

# Transcrição e Processamento de RNA

QBQ0104



INSTITUTO DE QUÍMICA  
USP

Bibliografia:

**Capítulo 8: Genome structure, chromatin, and the nucleosome**

**Capítulo 13: Mechanisms of transcription**

**Capítulo 14: RNA splicing**

Molecular Biology of the Gene. J.D. Watson, T.A. Baker, S. P. Bell, A. Gann, M. Levine, R. Losick. CSH Press, 7<sup>th</sup> edition (2013).

# O que é o genoma?

***Não é sinônimo de “código genético”...***

- Conjunto de **toda** a informação genética do organismo
  - Cromossomos
  - Plasmídeos (bactérias)
  - DNA das mitocôndrias e cloroplastos
  - *Inclui **genes**, regiões não codificadoras de proteínas, regiões regulatórias e mesmo aquelas que não se sabe a função*

O que é um gene do ponto de vista molecular?

**Um gene é um segmento de DNA que contém a informação para um produto final, que pode ser RNA ou proteína**

# Um gene corresponde a um transcrito de RNA

Região regulatória



Região não traduzida 5'

5' UTR



3' UTR

RNA

Região não traduzida 3'

Length of RNA defines region of gene

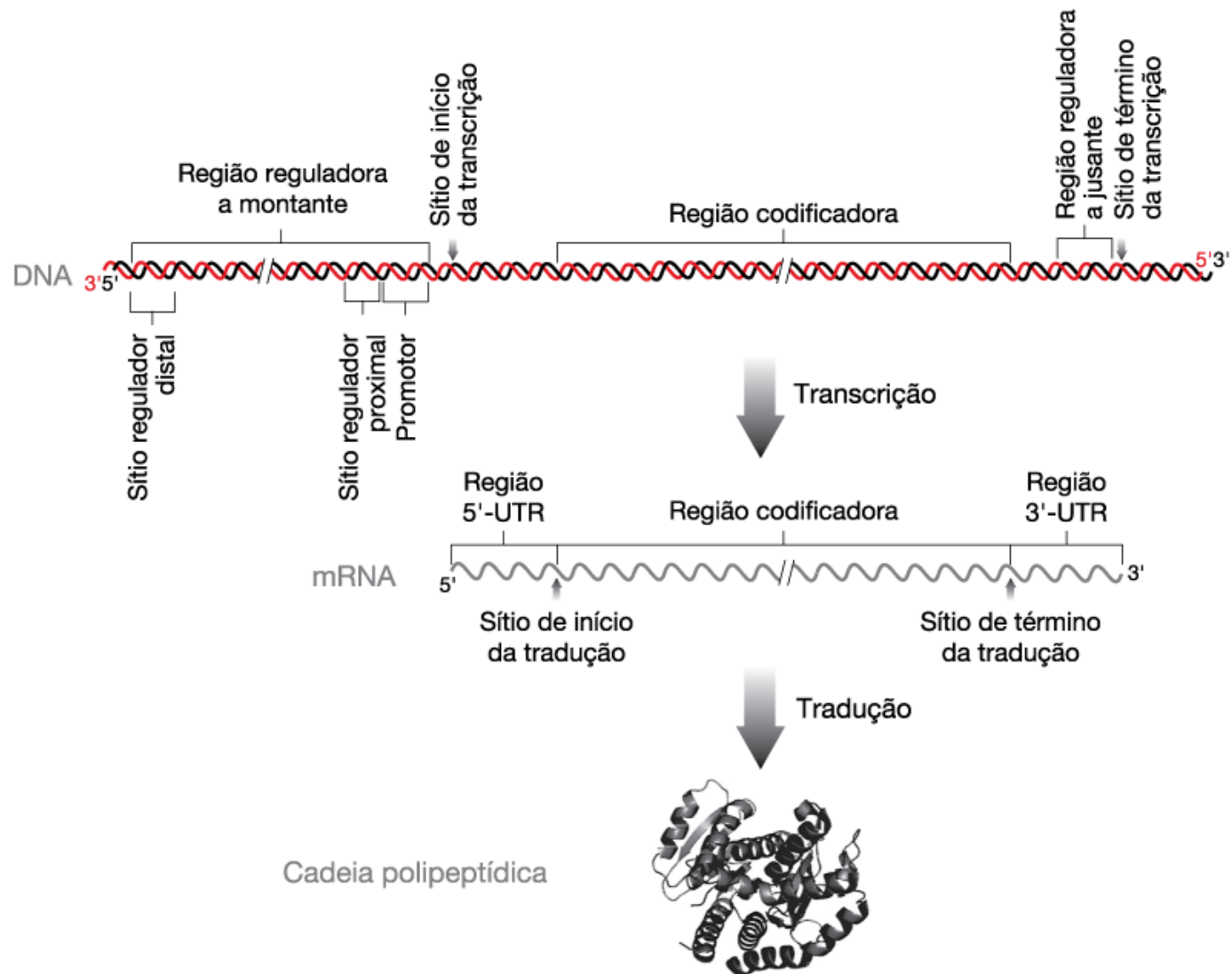


Protein defines coding region

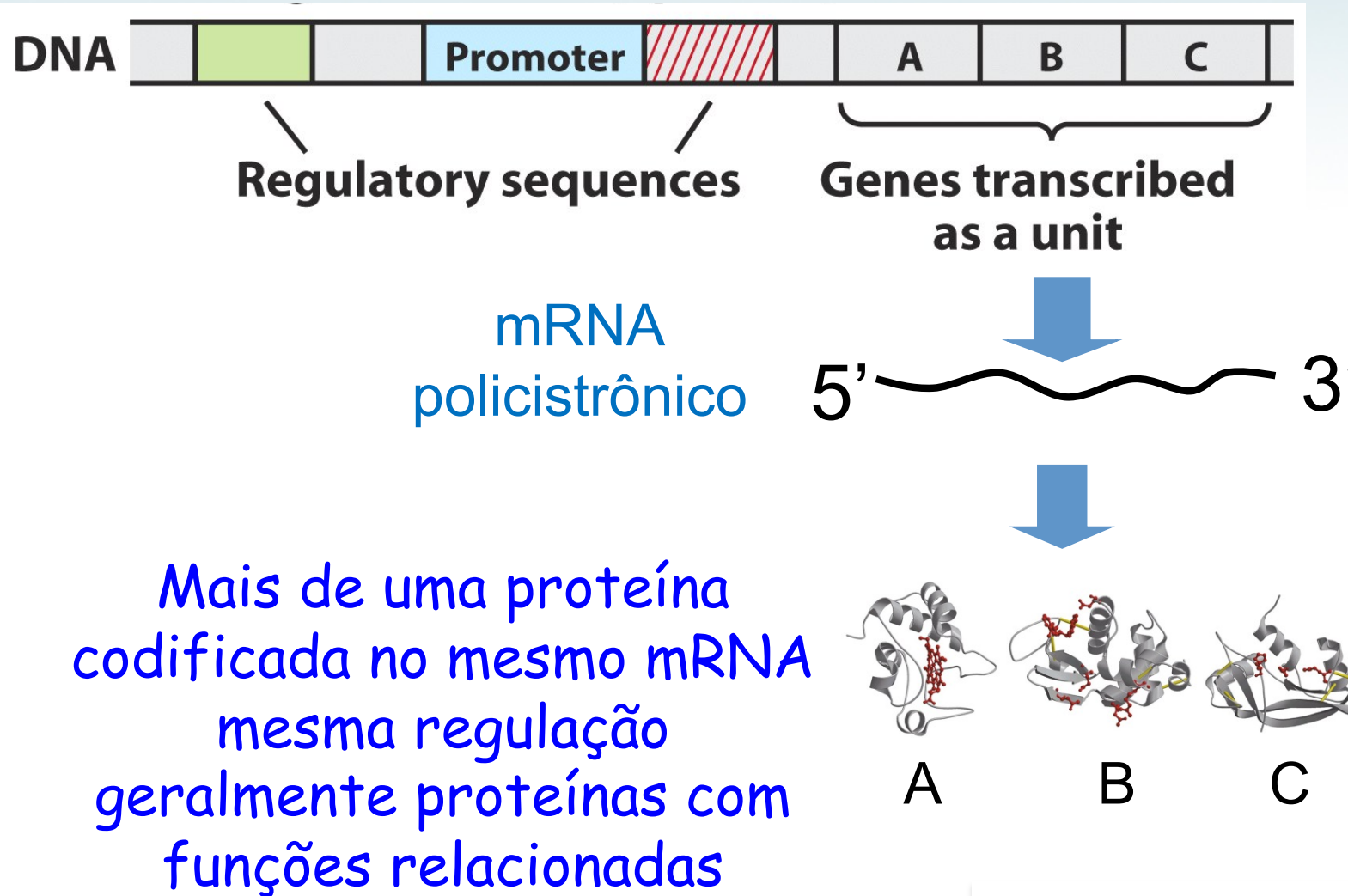
# Do que é composto um gene?

- Sequências codificadoras
  - informação do produto final (RNA ou proteína); exons (em eucariotos)
- Sequências não codificadoras
  - sequências não traduzidas (UTRs); introns (em eucariotos)
- Sequências regulatórias
  - indicam quando, onde e como o gene deve ser expresso

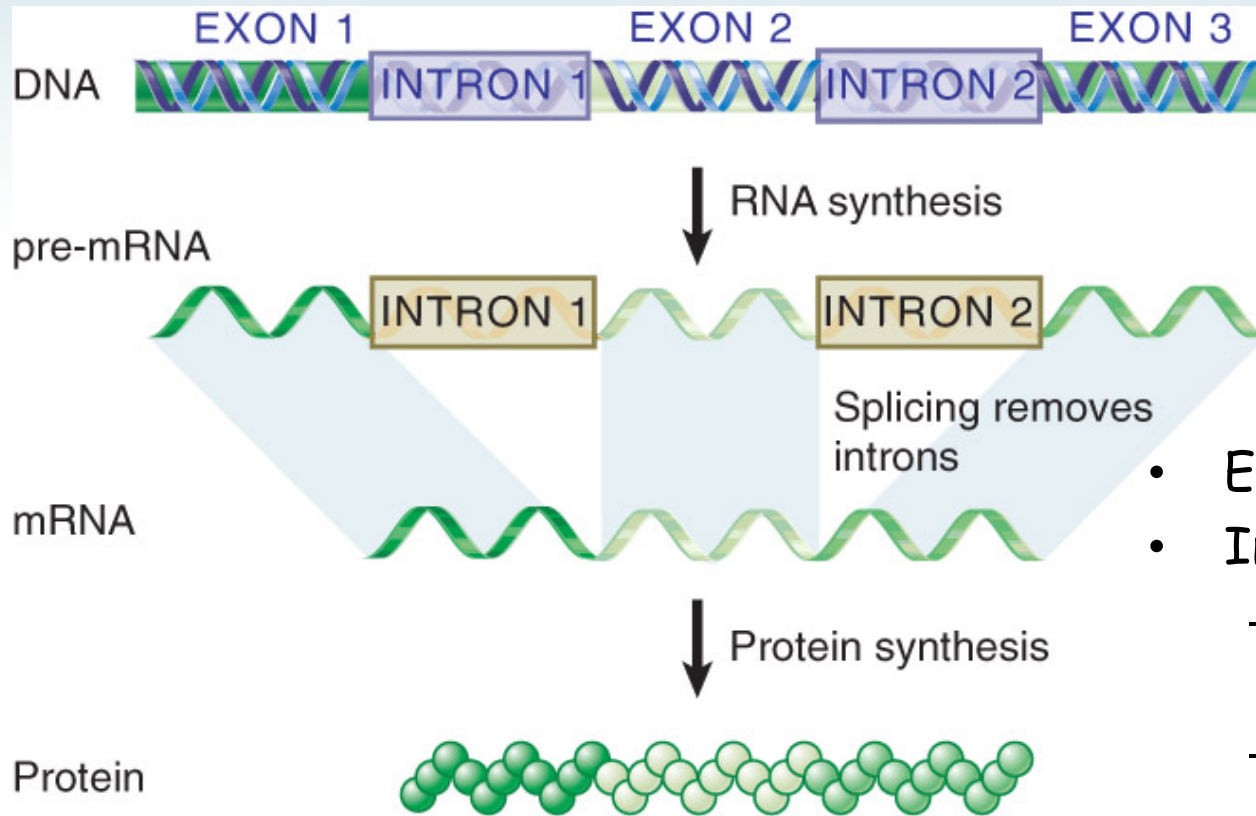
# Gene Procariótico



# Operon Bacteriano



# Eucariotos possuem genes interrompidos por introns



- Exons: regiões codificadoras
- Introns (intragenic regions):
  - presentes no DNA, mas não no mRNA
  - removidos do RNA precursor (splicing)

Length of precursor RNA (not mRNA) defines region of gene

Individual coding regions are separated in gene



# Replicação x Transcrição: Similaridades



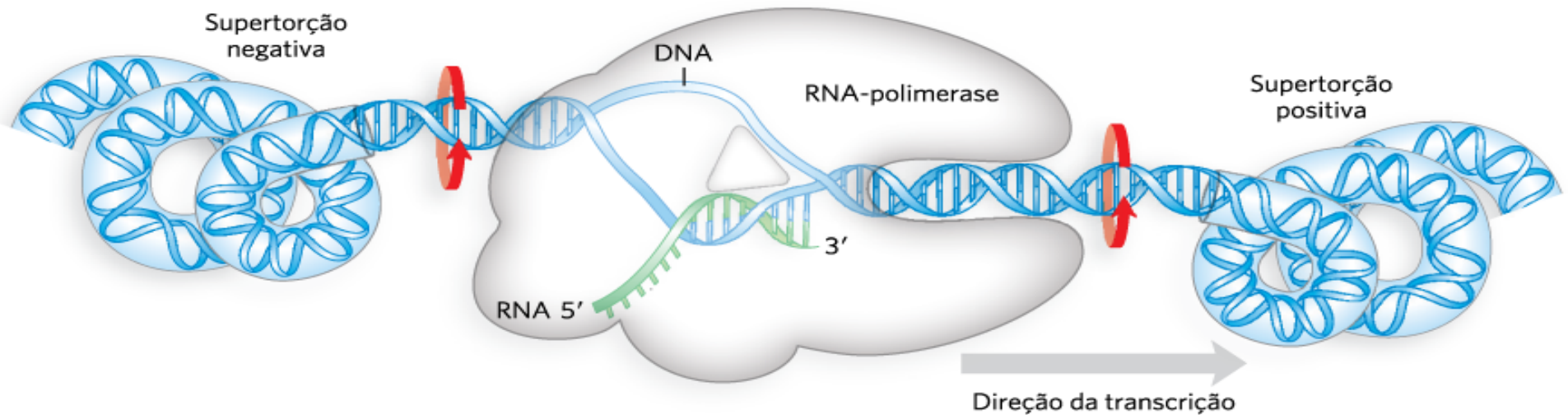
- Ambos sintetizam uma fita de polinucleotídeos complementar a uma fita de DNA molde antiparalela
- Ambos usam nucleotídeos trifosfatados como precursores
- A química da síntese é a mesma que ocorre sempre no **sentido 5'-3'**

# Replicação x Transcrição: diferenças



- Os substratos são **ribonucleosídeos** trifosfatados (**NTPs**) em vez de **desoxirribonucleosídeos** trifosfatados (**dNTPs**)
- A reação é catalisada por uma RNA polimerase dependente de DNA
- RNA polimerases não precisam de um primer
- A transcrição ocorre em uma bolha

# Transcrição

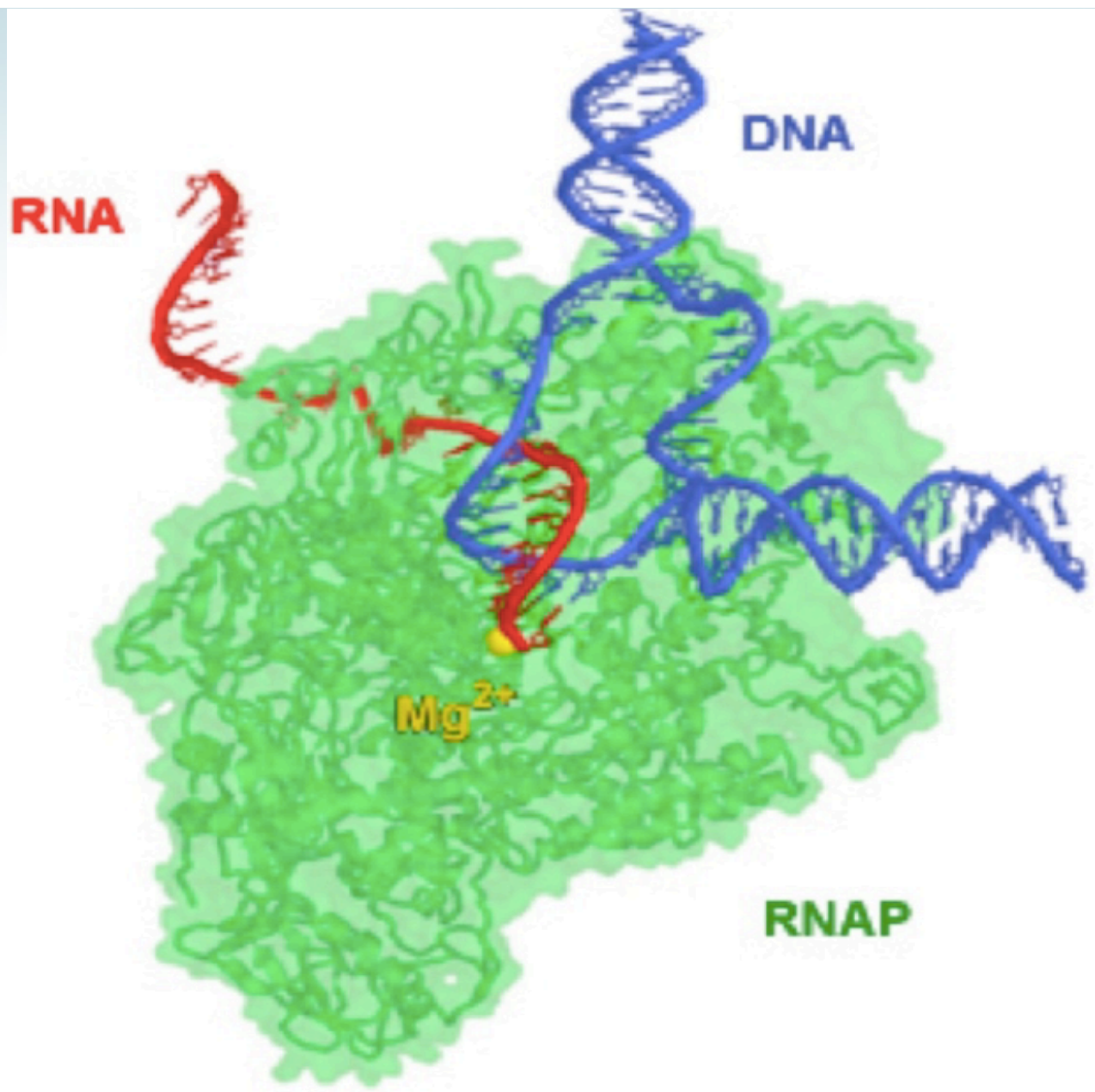


Um trecho de ~17 pares de base do DNA é desenrolado para formação da bolha de transcrição

# Replicação x Transcrição: Diferenças



- O transcrito de RNA não permanece associado à fita molde de DNA



# Replicação x Transcrição: Diferenças



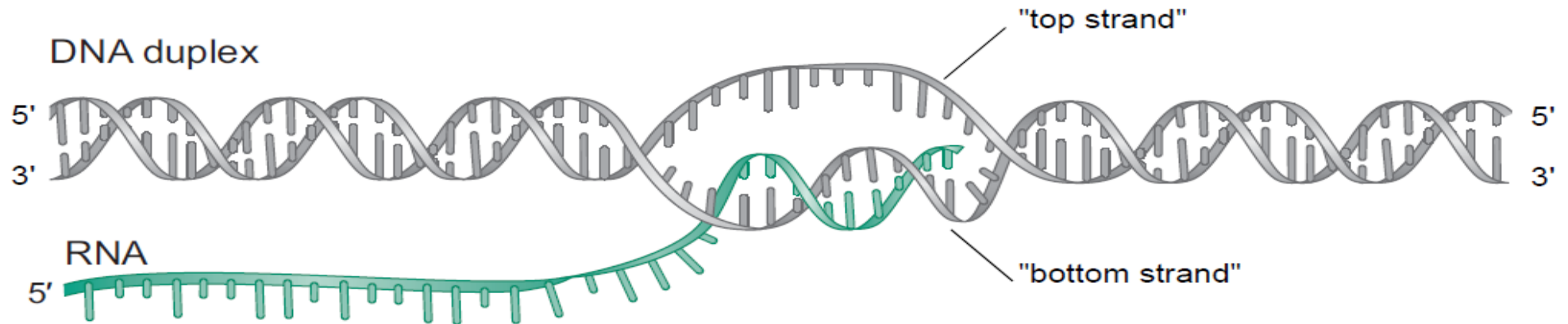
- O transcrito de RNA não permanece associado à fita molde de DNA
- Processo menos preciso que a replicação (incorporações de bases erradas =  $1 \times 10^4$ )

# Replicação x Transcrição: Diferenças



- O transcrito de RNA não permanece associado à fita molde de DNA
- Processo menos preciso que a replicação (incorporações de bases erradas =  $1 \times 10^4$ )
- **A transcrição é assimétrica**

# A Transcrição é assimétrica



- Apenas 1 das fitas de DNA é copiada
- Por isso uma das fitas de DNA é denominada de **fita codificadora** e a outra de **fita molde**

Qual é qual?



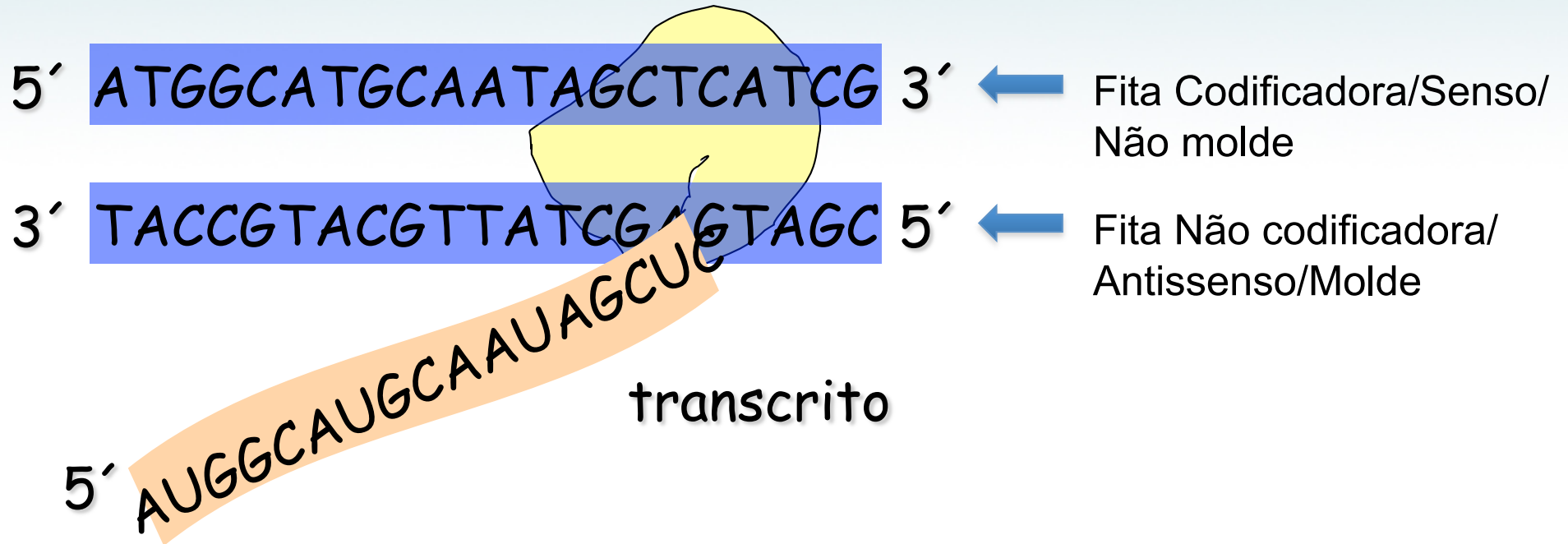
# A Transcrição é assimétrica

fita codificadora



- **Fita codificadora:** A fita de DNA que tem a mesma sequência do mRNA (exceto T vs U) e é relacionada pelo código genético à sequência proteica codificada pelo gene.
- **Fita molde:** A fita de DNA que tem sequência complementar a do mRNA

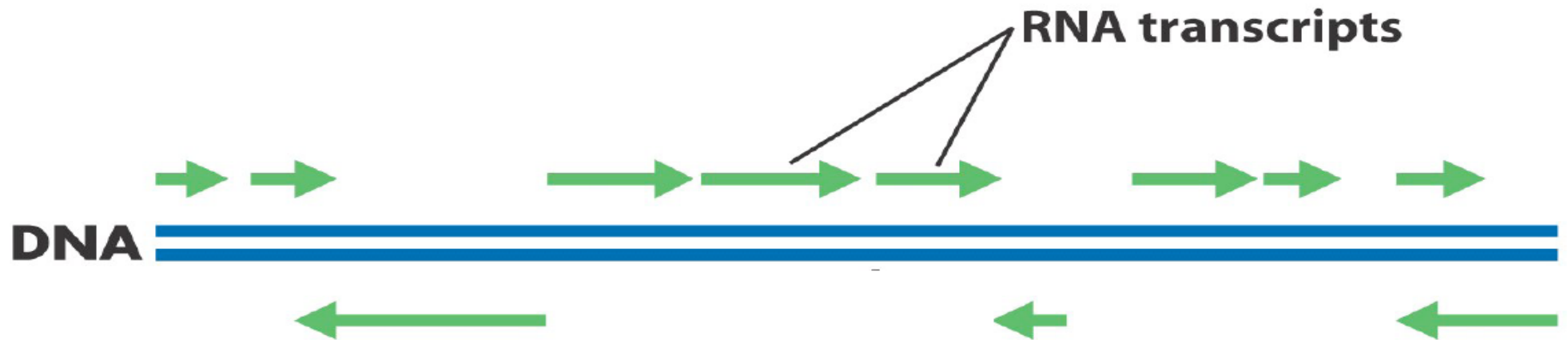
# Transcrição



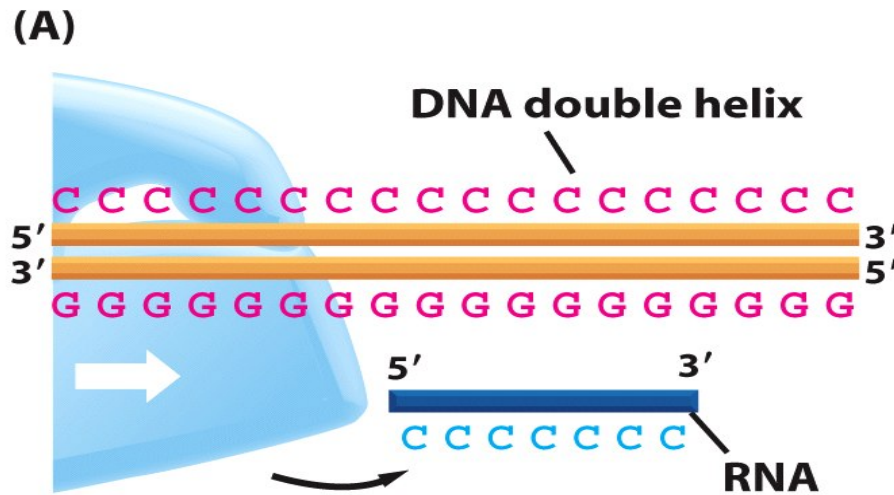
5' → 3'  
Direção de síntese da RNA polimerase

Notar a polaridade das espécies (fitas de DNA e o RNA)

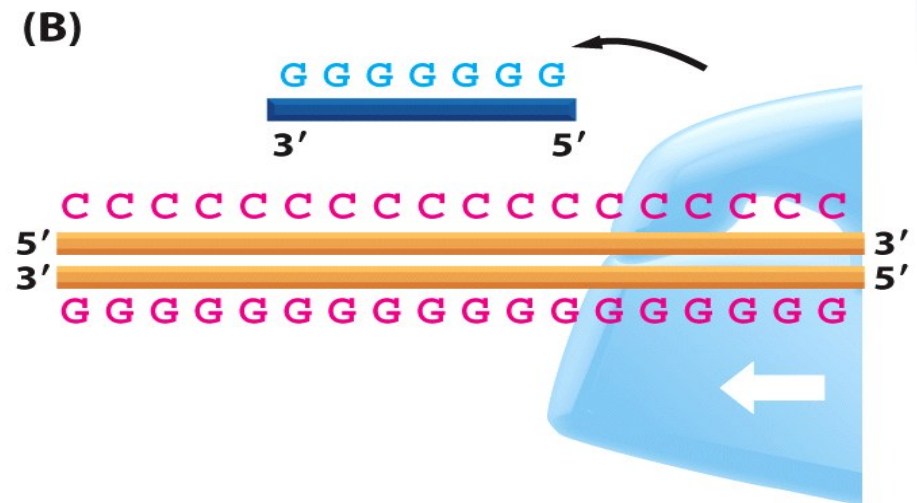
# A fita molde não é sempre a mesma



# A importância da orientação da RNA polimerase



**an RNA polymerase that moves from left to right makes RNA by using the bottom strand as a template**



**an RNA polymerase that moves from right to left makes RNA by using the top strand as a template**

# Replicação x Transcrição: Diferenças



- O transcrito de RNA não permanece associado à fita molde de DNA
- Processo menos preciso que a replicação (incorporações de bases erradas =  $1 \times 10^4$ )
- A transcrição é assimétrica
- Cópia seletivamente apenas partes do genoma com número de cópias variável
  - Escolha não rãndomica e regulável

**Como a RNAP sabe aonde o gene  
começa e onde termina?**

# Pontuação da transcrição

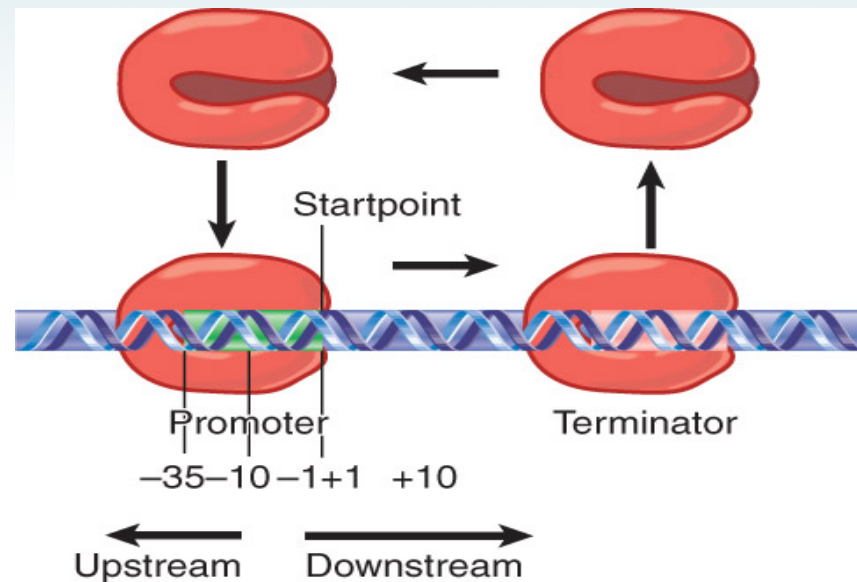


- **promotor** – Uma região do DNA aonde se liga a RNA polimerase para iniciar a transcrição.
- **Ponto de início de transcrição** – A posição no DNA correspondendo à primeira base incorporada no RNA.
- **terminador** – Uma sequência de DNA que induz a terminação da transcrição pela RNA polimerase.

# Transcrição

FIGURE 02: Promotores e terminadores definem a unidade de transcrição

- **Unidade de transcrição** – A sequência entre os sítios de iniciação e terminação pela RNA polimerase; pode incluir mais de um gene.
- **Upstream (a montante)** – Sequências na direção oposta da expressão (antes do ponto de iniciação).
- **Downstream (a jusante)** – Sequências na mesma direção da expressão (geralmente após o ponto de iniciação).





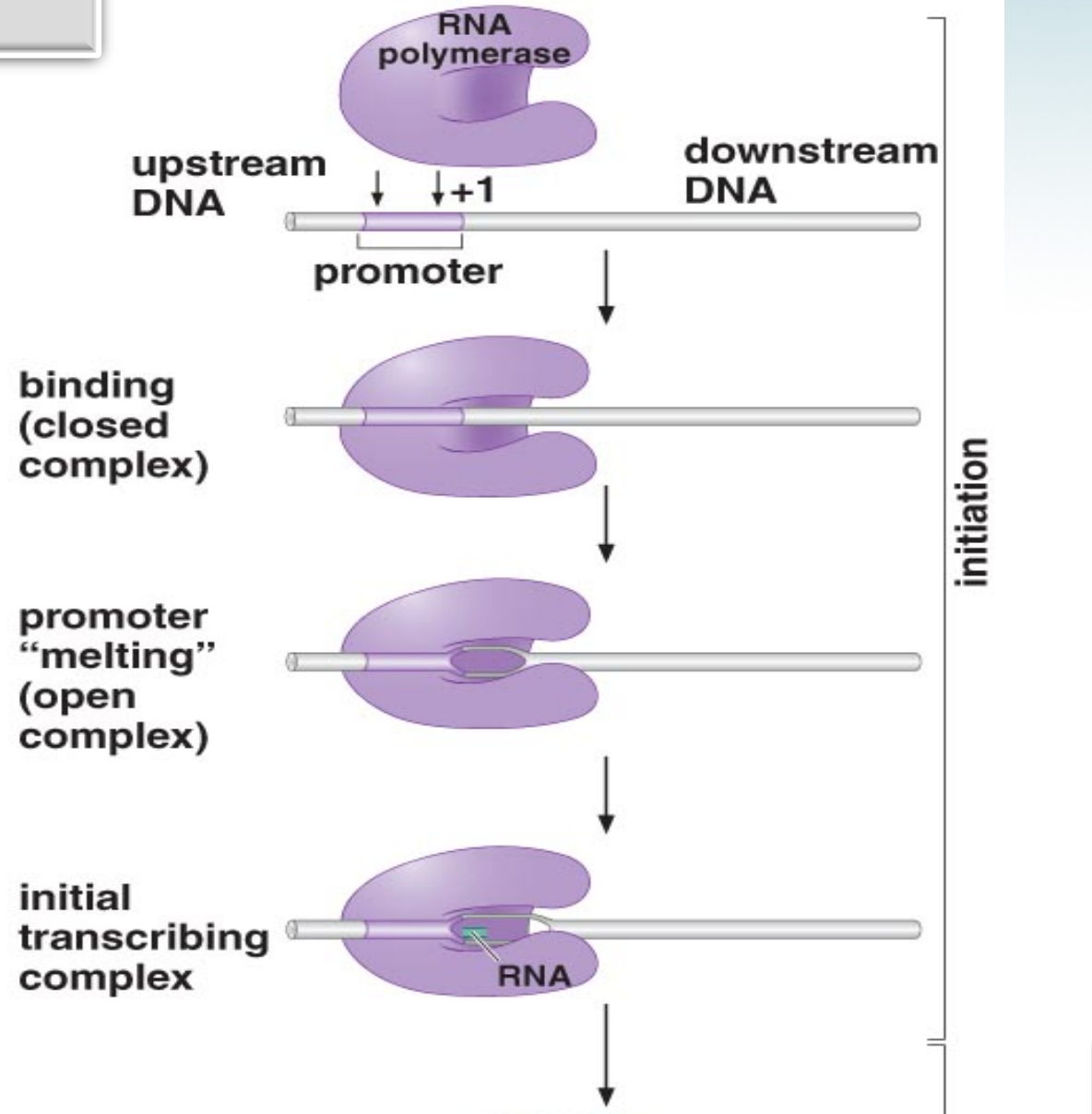
## Fases da transcrição

### Início

*Reconhecimento do promotor*

*Abertura da fita dupla de DNA*

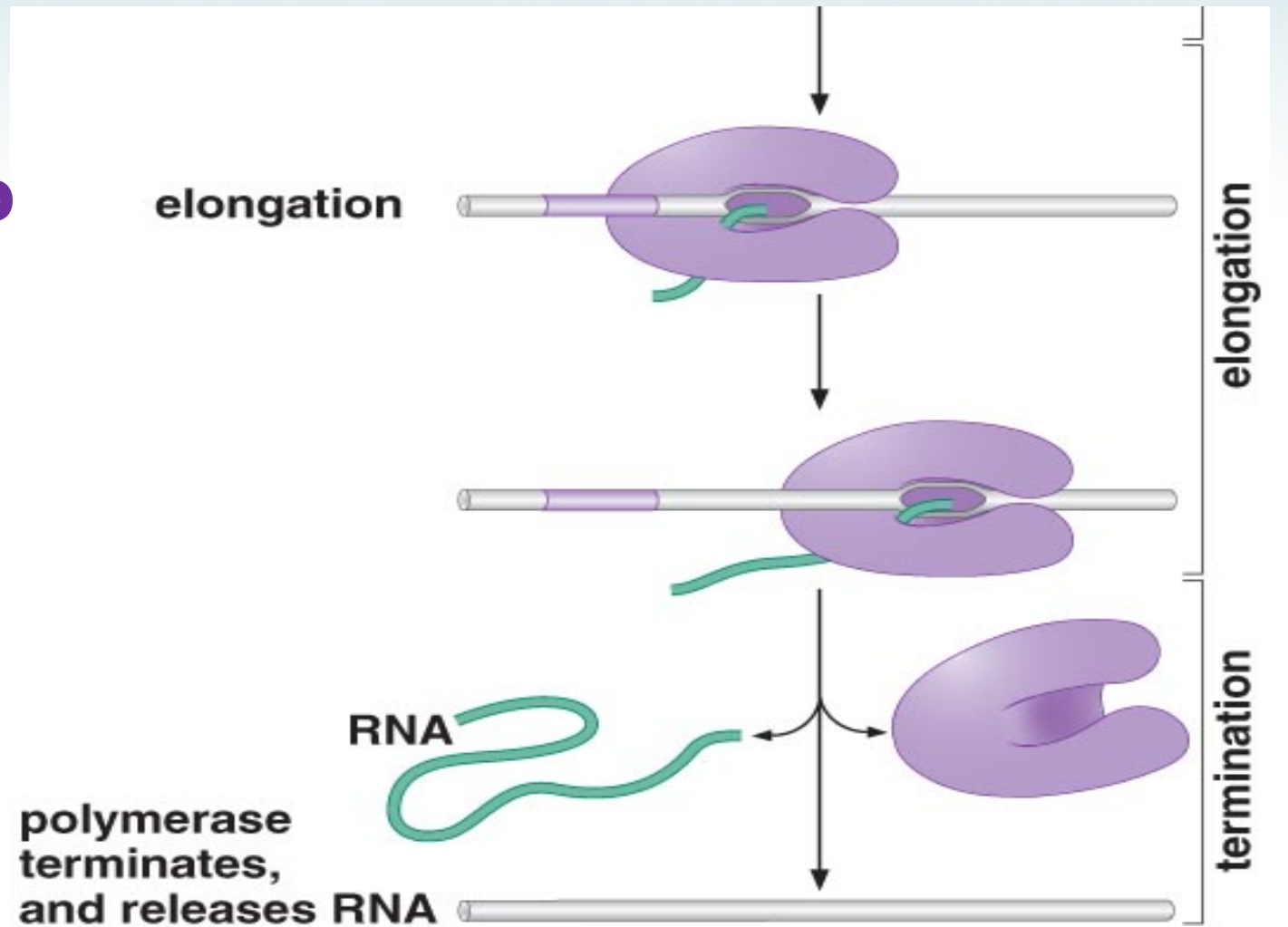
*Formação do complexo de início de transcrição*



## Fases da transcrição

**Alongamento  
da cadeia:  
escape da  
RNAP**

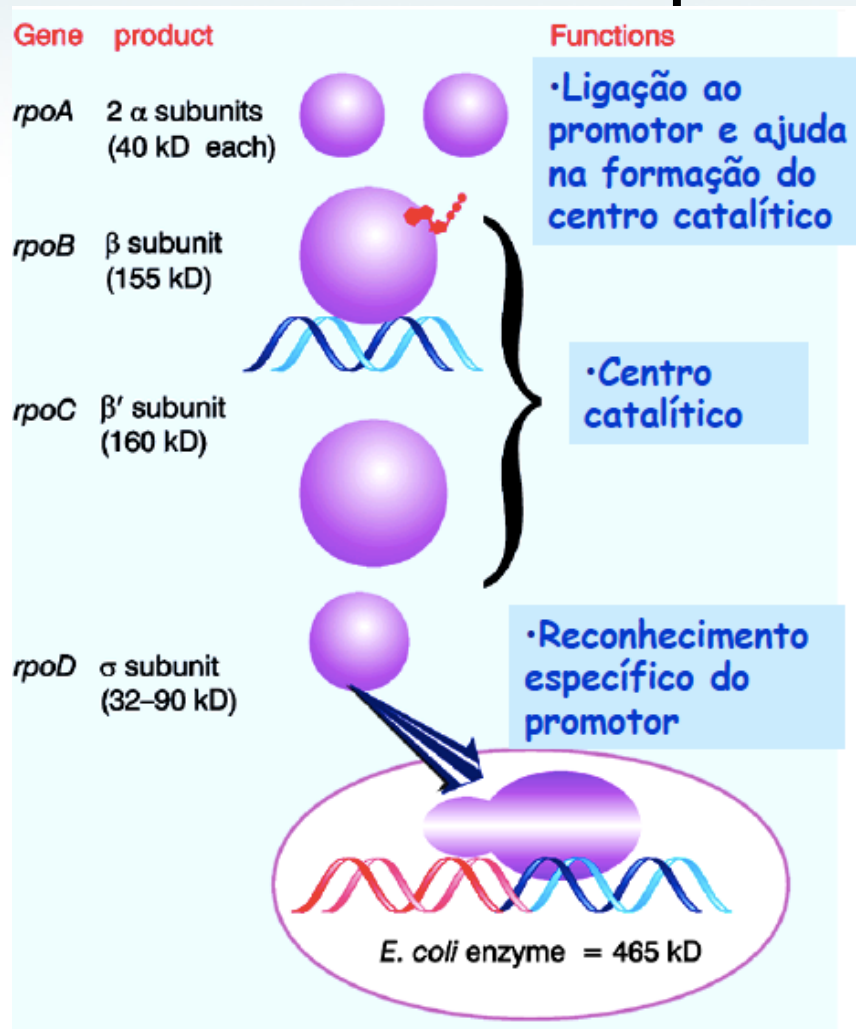
**Terminação**



# **Transcrição Bacteriana**



# A RNA Polimerase Bacteriana é Composta por Múltiplas Subunidades



- **holoenzima** – Forma completa da RNA polimerase competente para iniciar a transcrição. É composta pelas 5 subunidades do núcleo da enzima e o fator  $\sigma$ .

• Subunidade  $\omega$  (ômega) atua como chaperone na montagem da holoenzima

FIGURE 07: A RNA polimerase possui 4 tipos de

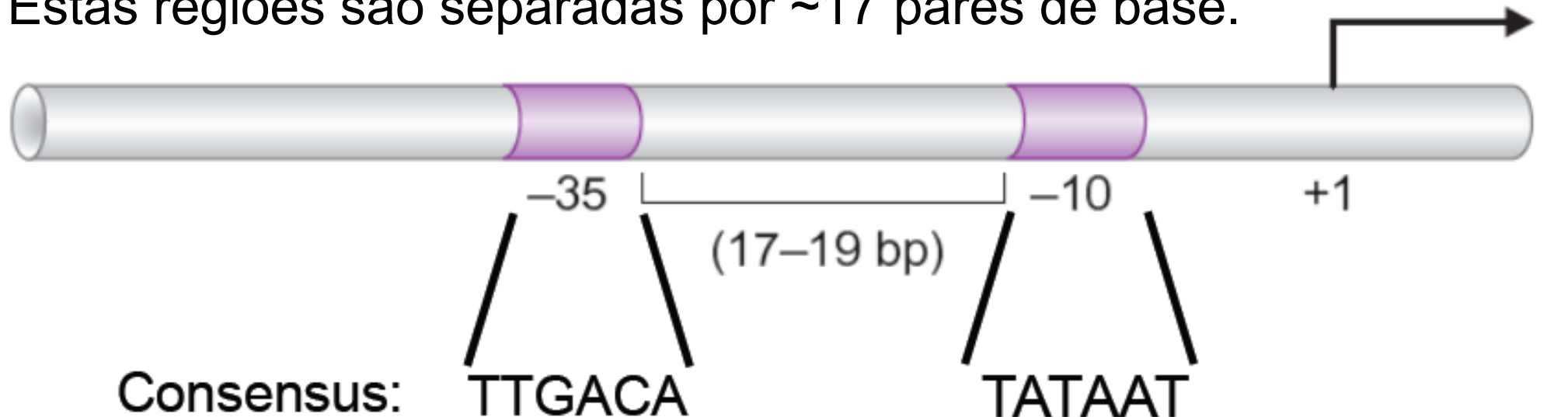
**Como a RNAP bacteriana reconhece o promotor?**

## Anatomia de um promotor bacteriano

Os promotores bacterianos são compostos de elementos de sequência conservada localizados a montante (antes) do sítio de início da transcrição.

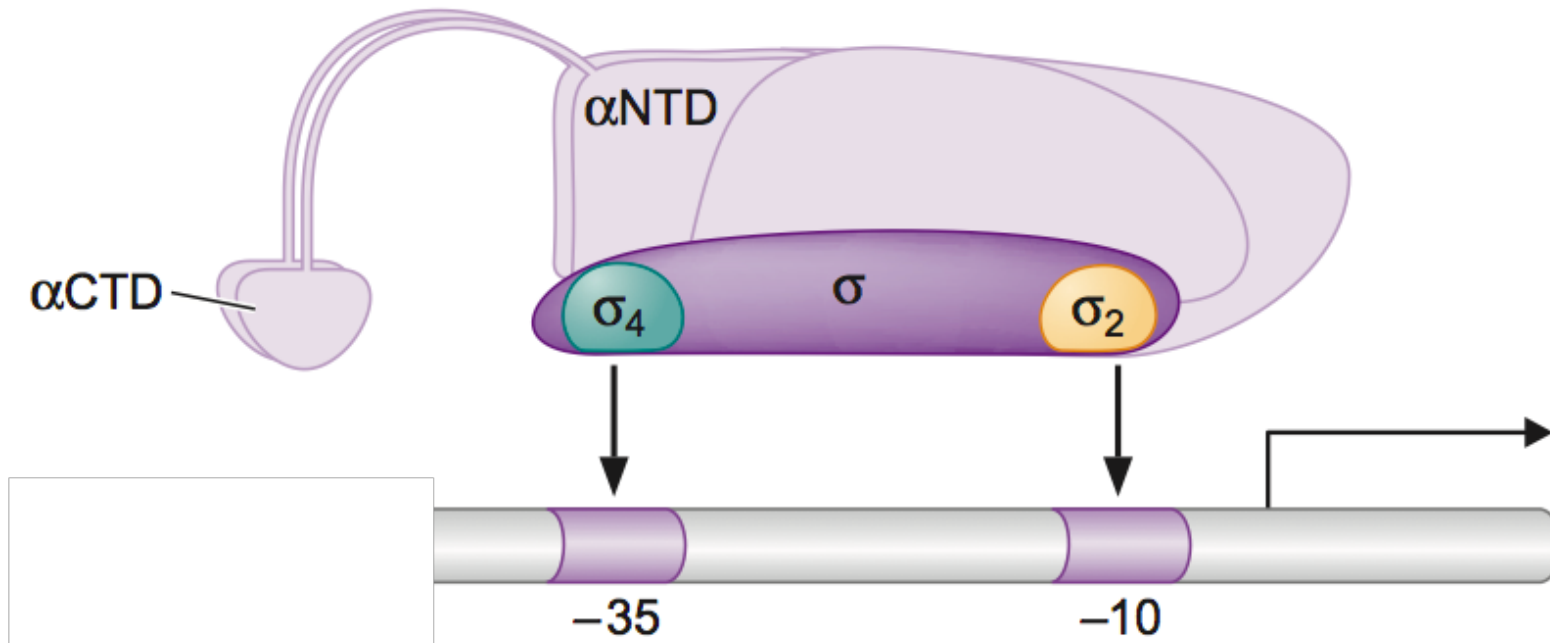
Os elementos mais importantes são as regiões a -10 e a -35.

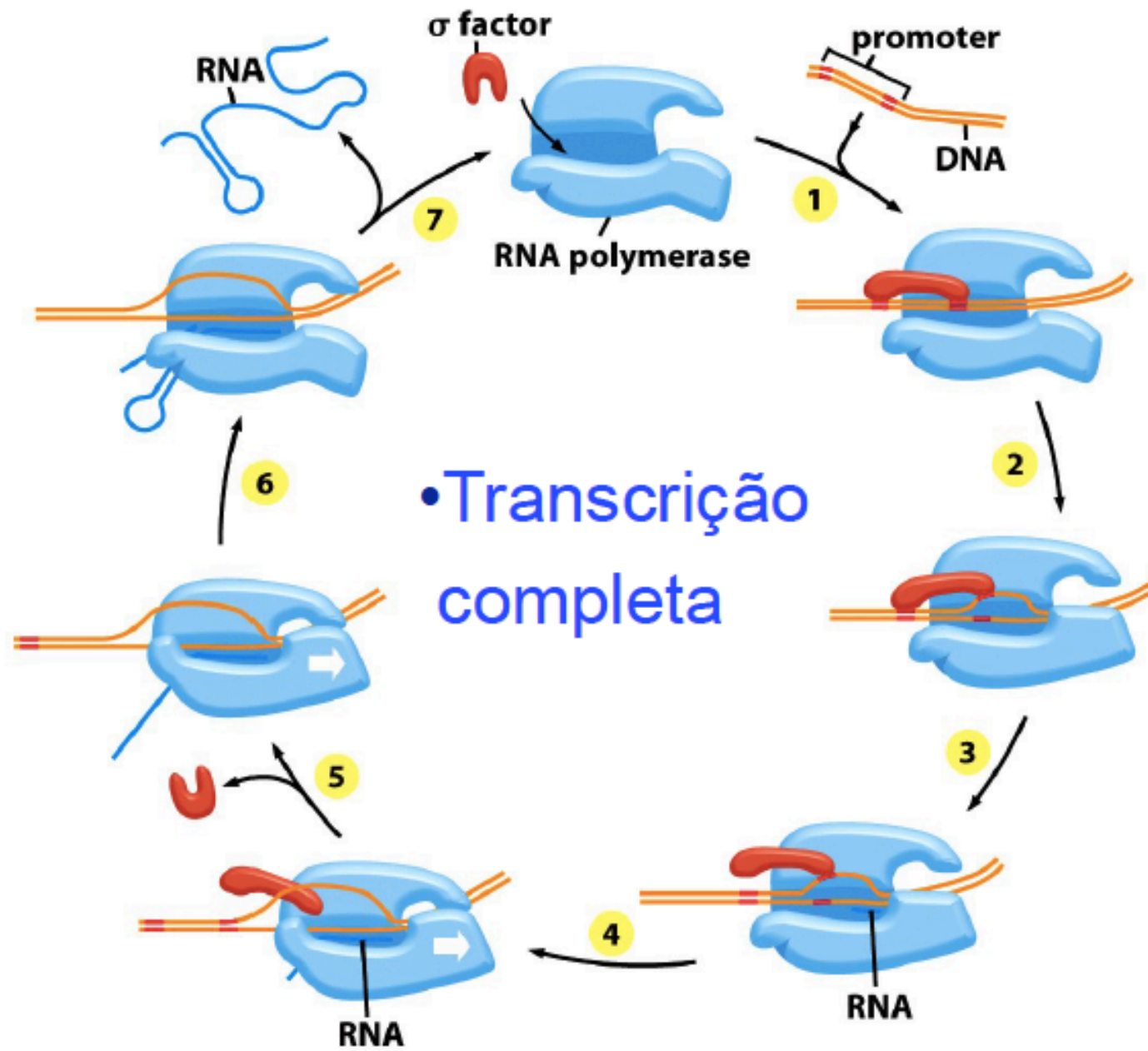
Estas regiões são separadas por ~17 pares de base.



# Como a RNAP bacteriana reconhece o promotor?

A subunidade  $\sigma$  da RNA polimerase se liga às regiões “-35” e “-10” do promotor .





• Transcrição completa



# **Terminação da transcrição bacteriana**

# A RNA polimerase bacteriana termina a transcrição em regiões definidas

Existem duas classes de terminadores:

- Aqueles reconhecidos somente pela RNA polimerase sem necessidade de quaisquer fatores celulares são chamados de **"terminadores intrínsecos"**.
- Outros requerem uma proteína celular chamada **rho** e são chamados de **"terminadores rho-dependentes."**

# Terminadores Intrínsecos (Rho-independentes)

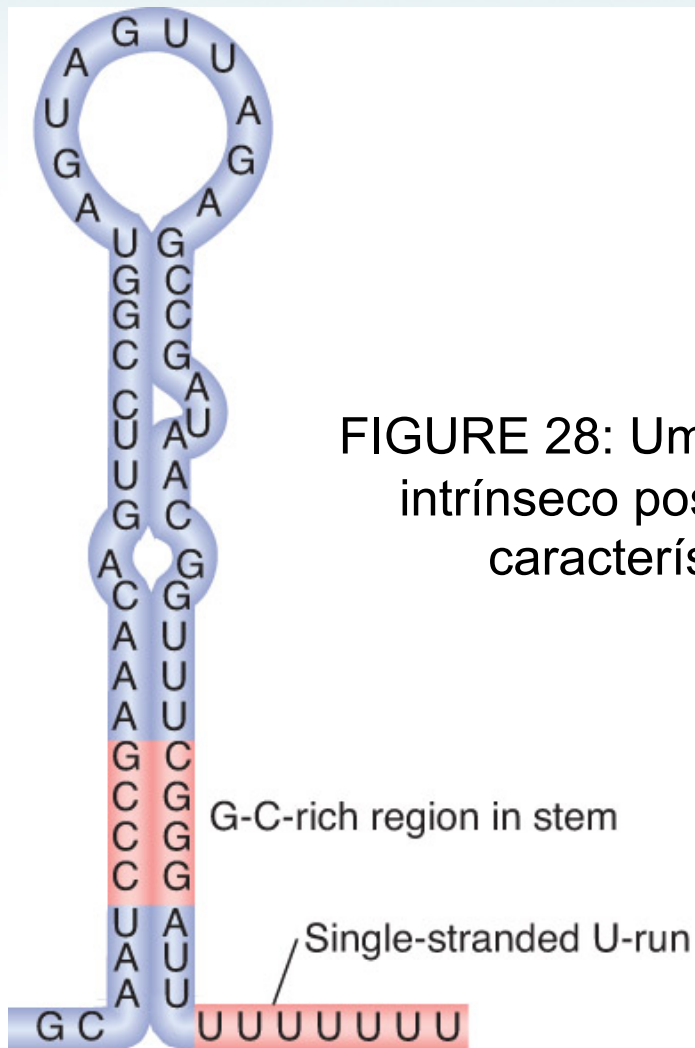
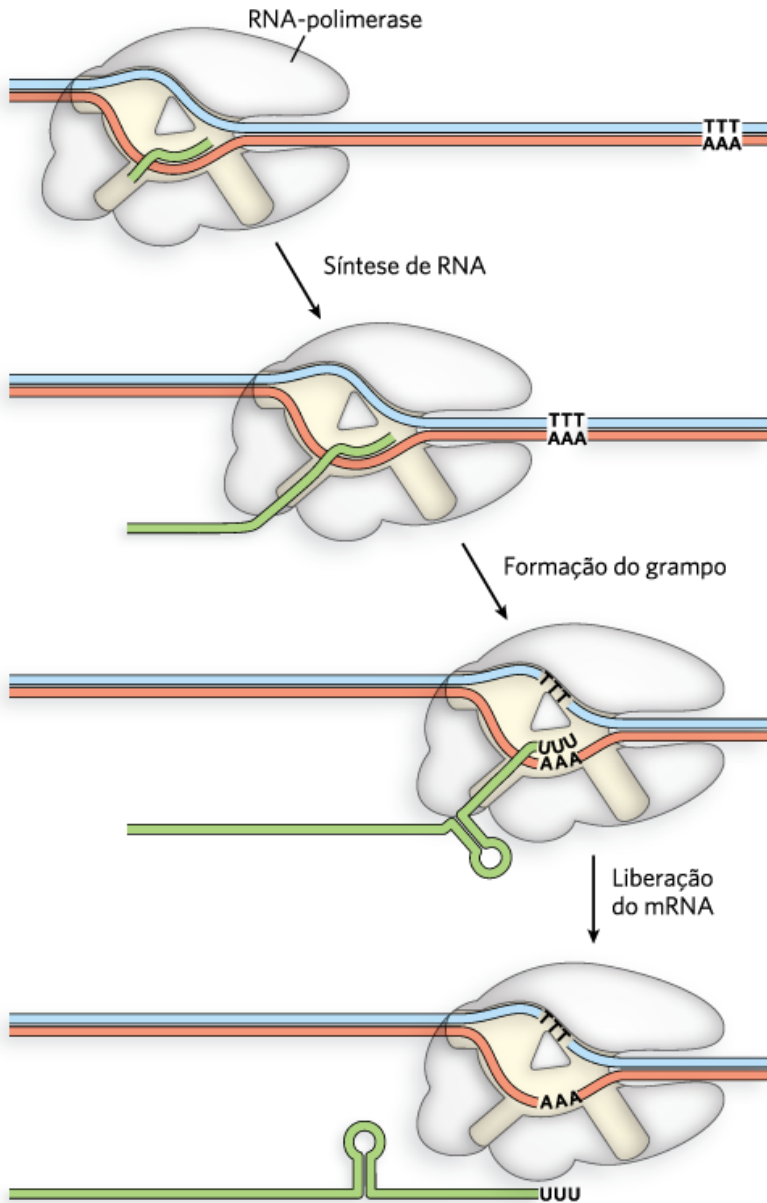


FIGURE 28: Um terminador intrínseco possui duas características

Requerem o reconhecimento de uma sequência terminadora no DNA que codifica uma estrutura de grampo (**hairpin**) no RNA.

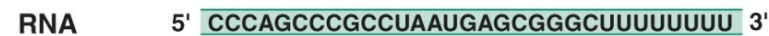
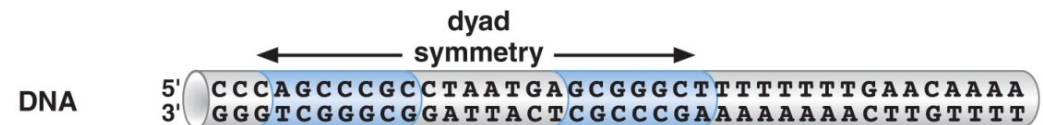
O sinal para terminação está contido dentro da sequência já transcrita pela RNA polimerase.

(a) Terminação independente de  $\rho$

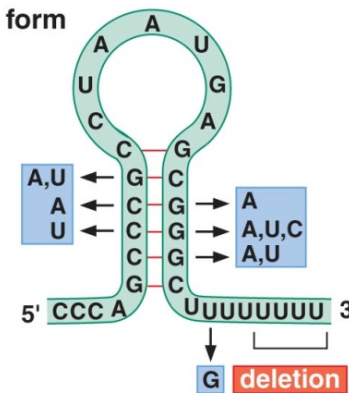


## Terminadores Intrínsecos (Rho-independentes)

Formação do grampo de terminação  $\rightarrow$  longa pausa  $\rightarrow$  desestabilização do complexo de transcrição



transcript folded to form termination hairpin



# **Transcrição em eucariotos**



# Transcrição em eucariotos

A transcrição em Eucariotos é mais complexa.

A sequência dos promotores é diferente.

Existem vários fatores de iniciação diferentes do fator  $\sigma$ .

Processo de terminação da transcrição é diferente.

Existe mais de uma RNA polimerase

# Transcrição em eucariotos

## Eucariotos tem 3 RNA polimerases

**TABLE 12-1** The Subunits of RNA Polymerases

Prokaryotic		Eukaryotic		
Bacterial	Archaeal	RNAP I	RNAP II	RNAP III
<b>Core</b>	<b>Core</b>	<b>(Pol I)</b>	<b>(Pol II)</b>	<b>(Pol III)</b>
$\beta'$	A'/A''	RPA1	RPB1	RPC1
$\beta$	B	RPA2	RPB2	RPC2
$\alpha^I$	D	RPC5	RPB3	RPC5
$\alpha^{II}$	L	RPC9	RPB11	RPC9
$\omega$	K	RPB6	RPB6	RPB6
	[+6 others]	[+9 others]	[+7 others]	[+11 others]

Adapted, with permission, from Ebricht R.H. 2000. *J. Mol. Biol.* 304: 687–698, Fig. 1, p. 688. © Elsevier. The subunits in each column are listed in order of decreasing molecular weight.

# Eucariotos tem 3 RNAs polimerases

1. A RNA pol I está localizada no nucléolo e transcreve genes de rRNAs

2. A RNA pol II está fora do nucléolo no núcleo e transcreve genes codificadores de proteínas e miRNAs

3. A RNA pol III está fora do nucléolo no núcleo e transcreve um gene de rRNA e genes de tRNAs



# Como promotores são reconhecidos pela RNA pol II?

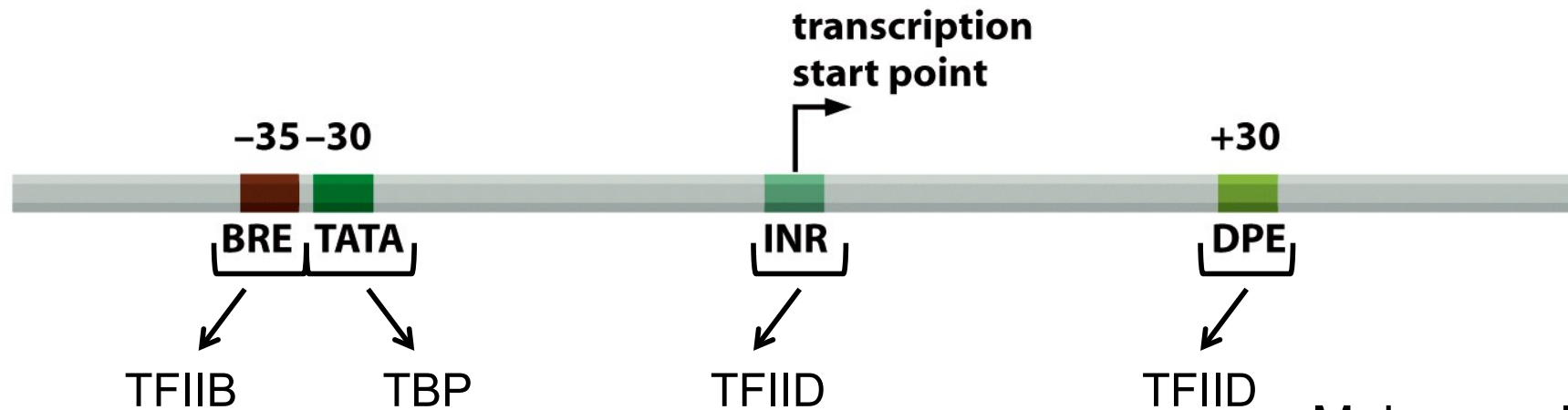
1. Ao contrário da RNA pol bacteriana, a RNA pol II não consegue reconhecer promotores sozinha
2. A RNA pol II requer proteínas adicionais chamadas **fatores de transcrição geral**

General Transcription Factor:	# of subunits:
TFIID (TBP + TAFs)	12
TFIIA	2
TFIIB	1
TFIIE	2
TFIIF	3
TFIIH	9

↓ TBP associated factors  
↓ Tata binding protein

# Promotor da RNA polimerase II

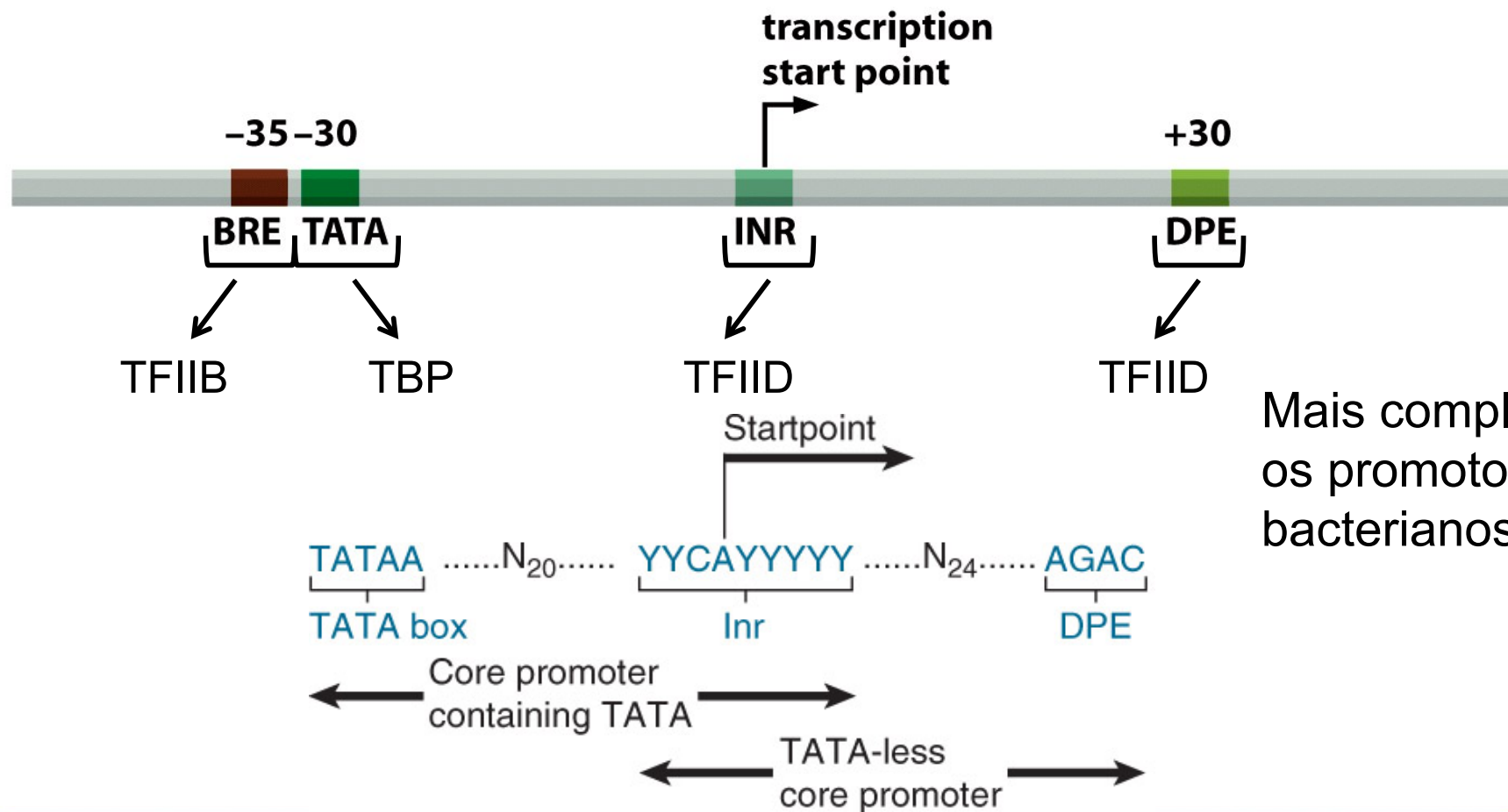
Sequências consenso encontradas na vizinhança dos sítios de início da transcrição da RNA polimerase II



Mais complexo que os promotores bacterianos.

# Promotor da RNA polimerase II

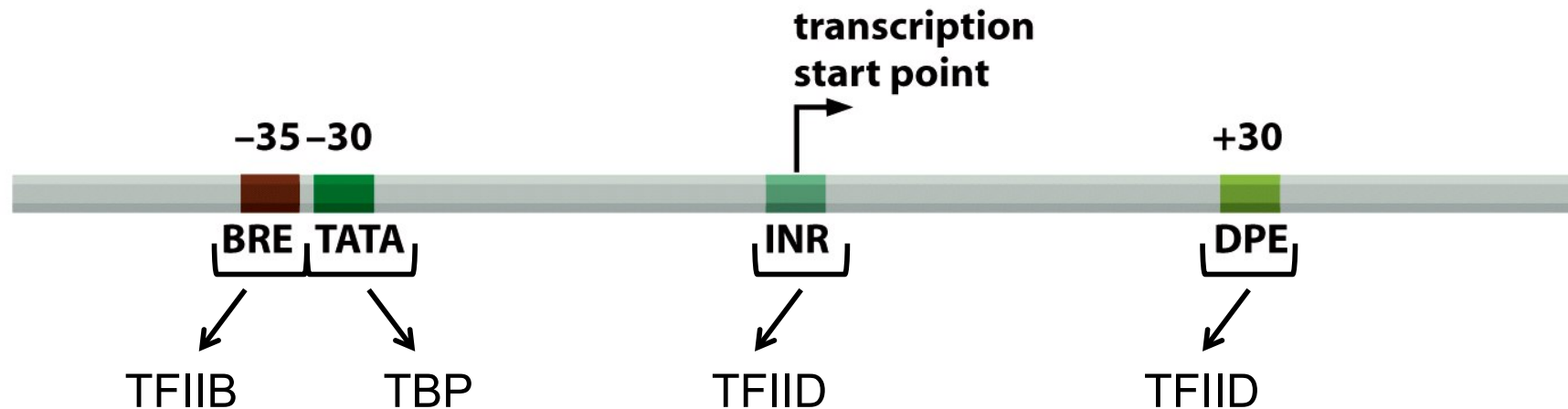
Sequências consenso encontradas na vizinhança dos sítios de início da transcrição da RNA polimerase II



Mais complexo que os promotores bacterianos.

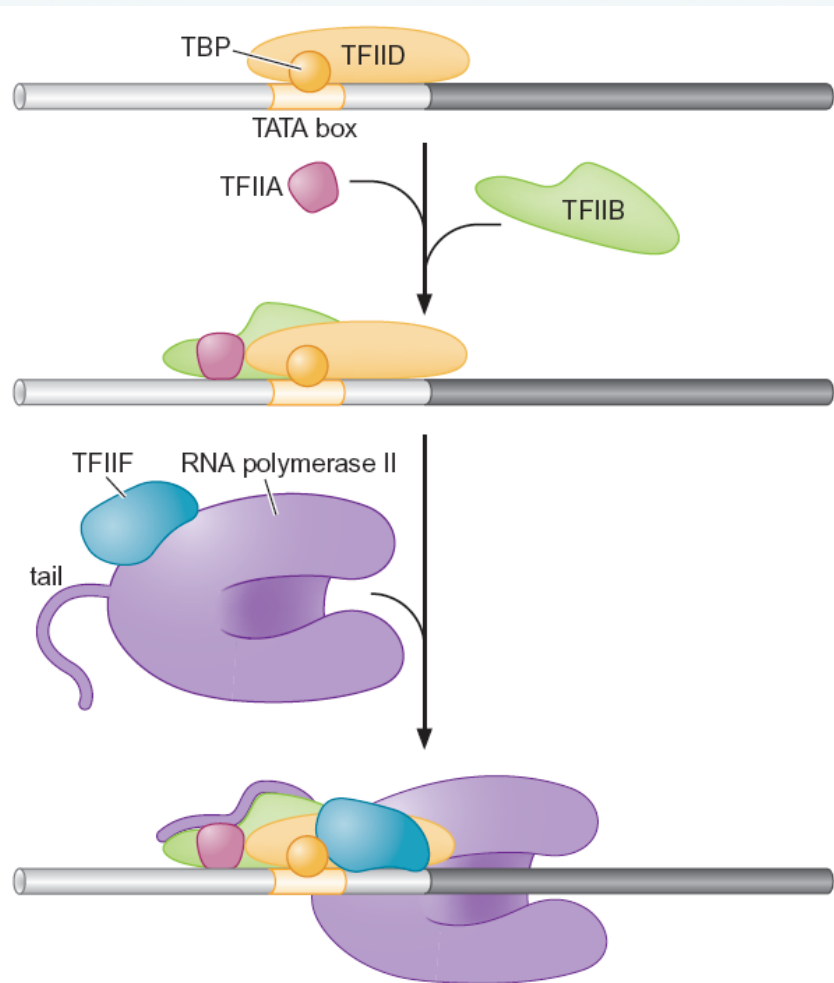
# Promotor da RNA polimerase II

Sequências consenso encontradas na vizinhança dos sítios de início da transcrição da RNA polimerase II



- O TATA box é o elemento mais fundamental do promotor eucariótico. A sequência consenso é: TATAAAA
- Localizado na posição -30 a -40 do início da transcrição
- TFIID se liga ao TATA box através da **Tata binding protein** ou **TBP**

# Montagem do complexo de transcrição basal

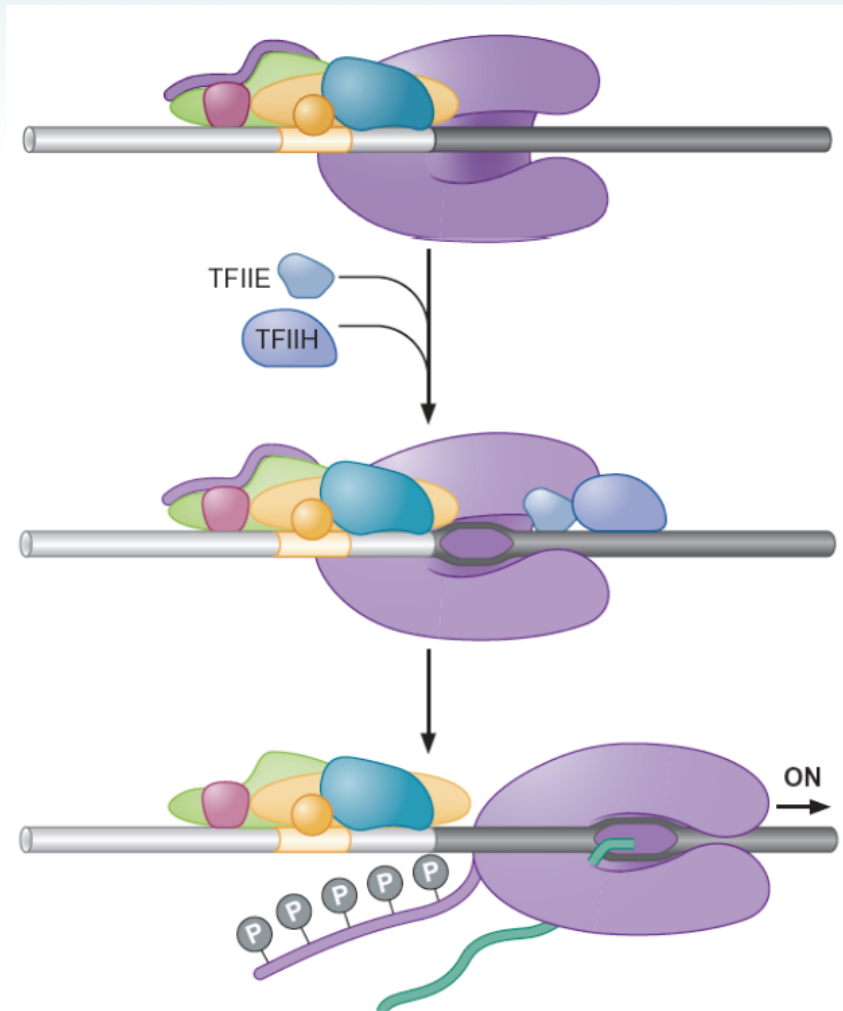


TFIID e sua subunidade TBP recrutam outros dois fatores de transcrição basais para o promotor, TFIIA e TFIIB.

Em seguida, o complexo TFIIA-TFIIB-TFIID recruta a RNAP para o promotor.

A RNA polimerase II forma um complexo de pré-iniciação com os fatores gerais de transcrição no promotor.

# Montagem do complexo de transcrição basal



Em seguida **TFIIE** e **TFIIH** se ligam ao complexo.

**TFIIH** tem duas atividades bioquímicas importantes: helicase e quinase.

**TFIIH** facilita a formação do complexo de transcrição aberto.

A RNA pol II se solta do complexo de iniciação após a **fosforilação** da sua cauda c-terminal

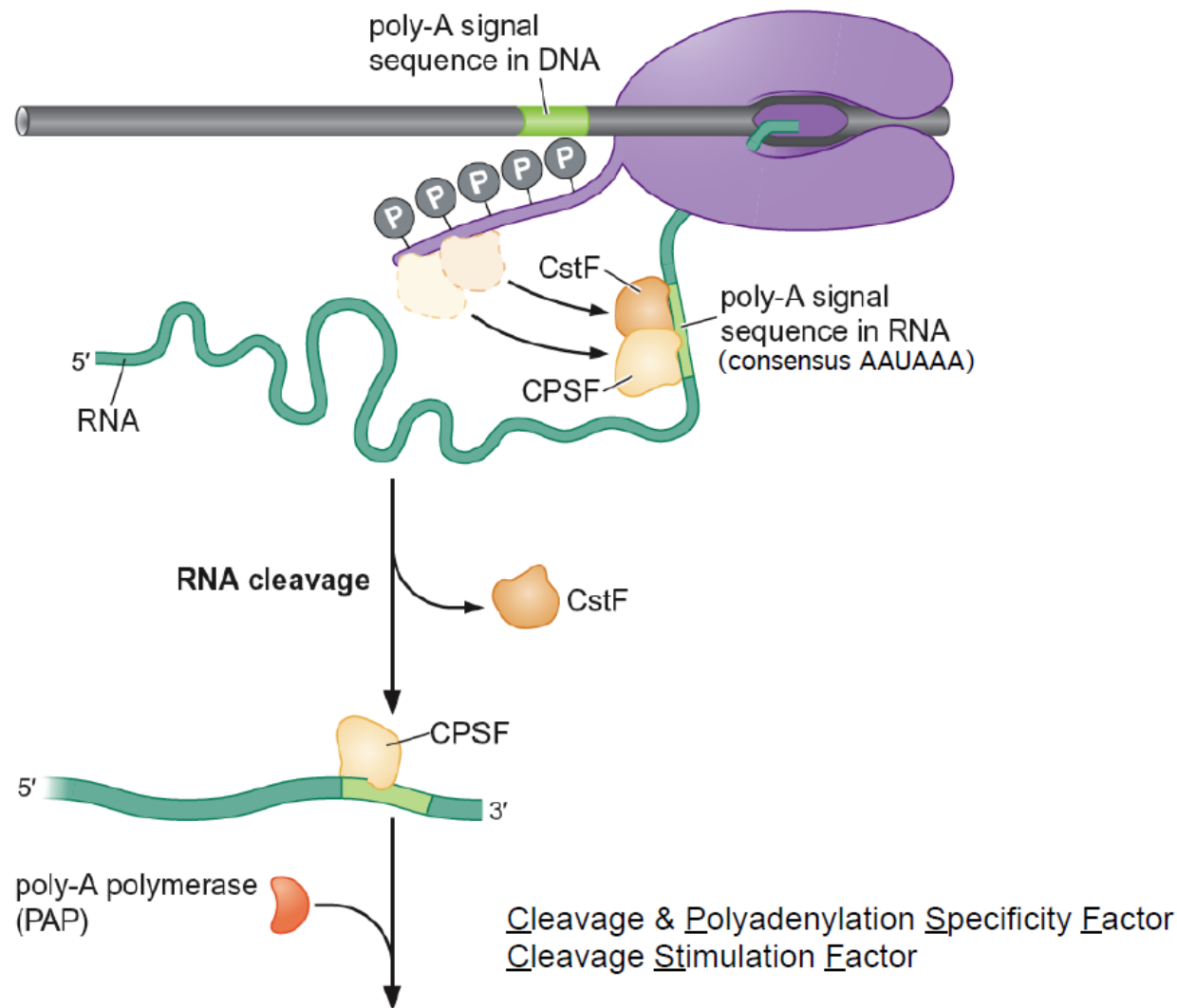
 **ELONGAÇÃO**

# Terminação da transcrição em eucariotos está ligada ao processo de poliadenilação

- O processo não é tão bem compreendido como em procariotos.
- O sinal de poliadenilação leva à clivagem do mRNA sendo transcrito

# Terminação da transcrição em eucariotos

## Polyadenylation Machinery



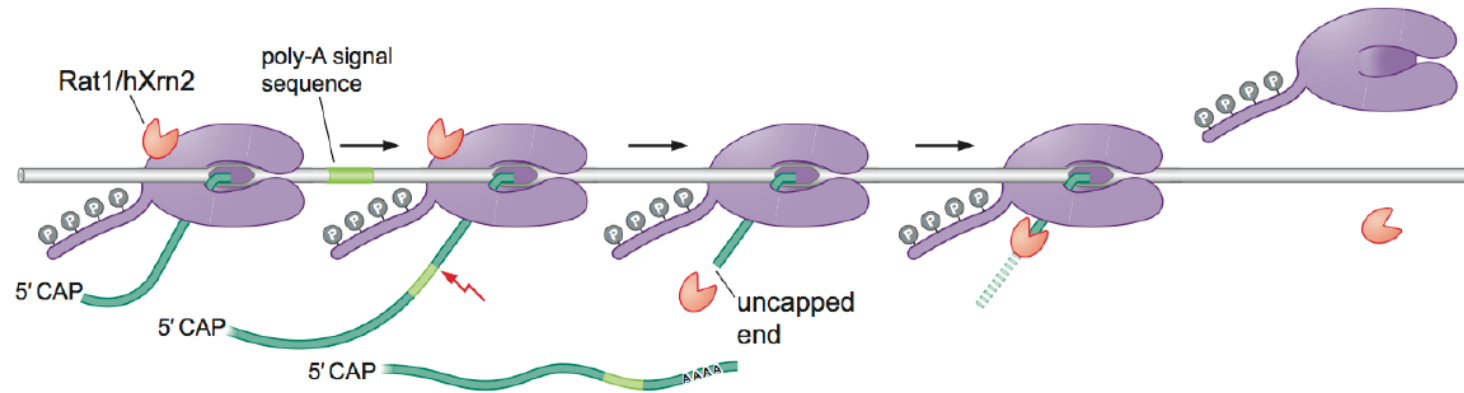


# Terminação da transcrição em eucariotos está ligada ao processo de poliadenilação

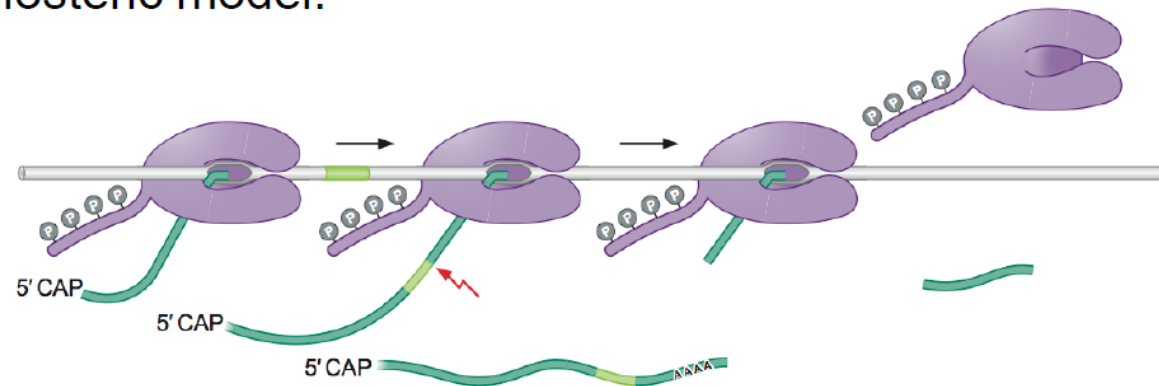
- O processo não é tão bem compreendido como em procariotos.
- O sinal de poliadenilação leva à clivagem do mRNA sendo transcrito
- Após a clivagem, a RNA pol continua transcrevendo, mas eventualmente se dissocia do molde terminando a transcrição. **Mas como?**

# Terminação da transcrição em eucariotos

The torpedo model:



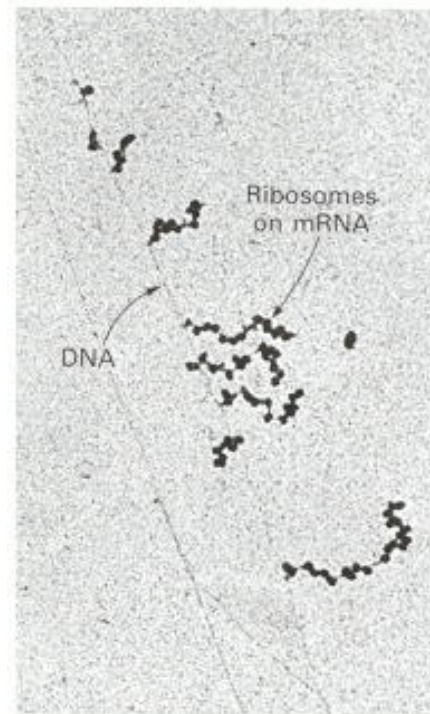
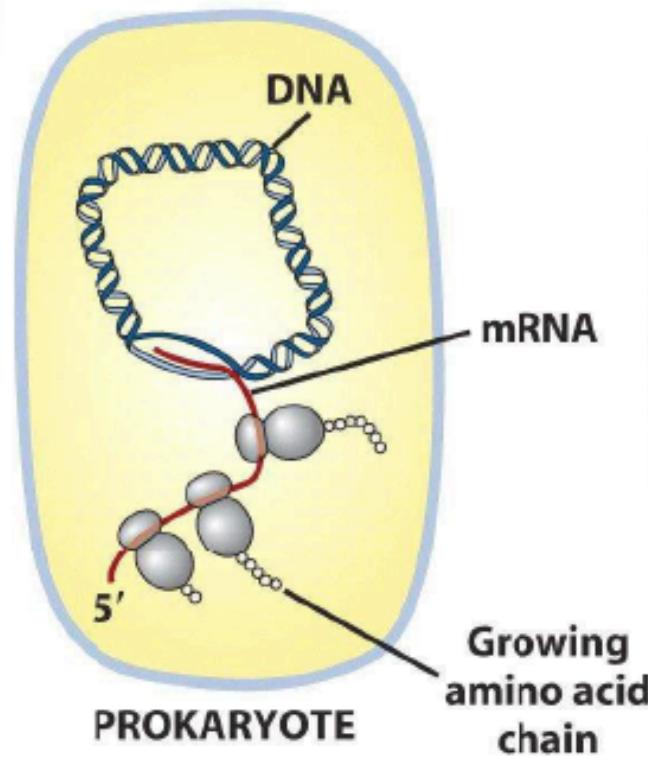
The allosteric model:



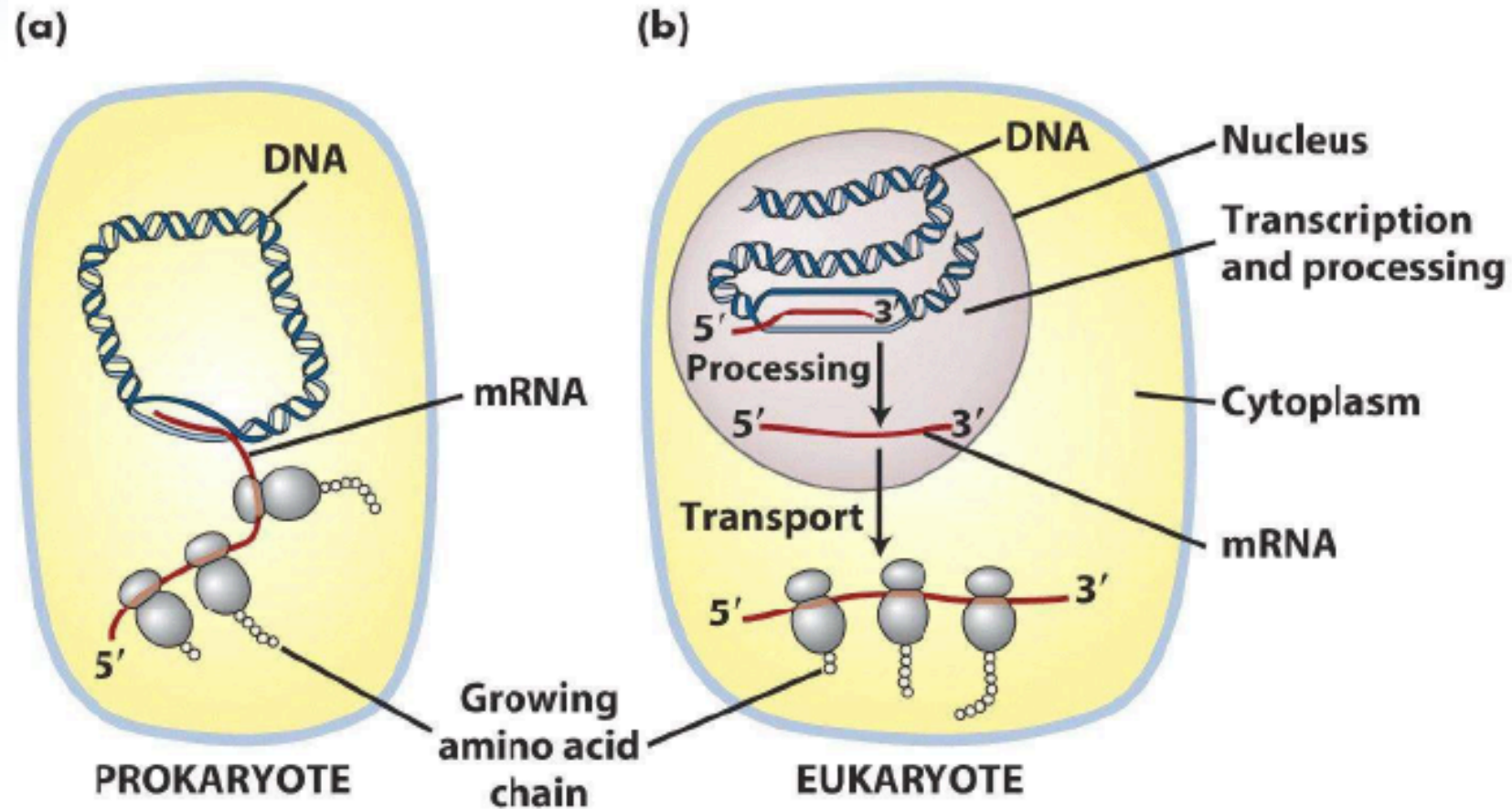
# Processamento de RNA (eucariotos)

# Processamento do RNA

(a)



# Processamento do RNA



# Processamento do RNA

Ocorre no núcleo durante a transcrição

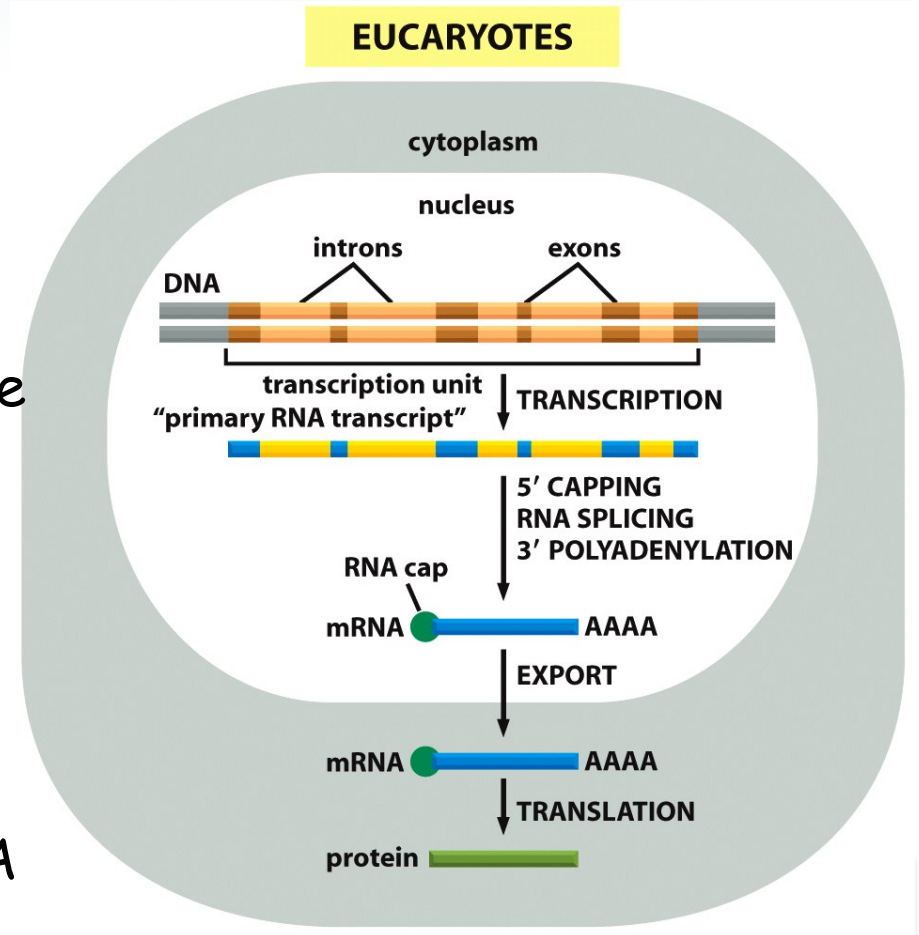
Etapas para a formação de um **mRNA maduro**

o Síntese de um pré-RNA (transcrito primário) no núcleo

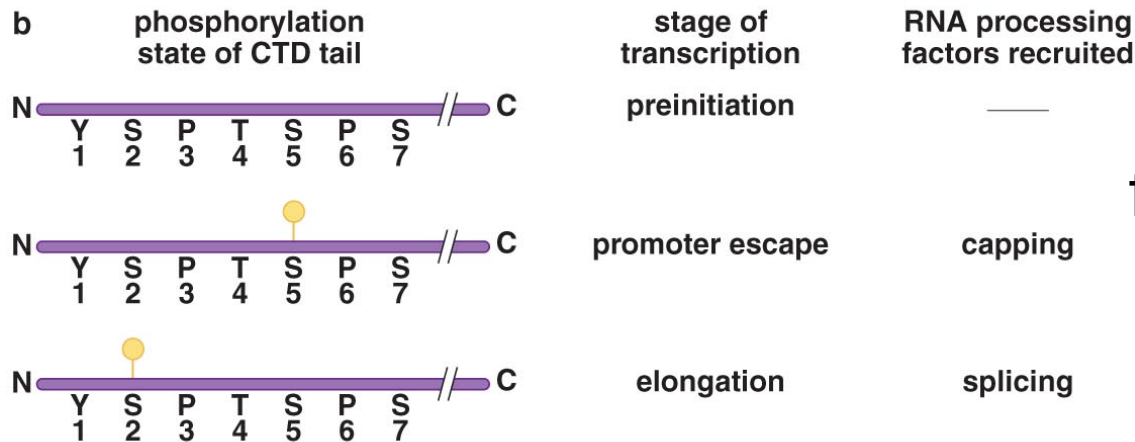
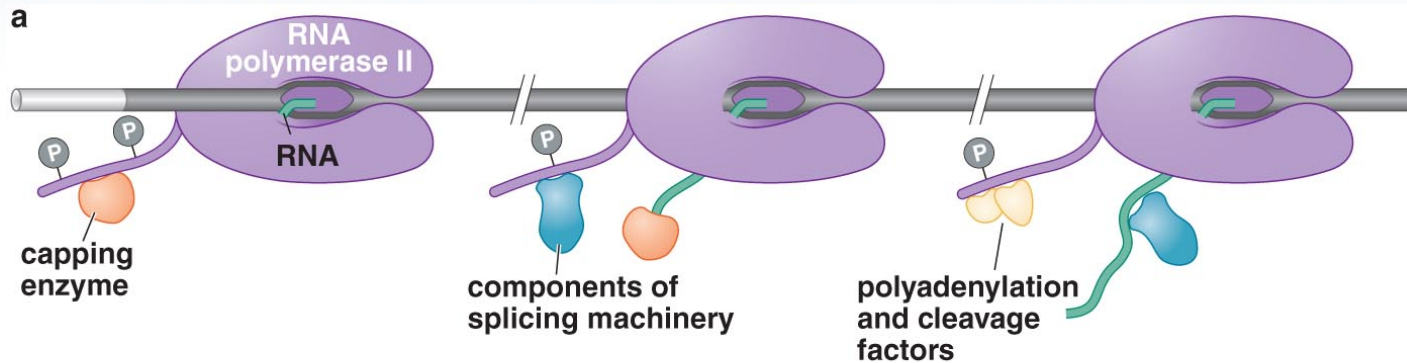
o Adição da estrutura CAP na extremidade 5' do RNA

o Remoção de introns (splicing)

o Poliadenilação na extremidade 3' do RNA



# Processamento do RNA depende do estado de fosforilação do CTD da RNAPII



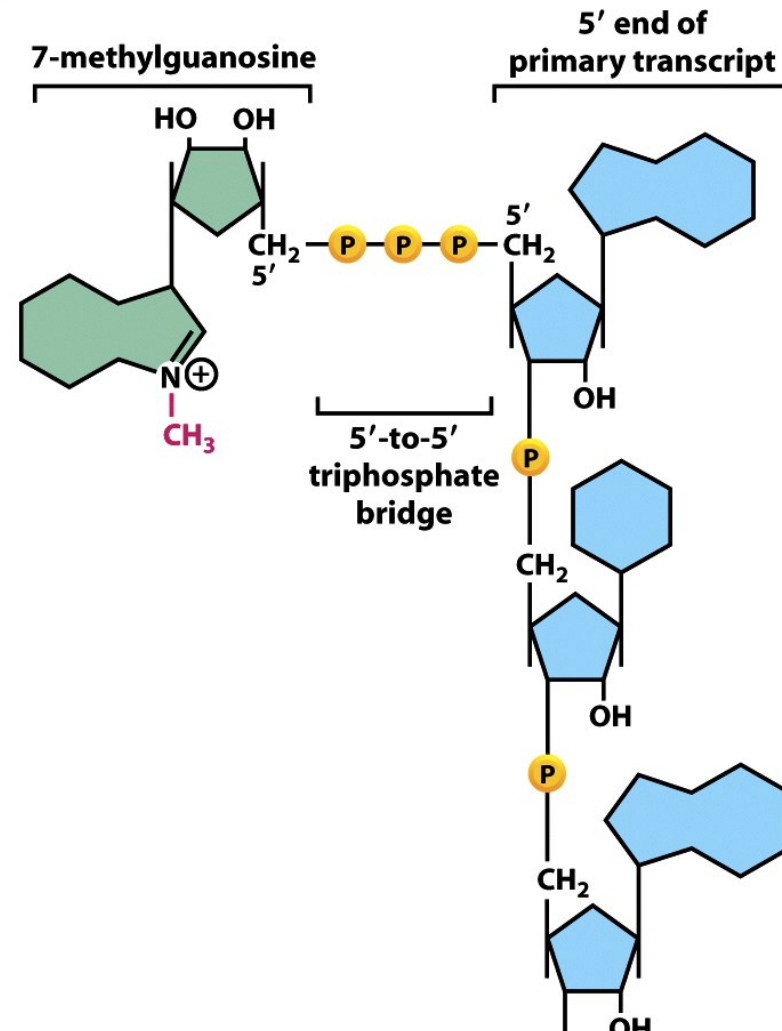
A cauda c-terminal fosforilada recruta a maquinaria de processamento

# Adição de CAP na extremidade 5' do mRNA

- Adição de uma guanosina à base terminal do transcrito *via* uma ligação trifosfato 5'-5' (Um G invertido).
- Metilação em N-7

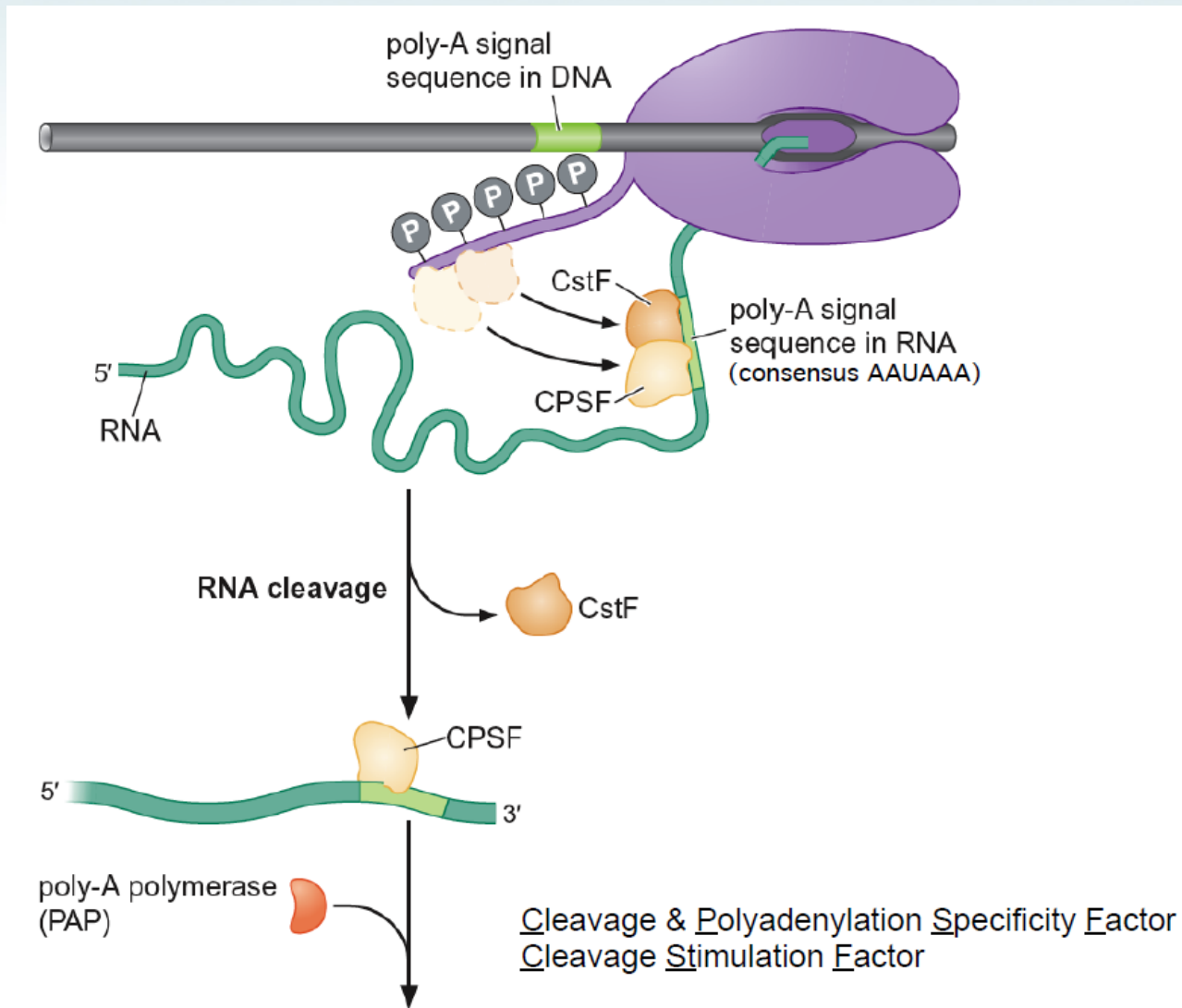
## Funções:

- Protege o mRNA de degradação
- Auxilia na transferência para o citoplasma
- Importante para a tradução



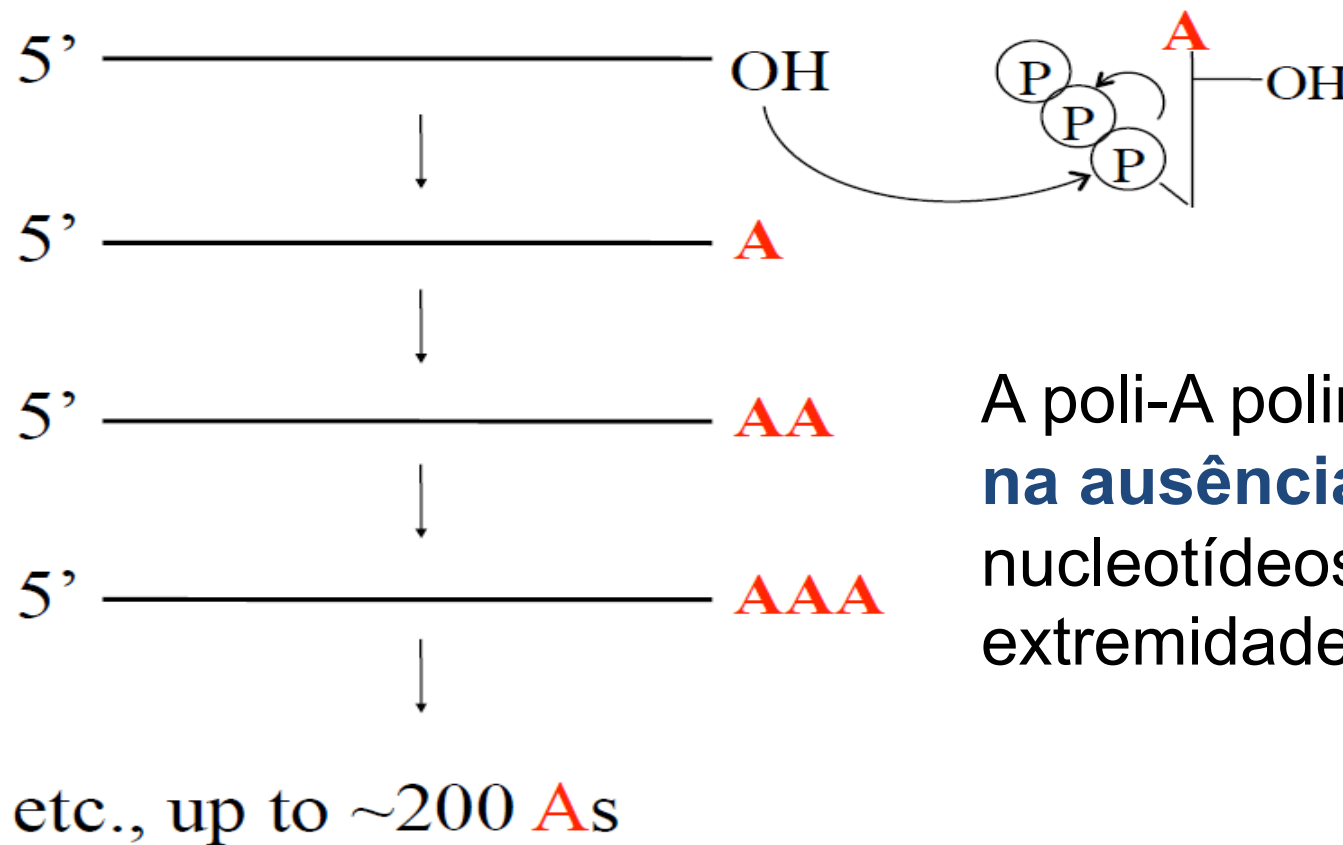


# Adição da cauda Poli-A na extremidade 3' do mRNA



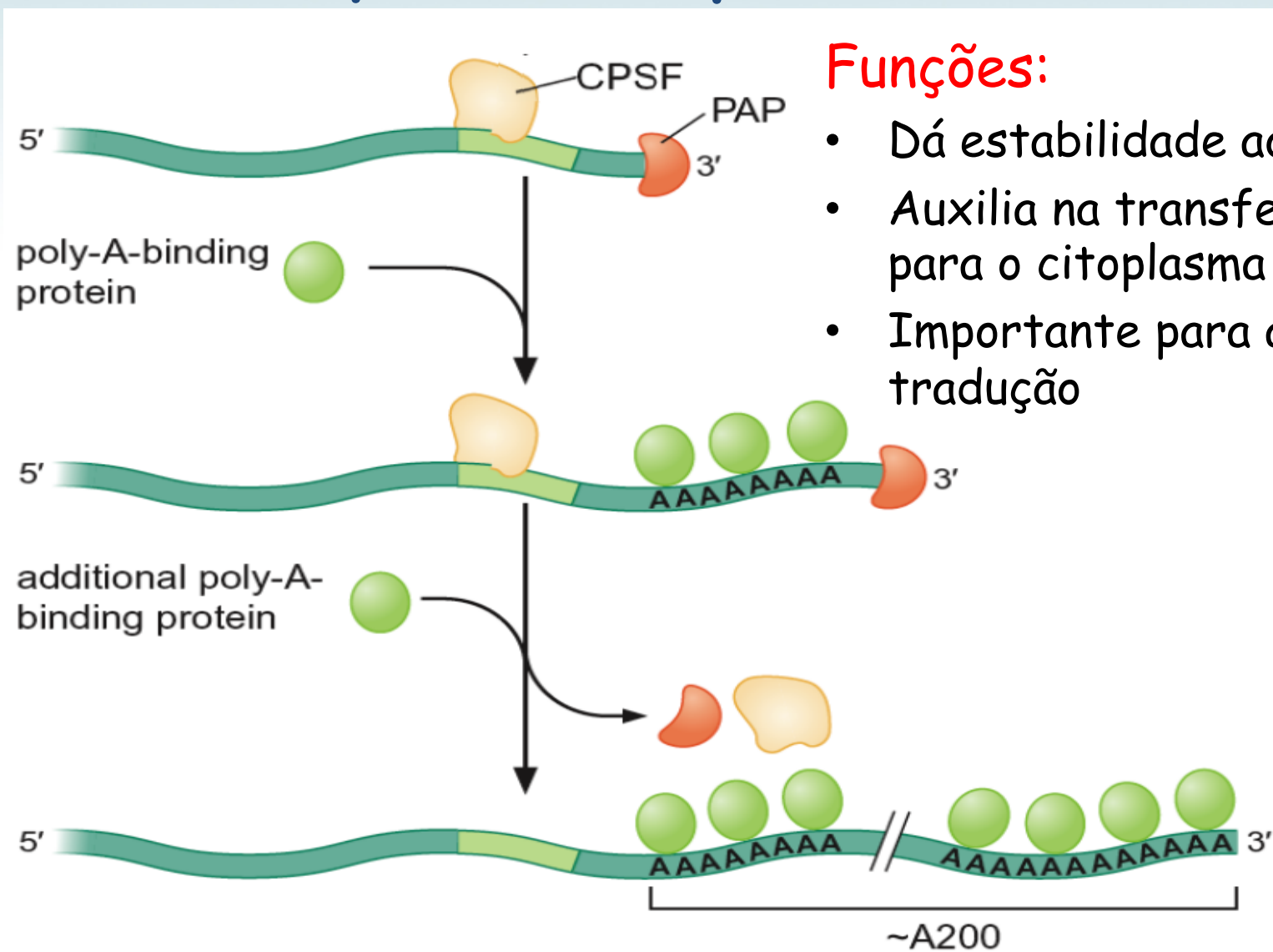
## A extremidade 3' do mRNA eucariótico sofre poliadenilação

- mRNAs eucarióticos adquirem uma cauda de adenosinas chamada de **cauda poli-A**.



A poli-A polimerase adiciona **na ausência de molde** nucleotídeos de adenosina à extremidade 3' dos mRNAs

# A maquinaria de poliadenilação

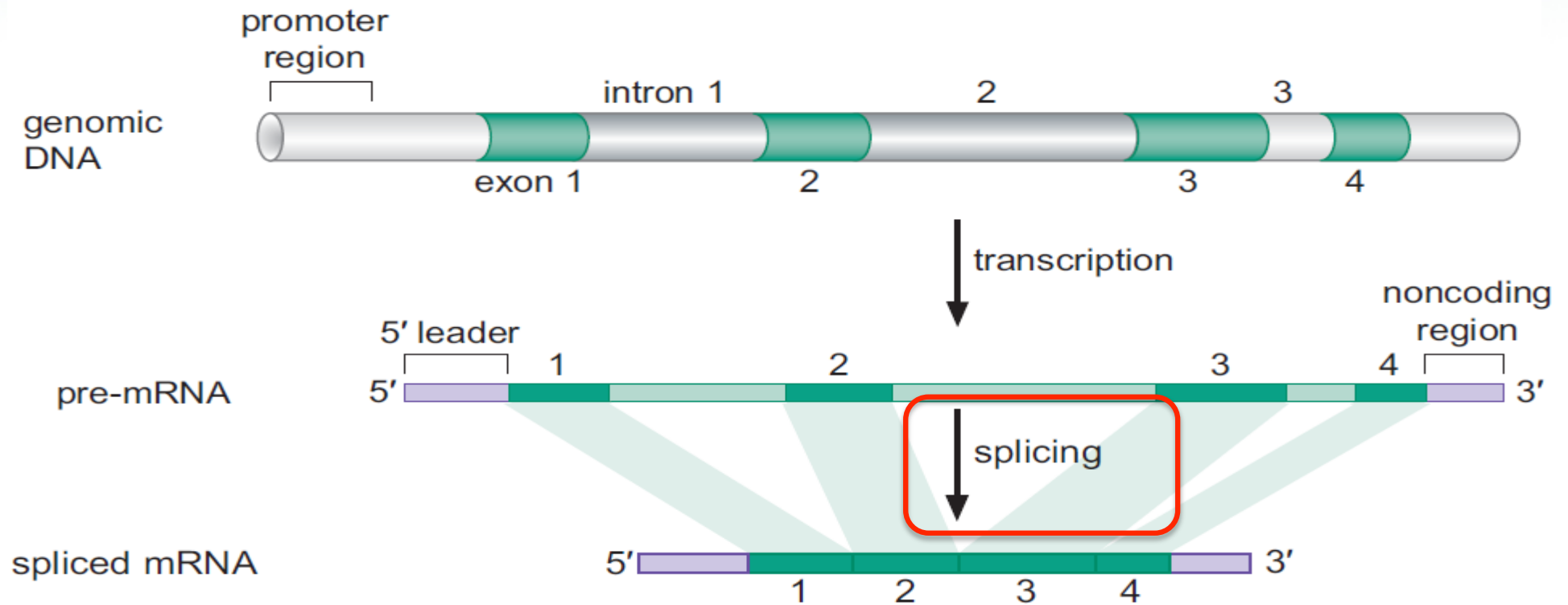


## Funções:

- Dá estabilidade ao mRNA
- Auxilia na transferência para o citoplasma
- Importante para a tradução

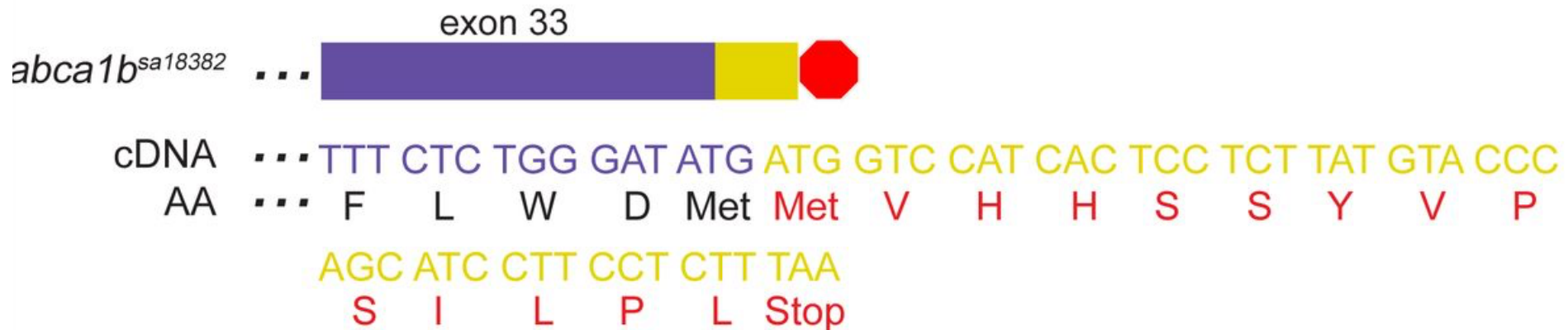
# Splicing

processo que remove os introns do transcrito primário (pré-mRNA)



Splicing é um processo que requer grande precisão, pois se a junção de dois exons for errada por apenas um nucleotídeo a sequência codificadora terá o quadro de leitura alterado

# Mutações podem alterar sítios de splicing e causar doenças genéticas



# Erros de splicing

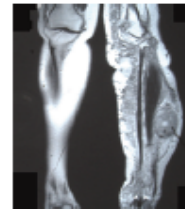
– Breast Cancer (*BRCA1*)



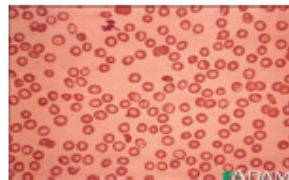
– Duchenne Muscular Dystrophy (*DMD*)



– Neurofibromatosis (*NF-1*)



– Thalassemias



– Ocular albinism (*OA-1*)



# Química e maquinaria de splicing

- O splicing é governado por **sequências conservadas** nas **junções exon/intron**
- O splicing é mediado por uma maquinaria denominada **spliceossomo**
- O spliceossomo é composto por ribonucleoproteínas chamadas **snRNPs (small nuclear ribonuclear proteins)**

A precisão da maquinaria de splicing é atingida pela complementaridade de sequência entre os snRNAs e o pré-mRNA

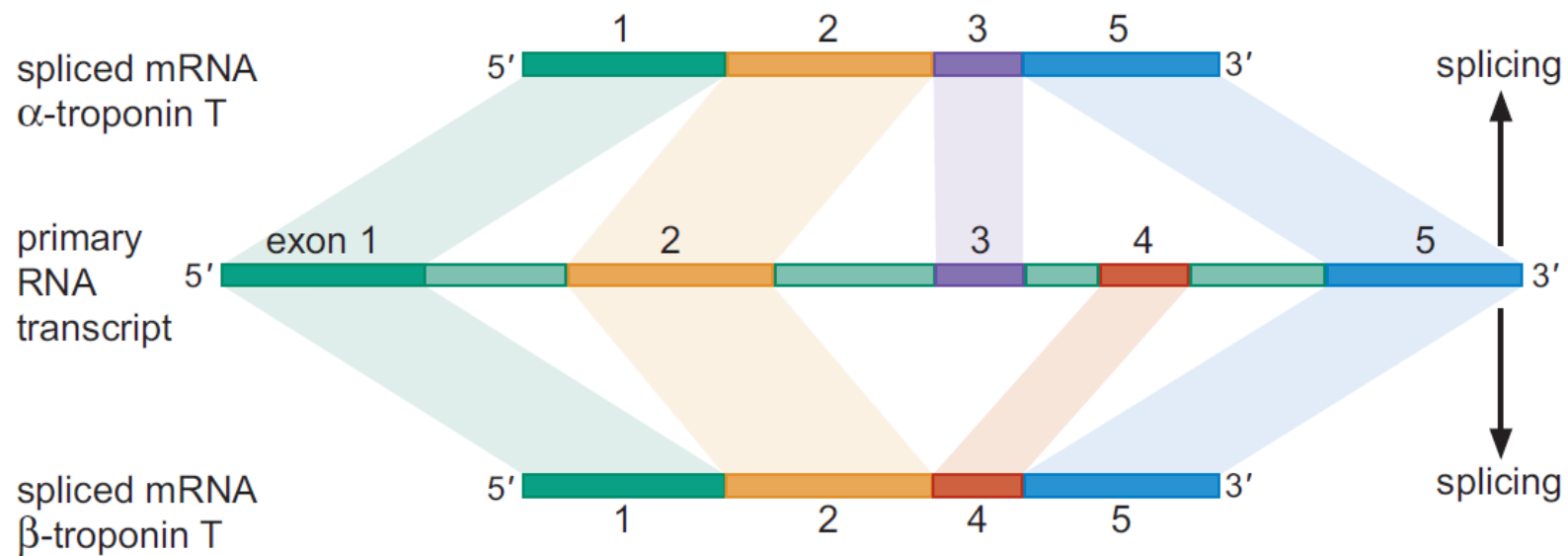


# Splicing alternativo

- Processo em que um gene dá origem a mais de 1 RNA. Os diferentes RNAs gerados são chamados de **isoformas**.
- Nem todos os pré-mRNAs são processados de maneira a capturar todos os exons
- Muitos são processados de diversas formas alternativas, gerando diferentes mRNAs maduros (cada um consistindo de uma diferente combinação de exons)

# Splicing alternativo

The gene for the muscle protein troponin generates two different proteins by alternative splicing:  
 **$\alpha$  troponin** and  **$\beta$  troponin**.



~~Um gene → um transcrito~~

***Splicing alternativo***

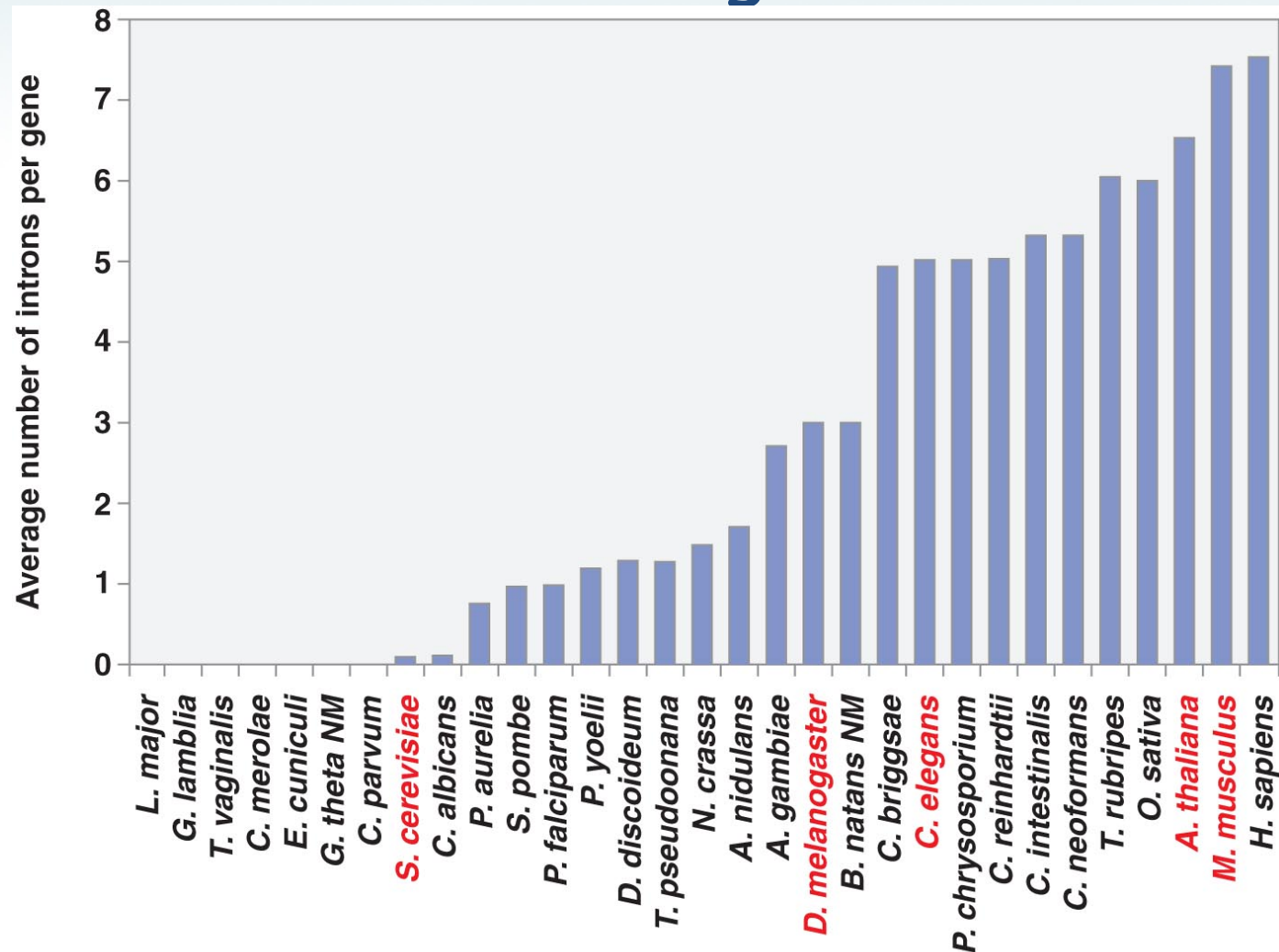


**Um gene → múltiplos transcritos  
(isoformas ou variantes)**

# Splicing alternativo

- Mecanismo combinatório que aumenta a diversidade do genoma
- Uma das soluções para explicar como apenas 20 - 25K genes podem gerar um ser humano

# Número médio de introns por gene em diferentes organismos



Copyright © 2008 Pearson Education, Inc., publishing as Pearson Benjamin Cummings.

Os introns variam em tamanho, mas em geral são maiores que os exons. Os