

## TEXTO 4

# TRANSCRIÇÃO E PROCESSAMENTO DO RNA

### Transcrição e Processamento do RNA

A Natureza Química do RNA

O Dogma Central da Biologia

As Classes de RNA

As Polimerases do RNA

A polimerase do RNA das bactérias

As polimerases do RNA dos eucariontes

A Transcrição do RNA

Terminologia referente à transcrição

Etapas da transcrição em procariontes

Reconhecimento do molde de DNA

Início da transcrição

Fase de alongamento

O término da transcrição

A Transcrição em Eucariontes

O início da transcrição em eucariontes

O RNA mensageiro dos eucariontes

Íntrons são removidos do pré-RNA<sub>m</sub> no *splicing*

Os RNA<sub>m</sub> sofrem outras modificações

O RNA Transportador

O RNA Ribossômico

A síntese de RNA<sub>r</sub>

Exercícios

**Parte A:** Revendo Conceitos Básicos

**Parte B:** Ligando Conceitos e Fatos

**Parte C:** Aplicando Conceitos

**Parte D:** Resolvendo Problemas

## Apresentação

O dogma central sintetiza o paradigma da biologia molecular: os genes se perpetuam como ácidos nucleicos, mas se expressam na forma de proteínas, cuja sequência de aminoácidos é determinada pela sequência de bases do RNA, o qual é transcrito a partir de uma das fitas da molécula de DNA

Essa apostila foi organizada pelos docentes do Instituto de Biociências da USP que ministraram e ministram as disciplinas “Biologia Molecular e de Microrganismos”, “Biologia Molecular” e “Fundamentos de Biologia Molecular” para os cursos de Bacharelado e Licenciatura em Ciências Biológicas. O texto foi adaptado e compilado de diversos livros didáticos e materiais instrucionais.

# Transcrição e Processamento do RNA

## A Natureza Química do RNA

As moléculas de ácido ribonucleico (RNA) são polímeros constituídos basicamente por quatro tipos de nucleotídeos que diferem quanto às bases nitrogenadas. Quando comparado ao DNA, o RNA apresenta duas diferenças básicas no que se refere à composição: enquanto o açúcar do DNA é a desoxirribose, o RNA apresenta em sua constituição a ribose, que possui um grupo hidroxila (OH) adicional, localizado no carbono 2'; no RNA, a base pirimídica uracila está presente no lugar da timina (Fig. 1).

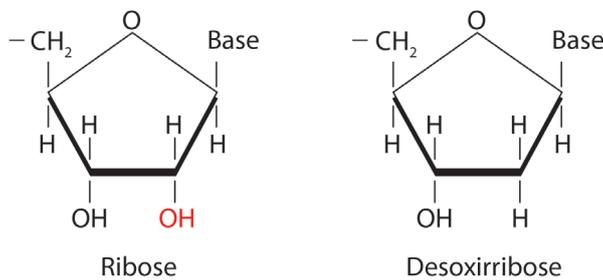


Figura 1. (a) Fórmula estrutural da ribose comparada à da desoxirribose,

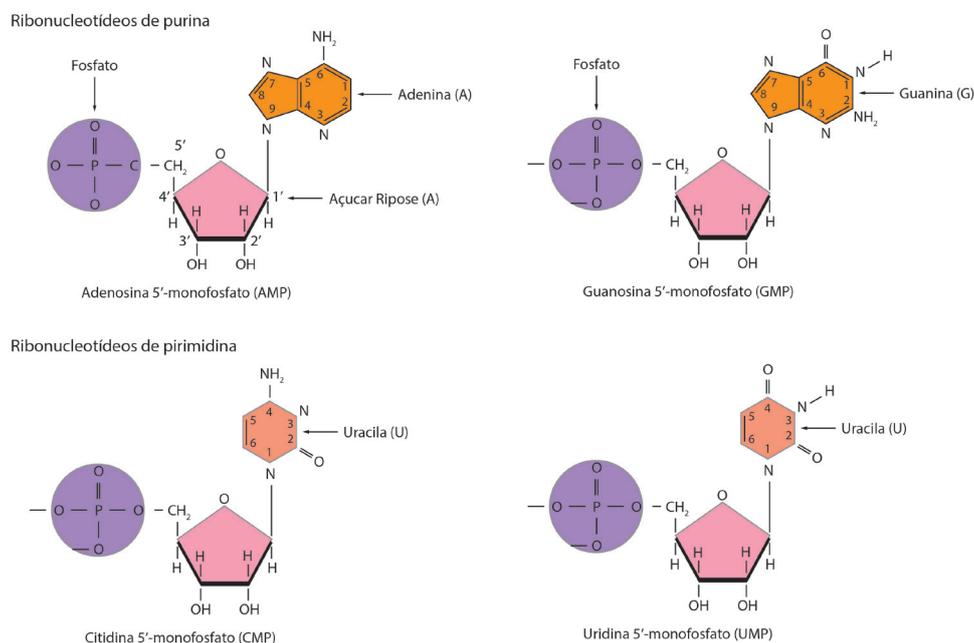
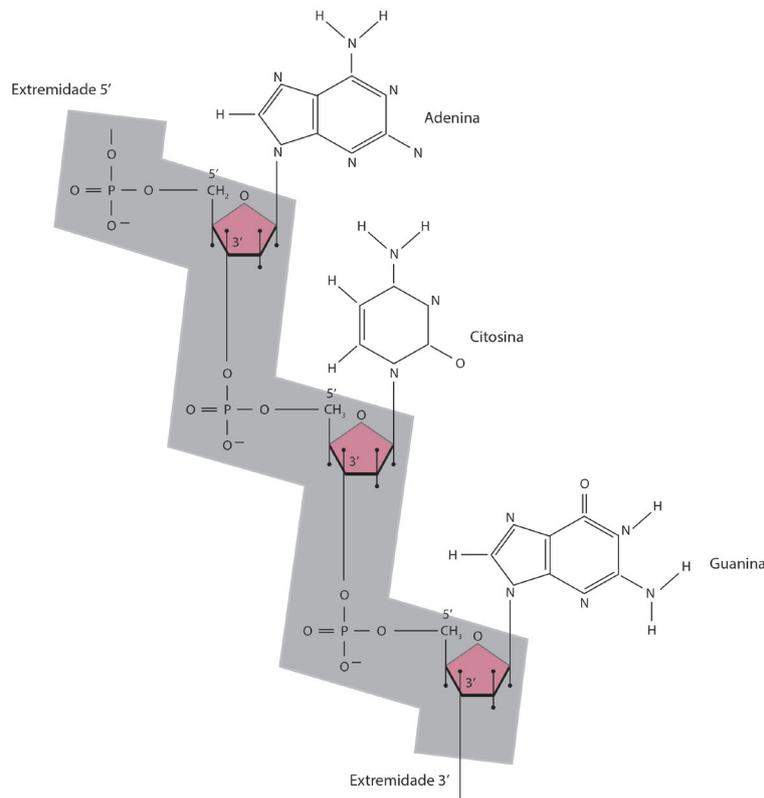


Figura 1. (b) Os quatro ribonucleotídeos encontrados no RNA.



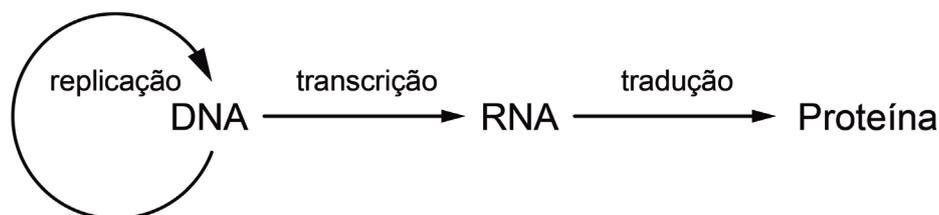
**Figura 1.** (c) Porção de uma molécula de RNA, mostrando as ligações fosfodiéster 3' - 5' que unem os ribonucleotídeos entre si.

A primeira indicação de que o RNA não possuía a estrutura regular de dupla hélice foi a observação de que não havia equivalência entre as quantidades de adenina-uracila e guanina-citosina na molécula. O RNA é basicamente uma molécula de fita simples onde os nucleotídeos estão covalentemente ligados entre si através de ligações fosfodiéster 3' - 5'. Entretanto, quando em solução, o RNA pode apresentar uma arquitetura tridimensional definida e complexa, determinada pelo dobramento da molécula sobre si mesma, formando regiões de dupla hélice em locais onde há sequências de bases complementares. Essa propriedade de formar regiões de dupla-hélice por meio de dobramentos é fundamental para uma série de atividades biológicas executadas pelas moléculas de RNA. Como veremos a seguir, esses dobramentos são muito importantes para a estrutura e função dos RNAs transportadores e ribossômicos.

É importante destacar que, diferente do DNA, moléculas de RNA têm a capacidade de catalisar reações biológicas. O termo **ribozima** é utilizado para descrever moléculas de RNA que funcionam como enzimas.

## O Dogma Central da Biologia

No final de 1953, os pesquisadores adotaram como hipótese de trabalho a idéia de que o DNA atuaria como molde para a síntese do RNA, cujas moléculas, nos eucariontes, migrariam para o citoplasma onde iriam determinar o arranjo dos aminoácidos nas proteínas. Este esquema para o fluxo de informação genética, abaixo esquematizado, foi denominado **dogma central** da Biologia, por Francis Crick em 1956.



**Figura 2.** O dogma central da biologia. A informação genética é perpetuada por meio da replicação do DNA. Essa informação é expressa em um processo de dois estágios: **1.** A transcrição origina um RNA de fita simples usando a molécula de DNA como molde, cuja sequência de bases é idêntica a uma das fitas desse DNA; **2.** A tradução converte a sequência de bases do DNA na sequência de aminoácidos de uma proteína.

Assim, o dogma central define o paradigma da Biologia Molecular, ou seja, que os genes são perpetuados como sequências de ácido nucleico e se expressam na forma de síntese de proteínas.

Sem dúvida, o dogma central da biologia precisou sofrer ajustes no que diz respeito ao fluxo da transmissão da informação genética. A descoberta das enzimas virais, as transcriptases reversas, capazes de sintetizar moléculas de DNA a partir de moldes de RNA, mostrou que não havia uma única maneira de se perpetuar e expressar a informação genética. Além disso, há classes de moléculas de RNA que não são traduzidas (ver item “As Classes de RNA”), mas que são importantes no funcionamento da célula. Contudo, o dogma central ainda corresponde ao resumo dos principais mecanismos de funcionamento gênico nas células.

## As Classes de RNA

Alguns autores agrupam os tipos de RNA em duas classes gerais, sendo uma a que contém a informação genética para a síntese de proteína, o **RNA mensageiro (RNAm)**. As demais classes de RNA são chamadas de **RNA funcionais**, pois essas moléculas exercem suas funções biológicas na forma de RNA, ou seja, o próprio RNA é o produto funcional final. São muitos os tipos de RNA funcional. Os mais conhecidos são o **RNA ribossômico (RNAr)** e o **RNA transportador (RNAt)**. Como veremos a seguir, as moléculas de **RNAr** são importantes constituintes estruturais e componentes catalíticos do **ribossomo**, atuando no processo de síntese de proteína. **Os RNAt** transportam aminoácidos do citoplasma para o RNAm nos ribossomos, no processo de tradução. Alguns outros RNA participam do processo de recomposição (*splicing*) do RNAm e são chamados de **pequenos RNA nucleares (snRNA)**. As demais classes de RNA, mais recentemente descobertas, como por exemplo, **microRNA**, curtos RNA de interferência (**siRNA**) e RNA longos não codificadores (**lncRNA**) são importantes na regulação da expressão gênica e na inibição da atividade de elementos de transposição.

## As Polimerases do RNA

As polimerases do RNA são enzimas que catalisam a síntese de RNA tendo como molde uma fita de DNA. Esse processo é denominado de **transcrição**. Ao contrário das DNA polimerases, as polimerases do RNA são capazes de iniciar uma nova cadeia de RNA a partir de um molde de DNA, sem precisar de iniciador (*primer*).

Os procariontes possuem apenas um tipo de polimerase do RNA, enquanto os eucariontes possuem três. As polimerases do RNA são, em geral, moléculas complexas formadas por múltiplas cadeias polipeptídicas e têm massa ao redor de 500 mil Dalton.

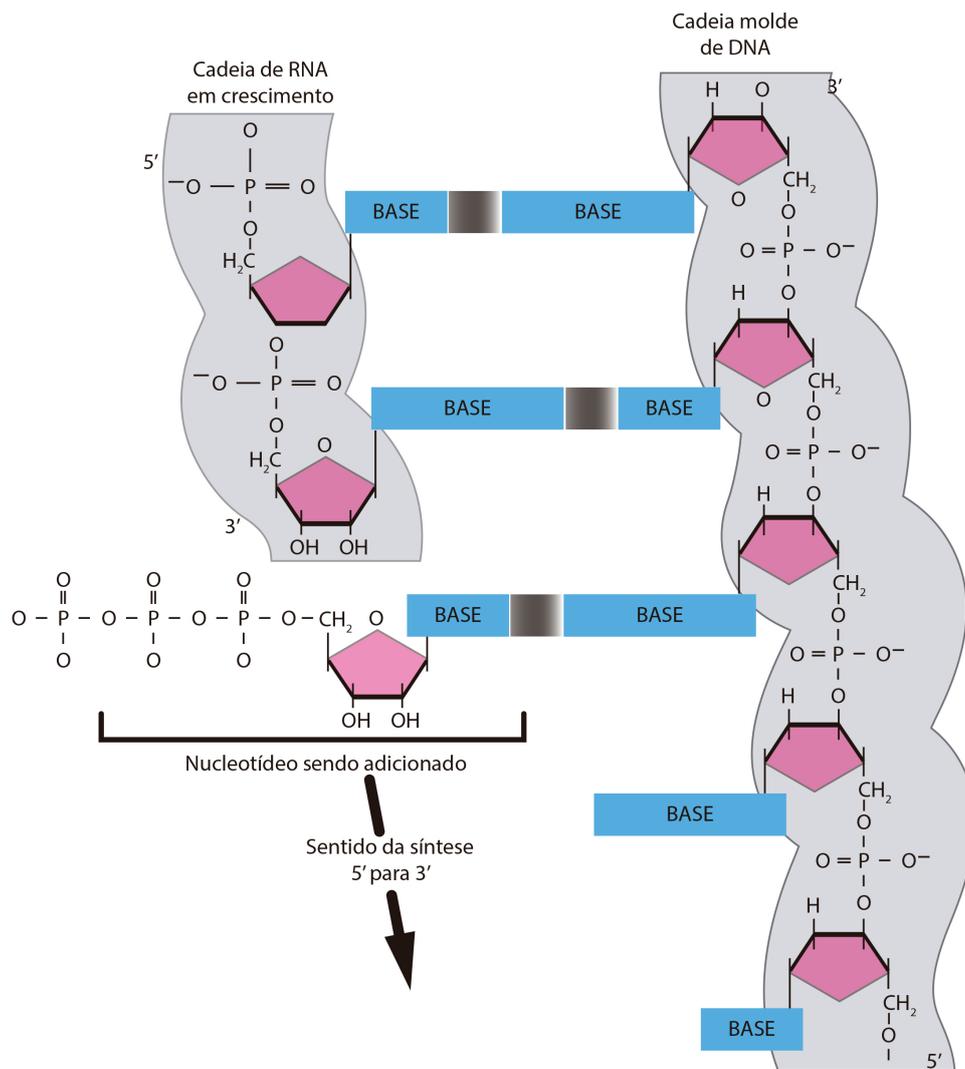
### A polimerase do RNA das bactérias

O núcleo enzimático da polimerase do RNA de *E. coli* contém quatro tipos de subunidades, alfa ( $\alpha$ ), beta ( $\beta$ ), beta' ( $\beta'$ ) e ômega ( $\omega$ ), sendo a subunidade alfa presente em duas cópias. A enzima completa, ou seja, a holoenzima, pode ser separada em dois componentes, um núcleo enzimático (do inglês *core*), responsável pela polimerização dos ribonucleotídeos trifosfatados (que inclui as cinco unidades descritas acima), e o **fator sigma ( $\sigma$ )**, necessário apenas para o reconhecimento do local de início da transcrição. A parte central (núcleo enzimático) da RNA polimerase é capaz de iniciar a transcrição *in vitro* em qualquer ponto de uma molécula de DNA. Porém, nas células, ela só inicia a transcrição no local correto, ou seja, nos promotores, se o fator sigma estiver presente.

Quando uma cadeia de RNA alcança cerca de 8-10 bases, o fator sigma se desprende do núcleo enzimático da polimerase, o qual continua a síntese do RNA tendo o DNA como molde (Fig. 3).

### As polimerases do RNA dos eucariontes

Embora o mecanismo de transcrição nos eucariontes seja basicamente o mesmo que nos procariontes (Figura 3), a maquinaria de síntese de RNA nas células nucleadas é consideravelmente mais complexa do que nas bactérias. Nos eucariontes há três tipos de polimerase do RNA, cada uma delas responsável pela transcrição de um conjunto específico de genes. Essas enzimas são estruturalmente semelhantes entre si e são denominadas **polimerases I, II e III do RNA**. Todas elas são mais complexas que a polimerase do RNA procariótica, sendo formadas por no mínimo 12 subunidades, sendo que algumas das subunidades são comuns a todas elas.



**Figura 3.** À esquerda, reação de alongamento de uma molécula de RNA catalisada pela polimerase do RNA. Em cada passo, o nucleotídeo trifosfatado a ser incorporado é selecionado em função de sua capacidade de se emparelhar com a base correspondente do molde de DNA. Observe que o sentido da síntese é o mesmo da replicação do DNA, ou seja, de 5' para 3'.

Enquanto a polimerase do RNA de *E. coli* é capaz de reconhecer e se ligar ao sítio promotor, iniciando a transcrição, as polimerases do RNA dos eucariontes necessitam da ajuda de várias proteínas adicionais chamadas **fatores de transcrição basais ou fatores gerais de transcrição (GTFs)**, as quais devem estar previamente ligadas ao promotor para possibilitar o posicionamento correto da polimerase e a iniciação eficiente da transcrição.

A polimerase II transcreve os genes cujos RNAs serão traduzidos em proteínas (RNAm), precursores de microRNA e lncRNA. As outras duas polimerases transcrevem os genes correspondentes a moléculas de RNA que têm função estrutural ou catalítica, como as que fazem parte da maquinaria de síntese de proteínas (RNA funcionais). A polimerase I sintetiza a molécula precursora de três dos quatro tipos de RNAr (28S, 18S e 5,8S) e a polimerase III sintetiza o RNAr 5S, os RNAt e uma grande variedade de snRNA.

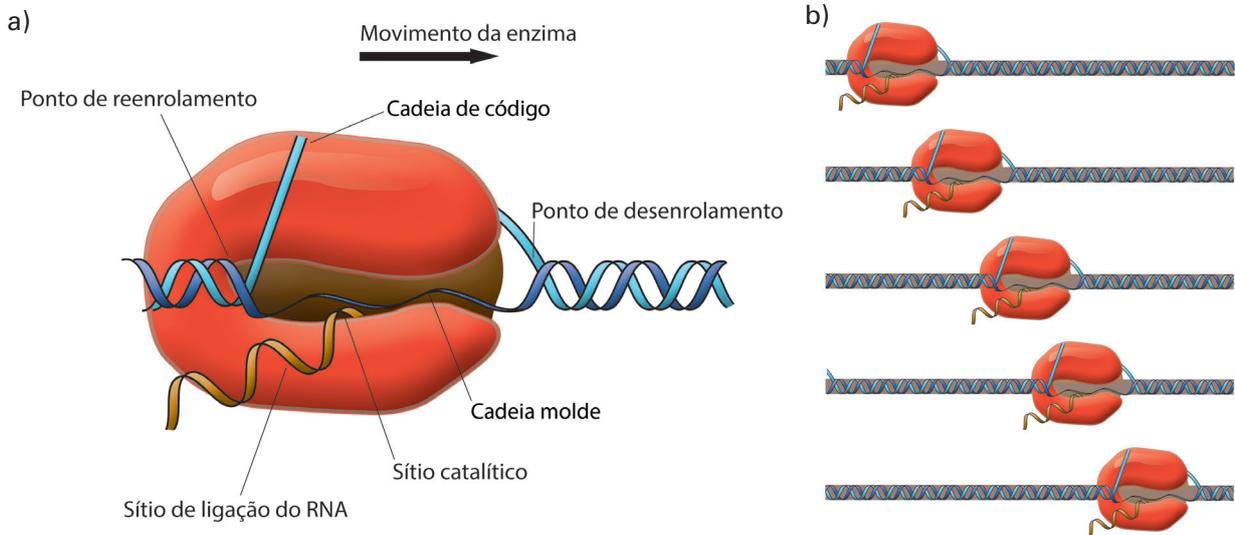
As células contêm grande quantidade de moléculas de polimerase do RNA. As de mamífero, por exemplo, contêm entre 20.000 e 40.000 moléculas de cada uma das polimerases do RNA, sendo que suas concentrações são reguladas individualmente de acordo com a taxa de crescimento da célula.

## A Transcrição do RNA

A transcrição do DNA em moléculas de RNA é um processo altamente seletivo, pois apenas uma pequena porção dos genes é copiada em RNA em um dado momento. Essa seletividade é de importância fundamental, uma vez que a transcrição é o primeiro estágio da expressão gênica e o passo principal em que essa expressão é controlada. O passo inicial na regulação de um gene, e às vezes o único, é a decisão se ele será transcrito ou não.

O processo de transcrição é baseado na complementaridade de bases. A síntese de RNA ocorre dentro da chamada bolha de transcrição, uma estrutura na qual cerca de 18 pares de bases do DNA são temporariamente

separados e uma das fitas da dupla hélice serve como molde para o RNA que está se formando. À medida que a polimerase do RNA se move, a bolha de transcrição move-se com ela, e a cadeia de RNA aumenta em tamanho. Após a passagem da polimerase, a cadeia de RNA se solta de seu molde e as duas cadeias do DNA voltam a se emparelhar, restabelecendo as pontes de hidrogênio rompidas durante a transcrição (Fig. 4).



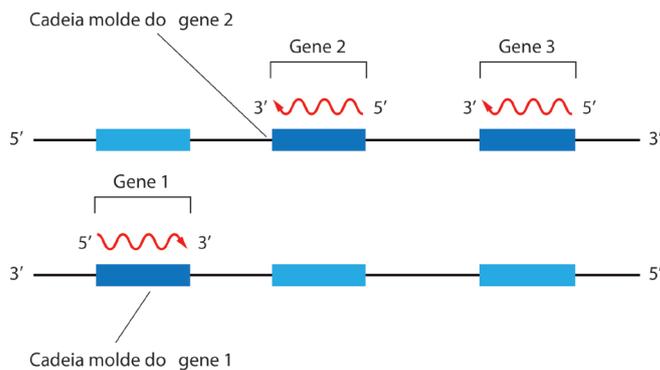
**Figura 4.** Em (a), a representação esquemática da organização do DNA na bolha de transcrição, onde uma das cadeias (cadeia molde) serve de molde para a transcrição da molécula de RNA. Em (b), a representação do processo de transcrição do DNA, à medida que a polimerase se desloca sobre o DNA.

O processo de transcrição tem início quando a polimerase do RNA liga-se a uma sequência especial de DNA, denominada **promotor**, localizada no início de um gene. Nessa sequência de bases existe uma região especial, denominada **sítio de iniciação**, que contém a primeira base do DNA a ser transcrita em RNA (chamada de +1). A partir desse ponto, a polimerase do RNA move-se ao longo do molde, sintetizando RNA, até alcançar uma **sequência de terminalização da transcrição**.

A sequência de DNA transcrita em uma única molécula de RNA, que teve início no promotor e término na região terminalizadora, constitui uma **unidade de transcrição**. Nos eucariontes uma unidade de transcrição é, em geral, composta por um único gene. Já nos procariontes, uma unidade de transcrição contém, em geral, instrução para a síntese de diversas cadeias polipeptídicas, ou seja, contém vários cistrons. O RNA codificado por tal unidade de transcrição é denominado **policistrônico**.

### Terminologia referente à transcrição

Consideremos a transcrição de um segmento de DNA que corresponde a um gene. As duas cadeias da dupla-hélice do DNA se separam e uma das cadeias separadas atua como **molde** para síntese do RNA. Em um cromossomo, em geral, ambas as cadeias do DNA podem ser usadas como molde, mas **em um determinado gene**, apenas uma das cadeias é utilizada como molde e, para esse gene, será sempre essa cadeia. Em outro gene, a outra cadeia pode ser utilizada como molde (Figura 5).



**Figura 5.** São apresentadas duas cadeias de DNA complementares e ilustrados três genes. Para cada gene, apenas uma das cadeias do DNA é usada como molde para a transcrição, mas essa cadeia pode variar entre genes. O sentido da transcrição será sempre o mesmo para qualquer gene e começa sempre na extremidade 3' da cadeia molde (5' do RNA recém-transcrito) e termina na ponta 5' do DNA molde (3' do RNA recém-transcrito). Assim, a cadeia de RNA cresce sempre do sentido 5' para o 3'.

Desse modo, em um gene, a cadeia do DNA que possui a mesma sequência de bases do RNA, exceto por possuir T ao invés de U, é denominada **cadeia de código** (*coding strand*). A outra cadeia do DNA, que serviu de molde para a síntese do RNA, é chamada **cadeia molde** (*template strand*). Assim, a sequência do RNA é a

mesma da fita molde do DNA (Fig. 6). É uma convenção que a sequência nucleotídica dos genes publicada em trabalhos científicos e depositada em bancos de dados tenha a sequência nucleotídica escrita como se apresenta na cadeia de código.

**DNA**

Cadeia de código

5' ...CCAGTACCTTGGATGACGATTAACGC...3'

3' ...GGTCATGGAACCTACTGCTAATTGCG...5'

Cadeia molde

**RNA**

5' ...CCAGUACCUUGGAUGACGAUUAACGC.3'

Figura 6. A molécula de RNA é complementar à cadeia de DNA a partir da qual ela foi sintetizada.

Em um segmento de DNA, as sequências localizadas anteriormente ao sítio de início de transcrição (ou *upstream*) são denominadas 5' e aquelas localizadas após o sítio de início (dentro da região transcrita ou após ela) são referidas como 3' (ou *downstream*).

As posições das bases em uma unidade de transcrição são numeradas em ambas as direções a partir do ponto de início da transcrição, a qual é assinalada como +1; os números na direção 5' são precedidos do sinal -, e aumentam à medida que se distanciam do sítio de início. A base posicionada imediatamente antes (direção 5') do sítio de início é numerada -1, enquanto a base imediatamente depois do sítio de início corresponde a +2 (direção 3').

Etapas da transcrição em procariontes

Reconhecimento do molde de DNA

As moléculas livres de polimerases do RNA colidem ao acaso com o DNA podendo ligar-se a ele, porém muito fracamente. A polimerase do RNA liga-se fortemente ao DNA apenas quando ela contacta uma região promotora. Nesse caso, a holoenzima liga-se fortemente ao promotor causando o desemparelhamento da dupla hélice (Fig.7).

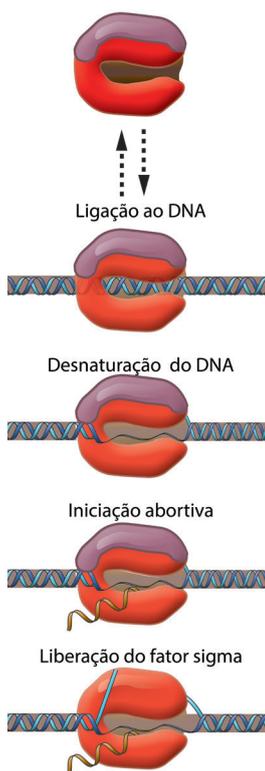


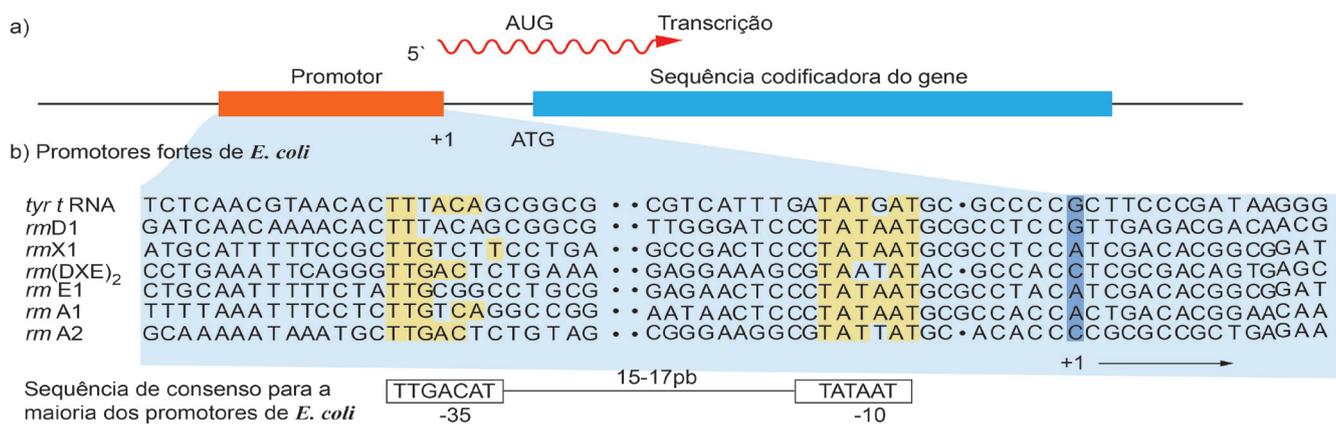
Figura 7. Representação esquemática do início da transcrição. A polimerase (em vermelho) reconhece um sítio promotor e se liga fortemente a ele. As hélices do DNA se separam, a transcrição tem início e o fator sigma (em roxo) é liberado.

Toda região promotora contém uma sequência de bases específica importante para marcar o local de início de síntese do RNA. No genoma das bactérias, a sequência mínima com capacidade de sinalizar o local de início de transcrição tem 12 pares de bases, subdivididos em dois módulos.

O sequenciamento de bases de diferentes promotores bacterianos mostrou que há sequências básicas que estão presente em todos eles. Por essa razão, elas são chamadas de **sequências consenso**. Essas sequências, no entanto, não possuem todas as bases iguais, havendo alguma variação entre elas. Assim, pode-se perguntar como as sequências devem ser analisadas para se determinar se elas são suficientemente semelhantes pra constituir um sinal reconhecível?

As sequências de DNA sinalizadoras podem ser definidas em termos de uma sequência ideal que representa cada uma das bases que estão presentes com maior frequência em cada uma das posições, mas não obrigatoriamente em todas as sequências. Todas as sequências conhecidas são então alinhadas para comparação. Para que uma sequência seja aceita como consenso, cada base em particular deve ser predominante em sua posição e a maioria dos exemplos conhecidos deve relacionar-se com o consenso possuindo apenas poucas substituições, não mais do que uma ou duas.

Em um promotor bacteriano há quatro características conservadas: (1) o **sítio de início de transcrição**, que em geral é uma purina, denominado de +1; (2) o **TATA box**, que é uma região de seis pares de bases ao redor do sítio -10 cuja sequência consenso é TATAAT; (3) uma sequência consenso TTGACA, localizada ao redor do sítio -35; (4) a distância entre os sítios -10 e -35, correspondente a 16 - 18 pares de bases, que é crítica para a ligação da polimerase do RNA (Figura 8).



**Figura 8.** Algumas sequências de bases de promotores da bactéria *E. coli*. A numeração é feita de acordo com o número de bases antes (-) ou após (+) o ponto de início da síntese do RNA (+1). Observe que a sequência codificadora do gene se inicia com a trinca ATG, que corresponde ao códon AUG no RNAm. Note a presença de regiões consenso, indicada nas caixas, ao redor da região -10 e da região -35.

A conservação de curtas sequências consenso em sítios reguladores do gene é uma característica tanto de procariontes quanto de eucariontes, mas as sequências e o número de sequências de consenso são diferentes entre procariontes e eucariontes.

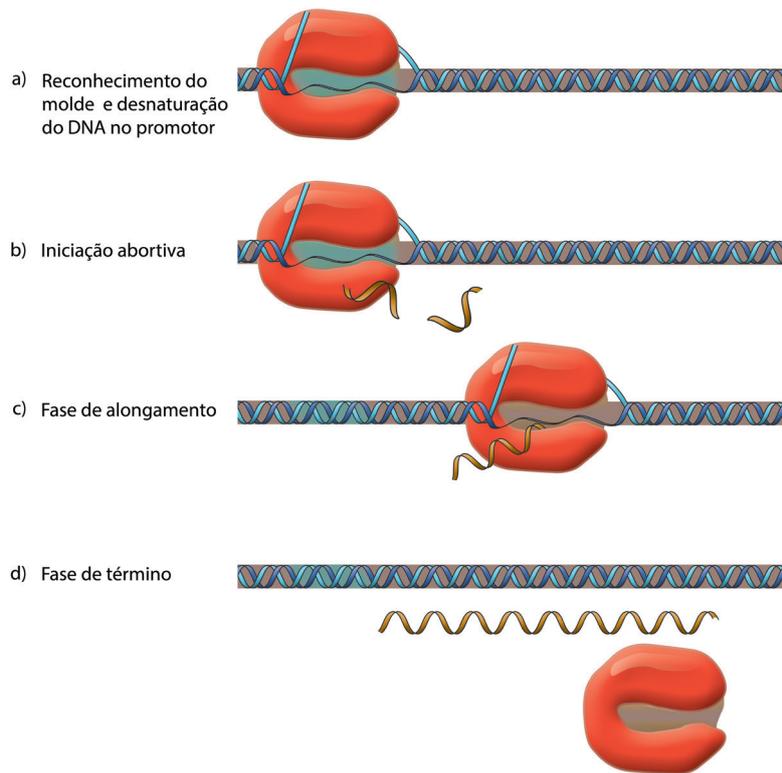
### Início da transcrição

A transcrição tem início quando um primeiro nucleotídeo trifosfatado é posicionado sobre a cadeia molde de DNA, por ação da polimerase do RNA.

A polimerase do RNA permanece ligada à região promotora enquanto são adicionados os primeiros nucleotídeos da cadeia de RNA sendo sintetizada. Essa fase pode ser prolongada pela ocorrência de eventos abortivos, nos quais a enzima sintetiza transcritos de tamanho inferior a dez nucleotídeos. Essa fase inicial, chamada **fase de iniciação**, termina quando a polimerase consegue estender a síntese além desse tamanho e prossegue para a fase de alongamento (Figura 9).

### Fase de alongamento

Após a síntese de cerca de nove a dez nucleotídeos da molécula de RNA, a subunidade sigma da polimerase do RNA dissocia-se e diversos **fatores de alongamento** se associam à polimerase, que passa a se deslocar ao longo do DNA, alongando a molécula de RNA. Essa é a **fase de alongamento** (Figura 9).



**Figura 9.** As quatro etapas (A - D) em que o processo de transcrição costuma ser dividido estão indicadas e envolvem diferentes tipos de interação entre o DNA e a polimerase do RNA.

À medida que a polimerase do RNA se move ao longo do DNA, ela desenrola a hélice, expondo um novo segmento da cadeia molde. Os nucleotídeos vão sendo covalentemente unidos à extremidade 3' da cadeia de RNA em crescimento, formando um híbrido DNA/RNA na região desenrolada do DNA. Esse híbrido tem uma extensão de cerca de oito ou nove nucleotídeos da cadeia crescente de RNA. Imediatamente atrás do local onde está ocorrendo a síntese, após a polimerase passar, a região do DNA já copiada volta a se emparelhar, refazendo a dupla hélice. O RNA, que vai se soltando do DNA, emerge finalmente como uma fita simples e livre.

A velocidade de síntese do RNA foi estimada em cerca de 40 nucleotídeos por segundo a 37°C, para a polimerase de bactéria. Essa síntese se dá, como no caso do DNA, pelo ataque nucleofílico da extremidade 3'OH do último nucleotídeo incorporado sobre a ligação do fosfato alfa de um ribonucleotídeo 5' trifosfatado emparelhado à cadeia molde de DNA. O nucleotídeo a ser incorporado perde seus dois grupos fosfatos terminais, liberados na forma de pirofosfato.

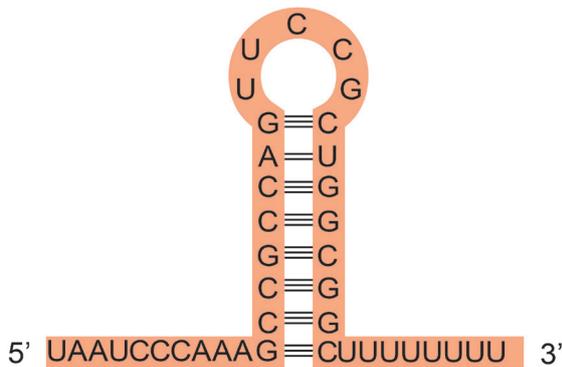
A etapa final da transcrição, chamada **fase de término**, envolve o reconhecimento de uma sequência sinalizadora, que indica que a transcrição deve parar. Quando a última base é adicionada ao RNA, a bolha de transcrição colapsa e o híbrido DNA/RNA é desfeito, as duas cadeias do DNA finalizam seu emparelhamento e o RNA e a enzima se soltam (Fig. 9).

A sequência de DNA que sinaliza o término da transcrição é chamada **região terminalizadora**.

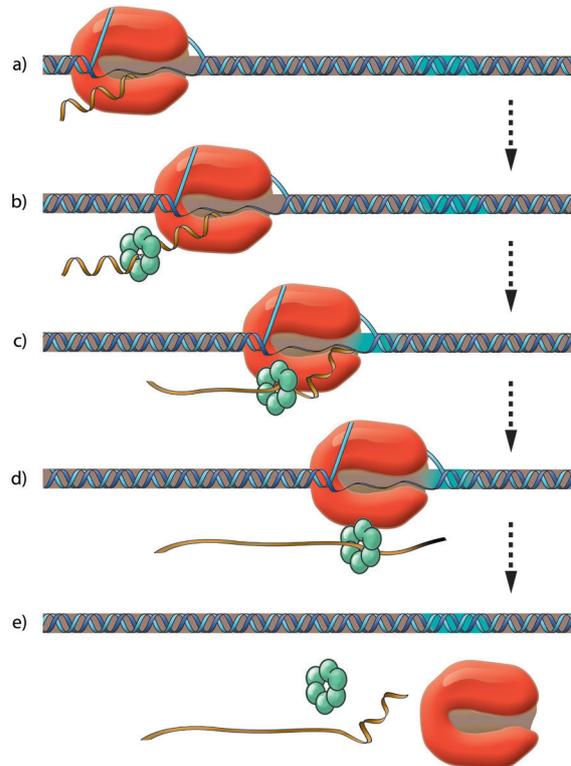
#### O término da transcrição

Na bactéria *E. coli* há dois mecanismos básicos de término de transcrição, ou seja, dois tipos de terminadores. No primeiro deles, a terminação é direta (intrínseca) e a sequência de DNA que sinaliza o término da transcrição contém cerca de 40 bases, sendo que na extremidade 3' desta sequência nucleotídica existe um segmento rico em C e G, seguido por seis ou mais timinas. No RNA transcrito, a sequência rica em C e G fica arranjada de modo a que a molécula de RNA possa formar, nessa região, uma alça em forma de grampo de cabelo (*hairpin*). Na molécula de RNA, essa alça é seguida por uma série de Us que são complementares aos resíduos de adenina da cadeia molde de DNA. A estrutura em forma de grampo, juntamente com a sequência de Us, provoca a parada e o despreendimento da polimerase do RNA com consequente término da transcrição (Fig. 10). Supõe-se que o grampo empurre a polimerase para a frente do RNA e do DNA e contribua para dissociar o RNA transcrito da polimerase.

O segundo tipo de término da transcrição tem a participação de uma proteína denominada **fator Rho** e é por isso chamado de **dependente de Rho**. Esse fator rho é uma proteína com aproximadamente 40 mil Daltons, mas atua na forma de um hexâmero com 275 mil Daltons. Essa proteína em forma de anel tem seis unidades idênticas e se liga ao RNA de fita simples assim que ele deixa a RNA polimerase. Acredita-se que o fator rho se ligue à molécula nascente de RNA em sítios com sequências específicas, chamados de sítios **rut** (Fig. 11). Esses sítios consistem em segmentos de cerca de 40 nucleotídeos que permanecem como fita simples no RNA. Há duas hipóteses sobre como o fator rho atua: ou ele empurra a polimerase para frente, liberando o RNA transcrito da polimerase ou induz uma modificação conformacional na polimerase, levando-a a pausar.



**Figura 10.** Os sinalizadores do término de transcrição incluem regiões palindrômicas, ricas em citosinas e guaninas, que formam alças tipo grampo de cabelo (*hairpins*) no RNA, as quais variam em comprimento entre 7 e 20 pares de bases, e uma sequência terminal rica em uracilas.



**Figura 11.** Esquema de um dos modelos do processo de terminação mediada pelo fator Rho. A polimerase do RNA transcreve o DNA (a). Ao surgir o sítio rut no transcrito, Rho liga-se a ele (b). A polimerase pausa e rho desenrola o híbrido DNA-RNA (d). Com o término da transcrição, todos os componentes são liberados (e).

Na ausência da polimerase do RNA, o fator Rho tem a capacidade de desfazer um híbrido DNA/RNA, hidrolisando ATP nessa reação. Isso sugere que o fator rho possa interagir diretamente com o segmento híbrido DNA/RNA dentro da bolha de transcrição, ocasionando o término da transcrição, ou como consequência do desemparelhamento ou por interação direta de Rho com a polimerase do RNA (Fig. 11). Estudos recentes sugerem que Rho se liga à polimerase durante todo o ciclo de transcrição, e não somente durante a etapa de terminação.

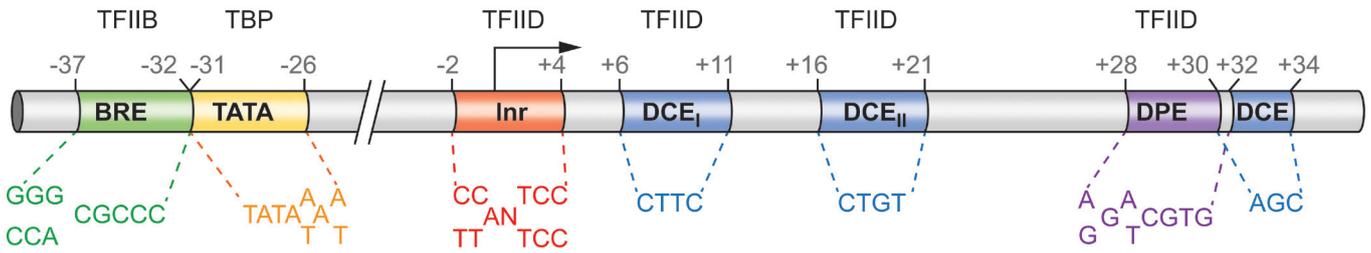
## A Transcrição em Eucariontes

### O início da transcrição em eucariontes

Conforme já explicado, a tarefa da transcrição dos eucariontes está dividida entre os três tipos de polimerase do RNA, chamadas de I, II e III. Os eventos que descreveremos a seguir referem-se ao mecanismo de início da transcrição executada pela polimerase do RNA II, aquela que sintetiza os RNAm.

Os genomas eucarióticos são grandes e complexos. Em organismos complexos, como os humanos, é uma fração reduzida do genoma que corresponde aos genes que serão transcritos em RNAm, e que ao serem traduzidos resultam em proteínas. A densidade de genes nos cromossomos é muito baixa em alguns desses organismos. Por isso, identificar o local correto de início da transcrição de um gene pode não ser trivial e regular a intensidade de transcrição de cada gene em um organismo multicelular torna-se uma tarefa com certeza sofisticada. Por isso, é de certo modo esperado que os eventos que culminam na transcrição eucariótica sejam complicados, assim como as sequências nos promotores.

A primeira diferença importante diz respeito à estrutura da região promotora eucariótica. Enquanto em bactérias todos os promotores apresentam as mesmas seqüências consenso nas posições -10 e -35, os promotores eucarióticos são variados no que diz respeito ao número e aos tipos de seqüências consenso que podem ocorrer em cada um deles. A figura 12 mostra um esquema de um promotor eucariótico hipotético.

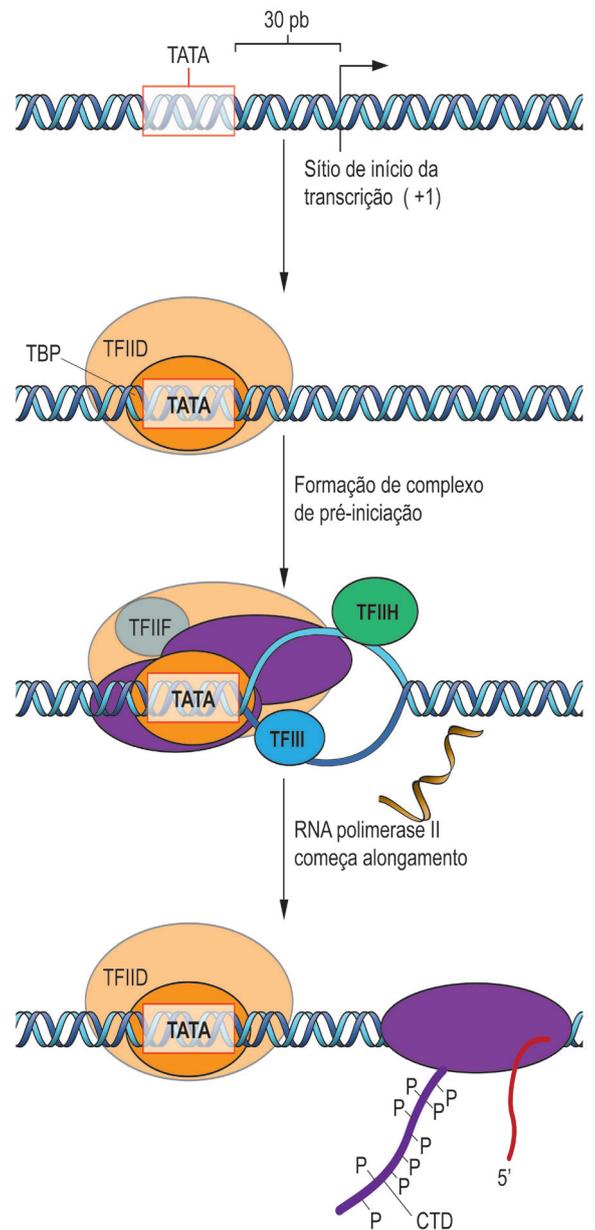


**Figura 12.** Esquema de um exemplo de promotor eucariótico. A figura mostra diversos elementos da região promotora e suas posições em relação ao sítio de início da transcrição. Esses elementos são os seguintes: elemento de reconhecimento de TFIIB (BRE), a seqüência TATA, o elemento iniciador (Inr), elemento promotor à jusante (DPE), o elemento central à montante (DCE). Estão também mostradas as seqüências de cada consenso e o nome do fator geral de transcrição geral que reconhece cada um deles.

A segunda diferença importante na iniciação eucariótica, quando comparada à procariótica, é a necessidade de ligação ao DNA de várias proteínas na região promotora, antes que a RNA polimerase possa se ligar. A polimerase não se liga diretamente ao DNA, mas liga-se a essas proteínas, somente depois que elas estiverem posicionadas corretamente na região promotora. Essas proteínas são **fatores de transcrição basais ou fatores gerais de transcrição (GTFs)**. Seu papel é de atrair a RNA polimerase II e posicioná-la no local correto onde deve ocorrer o início da transcrição. TFIIA, TFIIB são os nomes de alguns desses GTFs. Os GTFs acoplados à RNA polimerase constituem o chamado complexo de pré-iniciação. Esse complexo pode ser muito grande, pois é constituído de vários GTFs, cada um correspondendo a um complexo multiproteico, além do núcleo de RNA polimerase.

Quando se alinha a seqüência nucleotídica de promotores eucarióticos, observam-se seqüências conservadas. Uma delas é o TATA Box, cuja seqüência é rica em A e T e se localiza ao redor da posição -30. O consenso TATA Box é reconhecido por uma proteína chamada de TBP (*tata binding protein*), parte do complexo TFIID, um dos fatores gerais de transcrição. Um esquema da iniciação da transcrição em um promotor que contém TATA está apresentado na Figura 13. Deve ser lembrado, no entanto, que o esquema representa um exemplo de início da transcrição e que não são todos os promotores que contêm o TATA Box. Os promotores eucarióticos variam na extensão e na composição, no que diz respeito aos tipos de seqüências consensos que apresentam.

A RNA polimerase II contém uma cauda proteica, chamada CTD (*carboxi-terminal domain*) localizada estrategicamente no local onde emerge o RNA recém-transcrito. A fosforilação dessa cauda está relacionada ao término da fase de início e à transição para a fase de alongamento do transcrito. Conforme veremos a seguir, a cauda CTD é muito importante em outras fases da síntese e amadurecimento do RNAm.



**Figura 13.** Exemplo do processo de início da transcrição em um promotor eucariótico que contém TATA Box.

## O RNA mensageiro dos eucariontes

Os RNA que contêm a informação para a síntese de proteínas são denominados **RNA mensageiros (RNAm)**. Há importantes diferenças nos detalhes da síntese e da estrutura dos RNA mensageiros de pro e eucariontes. Nos procariontes, o RNAm é transcrito e traduzido no único compartimento celular e os dois processos ocorrem acopladamente. Na célula eucarionte, a síntese e a maturação dos RNAm ocorrem exclusivamente no núcleo. Apenas quando o RNAm está maduro ele é exportado para o citoplasma e traduzido.

Diferenças significativas na tradução dos RNAm de bactérias e de eucariontes são devidas a suas características estruturais e estabilidade. A diferença estrutural é que o RNAm de bactérias frequentemente codifica para várias proteínas, sendo chamado de **policistrônico**, enquanto o RNAm de eucariontes, na maioria dos casos, codifica para apenas uma cadeia polipeptídica, sendo chamado **monocistrônico** (Fig. 14). Uma diferença funcional é que o RNAm de bactérias geralmente é instável, e é por isso traduzido em proteínas durante um período de tempo muito curto, tipicamente poucos minutos. O RNAm dos eucariontes é mais estável e pode continuar a ser traduzido por várias horas ou mesmo dias.

São necessárias várias reações para produzir o RNA mensageiro maduro dos eucariontes. As moléculas mensageiras recém-sintetizadas pela polimerase II são denominadas **transcritos primários**. A população de tais transcritos no núcleo foi originalmente denominada **RNA heterogêneo nuclear (RNAhn)**, por causa da grande variação em tamanho que apresentava, contrastando com os RNAs mais uniformes e de menor tamanho realmente necessários para codificar uma proteína. Muita dessa variação em tamanho no RNA recém-sintetizado é devida à presença de longas sequências de nucleotídeos, denominadas **íntrons**, porque estão intercalados na sequência codificadora, a qual será expressa, formando blocos chamados **éxons**. Íntrons serão retirados do transcrito primário em um processo denominado *splicing* (recomposição). Além disso, outras reações importantes como a adição de uma extremidade **5'CAP** e da cauda de poli-A são fundamentais ao funcionamento do RNAm maduro (Fig. 14). Tem sido usado o termo **processamento** para descrever o conjunto de modificações co-transcricionais aos quais os transcritos primários são submetidos (adição de 5'-CAP, de cauda de Poli-A e *splicing*). Essas modificações que ocorrem no RNA são chamadas co-transcricionais, pois ocorrem imediatamente e simultaneamente à transcrição do RNAm.

A cauda CTD da polimerase eucariótica tem um papel central na coordenação dos eventos de processamento. Essa cauda tem sequência repetidas de sete aminoácidos que são sítios de ligação para proteínas importantes para a adição do CAP, para o *splicing*, clivagem e poliadenilação do RNAm. O estado de fosforilação de resíduos específicos da cauda CTD determina qual reação vai ocorrer com o RNAm a cada momento.

### Íntrons são removidos do pré-RNAm no *splicing*

Para a produção de um RNA mensageiro em células eucarióticas, o segmento inteiro do gene, incluindo íntrons e éxons, é primeiramente transcrito em uma longa molécula de RNA, o transcrito primário ou pré-RNAm. Antes que o RNA deixe o núcleo, todas as sequências correspondentes aos íntrons são retiradas e os éxons são unidos entre si. O resultado é uma molécula mais curta de RNA, que agora contém uma sequência codificadora ininterrupta. Quando esse passo, denominado *splicing* do pré-RNAm, é completado, temos uma molécula de RNAm maduro, funcional, que pode deixar o núcleo e ser traduzida em proteínas.

Como a célula determina quais as partes do transcrito primário devem ser removidas? Diferentemente da sequência de código de um éxon, a sequência exata de nucleotídeos da maioria dos íntrons parece não ser importante para a célula. Embora haja pouca semelhança entre as sequências de nucleotídeos de íntrons diferentes, cada íntron contém uma curta sequência de nucleotídeos que é importante para sua remoção (Fig. 15).

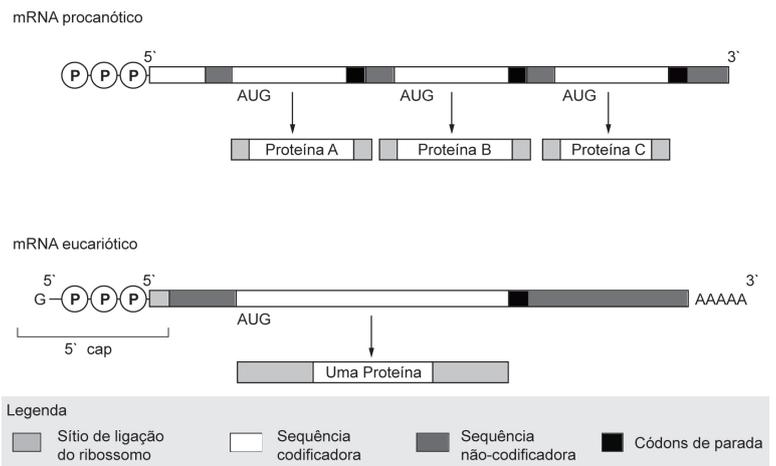
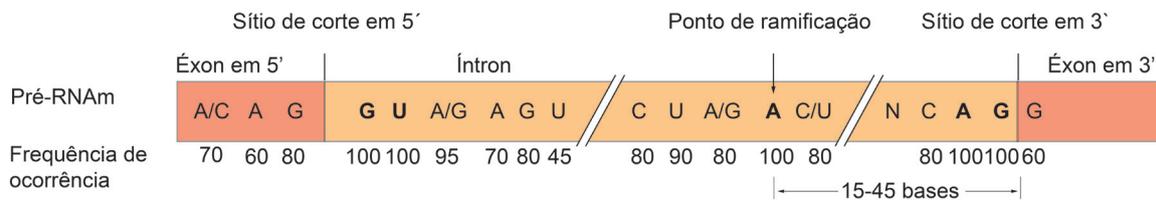
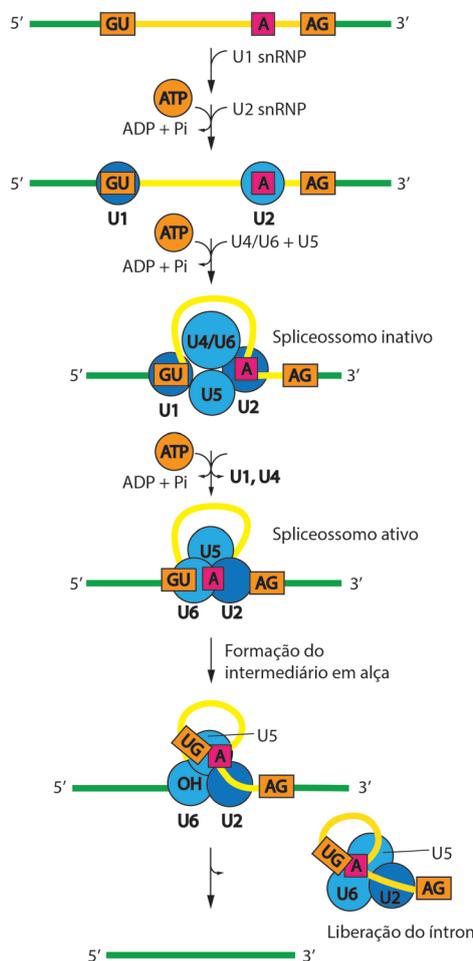


Figura 14. Comparação entre a estrutura de RNA mensageiros de procariontes e eucariontes.

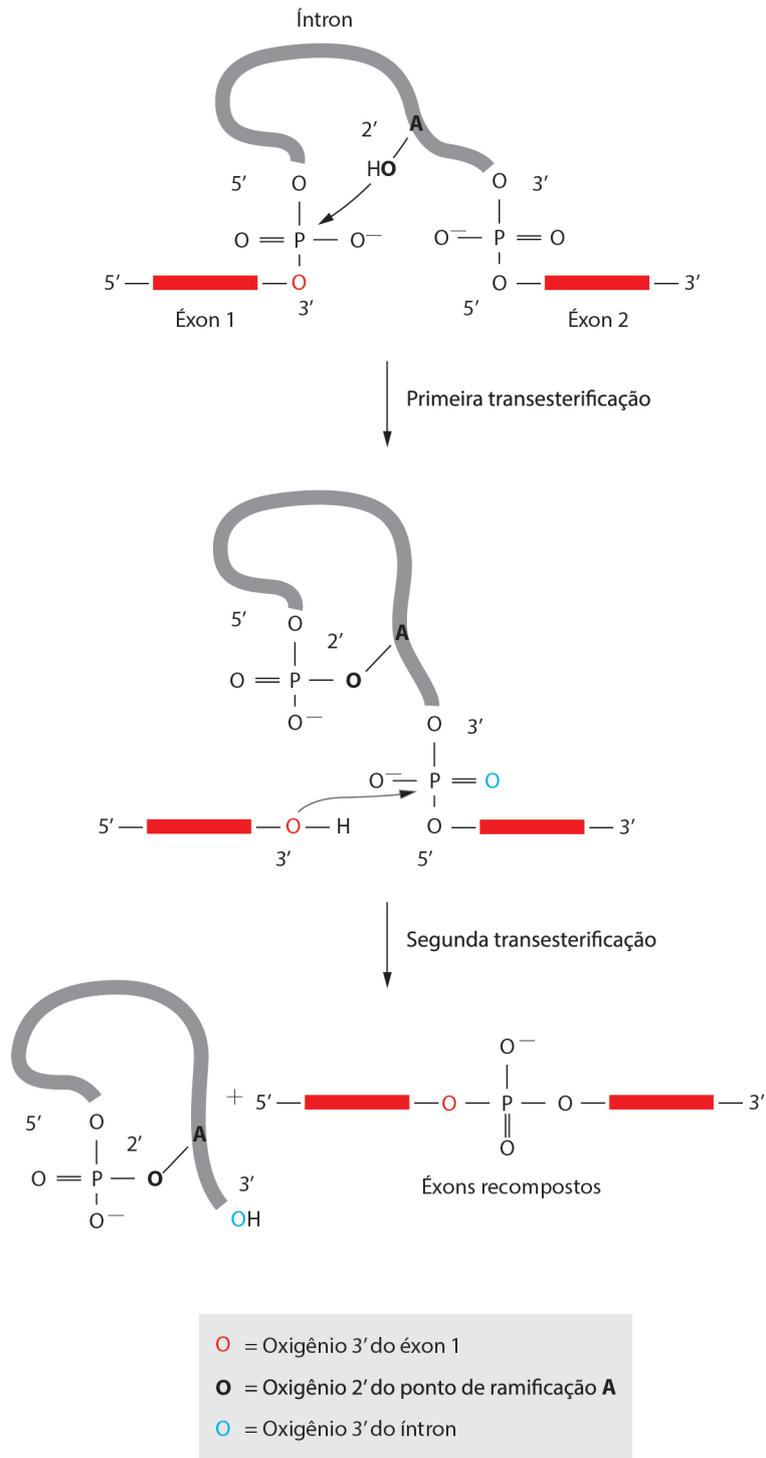


**Figura 15.** Sequências de nucleotídeos que sinalizam o início e o final de um íntron. As três sequências de nucleotídeos mostradas são necessárias para remover um íntron. As outras posições de um íntron podem ser ocupadas por qualquer nucleotídeo. As sequências especiais são reconhecidas por snRNPs, que clivam o RNA nos limites íntron-éxon e ligam covalentemente os éxons. A adenina (A) apontada pela seta forma o ponto de ramificação do laço produzido na reação de *splicing* (Figura 15) e está tipicamente localizada a cerca de 30 nucleotídeos da extremidade 3' do íntron.

Os íntrons são removidos do RNA por enzimas que, ao contrário de muitas outras enzimas, são compostas por um complexo de snRNA e proteína. Essas partículas encarregadas do *splicing* do RNA são denominadas snRNPs ou *snurps* (do inglês - *small nuclear ribonucleoprotein particles*). Há cinco tipos de snRNPs funcionais no *splicing* do pré-RNA. Cada uma carrega uma molécula de snRNA (U1, U2, U4, U5 ou U6) e dezenas de proteínas. Em cada íntron, duas snRNPs (U1 e U2) se associam ao pré-RNA, definindo os limites do íntron. Outras três snRNPs (U4 a U6), em complexo, interagem com as primeiras e o pré-RNA, modificando a configuração de RNAs e proteínas, formando um grande complexo macromolecular, observável à microscopia eletrônica, denominado spliceossomo. Através de atividade de ribozima dos snRNAs, o spliceossomo excisa (cliva) sequencialmente cada extremidade do íntron e o libera como um laço (Fig. 16 e 17). As duas reações químicas de transesterificação que ocorrem sequencialmente durante o *splicing* estão ilustradas na Figura 17. Um dos papéis do RNA presente nos snRNPs (*snurps*) é reconhecer e – por complementaridade de bases – emparelhar com a sequência de nucleotídeos que marcam o início e o ponto de ramificação de cada íntron (veja Fig.15). Desse modo, as *snurps* aproximam as duas extremidades do íntron permitindo a ocorrência do *splicing*. Embora as *snurps* tenham um papel central na reação de *splicing*, outras proteínas são também necessárias.

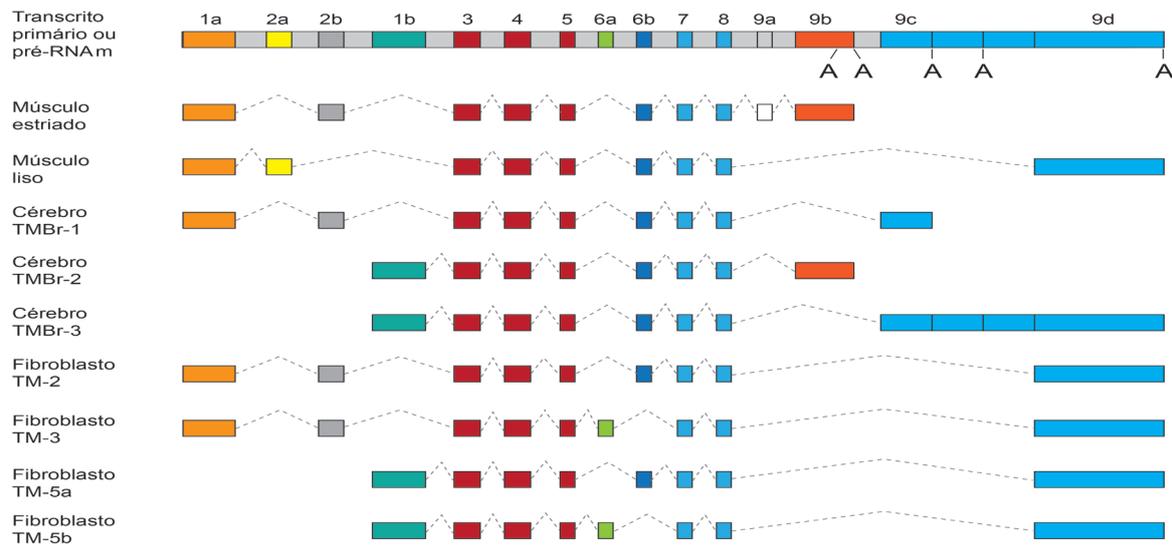


**Figura 16.** Mecanismo geral de *splicing* do pré-RNA. O *splicing* do pré-RNA é catalisado por um conjunto de snRNPs (mostrado como círculos) e de outras proteínas (não mostradas). Uma função do complexo snRNPs (U1 e U2) é definir a sequência do íntron. Após a entrada de outras três snRNPs (U4 a U6), aproximam-se as duas extremidades do íntron de modo a tornar possível a reação havendo a montagem do spliceossomo. Um nucleotídeo específico, que possui a base adenina no ponto de ramificação da sequência do íntron (indicada em rosa) liga-se ao sítio 5' de *splicing* e corta sua junção com o primeiro éxon. A extremidade 5' do íntron torna-se covalentemente ligada ao nucleotídeo adenina formando uma alça ou laço (lariat), na molécula de RNA (ver Figura 17). A extremidade 3' OH livre da sequência do primeiro éxon reage com o início da sequência do segundo éxon, cortando o íntron na sua extremidade livre 3' e, ao mesmo tempo, unindo os dois éxons. O produto dessa reação de *splicing* é a ligação de duas sequências de éxons mantendo a sequência de código contínua e o íntron é liberado em forma de laço e finalmente degradado.



**Figura 17.** Estrutura da cadeia ramificada do pré-RNA que se forma durante o *splicing*. O nucleotídeo apontado pela seta corresponde à adenina do ponto de ramificação, como marcado nas figuras 15 e 16. A ramificação se forma quando a extremidade 5' da sequência do íntron é covalentemente ligada ao grupo 2'OH da ribose da adenina. Esta estrutura não usual foi denominada intermediário em alça/laço/lariat. Esta é a primeira reação de transesterificação. A segunda reação permite a ligação entre os éxons e a liberação do íntron excisado, cuja cadeia ramificada permanece na sequência final (ver figura 16).

A existência do *splicing* do pré-RNA em eucariotos acarreta uma vantagem muito importante. Os transcritos primários de vários genes eucarióticos podem sofrer *splicing* de diferentes maneiras para produzir diferentes RNA (transcritos variáveis), às vezes, dependendo do tipo de célula no qual o gene está sendo expresso ou do estágio de desenvolvimento do organismo. Como consequência, diferentes proteínas podem ser produzidas a partir do mesmo gene (Fig. 18), denominadas isoformas proteicas, caracterizando o processo chamado de ***splicing alternativo***.



**Figura 18** Splicing alternativo do RNAm do gene da  $\alpha$ -tropomiosina de rato. Diferentes padrões de recomposição ocorrem nos diferentes tipos celulares. Os exons estão indicados por um número e uma letra (ex. 1a, 2a, etc.). Os sinais de poliadenilação estão representados pelas letras A. As regiões pontilhadas são removidas pelo *splicing*.

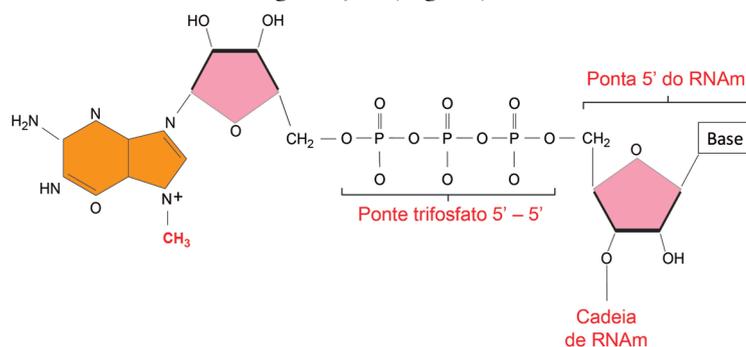
Com o acúmulo de informações sobre os genomas sequenciados de diversas espécies de microrganismos, plantas e animais (incluindo o homem), algumas revelações se mostraram intrigantes. Por exemplo, dada a complexidade do organismo e do proteoma humano, esperava-se, há algumas décadas, que deveríamos ter cerca de 100.000 genes codificadores de proteínas. No entanto, estima-se que temos pouco mais que 21.000 genes desta categoria. Foi muito intrigante avaliar que teríamos número semelhante de genes ao de algumas plantas, insetos ou vermes nematoides. A descoberta fundamental, ocorrida no fim da década de 70, de que os genes dos eucarióticos eram interrompidos, com sequências de éxons intercaladas com sequências dos introns, trouxe a explicação para essa aparente discrepância. Hoje sabemos que um mesmo RNAm primário pode sofrer *splicing* de diversas maneiras distintas. Apesar de termos somente 21.000 genes, são sintetizadas mais de 100.000 tipos de proteínas distintos. Quanto mais se estuda o transcriptoma dos diversos organismos, observa-se que o *splicing* alternativo é um fenômeno muito mais frequente do que inicialmente previsto, caracterizando quase uma regra, e não uma exceção, na expressão gênica dos organismos mais complexos.

Em conclusão, ao invés de ser um processo que parece acarretar desperdício, o *splicing* do RNA dos eucariotos permite o aumento do já enorme potencial de código de seus genomas.

### Os RNAm sofrem outras modificações

Os RNAs sintetizados pela polimerase II também são modificados em ambas as extremidades, ficando muito diferentes dos transcritos produzidos pelas polimerases bacterianas. Essas modificações são realizadas por proteínas que interagem com a cauda CTD da RNA polimerase II. Essas modificações serão utilizadas posteriormente, no citoplasma, como sinais para a estabilidade da molécula e de que esses RNAs devem ser traduzidos em proteínas.

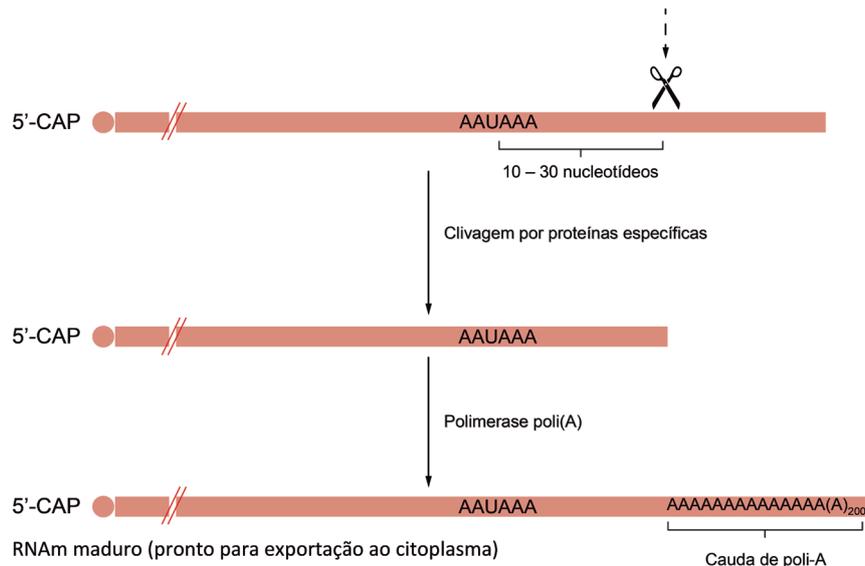
A extremidade 5' da molécula de RNA é “encapada” (do inglês *capped*) pela adição de um nucleotídeo contendo G, metilado. O encapamento ocorre no início da transcrição, quando o transcrito ainda é pequeno com poucos nucleotídeos de comprimento. Essa reação se dá pela ligação de uma molécula de 7-metilguanósina à extremidade 5' do transcrito. A extremidade 5' *cap* tem um importante papel nos passos iniciais da tradução do RNAm, além de proteger o RNA nascente de degradação (Fig. 19).



**Figura 19.** Estrutura do cap na extremidade 5' de uma molécula de RNAm de eucariote.

A extremidade 3' do transcrito também é modificada. O RNA transcrito é clivado logo após ser transcrita uma sequência específica da molécula (**senal de poliadenilação**). Em seguida, uma cadeia de nucleotídeos com base adenina (**cauda de poli-A**) é adicionada, na extremidade 3' clivada, por ação de uma polimerase específica.

O sítio específico de reconhecimento da região de clivagem e poliadenilação corresponde a uma sequência de bases AAUAAA (ou AUUAAA) localizada entre 10 e 30 nucleotídeos a 5' do sítio efetivo de corte. Imediatamente após o corte do transcrito, a enzima **polimerase poli-A** adiciona de 100 a 200 resíduos de nucleotídeos contendo adenina, na extremidade 3' da cadeia de RNAm, completando o transcrito primário (Fig. 20). Enquanto isto ocorre, a polimerase do RNA continua a transcrever por centenas ou milhares de nucleotídeos até que a enzima encontre a sinalização de parada da transcrição, até hoje ainda não completamente caracterizada. Este pedaço extra de RNA transcrito, sem função, é degradado, presumivelmente porque não possui a proteção do “encapamento”.



**Figura 20.** Mecanismo de formação da cauda de poli-A pela clivagem do transcrito nascente seguida de polidenilação por polimerase distinta (polimerase poli-A).

À cauda de poli-A atribuem-se várias funções: (1) ela é importante na exportação do RNAm maduro para o citoplasma; (2) acredita-se que ela afete a estabilidade dos RNAm no citoplasma; (3) faz parte de um sistema de sinalização para o reconhecimento pelo ribossomo durante a tradução. Esta última característica, associada ao encapamento, permitiria ao ribossomo reconhecer se o RNAm está intacto e apto à tradução, antes que energia e precursores sejam gastos na tradução.

A maioria dos RNAm é instável, ou seja, tem vida muito curta. Consequentemente, o RNAm no núcleo e o RNAm dele derivado, no citoplasma, constituem uma fração minoritária de moléculas de RNA na célula. Mesmo estando pouco representado na população de moléculas de RNA de uma célula, o RNAm pode ser facilmente isolado dos demais tipos por possuir a cauda de poli-A.

Quando uma amostra de RNA eucariótico total é passada através de uma coluna contendo curtas sequências de DNA de fita simples sintético contendo apenas a base timina (oligo-dT), ligadas a um suporte sólido, há emparelhamento entre as bases timina com a cadeia poli A dos RNAm. Com isso, os mensageiros ficam presos ao suporte sólido, enquanto todos os outros tipos de RNA passam livremente pela coluna. Em seguida, passa-se através da coluna uma solução com uma concentração iônica tal que as pontes de hidrogênio entre as caudas poliA e os oligos-dT são rompidas. Com isso as moléculas de RNAm se soltam e podem ser coletadas com alto grau de pureza.

## O RNA Transportador

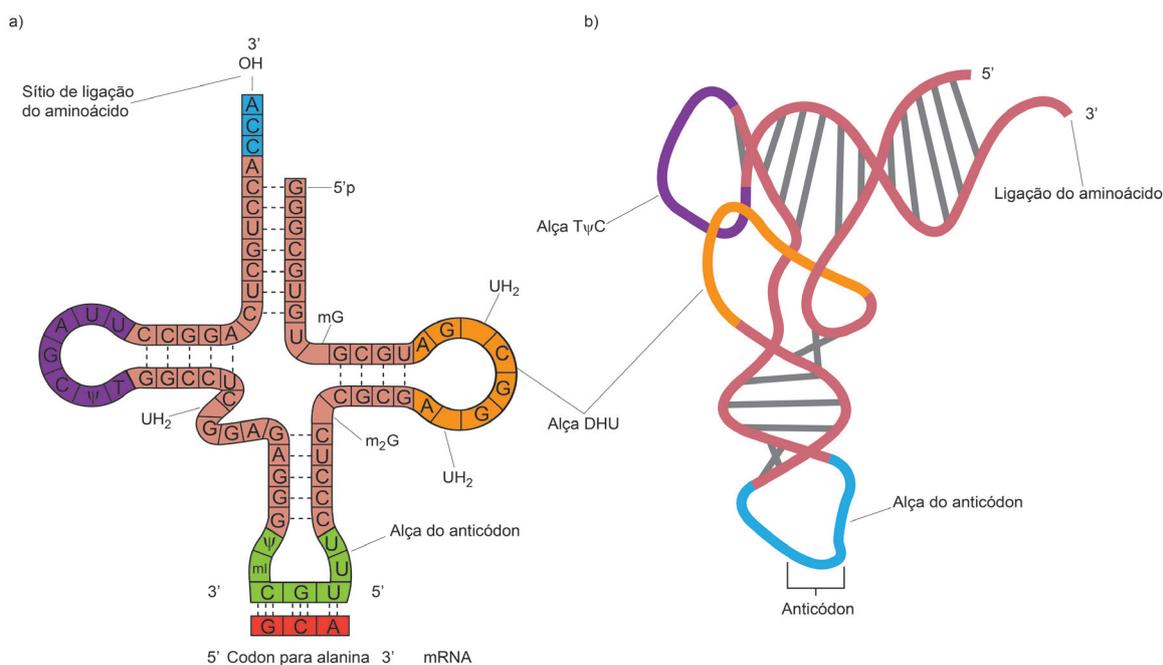
A primeira hipótese para explicar a atuação do RNA como intermediador na síntese de proteínas partia do princípio que as moléculas de RNAm se dobrariam de modo a formar concavidades nas quais os diferentes aminoácidos pudessem se encaixar. Esse dobramento do RNA seria específico para cada um dos 20 aminoácidos existentes na célula e, desse modo, o RNA forneceria a instrução para a ordenação dos aminoácidos durante a síntese de proteína.

Essa hipótese foi descartada por Francis Crick em 1956; seu argumento foi que não havia uma fundamentação química para a interação entre o RNA dobrado e um aminoácido. Mesmo que o dobramento proposto ocorresse, não haveria condições para a discriminação precisa entre os aminoácidos, sabidamente muito semelhantes entre si, como por exemplo, glicina e alanina ou valina e isoleucina, que diferem apenas pela presença de um grupo metil (CH<sub>3</sub>). Crick propôs, então, que antes de serem incorporados em proteínas, os aminoácidos deveriam ligar-se a moléculas adaptadoras específicas, e essas poderiam ligar-se especificamente às bases de um RNA que teria sido sintetizado tendo o DNA como molde. Hoje se sabe que o papel de molécula adaptadora é desempenhado pelos **RNAs transportadores (RNAt)**.

Os RNAt foram intensamente estudados e hoje se sabe que eles estão envolvidos em uma multiplicidade de reações. Para desempenhar certas reações é necessário que todos os tipos de RNAt possuam algumas características em comum mas, por outro lado, eles possuem pequenas diferenças que permitem a distinção entre os diferentes tipos de moléculas.

Todos os RNAt devem ser capazes de interagir com sítios específicos dos ribossomos durante o processo de síntese de proteína. Além disso, em uma de suas extremidades, eles devem se associar ao RNAm, através da interação entre bases complementares; na outra extremidade eles devem transportar um aminoácido. Para cumprir essas funções, todos os tipos de RNAt devem obedecer a uma regra geral e uniforme quanto à sua forma e tamanho.

Os RNAt consistem de um conjunto de pequenas moléculas de RNA, cada uma delas com cerca de 80 nucleotídeos. Diferentemente dos demais ácidos nucléicos, eles apresentam diversas **bases não usuais**, ou seja, diferentes de A, G, C ou U. Essas bases originam-se por modificações de bases usuais após sua incorporação na molécula de RNAt (Fig. 21).



**Figura 21.** Estrutura em forma de folha de trevo da molécula de RNAt no plano, à esquerda. À direita, uma visão da conformação tridimensional do RNAt determinada por difração de raios X. UH<sub>2</sub>, mG e m<sub>2</sub>G são bases nitrogenadas modificadas.

As bases das moléculas precursoras dos RNAt sofrem uma vasta gama de modificações, desde simples metilações até reestruturação do anel purínico. Tais modificações aumentam a versatilidade dos RNAt, o que é muito importante para as várias funções que desempenham. O efeito mais evidente da modificação de bases dos RNAt é a alteração em sua capacidade de emparelhamento com o RNAm. Nesse caso a modificação ocorre em bases da trinca (**anticódon**), por meio das quais o RNAt se emparelha com o RNA mensageiro durante a síntese de proteína, ou nas vizinhanças delas.

O sequenciamento de bases de centenas de RNAt de bactérias e eucariontes mostrou que todos têm a mesma estrutura secundária, uma forma de folha de trevo mantida por emparelhamento de bases entre pequenas regiões de complementaridade (Fig. 21).

Os quatro braços principais presentes na estrutura em forma de trevo receberam denominações relacionadas à sua estrutura ou à função que desempenham. O **braço aceitador** termina com uma sequência de bases não emparelhadas, cujos grupos hidroxila livres 2' ou 3' têm a capacidade de se ligar a aminoácidos; o **braço ou alça T**

é assim denominado porque contém timina; o **braço ou alça anticódon** contém a trinca do anticódon; o **braço ou alça D** contém a base dihidrouridina. Como mostrado na Figura 21, a estrutura secundária do RNAt se dobra, originando uma estrutura terciária mais compacta, em forma de L, com o anticódon em uma das extremidades e o braço acceptor na outra. A estrutura terciária é mantida por meio de pontes de hidrogênio entre bases que se encontram desemparelhadas na estrutura secundária.

Há uma tendência nas diferentes espécies de para que os genes dos RNAt fiquem agrupados em determinadas regiões do cromossomo. Em todos os casos conhecidos, uma molécula de RNAt é formada a partir de um RNA precursor, mais longo, do qual ele é liberado por clivagens e outras reações de processamento ou maturação. Em bactérias, os RNAt são formados a partir de grandes moléculas de RNA precursoras, contendo um ou mais RNAt, com sequências adicionais em ambas as extremidades, 5' e 3'. Uma única endorribonuclease, a enzima **RNase P**, é responsável pela formação da extremidade 5' de todas as moléculas de RNAt. A ribonuclease P é uma ribonucleoproteína, formada por um RNA com 375 nucleotídeos de comprimento e por um componente protéico. Ambos os componentes são necessários para a atividade catalítica dessa enzima, embora a responsabilidade básica de clivagem seja do RNA (**Ribozima** é o termo geral utilizado para descrever um RNA com atividade catalítica). Uma outra enzima, a **RNase D**, remove, um a um, os nucleotídeos da extremidade 3' da sequência, até chegar à sequência -CCA, comum a todos RNAt maduros.

Os RNAt recém-sintetizados presentes no citoplasma das células eucarióticas apresentam uma constante de sedimentação de aproximadamente 4,5S, correspondendo portanto a moléculas de 100 nucleotídeos. As formas maduras de 4S apresentam entre 70 e 80 nucleotídeos.

## O RNA Ribossômico

Os ribossomos são estruturas ribonucleoprotéicas, presentes no citoplasma de células pro e eucarióticas, que atuam na síntese de proteínas. Os ribossomos são compostos por duas subunidades de tamanhos diferentes, que se encaixam uma na outra formando um complexo com massa de vários milhões de Daltons. Mais da metade dessa massa corresponde a uma classe especial de RNA, o chamado **RNA ribossômico (RNAr)**. Existem evidências de que esse RNA tem atividade catalítica na síntese de proteínas (Fig. 22 e Fig. 23).

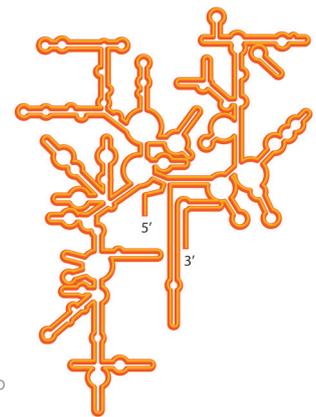


Figura 22. Estrutura do RNAr 16S presente na subunidade menor do ribossomo de *E. coli*.

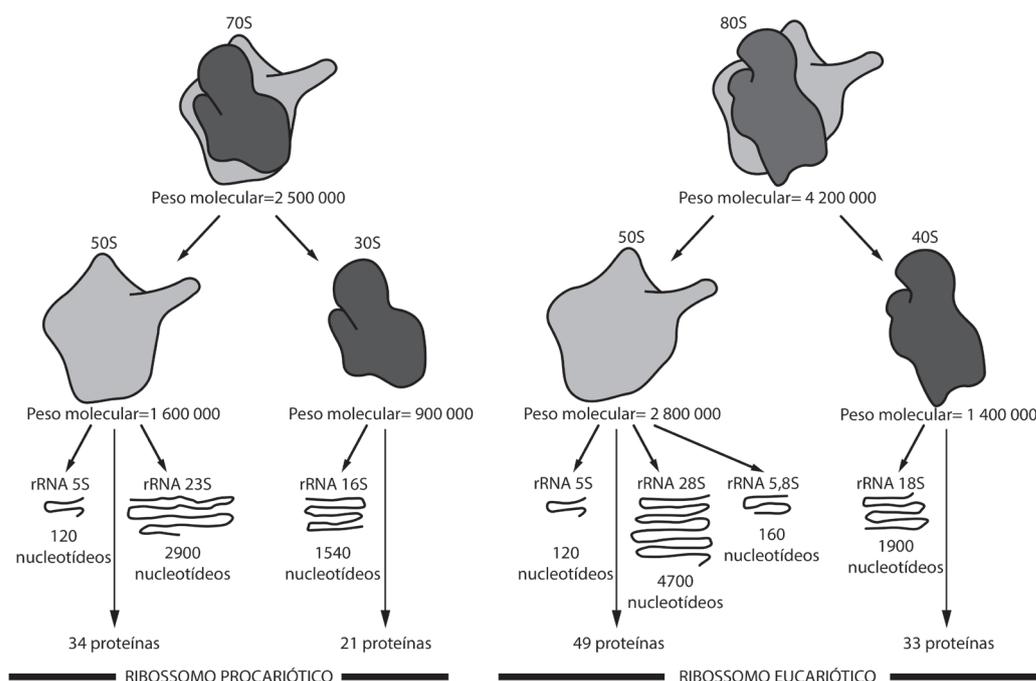


Figura 23. Tipos de RNA ribossômicos presentes nas diferentes subunidades de ribossomos bacterianos (à esquerda) e nas células eucarióticas (à direita).

As subunidades dos ribossomos, assim como as moléculas de RNAr, são designadas pelo seu coeficiente de sedimentação, expresso em unidades Svedberg (**S**). Svedberg é a medida da razão de sedimentação de partículas em suspensão, centrifugadas em gradiente de sacarose sob condições padronizadas.

A sequência de bases dos RNAr de centenas de organismos já foi determinada. Embora a sequência primária de bases varie consideravelmente, a estrutura geral do RNAr é semelhante na maioria das espécies, ou seja, há um emparelhamento de bases em várias regiões da molécula originando uma estrutura com muitas alças. (Figura 22)

Atualmente, há evidências de que o RNAr interage com o RNAm ou RNAt em cada um dos estágios da tradução, e que, de fato, as proteínas são necessárias para manter o RNAr na estrutura na qual ele seja capaz de realizar a atividade catalítica.

### A síntese de RNAr

Muitas das proteínas que são abundantes em células diferenciadas, tais como a actina e miosina de células musculares, são sintetizadas a partir de genes que estão presentes em cópia única no genoma (haploide). Essas proteínas são abundantes porque cada uma das muitas moléculas de RNAm transcritas serão traduzidas numa taxa de 10 moléculas de proteínas por minuto, o que normalmente produz, em uma geração celular, cerca de 10.000 moléculas de proteínas por molécula de RNAm. Entretanto, esse tipo de amplificação não é possível no caso dos genes que não codificam proteínas. Este é o caso, por exemplo, dos genes ribossomais, uma vez que o RNAr que compõe os ribossomos é o produto final desses genes. Uma célula de mamífero em crescimento deve sintetizar 10 milhões de cópias de cada tipo de molécula de RNAr em cada geração celular, para poder construir 10 milhões de ribossomos. Para que quantidades adequadas sejam produzidas, as células possuem múltiplas cópias de genes ribossomais.

Em *E. coli*, existem sete cópias dos genes ribossomais para suprir as necessidades da célula, enquanto em células humanas há cerca de 200 genes por genoma haplóide, espalhados em pequenos grupos em cinco diferentes cromossomos. Nos eucariontes, as cópias múltiplas dos **genes ribossomais** em um determinado cromossomo, ficam arranjadas em série, sendo que um gene é separado do outro por uma **região espaçadora**, que não é transcrita, e que varia em tamanho.

Por causa do grande número de genes ribossomais, eles são facilmente visualizados em preparações de microscopia eletrônica utilizando a técnica do espalhamento da cromatina.

Os genes para RNAr de uma célula eucariótica são transcritos pela polimerase I que produz uma molécula de **RNAr 45S**, com cerca de 13 mil nucleotídeos de comprimento. Antes de deixar o núcleo como parte do ribossomo, o RNAr de 45S é processado, dando origem a três moléculas distintas: um **RNAr 28S**, com cerca de 5 mil nucleotídeos; um **RNAr 18S**, com cerca de 2 mil nucleotídeos; e um **RNAr 5,8S**, com cerca de 160 nucleotídeos. A parte restante do transcrito de 45S, cerca de 6 mil nucleotídeos, é degradada no núcleo. Parte dessas sequências extras de RNA parece ter um papel transitório na montagem das subunidades do ribossomo. (Fig. 24)

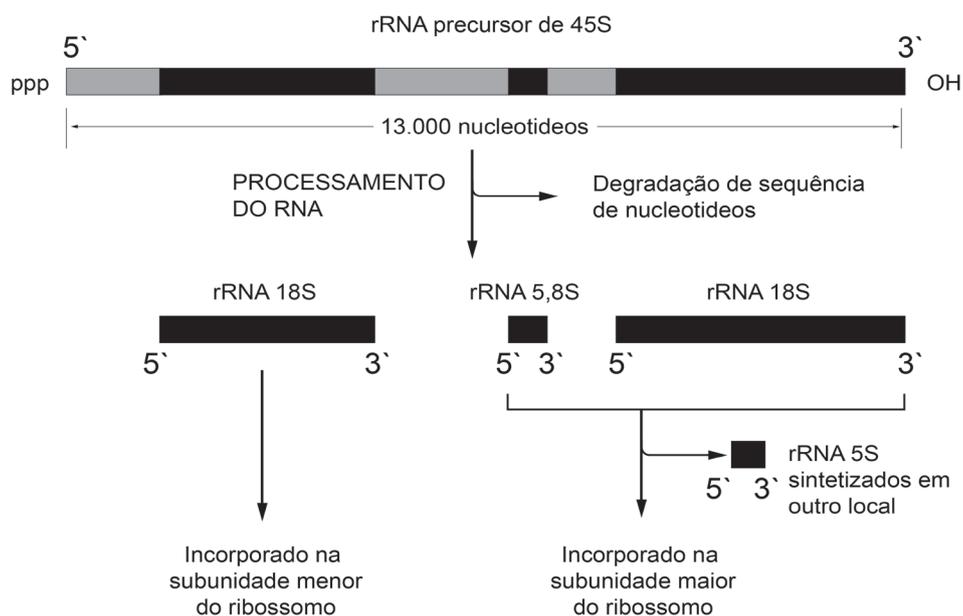


Figura 24. Esquema do processamento do RNA precursor do RNAr de eucariontes.

Outro conjunto de genes arranjados em sequência codifica para o **RNAr 5S** que faz parte da subunidade maior do ribossomo. Esse RNAr é o único que é transcrito separadamente dos demais, fora da **região organizadora do nucléolo**, como é chamada a região do cromossomo onde ficam localizados os demais genes para RNAr.

Os genes para o RNAr 5S têm cerca de 120 pares de nucleotídeos de comprimento e são transcritos pela polimerase III do RNA. Uma célula humana possui cerca de 2000 genes para RNAr 5S arranjados em série, longe dos demais genes RNAr.

## EXERCÍCIOS

### Parte A: Revendo Conceitos Básicos

Preencha os espaços em branco nas frases de 1 a 8 usando o termo abaixo mais apropriado:

- (a) fator sigma
- (b) cadeia de código
- (c) cadeia molde
- (d) promotor
- (e) bolha de transcrição
- (f) ligações fosfodiéster 3' 5'
- (g) polimerase do RNA
- (h) polimerase II do RNA

1. A ( ) copia uma fita de DNA em RNA num processo conhecido como transcrição.
2. ( ) é a região do DNA onde é iniciada a síntese de RNA.
3. O RNA é uma molécula formada por ribonucleotídeos que estão ligados entre si por ( ).
4. O/A ( ) é responsável pelo reconhecimento do local correto que sinaliza o início do processo de transcrição.
5. A ( ) é responsável pela transcrição de genes cujos RNAs serão traduzidos em proteínas.
6. Na(o) ( ) cerca de 18 pares de bases do DNA estão temporariamente desemparelhados e uma das fitas serve de molde para a síntese de RNA.
7. Na síntese de RNA, a fita de DNA que lhe serve de molde é denominada ( ).
8. A fita de DNA que possui a mesma sequência do RNA (exceto por possuir T no lugar de U) é a ( ).

Preencha os espaços em branco nas frases de 9 a 14 usando o termo abaixo mais apropriado:

- (a) fator rho
  - (b) nuclease
  - (c) 5' - 3'
  - (d) sequências consenso
  - (e) 3' - 5'
  - (f) 28S, 18S, 5,8S
  - (g) 5,8S e 5S
9. Nas posições -10 e -35 de um promotor existem ( ).

10. A síntese do RNA ocorre na direção ( ).
11. Além da sequência de bases que sinaliza o término da transcrição em bactérias, há genes nos quais é necessário o auxílio de uma proteína chamada ( ).
12. O produto de transcrição de um gene ribossomal corresponde a uma grande molécula precursora que é processada originando, nos eucariontes, os RNAr ( ).
13. Os RNAs ( ) fazem parte dos ribossomos, porém são sintetizados em diferentes locais do genoma, o primeiro na região organizadora do nucléolo e o outro fora dela.

## Parte B: Ligando Conceitos e Fatos

Para cada uma das frases listadas nas questões de 15 a 25 escreva no parêntesis: a letra V, caso a afirmação seja verdadeira, ou a letra F, no caso dela ser falsa.

14. Em uma determinada região da dupla hélice de DNA, em geral, apenas uma das fitas é usada como molde na transcrição. ( )
15. O término da transcrição nos eucariontes ocorre por ação do fator rho. ( )
16. Todas as células possuem três diferentes tipos de polimerases do RNA. ( )
17. Alguns genes codificam para proteínas, enquanto outros têm RNAs como produto final. ( )
18. De acordo com o dogma central da Biologia, a expressão da informação genética na célula é, em geral, bidirecional. ( )
19. Os transcritos primários correspondem em geral a produtos finais e não são modificados após a transcrição. ( )
20. Uma unidade de transcrição é composta por apenas um gene. ( )
21. A estrutura geral dos RNAr varia consideravelmente nas diferentes espécies. ( )
22. Todos os genes que codificam para proteínas possuem em seu final uma sequência de Ts que é transcrita e dá origem a cauda de poliA dos RNAm. ( )
23. A fração de RNA que contém os RNAm de uma célula é heterogênea quanto ao tamanho. ( )
24. Dentre as moléculas orgânicas, apenas as proteínas possuem atividade catalítica. ( )

Complete as frases de 26 a 30 com o termo mais apropriado dentre os listados abaixo.

- (a) cauda de poliA
  - (b) 5' cap
  - (c) unidade de transcrição
  - (d) RNAm
  - (e) RNAn
25. A molécula de polimerase do RNA dos eucariontes inicia e termina a transcrição em locais específicos do DNA; a região entre esses sítios é denominada ( ).
  26. No núcleo, os transcritos da polimerase II do RNA são conhecidos como moléculas de ( ) porque uma das primeiras características usadas para distingui-los dos outros RNAs foi sua heterogeneidade de tamanho.

27. Os transcritos da polimerase II do RNA deixam o núcleo na forma de ( ).
28. A adição de um nucleotídeo G metilado na extremidade 5' do transcrito forma o ( ), cuja função parece ser a de proteger o RNA que está sendo sintetizado da degradação, além disso ele parece ter papel importante na síntese das proteínas.
29. A extremidade 3' da maioria dos transcritos da polimerase II do RNA é definida por modificação, na qual o transcrito em crescimento é clivado em um sítio específico e uma ( ) é adicionada por uma outra polimerase.

Indique a alternativa mais apropriada para completar as frases de 30 a 31.

30. Uma sequência consenso
- é exclusiva de um determinado gene.
  - representa a sequência de bases do DNA de uma espécie.
  - corresponde a uma sequência conservada na evolução.
  - é particular de genes que codificam para os RNAr.
31. As moléculas precursoras de RNAr
- têm um tempo de vida curto.
  - são sintetizadas no citoplasma da célula.
  - tem cerca de 80S nos eucariontes.
  - são aquelas que dão origem ao RNAr 5S.

## Parte C: Aplicando Conceitos

32. Esquematize um segmento de DNA indicando a polaridade das fitas e a transcrição a partir de uma delas.
33. Construa um mapa de conceitos relacionando os seguintes termos: DNA / polimerase do RNA / holoenzima / fator sigma / transcrição do DNA / promotor/ sequência terminalizadora / transcrito.

## Parte D: Resolvendo Problemas

34. Um segmento de DNA tem a sequência abaixo esquematizada:



O seguinte RNA foi sintetizado a partir desse segmento:



Com base nos esquemas acima responda as questões abaixo:

- A molécula de RNA foi sintetizada tendo como molde qual das fitas do DNA?
  - Determine na cadeia molde o ponto de início e ponto de término da cadeia de RNA.
  - Na síntese do RNA houve necessidade de um oligonucleotídeo iniciador ou primer?
35. Em um sistema in vitro capaz de catalisar síntese de RNA, foi introduzida a cadeia molde de DNA representada abaixo. Sabendo-se que a primeira base a ser incorporada no RNA é complementar à base indicada em negrito, qual será a composição de bases do RNA transcrito?



36. Uridina tritiada foi adicionada ao meio de uma cultura de células. Cinco minutos após a adição, o RNA total da célula foi extraído e submetido a centrifugação em gradiente de sacarose para separar os diferentes tipos de moléculas de RNA produzidas. Após a centrifugação, o tubo de centrífuga foi perfurado em sua base e foram coletadas 100 amostras de igual volume. A concentração de cada amostra foi medida em um espectrofotômetro a 260nm e, em seguida, foi realizada uma medida da radioatividade incorporada. O seguinte gráfico foi obtido:

- Por que não há coincidência entre os picos de concentração e de radioatividade incorporada?
- A que tipo de RNA deve corresponder o pico de radioatividade polidisperso?
- Por que não existe uma concentração mensurável de RNA por densidade óptica correspondente ao pico de radioatividade polidispersa?

