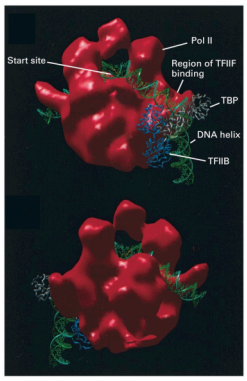
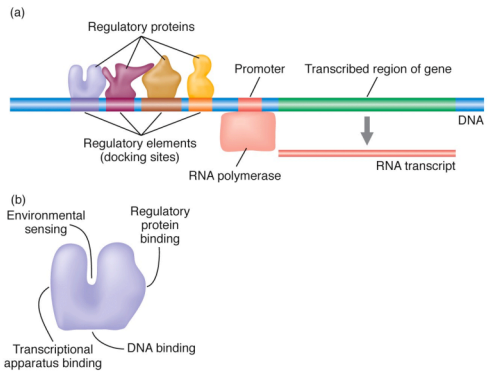
Chapter 16

Regulation of Gene Expression



Regulatory proteins

- Themselves product of transcription and translation
- Several regulatory domains
  - sequence-specific DNA-binding surface to recognize DNA docking (target) site
  - surface to recognize basal transcription apparatus
  - surface to interact with proteins binding to nearby docking sites (may or may not be present)
  - surface to act as sensor of environmental conditions (may or may not be present)
- Many variations and interactions



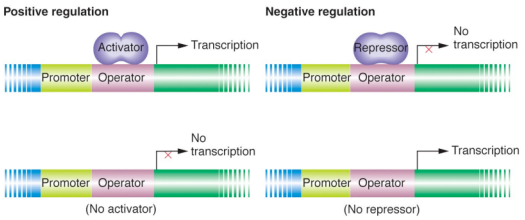
Prokaryotic transcriptional regulation

- Promoter
  - determines where transcription begins, including identification of template strand
  - binds RNA polymerase
- Operator
  - adjacent to or overlapping promoter
  - binds activator or repressor proteins
- In *positive regulation*, binding of activator to docking site is required for transcription
- In *negative regulation*, operator must be free of bound protein for transcription to occur

Table 16.2 Features of inducible and repressible operons with positive and negative control

Type of Control	Transcription Normally	Regulator Protein	Effect of Regulatory Protein	Action of Modulator
Negative inducible	Off	Active repressor	Inhibits transcription	Substrate makes repressor inactive
Negative repressible	On	Inactive repressor	Inhibits transcription	Product makes repressor active
Positive inducible	Off	Inactive activator	Stimulates transcription	Substrate makes activator active
Positive repressible	On	Active activator	Stimulates transcription	Product makes activator inactive

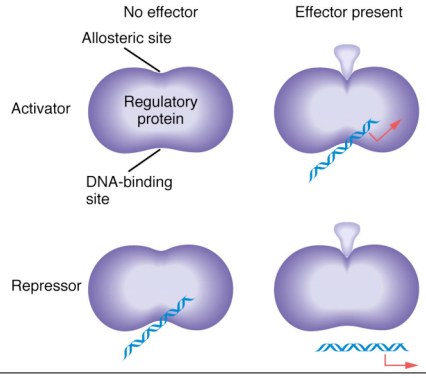
Positive and negative regulation



Activator and repressor proteins

- Usually have at least two sites
  - DNA-binding domain
  - allosteric site
    - binds effector molecule
    - toggles DNA-binding domain between functional and nonfunctional
    - activator usually has RNA polymerase-binding domain
- In some cases, effector must be bound to allosteric site for protein to bind to DNA
- In other cases, allosteric site must be empty for protein to bind to DNA

Allosteric effectors



The *lac* operon

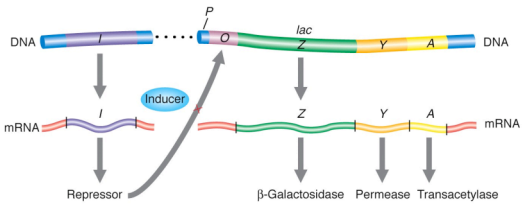
- A model of negative and positive regulation in *E. coli*
- Nobel prize for Jacob and Monod
- Regulates utilization of lactose sugar
  - lactose must be present
  - glucose (preferred sugar) must be absent
    - repression of *lac* operon by glucose is example of catabolite repression (an example of positive control)
  - example of negative regulation

The *lac* operon components

- *lac* structural genes
  - *Z*: β-galactosidase
  - *Y*: permease
  - *A*: transacetylase
  - transcribed as single, multigenic mRNA
- *lac* regulatory components
  - *I*: encodes lac repressor
  - *P*: promoter site for RNA polymerase binding
  - *O*: operator site to which Lac repressor binds

P, *O*, *Z*, *Y*, and *A* components constitute *lac* operon. Sometimes called *lacO*, *lacZ*, etc.

Lac operon



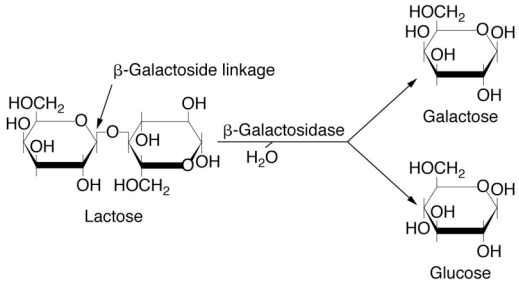
Effect of lactose

- Lac repressor has two recognition sites
  - exists as tetramer of identical subunits
  - single DNA-binding site to dock with *O*
    - binds to *O* site with high affinity
    - only one *O* site in genome
    - prevents transcription when bound to *O*
  - four allosteric sites that bind lactose and its analogs
  - in absence of lactose or analog, repressor is tightly bound to *O*, repressing transcription of *Z, Y,* and *A*
- Binding of lactose or analog to allosteric sites reduces affinity of repressor for *O*, resulting in its release from *O* and permitting transcription

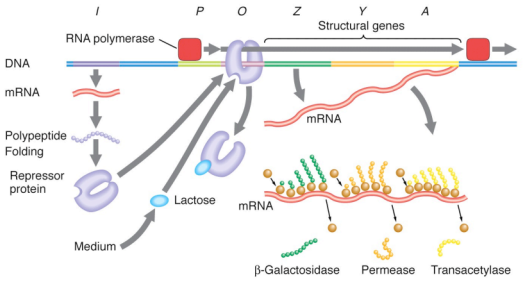
Induction

- *lac* operon is an example of an inducible system
- Induction
  - relief of repression by action of an inducer
  - inducer binds allosterically to repressor, altering its conformation and binding capacity
- During repression there is no transcription of structural genes *Z, Y,* and *A*
- After induction, these genes are transcribed

ß-galactosidase



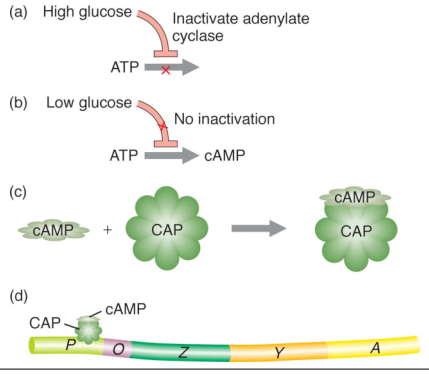
Regulation of the lac operon



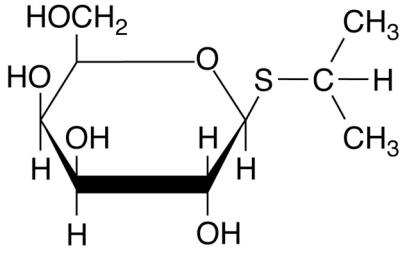
Catabolite repression

- Catabolite repression is superimposed on the *lac* and related operons
  - glucose is preferred carbon source
  - induction occurs only in absence of glucose
- Two components
  - cAMP (cyclic AMP)
    - when glucose level is low, cAMP concentration rises because AMP is not converted to ATP
  - catabolite activator protein (CAP)
    - CAP is activated allosterically by cAMP
    - active CAP binds to CAP site of *lac* operon, facilitating binding of RNA polymerase

Catabolite control



Non-substrate inducer



Isopropyl-β-D-thiogalactoside (IPTG)

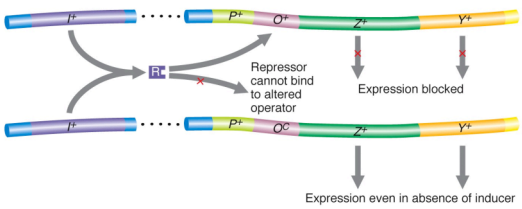
Genetic analysis (1)

- The *lac* operon was discovered through a combination of biochemical and genetic analyses
- Mutations in repressor and operator cause global misregulation of *lac* operon.
  - O^C mutations result in *constitutive* expression because repressor can not bind
    - said to be cis-acting (adjacent sequence)
  - I^- mutations result in *constitutive* expression because no functional repressor is produced
    - said to be trans-acting (not adjacent)

Genetic analysis (2)

- I^S mutations (super-repressors) result in permanent repression because inducer can not bind to repressor.
  - genetic evidence for allostery
- *P* region
  - between *I* and *O*
  - region for binding of RNA polymerase
- Repressor and CAP-cAMP binding sites
  - differ in nucleotide sequence
  - both show rotational twofold symmetry (partial palindrome)

Partial diploids



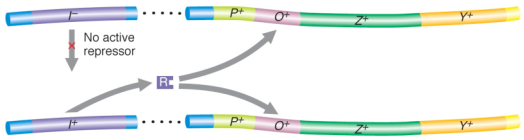
Mutations in the lac operon

TABLE 13-1 Synthesis of β-Galactosidase and Permease in Haploid and Heterozygous Diploid Operator Mutants

Strain	Genotype	β-GALACTOSIDASE (Z) Noninduced	β-GALACTOSIDASE (Z) Induced	PERMEASE (Y) Noninduced	PERMEASE (Y) Induced	Conclusions
1	$O^+Z^+Y^+$	−	+	−	+	Wild type is inducible
2	$O^+Z^+Y^+/F'\ O^+Z^-\ Y^+$	−	+	−	+	Z^+ is dominant to Z^-
3	$O^CZ^+Y^+$	+	+	+	+	O^C is constitutive
4	$O^+Z^-\ Y^+/F'\ O^CZ^+Y^-$	+	+	−	+	Operator is cis-acting

Note: Bacteria were grown in glycerol (no glucose present) with and without the inducer IPTG. The presence or absence of enzyme is indicated by + or −, respectively. All strains are I^+.

Partial diploid - I^-



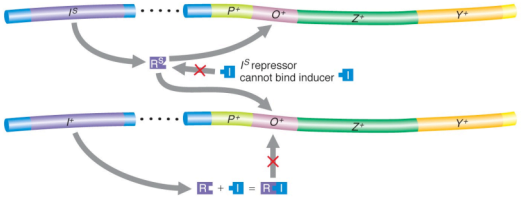
Mutations cont.

TABLE 13-2 Synthesis of β-Galactosidase and Permease in Haploid and Heterozygous Diploid Strains Carrying I^+ and I^-

Strain	Genotype	β-GALACTOSIDASE (Z) Noninduced	β-GALACTOSIDASE (Z) Induced	PERMEASE (Y) Noninduced	PERMEASE (Y) Induced	Conclusions
1	$I^+Z^+Y^+$	−	+	−	+	I^+ is inducible
2	$I^-Z^+Y^+$	+	+	+	+	I^- is constitutive
3	$I^+Z^-Y^+/F'\ I^-Z^+Y^+$	−	+	−	+	I^+ is dominant to I^-
4	$I^-Z^-Y^+/F'\ I^-Z^+Y^-$	−	+	−	+	I^+ is trans-acting

Note: Bacteria were grown in glycerol (no glucose present) and induced with IPTG. The presence of the maximal level of the enzyme is indicated by a plus sign; the absence or very low level of an enzyme is indicated by a minus sign. (All strains are O^+.)

I super repressor



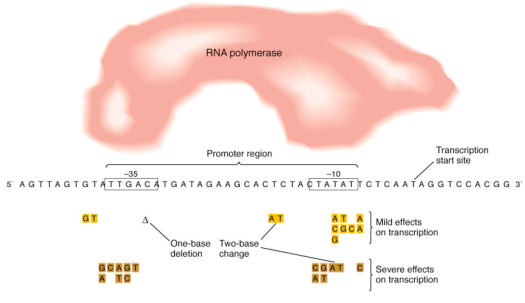
I super repressor

TABLE 13-3 Synthesis of β-Galactosidase and Permease by the Wild Type and by Strains Carrying Different Alleles of the *I* Gene

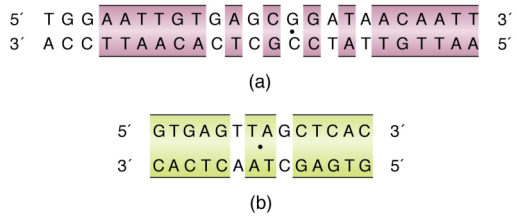
Strain	Genotype	β-GALACTOSIDASE (Z) Noninduced	β-GALACTOSIDASE (Z) Induced	PERMEASE (Y) Noninduced	PERMEASE (Y) Induced	Conclusions
1	$I^+Z^+Y^+$	−	+	−	+	I^+ is inducible
2	$I^S\ Z^+Y^+$	−	−	−	−	I^S is always repressed
3	$I^S\ Z^+Y^+/F'I^+Z^+Y^+$	−	−	−	−	I^S is dominant to I^+

Note: Bacteria were grown in glycerol (no glucose present) with and without the inducer IPTG. Presence of the indicated enzyme is represented by +; absence or low levels, by −.

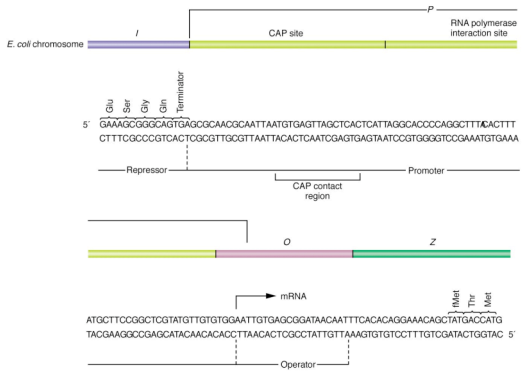
Promoter site



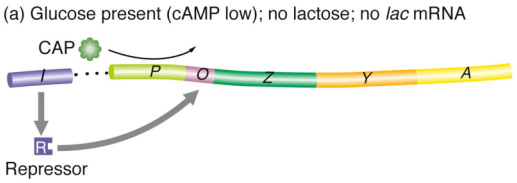
Lac operator and CAP



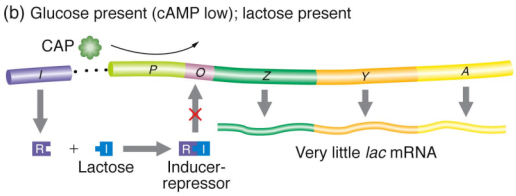
Upstream of the lac operon



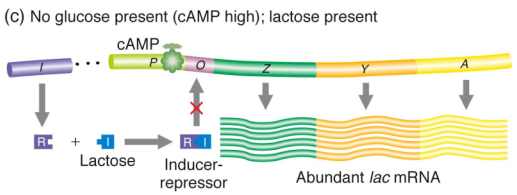
Glucose present-no lactose



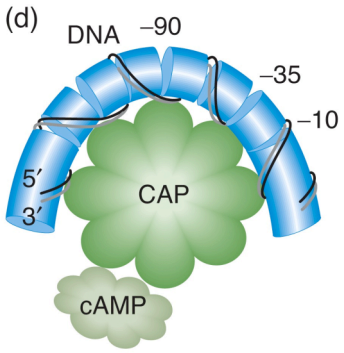
Glucose present-lactose present



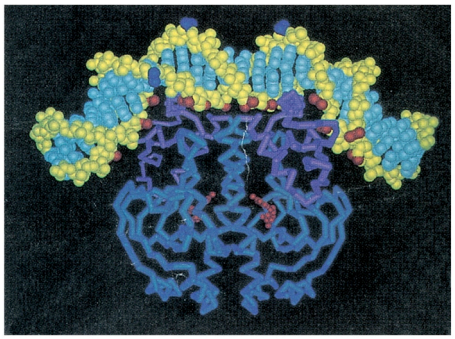
No glucose-lactose present



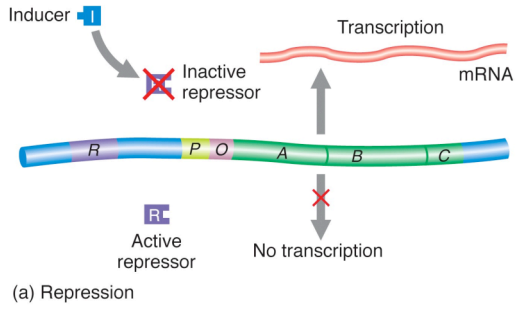
Cap binding



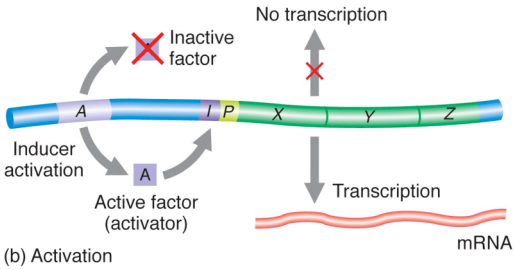
CAP binding



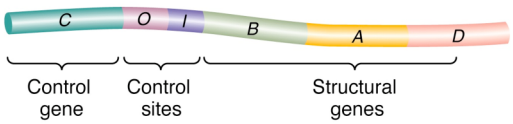
Repression



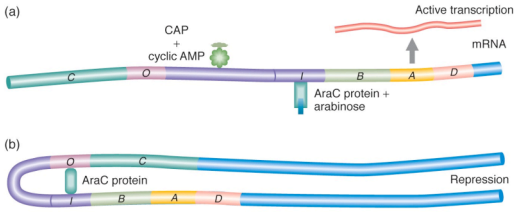
Activation



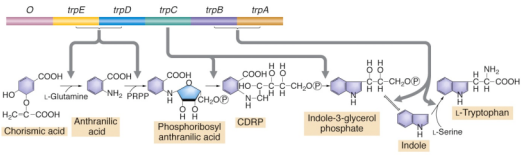
Arabinose operon

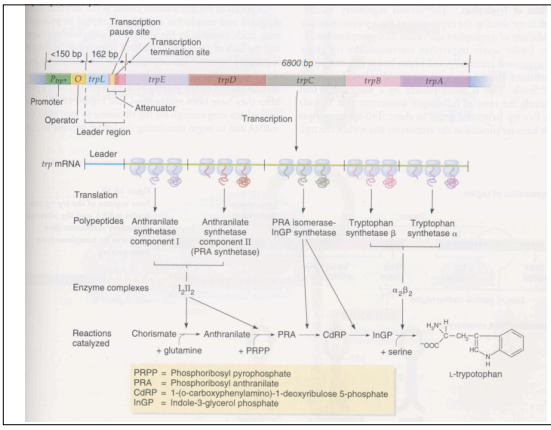


Control of the Arabinose operon

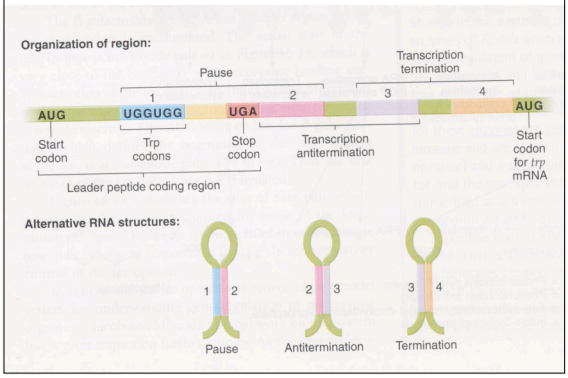


Tryptophan operon

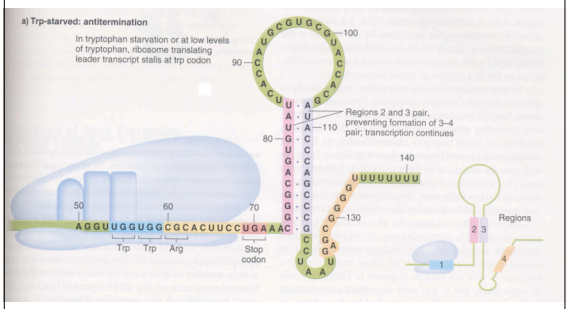




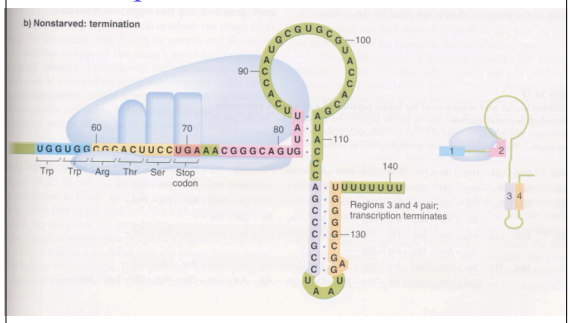
Trp organization

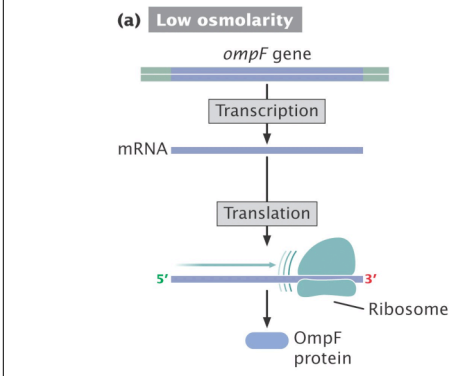


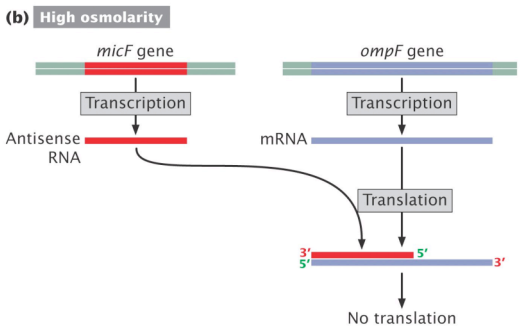
No trp - antitermination



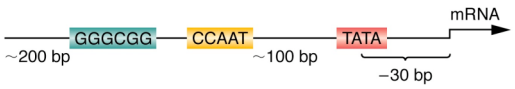
Trp available - termination



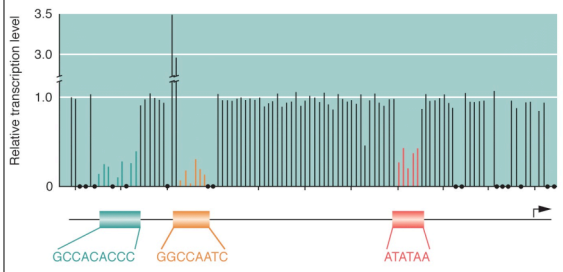




Eukaryotic upstream sites



Point mutations and transcription



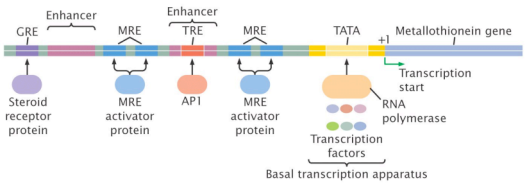
Eukaryotic transcriptional regulation

- No operon structure. Coordinately regulated genes usually are dispersed.
- No (or few) instances of multigenic mRNA
- Critical importance in cell cycle and in development
- Cis-acting sequences
  - core promoter and promoter-proximal elements
  - enhancers
  - silencers
- trans-acting control
  - numerous transcription factors encoded by linked and unlinked genes

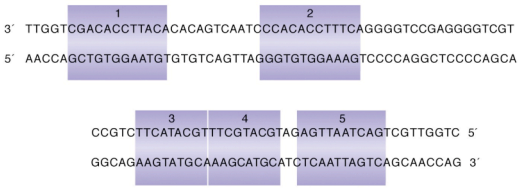
Table 16.4 A few response elements found in eukaryotic cells

Response Element	Responds to	Consensus Sequence
Heat-shock element	Heat and other stress	CNNGAANNTCCNNG
Glucocorticoid response element	Glucocorticoids	TGGTACAAATGTTCT
Phorbol ester response element	Phorbal esters	TGACTCA
Serum response element	Serum	CCATATTAGG

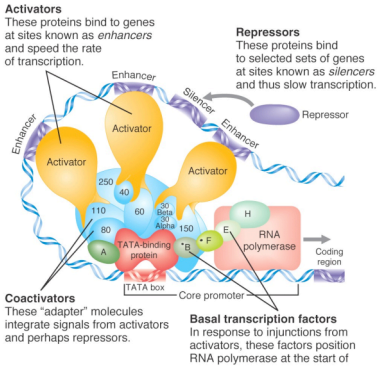
Source: Adapted from B. Lewin, *Genes IV* (Oxford: Oxford University Press, 1994), p. 880.



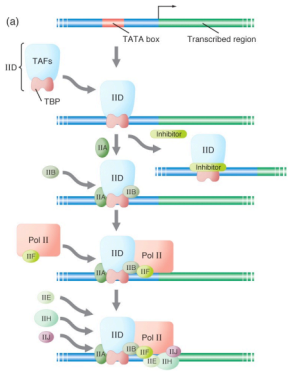
SV40 enhancer - cis acting



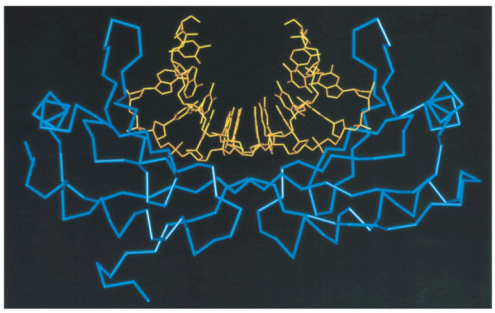
Eukaryotic transcription complex



RNA pol II initiation complex



TATA binding protein

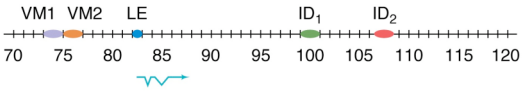


(b)

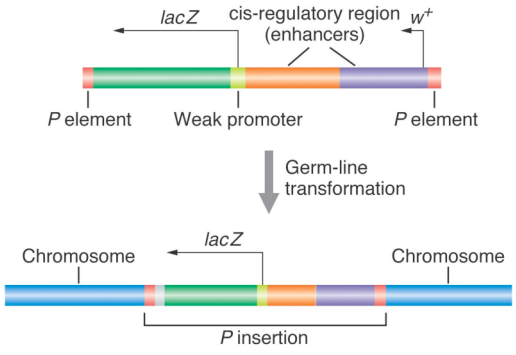
Dissecting eukaryotic regulatory elements

- Genetic analysis by classical means
- Use of transgenic reporter genes constructed by recombinant DNA technology
  - reporter gene, e.g., β-galactosidase fused to weak eukaryotic promoter
  - reporter construct fused to putative enhancer sequence (or sequences)
  - functional enhancer recognized by increased transcription of reporter gene
- Eventually clone gene encoding enhancer binding protein

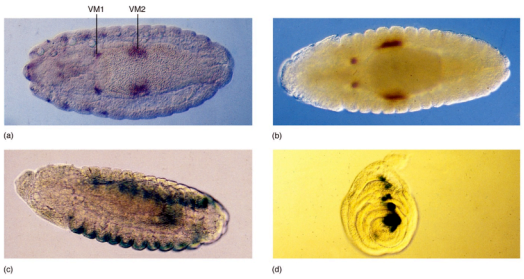
Tissue specific enhancers



Detecting enhancers

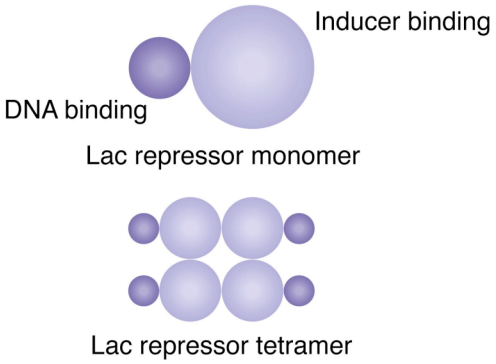


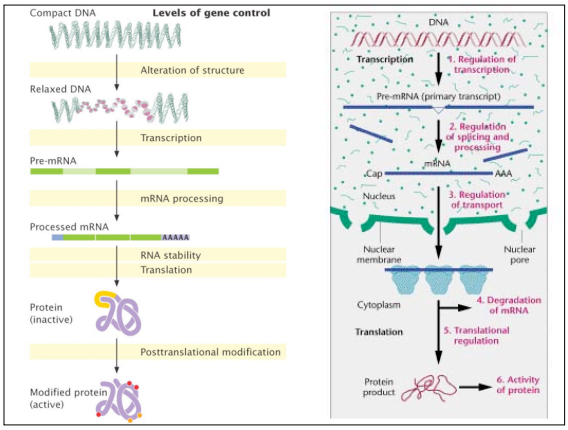
Blue areas have ß-galactosidase



Mouse embryo with expression marker

Lac repressor

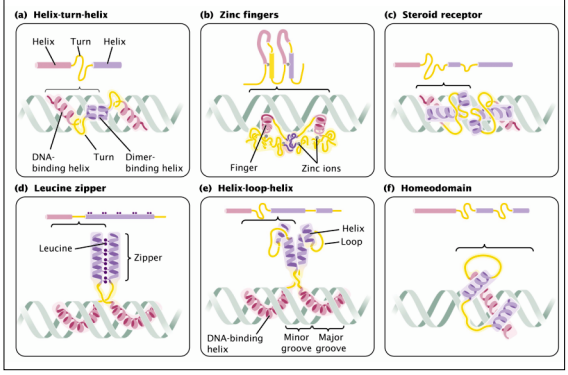




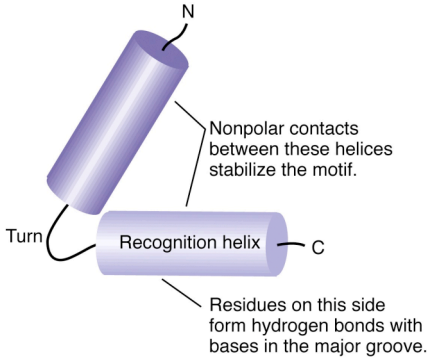
DNA-binding specificity

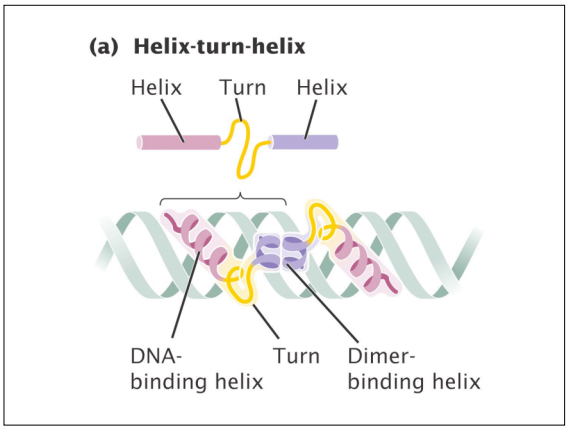
- DNA-binding regulatory proteins bind to specific nucleotide sequences (docking sequence)
- Several categories of DNA recognition motifs
  - helix-turn-helix
    - helices fit into major groove of DNA
    - common structure among regulatory proteins
  - zinc-finger protein
    - zinc atom is conjugated to two cysteines and two histidines
    - usually multiple zinc fingers per protein
  - Leucine zipper
- Coordinately regulated genes share promoter-proximal and/or enhancer sequences

DNA Binding motifs

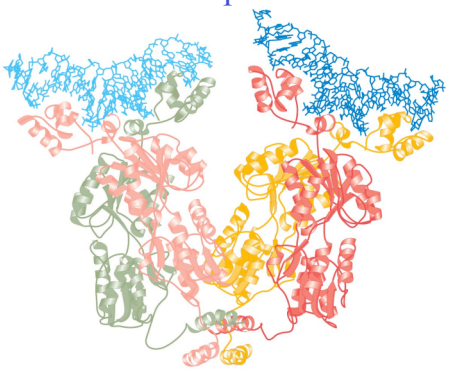


Helix turn helix

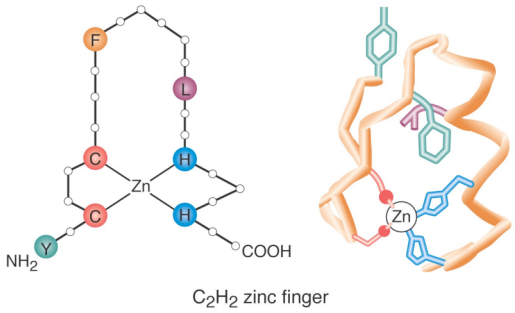


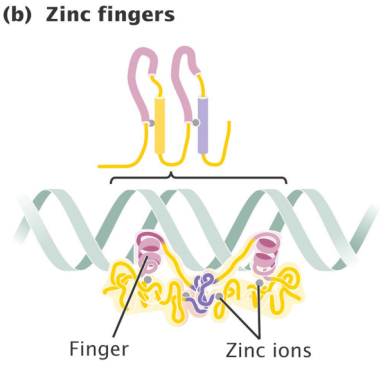


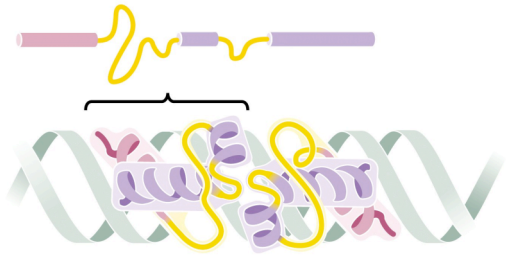
Lac repressor



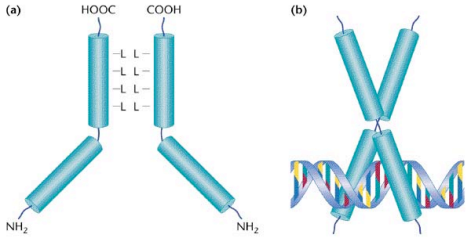
Zinc finger

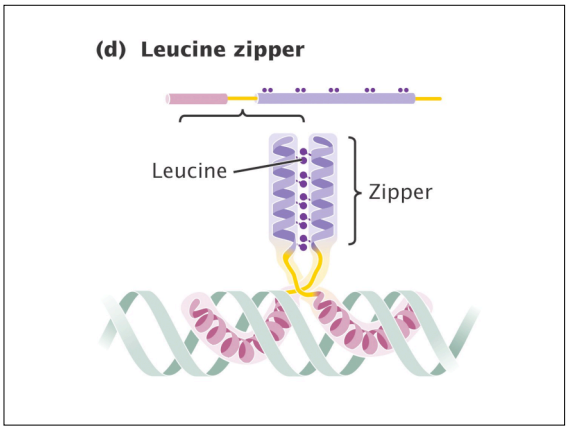


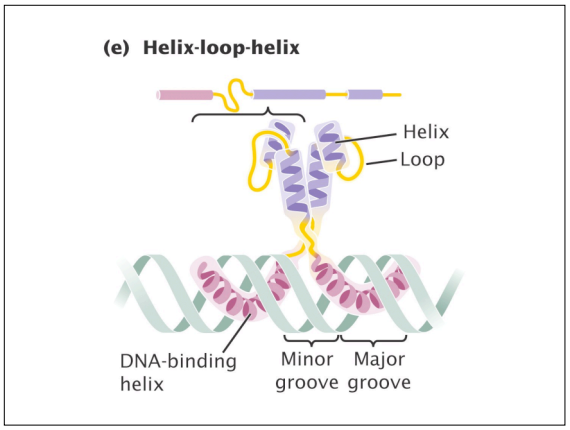




Leucine zipper







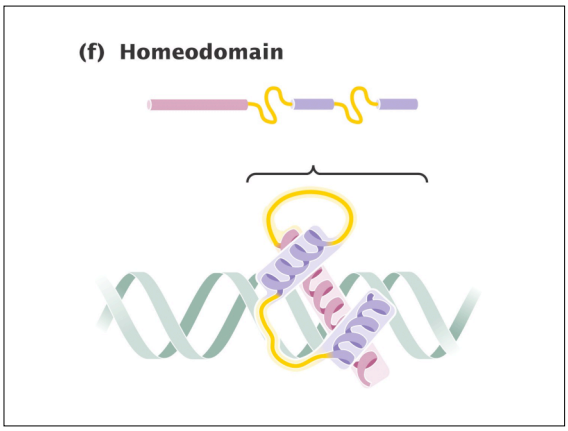
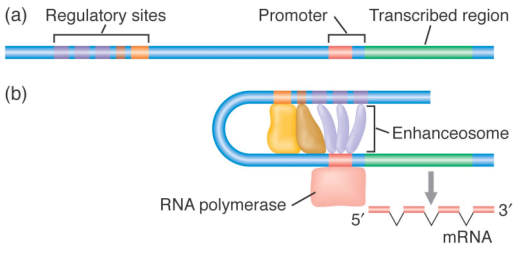


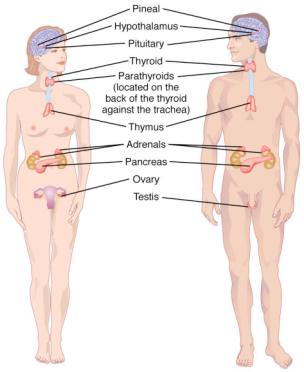
Table 16.1 Common DNA-binding motifs

Motif	Location	Characteristics	Binding Site in DNA
Helix-turn-helix	Bacterial regulatory proteins; related motifs in eukaryotic proteins	Two alpha helices	Major groove
Zinc-finger	Eukaryotic regulatory and other proteins	Loop of amino acids with zinc at base	Major groove
Steroid receptor	Eukaryotic proteins	Two perpendicular alpha helices with zinc surrounded by four cysteine residues	Major groove and DNA backbone
Leucine-zipper	Eukaryotic transcription factors	Helix of leucine residues and a basic arm; two leucine residues interdigitate	Two adjacent major grooves
Helix-loop-helix	Eukaryotic proteins	Two alpha helices separated by a loop of amino acids	Major groove
Homeodomain	Eukaryotic regulatory proteins	Three alpha helices	Major groove

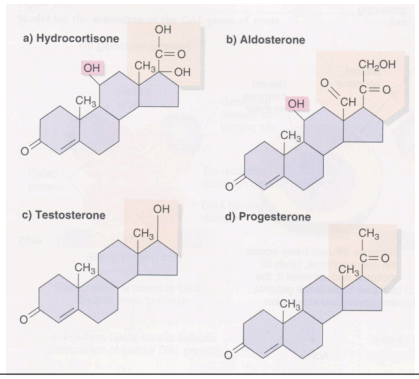
enhancers

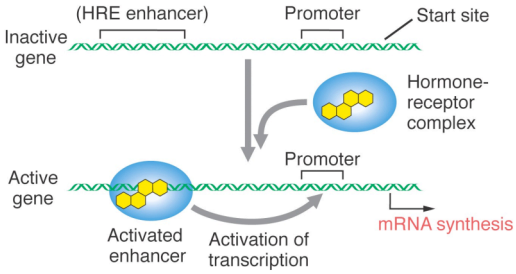


Endocrine system

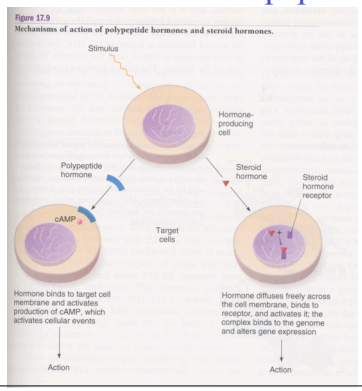


Steroids

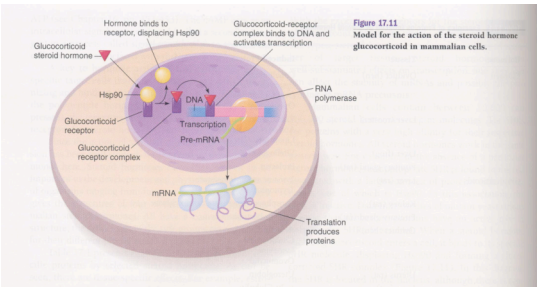




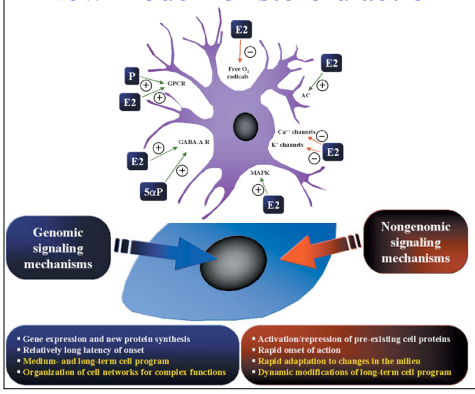
Steroid verses peptide



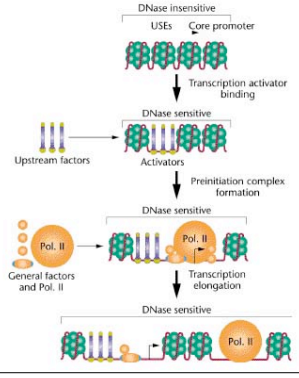
Steroid action

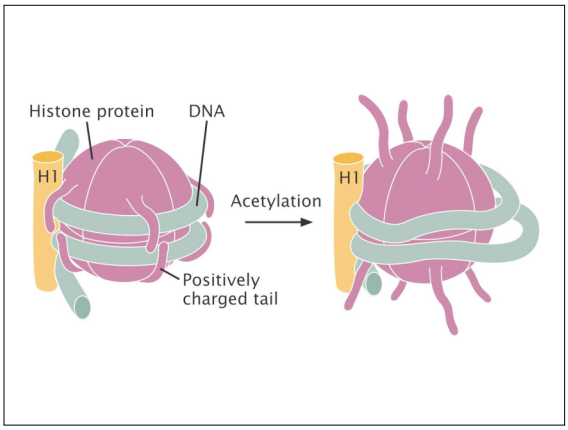


New model for steroid action

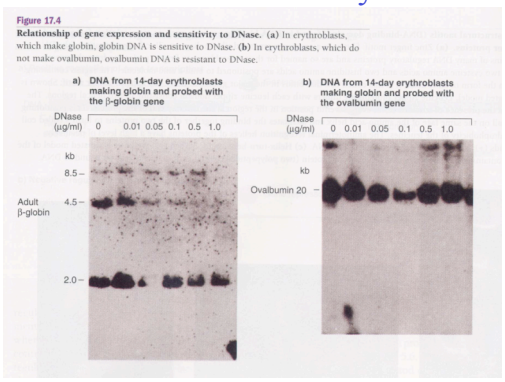


Nucleosomes inhibit transcription

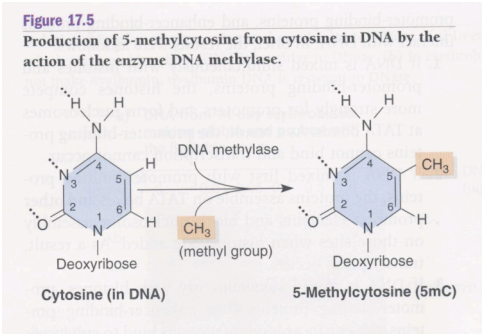




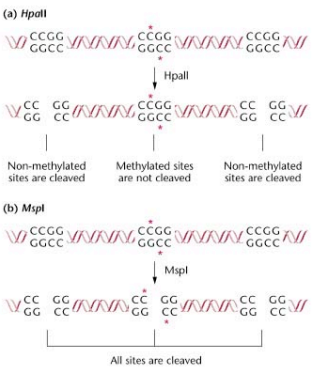
DNAse sensitivity



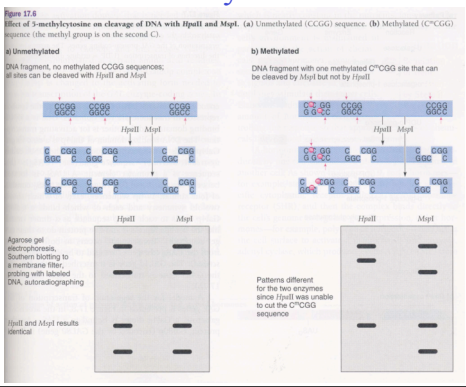
Methylation



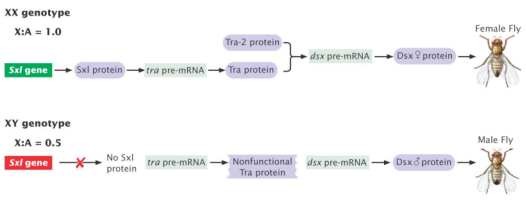
Methylation

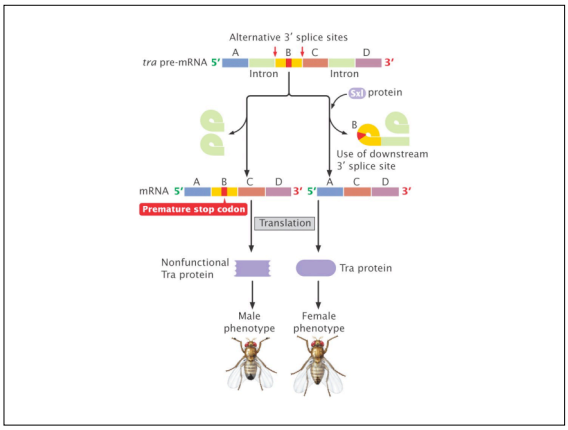


Methylation

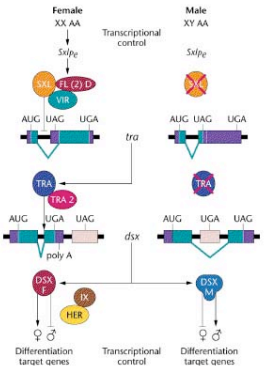


Alternate splicing

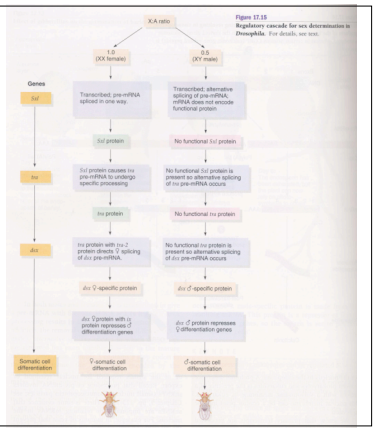




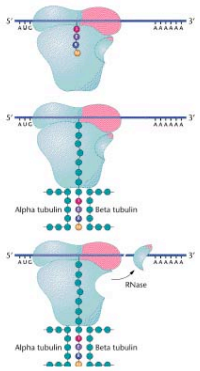
Alternative splicing



Alternative splicing



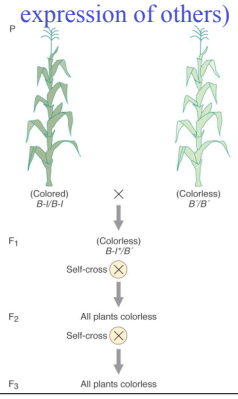
mRNA stability



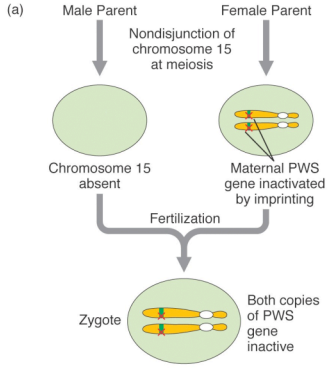
Epigenetic inheritance

- Heritable alterations in gene expression in which DNA sequence is unchanged
- Paramutation in plants
- Parental imprinting
  - genes rendered inactive depending upon parental source
  - Polydactlyly, Prader-Willi syndrome,
  - imprinted genes usually methylated by special methylases, demethylated by demethylases
    - active genes less methylated (hypomethylated)
  - Male genes -placenta development
  - Female genes - embryo development
- X-chromosome inactivation
  - random X chromosome inactivation in placental females
  - inactivation hereditary through cell division

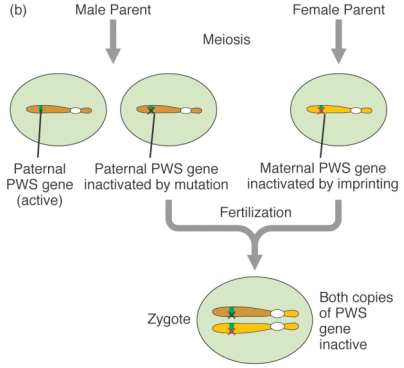
Paramutation in plants (one gene prevents the expression of others)



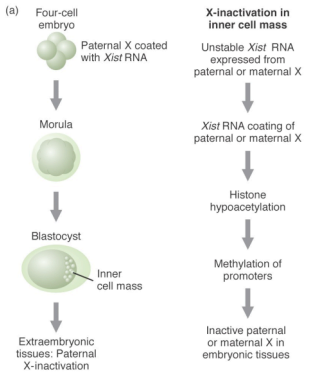
Prader-Willi syndrome nondisjunction



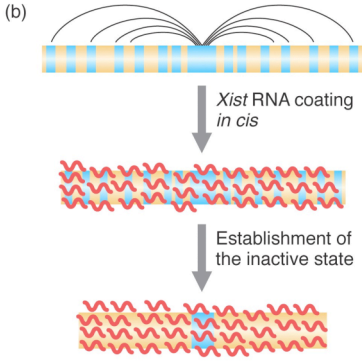
Prader-Willi syndrome imprinting



Methylation and acetylation

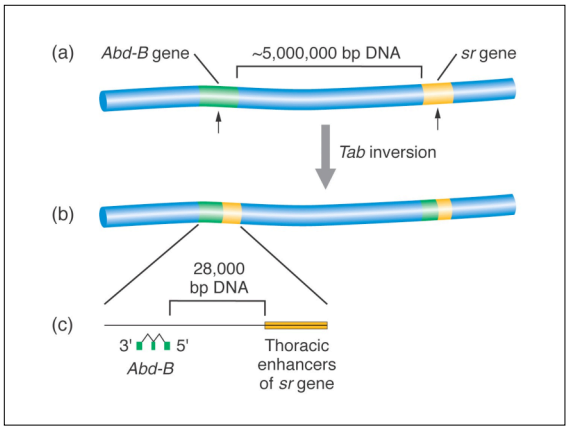


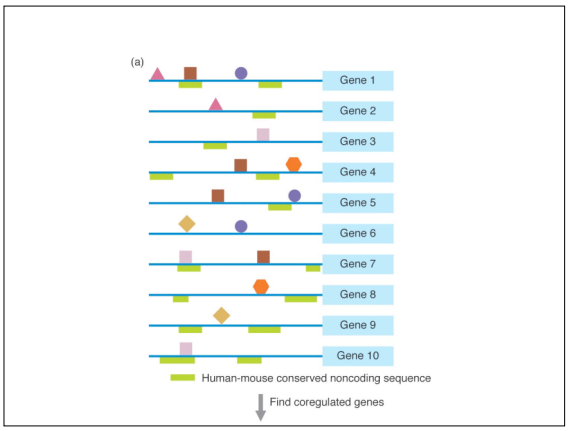
Xist RNA inactivation

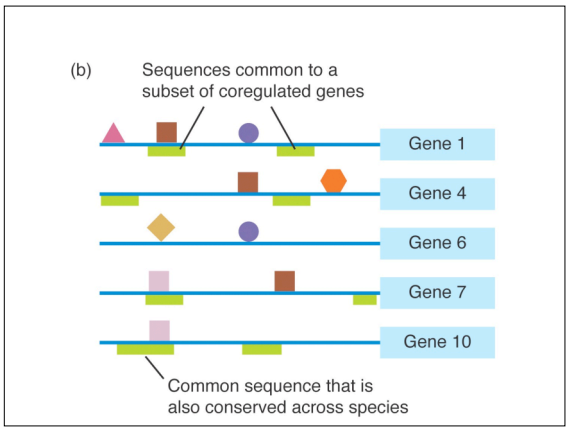


Overview

- Response to environmental signals may alter transcription rates of genes (signal transduction).
- In prokaryotes, coordinately regulated genes are clustered and transcribed as multigenic mRNAs.
- The *lac* system is a model of negative regulation.
- Positive regulation requires protein factors to activate transcription.
- Regulatory proteins share common features.
- In eukaryotes, regulatory sites may be at considerable distance from regulated gene.
- Some regulation is epigenetic, modulated by parentage or cell division history.


(a)
Abd-B gene
~5,000,000 bp DNA
sr gene
Tab inversion
(b)
28,000
bp DNA
(c)
3'
5'
Abd-B
Thoracic
enhancers
of sr gene


(a)
Gene 1
Gene 2
Gene 3
Gene 4
Gene 5
Gene 6
Gene 7
Gene 8
Gene 9
Gene 10
Human-mouse conserved noncoding sequence
Find coregulated genes


(b)
Sequences common to a
subset of coregulated genes
Gene 1
Gene 4
Gene 6
Gene 7
Gene 10
Common sequence that is
also conserved across species