

Computational Science and Engineering

Computational Cell Biology

Lecture 4

Case Study: Basic Modeling in Gene Expression

Yang Cao

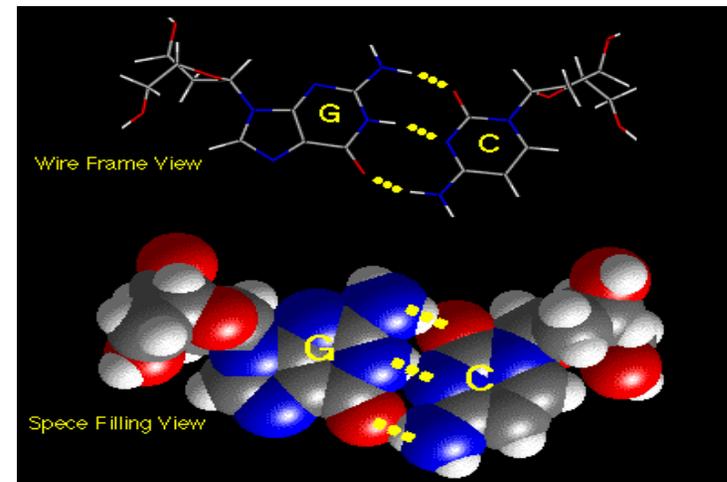
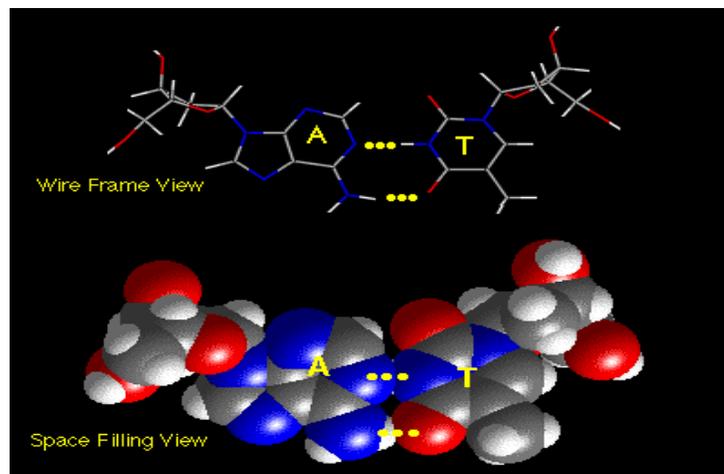
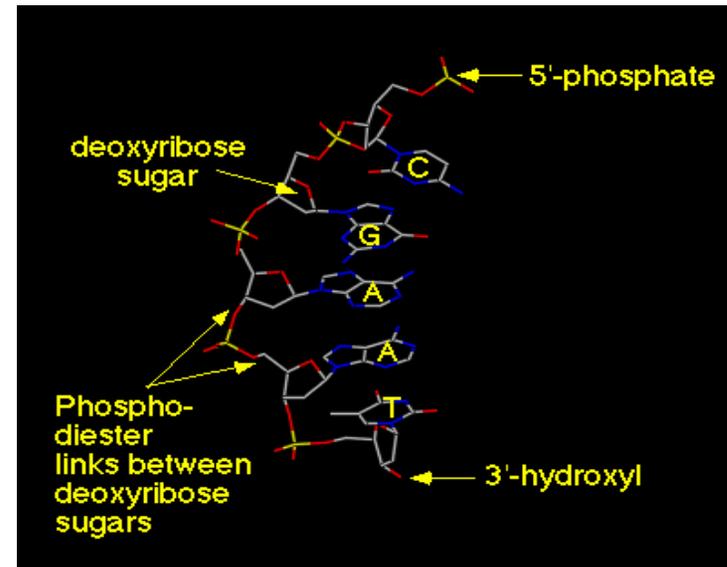
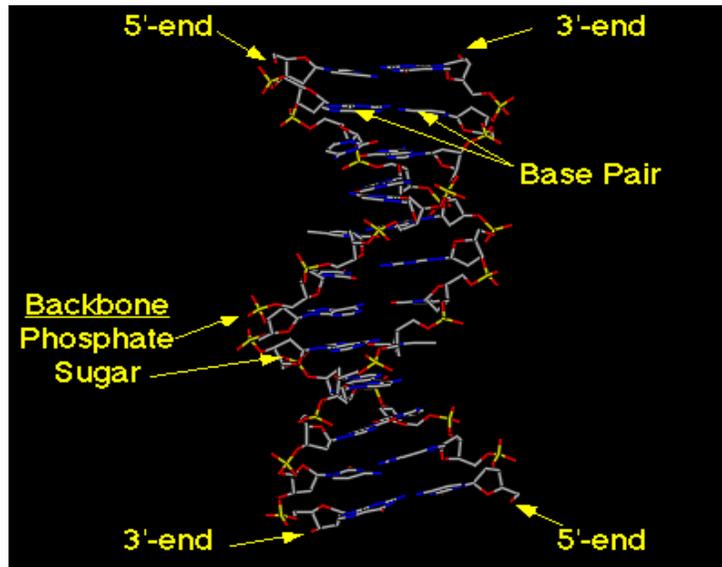
Department of Computer Science

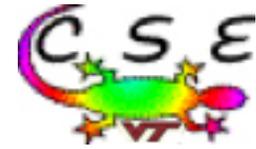




DNA Structure and Base Pair

Computational Science and Engineering

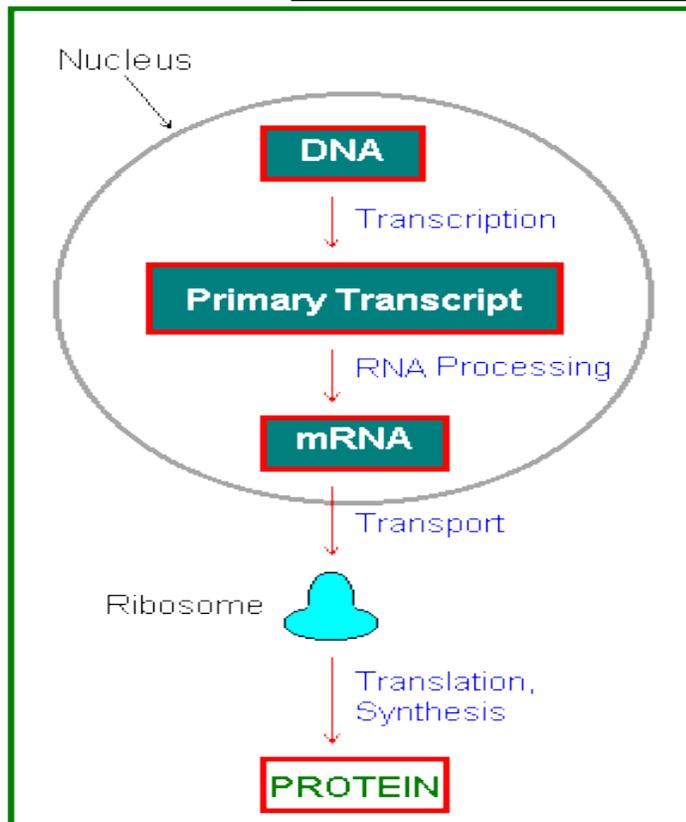




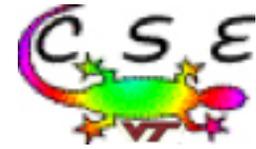
Gene Expression

Gene is just a small part of DNA. The gene expression follows the process of:

DNA → RNA → Protein



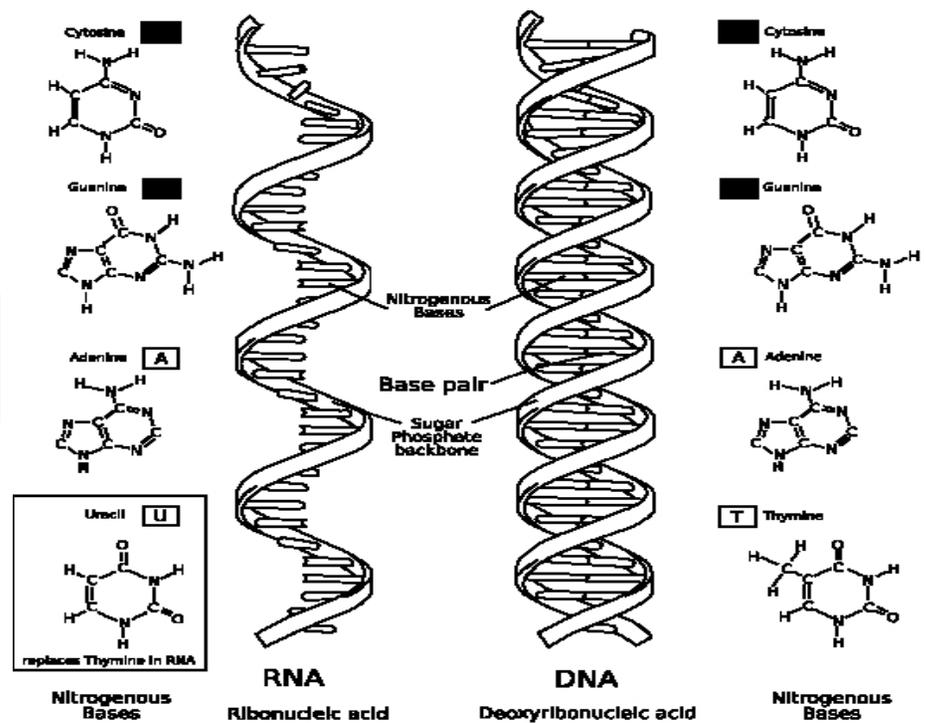
Gene expression shows big difference between prokaryotic and eukaryotic cells. Most of the models of gene networks proposed in literature are for prokaryotic cells.

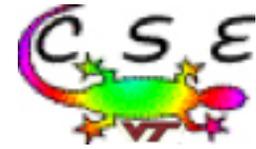


Transcription

Transcription from DNA to RNA is based on the base pair.
However RNA doesn't have "T", instead it has "U",
which pairs with "A" just as "T" does.

DNA → RNA

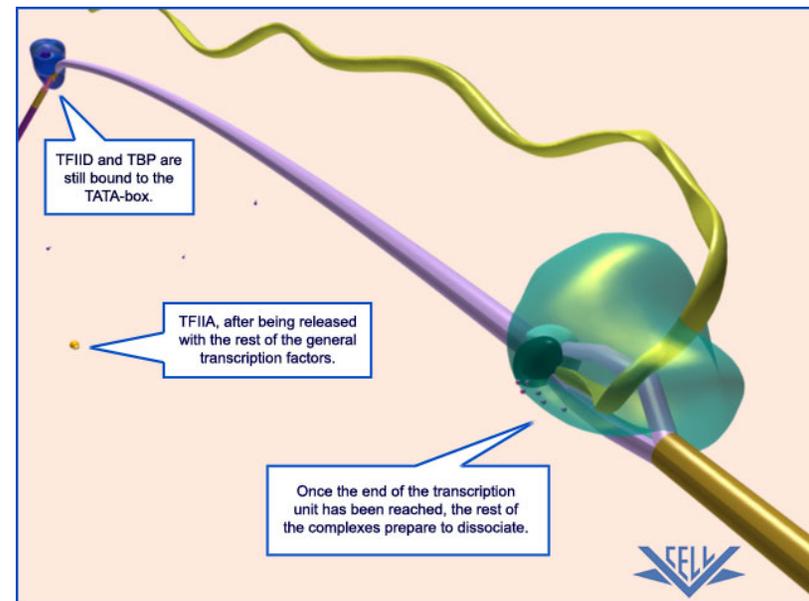
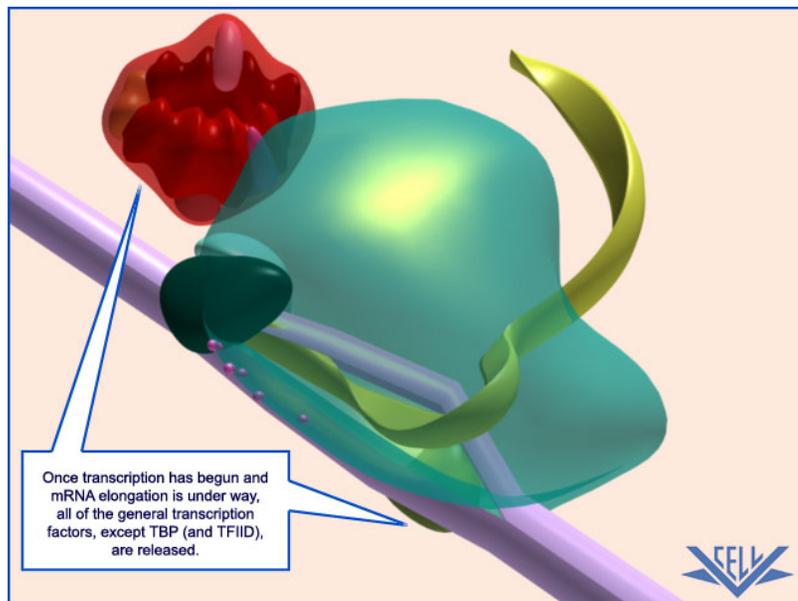




Transcription (Continue)

Computational Science and Engineering

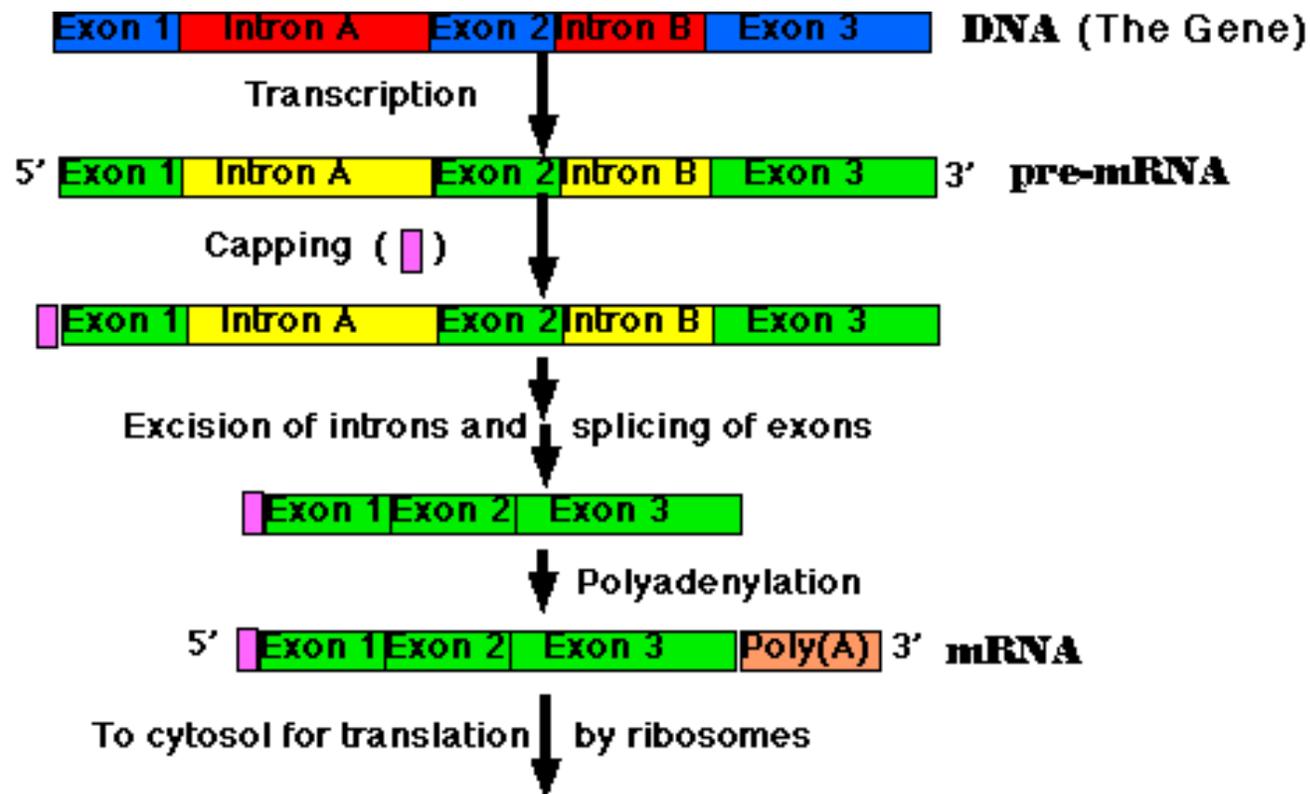
- Transcription is carried out by the enzyme RNA Polymerase (RNAP)
- Several types of RNA are produced
 - mRNAs
 - rRNAs
 - tRNAs
 - Small RNAs (can regulate transcription)
- Transcription occurs only on one strand of DNA

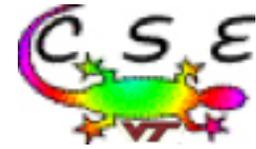




RNA Processing

- In Prokaryotic cells, mRNA can be immediately translated by ribosomes.
- In Eukaryotic cells, RNA has to be processed and then transported to cytoplasm.

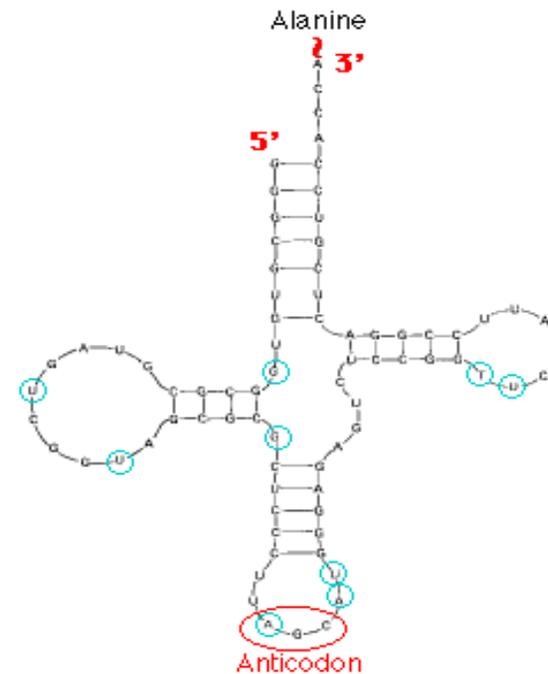
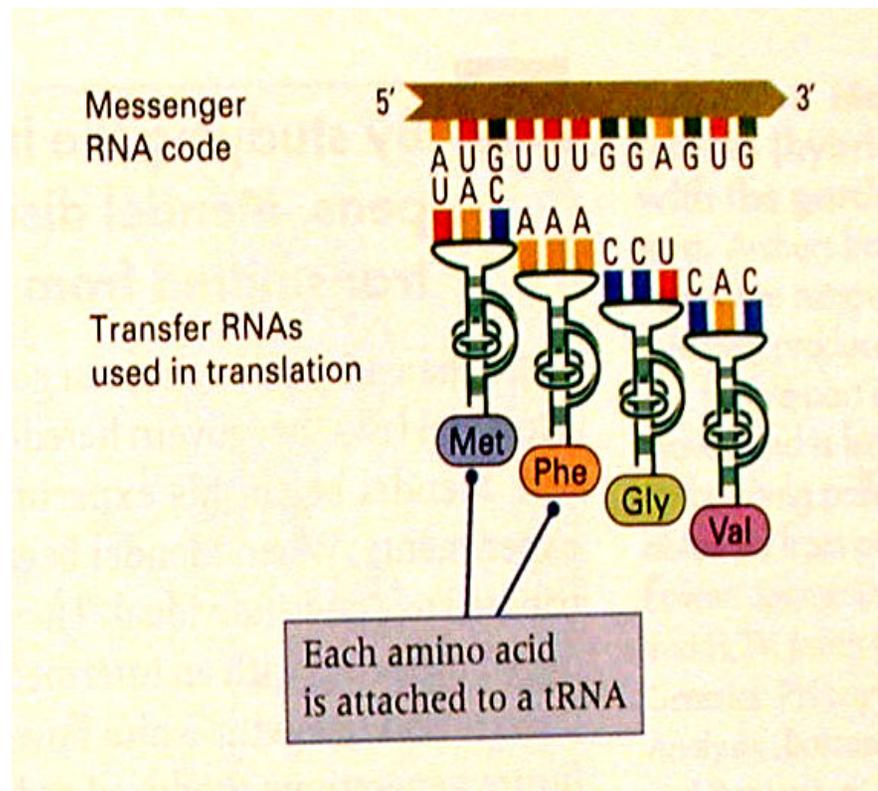




Translation: From RNA to Protein

Computational Science and Engineering

- An mRNA sequence is decoded in sets of three nucleotides, called codon.
- Amino acids are specified by codons (not one by one).
- Amino acids and codons are connected by tRNAs.



tRNA



Translation: From RNA to Protein

- Genetic Code is universal

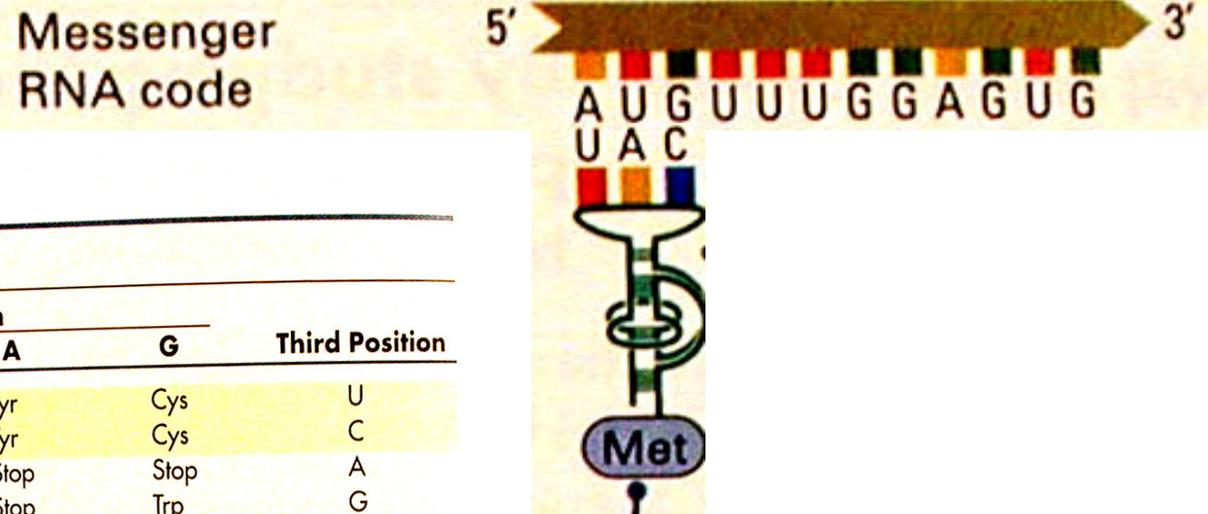


Table 6.1 The genetic code

First Position	Second Position				Third Position
	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Stop	Stop	A
	Leu	Ser	Stop	Trp	G
	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
C	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
A	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
G	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

AUG = Met

Note: A sequence of three nucleotides forms the nucleic acid codon for a single amino acid. The four nucleotides U, C, A, G can produce 64 different three-nucleotide combinations. All the amino acids except methionine (Met) and tryptophan (Trp) have more than one codon. The "stop" codons UAA, UAG, and UGA do not code for amino acids but signal the end of a protein. All proteins start with methionine.

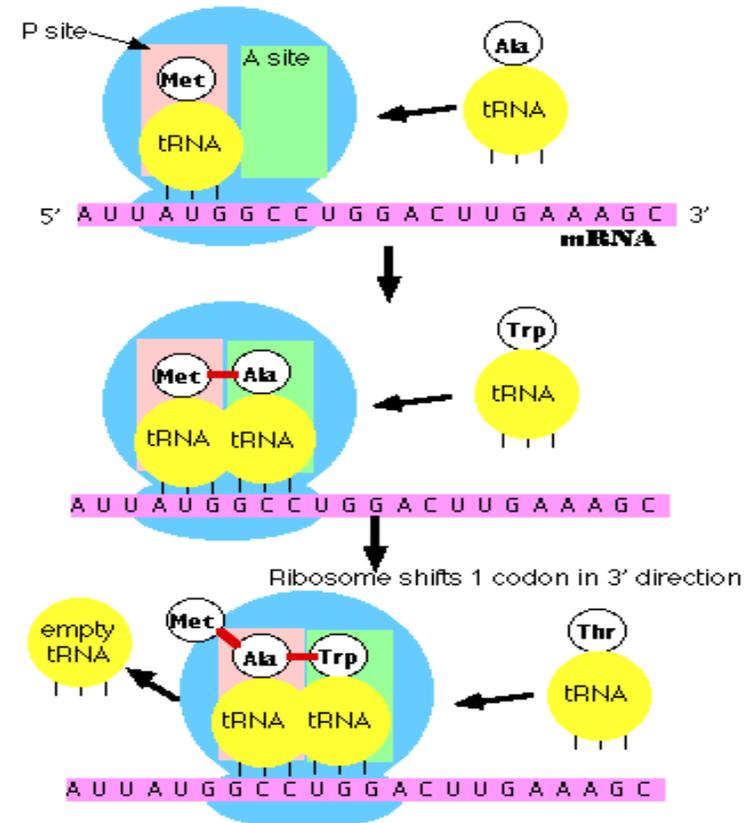
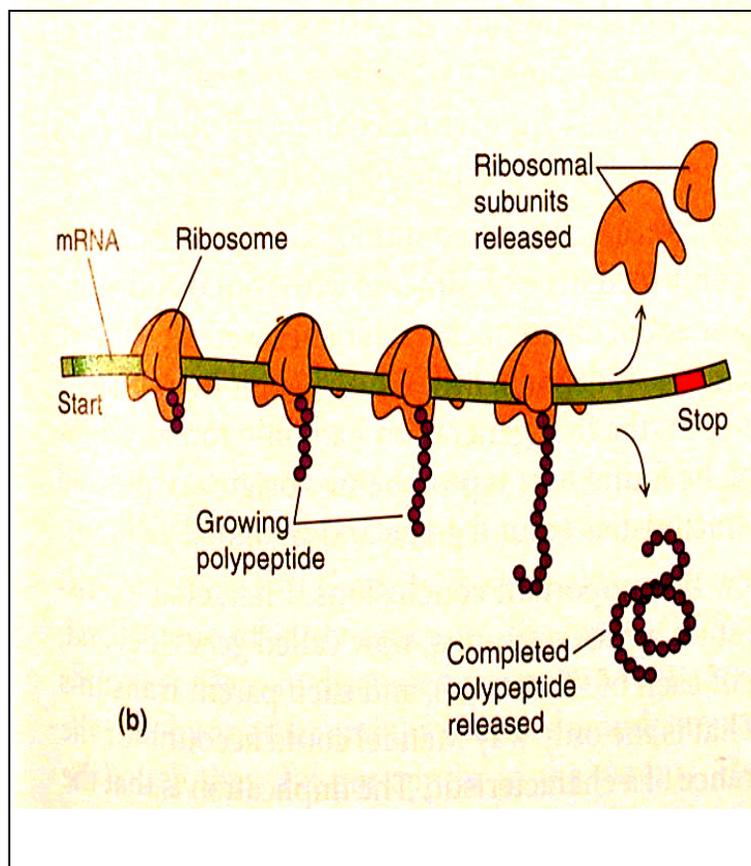
The codons are given as they appear in messenger RNA. The four bases in the nucleotides of ribonucleic acids are uracil (U), cytosine (C), adenosine (A), and guanine (G). The amino acids specified by the genetic code are alanine (Ala), arginine (Arg), asparagine (Asn), aspartic acid (Asp), cysteine (Cys), glycine (Gly), glutamine (Gln), glutamic acid (Glu), histidine (His), isoleucine (Ile), leucine (Leu), lysine (Lys), methionine (Met), phenylalanine (Phe), proline (Pro), serine (Ser), threonine (Thr), tryptophan (Trp), tyrosine (Tyr), and valine (Val).

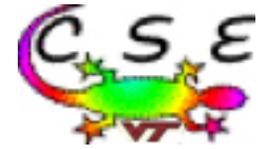


Translation: From RNA to Protein

Computational Science and Engineering

- RNA message is decoded by Ribosomes
- Initiation starts at binding site (prokaryotic) or “AUG”
- Elongation
- Termination

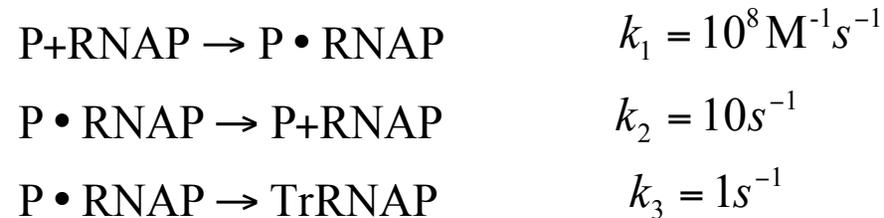




A Model for Prokaryotic Gene Expression

Computational Science and Engineering

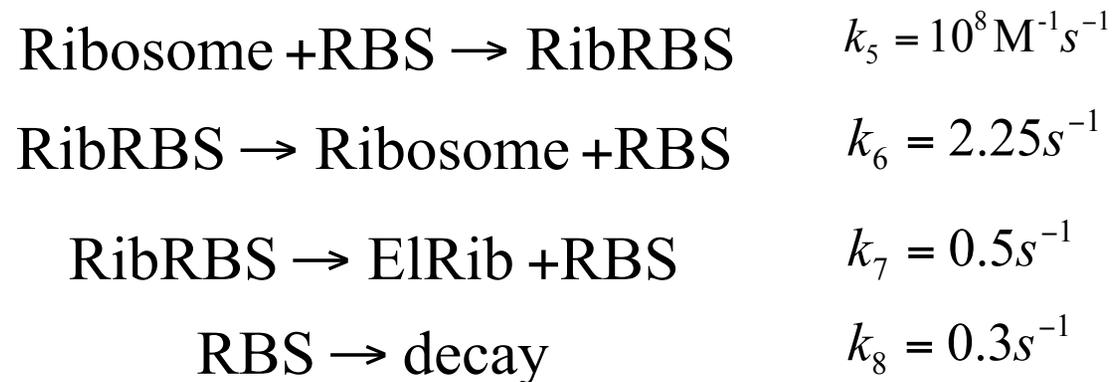
1. Transcription Initiation (the binding and initiation)



2. Elongation (RBS is available before elongation terminates)



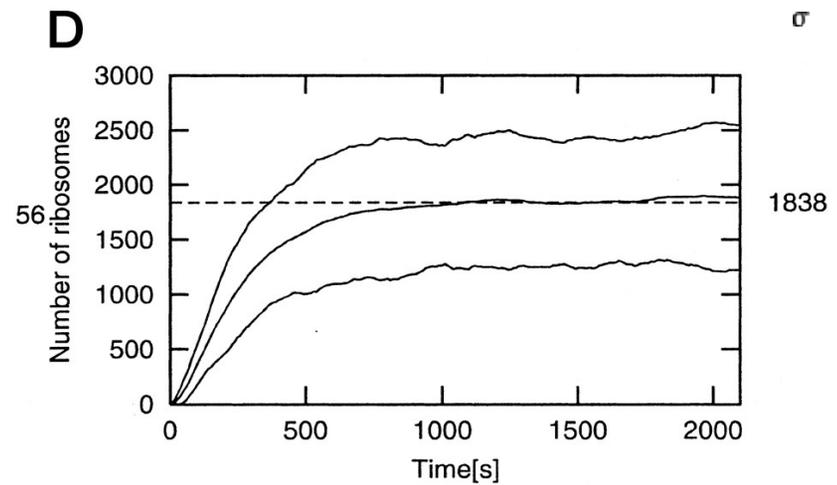
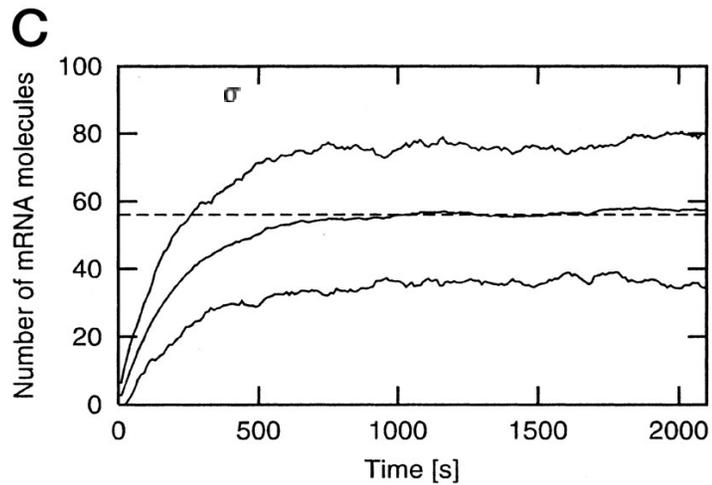
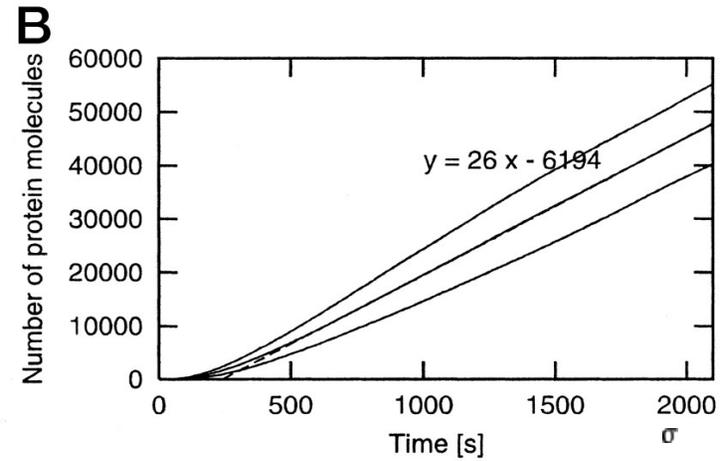
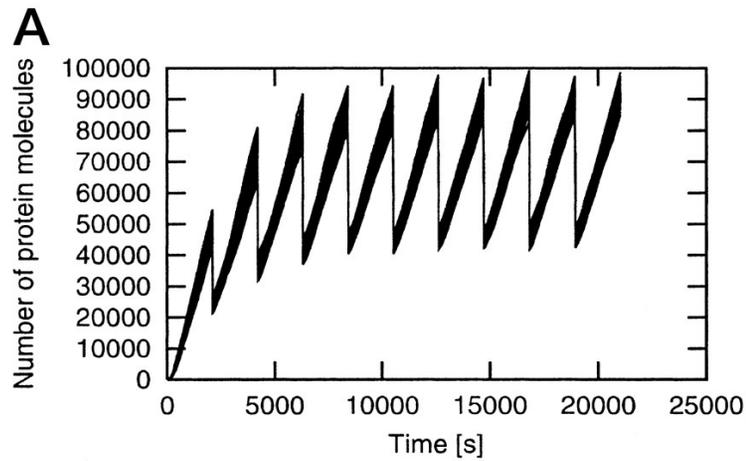
3. Translation Initiation

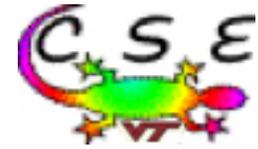


4. Elongation



Simulation Results



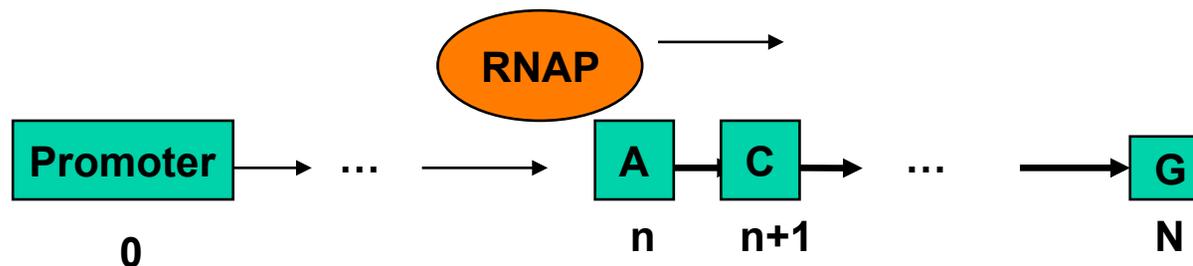
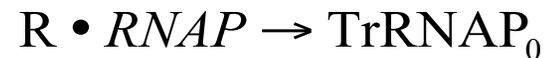


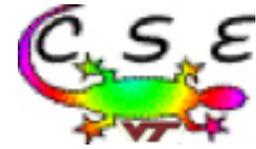
Some Further Discussion

- The elongation process



- A more detailed model for that





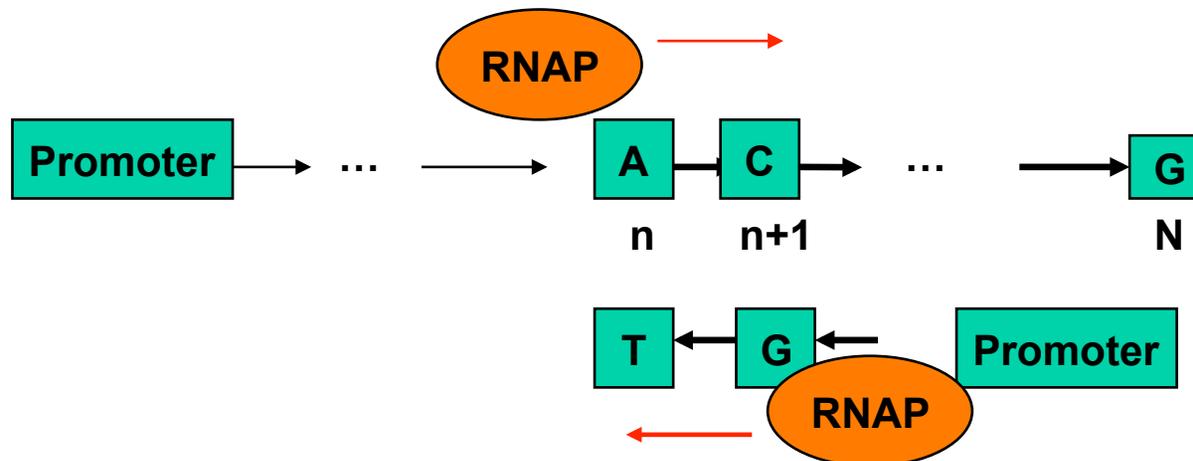
Model Difference

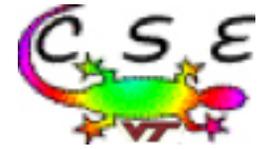
1. From Exponential distribution to Gamma distribution

$$A_0 \rightarrow A_1 \rightarrow \dots \rightarrow A_{N-1} \rightarrow A_N$$

$$t = \sum_{i=1}^N t_i, \quad \text{where } t_i \prec E(a), \text{ then } t \prec \Gamma(a, N) \approx \frac{1}{aN}$$

2. RNAP collision may happen

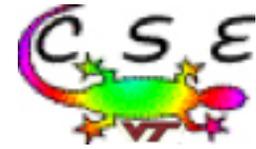




Summary

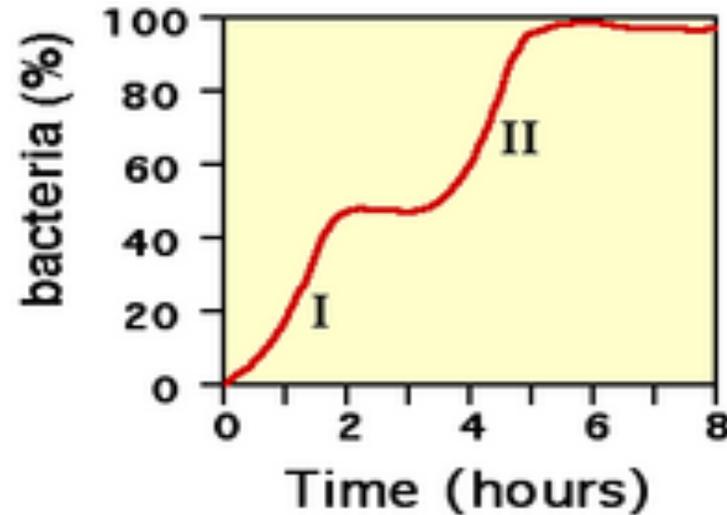
Computational Science and Engineering

- **Gene Expression**
 - Translation
 - Transcription
 - A biochemical Model
- **Gene Regulation**
 - Repressor
 - Activator
 - Feedback Control
 - Models for gene regulation network



Some History: Lac Operon

During World War II, Monod was testing the effects of combinations of sugars as nutrient sources for *E. coli*. He found that bacteria grown with two different sugars often displayed two phases of growth. For example, if glucose and lactose were both provided, glucose would be metabolized first (growth phase I, see Figure 2) and then lactose (growth phase II). But why was there a delay between the two growth phases?

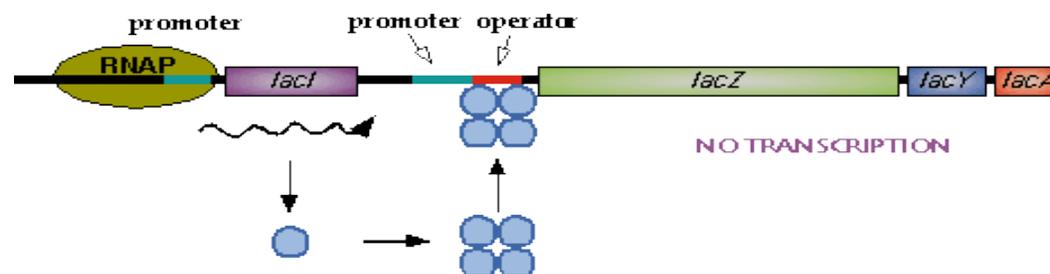
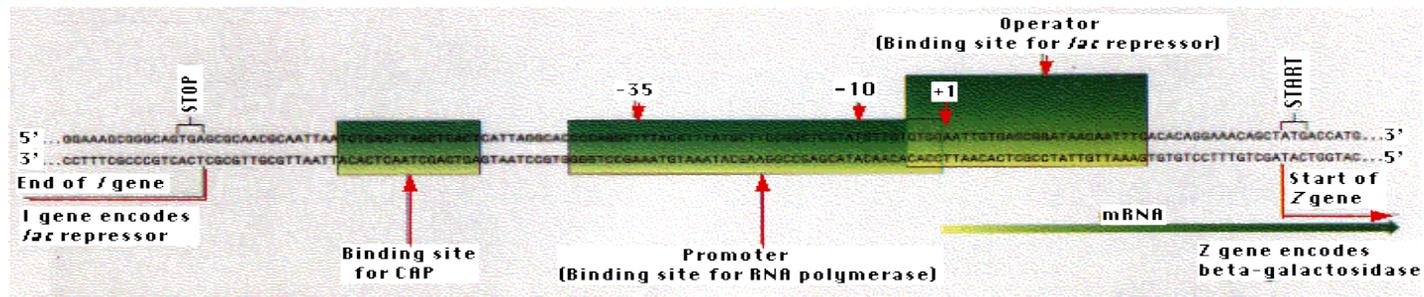
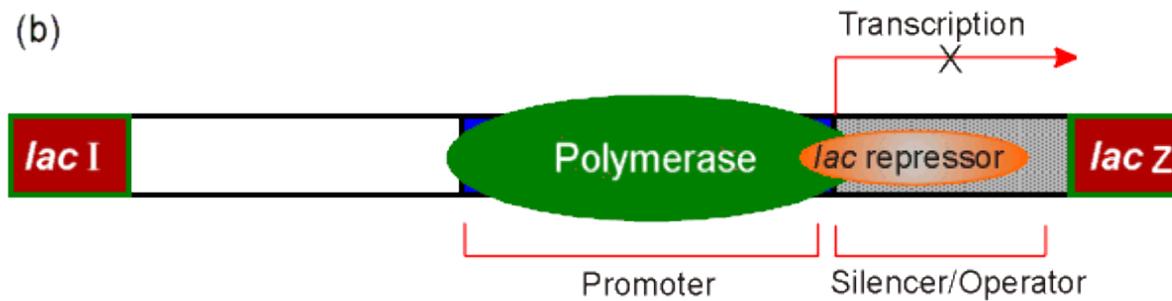


Discovered by Francois Jacob and Jacques Monod, They got Nobel Prize in Physiology or Medicine in 1965



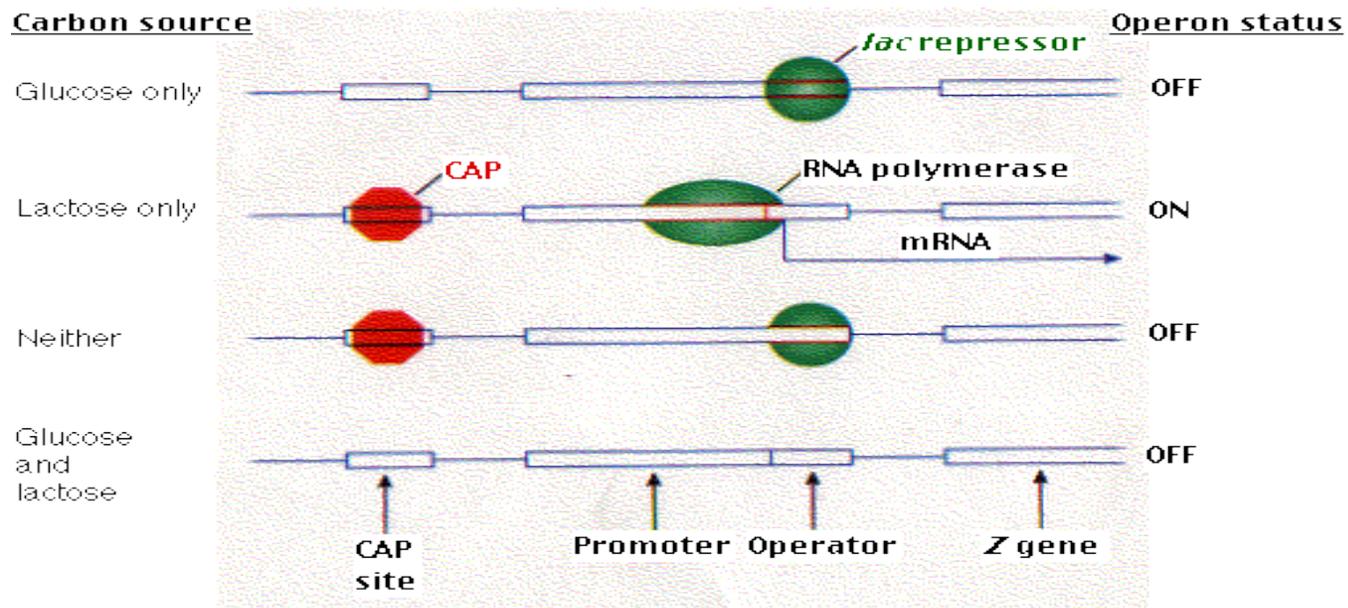
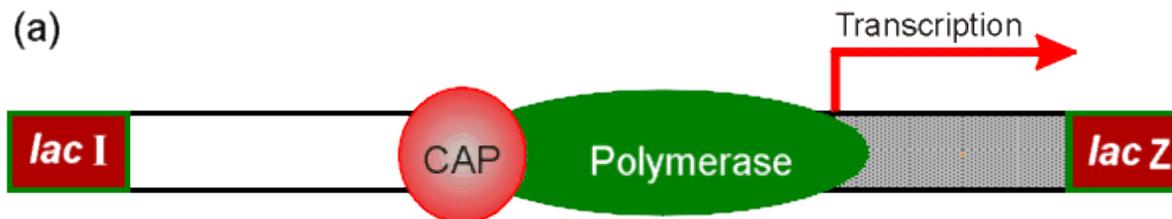
Gene Regulation

- Repressor

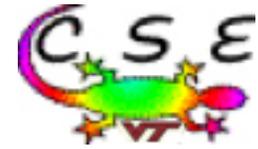


Gene Regulation

- **Activator**



An inducible operon

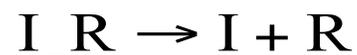
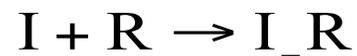
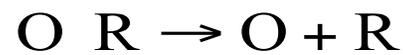
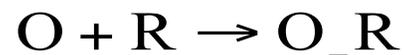


Regulation Reactions in Gene Models

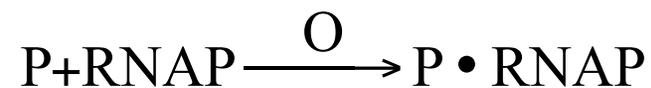
Computational Science and Engineering

Gene expression becomes interesting when regulation system is added into it.

Add the following reaction set



into the initiation of transcription



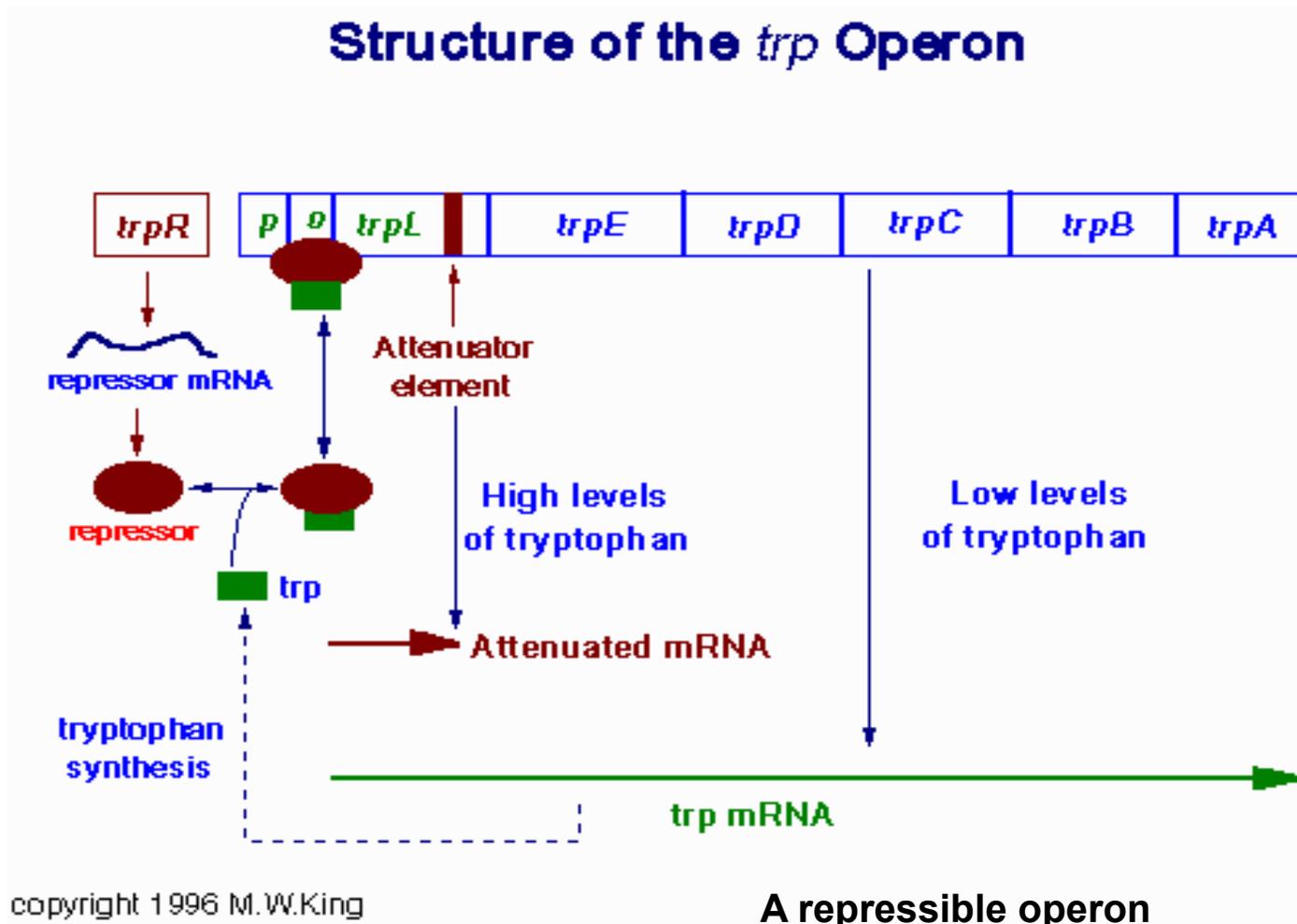
or written as:





Trp Coresspressor: A negative feedback system

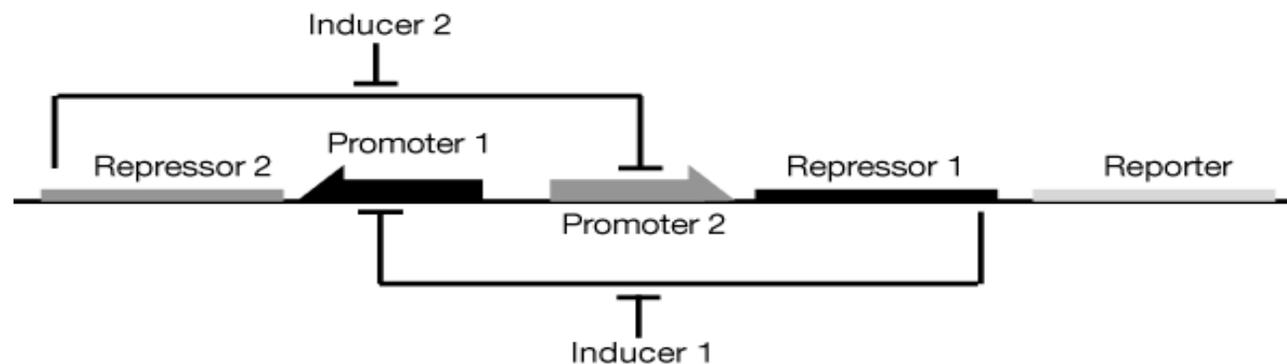
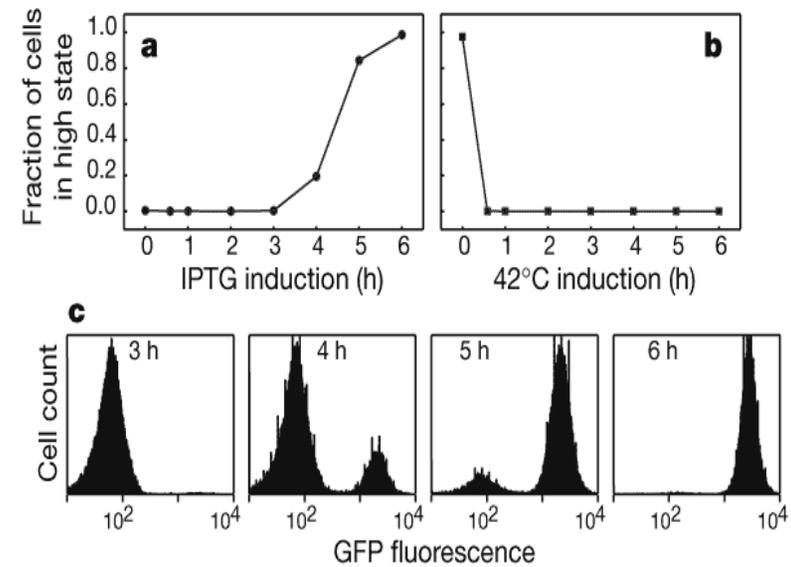
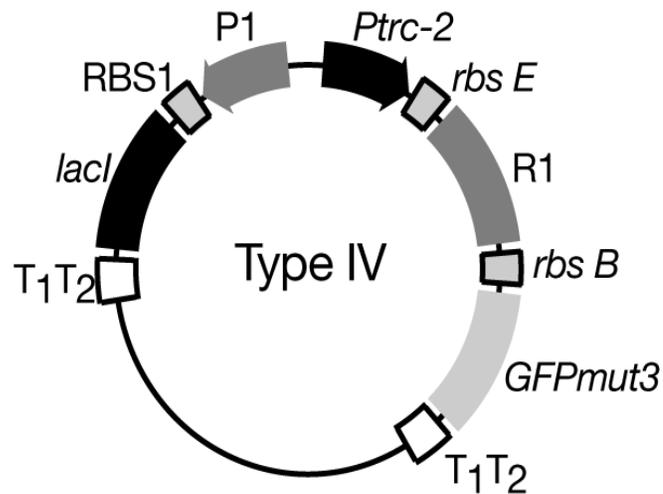
Computational Science and Engineering





Positive Feedback Regulation System

Computational Science and Engineering



Gardner TS, Cantor CR, Collins JJ, Construction of a genetic toggle switch in *Escherichia coli*, NATURE 403 (6767): 339-342 JAN 20 2000



A Simplified Mathematical Model

Computational Science and Engineering

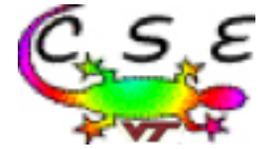
The toggle model

The behaviour of the toggle switch and the conditions for bistability can be understood using the following dimensionless model for the network:

$$\frac{du}{dt} = \frac{\alpha_1}{1 + v^\beta} - u \quad (1a)$$

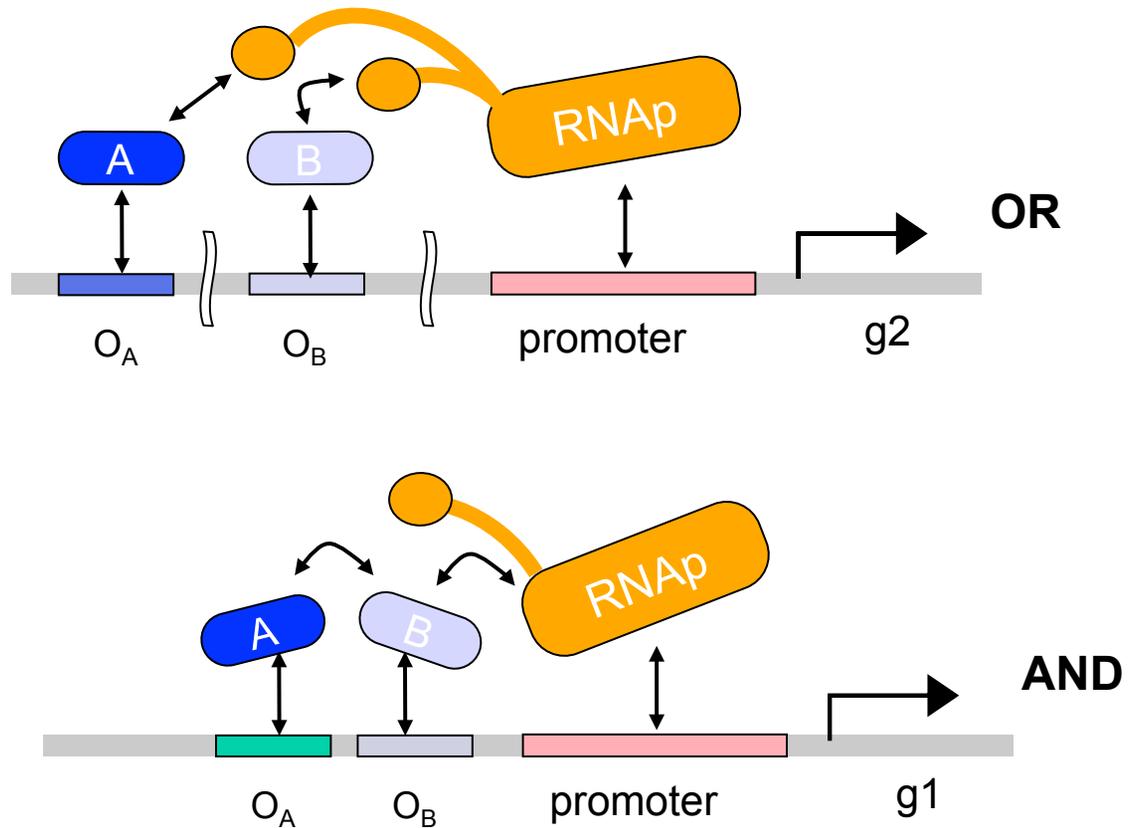
$$\frac{dv}{dt} = \frac{\alpha_2}{1 + u^\gamma} - v \quad (1b)$$

where u is the concentration of repressor 1, v is the concentration of repressor 2, α_1 is the effective rate of synthesis of repressor 1, α_2 is the effective rate of synthesis of repressor 2, β is the cooperativity of repression of promoter 2 and γ is the cooperativity of repression of promoter 1. The above model is derived from a biochemical rate equation formulation of gene expression²⁴⁻²⁷. The final form of the toggle equations preserves the two most fundamental aspects of the network: cooperative repression of constitutively transcribed promoters (the first term in each equation), and degradation/dilution of the repressors (the second term in each equation).



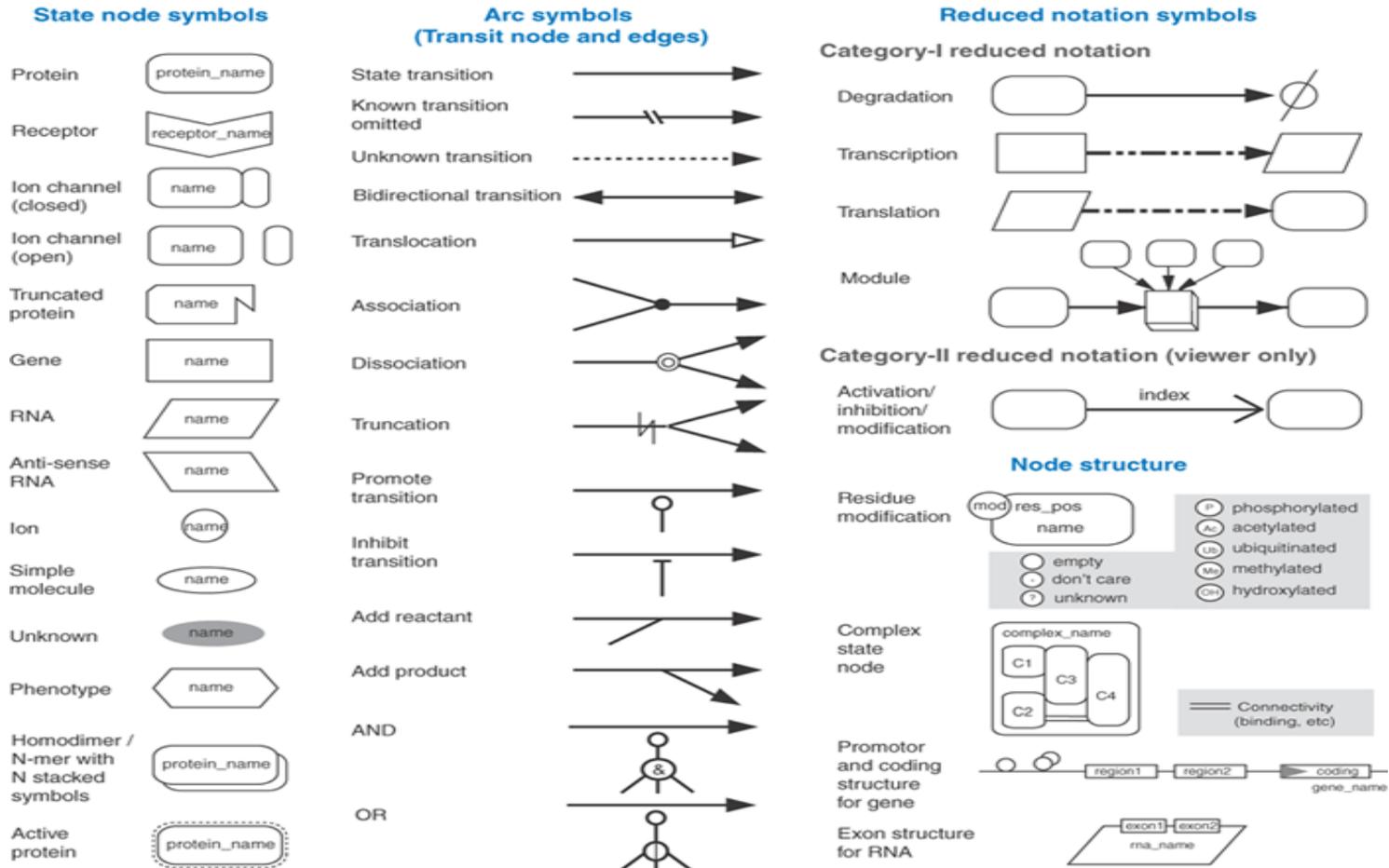
Simple Regulation in Biology – Circuits?

Computational Science and Engineering





Yes! Circuits!



Kitano H, Funahashi A, Matsuoka Y, et al., Using process diagrams for the graphical representation of biological networks, NATURE BIOTECHNOLOGY 23 (8): 961-966 AUG 2005