

Parte 3

Expressão do Gene



6

Transcrição e Regulação Gênica

Marilis do Valle Marques e Regina Lúcia Baldini

Neste capítulo será abordado o processo de transcrição, no qual um RNA é gerado com base em uma sequência de DNA. A síntese do RNA usando o DNA como molde é feita pela enzima RNA polimerase, que inicia a síntese em uma região do DNA chamada promotor, e termina em uma região chamada terminador. A afinidade da enzima pelo promotor pode ser aumentada pela ação de fatores ativadores, ou reprimida pela ação de fatores repressores de transcrição. Enquanto bactérias apresentam apenas uma RNA polimerase, eucariotos apresentam um número maior, com conjuntos de genes-alvos específicos. Os transcritos bacterianos são produzidos já em sua forma final, mas os transcritos primários de eucariotos geralmente passam por várias etapas de processamento desde sua síntese no núcleo até sua tradução no ribossomo. A regulação da expressão gênica torna possível a diferenciação celular e respostas diferentes das células a seu ambiente.

Mecanismos de transcrição

O processo de transcrição consiste na conversão de uma sequência de nucleotídeos do DNA em uma cópia dessa sequência sintetizada agora em uma molécula de RNA. O RNA é uma longa cadeia composta de ribonucleotídeos, e existe principalmente na forma de fita simples, o que possibilita o pareamento intramolecular de bases. Esse pareamento leva os RNA celulares a assumirem muitas formas distintas e interagirem com várias outras moléculas, atestando sua versatilidade. Todas as moléculas de RNA da célula, ou seja, os RNA transportadores (tRNA), os RNA ribossômicos (rRNA), os RNA mensageiros (mRNA) e os RNA regulatórios e catalíticos são obtidos por meio desse processo.

As primeiras evidências de que o RNA é derivado do DNA vieram do trabalho de Volkin e Astrachan em 1956, estudando a infecção da bactéria *Escherichia coli* pelo fago T2. Eles observaram que logo após a infecção havia uma grande síntese de RNA, e que este tinha uma composição de bases similar à do DNA do fago. Neste capítulo serão abordados os detalhes desse processo e como ele é regulado nas bactérias e nos organismos eucarióticos.

Sequências específicas no DNA delimitam unidades de transcrição

O DNA é constituído por uma longa sequência de nucleotídeos, podendo ser organizado em um ou mais cromossomos, ou unidades de replicação independente. A função do DNA é codificar as informações necessárias à manutenção da célula e da integridade do organismo, mas a molécula de DNA não tem atividade própria além de sua manutenção. A informação do DNA tem de ser convertida na molécula de RNA, e esta, por sua vez, poderá exercer as diferentes funções, diretamente como RNA mesmo ou codificando as proteínas celulares.

Embora boa parte do DNA seja importante para codificar essas informações, nem toda essa sequência é essencial. Como a célula pode, então, distinguir quais partes dessa sequência têm significado e devem ser transformadas em RNA? A resposta está na própria sequência. Os sistemas enzimáticos que realizarão a síntese do RNA reconhecem sequências específicas no DNA para iniciar o processo a partir daquele ponto, e outras sequências que informam quando o processo deve terminar.

Uma unidade de transcrição é delimitada por uma sequência no DNA chamada promotor, na qual se inicia a transcrição, e uma sequência chamada terminador, na qual se encerra a transcrição (Figura 6.1). Do promotor ao

terminador, uma única fita de RNA é sintetizada, correspondendo a essa região do DNA. Após sua síntese, esse RNA ainda pode sofrer processamentos que o modificam em tamanho ou sequência, dependendo de sua função, como será visto adiante.

Síntese de RNA é dividida em três etapas bem definidas

A transcrição pode ser dividida em três etapas: iniciação, elongação e terminação. A enzima que realiza a síntese do RNA tendo o DNA como molde é a RNA polimerase, descoberta em 1960 por Samuel Weiss e Jerard Hurwitz. As RNA polimerases bacterianas são distintas das enzimas eucarióticas, pois as bactérias apresentam uma única RNA polimerase, enquanto os eucariotos dispõem de pelo menos três RNA polimerases nucleares distintas, além de uma RNA polimerase mitocondrial, mais semelhante à de bactérias. Procariotos do domínio Arqueia apresentam apenas uma RNA polimerase com estrutura e mecanismo de ação mais parecida com a RNA polimerase II de eucariotos do que com a RNA polimerase bacteriana. As RNA polimerases de bactérias e as três polimerases conservadas em eucariotos serão abordadas separadamente mais adiante.

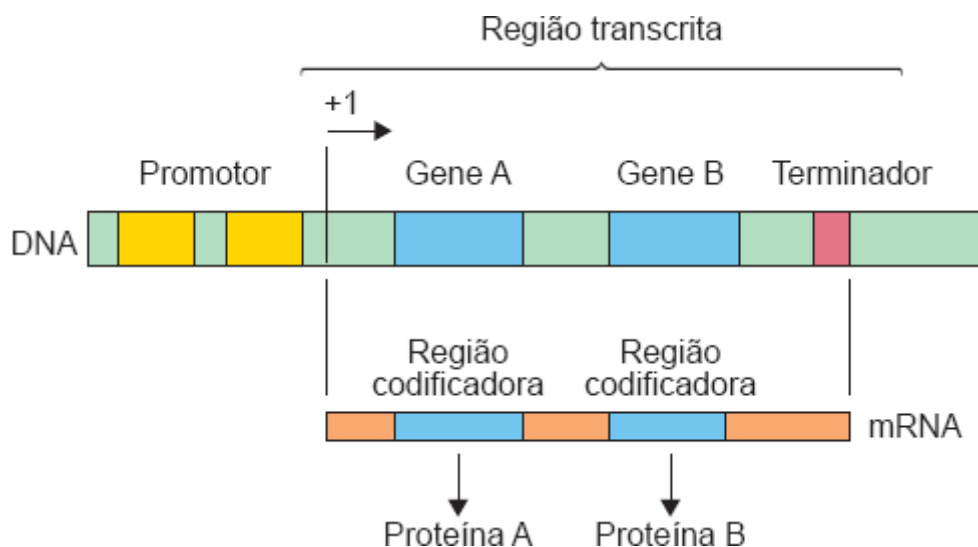


Figura 6.1 Regiões principais identificadas em um operon bacteriano. Um promotor bacteriano composto por duas sequências direciona a transcrição a partir da posição +1, e o mRNA produzido estende-se até um terminador. Dentro do transcrito pode haver uma ou mais regiões codificadoras, que tendem a ser reconhecidas pelos ribossomos de maneira independente, dando origem a cada proteína.

A transcrição em bactérias começa com o reconhecimento da sequência do promotor no DNA pela RNA polimerase, que se liga a essa região promotora formando um complexo fechado de transcrição (Figura 6.2). Na sequência, a enzima abre as fitas do DNA para possibilitar o reconhecimento da sequência de nucleotídeos, formando o chamado complexo aberto de transcrição. A região de DNA aberta é chamada bolha de transcrição. Uma vez abertas as fitas, a enzima começa a colocar os primeiros ribonucleotídeos e ocorre, então, a transição para a 2ª fase, a elongação.

Na elongação da transcrição, a RNA polimerase prossegue adicionando e ligando os ribonucleotídeos sequencialmente, gerando a molécula de RNA, sem se dissociar deste até o final. A RNA polimerase da bactéria *E. coli* tem uma taxa de elongação *in vivo* de 20 a 50 nucleotídeos por segundo a 37°C. Assim que uma RNA polimerase sai do promotor e entra na fase de elongação, outra enzima pode se ligar ao promotor e iniciar uma nova transcrição. Esse fato torna possível que RNA necessários em grande quantidade, como os RNA ribossômicos, sejam gerados rapidamente na célula. Durante a elongação, o movimento da RNA polimerase pelo DNA causa uma torção da dupla-hélice. O avanço da enzima aumenta o superenovelamento positivo à frente, e causa um desenovelamento do DNA atrás da bolha de transcrição. Essas alterações são corrigidas pela ação das topoisomerases de DNA (ver Capítulo 3).

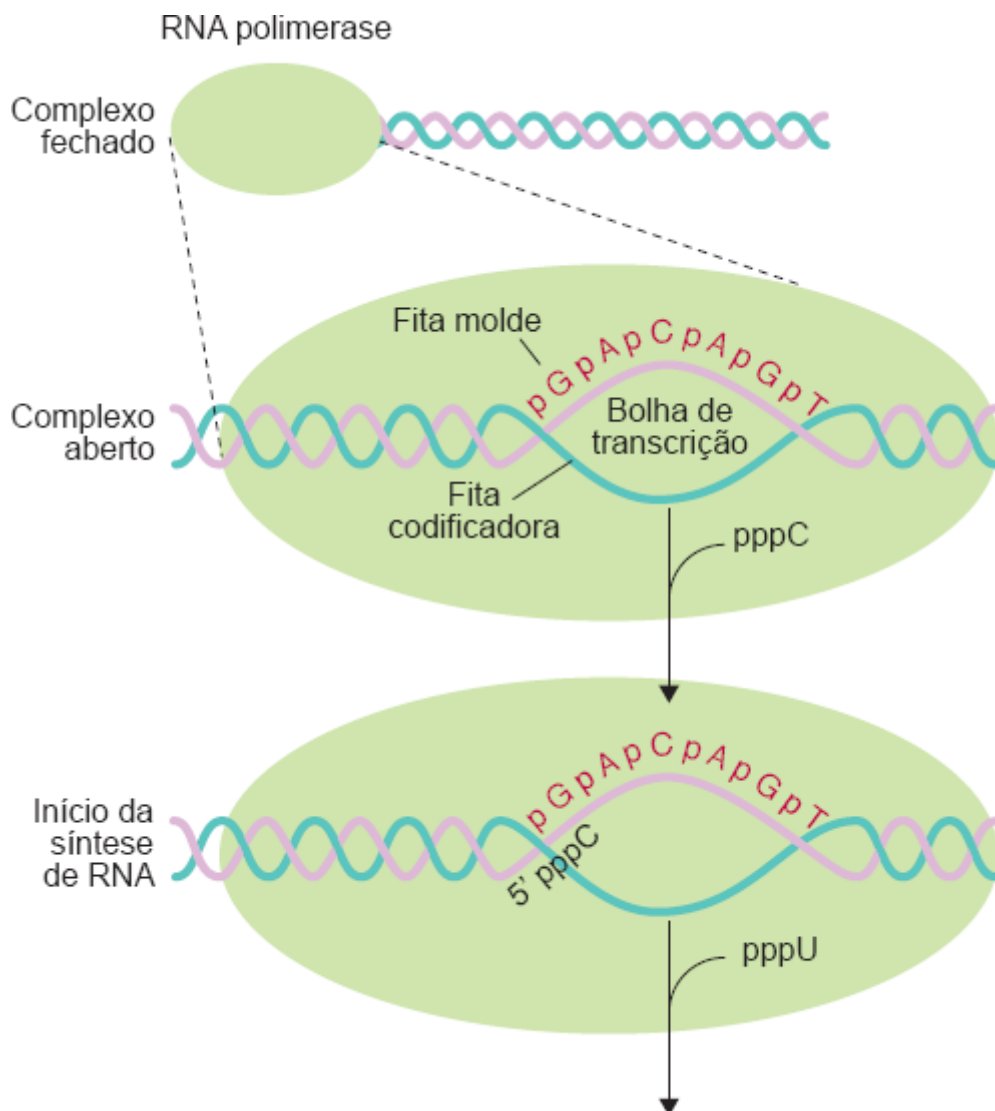


Figura 6.2 Etapas do início da transcrição. A RNA polimerase reconhece as sequências do promotor e se liga ao DNA, formando um complexo fechado de transcrição. Na sequência, a enzima abre as fitas do DNA, formando a chamada bolha de transcrição, e os primeiros ribonucleotídeos são colocados a partir da posição +1. A RNA polimerase movimenta-se mantendo as fitas do DNA separadas, e usa a fita-molde para colocar os ribonucleotídeos corretos em posição. A fita codificadora não participa do processo. A fita de RNA cresce pela adição dos ribonucleotídeos na extremidade 3', o que faz com que o mRNA seja sintetizado na direção 5'-3'.

A elongação da transcrição prossegue até que a RNA polimerase reconheça sinais que levem à dissociação do complexo de transcrição. Em bactérias, esses sinais são bem conhecidos, e podem ou não depender de proteínas acessórias. Em eucariotos, existe mais de um modelo de terminação de transcrição, e o ponto de término é mais impreciso. As particularidades de cada sistema serão abordadas posteriormente.

RNA polimerases usam DNA como molde para a síntese de RNA

Como já discutido, apenas partes do DNA são usadas como codificadoras da informação que deve ser passada aos RNA. Assim, a RNA polimerase posiciona-se no promotor de um gene para iniciar a sua transcrição, e segue sempre a partir daquele ponto. Mas como a enzima sabe qual das duas fitas do DNA é a que contém a informação?

A resposta está na estrutura da molécula de DNA, que apresenta cadeias antiparalelas, uma no sentido 5'-3' e a outra no sentido 3'-5'. A sequência reconhecida pela enzima como promotor encontra-se em uma das fitas, e é unidirecional, ou seja, a RNA polimerase prossegue somente daquele ponto na direção 3' da fita em que se encontra o promotor.

A RNA polimerase reconhece, então, a fita em que se encontra o promotor como a fita cujo código deve ser copiado. Essa é a chamada fita codificadora, e sua sequência é a mesma a ser transcrita no RNA, com a substituição de timinas por uracilas (ver Figura 6.1). Para sintetizar o RNA com essa sequência, a enzima faz uso da outra fita, chamada fita-molde, como informação para a colocação dos ribonucleotídeos na ordem correta. Assim, ao abrir as

fitas do DNA, a fita-molde é encaixada em uma fenda da RNA polimerase, e a enzima sintetizará o RNA colocando os ribonucleotídeos complementares aos dessa fita.

RNA polimerases catalisam a formação de ligações fosfodiéster entre ribonucleotídeos

A RNA polimerase tem uma propriedade muito importante: ela consegue iniciar a polimerização de uma fita de RNA sem nenhum iniciador, ou seja, colocando o 1º nucleotídeo. Essa capacidade é essencial na transcrição e também na replicação do DNA, pois a DNA polimerase não dispõe dessa capacidade, e apenas estende as fitas com uma extremidade 3'OH livre (necessitando, portanto, de um iniciador – *primer*). Assim, uma RNA polimerase chamada primase realiza a síntese de pequenos trechos de RNA para a DNA polimerase estender (ver Capítulo 3). No caso da transcrição, o 1º nucleotídeo é colocado a poucas bases de distância do promotor, sempre respeitando a complementaridade de bases com a fita-molde, e essa posição é chamada de posição +1.

A RNA polimerase catalisa a síntese do RNA usando como substrato ribonucleotídeos contendo um grupo trifosfato (ATP, UTP, CTP ou GTP) na posição 5'. Após a colocação do 1º nucleotídeo, o próximo nucleotídeo trifosfato a ser incorporado é ligado covalentemente ao 1º, e assim sucessivamente até o final da transcrição. A ligação entre os ribonucleotídeos ocorre com a remoção de 2 dos 3 grupos fosfato do NTP, e a ligação do 1º grupo fosfato (o fosfato alfa) do 2º nucleotídeo ao grupo OH da posição 3' do 1º (Figura 6.3). Esse tipo de ligação é chamado ligação fosfodiéster, pois o grupo fosfato forma uma ligação do tipo éster entre os dois nucleotídeos. Como a ligação dos próximos nucleotídeos ocorre na posição 3' na cadeia, a fita de RNA aumenta sempre na direção 5'-3', como toda a polimerização de ácidos nucleicos.

Transcrição em bactérias

Os genes bacterianos podem se organizar de duas maneiras quanto à transcrição: genes transcritos sozinhos (chamados monocistrônicos) ou mais de um gene cotranscrito (policistrônicos). Quando mais de um gene é transcrito em um único RNA, esse conjunto de genes é chamado de operon, e as proteínas codificadas por eles costumam ser relacionadas em suas funções, em geral uma mesma via metabólica da célula. A estrutura do operon foi identificada pela primeira vez pelos pesquisadores François Jacob e Jacques Monod para o operon de uso da lactose de *E. coli*, em 1961 (ver boxe “A descoberta da regulação gênica”). Esses pesquisadores ganharam o Prêmio Nobel por suas descobertas sobre a regulação desse operon, processo que será explicado mais adiante neste Capítulo. Uma das vantagens desse tipo de arranjo é que genes que realizam uma função comum podem ser corregulados, facilitando sua expressão conjunta.

Promotor bacteriano é composto por duas regiões discretas

Sejam monocistrônicos ou policistrônicos, todos os genes são transcritos a partir de uma região chamada promotor. Essa é a região reconhecida pela RNA polimerase para se ligar, abrir as fitas e iniciar a síntese do RNA. O promotor bacteriano é composto por duas regiões distintas no DNA, uma localizada cerca de 10 pb a montante do sítio de início de transcrição (chamada região –10) e outra localizada a cerca de 35 pb a montante do sítio de início de transcrição (chamada região –35) (Figura 6.4 A). O sítio de início de transcrição, no qual o 1º nucleotídeo do RNA é adicionado, é chamado de +1 (não há zero nessa escala, a posição antes de +1 é –1). A região –35 é o sítio de ligação reconhecido pela RNA polimerase para se ligar ao DNA. A região –10 geralmente é rica em A e T, e é o sítio no qual são abertas as fitas do DNA (o par A–T é mais fácil de ser separado que G–C, por conter apenas duas ligações de hidrogênio entre as bases). Além dessas, a região a montante da posição –35 também pode ter papel na estabilização da interação DNA–RNA polimerase.

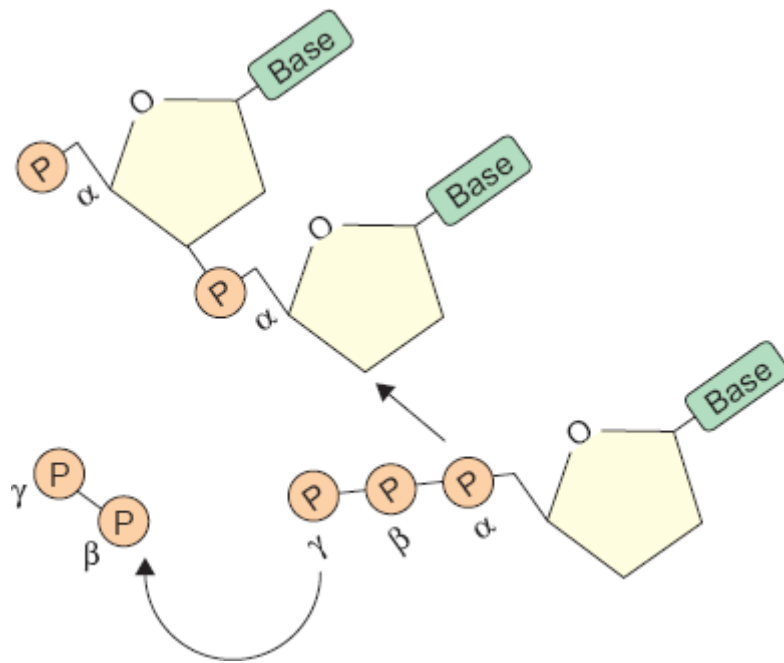


Figura 6.3 Incorporação dos nucleotídeos na cadeia. A RNA polimerase usa ribonucleotídeos trifosfato como substratos e, ao incorporá-los na cadeia nascente de RNA, remove os dois grupos fosfato mais externos (chamados fosfato beta e gama) na forma de pirofosfato. O fosfato mais interno (α) será usado na estrutura do RNA para fazer a ponte fosfodiéster entre os açúcares.

O promotor é sempre unidirecional, o que significa que as sequências -35 e -10 definem a direção da transcrição e a fita a ser transcrita. A RNA polimerase liga-se somente de uma maneira ao promotor, e a transcrição sempre ocorrerá na direção da posição $+1$ em diante.

Embora todos os genes sejam transcritos a partir de um promotor, existem promotores que recrutam mais eficientemente a RNA polimerase do que outros. Os promotores eficientes são chamados promotores fortes; e os menos eficientes, promotores fracos. O que define se um promotor é forte ou fraco são as sequências -10 e -35 , de maneira que, quanto mais a sequência de bases for bem reconhecida pelas subunidades da RNA polimerase, mais eficiente é o promotor.

Fatores sigma são necessários para identificar o promotor

RNA polimerase bacteriana é composta de duas subunidades alfa (40 kDa cada), uma subunidade beta (155 kDa), uma beta' (160 kDa), uma ômega (10 kDa), que são constantes, e um fator variável chamado sigma (Figura 6.4 B). As subunidades alfa, beta, beta' e ômega formam o cerne da RNA polimerase, que é a parte da enzima que realmente catalisa a síntese de RNA. Os fatores sigma são as subunidades que identificam as sequências do promotor e fazem com que a RNA polimerase se ligue nessa região. Cada fator sigma distinto reconhece um promotor de sequência diferente, o que significa que, variando o fator sigma, a célula pode controlar a expressão de jogos de genes distintos a cada momento.

A elongação da transcrição requer apenas o cerne da RNA polimerase

A RNA polimerase contendo o fator sigma é chamada holoenzima. A holoenzima liga-se ao promotor do gene e realiza a abertura das fitas do DNA entre as posições -10 e $+1$ para dar início à síntese do RNA. Entretanto, o fator sigma dificulta o movimento da RNA polimerase, por estar ligado ao promotor. Ao iniciar a síntese do RNA, a RNA polimerase separa-se do fator sigma e somente o cerne da RNA polimerase prossegue para a etapa de elongação da transcrição (Figura 6.5). A elongação da transcrição não ocorre de maneira uniforme: várias pausas na elongação ocorrem durante o caminho, e essas pausas têm papel também na regulação desse processo, como será visto posteriormente.

Em 2011 foi comemorado o cinquentenário da publicação do trabalho de François Jacob e Jacques Monod sobre o sistema de regulação da expressão dos genes de utilização da lactose em *E. coli*. Esse trabalho foi uma revisão de seus resultados anteriores, na qual propuseram que haveria genes regulatórios, codificadores de fatores que regulavam a expressão de outros genes de maneira coordenada. Esses fatores foram propostos como sendo inibidores da expressão, e foram chamados repressores (os ativadores só foram descritos mais tarde). O modelo de repressores explicava a regulação da expressão da enzima betagalactosidase e também o controle da lisogenia do fago lambda, dois objetos de pesquisa genética da época. Entretanto, nesse trabalho eles ainda não conheciam a natureza química dos repressores, e cogitaram ser RNA ou proteínas, pois ambos poderiam exercer essa função. Hoje já se sabe que RNA regulatórios são extremamente importantes e agem regulando a estabilidade do mRNA e sua tradução, tanto em procariotos como em eucariotos. O trabalho de Jacob e Monod previa essa possibilidade já há 50 anos. Nesse trabalho também foi lançada a ideia de operon, que seria uma série de genes que participariam na mesma via biossintética e seriam coregulados. Alguns anos depois, um trabalho desses pesquisadores definiu a existência do promotor, o sítio no qual se inicia a transcrição (Jacob *et al.*, 1964).

A publicação do trabalho de 1961 estabeleceu um paradigma revolucionário, haja vista que até então somente se conhecia a regulação da atividade ou síntese de proteínas, como enzimas. O reconhecimento de sua importância foi imediato, e poucos anos depois, em 1965, François Jacob, André Lwoff e Jacques Monod receberam o Prêmio Nobel de Fisiologia e Medicina.

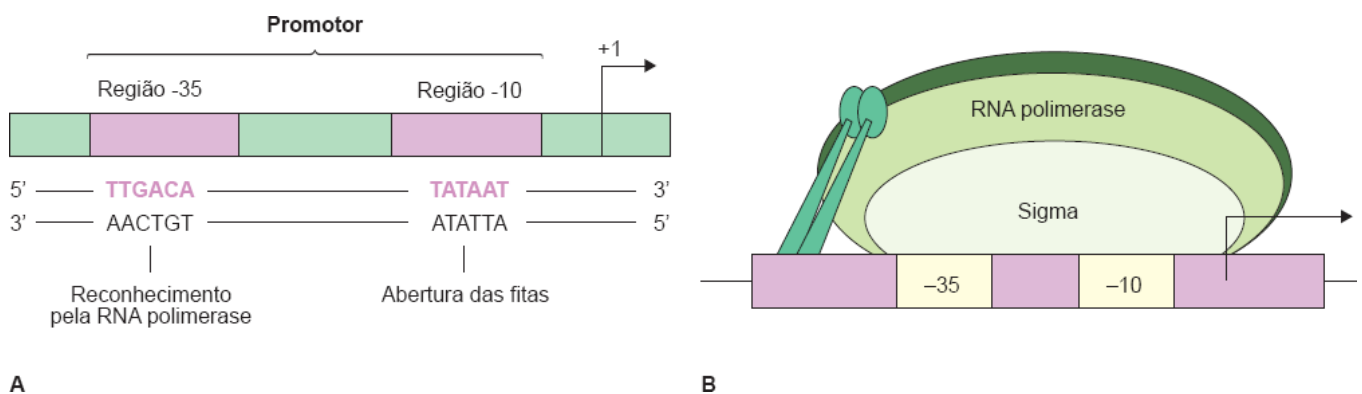


Figura 6.4 Estrutura de um promotor bacteriano típico. **A.** O promotor bacteriano é composto de duas sequências, uma localizada em torno da posição -35 , e outra, da posição -10 . A região -35 é o principal sítio de reconhecimento pela RNA polimerase, enquanto a posição -10 é o sítio de abertura das fitas. **B.** RNA polimerase é composta de duas subunidades alfa, uma beta, uma beta' e uma subunidade variável sigma. A subunidade sigma reconhece as sequências do promotor e posiciona a enzima. As subunidades alfa apresentam dois domínios, com o domínio aminoterminal contatando as subunidades beta e beta' e o domínio carboxiterminal reconhecendo sequências adicionais a montante do promotor, auxiliando a localização da enzima no DNA.

Outro aspecto importantíssimo a ser considerado sobre esse processo é que, por não apresentar membrana nuclear, o DNA é transcrito em mRNA e este é simultaneamente traduzido em proteínas. À medida que o mRNA surge da RNA polimerase que está caminhando no DNA, ele é imediatamente ligado por ribossomos que iniciam a síntese proteica (Figura 6.5 C). Assim, cada região codificadora no DNA pode gerar muitas cópias de transcrito, e a partir de cada transcrito são sintetizadas muitas cópias de proteína, amplificando o processo.

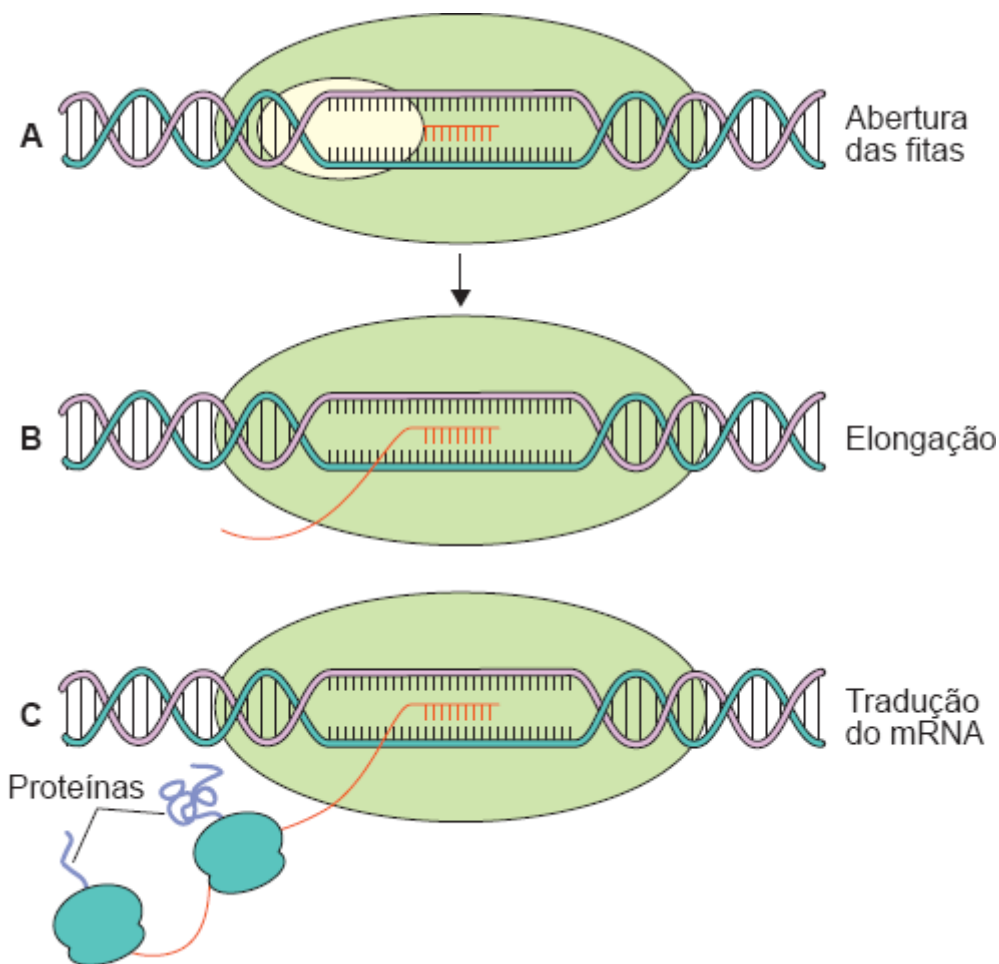


Figura 6.5 Elongação da transcrição e acoplamento com a tradução. Após a adição de alguns ribonucleotídeos (**A**), a RNA polimerase pode liberar a subunidade sigma (**B**) e realizar a elongação da transcrição somente com o cerne da enzima. Em bactérias, a transcrição da mensagem e sua leitura pelos ribossomos ocorrem simultaneamente, pela ausência de uma membrana nuclear (**C**).

Terminação da transcrição pode ou não requerer uma proteína terminadora

A elongação da transcrição prossegue até determinado ponto, que a RNA polimerase reconhece como sítio de término da transcrição, chamado terminador. Mas como a enzima sabe onde terminar? Existem dois tipos de sistemas para a terminação da transcrição em bactérias: um dependente de uma proteína, e outro, independente. Primeiramente será detalhado o sistema que não depende de uma proteína, a chamada terminação intrínseca.

Na terminação intrínseca, a informação para a RNA polimerase terminar a transcrição não está no DNA, mas sim no RNA. De fato, algumas sequências podem formar estruturas de grampes se houver suficiente complementaridade de bases dentro da mesma molécula (Figura 6.6 A). Essas estruturas de grampo são formadas assim que o RNA é transcrito, e causam uma pausa longa no movimento da RNA polimerase. Se logo após essa estrutura de grampo houver muitas uracilas no RNA, a ligação DNA–RNA fica muito instável, pois o pareamento A–U é bastante fraco. Toda essa situação faz com que a ligação entre RNA polimerase–DNA–RNA se desfça, terminando a transcrição. Assim, o terminador intrínseco é composto por uma estrutura de grampo seguida de várias uracilas no RNA.

Embora esse sistema seja bastante simples, ele é usado em boa parte dos genes bacterianos, e mesmo assim muitos outros genes não apresentam esse tipo de terminador. A transcrição desses genes, então, deve ser terminada por outro sistema, que requer o fator de terminação de transcrição Rho. A proteína Rho organiza-se como um hexâmero de subunidades iguais que forma uma estrutura em forma de anel aberto. Esse fator reconhece uma sequência de ligação no RNA chamada de sítio *rut*, posicionado um pouco antes do local em que a transcrição deve terminar. Ao mesmo tempo em que a RNA polimerase está estendendo o RNA, Rho liga-se ao *rut*, e hidrolisa ATP para obter energia para se mover sobre o RNA até encontrar a RNA polimerase no sítio de terminação. A RNA polimerase tende a fazer uma pausa longa nesse sítio, e, ao encontrar o complexo de transcrição, Rho causa sua dissociação, terminando a transcrição (Figura 6.6 B).

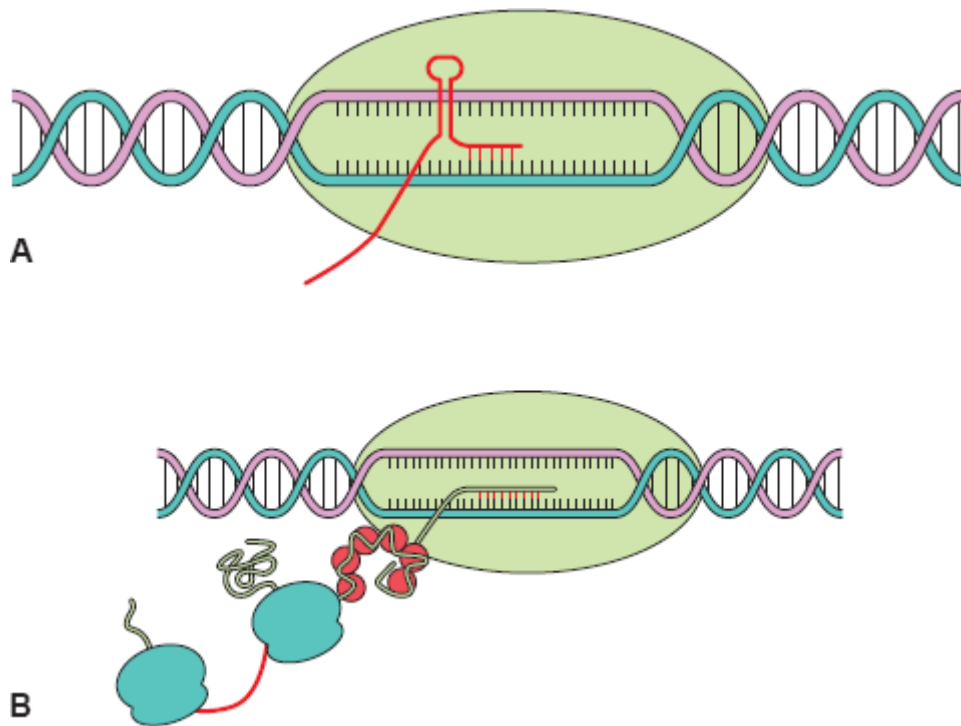


Figura 6.6 Estrutura de terminadores de transcrição. **A.** Um terminador intrínseco ocorre quando há formação de uma estrutura de alça e haste (formando um grampo) no mRNA seguida de várias uracilas. Essa estrutura, dentro do complexo DNA-RNA polimerase-mRNA, causa a separação da RNA polimerase e o término da transcrição. **B.** Um terminador dependente de Rho ocorre quando é transcrita uma sequência específica no mRNA chamada *rut*, que é reconhecida pela proteína Rho. Rho liga-se ao mRNA e percorre o mRNA na direção 5'-3', até alcançar a RNA polimerase e promover a dissolução do complexo de transcrição.

Regulação da transcrição em bactérias

As bactérias têm a economia como princípio básico. A expressão gênica é um processo que demanda uma quantidade razoável de energia, portanto só deverá ocorrer na situação em que houver necessidade daquele produto gênico. O início da transcrição é o ponto de maior regulação desse processo, pois, se não há necessidade do produto na célula, a transcrição do gene nem se inicia.

A regulação da transcrição é exercida em dois níveis: nas sequências regulatórias presentes no DNA (chamados fatores em *cis*, ou na mesma molécula, ou “próximos”) e nos fatores regulatórios que reconhecem essas sequências (chamados fatores em *trans*, ou podendo vir de outra molécula, ou “distantes”). A combinação desses fatores em *cis* e em *trans* possibilita a expressão regulada temporal e espacialmente, como será visto a seguir.

Fatores sigma competem pelo cerne da RNA polimerase

O primeiro mecanismo para se regular o início da transcrição é o uso de diferentes fatores sigma ligados ao cerne da RNA polimerase, como visto anteriormente. Os fatores sigma competem pelo cerne, que existe em quantidade limitada na célula e, portanto, a concentração de determinado sigma na célula pode definir qual conjunto de genes será coordenadamente expresso naquele momento. A célula pode alterar rapidamente quais genes serão transcritos simplesmente aumentando a expressão de um sigma e diminuindo a de outro. O conjunto de genes regulados por um mesmo fator de transcrição é chamado de regulon. Assim, o aumento da concentração do sigma B e a diminuição da concentração do sigma A na célula leva ao aumento da expressão do regulon do fator B e diminuição da expressão do regulon do fator A.

Fatores reguladores da transcrição auxiliam ou dificultam a ligação da RNA polimerase ao promotor

Outro mecanismo muito empregado para regular o início da transcrição é o uso de proteínas que auxiliam ou dificultam a ligação da RNA polimerase ao promotor. Essas proteínas são chamadas de fatores de transcrição, podendo aumentar a expressão (chamadas de ativadores) ou diminuir a expressão (chamadas de repressores), como se descreve a seguir.

Os promotores fracos podem ser transcritos mais eficientemente com a ajuda de fatores ativadores, que são proteínas que se ligam próximo ao promotor e ajudam a RNA polimerase a se ligar e iniciar a transcrição. Os fatores ativadores costumam se ligar a montante da região -35 , e podem fazer contato com a subunidade alfa ou sigma da RNA polimerase (Figura 6.7 A e B). A ligação do fator ativador ao seu sítio no DNA pode ocorrer sozinha, mas em alguns casos o ativador pode precisar se ligar a outra molécula para se tornar apto a ligar-se ao DNA. Essas pequenas moléculas são chamadas indutores, pois estimulam (induzem) a transcrição. A ligação dos ativadores é muito específica, isto é, as proteínas reconhecem uma sequência de nucleotídeos definida para se ligar, o que limita o número de genes do seu regulon.

A regulação gênica em bactérias também costuma fazer uso de proteínas que impedem a transcrição, chamadas de repressores. Os repressores tendem a se ligar à jusante da posição -35 , ou seja, na mesma posição de ligação da RNA polimerase, e impedem a ligação desta. Em alguns casos, podem se ligar à jusante da região $+1$, e impedir o movimento da RNA polimerase (Figura 6.7 C). O sítio em que um repressor se liga é chamado de operador, e também é uma sequência específica para aquele repressor. Assim como com os ativadores, o repressor pode necessitar da ligação de uma pequena molécula para poder se ligar ao DNA. Essa pequena molécula é chamada, então, de corepressor.

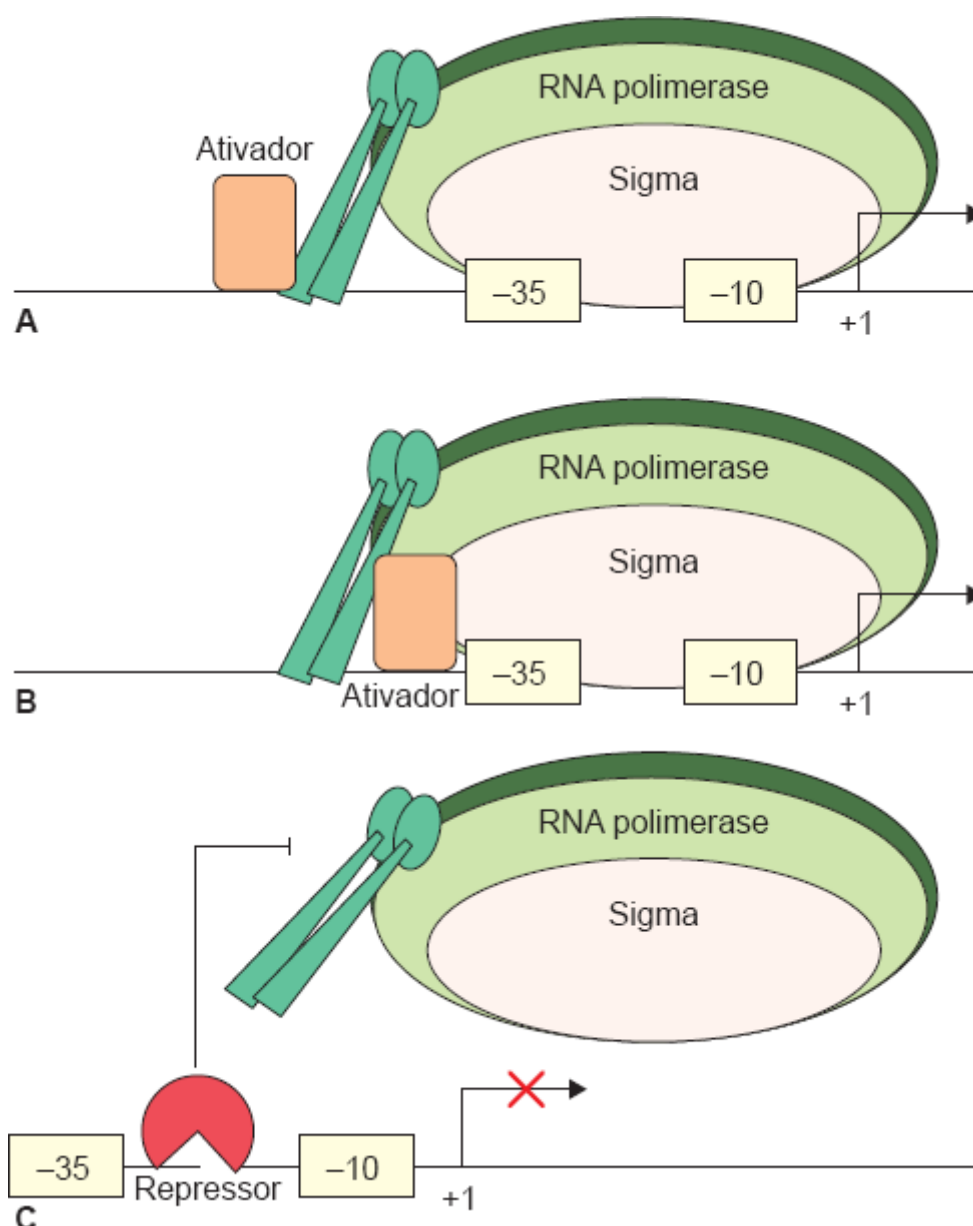


Figura 6.7 Posicionamento de fatores regulatórios da transcrição. Os ativadores de transcrição posicionam-se contatando a RNA polimerase, seja a subunidade alfa (A) ou sigma (B), e auxiliando a estabilização da RNA polimerase no promotor. Os repressores de transcrição, em sua maioria, reconhecem sequências (operador) sobrepostas ao promotor (C) e, ao se ligarem, impedem a ligação da RNA polimerase.

Quando pode ser usado o mecanismo de ativação ou repressão mediado por pequenas moléculas? Considere, por exemplo, um operon que codifique enzimas de uma via metabólica que degrade determinado substrato para usar o produto. Quando a expressão desse operon é necessária? Apenas quando houver o substrato a ser metabolizado! Assim, o substrato pode se ligar a um ativador e ajudá-lo a ligar ao DNA ou se ligar a um repressor e tirá-lo do operador. Em qualquer desses casos, o resultado é o estímulo da transcrição desse operon, ou a indução dos genes.

Agora considere outro operon que codifique enzimas para a síntese de um composto do qual a célula precise. Quando a expressão desse operon é necessária? Apenas quando não houver o produto na célula! Assim, o produto pode se ligar a um repressor, ajudando-o a se ligar (como um correpressor), ou pode se ligar a um ativador e removê-lo do DNA. Em qualquer dos casos, o resultado é a diminuição da transcrição desse operon, ou a repressão dos genes.

Operon *lac* é regulado simultaneamente por ativação e repressão da transcrição

O mecanismo de regulação gênica do operon *lac* foi o primeiro sistema de regulação estudado pelos pesquisadores Jacob e Monod na década de 1960, e a compreensão desse sistema promoveu um grande avanço no estudo da genética. O operon *lac* consiste em três genes, que juntos tornam possível que a célula bacteriana use o açúcar lactose como fonte de carbono e energia (Figura 6.8). O primeiro gene é o *lacZ*, que codifica a enzima betagalactosidase, o segundo gene (*lacY*) codifica uma proteína de transporte de lactose, e o terceiro (*lacA*) codifica uma enzima que modifica a lactose para ser internalizada. A enzima betagalactosidase cliva o dissacarídeo lactose nos monossacarídeos glicose e galactose, e a célula usa a glicose como fonte de energia.

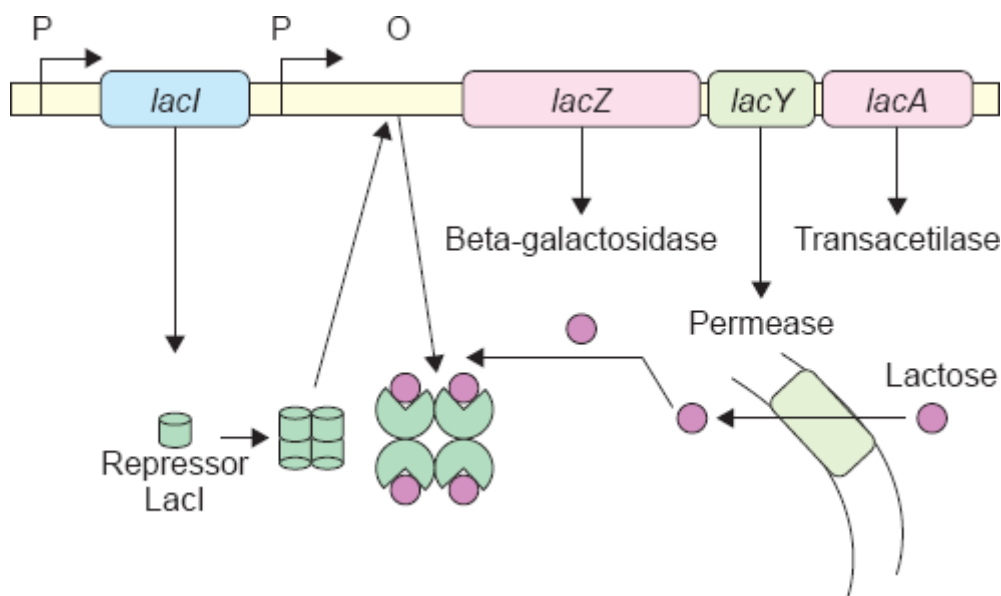


Figura 6.8 Esquema da região do operon *lac*. O operon *lac* é composto por três genes (*lacZ*, *lacY* e *lacA*) que são transcritos a partir do promotor P. A montante desse operon está o gene que codifica o repressor LacI. O repressor LacI pode ligar em operadores (O) no operon *lacZYA*, impedindo a transcrição. Quando há lactose no meio, ela se liga ao repressor LacI e leva ao desligamento do operador, promovendo o acesso da RNA polimerase.

Como discutido anteriormente, o *lac* é um operon de degradação de lactose para obter glicose, então a célula deve expressar esses genes quando o substrato (que é a lactose) estiver disponível. O que ocorre é que, quando não há lactose no meio, a expressão está reprimida pela ligação de um repressor (a proteína LacI) ao operador. O repressor LacI é um tetrâmero formado por subunidades iguais, que pode ligar em dois sítios operadores simultaneamente. Quando há lactose no meio, ela se liga ao repressor LacI e causa seu desligamento do operador, promovendo o acesso da RNA polimerase (Figura 6.8). Isso ocorre porque a ligação de lactose causa uma mudança de conformação da proteína LacI que separa os tetrâmeros.

A saída do repressor provoca uma indução perceptível do operon *lac*, mas o promotor é muito fraco, e os níveis absolutos de expressão mesmo sem o repressor ainda são baixos. Estudos realizados na década de 1950 mostraram que a glicose inibia a síntese de enzimas para uso de outros açúcares, e esse fenômeno foi chamado de repressão catabólica. Mais tarde foi observado que a repressão catabólica atinge vários outros operons de uso de açúcares, além do operon *lac*.

Como, então, a célula bacteriana percebe os níveis de glicose do meio? Essa resposta foi obtida por meio do estudo do sistema de transporte de glicose em *E. coli* (Figura 6.9). Nessa bactéria, a glicose é transportada pela membrana citoplasmática por um sistema de transporte (chamado de PTS), que fosforila a glicose durante o transporte, transformando-a em glicose-6-fosfato. Enquanto há glicose sendo transportada, o fosfato do sistema PTS é direcionado para a glicose. Quando não há mais glicose disponível, o fosfato passa a ser direcionado para uma enzima, chamada adenilato ciclase. Essa enzima converte ATP em uma molécula de AMP cíclico (Figura 6.9), que se acumula no citoplasma. O AMP cíclico é um coindutor para um ativador chamado CAP (*catabolite activator protein*), ajudando esse ativador a se ligar ao DNA. O ativador CAP ligado ao AMPc ativa a expressão de grande número de genes em resposta à carência de glicose. Esse nucleotídeo modificado é chamado de 2º mensageiro, pois leva a mensagem de uma situação extracelular (a carência de glicose) para dentro da célula. É importante notar que, embora bem estabelecido para *E. coli*, esse sistema de transdução de sinais pode ser diferente em outras bactérias.

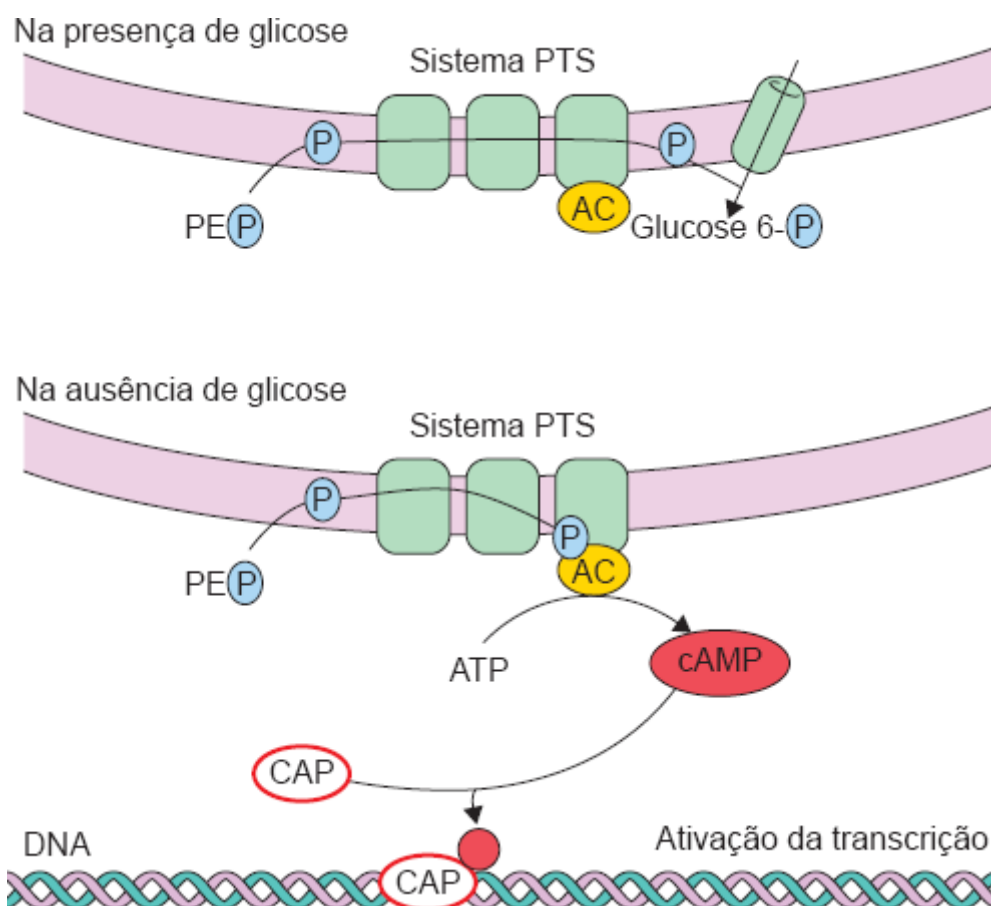


Figura 6.9 Sistema de transporte de glicose em *E. coli*. A glicose é transportada pela membrana citoplasmática pelo sistema de transporte PTS, que fosforila a glicose durante o transporte, transformando-a em glicose-6-fosfato. Nesse processo, a enzima adenilato ciclase está inativa. Quando não há mais glicose disponível, o fosfato passa a ser direcionado para a adenilato ciclase, ativando-a, gerando o 2º mensageiro AMP cíclico. O AMP cíclico liga-se ao ativador CAP (*catabolite activator protein*), ajudando-o a se ligar ao DNA e ativar a expressão de genes em resposta à carência de glicose.

O operon *lac* dispõe, na sua região a montante do promotor, de um sítio de ligação de CAP-AMPc. Quando não há glicose disponível, ativa-se a transcrição desse operon para promover o uso do açúcar lactose como fonte de energia (Figura 6.10). Assim, o operon *lac* sofre um grande aumento de expressão quando existe lactose no meio (que retira o repressor) e quando não existe glicose no meio (que causa a ligação do ativador). Deve-se ressaltar que, mesmo quando reprimido, o operon *lac* ainda apresenta um nível basal de transcrição, já que o sistema de transporte de lactose para dentro da célula precisa estar sempre disponível na membrana citoplasmática. A regulação modula os níveis de expressão de acordo com a disponibilidade de ambos os açúcares no meio.

Operon de síntese do triptofano é regulado na etapa da terminação da transcrição

O operon de síntese do aminoácido triptofano consiste em cinco genes que codificam as enzimas necessárias dessa via (Figura 6.11). O estudo desse operon em *E. coli* mostrou que ele só era expresso quando não havia triptofano no

meio, mas não na presença do aminoácido, devido à ligação de um repressor a um sítio operador. O pesquisador Charles Yanofsky estudou, em 1981, uma linhagem mutante que não apresentava o repressor de transcrição, e observou que mesmo assim ainda ocorria aumento na transcrição quando o triptofano era retirado do meio. Ele observou que mesmo quando havia triptofano no meio, se dava a transcrição de uma região inicial do operon. Na presença do aminoácido, ocorria uma terminação prematura da transcrição antes que os genes codificando as enzimas fossem transcritos, e esse fenômeno foi chamado por Yanofsky de “atenuação de transcrição”. Isso significa que existe um terminador de transcrição no início e outro no final do operon. Mas como a bactéria consegue discriminar qual terminador usar?

Observou-se que na região antes do primeiro terminador, chamada região líder, existe uma fase aberta de leitura que codifica um pequeno peptídeo. Ao final dessa região existe um terminador intrínseco, que forma uma estrutura de grampo (pareando as sequências C e D) seguida de uridinas no mRNA (Figura 6.11). A região antes do terminador apresenta outras 2 sequências (A e B) que podem se parear formando outro grampo. O interessante é que as sequências B e C também podem parear entre si, e as alças A/B e C/D deixam de existir e também, consequentemente, o terminador (Figura 6.11).

Assim, quando na presença de triptofano, as regiões B/C pareiam, o terminador C/D não é formado e a RNA polimerase passa por ele e transcreve o operon. Na ausência de triptofano, as regiões A/B e C/D pareiam, e o terminador atenua a transcrição na região líder. Mas como a presença de triptofano interfere na formação dos grampos?

Para compreender isso, vale lembrar que, em bactérias, a transcrição e a tradução são acopladas! Ao mesmo tempo em que a RNA polimerase está transcrevendo o operon, os ribossomos estão caminhando pelo mRNA. Também é preciso analisar a sequência de códons da região líder. Essa região do mRNA apresenta dois códons para triptofano, localizados ao final da sequência A. Se a célula está com triptofano abundante, os tRNA para triptofano estarão carregados com aminoácido, e os ribossomos passarão rapidamente pelos códons. Entretanto, se não houver muito triptofano disponível, os ribossomos farão pausas longas nesses códons. Se o ribossomo parar nos códons, a sequência A ficará presa, e a sequência B estará livre para parear com C. Isso impede a formação da alça de terminação CD, e a transcrição continua. Se o ribossomo passar rápido por A e B, C ficará livre para parear com D, e haverá a formação do terminador CD, interrompendo a transcrição (Figura 6.12). Assim, a atenuação de transcrição exercida pelos ribossomos só é possível em bactérias, pois em eucariotos esses processos ocorrem em compartimentos distintos da célula. Esse sistema de atenuação pelos ribossomos também é usado em operons que codificam a biossíntese de outros aminoácidos, como fenilalanina e histidina, e suas regiões-líder apresentam vários códons para esses aminoácidos.

Na bactéria Gram-positiva *Bacillus subtilis*, a regulação do operon do triptofano também ocorre por atenuação da transcrição, mas, nessa bactéria, quem regula a formação do grampo de terminação não são os ribossomos. Em *Bacillus*, uma proteína multimérica chamada TRAP (da sigla em inglês: *trp RNA-binding attenuation protein*) é capaz de se ligar à região líder do mRNA somente quando está ligada ao triptofano. A ligação de TRAP ao líder forma o grampo de terminação, impedindo a continuação da RNA polimerase. Quando não há triptofano, a TRAP não é capaz de se ligar ao mRNA e a transcrição continua.

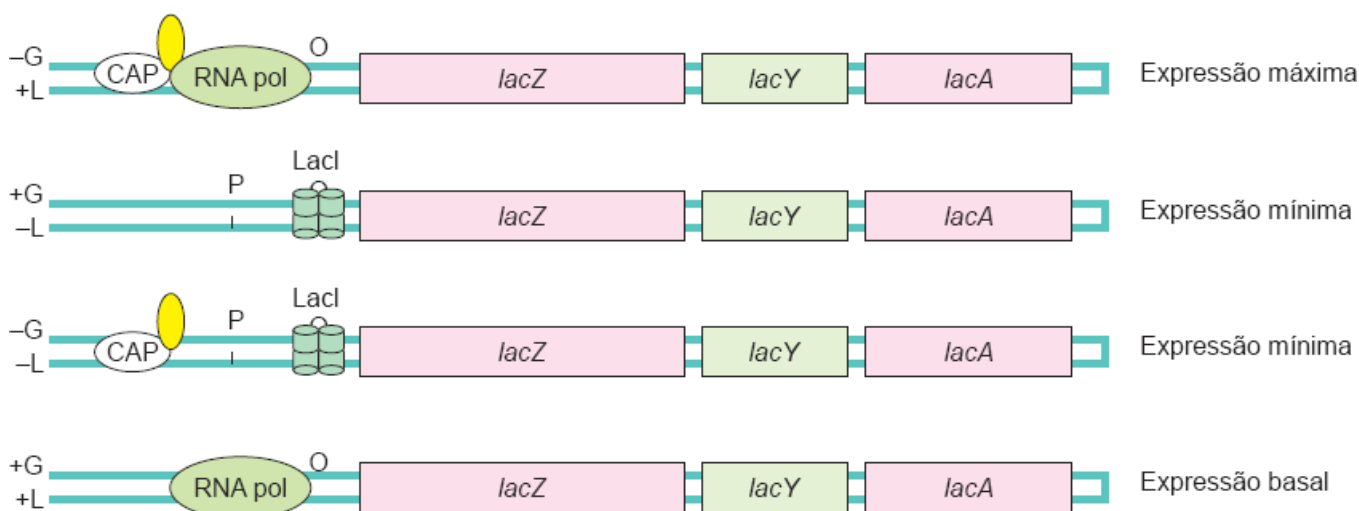


Figura 6.10 Interação dos sinais regulatórios do operon *lac* em *E. coli*. A expressão é modulada pela ligação do repressor LacI e do ativador CAP, em resposta à disponibilidade de lactose (L) e glicose (G). Quando não há lactose no meio, a ligação do repressor LacI possibilita apenas níveis mínimos de transcrição. Quando não há glicose no meio, CAP ligada ao AMPc ativa a transcrição. Assim, na presença de ambos os açúcares, a transcrição ocorre em nível basal, pois a célula prefere usar glicose antes de lactose. Quando não há glicose e há lactose, a expressão é máxima, e é possível usar a lactose.

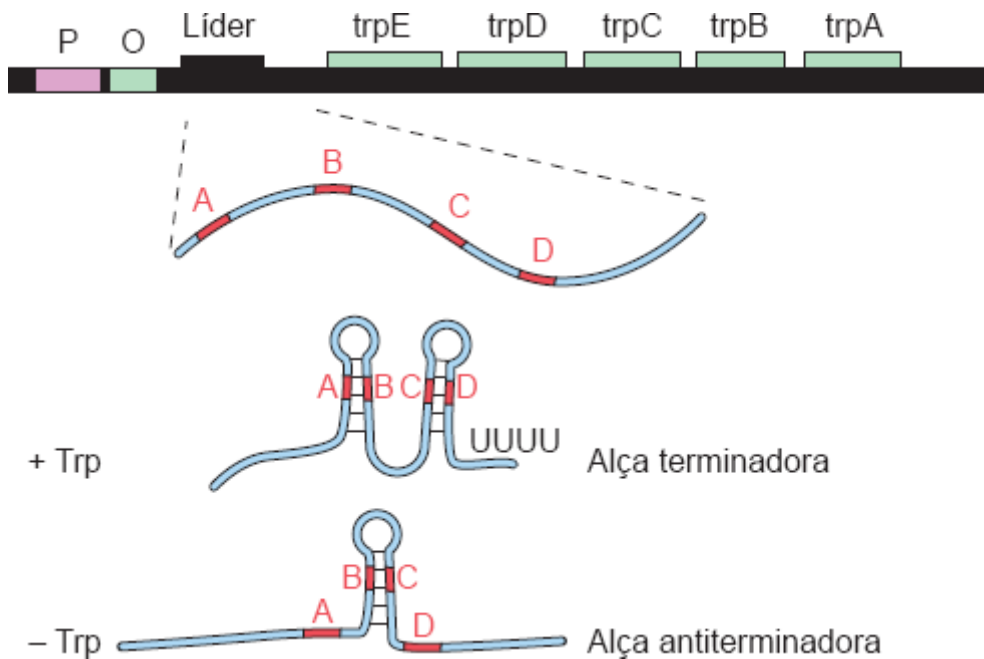


Figura 6.11 Esquema do operon de síntese do triptofano. O operon do triptofano é composto por cinco genes codificando enzimas para a síntese desse aminoácido. A transcrição começa no promotor P e pode estender-se até um terminador intrínseco no início do operon ou até um no final do operon. Um operador O é o sítio de ligação do repressor. A região transcrita até o 1º terminador é chamada de região-líder, e é traduzida em um pequeno peptídeo. Ao final da região líder existem quatro sequências que podem parear formando ou não um terminador intrínseco. Se as sequências A e B pareiam, C e D podem parear e forma-se o terminador. As sequências B e C também podem parear entre si, impedindo a formação das alças A/B e C/D, não formando o terminador. Quando há triptofano, as regiões B/C pareiam, o terminador C/D não é formado, e a RNA polimerase transcreve todo o operon. Na ausência de triptofano, as regiões A/B e C/D pareiam, e o terminador atenua a transcrição na região líder.



Figura 6.12 Papel dos ribossomos na atenuação da transcrição do operon *trp*. A região líder do mRNA apresenta dois códons para triptofano ao final da sequência A. Se há triptofano abundante, os ribossomos passarão rapidamente pelas sequências A e B, e haverá a formação do terminador CD, interrompendo a transcrição. Se não houver muito triptofano disponível, os ribossomos farão pausas longas nos códons, a sequência A ficará coberta pelo ribossomo e a sequência B estará livre para parear com C, impedindo a formação da alça de terminação CD, e continuando a transcrição.

Transcrição em eucariotos

O processo de transcrição dos genes das células eucariotas apresenta diversas peculiaridades em relação à transcrição em bactérias, apesar de o mecanismo de síntese de RNA usando o DNA como molde ser praticamente idêntico. Em primeiro lugar, o DNA está localizado no núcleo das células, enquanto a maioria das moléculas de

RNA desempenha seu papel no citoplasma. Nas células eucariotas, os RNA devem ser transportados do núcleo para o citoplasma. Assim, os processos de transcrição e tradução são completamente separados, ao contrário do acoplamento existente em bactérias. Em segundo lugar, os genes eucariotos que codificam proteínas são frequentemente interrompidos por sequências que não fazem parte da mensagem final, traduzida em proteínas. Essas sequências, denominadas introns, devem ser retiradas na maturação do mRNA, que também sofre outras modificações para garantir sua estabilidade na viagem com destino ao citoplasma. Em terceiro lugar, o DNA dos eucariotos está em um estado de condensação muito mais compacto que o DNA de procariotos, com a presença de nucleossomos e outras proteínas que compõem a cromatina. Para que as RNA polimerases tenham acesso ao DNA, ele precisa ser exposto. O processo de modificação da cromatina é altamente preciso, a fim de que os genes certos sejam expressos no tipo celular e no tempo certos. Finalmente, eucariotos apresentam múltiplas RNA polimerases especializadas na transcrição de classes distintas de genes, diferentemente da única RNA polimerase dependente de DNA presente em procariotos. Entretanto, todas as RNA polimerases apresentam subunidades homólogas entre si, demonstrando uma origem comum e a conservação de estrutura e mecanismos de catálise.

Dada a grande complexidade das RNA polimerases eucariotas e dos fatores de transcrição, como definir se determinada proteína faz parte da RNA polimerase em si ou de fatores acessórios? Para facilitar a leitura das seções subsequentes, serão apresentados alguns conceitos e definições importantes.

Fazem parte das RNA polimerases as subunidades que não se dissociam durante todo o processo de transcrição, da iniciação ao término. Além das subunidades envolvidas diretamente com a catálise da ligação e o reconhecimento do molde e dos substratos, há proteínas estruturais e com função de se associar a proteínas do complexo de transcrição basal e a proteínas regulatórias. Como o cerne da RNA polimerase de bactérias, as RNA polimerases de eucariotos não são capazes de reconhecer as regiões nas quais devem iniciar a transcrição. Para isso, são necessários os fatores de transcrição basal, específicos de cada polimerase e que reconhecem estruturas conservadas nos promotores, formando com a RNA polimerase o complexo de transcrição basal. Esses fatores, com função análoga aos fatores sigma de bactéria, são essenciais para o posicionamento correto das polimerases no promotor, mas se dissociam dela assim que a transcrição propriamente dita se inicia.

Os ativadores de transcrição são proteínas necessárias para aumentar a frequência com que os genes são transcritos e, na maioria dos casos, são também essenciais para a expressão gênica. Esses fatores reconhecem sequências específicas no DNA e podem se associar diretamente a proteínas do complexo basal de transcrição e à própria RNA polimerase, ligando-se a sítios de ativação encontrados nas proximidades do promotor. Alguns ativadores, entretanto, ligam-se a regiões no DNA distantes dos promotores e necessitam de proteínas que funcionam como uma ponte de ligação entre os ativadores e o complexo basal. Essas proteínas intermediárias (as pontes) são conhecidas como coativadores e não se ligam diretamente ao DNA, mas fazem diversas interações entre proteínas. Os coativadores podem apenas servir como intermediários entre proteínas distantes ou, ainda, podem modular a atividade dos ativadores de transcrição por modificações, como fosforilação ou desfosforilação, por exemplo.

Outra classe de complexos proteicos que tem papel importante na transcrição, e que é descrita no Capítulo 10, é composta pelos fatores de remodelamento da cromatina, que tornam os promotores mais ou menos acessíveis ao aparato de transcrição.

Diferentes RNA polimerases transcrevem diferentes classes de genes

Existem pelo menos três RNA polimerases nucleares, denominadas I, II e III, em todos os eucariotos estudados até o momento, além de RNA polimerases presentes em mitocôndrias e cloroplastos. A RNA polimerase I localiza-se no nucléolo e transcreve os genes dos rRNA 28S e 18S, sendo responsável pela maior atividade de síntese de RNA na célula em termos de quantidade de transcritos gerados. As RNA polimerases II e III são distribuídas no nucleoplasma. A RNA polimerase III sintetiza pequenos RNA não codificantes, envolvidos na regulação gênica, no processamento de outros RNA e na tradução, entre eles o rRNA 5S e os tRNA. A RNA polimerase II responde pela maior diversidade de transcritos. Todos o hnRNA (RNA nucleares heterogêneos), que compreendem os mRNA ainda não processados e diversos RNA não codificantes, são sintetizados pela RNA polimerase II. A expressão dessa enorme variabilidade de transcritos deve ser finamente regulada em um organismo com relação ao tempo e tipo celular, sendo de extrema importância a atuação de diferentes proteínas moduladoras da transcrição, como será abordado posteriormente neste capítulo.

Fatores de transcrição basal são necessários para o recrutamento das RNA polimerases aos promotores

Todas as RNA polimerases de eucariotos apresentam-se como complexos de alto peso molecular, formados por várias subunidades. Algumas dessas subunidades, como as que constituem o cerne estrutural e catalítico, são conservadas, sendo equivalentes às subunidades alfa e beta' em bactérias. Diferente da RNA polimerase bacteriana, as RNA polimerases I, II e III não apresentam um fator sigma, que confere especificidade ao promotor. Isso faz com que fatores de transcrição basal sejam essenciais para o recrutamento da polimerase à região promotora dos genes e, portanto, para o início da transcrição. Esses fatores ligam-se a regiões específicas no DNA e interagem com a RNA polimerase cognata, trazendo-a para a região onde a transcrição se inicia. O que faz com que cada RNA polimerase transcreva apenas seu conjunto particular de genes é a combinação das sequências em *cis* nos promotores com os fatores de transcrição basal que recrutam a RNA polimerase comprometida com aquele conjunto de genes. Apesar de essenciais para que ocorra a transcrição *in vivo*, os fatores de transcrição não são considerados parte integrante das RNA polimerases, haja vista que não são necessários após o início da transcrição.

Além das RNA polimerases I, II e III, presentes em todos os eucariotos e descritas com maiores detalhes a seguir, RNA polimerases alternativas foram descritas recentemente. Em plantas, as RNA polimerases IV e V, similares à RNA polimerase II, transcrevem pequenos RNA regulatórios envolvidos com o silenciamento de genes (ver Capítulo 11). A snRNAP-IV é uma RNA polimerase nuclear de mamíferos composta por apenas uma subunidade, semelhante à RNA polimerase mitocondrial e que parece também estar envolvida com funções regulatórias, além de transcrever genes estruturais em alguns tecidos.

Os promotores transcritos pela RNA polimerase I (Figura 6.13) apresentam um cerne e um elemento a montante, o UPE (*upstream promoter element*). O cerne do promotor localiza-se sobreposto ao sítio de início de transcrição, entre as posições -45 e +20, e costuma ser rico em GC, exceto por uma curta região rica em AT próxima ao início de transcrição. O UPE também é rico em GC e compreende a região -180 a -107. É nessa região que a proteína monomérica denominada *UBF* (*UPE-binding factor*) se liga à fenda menor da hélice do DNA, enrolando-o em uma volta de cerca de 360°, e promovendo a ligação de um segundo fator de transcrição basal, o *SL1*.

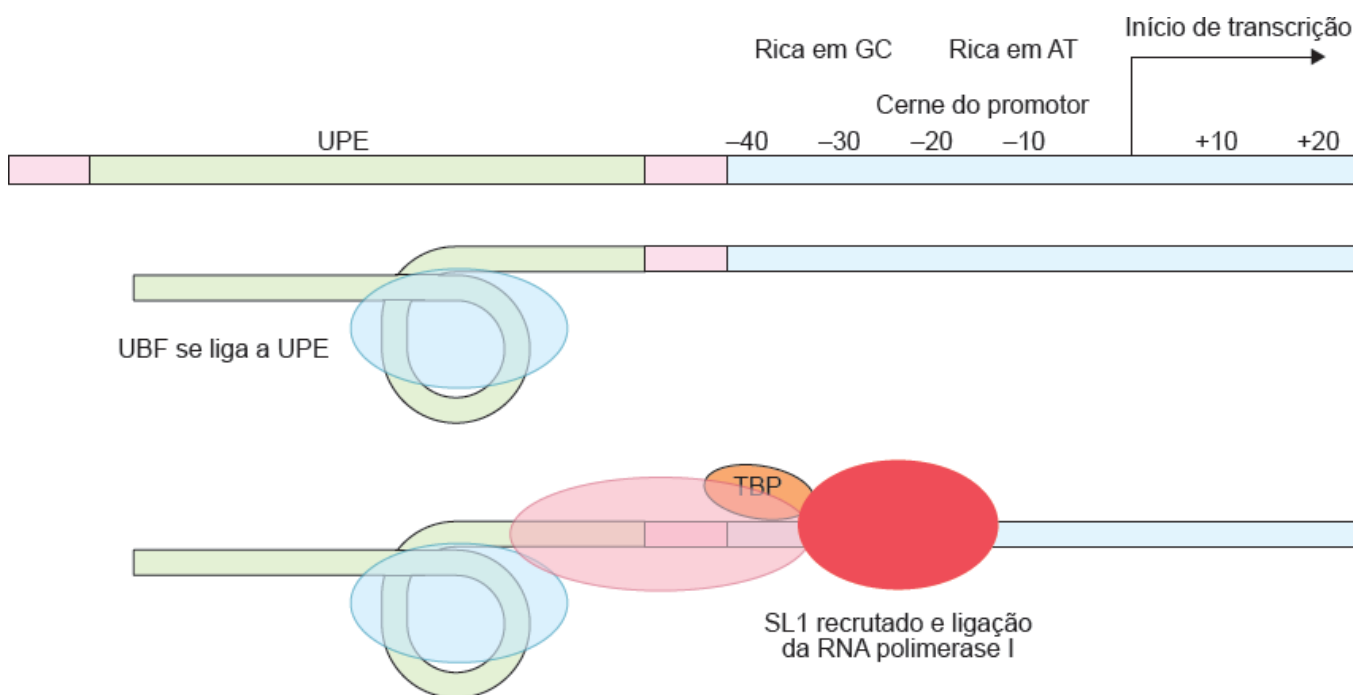


Figura 6.13 Estrutura do promotor de RNA polimerase I e sequência de ligação dos fatores do complexo basal. UBF (azul) é composto por duas subunidades e se liga à região UPE, ocasionando uma torção na dupla-fita do DNA e ligando o fator de posicionamento SL1 (rosa). SL1, do qual a proteína TBP faz parte, reconhece o cerne do promotor e traz a RNA polimerase I para iniciar a transcrição. Adaptada de Lewin *et al.*, 2011.

SL1 reconhece o cerne do promotor e é formado por quatro subunidades, entre elas a TBP, ou *TATA-binding-protein*, que é uma subunidade comum aos fatores de transcrição basal das RNA polimerases I, II e III e será descrita mais adiante (ver destaque na Figura 6.16). SL1 é o fator de posicionamento da RNA polimerase I, que se

associa a ela por meio de TBP, assegurando que a transcrição seja iniciada no ponto correto. Embora SL1 possa exercer sua função na ausência de UBF, a frequência de eventos de início da transcrição é bastante aumentada em sua presença.

Há três tipos de promotores de RNA polimerase III (Figura 6.14): os promotores tipo 1 e tipo 2 são internos, localizando-se na região que será efetivamente transcrita, a jusante do início de transcrição, e os do tipo 3 localizam-se completamente a montante do início de transcrição.

Promotores tipo 1 e 2, que são conservados de leveduras até humanos, são compostos por regiões denominadas boxA, boxB e boxC. BoxA está presente nos 2 tipos, enquanto boxC é característico de promotores do tipo 1, e boxB, do tipo 2. Os genes de 5S rRNA apresentam promotores do tipo 1, nos quais se ligam em sequência os fatores de transcrição basal TFIIIA e TFIIIC às regiões boxA e boxC. Após o desligamento de TFIIIA, o fator de posicionamento TFIIIB é recrutado e, por sua vez, se liga à região do DNA a montante e incluindo o início de transcrição. Após a liberação de TFIIIC, ocorre a ligação da RNA polimerase III a TFIIIB e o início da transcrição.

Promotores do tipo 2 são encontrados à frente de genes presentes em regiões não codificadoras do genoma e genes de tRNA. TFIIIC liga-se às regiões A e B desses promotores. A ligação de TFIIIB, com o recrutamento da RNA polimerase III, e o início da transcrição ocorrem de maneira análoga à dos promotores de classe 1.

TFIIIC é um grande complexo, formado por seis subunidades, enquanto TFIIIA compreende apenas uma proteína com motivo de ligação a DNA. TFIIIB também é composto por várias subunidades, dentre as quais se destaca a TBP, proteína assim nomeada por sua capacidade de ligação à região TATA-box de promotores de RNA polimerase II e III. Entretanto, apesar de fazer parte dos complexos de posicionamento das RNA polimerases I, II e III, a TBP não se liga diretamente ao DNA nos promotores de RNA polimerase I, nem aos promotores tipos 1 e 2 da RNA polimerase III. São outras proteínas desses complexos que fazem o contato com as regiões promotoras.

Promotores do tipo 3 são encontrados em plantas e animais, mas não em organismos unicelulares, como leveduras, tendo surgido durante a evolução da multicelularidade. Esses promotores assemelham-se aos da RNA polimerase II (descritos mais detalhadamente a seguir), contendo uma região TATA-box na qual a TBP efetivamente se liga, na posição -30 do início de transcrição, como parte de TFIIIB (Figura 6.14). Esses promotores, encontrados à frente de um conjunto de genes de RNA não codificadores, compreendem também um elemento distal (DSE) a cerca de 200 a 250 nucleotídeos a montante do início de transcrição e um elemento proximal (PSE) na posição -50. Esses elementos são sítios de reconhecimento de proteínas (Oct-1, Staf e PTF) e aumentam a transcrição a partir desses promotores.

A RNA polimerase II é responsável pela transcrição de todos os precursores de mRNA, entre outros. Dada a variabilidade dos transcritos gerados e da especificidade de muitos deles em relação ao tipo celular, ao estado fisiológico e à resposta da célula aos mais variados estímulos, os promotores de RNA polimerase II são os que apresentam maior diversidade em relação a fatores de transcrição necessários para a expressão de genes específicos, chamados mais genericamente de ativadores de transcrição. Esses ativadores são fundamentais para o recrutamento dos fatores de transcrição basal e, conseqüentemente, da própria RNA polimerase e serão abordados no item a seguir.

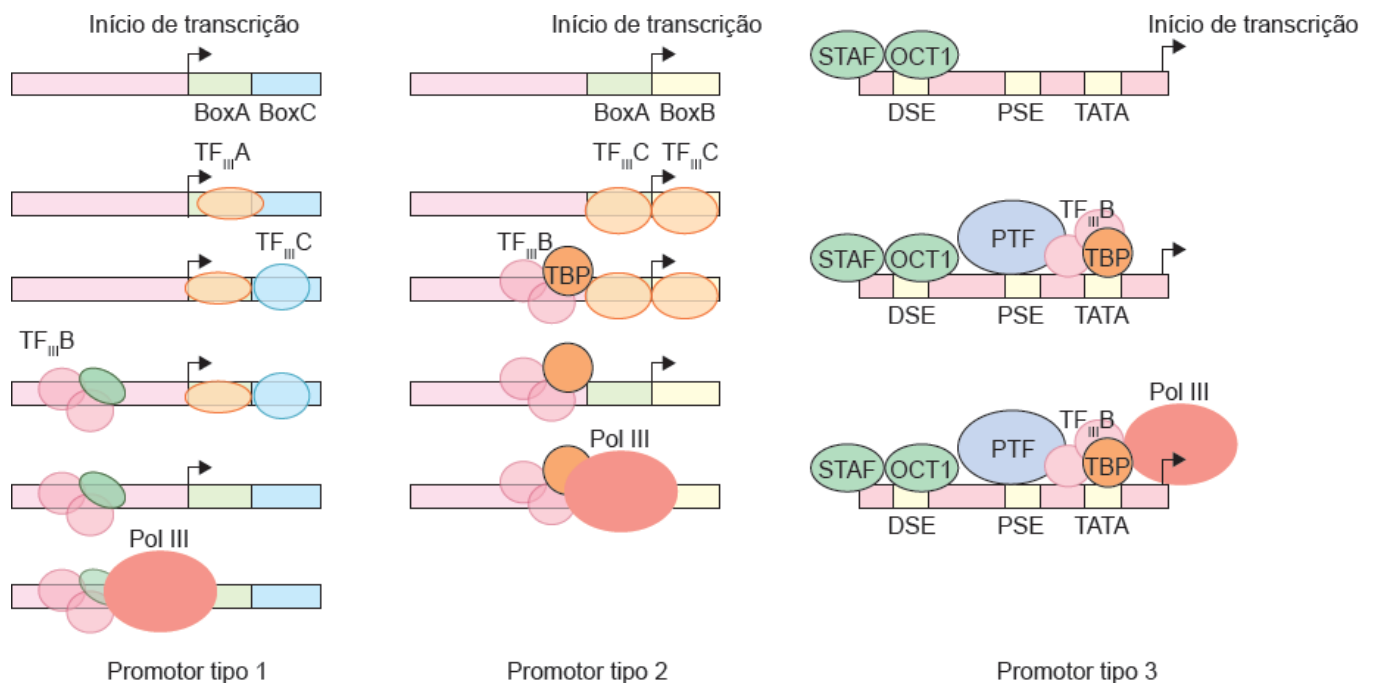


Figura 6.14 A. Estrutura dos promotores internos de RNA polimerase III e sequência de ligação dos fatores do complexo basal. As regiões de ligação dos complexos de montagem (TFIIIA para promotores tipo 1 e TFIIIC para tipo 2) localizam-se dentro da região a ser transcrita. O fator de posicionamento TFIIIB liga-se em boxB após os complexos de montagem já estarem no DNA e recrutam a RNA polimerase III para o promotor. Apesar de TBP fazer parte de TFIIIB, esta subunidade não se liga diretamente ao DNA. **B.** Estrutura dos promotores tipo 3 RNA polimerase III e fatores do complexo basal. Nos promotores tipo 3, todos os elementos de ligação dos fatores de transcrição basal localizam-se a montante do início de transcrição. A ligação de STAF e Oct1 ao sítio DSE aumenta fortemente a transcrição desses promotores, cooperando com a ligação do complexo PTF ao sítio PSE, que por sua vez interage com TFIIIB. A proteína TBP nesse tipo de promotor liga-se diretamente ao elemento TATA, posicionando a RNA polimerase III e possibilitando o início da transcrição. Adaptada de Lewin *et al.*, 2011; Dumay-Odelot *et al.*, 2010.

Apesar da grande diversidade de promotores de RNA polimerase II, é possível distinguir algumas regiões conservadas na maioria deles, ainda que sejam pequenas (Figura 6.15). Os promotores mínimos para a RNA polimerase II, ou seja, aqueles que atendem aos requisitos básicos para o início da transcrição, apresentam a região iniciadora (Inr), com consenso Py₂CAPy₅ (em que Py representa uma pirimidina), sendo A o início da transcrição. Existem os promotores com sequências TATAA, o TATA-box, geralmente na posição -25, relativo ao início de transcrição. É na sequência TATA-box que se liga a proteína TBP (*TATA binding protein*), parte do grande fator de transcrição basal TFIID (Figura 6.16).

TBP, que também faz parte dos complexos de posicionamento das RNA polimerases I e III, é uma pequena proteína dimérica em forma de sela que se associa à fenda menor da hélice do DNA em sua porção interna, ficando a externa exposta para se associar a outras proteínas (Figura 6.16) e faz parte, junto com outras 14 subunidades, do fator de transcrição basal TFIID, que reconhece os promotores da RNA polimerase II. Ao se ligar a TBP, o DNA sofre uma curvatura de cerca de 80°. Essa alteração estrutural desempenha um papel importante na organização do complexo de transcrição, facilitando a interação entre proteínas que se associam a diferentes motivos no DNA e tornando possível que a transcrição se inicie em um ponto específico.

Entretanto, grande parte dos promotores não apresentam TATA-box e geralmente contém outro elemento *cis*, denominado DPE (*downstream promoter element*), localizado entre +28 a +32 do início de transcrição. Nesses casos, outras subunidades de TFIID reconhecem as sequências Inr e/ou DPE, posicionando a RNA polimerase II no promotor, e a TBP não faz contato direto com o DNA. Esses promotores não apresentam um único início de transcrição bem definido, mas sim vários inícios próximos uns aos outros.

Além de TFIID, outros fatores de transcrição basal são necessários para o início de transcrição (Tabela 6.1 e Figura 6.15). Esses fatores ligam-se ao DNA e/ou a outros fatores em uma ordem determinada, recrutando a RNA polimerase, ativando-a. Dentre esses, podem-se destacar as funções de TF_{II}B, que determina a polaridade do promotor, e de TF_{II}H, que fosforila a subunidade maior da RNA polimerase II em sua porção carboxi-terminal (CTD, *carboxiterminal domain*), promovendo a saída da polimerase do promotor e a elongação. O CTD da RNA polimerase II é composto por repetições de uma sequência rica em serinas, prolinas e tirosinas, sendo uma

característica peculiar dessa polimerase e possibilitando sua fosforilação em diversos resíduos de aminoácidos. Diferentes estados de fosforilação do CTD modulam a interação com proteínas envolvidas no início, na elongação e no término da transcrição, bem como no processamento e transporte do mRNA, de maneira ordenada e regulada por diferentes proteínas quinases.

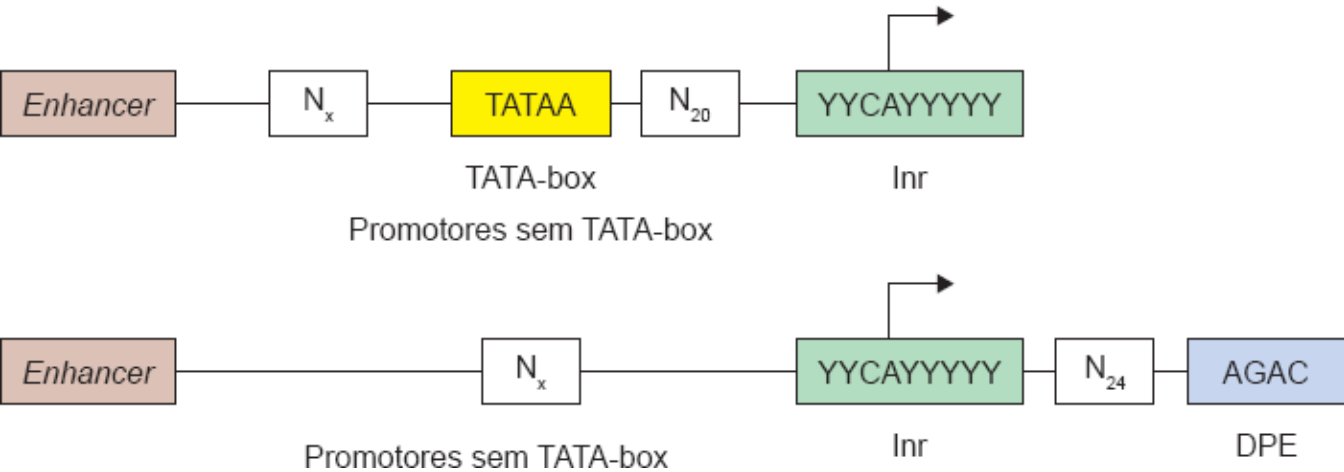


Figura 6.15 Elementos básicos de promotores de RNA polimerase II. Os promotores de RNA polimerase II contêm o elemento Inr, no qual efetivamente se inicia a transcrição e pode ou não conter uma TATA-box a montante do início e/ou um elemento DPE dentro da região transcrita. Os sítios de ligação de ativadores de transcrição, sequências denominadas *enhancers*, variam em número, sequência e localização. Adaptada de Lewin *et al.*, 2011.

Tabela 6.1 Comparação entre as RNA polimerases I, II e III.

RNA polimerase		Alvos de transcrição	Fatores de transcrição do complexo basal	Elemento no DNA (posição em relação ao início de transcrição)	Comentários
RNA polimerase I		rRNA (18S, 28S e 5.8S)	SL1 (contém TBP) UBF	Cerne (–45 a +20) UPE (–180 a –107)	Maior abundância de transcritos (cerca de 80%) Localizada no nucléolo
RNA polimerase II		mRNA e RNA não codificantes	TFIID (TBP) TFIIA, B, F e H	TATA-box (–25) Inr (+1) Várias outras sequências ativadoras	Maior variabilidade dos transcritos e das regiões promotoras; menor abundância de transcritos (3 a 5%)
RNA polimerase III		tRNA, rRNA 5S, pequenos RNA regulatórios	–	–	–
Tipo de promotor	Tipo 1	–	TFIIIA, TFIIB e TFIIC	boxA boxC	–
	Tipo 2	–	TFIIIA, TFIIB e TFIIC	boxA boxB	–
	Tipo 3	–	TFIIB (TBP), Oct-1, Staf e PTF	TATA-box (–30) DSE (+200 a +250) PSE (–50)	Promotores tipo 3 presentes apenas em eucariotos multicelulares

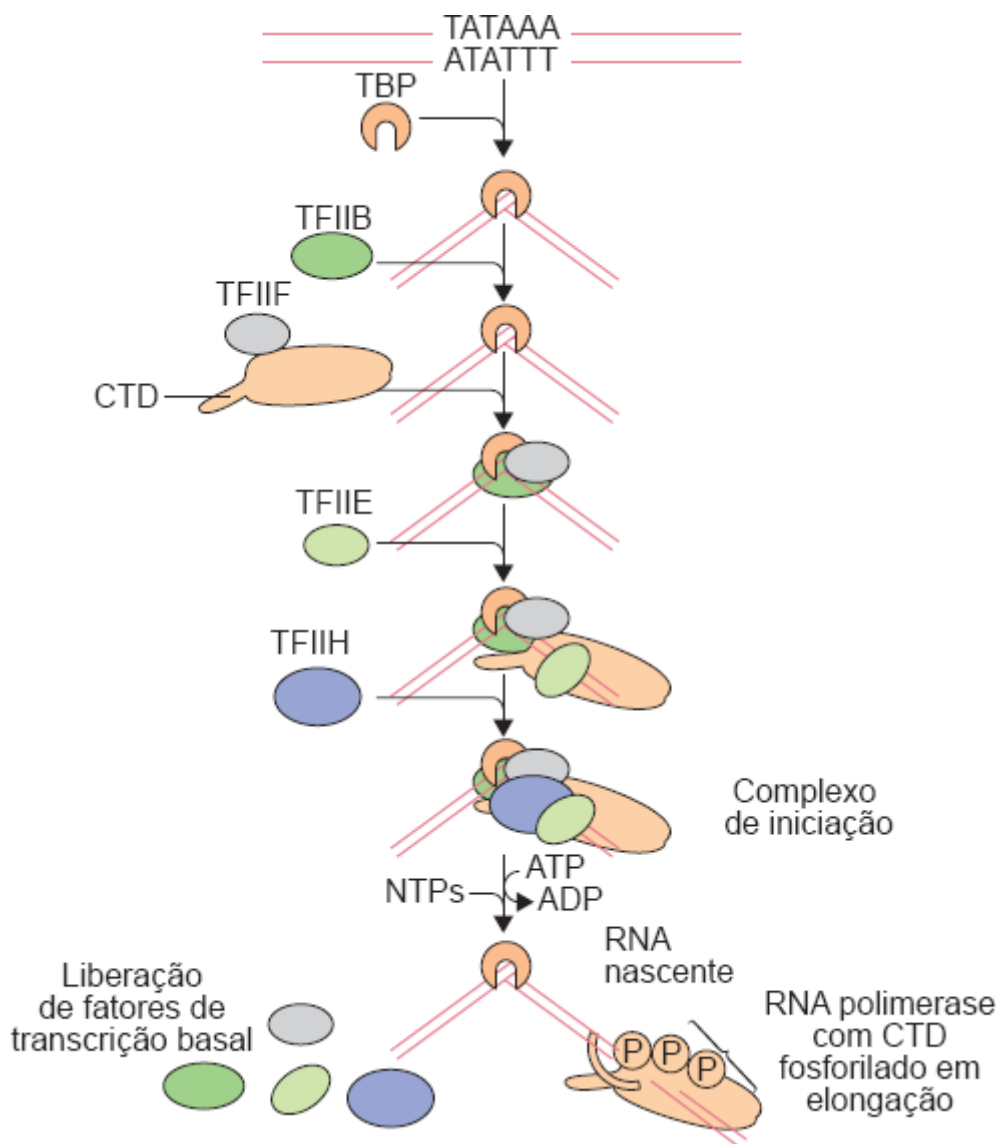


Figura 6.16 Ligação sequencial dos fatores que compõem o complexo de transcrição basal da RNA polimerase II. A montagem do complexo de transcrição basal se inicia com a ligação de TFIID, do qual a proteína dimérica TBP faz parte. TBP (*TATA-binding protein*) liga-se na fenda menor do DNA, na região conhecida como TATA-box nos promotores do tipo 3 da RNA polimerase III e em promotores da RNA polimerase II. A ligação de TBP causa uma curvatura de 80° no DNA, aproximando sítios que estão distantes na sequência e, conseqüentemente, favorecendo a interação entre proteínas que se ligam a esses sítios. Enquanto a região interna da proteína liga-se ao TATA-box, a externa mantém-se disponível para interagir com outras proteínas do complexo basal de transcrição. A transcrição tem início após a ligação da RNA polimerase II e da fosforilação de seu CTD por TFIIH. Os fatores de transcrição basal desligam-se e a RNA polimerase II pode prosseguir com a elongação. Adaptada de Lodish, 2008.

Os mRNA de eucariotos sofrem processamento e modificação

Enquanto os mRNA de bactérias podem ser policistrônicos, ou seja, podem conter informações que serão traduzidas em mais de um peptídeo, os mRNA de eucariotos carregam a informação necessária para apenas uma cadeia polipeptídica. Ainda assim, os transcritos eucariotos são frequentemente muito mais longos que o necessário para conter a informação essencial para sua tradução em proteínas. Isso se deve principalmente à presença de introns, que são sequências que não codificam proteínas, presentes entre segmentos que contêm os códons a serem traduzidos, denominados exons. Para que a informação correta seja lida nos ribossomos, os introns devem ser removidos ainda no núcleo da célula, durante o processo de *splicing*. A jornada de um pré-RNA (transcrito primário) do núcleo para o citoplasma, já na forma de mRNA maduro, envolve, ainda, outras modificações que possibilitam que este seja protegido da ação de nucleases, garantindo que a informação chegue intacta ao aparato de tradução (Figura 6.17). Essas modificações são reguladas no tempo e no espaço e dependem de várias proteínas, que são engajadas nos processos de modo que um passo aconteça à medida que o anterior se completa.

Assim que um gene vai sendo transcrito, a extremidade 5' do RNA, que retém o grupo 5' trifosfato, deve ser prontamente protegida para evitar a ação de nucleases e para que a elongação possa continuar após uma breve

caminhada da polimerase sobre o molde de DNA. Um nucleotídeo de guanina é ligado por uma ligação 5'-5' a esta extremidade do RNA nascente pela enzima guanilil transferase e é metilado pela ação de uma metiltransferase. Essa estrutura, denominada *cap* (Figura 6.17, no destaque), além de proteger a extremidade do RNA da ação de exonucleases, é um sítio de reconhecimento sequencial de proteínas que participarão do *splicing* e posteriormente do transporte e da tradução do mRNA. A adição de *cap* também estimula a retirada do 1º intron do pré-mRNA.

O segundo passo de processamento do pré-mRNA é o *splicing*, ou a remoção dos introns, por um complexo formado por ribonucleoproteínas denominado spliceossomo (Figura 6.18). Em suas extremidades, os introns apresentam pequenas sequências consenso, com um GU conservado na ponta 5', a extremidade doadora e um AG na ponta 3', que corresponde à extremidade aceptora. O primeiro passo para o *splicing* é a quebra da ligação fosfodiéster do sítio doador e da ligação da ponta 5' gerada no intron no 2'-OH de uma adenina no interior do intron, formando uma estrutura em forma de laço (*lariat*) (Figura 6.18). No segundo passo, a extremidade 3'-OH livre do sítio doador ataca a ligação no sítio aceptor, religando os exons e liberando o intron, ainda em forma de laço. Durante todo o processo, o spliceossomo mantém contato com o RNA, possibilitando sua organização espacial. Outras proteínas são recrutadas pelo spliceossomo e se associam ao mRNA, sendo deixadas como marcas de *splicing* e que terão papel nos passos subsequentes de exportação do núcleo e tradução.

Os processos de transcrição e de processamento são acoplados, garantindo que haja organização temporal dos eventos. Assim, o 1º intron é reconhecido e corretamente retirado antes que os demais sejam sequer transcritos, tornando possível que as extremidades 5' e 3' de diferentes introns não sejam reunidas de maneira aleatória. Entretanto, muitos genes podem gerar mRNA alternativos, pela remoção ou não de determinados introns, em um processo regulado dependendo do tipo celular, gerando produtos proteicos diferentes.

A última etapa de processamento do mRNA é a proteção da extremidade 3' com uma cauda poli-A. Para que isso aconteça, ocorre a clivagem do transcrito a cerca de 10 bp a jusante de uma sequência rica em AU (AAUAAA em mamíferos) pela ação de várias enzimas, entre elas uma endonuclease. Após a clivagem, a cauda poli-A, que consiste de uma longa sequência de adeninas (cerca de 200 resíduos), é adicionada pela enzima poli-A-polimerase (PAP) à extremidade 3'-OH livre.

A terminação da transcrição também é coordenada com a clivagem do mRNA. Um dos modelos prevê que o sinal de poliadenilação no mRNA promove a saída de fatores de alongação e o recrutamento de fatores de terminação. Outra alternativa é o modelo de torpedo, em que uma exonuclease 5' → 3' se acopla à extremidade 5' do mRNA a jusante do sítio de poliadenilação após a clivagem, degradando-o e “expulsando” a RNA polimerase de sua posição ao final do gene, evitando que a transcrição se estenda além dele. Em todos os passos descritos, o transcrito nunca fica desprovido de proteínas, formando grandes complexos nucleoproteicos.

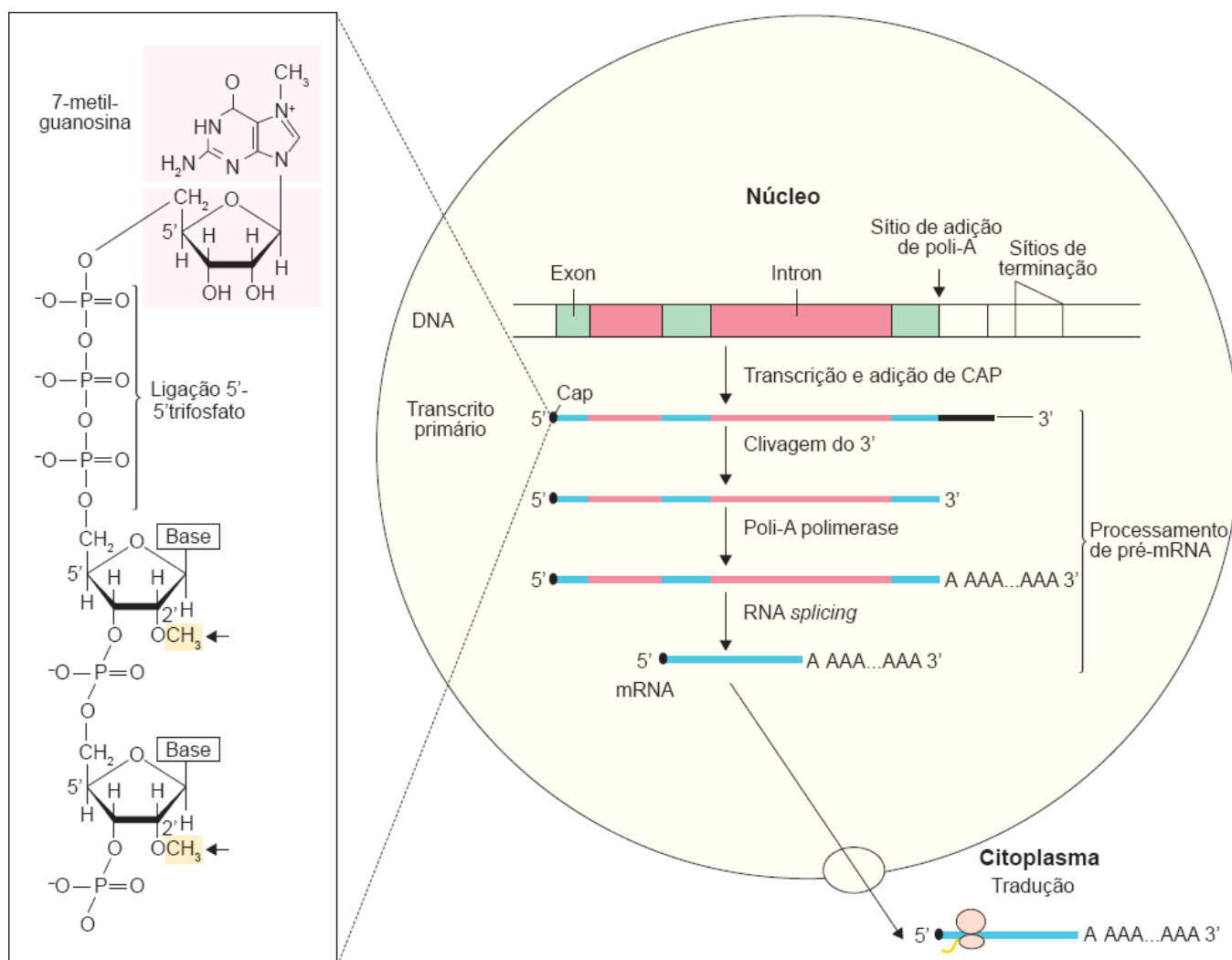


Figura 6.17 Etapas de processamento do RNA. Os genes transcritos a partir do DNA contêm sequências que codificam aminoácidos no produto final (introns) e sequências que devem ser removidas antes da tradução (exons). Após a transcrição e ainda no núcleo, o RNA sofre modificação nas extremidades, com adição do CAP e da cauda poli-A e a retirada dos introns (*splicing*). Apesar de a figura mostrar a etapa de adição da cauda poli-A como anterior ao *splicing*, ela pode ocorrer posteriormente a este, pois o sítio de adição de poli-A pode não ter sido transcrito ainda por ocasião do *splicing* (detalhes no texto). O mRNA maduro será, então, transportado para o citoplasma, no qual será traduzido nos ribossomos. No destaque, a estrutura do CAP, em que um nucleotídeo de 7-metil-guanosina é adicionado no 1º nucleotídeo do transcrito, em uma ligação 5'-5'fosfodiéster. As duas primeiras bases do RNA podem também ser metiladas na posição 2', conforme mostrado na figura (setas). Adaptada de Lodish, 2008.

Terminada a transcrição e o processamento do mRNA, é necessária a sua transferência para o citoplasma através dos poros nucleares, em um mecanismo ainda não totalmente elucidado, que envolve o ancoramento do complexo ribonucleoproteico ao poro, sua translocação propriamente dita e a liberação do mRNA. Proteínas que participam do transporte também interagem com a maquinaria de transcrição e/ou *splicing*, demonstrando mais uma vez a dinâmica e o acoplamento desses processos.

Regulação da expressão de genes transcritos pela RNA polimerase II

O controle da expressão gênica em eucariotos ocorre em diversos níveis e tem uma regulação peculiar a cada tipo celular, estado fisiológico e de desenvolvimento. Essa regulação pode se dar nos níveis de início da transcrição, do processamento do mRNA, no seu transporte para o citoplasma, na tradução da mensagem em proteína e ainda na regulação da atividade e degradação do produto gênico. Nesta seção, será tratado apenas do nível de regulação em nível de início da transcrição pela RNA polimerase II, responsável pela transcrição dos genes que codificam proteínas.

Em eucariotos, os genes estão embebidos no interior da cromatina (ver Capítulo 10), estando geralmente em um estado silencioso, devendo ser expostos para que sejam acessíveis aos ativadores de transcrição e à RNA polimerase

para serem transcritos. De maneira geral, genes em regiões em que o DNA apresenta um alto grau de metilação são menos transcritos, enquanto a acetilação de histonas favorece a transcrição.

Tanto o remodelamento da cromatina como a ligação de ativadores de transcrição ao DNA são necessários para que os genes sejam devidamente expressos. Esses eventos ocorrem de maneira coordenada, sendo a ordem de chegada dos fatores ditada pela natureza de cada região ou de cada promotor específico. Alguns ativadores de transcrição ligam-se ao DNA e recrutam fatores de remodelamento de cromatina, que agem expondo as regiões nas quais o complexo de transcrição basal deve se ligar. Em outros casos, a ação dos fatores de remodelamento é a responsável por promover que os ativadores encontrem seus sítios de ligação ao DNA, recrutando ou ativando o complexo de transcrição basal (Figura 6.19). De qualquer maneira, a ação de diferentes fatores faz com que o início da transcrição seja um passo regulado em vários níveis, assegurando que a expressão gênica seja finamente orquestrada em cada uma das células de um organismo.

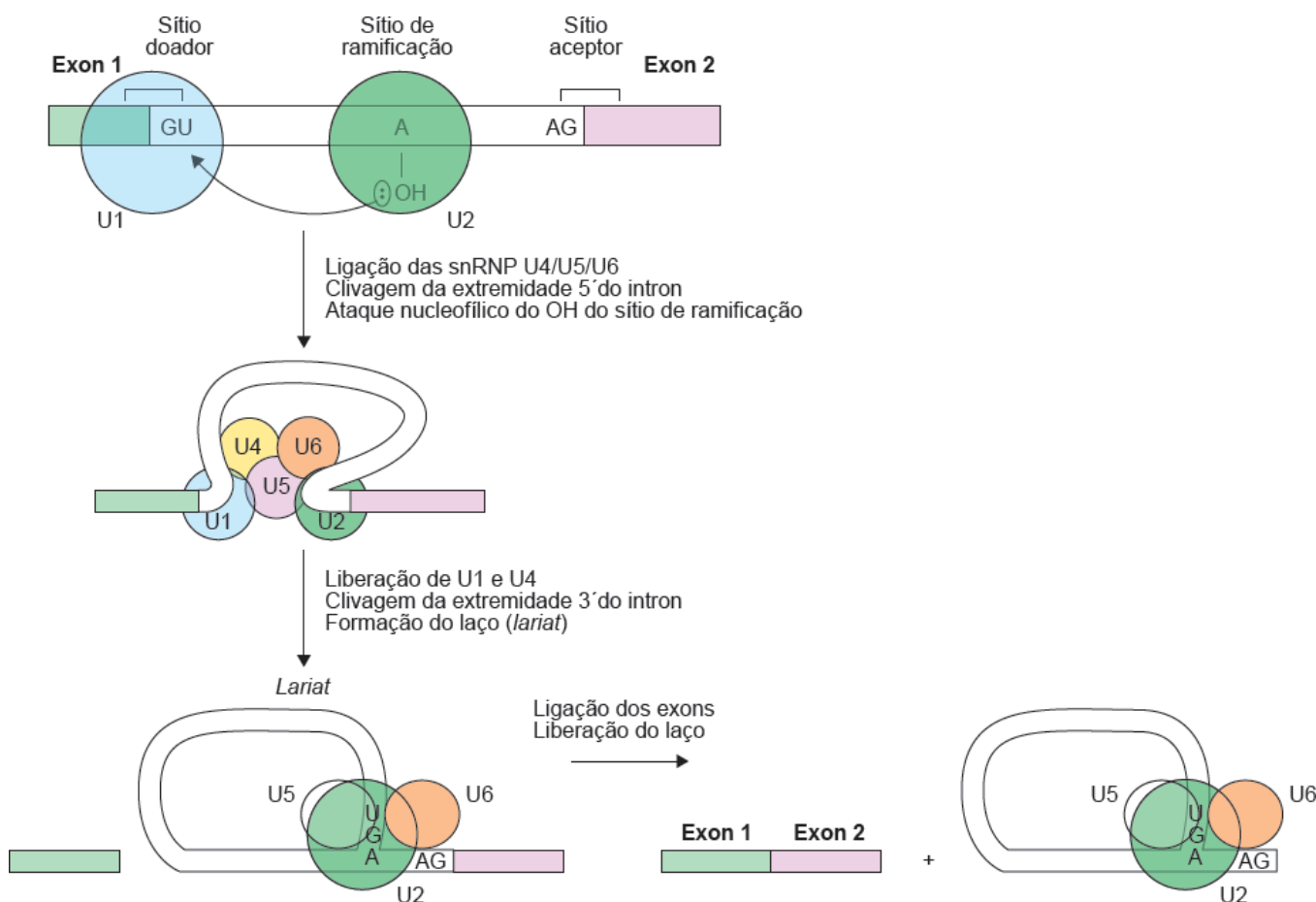


Figura 6.18 Mecanismo de *splicing* de introns tipo 1. As pequenas ribonucleoproteínas nucleares (snRNP) U1 e U2 ligam-se, respectivamente, aos sítios doador e de ramificação. Há a clivagem da extremidade 5' do intron, com ataque nucleofílico do sítio de ramificação. As snRNP U4/U5/U6 são recrutadas, aproximando as extremidades do intron e promovendo a clivagem da extremidade 3' do intron e ligação dos exons. Outras proteínas fazem parte do spliceossomo, mas foram omitidas da figura por motivo de clareza. Adaptada de Lewin *et al.*, 2011.

RNA polimerase II depende de fatores de ativação para iniciar a transcrição

Como a RNA polimerase II por si só não apresenta grande afinidade pelas regiões promotoras, os ativadores de transcrição são essenciais para que ela seja recrutada ou ativada para iniciar a transcrição. A ligação de ativadores ao DNA se dá tanto em regiões próximas aos promotores, como em regiões que podem estar a centenas ou até milhares de pares de base de distância dele. Essa é uma diferença interessante entre a regulação em eucariotos e em bactérias, em que os fatores regulatórios se ligam próximos ao promotor. Essas regiões distantes são denominadas *enhancers* em eucariotos superiores e podem se localizar a montante, a jusante ou mesmo no interior da sequência codificadora ou em introns. Em leveduras, os sítios próximos e a montante dos promotores são denominados UAS (*upstream activation sequences*). O número e as sequências de sítios de ligação de ativadores de transcrição variam muito de gene para gene, de acordo com a complexidade de sua regulação. Alguns ativadores de transcrição podem ser

responsáveis pela regulação de diversos genes, enquanto outros são mais específicos e têm espectro de ação mais restrito.

A atividade de um ativador de transcrição também pode ser modulada pela interação com outras proteínas e por modificações covalentes, como a fosforilação por proteínas quinases ou desfosforilação por fosfatases. Ativadores de transcrição geralmente apresentam dois domínios distintos: um domínio de ligação a DNA e outro que promove a interação com outras proteínas, tornando possível que a RNA polimerase II seja trazida ao promotor.

Como os ativadores podem se ligar a sítios no DNA distantes do promotor, proteínas denominadas coativadores são igualmente importantes, pois viabilizam a interação dos ativadores com o complexo de transcrição basal. Essas proteínas fazem parte de grandes complexos de mais de 20 polipeptídeos, sendo o complexo principal denominado *mediador* (Figura 6.19).

Repressores interferem no início da transcrição

Repressores de transcrição também estão presentes em eucariotos, embora o mecanismo molecular seja diferente do encontrado em bactérias. Nestas, proteínas repressoras como LacI, descrito anteriormente neste capítulo, impedem diretamente a ligação da RNA polimerase na região promotora. Como nos eucariotos a RNA polimerase não consegue se ligar por si só ao DNA, o papel dos repressores é indireto, interferindo nas proteínas moduladoras da transcrição, como ativadores, coativadores e proteínas remodeladoras da cromatina. O próximo item exemplifica a ação de várias proteínas que interferem na transcrição do regulon GAL de leveduras.

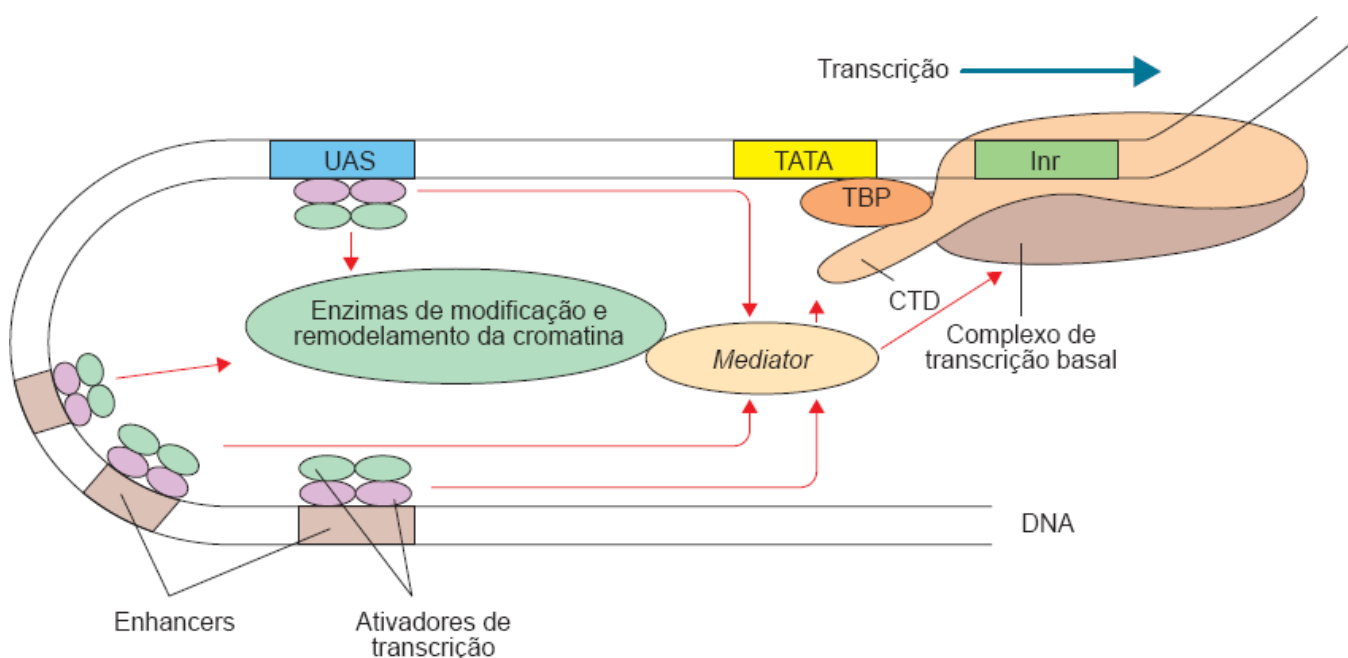


Figura 6.19 Visão geral das regiões do DNA e dos fatores necessários para a ativação da expressão de um gene eucarioto. Ativadores de transcrição ligam-se a *enhancers* ou regiões UAS e interagem com proteínas que remodelam e modificam a cromatina e com o complexo coativador *mediator*. Essas interações são facilitadas pelo dobramento do DNA, promovido por proteínas HMG, abundantes na cromatina. O complexo de transcrição basal é montado no promotor (são mostrados o TATA-box ligado a TBP e o sítio Inr). O CTD da RNA polimerase II é um sítio importante de interação com o complexo *mediator* e com outras proteínas. Adaptada de Lehninger *et al.*, 2008.

Exemplo de regulação gênica | Ativação e repressão dos genes de uso da lactose em leveduras

Na ausência de glicose, o substrato favorito, *Saccharomyces cerevisiae* pode usar galactose como fonte de carbono e energia. Os genes que codificam as enzimas responsáveis pela conversão de galactose a glicose-6-fosfato, bem como as proteínas regulatórias e as permeases, são denominados GAL e regulados em conjunto, formando um regulon, embora não estejam todos agrupados em uma mesma região do genoma. Apesar de, à primeira vista, esse modelo se assemelhar à regulação do operon *lac* em *E. coli*, uma análise do mecanismo molecular envolvido demonstra a maior complexidade do sistema de leveduras.

Uma rede de interações entre as proteínas Gal4p, Gal3p e Gal80p e pequenas moléculas dita o estado de ativação/repressão do regulon (Figura 6.20 A). Gal4p é um ativador de transcrição, que permanece no núcleo, ligado

a sítios ativadores no DNA. Gal80p pode ser encontrada no citoplasma ou no núcleo, interagindo com Gal4p e funcionando como um repressor na ausência de galactose, silenciando a transcrição dos genes GAL pelo bloqueio de Gal4p. Gal3p funciona como uma proteína transdutora de sinal: ela se localiza no citoplasma e interage com a Gal80p quando ligada a galactose e ATP. Dessa maneira, o complexo Gal4p/Gal80p é desfavorecido no núcleo, pelo sequestro de Gal80p no citoplasma. O ativador Gal4p livre da inibição de Gal80p pode assim interagir com o complexo *mediator* e com proteínas remodeladoras da cromatina, favorecendo a transcrição dos genes-alvo que resultarão no uso da galactose.

A repressão por glicose dos genes GAL se dá pela ação de outro repressor, Mig1, que é encontrado no núcleo em sua forma desfosforilada, inativa. A fosforilação de Mig1 é feita pela quinase Snf1, que é inativa na presença de glicose. Em resumo, a ausência de glicose na célula ativa uma via que leva à fosforilação de Mig1, que, ao se ligar ao DNA entre a região ativadora e o promotor, recruta fatores que impedem a transcrição dos genes GAL.

A ativação da transcrição pode ocorrer por mecanismos diversos

Uma classe de ativadores de transcrição importante na regulação da expressão gênica de eucariotos superiores é representada pelos receptores de hormônios esteroides (Figura 6.20 B). Ao contrário de hormônios peptídicos, como a insulina, os esteroides, entre eles o hormônio da tireoide e a progesterona, não interagem com proteínas receptoras na membrana das células. Por sua natureza hidrofóbica, os esteroides são capazes de atravessar a membrana e interagir diretamente com proteínas localizadas no citoplasma. Esses complexos hormônio-receptor são translocados para o núcleo, onde podem se ligar a sequências específicas no DNA, denominadas HRE (*hormone responsive elements*), alterando a expressão dos genes-alvo.

Alguns ativadores de transcrição têm sua atividade modulada por fosforilação, mediada por proteínas quinases. Frequentemente, os sinais para a ativação da expressão de um gene estão presentes no exterior da célula, e esse sinal precisa chegar ao núcleo para que a ativação ou repressão da transcrição aconteça. Nesse mecanismo, que faz parte de processos denominados genericamente *transdução de sinal*, estão envolvidas proteínas receptoras, pequenas moléculas denominadas segundo mensageiros e as proteínas quinases e fosfatases que modificam os ativadores de transcrição ou outras proteínas envolvidas no processo. Assim, ocorre a resposta necessária para ajustar a fisiologia da célula aos estímulos externos. Exemplos de transdução de sinal serão explorados em outros capítulos deste livro.

Conclusão e perspectivas | Regulação da expressão gênica, desenvolvimento e evolução

Com o sequenciamento do genoma de várias espécies, vem se tornando evidente que as sequências das proteínas de diferentes organismos não são tão diversas quanto se poderia imaginar. De maneira análoga, em um organismo multicelular complexo em que todas as células apresentam o mesmo genoma, como é possível que elas se diferenciem em tipos tão diversos? A resposta para essas aparentes discrepâncias está, em grande parte, relacionada com a expressão diferencial dos genes em cada célula. Neste capítulo foram mostrados os mecanismos básicos de ativação da transcrição, que se inter-relacionam com vias de transdução de sinal e regulação pós-transcricional, trazendo à tona toda a diversidade que está contida nos diferentes organismos. Sob esse ponto de vista, mutações em genes regulatórios podem ter uma influência sobre vários outros genes, resultando em um fenótipo com alterações mais evidentes do que se essas mutações tivessem se dado em genes de proteínas estruturais, por exemplo.

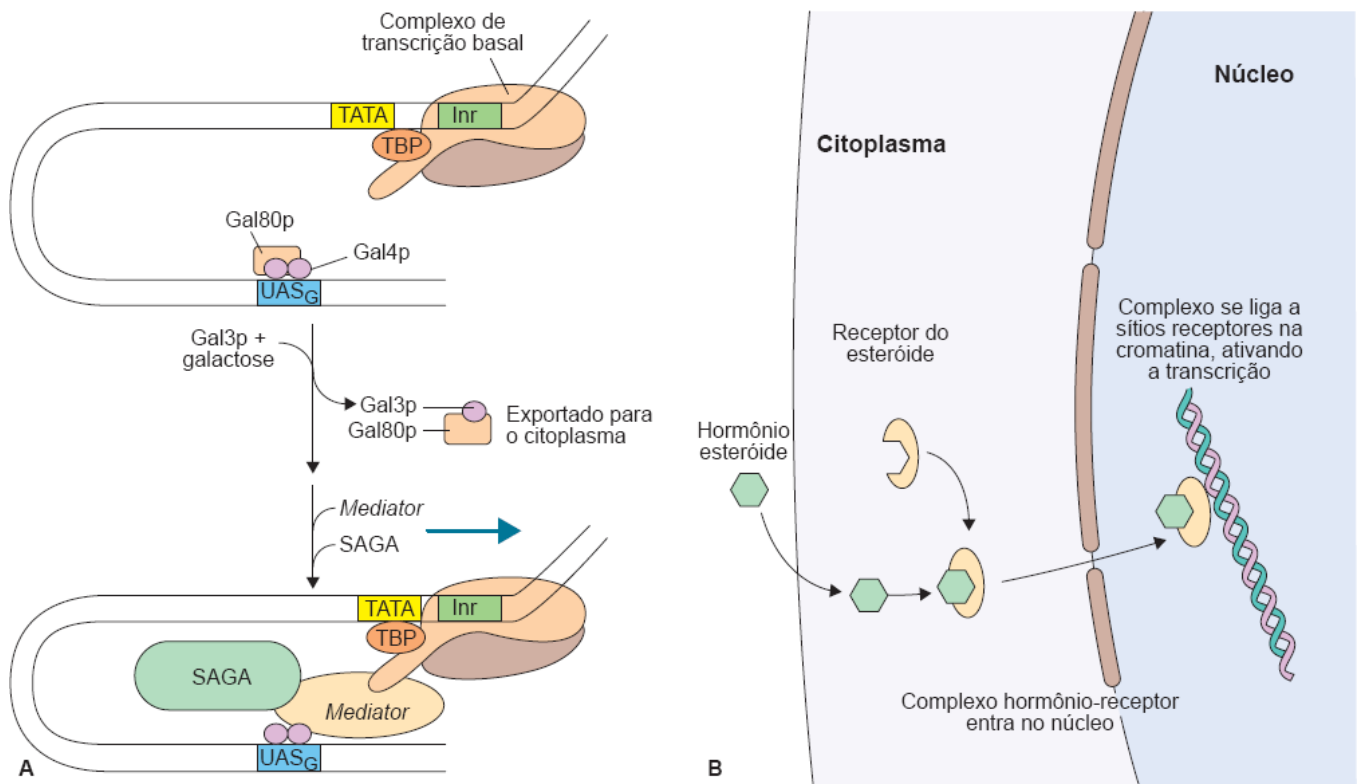


Figura 6.20 A. Ativação da transcrição dos genes GAL de levedura por galactose. Na ausência de galactose, o ativador de transcrição Gal4p mantém-se ligado ao sítio de ativação no DNA, porém sua atividade é bloqueada por Gal80p. A proteína Gal3p, encontrada tanto no núcleo quanto no citoplasma, liga-se à galactose presente na célula e interage com Gal80p no núcleo. O complexo Gal3p/galactose/Gal80p é exportado para o citossol e Gal4p pode agora interagir com o complexo de remodelamento da cromatina (SAGA) e com o complexo coativador *mediator*, ativando a transcrição pela RNA polimerase II. **B.** Ativação da transcrição por hormônios esteroides. As moléculas hidrofóbicas dos esteroides atravessam a membrana da célula e se ligam a receptores citoplasmáticos. O complexo hormônio-receptor é translocado para o núcleo, onde interage com sítios receptores na cromatina. Os domínios de ligação a DNA do receptor ligam-se ao DNA e ativam a transcrição dos genes-alvo. Adaptada de Lehninger *et al.*, 2008.

O avanço atual nas tecnologias de alto desempenho que, com a análise global de genomas, transcriptomas e proteomas aliada a técnicas que promovam a detecção de moléculas em células vivas em tempo real, certamente contribuirá cada vez mais para o entendimento dos processos de regulação da expressão gênica, com possíveis aplicações em saúde e biotecnologia.

Bibliografia

- Alberts B, Wilson JH, Hunt T. Molecular biology of the cell. 5. ed. New York: Garland Science; 2008.
- Browning DF, Busby SJ. The regulation of bacterial transcription initiation. *Nat Rev Microbiol.* 2004;2:57-65.
- Carroll SB, Prud'Homme B, Gompel N. O jogo da evolução. *Scientific American Brasil.* 2008;73.
- Christensen T, Oliveira A, Nielsen J. Reconstruction and logical modeling of glucose repression signaling pathways in *Saccharomyces cerevisiae*. *BMC Systems Biology.* 2009;3:7.
- Dumay-Odelot H, Durrieu-Gaillard S, Da Silva D, Rolder RG, Teichmann M. Cell growth- and differentiation-dependent regulation of RNA polymerase III transcription. *Cell Cycle.* 2010;9(18):3687-99.
- Görke B, Stülke J. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat Rev Microbiol.* 2008;6:613-24.
- Haag JR, Pikaard CS. Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. *Nat Rev Mol Cell Biol.* 2011;12:483-92.
- Jacob F, Monod J. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol.* 1961;3:318-56.
- Jacob F, Ullmann A, Monod J. The promoter, a genetic element necessary to the expression of an operon. *C R Hebd Seances Acad Sci.* 1964;258:3125-8.
- Juven-Gershon T, Hsu JY, Theisen JWM, Kadonaga JT. The RNA polymerase II core promoter – the gateway to transcription. *Curr Opin Cell Biol.* 2008;20:253-9.

Kravchenko JE, Rogozin IB, Koonin EV, Chumakov PM. Transcription of mammalian messenger RNAs by a nuclear RNA polymerase of mitochondrial origin. *Nature*. 2005;436:735-9.

Lehninger AL, Nelson DL, Cox MM. *Lehninger Principles of Biochemistry*. 5. ed. New York: W.H. Freeman; 2008.

Lewin B, Krebs JE, Kilpatrick ST, Goldstein ES. *Lewin's genes X*. 10. ed. Sudbury, EUA: Jones and Bartlett; 2011.

Lodish HF. *Molecular cell biology*. 6. ed. New York: W.H. Freeman; 2008.

Schramm L, Hernandez N. Recruitment of RNA polymerase III to its target promoters. *Genes & Development*. 2002;16:2593-620.

Yanofsky C. Attenuation in the control of expression of bacterial operons. *Nature*. 1981;289(5800):751-8.

7

Tradução do Código Genético | Processo de Síntese Proteica

Emy Tiyo Mano e Wellington Luiz de Araújo

Todas as informações herdadas que originam um organismo estão contidas na molécula de DNA, que, por sua vez, se envolve com os processos evolutivos. Entretanto, a maior parte do metabolismo e a estrutura celular são mantidos pelas proteínas, as quais são constituídas por sequências lineares de aminoácidos. Essa sequência peptídica é determinada pela informação contida na molécula de DNA, a qual deve ser lida e usada para a síntese de uma molécula de RNA (transcrição), usada pelos ribossomos para dirigir a síntese do peptídeo, que é revisada neste capítulo.

Conceito | Um gene, uma proteína

O conceito de que um gene é responsável pela produção de uma proteína específica foi proposto pela primeira vez em 1909, pelo físico inglês Archibald E. Garrod (1857-1936). Garrod tinha um grande interesse por doenças humanas, sendo o primeiro a usar a expressão “erros inatos do metabolismo”. Ele sugeriu que a alcaptonúria era resultante da mutação em um único gene, resultando em um defeito na via bioquímica associada à excreção de metabólitos pela urina. Foi observado que essa doença genética ocorreria pela presença de um alelo recessivo, e é rara, com baixa frequência na população (1:1.000.000), além de não apresentar predominância étnica. A alcaptonúria é uma condição herdada que provoca deficiência no metabolismo da fenilalanina e da tirosina, resultando em sintomas mais evidentes após a quarta década de vida, em decorrência da deposição do ácido homogentísico em diversos tecidos e aumento de sua excreção urinária. Embora Garrod tenha publicado um livro e vários artigos a respeito, seu trabalho foi ignorado até 1941, quando passou a ser citado pelos geneticistas americanos George W. Beadle (1903-1989) e Edward L. Tatum (1909-1975) em seus ensaios com esporos do fungo ascomiceto *Neurospora crassa* (o bolor laranja do pão). Beadle e Tatum realizaram vários experimentos para estudar a conexão entre genes e enzimas, e publicaram um trabalho no qual descreveram o conceito “um gene, uma enzima”. Esses resultados foram reconhecidos pela comunidade científica e, em 1958, os dois autores ganharam o Prêmio Nobel de Fisiologia/Medicina.

O trabalho de Beadle e Tatum foi desenvolvido com o fungo *N. crassa*, o qual apresenta um genoma haploide. Assim, mutações em um gene que participa de uma via biossintética essencial tornam a célula inviável, a menos que o produto (ou intermediário metabólico) dessa via seja fornecido ao organismo no meio de cultura. Pensando nisso, Beadle e Tatum avaliaram a relação 1:1 entre genes e enzimas por meio da geração de mutantes genéticos. Eles irradiaram os esporos do fungo com raios X para aumentar a frequência de mutação de genes, e colocaram esses esporos em diferentes meios de cultura. Por exemplo, para selecionar um mutante incapaz de produzir o aminoácido arginina (arg⁻) bastaria suplementar o meio mínimo (meio de cultura contendo apenas sais, uma fonte de carbono e uma fonte de nitrogênio) com arginina; dessa maneira, os mutantes usariam essa substância do meio e sobreviveriam à sua deficiência genética. Mas surgiu um problema: como separar os fungos selvagens (arg⁺) que não sofreram a mutação? A solução proposta por Beadle e Tatum foi a retirada de uma pequena amostra de cada fungo cultivado no meio suplementado, transferindo-a para o meio mínimo. Desse modo, os fungos que se

desenvolvessem nesse meio mínimo seriam certamente selvagens, e os que não se desenvolvessem seriam mutantes (Figura 7.1).

Com esses experimentos de Beadle e Tatum, consolidou-se a teoria “um gene, *uma enzima*”, que posteriormente foi ampliada para “um gene, *uma proteína*”, pois ficou claro que genes controlavam a síntese de qualquer proteína, e não apenas aquelas que apresentavam função enzimática. À medida que foi descoberto que uma proteína poderia ser formada por mais de uma cadeia polipeptídica, como no caso da hemoglobina humana, sendo cada uma delas codificada por um gene diferente, a teoria tornou-se ainda mais abrangente, passando a ser denominada teoria “um gene, *um polipeptídeo*”.

Desvendando o código genético

Os genes são unidades biológicas responsáveis por determinar as características de um organismo, funcionando provavelmente como um código capaz de direcionar a síntese de proteínas. Segundo o “dogma central”, proposto por Francis H. Crick (1916-2004), essa informação genética flui do DNA para proteína por meio de uma molécula carreadora, o RNA mensageiro (mRNA). Esse processo de transferência da informação, conhecido como tradução, é um dos mais conservados em todos os organismos e um dos eventos de maior custo energético; por exemplo, nas bactérias que têm o seu crescimento rápido, 80% da energia celular e 50% do seu peso seco destinam-se à síntese de proteínas.

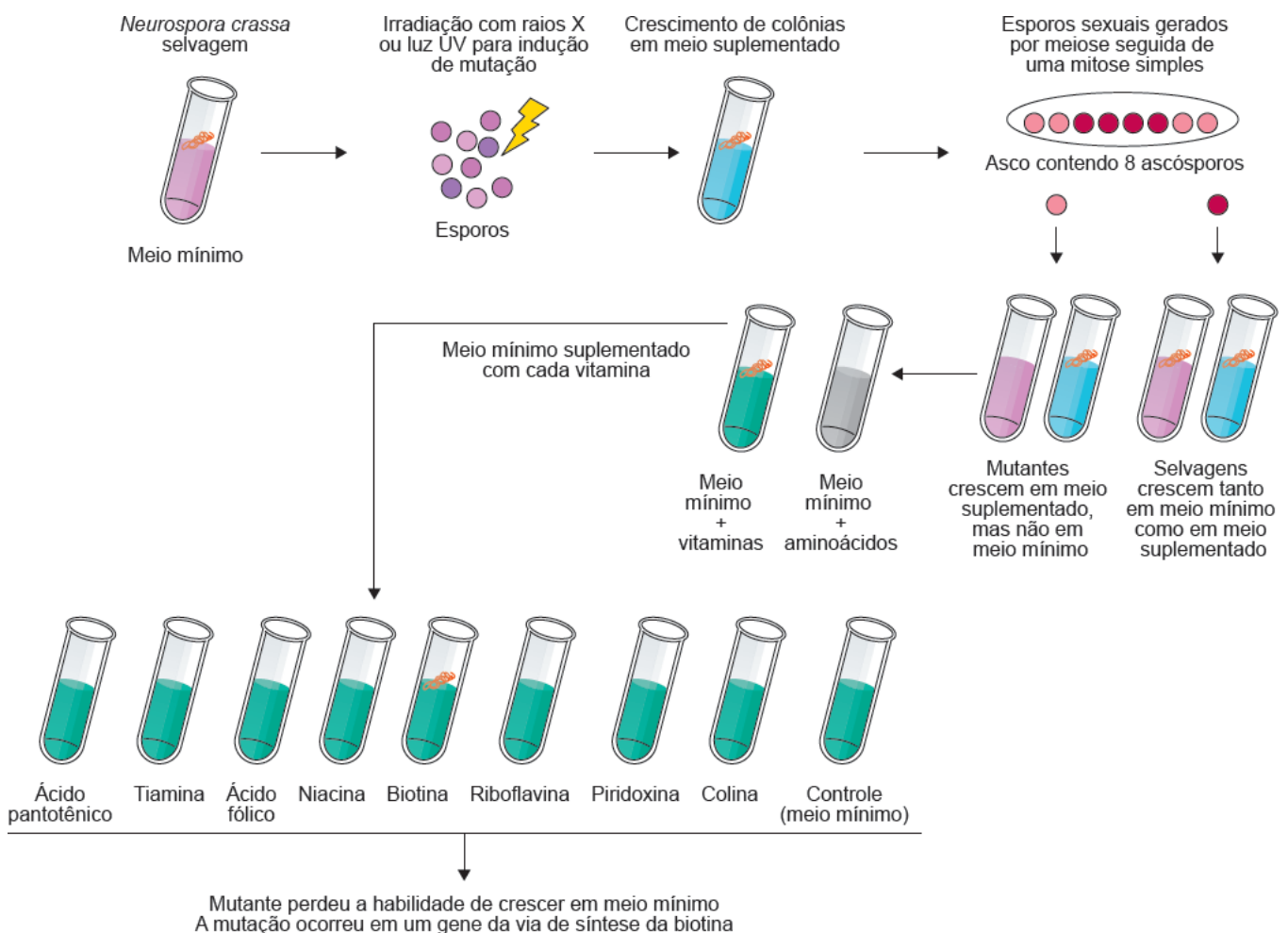


Figura 7.1 Ensaios desenvolvidos por Beadle e Tatum com esporos mutantes de *Neurospora crassa* semeados em diferentes meios de cultura.

Além desses atributos, a tradução é um processo desafiador no que se refere à transferência da informação, se comparado à transcrição do DNA em RNA. Diferentemente da complementaridade que existe entre o molde de DNA para os ribonucleotídeos do RNA, as cadeias laterais dos aminoácidos têm pouca ou nenhuma afinidade química pelas bases púricas e pirimídicas da cadeia ribonucleotídica. Mais improvável ainda seria se diferentes combinações de 3 bases de RNA pudessem formar superfícies com afinidades específicas por cada um dos aminoácidos.

Diante dessa característica intrigante a respeito do processo de tradução do mRNA para sequências de aminoácidos, em 1955, Francis Crick propôs que, antes da incorporação dos aminoácidos em polipeptídeos, os aminoácidos precisam associar-se a uma molécula adaptadora especial. Essa molécula adaptadora deveria possibilitar a interação entre essas duas moléculas quimicamente incompatíveis. Dessa maneira, Crick imaginou que muito provavelmente esse adaptador seria uma espécie de RNA, pois só assim ela reconheceria as regras de pareamento de bases do código de Watson e Crick.

Apenas dois anos mais tarde, Mahlon B. Hoagland (1921-2009) e Paul C. Zamecnik (1912-2009) demonstraram que, antes de sua incorporação às proteínas, os aminoácidos se ligam a uma classe de moléculas de RNA, no início chamada de RNA “solúvel”. Eles também identificaram as enzimas aminoacil-tRNA sintetases, que eram responsáveis por ligar um aminoácido específico a um RNA “solúvel” e transportá-lo para o ribossomo para a síntese proteica. Esse RNA “solúvel”, assim denominado por ser uma molécula de baixo peso molecular e solúvel, foi posteriormente renomeado de RNA transportador (tRNA ou RNA de transferência).

Outro detalhe curioso que precisava ser decifrado no código genético era a maneira como a sequência de quatro tipos de nucleotídeos do mRNA poderia ser traduzida nos 20 tipos diferentes de aminoácidos das proteínas. A proposta inicial foi de que a unidade de leitura (denominada códon) teria três nucleotídeos, de modo que sua combinação promoveria a existência de 64 códons, que seriam, portanto, redundantes; ou seja, haveria mais de um códon para cada um dos 20 aminoácidos. Se os códons fossem formados por um ou dois nucleotídeos, o número máximo de combinações entre eles seria de 16, insuficiente para os 20 aminoácidos. Da mesma maneira, pelo princípio da parcimônia, segundo o qual a solução mais simples geralmente é a certa, foi descartada a hipótese de um códon com quatro nucleotídeos.

Para demonstrar se realmente o mRNA que codifica uma proteína consistia de uma sequência ordenada de unidades trinucleotídicas, ou seja, cada códon formado por uma trinca de bases nucleotídicas determinaria um aminoácido, Marshall W. Nirenberg (1927-2010) e J. Heinrich Matthaei (1929-) iniciaram, em 1961, experimentos para testar a hipótese dos códons de trios. Eles usaram um extrato acelular de *Escherichia coli*, pois acreditavam que esse extrato teria todos os componentes necessários para traduzir o mRNA em proteínas. O extrato foi tratado com DNase para destruir todo o DNA que pudesse servir de molde para síntese de novos mRNA. Posteriormente, a esse extrato tratado foram adicionados mRNA sintéticos completamente compostos de uracila (poli-U RNA) e aminoácidos marcados radioativamente. Foi observado que nesse sistema acelular foram obtidos polipeptídeos inteiramente compostos do aminoácido fenilalanina (FEN). O resultado sugeriu fortemente que uma trinca UUU codificava uma fenilalanina, que seria, então, o primeiro códon para síntese de um aminoácido decifrado. Os próximos códons a serem testados foram as poli-C, resultando em uma cadeia de prolina (PRO); as poli-A, resultando em uma cadeia de lisinas (LIS); e poli-G, que não resultou em nenhuma cadeia. Esse resultado negativo para a sequência de poli-G foi explicado posteriormente por Nirenberg, Maxine F. Singer (1931-) e Bill Jones como decorrente do fato de que uma sequência de RNA rica em guanosina (G) forma tetrâmeros estáveis que impedem a ligação do ribossomo. Em seguida, foram produzidos mRNA sintéticos contendo dois tipos de nucleotídeos em grupos repetidos, por exemplo, tendo sequências (ACACAC)_n nas quais foram observadas a síntese de uma cadeia polipeptídica de histidina e treonina (Figura 7.2) – de onde se concluiu que realmente a síntese de aminoácidos era determinada a partir de uma trinca de bases do mRNA. Pela descoberta do código genético e a descrição de como ocorre a síntese proteica, Nirenberg, junto com Har G. Khorana (1922-2011) e Robert W. Holley (1922-1993), ganhou o Prêmio Nobel de Medicina/Fisiologia de 1968, sendo Matthaei excluído do prêmio.

Ainda restava uma dúvida, pois a combinação dos quatro tipos de nucleotídeos de RNA em uma trinca (4³) resultaria em 64 combinações, o que extrapolaria a quantidade de 20 diferentes aminoácidos disponíveis. Desse modo, apostou-se na hipótese de um código degenerado, ou seja, muitos códons diferentes especificariam para o mesmo aminoácido, e essa noção é válida ainda hoje. A teoria de que vários códons poderiam codificar para um mesmo aminoácido foi confirmada pela técnica de síntese de mRNA sintéticos, que testou todas as combinações de códons e seus aminoácidos correspondentes, embora alguns deles fossem difíceis de decifrar porque não se podia determinar a ordem bioquímica. Foi então que Phillip Leder (1934-) e Nirenberg determinaram o restante do código genético por meio da ativação do tRNA (Figura 7.3). Eles desenvolveram uma técnica usando um mRNA curto de sequência conhecida e adicionaram uma mistura de ribossomos e um aminoácido marcado radioativamente. Se determinado tRNA pareava com o pequeno mRNA sintético complexado ao ribossomo, ao ser passado pelo filtro, este era retido e identificado. Isolado cada tRNA ativado e avaliado o anticódon desse tRNA, não apenas todo o

código genético foi confirmado, como também se descobriu que existia um códon específico para o início da tradução (AUG) e códons de terminação responsáveis pelo fim da tradução (UAA, UAG e UGA). Assim, no final da década de 1960, todo o código genético já havia sido decifrado (Tabela 7.1).

Tradutor de códigos | RNA transportador

O processo central da síntese proteica se dá na tradução da informação contida na sequência de ribonucleotídeos do mRNA para aminoácidos. Papel fundamental para essa síntese é atribuído às moléculas de tRNA, que atuam como adaptadores entre os códons do mRNA com os aminoácidos correspondentes. Para tanto, existem vários tipos de tRNA, do qual se fosse necessário um pareamento perfeito de bases Watson-Crick entre códon e anticódon, as células deveriam conter exatamente 61 diferentes tipos de tRNA. No entanto, o que ocorre de modo geral, é um tRNA para cada aminoácido, com a maioria dos tRNA reconhecendo mais de uma combinação de trinca de bases do mRNA. Isso varia entre os organismos e ocorre em função de pareamentos não convencionais entre as bases na posição pendular, ou seja, a 3ª base (3') no códon do mRNA e sua correspondente, a 1ª base (5') no anticódon do tRNA. As 1ª e 2ª bases de um códon geralmente formam pareamento clássico com suas respectivas 3ª e 2ª bases do anticódon. Já as bases da região pendular (3ª base do códon e 1ª base do anticódon) podem fazer 4 interações não convencionais. Essas observações levaram Crick a concluir que a 3ª base da maioria dos códons parecia mais fracamente com a base correspondente de seu anticódon, a partir da qual ele propôs um conjunto de quatro relações chamadas de hipótese de oscilação:

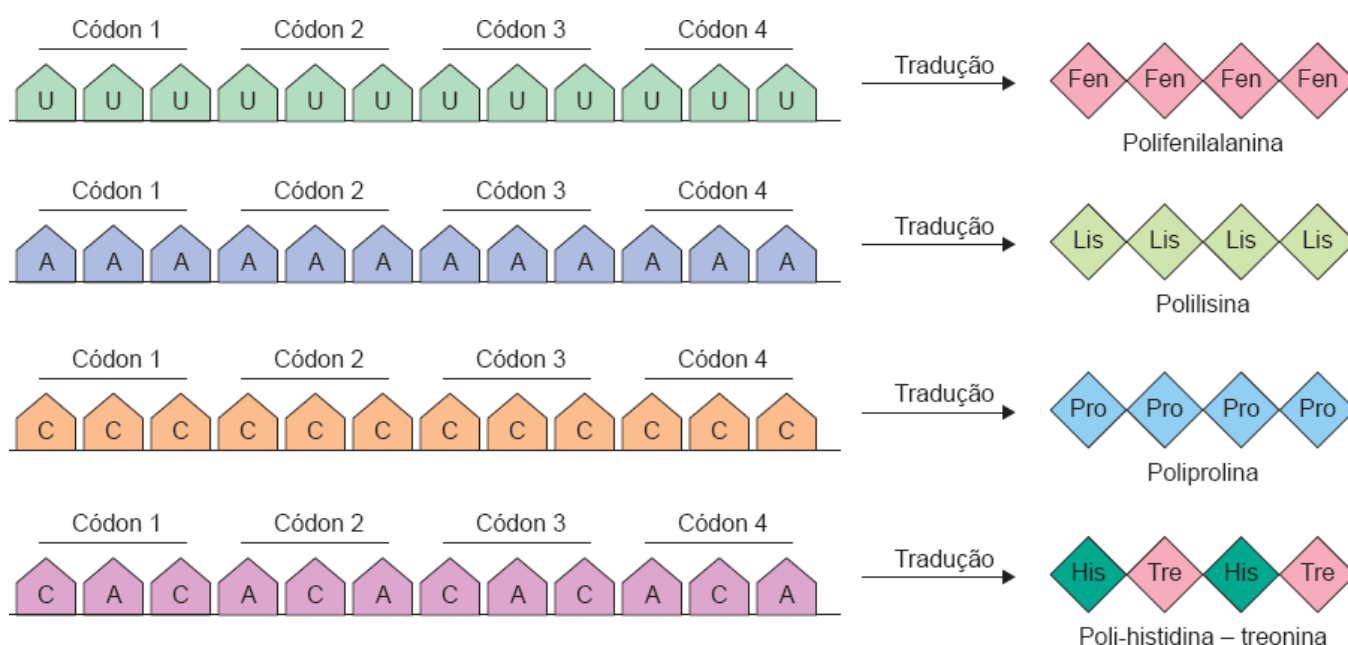


Figura 7.2 Experimentos de Nirenberg e Matthaei com mini-mRNA trinucleotídeos sintéticos que ajudaram a decifrar o código genético.

1. As primeiras bases de um códon do mRNA sempre formam pareamento de bases fortes Watson-Crick com as bases correspondentes do anticódon do tRNA, e conferem a maior parte da especificidade do código.
2. A 1ª base do anticódon (correspondente à 3ª base do códon) determina o número de códons reconhecidos pelo tRNA. Quando a 1ª base do anticódon for C ou A, o pareamento de bases é específico e apenas um códon é reconhecido por aquele tRNA. Quando a 1ª base do anticódon for U ou G, a ligação é menos específica e 2 códons diferentes podem ser lidos. Quando a inosina for o 1º nucleotídeo de um anticódon, 3 anticódons diferentes podem ser reconhecidos.

C	CUU	Leucina	CCU	Prolina	CAU	Histidina	CGU	Arginina	U
	CUC		CCC		CAC		CGC		C
	CUA		CCA		CAA	Glutamina	CGA		A
	CUG		CCG		CAG		CGG		G
A	AUU	Isoleucina	ACU	Treonina	AAU	Asparagina	AGU	Serina	U
	AUC		ACC		AAC		AGC		C
	AUA		ACA		AAA	Lisina	AGA	Arginina	A
	AUG	Metionina	ACG		AAG		AGG		G
G	GUU	Valina	GCU	Alanina	GAU	Ácido aspártico	GGU	Glicina	U
	GUC		GCC		GAC		GGC		C
	GUA		GCA		GAA	Ácido glutâmico	GGA		A
	GUG		GCG		GAG		GGG		G

Da mesma maneira, o tRNA também conserva uma estrutura secundária comum, apresentando regiões de fita simples e regiões de fita dupla, com aspecto semelhante ao uma folha de trevo (Figura 7.4), demonstrando haver regiões de complementaridade nos tRNA. Os principais aspectos desse tRNA em folha de trevo são:

- Uma haste aceptora, que é o sítio de ligação do aminoácido, formada pelo pareamento entre as extremidades 5' e 3' da molécula de tRNA, contendo uma sequência 5'-CCA-3' conservada na porção 3'
- Três hastes em forma de alças, chamadas de alça ψ U (*T-loop* – pela presença característica de uma base incomum pseudouridina); alça D (pela presença característica de di-hidrouridina); e uma alça anticódon (pela presença do anticódon, formado por 3 nucleotídeos, delimitada por uma purina na extremidade 3' e uma uridina na extremidade 5', responsável pelo seu pareamento ao códon do mRNA)
- Uma alça variável, que se situa entre a alça do anticódon e a alça ψ U. O tamanho dessa alça pode variar de 3 a 21 bases.

Já a estrutura terciária do tRNA foi determinada por cristalografia de raios X. Foi observada uma estrutura em forma de L invertido, com a haste aceptora do aminoácido em uma extremidade, e a alça do anticódon em outra extremidade (Figura 7.4). Três tipos de interações estabilizam essa estrutura tridimensional. A primeira são as pontes de hidrogênio entre bases de regiões helicoidais diferentes, que se aproximam pela estrutura terciária. A segunda constitui-se por interações entre as bases e o esqueleto de açúcar-fosfato. E a terceira interação ocorre pelo empilhamento das bases, causado pela formação de duas extensas regiões de pareamento entre bases, que também atuam estabilizando a molécula.

Aminoacil-tRNA sintetases

Durante a primeira etapa da síntese proteica, que ocorre no citosol, os 20 aminoácidos são esterificados aos seus tRNA correspondentes. Quando uma molécula de tRNA apresenta um aminoácido covalentemente ligado à região 5'-CCA-3' conservada na porção 3', é dito que esse tRNA está carregado, e, conseqüentemente, aquelas que não apresentam um aminoácido ligado estão descarregadas. A molécula enzimática responsável por fazer esse carregamento é a aminoacil-tRNA sintetase, que pode ser específica para cada aminoácido e/ou mais tRNA correspondentes. No entanto, sabe-se que podem existir menos de 20 aminoacil-tRNA sintetases, de modo que uma mesma enzima é responsável pelo acoplamento de mais de um aminoácido aos seus tRNA apropriados, como ocorre em algumas bactérias. Nesses casos, uma única sintetase posiciona o aminoácido idêntico em dois tipos diferentes de tRNA, mas apenas um deles tem o anticódon que combina com o aminoácido. Assim, uma segunda enzima

As aminoacil-tRNA sintetases são enzimas de acoplamento que ligam um aminoácido à hidroxila 2' ou 3' livre da adenosina na extremidade 3' terminal da molécula de tRNA, em uma reação dependente de ATP. Esse acoplamento é realizado em duas etapas enzimáticas no sítio ativo dessa enzima. Em uma primeira etapa, chamada adenililação, o grupo carbonila do aminoácido reage com o grupo fosforil do ATP para formar uma ligação anidrídica, com a liberação do pirofosfato. Na adenililação, o aminoácido é ligado ao ácido adenílico por uma ligação éster de alta energia. A segunda etapa envolve o carregamento, no qual o grupo aminoacila do aminoácido adenilado, que continua fortemente ligado à sintetase, é transferido do aminoacil-AMP para a extremidade 3' do tRNA (extremidade 5'-CCA- 3') correspondente, por meio da hidroxila 2'- ou 3'- e com liberação concomitante da AMP (Figura 7.5).

O diagrama ilustra a estrutura de uma tRNA em duas representações: A (plano de base) e B (estrutura 3D).

Diagrama A (Plano de Base): Mostra a tRNA em uma forma de T invertida. O **Sítio de ligação do aminoácido** está na extremidade 3', com a sequência ACC. O **Anticódon** está na extremidade oposta, com a sequência GAA. As **Alças** (loops) são rotuladas: Alça D (azul), Alça ψ U (rosa), Alça variável (verde claro) e Alça do anticódon (verde escuro). As bases nitrogenadas são mostradas em seus pares de base.

Diagrama B (Estrutura 3D): Mostra a tRNA em sua conformação tridimensional. As **Alças** são rotuladas: Alça ψ U (rosa), Alça D (azul) e **Anticódon** (verde escuro). A extremidade 3' termina em CCA, com um sítio de ligação do aminoácido (representado por uma esfera vermelha).

Figura 7.4 A. Estrutura semelhante a uma folha de trevo. **B.** Estrutura terciária do tRNA em forma de L invertido. Adaptada de University of Tokyo, 2011.

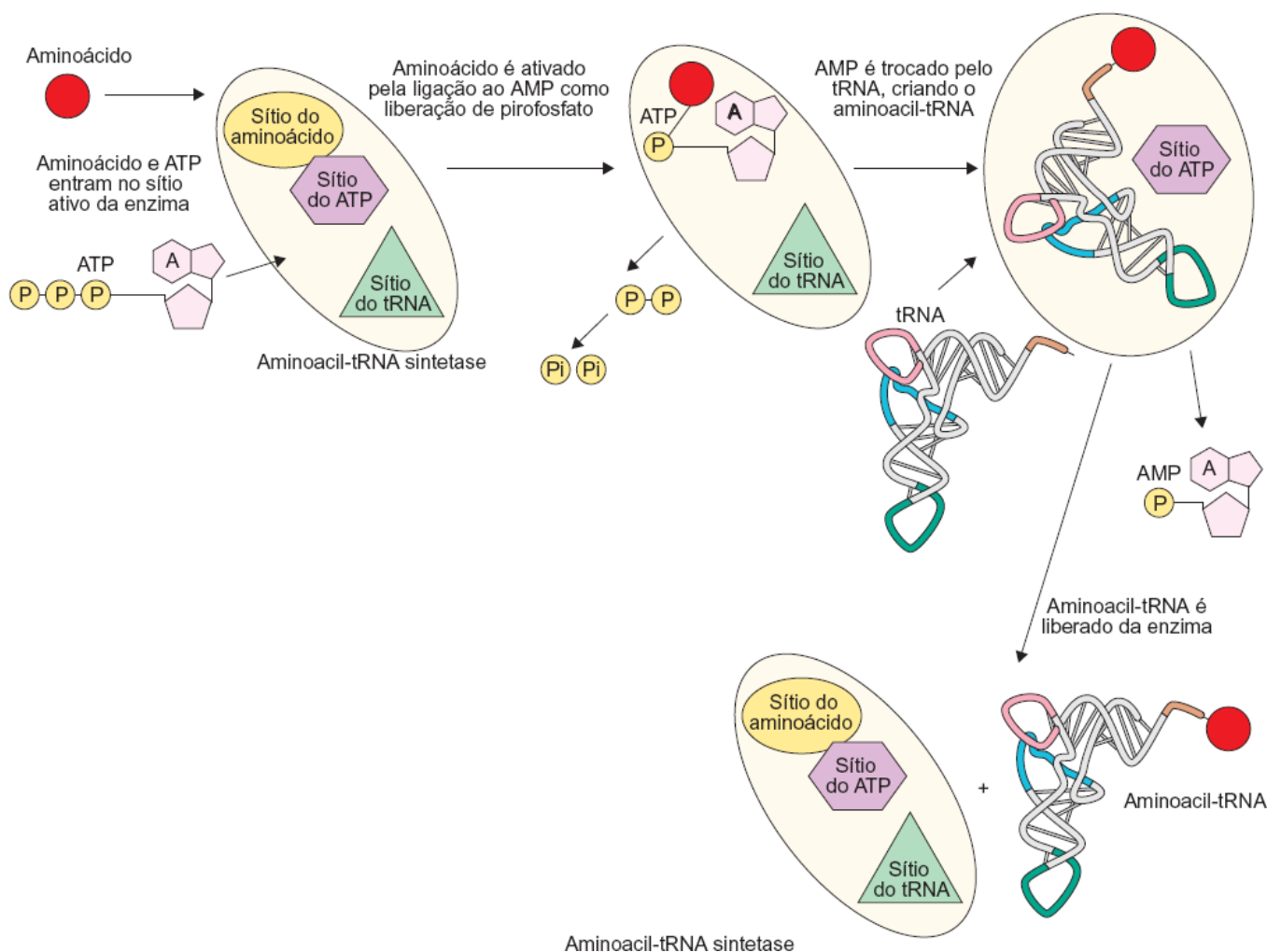


Figura 7.5 Etapa de carregamento do tRNA pela aminoacil-tRNA sintetase.

A taxa geral de erro de tradução é baixa, geralmente menor que 1 em 1.000, de maneira que em uma *E. coli*, por exemplo, é de aproximadamente em 1 códon a cada 50.000, porém sua frequência é maior se comparada à replicação do DNA. Falhas em uma proteína são eliminadas quando ela é degradada, como se verá posteriormente, e não são passadas para futuras gerações, tendo um significado biológico menor. No entanto, o grau de fidelidade na síntese de proteínas é suficiente para garantir que muitas proteínas não contenham erro e que a grande quantidade de energia requerida para sintetizar uma proteína raramente seja desperdiçada.

O ribossomo é uma maquinaria de leitura

Se os diversos componentes do sistema de tradução do mRNA tivessem de interagir livremente no citoplasma da célula, em solução, provavelmente esse mecanismo seria relativamente lento ou energeticamente desfavorável, uma vez que a ocorrência de colisões simultâneas seria tão limitada que a taxa de polimerização dos aminoácidos seria extremamente baixa.

O ribossomo é uma molécula composta de RNA ribossômico (rRNA) e um conjunto de proteínas ribossomais, organizadas em duas subunidades: a subunidade maior (50S para procariotos, e 60S para eucariotos) e a menor (30S para procariotos, e 40S para eucariotos), responsável por orientar a extensão de um polipeptídeo. O ribossomo aumenta a eficiência de tradução pela ligação do mRNA e de um tRNA de iniciação em sua subunidade menor, e, após o acoplamento da subunidade maior, se desloca da extremidade 5' sentido 3' do mRNA, carregando os tRNA seguintes, atuando como uma maquinaria de leitura códon a códon, a uma taxa de adição de 3 a 5 aminoácidos por segundo. Quando o ribossomo se depara com um códon de terminação, o peptídeo é liberado e o ribossomo se dissocia do mRNA, com as subunidades maior e menor agora separadas, ficando livres para a síntese de outro mRNA.

Embora um ribossomo só possa sintetizar um polipeptídeo por vez, cada mRNA pode ser traduzido por vários ribossomos ao mesmo tempo, processo pelo qual cada ribossomo interage com 30 nucleotídeos por vez, embora seu tamanho ocupe um espaço equivalente a 80 nucleotídeos. Esse mRNA portando vários ribossomos é chamado de

polirribossomo ou polissomo. Essa atividade simultânea de vários ribossomos foi observada por microscopia eletrônica usando técnicas desenvolvidas por Oscar L. Miller. A maioria das proteínas dobra-se na superfície do ribossomo durante sua síntese, entretanto, a fibroína permanece distendida nas condições usadas por esses pesquisadores, nas quais as cadeias polipeptídicas crescentes podem ser vistas ligadas aos ribossomos à medida que são examinadas da ponta 5' do mRNA para a ponta 3'. A fibroína é uma proteína de 200.000 daltons presente na glândula posterior do bicho-da-seda *Bombyx mori*, e é sintetizada em grandes polissomos, contendo de 50 a 80 ribossomos. Além disso, seja em procariotos ou eucariotos, o número de ribossomos está relacionado diretamente com a atividade de síntese de proteínas na célula, existindo cerca de 20.000 ribossomos em uma bactéria em crescimento.

Composição e estrutura dos ribossomos

Um ribossomo é composto por 3 (nas bactérias) ou 4 (em eucariotos) moléculas diferentes de rRNA e um conjunto de até 83 proteínas, distribuídas em duas subunidades, maior e menor. A subunidade maior contém o centro da peptidil-transferase, responsável pela formação da ligação peptídica. Já a subunidade menor contém o centro de decodificação, que se liga ao mRNA fazendo a decodificação dos tRNA carregados. Embora o número de moléculas proteicas exceda bastante o número de rRNA, os rRNA constituem aproximadamente 60% da massa do ribossomo.

As subunidades ribossômicas e as moléculas de rRNA normalmente são designadas em unidade de Svedberg (S – quanto maior o valor de S, maior a velocidade de sedimentação), uma medida de taxa de sedimentação de partículas em suspensão centrifugadas sob condições padrão, nome designado em homenagem ao químico sueco Theodor Svedberg (1884-1971), o inventor da centrífuga.

Em procariotos (bactérias e arqueias), a menor subunidade do ribossomo apresenta velocidade de sedimentação de 30 unidades de Svedberg, sendo, portanto, uma subunidade 30S, enquanto a subunidade maior é uma 50S. A subunidade menor é formada por uma única molécula de rRNA, referida como RNA pequeno (16S), e 21 proteínas ribossômicas, enquanto a subunidade maior contém uma molécula de rRNA grande (23S) mais uma molécula de rRNA 5S, além de 31 proteínas ribossômicas. O ribossomo procariótico como um todo é 70S, pois a velocidade de sedimentação é determinada pelo volume do material, e não por sua medida de massa. Já nas células eucarióticas, a subunidade menor do ribossomo é a 40S, e a maior 60S, sendo o ribossomo total 80S, com uma unidade menor formada pelo 18S e 33 proteínas ribossômicas, e a subunidade maior formada pelos rRNA 28S, 5S e 5,8S, além de 49 proteínas ribossômicas (Figura 7.6).

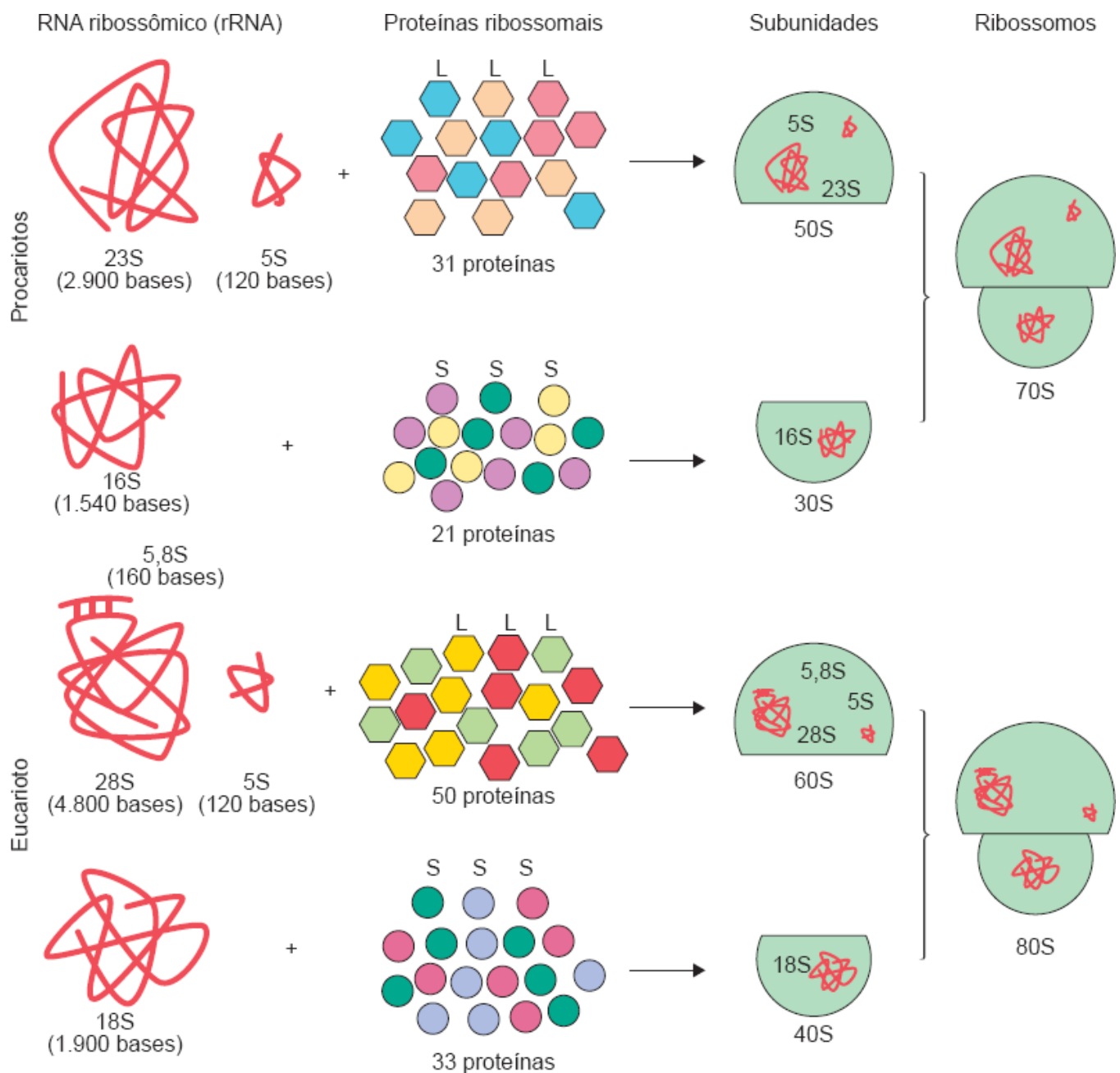


Figura 7.6 Composição dos ribossomos procarióticos e eucarióticos.

Apesar dessas diferenças, há uma enorme semelhança estrutural e funcional entre os ribossomos de todas as espécies, o que reflete a origem evolutiva comum dos constituintes básicos das células vivas. Justamente por esse motivo, a sequência dos genes que codificam o rRNA é usada em estudos filogenéticos, para se verificar a distância genética entre diferentes indivíduos.

Funções do RNA ribossômico

O rRNA não se trata apenas de um componente estrutural dos ribossomos, sendo também um determinante catalítico para a síntese proteica. Embora o ribossomo e suas funções básicas tenham sido descobertos há mais de 50 anos, apenas recentemente foi determinada sua estrutura tridimensional em alta definição, que ajudou na confirmação de uma hipótese levantada há décadas: o ribossomo é uma ribozima (molécula de RNA com atividade catalítica). Foi demonstrado que o rRNA participa diretamente das funções do ribossomo, de modo que as proteínas ribossômicas são elementos secundários e que não há proteínas dentro do centro ativo da formação da ligação peptídica.

As duas subunidades ribossômicas de formato irregular ajustam-se para formar uma fenda pela qual o mRNA passa à medida que o ribossomo se movimenta ao longo dele. Na subunidade maior dos ribossomos procarióticos, o centro da peptidil-transferase, já mencionado como responsável pela formação da ligação peptídica, é composto inteiramente por RNA. Na subunidade menor, o RNA também desempenha importante papel, na qual as alças dos

anticódons dos tRNA, assim como os códons do mRNA, interagem diretamente com o 16S rRNA, e não com as proteínas ribossômicas. As proteínas ribossômicas estão localizadas geralmente na periferia e não no interior do ribossomo, e as que alcançaram o núcleo da subunidade parecem estabilizar os rRNA compactados firmemente, protegendo-as das cargas negativas de seus esqueletos de açúcar-fosfato. As 55 proteínas ribossômicas nas bactérias apresentam tamanho e formato variados, e, embora a função de algumas delas não tenha sido completamente esclarecida, o papel estrutural na estabilização da molécula tem sido o mais evidente.

RNA ribossômicos são altamente conservados

Como já descrito, os ribossomos são partículas de riboproteínas complexas responsáveis pela síntese de proteínas em todas as células, seja de um eucarioto ou de um procarioto. Existe um volume de evidências que apontam que a essência da função ribossômica reside em seus componentes de rRNA. Uma vez que a função dessas moléculas é essencial para vida da célula e ocorre em todos os organismos, é esperado que sua sequência seja altamente conservada para garantir a estabilidade da estrutura do rRNA, basicamente composta por várias ramificações contendo *hairpin loops* e *loops* internos.

Os *hairpin loops* (em grampo), também conhecidos como *stem loop* ou *loops* externos, têm comprimento variável de 2 a 14 nucleotídeos que pareiam entre si dentro da alça e apresentam uma porção terminal da dupla-hélice não pareada. Já os *loops* internos são formados por dois segmentos de RNA dupla-hélice que não pareiam entre si, e podem ser simétricos se contiverem o mesmo número de nucleotídeos entre as fitas opostas, ou assimétricos. Ainda podem se apresentar na forma de bolhas ou protuberâncias laterais (Figura 7.7).

Regiões com resíduos conservados nos rRNA têm sido estudadas há muito tempo e, tanto para o gene 16S (procariotos) como para o 18S (eucariotos) rRNA existem múltiplos segmentos de 15 a 20 nucleotídeos que são exatamente os mesmos em todos os organismos. Esses segmentos são essenciais para a atividade catalítica do ribossomo e permaneceram inalterados por bilhões de anos durante a evolução dos seres vivos.

Um alinhamento preciso do sítio A-(aminoacil) com o P-(peptidil) no centro da peptidil-transferase é essencial para a função do ribossomo. A estrutura em cristal da subunidade 50S isolada revela que essa peptidiltransferase é um complexo terciário de elementos de RNA, principalmente do domínio V do 23S rRNA. No rRNA 16S foram identificados diversos nucleotídeos que também interagem com os sítios ribossômicos A e P. Esses resíduos são conservados em praticamente todos os organismos e aparecem como um elemento estrutural constante filogeneticamente.

Pareamento de RNA mensageiros e ribossomos

Como já descrito anteriormente, a informação para a síntese proteica está na forma de códons de 3 nucleotídeos, e cada códon especifica 1 aminoácido. As regiões que codificam 1 proteína de 1 mRNA são compostas por uma sucessão contínua de códons não sobrepostos, chamadas de fase aberta de leitura ou ORF (*open reading frame*), pois a leitura códon a códon segue continuamente do códon de iniciação até o de terminação, definindo uma única proteína. Assim, a tradução inicia-se na extremidade 5' do mRNA no códon de iniciação, que nas bactérias geralmente é um 5'-AUG-3', mas também pode ser 5'-GUG-3' e, mais raramente, 5'-UGG-3'; enquanto nas células eucarióticas o códon de iniciação é sempre 5'-AUG-3'. Esse códon de iniciação determina o primeiro aminoácido a ser incorporado à cadeia polipeptídica em crescimento, assim como também define a fase de leitura para todos os códons subsequentes – uma vez que a leitura pode ser feita em três fases de leitura diferentes. A tradução segue códon a códon sem sobreposição, até alcançar o códon de terminação, que pode ser uma das três opções (5'-UAG-3', 5'-UGA-3' e 5'-UAA-3'), no qual, então, termina a síntese do polipeptídeo.

A tradução ocorre pela leitura do mRNA na região entre o códon de iniciação e o de terminação (ORF), de 3 em 3 bases, sem sobreposição de bases. Para que isso ocorra, o primeiro evento necessário é a aproximação do ribossomo ao mRNA. O que facilita essa ligação são os procariotos de uma sequência que ocorre em várias fases de leitura a montante do códon de iniciação, chamada de sítio de ligação ao ribossomo ou RBS (*ribosome binding site*). Este sítio de ligação ao ribossomo, em bactérias, também é conhecido como sequência de Shine-Dalgarno, em homenagem aos cientistas australianos John Shine (1946-) e Lynn Dalgarno (1935-), que identificaram essa sequência por meio da comparação de vários mRNA. O sítio de ligação do ribossomo é localizado geralmente de 3 a 9 pares de bases em relação a 5' do códon de iniciação, sendo complementar a uma sequência localizada próxima a extremidade 3' de um dos componentes do rRNA 16S, de modo que o ribossomo faz, por meio desse rRNA

(sequência 5'-CCUCCU-3'), pontes de hidrogênio ao mRNA (sequência 5'-AGGAGG-3'), posicionando o ribossomo no início da fase aberta de leitura. Isso favorece o pareamento inicial da subunidade menor do ribossomo ao mRNA.

O grau de complementaridade e a distância entre o sítio de ligação ao ribossomo e o códon de iniciação exerce forte influência sobre a atividade de tradução de determinada fase de leitura, que pode ser ativa, se tiver uma alta complementaridade e o espaçamento apropriado; ou menos ativa, se a complementaridade for baixa ou o espaçamento limitado. Em procariotos, algumas ORF localizadas em mRNA policistrônicos (vários genes) não apresentam forte ligação ao ribossomo, no entanto, são traduzidas ativamente, pois o códon de iniciação se sobrepõe à extremidade 3' da fase de leitura adjacente. Nesse caso, o ribossomo que recém-completou a tradução de uma fase de leitura a montante já está adequadamente posicionado para iniciar a tradução a partir do códon de iniciação da fase de leitura a jusante, dispensando a necessidade de uma sequência Shine-Dalgarno para recrutar um ribossomo. Esse fenômeno de tradução contínua entre duas fases de leitura é conhecido como tradução acoplada. Apesar de o aspecto geral do processo de iniciação da tradução se assemelhar a procariotos, o método de reconhecimento do mRNA e do códon de iniciação nos eucariotos difere significativamente daquele observado em células procarióticas. Em eucariotos, a ligação do tRNA iniciador à subunidade menor do ribossomo sempre precede sua associação com o mRNA na extremidade 5'-*cap* em um sistema denominado cap-dependente, o qual envolve o acoplamento de uma série de fatores de iniciação que reconhecem essa região.

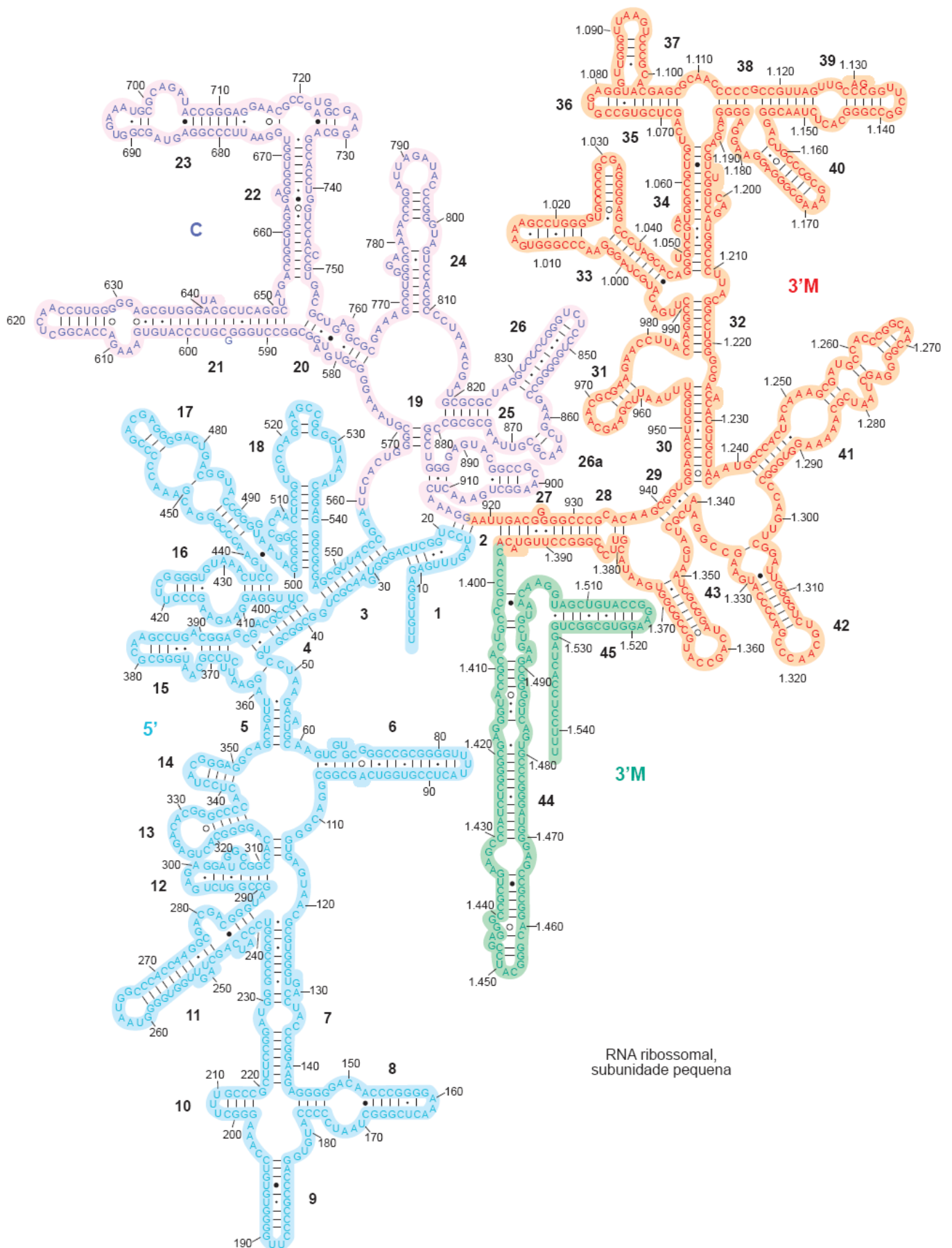


Figura 7.7 Estrutura secundária da subunidade menor do RNA ribossômico de bactéria (16S rRNA). Adaptada de Center for Molecular Biology of RNA.

Início de tradução do RNA mensageiro

O primeiro aminoácido é a formilmetionina em bactérias e a metionina em eucariontes.

Assim que a menor subunidade do ribossomo procariótico pareia sobre o mRNA, fixando-o na posição ideal, de modo que o códon de iniciação esteja em um sítio chamado P, a subunidade maior do ribossomo poderá se associar ao complexo. De fato, a associação da maior subunidade do ribossomo só ocorre imediatamente antes da formação da 1ª ligação peptídica. O caminho normal dos tRNA carregados é ingressar pelo sítio A e encaminhar para o sítio P após a síntese de uma ligação peptídica, porém, durante a iniciação da tradução, um tRNA carregado entra diretamente no sítio P do ribossomo, sendo este tRNA especial. Esse tRNA especial, conhecido como tRNA iniciador, pareia com o códon de iniciação (AUG ou GUG), e que, em vez de resultarem na leitura dos aminoácidos metionina e valina, respectivamente, são carregados por uma forma modificada de metionina, a N-formilmetionina (fMet-tRNA_i^{fMet}). Esse grupo formila não permanece no peptídeo final, pois é removido da extremidade da cadeia polipeptídica pela enzima desformilase, assim como muitas vezes a própria metionina N-terminal é removida por peptidases juntamente com 1 ou 2 aminoácidos, tanto durante como após a síntese final da cadeia polipeptídica. Diferentemente dos procariotos, o tRNA iniciador está carregado de metionina convencional, sendo, portanto, denominado de Met-tRNA_i^{Met}.

Os dois sítios de ligação de RNA transportador no ribossomo

Os ribossomos bacterianos dispõem de dois sítios para ligarem pelo menos 2 tRNA simultaneamente. *A* é o sítio de ligação para o aminoacil-tRNA (ligado em sua extremidade 3' à carboxila do aminoácido carregado), e *P* é o sítio de ligação para o peptidil-tRNA (ligado em sua extremidade 3' na posição C-terminal da cadeia polipeptídica crescente). Existe um 3º sítio, chamado *E* (de *exit*, saída), que é o sítio de ligação do tRNA liberado após a transferência da cadeia polipeptídica crescente para o aminoacil-tRNA. Esses sítios de ligação ao tRNA são formados na interface entre as subunidades maior e menor do ribossomo, podendo os tRNA permanecerem ligados a uma distância que abrange entre o centro da peptidil-transferase na subunidade maior e o centro de decodificação, na subunidade menor.

Após a formação do ribossomo com o tRNA iniciador carregado no sítio P, o primeiro aminoacil-tRNA carregado correto é trazido ao sítio A do ribossomo, de acordo com o códon do mRNA que está nesse sítio. A partir disso, uma ligação peptídica é formada entre o aminoacil-tRNA no sítio A e a cadeia peptídica ligada ao peptídeo-tRNA no P, ligação chamada de reação de peptidil-transferase, uma vez que a cadeia peptídica é transferida do P para o aminoácido do tRNA no A. Esse peptídeo-tRNA resultante no sítio A e o seu códon associado são translocados para P, de modo que A fica livre para o próximo ciclo de reconhecimento.

Fatores de iniciação, alongação e terminação

Os fatores de iniciação têm por função orientar o processo inicial da tradução, a fim de facilitá-lo. Nas células procarióticas, o processo de tradução iniciado na subunidade menor é catalisado por três fatores de iniciação, chamados de IF1, IF2 e IF3. Esses fatores de iniciação ligam-se a um dos três sítios de ligação de tRNA da subunidade menor, ou em seus proximidades. A IF1 liga-se diretamente à porção da subunidade menor que originará o sítio A, bloqueando a ligação de tRNA carregados no A; o IF2, que é uma GTPase, interage com três componentes-chave: a subunidade menor do ribossomo, a IF-1 sobre o sítio A, e o fMet-tRNA_i^{Met} do tRNA iniciador no sítio P, tendo como função impedir que outros tRNA carregados se associem à subunidade menor; e, por fim, a IF3, que se liga à subunidade menor em parte do sítio que dará origem ao sítio E, impedindo a sua reassociação com a subunidade maior ou a ligação de tRNA carregados, auxiliando na dissociação do ribossomo 70S, deixando-o livre para um novo ciclo de tradução.

Nas células eucarióticas, a dissociação das subunidades maior e menor do ribossomo 80S é realizada pelos fatores de iniciação eIF3 e eIF1A, análogos aos IF3 e IF1 procarióticos, e o recrutamento do tRNA iniciador carregado por duas proteínas de ligação ao GTP, a eIF2 e a eIF5B (análogo ao IF2-GTP procariótico) para formar o complexo de iniciação 43S (Figura 7.8). O eIF5B associa-se à subunidade menor de modo independente de eIF1A, e ajuda a recrutar um complexo de IF2-GTP e Met-tRNA_i^{Met} para a subunidade menor. Juntas, essas duas proteínas de ligação ao GTP posicionam a Met-tRNA_i^{Met} no futuro sítio P da subunidade menor. O complexo de pré-iniciação 43S faz o reconhecimento do 5'-cap na extremidade do mRNA eucariótico, sendo mediado por uma proteína de 3 subunidades, a eIF4F. Uma das três subunidades liga-se diretamente a 5'-cap e as outras duas ligam-se de modo inespecífico ao RNA associado. Quando esse complexo se une ao eIF4F, uma RNA-helicase é ativada, removendo estruturas secundárias que possam ter se formado na extremidade do mRNA, deixando-a livre para interações com a

[illegible]

Assim como no processo de início da tradução, a elongação também necessita de fatores para levar a termo o processo de extensão da cadeia. Os fatores de elongamento são um conjunto de proteínas especiais que recrutam os aminoacil-tRNA e auxiliam na exatidão do posicionamento do mRNA em relação à maquinaria de síntese polipeptídica, garantindo sua correta fase de leitura. Os aminoacil-tRNA não se ligam diretamente aos ribossomos, sendo direcionados pelo fator de elongamento eEF1A. Esse fator liga a extremidade 3' do aminoacil-tRNA, encobrindo o aminoácido ligado, impedido sua interação para formação de uma cadeia polipeptídica. Essa ligação do eEF1A ao aminoacil-tRNA depende da associação com uma molécula de GTP. Para que o eEF1A seja liberado do tRNA, deixando-o livre para polimerizar na cadeia polipeptídica, é necessária a hidrólise do GTP pela função

GTPase (Figura 7.9). A atividade da GTPase é muito importante para o mecanismo de seleção contra a adição de aminoacil-tRNA incorretos na cadeia, pois, uma vez ocorrido o pareamento incorreto de pelo menos uma base entre códon e anticódon, acontece uma redução drástica na atividade desta enzima.

Após essa reação da peptidil-transferase, o tRNA do sítio P é desacetilado (não mais ligado a um aminoácido), e a cadeia polipeptídica crescente é ligada ao tRNA no sítio A. Para que se dê um novo ciclo de alongamento, um conjunto de movimentos chamado de *translocação* deve ocorrer, o que envolve a transferência do tRNA desacetilado de P para E, do aminoacil-tRNA de A para P, e o deslocamento em 3 nucleotídeos do mRNA, para exposição do códon seguinte. O término desse processo de translocação requer a ação de um 2º fator de alongação, o *EF-G*, também dependente da ligação de GTP. Esse fator de alongamento EF-G liga-se a uma porção do sítio A após o deslocamento do tRNA desacetilado, ocorrendo a hidrólise do GTP ligado a ele, o que resulta na alteração da sua conformação para uma forma EF-G GDP. Isso desencadeia o conjunto de movimentos da translocação, que, após o término da translocação completa do ribossomo, perde sua afinidade e libera o fator de alongação (Figura 7.9).

A tradução é concluída quando ocorre a translocação do ribossomo, e o códon de terminação encontra-se em A. Para o término da tradução, fatores de liberação (RF) fazem a coordenação. O fator eRF1 (RF1 ou RF2 em procariotos) reconhece o códon de terminação e se liga ao ribossomo (sítio A). O eRF3 (RF3) liga-se a um GTP, e o complexo eRF3-GTP atua coordenado com o eRF1 para clivar a ligação peptidil-tRNA, fazendo a liberação do peptídeo e do tRNA (Figura 7.10).

Direção de leitura do RNA 5'-3'

A leitura do mRNA para a síntese proteica começa do códon de iniciação da extremidade 5', seguindo adiante até a extremidade 3'. Como descrito anteriormente, durante o início da tradução, a extremidade 5' do mRNA é ligada a fatores de iniciação, formando o chamado complexo de iniciação. Esse complexo provavelmente desliza e escaneia o mRNA à procura do códon de iniciação, que é o primeiro AUG a jusante na extremidade 5' na maioria dos RNA eucarióticos, e AUG ou GUG nos RNA procarióticos. Nos eucariotos, essa seleção do AUG de iniciação é auxiliada por nucleotídeos específicos que o flanqueiam, denominados sequência de Kozak, em homenagem à pesquisadora Marilyn Kozak, que identificou os nucleotídeos mais importantes que afetam a eficiência do início da tradução, tanto a jusante como a montante do códon de iniciação. Assim, com o correto posicionamento da maquinaria de síntese, a tarefa de adição de aminoácidos pode ser iniciada por meio da tradução do mRNA sentido 5' → 3'.

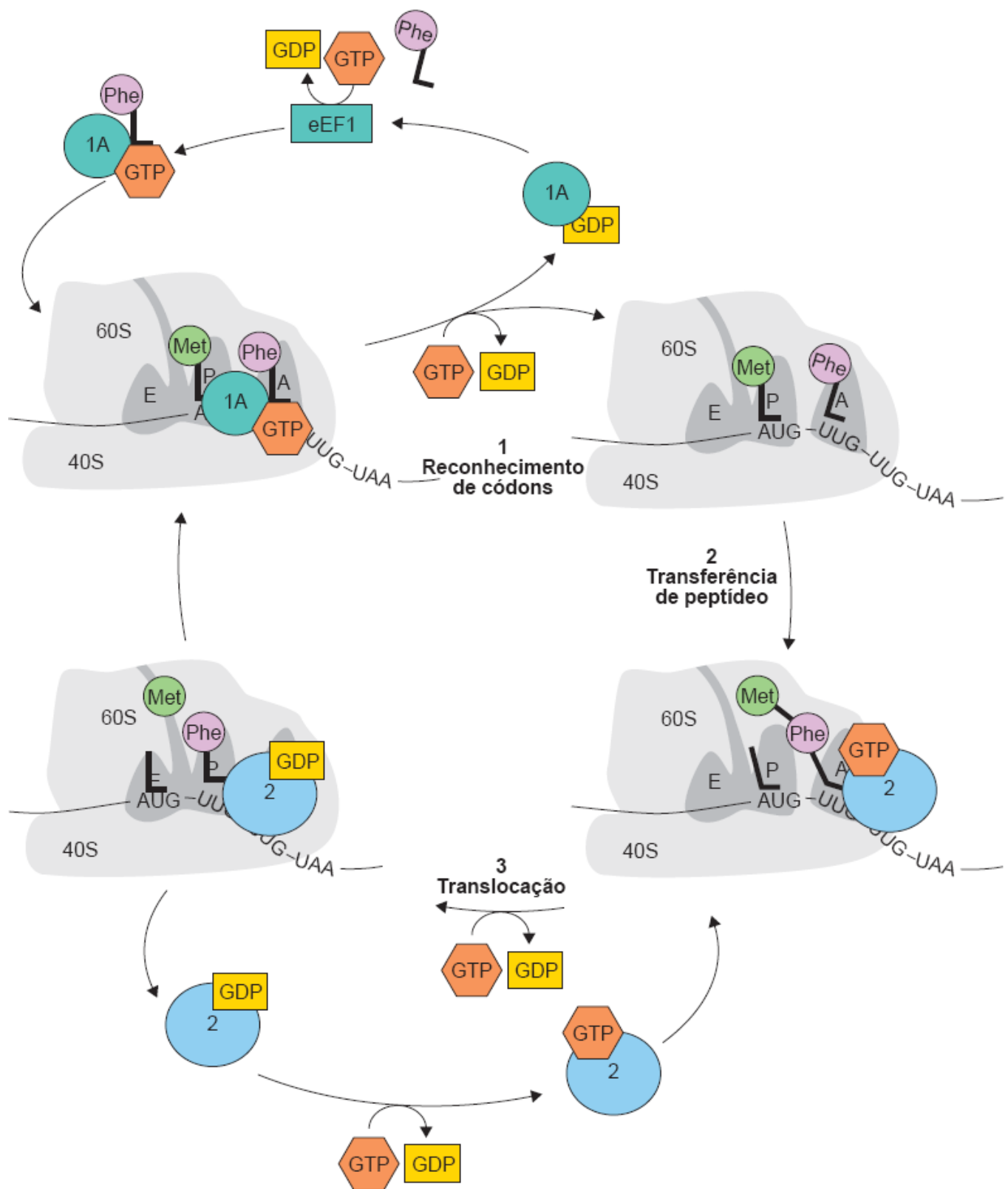


Figura 7.9 Fatores de alongamento para ligação do segundo aminoacil-tRNA.

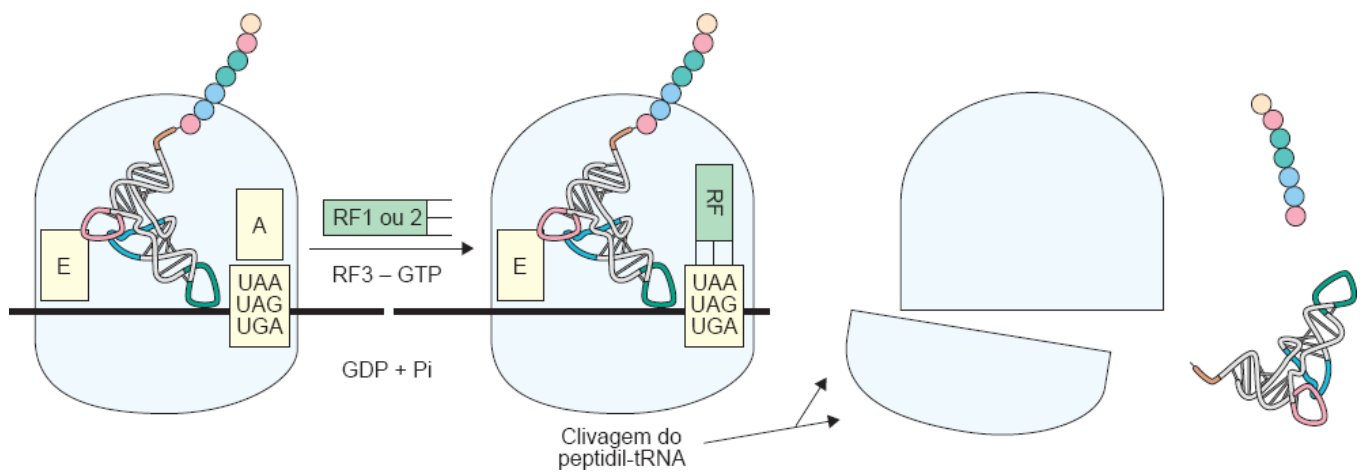


Figura 7.10 Término da tradução. A liberação do peptídeo inicia com o reconhecimento dos códons de término pelo fator de liberação.

Antibióticos podem atuar de forma específica na tradução

Antibióticos são compostos que interagem com microrganismos, matando-os ou inibindo o seu metabolismo e/ou sua reprodução, sendo, portanto, uma eficaz ferramenta no combate de doenças infecciosas. Como apresentam reduzido efeito sobre células eucarióticas, muitos antibióticos são utilizados no controle de inúmeras infecções que no século 20 não podiam ser tratadas, como a tuberculose, a pneumonia bacteriana, a sífilis e a gonorreia. Os antibióticos podem atuar em diferentes alvos na célula, como impedir a síntese de parede celular, alterar a permeabilidade da membrana, inibir enzimas que atuam na transcrição do DNA, bloquear a síntese de ácidos graxos, entre inúmeros outros mecanismos. Apesar desses diversos alvos, 40% dos antibióticos conhecidos são inibidores da maquinaria de tradução e inibem seu funcionamento. Além disso, dentro dos antibióticos que tem como alvo a inibição da síntese proteica, diferentes etapas da síntese podem ser alvo. Vários exemplos são encontrados para antibióticos que inibem algum processo da tradução, como a bem conhecida tetraciclina, utilizada no combate de infecções por *Chlamydia*, *Mycoplasma* e *Rickettsia*, além de acne e sífilis. A tetraciclina é um antibiótico bacteriostático de amplo espectro que tem como alvo molecular o sítio A da subunidade 30S, inibindo a ligação do aminoacil-tRNA ao sítio A. Os efeitos de outros antibióticos constam na Tabela 7.2.

Tabela 7.2 Efeitos dos antibióticos.

Antibiótico	Células-alvo	Alvo molecular	Consequência
Ácido fusídico	Procarióticas	EF-G	Impede a liberação do EF-G-GDP do ribossomo
Aminoglicosídeos	Procarióticas	Subunidade 30S do ribossomo	Paralisa o complexo de iniciação 30S (30S-mRNA-tRNA), inibindo a iniciação posterior. Também ocorre inibição da síntese proteica já iniciada e indução de erro de leitura do mRNA
Ciclohexamida	Eucarióticas	Centro da peptidil-transferase da subunidade 60S	Inibe a atividade da peptidil-transferase
Cloranfenicol	Procarióticas	Centro da peptidil-transferase da subunidade 50S	Bloqueia o posicionamento correto da aminoacil-tRNA no sítio A para a reação de transpeptidase
Eritromicina	Procarióticas	Canal de saída do peptídeo na subunidade 50S	Bloqueia a saída da cadeia polipeptídica crescente no ribossomo, suspendendo a tradução
Espectinomicina	Procarióticas	Subunidade 30S do ribossomo	Interfere reversivelmente na interação entre a subunidade 30S do ribossomo e o mRNA.

			Estruturalmente similar aos aminoglicósidos, mas não induz erros na leitura do mRNA
Higromicina B	Procarióticas e eucarióticas	Próximo ao sítio A da subunidade 30S	Impede a translocação do tRNA do sítio para o sítio P
Paromicina	Procarióticas e eucarióticas	No sítio A da subunidade 30S, adjacente ao sítio da interação códon-anticódon	Aumenta a taxa de erros durante a tradução, pela diminuição da seletividade entre códon e anticódons
Quirromicina	-	EF-Tu	Impede as alterações conformacionais associadas à hidrólise de GTP
Tioestreptona	Procarióticas	Centro de ligação do fator da subunidade 50S	Interfere na associação do IF-2 e do EF-G com o centro de ligação do fator
Toxina diftérica	Eucarióticas	Modificação química do EF-Tu	Inibe a função do EF-Tu

Um RNA mensageiro, várias proteínas

Conforme visto anteriormente, o RNA atua como um intermediário entre a informação codificada no DNA para especificar uma sequência de aminoácidos de uma proteína. Esse intermediário é o mRNA. Em procariotos, uma única molécula de mRNA pode codificar tanto para uma como para várias cadeias polipeptídicas. Quando uma molécula de mRNA transporta o código para apenas um polipeptídeo, o mRNA é chamado de monocistrônico (“cístron” refere-se a um gene); se codifica para dois ou mais polipeptídeos, o mRNA é então denominado policistrônico. Os mRNA policistrônicos contêm dois ou mais ORF e codificam para mais de uma cadeia polipeptídica. Cada gene desse mRNA tem seu códon de iniciação, mas geralmente é regulado por um único promotor. Ao contrário dos eucariotos, em que a maioria dos mRNA são monocistrônicos, nos procariotos a maioria dos mRNA são policistrônicos, normalmente incluindo proteínas com funções correlatas, por exemplo, proteínas que participam de uma mesma via de síntese. É importante ressaltar que tanto características morfológicas como fisiológicas (fenótipo) de uma célula estão diretamente relacionadas ao seu conteúdo de mRNA. Células que se multiplicam muito rapidamente necessitam de diferentes proteínas em um curto intervalo de tempo. Uma maneira de atender a essa exigência seria a expressão de vários genes em um único mRNA, sendo muito mais econômico em termos de energia e mais eficaz em termos de rapidez, principalmente porque a expressão gênica em bactérias é regulada com maior importância ao nível do mRNA.

A síntese proteica em bactérias ocorre simultaneamente à transcrição do RNA mensageiro

Nas bactérias, a tradução e a degradação de uma molécula de mRNA em geral começam antes de a síntese (transcrição DNA → mRNA) estar completa. Como universalmente as moléculas de mRNA são sintetizadas, traduzidas e degradadas no sentido 5' → 3', os três processos podem ocorrer simultaneamente na mesma molécula de RNA. Nos eucariotos, o sentido universal também é esse, porém a transcrição e a tradução ocorrem em locais diferentes dentro da célula, no núcleo e no citoplasma, respectivamente. Nos procariotos, a maquinaria de síntese de polipeptídeos não é separada por uma membrana nuclear do sítio de síntese do mRNA, de modo que, uma vez sintetizada a ponta 5' de um mRNA, ele pode imediatamente ser usado como molde para a síntese de polipeptídeo, sendo os dois processos acoplados. Para confirmar se a transcrição e a tradução poderiam ocorrer simultaneamente, Miller *et al.* desenvolveram técnicas de microscopia eletrônica, que tornaram possível o acoplamento entre esses dois processos diretamente em *E. coli*.

Modificações pós-traducionais de proteínas em eucariontes

A elongação da cadeia peptídica continua até que o ribossomo adicione o último aminoácido codificado pelo mRNA, que pode ser um códon UAA, UAG ou UGA, no processo chamado de terminação. Esse término é auxiliado por três fatores de terminação, ou de liberação, sendo as proteínas RF-1, RF-2 e RF-3 nos procariotos, e dois fatores, RF-1 e RF-3, nos eucariotos. Esses fatores de terminação ou liberação reconhecem os códons de terminação e contribuem para a hidrólise da ligação terminal do peptidil-tRNA, a liberação do polipeptídeo livre e o último tRNA agora descarregado, a partir de P, e a dissociação do ribossomo em sua subunidade maior e menor. Ao término da síntese proteica, a proteína ainda não está pronta para exercer sua função, sendo necessário que ela seja enovelada e processada até sua forma biológica ativa. Durante ou depois da sua síntese, o polipeptídeo assume progressivamente sua conformação tridimensional natural, com a formação de pontes de hidrogênio e as interações van der Waals e hidrofóbicas apropriadas. Outras proteínas, porém, mesmo depois de adquirir sua conformação tridimensional precisam passar por reações de processamento, chamadas modificações pós-tradução, para atingirem sua conformação biologicamente ativa. Por fim, as proteínas devem ser endereçadas para o seu sítio de ação por meio de sequências sinalizadoras.

Degradação de proteínas pelo sistema proteossoma

A degradação de proteínas é necessária para impedir a construção de proteínas anormais ou indesejadas, e possibilita a reciclagem de aminoácidos. A vida de proteínas eucarióticas é de trinta segundos a vários dias, porém, a maioria é degradada rapidamente se comparada à vida média de uma célula. As proteínas rapidamente degradadas incluem aquelas defeituosas, seja pela adição incorreta de aminoácidos ou por lesão acumulada durante o seu funcionamento normal, e também as proteínas enzimáticas que agem em pontos-chave da regulação nas vias metabólicas e geralmente degradam rapidamente. Para que essas proteínas sejam degradadas seletivamente, um conjunto de sinalizadores moleculares é embutido na estrutura de cada proteína, de modo análogo ao que ocorre com a sinalização para endereçamento das proteínas para o seu sítio de ação.

Essas proteínas defeituosas geralmente são degradadas em sistemas citosólicos dependentes de ATP, tanto em células procarióticas como eucarióticas. Nos procariotos, muitas proteínas são degradadas por uma protease dependente de ATP, a Lon (referindo-se à forma longa), a qual é ativada por essas proteínas defeituosas, nas quais a protease hidrolisa a ligação peptídica, clivando essas proteínas, a fim de reduzi-las em pequenos peptídeos inativos. Nas células eucarióticas, esse processo de degradação de proteínas dependente de ATP é bem diferente, e envolve uma proteína chamada de ubiquitina. A ubiquitina é encontrada em todos os eucariotos, sendo altamente conservada, e se liga covalentemente às proteínas destinadas à degradação por um sistema de três enzimas separadas (E1 – ativador da ubiquitina; E2 – conjugadores de ubiquitina; e E3 – ligases de ubiquitina-proteína), processo chamado de ubiquitinação. As proteínas multiubiquitinadas são degradadas pelo proteossoma 26S, uma protease multimérica dependente de ATP, com aproximadamente 2000 kDa, constituída por dois complexos regulatórios 19S e por um complexo catalítico de aproximadamente 700 kDa, designado proteossoma 20S. Os complexos regulatórios 19S, quando posicionados em ambas as extremidades do proteossoma 20S, composto por várias ATPases e outras subunidades, reconhecem as proteínas ubiquitinadas, conduzindo ao seu desenrolamento e translocamento para a câmara catalítica, geralmente degradando as proteínas a pequenos peptídeos (Figura 7.11).

A síntese da proteína ocorre no citoplasma

Apesar do DNA, que guarda a informação genética contida no núcleo, a síntese de proteínas ocorre nos ribossomos, dentro do citoplasma. Esse processo compartimentado ocorre nas células eucarióticas, pois elas apresentam o núcleo delimitado por uma membrana. Isso é possível graças ao mRNA, que é capaz de transportar a mensagem genética do núcleo ao citoplasma. Desse modo, é possível dizer que o processo de síntese de proteínas tem início no núcleo da célula, com a transcrição dos genes do DNA em mRNA, terminando no citoplasma, com a tradução do mRNA em proteínas nos ribossomos. O RNA foi evidenciado como candidato natural a ser responsável por esse transporte desde o início da década de 1950, por ser encontrado tanto no núcleo quanto no citoplasma, e também por ter sido observado que o aumento da síntese de proteínas era acompanhado pelo aumento na quantidade de RNA citoplasmático e na sua velocidade de renovação. Já nos procariotos, embora a localização do DNA genômico não seja ao acaso no citoplasma, este não se encontra delimitado por uma membrana nuclear como nos eucariotos. Dessa maneira, a transcrição também ocorre no citoplasma bacteriano, assim como a síntese de proteínas nos ribossomos.

Tráfego intracelular de proteínas

A maioria das proteínas, exceto aquelas sintetizadas em mitocôndrias e plastídeos, é sintetizada nos ribossomos no citosol. No entanto, deve-se pensar como eles são direcionados do citoplasma para o sítio de ação final, haja vista que a célula eucariótica é constituída de muitas estruturas, compartimentos e organelas. Proteínas que serão secretadas, integradas na membrana ou incluídas nos lisossomos geralmente compartilham as primeiras etapas de uma via que se inicia no retículo endoplasmático (RE). Já proteínas que se destinam para as mitocôndrias, os cloroplastos ou núcleo usam três mecanismos separados. E, por fim, as proteínas destinadas ao citoplasma simplesmente permanecem onde elas são sintetizadas.

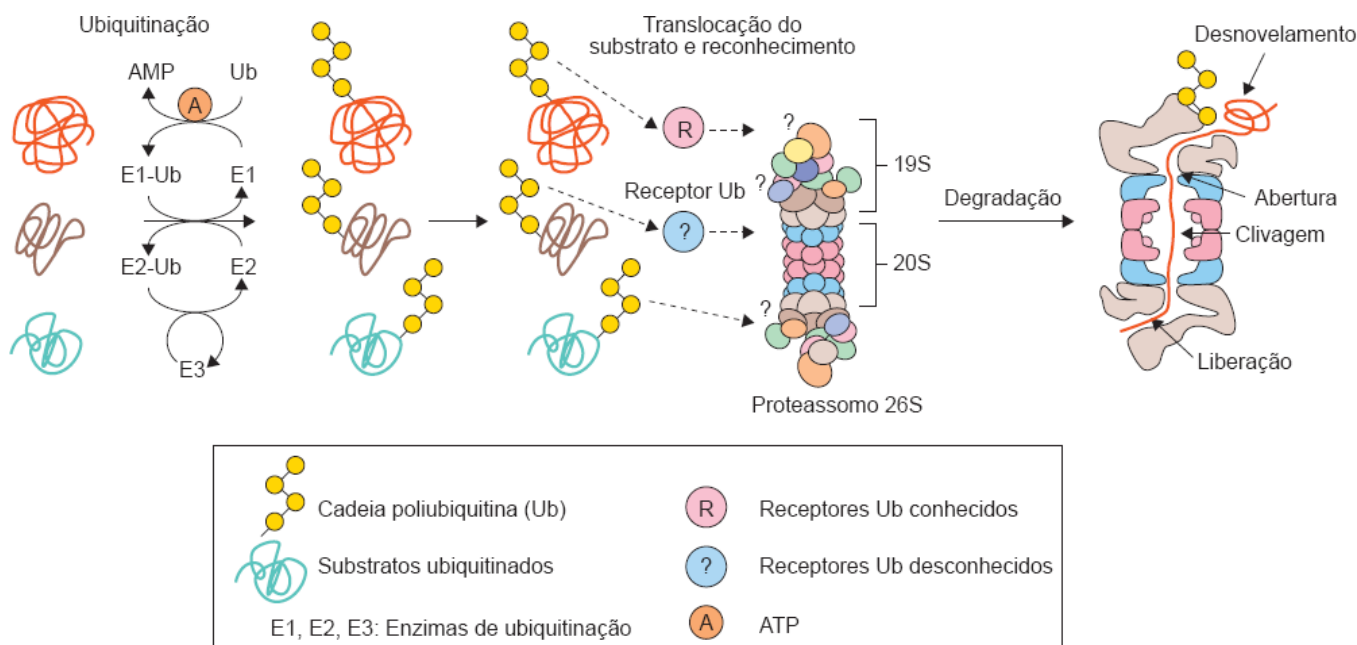


Figura 7.11 Ubiquitinação de proteínas como sinalização de proteínas a serem degradadas pelo proteossoma 26S. Adaptada de Huang.

De modo geral, o elemento mais importante para o endereçamento é a adição de uma sequência curta de aminoácidos chamada de *sequência sinalizadora*, cuja função foi primeiramente postulada por Blobel *et al.*, em 1979. Essa sequência sinalizadora direciona uma proteína para sua localização apropriada na célula e, para muitas proteínas, é removida durante o transporte ou depois que ela tenha alcançado seu destino final.

A maior parte das proteínas secretadas do lisossomo ou das membranas apresenta uma sequência sinalizadora no terminal amino que as marca para a translocação para dentro do lúmen do RE. Essa sequência sinalizadora dispõe de um terminal carboxila e é definido por um sítio de clivagem, de modo que uma protease remove a sequência de sinalização depois que a proteína foi importada para o interior do RE. Como demonstrado por George Palade (1912-2008), as proteínas com essas sequências sinalizadoras são sintetizadas nos ribossomos ligados ao RE (esse trabalho rendeu o Prêmio Nobel de Fisiologia e Medicina de 1999 a Palade). Essas sequências sinalizadoras sintetizadas nos ribossomos do RE são ligadas pela partícula de reconhecimento da sinalização (SRP), juntamente ao ribossomo livre que está sintetizando o polipeptídeo. Essa SRP, então, se liga ao GTP e para a elongação do polipeptídeo quando ele apresenta cerca de 70 aminoácidos, passando a direcionar o ribossomo (ainda ligado ao mRNA) e o polipeptídeo incompleto para os receptores de SRP ligada ao GTP na fase citosólica do RE. Assim, o polipeptídeo crescente é entregue ao complexo de translocação do peptídeo no RE, a SRP dissocia-se do ribossomo, acompanhada pela hidrólise do GTP, e o alongamento do polipeptídeo prossegue com o complexo de translocação direcionado por ATP, introduzindo o polipeptídeo crescente dentro do lúmen do RE até que a proteína completa tenha sido sintetizada (Figura 7.12). Assim, ao término da síntese, a sequência sinalizadora é removida por uma peptidase de sinalização dentro do lúmen do RE, e o ribossomo dissociado é reciclado.

Uma vez no lúmen do RE, as proteínas recém-sintetizadas são ainda modificadas de várias maneiras. Após a remoção da sequência sinalizadora, os polipeptídeos são enovelados, as ligações dissulfetos formadas e muitas proteínas glicosiladas para formar glicoproteínas, que, após adequadamente modificadas, podem ser movimentadas a uma variedade de destinos intracelulares. As proteínas viajam do RE para o complexo de Golgi em vesículas

transportadoras, e, no complexo de Golgi, esses oligossacarídeos são ligados a proteínas e posteriormente modificados. Da mesma maneira, por mecanismos desconhecidos, o complexo de Golgi também distribui as proteínas e as envia para o seu destino final, que se distingue por características estruturais entre essas proteínas.

Tendo como exemplo as hidrolases destinadas ao lisossomo, ao chegar ao complexo de Golgi, essa hidrolase é reconhecida por uma fosfotransferase que adiciona fosfato em certos resíduos de manose no oligossacarídeo. Um ou mais resíduos de manose 6-fosfato no oligossacarídeo N-ligado constitui a sinalização estrutural que direciona a proteína ao lisossomo.

Vias que endereçam proteínas à mitocôndria e aos cloroplastos também dependem de sequências sinalizadoras no terminal amino. Embora essas organelas tenham DNA próprio, a maior parte de suas proteínas é sintetizada pelo DNA nuclear e precisam ser endereçadas para a organela apropriada. Ao contrário das outras proteínas, a inserção da sequência sinalizadora nessas proteínas só ocorre depois que elas estão completamente sintetizadas e liberadas do ribossomo. Essas proteínas precursoras destinadas à mitocôndria e/ou cloroplastos são ligadas a proteínas chaperonas citosólicas e entregues aos receptores na superfície externa da organela-alvo. Mecanismos de translocação especializados transportam a proteína ao seu destino final na organela, e, ao final, a sequência sinalizadora também é removida.

As bactérias também podem endereçar proteínas para as suas membranas internas, para o espaço periplasmático entre essas membranas ou para o meio extracelular. Para isso, também usam sequências sinalizadoras no terminal amino, muito parecidas com a dos eucariotos. Ao iniciar seu lento enovelamento pós-síntese, a proteína chaperona solúvel SecB liga-se à sequência sinalizadora da proteína ou a outros sítios não enovelados de sua estrutura, de modo que, ao se associar, entrega essa proteína a SecA, uma proteína associada à superfície interna da membrana plasmática, na qual a SecB dissocia-se dessa ligação. A SecA atua como a ATPase de translocação dos eucariotos, sendo esse complexo de translocação constituído ainda de SecY, E e G. Assim, uma proteína é exportada pela membrana pela SecA localizada na superfície citoplasmática. Embora essa seja a via principal de translocamento de proteínas, elas podem usar uma via alternativa, que utiliza um reconhecimento de sinalização e proteínas receptoras homólogas aos componentes do SRP eucariótico e do receptor SRP.

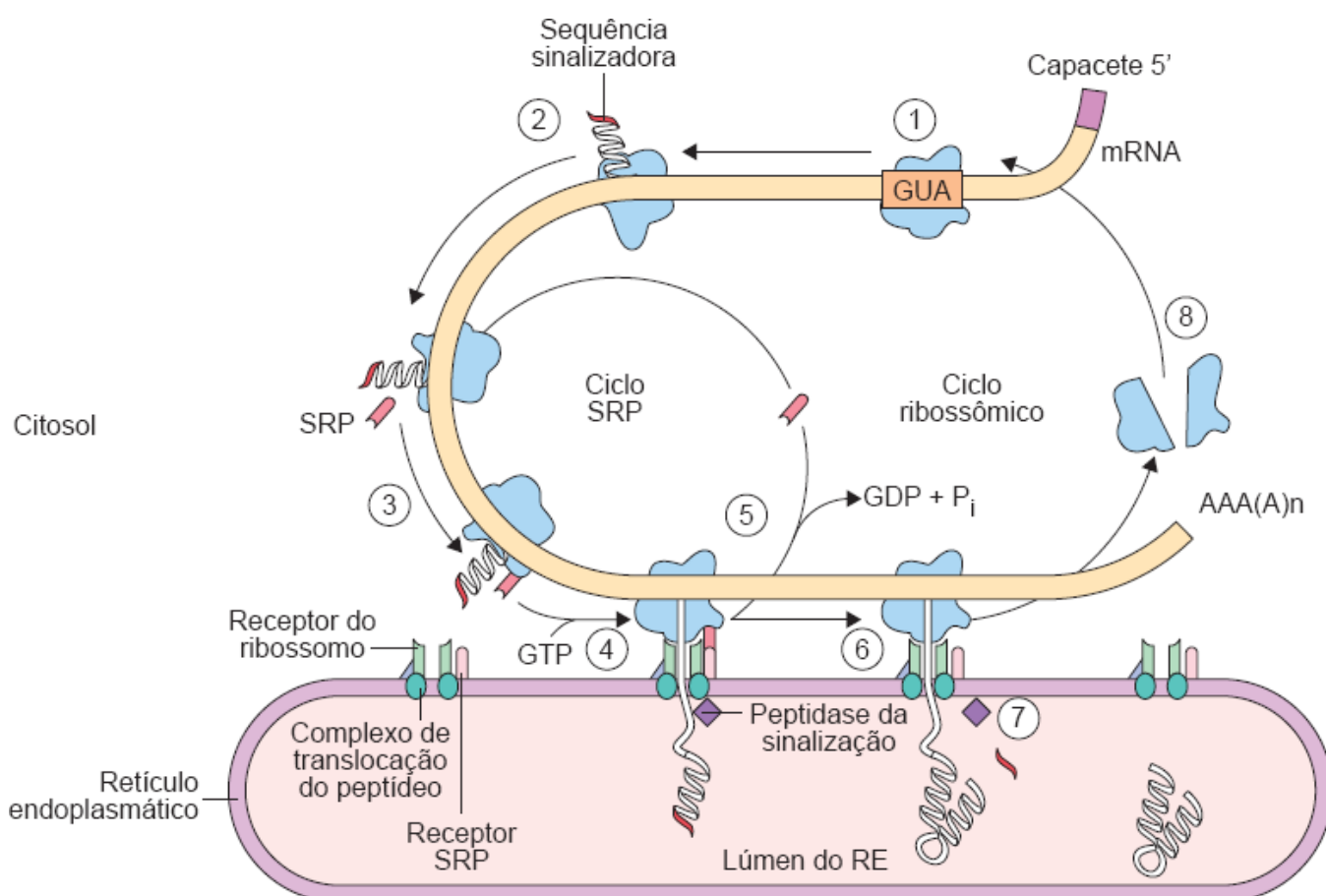


Figura 7.12 Direcionamento das proteínas eucarióticas com a sinalização apropriada para o retículo endoplasmático.

Modificações proteicas

Conforme já mencionado, apenas a tradução do mRNA para proteína pela maquinaria de síntese polipeptídica não é suficiente para que as proteínas sejam funcionais, sendo necessário um processo de modificação pós-traducional a fim de que as proteínas cheguem a sua conformação biologicamente ativa. Essas proteínas podem ser alteradas por uma ou mais reações de processamento, discutidas a seguir.

- Perda da sequência de sinalização: como visto anteriormente, embora as sequências de sinalização sejam importantes para o endereçamento das proteínas, elas não participam de sua estrutura ativa final, sendo removidas por peptidases específicas
- Modificações no terminal amino e no terminal carboxila: de acordo com o que já foi explicado, o 1º resíduo inserido em todos os polipeptídeos bacterianos é a N-formilmetionina ou metionina dos eucarióticos. Entretanto, o grupo formila geralmente é removido enzimaticamente na formação da proteína funcional final, e, em 50% das proteínas eucarióticas, o grupo amino do resíduo terminal amino é N-acetilado depois da tradução. A porção terminal carboxila às vezes é também modificada
- Processamento proteolítico: muitas proteínas são sintetizadas na forma de polipeptídeos e posteriormente são clivadas proteoliticamente em peptídeos menores e ativos, como o que ocorre com a pró-insulina, algumas proteínas virais e proteases, como o quimiotripsinogênio e o tripsinogênio
- Modificação de aminoácidos individuais: os grupos hidroxila de alguns resíduos podem ser enzimaticamente fosforilados por ATP, adicionando cargas negativas a esses polipeptídeos, o que pode ter um significado funcional variável de uma proteína para outra. Outras modificações podem incluir a adição de grupos carboxil extras, assim como a metilação de resíduos, removendo sua carga negativa
- Fixação das cadeias laterais de carboidratos: nessa modificação, cadeias laterais de carboidratos das glicoproteínas podem ser fixadas de maneira covalente durante ou depois da síntese de polipeptídeos, em resíduos de Asn, Ser e Thr. Exemplos de proteínas que contêm cadeias laterais de oligossacarídeos são as proteoglicanas que lubrificam as camadas mucosas das membranas e de muitas outras proteínas que funcionam extracelularmente
- Formação de pontes de dissulfeto: após o enovelamento das proteínas em sua estrutura tridimensional nativa, as proteínas podem formar pontes dissulfeto intra ou intercadeias entre resíduos de Cys, o que ajuda a proteger a conformação da molécula da desnaturação no ambiente extracelular quando exportada
- Adição de grupos prostéticos: muitas proteínas procarióticas e eucarióticas necessitam da adição de grupos prostéticos para se tornarem funcionais. Nesse grupo encontram-se a molécula de biotina da acetil-CoA carboxilase e o grupo heme da hemoglobina, ou do citocromo C
- Adição de grupos isoprenil: ocorre em proteínas eucarióticas pela adição de grupos derivados do isopreno (grupos isoprenil). Uma ligação tioéter é formada entre grupos isoprenil e um resíduo de Cys da proteína. São ocorrentes em proteínas Ras, produtos de oncogenes Ras e protooncogenes, e proteína G.

Perspectivas futuras

Um dos mais importantes achados do projeto Genoma ocorreu com a determinação da diversidade de DNA existente em diferentes indivíduos, que deu início a uma revolução no campo da genética, uma nova época denominada Era Genômica. Ultrapassada essa nova fronteira da ciência, o conhecimento da genômica foi complementado por meio da abordagem proteômica, auxiliando o estudo de proteínas expressas em determinados tipos de tecido ou células. Esses estudos demonstraram que a diversidade/quantidade de proteínas em um organismo é muito maior que a quantidade de genes, mostrando que existem mecanismos ainda pouco conhecidos de amplificação da informação contida nos genes, o que resulta como função diferente em proteínas.

Todo esse enfoque está sendo cada vez mais aplicado na descoberta de novos alvos para terapia gênica e/ou de fármacos para tratamento de doenças, na farmacogenômica, em estudos de envelhecimento, planejamento de vetores virais, administração de doenças, compreensão de mecanismos fisiológicos, biologia regenerativa, melhoramento genético, clonagem, diagnóstico, prognóstico, pesquisa clínica, reprodução, entre outros.

Dessa maneira, entender os mecanismos que regulam a síntese de proteínas é imperativo na busca de soluções nos campos ambiental, industrial, da medicina e da agronomia, visando ao aumento de produtividade de animais, plantas e microrganismos, além do tratamento de seres vivos e do meio ambiente.

Bibliografia

- Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. *Biologia molecular da célula*. 4. ed. Porto Alegre: Artmed; 2002.
- Blobel G, Walter P, Chang CN, Goldman B, Erickson A, Lingappa VR. Translocation of proteins across membranes: the signal hypothesis and beyond. In: Hopkins CR, Duncan CJ, editors. *Symposium of the Society of Experimental Biology (Great Britain)*. V. 33. London: Cambridge University Press; 1979. p. 9-36.
- Center of Molecular Biology of RNA. Ribosome images [online]. s. d. University of California, Santa Cruz. [Acesso em 14 jun 2016] Disponível em: http://rna.ucsc.edu/rnacenter/ribosome_images.html
- Hendrix DK, Brenner SE, Holbrook SR. RNA structural motifs: building blocks of a modular biomolecule. *Quarterly Reviews of Biophysics*. 2006;1:1-23.
- Hershlo A, Ciechavoner A, Varshavsky A. The ubiquitin system. *Nature Medicine*. 2000;69:617-50.
- Huang L. Functional proteomics of macromolecular machines [online]. Irvine, CA: University of California, Irvine; s.d. [Acesso em 15 jun 2016] Disponível em: <https://webfiles.uci.edu/lanhuang/www/>
- Krebs JE, Goldstein ES, Kilpatrick ST. *Lewin's Genes X*. New York: Oxford University Press; 2011.
- Lodish H, Berk A, Matsudaira P, Kaiser CA, Krieger M, Scoot MP, et al. *Biologia celular e molecular*. 5. ed. Porto Alegre: Artmed; 2005.
- Miller OL Jr, Hamkalo BA, Thomas CA Jr. Visualization of bacterial genes in action. *Science*. 1970;169(3943):392-5.
- Nelson DL, Cox MM. *Lehninger, princípios de bioquímica*. 4. ed. São Paulo: Sarvier; 2006.
- Smit S, Widmann J, Knight R. Evolutionary rates may vary among rRNA structural elements. *Nucleic Acids Research*. 2007;35:3339-54.
- Snustad DP, Simmons MJ. *Fundamentos de genética*. 2 ed. Rio de Janeiro: Guanabara Koogan; 2001.
- University of Tokyo. A comprehensive approach to Life Science: web textbook [online]. Aminoacyl-tRNA synthesis. 2011. [Acesso em 14 jun 2016] Disponível em: http://csls-text3.c.u-tokyo.ac.jp/active/09_01.html
- Watson JD, Baker TA, Bell SP, Gann A, Levine M, Losick R. *Biologia molecular do gene*. 5. ed. Porto Alegre: Artmed; 2006.