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

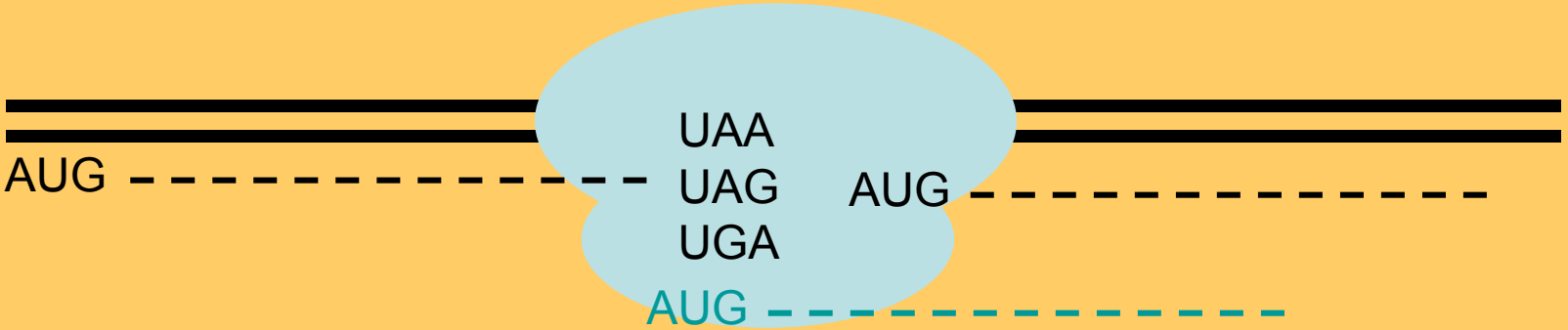
Surrounding of SD and AUG!!!

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**

AAA	AAA	AAA	UCA
Lys	Lys	Lys	Ser

AA	AAA	AAA	AUCA
Lys	Lys	Lys	Ile

Figure 7.1 All the triplet codons have meaning: 61 represent amino acids, and 3 cause termination (STOP).

First base → (rows: U, C, A, G)

Second base → (columns: U, C, A, G)

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

Figure 7.2 The number of codons for each amino acid does not correlate closely with its frequency of use in proteins.

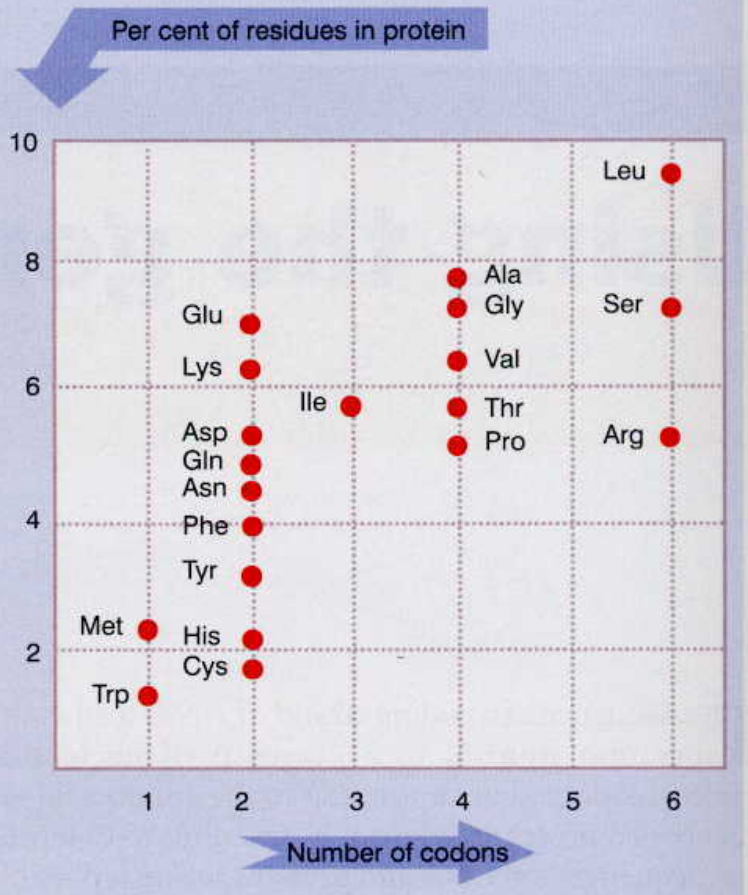


Figure 7.3 Third bases have the least influence on codon meanings. Boxes indicate groups of codons within which third-base degeneracy ensures that the meaning is the same.

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 CAA CAG	CGU CGC CGA CGG
AUU AUC AUA AUG	ACU ACC ACA ACG	AAU AAC AAA AAG	AGU AGC AGA AGG
GUU GUC GUA GUG	GCU GCC GCA GCG	GAU GAC GAA GAG	GGU GGC GGA GGG

Third base relationship	Third bases with same meaning	Number of codons
third base irrelevant	U, C, A, G	32
purines differ from pyrimidines	U or C	14
	A or G	10
unique definitions	U, C, A	3
	G only	2

Figure 7.4 Codon-anticodon pairing involves wobbling at the third position.

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

Universial 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.

Figure 6.1 Ribosomes are large ribonucleoprotein particles that contain more RNA than protein and dissociate into large and small subunits.

Ribosomes	rRNAs	Proteins
Bacterial		
50S	23S = 2904 bases	31
70S mass: 2.5×10^6 D 66% RNA	5S = 120 bases	
30S	16S = 1542 bases	21
Mammalian		
60S	28S = 4718 bases 5.8S = 160 bases 5S = 120 bases	49
80S mass: 4.2×10^6 D 60% RNA		
40S	18S = 1874 bases	33

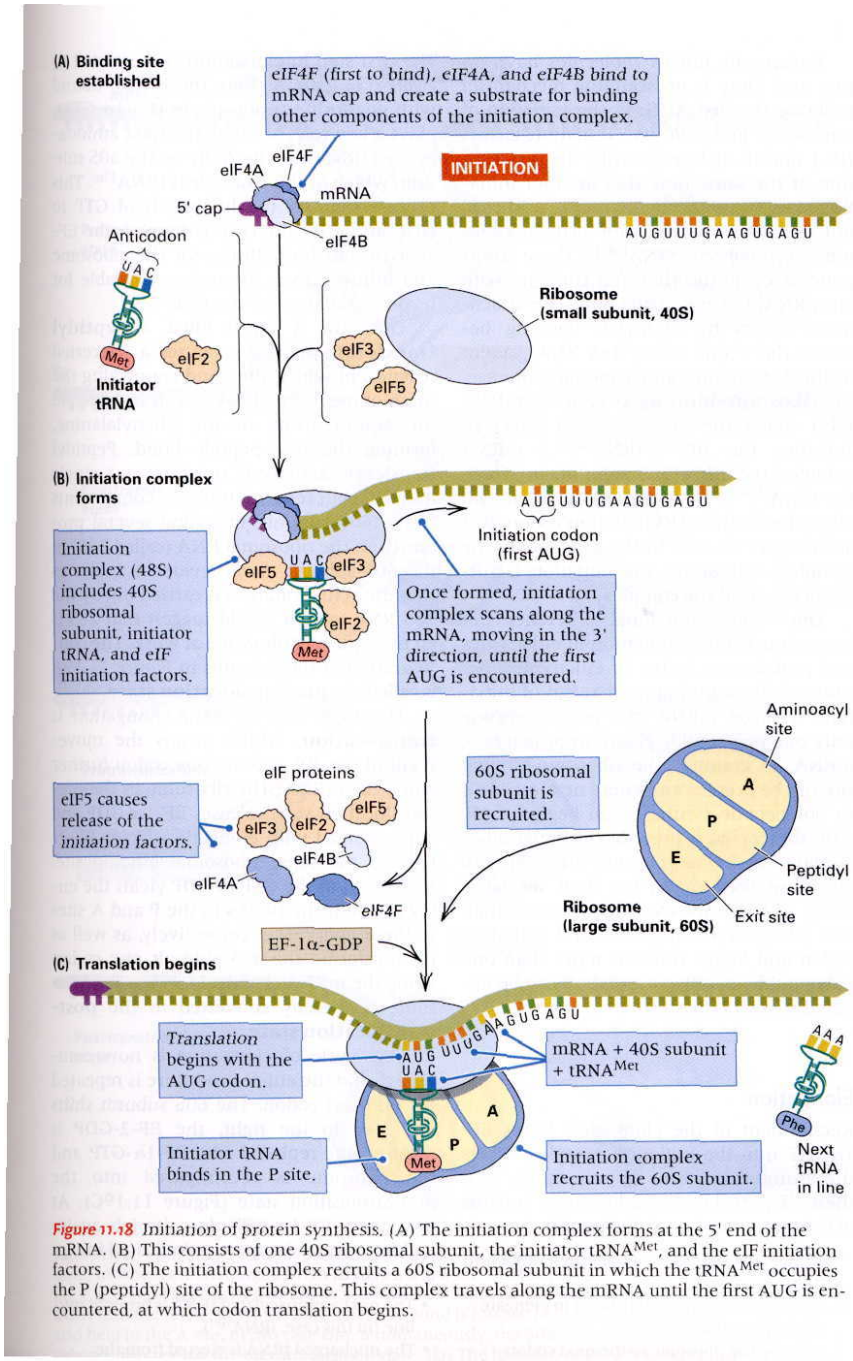


Figure 11.18 Initiation of protein synthesis. (A) The initiation complex forms at the 5' end of the mRNA. (B) This consists of one 40S ribosomal subunit, the initiator $tRNA^{Met}$, and the eIF initiation factors. (C) The initiation complex recruits a 60S ribosomal subunit in which the $tRNA^{Met}$ occupies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

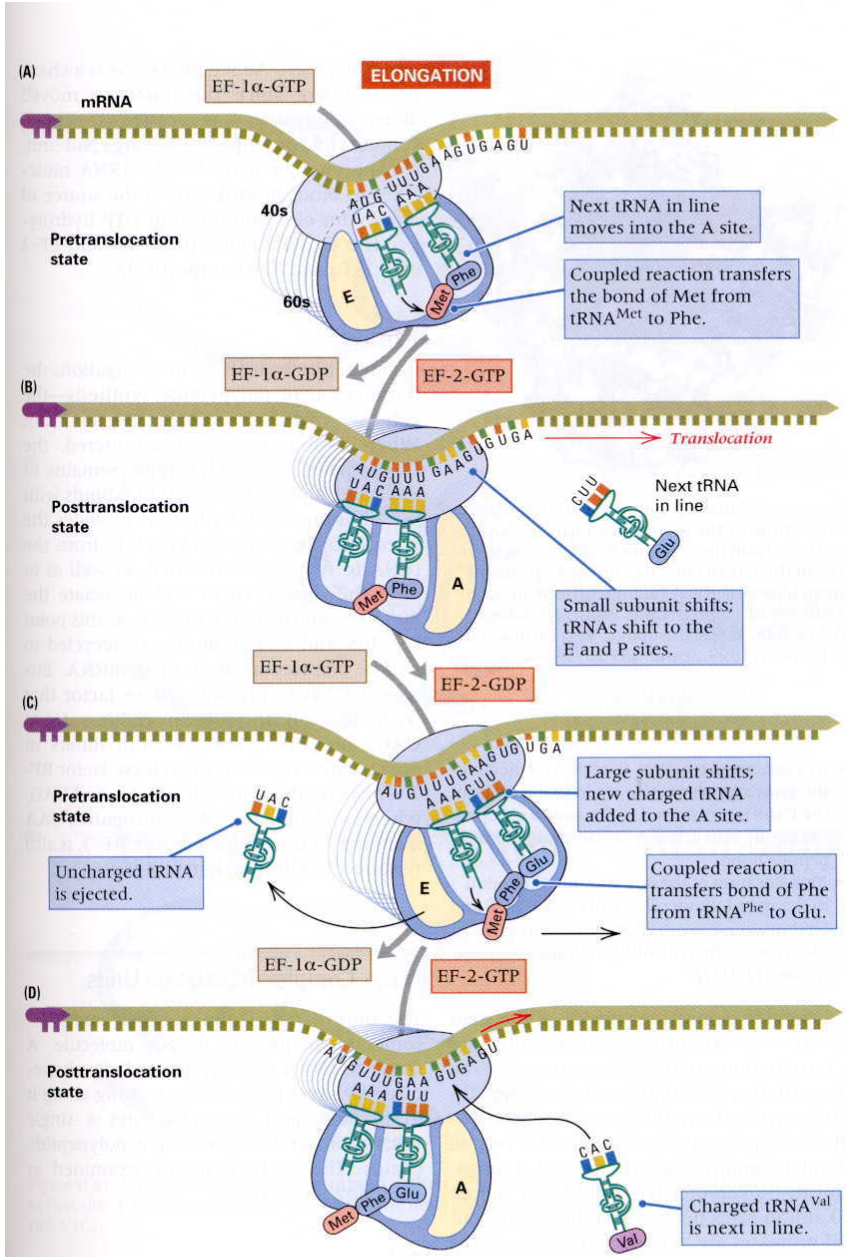


Figure 11.19 Elongation cycle in protein synthesis. (A) Pretranslocation state. (B) Posttranslocation state, in which an uncharged tRNA occupies the E site and the polypeptide is attached to the tRNA in the P site. (C) The function of EF-1α is to release the uncharged tRNA and bring the next charged 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.

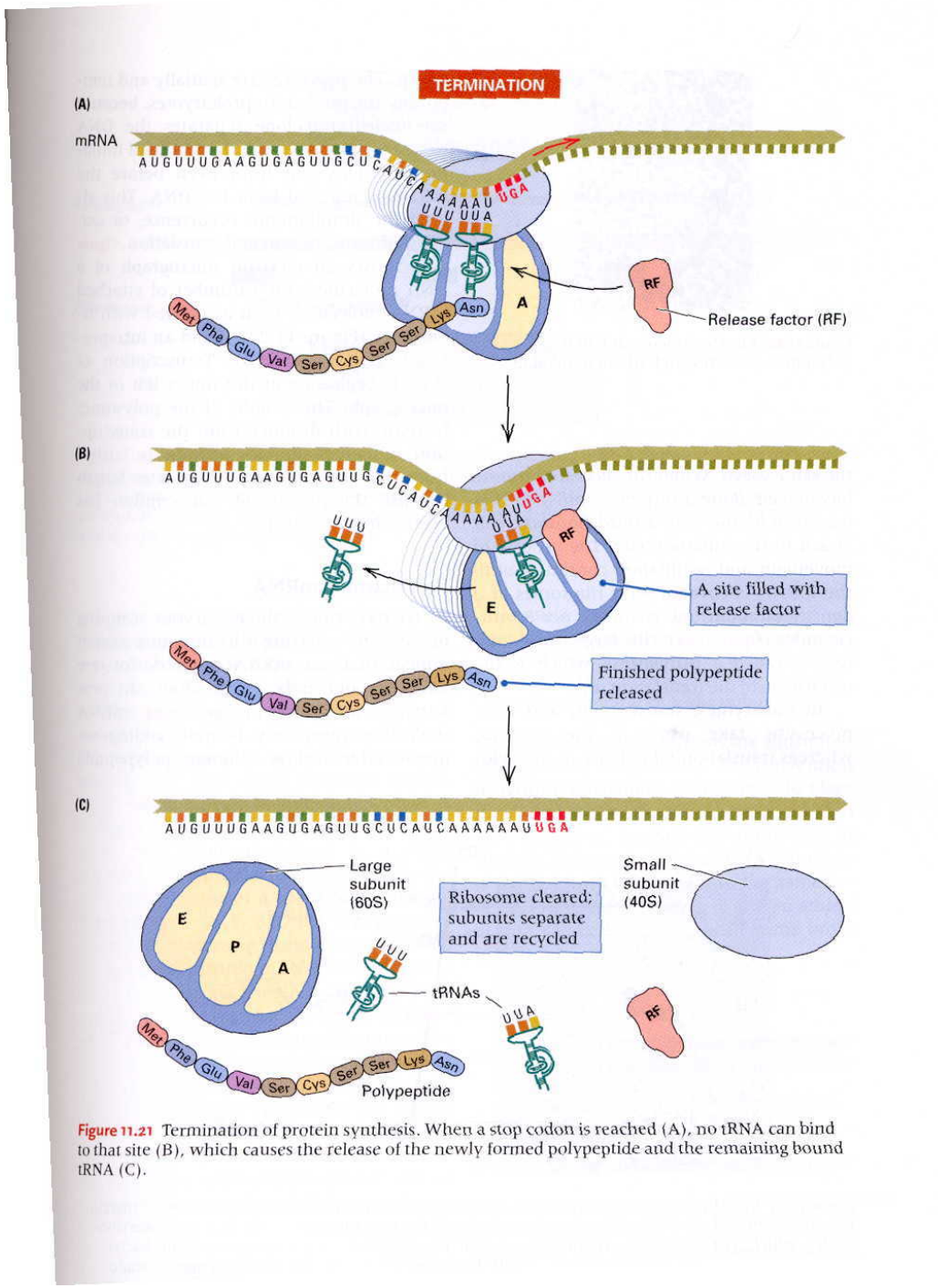


Figure 11.21 Termination of protein synthesis. When a stop codon is reached (A), no tRNA can bind to that site (B), which causes the release of the newly formed polypeptide and the remaining bound tRNA (C).

18.11.14

Regulation of Gene Expression

Prokaryotes

Escherichia coli

Lactose Metabolism

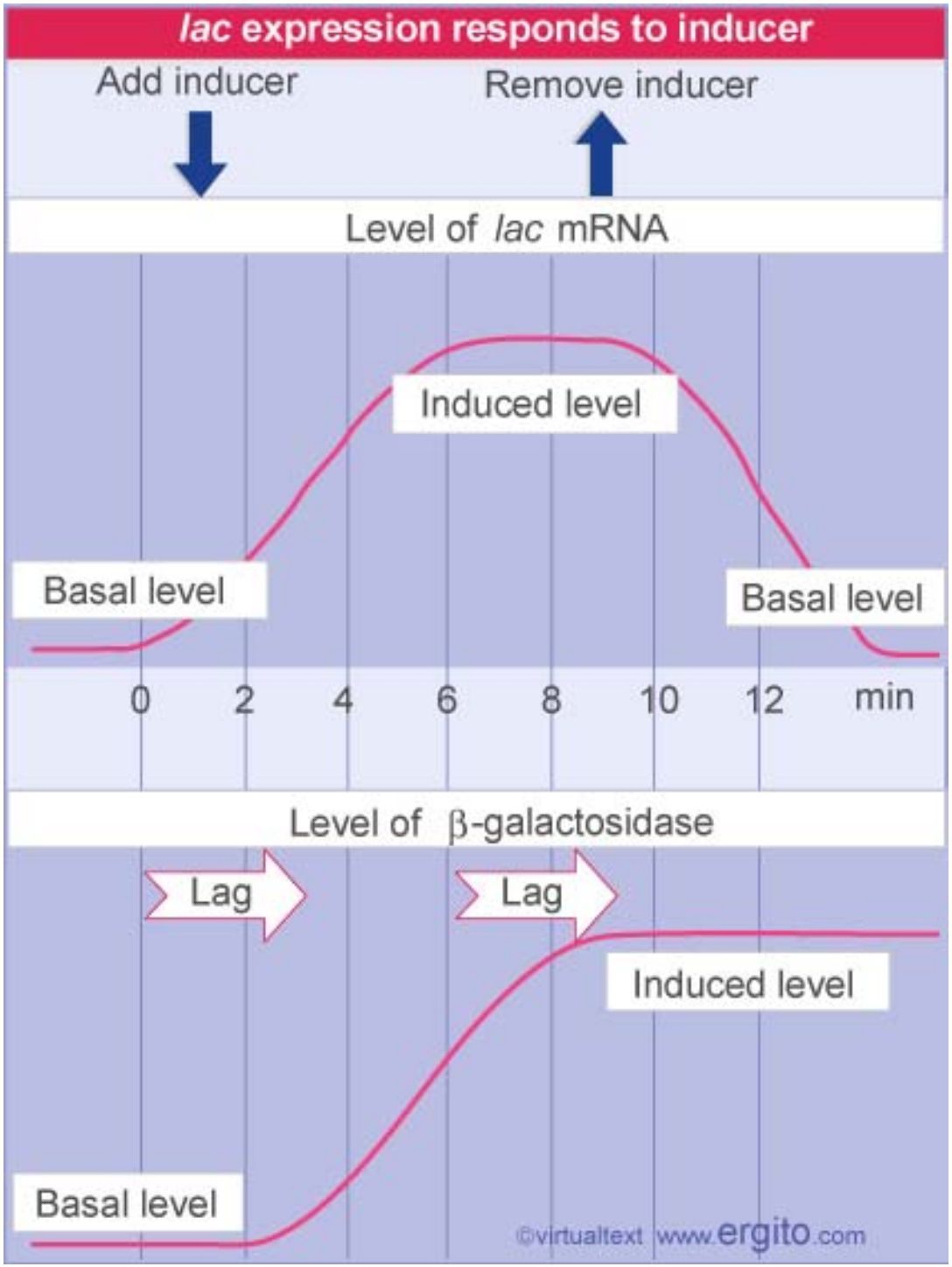
Absence of lactose → Only few molecules of ß-galactosidase per cell

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

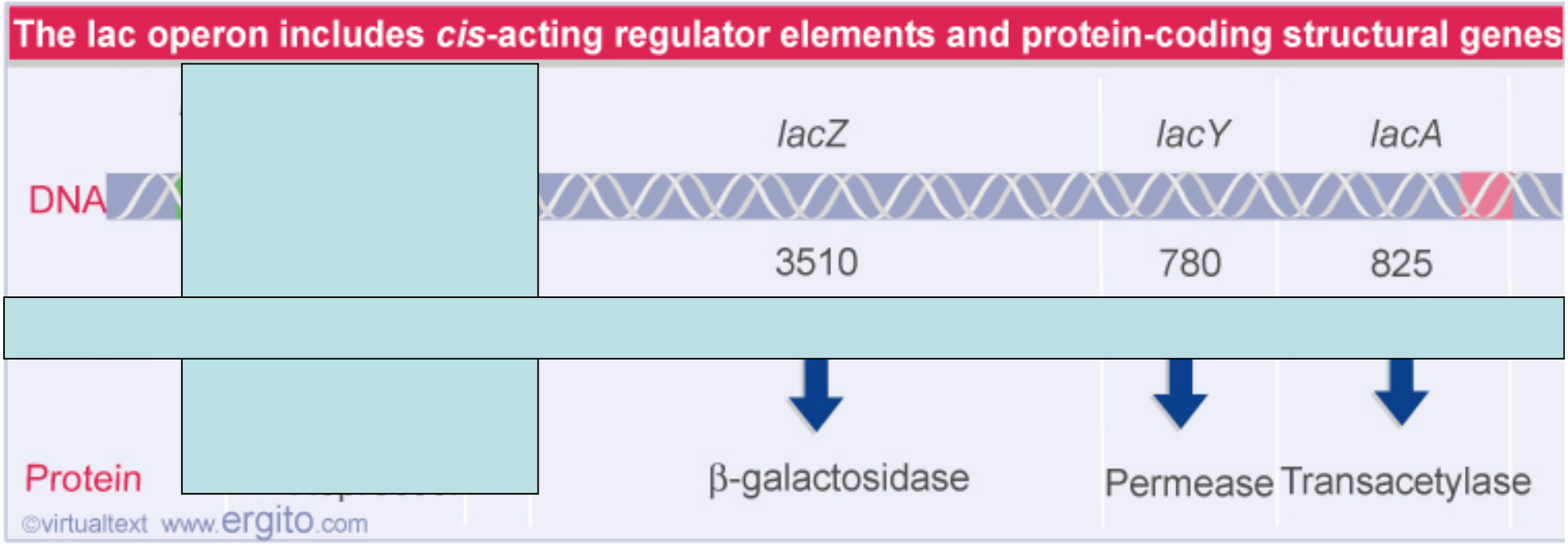
Not enzyme is inhibited, enzyme synthesis is affected

Detailed biochemical and genetic analysis

Jacob, Monod, Pardee → Nobel prize


lac expression responds to inducer
Add inducer
Remove inducer
Level of lac mRNA
Induced level
Basal level
Basal level
0
2
4
6
8
10
12
min
Level of β-galactosidase
Lag
Lag
Induced level
Basal level
©virtualtext www.ergito.com

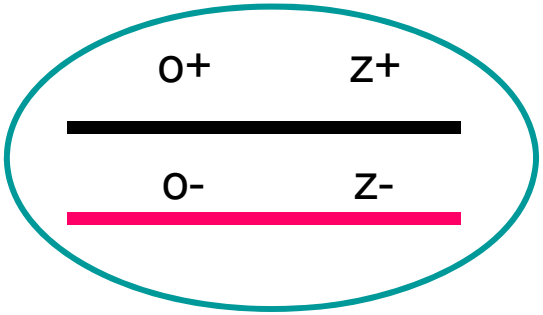
lac-Operon



Ort I

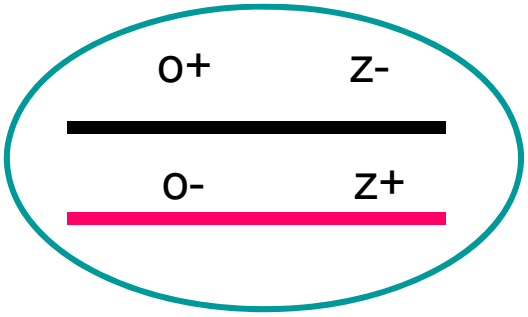
Ort O

Heterogenote analysis



Cis-configuration

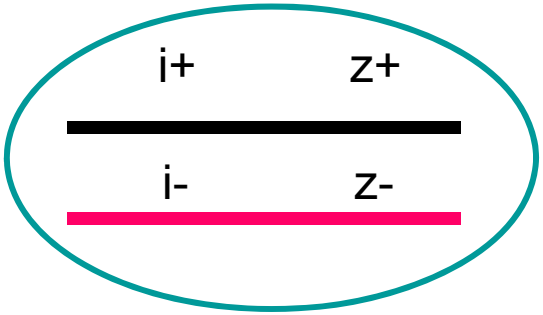
inducible



Trans-configuration

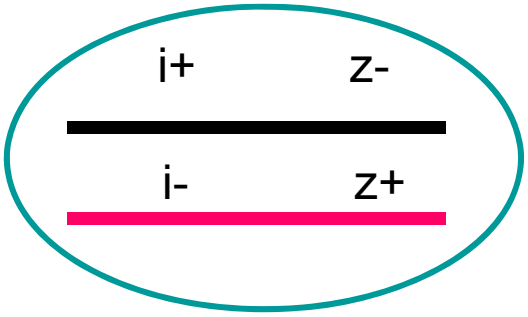
constitutive

Heterogenote analysis



Cis-configuration

inducible



Trans-configuration

inducible

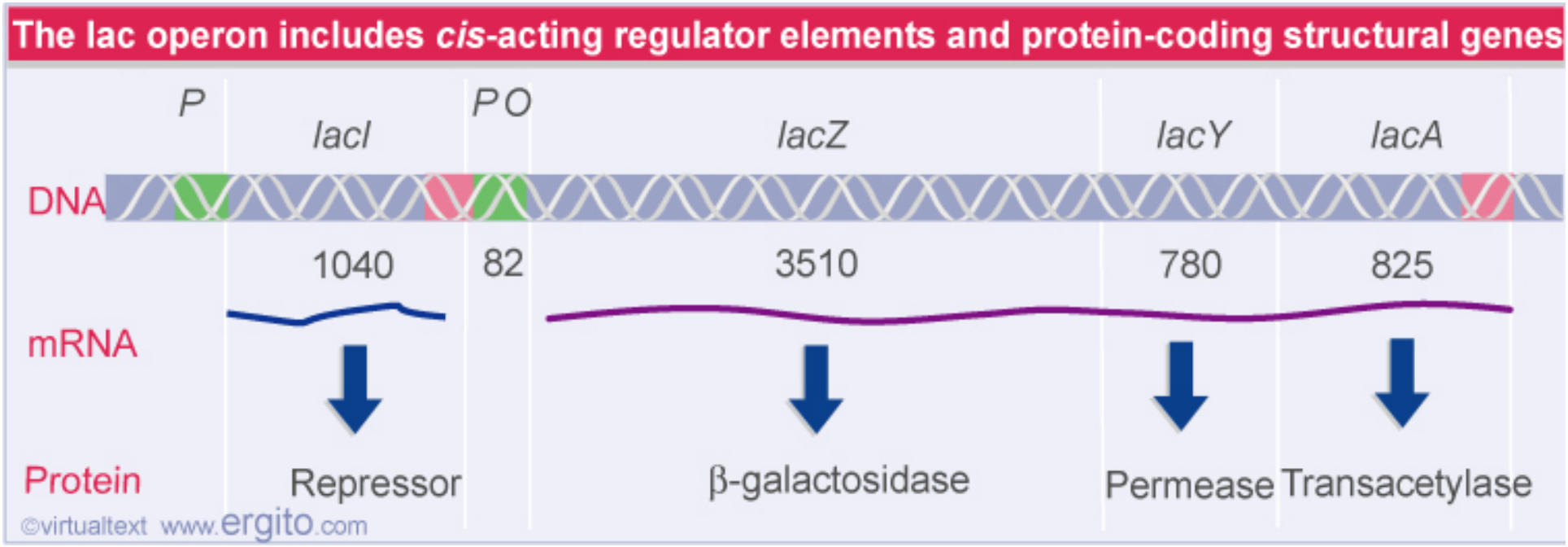
Model for behaviour of heterogenotes

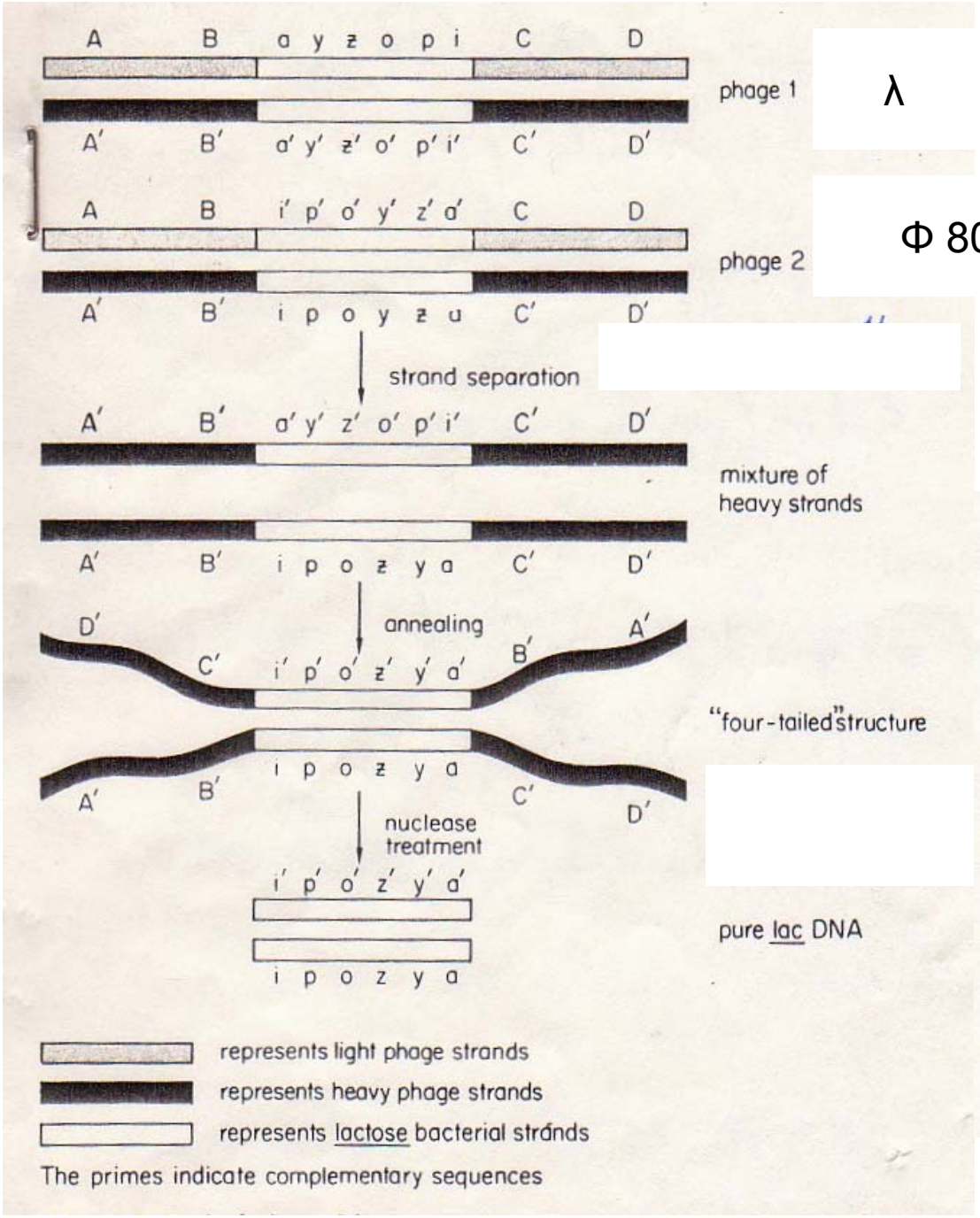
lacO → **located adjacent to *lacZ,* mutation in l*acO* results in loss of regulatory function when connected to *lacZ,* no complementation by wt-allele in trans**

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***


The lac operon includes *cis*-acting regulator elements and protein-coding structural genes
P
lacI
PO
lacZ
lacY
lacA
DNA
1040
82
3510
780
825
mRNA
Protein
Repressor
β-galactosidase
Permease
Transacetylase
©virtualtext www.ergito.com



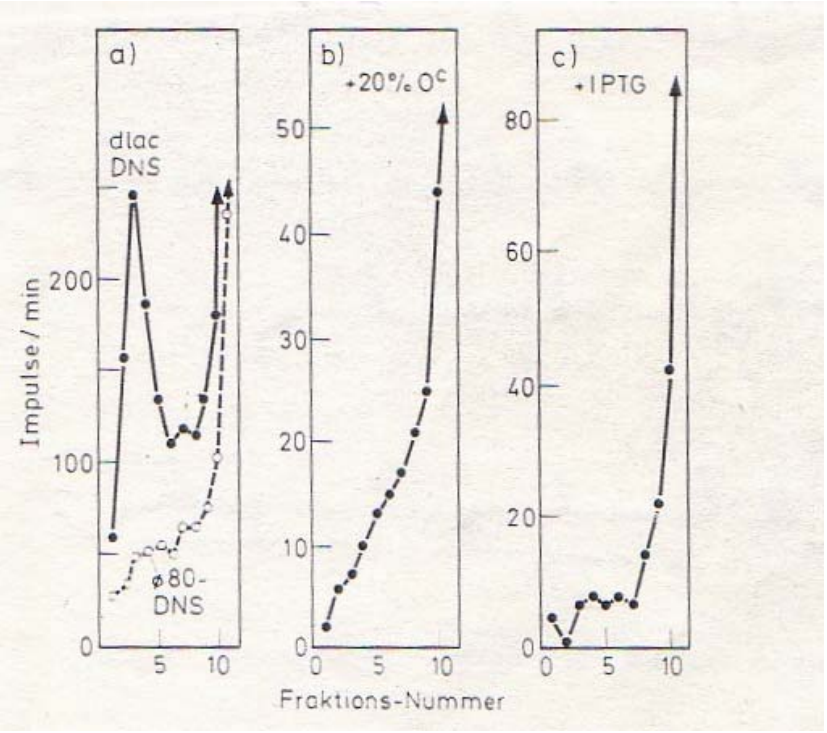
λ

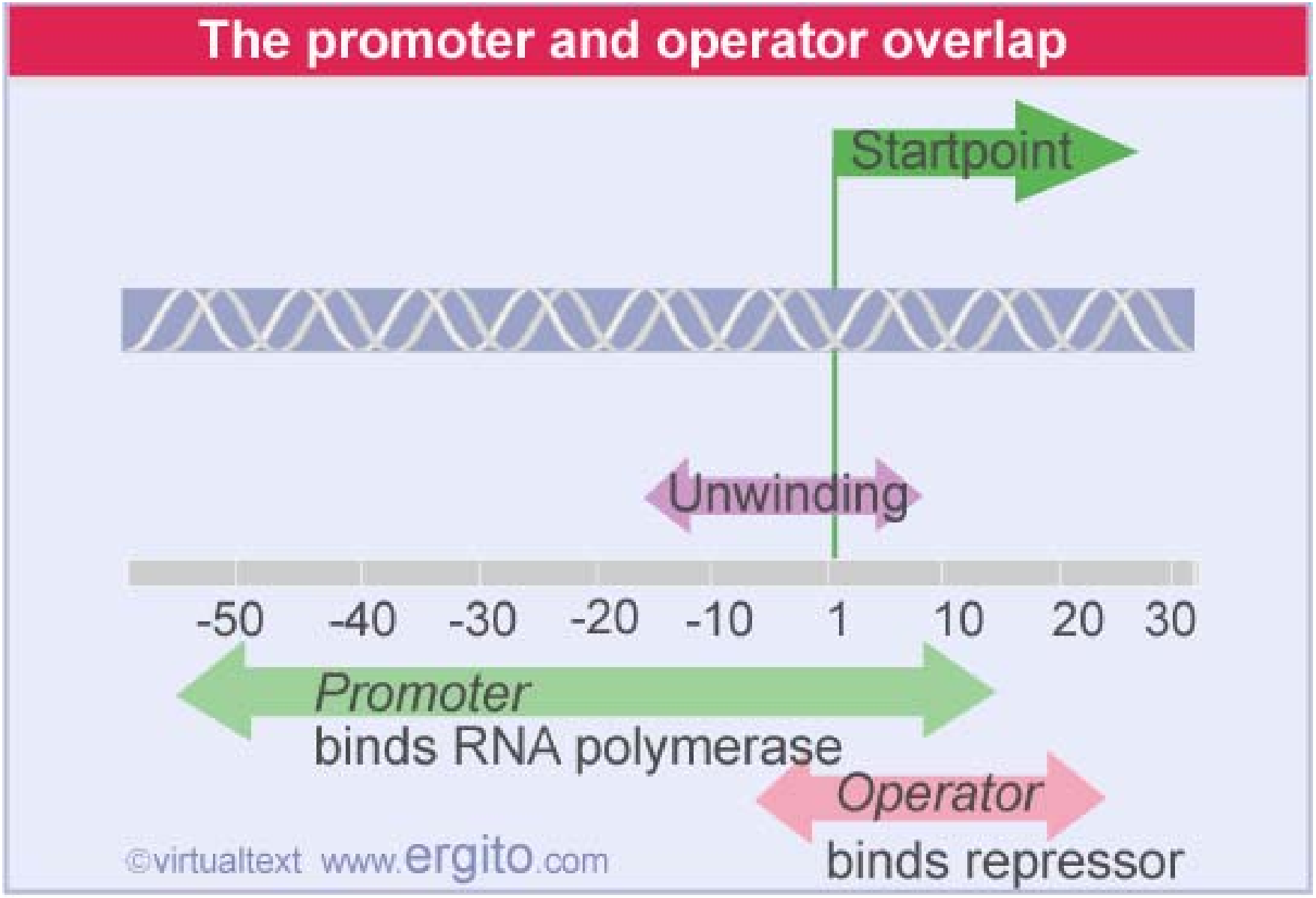
Φ 80

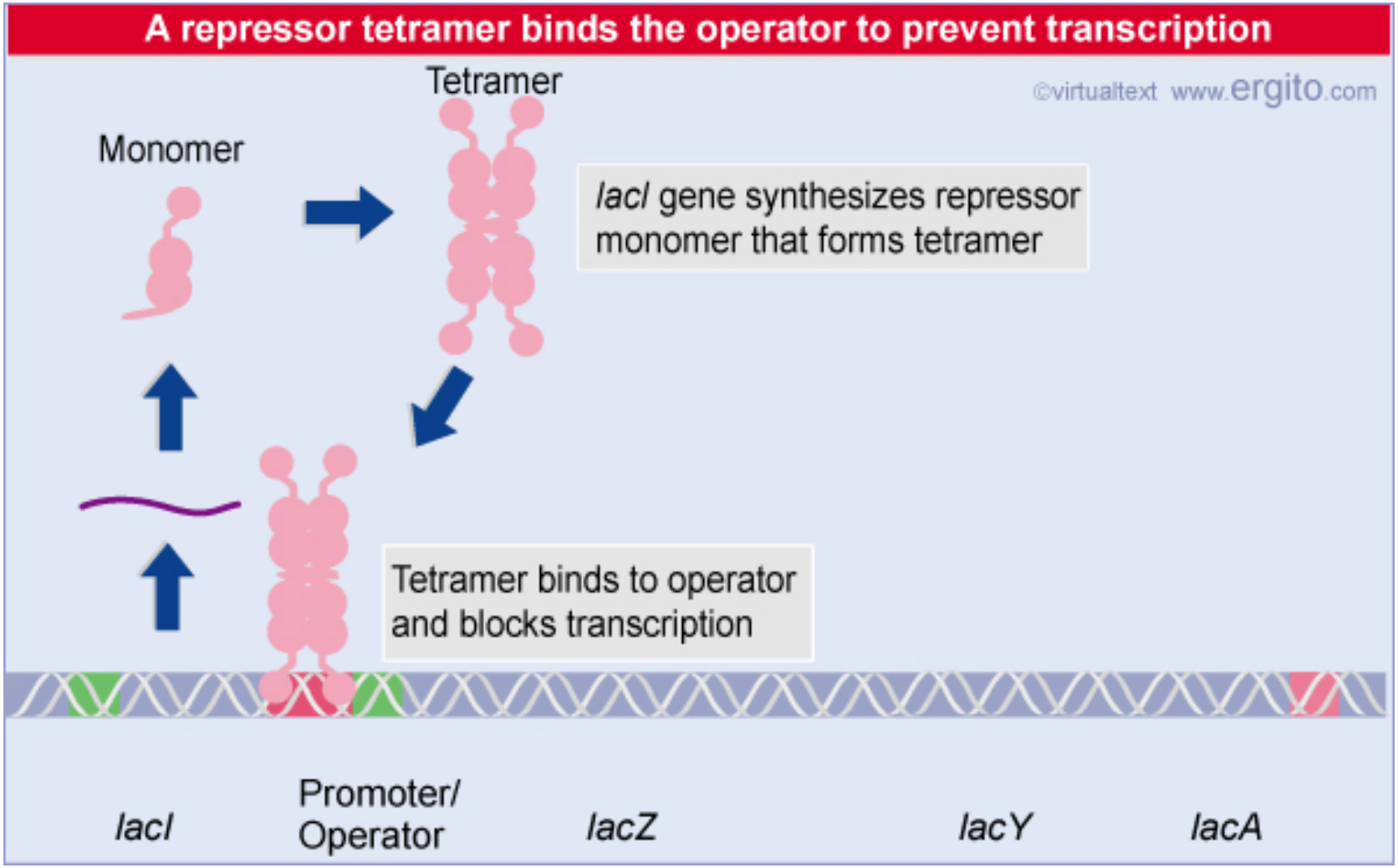
Gene isolation
lac operon

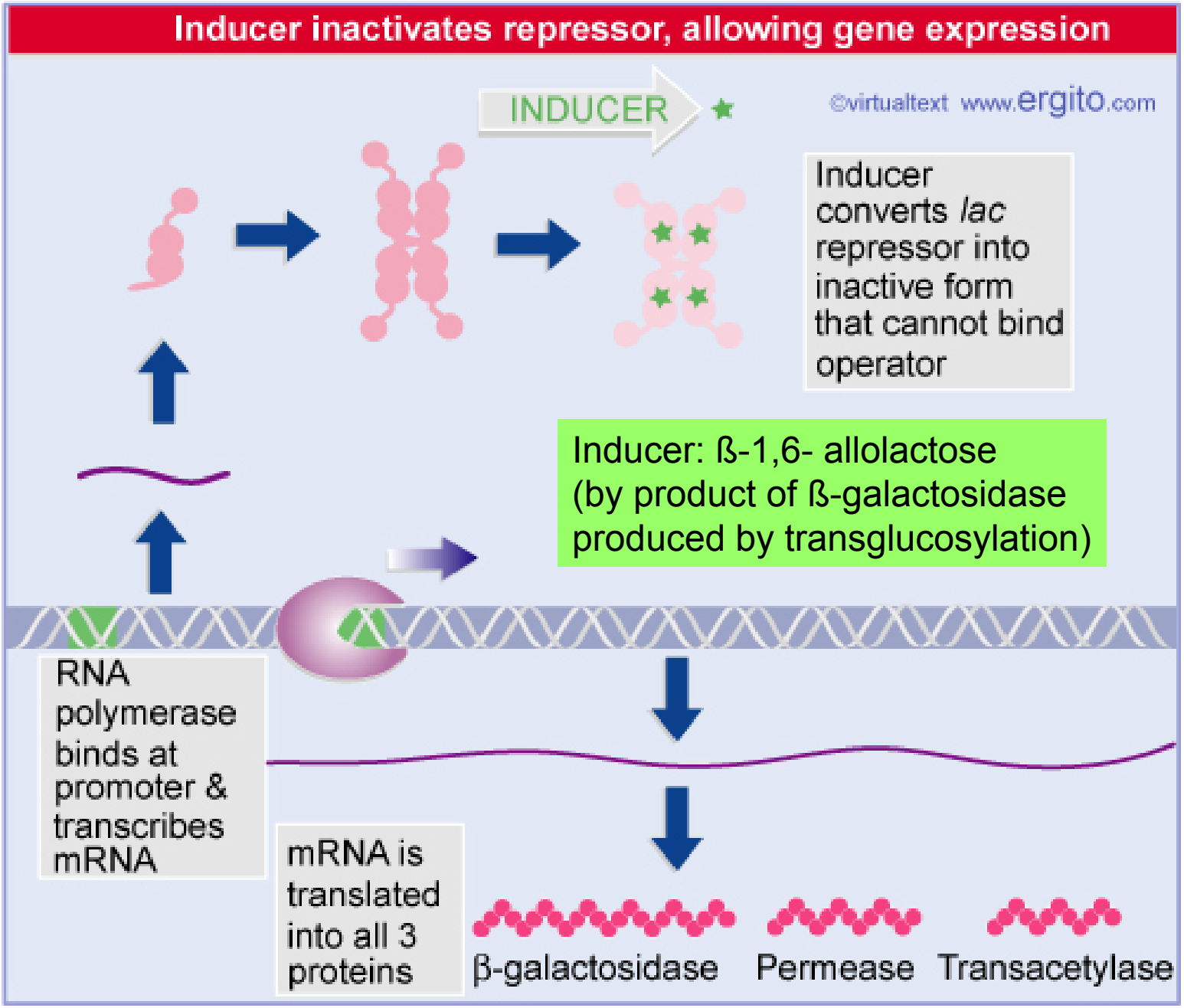
Isolation of Lac Repressor
lacI^{q} mutant

Binding studies



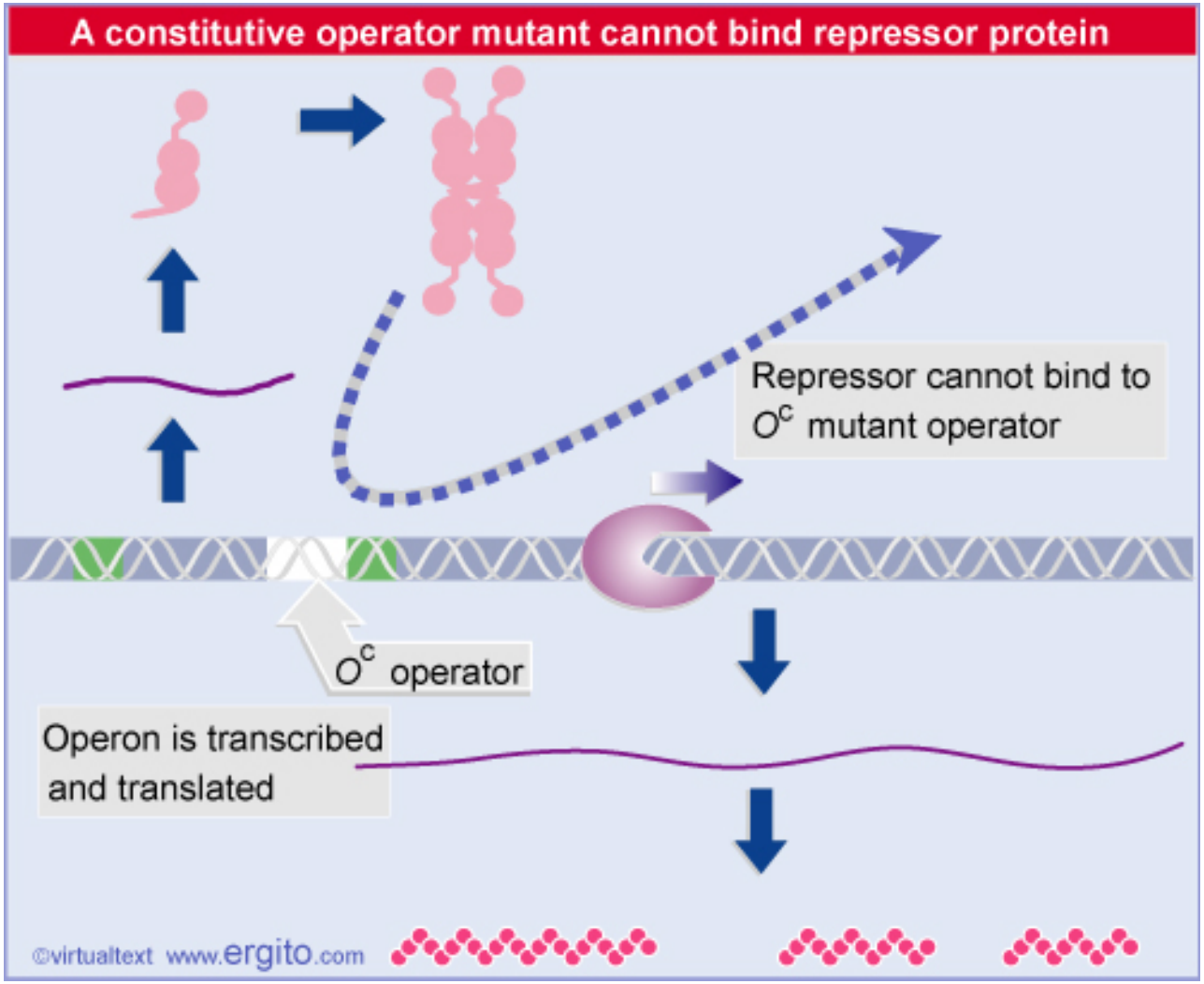

The promoter and operator overlap
Startpoint
Unwinding
-50 -40 -30 -20 -10 1 10 20 30
Promoter
binds RNA polymerase
Operator
binds repressor
©virtualtext www.ergito.com


A repressor tetramer binds the operator to prevent transcription
Tetramer
©virtualtext www.ergito.com
Monomer
lacI gene synthesizes repressor monomer that forms tetramer
Tetramer binds to operator and blocks transcription
lacI
Promoter/ Operator
lacZ
lacY
lacA


Inducer inactivates repressor, allowing gene expression
INDUCER
©virtualtext www.ergito.com
Inducer converts *lac* repressor into inactive form that cannot bind operator
Inducer: ß-1,6- allolactose (by product of ß-galactosidase produced by transglucosylation)
RNA polymerase binds at promoter & transcribes mRNA
mRNA is translated into all 3 proteins
β-galactosidase
Permease
Transacetylase

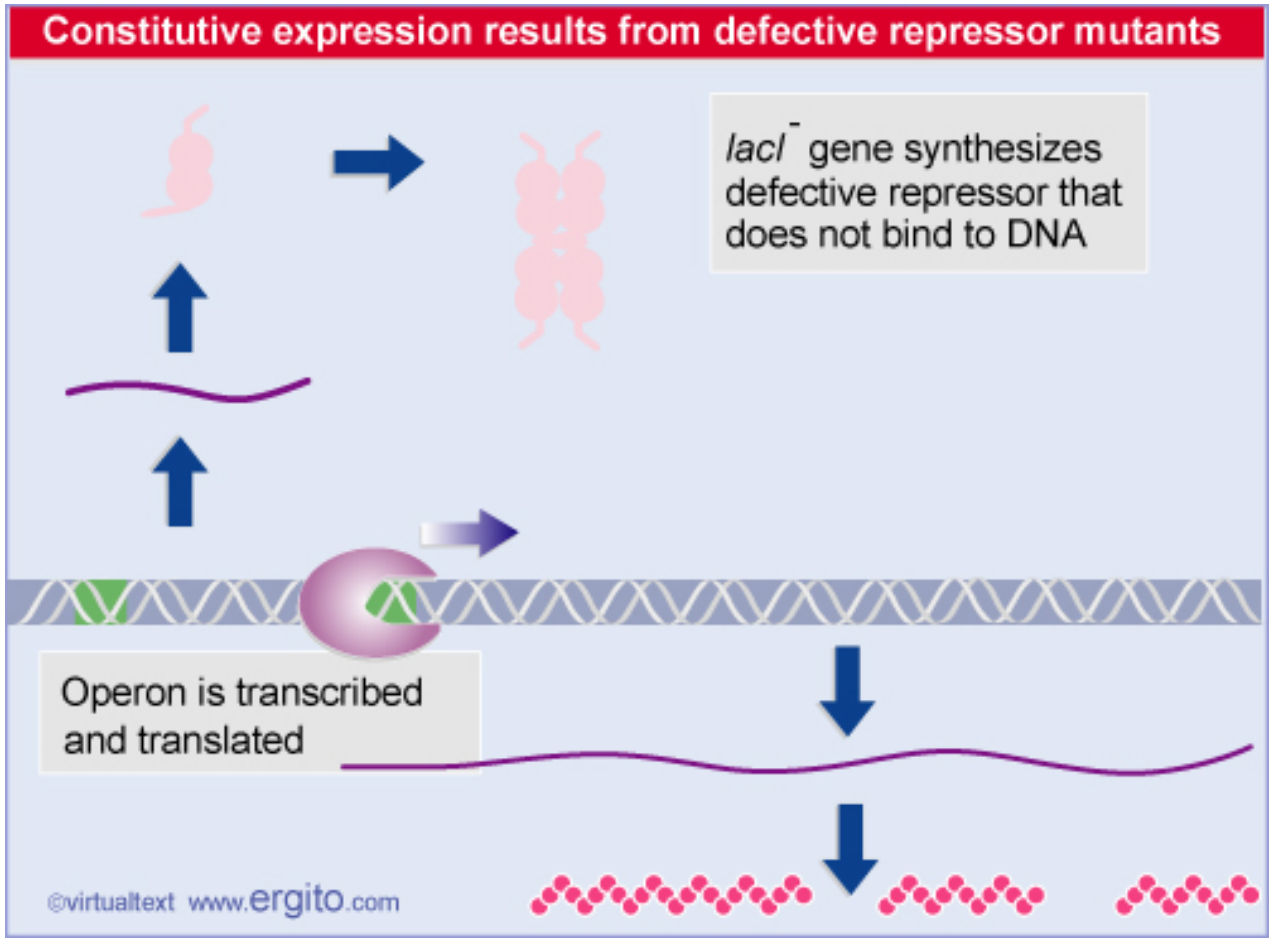
Mutant O^c

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



Mutant *I* $^{-}$

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



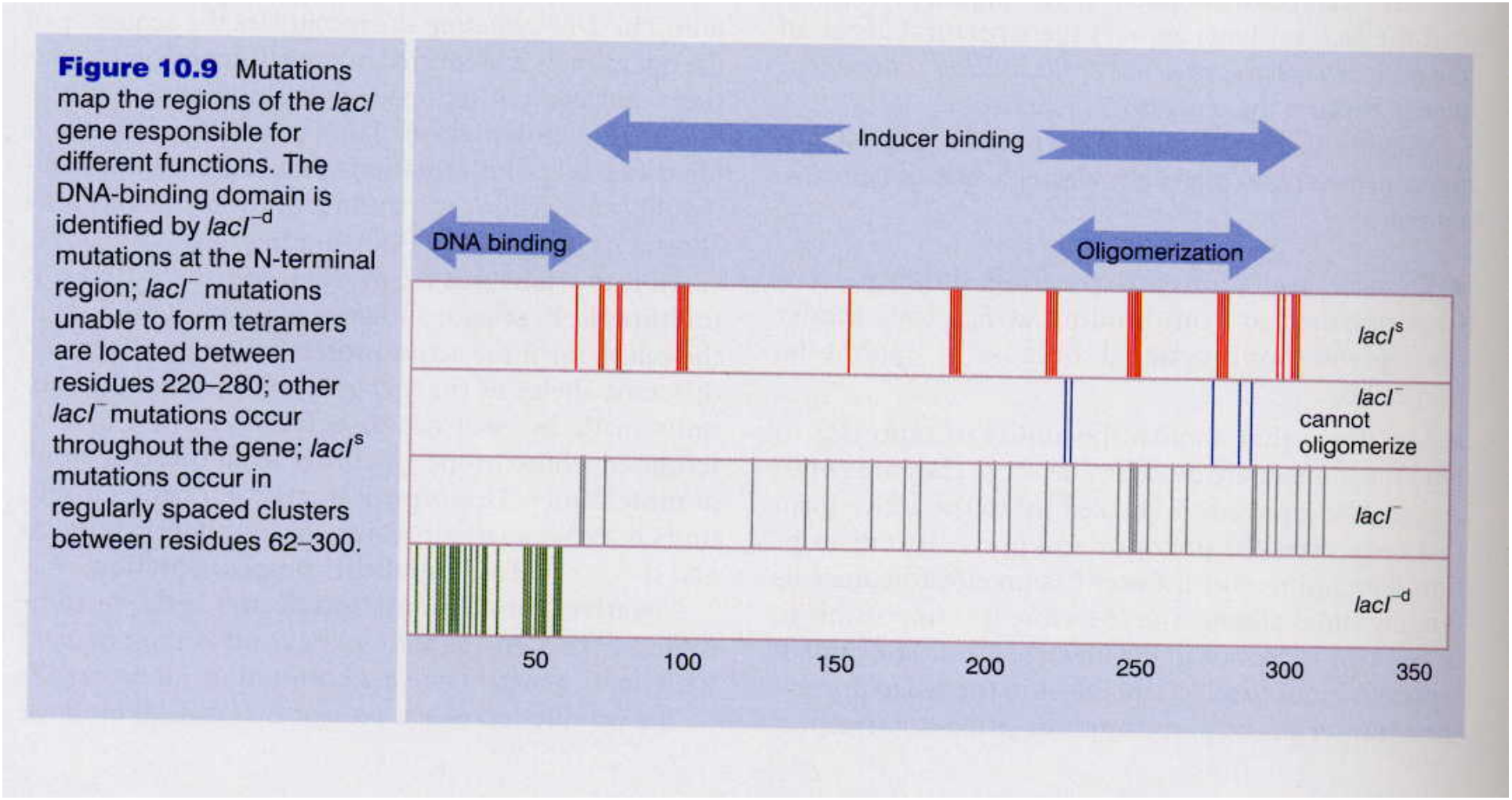
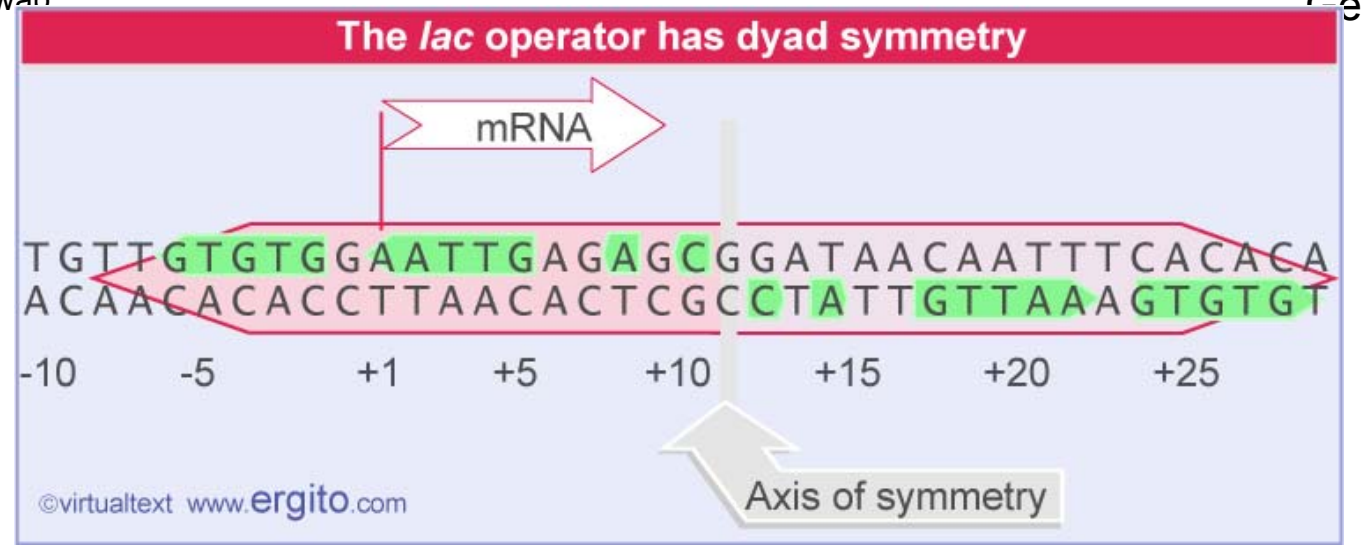
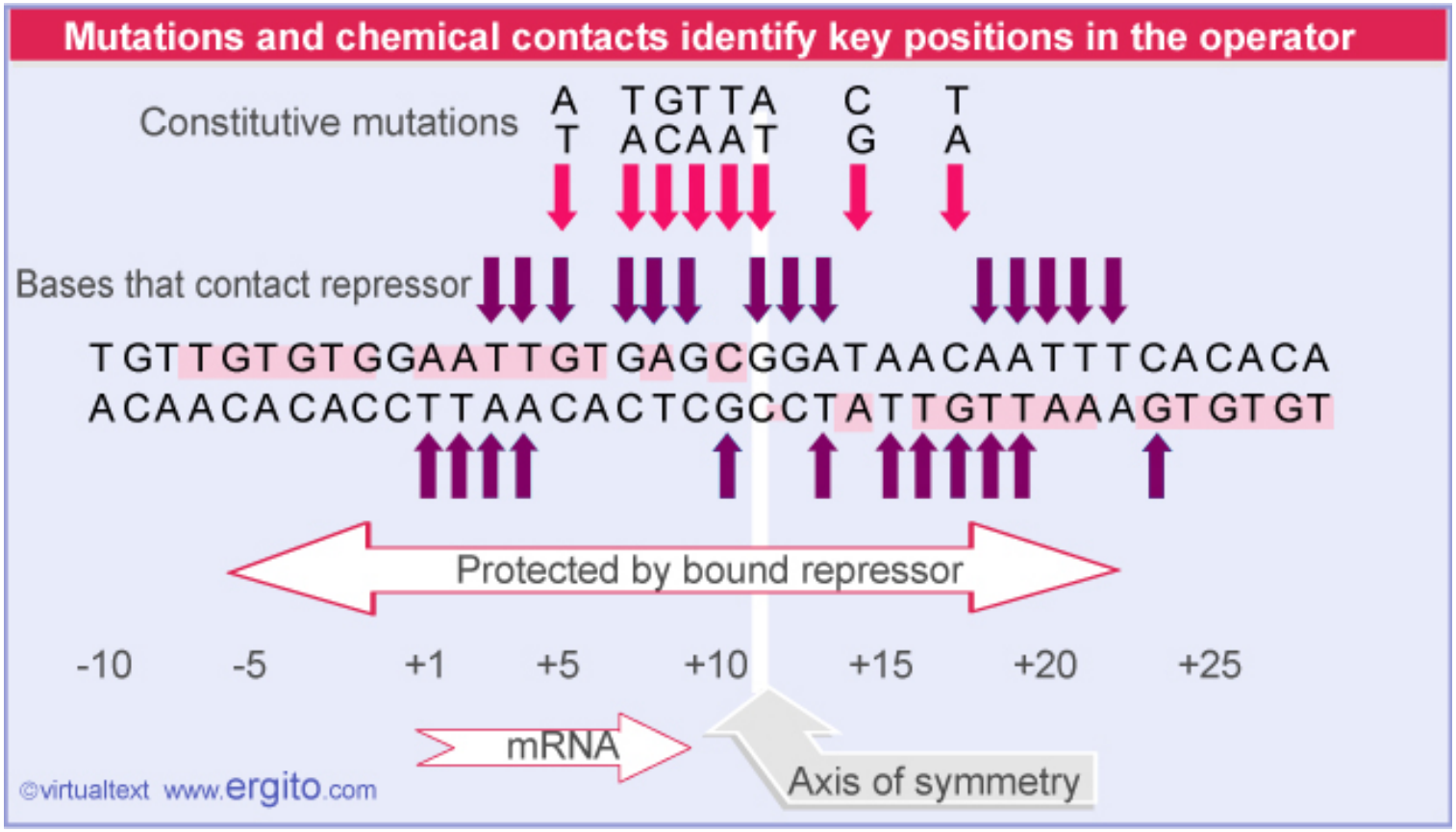
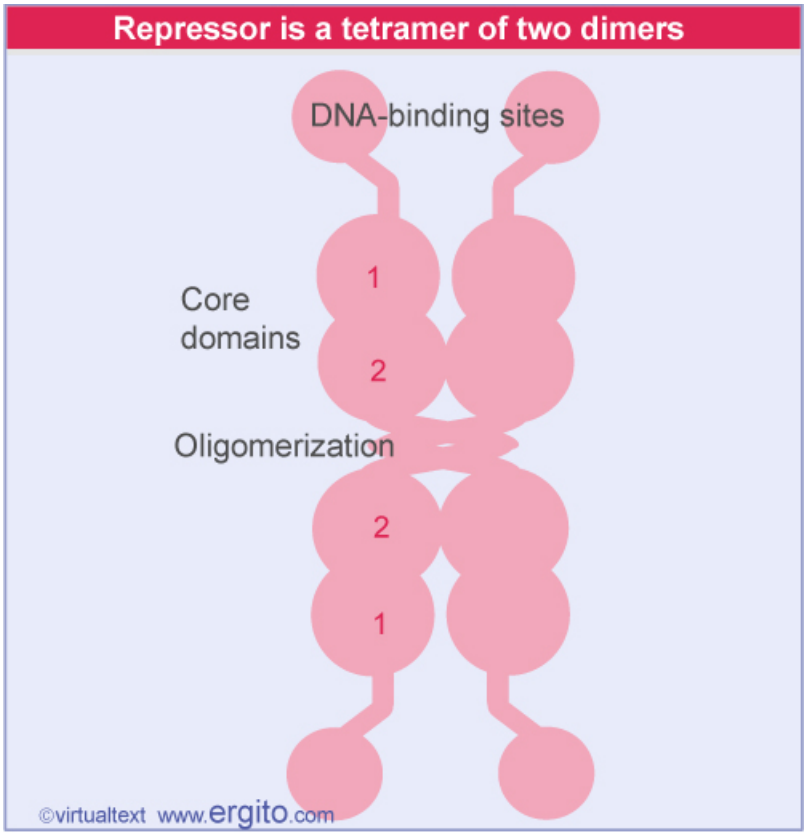
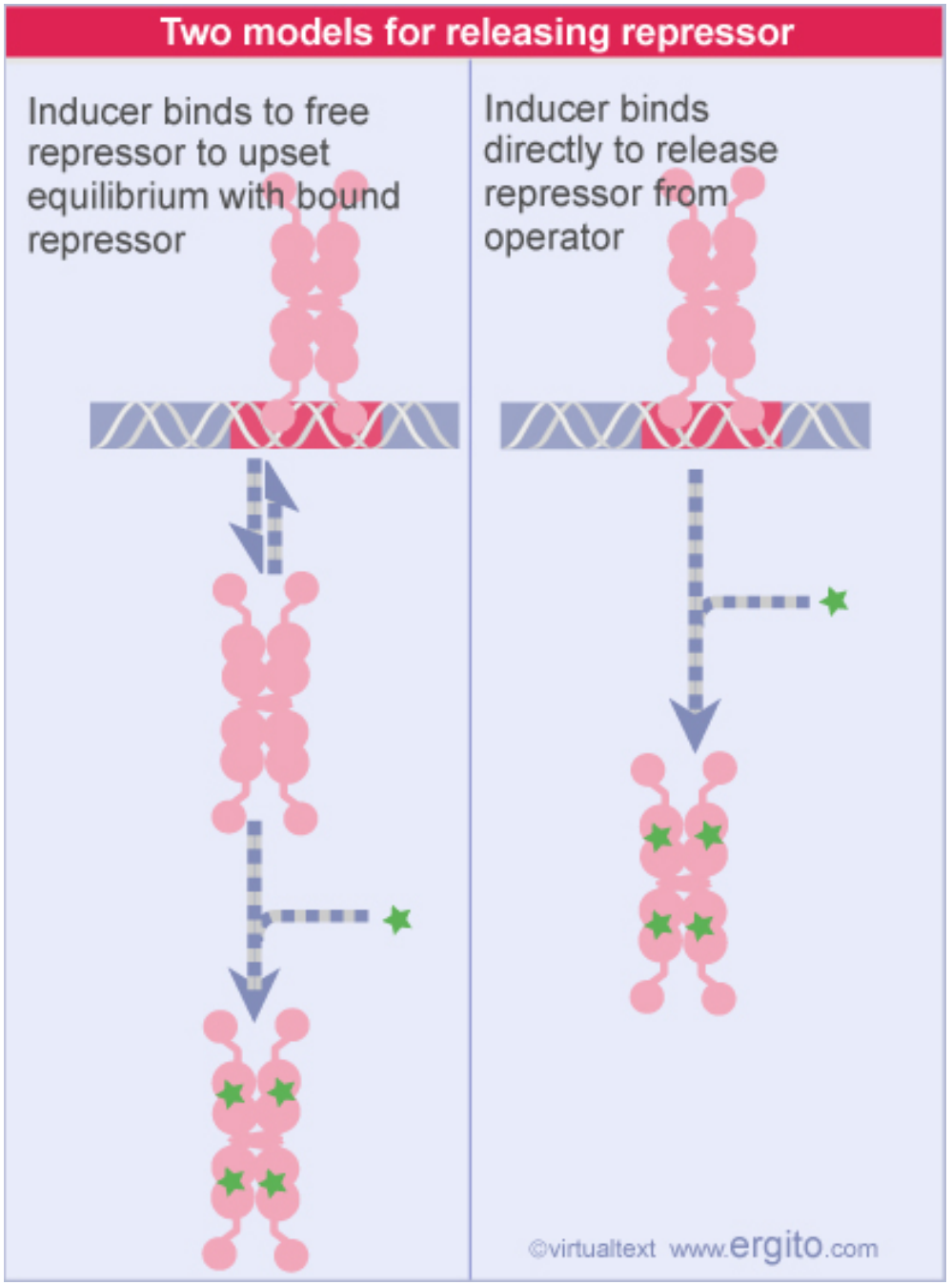


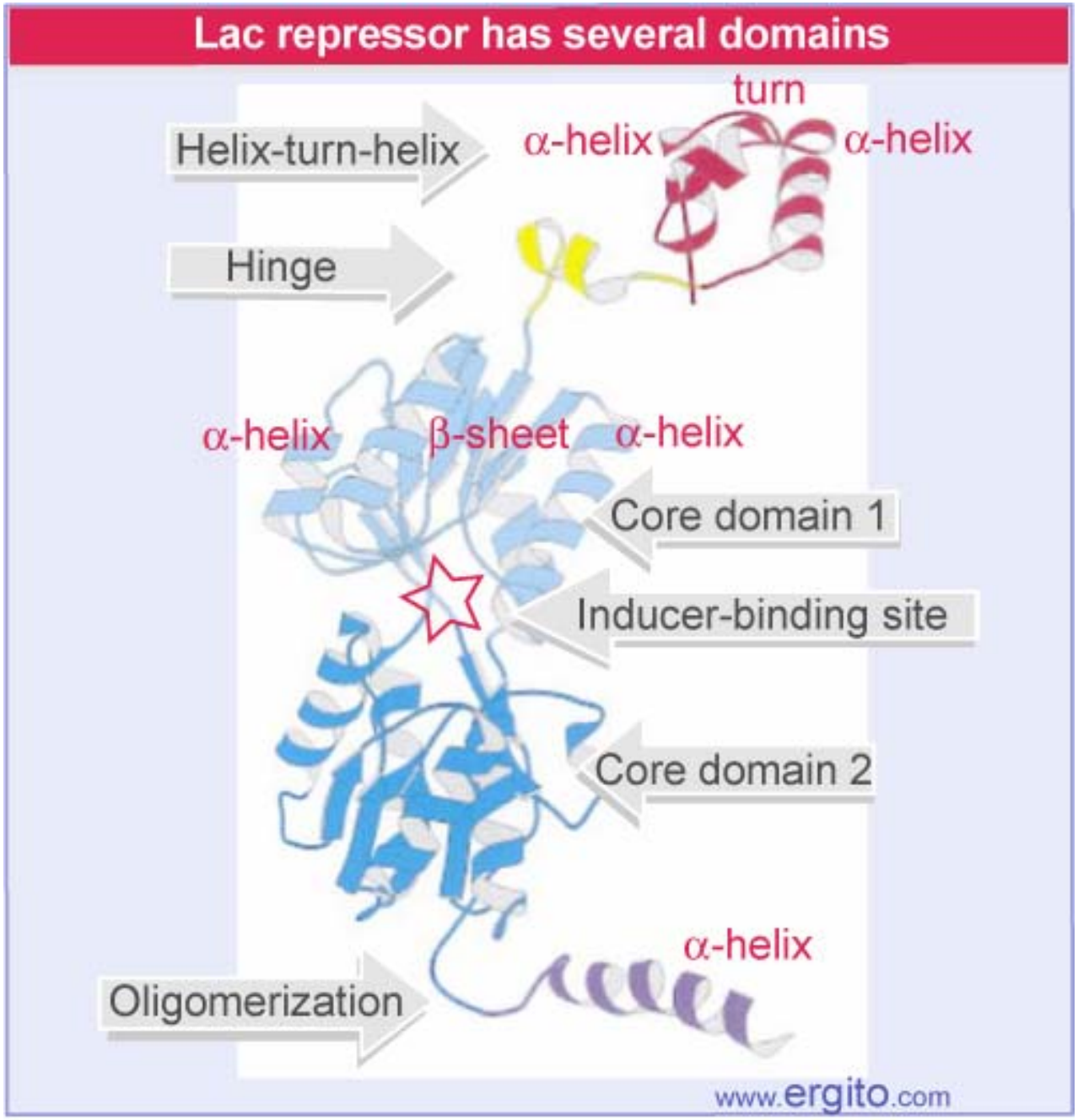
Figure 10.9 Mutations map the regions of the *lacI* gene responsible for different functions. The DNA-binding domain is identified by *lacI*$^{-d}$ mutations at the N-terminal region; *lacI*$^{-}$ mutations unable to form tetramers are located between residues 220–280; other *lacI*$^{-}$ mutations occur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

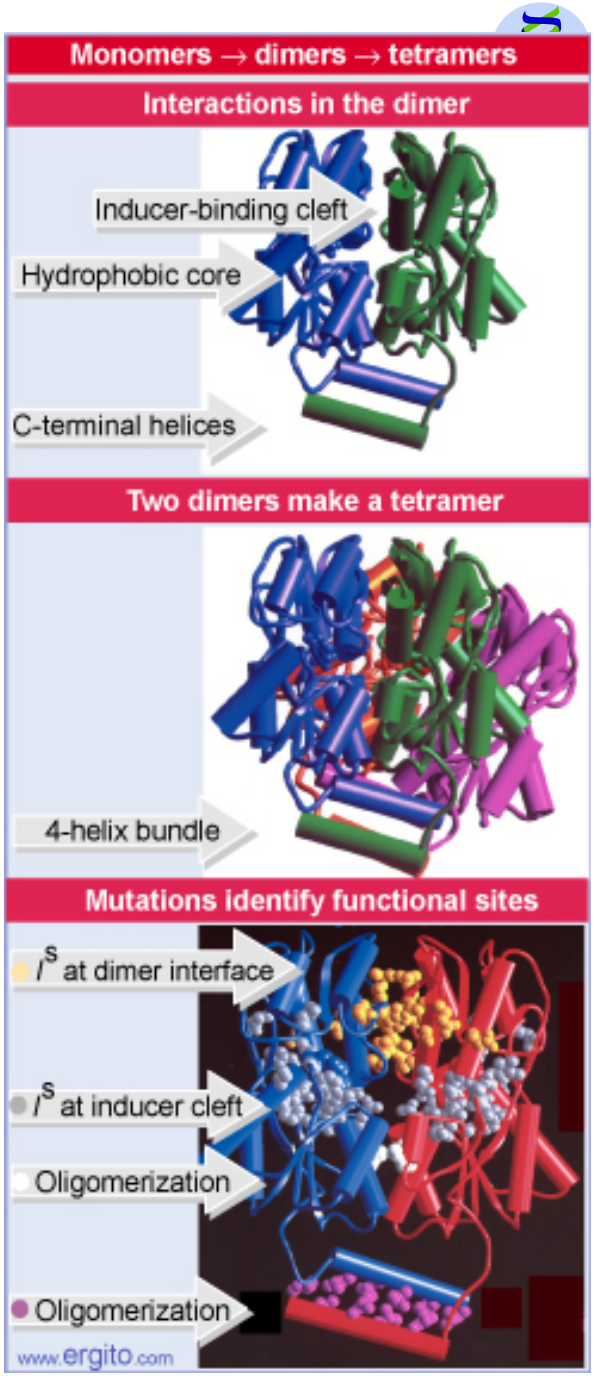

The lac operator has dyad symmetry
mRNA
TGTTGTGTGGAATTGAGAGCGGATAACAATTTCACACA
ACAACACACCTTAACACTCGCCTATTGTTAAAGTGTGT
-10 -5 +1 +5 +10 +15 +20 +25
©virtualtext www.ergito.com
Axis of symmetry


Mutations and chemical contacts identify key positions in the operator
Constitutive mutations
A T
TGTTA ACAAT
C G
T A
Bases that contact repressor
TGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACA
ACAACACACCTTAACACTCGCCTATTGTTAAAGTGTGT
Protected by bound repressor
-10 -5 +1 +5 +10 +15 +20 +25
mRNA
©virtualtext www.ergito.com
Axis of symmetry


Repressor is a tetramer of two dimers
DNA-binding sites
Core domains
1
2
Oligomerization
2
1
©virtualtext www.ergito.com


Two models for releasing repressor
Inducer binds to free repressor to upset equilibrium with bound repressor
Inducer binds directly to release repressor from operator
©virtualtext www.ergito.com


Lac repressor has several domains
turn
Helix-turn-helix
α-helix
α-helix
Hinge
α-helix
β-sheet
α-helix
Core domain 1
Inducer-binding site
Core domain 2
α-helix
Oligomerization
www.ergito.com


Monomers → dimers → tetramers
Interactions in the dimer
Inducer-binding cleft
Hydrophobic core
C-terminal helices
Two dimers make a tetramer
4-helix bundle
Mutations identify functional sites
I^S at dimer interface
I^S at inducer cleft
Oligomerization
Oligomerization
www.ergito.com

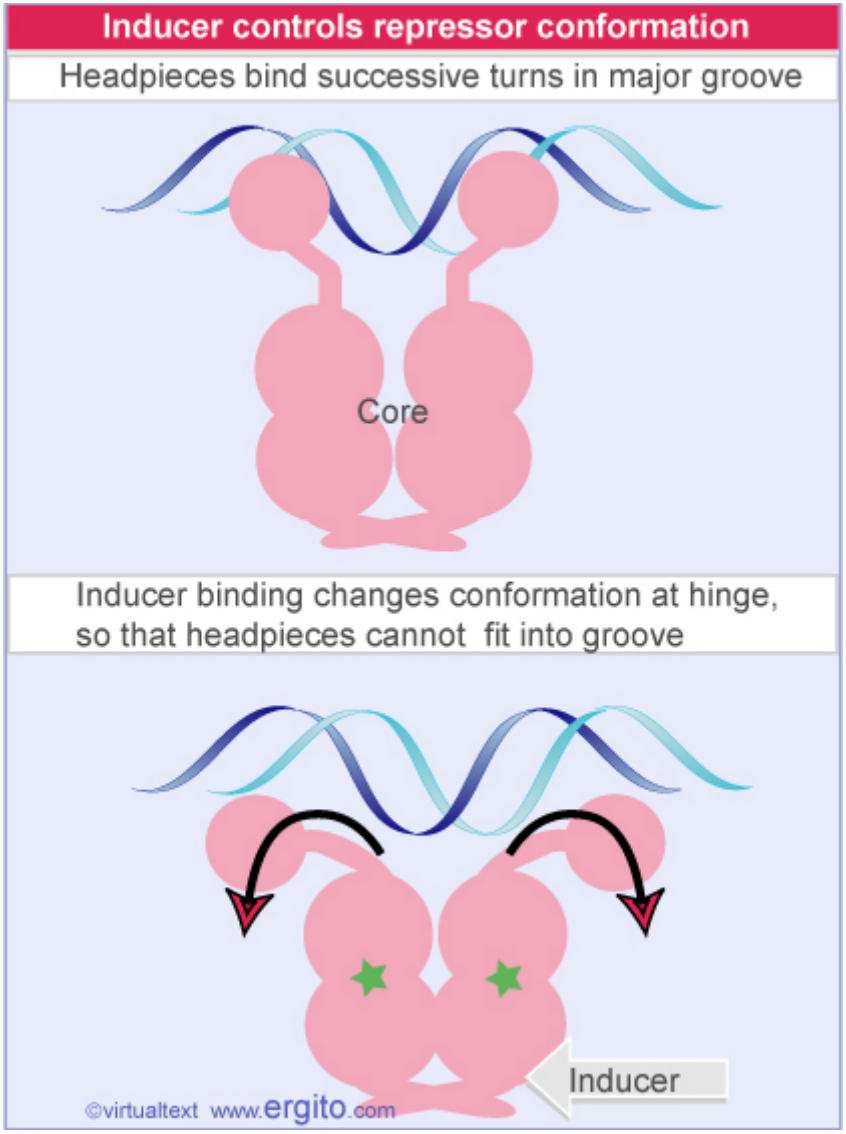
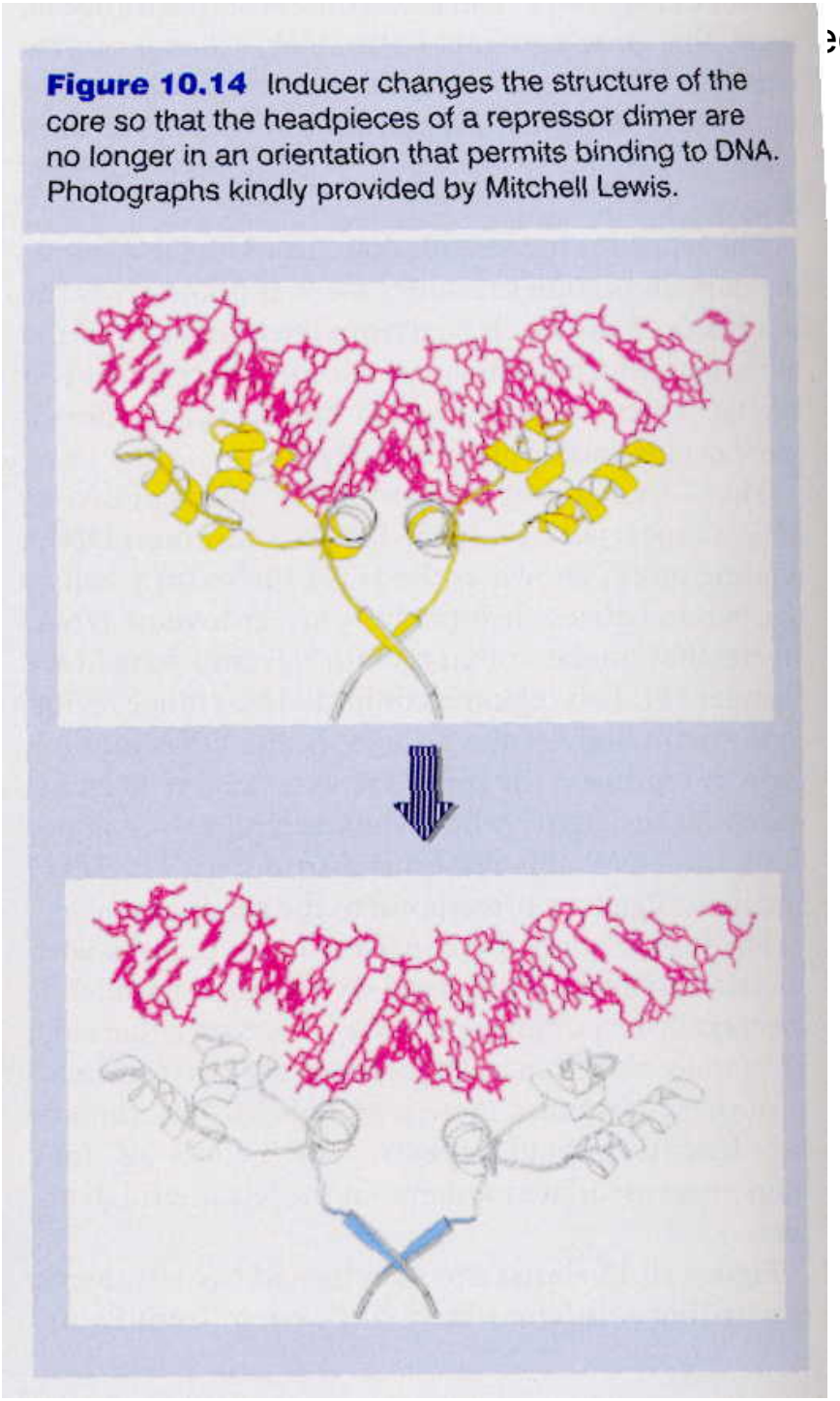
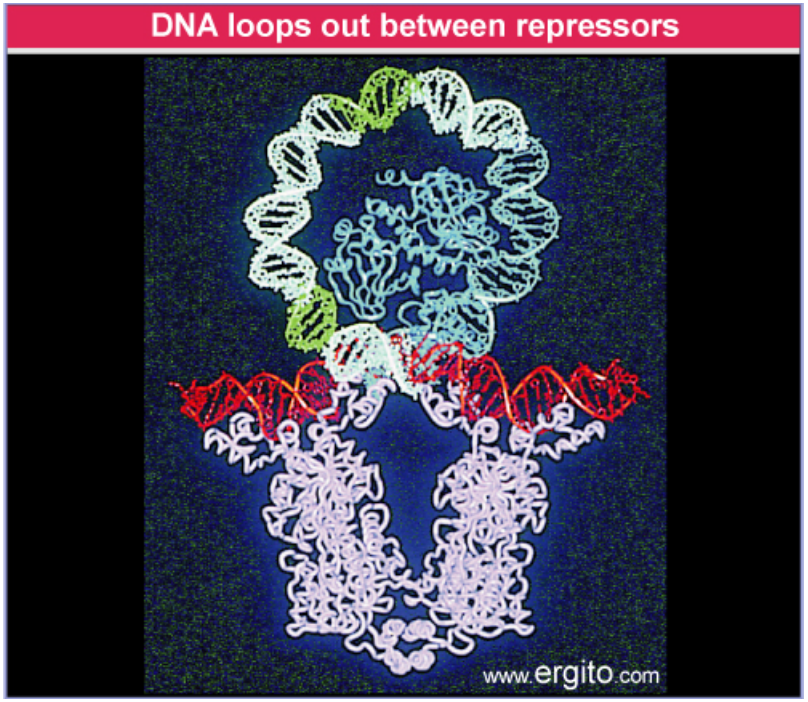
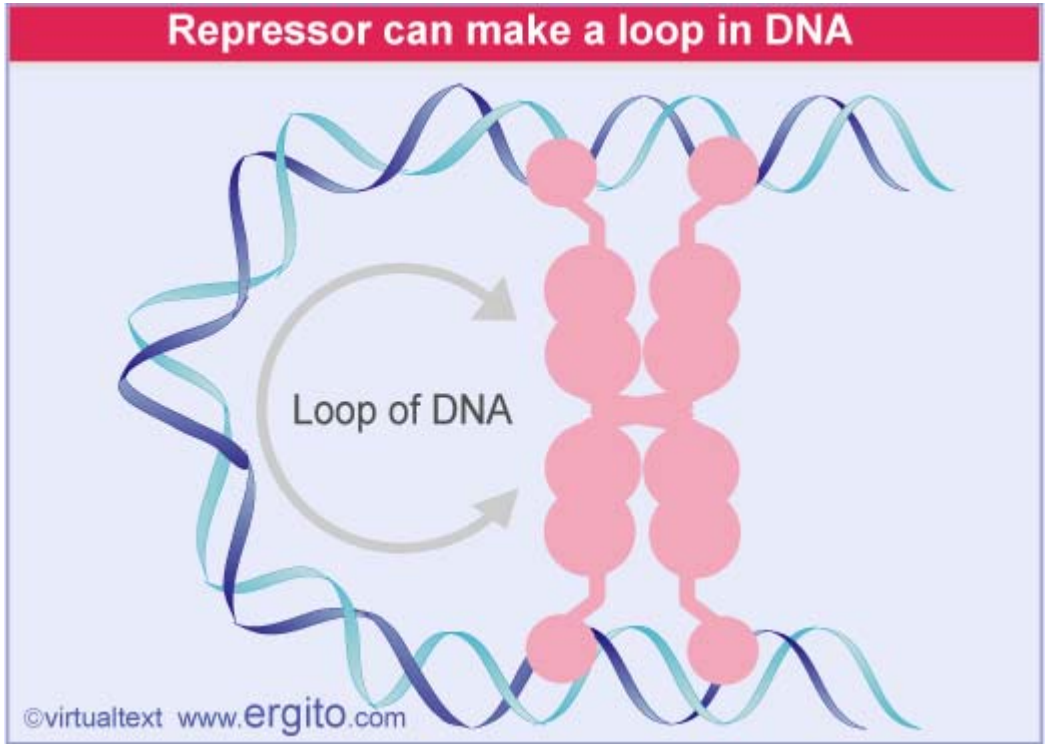
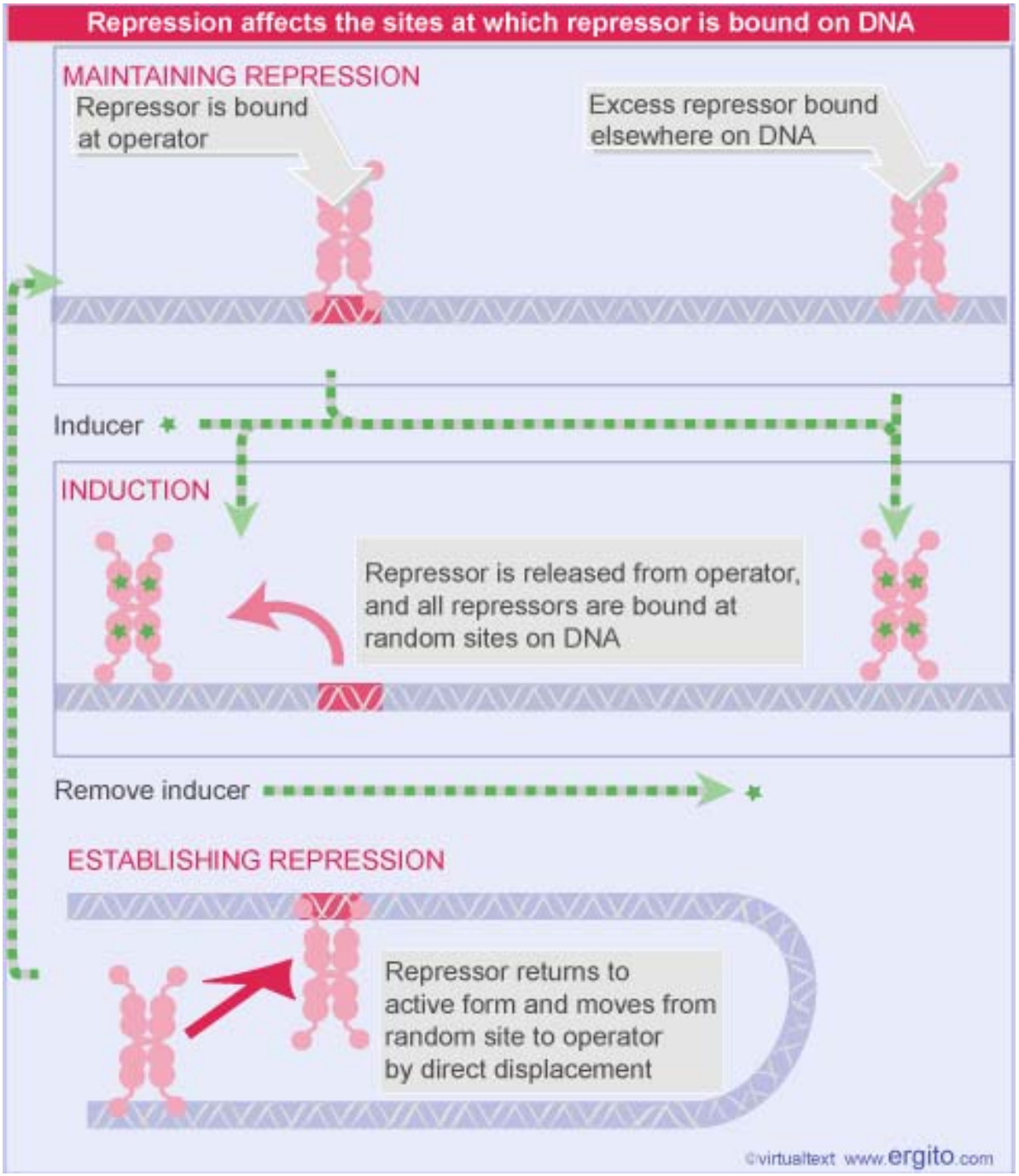


Figure 10.14 Inducer changes the structure of the core so that the headpieces of a repressor dimer are no longer in an orientation that permits binding to DNA. Photographs kindly provided by Mitchell Lewis.



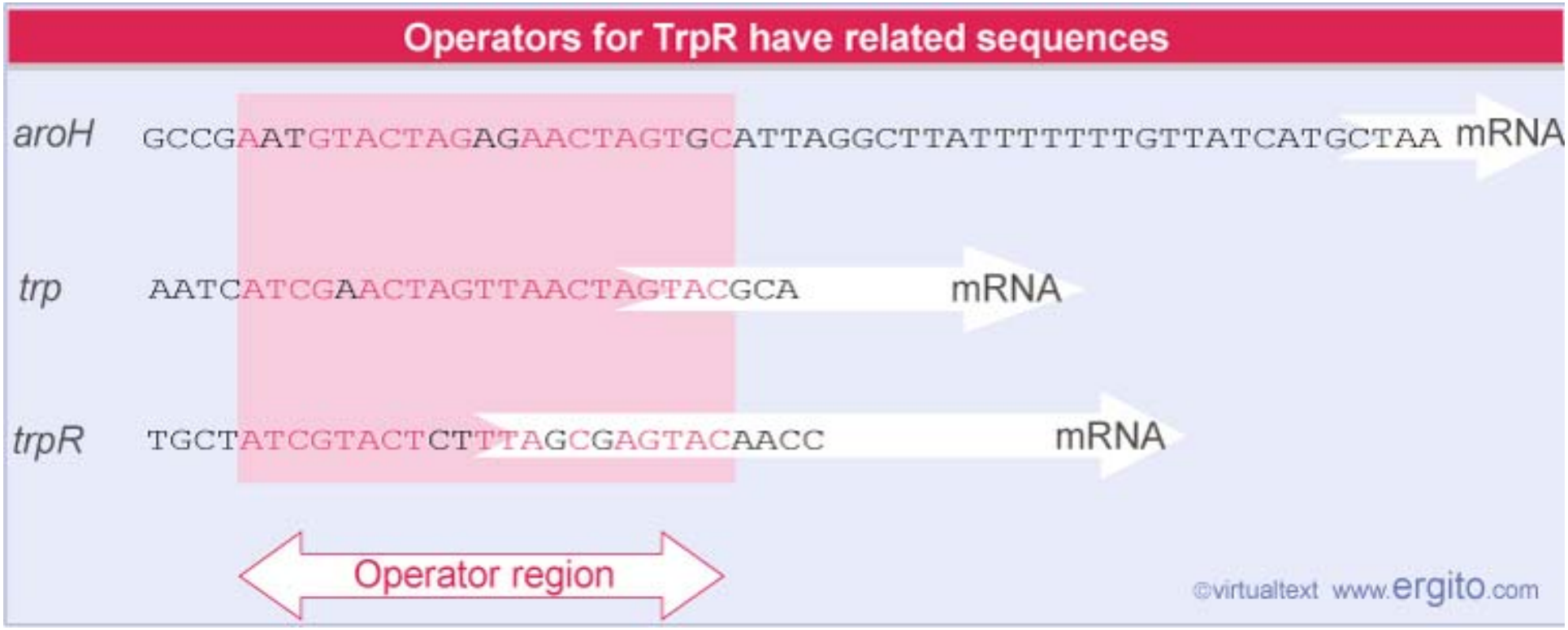

DNA loops out between repressors
www.ergito.com

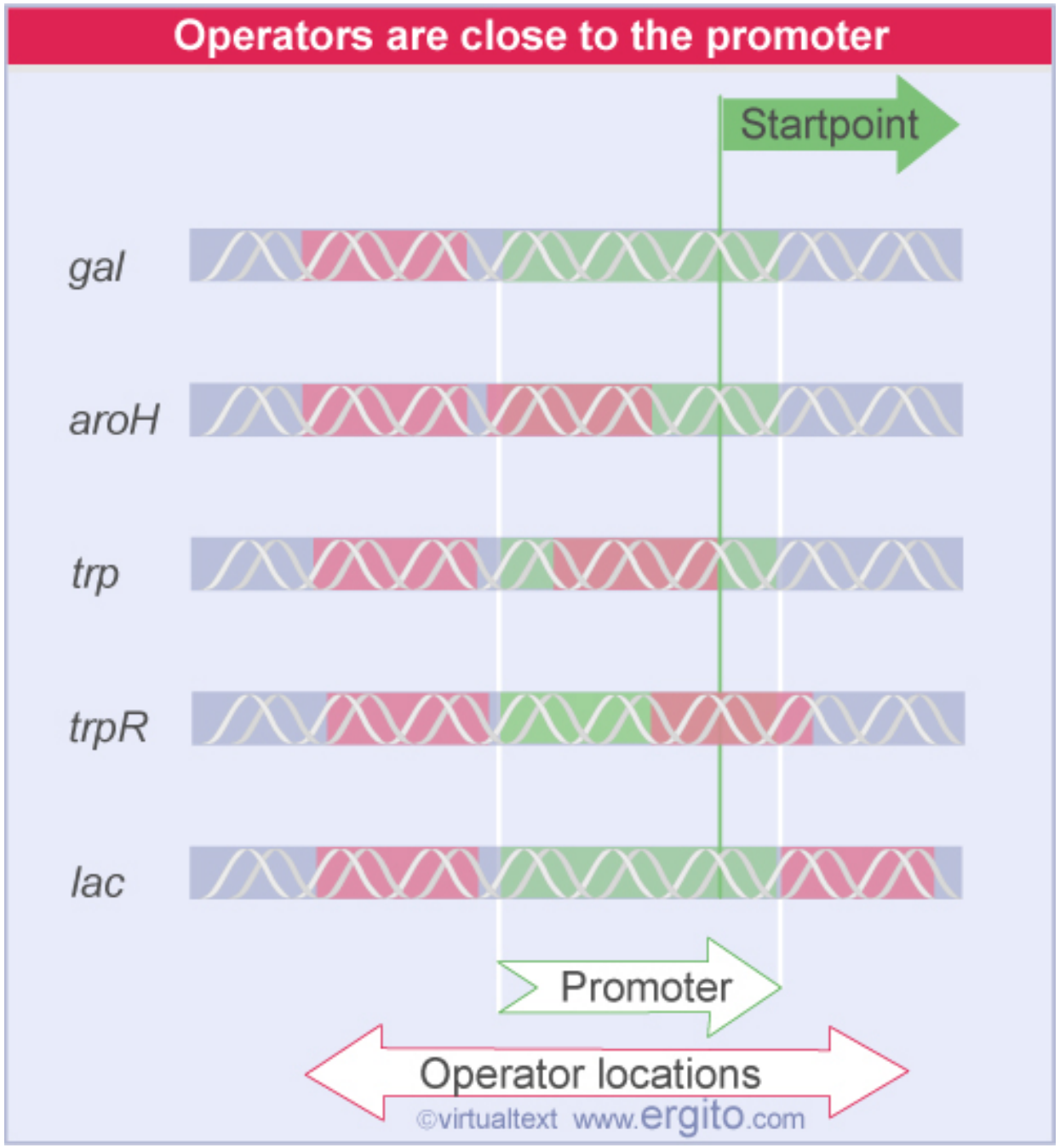

Repressor can make a loop in DNA
Loop of DNA
©virtualtext www.ergito.com



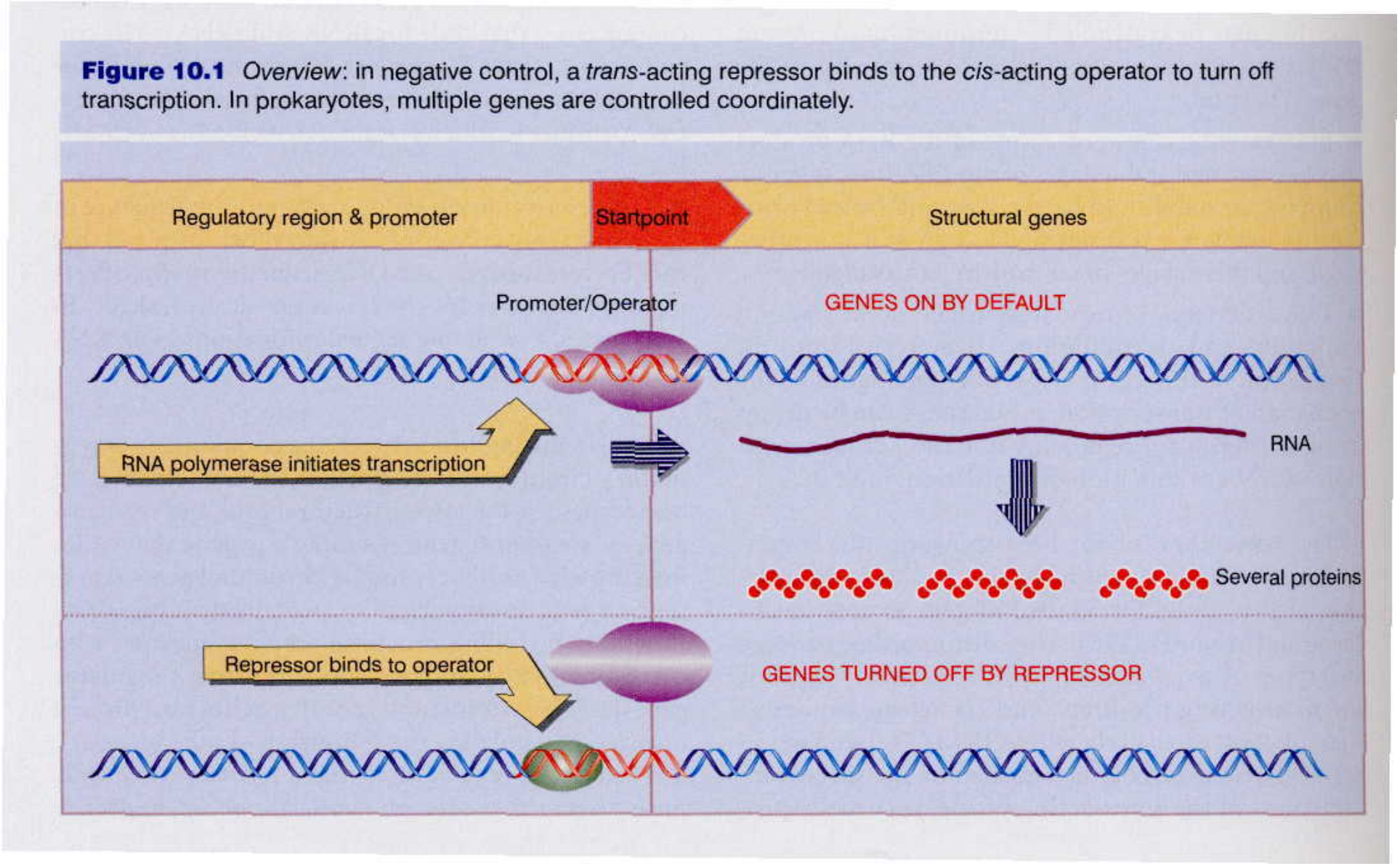
LacI repressor has general low affinity to DNA → Unspecific weak binding

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


Operators for TrpR have related sequences
aroH
GCCGAATGTACTAGAGAACTAGTGCATTAGGCTTATTTTTTTGTTATCATGCTAA
mRNA
trp
AATCATCGAACTAGTTAACTAGTACGCA
mRNA
trpR
TGCTATCGTACTCTTTAGCGAGTACAACC
mRNA
Operator region
©virtualtext www.ergito.com


Operators are close to the promoter
Startpoint
gal
aroH
trp
trpR
lac
Promoter
Operator locations
©virtualtext www.ergito.com

Negative Regulation



Positive Regulation

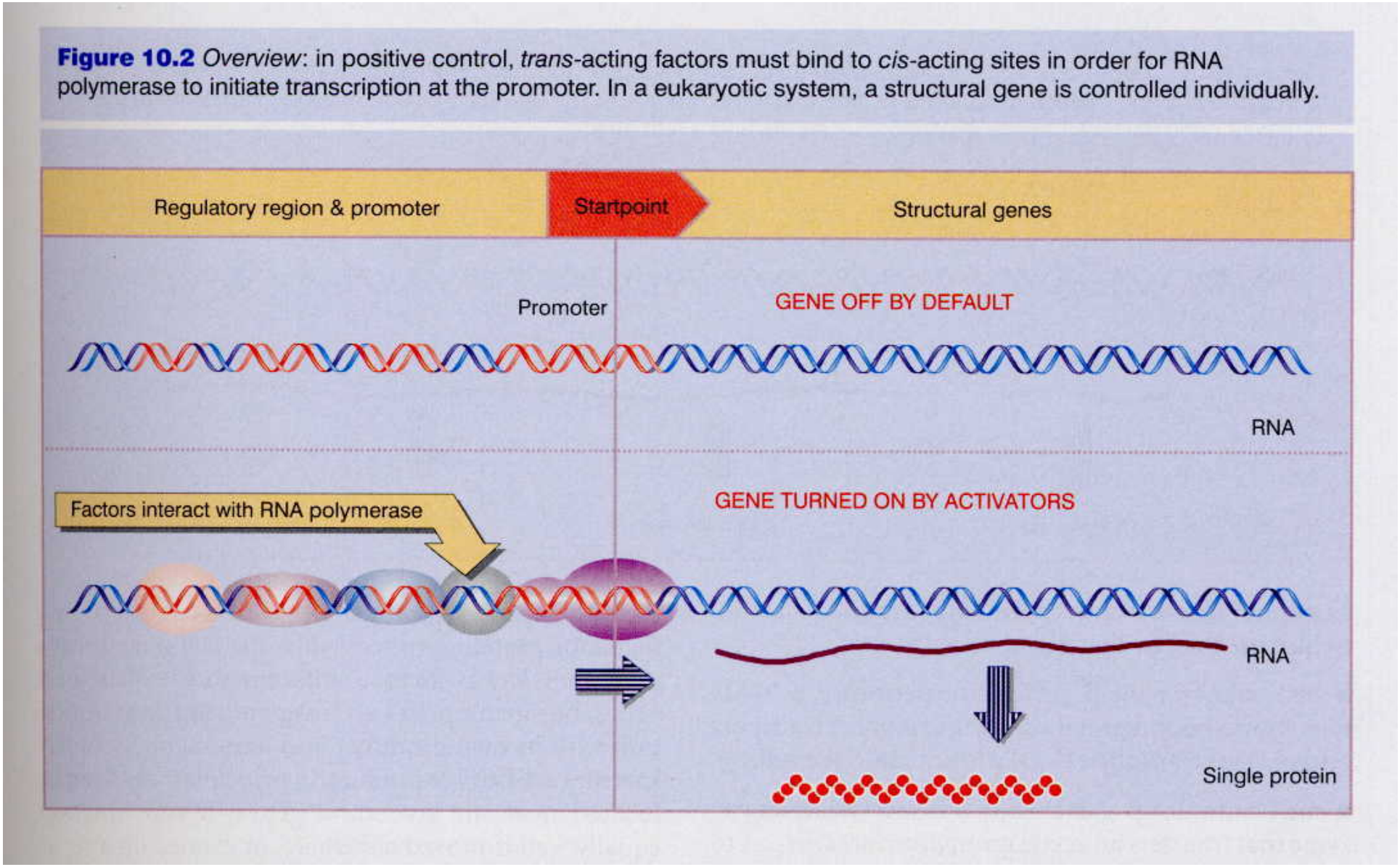
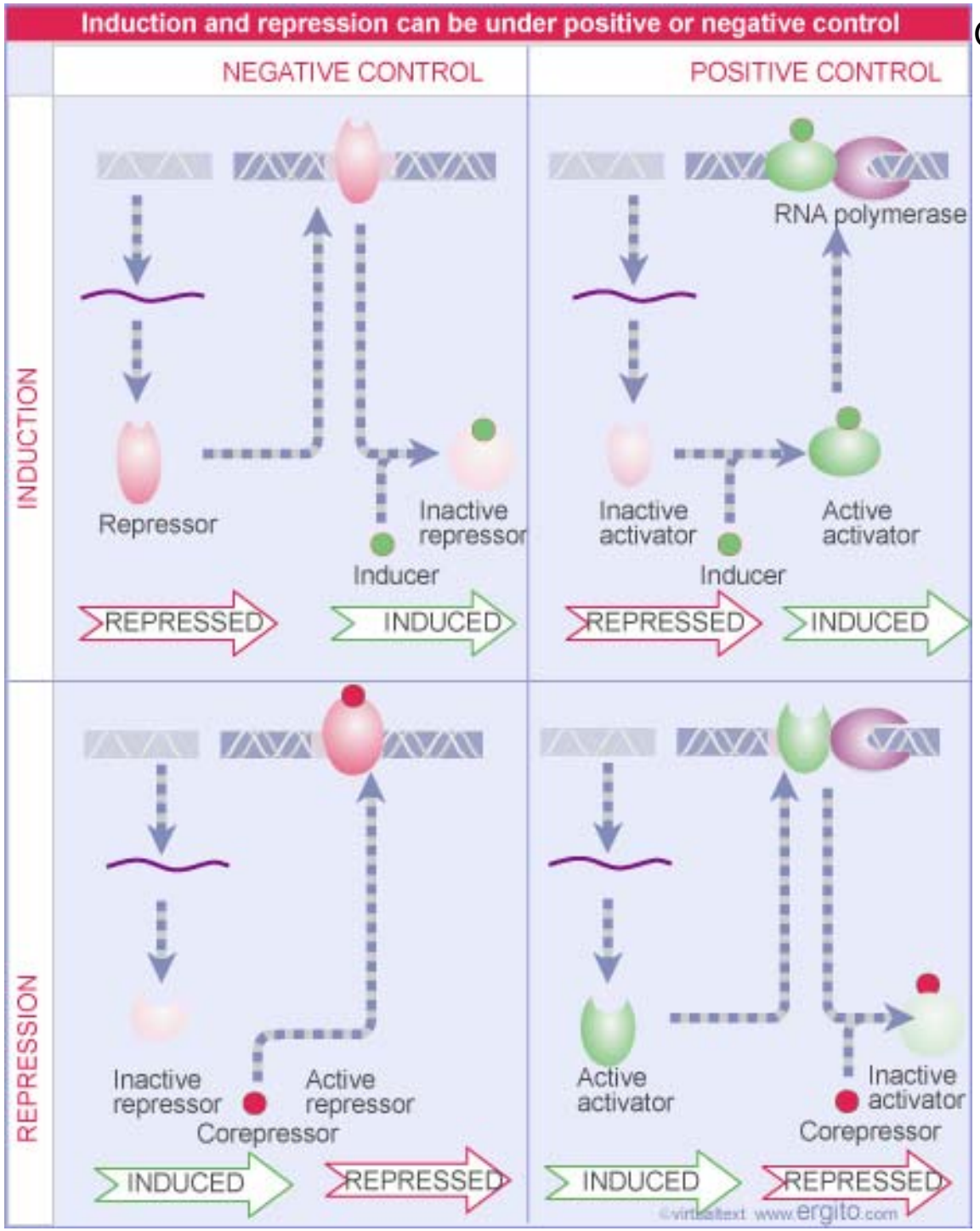
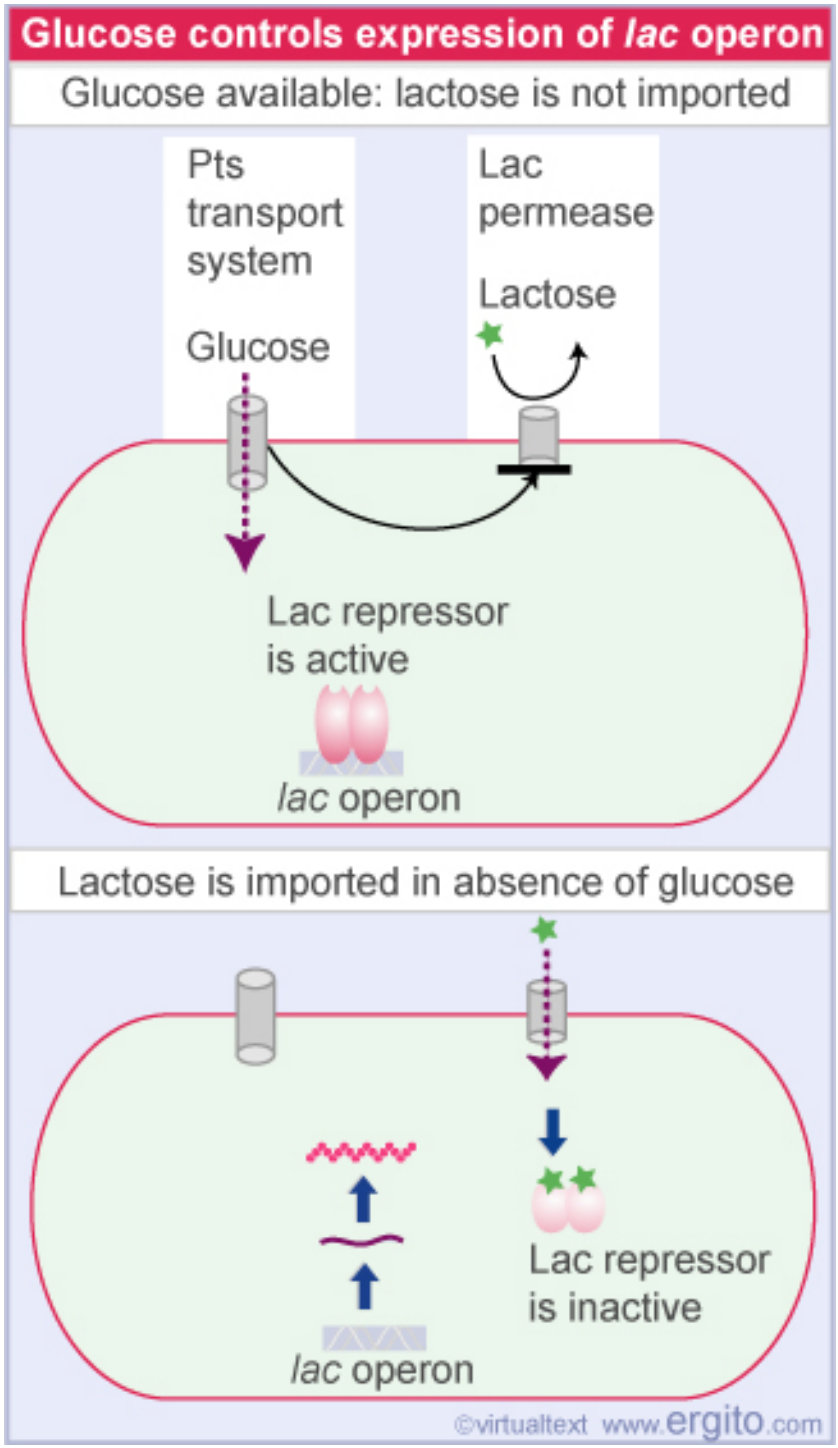


Figure 10.2 *Overview*: in positive control, *trans*-acting factors must bind to *cis*-acting sites in order for RNA polymerase to initiate transcription at the promoter. In a eukaryotic system, a structural gene is controlled individually.


Induction and repression can be under positive or negative control
NEGATIVE CONTROL
POSITIVE CONTROL
INDUCTION
REPRESSION
Repressor
Inactive repressor
Inducer
REPRESSED
INDUCED
RNA polymerase
Inactive activator
Active activator
Inducer
REPRESSED
INDUCED
Inactive repressor
Active repressor
Corepressor
INDUCED
REPRESSED
Active activator
Inactive activator
Corepressor
INDUCED
REPRESSED
www.ergito.com



Influence of Glucose on expression of *lac* Operon

Glucose controls import of lactose

and of other alternative carbon sources

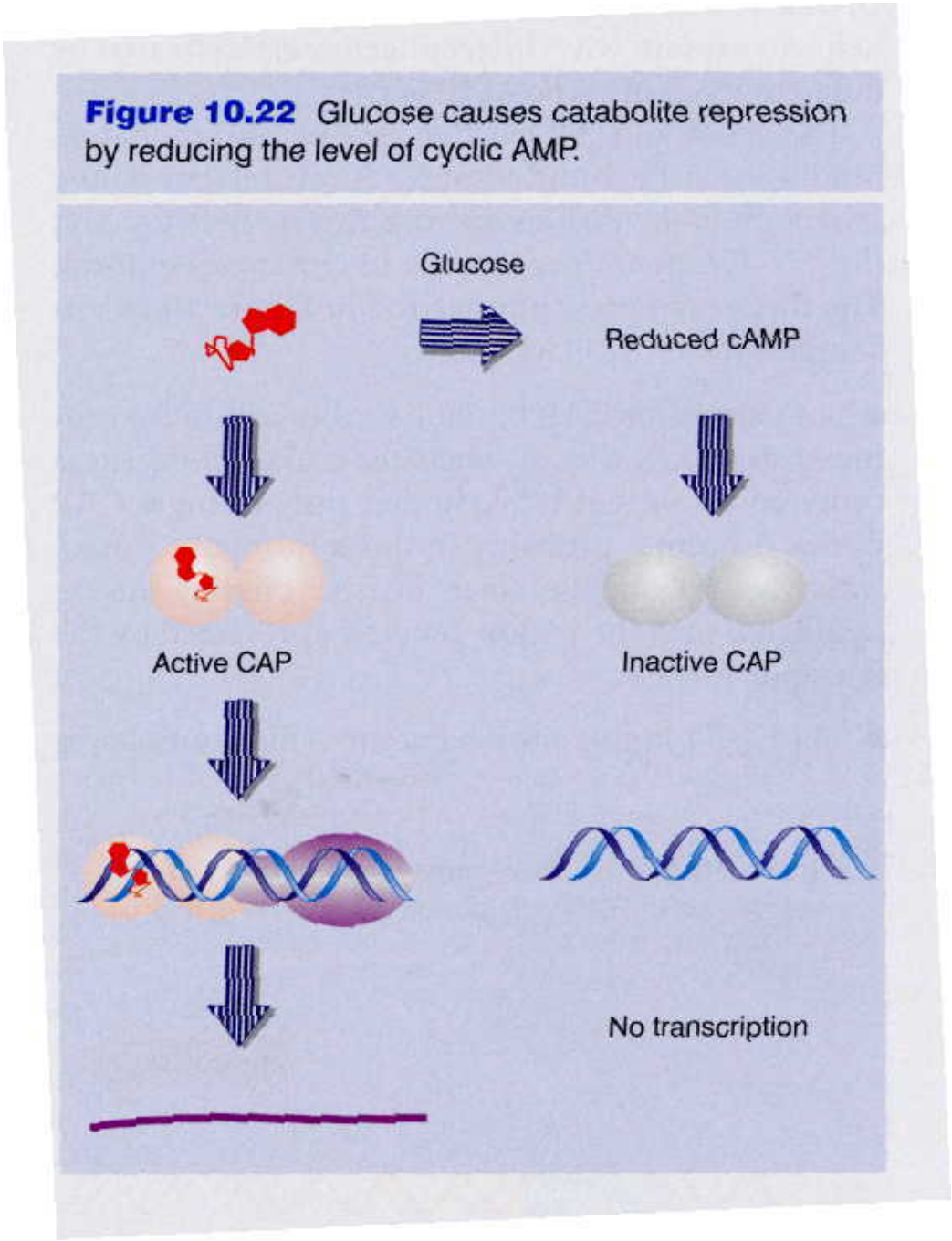


Figure 10.22 Glucose causes catabolite repression by reducing the level of cyclic AMP.

Carbon Catabolite Regulation

Cyclic AMP acts as an inducer

CAP (CRP) protein is a positive acting regulator protein

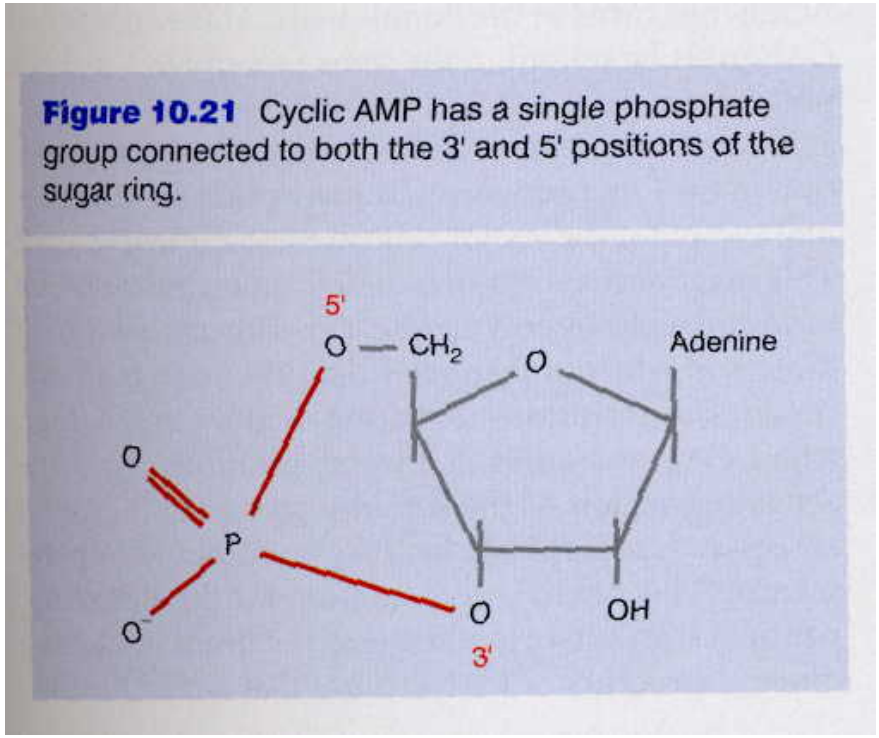


Figure 10.21 Cyclic AMP has a single phosphate group connected to both the 3' and 5' positions of the sugar ring.

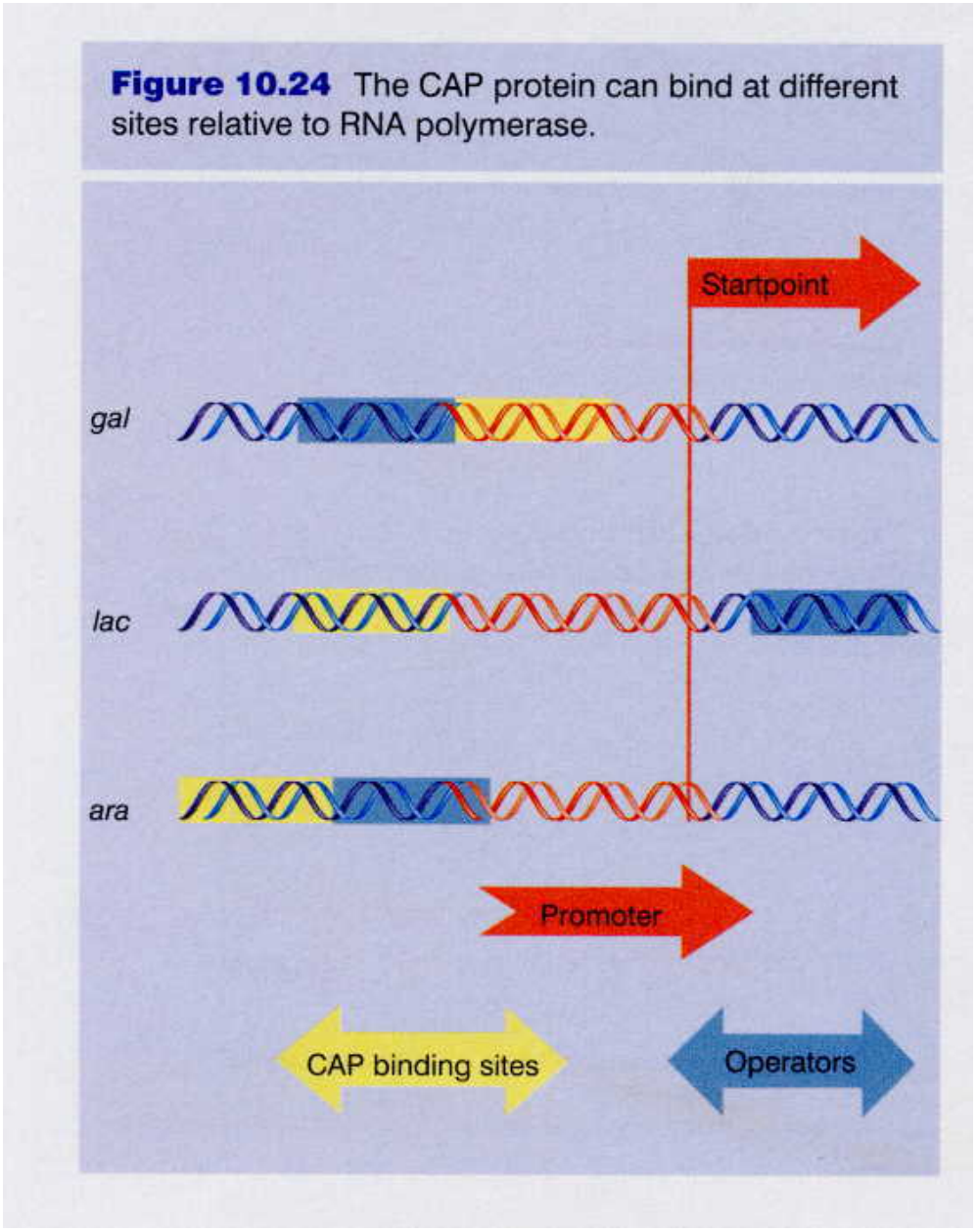


Figure 10.24 The CAP protein can bind at different sites relative to RNA polymerase.

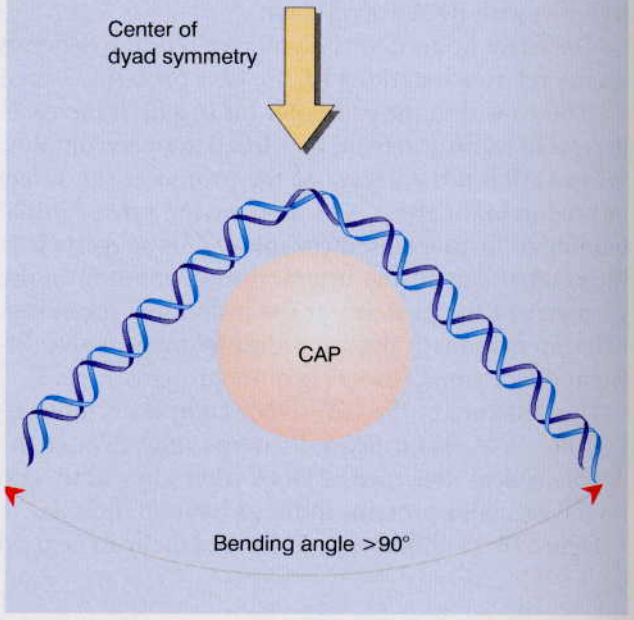


Figure 10.26 CAP bends DNA >90° around the center of symmetry.

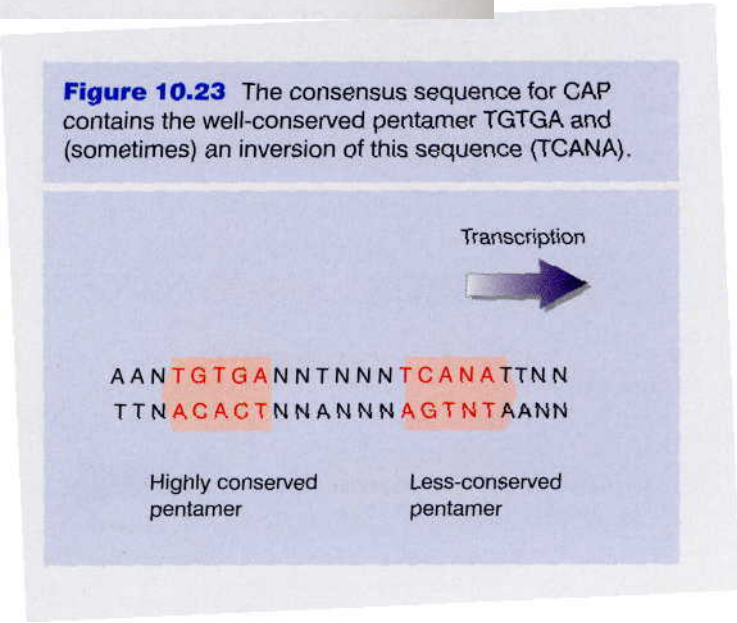


Figure 10.23 The consensus sequence for CAP contains the well-conserved pentamer TGTGA and (sometimes) an inversion of this sequence (TCANA).

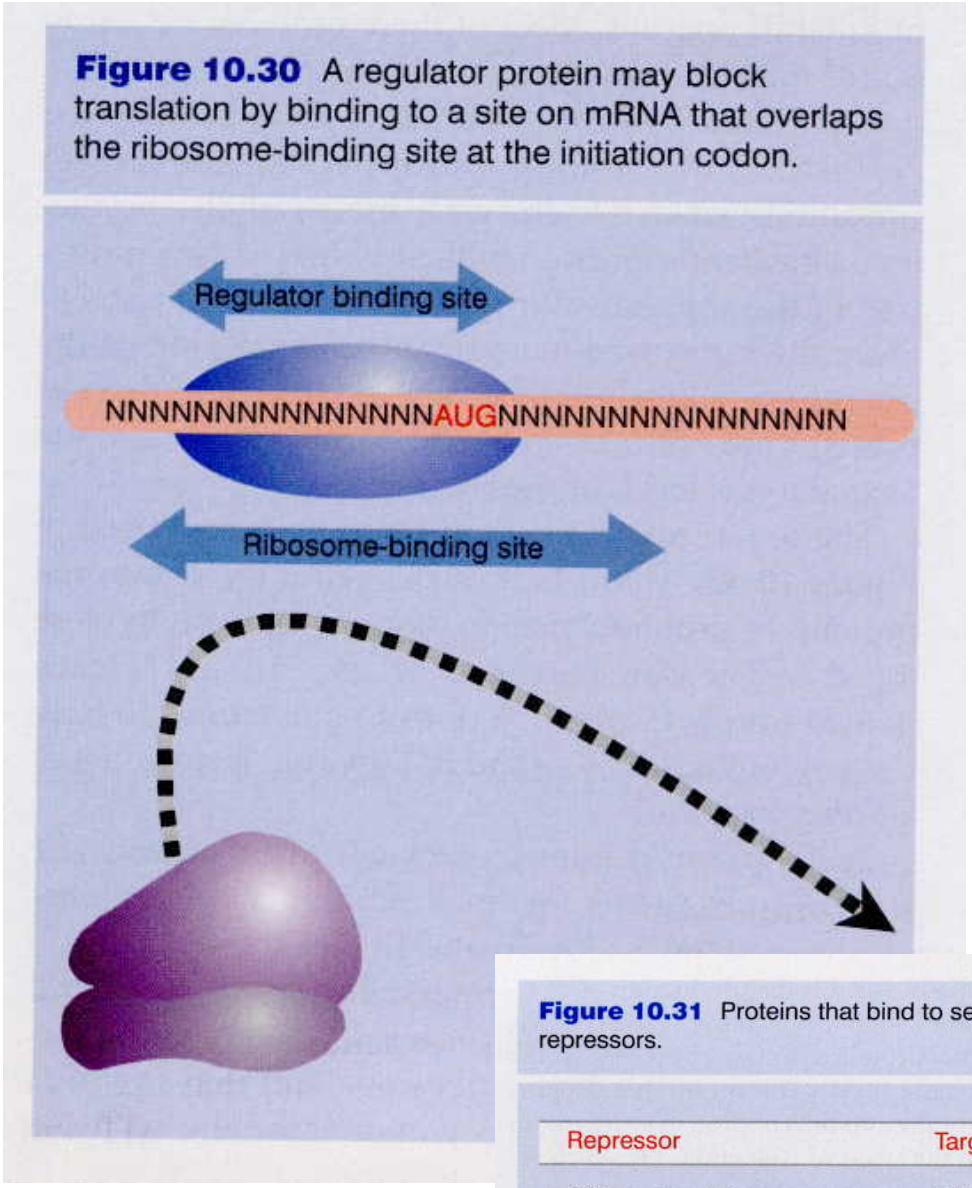


Figure 10.30 A regulator protein may block translation by binding to a site on mRNA that overlaps the ribosome-binding site at the initiation codon.

Figure 10.31 Proteins that bind to sequences within the initiation regions of mRNAs may function as translational repressors.

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

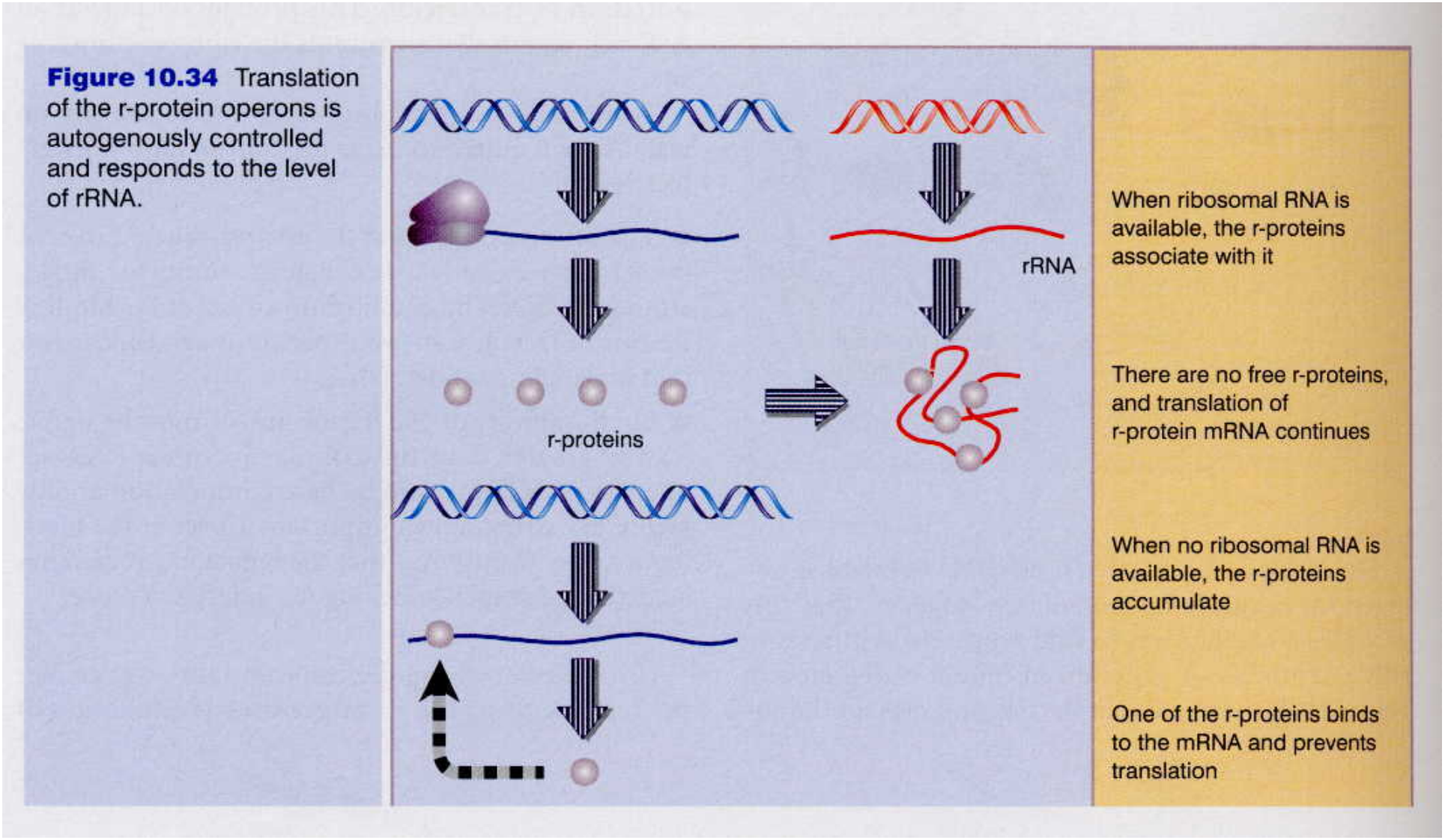


Figure 10.34 Translation of the r-protein operons is autogenously controlled and responds to the level of rRNA.

Attenuation

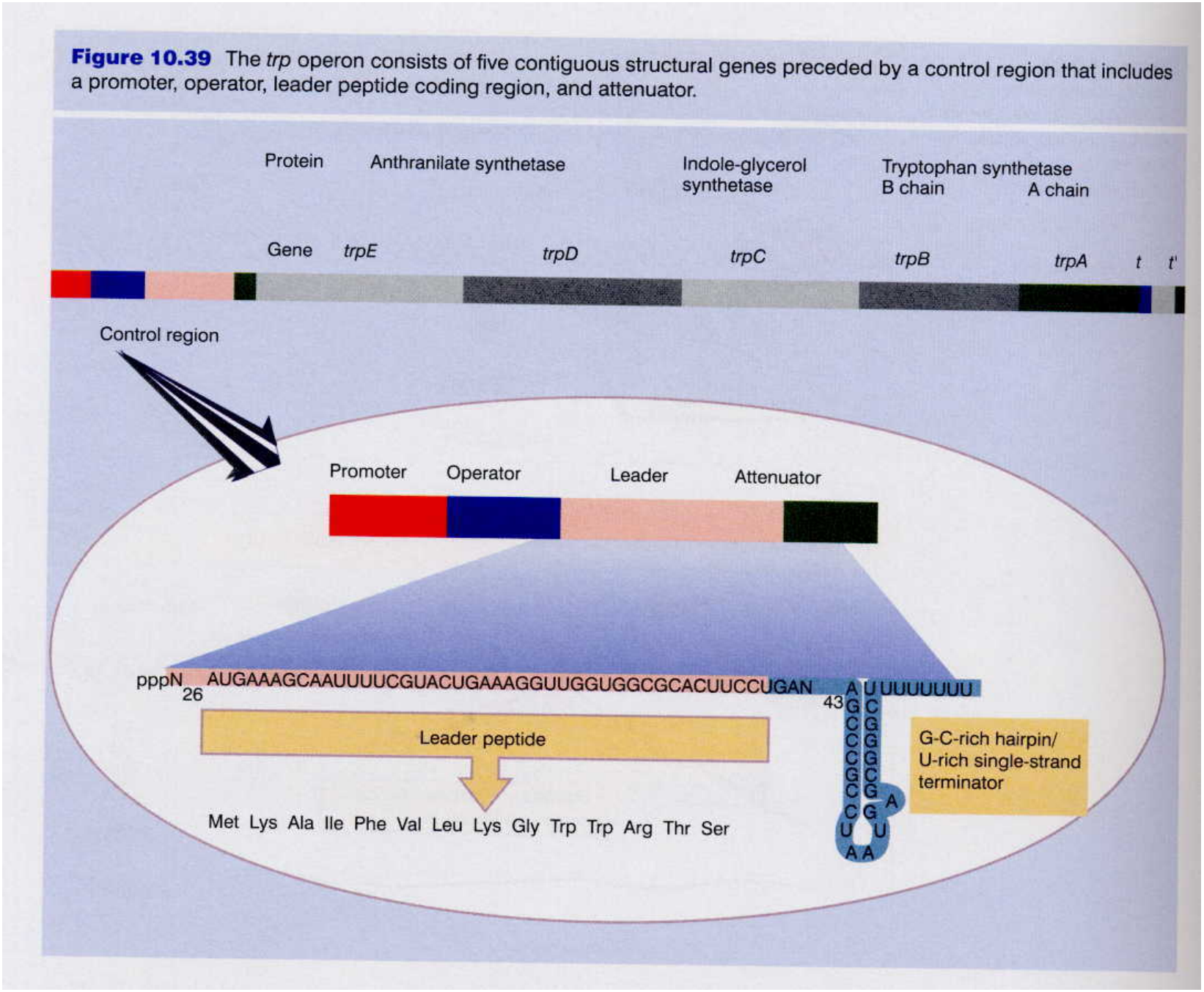
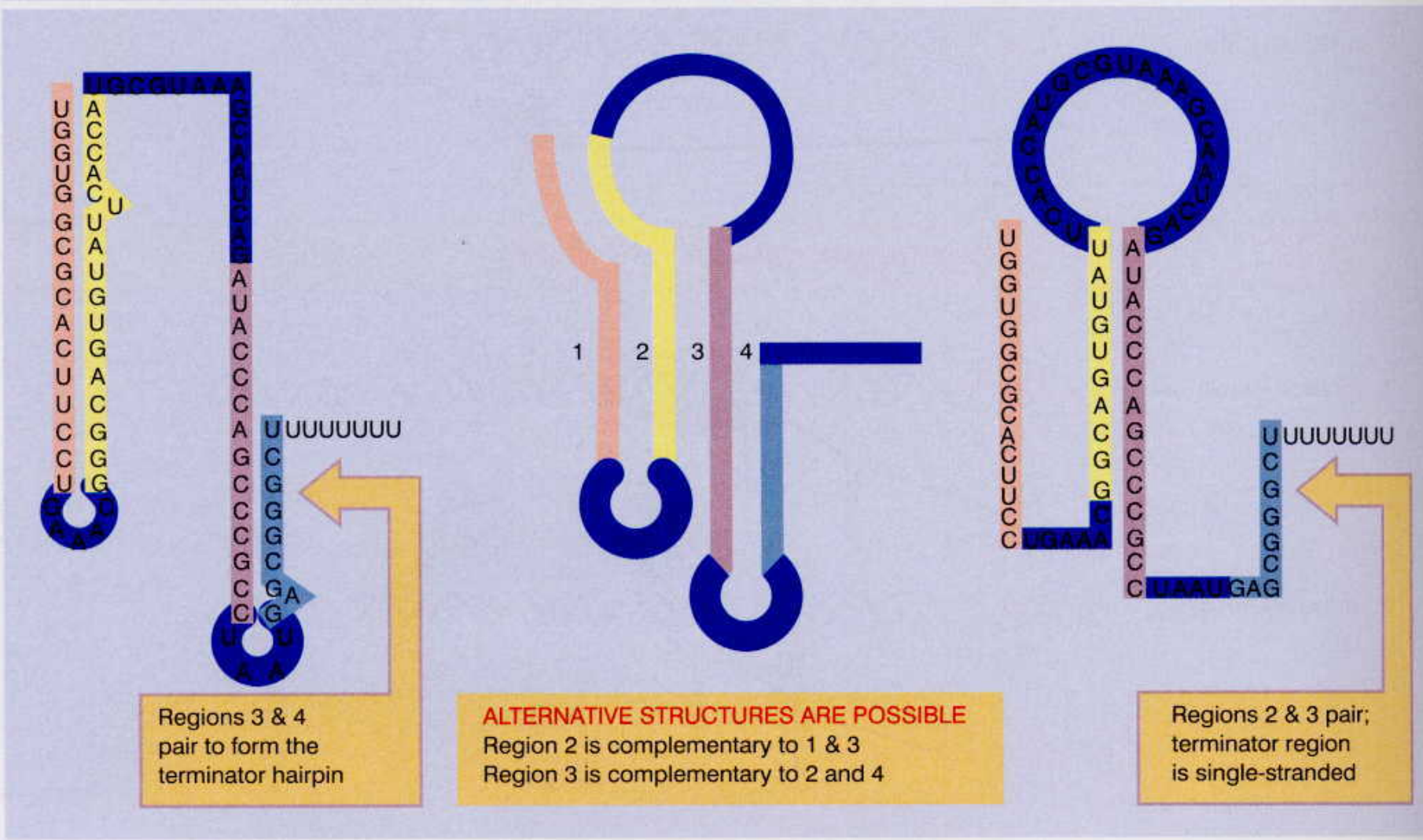


Figure 10.39 The *trp* operon consists of five contiguous structural genes preceded by a control region that includes a promoter, operator, leader peptide coding region, and attenuator.

Figure 10.41 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.



H.Schwab

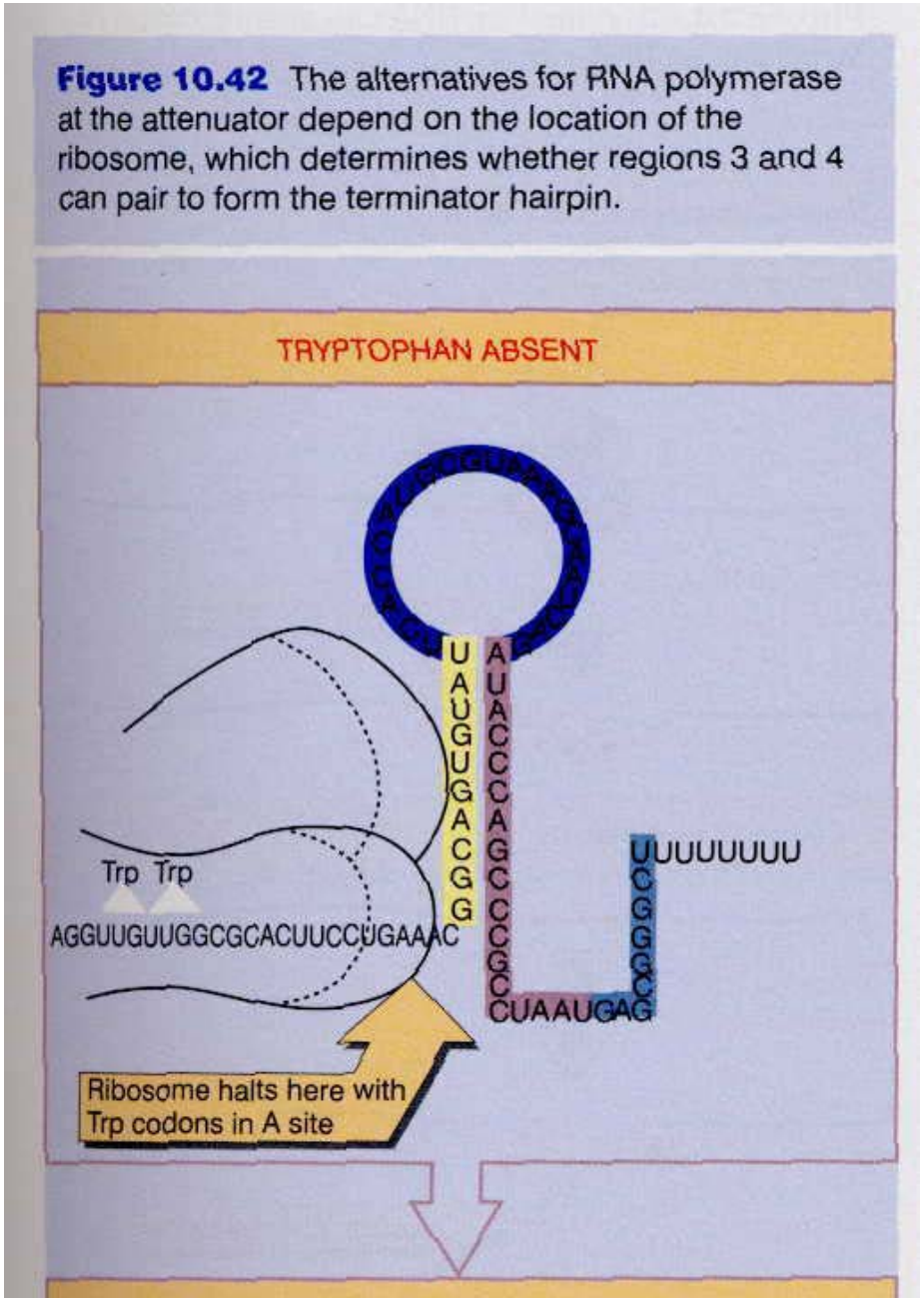
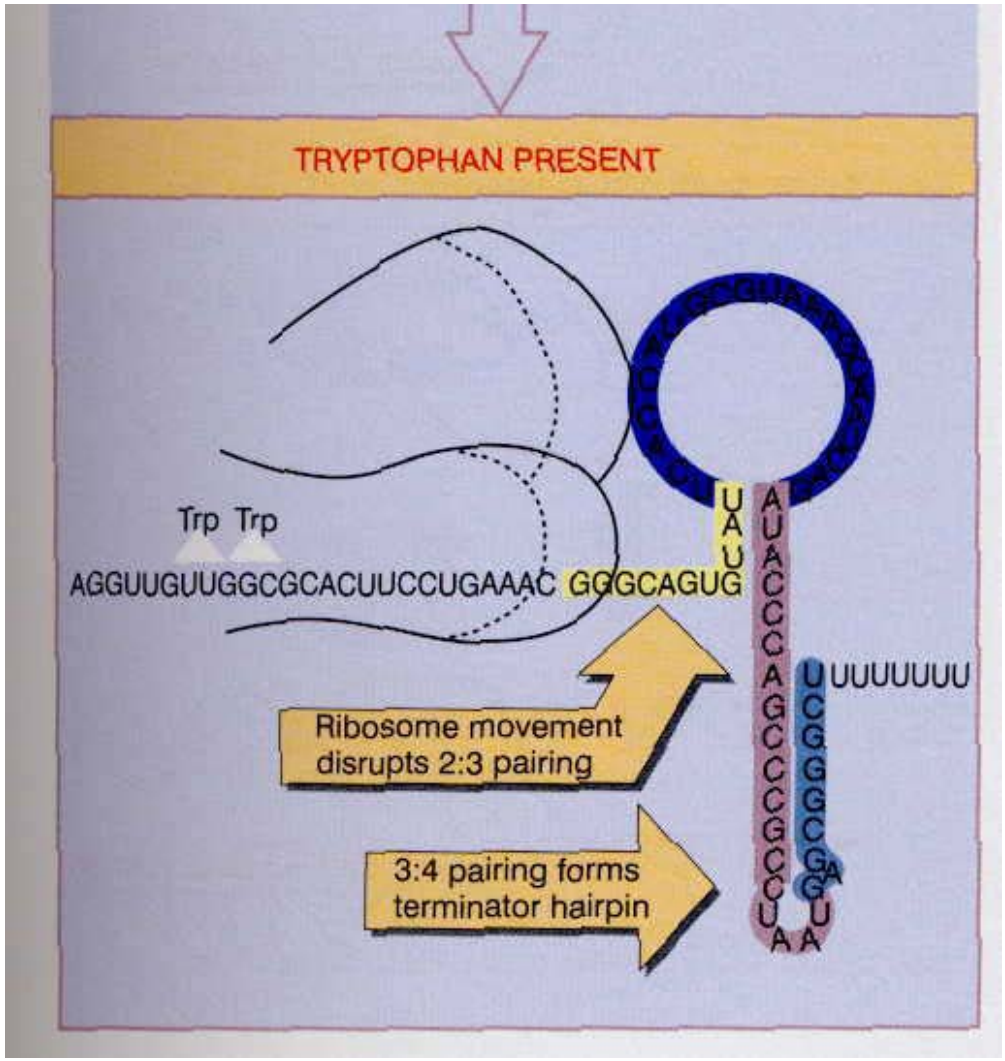


Figure 10.42 The alternatives for RNA polymerase at the attenuator depend on the location of the ribosome, which determines whether regions 3 and 4 can pair to form the terminator hairpin.



Antisense RNA

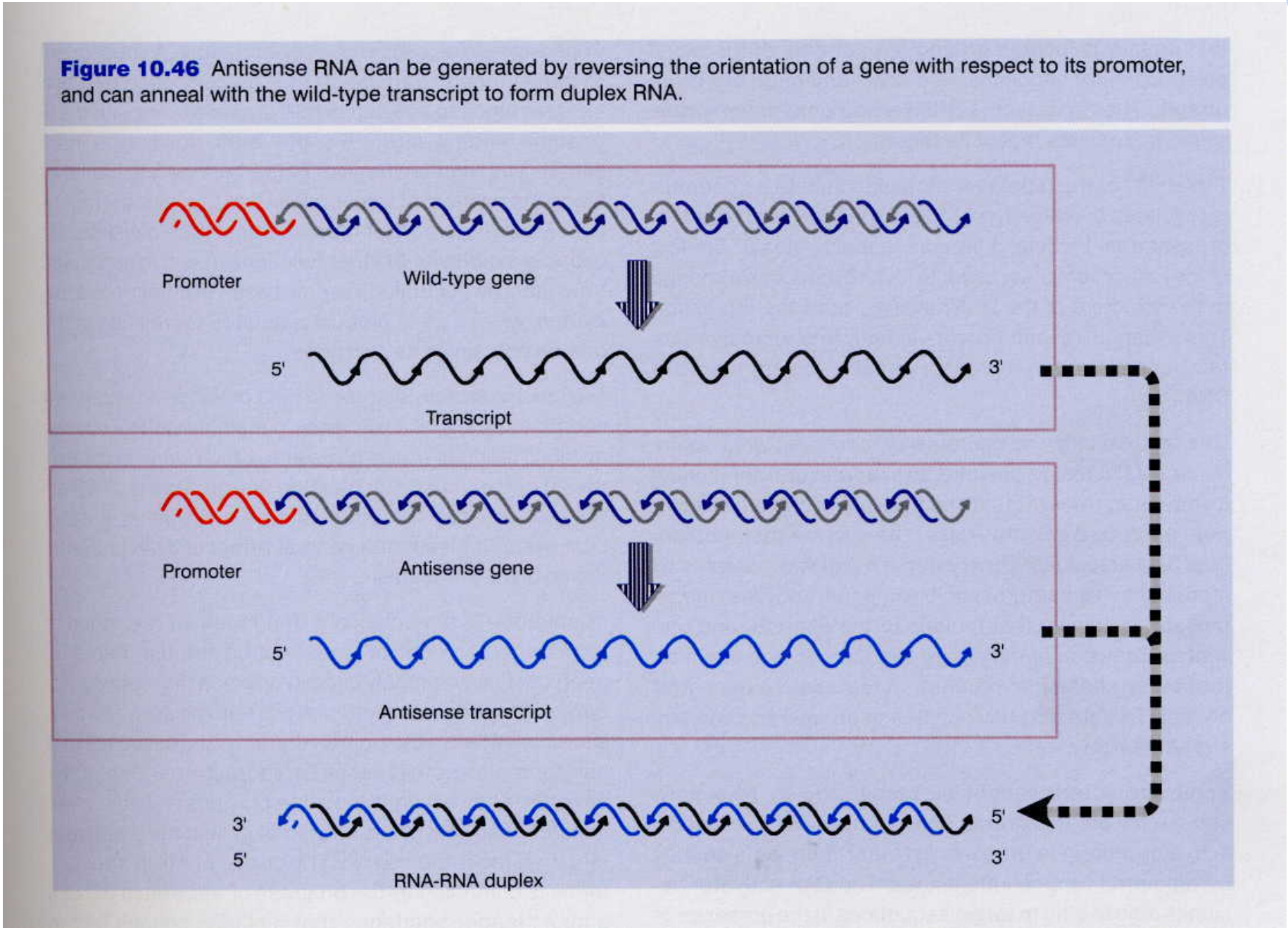


Figure 10.46 Antisense RNA can be generated by reversing the orientation of a gene with respect to its promoter, and can anneal with the wild-type transcript to form duplex RNA.

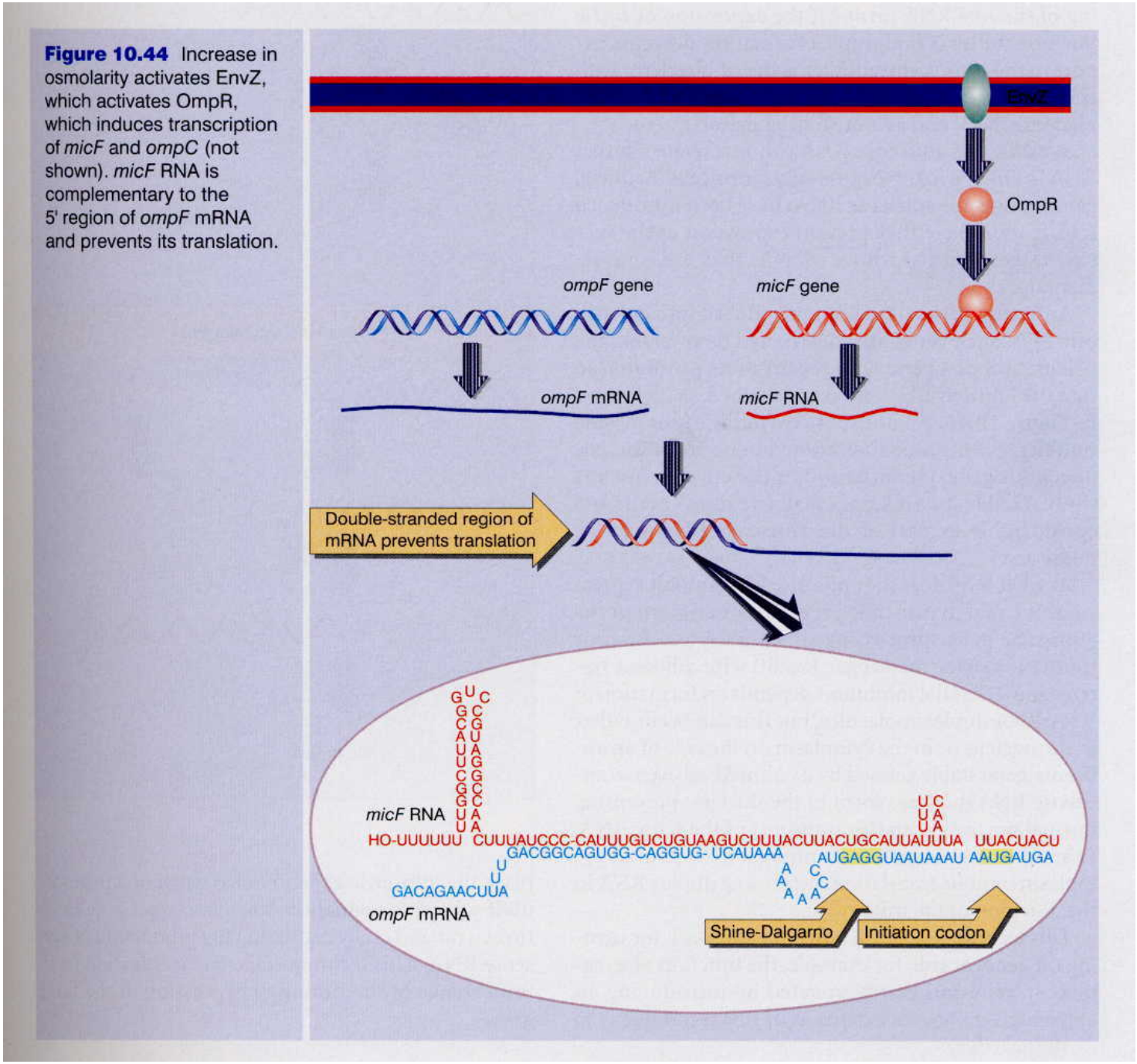
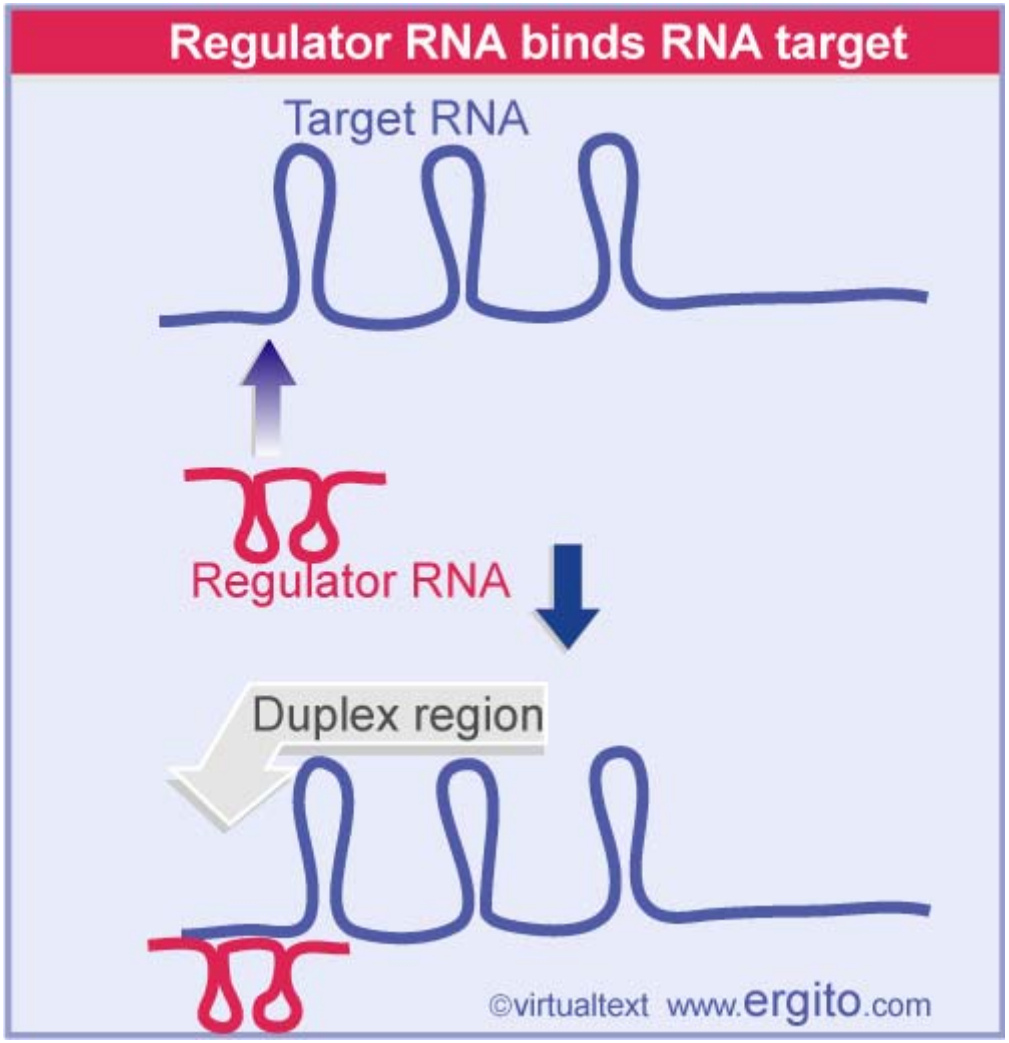
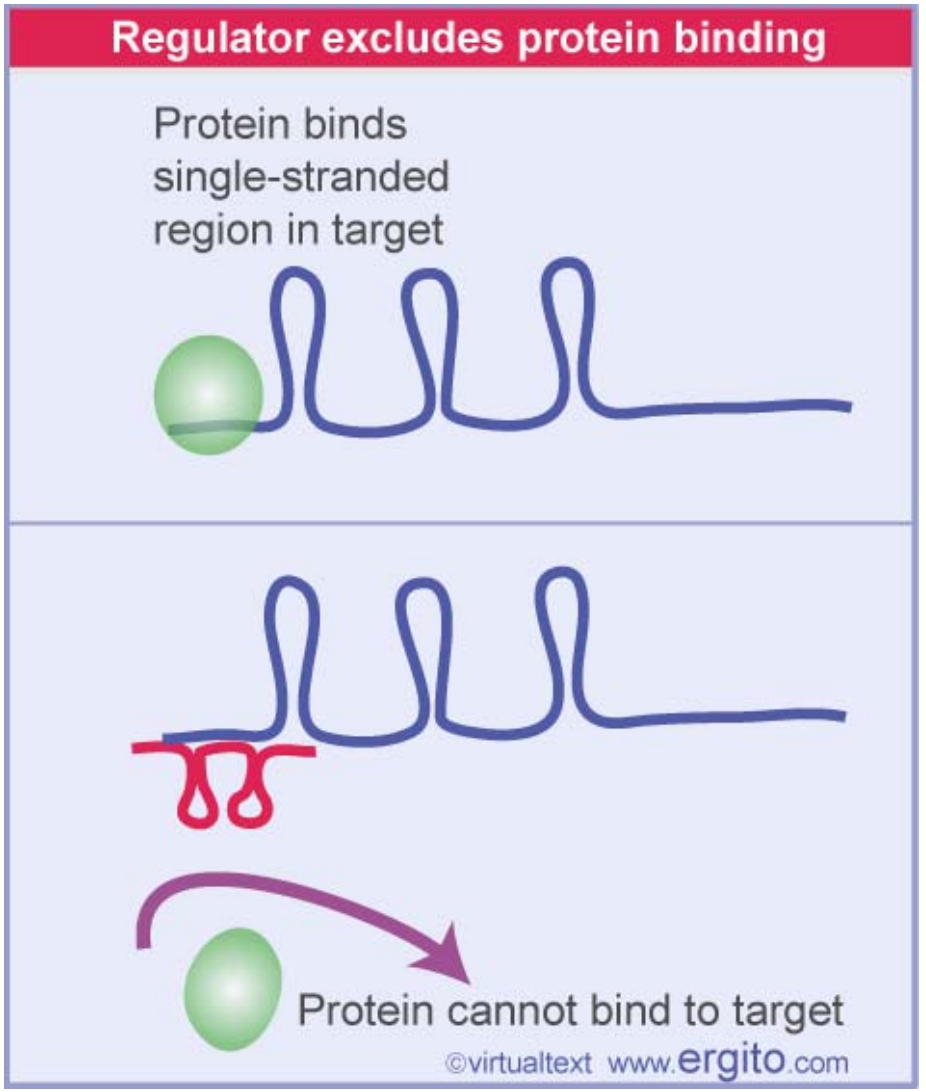
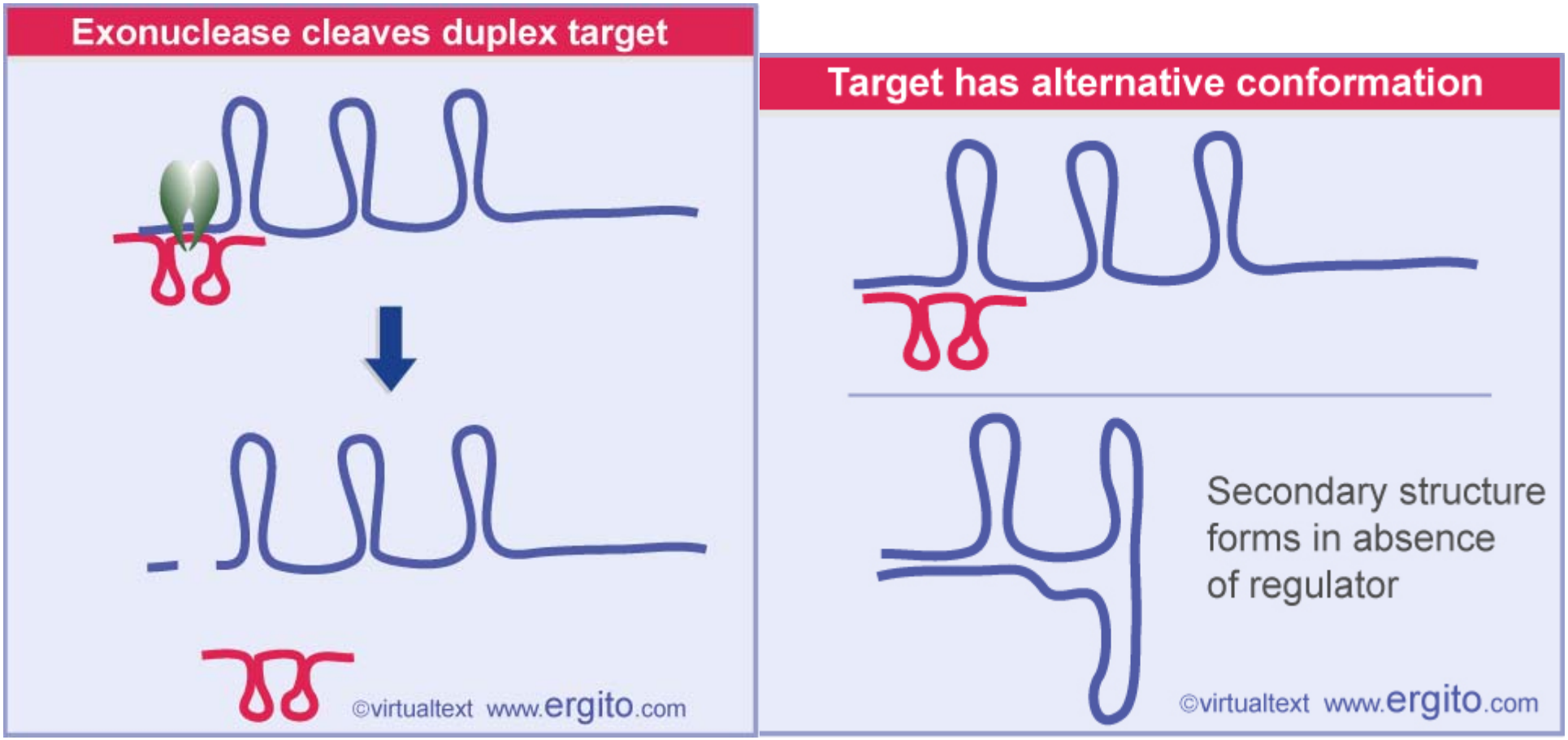
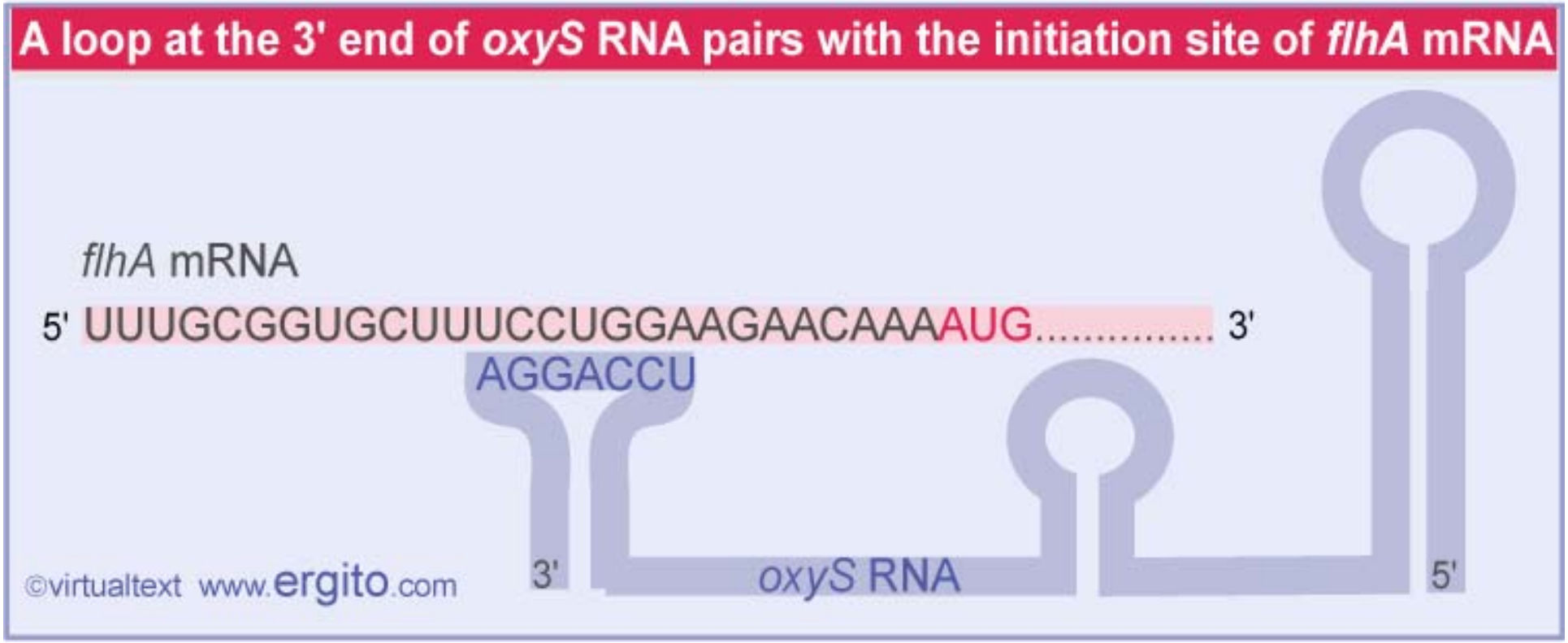


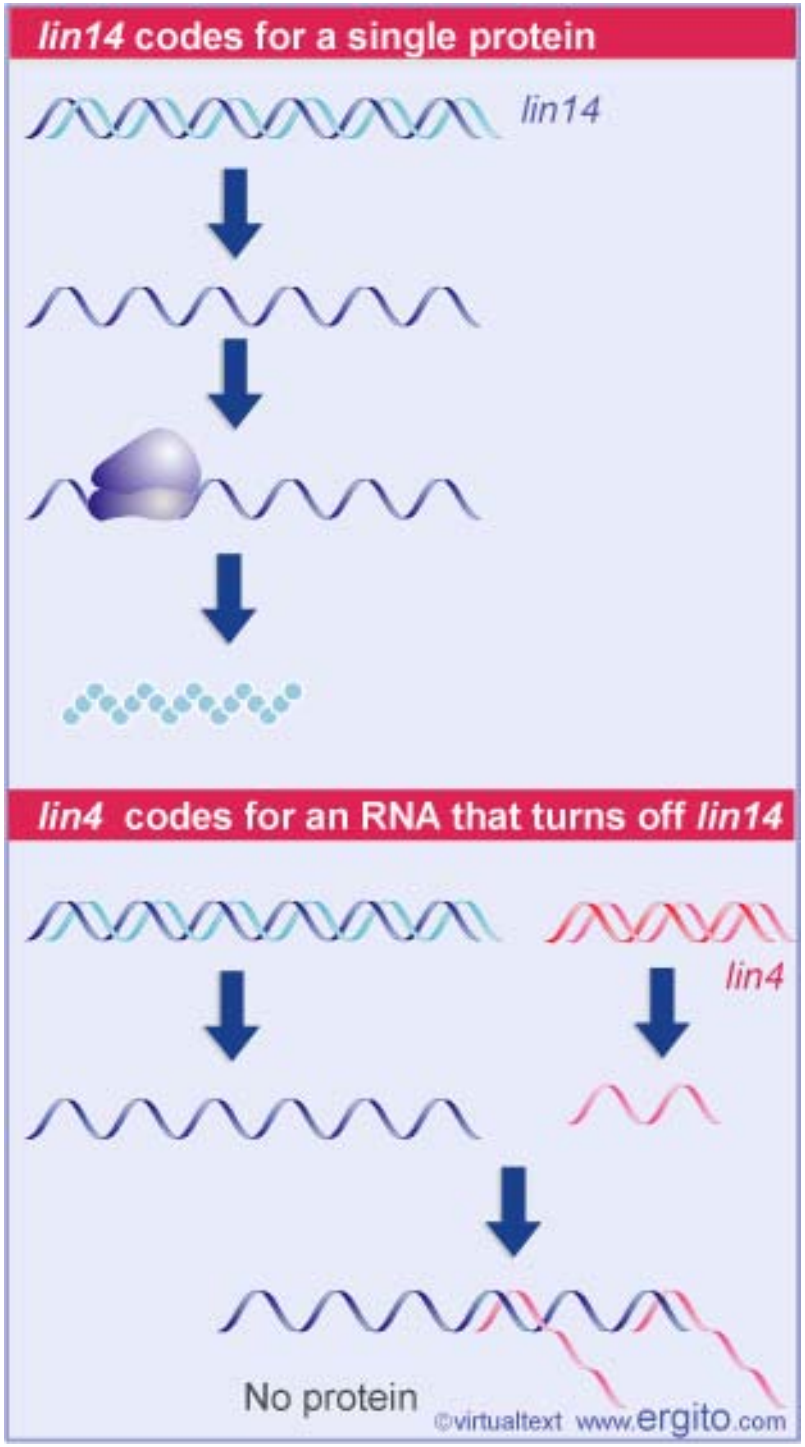
Figure 10.44 Increase in osmolarity activates EnvZ, which activates OmpR, which induces transcription of *micF* and *ompC* (not shown). *micF* RNA is complementary to the 5' region of *ompF* mRNA and prevents its translation.


Regulator RNA binds RNA target
Target RNA
Regulator RNA
Duplex region
©virtualtext www.ergito.com


Regulator excludes protein binding
Protein binds
single-stranded
region in target
Protein cannot bind to target
©virtualtext www.ergito.com


Exonuclease cleaves duplex target
©virtualtext www.ergito.com
Target has alternative conformation
Secondary structure forms in absence of regulator
©virtualtext www.ergito.com


A loop at the 3' end of oxyS RNA pairs with the initiation site of flhA mRNA
flhA mRNA
5' UUUGCGGUGCUUUCCUGGAAGAACAAAAUG..... 3'
AGGACCU
3'
oxyS RNA
5'
©virtualtext www.ergito.com


lin14 codes for a single protein
lin14
lin4 codes for an RNA that turns off lin14
lin4
No protein
©virtualtext www.ergito.com

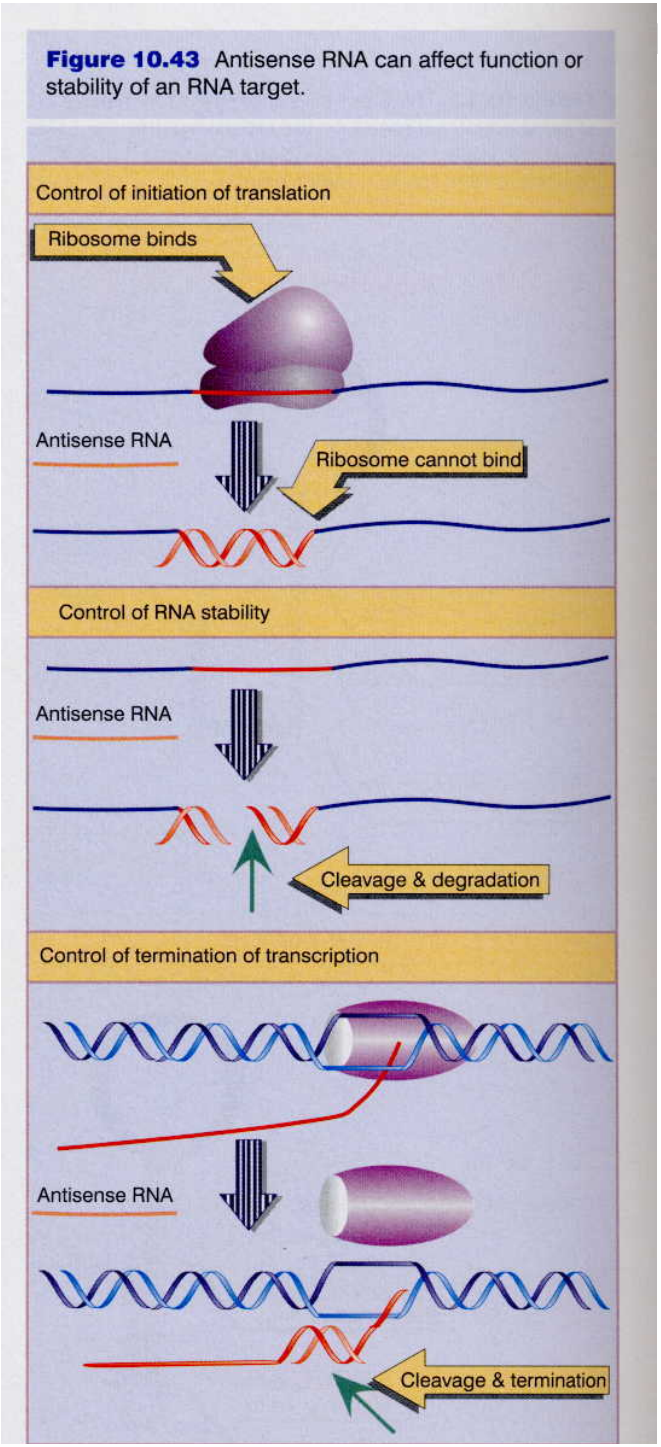
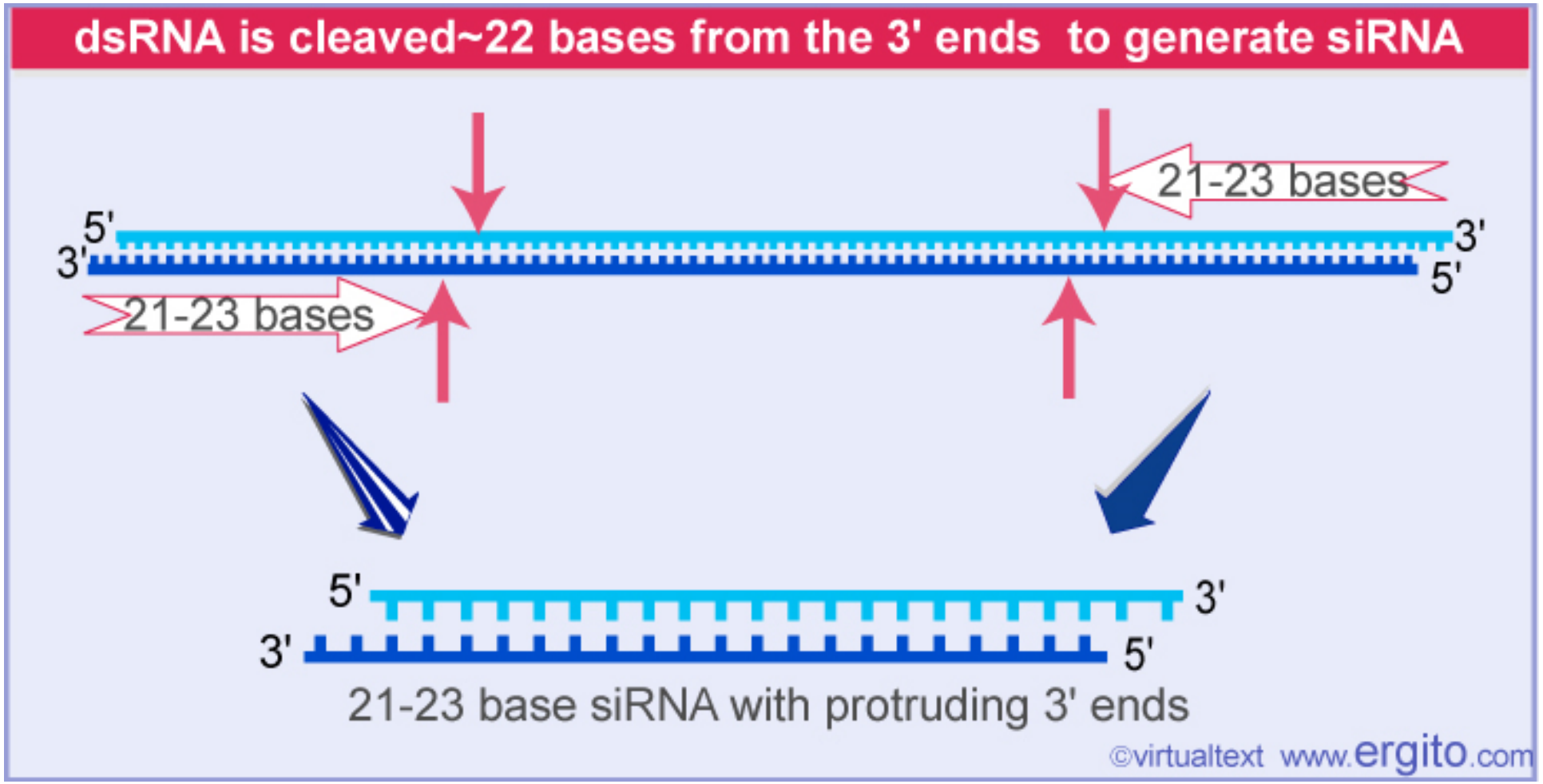
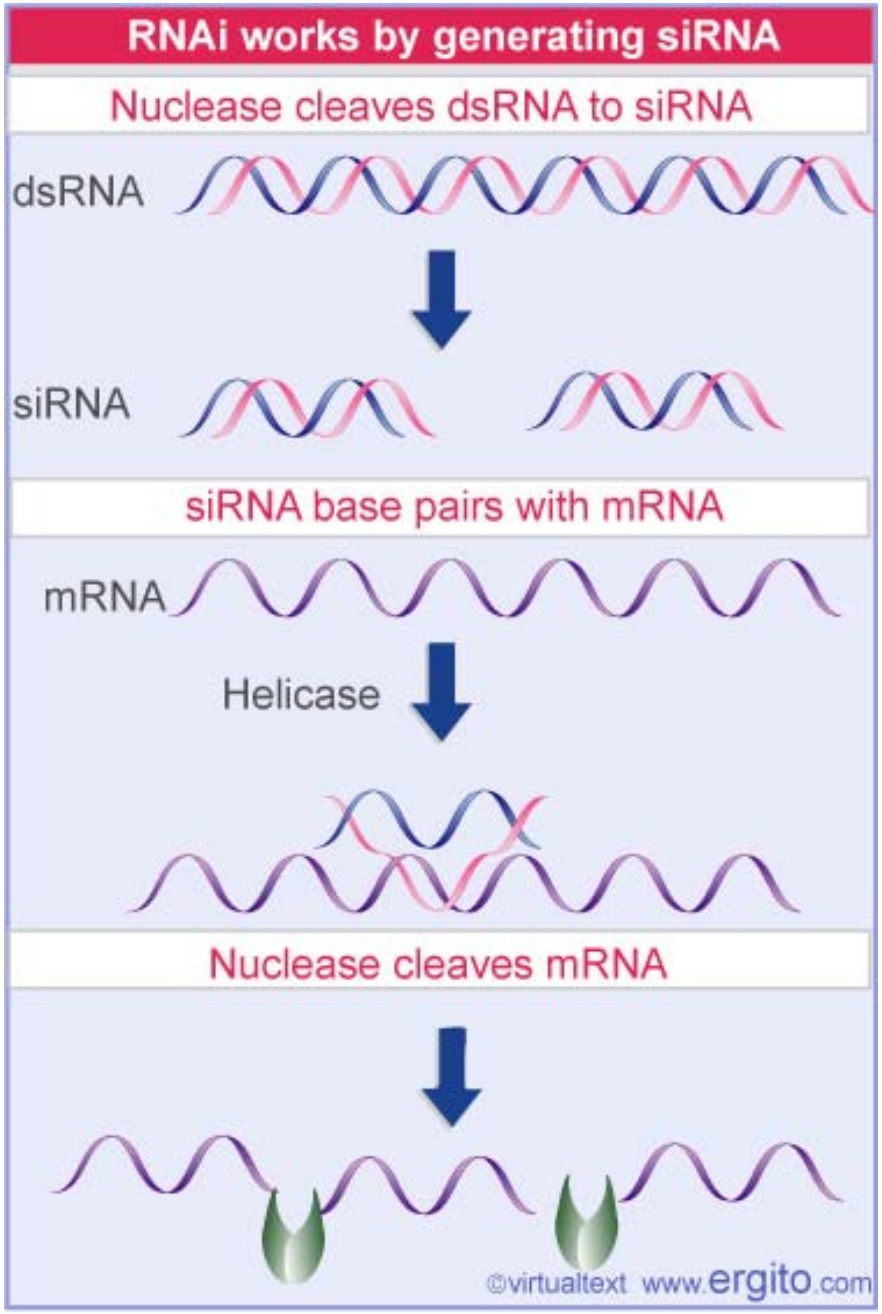
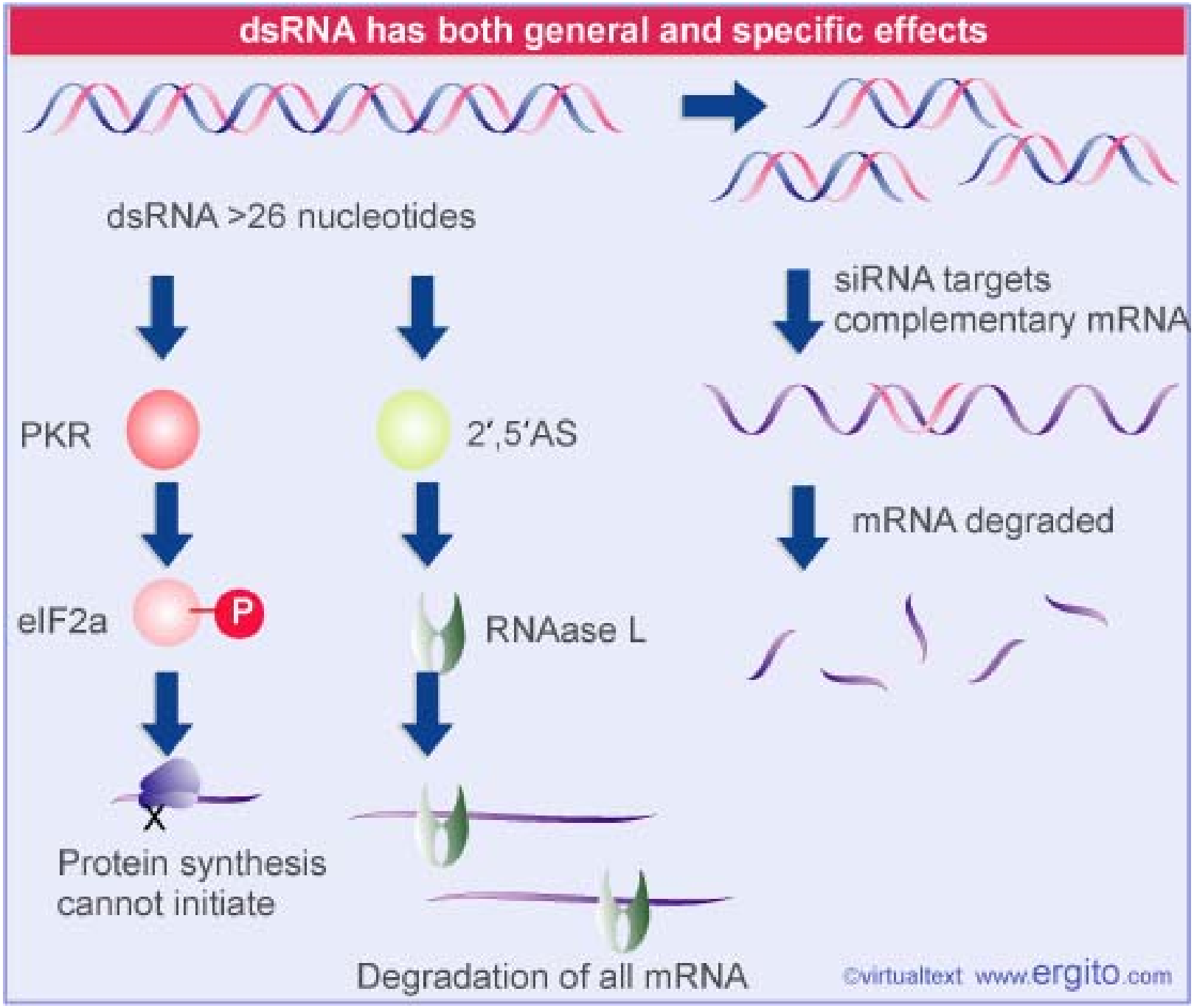


Figure 10.43 Antisense RNA can affect function or stability of an RNA target.
Control of initiation of translation
Ribosome binds
Antisense RNA
Ribosome cannot bind
Control of RNA stability
Antisense RNA
Cleavage & degradation
Control of termination of transcription
Antisense RNA
Cleavage & termination


dsRNA is cleaved~22 bases from the 3' ends to generate siRNA
21-23 bases
5'
3'
3'
5'
21-23 bases
5'
3'
3'
5'
21-23 base siRNA with protruding 3' ends
©virtualtext www.ergito.com


RNAi works by generating siRNA
Nuclease cleaves dsRNA to siRNA
dsRNA
siRNA
siRNA base pairs with mRNA
mRNA
Helicase
Nuclease cleaves mRNA
©virtualtext www.ergito.com


dsRNA has both general and specific effects
dsRNA >26 nucleotides
PKR
2',5'AS
siRNA targets complementary mRNA
mRNA degraded
eIF2a
P
RNAase L
X
Protein synthesis cannot initiate
Degradation of all mRNA
©virtualtext www.ergito.com

25.11.14