



# 4

## **GENÉTICA MOLECULAR**

*MIRNA DUARTE BARROS*

---



## Capítulo 4

# GENÉTICA MOLECULAR

*MIRNA DUARTE BARROS*

### INTRODUÇÃO

O dogma central da Genética Molecular define um paradigma que envolve polímeros de nucleotídeos, o ácido desoxirribonucléico, o DNA e o ácido ribonucléico, o RNA e polímeros de aminoácidos, as proteínas.

O dogma central pode ser resumido com esquema a seguir:



O DNA contém ao longo da seqüência das bases nitrogenadas de seus nucleotídeos, as mensagens para a vida, para o funcionamento do maquinário necessário para o funcionamento celular e do indivíduo. Essas mensagens não precisam e não são lidas simultaneamente. Assim como lemos um livro, palavra por palavra, parágrafo por parágrafo, o DNA também é lido por parte ao longo da vida.

Cada segmento de DNA que codifica uma proteína é denominado de gene. Neste momento, vários genes estão sendo transcritos e traduzidos em nossas células.

O processo de duplicação perpetua o material genético ao longo das gerações celulares do indivíduo, fornecendo às novas gerações “cópias idênticas” à de origem.

A transcrição fornece um ácido nucléico de cadeia simples, ao contrário do DNA, que é formado por uma cadeia dupla. O DNA, nos eucariontes, está restrito ao núcleo celular. O RNA, complementar ao DNA, leva ao citoplasma, local da síntese protéica, a mensagem genética que é traduzida da linguagem das bases nitrogenadas dos ácidos nucléicos, para a linguagem dos aminoácidos das proteínas, que irão efetuar as atividades celulares.

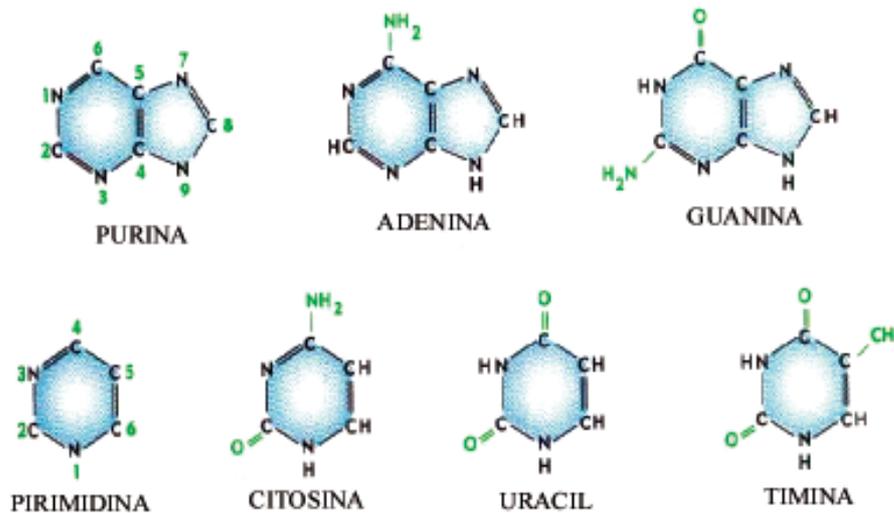
### ESTRUTURA DO DNA

O DNA é a principal força determinante de todo o ser vivo, pois a informação que ele contém dirige toda a complexidade da vida.

Apesar de sua importância, o DNA permaneceu desconhecido durante muito tempo; somente em 1869, Miescher detectou pela primeira vez o DNA. Ele usou como fonte de núcleos, os núcleos de glóbulos brancos contidos no pus, por serem grandes, facilitando seu isolamento. Sua análise revelou um composto de carácter ácido, rico em fósforo, organizado em grandes moléculas, que foi

denominado de **nucleína**. Em 1874, o próprio Miescher, utilizando técnicas mais avançadas, observou que estas grandes moléculas de nucleína eram ricas em nitrogênio, além do fósforo.

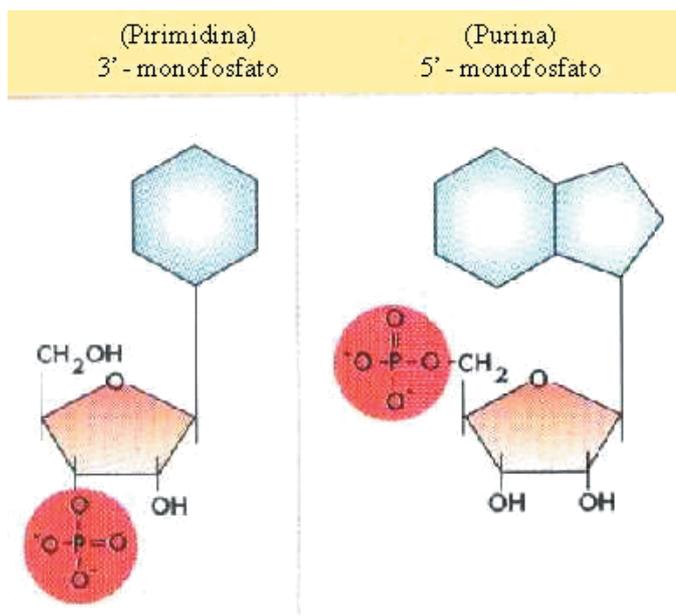
Em 1880, Kossel demonstrou que as nucleínas eram formadas por unidades estruturais: as **bases nitrogenadas**. Em 1889, com técnicas que permitiam isolar as nucleínas de forma pura, Altmann separou-as das proteínas e chamou-as de **ácidos nucléicos**. (Figura 4.1)



**Figura 4.1:** Bases Nitrogenadas, componentes dos nucleotídeos dos ácidos nucléicos. (Adaptado de Lewin, B. Gene V. Oxford Press, 1994).

Com os métodos de preparações puras de ácidos nucléicos, foi demonstrado que a degradação destes ácidos liberava quatro tipos de bases nitrogenadas, **adenina** e **guanina**, conhecidas como bases púricas e **citossina** e **timina**, denominadas de bases pirimídicas.

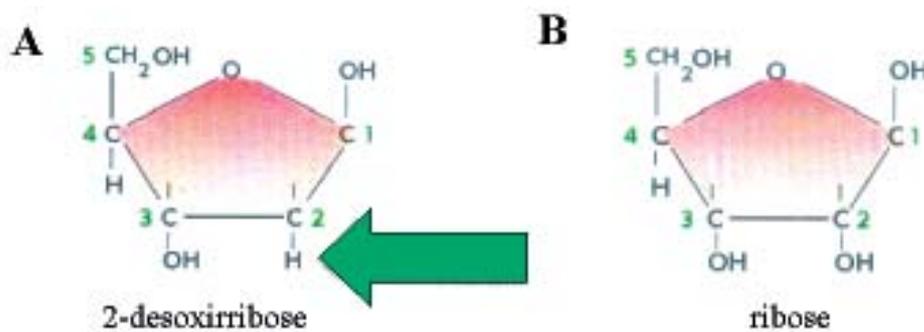
Em 1912, Levene e Jacobs determinaram que o componente básico dos ácidos nucléicos era uma estrutura composta de um açúcar com 5 carbonos, uma pentose, ligada a uma base nitrogenada e a um fosfato, que denominaram de **nucleotídeo**. (Figura 4.2).



**Figura 4.2:** O nucleotídeo é formado de uma pentose, uma base nitrogenada ligada ao carbono 1' e de um grupo fosfato, que pode ligar-se ao carbono 3' ou 5' da pentose. (Adaptado de Lewin, B. Gene V. Oxford Press, 1994).

Nesta mesma época, tornou-se evidente que existiam dois tipos de ácidos nucléicos: o ácido extraído de timo de bezerro e o extraído de células de levedura. Quando degradava-se por hidrólise (quebra molecular pela ação da água), ácidos nucléicos provenientes de levedura, observou-se a liberação de uma base nitrogenada diferente daquelas que apareciam até então, denominada de **uracila** e nunca havia a liberação de timina. O ácido nucléico do timo não tinha uracila e o de levedura não tinha timina; a pentose do ácido nucléico do timo tinha um oxigênio a menos do que a pentose do ácido da levedura, portanto o primeiro foi chamado de **ribose** e o segundo de **desoxirribose**.

O ácido nucléico que continha desoxirribose e timina foi denominado **ácido desoxirribonucléico** (DNA ou ADN), e o que apresentava uracila e ribose de **ácido ribonucléico** (RNA ou ARN). (Figura 4.3).



**Figura 4.3:** Estrutura molecular das pentoses dos ácidos nucléicos. **A-** desoxirribose (DNA). **B-** ribose (RNA). (Adaptado de Lewin, b. **Gene V**. Oxford Press, 1994).

Até a algumas décadas atrás, não se podia afirmar o papel exato do DNA das células. Sugeriu-se, através de evidências experimentais que o DNA seria a molécula responsável pela manutenção e passagem através das gerações das informações das características hereditárias. Mas restavam dúvidas, já que alguns pesquisadores acreditavam serem as proteínas as moléculas responsáveis pela hereditariedade.

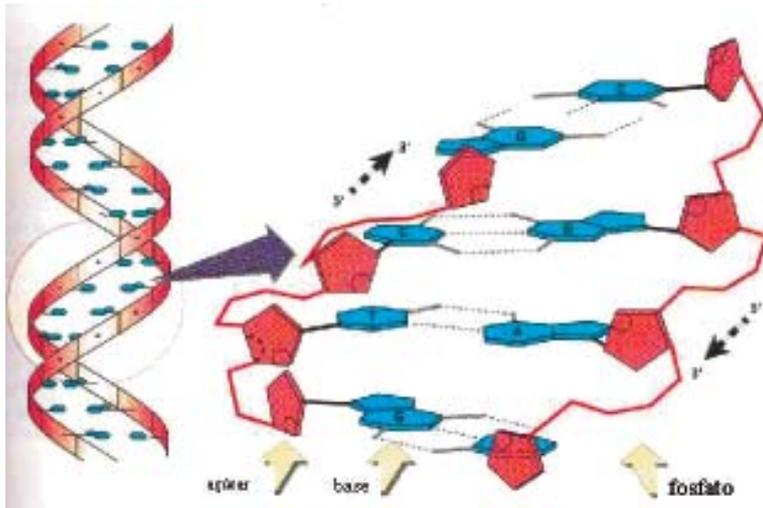
O experimento de Alfred Hershey e Martha Chase, em 1952, trouxe a confirmação de que a molécula de DNA era o material hereditário. Eles utilizaram a propriedade já conhecida das proteínas de não conterem fósforo (P) e do DNA não conter enxofre (S). Cultivaram bactérias em meio de cultura contendo fósforo radioativo ( $P^{32}$ ) e enxofre radioativo ( $S^{35}$ ). Em seguida, estas bactérias foram infectadas com vírus **bacteriófagos**. Os vírus que se desenvolveram nestas bactérias apresentavam suas proteínas marcadas com  $S^{35}$  e seus ácidos nucléicos marcados com  $P^{32}$ . Estes bacteriófagos marcados foram utilizados para infectar bactérias de culturas normais, não radioativas. Após a infecção, as culturas foram agitadas em um liquidificador, desfazendo as ligações das capas protéicas virais adsorvidas às paredes das bactérias.

Praticamente todo o fósforo radioativo incorporado no vírus era transmitido para o interior da bactéria que ele infectara, aparecendo na progênie produzida. O enxofre radioativo, por sua vez, não penetrava na bactéria e não aparecia na progênie viral. Concluiu-se portanto, que apenas o DNA do vírus penetra na bactéria quando da infecção e, a partir dele, a nova geração viral é produzida, com as informações que ele contém.

Com a confirmação do DNA como material hereditário, muitos pesquisadores interessaram-se no estudo desta molécula. Estudaram vários DNA de diferentes origens e observaram que a porcentagem de timina era semelhante a de adenina, enquanto a porcentagem de citosina era semelhante a de guanina.

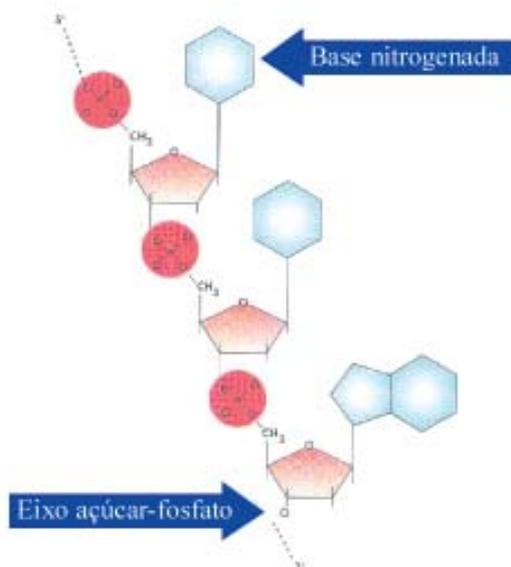
$$\frac{A}{T} = 1 \quad \frac{G}{C} = 1 \quad \frac{A+G}{T+C} = 1$$

Em 1953, Watson e Crick apresentaram um modelo espacial para a molécula de DNA, que é aceito praticamente na íntegra até hoje. (Figura 4.4).



**Figura 4.4:** Estrutura tridimensional da dupla hélice da molécula de DNA. (Adaptado de Lewin, B. *Gene V*. Oxford Press, 1994).

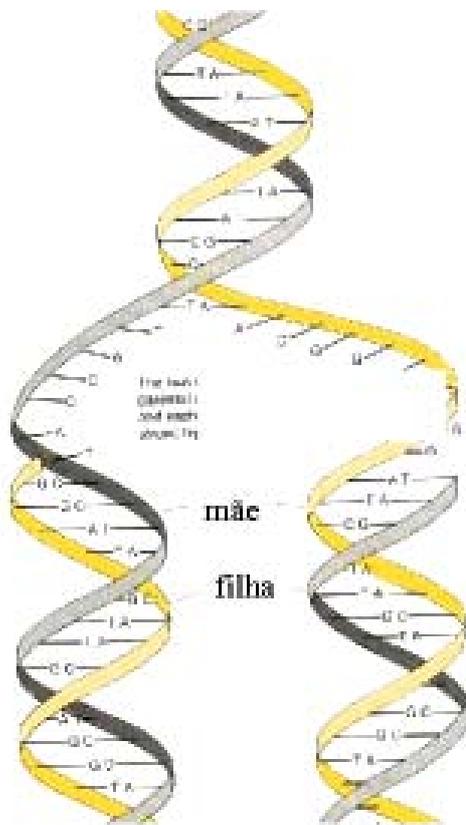
O modelo propôs que a molécula de DNA é constituída de duas cadeias helicoidais dispostas ao redor de um eixo central imaginário. Cada cadeia é constituída por unidades de desoxirribose unidas por ligações diéster-fosfato entre os seus carbonos 3' e 5'. Uma base nitrogenada está ligada ao carbono 1' da pentose e as duas cadeias mantêm-se unidas por pontes de hidrogênio entre as bases nitrogenadas: duas entre adenina e timina e três entre citosina e guanina (Figura 4.5).



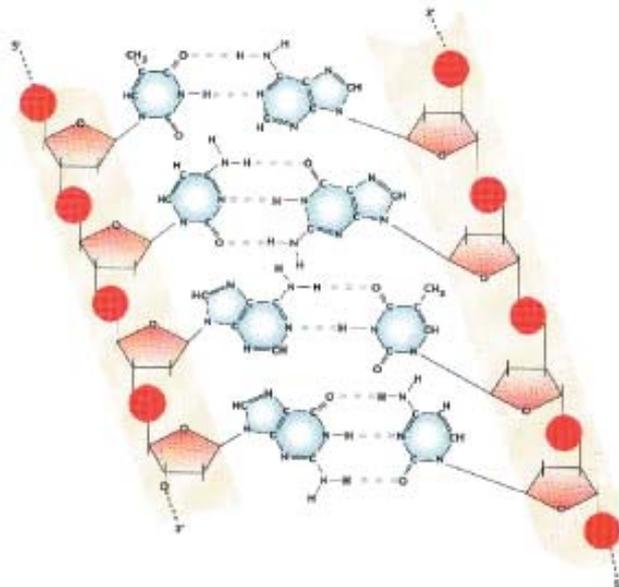
**Figura 4.5:** Cadeia polinucleotídica, com destaque para a ligação diéster-fosfato entre os carbonos 5' e 3' das pentoses. (Adaptado de Lewin, B. *Gene V*. Oxford Press, 1994).

As duas cadeias helicoidais da molécula de DNA são **complementares** entre si, ou seja, quando em uma cadeia a seqüência for adenina-citosina-timina-adenina, na outra cadeia a seqüência será obrigatoriamente timina-guanina-adenina-timina. Além de complementares entre si, as duas cadeias do DNA são **anti-paralelas**, ou seja correm em sentidos opostos. Quando em uma cadeia o carbono livre da pentose é o 3', o carbono livre da outra cadeia é o 5' e vice-versa.

Este modelo permitiu imaginar um mecanismo de autoduplicação para a molécula de DNA, onde as pontes de Hidrogênio rompidas, permitiriam o desenrolamento e a separação das cadeias helicoidais. Cada cadeia serve de molde para a síntese de uma nova cadeia helicoidal complementar ao molde, resultando em duas moléculas filhas, cada uma contendo uma cadeia proveniente da molécula mãe e uma cadeia recém sintetizada, configurando o que se denomina de duplicação **semiconservativa** do DNA (Figuras 4.6 e 4.7).



**Figura 4.6:** Duplicação semiconservativa do DNA. (Adaptado de Griffith A. J. *et al.* **An introduction to genetic analysis**. W. H. Freeman & Cia. 1993).



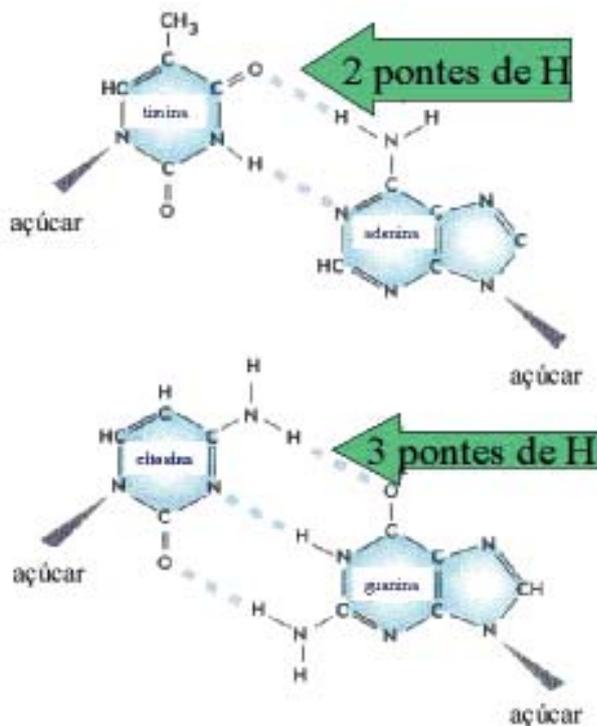
**Figura 4.7:** Esquema da estrutura das cadeias antiparalelas do DNA. Enquanto uma cadeia apresenta o sentido 5' - 3', a outra está no sentido 3' - 5'. (Adaptado de Lewin, B. **Gene V**. Oxford Press, 1994).

## ÁCIDOS NUCLÉICOS: COMPONENTES

Os ácidos nucleicos são grandes moléculas, constituídas de subunidades arranjadas seqüencialmente. Cada subunidade é um **nucleotídeo** constituído por uma **pentose** (um açúcar de 5 carbonos em forma de anel), uma **base nitrogenada** (um anel de átomos carbono e nitrogênio) e um grupo **fosfato**.

Os nucleotídeos são diferenciados pelo tipo de pentose ou pelo tipo de base nitrogenada que os constituem. No DNA, o açúcar é a **desoxirribose** e no RNA é a **ribose**, que diferem entre si somente pela presença de um grupo hidroxila no carbono 2 da ribose.

Quanto às bases nitrogenadas, elas podem ser classificadas em dois grupos: as **púricas** e as **pirimídicas**. As primeiras são formadas por anéis fundidos de 6 e 5 átomos, enquanto as últimas são constituídas por um anel de 6 átomos. Há duas purinas, **adenina** e **guanina**, encontradas tanto em ribonucleotídeos quanto em desoxirribonucleotídeos e três bases pirimídicas: **citossina**, que ocorre tanto no DNA quanto no RNA, **timina**, exclusiva do DNA e **uracil** ou **uracila**, exclusiva do RNA. Ao citar as bases nitrogenadas, usa-se apenas a inicial delas em maiúscula, então temos **A**, **G**, **C**, **T**, e **U**. (Figura 4.8).



**Figura 4.8:** Pontes de Hidrogênio entre adenina e Timina (duas) e Citosina e Guanina (três). (Adaptado de Lewin, B. *Gene V*. Oxford Press, 1994).

Os carbonos das pentoses são numerados de 1 a 5, iniciando pelo carbono que se liga covalentemente com o Nitrogênio 1 das bases pirimídicas e com o Nitrogênio 9 das púricas. Como os átomos dos anéis das bases nitrogenadas também são numeradas, a numeração dos carbonos das pentoses é seguida do símbolo (<sup>o</sup>), evitando ambigüidades. Portanto, a **base nitrogenada** está na posição 1', e o grupo fosfato na 5'. A ligação com o grupo fosfato com um **nucleosídeo** (pentose mais base nitrogenada) compõe um **nucleotídeo**.

Os nucleotídeos unidos entre si formam uma cadeia de **polinucleotídeos**. A ligação entre um nucleotídeo e seu vizinho faz-se sempre através de ligação covalente do tipo fosfo-diéster, envolvendo o carbono 5' de um nucleotídeo o 3' do outro. Em uma cadeia polinucleotídica, um dos nucleotídeos terminais apresenta o grupo 5' livre, enquanto na outra extremidade, o grupo 3' está livre. Observando a dupla hélice do DNA, uma apresenta a direção 5'→3' e a outra, a direção 3'→5'. Então, cadeias polinucleotídicas do DNA são **antiparalelas**, ou seja, colocam-se em direções opostas uma à outra. (Figura 4.7).

A estrutura molecular do DNA revelou o modelo de duplicação ou replicação da molécula, já que há complementaridade entre as bases nitrogenadas das cadeias da dupla hélice (para detalhes, veja o item seguinte deste capítulo). Além disso, a estrutura seqüencial das bases nitrogenadas condiciona a seqüência de aminoácidos das proteínas. Então, a seqüência de bases nitrogenadas revela um código de inscrição de informações que serão traduzidas para a linguagem de aminoácidos na **síntese protéica**.

A estrutura molecular proposta por Watson & Crick em 1953 para o DNA é aceita até hoje. O DNA é formado por duas cadeias que se dispõem **em dupla hélice** ao redor de um eixo central imaginário. Cada hélice ou cadeia é formada por subunidades: os nucleotídeos. Portanto, podemos afirmar que o DNA é um polímero de nucleotídeos. Esta estrutura espacial do DNA é independente da seqüência particular de nucleotídeos que a molécula contém, apesar da seqüência de nucleotídeos ser crucial na determinação da informação genética (Figura 4.6).

A molécula de DNA é constituída de uma hélice regular, que faz uma volta completa a cada 34 Å (3,4 nm). Como a distância média entre os nucleotídeos é de 3,4 Å, cada volta da molécula contém 10 nucleotídeos. A hélice desenvolve-se à direita, ou seja, no sentido horário.

As ligações entre nucleotídeos da mesma cadeia são covalentes do tipo fosfo-diéster, onde o grupo fosfato forma uma ponte entre os grupos -OH de duas desoxirriboses vizinhas.

A conformação espacial de dupla hélice é mantida em função das **pontes de Hidrogênio** entre bases nitrogenadas de ambas as cadeias. As pontes de Hidrogênio ocorrem entre os átomos de Hidrogênio com pequena carga positiva e um átomo receptor com uma pequena carga negativa. Cada átomo de Hidrogênio do grupo NH<sub>2</sub> é levemente positivo, pois o Nitrogênio tem a tendência de atrair os elétrons da ligação N-H, enquanto o de Oxigênio com dupla ligação tem elétrons ao seu redor, conferindo uma leve negatividade.

A ponte de Hidrogênio forma-se entre o **H** e o **O** dos pares de bases das duas cadeias de DNA. Apesar da ponte de Hidrogênio não ser uma ligação muito forte, quando comparada a uma ligação covalente, sustenta a dupla hélice de DNA e desempenha papel importante na hereditariedade. A ligação ocorre sempre entre adenina e timina com duas pontes de Hidrogênio, enquanto que entre guanina e citosina formam-se três pontes. O DNA que contém maior quantidade de pares **G - C** é mais estável do que o que contém predominantemente o par **A - T**. A ligação entre as bases nitrogenadas é denominada de **pareamento das bases** e é **complementar**, já que guanina só parecia com citosina e adenina com timina.

As ligações das bases nitrogenadas ocorrem sempre entre uma base púrica e uma pirimídica, proporcionando um diâmetro constante ao longo de toda a molécula. Ligações entre duas bases púricas resultariam em uma molécula extremamente espessa, enquanto que a ligação entre duas bases pirimídicas em uma molécula muito delgada.

Em uma visão tridimensional, as bases nitrogenadas dispõem-se formando degraus entre as cadeias do DNA, que aproximam-se umas das outras na formação helicoidal da molécula. A aproximação entre as bases nitrogenadas, como uma gaveta sobre outra, aumenta a estabilidade da molécula, pois exclui moléculas de H<sub>2</sub>O que ocupariam o espaço entre uma base e outra.

Podem ser identificadas duas formas de DNA por análise de difração, quanto ao grau de hidratação da molécula. A forma **A** é menos hidratada e mais compacta que a forma **B**. Esta última é a encontrada com maior frequência nas células.

## DUPLICAÇÃO DO DNA

A **duplicação do DNA** é a síntese de uma nova molécula de DNA, através de um processo de polimerização de nucleotídeos, mediada por enzimas específicas. O processo ocorre com a presença da molécula de DNA molde ou molécula mãe, já pré-existente na célula e precede a divisão celular, tanto nos procariontes como nos eucariontes, para a manutenção da quantidade de DNA. Isto só não ocorre na **meiose**, cujo produto final são células com a metade do DNA da célula inicial.

Nos **eucariontes**, a duplicação do DNA é um processo que ocorre principalmente no interior do núcleo, já que o DNA das mitocôndrias apesar de duplicar-se, acompanhando o DNA nuclear, representa pequena quantidade em relação ao DNA nuclear. Uma célula humana pode levar de seis a dezenas de horas para completar a duplicação de seu DNA.

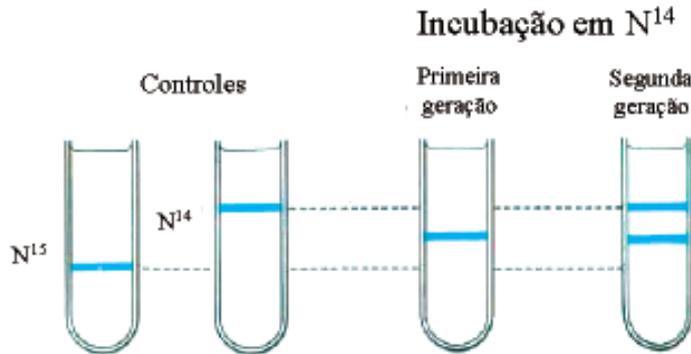
É primordial que o processo de duplicação da molécula de DNA seja preciso. Para tanto, um dos mecanismos é ter a molécula inicial como molde para a síntese da nova molécula. Na realidade, cada cadeia da dupla hélice serve de molde separadamente e é denominada de **template**. A separação das cadeias ocorre como uma abertura de um zíper, com a quebra das pontes de hidrogênio entre as bases nitrogenadas. Se as ligações entre as bases fossem covalentes, o processo exigiria uma quantidade maior de energia para rompê-las.

As moléculas filhas são sintetizadas complementares às cadeias da molécula mãe de DNA. Uma adenina na cadeia parental determina uma timina na cadeia filha; uma guanina na cadeia parental determina uma citosina na cadeia filha e assim por diante. Ao final da duplicação, existirão duas moléculas de DNA idênticas quanto à seqüência das bases nitrogenadas, mas cada molécula será constituída de uma cadeia proveniente da molécula parental e uma cadeia recém-sintetizada. Este modelo de duplicação do DNA é denominado de **semiconservativo**, pois a nova molécula de DNA mantém metade da estrutura da molécula parental. A ligação entre gerações seqüenciais do DNA é a permanência de uma das cadeias parentais na molécula da geração seguinte.

Este padrão de duplicação foi confirmado experimentalmente por Meselson & Stahl, em 1958. Eles cultivaram bactérias *Escherichia coli* em um meio com  $N^{15}$ , um isótopo pesado. O DNA destas bactérias era mais denso e após ser isolado e centrifugado, observava-se uma banda denominada de **pesada** ou **parental**. O DNA de bactérias cultivadas em meio com  $N^{14}$ , apresentava densidade menor e após centrifugação formava uma **banda leve**.

As bactérias do meio com  $N^{15}$  foram transferidas para um meio de cultivo com  $N^{14}$ . Após uma geração de duplicação bacteriana, o DNA foi centrifugado, resultando em uma **banda intermediária** entre a parental e a leve. O DNA da segunda geração apresentou uma banda híbrida e uma leve. Portanto, na primeira geração bacteriana, todo o DNA era constituído de uma cadeia com bases nitrogenadas contendo  $N^{15}$  e uma cadeia recém sintetizada, mais leve, com  $N^{14}$ , resultando na banda híbrida. Na segunda geração, parte do DNA ainda continha

uma cadeia com  $N^{15}$  e outra com  $N^{14}$  e parte com cadeias só com  $N^{14}$  (Figura 4.9). Esses resultados confirmaram que a síntese do DNA é semiconservativa e corroboraram o modelo estrutural e de duplicação do DNA proposto por Watson & Crick em 1953.



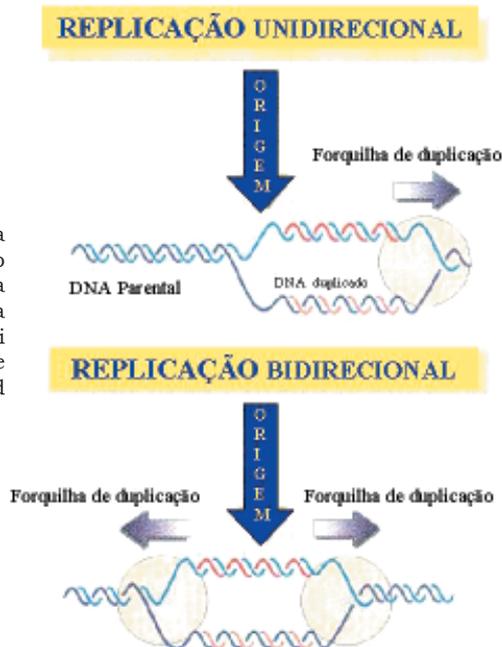
**Figura 4.9:** Centrifugação do DNA em um gradiente de cloreto de césio. DNA com  $N^{15}$  aparece como uma banda pesada; DNA com  $N^{14}$  encontra-se na banda leve; a banda intermediária representa o DNA híbrido. (Adaptado de Griffith A. J. *et al. An Introduction to Genetic Analysis*. W. H. Freeman & Cia. 1993).

A unidade ou porção do DNA que replica é chamada de **unidade de replicação** ou **replicon**. O replicon possui uma **origem**, onde a duplicação inicia e pode ter um **ponto de término**, onde ela finaliza. Cada replicon controla os elementos para a síntese da nova cadeia de DNA e funciona uma única vez em cada ciclo de divisão celular.

O genoma dos **procariontes** é composto por uma única molécula de DNA, portanto só há um replicon. Entretanto, cada cromossomo dos eucariontes contém um grande número de replicons.

Na região onde a duplicação inicia, há a separação das cadeias do DNA, formando uma figura em forma de **Y**. Esta região é chamada **forquilha de duplicação**, que move-se seqüencialmente ao longo do DNA, partindo do ponto de origem do replicon. A duplicação pode ser **uni** ou **bidirecional**. No primeiro caso, apenas uma forquilha de duplicação forma-se a partir da origem, enquanto que no segundo caso, duas forquilhas são formadas, ambas partindo da origem, mas em direções opostas (Figura 4.10).

**Figura 4.10:** Representação da separação da cadeia dupla do DNA durante a duplicação da molécula, formando a forquilha de duplicação, que pode ser uni ou bidirecional. (Adaptado de Lewin, B. **Gene V**. Oxford Press, 1994).



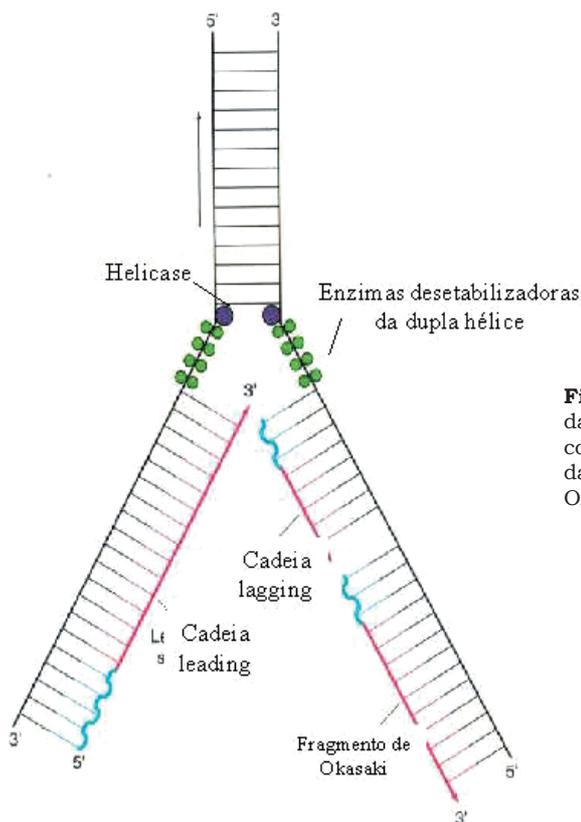
## ENZIMAS DA DUPLICAÇÃO DO DNA

A duplicação do DNA é um evento complexo, mediado por enzimas e envolve três estágios:

1. **Iniciação:** ocorrem o reconhecimento da origem do replicon, a abertura da dupla hélice e a estabilização da cadeia única do DNA.
2. **Elongação:** síntese da nova cadeia do DNA ao longo da forquilha de duplicação.
3. **Término ou Parada:** encerramento da síntese e separação das novas moléculas já duplicadas.

Para que haja o início da síntese das novas cadeias de DNA, é necessário romper as ligações que mantêm a dupla hélice. A enzima **helicase** age na forquilha de duplicação, rompendo as pontes de hidrogênio. A hidrólise de ATP libera energia, que modifica a conformação espacial da helicase, provocando o rompimento das pontes de hidrogênio e a separação da dupla hélice.

O DNA é estável quando a molécula forma a dupla hélice. Para que a cadeia simples seja mantida após a ação da helicase, as **enzimas desestabilizadoras da dupla hélice** ligam-se às cadeias simples do DNA, esticando-as e impedindo a reunião das cadeias, deixando espaço e tempo para a ação das **enzimas polimerizadoras**. (Figura 4.11)

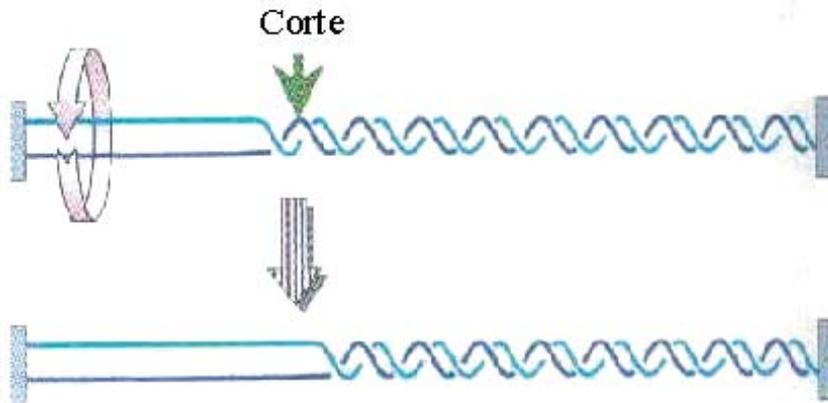


**Figura 4.11:** Enzimas que participam da abertura da cadeia dupla do DNA, no início de sua síntese, como a helicase e as enzimas desestabilizadoras da dupla hélice. (Adaptado de Lewin, B. **Gene V**. Oxford Press, 1994).

A contínua abertura da dupla hélice pela helicase, requer a rotação de uma cadeia sobre a outra. Como as extremidades do DNA não são livres a medida que há o desenrolamento a partir de um ponto, há um **superenrolamento** no restante da molécula. O superenrolamento é relaxado pela ação das enzimas

**topoisomerases**, que induzem a quebra de uma das cadeias do DNA, ligam-se à extremidade livre da cadeia rompida e giram ao redor da cadeia que ficou intacta, diminuindo a tensão da molécula. A enzima desliga-se e a cadeia é soldada novamente, permitindo a continuação da abertura da forquilha de duplicação. (Figura 4.12).

## Corte, Rotação e Ligação

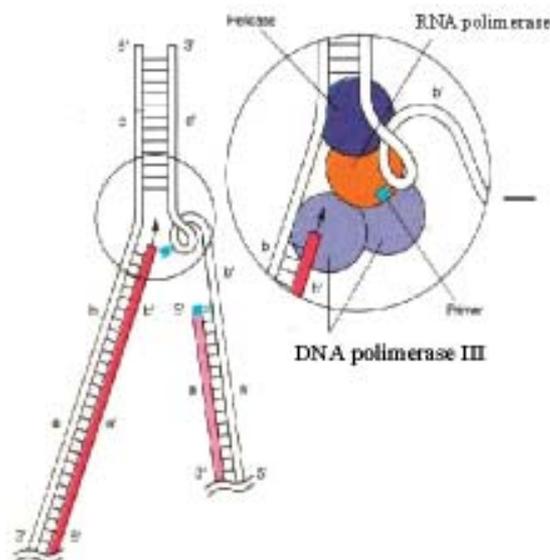


**Figura 4.12:** Ação das topoisomerases para diminuir a tensão de enrolamento da molécula de DNA. (Adaptado de Lewin, B. **Gene V**. Oxford Press, 1994).

As enzimas que efetivamente sintetizam a nova molécula de DNA são denominadas de **DNA polimerases**. A primeira DNA polimerase foi isolada e identificada em uma bactéria, *E. coli*, por Kornberg em 1956. Posteriormente, foram identificadas três DNA polimerase: **DNA pol I**, **DNA pol II** e **DNA pol III**.

Tanto procariontes, quanto eucariontes, possuem DNA polimerase com o mesmo tipo fundamental de atividade sintética, ou seja, todas as DNA pol acrescentam novos nucleotídeos na **extremidade 3'-OH** da desoxirribose. Portanto, o crescimento da molécula de DNA sempre ocorre no sentido **5' → 3'**.

Apesar das três DNA pol possuírem atividade sintética **5' → 3'**, a **DNA pol III** é a principal responsável pela duplicação do DNA. É uma enzima multimérica, ou seja, é composta por várias subunidades e age na forquilha de duplicação (Figura 4.13).

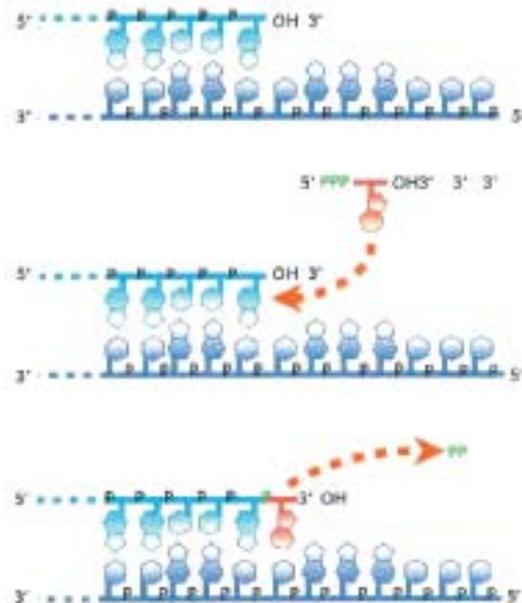


**Figura 4.13:** Esquema da ação das enzimas atuantes na forquilha de duplicação. (Adaptado de Lewin, B. **Gene V**. Oxford Press, 1994).

Tanto a DNA pol I, quanto a DNA pol II, apresentam a capacidade de retirar nucleotídeos tanto da extremidade 5' quanto da 3', por isso são denominadas **exonucleases 5'→3' e 3'→5'**. A DNA pol I está envolvida no reparo de quebras ou injúrias do DNA e possui papel subsidiário na duplicação. A DNA pol II é uma enzima de reparo do DNA.

Uma característica comum a todas as DNA pol é o fato de todas serem incapazes de iniciar a síntese da cadeia de DNA. Precisa haver um pequeno segmento já pronto, com a extremidade 3'-OH livre, denominado de **primer**. Na realidade, estas enzimas são capazes de alongar a cadeia, acrescentando novos nucleotídeos. Normalmente, o *primer* é constituído de um pequeno segmento de RNA, sintetizado pela **RNA polimerase**, ou **RNA primase**, que tem a capacidade de iniciar a síntese.

As DNA polimerases atuam sobre o trifosfato dos nucleotídeos. A base complementar à cadeia molde emparelha-se e a DNA pol age sobre o trifosfato, liberando um difosfato e adicionando o novo nucleotídeo à extremidade 3' - OH da cadeia, formando uma **ligação diestérica** (Figura 4.14).



**Figura 4.14:** ação da DNA polimerase na união de dois nucleotídeos, formando uma ligação diestérica. (Adaptado de Lewin, B. **Gene V**. Oxford Press, 1994).

A estrutura **antiparalela** da molécula de DNA constitui uma dificuldade para a duplicação, já que as polimerases conhecidas só conseguem adicionar nucleotídeos na extremidade 3' da cadeia, determinando um crescimento no sentido 5'→3'. Na forquilha de duplicação, uma das cadeias moldes está no sentido 3'→5', onde a polimerase poderá atuar sem dificuldades, mas a outra cadeia molde está no sentido 5'→3' e a polimerase não sintetizará uma cadeia complementar no sentido oposto (Figura 4.15).



**Figura 4.15:** Enquanto a cadeia leading é sintetizada continuamente, a cadeia lagging é duplicada em segmentos, caracterizando a duplicação semidescontínua do DNA. (Adaptado de Lewin, B. **Gene V**. Oxford Press, 1994).

Okasaki e colaboradores, em 1968, descobriram que a síntese de DNA não é contínua. No sentido  $5' \rightarrow 3'$ , a DNA pol III catalisa normalmente a polimerização dos nucleotídeos, mas na outra cadeia molde, a fita complementar é sintetizada em pequenos segmentos, denominados **fragmentos de Okasaki**, de 1000 a 2000 nucleotídeos, também no sentido  $5' \rightarrow 3'$ . Esses fragmentos são sintetizados na direção oposta ao da abertura da forquilha de duplicação. Sua síntese termina, quando encontram a extremidade do fragmento que o precedeu na síntese.

A síntese de cada fragmento de Okasaki representa uma nova iniciação na duplicação do DNA. Como nenhuma DNA pol é capaz de iniciar a síntese, cada fragmento possui em sua extremidade inicial um *primer* de aproximadamente 10 ribonucleotídeos, sintetizado pela RNA primase. O *primer* funciona como iniciador para a DNA pol III. O *primer* de RNA da extremidade 5' do segmento de Okasaki é posteriormente substituído por desoxirribonucleotídeos pela ação da DNA pol I, que através de sua atividade reparadora **exonucleotídica  $5' \rightarrow 3'$** , remove os ribonucleotídeos.

Os fragmentos já reparados são unidos entre si pela **DNA ligase**, que conecta a extremidade 3'-OH de um fragmento com a extremidade 5'-P do fragmento adjacente. Como a cadeia *leading* ( $5' \rightarrow 3'$ ) é duplicada continuamente, enquanto a cadeia *lagging* ( $3' \rightarrow 5'$ ) é sintetizada em fragmentos, uma das características da duplicação do DNA é ser **semi-descontínua**. (Figura 4.15).

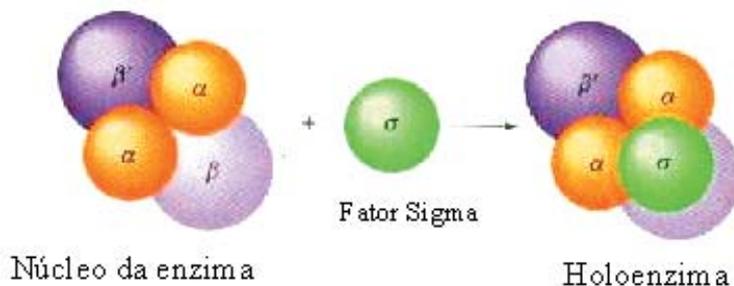
## SÍNTESE DE RNA – TRANSCRIÇÃO

O RNA é um polímero de ribonucleotídeos de cadeia simples e linear. O tamanho das cadeias é bastante variável, de 75 até 10.000 nucleotídeos.

A síntese do RNA tem como molde uma das cadeias do DNA. A seqüência de ribonucleotídeos é determinada pela complementaridade à seqüência de bases da cadeia molde de DNA. Este processo é denominado de **Transcrição**.

Embora, para cada gene, apenas uma das cadeias da dupla hélice do DNA sirva de molde para a transcrição, ao longo de todo o cromossomo não ocorre transcrição da mesma cadeia de DNA. Diferentes genes podem ser ativados em diferentes segmentos do cromossomo em tempos distintos, variando a cadeia de DNA a ser transcrita. Enquanto o RNA é complementar à cadeia molde do DNA, sua seqüência de ribonucleotídeos é a mesma da cadeia de DNA que não foi transcrita, exceto pelas uracilas que substituem as timinas no RNA.

A transcrição só ocorre na presença da enzima **RNA polimerase**, que é formada por 4 subunidades diferentes:  $\alpha$  (alfa),  $\beta$  (beta),  $\beta'$  (beta linha) e  $\sigma$  (sigma), sendo que cada conjunto enzimático possui um representante de cada subunidade e duas subunidades  $\alpha$ . (Figura 4.16).

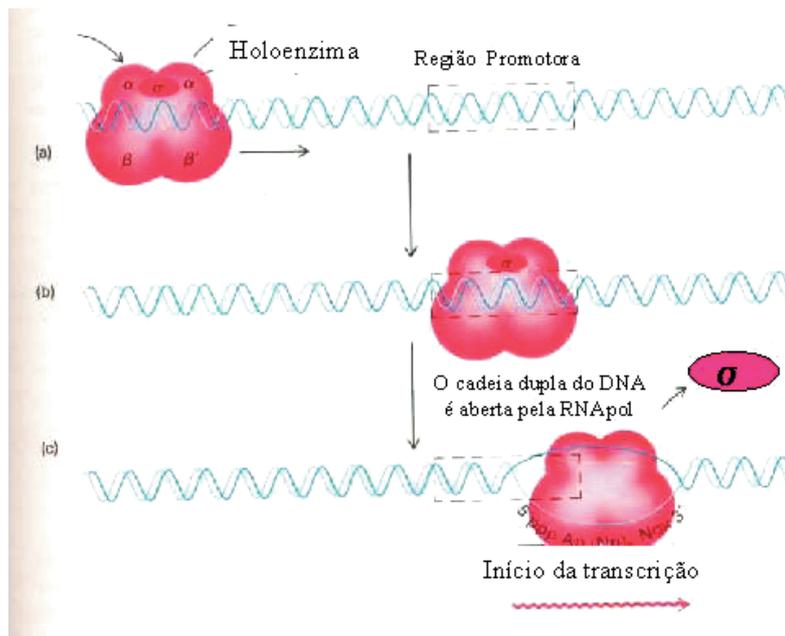


**Figura 4.16:** Sub-unidades constituintes da RNA polimerase. (Adaptado de Griffiths, A. J. **An Introduction of Genetic Analysis**. W. H. Freeman & Cia., 1993).

Podem ser distinguidas três fases da transcrição: **Iniciação**, **Elongação** e **Término** ou **Parada**.

- **Iniciação:**

A RNA pol reconhece e liga-se a uma região específica do DNA, denominada de **região promotora** ou **promotor**, através da subunidade  $\sigma$ . As regiões promotoras dos vários genes possuem homologias, ou seja, apresentam similaridade na seqüência de bases. Estes segmentos estão ao lado da primeira base a ser transcrita e são denominados de **sítios de iniciação**. (Figura 4.17).



**Figura 4.17:** Inicialização da transcrição. Após a localização da Região Promotora, há a liberação do fator  $\sigma$ . (Adaptado de Griffiths, A. J. *An Introduction of Genetic Analysis*. W. H. Freeman & Cia., 1993).

Após a ligação com a região promotora, a RNA pol causa uma desnaturação local no DNA, abrindo a dupla hélice. A presença do fator  $\sigma$  é primordial para a ocorrência destes eventos.

- **Elongação:**

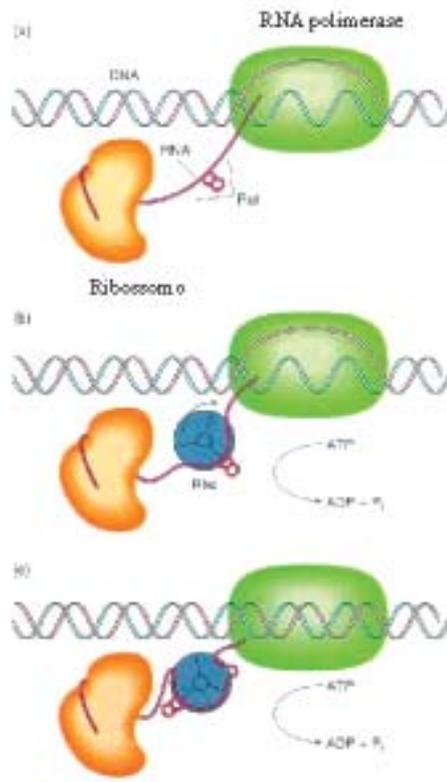
Com a abertura da dupla hélice do DNA, o fator sigma ( $\sigma$ ) dissocia-se da RNA pol. O RNA é sintetizado no **sentido 5' → 3'**, com a colocação de ribonucleotídeos com bases complementares às da cadeia molde do DNA. A energia da reação deriva da quebra do trifosfato dos ribonucleotídeos.

- **Término** ou **Parada:**

A elongação termina, quando a RNA pol reconhece o "**sinal de parada**". Há dois mecanismos envolvidos no **término**. No primeiro, uma seqüência no DNA molde, de aproximadamente 40 pb (pares de bases), com alta freqüência de G-C, seguidos de 6 ou mais adeninas é o **sinal de parada**. As seqüências repetidas de C-G no RNA formam uma cadeia dupla, resultando num *looping*, seguido de uma seqüência de uracilas, complementares à seqüência de A do DNA molde. O *looping* impede a continuação da ação da RNA polimerase e a transcrição termina.

No segundo mecanismo, há a necessidade da presença de uma proteína específica, o **fator  $\rho$**  (rho), que se liga a uma região específica do RNA recém sintetizado, deslocando-o do sítio de síntese da RNA pol, impedindo a continuação

da transcrição. Uma vez terminada a síntese, a molécula de RNA separa-se completamente do DNA, já que a cadeia híbrida DNA-RNA não é estável. (Figura 4.18).



**Figura 4.18:** Papel do Fator Rho no término da transcrição. (Adaptado de Griffiths, A. J. **An Introduction of Genetic Analysis**. W. H. Freeman & Cia., 1993).

## TIPOS DE RNA

Existem três tipos fundamentais de RNA: o transportador (**RNA<sup>t</sup>**), o ribossômico (**RNA<sup>r</sup>**) e o mensageiro (**RNA<sup>m</sup>**), cada um com características e funções específicas.

### • RNA transportador:

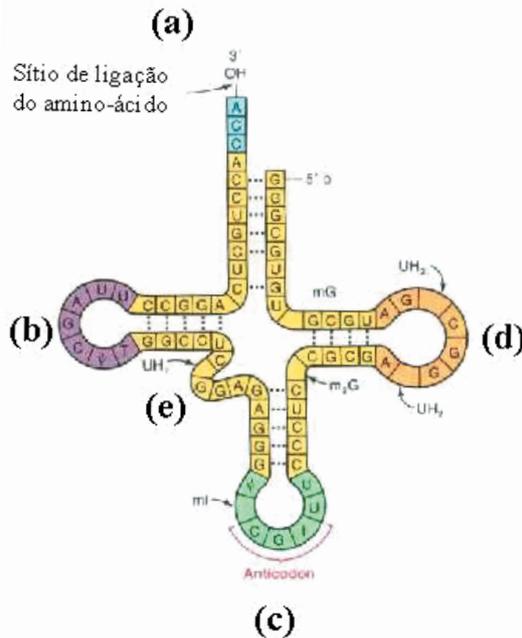
O RNA<sup>t</sup> atua na síntese protéica, levando os aminoácidos até o ribossomo, local da síntese protéica.

A estrutura terciária do RNA<sup>t</sup> confere a sua capacidade funcional. Seu tamanho está em torno de 74 a 95 bases, que em algumas regiões ligam-se por pontes de Hidrogênio, formando cadeia dupla.

Encontram-se ao longo da molécula, nucleotídeos com bases não usuais, modificadas como a **pseudouridina** ( $\psi$ ), a **dihidrouridina** (DHU) e outras. Nestas regiões, não há pareamento das bases, a cadeia permanece simples, formando alças que se continuam com as porções de cadeia dupla, permitindo que grupos livres (amino ou cetona) possam formar ligações secundárias. (Figura 4.19).

A seqüência de bases do RNA<sup>t</sup> dispõe-se, formando uma imagem que lembra uma folha de trevo, com 4 extremidades:

- a.** Extremidade receptora, formada por uma seqüência de bases pareadas, terminando com uma seqüência ACC, onde há a ligação com o amino-ácido.
- b.** Alça T $\psi$ C, que contém pseudouridina, formada por 7 bases não emparelhadas



**Figura 4.19:** Estrutura do RNA transportador. A. Extremidade ligada ao amino-ácido. B. Alça TψC. C. Alça do anti-códon. D. Alça DHU. E. Alça extra. (Adaptado de Griffiths, A. J. **An Introduction of Genetic Analysis**. W. H. Freeman & Cia., 1993).

e que se liga ao ribossomo na síntese protéica.

**c.** Alça anticódon, que se liga ao códon do RNAm durante a síntese protéica, formada por 7 bases não emparelhadas, sendo três destas bases as formadoras da trinca do anti-códon.

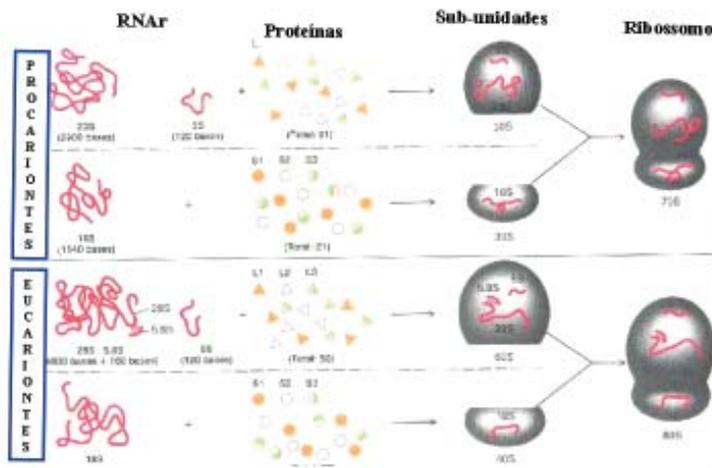
**d.** Alça extra, de ocorrência e tamanho variáveis, entre as alças TψC e do anticódon, com função desconhecida.

Dentre a população de RNAt, determinadas seqüências de bases são constantes, comuns a todos, dando a mesma conformação espacial a todas as moléculas.

Estudos com difração de raio X elucidaram a estrutura terciária do RNAt. Estes estudos mostraram que pontes de hidrogênio adicionais são responsáveis por dobras nas alças do trevo, conferindo à molécula uma estrutura terciária estável com forma aproximada de um **L**.

• **RNA ribossômico:**

O RNA ribossômico, juntamente com as **proteínas ribossomais**,



**Figura 4.20:** Tipos de RNA ribossômicos e seu papel na formação dos ribossomos de procariontes e de eucariontes. (Adaptado de Griffiths, A. J. **An Introduction of Genetic Analysis**. W. H. Freeman & Cia., 1993).

constituem uma organela não membranosa, o **ribossomo**, local onde ocorre a síntese protéica. (Figura 4.20).

A transcrição do DNA que codifica o RNAr, a **região organizadora de nucléolo** (RON) ou **organizador nucleolar** (ON), é composta de grupos de vários genes iguais repetidos e origina, nos procariontes, uma molécula de taxa de sedimentação de 30S (medida em Svedbergs, em um gradiente de densidade de sacarose, onde um alto valor de **S** indica uma molécula com alta densidade e massa). Após sua síntese, a molécula é clivada em três RNA ribossômicos menores, de índice de sedimentação **23S**, **16S** e **5S**.

Nos eucariontes, a molécula de RNAr, é maior que a dos procariontes, 45S e é clivada em segmentos **28S** e **18S**. Neste caso, o RNAr 5S não faz parte da molécula precursora 45S, derivando de outro gene a ser transcrito.

#### • RNA mensageiro:

Suspeitava-se da existência do RNAm, pois nos eucariontes há a necessidade de um intermediário que leve a informação genética do DNA nuclear para o citoplasma, onde as proteínas são sintetizadas.

O RNAm representa uma seqüência de nucleotídeos feita a partir do DNA e que corresponderá à seqüência de aminoácidos de uma proteína. Portanto, as moléculas de RNAm são bastante **heterogêneas** quanto ao seu tamanho e à seqüência de bases nitrogenadas que os compõem.

Existem diferenças marcantes entre os RNAm de procariontes e de eucariontes.

#### **Procariontes:**

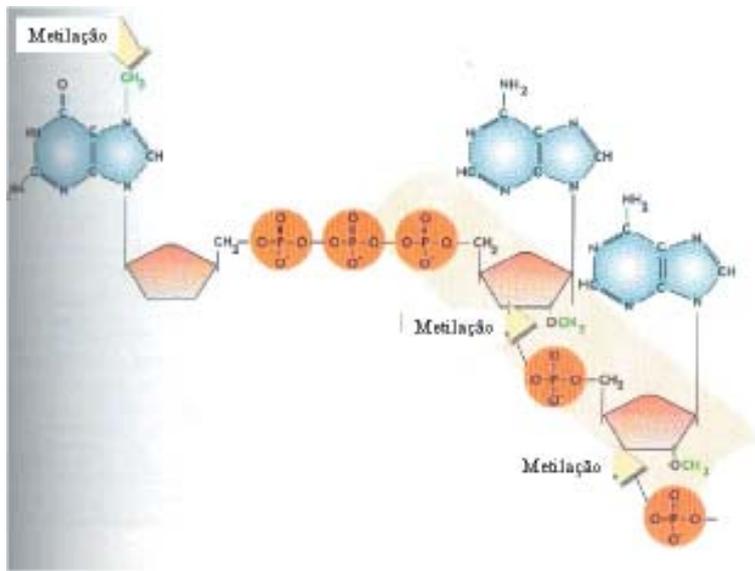
1. Como não há separação física entre o núcleo e o citoplasma, o RNAm é transcrito e traduzido no mesmo compartimento e de maneira simultânea.
2. O RNAm é extremamente **instável**, com vida média de 1 a 2 minutos. Enquanto a extremidade 3' ainda está sendo transcrita, a extremidade 5' já é degradada.
3. O RNAm é, na maioria das vezes, **policistrônico**, ou seja, representa vários genes vizinhos e ao ser traduzido, gera várias proteínas diferentes, cada uma correspondente a um dos genes que foi transcrito.
4. A extremidade 5' do RNAm contém uma trinca de iniciação, AUG (ou GUG, ou UUG), que acrescida de mais quatro bases, compõem uma seqüência de 7 bases conhecida como **seqüência de Shine-Dalgarno**. Por esta extremidade que o RNAm liga-se ao RNAr 16S da sub-unidade menor do ribossomo, para dar início à tradução.

#### **Eucariontes:**

1. A transcrição e a maturação do RNAm ocorrem exclusivamente no núcleo, enquanto que a tradução é restrita ao citoplasma.
2. RNAm é **monocistrônico**. Representa apenas um gene e, transcrito, origina um único tipo protéico.
3. RNAm é mais **estável**, durando por várias horas, pois há síntese de maior quantidade de proteína e as proteínas são mais complexas, aumentando o tempo da tradução.
4. À extremidade 3' do RNAm já processado, é acrescida uma seqüência de aproximadamente 200 ribonucleotídeos de adenina, denominado de **poli-A**.

A presença da poli-A está relacionada com a estabilidade da molécula RNAm, pois a degradação da molécula só ocorre após a sua remoção.

5. A extremidade 5' do RNA está ocupada por uma purina, A ou G. Ao invés da extremidade desta base apresentar como radical um trifosfato, há a ligação com uma **guanina trifosfatada**, o que ocorre após a transcrição. Enquanto todos os nucleotídeos do RNAm estão unidos por ligações 5'→3' fosfo-diéster, a Gppp liga-se ao último nucleotídeo do RNAm pela extremidade 5', com uma orientação reversa do restante da molécula (5'→5'). Esta conformação é chamada de **cap** ou **capuz** e freqüentemente é metilada. O grupamento inicial **5'→cap** é imprescindível no reconhecimento e ligação com a sub-unidade menor do ribossomo. (Figura 4.21).



**Figura 4.21:** Nos eucariontes o grupamento inicial 5' -cap facilita o reconhecimento e ligação do RNAm com a sub-unidade menor do ribossomo. Outros locais de metilação também podem existir. (Adaptado de Lewis B. **Genes V**. Oxford Press, 1994).

#### • Processamento do RNA em Eucariontes:

Apesar do RNAm possuir uma seqüência de nucleotídeos que corresponde exatamente à seqüência de aminoácidos da proteína que traduz, o segmento de DNA que originou o RNAm, possui seqüências de nucleotídeos que não estão representadas na proteína. Esta discrepância é devida a função diferente de duas porções do gene: o **exon** e o **intron**.

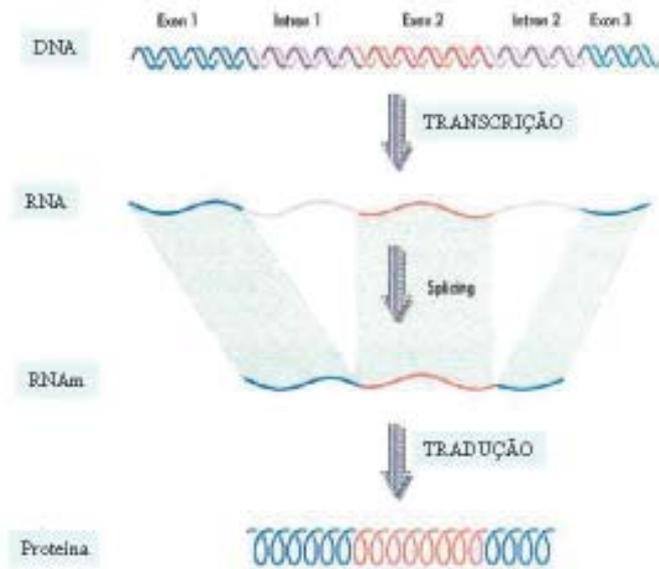
Exon: regiões do DNA que são transcritas e traduzidas.

Intron: regiões do DNA que são transcritas, mas são retiradas do RNAm antes da tradução, por um processo denominado **splicing**, que ocorre ainda no núcleo. (Figura 4.22).

Após o *splicing*, os exons são unidos na mesma ordem ditada pela seqüência no DNA. O RNAm que ainda contém os introns é chamado de **RNAm precursor** ou **pré-RNA**.

O *splicing* é um processo mediado por enzimas, as **endonucleases** que reconhecem os introns e os cortam, retirando-os da seqüência a ser traduzida. É um processo extremamente preciso, pois cortes não exatos gerarão RNAm diversos e proteínas alteradas.

Enquanto há **colinearidade** entre as seqüências de bases do DNA e a seqüência de aminoácidos da proteína nos procariontes, esta característica é interrompida nos eucariontes pela presença dos introns, que não são traduzidos.



**Figura 4.22:** Representação da retirada de segmentos do RNAm de eucariontes, *splicing*, correspondentes aos introns do DNA. (Adaptado de Lewis B. **Genes V**. Oxford Press, 1994).

## CÓDIGO GENÉTICO

O **gene** é uma unidade funcional que codifica um polipeptídeo. A unidade funcional corresponde a um segmento do DNA nuclear. A seqüência de aminoácidos da proteína corresponde à seqüência de bases nitrogenadas do DNA. Alterações das bases do DNA determinam alterações na seqüência de aminoácidos da proteína.

A informação contida na molécula de DNA, em forma de seqüência de bases nitrogenadas, é transcrita para a seqüência de bases nitrogenadas do RNAm, que é traduzida para uma linguagem de aminoácidos durante a síntese protéica.

Em 1954, George Gamov, um físico teórico, propôs que cada amino-ácido seria codificado por um conjunto de três nucleotídeos. Como são conhecidos 20 aminoácidos diferentes e apenas 4 tipos de nucleotídeos, o número mínimo de combinações de nucleotídeos para dar um número igual ou maior que 20, seriam combinações de três nucleotídeos. Esta hipótese foi confirmada por Crick e colaboradores, em 1961, trabalhando com um vírus, o fago  $T_4$ .

Durante a síntese protéica, a leitura do RNAm tem início de uma extremidade da molécula, deslocando-se em blocos consecutivos de três bases, em direção ao extremo oposto.

Em 1961, Nurenberg & Mattei sintetizaram um RNA artificial **poli-U** e *in vitro* sintetizaram uma proteína constituída exclusivamente pelo amino-ácido fenilalanina. Concluíram que cada amino-ácido é codificado por uma trinca de nucleotídeos que foi chamada de **códon**. Em 1964, decifraram todo o **Código Genético**, com o quadro completo de correspondência entre os 20 aminoácidos e os códons. Portanto, durante a tradução, cada trinca de bases do RNAm, o códon, corresponde a um amino-ácido específico na proteína que está sendo sintetizada. A leitura dos códons é seqüencial e sem sobreposição. Um RNA com a seqüência AUCAUUCAG tem três códons a serem traduzidos que são AUC, AUU e CAG.

Qualquer uma das quatro bases nitrogenadas pode ocupar cada uma das três posições do códon. Assim, são  $4^3 = 64$  combinações possíveis. Como são 20 aminoácidos que constituem as proteínas, há mais códons do que aminoácidos, determinando que quase todos os aminoácidos sejam representados por mais de um códon. Por esse motivo, diz-se que o Código Genético é **degenerado**.

Códons que representam o mesmo amino-ácido são denominados **sinônimos**. Somente triptofano e metionina são aminoácidos relacionados com apenas um códon cada. Há uma tendência de que quanto mais comum o amino-ácido, maior a quantidade de códons que o representam. Além disso, códons sinônimos são similares em suas seqüências de bases, minimizando os efeitos das mutações.

Três códons, UAA, UAG e UGA não representam qualquer amino-ácido. Eles participam do término da síntese protéica como **códons de parada**. (Figura 4.23).

	Segunda base					
	U	C	A	G		
U	UUU } Phe UUC } UUA } Leu UUG }	UCU } Ser UCC } UCA } UCG }	UAU } Tyr UAC } UAA } TERM UAG }	UGU } Cys UGC } UGA } TERM UGG } Trp		
	C	CUU } Leu CUC } CUA } CUG }	CCU } Pro CCC } CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } Arg CGC } CGA } CGG }	
		A	AUU } Ile AUC } AUA } Met AUG }	AUU } Thr AUC } AUA } AUG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }
			G	GUU } Val GUC } GUA } GUG }	GCU } Ala GCC } GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }

**Figura 4.23:** Trinca de bases nitrogenadas dos códons e a correspondência com os respectivos amino-ácidos que determina. As trinca seguidas de TERM, não codificam amino-ácido algum, indicando a parada da síntese protéica. (adaptado de Lewis, B. **Genes V**. Oxford Press, 1994).

Comparações entre a seqüência de bases do DNA e a proteína que gera demonstraram que a correspondência dos códons com os aminoácidos é a mesma tanto nas bactérias quanto nos eucariontes. Um RNAm de uma espécie pode ser traduzido por um aparato enzimático de outra espécie, gerando uma proteína idêntica à original. O Código Genético é **universal**, sugerindo seu estabelecimento precoce na história evolutiva dos seres vivos.

## TRADUÇÃO OU SÍNTESE PROTÉICA

A tradução ocorre com a leitura da série de códons do RNAm, no **sentido 5' → 3'**, produzindo uma proteína.

A polimerização dos aminoácidos em proteínas ocorre no ribossomo. Todos os ribossomos podem ser dissociados em duas subunidades, uma maior e uma menor, cada uma constituída por um RNAr e várias proteínas diferentes, as **proteínas r**, que coordenam as atividades da tradução e estabelecem a estrutura dos sítios ativos do ribossomo.



### • Iniciação:

Envolve a reação do RNAm ao seu sítio de ligação na subunidade menor do ribossomo, formando o **complexo de iniciação**, ao qual a subunidade maior vem acoplar-se.

A formação do complexo de iniciação é acionada por proteínas denominadas **Fatores de iniciação (IF)**, que permanecem ligadas ao ribossomo somente durante a fase de iniciação.

Nos procariontes, a porção do RNAm que liga-se ao ribossomo é a seqüência de Shine-Dalgarno. Nos eucariontes, existem vários tipos de fatores de iniciação e, alguns deles reconhecem o capuz (cap) da extremidade 5' do RNAm, ligando-a à subunidade 40S do ribossomo.

O primeiro códon a ser lido é o referente à metionina, AUG.

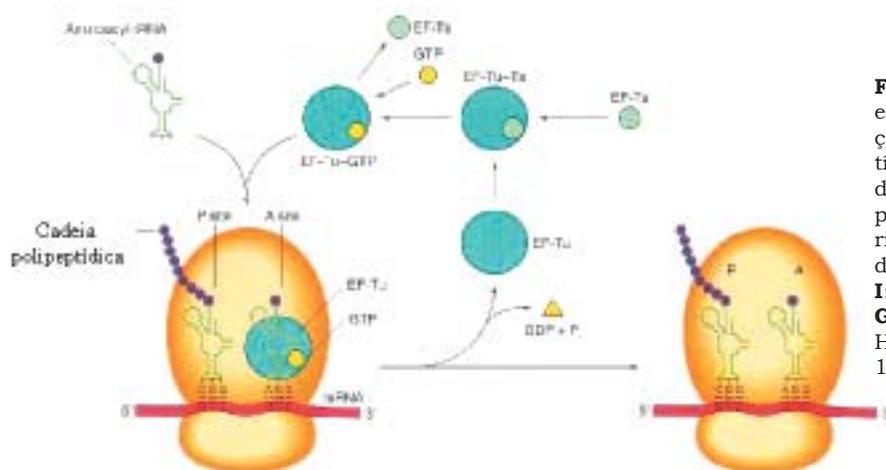
### • Elongação:

Nesta fase, ocorrem as reações de síntese do polipeptídeo. Os aminoácidos são acoplados através de **ligações peptídicas**.

A subunidade maior do ribossomo possui dois sítios de ligação: **A** e **P**, assim denominados, pois o **A** abriga o **aminoacil-RNAt** e o **P** o **polipeptídeo-RNAt**. O sítio A tem afinidade pelo aminoacil-RNAt, cujo anticódon é complementar ao códon que está exposto no sítio A.

O **fator EF-tu**, nos procariontes e o **EF-1** nos eucariontes, são proteínas ligadas a um GTP com a função de carregar o aminoacil-RNAt ao sítio A do ribossomo. Após a colocação do aminoacil-RNAt, o GTP é clivado, o fator EF-tu desliga-se e a energia liberada faz com que o ribossomo exponha o sítio P, para onde o aminoacil-RNAt vai ser transferido. Com o sítio A livre, outro aminoacil-RNAt é carregado pelo fator EF. Há, então, a ligação do amino-ácido do sítio P com o do sítio A mediada pela enzima **peptidil transferase** e seu desprendimento do RNAt, que perde afinidade pelo sítio P e sai do ribossomo. O RNAt, ligado ao dipeptídeo, passa a ocupar o sítio P e é denominado **polipeptidil-RNAt**.

O processo de elongação continua com a entrada de um novo aminoacil-RNAt no sítio A, complementar ao códon seguinte. Uma nova ligação peptídica entre os aminoácidos adjacentes é processada, há liberação do sítio P, migração do polipeptidil-RNAt do sítio A para o P e entrada de novo aminoacil-RNAt. (Figura 4.25).



**Figura 4.25:** Fatores envolvidos na elongação da cadeia polipeptídica e a transferência do RNAt do sítio A para o sítio P do ribossomo. (Adaptado de Griffiths, A. J. **An Introduction of Genetic Analysis**. W. H. Freeman & Cia., 1993).

• **Término** ou **Parada**:

A tradução prossegue até que um códon de término ou de parada seja exposto no sítio A do ribossomo. Os **códons UAG, UAA e UGA** não são complementares a nenhum anticódon, portanto não representam qualquer dos aminoácidos.

As reações de término envolvem a liberação do polipeptídeo do último RNAt, retirada deste RNAt do sítio P e a dissociação do ribossomo do RNAm.

Vários ribossomos são percorridos por uma molécula de RNAm simultaneamente. Isto determina que diversas cadeias polipeptídicas, em diferentes fases de crescimento, possam ser encontradas ao longo de uma mesma molécula mensageira. Ao conjunto constituído pela molécula de RNAm, pelos ribossomos a ela acoplados e pelas cadeias polipeptídicas em formação, que estão sendo traduzidas pelo conjunto, denomina-se **polissomo** ou **polirribossomo**. Em uma atividade de síntese intensa pode ser encontrado o máximo de um ribossomo para cada 80 nucleotídeos do RNAm. Isto permite que muitas proteínas formem-se a partir de uma única cadeia de RNAm e explica a baixa porcentagem que a fração de RNAm representa no total de RNA de uma célula.

## TECNOLOGIA DO DNA RECOMBINANTE

O uso da tecnologia do DNA recombinante em pesquisas na área de saúde e, recentemente, na prática médica, constitui um revolucionário recurso tanto para o estudo de doenças, genéticas ou não, quanto para a pesquisa do genoma humano.

No início da década de 70, os trabalhos de Davis & Mertz sobre enzimas de restrição e a introdução da técnica de transferência de Southern (*Southern blot*) foram os passos fundamentais para o rápido desenvolvimento da biologia molecular e os seus aspectos aplicados, tanto para a medicina, quanto para a genética populacional, a agricultura e a veterinária.

Atualmente, pode-se isolar fragmentos distintos de DNA de qualquer organismo e recombiná-los em pequenos genomas, que são mais fáceis para serem analisados e manipulados. O DNA recombinante pode ser feito a partir de segmentos de DNA não homólogo e de origem diversa ou de diferentes espécies. Pode ser amplificado, ou seja, duplicado por vários ciclos, produzindo um grande número de cópias, facilitando a análise de detalhes da molécula.

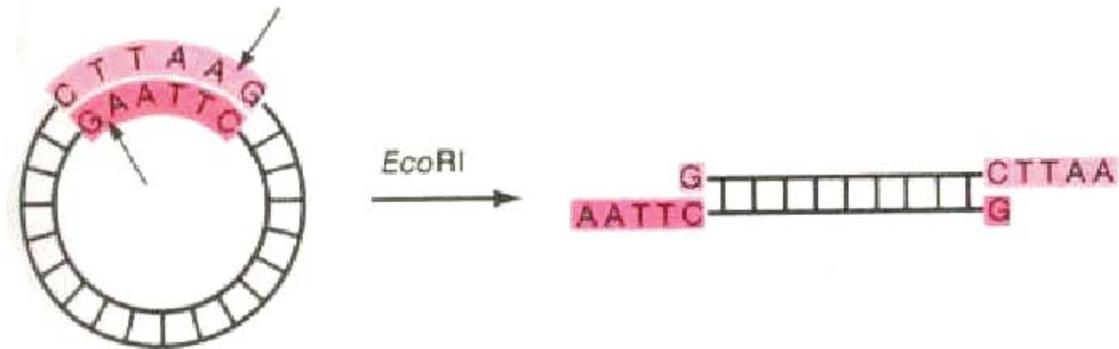
• **Enzimas de Restrição**:

As enzimas de restrição são **endonucleases bacterianas** que cortam o DNA em fragmentos e representam o principal instrumento da tecnologia do DNA recombinante. Reconhecem uma seqüência específica de nucleotídeos, como CTGCA e cortam o DNA onde esta seqüência ocorre, produzindo uma população de fragmentos heterogêneos com extremidades idênticas.

As enzimas de restrição são proteínas de ocorrência natural nas bactérias, da onde foram isoladas e identificadas. Constituem um **sistema imune bacteriano primitivo**, pois reconhecem um DNA exógeno, de bacteriófagos ou de outros vírus que invadem as bactérias. Cortam o DNA viral, tornando-os inofensivos.

O DNA bacteriano está protegido contra a ação de suas próprias endonucleases de restrição, através de modificações químicas, como a metilação, que se faz nas bases nitrogenadas.

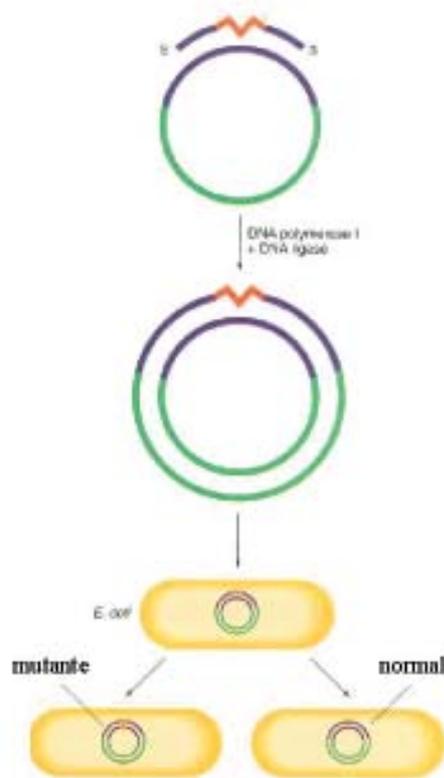
São conhecidas dezenas de enzimas de restrição diferentes, que constituem uma potente ferramenta de análise, pois localizam uma seqüência específica no DNA, cortando-o. (Figura 4.26).



**Figura 4.26:** Ação da enzima de restrição EcoRI sobre um DNA circular bacteriano, resultando em um DNA linear. (Adaptado de Griffiths, A. J. **An Introduction of Genetic Analysis**. W. H. freeman & Cia., 1993).

#### • Metodologia do DNA recombinante:

Um dos objetivos da obtenção do DNA recombinante é isolar e recombinar um segmento específico do DNA com um genoma pequeno, como o de um **vírus** ou **plasmídeo**, que ao duplicar-se, gera numerosas cópias, amplificando o segmento de DNA de interesse. Este processo é denominado de **clonagem**. (Figura 4.27).



**Figura 4.27:** Introdução de um segmento de DNA exógeno (laranja), no genoma de uma bactéria, dando origem a uma linhagem mutante. (Adaptado de Griffiths, A. J. **An Introduction of Genetic Analysis**. W. H. freeman & Cia., 1993).

Os veículos para a clonagem são fragmentos de DNA, denominados **vetores**, que contêm uma origem para replicação e têm a capacidade de duplicar-se juntamente ao fragmento de DNA inserido. Podem ser utilizados como vetores os **plasmídeos**, o **fago  $\lambda$**  e os **cosmídeos**.

**Plasmídeos:** Moléculas pequenas e circulares de DNA, que podem ser introduzidas em bactérias ou outras células. São considerados minicromossomos, pois autoduplicam-se. Permitem a amplificação do DNA, pois podem existir até 1000 plasmídeos por ciclo celular.

**Fago  $\lambda$ :** Um vírus que reproduz-se introduzindo seu DNA em bactérias. É utilizado para clonar fragmentos grandes de DNA, de 15 a 20 kilobase (kb=1000 bases).

**Cosmídeo:** São híbridos de fagos e plasmídeos, que foram desenvolvidos para a clonagem de fragmentos maiores de DNA, com cerca de 45kb.

#### • **Detecção de genes clonados:**

Uma das técnicas mais valiosas para a identificação de genes clonados foi desenvolvida por E. M. Southern e, por isso, foi denominada de **transferência de Southern** ou **Southern blot**.

Nesta técnica, fragmentos de DNA resultantes da ação de enzimas de restrição são separados por diferença de tamanho em um gel de agarose, submetido a **eletroforese**. Eles são transferidos do gel para um **filtro de nitrocelulose** carregado eletrostaticamente.

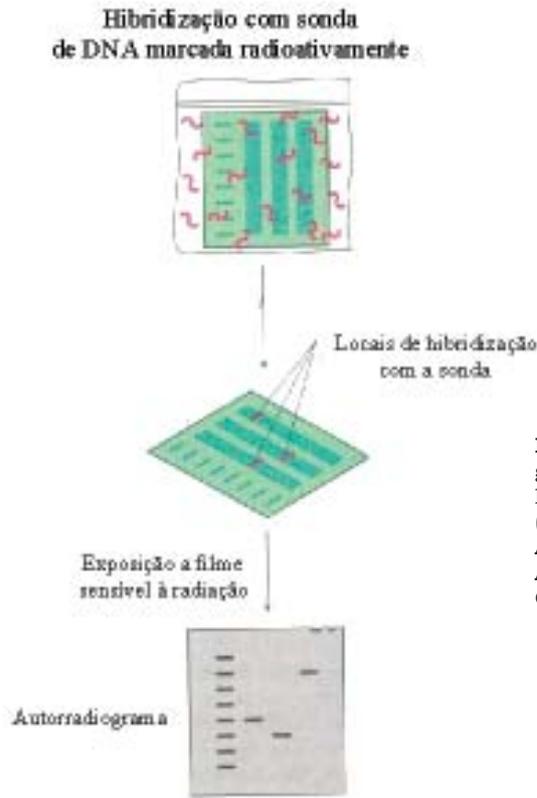
A transferência de Southern pode ser estendida para fragmentos de RNA e então é denominada **transferência de Northern** ou **Northern blot**. Utilizando-se os mesmos passos descritos da transferência de Southern, na transferência de Northern a análise torna-se mais difícil, pois o RNAm é uma molécula mais instável do que o DNA, mas apresenta a vantagem de revelar os genes que estão sendo transcritos, ou seja que estão em funcionamento na célula no momento da coleta dos RNA.

A enzima **transcriptase reversa** tem a capacidade de sintetizar um DNA a partir de um RNAm. Este DNA é denominado **DNA complementar** ou **cdNA**. Este processo facilita a obtenção dos segmentos de DNA que estão transcrevendo naquele momento na célula, sem apresentar o problema da instabilidade do RNAm. A população de cdNA pode ser unida a plasmídeos, cada plasmídeo com um tipo de cdNA e clones de cada população podem ser obtidos.

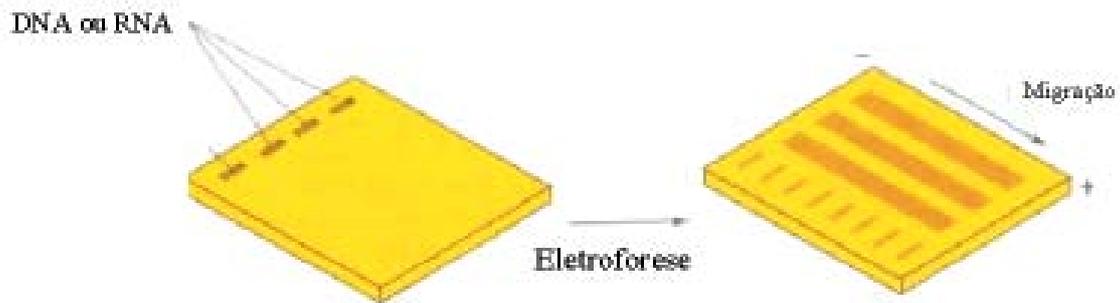
Tanto os fragmentos de DNA quanto os de RNA são revelados por hibridização com **sondas marcadas radioativamente**. Estas sondas, complementares ao DNA de interesse, juntam-se a ele, por ação da **complementaridade** de suas bases. O filtro é exposto a um filme de Raio X e no **autorradiograma** as bandas híbridas, DNA mais sonda radioativa, são reveladas (Figura 4.28).

Existem métodos frios de marcação de DNA, que utilizam nucleotídeos marcados quimicamente, com **biotina**, **fosfatase alcalina** ou **substâncias quimioluminescentes**, que apresentam a vantagem de serem mais estáveis do que as sondas radioativas.

Na **transferência de Western** ou **Western blot** há a transferência de proteínas de um gel de agarose, após eletroforese, para a membrana de nitrocelulose, onde podem ser reveladas por ação de **anticorpos específicos**. (Figura 4.29 e 4.30)



**Figura 4.28:** Autorradiograma revelando bandas híbridas com a sonda de DNA. (Adaptado de Griffiths, A. J. **An Introduction of Genetic Analysis**. W. H. freeman & Cia., 1993).

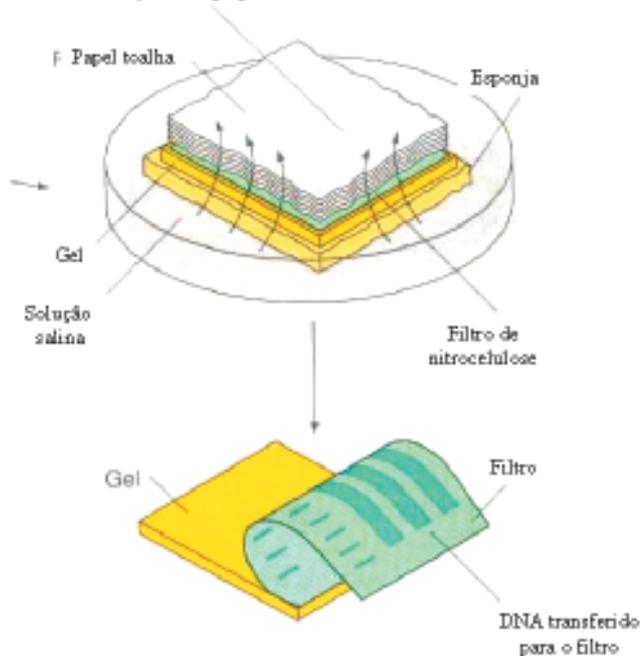


**Figura 4.29:** Eletroforese de fragmentos de DNA ou de RNA para realização da transferência de Southern e de Northern, respectivamente. (Adaptado de Griffiths, A. J. **An Introduction of Genetic Analysis**. W. H. freeman & Cia., 1993).

• **A Reação em Cadeia da Polimerase – PCR:**

O método da reação em cadeia da polimerase tem como principal objetivo a amplificação de segmentos de DNA de interesse para pesquisa, ou para saúde ou outros. Foi estabelecido por Kary Mullis, em 1983, que ganharia o Prêmio Nobel em 1994.

A solução salina passa para o filtro,  
em direção ao papel toalha



**Figura 4.30:** Transferência de fragmentos de DNA ou RNA para o filtro de nitrocelulose. (Adaptado de Griffiths, A. J. **An Introduction of Genetic Analysis**. W. H. Freeman & Cia., 1993).

O PCR apresenta três etapas distintas:

- A)** Síntese dos *primers*: síntese de dois segmentos de 20 pb, complementares às extremidades do fragmento de DNA de interesse.
- B)** Reunião dos *primers* com nucleotídeos livres, o fragmento de DNA a ser amplificado e uma DNA polimerase especial, a **Taq polimerase**.
- C)** Três ciclos de temperatura:
  1. 95°C: desnaturação da cadeia dupla do DNA a ser amplificado.
  2. 55°C: com o resfriamento há a ligação dos *primers* com sítios complementares do DNA de cadeia simples.
  3. 75°C: a Taq polimerase promove a síntese de uma nova cadeia do DNA, a partir dos *primers*, no sentido 5'→3'. Cópias novas do DNA inicial foram criadas a partir da região de ligação com os *primers*.

A etapa C é repetida várias vezes e, com a utilização de aparelhos automatizados, como o termociclador, que após 30 ciclos obtém 1 milhão de cópias do DNA inicial.

A Taq polimerase é uma polimerase especial, pois atua sintetizando o DNA a altas temperaturas (75°C) e não desnatura em temperaturas mais altas, como 95°C da etapa de desnaturação. Esta enzima é de ocorrência natural e foi isolada de um microrganismo aquático, o *Thermus aquaticus*, que vive nas fontes termais do Parque Nacional de Yellowstone.

#### • Aplicações da tecnologia do DNA recombinante:

##### **A)** Clonagem gênica:

Um segmento de DNA pode ser isolado e clonado. Este segmento é obtido pela ação de enzimas de restrição e é inserido em um vetor, como um plasmídeo.

O plasmídeo portador do segmento de DNA multiplica-se dentro de bactérias. O DNA multiplicado é removido dos plasmídeos e pode ser utilizado para pesquisa ou na detecção do produto deste gene.

**B) Animais transgênicos:**

Resultado da inserção de um gene exógeno na fase embrionária do animal. O gene exógeno passa a fazer parte do material genético do animal hospedeiro. Se o animal for fértil o **transgene** pode ser herdado pela futura progênie.

A técnica envolve a escolha do gene exógeno, a colocação do gene exógeno na forma de um **construto**, que orientará a inserção do gene exógeno no genoma do animal e estimulará sua expressão e a injeção do construto em um zigoto, ou em um embrião nas primeiras fases de desenvolvimento do animal hospedeiro.

O animal transgênico pode expressar o gene exógeno em apenas alguns dos seus tecidos, em determinadas ocasiões de sua vida.

Os animais transgênicos podem servir de modelos animais para o estudo de doenças humanas, como tumores malignos. Também servem como modelos para terapia gênica. Até agora os animais mais usados como transgênicos são os camundongos, mas pode se utilizar em breve outros animais, como vacas transgênicas que secretam substâncias farmacêuticas importantes em seu leite.

**C) Detecção de doenças infecciosas:**

A aplicação da tecnologia do DNA recombinante a doenças infecciosas é direto, pois detecta-se uma seqüência de DNA de origem não humana, no interior de células. O potencial do diagnóstico molecular para a identificação de microrganismos é ilimitado.

A detecção de DNA ou RNA microbianos podem ser identificados por uma sonda que hibridiza com a seqüência alvo. A sonda é marcada com radioatividade, quimioluminescência ou marcadores enzimáticos, facilitando a visualização do híbrido.

O PCR é uma alternativa ao método da hibridização direta, podendo amplificar um pequeno segmento do DNA microbiano até que seja mais facilmente detectável.

Muitos *kits* de detecção de microrganismos usando hibridização direta ou PCR, já estão disponíveis e são utilizados de rotina em laboratórios clínicos, como a detecção do vírus da herpes simples (HSV), o citomegalovírus (CMV), o vírus da hepatite B (HBV), a *Chlamydia trachomatis*, causador de doença sexualmente transmissível e a detecção do vírus da AIDS, o HIV, que é realizado pela técnica de PCR.

**D) Detecção de doenças genéticas:**

Uma das principais técnicas moleculares para a detecção de doenças genéticas é o método do **polimorfismo do tamanho do fragmento de restrição (RFLP)**, que utiliza a hibridização por transferência de Southern.

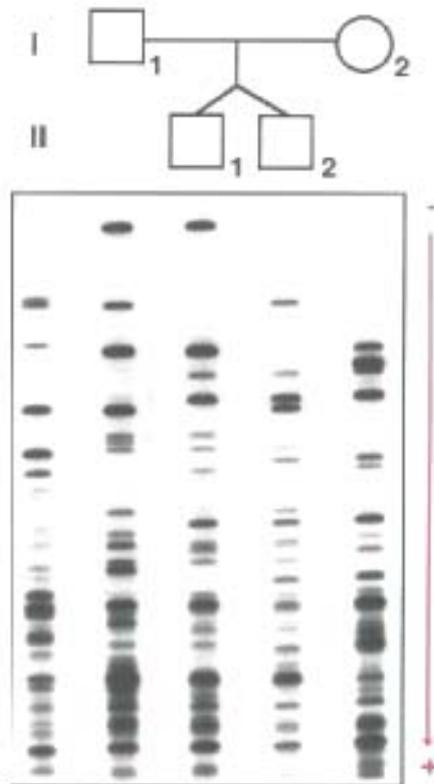
Um polimorfismo é uma variação accidental do genoma, que ocorre entre os genes e não afeta a expressão dos mesmos. Um RFLP serve como marcador genético útil, que permite estudar a hereditariedade de genes próximos. Um polimorfismo pode auxiliar, indicando que cromossomo uma criança herdou de que progenitor. Portanto, tenta-se descobrir marcadores genéticos polimórficos ligado a uma doença genética.

Diversos defeitos bioquímicos hereditários podem ser detectados ainda na fase pré-natal, por biópsias fetais. A condição de portador de alelos recessivos para doenças graves como a Fibrose cística pode ser determinada nos pais, assim como o estudo das bases genéticas de outras doenças como a hipercolesterolemia.

**E) *Fingerprint* do DNA e teste de paternidade:**

Um dos maiores desafios da medicina legal e da justiça é estabelecer com segurança a identidade de um indivíduo, vivo ou morto. Os testes do DNA recombinante têm a capacidade de substituir as impressões digitais convencionais usadas até então.

Dos 6 bilhões de pb que constituem o DNA de uma célula humana, uma pessoa difere da outra em 0,05% de seu genoma, Tais alterações são os polimorfismos mencionados no item anterior. Muitos dos polimorfismos ocorrem no **DNA espaçador**, que corresponde a segmentos curtos de DNA que se repetem no genoma. O zigoto é formado pela reunião ao acaso de metade dos elementos espaçadores de seu pai com metade dos elementos espaçadores de sua mãe.. Isto forma a base do *fingerprint* do DNA (Figura 4.31).



**Figura 4.31:** Autorradiograma de *fingerprint* de DNA de um casal e de seus dois filhos. A coluna mais da direita corresponde a um indivíduo não aparentado. Cada banda que ocorre nos filhos está presente em pelo menos um dos pais. (Adaptado de Muller & Young. **Emergy's Elements of Medical Genetics**. Churchill Livingstone, 1996).

O DNA, ao contrário das impressões digitais é muito estável. Uma amostra de sangue seco, um fio de cabelo, ou um pequeno fragmento de pele, pode ser testado tanto pela transferência de Southern quanto pelo PCR.

O *fingerprint* do DNA pode ser usado tanto em aplicações policiais, detectando o autor de crimes violentos como estupros, como identificando vítimas de acidentes e guerras. É amplamente empregado em testes de paternidade duvidosa, em que o polimorfismo do filho é comparado ao do suposto pai, com alto grau de confiabilidade.

O DNA de um indivíduo que possui um polimorfismo pode ser reconhecido pela transferência de Southern.

---

**REFERÊNCIAS BIBLIOGRÁFICAS:**

---

1. ALBERTS, B.; Bray, D.; Lewis, J.; Raff, M.; Roberts, K. & Watson, J.D. **Molecular Biology of the cell**. Garland Publishing Inc., N. York, 1994, 1361 p.;
2. BROWN, T.A. **DNA sequencing: the basics**. Oxford University Press, Oxford, 1ª edição, 101 p.;
3. DULBECCO, R. **Os genes e o nosso futuro**. Editora Best Seller, São Paulo, 1ª edição, 101 p.;
4. GRIFFITHS, A.J.F.; Miller, J.H.; Suzuki, D.T.; Levontin, R.C. & Gelbart, W.M. **An Introduction to genetic analysis**. W.H. Freeman & Company, N. York, 5ª edição, 1993, 840 p.;
5. LEWIN, B. **Genes V**. Oxford University Press, Oxford, 1994, 1272 p.;
6. MUELLER, R.F. & Young, I.D. **Emery's elements of medical genetics**. Churchill Livingstone, Edimburgo, 1995, 317 p.;
7. ROSS, D.W. **Introdução à medicina molecular**. Interlivros, Rio de Janeiro, 2ª edição, 1997, 204 p.;
8. WATSON, J.D. & Crick, F.H.C. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. **Nature**, **171**: 737-738, 1953;
9. WATSON, J.D. & Crick, F.H.C. Genetical implications of the structure of deoxyribonucleic acid. **Nature**, **171**: 964-967, 1953.