

## Unité de Biologie Moléculaire

Cours général (Transcription et Traduction)  
C. Jourlin-Castelli

Biologie Moléculaire Procaryote (Cours + TD)  
C.C. Zhang et C. Jourlin-Castelli

Biologie Moléculaire Eucaryote (Cours + TD)  
D. Aragnol et C. Maurel-Zaffran

## Le dogme central de la Biologie Moléculaire

Acide DéoxyriboNucléique (ADN)

↓ **Transcription**

Acide RiboNucléique (ARN)

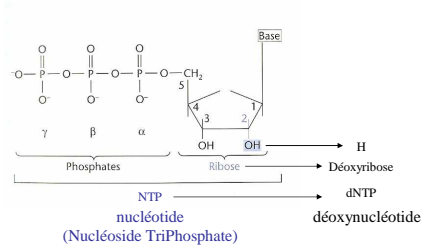
↓ **Traduction**

Protéine

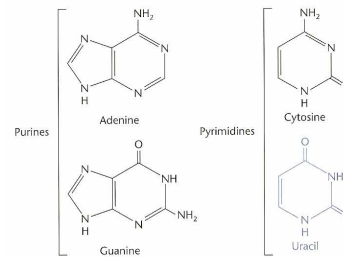
## L'ARN

ARN : Chaîne de nucléotides contenant du ribose

ADN : Chaîne de nucléotides contenant du déoxyribose



## L'ARN : les bases



Différence par rapport à l'ADN : l'uracile remplace la thymine

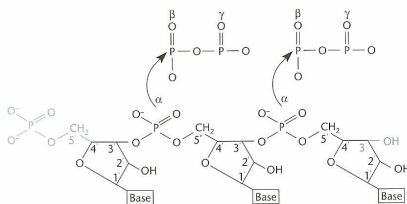
→ Appariement : A-U et G-C

## L'ARN : orientation 5' → 3'

ARN synthétisé et représenté dans le sens 5' → 3'

Extrémité 5' : phosphate

Extrémité 3' : hydroxyle



La chaîne (ou brin) d'ARN = enchaînement de **ribonucléotides** GMP, AMP, CMP et UMP reliés par des liaisons **phosphodiester** entre le carbone 3' d'un nucléotide et le carbone 5' du suivant.

## Les différents types d'ARN

Codants (traduits en protéines) :

Les ARN messagers : ARN<sub>m</sub>

→ Ne représentent que ~ 5% des ARNs de la cellule

Non codants :

- Les ARN ribosomiques : ARN<sub>r</sub>

- Les ARN de transfert : ARN<sub>t</sub>

} Classes les plus abondantes d'ARNs

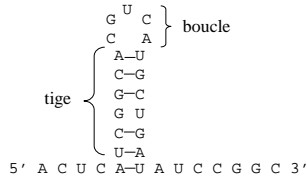
### L'ARN : structure

Structure primaire : séquence en nucléotide dans l'ARN

Structure secondaire :

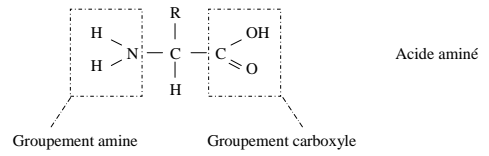
- synthétisé sous forme simple brin
- repliement sur lui-même (appariement des bases)

→ Région double brin : structure tige-boucle (ou épingle à cheveux)



### Les protéines

→ Chaînes d'acides aminés

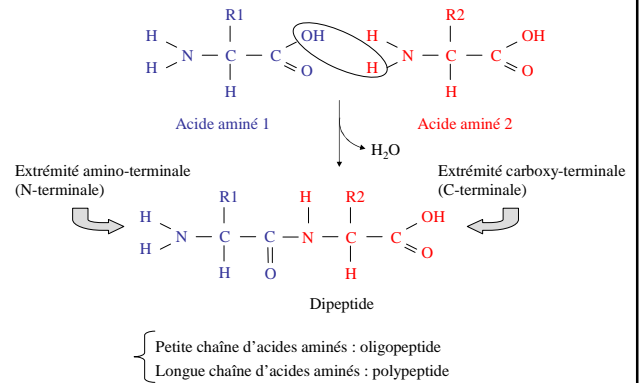


- R : chaîne latérale
- R diffère d'un acide aminé à l'autre
- 20 acides aminés

Acide aminé	Abréviation	
	3 lettres	1 lettre
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Acide Aspartique	Asp	D
Cystéine	Cys	C
Acide Glutamique	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Méthionine	Met	M
Phénylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Thréonine	Thr	T
Tryptophane	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

### La structure des protéines

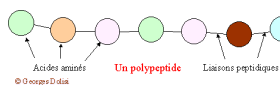
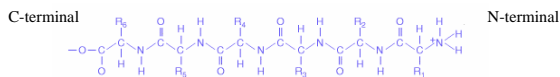
La liaison peptidique



### La structure des protéines

La structure primaire

→ la séquence en acides aminés

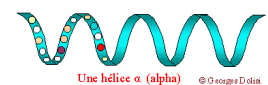
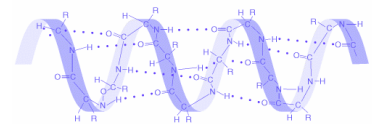
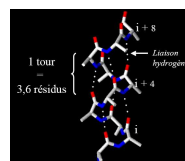


### La structure des protéines

La structure secondaire

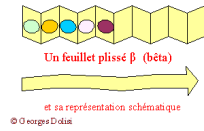
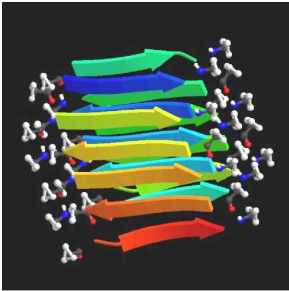
→ liaison hydrogène entre les acides aminés de la chaîne polypeptidique  
→ 2 formes principales : hélice  $\alpha$  et feuillet  $\beta$

→ L'hélice  $\alpha$



## Structure secondaire des protéines

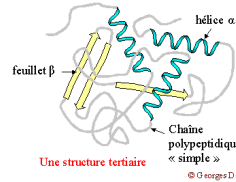
### Le feuillet $\beta$



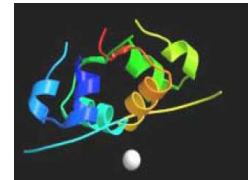
© Georges D'Almeida

## La structure tertiaire

→ repliement de la prot\u00e9ine sur elle-m\u00eame



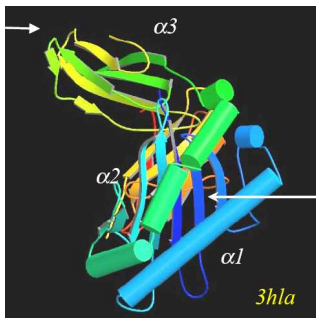
© Georges D'Almeida



## La structure quaternaire

→ protéines form\u00e9es de plusieurs cha\u00eenes polypeptidiques  
→ multim\u00e8res

Homomultim\u00e8res : un polypeptide en plusieurs exemplaires  
H\u00e9teromultim\u00e8res : diff\u00e9rents polypeptides (\u00e9galement appel\u00e9s sous-unit\u00e9s)



## La transcription

La s\u00e9quence informative de l'ADN, pour \u00eatre convertie en une s\u00e9quence prot\u00e9ique, doit \u00eatre r\u00e9\u00e9crite (transcrite) en une s\u00e9quence d'ARN.

Transcription = Processus de synth\u00e8se d'ARN \u00e0 partir d'une matrice ADN

Cha\u00eene d'ARN synth\u00e9tis\u00e9e:

identique \u00e0 un brin de l'ADN (brin codant ou brin compl\u00e9mentaire)  
compl\u00e9mentaire de l'autre brin d'ADN (brin matrice)

ADN  $\left\{ \begin{array}{l} 5' \text{ ATTACGACCTACGCAT } 3' \text{ Brin codant} \\ \text{ (Brin compl\u00e9mentaire)} \\ 3' \text{ TAATGCTGGATGCGTA } 5' \text{ Brin matrice} \end{array} \right.$

ARN  $5' \text{ AUUACGACCUACGCAU } 3'$

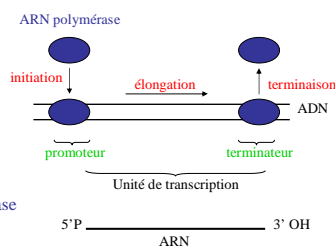
## La transcription

Ne concerne qu'une portion de l'ADN

D\u00e9marre \u00e0 un « promoteur »  
S'arr\u00eate \u00e0 un « terminateur » (pas chez les eucaryotes)

3 grandes \u00e9tapes :

initiation (d\u00e9but)  
\u00e9longation  
terminaison



Enzyme responsable : ARN polym\u00e9rase

## Les ARN polym\u00e9rases

Synth\u00e9tisent l'ARN dans le sens 5' vers 3'

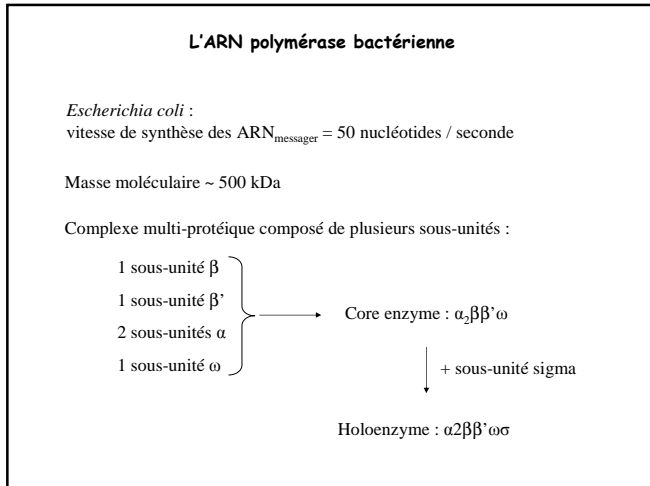
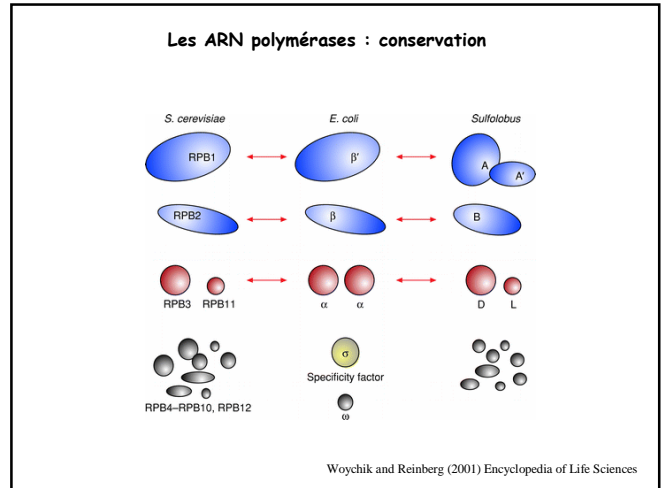
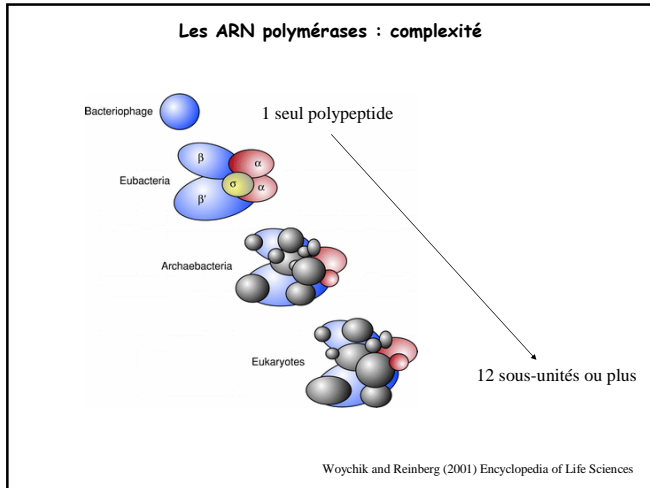
Ne n\u00e9cessitent pas d'amorces

Se d\u00e9placent le long du brin matrice dans le sens 3' vers 5'

Structure en « pince de crabe »

1 seule ARN polym\u00e9rase chez les procaryotes

3 ARN polym\u00e9rases chez les eucaryotes : I, II et III



### L'ARN polymérase bactérienne : les sous-unités $\beta\beta'$

$\beta$  : 1342 résidus et  $\beta'$  : 1407 résidus (*Escherichia coli*)

Contiennent le site catalytique de l'enzyme

Interagissent directement avec l'ADN

Portent les sites de fixation des nucléosides triphosphates (NTP)

Nécessaires au démarrage de la transcription et à la formation de la chaîne ARN

Cibles de certains antibiotiques

### L'ARN polymérase bactérienne : la sous-unité $\alpha$

329 résidus (*Escherichia coli*)

Possède 2 domaines capables de se replier indépendamment et liés par un peptide d'environ 20 résidus

domaine amino-terminal :  $\alpha$ NTD (résidus 1 à 235)

- dimérisation des sous-unités  $\alpha$
- assemblage avec les sous-unités  $\beta$  et  $\beta'$

domaine carboxy-terminal :  $\alpha$ CTD (résidus 250 à 329)

- module de fixation à l'ADN (séquence « UP »)

### L'ARN polymérase bactérienne : la sous-unité $\omega$

91 résidus (*Escherichia coli*)

Apparemment pas de rôle direct dans la transcription

Agirait comme protéine chaperonne

Faciliterait le repliement correct de la sous unité  $\beta'$

### L'ARN polymérase bactérienne : la sous-unité $\sigma$

Taille variable : 20 à 70 kDa

Sous-unité dissociable de l'ARN polymérase

S'associe de manière réversible au « core » enzyme : holoenzyme ( $\alpha 2\beta\beta' \sigma$ )

Permet à l'ARN polymérase d'initier la transcription

Plusieurs facteurs  $\sigma$  dans une même bactérie :

- 1 facteur  $\sigma$  principal responsable de la majorité de la transcription
- d'autres facteurs  $\sigma$  dits « alternatifs » ayant des fonctions spécialisées

Chez *Escherichia coli*, 7 facteurs  $\sigma$  différents :

$\sigma^{70}$ (RpoD)	$\sigma^{54}$ (RpoN)	$\sigma^{28}$ (FlhA)
« principal »	$\sigma^{38}$ (RpoS)	$\sigma^{24}$ (RpoE)
	$\sigma^{32}$ (RpoH)	$\sigma^{19}$ (FecI)
	« alternatifs »	

### Le cycle de transcription

Transcription = réaction très conservée entre procaryotes et eucaryotes

Fixation de l'ARN polymérase sur les éléments du promoteur

→ Formation d'un complexe stable : **complexe fermé**

Enroulement de l'ADN correspondant au promoteur autour de l'ARN polymérase

→ **Complexe intermédiaire**

Séparation des brins d'ADN : formation de la « bulle » de transcription

→ **Complexe ouvert**

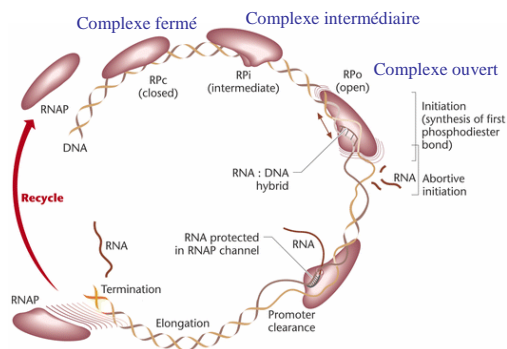
Initiation de la synthèse de l'ARN en présence de NTP

→ Formation d'un hybride ARN-ADN

Libération du promoteur et **élongation** (départ du facteur sigma chez les procaryotes) : progression de l'ARN polymérase le long de l'ADN

**Terminaison** et dissociation du complexe ADN-ARN-ARN polymérase

### Le cycle de transcription



Drapkin and Reinberg (2002) Encyclopedia of Life Sciences

### Initiation de la transcription : les éléments sur l'ADN

+1 = nucléotide où commence la transcription

→ site de démarrage de la transcription

Promoteur :

signal pour initier la transcription

localisé en amont (avant) du site +1

n'est pas transcrit

### Les séquences consensus du promoteur chez les procaryotes

Comparaison de plusieurs promoteurs

→ 2 régions conservées : région -35 et région -10

		Site de démarrage →					
GTGCGTG	TTGACT	ATTTTA	CCTCTGGCGGT	GATAAATGG	TTGCAT	TGTACTAAGGA	$\lambda P_g$
GGCGGT	TTGACA	TAAATA	CCACTGGCGGT	GATACTGA	GCACAT	TCAGCAGGACG	$\lambda P_L$
TGAGCTG	TTGACA	ATTAAT	CATCGAACTAG	TAACTAG	TACGCA	AGTTCACTGAA	<i>trp</i>
CCCAAGGC	TTTACA	CTTTAT	GCTTCCGGCTCG	TATGTTGT	GTGGAA	ATTGTGAGCGG	<i>lac</i>
CCCAAGGC	TTTACA	CTTTAT	GCTTCCGGCTCG	TATAATGT	GTGGAA	ATTGTGAGCGG	<i>lacUV5</i>
ATCCTAC	CTGACG	CTTTTT	ATCGCAACTCTC	TACTGT	TTCAT	ATCCGTTTTTTT	<i>araBAD</i>
TTTCCTC	TTGTCG	AGGCCGG	AATAACTCCC	TATAAT	GGCCACC	CTGACACGGAA	<i>rnmA1</i>
TAAATGC	TTGACT	CTGTAG	CGGGAAGGCG	TATTATGC	ACACC	CGCCCGCTGA	<i>rnmA2</i>
TCCATGT	CACACT	TTTCGCATCTT	TGTTATGC	TATGGT	TAT	TTCAT	TACCATAAGCC
TTATTC	ATGTCA	CACATT	TCGCATCTTGT	TATGCT	AT	GTTA	TTTCATACCAT
Consensus sequence:		TTGACA		TATAAT		+1	
		-35 region		-10 region			
		35 pb en amont du +1		10 pb en amont du +1			

### Efficacité des promoteurs

Aucun promoteur naturel contenant tous les éléments parfaits

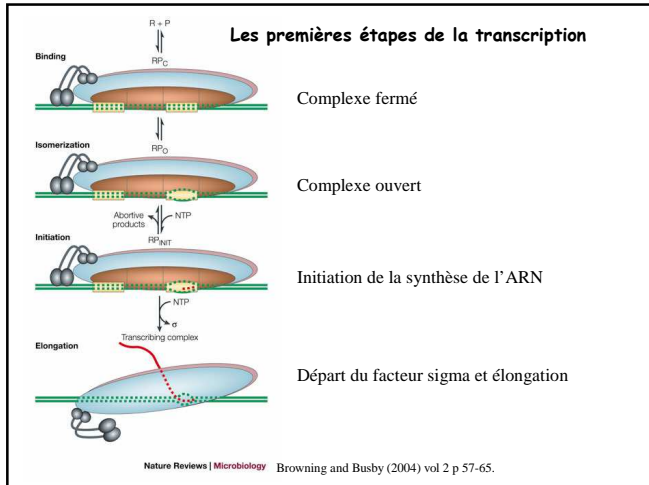
Promoteurs avec séquences proches des consensus = très efficaces

→ Promoteurs « forts »

Promoteurs avec séquences éloignées des consensus = peu efficaces

→ Promoteurs « faibles »





### Le complexe ouvert

Après la fixation de l'ARN polymérase :

- Séparation (dénaturation) des deux brins d'ADN
- Région dénaturée (environ de la position -12 à +3) inclut la région -10
- Formation d'une bulle de transcription

Mécanisme : isomérisation

- processus mal compris
- le brin non-matrice est lié au domaine 2 du facteur  $\sigma$  de l'ARN polymérase (la région 2.3)

### Synthèse de transcrits « avortés »

Avant passage à l'étape d'élongation :

- synthèse et relargage par l'ARN polymérase d'un ensemble de transcrits « avortés »
- longueur de ces transcrits : souvent entre 2 et 12 nucléotides, mais parfois jusqu'à 15 ou 17 nucléotides
- pas de dissociation de l'ARN polymérase pendant cette initiation « avortée »

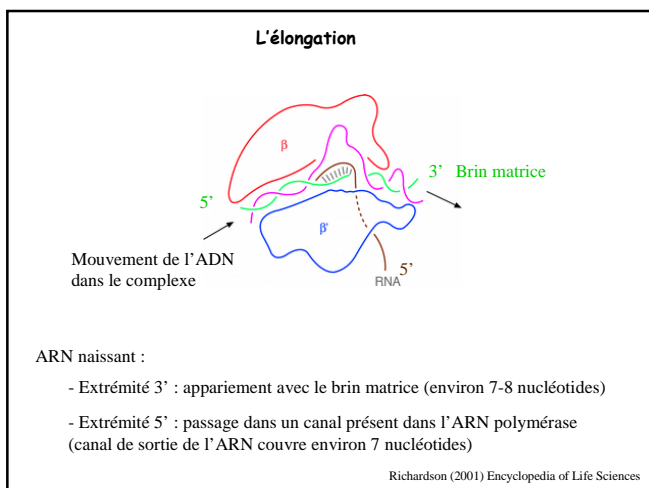
### Libération du promoteur

Dernière étape de l'initiation de la transcription

ARN polymérase se désengage du promoteur

Relargage du facteur sigma chez les procaryotes

Démarrage de la phase d'élongation : progression de l'ARN polymérase le long de l'ADN



### Correction des erreurs

Incorporation d'un mauvais nucléotide (cf mauvais appariement)

- blocage de l'élongation (pause)
- activation d'une activité nucléase de l'ARN polymérase

Deux protéines facilitant la reconnaissance de l'extrémité 3' du transcrit mal apparié au brin d'ADN matrice : GreA et GreB

- fixation sur l'ARN polymérase et contact avec l'ARN naissant
- activation de l'activité nucléase de l'ARN polymérase
- clivage de 2-3 nucléotides (pour GreA) jusqu'à 18 nucléotides (pour GreB)

### Le facteur NusA

Facteur d'élongation conservé chez les procaryotes

Présente des homologies structurales avec le facteur  $\sigma^{70}$

Effet antagoniste en fonction du contexte (structure ARN, protéines associées) :

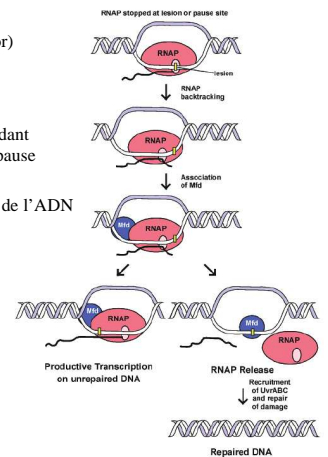
- Stimulation de l'élongation, rôle antiterminateur à certains sites (e.g. *nut*) notamment au niveau de terminateurs rho-dépendants
- Stimulation de la pause de la RNAP à certains sites (e.g. opérons *his* et *trp*) et de la terminaison rho-indépendante

### Le facteur Mfd

(Transcription Repair coupling factor)

Permet de réactiver ou de recycler pendant l'élongation une ARN polymérase en pause

Recrutement du système de réparation de l'ADN



### La terminaison de la transcription (chez les procaryotes)

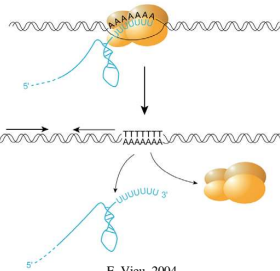
Deux mécanismes :

- Terminaison dépendante d'un facteur intrinsèque :
  - Formation dans le transcrite d'une tige boucle riche en G et C, suivie d'une série de U (= terminateur)
  - Pause de l'ARN polymérase puis décrochage → arrêt de la transcription
- Terminaison dépendante d'un facteur extrinsèque :
  - Fixation de la protéine Rho sur une séquence spécifique du transcrite
  - Progression de Rho le long du transcrite
  - Contact entre Rho et ARN polymérase → arrêt de la transcription

### La terminaison de la transcription dépendante d'un facteur intrinsèque (Rho-indépendante)



Séquences de terminateurs



Formation d'une tige boucle

Déstabilisation du complexe ADN-ARN-RNAP

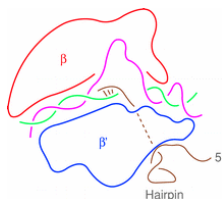
Arrêt de la transcription

E. Vieu, 2004

### La terminaison de la transcription dépendante d'un facteur intrinsèque (Rho-indépendante)

Formation de la structure tige boucle

- structure émerge du canal de sortie
- interaction faible entre extrémité 3' de l'ARN et le brin ADN matrice



Richardson (2001) Encyclopedia of Life Sciences

### La terminaison de la transcription dépendante d'un facteur extrinsèque (Rho-dépendante)

Protéine Rho : 6 sous unités identiques formant un anneau

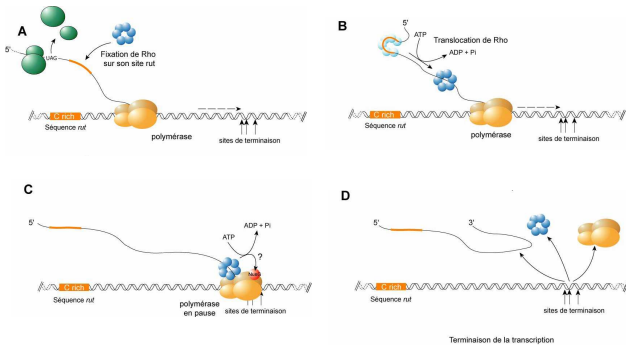
Fixation de Rho sur un site appelé *rut* (rho utilization site)

→ site *rut* : région de 40 à 60 nucléotides sur le transcrite, riche en C, simple brin

Progression de Rho le long du transcrite (nécessite hydrolyse ATP)

Pas de point précis de terminaison : un site où l'ARN polymérase pause

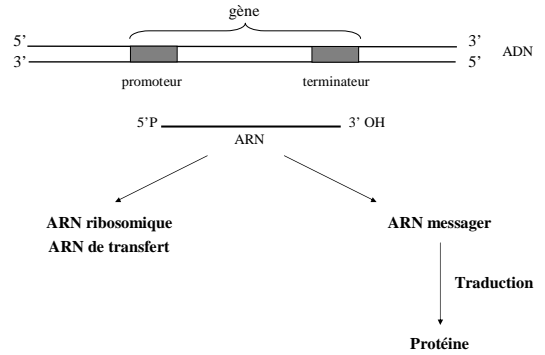
### La terminaison de la transcription dépendante d'un facteur extrinsèque (Rho-dépendante)



E. Vieu, 2004

### La notion de gène

Gène : région d'ADN portant l'information pour la synthèse d'un ARN ou d'une protéine



### Le symbolisme chez les procaryotes

→ trois lettres (se rapportant en général à la fonction du gène)

Gène : italique ou souligné  
Ex : *hisA* ou hisA

Protéine : 1<sup>ère</sup> lettre en majuscule  
Ex : HisA

### La notion d'opéron chez les procaryotes

Opéron : région d'ADN transcrite en un seul ARN<sub>m</sub> mais contenant l'information nécessaire à la synthèse de plusieurs protéines

