Gene Regulation

Gene regulation

Transcriptional regulation.

1. **Transcriptional activators and repressors**
2. **Chromatin remodeling factors**
3. **DNA methylation (e.g. imprinting)**

mRNA splicing

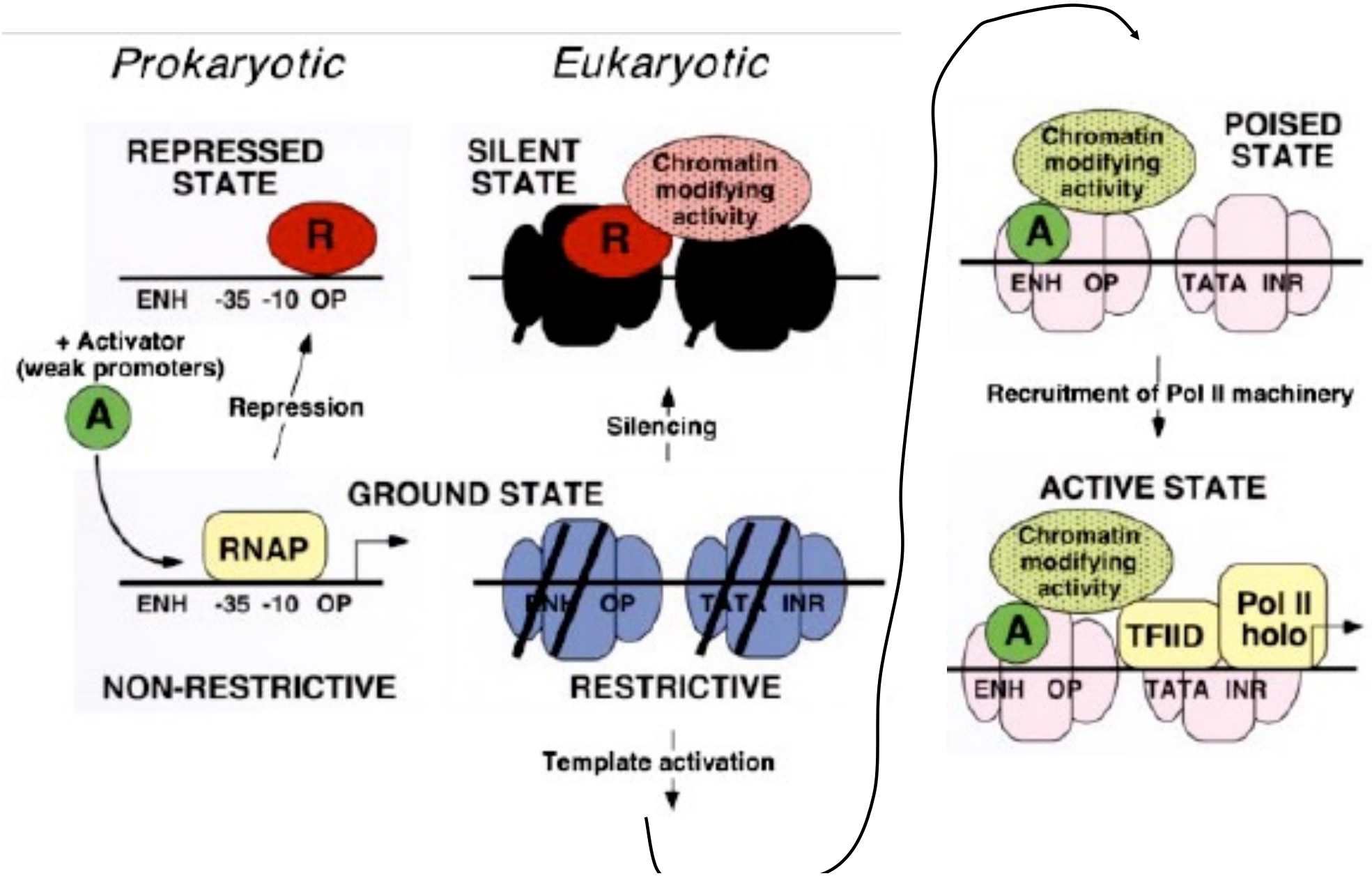
mRNA degradation

Translational regulation

Protein stability

Protein modification by other trans acting factors (e.g. phosphorylation).

Prokaryotic DNA is primed for transcription, which can be actively enhanced, or repressed. Eukaryotic chromatin is transcriptionally off, until turned on.



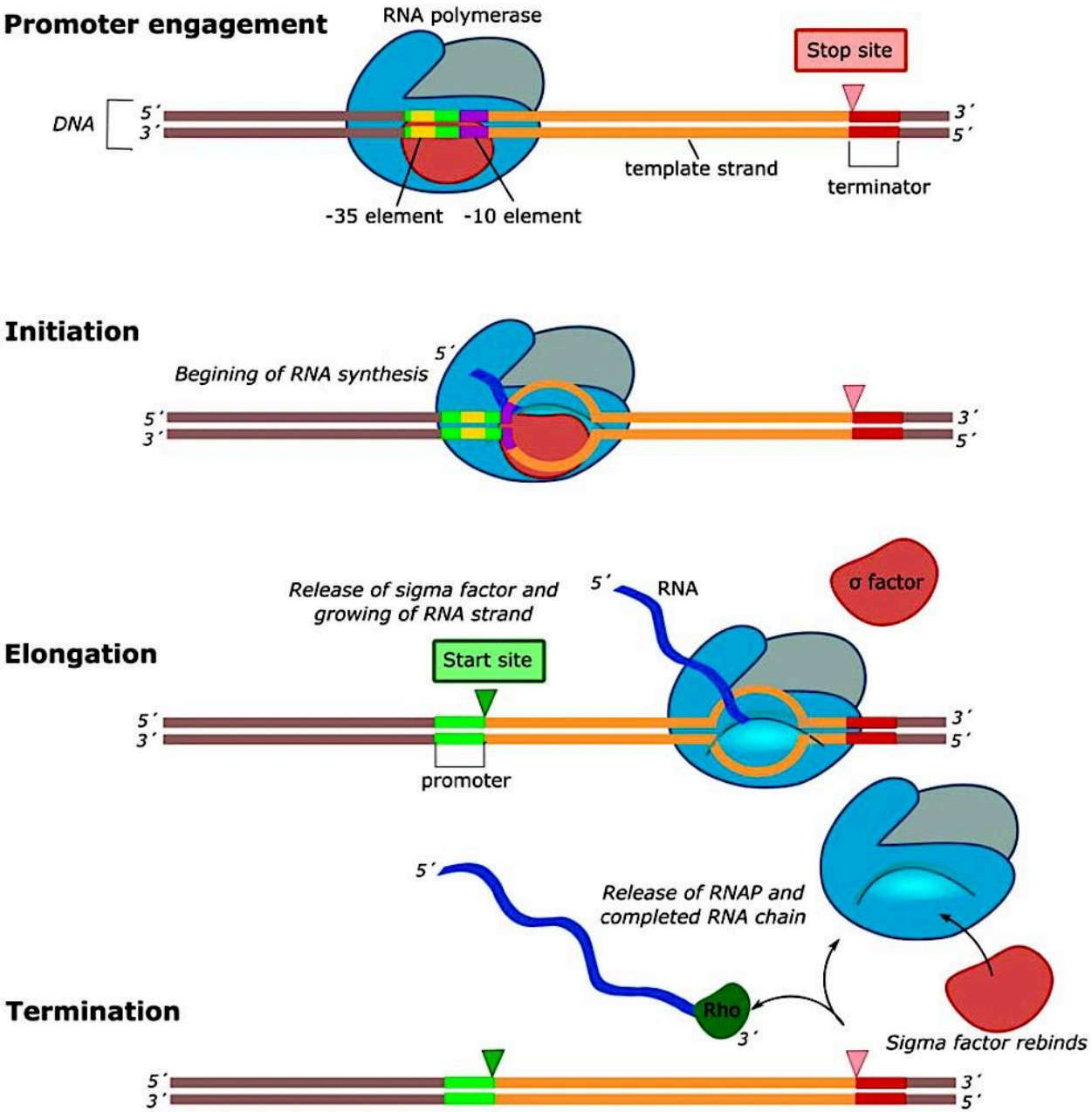
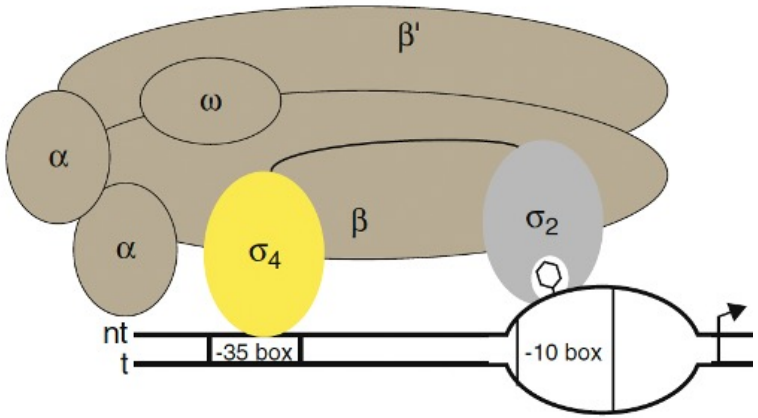
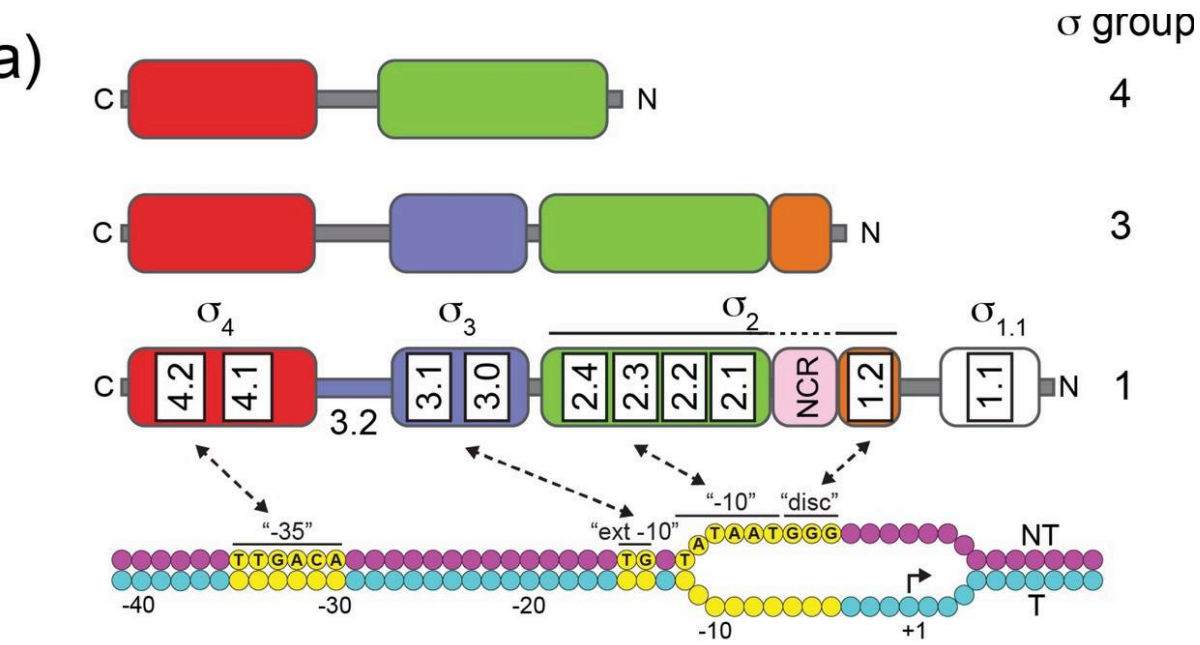


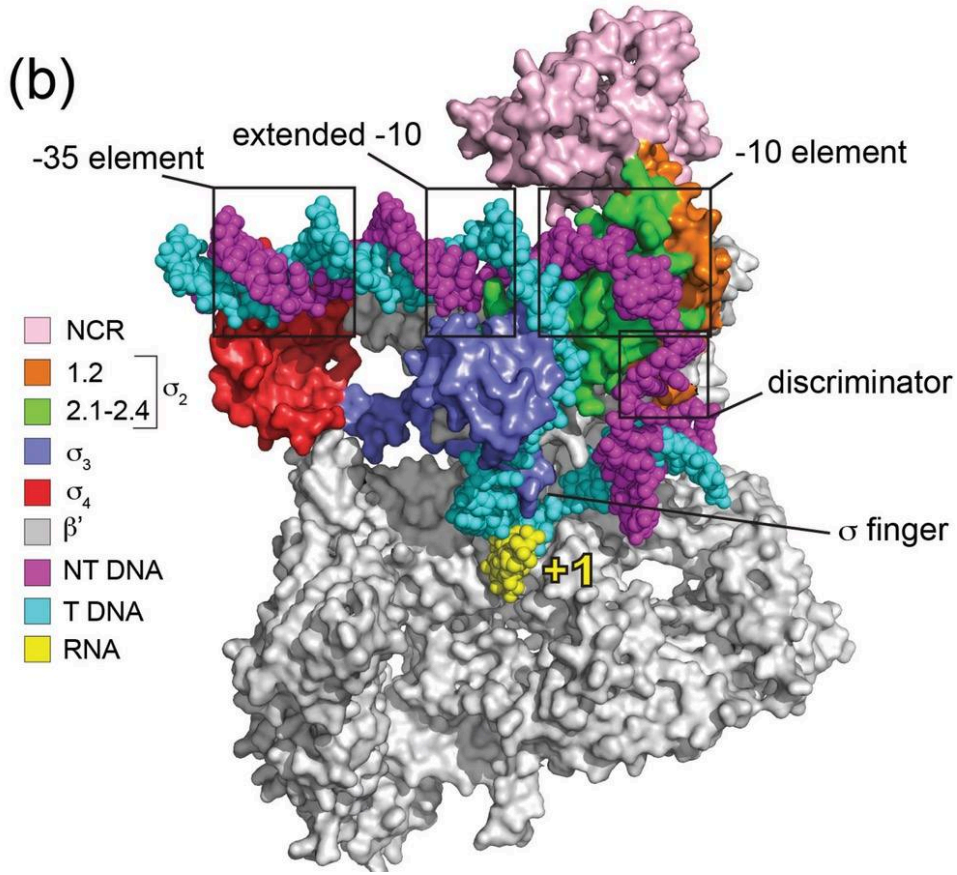
Fig. 1 Graphic representation of the bacterial transcription process. Transcription begins with the recognition of the −10 Schaller-Pribnow box (purple rectangle), and the −35 element (yellow box) by the sigma factor (reddish oval) leading to the promoter (green rectangle) engagement. Transcription initiation requires formation of the transcription bubble, with RNA (dark blue wavy line) synthesis initiating at the start site (green triangle). Sigma factor falls off during RNA elongation, allowing the RNAP to continue copying the DNA (thick gray line) into RNA. Rho factor (dark green oval) moves along the forming RNA chain, following the RNAP complex until it reaches the terminator (thick red line). This completes transcription of the gene (thick orange line), Rho unwinds the DNA-RNA hybrid within the transcription bubble, and the RNAP, Rho, and newly formed RNA strand are released

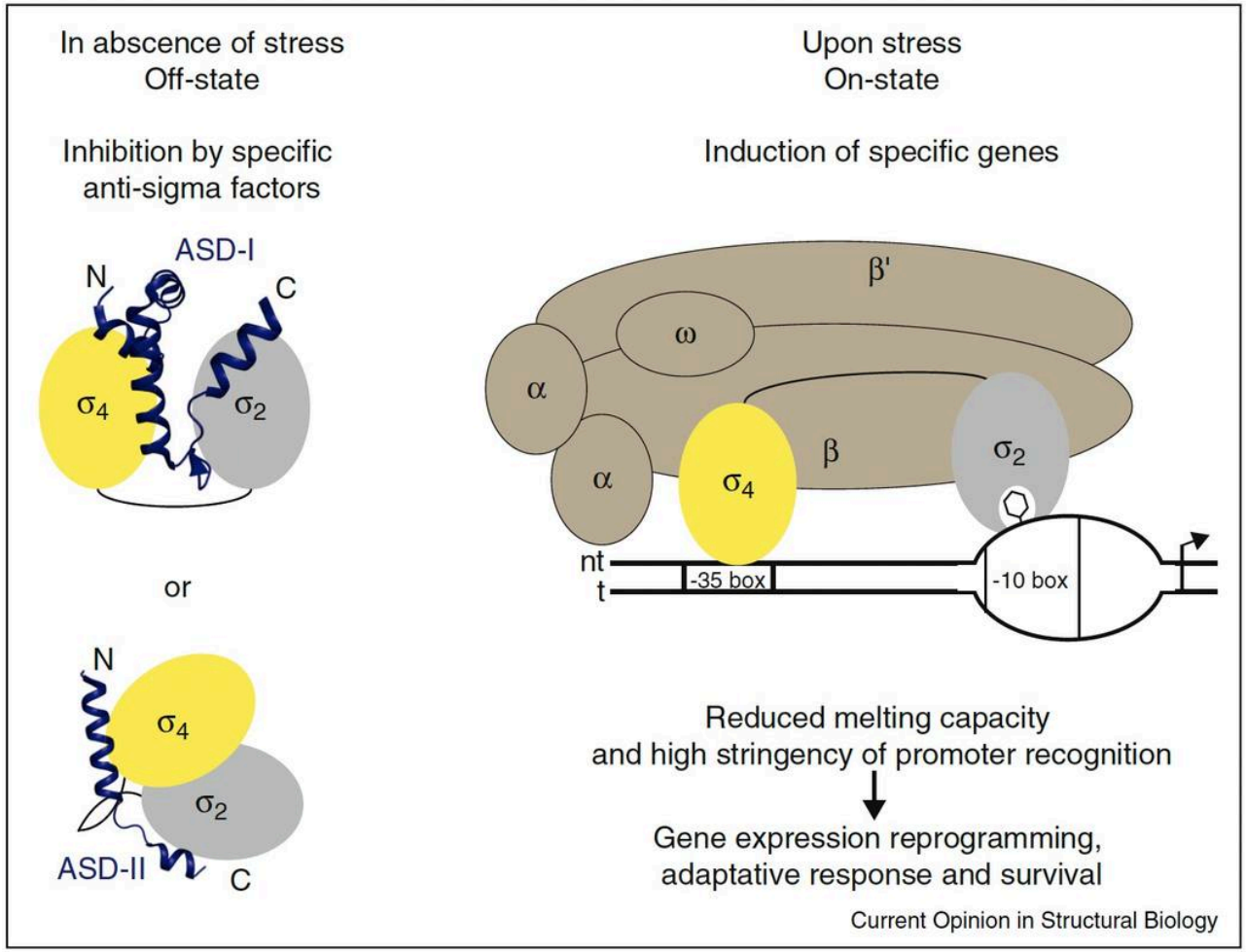
Multiple sigma factors within the cell compete for binding to a limiting amount of RNA polymerase.

Each sigma factor binds near multiple genes, resulting in different multi-gene transcriptional responses

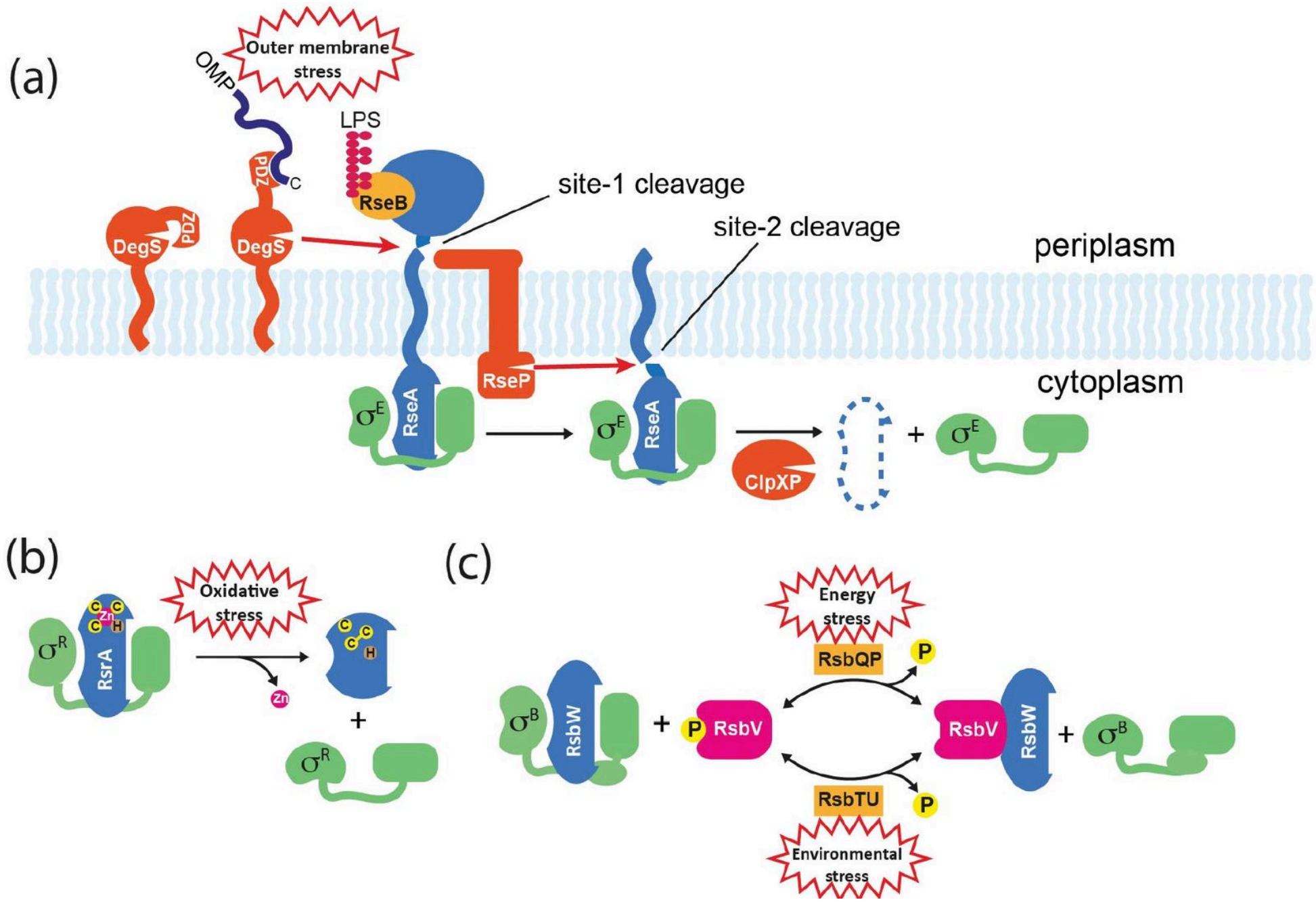








Mode of action of σ^{ECF}. In the left panel, the inactivation of σ^{ECF} by anti-sigma factors is schematically represented. The sigma factor subdomains σ_2 and σ_4 are colored in gray and yellow, respectively, whereas the anti-sigma factors are colored in blue. ASD stands for anti-sigma factor domain, which can be of class I or II. The right panel schematically illustrates the holoenzyme containing a σ^{ECF}. The RNAP catalytic core is colored in brown. On the promoter, the position of the -35 and -10 boxes are indicated. At the level of the -10 promoter element, the transcription bubble is illustrated with the base flipped out from the -10 element non-template strand. The labels t and nt refer to template and non-template strands, respectively.


(a)
Outer membrane stress
OMP
LPS
PDZ
C
RseB
site-1 cleavage
site-2 cleavage
periplasm
DegS
PDZ
DegS
RseP
cytoplasm
σE
RseA
σE
RseA
ClpXP
+
σE
(b)
Oxidative stress
σR
RsrA
C
C
C
Zn
H
Zn
C
C
H
+
σR
(c)
Energy stress
RsbQP
P
σB
RsbW
+
P
RsbV
RsbV
RsbW
+
σB
RsbTU
P
Environmental stress

Lactose is broken down into two sugars and modified into allolactose

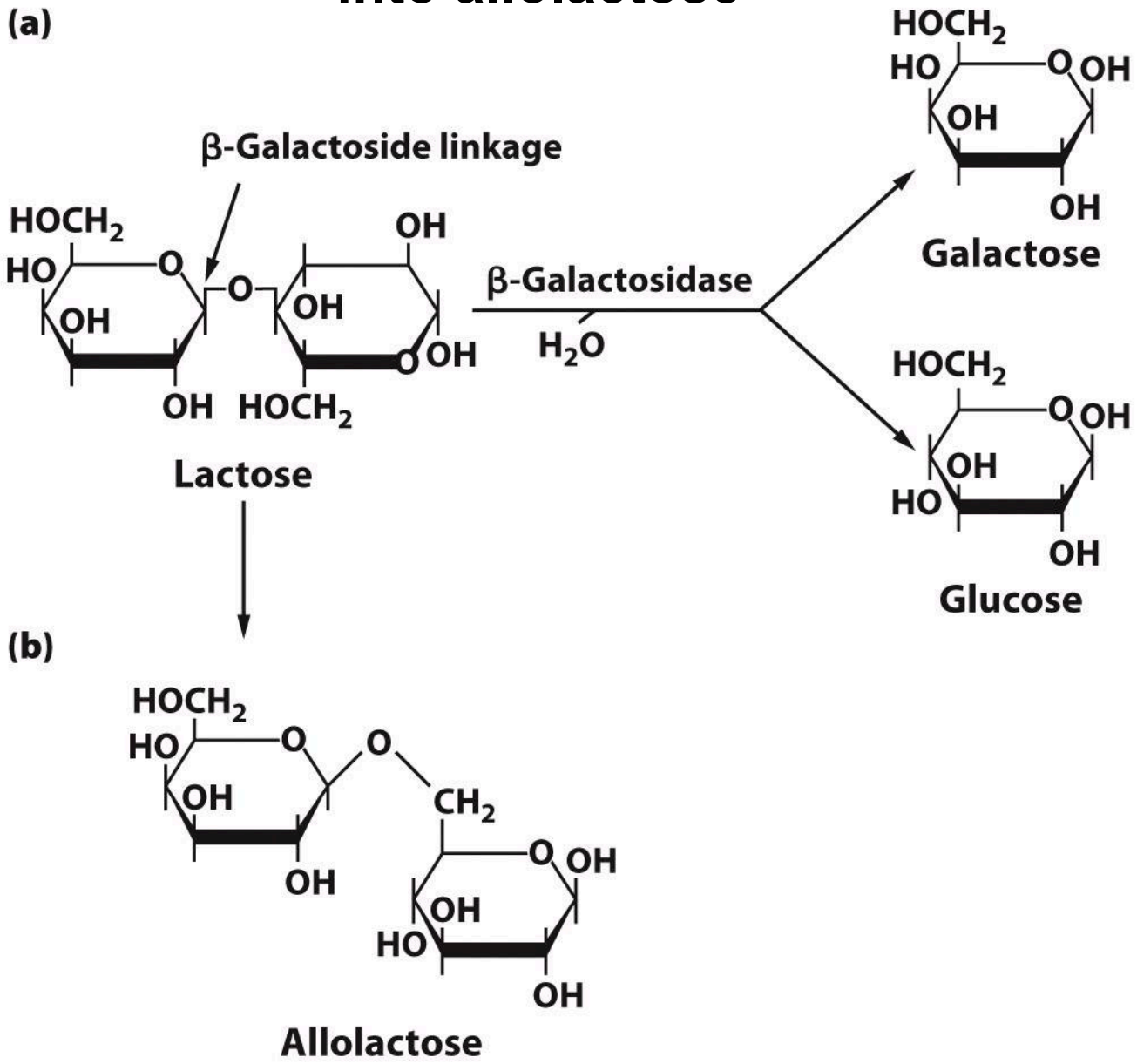


Figure 11-5
Introduction to Genetic Analysis, Twelfth Edition

Regulatory proteins control transcription

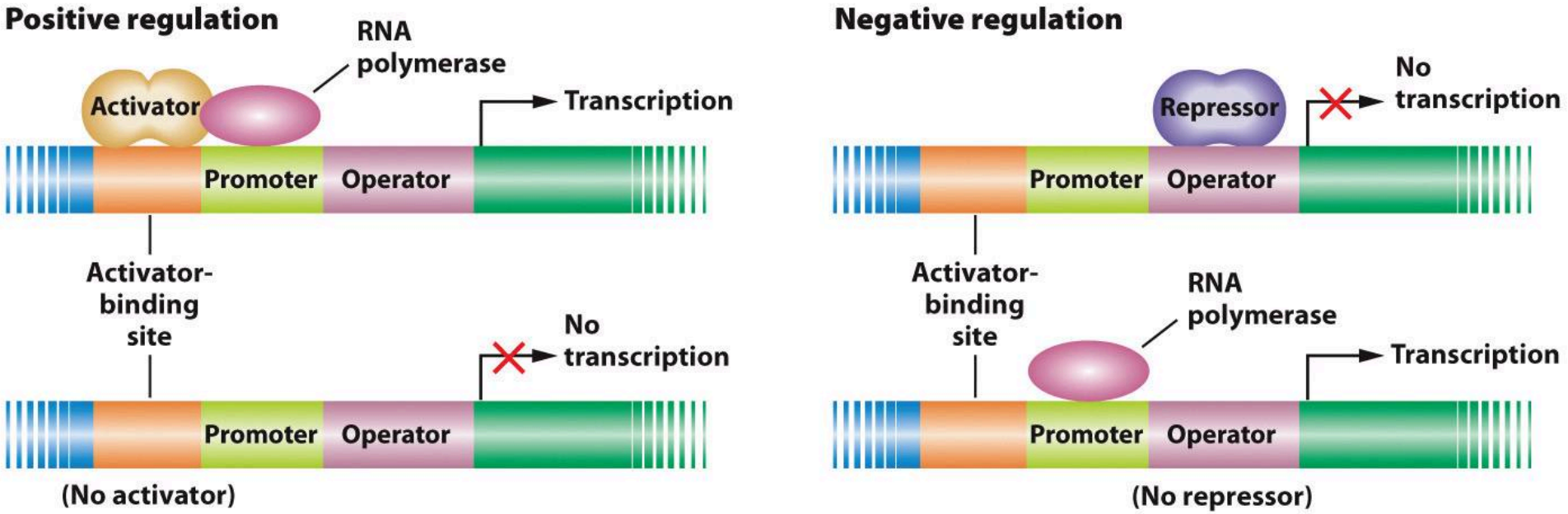
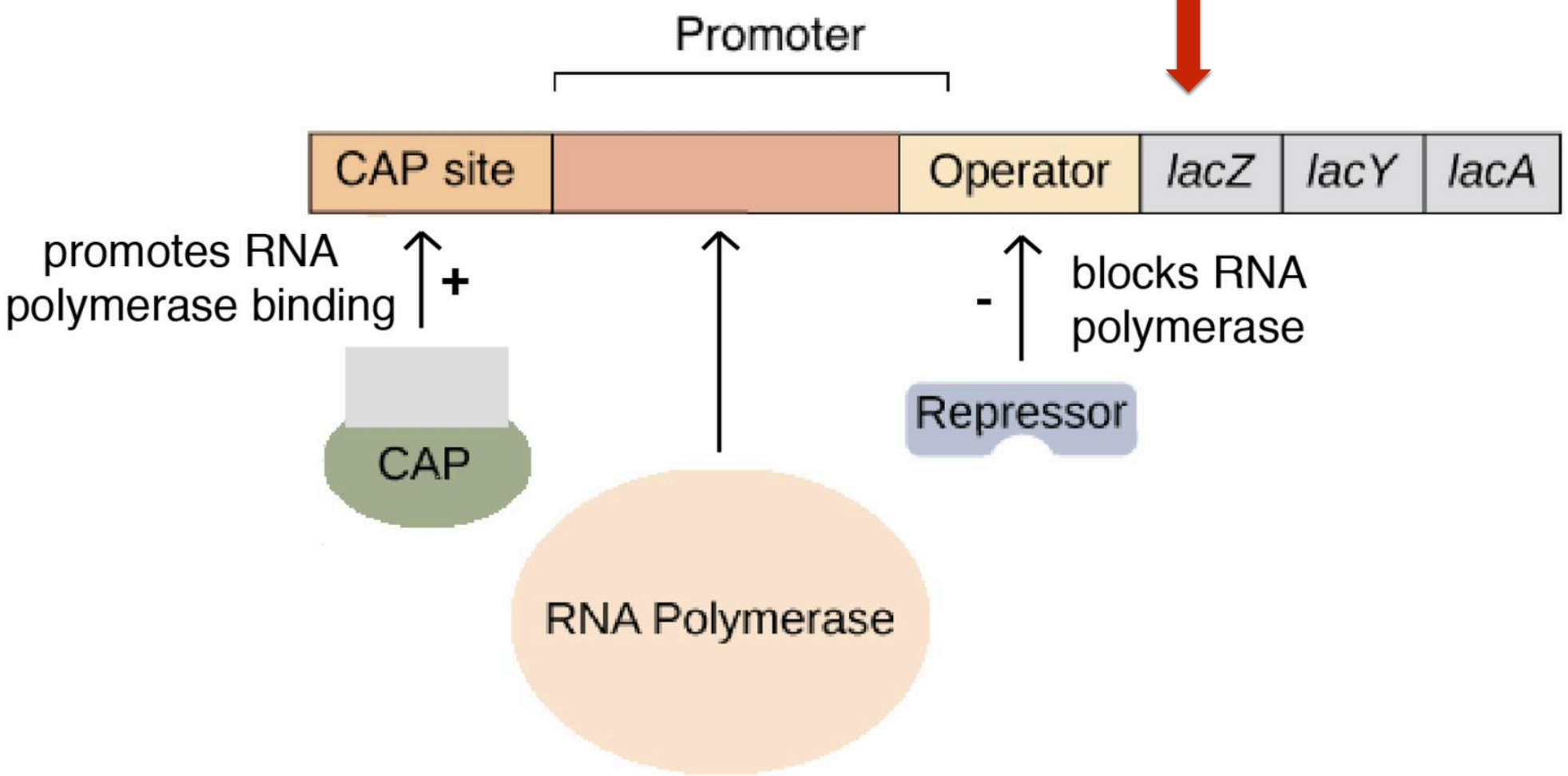


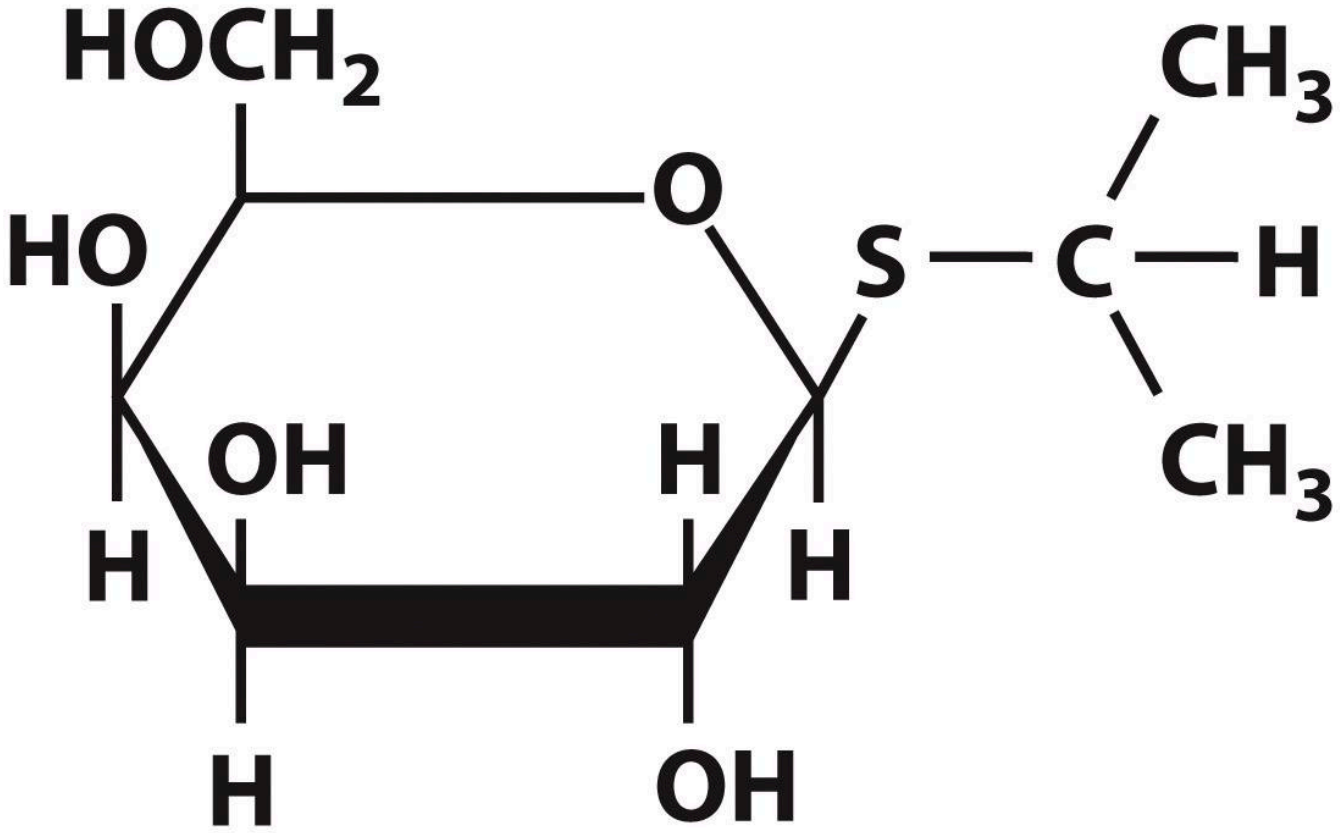
Figure 11-2
Introduction to Genetic Analysis, Twelfth Edition

lacZ gene produces
B-galactosidase

The *lac* operon:



Structure of IPTG, a lactose analog that is not broken down to glucose



Isopropyl-β-D-thiogalactoside (IPTG)

Figure 11-7
Introduction to Genetic Analysis, Twelfth Edition

Kinetics of growth in presence of glucose and galactose

growth

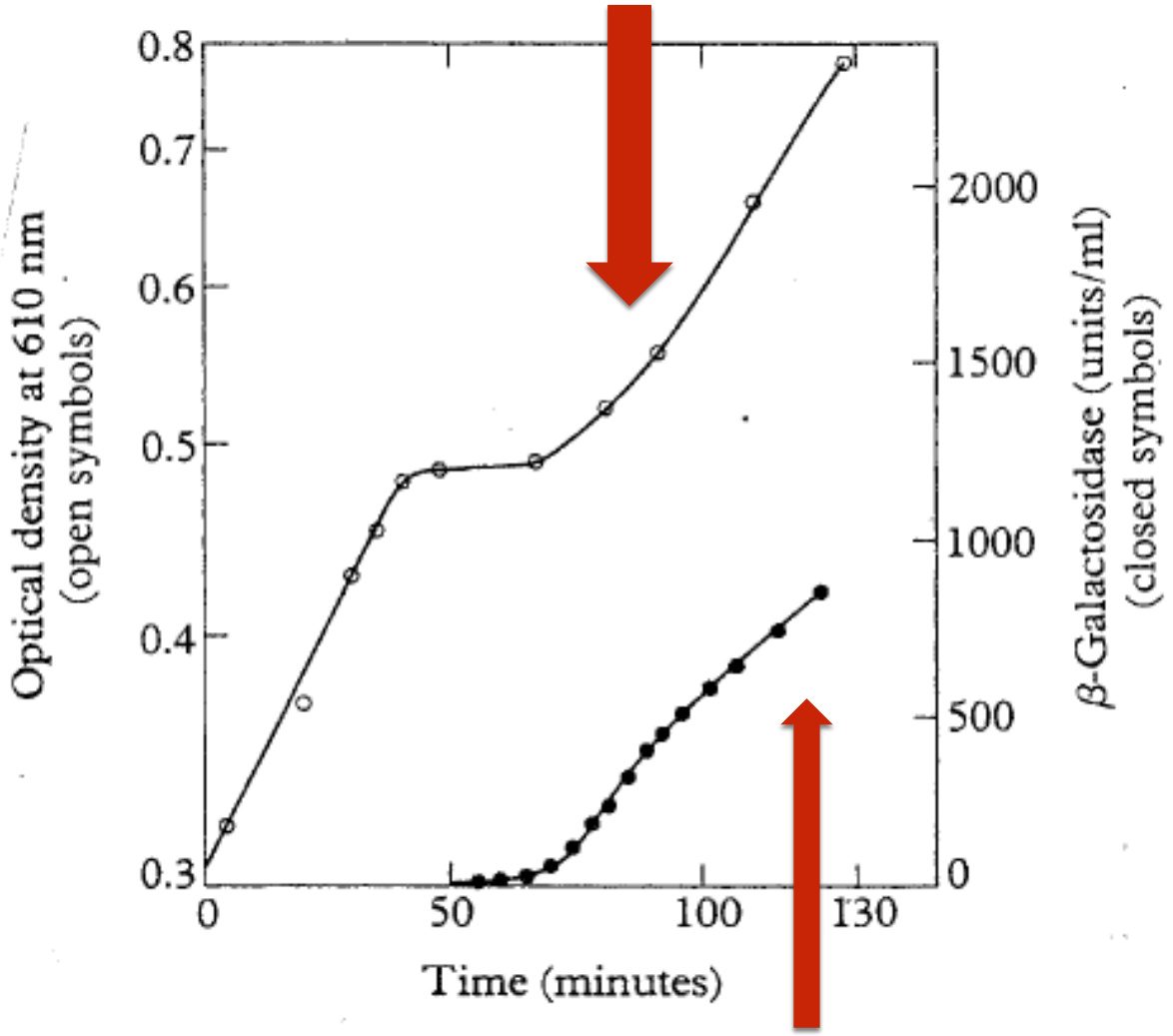
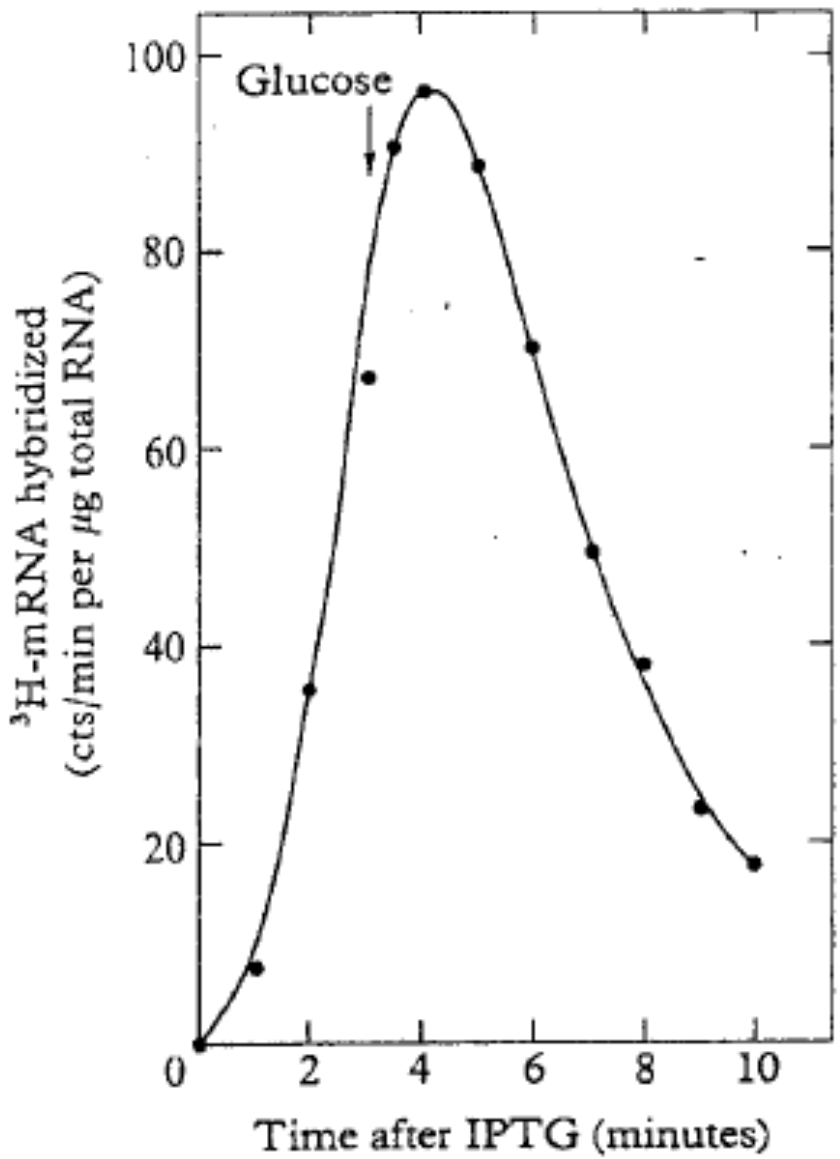


FIGURE 20-10
Delayed appearance of β-galactosidase in *E. coli* growing in a medium containing initially 0.4 mg/ml of glucose and 2 mg/ml of lactose. The left-hand ordinate indicates the cell density of the growing culture. [From W. Epstein, S. Naono, and F. Gros, *Biochem. Biophys. Res. Commun.* **24**, 588 (1966).]

Beta-galactosidase

Glucose inhibits expression of the lac operon

FIGURE 20-8
Kinetics of synthesis of *lac* operon messenger RNA following induction, and of its degradation following addition of glucose. *E. coli* was grown in a synthetic medium containing glycerol as the carbon source and ^{3}H-uridine to label mRNA. IPTG was added to the culture to induce synthesis of *lac* enzymes, and 3 min later, glucose was added to stop synthesis of these enzymes. RNA was extracted from samples of the culture at the times indicated on the abscissa and assayed for ^{3}H-labeled *lac* mRNA, by hybridization to the DNA of *lambda* transducing phage carrying the *lac*Y and *lac*Z genes. The ordinate presents the relative amount of ^{3}H-labeled *lac* operon mRNA hybridized, in cts/min per μg total RNA in the extract. [After M. Adesnik and C. Levinthal, *Cold Spring Harbor Symp. Quant. Biol.* **35**, 451 (1970).]



The *lac* operon is transcribed only in the presence of lactose

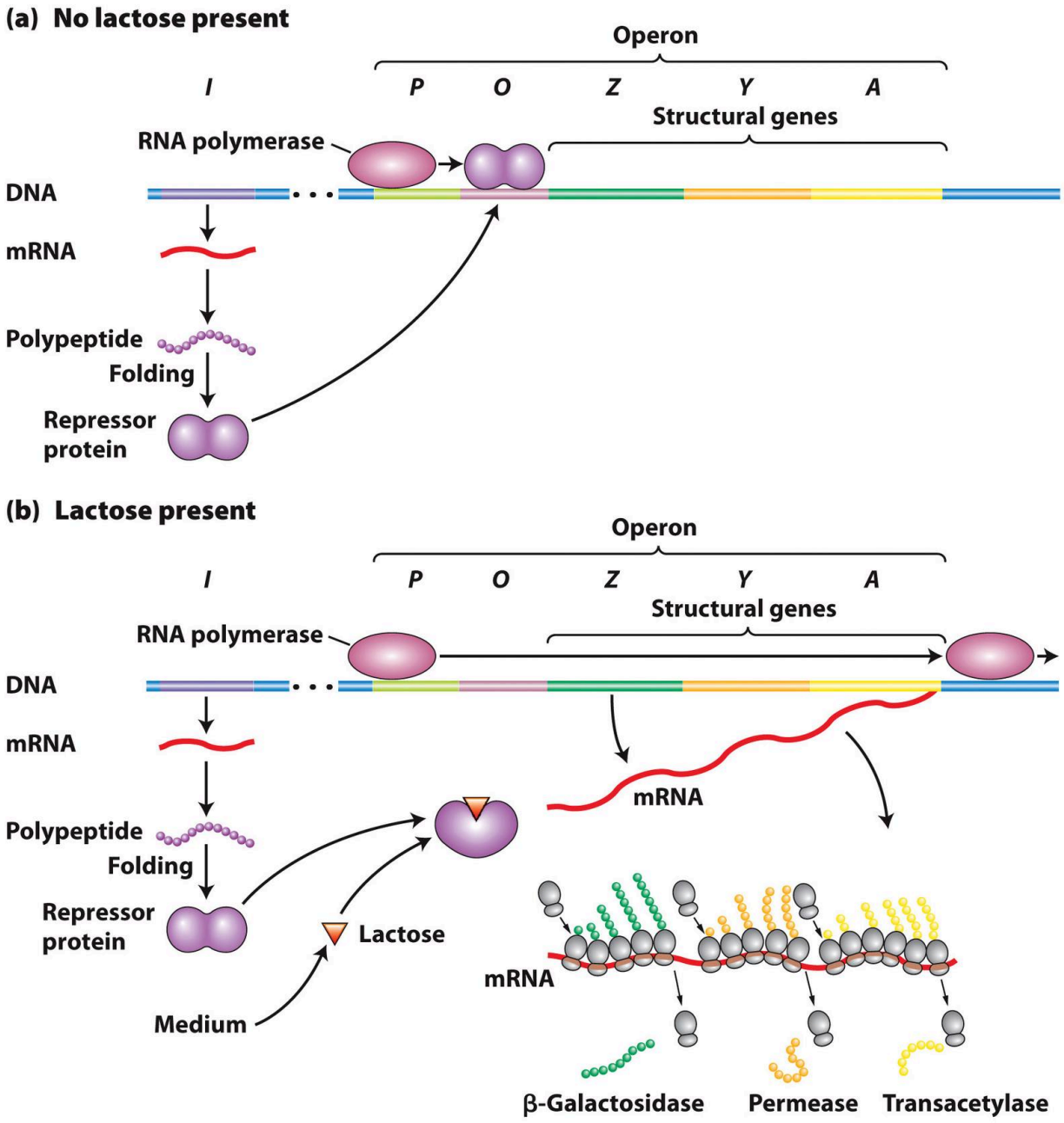


Figure 11-6
Introduction to Genetic Analysis, Twelfth Edition

Allosteric effectors bind to regulatory proteins

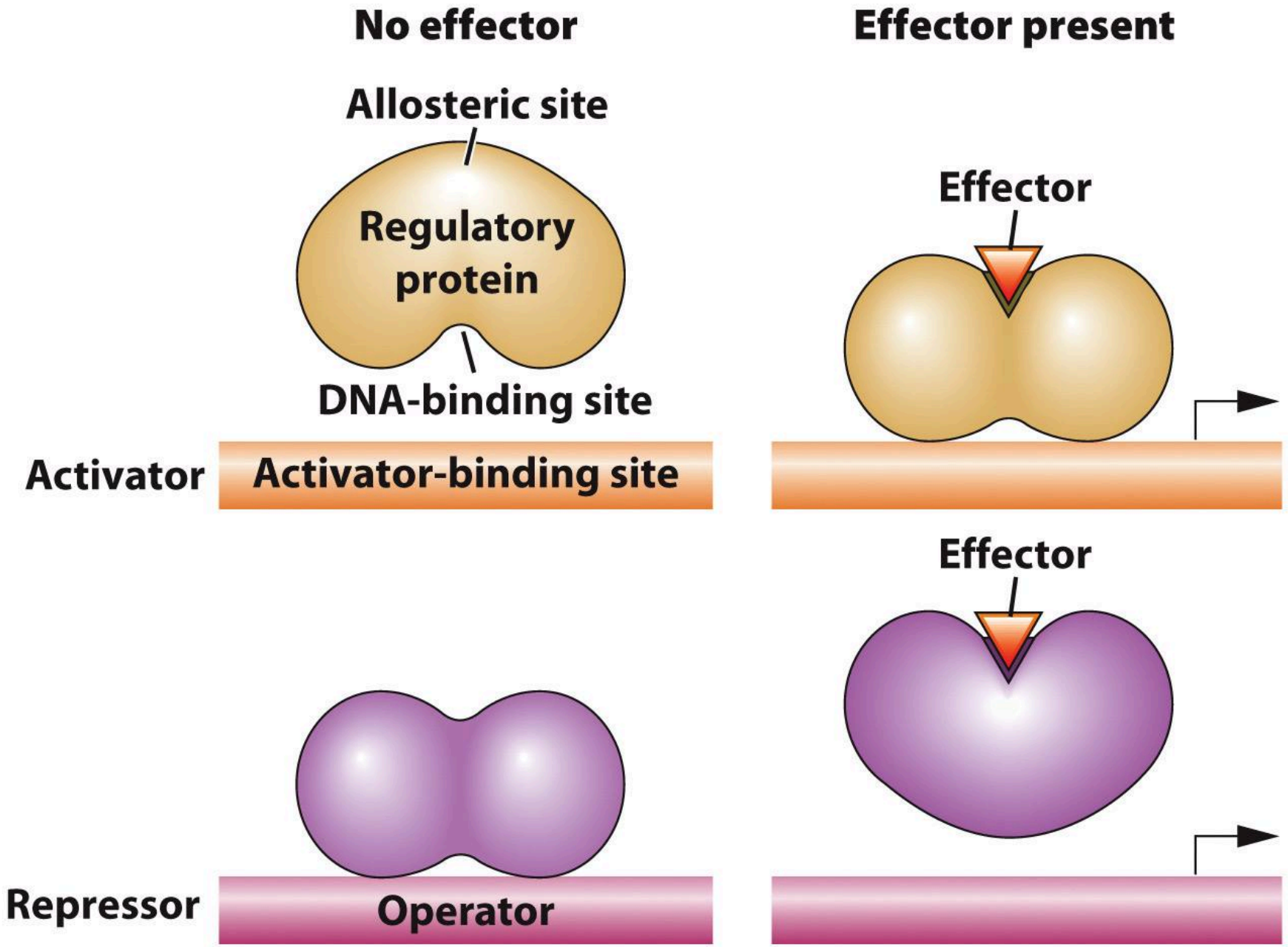
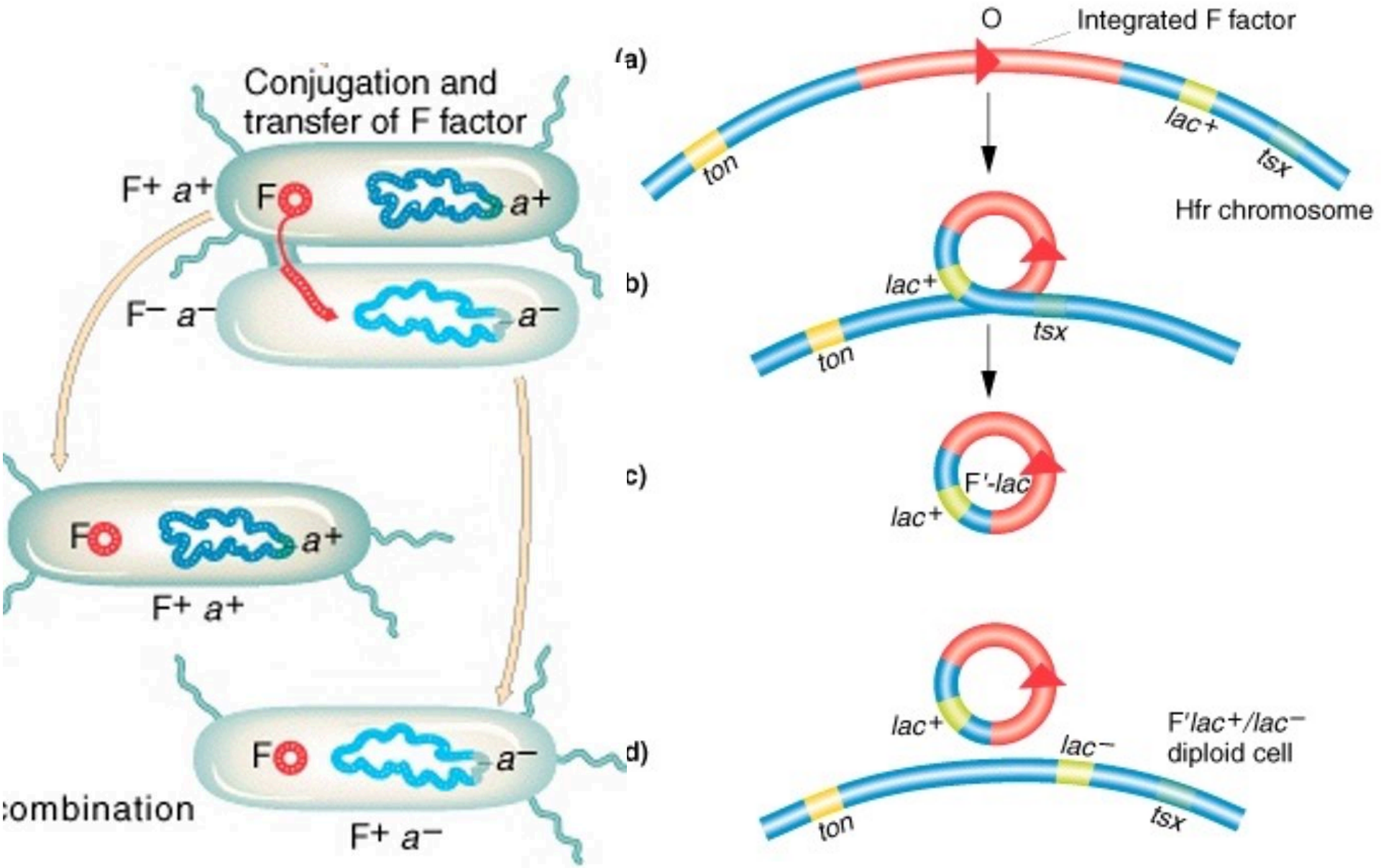


Figure 11-3
Introduction to Genetic Analysis, Twelfth Edition

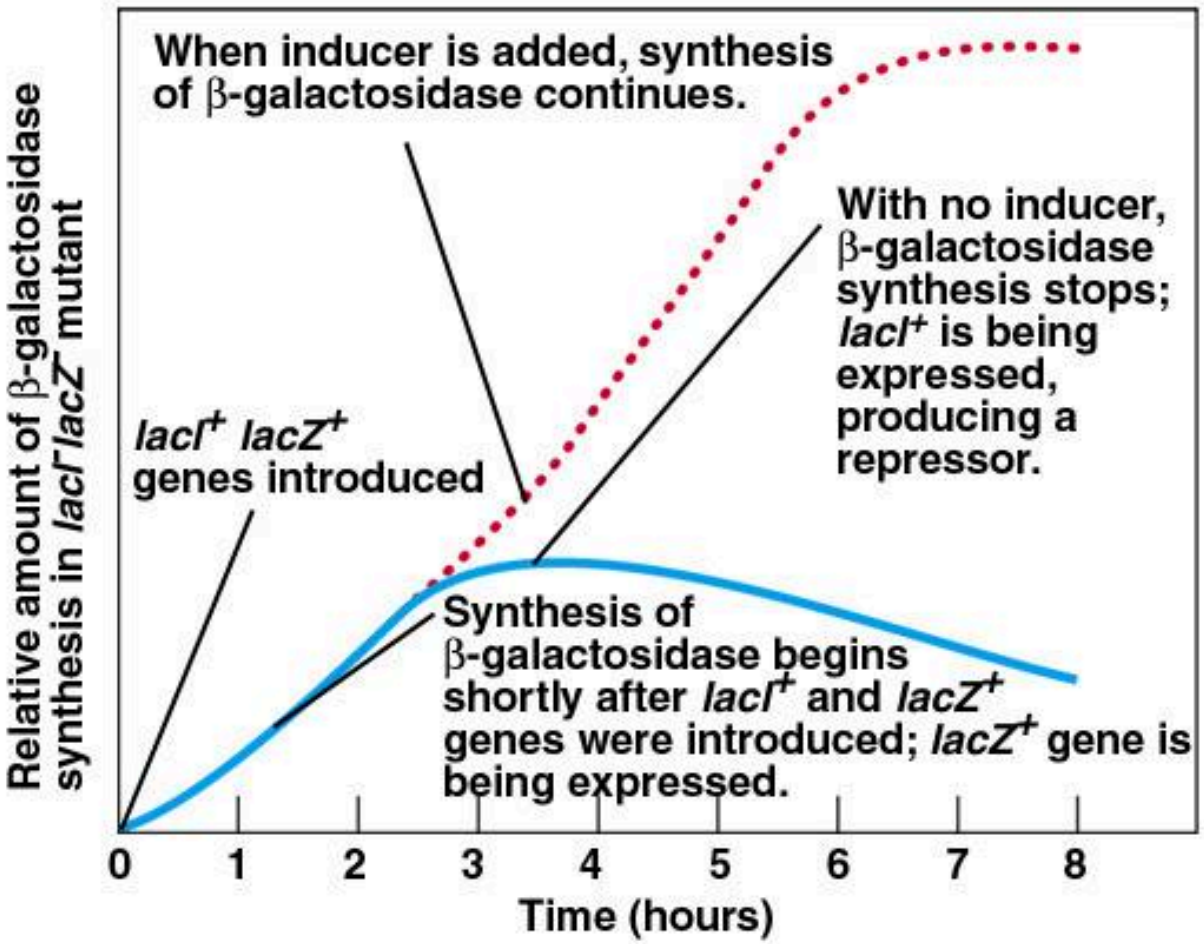
The F'. Infectious transfer of an F element that has picked up a small number of cellular genes during excision from the bacterial chromosome.



Experiment # 1

Str^S, tsx^S, lacI+ z+ (Hfr) X (F-) lacI- z- Str^r, tsx^r

The PaJaMo experiment



When DNA carrying lacI+ lacZ+ genes was introduced into a lacI-lacZ- cell through conjugation, Bgal was synthesized from the introduced lacZ+ gene initially. However, as repressor (made by the introduced lacI+ copy of the gene) accumulates, the synthesis of Bgal stops. The levels Of Bgal slowly decay as the protein is degraded. If inducer is added (dotted line), the synthesis of Bgal resumes. The logic behind the strains Used is the following: After enough time to allow for transfer of the lac genes from the donor, addition of T6 phage and streptomycin kills the Donor cells selectively. Inducer is not added until these cells have been killed. Otherwise Bgal would be induced in the donor. We are only Interested in the fate of the lacZ gene when it is introduced into an environment in which there is initially no lacI protein.

Repressors are trans-acting

TABLE 11-2 Synthesis of β-Galactosidase and Permease in Haploid and Heterozygous Diploid Strains Carrying I^+ and I^-

Strain	Genotype	β-Galactosidase (Z)		Permease (Y)		Conclusion
		Noninduced	Induced	Noninduced	Induced	
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) with and without the inducer IPTG. Expression of maximal enzyme levels is indicated by +. Absence or very low levels of enzyme activity is indicated by –. All strains are O^+.

Table
Introdu
© 202(

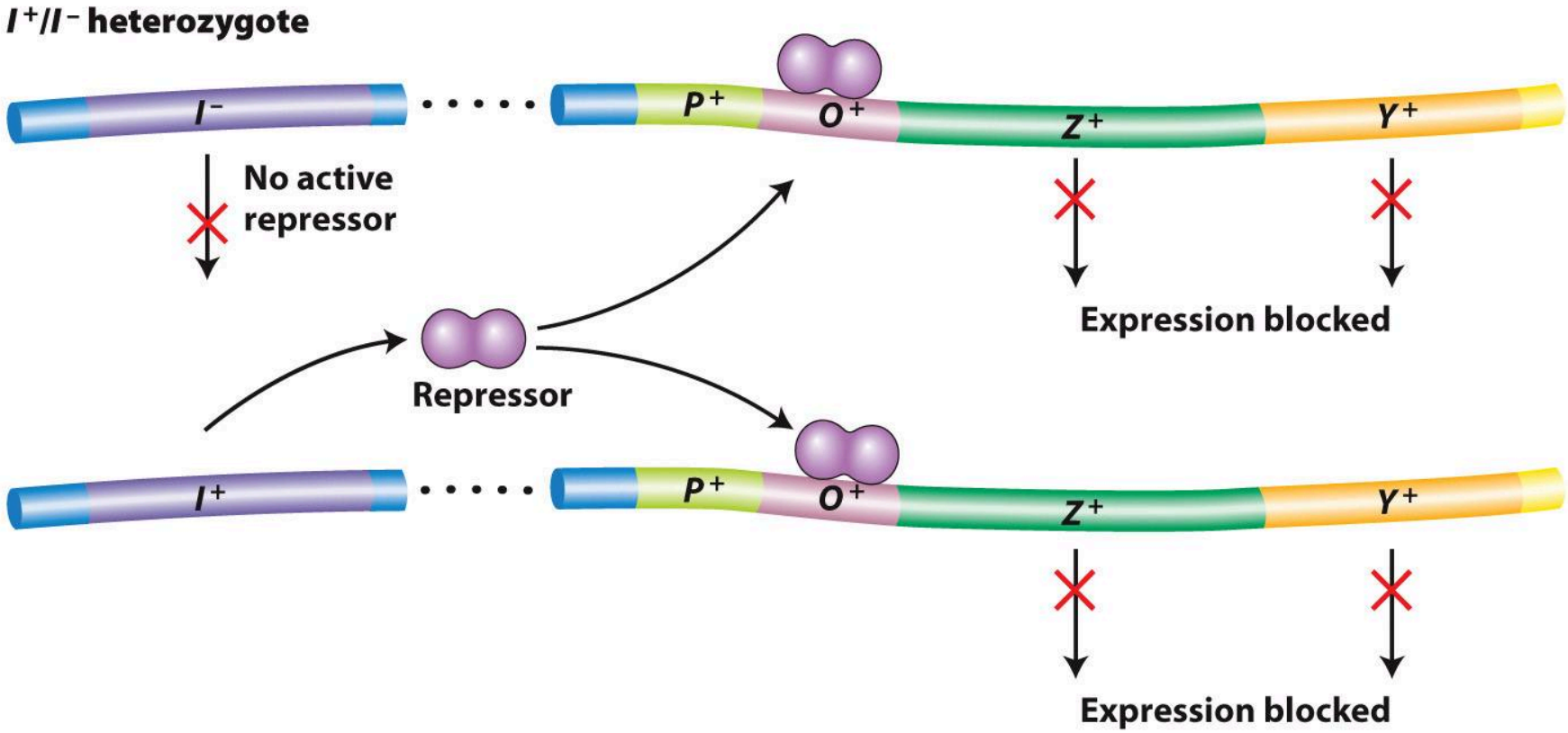


Figure 11-9
Introduction to Genetic Analysis, Twelfth Edition

Operators are cis-acting

TABLE 11-1 Synthesis of and Permease in Haploid and Heterozygous Diploid Operator Mutants

Strain	Genotype	β-*Galactosidase* (*Z*)		*Permease* (*Y*)		Conclusion
		Noninduced	Induced	Noninduced	Induced	
1	$O^+ Z^+ Y^+$	–	+	–	+	Wild type is inducible
2	$O^+ Z^+ Y^+ / F' O^+ Z^- Y^+$	+	+	–	+	Z^+ is dominant to Z^-
3	$O^C Z^+ Y^+$	+	+	+	+	O^C is constitutive
4	$O^+ Z^- Y^+ / F' O^C Z^+ Y^-$	+	+	–	+	Operator is cis-acting

Note: Bacteria were grown in glycerol (no glucose present) with and without the inducer IPTG. Expression of maximal enzyme levels is indicated by +. Absence or very low levels of enzyme activity is indicated by –. All strains are I^+.

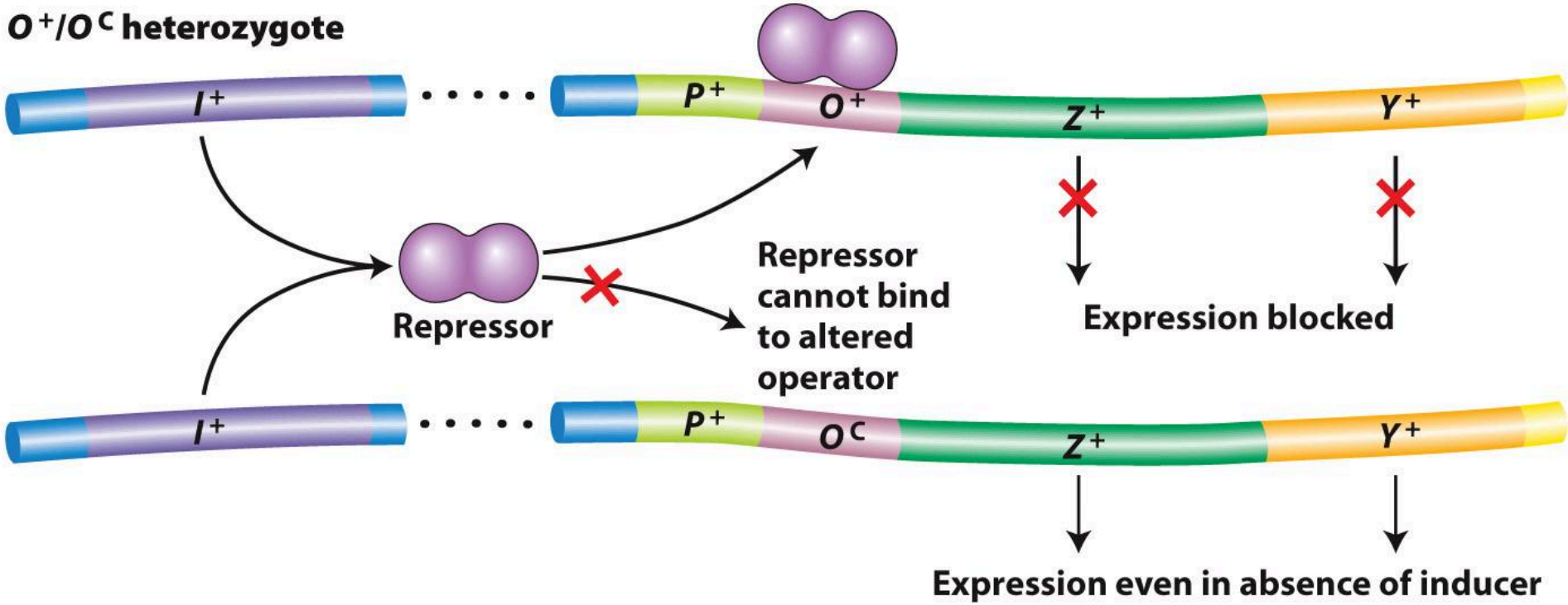


Figure 11-8
Introduction to Genetic Analysis, Twelfth Edition

The repressor contains an allosteric lactose-binding site

TABLE 11-3 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*)		*Permease* (*Y*)		Conclusion
		Noninduced	Induced	Noninduced	Induced	
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. Expression of maximal enzyme levels is indicated by +. Absence or very low levels of enzyme activity is indicated by –.

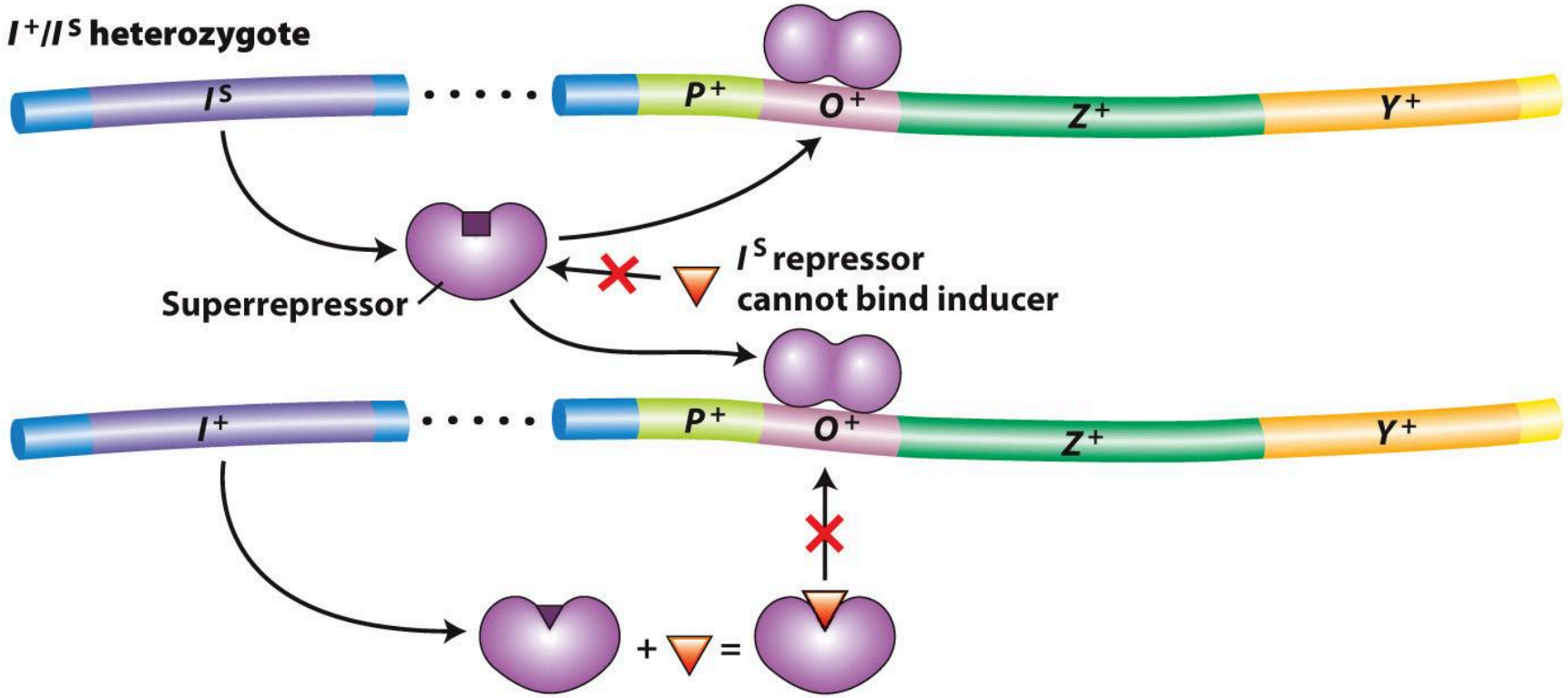


Figure 11-10
Introduction to Genetic Analysis, Twelfth Edition

The *lac* operon is transcribed only in the presence of lactose

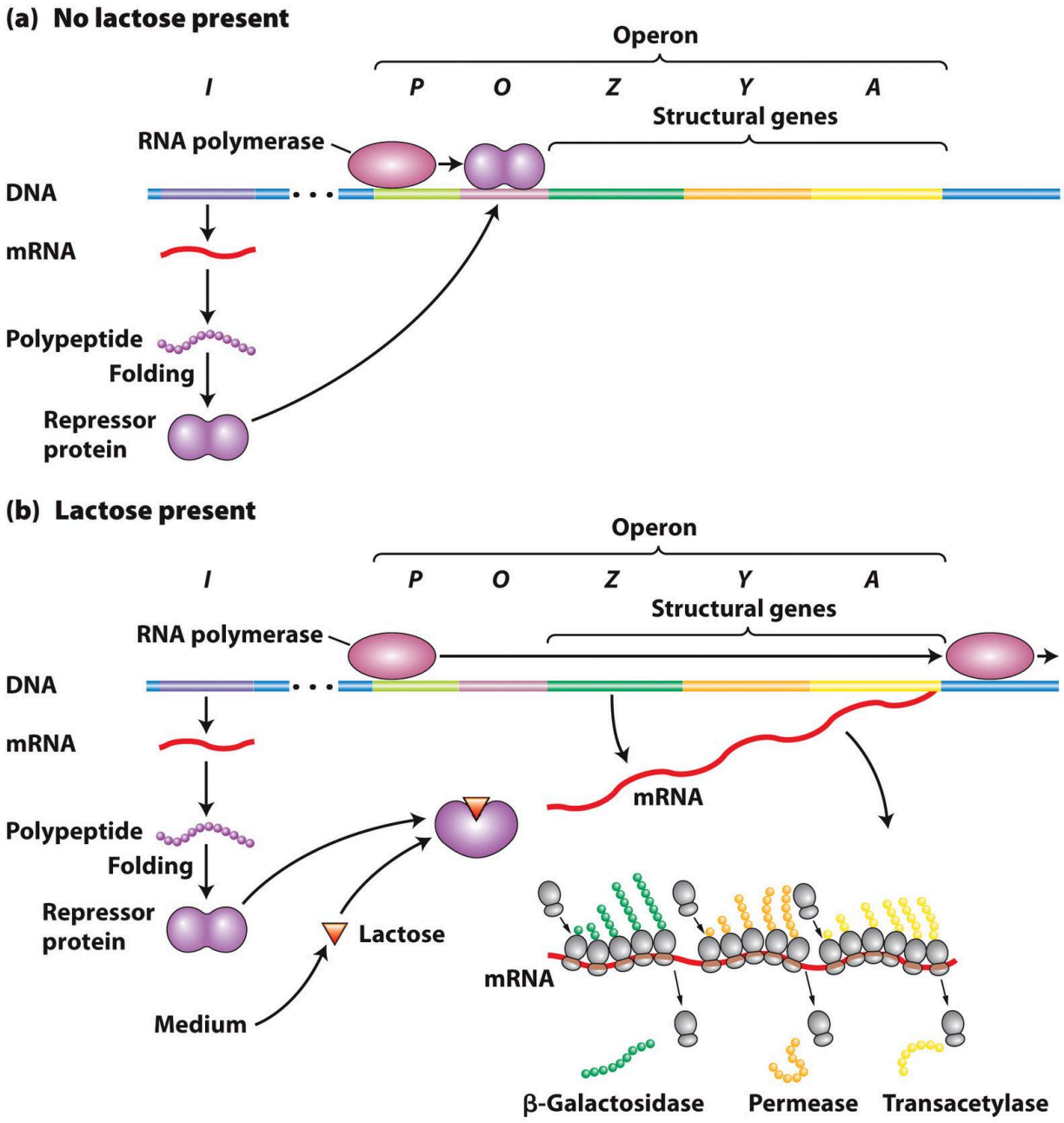


Figure 11-6
Introduction to Genetic Analysis, Twelfth Edition

Kinetics of growth in presence of glucose and galactose

growth

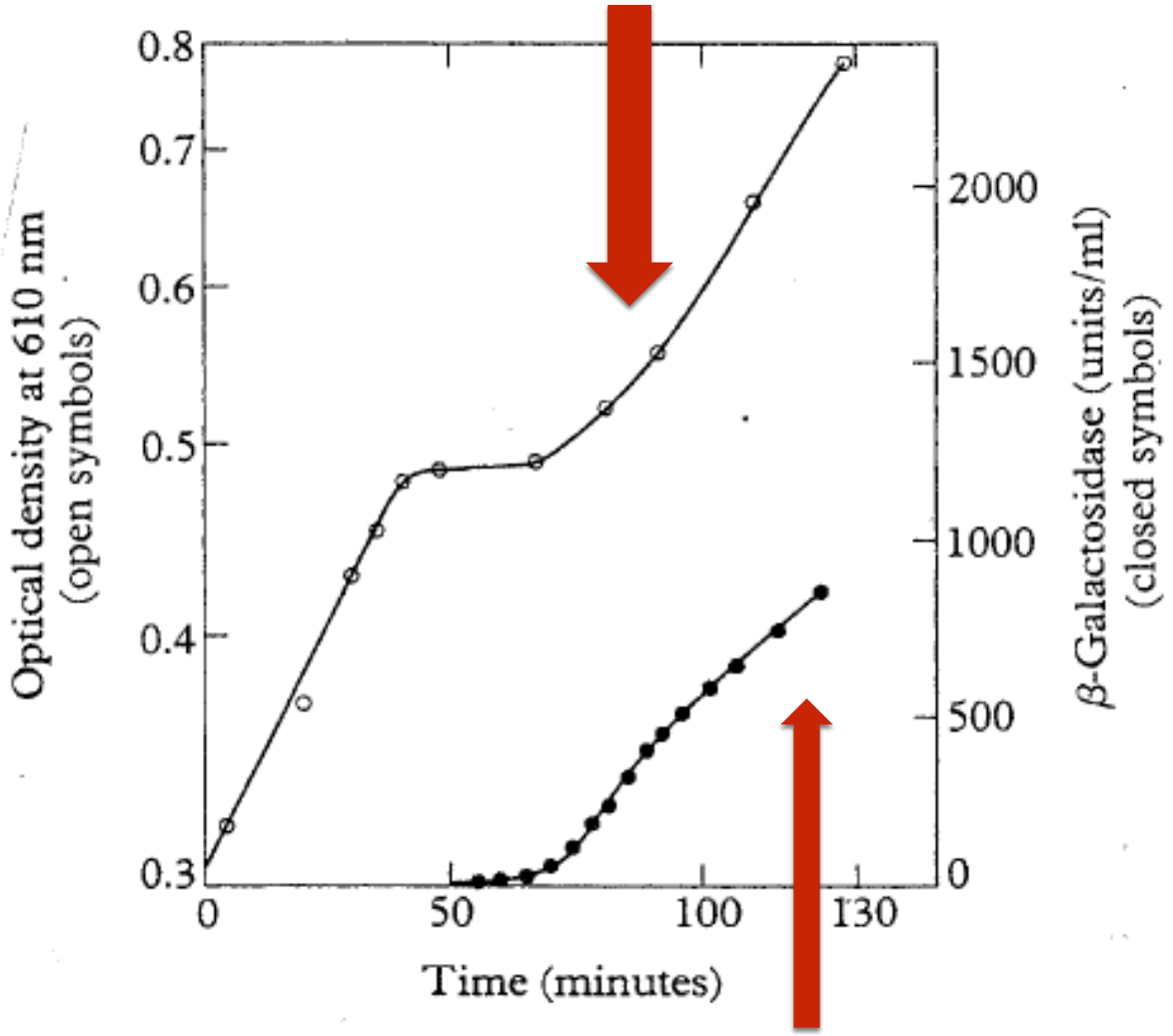


FIGURE 20-10
Delayed appearance of β-galactosidase in *E. coli* growing in a medium containing initially 0.4 mg/ml of glucose and 2 mg/ml of lactose. The left-hand ordinate indicates the cell density of the growing culture. [From W. Epstein, S. Naono, and F. Gros, *Biochem. Biophys. Res. Commun.* **24**, 588 (1966).]

Beta-galactosidase

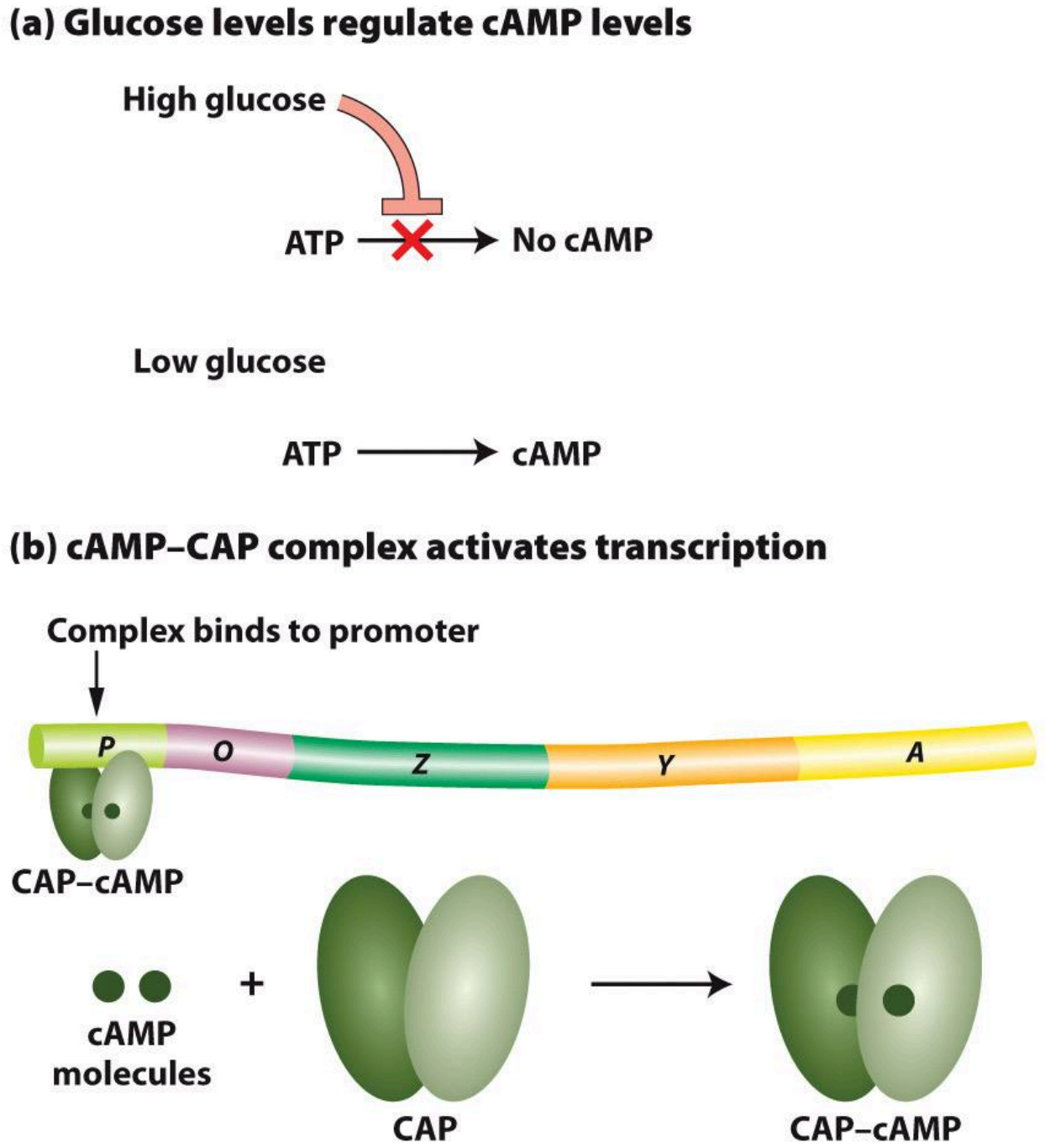


Figure 11-13
Introduction to Genetic Analysis, Twelfth Edition

The lac operon: synthesizing B-gal, but only when glucose is absent, and lactose is present

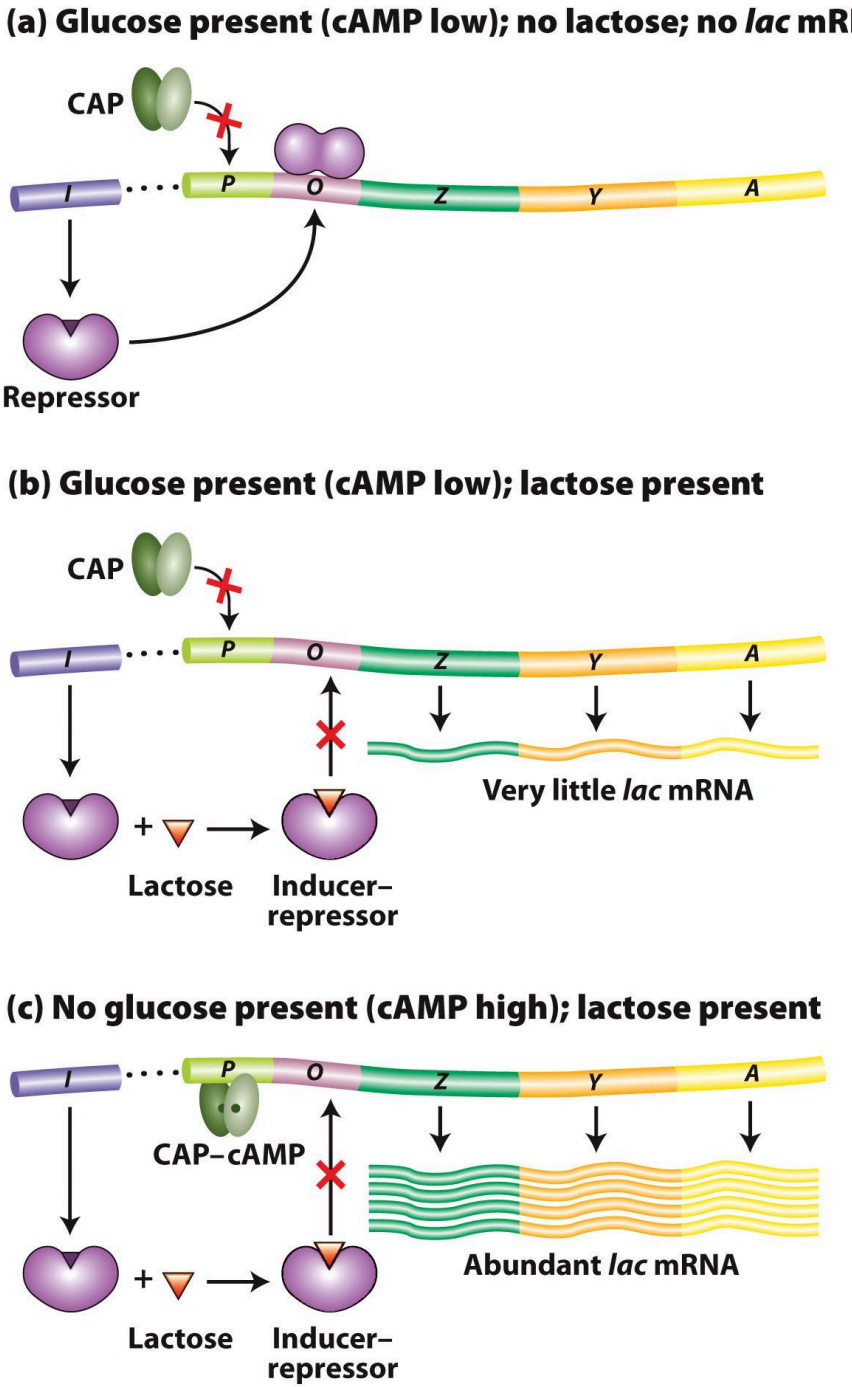
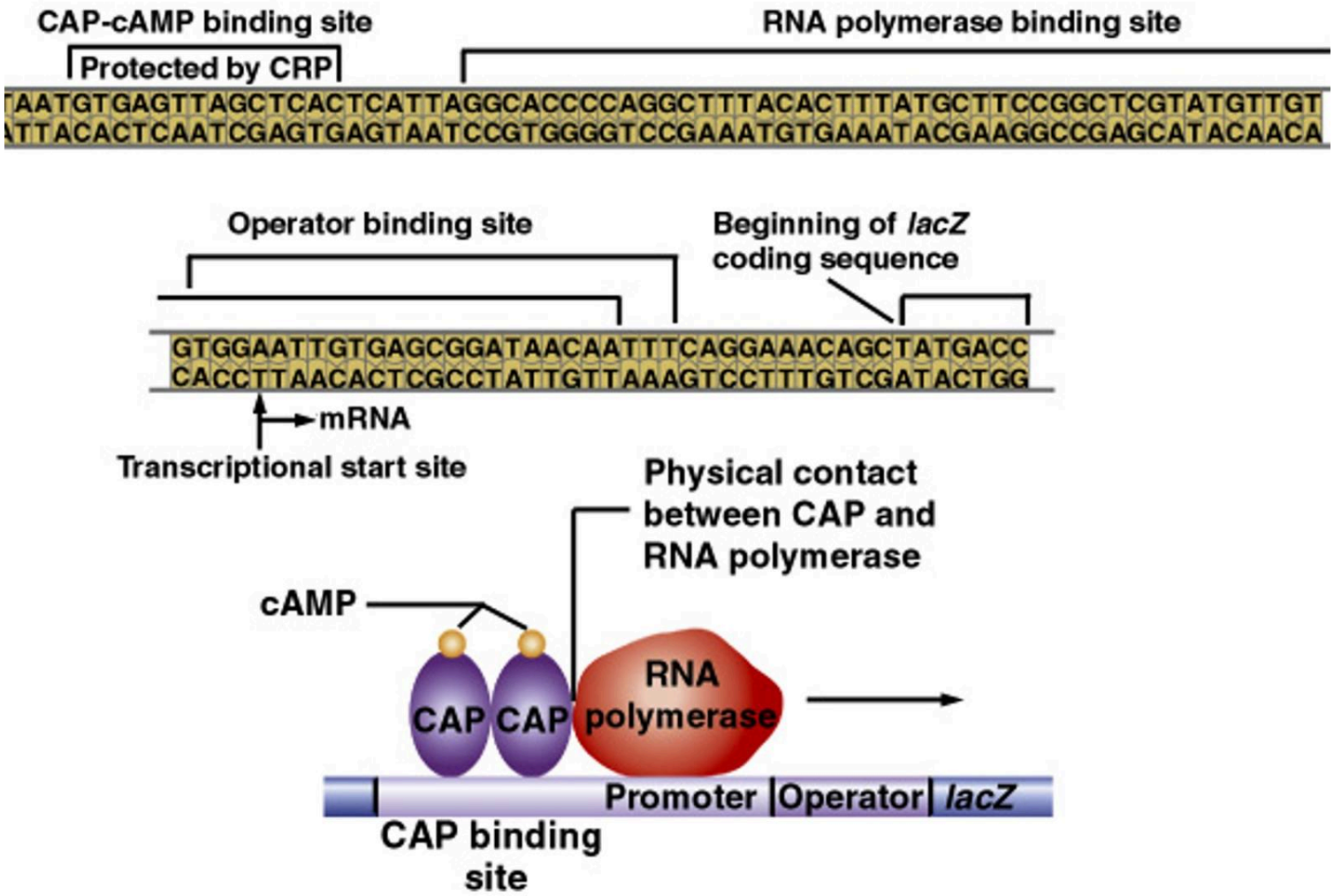


Figure 11-17
Introduction to Genetic Analysis, Twelfth Edition

The lac operon: synthesizing B-gal, but only when glucose is absent, and lactose is present



Regulatory proteins modulate the ability of RNA polymerase to carry out transcription. The lac repressor bound to the operator prevents RNA polymerase from binding. The binding sites for RNA polymerase and repressor show that there is overlap between the sites. The CAP-cAMP complex contacts RNA polymerase directly to help in transcription initiation. RNA polymerase can bind to a promoter in the absence of a positive regulator, but the positive regulator stabilizes this binding, effectively making it more likely that the polymerase will Initiate transcription rather than fall off.

Gene order in the trp operon corresponds to reaction order in the biosynthetic pathway

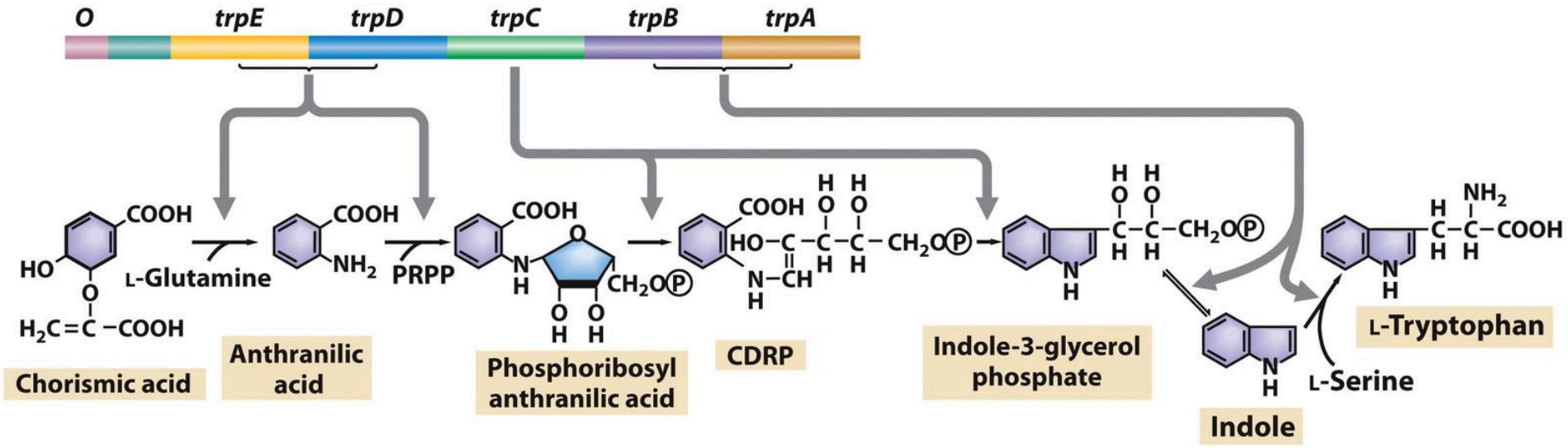


Figure 11-21
Introduction to Genetic Analysis, Twelfth Edition

The *trp* mRNA leader sequence contains an attenuator region and two tryptophan codons

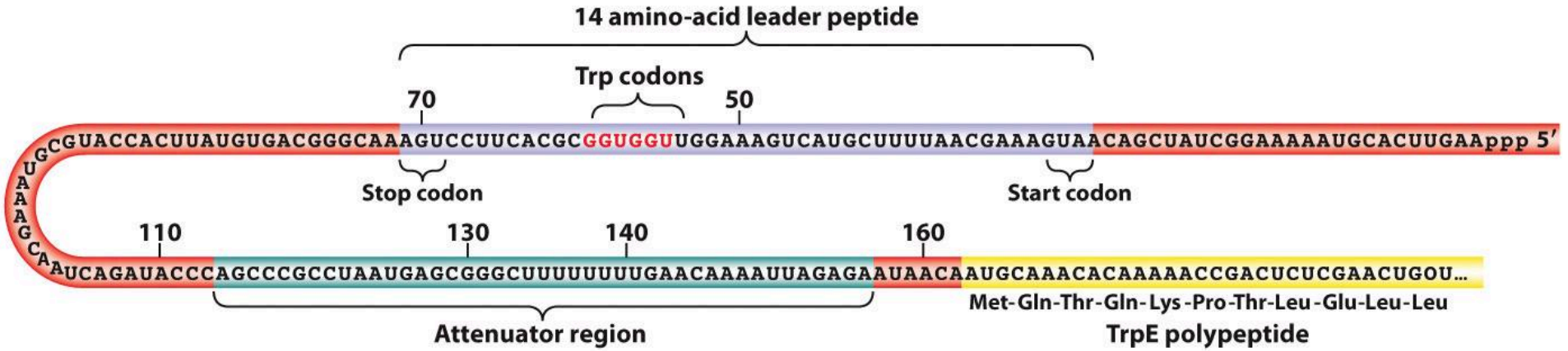


Figure 11-22
Introduction to Genetic Analysis, Twelfth Edition

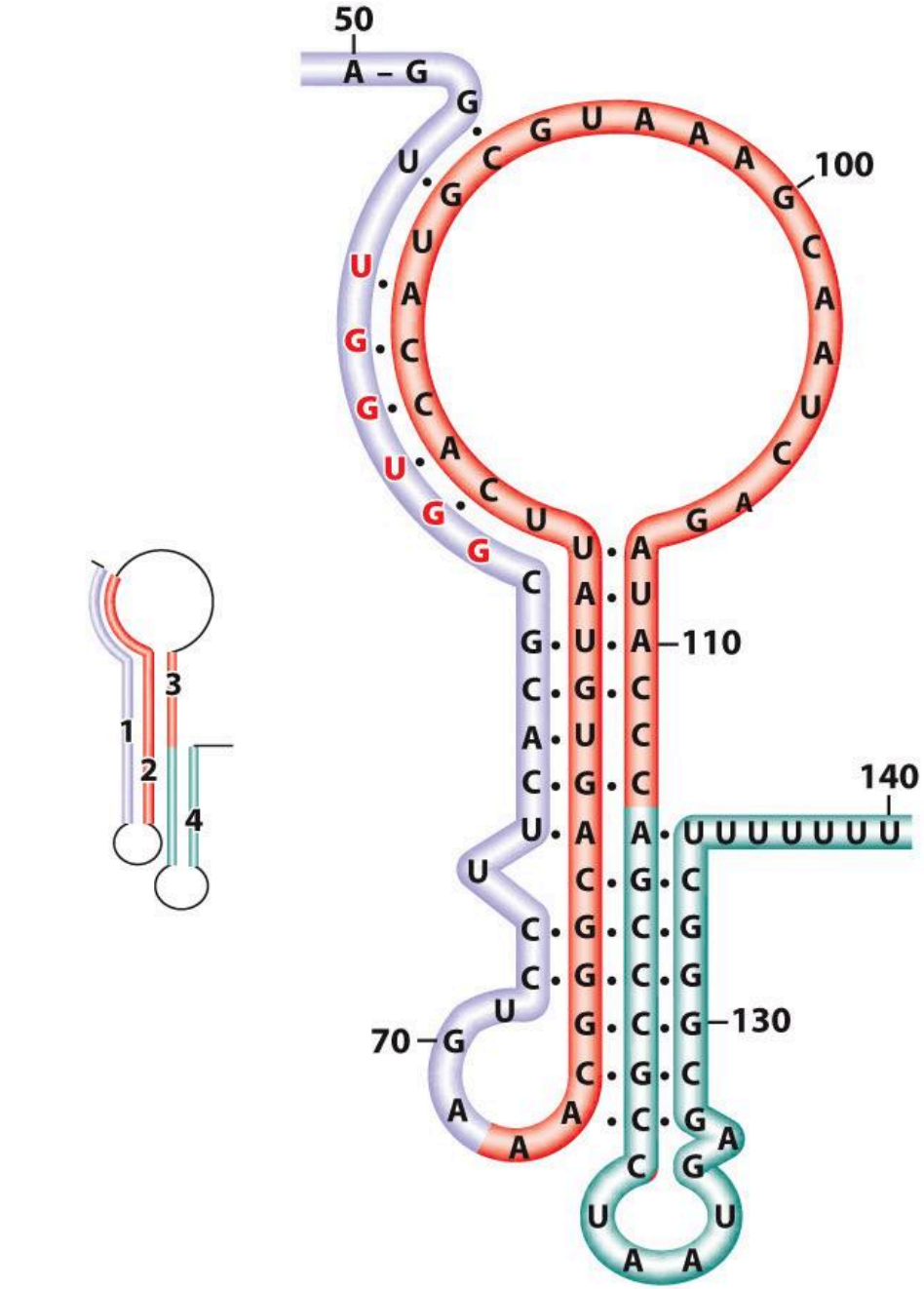
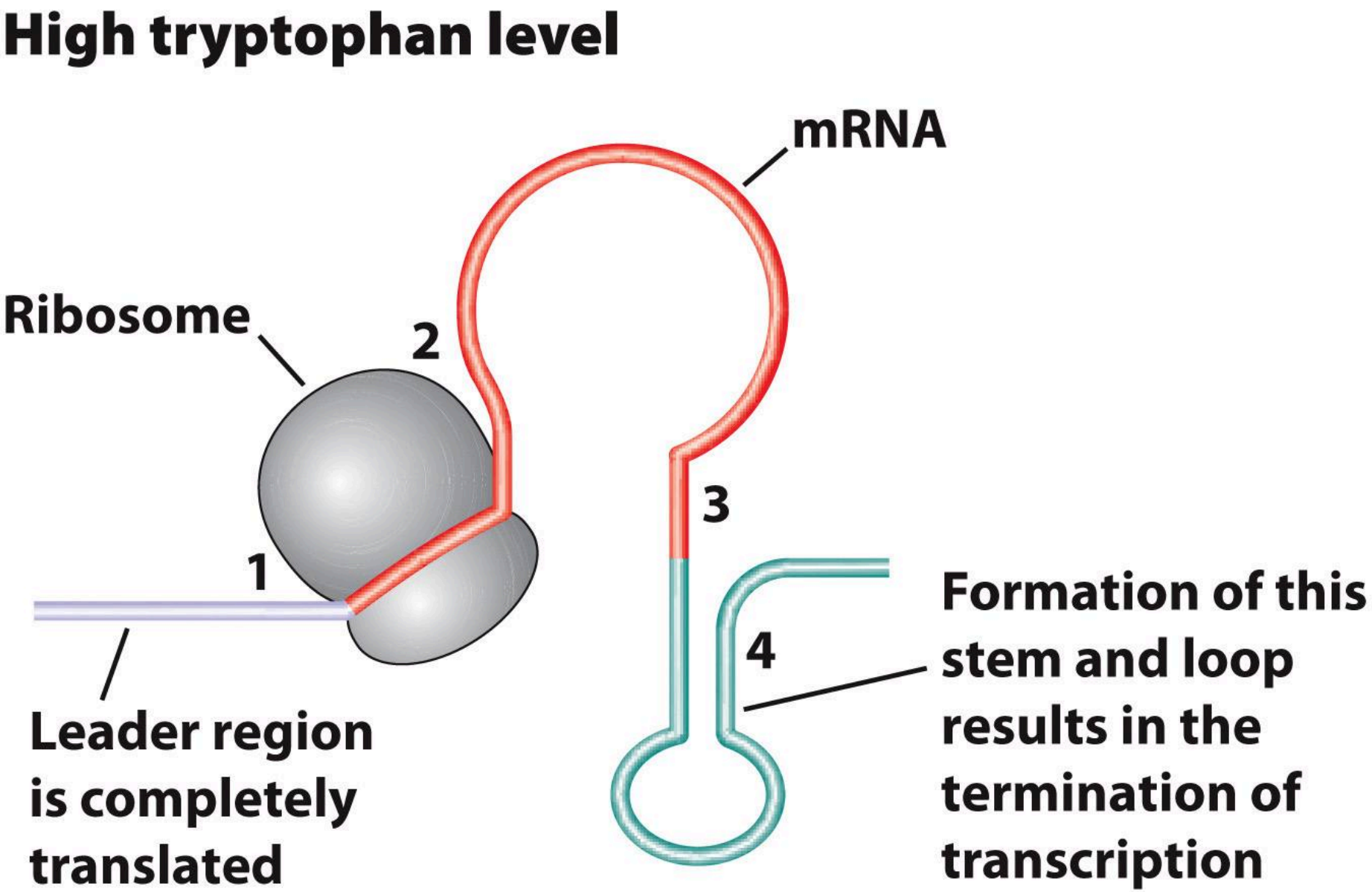
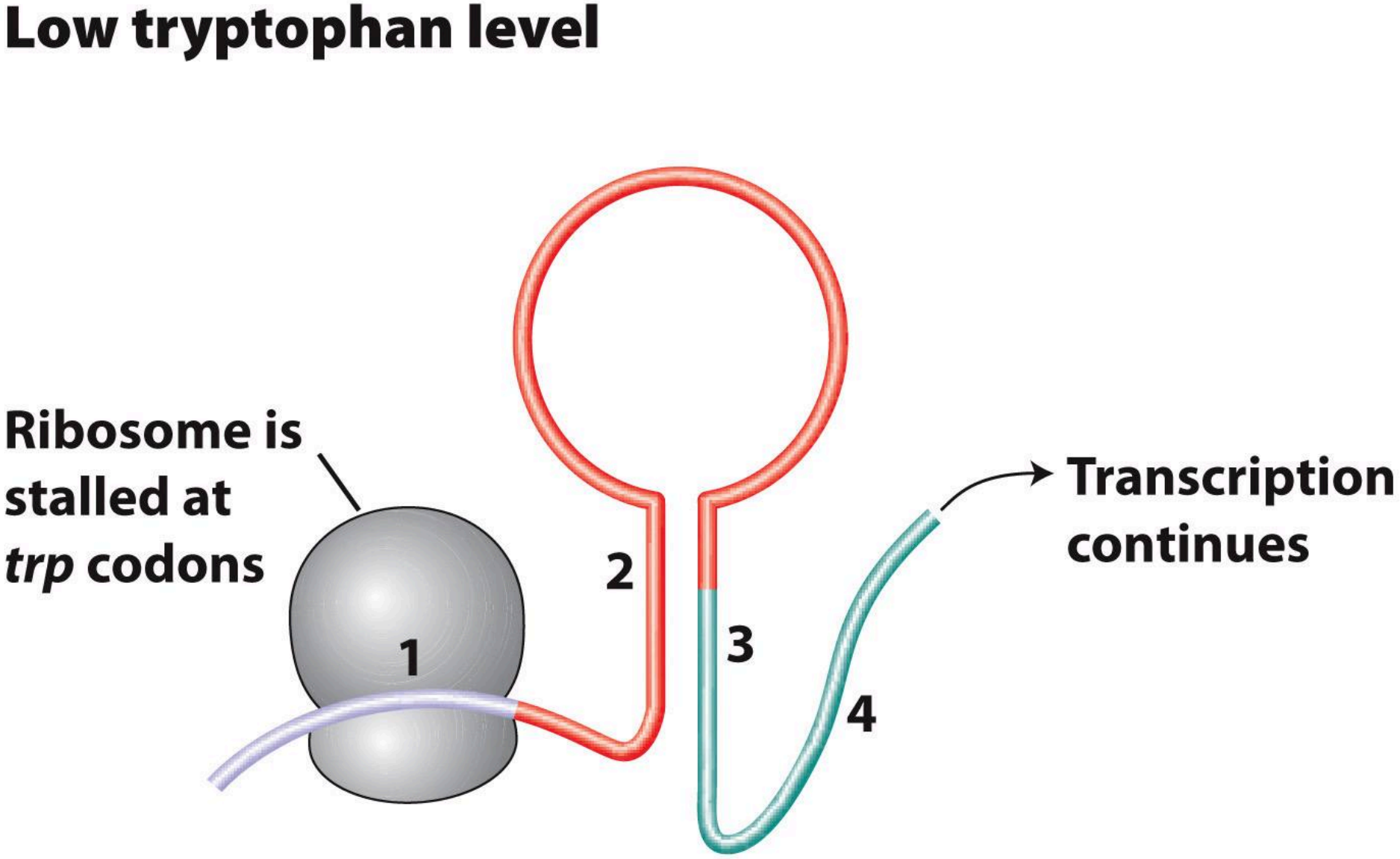


Figure 11-23a
Introduction to Genetic Analysis, Twelfth Edition

Abundant tryptophan attenuates transcription of the *trp* operon



Low tryptophan attenuates transcription of the *trp* operon



Data from D. L. Oxender, G. Zurawski, and C. Yanofsky, *Proc. Natl. Acad. Sci. USA* 76, 1979, 5524

Figure 11-23c
Introduction to Genetic Analysis, Twelfth Edition

Cis-acting mutants can have diverse phenotypes

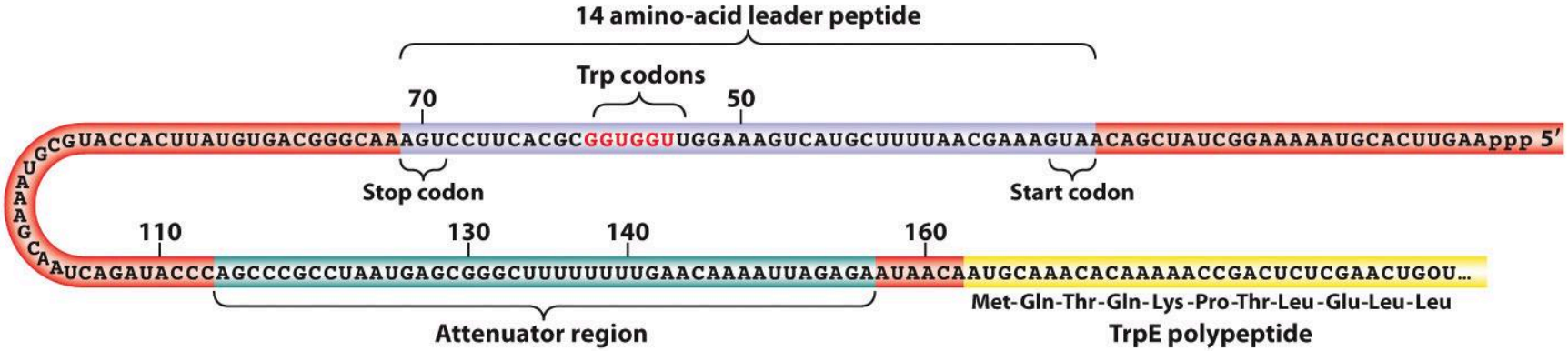
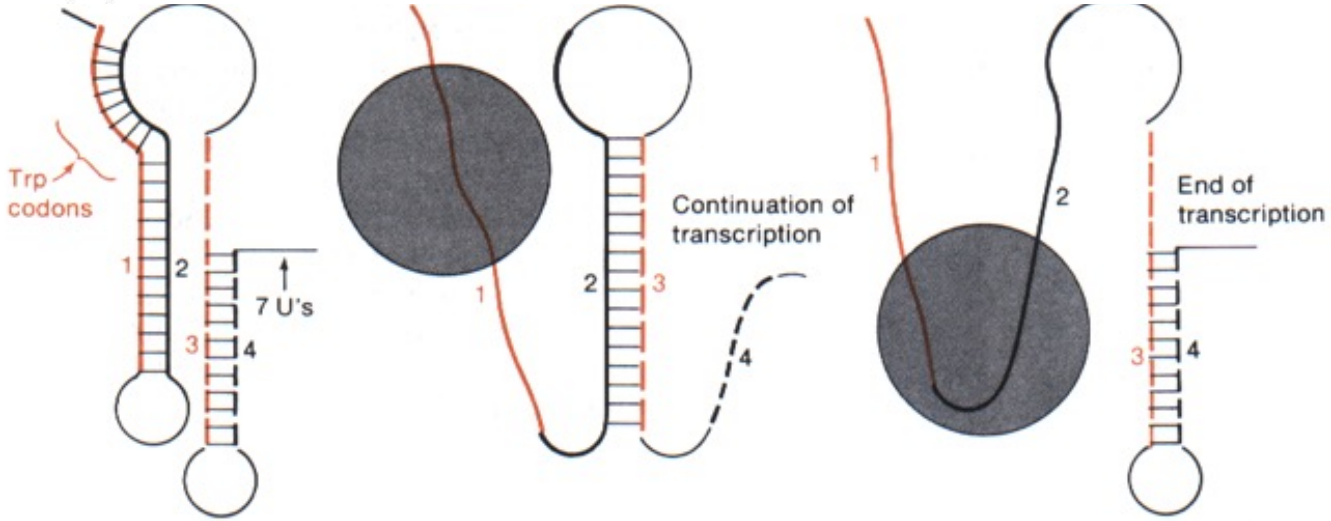


Figure 11-22
Introduction to Genetic Analysis, Twelfth Edition



(a) Free mRNA. Base pairs between 1 and 2 and between 3 and 4.

(b) Low concentration of tryptophan. Ribosome stalled in region 1 permits formation of 2–3 before transcription of region 4 is completed.

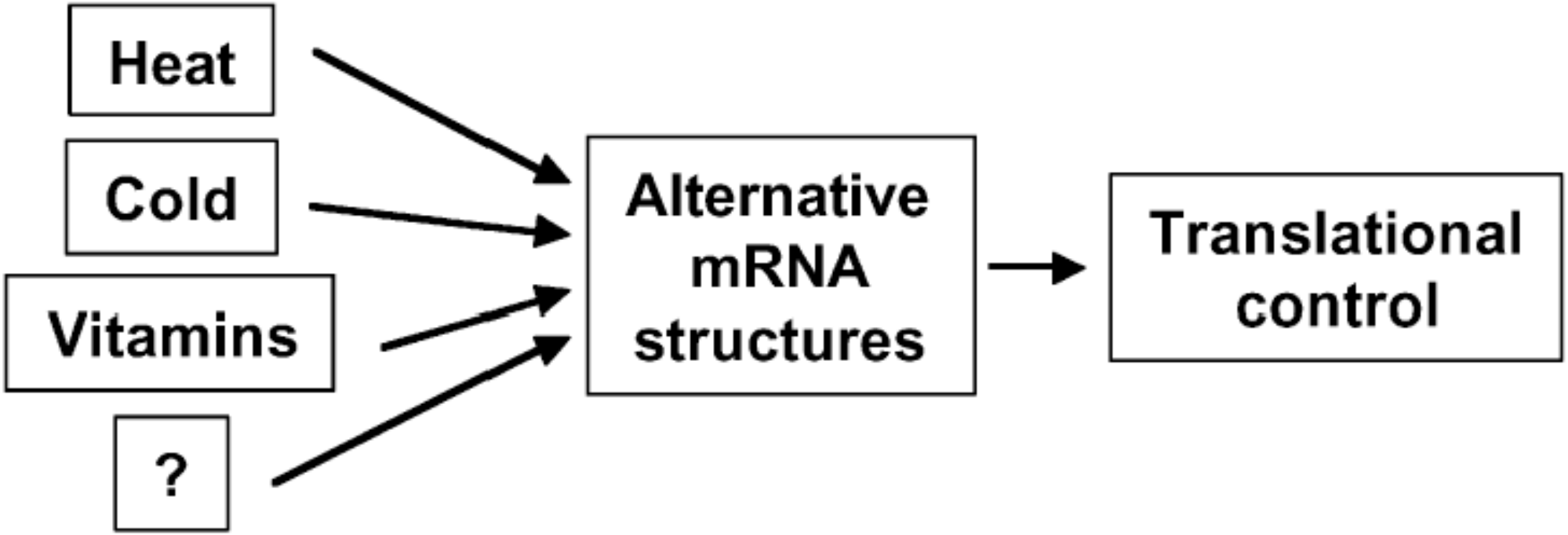
(c) High concentration of tryptophan. Ribosome reaches region 2 before region 4 is completed, and permits formation of 3–4.

Constitutively active mutants

1. Mutations that eliminate basepairing in regions 3 and 4
2. Deletion of the attenuator sequence

Uninducible mutants = an increase in termination

1. Destabilize pairing between regions 2 and 3
2. Mutation of the ATG of the leader peptide


Heat
Cold
Vitamins
?
Alternative mRNA structures
Translational control

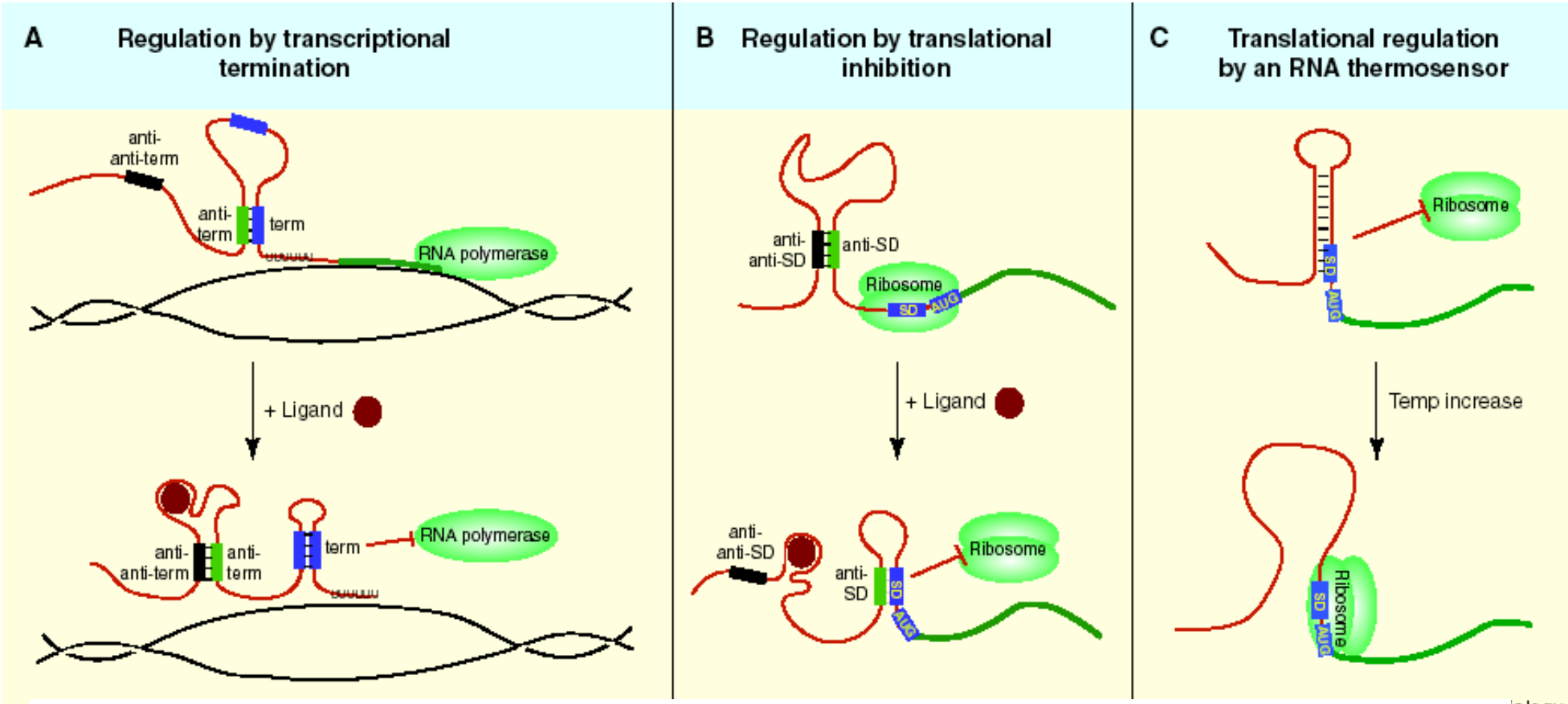
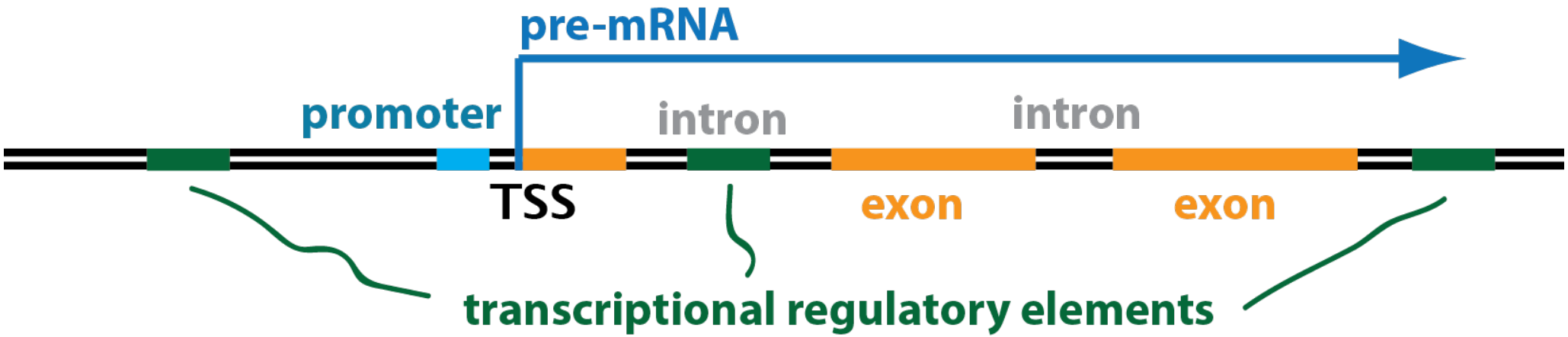


Figure 1. Diverse strategies for gene regulation linked to RNA sensors.

Small molecule-regulated riboswitches have been linked to both transcriptional attenuation (A) and translational inhibition (B) mechanisms, while characterized thermosensors use a translational inhibition strategy (C). (A) Transcriptional termination mechanism. DNA is depicted as a black double helix, the nascent mRNA transcript as a colored line (red is untranslated leader, green is coding sequence), and RNA polymerase as a green oval. (Top) In the absence of the relevant metabolite, part of the terminator (two blue boxes) is bound by an anti-terminator (green box) and is non-functional. Transcription proceeds past the poly-uridine stretch following the terminator sequence and extends into coding sequence. (Bottom) In the presence of the relevant metabolite (red circle), ligand-bound RNA adopts an alternative conformation in which the anti-terminator is bound by an anti-anti-terminator (black box). This allows formation of the structurally distinctive terminator hairpin, which induces release of RNA polymerase following the poly-uridine tract. (B) Translational inhibition mechanism. The mRNA is depicted with the colored line; untranslated leader in red and the coding region in green. (Top) In the absence of cofactor-binding, the anti-Shine-Dalgarno (anti-SD) ribosome binding sequence is sequestered by an anti-anti-SD sequence; this conformation allows ribosome (double green oval) binding to the SD box and translation to occur. (Bottom) Cofactor binding (red circle) restructures the RNA so that the anti-SD sequence is allowed to pair with the SD box, thus inhibiting translation. Note that variations on both mechanisms exist, as may dual regulatory mechanisms involving both transcriptional attenuation and translational inhibition (see text). (C) A thermosensor from *L. monocytogenes prfA*. (Top) A stem structure adjacent to the SD box prevents translation at lower temperatures. (Bottom) The increase in ambient temperature following host infection melts this structure, allowing the ribosome to access the SD box and translate *prfA*.

Eukaryote



A simple gene. TSS, transcriptional start site.

DNA transcription regulatory sequences and factors

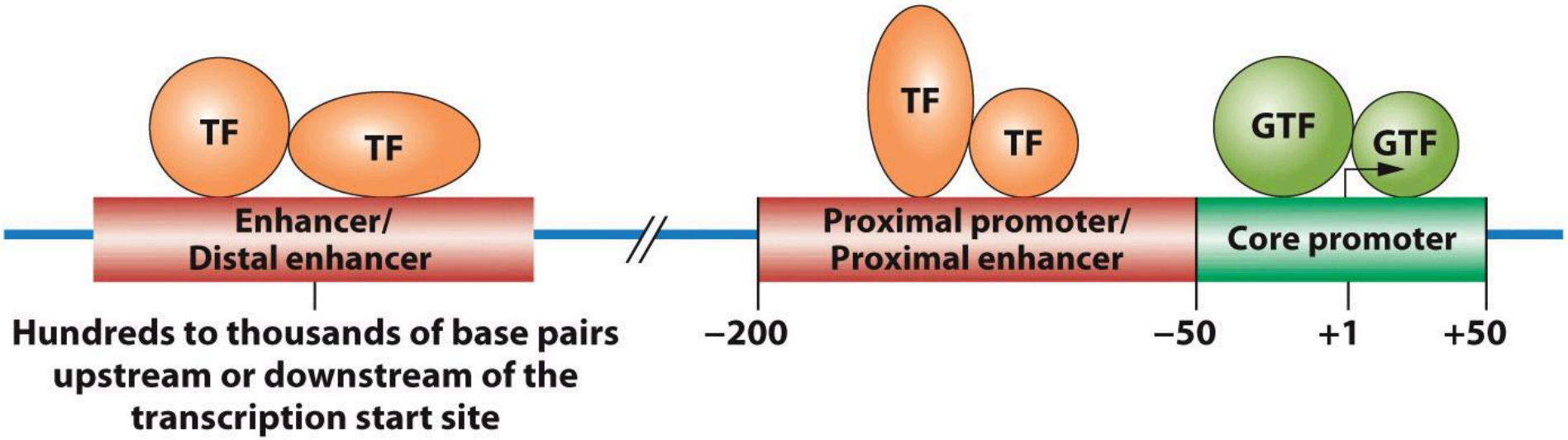


Figure 12-1
Introduction to Genetic Analysis, Twelfth Edition

Coactivators bridge interactions between transcription factors and RNA polymerase II

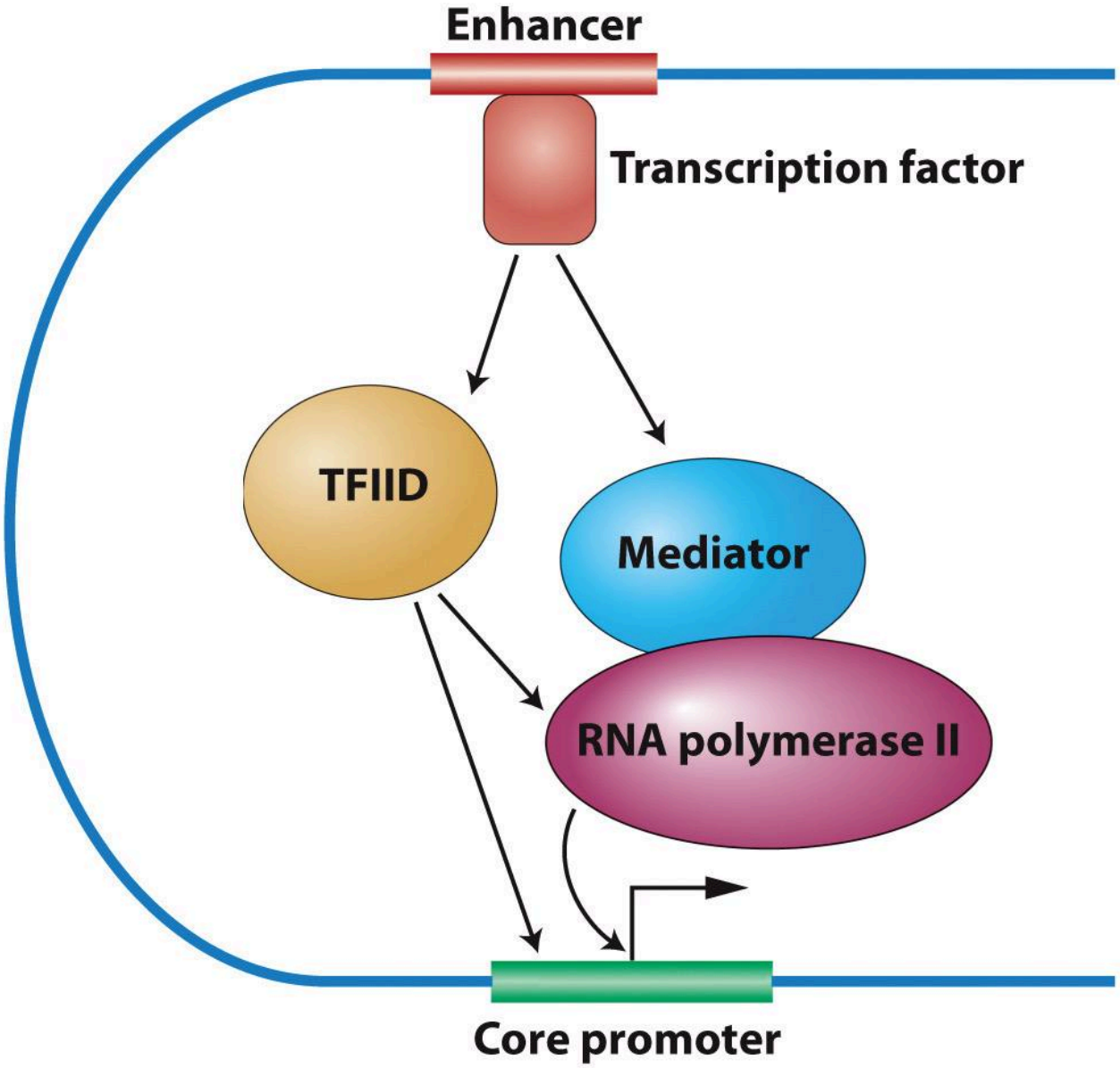
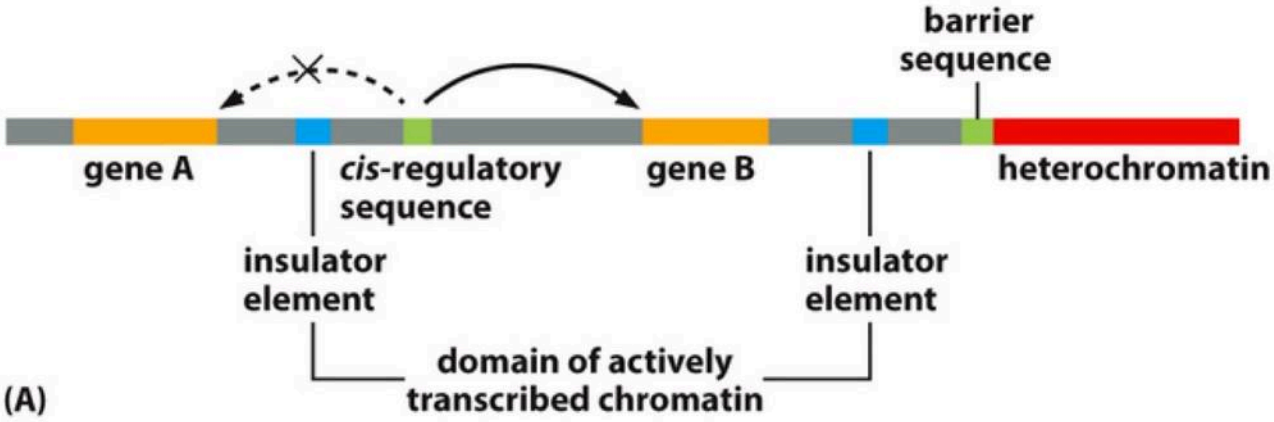


Figure 12-2
Introduction to Genetic Analysis, Twelfth Edition

More cis regulatory sequences and trans acting factors



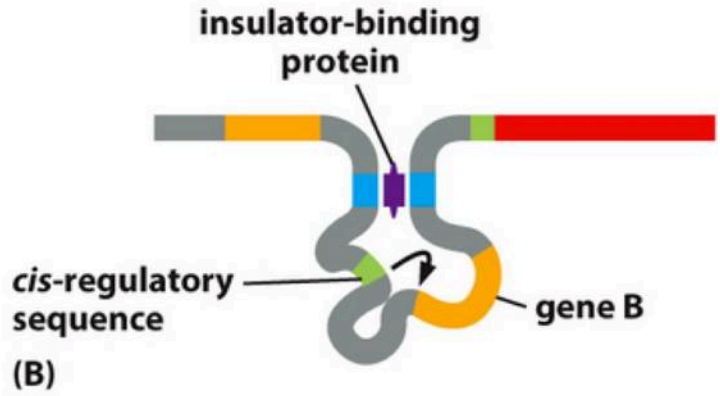
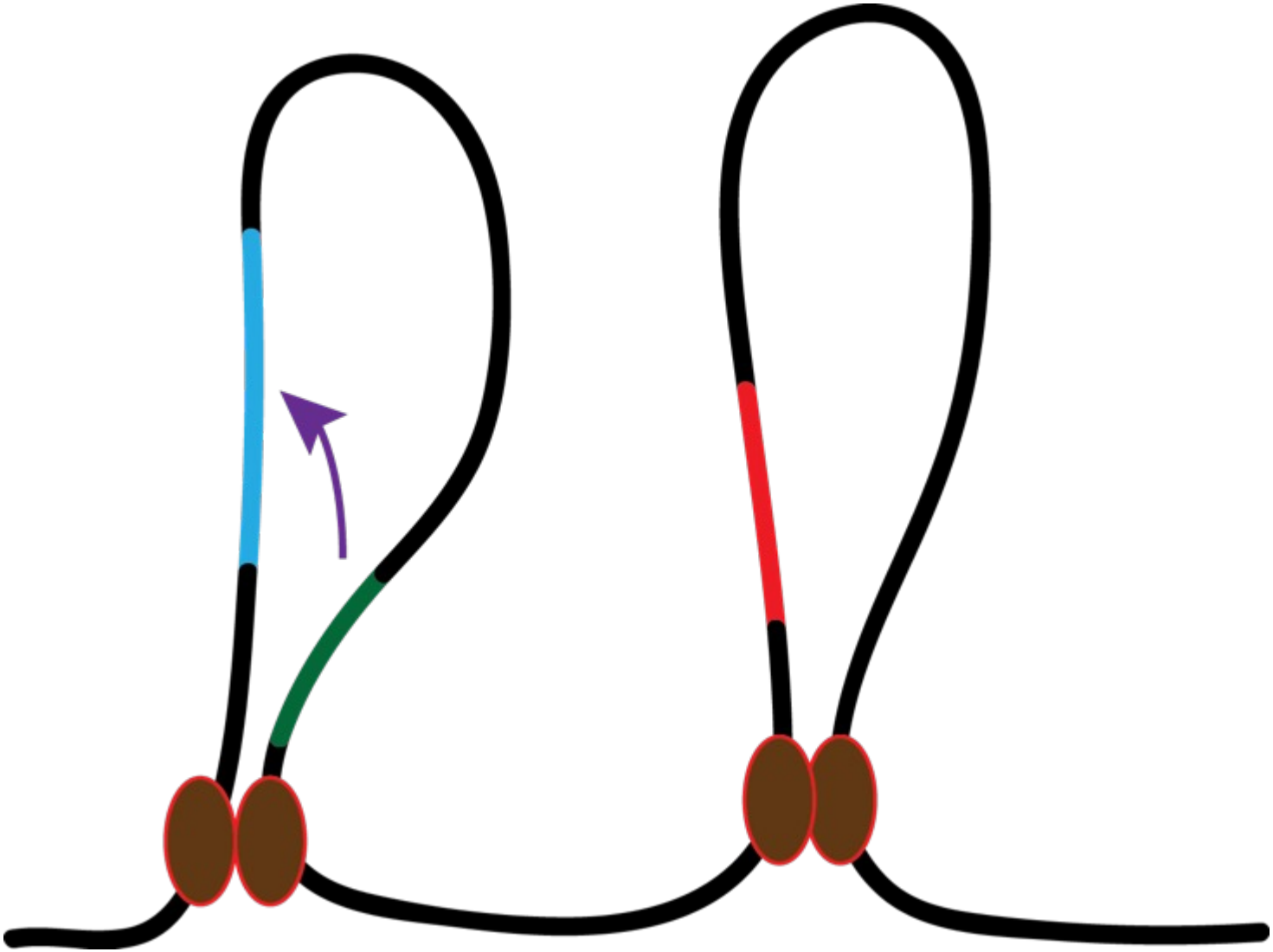
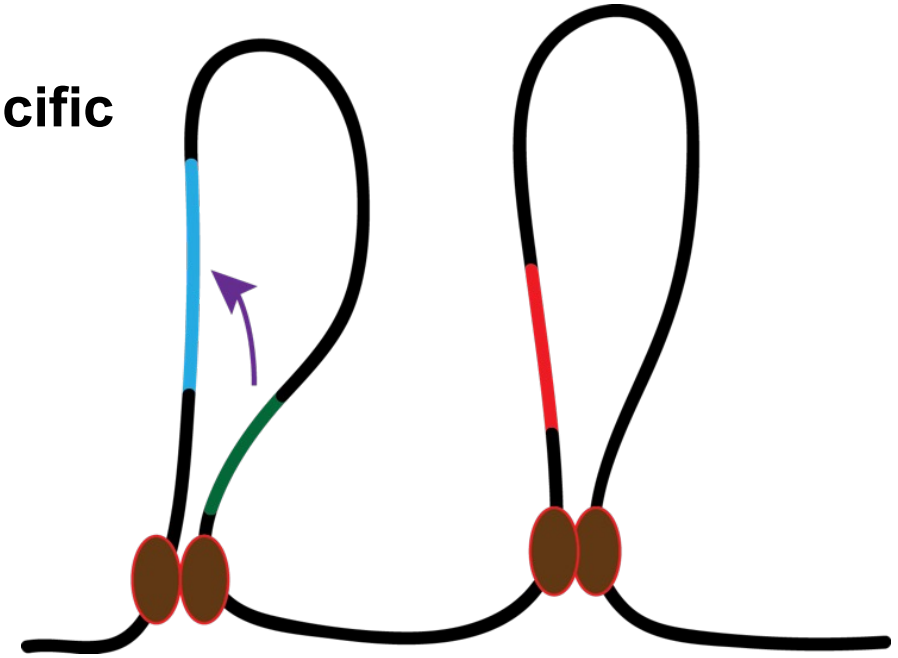


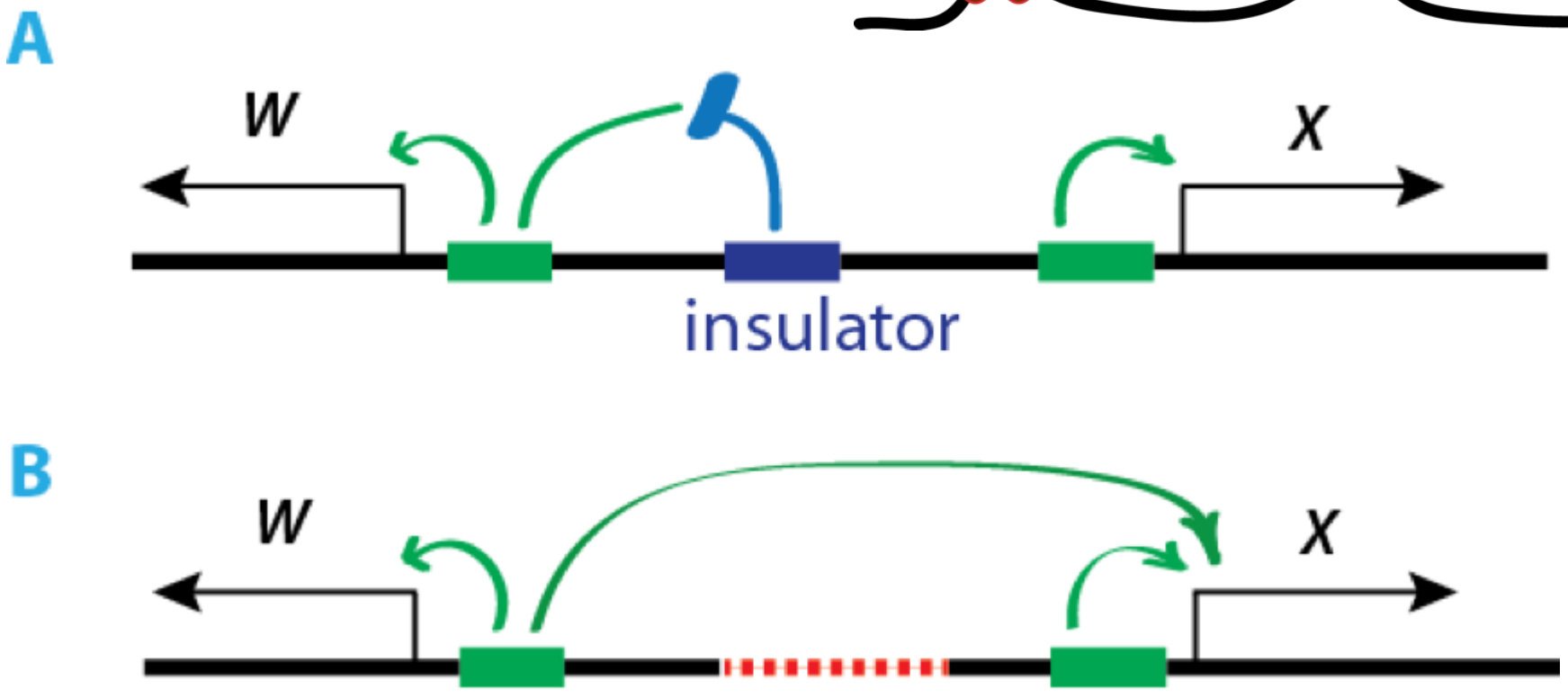
Figure 7-24 Molecular Biology of the Cell 6e (© Garland Science 2015)

Interactions between sequences (via their bound factors) occurs within loop regions, but not between loops



Insulators (sequences bound by specific factors) prevent inappropriate enhancer-promoter interactions



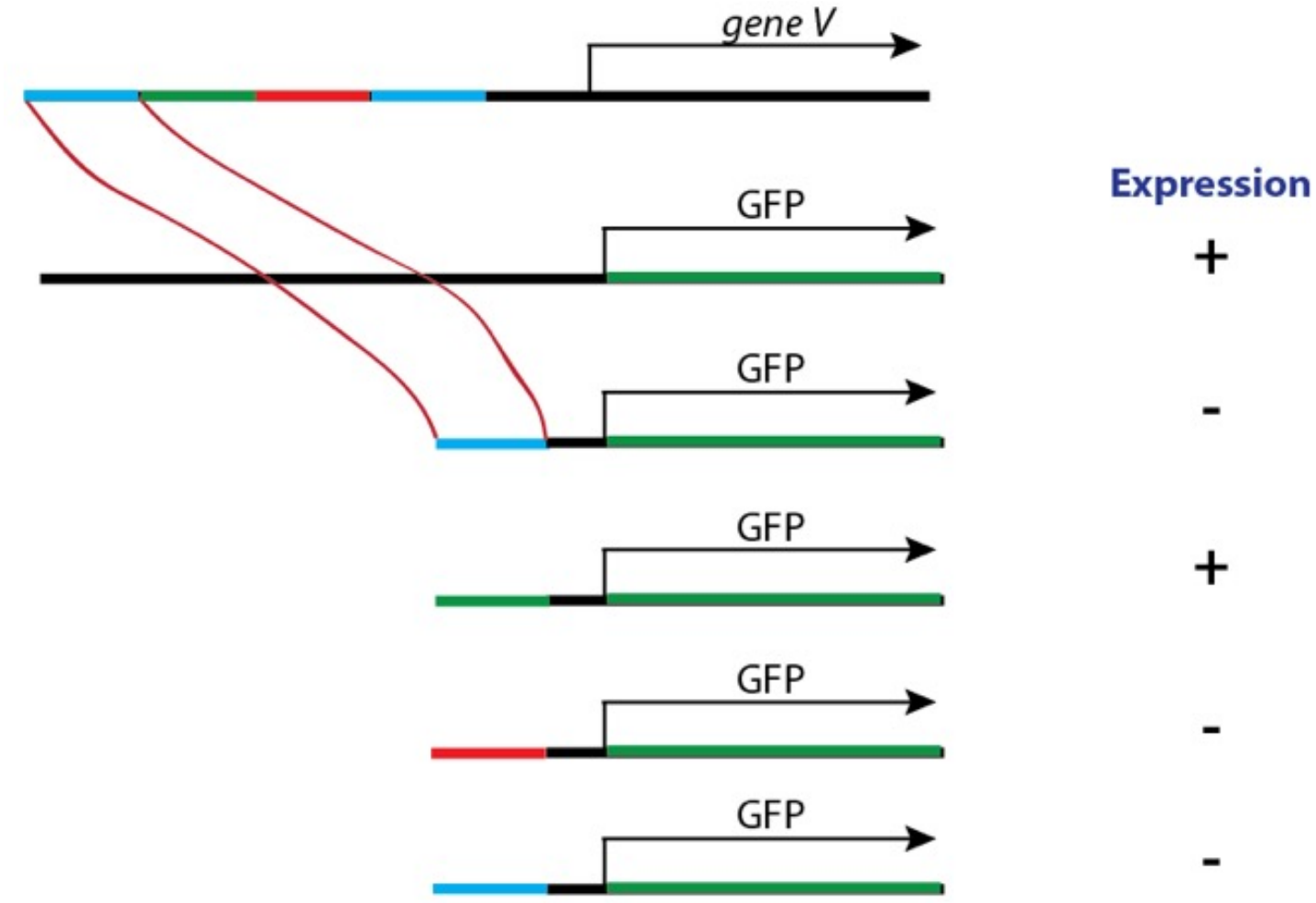


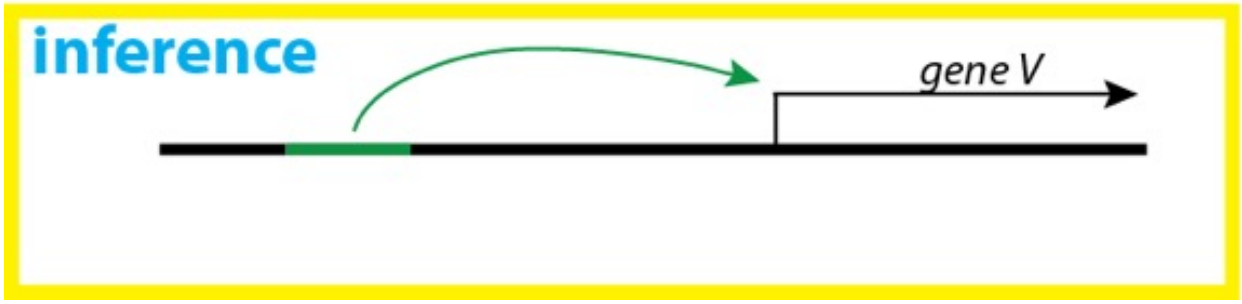
A cell can decode regulatory sequences; why can't we?

we want to predict gene expression and regulation from DNA sequence

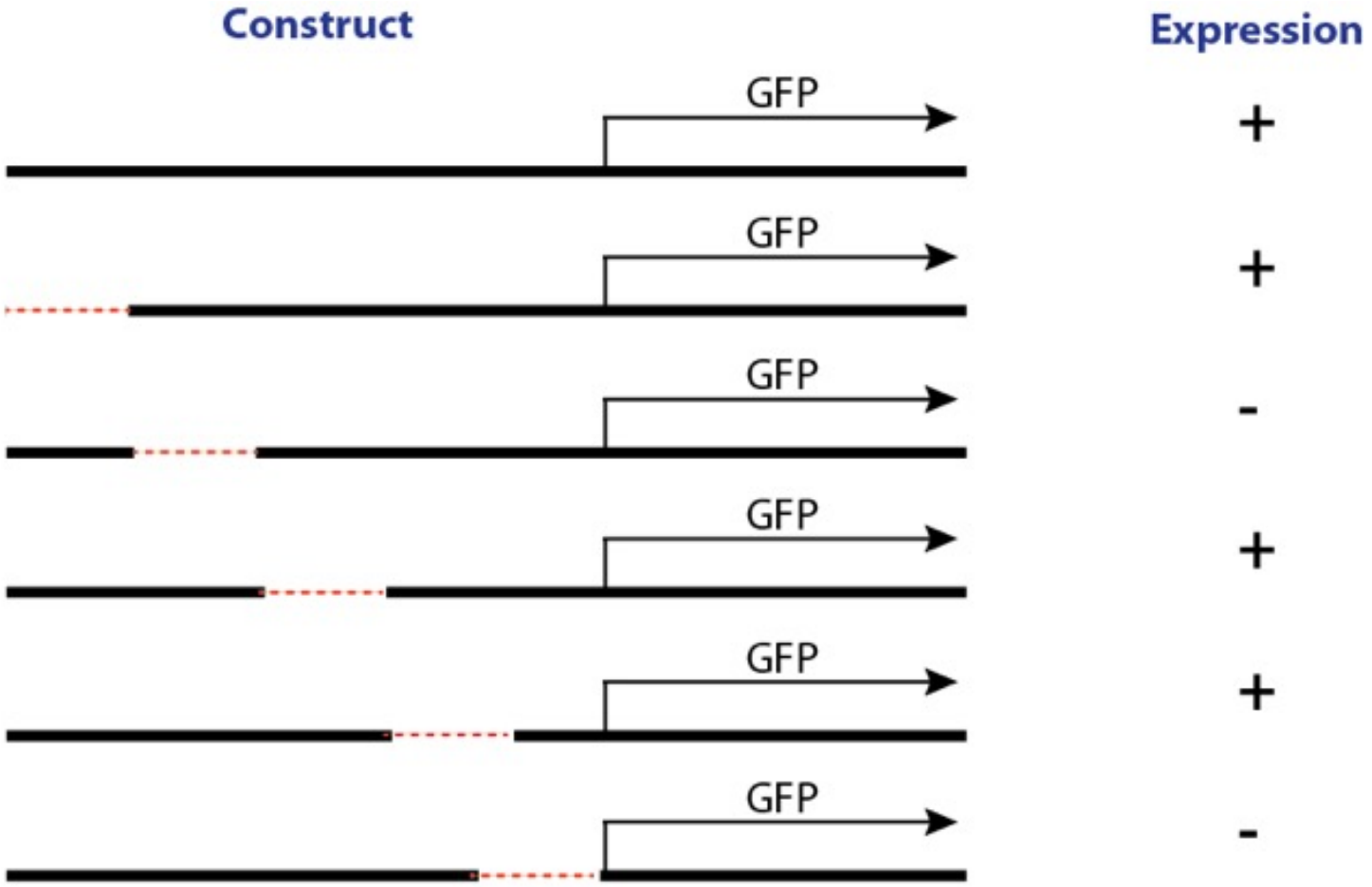
What is the consequence of a non-coding variant in your genome?

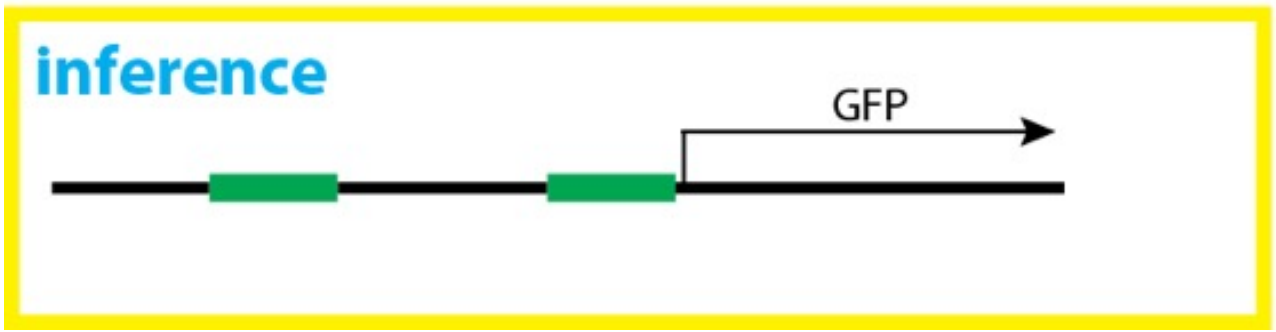
Sufficiency test with transgenics. Keep in mind these each mediate their biology through trans-acting factors



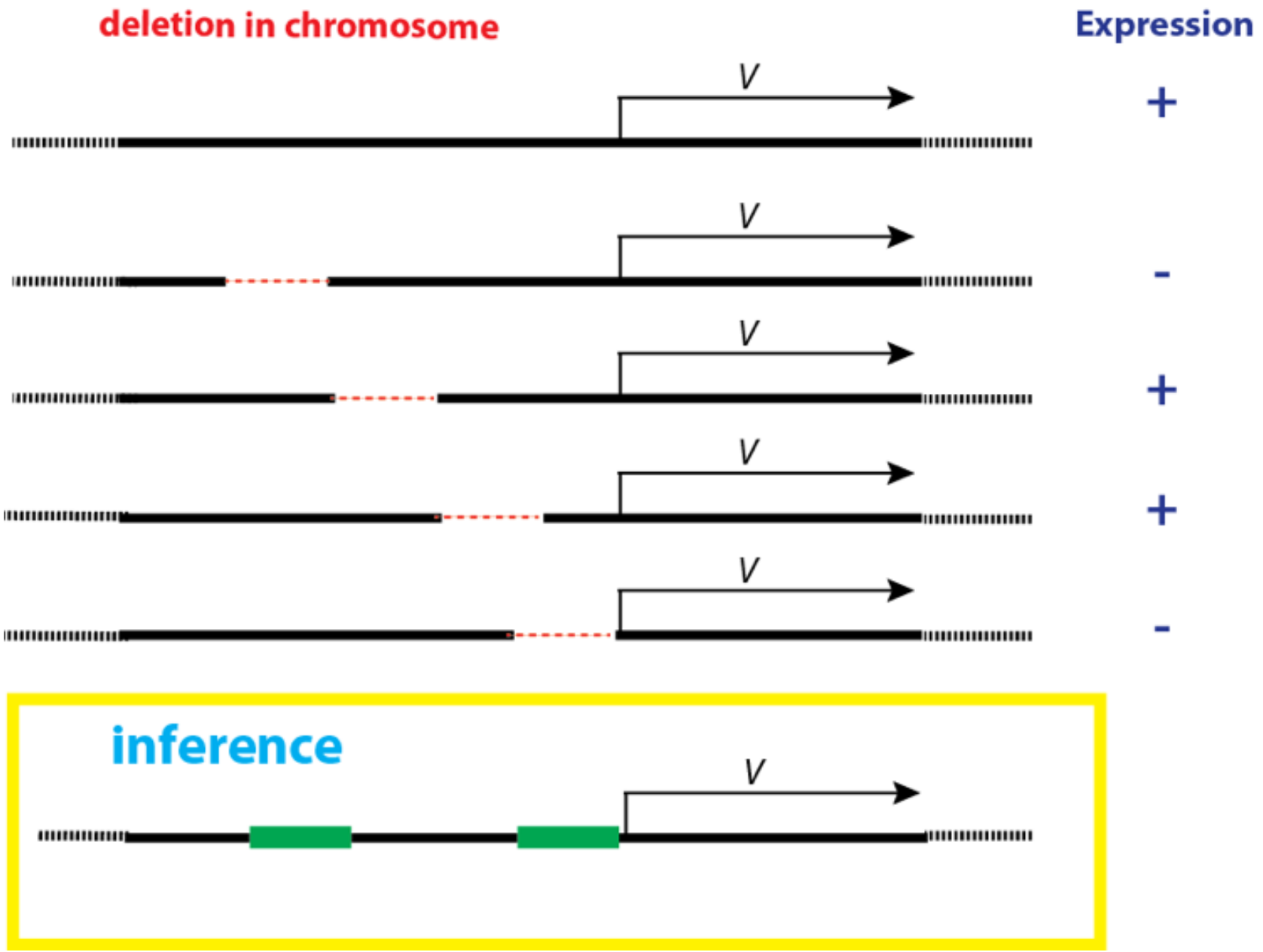


Necessity test with transgenics



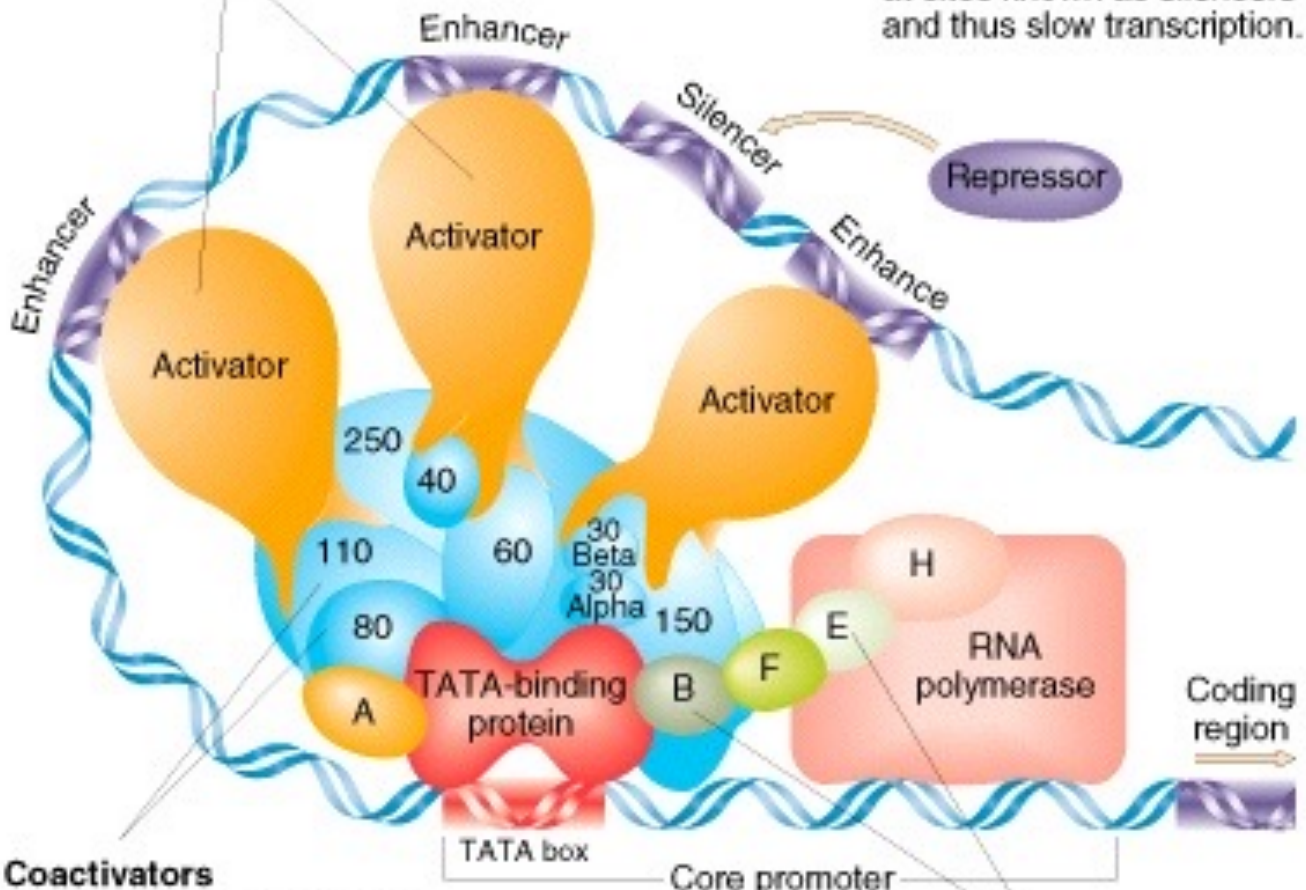


Can also use gene targeting to test necessity and sufficiency at endogenous locus



Activators
These proteins bind to genes at sites known as *enhancers* and speed the rate of transcription.

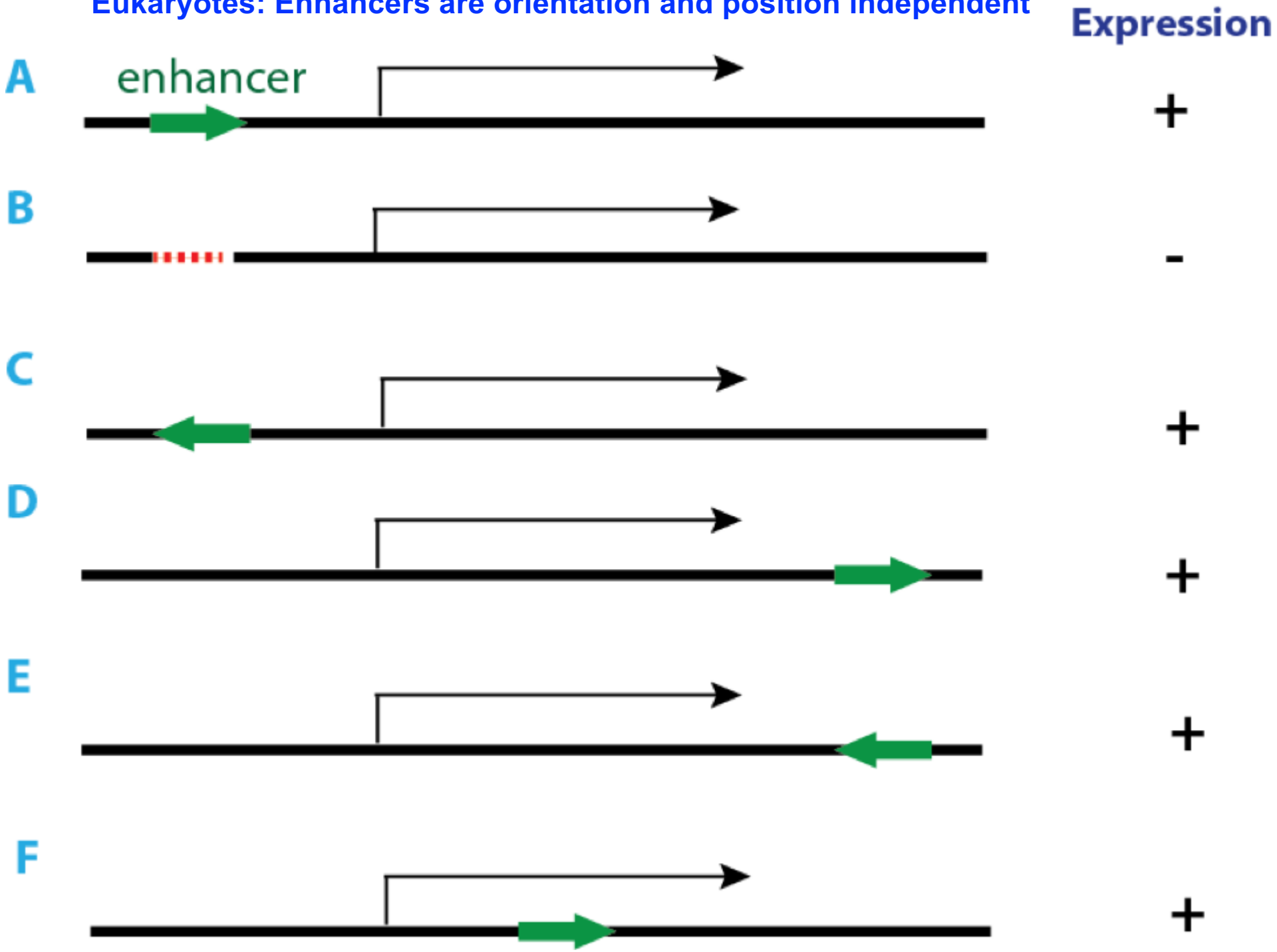
Repressors
These proteins bind to selected sets of genes at sites known as *silencers* and thus slow transcription.



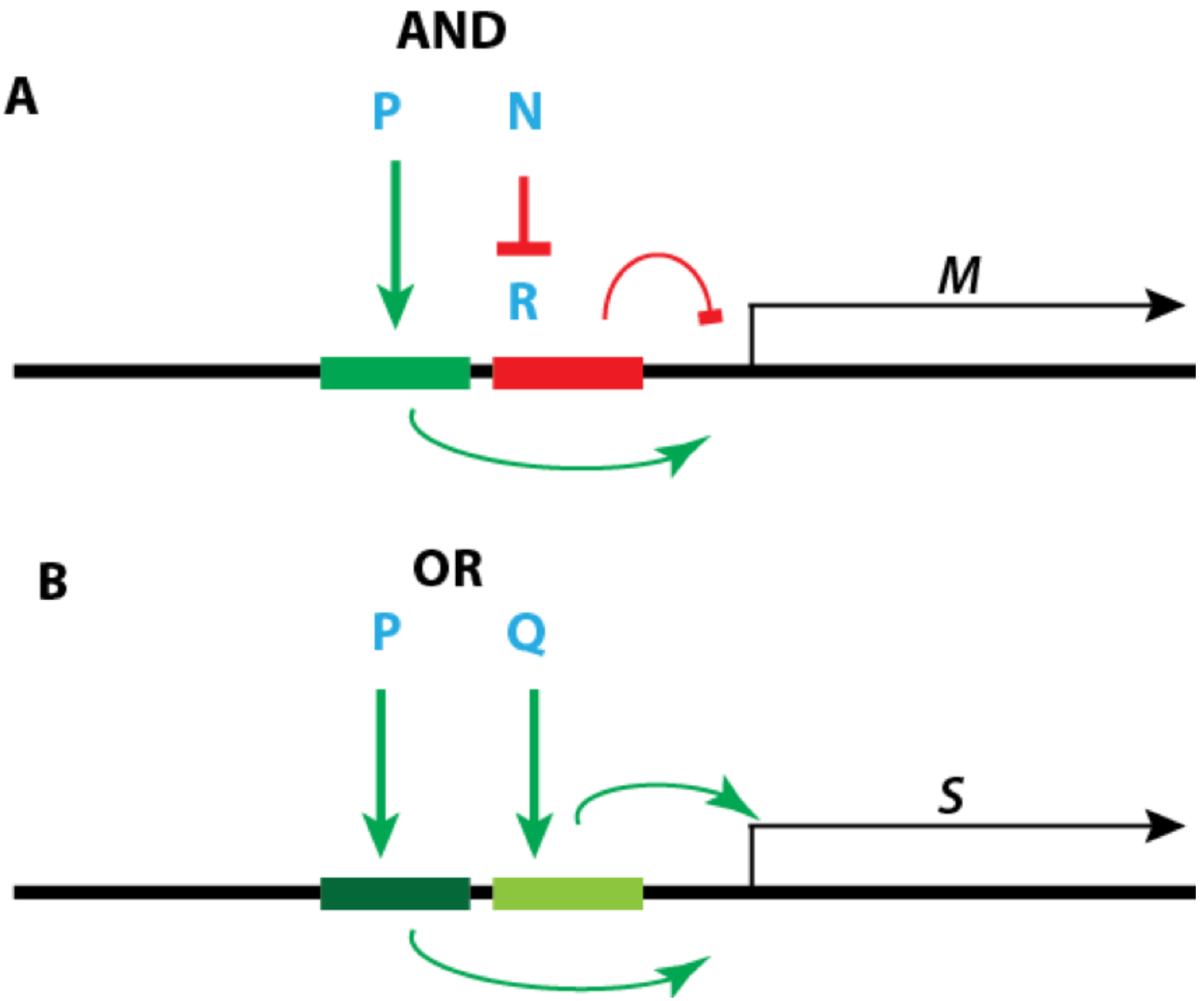
Coactivators
These "adapter" molecules integrate signals from activators and perhaps repressors.

Basal transcription factors
In response to injunctions from activators, these factors position RNA polymerase at the start of transcription and initiate the transcription process.

Eukaryotes: Enhancers are orientation and position independent

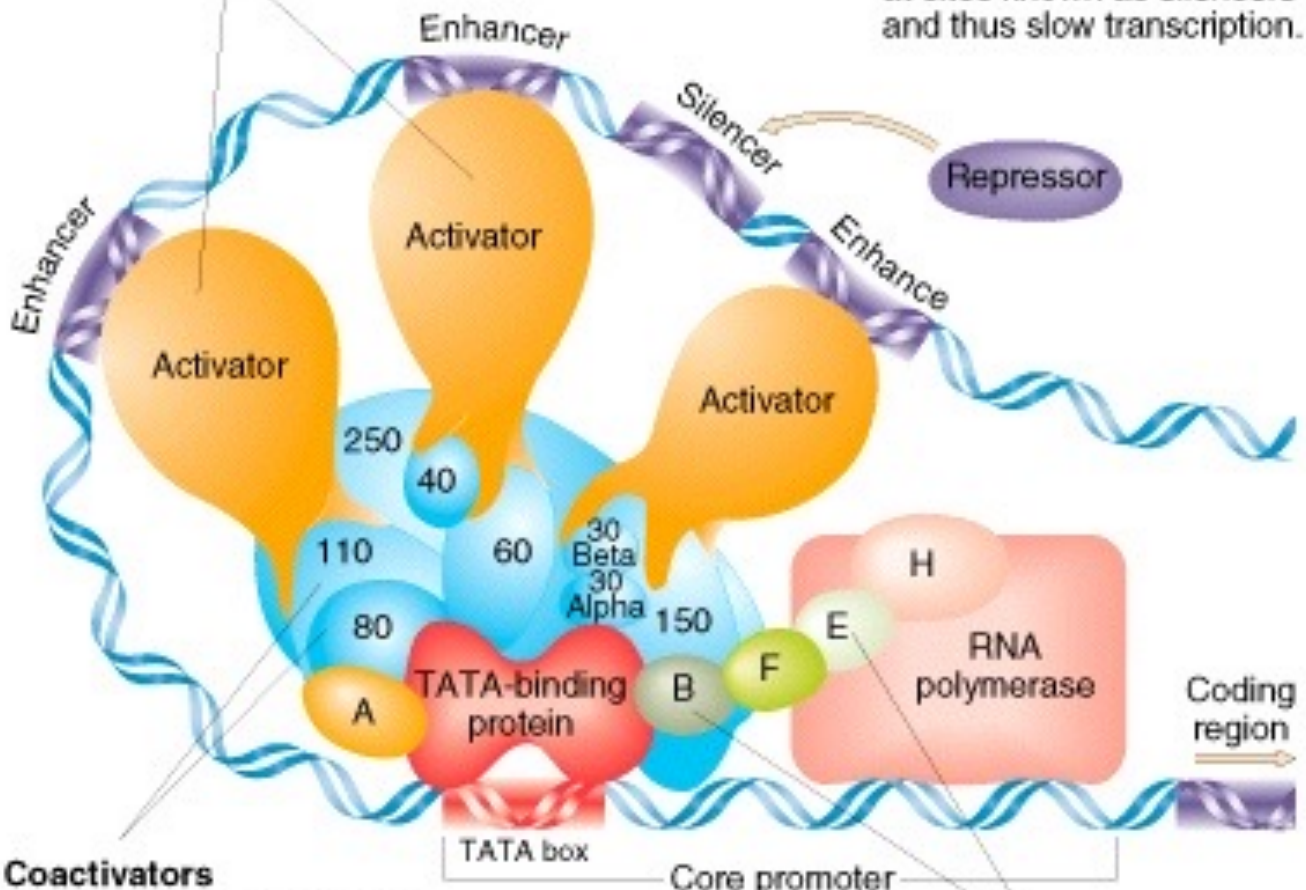


AND and OR logic. A. P and N are necessary for expression of gene *M*. B. Either P or Q are sufficient for expression of gene S



Activators
These proteins bind to genes at sites known as *enhancers* and speed the rate of transcription.

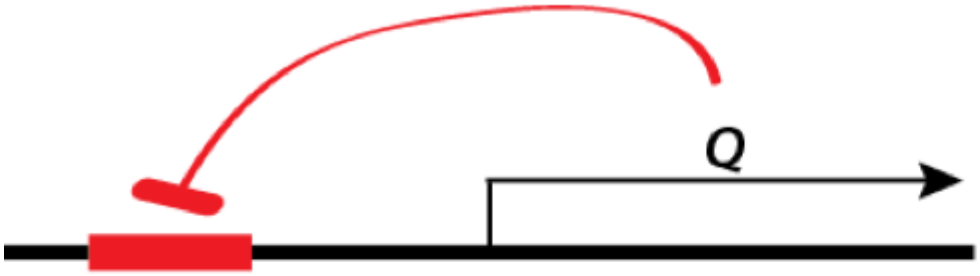
Repressors
These proteins bind to selected sets of genes at sites known as *silencers* and thus slow transcription.



Coactivators
These "adapter" molecules integrate signals from activators and perhaps repressors.

Basal transcription factors
In response to injunctions from activators, these factors position RNA polymerase at the start of transcription and initiate the transcription process.

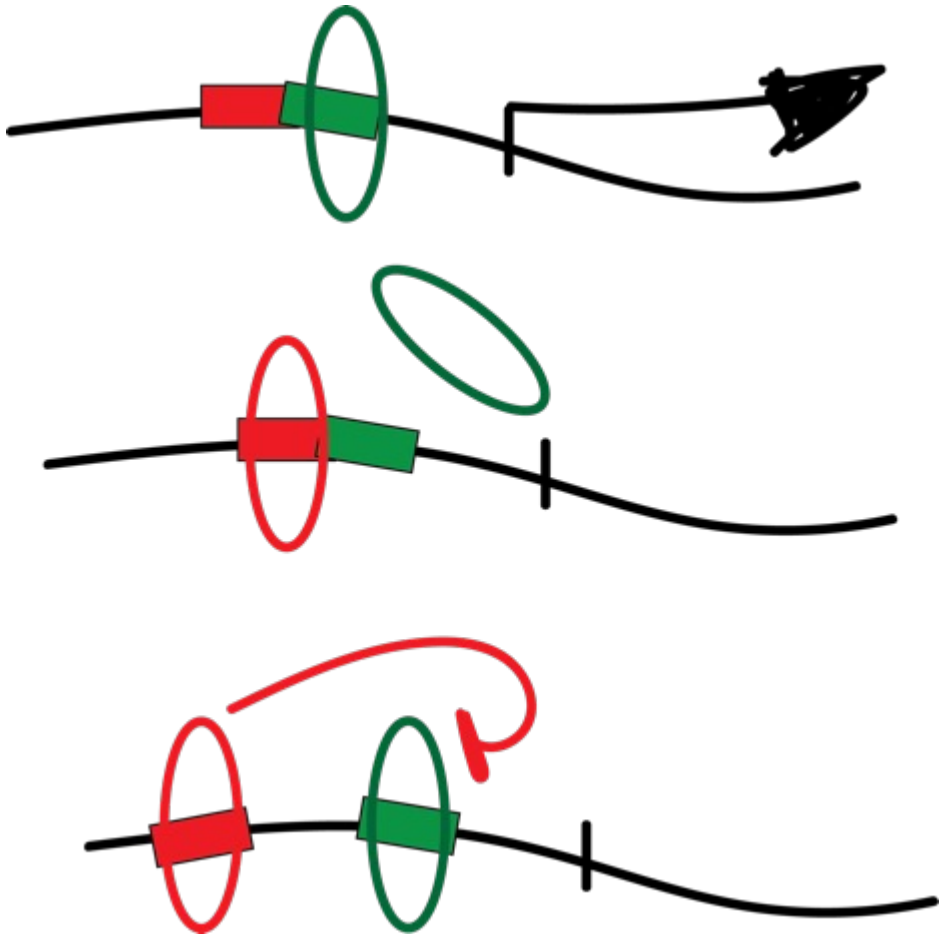
A. Negative autoregulation



Expression	
Q_+	Q_{lf}
+	+++

B. Positive autoregulation

R_+	R_{lf}
+	-



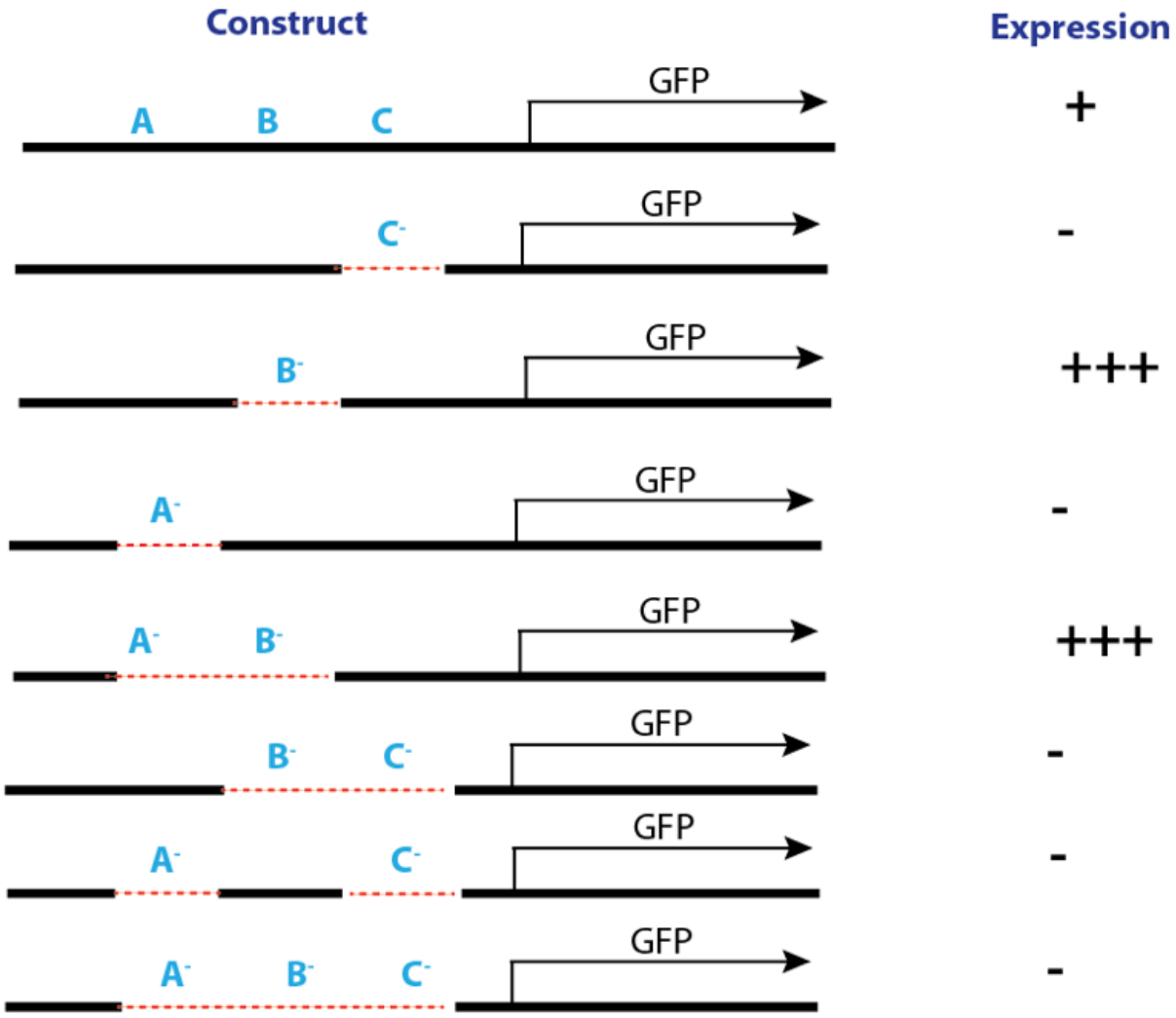
Overlap of positive and negative binding sites.

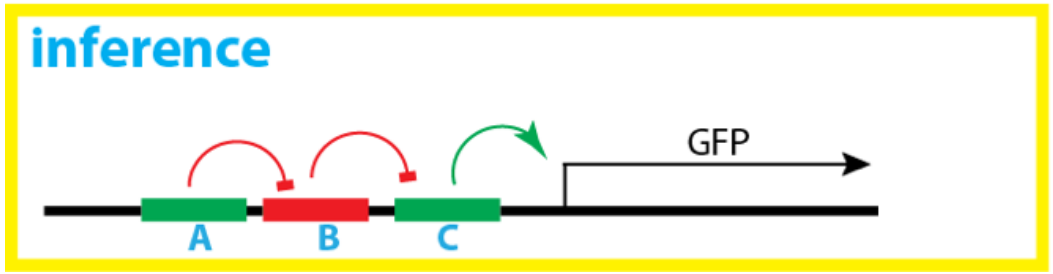
A. Positive factor shown in green activates transcription.

B. Repressor binding to red site preclude activator from binding green site and there is no transcription.

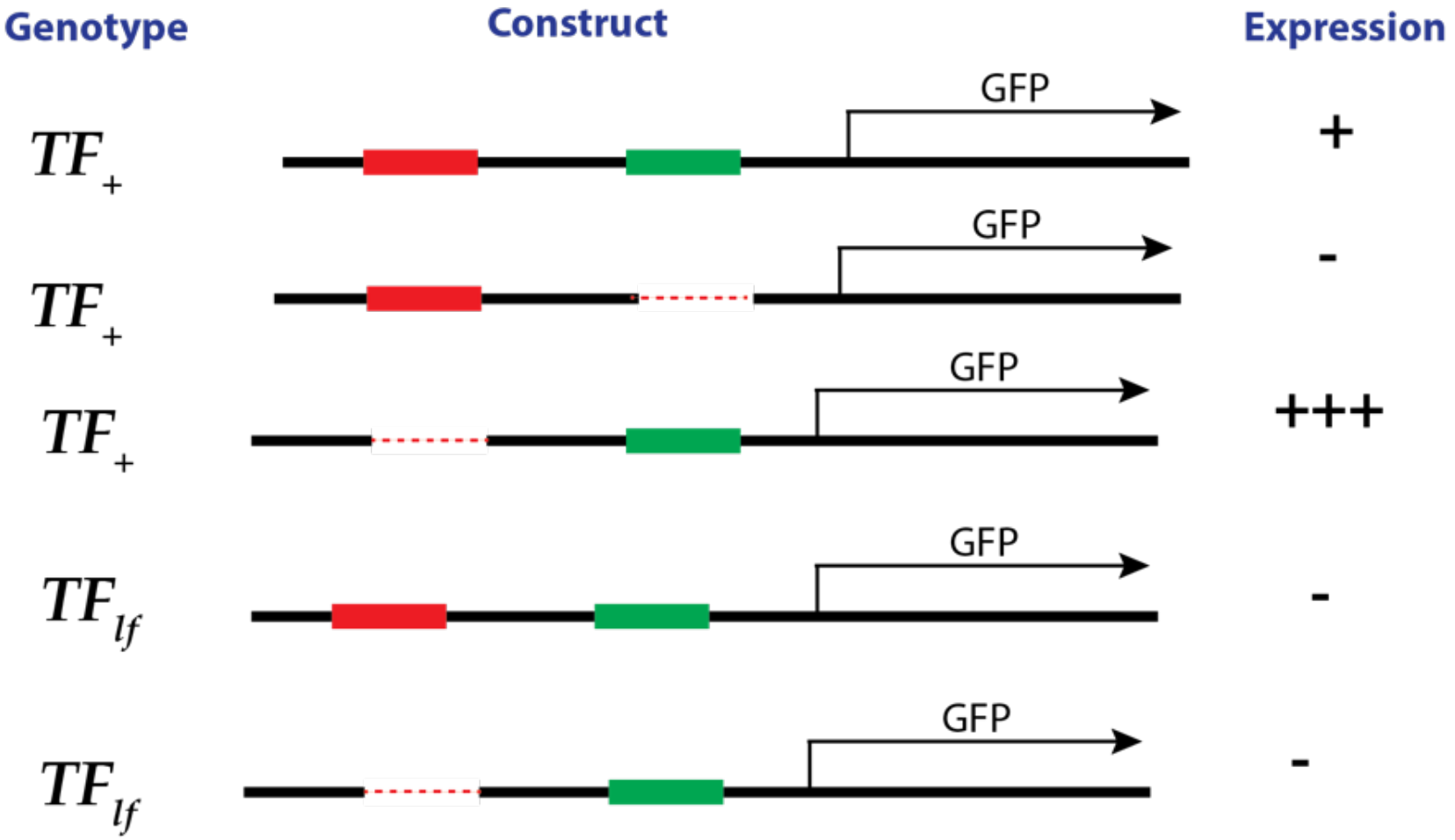
C. An alternative possibility in which both the repressor and the activator can bind and the repressor prevents the activator from positively regulation transcription.

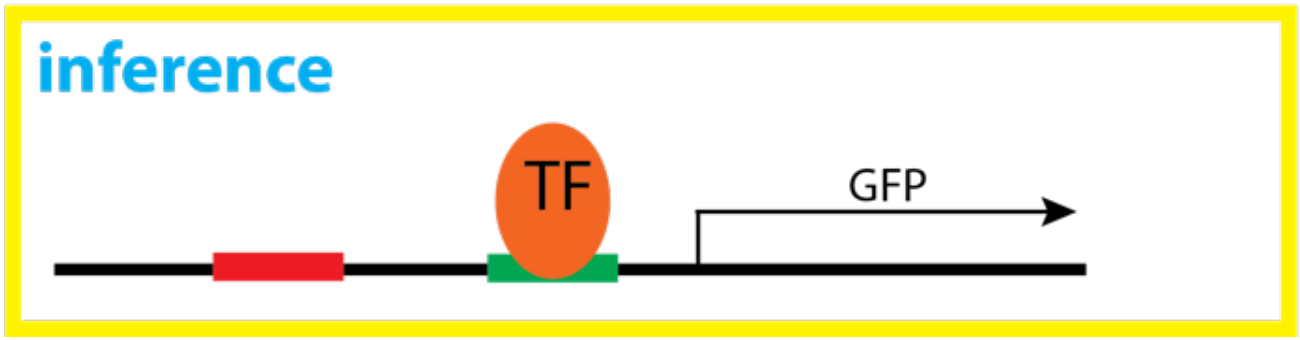
Epistasis between positive and negative acting elements



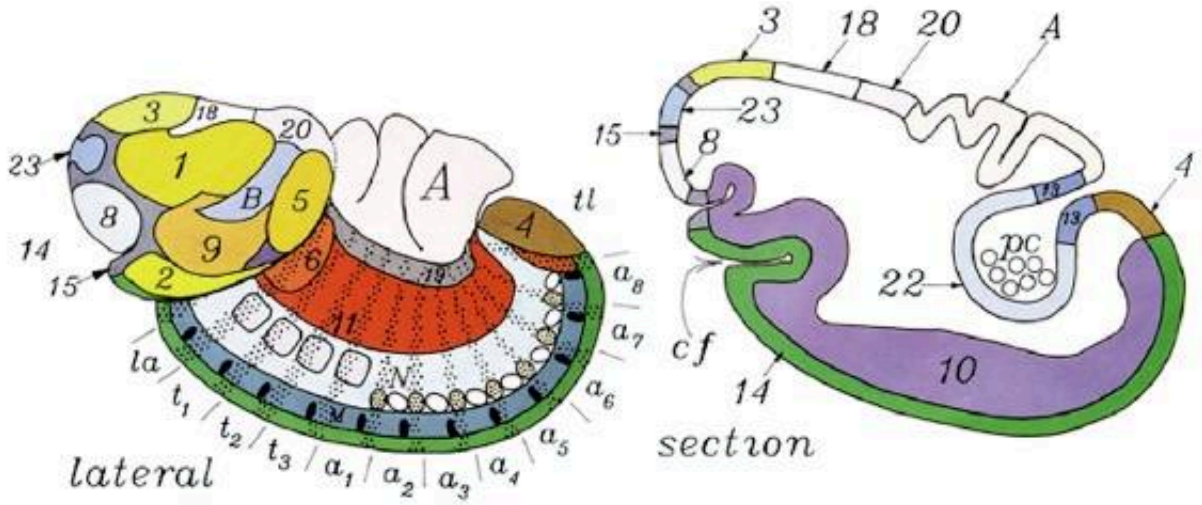


Epistasis of cis and trans (TF = transcription factor)

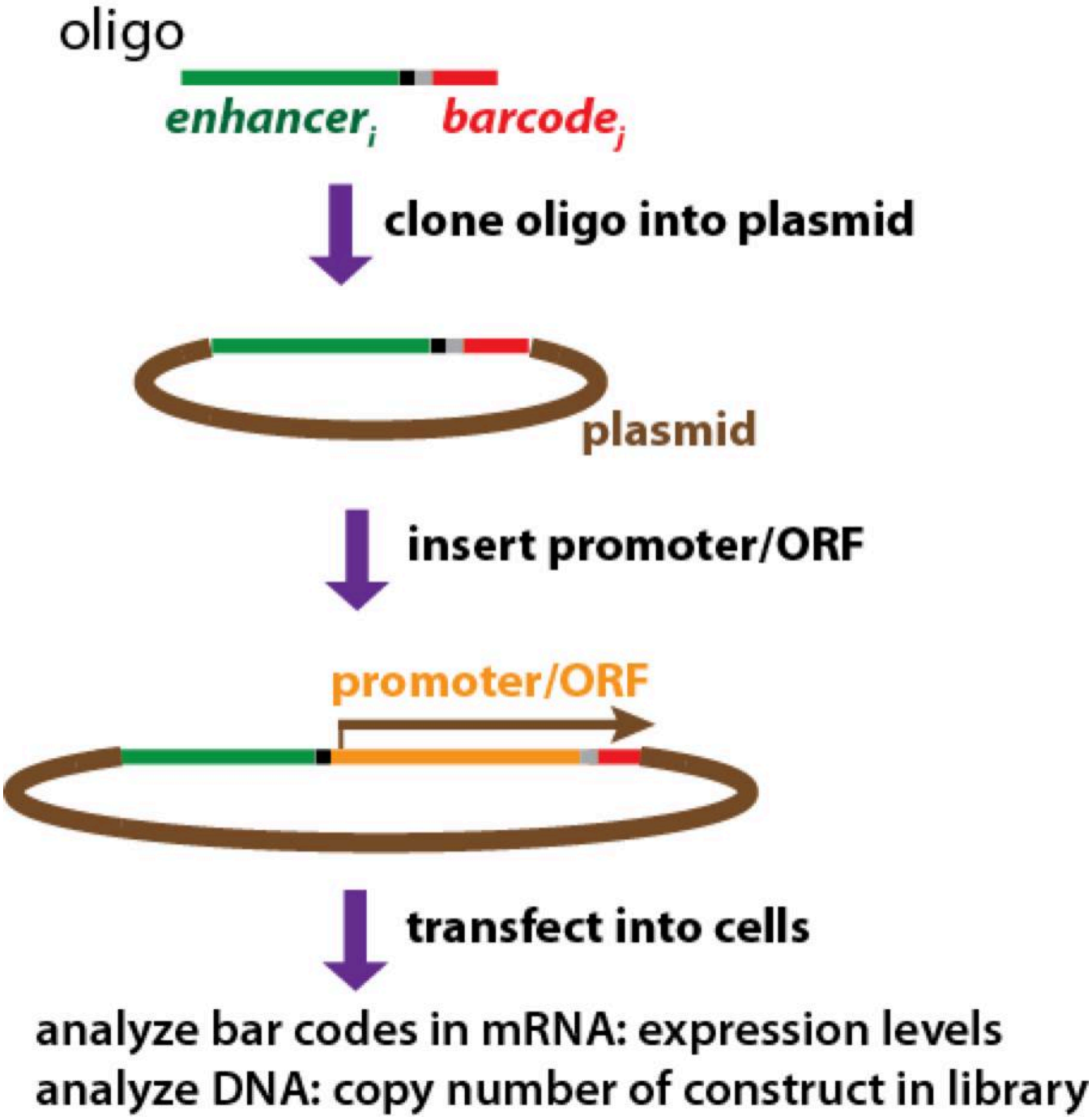


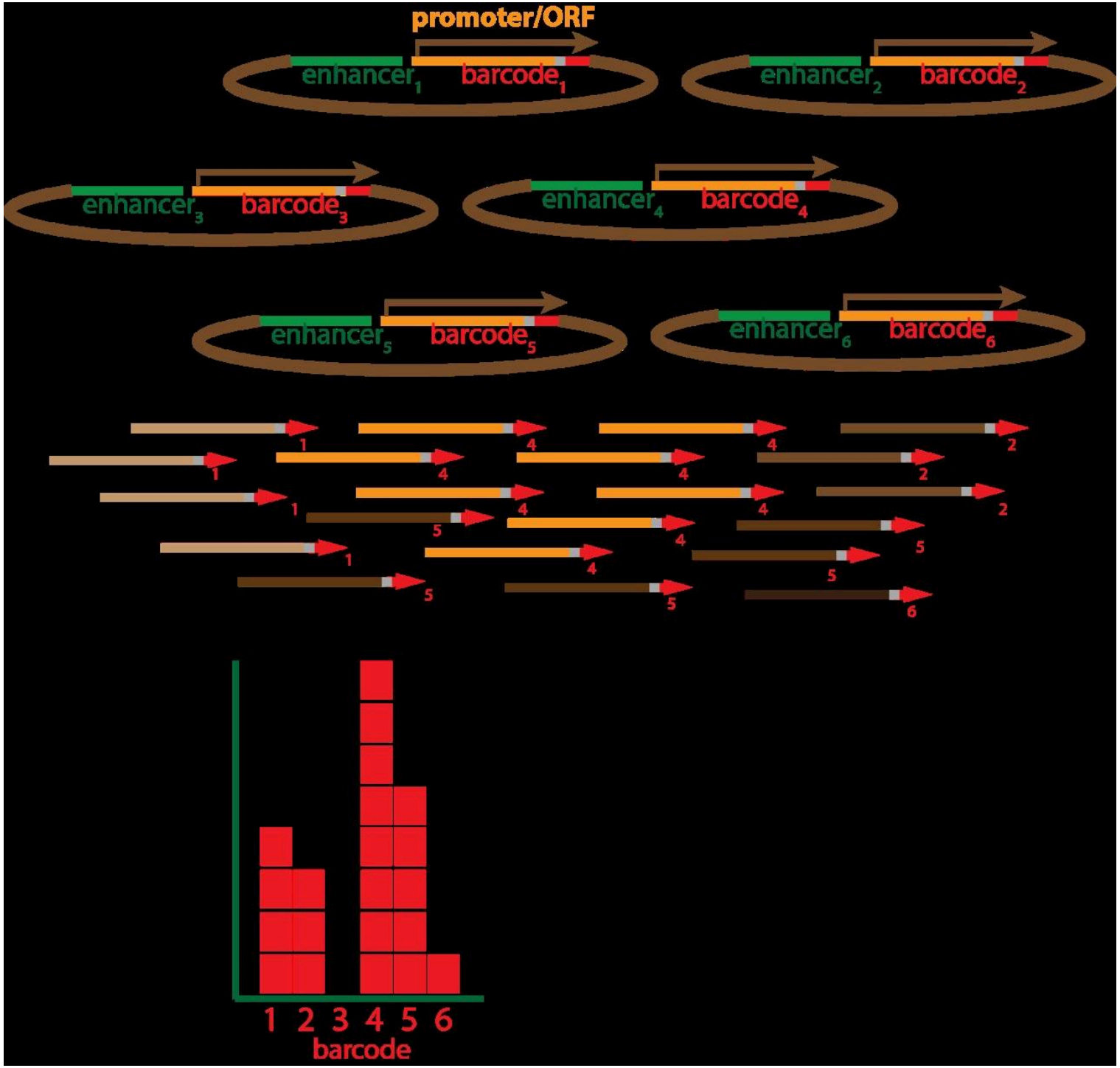


Enhancers are often modular. A short stretch of DNA is sufficient to make up a CNS enhancer. String is protein involved in Mitosis. Different cell types express different proteins, each of which binds a different string enhancer, resulting in String transcription



massively parallel enhancer assays




promoter/ORF
enhancer1
barcode1
enhancer2
barcode2
enhancer3
barcode3
enhancer4
barcode4
enhancer5
barcode5
enhancer6
barcode6
1 2 3 4 5 6
barcode

Connecting transcriptional regulators to DNA and vice versa.

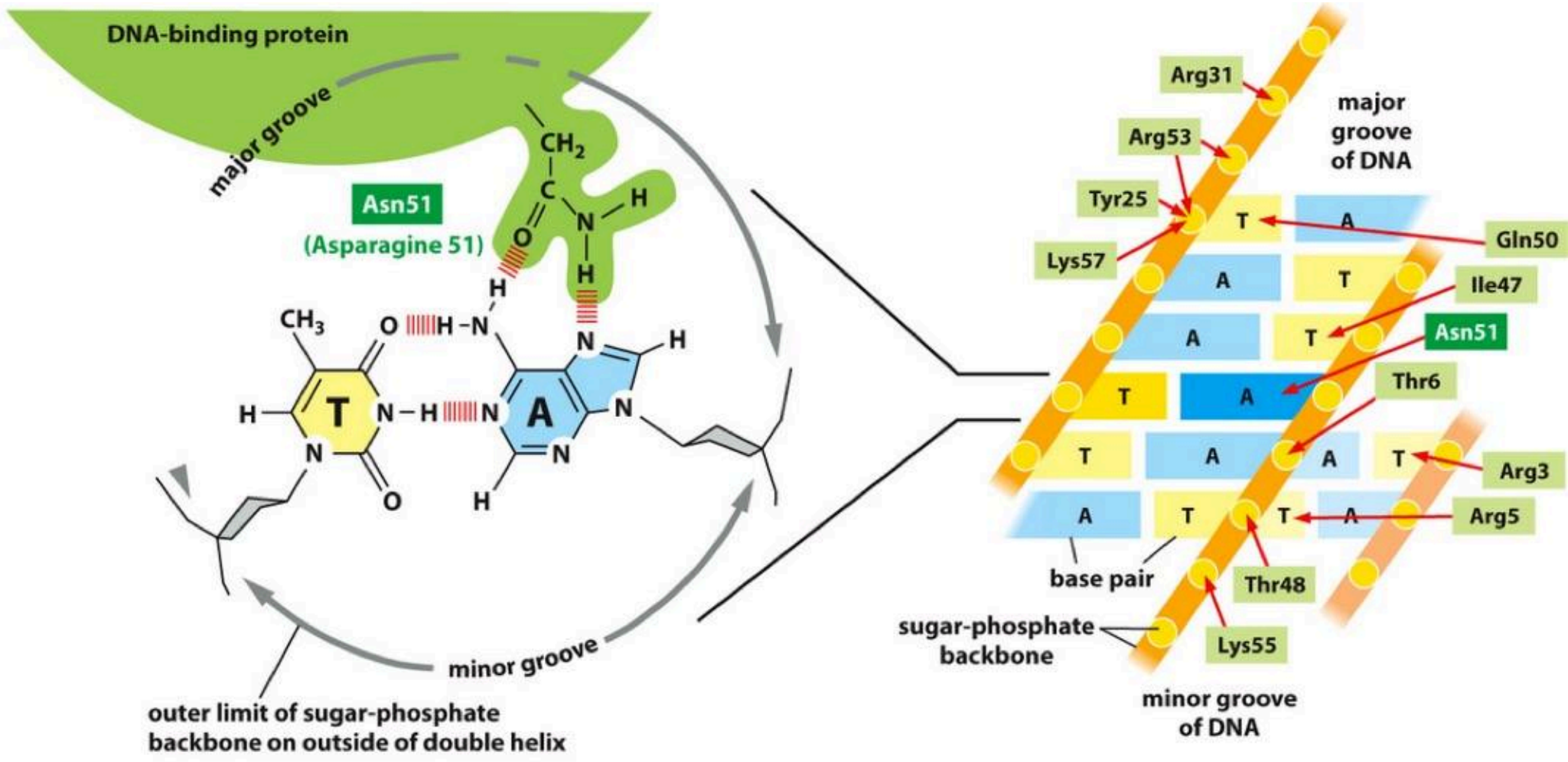
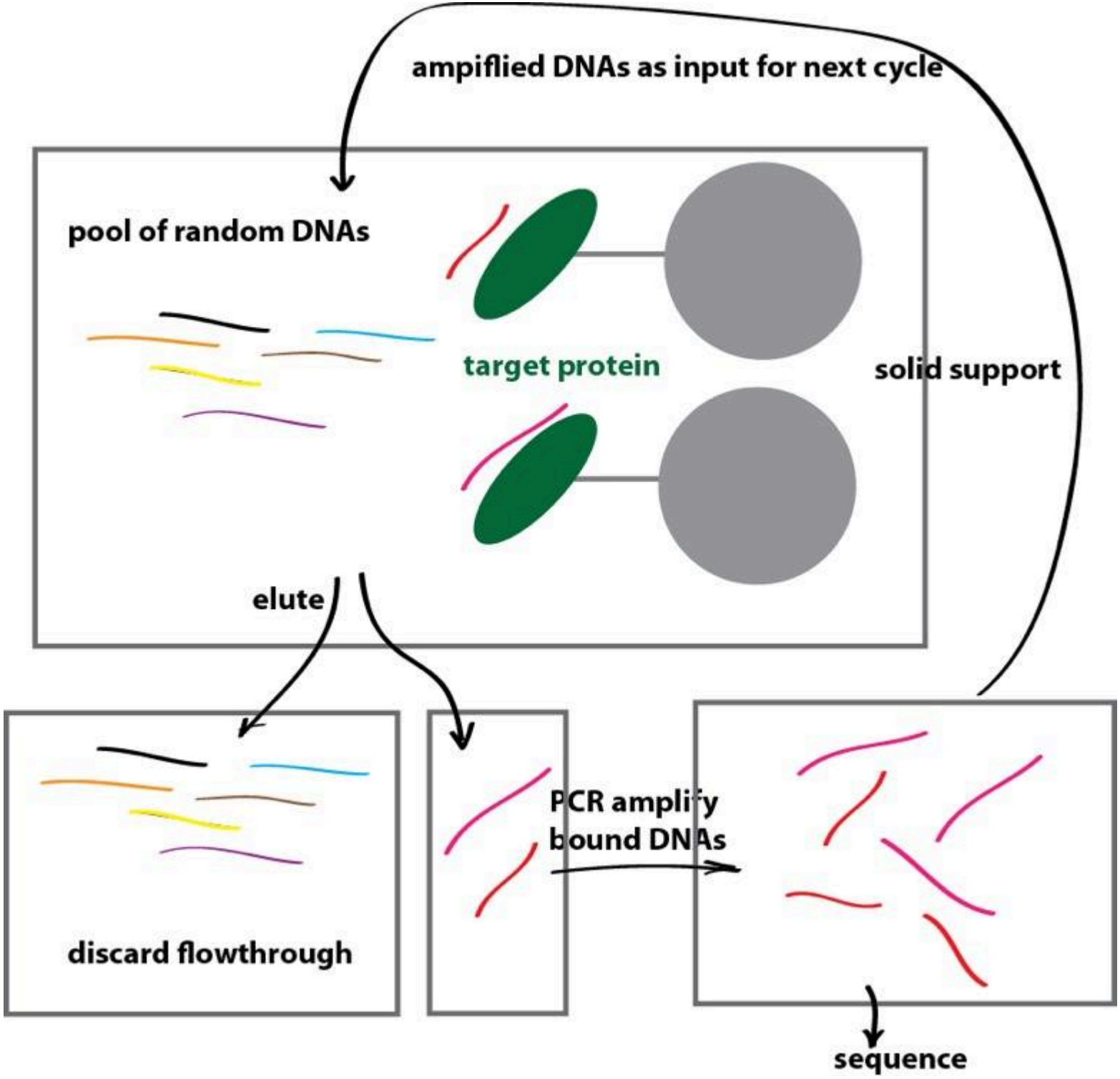
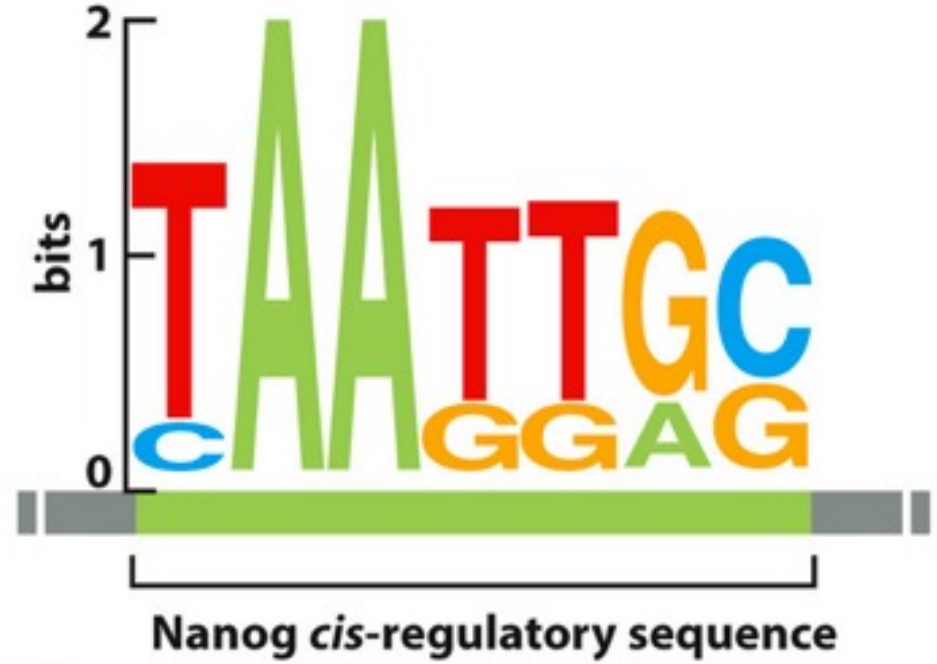


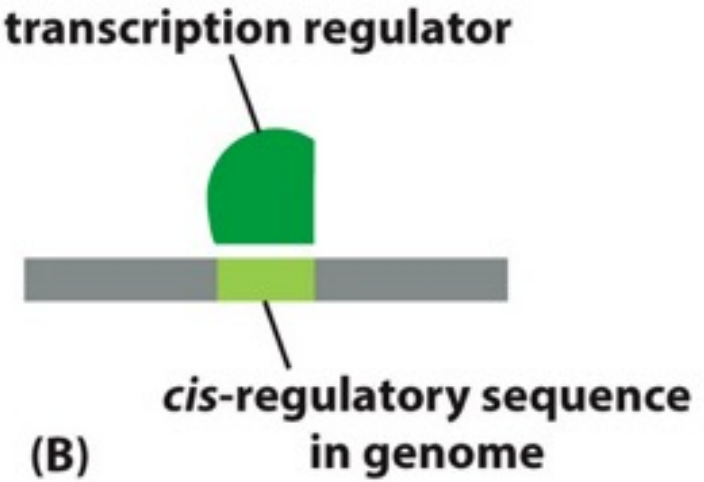
Figure 7-8 Molecular Biology of the Cell 6e (© Garland Science 2015)

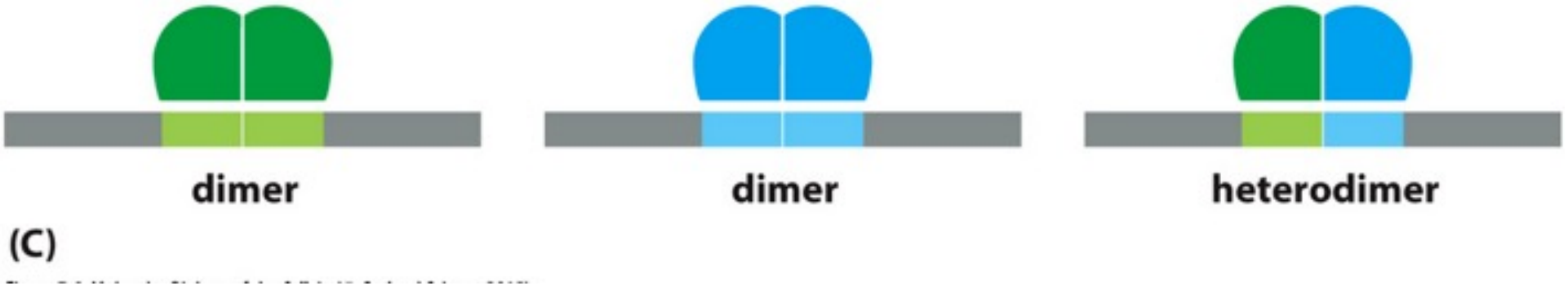
SELEX.
pool of sequences, bind separate bound from unbound, amplify bound with PCR. Iterate, and sequence to analyze



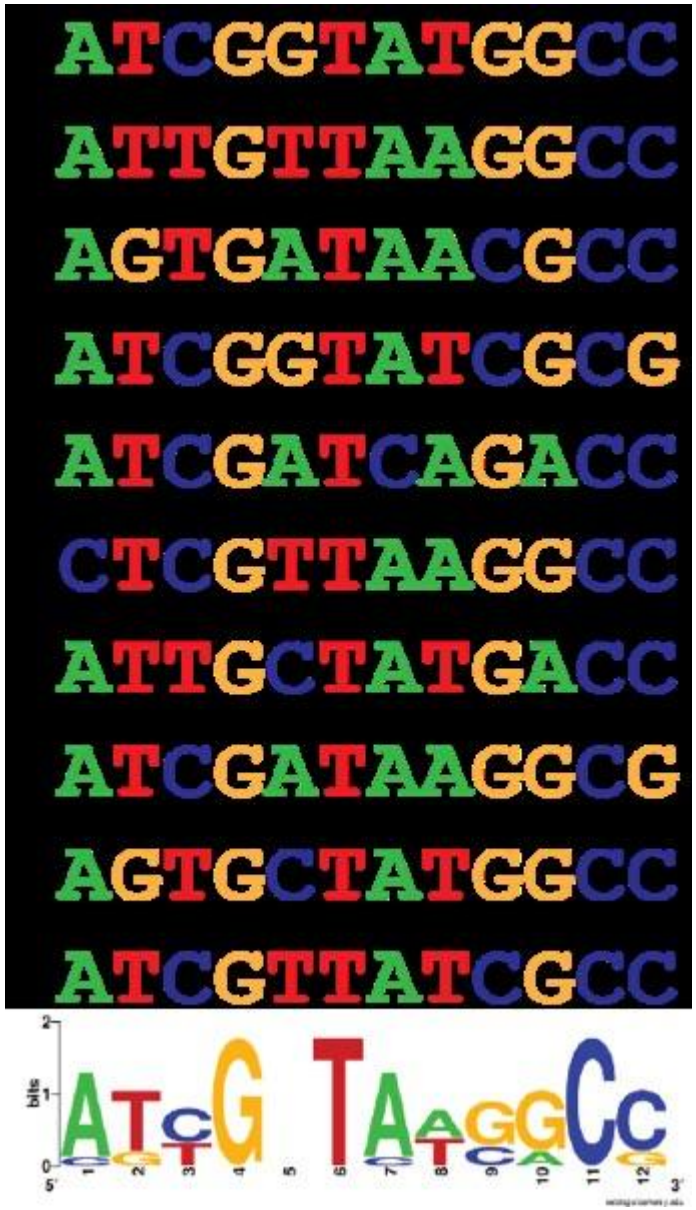
Regulatory sequence binding motif

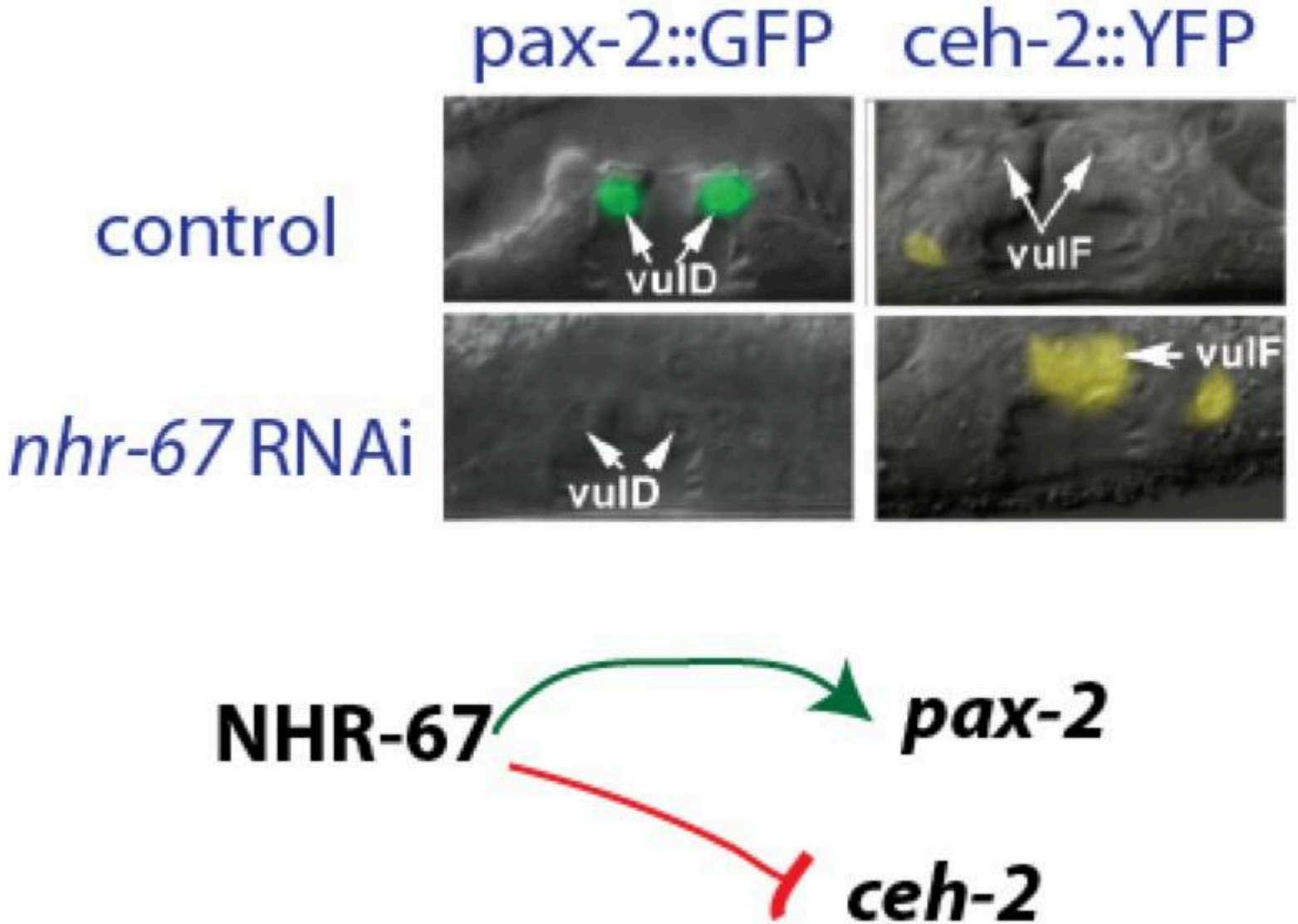


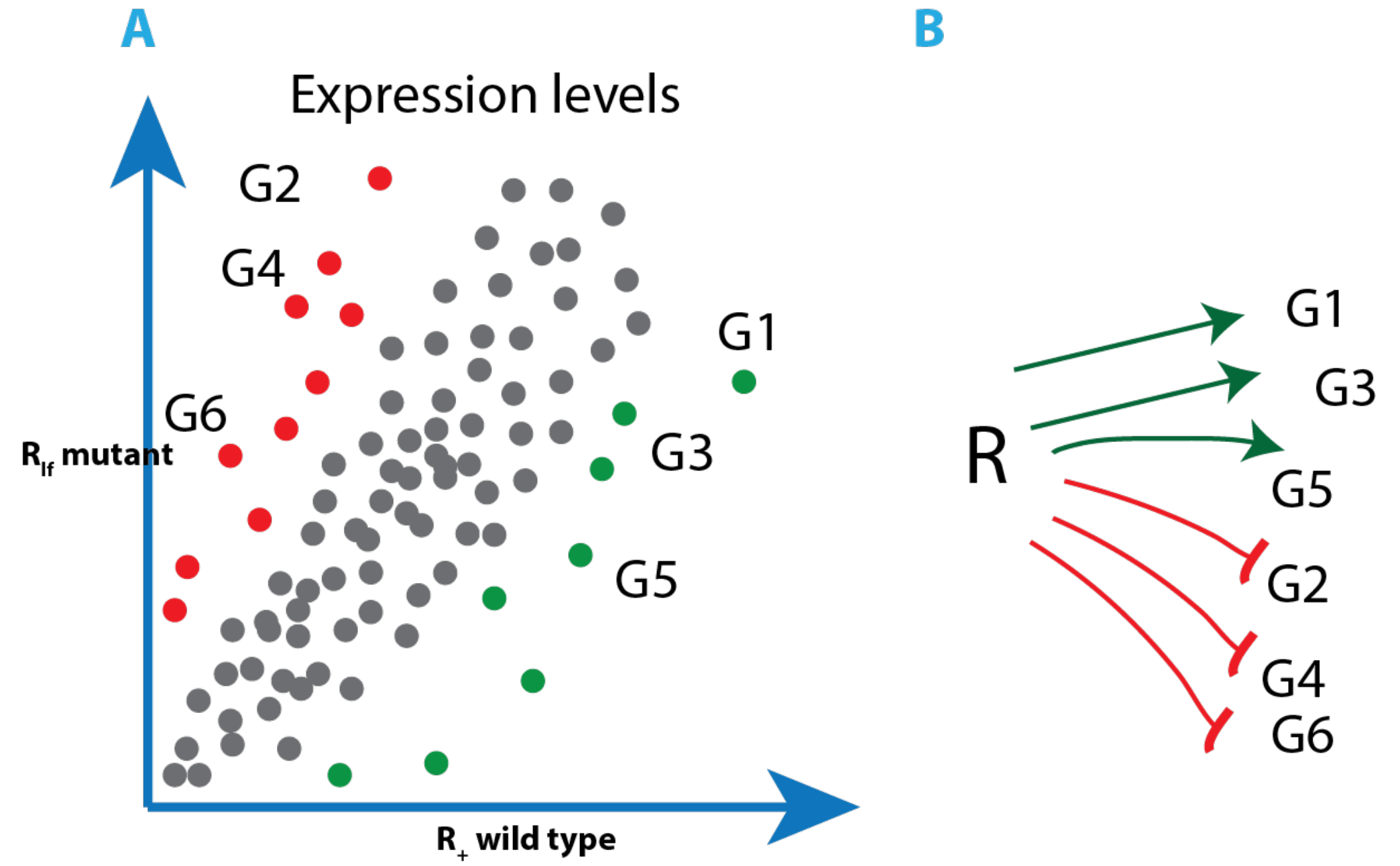




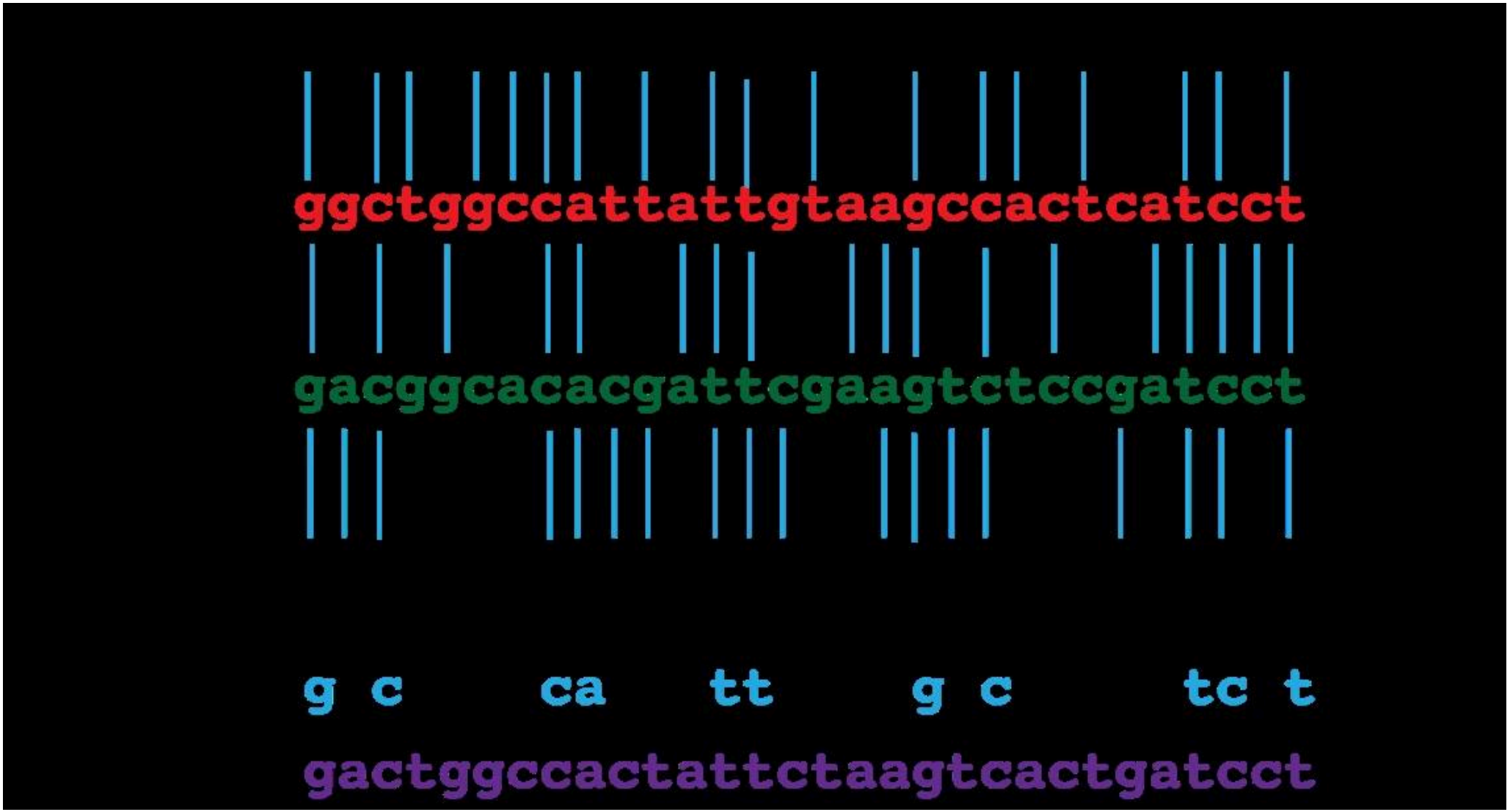
Deriving sequence logos




pax-2::GFP
ceh-2::YFP
control
nhr-67 RNAi
vulD
vulF
vulD
vulF
NHR-67
pax-2
ceh-2

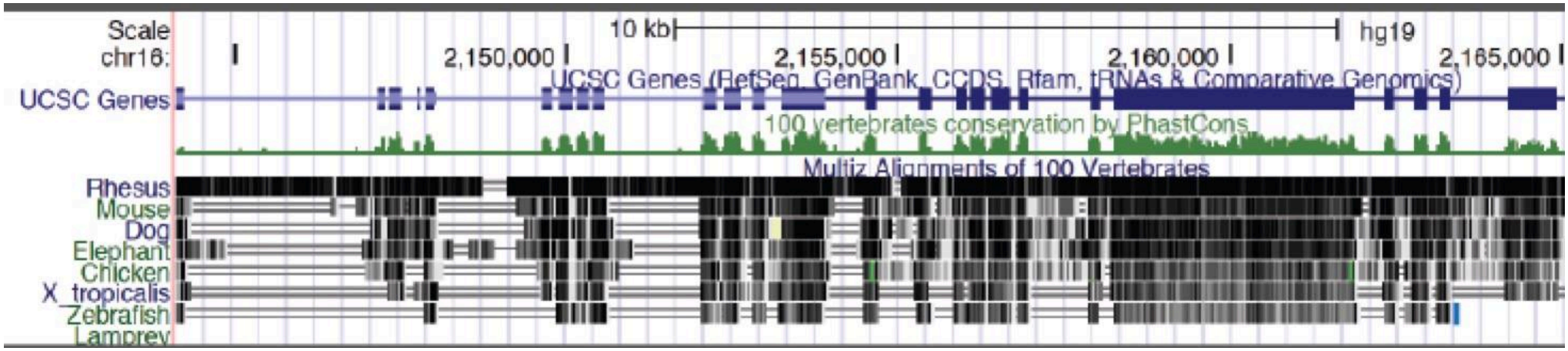

A
Expression levels
G2
G4
G1
G6
G3
G5
R_{lf} mutant
R_+ wild type
B
R
G1
G3
G5
G2
G4
G6

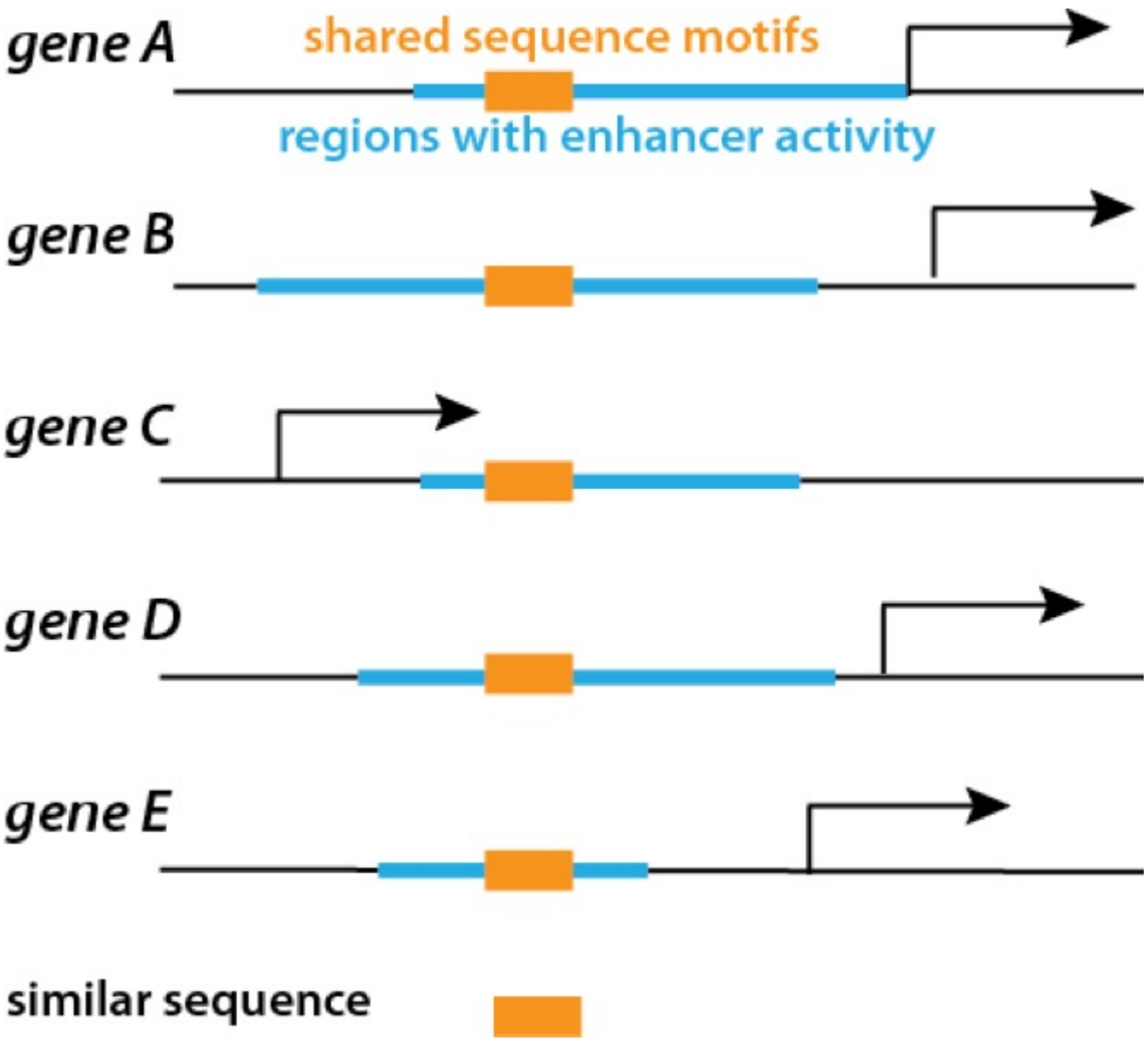
Multiple alignments allow you to identify invariant sequences among different sequences

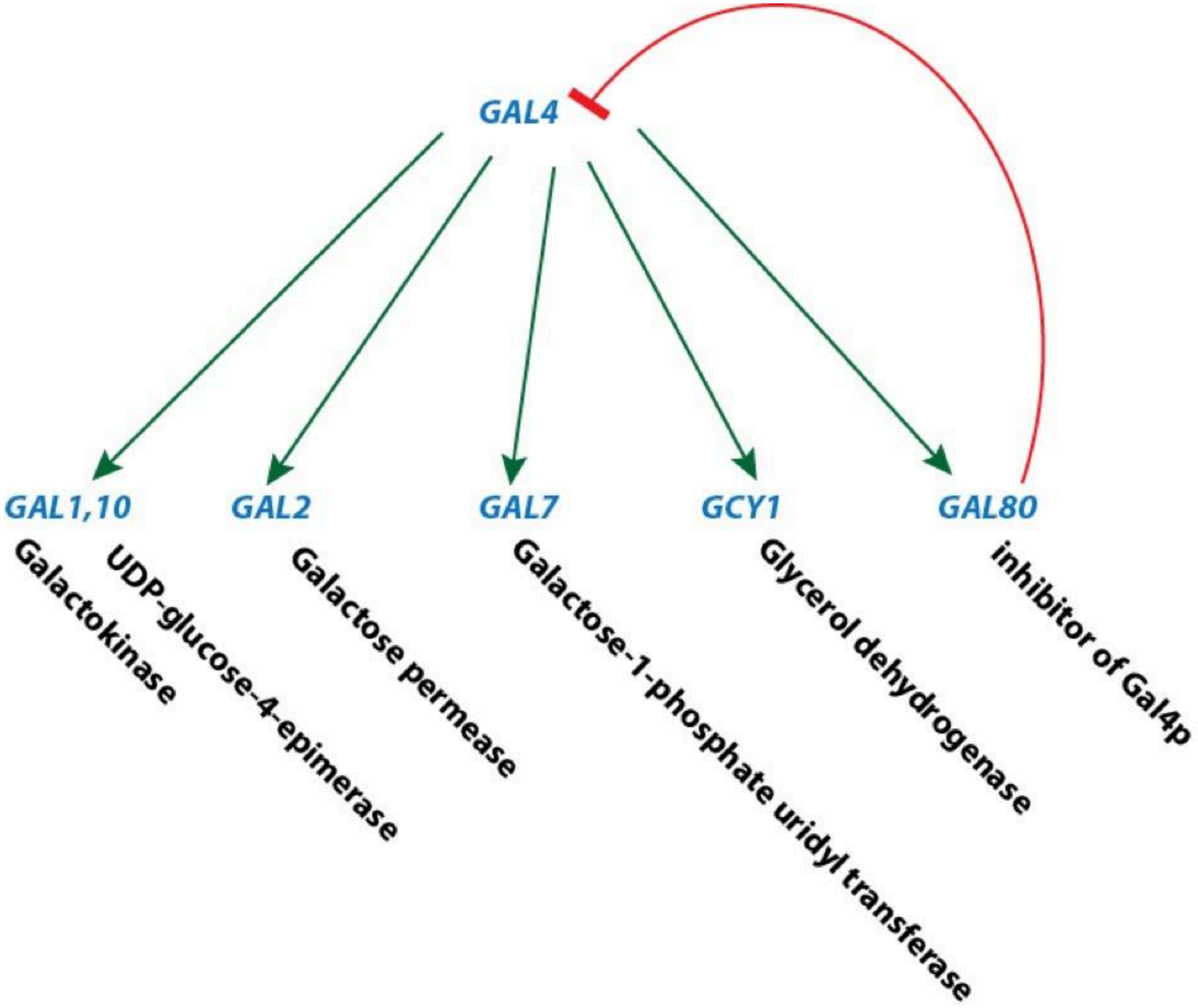


Same gene. Phylogenetic footprinting. In red are the numbers of differences from the first sequence.

Example of conservation. The green shows the phastCons Score.
The tracks below show the similarity of the human sequence with different vertebrates.




gene A
shared sequence motifs
regions with enhancer activity
gene B
gene C
gene D
gene E
similar sequence


GAL4
GAL1,10
GAL2
GAL7
GCY1
GAL80
Galactokinase
UDP-glucose-4-epimerase
Galactose permease
Galactose-1-phosphate uridyl transferase
Glycerol dehydrogenase
inhibitor of Gal4p

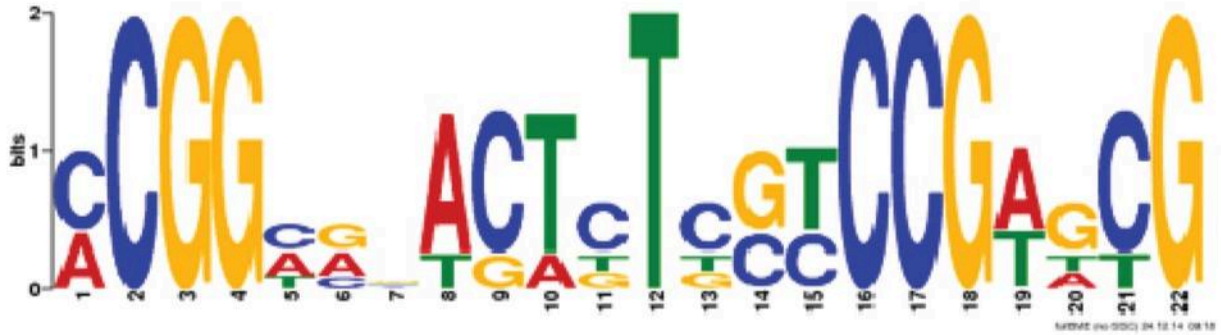
GAL7
aacactgtatagccagtccttcggatcacggtcaacagttgtccgagcgcttttggacccttcccttatttttgggtt

GAL80
taatatcctatattttcttcatttaccggcgcactctcgcccgaacgacctcaaaatgtctgctacattcataataacca

GAL1,10
gaagtacggattagaagccgccgagcgggcgacagccctccgacggaagactctcctccgtgcgtcctcgtcttcaccgg

GAL2
cgaaaggggcggttgcctcaggaaggcaccggcggtctttcgtccgtgcggagatatctgcgccgttcaggggtccatgt

GCY1
acttcctatggagtgtataagaattgtaagttataacaccggcgaacaatcggggcagactattccggggaagaacaagg

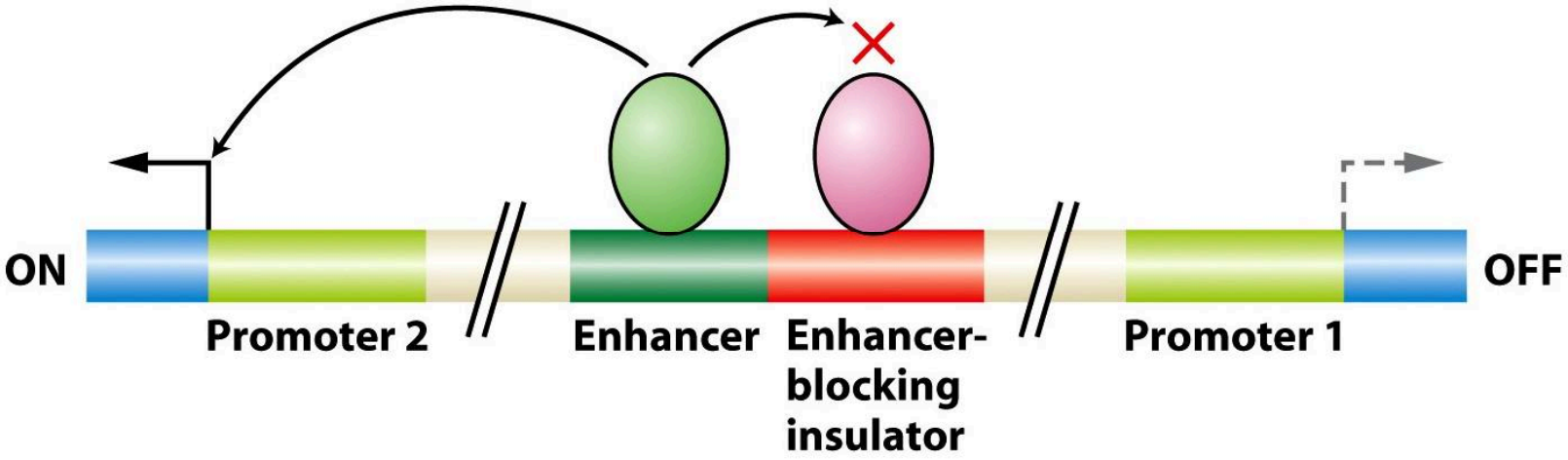


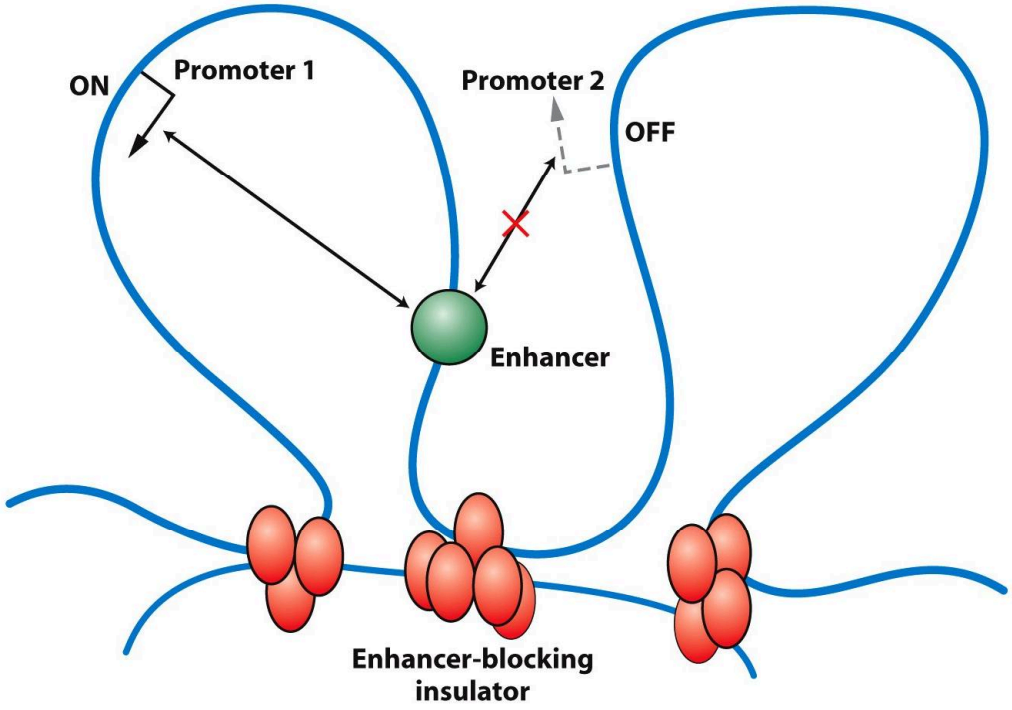
opposite strand

E-value 2.9e-002 = 0.029
observe 5 sites

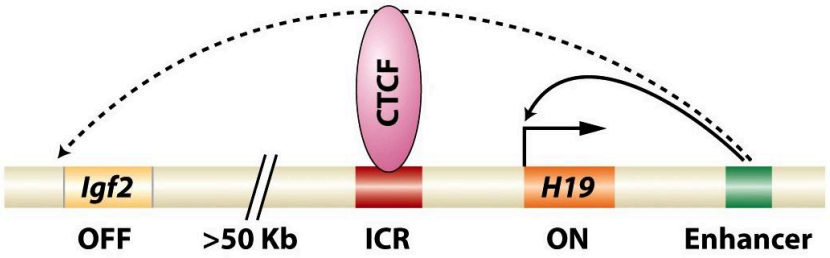
Name	Strand	Start	*p*-value	Sites
GAL1,10	+	43	8.74e-12	CAGCCCTCCG ACGGAAGACTCTCCTCCGTGCG TCCTCGTCTT
GAL80	+	26	1.30e-10	TCTTCATTTA CCGGCGCACTCTCGCCCGAACG ACCTCAAAAT
GAL2	+	29	3.71e-10	CAGGAAGGCA CCGGCGGTCTTTCGTCCGTGCG GAGATATCTG
GAL7	+	28	6.36e-09	CCTTCGGATC ACGGTCAACAGTTGTCCGAGCG CTTTTTGGAC
GCY1	-	47	7.85e-09	TTGTTCTTCC CCGGAATAGTCTGCCCCGATTG TTCGCCGGTG

Insulators create isolated chromatin domains, which constrain the activity of enhancers to promoters in the same domain

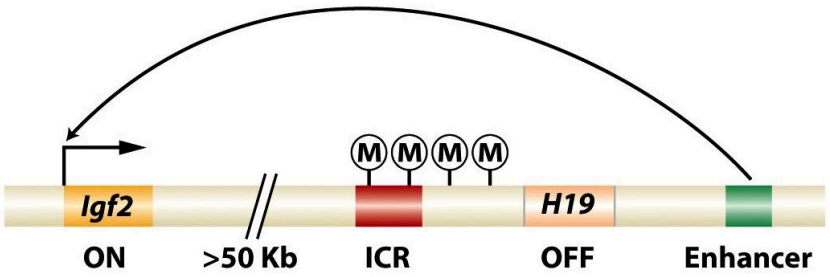


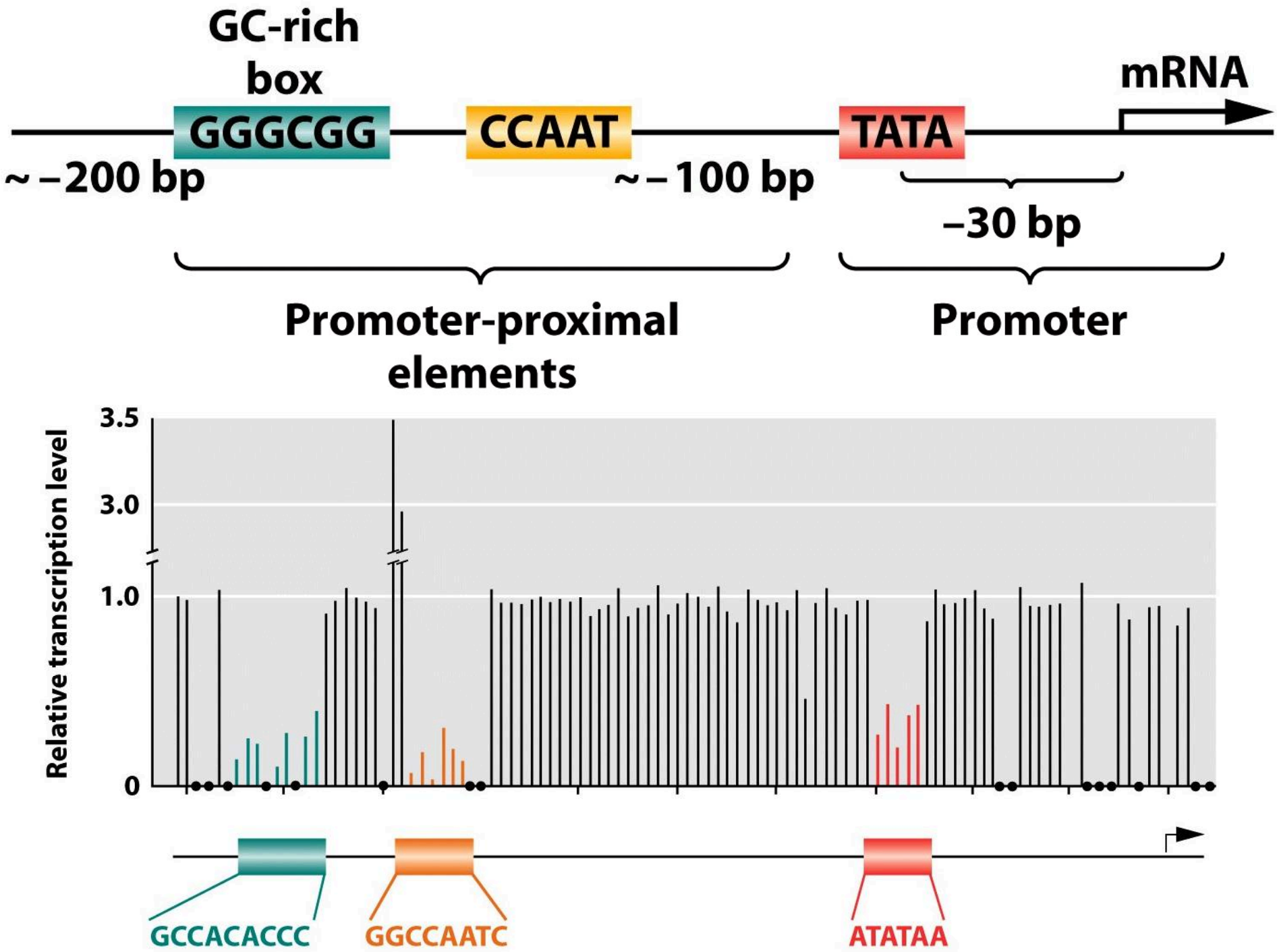


♀ Maternal allele



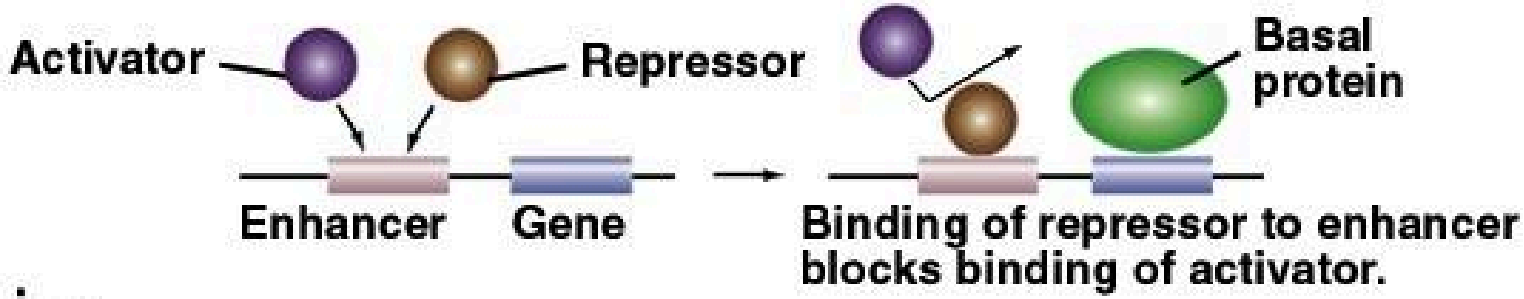
♂ Paternal allele




GC-rich box
GGGCGG
CCAAT
TATA
mRNA
~ –200 bp
~ –100 bp
–30 bp
Promoter-proximal elements
Promoter
Relative transcription level
3.5
3.0
1.0
0
GCCACACCC
GGCCAATC
ATATAA

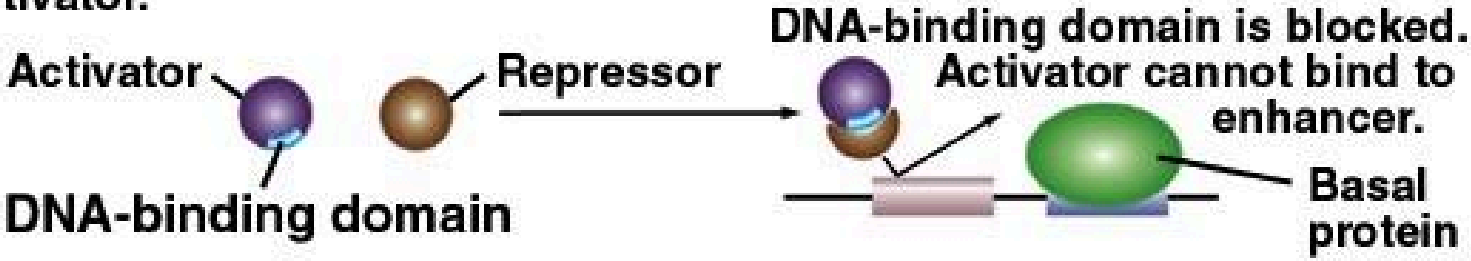
How repressors work

(a)
Competition for binding between repressor and activator proteins

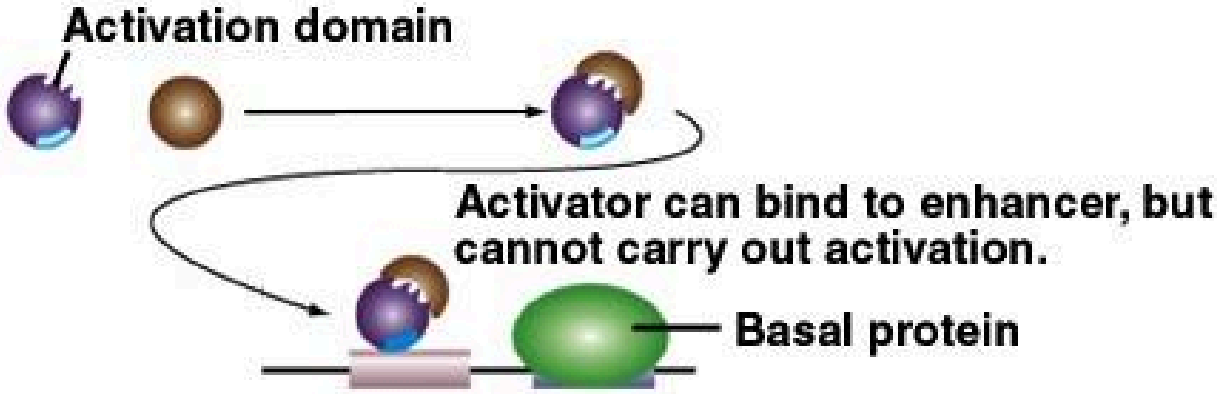


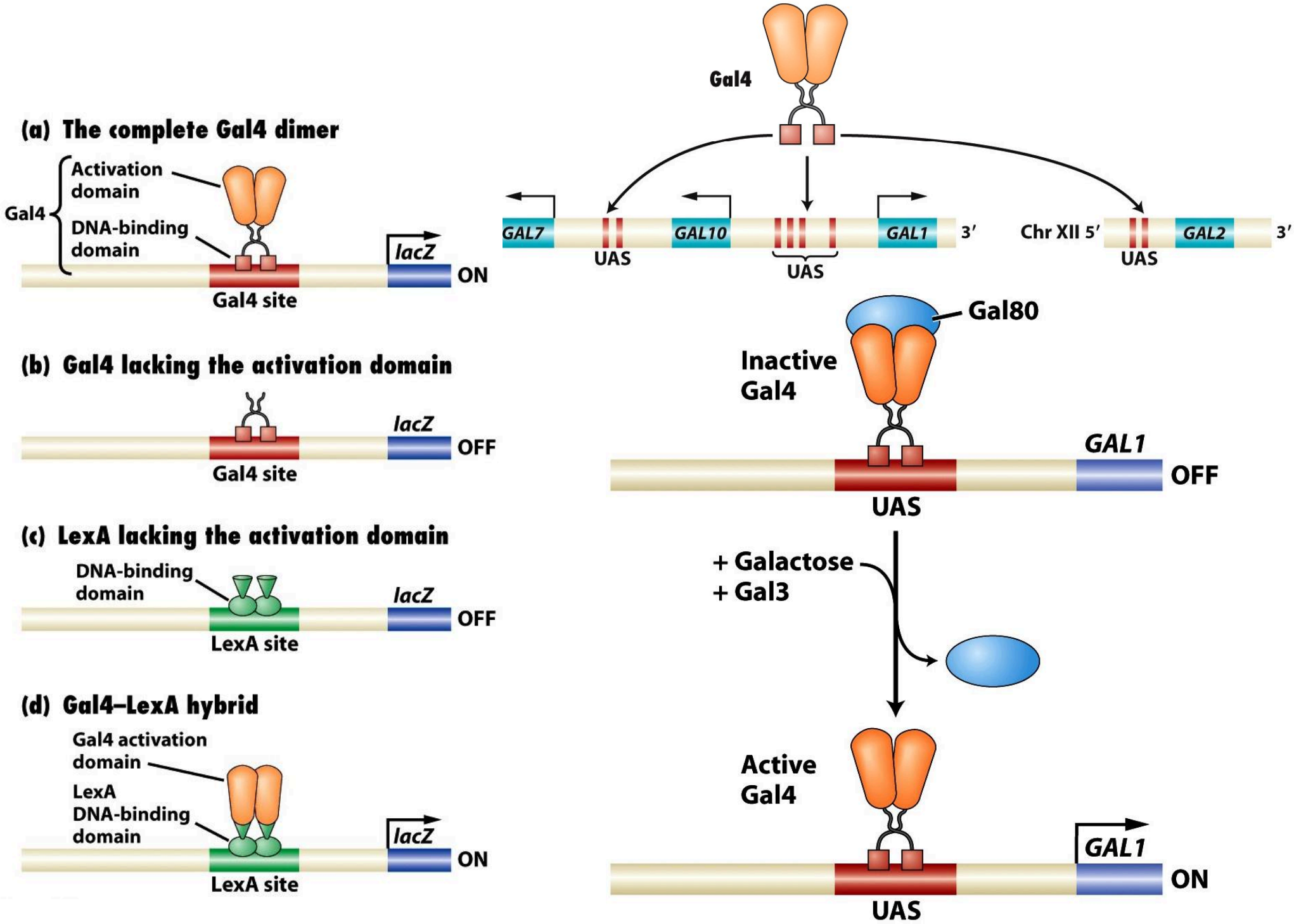
(b)
Quenching

Type I: Repressor binds to and blocks the DNA-binding region of an activator.



Type II: Repressor binds to and blocks the activation domain of an activator.




(a) The complete Gal4 dimer
Gal4
Activation domain
DNA-binding domain
lacZ
ON
Gal4 site
(b) Gal4 lacking the activation domain
lacZ
OFF
Gal4 site
(c) LexA lacking the activation domain
DNA-binding domain
lacZ
OFF
LexA site
(d) Gal4–LexA hybrid
Gal4 activation domain
LexA DNA-binding domain
lacZ
ON
LexA site
Gal4
GAL7
GAL10
GAL1
3′
UAS
UAS
Chr XII 5′
UAS
GAL2
3′
Gal80
Inactive Gal4
GAL1
OFF
UAS
+ Galactose
+ Gal3
Active Gal4
GAL1
ON
UAS

Histones are modified by acetyltransferases and methyltransferases, Which promote or inhibit nucleosome slippage