

## CAPÍTULO 7

### Do DNA à Proteína: Como as Células Leem o Genoma

Desde que a estrutura do DNA (ácido desoxirribonucleico) foi determinada no início da década de 1950, tornou-se claro que a informação genética nas células estava codificada na sequência de nucleotídeos do DNA. Vimos, no Capítulo 6, como essa informação pode ser transmitida, de modo conservado, de uma célula aos seus descendentes pelo processo de replicação do DNA. Como uma célula decodifica e usa essa informação? Como as instruções genéticas escritas sob a forma de um alfabeto de apenas quatro “letras” – os quatro diferentes nucleotídeos do DNA – podem levar à formação de uma bactéria, uma mosca-das-frutas ou um ser humano? Se ainda temos muito a aprender a respeito de como a informação estocada nos genes de um organismo leva à produção, mesmo da mais simples das bactérias, o que não dizer de como ela pode direcionar o desenvolvimento de organismos multicelulares complexos, como nós mesmos? No entanto, o código de DNA *per se* já foi decifrado, e a linguagem dos genes pode ser lida.

Mesmo antes da decifração do código de DNA, sabíamos que a informação contida nos genes, de alguma forma, era responsável pelo direcionamento da síntese de proteínas. As proteínas são os principais constituintes das células e determinam não apenas a sua estrutura, mas também o seu funcionamento. Nos capítulos anteriores, deparamo-nos com alguns dos milhares de tipos diferentes de proteínas que podem ser produzidas pelas células. Vimos no Capítulo 4 que as propriedades e funções de uma molécula proteica são determinadas pelo ordenamento linear – *sequência* – das diferentes subunidades de aminoácidos que compõem sua cadeia polipeptídica: cada tipo de proteína possui a sua sequência específica de aminoácidos, e essa sequência dita como a cadeia se dobrará para dar à molécula a sua forma e características químicas específicas. As instruções genéticas transportadas pelo DNA devem, portanto, especificar a sequência de aminoácidos das proteínas. Veremos, no presente capítulo, como isso acontece realmente.

O DNA não dirige a síntese proteica diretamente, mas age como um coordenador, delegando as diferentes tarefas necessárias para a realização do

DO DNA AO RNA

DO RNA À PROTEÍNA

RNA E A ORIGEM DA VIDA

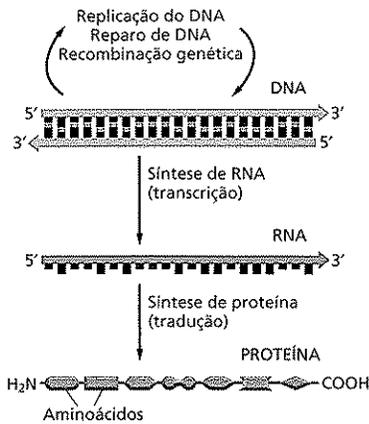


Figura 7-1 A informação genética direciona a síntese de proteínas. O fluxo de informação genética do DNA ao RNA (transcrição) e do RNA à proteína (tradução) ocorre em todas as células vivas.

trabalho a um grupo de colaboradores. Quando uma determinada proteína é necessária na célula, a sequência nucleotídica da porção apropriada da longa molécula de DNA de um cromossomo será inicialmente copiada em um outro tipo de ácido nucleico – RNA (*ácido ribonucleico*). Essas cópias de RNA, formadas a partir de curtos segmentos do DNA, serão usadas como molde para conduzir a síntese da proteína. Milhares dessas conversões de DNA para proteína ocorrem a cada segundo em cada uma das células de nosso organismo. O fluxo de informação genética nas células segue, portanto, uma rota do DNA para o RNA para a proteína (Figura 7-1). Todas as células, de bactérias a seres humanos, expressam suas informações genéticas dessa forma – um princípio tão fundamental que foi denominado *dogma central* da biologia molecular.

Neste capítulo, abordaremos os mecanismos pelos quais as células copiam o DNA em RNA (um processo denominado *transcrição*) e, a seguir, utilizam a informação presente no RNA para a produção de proteína (um processo denominado *tradução*). Introduziremos também algumas das principais variações que ocorrem nesse esquema básico. A principal dessas variações é o *splicing de RNA*, um processo ao longo do qual os transcritos de RNA são clivados e reunidos antes que as células eucarióticas os traduzam em proteínas. Essas alterações podem alterar a mensagem transportada pelas moléculas de RNA e, conseqüentemente, são essenciais para a compreensão dos mecanismos de decodificação do genoma pelas células. Na seção final deste capítulo, consideraremos como o esquema de estoque de informação, de transcrição e de tradução atual deve ter-se originado a partir de sistemas mais simples, nos estágios iniciais da evolução celular.

## DO DNA AO RNA

A transcrição e a tradução são as maneiras pelas quais as células leem, ou *expressam*, as suas instruções genéticas – os seus *genes*. Várias cópias idênticas de RNA podem ser feitas a partir de um mesmo gene, e cada molécula de RNA pode direcionar a síntese de várias cópias idênticas de uma molécula proteica. Visto que cada célula contém apenas uma ou duas cópias de um gene determinado, essa amplificação sucessiva permite que as células sejam capazes de sintetizar rapidamente grandes quantidades de uma proteína quando necessário. Ao mesmo tempo, cada gene pode ser transcrito e traduzido em diferentes taxas, possibilitando que a célula produza enormes quantidades de algumas proteínas e, simultaneamente, quantidades minúsculas de outras (Figura 7-2). Além disso, como veremos no Capítulo 8, uma célula pode alterar (ou regular) a expressão de cada um dos seus genes de acordo com as necessidades do momento. Nesta seção, discutiremos a produção do RNA – a primeira etapa da *expressão gênica*.

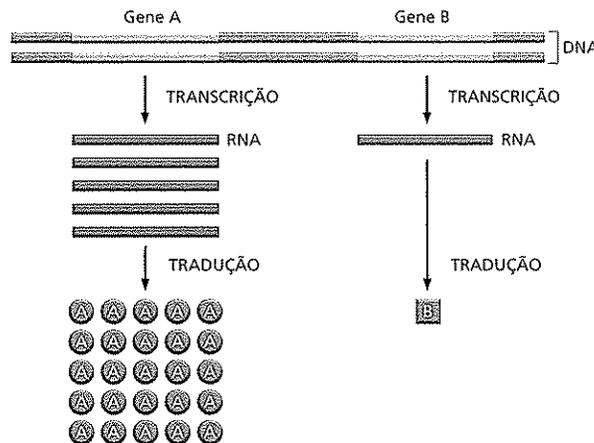


Figura 7-2 Os genes podem ser expressos sob diferentes taxas. O gene A é transcrito e traduzido muito mais eficientemente do que o gene B. Isso permite que a quantidade de proteína A na célula seja muito maior do que a quantidade de proteína B. Nesta figura, e em figuras posteriores, as regiões não transcritas do DNA são representadas em cinza.

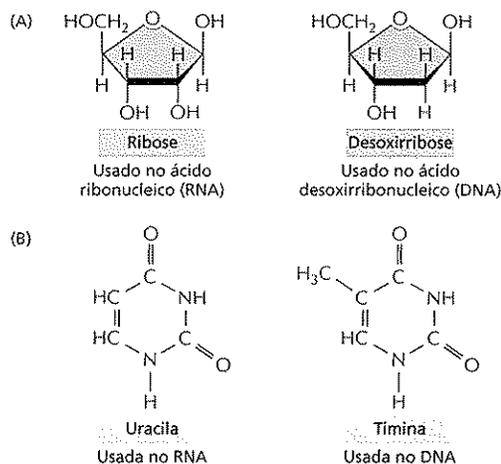
Os s  
O pri  
gene  
proc  
uma  
guag  
merc  
unid  
em t  
cleo  
boni  
as b  
da ti  
de t  
prop  
tam  
  
trut  
oco  
con  
Vist  
adq  
um.  
DN.  
nes  
xas  
de:  
atu  
for  
fun  
  
Fig  
O I  
pré  
a b  
RN

## Os segmentos da sequência de DNA são transcritos em RNA

O primeiro passo que uma célula dá para a leitura de um dos seus milhares de genes é a cópia da sequência nucleotídica desse gene sob a forma de RNA. Esse processo é denominado **transcrição**, pois a informação, apesar de copiada sob uma nova forma química, permanece escrita essencialmente sob a mesma linguagem – a linguagem dos nucleotídeos. Assim como o DNA, o **RNA** é um polímero linear composto por quatro diferentes tipos de subunidades nucleotídicas unidas entre si por ligações fosfodiéster (Figura 7-3). Ele se diferencia do DNA, em termos químicos, sob dois aspectos: (1) os nucleotídeos no RNA são *ribonucleotídeos* – ou seja, eles contêm o açúcar ribose (origem do nome ácido ribonucleico), em vez de desoxirribose; (2) embora, como o DNA, o RNA contenha as bases adenina (A), guanina (G) e citosina (C), ele contém uracila (U) em vez da timina (T) encontrada no DNA. Visto que U, assim como T, pode formar pares de bases pelo estabelecimento de pontes de hidrogênio com A (Figura 7-4), as propriedades de complementaridade de base descritas para o DNA no Capítulo 5 também se aplicam ao RNA.

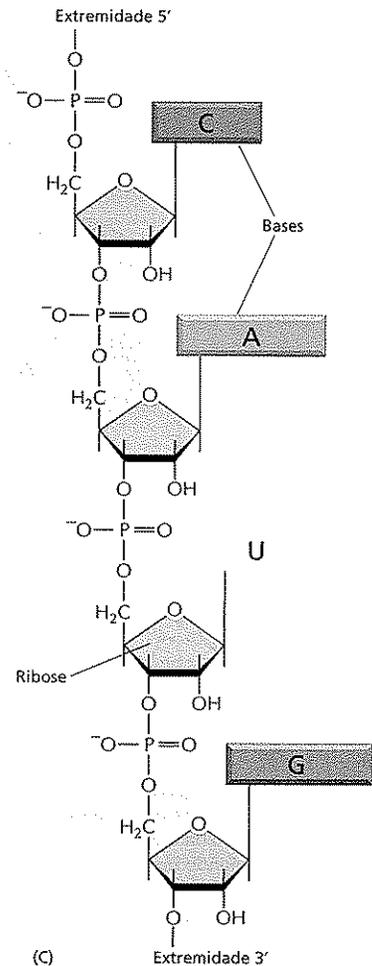
Apesar de apresentarem composição química bastante semelhante, a estrutura geral do DNA e do RNA difere drasticamente. Enquanto o DNA sempre ocorre nas células sob a forma de uma hélice de fita-dupla, o RNA se apresenta como fita simples. Essa diferença tem importantes consequências funcionais. Visto que a cadeia de RNA é de fita simples, ela pode dobrar-se sobre ela própria adquirindo diferentes formas, exatamente como ocorre com o dobramento de uma cadeia polipeptídica na estruturação final de uma proteína (Figura 7-5); o DNA de fita dupla não pode dobrar-se desse modo. Como veremos, mais adiante neste capítulo, a capacidade de dobrar-se em estruturas tridimensionais complexas permite que o RNA desempenhe funções na célula que vão muito além das de simples intermediário de informações entre DNA e proteína. Enquanto o DNA atua unicamente no estoque de informações, o RNA se apresenta sob diversas formas, algumas com funções estruturais ou mesmo capazes de desempenhar funções catalíticas.

Figura 7-3 A estrutura química do RNA se diferencia ligeiramente da estrutura do DNA. (A) O RNA contém o açúcar ribose, o qual difere da desoxirribose, o açúcar utilizado no DNA pela presença de um grupo  $-OH$  adicional. (B) O RNA contém a base uracila, a qual difere da timina, a base equivalente no DNA, pela ausência de um grupo  $-CH_3$ . (C) Um pequeno fragmento de RNA. A ligação química entre nucleotídeos no RNA é a mesma que ocorre no DNA.



### QUESTÃO 7-1

Considere a expressão "dogma central", referente ao fluxo da informação genética do DNA para o RNA e, a seguir, para proteína. A palavra "dogma" é apropriada nesse contexto?



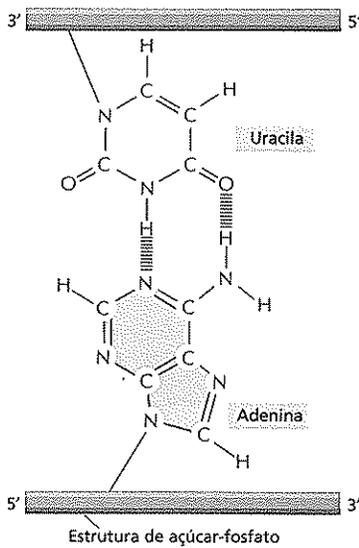


Figura 7-4 A uracila forma pares de bases com a adenina. Apesar da ausência de um grupo metila, a uracila apresenta as mesmas propriedades de pareamento da timina. Assim, pares de bases U-A se assemelham bastante a pares de bases T-A (ver Figura 5-6A).

### A transcrição produz um RNA complementar a uma das fitas do DNA

Todo o RNA de uma célula é produzido a partir da transcrição, um processo que apresenta certas similaridades com a replicação do DNA (discutida no Capítulo 6). A transcrição tem início com a abertura e o desespiralamento de uma pequena porção da dupla-hélice de DNA para que sejam expostas as bases de ambas as fitas do DNA. A seguir, uma das duas fitas do DNA de dupla-hélice atuará como molde para a síntese do RNA. Os ribonucleotídeos são adicionados, um a um, à cadeia de RNA em crescimento, e, da mesma forma que ocorre na replicação do DNA, a sequência nucleotídica da cadeia de RNA é determinada pelo pareamento por complementaridade de bases com o DNA-molde. Quando um pareamento correto é feito, o ribonucleotídeo recém-chegado é ligado covalentemente à fita de RNA em crescimento por meio de uma reação catalisada enzimaticamente. A cadeia de RNA produzida pela transcrição – o *transcrito* – é, desse modo, estendida nucleotídeo a nucleotídeo e apresenta sequência nucleotídica exatamente complementar à fita de DNA usada como molde (Figura 7-6).

A transcrição, no entanto, difere da replicação do DNA em vários pontos essenciais. Diferentemente de uma fita de DNA recém-formada, a fita de RNA não permanece ligada à fita de DNA-molde por meio de pontes de hidrogênio. Em vez disso, em uma região imediatamente além da região onde os ribonucleotídeos estão sendo inseridos, a cadeia de RNA é deslocada, e a hélice de DNA é reestruturada. Por essa razão – e considerando que apenas uma fita da molécula de DNA é transcrita –, as moléculas de RNA são de fita simples. Além disso, como os RNAs são copiados a partir de uma região delimitada do DNA, essas moléculas são muito mais curtas do que as moléculas de DNA; as moléculas de DNA em um cromossomo humano podem alcançar um comprimento de até 250 milhões de pares de nucleotídeos, ao passo que a maioria dos RNAs não possui um comprimento maior do que uns poucos milhares de nucleotídeos, sendo muitos deles ainda bem menores do que isso.

As enzimas que realizam a transcrição são chamadas de **RNA-polimerases**. Assim como a DNA-polimerase que catalisa a replicação do DNA (discutida

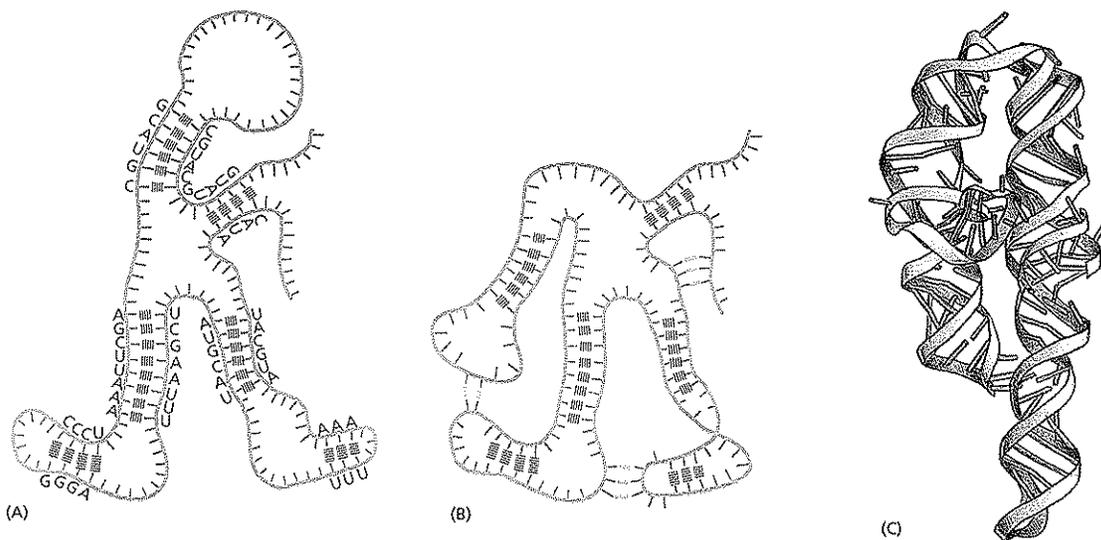


Figura 7-5 O RNA pode formar pares de bases intramolecularmente e se dobrar em estruturas específicas. O RNA é uma fita simples, mas frequentemente contém pequenos segmentos de nucleotídeos que podem sofrer pareamento com sequências complementares encontradas em outras regiões da mesma molécula. Tais interações, em adição a pareamentos "não convencionais", permitem que uma molécula de RNA se dobre, adquirindo uma estrutura tridimensional, a qual é determinada por sua sequência de nucleotídeos. (A) Diagrama de uma estrutura hipotética de RNA dobrada, ilustrando somente interações convencionais (ou seja, Watson-Crick) entre pares de bases; (B) a mesma estrutura hipotética com ambos os tipos de pareamento, convencionais (vermelho) e não convencionais (p. ex., A-G) (verde); (C) estrutura de um RNA real, uma molécula envolvida no *splicing* de RNA. Cada pareamento convencional está indicado por um "traço" na dupla-hélice. Bases em outras configurações estão indicadas por traços quebrados. Para visualização adicional da estrutura do RNA, ver Animação 7.1.

no Capítulo 6), as RNA-polimerases catalisam a formação de ligações fosfodiéster que unem os nucleotídeos e formam o esqueleto açúcar-fosfato de uma cadeia de RNA. A RNA-polimerase se move paulatinamente sobre o DNA, desenrolando a hélice de DNA à sua frente e expondo a nova região de fita-molde para que ocorra o pareamento por complementaridade de bases. Desse modo, a cadeia de RNA em crescimento é estendida nucleotídeo a nucleotídeo na direção 5'– 3' (Figura 7-7). A enzima utiliza ribonucleosídeos trifosfato (ATP, CTP, UTP e GTP), cujas ligações de alta energia fornecem a energia que direciona a continuidade da reação (ver Figura 6-10).

A liberação quase imediata da fita de RNA recém-sintetizada da fita de DNA-molde faz com que muitas cópias de RNA possam ser feitas a partir de um único gene, em um intervalo de tempo relativamente curto; a síntese do próximo RNA é geralmente iniciada antes que a primeira cópia de RNA esteja completa (Figura 7-8). Um gene de tamanho mediano (digamos, 1.500 pares de nucleotídeos) leva aproximadamente 50 segundos para ser transcrito por uma molécula de RNA-polimerase (Animação 7.2). Em um momento específico qualquer, podem existir até 15 polimerases percorrendo esse pequeno segmento de DNA, umas nos calcanhares das outras, o que permite a síntese de mais de 1.000 transcritos em um período de uma hora. Na maioria dos genes, entretanto, a taxa de transcrição é bem mais baixa do que essa.

Apesar de a RNA-polimerase catalisar essencialmente a mesma reação química que a DNA-polimerase, existem algumas diferenças importantes entre essas duas enzimas. A primeira, e mais óbvia, é que a RNA-polimerase catalisa a ligação de ribonucleotídeos, em vez de desoxirribonucleotídeos. A segunda é que, contrariamente à DNA-polimerase, envolvida na replicação de DNA, as RNA-polimerases podem dar início à síntese de uma cadeia de RNA na ausência de um iniciador. Essa diferença ocorre porque a transcrição não precisa acontecer de forma tão exata quanto a replicação do DNA; diferentemente do DNA, o RNA não é usado como forma de estoque permanente da informação genética nas células. Assim, pequenos erros nos transcritos de RNA levam a consequências relativamente desprezíveis. As RNA-polimerases possuem uma taxa de erro de aproximadamente um em cada  $10^4$  nucleotídeos copiados, comparada com uma taxa de erro da DNA-polimerase de aproximadamente um em cada  $10^7$  nucleotídeos.

### Os diversos tipos de RNAs são produzidos nas células

A grande maioria dos genes que existem no DNA de uma célula codifica as seqüências de aminoácidos de proteínas, e as moléculas de RNA copiadas a partir desses genes (e que direcionam a síntese das proteínas) são coletivamente chamadas de **RNA mensageiro (mRNA)**. Em eucariotos, cada mRNA tipicamente possui informação transcrita a partir de um único gene e codifica uma única proteína. Em bactérias, um conjunto de genes adjacentes é frequentemente transcrito em um único mRNA que, consequentemente, possui informação para a produção de diferentes proteínas.

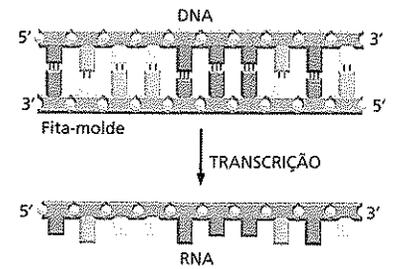
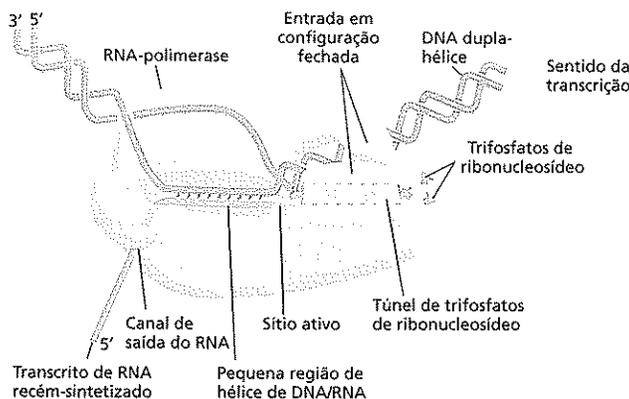
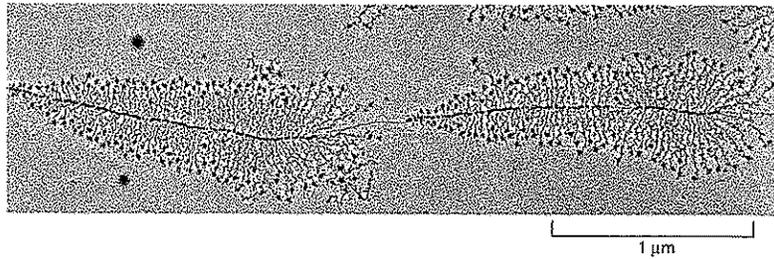


Figura 7-6 A transcrição produz uma molécula de RNA complementar a uma das fitas do DNA. A fita de DNA não molde (a fita superior neste exemplo) é algumas vezes denominada *fita codificadora*, pois sua seqüência é equivalente à seqüência do RNA produzido.

Figura 7-7 O DNA é transcrito pela enzima RNA-polimerase. A RNA-polimerase (azul-clara) se move paulatinamente sobre o DNA, desespiralizando a hélice de DNA à sua frente. Conforme avança, a polimerase adiciona nucleotídeos (pequenos objetos representados em "T") um a um à cadeia de RNA, no sítio de polimerização, usando uma fita de DNA exposta como molde. Consequentemente, o RNA transcrito é uma cópia complementar da fita simples de uma das duas fitas do DNA. Conforme se move sobre o DNA-molde, a polimerase desloca o RNA recém-formado, permitindo que as duas fitas de DNA se reassociem na região posterior da mesma. Uma região curta de hélice híbrida DNA/RNA (de aproximadamente nove nucleotídeos de comprimento) é formada temporariamente, fazendo com que uma "janela" da hélice DNA/RNA se mova ao longo do DNA junto à polimerase (Animação 7.2).

**Figura 7-8** A transcrição pode ser visualizada sob microscopia eletrônica. A microfotografia mostra diversas moléculas de RNA-polimerase transcrevendo simultaneamente dois genes adjacentes. As moléculas de RNA-polimerase são visualizadas como uma série de pontos ao longo do DNA, com os transcritos (filamentos finos) ligados a elas. As moléculas de RNA (denominadas rRNAs) transcritas a partir dos genes ilustrados neste exemplo não são traduzidas em proteína, sendo utilizadas diretamente como componentes dos ribossomos, as máquinas onde a tradução ocorre. Acredita-se que as partículas na extremidade 5' (a extremidade livre) de cada transcrito rRNA sejam proteínas ribossomais que se associaram ao rRNA. (Cortesia de Ulrich Scheer.)



O produto final de outros genes, no entanto, é o próprio RNA (Tabela 7-1). Como veremos em seções posteriores do presente capítulo, esses RNAs não mensageiros, assim como as proteínas, atuam como componentes regulatórios, estruturais e enzimáticos das células e desempenham papéis essenciais na tradução da mensagem genética em proteína. O *RNA ribossomal (rRNA)* forma a região central dos ribossomos, na qual o mRNA é traduzido em proteínas, e o *RNA transportador (tRNA)* forma os adaptadores que selecionam os aminoácidos e os posicionam no local adequado do ribossomo para que eles sejam incorporados em uma proteína. Outros pequenos RNAs, denominados *microRNAs (miRNAs)*, atuam como importantes reguladores na expressão gênica em eucariotos, como discutiremos no Capítulo 8.

Em seu sentido mais amplo, o termo **expressão gênica** se refere ao processo pelo qual a informação codificada na sequência de DNA é traduzida em um produto que desencadeia um efeito determinado sobre uma célula ou organismo. Nos casos em que o produto final do gene é uma proteína, a expressão gênica inclui tanto a transcrição quanto a tradução. Quando uma molécula de RNA é o produto final do gene, entretanto, a expressão gênica não requer a tradução.

**QUESTÃO 7-2**  
Na microfotografia eletrônica da Figura 7-8, as moléculas de RNA-polimerase estão se movendo da direita para a esquerda ou da esquerda para a direita? Por que os transcritos de RNA são muito mais curtos do que o DNA que os codifica?

**Sinais no DNA indicam os pontos de início e de término para a RNA-polimerase**

A iniciação da transcrição é um processo especialmente importante, pois é o principal momento no qual a célula pode selecionar quais proteínas ou RNAs que deverão ser produzidos e a sua taxa de síntese. Para dar início à transcrição, a RNA-polimerase deve ser capaz de reconhecer o início de um gene e ligar-se firmemente ao DNA sobre esse ponto. O modo pelo qual as RNA-polimerases reconhecem o sítio de início de transcrição difere consideravelmente entre bactérias e eucariotos. Visto que essa situação é mais simples em bactérias, inicialmente nos deteremos no sistema procaríoto.

Quando uma RNA-polimerase colide aleatoriamente com uma porção de DNA, ela se liga fracamente à dupla-hélice e começa a deslizar rapidamente sobre o DNA. A enzima se agarra fortemente ao DNA apenas quando encontra uma região denominada **promotor**, a qual contém uma sequência específica de nucleotídeos indicadora do ponto de iniciação para a síntese de RNA. Após ter feito contato com o promotor e ter-se fixado fortemente ao DNA, a RNA-polimerase

**TABELA 7-1** Tipos de RNAs produzidos nas células

Tipo de RNA	Função
mRNAs	Codificam proteínas
rRNAs	Formam a região central do ribossomo e catalisam a síntese proteica
miRNAs	Regulam a expressão de genes
tRNAs	Usados como adaptadores entre o mRNA e os aminoácidos durante a síntese proteica
Outros pequenos RNAs	Usados no <i>splicing</i> do mRNA, na manutenção de telômeros e em diversos outros processos celulares

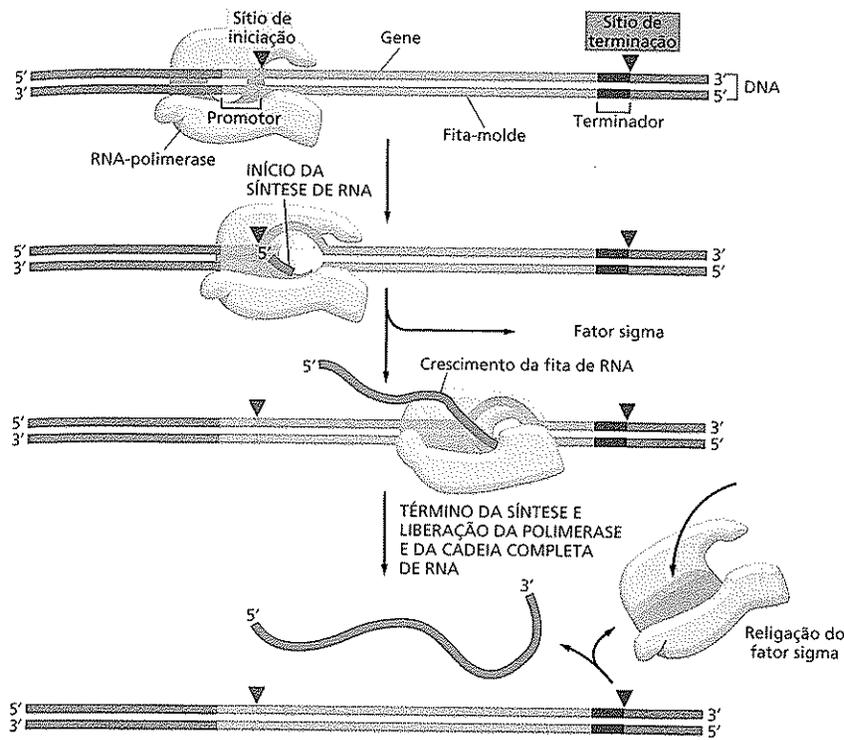


Figura 7-9 Sinais na sequência de um gene indicam à RNA-polimerase bacteriana onde iniciar e terminar a transcrição. A RNA-polimerase bacteriana (azul-claro) contém uma subunidade denominada fator sigma (amarelo) que reconhece o promotor (verde) sobre o DNA. Após o início da transcrição, o fator sigma é liberado, e a polimerase dá continuidade à síntese do RNA sem a presença do fator. A extensão da cadeia continua até que a polimerase encontre um sinal de terminação (vermelho) sobre o DNA. Nesse ponto, a enzima se detém e libera tanto a fita-molde de DNA quanto o rRNA recém-sintetizado. A seguir, a polimerase se reassocia a um fator sigma livre e recomeça a busca por outro promotor para reiniciar o processo.

abre a dupla-hélice imediatamente à sua frente e expõe os nucleotídeos de ambas as fitas, ao longo de uma curta extensão da molécula de DNA (Figura 7-9). A seguir, uma das duas fitas expostas do DNA atua como molde para o pareamento de bases por complementaridade a ser realizado com ribonucleotídeos que estão chegando, dois dos quais são unidos entre si pela polimerase para dar início à cadeia de RNA. Por meio desse sistema, a extensão da cadeia continua até que a enzima encontre um segundo sinal sobre o DNA, o *terminador* (ou sítio de parada), onde a polimerase se detém e libera tanto o DNA-molde quanto a cadeia de RNA recém-sintetizada (ver Figura 7-9).

Em bactérias, uma subunidade da polimerase, denominada *fator sigma* ( $\sigma$ ), é a principal responsável pelo reconhecimento da sequência promotora sobre o DNA. Após ter fixado fortemente a polimerase sobre o DNA nesse sítio, e essa ter sintetizado aproximadamente 10 nucleotídeos de RNA, o fator sigma é liberado, permitindo que a polimerase avance e continue a transcrição sem sua presença. Após ter sido liberada na região terminadora, a polimerase se reassocia a um fator sigma livre e parte em busca de um promotor, onde pode dar novo início ao processo de transcrição.

A proteína polimerase pode reconhecer o promotor, mesmo considerando-se que o DNA esteja sob sua forma de dupla-hélice, por meio do estabelecimento de contatos específicos com porções de bases que estão expostas na superfície externa da hélice. As sequências nucleotídicas de um promotor típico – e de um terminador típico – são apresentadas na Figura 7-10.

Visto que o DNA é uma fita dupla, um promotor pode, em princípio, direcionar a síntese de dois transcritos diferentes de RNA, um a ser transcrito rumo à esquerda e outro rumo à direita. No entanto, o promotor é assimétrico e se associa à polimerase sob uma orientação única; assim, uma vez adequadamente posicionada sobre um promotor, a RNA-polimerase não tem escolha a não ser transcrever a fita apropriada de DNA, pois a transcrição só ocorre no sentido de 5' – 3'. O sentido de transcrição, no que diz respeito ao cromossomo como um todo, pode variar de acordo com o gene que está sendo considerado (Figura 7-11). A necessidade de forte ligação da RNA-polimerase ao DNA antes que essa possa dar início à transcrição significa que um segmento de DNA só po-



**TABELA 7-2** As três RNA-polimerases de células eucarióticas

Tipo de polimerase	Genes transcritos
RNA-polimerase I	A maioria dos genes de rRNAs
RNA-polimerase II	Genes codificadores de proteínas, genes de miRNAs e genes para alguns pequenos RNAs (p. ex., aqueles referentes aos spliceossomos)
RNA-polimerase III	Genes de tRNAs Gene do rRNA 5S Genes de diversos outros pequenos RNAs

tram amplamente dispersos, existindo regiões de DNA com comprimento de até 100.000 pares de nucleotídeos entre um gene e o seguinte. Essa arquitetura permite que um único gene esteja sob o controle de um número praticamente ilimitado de sequências regulatórias espalhadas no DNA, permitindo que os eucariotos se engajem em sistemas de regulação de transcrição muito mais complexos do que os existentes nas bactérias.

- Por último, mas não menos importante, a iniciação da transcrição em eucariotos deve levar em consideração o empacotamento do DNA em nucleossomos e em formas mais compactas da estrutura de cromatina, como descrito no Capítulo 8.

Agora, voltaremos nossa atenção aos fatores gerais de transcrição e discutiremos como eles auxiliam a RNA-polimerase II dos eucariotos na iniciação da transcrição.

### A RNA-polimerase de eucariotos necessita de fatores gerais de transcrição

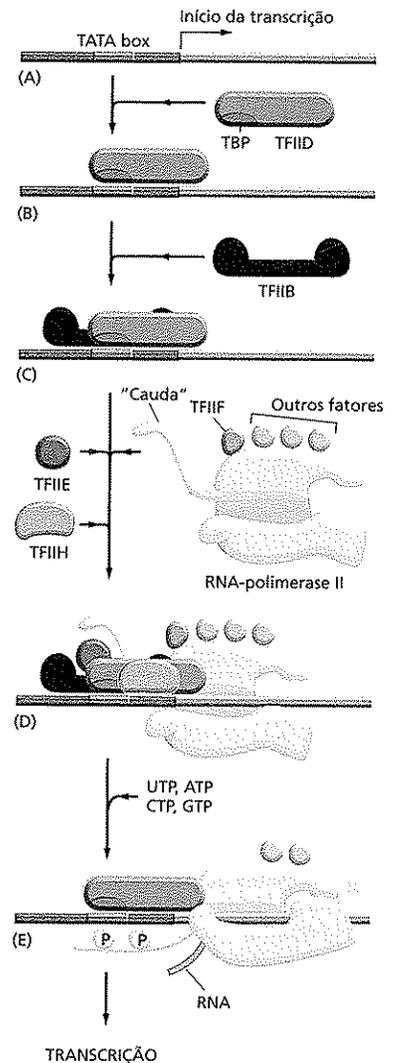
A descoberta inicial demonstrando que a RNA-polimerase II purificada de eucariotos, diferentemente da RNA-polimerase bacteriana, não era capaz de iniciar independentemente a transcrição *in vitro* levou à descoberta e à purificação dos **fatores gerais de transcrição**. Essas proteínas acessórias se ligam ao promotor, onde posicionam a RNA-polimerase, deslocam e separam a dupla-hélice para expor a fita-molde e direcionam a RNA-polimerase para dar início ao processo de transcrição.

A Figura 7-12 ilustra como ocorre a montagem e a ligação dos fatores gerais de transcrição sobre um promotor reconhecido pela RNA-polimerase II. Esse processo de montagem começa tipicamente por meio da ligação do fator geral de transcrição TFIID a uma pequena sequência de DNA de dupla-hélice composta predominantemente por nucleotídeos T e A; devido à sua composição, essa sequência é conhecida como sequência TATA ou **TATA box**. Por meio de sua ligação ao DNA, o TFIID provoca uma grande distorção local no DNA (Figura 7-13), a qual atua como uma marca sinalizadora para a subsequente montagem e agregação de outras proteínas sobre o promotor. O TATA box é um componente crucial de diversos promotores reconhecidos pela RNA-polimerase II e se encontra geralmente localizado a uma distância de 25 nucleotídeos antes do sítio de início

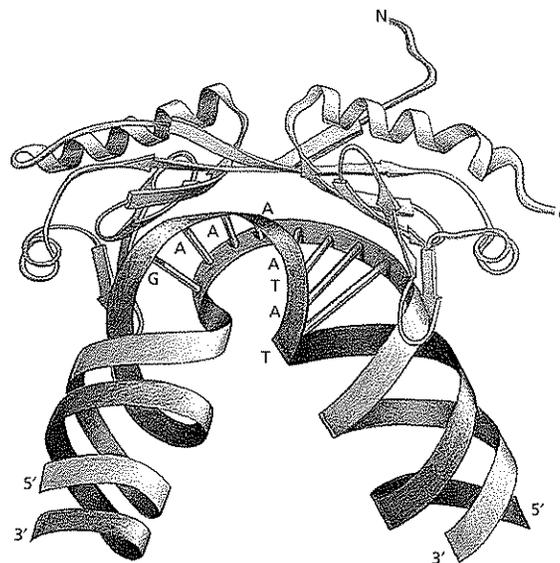
Figura 7-12 Para dar início à transcrição, a RNA-polimerase eucariótica necessita de um conjunto de fatores gerais de transcrição. Esses fatores de transcrição são denominados TFIIA, TFIIB e assim por diante. (A) Diversos promotores contêm uma sequência de DNA chamada de TATA box. (B) O TATA box é reconhecido pelo fator de transcrição TFIID que se liga a ele e permite a ligação subsequente do TFIIB. (C) Por questões de simplificação, a distorção do DNA produzida pela ligação do TFIID (ver Figura 7-13) não foi ilustrada. (D) Os demais fatores gerais de transcrição, assim como a própria RNA-polimerase, se associam ao promotor. (E) A seguir, o TFIIF separa as fitas da dupla-hélice no ponto de início da transcrição, usando energia da hidrólise de ATP, e faz com que a fita-molde seja exposta. O TFIIF também fosforila a RNA-polimerase II, liberando-a dos fatores gerais de transcrição, o que permite que tenha início a fase de extensão da transcrição. O sítio de fosforilação consiste em uma longa "cauda" polipeptídica que se estende a partir da molécula polimerase.

### QUESTÃO 7-3

Pode a RNA-polimerase usada na transcrição ser usada como polimerase para produzir o iniciador necessário para a replicação (discutida no Capítulo 6)?



**Figura 7-13** A proteína de ligação ao TATA (TBP) se liga a sequências TATA box e provoca distorções no DNA. A TBP é a subunidade do fator geral de transcrição TFIID responsável pelo reconhecimento e ligação à sequência TATA box (indicada por letras). A flexão característica do DNA, provocada pela TBP – duas torções na dupla-hélice separadas por DNA parcialmente desespiralizado –, pode provavelmente atrair os outros fatores gerais de transcrição. A TBP é uma cadeia polipeptídica única dobrada em dois domínios bastante similares (azul e verde). Seus oito folhetos β se deitam sobre a hélice do DNA como uma sela sobre um cavalo (Animação 7.4). (Adaptado de J.L. Kim et al., *Nature* 365:520-527, 1993. Com permissão de Macmillan Publishers Ltd.)



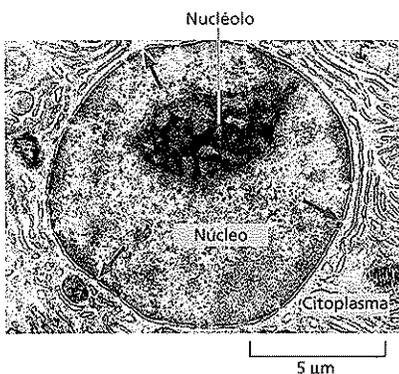
da transcrição. Uma vez que o primeiro fator geral de transcrição se tenha associado a esse sítio do DNA, os outros fatores e a RNA-polimerase II também são associados, para a formação de um *complexo de iniciação de transcrição* completo. Embora Figura 7-12 ilustre os fatores gerais de transcrição associando-se sobre o promotor em uma ordem determinada, a sequência exata de montagem provavelmente é diferente entre diferentes promotores presentes em uma célula.

Depois de a RNA-polimerase ter sido conectada ao DNA promotor por meio do complexo de iniciação de transcrição, ela deve ser liberada do complexo de fatores de transcrição para que possa dar início à sua tarefa de produção de uma molécula de RNA. Uma etapa essencial para essa liberação consiste na adição de grupos fosfato à "cauda" da RNA-polimerase, ação essa desempenhada pelo fator geral de transcrição TFIIF, o qual contém uma das suas subunidades com atividade de enzima proteína-cinase (ver Figura 7-12E). Acredita-se que essa fosforilação auxilie a dissociação da polimerase do grupo de fatores de transcrição, permitindo que a transcrição tenha início. Uma vez a transcrição iniciada, a maioria dos fatores de transcrição é liberada do DNA, de tal forma que essas moléculas estarão disponíveis para iniciar outro ciclo de transcrição com outra molécula de RNA-polimerase. Quando a RNA-polimerase II finaliza a transcrição, ela é liberada do DNA, os fosfatos de sua cauda são eliminados por meio de fosfatases, e ela pode reiniciar outro processo de transcrição. Apenas a forma defosforilada da RNA-polimerase II é capaz de dar início à síntese de RNA sobre um promotor.

### Os RNAs eucarióticos são transcritos e processados simultaneamente no núcleo

Apesar de o princípio de molde pelo qual o DNA é transcrito em RNA ser o mesmo em todos os organismos, o modo segundo o qual os transcritos são manipulados antes de poderem ser utilizados pela célula difere bastante entre bactérias e eucariotos. O DNA bacteriano permanece exposto diretamente no citoplasma, onde se localizam os *ribossomos* nos quais a síntese proteica ocorre. Conforme as moléculas de RNA bacteriano estão sendo transcritas, os ribossomos imediatamente se ligam à extremidade 5' livre do transcrito de RNA e dão início à síntese proteica.

Nas células eucarióticas, em contraste, o DNA está isolado dentro do *núcleo*. A transcrição ocorre no núcleo, mas a síntese de proteínas ocorre nos ribossomos que se encontram no citoplasma. Desse modo, antes que um mRNA eucariótico possa ser traduzido, ele deverá ser transportado para fora do núcleo por pequenos poros existentes no envelope nuclear (Figura 7-14). No entanto, antes que um mRNA



**Figura 7-14** Antes de poderem ser traduzidas, as moléculas de mRNA produzidas no núcleo migram para o citoplasma pelos poros presentes no envelope nuclear (setas). Aqui está ilustrada uma seção do núcleo de um hepatócito. (De D.W. Fawcett, *A Textbook of Histology*, 11. ed. Filadélfia: Saunders, 1986. Com permissão de Elsevier.)

eucariótico saia do núcleo, ele deverá passar por uma série de diferentes etapas que compõem o **processamento do RNA**. Essas reações ocorrem enquanto o RNA ainda está sendo transcrito. As enzimas responsáveis pelo processamento do RNA se associam à "cauda" da RNA-polimerase eucariótica quando esta está transcrevendo um RNA, podendo saltar sobre a molécula de RNA nascente e começar a processá-la no momento em que essa emerge da RNA-polimerase (Figura 7-15).

Dependendo do tipo de RNA que está sendo produzido – um mRNA ou algum outro tipo –, os transcritos são processados de um modo específico antes de deixar o núcleo. Duas etapas do processamento que ocorrem apenas em transcritos destinados a tornarem-se moléculas mRNA são o *capeamento* e a *poliadenilação* (Figura 7-16):

1. O *capeamento* do RNA envolve uma modificação da extremidade 5' do transcrito de mRNA, a extremidade que é sintetizada em primeiro lugar durante a transcrição. O RNA é capeado pela adição de um nucleotídeo atípico – um nucleotídeo guanina (G) com um grupo metila associado. Esse *capeamento* ocorre após a RNA-polimerase ter sintetizado um fragmento de aproximadamente 25 nucleotídeos de RNA, muito antes que a transcrição do gene esteja completa.
2. A *poliadenilação* provê uma estrutura especial, ou cauda, para a extremidade de 3' da maioria dos mRNAs recém-sintetizados. Em contraste com as bactérias, onde a extremidade 3' de um mRNA é simplesmente o final da cadeia sintetizada pela RNA-polimerase, as extremidades 3' de RNAs eucarióticos são inicialmente clivadas por uma enzima que corta o RNA em uma sequência particular de nucleotídeos e, a seguir, são complementadas por uma segunda enzima que adiciona uma série repetitiva de nucleotídeos adenina (A) (uma cauda poli-A) sobre a extremidade clivada. Essa *cauda poli-A* geralmente possui um comprimento de algumas centenas de nucleotídeos.

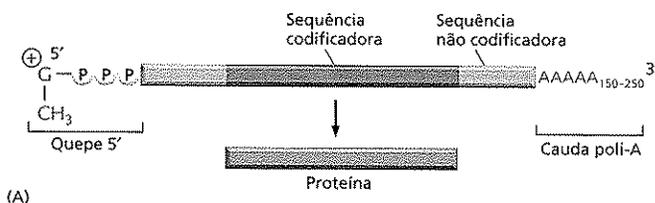
Acredita-se que essas duas modificações – *capeamento* e *poliadenilação* – aumentem a estabilidade da molécula de mRNA eucariótica, auxiliando sua exportação do núcleo para o citoplasma e sua identificação geral como mRNA. Elas também são utilizadas pela maquinaria de síntese de proteínas, antes do início da síntese, como um indicador de que ambas as extremidades do mRNA estão presentes e, conseqüentemente, de que essa mensagem está completa.

### Os genes eucarióticos são interrompidos por sequências não codificadoras

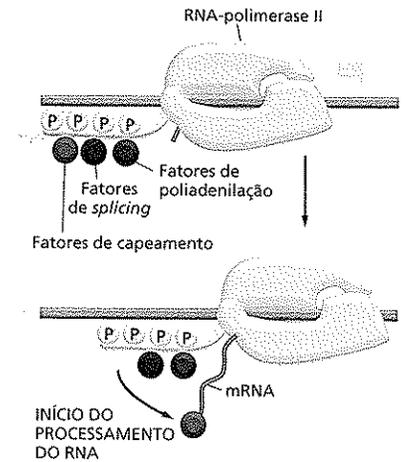
A maioria dos RNAs eucarióticos passa por uma etapa adicional de processamento antes de tornarem-se uma molécula funcional. Essa etapa envolve uma

**Figura 7-16** As moléculas de mRNA eucariótico são modificadas pelo *capeamento* e pela *poliadenilação*. (A) As extremidades de um mRNA eucariótico são modificadas por meio da adição de um quepe na extremidade 5' e pela clivagem do transcrito primário e adição de uma cauda poli-A na extremidade 3'. (B) A estrutura do quepe na extremidade 5' de moléculas de mRNA eucarióticas. Observe a ligação incomum 5'-5' do 7-metil G ao restante do RNA. Diversos quepes de mRNA eucarióticos apresentam uma modificação adicional: a metilação de um grupo 2'-hidroxila do segundo açúcar-ribose no mRNA (não ilustrada).

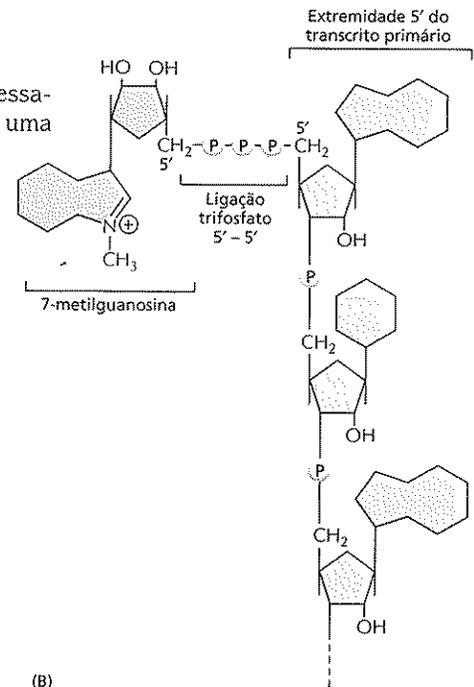
**Capeamento e poliadenilação do RNA**



(A)

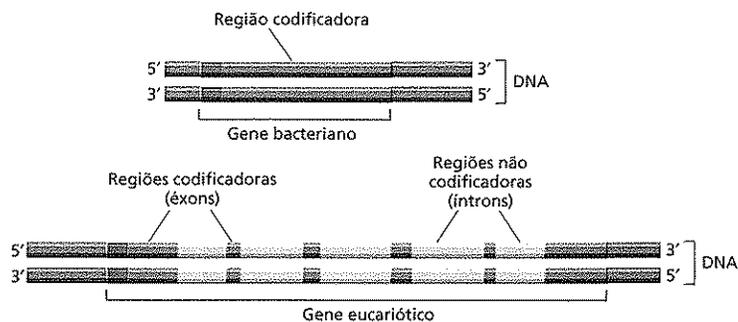


**Figura 7-15** A fosforilação da RNA-polimerase II permite que proteínas de processamento de RNA se associem à sua cauda. A polimerase não apenas transcreve o DNA, como também carrega as proteínas responsáveis pelo processamento do RNA, as quais atuam sobre o RNA recém-sintetizado. Algumas proteínas de processamento do RNA se ligam à cauda da RNA-polimerase quando essa é fosforilada no período tardio do processo de iniciação da transcrição (ver Figura 7-12). O *capeamento*, a *poliadenilação* e o *splicing* (descritos mais tarde neste capítulo) são modificações efetuadas sobre o RNA durante o processamento.



(B)

**Figura 7-17** Genes eucarióticos e bacterianos são organizados de forma distinta. Um gene bacteriano consiste em uma única porção de sequência nucleotídica não interrompida que codifica a sequência de aminoácidos de uma proteína. Em contraste, as sequências codificadoras (éxons) da maioria dos genes em eucariotos são interrompidas por sequências 5' não codificadoras (íntrons). Os promotores de transcrição estão indicados em verde.



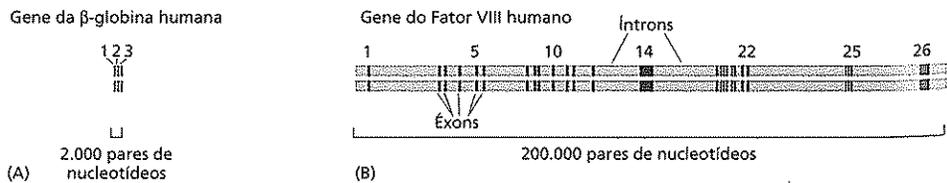
modificação bem mais radical sobre o transcrito primário de RNA se comparada ao capeamento e à poliadenilação, anteriormente citados, e é o resultado de uma característica surpreendente de arranjo dos genes eucarióticos. Em bactérias, a maior parte das proteínas é codificada a partir de um segmento não interrompido da sequência de DNA, o qual é transcrito em RNA que, sem qualquer outro processamento, pode atuar como mRNA. A maioria dos genes eucarióticos, no entanto, apresenta as suas sequências codificadoras interrompidas por longas *sequências intervenientes* não codificadoras, denominadas **íntrons** (Figura 7-17). As porções esparsas de sequências codificadoras, ou *sequências expressas*, denominadas **éxons**, são geralmente mais curtas do que as sequências de íntrons, sendo a porção codificadora de um gene eucariótico frequentemente apenas uma pequena fração do comprimento total do gene. O comprimento dos íntrons varia de um único nucleotídeo a mais de 10.000 nucleotídeos. Alguns genes eucarióticos não possuem qualquer íntron, e outros possuem apenas uns poucos, mas a maioria apresenta diversos íntrons (Figura 7-18).

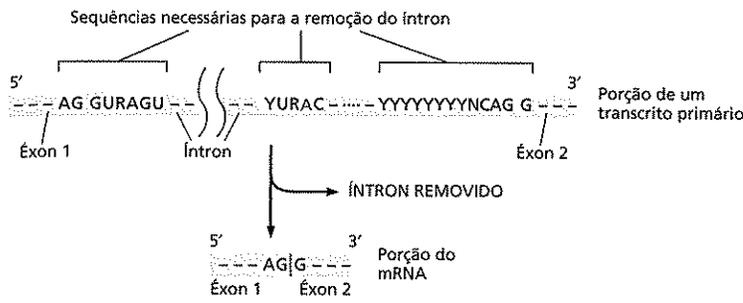
### Os íntrons são removidos pelo *splicing* de RNA

Para produzir um mRNA em uma célula eucariótica, o gene, em sua totalidade, incluindo tanto íntrons quanto éxons, é transcrito em RNA. Após o capeamento, e conforme a RNA-polimerase continua a transcrição do gene, tem início o processo de ***splicing* do RNA**, durante o qual as sequências íntron são removidas do RNA recém-sintetizado e as sequências éxon são unidas umas às outras. Finalmente, cada transcrito recebe uma cauda poli-A; em alguns casos, essa etapa ocorre após o *splicing*, ao passo que em outros casos, essa etapa ocorre antes que as reações de *splicing* estejam completas. Se um transcrito já sofreu *splicing* e ambas as extremidades 5' e 3' foram modificadas, esse RNA é uma molécula funcional que pode então deixar o núcleo e ser traduzida em proteína.

Como a célula determina quais segmentos do transcrito primário serão removidos durante o *splicing*? Diferentemente da sequência codificadora de um éxon, a sequência nucleotídica da maior parte de um íntron parece não ser importante. Apesar de existir pouca semelhança, em geral, entre as sequências nucleotídicas de diferentes íntrons, cada íntron contém umas poucas sequências nucleotídicas curtas essenciais que direcionam sua remoção. Essas sequências se encontram nos limites do íntron ou próximas a eles e são idênticas ou bastante similares entre todos os íntrons (Figura 7-19). Guiada por essas sequências, uma elaborada maquinaria de *splicing* remove o íntron sob a forma de uma estrutura em "laço" (Figura 7-20) produzida a partir da reação do "A" salientado em vermelho nas Figuras 7-19 e 7-20.

**Figura 7-18** Os genes humanos, em sua maioria, estão divididos em éxons e íntrons. (A) O gene da  $\beta$ -globina, que codifica uma das subunidades da proteína transportadora de oxigênio hemoglobina, contém 3 éxons. (B) O gene do Fator VIII codifica uma proteína (Fator VIII) que opera na via da coagulação sanguínea e contém 26 éxons. Mutações nesse grande gene são responsáveis pela forma mais prevalente de hemofilia.





Apesar de não nos determos em detalhes na maquinaria do *splicing*, é importante observar que, diferentemente do que ocorre nas outras etapas de produção do mRNA que foram discutidas, o *splicing* do RNA é realizado predominantemente por moléculas de RNA, em vez de proteínas. Moléculas de RNA reconhecem os limites entre éxon-intron (por meio de pareamento de bases por complementaridade) e participam intimamente das reações químicas do *splicing*. Essas moléculas de RNA, denominadas **pequenos RNAs nucleares (snRNAs, de small nuclear RNAs)** estão agregadas a proteínas adicionais para formar as *pequenas partículas ribonucleoproteicas nucleares (snRNPs, de small nuclear ribonucleoprotein particles)*. Esses snRNPs formam o núcleo do **spliceossomo**, o grande arranjo de RNA e de moléculas proteicas que realiza o *splicing* na célula. Para visualizar o spliceossomo em ação, ver Animação 7.5.

O tipo de arranjo que inclui éxons-introns em eucariotos pode parecer inicialmente um grande gasto desnecessário; no entanto, ele leva a consequências positivas. Em primeiro lugar, os transcritos de diversos genes eucarióticos podem ser processados por *splicing* sob diferentes formas, cada uma delas levando à produção de uma proteína distinta. Esse tipo de ***splicing* alternativo** permite, portanto, que diferentes proteínas sejam produzidas a partir de um mesmo gene (Figura 7-21). Estima-se que aproximadamente 60% dos genes humanos sejam capazes de realizar esse tipo de *splicing* alternativo. Dessa forma, o *splicing* do RNA permite que os eucariotos elevem astronomicamente o potencial de codificação de seus genomas.

O *splicing* do RNA também fornece outra vantagem aos eucariotos, uma que provavelmente desempenhou um papel extremamente importante na história evolutiva inicial dos genes. Como será discutido detalhadamente no Capítulo 9, acredita-se que a estruturação em éxons-introns tenha acelerado o surgimento de proteínas novas e úteis. A existência de introns longos torna a recombinação genética entre éxons de diferentes genes mais provável. Isso significa que genes para novas proteínas podem ter evoluído de forma bastante rápida por meio da combinação de segmentos de genes preexistentes, um mecanismo semelhante à montagem de um novo tipo de máquina a partir de um *kit* que contenha uma série de componentes funcionais pré-testados. Várias proteínas de células atuais se assemelham a uma colcha-de-retalhos composta a partir de um conjunto-padrão de peças proteicas, denominadas *domínios* proteicos (ver Figura 4-16).

### Os mRNAs eucarióticos maduros são seletivamente exportados do núcleo

Vimos como acontecem, ordenadamente, a síntese e o processamento do mRNA eucariótico dentro do núcleo da célula. No entanto, esses eventos geram um problema específico para as células eucarióticas: do mRNA total que é sintetizado, apenas uma pequena fração – o mRNA maduro – é útil para a célula. Os fragmentos restantes de RNA – introns excisados, RNAs quebrados e transcritos que sofreram *splicing* de forma inadequada – não apenas são inúteis, como podem também ser perigosos para a célula se não forem destruídos. Como, então, a célula distingue entre as moléculas relativamente raras de mRNA maduro que ela necessita manter e a enorme quantidade de fragmentos gerados pelo processamento de RNA?

Figura 7-19 Sequências especiais de nucleotídeos sinalizam o início e o final de um intron. Apenas as sequências nucleotídicas mostradas são necessárias para a remoção de um intron. As demais posições em um intron podem ser ocupadas por qualquer nucleotídeo. Essas sequências especiais são reconhecidas por pequenas ribonucleoproteínas nucleares (snRNPs), que clivam o RNA na interface entre introns-éxons e unem covalentemente os éxons. Na ilustração, R pode representar tanto um A quanto um G; Y representa tanto C quanto U; N se refere a qualquer um dos aminoácidos. O A salientado em *vermelho* forma o ponto de forquilha da alça produzida na reação de *splicing* (ver Figura 7-20). As distâncias sobre o RNA, entre as três sequências de *splicing*, são extremamente variáveis; entretanto, a distância entre o ponto de forquilha e a junção 5' do *splicing* é caracteristicamente muito maior do que a distância entre a junção 3' do *splicing* e o ponto de forquilha. As sequências de *splicing* ilustradas se referem a humanos; sequências similares direcionam o *splicing* do RNA em outros organismos eucarióticos.

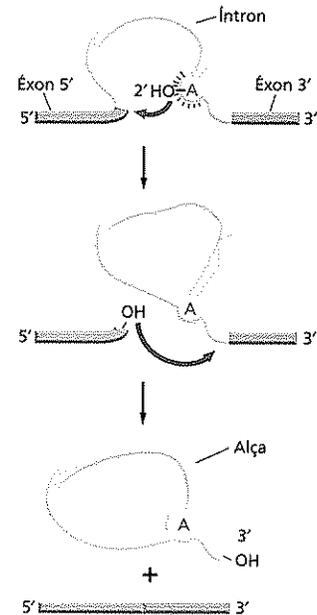
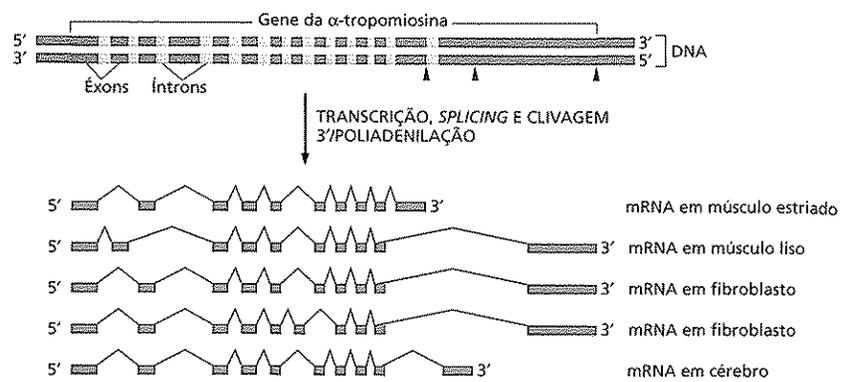


Figura 7-20 Um intron forma uma estrutura ramificada durante o *splicing*. No primeiro passo, a adenina do ponto de forquilha (A em *vermelho*), na sequência do intron, ataca o sítio 5' de *splicing* e corta a estrutura de açúcar-fosfato do RNA nesse ponto (essa é a mesma adenina que está salientada em *vermelho* na Figura 7-19). A extremidade 5' cortada do intron se torna covalentemente ligada ao grupo 2'-OH da ribose do A formando uma estrutura ramificada. A seguir, a extremidade 3'-OH livre, da sequência do éxon, reage com a sequência inicial do éxon seguinte, o que une os dois éxons em uma sequência codificadora contínua e libera o intron sob a forma de um laço, o qual é então degradado.

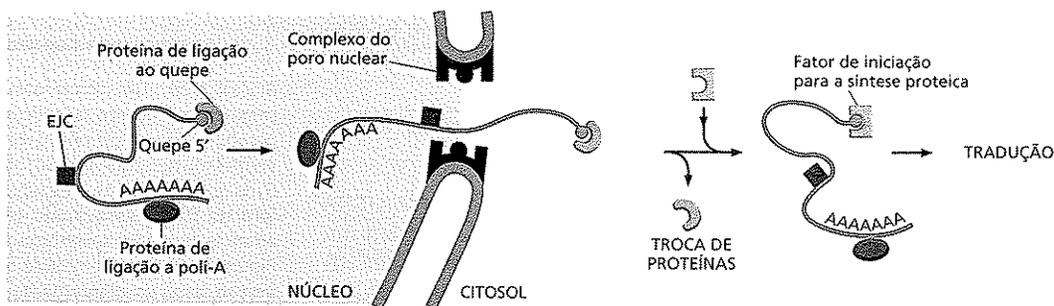
**Figura 7-21** O gene da  $\alpha$ -tropomiosina pode sofrer *splicing* de diferentes maneiras. A  $\alpha$ -tropomiosina é uma proteína supertorcida (ver Figura 4-13) que regula a contração em células musculares. O seu transcrito primário pode sofrer *splicing* em diferentes formas, como indicado na figura, produzindo mRNAs distintos que, a seguir, darão origem a variantes proteicas. Alguns dos padrões de *splicing* são específicos de determinados tipos de células. Por exemplo, a  $\alpha$ -tropomiosina produzida em músculo estriado é diferente daquela produzida a partir do mesmo gene na musculatura lisa. As setas vermelhas, na região superior da figura, representam os sítios onde a adição de poli-A pode ocorrer.



A resposta é que o transporte do mRNA do núcleo para o citoplasma, onde ele será traduzido em proteína, é altamente seletivo: apenas RNAs adequadamente processados podem ser transportados. Esse acoplamento entre um processamento adequado e o transporte do RNA é mediado pelo *complexo do poro nuclear*, que reconhece e transporta apenas mRNAs finalizados. Esses poros aquosos conectam o nucleoplasma ao citosol e, como será discutido no Capítulo 15, atuam como portões que controlam quais macromoléculas podem entrar ou sair do núcleo. Para estar "pronta para a exportação", uma molécula de mRNA deve estar ligada a um conjunto apropriado de proteínas, onde todas sinalizam que o mRNA foi corretamente processado. Dentre essas proteínas podemos citar as proteínas de ligação à poli-A, um complexo de ligação ao quepe e proteínas que marcam RNAs que sofreram *splicing* total (Figura 7-22). É provavelmente o conjunto completo de proteínas ligadas, muito mais do que uma proteína única específica, que determina, em última instância, se uma molécula de RNA poderá deixar o núcleo. Os "RNAs lixo" que permanecem no núcleo são degradados, e suas unidades de construção são reutilizadas na transcrição.

### As moléculas de mRNA são finalmente degradadas pela célula

Visto que uma única molécula de mRNA pode ser traduzida muitas vezes (ver Figura 7-2), o intervalo de tempo que uma molécula madura de mRNA permanece na célula afeta a quantidade de proteína produzida. Todas as moléculas de mRNA são finalmente degradadas em nucleotídeos por RNases celulares, mas o tempo de vida das moléculas de mRNA difere consideravelmente – dependendo da sequência nucleotídica do mRNA e do tipo de célula na qual o mRNA é produzido. A maioria dos mRNAs produzidos em bactérias é rapidamente degradada,



**Figura 7-22** Um conjunto especializado de proteínas de ligação ao RNA sinaliza que o mRNA maduro está pronto para ser exportado para o citoplasma. Como indicado à esquerda, o quepe e a cauda poli-A de uma molécula madura de mRNA estão "marcadas" por proteínas que reconhecem essas modificações. Além disso, um grupo de proteínas denominadas *complexo de junção do éxon (EJC)* é depositado sobre o mRNA após o RNA ter sofrido o *splicing* adequadamente. Quando o mRNA é considerado "pronto para exportação", um receptor de transporte nuclear (discutido no Capítulo 15) se associa a ele guiando-o através do poro nuclear. Uma vez no citosol, o mRNA pode perder algumas proteínas anteriormente ligadas e adquirir outras.

apresentando caracteristicamente um tempo médio de vida de aproximadamente 3 minutos. Os mRNAs de células eucarióticas geralmente persistem por períodos mais longos. Alguns transcritos, como aqueles que codificam a  $\beta$ -globina, apresentam um tempo médio de vida de mais de 10 horas, ao passo que outros mRNAs eucarióticos possuem um tempo médio inferior a 30 minutos.

Essas diferenças no tempo de vida são em parte controladas por sequências nucleotídicas que se encontram sobre o próprio mRNA, mais frequentemente na porção do RNA denominada região 3' não traduzida, que se localiza entre a extremidade 3' da sequência codificadora e a cauda poli-A. As diferenças na duração do mRNA auxiliam a célula a determinar a quantidade de cada proteína que é sintetizada. Em geral, proteínas produzidas em grandes quantidades, como a  $\beta$ -globina, são traduzidas a partir de mRNA que possuem longa duração, ao passo que aquelas proteínas presentes em baixos níveis, ou aquelas cujos níveis devem ser alterados rapidamente em resposta à sinalização, são tipicamente sintetizadas a partir de mRNAs de curta duração. Essas diferenças no tempo de vida do mRNA são o auge da sintonia fina evolutiva, em que a estabilidade dos mRNAs está atrelada às necessidades da célula.

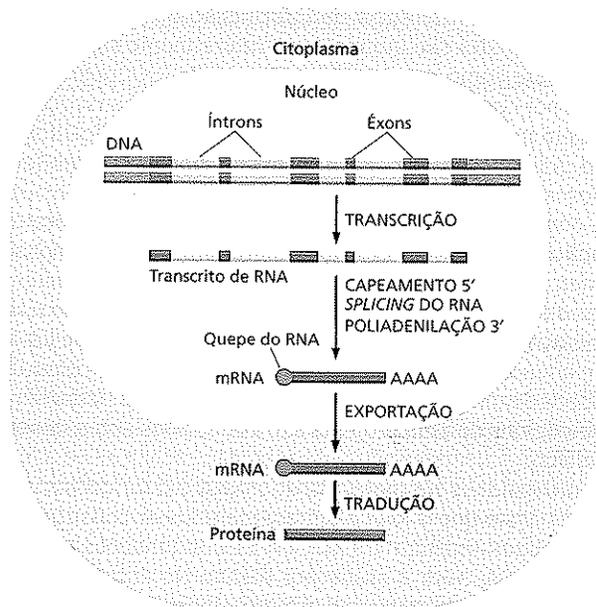
### As primeiras células devem ter possuído íntrons em seus genes

O processo de transcrição é universal: todas as células usam RNA-polimerase associada ao sistema de complementaridade de bases, para sintetizar RNA a partir de DNA. Além disso, as RNA-polimerases bacterianas e eucarióticas são praticamente idênticas em termos de estrutura geral e certamente evoluíram a partir de uma polimerase ancestral comum. É, portanto, intrigante que o transcrito resultante seja tratado de forma tão diferente em eucariotos e procariotos (Figura 7-23). Em particular, o *splicing* do RNA parece indicar uma diferença fundamental entre esses dois tipos de células; mas, afinal, como teve origem essa diferença tão marcante?

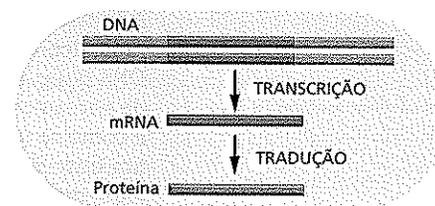
Como vimos, o *splicing* do RNA permite aos eucariotos a capacidade de produzir uma ampla variedade de proteínas a partir de um único gene e confere a eles um certo grau de flexibilidade evolutiva. No entanto, essas vantagens têm um custo: a célula deve manter um grande genoma e desprezar uma grande parcela do RNA que sintetiza. De acordo com uma linha de pensamento, as

Figura 7-23 Procariotos e eucariotos "manipulam" diferentemente seus transcritos de RNA. (A) Em células eucarióticas, a molécula inicial de RNA, produzida pela transcrição, contém tanto sequências de íntrons quanto de éxons. Suas duas extremidades são modificadas, e os íntrons são removidos por uma reação enzimaticamente catalisada de *splicing* do RNA. A seguir, o mRNA resultante é transportado do núcleo para o citoplasma, onde é traduzido em proteína. Apesar de esses passos estarem ilustrados como se ocorressem um de cada vez, na realidade eles ocorrem simultaneamente. Por exemplo, o quepe do RNA é caracteristicamente adicionado, e o *splicing* inicia antes que o transcrito esteja completo. Em virtude desse acoplamento, transcritos primários completos (que incluem todos éxons e íntrons) não são geralmente encontrados nas células. (B) Em procariotos, a produção de moléculas de mRNA é mais simples. A extremidade 5' de uma molécula de RNA é produzida na iniciação da transcrição pela RNA-polimerase, e a extremidade 3' é produzida pelo término da transcrição. Visto que células procarióticas não possuem núcleo, a transcrição e a tradução ocorrem em um mesmo compartimento. Assim, a tradução de um mRNA bacteriano tem início antes que sua síntese esteja completa. A quantidade de proteína em uma célula depende da eficiência de cada um desses passos e da taxa de degradação das moléculas de RNA e proteína.

(A) EUCARIOTOS



(B) PROCARIOTOS



células primordiais – os ancestrais comuns de procariotos e eucariotos – continham íntrons que foram perdidos pelos procariotos ao longo de sua evolução subsequente. Pelo abandono de seus íntrons e adoção de um genoma menor, mais fluido, os procariotos foram capazes de se reproduzir mais rápida e eficientemente. Corroborando essa ideia, eucariotos simples que se reproduzem rapidamente (p. ex., algumas leveduras) possuem relativamente poucos íntrons, e esses íntrons são em geral muito menores do que aqueles encontrados em eucariotos superiores.

Por outro lado, alguns argumentam que os íntrons se originaram a partir de elementos genéticos móveis parasitas (discutidos no Capítulo 6) que invadiram um ancestral eucariótico primordial e colonizaram seu genoma. Essas células hospedeiras então, inadvertidamente, replicaram esses fragmentos egoístas de nucleotídeos junto a seu próprio DNA, e os eucariotos modernos nunca se preocuparam em eliminar ou arrumar a bagunça genética deixada por essas infecções ancestrais. A questão, no entanto, está longe de ser definida; se os íntrons evoluíram no início – e foram perdidos pelos procariotos – ou evoluíram tardiamente nos eucariotos, é uma questão atual de debate científico, sobre a qual retornaremos no Capítulo 9.

## DO RNA À PROTEÍNA

No final da década de 1950, os biólogos haviam demonstrado que a informação codificada no DNA era inicialmente copiada em RNA e a seguir em proteína. O debate estava centrado no “problema da codificação”: como uma informação sob a forma de uma sequência linear de nucleotídeos no RNA era traduzida para a forma de uma sequência linear de um conjunto de subunidades quimicamente tão distintas – os aminoácidos das proteínas? Essa fascinante questão excitou e estimulou enormemente os cientistas da época. Esse era um quebra-cabeça proposto pela natureza que, após mais de 3 bilhões de anos de evolução, poderia ser resolvido por um dos produtos dessa evolução – os seres humanos! E foi o que aconteceu, não apenas o código foi finalmente decifrado e compreendido em nível molecular, como foram também estabelecidas as principais características da maquinaria por meio da qual as células leem esse código.

### Uma sequência de mRNA é decodificada em grupos de três nucleotídeos

A transcrição como forma de transferência de informação é simples de compreender, visto que o DNA e o RNA são química e estruturalmente similares, e o DNA pode atuar diretamente como molde para a síntese de RNA pelo sistema de pareamento de bases por complementaridade. Como o termo transcrição diz, é como se uma mensagem manuscrita estivesse sendo convertida, digamos, em um texto datilografado. A linguagem *per se* e a forma da mensagem não foram alteradas, e os símbolos utilizados são bastante semelhantes.

Em contraste, a conversão da informação contida no RNA para proteína representa uma **tradução** da informação em outra linguagem, composta por símbolos bastante diferentes. Visto existirem apenas quatro nucleotídeos diferentes no mRNA, mas 20 tipos diferentes de aminoácidos em uma proteína, essa tradução não pode acontecer por um sistema de correspondência direto entre um nucleotídeo no RNA e um aminoácido na proteína. As regras que ditam como uma sequência de nucleotídeos de um gene, por intermédio do mRNA, é traduzida em uma sequência de aminoácidos de uma proteína são conhecidas sob a denominação de **código genético**.

A sequência de nucleotídeos na molécula de mRNA é lida, consecutivamente, em grupos de três. Visto que o RNA é um polímero linear constituído de quatro diferentes nucleotídeos, existem  $4 \times 4 \times 4 = 64$  combinações possíveis de três nucleotídeos: AAA, AUA, AUG, e assim por diante. No entanto, apenas 20 aminoácidos são geralmente encontrados em proteínas. Dessa forma, ou



## COMO SABEMOS: DECIFRANDO O CÓDIGO GENÉTICO

No início da década de 1960, o *dogma central* havia sido aceito como representativo da via através da qual a informação fluía do gene para a proteína. Estava claro que os genes codificavam proteínas, que os genes eram feitos de DNA e que o mRNA funcionava como um intermediário, transportando a informação do núcleo – onde o DNA está estocado – para o citoplasma, onde acontece a tradução das proteínas.

Até mesmo o formato geral do código genético estava compreendido: cada um dos 20 aminoácidos encontrados nas proteínas é representado por um códon (tripleto) sobre uma molécula de mRNA. No entanto, um desafio ainda maior permanecia: biólogos, químicos e mesmo físicos concentravam seus esforços buscando a quebra do código – tentando desvendar qual aminoácido era codificado por cada um dos 64 possíveis tripletes de nucleotídeos. A via mais segura para a solução dessa questão seria a comparação da sequência de um segmento de DNA ou mRNA com seu produto polipeptídico correspondente. No entanto, as técnicas de sequenciamento de ácidos nucleicos só ficariam disponíveis no final da década de 1960.

Assim, os cientistas decidiram que, para decifrar o código genético, eles teriam de sintetizar suas próprias mensagens simples. Se eles pudessem direcionar essas moléculas para os ribossomos – as máquinas produtoras de proteínas – e a seguir analisar o produto proteico resultante, então estariam no rumo certo em relação à compreensão dos tripletes que correspondiam aos aminoácidos.

### Abandonando as células

Antes de os pesquisadores começarem a preparar seus mRNAs sintéticos, eles teriam de aperfeiçoar um sistema livre de células para a síntese de proteínas. Isso permitiria que eles traduzissem as mensagens em polipeptídeos dentro de tubos de ensaio. (De modo geral, quando se trabalha em laboratório, quanto mais simples o sistema utilizado, mais claros e fáceis de interpretar são os resultados.) Para isolar a maquinaria molecular que necessitavam para um dado sistema de tradução livre de células, os pesquisadores disrupturaram células de *E. coli* e colocaram o seu conteúdo em uma centrífuga. A centrifugação dessas amostras, em alta velocidade, fazia com que as membranas e outros grandes fragmentos celulares fossem levados para o fundo do tubo; os componentes celulares necessários para a síntese de proteínas, mais leves – como mRNA, tRNA, ribossomos, enzimas e outras moléculas pequenas –, permaneciam em suspensão no sobrenadante. Os pesquisadores descobriram que a simples adição de aminoácidos radioativos a essa “sopa” celular poderia induzir a produção de proteínas radiomarcadas. Por meio de uma nova centrifugação desse sobrenadante, sob uma velocidade um pouco maior, era possível forçar a deposição de ribossomos e dos peptídeos recém-sintetizados no fundo do tubo; os polipeptídeos marcados podiam então ser detectados pela medida da radioatividade remanescente no sedimento do tubo após descarte da fase aquosa superior.

O problema desse sistema em particular era que ele produzia as proteínas codificadas pelos mRNAs celulares que já estavam presentes no extrato, e os pesquisadores

queriam usar suas próprias mensagens sintéticas para direcionar a síntese de proteínas. Esse problema foi resolvido quando Marshall Nirenberg descobriu que poderia destruir RNA celular no extrato pela adição de uma pequena quantidade de ribonuclease – uma enzima que degrada RNA. A partir desse ponto, a próxima etapa era a produção de grandes quantidades da mensagem que ele adicionaria, a colocação desse mRNA sintético sobre o sistema livre de células e a análise dos peptídeos resultantes.

### Falsificando a mensagem

A produção de polinucleotídeos com uma sequência definida não foi tão simples como se pretendia. Novamente, seriam necessários anos antes que os químicos desenvolvessem as técnicas que pudessem ser utilizadas para a síntese de uma fita definida de nucleotídeos. Nirenberg decidiu utilizar a polinucleotídeo-fosforilase, uma enzima que unia ribonucleotídeos entre eles, sem a presença de um molde. Assim, a sequência de RNA resultante dependeria exclusivamente dos nucleotídeos que estivessem disponíveis para a enzima. Uma mistura de nucleotídeos seria sintetizada formando uma sequência absolutamente aleatória; mas se um único nucleotídeo estivesse presente, o resultado seria uma fita polimérica contendo esse único nucleotídeo. Assim, Nirenberg, em associação com seu colaborador Heinrich Matthaei, produziu inicialmente um mRNA composto inteiramente de uracila – poli-U.

A seguir, esses pesquisadores colocaram esse poli-U sobre o sistema de tradução livre de células. Eles então adicionaram aminoácidos radioativamente marcados sobre a mistura. Após testarem cada aminoácido – um de cada vez, em 20 experimentos diferentes –, eles determinaram que o poli-U induzia a síntese de um peptídeo que continha apenas fenilalanina (Figura 7-26). Visto que UUU é o único códon tripleto possível nessa mensagem, eles chegaram à conclusão que UUU codificava fenilalanina. A primeira palavra do código genético havia sido decifrada.

Nirenberg e Matthaei repetiram esse experimento com poli-A e poli-C e determinaram que AAA codificava lisina e CCC codificava prolina. O significado do poli-G não pode ser determinado por esse método, pois esse polinucleotídeo formava uma estranha hélice-fita tripla que não conseguia ser utilizada como molde no sistema livre de células.

Alimentar ribossomos com mRNAs sintéticos parecia uma técnica bastante interessante. No entanto, as possibilidades de uso de um único nucleotídeo estavam esgotadas, e os pesquisadores haviam definido três códons, restando ainda 61 possibilidades. As outras combinações de códons, no entanto, eram mais difíceis de serem construídas, e uma nova abordagem de síntese seria necessária. Na década de 1950, o químico orgânico Gobind Khorana havia desenvolvido métodos de preparo de misturas polinucleotídicas com sequência definida – mas essa técnica funcionava apenas com DNA. Quando ficou sabendo do trabalho de Nirenberg com mensagens sintéticas, Khorana redirecionou sua capacidade e esforço de trabalho para a produção de RNAs. Ele descobriu que se fizesse DNAs com uma sequência definida, poderia, a seguir, utilizar uma RNA-polimerase para



cadeia polinucleotídica pode formar uma associação relativamente forte com a sequência 5'-GAGC-3' presente em outra região da mesma molécula. A folha de trevo sofre outros dobramentos, originando uma estrutura compacta em forma de L que se mantém por pontes de hidrogênio adicionais entre as diferentes regiões da molécula (ver Figura 7-28B e C).

Duas regiões nucleotídicas não pareadas situadas cada uma em uma das extremidades da molécula estruturada em L são essenciais para o funcionamento do tRNA durante a síntese proteica. Uma dessas regiões forma o **anticódon**, um conjunto de três nucleotídeos consecutivos que sofre pareamento com o códon complementar sobre a molécula de um mRNA. A outra é uma região curta, de fita simples, que se situa na extremidade 3' da molécula; esse é o sítio onde o aminoácido que é codificado pelo códon se liga ao tRNA.

Vimos na seção anterior que o código genético é redundante; ou seja, vários códons diferentes podem determinar um mesmo aminoácido (ver Figura 7-24). Essa redundância implica que ou existe mais de um tRNA para muitos dos aminoácidos ou que algumas moléculas de tRNA podem ligar-se por complementaridade de bases a mais de um códon. Na verdade, ambas as situações ocorrem. Alguns aminoácidos possuem mais de um tRNA e alguns tRNAs são construídos de tal forma que eles exigem um pareamento de bases exato apenas às duas primeiras posições do códon, podendo tolerar um pareamento inexato (ou *oscilação*) sobre a terceira posição. Esse pareamento oscilante pode explicar por que tantos dentre os códons alternativos de um aminoácido diferem apenas em seus nucleotídeos da terceira posição (ver Figura 7-24). Os pareamentos oscilantes fazem com que seja possível adaptar os 20 aminoácidos a seus 61 códons com apenas 31 tipos diferentes de moléculas de tRNA. O número exato de diferentes tipos de tRNA, entretanto, difere entre as espécies. Por exemplo, humanos possuem quase 500 diferentes genes de tRNA, mas, entre esses, apenas 48 anticódons estão representados.

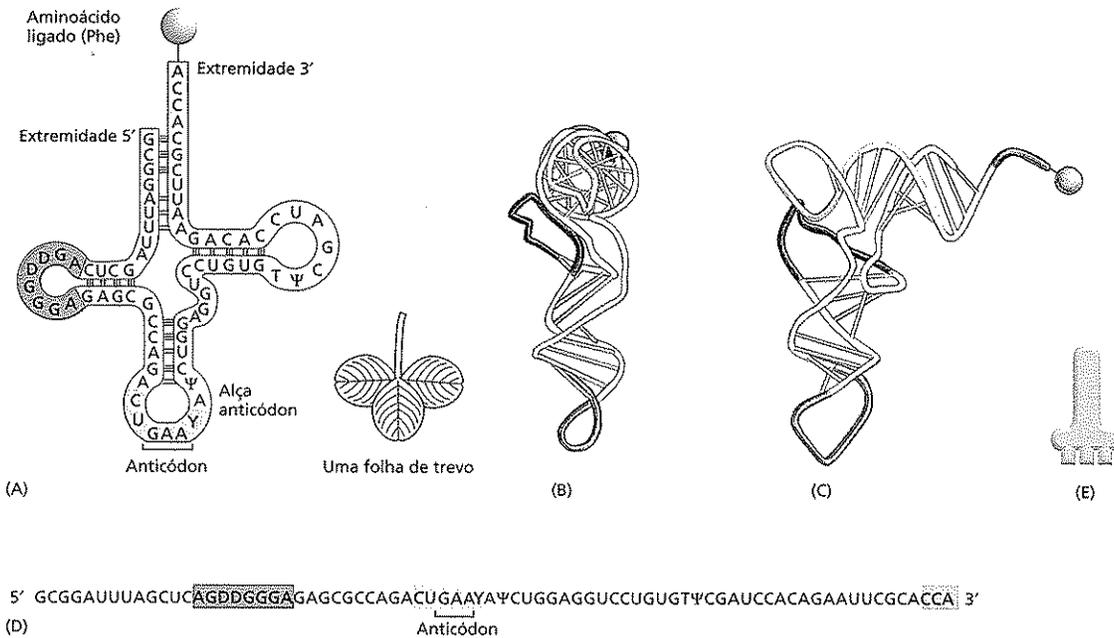
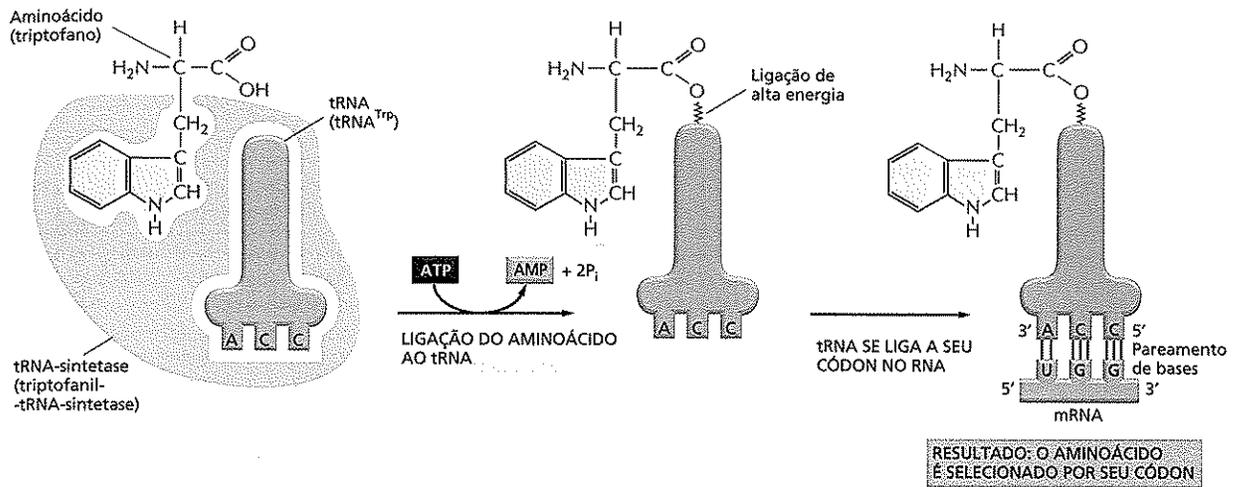


Figura 7-28 Moléculas de tRNA são adaptadores moleculares, que conectam os aminoácidos aos códons. Nessa série de diagramas, a mesma molécula de tRNA – neste caso, um tRNA específico para o aminoácido fenilalanina (Phe) – é ilustrada sob diferentes representações. (A) A estrutura em forma de folha de trevo, uma convenção utilizada para mostrar a complementaridade do pareamento de bases (linhas vermelhas) que cria as regiões de dupla-hélice na molécula. O anticódon (vermelho) é a sequência de três nucleotídeos que forma pares de bases com o códon no mRNA. O aminoácido correspondente ao par códon-anticódon está ligado à extremidade 3' do tRNA. Os tRNAs contêm algumas bases não usuais, as quais são produzidas por alterações químicas após a síntese do tRNA. As bases identificadas como  $\psi$  (de pseudouridina) e D (de di-hidrouridina) são derivadas da uracila. (B e C) Vistas da molécula real em forma de L, com base em análise de difração de raio X. Essas imagens apresentam um ângulo de rotação de 90°, uma em relação à outra. (D) A sequência nucleotídica linear da molécula, colorida de acordo com A, B e C. (E) Representação esquemática do tRNA que será usada nas próximas figuras e que enfatiza o anticódon.



### As enzimas específicas acoplam os tRNAs aos aminoácidos corretos

Com o objetivo de ler o código genético que se encontra sob a forma de DNA, as células produzem vários tRNAs diferentes. Agora, consideraremos como cada molécula de tRNA se torna *carregada* – ligada ao aminoácido que, dentre os 20 disponíveis, é seu companheiro mais adequado. O reconhecimento e a ligação do aminoácido correto é dependente de enzimas denominadas **aminoacil-tRNA sintetases**, que acoplam covalentemente cada aminoácido ao seu conjunto adequado de moléculas de tRNA. Na maioria dos organismos, existe uma enzima sintetase diferente para cada aminoácido (ou seja, existem no total 20 sintetases). Uma dessas acopla glicina a todos os tRNAs que reconhecem códons de glicina, outra acopla fenilalanina a todos os tRNAs que reconhecem códons de fenilalanina, e assim por diante. Nucleotídeos específicos, tanto no anticódon quanto no braço aceptor de aminoácidos, permitem que o tRNA correto seja reconhecido pela enzima sintetase (Animação 7.6). As sintetases são tão importantes quanto os tRNAs no processo de decodificação, pois é a ação combinada de sintetases e tRNAs que permite que cada códon presente sobre a molécula de mRNA promova a associação do aminoácido adequado (Figura 7-29).

A reação catalisada por sintetases que acopla o aminoácido à extremidade 3' do tRNA é uma das várias reações celulares acopladas à hidrólise de ATP liberadora de energia (ver Figura 3-30), produzindo uma ligação de alta energia entre o tRNA carregado e o aminoácido. A energia dessa ligação é usada em um estágio posterior da síntese proteica para ligar covalentemente o aminoácido à cadeia polipeptídica em crescimento.

### A mensagem do RNA é decodificada nos ribossomos

O reconhecimento de um códon pelo anticódon presente sobre uma molécula de tRNA depende do mesmo tipo de pareamento de bases por complementaridade usado na replicação do DNA e na transcrição. No entanto, a rápida e exata tradução do mRNA em proteínas requer uma maquinaria molecular grande que deslize sobre a cadeia de mRNA, capturando as moléculas de tRNA complementares, colocando-as em posição e ligando covalentemente os aminoácidos que elas carregam para que seja formada a cadeia proteica. Essa máquina produtora de proteínas é o **ribossomo** – um grande complexo composto por mais de 50 diferentes tipos de proteínas (as *proteínas ribossomais*) e diversas moléculas de RNA denominadas **RNAs ribossomais (rRNAs)**. Uma célula típica contém milhões de ribossomos em seu citoplasma (Figura 7-30).

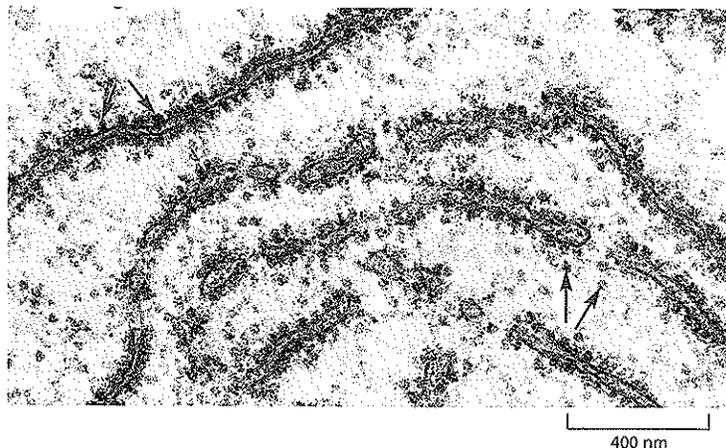
Os ribossomos eucarióticos e procarióticos são bastante similares, tanto em sua estrutura quanto em seu funcionamento. Ambos são compostos por uma subunidade grande e uma subunidade pequena que se encaixam para a formação do ribossomo completo, o qual possui uma massa de vários milhões

Figura 7-29 O código genético é traduzido por meio de dois adaptadores que atuam um após o outro. O primeiro adaptador é a aminoacil-tRNA-sintetase, que une um aminoácido específico ao seu tRNA correspondente; esse processo de acoplamento é denominado carregamento. O segundo adaptador é a própria molécula de tRNA, cujo anticódon forma pares de bases com o códon apropriado no mRNA. Um tRNA acoplado a seu aminoácido é também denominado tRNA carregado. Um erro em qualquer desses passos – no carregamento ou na ligação do tRNA carregado ao seu códon – provocará a incorporação de um aminoácido errado na cadeia de proteína. Na sequência de eventos ilustrada, o aminoácido triptofano (Trp) é selecionado pelo códon UGG no mRNA.

#### QUESTÃO 7-4

Em um inteligente experimento realizado em 1962, uma cisteína já associada ao seu tRNA foi quimicamente convertida em alanina. Essas moléculas tRNA "híbridas" foram adicionadas a um sistema de tradução livre de células do qual tRNAs-cisteína normais haviam sido removidos. Quando a proteína resultante foi analisada, determinou-se que havia sido inserida alanina em todos os pontos da cadeia proteica onde deveria existir uma cisteína. Discuta o que esse experimento nos revela sobre a função das aminoacil-tRNA-sintetases na tradução normal do código genético.

Figura 7-30 Ribossomos são encontrados no citoplasma de uma célula eucariótica. Essa microfotografia eletrônica mostra uma fina secção de uma pequena região do citoplasma. Os ribossomos aparecem como pontos pretos (setas vermelhas). Alguns estão livres no citosol; outros estão ligados a membranas do retículo endoplasmático. (Cortesia de George Palade.)



de dóltons (Figura 7-31); para comparação, uma proteína de tamanho médio possui uma massa igual a 30.000 dóltons. A subunidade pequena parece os tRNAs aos códons do mRNA, ao passo que a subunidade grande catalisa a formação das ligações peptídicas que unem os aminoácidos uns aos outros, formando a cadeia polipeptídica. As duas subunidades se associam sobre uma molécula de mRNA, geralmente próximo a seu início (a extremidade 5'), para que seja iniciada a síntese de uma proteína. Conforme o mRNA se move através do ribossomo, este traduz a sequência nucleotídica em uma sequência de aminoácidos, um códon por vez, usando os tRNAs como adaptadores. Dessa forma, cada aminoácido é adicionado, na ordem correta, à extremidade da cadeia polipeptídica em crescimento (Animação 7.7). Finalmente, as duas subunidades do ribossomo se

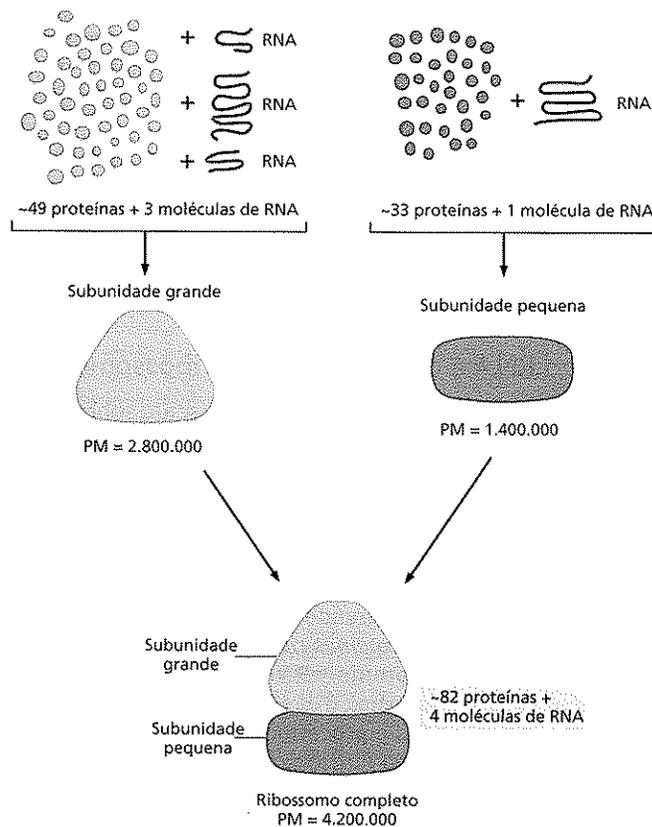


Figura 7-31 Um ribossomo é um grande complexo de quatro RNAs e mais de 80 proteínas. Estão ilustrados os componentes de ribossomos eucarióticos. Ribossomos procarióticos são bastante semelhantes. Apesar de as proteínas estarem em número muito superior aos RNAs ribossomais, os RNAs são responsáveis por mais de metade da massa do ribossomo.

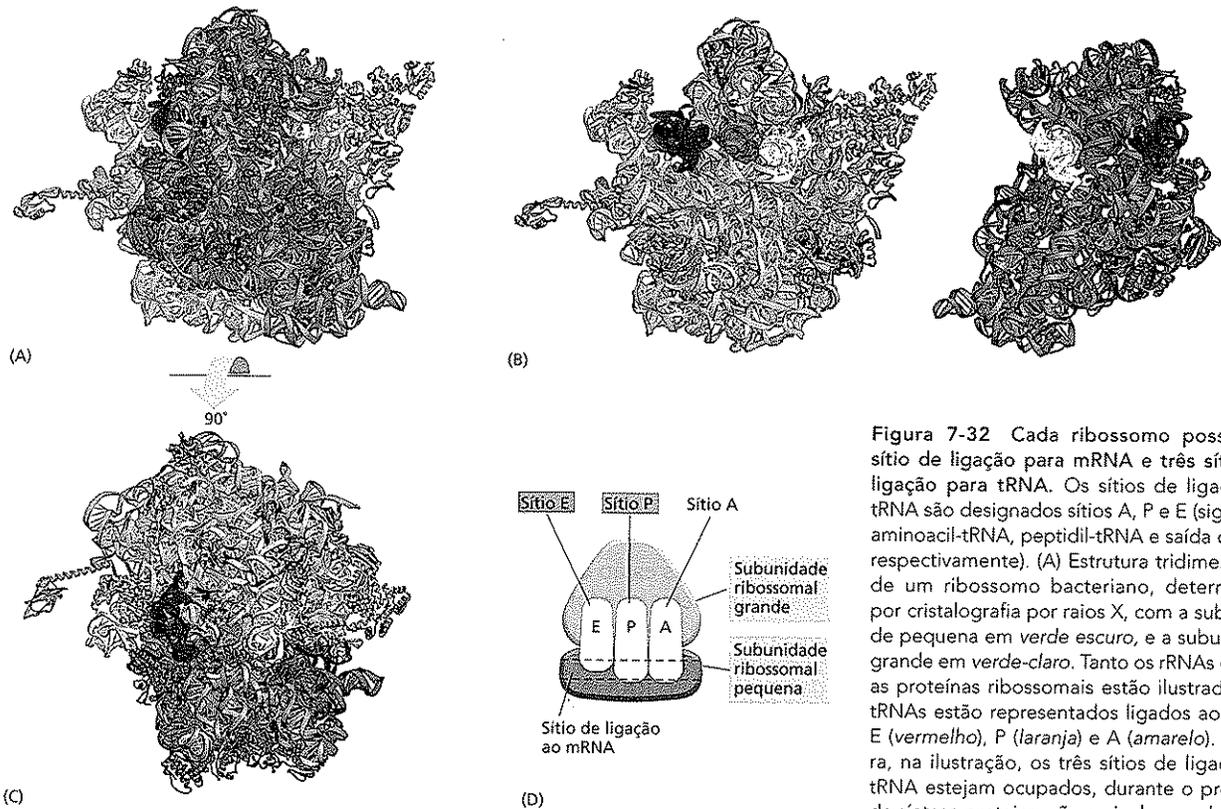


Figura 7-32 Cada ribossomo possui um sítio de ligação para mRNA e três sítios de ligação para tRNA. Os sítios de ligação ao tRNA são designados sítios A, P e E (sigla para aminoacil-tRNA, peptidil-tRNA e saída de exit, respectivamente). (A) Estrutura tridimensional de um ribossomo bacteriano, determinada por cristalografia por raios X, com a subunidade pequena em verde escuro, e a subunidade grande em verde-claro. Tanto os rRNAs quanto as proteínas ribossomais estão ilustrados. Os tRNAs estão representados ligados aos sítios E (vermelho), P (laranja) e A (amarelo). Embora, na ilustração, os três sítios de ligação de tRNA estejam ocupados, durante o processo de síntese proteica, não mais do que dois desses sítios contêm moléculas de tRNA simultaneamente (ver Figura 7-33). (B) A estrutura das subunidades ribossomais grande (esquerda) e pequena (direita), como provavelmente se organizam no ribossomo em (A), foi aberta como um livro. (C) Estrutura do ribossomo em (A) visto de cima. (D) Representação bastante esquemática de um ribossomo (na mesma orientação de C), que será utilizada nas figuras subsequentes. (A, B, e C, adaptada de M.M.Yusupov et al., *Science* 292: 883-896, 2001, com permissão de AAAS. Cortesia de Albion Bausom e Harry Noller.)

separam quando a síntese proteica está terminada. Os ribossomos operam por um sistema incrivelmente eficiente; em um intervalo de um segundo, um ribossomo de uma célula eucariótica adiciona aproximadamente dois aminoácidos à cadeia polipeptídica. Ribossomos bacterianos operam ainda mais rapidamente, sob uma taxa de cerca de 20 aminoácidos por segundo.

Como os ribossomos conseguem orquestrar todos os movimentos necessários para a tradução? Cada ribossomo contém um sítio de ligação para a molécula de mRNA e três sítios de ligação para moléculas de tRNA, denominados sítio A, sítio P e sítio E (Figura 7-32). Para adicionar um aminoácido à cadeia polipeptídica em crescimento, o tRNA adequadamente carregado penetra no sítio A por pareamento de bases com o códon complementar que se encontra na molécula de mRNA. Seu aminoácido é então ligado à cadeia polipeptídica que está associada ao tRNA presente no sítio P vizinho a ele. A seguir, o ribossomo se movimenta, e o tRNA vazio é posicionado sobre o sítio E antes de ser ejetado (Figura 7-33). Esse ciclo de reações é repetido cada vez que um aminoácido é adicionado à cadeia polipeptídica, fazendo com que essa cadeia cresça de sua extremidade amino rumo à sua extremidade carboxila até o momento em que um códon de terminação seja encontrado.

### O ribossomo é uma ribozima

O ribossomo é uma das maiores e mais complexas estruturas da célula, sendo composto por dois terços de RNA e um terço de proteína. A determinação, no ano 2000, da estrutura tridimensional completa de suas subunidades grande e pequena foi um dos grandes triunfos da biologia moderna. Essa determinação confirmou fortemente evidências anteriores de que os rRNAs – e não proteínas – eram responsáveis pela estrutura geral do ribossomo e por sua capacidade de orquestrar a síntese proteica.

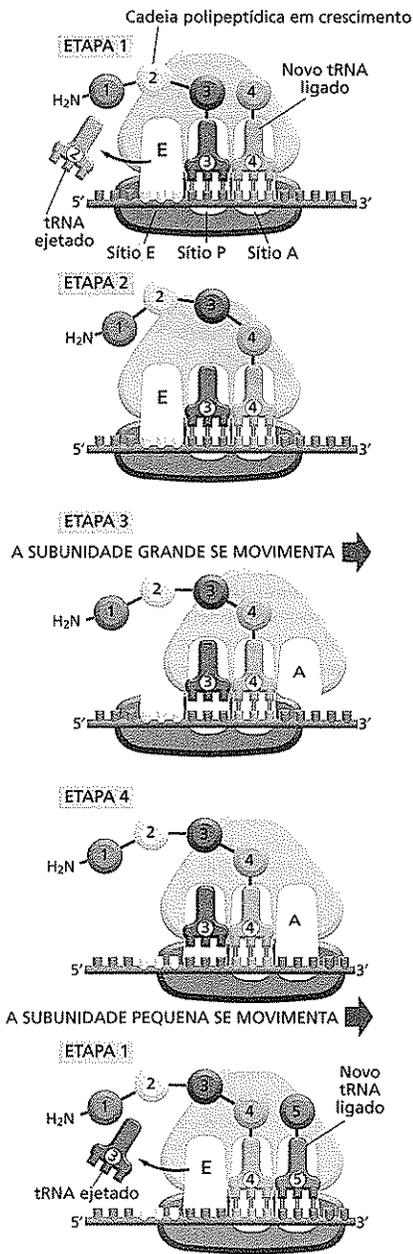


Figura 7-33 A tradução ocorre em um ciclo de quatro etapas. Este ciclo é repetido muitas e muitas vezes durante a síntese de uma cadeia proteica. Na etapa 1, um tRNA que transporta o próximo aminoácido da cadeia se liga ao sítio A livre no ribossomo pela formação de pares de bases com o códon que está ali exposto. Visto que apenas um dos diversos tipos de moléculas tRNA na célula pode formar pares de bases com cada códon, esse códon determina o aminoácido específico que será adicionado à cadeia polipeptídica em crescimento. Os sítios A e P estão suficientemente próximos para que as duas moléculas de tRNA ali fixadas sejam forçadas a formar pares de bases com códon contíguos, sem que fique qualquer base entre eles. Esse posicionamento dos tRNAs assegura que a fase de leitura correta seja mantida ao longo de toda a síntese da proteína. Na etapa 2, a extremidade carboxila da cadeia polipeptídica é separada do tRNA presente no sítio P e unida por ligações peptídicas ao grupo amino livre do aminoácido que se encontra ligado ao tRNA presente no sítio A. Essa reação é catalisada por um sítio enzimático na subunidade grande. Na etapa 3, um movimento da subunidade grande em relação à subunidade pequena faz com que os dois tRNAs sejam posicionados nos sítios E e P da subunidade grande. Na etapa 4, a subunidade pequena se move exatamente três nucleotídeos sobre a molécula de mRNA, o que a reposiciona novamente sobre sua posição original em relação à subunidade grande. Esse movimento reinicializa o ribossomo que apresenta um sítio A vazio de tal forma que o próximo aminoácil-tRNA pode ligar-se (Animação 7.8). Como indicado, o mRNA é traduzido no sentido 5' - 3', e a extremidade N-terminal de uma proteína é sintetizada primeiro, cada ciclo adiciona um novo aminoácido à extremidade C-terminal da cadeia polipeptídica. Para visualizar o ciclo da tradução em detalhes, ver Animação 7.9.

Os rRNAs estão dobrados em estruturas tridimensionais altamente precisas e compactas que formam o cerne do ribossomo (Figura 7-34). Em marcante contraste ao posicionamento central do rRNA, as proteínas ribossomais estão geralmente localizadas na superfície, onde preenchem as fendas e frestas do dobramento do RNA. A principal função das proteínas ribossomais parece ser a manutenção da estrutura e da estabilidade do núcleo de RNA, permitindo ainda que aconteçam as alterações na conformação do rRNA que são necessárias para que esse RNA catalise eficientemente a síntese proteica.

Não apenas os três sítios de ligação aos tRNAs (os sítios A, P e E) do ribossomo são formados principalmente por rRNA, mas também o sítio catalítico para a formação da ligação peptídica é formado pelo RNA 23S da subunidade grande. O aminoácido mais próximo está localizado a uma tal distância que impossibilita que esse estabeleça contatos, seja com o aminoácil-tRNA que chega, seja com a cadeia polipeptídica em crescimento. O sítio catalítico nessa peptidil-transferase com base em RNA é semelhante em diversos aspectos àquele encontrado em algumas enzimas proteicas: é uma fenda altamente estruturada que orienta precisamente os dois reagentes, o peptídeo que está sendo estendido e o tRNA carregado, aumentando, assim, as chances de ocorrência de uma reação produtiva.

Moléculas de RNA que possuem atividade catalítica são denominadas **ribozimas**. Na seção final deste capítulo, consideraremos outras ribozimas e discutiremos o que a catálise com base em RNA deve ter significado para a evolução inicial da vida na Terra. No momento, devemos apenas salientar que existe uma boa razão para suspeitar que RNAs, em vez de moléculas proteicas, atuaram como os primeiros catalisadores em células vivas. Se isso é verdade, o ribossomo, com seu núcleo de RNA, pode ser considerado como uma relíquia do período inicial da história da vida, quando as células eram conduzidas quase que exclusivamente por ribozimas.

### Os códon do mRNA sinalizam onde a síntese proteica deve iniciar e terminar

Apesar de ser possível forçar os ribossomos a traduzir qualquer RNA em tubos de ensaio (ver Como Sabemos, p. 248-249), nas células, a tradução requer um sinal de início específico. O ponto sobre o qual a síntese proteica tem início no mRNA é essencial, pois ele determina a fase de leitura que será seguida em toda a extensão da mensagem. Um erro de um nucleotídeo em qualquer dos sentidos, nessa etapa, fará com que cada códon subsequente da mensagem seja erroneamente lido, de tal forma que será sintetizada uma proteína não funcional, composta por uma sequência equivocada de aminoácidos (ver Figura 7-27).

#### QUESTÃO 7-5

A seguinte sequência de nucleotídeos de uma fita de DNA foi usada como molde para a síntese de um mRNA que foi então traduzido em proteína: 5'-TTAACGGCTTTTTTC-3'. Determine o aminoácido C-terminal e o aminoácido N-terminal do polipeptídeo resultante. Assuma que o mRNA é traduzido sem a necessidade de um códon de iniciação.

pecto  
ser ti  
sinte

é nec  
o am  
tion  
poss  
a ext  
geral  
inici

ment  
nas i  
Dent  
é cap  
quer  
de ui  
queç  
pequ  
do p  
inici  
asso  
ribo:  
prot  
amir

rias.

Figur  
inicia  
de tr  
queç  
bas a  
a prc

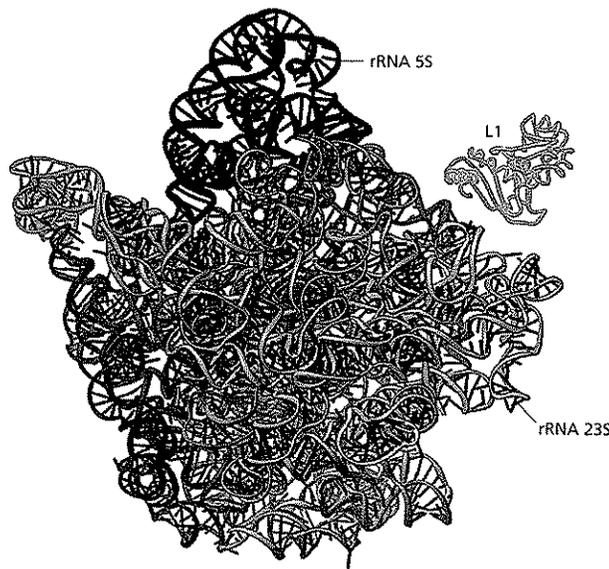


Figura 7-34 Os RNAs ribossomais conferem ao ribossomo sua forma geral. São apresentadas as estruturas detalhadas de dois rRNAs, o rRNA 23S (azul) e o rRNA 5S (púrpura), que formam o núcleo da subunidade grande dos ribossomos. Uma das subunidades proteicas do ribossomo (L1) está incluída como ponto de referência, visto que essa proteína é responsável por uma protuberância característica na superfície ribossomal. Os componentes ribossomais são geralmente designados por seu "valor S", o qual se refere à taxa de sedimentação em ultracentrifugação. (Adaptada de N. Ban et al., *Science* 289:905-920, 2000. Com permissão de AAAS.)

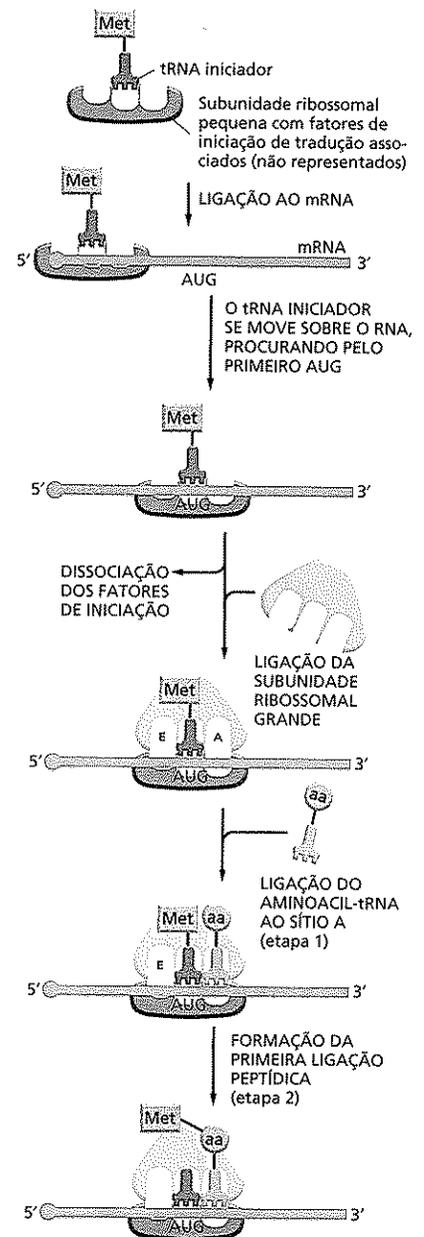
A etapa de iniciação é também extremamente importante sob outro aspecto: esse é o último momento no qual a célula pode decidir se um mRNA deve ser traduzido; assim, a taxa de iniciação determina a taxa na qual a proteína é sintetizada a partir do RNA.

A tradução de um mRNA tem início com o códon AUG, e um tRNA especial é necessário para a iniciação da tradução. Esse **tRNA iniciador** sempre carrega o aminoácido metionina (ou uma forma modificada de metionina, a formilmetionina, em bactérias) de tal forma que todas as proteínas recém-sintetizadas possuem metionina como seu primeiro aminoácido, na extremidade N-terminal, a extremidade proteica que é primeiramente sintetizada. Essa metionina é em geral removida posteriormente pela ação de uma protease específica. O tRNA iniciador é distinto do tRNAs que normalmente transporta metionina.

Em eucariotos, o tRNAs iniciador que está acoplado à metionina é inicialmente posicionado sobre a subunidade ribossomal pequena junto com proteínas adicionais denominadas **fatores de iniciação de tradução** (Figura 7-35). Dentre todos os tRNAs carregados da célula, apenas o tRNA iniciador carregado é capaz de se associar fortemente ao sítio P de uma subunidade ribossomal pequena. A seguir, essa subunidade ribossomal carregada se liga à extremidade 5' de uma molécula de mRNA, a qual é em parte reconhecida pela presença de seu quepe típico de mRNAs eucarióticos (ver Figura 7-16). A subunidade ribossomal pequena então se move para a frente (de 5' para 3') sobre o mRNA, à procura do primeiro códon AUG. Quando esse AUG é encontrado, diversos fatores de iniciação se dissociam da subunidade ribossomal pequena, dando espaço para a associação da subunidade ribossomal grande e consequente montagem de um ribossomo completo. Visto que o tRNA iniciador está ligado ao sítio P, a síntese proteica está pronta para ter início pela adição do próximo tRNA acoplado a seu aminoácido sobre o sítio A (ver Figura 7-35).

O mecanismo de seleção de um códon de iniciação é diferente em bactérias. Os mRNAs bacterianos não possuem o quepe 5' para indicar ao ribossomo

Figura 7-35 A fase de iniciação da síntese de proteínas em eucariotos requer fatores de iniciação e um tRNA iniciador especial. Apesar de não estarem aqui ilustradas, uma iniciação de tradução eficiente também requer proteínas adicionais (ilustradas na Figura 7-22) ligadas ao quepe 5' e à cauda poli-A do mRNA. Dessa maneira, o aparato de tradução se certifica que ambas as extremidades do mRNA estejam intactas antes da iniciação da tradução. Após a iniciação, a proteína é sintetizada pelas reações descritas na Figura 7-33.



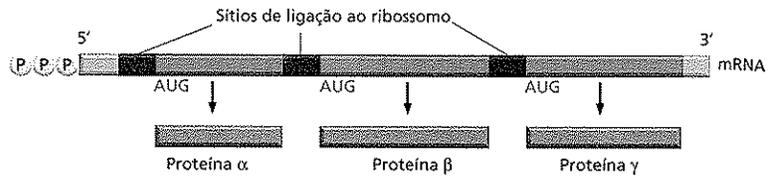
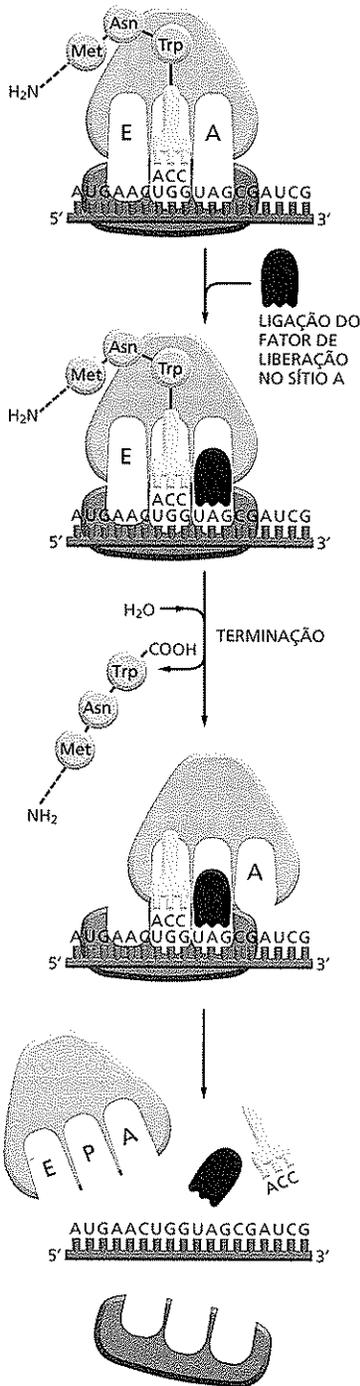


Figura 7-36 Uma única molécula de mRNA procariótico pode codificar várias proteínas diferentes. Em procariotos, genes envolvidos em diferentes passos de um mesmo processo se encontram frequentemente organizados em grupos (óperons) que são transcritos juntos sob a forma de um único mRNA. Diferentemente dos ribossomos eucarióticos, os quais reconhecem um quepe 5', os ribossomos procarióticos iniciam a transcrição em sítios de ligação ao ribossomo (vermelho), os quais podem estar localizados no interior de uma molécula de mRNA. Essa característica permite que procariotos sintetizem várias proteínas distintas a partir de uma única molécula de mRNA.



onde iniciar a busca pelo ponto de início da tradução. Em vez disso, eles contêm sequências específicas de ligação a ribossomos, com comprimento de até seis nucleotídeos que estão localizadas poucos nucleotídeos à montante dos AUGs sobre os quais a tradução deve ter início. Diferentemente de um ribossomo eucariótico, um ribossomo procariótico pode, com facilidade, ligar-se diretamente a um códon de iniciação que esteja localizado no interior de uma molécula de mRNA. Tais sequências de ligação ao ribossomo são necessárias em bactérias, pois os mRNAs procarióticos são frequentemente *policistrônicos* – ou seja, eles codificam várias proteínas diferentes, todas elas sendo traduzidas a partir da mesma molécula de mRNA (Figura 7-36). Em contraste, um mRNA eucariótico geralmente transporta informação referente a uma única proteína.

O fim de uma mensagem codificadora de proteína, tanto em procariotos quanto em eucariotos, é sinalizado pela presença de um dos diversos códonos chamados de *códonos de terminação* (ver Figura 7-26). Esses códonos especiais (UAA, UAG e UGA) não são reconhecidos por um tRNA e não determinam qualquer aminoácido, sinalizando para o ribossomo o término da tradução. Proteínas conhecidas como *fatores de liberação* se ligam a qualquer códon de terminação que chegue a um sítio A do ribossomo, e essa ligação altera a atividade peptidil-transferase no ribossomo, fazendo com que ela catalise a adição de uma molécula de água, em vez de um aminoácido ao peptidil-tRNA (Figura 7-37). Essa reação libera a extremidade carboxila da cadeia polipeptídica de sua conexão à molécula de tRNA, e, considerando-se que esse era o único elo de ligação que mantinha o polipeptídeo em crescimento associado ao ribossomo, a cadeia proteica completa é imediatamente liberada no citosol. O ribossomo libera o mRNA e se dissocia em suas duas subunidades, as quais podem associar-se novamente sobre outra molécula de mRNA para dar início a um novo ciclo de síntese proteica.

Vimos, no Capítulo 4, que muitas proteínas podem dobrar-se espontaneamente adquirindo uma estrutura tridimensional definida e que algumas realizam esse dobramento conforme estão sendo sintetizadas no ribossomo. No entanto, a maior parte das proteínas requer o auxílio de *chaperonas moleculares* para executar o dobramento adequado na célula. As chaperonas mais bem estudadas usam ciclos sucessivos de hidrólise de ATP para ligar e liberar continuamente proteínas recém-sintetizadas até que essas estejam adequadamente dobradas. Essa manipulação direciona as proteínas rumo vias de dobramento produtivo, evitando que elas formem agregados ou que adquiram conformações inadequadamente dobradas. As proteínas recém-sintetizadas são frequentemente interceptadas por suas chaperonas conforme emergem do ribossomo.

Figura 7-37 Na fase final da síntese proteica, a ligação de um fator de liberação no sítio A sobre um códon de terminação finaliza a tradução. O polipeptídeo completo é liberado, e o ribossomo se dissocia liberando suas duas subunidades.

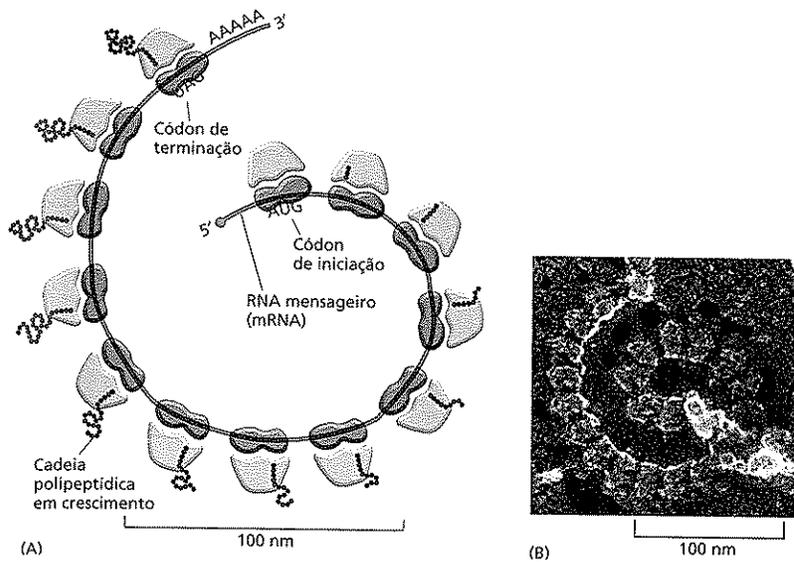


Figura 7-38 As proteínas são traduzidas por polirribossomos. (A) Desenho esquemático mostrando como uma série de ribossomos pode traduzir simultaneamente a mesma molécula de mRNA (Animação 7.10). (B) Microfotografia eletrônica de um polirribossomo de uma célula eucariótica. (B, cortesia de John Heuser.)

### As proteínas são produzidas em polirribossomos

A síntese da maior parte das moléculas proteicas leva entre 20 segundos e alguns minutos. No entanto, mesmo durante esse curto intervalo de tempo, geralmente ocorrem múltiplas iniciações sobre cada molécula de mRNA que está sendo traduzida. Se o mRNA está sendo traduzido eficientemente, um novo ribossomo é montado sobre a extremidade 5' de uma molécula de mRNA quase que imediatamente após o ribossomo precedente ter traduzido uma sequência nucleotídica longa o suficiente para não mais o atrapalhar. Consequentemente, as moléculas de mRNA que estão sendo traduzidas são, geralmente, encontradas sob a forma de *polirribossomos* (também denominados *polissomos*), grandes agregados citoplasmáticos compostos por vários ribossomos espaçados por segmentos tão curtos quanto 80 nucleotídeos, sobre uma única molécula de mRNA (Figura 7-38). Essas iniciações múltiplas significam que muito mais moléculas de proteína podem ser feitas em um dado período do que seria possível se cada molécula tivesse de ser concluída antes que a próxima fosse iniciada.

Tanto bactérias quanto eucariotos utilizam os polissomos, mas as bactérias podem tornar a síntese proteica ainda mais rápida. Visto que o mRNA bacteriano não precisa ser processado e também se encontra fisicamente acessível aos ribossomos mesmo durante a síntese, os ribossomos se ligarão à extremidade livre de uma molécula de mRNA bacteriano e começarão a traduzi-la antes mesmo do término da transcrição do RNA. Esses ribossomos seguem o encaixe da RNA-polimerase, conforme essa se move sobre o DNA.

### Os inibidores da síntese proteica de procariotos são utilizados como antibióticos

A capacidade de traduzir eficientemente mRNA em proteínas é uma característica fundamental de toda a vida na Terra. Apesar de os ribossomos e outras moléculas que executam essa fantástica tarefa serem muito similares quando comparados a diferentes organismos, existem algumas diferenças sutis, como vimos, no mecanismo utilizado para a síntese de proteínas por bactérias e eucariotos. Apesar de uma peculiaridade evolutiva, essas diferenças formam a base de um dos mais importantes avanços da medicina moderna.

Vários dos mais efetivos antibióticos são compostos que atuam pela inibição da síntese proteica bacteriana, mas não afetam a síntese proteica eucariótica. Alguns desses fármacos exploram pequenas diferenças estruturais e funcionais existentes entre os ribossomos bacterianos e eucarióticos, de tal forma que in-

**TABELA 7-3** Antibióticos que inibem a síntese proteica ou de RNA

Antibiótico	Efeito específico
Tetraciclina	Bloqueia a ligação da aminoacil-tRNA ao sítio A do ribossomo (etapa 1 na Figura 7-33)
Streptomicina	Evita a transição do complexo de iniciação para um ribossomo capaz de extensão de cadeia (ver Figura 7-35); também pode causar erros de decodificação
Cloranfenicol	Bloqueia a reação de peptidil-transferase nos ribossomos (etapa 2 na Figura 7-33)
Cicloexamida	Bloqueia a reação de translocação nos ribossomos (etapa 3 na Figura 7-33)
Rifamicina	Bloqueia a iniciação das cadeias de RNA por meio de ligação à RNA-polimerase

terfere preferencialmente na síntese proteica bacteriana. Assim, esses compostos podem ser administrados em doses relativamente altas sem que apresentem toxicidade em seres humanos. Visto que diferentes antibióticos se ligam a diferentes regiões sobre o ribossomo bacteriano, essas medicações frequentemente inibem passos diferentes do processo de síntese proteica. Alguns dentre os mais comuns antibióticos dessa classe estão listados na Tabela 7-3.

Diversos antibióticos comumente utilizados foram inicialmente isolados de fungos. Fungos e bactérias frequentemente ocupam os mesmos nichos ecológicos; para ter uma vantagem competitiva, os fungos desenvolveram, ao longo do tempo, toxinas potentes que matam bactérias, mas são inócuas a eles próprios. Visto que fungos e seres humanos são eucariotos e conseqüentemente são mais intimamente relacionados entre si do que qualquer um deles o é em relação às bactérias (ver Figura 1-29), podemos pegar emprestado esses compostos para combater nossos próprios inimigos bacterianos.

### Uma degradação proteica cuidadosamente controlada ajuda a regular a quantidade de cada proteína na célula

Após uma proteína ser liberada do ribossomo, ela fica sujeita a uma série de controles efetuados pela célula. O número de cópias de uma proteína em uma célula depende, assim como ocorre em uma população humana, não apenas de quão rápido novos indivíduos podem ser gerados, mas também de quanto tempo eles sobreviverão. Assim, a degradação de proteínas em seus aminoácidos constituintes, mediada pela célula, é uma forma de regular a quantidade de uma dada proteína presente em um momento determinado. As proteínas diferem enormemente em relação à sua duração, ou tempo de vida. Proteínas estruturais que fazem parte de tecidos relativamente permanentes, como ossos ou músculos, podem durar meses ou mesmo anos, ao passo que outras proteínas, como enzimas metabólicas e proteínas que regulam o ciclo de crescimento, a mitose e a divisão celular (discutidos no Capítulo 18), duram apenas alguns dias, horas ou mesmo segundos. Como a célula controla essa duração?

As células possuem vias especializadas para a quebra, ou digestão, enzimática de proteínas em seus aminoácidos constituintes (um processo denominado *proteólise*). As enzimas que degradam proteínas, inicialmente em peptídeos pequenos e finalmente nos aminoácidos individuais, são coletivamente denominadas **proteases**. As proteases atuam pela clivagem (hidrólise) das ligações peptídicas entre os aminoácidos (ver Painel 2-5, p. 72-73). Uma das funções das vias proteolíticas é a rápida degradação das proteínas que devem ter curta duração. Outra é o reconhecimento e a eliminação de proteínas que estejam danificadas ou erroneamente dobradas. A eliminação de proteínas com conformações errôneas é essencial para um organismo, como pode ser observado pelo fato de doenças neurodegenerativas, como as doenças de

Hunti  
prote:  
car sé  
a mai  
de de  
res et  
eucar  
nadas  
por p  
câma:  
protei:  
cas (F  
digesti  
dade  
protei:  
libera  
câma:  
correi

das no  
sobre  
uma p  
as pr  
essas  
las pr

cia cu  
conse  
ou er:  
dados  
degra  
que a  
expos  
co - p  
estão

Exist

Vimos  
são ne  
em ur  
la, coi  
etapa:

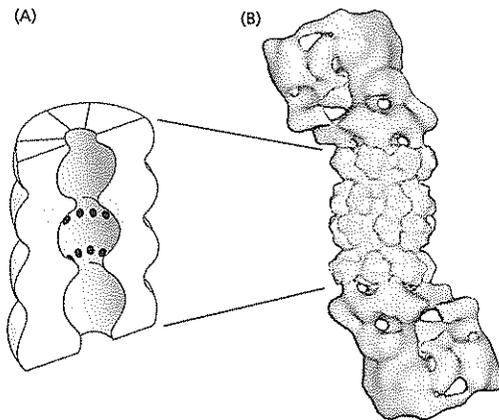


Figura 7-39 O proteassomo degrada proteínas de curta duração ou proteínas dobradas erroneamente. (A) Uma visão em corte da estrutura do cilindro central do proteassomo, determinada a partir de cristalografia por raios X, com os sítios ativos das proteases indicados por pontos vermelhos. (B) A estrutura completa do proteassomo, na qual o cilindro central (amarelo) é acrescido de um quepe (púrpura) em cada extremidade. (B, de W. Baumeister et al., *Cell* 92:367-380, 1998. Com permissão de Elsevier.)

Huntington, Alzheimer e Creutzfeldt-Jacob serem causadas pela agregação de proteínas dobradas erroneamente. Esses agregados proteicos podem provocar sérios danos a células e tecidos e mesmo levar à morte celular. Apesar de a maioria das proteínas danificadas ser digerida no citosol, importantes vias de degradação proteica também operam em outros compartimentos celulares eucarióticos, como os lisossomos (discutidos no Capítulo 15). Em células eucarióticas, a maior parte das proteínas é degradada por estruturas denominadas **proteassomos**. Um proteassomo contém um cilindro central formado por proteases cujos sítios ativos estão dirigidos para a face interna de uma câmara. Cada extremidade desse cilindro é tampada por um grande complexo proteico formado por pelo menos 10 diferentes tipos de subunidades proteicas (Figura 7-39). Essas tampas proteicas se ligam às proteínas destinadas à digestão e então – usando a hidrólise de ATP como combustível para sua atividade – as inserem no interior da câmara do cilindro. Nesse compartimento, as proteases cortam as proteínas em pequenos peptídeos, os quais são a seguir liberados pela extremidade do proteassomo. Manter as proteases dentro dessa câmara de destruição parece fazer sentido, pois evita que elas fiquem livres, correndo desenfreadamente pela célula.

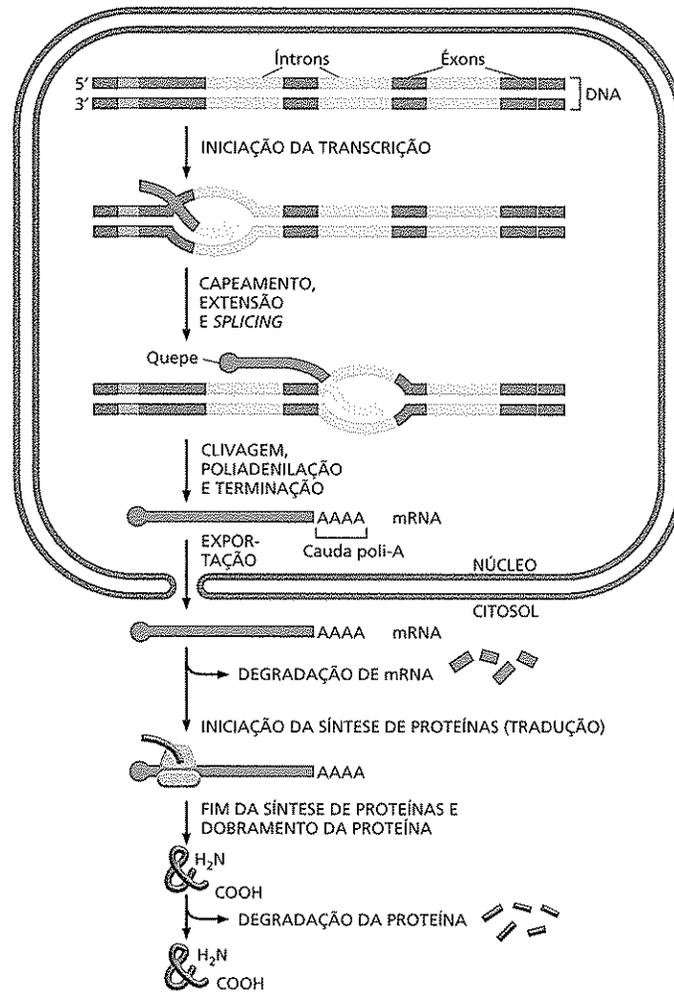
Como o proteassomo seleciona quais proteínas da célula devem ser inseridas no cilindro para sofrer degradação? Os proteassomos atuam principalmente sobre proteínas que foram marcadas para destruição pela ligação covalente a uma pequena proteína denominada *ubiquitina*. Enzimas especializadas marcam as proteínas selecionadas com uma ou mais moléculas de ubiquitina; a seguir, essas proteínas ubiquitinadas são reconhecidas e inseridas no proteassomo pelas proteínas da tampa.

Proteínas que devem ter vida curta frequentemente ostentam uma sequência curta de aminoácidos que as identificam como devendo ser ubiquitinadas e, conseqüentemente, degradadas pelo proteassomo. Proteínas desnaturadas ou erroneamente dobradas, bem como proteínas que contêm aminoácidos oxidados ou alterados de qualquer outro modo, são também reconhecidas e degradadas por esse sistema proteolítico dependente de ubiquitina. As enzimas que adicionam ubiquitina a essas proteínas reconhecem sinais que se tornam expostos como resultado do dobramento inadequado ou devido a dano químico – por exemplo, sequências de aminoácidos ou motivos conformacionais que estão normalmente escondidos e inacessíveis.

### Existem várias etapas entre DNA e proteína

Vimos até o momento, neste capítulo, que diferentes tipos de reações químicas são necessários para a produção de uma proteína a partir da informação contida em um gene (Figura 7-40). A concentração final de uma proteína em uma célula, conseqüentemente, depende da eficiência na qual cada uma dessas diversas etapas é realizada. Além disso, diversas proteínas – depois de serem liberadas

Figura 7-40 A produção de uma proteína em uma célula eucariótica requer diversas etapas. A concentração final de cada proteína depende da eficiência de cada etapa indicada. Mesmo após uma proteína ter sido traduzida, sua concentração pode ser regulada por degradação, e sua atividade pode ser regulada por meio de ligação a pequenas moléculas, ou via outras alterações pós-traducionais (ver Figura 7-41).



no ribossomo – necessitam de outras alterações para que possam ser úteis às células. Por exemplo, modificações covalentes (como a fosforilação), a ligação a pequenas moléculas cofatores ou a associação a outras subunidades proteicas são frequentemente necessárias para a ativação de proteínas recém-sintetizadas (Figura 7-41).

Veremos, no próximo capítulo, que as células possuem a capacidade de alterar os níveis da maior parte de suas proteínas de acordo com suas necessidades. Em princípio, todas as etapas da Figura 7-40 podem ser reguladas pela célula. No entanto, a iniciação da transcrição é o ponto mais comum usado pelas células para regular a expressão de cada um de seus genes.

A transcrição e a tradução são processos universais que se localizam na base da vida. No entanto, quando os cientistas começaram a considerar como o fluxo de informações do DNA para proteína deve ter-se originado, surgiram algumas conclusões inesperadas.

Figura 7-41 Diversas proteínas necessitam de modificações adicionais para tornarem-se plenamente funcionais. Para ser útil à célula, um polipeptídeo completo deve estar adequadamente dobrado sob sua conformação tridimensional, ligado aos cofatores necessários e agrupado a suas proteínas acessórias. A formação de ligações covalentes conduz essas alterações. Diversas proteínas também necessitam de modificações covalentes para tornarem-se ativas. Apesar de a fosforilação e a glicosilação serem as alterações mais comuns, mais de 100 diferentes tipos de modificações covalentes são conhecidas para as proteínas.

## RNA E A ORIGEM DA VIDA

Para compreender plenamente os processos que ocorrem nas células atuais, devemos considerar como as células evoluíram. No entanto, um processo essencial – a expressão gênica – nos confronta a um intrigante desafio: se são necessários ácidos nucleicos para direcionar a síntese de proteínas e se proteínas são necessárias para a síntese de ácidos nucleicos, como esse sistema de componentes interdependentes pode ter-se originado? Uma das hipóteses é a que considera a existência de um *mundo de RNA* na Terra antes do surgimento das células modernas (Figura 7-42). De acordo com essa hipótese, o RNA – que hoje serve de intermediário entre os genes e as proteínas – tanto estocava informação genética quanto catalisava reações químicas nas células primitivas. Apenas tardiamente, em termos evolutivos, o DNA suplantou o RNA como material genético, e as proteínas se tornaram os principais componentes catalisadores e estruturais das células. Se essa ideia está correta, então, a transição do mundo de RNA nunca foi completa; como vimos, neste capítulo, o RNA ainda catalisa várias reações fundamentais nas células atuais. Esses RNA catalisadores, inclusive o ribossomo e a maquinaria de *splicing* de RNA, podem então ser considerados como fósseis moleculares de um mundo anterior.

### A vida requer autocatálise

A origem da vida necessitou de moléculas que possuíssem, pelo menos em um determinado nível, uma propriedade essencial: a capacidade de catalisar reações que levassem – direta ou indiretamente – à produção de mais moléculas idênticas a elas. Catalisadores com essa propriedade especial de autopromoção, a partir do surgimento ao acaso, poderiam se reproduzir e, desse modo, gerar matéria-prima para a produção de outras substâncias. Dessa maneira, podemos imaginar o desenvolvimento gradual de um sistema químico de crescente complexidade composto de monômeros e polímeros orgânicos que funcionariam em conjunto para a geração de mais moléculas semelhantes, abastecido por um suplemento de matérias-primas simples presentes no ambiente. Um sistema *autocatalítico* dessa natureza deveria ter muitas das propriedades que consideramos como características de matéria viva: o sistema deve conter uma seleção não aleatória de moléculas interativas; ele deve tender à própria reprodução; deve competir com outros sistemas dependentes das mesmas matérias-primas, e, se privado de suas matérias-primas ou mantido a uma temperatura que provoque um distúrbio no balanço das taxas de reação, deve decair rumo a um equilíbrio químico e “morrer”.

Que moléculas podem ter apresentado tais propriedades autocatalíticas? Nas células vivas atuais, os catalisadores mais versáteis são as proteínas – capazes de adotar diferentes formas tridimensionais que atuam como sítios quimicamente reativos. No entanto, não é conhecido qualquer meio pelo qual uma proteína possa reproduzir a si própria diretamente. Ácidos nucleicos, no entanto, podem desempenhar todas essas tarefas.

### O RNA pode tanto estocar informação como catalisar reações químicas

O sistema de complementaridade de pares de bases permite que um ácido nucleico atue como molde para a formação de outro. Assim, uma fita simples de RNA ou DNA pode determinar a sequência de um polinucleotídeo complementar,

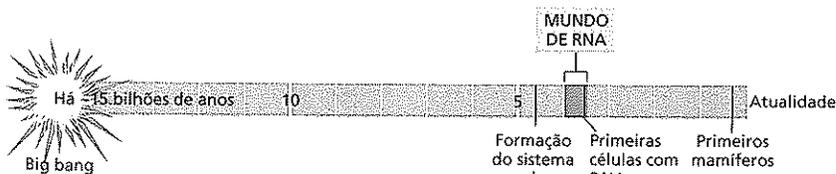
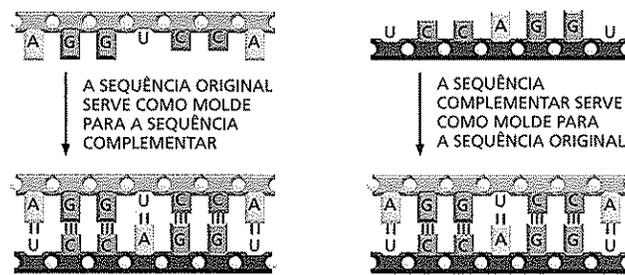


Figura 7-42 Um mundo de RNA pode ter existido antes do surgimento das células modernas.

Figura 7-43 Uma molécula de RNA pode, em princípio, direcionar a formação de uma cópia exatamente igual a ela. Na primeira etapa, a molécula original de RNA atua como molde para formar uma molécula de RNA de sequência complementar. Na segunda etapa, essa molécula de RNA complementar atua, por sua vez, como molde, dando origem a moléculas de RNA com a sequência original. Visto que cada molécula-molde pode produzir diversas cópias da fita complementar, essas reações podem resultar em "multiplicação" da sequência original.



o qual, por sua vez, pode determinar a sequência da molécula original, permitindo que o ácido nucleico original seja replicado (Figura 7-43). Esses mecanismos de molde por complementaridade formam a base da replicação do DNA e da transcrição nas células atuais.

No entanto, a síntese eficiente de polinucleotídeos por meio de tais mecanismos de molde por complementaridade também necessita de catalisadores que promovam a reação de polimerização: sem catalisadores, a formação do polímero é lenta, sujeita a erros e ineficiente. Atualmente, a polimerização de nucleotídeos é rapidamente catalisada por enzimas proteicas – como a DNA e a RNA-polimerases. No entanto, como poderiam ter sido catalisadas essas reações antes da existência de proteínas que contivessem a especificidade catalítica adequada? O começo de uma resposta para essa questão foi obtido em 1982, quando foi descoberto que moléculas de RNA *per se* podem atuar como catalisadoras. Vimos, por exemplo, que uma molécula de RNA catalisa a reação peptidil-transferase que ocorre no ribossomo. Acredita-se que o potencial sem igual das moléculas de RNA, que as torna capazes de atuar tanto como carreadoras de informação quanto como catalisadoras, tenha permitido que elas desempenhassem um papel central na origem da vida.

Nas células atuais, o RNA é sintetizado sob a forma de uma molécula de fita simples, e vimos que pareamentos por complementaridade de bases podem ocorrer entre nucleotídeos sobre a própria fita (ver Figura 7-5). Esses pareamentos de base, junto a pontes de hidrogênio "não convencionais", podem fazer com que a molécula de RNA se dobre adquirindo uma estrutura específica, a qual é determinada por sua sequência nucleotídica. Tais associações produzem padrões tridimensionais complexos de dobramento, onde a molécula como um todo adota uma forma especial.

Como vimos no Capítulo 4, enzimas proteicas são capazes de catalisar reações bioquímicas, pois possuem uma superfície com contornos característicos e propriedades químicas sobre as quais um dado substrato pode reagir. Da mesma forma, as moléculas de RNA, com suas formas caracteristicamente dobradas, podem atuar como enzimas (Figura 7-44), embora o fato de serem construídas a partir de apenas quatro diferentes subunidades limite sua eficiência catalítica e o espectro de reações químicas que podem por elas ser catalisadas, quando comparado às proteínas. Apesar dessas limitações, ribozimas podem catalisar um espectro bastante variado de reações químicas. A maioria das ribozimas já estudadas foi sintetizada em laboratórios (Tabela 7-4), e relativamente poucos desses RNAs catalíticos existem em células atuais. No entanto, os processos nos quais RNAs catalíticos ainda desempenham um papel importante incluem algumas das etapas mais fundamentais da expressão da informação genética – espe-

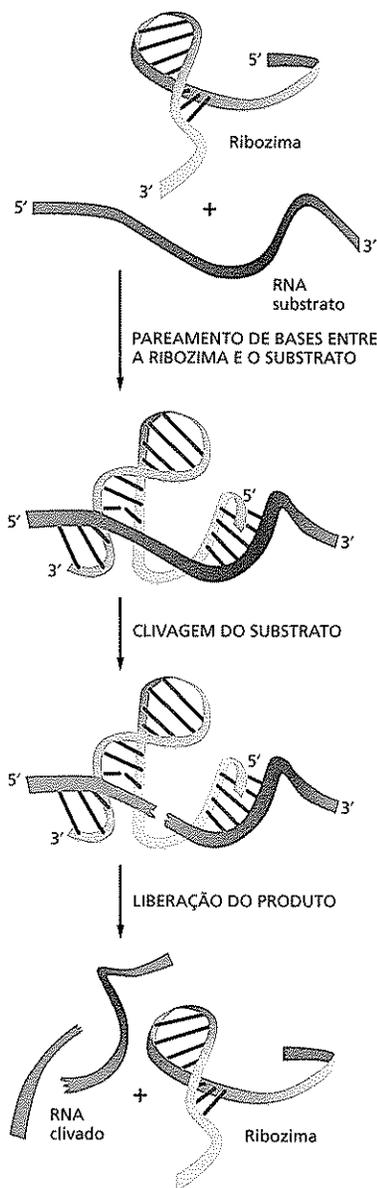


Figura 7-44 Uma ribozima é uma molécula de RNA que possui atividade catalítica. A molécula de RNA simples ilustrada catalisa a clivagem de uma segunda molécula de RNA em um sítio específico. Essa ribozima é encontrada inserida em grandes genomas de RNA – denominados viroides – que infectam plantas. A clivagem, que ocorre na natureza em uma posição distante, sobre a própria molécula de RNA que contém a ribozima, é uma das etapas da replicação do genoma RNA. Essa reação também necessita de um íon magnésio (não representado), que é posicionado próximo ao sítio de clivagem, sobre o substrato. (Adaptado de T.R. Cech e O.C. Uhlenbeck, *Nature* 372:39-40, 1994. Com permissão de Macmillan Publishers Ltd.)

**TABELA 7-4** Reações bioquímicas que podem ser catalisadas por ribozimas

Atividade	Ribozimas
Clivagem de RNA, ligação de RNA	RNAs de autosplicing
Clivagem de DNA	RNAs de autosplicing
Formação da ligação peptídica na síntese de proteínas	RNA ribossomal
Ligação de DNA	RNA selecionado <i>in vitro</i>
Splicing de RNA	RNAs de autosplicing, RNAs do spliceossomo (?),
Polimerização de RNA	RNA selecionado <i>in vitro</i>
Fosforilação de RNA	RNA selecionado <i>in vitro</i>
Aminoacilação de RNA	RNA selecionado <i>in vitro</i>
Alquilação de RNA	RNA selecionado <i>in vitro</i>
Rotação de ligação C-C (isomerização)	RNA selecionado <i>in vitro</i>

cialmente aquelas etapas onde as próprias moléculas de RNA sofrem *splicing* ou são traduzidas em proteínas.

Assim, o RNA possui todas as características necessárias a uma molécula que possa catalisar sua própria síntese (Figura 7-45). Sistemas de autorreplicação de moléculas de RNA não foram encontrados na natureza, mesmo assim cientistas estão confiantes de que eles possam ser construídos em laboratório. Apesar de essa demonstração não provar que moléculas de RNA autorreplicadoras foram essenciais para a origem da vida na Terra, ela certamente demonstraria que um cenário assim é possível.

### O RNA provavelmente antecedeu o DNA na evolução

As primeiras células na Terra presumivelmente eram bem menos complexas e menos eficientes na reprodução de si mesmas se comparadas mesmo à mais simples das células atuais. Elas devem ter sido compostas por pouco mais do que uma simples membrana delimitando um conjunto de moléculas de autorreplicação e alguns outros poucos componentes necessários para fornecer materiais e energia para sua replicação. Se as especulações evolutivas sobre o RNA salientadas estão corretas, essas células primordiais devem também ter apresentado diferenças fundamentais em relação às células que conhecemos atualmente por estocarem sua informação hereditária sob a forma de RNA, e não DNA.

Evidências que indicam o surgimento do RNA antes do DNA na evolução podem ser encontradas nas diferenças químicas existentes entre eles. A ribose (ver Figura 7-3A), assim como a glicose e outros carboidratos simples, é facilmente formada a partir de formaldeído (HCHO), o qual é um dos principais produtos formados em experimentos que simulam as condições da Terra primitiva. O açúcar desoxirribose é mais difícil de ser obtido e, nas células atuais, é produzido a partir da ribose por uma reação catalisada por uma enzima proteica, sugerindo que a ribose antecedeu a desoxirribose nas células. Presumivelmente, o DNA surgiu tardiamente em cena e então se mostrou mais adaptado do que o RNA como

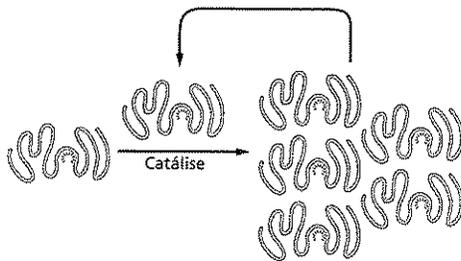
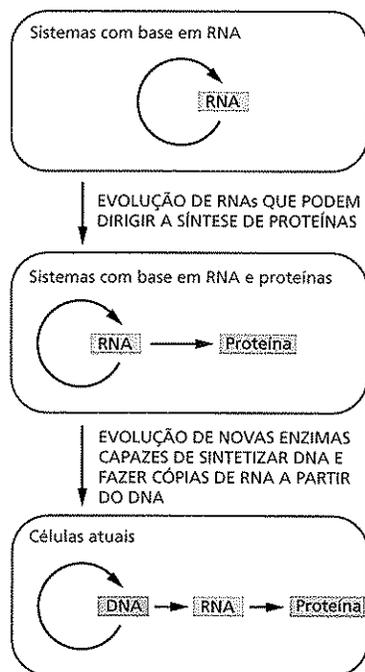


Figura 7-45 Pode uma molécula de RNA catalisar sua própria síntese? Esse processo hipotético necessitaria da catálise de ambas as etapas ilustradas na Figura 7-43. Os raios vermelhos representam o sítio ativo dessa enzima RNA.

**QUESTÃO 7-6**

Discuta a seguinte afirmação: "Ao longo da evolução da vida na Terra, o RNA foi rebaixado de sua gloriosa posição de primeiro catalisador de autorreplicação. Sua função atual é de mero mensageiro no fluxo de informações entre o DNA e as proteínas".

**Figura 7-46** O RNA pode ter antecedido o DNA e as proteínas na evolução. De acordo com essa ideia, as moléculas de RNA proviam funções genéticas, estruturais e catalíticas para as primeiras células. Atualmente, o DNA é o repositório de informação genética, e as proteínas desempenham a quase totalidade das funções catalíticas nas células. O RNA funciona, atualmente, de forma principal como um intermediário na síntese de proteínas, embora permaneça atuando como catalisador em poucas reações essenciais.

repositório permanente de informação genética. Em particular, a desoxirribose com sua estrutura açúcar-fosfato torna as cadeias de DNA muito mais estáveis quimicamente do que as cadeias de RNA, de modo que moléculas mais longas de DNA podem ser mantidas sem que ocorra quebra.

As outras diferenças entre RNA e DNA – a estrutura em dupla-hélice do DNA e o uso de timina em vez de uracila – aumentam ainda mais a estabilidade do DNA ao fazer essa molécula mais fácil de reparar. Vimos, no Capítulo 6, que um nucleotídeo lesado sobre uma fita do DNA dupla-hélice pode ser reparado usando-se a outra fita como molde. Além disso, a desaminação, uma das alterações químicas deletérias mais comuns que ocorrem em polinucleotídeos, é mais fácil de ser detectada e reparada no DNA do que no RNA (ver Figura 6-23). Isso acontece porque o produto da desaminação de citosina é, por uma questão de sorte, a uracila, a qual ocorre normalmente no RNA, de tal forma que seria impossível para enzimas de reparo detectarem tal alteração na molécula de RNA. No entanto, no DNA, que possui timina em vez de uracila, qualquer uracila produzida pela degradação accidental de citosina é facilmente detectada e reparada.

Em conjunto, as evidências que estivessem discutindo sustentam a ideia de que o RNA, possuindo tanto propriedades genéticas quanto catalíticas, antecedeu o DNA na evolução. Acredita-se que conforme células mais semelhantes às células atuais apareceram, muitas das funções originalmente desempenhadas pelo RNA foram sendo direcionadas para moléculas mais especificamente adaptadas para as tarefas em questão. Em um determinado ponto, o DNA se sobrepôs na função genética principal, e as proteínas se tornaram as principais catalisadoras, e o RNA permaneceu predominantemente como um intermediário, conectando-os (Figura 7-46). Com o advento do DNA, as células foram capazes de se tornar mais complexas, pois eram capazes de transportar e transmitir mais informação genética do que poderia anteriormente ser mantido de forma estável em uma molécula de RNA. Tendo em vista a maior complexidade química das proteínas e a maior diversidade das reações químicas que elas podiam catalisar, a troca (apesar de incompleta) do RNA pelas proteínas também forneceu uma fonte muito mais rica de componentes estruturais e enzimas. Isso permitiu que as células evoluíssem sob a ampla diversidade estrutural e funcional que vemos nos seres vivos da atualidade.

**CONCEITOS ESSENCIAIS**

- O fluxo de informação genética em todas as células vivas é DNA → RNA → proteína. A conversão das instruções genéticas do DNA para RNA e proteína é denominada expressão gênica.
- Para expressar a informação genética transportada no DNA, a sequência nucleotídica de um gene é inicialmente transcrita em RNA. A transcrição é catalisada por uma enzima RNA-polimerase. Sequências nucleotídicas na molécula de DNA indicam para a RNA-polimerase onde iniciar e terminar a transcrição.
- O RNA difere do DNA em diversos aspectos. Ele contém o açúcar ribose, em vez de desoxirribose, e a base uracila (U), em vez de timina (T). Os RNAs celulares são sintetizados sob a forma de moléculas de fita simples, as quais frequentemente se dobram adquirindo estruturas tridimensionais bem determinadas.
- As células produzem vários tipos funcionalmente diferentes de RNA, como o RNA mensageiro (mRNA), que transporta as instruções de como

fazer proteínas; o RNA ribossomal (rRNA), que é um dos componentes dos ribossomos, e o RNA transportador (tRNA), que atua como molécula adaptadora na síntese de proteínas.

- A transcrição é iniciada em regiões do DNA denominadas promotores. Para a iniciação da transcrição, as polimerases eucarióticas necessitam da montagem de um complexo de fatores gerais de transcrição sobre o promotor, ao passo que a RNA-polimerase bacteriana necessita apenas de uma subunidade adicional, denominada fator sigma.
- No DNA eucariótico, a maioria dos genes é composta por uma série de pequenas regiões codificadoras menores (éxons) interespaçadas por regiões não codificadoras (íntrons). Quando um gene eucarioto é transcrito do DNA para RNA, tanto os éxons quanto os íntrons são copiados.
- Os íntrons são removidos dos transcritos de RNA no núcleo pelo processo de *splicing* do RNA. Em uma reação catalisada por pequenos complexos ribonucleoproteicos conhecidos como snRNPs, os íntrons são excisados do transcrito primário, e os éxons são unidos uns aos outros.
- Os mRNAs eucarióticos passam por uma série de etapas adicionais de processamento antes de deixarem o núcleo, como o capeamento do RNA e a poliadenilação. Essas reações, juntamente com o *splicing*, ocorrem conforme o RNA está sendo transcrito. A seguir, o mRNA maduro é transportado para o citoplasma.
- A tradução da sequência de nucleotídeos do mRNA em proteína ocorre no citoplasma em grandes agregados ribonucleoproteicos denominados ribossomos. Conforme o mRNA desliza pelo ribossomo, sua mensagem é traduzida em proteína.
- A sequência de nucleotídeos do mRNA é lida em grupos de três nucleotídeos (códon), cada códon correspondendo a um aminoácido.
- A correspondência entre os aminoácidos e os códon é determinada pelo código genético. As possíveis combinações de 4 diferentes nucleotídeos no RNA originam 64 diferentes códon no código genético. A maioria dos aminoácidos é determinada por mais de um códon.
- Os tRNAs atuam como molécula adaptadora na síntese proteica. Enzimas denominadas aminoacil-tRNA-sintetases acoplam os aminoácidos aos tRNAs adequados. Cada tRNA contém uma sequência de três nucleotídeos, o anticódon, que se associa ao códon sobre o mRNA pelo pareamento de bases complementares entre o códon e o anticódon.
- A síntese de proteínas tem início quando um ribossomo se organiza sobre um códon de iniciação (AUG) no mRNA – um processo que é regulado por proteínas denominadas fatores de iniciação da tradução. A cadeia proteica completa é liberada do ribossomo quando um códon de terminação (UAA, UAG ou UGA) é alcançado.
- A ligação sequencial de aminoácidos originando uma cadeia polipeptídica é catalisada por uma molécula de rRNA na subunidade grande ribossomal. Assim, o ribossomo é um exemplo de ribozima, uma molécula de RNA capaz de catalisar uma reação química.
- A degradação de proteínas nas células é cuidadosamente controlada. Algumas proteínas são degradadas no citosol por grandes complexos proteicos denominados proteassomos.
- Considerando nosso conhecimento dos organismos atuais e das moléculas que eles contêm, parece razoável afirmar que os sistemas vivos se originaram a partir da evolução de moléculas RNA capazes de catalisar sua própria replicação.
- Foi proposto que, conforme as células evoluíram, a dupla-hélice de DNA substituiu o RNA como uma molécula mais estável para o armazenamento de informação genética, e as proteínas substituíram os RNAs

como principais componentes catalíticos e estruturais. No entanto, reações importantes, como a formação das ligações peptídicas, são ainda catalizadas por RNA; o que, podemos pensar, nos permite vislumbrar uma parcela do antigo mundo com base em RNA.

#### TERMOS-CHAVE

Splicing alternativo	Fase de leitura
Aminoacil-tRNA sintetase	RNA ribossomal (rRNA)
Anticódon	Ribossomo
Codon	Ribozima
Éxon	RNA
Expressão gênica	RNA-polimerase
Código genético	Processamento de RNA
Fatores gerais de transcrição	Splicing de RNA
tRNA iniciador	Pequenos RNAs nucleares (snRNAs)
Intron	Spliceossomo
RNA mensageiro (mRNA)	Transcrição
Promotor	RNA transportadores (tRNA)
Protease	Tradução
Proteossomo	Fator iniciador da tradução

## TESTE SEU CONHECIMENTO

### QUESTÃO 7-7

Quais das seguintes afirmações estão corretas? Justifique suas respostas.

- Um determinado ribossomo pode fazer apenas um tipo de proteína.
- Todos os mRNAs se dobram, adquirindo estruturas tridimensionais particulares, as quais são necessárias para sua tradução.
- As subunidades grande e pequena de um dado ribossomo permanecem sempre unidas entre elas e nunca substituem a subunidade acompanhante.
- Os ribossomos são organelas citoplasmáticas encapsuladas por uma membrana única.
- Visto que as duas fitas do DNA são complementares, o mRNA de um dado gene pode ser sintetizado utilizando-se qualquer uma das duas fitas como molde.
- Um mRNA deve conter a sequência ATTGACCCCGGTCAA.
- A quantidade de proteína presente em uma célula em um dado estado depende da taxa de síntese dessa proteína, de sua atividade catalítica e de sua taxa de degradação.

### QUESTÃO 7-8

A proteína Lacheinmal é uma proteína hipotética que faz as pessoas sorrírem mais frequentemente. Ela se encontra inativa em muitos indivíduos cronicamente infelizes. O mRNA isolado a partir de diferentes pessoas infelizes pertencentes à mesma família revelou a ausência de um segmento interno ao gene de 173 nucleotídeos, o qual estava presente no mRNA Lacheinmal isolado a partir de um grupo-controle de pessoas geralmente felizes. A sequência do DNA do gene *Lacheinmal* de famílias felizes e infelizes foi determinada e comparada. Essas sequências diferiam em

apenas um nucleotídeo, que se encontrava em um intron. O que pode ser sugerido a respeito da base molecular da infelicidade nessa família?

(Dicas: [1] Você pode sugerir um mecanismo molecular pelo qual uma alteração em um único nucleotídeo de um gene pudesse causar a deleção observada no mRNA? Observe que essa é uma deleção *interna* ao mRNA. [2] Assumindo-se que as 173 bases deletadas removem sequências codificadoras do mRNA Lacheinmal, quais seriam as diferenças entre a proteína Lacheinmal de pessoas felizes e a de pessoas infelizes?)

### QUESTÃO 7-9

Utilize o código genético ilustrado na Figura 7-24 para identificar quais das seguintes sequências nucleotídicas codificarão uma sequência polipeptídica arginina-glicina-aspartato:

- 5'-AGA-GGA-GAU-3'
- 5'-ACA-CCC-ACU-3'
- 5'-GGG-AAA-UUU-3'
- 5'-CGG-GGU-GAC-3'

### QUESTÃO 7-10

"As ligações que se formam entre o anticódon de uma molécula tRNA e os três nucleotídeos de um códon sobre o mRNA são \_\_\_\_\_." Complete essa sentença com cada uma das opções seguintes e explique o porquê de as frases estarem corretas ou incorretas.

- Ligações covalentes formadas por hidrólise de GTP.
- Pontes de hidrogênio que se formam quando o tRNA está no sítio A.
- Quebradas pela translocação do ribossomo sobre o mRNA.

### QUESTÃO 7-11

Liste as definições comuns encontradas em dicionário para os termos *replicação*, *transcrição* e *tradução*. Ao lado de cada de-

finalização, liste o significado específico de cada um desses termos aplicado a células vivas.

#### QUESTÃO 7-12

Em um mundo alienígena, o código genético é escrito em pares de nucleotídeos. Quantos aminoácidos esse código pode determinar? Em um mundo diferente, um código de tripletes é usado, mas a sequência dos nucleotídeos não é importante, sendo apenas considerada a presença ou ausência de cada tipo de nucleotídeo. Quantos aminoácidos esse código genético poderia determinar? Você poderia imaginar algum problema referente à tradução desses códigos?

#### QUESTÃO 7-13

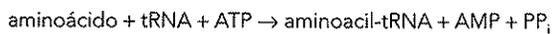
Uma característica impressionante do código genético é o fato de aminoácidos que apresentam propriedades químicas similares frequentemente possuírem códons similares. Desse modo, códons com U ou C como segundo nucleotídeo tendem a especificar aminoácidos hidrofóbicos. Você pode sugerir uma explicação plausível para esse fenômeno considerando a evolução inicial da maquinaria de síntese proteica?

#### QUESTÃO 7-14

Uma mutação no DNA gera um códon de terminação UGA no meio de um RNA que codifica uma determinada proteína. Uma segunda mutação nessa célula leva à alteração de um único nucleotídeo sobre um tRNA e permite a tradução correta da proteína; ou seja, essa segunda mutação "suprime" o defeito causado pela primeira. O tRNA alterado traduz o UGA como triptofano. Que alteração nucleotídica provavelmente ocorreu na molécula mutante de tRNA? Quais as consequências potenciais da presença de tal tRNA mutado na tradução dos genes normais dessa célula?

#### QUESTÃO 7-15

O carregamento de um tRNA com um aminoácido pode ser representado pela seguinte equação:



onde  $\text{PP}_i$  representa o pirofosfato (ver Figura 3-40). No aminoacil-tRNA, o aminoácido e o tRNA estão ligados por uma ligação de alta energia; uma grande parte da energia derivada da hidrólise de ATP é estocada desse modo nessa ligação e está disponível para conduzir a formação de ligações peptídicas em estágios posteriores da síntese proteica. A alteração de energia livre da reação de carregamento ilustrada na equação é próxima de zero e, conseqüentemente, não se esperaria que favorecesse

a associação do aminoácido ao tRNA. Que etapa suplementar que direcionaria a ocorrência da reação pode ser sugerida?

#### QUESTÃO 7-16

- O peso molecular médio das proteínas de uma célula é de aproximadamente 30.000 dáltons. Algumas proteínas, no entanto, são muito maiores. A maior cadeia polipeptídica conhecida produzida por uma célula se refere à proteína denominada titina (produzida em células musculares de mamíferos), que possui um peso molecular de 3.000.000 de dáltons. Estime o tempo necessário para que uma célula muscular traduza uma molécula de mRNA que codifica a titina (considere o peso molecular médio de um aminoácido igual a 120 e uma taxa de tradução de dois aminoácidos por segundo para células eucarióticas).
- A síntese proteica é bastante exata: é cometido apenas um erro a cada 10.000 aminoácidos unidos. Qual é a fração de moléculas proteicas de tamanho médio e de moléculas de titina que são sintetizadas sem qualquer erro? (Dica: a probabilidade  $P$  de obter uma proteína livre de erros é dada por  $P = (1 - E)^n$ , onde  $E$  é a taxa de erros e  $n$  o número de aminoácidos.)
- O peso molecular combinado de todas as proteínas ribossomais eucarióticas é de aproximadamente  $2,5 \times 10^6$  dáltons. Seria vantajoso sintetizá-las sob a forma de uma única proteína?
- A transcrição ocorre a uma taxa de aproximadamente 30 nucleotídeos por segundo. Seria possível calcular o tempo necessário para a síntese do mRNA da titina a partir das informações anteriormente fornecidas?

#### QUESTÃO 7-17

Quais das seguintes alterações mutacionais podem ser consideradas como deletérias para um organismo? Justifique suas respostas.

- Inserção de um único nucleotídeo próximo ao fim da sequência codificadora.
- Deleção de um único nucleotídeo próximo ao início da sequência codificadora.
- Deleção de três nucleotídeos consecutivos na região mediana da sequência codificadora.
- Deleção de quatro nucleotídeos consecutivos na região mediana da sequência codificadora.
- Substituição de um nucleotídeo por outro na região mediana da sequência codificadora.