Translation - Prokaryotes

Shine-Dalgarno (SD) Sequence

rRNA 3‘-GAUACCAUCCUCCUUA-5‘

mRNA ....G**GAGG**..(5-7bp)...**AUG**

Influences:

Secondary structure!! SD and AUG in unstructured region

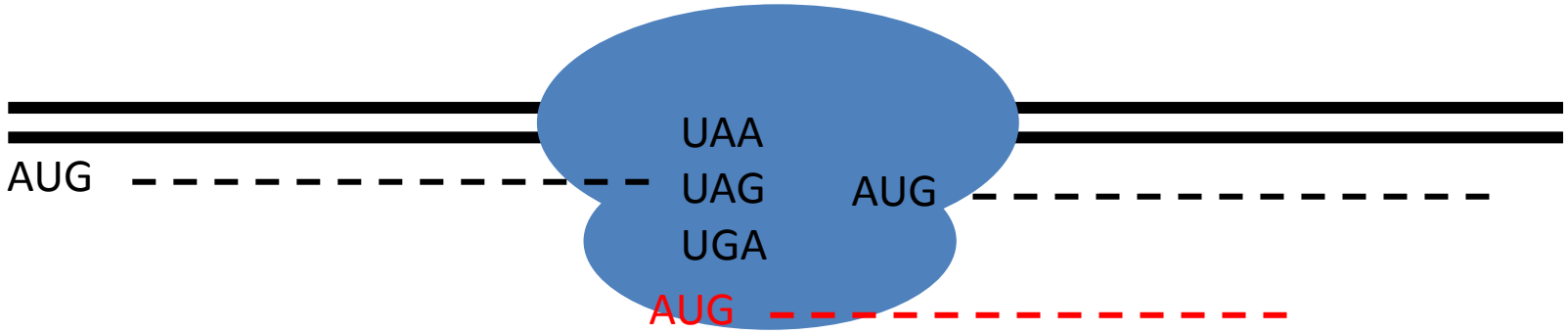
Surrounding of SD and AUG!!!

Start

AUG	91%
GUG	8
UUG	1

Ribosomal protein S1: present only in Gram-negatives (not in Gram-positives):
→ binds to AU-rich sequences found in many prokaryotic mRNAs 15-30 nucleotides upstream of start-codon

Translational coupling



Translation - Eukaryotes

Start Codon

mRNA 5‘-CAP.....**AUG**

Influences:

Surrounding of AUG!!!

Kozak Consensus

.....CC$^{A}/_{G}$CCAUGG..... **mammalian**

..... $^{A}/_{T}A^{A}/_{C}A^{A}/_{C}$AAUGTC$^{T}/_{C}$..... **Yeast**

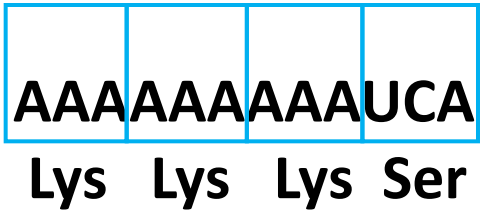
..... gccgcc(A/G)ccAUGG ..... **Wikipedia**

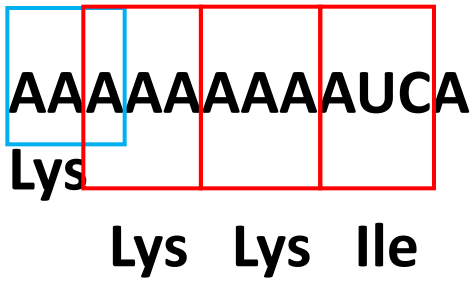
Translation elongation

- Codon usage

- Secondary structures

- Codon structure – translational frameshifting





First base	Second base: U	C	A	G
U	UUU Phe F	UCU Ser S	UAU Tyr Y	UGU Cys C
	UUC	UCC	UAC	UGC
	UUA Leu L	UCA	UAA STOP	UGA STOP
	UUG	UCG	UAG	UGG Trp W
C	CUU Leu L	CCU Pro P	CAU His H	CGU Arg R
	CUC	CCC	CAC	CGC
	CUA	CCA	CAA Gln Q	CGA
	CUG	CCG	CAG	CGG
A	AUU Ile I	ACU Thr T	AAU Asn N	AGU Ser S
	AUC	ACC	AAC	AGC
	AUA	ACA	AAA Lys K	AGA Arg R
	AUG Met M	ACG	AAG	AGG
G	GUU Val V	GCU Ala A	GAU Asp D	GGU Gly G
	GUC	GCC	GAC	GGC
	GUA	GCA	GAA Glu E	GGA
	GUG	GCG	GAG	GGG

FIGURE 25.1 All the triplet codons have meaning: 61 represent amino acids and 3 cause termination (stop codons).

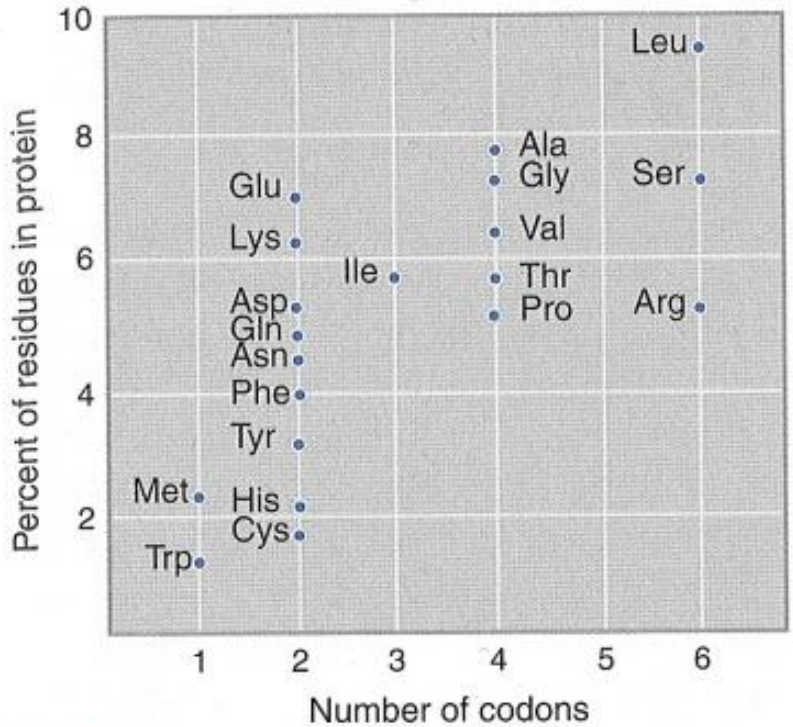


FIGURE 25.2 Some correlation of the frequency of amino acid use in proteins with the number of codons specifying the amino acid is observed. An exception is found for amino acids specified by two codons, which occur with a wide variety of frequencies.

Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

Third-base relationship		Third bases with same meaning	Codon number
UUU UUC	UCU UCC	UAU UAC	UGU UGC
UUA UUG	UCA UCG	UAA UAG	UGA UGG
CUU CUC CUA CUG	CCU CCC CCA CCG	CAU CAC	CGU CGC CGA CGG
		CAA CAG	
AUU AUC AUA	ACU ACC ACA ACG	AAU AAC	AGU AGC
AUG		AAA AAG	AGA AGG
GUU GUC GUA GUG	GCU GCC GCA GCG	GAU GAC	GGU GGC GGA GGG
		GAA GAG	

Third base irrelevant		U, C, A, G	32
		U, C, A	3
Purines differ from pyrimidines		A or G	14
		U or C	10
Unique		G only	2

FIGURE 25.3 Third bases have the least influence on codon meanings. Boxes indicate groups of codons within which third-base degeneracy ensures that the meaning is the same.

Base in first position of anticodon	Base(s) recognized in third position of codon
U	A or G
C	G only
A	U only
G	C or U

FIGURE 25.5 Codon–anticodon pairing involves wobbling at the third position.

Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

Universal Triplet Code → rare exemptions

		Other mitochondrial codes			Other codes in cellular chromosomes		
Codon	**Universal code**	Mycoplasma	Paramecium	Euplotes	Yeast	Protozoa	Mammals
UGA	Stop	Tryptophan	Stop	Cysteine	Tryptophan	Tryptophan	Tryptophan
UAA/UAG	Stop	Stop	Glutamine	Stop	Stop	Stop	Stop
AUA	Isoleucine	Isoleucine	Isoleucine	Isoleucine	Methionine	Methionine	Methionine
CUA	Leucine	Leucine	Leucine	Leucine	Threonine	Leucine	Leucine
AGA/AGG	Arginine	Arginine	Arginine	Arginine	Arginine	Arginine	Stop

The universal genetic code is used in the chromosomes of most cells, chloroplasts, plant mitochondria, and their viruses and plasmids. A few organisms use slightly different codes in their chromosomes (in the nucleus). The examples of these other nuclear codes are from Mycoplasma (Bacteria) and two different ciii ated protozoa (Eukarya). All nonplant mitochondria use variations of the universal code, whereas plant mitochondria use the universal code. The examples here are only a few of the different types known.

1

Ribosomes		rRNAs	r-proteins
Bacterial (70S) mass: 2.5 MDa 66% RNA	50S	23S = 2904 bases 5S = 120 bases	31
	30S	16S = 1542 bases	21
Mammalian (80S) mass: 4.2 MDa 60% RNA	60S	28S = 4718 bases 5.8S = 160 bases 5S = 120 bases	49
	40S	18S = 1874 bases	33

FIGURE 24.2 Ribosomes are large ribonucleoprotein particles that contain more RNA than protein and dissociate into large and small subunits.

Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

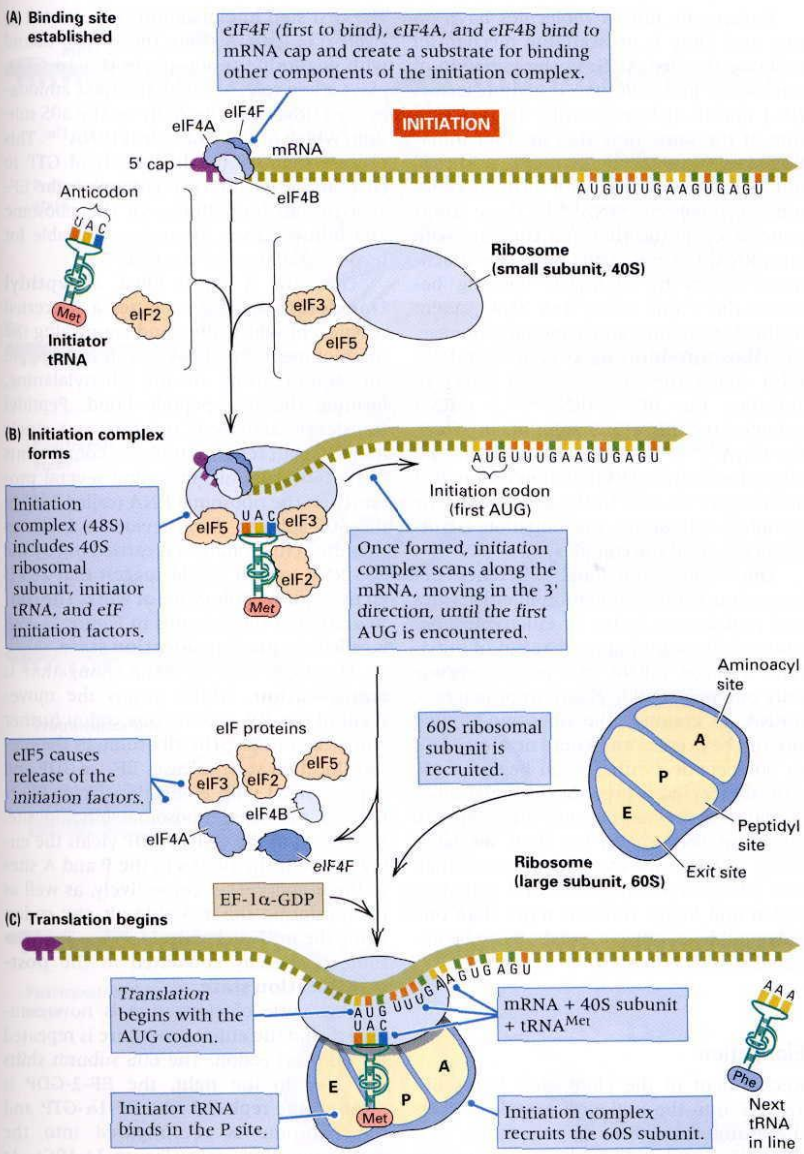


Figure 11.18 Initiation of protein synthesis. (A) The initiation complex forms at the 5' end of the mRNA. (B) This consists of one 40S ribosomal subunit, the initiator $tRNA^{Met}$, and the eIF initiation factors. (C) The initiation complex recruits a 60S ribosomal subunit in which the $tRNA^{Met}$ occupies the P (peptidyl) site of the ribosome. This complex travels along the mRNA until the first AUG is encountered, at which codon translation begins.

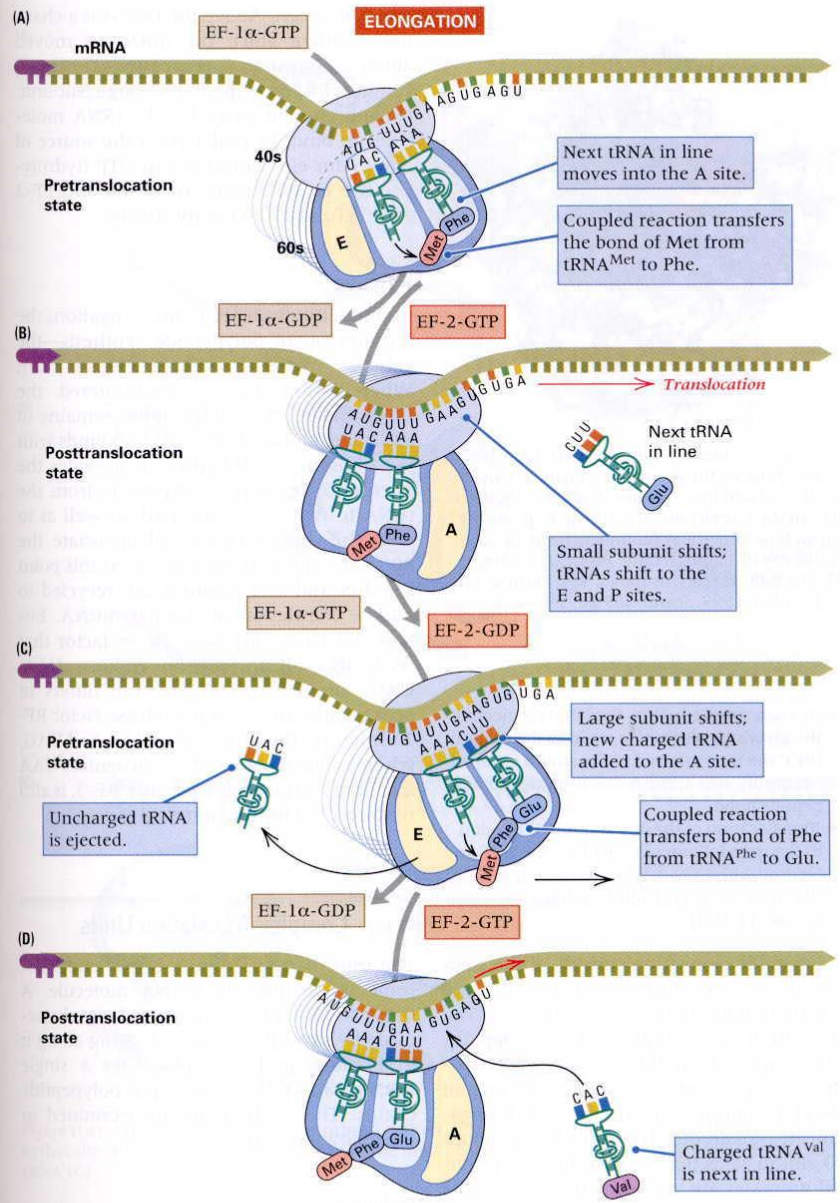


Figure 11.19 Elongation cycle in protein synthesis. (A) Pretranslocation state. (B) Posttranslocation state, in which an uncharged tRNA occupies the E site and the polypeptide is attached to the tRNA in the P site. (C) The function of EF-1α is to release the uncharged tRNA and bring the next charged tRNA into the A site, at which time a peptide bond is formed between the polypeptide and the amino acid held in the A site, in this case Glu. Simultaneously, the 60S subunit is shifted relative to the 40S subunit, re-creating the pretranslocation state. (D) The function of EF-2 is to translocate the 40S ribosome to the next codon, once again generating the posttranslocation state.

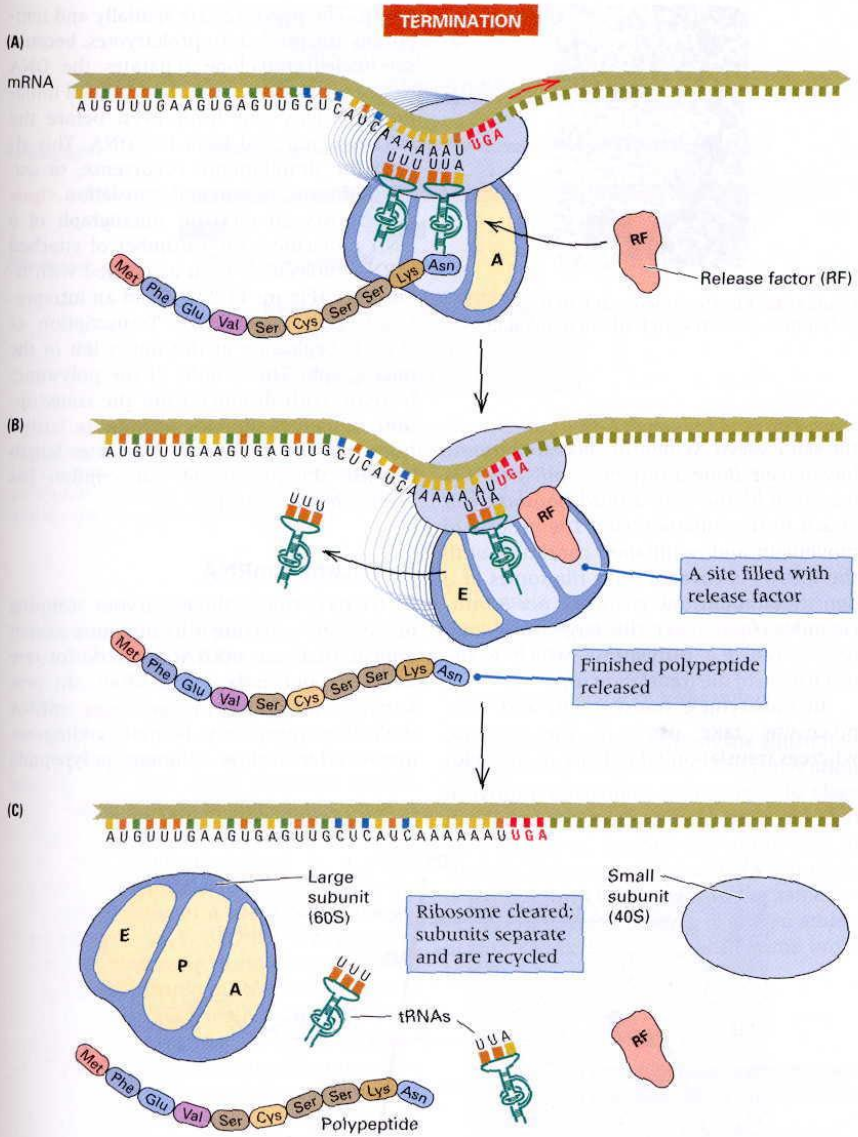


Figure 11.21 Termination of protein synthesis. When a stop codon is reached (A), no tRNA can bind to that site (B), which causes the release of the newly formed polypeptide and the remaining bound tRNA (C).

Regulation of Gene Expression

Prokaryotes

Escherichia coli

Lactose Metabolism

Absence of lactose → Only few molecules of ß-galactosidase per cell

Presence of lactose → about 5000 molecules of ß-galactosidase per cell

Not enzyme is inhibited, enzyme synthesis is affected

Detailed biochemical and genetic analysis

Jacob, Monod, Pardee → Nobel prize

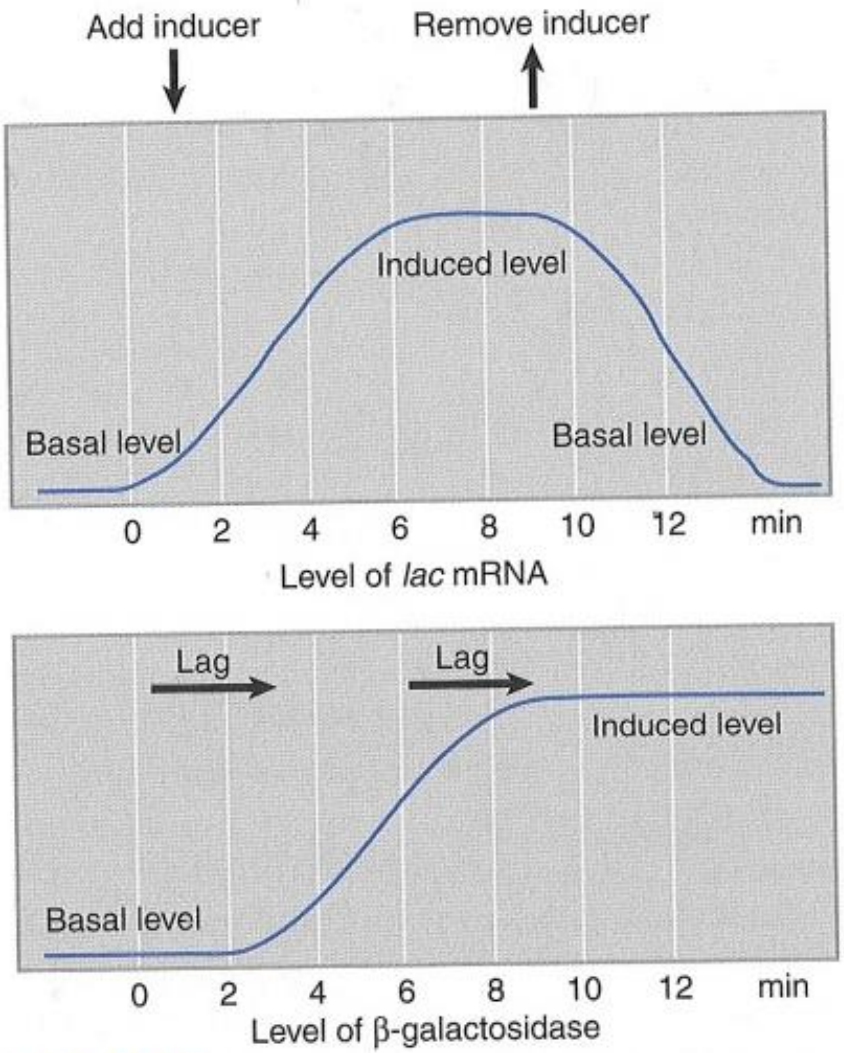
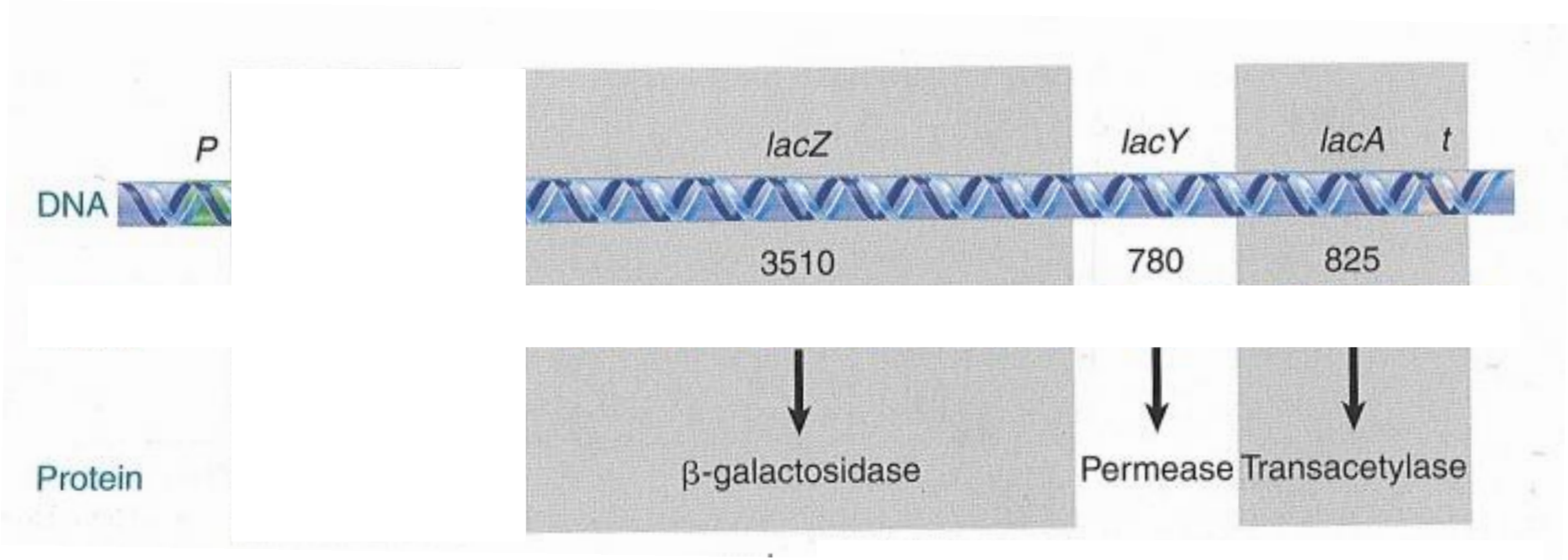


FIGURE 26.7 Addition of inducer results in rapid induction of *lac* mRNA and is followed after a short lag by synthesis of the enzymes; removal of inducer is followed by rapid cessation of synthesis.

Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

lac-Operon

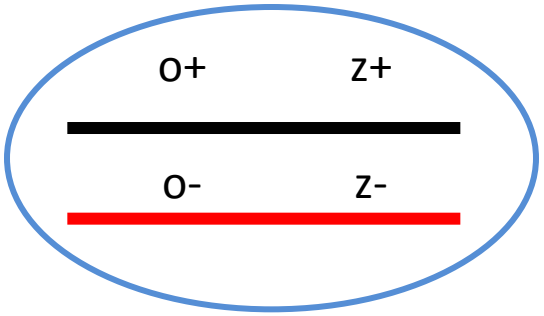


Ort I

Ort O

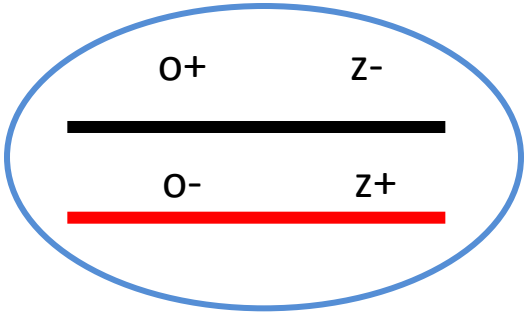
Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

Heterogenote analysis



Cis-configuration

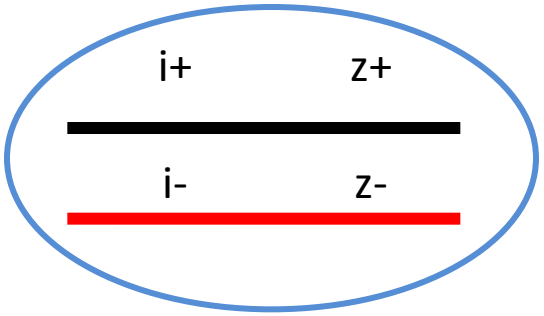
inducible



Trans-configuration

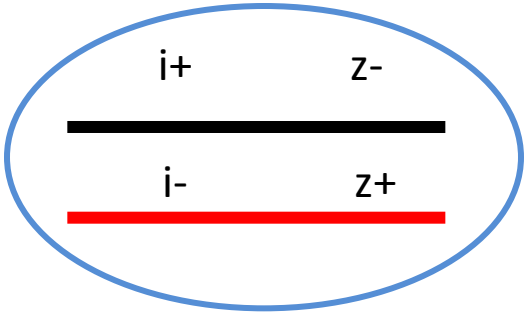
constitutive

Heterogenote analysis



Cis-configuration

inducible



Trans-configuration

inducible

Model for behaviour of heterogenotes

lacO → located adjacent to *lacZ,* mutation in l*acO* results in loss of regulatory function when connected to *lacZ,* no complementation by wt-allele in trans

lacI → located upstream of *lacZ,* mutation in *lacI* results in maintenance of regulatory function in both configurations to *lacZ* complementation by wt-allele

lacO → DNA locus, mobile factor binds there and represses synthesis

lacI → encodes a mobile factor (= protein) which binds at *lacO*

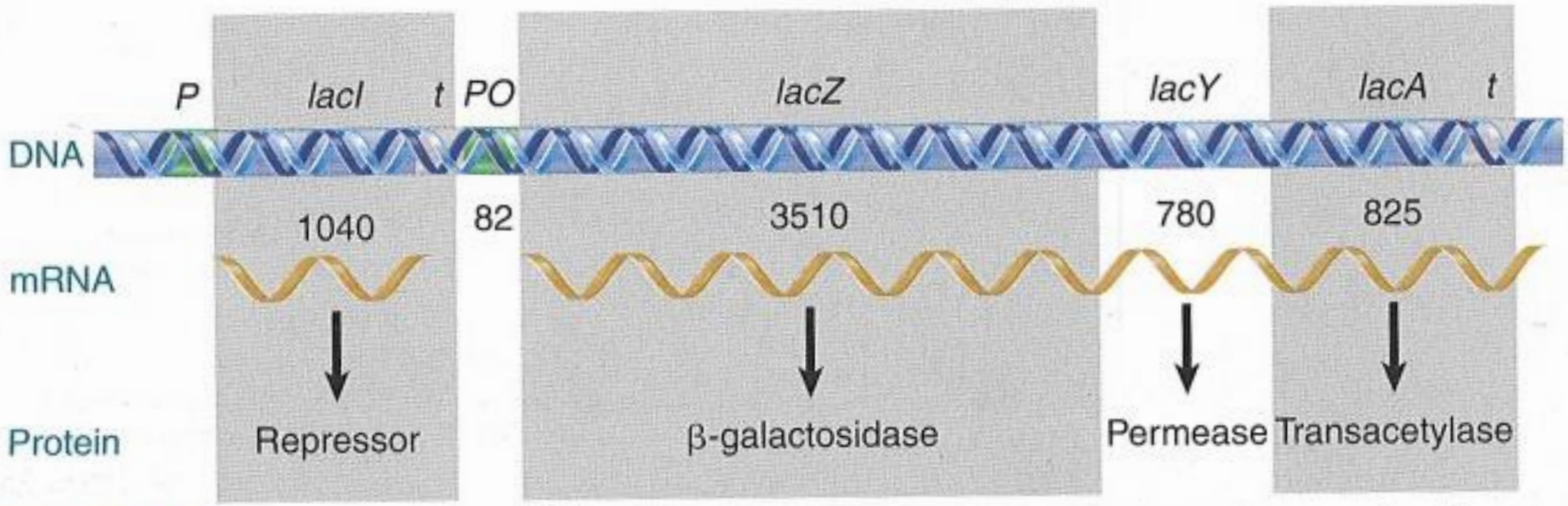
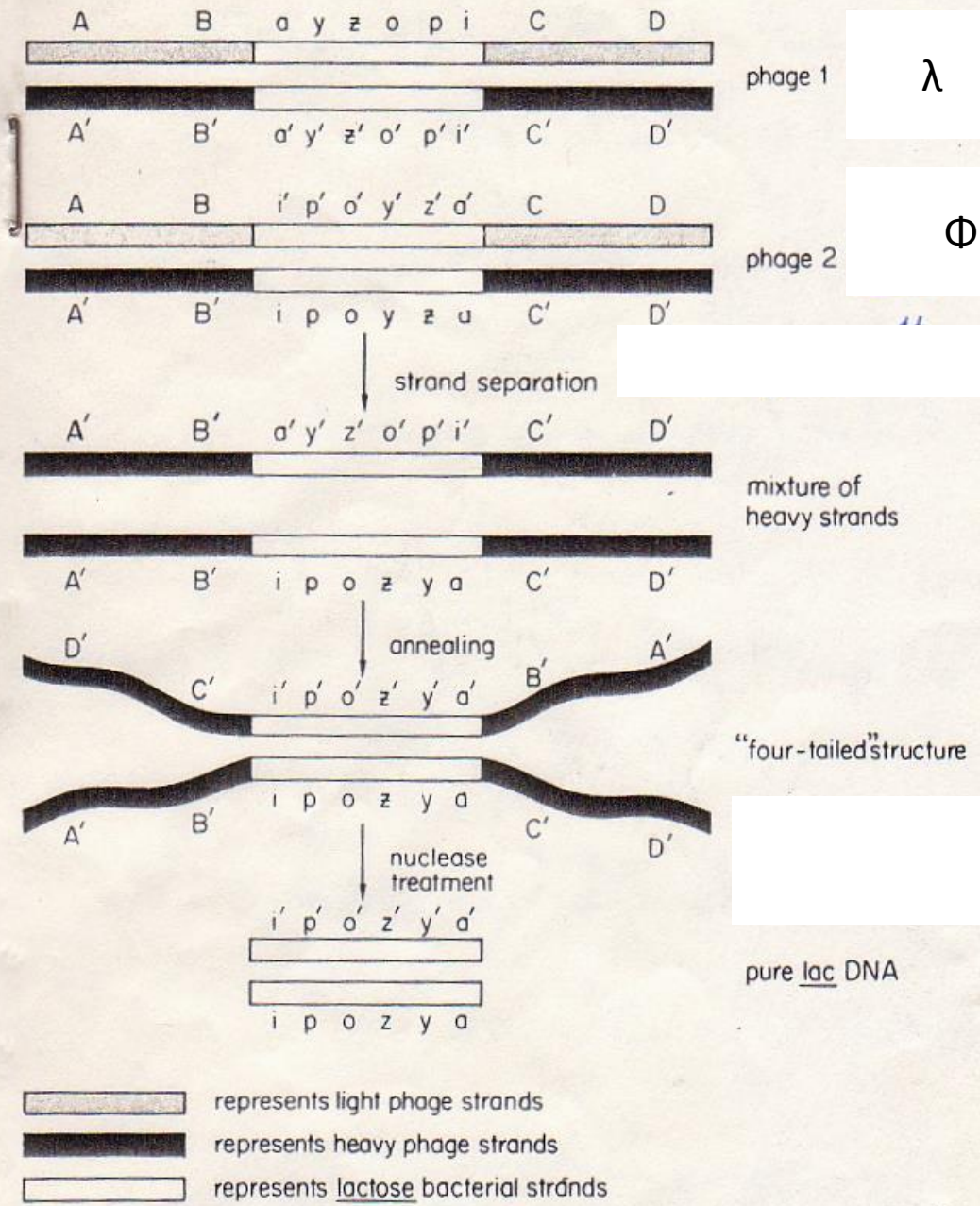


FIGURE 26.5 The *lac* operon occupies ~6000 bp of DNA. At the left the *lacI* gene has its own promoter and terminator. The end of the *lacI* region is adjacent to the *lacZYA* promoter, *P*. Its operator, *O*, occupies the first 26 bp of the transcription unit. The long *lacZ* gene starts at base 39, and is followed by the *lacY* and *lacA* genes and a terminator.

Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning



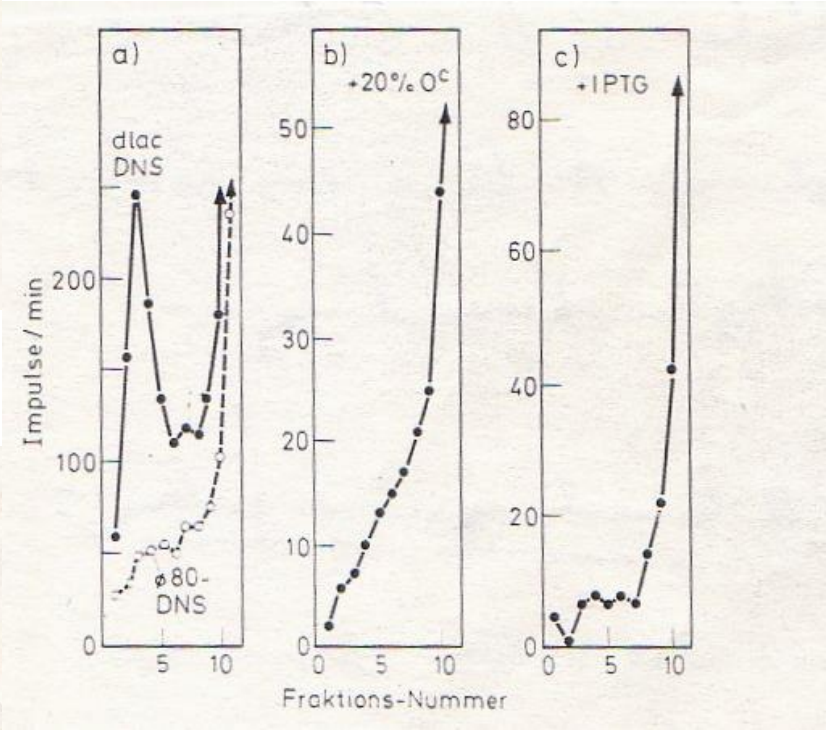
λ

Φ 80

Gene isolation
lac operon

Isolation of Lac Repressor
$lacI^q$ mutant

Binding studies



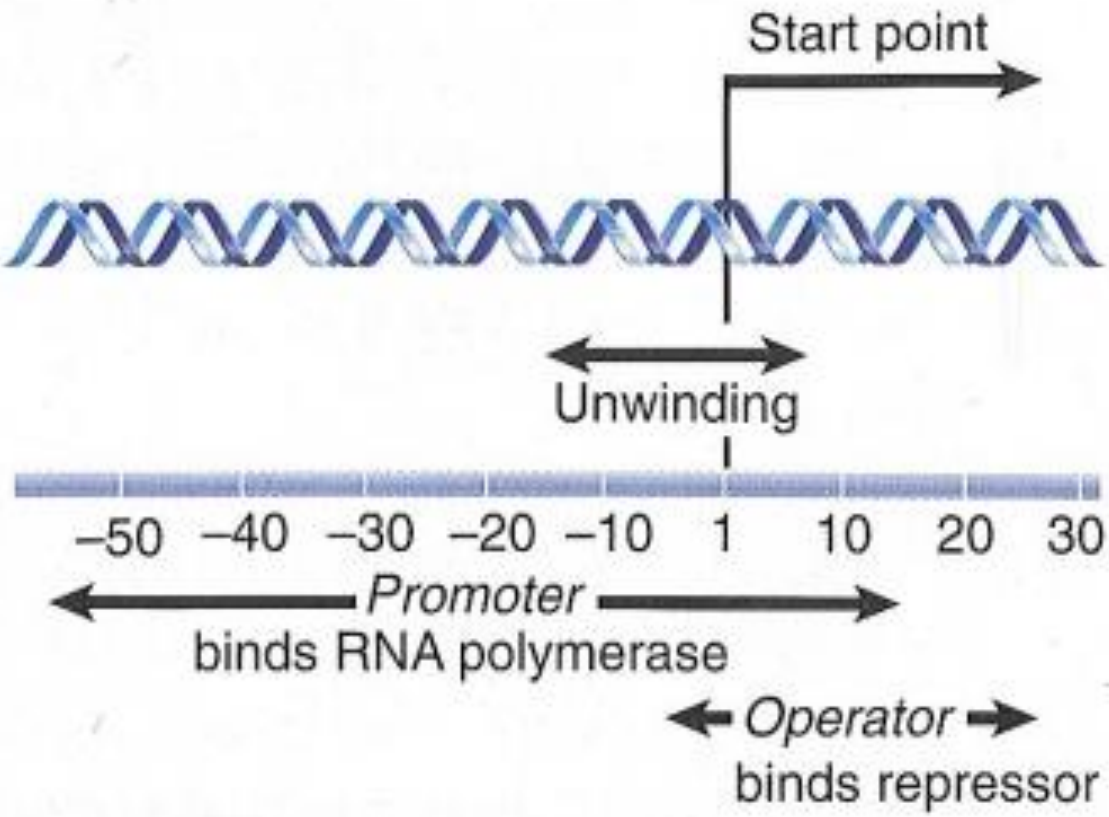
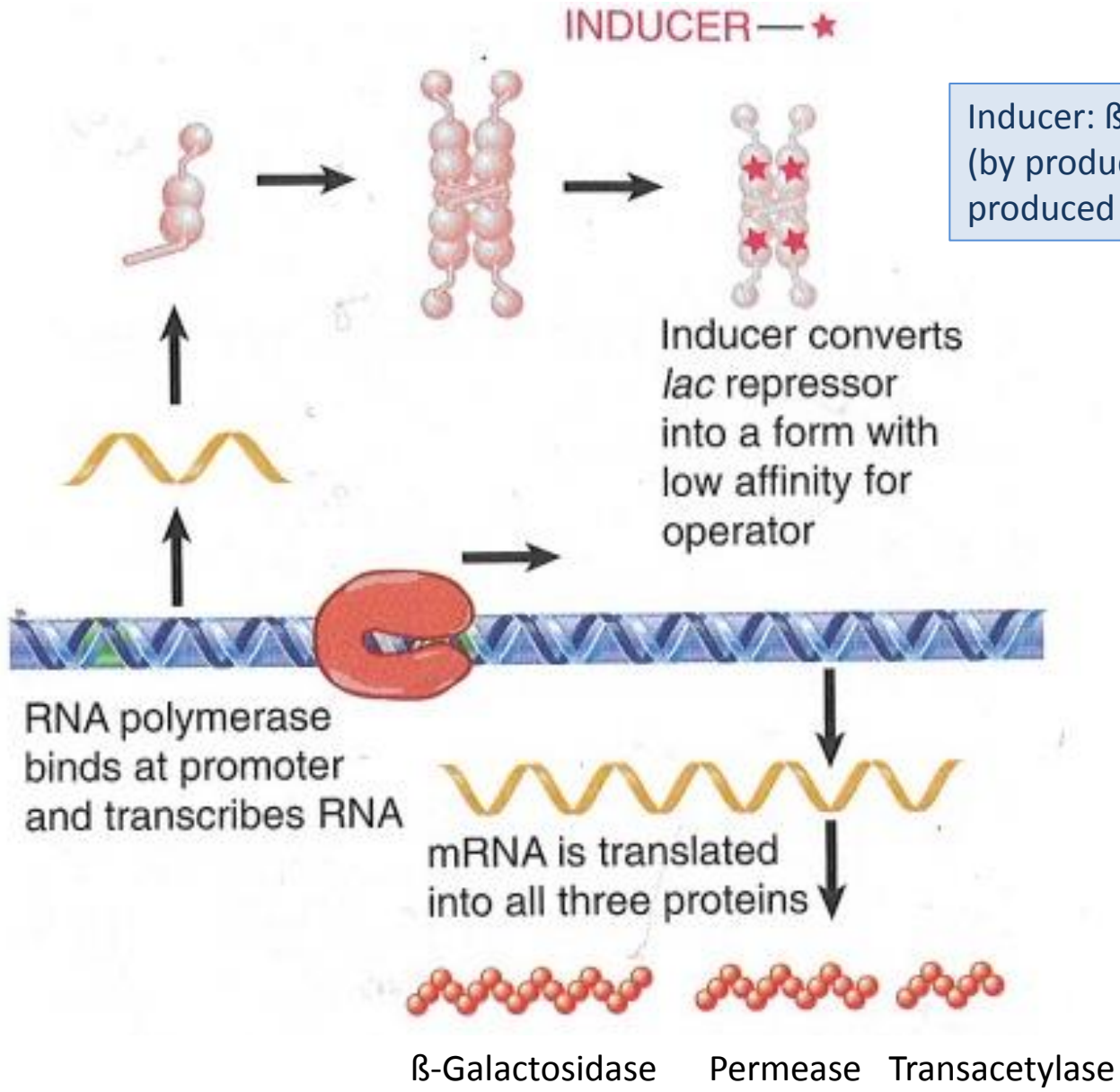


FIGURE 26.6 *lac* repressor and RNA polymerase bind at sites that overlap around the transcription start point of the *lac* operon.

Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

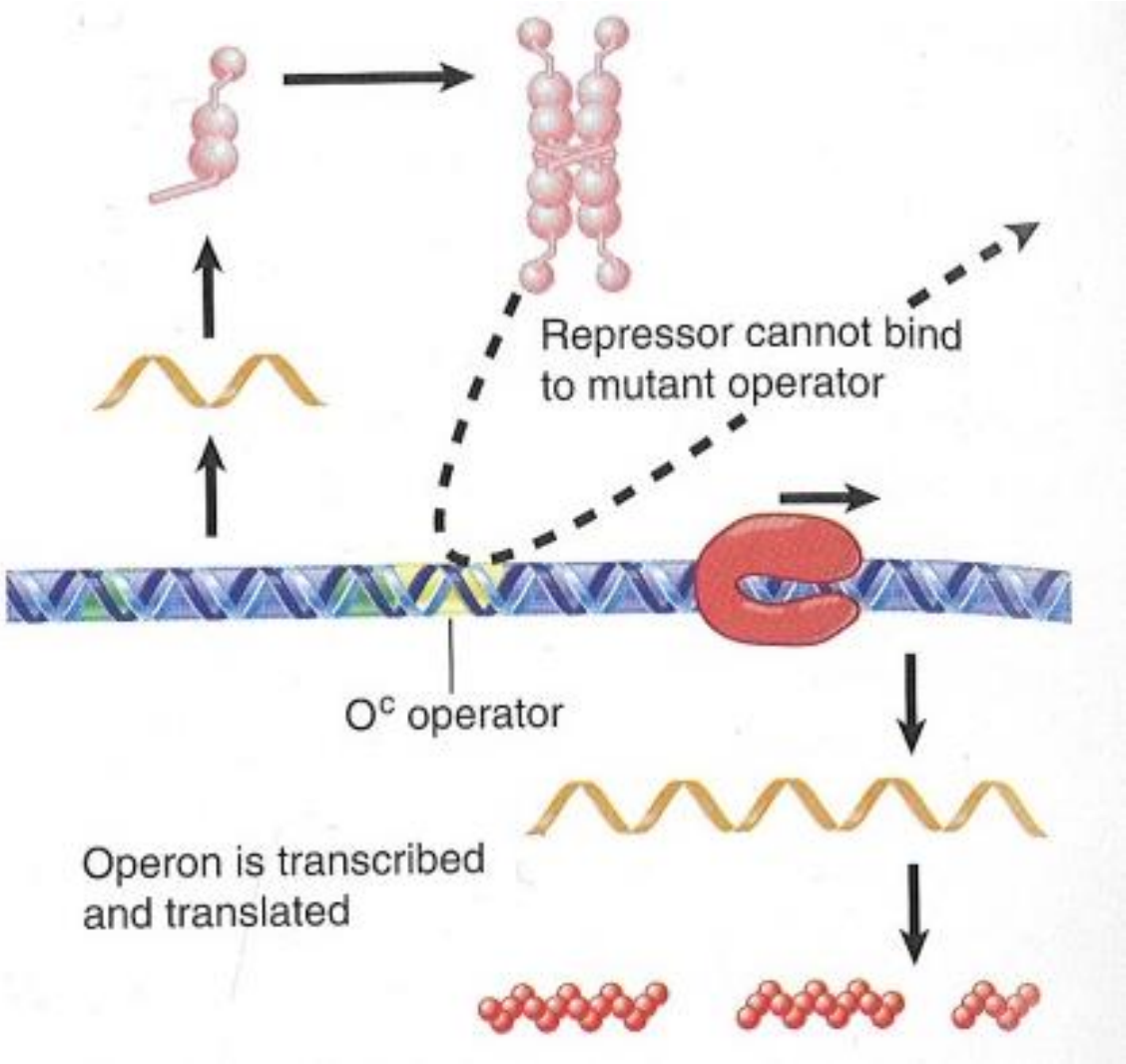


Inducer: ß-1,6- allolactose
(by product of ß-galactosidase produced by transglucosylation)

Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

Mutant O^c

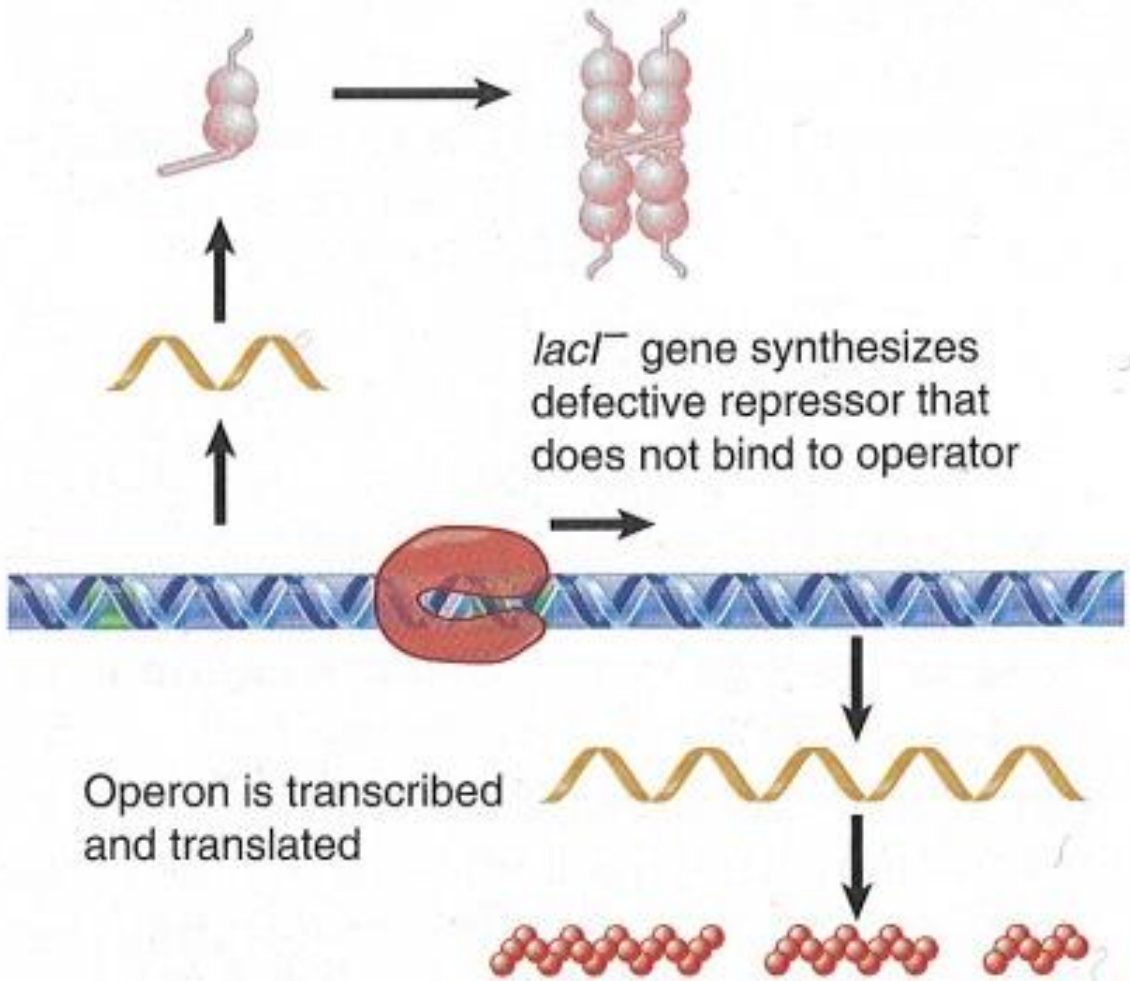
Mutation in *lacO* prevents binding of LacI Repressor protein to Operator



Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

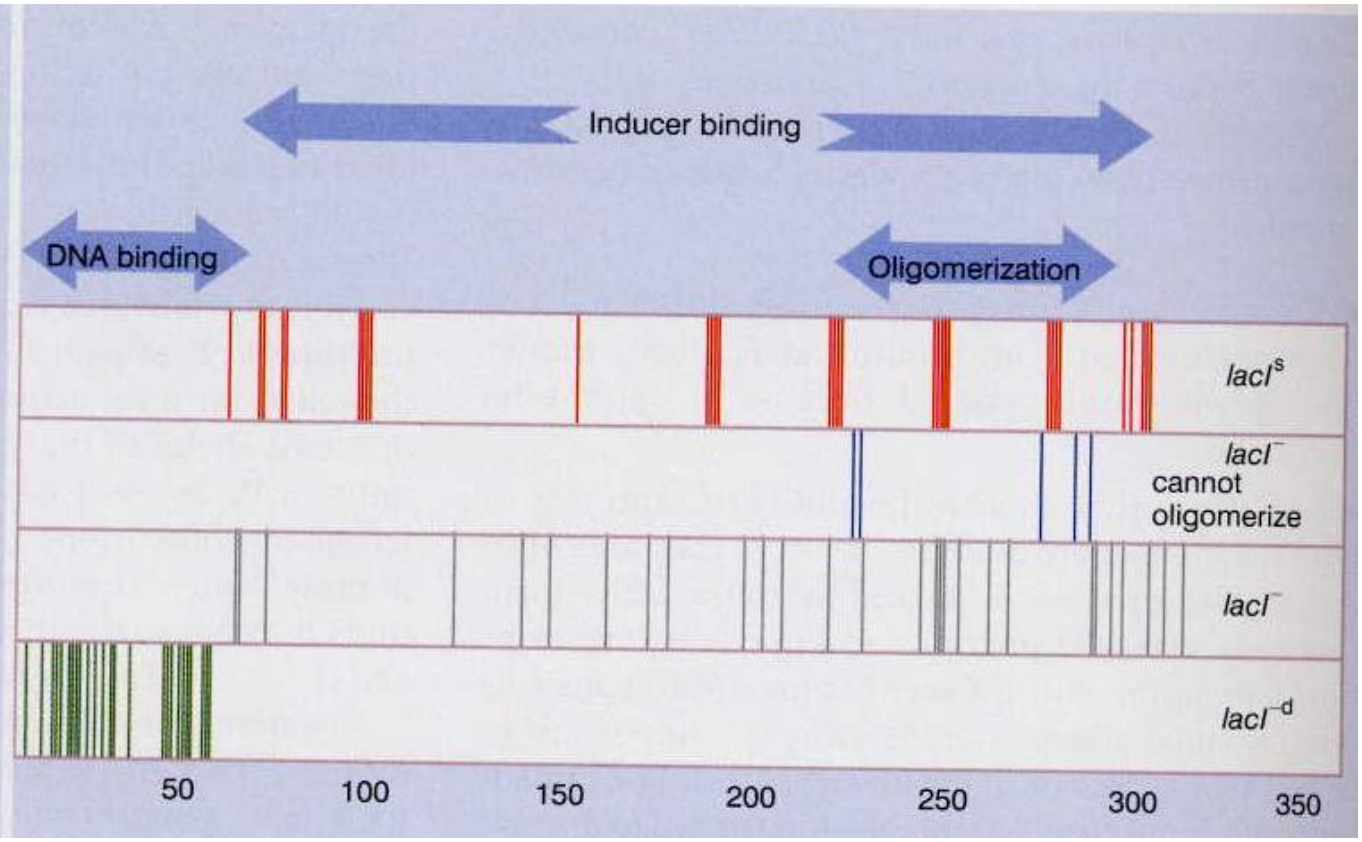
Mutant *I* $^{-}$

Mutation in *lacI* no binding capacity of LacI repressor protein



Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

Figure 10.9 Mutations map the regions of the *lacI* gene responsible for different functions. The DNA-binding domain is identified by *lacI*$^{-d}$ mutations at the N-terminal region; *lacI*$^{-}$ mutations unable to form tetramers are located between residues 220–280; other *lacI*$^{-}$ mutations occur throughout the gene; *lacI*^{s} mutations occur in regularly spaced clusters between residues 62–300.



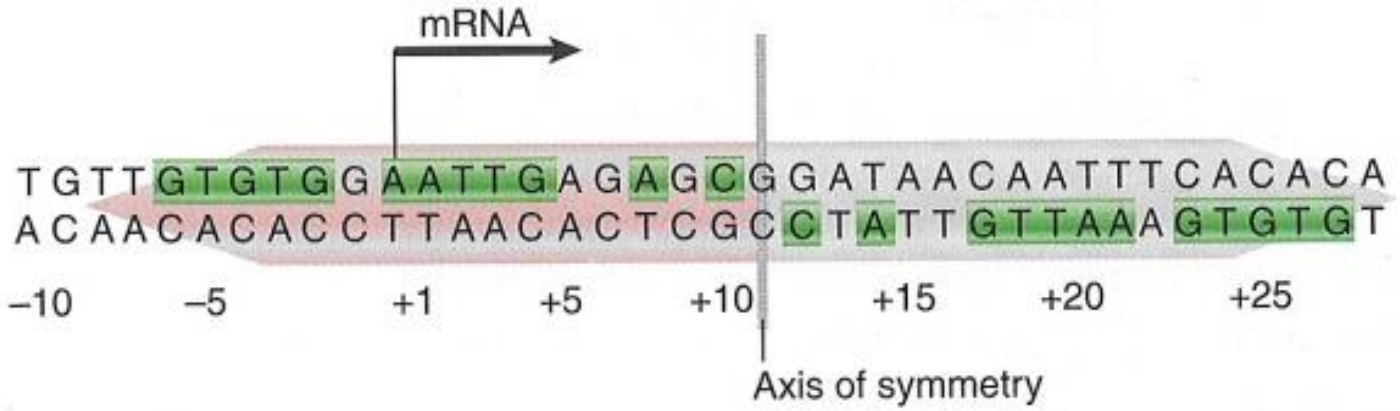


FIGURE 26.17 The *lac* operator has a symmetrical sequence. The sequence is numbered relative to the start point for transcription at +1. The pink arrows to the left and to the right identify the two dyad repeats. The green blocks indicate the positions of identity.

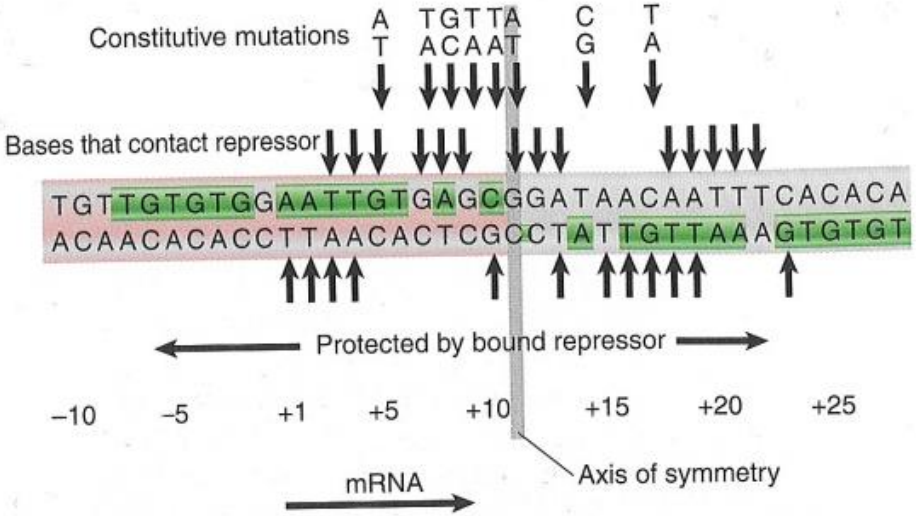


FIGURE 26.19 Bases that contact the repressor can be identified by chemical crosslinking or by experiments to see whether modifications prevent binding. They identify positions on both strands of DNA extending from +1 to +23. Constitutive mutations occur at 8 positions in the operator between +5 and +17.

Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

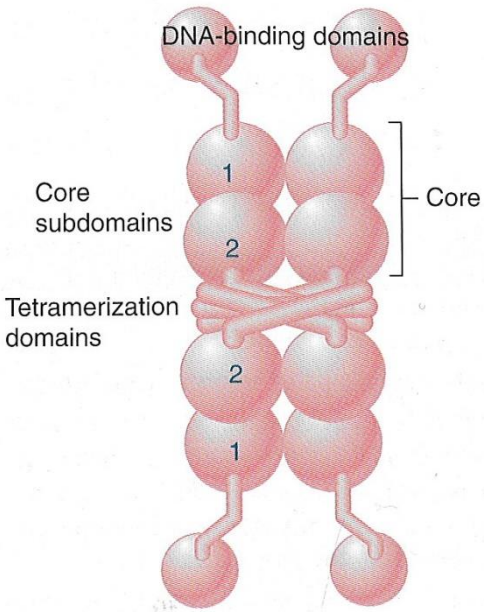


FIGURE 26.15 The repressor tetramer consists of two dimers. Dimers are held together by contacts involving core subdomains 1 and 2 as well as by the tetramerization helix. The dimers are linked into the tetramer by the tetramerization interface.

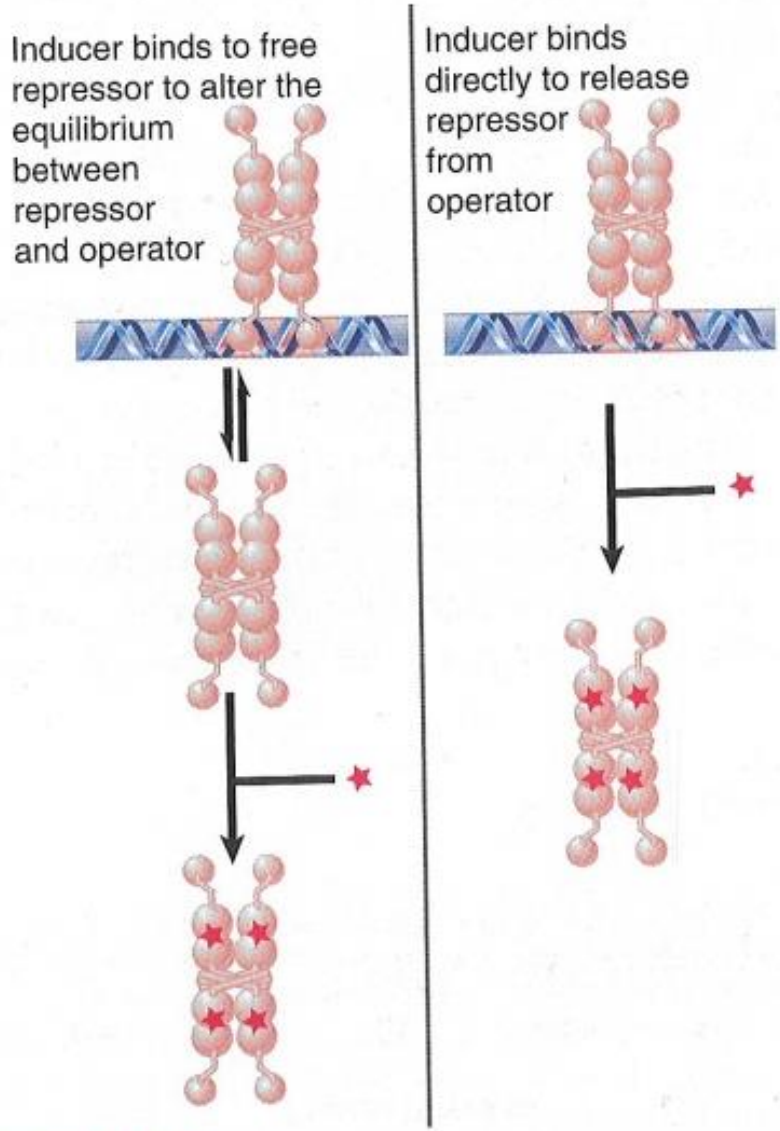


FIGURE 26.20 Does the inducer bind to the free repressor to upset an equilibrium (left) or directly to repressor bound at the operator (right)?

Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

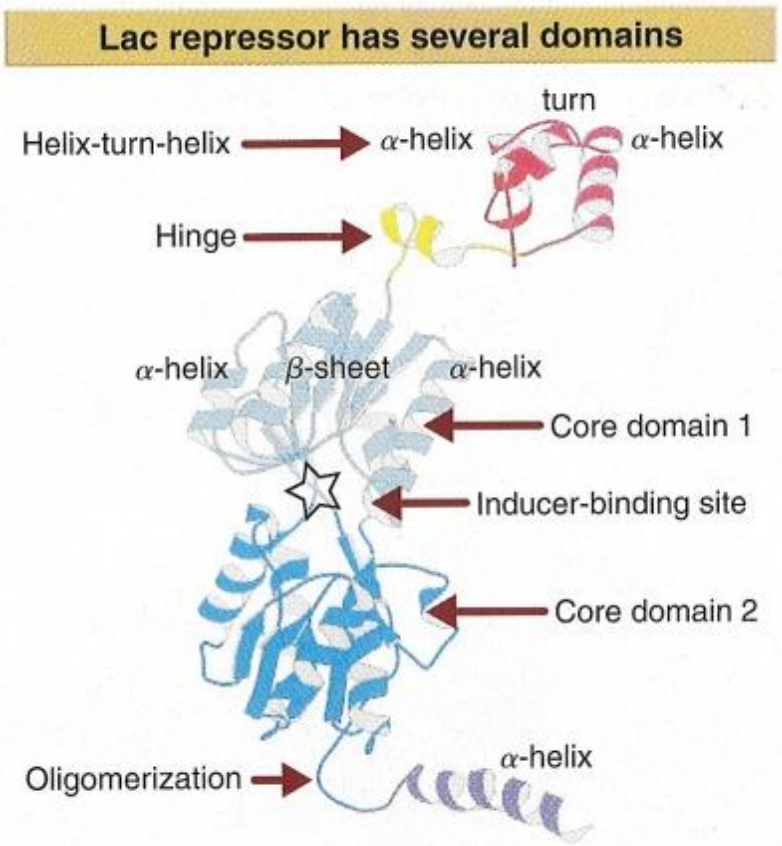


Figure 12.11 The structure of a monomer of Lac repressor identifies several independent domains. Photograph kindly provided by Mitchell Lewis, Dept. of Biochemistry & Biophysics, University of Pennsylvania.

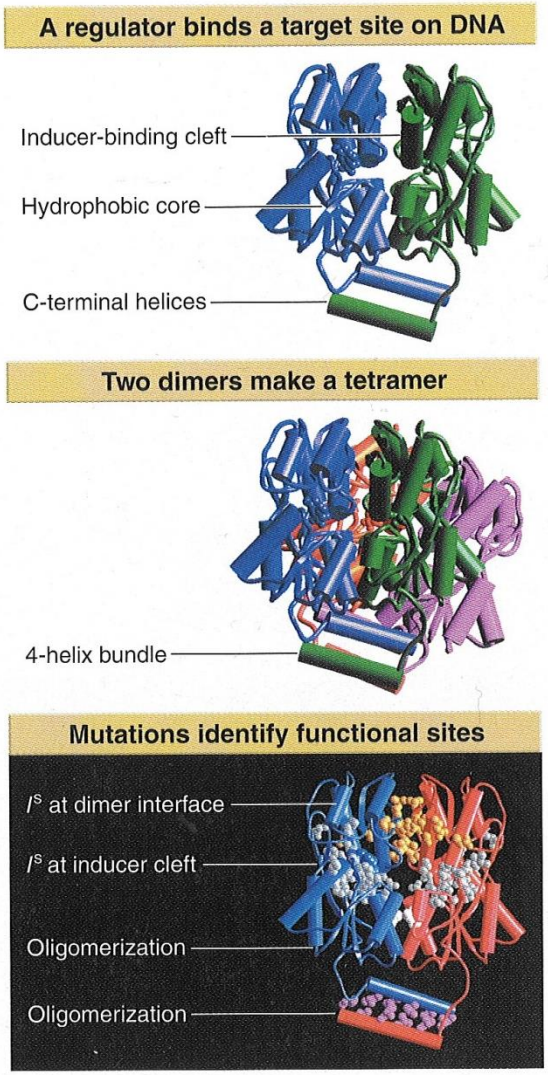


Figure 12.12 The crystal structure of the core region of Lac repressor identifies the interactions between monomers in the tetramer. Each monomer is identified by a different color. Mutations are colored as: dimer interface—yellow; inducer-binding—blue; oligomerization—white and purple. Photographs kindly provided by Alan Friedman.

Taken from: B. Lewin, Essential Genes, Pearson Ed. International

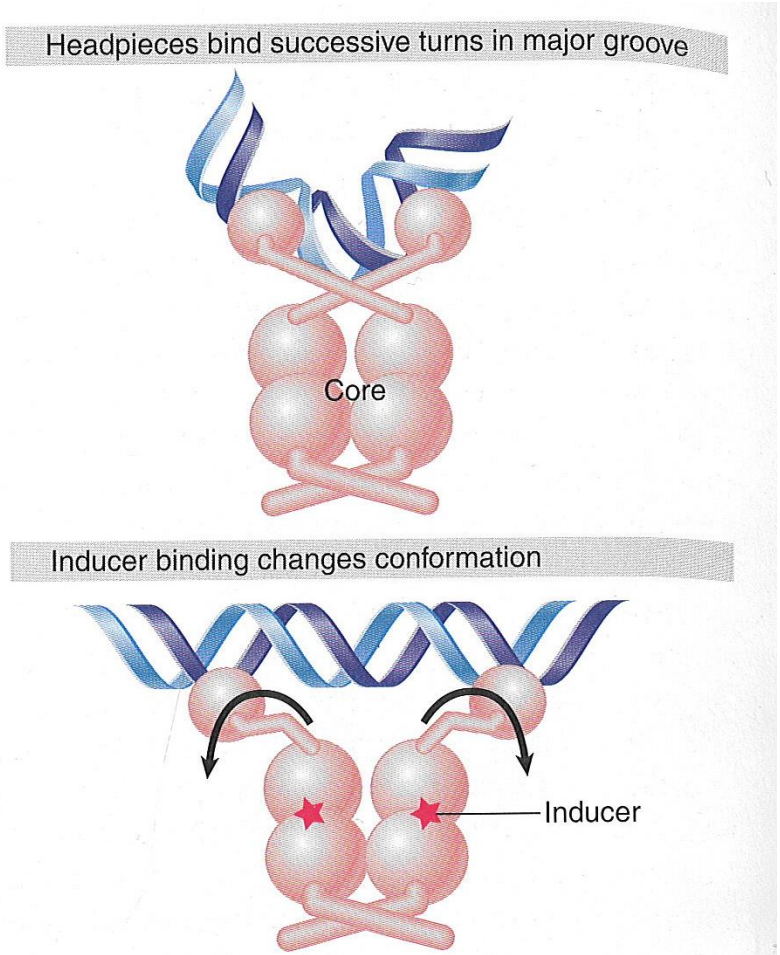
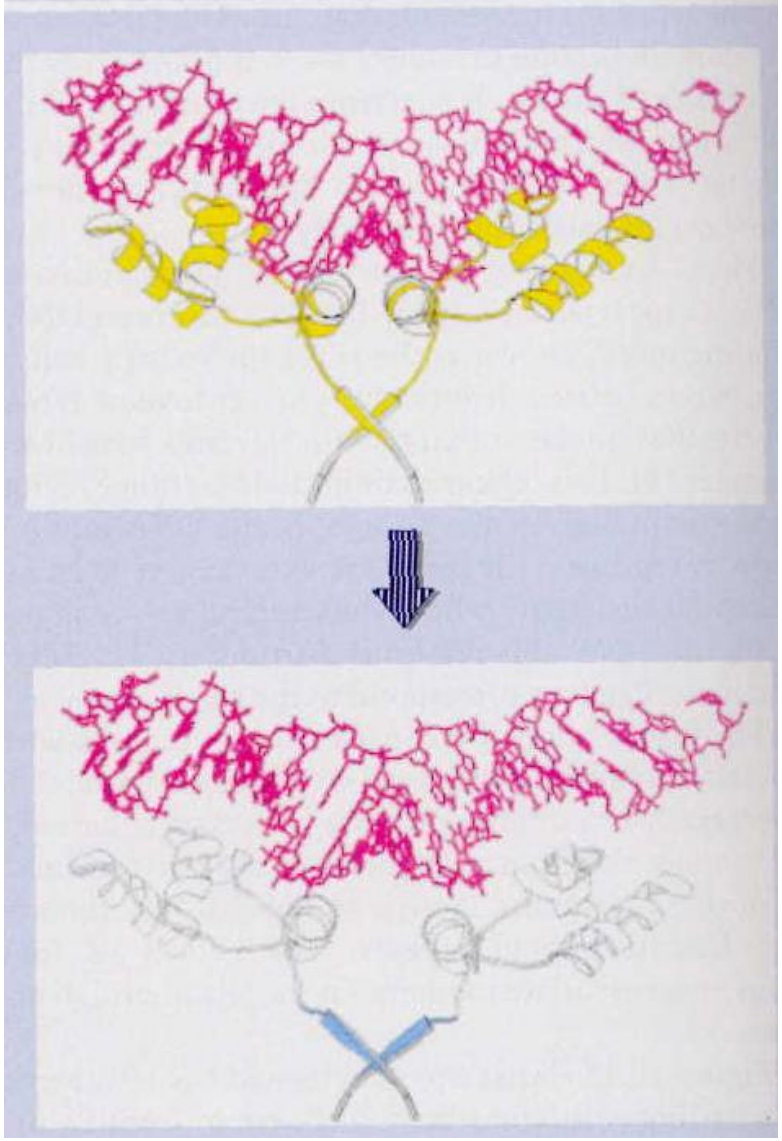


FIGURE 26.18 The inducer changes the structure of the core so that the headpieces of a repressor dimer are no longer in an orientation with high affinity for the operator.

Figure 10.14 Inducer changes the structure of the core so that the headpieces of a repressor dimer are no longer in an orientation that permits binding to DNA. Photographs kindly provided by Mitchell Lewis.



Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

FIGURE 26.22 When a repressor tetramer binds to two operators, the stretch of DNA between them is forced into a tight loop. (The blue structure in the center of the looped DNA represents CRP, which is another regulator protein that binds in this region.) Reproduced from M. Lewis et al., *Science* 271 (1996): 1247–1254 [http://www.sciencemag.org]. Reprinted with permission from AAAS. Photo courtesy of Ponzy Lu, University of Pennsylvania.

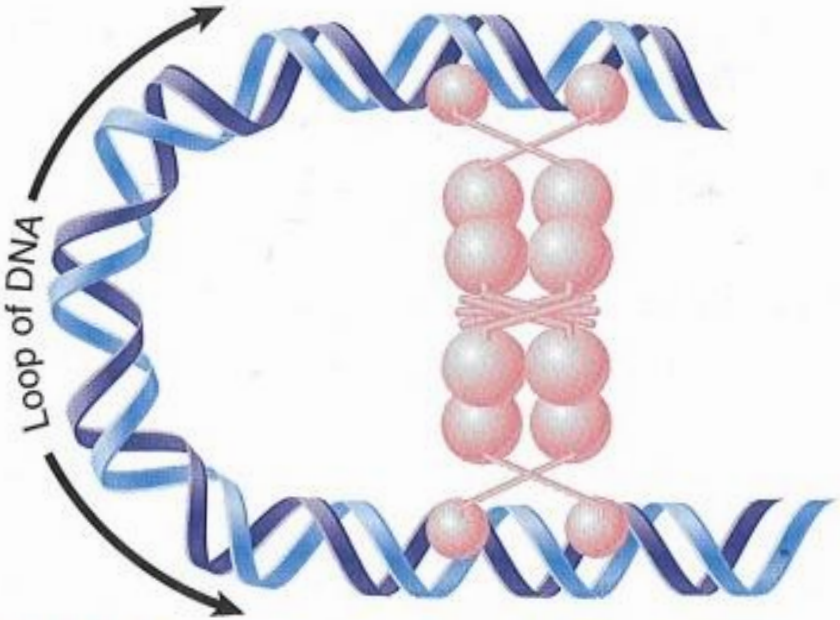


FIGURE 26.21 If both dimers in a repressor tetramer bind to DNA, the DNA between the two binding sites is held in a loop.

Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

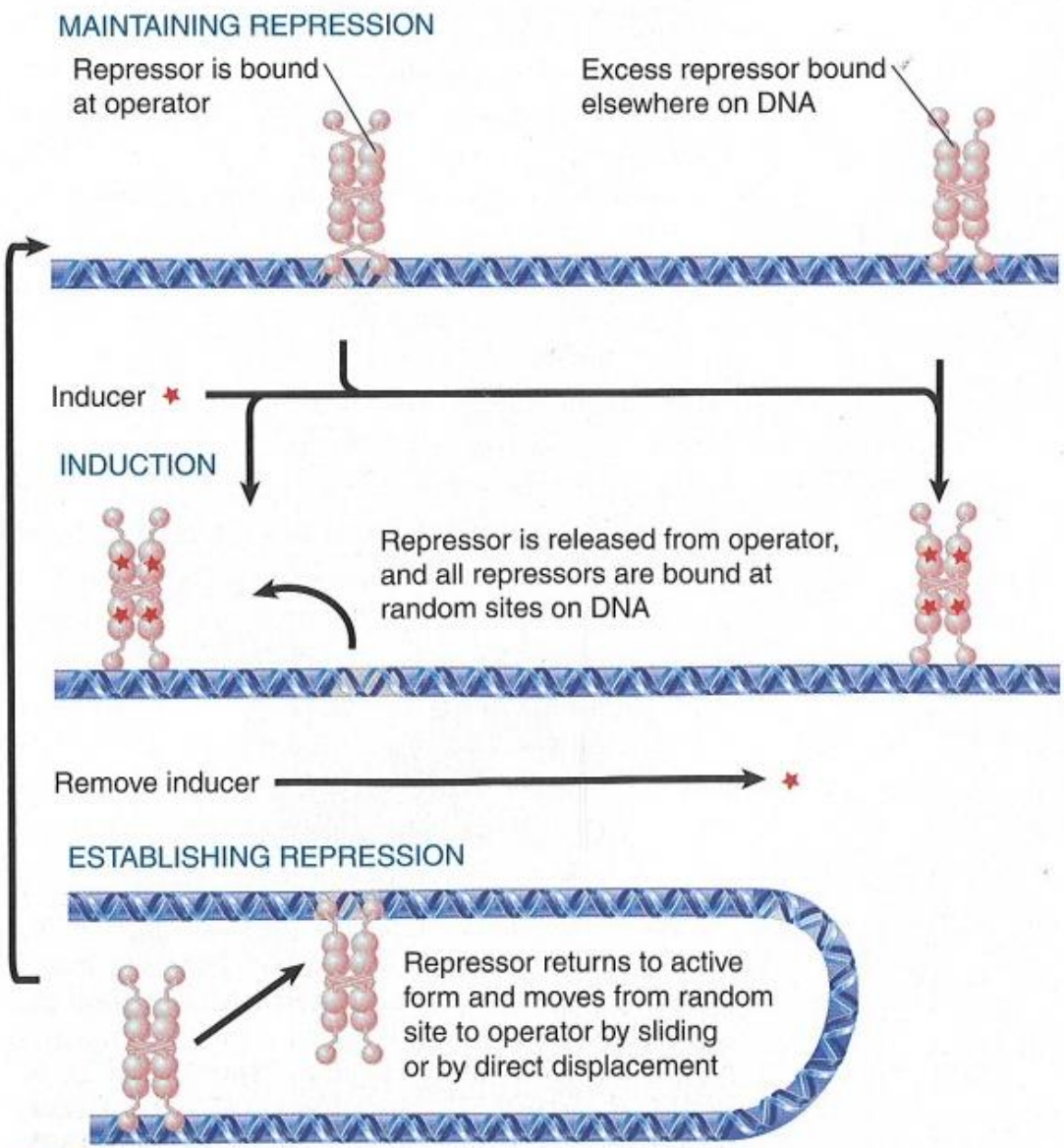


FIGURE 26.24 Virtually all the repressor in the cell is bound to DNA.

LacI repressor has general low affinity to DNA → Unspecific weak binding

LacI repressor has high affinity to specific operon Region on DNA → Specific strong binding

Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

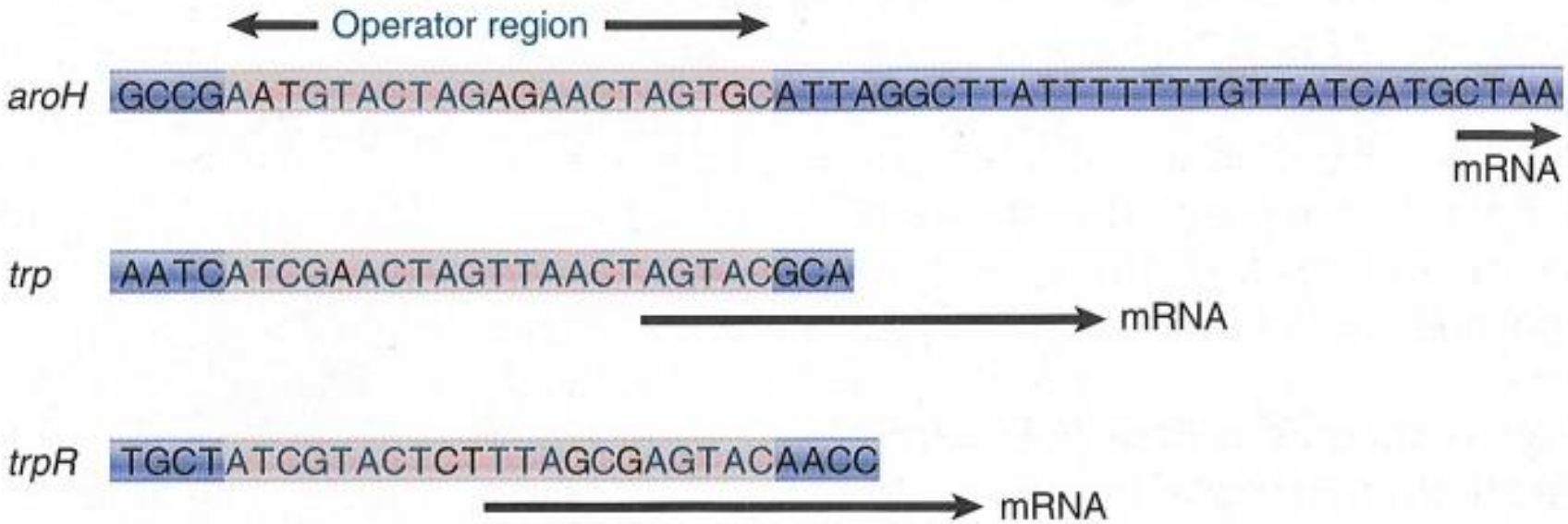


FIGURE 26.31 The *trp* repressor recognizes operators at three loci. Conserved bases are shown in red. The location of the start point and mRNA varies, as indicated by the black arrows.

Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

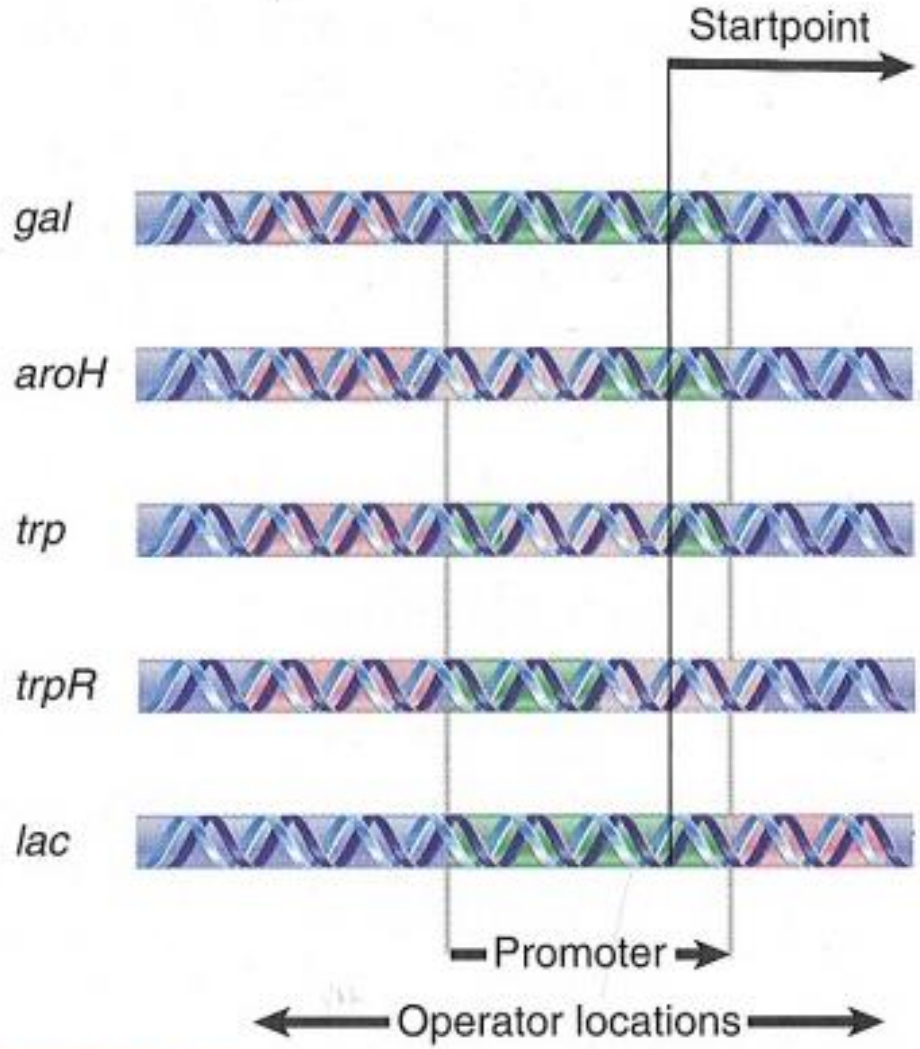
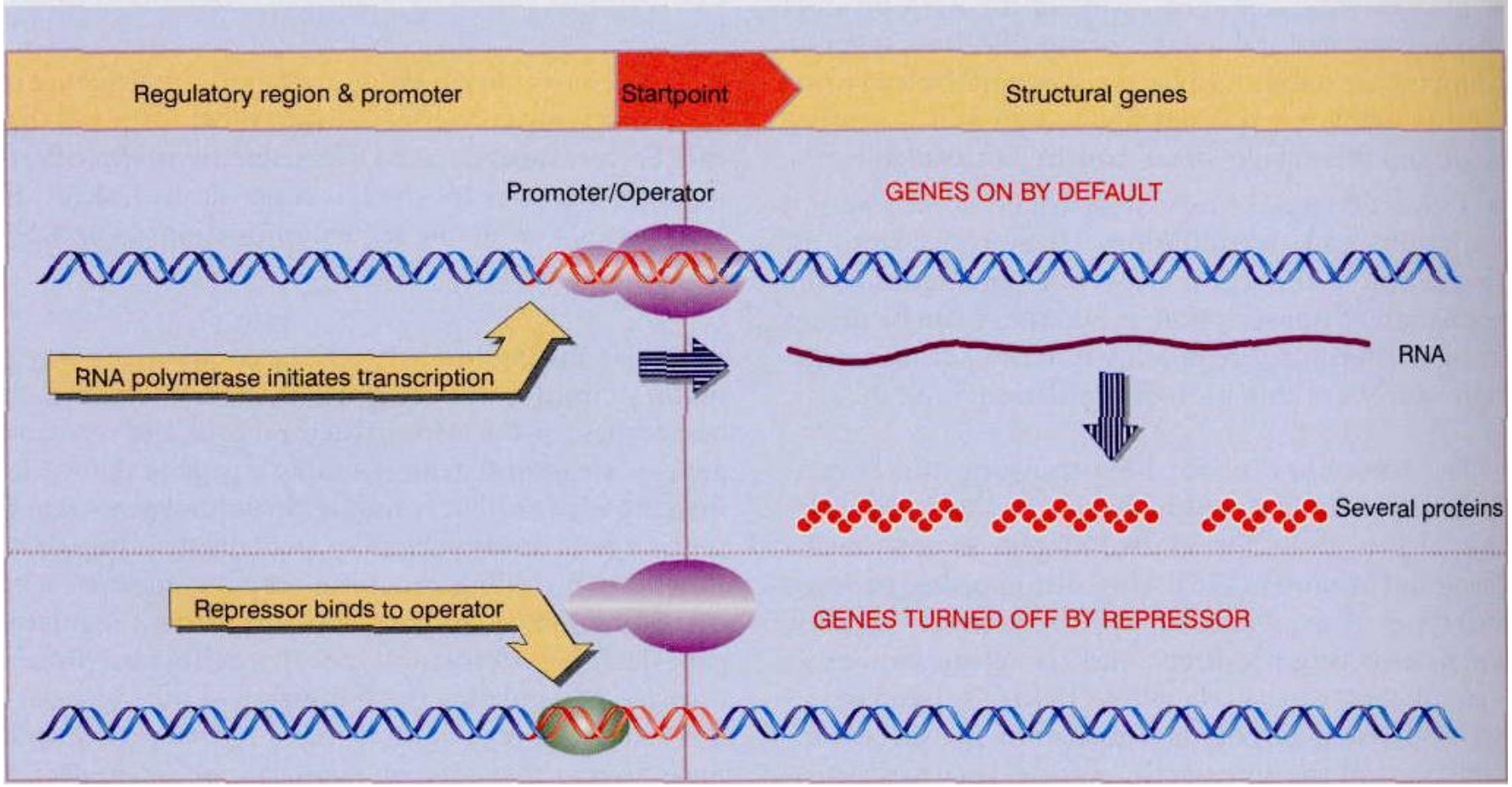


FIGURE 26.32 Operators may lie at various positions relative to the promoter.

Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

Negative Regulation

Figure 10.1 *Overview*: in negative control, a *trans*-acting repressor binds to the *cis*-acting operator to turn off transcription. In prokaryotes, multiple genes are controlled coordinately.



1.12.15

Negative Regulation

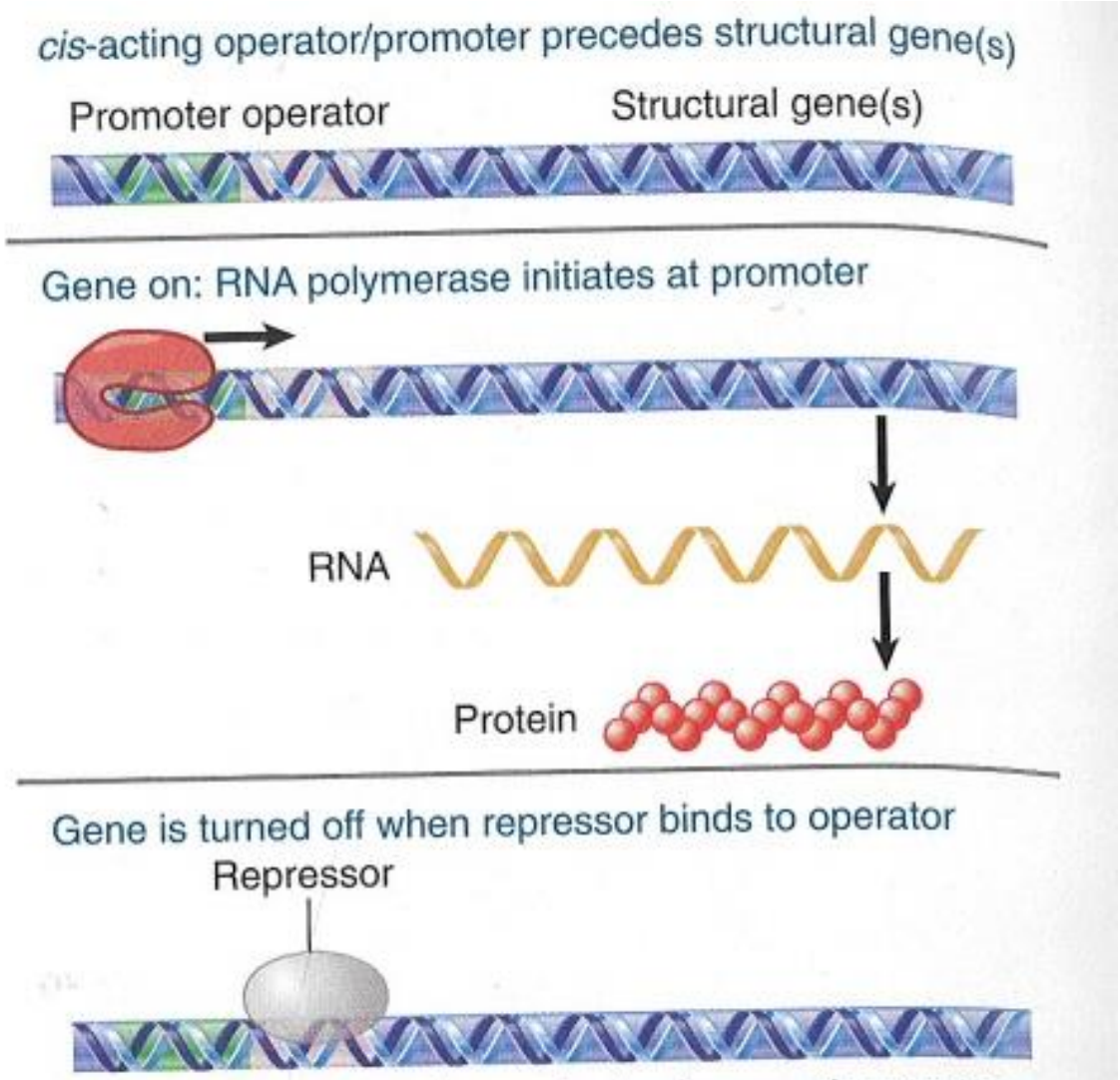


FIGURE 26.2 In negative control, a *trans*-acting repressor binds to the *cis*-acting operator to turn off transcription.

Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

Positive Regulation

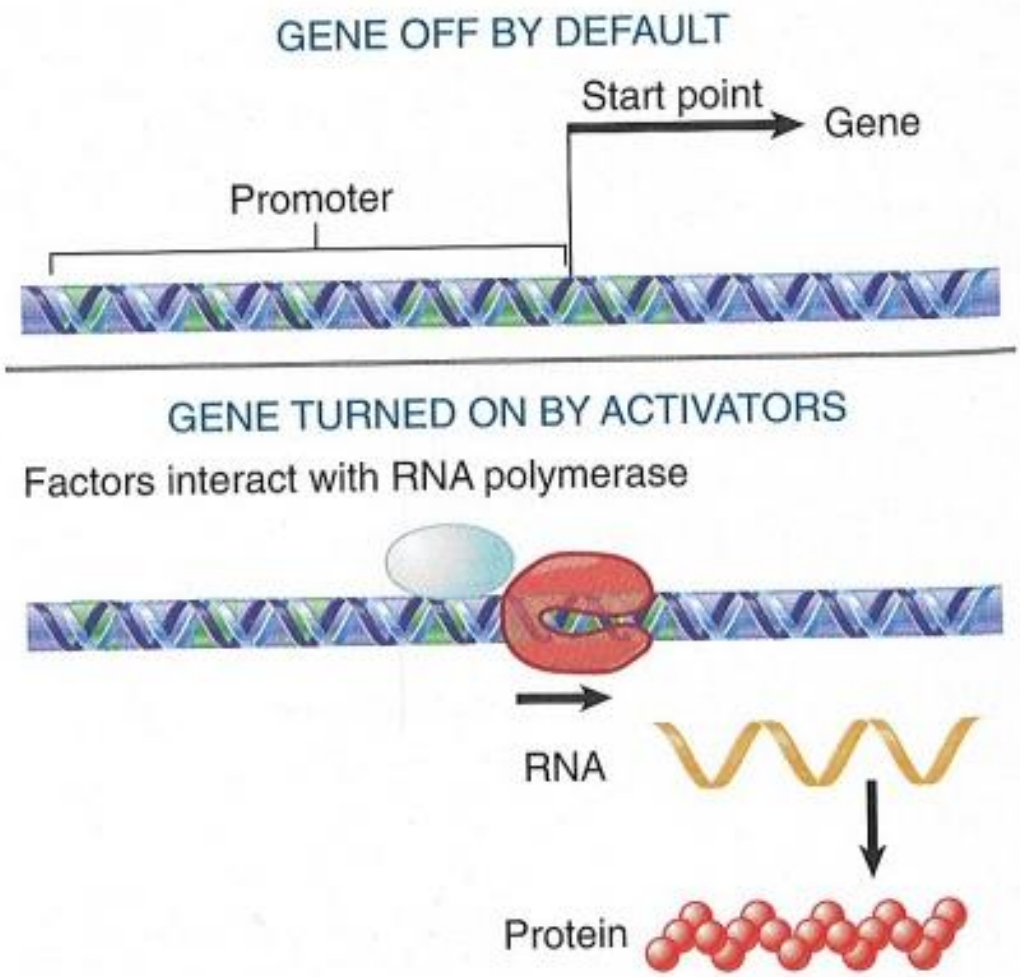


FIGURE 26.3 In positive control, a *trans*-acting factor must bind to the *cis*-acting site in order for RNA polymerase to initiate transcription at the promoter.

Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

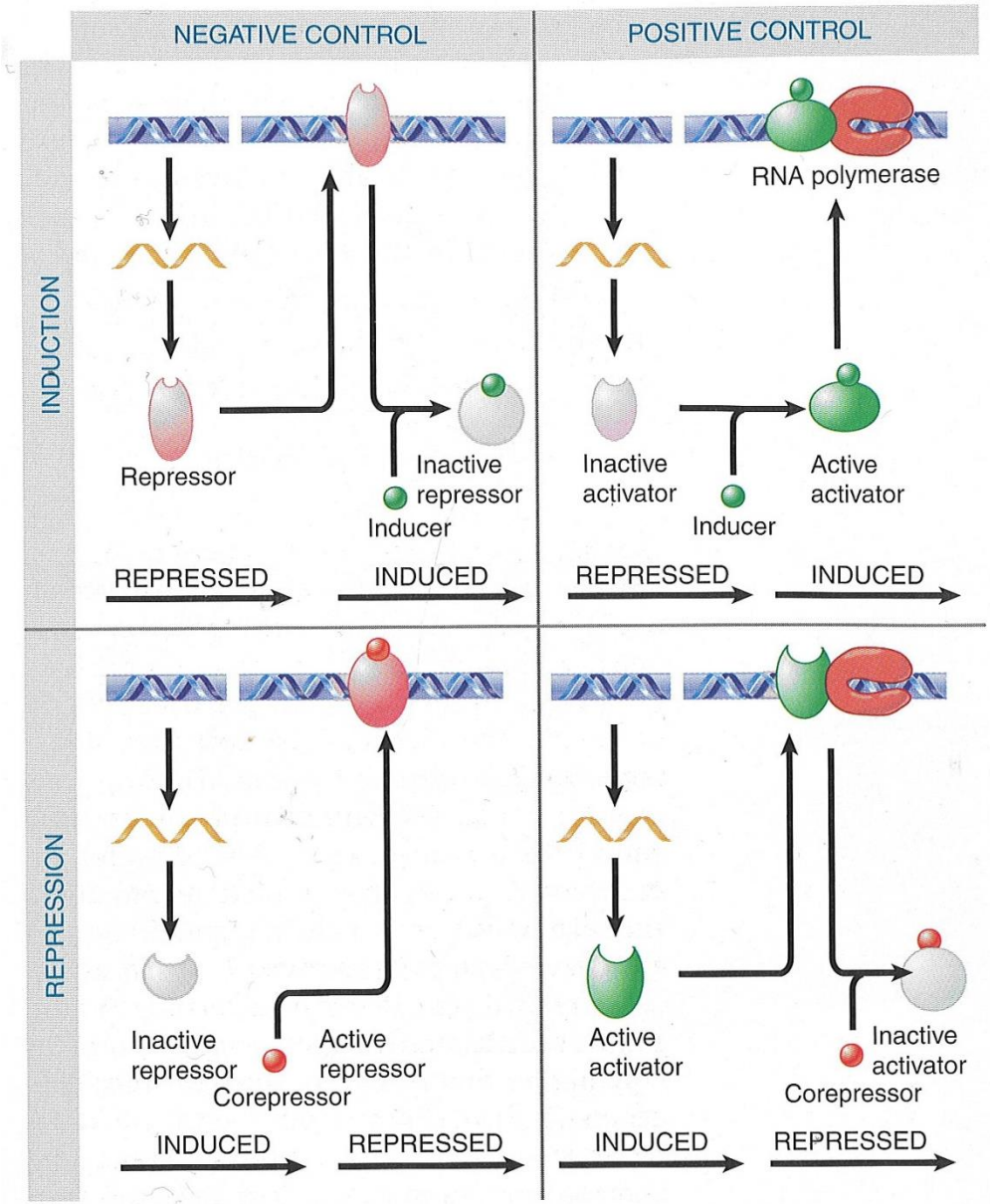


FIGURE 26.4 Regulatory circuits can be designed from all possible combinations of positive and negative control with inducible and repressible control.

Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

Influence of Glucose on expression of *lac* Operon

Glucose controls import of lactose

and of other alternative carbon sources

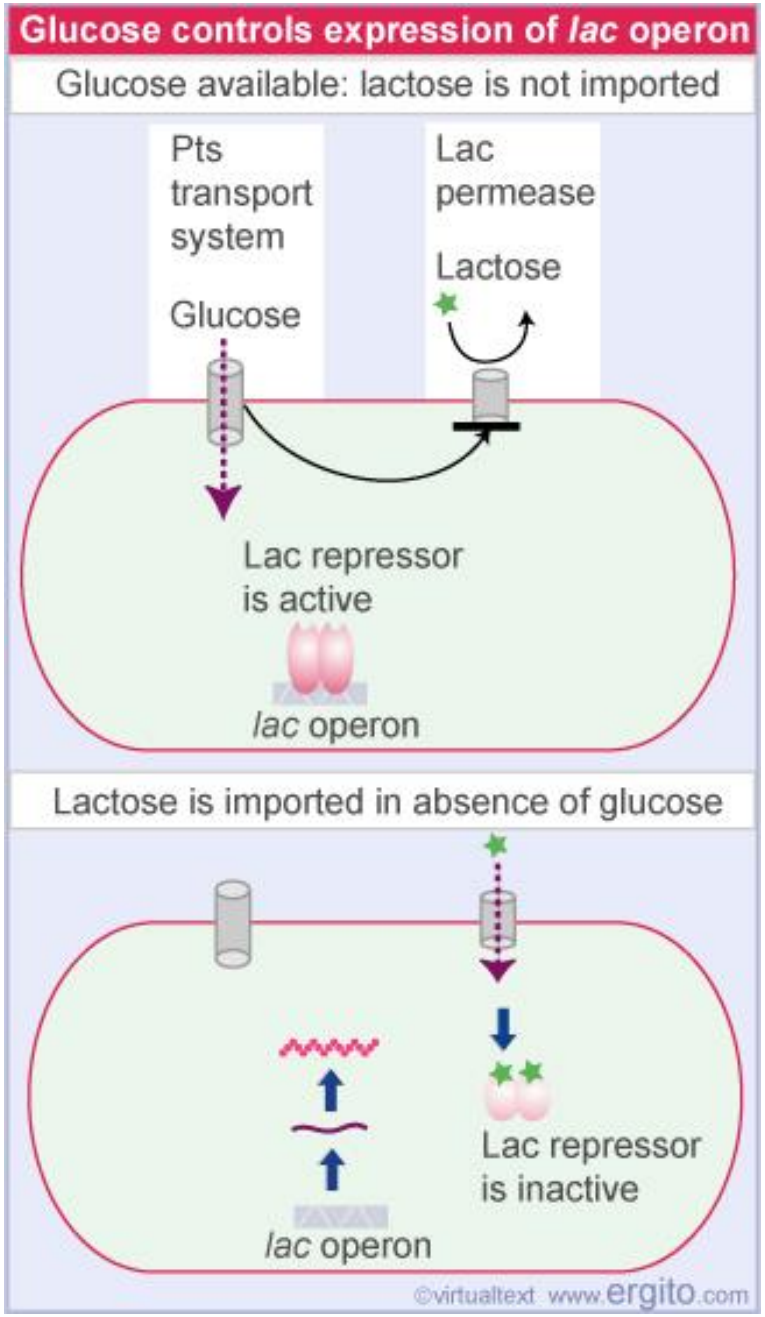
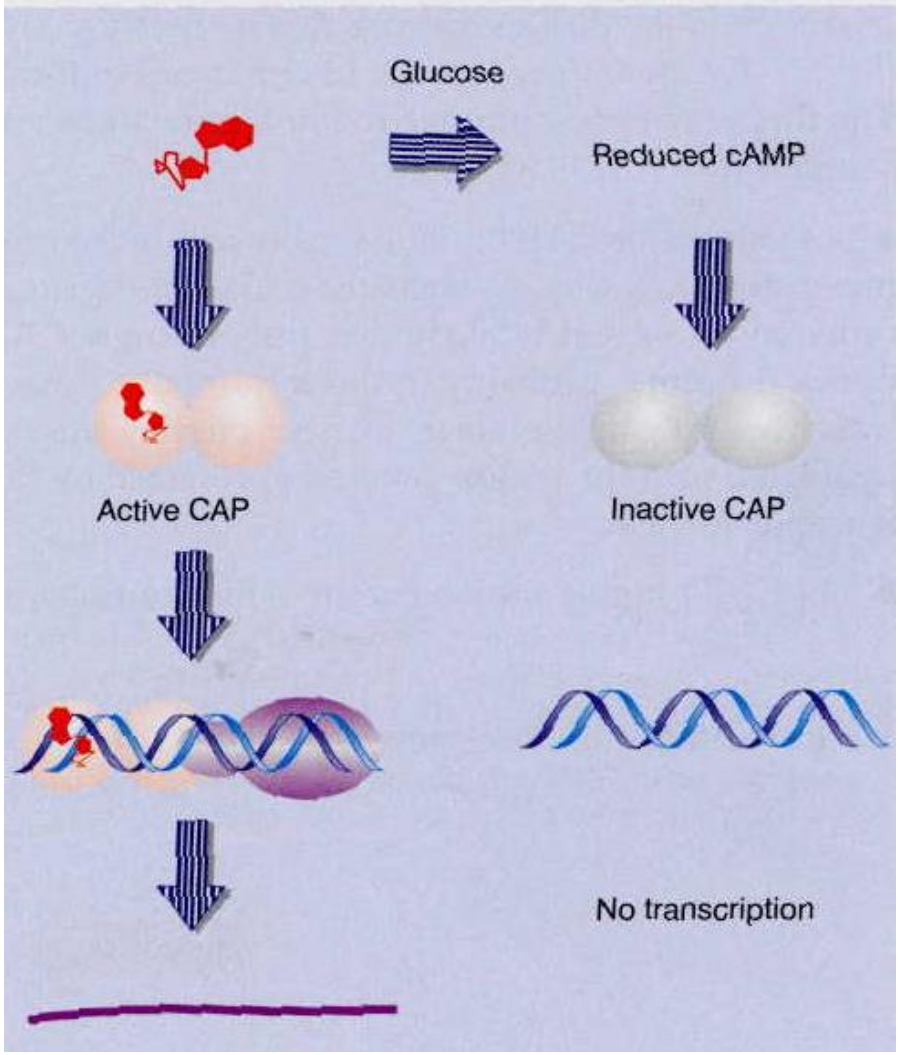


Figure 10.22 Glucose causes catabolite repression by reducing the level of cyclic AMP.



Carbon Catabolite Regulation

Cyclic AMP acts as an inducer

CAP (CRP) protein is a positive acting regulator protein

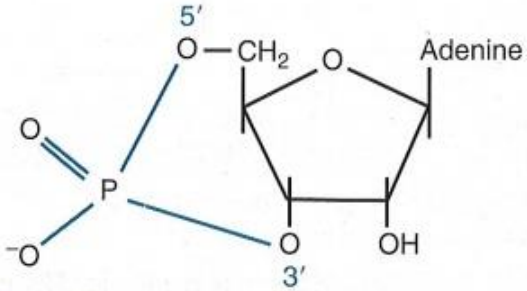


FIGURE 26.26 Cyclic AMP has a single phosphate group connected to both the 3′ and 5′ positions of the sugar ring.

Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

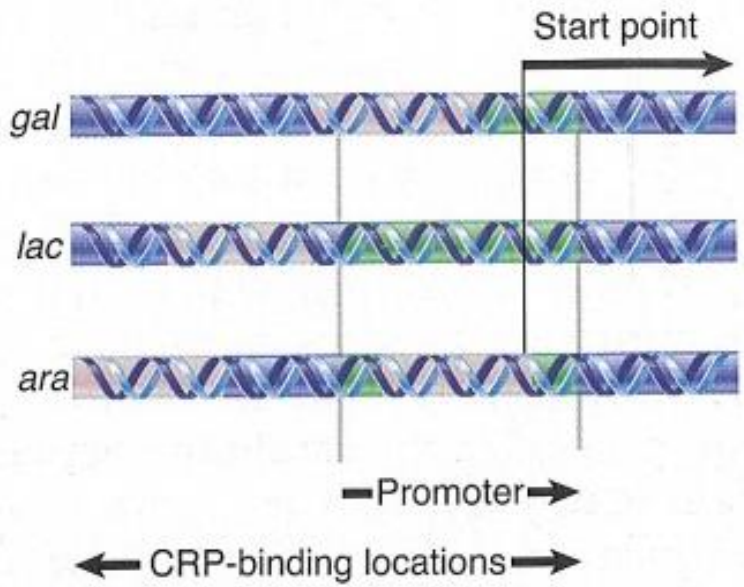


FIGURE 26.30 The CRP protein can bind at different sites relative to RNA polymerase.

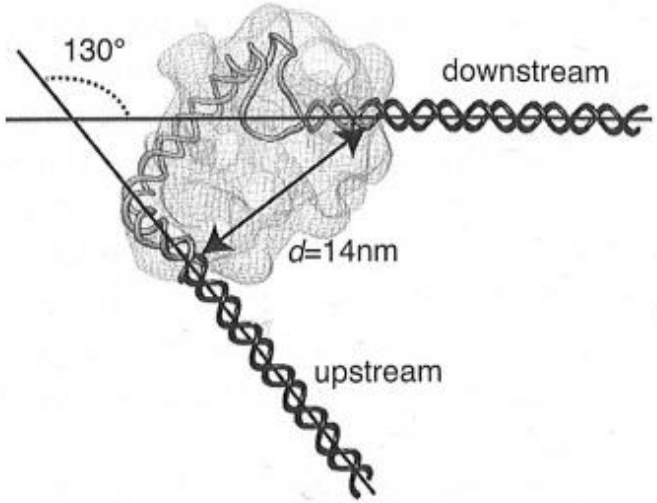


FIGURE 26.29 CRP bends DNA >90° around the center of symmetry. Class I CAP-RNAP-promoter complex electron microscopy (EM) reconstruction and fitted model: inferred path of DNA. Reproduced from H. P. Hudson, et al., Proc. Natl. Acad. Sci. USA 47 (2009): 19830-19835.

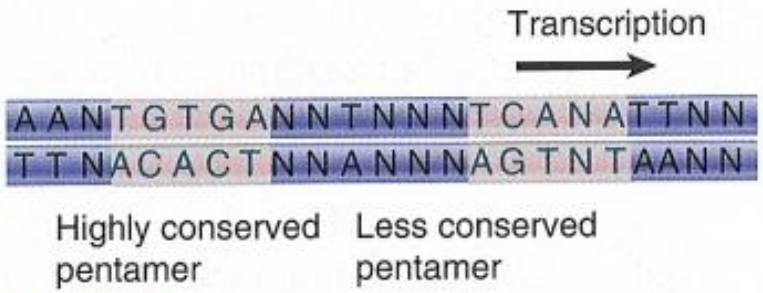


FIGURE 26.28 The consensus sequence for CRP contains the well conserved pentamer TGTGA and (sometimes) an inversion of this sequence (TCANA).

Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

Regulation at translation level

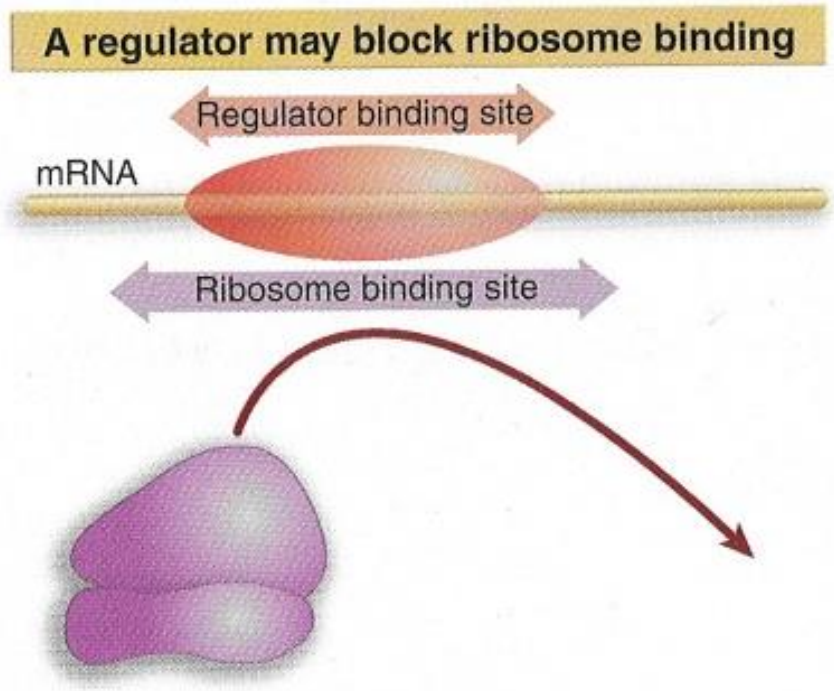


Figure 12.27 A regulator protein may block translation by binding to a site on mRNA that overlaps the ribosome-binding site at the initiation codon.

Translational repressors bind to mRNA

Repressor	Target gene	Site of action
R17 coat protein	R17 replicase	Hairpin that includes ribosome binding site
T4 RegA	Early T4 mRNAs	Various sequences including initiation codon
T4 DNA polymerase	T4 DNA polymerase	Shine-Dalgarno sequence
T4 p32	Gene 32	Single-stranded 5′ leader

Figure 12.28 Proteins that bind to sequences within the initiation regions of mRNAs may function as translational repressors.

Taken from: B. Lewin, Essential Genes, Pearson Ed. International

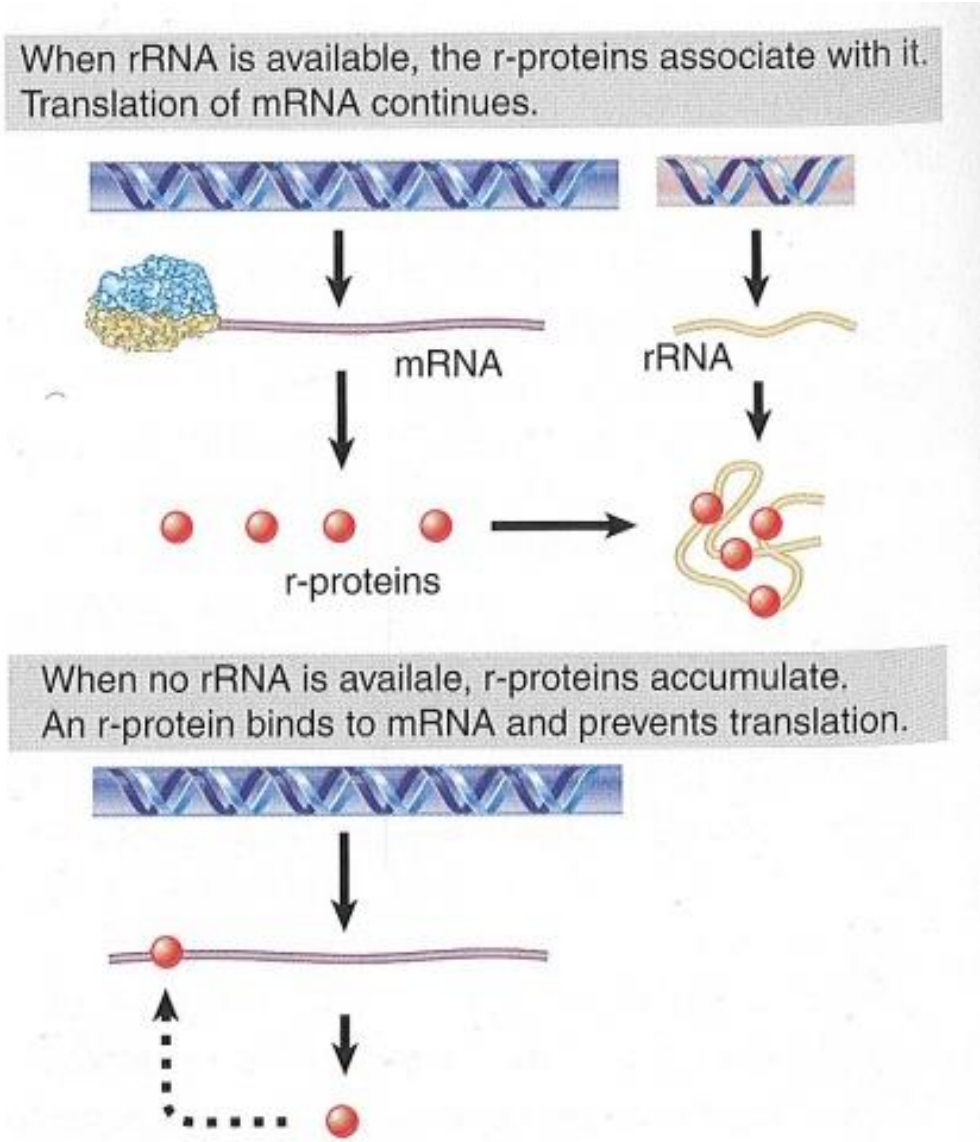


FIGURE 26.42 Translation of the r-protein operons is autogenously controlled and responds to the level of rRNA.

Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

Attenuation

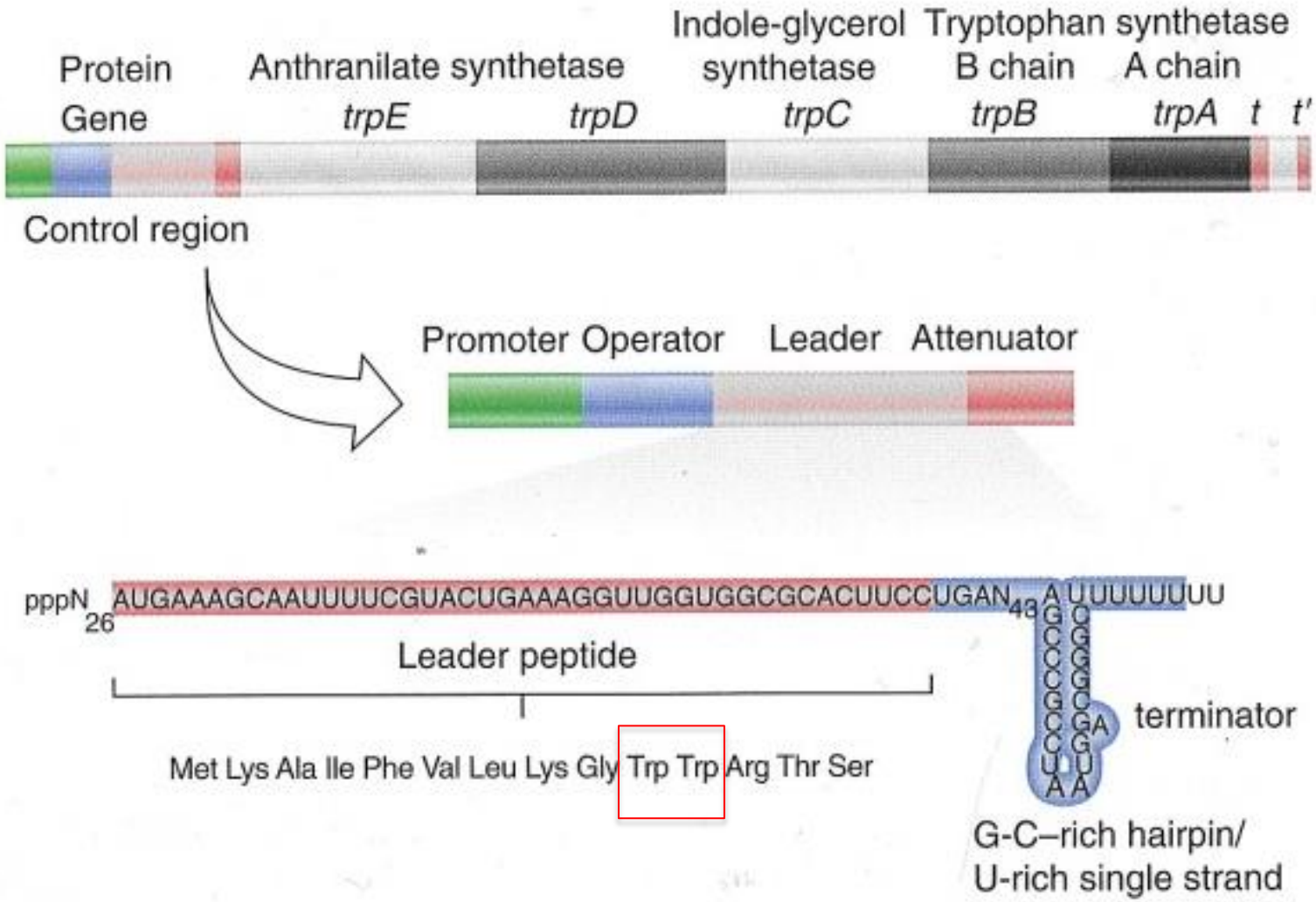
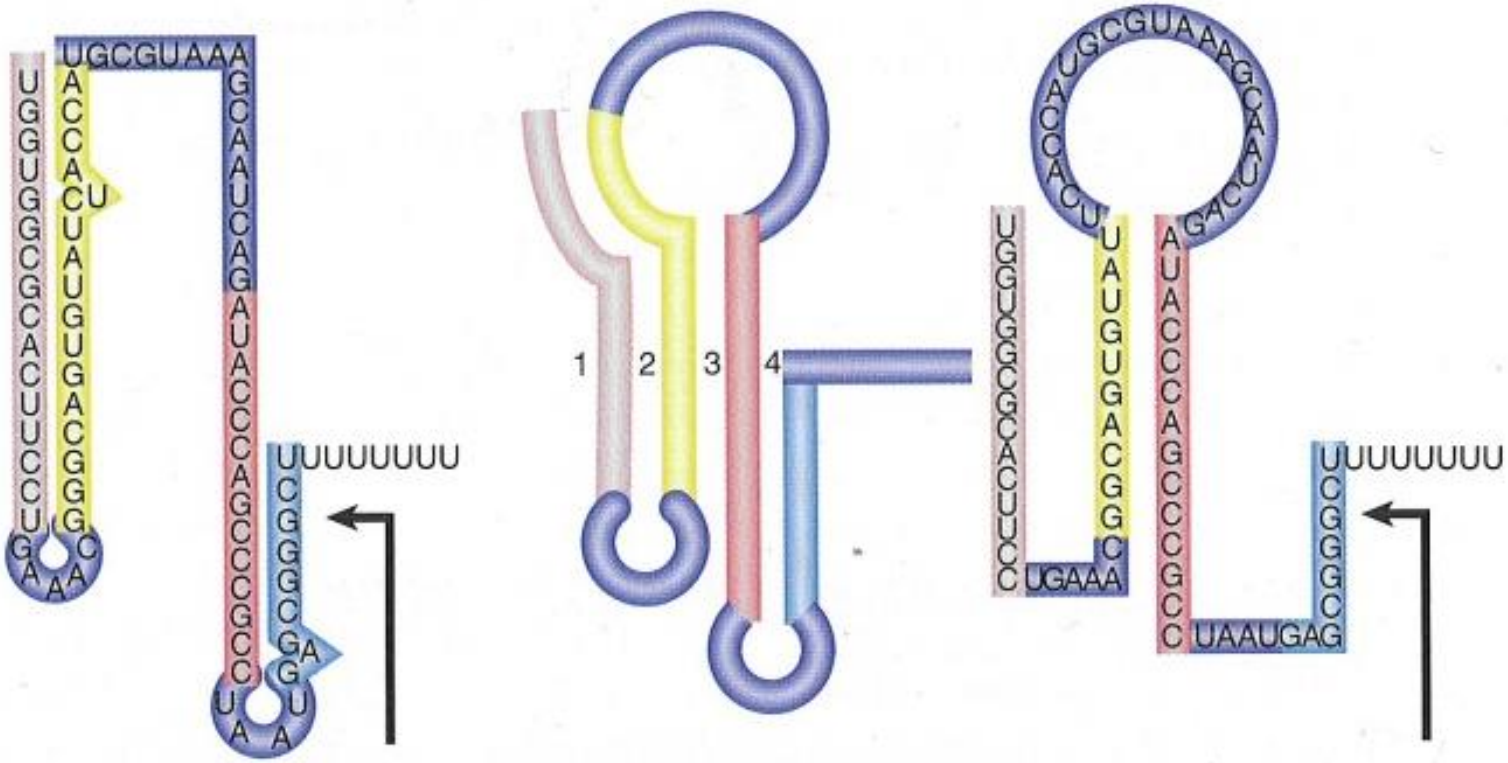


FIGURE 26.35 The *trp* operon has a short sequence coding for a leader peptide that is located between the operator and the attenuator.

Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning



Regions 3 and 4 pair to form the terminator hairpin

ALTERNATIVE STRUCTURES
Region 2 is complementary to 1 and 3
Region 3 is complementary to 2 and 4

Regions 2 and 3 pair; terminator region is single stranded

FIGURE 26.36 The *trp* leader region can exist in alternative base-paired conformations. The center shows the four regions that can base pair. Region 1 is complementary to region 2, which is complementary to region 3, which is complementary to region 4. On the left is the conformation produced when region 1 pairs with region 2 and region 3 pairs with region 4. On the right is the conformation when region 2 pairs with region 3, leaving regions 1 and 4 unpaired.

Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

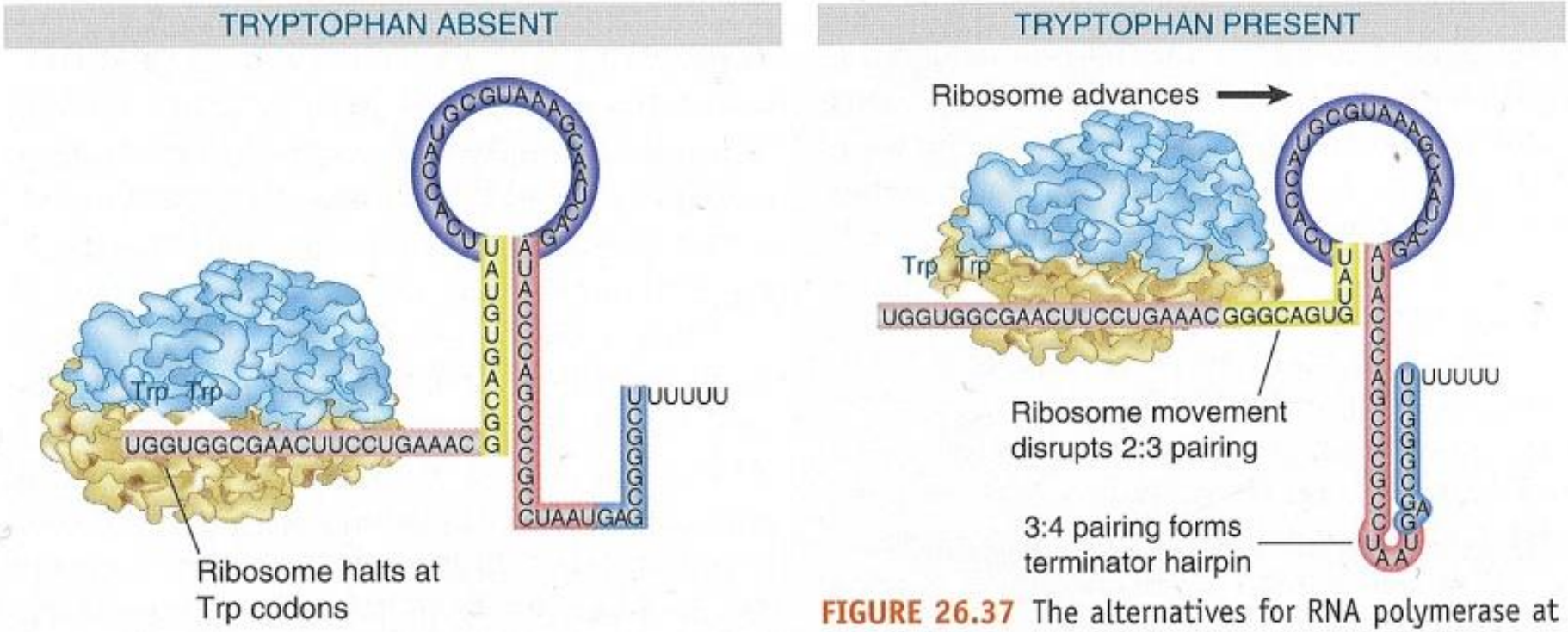
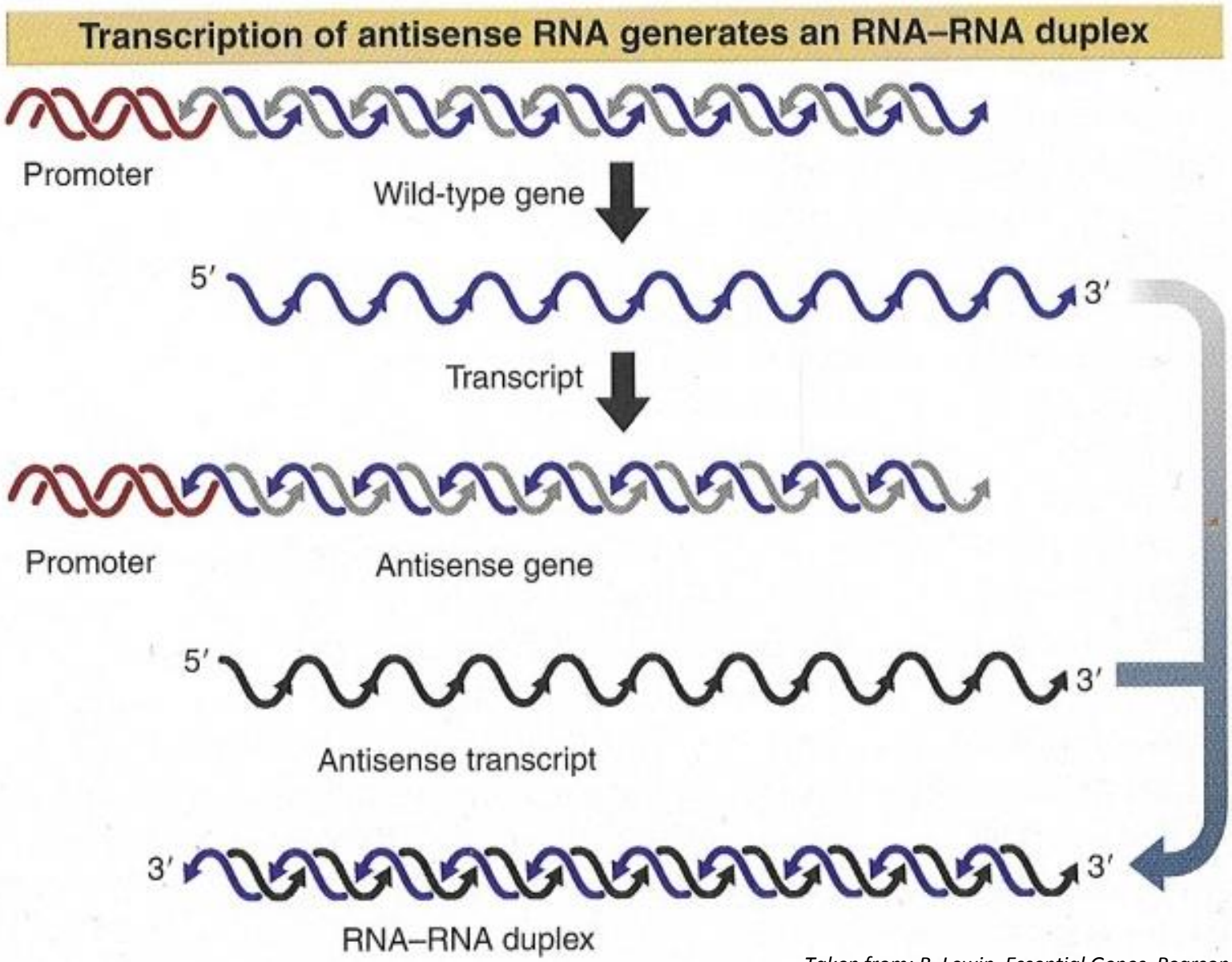


FIGURE 26.37 The alternatives for RNA polymerase at the attenuator depend on the location of the ribosome, which determines whether regions 3 and 4 can pair to form the terminator hairpin.

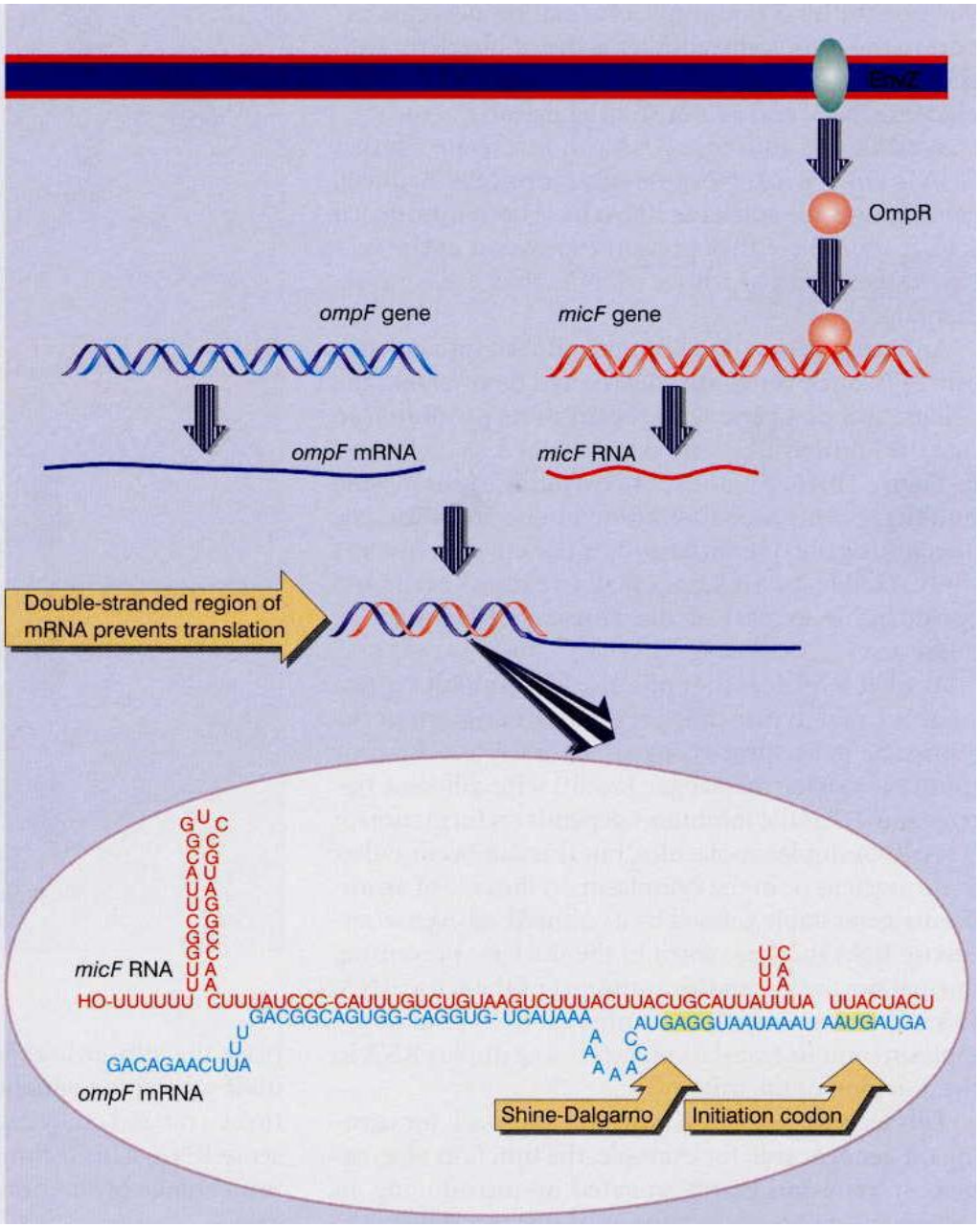
Taken from: J.E. Krebs, E.S. Goldstein, S.T. Kilpatrick; „Lewin´s Genes XI"; Jones&Bartlett Learning

Antisense RNA



Taken from: B. Lewin, Essential Genes, Pearson Ed. International

Figure 10.44 Increase in osmolarity activates EnvZ, which activates OmpR, which induces transcription of *micF* and *ompC* (not shown). *micF* RNA is complementary to the 5' region of *ompF* mRNA and prevents its translation.



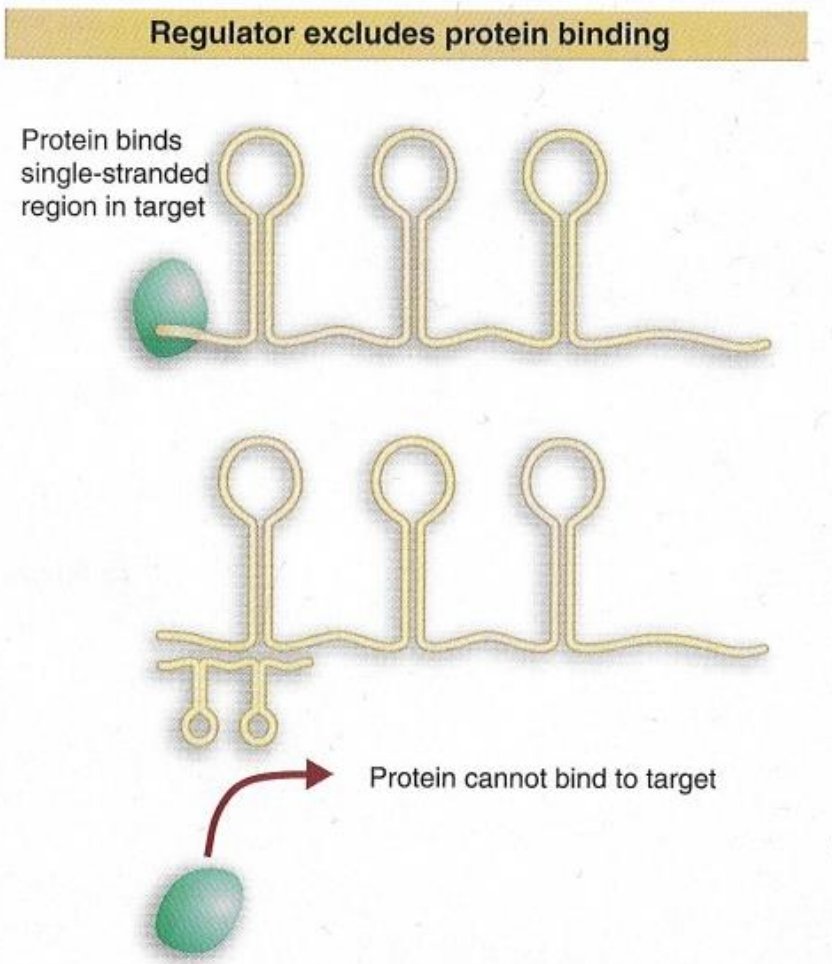


Figure 13.13 A protein that binds to a single-stranded region in a target RNA could be excluded by a regulator RNA that forms a duplex in this region.

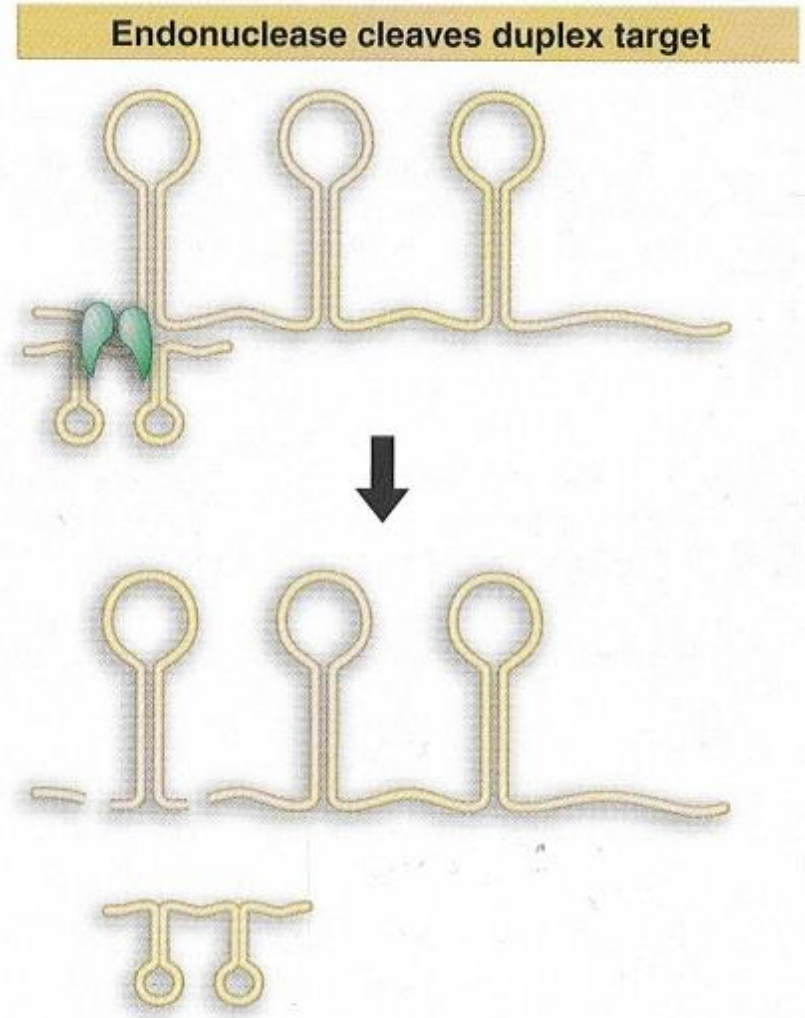


Figure 13.14 By binding to a target RNA to form a duplex region, a regulator RNA may create a site that is attacked by a nuclease.

Taken from: B. Lewin, Essential Genes, Pearson Ed. International

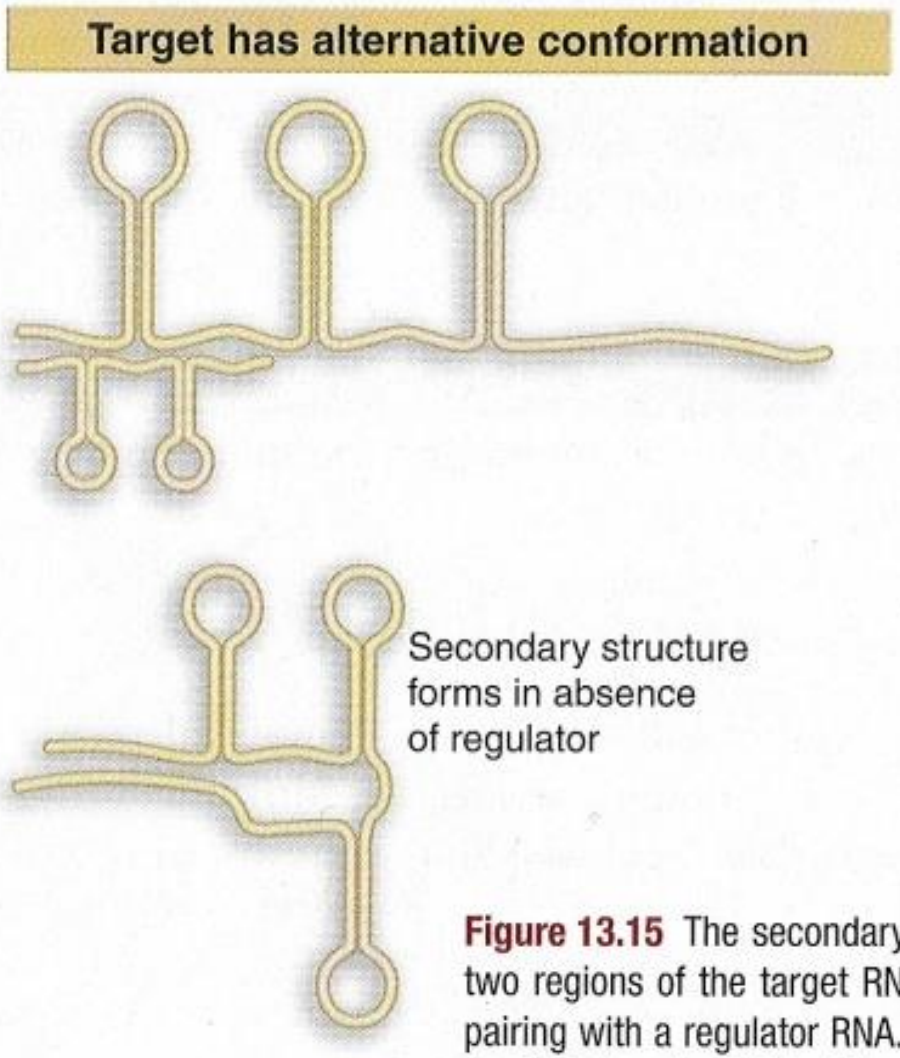


Figure 13.15 The secondary structure formed by base pairing between two regions of the target RNA may be prevented from forming by base pairing with a regulator RNA. In this example, the ability of the 3′ end of the RNA to pair with the 5′ end is prevented by the regulator.

Taken from: B. Lewin, Essential Genes, Pearson Ed. International

A 3′ terminal loop in *oxyS* RNA pairs with the initiation site of *flhA* mRNA

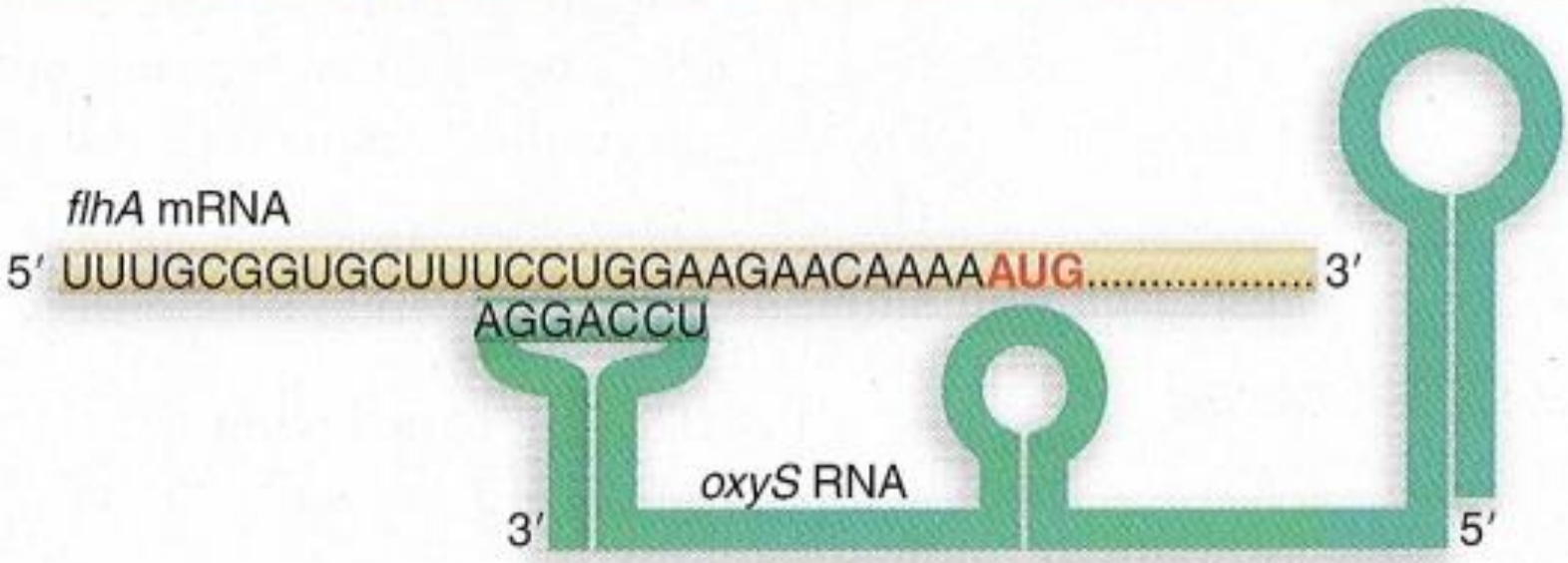


Figure 13.18 *oxyS* RNA inhibits translation of *flhA* mRNA by base pairing with a sequence just upstream of the AUG initiation codon.

Taken from: B. Lewin, Essential Genes, Pearson Ed. International

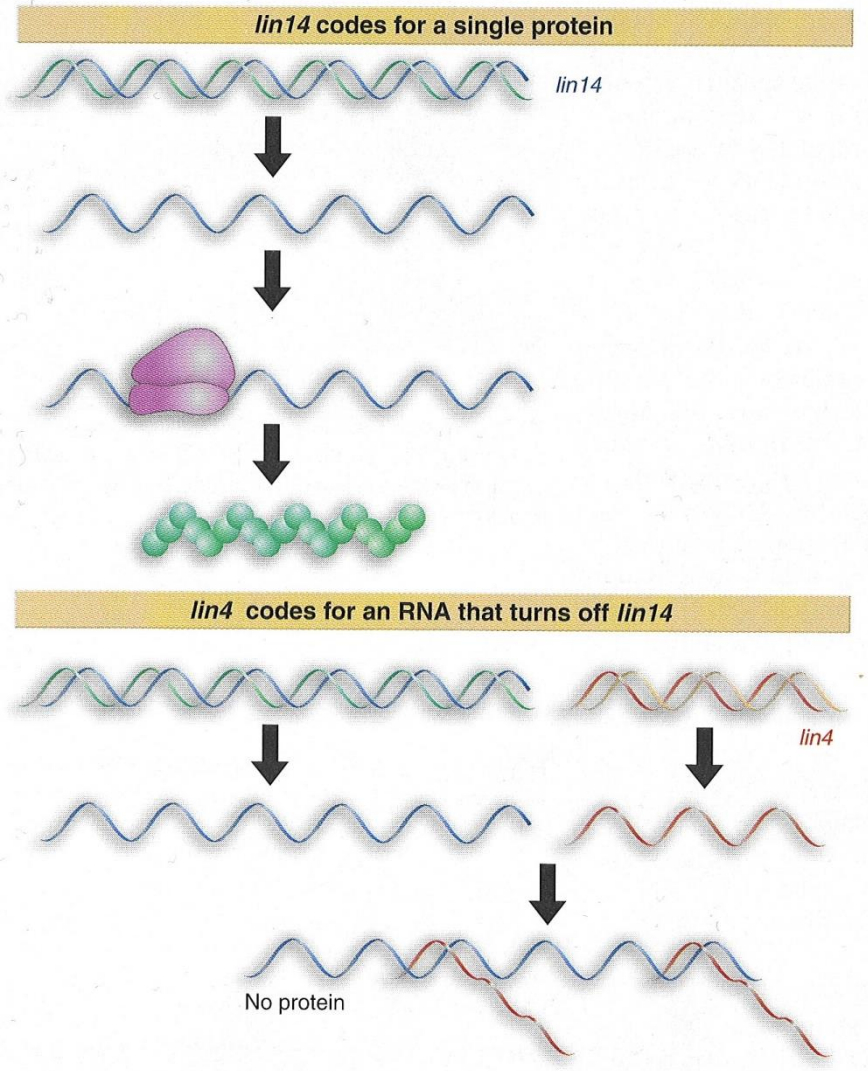
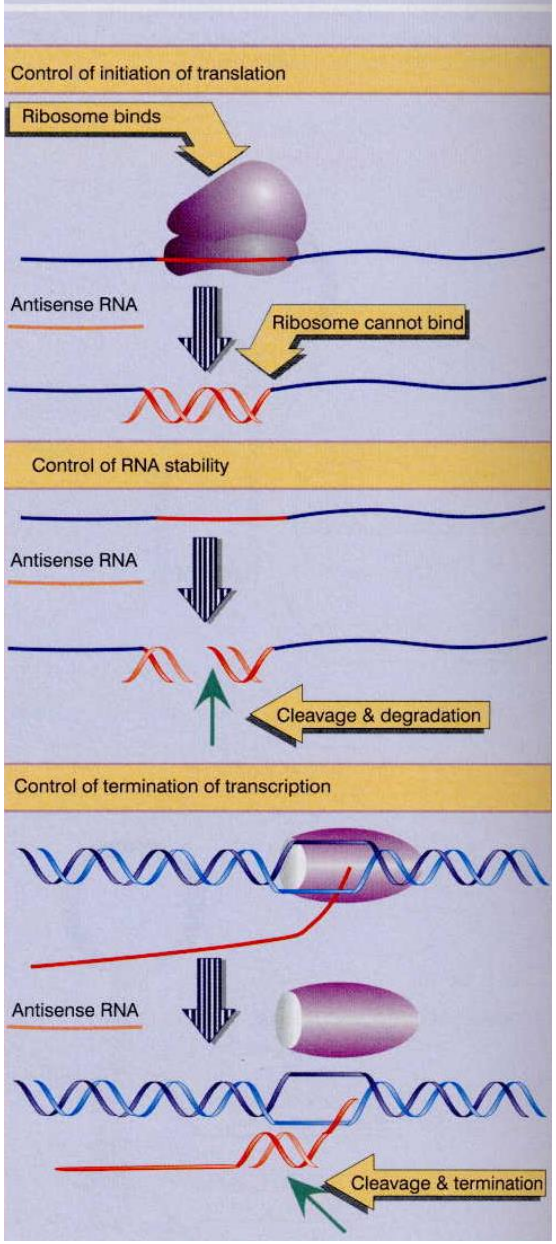


Figure 13.19 *lin4* RNA regulates expression of *lin14* by binding to the 3′ nontranslated region.

Figure 10.43 Antisense RNA can affect function or stability of an RNA target.



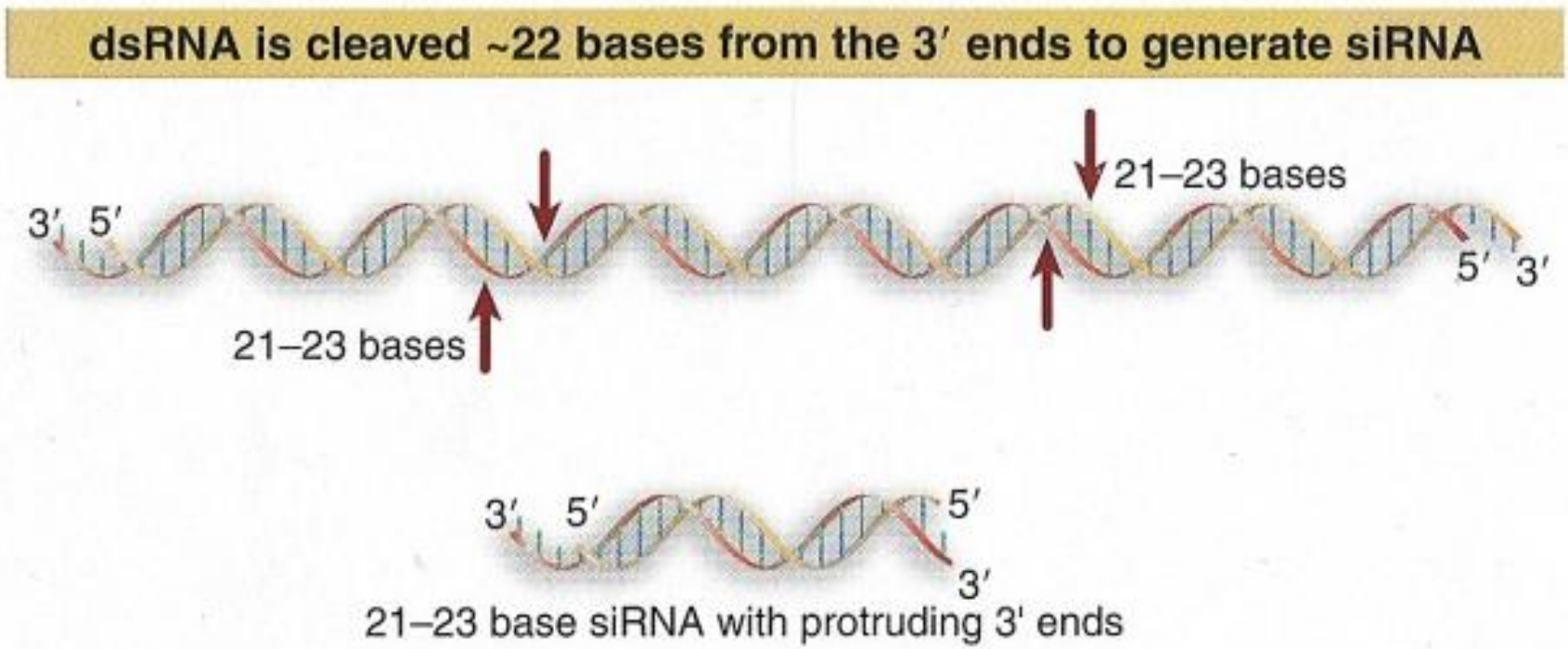


Figure 13.20 siRNA that mediates RNA interference is generated by cleaving dsRNA into smaller fragments. The cleavage reaction occurs 21–23 nucleotides from a 3′ end. The siRNA product has protruding bases on its 3′ ends.

Taken from: B. Lewin, Essential Genes, Pearson Ed. International

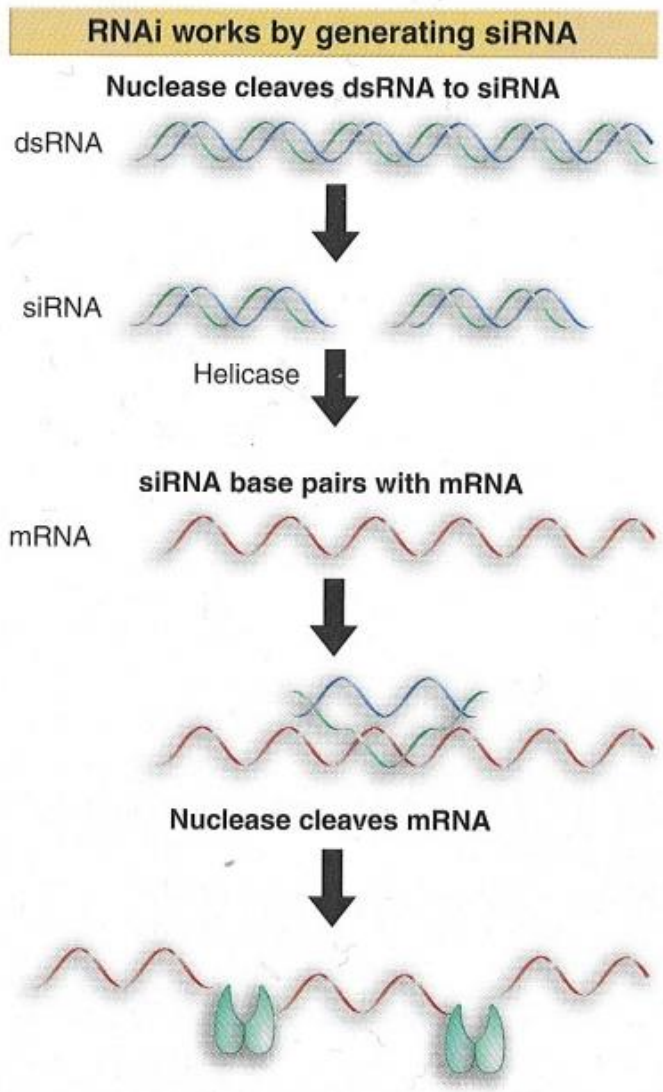


Figure 13.21 RNAi is generated when a dsRNA is cleaved into fragments that direct cleavage of the corresponding mRNA.

Taken from: B. Lewin, Essential Genes, Pearson Ed. International

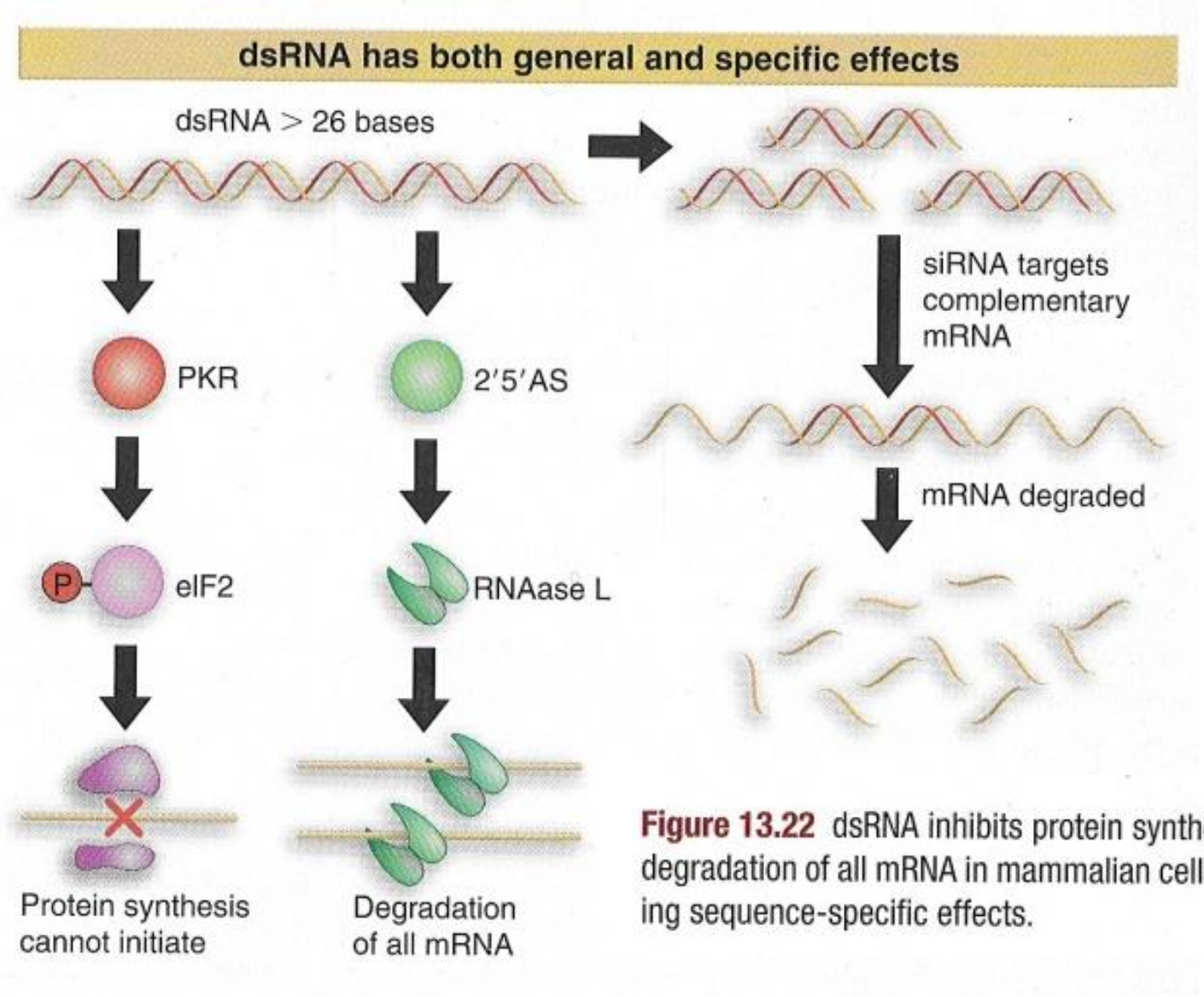


Figure 13.22 dsRNA inhibits protein synthesis and triggers degradation of all mRNA in mammalian cells as well as having sequence-specific effects.

Taken from: B. Lewin, Essential Genes, Pearson Ed. International