Regulation of Gene Expression

http://www.mun.ca/biology/desmid/brian/BIOL2060/BIOL2060-23/CB23.html

TENDAI WALTER SANYIKA
CUT, DEPARTMENT OF BIOTECHNOLOGY
BLOCK 11, ROOM 10

waltersanyika@yahoo.com; wsanyika@cut.ac.zw
Objectives

1. To explain the differences between eukaryotic and prokaryotic gene expression systems.

2. To explain the mechanisms of control of gene expression.

3. To explain how the environment and internal requirements “regulate” the expression of genes in cells.

4. To be able to identify stages at which gene expression is regulated and explain the processes involved.
   - For both prokaryotes & eukaryotes.
Gene Expression

Characteristics of gene expression:

• Spatial.
  – Varies with space.
    – Not every gene product needed in every cell type.
    – Eukaryotic genome contains a lot of non-coding DNA sequences.
      » Known as “JUNK DNA”.
      » Actually it is not junk but DNA that is not expressed at a particular time.

• Temporal.
  – Different genes expressed at different times in response to internal or external stimuli.
    • Environmental stimuli varies over time.
    • Hormones expressed at different stages.
      » Especially seen in development- formation of tissues and organs over time or during cellular development.
Gene Expression

Characteristics of gene expression:

- **Prokaryotic** organisms are unicellular.
- **Prokaryotes** regulate gene expression in response to their environment.
- **Eukaryotic** cells regulate gene expression to maintain homeostasis in the organism.
Levels of Gene Expression Regulation

1. Genome
   - Possible gene amplification or deletion (rare)
   - Possible DNA rearrangements (rare)
   - Chromatin decondensation
   - DNA methylation
   - Histone acetylation, changes in HMG proteins, nuclear matrix

2. Transcription
   - Transcription (control by transcription factors)
   - Primary RNA transcript (pre-mRNA)
   - RNA splicing and other processing events

3. RNA processing and nuclear export
   - mRNA in nucleus
   - Transport of mRNA to cytoplasm

4. Translation
   - mRNA degradation (turnover)
   - Translation (control by initiation factors and translational repressors, including microRNAs)
   - Polypeptide product in cytosol or ER

5. Posttranslation
   - Protein folding and assembly
   - Possible polypeptide cleavage
   - Possible modification
   - Possible import into organelles
   - Functional protein
   - Protein degradation (turnover)
Levels of Gene Expression Regulation

1. DNA → RNA transcript (transcriptional control)
2. RNA transcript → mRNA (RNA processing control)
3. mRNA → mRNA (RNA transport control)
4. mRNA → protein (translation control)
5. Protein → inactive protein (protein activity control)

NUCLEUS

CYTOSOL

nuclear pore

nuclear envelope
Induction of transcription

• **Induction** of transcription often occurs in prokaryotes.
  – Also occurs to a lesser extent in eukaryotes.
    – Induction occurs less often in eukaryotes.

• **Induction** can work by a **stimulus**:
  – Temperature (heat).
  – Light.
  – Hormones.
  – Substrates.
  – Etc.
Induction of Transcription

• Temperature.
  – Activate heat shock proteins (HSPs).
    » Transcriptional regulation – stress of high heat signals HSPs to be transcribed.
    » Remember the sigma factor belongs to this family of proteins and it activates transcription of a family of most genes in prokaryotes.

• Light.
  – RBC (ribulose 1,5 bisphosphate carboxylase)
    » Produced when plants are exposed to light.

• Hormones.
  – Signal molecules (eg., secondary messengers).
  – 2 classes of hormones that activate transcription.
    » Steroid hormones.
    » Peptide hormones.
Steroid Hormones

• Small, lipid molecules derived from cholesterol.
  – Easily pass through cell membranes.

• Examples:
  – Estrogen.
  – Progesterone.
  – Testosterone.
  – Glucocorticoids.
Peptide Hormones

- Cannot pass through cell membrane easily.
  - Convey **signals** through **membrane bound receptors**.

- Require signal transduction to communicate the message.
  1. Hormone binds receptor on cell surface.
  2. Signal gets internalized.
  3. Cascade of events begins
Cell signaling cascades activate expression of certain genes.
- Control gene expression, e.g., cAMP in lactose metabolism.
Activation of Gene Transcription by cAMP

Cyclic AMP

Stimulates

Protein kinase A

DNA

CRE

Cyclic AMP response element

CREB

ATP

ADP

CREB

P

CBP

Transcription is activated

Copyright © 2005 Pearson Education, Inc. publishing as Benjamin Cummings
Hormone Response Elements

• Certain DNA sequences are required to bind hormones and initiate transcription.

• Known as Hormone Response Elements (HREs).
  – A response element for hormones.
  – A short sequence of DNA within the promoter of a gene.
  – Able to bind a specific hormone-receptor complex and regulate transcription.

• A gene may have many different response elements.
  – Allows hormone-receptor complex control over the level and rate of transcription.
Comparison of the DNA Sequence of Several Hormone Response Elements

Response element for:

- **Glucocorticoids**
  
  5’ AGAACCAnnnnTGTTTCT 3’
  3’ TCTTGTnmmmACAAAGA 5’

- **Estrogen**
  
  5’ AGGTCAAnnnnTGACCT 3’
  3’ TCCAGTnmmnACTGGGA 5’

- **Thyroid hormone**
  
  5’ AGGTCACTGACCT 3’
  3’ TCCAGTACTGGGA 5’
Hormone receptors form the hormone-receptor complex that binds to DNA at the HRE.

- Note the DNA binding motifs.
Activation of Gene Transcription by Glucocorticoid Receptors

Steroid hormone (cortisol)

Receptor-Hsp complex

CYTOPLASM

Hsp proteins

NUCLEUS

DNA

Transcription activated

RNA

Glucocorticoid response element
Prokaryotic Regulation

- Control of transcription initiation can be:
  - **Positive control** – Increases transcription when activators bind DNA.
  - **Negative control** – Reduces transcription when repressors bind to DNA regulatory regions called operators.

- Prokaryotic cells often respond to their environment by changes in gene expression.
  - Genes involved in the same metabolic pathway are organized in **operons**.
    - Some operons are **induced** when the metabolic pathway is needed.
    - Some operons are **repressed** when the metabolic pathway is no longer needed.
Prokaryotic Regulation: *lac* Operon
Typical Example For Regulation of Energy

- Contains genes for the **metabolism of lactose as an energy source**.
- Regulatory regions of the operon:
  - Promoter.
  - Operator.
  - CAP binding site.
- The coding region contains genes for 3 enzymes:
  - β-galactosidase.
    - Converts the substrate, lactose into glucose and galactose.
  - Permease.
    - Transports lactose into the cell.
  - Transacetylase.
    - Function unknown (acetylation?).
Prokaryotic Regulation: *lac* Operon

- Gene for repressor protein
- Promoter for *l* gene
- CAP-binding site
- Operator
- Promoter for *lac* operon
- Gene for β-galactosidase
- Gene for permease
- Gene for transacetylase

**Regulatory region**

**Coding region**

*lac* Control system
Figure 3.10  Alternate forms of the repressor, in the absence or presence of inducer, inhibit or permit, respectively, transcription of the β-galactosidase gene.
Figure 3.11  Transcription of the $\beta$-galactosidase gene requires an activator.
(a) Lactose absent, repressor active, operon off

(b) Lactose present, repressor inactive, operon on
The Lac Operon: CAP Site

• The CAP site
  – DNA Binding Site for a complex between the Catabolite Activator Protein (CAP) and cyclic AMP (cAMP).
    – Forms a CAP-cAMP complex.
  – The binding of the CAP-cAMP complex to the promoter site is required for transcription of the lac operon.
  – The presence of this complex is closely associated with the presence and levels of glucose in the cell.
  – But as the concentration of glucose increases the amount of cAMP decreases (inverse relationship).
  – As the cAMP decreases, the amount of CAP-cAMP complex decreases.
    – Decrease in the complex inactivates the promoter, and the lac operon is turned off.
      » Because the CAP-cAMP complex is needed for transcription, the complex exerts a positive control over the expression of the lac operon.
Prokaryotic Regulation: lac Operon

- The lac operon is negatively regulated by a repressor protein (always under repression):
  - lac repressor binds to the operator to block transcription.
  - In the presence of lactose, an inducer molecule that binds to the repressor protein.
    - Repressor will no longer bind to operator.
      » Promoter available for RNA polymerase binding.
      » Transcription proceeds.

- In the presence of both glucose and lactose.
  - Bacterial cells prefer to use glucose.
    - Phenomenon known as Catabolite Repression.

- Glucose prevents induction of the lac operon.
  - Process regulated by cyclic AMP (cAMP).
  - Binding of CAP-cAMP complex to the CAP binding site is required for induction of the lac operon.
    » Low glucose levels $\rightarrow$ high cAMP $\rightarrow$ cAMP binds CAP protein $\rightarrow$ cAMP-CAP complex binds CAP site on DNA.
    » High glucose levels cause low cAMP levels $\rightarrow$ low cAMP $\rightarrow$ no induction of lac operon because glucose is present.
  - Thus use of other sugars is an alternative required only when glucose is not available. Bacteria preferentially selects amongst different sugars for metabolism.
The Tryptophan (Trp) Operon
Trp Operon

- *E. coli* uses several proteins (enzymes) encoded by a cluster of **5 genes** to manufacture/synthesize the amino acid **tryptophan**
- All 5 genes are transcribed together as a unit called an **operon**.
  - Which produces a single long piece of mRNA for all the genes
- RNA polymerase binds to a promoter located at the beginning of the first gene and proceeds down the DNA transcribing the genes in sequence.
- The **trp operon**
  - Encodes genes for the biosynthesis of tryptophan when necessary.
  - The operon is not expressed when the cell contains sufficient amounts of tryptophan.
  - The operon is expressed when levels of tryptophan are low.
Prokaryotic Regulation: *Trp* Operon

Typical Operon For Amino Acid Metabolism

- The *trp* operon is **negatively regulated** by the *trp* repressor protein.
  - *trp* repressor binds to the operator to block transcription.
  - Binding of repressor to the operator **requires** a **corepressor** which is **tryptophan**.
  - Low levels of tryptophan prevent the repressor from binding to the operator and vice versa.
- Regulation uses two mechanisms:
  - Negative feedback by tryptophan (as above).
  - Attenuation, which relies on the ribosome.
Figure 11.1C  Two types of repressor-controlled operons
Fig. 16.6

[Diagram of the trp Operon Control]

- **Inactive repressor**
- **RNA polymerase**
- **mRNA synthesis**
- **Genes are ON** → Tryptophan is synthesized
- **Genes are OFF** → Tryptophan is not synthesized

**trp Operon Control**

- **Promoter**
- **Operator**
- **Tryptophan present**
- **Tryptophan absent**

[Diagram showing the interaction of tryptophan with the repressor and RNA polymerase, leading to mRNA synthesis when tryptophan is absent and repression when tryptophan is present.]
Trp operon

regulatory regions

structural gene

Trp mRNA

attenuated mRNA

Trp mRNA
Attenuation Mechanism

- In addition to the repressor regulation, trp gene has a fine tuning mechanism called **attenuation**.
- Remember transcription & translation are coupled.
- The trp operon is transcriptionally **regulated** using attenuation mechanism at the **translation level** (involves both the ribosome and RNA polymerase).
  - Termination is controlled via changes in RNA secondary structure that are determined by ribosome movement.
    - Termination occurs when the ribosome translates a leader segment of the mRNA to form a termination hairpin loop at the attenuator.
    - When the ribosome translates the leader region, a termination hairpin forms at terminator.
    - When the ribosome is prevented from translating the leader, the termination hairpin does not form, and RNA polymerase transcribes the coding region.
The *Trp* Operon: Leader Sequence

- Transcription termination at the attenuator responds to the levels of tryptophan.
- Leader sequence contains the following:
  - Ribosome binding site, AUG (stop) codon:
    - Followed by a short coding region that contains **two successive tryptophan codons** in Domain 1/Region 1.
  - When the cell runs out of tryptophan:
    - Ribosomes initiate translation of the leader peptide, but process **stops** when it reaches the **2 Trp codons**.
      » **Ribosome stalling** influences termination of leader peptide synthesis at the attenuator but enables the transcription of structural genes.
- The **leader sequence** can form **alternative base-paired structures**.
  » Affects the **ability of the ribosome to proceed through leader region** (results in **attenuation** or **transcription** of structural genes).
Attenuation By Leader sequence

mRNA  pppAAGUUCACGUA........AGGUACGACAAUGAAAGCAAUUUCGUACUGCA

Leader peptide
Met-Lys-Ala-Ile-Phe-Val-Leu-

(stop)-Ser-Thr-Arg-Trp-Trp-

UACCCAGCCCGCCUUAAUGAGCGGGCUU....UUGAACA...UAACAAUGCGAAACA

Site of transcription attenuation

139

End of leader region (trpL)

162

TrpE polypeptide

4
Leader Peptide
3 Possible Hairpins

1/2 and 3/4 hairpin structures:
Terminators

2/3 hairpin Structure:
Anti-terminator
Attenuation

• *Trp* operon contains the **leader sequence (L)**.
  – Occurs immediately upstream (5′) of the *trpE* gene.
  – This sequence about 160 bp in size.
• Controls the expression of the operon through attenuation.
• Leader Sequence has **four domains (1-4)**.
  – Domain 3 (nucleotides 108-121) of the mRNA can base pair with either domain 2 (nucleotides 74-94) or domain 4 (nucleotides 126-134).
  – If domain 3 pairs with domain 4, a stem and loop structure forms on the mRNA and transcription stops.
    – Domain 4 is the attenuator.
  – This structure forms when the level of tryptophan is high in the cell.
Attenuation: The attenuator

• An attenuator.
  – Controls the progression of RNA polymerase into the *trp* structural genes.

• RNA polymerase.
  – Initiates at the promoter.
  – Proceeds to position 90.
    – Where it *pauses* before proceeding to the attenuator at position 140.
  – In the absence of tryptophan.
    – The polymerase continues into the structural genes (*trpE* starts at +163).
  – In the presence of tryptophan.
    – There is ~90% probability of termination to release the 140-base leader RNA without transcribing the structural proteins.
Attenuation

• If domain 3 pairs with domain 2.
  – The stem and loop structure does not form.
  – Transcription continues throughout the operon.
  – All of the enzymes required for tryptophan biosynthesis are produced.
  – These events occur when tryptophan is low in the cell.

• Domain 1
  – Also important for attenuation.
    – Section of the leader sequence encodes a 14 amino acid peptide that has two tryptophan residues.
      » The 2 trp residues can also act as terminators of the leader peptide sequence translation by the ribosome.
    – Function depends on conditions of high and low tryptophan.
Possible Hairpins

1/2 and 3/4 hairpin structure

2/3 hairpin structure
Attenuation
Leader Translational Speed Determines Loop Outcome

- Translation termination at the attenuator responds to the level of tryptophan.
  - Relies on the movement of the ribosome.
- In the presence of adequate amounts of tryptophan.
  - Termination of translation is efficient.
    - Leader sequence completely translated by ribosome.
      » Secondary structure formed terminates transcription of structural genes by polymerase.
- In the absence of tryptophan or when levels are low.
  - RNA polymerase can continue into the structural genes.
    - Ribosomal stalling (pausing) occurs at the 2 trp codons as the ribosome will have high affinity for them.
      » Stalling allows formation of a 2° structure that favours polymerase read-through to transcribe structural genes & tryptophan synthesis.
At High $Trp$ Concentration
At Low $Trp$ concentration
Attenuation
Leader Translational Speed Determines Loop Outcome

Attenuation of the \textit{trp} operon mRNA

- Domain 1
- Domain 2
- Domain 3
- Domain 4

\textit{trp} leader region

Low Tryptophan Levels
- Slow translation of Domain 1 peptide
- Domain 2-3 pairing occurs
- Normal full gene transcription

High Tryptophan Levels
- Fast translation of domain 1 peptide
- Domain 2 blocked by ribosome
- Domain 3-4 pairing occurs
- Attenuation of transcription occurs
- Only 10\% of normal mRNAs made
Attenuation

- The **region 3+4 hairpin** structure acts as a transcription termination signal (attenuator).
  - As soon as it forms, the RNA and the RNA polymerase are released from the DNA.

- **During periods of tryptophan scarcity.**
  - A ribosome translating the coding sequence for the leader peptide may stall when it encounters the two tryptophan (trp) codons because of the shortage of tryptophan-carrying tRNA molecules (it would somehow prefer those codons).

- **Stalled ribosome at this site blocks region 1.**
  - A **region 1+2** hairpin cannot form.
  - An alternative, **region 2+3** hairpin forms instead.
    - This prevents formation of the **region 3+4** transcription termination hairpin and therefore RNA polymerase can move on to transcribe the entire operon to produce enzymes that will synthesize tryptophan.
(a) The most stable secondary structure for trp leader mRNA. Attenuation depends on the ability of regions 1 and 2 and regions 3 and 4 of the trp leader sequence to base-pair, forming hairpin secondary structures. The 3–4 hairpin structure acts as a transcription termination signal.

(b) When tryptophan is scarce the ribosome stalls, allowing a 2–3 “antiterminator” hairpin to form. The ribosome stalls when it encounters the two tryptophan (Trp) codons due to a shortage of tryptophan-carrying tRNA molecules. The stalled ribosome blocks region 1, so a 1–2 hairpin cannot form. Instead an alternative 2–3 hairpin is created, which prevents formation of the 3–4 termination hairpin. Therefore RNA polymerase can move on to transcribe the entire operon.

(c) When tryptophan is plentiful the ribosome continues, allowing the 3–4 transcription termination signal to form. The moving ribosome completes translation of the leader peptide and pauses at the stop codon, blocking region 2. As a result, the 3–4 structure forms and terminates transcription near the end of the leader sequence.
**Attenuation**

- When tryptophan is readily available.
  - Ribosome can complete translation of the leader peptide without stalling.
  - As it pauses at the stop codon, it blocks region 2, preventing it from base pairing.
  - As a result, the region 3+4 structure forms and terminates transcription near the end of the leader sequence.
    - Structural genes of the operon are not transcribed (nor translated).

- The signal sequence is example of a “riboswitch” mechanism.
  - Mechanism which can control transcription and translation through interactions of molecules with the mRNA.
  - Riboswitch sequences (in the mRNA leader transcript) bind molecules such as amino acids, nucleotides, sugars, vitamins, metal ions and other small ligands.
In Eukaryotes
Opportunities For Controlling Gene Expression in Eukaryotic cells

- Enhancers (distal control elements)
- Proximal control elements
- Promoter
- Transcription
- RNA processing: Cap and tail added; introns excised and exons spliced together
- Translation

DNA: Enhancers (distal control elements) upstream, Proximal control elements, Promoter, Exon, Intron, Exon, Intron, Exon, Terminator, Downstream

Primary transcript (pre-mRNA) with regions labeled 5' and 3', showing processes of transcription and RNA processing.

mRNA: Coding segment with 5' cap, leader, start codon, trailer, poly(A) tail.

Ribosome: GTP-Pi-Pi-Pi, Polypeptide: 5' Cap, Leader, Start codon, Stop codon, Poly(A) tail.
Gene Expression & Chromosomes

- DNA needs to be accessible to RNA polymerase for transcription initiation
  - Gene expression affected by position on chromosome.

- Gene expression influenced by chromosomal structure
Chemical Modification of Histone Proteins control Eukaryotic Gene Expression

- Eukaryotic DNA is packaged into chromatin.
  - Chromatin structure is directly related to the control of gene expression.
- Chromatin structure begins with the organization of the DNA into nucleosomes.
- Nucleosomes may block RNA polymerase II from gaining access to promoters.
Chemical Modification of Histone Proteins control Eukaryotic Gene expression

- **Methylation** (the addition of –CH$_3$) of DNA or histone proteins is associated with the control of gene expression.

- Clusters of methylated cytosine nucleotides bind to a protein that prevents activators from binding to DNA.

- Methylated histone proteins are associated with **inactive regions** of chromatin.
  - Methylation shuts off gene expression.

- **Acetylation** (addition of acetyl group) causes unwinding of histone proteins.
  - Allows region of DNA to be transcribed.
  - Allows gene expression.
• Histone modifications
  – Methylation
  – Acetylation
  – Etc
Nucleosomes block the binding of RNA polymerase II to the promoter.

Condensed solenoid

Amino acid histone tail

N-terminus

Addition of acetyl groups to histone tails remodel the solenoid so that DNA is accessible for transcription.

Acetyl group

DNA available for transcription
Eukaryotic Gene Structure & Regulation of Expression

• Eukaryotic RNA processing
  – 5' cap - stabilizes mRNA
  – Untranslated regions (5' & 3' UTR) – stabilize mRNA from degradation.
  – 3' Poly A tail - for localization & transportation of mRNA.

• Prokaryotic cells more regulated than eukaryotic cells.

• mRNA stability - Half-life:
  – Long vs. short lived mRNAs
  – Long- many rounds protein synthesis from one mRNA
  – Short – rapidly degraded, needs more transcription to replenish (half-life)

• Prokaryotic mRNA is not modified.
  – Transcription and translation are coupled.
    » mRNA degraded as it is transcribed.
    » Eukaryotic mRNA persists and translated separately.
Also see Role of Epigenetic Mechanisms in Regulation of Gene Expression
Epigenetic Mechanisms of Regulating Gene Expression

• Position effect variegation.
• Paramutation.
• Genetic imprinting.
• Histone modifications and the histone code.
  – Acetylation, phosphorylation, methylation.
• DNA methylation.
• sRNA/ RNA interference.
• Ubiquitination.
• SUMOylation.
• Chromosome inactivation.
• Etc.
Summary

• Gene expression is in response to cell necessity.
  – Regulation requirements and mechanisms differ between prokaryotes and eukaryotes.

• The environment transmits signals that regulate the expression of certain genes.
  – How internal cell mechanisms catalyze and regulate the processes.
    » Transcription regulation is the major level at which expression is regulated by controlling the initiation (strength & rate) and therefore, levels of mRNA produced.
  – How gene expression is regulated at:
    » mRNA level (transcription)
    » mRNA level (post transcription)
    » Protein level (degradation).
END