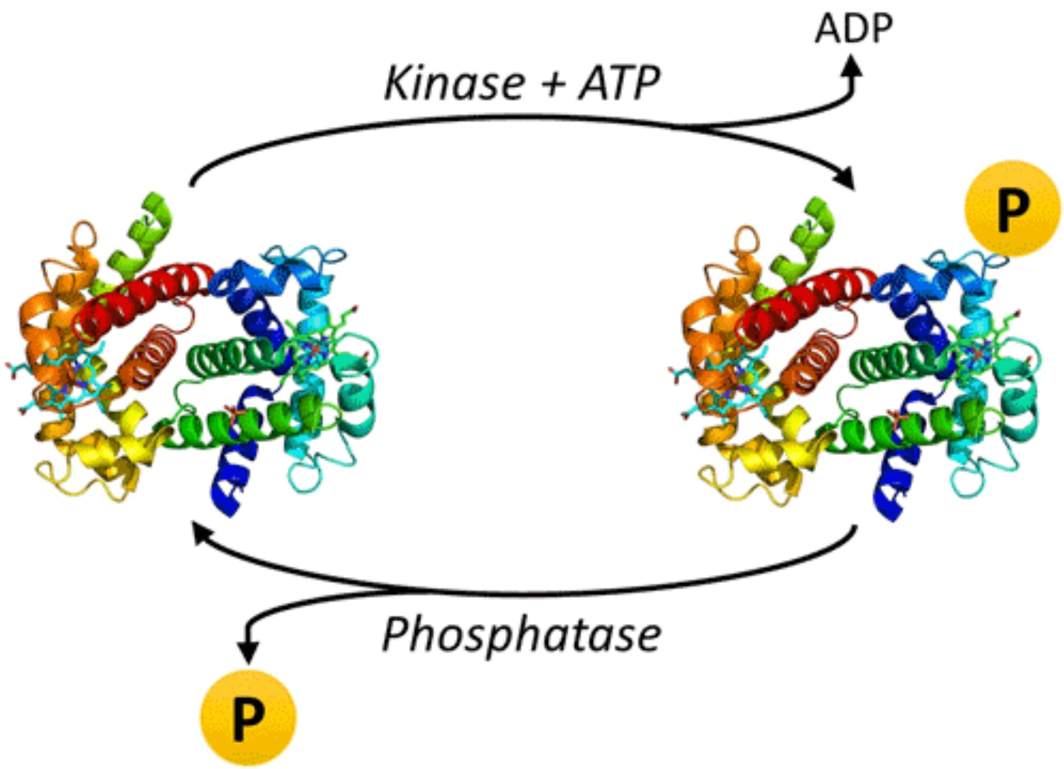
Beta-Hairpin-Ribbon Group
Transcription Factor T-Domain (TBX21)



Helix-Turn-Helix Group
Homeodomain Family (Pax8)

Two of many transcription factor proteins that exist

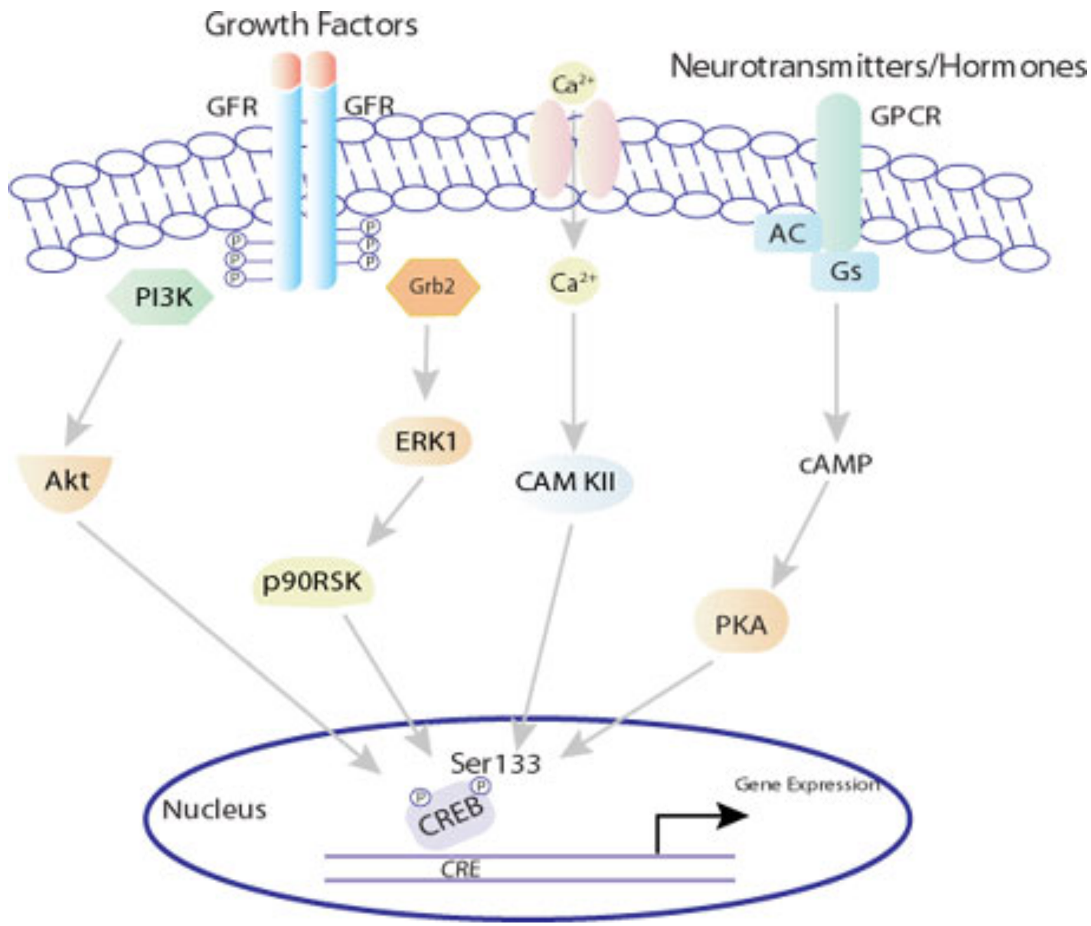
Transcription Factors are Often Activated by Phosphorylation



A kinase is an enzyme that phosphorylates a substrate.
A phosphatase is an enzyme that dephosphorylates a substrate.

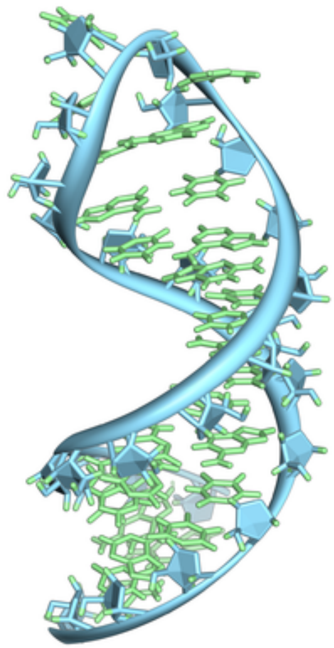
Example of Transcription Factor Activation by Inducers

Akt=protein kinase B

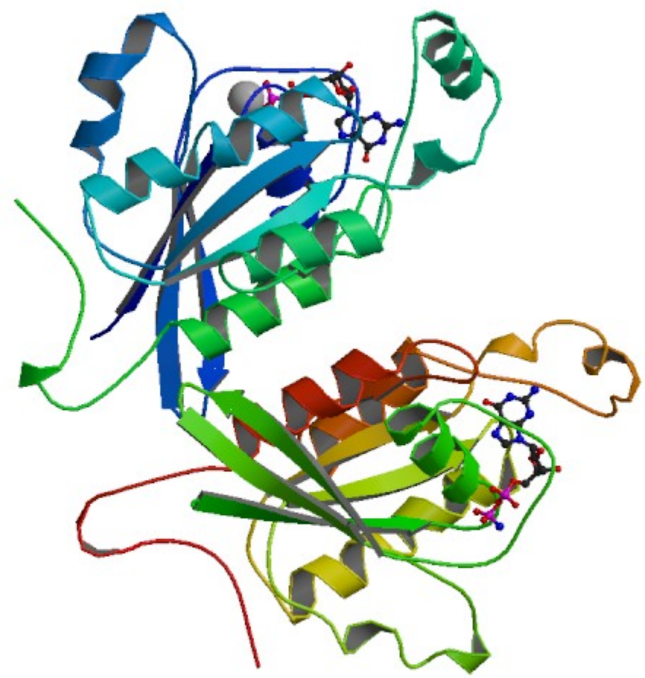


Cyclic AMP Response Element Binding Protein (CREB) activation is induced by phosphorylation. Akt, p90RSK, CAM KII, PKA are inducers.

Translation Outside the Nucleus Converts mRNA Into a Protein



Translation

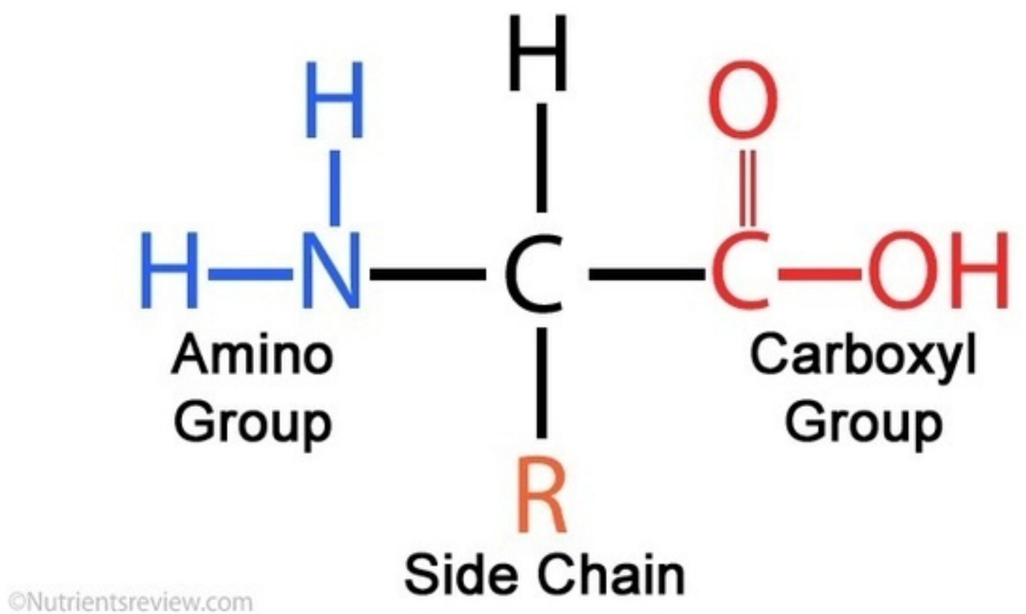


Ribonucleic Acid (RNA)

Proteins

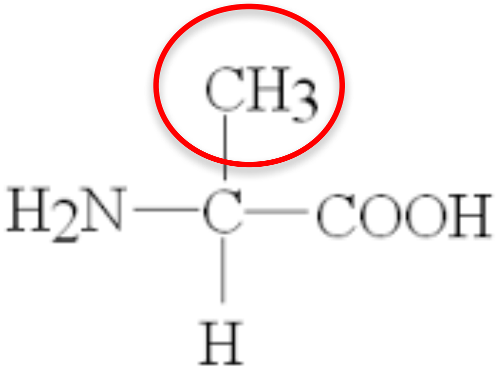
Different Codons in the mRNA Codes for Different Amino Acids

An amino acid is a small organic molecule that is the basic building block of all proteins.

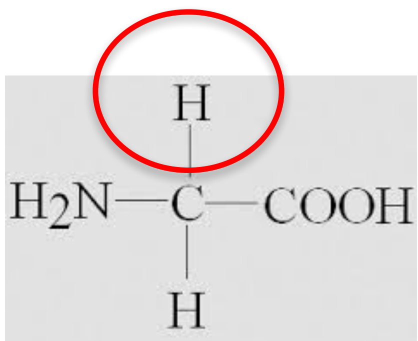


20 different side chains provide the 20 different types of amino acids used in proteins.

The 20 Amino Acids Found in Proteins

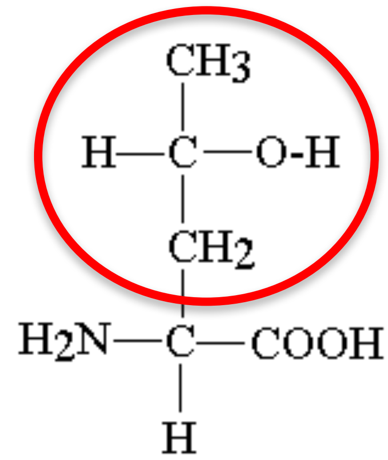


alanine

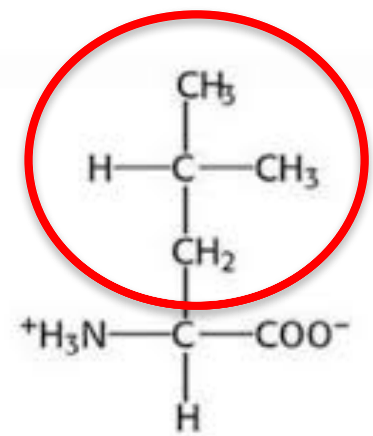


glycine

- alanine - ala - A
- arginine - arg - R
- asparagine - asn - N
- aspartic acid - asp - D
- cysteine - cys - C
- glutamine - gln - Q
- glutamic acid - glu - E
- glycine - gly - G
- histidine - his - H
- isoleucine - ile - I
- leucine - leu - L
- lysine - lys - K
- methionine - met - M
- phenylalanine - phe - F
- proline - pro - P
- serine - ser - S
- threonine - thr - T
- tryptophan - trp - W
- tyrosine - tyr - Y
- valine - val - V

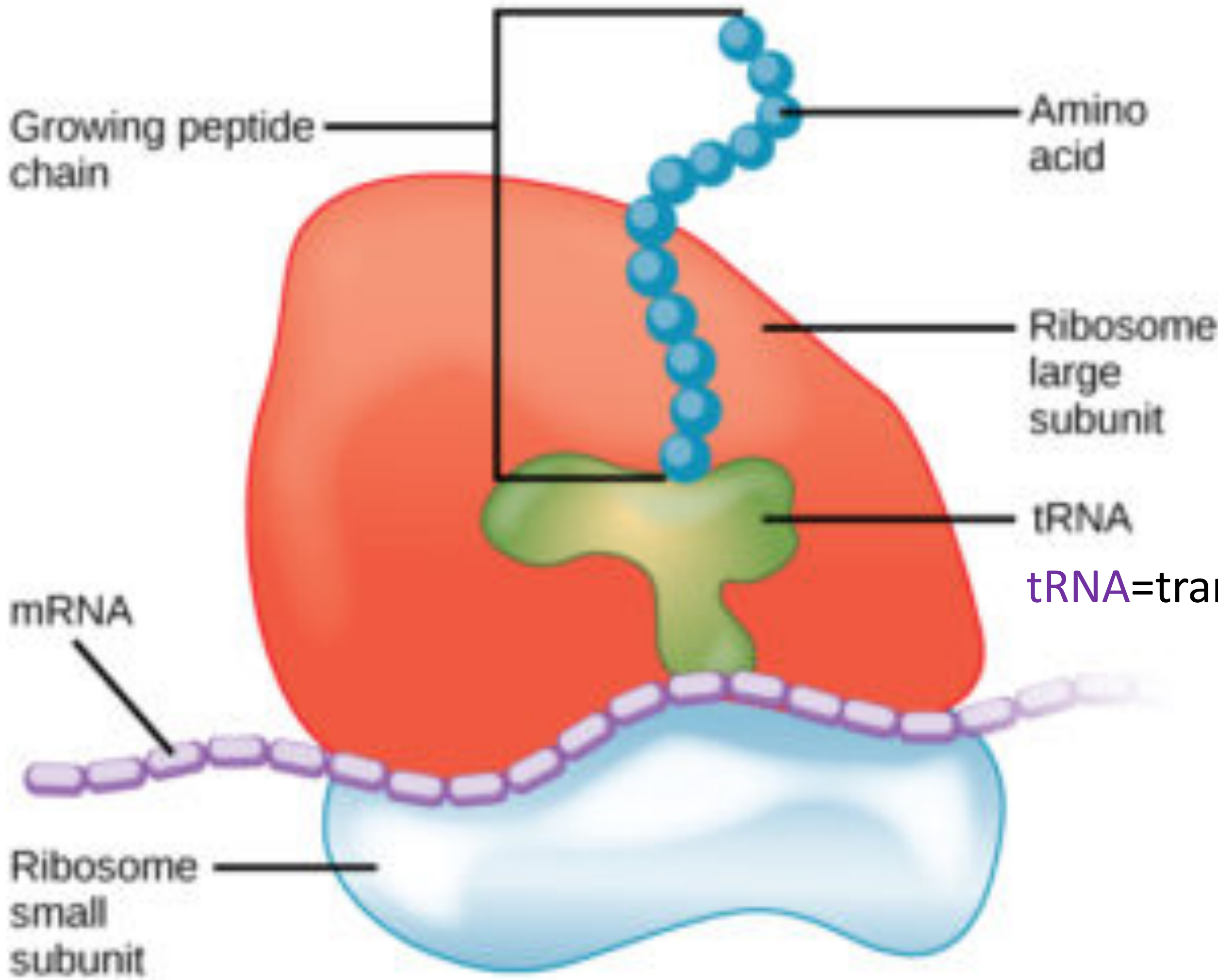


threonine



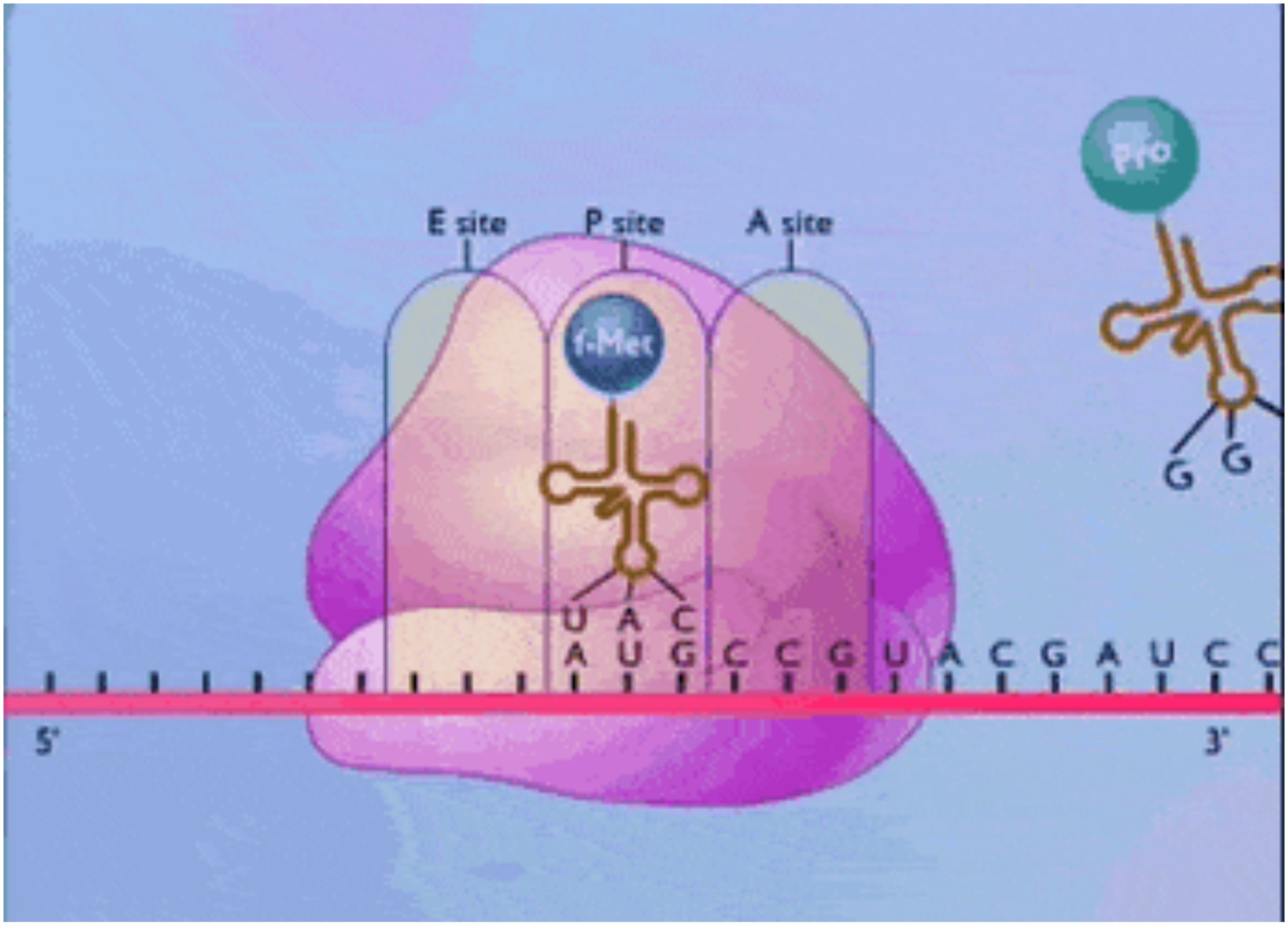
Leucine
(Leu, L)

Ribosomes Translate mRNA Into a Protein



tRNA=transfer RNA

Ribosomes Translate mRNA Into a Protein



Time Scales

Fast

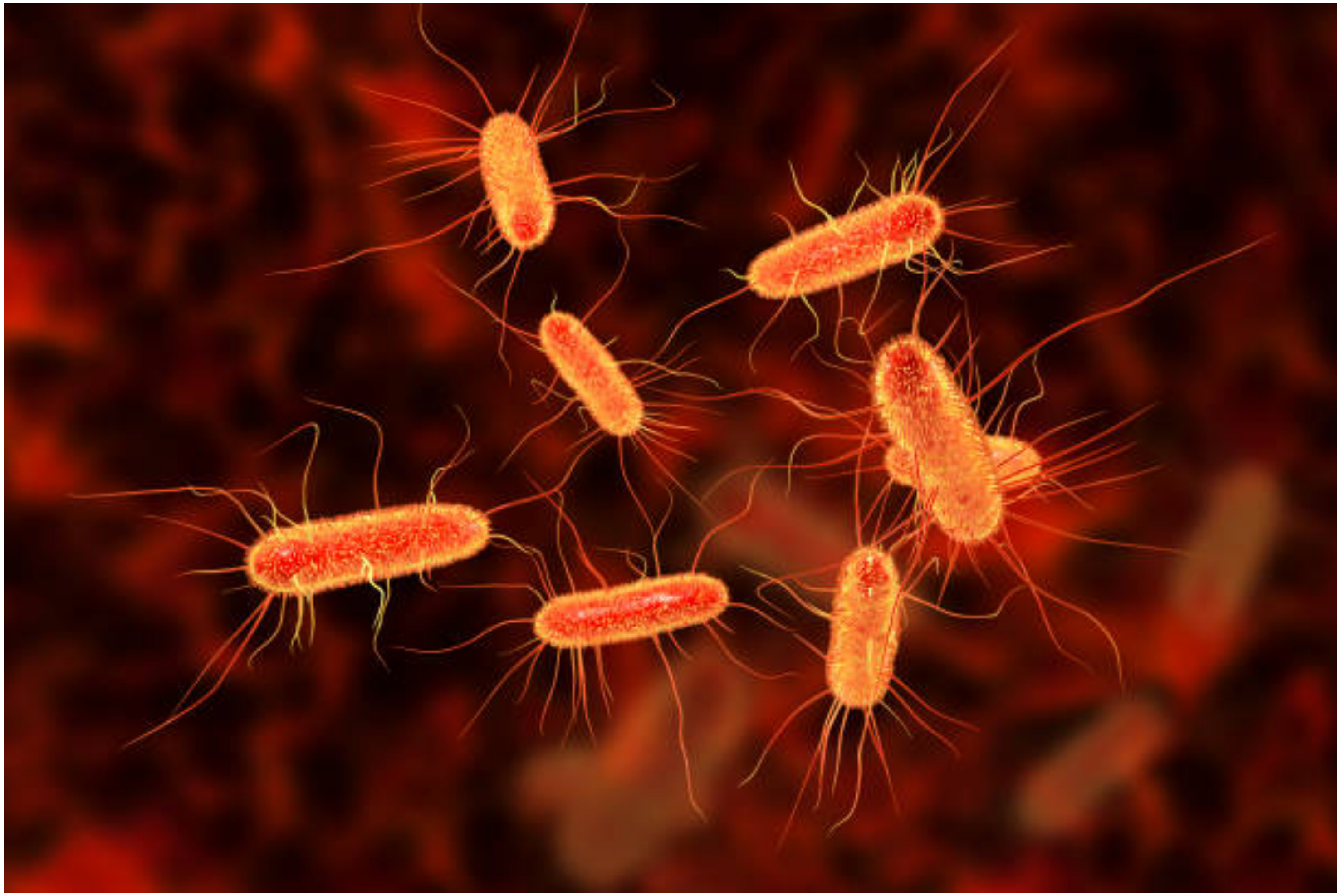
Transcription factor activation (1 sec)

Transcription and translation (5 min)

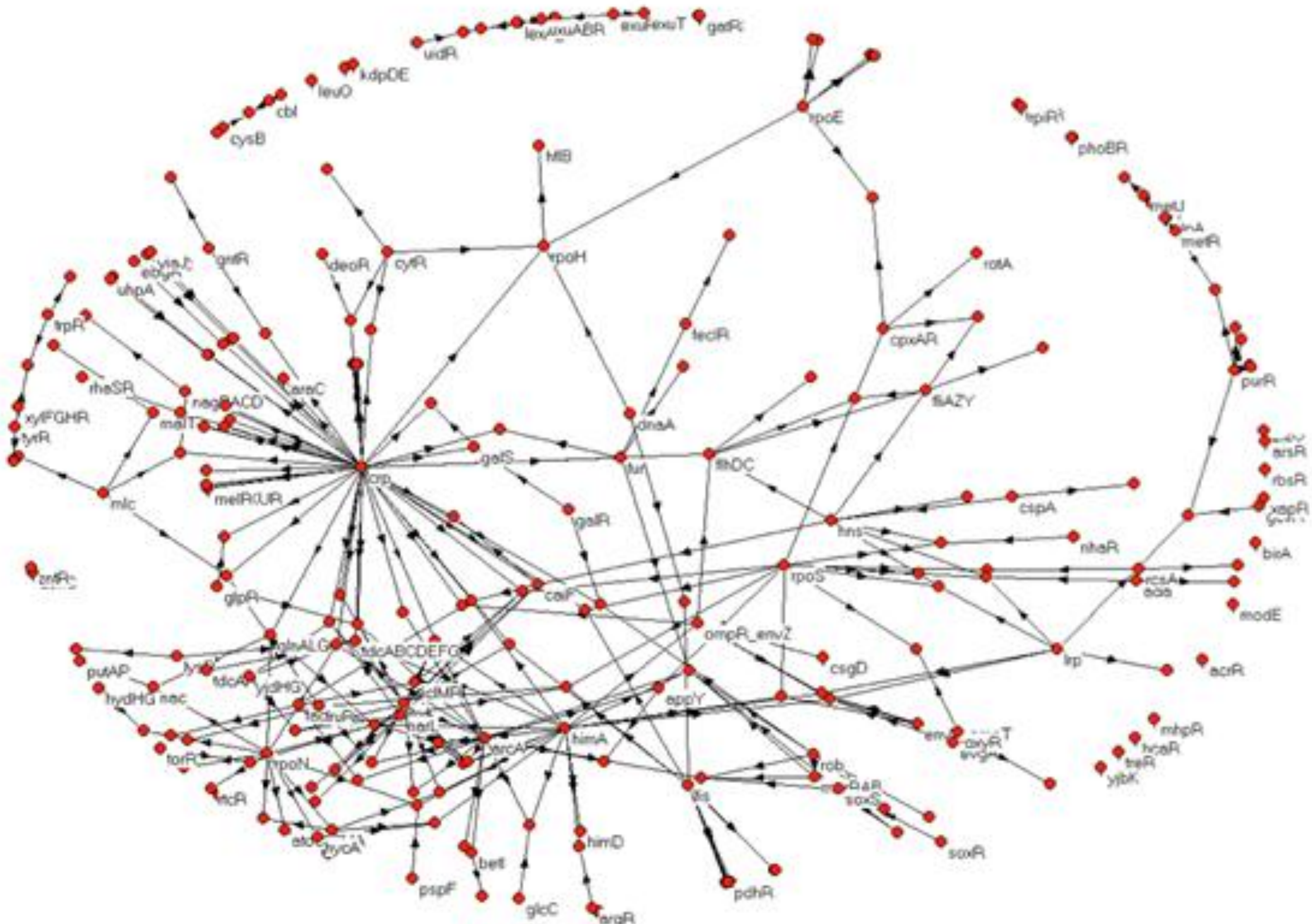
Slow

50% concentration change of gene product (1 hr)

Much Has Been Done with E. coli



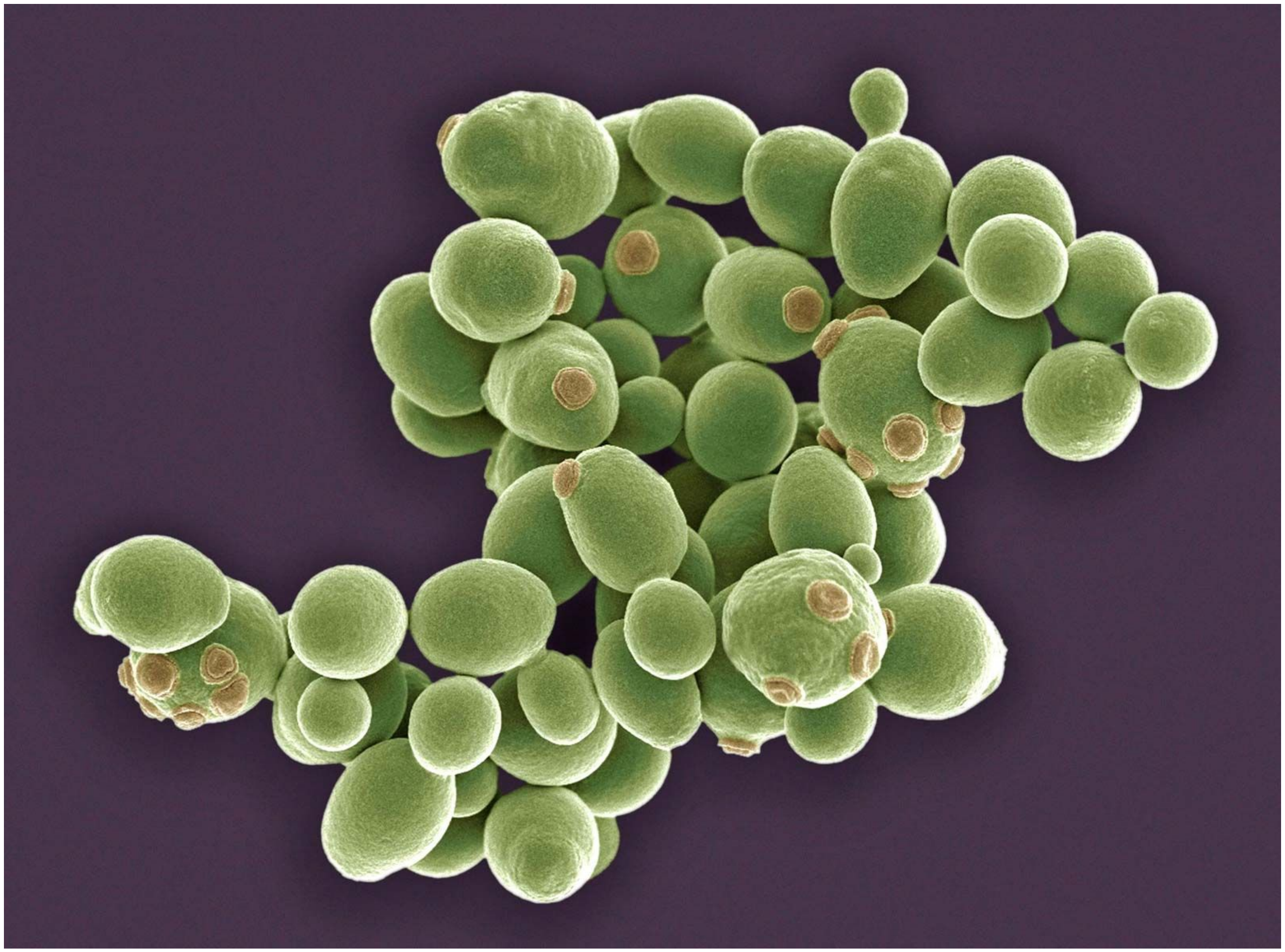
Transcription Network from E. coli



Nodes=genes

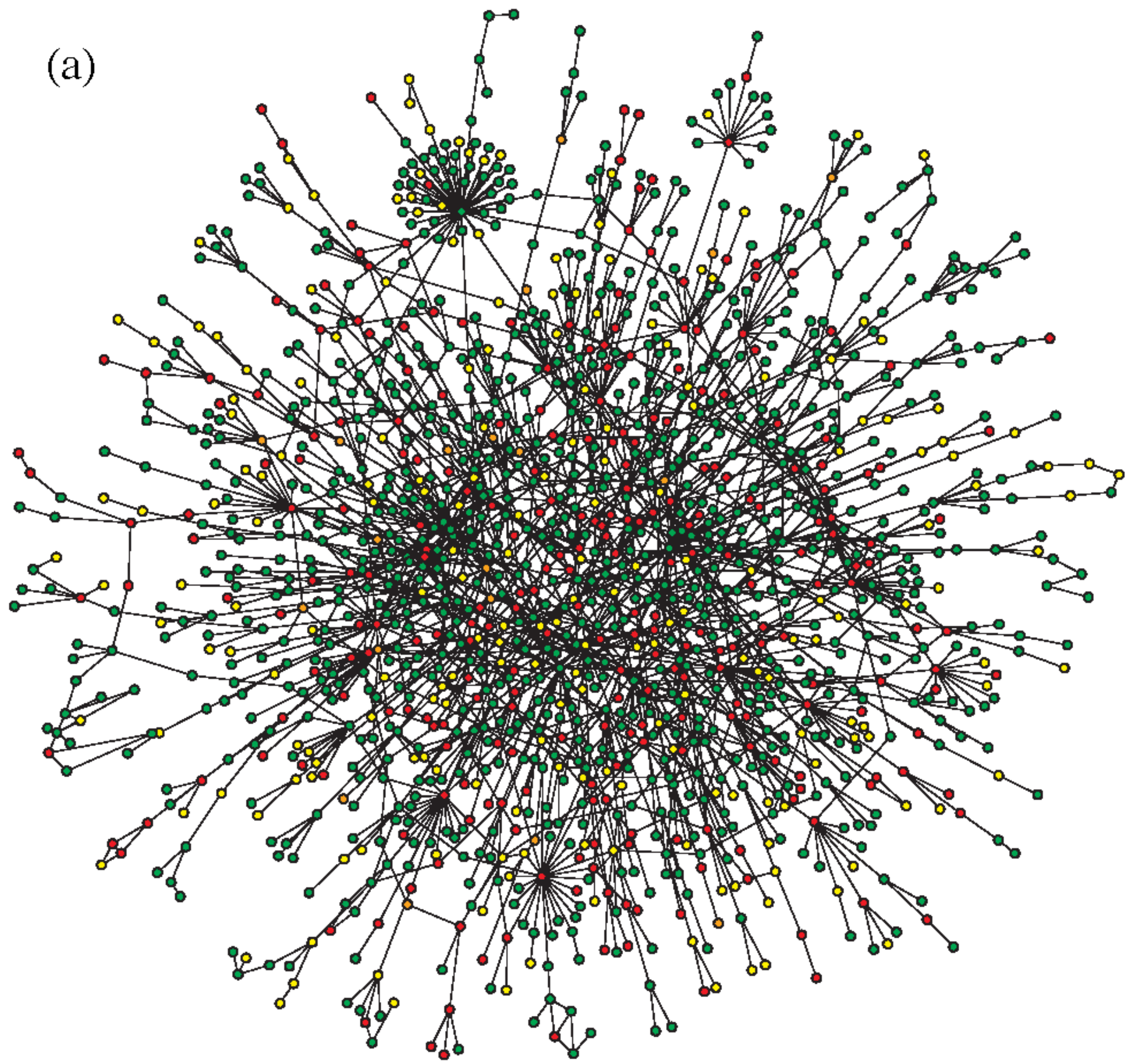
Directed edges=actions of gene products

Much Has Been Done with Yeast



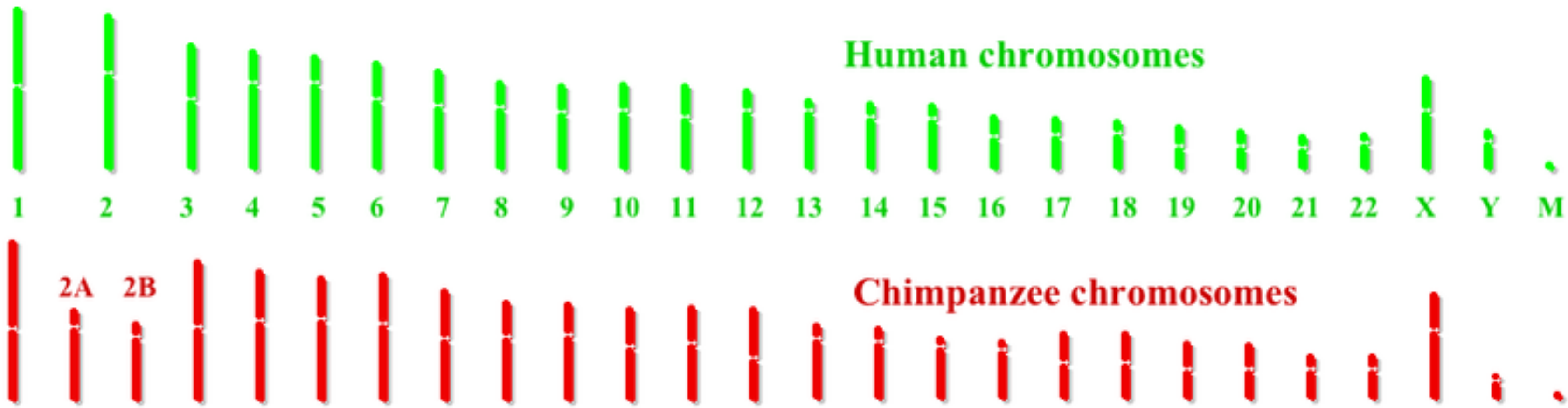
Yeast Transcription Network

(a)



Roughly 1500 genes (nodes) and 1800 interactions (edges).
From Jeong et al., Nature, 411:41, 2001.

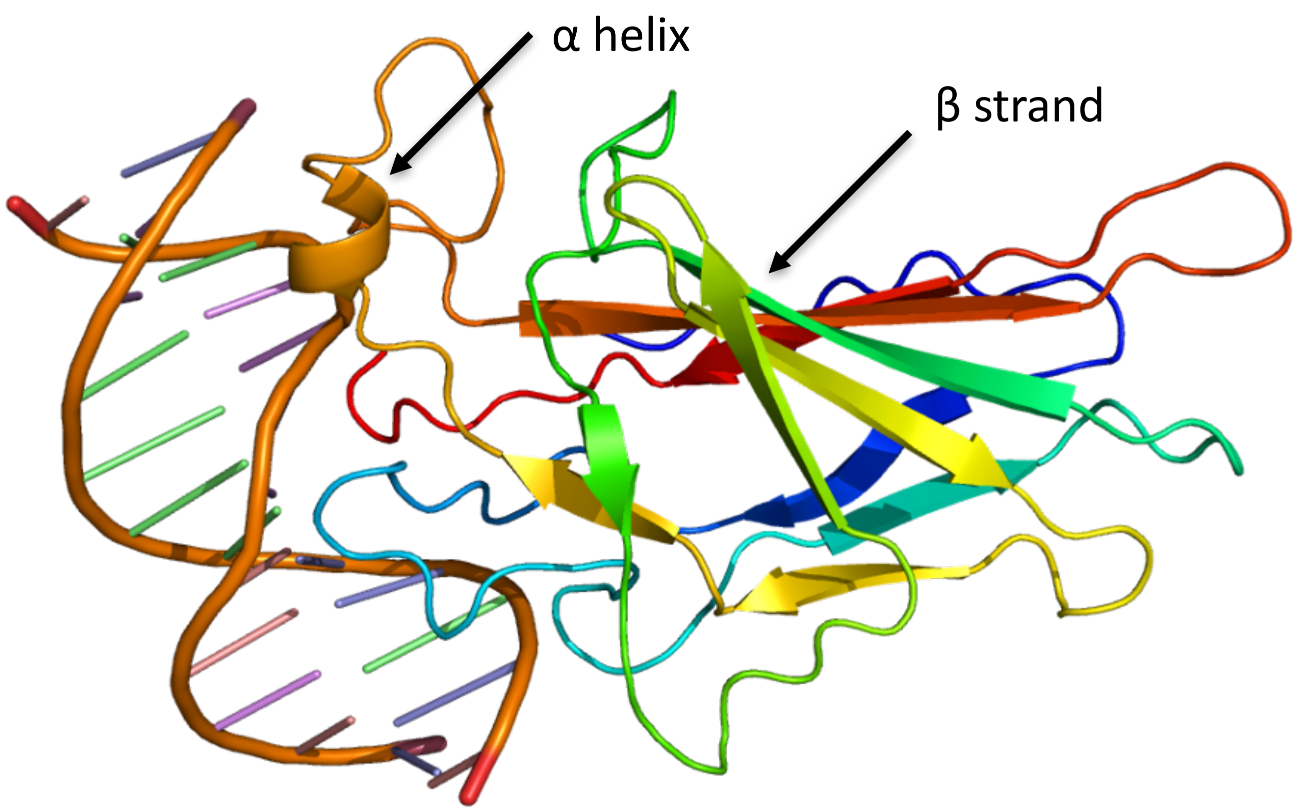
Similarity of Genomes Across Species



Genomes of humans and chimpanzee are 96% identical

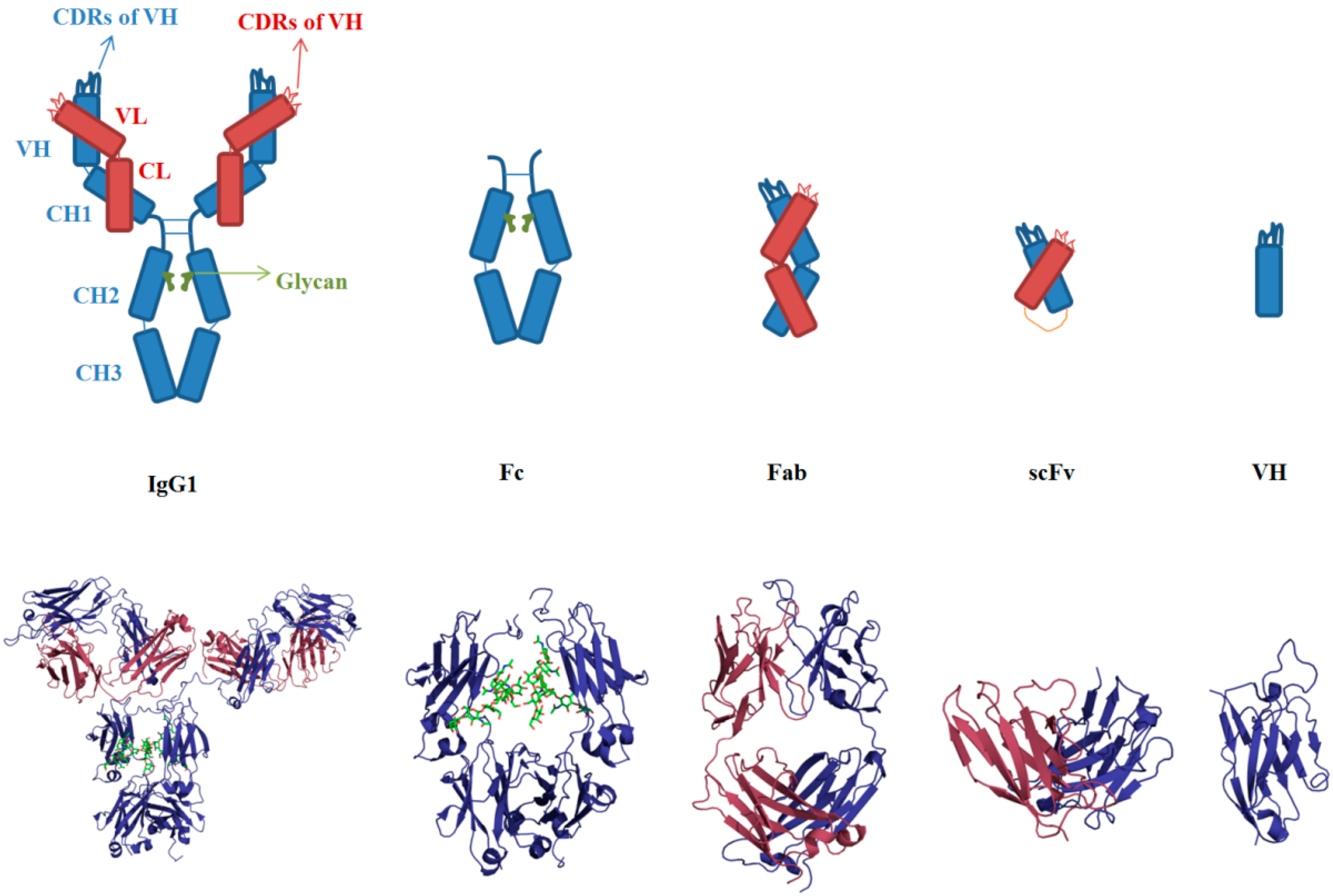
The main difference between us and them is not the genomes, but the sequence of transcription factors that are activated to convert genes to proteins.

Proteins Can Be Transcription Factors



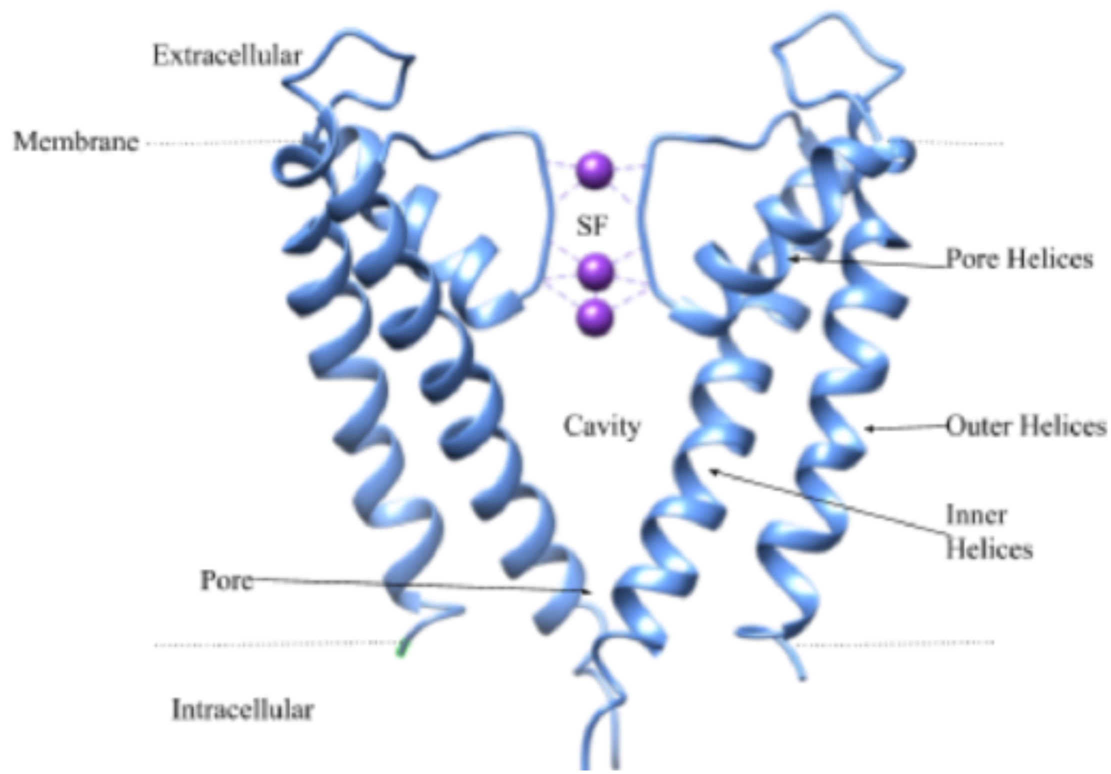
NFATC1 transcription factor

Proteins Are Used in Antibodies



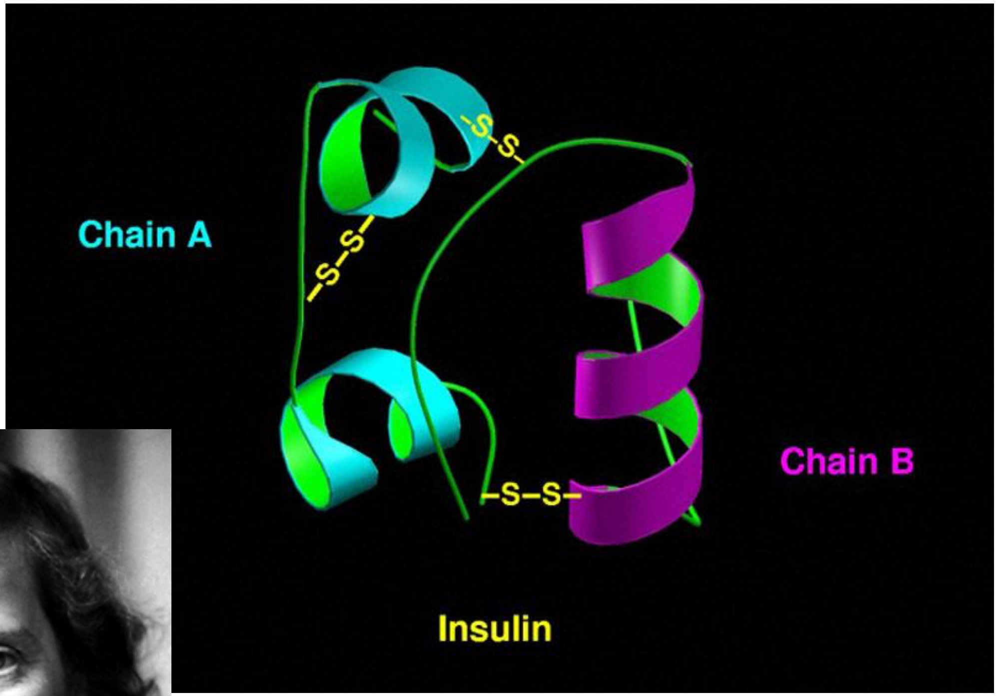
Some antibodies, composed of protein complexes

Proteins Form Ion Channels



Ribbon diagram of two subunits of the transmembrane domain of a KcsA K^+ channel (the full domain is a tetramer).

Proteins Can Be Used As Hormones

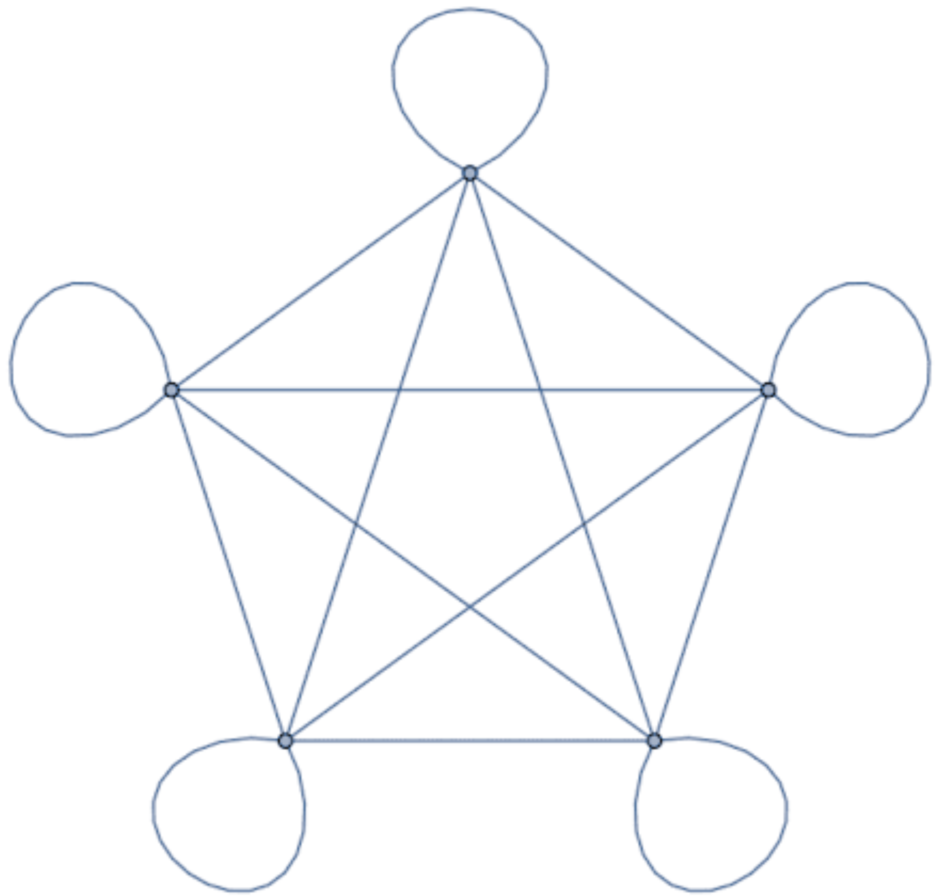


Two chains, total of 51 amino acids

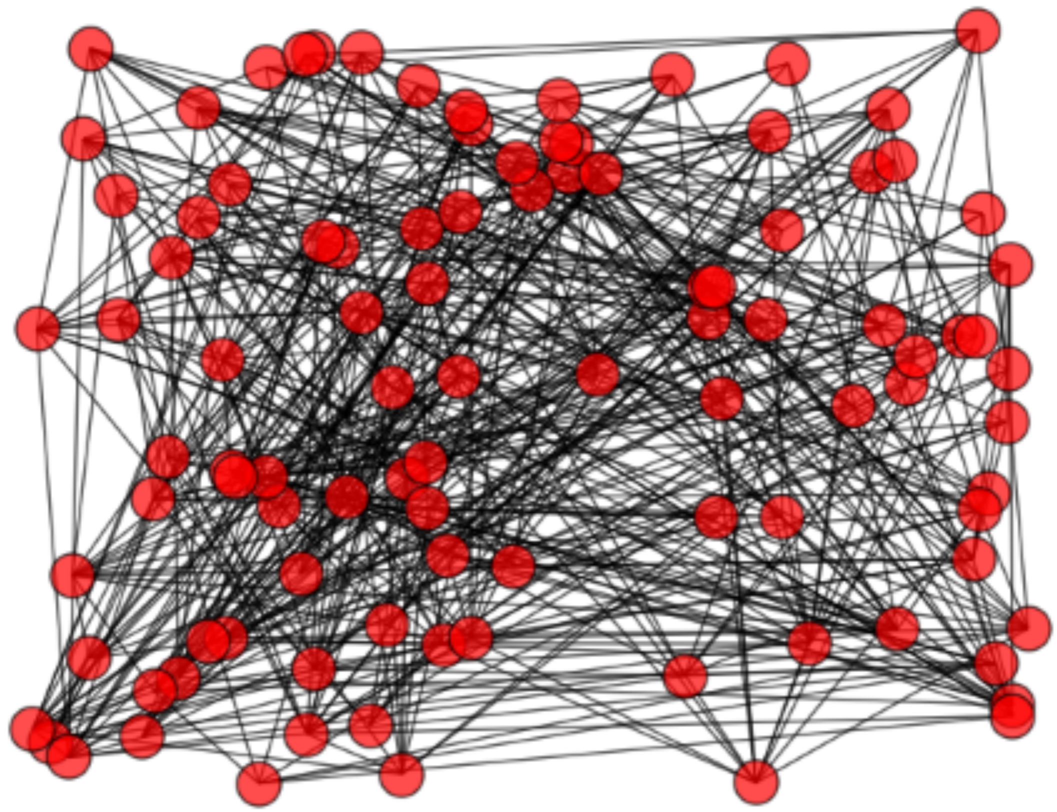
Dorothy Hodgkin determined the atomic structure of insulin, Nobel Prize in 1964

Biological Network Motifs

Example of a Graph with Self-Edges



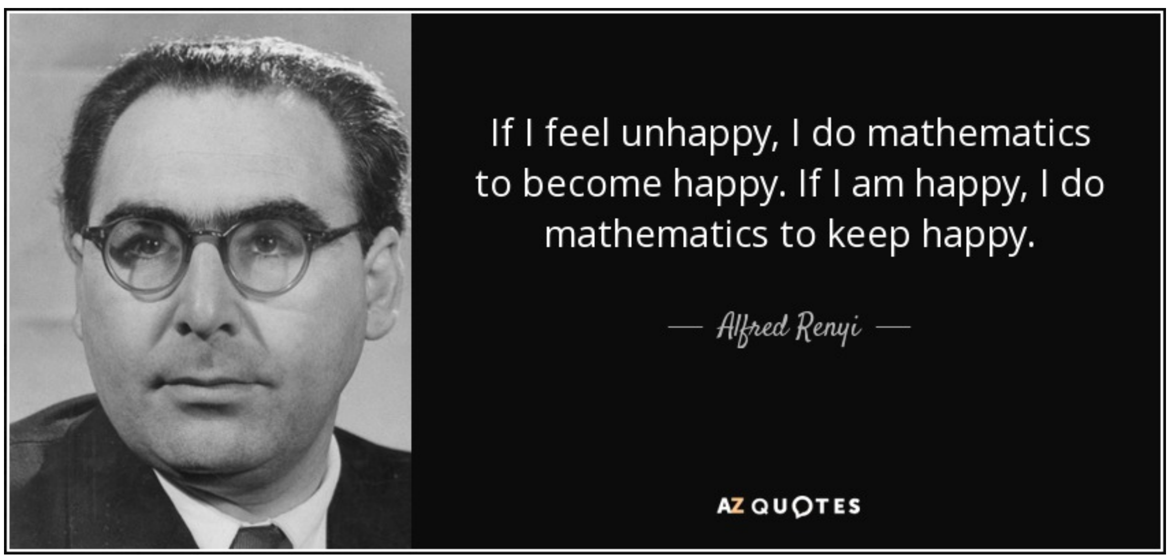
Example of a Simple Random Graph



Simple means no self-edges

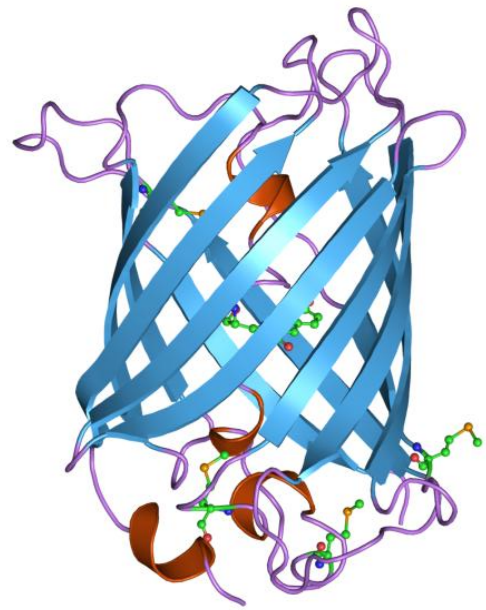
Random Graph Pioneers

Paul Erdós (1913-1996) was a Hungarian mathematician who published around 1500 mathematical papers over his long career. He had no home, but was not homeless.



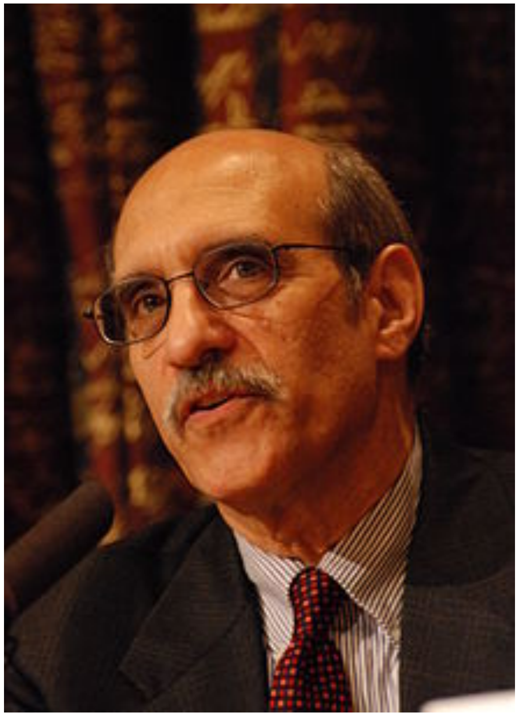
Alfréd Rényi (1921-1970) was a frequent collaborator, and therefore has an Erdós number of 1.

Green Fluorescent Protein from Jellyfish

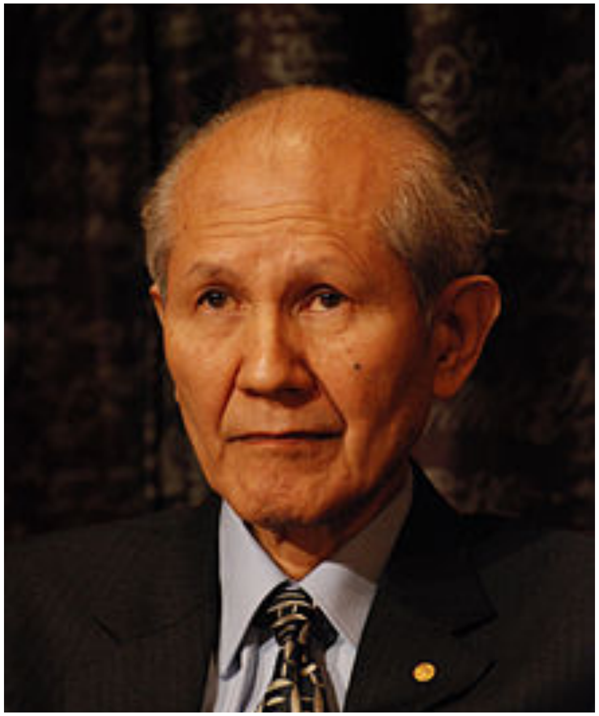


Atomic structure of GFP

2008 Nobel Prize for Discovery and Development of GFP



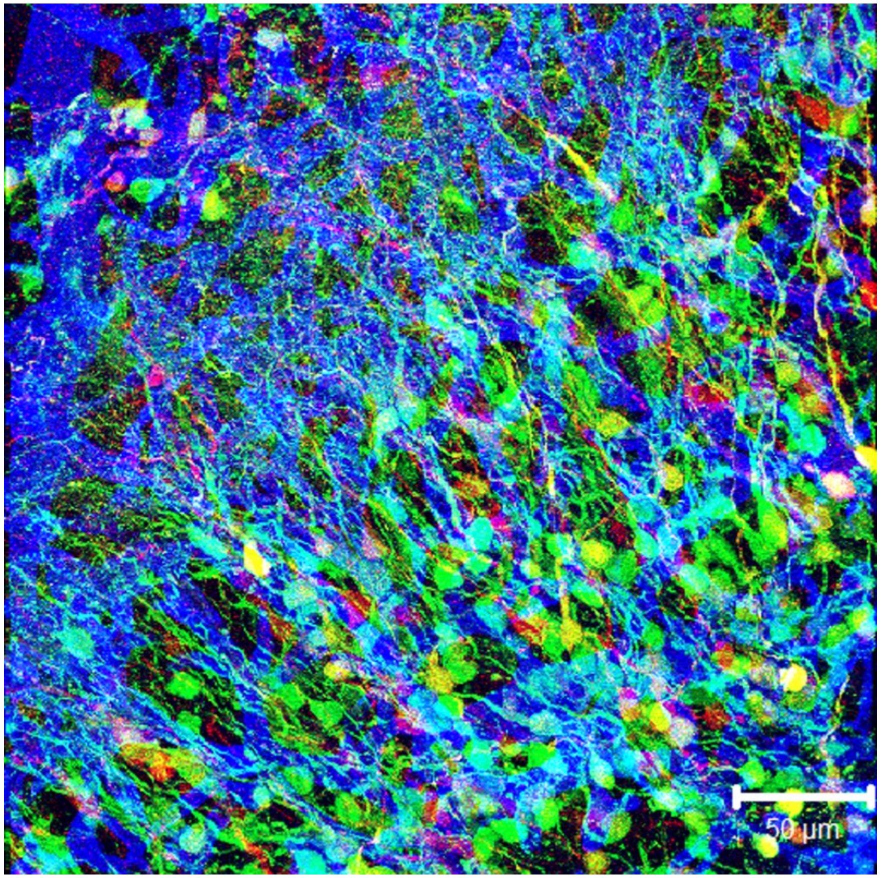
Martin Chalfie
(Columbia University)



Osama Shimomura
(Boston University)

Roger Tsien
(UC San Diego)

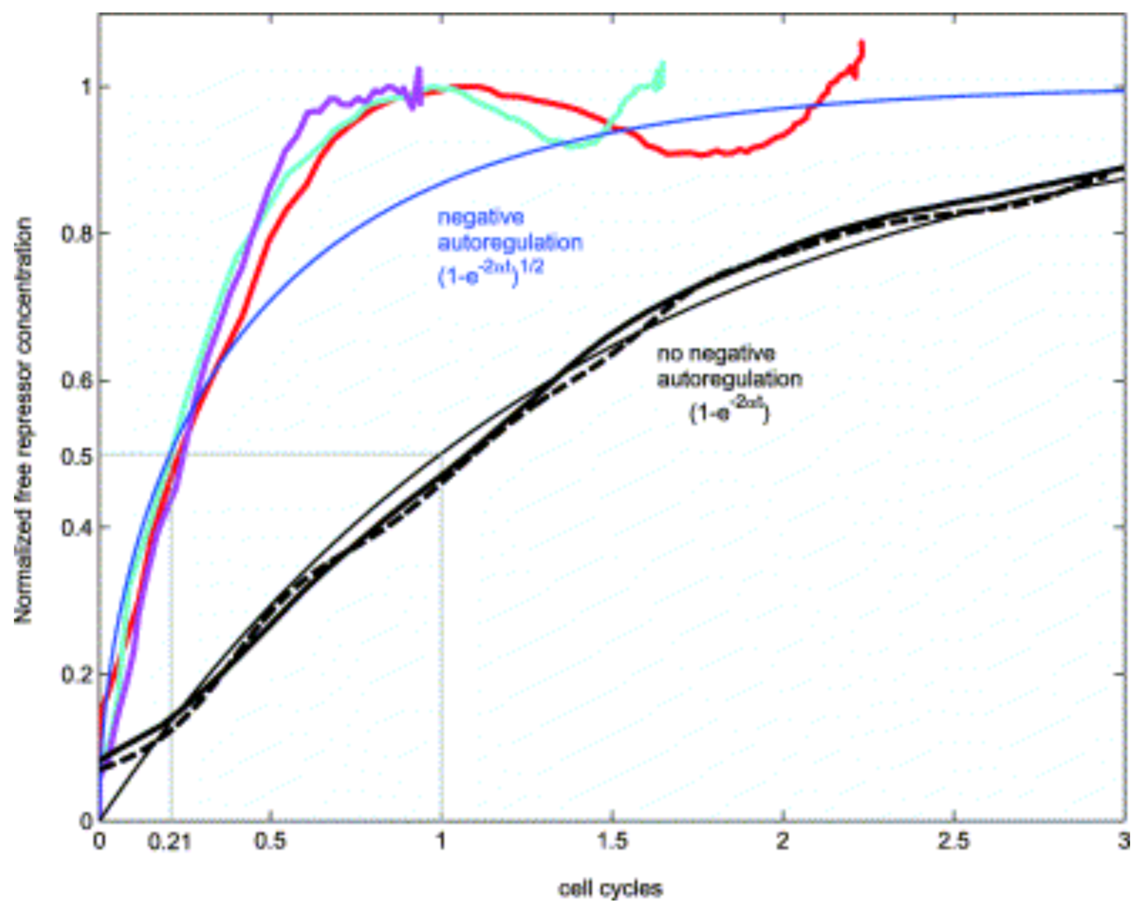
Widespread Use of GFP in Biology



Neural progenitor cells labeled by GFP in olfactory bulb

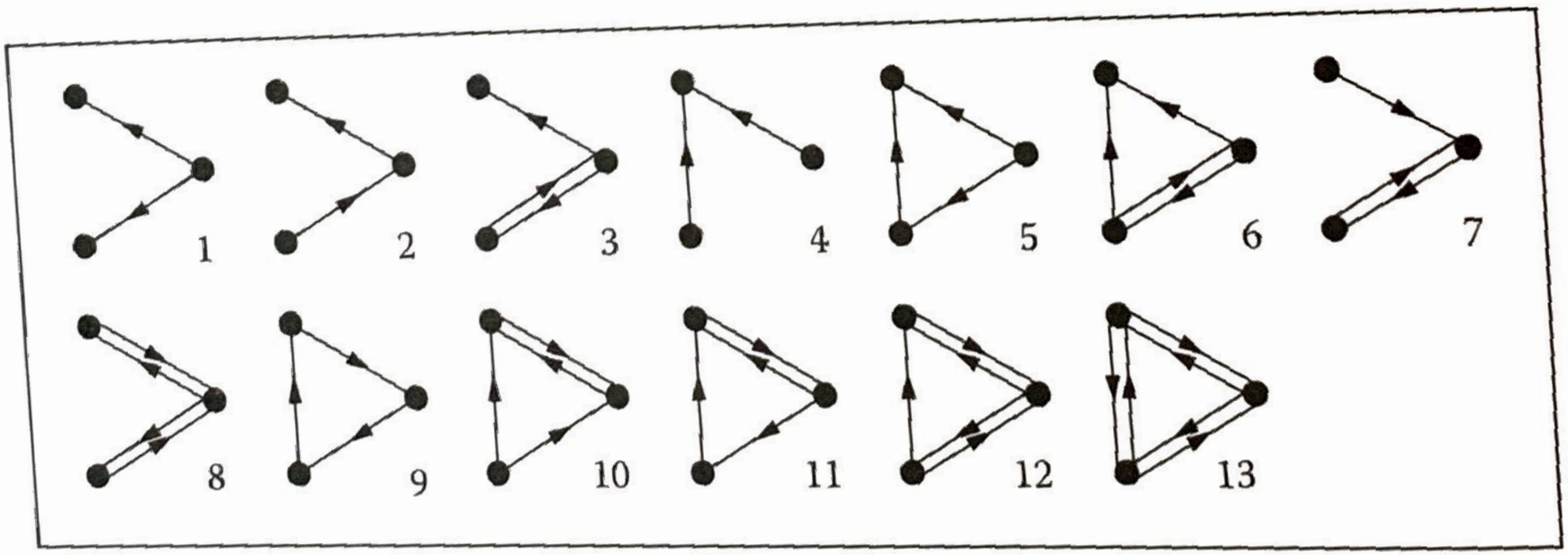
Widespread Use of GFP in Biology

Faster Response with Negative Autoregulation

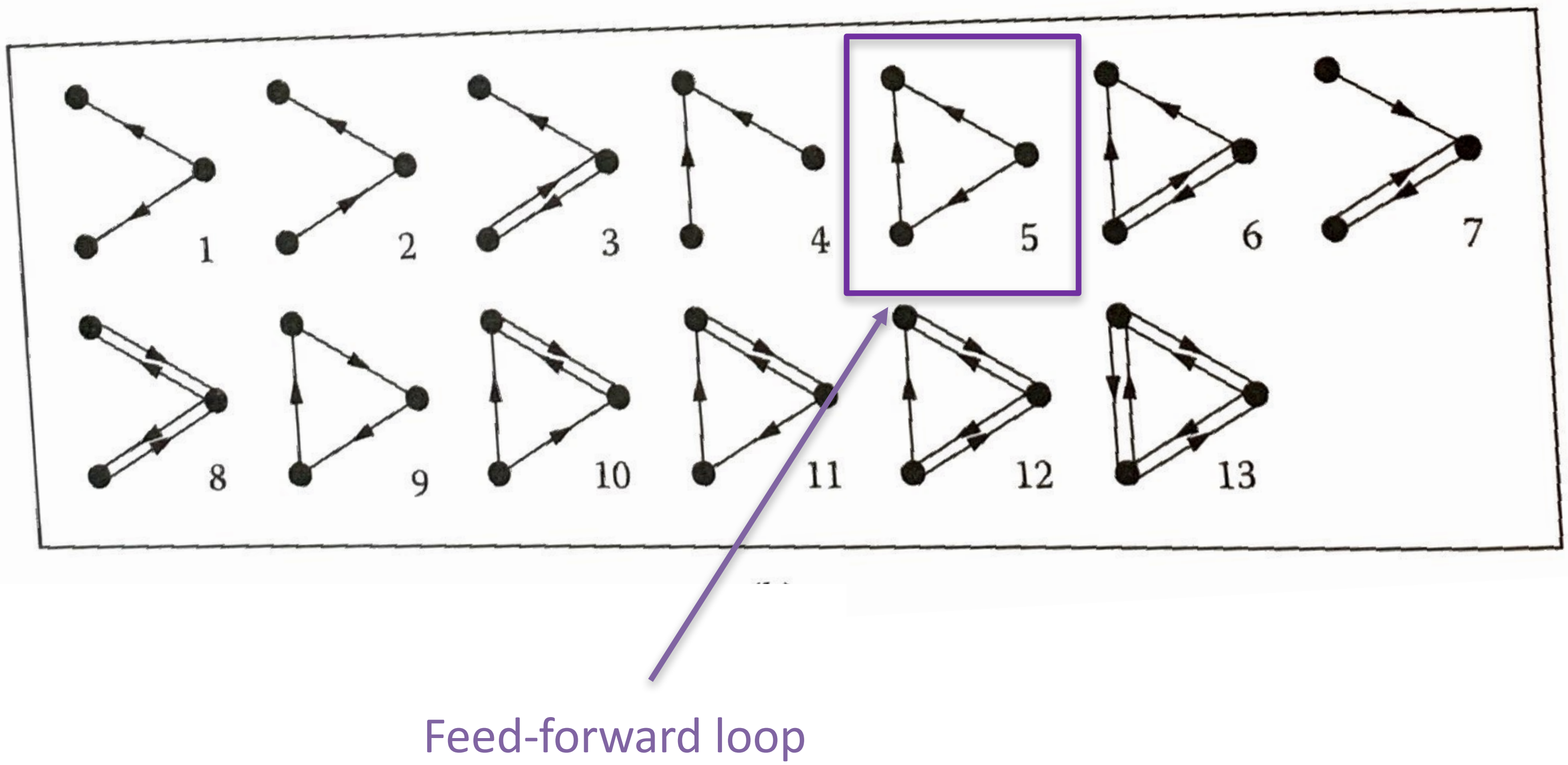


Measured in E. coli using green fluorescent protein. Normalized to steady state level. (From Rosenfeld et al., 2002)

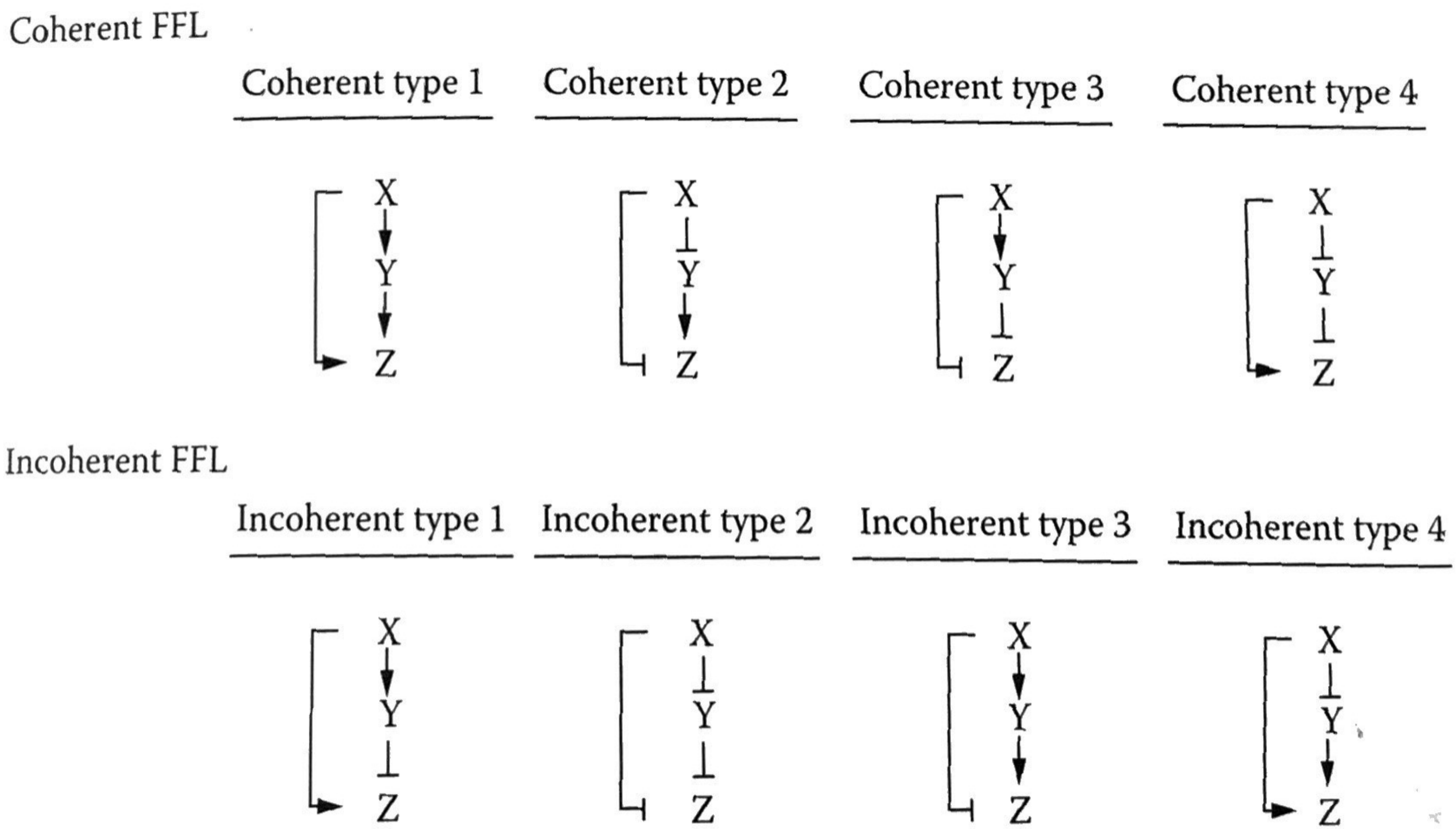
13 Possible Connected Subgraphs with 3-Node Network



13 Possible Connected Subgraphs with 3-Node Network

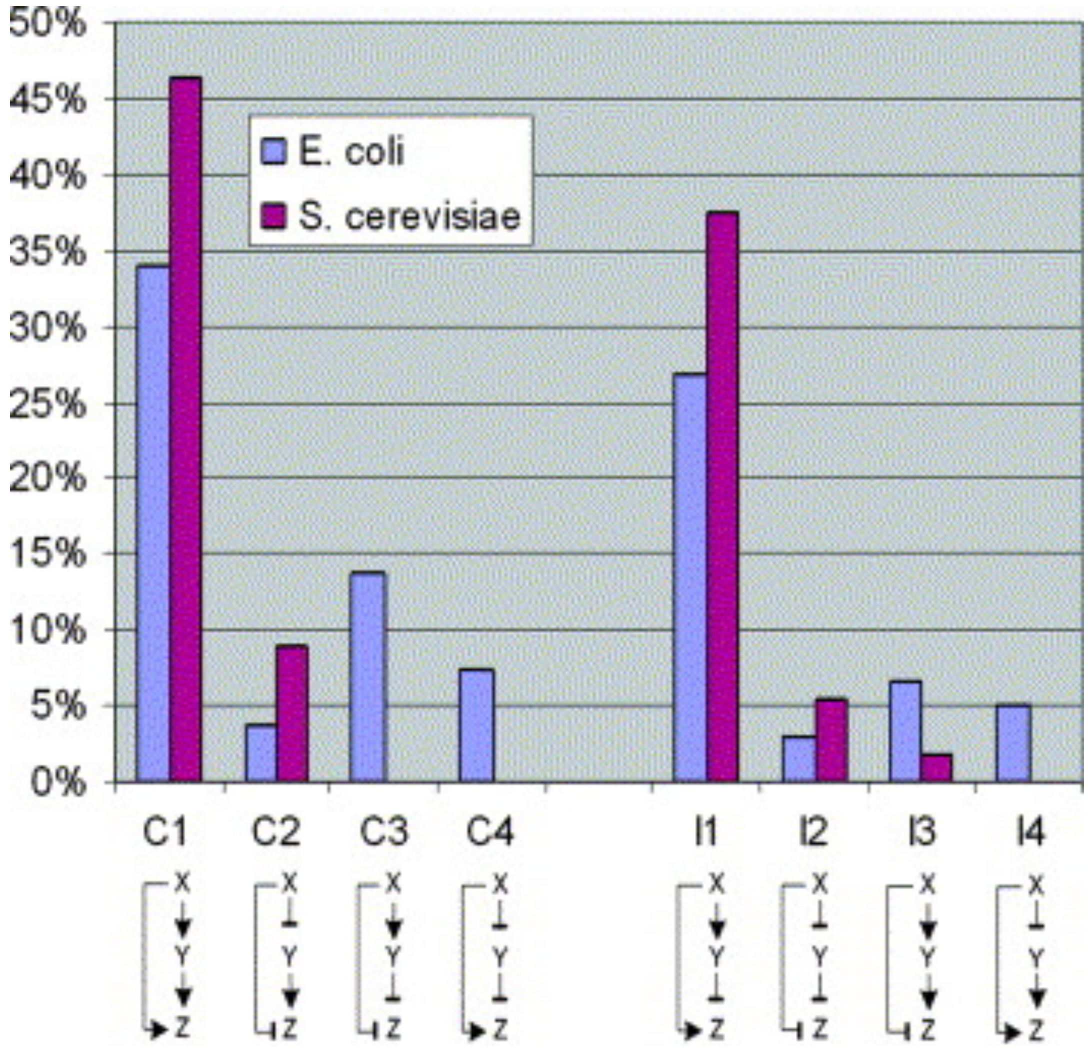


Eight Possible Feed-Forward Loops



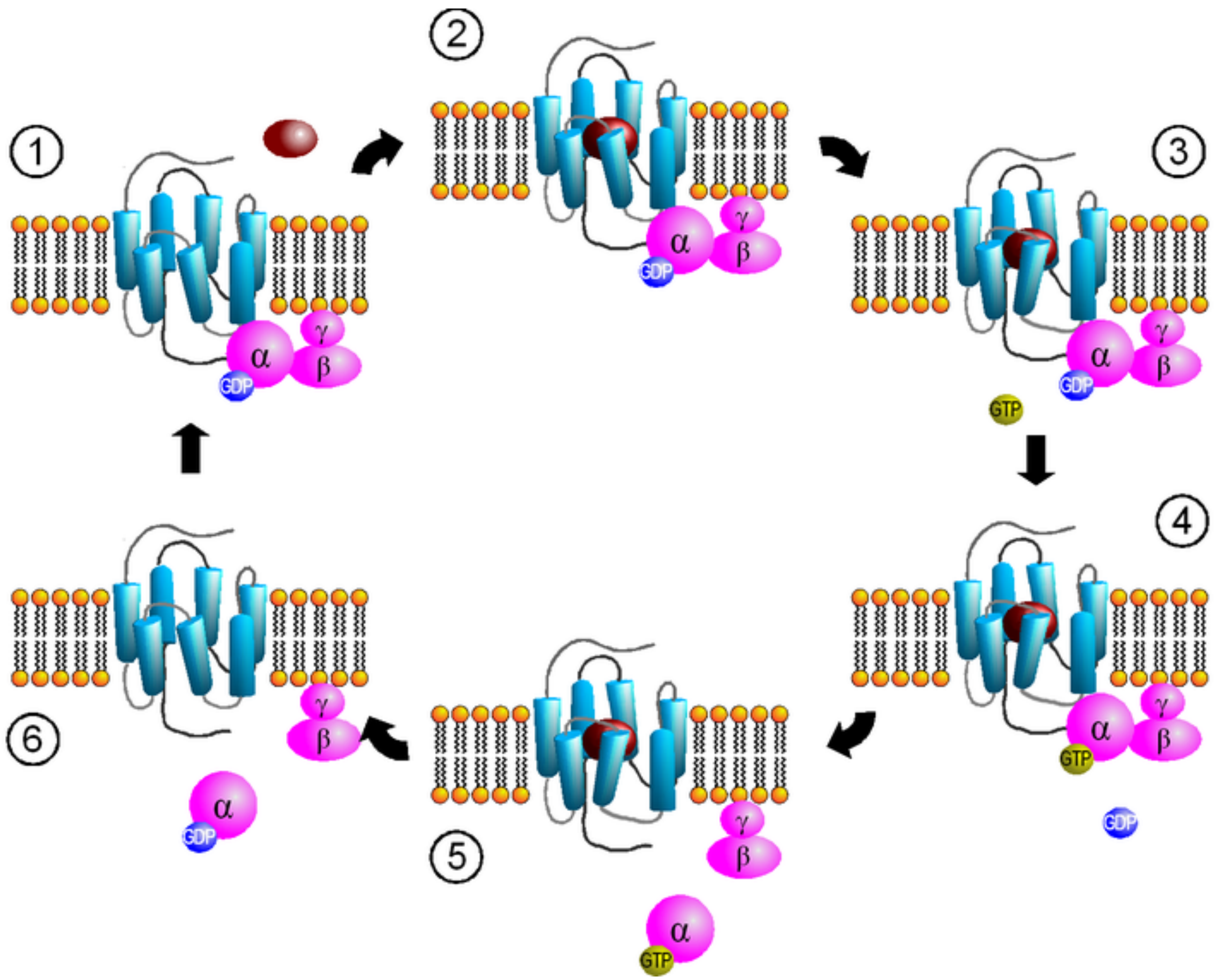
arrow=activation, flat line=repression

The Abundance of the 8 Possible FFLs in Bacteria and Yeast



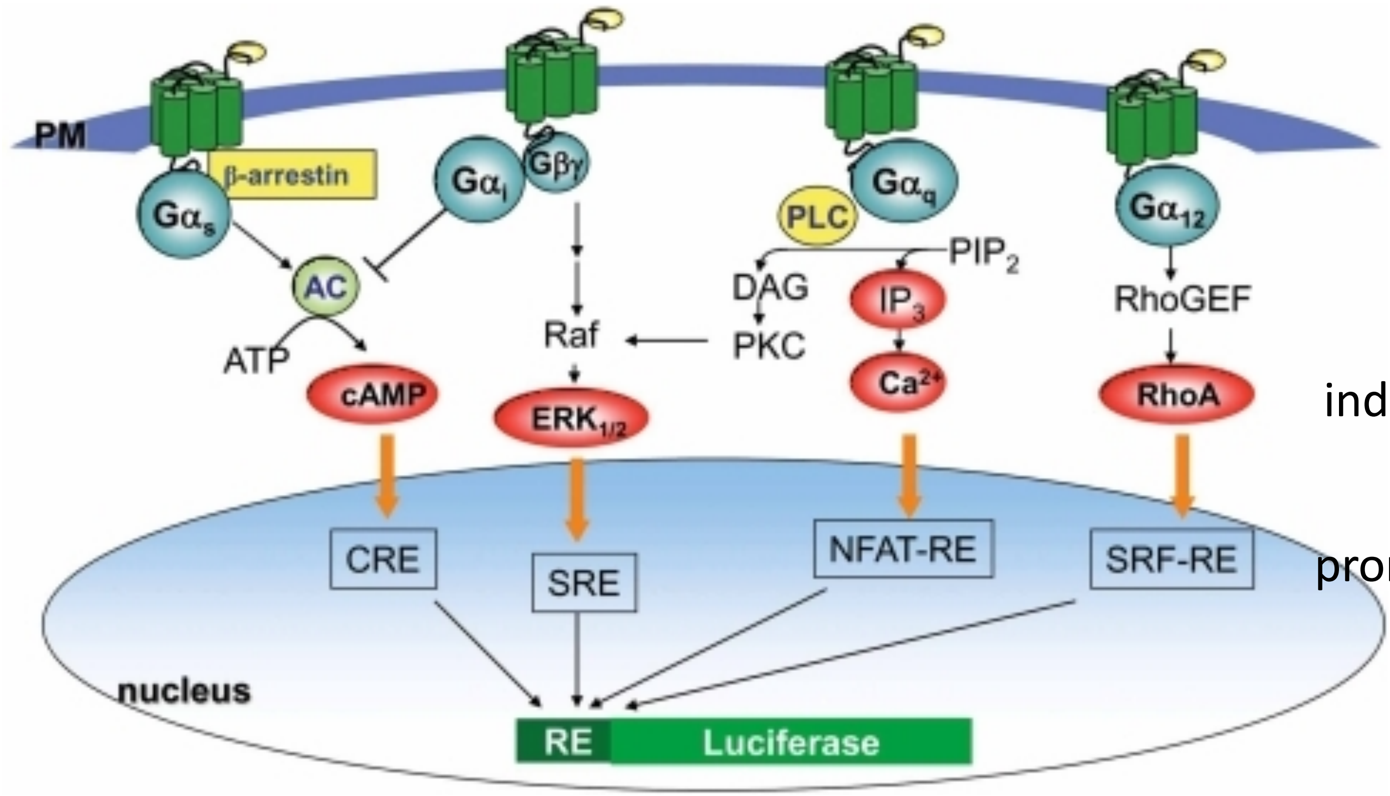
Relative to total number of FFLs found in E. Coli (138) and S. cerevisiae (brewers yeast, 56). From Mangan et al., 2006

G-Protein Activation



The Gα and Gβγ subunits are both active when not part of a trimer. From Wikipedia.

G-Protein Signaling Pathways



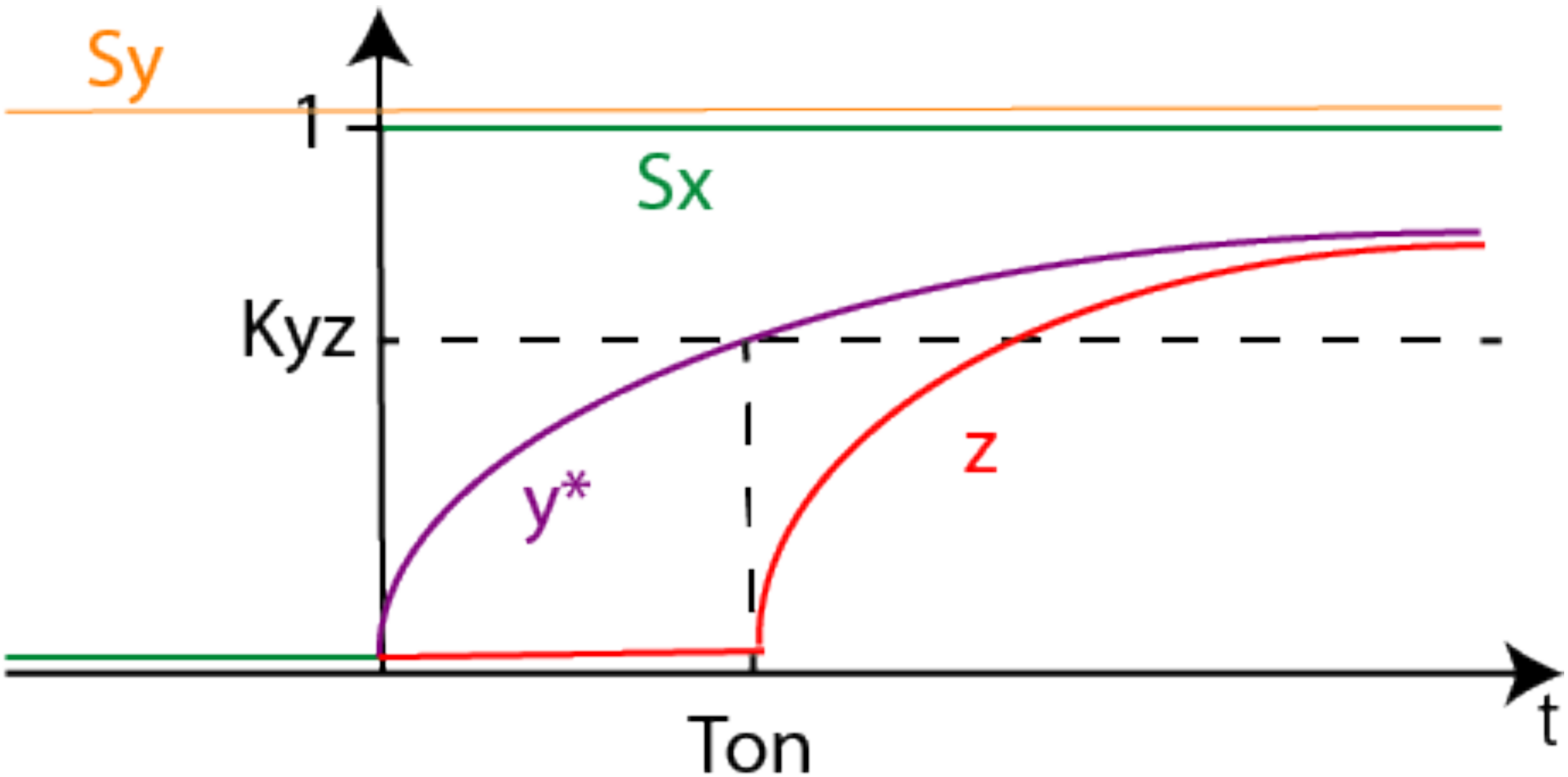
inducers

promoters

Inducers are cAMP, ERK, Ca^{2+}, and RhoA. Endpoints of all four pathways are transcription factors that bind to the **promoters** CRE, SRE, NFAT-RE, and SRF-RE. From Cheng et al., 2010.

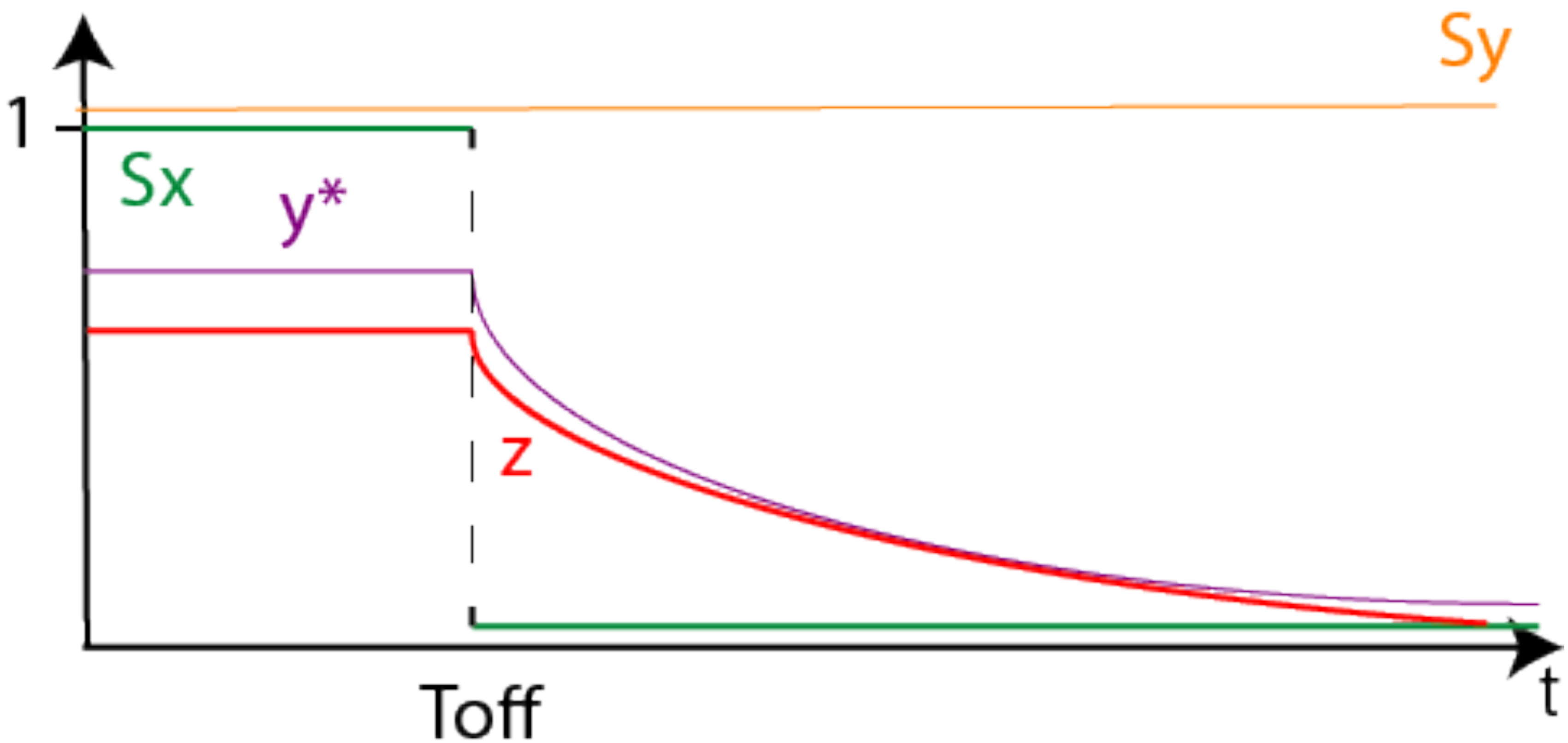
Coherent Type 1-AND Motif Dynamics

Delayed ON response

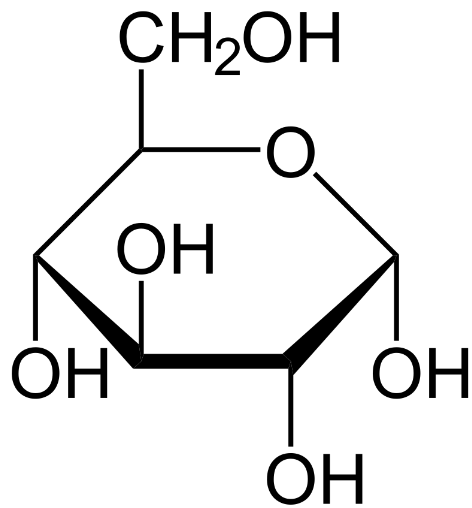


Coherent Type 1-AND Motif Dynamics

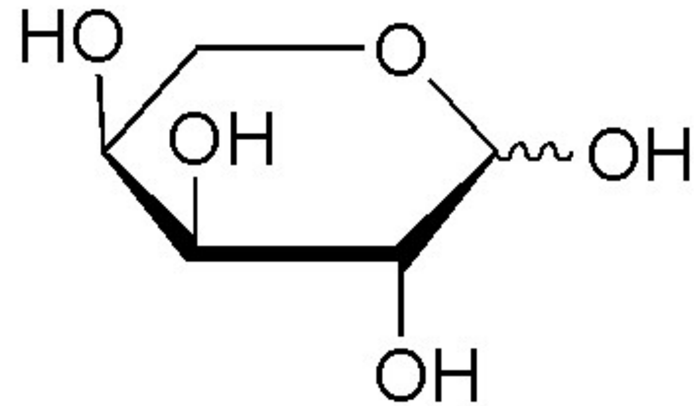
Immediate OFF response



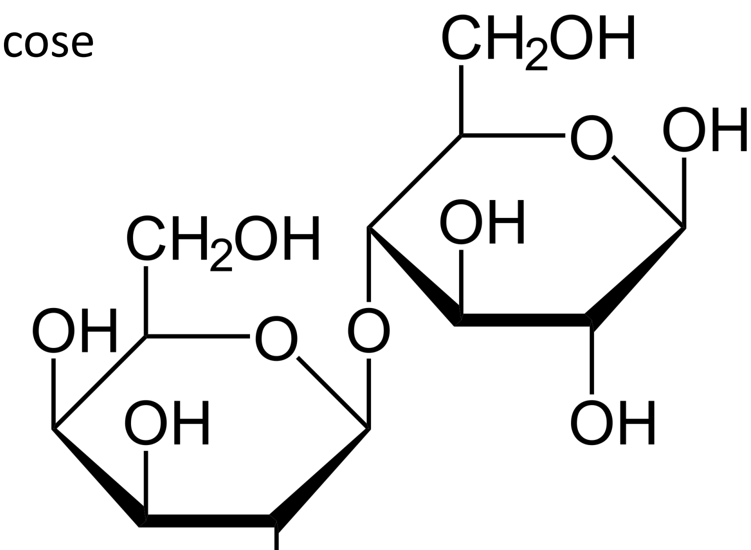
Three Sugars That E. coli Can Metabolize



Glucose

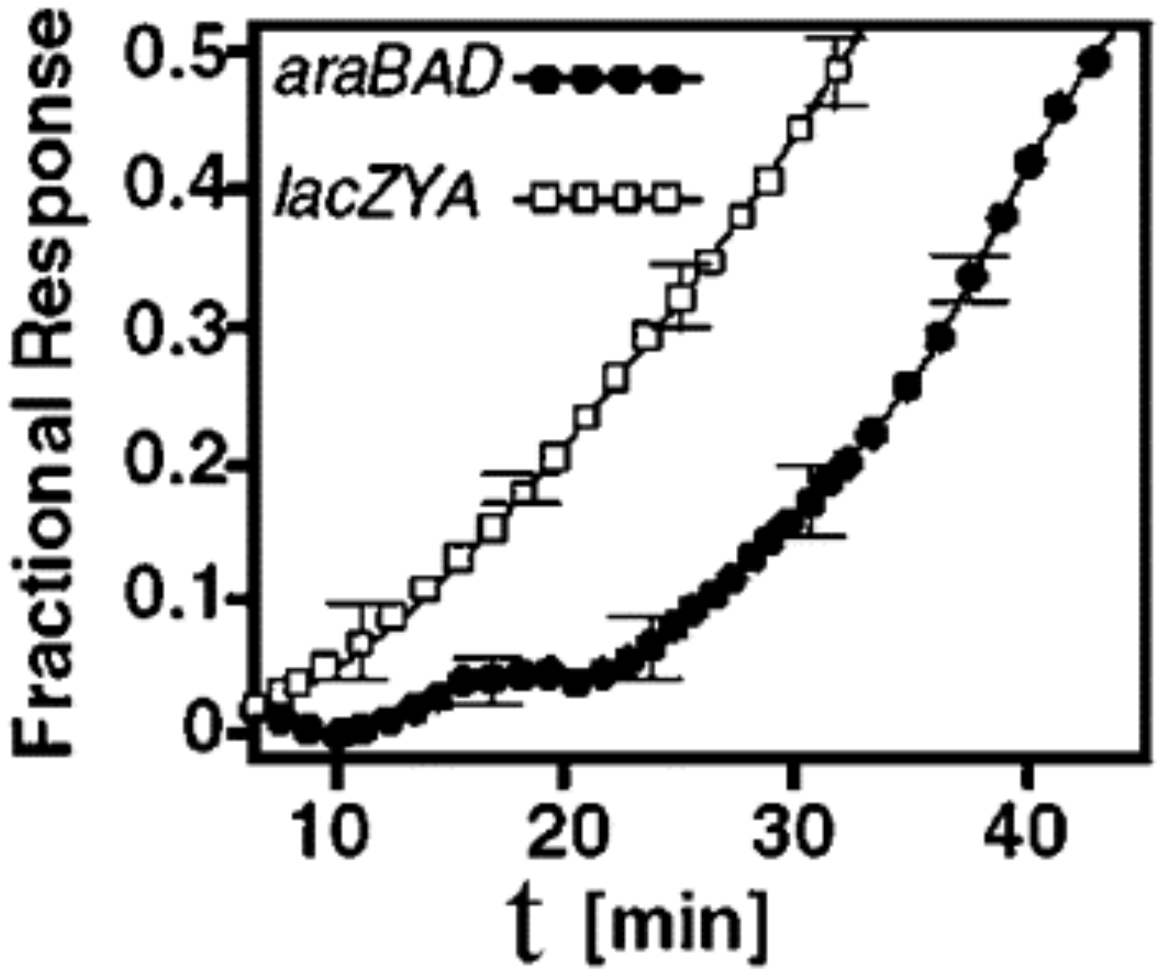


Arabinose



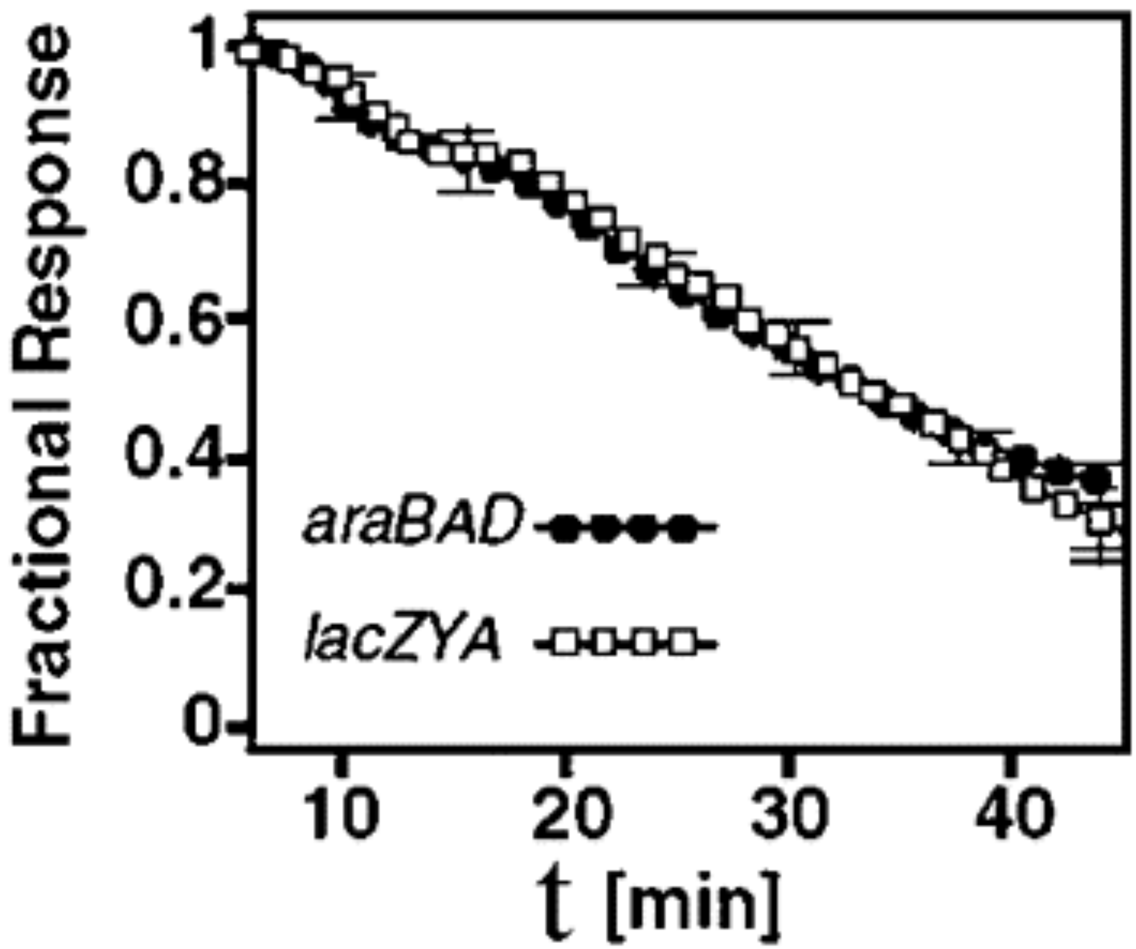
Lactose

Delayed ON response to cAMP with C1-FFL-AND



lacZYA activation through simple regulation, araBAD through C1-FFL-AND.
From Mangan et al., 2003. Measured using GFP.

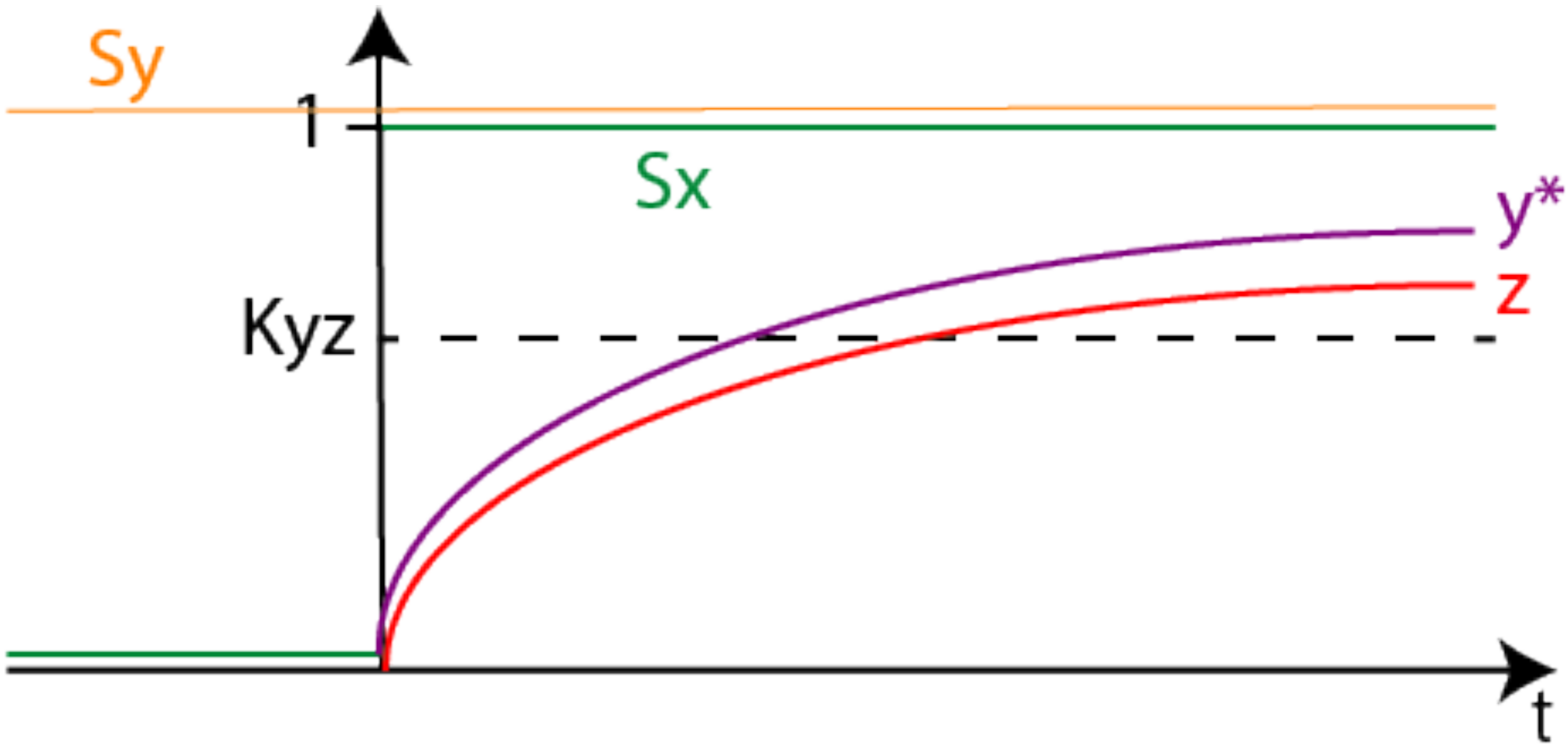
Immediate OFF response to cAMP removal with C1-FFL-AND



lacZYA activation through simple regulation, araBAD through C1-FFL-AND.
From Mangan et al., 2003.

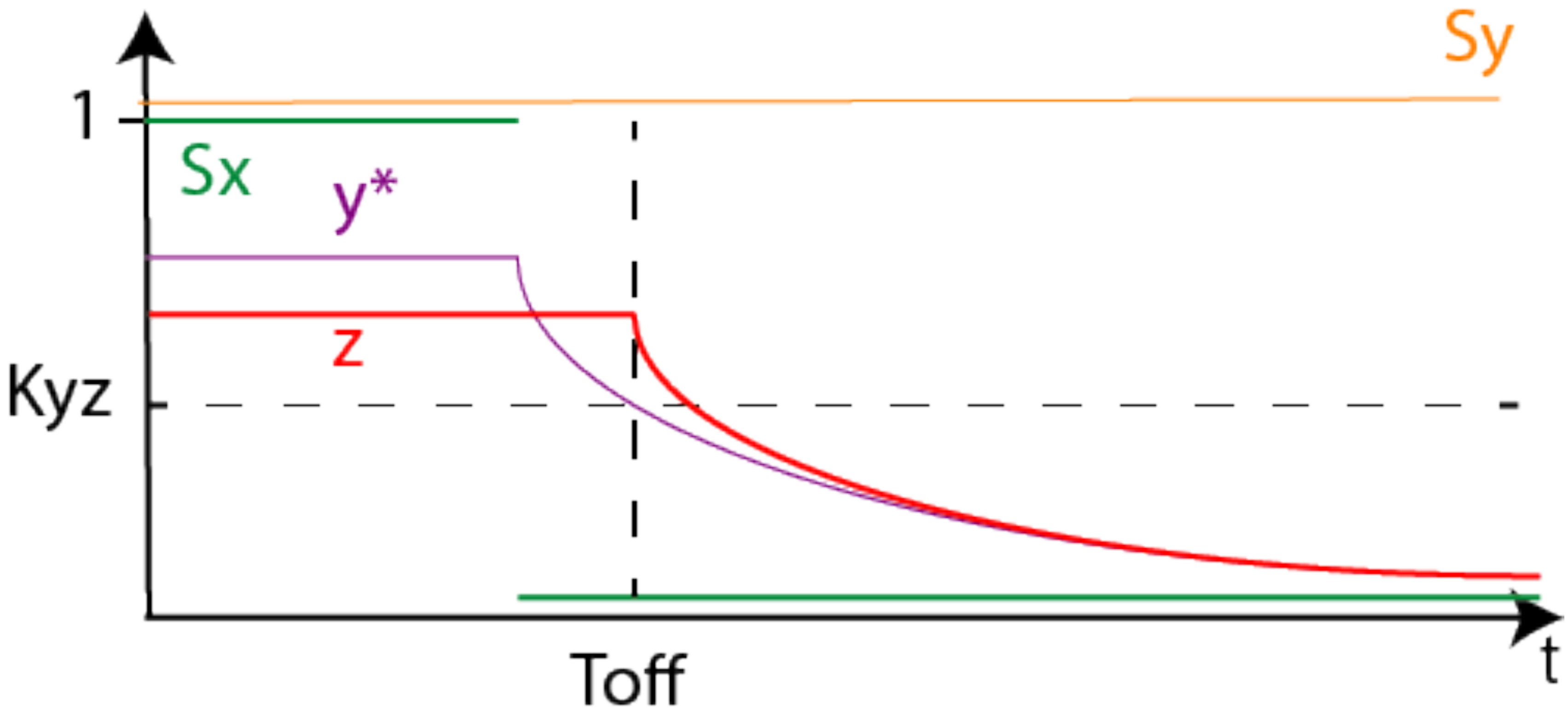
Coherent Type 1-OR Motif Dynamics

Immediate ON response

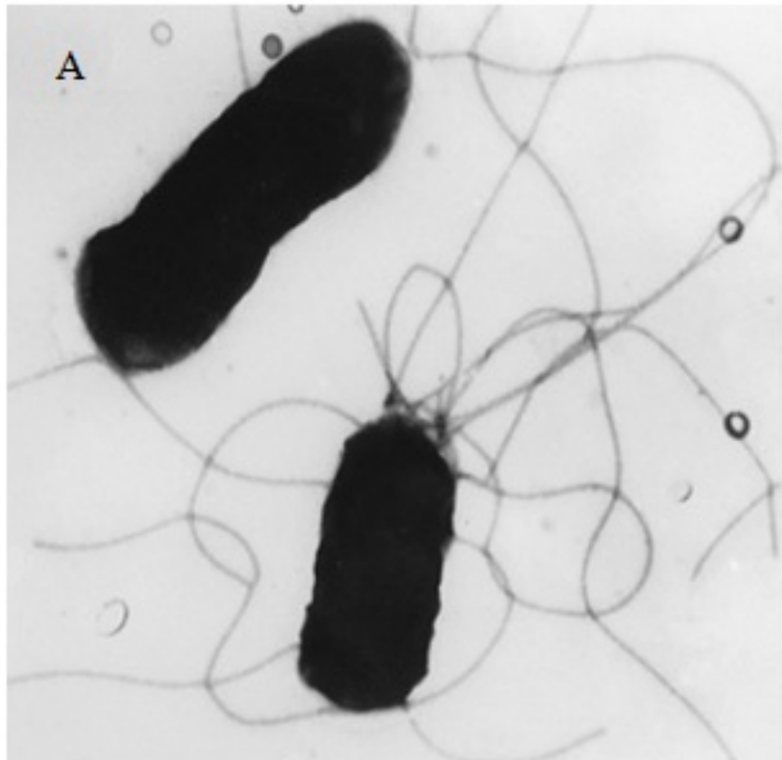


Coherent Type 1-OR Motif Dynamics

Delayed OFF response

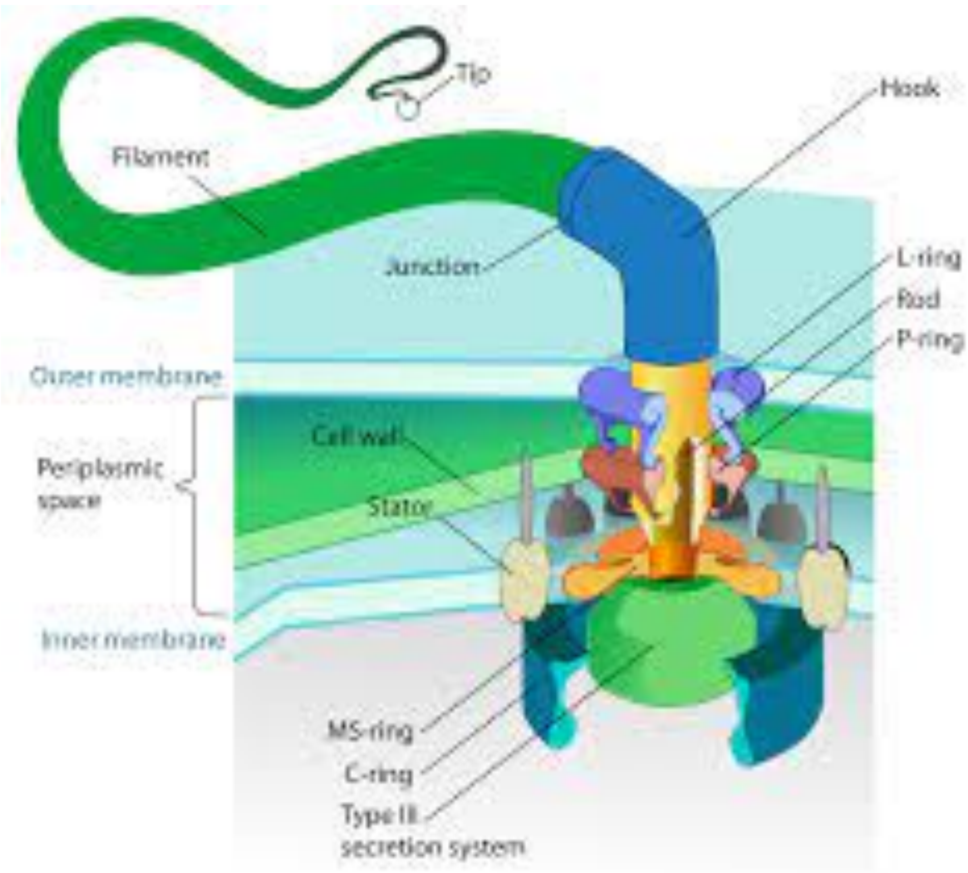


E. Coli Flagella



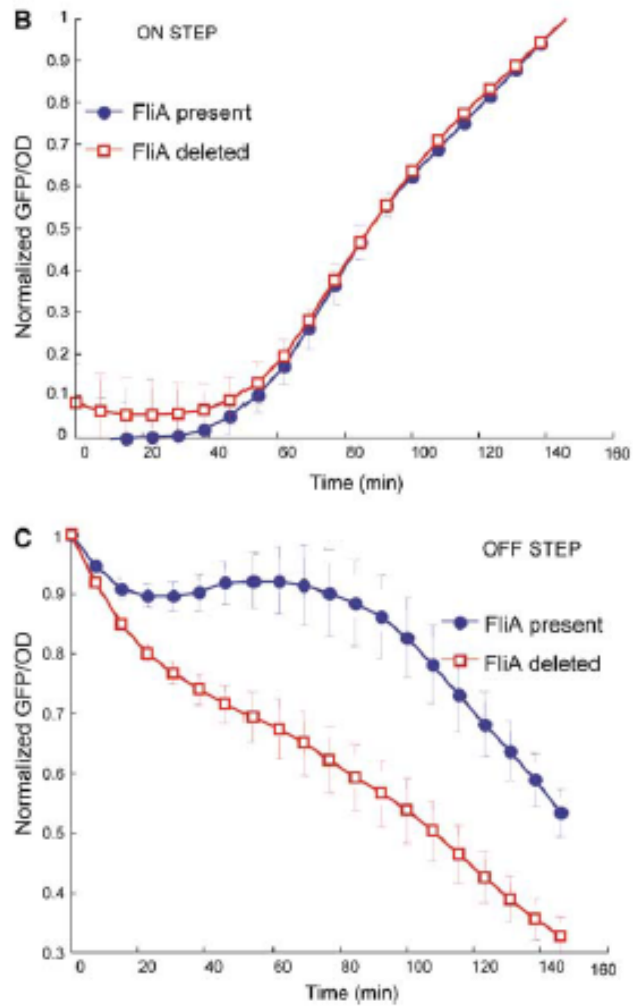
http://www.rowland.harvard.edu/labs/bacteria/movies/ecoli.php

E. Coli Flagella



Delayed OFF Response in Flagella Motor Protein Transcription

Blue: wild-type
Red: FIA deleted



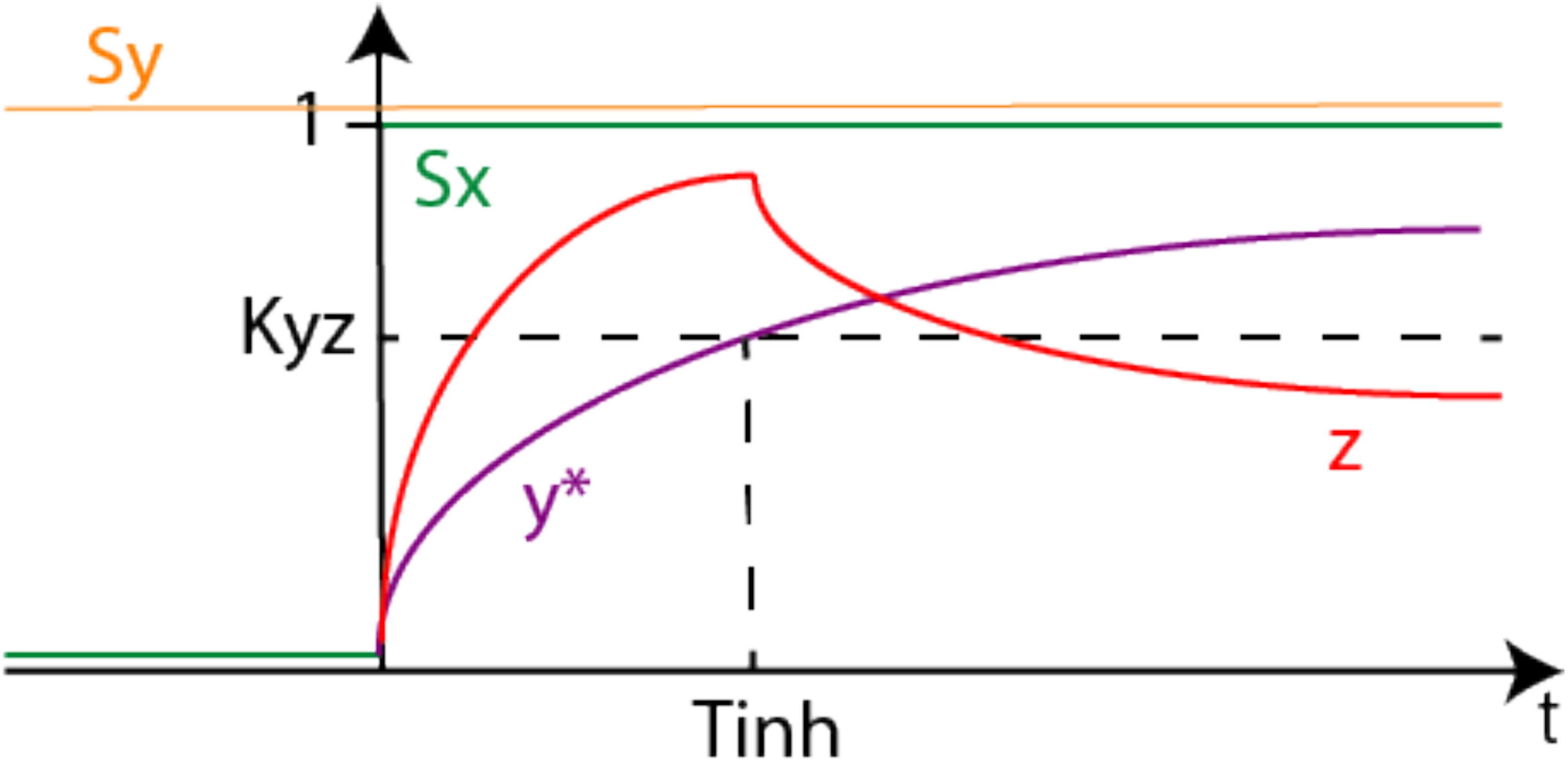
ON response

OFF response

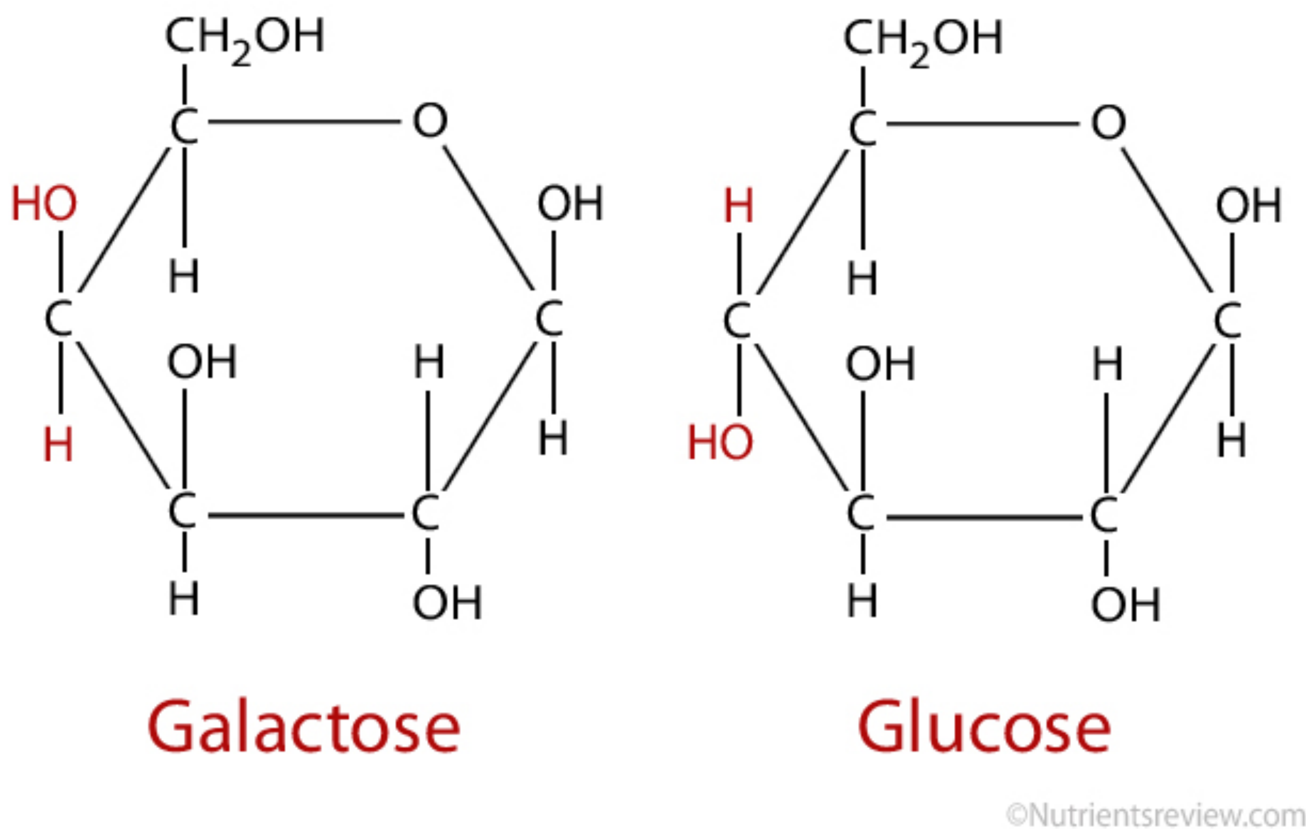
ON and OFF responses to signal in E. coli with C1-FFL-OR motif.
From Kalir et al., 2005.

Incoherent Type 1 Motif Dynamics

Pulsed ON response

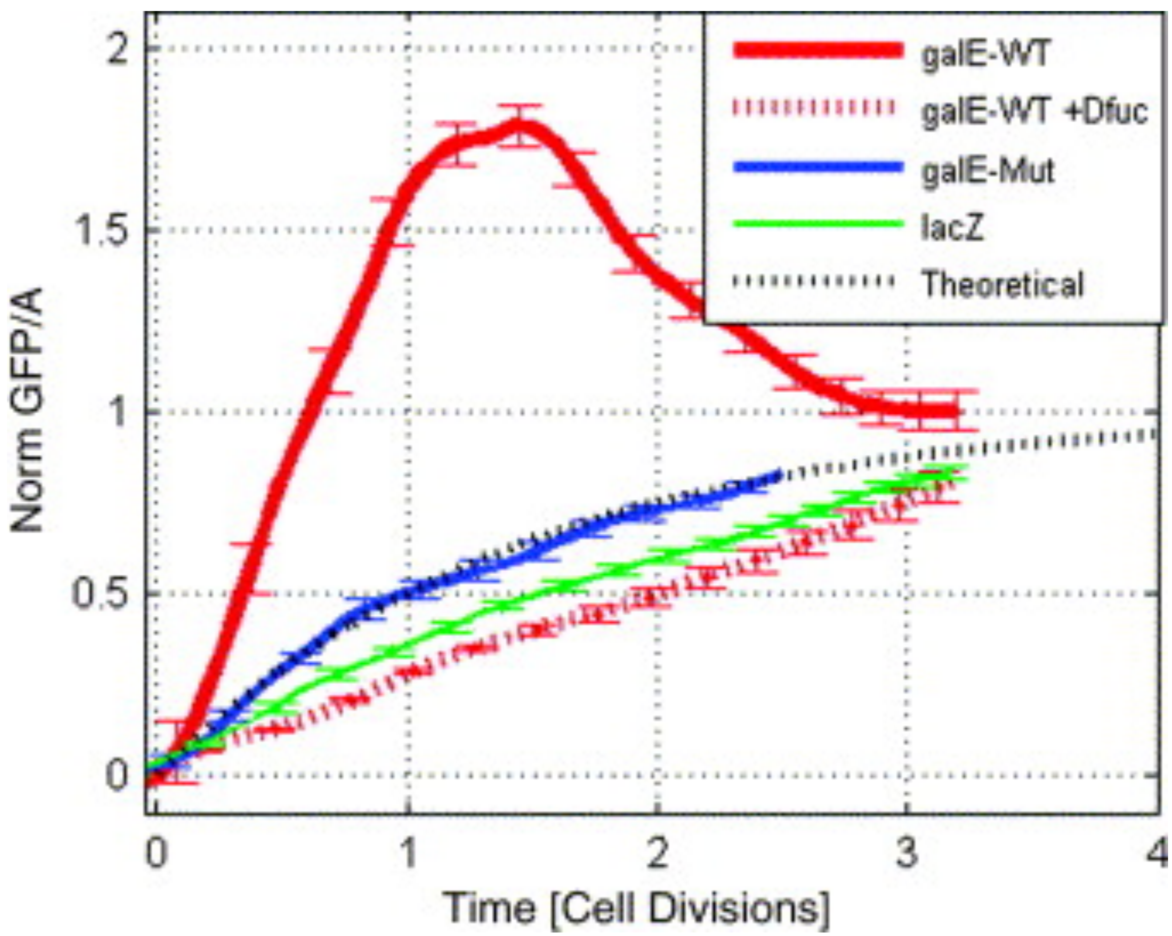


Galactose vs. Glucose



Galactose is about as sweet as glucose. When galactose is linked with glucose it forms lactose.

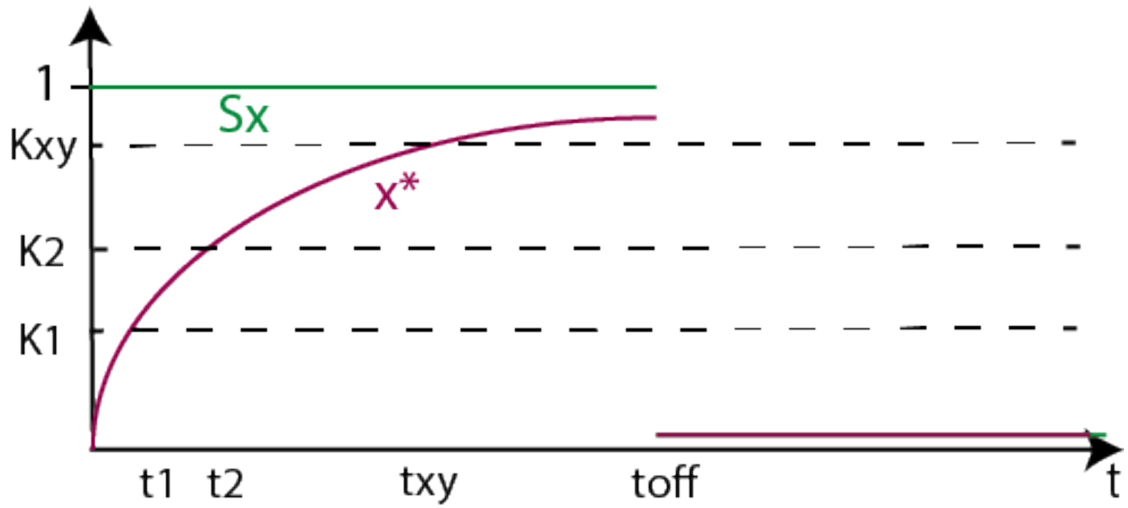
On-Pulse in the Galactose System in E. coli



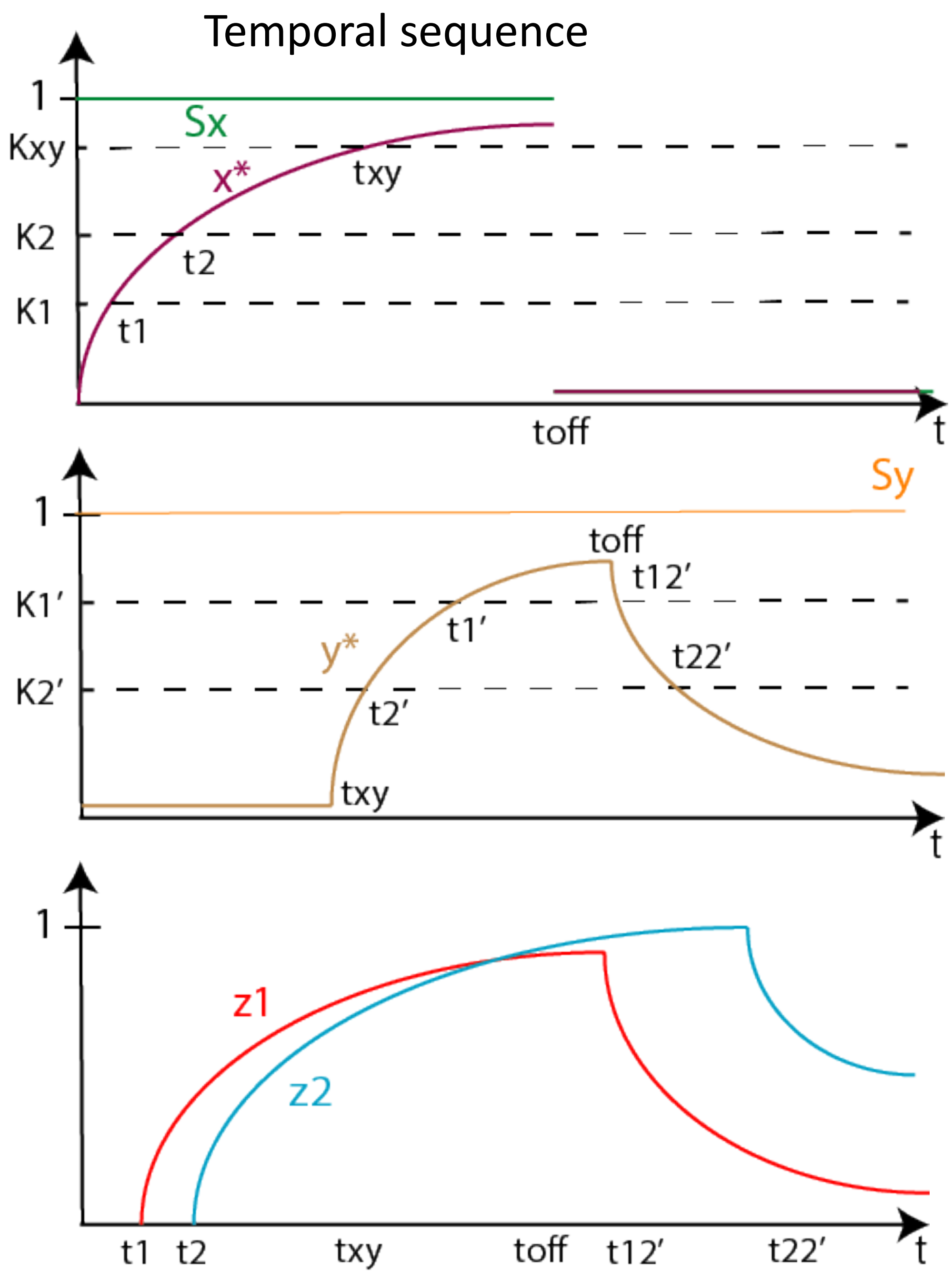
Wild type I1-FFL rises faster and exhibits a pulse, compared to a mutant with repressor not expressed and a simple transcription model. From Mangan et al., 2006.

Multi-Output Motif Dynamics

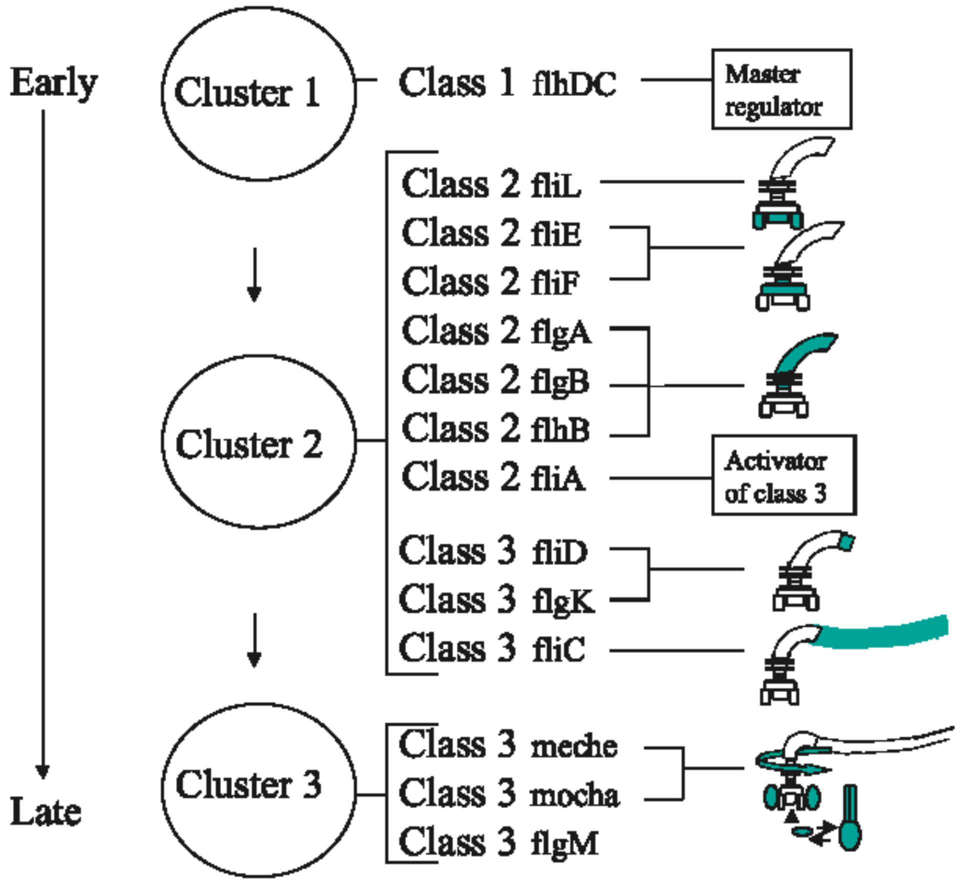
Temporal sequence



Multi-Output Motif Dynamics

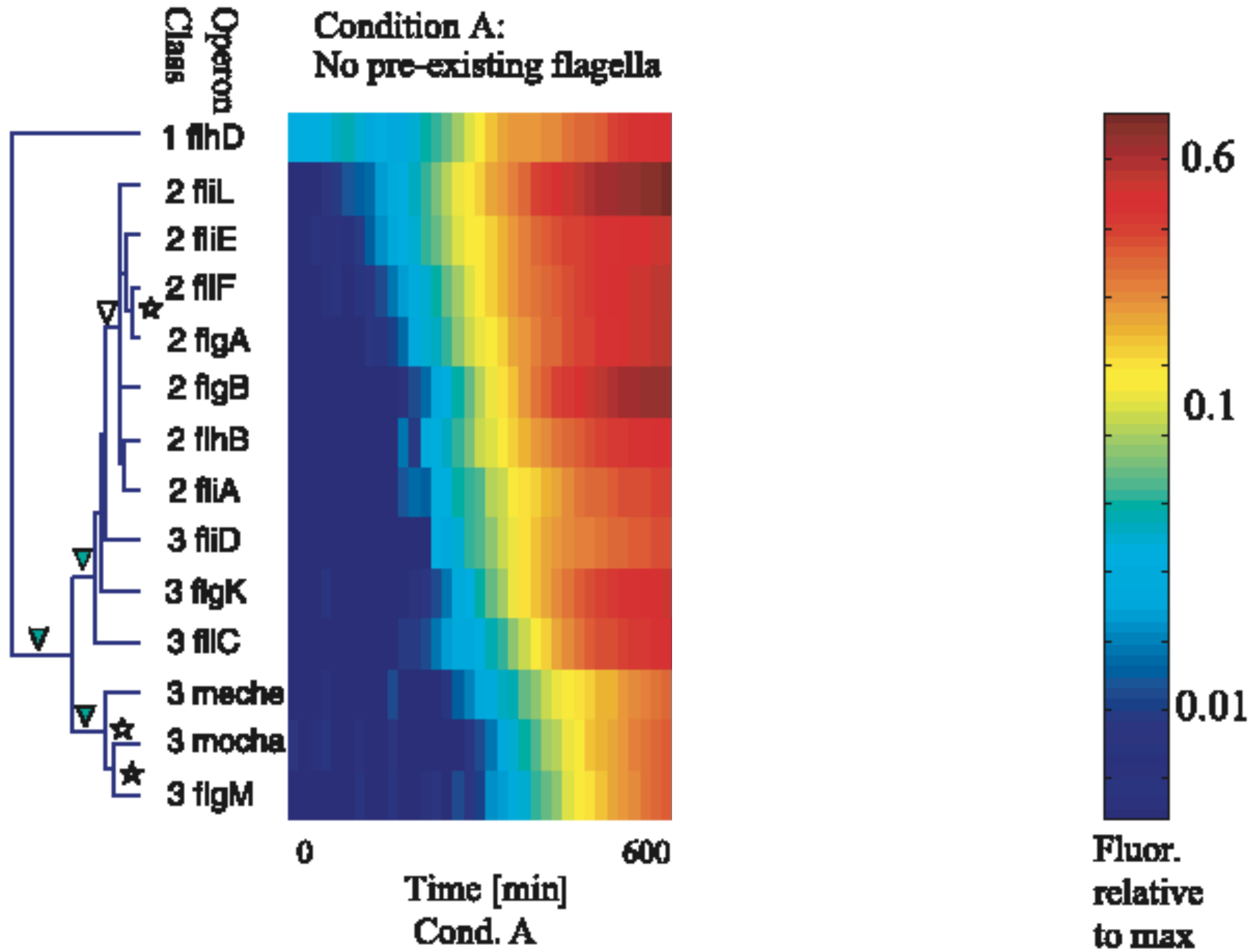


Sequential Production of Flagella Proteins



Sequential production of flagella proteins. Cyan indicates gene product. From Kalir et al., 2001.

Sequential Activation of Flagella Proteins



Sequential production of flagella proteins. From Kalir et al., 2001.

Genetic Toggle Switch

Goal: Design a genetic circuit that will exhibit bistability. A system with bistability is like a “toggle switch”.

This first example of synthetic biology was published in *Nature* in 2000.

Genetic Toggle Switch

letters to nature

Construction of a genetic toggle switch in *Escherichia coli*

Timothy S. Gardner*†, Charles R. Cantor* & James J. Collins*†

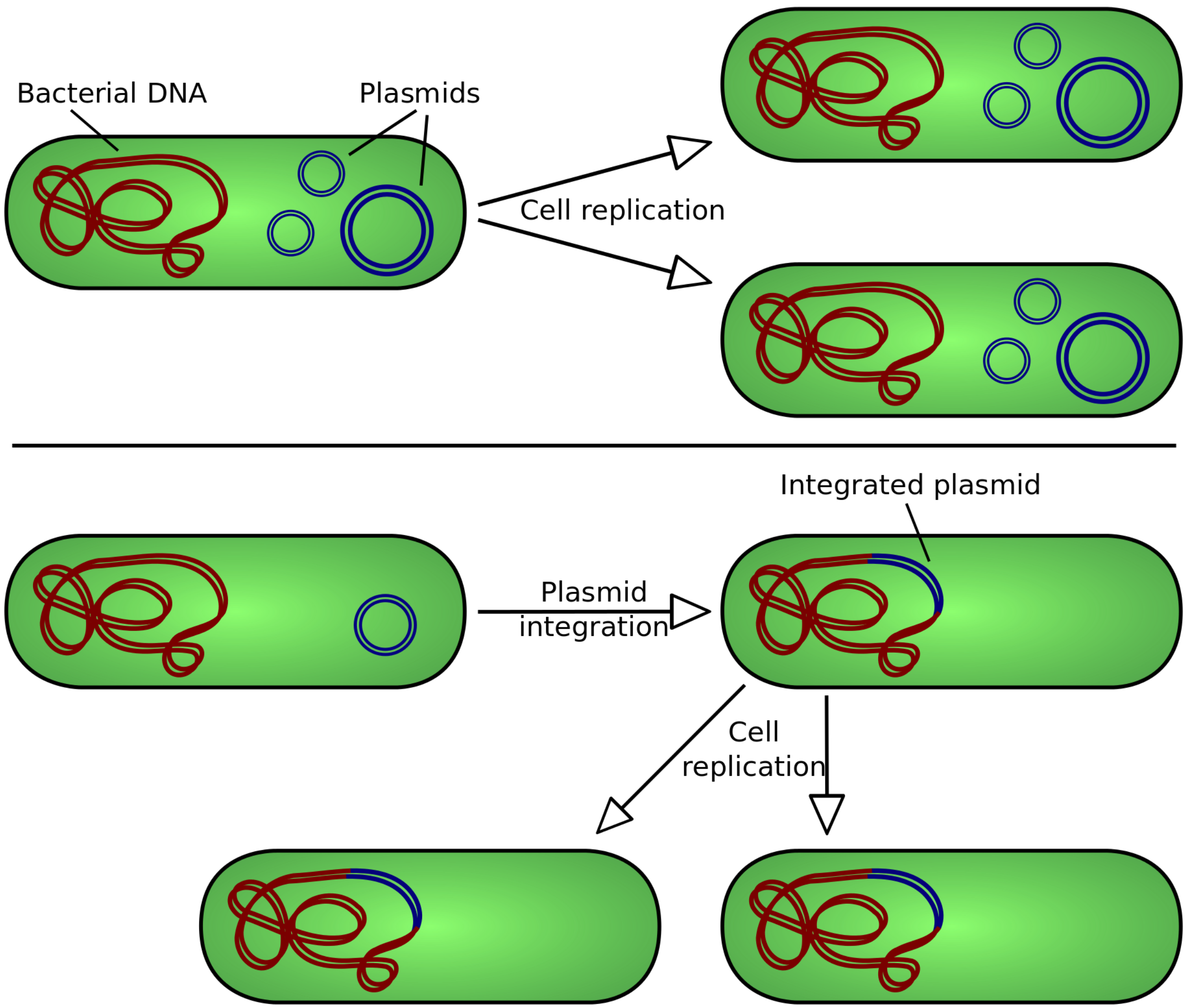
** Department of Biomedical Engineering, † Center for BioDynamics and ‡ Center for Advanced Biotechnology, Boston University, 44 Cummington Street, Boston, Massachusetts 02215, USA*

robust and more difficult to tune experimentally. In addition, the chosen toggle design does not require any specialized promoters, such as the P_R/P_{RM} promoter of bacteriophage λ. Bistability is possible with any set of promoters and repressors as long as they fulfil the minimum set of conditions described in Box 1 and Fig. 2.

The bistability of the toggle arises from the mutually inhibitory arrangement of the repressor genes. In the absence of inducers, two stable states are possible: one in which promoter 1 transcribes repressor 2, and one in which promoter 2 transcribes repressor 1. Switching is accomplished by transiently introducing an inducer of the currently active repressor. The inducer permits the opposing

Gardner et al., Nature, 403:339, 2000

Genes Introduced Into Bacteria Using Plasmids



Non-integrating plasmids

Episomes

Plasmid: a small, circular piece of DNA that is separate from chromosomal DNA and can replicate independently

Current Applications of Synthetic Biology

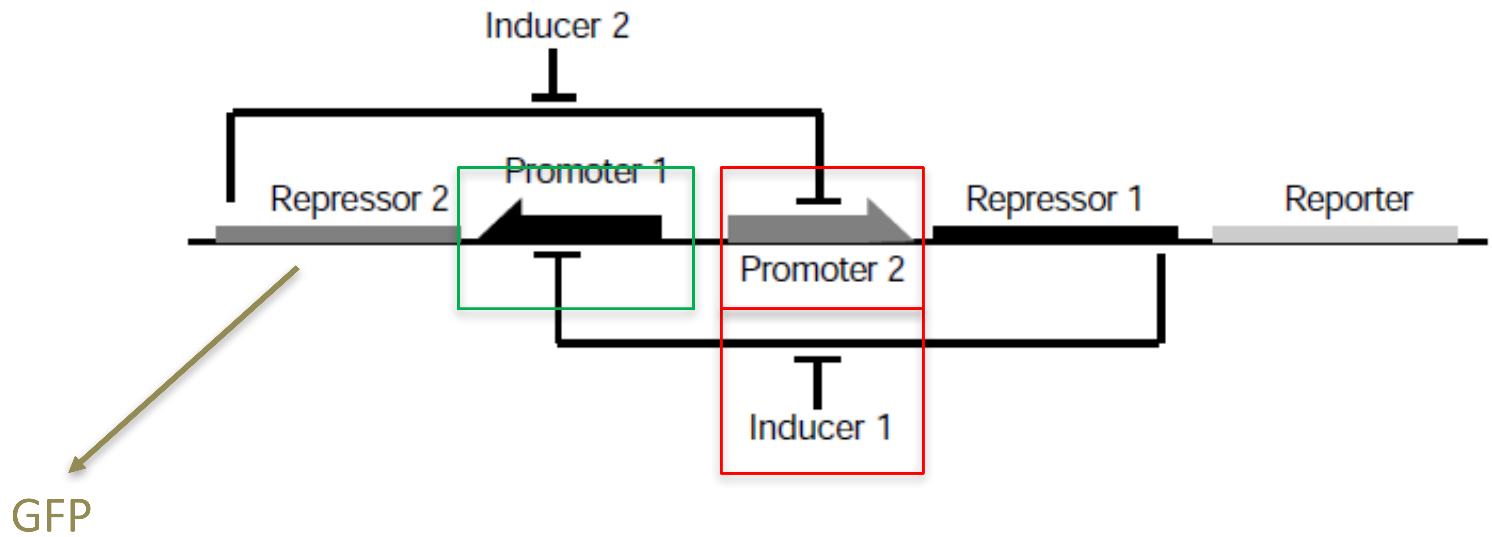
Production of chemical and industrial enzymes for human medicine, animal health, and crop protection from insects.

Biofuels

Creation of vaccines

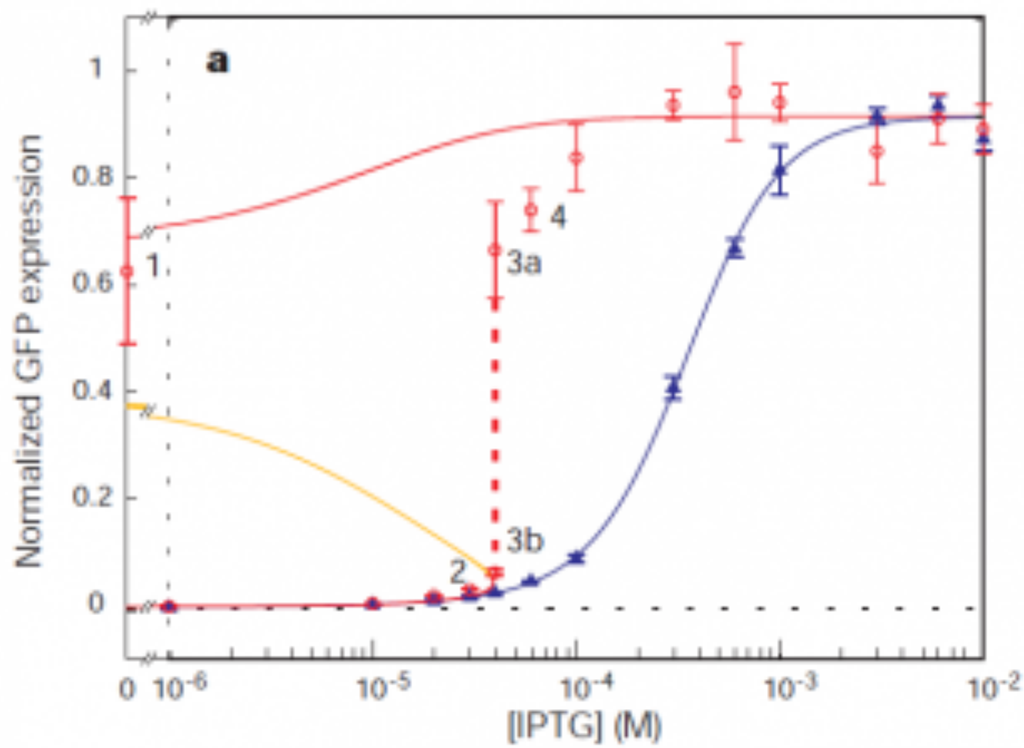
Chemicals for fragrances and flavours

Genetic Toggle Switch Design



From Gardner et al., Nature, 403:339, 2000

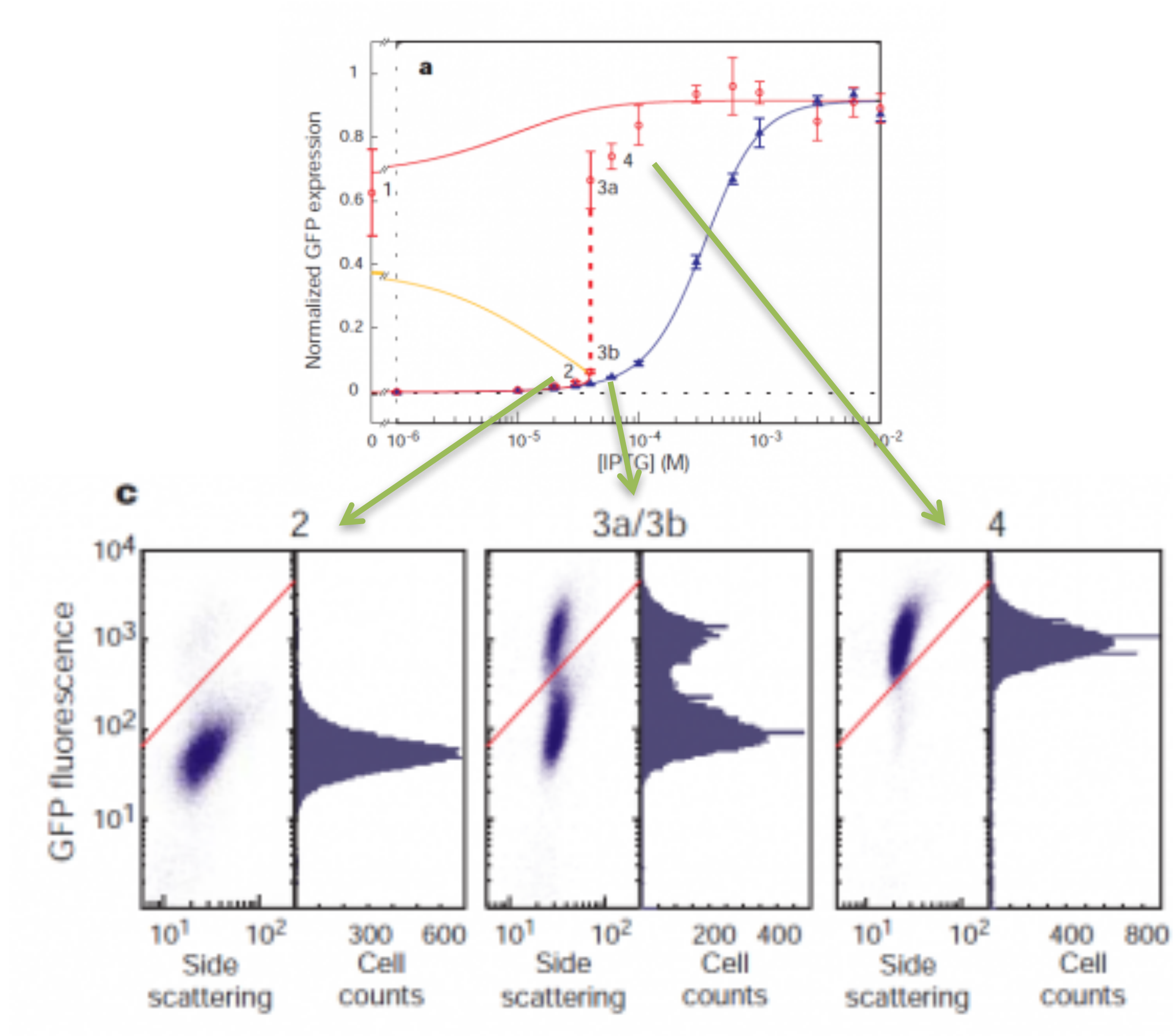
Experimental Evidence for Toggle Switch



IPTG = inducer 1
Red points = repressor 2 (y)
Red curves = model lower and upper branches of stable equilibria
Yellow curve = model middle equlibrium branch
Blue curve = standard sigmodial response to inducer (without bistability)

From Gardner et al., Nature, 403:339, 2000

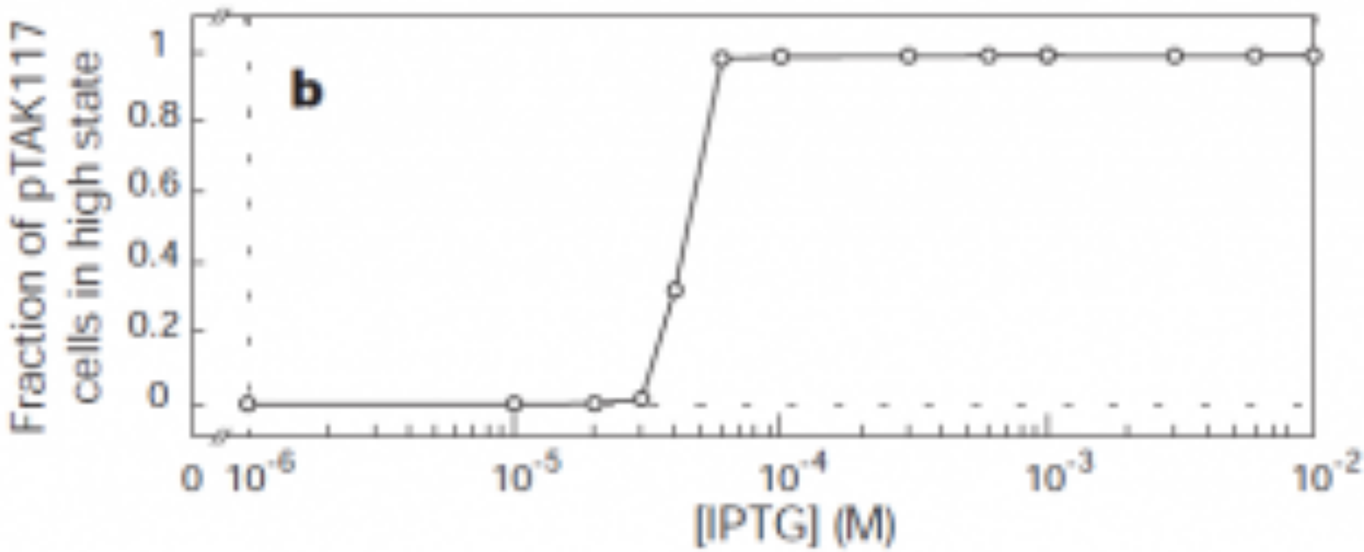
Experimental Evidence for Toggle Switch



Distribution of cells expressing different levels of GFP fluorescence.

From Gardner et al., Nature, 403:339, 2000

Experimental Evidence for Toggle Switch



Another view of the discontinuous jump, showing fraction of cells in the high state.

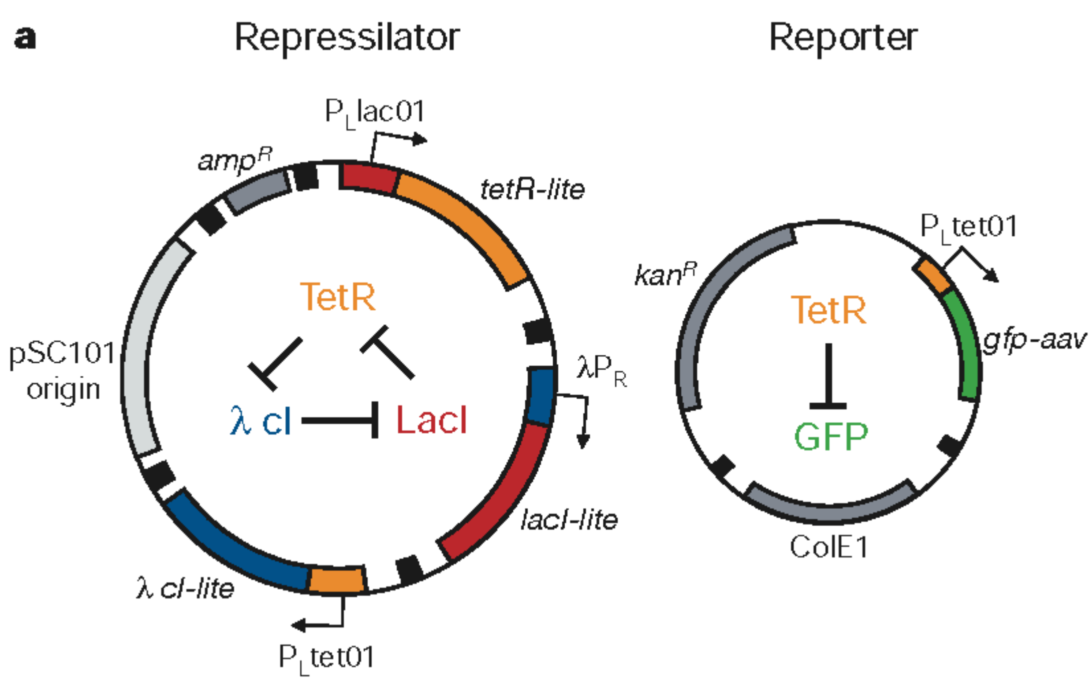
From Gardner et al., Nature, 403:339, 2000

The Repressilator

Goal: Construct a genetic circuit that will produce sustained oscillations.

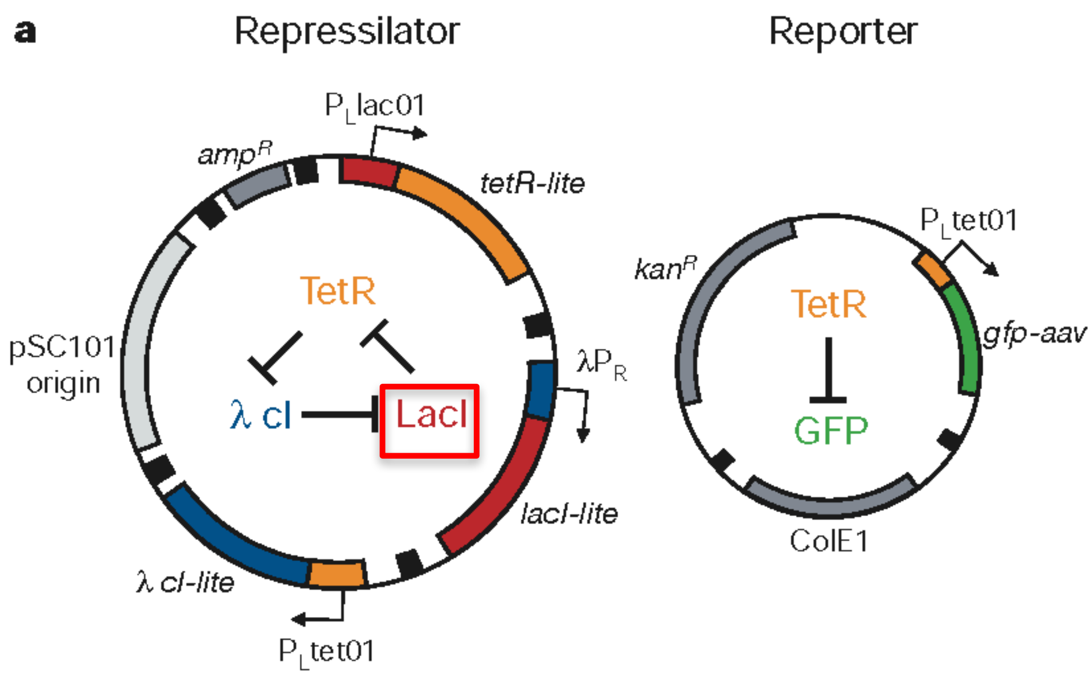
This second example of synthetic biology was published in the same issue of *Nature* as the first example, in the year 2000.

Genetic Design of the Repressilator



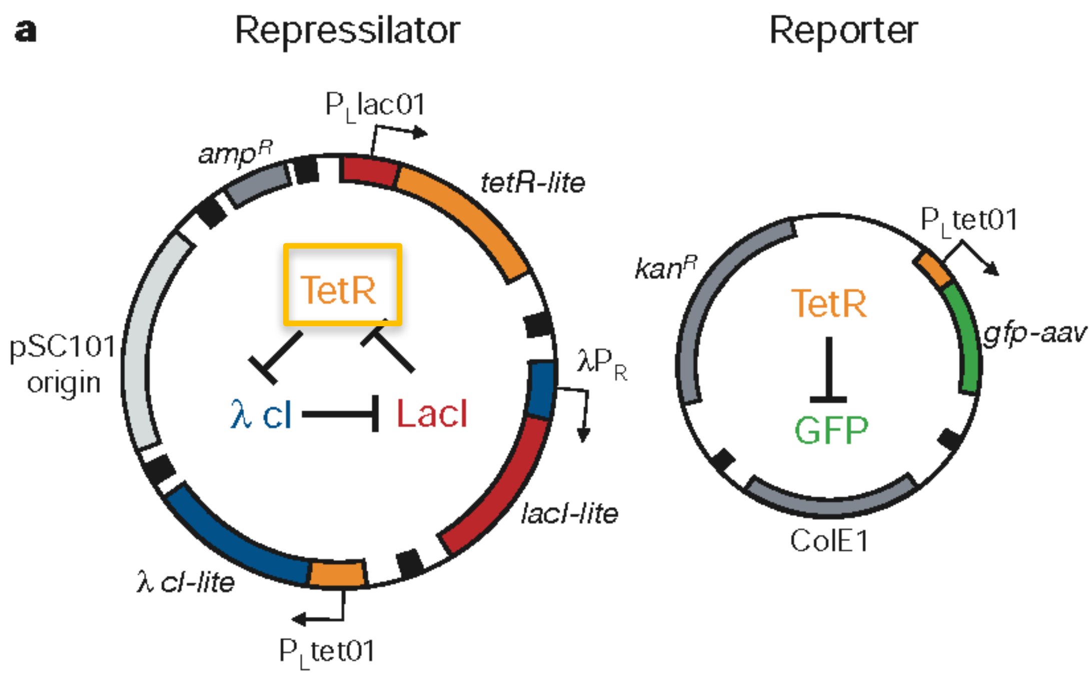
Three repressor genes and a reporter gene transfected into *E. coli.*
From Elowitz and Leibler, Nature, 403:335, 2000.

Genetic Design of the Repressilator



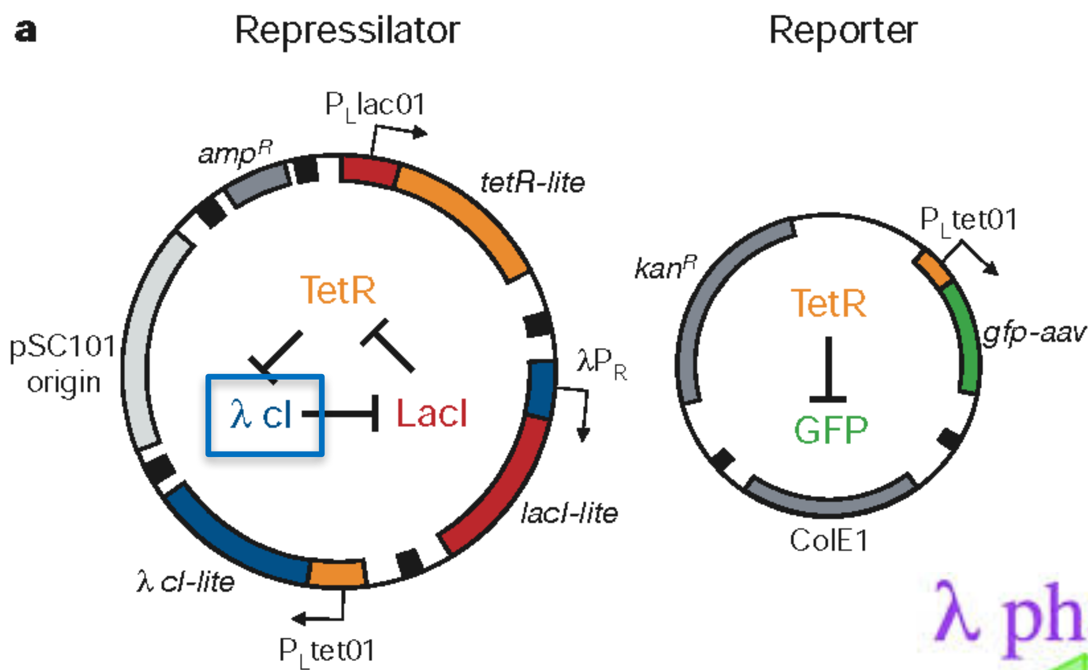
LacI – A repressor from *E. coli* involved in the metabolism of lactose and galactose.

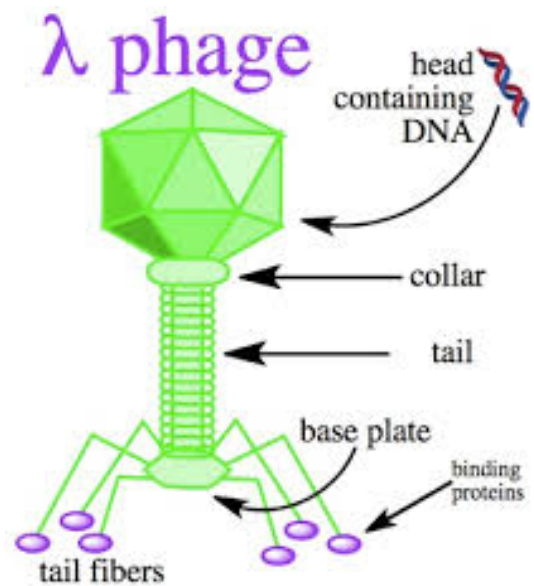
Genetic Design of the Repressilator



tetR – A repressor from the Tn10 transposon. A transposon is a sequence of DNA whose gene product changes the location of the gene in the DNA by cutting, followed by diffusion, followed by random pasting. Confers antibacterial resistance. Discovery by Barbara McClintock, who won a 1983 Nobel Prize.

Genetic Design of the Repressilator





λcI – A repressor from λ phage (a bacteriophage for *E. coli*).

The Mathematical Model

6 coupled ODEs

$$\frac{dm_i}{dt} = -m_i + \frac{\alpha}{(1+p_j^n)} + \alpha_0 \qquad \left(\begin{array}{l} i = lacI, tetR, cI \\ j = cI, lacI, tetR \end{array}\right)$$

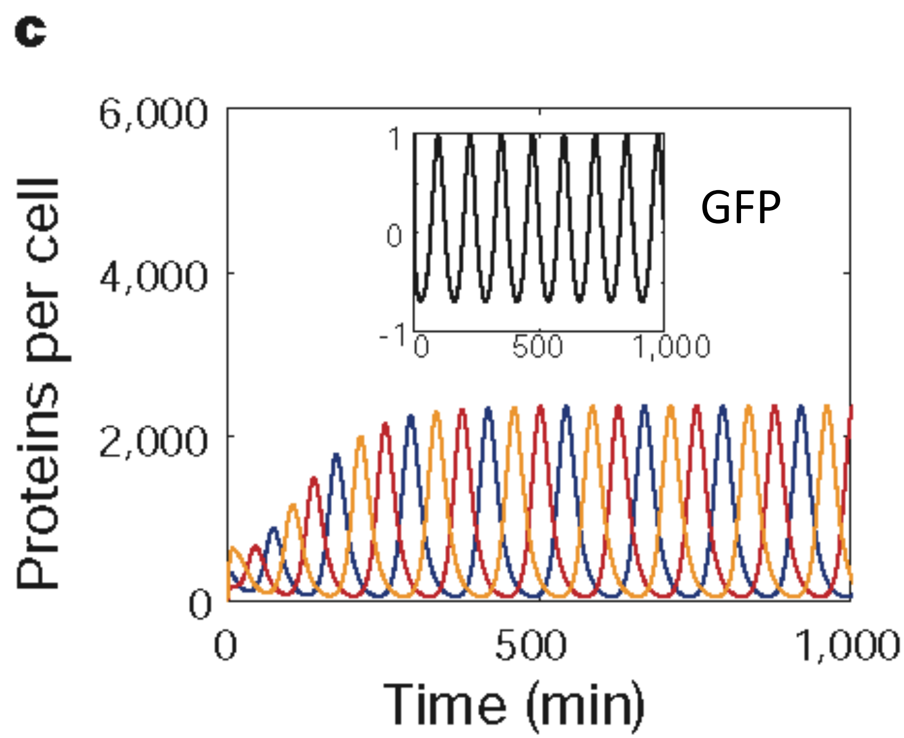
$$\frac{dp_i}{dt} = -\beta(p_i - m_i)$$

m = mRNA
p = protein

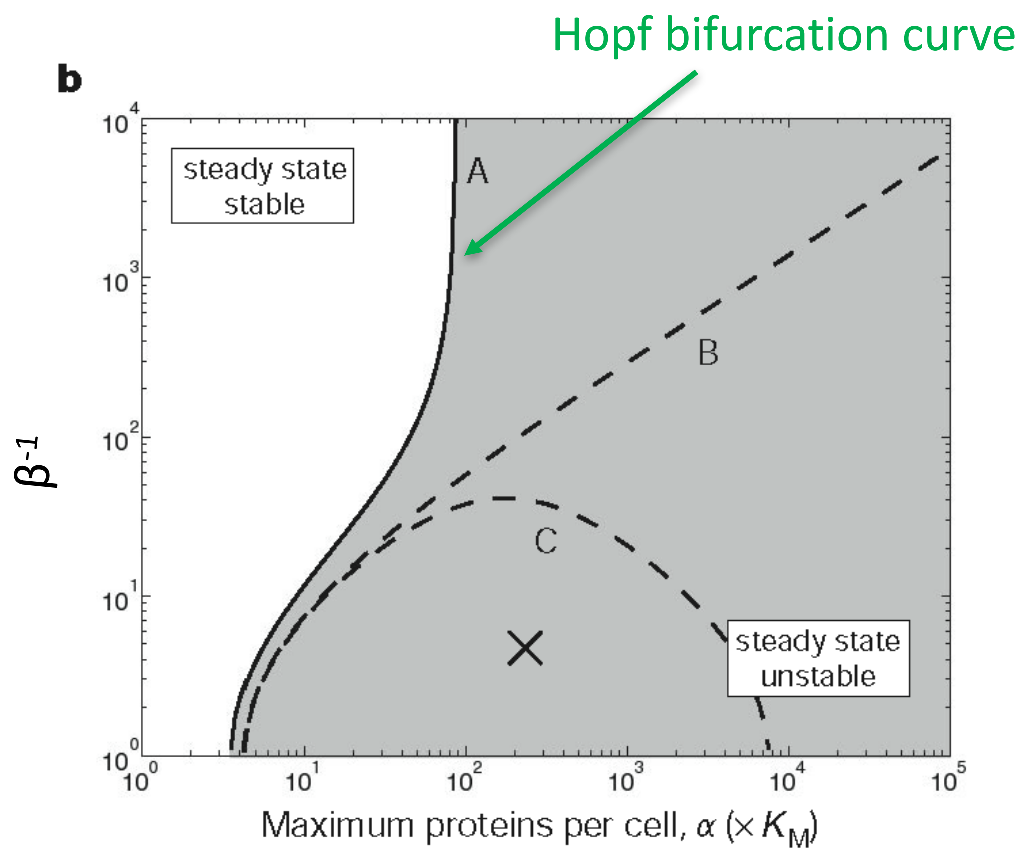
α = transcription rate for component subject to repression
α_o = basal transcription rate
β = protein degradation rate divided by mRNA degradation rate
n = cooperativity

Oscillations are Possible with the Model

LacI
TetR
λ cI



Two-Parameter Bifurcation Diagram



$$\frac{dm_i}{dt} = -m_i + \frac{\alpha}{(1+p_j^n)} + \alpha_0 \qquad \begin{pmatrix} i = lacI, tetR, cI \\ j = cI, lacI, tetR \end{pmatrix}$$

$$\frac{dp_i}{dt} = -\beta(p_i - m_i)$$

A: n=2.1, α_o=0
B: n=2, α_o=0
C: n=2, $\alpha_o/\alpha=10^{-3}$

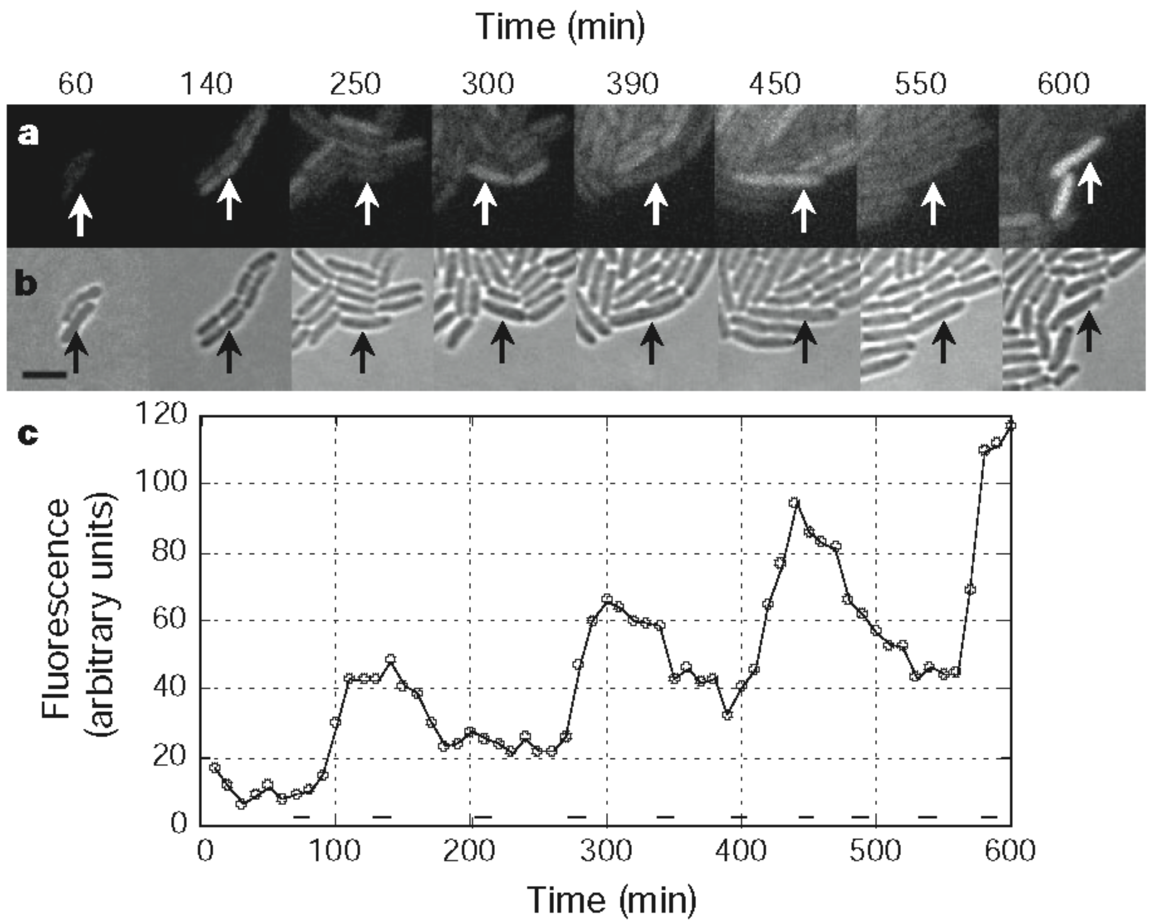
Modeling Insights

- α_o/α should be small, so tightly repressible promoters are used (minimize transcriptional leakiness).
- β^{-1} should be small, which means protein lifetime should be short. This was achieved by adding a tag onto the mRNA of each repressor gene, so that the gene product is flagged for rapid degradation by proteases.

GFP Oscillations in a Single Cell

Fluorescence

Bright field



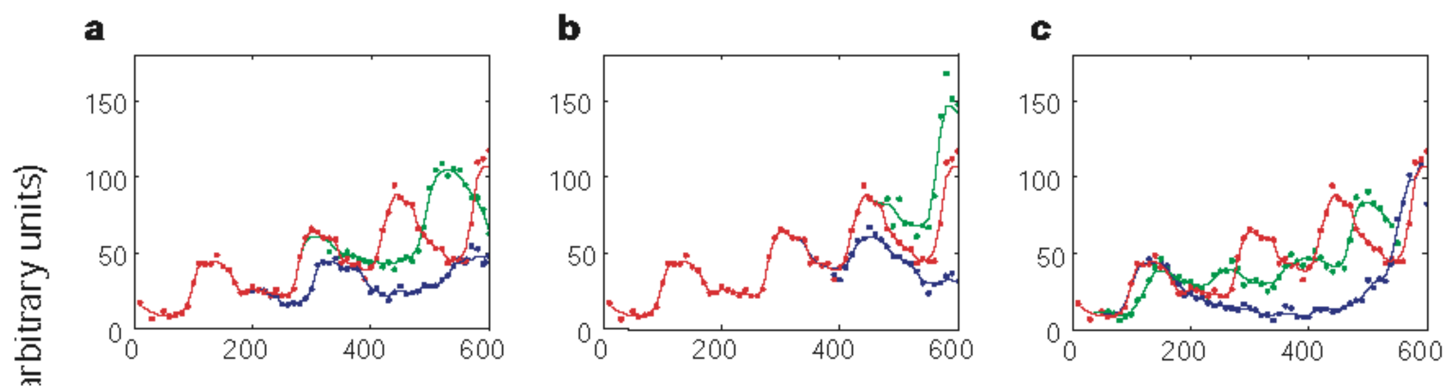
Each dash represents cell division (septation).

From Elowitz and Leibler, Nature, 403:335, 2000.

Observations

- Oscillations occur in the transfected E. coli!
- The oscillation period is roughly three times that of the time required for cell division. This means that the network that maintains oscillations is inherited by the progeny cells.

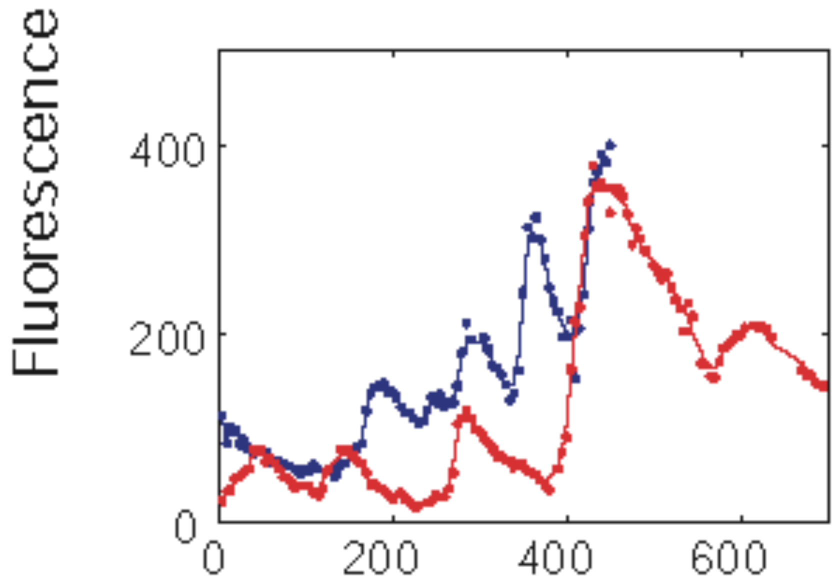
Oscillations Among Progeny are Sometimes Correlated, Sometimes Not



Red: Original cell
Blue and green: Sister cells

From Elowitz and Leibler, Nature, 403:335, 2000.

Cells from Different Experiments Exhibit Different Oscillations



Red and blue: cells from different batches of bacteria

From Elowitz and Leibler, Nature, 403:335, 2000.