

FUNDAMENTOS DA BIOLOGIA CELULAR

4ª Edição

ALBERTS • BRAY • HOPKIN • JOHNSON • LEWIS • RAFF • ROBERTS • WALTER



7

Do DNA à proteína: como as células leem o genoma

Uma vez que a estrutura de dupla-hélice do DNA (ácido desoxirribonucleico) foi determinada no início da década de 1950, tornou-se claro que a informação hereditária nas células está codificada na ordem linear – ou *sequência* – das quatro subunidades de nucleotídeos diferentes que compõem o DNA. Vimos, no Capítulo 6, como essa informação pode ser transmitida, de modo conservado, de uma célula às suas descendentes pelo processo de replicação do DNA. No entanto, como uma célula decodifica e usa essa informação? Como as instruções genéticas escritas sob a forma de um alfabeto de apenas quatro “letras” podem levar à formação de uma bactéria, uma mosca-da-fruta 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 até da mais simples bactéria unicelular, o que não dizer de como ela pode direcionar o desenvolvimento de organismos multicelulares complexos, como nós mesmos? Mas o próprio código do DNA foi decifrado, e já percorremos um longo caminho na compreensão de como as células o leem.

Mesmo antes de termos decifrado o código do 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 estrutura celular, mas também as suas funções. Nos capítulos anteriores, deparamo-nos com alguns dos milhares de tipos diferentes de proteínas que podem ser produzidos pelas células. Vimos, no Capítulo 4, que as propriedades e funções de uma molécula de proteína são determinadas pela sequência das 20 diferentes subunidades de aminoácidos em sua cadeia polipeptídica: cada tipo de proteína tem a sua sequência de aminoácidos característica, que dita como a cadeia vai dobrar-se para dar origem a uma molécula com forma e características químicas definidas. As instruções genéticas transportadas pelo DNA devem, portanto, especificar a sequência dos aminoácidos nas proteínas. No presente capítulo, vamos ver como isso realmente acontece.

O DNA *per se* não sintetiza proteínas, mas atua como um gerente, delegando as diferentes tarefas a uma equipe de trabalhadores. Quando uma determinada proteína é necessária para a célula, a sequência de nucleotídeos do segmento apropriado de uma molécula de DNA é inicialmente copiada para outra forma de ácido nucleico – o RNA (*ácido ribonucleico*). Esse segmento de DNA é denominado **gene**, e as cópias de RNA resultantes são utilizadas para dirigir a síntese da proteína. Milhares dessas conversões de DNA para proteína ocorrem a cada segundo em cada uma das células do nosso organismo. O fluxo da informação genética nas células segue, portanto, uma rota do DNA para o RNA e deste 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.

DO DNA AO RNA

DO RNA À PROTEÍNA

RNA E A ORIGEM DA VIDA

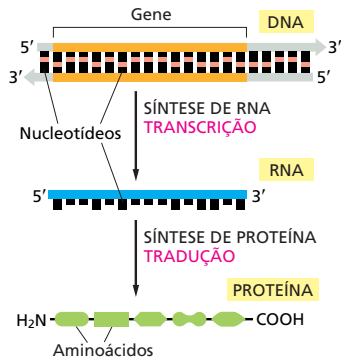


Figura 7-1 A informação genética direciona a síntese de proteínas. O fluxo de informação genética do DNA ao RNA (transcrição) e do RNA à proteína (tradução) ocorre em todas as células vivas. Foi Francis Crick que apelidou esse fluxo de informação de “o dogma central”. Os segmentos de DNA que são transcritos em RNA são chamados de genes.

Neste capítulo, abordamos 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*). Discutimos também algumas das principais variações que ocorrem nesse esquema básico. Destaca-se entre elas o *splicing do RNA* (ou, em português, encadeamento do RNA), um processo, nas células eucarióticas, em que segmentos de um *transcrito de RNA* são removidos – e os segmentos restantes são unidos entre si – antes que o RNA seja traduzido em proteína. Na seção final, consideramos como o esquema de estoque de informação, de transcrição e de tradução atual deve ter se originado a partir de sistemas muito mais simples, nos estágios iniciais da evolução celular.

DO DNA AO RNA

A transcrição e a tradução são os processos pelos quais as células leem, ou *expressam*, as instruções escritas em 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. Essa amplificação sucessiva permite que as células sintetizem rapidamente, e no momento necessário, grandes quantidades de proteína. Ao mesmo tempo, cada gene pode ser transcrito, e seu RNA traduzido, a diferentes taxas, possibilitando que a célula produza grandes quantidades de algumas proteínas e pequenas quantidades de outras (Figura 7-2). Além disso, como discutido 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, discutimos a produção do RNA – a primeira etapa da *expressão gênica*.

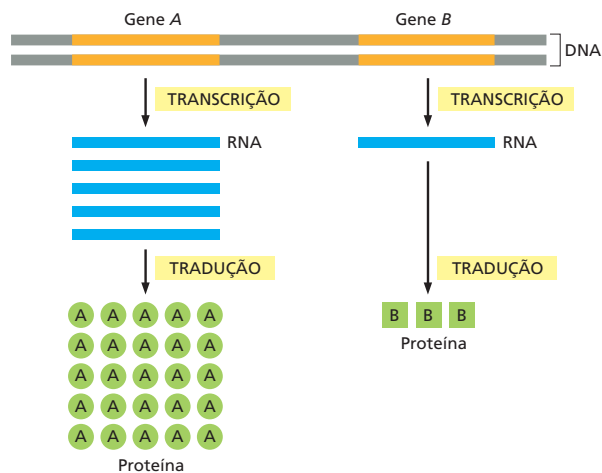
Segmentos da sequência de DNA são transcritos em RNA

O primeiro passo que uma célula dá para expressar 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 na mesma linguagem – a linguagem dos nucleotídeos. Assim como o DNA, o RNA é um polímero linear composto por quatro diferentes subunidades nucleotídicas unidas entre si por ligações fosfodiéster. 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 (Figura

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 7-2 Uma célula pode expressar diferentes genes em diferentes taxas. Nesta figura, e nas seguintes, as porções não transcritas do DNA são mostradas em cinza.



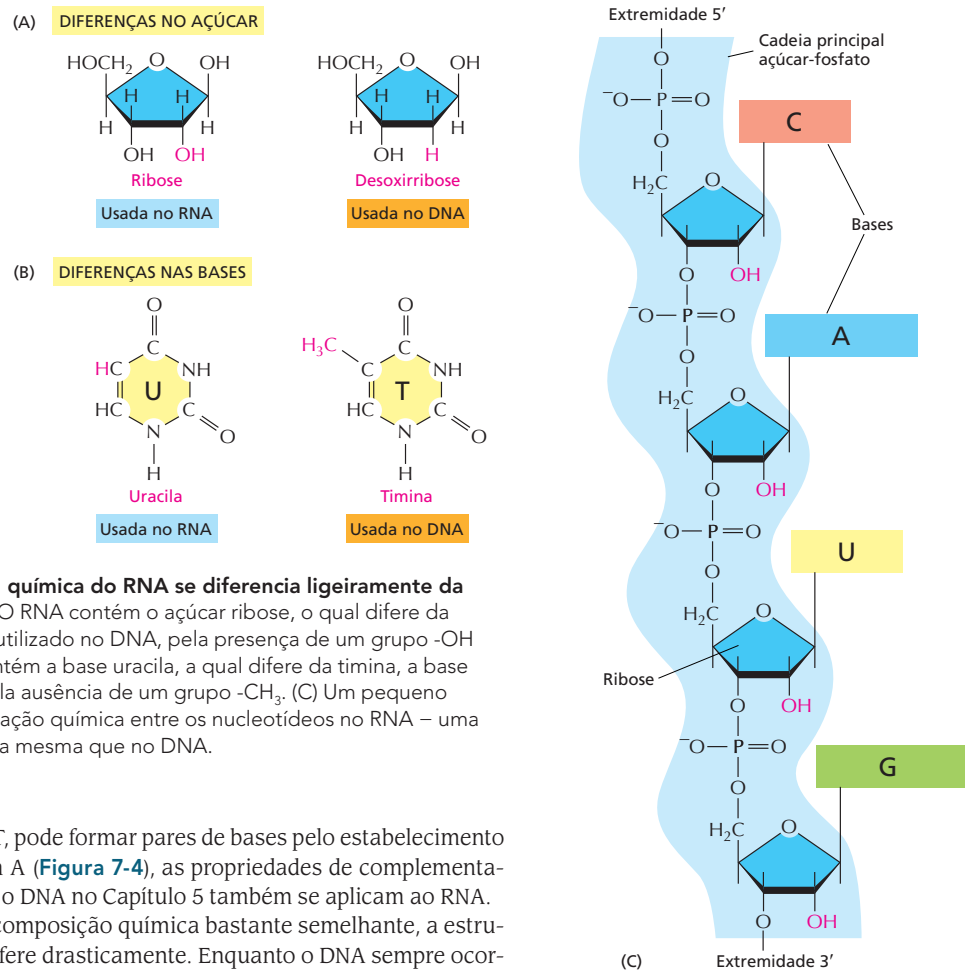


Figura 7-3 A estrutura química do RNA se diferencia ligeiramente da estrutura do DNA. (A) O RNA contém o açúcar ribose, o qual difere da desoxirribose, o açúcar utilizado no DNA, pela presença de um grupo -OH adicional. (B) O RNA contém a base uracila, a qual difere da timina, a base equivalente no DNA, pela ausência de um grupo -CH₃. (C) Um pequeno segmento de RNA. A ligação química entre os nucleotídeos no RNA – uma ligação fosfodiéster – é a mesma que no DNA.

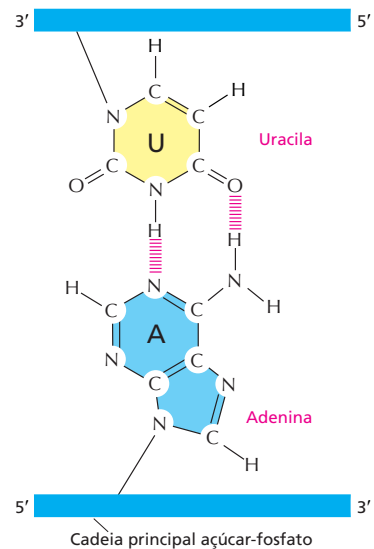
7-3). Visto que U, assim como T, pode formar pares de bases pelo estabelecimento de ligações de hidrogênio com A (Figura 7-4), as propriedades de complementaridade de bases descritas para o DNA no Capítulo 5 também se aplicam ao RNA.

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 conformações, exatamente como ocorre com o dobramento de uma cadeia polipeptídica na estruturação fina de uma proteína (Figura 7-5); o DNA de fita dupla não pode dobrar-se desse modo. Como discutimos mais adiante, a capacidade de dobrar-se em estruturas tridimensionais complexas permite que o RNA desempenhe várias funções na célula que vão muito além das de simples intermediário de informações entre DNA e proteína. Enquanto as funções do DNA limitam-se ao estoque de informação, alguns RNAs possuem funções estruturais, reguladoras ou catalíticas.

A transcrição produz um RNA que é 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 Cap. 6). A transcrição tem início com a abertura e a desespiralização de uma pequena

Figura 7-4 A uracila forma pares de bases com a adenina. As ligações de hidrogênio que mantêm unido um par de bases são mostradas em *vermelho*. A uracila tem as mesmas propriedades de pareamento de bases que a timina. Assim, pares de bases U-A no RNA assemelham-se a pares de bases A-T no DNA (ver Figura 5-6A).



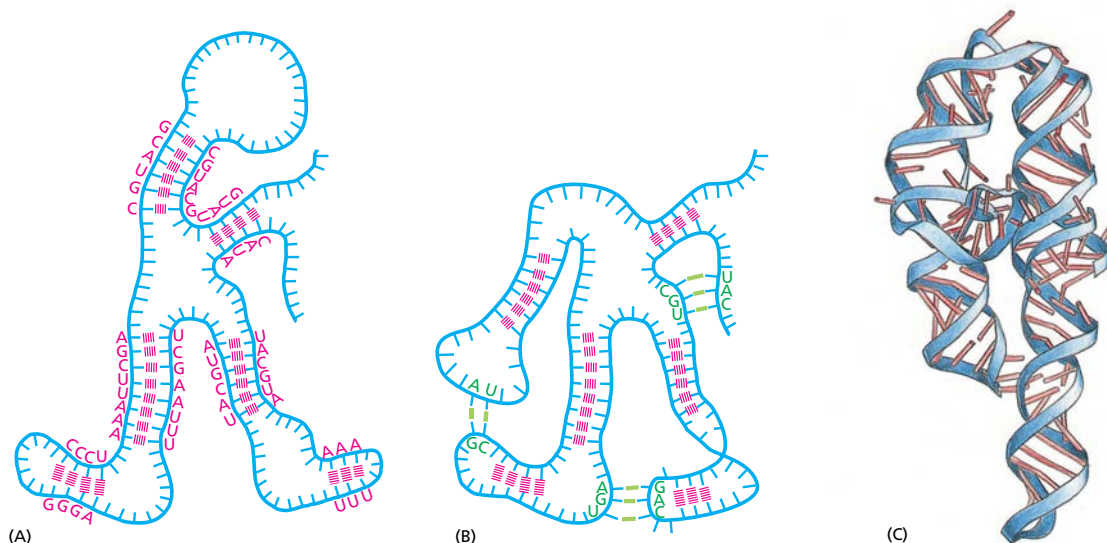


Figura 7-5 As moléculas de RNA podem formar pares de bases intramolecularmente e se dobrar em estruturas específicas. O RNA é uma fita simples, mas frequentemente contém pequenos segmentos de nucleotídeos que podem sofrer pareamento com sequências complementares encontradas em outras regiões da mesma molécula. Essas interações, junto a algumas interações de pares de base “não convencionais” (p. ex., A-G), permitem que uma molécula de RNA se dobre em uma estrutura tridimensional que é determinada pela sua sequência de nucleotídeos. (A) Um diagrama de uma estrutura de RNA hipotética dobrada, mostrando apenas interações convencionais de pares de bases (G-C e A-U). (B) A incorporação de interações de pares de bases não convencionais (verde) altera a estrutura do RNA hipotético ilustrado em (A). (C) Estrutura de uma molécula de RNA real que está envolvida no *splicing* de RNA. Esse RNA contém uma quantidade considerável de estruturas em dupla-hélice. A cadeia principal de açúcar-fosfato está indicada em azul e as bases em vermelho; as interações convencionais de pares de bases estão indicadas por “degraus” vermelhos contínuos, e os pares de bases não convencionais estão indicados por degraus vermelhos interrompidos. Para visualização adicional da estrutura do RNA, ver [Animação 7.1](#).

porção da dupla-hélice de DNA para que as bases de ambas as fitas do DNA sejam expostas. 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; da mesma forma que ocorre na replicação do DNA, a sequência nucleotídica da cadeia 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 cadeia de RNA em crescimento pela enzima *RNA-polimerase*. A cadeia de RNA produzida pela transcrição – o **transcrito de RNA** – é, desse modo, estendida nucleotídeo a nucleotídeo e apresenta sequência nucleotídica exatamente complementar à fita de DNA usada como molde ([Figura 7-6](#)).

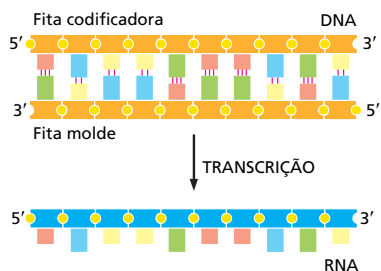


Figura 7-6 A transcrição de um gene produz uma molécula de RNA complementar a uma das fitas do DNA. A fita transcrita do gene, a fita *inferior* neste exemplo, é chamada de fita molde. A fita não molde do gene (neste caso, mostrada na parte *superior*) é muitas vezes chamada de *fita codificadora*, pois a sua sequência é equivalente ao produto de RNA, como ilustrado. A fita de DNA que serve como molde varia, dependendo do gene, conforme discutido mais adiante. Para fins de convenção, uma molécula de RNA é sempre escrita ou desenhada com a sua extremidade 5' – a primeira porção a ser sintetizada – à esquerda.

A transcrição difere da replicação de DNA em vários aspectos essenciais. Diferentemente de uma fita de DNA recém-formada, a fita de RNA não permanece ligada à fita de DNA molde. 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 das fitas da molécula de DNA é transcrita, as moléculas de RNA são constituídas de fita simples. Além disso, como os RNAs são copiados somente a partir de uma região definida do DNA, as moléculas de RNA 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 poucos milhares de nucleotídeos, sendo muitos deles ainda bem menores do que isso.

Assim como a DNA-polimerase que catalisa a replicação do DNA (discutida no Cap. 6), as **RNA-polimerases** catalisam a formação de ligações fosfodiéster que unem os nucleotídeos e formam a cadeia principal de açúcar-fosfato de uma cadeia de RNA (ver [Figura 7-3](#)). A RNA-polimerase se move paulatinamente sobre o DNA, desenrolando a hélice de DNA à sua frente e expondo a nova região da fita molde para que ocorra o pareamento por complementaridade de bases. Des-

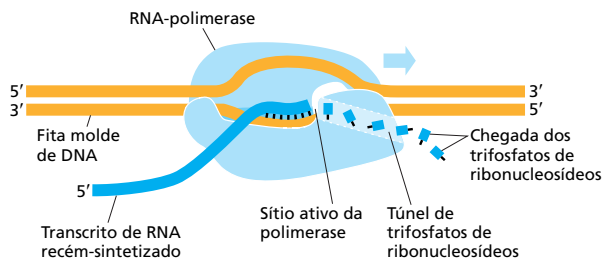


Figura 7-7 O DNA é transcrito em RNA pela enzima RNA-polimerase. A RNA-polimerase (azul-claro) se move paulatinamente ao longo do DNA, desespiralizando a hélice de DNA à sua frente. À medida que avança, a polimerase adiciona ribonucleotídeos, um a um, à cadeia de RNA, utilizando uma fita exposta do DNA como molde. O transcrito de RNA resultante é, portanto, uma fita simples e complementar a essa fita molde (ver Figura 7-6). Conforme a polimerase se move ao longo do DNA molde (no sentido 3' para 5'), ela desloca o RNA recém-formado, permitindo que as duas fitas de DNA atrás da polimerase se reassociem. Uma região curta de hélice híbrida DNA/RNA (com cerca de 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).

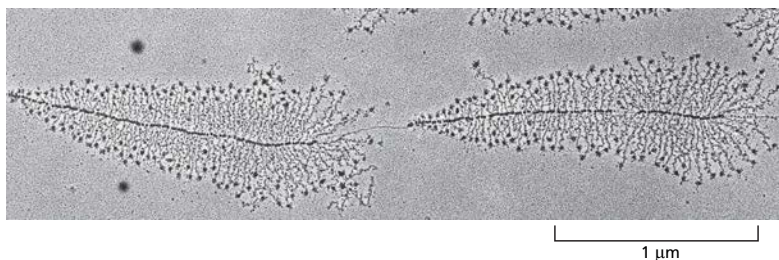
se modo, a cadeia de RNA em crescimento é estendida nucleotídeo a nucleotídeo na direção de 5' para 3' (Figura 7-7). Os trifosfatos de ribonucleosídeos recém-chegados (ATP, CTP, UTP e GTP) fornecem a energia necessária para a continuidade da reação (ver Figura 6-11).

A liberação quase imediata da fita de RNA recém-sintetizada da fita de DNA molde permite 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 tenha sido completada (Figura 7-8). Um gene de tamanho mediano – digamos, de 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 dúzias de polimerases percorrendo esse pequeno segmento de DNA, umas nos calcanhares das outras, o que permite a síntese de mais de 1.000 transcritos no período de uma hora. Na maioria dos genes, entretanto, a taxa de transcrição é bem mais baixa do que essa.

Embora a RNA-polimerase catalise 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 usa ribonucleosídeos como substrato para fosfatos e, portanto, ela catalisa a ligação de ribonucleotídeos, e não 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 provavelmente evoluiu porque a transcrição não precisa ser tão exata quanto a replicação do DNA; diferentemente do DNA, o RNA não é usado como a forma de armazenamento permanente de informação genética nas células, de modo que erros em transcritos de RNA apresentarão consequências relativamente menores para uma célula. As RNA-polimerases cometem aproximadamente um erro a cada 10^4 nucleotídeos copiados em RNA, ao passo que as DNA-polimerases cometem apenas um erro a cada 10^7 nucleotídeos copiados.

As células produzem vários tipos de RNA

A grande maioria dos genes presentes no DNA de uma célula específica as sequências de aminoácidos das proteínas. As moléculas de RNA codificadas por esses genes – que em última instância dirigem a síntese das proteínas – são chamadas de **RNAs mensageiros (mRNAs)**. Em eucariotos, cada mRNA geralmente contém a informação transcrita a partir de um único gene, que codifica uma única proteína. Em bactérias, um conjunto de genes adjacentes é frequentemente



QUESTÃO 7-2

Na micrografia 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 os segmentos de DNA (genes) que os codificam?

Figura 7-8 A transcrição pode ser visualizada sob microscopia eletrônica. A micrografia mostra muitas moléculas de RNA-polimerase transcrevendo simultaneamente dois genes ribossômicos adjacentes em uma única molécula de DNA. As moléculas de RNA-polimerase estão pouco visíveis, como uma série de pequenos pontos ao longo da coluna da molécula de DNA; cada polimerase produz um transcrito de RNA (uma linha fina e curta) que irradia dela. As moléculas de RNA que estão sendo transcritas a partir dos dois genes ribossômicos – RNAs ribossômicos (rRNAs) – não são traduzidas em proteína, mas serão usadas diretamente como componentes dos ribossomos, máquinas macromoleculares feitas de RNAs e proteínas. Acredita-se que as partículas grandes que podem ser vistas nas extremidades 5' livres de cada transcrito de rRNA sejam proteínas ribossômicas que se associaram às extremidades dos transcritos em crescimento. (Cortesia de Ulrich Scheer.)

transcrito em um único mRNA que, conseqüentemente, possui informação para a produção de diferentes proteínas.

O produto final de outros genes, no entanto, é o próprio RNA. Como veremos mais adiante, esses RNAs não mensageiros, assim como as proteínas, têm várias funções, atuando como componentes reguladores, estruturais e catalíticos das células. Eles desempenham papéis fundamentais, por exemplo, na tradução da mensagem genética em proteína: os *RNAs ribossômicos (rRNAs)* formam o núcleo estrutural e catalítico dos ribossomos, que traduzem os mRNAs em proteína, e os *RNAs transportadores (tRNAs)* atuam como adaptadores que selecionam aminoácidos específicos e os posicionam adequadamente sobre um ribossomo para serem incorporados em uma proteína. Outros pequenos RNAs, denominados *microRNAs (miRNAs)*, atuam como importantes reguladores da expressão gênica em eucariotos, como discutiremos no Capítulo 8. Os tipos mais comuns de RNAs estão listados na **Tabela 7-1**.

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 em 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.

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

A iniciação da transcrição é um processo especialmente importante, pois é o principal momento no qual a célula seleciona quais proteínas ou RNAs deverão ser produzidos. 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* de um gene difere consideravelmente entre bactérias e eucariotos. Visto que essa situação é mais simples em bactérias, inicialmente nos deteremos no sistema procariótico.

Quando uma RNA-polimerase colide aleatoriamente com uma molécula de DNA, a enzima adere fracamente à dupla-hélice e, em seguida, desliza rapidamente sobre ela. A RNA-polimerase adere fortemente ao DNA somente após ter encontrado uma região do gene chamada de **promotor**, que contém uma sequência específica de nucleotídeos posicionada imediatamente a montante do ponto de início para a síntese do RNA. Uma vez ligada firmemente a essa sequência, a RNA-polimerase separa a dupla-hélice imediatamente em frente ao promotor para expor os nucleotídeos de um segmento curto de cada fita de DNA. Uma das duas fitas de DNA expostas funciona como um molde para o pareamento de bases complementares com os trifosfatos de ribonucleotídeos que aí chegam, e dois desses ribonucleotídeos são unidos pela polimerase para dar início à síntese da cadeia de RNA. Por meio desse sistema, a extensão, ou alongamento, da cadeia

TABELA 7-1 Tipos de RNAs produzidos nas células

Tipo de RNA	Função
RNAs mensageiros (mRNAs)	Codificam proteínas
RNAs ribossômicos (rRNAs)	Formam a região central da estrutura do ribossomo e catalisam a síntese proteica
microRNAs (miRNAs)	Regulam a expressão dos genes
RNAs transportadores (tRNAs)	Usados como adaptadores entre o mRNA e os aminoácidos durante a síntese proteica
Outros RNAs não codificadores	Usados no <i>splicing</i> do mRNA, na regulação gênica, na manutenção de telômeros e em diversos outros processos celulares

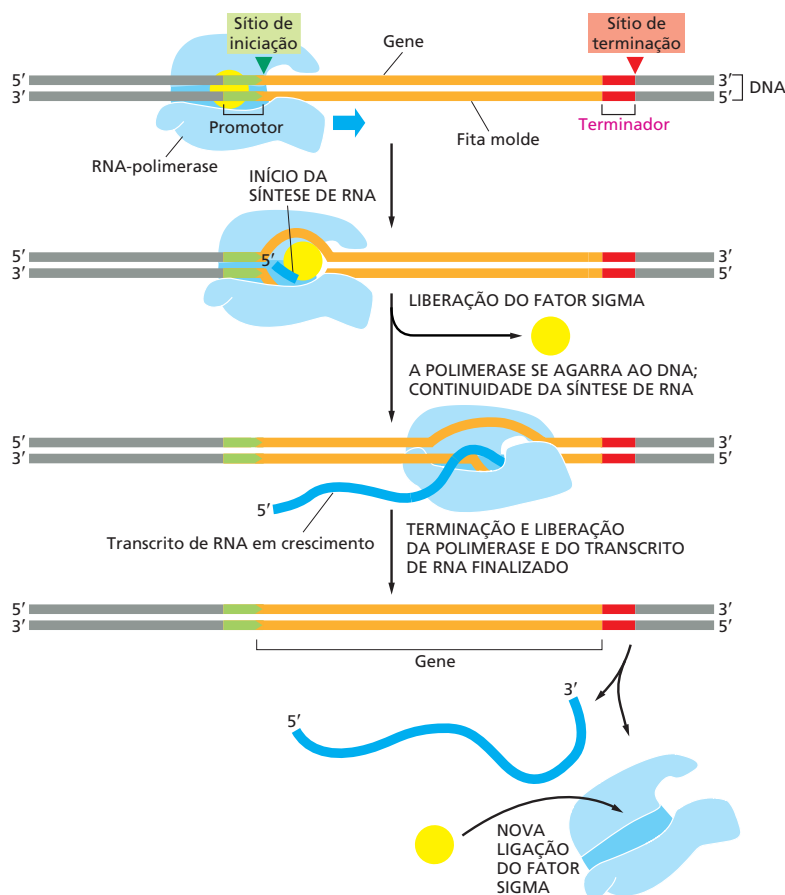


Figura 7-9 Sinais na sequência de nucleotídeos de um gene indicam à RNA-polimerase bacteriana onde iniciar e terminar a transcrição. A RNA-polimerase bacteriana (azul-clara) contém uma subunidade denominada fator sigma (amarelo), que reconhece o promotor de um gene (verde). Uma vez que a transcrição tenha iniciado, o fator sigma é liberado e a polimerase move-se para frente e continua a sintetizar RNA. O alongamento da cadeia prossegue até que a polimerase encontre uma sequência, no gene, denominada terminador (vermelho). Nesse ponto, a enzima para e libera tanto o DNA molde quanto o transcrito de RNA recém-sintetizado. A seguir, a polimerase se reassocia a um fator sigma livre e recomeça a busca por outro promotor para reiniciar o processo.

continua até que a enzima encontre um segundo sinal sobre o DNA, o *terminador* (sítio de parada, terminação ou término), onde a polimerase se detém e libera tanto o DNA molde quanto o transcrito de RNA recém-sintetizado (Figura 7-9). Essa sequência terminadora está contida no gene e é transcrita na extremidade 3' do RNA recém-sintetizado.

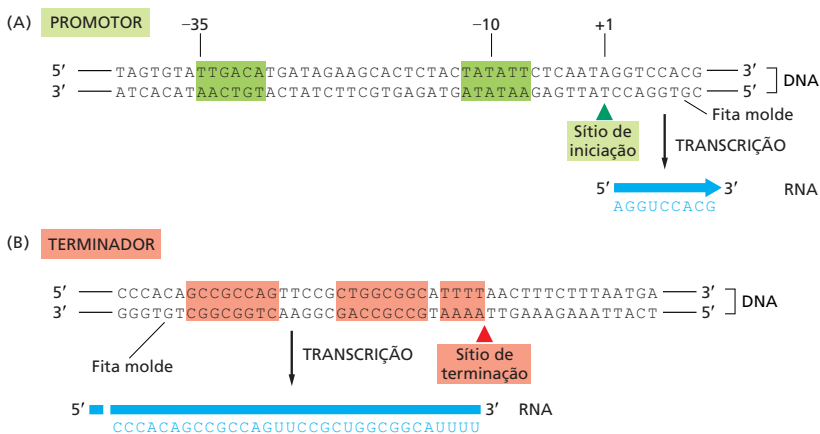
Visto que a polimerase deve ligar-se fortemente antes de poder começar a transcrição, um segmento de DNA só será transcrito se for precedido por um promotor. Isso assegura a transcrição em RNA das porções da molécula de DNA que contêm um gene. As sequências nucleotídicas de um promotor típico – e de um terminador típico – estão representadas na Figura 7-10.

Em bactérias, uma subunidade da RNA-polimerase, o *fator sigma* (σ) (ver Figura 7-9), é o principal responsável pelo reconhecimento da sequência promotora no DNA. Mas como pode esse fator “ver” o promotor, se os pares de bases em questão estão escondidos no interior da dupla-hélice do DNA? O fato é que cada base apresenta características específicas na porção voltada para o exterior da dupla-hélice, permitindo que o fator sigma encontre a sequência do promotor sem que haja a necessidade de separar as fitas espiralizadas do DNA.

O problema seguinte que uma RNA-polimerase enfrenta é determinar qual das duas fitas de DNA utilizar como molde para a transcrição: cada fita tem uma sequência diferente de nucleotídeos e irá produzir um transcrito de RNA distinto. O segredo da escolha reside na estrutura do próprio promotor. Cada promotor tem uma polaridade determinada: ele contém duas sequências diferentes de nucleotídeos, a montante do sítio de início da transcrição, que posicionam a RNA-polimerase, assegurando que ela se ligue ao promotor sob uma única orientação (ver Figura 7-10A). Visto que a polimerase só pode sintetizar RNA na direção 5' para 3', ao ser posicionada a enzima deverá necessariamente usar a fita de DNA orientada no sentido 3' para 5' como molde.

Figura 7-10 Promotores e terminadores bacterianos possuem seqüências nucleotídicas específicas que são reconhecidas pela RNA-polimerase.

(A) As regiões sombreadas em verde representam as seqüências de nucleotídeos que especificam um promotor. Os números acima do DNA indicam a posição do nucleotídeo, contada a partir do primeiro nucleotídeo transcrito, o qual é denominado +1. A polaridade do promotor orienta a polimerase e determina a fita de DNA a ser transcrita. Todos os promotores bacterianos possuem seqüências de DNA a -10 e a -35 que se assemelham bastante às aqui ilustradas. (B) As regiões sombreadas em vermelho representam as seqüências no gene que sinalizam o término da transcrição para a RNA-polimerase. Observe que as regiões transcritas em RNA incluem o terminador, mas não as seqüências de nucleotídeos do promotor. Por questão de convenção, a seqüência de um gene é aquela referente à fita não molde, visto que essa fita apresenta a mesma seqüência do RNA transcrito (com T nos sítios referentes a U).



Essa seleção de uma fita molde não significa que, em um dado cromossomo, a transcrição procederá sempre na mesma direção. No que diz respeito ao cromossomo como um todo, a direção da transcrição varia de gene para gene. No entanto, como cada gene possui geralmente apenas um promotor, a orientação do seu promotor determina em que direção o gene será transcrito e, por conseguinte, qual das duas fitas é a fita molde (Figura 7-11).

A iniciação da transcrição gênica em eucariotos é um processo complexo

Muitos dos princípios que descrevemos até o momento para a transcrição em bactérias também se aplicam aos eucariotos. Contudo, a iniciação da transcrição em eucariotos se diferencia da de bactérias em uma série de pontos importantes:

- A primeira diferença reside nas próprias RNA-polimerases. Enquanto as bactérias contêm um único tipo de RNA-polimerase, as células de eucariotos possuem três: *RNA-polimerase I*, *RNA-polimerase II* e *RNA-polimerase III*. Essas polimerases são responsáveis pela transcrição de diferentes tipos de genes. As RNA-polimerases I e III transcrevem os genes que codificam os RNAs transportadores, os RNAs ribossômicos e vários outros RNAs que desempenham papéis estruturais e catalíticos nas células (Tabela 7-2). A RNA-polimerase II transcreve a ampla maioria dos genes de eucariotos, incluindo todos aqueles que codificam proteínas e miRNAs (Animação 7.3). Nossa discussão subsequente tem como foco, portanto, a RNA-polimerase II.
- Uma segunda diferença é que, enquanto a RNA-polimerase bacteriana (em conjunto à sua subunidade sigma) é capaz de dar início ao processo de transcrição de forma independente, as RNA-polimerases de eucariotos necessitam da assistência de um grande número de proteínas acessórias. Entre essas proteínas acessórias, são essenciais os *fatores gerais de transcrição*, que devem associar-se a cada promotor, em conjunto com a polimerase, antes que essa enzima possa iniciar a transcrição.
- Uma terceira característica que diferencia a transcrição em eucariotos é que os mecanismos que controlam a sua iniciação são muito mais complexos do que os existentes em procariotos – um ponto que será amplamente discutido no Ca-

Figura 7-11 Em um cromossomo, alguns genes são transcritos usando uma das fitas do DNA como molde, enquanto outros são transcritos a partir da outra fita do DNA. A RNA-polimerase sempre se move no sentido de 3' para 5' e a escolha da fita molde é determinada pela orientação do promotor (pontas de seta verdes) no início de cada gene. Assim, os genes transcritos da esquerda para a direita utilizam a fita inferior de DNA como molde (ver Figura 7-10); aqueles transcritos da direita para a esquerda utilizam a fita superior como molde.

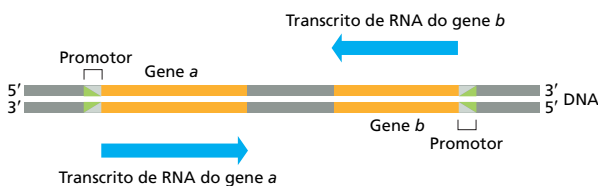


TABELA 7-2 As três RNA-polimerases de células eucarióticas	
Tipo de polimerase	Genes transcritos
RNA-polimerase I	A maioria dos genes de rRNAs
RNA-polimerase II	Todos os genes codificadores de proteínas, os genes de miRNA, além dos genes para outros RNAs não codificadores (p. ex., aqueles do spliceossomo)
RNA-polimerase III	Genes do tRNA Gene do rRNA 5S Genes de diversos outros pequenos RNAs

pítulo 8. Em bactérias, os genes tendem a se organizar sobre o DNA próximos uns dos outros, com apenas pequenas regiões de DNA não transcritas entre eles. No entanto, tanto no DNA vegetal quanto no de animais, inclusive em seres humanos, os genes se encontram amplamente dispersos, existindo regiões de DNA não transcrito com comprimento de até 100.000 pares de nucleotídeos entre um gene e o seguinte. Essa arquitetura permite que um único gene seja controlado por uma ampla variedade de *sequências de DNA reguladoras* distribuídas ao longo do DNA, e permite que os eucariotos utilizem formas de regulação transcricional muito mais complexas do que as das 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 de cromatina estruturalmente mais compactas, como descrito no Capítulo 8.

Agora, voltamos aos fatores gerais de transcrição e discutimos como eles auxiliam a RNA-polimerase II dos eucariotos na iniciação da transcrição.

A RNA-polimerase de eucariotos requer fatores gerais de transcrição

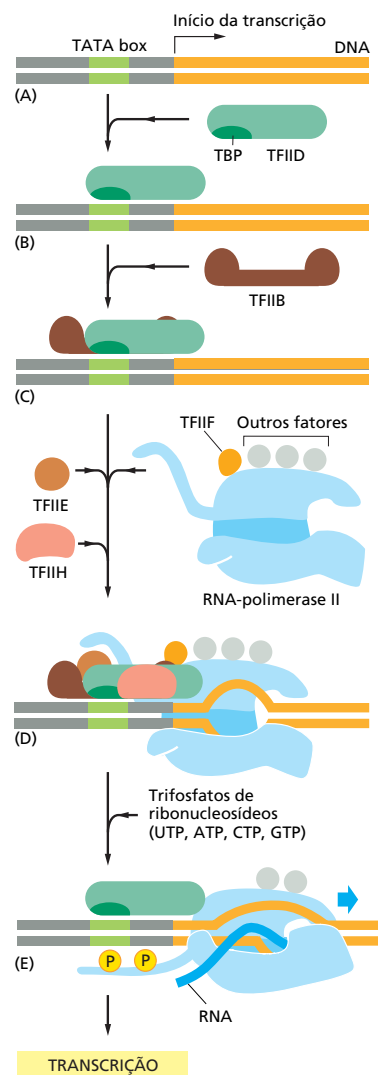
A descoberta inicial de que, ao contrário da RNA-polimerase bacteriana, uma RNA-polimerase II eucariótica purificada não podia, sozinha, iniciar a transcrição em um tubo de ensaio levou à descoberta e à purificação dos **fatores gerais de transcrição**. Essas proteínas acessórias organizam-se sobre o promotor, onde posicionam a RNA-polimerase, e separam a dupla-hélice do DNA para expor a fita molde, permitindo que a polimerase inicie a transcrição. Assim, os fatores gerais de transcrição têm um papel na transcrição eucariótica semelhante ao do fator sigma na transcrição bacteriana.

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. O proces-

Figura 7-12 Para dar início à transcrição, a RNA-polimerase II eucariótica necessita de um conjunto de fatores gerais de transcrição. Esses fatores de transcrição são denominados TFIIIB, TFIIID e assim por diante. (A) Diversos promotores eucariotos contêm uma sequência de DNA chamada de TATA box. (B) O TATA box é reconhecido por uma subunidade do fator geral de transcrição TFIIID, denominada proteína de ligação ao TATA (TBP). Por questões de simplificação, a distorção do DNA produzida pela ligação do TBP (ver Figura 7-13) não foi ilustrada. (C) A ligação do TFIIID permite a ligação adjacente do TFIIIB. Os demais fatores gerais de transcrição, assim como a própria RNA-polimerase, se associam ao promotor. (E) Em seguida, o TFIIH separa a dupla-hélice no ponto de iniciação da transcrição, utilizando a energia da hidrólise do ATP, o que expõe a fita molde do gene (não ilustrada). TFIIH também fosforila a RNA-polimerase II, liberando a maioria dos fatores gerais de transcrição, para que ela possa dar início à transcrição. O sítio de fosforilação consiste em uma longa "cauda" polipeptídica que se estende a partir da polimerase.

QUESTÃO 7-3

Poderia a RNA-polimerase usada na transcrição ser usada como polimerase para produzir o iniciador de RNA necessário para a replicação de DNA (discutida no Cap. 6)?



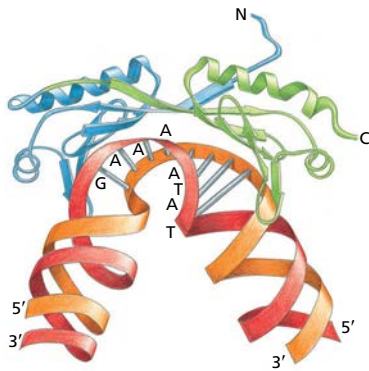


Figura 7-13 A proteína de ligação ao TATA (TBP) liga-se ao TATA box (indicado por letras) e flexiona a dupla-hélice do DNA. Essa distorção única do DNA causada pela TBP, que é uma subunidade do TFIID (ver Figura 7-12), ajuda a atrair os demais fatores gerais de transcrição. A TBP é uma cadeia polipeptídica única dobrada em dois domínios bastante similares (azul e verde). A proteína posiciona-se sobre a dupla-hélice do DNA como uma sela em um cavalo de rodeio (**Animação 7.4**). (Adaptada de J.L. Kim et al., *Nature* 365:520–527, 1993. Com permissão de Macmillan Publishers Ltd.)

so de montagem começa geralmente com a ligação do fator geral de transcrição TFIID a um segmento curto da dupla-hélice do DNA constituído principalmente por nucleotídeos T e A; devido à sua composição, essa porção do promotor é conhecida como *TATA box*. Por meio de sua ligação ao DNA, o TFIID provoca uma grande distorção local na dupla-hélice de 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 da transcrição. Uma vez que o TFIID tenha se ligado ao TATA box, os demais fatores se organizam, juntamente com a RNA-polimerase II, para formar um *complexo de iniciação de transcrição* completo. Embora a Figura 7-12 ilustre os fatores gerais de transcrição associando-se sobre o promotor em uma ordem determinada, a sequência exata de montagem é provavelmente diferente em distintos promotores.

Depois de a RNA-polimerase II ter sido posicionada no promotor, ela deve ser liberada do complexo de fatores gerais de transcrição para começar a sua tarefa de síntese de uma molécula de RNA. Um passo essencial na liberação da RNA-polimerase é a adição de grupos fosfato à sua “cauda” (ver Figura 7-12E). Essa liberação é iniciada pelo fator geral de transcrição TFIIF, que contém uma de suas subunidades com atividade de proteína-quinase. Uma vez que a transcrição tenha começado, muitos dos fatores gerais de transcrição irão se dissociar do DNA, tornando-se, em seguida, disponíveis para iniciar outro ciclo de transcrição com uma nova molécula de RNA-polimerase. Quando a RNA-polimerase II termina a transcrição de um gene, ela é também liberada do DNA; os fosfatos na sua cauda são removidos por proteínas-fosfatase, e a polimerase está disponível para buscar um novo promotor. Apenas a forma desfosforilada da RNA-polimerase II é capaz de dar início à síntese de RNA.

Os mRNAs eucarióticos são processados no núcleo

Apesar de todos os organismos utilizarem o mesmo princípio de molde de DNA para a transcrição do RNA, o modo segundo o qual os transcritos são manipulados antes de poderem ser utilizados pela célula para a síntese proteica difere bastante entre bactérias e eucariotos. O DNA bacteriano é diretamente exposto no citoplasma, onde se localizam os *ribossomos* nos quais a síntese proteica ocorre. Em uma bactéria, assim que uma molécula de mRNA começa a ser sintetizada, os ribossomos imediatamente se ligam à extremidade 5' livre do transcrito de RNA e começam a tradução da proteína.

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 em proteína, ele deverá ser transportado para fora do núcleo por pequenos poros existentes no envelope nuclear (**Figura 7-14**). Antes que possa ser exportado para o citoplasma, no entanto, um RNA eucariótico deve passar por várias etapas de **processamento do RNA**, que incluem o *capeamento*, o *splicing* (*encadeamento*) e a *poliadenilação*, como discutimos a seguir. Essas etapas acontecem enquanto o RNA está sendo sintetizado. As enzimas responsáveis pelo processamento do RNA são transportadas sobre a cauda fosforilada da RNA-polimerase II eucariótica conforme ela sintetiza uma molécula de RNA (ver Figura 7-12), e processam o transcrito à medida que ele emerge da polimerase (**Figura 7-15**).

Diferentes tipos de RNA são processados de diferentes formas antes de sair do núcleo. Duas etapas do processamento, o capeamento e a poliadenilação, ocorrem somente em transcritos de RNA destinados a se tornarem moléculas de mRNA (denominadas *mRNAs precursores*, ou *pré-mRNAs*).

1. O **capeamento do RNA** modifica a extremidade 5' do transcrito de RNA, que é a primeira a ser sintetizada. O RNA é capeado pela adição de um nucleotídeo atípico – um nucleotídeo guanina (G) contendo um grupo metila, que

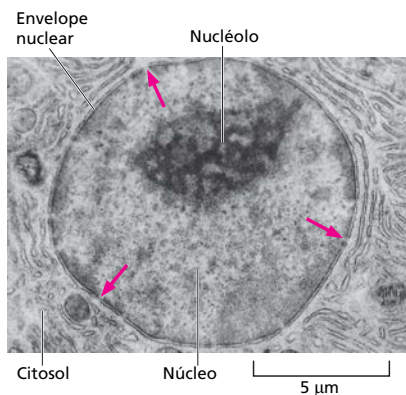


Figura 7-14 Antes de serem traduzidas, as moléculas de mRNA sintetizadas no núcleo devem ser exportadas para o citoplasma pelos poros no envelope nuclear (*setas vermelhas*). Aqui está ilustrada uma seção do núcleo de um hepatócito. O nucléolo é o local da síntese dos RNAs ribossômicos e também o local onde esses são combinados com proteínas para formar os ribossomos, que posteriormente são exportados para o citoplasma. (De D.W. Fawcett, *A Textbook of Histology*, 11th ed. Philadelphia: Saunders, 1986. Com permissão de Elsevier.)

é ligado à extremidade 5' do RNA de uma forma não habitual (**Figura 7-16**). Esse capeamento ocorre após a RNA-polimerase II ter sintetizado cerca de 25 nucleotídeos do RNA, muito antes de completar a transcrição de todo o gene.

2. A **poliadenilação** insere uma estrutura especial na extremidade 3' do mRNA recentemente transcrito. Em contraste com as bactérias, em que a extremidade 3' de um mRNA é simplesmente o final da cadeia sintetizada pela RNA-polimerase, a extremidade 3' de um mRNA eucarioto é inicialmente clivada por uma enzima que corta a cadeia de RNA em uma sequência determinada de nucleotídeos. O transcrito é então modificado por uma segunda enzima que adiciona uma série de repetições de nucleotídeos adenina (A) à extremidade cortada. Essa *cauda poli-A* geralmente possui um comprimento de algumas centenas de nucleotídeos (ver Figura 7-16A).

Estas duas modificações – capeamento e poliadenilação – aumentam a estabilidade de uma molécula de mRNA eucariótico, facilitando a sua exportação do núcleo para o citoplasma, e geralmente identificam a molécula de RNA como um 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.

Em eucariotos, genes codificadores de proteínas são interrompidos por sequências não codificadoras denominadas íntrons

A maioria dos pré-mRNAs eucarióticos passa por uma etapa adicional de processamento antes de se tornar mRNA funcional. Essa etapa envolve uma modificação muito mais radical do transcrito de pré-mRNA do que simplesmente o capeamento e a poliadenilação, e é a consequência de uma característica surpreendente da maioria dos genes eucarióticos. Em bactérias, a maioria das proteínas é codificada por um segmento ininterrupto da sequência de DNA que é transcrito para um mRNA e que, sem qualquer processamento adicional, pode ser traduzido em proteína. A maioria dos genes eucarióticos que codificam proteínas, ao contrário, tem suas sequências codificadoras interrompidas por *sequências intervenientes* longas e não codificadoras, chamadas de **íntrons** (do inglês, *intervening sequences*). As porções codificadoras dispersas – chamadas de *sequências expressas* ou **éxons** (do inglês, *expressed sequences*) – são geralmente mais curtas do que os íntrons, e muitas vezes representam apenas uma pequena fração do comprimento total do gene (**Figura 7-17**). Os íntrons variam em comprimento de um único nucleotídeo a mais de 10.000 nucleotídeos. Alguns genes eucarióticos que codificam proteínas não possuem íntrons, e alguns têm apenas poucos íntrons; mas a maioria desses genes possui numerosos íntrons (**Figura 7-18**). Observe que os termos “éxon” e “íntron” aplicam-se tanto ao DNA quanto às sequências correspondentes no RNA.

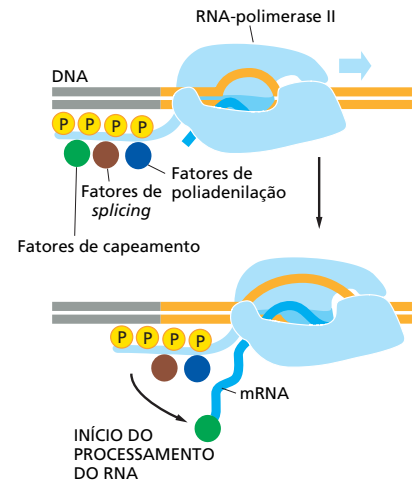


Figura 7-15 A fosforilação da cauda da RNA-polimerase II permite o arranjo das proteínas de processamento do RNA. Observe que os fosfatos aqui ilustrados são suplementares aos necessários para a iniciação da transcrição (ver Figura 7-12). Capeamento, poliadenilação e *splicing* são modificações que ocorrem durante o processamento do RNA no núcleo.

Figura 7-16 As moléculas do pré-mRNA eucariótico são modificadas pelo capeamento e pela poliadenilação. (A) Um mRNA eucariótico possui um quepe na extremidade 5' e uma cauda poli-A na extremidade 3'. Lembre-se de que nem todos os transcritos de RNA codificam proteínas. (B) A estrutura do quepe 5'. Diversos quepes de mRNAs 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

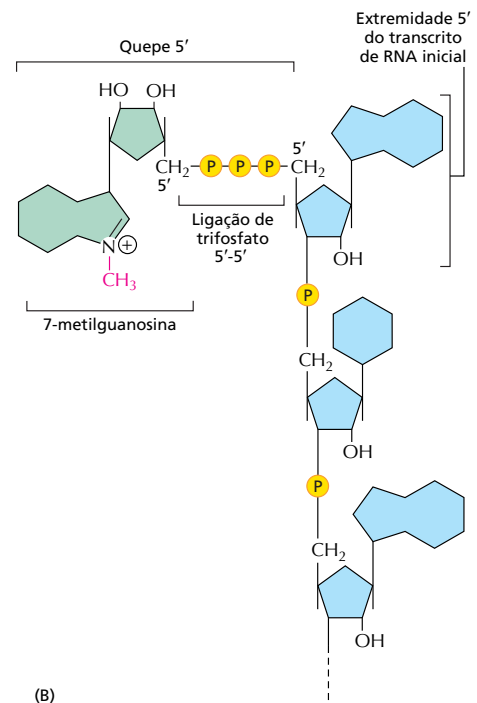
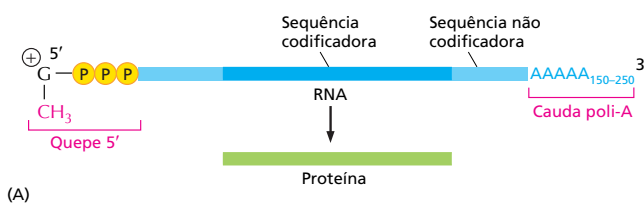
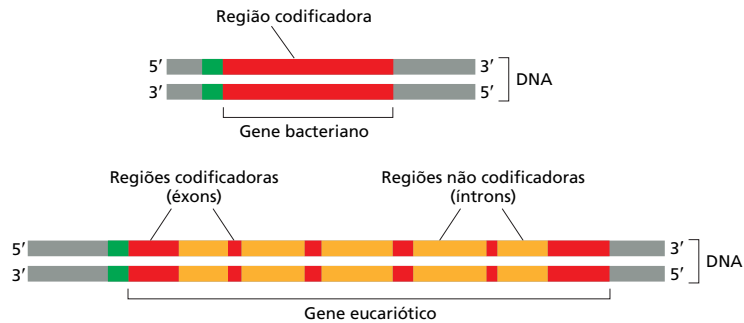


Figura 7-17 Genes eucarióticos e bacterianos são organizados de forma distinta. Um gene bacteriano consiste em um único segmento de sequência nucleotídica não interrompido que codifica a sequência de aminoácidos de uma proteína (ou mais de uma). Em contraste, as sequências codificadoras de proteínas da maioria dos genes eucarióticos (éxons) são interrompidas por sequências não codificadoras (íntrons). Os promotores de transcrição estão indicados em verde.



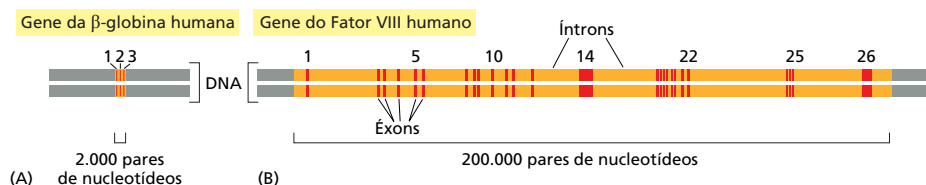
Os íntrons são removidos de pré-mRNAs pelo *splicing* do RNA

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 II continua a transcrever o gene, tem início o processo de *splicing* (ou **encadeamento**) do RNA, durante o qual os íntrons são removidos do RNA recém-sintetizado e seus éxons são unidos. 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 agora uma molécula funcional que pode então deixar o núcleo e ser traduzida em proteína.

Como a célula determina quais segmentos do transcrito de RNA serão removidos durante o *splicing*? Diferentemente da sequência codificadora de um éxon, a maior parte da sequência nucleotídica de um íntron parece não ser importante. Apesar de, em geral, existir pouca semelhança entre as sequências nucleotídicas de diferentes íntrons, cada íntron contém poucas sequências nucleotídicas curtas essenciais que direcionam sua remoção do pré-mRNA. Essas sequências especiais 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 nucleotídeo “A” salientado em vermelho nas Figuras 7-19 e 7-20.

Não vamos descrever a maquinaria do *splicing* em detalhes, mas vale a pena notar que, ao contrário das outras etapas de produção do mRNA que discutimos, o *splicing* do RNA é realizado em grande parte por moléculas de RNA, em vez de proteínas. Essas moléculas de RNA, chamadas de **pequenos RNAs nucleares (snRNAs)**, estão unidas a proteínas adicionais para formar *pequenas ribonucleoproteínas nucleares (snRNPs)*. As snRNPs (cuja pronúncia é “snurps”) reconhecem sequências de sítios de *splicing* pelo pareamento de bases complementares entre os seus componentes de RNA e as sequências no pré-mRNA, e também participam intimamente na química do *splicing* (Figura 7-21). Juntas, essas snRNPs formam a região central do **spliceossomo** (ou **encadeossomo**), o grande arranjo

Figura 7-18 A maioria dos genes humanos codificadores de proteína é dividida em múltiplos éxons e íntrons. (A) O gene da β -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 do distúrbio sanguíneo hemofilia.



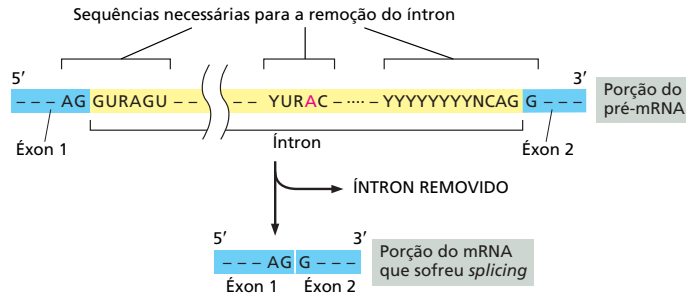


Figura 7-19 Sequências especiais de nucleotídeos em um transcrito de pré-mRNA sinalizam o início e o final de um íntron. Apenas as sequências de nucleotídeos apresentadas são necessárias para a remoção de um íntron; as demais posições em um íntron podem ser ocupadas por quaisquer nucleotídeos. As sequências especiais são reconhecidas principalmente por pequenas ribonucleoproteínas nucleares (snRNPs), que dirigem a clivagem do RNA nos limites éxon-íntron e catalisam a ligação covalente das sequências dos éxons. Aqui, além dos símbolos padrão dos nucleotídeos (A, C, G, U), R significa A ou G; Y significa C ou U; e N representa qualquer nucleotídeo. O A ilustrado em vermelho forma o ponto de forquilha do laço produzido na reação de *splicing* mostrada na 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 (ver Figura 7-20). As sequências de *splicing* ilustradas se referem a humanos; sequências similares direcionam o *splicing* do RNA em outros organismos eucarióticos.

de RNA e de moléculas proteicas que realiza o *splicing* no núcleo. Para visualizar o spliceossomo em ação, veja a **Animação 7.5**.

Esse tipo de arranjo do gene em eucariotos envolvendo íntrons e éxons pode, à primeira vista, parecer um processo dispendioso e desnecessário. No entanto, ele contém uma série de importantes benefícios. 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 7-22**). Acredita-se que cerca de 95% dos genes humanos sofram *splicing* alternativo. Dessa forma, o *splicing* do RNA permite que os eucariotos elevem astronômicamente 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 discutimos em detalhes no Capítulo 9, acredita-se que a estrutura íntron-éxon dos genes tenha acelerado o surgimento de proteínas novas e úteis: novas proteínas parecem ter surgido pela mistura e recombinação de diferentes éxons de genes preexistentes, analogamente à montagem de um novo tipo de máquina a partir de um *kit* de componentes funcionais preexistentes. Efetivamente, várias proteínas presentes nas células atuais se assemelham a uma colcha de retalhos, composta a partir de um conjunto padrão de peças proteicas, denominadas *domínios proteicos* (ver Figura 4-51).

Os mRNAs eucarióticos maduros são exportados do núcleo

Vimos como a síntese e o processamento do pré-mRNA ocorrem de forma ordenada dentro do núcleo da célula. No entanto, esses eventos criam um problema específico para as células eucarióticas: do número total de transcritos do pré-mRNA que são sintetizados, apenas uma pequena fração – o mRNA maduro – será útil para a célula. Os fragmentos remanescentes de RNA – íntrons excisados, RNAs quebrados e transcritos erroneamente unidos – não são apenas inúteis, mas podem ser perigosos para a célula se forem autorizados a sair do núcleo.

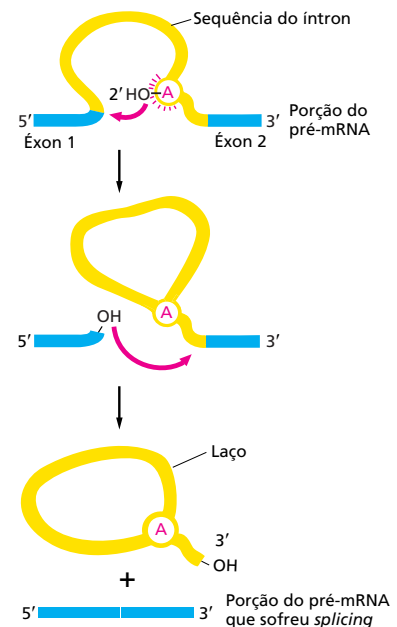
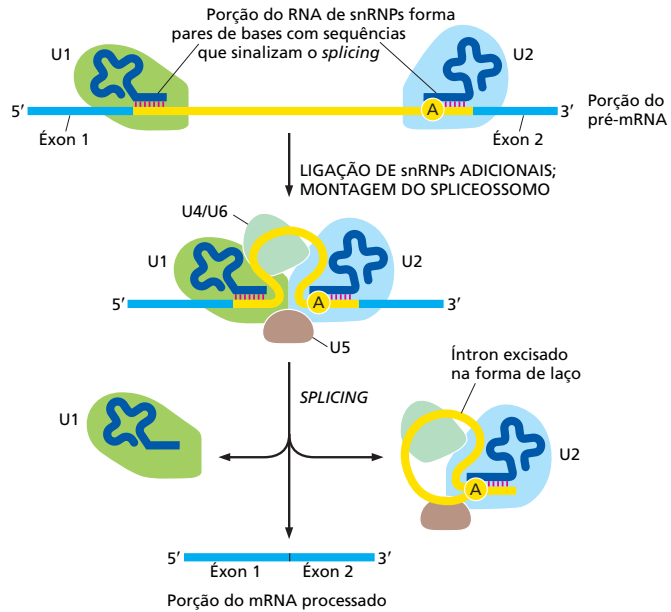


Figura 7-20 Um íntron em uma molécula de pré-mRNA forma uma estrutura ramificada durante o *splicing* do RNA.

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 cadeia principal de açúcar-fosfato do RNA nesse ponto (essa é a mesma adenina que está salientada em vermelho na Figura 7-19). Neste processo, a extremidade 5' cortada do intron é covalentemente ligada ao grupo 2'-OH da ribose do nucleotídeo A para formar uma estrutura em forquilha. 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 no núcleo.

Figura 7-21 O *splicing* é realizado por uma série de complexos RNA-proteína chamados de snRNPs. Há cinco snRNPs, denominadas U1, U2, U4, U5 e U6. Como mostrado aqui, U1 e U2 ligam-se ao sítio 5' de *splicing* (U1) e ao ponto de forquilha do laço (U2) por meio de pareamento por complementaridade de bases. Outras snRNPs são atraídas para o sítio de *splicing*, e interações entre seus componentes proteicos dirigem a montagem completa do spliceossomo. A seguir, rearranjos dos pares de bases que unem as snRNPs e o transcrito de RNA reorganizam o spliceossomo para formar o sítio ativo que excisa o íntron, dando origem ao mRNA processado (ver também Figura 7-20).



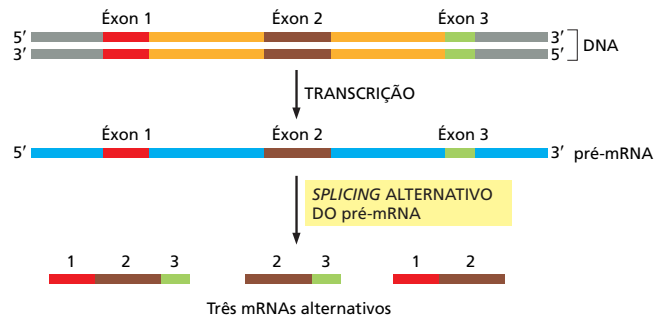
Como, então, a célula distingue entre as moléculas de mRNA maduras relativamente raras que devem ser exportadas para o citosol e a enorme quantidade de detritos gerados pelo processamento do RNA?

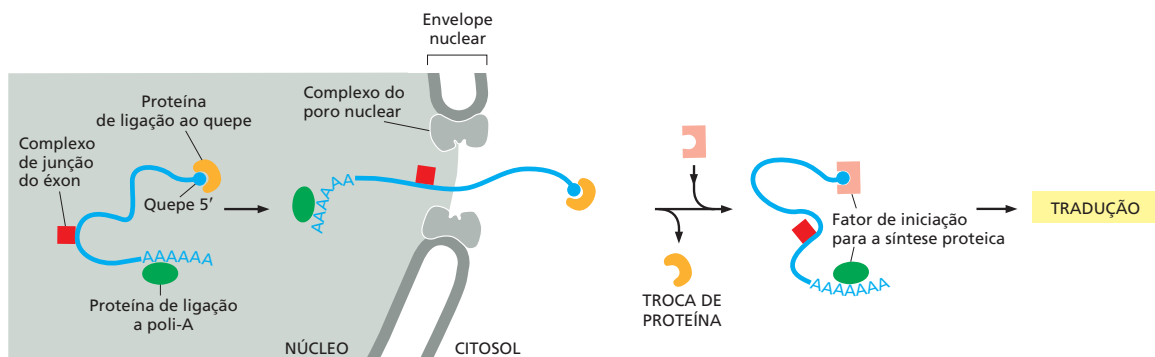
A resposta é que o transporte de mRNA do núcleo para o citoplasma, onde os mRNAs são traduzidos em proteína, é altamente seletivo: apenas mRNAs processados corretamente são exportados. Esse transporte seletivo é mediado por *complexos do poro nuclear*, que conectam o nucleoplasma ao citosol e agem como portões que controlam as macromoléculas que podem entrar ou sair do núcleo (discutido no Cap. 15). Para estar “pronta para exportação”, uma molécula de mRNA deve estar ligada a um conjunto apropriado de proteínas, cada qual capaz de reconhecer partes diferentes de uma molécula madura de mRNA. Essas proteínas incluem proteínas de ligação à poli-A, um complexo de ligação ao quepe, e proteínas que se ligam a mRNAs que tenham sofrido *splicing* correto (Figura 7-23). É o conjunto completo de proteínas ligadas, e não uma única proteína, que, em última instância, determina se uma molécula de mRNA deixará o núcleo. Os “resíduos de RNAs” que ficam para trás no núcleo são degradados, e seus “blocos construtores”, os nucleotídeos, são reutilizados para a transcrição.

As moléculas de mRNA são finalmente degradadas no citosol

Visto que uma única molécula de mRNA pode ser traduzida diversas vezes, gerando várias cópias da proteína (ver Figura 7-2), o intervalo de tempo que uma molécula madura de mRNA permanece na célula afeta a quantidade de proteína produzida.

Figura 7-22 Alguns pré-mRNAs sofrem *splicing* alternativo do RNA para produzir vários mRNAs e proteínas a partir do mesmo gene. Apesar de todos os éxons estarem presentes em um pré-mRNA, alguns podem ser excluídos da molécula de mRNA final. Neste exemplo, três de quatro possíveis mRNAs são produzidos. Os quepes 5' e as caudas poli-A dos mRNAs não estão ilustrados.





Cada molécula de mRNA, finalmente, é degradada em nucleotídeos por ribonucleases (RNases) presentes no citosol, mas o tempo de vida médio de moléculas de mRNA difere consideravelmente, dependendo da sequência de nucleotídeos do mRNA e do tipo de célula. Em bactérias, a maioria dos mRNAs é rapidamente degradada, apresentando um tempo de vida médio típico de cerca de 3 minutos. Os mRNAs em células eucarióticas normalmente persistem por mais tempo: alguns, como os que codificam a β -globina, têm tempo de vida médio de mais de 10 horas, ao passo que outros têm tempo de vida médio inferior a 30 minutos.

Esses tempos de vida média distintos são, em parte, controlados por sequências de nucleotídeos presentes no próprio mRNA, frequentemente localizadas na porção do RNA denominada *região 3' não traduzida*, que se situa entre a extremidade 3' da sequência codificadora e a cauda poli-A. Os diferentes tempos de vida média de mRNAs ajudam a célula a controlar a quantidade de cada proteína que ela sintetiza. Em geral, as proteínas produzidas em grande quantidade, como a β -globina, são traduzidas a partir de mRNAs que apresentam tempos de vida média longos, enquanto as proteínas feitas em quantidades menores, ou cujos níveis devem ser rapidamente alterados em resposta a sinais específicos, são em geral sintetizadas a partir de mRNAs de curta duração.

As primeiras células devem ter possuído íntrons em seus genes

O processo de transcrição é universal: todas as células usam RNA-polimerase e o 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 gerais de suas estruturas, e certamente evoluíram a partir de uma polimerase ancestral comum. Portanto, pode ter parecido desafiador explicar por que os transcritos de RNA resultantes são tratados de maneira tão diferente em eucariotos e em procariotos (Figura 7-24). Em particular, o *splicing* do RNA parece marcar uma diferença fundamental entre esses dois tipos de células. Mas, afinal, como teve origem essa diferença tão marcante?

Como vimos, o *splicing* do RNA proporciona aos eucariotos a capacidade de produzir uma ampla variedade de proteínas a partir de um único gene. Isso também lhes permite desenvolver novos genes por meio da mistura e recombinação de éxons de genes preexistentes, como discutimos no Capítulo 9. No entanto, essas vantagens vêm acompanhadas de um custo: a célula deve manter um genoma maior e descartar uma grande fração do RNA por ela sintetizado sem tê-lo usado. De acordo com uma linha de pensamento, as 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 rapidamente e de maneira eficiente. Corroborando essa ideia, eucariotos simples que se reproduzem rapidamente (p. ex., algumas

Figura 7-23 Um conjunto especializado de proteínas de ligação ao RNA sinaliza que o mRNA maduro está pronto para ser exportado para o citosol. Como indicado à esquerda, o quepe e a cauda poli-A de uma molécula madura de mRNA estão “marcados” por proteínas que reconhecem essas modificações. Além disso, um grupo de proteínas chamado de complexo de junção do éxon é depositado sobre o pré-mRNA após a ocorrência de cada *splicing* bem-sucedido. Quando o mRNA é considerado “pronto para exportação”, um receptor de transporte nuclear (discutido no Cap.15) se associa ao mRNA, guiando-o pelo poro nuclear. No citosol, o mRNA pode perder algumas dessas proteínas e ligar-se a novas, que, com a proteína de ligação à poli-A, atuam como fatores de iniciação para a síntese de proteínas, como discutimos adiante.

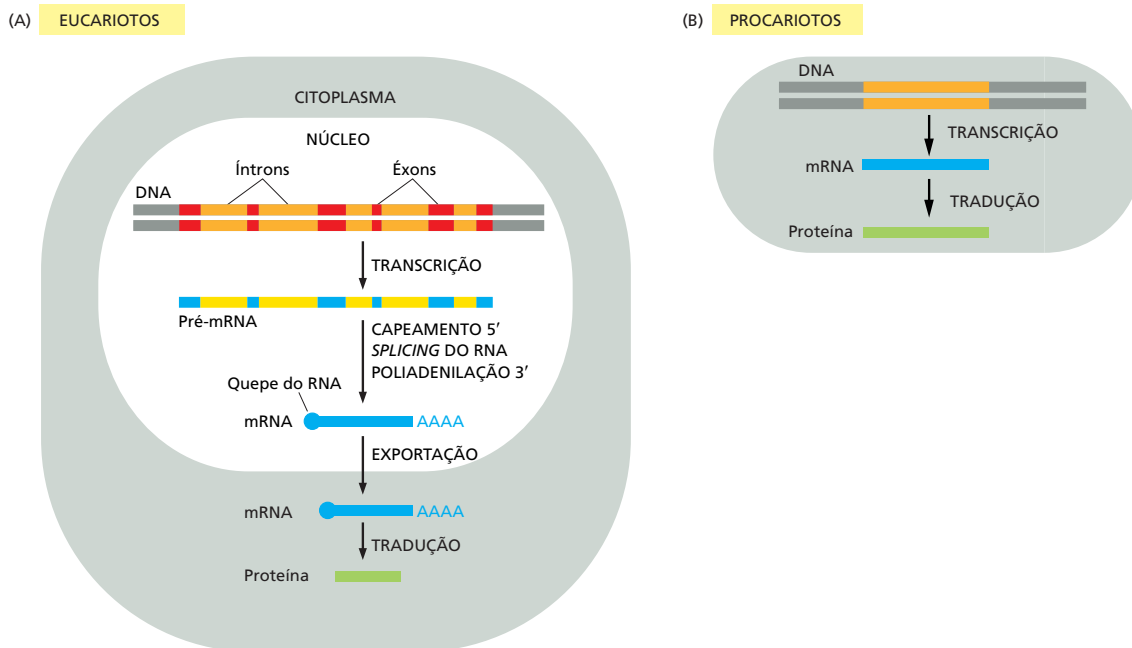


Figura 7-24 Procariotos e eucariotos “manipulam” diferentemente seus transcritos de RNA. (A) Em células eucarióticas, a molécula de pré-RNA, produzida pela transcrição, contém tanto seqüências de íntrons quanto de éxons. Suas duas extremidades sofrem modificações, e os íntrons são removidos pelo *splicing* do RNA. A seguir, o mRNA resultante é transportado do núcleo para o citoplasma, onde é traduzido em proteína. Embora esses passos estejam representados como ocorrendo em seqüência, um de cada vez, na realidade eles ocorrem simultaneamente. Por exemplo, a adição do quepe do RNA costuma ocorrer antes de a transcrição estar concluída. O mesmo se dá com o início do *splicing* do RNA. Devido a essa sobreposição, transcritos inteiros do gene (incluindo todos os íntrons e éxons) em geral não são encontrados na célula. (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. Tanto em procariotos quanto em eucariotos, a quantidade de uma proteína em uma célula depende das taxas de cada uma dessas etapas, bem como das taxas de degradação do mRNA e das moléculas de proteína.

leveduras) possuem relativamente poucos íntrons, e esses íntrons são em geral muito menores do que os encontrados em eucariotos superiores.

Por outro lado, alguns argumentam que os íntrons se originaram a partir de elementos genéticos móveis parasitas (discutidos no Cap. 9) que invadiram um ancestral eucariótico primordial e colonizaram seu genoma. Essas células hospedeiras replicaram involuntariamente as seqüências “clandestinas” de nucleotídeos juntamente com o seu próprio DNA, e os eucariotos modernos simplesmente não se preocupam em eliminar a desordem genética deixada por essas infecções antigas. 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 ainda 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 seqüência linear de nucleotídeos em uma molécula de RNA era traduzida para a forma de uma seqüência linear de um conjunto de subunidades quimicamente tão distintas – os aminoácidos – em uma proteína? Essa fascinante questão intrigava os cientistas da época. Ali estava 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 seqüência de mRNA é decodificada em grupos de três nucleotídeos

A transcrição como forma de transferência de informação é simples de compreender: 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 para-

Códons	AGA	AGG	GCA	GCC	GCG	GCU	ACA	ACG	ACC	ACU	UGG	UUA	UUG	CUA	CUC	CUG	CUU	AAA	AAG	AUG	UUU	UUC	CUA	UCA	UCC	UCG	CCU	CCA	CCC	CCG	CCU	ACA	ACC	ACG	ACU	UAC	UAU	GUA	GUC	GUG	GUU	UAA	UAG	UGA	
Amino-ácidos	Ala	Arg	Asp	Asn	Cys	Glu	Gln	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val	Terminação																								
Amino-ácidos	A	R	D	N	C	E	Q	G	H	I	L	K	M	F	P	S	T	W	Y	V																									

Figura 7-25 A sequência nucleotídica de um mRNA é traduzida para a sequência de aminoácidos de uma proteína pelo uso de um código genético. Todos os códons de três nucleotídeos em mRNAs que determinam um dado aminoácido estão listados sobre esse aminoácido, o qual está representado tanto pela abreviação de uma letra quanto pela abreviação em sigla de três letras (ver Painel 2-5, p. 74-75 para o nome completo de cada aminoácido e a sua estrutura). Como as moléculas de RNA, os códons são sempre escritos com o nucleotídeo 5'-terminal à esquerda. Observe que a maioria dos aminoácidos é representada por mais de um códon e que existem algumas regularidades no conjunto de códons que especificam cada aminoácido. Códons para o mesmo aminoácido contêm, em geral, os mesmos nucleotídeos na primeira e na segunda posições, e podem variar em sua terceira posição. Existem três códons que não especificam aminoácidos, mas atuam como sítios de terminação (*códons de terminação*), sinalizando o final da sequência codificadora de proteína em um mRNA. Um códon – AUG – age tanto como códon de iniciação, sinalizando o início de uma mensagem que codifica uma proteína, quanto como códon que especifica o aminoácido metionina.

mento 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 diferentes. Tendo em vista que existem 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 pelas quais a sequência de nucleotídeos de um gene, passando por uma molécula intermediária de mRNA, é traduzida na sequência de aminoácidos de uma proteína são conhecidas como o **código genético**.

Em 1961, foi descoberto que a sequência de nucleotídeos em uma molécula de mRNA é lida, consecutivamente, em grupos de três. Visto que o RNA é 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 diferentes são geralmente encontrados em proteínas. Dessa forma, ou alguns tripletes de nucleotídeos nunca são usados, ou o código é redundante, com alguns aminoácidos sendo especificados por mais de um triplete. A segunda possibilidade mostrou-se correta, conforme mostrado no código genético completamente decifrado apresentado na **Figura 7-25**. Cada grupo de três nucleotídeos consecutivos sobre o RNA é denominado **códon**, e cada um desses códons especifica um aminoácido. A estratégia utilizada para decifrar esse código está descrita em **Como Sabemos**, p. 240-241.

O mesmo código genético é usado por quase todos os organismos da atualidade. Apesar de algumas diferenças terem sido descritas, estas ocorrem principalmente no mRNA mitocondrial e em alguns fungos e protozoários. As mitocôndrias têm as suas próprias maquinarias de replicação de DNA, de transcrição e de síntese proteica, que operam independentemente das maquinarias correspondentes do restante da célula (discutido no Cap. 14), e essas organelas foram capazes de acomodar pequenas alterações àquele que é um código genético praticamente universal. Mesmo no caso de fungos e protozoários, as similaridades do código são muito superiores às poucas diferenças.

Em princípio, uma sequência de mRNA pode ser traduzida em qualquer uma de três diferentes **fases de leitura**, dependendo do ponto de início do processo de decodificação (**Figura 7-26**). No entanto, apenas uma das três possíveis fases de leitura sobre um mRNA codifica a proteína correta. Discutimos adiante como um sinal de pontuação especial, no início de cada molécula de mRNA, determina a fase de leitura correta.

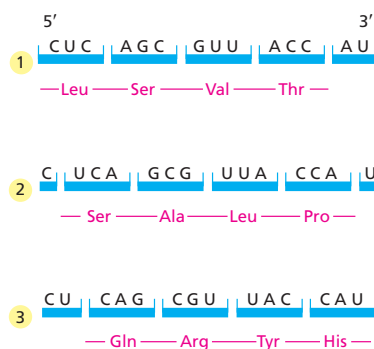


Figura 7-26 Uma molécula de mRNA pode ser traduzida em três fases de leitura diferentes. No processo de tradução de uma sequência nucleotídica (azul) em uma sequência de aminoácidos (vermelho), a sequência de nucleotídeos na molécula de mRNA é lida da extremidade 5' para a 3' em grupos consecutivos de três nucleotídeos. Em princípio, portanto, a mesma sequência de mRNA pode determinar três sequências de aminoácidos completamente diferentes, dependendo do local onde começa a tradução, isto é, da fase de leitura utilizada. Na realidade, porém, apenas uma dessas fases de leitura codifica a mensagem real e, portanto, é usada na tradução, como discutimos adiante.

DECIFRANDO O CÓDIGO GENÉTICO

No início da década de 1960, o *dogma central* havia sido aceito como representativo da via pela qual a informação fluía do gene para a proteína. Estava claro que os genes codificavam as proteínas, que os genes eram feitos de DNA, e que o mRNA atuava como um intermediário, levando a informação do DNA para o ribossomo, no qual o RNA era traduzido em proteína.

Até mesmo o formato geral do código genético estava compreendido: cada um dos 20 aminoácidos encontrados nas proteínas é representado por um códon (tripleto) em uma molécula de mRNA. No entanto, um desafio ainda maior permanecia: biólogos, químicos e mesmo físicos concentravam seus esforços tentando decifrar o código – tentando desvendar qual aminoácido era codificado por cada um dos 64 possíveis tripletos 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. Contudo, 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 moléculas simples de RNA. Se pudessem direcionar essas moléculas de RNA 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 tripletos que correspondiam aos aminoácidos.

Abandonando as células

Antes que os pesquisadores pudessem testar seus mRNAs sintéticos, eles precisariam aperfeiçoar um sistema livre de células para a síntese proteica. 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 de que necessitavam para um dado sistema de tradução livre de células, os pesquisadores romperam 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 polipeptídeos radiomarcados. Por meio de uma nova centrifugação desse sobrenadante, sob uma velocidade maior, era possível forçar a deposição dos ribossomos, e dos peptídeos recém-sintetizados a eles conectados, no fundo do tubo; os polipeptídeos marcados podiam então ser detectados, medindo-se a radioatividade remanescente no sedimento do tubo após descarte da fase aquosa superior.

O problema com esse sistema específico era que ele produzia proteínas codificadas por mRNAs próprios da célula, já presentes no extrato, e os pesquisadores queriam usar suas próprias mensagens sintéticas para dirigir a síntese de proteínas. Esse problema foi resolvido quando Marshall Nirenberg descobriu que poderia destruir o RNA celular presente no extrato pela adição de uma pequena quantidade de ribonuclease – uma enzima que degrada o RNA. Agora tudo o que ele precisava fazer era preparar grandes quantidades de mRNA sintético, adicioná-las ao sistema livre de células, e analisar os peptídeos resultantes.

Falsificando a mensagem

A produção de polinucleotídeos com uma sequência definida não foi tão simples como se pretendia. Mais uma vez, seriam necessários anos até que químicos e bioengenheiros desenvolvessem máquinas capazes de sintetizar uma determinada sequência de ácidos nucleicos de forma rápida e barata. 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 poderia ser sintetizada em uma sequência aleatória; mas um único tipo de nucleotídeo produziria um polímero homogêneo, contendo apenas um nucleotídeo. Assim, Nirenberg, trabalhando com seu colaborador Heinrich Matthaei, produziu inicialmente mRNAs sintéticos feitos inteiramente de uracila – poli-U.

Juntos, esses pesquisadores colocaram esse poli-U sobre o sistema de tradução livre de células. Eles então adicionaram aminoácidos radioativamente marcados sobre a mistura. Depois de testar cada aminoácido – um de cada vez, em 20 diferentes experimentos – eles determinaram que poli-U dirige a síntese de um polipeptídeo contendo apenas fenilalanina (**Figura 7-27**). Com esse resultado eletrizante, a primeira palavra do código genético foi decifrada (ver Figura 7-25).

Nirenberg e Matthaei repetiram esse experimento com poli-A e poli-C, e determinaram que AAA codificava lisina e CCC codificava prolina. O significado de poli-G não pôde ser determinado por tal método, porque esse polinucleotídeo forma uma hélice de cadeia tripla incomum, que não atua como molde em um sistema livre de células.

Alimentar ribossomos com RNAs sintéticos parecia uma técnica bastante interessante. Mas, com as possibilidades de nucleotídeos únicos esgotadas, os pesquisadores haviam decifrado apenas três códons; e eles ainda tinham outros 61 para descobrir. Os outros códons, no entanto, eram mais difíceis de decifrar, e foi necessária uma nova abordagem sintética. 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 soube do trabalho de Nirenberg com RNAs sintéticos, Khorana direcionou suas energias e ha-

As moléculas de tRNA conectam os aminoácidos e os códons no mRNA

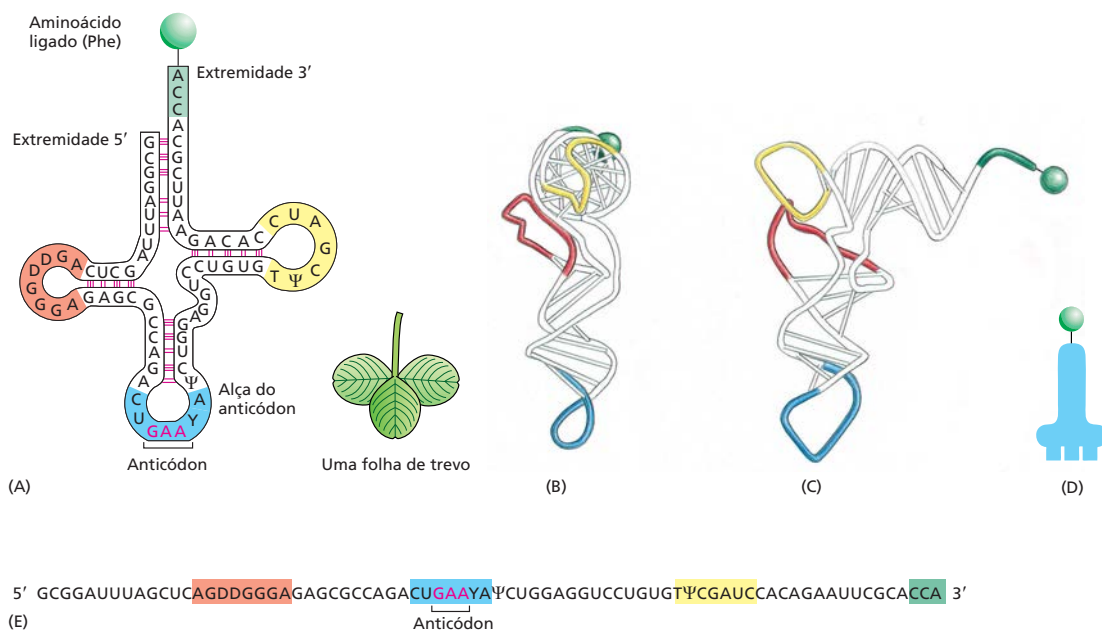
Os códons de uma molécula de mRNA não reconhecem diretamente os aminoácidos por eles codificados: o grupo de três nucleotídeos não se liga diretamente ao aminoácido, por exemplo. Em vez disso, a tradução do mRNA em proteína depende de moléculas adaptadoras que podem reconhecer e ligar-se ao códon, por um sítio sobre sua superfície, e ao aminoácido, por um outro sítio. Esses adaptadores consistem em um conjunto de pequenas moléculas de RNA conhecidas como **RNAs transportadores (tRNAs)**, cada uma com aproximadamente 80 nucleotídeos de comprimento.

Vimos antes que uma molécula de RNA costuma se dobrar em uma estrutura tridimensional por intermédio do pareamento de bases entre diferentes regiões da molécula. Se as regiões de pareamento de bases forem suficientemente extensas, promoverão o dobramento da molécula e a formação de uma estrutura de dupla-hélice, semelhante à dupla fita do DNA. As moléculas de tRNA fornecem o exemplo mais impressionante desse fenômeno. Quatro pequenos segmentos de tRNA adquirem estrutura de dupla-hélice, produzindo uma molécula que se assemelha a uma folha de trevo quando desenhada esquematicamente (**Figura 7-29A**). Por exemplo, uma sequência 5'-GCUC-3' em uma parte de uma cadeia polinucleotídica pode formar pares com uma sequência 5'-GAGC-3' presente em uma outra região dessa mesma molécula. A folha de trevo sofre outros dobramentos, originando uma estrutura compacta em forma de L que se mantém por ligações de hidrogênio adicionais entre as diferentes regiões da molécula (**Figura 7-29B e C**).

Dois regiões nucleotídicas não pareadas, situadas cada uma em uma das extremidades da molécula de tRNA estruturada em L, são essenciais para o funcionamento dos tRNAs 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 covalentemente ao tRNA.

Vimos, na seção anterior, que o código genético é redundante; ou seja, vários códons diferentes podem determinar um mesmo aminoácido (ver **Figura 7-25**). Essa redundância implica ou que exista mais de um tRNA para muitos dos ami-

Figura 7-29 Moléculas de tRNA são adaptadores moleculares, que conectam os aminoácidos aos códons. Nesta 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 convencional em “folha de trevo” mostra o pareamento por complementaridade de bases (*linhas vermelhas*) que cria as regiões em dupla-hélice da molécula. A alça do anticódon (*azul*) contém a sequência de três nucleotídeos (*letras vermelhas*) que forma pares de bases com um códon no mRNA. O aminoácido correspondente ao par códon-anticódon está ligado à extremidade 3' do tRNA. Os tRNAs contêm algumas bases incomuns, as quais são produzidas por alterações químicas após a síntese do tRNA. As bases identificadas como Ψ (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 raios X. Estas duas imagens estão posicionadas em ângulo de 90° uma em relação à outra. (D) Representação esquemática do tRNA, enfatizando o anticódon, que será utilizada nas figuras subsequentes. (E) A sequência linear de nucleotídeos da molécula de tRNA, com as regiões no mesmo código de cores usado em A, B e C.



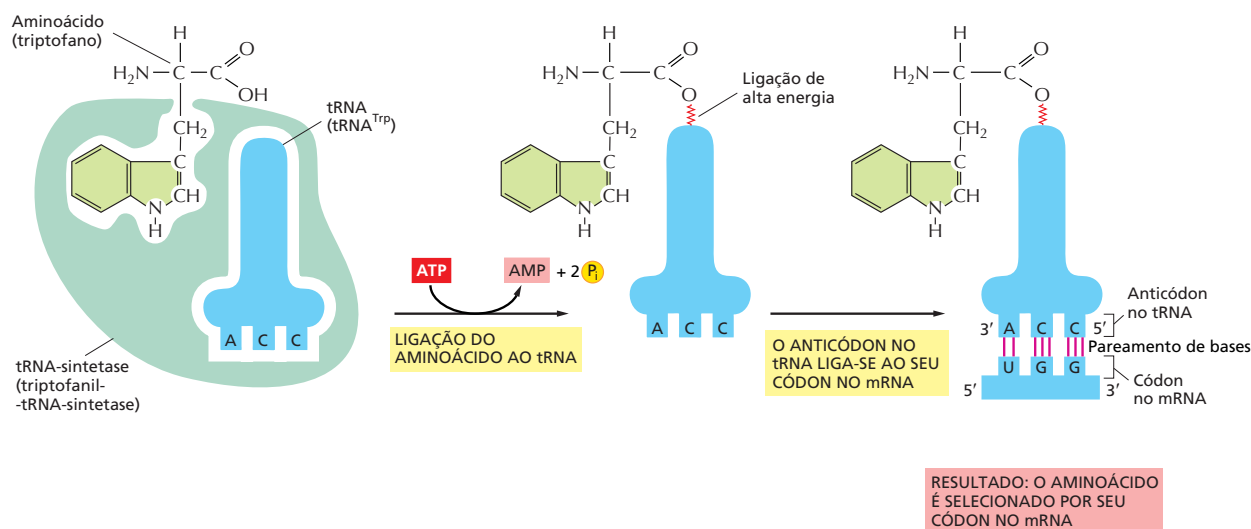
noácidos, ou que algumas moléculas de tRNA possam ligar-se por complementaridade de bases a mais de um códon. Na verdade, ambas as situações ocorrem. Alguns aminoácidos possuem mais de um tRNA e alguns tRNAs são construídos de tal forma que eles exigem um pareamento de bases exato apenas às duas primeiras posições do códon, podendo tolerar um pareamento inexato (ou *oscilação*) sobre a terceira posição. Esse pareamento oscilante pode explicar por que tantos entre os códons alternativos de um aminoácido diferem apenas em seus nucleotídeos da terceira posição (ver Figura 7-25). Os pareamentos oscilantes tornam 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.

Enzimas específicas acoplam os tRNAs aos aminoácidos corretos

Para que uma molécula de tRNA desempenhe o seu papel como um adaptador, ela deve ser ligada a – ou carregada com – um aminoácido correto. De que forma cada molécula de tRNA reconhece, entre os 20 aminoácidos possíveis, seu parceiro correto e ideal? 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, há uma enzima sintetase diferente para cada aminoácido. Isso significa que existem 20 sintetases ao todo: uma liga a glicina a todos os tRNAs que reconhecem os códons para a glicina, outra liga a fenilalanina a todos os tRNAs que reconhecem códons para a fenilalanina, e assim por diante. Cada enzima sintetase reconhece nucleotídeos específicos, tanto no anticódon quanto no braço aceptor de aminoácidos do tRNA correto (**Animação 7.6**). As sintetases são, portanto, tão importantes quanto os tRNAs no processo de decodificação, pois é a ação combinada das sintetases e dos tRNAs que permite que cada códon na molécula de mRNA especifique o seu aminoácido adequado (**Figura 7-30**).

A reação catalisada pela sintetase que liga o aminoácido à extremidade 3' do tRNA é uma das muitas reações celulares acoplada à liberação de energia pela hidrólise de ATP (ver Figura 3-33). A reação produz uma ligação de alta energia entre o tRNA carregado e o aminoácido. A energia dessa ligação é posteriormente usada para ligar o aminoácido covalentemente à cadeia polipeptídica em crescimento.

Figura 7-30 O código genético é traduzido pela atuação conjunta de dois adaptadores: aminoacil-tRNA-sintetases e tRNAs. Cada sintetase acopla um aminoácido específico a seus tRNAs correspondentes, em um processo chamado de carregamento. O anticódon da molécula de tRNA carregada forma pares de bases com o códon apropriado no mRNA. Um erro, seja na etapa de carregamento, seja na etapa de ligação do tRNA carregado ao seu códon, fará com que o aminoácido errado seja incorporado em uma cadeia de proteína. Na sequência de eventos ilustrada, o aminoácido triptofano (Trp) é selecionado pelo códon UGG no mRNA.



A mensagem do mRNA é decodificada por 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 tradução precisa e rápida do mRNA em proteína requer uma maquinaria molecular capaz de mover-se ao longo do mRNA, capturar moléculas de tRNA complementares, manter os tRNAs em posição, e ainda ligar covalentemente os aminoácidos por eles transportados para formar uma cadeia polipeptídica. Tanto em procariotos quanto em eucariotos, a maquinaria que dá início ao processo é o **ribossomo** – um grande complexo composto por dezenas de pequenas proteínas (as *proteínas ribossômicas*) e várias moléculas essenciais de RNA, chamadas de **RNAs ribossômicos (rRNAs)**. Uma célula eucariótica típica contém milhões de ribossomos em seu citoplasma (**Figura 7-31**).

QUESTÃO 7-4

Em um inteligente experimento realizado em 1962, uma cisteína já associada ao seu tRNA foi quimicamente convertida em alanina. Essas moléculas de tRNA “híbridas” foram adicionadas depois 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 polipeptídica 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.

Os ribossomos de eucariotos e procariotos são bastante semelhantes em estrutura e função. 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 de daltons (**Figura 7-32**); para comparação, uma proteína de tamanho médio possui uma massa igual a 30.000 daltons. A subunidade ribossômica pequena pareia 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. Essas duas subunidades se reúnem sobre uma molécula de mRNA, próximo de sua extremidade 5', para iniciar a síntese de uma proteína. O mRNA é então puxado ao longo do ribossomo como uma longa fita. Conforme o mRNA avança na direção de 5' para 3', o ribossomo traduz a sua sequência de nucleotídeos em uma sequência de aminoácidos, um códon de cada vez, utilizando os tRNAs como adaptadores. Cada aminoácido é acrescentado, na sequência correta, à extremidade final da cadeia polipeptídica em crescimento (**Animação 7.7**). Quando a síntese da proteína é finalizada, as duas subunidades do ribossomo se separam. Os ribossomos operam com uma incrível eficiência: um ribossomo eucarioto adiciona cerca de 2 aminoácidos por segundo a uma cadeia polipeptídica; um ribossomo bacteriano opera ainda mais rapidamente, adicionando cerca de 20 aminoácidos por segundo.

Como os ribossomos conseguem orquestrar todos os movimentos necessários para a tradução? Além de um sítio de ligação para uma molécula de mRNA, cada ribossomo contém três sítios de ligação para moléculas de tRNA, denominados sítio A, sítio P e sítio E (**Figura 7-33**). Para adicionar um aminoácido à

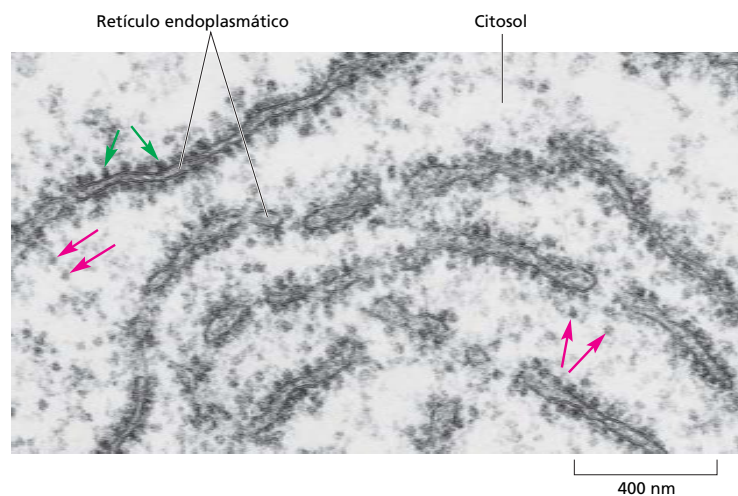


Figura 7-31 Os ribossomos estão localizados no citoplasma das células eucarióticas. Esta micrografia eletrônica mostra uma fina secção de uma pequena região do citoplasma. Os ribossomos aparecem como pequenas bolhas cinza. Alguns estão livres no citosol (setas vermelhas); outros estão ligados a membranas do retículo endoplasmático (setas verdes). (Cortesia de George Palade.)

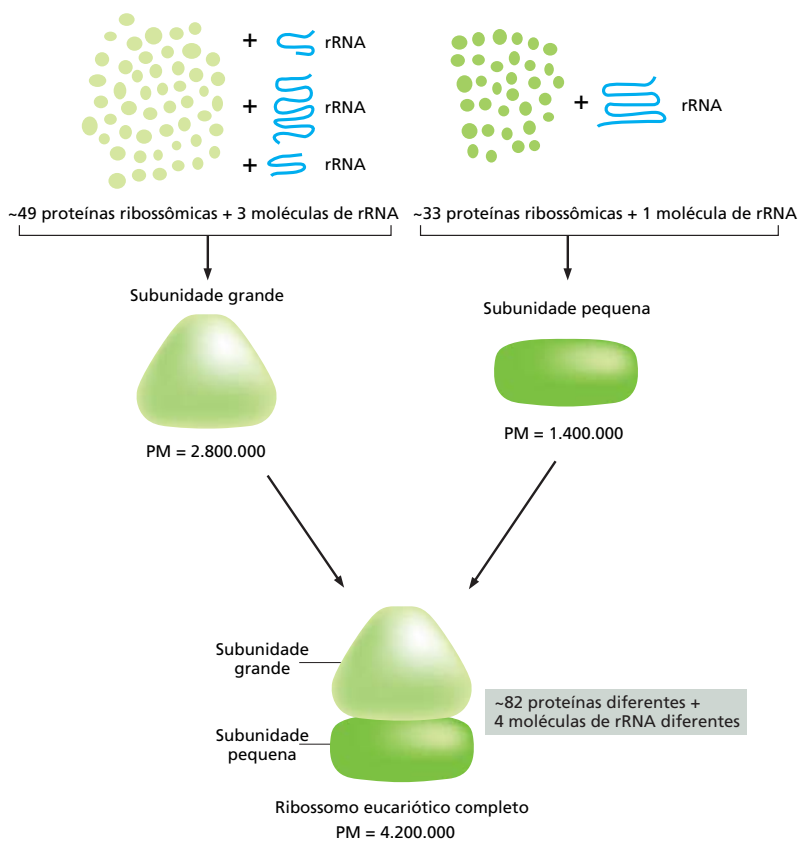


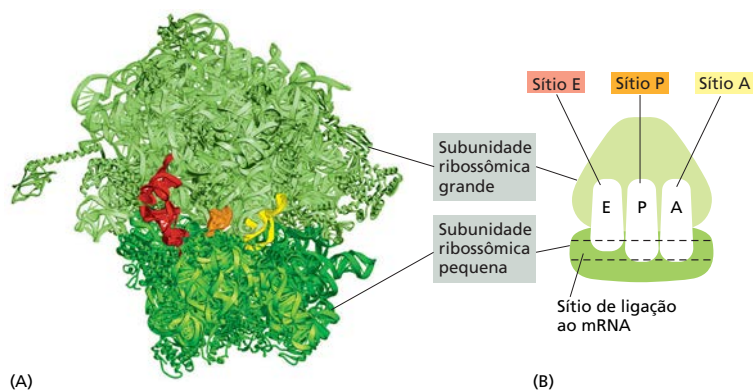
Figura 7-32 O ribossomo eucarioto é um grande complexo de quatro rRNAs e mais de 80 pequenas proteínas.

Os ribossomos procarióticos são muito semelhantes: ambos são formados a partir de uma subunidade grande e uma subunidade pequena, que só se unem após a pequena subunidade estar ligada a um mRNA. Embora as proteínas ribossômicas sejam mais numerosas do que os rRNAs, os RNAs são responsáveis pela maior parte da massa do ribossomo e por sua forma e estrutura gerais.

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 peptídica mantida em posição pelo tRNA que se encontra no sítio P adjacente. Em seguida, a subunidade ribossômica grande desloca-se para frente, movendo o tRNA usado para o sítio E antes de ejetá-lo (Figura 7-34). Esse ciclo de reações é repetido cada vez que um aminoácido é adicionado à cadeia polipeptídica, com a nova proteína crescendo de sua extremidade amino para sua extremidade carboxila até que um códon de terminação seja encontrado no mRNA.

Figura 7-33 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 – exit, respectivamente). (A) Estrutura tridimensional de um ribossomo bacteriano, determinada por cristalografia de raios X, com a subunidade pequena em verde-escuro, e a subunidade grande em verde-claro. Tanto rRNAs quanto proteínas ribossômicas estão ilustrados em verde. Os tRNAs estão representados ligados ao sítio E (vermelho), ao sítio P (laranja) e ao sítio 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 7-34). (B) Representação altamente esquemática de um ribossomo (na mesma orientação de A), que será utilizada nas figuras subsequentes. Observe que tanto a subunidade pequena quanto a grande estão envolvidas na formação dos sítios A, P e E, enquanto apenas a subunidade pequena forma o sítio de ligação para um mRNA. (B, adaptada de M.M. Yusupov et al., *Science* 292:883–896, 2001, com permissão de AAAS. Cortesia de Albion Baucom e Harry Noller.)



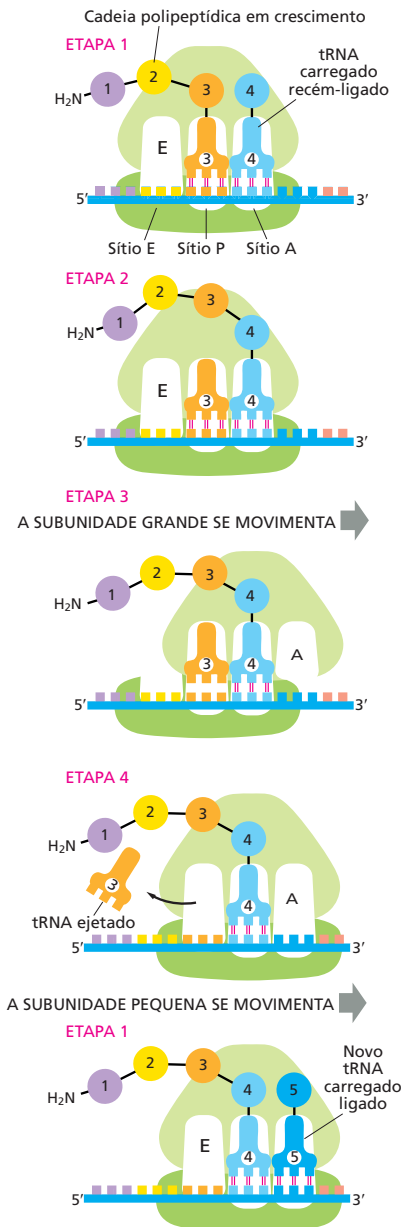


Figura 7-35 Os RNAs ribossômicos conferem ao ribossomo sua forma geral. São ilustradas as estruturas detalhadas dos dois rRNAs que formam a região central da subunidade grande de um ribossomo bacteriano – rRNA 23S (azul) e rRNA 5S (roxo). Uma das subunidades proteicas do ribossomo (L1) está incluída como ponto de referência, visto que essa proteína é responsável por uma protuberância característica na superfície ribossômica. Os componentes ribossômicos 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.)

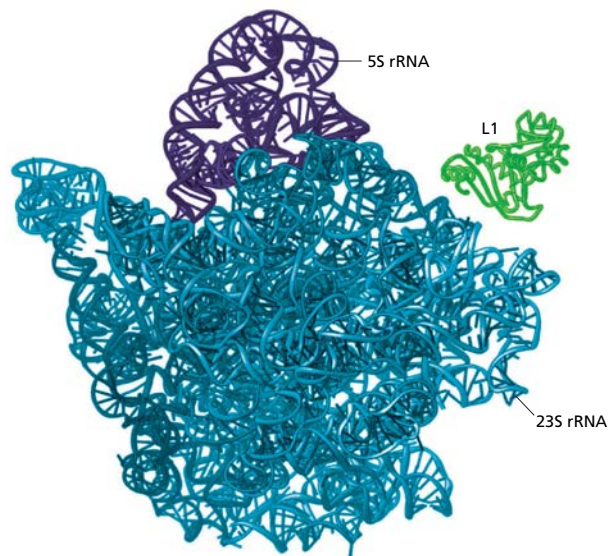
Figura 7-34 A tradução ocorre em um ciclo de quatro etapas. Este ciclo é repetido muitas e muitas vezes durante a síntese de uma cadeia proteica. Na etapa 1, um tRNA carregado, transportando o próximo aminoácido a ser adicionado à cadeia polipeptídica, se liga ao sítio A vazio sobre o ribossomo pela formação de pares de bases com o códon do mRNA ali exposto. Como apenas as moléculas de tRNA apropriadas podem formar pares de bases com um dado códon, esse códon determina o aminoácido específico a ser incorporado. 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 códons 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 (aminoácido 3 na etapa 1) 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 leva os dois tRNAs a se posicionarem 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 em sua conformação original em relação à subunidade grande. Esse movimento ejet o tRNA usado e reinicializa o ribossomo com um sítio A vazio, de tal forma que uma nova molécula de tRNA carregada possa se ligar (**Animação 7.8**).

Como indicado, o mRNA é traduzido no sentido 5' para 3', e a extremidade N-terminal de uma proteína é sintetizada primeiro, com cada ciclo adicionando um novo aminoácido à extremidade C-terminal da cadeia polipeptídica. Para visualizar o ciclo da tradução em detalhes, ver **Animação 7.9**.

O ribossomo é uma ribozima

O ribossomo é uma das maiores e mais complexas estruturas da célula, sendo seu peso 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. A estrutura confirmou evidências anteriores de que os rRNAs – e não as proteínas – são responsáveis pela estrutura geral do ribossomo e por sua capacidade de coreografar e catalisar a síntese de proteínas.

Os rRNAs estão dobrados em estruturas tridimensionais altamente precisas e compactas que formam o cerne do ribossomo (**Figura 7-35**). Em marcante contraste ao posicionamento central do rRNA, as proteínas ribossômicas estão geralmente localizadas na superfície, onde preenchem as fendas e frestas do RNA dobrado. A principal função das proteínas ribossômicas parece ser auxiliar na



manutenção da estrutura e da estabilidade do núcleo de RNA, permitindo ainda que aconteçam as alterações na conformação do rRNA necessárias para que esse RNA catalise de maneira eficiente a síntese proteica.

Não apenas os três sítios de ligação do tRNA (os sítios A, P e E) no ribossomo são primariamente formados por rRNAs, mas também o sítio catalítico para a formação da ligação peptídica é formado pelo rRNA 23S da subunidade grande; a proteína ribossômica mais próxima está localizada a uma distância muito grande, o que impede que ela faça contato com o tRNA carregado recém-chegado ou com a cadeia polipeptídica em crescimento. O sítio catalítico nesse rRNA – uma peptidil-transferase – é, em muitos aspectos, semelhante ao encontrado em algumas enzimas proteicas: consiste em uma fenda altamente estruturada que orienta precisamente os dois reagentes – o polipeptídeo em crescimento e o tRNA carregado –, dessa forma fortemente incrementando a probabilidade de uma reação produtiva.

As moléculas de RNA que possuem atividade catalítica são denominadas **ribozimas**. Mais tarde, na seção final deste capítulo, consideramos outras ribozimas e discutimos o que a catálise com base em RNA deve ter significado para a evolução inicial da vida na Terra. No momento, iremos 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 for verdade, o ribossomo, com seu núcleo de RNA catalítico, pode ser considerado como uma relíquia do período inicial da história da vida, quando as células eram dirigidas quase exclusivamente por ribozimas.

Códons específicos no mRNA sinalizam para o ribossomo os pontos de início e final da síntese proteica

Em um tubo de ensaio, os ribossomos podem ser forçados a traduzir qualquer molécula de RNA (ver Como Sabemos, p. 240-241). Em uma célula, no entanto, um sinal específico é necessário para a iniciação da tradução. 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. Nessa etapa, um erro de um nucleotídeo em qualquer dos sentidos 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 7-26). A taxa de iniciação determina a taxa na qual a proteína é sintetizada a partir do mRNA.

A tradução de um mRNA tem início com o códon AUG, e um tRNA especialmente carregado é necessário para a iniciação da tradução. Esse **tRNA iniciador** sempre carrega o aminoácido metionina (ou uma forma modificada da metionina, a formil-metionina, em bactérias). Assim, todas as proteínas recentemente sintetizadas possuem uma metionina como o primeiro aminoácido na sua extremidade N-terminal, a extremidade onde é iniciada a síntese de uma proteína. Essa metionina é, em geral, removida posteriormente pela ação de uma protease específica.

Em eucariotos, um tRNA iniciador, carregado com metionina, é inicialmente inserido no sítio P da subunidade ribossômica pequena, juntamente com proteínas adicionais denominadas **fatores de iniciação da tradução** (Figura 7-36). O tRNA iniciador é distinto dos tRNAs que normalmente transportam metionina. De todos os tRNAs na célula, apenas uma molécula de tRNA iniciador carregada é capaz de se ligar firmemente ao sítio P na ausência da subunidade ribossômica

QUESTÃO 7-5

Uma fita de DNA com a seguinte sequência de nucleotídeos – 5'-TTA-ACGGCTTTTTC-3' – foi usada como molde para a síntese de um mRNA que, a seguir, foi traduzido em proteína. 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.

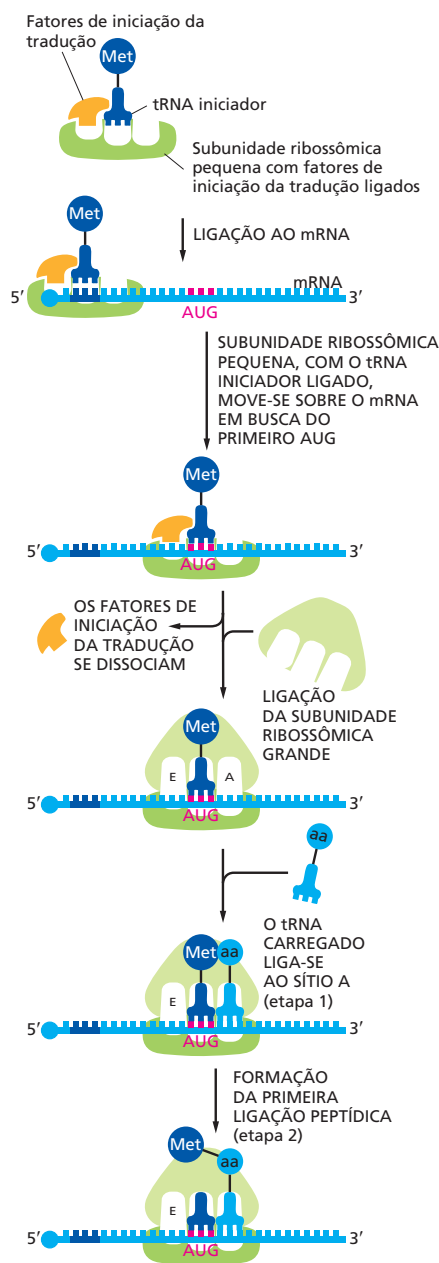


Figura 7-36 A iniciação da síntese de proteínas em eucariotos requer fatores de iniciação da tradução e um tRNA iniciador especial. Apesar de não estarem aqui ilustradas, uma iniciação de tradução eficiente também requer proteínas adicionais ligadas ao quepe 5' e à cauda poli-A do mRNA (ver Figura 7-23). Dessa maneira, o aparato de tradução se certifica de 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 7-34.

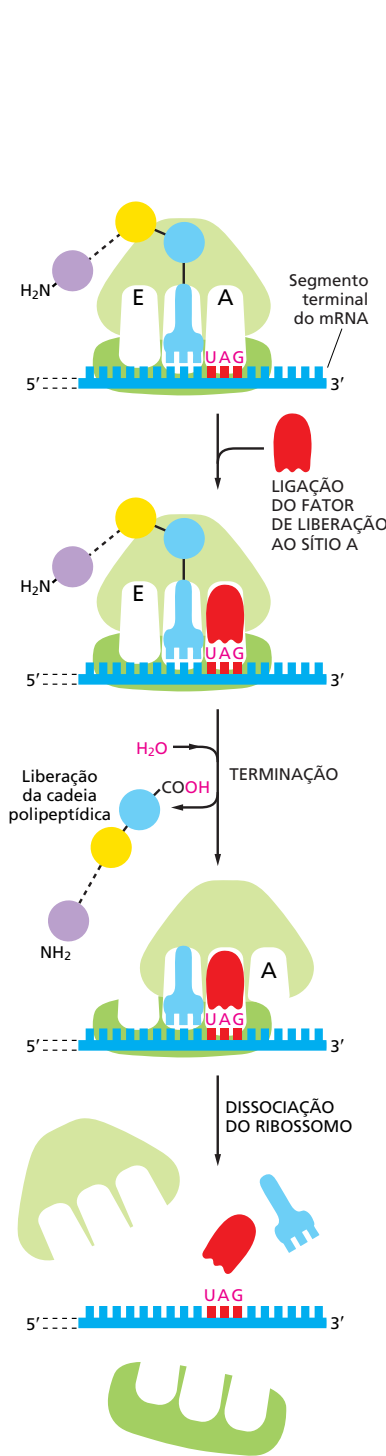


Figura 7-38 A tradução é interrompida em um códon de terminação. Na fase final da síntese proteica, a ligação do fator de liberação a um sítio A que contém um códon de terminação finaliza a tradução de uma molécula de mRNA. O polipeptídeo completo é liberado e o ribossomo se dissocia, liberando suas duas subunidades. Observe que apenas a extremidade 3' da molécula de mRNA está ilustrada.

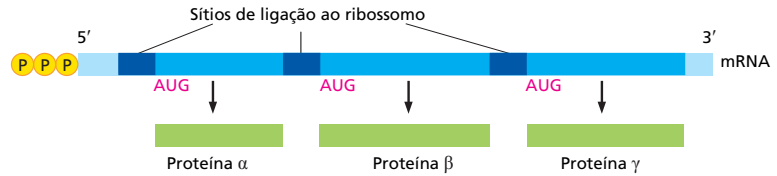


Figura 7-37 Uma única molécula de mRNA procariótico pode codificar várias proteínas diferentes. Em procariotos, genes envolvidos em diferentes passos de um mesmo processo se encontram frequentemente organizados em grupos (óperons) que são transcritos em conjunto sob a forma de um único mRNA. Um mRNA procariótico não tem o mesmo tipo de quepe 5' que está presente nos mRNAs eucarióticos, apresentando em seu lugar um trifosfato na extremidade 5'. Os ribossomos procarióticos iniciam a tradução em sítios de ligação ao ribossomo (azul-escuro), que podem estar localizados no interior de uma molécula de mRNA. Essa característica permite que os procariotos sintetizem diferentes proteínas a partir de uma única molécula de mRNA, sendo cada proteína produzida por um ribossomo diferente.

grande. Em seguida, a subunidade ribossômica pequena carregada com o tRNA iniciador liga-se à extremidade 5' de uma molécula de mRNA, que está identificada pelo quepe 5' presente em todos os mRNAs eucarióticos (ver Figura 7-16). A subunidade ribossômica pequena então se move para frente (de 5' para 3') sobre o mRNA, à procura do primeiro códon AUG. Quando esse AUG é encontrado e reconhecido pelo tRNA iniciador, vários fatores de iniciação dissociam-se da subunidade ribossômica pequena abrindo caminho para a ligação da subunidade ribossômica grande e para a montagem completa do ribossomo. Estando o tRNA iniciador 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 7-34).

Em bactérias, o mecanismo de seleção de um códon de iniciação é diferente. Os mRNAs bacterianos não possuem o quepe 5' para indicar ao ribossomo onde iniciar a busca pelo ponto de início da tradução. Em vez dessa estrutura, 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 localizado no interior de um mRNA, desde que um sítio de ligação ao ribossomo o preceda em vários nucleotídeos. 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 sendo traduzidas a partir da mesma molécula de mRNA (Figura 7-37). Em contraste, um mRNA eucariótico geralmente transporta informação referente a uma única proteína.

O fim da tradução tanto em procariotos quanto em eucariotos é sinalizado pela presença de um de vários códons, denominados *códons de terminação*, presentes no mRNA (ver Figura 7-25). Os códons de terminação – UAA, UAG e UGA – não são reconhecidos por um tRNA e não especificam um aminoácido, mas, em vez disso, sinalizam o término da tradução para o ribossomo. Proteínas conhecidas como *fatores de liberação* ligam-se a qualquer códon de terminação que chegue a um sítio A do ribossomo, e essa ligação altera a atividade da peptidil-transferase no ribossomo, fazendo com que seja catalisada a adição de uma molécula de água, em vez de um aminoácido ao peptidil-tRNA (Figura 7-38). Essa reação libera a extremidade carboxila da cadeia polipeptídica de sua conexão à molécula de tRNA; 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. Nesse momento, o ribossomo também libera o mRNA e dissocia suas duas subunidades, que poderão, posteriormente, unir-se 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 enquanto ainda estão sendo sintetizadas no ribossomo. A maio-

ria das proteínas, no entanto, requer *proteínas chaperonas* para ajudá-las a dobrar corretamente na célula. As chaperonas podem “ciceronear” as proteínas ao longo de vias produtivas de dobramento e impedir que ocorra agregação dentro da célula (ver Figuras 4-9 e 4-10). As proteínas recém-sintetizadas são frequentemente interceptadas por suas chaperonas conforme emergem do ribossomo.

As proteínas são produzidas em polirribossomos

A síntese da maior parte das moléculas proteicas leva entre 20 segundos e alguns minutos. Mas mesmo durante este curto período, vários ribossomos normalmente se ligam a cada molécula de mRNA a ser traduzida. Se o mRNA está sendo traduzido de maneira eficiente, um novo ribossomo é montado sobre a extremidade 5' de uma molécula de mRNA quase imediatamente após o ribossomo precedente ter traduzido uma sequência nucleotídica longa o suficiente para não mais o atrapalhar. As moléculas de mRNA que estão sendo traduzidas são, por conseguinte, normalmente encontradas sob a forma de *polirribossomos*, também conhecidos como *polissomos*. Esses grandes arranjos citoplasmáticos são compostos de muitos ribossomos espaçados por aproximadamente 80 nucleotídeos ao longo de uma única molécula de mRNA (**Figura 7-39**). Visto que múltiplos ribossomos podem atuar simultaneamente sobre um único mRNA, muito mais moléculas de proteína podem ser feitas em um dado tempo do que seria possível se fosse necessário completar cada polipeptídeo antes de poder-se iniciar a síntese do polipeptídeo seguinte.

Os polissomos operam tanto em bactérias quanto em eucariotos, mas as bactérias podem acelerar ainda mais a taxa de síntese proteica. 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 encaço da RNA-polimerase, conforme esta se move sobre o DNA.

Os inibidores da síntese proteica de procaríotos são utilizados como antibióticos

A capacidade de traduzir de maneira eficiente o mRNA em proteínas é uma característica fundamental à toda a vida na Terra. Embora o ribossomo e outras moléculas que levam a cabo essa tarefa complexa sejam muito semelhantes entre os organismos, vimos que existem algumas diferenças sutis na forma como as bactérias e os eucariotos sintetizam RNA e proteínas. Apesar de ser uma peculiaridade evolutiva, essas diferenças formam a base de um dos mais importantes avanços da medicina moderna.

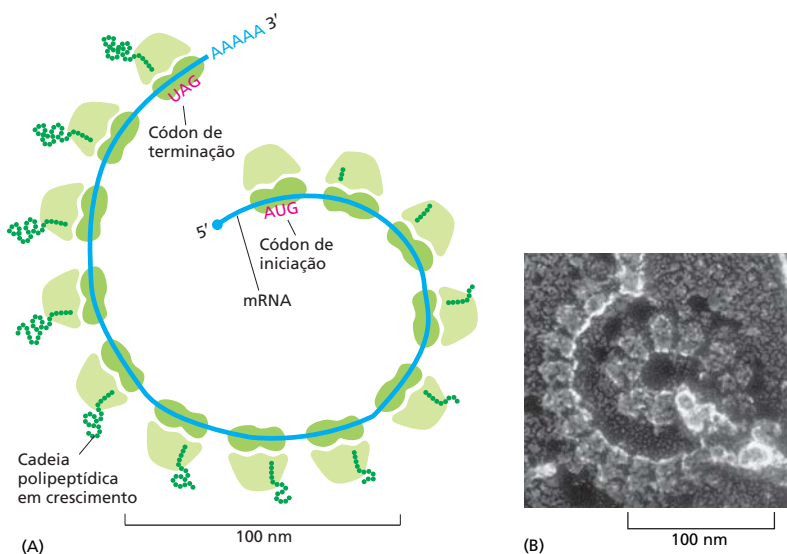


Figura 7-39 As proteínas são sintetizadas em polirribossomos. (A) Desenho esquemático mostrando como uma série de ribossomos pode traduzir simultaneamente a mesma molécula de mRNA (**Animação 7.10**). (B) Micrografia eletrônica de um polirribossomo no citosol de uma célula eucariótica. (B, cortesia de John Heuser.)

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

Antibiótico	Efeito específico
Tetraciclina	Bloqueia a ligação da aminoacil-tRNA ao sítio A do ribossomo (etapa 1 na Figura 7-34)
Estreptomicina	Impede a transição do complexo de iniciação para o alongamento da cadeia (ver Figura 7-36); também pode causar erros de decodificação
Cloranfenicol	Bloqueia a reação da peptidil-transferase nos ribossomos (etapa 2 na Figura 7-34)
Ciclo-heximida	Bloqueia a reação de movimentação (translocação) nos ribossomos (etapa 3 na Figura 7-34)
Rifamicina	Bloqueia a iniciação da transcrição por meio de ligação à RNA-polimerase

Muitos dos nossos antibióticos mais eficazes são compostos que atuam na inibição da síntese proteica e na síntese do RNA bacteriano, mas não do eucariótico. Alguns desses fármacos exploram pequenas diferenças estruturais e funcionais existentes entre os ribossomos bacterianos e eucarióticos, de tal forma que interferem preferencialmente na síntese proteica bacteriana. Esses compostos podem, assim, ser ingeridos em doses suficientemente elevadas para matar as bactérias sem que apresentem toxicidade para os 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 dos antibióticos que inibem a síntese proteica e de RNA em bactérias estão listados na **Tabela 7-3**.

Diversos antibióticos comumente utilizados foram inicialmente isolados de fungos. Fungos e bactérias em geral 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. Como os fungos e os seres humanos são ambos eucariotos e, portanto, mais intimamente relacionados entre si do que às bactérias (ver Figura 1-28), tivemos a oportunidade de tomar emprestadas essas armas para combater os nossos próprios inimigos bacterianos.

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

Depois de uma proteína ser liberada do ribossomo, uma célula pode controlar a sua atividade e longevidade de diversas formas. O número de cópias de uma proteína em uma célula depende, assim como ocorre em uma população humana, não apenas de quão rapidamente novos indivíduos podem ser gerados, mas também de quanto tempo eles sobreviverão. Assim, o controle da degradação das proteínas em seus aminoácidos constituintes ajuda as células a regular a quantidade de cada proteína em particular. As proteínas diferem enormemente em relação à sua duração, ou tempo médio de vida. As proteínas estruturais que passam a fazer parte de um tecido relativamente estável, como o tecido ósseo ou o tecido muscular, podem durar meses ou mesmo anos, ao passo que outras proteínas, como enzimas metabólicas ou aquelas que regulam o crescimento e a divisão celular (discutida no Cap. 18), só duram alguns dias, horas ou mesmo poucos 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 por fim nos aminoácidos individuais, são coletivamente denominadas

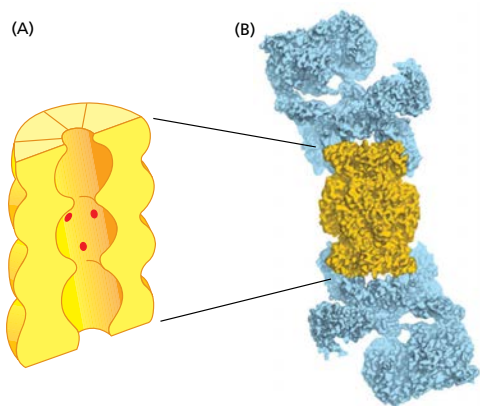


Figura 7-40 Um proteassomo degrada proteínas de curta duração ou proteínas erroneamente dobradas. As estruturas mostradas foram determinadas por cristalografia de raios X. (A) Uma vista em corte do cilindro central do proteassomo, com os sítios ativos das proteases indicados por pontos vermelhos. (B) A estrutura completa do proteassomo, na qual o acesso ao cilindro central (amarelo) é regulado por uma “rolha” (azul) em cada extremidade. (B, adaptado de P.C.A. da Fonseca et al., *Mol. Cell* 46:54–66, 2012.)

proteases. As proteases atuam pela clivagem (hidrólise) das ligações peptídicas entre os aminoácidos (ver Painel 2-5, p. 74-75). Uma das funções das vias proteolíticas é a rápida degradação das proteínas que devem ter curta duração. Outra função envolve o reconhecimento e a remoção de proteínas que estejam danificadas ou erroneamente dobradas. A eliminação das proteínas incorretamente dobradas é crítica para um organismo, pois proteínas deformadas tendem a se agregar, e agregados de proteína podem danificar as células ou mesmo desencadear a morte celular. Finalmente, todas as proteínas, até mesmo as de longa duração, acumulam danos e são degradadas por proteólise.

Em células eucarióticas, as proteínas são quebradas por grandes máquinas proteicas denominadas **proteassomos**, presentes tanto no citosol quanto no núcleo. 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 tipos de subunidades proteicas (**Figura 7-40**). Essas “rolhas proteicas” se ligam às proteínas destinadas à degradação e, em seguida, utilizando a energia da hidrólise de ATP, desdobram as proteínas condenadas à degradação e as inserem na câmara interna do cilindro. Uma vez no interior da câmara, as proteínas são clivadas por proteases em pequenos peptídeos que, a seguir, serão ejetados por ambas as extremidades do proteassomo. O uso de câmaras de destruição com proteases isoladas em seu interior impede que essas enzimas atuem inespecificamente e promovam danos na célula.

Como os proteassomos selecionam quais proteínas celulares deverão sofrer degradação? Em eucariotos, 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 cadeia curta de moléculas de ubiquitina; estas proteínas ubiquitinadas são, então, reconhecidas, desdobradas e entregues aos proteassomos por proteínas específicas da “rolha” (*stopper*) dessa estrutura (**Figura 7-41**).

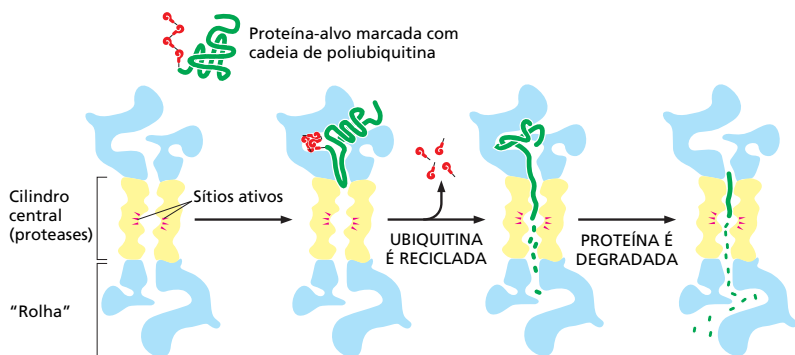


Figura 7-41 As proteínas marcadas por uma cadeia de poliubiquitina são degradadas pelo proteassomo. Proteínas na rolha do proteassomo (azul) reconhecem as proteínas-alvo marcadas por um tipo específico de cadeia de poliubiquitina. A rolha, em seguida, desenrola a proteína-alvo e insere-a no cilindro central do proteassomo (amarelo), que é revestido por proteases que cortam a proteína em pequenos pedaços.

As proteínas destinadas a ter curta duração, muitas vezes, contêm uma curta sequência de aminoácidos que as identifica como um alvo a ser ubiquitinado e degradado nos proteassomos. Proteínas danificadas 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 uma cadeia de poliubiquitina a essas proteínas reconhecem sinais que ficam expostos como consequência do dobramento incorreto ou de danos químicos – por exemplo, sequências de aminoácidos ou motivos conformacionais internos ou inacessíveis na proteína normal “saudável”.

Existem várias etapas entre o DNA e a proteína

Já vimos que muitos tipos de reações químicas são necessários para a produção de uma proteína a partir da informação contida em um gene. Portanto, a concentração final de uma proteína em uma célula depende da taxa de reação de cada um dos diversos passos (Figura 7-42). Além disso, várias proteínas – depois de serem liberadas no ribossomo – precisam de outras alterações para que possam ser úteis às células. Exemplos de tais *modificações pós-traducionais* incluem as modificações covalentes (como a fosforilação), a ligação de cofatores de pequenas moléculas, ou a associação com outras subunidades proteicas, muitas vezes necessárias para que uma proteína recém-sintetizada se torne plenamente funcional (Figura 7-43).

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, todos os passos da Figura 7-42 podem ser regulados pela célula – e,

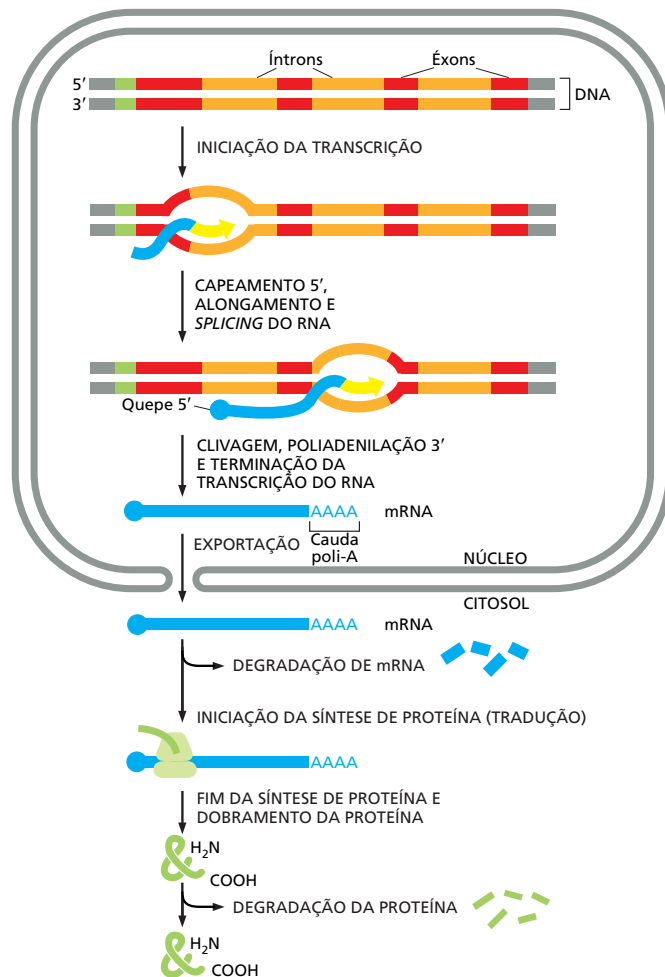


Figura 7-42 A produção de uma proteína em uma célula eucariótica requer diversas etapas. A concentração final de cada proteína depende da velocidade de cada etapa indicada. Mesmo após a produção de um mRNA e da sua proteína correspondente, suas concentrações podem ser reguladas via degradação. Embora não ilustradas, a atividade da proteína pode também ser regulada por outras modificações pós-traducionais ou pela ligação de pequenas moléculas (ver Figura 7-43).

de fato, muitos deles o são. Todavia, como veremos no próximo capítulo, a iniciação da transcrição é o ponto mais comum usado pelas células para regular a expressão dos 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.

RNA E A ORIGEM DA VIDA

O dogma central – de que o DNA dá origem ao RNA que, por sua vez, dá origem às proteínas – apresenta um intrincado e paradoxal quebra-cabeça aos biólogos evolutivos: se os ácidos nucleicos são necessários para dirigir a síntese das proteínas, e as proteínas são necessárias para sintetizar os ácidos nucleicos, como poderia esse sistema de componentes interdependentes ter surgido? Uma visão alternativa sugere a existência de um **mundo de RNA** na Terra antes do aparecimento das células que contêm DNA e proteínas. De acordo com essa hipótese, o RNA – que hoje atua predominantemente como 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 (Figura 7-44). Se essa ideia estiver correta, então, a transição do mundo de RNA nunca foi completa; como vimos, o RNA ainda catalisa várias reações fundamentais nas células atuais. Esses RNA catalíticos, ou ribozimas, incluindo os que operam no ribossomo e na maquinaria de *splicing* de RNA, podem, assim, ser considerados fósseis moleculares de um mundo ancestral.

A vida requer autocatálise

A origem da vida necessitou de moléculas que possuíssem, pelo menos em certo 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 de autorreprodução, uma vez surgidos por acaso, poderiam utilizar matérias-primas provenientes da produção de outras substâncias para fazer cópias de si próprios. 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 primitivo da Terra. Um sistema *autocatalítico* dessa natureza deveria ter muitas das propriedades que consideramos como características da 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 ao equilíbrio químico e “morrer”.

Quais moléculas podem ter apresentado tais propriedades autocatalíticas? Nas células vivas atuais, os catalisadores mais versáteis são as proteínas – capazes de adotar diferentes conformações tridimensionais que formam sítios quimicamente



Figura 7-43 Muitas proteínas precisam de diversas modificações antes de se tornarem totalmente funcionais. Para ser útil para a célula, um polipeptídeo sintetizado deve dobrar-se corretamente, assumindo a sua conformação tridimensional exata e, em seguida, ligar-se a eventuais cofatores necessários (vermelho) ou outras proteínas – por meio de ligações não covalentes. Muitas proteínas também necessitam de uma ou mais modificações covalentes para se tornarem ativas – ou para serem recrutadas para membranas ou organelas específicas (não ilustrado). Apesar de a fosforilação e a glicosilação serem as alterações mais comuns, mais de 100 diferentes tipos de modificações covalentes são conhecidos para as proteínas.

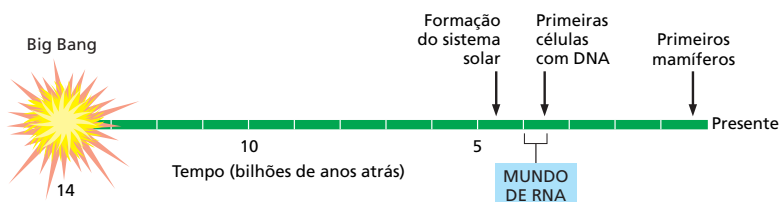


Figura 7-44 Um mundo de RNA pode ter existido antes da existência das células modernas com DNA e proteínas.

reativos em sua superfície. No entanto, não é conhecido qualquer meio pelo qual uma proteína possa reproduzir a si própria diretamente. As moléculas de RNA, em contrapartida, podem – pelo menos teoricamente – catalisar sua própria síntese.

O RNA pode tanto estocar informação como catalisar reações químicas

Vimos que o pareamento de bases complementares permite que um ácido nucleico atue como molde para a formação de outro ácido nucleico. Assim, uma fita simples de RNA ou DNA pode determinar a sequência de um polinucleotídeo complementar, o qual, por sua vez, pode determinar a sequência da molécula original, permitindo que o ácido nucleico original seja replicado (**Figura 7-45**). Esses mecanismos de molde por complementaridade formam a base da replicação do DNA e da transcrição nas células atuais.

Contudo, 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 catalisada por proteínas enzimáticas – como as DNA-polimerases e as 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 foi obtido em 1982, quando foi descoberto que as próprias moléculas de RNA podem atuar como catalisadoras. Acredita-se que o potencial sem igual das moléculas de RNA, tornando-as capazes de atuar como carreadoras de informação e como catalisadoras, tenha permitido que essas moléculas 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 pertencentes à própria fita. Esses pareamentos de base, em conjunto a ligações de hidrogênio não convencionais, podem levar cada molécula de RNA a se dobrar na forma de uma estrutura específica, que é determinada por sua sequência nucleotídica (ver Figura 7-5). Tais associações produzem conformações tridimensionais complexas.

Como discutimos no Capítulo 4, as enzimas são proteínas capazes de catalisar reações bioquímicas, pois possuem uma superfície com contornos específicos e propriedades químicas características. Do mesmo modo, as moléculas de RNA, com as suas formas dobradas características, podem atuar como catalisadoras (**Figura 7-46**). Os RNAs não apresentam a mesma diversidade estrutural e funcional das enzimas proteicas; afinal, eles são constituídos a partir de apenas quatro subunidades diferentes. Apesar dessa limitação, as ribozimas podem catalisar um espectro bastante variado de reações químicas. Visto que existe um número relativamente pequeno de RNAs catalíticos nas células atuais, a maioria das ribozimas estudadas foi produzida em laboratório e selecionada com base em sua atividade catalítica em tubos de ensaio (**Tabela 7-4**). No entanto, os processos em que os RNAs catalíticos ainda parecem desempenhar um papel importante incluem algumas das etapas mais fundamentais da expressão da informa-

Figura 7-45 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 recém-produzida atua como molde para formar uma molécula de RNA com a sequência original. Visto que cada molécula molde pode produzir diversas cópias da fita complementar, essas reações podem resultar na multiplicação da sequência original.

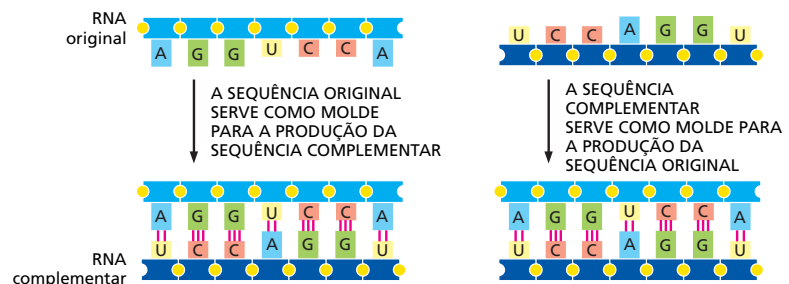


TABELA 7-4 Reações bioquímicas que podem ser catalisadas por ribozimas	
Atividade	Ribozimas
Formação da ligação peptídica na síntese de proteínas	RNA ribossômico (rRNA)
Ligação de DNA	RNA selecionado <i>in vitro</i>
Splicing do RNA	RNAs de <i>autosplicing</i> , pequenos RNAs nucleares (snRNAs)
Polimerização 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>

ção genética – especialmente 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 pode catalisar sua própria síntese (Figura 7-47). Embora os sistemas de autorreplicação de moléculas de RNA não tenham sido encontrados na natureza, mesmo assim os cientistas estão confiantes de que eles possam ser construídos em laboratório. Ainda que essa demonstração não prove que as moléculas de RNA autorreplicadoras foram essenciais para a origem da vida na Terra, ela certamente estabeleceria que um cenário assim é possível.

O RNA provavelmente antecedeu o DNA na evolução

As primeiras células na Terra presumivelmente devem ter sido bem menos complexas e menos eficientes na sua reprodução, se comparadas mesmo às 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 componentes necessários para fornecer materiais e energia para essa replicação autocatalítica. Se o papel evolutivo proposto anteriormente para o RNA estiver correto, essas células ancestrais também difeririam fundamentalmente das células que conhecemos hoje, pelo fato de terem suas informações hereditárias armazenadas no RNA, e não no 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

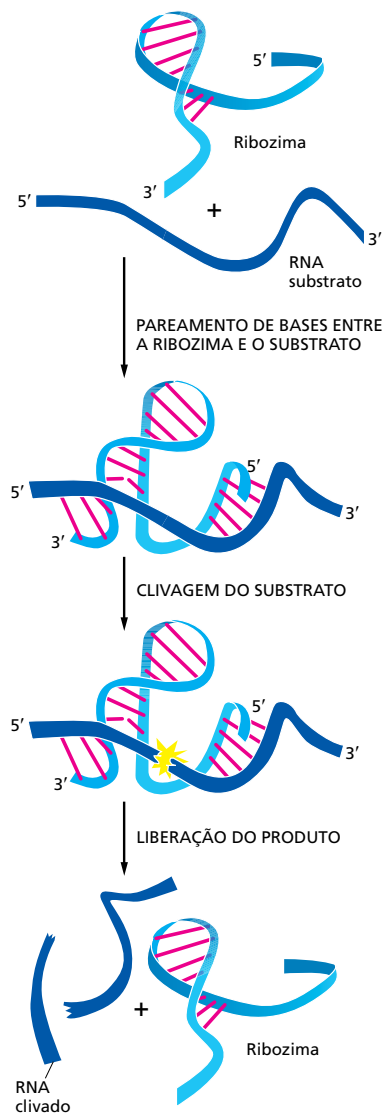


Figura 7-46 Uma ribozima é uma molécula de RNA que possui atividade catalítica. A molécula de RNA ilustrada catalisa a clivagem de uma segunda molécula de RNA em um sítio específico. Ribozimas similares são encontradas incorporadas em grandes genomas de RNA – chamados de viroides – que infectam plantas, nos quais a reação de clivagem é uma das etapas para a replicação do viroide. (Adaptada de T.R. Cech e O.C. Uhlenbeck, *Nature* 372:39–40, 1994. Com permissão de Macmillan Publishers Ltd.)

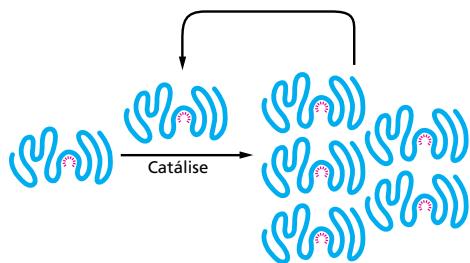


Figura 7-47 Pode uma molécula de RNA catalisar sua própria síntese? Esse processo hipotético exigiria que o RNA catalisasse ambas as etapas mostradas na Figura 7-45. Os raios vermelhos representam o sítio ativo dessa ribozima.

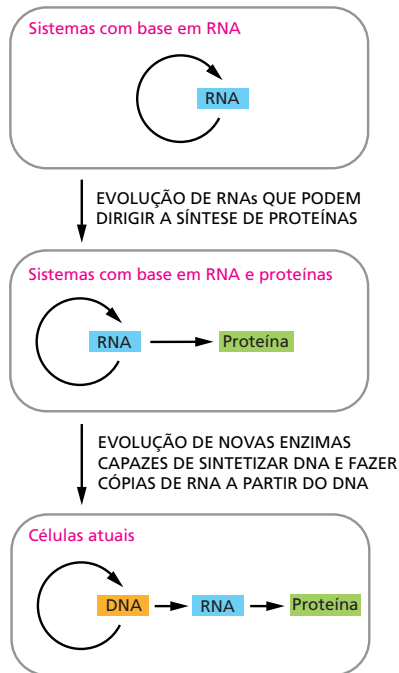


Figura 7-48 O RNA pode ter antecedido o DNA e as proteínas na evolução.

De acordo com essa hipótese, 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 algumas reações essenciais (incluindo a síntese de proteínas).

QUESTÃO 7-6

Discuta a seguinte afirmação: "Ao longo da evolução da vida na Terra, o RNA perdeu 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 7-3A), assim como a glicose e outros carboidratos simples, é facilmente formada a partir de formaldeído (HCHO), que é um dos principais produtos formados em experimentos que simulam as condições na 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 entrou em cena depois do RNA, e então se mostrou mais adaptado do que o RNA como repositório permanente de informação genética. Em particular, a desoxirribose, em sua cadeia principal de 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 ocorram quebras.

As outras diferenças entre o RNA e o 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 tornarem essa molécula mais fácil de reparar. Vimos, no Capítulo 6, que um nucleotídeo lesado sobre uma fita da dupla-hélice do DNA 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 da citosina é, por acaso, a uracila, a qual ocorre normalmente no RNA, portanto seria impossível que as enzimas de reparo detectassem tal alteração na molécula de RNA. No entanto, no DNA, que possui timina em vez de uracila, qualquer uracila produzida pela degradação acidental de citosina é facilmente detectada e reparada.

Em conjunto, as evidências que discutimos apoiam a ideia de que o RNA, com a sua capacidade para desempenhar funções genéticas, estruturais e catalíticas, tenha precedido o DNA na evolução. É possível que, conforme células mais semelhantes às células atuais foram surgindo, muitas das funções desempenhadas originalmente pelo RNA tenham sido assumidas pelo DNA e pelas proteínas: o DNA se sobrepôs na função genética principal, as proteínas se tornaram as principais catalisadoras, e o RNA permaneceu predominantemente como um intermediário, conectando-os (Figura 7-48). Com o advento do DNA, as células puderam tornar-se mais complexas, pois podiam conter e transmitir uma maior quantidade de informação genética, quando comparada àquela que poderia ser mantida de forma estável unicamente pelo RNA. Tendo em vista a maior complexidade química das proteínas e a maior diversidade das reações químicas que elas podiam catalisar, a substituição (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 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 os RNAs e proteínas 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 pela enzima RNA-polimerase, que utiliza sequências de nucleotídeos presentes nas moléculas de DNA para determinar qual fita será usada como molde, e quais serão os pontos de início e término da 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, assumindo estruturas tridimensionais complexas.
- As células produzem diversos tipos funcionais de RNAs, incluindo RNAs mensageiros (mRNAs), que carregam as instruções para fazer proteínas; RNAs ribossômicos (rRNAs), que são componentes essenciais dos ribossomos; e RNAs

transportadores (tRNAs), que agem como moléculas adaptadoras na síntese de proteínas.

- Para dar início à transcrição, a RNA-polimerase se liga a sítios específicos sobre o DNA, denominados promotores, situados imediatamente à montante dos genes. Para a iniciação da transcrição, as RNA-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.
- Em células eucarióticas, a maioria dos genes codificadores de proteínas é composta de regiões codificadoras, denominadas éxons, intercaladas com regiões não codificadoras maiores, chamadas de íntrons. Quando um gene eucarioto é transcrito do DNA para o RNA, tanto os éxons quanto os íntrons são copiados.
- Os íntrons são removidos dos transcritos de RNA no núcleo por *splicing* do RNA, em uma reação catalisada por pequenos complexos ribonucleoproteicos conhecidos como snRNPs. O *splicing* remove os íntrons do RNA e une os éxons – frequentemente em diferentes combinações, permitindo que múltiplas proteínas sejam produzidas a partir do mesmo gene.
- Pré-mRNAs eucarióticos passam por várias etapas adicionais de processamento do RNA antes de saírem do núcleo como mRNAs, incluindo o capeamento 5' do RNA e a poliadenilação da extremidade 3'. Essas reações, junto ao *splicing*, ocorrem conforme o pré-mRNA está sendo transcrito.
- A tradução da sequência de nucleotídeos do mRNA em proteína ocorre no citoplasma em grandes agregados ribonucleoproteicos denominados ribossomos. À medida que o mRNA se move pelo ribossomo, a 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 dos 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 moléculas adaptadoras na síntese proteica. Enzimas denominadas aminoacil-tRNA-sintetases acoplam covalentemente os aminoácidos aos tRNAs adequados. Cada tRNA contém uma sequência de três nucleotídeos, o anticódon, que reconhece um códon no mRNA pelo pareamento por complementaridade de bases.
- A síntese proteica inicia-se quando um ribossomo é organizado sobre um códon de iniciação (AUG) de uma molécula de mRNA, em um processo que depende de proteínas conhecidas como fatores de iniciação da tradução. A cadeia proteica completa é liberada do ribossomo quando um códon de terminação (UAA, UAG ou UGA) no mRNA é alcançado.
- A ligação sucessiva e passo a passo de aminoácidos em uma cadeia polipeptídica é catalisada por uma molécula de rRNA da subunidade ribossômica grande, que atua assim como uma ribozima.
- A concentração de uma proteína em uma célula depende das taxas nas quais o mRNA e a proteína são sintetizados e degradados. A degradação de proteína no citoplasma e no núcleo ocorre no interior de grandes complexos proteicos chamados de proteassomos.
- Considerando nosso conhecimento dos organismos atuais e das moléculas que eles contêm, parece razoável afirmar que a vida na Terra se originou a partir da evolução de moléculas de RNA capazes de catalisar sua própria replicação.
- Foi proposto que, nas primeiras células, o RNA tenha atuado tanto como genoma quanto como catalisador, antes que o DNA o substituísse como uma molécula mais estável para o armazenamento da informação genética, e que as proteínas substituíssem os RNAs como os principais componentes estruturais e catalíticos. Acredita-se que os RNAs catalíticos das células modernas possam nos dar um vislumbre desse mundo antigo, baseado em RNA.

TERMOS-CHAVE

aminoacil-tRNA-sintetase	mundo de RNA	RNA ribossômico (rRNA)
anticódon	pequenos RNAs nucleares (snRNA)	RNA transportador (tRNA)
capeamento do RNA	poliadenilação	spliceossomo
código genético	processamento do RNA	<i>splicing</i> (ou encadeamento) alternativo
códon	promotor	<i>splicing</i> (ou encadeamento) do RNA
éxon	protease	tradução
expressão gênica	proteassomo	transcrição de RNA
fase de leitura	ribossomo	transcrição
fator de iniciação da tradução	ribozima	tRNA iniciador
fator geral de transcrição	RNA	
gene	RNA-polimerase	
íntron	RNA mensageiro (mRNA)	

TESTE SEU CONHECIMENTO

QUESTÃO 7-7

Quais das seguintes afirmativas estão corretas? Explique 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 pode conter a sequência **ATTGACCCCGTCAA**.
- A quantidade de proteína presente em uma célula 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 sorrir mais frequentemente. Ela se encontra inativa em muitos indivíduos cronicamente infelizes. O mRNA isolado a partir de vários diferentes indivíduos infelizes da mesma família revelou a ausência de um segmento interno de 173 nucleotídeos, que estava presente no mRNA Lacheinmal isolado dos membros felizes da mesma família. As sequências do DNA dos genes *Lacheinmal* de membros felizes e infelizes dessa família foram determinadas e comparadas. Essas sequências diferiam em apenas um nucleotídeo, que se encontrava em um íntron. 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 7-25 para identificar quais das seguintes sequências nucleotídicas codificarão uma sequência de arginina-glicina-ácido aspártico:

- 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 de 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.
- Ligações de hidrogênio que se formam quando o tRNA está no sítio A.
- Quebradas pela movimentação do ribossomo ao longo do mRNA.

QUESTÃO 7-11

Liste as definições comuns encontradas em dicionário para os termos *replicação*, *transcrição* e *tradução*. Ao lado de cada definição, 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 outro mundo, um código de tripletes é usado, mas a sequência dos nucleotídeos não é importante, somente importando saber quais nucleotídeos estão presentes. Quantos aminoácidos esse código genético poderia determinar? Você poderia imaginar algum problema referente à tradução desses códigos?

QUESTÃO 7-13

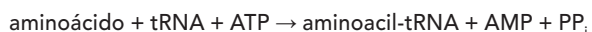
Uma característica impressionante do código genético é o fato de aminoácidos que apresentam propriedades químicas similares frequentemente possuem códons similares. Desse modo, códons com U ou C como segundo nucleotídeo tendem a especificar aminoácidos hidrofóbicos. Você pode sugerir uma explicação plausível 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 mRNA que codifica uma determinada proteína. Uma segunda mutação no DNA da célula leva à alteração de um único nucleotídeo em um tRNA, que 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 PP_i representa o pirofosfato (ver Figura 3-40). No aminoacil-tRNA, o aminoácido e o tRNA estão ligados por uma ligação covalente 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 da 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. Você pode sugerir a etapa suplementar que direcionaria a ocorrência da reação completa?

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 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. Quais são as frações 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 ribossômicas 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 dos seguintes tipos de mutações podem ser considerados como deletérios para um organismo? Explique 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.