

L'information génétique et les flots d'information dans la cellule

Table des matières

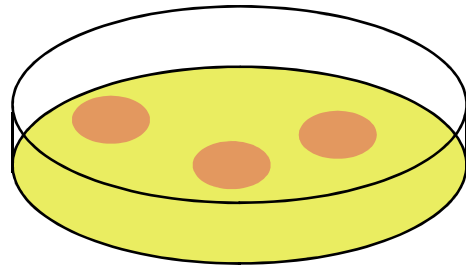
- Expression génétique : du promoteur a la protéine
 - La molécule d'ADN et l'information génétique
 - Organisation generale de l'expression génétique
 - Transcription
 - Epissage et modifications de l'ARNm
 - Traduction
 - Promoteurs et concepts de regulation génétique

Le principe transformant

- 1928, Griffith

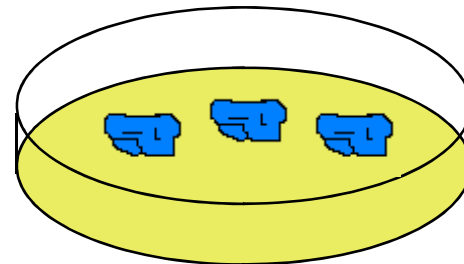
Streptococcus pneumoniae

Colonies lisses
virulentes



S ("smooth")

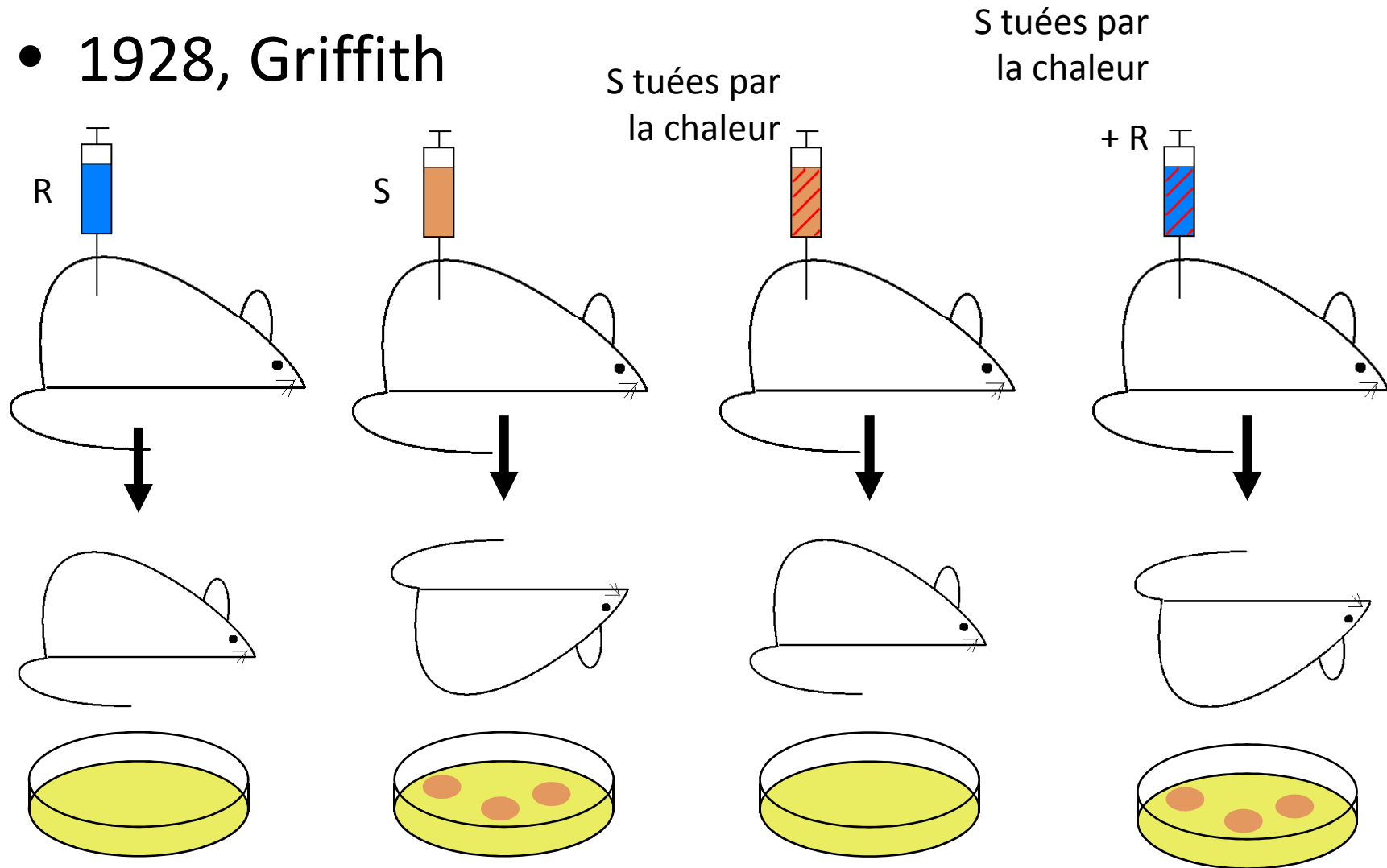
Colonies rugueuses
non virulentes



R ("rough")

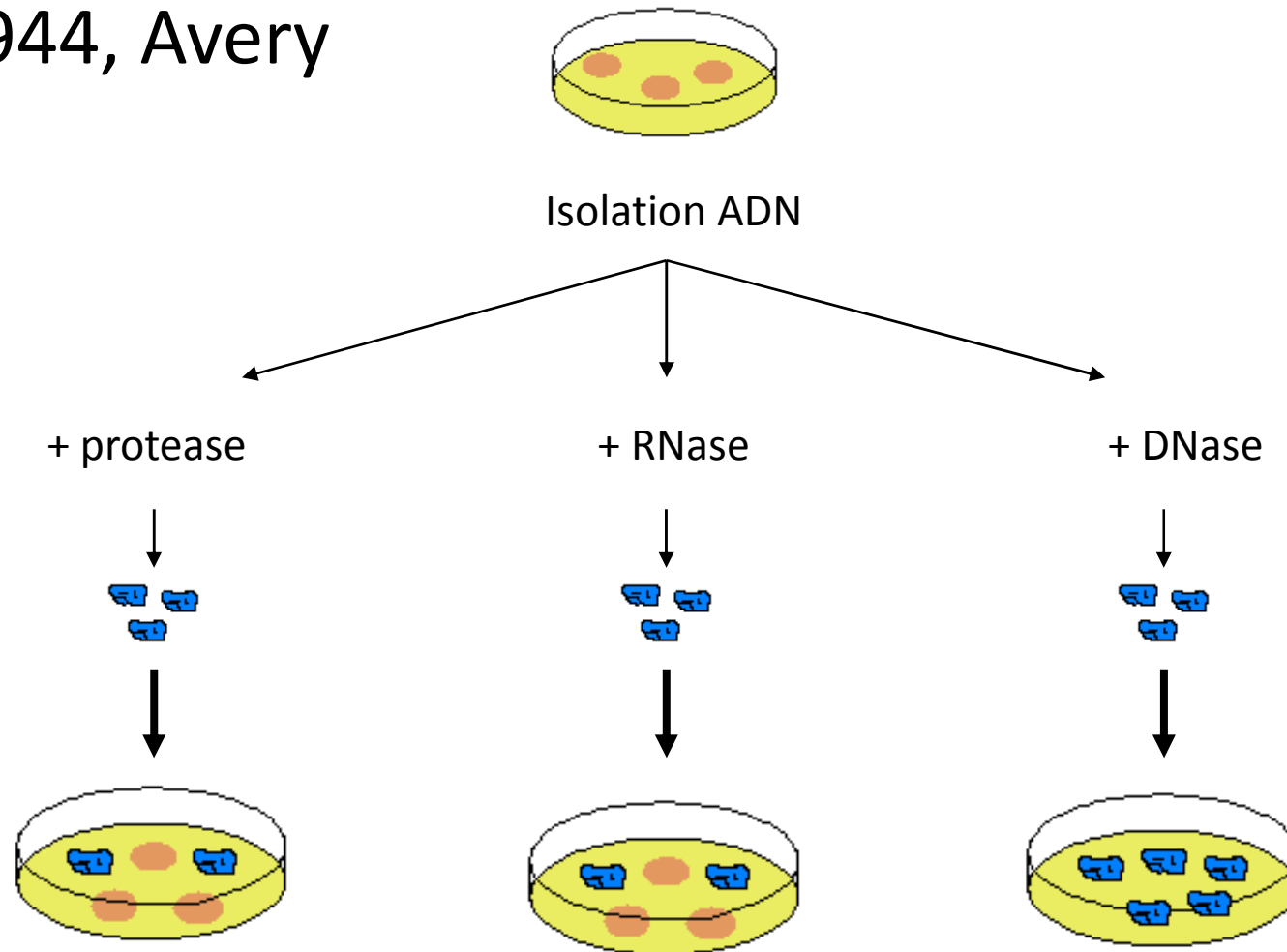
Le principe transformant

- 1928, Griffith

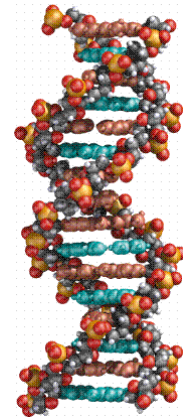


Le principe transformant

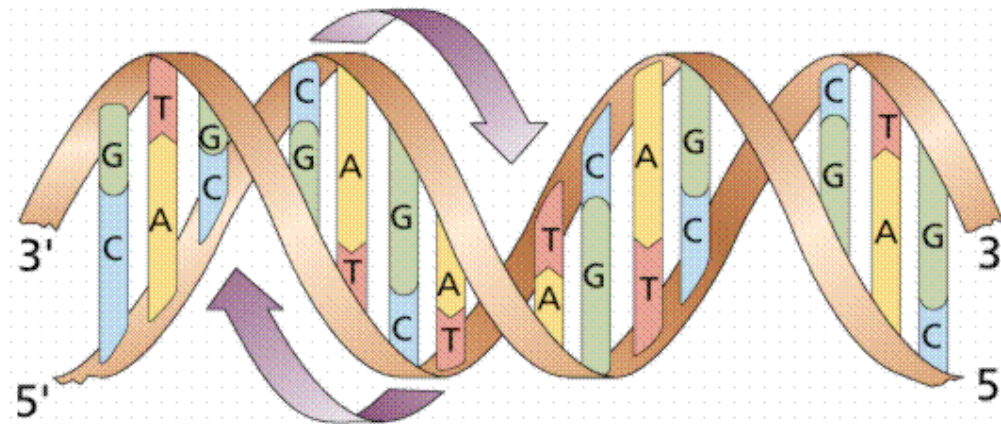
- 1944, Avery



Structure spatiale de l'ADN

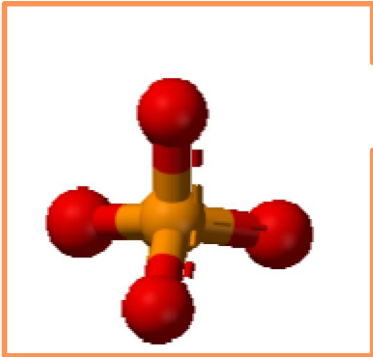


- L'ADN est un polymère de nucléotides
 - 1953, Watson et Crick
- modèle de la double hélice

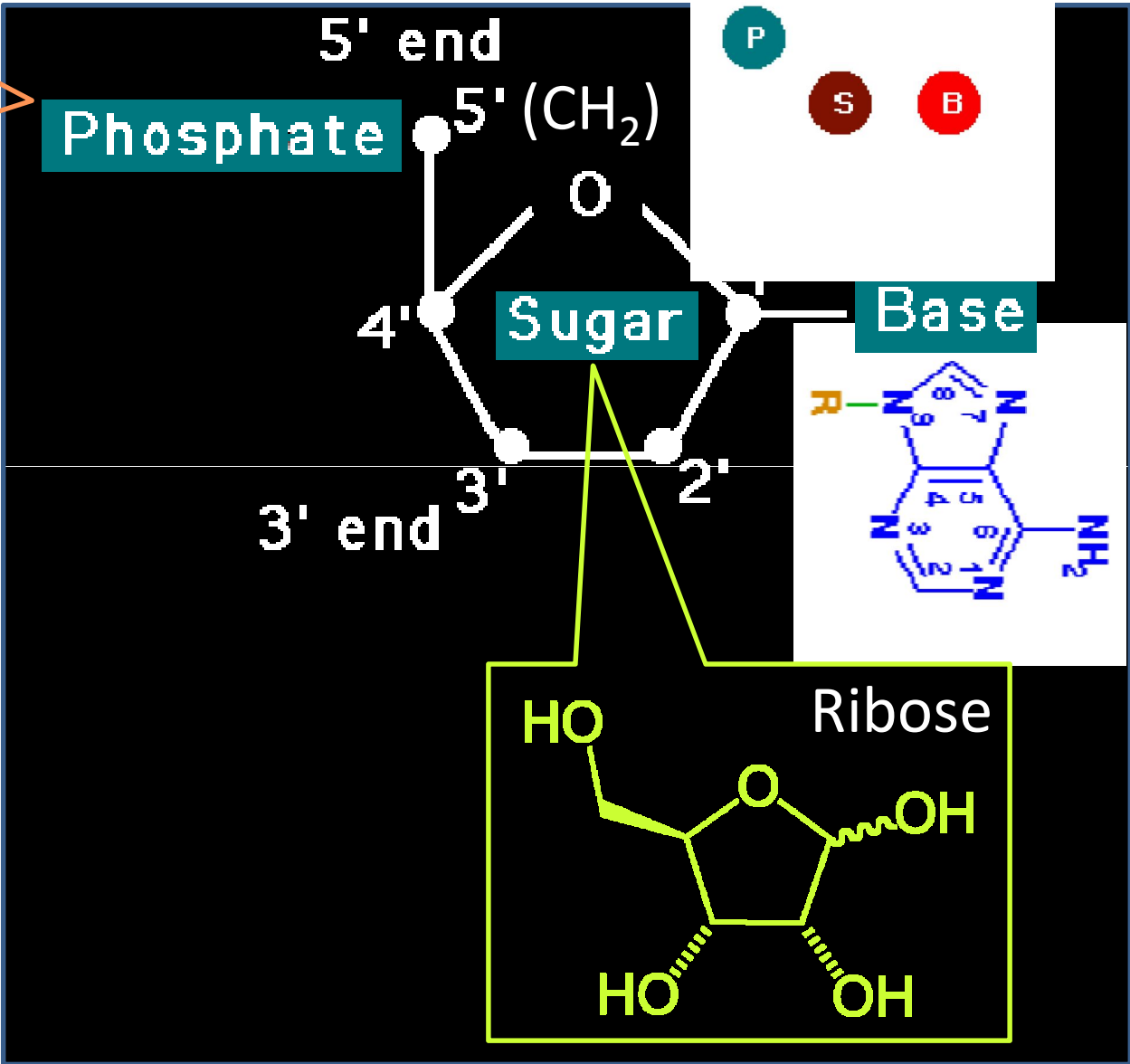


Les deux brins sont **antiparallèles** et **complémentaires**.

Nucléotides



- Phosphate
(=> acidité)
- Sucre
- Base



Acides Nucléiques

- Les bases de l'ADN et l'ARN

Adénine (A), Thymin (T), Guanine (G),
Cytosine (C), Uracil (U)

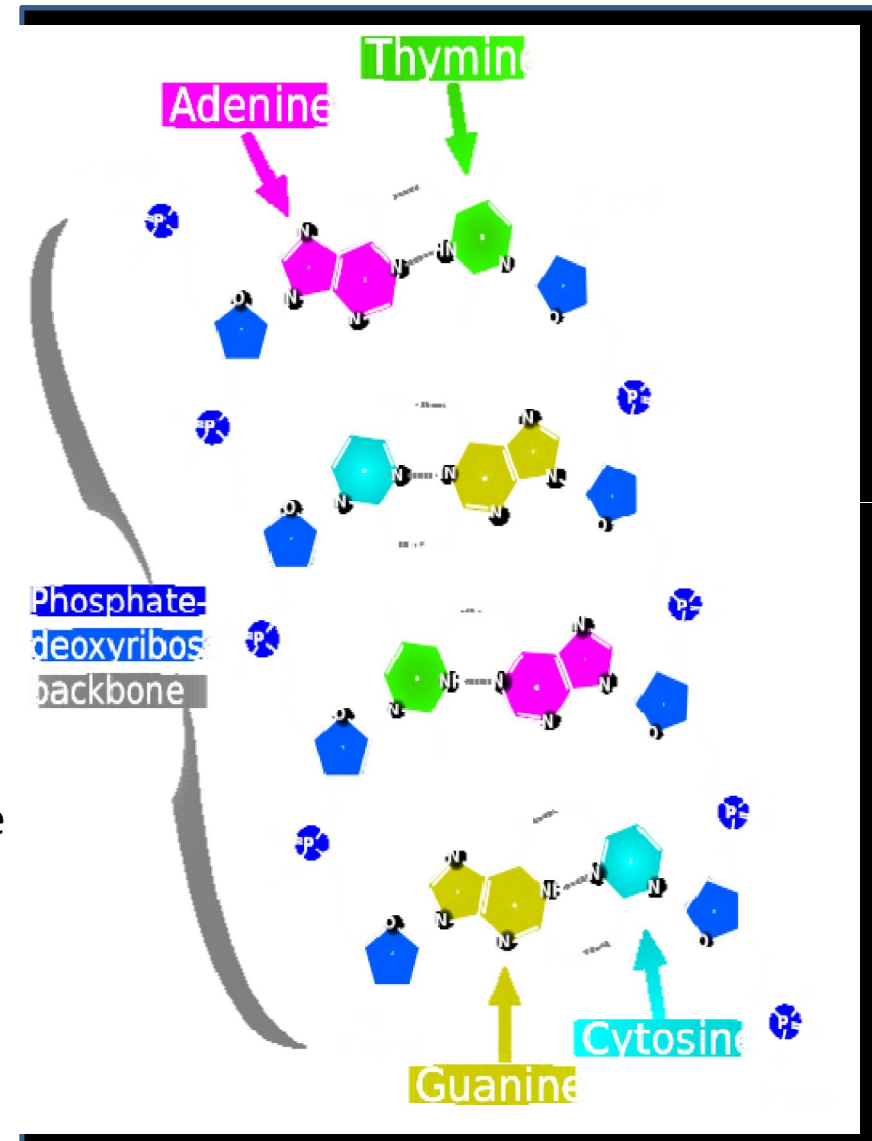
L'ADN a seulement A, T, G et C.

- Les bases créent des liaisons hydrogène entre elles

A \leftrightarrow T, A \leftrightarrow U, G \leftrightarrow C sont les liens stables : c'est l'appariement des bases

Conduit à la double hélice d'ADN et au repliement de l'ARN

Un brin d'ADN nu évolue thermodynamiquement vers un double brin



1949, la règle de Chargaff

Composition en bases (%) de l'ADN

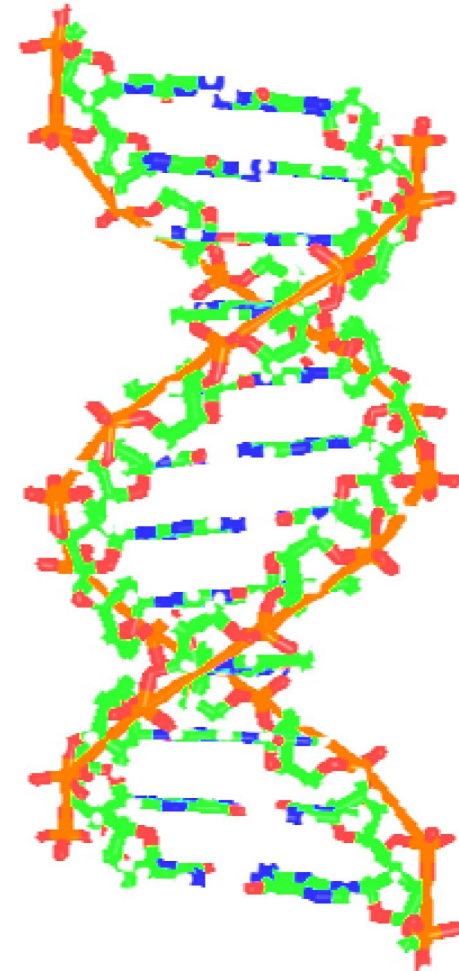
Organisme	A	T	G	C	$\frac{A+T}{G+C}$
Homme	31.0	31.5	19.1	18.4	1.67
Drosophile	27.3	27.6	22.5	22.5	1.22
Levure	31.3	32.9	18.7	17.1	1.79
Rat	28.6	28.4	21.4	21.5	1.32
E. coli	26.0	24.0	25.0	25.0	1.00
M. tuberculosis	15.1	14.6	34.9	35.4	0.42

$$A = T$$

$$G = C$$

ADN et ARN

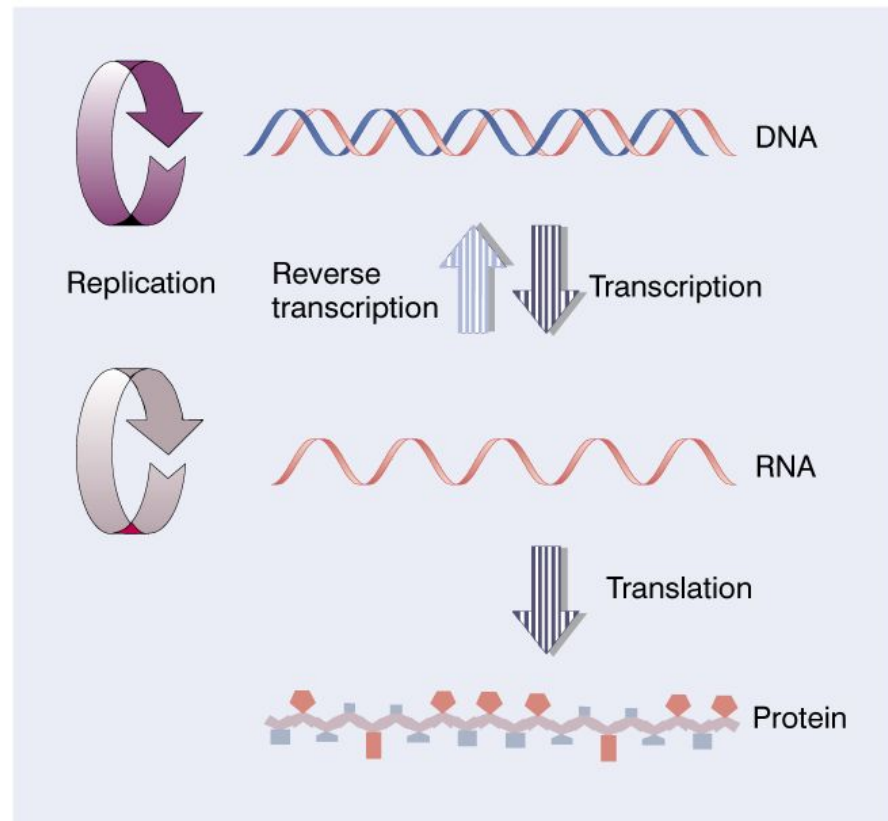
- L'ADN a une structure en double hélice et est plus stable:
 - Il peut former de longues chaînes
 - Il agit comme stockage à long terme de l'information
 - La complémentarité des brins permet la réplication
- L'ARN est plus court, en simple ou double brin, et plus réactif:
 - C'est un intermédiaire lors de l'expression génétique, synthétisé à partir de l'ADN



Dogme central de la biologie

- Flot unidirectionnel
- Simpliste
- Reste un bon modèle

Figure 1.34 The central dogma states that information in nucleic acid can be perpetuated or transferred, but the transfer of information into protein is irreversible.



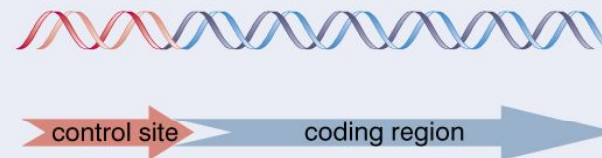
De l'ADN non-codant

On distingue:

- Les régions codantes de l'ADN, qui vont être transcrites en ARN
- Les régions non-codantes, qui servent de signaux de régulation

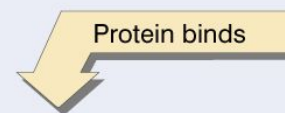
Figure 1.31 Control sites in DNA provide binding sites for proteins; coding regions are expressed via the synthesis of RNA.

DNA contains two types of sequences



Protein binding at control site is required for RNA synthesis

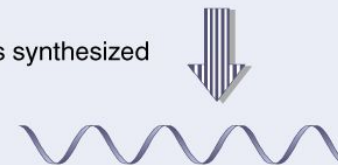
Protein binds



A yellow arrow points down to a grey protein molecule binding to the red control site of the DNA double helix.



RNA is synthesized

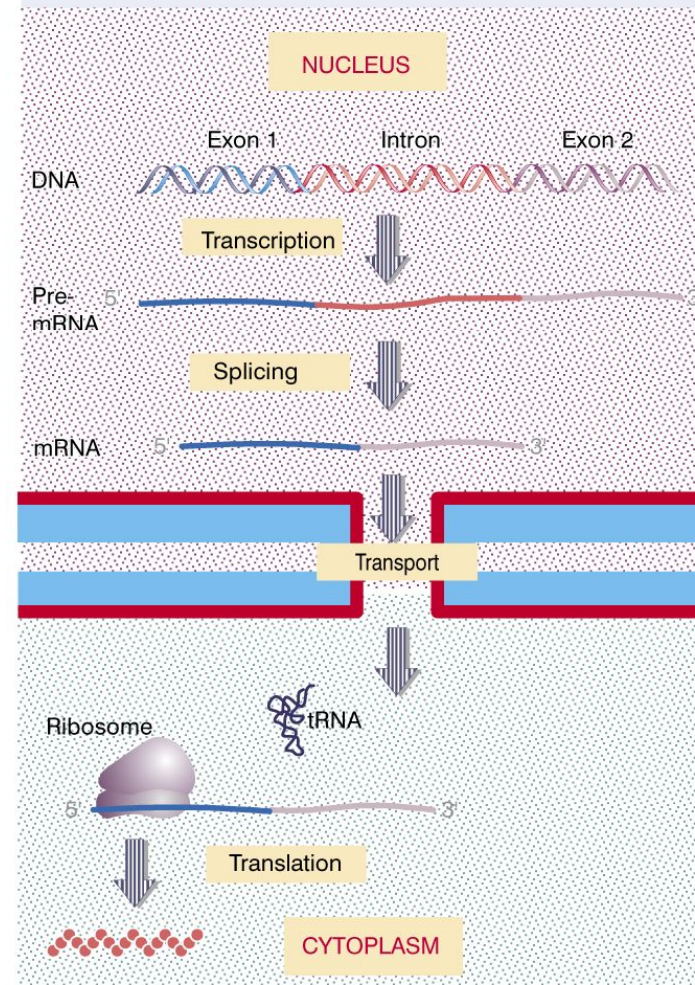


Un processus multi-étapes

Chez les cellules procaryotes on a seulement transcription et traduction.

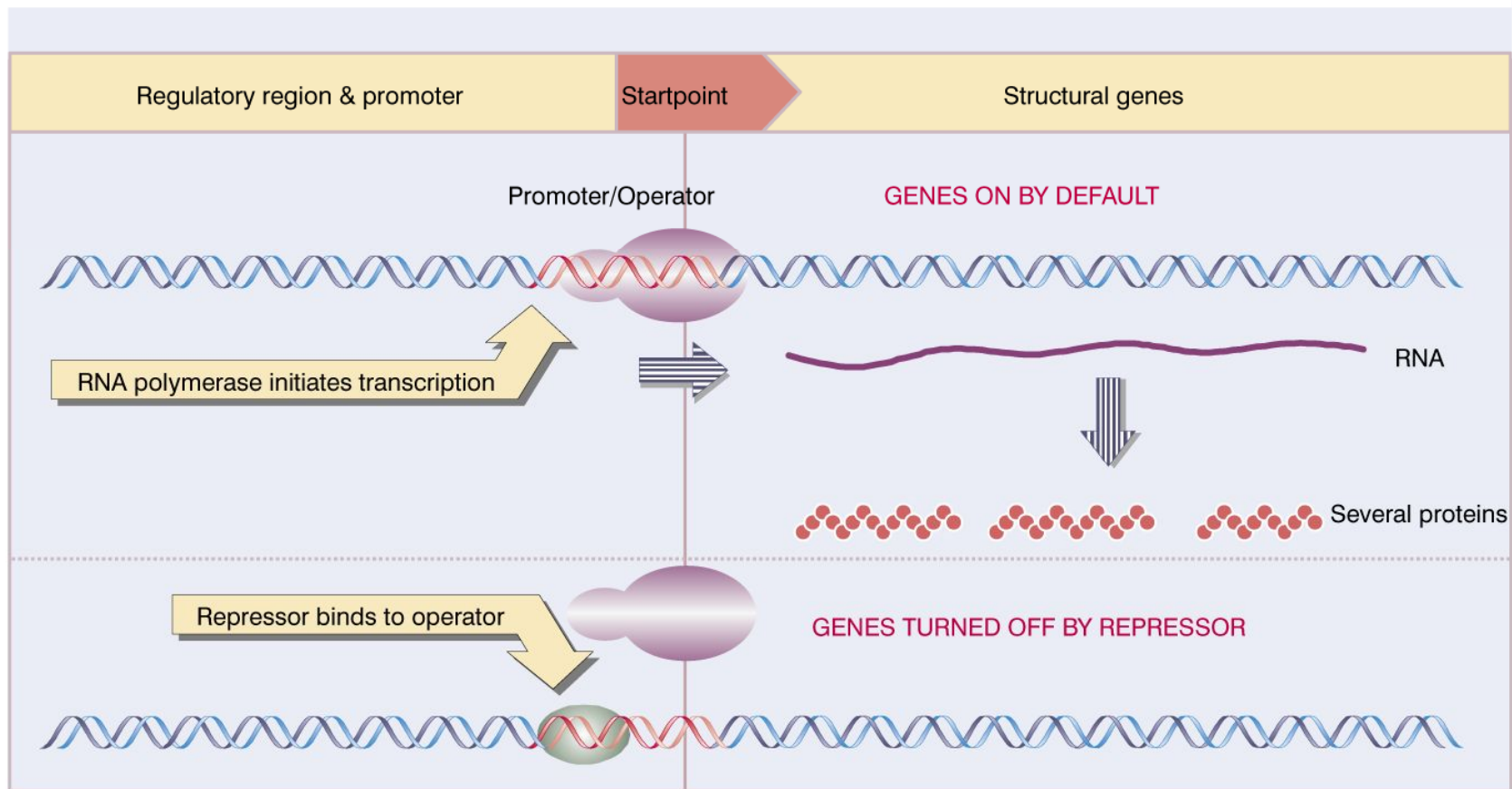
- Chez les eucaryotes il existe plusieurs étapes intermédiaires.

Figure 1.30 Gene expression is a multistage process.



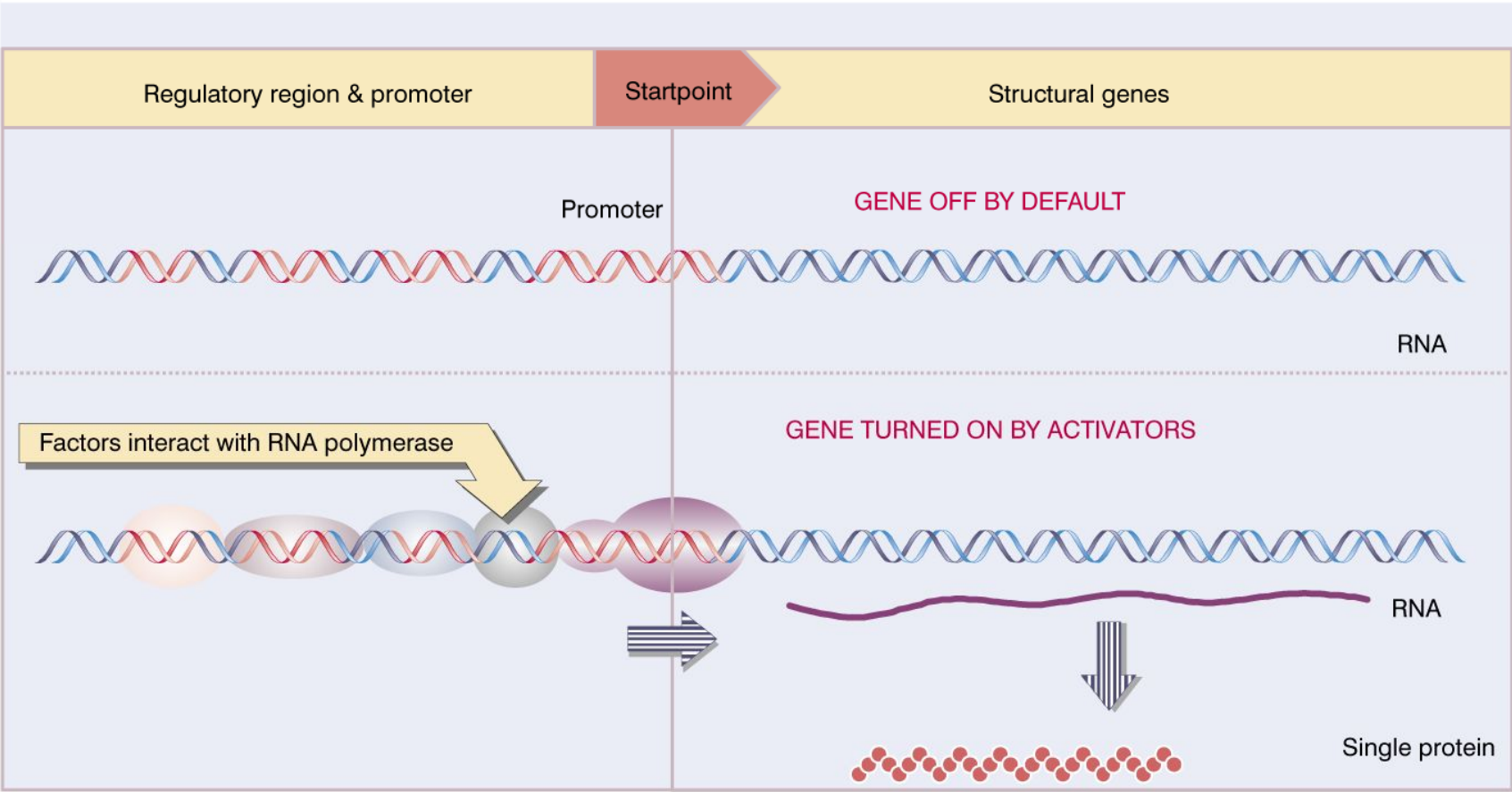
Contrôles négatifs

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.



Contrôles positifs

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.

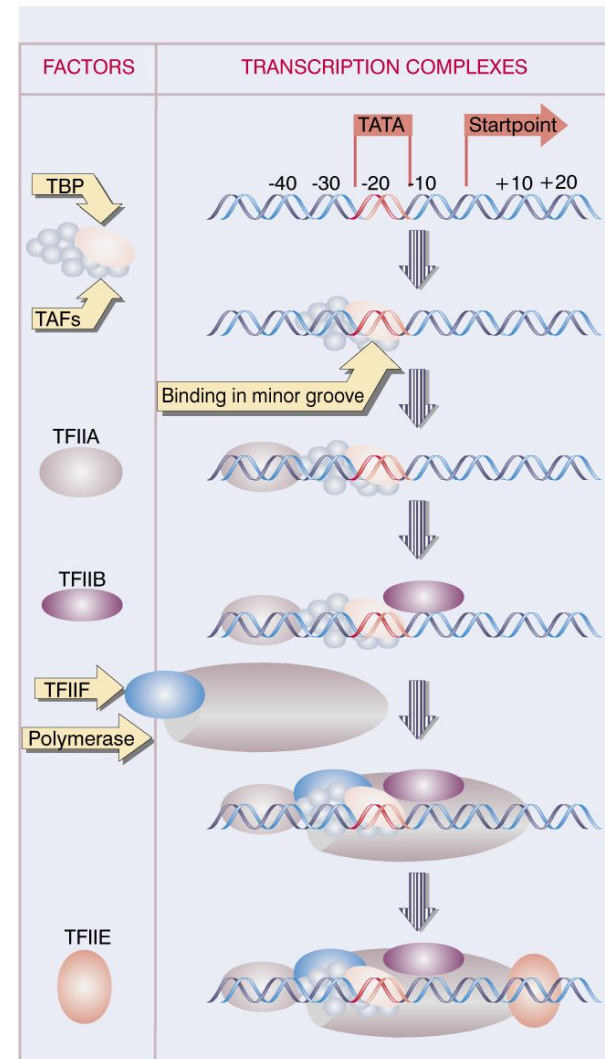


Initiation de la transcription

La transcription nécessite la formation d'un complexe pour arrimer l'ARN polymérase

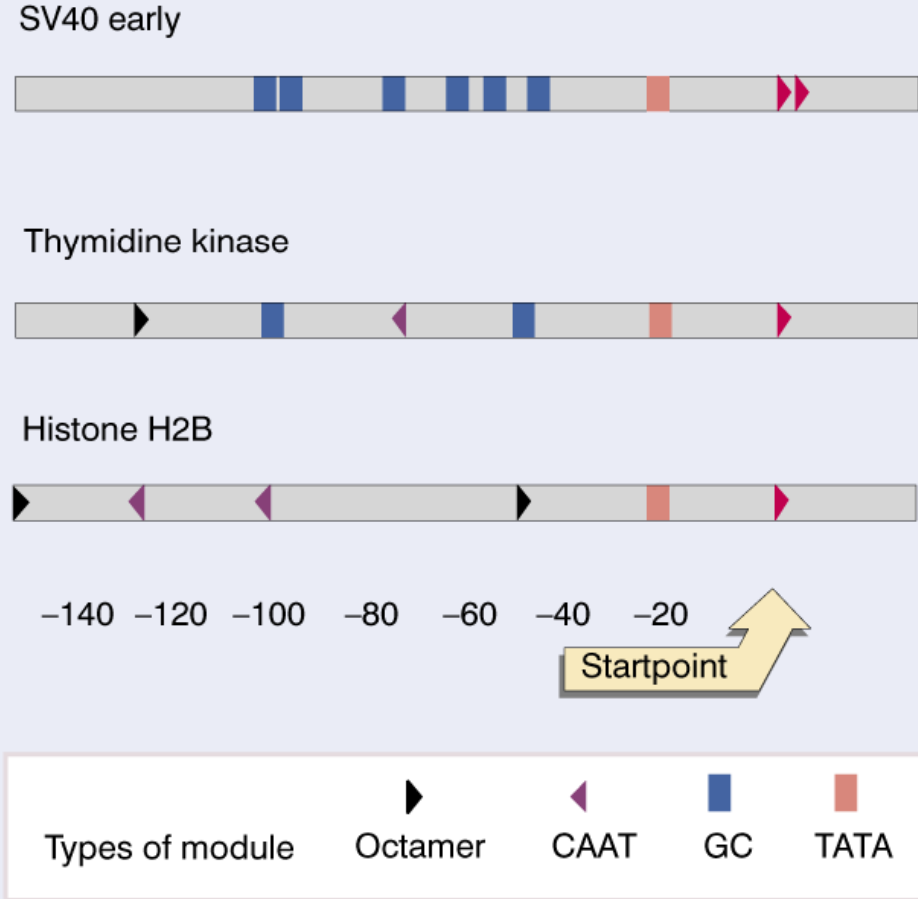
Ce complexe est formé d'un ensemble de protéines pour tous les gènes, ainsi que de facteurs de transcription spécifiques au gène transcrit.

Figure 20.11 An initiation complex assembles at promoters for RNA polymerase II by an ordered sequence of association with transcription factors.



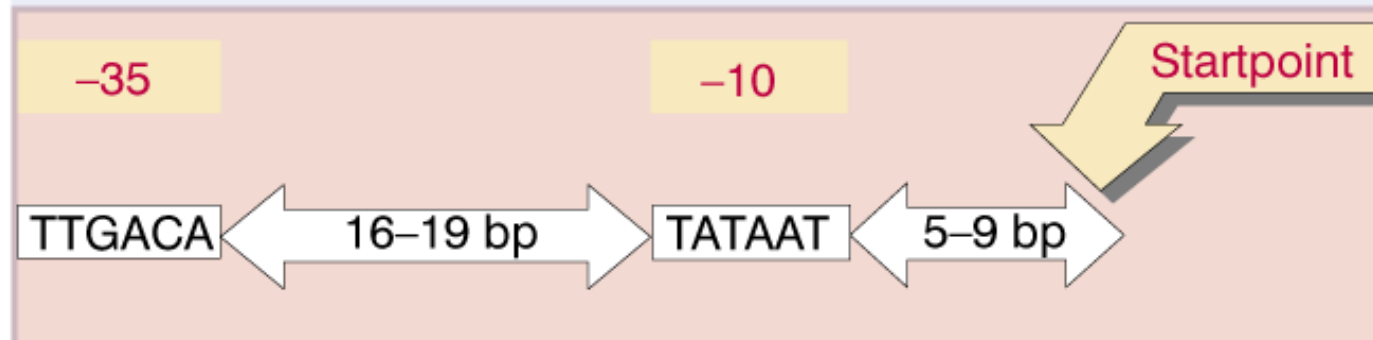
Promoteurs

Figure 20.17 Promoters contain different combinations of TATA boxes, CAAT boxes, GC boxes, and other elements.



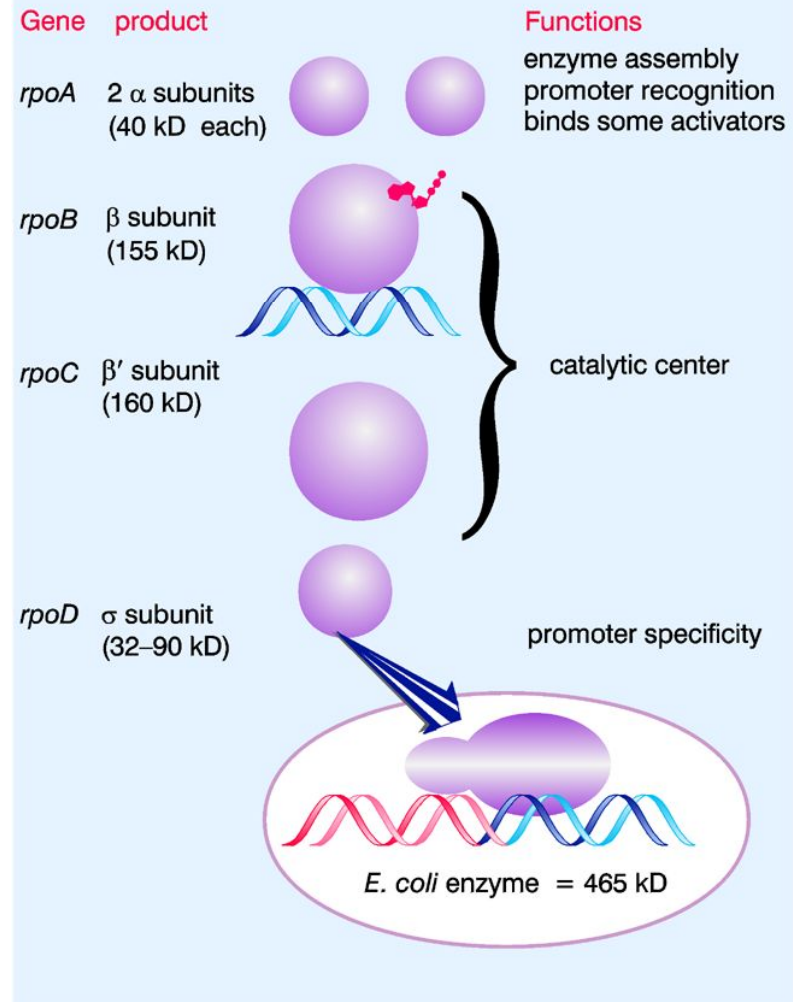
Séquences consensus

Figure 9.15 A typical promoter has three components, consisting of consensus sequences at -35 and -10, and the startpoint.



L'ARN polymérase bactérienne

Figure 9.9 Eubacterial RNA polymerases have four types of subunit; α , β , and β' have rather constant sizes in different bacterial species, but σ varies more widely.



Les facteurs sigma

Figure 9.19 *E. coli* sigma factors recognize promoters with different consensus sequences. (Numbers in the name of a factor indicate its mass.)

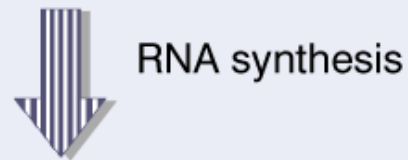
Gene	Factor	Use	-35 Sequence	Separation	-10 Sequence
<i>rpoD</i>	σ^{70}	general	TTGACA	16–18 bp	TATAAT
<i>rpoH</i>	σ^{32}	heat shock	CCCTTGAA	13–15 bp	CCCGATNT
<i>rpoE</i>	σ^E	heat shock	not known	not known	not known
<i>rpoN</i>	σ^{54}	nitrogen	CTGGNA	6 bp	TTGCA
<i>fliA</i>	σ^F	flagellar	CTAAA	15 bp	GCCGATAA

Synthèse de l'ARN

Figure 1.28 RNA is synthesized by using one strand of DNA as a template for complementary base pairing.

DNA consists of two base-paired strands

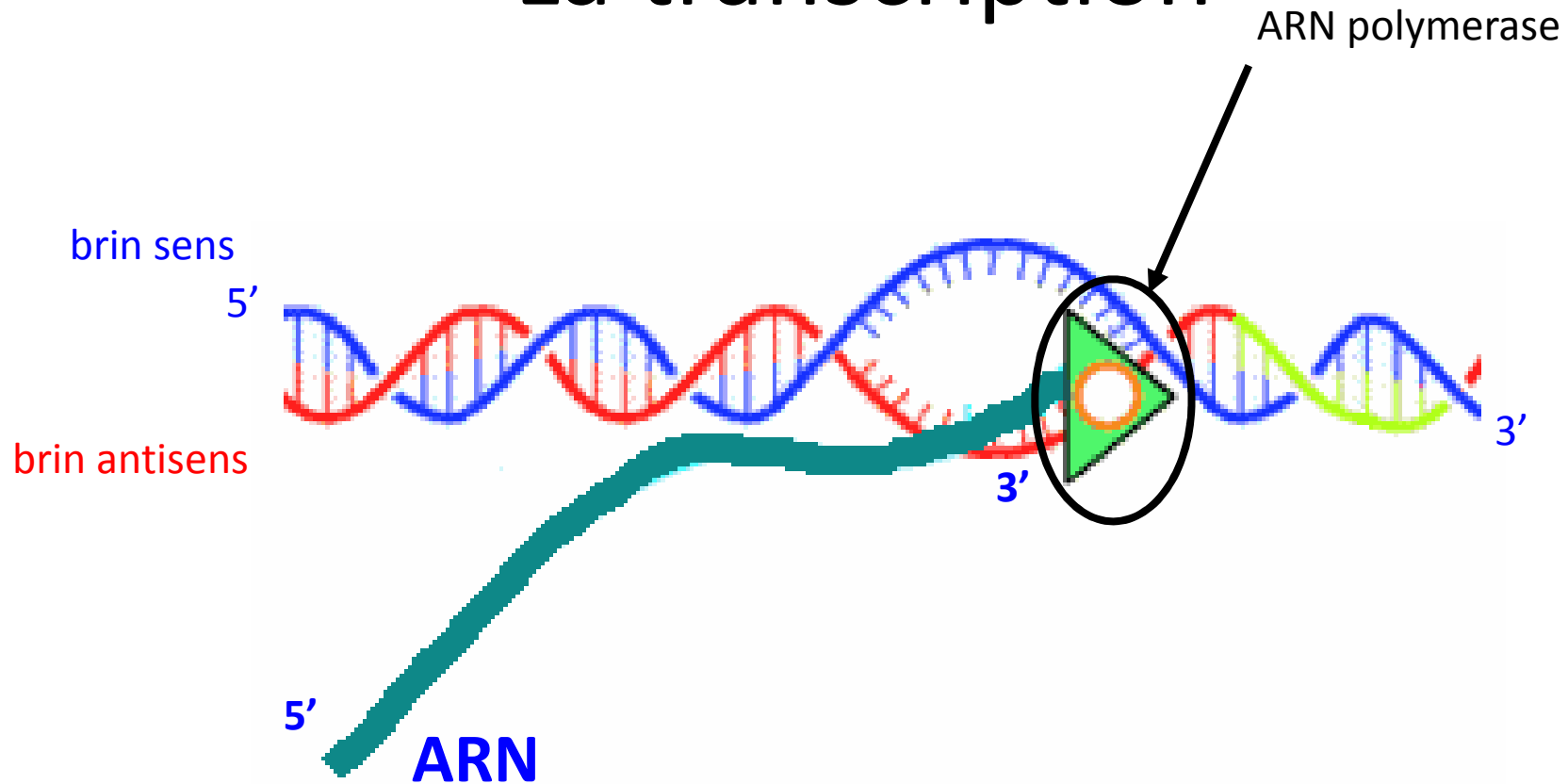
5' ATGCCGTTAGACCGTTAGCGGACCTGAC top strand
3' TACGGCAATCTGGCAATCGCCTGGACTG bottom strand



5' AUGCCGUUAGACCGUUAGCGGACCUGAC 3'

RNA has same sequence as DNA top strand;
is complementary to DNA bottom strand

La transcription

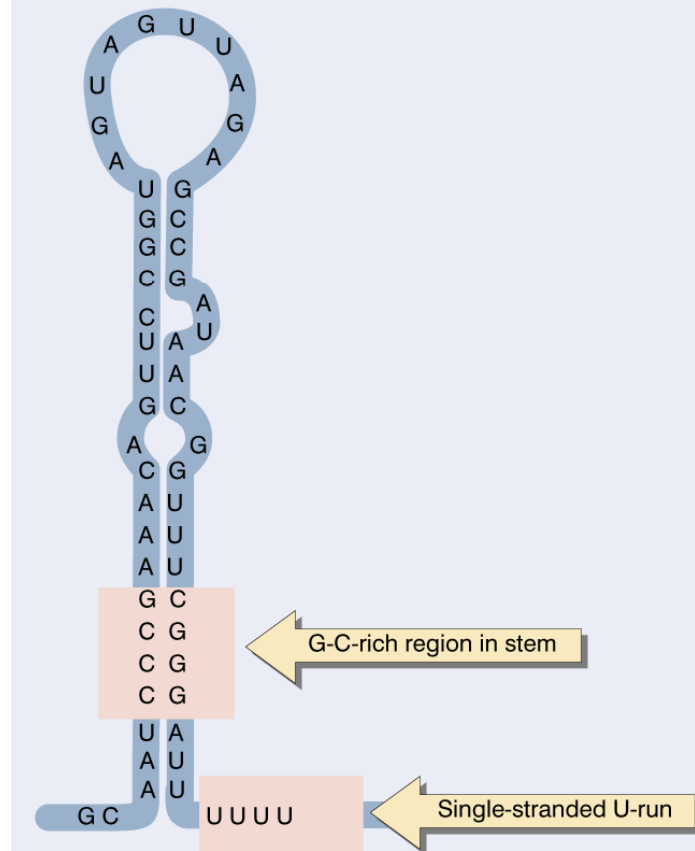


La séquence du « gène » est la même (T->U) que celle de l'ARN messenger, car la polymérase lit son complémentaire

Les terminateurs

- La transcription s'arrête quand la polymérase « bute » sur un terminateur
- Il existe deux types de terminateurs : les terminateurs naturels et les terminateurs rho-dépendants

Figure 9.27 Intrinsic terminators include palindromic regions that form hairpins varying in length from 7 to 20 bp. The stem-loop structure includes a G-C-rich region and is followed by a run of U residues.



Les terminateurs rho-dépendants

Figure 9.28 A rho-dependent terminator has a sequence rich in C and poor in G preceding the actual site(s) of termination.

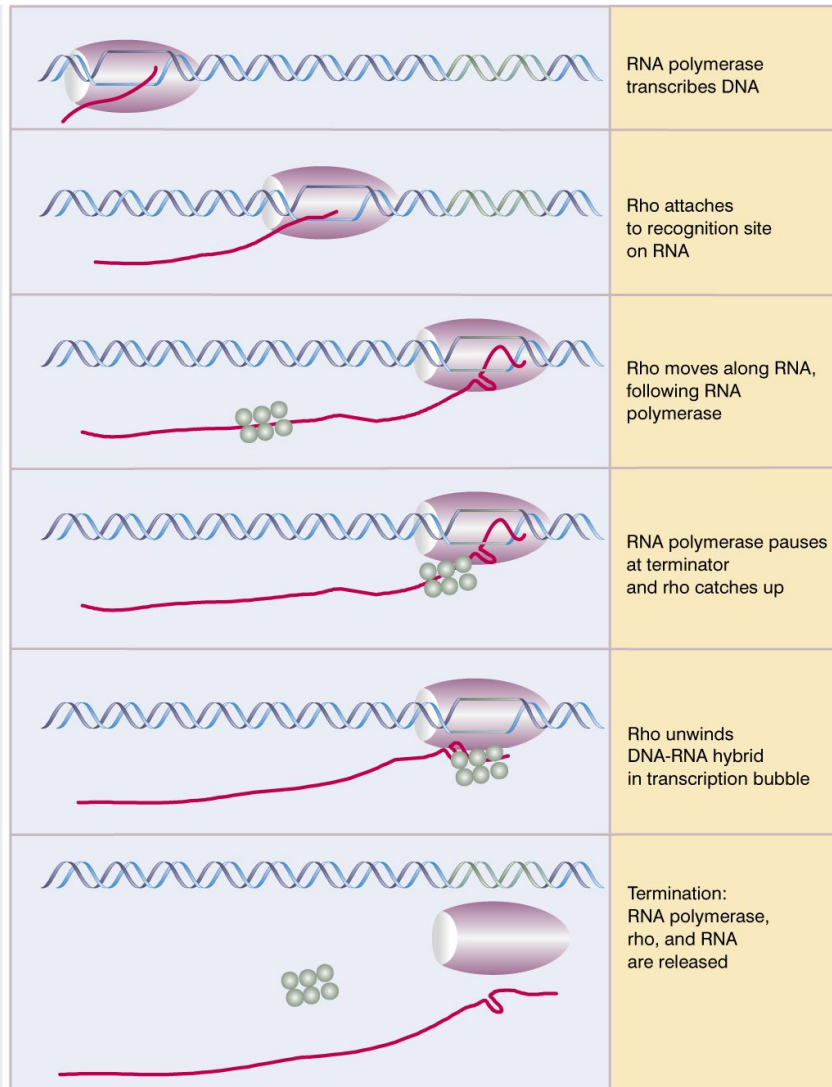
AUCGCUACCUCUAUAUCCGCACCUCUCAAACGCUACCUCGACCAGAAAGGCGUCUCUU

Bases	
C	41%
A	25%
U	20%
G	14%



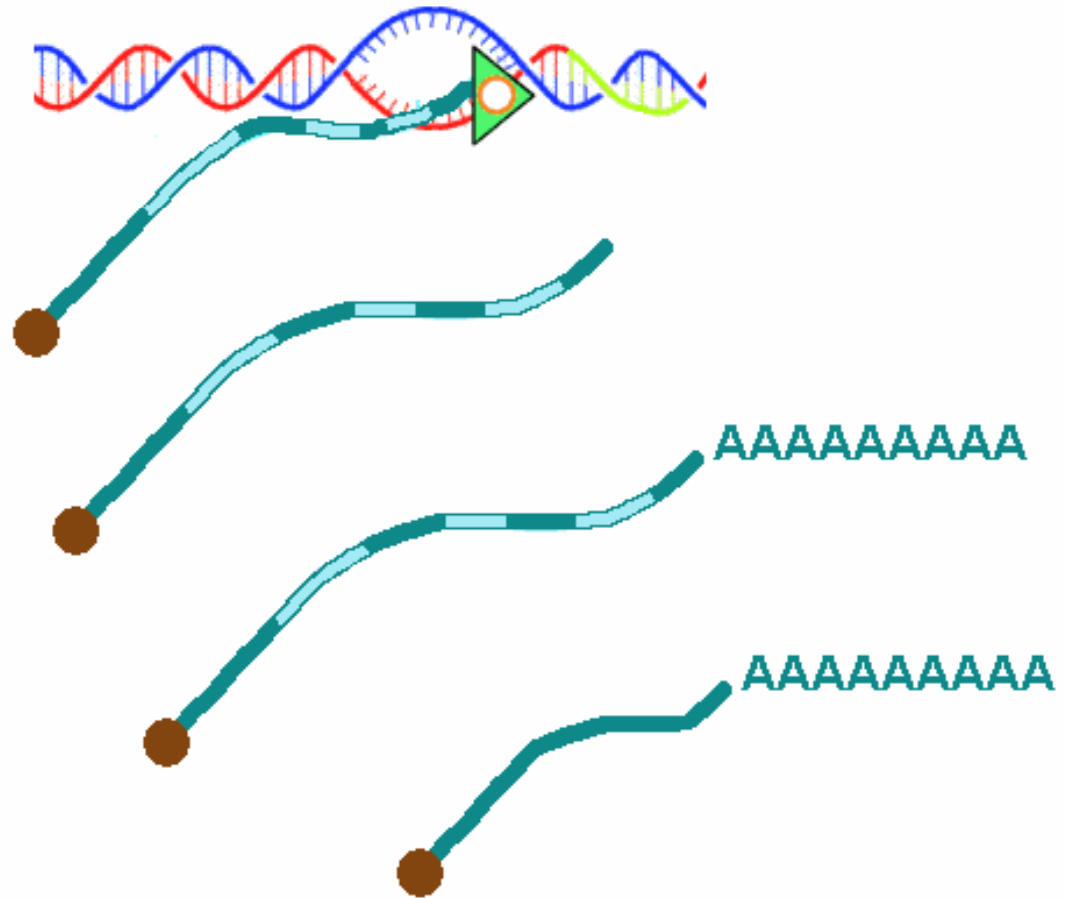
La terminaison rho

Figure 9.29 Rho factor pursues RNA polymerase along the RNA and can cause termination when it catches the enzyme pausing at a rho-dependent terminator.



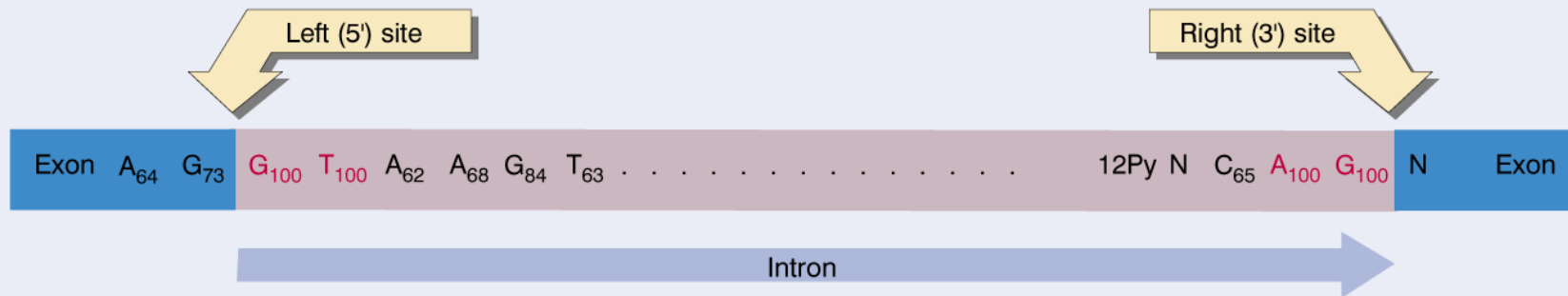
La maturation des ARN eucaryotes

1. Addition d'une coiffe en 5'
2. Polyadénylation en 3'
3. Épissage



Les introns

Figure 22.3 The ends of nuclear introns are defined by the GT-AG rule.



Epissage

Figure 22.6 Splicing occurs in two stages, in which the 5' exon is separated and then is joined to the 3' exon.

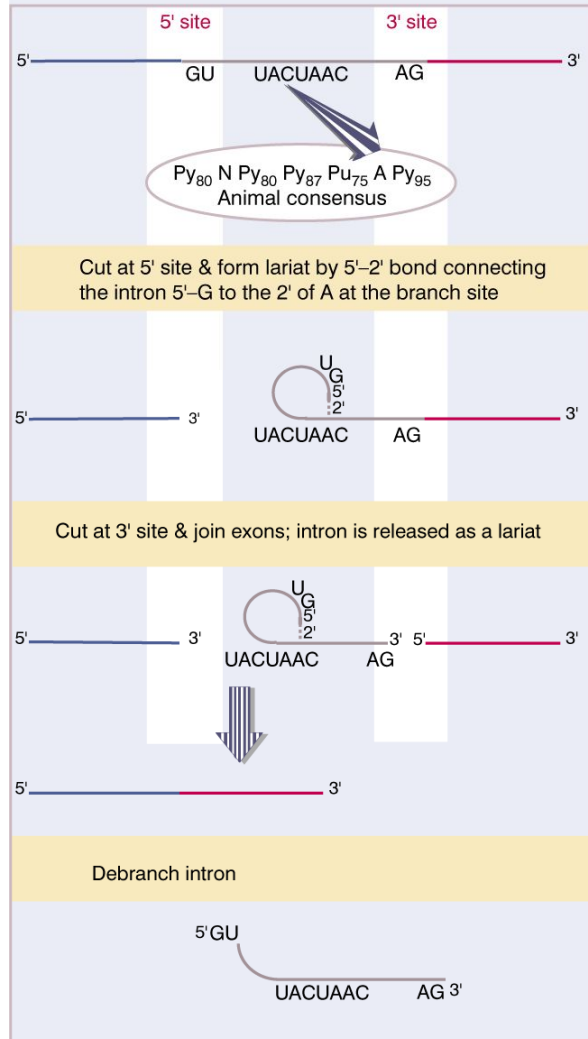


Figure 22.18 Alternative forms of splicing may generate a variety of protein products from an individual gene. Changing the splice sites may introduce termination codons (shown by asterisks) or change reading frames.

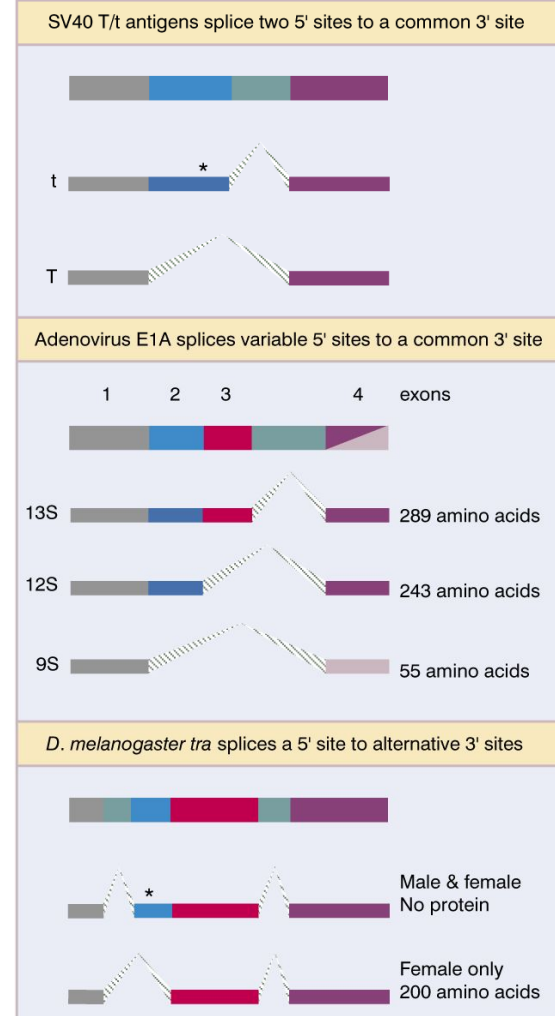
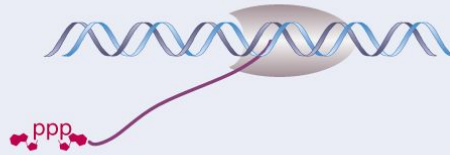
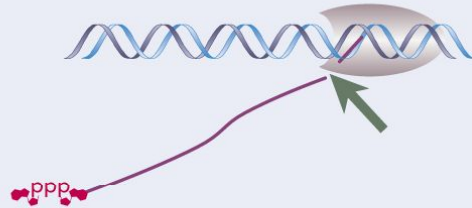


Figure 5.16 Overview: expression of mRNA in animal cells requires transcription, modification, processing, nucleocytoplasmic transport, and translation.

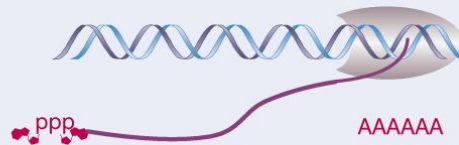
< 1 min Transcription starts: 5' end is modified



6 min 3' end of mRNA is released by cleavage



20 min 3' end is polyadenylated



25 min mRNA is transported to cytoplasm

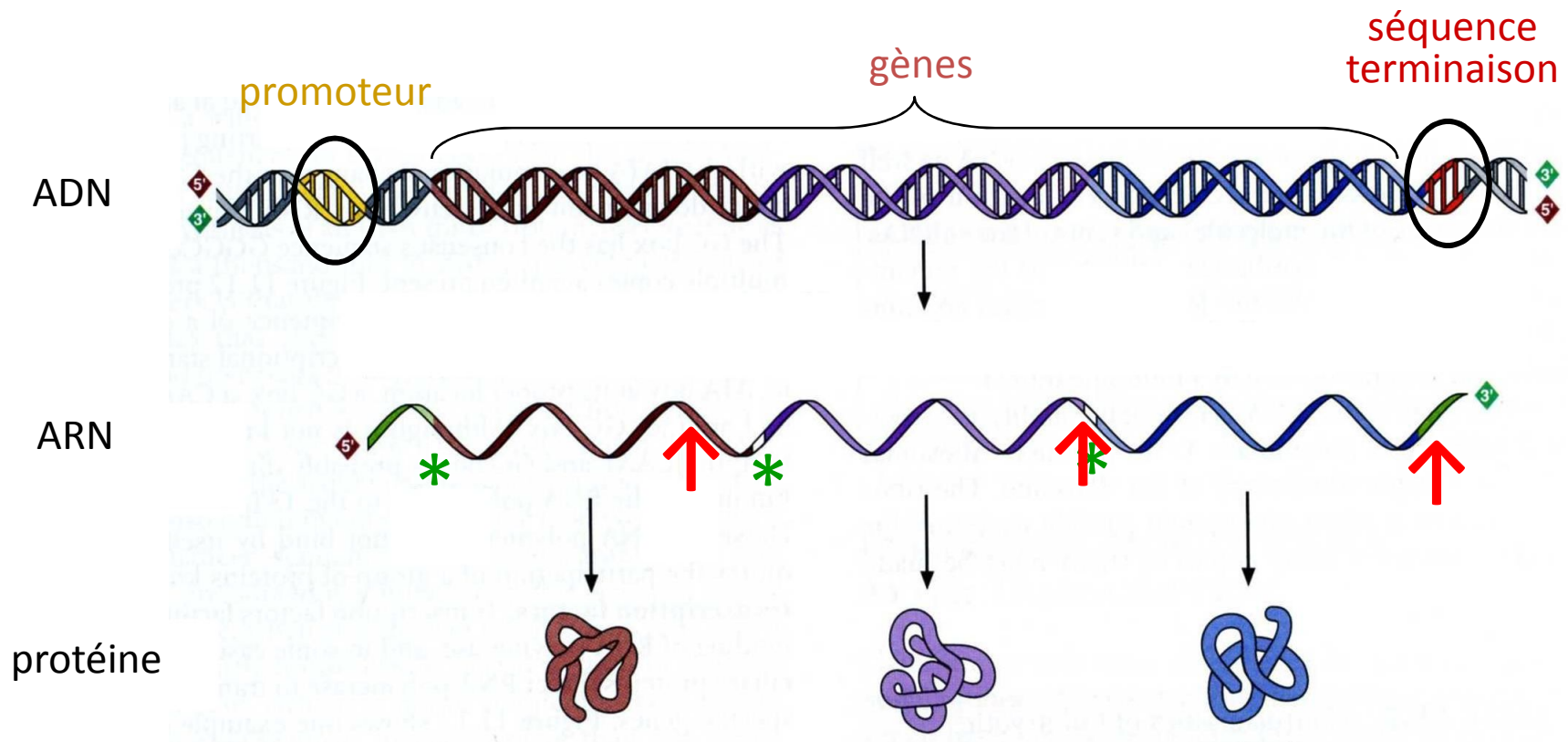


> 4 hr Ribosomes translate mRNA



Récapitulatif : Spécificités procaryotes

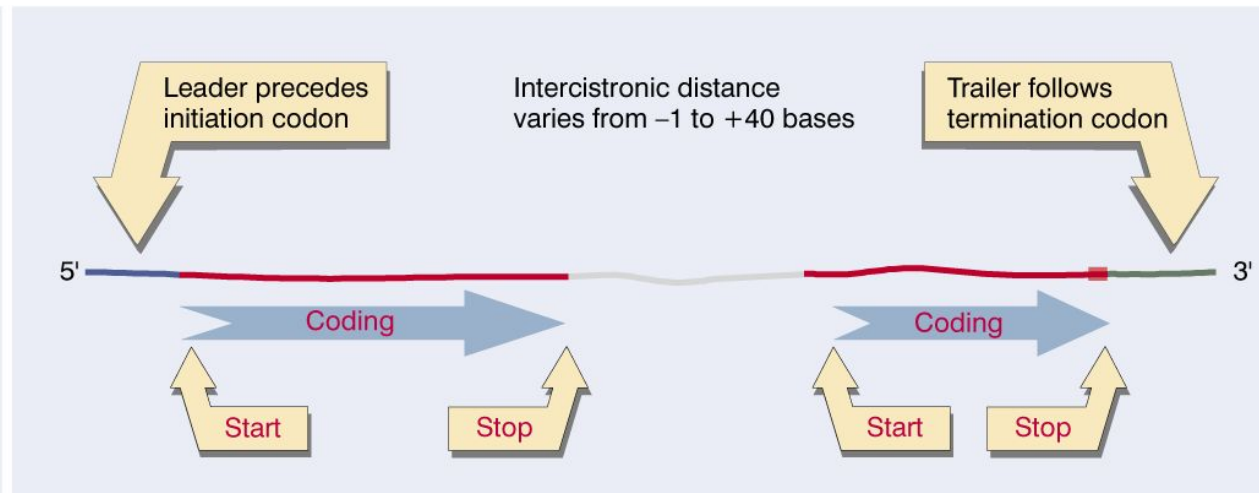
- Transcription dans le cytoplasme
- ARN polycistronique



Transcrits polycistroniques (opérons)

- Un opéron est une succession de gènes proches transcrits séquentiellement, chez les procaryotes
- Il n'y a formation que d'un seul complexe d'initiation de la transcription
- Il n'y a production que d'un seul ARNm
- Les différents gènes ont souvent des fonctions liées

Figure 5.14 Bacterial mRNA includes non-translated as well as translated regions. Each coding region has its own initiation and termination signals. A typical mRNA may have several coding regions.



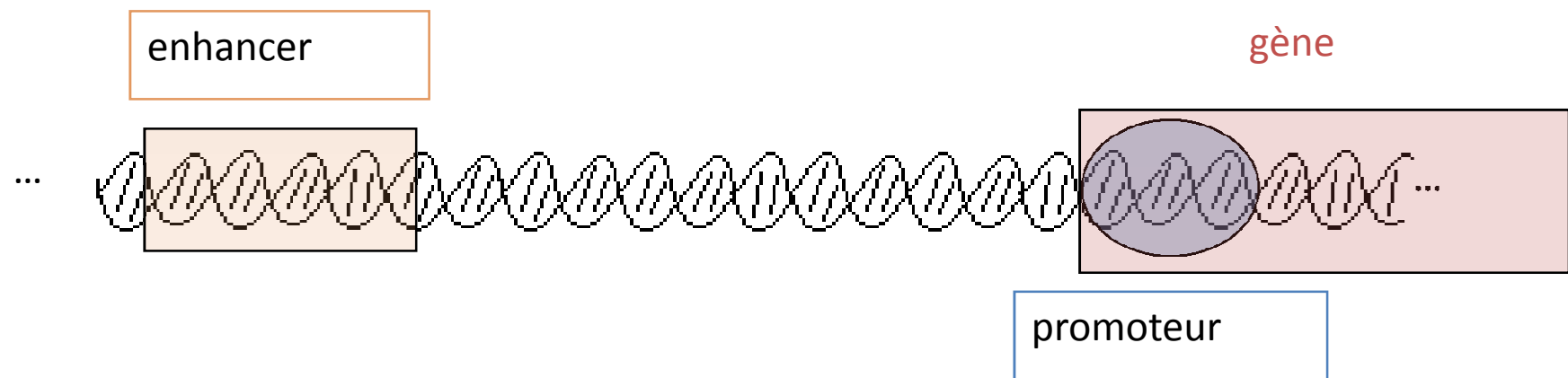
L'opéron lac

- Il est composé de **trois gènes** impliqués dans le processing du sucre complexe lactose
- L'un des gènes code pour une enzyme, la **β -galactosidase, qui clive le lactose en glucose et galactose**
- Un autre code pour une perméase qui importe le lactose dans la cellule
- Le dernier code pour un cofacteur régulant ce processus

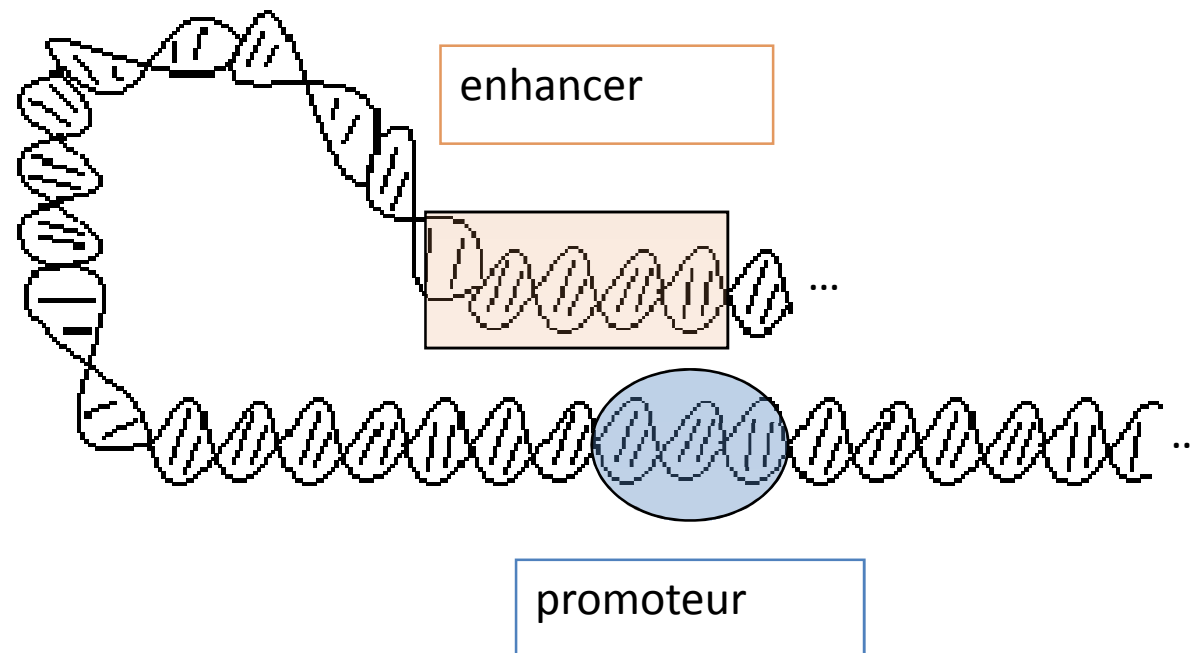
Récapitulatif : Spécificités eucaryotes

- Pas de transcrits polycistroniques
- Transcription dans le noyau
- Facteurs de transcription plus employés
- Enhancers et silencers
- Maturation des transcrits

Spécificités eucaryotes

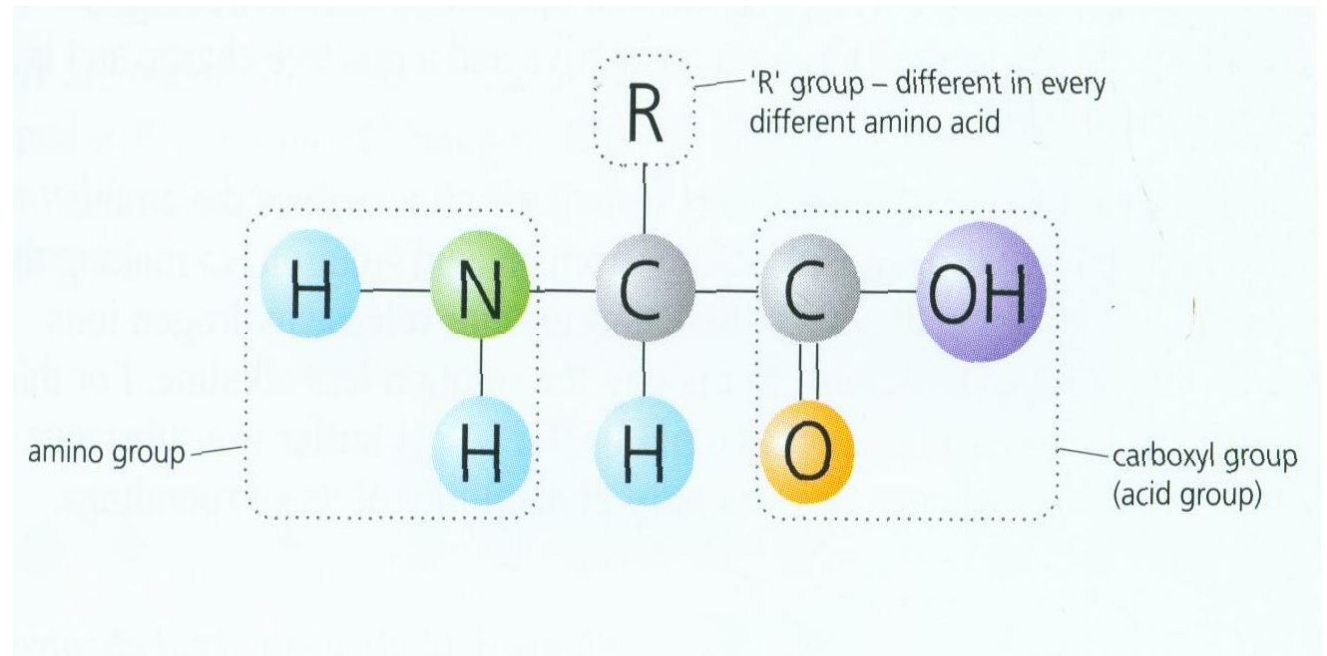


Spécificités eucaryotes



Acides aminés

- Formule générale:



- Contiennent

- Acide

- Carboxylique

- Amine

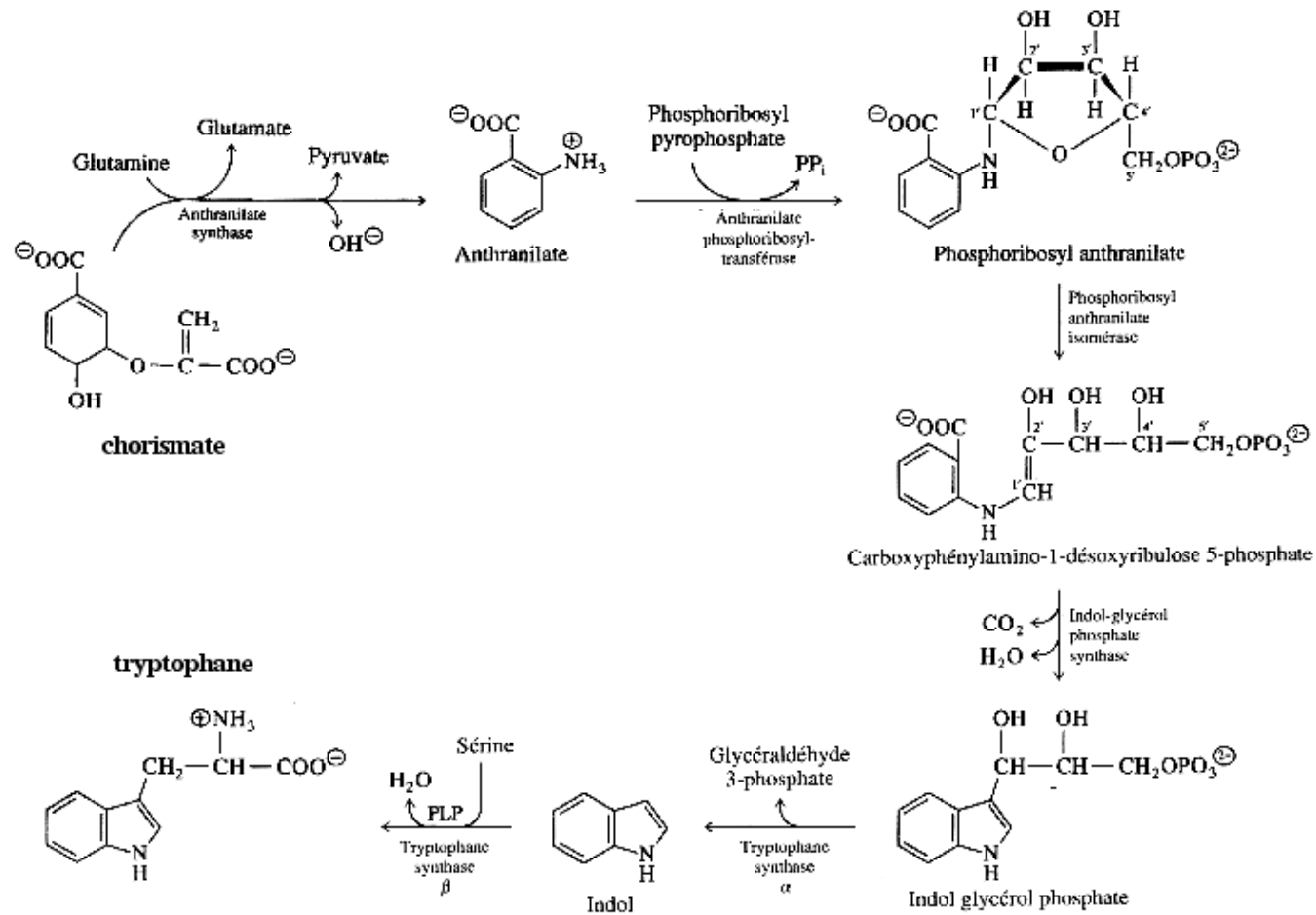
- Radical “R”

- Très grande variété
- Propriétés physico-chimiques variées

Acides aminés

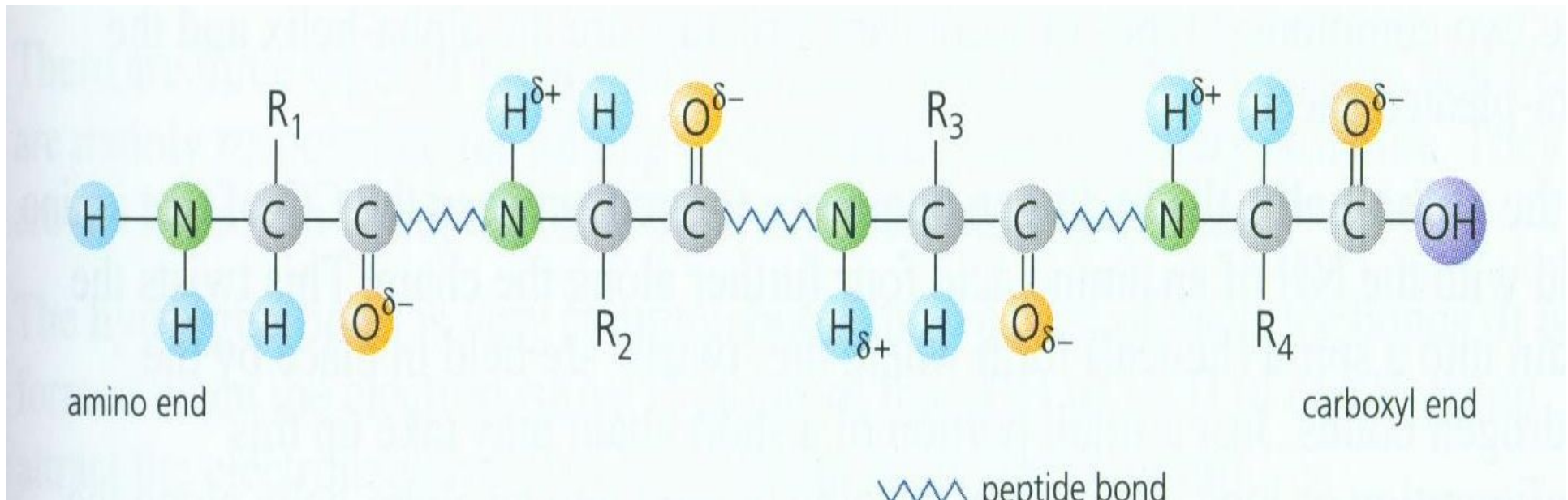
- Certains acides aminés sont le fruit d'un processus métabolique (par exemple le tryptophane) : on parle de biosynthèse
- Les autres doivent être obtenus à partir de l'extérieur : ce sont les acides aminés essentiels
- Ce qui est essentiel pour une espèce ne l'est pas pour l'autre.

Biosynthèse du tryptophane



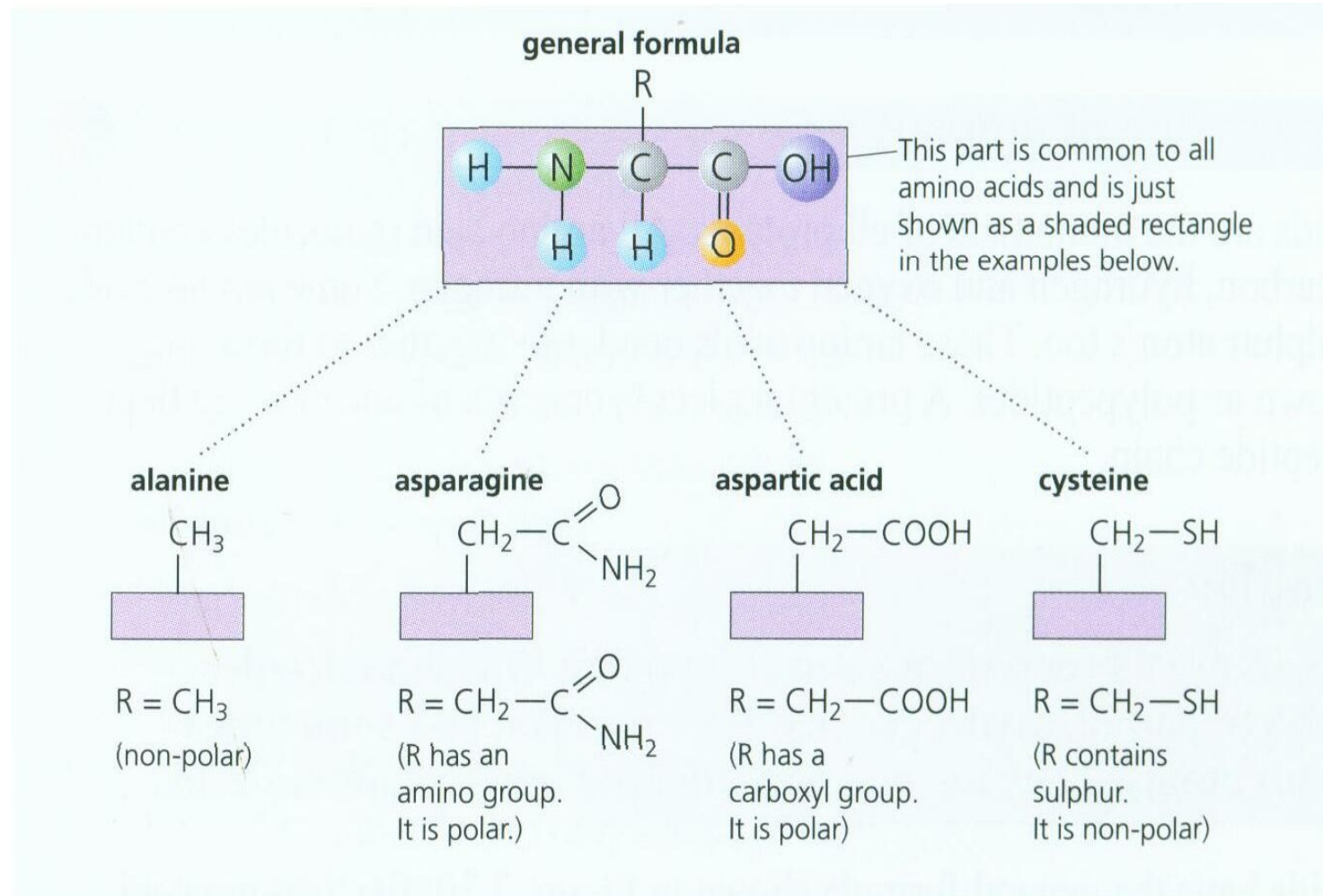
Formation des polypeptides

- Des réactions de condensation forment les polypeptides
 - Une extrémité amine et une acide carboxylique
 - Pas d'ordre obligatoire



Polypeptides

- On trouve 20 (+3) acides aminés différents chez les êtres vivants
- Un polypeptide peut être aussi long que nécessaire



Le code génétique

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

Third letter

Les ARN de transfert

Le code génétique est représenté au niveau moléculaire par les ARN de transfert:

- petits ARNs
- « lus » par une ARN polymérase spéciale
- forme de feuille de trèfle
- un anticodon
- un site porteur d'acide aminé très spécifique

Figure 5.2 A tRNA has the dual properties of an adaptor that recognizes both the amino acid and codon. The 3' adenosine is covalently linked to an amino acid. The anticodon base pairs with the codon on mRNA.

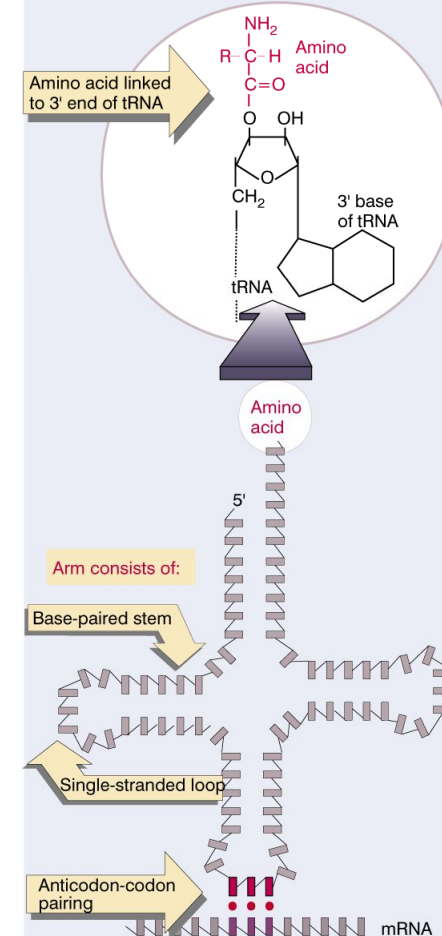
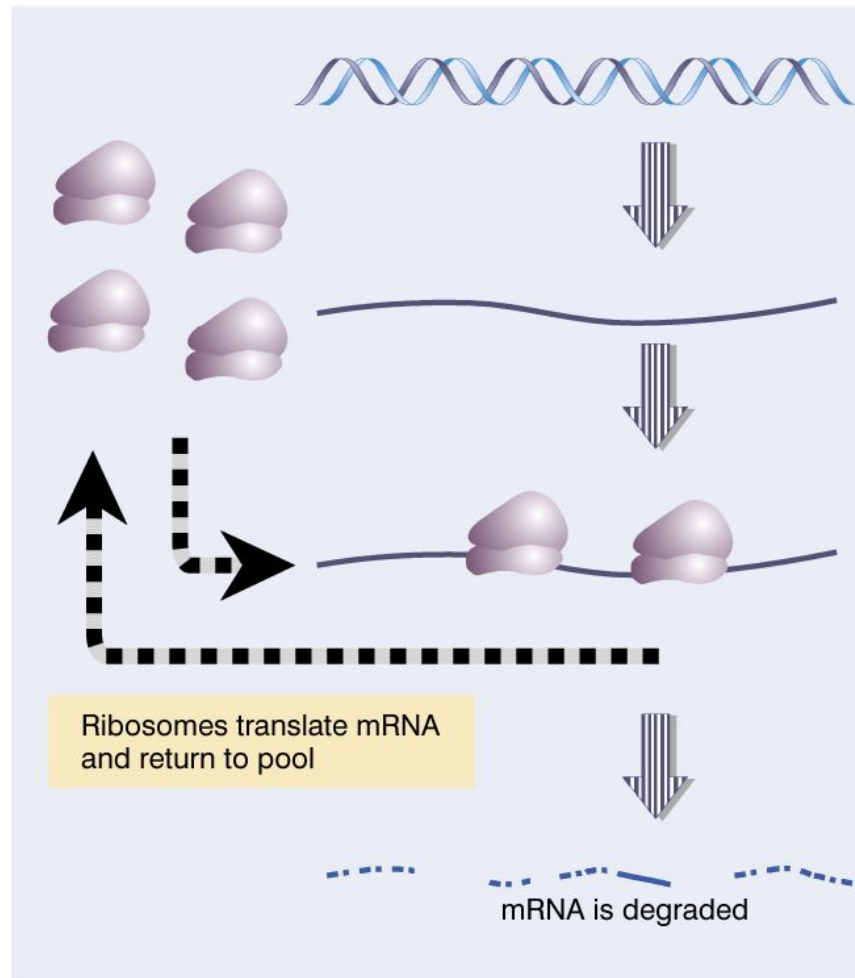


Figure 5.5 A space-filling model shows that tRNA^{Phe} tertiary structure is compact. The two views of tRNA are rotated by 90°. The top view corresponds with the bottom panel in Figure 5.4. Photograph kindly provided by S. H. Kim.



La traduction

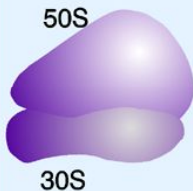
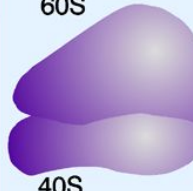
Figure 5.10 Messenger RNA is translated by ribosomes that cycle through a pool.



Le retour des ribosomes

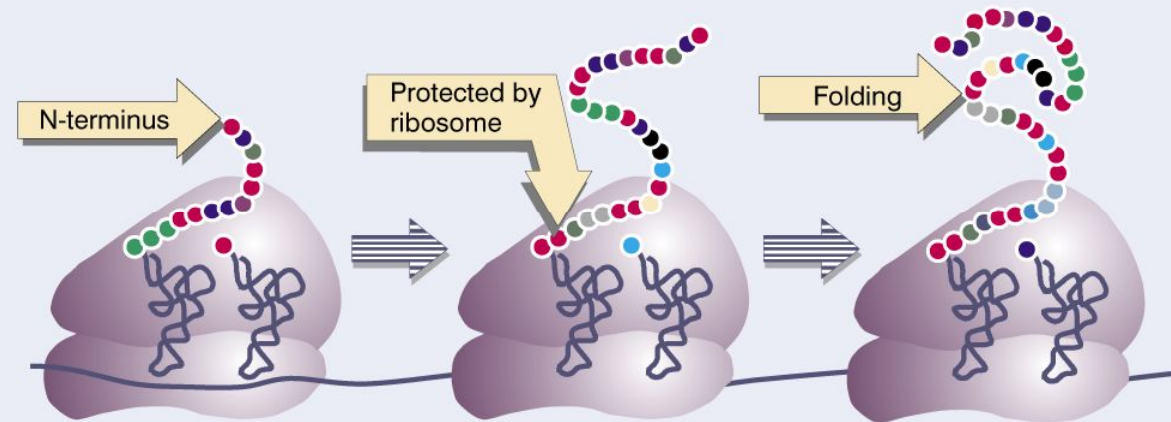
Chez les bactéries, les ribosomes peuvent représenter plus du quart du poids total de la cellule, ce qui montre l'importance relative en termes de coût de la phase de traduction pour ces organismes.

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

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

Le polysome

Figure 5.8 A polyribosome consists of an mRNA being translated simultaneously by several ribosomes moving in the direction from 5' to 3'. Each ribosome has two tRNA molecules: one carrying the nascent protein, the second carrying the next amino acid to be added.



Mécanique du ribosome

Figure 6.5 tRNA and mRNA move through the ribosome in the same direction.

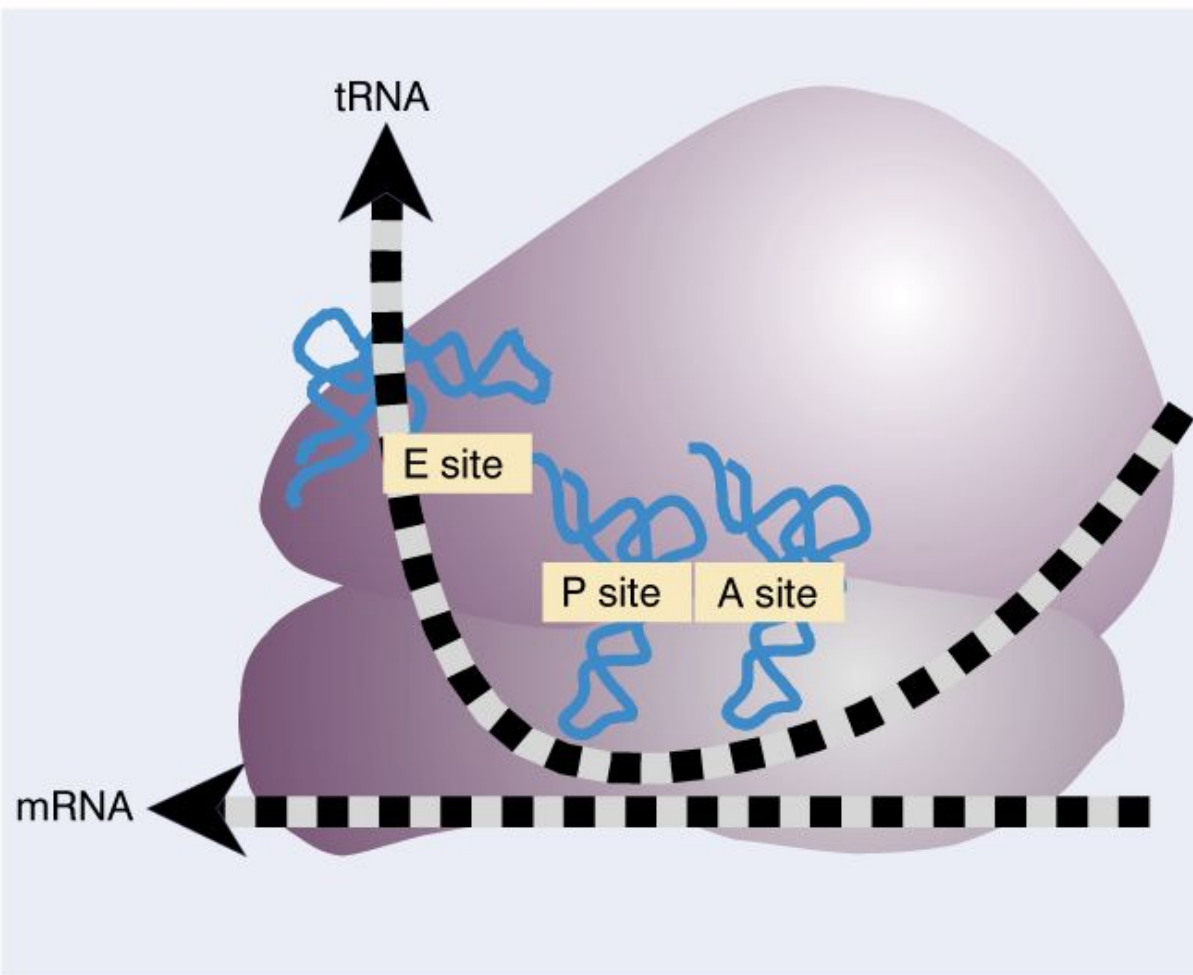


Figure 6.22 Models for translocation involve two stages. First, at peptide bond formation the aminoacyl end of the tRNA in the A site becomes located in the P site. Second, the anticodon end of the tRNA becomes located in the P site.

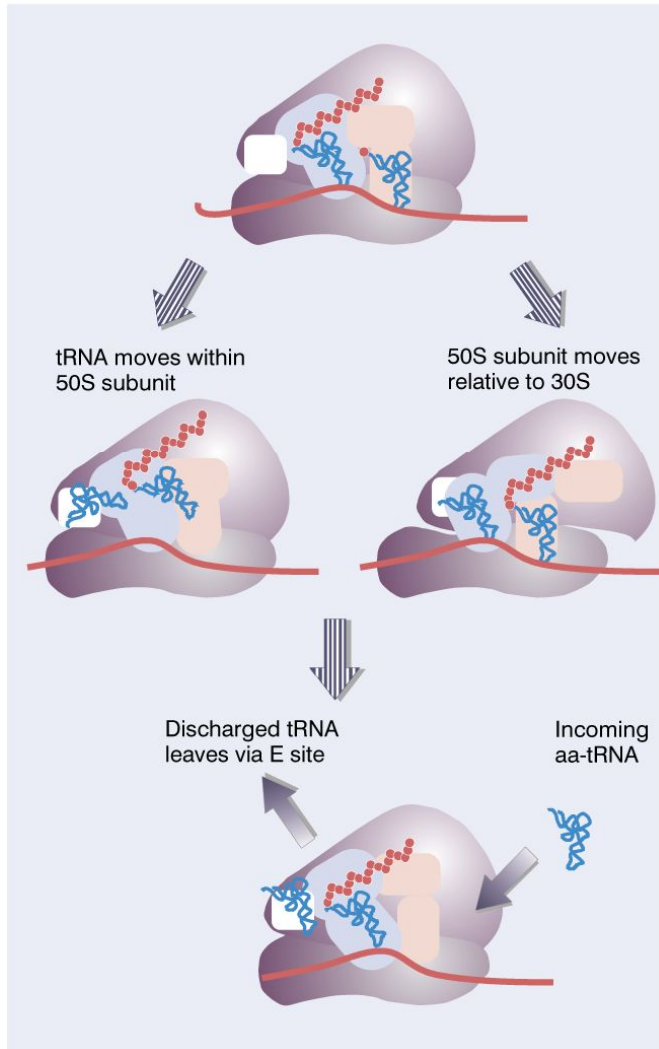
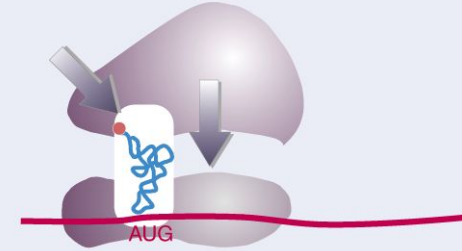


Figure 6.6 Protein synthesis falls into three stages.

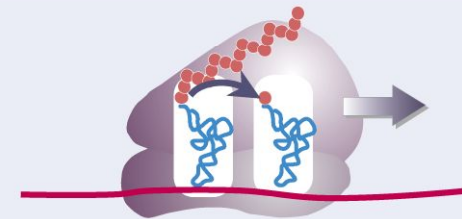
Initiation

30S subunit on mRNA binding site is joined by 50S subunit and aminoacyl-tRNA binds



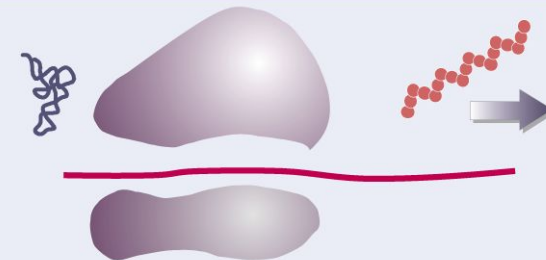
Elongation

Ribosome moves along mRNA and length of protein chain extends by transfer from peptidyl-tRNA to aminoacyl-tRNA



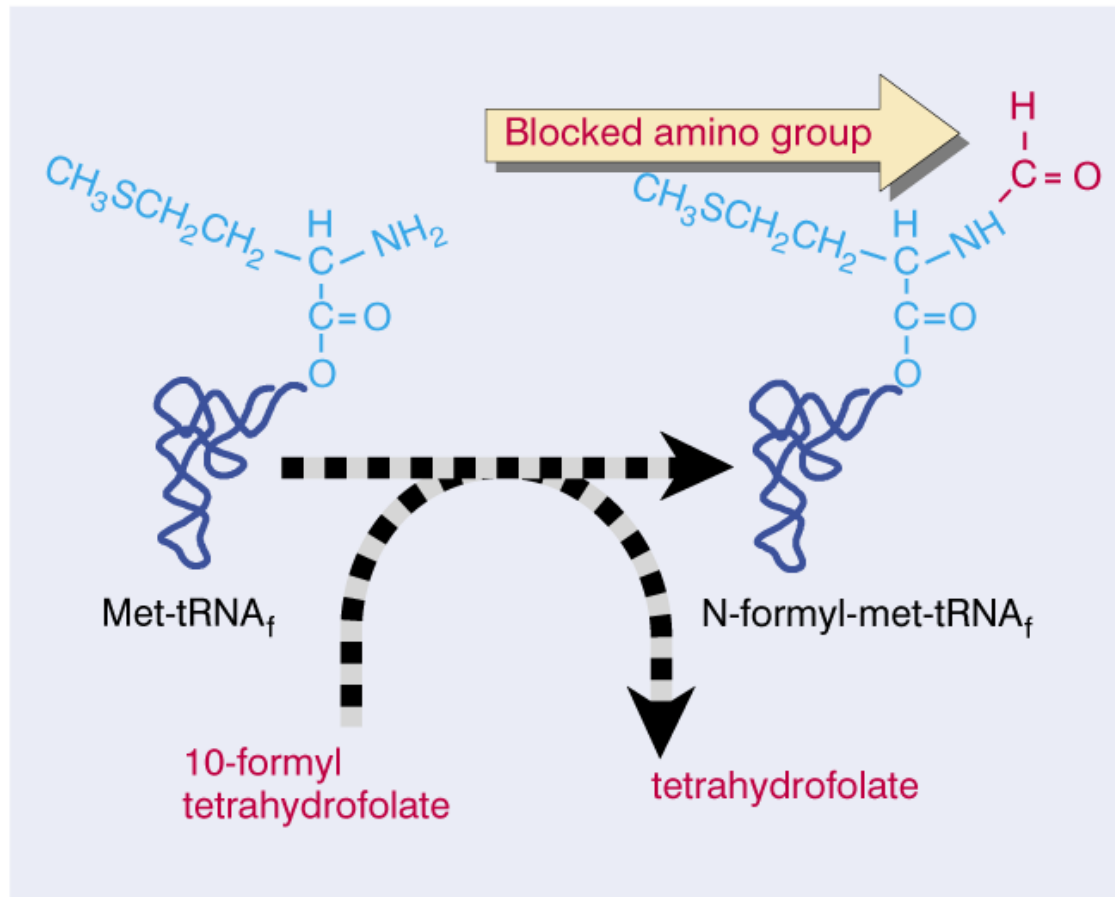
Termination

Polypeptide chain is released from tRNA, and ribosome dissociates from mRNA



ARNt initiateurs

Figure 6.9 The initiator N-formyl-methionyl-tRNA (fMet-tRNA_f) is generated by formylation of methionyl-tRNA, using formyl-tetrahydrofolate as cofactor.



Les erreurs

Figure 7.18 Missense suppression occurs when the anticodon of tRNA is mutated so that it responds to the 'wrong' codon. The suppression is only partial because both the wild-type tRNA and the suppressor tRNA can respond to AGA.

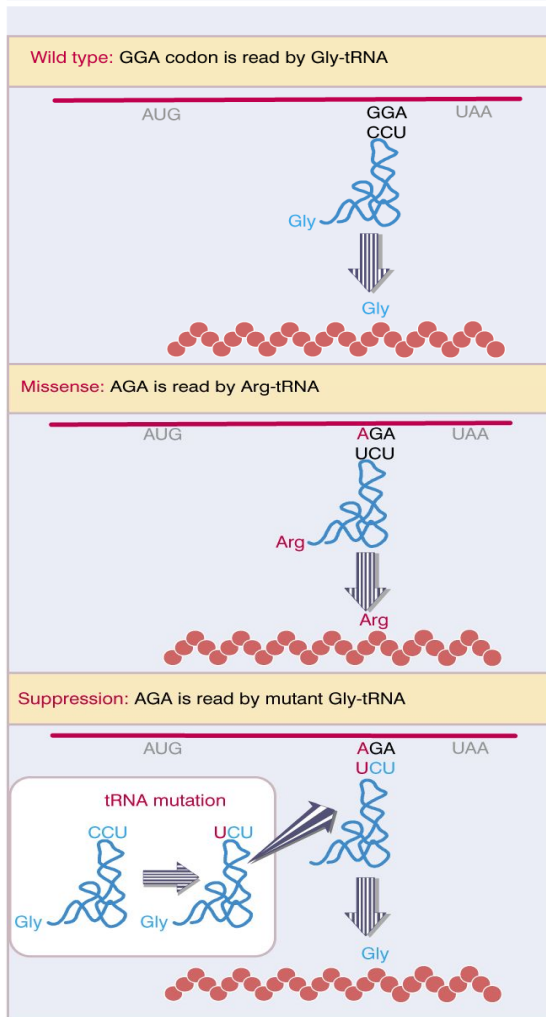


Figure 7.19 Nonsense suppressors also read through natural termination codons, synthesizing proteins that are longer than wild-type.

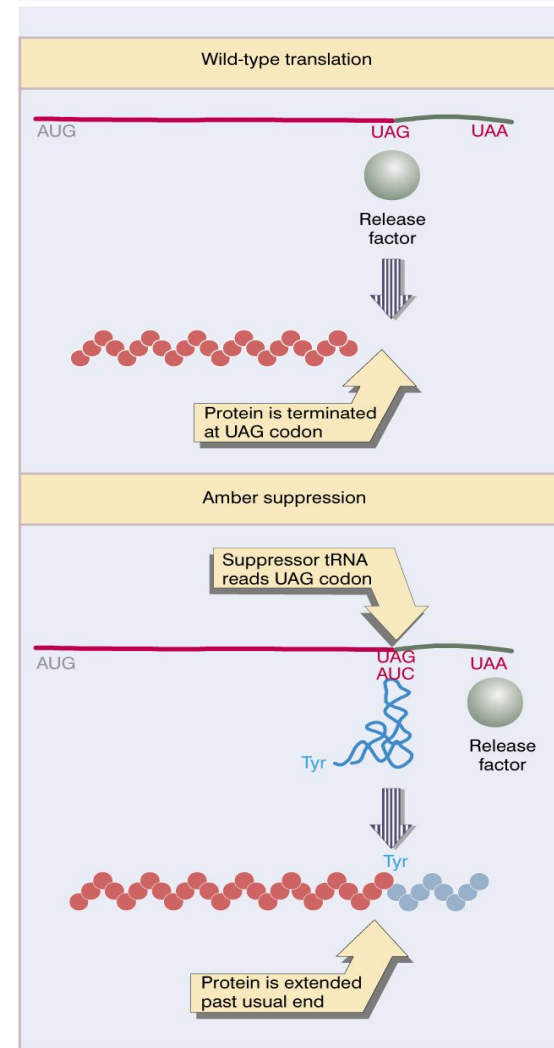
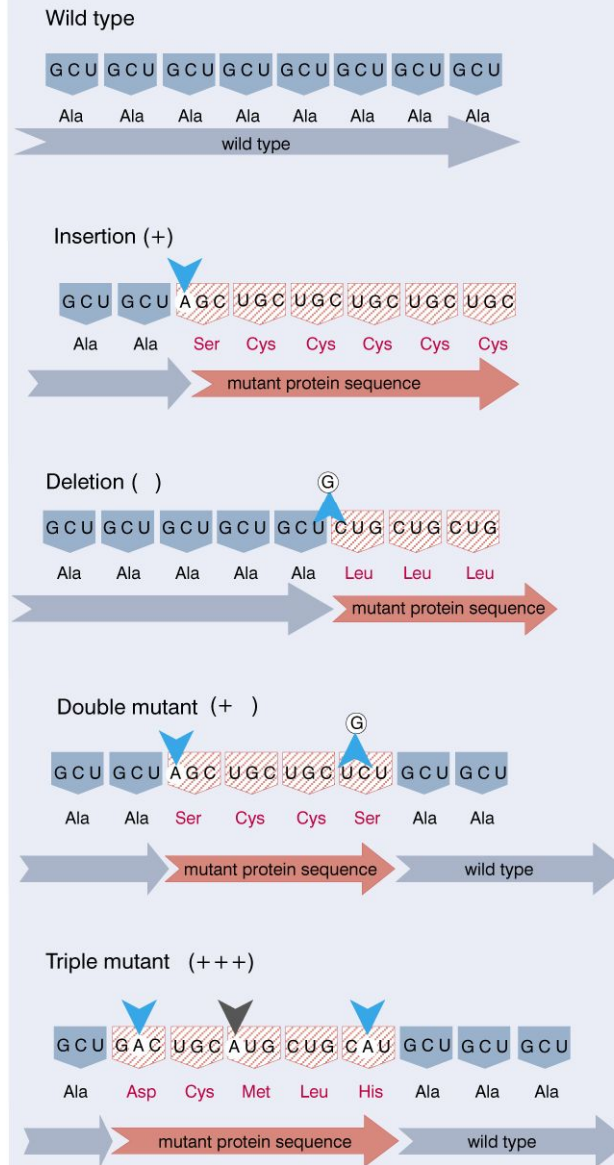
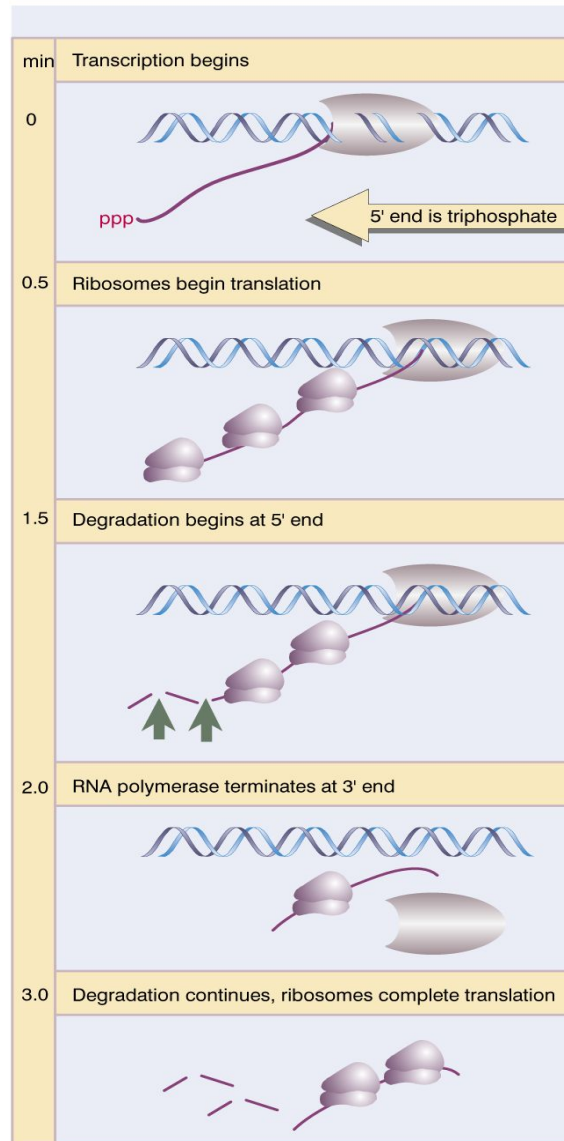


Figure 1.25 Frameshift mutations show that the genetic code is read in triplets from a fixed starting point.



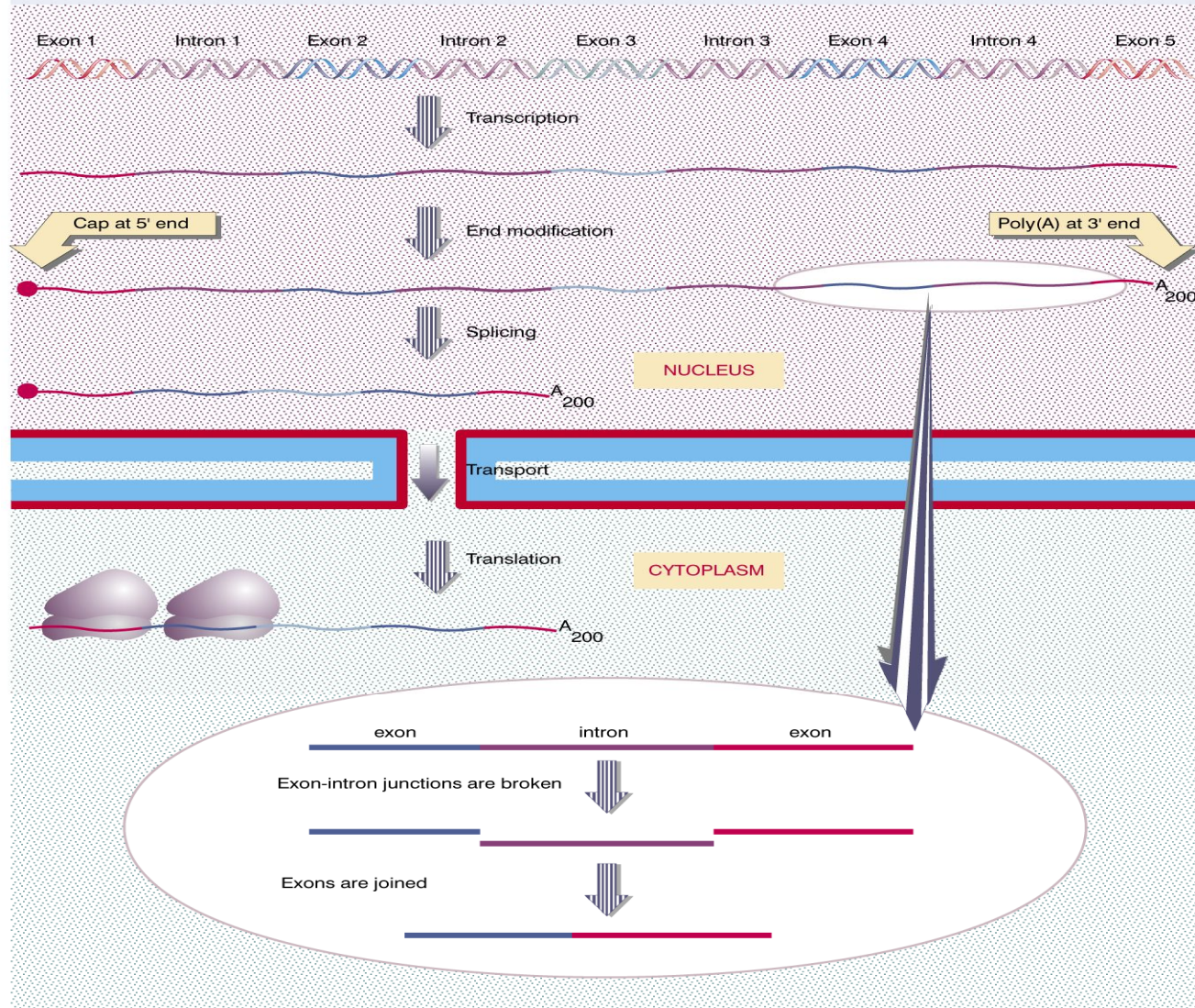
Récapitulatif bactérien

Figure 5.12 Overview: mRNA is transcribed, translated, and degraded simultaneously in bacteria.



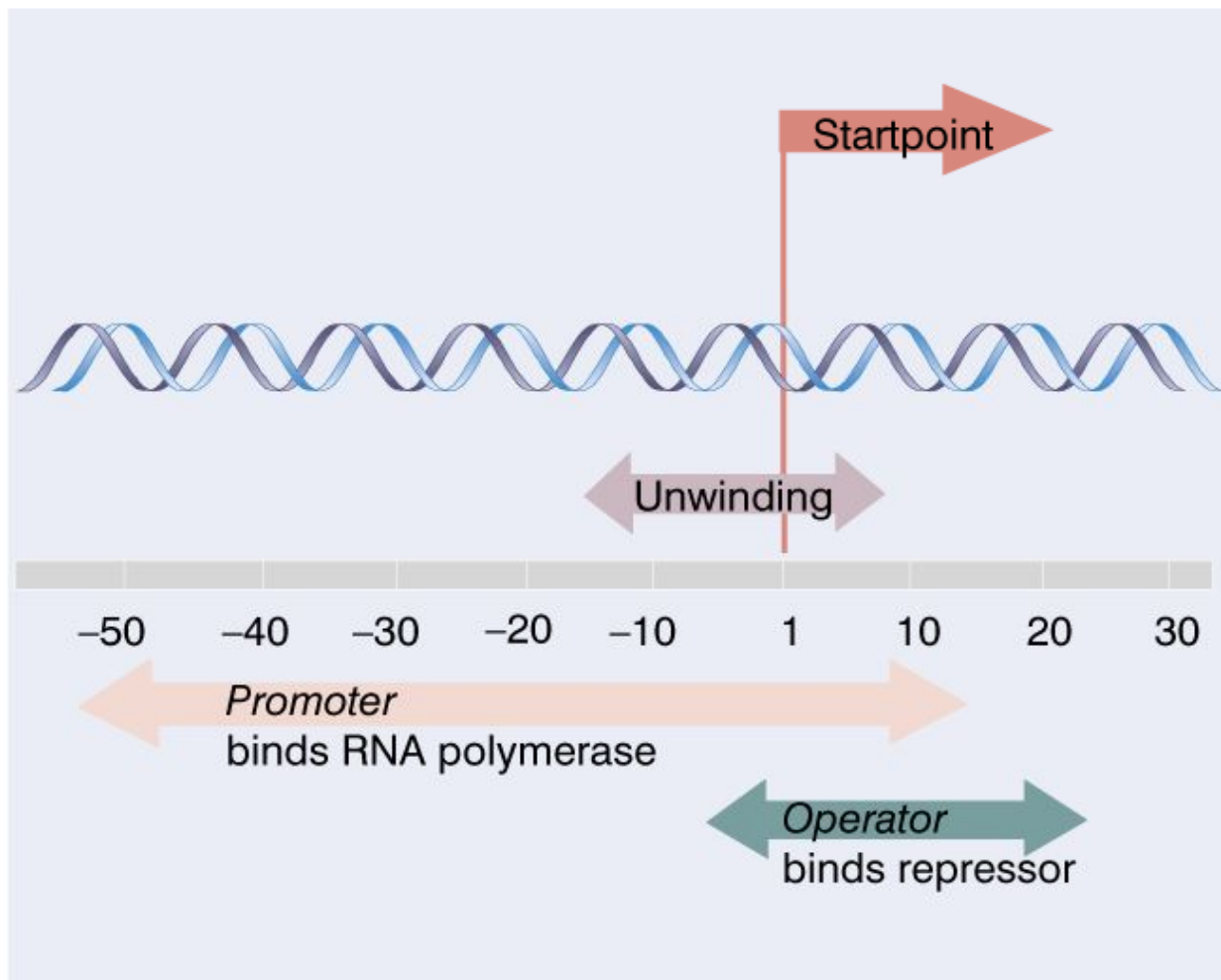
Récapitulatif eucaryote

Figure 22.2 Overview: RNA is modified in the nucleus by additions to the 5' and 3' ends and by splicing to remove the introns. The splicing event requires breakage of the exon-intron junctions and joining of the ends of the exons; the expanded illustration shows the principle schematically, but not the actual order of events. Mature mRNA is transported through nuclear pores to the cytoplasm, where it is translated.



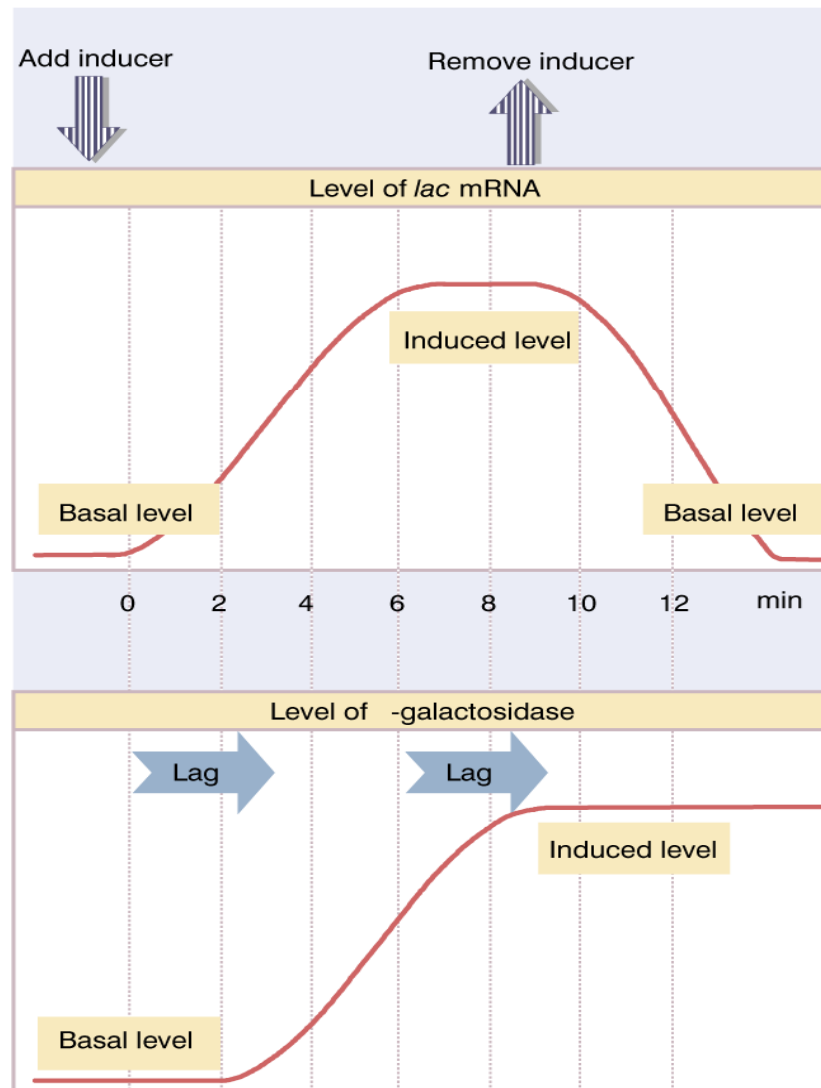
L'opéron lactose

Figure 10.4 Repressor and RNA polymerase bind at sites that overlap around the startpoint of the *lac* operon.



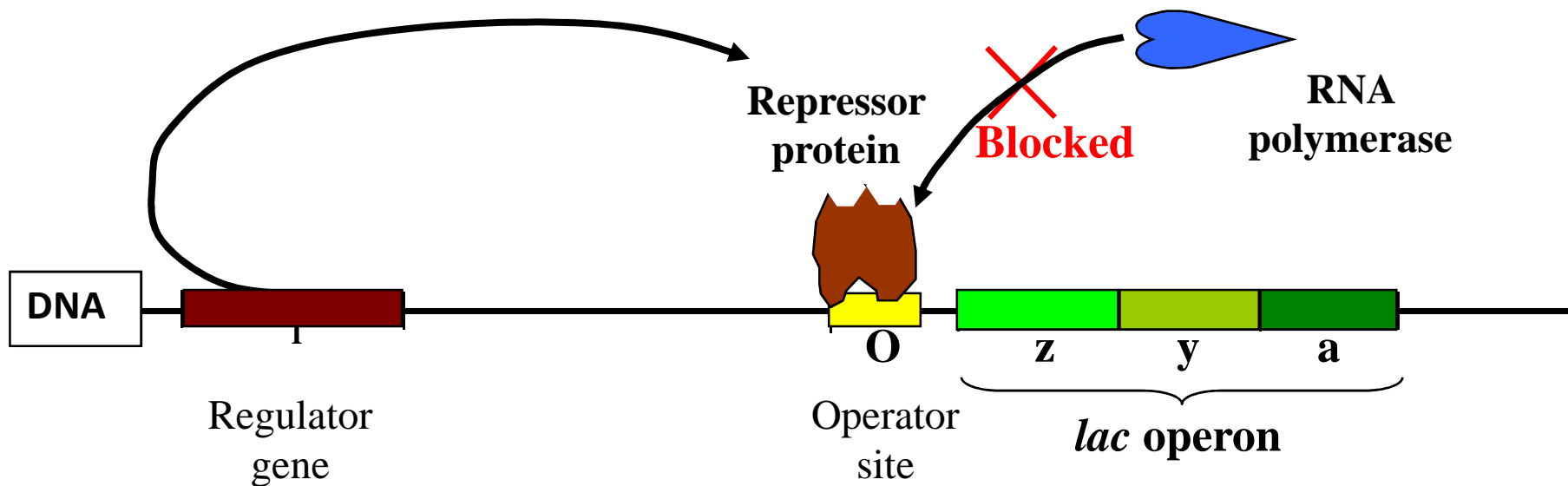
L'opéron lactose

Figure 10.5 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.



Quand le lactose est absent

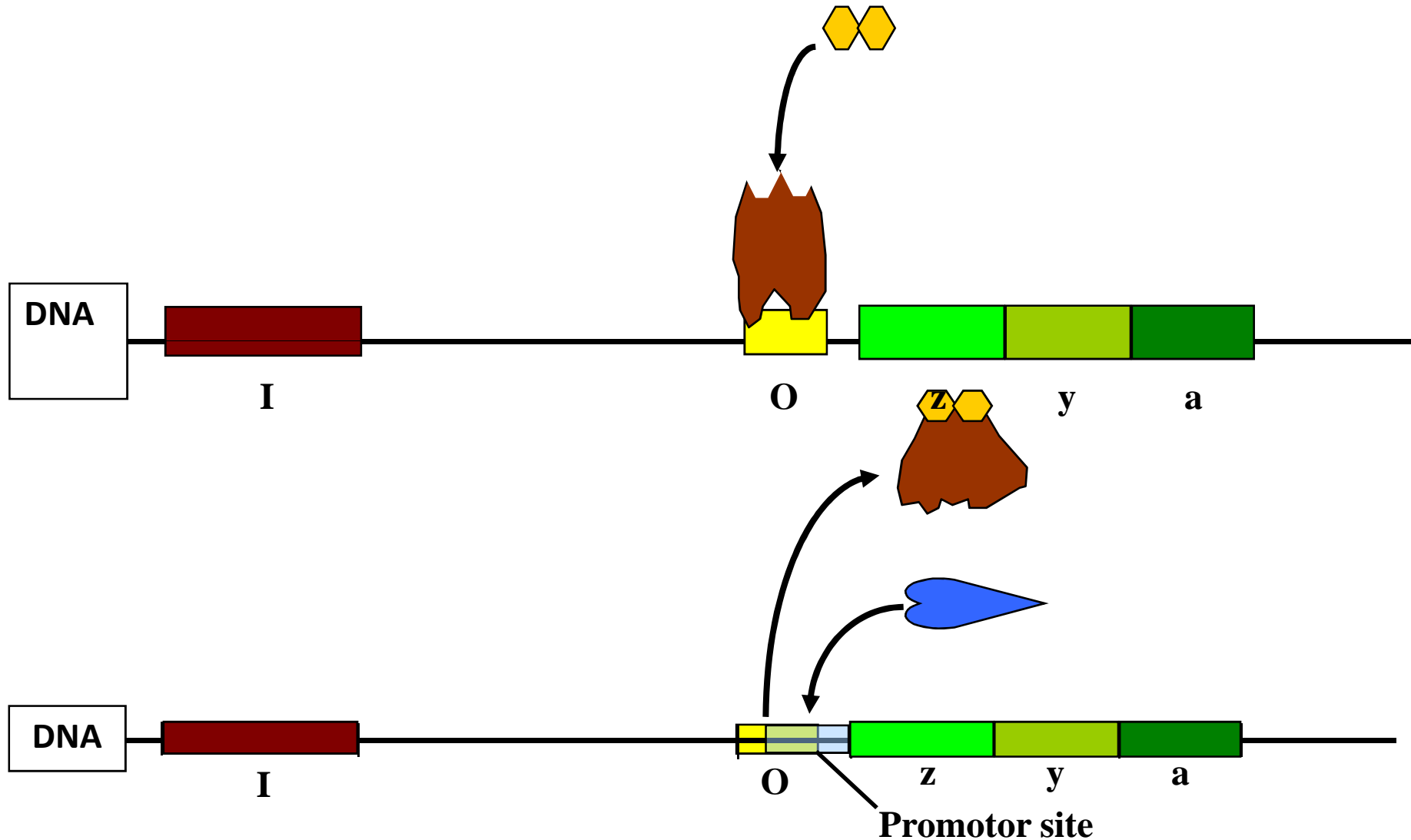
- Un répresseur (une protéine) est présent en permanence sur le site opérateur, en amont de l'opéron.
- Ce répresseur bloque la transcription en empêchant l'ARN polymérase de la débiter.



Quand le lactose est présent

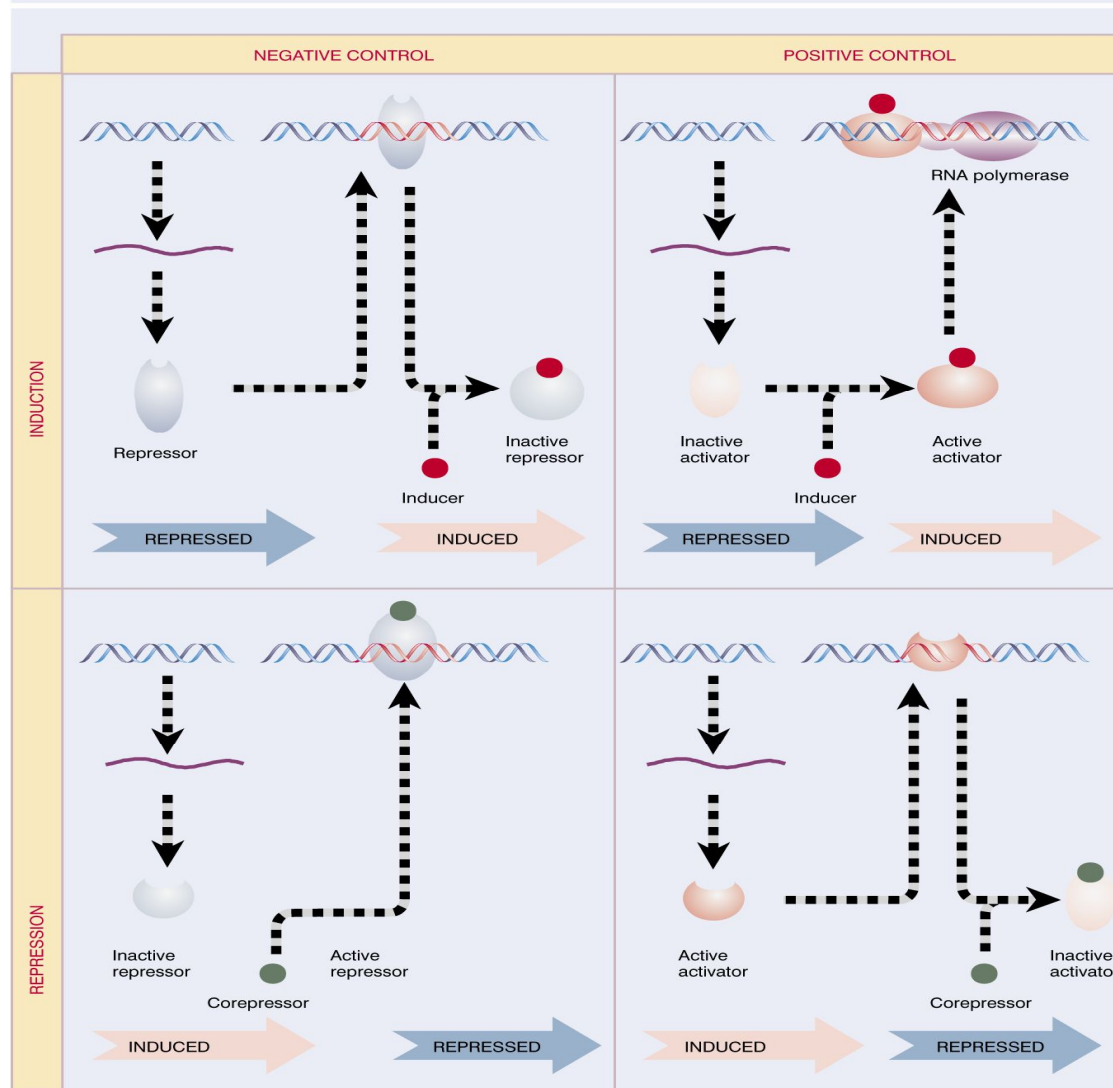
- Une faible quantité d'allolactose est formée dans la cellule. Ce sucre va se fixer sur le répresseur sur un site actif, dit allostérique car ce n'est pas celui par lequel le répresseur se fixe sur l'ADN.
- Ceci induit un changement de conformation du répresseur, qui se détache alors du site opérateur. L'ARN polymérase peut alors initier la transcription et l'opéron est exprimé.

Quand le lactose est présent

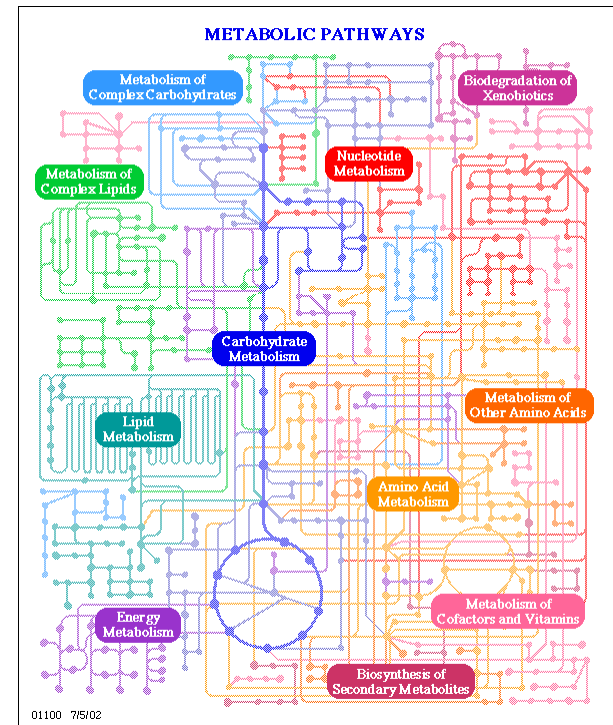
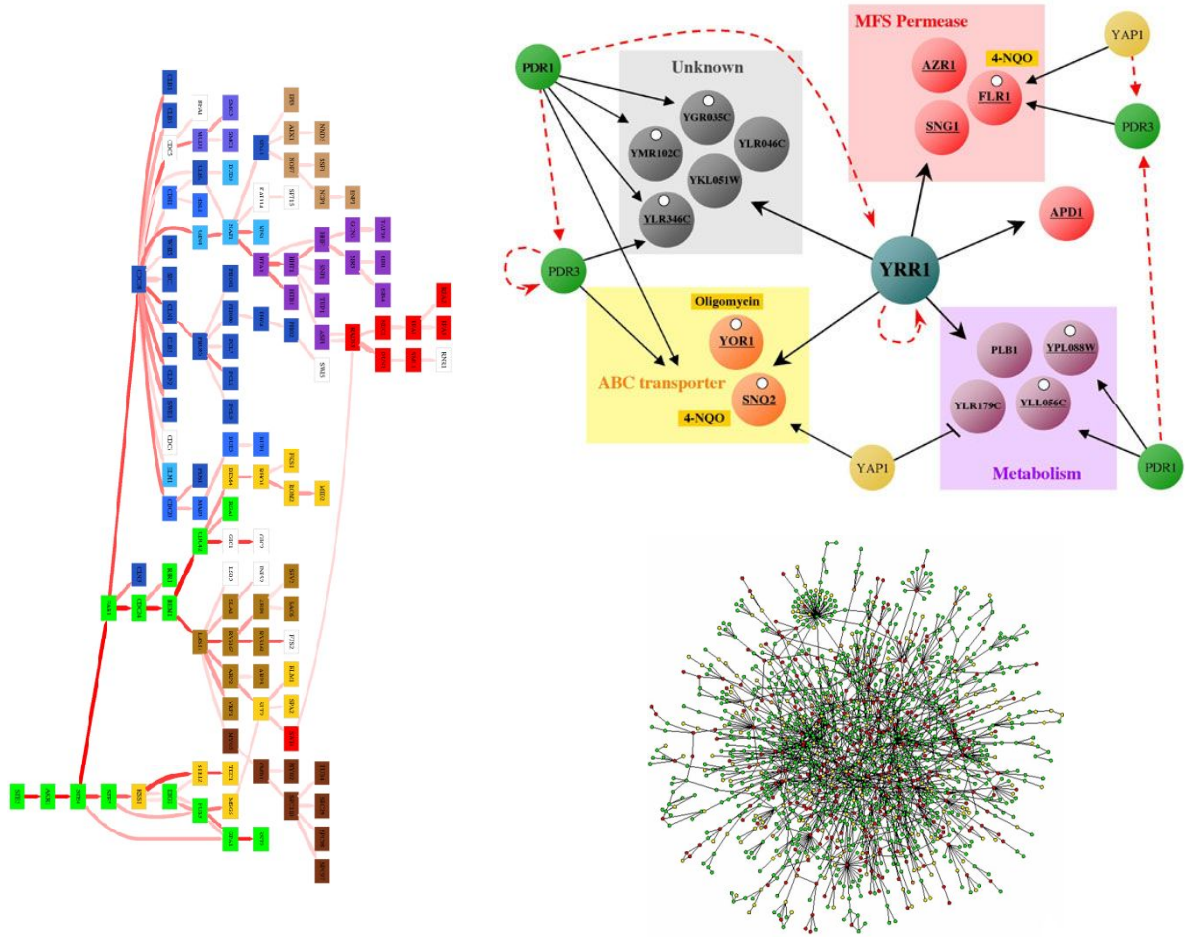


Différents types de contrôle

Figure 10.20 Control circuits are versatile and can be designed to allow positive or negative control of induction or repression.



Réseaux biologiques



Modèles graphiques et réseaux

- Très utilisés en biologie avec les grandes quantités de données disponibles
- Liens avec les télécom, la théorie des graphes et la physique statistique
- Représentation probabiliste des interactions entre composants biologiques
- Les problèmes étudiés vont de la structure globale aux propriétés locales

Méthodes et problèmes de réseaux

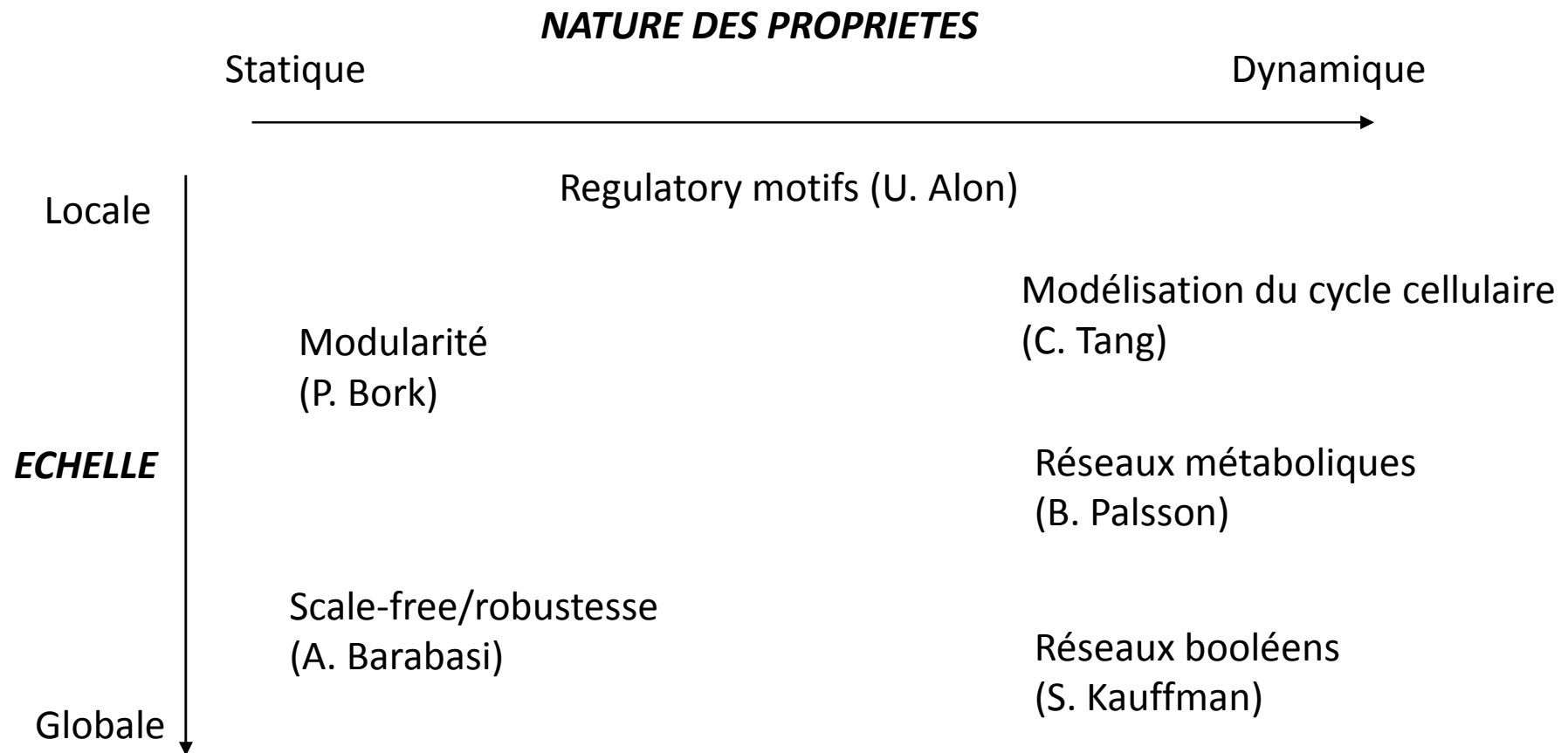
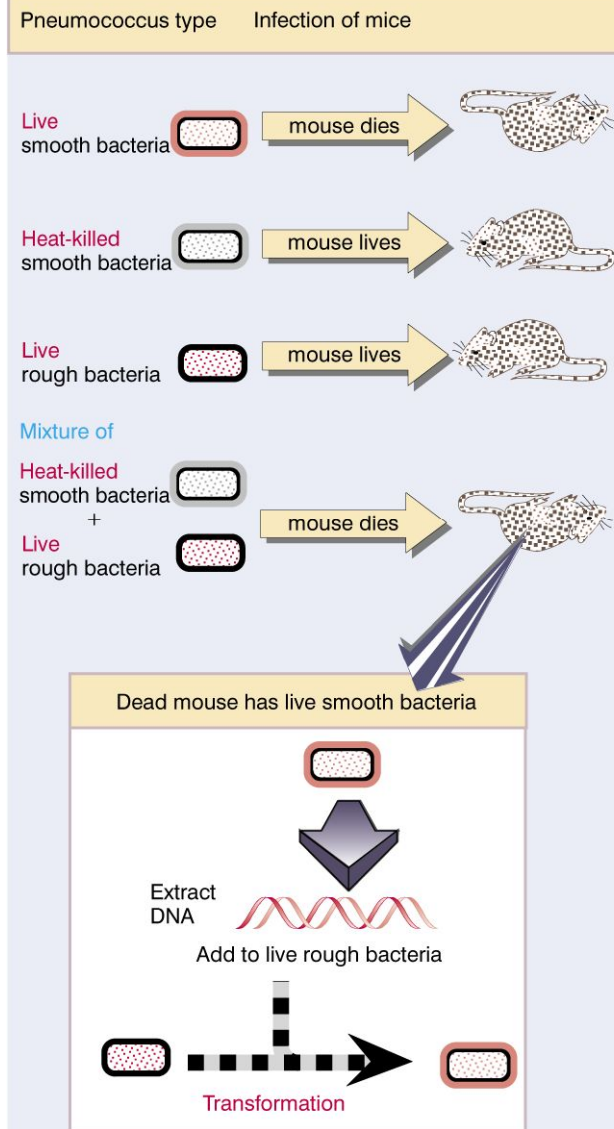


Figure 1.2 The transforming principle is DNA.



Nucleic Acids

- Nucleic Acids are polymers of nucleotides

Different nucleotides link together

Phosphate at 5' of one nucleotide links to 3' Carbon of another nucleotide

Called Phosphodiester bridge

- Common nucleic acids

RNA

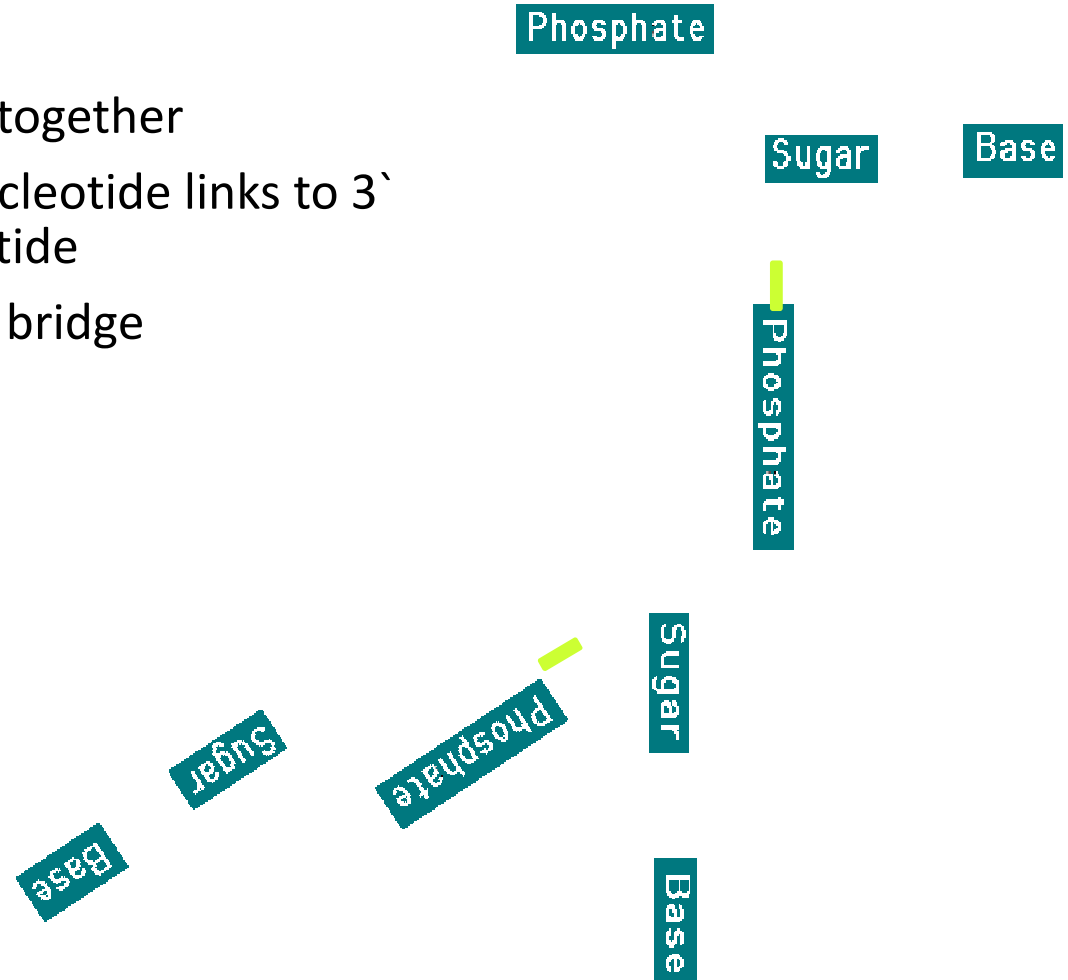
Ribonucleic acid

Sugar is ribose

DNA

Deoxyribonucleic acid

Sugar is deoxyribose



Nucleic Acids

- A sequence of 3 bases attracts a specific amino acid
 - AGC->Serine, AGA->Arginine etc.
 - Such a sequence is called a codon
 - Sequence of codons can assemble multiple amino acids into proteins
 - This is how protein structure is coded in nucleic acid
 - These proteins are manufactured during biosynthesis



Figure 9.3 Transcription takes place in a 'bubble', in which RNA is synthesized by base pairing with one strand of DNA in the transiently unwound region. As the bubble progresses, the DNA duplex reforms behind it, displacing the RNA in the form of a single polynucleotide chain.

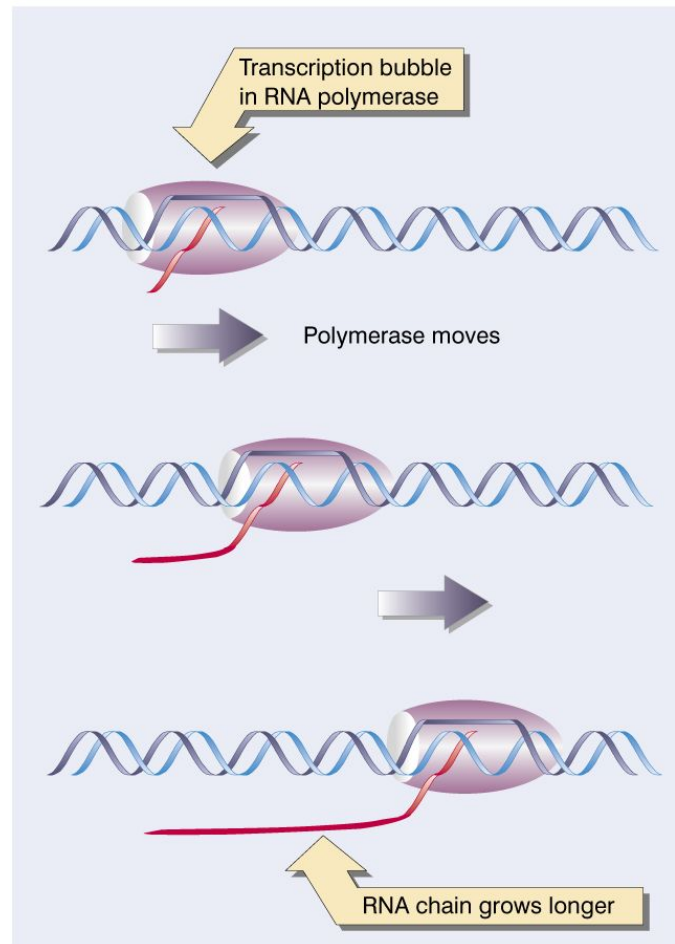
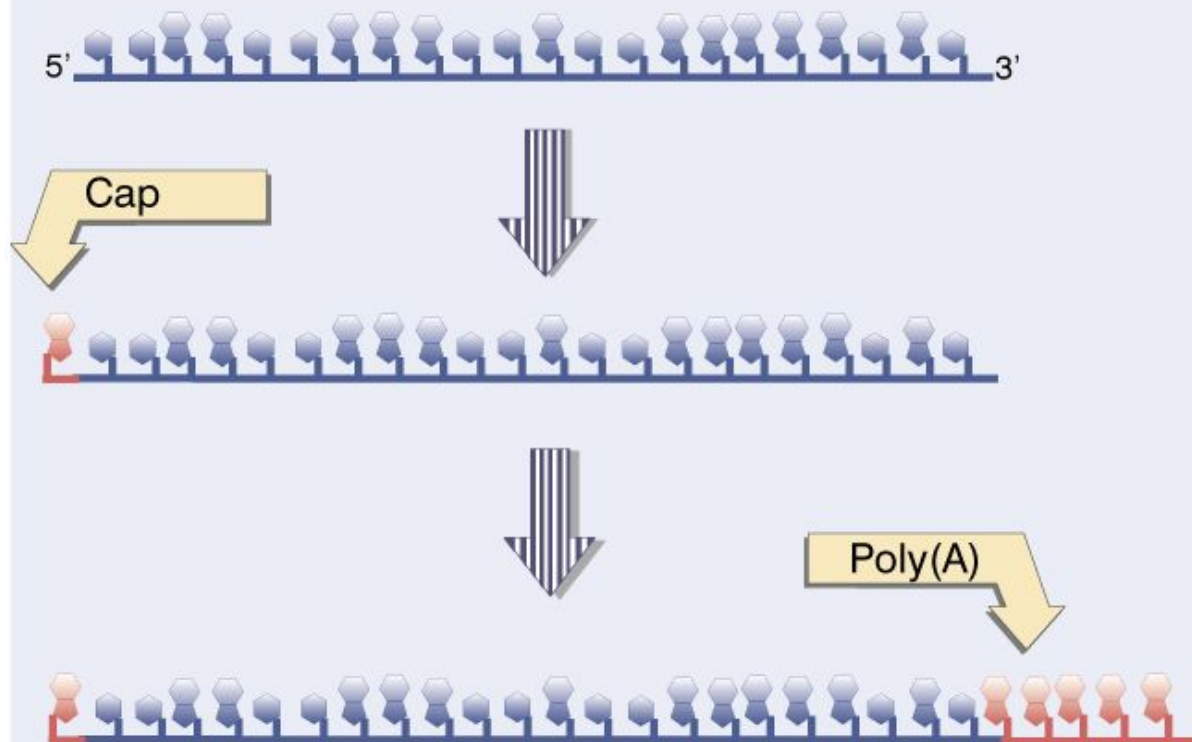


Figure 5.15 Eukaryotic mRNA is modified by addition of a cap to the 5' end and poly(A) to the 3' end.



Condensation of Amino Acids

- Two amino acids condense together forming a dipeptide (peptide bond)

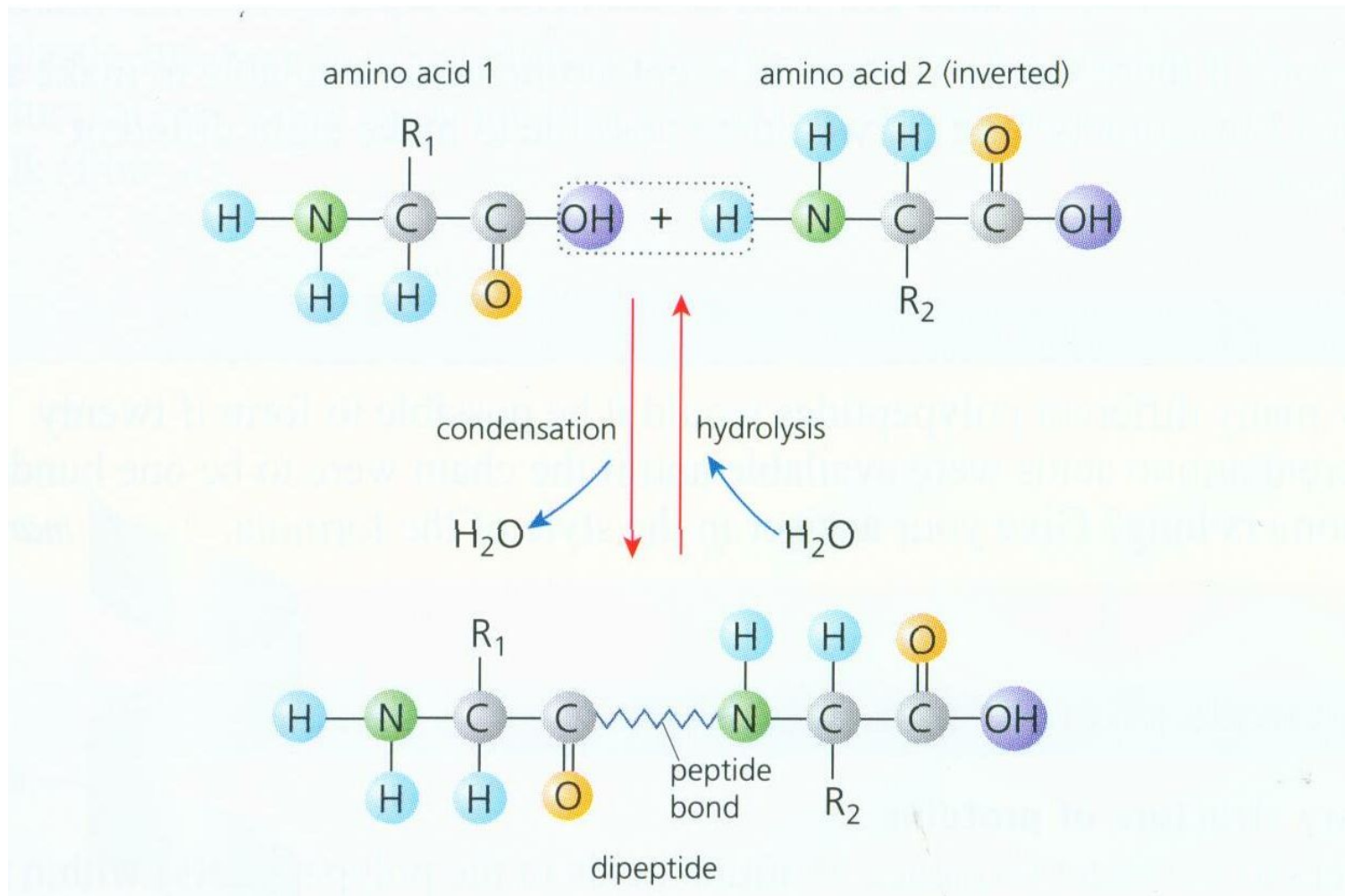


Figure 5.3 The tRNA cloverleaf has invariant and semi-invariant bases, and a conserved set of base pairing interactions.

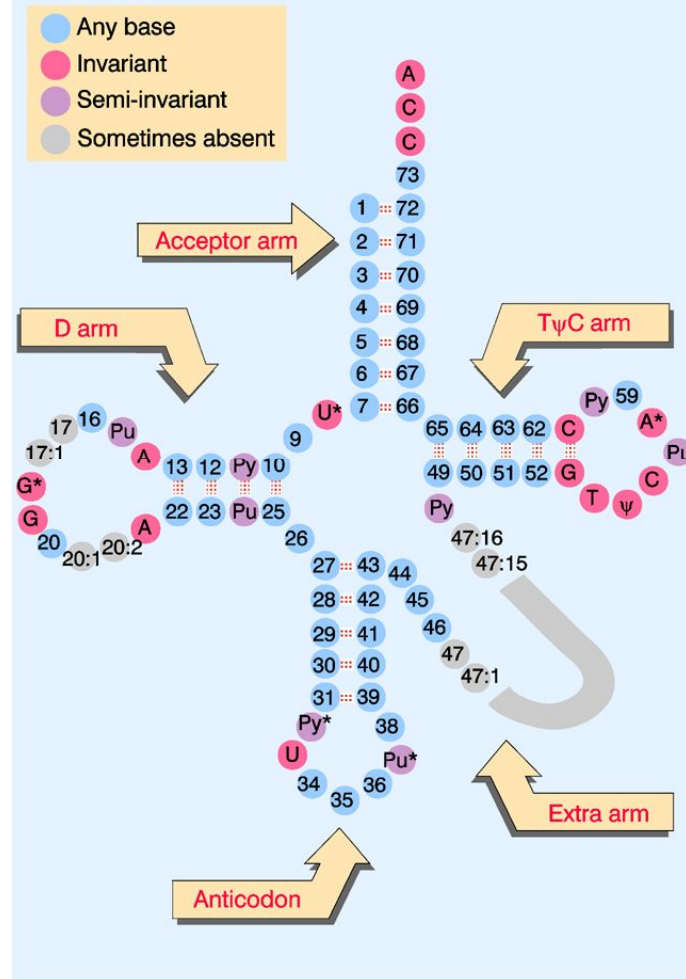


Figure 5.7 A ribosome consists of two subunits.

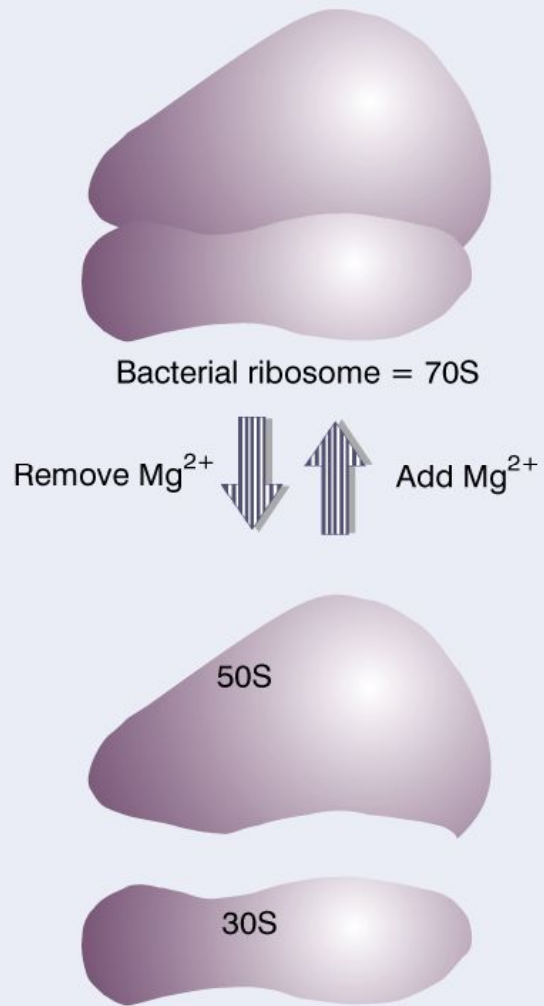


Figure 10.16 Lac repressor binds strongly and specifically to its operator, but is released by inducer. All equilibrium constants are in M^{-1} .

DNA	Repressor	Repressor + inducer
Operator	2×10^{13}	2×10^{10}
Other DNA	2×10^6	2×10^6
Specificity	10^7	10^4

Biological networks

- Regulation network
- Interaction network
- Signalling network
- Metabolic network
- Phylogenetic network
- Mixt networks...

