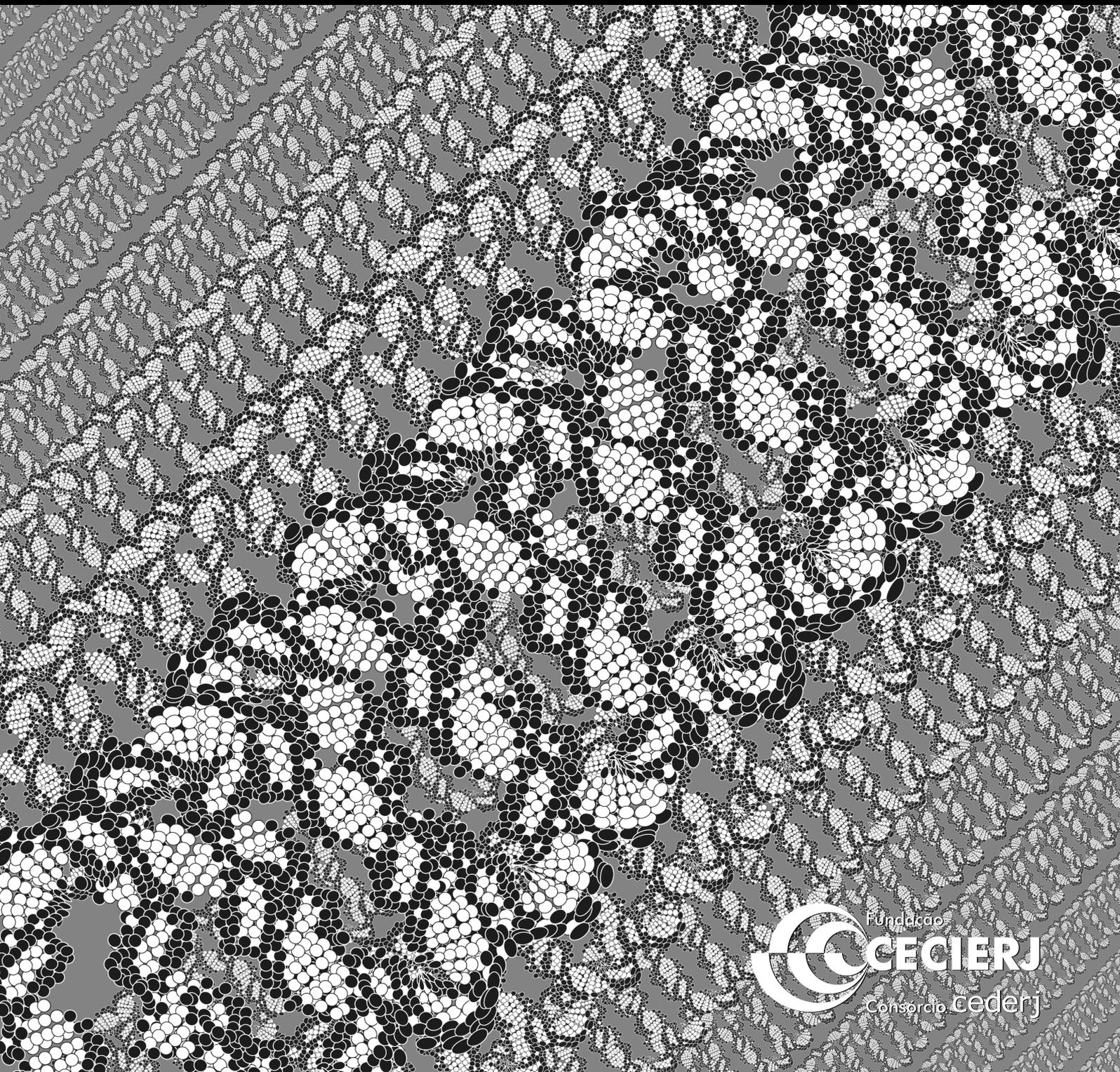


Gonçalo A. De Souza Filho  
Jaclyra M. B. Macedo

Biologia Molecular







Fundação

**CECIERJ**

Consórcio **cederj**

Centro de Educação Superior a Distância do Estado do Rio de Janeiro

## Biologia Molecular

Volume 3 - Módulos 3 e 4

Gonçalo A. de Souza Filho

Jacyara M. B. Macedo



SECRETARIA DE  
CIÊNCIA E TECNOLOGIA

Ministério  
da Educação



Apoio:



Fundação Carlos Chagas Filho de Amparo  
à Pesquisa do Estado do Rio de Janeiro

# Fundação Cecierj / Consórcio Cederj

Rua Visconde de Niterói, 1364 – Mangueira – Rio de Janeiro, RJ – CEP 20943-001

Tel.: (21) 2299-4565 Fax: (21) 2568-0725

## Presidente

Masako Oya Masuda

## Vice-presidente

Mirian Crapez

## Coordenação do Curso de Biologia

UENF - Milton Kanashiro

UFRJ - Ricardo Iglesias Rios

UERJ - Cibele Schwanke

## Material Didático

### ELABORAÇÃO DE CONTEÚDO

Gonçalo A. de Souza Filho

Jacyara M. B. Macedo

### COORDENAÇÃO DE DESENVOLVIMENTO INSTRUCIONAL

Cristine Costa Barreto

### DESENVOLVIMENTO INSTRUCIONAL E REVISÃO

Ana Tereza de Andrade

Anna Maria Osborne

Carmen Irene Correia de Oliveira

### COORDENAÇÃO DE LINGUAGEM

Maria Angélica Alves

### REVISÃO TÉCNICA

Marta Abdala

## Departamento de Produção

### EDITORA

Tereza Queiroz

### COPIDESQUE

Jane Castellani

### REVISÃO TIPOGRÁFICA

Kátia Ferreira dos Santos

Patrícia Paula

### COORDENAÇÃO DE PRODUÇÃO

Jorge Moura

### PROGRAMAÇÃO VISUAL

Andréa Dias Fiães

Márcia Valéria de Almeida

### ILUSTRAÇÃO

Fábio Muniz

### CAPA

Fábio Muniz

### PRODUÇÃO GRÁFICA

Andréa Dias Fiães

Fábio Rapello Alencar

Copyright © 2005, Fundação Cecierj / Consórcio Cederj

Nenhuma parte deste material poderá ser reproduzida, transmitida e gravada, por qualquer meio eletrônico, mecânico, por fotocópia e outros, sem a prévia autorização, por escrito, da Fundação.

S729b

Souza, Gonçalo A. de.

Biologia molecular. v. 3. - Gonçalo A. de Souza. – Rio de Janeiro : Fundação CECIERJ, 2008.

158p.; 19 x 26,5 cm.

ISBN: 85-89200-77-9

1. Informação genética. 2. Processamento de proteínas. 3. Genomas. I. Macedo, Jacyara M.B. II. Título.

CDD: 572.8

# Governo do Estado do Rio de Janeiro

**Governador**  
Sérgio Cabral Filho

**Secretário de Estado de Ciência e Tecnologia**  
Alexandre Cardoso

## Universidades Consorciadas

**UENF - UNIVERSIDADE ESTADUAL DO  
NORTE FLUMINENSE DARCY RIBEIRO**  
Reitor: Almy Junior Cordeiro de Carvalho

**UERJ - UNIVERSIDADE DO ESTADO DO  
RIO DE JANEIRO**  
Reitor: Ricardo Vieiralves

**UFF - UNIVERSIDADE FEDERAL FLUMINENSE**  
Reitor: Roberto de Souza Salles

**UFRJ - UNIVERSIDADE FEDERAL DO  
RIO DE JANEIRO**  
Reitor: Aloísio Teixeira

**UFRRJ - UNIVERSIDADE FEDERAL RURAL  
DO RIO DE JANEIRO**  
Reitor: Ricardo Motta Miranda

**UNIRIO - UNIVERSIDADE FEDERAL DO ESTADO  
DO RIO DE JANEIRO**  
Reitora: Malvina Tania Tuttman



### SUMÁRIO

#### Módulo 3

- Aula 24** - Regulação da expressão gênica em eucariotos I \_\_\_\_\_ **7**  
*Gonçalo A. de Souza Filho*
- Aula 25** - Regulação da transcrição em eucariotos \_\_\_\_\_ **25**  
*Gonçalo A. de Souza Filho*
- Aula 26** - Fluxo da informação genética – tradução I \_\_\_\_\_ **41**  
*Jacyara M. B. Macedo*
- Aula 27** - Fluxo da informação genética – tradução II \_\_\_\_\_ **59**  
*Jacyara M. B. Macedo*
- Aula 28** - Fluxo da informação genética – tradução III \_\_\_\_\_ **79**  
*Jacyara M. B. Macedo*

#### Módulo 4

- Aula 29** - Processamento e endereçamento de proteínas \_\_\_\_\_ **101**  
*Jacyara M. B. Macedo*
- Aula 30** - Complexidade dos genomas I \_\_\_\_\_ **115**  
*Gonçalo A. de Souza Filho*
- Aula 31** - Complexidade dos genomas II \_\_\_\_\_ **133**  
*Gonçalo A. de Souza Filho*
- Gabarito** \_\_\_\_\_ **145**



# Regulação da expressão gênica em eucariotos I

AULA

# 24

## objetivos

Ao final desta aula, você deverá ser capaz de:

- Entender as diferenças na regulação da expressão gênica entre procariotos e eucariotos.
- Estudar as diferenças etapas nas quais a expressão gênica de eucariotos pode ser regulada.

### Pré-requisitos

Para acompanhar mais facilmente esta aula, é importante que você tenha assimilado as informações sobre Regulação da expressão gênica em procariotos, Aula 23.

Adicionalmente, reveja os conceitos sobre Fluxo da informação gênica – transcrição e processamento do RNA, apresentados nas Aulas 20, 21 e 22.

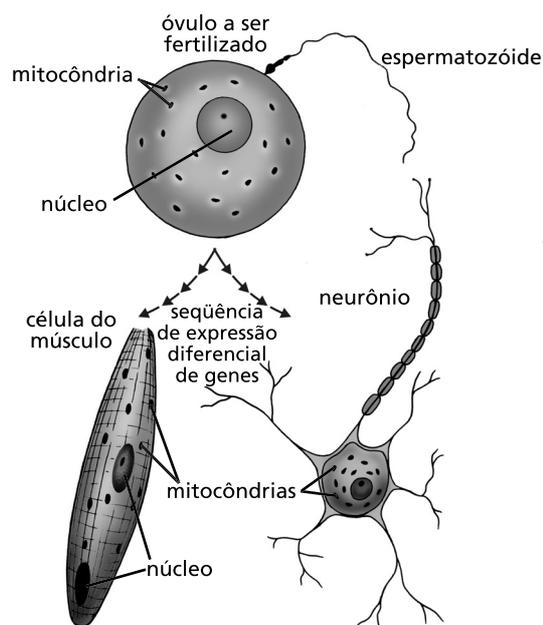
## INTRODUÇÃO

Na Aula 23, você teve a oportunidade de estudar os mecanismos que regulam a expressão gênica em procariotos. Nela, você viu que é muito útil à célula procariótica a capacidade de responder ao ambiente, através da ativação da transcrição de determinados genes e/ou a inativação da transcrição de outros. Tais mecanismos permitem à célula adaptar-se aos diferentes ambientes e, simultaneamente, otimizar seu gasto de energia e nutrientes, mantendo desligados aqueles genes cujos produtos sejam desnecessários.

Os mecanismos que você estudou na Aula 23 estão envolvidos na resposta de organismos unicelulares (procariotos) ao ambiente. Nesta aula, você terá a oportunidade de estudar os mecanismos envolvidos na regulação da expressão gênica em eucariotos. Muitos eucariotos são organismos multicelulares, cujos diversos tipos celulares estão organizados em tecidos e órgãos. Fica evidente, portanto, que se trata de um sistema mais sofisticado de regulação da expressão gênica.

## REGULAÇÃO DA EXPRESSÃO GÊNICA EM ORGANISMOS MULTICELULARES

Um organismo eucarioto multicelular apresenta diferentes tipos de células. Assim, se compararmos células do fígado, pele, neurônios, músculos e outros tecidos de um mesmo animal, perceberemos que elas são muito distintas, tanto nos aspectos morfológicos quanto funcionais. Adicionalmente, a constituição de cada tipo celular é específica (Figura 24.1).



**Figura 24.1:** Esquema ilustrativo do “dramático” efeito da expressão gênica diferencial durante o desenvolvimento em animais superiores. Uma célula, zigoto, dá origem a uma vasta gama de fenótipos celulares diferentes no organismo adulto. As células de músculo e de neurônio são apenas exemplos dentre os muitos tipos celulares que exibem fenótipos altamente divergentes em animais.

Sabemos que as células somáticas de um indivíduo possuem o mesmo conjunto de genes, pois todas se originaram a partir do mesmo zigoto, através de uma sucessão de divisões mitóticas equacionais. Assim, se o genótipo é o mesmo em todas essas células, como podemos explicar a grande diversidade de fenótipos observada?

A resposta geral que podemos apresentar para tal questão é que nem todos os genes presentes no núcleo se expressam ao mesmo tempo. De fato, estima-se que, em organismos eucariotos superiores, apenas 10% dos genes se expressam em todas as células durante todo o tempo. Os demais genes são ativados em determinados tecidos ou em momentos específicos. Existe, portanto, uma regulação da expressão gênica, que define os genes que serão ativados em um determinado tecido. Adicionalmente, muitos genes são ativados apenas em certos momentos ou períodos do desenvolvimento.

Um ótimo exemplo para compreendermos a complexidade da regulação da expressão gênica em eucariotos superiores é o processo de diferenciação celular. Além de os diferentes tipos celulares apresentarem morfologias altamente divergentes, alguns são altamente especializados, com funções metabólicas específicas. Como exemplos, podemos citar as hemácias, altamente especializadas em sintetizar e estocar hemoglobina, e os neurônios, capazes de produzir neurotransmissores. Que mecanismos definem os genes que estarão ativos em um dado tipo celular e inativos em outros?

## **CASCATAS DE REGULAÇÃO DA EXPRESSÃO GÊNICA**

A partir da fecundação, o processo de desenvolvimento embrionário segue passos claramente definidos de multiplicação e diferenciação celular. De tal modo, após poucas divisões mitóticas, os conjuntos celulares começam a se diferenciar para formar tecidos e órgãos distintos que, ao final, resultarão em um organismo completo e funcional. Tal processo é denominado morfogênese, e envolve um intrincado sistema de regulação gênica, cujos componentes permanecem pouco esclarecidos.

Diversos estudos têm revelado que os processos de diferenciação celular envolvem a participação de “genes reguladores chave”, que controlam a expressão de genes específicos para a formação de cada

órgão ou tecido. Dentre os genes ativados, por sua vez, existem alguns que também possuem a função de regular novos genes, e assim sucessivamente. Você pode concluir, com isso, que existe uma “cascata” de regulação da expressão gênica, na qual um determinado gene pode ativar a expressão de vários outros, dentre os quais novos genes que passarão a regular os futuros passos do processo. A morfogênese, portanto, trata-se de um processo coordenado de regulação seqüencial da expressão gênica.

As informações anteriores demonstram que os vários caminhos de diferenciação celular, desenvolvidos a partir do zigoto, seguem um “circuito pré-programado de expressão gênica”, no qual um evento inicial (fecundação) ativou um conjunto de genes. Dentre os genes ativados, alguns têm função reguladora, atuando tanto na ativação de um segundo conjunto de genes quanto na inativação de genes do primeiro conjunto. Desse segundo conjunto de genes, alguns são também reguladores, ativando a expressão de um terceiro conjunto de genes, e assim por diante. A ordem seqüencial da expressão desses genes é, portanto, geneticamente pré-programada.

Outros exemplos de processos que utilizam “cascatas de expressão gênica” são os mecanismos regulados por hormônios. Nesses casos, a presença de um dado hormônio ativa a expressão seqüencial de genes específicos (cascatas específicas de resposta).

Observe que os mecanismos de expressão seqüencial de genes não ocorrem somente durante o desenvolvimento embrionário. Cascatas de expressão gênica continuam ativas em cada indivíduo durante toda a sua vida. Por exemplo, determinados conjuntos de genes permanecem inativos em nós durante toda a nossa infância, mas são ativados a partir da puberdade. Tal evento estava geneticamente pré-programado, e é uma interessante demonstração de que as cascatas de expressão gênica são mecanismos essenciais em eucariotos.

## CLASSIFICAÇÃO DOS GENES QUANTO À SUA REGULAÇÃO

Determinados genes são continuamente expressos na maioria das células. Tais genes são responsáveis pela manutenção das funções essenciais de qualquer célula, tais como transcrição, tradução, replicação etc. A expressão destes genes é denominada expressão constitutiva e os genes são denominados genes constitutivos.

Os demais genes são, portanto, regulados. Apesar de muitos genes comporem “circuitos de expressão geneticamente pré-programados” (cascatas de regulação), diversos outros são ativados ou reprimidos em resposta a estímulos ambientais, tais como frio, calor, estresse, ferimentos etc. Por terem sua expressão passível de indução pelo ambiente, são denominados genes induzíveis.

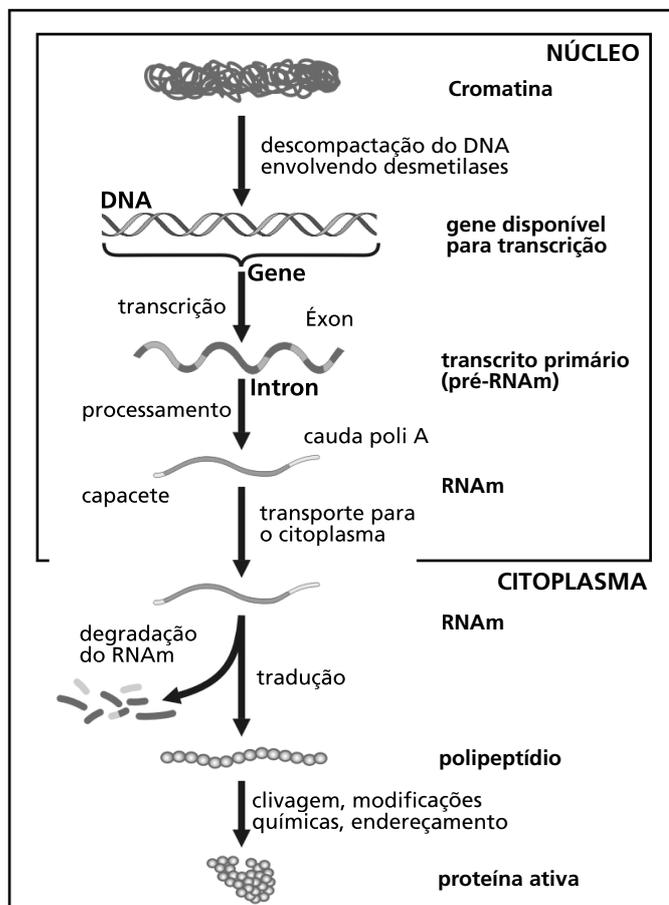
Os genes que se expressam apenas em determinados tecidos são chamados genes tecido-específicos e estão sob regulação tecido-específica. Aqueles genes que se expressam apenas durante determinados períodos são controlados através de regulação temporal.

## ETAPAS NAS QUAIS A EXPRESSÃO GÊNICA PODE SER REGULADA

Agora que você já teve a oportunidade de ver que a expressão gênica em eucariotos é regulada de diferentes maneiras para atender a objetivos específicos, é importante saber como a célula define quais genes devem ser ativados em um dado tecido, enquanto outros devem permanecer desligados.

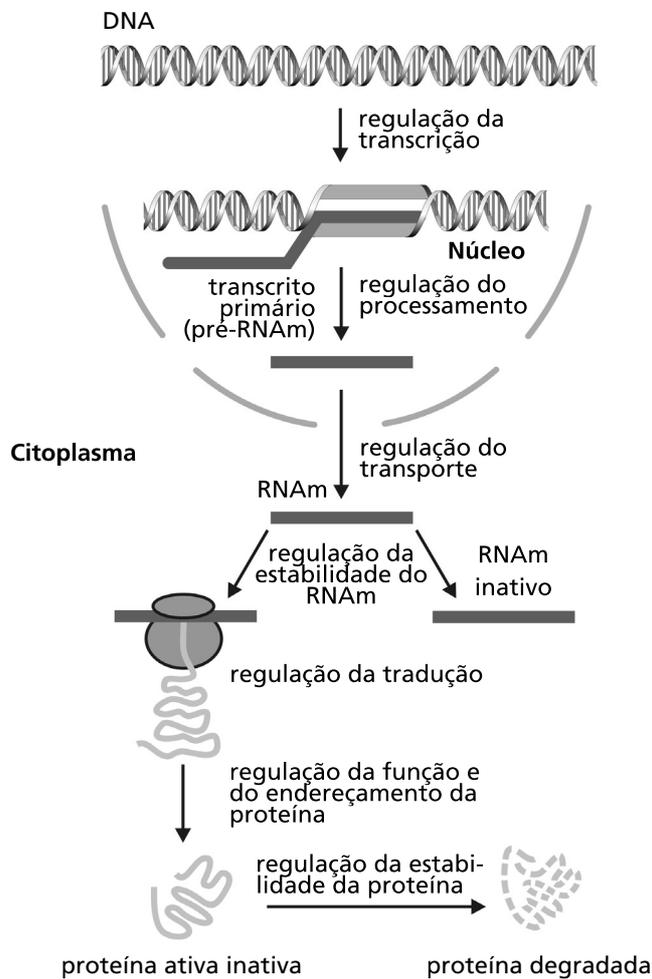
Antes, porém, é importante relembrar um conceito sobre expressão gênica que foi visto na Aula 23. A expressão de um gene só se verifica se o polipeptídeo por ele codificado for produzido e estiver funcional. Portanto, para que um gene presente no núcleo se expresse, devem ocorrer as seguintes etapas: transcrição, processamento do transcrito, transporte para o citoplasma, tradução, endereçamento e ativação do polipeptídeo formado (Figura 24.2).

**Figura 24.2:** Esquema ilustrando as várias etapas envolvidas na expressão gênica em eucariotos. A cromatina deve ser descompactada para que os genes sejam transcritos. Em seguida, o pré-RNAm deve ser processado para a produção do RNAm que, após ser transportado para o citoplasma, será interpretado para a tradução. O polipeptídeo formado durante a tradução é, então, submetido a modificações (clivagens e modificações químicas) e endereçado ao compartimento celular, onde comporá a proteína com atividade biológica.



Esses diversos passos, essenciais para a expressão de cada gene, podem ser individualmente regulados. Em outras palavras, se algum destes passos sofrer interferência, a quantidade ou a funcionalidade do produto final do gene em questão serão também afetadas.

Para desenvolver a regulação gênica, a célula pode interferir em qualquer um dos passos já descritos (Figura 24.3). Cada gene pode ser regulado de maneira diferente quanto à etapa do processo a ser ativada ou reprimida. Determinados genes sofrem regulação em várias etapas de sua expressão. A seguir, estudaremos os mecanismos que podem ser utilizados pela célula para regular cada uma destas etapas.



**Figura 24.3:** Esquema ilustrativo das diversas etapas nas quais a expressão de um gene de eucarioto pode ser regulada.

## REGULAÇÃO DA TRANSCRIÇÃO

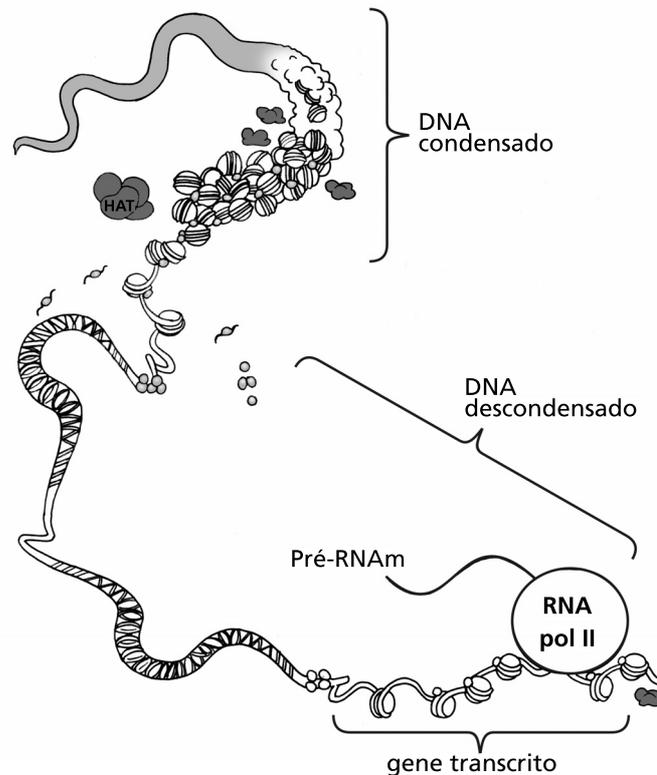
Este é o principal nível em que a expressão dos genes pode ser regulada. Para a célula, regular a expressão gênica através do controle da transcrição é mais econômico, pois, se o gene não for transcrito, nenhum dos passos posteriores ocorrerá. Devido a sua importância, a regulação transcricional da expressão gênica será abordada em maior profundidade na próxima aula (Aula 25). A seguir, veremos apenas os aspectos gerais dos principais mecanismos.

A regulação da transcrição pode ocorrer através de dois processos distintos: Remodelamento da cromatina e Ligação de proteínas reguladoras, descritos a seguir.

### Regulação da transcrição por Remodelamento da Cromatina

Na Aula 8 (Módulo 1), você teve a oportunidade de aprender que o DNA de eucariotos está associado a proteínas chamadas histonas, formando estruturas denominadas nucleossomos. A interação entre proteínas histonas e outras proteínas (não-histônicas), permite que vários nucleossomos se associem, formando estruturas maiores. Assim, regiões do DNA de eucariotos que contenham grande quantidade de nucleossomos associados serão mais condensadas que as demais regiões. Em determinadas fases do desenvolvimento celular, é interessante que o DNA se encontre altamente condensado, por exemplo, durante a divisão celular.

Regiões condensadas do genoma são, normalmente, transcricionalmente inativas. Portanto, para que ocorra a transcrição dos genes presentes, é necessário que o DNA esteja descompactado (Figura 24.4). O nível de condensação das diferentes regiões dos cromossomos pode ser controlado e, dessa forma, podem ser definidas as regiões que estarão disponíveis para a transcrição em determinado tecido, ou estágio de desenvolvimento do organismo. Portanto, a célula controla a disponibilidade de tais regiões através de mecanismos de remodelamento da cromatina.



**Figura 24.4:** Esquema ilustrando a correlação entre o nível de compactação da cromatina e a ocorrência de transcrição. A região do genoma esquematizada contém segmentos em diferentes níveis de condensação. Em regiões altamente compactadas não se verifica transcrição. A transcrição é observada em regiões onde o DNA não está compactado (DNA descondensado).

Embora os mecanismos através dos quais a célula controla os níveis de condensação do cromossomo ainda permaneçam pouco elucidados, diversos trabalhos têm demonstrado a ocorrência de modificações na estrutura do DNA e das proteínas histonas capazes de provocar tal controle. Tais modificações correspondem à adição, ou remoção, de radicais à molécula de DNA ou às proteínas histonas a ele associadas. Dentre essas modificações, destacam-se os processos de metilação do DNA e metilação, acetilação e fosforilação das proteínas histonas.

Você terá a oportunidade de estudar mais detalhadamente os mecanismos de remodelamento da cromatina na próxima aula (Aula 25).

### Regulação da Transcrição por proteínas reguladoras

Como você viu nas Aulas 20 e 21, o processo de transcrição em eucariotos é realizado por três RNA polimerases distintas (I, II e III). Dentre elas, a RNA polimerase II é a responsável pela produção dos RNA mensageiros que, por sua vez, serão traduzidos para a produção

de proteínas. Na Aula 21, foi abordado que a RNA polimerase II se acopla às regiões promotoras dos genes (caixa TATA e caixa CAAT, **Figura 21.1**) com o auxílio de vários fatores de transcrição. A partir de então, tem início a fase de alongação da transcrição.

Nesta aula, já vimos que em cada tecido ou condição do organismo eucarioto um conjunto específico de genes será transcrito, enquanto os demais genes permanecerão inativos. Se as seqüências conservadas, “Caixa TATA” e “Caixa CAAT”, estão presentes nos promotores de, praticamente, todos os genes, como o complexo da transcrição define os genes que devem ser transcritos num dado tecido?

A resposta para tal questão se baseia no fato de que o complexo da RNA polimerase II (incluindo os fatores de transcrição) não é capaz de reconhecer e se acoplar ao promotor de um dado gene sem a ajuda de outros componentes acessórios. Tais componentes acessórios são as proteínas reguladoras.

Elas são moléculas capazes de reconhecer seqüências específicas de um dado promotor e se ligar a ele. A partir de então, tais proteínas realizam interações com o complexo da RNA polimerase, viabilizando o acoplamento dos “fatores de transcrição” às seqüências conservadas previamente discutidas (“caixa TATA” e “caixa CAAT”). Assim, apenas os genes cujos promotores estão associados a proteínas reguladoras são reconhecidos e transcritos.

As seqüências reconhecidas por esse tipo de proteína são específicas para cada promotor, e são chamadas “sítios de ligação”. Cada promotor contém sítios específicos para a ligação de proteínas reguladoras diferentes. Conseqüentemente, podemos concluir que existem muitas proteínas reguladoras diferentes em um organismo.

O processo de regulação se verifica pelo fato de que um dado gene só será transcrito nas células em que a proteína reguladora, capaz de reconhecer seu promotor, estiver presente. Nos demais tecidos, o gene permanecerá desligado.

O mecanismo de controle da transcrição via proteínas reguladoras é o principal sistema de regulação da expressão gênica em eucariotos. Na Aula 25, esse assunto será discutido mais detalhadamente, e abordaremos as diferentes proteínas reguladoras, seus mecanismos de interação com os promotores e com o complexo da transcrição.

## REGULAÇÃO PÓS-TRANSCRICIONAL

Conforme você estudou na Aula 21, o processo da transcrição resulta na produção de um transcrito primário denominado pré-RNAm. Em seguida, tal molécula sofre várias modificações para formar o RNAm que, após ser exportado para o citoplasma, será interpretado para a tradução. Estes passos são essenciais para que o gene se expresse e, portanto, qualquer mecanismo que interfira nesse processo, inibindo ou induzindo, estará regulando a expressão do gene. Quando a expressão de um dado gene é regulada através de mecanismos que interfiram em etapas posteriores à transcrição, o processo é denominado regulação pós-transcricional.

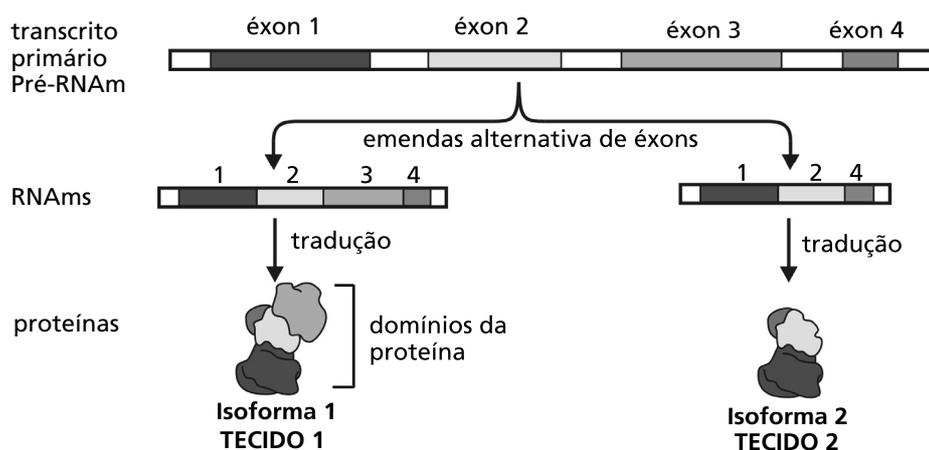
Os mecanismos de regulação pós-transcricional ainda são pouco esclarecidos. Porém, o grande número de pesquisas desenvolvidas sobre este assunto, nos últimos anos, tem revelado a existência de diferentes mecanismos de regulação pós-transcricional. Dentre eles, podemos destacar três mecanismos principais: regulação do processo de edição do RNAm, regulação do sistema de “retirada dos íntrons e emenda dos éxons” (emenda alternativa dos éxons) e regulação da estabilidade do RNAm. Cada um desses mecanismos de regulação pós-transcricional será discutido individualmente a seguir.

### Regulação do processo de “emenda alternativa de éxons”

A retirada de íntrons e emenda dos éxons compõem um passo crucial do processamento do RNA em eucariotos, conforme você aprendeu na Aula 22. Tal processo permite a eliminação das seqüências correspondentes aos íntrons do transcrito primário (pré-RNAm) e a subsequente reunião (emenda) dos éxons.

Muitos genes de eucariotos apresentam um elevado número de íntrons. Teoricamente, todos eles devem ser retirados durante o processamento do pré-RNAm. Assim, no decorrer do processo, os diversos éxons deverão ser emendados. Imagine o que aconteceria se, em uma dada condição, um dos éxons fosse eliminado juntamente com os íntrons. Certamente, o RNAm formado seria menor que o convencional e, por conseqüência, a proteína formada não conteria a região codificada por tal éxon. Na realidade, as células utilizam tal processo como uma importante forma de regulação da expressão gênica. Tal mecanismo

é denominado emenda alternativa de éxons e consiste em selecionar os éxons que serão emendados para compor o RNAm final. Dessa forma, um mesmo pré-RNAm pode dar origem a duas ou mais formas alternativas de RNAm, em função dos éxons utilizados na montagem de cada uma. Conseqüentemente, cada um desses RNAm variantes dará origem a uma diferente forma da proteína, na qual domínios estarão presentes, ou ausentes, em função dos éxons que compõem os respectivos RNAs (Figura 24.5).



**Figura 24.5:** Esquema ilustrando o mecanismo de emenda alternativa de éxons. A partir de um mesmo transcrito primário (pré-RNAm) dois diferentes RNAs podem ser formados. Como resultado, o processo de tradução desses RNAs gerará isoformas distintas da proteínas codificada, onde a isoforma 2 não contém a região codificada pelo éxon 3.

Alguns mecanismos de emenda alternativa parecem ser constitutivos, com os RNAs variantes coexistindo em proporções constantes na mesma célula. Outros mecanismos são regulados em resposta a estímulos ambientais ou de desenvolvimento. O controle dos éxons que serão utilizados na montagem do RNAm em um dado tecido, ou condição, é realizado através de proteínas que interagem com os complexos **EMENDOSSOMAS** (lembra-se da Aula 22?). Portanto, se em um determinado tecido ou condição tais proteínas estiverem presentes, um dado conjunto de éxons pode ser selecionado. Porém, se em outro tecido ou condição tais proteínas estiverem ausentes, ou forem substituídas por outras, um novo conjunto de éxons pode ser emendado, dando origem à forma alternativa do mRNA (Figura 24.5).

Podemos notar que, através dos mecanismos de “emenda alternativa de éxons”, um mesmo gene pode dar origem a diferentes proteínas. Por exemplo, consideremos uma proteína na qual uma determinada região seja responsável por sua atividade enzimática.

#### EMENDOSSOMAS

São os complexos compostos por diversas proteínas e pequenos RNAs nucleares responsáveis pelo processo de retirada de íntrons e emenda dos éxons. Em alguns livros, tais estruturas são denominadas spliceossomos.

Se o éxon que codifica para tal região for eliminado durante os processos de emenda, a proteína final não conterà tal região e, conseqüentemente, a atividade enzimática será abolida. Similarmente, determinadas proteínas são endereçadas para organelas específicas após a sua tradução. Se o éxon que codifica para o domínio de endereçamento celular da proteína estiver ausente, a proteína passará a ocupar outro compartimento celular ou ficar no citosol.

Portanto, a emenda alternativa de pré-RNAs é um poderoso e versátil mecanismo regulatório, que pode efetuar o controle da expressão gênica e a diversificação funcional de proteínas. Dessa forma, a expressão de um gene pode resultar em diferentes isoformas da proteína, com distintas estruturas e funções.

### **Regulação do processo de edição do RNAm**

No último tópico da Aula 21, você viu que, durante o processamento, determinados transcritos podem receber a inserção, deleção ou modificação de nucleotídeos em sua seqüência. Tal processo é denominado edição do RNA. Também naquele capítulo, você estudou o exemplo do RNAm para apolipoproteína-B que, no fígado, codifica uma proteína que contém 4.563 aminoácidos. Em contrapartida, no intestino, a mesma proteína contém somente 2.153 aminoácidos, como resultado de um processo de edição do RNAm que resulta na tradução de uma proteína mais curta. Este é um excelente exemplo de regulação pós-transcricional realizada via edição de RNAm, pois permite que um mesmo gene codifique para proteínas diferentes, em função do tecido onde se expressa. A regulação do processo, portanto, dependerá da presença ou ausência das enzimas responsáveis pela edição daquele transcrito, em cada tecido.

Embora o exemplo do RNAm para apolipoproteína-B seja bastante ilustrativo, é importante observar que a edição do RNA tem sido verificada para muitos outros RNAs, além de RNAs e RNAs, de diversas espécies.

### **Regulação da estabilidade do RNAm**

O processo de transcrição de um dado gene resulta no acúmulo do RNAm correspondente no citoplasma da célula. Assim, essas moléculas estarão disponíveis para a tradução. Entretanto, RNAs possuem “vida

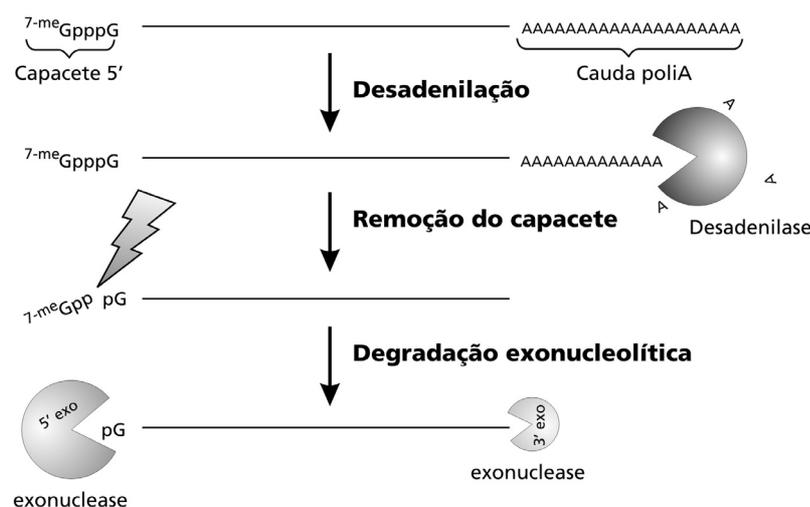
útil”, ou seja, após certo tempo, tais moléculas precisam ser eliminadas. Isto é bastante aceitável, pois se os RNAs fossem disponíveis no citoplasma por tempo indefinido, sua tradução continuaria ocorrendo também indefinidamente, mesmo que a transcrição do gene já não mais ocorresse. Isso seria um problema para a célula que necessita de determinadas proteínas apenas em momentos específicos. Adicionalmente, se os RNAs não fossem degradados, tais moléculas se acumulariam no citoplasma e, com a continuação do processo de transcrição dos milhares de genes durante a vida da célula, a quantidade de RNA no citoplasma atingiria níveis altíssimos, inviabilizando o funcionamento celular.

Cada molécula de RNA permanece no citoplasma celular durante um determinado período. Em seguida, é degradada por enzimas. Inicialmente, acreditava-se que todos os RNAs possuíam tempos similares de permanência no citoplasma. Entretanto, diversas pesquisas têm demonstrado que os transcritos produzidos por determinados genes têm um período de duração no citoplasma muito maior que os demais. Este período de duração no citoplasma é também denominado estabilidade do RNA. Diante de tal informação, podemos concluir que, entre dois genes diferentes, que possuam taxas de transcrição semelhantes, aquele cujos RNAs apresentem maior estabilidade, gerará maior acúmulo dos transcritos no citoplasma celular. A tradução de tais transcritos resultará em maior produção da proteína codificada. Portanto, o controle da estabilidade do RNA de um dado gene pode ser um importante mecanismo de regulação de sua expressão.

De fato, genes distintos apresentam diferentes estabilidades para seus RNAs. Adicionalmente, os transcritos produzidos por um mesmo gene podem apresentar diferentes estabilidades, em função do tecido onde são produzidos. Como a célula faz para que RNAs diferentes apresentem estabilidades distintas?

Conforme você viu na Aula 22, praticamente todos os RNAs recebem uma estrutura denominada capicete, em sua extremidade 5', e uma cauda poli-A, na extremidade 3'. Acredita-se que ambas estejam envolvidas na estabilidade do RNA. Atualmente, várias pesquisas têm confirmado tal hipótese. Ao chegarem no citoplasma, as moléculas de RNA passam a estar expostas à ação das enzimas de degradação. O principal processo de degradação do RNA, descrito em leveduras e eucariotos superiores, é iniciado pela remoção da cauda poli-A e é

denominado desadenilação. As enzimas responsáveis pelos processos de desadenilação são chamadas desadenilases. Em seguida ao processo de desadenilação, ocorre a remoção do “capacete”, presente na extremidade 5’ do RNAm, realizada por um conjunto de proteínas e fatores. Após a retirada das estruturas de proteção (capacete e cauda poli-A), o transcrito é rapidamente degradado por enzimas denominadas exonucleases (Figura 24.6).



**Figura 24.6:** Esquema ilustrando o mecanismo de degradação de RNAm dependente de desadenilação. Quando o processamento do RNAm é concluído, ele contém uma estrutura capacete (5') e uma cauda Poli-A (3') que o protegem de degradação exonucleolítica. O primeiro passo no processo de degradação é a redução da cauda Poli-A realizada por enzimas denominadas desadenilases. Após a degradação da cauda Poli-A, a estrutura capacete 5' é rapidamente removida e o resto do RNAm é rapidamente atacado por exonuclease.

A regulação dos processos de degradação do RNAm tem sido amplamente estudada e, embora não esteja completamente esclarecida, já permite a elaboração de modelos que expliquem os processos. De maneira geral, os modelos sugerem a existência de proteínas capazes de se associar às extremidades “capacete” e “cauda Poli-A” dos RNAm, inibindo ou promovendo a ação das enzimas de degradação. Existem diversas classes de tais proteínas, capazes de se associar preferencialmente a determinados RNAm. Assim, se as proteínas associadas aos transcritos de um dado gene forem estabilizadoras, o processo de degradação será inibido, resultando em maior estabilidade do mRNA. Entretanto, se as proteínas associadas forem desestabilizadoras, a degradação será induzida, resultando em menor “vida-média” daqueles transcritos. Tal modelo explica como os transcritos de um dado gene apresentam estabilidade “média” maior, ou menor, que os demais.

## REGULAÇÃO DO PROCESSO DE TRADUÇÃO

A partir do momento em que um RNAm se encontra no citoplasma, seu processo de tradução também pode ser regulado. Os vários mecanismos de regulação traducional não são completamente elucidados, mas existem muitos exemplos documentados (particularmente durante o desenvolvimento embrionário) de moléculas de RNAm que são rotineiramente encontradas no citoplasma, mas são traduzidas apenas sob determinadas circunstâncias. Um bom exemplo de tal processo, pode ser verificado em espécies animais nas quais grande quantidade de RNAm é estocada em óvulos. Porém, o processo de tradução só é desencadeado a partir da fertilização.

Dentre os mecanismos de controle de tradução, a regulação pela presença de regiões codificadoras adicionais tem sido bem caracterizada. O processo de tradução de um polipeptídeo, a partir de um RNAm, tem início no “códon de iniciação da tradução”, localizado na região próxima à extremidade 5’ (capacete). Isso ocorre porque a “maquinaria” responsável pela tradução percorre a fita de RNAm a partir da extremidade “capacete” (5’) em direção à extremidade 3’, em busca do “códon de iniciação da tradução”. Entretanto, certas classes de RNAm contêm regiões codificadoras de pequenos peptídeos, antes da região que codifica para o polipeptídeo principal. Em vários casos, tem sido estabelecido que tais regiões afetam a expressão gênica no âmbito traducional, prejudicando a tradução do polipeptídeo principal.

## REGULAÇÃO DOS PROCESSOS DE MODIFICAÇÃO PÓS-TRADUCIONAL

Após a tradução, muitos polipeptídeos precisam passar por um processamento que lhes confira a estrutura necessária ao desempenho de sua função biológica. Tal processamento consiste em alterações promovidas nas propriedades da proteína, através da clivagem proteolítica ou da adição de grupos modificadores (radicais fosfato, açúcares etc.) a um ou mais aminoácidos. Tal conjunto de alterações recebe o nome de modificações pós-traducionais.

Em muitas proteínas de eucariotos, a presença ou ausência de tais modificações pode definir o estado de atividade da proteína, sua localização celular, taxa de degradação e a interação com outras proteínas.

Assim, os mecanismos de modificação pós-traducional desempenham um importante papel biológico. A regulação desses processos é realizada através da atividade de enzimas modificadoras que atuam sobre substratos específicos. Como resultado, o controle da ocorrência de tais modificações na proteína codificada por um dado gene pode modular sua atividade e, conseqüentemente, representar uma forma de regulação da expressão gênica.

### **CONTROLE DA ESTABILIDADE DA PROTEÍNA**

O tempo decorrido entre o final do processo de tradução e o momento da degradação de um polipeptídeo define a “longevidade” desta molécula no citoplasma. Portanto, a indução ou a repressão do processo de degradação dos polipeptídeos codificados por um dado gene representam uma forma de controle da expressão gênica, afetando a estabilidade de tais moléculas. Cabe ressaltar que o processo é seletivo, pois é observada grande heterogeneidade entre as taxas de degradação das diferentes proteínas de uma célula.

Muitos aspectos do desenvolvimento de eucariotos dependem de mecanismos regulados de degradação de proteínas. Dentre eles, um importante papel é desempenhado pelas “vias de degradação ubiquitina-proteassomo”. Nesse mecanismo, as proteínas a serem degradadas são marcadas através da ligação covalente da ubiquitina, realizada por enzimas específicas. Em seguida, os complexos ubiquitina/proteína são submetidos à degradação pelo proteassomo 26S, uma estrutura composta por diversas enzimas que possuem função proteolítica (capacidade de degradar proteínas).

### **REGULAÇÃO DOS MECANISMOS DE ENDEREÇAMENTO DAS PROTEÍNAS**

A célula eucariótica é dividida em vários compartimentos. Cada organela, por exemplo, contém proteínas específicas, cuja função é característica daquele compartimento celular. Assim, fica claro que muitas proteínas, após sua tradução, precisam ser corretamente endereçadas para o compartimento celular onde desempenharão seu papel biológico. Você estudará, em aulas futuras, os mecanismos responsáveis pelo endereçamento de proteínas. Entretanto, com as informações que você já detém, é possível perceber que o controle dos mecanismos responsáveis pelo endereçamento de uma proteína representa uma forma de regulação da expressão gênica.



Nesta aula, você teve a oportunidade de estudar os diversos mecanismos que podem regular a expressão de um gene. Entretanto, é importante ressaltar que nem todos os mecanismos são utilizados pela célula para controlar a expressão de um dado gene. Em outras palavras, cada um dos milhares de genes de um organismo pode ser regulado de uma maneira distinta. Embora a grande maioria dos genes sofra regulação transcricional, nem todos são regulados pós-transcricionalmente ou pós-traducionalmente. Similarmente, a opção de qual dos mecanismos de regulação será ativado também é específica para cada gene. Por exemplo, apenas alguns genes sofrem regulação por emenda alternativa de éxons, enquanto outros podem ser regulados quanto à estabilidade do RNAm, ou quanto ao endereçamento do proteína etc.

## RESUMO

Nesta aula, você teve a oportunidade de aprender que a regulação da expressão gênica em eucariotos apresenta mecanismos mais sofisticados que o observado para células procaríotas. Vimos que o fato de diferentes células de um organismo multicelular apresentarem diferentes fenótipos, apesar de conterem os mesmos genes, pode ser explicado pela expressão diferencial de genes. Vimos também que determinados genes são expressos constitutivamente enquanto outros sofrem regulação temporal e/ou tecido-específica. Em seguida, foram apresentados os diferentes mecanismos que podem regular a expressão de um gene: regulação transcricional (por metilação e por proteínas reguladoras), regulação pós-transcricional (processamento do transcrito, transporte do RNAm para o citoplasma e a estabilidade do RNAm), regulação traducional e regulação pós-traducional (controle do endereçamento da proteína e do seu endereçamento).

## EXERCÍCIOS

1. Se todas as células somáticas de um indivíduo eucarioto multicelular possuem os mesmos genes, como podemos explicar as diferenças fenotípicas observadas entre os diversos tipos celulares?
2. O que são “cascatas de regulação da expressão gênica”?
3. Como os genes podem ser classificados quanto à sua regulação?
4. Quais as etapas em que a expressão de um gene de eucarioto pode ser regulada?

5. Qual o mais importante mecanismo de regulação da expressão gênica em eucariotos?
6. Quais os mecanismos de regulação pós-transcricional?
7. Quais os mecanismos de regulação pós-traducional?

### **AUTO-AVALIAÇÃO**

Se você compreendeu o conteúdo desta aula não deve ter encontrado dificuldades para fazer os exercícios. Lembre-se de que a visão geral acerca da importância da regulação gênica e os níveis em que a regulação pode ocorrer é mais importante que os detalhes de funcionamento de cada mecanismo específico. Portanto, certifique-se de que você assimilou os conceitos gerais de regulação em eucariotos e os níveis nos quais a regulação pode ocorrer, pois tais informações serão importantes para a compreensão da próxima aula.

## Regulação da transcrição em eucariotos

AULA

# 25

## objetivos

Ao final desta aula, você deverá ser capaz de:

- Entender os mecanismos envolvidos na regulação transcricional da expressão gênica.
- Estudar mecanismos envolvidos no controle dos processos de remodelamento da cromatina.
- Estudar os mecanismos envolvidos na ação das proteínas reguladoras para o controle da expressão gênica.

### **Pré-requisitos**

Regulação da expressão gênica em eucariotos, Aula 24.

"Fluxo da informação gênica – transcrição" e "Processamento do RNA",  
apresentados nas Aulas 20, 21 e 22.

Controle da expressão gênica em procariotos, Aula 23.

Estrutura dos cromossomos de eucariotos, Aula 8.

## INTRODUÇÃO

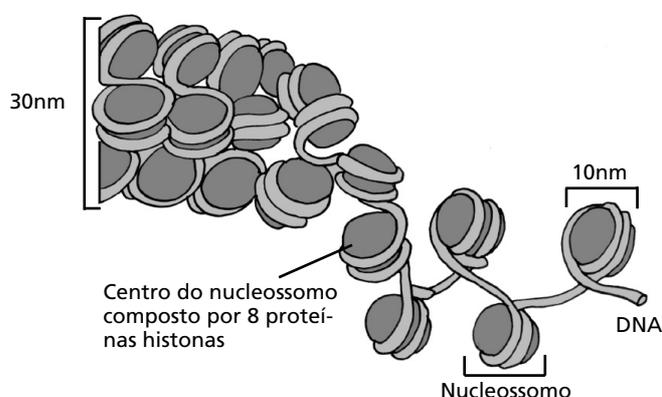
Na Aula 24, você teve a oportunidade de estudar os mecanismos que regulam a expressão gênica em eucariotos. Naquela aula, foram abordadas as diversas etapas da expressão gênica de eucariotos passíveis de regulação. Dentre tais etapas, a principal forma de regulação é o controle da transcrição, que pode ser realizado através de dois mecanismos distintos: remodelamento da cromatina e ligação de proteínas reguladoras.

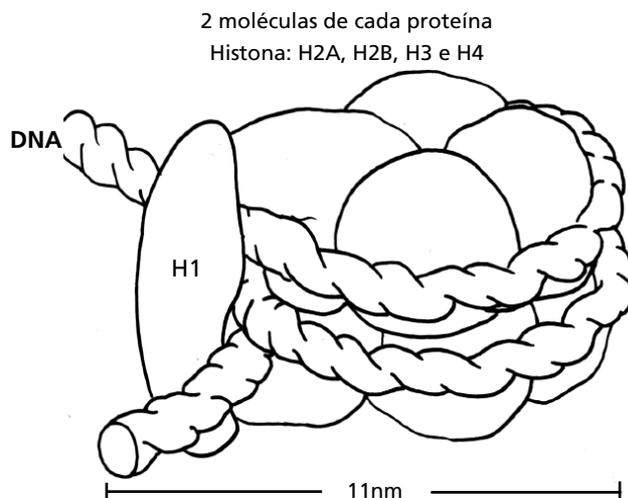
Nesta aula, teremos a oportunidade de abordar, com maior detalhamento, os processos que regulam a transcrição em eucariotos, estudando os mecanismos que controlam a disponibilização de regiões genômicas para a transcrição (remodelamento da cromatina) e as proteínas que regulam a ativação de cada gene (proteínas reguladoras).

## CONTROLE DA TRANSCRIÇÃO POR REMODELAMENTO DA CROMATINA

Você já aprendeu na Aula 8 (Módulo 1) que os cromossomos metafásicos encontram-se muito mais condensados que os cromossomos encontrados na fase da intérfase. Viu também que a cromatina pode ser encontrada em diferentes níveis de condensação, segundo um mecanismo geral de empacotamento. As unidades básicas desse mecanismo são os nucleossomos, compostos por proteínas histonas. Estes nucleossomos são dispostos a cada 160 nucleotídeos ao longo da molécula de DNA, envolvendo 145-147 pares de bases em torno de sua estrutura. Como resultado, a cromatina assume uma disposição em forma de colar de contas (Figura 25.1). Cada nucleossomo é composto por oito moléculas de proteínas histonas (Figura 25.2).

**Figura 25.1:** Esquema ilustrativo de uma região da cromatina, apresentado os nucleossomos dispostos seqüencialmente ao longo da molécula de DNA, gerando uma estrutura em forma de "colar de contas".

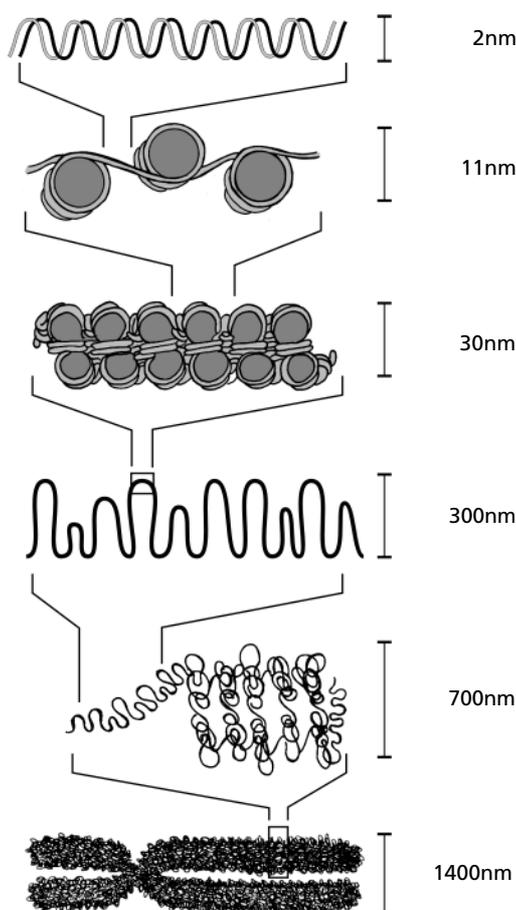




**Figura 25.2:** Esquema ilustrativo de um nucleossomo. O centro do nucleossomo é composto por oito moléculas de proteínas histonas. Externamente, encontra-se a histona H1.

As proteínas histonas dos nucleossomos, juntamente com proteínas cromossomais não histonas, são capazes de realizar interações que resultam em níveis mais intensos de empacotamento, conforme resumido na **Figura 25.3**. Dessa forma, a célula utiliza as histonas para comprimir o DNA em uma estrutura compacta denominada cromatina, cuja função é organizar e empacotar o genoma dentro do núcleo celular.

Na Aula 24, discutimos que as diversas regiões do genoma apresentam níveis de compactação distintos. Adicionalmente, determinadas regiões se apresentam compactadas em determinados tecidos, ou fase do desenvolvimento, e descompactadas em outros. Paralelamente, determinadas regiões permanecem sempre compactadas em qualquer tecido ou condição.



**Figura 25.3:** Esquema ilustrativo dos diferentes níveis de compactação da cromatina. Na ausência de compactação, a espessura da molécula linear de DNA é de aproximadamente 2  $\eta$ M. No cromossomo metafásico, entretanto, a espessura da cromatina pode chegar a 1400  $\eta$ M, o que representa uma condensação de 700 vezes em relação ao DNA linear.

Durante a divisão celular, entretanto, todos os cromossomos se apresentam altamente condensados. O nível de condensação de uma região cromossômica está inversamente relacionado com o nível de transcrição de seus genes. Em outras palavras, regiões compactadas do genoma são, normalmente, transcricionalmente inativas.

Com base nas informações apresentadas, você pode concluir que o controle do nível de condensação da cromatina representa uma forma de regulação da transcrição, pois define as regiões disponíveis para serem transcritas. Sendo assim, como a célula regula os níveis de compactação de tais regiões?

A resposta para a questão anterior foi parcialmente abordada na Aula 24. Vimos que os mecanismos associados a tal controle estão relacionados à presença ou ausência de radicais na estrutura do DNA ou das proteínas histonas. A ocorrência destas modificações permite que regiões genômicas idênticas apresentem diferentes estados de atividade, em função dos tecidos ou fase de desenvolvimento em que a célula se encontre. Portanto, a atividade transcricional de um gene pode ser determinada pela presença ou ausência de tais marcas, ou radicais, na região do genoma que o contém.

Um interessante exemplo do papel do remodelamento da cromatina na regulação da expressão gênica é observado nas fases iniciais do desenvolvimento embrionário. Imediatamente após a fecundação, os cromossomos do zigoto recém-formado encontram-se altamente condensados. Nessa condição, portanto, os níveis de expressão gênica são muito baixos. A partir daí, tem início um processo de descompactação seqüencial do genoma, realizado por enzimas que promovem a adição ou remoção de radicais (metil, acetil e fosfato). Conseqüentemente, passa a ocorrer a ativação transcricional de tais regiões. Várias pesquisas sugerem que o processo de descompactação do DNA segue padrões tecido-específicos, através de enzimas específicas para cada tecido. Como resultado, algumas regiões do genoma estarão disponíveis para a transcrição apenas em determinados tecidos, nos quais se encontram descompactadas.

## Metilação do DNA

No caso do DNA, o principal mecanismo de adição de radicais é denominado metilação. Este processo consiste na adição de radicais metil a determinadas bases nitrogenadas citosinas, ao longo da molécula de DNA. A metilação do DNA, conforme discutido na Aula 24, tem sido intensamente estudada nas últimas décadas e, atualmente, sabe-se que regiões altamente metiladas (regiões hipermetiladas) não são passíveis de transcrição. A regulação dos níveis de metilação do DNA

ao longo do genoma é controlada pela ação coordenada de metilases (DNA metiltransferases) (responsáveis pela adição dos radicais metil) e as desmetilases (responsáveis pela remoção dos mesmos). O processo de metilação ocorre imediatamente após a síntese do DNA e atinge considerável parcela das citosinas do genoma. Estima-se que entre 5 e 7% de todas as citosinas de eucariotos sejam metiladas.

Você já estudou, na Aula 24, que determinadas regiões do genoma apresentam diferentes níveis de metilação do DNA, em função do tecido ou fase do desenvolvimento em que se encontram. Tais diferenças decorrem da ação de desmetilases tecido-específicas sobre o genoma que, no início do desenvolvimento embrionário, encontra-se altamente metilado. A ação diferencial de desmetilases tecido-específicas ao longo do desenvolvimento do organismo, portanto, determina os níveis de metilação presentes nas células do organismo adulto.

## METILAÇÃO, ACETILAÇÃO E FOSFORILAÇÃO DAS HISTONAS

Diferentemente dos conhecimentos acerca da metilação do DNA, os processos de **MODIFICAÇÕES EPIGENÉTICAS** que atuam sobre as proteínas histonas e os mecanismos por eles regulados são pouco esclarecidos atualmente. Porém, pesquisas recentes têm revelado que a presença de tais modificações nas histonas são cruciais à determinação do nível de compactação da cromatina. Os três principais mecanismos de modificações epigenéticas em histonas são metilação, acetilação e fosforilação.

As proteínas histonas centrais do nucleossomo (duas cópias de cada uma das proteínas H2A, H2B, H3 e H4) possuem longas caudas amino-terminais que se estendem para fora da estrutura, conforme esquematizado na **Figura 25.4**. Tais caudas são o principal alvo para a gama de modificações químicas já citadas, que resultam na alteração das cadeias laterais de aminoácidos específicos. Essas modificações nas proteínas histonas afetam o acesso de proteínas reguladoras e complexos protéicos à cromatina e, dessa forma, influenciam a expressão gênica. O papel de tais modificações nos níveis de compactação da cromatina está associado ao fato de que a presença dos radicais modificadores afeta a capacidade de interação entre nucleossomos adjacentes, conforme ilustrado na **Figura 25.5**. Quanto mais intensa for tal associação entre os nucleossomos, maior a tendência à compactação da cromatina. Na **Figura 25.6** são apresentadas as principais modificações caracterizadas em proteínas histonas: acetilação, fosforilação e metilação.

### MODIFICAÇÕES EPIGENÉTICAS

São modificações ocorridas após o processo de criação de qualquer material (gênese). No caso de DNA e proteínas, este termo é utilizado para definir as alterações sofridas pelas moléculas após a sua produção que, no caso do DNA, ocorre durante a replicação, e no caso das proteínas, ocorre através da tradução. Assim, mecanismos que adicionem, ou retirem, radicais a estas moléculas, resultam em modificações epigenéticas.

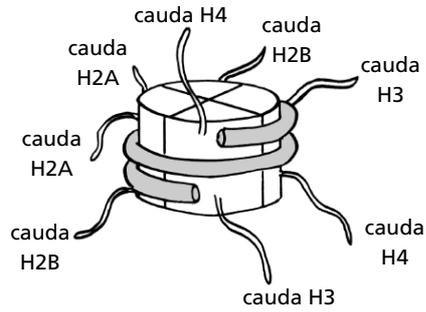


Figura 25.4: Esquema ilustrando a presença de caudas amino-terminais das proteínas histonas que se projetam para fora dos nucleossomos.

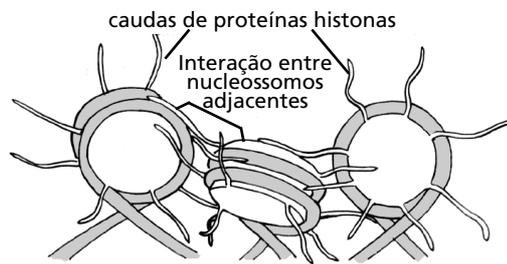
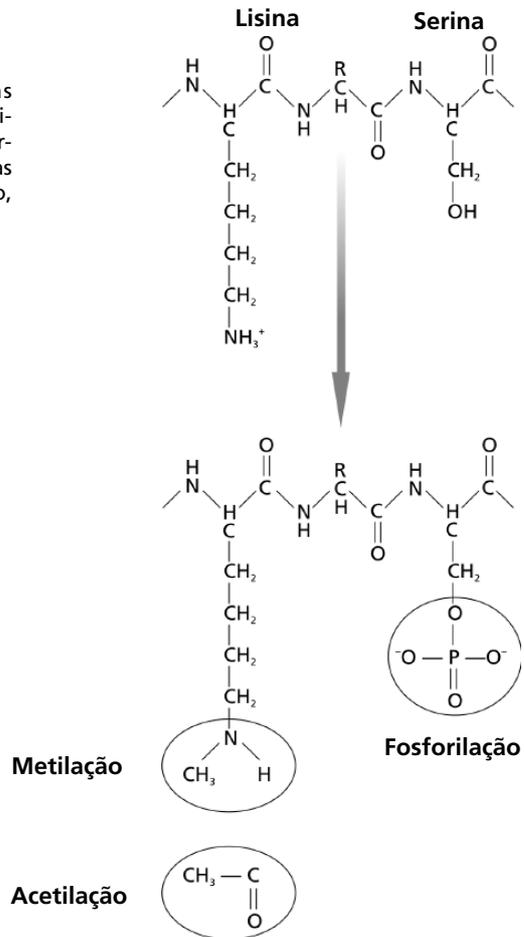


Figura 25.5: Esquema ilustrando a associação entre nucleossomos adjacentes, provida pela interação entre as caudas das proteínas histonas.

Figura 25.6: Esquema das principais modificações verificadas nas caudas amino-terminais das proteínas histonas de nucleossomos: acetilação, fosforilação e metilação.



Dentre os processos de modificação de histonas, os mecanismos de acetilação e desacetilação são os mais bem caracterizados. A adição ou retirada de radicais acetil em aminoácidos lisina conservados na cauda de proteínas histonas tem sido amplamente correlacionada com a atividade transcricional. Histonas acetiladas são comumente associadas à cromatina transcricionalmente ativa, e histonas desacetiladas à cromatina inativa para a transcrição. Os processos de acetilação e desacetilação são desempenhados por enzimas denominadas acetiltransferases e desacetilases, respectivamente.

Acredita-se que a neutralização da carga básica da cauda das histonas, promovida pela acetilação, resulte em três alterações funcionais: a) redução da afinidade da histona pelo DNA; b) alteração nas interações entre histonas de nucleossomos adjacentes c) alteração nas interações entre as histonas e proteínas reguladoras. Segundo tal hipótese, acredita-se que tais alterações confirmam à cromatina um ambiente permissivo para a transcrição.

É importante ressaltar que, além de seu papel na regulação transcricional, a acetilação de histonas também está envolvida nos processos de replicação, montagem dos nucleossomos, empacotamento dos cromossomos e a interação entre os nucleossomos e proteínas não histônicas.

Em contraste ao considerável volume de informações disponíveis acerca dos processos de acetilação e desacetilação, os conhecimentos sobre os mecanismos e enzimas envolvidos com as demais modificações nas histonas são relativamente escassos. Importantes progressos, entretanto, têm sido obtidos na tentativa de entender o papel dos mecanismos de fosforilação de histonas em processos como a transcrição, reparo do DNA e condensação dos cromossomos. Por exemplo, a fosforilação do aminoácido serina 10 da histona H3 tem demonstrado correlação com a ativação gênica em células de mamíferos. Embora os mecanismos através dos quais a fosforilação das histonas contribui com a ativação transcricional não estejam bem elucidados, a adição de grupos fosfatos, negativamente carregados às caudas das histonas neutralizam sua carga básica e, portanto, acredita-se que reduzam sua afinidade pelo DNA. Adicionalmente, tem sido demonstrado que a fosforilação de aminoácidos promove o incremento da atividade de processos de acetilação subsequentes.

A metilação de histonas foi descrita pela primeira vez em 1964. Porém, apenas 35 anos depois a sua correlação com a regulação da transcrição foi demonstrada. Embora os mecanismos sejam pouco elucidados, algumas pesquisas têm demonstrado que as histonas H3 e H4 são preferencialmente metiladas em aminoácidos lisinas, presentes em locais específicos de cada proteína. O processo é realizado por enzimas metiltransferases específicas.

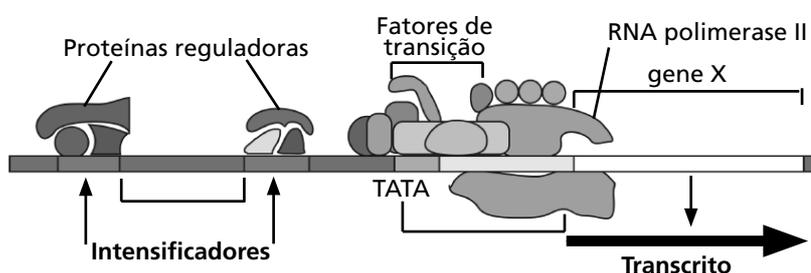
É importante destacar que os processos de modificação das proteínas histonas e do DNA atuam de maneira coordenada para o controle da transcrição, bem como dos níveis de condensação da cromatina. Os conhecimentos acerca do assunto são bastante superficiais, mas a sua relevância para tais controles tem sido reforçada gradualmente por novas descobertas científicas. A presença, ou ausência, de tais modificações (no DNA e nas proteínas histonas) pode afetar a expressão de um dado gene por duas maneiras distintas; a) alteram a compactação do DNA e a sua disponibilidade para transcrição; b) alteram os níveis de interação entre a região do genoma e proteínas reguladoras específicas.

### **Regulação da transcrição por proteínas reguladoras**

Na Aula 24, você teve a oportunidade de aprender que cada gene possui uma região reguladora, também chamada região promotora. Trata-se de uma seqüência localizada na porção 5' adjacente ao gene. Nela, estão presentes as seqüências conservadas “caixa TATA” e “caixa CAAT”, que estão envolvidas no acoplamento do complexo da transcrição. Além dessas seqüências, conservadas nos promotores de praticamente todos os genes, existem outras regiões que influenciam na regulação de cada gene. Tais regiões contêm seqüências específicas para cada promotor, que influenciam o nível de transcrição dos genes. Tais regiões são denominadas “intensificadores”.

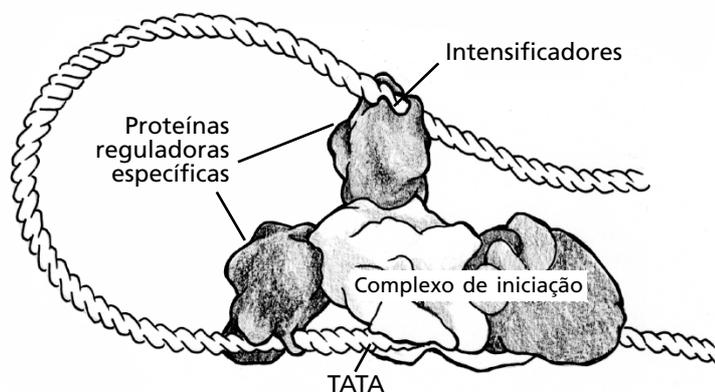
Os intensificadores correspondem a sítios de ligação de proteínas reguladoras. Conforme discutido na Aula 24, proteínas reguladoras são moléculas capazes de reconhecer seqüências específicas de um dado promotor e se ligar a ele. Em seguida, realizam interações com o complexo da RNA polimerase, viabilizando o acoplamento dos “fatores de transcrição” às seqüências conservadas previamente discutidas (“caixa TATA” e “caixa CAAT”).

Uma dada região promotora pode conter vários intensificadores distintos. Portanto, várias proteínas reguladoras diferentes podem controlar a expressão de um gene. A **Figura 25.7** ilustra um exemplo da região reguladora de um gene X. Nela, podemos destacar as regiões do DNA que afetam a expressão do gene (seqüências consenso TATA e as seqüências intensificadoras) e as proteínas que se associam a tais regiões: complexo da transcrição (Fatores de transcrição+ RNA polimerase II) e as proteínas reguladoras que se ligam aos intensificadores.



**Figura 25.7:** Esquema ilustrativo da região reguladora da transcrição de um gene X. Estão representadas as regiões do DNA que afetam a expressão do gene (seqüências consenso TATA e seqüências intensificadoras) e as proteínas que se associam a tais regiões: complexo da transcrição (fatores de transcrição+ RNA polimerase II) e as proteínas reguladoras que se ligam aos intensificadores.

A molécula de DNA é uma estrutura muito delgada. Diversas pesquisas têm demonstrado que a ligação de várias proteínas a uma região reguladora e a interação entre tais proteínas promovem o dobramento da molécula de DNA, viabilizando a associação entre proteínas que se ligam a sítios separados. Na **Figura 25.8** você pode ver um esquema ilustrando tal dobramento.



**Figura 25.8:** Ilustração do dobramento da molécula de DNA decorrente da interação entre proteínas reguladoras, ligadas aos intensificadores, com os elementos componentes do complexo de iniciação, acoplados à região "Caixa TATA".

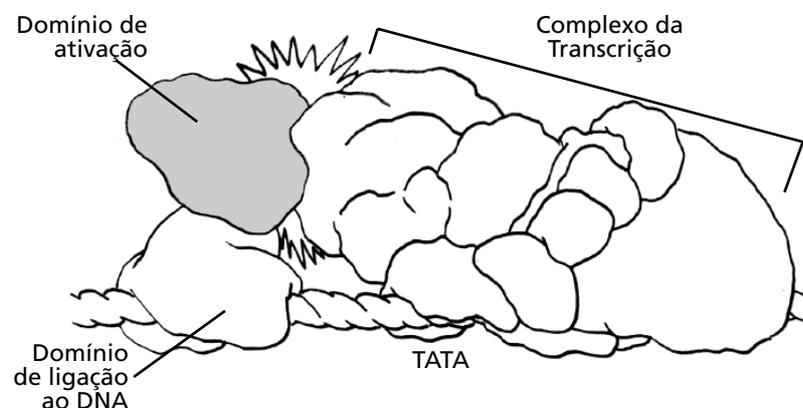
## ATUAÇÃO DAS PROTEÍNAS REGULADORAS

A ligação de proteínas reguladoras pode exercer dois papéis contrastantes sobre a transcrição de um gene. Muitas dessas proteínas ativam o processo de transcrição, pois facilitam o acesso do complexo da transcrição ao promotor. Nesses casos, tais proteínas são denominadas ativadores (ou proteínas ativadoras) da transcrição. Entretanto, determinadas proteínas reguladoras atuam prejudicando a transcrição de gene, atenuando ou inviabilizando o acesso do complexo da RNA polimerase. Tais proteínas são denominadas repressores (ou proteínas repressoras) da transcrição.

Proteínas ativadoras da transcrição são proteínas modulares compostas por dois **DOMÍNIOS FUNCIONAIS** distintos. O primeiro domínio corresponde à região capaz de reconhecer e se ligar a uma seqüência específica de DNA (sítio de ligação), sendo denominado “domínio de ligação ao DNA”. O outro domínio corresponde à região que estabelece interação proteína-proteína com o complexo da transcrição, denominado “domínio de ativação”. A **Figura 25.9** ilustra a ação de tais domínios.

### DOMÍNIO FUNCIONAL

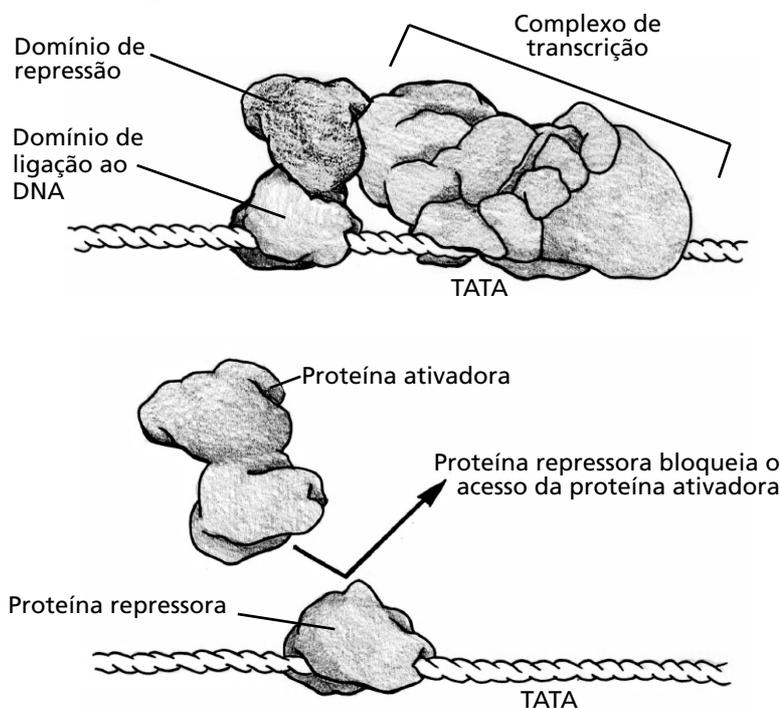
de uma proteína corresponde à região da molécula responsável pela atividade biológica a ela associada. Por exemplo, o domínio funcional de uma enzima é a porção da proteína responsável pelo reconhecimento e modificação do substrato. Determinadas proteínas possuem mais de um domínio funcional localizado, portanto, em regiões diferentes da cadeia polipeptídica.



**Figura 25.9:** Ilustração da ação dos domínios componentes de uma proteína reguladora da transcrição. O domínio de ligação ao DNA é responsável pelo reconhecimento e ligação a seqüências específicas presentes no DNA (intensificadores). O domínio de ativação é responsável pela interação com o complexo da transcrição.

Proteínas repressoras podem desempenhar seu papel através de dois processos diferentes. Alguns repressores atuam de maneira ativa para reprimir a transcrição. A **Figura 25.10.a** apresenta um exemplo de repressão ativa, onde a proteína repressora possui, além do domínio de ligação ao DNA, um “domínio de repressão”. Nesses casos, a proteína interage com fatores do complexo da transcrição para realizar a repressão

transcricional. Em outros casos, a repressão da transcrição pode ser realizada de forma passiva, quando a ligação do repressor ao seu sítio no DNA prejudica, ou bloqueia, o acesso de proteínas ativadoras, conforme ilustrado na Figura 25.10.b.



**Figura 25.10:** Esquema ilustrativo dos mecanismos de atuação de proteínas represoras da transcrição. (a) Repressão ativa da transcrição. Nesse caso, a proteína possui um domínio de repressão que interage com o complexo da transcrição prejudicando o início da transcrição; (b) repressão passiva da transcrição. Nesse caso, a ligação da proteína repressora ao seu sítio no DNA dificulta, ou bloqueia, o acesso da proteína ativadora.

### DOMÍNIOS DE LIGAÇÃO AO DNA PODEM SER CLASSIFICADOS EM DIVERSOS TIPOS ESTRUTURAIS

As regiões reguladoras dos numerosos genes de um organismo eucarioto possuem diferentes intensificadores. Conforme discutimos anteriormente, cada intensificador possui seqüência específica sendo, portanto, reconhecido por uma proteína reguladora também específica. Assim, podemos concluir que existe um grande número de proteínas reguladoras diferentes, capazes de reconhecer cada um dos intensificadores.

O reconhecimento e ligação de uma proteína ao seu sítio de ligação no DNA dependem de compatibilidade conformacional entre as estruturas, além da interação entre cargas dos átomos presentes nas

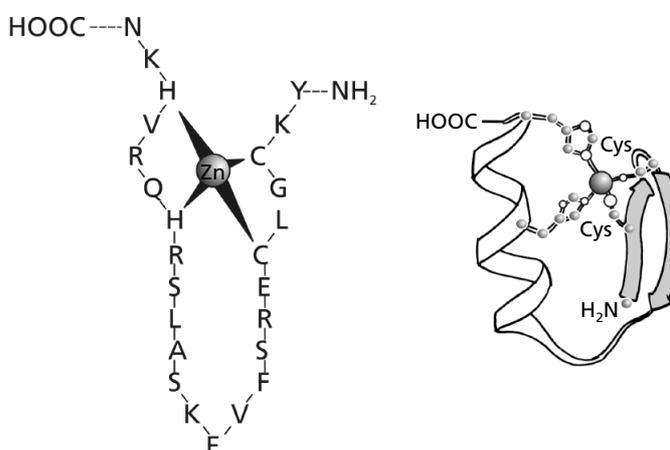
moléculas. O estudo e a comparação das seqüências e das estruturas de milhares de proteínas reguladoras, identificadas em diversas espécies, têm revelado a existência de classes de proteínas reguladoras que apresentam estruturas similares em seu domínio de ligação ao DNA. A seqüência de aminoácidos do domínio de ligação varia entre proteínas de uma mesma classe estrutural, resultando no reconhecimento de sítios de ligação diferentes. A maioria das proteínas reguladoras reconhece seqüências de DNA que variam entre 10 e 18 pares de bases.

Os domínios de ligação ao DNA possuem balanço geral de cargas positivo, o que facilita sua interação com a molécula de DNA, cujo balanço de cargas é negativo. A disposição do domínio de ligação na estrutura da proteína varia entre as diferentes classes estruturais. A seguir, abordaremos alguns exemplos de classes estruturais de proteínas reguladoras da transcrição.

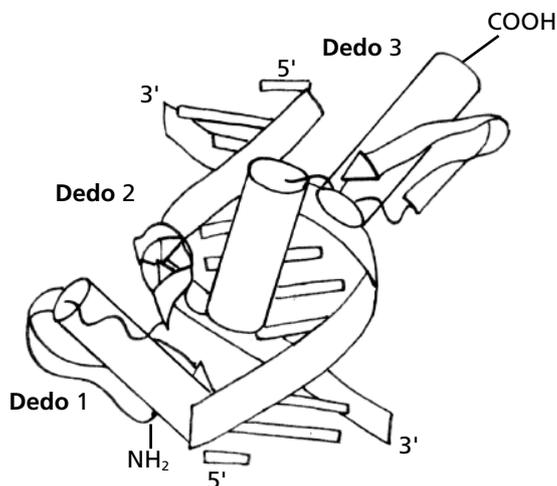
### Proteínas tipo “dedos de zinco”

Esta classe de proteínas recebe tal nome em função da estrutura apresentada na Figura 25.11, na qual um pequeno grupo de aminoácidos conservados liga-se a um íon zinco, formando um domínio relativamente independente dentro da proteína, denominado “dedo de zinco”. A região carboxi-terminal de cada dedo forma uma  $\alpha$ -hélice que se liga ao DNA. A região amino-terminal forma duas estruturas  $\beta$ -pregueadas. Proteínas desta classe são tipicamente compostas por várias estruturas do tipo “dedo de zinco” dispostas em série. O número de dedos de zinco varia entre as diversas proteínas desta classe, podendo ocorrer apenas dois em determinadas proteínas ou até nove em outras.

**Figura 25.11:** Ilustração da estrutura denominada dedo de zinco. (a) Um pequeno grupo de aminoácidos conservados liga-se a um átomo de zinco, formando um domínio dentro da proteína em formato de dedo; (b) os aminoácidos que compõem este domínio formam uma estrutura em  $\alpha$ -hélice e duas estruturas  $\beta$ -pregueadas.



Dados de **CRISTALOGRAFIA DE RAIOS X**, obtidos da interação entre uma proteína contendo três dedos de zinco e o seu sítio no DNA, sugeriram a estrutura ilustrada esquematicamente na **Figura 25.12**.



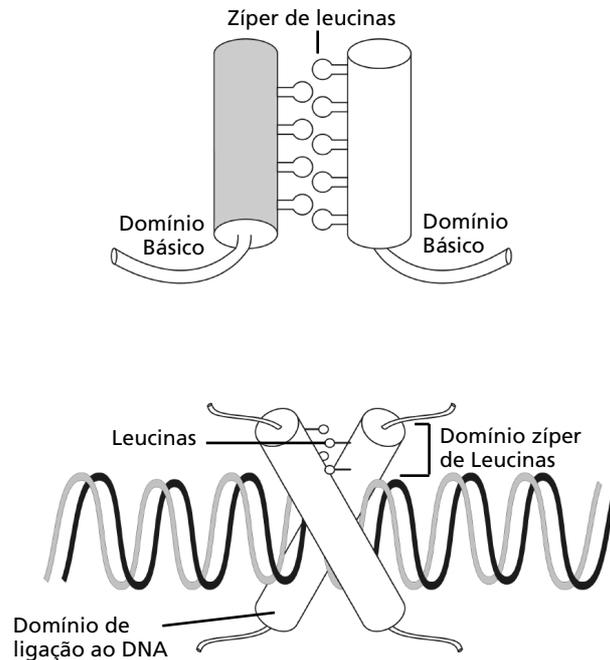
**Figura 25.12:** Esquema ilustrativo da interação entre uma proteína reguladora do tipo "dedo de zinco" e o seu sítio de ligação ao DNA.

### Proteínas tipo "zíper de leucina"

Proteínas reguladoras do tipo "zíper de leucina" participam de um sistema sofisticado de regulação da expressão gênica. O "zíper de leucina" corresponde a uma região da proteína composta rica em aminoácidos leucina, dispostos a cada sete aminoácidos e voltados para o mesmo lado da proteína. O número e a disposição dos aminoácidos leucina confere à proteína a capacidade de formar dímeros com outras proteínas que apresentem regiões com características similares. A **Figura 25.13.a** apresenta um esquema da interação entre duas proteínas distintas que, através de zíperes de leucina, formam dímeros. Entretanto, cabe ressaltar que, freqüentemente, proteínas idênticas se unem para formar dímeros (homodímeros) enquanto a formação de dímeros entre proteínas diferentes (heterodímeros) é menos freqüente.

Proteínas zíper de leucina só reconhecem o seu sítio de ligação no DNA quando se encontram na forma de dímeros. A interação com o DNA ocorre através do domínio positivamente carregado da proteína (domínio básico), conforme ilustrado na **Figura 25.13.b**. Diferenças presentes na seqüência de aminoácidos do domínio básico definem os sítios de ligação das diferentes proteínas desta classe.

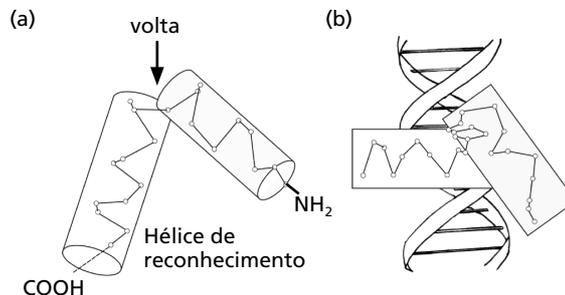
**CRISTALOGRAFIA** de raios X é uma das abordagens utilizadas para o estudo estruturais em moléculas. Para tanto, o material em estudo necessita ser previamente cristalizado para, em seguida, ter o arranjo tridimensional de seus átomos determinado via difração de raios X, emitidos através do cristal.



**Figura 25.13:** Ilustração da estrutura de proteínas reguladoras do tipo “zíper de leucina” e de sua ligação ao DNA. (a) Esquema da interação entre duas moléculas de proteína para formação de um dímero (heterodímero). A interação ocorre através do pareamento propiciado pelos domínios de leucinas; (b) interação entre um dímero de proteínas reguladoras do tipo “zíper de leucina” e o seu sítio de ligação ao DNA.

### Proteínas tipo “hélice-volta-hélice”

Proteínas desta classe se caracterizam por possuírem duas  $\alpha$ -hélices separadas por uma curta seqüência de aminoácidos que promove uma “volta”, ou dobra, na molécula. Na **Figura 25.14.a**, você pode ver um esquema ilustrativo de tal estrutura. Dentre as duas hélices, a que se localiza mais próximo à extremidade carboxi-terminal é a responsável pela ligação ao DNA, conforme ilustrado pela **Figura 25.14.b**.



**Figura 25.14:** Ilustração da estrutura de proteínas reguladoras do tipo “hélice-volta-hélice” e de sua ligação ao DNA. (a) A estrutura é composta por duas  $\alpha$ -hélices separadas por uma curta seqüência de aminoácidos que promove uma “volta”, ou dobra, na molécula; (b) interação entre uma proteína reguladora do tipo “hélice-volta-hélice” e o seu sítio de ligação ao DNA.

## Proteínas reguladoras e conformação da cromatina

Embora tenhamos abordado separadamente os processos de regulação da transcrição por remodelamento da cromatina e o realizado através de proteínas reguladoras, é importante destacar que ambos os processos ocorrem simultaneamente. Muitas vezes a ativação de um dado gene necessita da atuação acoplada de proteínas reguladoras e dos processos que promovem o remodelamento da cromatina.

Determinadas proteínas reguladoras promovem a desorganização dos nucleossomos ou afetam seu posicionamento no DNA. Em outros casos, proteínas reguladoras promovem o acesso de enzimas modificadoras como transacetilases, desmetilases etc., que alteram covalentemente a estrutura das histonas e, conseqüentemente, as estrutura da cromatina. Portanto, a regulação da expressão gênica ocorre de maneira coordenada, utilizando diferentes componentes de mecanismos específicos e de mecanismos gerais de controle da expressão gênica.

### RESUMO

Nesta aula, você teve a oportunidade de estudar mais detalhadamente os mecanismos de regulação da transcrição em eucariotos. Tal regulação ocorre através de dois mecanismos principais: remodelamento da cromatina e atuação de proteínas reguladoras. Os mecanismos de remodelamento da cromatina se baseiam em processos de modificação (adição de radicais) à molécula do DNA ou às proteínas histonas. Dentre as modificações na molécula do DNA, a metilação é o processo mais comum. A enzima responsável pela adição do radical metil às citosinas do DNA é denominada DNA metil transferase.

No caso das histonas, as principais modificações observadas são acetilação, fosforilação e metilação. A regulação da transcrição promovida por proteínas reguladoras é o principal processo de controle da expressão de genes específicos de um organismo. Tais proteínas possuem estrutura modular e apresentam dois

domínios principais: domínio de ligação ao DNA e domínio de interação com o complexo da RNA polimerase. Proteínas reguladoras possuem a capacidade de reconhecer sítios específicos, em determinadas regiões reguladoras, e influenciar a transcrição dos genes. Os diversos domínios de ligação ao DNA podem ser agrupados em classes estruturais distintas, dentre as quais as mais amplamente estudadas são “zíper de leucina”, “dedos de zinco” e “hélice-volta-hélice”.

As proteínas reguladoras atuam coordenadamente com os processos responsáveis pelo remodelamento da cromatina. Dessa forma, as regiões reguladoras de muitos genes devem ser simultaneamente reconhecidas por proteínas reguladoras específicas e por enzimas capazes de promover modificações no DNA e nas proteínas histonas.

## EXERCÍCIOS

1. Como o remodelamento da cromatina pode influenciar a transcrição de um gene presente em uma dada região do genoma?
2. Qual o principal mecanismo de modificação da estrutura do DNA que influencia o nível de compactação da cromatina, e qual a sua correlação com a atividade transcricional?
3. Quais os principais mecanismos de modificação das proteínas histonas e como eles podem influenciar a atividade transcricional?
4. Quais os papéis que proteínas reguladoras podem exercer sobre a transcrição de um gene?
5. Quais são os principais domínios das proteínas reguladoras e qual o papel de cada um deles?

## AUTO-AVALIAÇÃO

Se você compreendeu o conteúdo desta aula, não deve ter encontrado dificuldades para responder aos exercícios. Esta aula enfocou detalhadamente o principal mecanismo de regulação da expressão gênica, o controle da transcrição. Entretanto, lembre-se de que tais informações fazem parte de um conjunto mais amplo de mecanismos, apresentado na Aula 24.

# Fluxo da informação genética – tradução I

AULA

# 26

## objetivos

Nesta aula, você terá oportunidade de:

- Acompanhar o experimento que indicou a relação entre um gene e uma enzima.
- Compreender o conceito de gene e saber aplicá-lo.
- Acompanhar alguns dos experimentos que resultaram na descoberta do código genético.
- Conhecer o código genético e suas peculiaridades, além de saber utilizá-lo.

### Pré-requisitos

Você acompanhará esta aula mais facilmente se não tiver dúvidas em relação às aulas sobre RNA (Aula 5) e aos conceitos de proteína vistos nas aulas de Bioquímica. Também serão necessários os conceitos básicos de mutação (Aula 16).

## INTRODUÇÃO

Chegamos, enfim, ao último módulo da nossa disciplina. Até agora você estudou a importância do DNA como molécula de hereditariedade e sua capacidade de se acomodar em uma célula ou em um compartimento celular, apesar de seu enorme tamanho. Você também teve a oportunidade de estudar o complexo processo que envolve a replicação do DNA, permitindo, assim, que a informação genética seja transmitida da célula-mãe para as células-filhas. Embora existam inúmeros sistemas encarregados pelo reparo dos possíveis danos sofridos pela molécula de DNA, você viu como mutação, recombinação e transposição podem contribuir para a variabilidade existente entre os indivíduos de uma população.

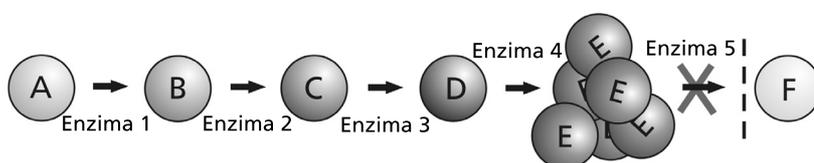
Toda a informação contida no genoma dos mais diferentes organismos é utilizada para produzir moléculas que desempenham as mais diversas funções celulares. São essas moléculas que determinam o fenótipo de uma célula ou indivíduo. Você já começou a estudar o fluxo da informação genética nas aulas do Módulo 3, nas quais foram abordadas as particularidades da transcrição. Nessa primeira aula do Módulo 4, começaremos pela descrição do experimento que confirmou a relação entre o gene e uma enzima, para que você possa, então, começar a estudar as particularidades da tradução, um processo que se caracteriza pelo seu alto grau de complexidade. Tenha certeza de que ao final das próximas quatro aulas esse processo complexo passará a ser visto com outros olhos por você.

## EXPERIMENTO DE BEADLE E TATUM (1940) – DEFINIÇÃO MOLECULAR DE GENE

Até a década de 1940, não se sabia exatamente o que eram os genes e tampouco o que eles “faziam”. A pergunta era: “Qual o significado da herança biológica?” Sabia-se que a herança biológica era uma potencialidade, ou seja, a capacidade de se desenvolver de maneira específica. Outra pergunta então surgia: “Como essas regiões dos cromossomos eram capazes de armazenar tantas e tão complexas informações?” Mais ainda: “Quais os mecanismos envolvidos na manifestação dessas informações?”

Pois saiba que o conceito de gene mudou muito ao longo dos últimos anos. Tradicionalmente, o gene era definido como uma região do cromossomo que determinava ou afetava uma única característica ou fenótipo. Já em 1908, um médico inglês, Archibald Garrod, propôs um novo conceito para as doenças humanas, que são os erros inatos

do metabolismo. Essas doenças se caracterizam por bloqueio de uma via metabólica em um ponto específico. Repare a **Figura 26.1**, na qual está representada a via metabólica de síntese do composto fictício F. O composto precursor A sofre a ação da enzima 1, sendo convertido em B, que pela ação da enzima 2 é convertido em C, e assim por diante, até formar F. Neste exemplo, a enzima 5, que catalisa a conversão de E em F, está ausente ou inativa, de forma que F não é formado e E se acumula no organismo. Sem saber de sua importância para o desenvolvimento da Genética Molecular, Garrod foi o primeiro a postular a existência da relação entre o gene e a química do corpo ao reconhecer que essas doenças, caracterizadas por um erro metabólico, eram herdadas.



**Figura 26.1: Bloqueio de uma via metabólica fictícia.** Neste exemplo, o composto A é precursor do composto F sintetizado em uma via metabólica da qual participam cinco enzimas. A ausência ou a inativação da enzima 5 impede a formação do composto F e determina o acúmulo do composto E.

A primeira definição molecular de gene, entretanto, só foi proposta, em 1940, por George Beadle e Edward Tatum. Esses dois pesquisadores observaram que mutantes selecionados do fungo *Neurospora crassa* (fungo laranja-rosado que se desenvolve normalmente em pão) eram afetados em pontos diferentes da via biossintética de arginina, ou seja, eram deficientes em enzimas específicas dessa via metabólica. Vamos entender o que eles fizeram acompanhando uma série de experimentos apresentados, de forma resumida, nas **Figuras 26.2, 26.3 e 26.4**.

Inicialmente, após experimentos clássicos de Genética envolvendo exposição de **ESPOROS** de *Neurospora crassa* a raios X e diversos cruzamentos, foi feita a seleção de mutantes **AUXOTRÓFICOS** para a arginina, ou seja, mutantes com capacidade de crescimento apenas em **MEIO MÍNIMO** acrescido de arginina. Acompanhe como isso foi feito na **Figura 26.2**.

O tratamento com raios X foi feito no sentido de induzir mutações (**Figura 26.2.a**). Nesta etapa, é de se esperar que diferentes mutações, que podem bloquear uma ou várias vias metabólicas, ocorram nas diferentes células. Os esporos irradiados foram, então, cultivados em meio completo (**Figura 26.2.b**), permitindo, assim, o crescimento de todas as células, mutantes

### ESPOROS

Células reprodutivas encontradas especialmente em fungos. Nestes, os esporos assexuais são células somáticas que se comportam como gametas ou como células iniciais de novos indivíduos haplóides (indivíduos com metade do conjunto de cromossomos). Não só em fungos como também em plantas, existem os esporos sexuais, que são células diplóides.

### AUXOTRÓFICOS

Denomina-se auxotrófico uma linhagem ou um mutante incapaz de sintetizar uma ou mais substâncias específicas. Para seu crescimento é necessário meio completo (fonte de carbono e todos os fatores nutricionais) ou meio mínimo simples (fonte de carbono e sais inorgânicos) acrescido de requerimento nutricional específico.

Os auxotróficos diferem dos prototróficos, que são linhagens capazes de crescer em meio mínimo simples por sintetizarem substâncias específicas, a partir dos componentes deste meio.

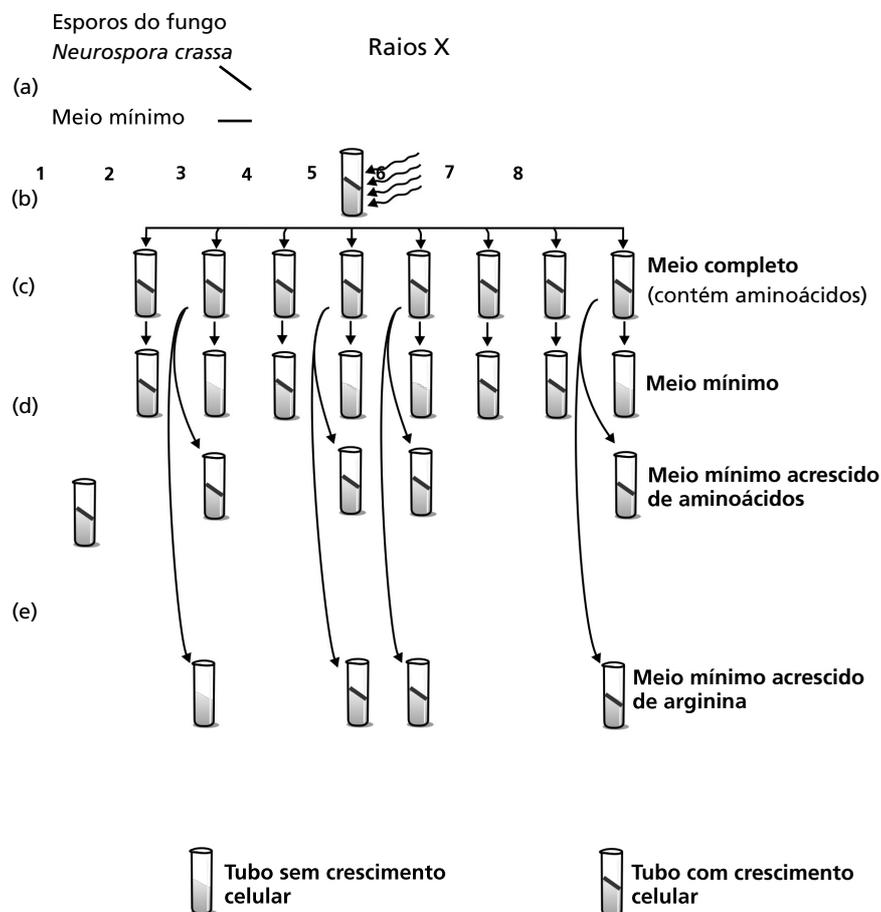
### MEIO MÍNIMO

Meio de cultura contendo compostos simples, fonte de carbono e sais inorgânicos (incluindo fonte de nitrogênio), a partir dos quais os compostos necessários ao crescimento são sintetizados.

ou não, no sentido de se certificar da viabilidade celular. As células viáveis foram transferidas para meio mínimo (Figura 26.2.c), de forma que aquelas que apresentavam bloqueio na via biossintética de qualquer composto essencial ao crescimento foram impedidas de crescer. Em um experimento dessa natureza, o número de células com essa característica corresponde, em média, a 1-2% das células irradiadas, e são ditas mutantes auxotróficos.

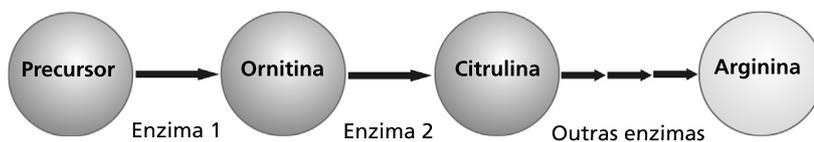
É importante ressaltar que apenas com esse experimento não seria possível saber qual o nutriente requerido para crescimento desses mutantes. Portanto, na etapa seguinte, as células viáveis correspondentes aos tubos sem crescimento (tubos 2, 4, 5 e 8) foram, então, transferidas para meio mínimo, acrescido de aminoácidos (Figura 26.2.d). O crescimento celular foi observado nos quatro tubos, comprovando que estes mutantes eram auxotróficos para aminoácidos em geral. A seleção dos mutantes auxotróficos para arginina foi feita com o cultivo das células viáveis correspondentes aos tubos 2, 4, 5 e 8 em um novo meio de cultura constituído de meio mínimo, acrescido de arginina (Figura 26.2.e). As células selecionadas foram aquelas que apresentaram crescimento neste meio de cultura, ou seja, as correspondentes aos tubos 4, 5 e 8.

**Figura 26.2: Isolamento de mutantes auxotróficos para arginina.** (a) Exposição de esporos de *Neurospora crassa* a raios X. (b) Cultivo dos esporos irradiados em meio completo. (c) Transferência das células crescidas para meio mínimo. (d) Cultivo das células correspondentes aos tubos sem crescimento em meio mínimo, acrescido de aminoácidos. (e) Cultivo das células correspondentes aos tubos sem crescimento na etapa (c) em meio mínimo, acrescido de arginina. No esquema, as células dos tubos 2, 4, 5 e 8 são auxotróficas para aminoácidos em geral. Apenas as células dos tubos 4, 5 e 8 são auxotróficas para arginina, uma vez que cresceram em meio completo, não cresceram em meio mínimo, mas cresceram em meio mínimo acrescido de todos os aminoácidos, e em meio mínimo adicionado de arginina.



O experimento de Beadle e Tatum é um exemplo de aplicação de agentes mutagênicos na área da ciência básica. Nesse caso específico, a possibilidade de se utilizar raios X para a geração de mutantes foi fundamental, acelerando o processo de obtenção de mutantes auxotróficos que contribuíram para a proposição da hipótese *um gene-uma enzima*.

Com base em uma via hipotética da biossíntese de arginina (Figura 26.3), na qual um precursor simples é convertido em ornitina, que, por sua vez, converte-se em citrulina, que, então, é convertida em arginina; na etapa seguinte, foi avaliada a capacidade desses mutantes de crescerem em meio mínimo e em vários meios adicionados de diferentes suplementos, no sentido de se determinar o local de bloqueio metabólico específico.



**Figura 26.3: Via hipotética de biossíntese de arginina.** A síntese de arginina se iniciaria com a conversão de uma molécula precursora em ornitina, que seria convertida em citrulina, que, então, pela ação de uma ou mais enzimas se converteria em arginina.

Para você entender como o experimento foi realizado, leia o texto e acompanhe a Figura 26.4, na qual está representada a seleção de três mutantes de *Neurospora* dependentes de arginina, cada qual com perda de uma enzima particular necessária para a síntese deste aminoácido.

	Meio mínimo	Meio mínimo + Ornitina	Meio mínimo + Citrulina	Meio mínimo + Arginina	RESULTADO	CONCLUSÃO
Mutante "arg1"					Crescimento do mutante arg1 em presença de qualquer um dos precursores	A mutação deve inativar uma enzima que atua no início da via metabólica
Mutante "arg2"					Crescimento do mutante arg2 em presença de citrulina ou arginina, mas não em presença de ornitina	A mutação deve inativar uma enzima que atua entre ornitina e citrulina na via metabólica
Mutante "arg3"					Crescimento do mutante arg3 em presença apenas de arginina	A mutação deve inativar uma enzima que atua na formação de arginina ao final da via metabólica

**Figura 26.4: Localização do bloqueio metabólico associado a cada um dos mutantes.** Três mutantes previamente selecionados foram cultivados em quatro meios de cultura distintos: meio mínimo e meio mínimo acrescido de ornitina, citrulina ou arginina.

Cada mutante selecionado na etapa anterior como auxotrófico para arginina foi cultivado em tubos contendo meio mínimo, acrescido de um composto que participa da biossíntese de arginina. Assim, em um dos tubos contendo meio mínimo foi adicionada ornitina, no segundo foi adicionada citrulina e no terceiro foi feita a adição de arginina. Confira os resultados observados nos cultivos de cada mutante e as conclusões que podem ser tiradas dessa série de experimentos.

Os três mutantes selecionados não deveriam apresentar crescimento em meio mínimo e todos deveriam crescer em meio adicionado de arginina. Concorda? De fato isso ocorreu! Vamos, agora, acompanhar o comportamento do mutante “arg1”, cujo crescimento foi observado em presença dos três suplementos analisados, ornitina, citrulina e arginina. Arginina é o aminoácido necessário para o crescimento deste mutante e não é relevante para a nossa análise. Entretanto, o crescimento em presença de citrulina indicou que “arg1” contém a enzima capaz de converter este composto em arginina. Certo? Com o crescimento em ornitina podemos também descartar a possibilidade de um bloqueio na etapa de conversão de ornitina em citrulina. Podemos concluir que a incapacidade deste mutante de crescer em meio mínimo se deve a um bloqueio na via biossintética anterior à ornitina. Podemos inferir, portanto, que a via metabólica está bloqueada em relação à conversão do precursor em ornitina. Ficou mais claro?

Vamos pensar juntos sobre o que ocorreu com os outros dois mutantes. Com relação ao mutante “arg2”, observou-se crescimento apenas nos tubos contendo citrulina e arginina. A ausência de crescimento no tubo contendo ornitina indica que este mutante tem a via biossintética de arginina bloqueada na etapa de conversão de ornitina em citrulina. Já para o mutante “arg3” só se observou crescimento quando a arginina foi adicionada, indicando que pelo menos a etapa de conversão de citrulina em arginina está comprometida.

Dessa forma, Beadle e Tatum conseguiram, então, provar que a perda de uma função enzimática era o resultado de mutação em um único gene. Foi proposta, então, a hipótese *um gene–uma enzima*, ou seja, que um gene determina a síntese de uma enzima específica.

Esta hipótese foi imediatamente aceita e eventualmente modificada, pois, em seguida, observou-se que muitos genes especificam a síntese de proteínas, sendo que apenas algumas têm atividade catalítica.

Nos anos que se seguiram, este conceito foi ainda mais ampliado, já que algumas proteínas são formadas por duas ou mais subunidades que, por sua vez, podem ser codificadas por diversos genes. Passou-se, então, a aceitar a relação entre gene e polipeptídeo. Hoje em dia, sabe-se que dois grupos de genes especificam os tRNAs e os rRNAs, o que levou à maior ampliação do conceito de gene. O conceito atualmente aceito é o de que um gene especifica a síntese de uma molécula funcional.

A grande maioria dos genes contém informação para a síntese de proteínas e você estudou, nas aulas do Módulo 3, que a transcrição é uma etapa fundamental para a expressão gênica. Essa primeira etapa pode ser entendida como um mecanismo de proteção da molécula da hereditariedade, uma vez que poderia ser desastroso expô-la inteiramente cada vez que um polipeptídeo precisasse ser sintetizado. Como na célula, continuamente, ocorre degradação de proteínas e os aminoácidos resultantes são utilizados na síntese de novas moléculas protéicas, você pode reconhecer a importância da transcrição como uma etapa intermediária para a síntese de proteínas. Assim, com parte do mistério envolvendo a síntese de proteínas esclarecida, o processo, tal qual como ele é, parece bastante óbvio. Não é à toa que dizemos que a natureza é sábia!

Lembre-se de que a transcrição é a única etapa para a síntese de rRNAs e tRNAs, por processos catalisados por RNA polimerases distintas daquela responsável pela síntese de mRNA.

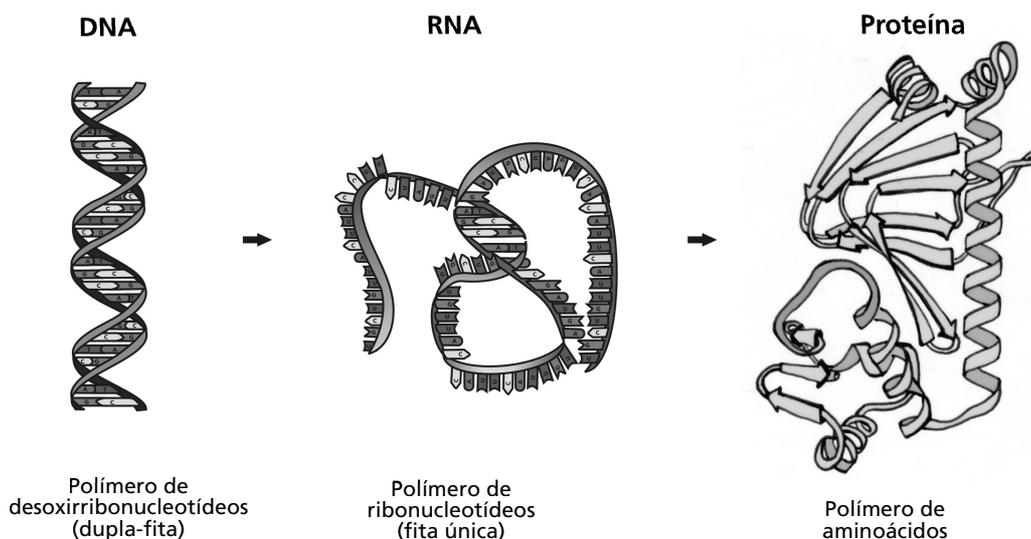
Hoje em dia, sabemos que apenas o gene a ser expresso é utilizado como molde, no processo conhecido por transcrição, para a síntese de uma molécula de RNA, que, então, carrega a informação para a síntese de polipeptídeos, o que ocorre em uma etapa denominada tradução. Já comentamos sobre o fluxo da informação genética no início da Aula 5. Releia essa aula e reveja principalmente a **Figura 5.1**, na qual estão representados os principais processos celulares envolvendo perpetuação do material genético e fluxo da informação genética.

### **POR QUE ESTE PROCESSO SE CHAMA TRADUÇÃO?**

Após a elucidação da estrutura do DNA, a proposição do mecanismo de replicação e descoberta da relação direta entre um gene e uma proteína, uma pergunta ainda pairava no ar: “Como o DNA codifica a informação genética para a síntese de proteínas?”

Primeiramente é importante lembrar que a molécula de DNA é constituída de quatro nucleotídeos diferentes. Além disso, na etapa de transcrição, a informação contida no DNA é transcrita em um formato diferente, ou seja, em RNA, constituído também pela combinação de quatro nucleotídeos. Lembre-se de que para a síntese de RNA, o fator determinante é justamente a complementaridade das bases, ou seja, para cada A, C, G e T no molde de DNA, são incorporados U, G, C e A na molécula de RNA que está sendo sintetizada.

O processo celular envolvido na síntese de proteínas é chamado tradução porque o “alfabeto” de quatro letras dos ácidos nucléicos é literalmente traduzido em um “alfabeto” inteiramente diferente. Lembre-se de que os polipeptídeos são sintetizados a partir da combinação dos vinte aminoácidos-padrão. Em outras palavras, o que ocorre é a conversão da informação “escrita” na linguagem de um ácido nucléico em linguagem de uma proteína, tal qual a conversão de um texto escrito, por exemplo, do português para o chinês, idiomas que utilizam, inclusive, caracteres diferentes. Esta conversão só é possível devido à existência de um código genético que, como você verá em detalhes nesta aula, relaciona a seqüência de bases constituintes da molécula de DNA (ou mRNA) com a seqüência dos aminoácidos que formam a proteína. Repare na **Figura 26.5** as características estruturais das principais moléculas envolvidas no fluxo da informação genética.



**Figura 26.5: DNA, RNA e proteínas.** Representação das estruturas das três principais moléculas envolvidas no fluxo da informação genética.

## A DESCOBERTA DO CÓDIGO GENÉTICO

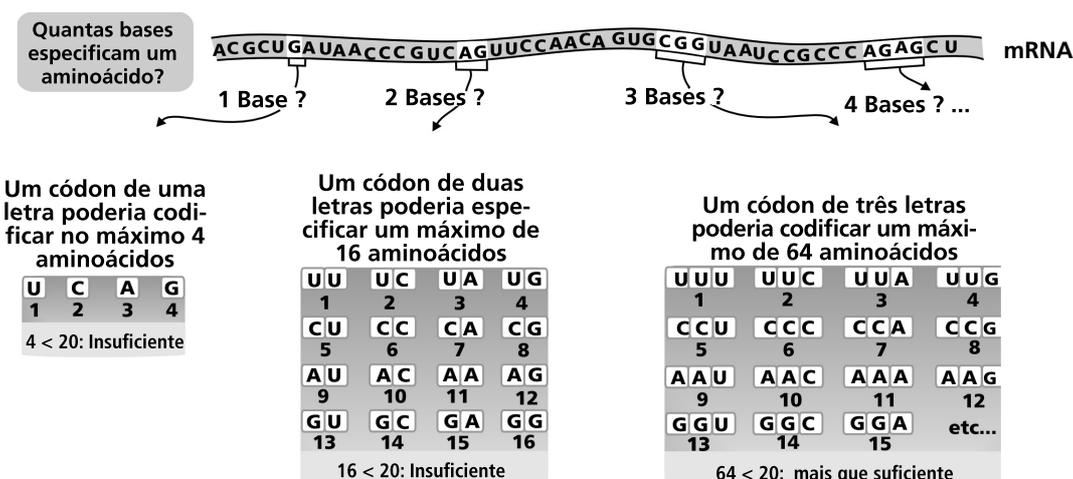
Todos os esforços feitos no sentido de se descobrir o código genético tiveram início com a busca de resposta para uma primeira pergunta: “Que tipo de combinação dos nucleotídeos poderia codificar um aminoácido?”

As moléculas de DNA são constituídas de quatro nucleotídeos, representados pelas bases A, C, T e G, que codificam os vinte aminoácidos que participam da síntese protéica. Deve haver, então, uma combinação desses quatro nucleotídeos, de forma a especificar a seqüência de aminoácidos no polipeptídeo formado. Certo?

Vamos entender o que se pensava antes de se ter conhecimento da relação entre os nucleotídeos presentes nos ácidos nucléicos e os aminoácidos que compõem as proteínas. Leia o texto acompanhando a **Figura 26.6**. A existência de quatro bases permite quatro combinações, que podem ser expressas como  $4^n$ , onde

$n$  = número de bases combinadas (1, 2, 3 ou 4).

A primeira combinação,  $4^1$  (quatro bases combinadas, uma a uma), indica que cada base codifica um aminoácido, havendo a possibilidade de se codificar apenas quatro aminoácidos ( $4^1 = 4$ ). Já a segunda combinação,  $4^2$  (quatro bases combinadas, duas a duas), explica a codificação de 16 aminoácidos ( $4^2 = 16$ ), ainda um número inferior a 20. A terceira combinação,  $4^3$  (quatro bases combinadas, três a três), permitiria a codificação de 64 aminoácidos ( $4^3 = 64$ ), número superior ao número real de aminoácidos incorporados durante a síntese de proteínas. Parecia, portanto, razoável que, pelo menos, três nucleotídeos fossem necessários para codificar cada um dos vinte aminoácidos.



**Figura 26.6: Possibilidades de combinação das bases para formação do códon.** Códon formado por combinação de uma ou duas bases resultam em número insuficiente para codificar os vinte aminoácidos-padrão. A combinação das quatro bases, três a três, parece razoável, uma vez que permite a formação de 64 códon, número mais do que suficiente para codificar os vinte aminoácidos-padrão.

Estudos genéticos indicaram que o código genético era de fato composto por trincas de nucleotídeos, conhecidas como códon, que correspondiam aos aminoácidos. Por exemplo, observou-se que adição ou deleção de um ou dois nucleotídeos causavam mutação *frameshift*, ou seja, mudança no quadro de leitura. Lembra da Aula 16? Volte a ela se tiver dúvidas neste assunto. No entanto, adição ou deleção de três ou múltiplos de três nucleotídeos mantinham o quadro de leitura.

Foi, então, a partir de uma série de experimentos, usando moldes sintéticos de mRNA, que o código genético foi desvendado.

A primeira descoberta foi descrita, no início da década de 1960, por Marshall Nirenberg e Heinrich Matthaei. O experimento desses dois pesquisadores consistiu na incubação de um polinucleotídeo de uracila (poli-U) com extrato celular da bactéria *E. coli*, GTP, ATP e uma mistura dos vinte aminoácidos em 20 tubos diferentes, sendo que, em cada um dos tubos, havia um aminoácido diferente marcado radioativamente. O que eles estavam buscando com esse experimento?

Vamos pensar juntos! A molécula de RNA formada apenas por nucleotídeo de uracila (poli-U) contém apenas um tipo de códon – UUU. Portanto, se havia uma correspondência única entre uma trinca de nucleotídeos e um aminoácido, era de se esperar que a molécula poli-U servisse de molde para a síntese de um polipeptídeo formado por apenas um tipo de aminoácido. Certo? Mas qual aminoácido?

A presença de apenas um aminoácido, marcado radioativamente, em cada um dos tubos foi fundamental para a interpretação dos resultados desse experimento. Se tudo o que se supunha era verdade, em apenas um tubo poderia se observar a formação de um polipeptídeo marcado radioativamente, ou seja, o tubo contendo o aminoácido radioativo codificado pela trinca UUU. E foi exatamente isso que eles observaram. O polipeptídeo marcado radioativamente foi observado em apenas um tubo, aquele que continha fenilalanina radioativa. Os pesquisadores, então, concluíram que o códon UUU codifica fenilalanina.

O uso de estratégia semelhante revelou que um polinucleotídeo de citosina (poli-C) codificava um polipeptídeo de prolina, ou seja, que o códon CCC codifica prolina, e que um polinucleotídeo de adenina (poli-A) codificava um polipeptídeo de lisina, indicando que o códon AAA codifica lisina.

Inúmeros mRNAs sintéticos foram utilizados em diversos experimentos, permitindo, assim, a identificação das composições de bases nas trincas que codificam quase todos os aminoácidos.

Uma limitação dessa estratégia era o tipo de mRNA utilizado como molde para a síntese dos polipeptídeos. Nesta série de experimentos, os mRNAs foram sintetizados em reações utilizando a enzima polinucleotídeo fosforilase, que catalisa a formação de polímeros de RNA a partir de ADP, CDP, GDP e UDP. Esta enzima não requer molde e os polímeros formados apresentam uma composição de bases de acordo com as concentrações relativas dos nucleosídeos 5'-difosfatos que servem como substratos, o que dificulta a definição da seqüência da molécula de RNA sintetizada.

Outras estratégias foram empregadas para se desvendar todos os mistérios do código genético. Por exemplo, os experimentos realizados por Nirenberg e Philip Leder, em 1964, envolvendo a associação de ribossomos, mRNAs e aminoacil-tRNAs foram de fundamental importância. Neste caso, foi possível determinar que tipo de aminoacil-tRNA se liga a cada uma das quase 50 das 64 possíveis trincas. Observou-se, também, que alguns códons não são reconhecidos por nenhum dos aminoacil-tRNAs, enquanto alguns aminoacil-tRNAs podem reconhecer mais de um códon.

O desenvolvimento de métodos químicos para a síntese de polirribonucleotídeos, por H. Gobind Khorana, foi decisivo para a elucidação do código genético, uma vez que permitiu a síntese de moléculas com seqüências repetidas e definidas de duas a quatro bases. Os polipeptídeos, formados a partir destes mRNAs, apresentavam padrões definidos de um ou poucos aminoácidos. Estes padrões foram analisados juntamente com todos os dados obtidos até então e, em 1966, foi estabelecido o significado de todos os 64 possíveis códons.

## O CÓDIGO GENÉTICO

O código genético pode ser visto como a correspondência entre a informação armazenada nas linguagens de ácido nucléico e de proteína. Mais especificamente a correspondência entre códons, trincas de nucleotídeos no mRNA (“lidas” na direção 5' para 3') e aminoácidos na proteína (“lidos” na direção N-terminal para C-terminal).

Na **Figura 26.7**, você pode observar as 64 trincas de nucleotídeos e o que elas especificam. À esquerda estão representadas as quatro bases que podem estar presentes na primeira posição do códon. No topo da figura estão representadas as bases que ocupam a segunda posição. Finalmente, à direita estão representadas as bases da terceira posição. Confuso? Nem tanto. Vamos entender como a gente “lê” esse código!

		Segunda base				
		U	C	A	G	
Primeira base (próxima à extremidade 5')	U	UUU } Fenilalanina (Phe ou F) UUC } UUA } Leucina (Leu ou L) UUG }	UCU } Serina (Ser ou S) UCC } UCA } UCG }	UAU } Tirosina (Tyr ou Y) UAC } UAA } <b>Stop codon</b> UAG } <b>Stop codon</b>	UGU } Cisteína (Cys ou C) UGC } UGA } <b>Stop codon</b> UGG } Triptofano (Trp ou W)	Terceira base (próxima à extremidade 3')
	C	CUU } Leucina (Leu ou L) CUC } CUA } CUG }	CCU } Prolina (Pro ou P) CCC } CCA } CCG }	CAU } Histidina (His ou H) CAC } CAA } Glutamina (Gln ou Q) CAG }	CGU } Arginina (Arg ou R) CGC } CGA } CGG }	
	A	AUU } Isoleucina (Ile ou I) AUC } AUA } Metionina* (Met ou M) <b>AUG</b> }	ACU } Treonina (Thr ou T) ACC } ACA } ACG }	AAU } Asparagina (Asn ou N) AAC } AAA } Lisina (Lys ou K) AAG }	AGU } Serina (Ser ou S) AGC } AGA } Arginina (Arg ou R) AGG }	
	G	GUU } Valina (Val ou V) GUC } GUA } GUG }	GCU } Alanina (Ala ou A) GCC } GCA } GCG }	GAU } Ácido aspártico (Asp ou D) GAC } GAA } Ácido glutâmico (Glu ou E) GAG }	GGU } Glicina (Gly ou G) GGC } GGA } GGG }	

\* AUG também serve como códon de iniciação na síntese de proteína.

**Figura 26.7: O código genético.** Os códons são escritos na direção 5'→3'. Em destaque, estão representados os códons de terminação da síntese protéica (stop codons) e os aminoácidos codificados por apenas um códon, metionina (AUG) e triptofano (UGG). AUG também serve como códon de iniciação na síntese de proteína.

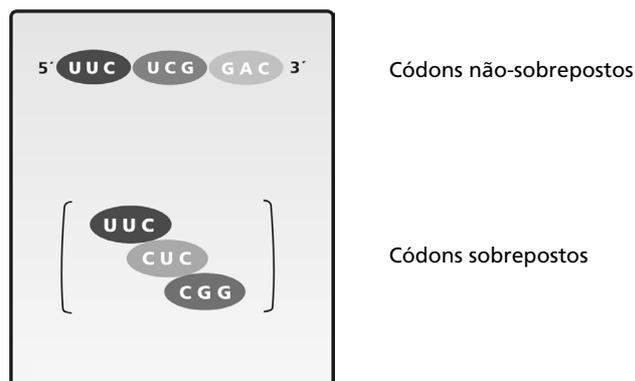
Vamos procurar na **Figura 26.7** o aminoácido codificado pela trinca ACA. A primeira letra do códon é A; portanto, procure esta letra à esquerda da figura. Achou? Todas as trincas que se iniciam com a letra A estão posicionadas na terceira linha. Concorda? O códon que estamos procurando tem C como segunda letra. Logo, devemos procurar na parte superior da figura qual a coluna que corresponde a esta letra na segunda posição. Confirme que, na segunda coluna, encontram-se todas as trincas que têm C como segunda letra. Certo? Então o códon que procuramos se encontra na terceira linha e na segunda coluna. Agora resta localizar nesta célula o códon ACA, que é o terceiro de cima para baixo. O que você encontrou? Espero que você tenha achado o códon certo. Apenas para confirmar, o códon ACA codifica o aminoácido treonina. Não é tão difícil assim, é? Talvez você precise treinar um pouco. Nos exercícios, você terá a oportunidade de checar se entendeu tudo bem direitinho.



- Um códon – AUG – codifica, normalmente, o início da tradução, sendo denominado códon de iniciação. Esse é o códon que especifica o aminoácido metionina. Outros códons de iniciação também podem ser encontrados, como por exemplo GUG. Procure o códon AUG na **Figura 26.7**.
- Três códons (UAA, UAG e UGA) não especificam nenhum aminoácido, mas atuam como códons de parada ou terminação da etapa de tradução. O significado dos códons de terminação é determinado diretamente por fatores protéicos. Localize esses códons na **Figura 26.7**.
- O código genético é usado em todos os organismos. Entretanto, o código genético “padrão” é comum, mas não é universal. Apesar de os códons serem praticamente idênticos em quase todos os seres vivos, existem diferenças no código genético de mitocôndrias e cloroplastos. Confira no **Quadro 26.1**. O códon UGA, por exemplo, no código padrão, corresponde a um códon de término de tradução. Entretanto, no código mitocondrial, UGA codifica triptofano. Já AUA, que normalmente codifica isoleucina, na mitocôndria codifica metionina.
- Os códons não se sobrepõem. Na **Figura 26.9**, você pode observar o efeito que teria a sobreposição dos códons, o que ainda não foi observado.

**Quadro 26.1: Códons que não fazem parte do código genético padrão.**

Códons	Código padrão	Código mitocondrial
UGA	Stop	Trp
UGG	Trp	Trp
AUA	Ile	Met
AUG	Met	Met
AGA	Arg	Stop
AGG	Arg	Stop



**Figura 26.9:** Não-sobreposição *versus* sobreposição de códons.

- Não existe qualquer tipo de pontuação entre os códons, ou seja, estabelecida a primeira trinca (três bases), cada trinca seguinte codifica um aminoácido diferente até que se atinja o(s) códon(s) de terminação da síntese protéica. Em outras palavras, estabelecido o início da tradução, não existe “base ociosa” que não faça parte de um códon. O códon iniciador, a primeira trinca, estabelece o quadro (ou fase) de leitura. Essa característica, que ficará mais bem entendida no item seguinte, você pode visualizar na **Figura 26.10**.
- Em uma dada seqüência, é possível estabelecer três quadros (ou fases) de leitura. Por tudo que já estudamos até agora, é fácil você entender o porquê, não? Observe a **Figura 26.10**. Uma mesma seqüência de nucleotídeos de um RNA pode ser traduzida em qualquer um dos três quadros de leitura, dependendo de onde o processo de descodificação se inicia. Esses quadros de leitura se estabelecem de acordo com o códon iniciador. Por exemplo, o quadro de leitura 1 se inicia na primeira base da seqüência, tendo como códon iniciador o CUC. O quadro de leitura 2 tem início na segunda base, com o códon iniciador UCA. Já o quadro de leitura 3 começa com a terceira base, que forma o códon iniciador CAG. Cada quadro de leitura apresenta uma seqüência distinta de códons e, obviamente, codificará uma proteína diferente. Em geral, apenas uma das três possíveis fases de leitura em um mRNA codifica uma proteína específica.

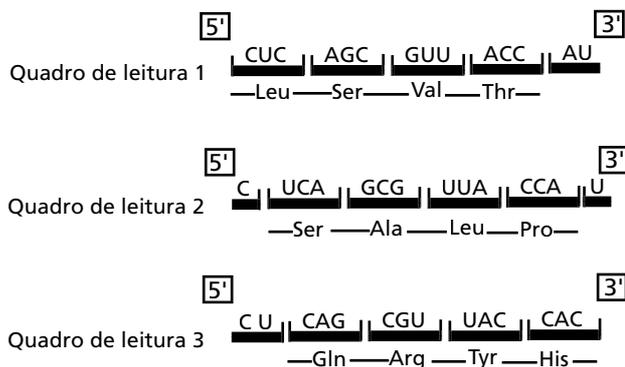


Figura 26.10: Os três possíveis quadros (ou fases) de leitura para uma seqüência de RNA. Cada um dos quadros de leitura codifica uma proteína diferente.

O termo quadro ou a fase de leitura aberta – *Open Reading Frame* (ORF) – é muito empregado na área da genômica. ORF designa uma região do genoma que apresenta cinquenta ou mais códons desde um códon de iniciação até um códon de terminação, podendo estar representando um gene.

## RESUMO

Nesta aula, você teve a oportunidade de acompanhar o experimento clássico de Beadle e Tatum, no qual foi possível confirmar a relação entre um gene e uma enzima. Esta foi uma grande conquista na área da Genética, uma vez que comprovou que a herança não é apenas uma “potencialidade”, mas algo que se manifesta através das moléculas funcionais codificadas pelos diferentes genes. A descoberta do código genético também foi um marco na história da Ciência Biológica, vindo a estabelecer a relação direta entre a seqüência de nucleotídeos em um DNA (ou mRNA) e a seqüência de aminoácidos no polipeptídeo por ele codificado. O código genético apresenta diversas características, destacando-se o fato de o códon ser constituído de uma trinca de nucleotídeos, de existir um códon que determina o início da tradução, e que estabelece o quadro de leitura, e três códons que determinam o término da síntese protéica. Podemos também dizer que o código genético é degenerado e comum e não universal – como se pensava inicialmente. A disposição dos códons em uma molécula de RNA não apresenta sobreposição e nem pontuação, ou seja, a partir do códon de início, a cada três nucleotídeos se tem um novo códon. Portanto, para um mesmo RNA, dependendo da posição do códon inicial, é possível se estabelecer três quadros (ou fases) de leitura.

## EXERCÍCIOS

1. O conceito de gene teve de ser mudado à medida que novos conhecimentos foram adquiridos. Qual o conceito inicial? Que descobertas determinaram a mudança do conceito inicial? Qual o conceito aceito atualmente?
2. Como a utilização de raios X contribuiu para a proposição, por Beadle e Tatum, da hipótese um gene–uma enzima?
3. Cite os aminoácidos codificados pelos seguintes códons: CGG, UAC, GGA, ACU, UCA, GCG, UUA e AAU.
4. Cite o(s) códon(s) que codifica(m) os seguintes aminoácidos: arginina, histidina, metionina, prolina, triptofano e valina.
5. Quantos pares de bases no DNA devem ser deletados em um processo mutacional para que sejam eliminados três aminoácidos da cadeia polipeptídica sem causar alteração do restante da seqüência protéica? Explique.

## AUTO-AVALIAÇÃO

Você já deve estar percebendo que, apesar de dividirmos as disciplinas em tópicos, ao final todos os temas são necessários para o entendimento global. No exercício 1, por exemplo, você teve de resgatar um pouco do que aprendeu nas Aulas 1 e 2. Já o exercício 2 exigiu seus conhecimentos sobre Mutação (Aula 16). Os exercícios 3 e 4 fizeram com que você brincasse um pouco com o código genético, permitindo que você se familiarizasse com a relação códon–aminoácido. Por fim, para responder ao exercício 5, muito certamente você teve de voltar à parte que discute as características do código genético. Caso tenha tido dificuldade para resolver os exercícios, releia a aula e recorra aos tutores, que estarão sempre nos pólos dispostos a ajudar.

## FIQUE ATENTO!

Na próxima aula, discutiremos as características do processo de tradução, as moléculas envolvidas e uma etapa importante conhecida como ativação dos aminoácidos. Até lá!



## Fluxo da informação genética – tradução II

AULA

# 27

## objetivos

Nesta aula, terá oportunidade de:

- Reconhecer as particularidades do processo de tradução.
- Listar os principais elementos da maquinaria envolvida na síntese de proteínas e suas peculiaridades.
- Saber descrever a etapa de ativação dos aminoácidos.

### Pré-requisitos

Esta aula será mais facilmente assimilada se você não tiver dúvidas em relação às aulas sobre RNA (Aula 5) e transcrição (aulas do Módulo 3). É claro que ter entendido a Aula 26, sobre a parte introdutória da tradução, também é de suma importância.

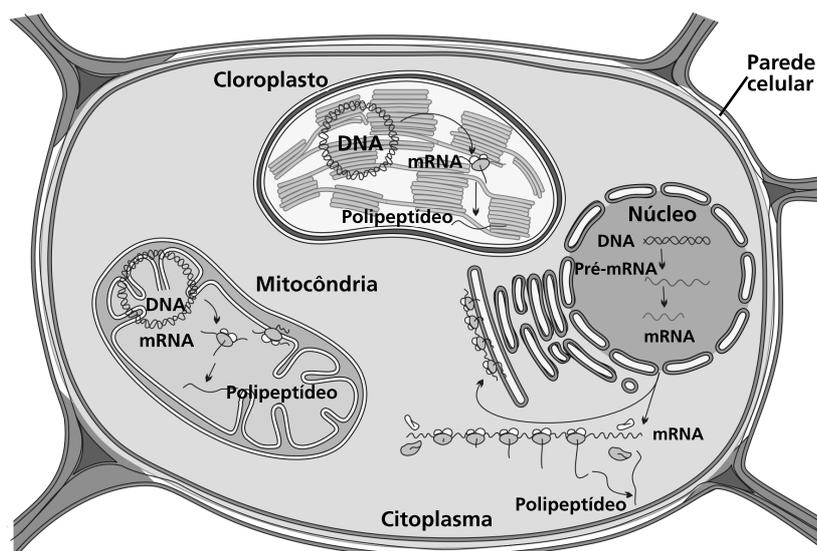
## INTRODUÇÃO

Na Aula 26, foram discutidos alguns experimentos que culminaram na proposição da relação um gene–uma enzima e na descoberta do código genético. Nesta aula, você estudará as peculiaridades da síntese protéica e as moléculas que participam desse processo. Você estudará, ainda, uma primeira etapa, anterior à tradução propriamente dita, que envolve a ativação de aminoácidos. Como discutiremos mais adiante, a participação dos aminoácidos na síntese protéica depende essencialmente desse primeiro ciclo de reações.

Vamos começar!

## PECULIARIDADES DA SÍNTESE DE PROTEÍNA

Antes de começarmos, é importante que você saiba que a síntese de proteína não está limitada ao citoplasma. Ela ocorre também nas mitocôndrias e nos cloroplastos, que possuem material genético próprio e funcionam de forma independente em relação à transcrição e à tradução de seus genes. Para você fixar melhor essa informação, busque a **Figura 27.1**, na qual você encontrará, como exemplo, um esquema de uma célula vegetal, que você já deve ter estudado em Biologia Celular, com destaque para os locais onde ocorre tradução. Nesta aula, entretanto, focaremos apenas a síntese de proteína que ocorre no citoplasma.



**Figura 27.1:** Locais de síntese de proteínas em uma célula vegetal. Esquema de uma célula vegetal indicando os locais onde ocorre síntese protéica, citoplasma, mitocôndria e cloroplasto.

A síntese de proteínas engloba a tradução propriamente dita, ou seja, a síntese de um polipeptídeo, e outras etapas que garantem a funcionalidade da molécula sintetizada. A tradução ocorre em três estágios: iniciação, alongamento e terminação, que serão estudados na Aula 28. As outras duas etapas, fundamentais para a síntese correta de uma proteína, são a ativação dos aminoácidos, assunto desta aula, e o processamento pós-traducional dos polipeptídeos recém-sintetizados, um dos temas da Aula 29. Uma vez terminada a síntese, as proteínas devem, ainda, ser transportadas para os locais em que desempenharão suas funções. Este processo é conhecido por endereçamento de proteínas e será discutido na Aula 29.

Pensamos, primeiramente, em iniciar esta aula apresentando os principais aspectos que tornam a síntese protéica um processo peculiar, de forma que eles recebessem o merecido destaque. Entretanto, entendemos que você assimilaria melhor cada uma dessas características ao longo desta aula e da Aula 28. Portanto, fique atento ao texto, em particular aos “boxes”, com informações relacionadas:

- à complexidade do processo;
- ao local onde ocorre a tradução;
- à molécula que serve de molde para a biossíntese de proteína;
- à direção da “leitura” do mRNA e da síntese do polipeptídeo;
- à importância dos tRNAs como adaptadores desse processo;
- à velocidade da síntese de proteína;
- ao gasto de energia para sintetizar uma proteína;
- à regulação da síntese protéica.

## MAQUINARIA DA SÍNTESE PROTÉICA

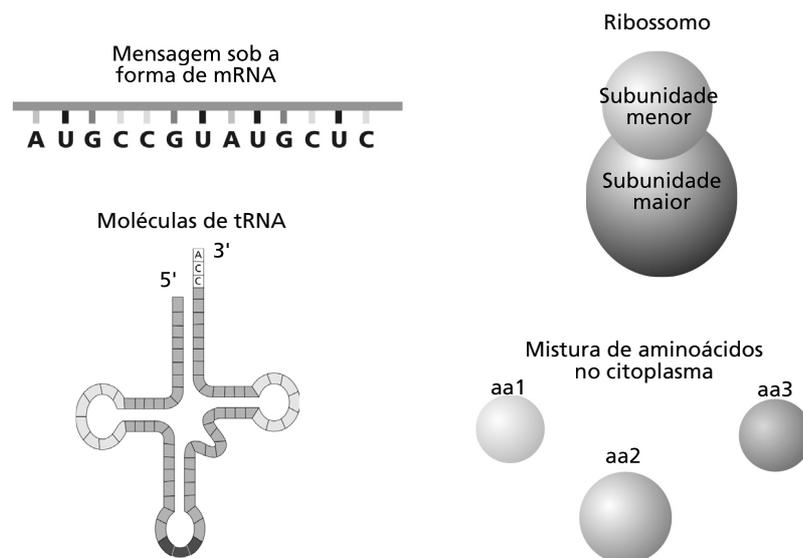
Podemos dizer que a tradução é um processo muito mais complexo que a transcrição e a replicação, pois não se baseia no pareamento comum de bases nitrogenadas. A tradução necessita da participação coordenada de mais de 300 macromoléculas, e muitas delas estão organizadas na estrutura tridimensional complexa do ribossomo. Essas macromoléculas incluem, além do mRNA, diferentes moléculas de tRNA, rRNA e inúmeras proteínas.

Para você ter uma idéia da complexidade desse processo, em células eucarióticas, a síntese de uma proteína conta com a participação de mais

de setenta proteínas ribossomais diferentes; vinte ou mais enzimas para ativar os aminoácidos; uma dúzia ou mais de enzimas auxiliares e outros fatores protéicos específicos para iniciação, alongamento e terminação da tradução e cerca de cem ou mais enzimas adicionais para o processamento final dos mais diferentes tipos de proteínas. Além das proteínas, são necessários um mRNA e quarenta ou mais tipos de tRNA e rRNA.

Na bactéria *E. coli*, o total dos diferentes tipos de moléculas de proteínas e RNAs envolvidos na síntese protéica é similar ao que se observa nas células eucarióticas. Os organismos procarióticos e eucarióticos contêm milhares de cópias de cada proteína e tipo de RNA por célula, o que equivale a aproximadamente 20.000 ribossomos, 100.000 fatores protéicos e enzimas e 20.000 tRNAs. Para uma bactéria, as moléculas envolvidas na síntese de proteína podem representar mais de 35% do peso seco da célula.

Apesar da grande variedade de moléculas que participam desse processo, os principais elementos necessários à tradução, que serão estudados nesta aula, são: a molécula de mRNA, que contém a informação genética, os ribossomos, os tRNAs e os aminoácidos. Confira na **Figura 27.2**. Muito do que se conhece a respeito dessas moléculas já foi abordado em aulas da nossa disciplina (Aula 5) ou de Bioquímica. Não iremos repetir o conteúdo anteriormente estudado, pois nosso objetivo aqui é destacar apenas os aspectos importantes para o processo de tradução.

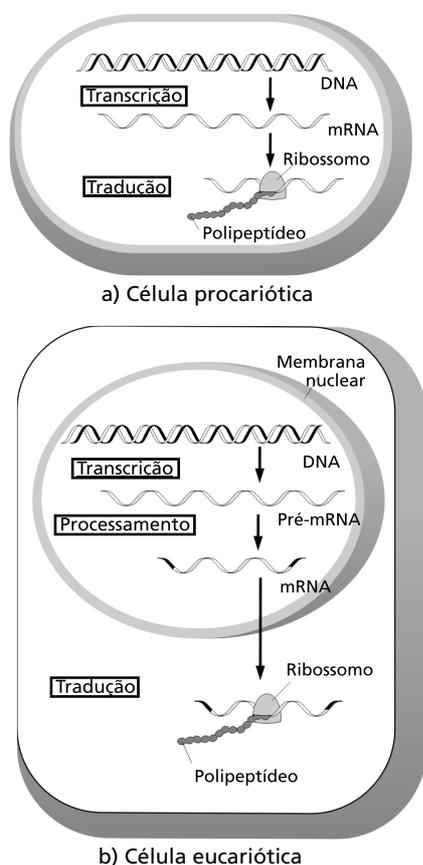


**Figura 27.2:** Principais elementos necessários à tradução. Mensagem sob a forma de mRNA, ribossomo, moléculas de tRNA e mistura de aminoácidos no citoplasma.

## mRNA – MOLÉCULA CARREADORA DA INFORMAÇÃO GENÉTICA

Já comentamos anteriormente que o mRNA é a molécula que leva a informação genética armazenada sob a forma de DNA para o ribossomo, no qual será utilizada como molde para a síntese protéica.

Em células procarióticas, que não apresentam núcleo, a transcrição e a tradução ocorrem simultaneamente, ou seja, a tradução inicia antes mesmo que o transcrito, a molécula de mRNA, esteja completamente sintetizado. Já em células eucarióticas, para a grande maioria das moléculas, a transcrição acontece no núcleo, e o transcrito primário (pré-mRNA) é processado. O mRNA formado é, então, exportado para o citoplasma, onde se encontram os ribossomos. Confira na **Figura 27.3**.

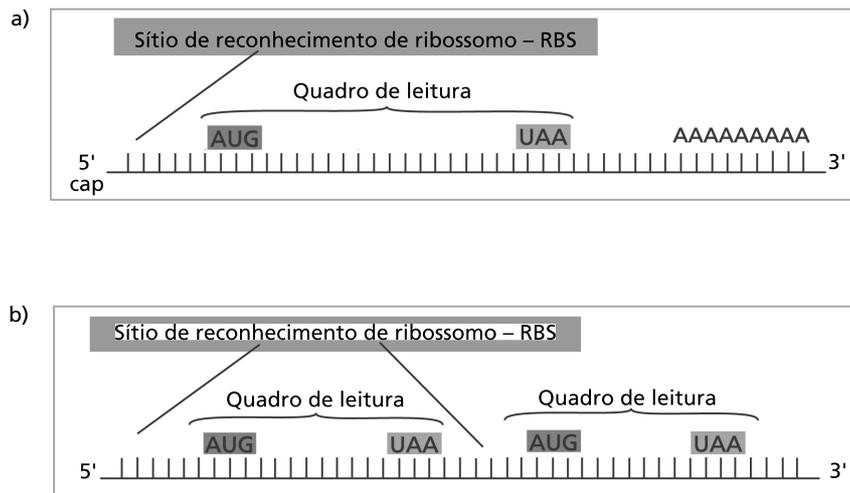


**Figura 27.3:** Fluxo da informação genética em procarioto e eucarioto.

Em caso de dúvidas, volte às aulas do Módulo 3, que tratam justamente da síntese de RNA pelo processo conhecido como transcrição. Lembra?

A molécula de mRNA é responsável por levar a mensagem genética armazenada no DNA para o ribossomo, onde funcionará como molde para a síntese de proteínas através de suas trincas de nucleotídeos que especificam a ordem dos aminoácidos na cadeia polipeptídica.

Os mRNAs de procariotos e eucariotos apresentam características marcantes que conferem estabilidade, favorecem o transporte do núcleo para o citoplasma, no caso de organismo eucariótico, e/ou são de suma importância para o processo de tradução. Veja na **Figura 27.4** as regiões em destaque para os mRNAs procarióticos e eucarióticos.



**Figura 27.4:** mRNAs de eucariotos e procariotos. (a) mRNA eucariótico. (b) mRNA procariótico.

Vamos relembrar primeiramente os detalhes do mRNA eucariótico (**Figura 27.4.a**), que contém informação para a síntese de uma única cadeia polipeptídica! Além da região codificadora, determinada por um códon de iniciação de tradução, em geral AUG, e um ou mais códons de terminação, UAA, UAG ou UGA, o mRNA contém duas regiões não traduzidas, uma localizada na extremidade 5' – 5'UTR (*UnTranslated Region*) e outra localizada na extremidade 3' – 3'UTR. O mRNA transportado para o citoplasma apresenta um nucleotídeo modificado (7-metil guanosina) na extremidade 5', gerando a extremidade 5'cap, e uma seqüência de nucleotídeos na extremidade 3', conhecida como cauda poli-A. Você se lembra desses detalhes? É justamente na região envolvendo a extremidade 5'cap no mRNA de eucarioto que se localiza o sítio de reconhecimento do ribossomo – RBS (*Ribosome Recognition Site*) que, como seu próprio nome diz, é o local onde o ribossomo se liga ao mRNA para iniciar a tradução. Você verá em detalhes como essa ligação ocorre na Aula 28.

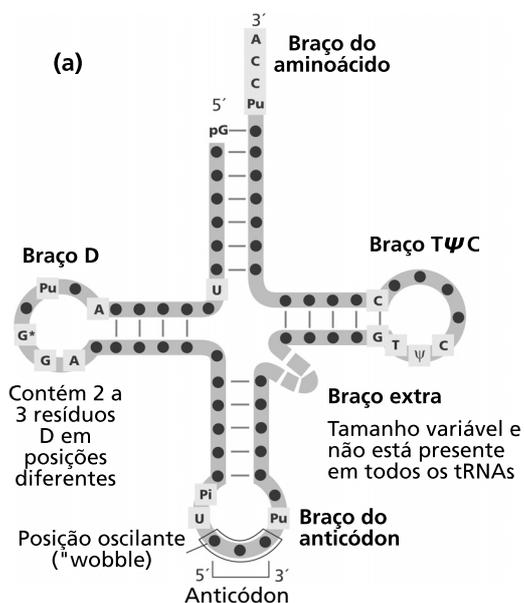


## tRNA E AMINOÁCIDOS

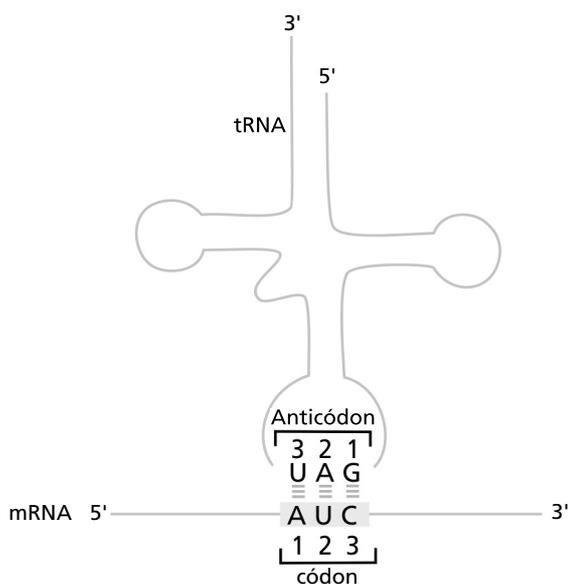
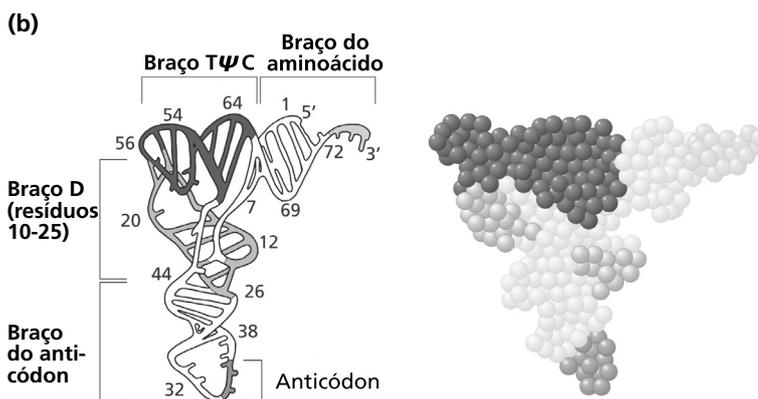
Os códons não são reconhecidos diretamente pelos aminoácidos que eles especificam. Esta etapa depende de moléculas adaptadoras, os tRNAs, que reconhecem e se ligam tanto ao códon quanto ao aminoácido. Vamos entender como isso é possível!

Você já estudou, na Aula 5, que, por pareamento de bases entre seqüências complementares intracadeias, as moléculas de tRNA formam estruturas secundárias e se assemelham à folha de trevo. Essa estrutura sofre dobramento posterior para formar a estrutura tridimensional em forma de L compacta. Lembra? Reveja essas estruturas na **Figura 27.6**. Duas regiões de nucleotídeos não emparelhados localizados em terminais opostos da molécula em forma de L são fundamentais para a função do tRNA na síntese de proteínas. Uma dessas regiões é um segmento curto em fita simples no terminal 3' da molécula e está do lado em que o aminoácido correspondente ao códon se liga. A outra região, o anticódon, é complementar à seqüência específica do códon na molécula de mRNA. Nesse caso, o pareamento entre códon e anticódon ocorre devido ao alinhamento antiparalelo dos dois RNAs, como pode ser visto na **Figura 27.7**. Como a relação entre um códon e um aminoácido é determinada pelo tRNA ligado a ele, podemos, então, dizer que os tRNAs são moléculas adaptadoras que fazem a correspondência entre os aminoácidos e os códons no mRNA.

Os aminoácidos “livres” não participam da tradução. Em lugar disso, os aminoácidos são levados para o ribossomo ligados a tRNAs, que são os adaptadores desse processo. Nesse caso, os precursores ativados são os aminoacil-tRNAs, nos quais o grupo carboxila de um aminoácido é ligado ao terminal 3'-OH do tRNA, em uma reação catalisada por aminoacil-tRNA sintetases. Esta etapa é chamada ativação de aminoácido, e será estudada ainda nesta aula.



**Figura 27.6:** Estrutura do tRNA. (a) Estrutura secundária destacando as regiões mais importantes. (b) Estrutura tridimensional em forma de L, indicando o posicionamento do braço do aminoácido e do braço do anticódon. D- nucleotídeo de dihidroxi-uracila.



**Figura 27.7:** Pareamento códon-anticódon. O tRNA foi representado por sua estrutura secundária. Note as posições das extremidades 5' e 3' das duas moléculas de RNA, tRNA e mRNA, o que confirma o alinhamento antiparalelo e o pareamento entre bases que compõem o códon e o anticódon.

Os tRNAs são específicos para cada aminoácido, existindo pelo menos um tipo de tRNA correspondente a cada um dos vinte tipos de aminoácidos, conhecidos como aminoácidos-padrão, que são incorporados nas proteínas. Para que a tradução ocorra de forma precisa, os tRNAs têm de apresentar duas características importantes. Primeiramente, devem ser distinguíveis uns dos outros pelas moléculas que reconhecem tipos específicos de tRNAs. Além disso, devem ser reconhecidos pelas moléculas que interagem com todos os tRNAs.

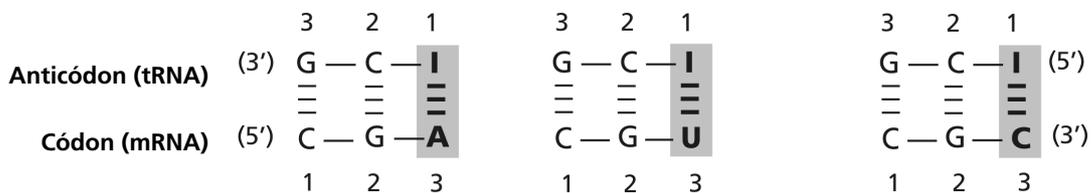
Você já estudou alguns aspectos dos tRNAs na Aula 5. Volte a ela se achar necessário!

**tRNA<sup>SER</sup>**

A notação simplificada de um aminoacil-tRNA é tRNA<sup>símbolo de três letras do aminoácido</sup>. Por exemplo, tRNA<sup>Ser</sup> representa o seril-tRNA, o tRNA<sup>Val</sup> representa o valil-tRNA, e assim por diante.

Alguns tRNAs correspondem a um único tipo de aminoácido, sendo conhecidos como tRNAs cognatos. Por exemplo, o tRNA<sup>SER</sup> designa o tRNA que corresponde ao aminoácido serina.

Outros aminoácidos têm diversos tRNAs cognatos, que são chamados tRNAs isoceptores. Os anticódons destes tRNAs contêm, na extremidade 5', o nucleotídeo inosinato, representado por I, constituído pela base rara hipoxantina (veja sua estrutura na Figura 5.7 da Aula 5). O inosinato é capaz de formar pontes de hidrogênio com três nucleotídeos diferentes (U, C e A), embora essas interações sejam mais fracas do que as pontes de hidrogênio entre os pares de Watson e Crick (Figura 27.8). Por isso, esta posição, que pode ser confirmada na Figura 27.6.a, é chamada oscilante (*wobble*). Assim, as duas primeiras bases dos códons CGA, CGU e CGC formam pares de Watson-Crick fortes com as bases correspondentes no anticódon do tRNA, enquanto a terceira base, A, U ou C, interage com o inosinato (*wobble*), permitindo que um único tRNA reconheça mais do que um único códon.



**Figura 27.8:** Possíveis pareamentos quando o anticódon do tRNA possui inosinato. O mesmo tRNA é capaz de parear com três códons distintos, devido à presença do inosinato na posição oscilante.

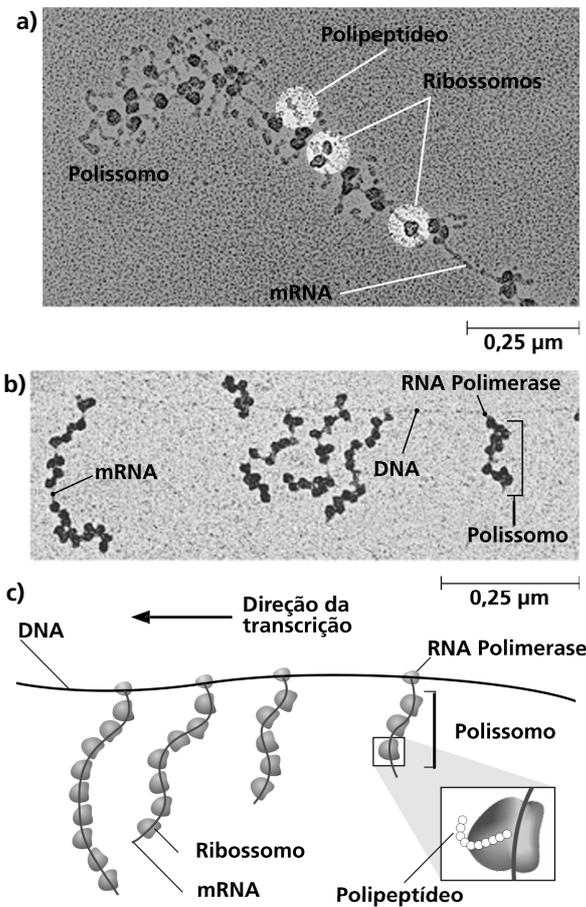
Apesar de a terceira base contribuir para a especificidade, o fato de sua interação com o anticódon ser fraca permite a rápida dissociação do tRNA de seu códon durante a tradução. Provavelmente, se as três pontes de hidrogênio formadas entre as bases do códon e do anticódon fossem fortes, os tRNAs se dissociariam muito lentamente, o que poderia comprometer a velocidade do processo. Portanto, podemos concluir que as interações códon/anticódon devem atender não apenas à acurácia como também à velocidade da síntese de proteína.

## RIBOSSOMO

Como já foi dito em vários momentos da nossa disciplina, os ribossomos constituem o local de síntese protéica. Os detalhes estruturais dos ribossomos foram discutidos na Aula 5, mas vale a pena relembrar que eles são constituídos de diferentes tipos de rRNAs e inúmeras proteínas, chamadas proteínas ribossomais. Todos os ribossomos, procarióticos e eucarióticos, são formados por duas subunidades, menor e maior, lembra?

Também vimos que os rRNAs apresentam estruturas tridimensionais específicas determinadas por pareamento intracadeia das bases nitrogenadas. Possivelmente, os rRNAs servem como uma armação à qual as proteínas ribossomais estão ligadas. Cada uma dessas proteínas deve desempenhar um papel importante na síntese de polipeptídeos, ou como uma enzima, ou como um componente estrutural do processo global.

É importante ressaltar que, apesar do alto grau de complexidade, as proteínas são sintetizadas a uma velocidade surpreendente. Isso ocorre devido à formação, tanto em células procarióticas quanto em eucarióticas, de grupos de dez a cem ribossomos ligados a uma única molécula de mRNA, denominados polirribossomos, ou simplesmente polissomos, que podem ser vistos com auxílio de microscópio eletrônico. Observe na Figura 27.9.a que a fita que conecta os ribossomos é a molécula de mRNA que está sendo, simultaneamente, traduzida pelo conjunto de ribossomos, garantindo o uso eficiente do mRNA. Os ribossomos individuais nesses polissomos estão separados de tal forma que a densidade máxima no mRNA é de aproximadamente um ribossomo por cada oitenta nucleotídeos. Os polissomos se originam porque, uma vez que um ribossomo ativo tenha liberado o sítio de iniciação, um segundo ribossomo pode se ligar ao mesmo sítio. Esse processo será discutido mais extensivamente na Aula 28.



**Figura 27.9:** Polissomos. (a) A estrutura do polissomo é formada por uma molécula de mRNA ligada a vários ribossomos, que simultaneamente participam da tradução. (b) Nas células procarióticas, a tradução se inicia antes mesmo de terminada a transcrição. (c) Diagrama explicativo da parte b.

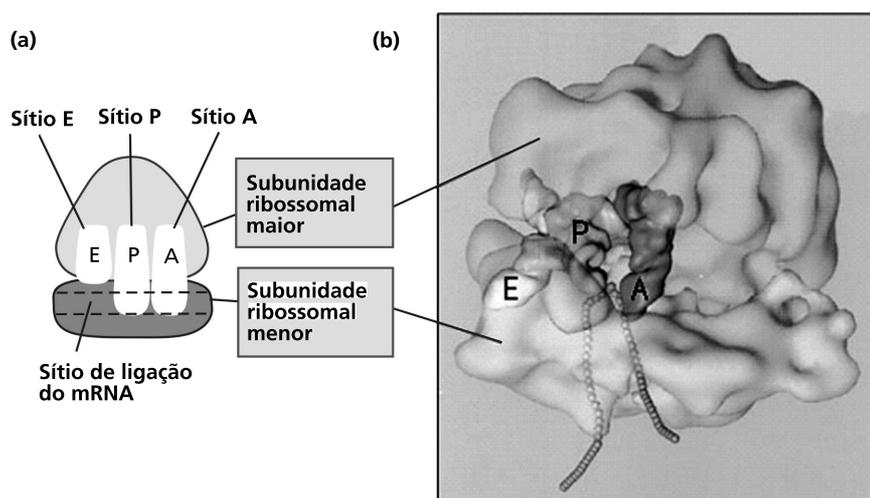
Além disso, em organismos procarióticos, a transcrição e a tradução ocorrem simultaneamente, o que também pode ser visualizado por microscopia eletrônica, como na **Figura 27.9.b**. Na **Figura 27.9.c**, você encontra um diagrama explicativo da **Figura 27.9.b**.

Na bactéria *E. coli*, a síntese completa de uma cadeia polipeptídica de 100 resíduos a 37°C ocorre em, aproximadamente, 5 segundos. Em média, os ribossomos eucarióticos e procarióticos podem adicionar, respectivamente, 2 e 20 aminoácidos por segundo. É uma velocidade muito alta se considerarmos a complexidade do processo, não?

Uma característica importante dos ribossomos é o fato de suas duas subunidades apresentarem formas irregulares e se encaixarem, formando uma fenda através da qual o mRNA passa quando o ribossomo se move ao longo do mesmo durante o processo de tradução, e a partir do qual a cadeia polipeptídica crescente emerge. As duas subunidades ribossomais permanecem separadas,

juntando-se à molécula de mRNA, nas proximidades do terminal 5', apenas para dar início à síntese de uma proteína. A subunidade menor combina os tRNAs aos códons do mRNA, enquanto a subunidade maior catalisa a formação das ligações peptídicas. Durante a etapa de alongamento da tradução, os ribossomos se movem ao longo do mRNA, um códon por vez. Ao final do processo, as duas subunidades do ribossomo se separam e ficam livres para iniciar uma nova síntese protéica. Você estudará detalhadamente todo esse processo na Aula 28. Aguarde!

Além do sítio de ligação do mRNA, os ribossomos têm três outros sítios importantes: o sítio aminoacil ou A (*Aminoacyl*), no qual o aminoacil-tRNA se liga trazendo o aminoácido a ser incorporado à cadeia; o sítio peptidil ou P (*Peptidyl*), em que se encontra o peptidil-tRNA, ou seja, o tRNA ligado à cadeia polipeptídica nascente; e o sítio de saída ou E (*Exit*), onde o tRNA não carregado, ou livre, posiciona-se antes de ser liberado durante a síntese. Não mais do que dois desses sítios contêm tRNAs em um dado momento. Na **Figura 27.10**, você encontra um esquema do ribossomo com destaque para os sítios A, E e P. Na Aula 28, voltaremos a citar esses sítios quando estudarmos as etapas de tradução.



**Figura 27.10:** Sítios A, P e E no ribossomo. (a) Representação esquemática do ribossomo, com destaque para as subunidades maior e menor, sítio de ligação do mRNA e os sítios A, P e E. (b) Modelo tridimensional do ribossomo com destaque para os sítios A, P e E.

## ATIVAÇÃO DE AMINOÁCIDO

A ativação dos aminoácidos é uma etapa fundamental, pois, como citado anteriormente, os aminoácidos livres no citoplasma não participam da síntese de proteínas do ribossomo. Os precursores ativados são os aminoacil-tRNAs, nos quais o grupo carboxila de um aminoácido está ligado covalentemente ao terminal 3'OH do seu tRNA cognato.

Durante esta etapa, que ocorre no citosol e não nos ribossomos, cada um dos vinte aminoácidos-padrão se liga ao tRNA específico à custa de energia sob a forma de ATP. Essas reações são catalisadas por um grupo de enzimas ativadoras dependentes de  $Mg^{2+}$  chamadas, genericamente, aminoacil-tRNA sintetases. Por exemplo, a fenilalanil-tRNA sintetase, cuja notação simplificada é **Phe-tRNA sintetase**, é a enzima que catalisa a reação de formação do tRNA<sup>Phe</sup>.

### PHE-tRNA SINTETASE

A notação simplificada de uma aminoacil-tRNA sintetase é – Símbolo de três letras do aminoácido-tRNA sintetase. Assim, Phe-tRNA sintetase é a fenilalanil-tRNA sintetase, Leu-tRNA sintetase é a leucil-tRNA sintetase, e assim por diante.

É importante que você saiba que para cada aminoácido e seus tRNAs correspondentes existe uma enzima aminoacil-tRNA sintetase, ou seja, quando existem dois ou mais tRNAs para um mesmo aminoácido, geralmente uma única enzima catalisa a reação de todos eles.

A importância da ativação se deve ao fato de que a síntese correta de uma proteína requer duas etapas igualmente importantes de reconhecimento de moléculas. Primeiramente, deve ser feita a escolha do aminoácido correto para a ligação covalente a um tRNA. Em seguida, deve ser feita a seleção do tRNA carregado de um aminoácido especificado pelo códon presente no mRNA.

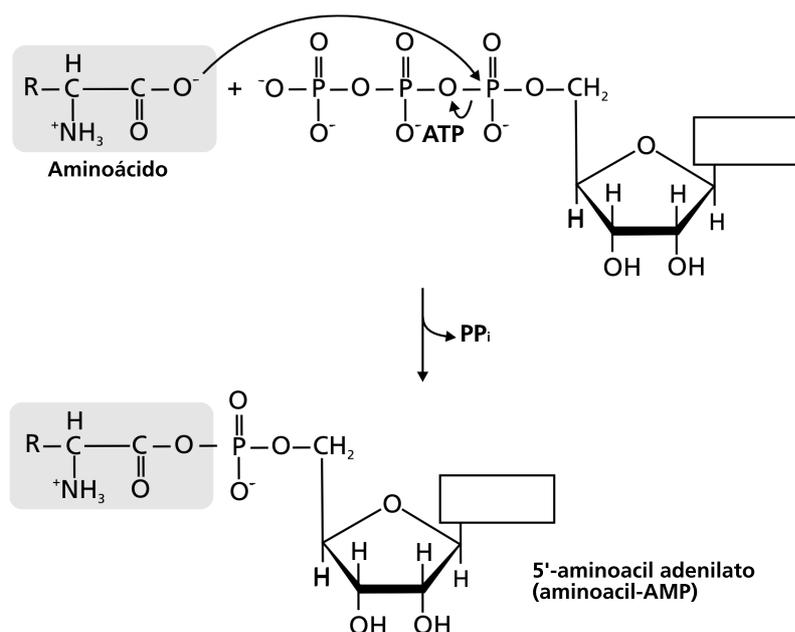
Chegou a hora de você entender como ocorre a ativação dos aminoácidos, o que será mais fácil se você ler o texto acompanhando os esquemas mais detalhados representados nas Figuras 27.11, 27.12 e um esquema mais simplificado de todo o processo representado na Figura 27.13.

A ativação ocorre em duas reações seqüenciais catalisadas por uma mesma enzima. O processo, que é energeticamente desfavorável, torna-se possível apenas mediante a hidrólise de uma molécula de ATP. Isso ocorre em uma primeira etapa, na qual a aminoacil-tRNA sintetase reconhece seu aminoácido correspondente e seu segundo substrato, um ATP. Forma-se, então, uma ligação entre o grupo carboxila do aminoácido e o fosfato do AMP, originando o AMINOACIL-AMP, com liberação de fosfato inorgânico (PPi), como você pode conferir na Figura 27.11.

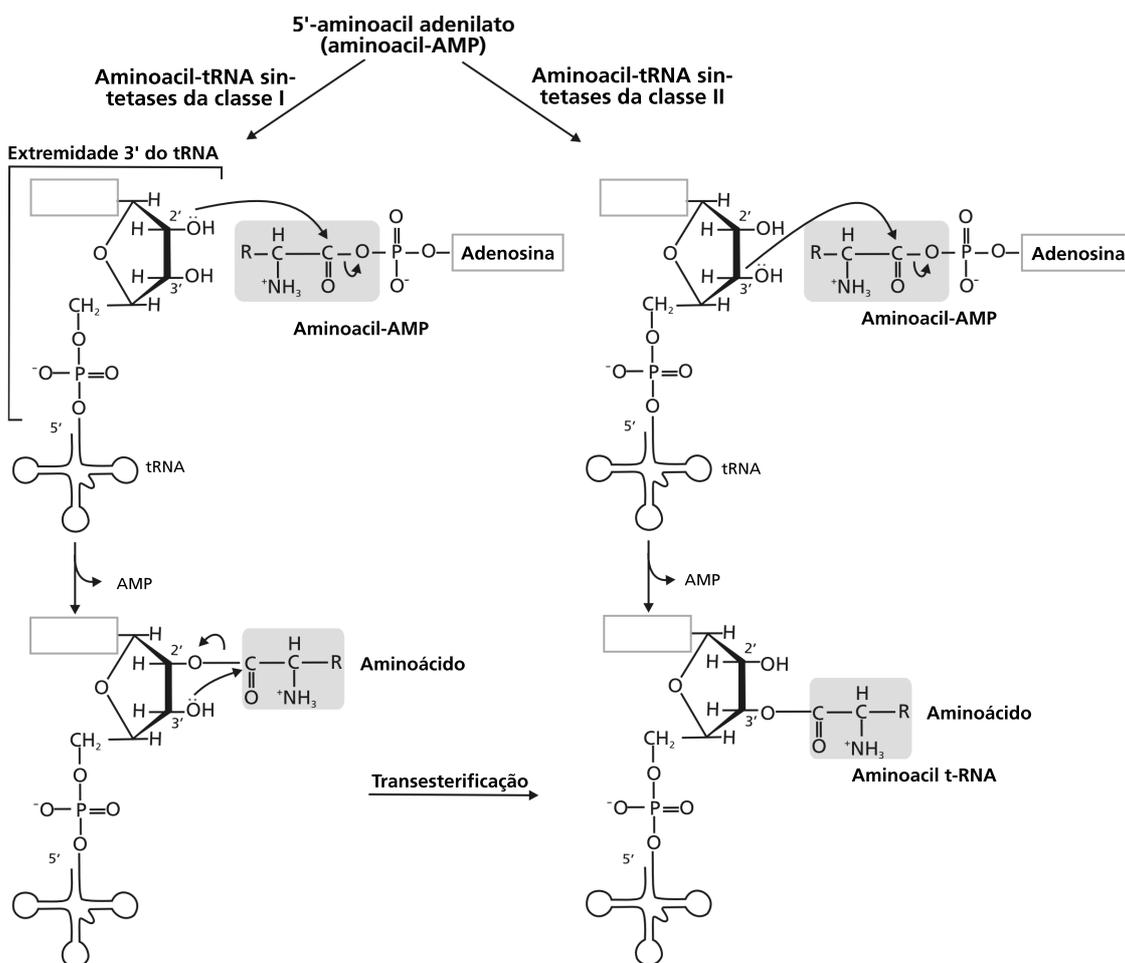
#### AMINOACIL-AMP

Aminoacil-AMP é um termo genérico. De uma forma geral, o sufixo *ina*, presente nos nomes da maioria dos aminoácidos, é substituído por *il* para denominar o composto formado. Por exemplo, o composto formado pela ligação do fenilalanina com o ATP é chamado fenilalanil-AMP. O correspondente da metionina é o metionil-AMP, o da leucina é o leucil-AMP, e assim por diante. A mesma regra se aplica aos termos genéricos aminoacil-tRNA e aminoacil tRNA sintetase.

**Figura 27.11:** Ativação de aminoácido – 1ª etapa. Ligação entre um aminoácido e o ATP, com formação de um aminoacil-AMP. Reação catalisada por aminoacil-tRNA sintetases.



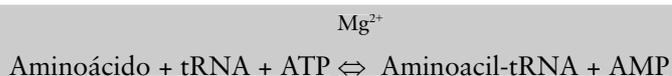
A segunda etapa produz uma ligação rica em energia entre o aminoácido e a ribose do nucleotídeo posicionado na extremidade 3' do tRNA, de forma que, quando este aminoácido é adicionado à cadeia polipeptídica crescente, esta energia é liberada para a formação da ligação peptídica. Confira, também, na **Figura 27.6.a**, que a seqüência de nucleotídeos próxima a essa extremidade é altamente conservada, Pu (A ou G), C, C e A. Portanto, a reação será sempre com um nucleotídeo derivado da . Na **Figura 27.12**, você pode observar que duas classes de enzimas aminoacil-tRNA sintetases podem catalisar a reação a partir do aminoacil-AMP. As enzimas pertencentes às classes I e II catalisam a ligação do aminoácido, respectivamente, ao carbono 2' e ao carbono 3' da ribose. O produto da reação catalisada por enzimas da classe I sofre ainda transesterificação, sendo convertido ao composto resultante da ação de enzimas da classe II.



**Figura 27.12:** Ativação de aminoácido – 2ª etapa. Transferência do grupo aminoacil, presente em um aminoacil-AMP, para o tRNA específico. Reação catalisada por aminoacil-tRNA sintetases, com formação de produtos diferentes de acordo com a classe da enzima.

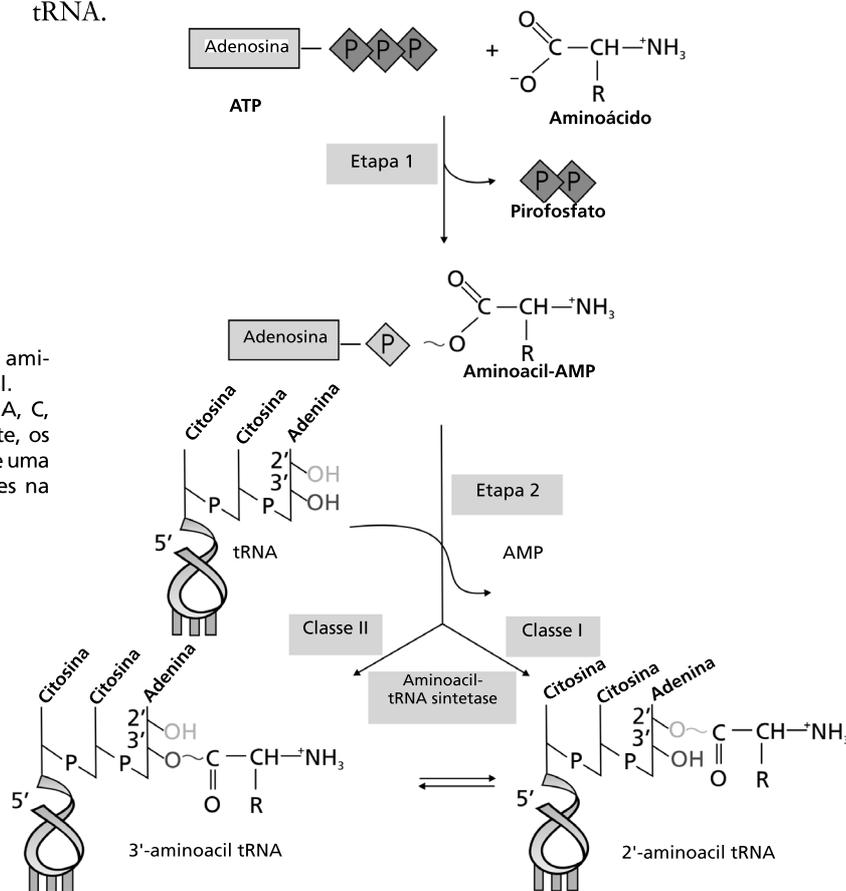
Nos dois esquemas detalhados, você pôde visualizar as fórmulas estruturais dos compostos que participam do processo de ativação de um aminoácido. Na **Figura 27.11**, por exemplo, você encontrou os detalhes do ATP e do aminoácido, que reagem entre si para formar o aminoacil-AMP na primeira etapa de ativação. Já na **Figura 27.12** você pôde observar os detalhes do que ocorre pela ação das classes I e II da enzima aminoacil-tRNA sintetases sobre o aminoacil-AMP.

A reação global catalisada por estas enzimas é:

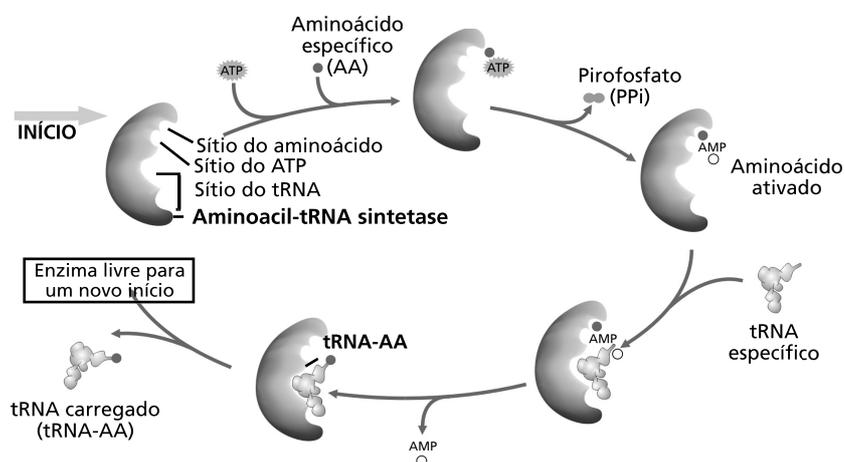


O esquema simplificado apresentado na **Figura 27.13** tem por objetivo facilitar o aprendizado do processo geral. Note que as duas etapas são catalisadas por uma mesma enzima. Portanto, a ação das duas classes de aminoacil-tRNA sintetases, na primeira etapa, resulta em formação de aminoacil-AMP. A partir desse composto, dois grupos de aminoacil-tRNA sintetases, classes I e II, catalisam a reação de ligação do aminoácido, respectivamente, ao C<sub>3'</sub> e ao C<sub>2'</sub> do nucleotídeo A da extremidade 3' do tRNA.

**Figura 27.13:** Ativação de aminoácido – processo global. Na representação do tRNA, C, C e A são, respectivamente, os símbolos de duas citosinas e uma adenina, sempre presentes na extremidade 3'.



Na **Figura 27.14**, você encontra um outro esquema, com destaque para a participação direta da enzima, sua ligação com os substratos e a formação do aminoacil-tRNA. Repare que, ao final do processo, a enzima está livre para reiniciar um novo ciclo de reações.



**Figura 27.14:** Ativação de aminoácido – participação da aminoacil-tRNA sintetase. Destaque para a enzima aminoacil-tRNA sintetase, seus sítios de ligação com o ATP, o aminoácido e o tRNA.

### ATIVIDADE DE REVISÃO DAS AMINOACIL-tRNA SINTETASES

A taxa de erro durante a síntese protéica é de aproximadamente 1 erro/ $10^4$  aminoácidos incorporados, um valor não tão baixo quanto o observado na replicação do DNA. Não parece mesmo necessário uma taxa tão baixa, uma vez que um erro na proteína pode ser “apagado” pela sua destruição, não sendo transmitido para as gerações futuras. Concorda? O que se sabe é que este grau de fidelidade é suficiente para assegurar que a maioria das proteínas não contenha erro e que a grande quantidade de energia despendida na síntese protéica raramente seja desperdiçada.

O que se sabe também é que a capacidade desse grupo de enzimas de se ligar a um aminoácido errado é mais baixa do que seria na verdade o limitado pela energia de ligação disponível derivada de interações entre a enzima e o substrato. Isso se deve ao fato de algumas aminoacil-tRNA sintetases apresentarem uma função revisora. Por exemplo, todo aminoacil-AMP produzido pela Ile-tRNA sintetase (isoleucil-tRNA sintetase) é checado em um segundo sítio ativo da mesma enzima, de forma que os compostos formados incorretamente na primeira etapa do processo sejam hidrolisados. Além disso, grande parte das aminoacil-tRNA sintetases pode hidrolisar a ligação entre aminoácido e tRNA, o que é acelerado em presença de um aminoacil-tRNA incorreto.

A atividade revisora é fundamental para o reconhecimento correto de um aminoácido que apresenta similaridade estrutural com outro(s). Volte às aulas de Bioquímica para rever as estruturas dos aminoácidos. Você verá que, geralmente, vários deles são estruturalmente relacionados. Portanto, parece óbvio que as aminoacil-tRNA sintetases que ativam aminoácidos estruturalmente distintos de qualquer outro não exibam atividade de revisão. Nesses casos, acredita-se que o sítio ativo de aminoacilação possa discriminar entre o seu substrato e um aminoácido incorreto.

Como cada enzima desse grupo tem o potencial para discriminar entre dois substratos diferentes, muitas vezes a interação entre aminoacil-tRNA sintetases e tRNAs é chamada segundo código genético, para ressaltar seu papel crítico na manutenção da acurácia da síntese protéica.

## RESUMO

A síntese protéica é um processo bastante complexo que envolve a ativação dos aminoácidos, a tradução propriamente dita, que ocorre em três estágios, e o processamento pós-traducional dos polipeptídeos recém-sintetizados. A síntese de proteínas, representa um alto custo energético, mas entretanto, a velocidade de síntese de proteína é surpreendentemente alta.

Os ribossomos encontrados no citoplasma representam o principal local de síntese de cadeias polipeptídicas em uma célula. Nesse processo, o mRNA desempenha o papel de molde, uma vez que carrega a informação armazenada no DNA. Em sua estrutura, encontram-se os códons, que especificam os aminoácidos na proteína, e regiões importantes para que o processo se inicie. Já os tRNAs têm a função de transportar os aminoácidos para o local de síntese. Suas duas principais características são: a extremidade 3', na qual um aminoácido específico se liga, e o anticódon, complementar ao códon, no mRNA, que codifica o aminoácido ligado.

A ligação dos aminoácidos aos seus tRNAs cognatos ocorre no citoplasma, sendo catalisada por enzimas denominadas aminoacil-tRNA sintetases. Esse processo ocorre em duas etapas: formação de aminoacil-AMP e ligação do aminoácido ao tRNA.

Algumas enzimas desse grupo exibem atividade de revisão, ou seja, o aminoacil-AMP é checado, o que contribui para que a taxa de incorporação de aminoácidos errados seja baixa o suficiente para garantir que grande parte das proteínas não apresente erro e que o elevado gasto de energia não seja desperdiçado.

## EXERCÍCIOS

1. Quais as etapas envolvidas no processo de síntese de proteínas?
2. Leia com bastante atenção as sentenças a seguir e assinale as verdadeiras (V) e as falsas (F). No caso de considerá-las falsas, reescreva-as corretamente ou explique o porquê.
  - a) ( ) A seqüência de aminoácidos em uma cadeia polipeptídica é determinada pela seqüência de códons no mRNA em um processo chamado transcrição.
  - b) ( ) Uma parte da molécula de tRNA contém um anticódon de duas bases que é complementar ao códon apropriado no mRNA.
  - c) ( ) Cada ribossomo é constituído de duas subunidades, sendo que cada uma delas contém mRNA e muitas proteínas diferentes.
  - d) ( ) Os ribossomos contêm um sítio de ligação do mRNA e três sítios de ligação para tRNA.
3. Cite duas diferenças entre os mRNA procarióticos e eucarióticos importantes para a síntese de proteínas.
4. Qual a importância dos tRNAs para a síntese de proteínas?
5. Descreva resumidamente a etapa de ativação dos aminoácidos.

## AUTO-AVALIAÇÃO

Os exercícios desta aula têm por objetivo ajudar na sedimentação dos principais pontos abordados. Se você não estiver acumulando dúvidas em nossa disciplina, certamente não terá encontrado dificuldades para responder às questões. Lembre-se de que haverá sempre um tutor disponível no pólo para ajudá-lo!

## FIQUE ATENTO!

Esperamos que, ao final desta aula, você se sinta bastante familiarizado com todas as moléculas que serão mencionadas na Aula 28, na qual abordaremos os mecanismos pelos quais a cadeia polipeptídica é sintetizada.



# Fluxo da informação genética – tradução III

AULA

# 28

## objetivos

Nesta aula, você terá a oportunidade de:

- Descrever cada uma das etapas da tradução: iniciação, alongamento e terminação.
- Relatar alguns mecanismos envolvidos no controle da expressão gênica ao nível traducional e alguns antibióticos que interferem na tradução.

### **Pré-requisito**

Para o bom acompanhamento desta aula é fundamental que você tenha entendido muito bem as Aulas 26 e 27.

## INTRODUÇÃO

Na Aula 26, você estudou o código genético que correlaciona os códons no mRNA e os aminoácidos na proteína. Já na Aula 27 você teve a oportunidade de conhecer os principais aspectos gerais da síntese protéica e o processo de ativação dos aminoácidos precursores. Nesta aula, você verá em detalhes como ocorre a tradução propriamente dita, um processo que ocorre nos ribossomos e conta com a participação de inúmeras moléculas, sendo que as principais já foram apresentadas a você nas aulas anteriores. Como a síntese de biomoléculas poliméricas, em geral, a tradução consiste em três estágios: iniciação, alongamento e terminação. Estudaremos detalhadamente cada um deles.

Após o aprendizado da síntese de cadeias polipeptídicas, discutiremos também os principais processos envolvidos no controle da expressão gênica ao nível de tradução.

As três etapas que constituem a tradução propriamente dita são iniciação, alongamento e terminação. Vamos, então, entender cada uma delas!

## INICIAÇÃO

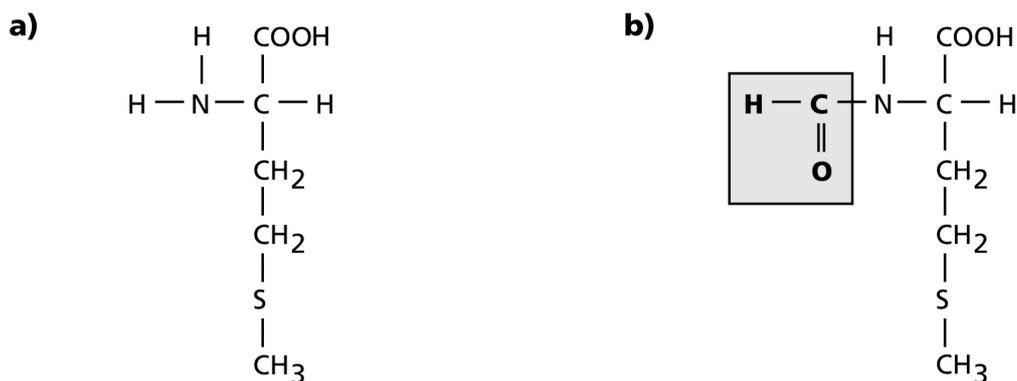
Primeiramente, é importante que você saiba que a tradução de toda proteína se inicia com a incorporação de metionina, modificada ou não. Para isso, é necessária a formação de um complexo, denominado complexo de iniciação de tradução. Esta etapa ocorre de maneira distinta entre os organismos procarióticos e eucarióticos. Entretanto, em linhas gerais, o mRNA, contendo informação para o polipeptídeo a ser sintetizado, se liga à subunidade menor do ribossomo. Em seguida, ocorre a ligação entre o aminoacil-tRNA iniciador e a subunidade ribossomal maior para formar o complexo de iniciação. Este tRNA iniciador ocupa o sítio P no ribossomo (já falamos dele na Aula 27, você se lembra?) e seu pareamento, através do anticódon, com o códon AUG do mRNA sinaliza o começo da síntese da cadeia polipeptídica. Esta etapa inicial requer a hidrólise do GTP e é promovida por proteínas citosólicas específicas chamadas fatores de iniciação – IF (Initiation Factor).

Você deve estar se perguntando quem é este tRNA iniciador, não?

Pois saiba que, embora exista apenas um códon para metionina (AUG), existem dois tRNAs para metionina em todos os organismos. Um deles é exclusivamente usado quando AUG representa o códon de iniciação da síntese protéica. O outro é usado quando a metionina é adicionada em uma posição interna de um polipeptídeo. Vamos conhecer um pouco mais sobre esses tRNAs!

## Aminoacil-tRNA iniciador

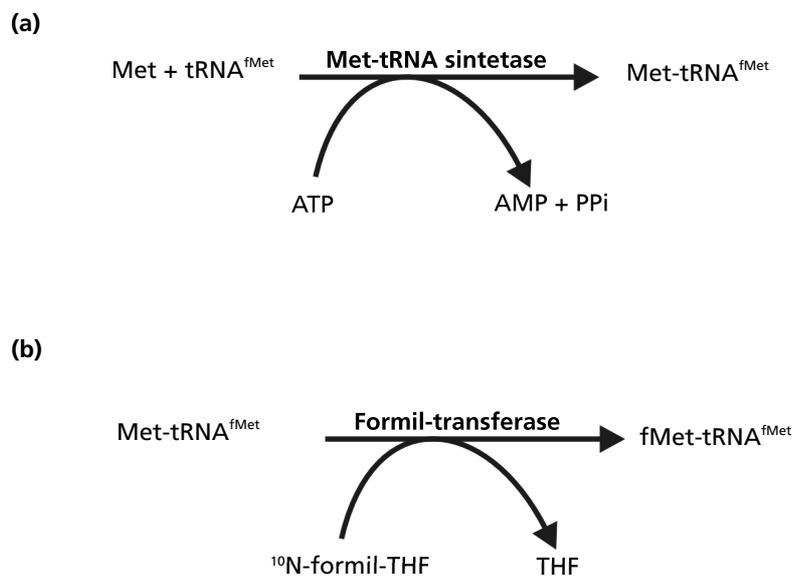
Em bactéria, as duas classes distintas de tRNA específico para metionina são designadas  $\text{tRNA}^{\text{Met}}$  e  $\text{tRNA}^{\text{fMet}}$ . O aminoácido iniciador no terminal amino é N-formil-metionina (fMet). Observe sua fórmula estrutural na **Figura 28.1**.



**Figura 28.1:** Metionina e N-formil-metionina. Fórmulas estruturais da metionina (a), incorporada internamente na cadeia polipeptídica, e da N-formil-metionina (b), que dá início à cadeia polipeptídica.

A entrada do aminoácido iniciador no ribossomo se dá como N-formil-metionil-tRNA<sup>fMet</sup> (fMet-tRNA<sup>fMet</sup>), que é formado em duas reações sucessivas. A primeira envolve a ligação da metionina com o tRNA<sup>fMet</sup> pela ação da metionil-tRNA sintetase, tal qual ocorre na ativação de todos os aminoácidos. Interessante é que a mesma Met-tRNA sintetase que aminoacila tRNA<sup>Met</sup> catalisa a reação com o tRNA<sup>fMet</sup>. Essa etapa está representada na **Figura 28.2.a**.

Em seguida, como pode ser observado na **Figura 28.2.b**, um grupo formil é transferido para o grupo amino da metionina em uma reação catalisada pela enzima transformilase, ou formil-transferase, tendo como doador de formil o <sup>10</sup>N-formil-tetraidrofolato.



**Figura 28.2:** Formação de N-formil-metionil-tRNA<sup>Met</sup> (fMet-tRNA<sup>fMet</sup>). Ligação da metionina ao tRNA<sup>fMet</sup> (a) e, então, adição do grupamento formil ao grupo amino da metionina (b). THF – tetraidrofolato.

Essa reação assegura que o aminoacil-tRNA iniciador não participará das reações de alongamento. A transformilase é mais seletiva que a Met-tRNA sintetase, não sendo capaz de formilar nem metionina livre, nem metionina ligada ao tRNA<sup>Met</sup>. Esta enzima é específica para resíduos de metionina ligados ao tRNA<sup>fMet</sup>, provavelmente por reconhecimento de alguma característica estrutural única desse RNA.

Também é importante ressaltar que o bloqueio do grupo amino (NH<sub>2</sub>) da metionina pelo N-formil serve para impedir sua entrada em posições internas da cadeia, além de permitir que fMet-tRNA<sup>fMet</sup> se ligue ao sítio de iniciação específico no ribossomo, que, aliás, não aceita Met-tRNA<sup>Met</sup> ou qualquer outro aminoacil-tRNA.

Ao contrário do que ocorre em bactérias, em eucariotos todos os polipeptídeos sintetizados pelos ribossomos citosólicos começam com um resíduo de metionina em lugar de N-formil-metionina. Entretanto, também um tRNA iniciador, tRNA<sub>i</sub><sup>Met</sup>, que se difere do tRNA usado para carregar metionina nas posições internas da cadeia polipeptídica, participa dessa etapa. A mesma Met-tRNA sintetase catalisa a reação de ligação de metionina aos dois tRNAs, tRNA<sup>Met</sup> e tRNA<sub>i</sub><sup>Met</sup>.

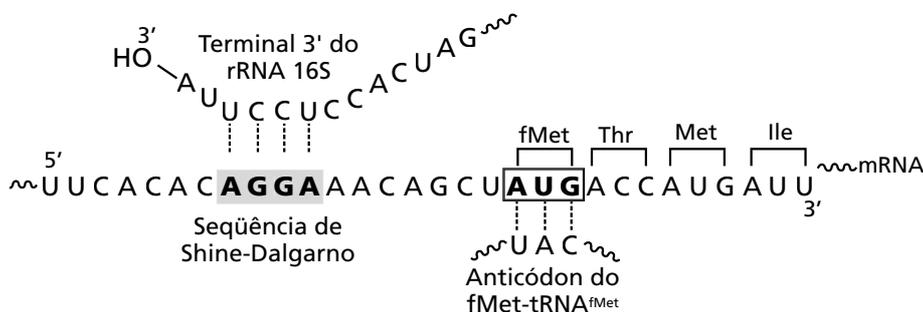
A discriminação funcional entre os dois tRNAs isoceptores é realizada pelos fatores de iniciação e de alongamento. Os fatores de iniciação reconhecem tRNA<sub>i</sub><sup>Met</sup> enquanto os fatores de alongamento reconhecem tRNA<sup>Met</sup>.

De forma distinta dos polipeptídeos sintetizados pelos ribossomos, os sintetizados nas mitocôndrias e nos cloroplastos começam com N-formil-metionina.

### Identificação do códon iniciador

Uma exigência da síntese protéica é selecionar o códon de iniciação apropriado, geralmente AUG, para a tradução. Isto é realizado ao nível do ribossomo pela ligação da unidade ribossomal menor ao mRNA. O reconhecimento do códon de iniciação é realizado de maneiras diferentes em procariotos e eucariotos. Esta diferença no mecanismo de iniciação é responsável pelo fato de que cada mRNA eucariótico, geralmente, codifica só uma proteína, enquanto os mRNAs procarióticos codificam mais de uma.

Em procariotos, o códon AUG iniciador pode ocorrer em qualquer ponto dentro do mRNA e pode existir mais de um sítio de iniciação dentro do mesmo mRNA. Nesses organismos, o códon iniciador AUG é direcionado para a posição correta na subunidade 30S por um sinal de iniciação chamado seqüência Shine-Dalgarno, no mRNA. Lembre que esta seqüência parece com uma região complementar, rica em pirimidina, próxima ao terminal 3' do RNA 16S da subunidade 30S (volte à Aula 27 se julgar necessário). O pareamento de bases entre o mRNA e o rRNA fixa o mRNA, de forma que o AUG é corretamente posicionado para iniciar a tradução. O AUG específico, no qual se liga o fMet-tRNA<sup>fMet</sup>, é, portanto, distinguido dos códons de metionina internos por sua proximidade à seqüência Shine-Dalgarno no mRNA. Observe a **Figura 28.3**, na qual está representada a identificação do códon iniciador em procariotos.



**Figura 28.3:** Identificação do códon iniciador em procariotos. A seqüência de Shine-Dalgarno direciona o códon iniciador AUG para sua posição na subunidade menor do ribossomo.

A importância da sequência Shine-Dalgarno é corroborada pelo fato de que mutações nesta região que prejudicam o pareamento de base entre mRNA e rRNA diminuem também a eficiência de tradução. Em contrapartida, aquelas que favorecem o pareamento de base fazem com que o mRNA seja mais eficientemente traduzido.

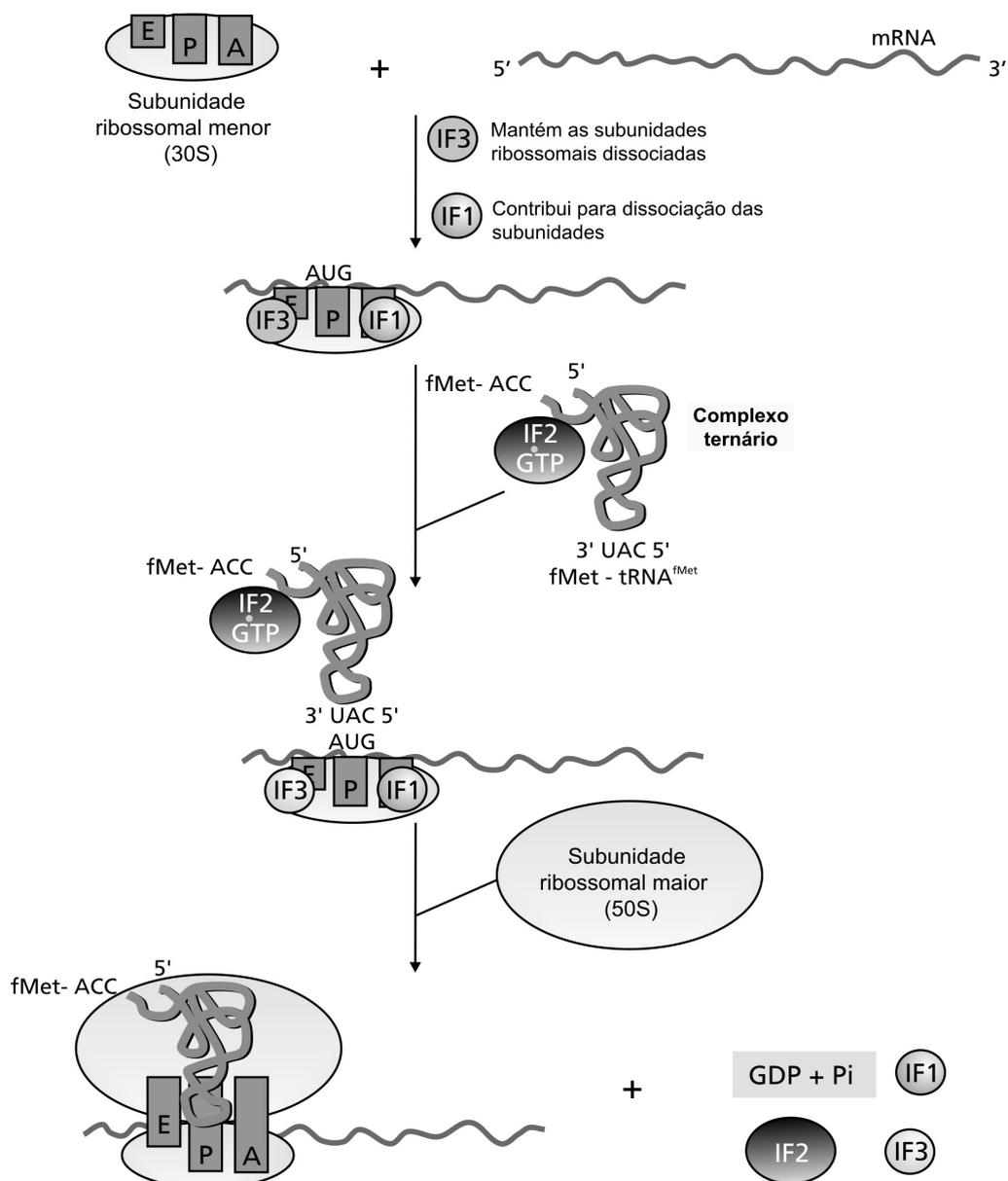
Diferentemente do que ocorre nos organismos procarióticos, em eucariotos, a sequência AUG mais próxima do terminal 5' do mRNA geralmente serve como códon de início para a única proteína codificada por cada mRNA. Nesses organismos, a subunidade ribossomal menor se liga ao terminal 5' do mRNA e se move ao longo da fita até encontrar uma sequência AUG, que é reconhecida pelo pareamento de bases com o anticódon do Met-tRNA<sup>Met</sup>. Em organismos superiores, mas não em levedura, o reconhecimento do AUG iniciador é facilitado por uma sequência de bases flanqueadoras, conhecida como contexto de iniciação. A sequência consenso, ou seja, que normalmente é encontrada, é PurNNAUGG, onde Pur é qualquer purina (A ou G), N é qualquer base e AUG, em negrito, é o códon de iniciação. Sabe-se que a sequência preferida é GCCGCCPurCCAUGG, mas ainda não se sabe como ela é reconhecida. No entanto, sabemos que, quando o contexto de iniciação do primeiro AUG (códon de iniciação) é muito diferente dessa sequência, a subunidade menor, 40S, pode passar por ele e iniciar a tradução no próximo AUG, localizado abaixo do códon de iniciação. Você já pode imaginar o prejuízo para a célula se isso ocorre, não?

### Formação do complexo de iniciação em organismos procarióticos

A formação do complexo de iniciação ocorre em três etapas. Vamos começar discutindo os detalhes do que ocorre em procariotos, nos quais a subunidade ribossomal 30S forma um complexo com os três fatores de iniciação, IF-1, IF-2 e IF-3. Comece observando a função de cada um deles no **Quadro 28.1**. Em seguida, leia os parágrafos seguintes acompanhando o esquema representado na **Figura 28.4**.

**Quadro 28.1:** Fatores de iniciação em bactéria.

Fator	Função
IF-1	Aumenta a taxa de dissociação das duas subunidades ribossomais, provavelmente contribuindo com a ligação do IF-3.
IF-2	Facilita a ligação do fMet-tRNA <sup>fMet</sup> à subunidade ribossomal 30S.
IF-3	Liga-se à subunidade ribossomal 30S. Impede a associação prematura das subunidades 30S e 50S. Aumenta a especificidade do sítio P ao fMet-tRNA <sup>fMet</sup> .



Complexo de iniciação 70S, contendo o mRNA e o iniciador fMet-tRNA<sup>fMet</sup>

**Figura 28.4:** Formação do complexo de iniciação da tradução em procariotos. A explicação do esquema encontra-se no texto.

Inicialmente, ocorre a ligação da subunidade 30S com o fator de iniciação IF-3, o que impede a combinação prematura das subunidades 30S e 50S. A dissociação das subunidades ribossomais é também favorecida pela ligação de IF-1 à porção da subunidade ribossomal menor correspondente ao sítio A. Por fim ocorre a ligação do mRNA

à subunidade 30S, com o posicionamento do códon de iniciação em um local específico da subunidade menor. Lembra da importância da seqüência de Shine-Dalgarno?

As duas subunidades ribossomais, 30S e 50S, contribuem para a característica dos sítios de ligação do aminoacil-tRNA. Entretanto, no estágio inicial de formação do complexo de iniciação da tradução, o AUG iniciador se posiciona na subunidade 30S na parte correspondente ao sítio P, que é o único sítio ao qual fMet-tRNA<sup>fMet</sup> pode se ligar. É importante lembrar que fMet-tRNA<sup>fMet</sup> é uma exceção, pois, como você verá mais adiante nesta aula, durante o estágio de alongamento subsequente, todos os outros aminoacil-tRNAs que chegam, incluindo Met-tRNA<sup>Met</sup>, que se liga aos AUGs internos da fase de leitura do mRNA, ligam-se ao sítio A.

Em seguida, o complexo constituído pela subunidade 30S, por IF-3 e pelo mRNA forma um complexo ainda maior com IF-2 ligado ao GTP e fMet-tRNA<sup>fMet</sup>. Finalmente, antes da liberação de IF-3, esse complexo grande se combina com a subunidade ribossomal 50S, o que estimula a hidrólise de GTP a GDP e Pi pelo IF-2. Essa reação irreversível fornece energia para o rearranjo conformacional das duas subunidades e libera IF-1 e IF-3 para participação em outras reações de iniciação.

Em bactéria, o resultado dessa etapa é um ribossomo 70S funcional chamado complexo de iniciação, contendo o mRNA e o iniciador fMet-tRNA<sup>fMet</sup>. A correta ligação do fMet-tRNA<sup>fMet</sup> ao sítio P no complexo de iniciação 70S completo é assegurada por dois pontos de reconhecimento e ligação: interação entre códon do mRNA e anticódon do tRNA, envolvendo o AUG iniciador fixado no sítio P, e interações entre o sítio P e o fMet-tRNA<sup>fMet</sup>.

Resumindo, a iniciação resulta na formação de um complexo fMet-tRNA<sup>fMet</sup>•mRNA•ribossomo em que fMet-tRNA<sup>fMet</sup> ocupa o sítio P do ribossomo, enquanto o sítio A está posicionado para receber o aminoacil-tRNA carregando o aminoácido seguinte da cadeia.

### **Formação do complexo de iniciação em organismos eucarióticos**

Como em organismos procarióticos, o início da tradução em eucariotos se dá com as duas subunidades ribossomais separadas, o que é mantido por um fator antiassociação, e termina com a hidrólise de GTP e

associação da subunidade maior. Entretanto, a etapa inicial de formação do complexo de iniciação é diferente em eucariotos provavelmente devido às características do mRNA. Estudaremos essa etapa em detalhes mais adiante. Uma outra diferença entre procariotos e eucariotos é a existência de um número maior de fatores de iniciação eucarióticos. Alguns desses fatores e suas funções estão listados no **Quadro 28.2**.

**Quadro 28.2:** Fatores de iniciação em eucariotos.

Fator*	Função
eIF-2	Facilita a ligação do Met-tRNA <sub>i</sub> <sup>Met</sup> à subunidade ribossomal 40S.
eIF-2B, eIF-3	Primeiros fatores que se ligam à subunidade 40S, facilitando as etapas subsequentes.
eIF-4A	A atividade RNA helicase remove a estrutura secundária no mRNA, permitindo a ligação da subunidade 40S. Faz parte do complexo eIF-4F.
eIF-4B	Liga-se ao mRNA facilitando o movimento da subunidade 40S ao longo do mRNA para localizar o primeiro AUG.
eIF-4E	Liga-se ao 5' cap do mRNA. É parte integrante do complexo eIF-4F.
eIF-4G	Liga-se ao eIF-4E e à proteína de ligação à cauda poli-A. Faz parte do complexo eIF-4F.
eIF-5	Promove a dissociação de outros fatores de iniciação e da subunidade 40S como um início para a associação da subunidade 60S para formar o complexo de iniciação 80S.
eIF-6	Facilita a dissociação do ribossomo 80S inativo nas subunidades 40S e 60S.

\* O prefixo "e" identifica esses fatores como eucarióticos.

Para um melhor entendimento do processo de formação do complexo de iniciação em eucarioto, leia o texto a seguir acompanhando a **Figura 28.5**.

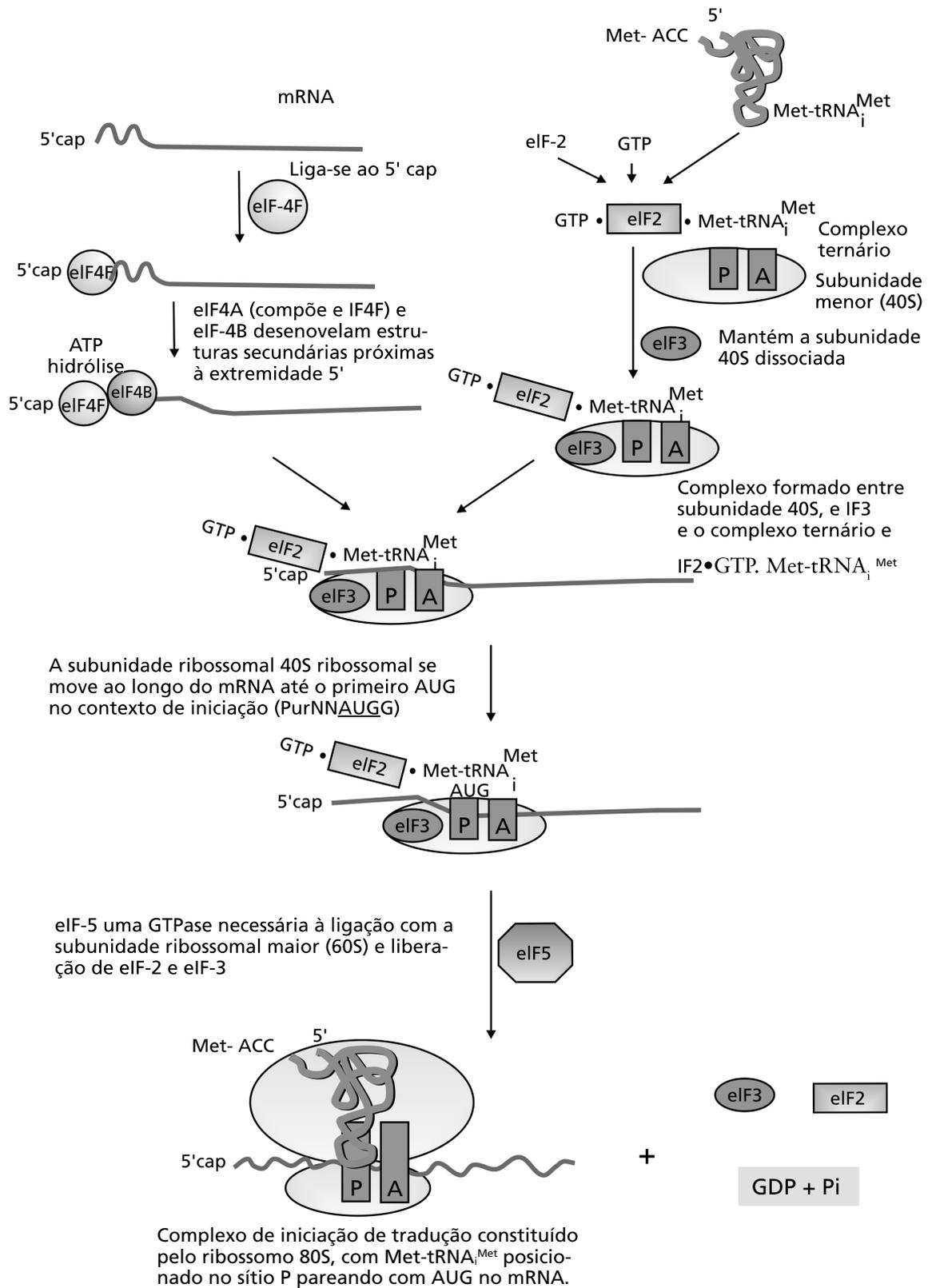


Figura 28.5: Formação do complexo de iniciação da tradução em eucariotos. A explicação do esquema encontra-se no texto.

O processo se inicia com o complexo eIF-4F, formado pelas proteínas eIF-4E, eIF-4G e eIF-4A, ligando-se à extremidade 5' cap do mRNA e promovendo o desenovelamento de estruturas secundárias, uma vez que eIF-4A exibe atividade de RNA helicase. A ligação de eIF-4B contribui ainda mais para o desenovelamento.

Paralelamente, o fator eIF-3, necessário para manter a subunidade 40S dissociada, promove sua ligação, através da região correspondente ao sítio P, com o complexo ternário constituído do primeiro Met-tRNA<sub>i</sub><sup>Met</sup> e eIF-2 ligado ao GTP. O complexo maior formado pela associação da subunidade 40S e eIF-2•GTP•Met-tRNA<sub>i</sub><sup>Met</sup> se liga, então, ao terminal 5' do mRNA com o auxílio de vários fatores. Um deles, eIF-4E, também chamado proteína de ligação ou CBP (Cap Binding Protein) e que faz parte do complexo eIF-4F, liga-se especificamente ao terminal 5' cap do mRNA, mas não está representado na figura. O mRNA é, então, explorado para localização do primeiro AUG que sinaliza o início da fase de leitura. Lembre que, em eucarioto, o mRNA não tem as seqüências complementares para se ligar ao rRNA 16S, tal qual a seqüência de Shine-Dalgarno, mas sim o contexto de iniciação.

Em resumo, nessas duas primeiras etapas, a ligação do mRNA à subunidade menor e a exploração do primeiro códon posiciona a subunidade 40S do ribossomo no AUG iniciador na seqüência apropriada de nucleotídeos flanqueadores.

Na última etapa, com a hidrólise de GTP a GDP e Pi pela ação de GTPase (eIF-5), ocorre a montagem do complexo de iniciação completo, constituído do ribossomo 80S ligado ao Met-tRNA<sub>i</sub><sup>Met</sup> e ao mRNA, que se encontra pronto para dar início à etapa de alongamento de cadeia.

Os vários fatores de iniciação adicionais não mencionados no texto e nem representados na figura são necessários para a etapa de exploração do mRNA e para a montagem do ribossomo 80S.

## ALONGAMENTO

Ao final do processo de iniciação, o ribossomo começa a tradução da fase de leitura associada ao códon de iniciação. A tradução dos códons seqüenciais no mRNA é realizada pela repetição seqüencial de três reações para cada aminoácido. Dos três estágios do alongamento, que são similares tanto em procariotos como em eucariotos, apenas dois requerem proteínas não ribossomais conhecidas como fatores de

alongamento – Elongation Factor (EF). A formação da ligação peptídica não requer nenhum fator e é a única reação da síntese protéica catalisada pelo próprio ribossomo.

#### Direção da “leitura” do mRNA e da síntese do polipeptídeo

Traçando um paralelo com a replicação e a transcrição – que ocorrem em uma direção determinada, ou seja, de 5' (terminal fosfato livre) para 3' (terminal hidroxila livre) – um polipeptídeo é sempre sintetizado na direção do amino terminal (-NH<sub>2</sub>) para o carboxi terminal (-COOH). Isto ocorre pela adição seqüencial de aminoácidos ao terminal carboxila da cadeia polipeptídica crescente. A seqüência de aminoácidos, por sua vez, é especificada pela seqüência de códons no mRNA, que sempre é “lido” na direção 5'→3'.

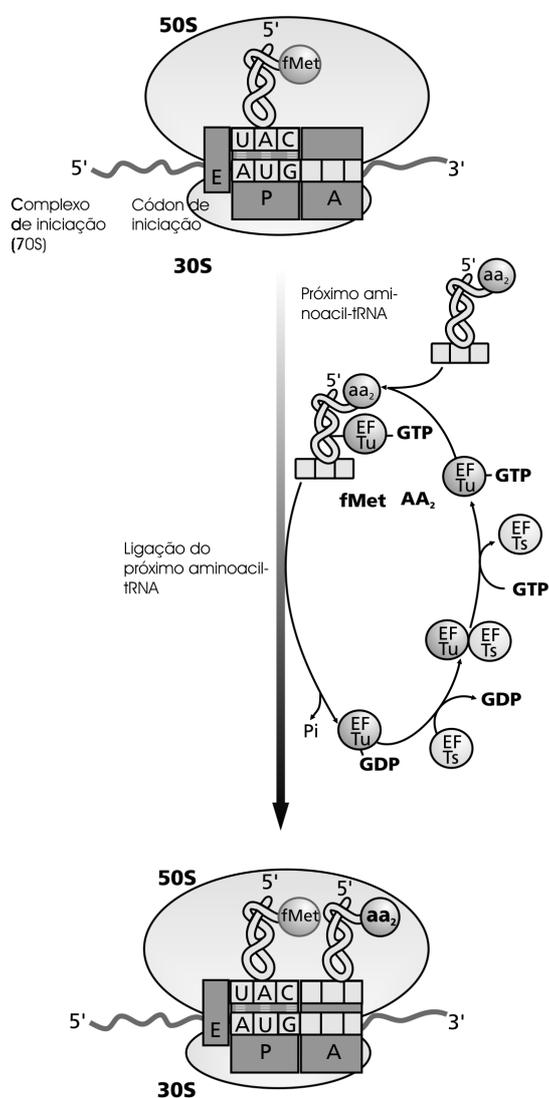
Em linhas gerais, a extensão da cadeia polipeptídica é devida a sucessivas ligações covalentes de unidades de aminoácidos, cada uma levada ao ribossomo e corretamente posicionada por seu tRNA, que parecia com seu códon correspondente no mRNA. O alongamento da cadeia se inicia com a ligação de um aminoacil-tRNA ao sítio A do ribossomo. Uma ligação peptídica, então, se forma entre o grupo amino do aminoacil-tRNA a ser incorporado e o grupo carboxil da formil-metionina carregada pelo tRNA iniciador. Ocorre translocação do ribossomo ao longo do mRNA, de forma a mover o dipeptidil-tRNA resultante do sítio A para o sítio P e promover a liberação do tRNA iniciador. Um novo aminoacil-tRNA se liga ao sítio A vazio para começar um outro ciclo de extensão de cadeia, que prossegue como já descrito. A ligação de cada aminoacil-tRNA que chega ao ribossomo, o movimento do peptidil-tRNA do sítio A para o sítio P e o movimento associado do mRNA, em relação ao ribossomo, para o códon seguinte são promovidos pela hidrólise de duas moléculas de GTP para cada resíduo adicionado à cadeia polipeptídica crescente. Você pode imaginar o alto custo energético que esse processo representa para a célula, não?

Sintetizar uma seqüência definida de aminoácidos requer um alto gasto energético para a célula. Para você ter uma idéia, a síntese de um polipeptídeo que apresenta um total de  $n$  resíduos requer a hidrólise de, no mínimo,  $4n$  fosfatos ricos em energia, tais como ATP ou GTP, o que corresponde à utilização de até 90% da energia química gasta por uma célula para todas as reações biossintéticas.

Chegou a hora de discutirmos o processo com a participação dos fatores de alongamento. Discutiremos o processo que ocorre em procarioto e comentaremos as devidas diferenças em relação aos eucariotos. Para isso, leia o texto acompanhando as **Figuras 28.6, 28.7 e 28.8**, que correspondem, respectivamente, à ligação do próximo aminoacil-tRNA ao sítio A, à formação da ligação peptídica e à translocação.

No primeiro estágio do ciclo de alongamento, representado na **Figura 28.6**, o próximo aminoacil-tRNA especificado pelo códon imediatamente adjacente ao códon iniciador liga-se primeiramente ao complexo formado por EF-Tu (EF-1 $\alpha$ , em eucarioto) e GTP. Esse complexo ternário se liga ao ribossomo e, com a hidrólise de GTP a GDP e Pi, o aminoacil-tRNA se liga ao sítio A, através da interação entre códon (mRNA) e anticódon (tRNA), e EF-Tu•GDP é liberado do ribossomo 70S. O complexo EF-Tu•GTP é regenerado com a participação do fator de alongamento EF-Ts. Primeiramente, o fator EF-Ts desloca o GDP do complexo EF-Tu•GDP, formando EF-Tu•EF-Ts. Em seguida, uma nova molécula de GTP desloca o fator EF-Ts, havendo a formação de EF-Tu•GTP, pronto para se ligar a um novo aminoacil-tRNA.

Aminoacil-tRNAs podem se ligar ao sítio A sem a mediação de EF-Tu, mas a uma taxa muito baixa para suportar o crescimento celular. Devido ao grande número de cópias da proteína EF-Tu em *E. coli*, as moléculas de aminoacil-tRNAs da célula são quase que inteiramente seqüestradas por EF-Tu. Além disso, fator EF-Tu não se liga nem ao fMet-tRNA<sup>fMet</sup> formilado nem ao não-formilado, por isso o tRNA iniciador nunca “lê” códons AUG ou GUG internos.

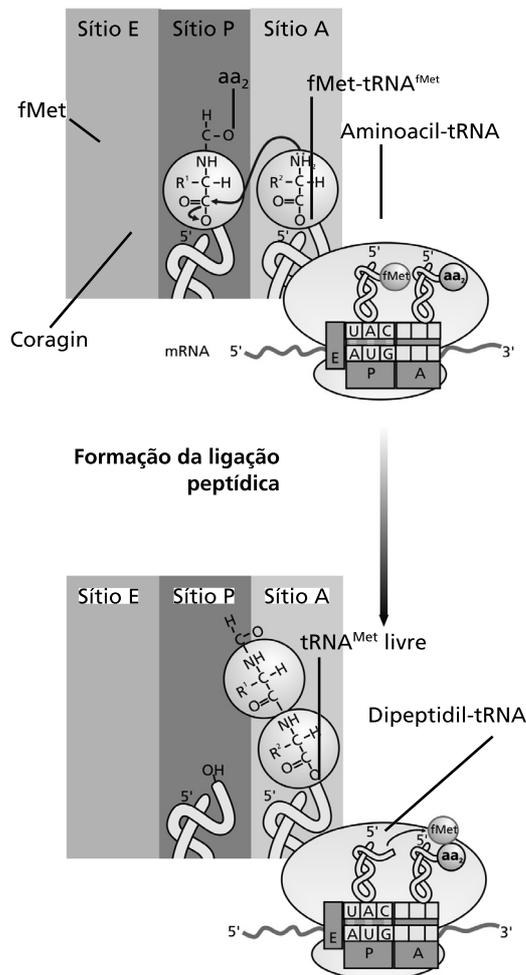


**Figura 28.6:** Alongamento da cadeia polipeptídica – ligação do segundo aminoacil-tRNA. A explicação do esquema se encontra no texto.

Por tudo que já foi mencionado, é importante enfatizar que o tRNA posicionado no sítio A deve ter o anticódon correto para parear com o códon do mRNA também posicionado no sítio A. A ligação do tRNA causa uma mudança conformacional da subunidade ribossomal menor, que favorece a interação entre o rRNA 16S com as duas primeiras bases do complexo formado pelo códon e o anticódon, assegurando que somente o tRNA possa se ligar. Caso a conformação do ribossomo não seja estabilizada pela interação do tRNA no sítio A, ocorre liberação do aminoacil-tRNA antes da formação da ligação peptídica, o que é mediado pelo fator EF-Tu.

Os dois complexos EF-Tu•GTP e EF-Tu•GDP levam alguns segundos para se dissociar. Esse tempo é fundamental para que seja feita a revisão nas

interações entre códon e anticódon, de tal maneira que um aminoacil-tRNA incorreto normalmente se dissocia do sítio A durante este período. Esse mecanismo de revisão apenas confirma se o pareamento entre códon e anticódon está correto, não sendo capaz de identificar aminoácidos incorretos ligados aos tRNAs posicionados no sítio A. Podemos dizer, então, que o fator EF-Tu contribui para a taxa e a fidelidade do processo global de síntese.

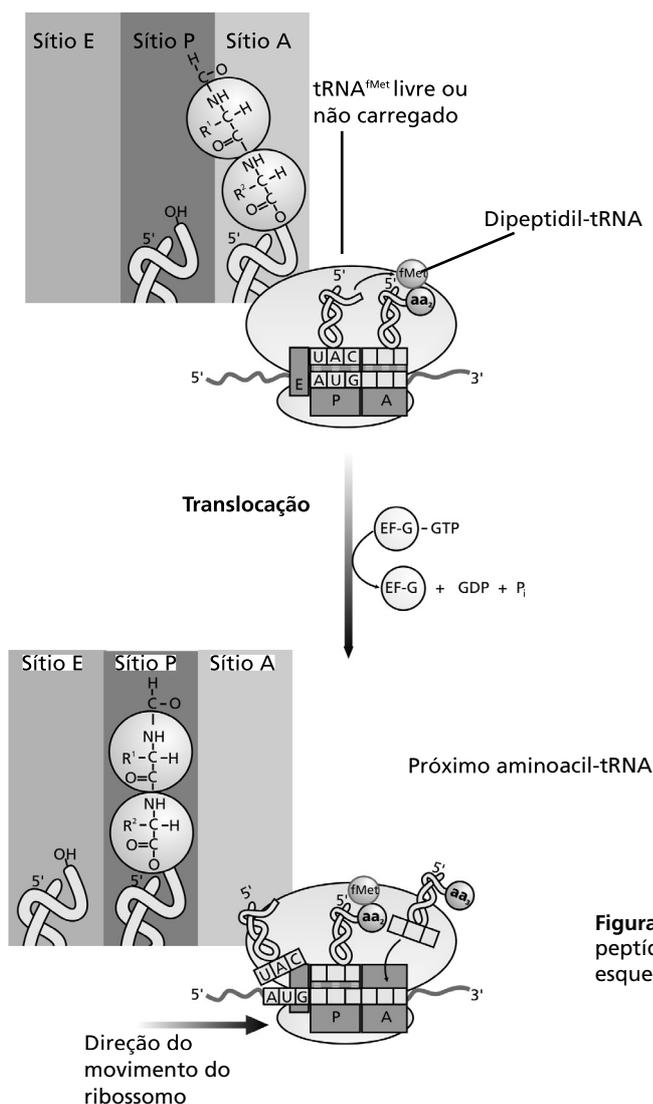


No segundo estágio, representado na **Figura 28.7**, forma-se uma ligação peptídica entre os aminoácidos ligados por seus tRNAs aos sítios A e P do ribossomo. O que ocorre, de fato, é a transferência do grupo N-formil-metionil do seu tRNA, localizado no sítio P, para o grupo amino do segundo aminoácido da cadeia, que se apresenta ligado ao tRNA posicionado no sítio A. Essa reação tem como produto um dipeptidil-tRNA, no sítio A, e o tRNA<sup>Met</sup> livre ou não carregado, no sítio P. Uma das funções da ligação do aminoácido correto ao seu tRNA é ativar o aminoácido, tornando a formação da ligação peptídica, durante a síntese de proteínas, no ribossomo, energeticamente favorável.

**Figura 28.7:** Alongamento da cadeia polipeptídica – formação da ligação peptídica. A explicação do esquema se encontra no texto.

Até pouco tempo atrás, acreditava-se que esta reação fosse catalisada por uma enzima componente do ribossomo, ao qual se denominou peptidil transferase. Entretanto, existem fortes evidências de que, de fato, a reação seja catalisada pelo rRNA 23S, que compõe a subunidade ribossomal maior. Este seria mais um dos casos de moléculas de RNA exibindo atividade catalítica.

O estágio final do alongamento, representado na **Figura 28.8**, consiste na translocação, que é simplesmente o movimento do ribossomo de um códon em direção à extremidade 3' do mRNA. Isso resulta em mudança de posição do anticódon do dipeptidil-tRNA, ainda ligado ao segundo códon do mRNA, do sítio A para o sítio P e do tRNA<sup>fMet</sup> livre do sítio P para o sítio E, ocorrendo, então, liberação do tRNA<sup>fMet</sup> livre para o citosol. Nesse momento, o terceiro códon do mRNA está posicionado no sítio A e o segundo códon está localizado no sítio P.



**Figura 28.8:** Alongamento da cadeia polipeptídica – translocação. A explicação do esquema se encontra no texto.

O movimento do ribossomo ao longo do mRNA requer o fator de alongamento EF-G, denominado translocase, e a energia proveniente da hidrólise de GTP a GDP e Pi, sendo favorecido pela mudança de conformação do ribossomo inteiro. Repare, ainda, na **Figura 28.9**, a similaridade entre o complexo EF-Tu•tRNA e o fator EF-G, sugerindo uma possível ligação do EF-G ao sítio A deslocando o peptidil-tRNA.



**Figura 28.9:** Estruturas tridimensionais do complexo EF-T•tRNA•GTP e do fator EF-G.

Após esse primeiro ciclo de eventos associados à extensão de cadeia, o ribossomo está pronto para iniciar um novo ciclo para incorporação do terceiro aminoácido. Essa etapa, como as subseqüentes que envolverão a ligação do quarto, quinto, sexto até o último aminoácido da cadeia polipeptídica, ocorre tal qual descrito anteriormente e representado nas **Figuras 28.6, 28.7 e 28.8**. A diferença é que, a partir do segundo ciclo, o sítio P será sempre ocupado por um peptidil-tRNA e o sítio A receberá sempre o aminoacil-tRNA trazendo o aminoácido a ser incorporado na cadeia.

O ciclo de eventos de alongamento em eucariotos difere muito pouco do que se observa em procariotos e que você acabou de estudar. Uma dessas diferenças são os fatores eucarióticos, EF-1 $\alpha$ , EF-1 $\beta\gamma$  e EF-2, que têm funções análogas às dos fatores procarióticos EF-Tu, EF-Ts e Ef-G, respectivamente. Além disso, pelo fato de os ribossomos eucarióticos não apresentarem sítio E, os tRNAs não carregados são liberados diretamente do sítio P.

## TERMINAÇÃO

A etapa de alongamento é um processo cíclico que se repete até que o ribossomo adiciona o último aminoácido codificado pelo mRNA. A terminação da cadeia polipeptídica é sinalizada por um códon de terminação, UAA, UAG ou UGA, localizado imediatamente após o códon que codifica o último aminoácido. A cadeia polipeptídica é, então, liberada do ribossomo, com o auxílio de proteínas chamadas fatores de liberação – Release Factor (RF). Vamos conhecer como isso acontece nas bactérias, lendo o texto e buscando o esquema representado na **Figura 28.10**.

Quando um códon de terminação ocupa o sítio A, três fatores de terminação procarióticos, RF1, RF2 e RF3, promovem três eventos distintos: a hidrólise da ligação do polipeptidil-tRNA; liberação, do sítio P, do polipeptídeo livre e do último tRNA não carregado; dissociação do ribossomo 70S em suas subunidades 30S e 50S, que podem, então, dar início a uma nova síntese.

Apesar de não se conhecer a verdadeira função de RF3, possivelmente esse fator está associado à dissociação do ribossomo. Já os fatores RF1 e RF2 reconhecem códon de terminação distintos, UAG/UAA e UGA/UAA, respectivamente. Quando um desses fatores protéicos se liga ao códon de terminação, a atividade de peptidil transferase induz a transferência do polipeptídeo recém-sintetizado para uma molécula de água, em lugar de um novo aminoácido. Curiosamente, em eucariotos, um único fator de terminação, denominado eRF, reconhece os três códon de terminação.

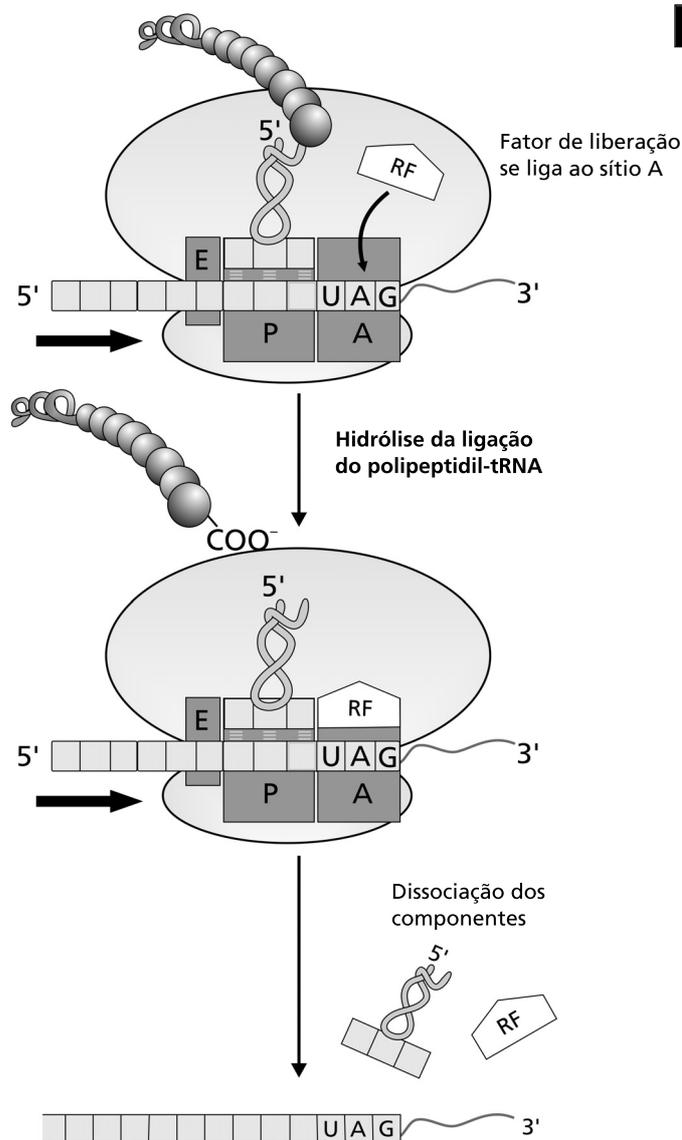


Figura 28.10: Terminação da cadeia polipeptídica. A explicação do esquema se encontra no texto.

## REGULAÇÃO DA TRADUÇÃO E AÇÃO DE ANTIBIÓTICOS

A síntese de milhares de proteínas diferentes em cada célula é regulada de tal forma que somente o número necessário de moléculas de cada uma delas é produzido em qualquer circunstância metabólica. Vale ressaltar que, para manter a mistura apropriada e a concentração de proteínas em

uma célula, os processos de endereçamento e degradação, que serão estudados na Aula 29, devem ocorrer de forma coordenada com a síntese. Qualquer descontrole em um desses três processos pode resultar em comprometimento das funções vitais da célula.

Com relação à regulação da síntese protéica, você já estudou, nas aulas do Módulo 3, que o principal mecanismo de controle de expressão gênica é ao nível transcricional, ou seja, ocorre na transcrição. Entretanto, esse controle também pode ocorrer ao nível pós-transcricional.

A estabilidade dos mRNAs, por exemplo, é bastante variável, com a região 3' não traduzida exercendo um papel importante na estabilidade dessas moléculas. Ocorre que o equilíbrio entre a degradação e a síntese de mRNA é o fator determinante para a quantidade de mRNAs individuais nas células, que, por sua vez, determina a taxa de síntese das proteínas.

A regulação da tradução geralmente é feita no estágio inicial. Uma das formas de controle está associada à estrutura terciária do mRNA, que pode prevenir sua ligação à subunidade ribossomal. A presença de proteínas que podem se ligar ao mRNA, bloqueando a iniciação, também é importante para o controle da tradução. Como exemplo podemos citar proteínas que se ligam à seqüência de Shine-Dalgarno, em mRNAs procarióticos, e a presença de um **RNA ANTI-SENSO** capaz de impedir a formação do complexo de iniciação.

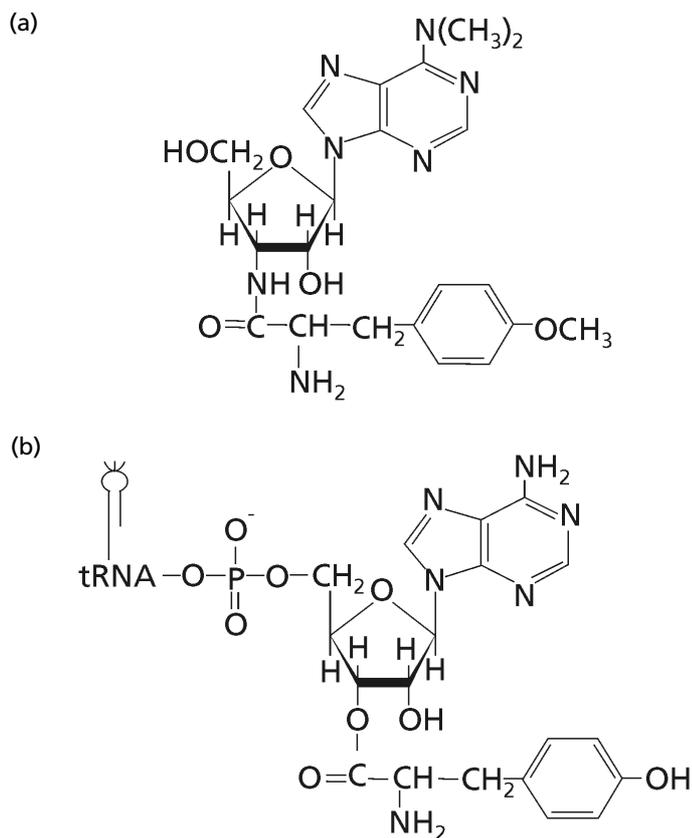
#### **RNA ANTI-SENSO**

O RNA anti-senso apresenta seqüências complementares à região de um mRNA que contém um códon de iniciação, podendo parear com essa região e impedir o início da tradução. Esse mecanismo de controle ao nível da tradução é comum em bactérias e envolve o bloqueio do reconhecimento do códon de iniciação e da ligação da subunidade ribossomal 30S à seqüência de Shine-Dalgarno.

Alguns fatores podem interferir diretamente no processo de tradução, destacando-se vários antibióticos e toxinas comumente encontrados na natureza. Os antibióticos, na sua grande maioria, são inibidores de síntese protéica em bactérias. As pequenas diferenças observadas entre a síntese de proteínas em eucarioto e procarioto são suficientes para que a maioria dos antibióticos não exerça qualquer efeito em células eucarióticas.

O bloqueio da tradução por alguns antibióticos é observado em diferentes estágios desse processo. A estreptomicina, por exemplo, liga-se à subunidade menor do ribossomo bacteriano, impedindo a formação do complexo de iniciação. As tetraciclinas inibem a síntese protéica por se ligarem reversivelmente à subunidade 30S do ribossomo de bactérias sensíveis, impedindo, assim, a ligação do aminoacil-tRNA. Já o cloranfenicol, pertencente a um grupo de antibióticos que se ligam à subunidade 50S do ribossomo bacteriano, impede a ação da

peptidil-transferase, enquanto a eritromicina bloqueia a translocação. A puromicina, por sua vez, apresenta estrutura similar ao tirosinil-tRNA (Figura 28.11) e, por isso, incorpora-se na cadeia polipeptídica crescente causando uma terminação prematura.



**Figura 28.11:** Estruturas comparativas do antibiótico puromicina (a) e a porção da extremidade 3' do tirosinil-tRNA (b).

## RESUMO

A primeira etapa da tradução consiste em formação do complexo de iniciação. Esse processo ocorre de forma diferente nos organismos procarióticos e eucarióticos, provavelmente devido às diferenças entre os mRNAs. Entretanto, tanto em procarioto quanto em eucarioto, o mRNA se liga à subunidade ribossomal menor e ao aminoacil-tRNA iniciador e, em seguida, ocorre a ligação da subunidade ribossomal maior. Inúmeros fatores protéicos, conhecidos como fatores de iniciação, participam dessa etapa.

A etapa de alongamento se caracteriza por três eventos: ligação de um aminoacil-tRNA ao sítio A, formação da ligação peptídica entre o aminoácido ligado a este aminoacil-tRNA e o peptídeo presente no peptidil-tRNA localizado no sítio P e translocação, que posiciona o sítio A no próximo códon para dar início a um novo ciclo de eventos.

A etapa de terminação conta com a participação de fatores protéicos que, ao reconhecerem um códon de terminação no sítio A, promovem a hidrólise da ligação do peptidil-tRNA, a liberação do polipeptídeo recém-sintetizado e a dissociação das duas subunidades ribossomais.

O estágio inicial da tradução é o principal alvo de controle desse processo. Diferentes estruturas tridimensionais que podem ser assumidas pelo mRNA e a presença de proteínas ou de RNAs anti-senso que se ligam ao mRNA podem impedir a formação do complexo de iniciação de tradução.

O princípio de ação de diversos antibióticos é justamente impedir a tradução nas bactérias resistentes, o que pode ser feito através do bloqueio de várias etapas do processo.

## EXERCÍCIOS

1. Descreva resumidamente a diferença entre o processo de formação do complexo de iniciação em procariotos e eucariotos.
2. Marque V (verdadeiro) ou F (falso). Durante a etapa de iniciação da tradução:
  - a) ( ) metionil-tRNA aparece no sítio A do complexo de iniciação 80S.
  - b) ( ) eIF-3 e a subunidade ribossomal 40S participam da formação do complexo de iniciação.

c) ( ) o mesmo metionil-tRNA é usado nas etapas de iniciação e de alongamento de cadeia.

d) ( ) um complexo constituído de mRNA, subunidade ribossomal 60S e certos fatores de iniciação é formado.

3. Qual dos itens a seguir não é necessário para a síntese de proteínas em eucariotos?

a) mRNA

b) Ribossomos

c) GTP

d) 20 aminoácidos diferentes sob a forma de aminoacil-tRNAs

e) fMet-tRNA<sub>i</sub><sup>Met</sup>

4. Marque V (verdadeiro) ou F (falso). Durante a etapa de alongamento de cadeia em eucarioto:

a) ( ) o aminoacil-tRNA que chega ao ribossomo se liga ao sítio P.

b) ( ) uma nova ligação peptídica é formada pelo sítio de peptidil transferase na subunidade ribossomal maior em uma reação GTP-dependente.

c) ( ) o peptídeo, ainda ligado à molécula de tRNA, é translocado para um sítio diferente no ribossomo.

d) ( ) estreptomicina pode causar liberação prematura do peptídeo incompleto.

5. Por que o antibiótico puromicina interfere na síntese de proteína?

### AUTO-AVALIAÇÃO

Faça as questões de múltipla escolha lendo e relendo cuidadosamente as alternativas. Dessa forma, você poderá usufruir muito mais desse tipo de exercício. Em relação às outras questões, você deve retornar à aula para rever os pontos que porventura ainda não foram bem entendidos. E não se esqueça – sempre haverá um tutor disponível no pólo para ajudá-lo!

**FIQUE ATENTO!**

Na Aula 29, você estudará como os polipeptídeos recém-sintetizados são processados para gerar uma molécula funcional e como ocorre seu endereçamento para o local onde deverá desempenhar sua função na célula.

# Processamento e endereçamento de proteínas

AULA

# 29

## objetivos

Nesta aula, você a oportunidade de:

- Reconhecer a importância do processamento de proteínas.
- Descrever algumas das modificações pós-traducionais.
- Reconhecer a importância dos processos envolvidos no endereçamento de proteínas.
- Descrever o endereçamento de proteínas para o retículo endoplasmático.

### **Pré-requisitos**

Esta aula será mais facilmente assimilada se você não tiver dúvidas em relação às Aulas 26, 27 e 28. Provavelmente você também terá de rever alguns conceitos de Biologia Celular.

## INTRODUÇÃO

Para obter sua forma biologicamente ativa, o polipeptídeo deve se dobrar em sua conformação tridimensional própria, o que é chamado enovelamento. O novo polipeptídeo pode sofrer ainda processamento enzimático. Um fato interessante é que muitas vezes esta etapa se inicia durante a síntese do polipeptídeo e só termina durante o transporte da proteína para o seu local de ação. Além disso, esse processamento ocorre de acordo com o destino final dessas proteínas, que pode ser o interior ou o exterior da célula, no qual elas desempenharão suas funções. Por isso, muitas vezes, ao falarmos do processamento de uma proteína, já estaremos abordando o seu endereçamento. Se estivermos tratando de endereçamento, por vezes teremos de comentar as modificações que um polipeptídeo sofre durante o caminho para o seu destino final na célula. Vamos entender como tudo isso ocorre!

## QUAL A IMPORTÂNCIA DO PROCESSAMENTO DE PROTEÍNAS?

Você já estudou que apenas os vinte aminoácidos são especificados pelo código genético, certo? Entretanto, a análise química de inúmeras proteínas funcionais revela centenas de aminoácidos diferentes, sendo todos variantes estruturais dos vinte aminoácidos-padrão. Como isso é possível?

Essa diversidade estrutural é resultado de modificações pós-traducionais das cadeias polipeptídicas sintetizadas durante o processo de tradução, que, como você sabe, ocorre nos ribossomos. Não só as cadeias laterais dos aminoácidos se apresentam modificadas, como também, inúmeras vezes, partes das cadeias polipeptídicas são removidas durante o processamento das proteínas. Em linhas gerais, essa etapa se destina à remoção de um ou mais aminoácidos do terminal amino, à adição de grupamentos, tais como acetil, fosfato, metil e carboxil, a certos resíduos de aminoácidos, à sua clivagem proteolítica ou à ligação de oligossacarídeos ou grupos prostéticos. Você deve se lembrar desses termos, pois eles já foram estudados nas aulas de Bioquímica!

Será que o tipo de processamento é o mesmo para todas as proteínas? Parece lógico que não seja, concorda? É importante que você saiba que as modificações pós-traducionais são determinadas por informações presentes na própria estrutura da proteína. Nesse caso, a seqüência de aminoácidos e a conformação do polipeptídeo determinam se a proteína será modificada enzimaticamente e/ou o local intra ou

extracelular para onde ela será endereçada. Essas alterações podem resultar, portanto, não apenas em conversão na forma funcional da proteína, mas como também em seu direcionamento ao compartimento celular, no qual a proteína desempenhará sua função, em sua secreção da célula, ou em mudanças na atividade ou na estabilidade protéica.

Algumas proteínas assumem sua conformação funcional espontaneamente. No entanto, muitas vezes o enovelamento de proteínas conta com a participação de proteínas específicas, conhecidas como chaperonas moleculares. As chaperonas constituintes da família Hsp70 têm como função impedir a formação de estruturas indesejadas de um polipeptídeo recém-sintetizado. Esse grupo de proteínas também está envolvido na inserção dos polipeptídeos em membranas e no seu transporte através delas. Em contrapartida, outro grupo de proteínas, as chaperoninas, participa mais diretamente do dobramento dos polipeptídeos.

### QUANDO O PROCESSAMENTO SE INICIA?

Você pode estar pensando que as modificações só ocorrem nos polipeptídeos já sintetizados. O que se sabe, de fato, é que o processamento freqüentemente se inicia durante a tradução, terminando algum tempo após o polipeptídeo ter sido liberado dos ribossomos.

Em eucariotos, as proteínas sintetizadas nos polissomos livres podem ter dois destinos: permanecer no citosol ou ser endereçadas ao núcleo, ao cloroplasto ou à mitocôndria. Aquelas que permanecem no citosol estão prontas para funcionar tão logo sejam liberadas do ribossomo. Já aquelas que devem ser transportadas para um outro local, geralmente, sofrem modificações substanciais durante o processo de transporte.

Muitas proteínas, no entanto, são sintetizadas em polissomos ligados à membrana em lugar de polissomos livres. Em bactéria, os polissomos estão associados à membrana plasmática interna. Já em eucariotos os polissomos são encontrados em associação com o retículo endoplasmático, e as proteínas sintetizadas permanecem neste local ou se destinam ao complexo de Golgi, aos grânulos secretores, à membrana plasmática ou aos lisossomos. Se tiver dúvida em alguns desses termos, volte às aulas de Biologia Celular! Em geral, as proteínas sintetizadas em ribossomos ligados à membrana sofrem modificações mais extensas antes de atingirem seu destino final e se tornarem totalmente funcionais.

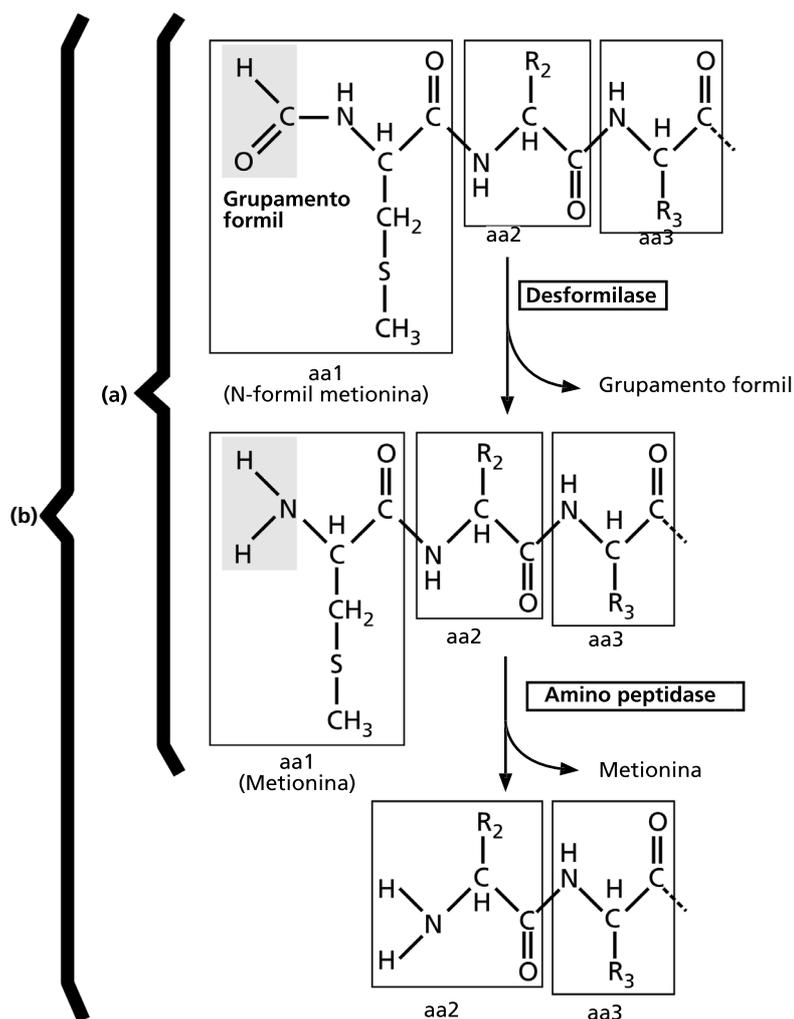
Voltaremos a esse tema quando discutirmos um mecanismo envolvido no endereçamento de proteínas.

## EXEMPLOS DE MODIFICAÇÕES PÓS-TRADUCIONAIS

Você já viu que as modificações pós-traducionais, que podem ser de vários tipos, envolvem diferentes porções do polipeptídeo a ser processado. Tanto os terminais amino e carboxi quanto os grupamentos laterais de inúmeros aminoácidos podem sofrer reações de adição de grupamentos específicos. Além disso, a remoção proteolítica de alguns resíduos de aminoácidos se faz necessária para que a proteína adquira sua conformação funcional. No processamento, pode ocorrer ainda a adição de grupamentos prostéticos e formação de pontes dissulfeto intra ou intercadeia entre os resíduos de cisteína. Apenas para ilustrarmos essas modificações, discutiremos as clivagens proteolíticas e algumas reações para adição de grupamentos específicos.

### Clivagem proteolítica

Provavelmente todas as proteínas maduras sofreram clivagem proteolítica. Vamos conferir um exemplo bem simples, que está representado na **Figura 29.1!** Você já viu na Aula 28 que, em procariotos, a síntese de polipeptídeos sempre se inicia com a N-formil-metionina. Entretanto, ao término da tradução, o grupamento formil é sempre retirado por uma desformilase. Será que todas as proteínas de procariotos apresentam um resíduo de metionina no terminal amino? Sabemos que isso não é verdade, pois, para alguns polipeptídeos, não apenas o grupamento formil, mas também o resíduo inicial de N-formil-metionina, sofre remoção proteolítica, pela ação de uma amino peptidase, ao término da tradução. Algumas vezes dois ou mais resíduos de aminoácidos são retirados pela amino peptidase.



**Figura 29.1:** Retirada do grupamento formil (a) e da N-formil metionina (b) dos polipeptídeos recém-sintetizados em bactérias. Apenas a porção inicial do polipeptídeo está representada. R2 e R3 representam, respectivamente, as cadeias laterais do segundo e do terceiro aminoácidos incorporados na cadeia polipeptídica. Já aa1, aa2 e aa3 representam os aminoácidos 1, 2 e 3, respectivamente.

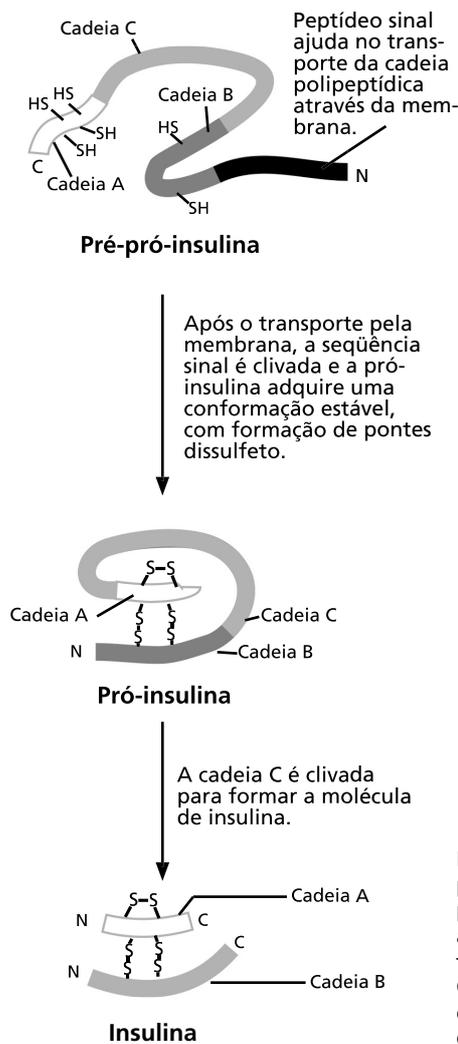
Muitas proteínas envolvidas em uma infinidade de processos biológicos são sintetizadas como precursores inativos que são ativados em condições próprias por proteólise. Como exemplos podemos citar as enzimas digestivas tripsinogênio e quimiotripsinogênio, produzidas pelo pâncreas, que sofrem clivagens de ligações peptídicas específicas para gerar as formas ativas, tripsina e quimiotripsina, respectivamente.

Todas as proteínas inativas que são ativadas por remoção de polipeptídeos são chamadas pró-proteínas. Já as enzimas sintetizadas sob sua forma inativa são mais comumente chamadas zimogênios; como, por exemplo, o tripsinogênio, citado no texto, que é a forma inativa da tripsina.

Algumas proteínas são sintetizadas como segmentos de poliproteínas, que contêm seqüências de duas ou mais proteínas. Nesse caso, a clivagem da poliproteína é necessária para a geração das proteínas funcionais, como é o caso dos hormônios polipeptídeos e das proteínas sintetizadas por muitos vírus.

Alguns polipeptídeos recém-sintetizados, tais como a pré-pró-insulina e o pré-pró-colágeno, estão sujeitos a três séries de clivagens proteolíticas: deleção do resíduo de metionina iniciadora; remoção de seus peptídeos sinais e excisão de seus pró-peptídeos. Vamos entender o processamento da pré-pró-insulina lendo o texto e acompanhando o esquema representado na **Figura 29.2**.

Muitos hormônios protéicos são sintetizados sob a forma de pré-proteína. Por exemplo, o mRNA de insulina é traduzido em uma única cadeia polipeptídica, reconhecida como pré-pró-insulina (110



aminoácidos), com 86 resíduos correspondentes à pró-insulina e 24 resíduos correspondendo ao peptídeo sinal. Como você deve saber, a insulina é produzida no pâncreas e é neste local que a seqüência amino terminal é removida ainda durante a tradução, direcionando a pró-insulina ao complexo de Golgi. As duas pontes dissulfeto ligando os terminais da molécula se formam durante esse processo. Em seguida, a cadeia C da pró-insulina é removida para produzir a forma circulante de insulina, que apresenta 51 resíduos em dois peptídeos ligados por duas pontes dissulfeto.

**Figura 29.2:** Processamento da pré-pró-insulina para formação da molécula de insulina. A pré-pró-insulina é composta pelo peptídeo sinal (24 aminoácidos) e as cadeias B, C e A, que apresentam 30, 35 e 21 aminoácidos, respectivamente. Os grupamentos SH representados no esquema estão presentes nas cadeias laterais dos resíduos de cisteína existentes no polipeptídeo.

## Modificação covalente

As proteínas estão sujeitas a modificações químicas dos grupamentos funcionais das cadeias laterais e dos grupamentos amino e carboxi-terminais. Mais de 150 tipos de modificações de cadeia lateral envolvendo todos os resíduos de aminoácidos, exceto alanina (Ala), glicina (Gly), isoleucina (Ile), leucina (Leu), metionina (Met) e valina (Val), são conhecidos, incluindo reações de **ACETILAÇÃO, GLICOSILAÇÃO, HIDROXILAÇÃO, METILAÇÃO, NUCLEOTIDILAÇÃO** e **FOSFORILAÇÃO**.

Algumas dessas modificações estão associadas a mecanismos de controle da atividade. Como, por exemplo, a fosforilação de inúmeras proteínas envolvidas no controle do ciclo celular. Outras modificações nas cadeias laterais favorecem a ligação covalente de co-fatores a enzimas, possivelmente para aumentar sua eficiência catalítica.

A glicosilação é uma modificação visível de proteínas que passam através do complexo de Golgi. Muitas das proteínas de superfície celular e de excreção produzidas nesse local são glicosiladas. Devido à diversidade de reações de glicosilação que ocorrem no complexo de Golgi e ao fato de essa organela ser o local principal do transporte de proteína, parece provável que os teores de carboidratos ligados a proteínas sejam responsáveis por direcioná-las aos seus destinos finais. A ligação de carboidratos complexos, que ocorrem em uma variedade quase infinita, altera as propriedades estruturais das proteínas e forma marcadores de reconhecimento em vários tipos de interações célula-célula e de direcionamento. Saiba ainda que as proteínas glicosiladas ou glicoproteínas, como são comumente chamadas, freqüentemente estão ligadas aos seus oligossacarídeos através de resíduos de asparagina. Além disso, a glicosilação é de grande importância para a formação da proteína funcional, o que é corroborado pelo fato de vários antibióticos interferirem em uma ou mais etapas deste processo.

## ENDEREÇAMENTO DE PROTEÍNAS

As vias envolvidas na classificação e transporte das proteínas para o seu local celular próprio são freqüentemente chamadas vias de endereçamento de proteínas. No entanto, como a célula “sabe” para onde a proteína será endereçada? Já parou para pensar nisso? Pois saiba que em todos os sistemas de endereçamento em eucariotos, com exceção daqueles

### **ACETILAÇÃO, GLICOSILAÇÃO, HIDROXILAÇÃO, METILAÇÃO, NUCLEOTIDILAÇÃO E FOSFORILAÇÃO**

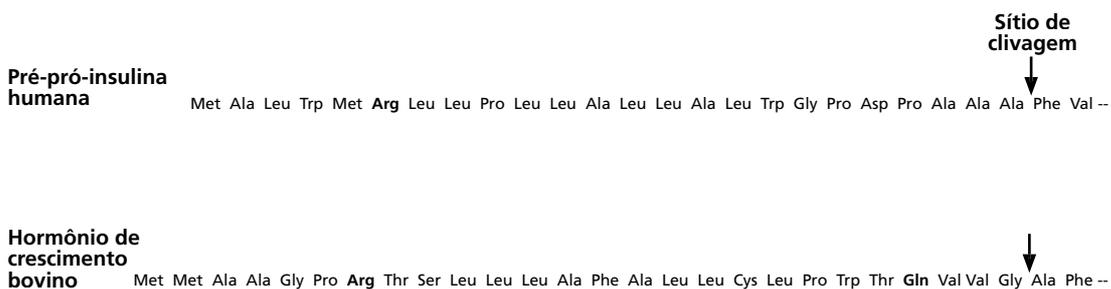
Reações que envolvem a adição de grupamentos acetil, glicosil (derivado de glicídios), hidroxila, metil, nucleotidil (derivado de nucleotídeos) e fosfato, respectivamente.

que envolvem proteínas citosólicas e nucleares, o elemento de sinalização mais importante é uma seqüência pequena de aminoácidos no terminal amino de um polipeptídeo recém-sintetizado chamada seqüência sinal (ou líder) ou peptídeo sinal. Essa seqüência, que direciona o polipeptídeo ao seu sistema de transporte apropriado e, portanto, ao seu local específico na célula, é geralmente removida durante o processo de transporte ou quando o polipeptídeo atinge seu destino final.

Vamos conhecer um pouco mais sobre essas seqüências sinais!

### SEQÜÊNCIAS SINALIZADORAS

As seqüências sinalizadoras, em eucariotos, localizadas perto do ou mesmo no terminal amino, variam em tamanho, apresentando entre 13 e 36 resíduos de aminoácidos, e são geralmente caracterizadas pela ocorrência de 10 a 15 resíduos hidrofóbicos. A presença desse tipo de resíduo reflete o fato de o transporte de proteínas constituir um problema com relação ao movimento das proteínas através de membranas lipídicas. Portanto, os resíduos hidrofóbicos tanto direcionam a proteína ao compartimento apropriado como iniciam sua penetração nas membranas. Uma outra característica interessante é que, geralmente, um ou mais aminoácidos carregados positivamente precedem a seqüência hidrofóbica. Uma série de aminoácidos com cadeias laterais pequenas, tais como a alanina, normalmente antecede o sítio de clivagem. Observe a **Figura 29.3**.



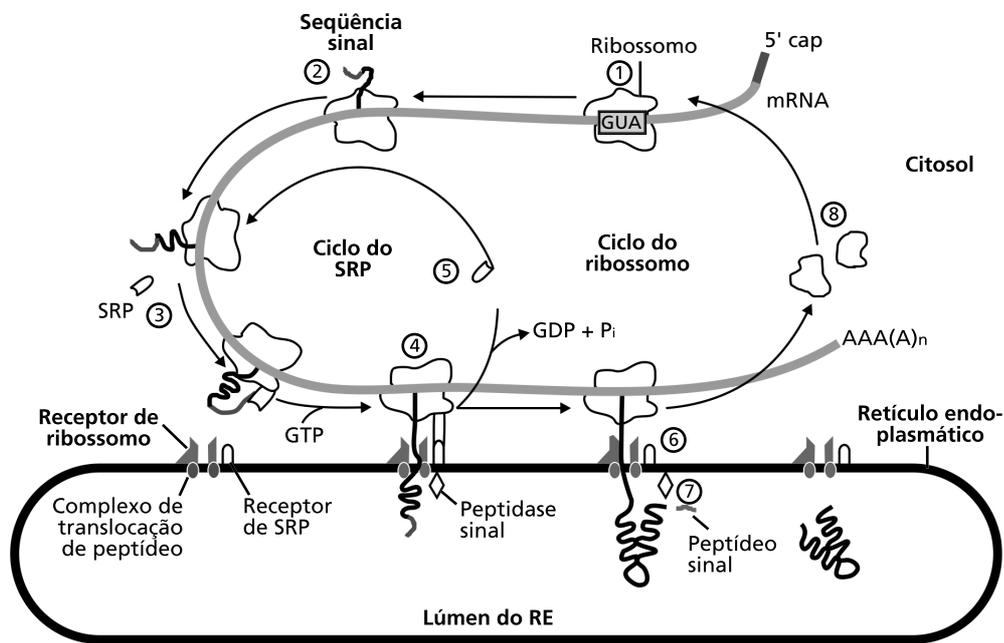
**Figura 29.3:** Seqüências sinais de proteína eucarióticas. Repare nos resíduos hidrofóbicos precedidos por um ou mais resíduos de aminoácidos carregados positivamente. Precedendo o sítio de clivagem encontram-se aminoácidos com cadeia lateral pequena.

As regras de identificação de seqüências sinais pelos compartimentos celulares individuais não são conhecidas. O que se sabe é que a natureza do reconhecimento é altamente conservada entre as espécies, uma vez que as seqüências sinais geralmente direcionam proteínas ao compartimento apropriado, mesmo em espécies diferentes. A especificidade do sinal também é independente da proteína a ser transportada, já que proteínas citoplasmáticas, por exemplo, podem ser direcionadas a locais na membrana pela adição de seqüências sinais específicas.

Em bactérias, o endereçamento de proteínas às membranas interna e externa, ao espaço periplásmico entre essas membranas e ao meio extracelular também usa seqüências sinais no terminal amino, de forma análoga ao que se observa em eucariotos no endereçamento de proteínas ao retículo endoplasmático, mitocôndria e cloroplastos. A maioria das proteínas não-citoplasmáticas de bactéria é direcionada através ou para dentro da membrana plasmática enquanto está sendo sintetizada no ribossomo, por um processo conhecido como transporte co-traducional, ou seja, que ocorre durante a tradução.

## ENDEREÇAMENTO DE PROTEÍNAS PARA O RETÍCULO ENDOPLASMÁTICO

Em uma célula eucariótica típica, o retículo endoplasmático é a maior organela ligada à membrana, sendo que sua maior parte contém ribossomos, constituindo o retículo endoplasmático rugoso (RER). Você se lembra das aulas de Biologia Celular? Os ribossomos do RER são o sítio de síntese de dois tipos de proteínas: as de membrana e as excretadas, sendo que para estas últimas representam o ponto de partida para as vias de secreção de proteínas. Na membrana e no lúmen do retículo endoplasmático, encontra-se um conjunto característico de proteínas que atua no processamento de proteínas secretadas e de membrana. Após um período curto no RER, proteínas a serem secretadas são transportadas por um processo de brotamento de vesícula e fusão para o complexo de Golgi, e de lá para a superfície celular, grânulos excretores ou lisossomas. Esse processo pode ser mais bem entendido se você ler o texto acompanhando a **Figura 29.4**.



**Figura 29.4:** Endereçamento de proteínas eucarióticas para o retículo endoplasmático. A explicação do esquema se encontra no texto.

Na figura, a etapa 1 representa o início da tradução, e a etapa 2 representa um estágio mais avançado no qual já se encontra formado o peptídeo sinal. O direcionamento inicial de polipeptídeos nascentes à membrana do retículo endoplasmático é resultado do reconhecimento co-traducional, ou seja, que ocorre simultaneamente ao processo de tradução, da seqüência sinal por um complexo ribonucleoprotéico. Esse complexo, chamado partícula de reconhecimento sinal – Signal Recognition Particle (SRP) –, é constituído de seis proteínas diferentes e uma molécula de RNA, de 300 resíduos, designada RNA 7S. Essa estrutura reconhece e se liga à seqüência sinal quando a cadeia nascente total atinge aproximadamente 90 resíduos de aminoácidos, o que só é possível porque a seqüência sinal está no amino-terminal (etapa 3). Ao se ligar, a SRP promove a interrupção da tradução da cadeia nascente, ao mesmo tempo que direciona o ribossomo com o polipeptídeo incompleto a um receptor específico de SRP, que é uma proteína de membrana do retículo endoplasmático (etapa 4). Após a ligação do polipeptídeo nascente, juntamente com o ribossomo, ao receptor de SRP no retículo endoplasmático, o SRP se dissocia do ribossomo (etapa 5), a síntese de proteína é reiniciada e a translocação do polipeptídeo nascente através da

membrana se inicia (etapa 6). A etapa de ancoramento envolve o consumo de GTP, sendo que a hidrólise dessa molécula é provavelmente essencial à liberação subsequente do SRP. Esse processo de translocação promove a formação do peptídeo no lúmen do retículo endoplasmático. Logo após a passagem do peptídeo sinal pela membrana ou após o término da síntese da proteína, a seqüência sinal é clivada do polipeptídeo recém-sintetizado pela ação de uma peptidase sinal ligada à membrana (etapa 7), o ribossomo se dissocia do retículo endoplasmático (etapa 8) e, no lúmen, as proteínas recém-sintetizadas são modificadas de diferentes maneiras.

O peptídeo sinal é reconhecido pelo SRP, que liga o polipeptídeo nascente, juntamente com o ribossomo, a um receptor na membrana do RER, em eucarioto, e à membrana plasmática, em bactéria, conduzindo o peptídeo sinal e seu polipeptídeo nascente através dela.

Outros processos poderiam ser estudados, tais como os envolvidos no endereçamento de proteínas que atuam nas mitocôndrias, nos cloroplastos ou no núcleo. Entretanto, o importante desta aula é você saber da existência e da importância de tais processos, de forma que você possa buscar essas informações quando elas se fizerem necessárias.

Para finalizar, busque a **Figura 29.5**, na qual está representado um esquema do metabolismo de proteínas em eucariotos. A seqüência de nucleotídeos no DNA que codifica a seqüência de aminoácidos de uma proteína é primeiramente transcrita, havendo formação de uma molécula de pré-mRNA. Este é devidamente processado, convertendo-se em uma molécula de mRNA maduro contendo a informação necessária para a síntese da proteína. Essa molécula de mRNA é, então, transportada ao citosol, local onde participa do processo de tradução nos ribossomos. O polipeptídeo recém-sintetizado pode sofrer modificações pós-traducionais antes ou após seu enovelamento e é endereçado a um local específico na célula. Após desempenhar sua função biológica, a proteína é degradada regenerando aminoácidos para participação na síntese de novas moléculas protéicas.

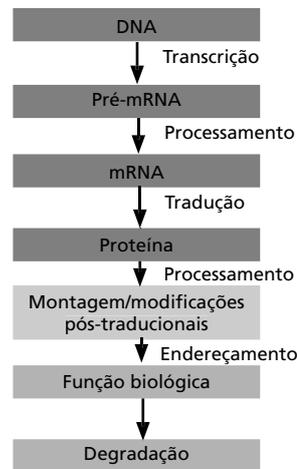


Figura 29.5: Metabolismo de proteínas. A explicação do esquema se encontra no texto.

## RESUMO

Os polipeptídeos liberados pelos ribossomos são processados e endereçados ao local onde desempenharão suas funções. O processamento envolve a adição de grupamentos às cadeias laterais de diversos resíduos de aminoácidos, a clivagem proteolítica ou a ligação de açúcares ou grupos prostéticos. Essa etapa é fundamental para a conversão na forma funcional da proteína e seu direcionamento ao compartimento celular, no qual desempenhará sua função, ou ainda para mudanças associadas a sua atividade e estabilidade.

Seqüências específicas perto do ou mesmo no terminal amino, conhecidas como seqüências sinal ou líder, direcionam o polipeptídeo a um sistema de transporte e, portanto, ao seu local específico na célula. No endereçamento de proteínas para o retículo endoplasmático, por exemplo, o reconhecimento do peptídeo sinal pela SRP ocorre durante a tradução. Isso resulta em uma interrupção na síntese do polipeptídeo e a ligação do complexo formado a um receptor de membrana. A SRP é, então, liberada, e a síntese do polipeptídeo é reiniciada. Com o término da síntese, uma peptidase sinal cliva a seqüência sinal, o polipeptídeo recém-sintetizado é liberado no lúmen do retículo endoplasmático e as subunidades ribossomais são liberadas no citosol, prontas para reiniciar a síntese de uma nova molécula de proteína.

## EXERCÍCIOS

1. Sabemos que a síntese das proteínas bacterianas se inicia com o resíduo de aminoácido N-formil metionina. Será que todas as proteínas funcionais apresentam como resíduo amino terminal a metionina? Escreva um pequeno texto comentando essa afirmativa com suas palavras.
2. Para a formação da insulina madura são necessárias todas as etapas listadas a seguir, exceto uma. Diga qual é e justifique.
  - (a) Remoção de um peptídeo sinal.
  - (b) Remoção de um peptídeo de uma região interna da cadeia.
  - (c) Dobramento da estrutura resultando na conformação funcional.
  - (d) Formação de ponte dissulfeto.
  - (e) Formação de  $\gamma$ -carboxiglutamato a partir de resíduos de glutamato.
3. Como uma célula "sabe" para onde endereçar uma proteína?
4. Como você explica a existência dos resíduos hidrofóbicos nas seqüências sinais?

## AUTO-AVALIAÇÃO

Ao final desta aula, você deve se sentir capaz de reconhecer a importância do processamento e do endereçamento dos polipeptídeos recém-sintetizados na formação de uma proteína funcional e o seu direcionamento para o local onde desempenhará sua função. Os exercícios têm por objetivo avaliar o quanto você aprendeu do tema em questão. Assim, caso encontre dificuldades para resolvê-los, releia a aula e recorra aos tutores, que estarão sempre nos pólos dispostos a ajudá-lo.

## FIQUE ATENTO!

O tema das próximas duas aulas é complexidade genômica. Até lá!



# Complexidade dos genomas I

AULA

# 30

## objetivos

Ao final desta aula, você deverá ser capaz de:

- Diferenciar os conceitos relativos à complexidade e ao tamanho dos genomas.
- Estabelecer comparações entre variações no tamanho do genoma haplóide das diversas espécies e sua correlação com a complexidade morfológica.
- Conceituar e enumerar os diferentes padrões de repetição das seqüências que compõem o genoma de eucariotos.
- Adquirir conhecimentos gerais sobre o genoma de mitocôndrias e cloroplastos.

### **Pré-requisitos**

DNA – aspectos funcionais e estruturais, Aula 4.

Estrutura dos cromossomos de vírus e procariotos, Aula 6.

Estrutura dos cromossomos de eucariotos, Aula 8.

Fluxo da informação gênica – transcrição e processamento do RNA, Aulas 20, 21 e 22.

## INTRODUÇÃO

O conhecimento dos genomas tem ganhado crescente importância no cenário científico. Através destes estudos, tem sido possível entender inúmeros processos biológicos, a relação filogenética entre as espécies, além da identificação de genes que controlam características de interesse.

Você já teve a oportunidade de estudar, nas aulas anteriores, que cada espécie se caracteriza por possuir um conjunto específico de genes. Estes genes se encontram no DNA que, por sua vez, é componente dos cromossomos. Nesta aula, discutiremos de maneira mais aprofundada a organização geral dos genomas das diferentes espécies, as características dos diferentes tipos de seqüência e a importância destes aspectos para o tamanho de cada genoma e sua complexidade.

## DEFININDO GENOMA

O genoma de uma espécie é freqüentemente definido como o conjunto de genes que a caracteriza. Porém, em nossos cromossomos, existem diversas regiões que não codificam nada, não são transcritas e, portanto, não correspondem a genes. Assim, uma definição mais apropriada seria a de que o genoma é o conjunto de todo o material genético contido nas células de uma espécie.

Na aula sobre estrutura dos cromossomos de vírus e procariotos (Aula 6 – Módulo 1), você teve a oportunidade de aprender que o genoma de bactérias é formado por uma grande molécula de DNA, fita dupla e circular. Muitas bactérias apresentam moléculas adicionais de DNA circular, denominadas plasmídeos. Considerando nossa definição anterior, o genoma dessas bactérias é composto pelo DNA cromossomal e pelo DNA plasmidial.

No caso dos genomas virais, você também teve a oportunidade de aprender, na Aula 6, que o material genético de vírus pode ser DNA ou RNA. Esse material genético, ou genoma, pode ainda se apresentar na forma de fita simples ou fita dupla. Enquanto permanece dentro do capsídeo viral, o material genético permanece na forma linear.

Em eucariotos, entretanto, diversas diferenças devem ser consideradas. O material genético se encontra dividido em vários cromossomos lineares, cujo número e tamanho variam de espécie para espécie. A maioria dos eucariotos possui mais de uma cópia de cada cromossomo. Esse número de cópias pode variar entre as espécies, e é definido como **ploidia**, representada pela letra “n” (conforme você aprendeu na Aula 7 da disciplina Genética Básica). Quando apenas uma cópia de cada cromossomo está presente, a espécie (ou célula)

é denominada haplóide (“n”). Quanto duas ou mais cópias estão presentes, pode tratar-se de uma espécie diplóide (2n) (duas cópias de cada cromossomo), triplóide (3n) (três cópias de cada cromossomo), tetraplóide (4n) (quatro cópias de cada cromossomo), e assim sucessivamente. As células somáticas da espécie humana, por exemplo, são diplóides, havendo duas cópias de cada cromossomo por célula. Os gametas de humanos, entretanto, possuem apenas uma cópia de cada cromossomo, representando células haplóides (n).

No decorrer desta aula, o termo genoma será utilizado para definir apenas um conjunto de cromossomos. Em outras palavras, estaremos sempre nos reportando ao genoma haplóide (uma cópia de cada cromossomo). Por exemplo, no caso da espécie humana, o genoma haplóide corresponde ao conjunto de 24 cromossomos (22 cromossomos autossômicos + X + Y).

## O CONTEÚDO DE DNA E A COMPLEXIDADE DOS ORGANISMOS

A quantidade total de DNA do genoma haplóide é característica de cada espécie e é conhecido como Valor-C. Tal valor varia amplamente entre as diversas espécies, entre valores inferiores a 600 mil pares de bases, em **micoplasmas**, e valores superiores a 670 bilhões de pares de bases, em determinadas espécies de amebas.

A análise do tamanho do genoma de milhares de espécies tem revelado um dado interessante. Em procariotos, o tamanho dos genomas (quantidade de DNA) está diretamente correlacionado com a complexidade funcional das espécies. Assim, espécies de micoplasma, que apresentam os menores genomas dentre procariotos, possuem pouco mais de 500 mil pb em seu genoma, para codificar 480 proteínas. A bactéria *Nostoc punctiforme*, entretanto, possui um genoma de aproximadamente 10 milhões de pares de bases, capazes de codificar cerca de 5.000 proteínas. Essa bactéria (cianobactéria) apresenta grande complexidade funcional dentre os procariotos, possuindo, inclusive, as capacidades de realizar fotossíntese e de fixar nitrogênio.

A correlação entre tamanho do genoma e complexidade funcional é algo bastante lógico. Se considerarmos que organismos funcionalmente mais complexos necessitam de um maior número de genes que coordenem tais funções, seus genomas teriam de apresentar maior quantidade de DNA para abrigar tais genes.

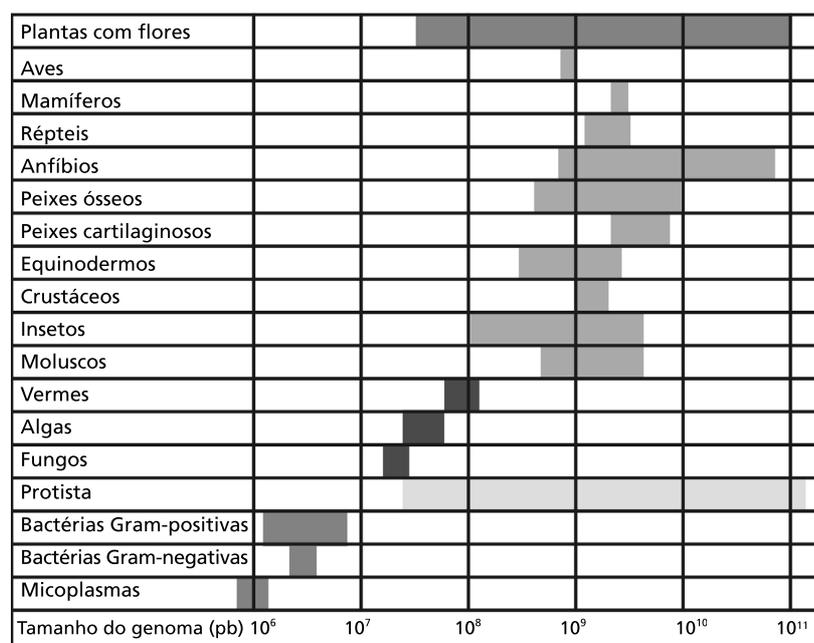
### MICOPLASMAS

Organismos procariotos muito simples que apresentam como uma de suas principais características a ausência de membrana plasmática. Determinadas doenças humanas, como a pneumonia, podem ser causadas por micoplasmas.

Entretanto, a correlação entre complexidade dos organismos e conteúdo total de DNA do genoma só é válida em procariotos. A análise dos genomas de eucariotos tem fornecido grandes surpresas. Embora tais genomas apresentem, em sua grande maioria, tamanhos bem superiores aos observados em procariotos, não é possível estabelecer correlação entre o conteúdo de DNA e a complexidade funcional dos organismos.

No **Gráfico 30.1**, são apresentadas as faixas de tamanhos estimados para os genomas de espécies dos diversos filos. Os tamanhos de genomas dos organismos eucariotos variam desde 12 milhões de nucleotídeos, em fungos, até 670 bilhões de nucleotídeos, em amebas (filo Protista). Isso representa uma variação de aproximadamente 55 mil vezes. Dentre a grande amplitude no tamanho dos genomas, você pode observar, na figura, que os genomas de todas as espécies de mamíferos têm tamanhos situados entre 2 e 4 bilhões de nucleotídeos. Esse valor é aproximadamente 200 vezes menor que o genoma de uma espécie de ameba (670 bilhões), e 50 vezes menor que o genoma de determinadas espécies de anfíbios. Além disso, dentro de um mesmo filo, podemos verificar grande amplitude de tamanho dos genomas entre as diversas espécies. Determinadas espécies de anfíbios, por exemplo, têm genomas 100 vezes maiores que outros anfíbios. Em plantas, tal discrepância é ainda maior, com algumas espécies apresentando genomas mil vezes maiores que o de outras.

**Gráfico 30.1:** O conteúdo de DNA do genoma haplóide é relacionado como a complexidade morfológica em procariotos, mas varia amplamente entre eucariotos. A gama de tamanhos de genomas das espécies que compõem cada filo é representada por barras horizontais.



As observações anteriores geram alguns questionamentos:

1- Que motivos justificariam o fato de uma ameba possuir um genoma 200 vezes maior que o genoma humano?

2- Dentro de um mesmo filo, como uma espécie de anfíbio poderia ter o genoma 100 vezes maior que o de outro anfíbio?

Antes de começarmos a discutir as questões anteriores, precisamos estabelecer certos conceitos gerais. Devemos considerar o fato de que determinados organismos são mais complexos que outros. Por exemplo, um organismo mamífero é morfológica e funcionalmente mais complexo que um fungo. Portanto, ele deverá possuir um maior número de genes em seu genoma, capazes de coordenar tais características. Assim, podemos estabelecer que um organismo que possua um maior número de genes distintos em seu genoma, terá um genoma mais complexo. Portanto, a complexidade do genoma está relacionada com a quantidade de genes diferentes presentes, enquanto o tamanho do genoma se refere ao conteúdo de DNA, em pares de bases (Valor-C). Você pode concluir, então, que uma determinada espécie de pássaros possui um genoma mais complexo (número de genes distintos) que o de uma espécie de molusco, embora os tamanhos de seus genomas possam ser similares.

Lembre-se, a complexidade do genoma está relacionada com o número de genes diferentes presentes, enquanto o tamanho do genoma se refere ao conteúdo de DNA, em pares de bases (Valor-C).

## COMPLEXIDADE DE SEQÜÊNCIAS DO GENOMA

A grande variação no tamanho do genoma de eucariotos permite observarmos que, embora determinadas espécies apresentem níveis equivalentes de complexidade (morfológica e funcional), muitas vezes seus genomas diferem muito quanto ao tamanho (conteúdo de DNA). Tal observação permite que formulemos três perguntas: 1) Genomas maiores possuem um maior número de diferentes genes? 2) Genomas maiores contêm maior número de cópias dos mesmos genes observados nos genomas menores? 3) Existem outros tipos de seqüência, que não sejam genes, responsáveis pelas diferenças de tamanho entre genomas?

Pare responder às questões anteriores, devemos considerar que, se o número de diferentes genes for maior nos genomas maiores, poderemos esperar que a quantidade de seqüências únicas de DNA do genoma

também aumente. Porém, se o aumento do tamanho do genoma resulta de maior número de cópias dos mesmos genes, o número de seqüências únicas no genoma não aumentará.

## ENSAIOS DE CINÉTICA DE REASSOCIAÇÃO E COMPLEXIDADE DE SEQÜÊNCIAS

A complexidade de seqüências de um genoma refere-se à quantidade de seqüências únicas (seqüências diferentes) nele presentes. Assim, quanto maior o número de genes distintos de um genoma, maior a quantidade de seqüências únicas.

Uma importante ferramenta para o estudo da complexidade dos genomas é a utilização de ensaios de cinética de reassociação. Para entendermos tais estudos, vamos lembrar algumas características do DNA, vistas na Aula 4. Cada molécula de DNA é composta por duas fitas de nucleotídeos complementares e antiparalelas. As fitas complementares são unidas por pontes de hidrogênio que, embora confirmam considerável estabilidade à molécula, podem ser desfeitas.

Dentre os componentes que podem desfazer as pontes de hidrogênio entre as fitas de DNA, estão o calor e o pH elevado. Ao processo de separação das fitas dá-se o nome de desnaturação do DNA. Tal processo é reversível, ou seja, caso o agente desnaturante seja removido, as duas fitas complementares voltam a estabelecer as pontes de hidrogênio, promovendo a renaturação da molécula de DNA. Você deve sempre considerar que a renaturação só é possível entre fitas complementares de DNA.

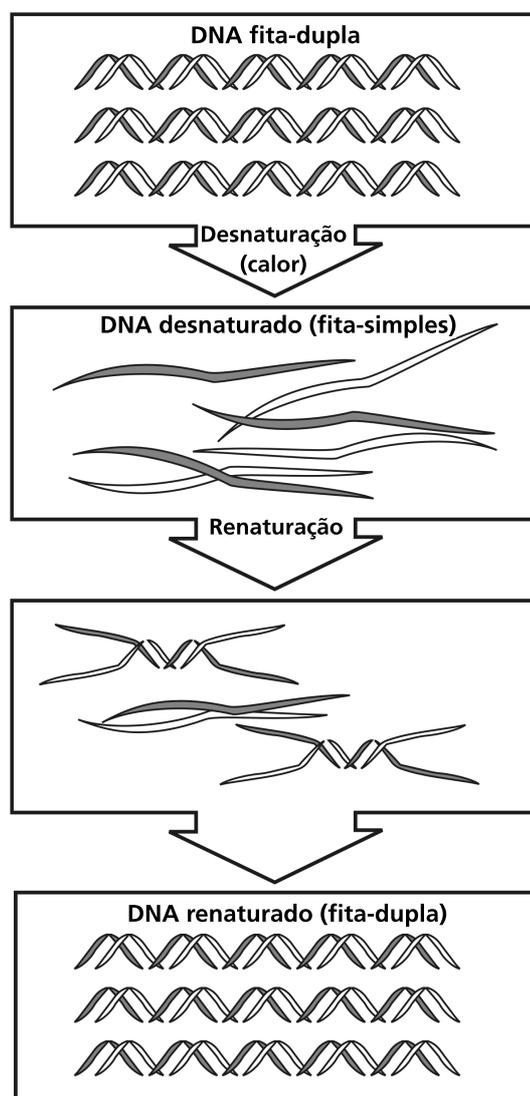
As características de desnaturação e renaturação das moléculas de DNA são essenciais aos trabalhos de cinética de reassociação. O processo, ilustrado na **Figura 30.1**, baseia-se na fragmentação do genoma em milhares de pequenos fragmentos e, em seguida, na exposição da solução ao calor. Como resultado, as pontes de hidrogênio entre as fitas serão desfeitas. Em seguida, o tubo é deixado resfriar à temperatura ambiente. Com a queda da temperatura, as fitas complementares tendem a se reassociar, recompondo as pontes de hidrogênio e, conseqüentemente, renaturando as moléculas de DNA.

A proporção de moléculas desnaturadas (fita simples) e renaturadas (fita dupla) dentro de uma solução pode ser monitorada, pois, quando expostas a um feixe de luz ultravioleta, moléculas desnaturadas apresentam níveis de absorção da luz diferentes dos observados para moléculas renaturadas. Assim, após a desnaturação térmica de uma amostra de DNA, é possível monitorar a velocidade de reassociação (renaturação) das moléculas. Tal ensaio é denominado cinética de reassociação.

Devemos considerar o fato de que, depois de dissociadas, duas fitas de DNA complementares terão de se reencontrar para que a reassociação ocorra. Se todos os fragmentos de DNA utilizados forem diferentes, a probabilidade de reassociação será a mesma para todos eles. Imagine, porém, se dentro da amostra houvesse várias cópias de um mesmo fragmento. Durante a fase de reassociação, a probabilidade de cada fita de DNA encontrar uma fita complementar seria muito maior, pois haveria várias cópias delas. Conseqüentemente, a formação de moléculas renaturadas seria muito mais rápida.

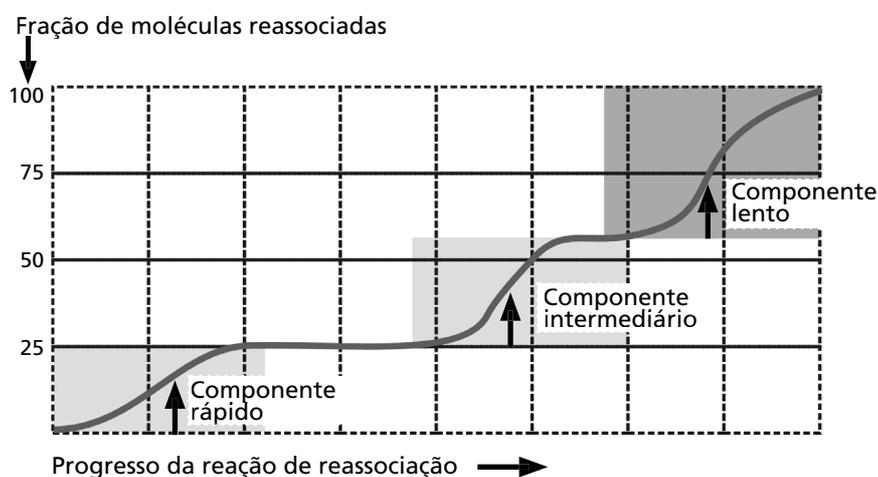
Dentro de um genoma pode haver regiões repetidas, ou seja, várias cópias de uma dada seqüência podem ser encontradas. Nesses casos, ao realizarmos um ensaio de cinética de reassociação, tais regiões tenderão a se reassociar muito mais rapidamente que as demais.

Agora que já discutimos os trabalhos de cinética de reassociação, podemos retomar a discussão acerca da complexidade dos genomas. Lembre-se de que, anteriormente, uma de nossas dúvidas acerca das diferenças entre os tamanhos dos genomas seria a presença de repetições de seqüências.



**Figura 30.1:** Processo de desnaturação térmica de fragmentos de DNA, seguida de renaturação.

Quando o DNA de um genoma de eucarioto é analisado através de cinética de reassociação, observa-se que a velocidade de reassociação das moléculas não ocorre de maneira homogênea. Uma parcela das moléculas se reassocia muito rapidamente (componente rápido), uma segunda parcela apresenta velocidade moderada de reassociação (componente moderado), e a terceira parcela apresenta reassociação lenta (componente lento), conforme apresentado no **Gráfico 30.2**. Quando estas velocidades de reassociação são comparadas com uma amostra de DNA composta apenas por seqüências não repetidas, observa-se que a velocidade é compatível apenas com o componente lento da reação.



**Gráfico 30.2:** A cinética de reassociação é apresentada sob a forma de um gráfico. Em genomas de eucariotos, são identificados três componentes, com velocidades de reassociação distintas. A fração de moléculas correspondentes a cada componente pode ser estimada.

As observações anteriores permitiram estimar que a parcela de fragmentos do genoma que se reassocia lentamente é composta por seqüências não repetidas. A fração de velocidade intermediária de reassociação é composta por seqüências moderadamente repetidas. O componente rápido é composto por seqüências altamente repetidas ao longo do genoma. No decorrer desta aula, vamos nos referir a tais frações como DNA não-repetitivo, DNA moderadamente repetitivo e DNA altamente repetitivo, respectivamente.

Os trabalhos de cinética de reassociação demonstraram que cada espécie de eucariotos apresenta proporções específicas de DNA repetitivo e não-repetitivo.

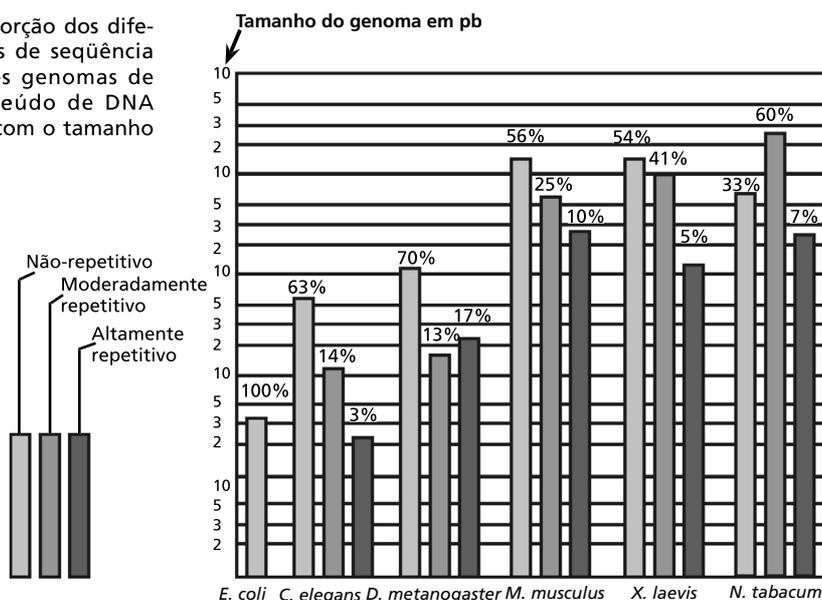
## A GRANDE MAIORIA DOS GENES ESTÁ PRESENTE NO DNA NÃO-REPETITIVO

A análise detalhada das diferentes frações de DNA tem revelado que a maioria dos genes ocorre na fração de DNA não-repetitivo. Em outras palavras, a maioria dos genes de eucariotos ocorre apenas uma vez no genoma haplóide. Alguns genes são repetidos, porém em pequeno número de cópias. Raros genes possuem dezenas de cópias ao longo do genoma. Enfim, a conclusão geral destas análises é que a maioria dos genes compõe o DNA não-repetitivo que, portanto, é a fração do genoma responsável pela complexidade funcional do organismo.

## O DNA REPETITIVO E O TAMANHO DOS GENOMAS

A proporção do genoma ocupada por DNA repetitivo e DNA não-repetitivo varia amplamente entre as espécies. O Gráfico 30.3 resume as proporções de DNA não-repetitivo, moderadamente repetitivo e altamente repetitivo em algumas espécies. Procariotos contêm apenas DNA não-repetitivo. Para eucariotos inferiores, a maior parte do DNA é não-repetitivo, restando menos de 20% do DNA dentro da fração moderadamente repetitiva. Em eucariotos superiores, entretanto, as proporções se invertem. Em animais, mais de 50% do genoma é composto por DNA repetitivo e, em plantas, a quantidade de DNA repetitivo pode ultrapassar 80% de todo o genoma, em determinadas espécies.

**Gráfico 30.3:** A proporção dos diferentes componentes de seqüência varia nos diferentes genomas de eucariotos. O conteúdo de DNA repetitivo aumenta com o tamanho dos genomas.



Ao compararmos estas informações com a amplitude de tamanho do genoma nas diferentes espécies, podemos então entender o fato de que a correlação entre tamanho do genoma e complexidade só é válida para organismos procariotos. Nessas espécies, os genomas são compostos por DNA não-repetitivo, que contém os genes. Em eucariotos, a presença de DNA repetitivo, que não contém genes, pode resultar no aumento do tamanho do genoma, sem que ocorra um aumento no número de genes (complexidade). Isso explica o fato de existirem espécies com níveis de complexidade similares e tamanhos de genomas completamente distintos. A diferença de conteúdo de DNA deve-se à presença de maior quantidade de DNA repetitivo em uma delas. Ou seja, em eucariotos superiores, o tamanho do genoma é fortemente afetado pelo conteúdo de DNA repetitivo.

Assim, o fato de uma determinada espécie de anfíbio possuir um genoma 50 vezes maior que o genoma humano pode ser explicado pela presença de grande quantidade de DNA repetitivo no genoma daquela espécie. Um outro exemplo que ilustra esta observação é o fato de que o genoma de mosca doméstica (*Musca domestica*) é seis vezes maior que o genoma de mosca das frutas (*Drosophila melanogaster*). Considerando a grande proximidade genética entre estas espécies, fica evidente que a complexidade de seus genomas é próxima. Assim, o que difere os tamanhos dos genomas é a presença de grande quantidade de DNA repetitivo em mosca doméstica.

Na **Tabela 30.1** estão listadas várias espécies de organismos, com os tamanhos de seus genomas e o número de genes presentes. Por exemplo, observe que o milho possui um genoma 5 vezes maior que o genoma humano, mas o número de genes é semelhante. Tais informações, mais uma vez, evidenciam a falta de correlação entre o tamanho do genoma (conteúdo de DNA) e o número de genes. Enfim, a complexidade funcional será dependente dos genes presentes e não do tamanho do genoma do organismo.

**Tabela 30.1:** O número de genes não está correlacionado com o tamanho do genoma.

Espécies	Tamanho do genoma	Número de genes
Humano	3,3 bilhões de pares de bases	35.000
Mosca-das-frutas ( <i>Drosophila melanogaster</i> )	120 milhões de pares de bases	13.601
Levedura ( <i>Saccharomyces cerevisiae</i> )	12 milhões de pares de bases	6.275
Nematóide (verme) ( <i>Caenorhabditis elegans</i> )	97 milhões de pares de bases	19.000
Milho ( <i>Zea mays</i> )	15 bilhões de pares de bases	30.000
<i>E. coli</i>	4,1 milhões de pares de bases	4.800
Arabidopsis ( <i>Arabidopsis thaliana</i> )	125 milhões de pares de bases	25.000

## O PAPEL DO DNA REPETITIVO NO GENOMA DE EUCARIOTOS

Ao considerarmos que os genes se encontram principalmente localizados na fração de DNA não-repetitivo, qual seria o papel das frações moderadamente repetitiva e altamente repetitiva do genoma?

A resposta para tal questão ainda permanece obscura. Sabemos que o DNA moderadamente repetitivo contém alguns genes (genes que ocorrem em várias cópias no genoma). Além disso, você teve a oportunidade de aprender, na Aula 19, que os elementos de transposição são seqüências capazes de produzir muitas cópias de si mesmas, que vão se inserir aleatoriamente ao longo do genoma. Diversas pesquisas têm demonstrado que grande parte do DNA repetitivo ao longo dos genomas resultou da presença de cópias de elementos de transposição. Existem diversas outras seqüências moderadamente repetidas ao longo do genoma cujas características e funções são pobremente elucidadas.

O papel do DNA altamente repetitivo é ainda menos esclarecido. Sabemos, entretanto, que determinadas regiões essenciais de nossos cromossomos são compostas por seqüências altamente repetidas. Dentre tais regiões, destacamos o centrômero e os telômeros dos cromossomos.

Essas regiões, embora não contenham genes, possuem importante função estrutural. Além de tais seqüências, muitas outras de função desconhecida, de diferentes tamanhos, também ocorrem em elevado número de cópias ao longo do genoma.

É importante destacar que, embora haja muitas regiões compostas exclusivamente por DNA repetitivo ou não-repetitivo, é freqüente encontrarmos porções dos cromossomos nos quais os dois tipos de seqüências se intercalam ao longo da molécula de DNA. A distinção entre tais seqüências dependeria de sabermos se existem outras cópias delas ao longo do genoma (DNA repetitivo) ou não (DNA não-repetitivo).

### **NEM TODO O DNA NÃO-REPETITIVO É COMPOSTO POR GENES**

O fato de sabermos que a grande maioria dos genes corresponde a seqüências únicas no genoma pode causar a falsa idéia de que todo o DNA não-repetitivo é constituído exclusivamente por genes. Porém, isso não é verdade. Existem muitas seqüências únicas do genoma que não são regiões codificadoras. Dentre essas regiões, podemos destacar as seqüências intergênicas, ou seja, regiões de DNA não-repetitivo que ocupam o espaço compreendido entre dois genes adjacentes, ao longo de uma molécula de DNA.

### **O TAMANHO MÉDIO DOS GENES VARIA ENTRE AS ESPÉCIES**

Um dos fatores que contribuem para as diferenças entre os tamanhos dos genomas das diversas espécies é o tamanho médio dos genes. Por exemplo, em leveduras, o tamanho dos genes normalmente não ultrapassa 10 kilobases, enquanto em mamíferos, determinados genes chegam a possuir mais de 100 kilobases de extensão. As diferenças no tamanho médio dos genes resultam, principalmente, da presença de íntrons. Conforme você teve a oportunidade de aprender, na Aula 22, a maioria dos genes de eucariotos superiores possuem a seqüência interrompida por íntrons. Aproximadamente 94% dos genes de mamíferos possuem, no mínimo, um íntron. Em eucariotos inferiores, como *Saccharomyces cerevisiae*, entretanto, aproximadamente 95% dos genes não possuem íntrons.

Em procariotos, por sua vez, não se verificam íntrons nos genes. Determinados genes de mamíferos chegam a possuir mais de 60 íntrons. Como conseqüência, a extensão de tais genes é sempre grande. Você não deve esquecer de que, após a transcrição, tais íntrons são retirados do transcrito durante o processamento e, portanto, não estarão representados nas proteínas codificadas.

A partir do exposto anteriormente, você pode concluir que uma porção considerável do DNA não-repetitivo do genoma de eucariotos superiores é composta pelos íntrons.

A bactéria *E. coli* possui aproximadamente 4.800 genes em um genoma de 4.000kb. O genoma humano possui aproximadamente 35.000 genes. Assim, seria de se esperar que o genoma humano fosse 10 vezes maior. Entretanto, o genoma humano é 1.200 vezes maior. Com o que você aprendeu nesta aula, podemos explicar tal discrepância com os seguintes argumentos: a) O genoma de procariotos não apresenta DNA repetitivo, enquanto eucariotos possuem grande quantidade de DNA repetitivo. b) Os genes de humanos são, em média, 10 vezes maiores que os de bactéria. c) Os genes humanos apresentam vários íntrons, enquanto os genes de procariotos não possuem íntrons.

## GENOMAS DE ORGANELAS

Embora a grande maioria dos genes de eucariotos esteja localizada nos cromossomos nucleares, desde o início do século XX, diversas evidências têm sugerido a existência de genes no citoplasma. Como você teve a oportunidade de aprender na disciplina Genética Básica, o fenômeno de herança materna resulta da presença de genes no citoplasma do óvulo. A herança materna é, portanto, uma das formas de herança não-mendeliana.

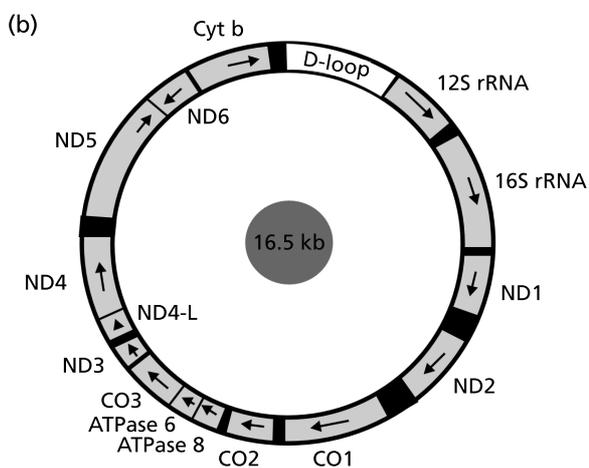
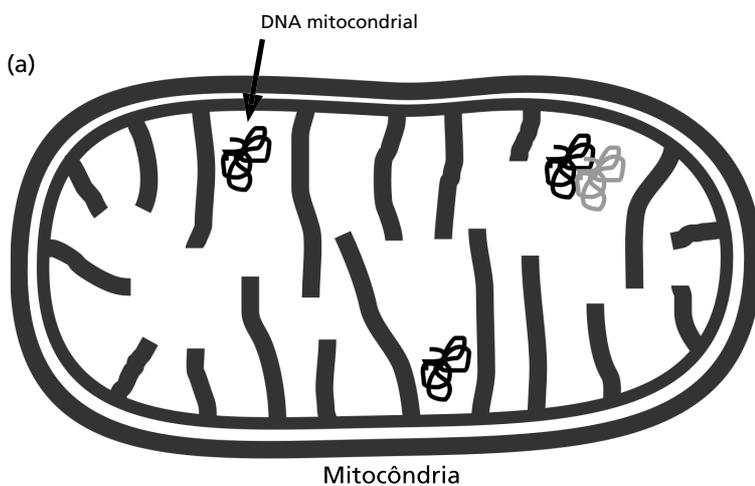
Tal tipo de herança resulta do fato de existir material genético dentro de mitocôndrias e cloroplastos. Em outras palavras, tais organelas possuem DNA, contendo genes que, portanto, são herdados independentemente do genoma nuclear. Dessa forma, os genes presentes no genoma de tais organelas encontram-se confinados, sujeitos a formas de expressão e regulação específicos.

Os genomas de mitocôndrias e cloroplastos são pequenos, circulares e possuem um número de genes relativamente restrito. Usualmente, existem várias cópias do genoma em cada organela. Considerando-se que existem diversas organelas em cada célula, podemos concluir que existem múltiplas cópias de genomas de organela por célula.

Os genes presentes codificam todos os RNAs ribossômicos e transportadores, porém, apenas algumas das proteínas necessárias pela organela são codificadas. As demais proteínas necessárias são codificadas por genes presentes no núcleo e, após serem sintetizadas, são endereçadas para a organela. A organela, portanto, possui seus próprios ribossomos e RNAs transportadores, além das enzimas envolvidas na transcrição e tradução. Como consequência, cloroplastos e mitocôndrias possuem o aparato necessário para a expressão de seus próprios genes. A seguir, vamos discutir separadamente os genomas de cada uma dessas organelas.

Genomas de cloroplastos são relativamente extensos, com aproximadamente 140kb em plantas superiores, e mais de 200kb em eucariotos inferiores. Em plantas superiores, o número de cópias do genoma por organela varia entre 20 e 40. Adicionalmente, existem entre 20 e 40 organelas por célula em tais plantas. O genoma de cloroplastos codifica para 4 RNAs ribossômicos, 30 RNAs transportadores e 50 proteínas.

O genoma mitocondrial varia amplamente em tamanho entre as espécies. Em mamíferos, seu tamanho é de aproximadamente 16,5kb. Em leveduras, tal genoma possui aproximadamente 84kb, enquanto em plantas superiores, o tamanho do genoma de mitocôndrias pode atingir 570kb. O número de mitocôndrias por célula varia entre as espécies. Em *S. cerevisiae*, existem aproximadamente 22 mitocôndrias por célula, com quatro cópias do genoma por organela. O número de proteínas codificadas pelo genoma mitocondrial é pequeno. Em humanos, o DNA mitocondrial contém 22 genes para RNAs transportadores, 2 genes para RNA ribossômico e 13 genes que codificam proteínas. As demais proteínas que compõem a organela são codificadas por genes nucleares. Na **Figura 30.2.a**, é apresentado um esquema ilustrativo de uma mitocôndria, contendo várias cópias de DNA circular. Na **Figura 30.2.b**, está esquematizado o genoma mitocondrial de mamíferos, com os principais genes indicados.



- Genes para tRNA
- Regiões codificadoras
- Indica a direção do gene
- CO Citocromo Oxidase
- ND NADH desidrogenase

**Figura 30.2:** (a) Esquema ilustrativo de uma mitocôndria contendo várias cópias do material genético. (b) Representação esquemática do genoma mitocondrial de mamíferos.

## RESUMO

Nesta aula, você teve a oportunidade de aprender os conceitos e características inerentes à organização geral dos genomas dos organismos. Para tanto, foram abordadas as diferenças gerais entre os genomas de bactérias, vírus e eucariotos. Através da análise dos tamanhos dos genomas de espécies dos diversos filos, foi evidenciada a existência de correlação entre conteúdo de DNA e a complexidade funcional dos organismos. Porém, vimos que tal correlação só é mantida em procariotos. Em eucariotos, tal correlação não foi verificada, havendo organismos com níveis de complexidade similares e tamanhos de genomas completamente distintos. Em seguida, foi abordada a ocorrência de seqüências repetidas ao longo dos genomas. Para isso, vimos os princípios inerentes à principal técnica utilizada para o estudo de complexidade de seqüências, a cinética de reassociação. Em seguida, discutimos que grande parte do genoma de eucariotos superiores é composta por DNA repetitivo, enquanto em procariotos tais seqüências inexistem. Vimos, também, que a grande maioria dos genes são seqüências únicas no genoma de eucariotos, constituindo, portanto, o DNA não-repetitivo. Dentre o DNA não-repetitivo, existem seqüências que não codificam proteínas, sendo representadas, principalmente, pelas regiões intergênicas e pelos íntrons dos genes. Ao final, tivemos a oportunidade de discutir a existência de material genético em mitocôndrias e cloroplastos. As características gerais inerentes aos genomas de tais organelas foram também abordadas.

## EXERCÍCIOS

1. Por que a palavra genoma não pode ser definida como o conjunto de genes que caracteriza uma espécie?
2. Quais as diferenças entre os termos "tamanho do genoma" e "complexidade do genoma"?
3. Qual a correlação entre conteúdo de DNA do genoma haplóide e complexidade funcional das espécies nos diferentes filos?
4. Como os ensaios de cinética de reassociação permitem estimar a presença de DNA repetitivo em uma amostra de DNA?

5. Qual a diferença entre os genomas de eucariotos superiores, eucariotos inferiores e procariotos quanto à presença de DNA repetitivo?
6. Todo o DNA não-repetitivo é composto por genes?
7. Quais as principais características dos genomas de cloroplastos e mitocôndrias?

### **AUTO-AVALIAÇÃO**

Se você compreendeu o conteúdo da aula de hoje, não deve ter encontrado dificuldades para responder aos exercícios. Lembre-se de que a visão geral acerca da organização dos genomas, os conceitos de tamanho de genoma (conteúdo de DNA) e sua correlação com a complexidade dos organismos são aspectos essenciais, pois tais informações serão importantes para a compreensão da próxima aula.



## Complexidade dos genomas II

AULA

# 31

## objetivos

Ao final desta aula, você deverá ser capaz de:

- Explicar os fatores envolvidos com a complexidade dos genomas de eucariotos.
- Descrever as principais características inerentes ao genoma humano.
- Entender importância da presença de elementos de transposição para o conteúdo de DNA do genoma de eucariotos.
- Entender a importância da emenda alternativa de éxons para a complexidade funcional dos organismos eucariotos.
- Absorver os conceitos sobre projetos de seqüenciamento e análise de genomas.

### Pré-requisitos

Estrutura dos cromossomos de vírus e procariotos (Aula 6 – Módulo 1).

Regulação da expressão gênica (Aula 24).

Elementos de transposição em eucariotos (Aula 19).

Complexidade dos genomas (Aula 30).

## INTRODUÇÃO

Na Aula 30, você teve a oportunidade de estudar os aspectos gerais que caracterizam a organização geral dos genomas e sua complexidade. Nela, foram abordados os conceitos relativos ao tamanho dos genomas (conteúdo de DNA) e à complexidade dos mesmos. Vimos também que os genomas das diferentes espécies podem diferir amplamente quanto à presença de DNA repetitivo. Ao final, discutimos as características dos materiais genéticos presentes em mitocôndrias e cloroplastos.

Apesar do grande volume de pesquisas relativas à caracterização dos genomas, realizadas previamente, um forte impulso dos conhecimentos foi gerado a partir do seqüenciamento do genoma humano, anunciado em 2000. A disponibilidade de informações acerca dos 3,3 bilhões de nucleotídeos que compõem o genoma humano tem permitido a elucidação de diversas perguntas relativas à organização do genoma, número e tamanho dos genes, proporção de DNA repetitivo, dentre muitas outras.

Nesta Aula, nos dedicaremos ao estudo mais detalhado da complexidade dos genomas, com ênfase em eucariotos, tomando como base as informações geradas pelo seqüenciamento do genoma humano.

## A ORIGEM DA COMPLEXIDADE

Na Aula 30, discutimos o fato de que organismos que apresentam maior complexidade funcional devem possuir um maior número de genes, capazes de coordenar tais funções. Tal afirmação é facilmente reafirmada ao compararmos os genomas de organismos que apresentem níveis de complexidade funcional contrastantes, tais como humanos e bactérias. O genoma da bactéria *E. coli*, por exemplo, possui aproximadamente 4.000 genes, enquanto o genoma humano apresenta cerca de 35.000 genes, 10 vezes mais.

Quando comparamos o número de genes dos genomas de organismos filogeneticamente muito distantes, tais como bactérias e humanos, torna-se fácil entender a correlação entre número de genes e complexidade funcional. Entretanto, quando a comparação é realizada entre organismos evolutivamente próximos, esta análise torna-se difícil. O número de genes é muito próximo entre espécies que apresentam complexidades funcionais consideravelmente distintas. Por exemplo, podemos analisar os primatas sagüi-pigmeu (*Cebuella pygmaea*) e chimpanzé (*Pan troglodytes*). Na **Figura 31.1**, você pode observar que,

embora ambos sejam primatas, o primeiro possui poucos centímetros de tamanho, enquanto um chimpanzé pode chegar a pesar 90 kg e medir 1,3 metro. A análise preliminar de seus genomas, entretanto, não revela diferenças consideráveis quanto ao número de genes. Assim, uma outra perspectiva deve ser considerada para explicar tal fato.



**Chimpanzé (*Pan troglodytes*)**

Tamanho: 1,3 metros  
Peso: 90 kg  
Tamanho do Genoma: 3,8 bilhões de pb



**Sagüi-pigmeu (*Cebuella pygmaea*)**

Tamanho: 15 centímetros  
Peso: 100 kg  
Tamanho do Genoma: 3,5 bilhões de pb

**Figura 31.1:** Exemplos de espécies geneticamente próximas que apresentam grandes diferenças morfológicas.

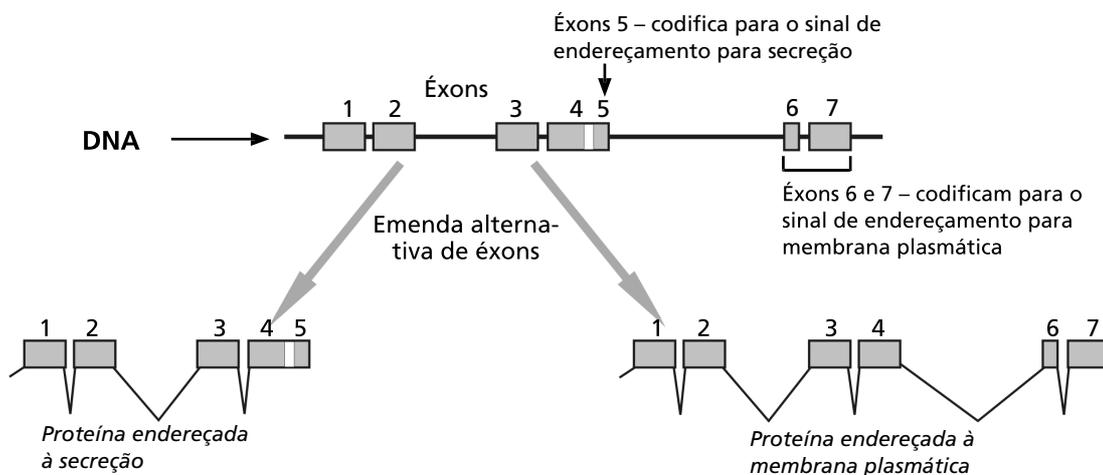
Nas Aulas 24 e 25 você teve a oportunidade de conhecer os diversos mecanismos de regulação da expressão gênica em eucariotos. Em tais aulas, discutimos que, embora todas as células somáticas de um organismo eucarioto possuam os mesmos genes, estes podem ser diferencialmente regulados nos diferentes tecidos, condições ambientais e fases do desenvolvimento. A expressão de conjuntos específicos de genes, durante o desenvolvimento embrionário, é responsável pela morfogênese de cada indivíduo. Tal processo é controlado por genes reguladores. Uma hipótese para explicar a grande diversidade morfológica observada entre indivíduos que apresentem número de genes parecidos, como no exemplo citado anteriormente, seria a presença e a ativação diferencial de genes reguladores. Assim, a presença de uns poucos genes reguladores (que controlam a formação dos órgãos e tecidos) em uma espécie, pode conferir grandes diferenças morfológicas e funcionais em relação a uma outra espécie que não os possua. Este mesmo raciocínio tem sido utilizado para explicar as diferenças observadas entre humanos e chimpanzés, que apresentam números de genes praticamente idênticos.

## COMPLEXIDADE FUNCIONAL E EMENDA DIFERENCIAL DE GENES

Um dos aspectos que tem intrigado os pesquisadores é o baixo número de genes, revelado durante o projeto de seqüenciamento do genoma humano. As previsões anteriores estimavam que seriam necessários entre 60 e 120 mil genes diferentes para atender a todas as funções já caracterizadas em seres humanos. Entretanto, o número encontrado é de aproximadamente 35.000 genes. Assim, que mecanismos o genoma utilizaria para atender ao grande número de funções requeridas?

Dentre as hipóteses que visam a responder à questão anterior, uma se baseia no sistema de regulação da expressão gênica. Conforme você aprendeu nas Aulas 24 e 25, a expressão diferencial de genes pode resultar em fenótipos distintos. Em outras palavras, um mesmo conjunto de genes pode ser diferencialmente ativado para cada tecido ou condição, de modo a resultar em características funcionais distintas.

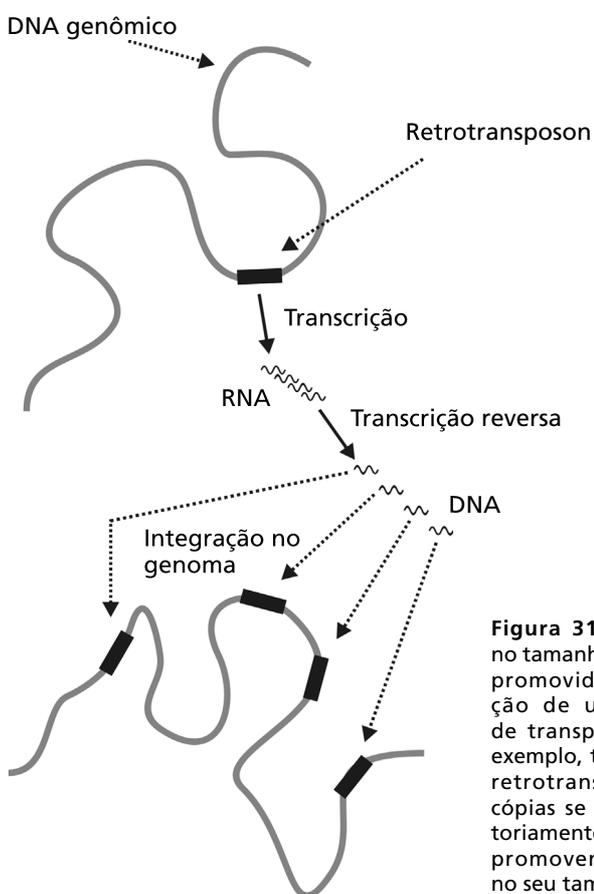
Uma outra hipótese para responder a tal questão é a existência de mecanismos de emenda diferencial de éxons. Conforme você aprendeu na Aula 30, a grande maioria dos genes humanos possui, no mínimo, um íntron. Dessa forma, existem muitos genes compostos por vários éxons, cujos transcritos necessitam ser emendados durante a fase de processamento do RNA (lembra-se da Aula 22?). O mecanismo de emenda alternativa de éxons, que você teve a oportunidade de aprender na Aula 24, corresponde a um sistema de regulação da expressão gênica que permite a formação de diferentes transcritos a partir de um mesmo gene. Na **Figura 31.2**, é apresentado um esquema ilustrativo no qual um mesmo gene pode produzir proteínas endereçadas a locais distintos da célula, em função da emenda alternativa de éxons. Estima-se que 40% dos genes de humanos estejam sujeitos à regulação por mecanismos de emenda alternativa de éxons. Portanto, esta se torna uma interessante explicação para a discrepância observada entre o baixo número de genes presentes e o elevado número de diferentes proteínas codificadas. A resposta pode ser a existência de vários produtos (polipeptídeos) codificados a partir de cada gene.



**Figura 31.2:** Esquema ilustrativo da geração de proteínas distintas a partir da emenda alternativa dos éxons codificados por um gene. O gene apresentado é composto por 7 éxons, dentre os quais o éxon 5 – que codifica para um sinal de endereçamento da proteína para secreção –, enquanto os éxons 6 e 7 codificam para o sinal que endereça a proteína para a membrana plasmática. Neste exemplo, a emenda alternativa de éxons resulta em proteínas endereçadas a locais distintos (membrana e secreção).

### ELEMENTOS DE TRANSPOSIÇÃO E CONTEÚDO DE DNA REPETITIVO

Conforme abordado na Aula 30, as espécies de eucariotos superiores diferem amplamente quanto ao conteúdo de DNA repetitivo. A presença de tal DNA é a principal responsável pelas diferenças observadas no conteúdo de DNA do genoma de espécies de eucariotos superiores. Uma considerável parcela do DNA repetitivo das espécies é representada pelos elementos de transposição. Na **Figura 31.3**, você pode observar uma ilustração de como a ativação de um elemento de transposição (do tipo retrotransposon) pode resultar no aumento do tamanho do genoma.



**Figura 31.3:** Aumento no tamanho do genoma, promovido pela ativação de um elemento de transposição. Neste exemplo, trata-se de um retrotransposon, cujas cópias se inserem aleatoriamente no genoma, promovendo aumento no seu tamanho.

**ARABIDOPSIS  
(*ARABIDOPSIS  
THALIANA*)**

Espécie de planta herbácea que cresce em clima temperado. Devido ao fato de estas plantas apresentarem estrutura muito simples e ciclo de vida curto (aproximadamente 60 dias), elas são utilizadas como modelo de estudos para cientistas.

A conclusão do seqüenciamento do genoma humano revelou um dado interessante: aproximadamente metade de nosso genoma é composta por seqüências remanescentes de elementos de transposição. De maneira similar, a comparação entre os genomas das plantas **ARABIDOPSIS** e arroz tem revelado que, embora possuam números de genes similares, o conteúdo de DNA de seus genomas é bastante discrepante, pois o genoma de arroz é aproximadamente quatro vezes maior que o outro. A análise destes genomas revelou, conforme esperado, uma quantidade muito maior de DNA repetitivo no genoma de arroz. O estudo da presença de elementos de transposição nesses genomas demonstrou a existência de uma grande quantidade de transposons no genoma de arroz, enquanto em arábido *arabidopsis* o número de tais seqüências foi bastante reduzido. Juntas, as informações anteriores permitiram aos pesquisadores concluir que a participação dos elementos de transposição no conteúdo de DNA do genoma de eucariotos é bem superior ao que se supunha previamente.

### **DENSIDADE DOS GENES AO LONGO DOS CROMOSSOMOS DE EUCARIOTOS**

A afirmação de que os genes de uma espécie encontram-se distribuídos ao longo de seus cromossomos pode gerar a falsa idéia de que tal distribuição é homogênea. Entretanto, a idéia mais apropriada é que os genomas de eucariotos superiores consistem em “oásis de genes separados por desertos vazios”. Em outras palavras, os genes estão concentrados em determinadas regiões dos cromossomos, separadas umas das outras por vastas regiões de seqüências que não codificam proteínas. Você deve se lembrar sempre de que a grande maioria de nosso genoma é composta por seqüências que não contêm genes.

A análise do genoma humano revelou que 75% do nosso genoma é composto por regiões que não contêm genes, 24% do genoma é representado por íntrons e apenas 1,1% é composto por éxons. Portanto, apenas 1,1 % do genoma de humanos codifica proteínas, enquanto os demais 98,9% são compostos por íntrons e regiões não-codificadoras. É importante ressaltar que muitas regiões, embora não codifiquem proteínas, atuam como componentes funcionais, ao desempenharem papel nos mecanismos de regulação da expressão gênica, por exemplo, as regiões promotoras dos genes. Outras seqüências, também não-codificadoras, desempenham papel estrutural nos cromossomos, tais como os centrômeros e telômeros.

A densidade de genes nos cromossomos humanos também não é homogênea. Determinados cromossomos possuem grande número de genes, enquanto outros apresentam poucos genes separados por uma grande quantidade de DNA repetitivo.

## **SIMILARIDADE GENÔMICA ENTRE INDIVÍDUOS DA MESMA ESPÉCIE**

Um interessante assunto que podemos abordar com relação ao genoma é a comparação da similaridade genética entre indivíduos de mesma espécie. Conforme você sabe, as diferenças fenotípicas entre as pessoas é, freqüentemente, resultante das características genéticas que elas possuem. Como consequência, durante a história da humanidade, as diferenças étnicas entre os diferentes grupos conduziram ao equívoco de tentar agrupar tais indivíduos em raças. Porém, tal agrupamento foi normalmente fundamentado em alguns poucos caracteres fenotípicos, tais como a cor da pele, o tipo de cabelo, o formato dos olhos etc.

O estudo dos genomas revelou um resultado surpreendente. A comparação entre os genomas de indivíduos das diferentes raças, previamente propostas, revelou que as seqüências são idênticas em 99,9% dos nucleotídeos. Quando a comparação foi realizada entre indivíduos da mesma “raça”, o nível de semelhança foi o mesmo, 99,9%. Estes resultados conduzem às seguintes conclusões:

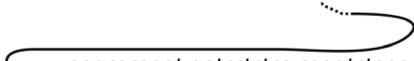
- 1) as diferenças genéticas, que observamos entre indivíduos da espécie humana, correspondem a apenas 0,1% de todo o DNA;
- 2) o conceito de raças na espécie humana não faz qualquer sentido, pois dois indivíduos de uma mesma “raça” são tão similares entre si, 99,9%, quanto dois indivíduos de “raças” diferentes.

As diferenças entre os indivíduos da espécie humana são da ordem de 0,1% dos nucleotídeos. Portanto, podemos perceber que estas poucas diferenças são suficientes para conferir a cada pessoa características fenotípicas exclusivas.

## O CONHECIMENTO DAS SEQÜÊNCIAS DOS GENOMAS E DA FUNÇÃO DE SEUS GENES

Atualmente, os pesquisadores têm dedicado grande esforço ao seqüenciamento e à caracterização dos genomas de diversas espécies de procariotos e eucariotos. Devido ao tamanho reduzido, os genomas de centenas de espécies de procariotos já foram seqüenciados até o momento. No caso de eucariotos superiores, entretanto, esse número se restringe a poucas espécies de animais, incluindo o homem, e de plantas.

Aos trabalhos de seqüenciamento do material genético e identificação das seqüências codificadoras presentes em uma espécie dá-se o nome de Projeto Genoma. O Projeto Genoma Humano, por exemplo, dedicou-se ao seqüenciamento dos nucleotídeos presentes nos 24 cromossomos que caracterizam a espécie (22 cromossomos autossômicos + X + Y). Na **Figura 31.4**, é apresentada a seqüência de nucleotídeos de uma pequena porção do genoma (lembre-se de que o genoma humano possui aproximadamente 3,3 bilhões de nucleotídeos!). Em seguida, foram desenvolvidos os trabalhos que visaram a localizar as regiões codificadoras ao longo da seqüência. Na **Figura 31.5**, é apresentado um esquema ilustrando as regiões codificadoras encontradas em uma seqüência de nucleotídeos. Porém, em humanos, a distância entre genes é muito maior que o apresentado na figura. Além disso, a maioria das regiões codificadoras de humanos é interrompida por íntrons. O próximo passo de um projeto genoma é a caracterização da função de cada um dos aproximadamente 35.000 genes identificados. Esta fase é a mais complexa e, embora muitos genes já tenham sido caracterizados, a função da maioria deles ainda permanece pouco elucidada, ou completamente obscura. A determinação da função de todos os genes da espécie humana permanece um grande desafio a ser vencido ao longo dos próximos anos.



```

aagcacaat gatcctca caagtctaaa tgcgctagag
4441 gactttgagc ggatcgatct cttgaatggt gtctttggag tgggtgctat
4501 ttgaatgaga agtagtgaat gcttggatgg ttggagtga agtggttggg
4561 agcccctaac caccaattca accgttgggg caggctgcta tcgatgggcg
4621 tgtccggtgc gcmagccacg tcaccaatc gttagggttc tgacggttc
4681 gctttgtctt cagctgcac cggatagtcc ggtgccgcac cggacagga
4741 tttggtgcbc cttctacgtc tgcttgact tctacgcgaa ctgtccgcgc
4801 tgcaggcac cgttgctctg gctagccgtt gctccgctgg tgcaccggat
4861 cacacggaca agttcgggta attatagcgg agcaacgctt ctgaaaccca
4921 gtttgaagtc gtacggcctt ggtgcaccgg acaactgccg tgcgccagac
4981 tcaatttctt ttactcttt
    
```



**Figura 31.4:** Esquema ilustrando a seqüência de nucleotídeos de uma pequena porção do genoma. Embora saibamos que o DNA é fita-dupla, apenas uma das fitas tem sua seqüência apresentada.



No caso da espécie humana, espera-se que o conhecimento das funções de todos os genes permita a identificação das regiões responsáveis por todas as doenças genéticas, bem como os genes envolvidos com a suscetibilidade ou resistência a doenças infecciosas. Estes resultados permitirão o desenvolvimento de terapias capazes de curar tais doenças.

Adicionalmente, o conhecimento dos genomas de um considerável número de espécies dos diversos filos permitirá a análise comparativa de todas estas seqüências, visando a identificar os genes presentes, ou ausentes, em cada uma delas. Assim, espera-se determinar os genes envolvidos com as características morfológicas e funcionais que definem cada espécie.

## RESUMO

Nesta aula, você teve a oportunidade de aprender alguns aspectos acerca da organização e complexidade dos genomas. Vimos que a maior complexidade funcional dos genomas de eucariotos, em relação aos procariotos, resulta do maior número de genes presentes. Adicionalmente, vimos que, em espécies filogeneticamente próximas, o número de genes é freqüentemente muito parecido, apesar de grandes diferenças morfológicas serem observadas. Nesses casos, as diferenças não decorrem do número de genes, mas da expressão diferencial dos mesmos.

As análises do genoma humano revelaram um número relativamente pequeno de genes, aproximadamente 35.000. Dentre as hipóteses formuladas para explicar a discrepância entre o número de genes e o número de funções requeridas, vimos que a expressão diferencial de genes pode explicar, em parte, tal discrepância. Uma segunda hipótese é a existência de emenda alternativa de éxons em muitas espécies.

No âmbito do conteúdo de DNA, você teve a oportunidade de aprender que uma considerável porção do DNA repetitivo observado em eucariotos superiores é composta por elementos de transposição ou seqüências remanescentes destes.

A distribuição dos genes não é homogênea ao longo do genoma humano, estando concentrados em determinadas regiões dos cromossomos, separadas por vastas regiões não-codificadoras. Aproximadamente 75% do genoma humano é composto por regiões que não contêm genes, 24% é composto por 1,1% é

composto por éxons. A comparação de seqüências do genoma de indivíduos distintos da espécie humana revelou que o nível de diferenças genéticas é de aproximadamente 0,1%.

O conhecimento dos genomas se encontra em fase inicial. Após o conhecimento das seqüências e identificação das regiões codificadoras, tem início um grande desafio, representado pelo estudo da regulação e da função de cada gene. Paralelamente, têm sido desenvolvidos projetos transcrito e proteoma para muitas espécies.

### EXERCÍCIOS

1. Eucariotos superiores apresentam complexidade morfológica e funcional muito maior que a observada em procariotos. Explique a correlação entre tal diferença e o número de genes presentes.
2. Indivíduos de espécies filogeneticamente muito próximas possuem número de genes muito similar. Como você poderia explicar as grandes diferenças morfológicas freqüentemente observadas?
3. O genoma humano possui um número de genes relativamente pequeno em relação às funções requeridas pelo organismo. Qual o papel da emenda diferencial de genes na complexidade funcional de humanos?
4. Qual a correlação entre a presença de elementos de transposição e a quantidade de DNA repetitivo do genoma de eucariotos?
5. Qual a proporção do genoma humano composta por regiões codificadoras?
6. Por que a análise dos dados do genoma humano revelou que não faz sentido agrupar os indivíduos da espécie humana em diferentes raças?
7. Além do conhecimento da seqüência de nucleotídeos de um genoma e da localização de seus genes, que outras informações devem ser obtidas para o conhecimento da complexidade funcional de uma espécie?

### **AUTO-AVALIAÇÃO**

Se você compreendeu o conteúdo desta aula, não deve ter encontrado dificuldades para responder aos exercícios. Esta aula enfocou detalhadamente os principais aspectos inerentes ao genoma de eucariotos, com ênfase no genoma humano. Entretanto, lembre-se de que tais informações correspondem a uma pequena parte de um conjunto imenso de descobertas realizadas atualmente. A cada dia, novas informações têm surgido acerca dos genomas, pois a ciência não pára.

**Biologia Molecular**

Gabbarito

1. Embora todas as células somáticas de um organismo eucarioto apresentem a mesma constituição genética, nem todos os genes se expressam simultaneamente em todas elas. Estima-se que apenas 10% dos genes se expressam em todos os tipos celulares durante todo o tempo. Tais genes são denominados genes constitutivos. Os demais têm sua expressão regulada e, portanto, expressam-se apenas em determinados tecidos ou condições. Assim, em cada tipo celular, um conjunto específico de genes se expressa, enquanto os demais permanecem inativos. A expressão de genes específicos em cada tipo celular resulta em características morfológicas, funcionais e de constituição distinta.

Comentário: este exercício fará com que você revise os conceitos relativos à expressão diferencial de genes e sua regulação em função dos níveis, locais e períodos de expressão.

2. Cascatas de regulação da expressão gênica são processos coordenados de regulação seqüencial da expressão de genes. Tal mecanismo é verificado para conjuntos de genes que compõem circuitos pré-programados de expressão, em que um evento inicial ativa um conjunto de genes, dentre os quais alguns têm função reguladora. Em seguida, tais genes reguladores atuam tanto na ativação de um segundo conjunto de genes quanto na inativação de genes do primeiro conjunto. Desse segundo conjunto de genes, alguns são também reguladores, ativando a expressão de um terceiro conjunto de genes, e assim por diante. A ordem seqüencial da expressão destes genes é, portanto, geneticamente pré-programada.

Comentário: este exercício fará com que você revise as informações relativas a expressão seqüencial de genes e, conseqüentemente, estabeleça o conceito de cascatas de expressão gênica.

3. Quanto à regulação da expressão, os genes podem ser classificados em genes de expressão constitutiva, genes de expressão tecido-específica, genes de expressão temporal e genes de expressão induzível. Genes de expressão constitutiva são aqueles cuja expressão pode ser verificada em todos os tipos celulares, o tempo todo. Genes de expressão tecido-específica apresentam sua expressão restrita a determinados tecidos ou tipos celulares. Genes de expressão temporal expressam-se apenas durante determinado período. Genes de expressão induzível são aqueles cuja expressão é ativada em resposta a fatores ambientais, tais como frio, calor, ferimento, infecções etc.

Comentário: este exercício visa a que você fixe as informações relativas ao agrupamento dos genes quanto ao seu padrão de regulação da expressão.

4. Para que ocorra a expressão de um gene de eucarioto, uma seqüência de várias etapas deve ser concluída: transcrição, processamento do transcrito, transporte para o citoplasma, tradução, modificações do polipeptídeo formado e endereçamento da proteína para o compartimento celular apropriado. Mecanismos celulares que afetem a eficiência de qualquer uma dessas etapas serão, portanto, formas de regulação da expressão gênica.

Comentário: este exercício fará com que você observe que a regulação da expressão gênica de eucariotos é composta de várias etapas distintas e que, em diferentes níveis, todas elas podem ser reguladas.

5. O principal mecanismo de regulação da expressão gênica em eucariotos é o controle da transcrição. Esta representa a forma mais econômica de regulação, pois, caso o gene não seja transcrito, as demais etapas da expressão gênica não se verificam. Existem dois mecanismos principais de regulação transcricional: o remodelamento da cromatina, que define as regiões do genoma que estarão disponíveis para a transcrição e a regulação através de proteínas reguladoras, que permite o controle da expressão de cada gene, individualmente.

Comentário: com este exercício você terá a oportunidade de revisar os conceitos gerais de regulação, de modo a observar a importância da regulação da transcrição.

6. Após a transcrição, o transcrito primário (pré-RNA<sub>m</sub>) deve ser processado, transportado para o citoplasma e utilizado para a tradução. Tais etapas envolvem vários processos que podem ser regulados. O processamento do transcrito envolve vários eventos distintos, dentre os quais podemos destacar a adição do capacet e da cauda poliA, a retirada de íntrons e emenda dos éxons e a edição do transcritos. A ocorrência de regulação destas etapas tem sido reportada, com ênfase aos mecanismos que permitem a “emenda alternativa de éxons” e os sistemas de edição do transcrito. Tais mecanismos permitem a um mesmo gene a capacidade de codificar diferentes formas da proteína. O transporte do RNA<sub>m</sub> formado para o citoplasma também pode ser regulado. Ao chegar no citoplasma, cada RNA<sub>m</sub> possui uma longevidade específica (estabilidade). Mecanismos que aumentem ou reduzam a suscetibilidade de um dado transcrito à degradação representam também uma forma de regulação pós-transcricional.

Comentário: Neste exercício você terá a oportunidade de revisar os diferentes mecanismos que atuam sobre a molécula de RNA após a sua síntese e, conseqüentemente, as formas através das quais os mesmos podem ser regulados.

7. Após a tradução (que também pode ser regulada), o polipeptídeo formado precisa ser corretamente processado para que seja endereçado ao compartimento celular apropriado e constitua uma proteína biologicamente ativa. Esse processo envolve vários mecanismos distintos: a clivagem de regiões da proteína (sob a ação de proteases específicas), o endereçamento para os compartimentos e a adição de modificações estruturais (adição de radicais). Todas essas etapas podem ser reguladas, afetando a produção da proteína ativa e, conseqüentemente, a expressão do gene que a codifica. As proteínas também possuem longevidade (estabilidade), período após o qual são degradadas. Mecanismos que afetam a estabilidade da proteína são formas de regulação pós-traducional da expressão gênica.

Comentário: este exercício permitirá a você revisar os vários mecanismos de regulação que podem atuar após a síntese dos polipeptídios.

## Aula 25

---

1. Para que um gene seja transcrito, é necessário que sua região reguladora esteja acessível ao complexo da transcrição (fatores de transcrição e RNA polimerase II). Quando uma região do genoma se encontra altamente condensada, o acesso do referido complexo protéico é inviabilizado. O remodelamento da cromatina, portanto, é capaz de definir quais regiões do genoma são passíveis de serem transcritas. O remodelamento da cromatina não está relacionado apenas com os níveis mais elevados de compactação da cromatina. A alteração dos níveis de interação entre nucleossomos adjacentes também resulta em remodelamento da cromatina, influenciando a expressão do gene.

Comentários: para responder a tal questão, você necessitará dominar as informações fornecidas sobre os níveis de compactação da cromatina e sua correlação com os níveis de transcrição.

2. O principal mecanismo de modificação estrutural do DNA é o processo de metilação. A metilação consiste na adição de radicais metil às bases nitrogenadas citosinas do DNA. O processo é desempenhado por enzimas DNA metiltransferases. Regiões do genoma onde o DNA é altamente metilado normalmente encontram-se condensadas.

Existe forte correlação entre nível de metilação de uma região do genoma e o nível de atividade transcricional. De maneira geral, regiões altamente metiladas são transcricionalmente inativas.

Comentários: o objetivo deste exercício é que você correlacione os níveis de metilação do DNA, o remodelamento da cromatina e a atividade transcricional.

3. Os três principais mecanismos de modificação das proteínas histonas são a acetilação, fosforilação e metilação que consistem na adição de radicais acetil, fosfato ou metil, respectivamente. Tais modificações ocorrem em caudas aminotermiais das histonas que se projetam para fora da estrutura do nucleossomo. A presença, ou ausência, de tais modificações altera a capacidade de interação entre nucleossomos adjacentes e, conseqüentemente, afeta o nível de compactação da cromatina. Paralelamente, a presença, ou ausência, de tais modificações afeta o acesso de proteínas reguladoras a uma dada região do genoma. A atividade transcricional é, portanto, afetada por duas vias: a primeira está relacionada com os níveis de compactação da cromatina, que pode facilitar ou dificultar o acesso do complexo da transcrição em função de seu nível de condensação; a segunda via está relacionada com a interação com proteínas reguladoras, capazes de ativar ou reprimir o acoplamento do complexo da transcrição.

Comentário: este exercício fará com que você relembre os mecanismos de modificação das proteínas histonas e correlacione tais processos com a conformação da cromatina e a ativação da transcrição.

4. A ligação de proteínas reguladoras pode exercer dois papéis contrastantes sobre a transcrição de um gene. Muitas dessas proteínas ativam o processo de transcrição, pois facilitam o acesso do complexo da transcrição ao promotor. Nesses casos, tais proteínas são denominadas ativadores (ou proteínas ativadoras) da transcrição. Entretanto, determinadas proteínas reguladoras atuam prejudicando a transcrição de gene, atenuando ou inviabilizando o acesso do complexo da RNA polimerase. Tais proteínas são denominadas repressores (ou proteínas repressoras) da transcrição.

Comentário: este exercício fará com que você estabeleça o conceito de que proteínas reguladoras podem ativar ou reprimir a expressão gênica.

5. As proteínas reguladoras possuem dois domínios principais denominados "Domínio de ligação ao DNA" e "Domínio de ativação (ou repressão)". O domínio de ligação ao DNA desempenha o papel de reconhecer seqüências de DNA específicas das regiões reguladoras de determinados genes. Tais seqüências são denominadas "sítios de ligação". O papel do "domínio de ativação" é realizar a interação com proteínas

do complexo da transcrição (fatores de transcrição), facilitando (no caso de proteínas ativadoras) ou dificultando (no caso de proteínas repressoras) o acoplamento do complexo da DNA polimerase II.

Comentário: este exercício fará com que você revise o assunto relacionado aos domínios que caracterizam as proteínas reguladoras.

## Aula 26

---

1. Após a descoberta de sua natureza física, ou seja, de sua localização, o gene foi definido como uma região específica do cromossomo capaz de determinar ou afetar uma única característica ou fenótipo.

Esse conceito sofreu mudanças com o desenvolvimento da ciência. Primeiramente, a proposição de Garrod, em 1908, de que os erros inatos do metabolismo eram herdados, o que sugeria a existência de uma relação entre o gene e a parte química do nosso organismo. Anos mais tarde, em 1940, a partir de uma série de experimentos usando o fungo *Neurospora crassa* como modelo, Beadle e Tatum propuseram a relação um gene–uma enzima. Esta hipótese foi prontamente aceita, mas foi alterada para um gene–uma proteína e depois para um gene–um polipeptídeo. Hoje em dia o conceito aceito é o de que gene é uma seqüência de nucleotídeos que determina a síntese de moléculas com função celular específica.

2. Sabemos que os eventos mutacionais ocorrem espontaneamente, mas a uma taxa muito baixa, o que reduz a probabilidade de isolamento de mutantes de qualquer organismo com características desejadas. O estudo de Beadle e Tatum dificilmente teria sido realizado sem a exposição de esporos do fungo *Neurospora* a raios X. De fato, a utilização de agentes mutagênicos contribuiu consideravelmente para as conquistas na área da Genética, uma vez que, ao aumentar a taxa de mutação, aumenta a probabilidade de se isolar mutantes, que foram e ainda são fundamentais para o entendimento dos processos celulares.

3. A partir da Figura 26.7 é possível se saber que aminoácidos são codificados por todos os códons. A Figura 26.8 também poderia ser utilizada, mas é menos indicada para a resolução deste exercício.

CGG – arginina (Arg ou R)

UAC – tirosina (Tyr ou Y)

GGA – glicina (Gly ou G)

ACU – treonina (Thr ou T)

UCA – serina (Ser ou S)

GCG – alanina (Ala ou A)

UUA – leucina (Leu ou L)

AAU – asparagina (Asn ou N)

4. A Figura 26.8 é a mais indicada para se saber os códons que codificam um aminoácido específico. Entretanto, a Figura 26.7 também pode ser utilizada.

arginina – AGA, AGG, CGA, CGC, CGG, CGU

histidina – CAC, CAU

metionina – AUG

prolina – CCA, CCC, CCG, CCU

triptofano – UGG

valina – GUA, GUC, GUG, GUU

5. Cada aminoácido é codificado por uma trinca de nucleotídeos presente no mRNA, que é sintetizado a partir de um molde de DNA. Portanto, quando se pretende introduzir qualquer alteração em uma proteína, deve-se introduzir uma alteração no gene (DNA) que a codifica. É importante lembrar que o mRNA é apenas o carreador da informação genética, armazenada sob a forma de DNA, do núcleo das células eucarióticas para o citoplasma, onde ocorre a síntese protéica.

Se cada aminoácido é codificado por uma trinca de bases e se o objetivo é eliminar três aminoácidos da cadeia polipeptídica, será necessária a deleção de três trincas correspondentes aos aminoácidos a serem eliminados, ou seja, de nove pares de bases do DNA. Este DNA servirá de molde para a síntese de uma molécula de mRNA com três trincas a menos, que, então, determinará a síntese de uma proteína com menos três aminoácidos. Qualquer número diferente de três ou múltiplo de três causará uma mudança no quadro de leitura, podendo resultar em síntese de uma proteína menor devido ao surgimento precoce de um códon de terminação. Em geral, nesses casos, a proteína é inativa, portanto, não-funcional.

1. Consideramos síntese de proteínas o processo que resulta em formação de uma molécula funcional. Isso porque muitas proteínas funcionais são compostas por mais de uma cadeia polipeptídica, além de necessitarem de um arranjo espacial específico dessas cadeias. Portanto, a síntese de proteínas engloba a ativação dos aminoácidos, o processo de tradução – que consiste na síntese da cadeia polipeptídica propriamente dita e ocorre em três estágios, iniciação, alongamento e terminação – e o processamento pós-traducional.

2. a) F. A seqüência de aminoácidos em uma cadeia polipeptídica é determinada pela seqüência de códons no mRNA em um processo chamado **tradução**.

b) F. Uma parte da molécula de tRNA contém um anticódon de **três** bases que é complementar ao códon apropriado no mRNA.

c) V.

d) V.

3. O mRNA eucariótico codifica apenas uma proteína e apresenta a extremidade 5'cap, caracterizada pela presença da 7-metil guanosina, que representa o sítio de ligação do ribossomo. Já o mRNA procariótico pode codificar várias proteínas e apresenta seqüências de Shine-Dalgarno, uma para cada uma das proteínas codificadas, que correspondem ao sítio de ligação do ribossomo.

4. Os tRNAs são conhecidos como os adaptadores do processo de tradução. Isso porque são as moléculas que levam o aminoácido correto para o local de síntese. Para desempenhar esse papel, um tRNA deve ser reconhecido por uma enzima específica, que adiciona o aminoácido ativado correto, e pelos ribossomos, para se posicionar corretamente em relação ao mRNA, a fim de parear com o códon que especifica o aminoácido ligado a ele.

5. A ativação dos aminoácidos consiste em duas reações catalisadas pelas aminoacil-tRNA sintetases. A primeira envolve a ligação entre um aminoácido, o AMP, com formação de aminoacil-AMP. A segunda tem como substratos o aminoacil-AMP e o tRNA específico a este aminoácido, resultando em 3' ou 2'-aminoacil-tRNA. O esquema resumido encontra-se na Figura 27.13.

1. O início da tradução sempre se dá com as duas subunidades ribossomais separadas, o que é possível com a participação de fatores específicos, e termina com a hidrólise de GTP e associação da subunidade maior.

As características marcantes dos mRNAs de procariotos e eucariotos determinam as diferenças observadas na etapa inicial de formação do complexo de iniciação nesses organismos. Além disso, em eucariotos, o processo conta com a participação de um número maior de fatores de iniciação.

Em procariotos, o processo se inicia com a ligação da subunidade 30S com o fator de iniciação IF-3, o que impede a ligação precoce entre as subunidades 30S e 50S. A dissociação das duas subunidades ribossomais é, ainda, favorecida pela ligação de IF-1 à porção da subunidade 30S correspondente ao sítio A. Em seguida, observa-se a ligação do mRNA à subunidade menor do ribossoma, com o posicionamento correto do códon de iniciação na parte da subunidade 30S correspondente ao sítio P e a formação do complexo constituído pela subunidade 30S, por IF-3 e pelo mRNA. Esse complexo se torna ainda maior com a ligação de IF-2 previamente ligado ao GTP e ao fMet-tRNA<sup>fMet</sup>. Por fim, esse complexo grande se associa à subunidade ribossomal 50S, estimulando a hidrólise de GTP pelo IF2 e liberação de IF-1 e IF-3. Dessa forma, em procarioto, o complexo de iniciação é formado pelo ribossomo 70S, contendo o mRNA e o iniciador fMet-tRNA<sup>fMet</sup>, que ocupa o sítio P. O sítio A está posicionado para receber o aminoacil-tRNA carregando o segundo aminoácido da cadeia polipeptídica.

Já em eucariotos, o processo se inicia com a ligação do complexo eIF-4F à extremidade 5' cap do mRNA, o que contribui para o desenovelamento de estruturas secundárias dessa molécula. Outro fator, o eIF-3, mantém a subunidade 40S dissociada e promove sua ligação, através da região correspondente ao sítio P, com o complexo ternário formado pelo Met-tRNA<sub>i</sub><sup>Met</sup> e eIF-2 ligado ao GTP. Esse complexo maior, constituído pela subunidade 40S, eIF-s, GTP e Met-tRNA<sub>i</sub><sup>Met</sup>, liga-se à extremidade 5' do mRNA com o auxílio de inúmeros fatores. O mRNA é, então, explorado para localização do primeiro AUG, que sinaliza o início da fase de leitura. Assim, a subunidade 40S do ribossomo se posiciona no AUG iniciador na seqüência apropriada de nucleotídeos flanqueadores. Por fim, com a hidrólise do GTP, ocorre a montagem do complexo de iniciação completo, constituído pelo ribossomo 80S ligado ao Met-tRNA<sub>i</sub><sup>Met</sup> e ao mRNA.

2. a) F. O metionil-tRNA<sub>i</sub><sup>Met</sup> aparece no sítio P.
- b) V.
- c) F. O metionil-tRNA<sup>Met</sup> é usado na incorporação de metionina posicionado na parte interna da cadeia polipeptídica.
- d) F. O mRNA se associa primeiramente com a subunidade 40S.
3. E. O fMet-tRNA<sub>i</sub><sup>Met</sup> é específico para procariotos. Todos os outros itens são necessários para a tradução em procariotos e eucariotos.
4. a) F. O aminoacil-tRNA carregando o aminoácido a ser incorporado na cadeia polipeptídica se liga ao sítio A.
- b) F. A reação de formação da ligação peptídica não requer energia oriunda da hidrólise de GTP.
- c) V.
- d) F. A estreptomicina bloqueia a formação do complexo de iniciação de tradução.
5. Porque a puromicina apresenta uma estrutura similar à tirosinil-tRNA, podendo ser incorporada erroneamente durante a tradução, o que causa uma terminação prematura.

## Aula 29

---

1. Vimos que, após a tradução, as proteínas sofrem inúmeras modificações, incluindo clivagens proteolíticas. Todos os polipeptídeos recém-sintetizados em procariotos têm o grupamento formil retirado por uma desformilase. Muitos deles sofrem ainda a ação de uma amino peptidase que retira um ou poucos resíduos do terminal amino. Assim, nem todas as proteínas bacterianas apresentam metionina neste terminal.
2. E. Durante o processamento da pré-pró-insulina, as seguintes etapas ocorrem, nesta seqüência: remoção do peptídeo sinal, formação de ponte dissulfeto, remoção de um peptídeo de uma região interna da cadeia e dobramento da estrutura resultando na conformação funcional. A formação de  $\gamma$ -carboxiglutamato é característica da síntese de alguns fatores de coagulação.

3. O principal elemento de sinalização é a seqüência sinal (ou líder), também chamada peptídeo sinal. Essa seqüência sinalizadora é responsável pelo direcionamento do polipeptídeo a um local específico na célula, sendo geralmente removida durante o processo de transporte ou quando o polipeptídeo atinge seu destino final.

4. Relembrando as aulas de Biologia Celular sobre a constituição das membranas celulares, entendemos que a presença dos resíduos hidrofóbicos é fundamental para o transporte das proteínas. Esses resíduos podem, portanto, direcionar a proteína ao compartimento apropriado e iniciar sua penetração nas membranas.

## Aula 30

---

1. O genoma de uma espécie corresponde ao conjunto de todo o material genético presente em suas células. O genoma, portanto, inclui não apenas os genes, mas também diversas regiões não-codificadoras, que são componentes do material genético. Assim, a definição apropriada para genoma é o conjunto de todo o material genético que caracteriza uma espécie.

Comentário: este exercício fará com que você revise a definição do termo genoma, considerando a existência de muitas outras seqüências de DNA, além dos genes, dentre o material genético de cada organismo.

2. O tamanho do genoma de uma espécie refere-se ao conteúdo de DNA (Valor-C) que ele possui. Tal valor pode ser medido em pares de bases. A complexidade de um genoma refere-se ao número de genes diferentes presentes nele.

Comentário: para responder a esta questão, você deve ter assimilado as diferenças entre os dois termos conforme apresentados no início desta aula.

3. A correlação entre o conteúdo de DNA do genoma haplóide e a complexidade funcional dos organismos pode ser facilmente verificada em procariotos. Ao analisarmos o conteúdo de DNA do genoma das espécies pertencentes a tais filós, percebemos uma correlação direta entre complexidade funcional dos organismos e conteúdo de DNA de seus genomas.

Em eucariotos, entretanto, tal correlação não é verificada, pois espécies com níveis similares de complexidade funcional podem apresentar conteúdos de DNA completamente distintos em seus genomas. Paralelamente, espécies com níveis de complexidade funcional completamente distintos podem apresentar conteúdos de DNA similares em seus genomas.

Comentário: ao responder a esta questão, você terá a oportunidade de revisar as correlações entre a complexidade funcional dos organismos e o conteúdo de DNA em seus genomas, nos diferentes filos.

4. Os ensaios de cinética de reassociação baseiam-se no princípio de que, em uma mistura de fragmentos de DNA, previamente desnaturados termicamente, a posterior reassociação das moléculas dependerá da probabilidade de cada fragmento fita-simples encontrar sua fita complementar. Tal evento ocorre ao acaso. Se, entretanto, houver várias cópias de um dado fragmento, as chances de que suas fitas encontrem uma fita complementar será aumentada, elevando, com isso, a velocidade de reassociação. Em outras palavras, quanto maior a quantidade de cópias de um dado fragmento, maior será sua velocidade de reassociação. Portanto, tais ensaios permitem estimar a presença de DNA repetitivo em um genoma, pois a fração de fragmentos que possuem várias cópias apresentará maior velocidade de reassociação.

Comentário: este exercício fará com que você revise os princípios envolvidos nos ensaios de cinética de reassociação de fragmentos de DNA, de modo a entender o método de determinação de frações de DNA repetitivo nos diferentes genomas.

5. Uma considerável parcela do genoma de eucariotos superiores é composta por DNA repetitivo. Determinadas espécies chegam a apresentar mais de 80% de DNA repetitivo em seu genoma. Em eucariotos inferiores, a proporção de DNA repetitivo é bem menor, não ultrapassando 20% do genoma. Organismos procariotos não possuem DNA repetitivo.

Comentário: para responder a esta questão, você deverá revisar as informações referentes à presença de DNA repetitivo nos genomas dos diferentes organismos.

6. Não. Embora a grande maioria dos genes de eucariotos esteja contida na fração de DNA não-repetitivo, existem diversas seqüências de DNA não-repetitivo que não são genes. Tais seqüências podem ser regiões intergênicas, ou seja, seqüências de DNA localizadas entre dois genes adjacentes. No genoma de procariotos, por exemplo, todo o genoma é composto por DNA não-repetitivo, mas apenas uma parcela dele é representada por genes.

Comentário: neste exercício, você deverá revisar as informações acerca da fração de DNA não-repetitivo dos genomas.

7. Mitocôndrias e cloroplastos possuem genomas circulares, existindo várias cópias do genoma por organela. Os genes presentes codificam para os RNAs ribossômicos e transportadores e para apenas algumas das proteínas requeridas pela organela.

As demais proteínas são codificadas pelo núcleo, sendo endereçadas para a organela após serem sintetizadas.

Genomas de mitocôndrias variam amplamente de tamanho entre 16,5kb, em mamíferos, e 570kb, em plantas superiores. Genomas de cloroplastos variam entre 120 e 200kb.

Comentário: neste exercício, você deverá rever as informações relativas às características gerais dos genomas de cloroplastos e mitocôndrias.

## Aula 31

---

1. Para que um organismo apresente maior complexidade funcional e morfológica, ele deverá possuir genes que coordenem tais características. De fato, o número de genes presentes em eucariotos superiores é muito maior que o observado em procariotos. Por exemplo, o genoma humano possui aproximadamente 10 vezes mais genes que o genoma da bactéria *E. coli*.

Comentário: este exercício fará com que você revise as informações referentes à correlação entre número e tamanho médio de genes e à complexidade funcional dos organismos.

2. As diferenças morfológicas não dependem apenas da presença ou ausência de genes. Muitas vezes, a regulação diferencial de um mesmo conjunto de genes pode resultar em fenótipos completamente distintos. Nesses casos, a ação de proteínas regulatórias é crucial para a morfogênese. Assim, a presença, ou ausência, de uns poucos genes reguladores pode resultar em grande variação fenotípica e, muitas vezes, funcional entre duas espécies.

Comentário: para responder a esta questão, você deve revisar os conceitos relativos à expressão diferencial de genes e seu papel na morfologia dos indivíduos.

3. O número restrito de genes no genoma humano exige a existência de mecanismos que elevem sua complexidade funcional. Assim, a regulação diferencial dos genes presentes passa a ter um papel importantíssimo. Dentre estes mecanismos, a emenda alternativa de éxons desempenha um papel crucial. A análise do genoma humano tem permitido estimar que aproximadamente 40% dos genes sejam regulados por emenda alternativa de éxons. Assim, tais genes podem gerar transcritos e, conseqüentemente, proteínas diferentes em função do tecido, condição ou estágio de desenvolvimento. Este processo permite que diferentes funções sejam desempenhadas por cada gene, resultando na amplificação do número de proteínas distintas codificadas pelo genoma.

Comentário: este exercício fará com que você revise as informações referentes aos mecanismos de emenda alternativa de éxons e o seu papel na geração de proteínas distintas a partir de um mesmo gene. Em seguida, você deve associar tais informações com a complexidade funcional dos organismos.

4. Dados fornecidos pela análise do genoma humano, bem como dos genomas de plantas, têm revelado que grande parte do DNA repetitivo de eucariotos é composta por elementos de transposição ou seqüências remanescentes destes. Na espécie humana, por exemplo, 50% do genoma é composto por tais seqüências.

Comentário: para responder a este exercício, você deverá revisar os conceitos relativos a elementos de transposição, bem como a participação destes na composição dos genomas de eucariotos.

5. A análise do genoma humano revelou que 75% do nosso genoma é composto por regiões que não contêm genes, 24% do genoma é representado por íntrons e apenas 1,1% do genoma é composto por éxons. Portanto, apenas 1,1 % do genoma de humanos codifica proteínas, enquanto os demais 98,9% são compostos por íntrons e regiões não-codificadoras.

Comentário: este exercício fará com que você fixe as informações relativas às proporções entre regiões codificadoras e não-codificadoras do genoma humano.

6. A comparação das seqüências de DNA dos genomas de indivíduos da espécie humana revelou níveis de diferenças inferiores a 0,1% dos nucleotídeos. Níveis similares foram obtidos quando os indivíduos analisados pertenciam a uma mesma "raça" ou quando eram de "raças" diferentes. Assim, ficou evidenciado que o conceito de "raça" não se aplica à espécie humana, pois indivíduos da mesma raça são tão parecidos entre si (99,9% de identidade de seqüência de nucleotídeos) quanto indivíduos de raças distintas.

Comentário: para desenvolver este exercício, você deverá rever as informações relativas à similaridade genômica entre indivíduos de uma mesma espécie.

7. Após a obtenção da seqüência de nucleotídeos de um genoma e a localização das regiões codificadoras, um importante passo é a caracterização da regulação e da função de cada gene. A determinação de tais informações será essencial para o estudo da expressão coordenada destes genes, bem como a interação entre seus produtos e o ambiente.

Comentário: este exercício fará com que você revise a importância de se conhecer as seqüências de nucleotídeos dos genomas e a função dos genes neles presentes.



ISBN 85-89200-77-9



9 788589 200776



**UENF**  
Universidade Estadual  
do Norte Fluminense



Universidade Federal Fluminense



Fundação Carlos Chagas Filho de Amparo  
à Pesquisa do Estado do Rio de Janeiro



**GOVERNO DO  
Rio de Janeiro**

SECRETARIA DE  
CIÊNCIA E TECNOLOGIA

Ministério  
da Educação

