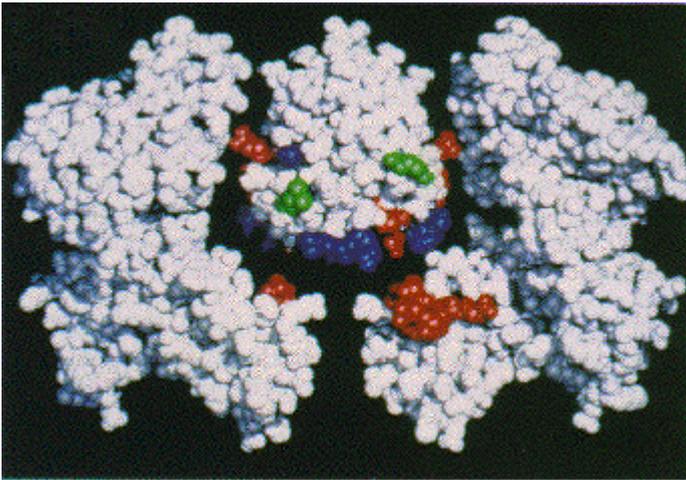


1 A GENÉTICA MOLECULAR DA CÉLULA NÃO PATOLÓGICA.1.2

1.1	CONCEITOS BÁSICOS DE GENÉTICA MOLECULAR.....	1.4
1.1.1	<i>O material genético.....</i>	<i>1.4</i>
1.1.1.1	Ácidos Nucleicos (DNA e RNA).....	1.4
1.1.1.1.1	A estrutura do DNA.....	1.6
1.1.1.1.2	A estrutura do RNA.....	1.16
1.1.1.2	A ORGANIZAÇÃO DOS GENES NO GENOMA.....	1.18
1.1.1.3	Estrutura de um Gene.....	1.20
1.1.1.4	O CÓDIGO GENÉTICO.....	1.27
1.1.1.5	Mutações e Polimorfismos.....	1.29
1.1.2	- A FISIOLOGIA DO GENE.....	1.31
1.1.2.1	Transcrição e transcrição reversa.....	1.31
1.1.2.2	Tradução.....	1.36
1.1.2.3	Replicação.....	1.42
1.1.2.4	Mecanismos de Mutação do DNA.....	1.47
1.1.2.4.1	Erros na replicação do DNA.....	1.47
1.1.2.4.2	Lesões espontâneas.....	1.48
1.1.2.4.3	Recombinação genética.....	1.50
1.1.2.4.3.1	Recombinação por crossing-over.....	1.50
1.1.2.4.3.2	Recombinação somática.....	1.52
1.1.2.4.4	Mutações mediadas por elementos de transposição.....	1.53
1.1.2.5	Mecanismos de reparação do DNA.....	1.61
1.1.2.5.1	Mecanismos não excisativos.....	1.61
1.1.2.5.2	Mecanismos excisativos.....	1.63
1.1.2.5.3	Mecanismos pós-replicativos.....	1.65
1.2	O CICLO CELULAR.....	1.66
1.2.1	<i>Divisão celular e anomalias genéticas.....</i>	<i>1.66</i>
1.2.2	<i>Ciclínas e CDK.....</i>	<i>1.68</i>
1.2.3	<i>Apoptose.....</i>	<i>1.76</i>
1.3	ACTIVIDADE CELULAR.....	1.78
1.3.1	<i>Componentes dos sistemas de sinalização.....</i>	<i>1.78</i>
1.3.2	<i>Regulação da sinalização por modulação da conformação proteica.....</i>	<i>1.79</i>
1.3.3	<i>A Superfamília dos receptores das hormonas esteróides.....</i>	<i>1.79</i>
1.3.4	<i>Receptores transmembranares; vias de transdução de sinal....</i>	<i>1.80</i>

Capítulo 1

1 A genética Molecular da Célula não Patológica



1.1 Conceitos básicos de Genética Molecular

1.1.1 O material genético

1.1.1.1 Ácidos Nucleicos (DNA e RNA)

Foi apenas em 1944 que Griffith demonstrou que a hereditariedade era transmitida pelos ácidos nucleicos. A experiência realizada demonstrou que a capacidade de matar um ratinho era conferida a uma estirpe bacteriana não virulenta, pelo DNA de uma outra estirpe bacteriana virulenta (fig.1.1).

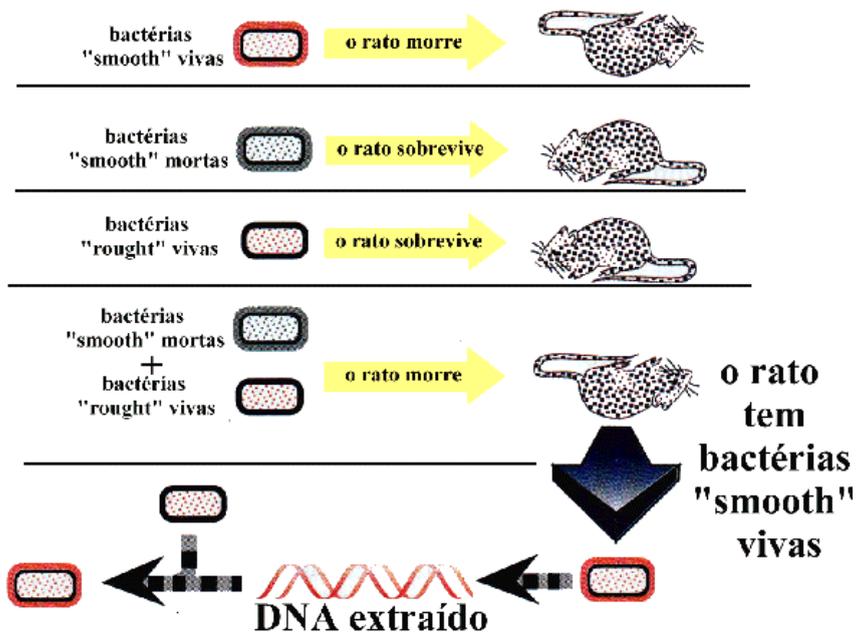


Fig. 1.1 – Experiência de Griffith demonstrando que a virulência bacteriana era uma característica genética transmitida pelos ácidos nucleicos.

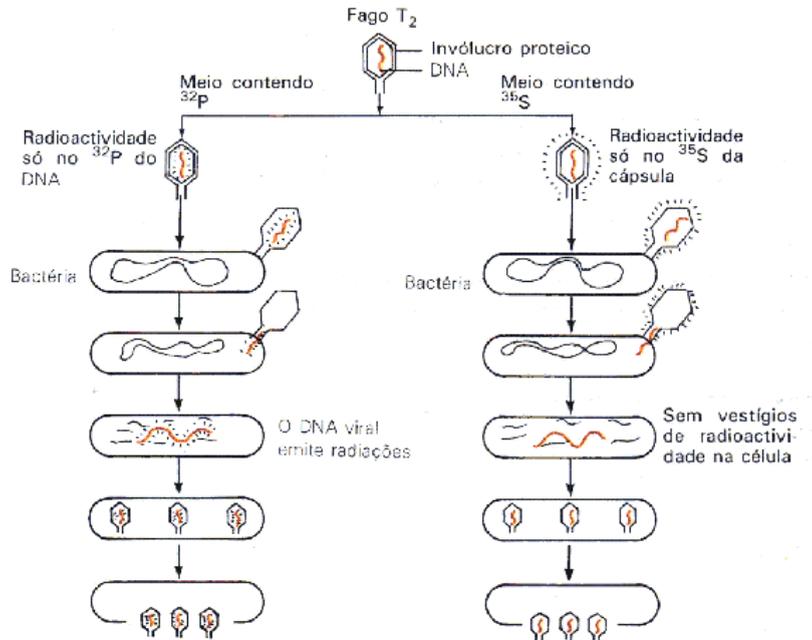


Figura 1.2 – Experiência de Hershley e Chase demonstrando que a transmissão genética nos fagos T2 estava associada ao DNA.

Em 1952, Hershley e Chase demonstraram que o fago T2 (um vírus que infecta bactérias) transmite os seu código genético à bactéria infectada, através da injeção do seu DNA na bactéria, alargando assim o número de organismos que demonstradamente utilizam o DNA como registo genético (fig.1.2).

Sabemos hoje que com a exceção de alguns tipos de vírus, todos os organismos utilizam o DNA como portador da sua informação genética. Os vírus que fogem a esta regra utilizam o RNA para o mesmo efeito.

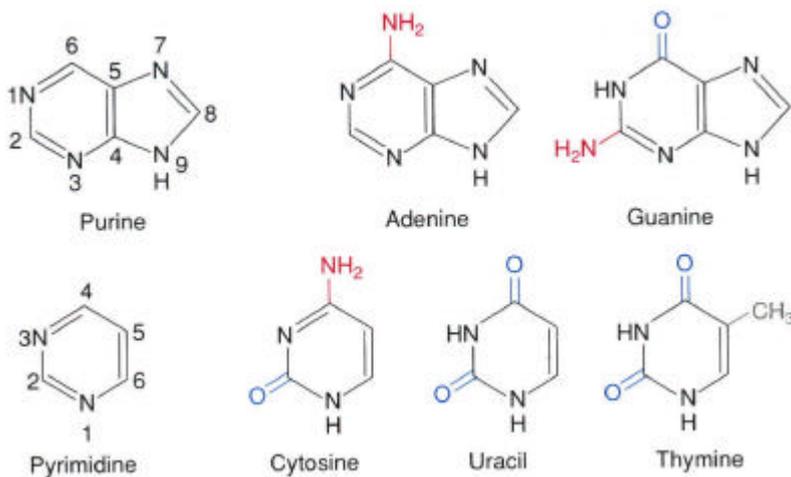


Figura 1.3 - As bases azotadas que entram na composição dos nucleótidos.

1.1.1.1 A estrutura do DNA

Com a aceitação generalizada por volta dos anos 50 de que a informação genética residia no DNA, a grande questão passou a ser o mecanismo de armazenamento dessa informação. Com efeito, nesta altura não se compreendia como é que um polímero tão simples (apenas constituído por 4 tipos de unidades diferentes) e que se pensava ser homogéneo em toda a sua extensão podia codificar a enorme variedade de proteínas que compunham os organismos. Para tal, houve necessidade de elucidar de forma precisa a estrutura dos ácidos nucleicos.

Sabemos hoje que os ácidos nucleicos são polímeros de nucleótidos. Cada nucleótido contém um anel heterocíclico de carbono com 5 átomos de azoto (a base nitrogenada), 1 anel de 5 carbonos (uma pentose) e um grupo fosfato. As bases nitrogenadas são de 2 tipos:

purinas e pirimidinas, sendo o número total de bases disponível de cinco (fig. 1.3). No entanto, cada tipo de ácido nucleico utiliza apenas 4 das cinco bases: o DNA contém Adeninas (A), Timidinas (T), Guaninas (G) e Citosinas (C); enquanto o RNA contém Adeninas (A), Uracilos (U), Guaninas (G) e Citosinas (C). As pentoses encontradas nos ácidos nucleicos são de 2 tipos: 2-desoxirriboses e riboses (Fig. 1.4) dando origem ao Ácido Desoxirribonucleico (DNA) e ao Ácido Ribonucleico (RNA).

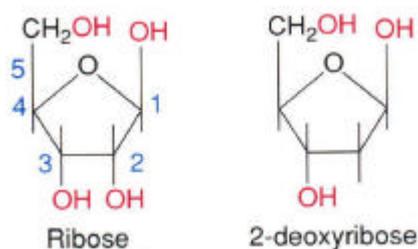


Figura 1.4 - As pentoses são o componente dos ácidos nucleicos que definem o seu tipo. A desoxirribose entra na composição do DNA, enquanto a ribose compõe o RNA

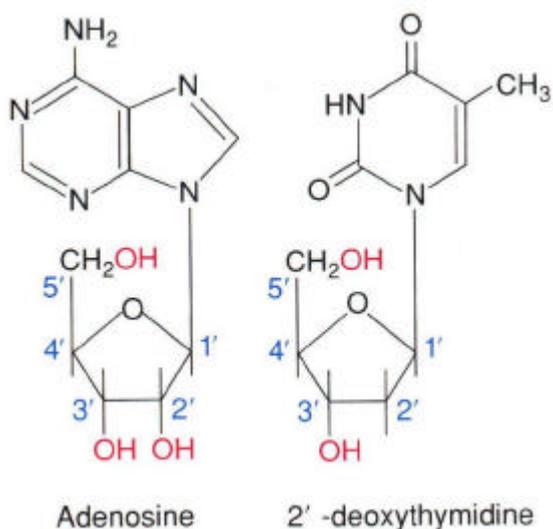


Figura 1.5 - Os nucleósidos são compostos por uma base e uma pentose.

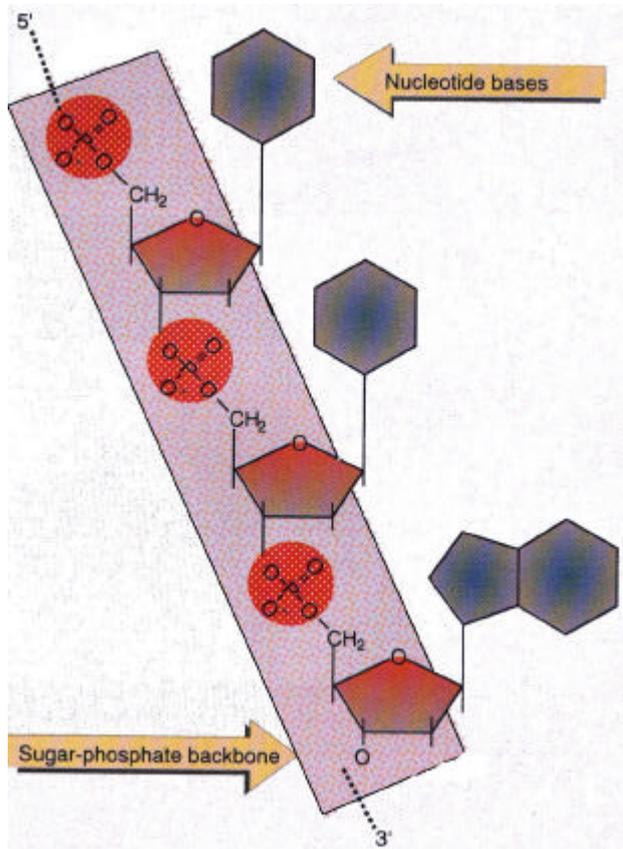


Figura 1.6 - Os nucleótidos unem-se por ligações fosfodiéster para dar origem aos ácidos nucleicos.

Os nucleósidos são compostos por uma base nitrogenada e uma pentose (fig. 1.5). Um nucleósido com um grupo fosfato no carbono 5 chama-se nucleótido. Os nucleótidos são as unidades de construção dos ácidos nucleicos, sendo unidas por uma ligação 5'-3': o carbono 5' da pentose de um nucleótido une-se ao carbono 3' da pentose do

nucleótido seguinte por uma ponte fosfodiéster, ficando a base nitrogenada exterior ao esqueleto da ligação (Fig.1.6).

Em 1953, uma importante descoberta realizada por Watson & Crick transformou a visão do material genético. Dados de difracção de raios X mostraram que o DNA tem a forma de uma hélice regular. Das dimensões obtidas para a hélice na difracção de raios X, e da densidade do DNA, inferiu-se então que a hélice era composta por duas cadeias polinucleotídicas, com as bases de cada cadeia viradas para o interior da hélice. As bases de cada hélice emparelham de tal modo que uma purina se opõe sempre a uma pirimidina. Estes dados, conjugados com a observação anterior de Chargaff indicando que independentemente da quantidade de cada base, a proporção G:C e A:T é sempre a mesma no DNA, indicam que G emparelha com C e A com T na dupla hélice do DNA. Watson & Crick propuseram que o emparelhamento não se realizava por ligação covalente, mas por pontes de hidrogénio entre as bases nitrogenadas (Fig. 1.7). Para tal, as 2 cadeias devem orientar-se de modo antiparalelo (Fig. 1.7). Obteve-se assim para o DNA o modelo ilustrado na figura 1.8.

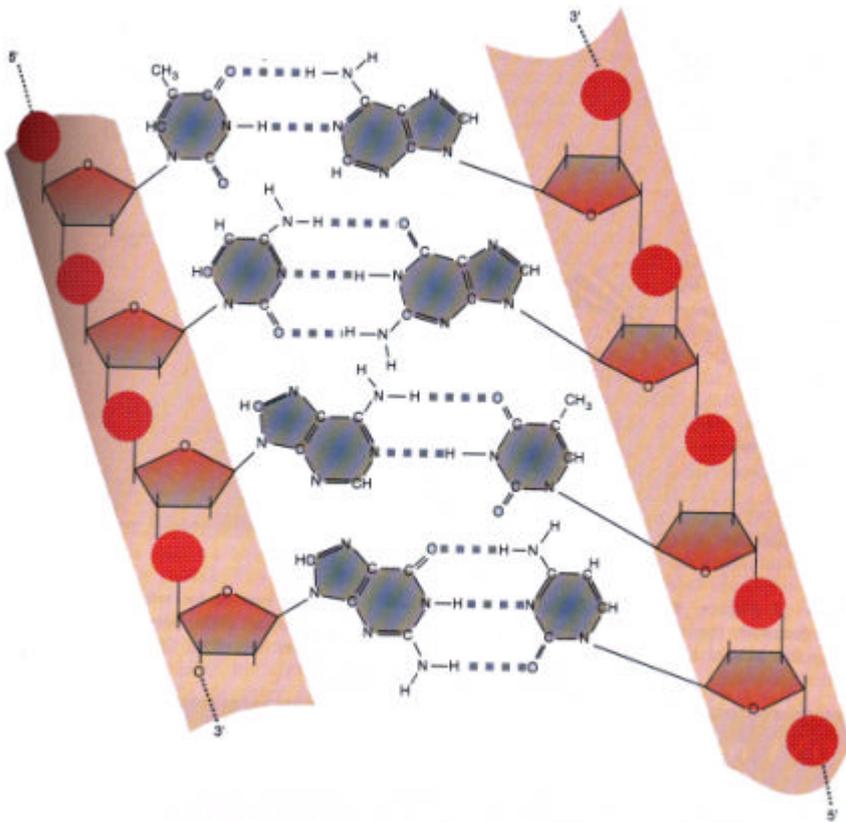


Figura 1.7 - O DNA é formado por duas cadeias com orientação antiparalela, com as bases de cada uma das cadeias a hibridizarem entre si.

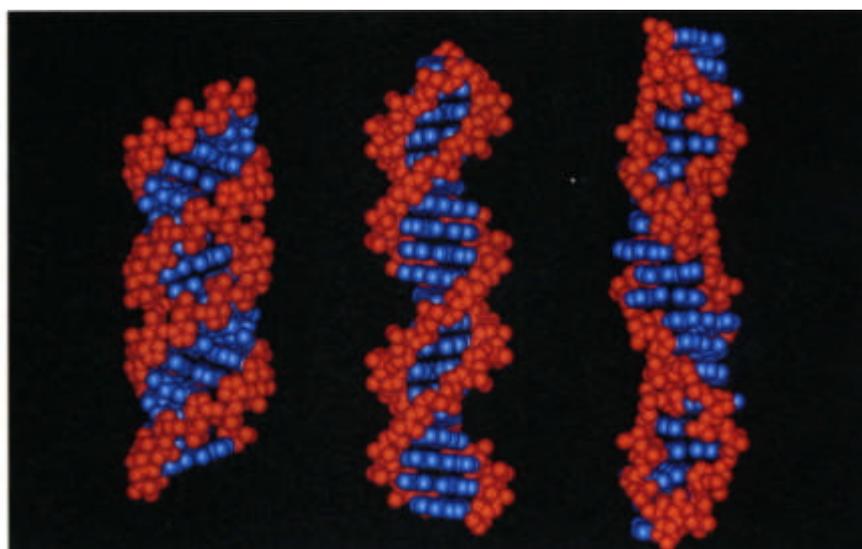


Figura 1.8 - Estruturas possíveis para a dupla hélice do DNA. O modelo da esquerda é o da forma A, o do meio o da forma B e do da direita o da forma Z.

A estrutura do DNA identificada por Watson & Crick, e ilustrada na figura 1.8B (forma B) é a que em situações fisiológicas é mais frequente. No entanto nem todo o DNA da célula se encontra nesta estrutura, e certamente *in vitro* é possível manipular as condições do meio, favorecendo outras conformações. Na tabela 1 encontram-se sumariadas as características das 4 conformações teoricamente possíveis para a conformação dos ácidos nucleicos, podendo na figura 1.8 ver-se comparativamente a conformação prevista.

Tabela 1.1 - Características dos tipos de hélice que o DNA pode tomar

Tipo de Hélice	N.º Bases por Volta	Rotação por par de Bases	Elevação por par de bases	Diâmetro da hélice
A	11	+32.7°	2.56 Å	23 Å
B	10	+36°	3.38 Å	19 Å
C	9.33	+38.6°	3.32 Å	19 Å
Z	12	-30°	3.71 Å	18 Å

A dimensão do material genético no Homem coloca o problema de como conseguir compactar 1,8m de DNA num núcleo que pode ser tão pequeno como 6µm (6×10^{-6} m). Este empacotamento tem ainda que ser flexível já que deve mudar ao longo do ciclo celular, aumentando durante as mitoses de tal modo que os cromossomas se tornam individualizados e visíveis ao microscópio óptico.

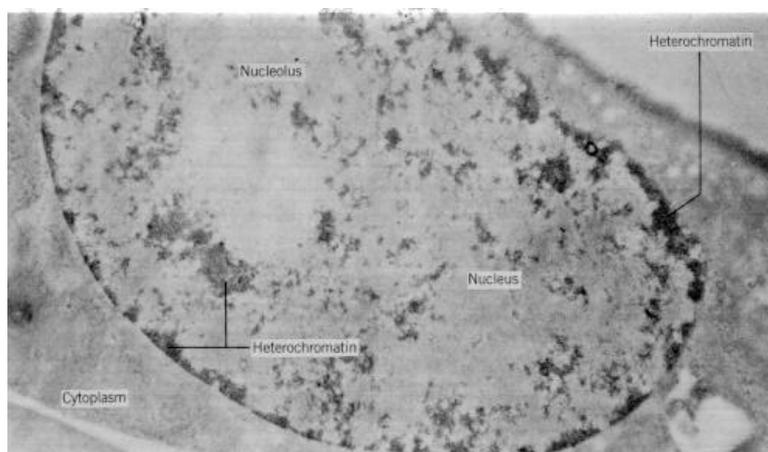


Figura 1.9 – Tipos de cromatina observáveis em interfase.

Durante a maior parte do ciclo celular a cromatina pode ser dividida em dois tipos de material genético (Fig. 1.9): a **euromatina** é a que ocupa a maior região do núcleo, sendo composta por material muito menos compactado que os cromossomas; a **heterocromatina** é composta por material muito compactado, formando fibras (encontra-se num estado intermédio entre a compactação dos cromossomas e a relativa descompactação da euromatina). A heterocromatina e a euromatina não representam fibras de DNA diferentes, já que as mesmas fibras passam pelas duas zonas do núcleo. Constituem assim partes das fibras com diferentes estados de condensação: a heterocromatina é constituída por regiões do DNA que não são habitualmente expressas na célula em causa, enquanto os genes expressos se localizam na euromatina (muito embora os genes na euromatina não estejam todos a ser expressos).

Em 1974, foi descoberta a estrutura básica de organização da cromatina em todos os eucariotas. Esta subunidade organizativa básica (nucleosoma) contém cerca de 200 bp de DNA, organizados por um octâmero de proteínas pequenas e básicas (histonas) numa estrutura tipo rosário em que o DNA se localiza no exterior das “contas”, e as proteínas no seu interior (Fig.1.10). Esta organização explica porque os locais de ligação a proteínas se encontram por vezes tão espaçados na

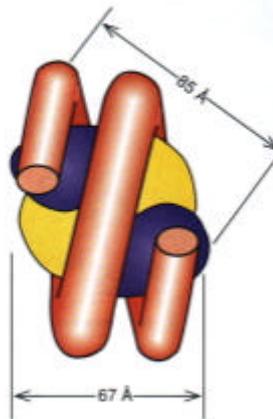


Figura 1.10 - A dupla hélice de DNA dá duas voltas ao núcleo central de proteínas do nucleosoma

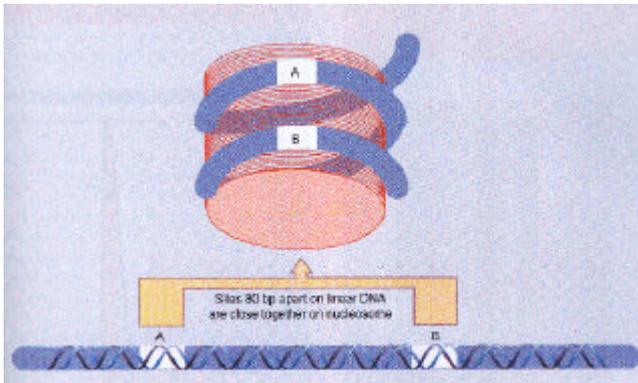


Figura 1.11 - A organização do DNA nos nucleosomas coloca próximas seqüências de DNA distantes na seqüência linear

seqüência do DNA (Fig. 1.11). O octâmero de histonas é constituído por 2 cadeias de cada uma das histonas H1, H2A, H2B e H3, existindo ainda, por vezes, uma 5ª histona (H1) a estabilizar as 2 voltas de DNA ao octâmero (Fig. 1.12).

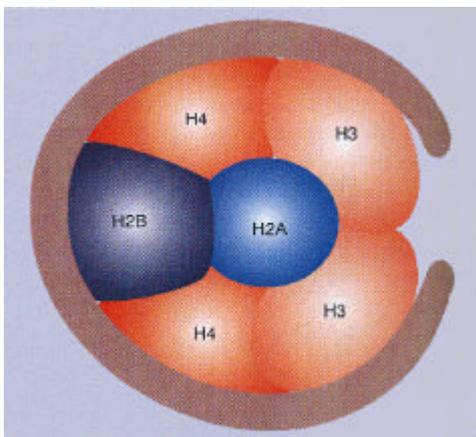


Figura 1.12 – Vista de cima da organização das histonas no nucleosoma. Existe ainda um par H2A/H2B por baixo, e por vezes uma histona H1 exterior ao nucleosoma

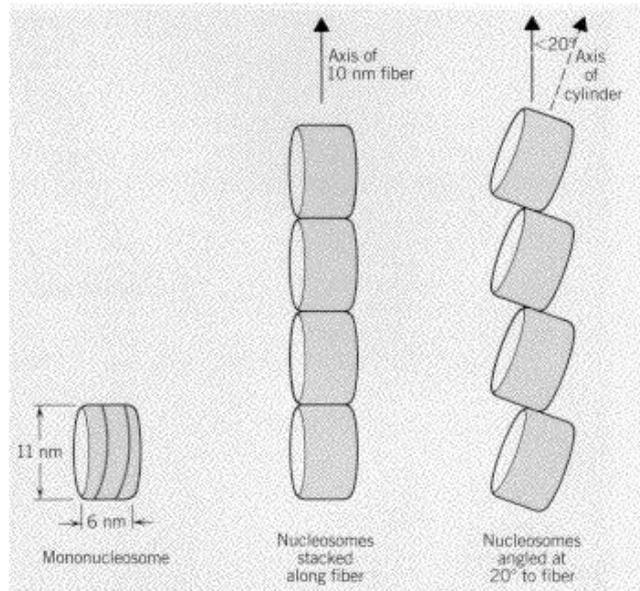
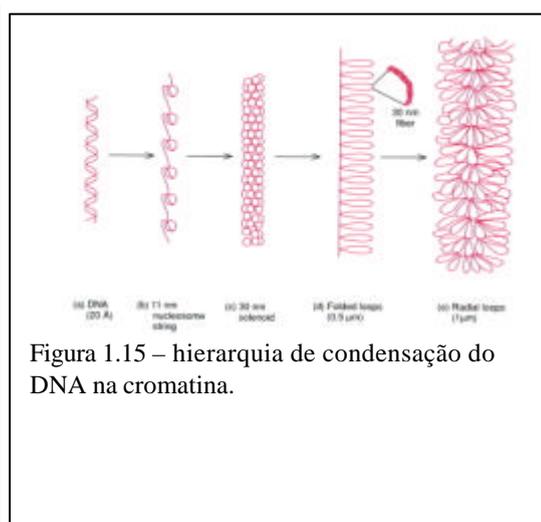
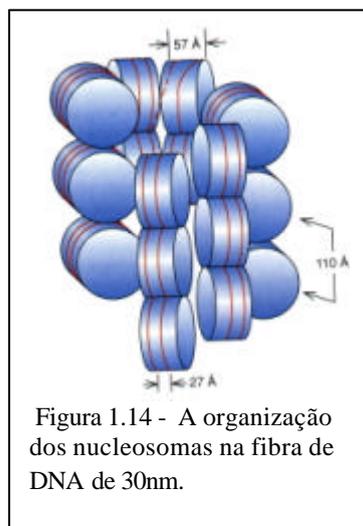


Figura 1.13 - A compactação das histonas na fibra de DNA de 10nm

A análise da cromatina ao microscópio electrónico revelou a existência de 2 tipos de fibras: a fibra de 10nm e a 30nm. A fibra de 10nm é essencialmente 1 sequência continua de nucleosomas (Fig. 1.13). Esta fibra ocorre em condições de baixa força iónica, e na ausência de histonas H1. Em condições de alta força iónica e na presença da histona H1, forma-se a fibra de 30nm, a qual é essencialmente um enrolamento de 6 nucleosomas por volta (Fig. 1.14).

A transcrição (cópia dos genes em mRNA), como veremos no capítulo 1.1.2.1, envolve a deslocação no DNA de uma complexa maquinaria enzimática, e inclui a abertura da dupla cadeia do DNA. Este facto, não é compatível com um elevado grau de empacotamento das fibras do DNA, pelo que se compreende que os genes transcripcionalmente activos se localizem na eucromatina. No entanto, os resultados experimentais indicam que a estrutura dos genes transcripcionalmente

activos envolve o empacotamento em nucleosomas, ainda que seja necessário admitir que durante a transcrição estes sejam temporariamente “desmontados” pela maquinaria enzimática.

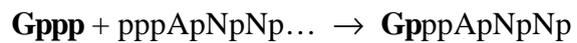


1.1.1.1.2 A estrutura do RNA

Na célula existem várias formas de RNA, as quais possuem estruturas e funções diferentes:

- **mRNA** - O mRNA ou RNA mensageiro, é a espécie de RNA que transporta a informação para a síntese das proteínas no ribossoma. O mRNA é formado no núcleo na transcrição do DNA, passando ainda por uma fase de processamento antes de atingir o citoplasma na forma madura (mRNA). O processamento efectuado inclui o “Splicing”, isto é a remoção das sequências não codificantes ou introns. Outras alterações introduzidas no processamento que ocorre no núcleo consistem na adição de uma cauda poli-adenina à extremidade 3’, e

metilação CAP da extremidade 5'. A estrutura CAP resulta da ligação de um G à purina com que a transcrição habitualmente se inicia, ficando este G na orientação inversa, e ligado pelo trifosfato deixado livre pela purina:



G sofre então uma ou mais metilações.

- **tRNA** - o tRNA ou RNA de transporte é um tipo de RNA que se encontra covalentemente ligado a um aminoácido, tendo como função o transporte do aminoácido para o ribossoma, onde este vai ser posicionado com precisão, sempre que o ribossoma estiver a ler um codão complementar do triplete que o tRNA possui (anticodão). As 64 espécies de tRNA (correspondentes aos 64 codões), possuem uma estrutura básica semelhante.

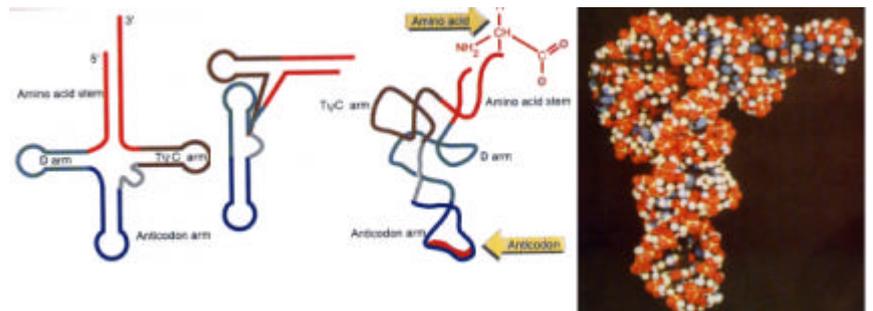


Figura 1.16 – Estrutura do tRNA

rRNA - trata-se do RNA ribossomal, o qual é como o nome indica um dos componentes dos ribossomas. O rRNA constitui a maior parte da massa do ribossoma, e provavelmente todas as proteínas do ribossoma se associam ao rRNA. Assim, o rRNA forma como que o esqueleto do ribossoma, determinando a posição das várias subunidades proteicas.

1.1.1.2 A ORGANIZAÇÃO DOS GENES NO GENOMA

O genoma pode, de uma forma genérica, ser classificado em DNA não repetitivo e DNA repetitivo. A abundância relativa dos dois tipos de DNA podem ser experimentalmente determinados, tendo por base a diferente cinética de re-hibridação (DNA repetitivo encontra mais rapidamente uma sequência complementar com quem pode hibridar). O DNA não repetitivo representa sequências únicas, ou seja genes de cópia única no genoma. O DNA repetitivo é constituído por DNA moderadamente repetitivo, representando genes com várias cópias no genoma, e DNA altamente repetitivo. A função do DNA altamente repetitivo permanece até ao momento uma incógnita. Como já vimos, um exemplo deste tipo de DNA é o que existe nos telómeros, onde provavelmente tem a função de estabilizar o cromossoma. Existem no entanto, repetições de pequenas unidades de sequências espalhadas pelo genoma (mini e microsatélites), os quais constituem em pequenas sequências, repetidas um determinado número de vezes. O número de repetições é em muitos casos altamente polimórfico, pelo que estas sequências têm sido utilizados como marcadores no mapeamento genético.

Sequências moderadamente repetitivas: Nos genomas eucarióticos, os genes que existem em cópia única são poucos. Na maior parte dos casos, existem sequências com alguma similaridade, algumas das quais não funcionais (os pseudogenes). A vantagem da existência de mais que uma cópia dos genes é óbvia já que assim os organismos podem conservar uma cópia intacta do gene, mutando a outra, numa tentativa de evoluir. Neste processo de evolução, algumas cópias ficam com a sua funcionalidade comprometida, tornando-se **pseudogenes**. No entanto, uma vez que uma outra cópia funcional existe, nenhum efeito nefasto daí ocorre para o organismo.

Um conjunto de genes que descende por duplicação e variação de um gene ancestral é chamado de família génica. Os seus membros podem estar arrançados em grupos sequenciais (“gene clusters”), dispersos no

genoma (muitas vezes mesmo em cromossomas diferentes), ou numa combinação de ambos os arranjos. Os “gene clusters” podem conter desde 2 até centenas de genes idênticos, alinhados em sequência. A dispersão dos genes ocorre por translocação de um gene após a duplicação. Os membros de um “gene cluster” têm função similar, mas podem ser expressos em tipos celulares diferentes ou em diferentes condições (Ex. Gene da globina). Em alguns casos, o gene cluster responde à grande necessidade de proteínas ou de RNA (ex.: rRNA e histonas).

Sequências altamente repetitivas: As sequências altamente repetitivas tomam a forma de sequências muito curtas, repetidas muitas vezes em sequência. Formam-se assim blocos de material genómico, consistindo cada bloco em longas repetições de uma unidade. Em alguns casos as unidades são rigorosamente iguais, noutras são relacionadas. A repetição sequencial de unidades de sequência forma blocos de DNA com características físicas distintas do resto do genoma, o que pode ser utilizado para as isolar. Uma das propriedades físicas do DNA que depende da sequência é a densidade, a qual depende do conteúdo GC. A densidade é habitualmente determinada mediante a centrifugação do DNA num gradiente de Cloreto de Césio (CsCl). O DNA forma assim bandas correspondentes a sua própria densidade. Quando este procedimento é realizado para DNA genómico eucariota, forma um pico algo largo, consistindo numa mistura de sequências com densidades próximas (a banda principal). Por vezes forma-se ainda um ou mais picos adicionais, de menor intensidade. A este material chama-se o DNA satélite.

O DNA satélite existe no genoma de várias espécies eucariotas, pode ter uma densidade superior ou inferior à banda principal, mas representa habitualmente menos de 5% do genoma total.

O DNA satélite encontra-se frequentemente localizado na heterocromatina, não sendo habitualmente possível encontrar as suas sequências entre o RNA.

Nos mamíferos, as sequências que compõem cada satélite mostram divergência apreciável entre as repetições de cada. Habitualmente existem sequências curtas predominantes, mas outras relacionadas com estas, mas contendo adições, substituições e deleções formam o restante satélite. Frequentemente, pode observar-se uma hierarquia nas repetições dos satélites, com uma sequência base a repetir-se, a qual por vezes sofre modificações, as quais por sua vez se repetem também de forma mais ou menos cíclica. Este facto originou uma hierarquia de nomenclatura: DNA satélite, minisatélites, microsátélites.

1.1.1.3 Estrutura de um Gene

A comparação directa entre a sequência do DNA de um gene, e a sequência da proteína respectiva, permite determinar se o gene e a proteína são ou não colineares: se a sequência do gene corresponde exactamente a sequência de aminoácidos da proteína. Nas bactérias e vírus, a equivalência é perfeita: cada gene contém uma sequência continua de nucleótidos, cuja sequência e comprimento está directamente relacionada com a da proteína. Quando falamos em correspondência entre o gene e a proteína, estamos no entanto a simplificar o que realmente se passa. Como veremos mais tarde, um gene não codifica directamente uma proteína, já que a informação tem que passar por um estado intermédio: o RNA. Assim, mesmo o mais simples dos genes tem que conter sequências de vários tipos:

- **sequências reguladoras ou não codificantes:** sequências que permitem à célula controlar que genes estão activos em cada momento, possibilitando assim uma resposta diferenciada dependente das necessidades de cada momento. As sequências reguladoras podem existir em cada extremidade do gene, e em

alguns casos estar mesmo bastante distanciadas das sequências codificantes.

- **sequências codificantes:** sequências que são directamente transcritas para RNA, e deste codificadas em proteínas. Note-se que enquanto o DNA é de cadeia dupla, o RNA é de cadeia simples, pelo que apenas uma das cadeias do DNA pode ser idêntica a do RNA (codificante ou +), sendo a outra cadeia complementar do RNA (-).

Como acima foi dito, o gene não é no entanto tão simples nos eucariotas. Ao contrario das bactérias e vírus, nos organismos eucariotas, os genes e as proteínas não são colineares, isto é, a região codificante dos genes (exons) é interrompida a espaços irregulares por sequências não codificantes (introns). Este facto faz com que nos eucariotas, a expressão genica envolva um passo adicional: o splicing do RNA, ou processamento do RNA (com exons e introns) em mRNA.

Um promotor é uma sequência de DNA, habitualmente na extremidade 5' de um gene, com a função de se ligar a proteínas, e controlar a iniciação da transcrição. As proteínas a que um promotor se deve ligar, são várias, disso dependendo a sua dinâmica funcional. Genericamente pode falar-se de proteínas repressoras, proteínas activadoras, e da RNA polimerase. As proteínas repressoras, ao ligar-se ao promotor impedem a ligação da RNA polimerase, impedindo assim o iniciar da transcrição, enquanto a ligação das proteínas activadoras tem o efeito inverso. As propriedades do promotor que lhe conferem afinidade para as diversas proteínas dependem da sua sequência, pelo que esta varia de gene para gene, conferindo aos diversos genes características de regulação diferentes. No entanto, a ligação a polimerase do RNA é universalmente necessária, pelo que deve ser possível encontrar uma sequência "consenso" para os promotores. Esta sequência consenso

consiste na sequência mínima comum entre os vários promotores, e deve incluir a sequência absolutamente necessária para a ligação a polimerase do RNA.

Para os procariotas foi possível definir a região 44-50bp "upstream" do ponto de iniciação até 20bp "downstream" com sendo a região que interaccua com a polimerase do RNA, tendo sido definida uma sequência consenso consistindo de vários padrões:

Pribnow box ou sequência -10- imediatamente upstream do ponto de iniciação (-18 a -12) existe uma região com a sequência T₈A₉T₄A₆A₅T₉ (os números representam a frequência com que as bases ocorrem). A função desta sequência parece ser a de permitir que após a ligação da polimerase do RNA esta possa iniciar a sua evolução ao longo do gene, possivelmente por permitir a iniciação da abertura da cadeia do DNA (o facto de ter alto conteúdo AT facilita a abertura da dupla hélice).

Sequência de reconhecimento ou Sequência -35 - O seu nome deriva do facto de esta ser parte da sequência que a polimerase tem que reconhecer, mas que não fica fortemente ligada a esta. A sequência consenso é: T₈T₈G₇A₆C₅A₄. A função desta região parece ser a de conferir a capacidade de ligação a polimerase do RNA.

No caso de organismos eucariotas, o estudo dos promotores é bem mais complexo, já que existem não uma RNA polimerase, mas três. A acrescentar a esta dificuldade, está o facto de não se conhecer com precisão todos os componentes da maquinaria de transcrição eucariota, pelo que os estudos *In viro* ficam comprometidos.

À partida 2 particularidades existem nos eucariotas, relativamente ao que se passa nos procariotas: 1) o promotor da polimerase III fica localizado downstream do gene; 2) não é possível conhecer as particularidades do promotor da polimerase I, já que esta transcreve apenas os genes dos rRNA os quais são todos idênticos.

No entanto o promotor da RNA polimerase II, a responsável pela transcrição da maioria dos genes nos eucariotas são conhecidos com alguma profundidade. As principais sequências consenso identificadas nos promotores da RNA polimerase II dos eucariotas são:

TATA BOX - sequência consenso: $T_{82}A_{97}\frac{A_{63}}{T_{37}}A_{83}\frac{A_{50}}{T_{37}}$ Também

conhecida por Hogness box. Trata-se de uma sequência quase universalmente presente em mamíferos, aves, anfíbios e insectos. Posiciona-se a uma distância do ponto de iniciação entre 19 e 27bp. Como pode ver-se da sequência consenso, a TATA Box é constituída quase exclusivamente por AT, sendo as mutações que inserem um GC muito raras. Esta sequência é habitualmente rodeada por sequências ricas em GC, o que pode ser importante para a sua função.

CAAT BOX - sequência consenso $GG\frac{T}{C}CAATCT$. Esta sequência esta presente em alguns promotores, mas não em todos. A sua distancia ao ponto de iniciação ronda os 70 a 80bp.

As análises *In vitro* identificaram uma estrutura semelhante ao promotor bacteriano, imediatamente upstream do ponto de iniciação. No entanto, estudos *In vivo* revelaram a dependência de zonas ainda mais upstream da TATA box. Este componente pode consistir em duas regiões, uma entre -80 e -110 e a outra entre -50 e -70. Esta última pode ou não conter a CAAT box. Juntos, estas duas regiões têm uma forte influência na frequência de iniciação, possivelmente por influenciar a ligação da RNA polimerase II.

Junto ao ponto de iniciação, em redor da TATA box existe um componente que parece não ter influência na frequência de iniciação, antes de terminando o ponto de iniciação. Na ausência deste elemento, a transcrição tem uma iniciação errática.

Os promotores eucarióticos são bem mais complexos que os procariotas. Ao contrário dos promotores procarióticos, e contrariamente ao que até agora assumimos, um promotor eucariótico não funciona só. A sua actividade é enormemente aumentada de acordo com a regulação efectuada por outro tipo de sequências reguladoras: os “enhancers”. Estas sequências são distinguíveis dos promotores devido a duas características essenciais:

- a sua posição relativamente ao promotor é muito variável, podendo ser considerável, e funcionando em qualquer sentido (“upstream” ou “downstream”) e orientação.
- Um enhancer não actua apenas num promotor, podendo interactuar com qualquer promotor colocado na sua área de influência.

Vários vírus contêm enhancers. Destes, os mais perigosos para a célula que o vírus infecta são os enhancers presentes nos retrovírus. Como estes vírus se integram no genoma da célula infectada, a presença de enhancers pode levar à activação de um ou mais genes celulares que de outra forma estariam silenciosos. Desta forma, os retrovírus podem de forma indirecta, originar patologias, mesmo no seu estado “dormente”, já que mesmo na ausência de transcrição viral, podem induzir alterações no programa genético da célula infectada.

O modo de funcionamento dos enhancers permanece desconhecido. Foram no entanto colocadas várias possibilidades, entre as quais:

- **Formação de estrutura no DNA em cadeia Z** (ver figura 5). Os enhancers contêm habitualmente uma sequência alternada de pirimidinas-purinas. Esta sequência tem elevada probabilidade de formar uma estrutura em zDNA. Se, por um lado, o modo como esta estrutura poderia afectar a transcrição não está esclarecido, por outro lado, este

mecanismo poderia explicar porque os enhancers funcionam independentemente da sua orientação.

- Ligação do DNA a uma estrutura como a matriz nuclear
- ligação directa à polimerase

Quando a polimerase do RNA inicia a transcrição, esta prossegue com o complexo enzimático a percorrer o DNA, até que a enzima encontra um sinal para cessar a actividade. Neste ponto, a enzima pára de adicionar nucleótidos, liberta a cadeia de RNA nascente e dissocia-se do DNA. Assim, a terminação envolve a quebra de todas as pontes de hidrogénio entre o DNA e o RNA, e a reassociação da dupla hélice do DNA. A sequência de DNA que dá o sinal para que este processo ocorra chama-se terminador (ou abreviadamente **t**). Em alguns genes procarióticos, existem factores denominados anti-terminadores, que permitem à polimerase continuar a transcrição passando por um terminador, num processo chamado de “read-through”). Assim, a terminação não constitui simplesmente uma forma de terminar a transcrição, mas também uma forma de controlar esta, já que a existência dos anti-terminadores pode determinar a transcrição ou não de determinados genes que se encontrem após o terminador.

Pouco se sabe dos terminadores dos genes eucarióticos. A principal dificuldade no estudo dos terminadores em eucarióticos é a incerteza quanto ao local de terminação da transcrição. Ainda que a maior parte das espécies de mRNA eucarióticas conhecidas possuam extremidades 3' bem definidas, é muito difícil saber se esta extremidade foi produzida por terminação ou por processamento. No caso dos produtos da polimerase II, o problema é exacerbado pelo extenso processamento que ocorre com a adição da cauda poli-A. Pelo menos em alguns casos foi possível determinar que a extremidade 3' observada no RNA é de facto originada por corte de uma cadeia de RNA mais longa.

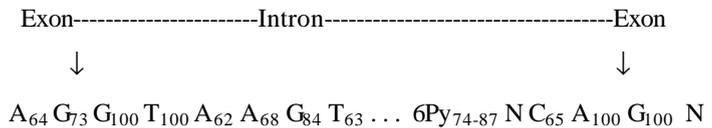
Estudos efectuados com sequências de histonas (não poliadeniladas), permitiram verificar que o mRNA termina numa estrutura semicircular (“stem-loop”). Com efeito, mutações que impeçam a formação desta estrutura, impedem a terminação, enquanto que outras mutações que revertam a mesma estrutura, embora com uma sequência diferente, restauram a terminação. Assim, a estrutura parece mais importante que a sequência que a determina.

Os genes eucarióticos e procarióticos diferem numa característica essencial. Ao contrário dos genes procarióticos, os genes dos organismos eucarióticos não são contínuos, mas interrompidos. Significa isto, que no meio das sequências codificantes, surgem sequências que têm que ser retiradas do RNA, antes de este poder servir de molde à construção das proteínas. Este processo de transformação que o RNA sofre nos organismos eucarióticos é chamado de processamento, ocorre no núcleo, e como veremos envolve não apenas a remoção das sequências extra (“splicing”) como outras transformações químicas.

Os genes eucarióticos são assim formados por dois tipos de sequências transcritas (isto é copiáveis para RNA) os exons e os introns (também chamados de intervening sequences). Os primeiros compõem as sequências que estarão presentes no RNA maduro, sendo os segundos as sequências que serão removidas durante o splicing.

A comparação das sequências nucleotídicas nas extremidades dos exons permite descrever as suas características:

- Não existe homologia extensa entre as duas extremidades de um intron, o que exclui a possibilidade da formação de uma estrutura secundária que determine os pontos de corte.
- As junções possuem uma sequência consenso conservada mas curta, a qual pode estar envolvida no processo de splicing:



1.1.1.4 O CÓDIGO GENÉTICO

A descoberta do código genético pretendeu responder à questão já por nós formulada (Cap. 1.1) sobre o mecanismo que permite aos ácidos nucleicos, com uma estrutura baseada em apenas quatro tipos de nucleótidos, conter a informação que codifica um enorme número de proteínas, as quais possuem 20 tipos de aminoácidos. A elucidação do código genético pretendeu ainda explicar como é que a expressão génica é regulada. No entanto, antes de esta questão poder ser estudada, era necessário estabelecer definitivamente a veracidade do dogma central da genética: Um gene - uma cadeia polipeptídica. Uma característica essencial do DNA é que a sua estrutura básica é independente da sequência (ao contrário das proteínas cuja conformação é directamente dependente da sequência). Assim, a sequência do DNA não parece ser importante devido à conformação, mas porque codifica uma sequência bem definida de aminoácidos. Note-se que o próprio conceito de que uma proteína contém sequências bem definidas de aminoácidos data dos anos 50 (a caracterização da insulina por Sanger), e portanto é estabelecida sensivelmente na mesma altura que se estuda a informação genética. A esta relação entre a sequência do DNA e a sequência proteica correspondente chamou-se código genético.

Como vimos, a sequência nucleotídica tem que conter informação suficiente para codificar aminoácidos diferentes. Como só há quatro tipos de nucleótidos no DNA, um cálculo simples indica que são necessários 3 nucleótidos (um tripleto ou codão) para codificar um aminoácido. As combinações possíveis com três nucleótidos são $4^3=64$, pelo que o código genético é degenerado, isto é, mais do que um tripleto deve codificar o mesmo aminoácido (tabela 1.2).

Tabela 1.2 - Código genético: significado dos 64 codons

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

Podem agrupar-se os códons segundo o aminoácido que codificam (tabela 1.2). Quando tal é realizado, pode observar-se que com frequência, a base na 3ª posição não é significativa, porque os 4 códons com as mesmas 1ª e 2ª bases codificam o mesmo aminoácido (Tabela 1.2). Por vezes apenas distingue entre uma pirimidina e uma purina a 3ª posição. A esta especificidade reduzida na 3ª base chama-se degenerância da 3ª base. Esta característica, em conjunto com a tendência para aminoácidos semelhantes (isto é polares, hidrofóbicos, etc.) serem codificados por códons relacionados minimiza o efeito das mutações.

Três codões não codificam aminoácidos. Como se pode observar na tabela 2, estes codões (UUA, UAG e UGA) indicam o fim da sequência génica, sendo por isso chamados de codões stop.

O código genético foi inicialmente estudado na bactéria *E.Coli*, pelo que a universalidade deste necessitou de extenso estudo. Sabemos hoje, que genericamente o código genético é similar em todos os organismos vivos estudados. As exceções conhecidas são representadas por pequenas alterações em algumas espécies de microorganismos, e no código genético mitocondrial, o qual possui algumas particularidades em alguns organismos (Tabela 1.3).

Tabela 1.3 - Exemplos de exceções à universalidade do código genético

Organismo	Codon	Significado Provável na mitocôndria	Significado habitual
Todos	UGA	Triptofano	Terminação
Levedura	CUA	Treonina	Leucina
Mosca da fruta	AGA	Serina	Arginina
Mamíferos	AGA	Terminação	Arginina
	AUA	Metionina	Isoleucina

1.1.1.5 Mutações e Polimorfismos

Uma vez que o código genético é lido em tripletos não sobreponíveis, a inserção ou remoção de um nucleótido causa uma alteração na fase de leitura, alterando os codões subsequentes. Este tipo de mutação é denominado em Inglês “**frameshift**”. Mutações deste tipo são susceptíveis de reverterem através da mutação inversa, isto é, se a primeira mutação foi uma inserção e a segunda uma deleção, ou vice-versa, apenas a zona do gene situada entre as duas mutações se encontra mutada. A segunda mutação, é denominada supressora, já que suprime o efeito da primeira, limitando a zona atingida (fig. 1.17)

As **mutações pontuais** são mutações que ocorrem devido à substituição de um nucleótido por outro. A forma mais frequente de mutações pontuais é a **transição**, a qual ocorre quando uma pirimidina é substituída por outra, ou uma purina por outra. A **transversão** é menos frequente, e implica a substituição de uma pirimidina por uma purina, ou vice-versa. As mutações pontuais podem ser de 3 tipos, de acordo com o efeito que provocam no aminoácido codificado. Se não afectam o aminoácido codificado são chamadas **silenciosas**, se mudam o aminoácido codificado são chamadas **missense**, e se transformam o codão num codão stop são chamadas **nonsense**.

As mutações pontuais foram durante muito tempo consideradas as principais causas de mutações. Sabe-se no entanto hoje, que as deleções são também muito frequentes, representando uma significativa porção das mutações identificadas.

As mutações podem ser vantajosas, desvantajosas ou neutras, segundo as consequências funcionais que provocam. As mutações neutras, apesar de causarem alteração na sequência não ocasionam mudança funcional. Neste caso, deve falar-se em polimorfismo e não em mutação.

Selvagem	GCU	GCU	GCU	GCU	GCU	GCU	GCU	GCU	GCU	
	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	
Inserção (+)	GCU	GCU	AGC	UGC	UGC	UGC	UGC	UGC	UGC	U
	Ala	Ala	Ser	Cys	Cys	Cys	Cys	Cys	Cys	
Deleção (-)	GCU	GCU	GCU	GCU	GCU	_CUG	CUG	CU		
	Ala	Ala	Ala	Ala	Ala	Leu	Leu			
Duplo mutante (+ -)	GCU	GCU	AGC	UGC	UGC	_UCU	GCU	GCU		
	Ala	Ala	Ser	Cys	Cys	Ser	Ala	Ala		
triplo mutante (+ + +)	GCU	GAC	UGC	AUG	CUG	CAU	GCU	GCU	GCU	
	Ala	Asp	Cys	Met	Leu	His	Ala	Ala	Ala	
triplo mutante (- - -)	GCU	_CUG	CU_C	UGC	U_CU	GCU	GCU			
	Ala	Leu	Leu	Cys	Ser	Ala	Ala			

Figura 1.17 - Mutações frameshift e seus efeitos. Note-se que as inserções e as deleções podem anular-se mutuamente, fora da zona entre as duas mutações

1.1.2 - A FISILOGIA DO GENE

1.1.2.1 Transcrição e transcrição reversa

O RNA é a espécie de ácido nucleico com um papel mais alargado na genética molecular dos organismos. Não só o RNA tem o papel mais “mediático” de mensageiro, mas é também a espécie que assegura a descodificação da informação genética (o rRNA e o tRNA). Para além destes papéis centrais em todos os organismos, existem ainda vírus que utilizam o RNA como material de armazenamento de informação genética (os retrovírus). A produção do RNA tem habitualmente uma origem comum: a transcrição do DNA. No caso do mRNA, o produto formado é um intermediário cuja função requer ainda a tradução. No caso do tRNA e do rRNA, o produto formado é o efector da função a que se destina.

A transcrição é talvez o passo por excelência para a regulação da expressão génica. A decisão principal na regulação de um

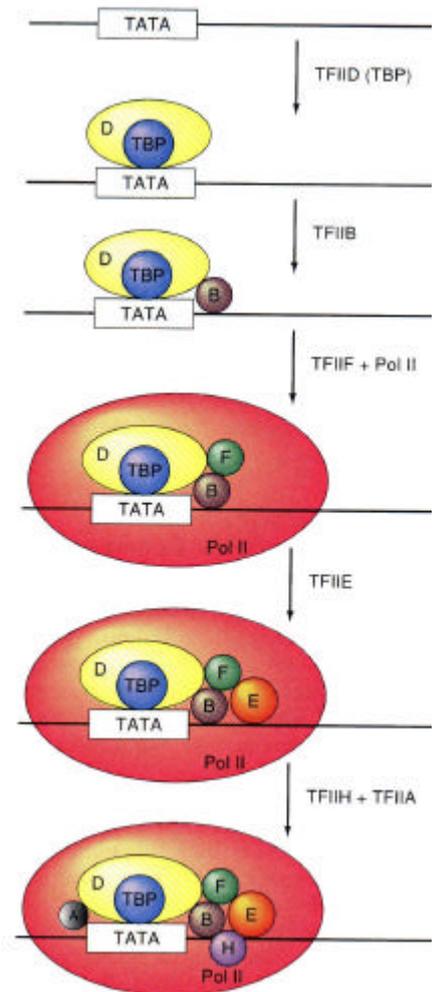


Figura 1.18 – Passos da iniciação da transcrição

gene, é habitualmente a decisão de transcrever ou não esse mesmo gene. O que se traduz possivelmente numa necessidade de economia de energia e materiais por parte da célula.

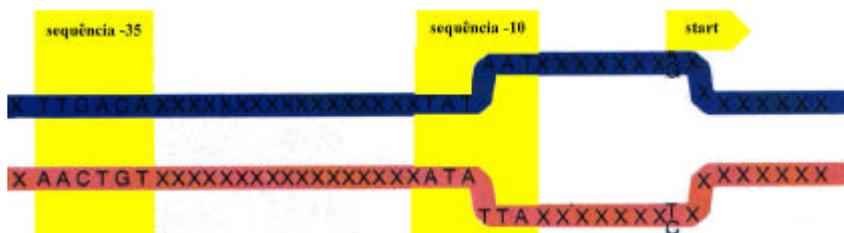


Figura 1.19 – Elementos reguladores da iniciação da transcrição eucariótica

A transcrição é catalisada pela RNA polimerase, e envolve a síntese de uma cadeia de RNA complementar da cadeia molde do DNA (a outra cadeia do DNA é a imagem do RNA, isto é a sua sequência é equivalente à do RNA, excepto no facto de em vez de possuir Uracilos possui Timidinas). A transcrição ocorre pelo processo habitual de emparelhamento de bases num processo altamente regulado e encadeado. Em primeiro lugar, a polimerase deve ligar-se ao DNA de cadeia dupla num processo complexo que envolve co-factores (fig. 1.18). Em seguida, as duas cadeias do DNA devem ser separadas (fig.1.19), para tornar a cadeia complementar acessível à maquinaria de transcrição. A abertura da hélice do DNA é um processo localizado, e à medida que a transcrição prossegue, novas zonas do DNA vão ficando acessíveis, enquanto as zonas já transcritas se vão emparelhando de novo, por forma a preservar a dupla hélice. A fase de **iniciação da transcrição** envolve assim, o reconhecimento do DNA pela polimerase, a abertura da hélice do DNA, e a incorporação do primeiro nucleótido na cadeia do RNA nascente. O local do gene onde se processa todo este processo é naturalmente o **promotor** (fig. 1.19). O

local da incorporação do primeiro nucleótido é designado “**start site**” ou “**startpoint**” (fig.1.19).

Depois da fase de iniciação inicia-se a **fase de alongação**, a qual produz a extensão da cadeia de RNA nascente, originando um híbrido de emparelhamento DNA-RNA. No entanto, à medida que a alongação se processa, a polimerase caminha para novas regiões do DNA, abrindo a hélice noutras zonas do gene, e fechando nas regiões já transcritas, o que implica o desemparelhamento DNA-RNA.

A **terminação** envolve o reconhecimento de um sinal indicando que não devem ser adicionados mais nucleótidos. Nesta fase, termina a ligação DNA-RNA da cadeia nascente, com libertação da polimerase e da molécula de RNA.

Desta descrição se pode inferir que a polimerase do RNA (a enzima que catalisa a adição de nucleótidos à cadeia de RNA nascente) não funciona só, necessitando de um conjunto de outros componentes com funções essencialmente reguladoras e assessórias. Assim, quer a iniciação quer a abertura do DNA, quer a terminação são exemplos de processos em que intervêm outros factores para a progressão organizada e controlada da expressão génica. A maquinaria de transcrição das células eucarióticas é mais complexa e menos bem definida que a dos procariotas. Existem 3 polimerases nucleares, as quais ocupam diferentes locais do núcleo, e são cada qual composta por várias subunidades. Para complicar ainda mais o problema, existem ainda outras polimerases do RNA em mitocondrias e cloroplastos.

A maior parte da actividade de polimerase do RNA é realizada, nos eucariotas, pela RNA polimerase I, a qual se encontra no nucléolo, e é responsável pela transcrição dos genes codificando os rRNA (cerca de 50-70% do RNA total sintetizado). A segunda enzima, é a RNA polimerase II (20-40% da actividade total de síntese de RNA), e é responsável pela síntese do RNA heterogéneo (hnRNA), o precursor

do mRNA. A RNA Polimerase III é responsável pela restante actividade de produção de RNA (até 10% do total), tem localização nucleoplasmática e é responsável pela produção dos tRNA e muitos dos “small nuclear RNA” (snRNA).

Nenhum mecanismo de controlo da transcrição pode ser uma forma eficaz de controlar a expressão génica, se o produto da transcrição (o mRNA) não tivesse uma vida curta. Se assim não fosse, previsivelmente ocorreria uma acumulação de mensageiro, ou pelo menos o mensageiro formado permaneceria activo tanto tempo que não seria possível parar de sintetizar a respectiva proteína. Na realidade, a instabilidade do mRNA é muito acentuada. As duas formas de determinar a instabilidade do DNA baseiam-se ambas no bloquear da síntese *de novo* do mRNA (transcrição), medindo então a sua capacidade para servir na síntese proteica (semi-vida funcional), ou a sua capacidade para hibridar com uma sonda (semi-vida química). De modo geral, a semi-vida funcional é ligeiramente inferior à semi-vida química, o que sugere que pequenas degradações como um simples corte poderão ser suficientes para a inactivação biológica do mRNA. Verifica-se que este primeiro passo inicial é seguido da degradação do mRNA nos seus nucleótidos componentes, de forma mais ou menos sequencial na direcção 5'→3'.

O dogma central da genética molecular afirma que os genes são unidades que se perpetuam a si próprios, e que funcionam através da sua expressão em proteínas, através de um intermediário de RNA. Note-se que o dogma, na sua versão original define um paradigma que considera que a informação genética é transmitida unidirecionalmente: DNA→RNA→Proteína.

Hoje em dia, sabemos que a restrição do dogma central não é absoluta. Efectivamente, a informação genética pode ser transmitida de forma diferente da acima prevista. Alguns vírus de RNA, utilizam o RNA para a propagação da sua informação genética. Se esta pode parecer

uma extensão relativamente pequena do dogma central, já a existência nos retrovírus (vírus de RNA de cadeia simples que utilizam o DNA de cadeia dupla como intermediária na sua replicação) de transcriptases reversas constitui uma grande mudança no paradigma da genética molecular. As transcriptases reversas são enzimas que catalisam a síntese de um DNA de cadeia simples a partir de uma cadeia de RNA. Esta cadeia de DNA pode então ser utilizada para sintetizar DNA de cadeia dupla, utilizando a maquinaria habitual da célula, efectivamente revertendo um dos passos acima indicado: RNA→DNA. Este facto tem implicações profundas não só na forma de pensar a genética, mas também na biologia da infecção viral, já que este DNA de cadeia dupla formado, e que é uma cópia do RNA viral, vai agora integrar-se no genoma da célula, fazendo com que a infecção se propague de forma mais ou menos inofensiva à progenia da célula infectada (a integração no genoma celular é uma parte normal do ciclo de vida do vírus sendo necessária à transcrição dos genes virais). Uma outra implicação deste mecanismo é a possibilidade de uma infecção de vírus deste tipo poder mediar a inserção de mRNA celular no genoma, como se de RNA viral se tratasse, originando duplicação génica, e/ou inserção de uma cópia do gene sob a acção de um promotor diferente, efectivamente alterando o programa genético da célula. Uma outra implicação da infecção por este tipo de vírus, foi já por nós abordada aquando da discussão da existência de enhancers, e constitui na possibilidade de colocar genes celulares sob a acção de enhancers virais, uma vez mais alterando o programa genético da célula infectada.

Os tipos de retrovírus de que existe mais informação disponível são os que originam as **partículas tipo C** em aves e mamíferos. Estes vírus contêm duas cópias de RNA em cada virião. Assim, quando uma célula é infectada por dois viriões diferentes, podem-se originar viriões heterozigóticos, o que pode ser importante na aquisição de sequências celulares por parte do vírus, já que mesmo que em contrapartida perca

algumas sequências do seu genoma, a restante cópia do RNA viral permite-lhe continuar a ser capaz de efectuar uma infecção eficaz.

1.1.2.2 Tradução

A síntese proteica efectua-se no citoplasma, envolvendo uma complexa maquinaria genética centrada no ribossoma (fig. 1.20). Esta maquinaria genética pode ser vista como migrando ao longo do mRNA, lendo-o e utilizando a informação nele contida para alinhar com precisão cada aminoacil-tRNA, promovendo a ligação peptídica entre este e a cadeia peptídica nascente.

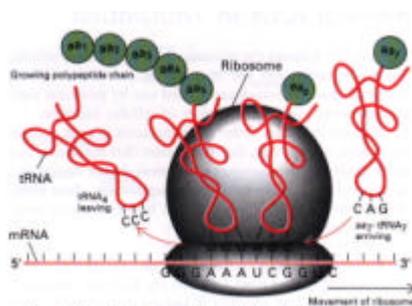


Figura 1.20 – O ribossoma interagindo com o mRNA, o tRNA e a cadeia polipeptídica nascente.

O ribossoma é assim um altamente elaborado e preciso complexo enzimático com diversificados componentes e vários centros activos, que requer vários cofactores para a sua actividade, e que obtém a energia química que necessita com a hidrólise de GTP.

Um ribossoma é composto por duas unidades (60S e 40S nos eucariotas) as quais, apesar de funcionarem em conjunto medeiam reacções diferentes na síntese proteica. O mRNA associa-se à subunidade menor, ficando associado a este por cerca de 30-40 nucleótidos. Apenas 2 moléculas de tRNA se podem associar ao ribossoma em cada momento, pelo que apenas 2 dos cerca de 30 codons associados ao ribossoma se encontram a ser processados em cada momento.

Cada tRNA liga-se ao ribossoma num local diferente deste, tendo cada um dos dois locais de ligação propriedades diferentes. Apenas o Local

A (local de entrada) pode receber um aminoacil-tRNA. Antes da entrada do aminoacil-tRNA, este local expõe o codon a ser decodificado. O último dos codons já decodificados encontra-se no local P (local dador), sendo este local ocupado pelo peptidil-tRNA (um tRNA contendo o aminoácido já covalentemente ligado por uma ligação peptídica à restante cadeia polipeptídica nascente). Quando estes locais (A e P) estão ambos ocupados ocorre a formação da ligação peptídica com transferência do polipéptido nascente para o tRNA do local A. O ribossoma desloca-se então no mRNA libertando o tRNA do local P e transferindo para este local o peptidil-tRNA do local A, e expondo um novo codon no local A.

A síntese proteica pode ser dividida em várias fases:

Iniciação: envolve as reacções que precedem a formação da ligação peptídica. Requer a ligação do ribossoma ao mRNA, a formação de um complexo de iniciação contendo o primeiro aminoacil-tRNA. Trata-se de um processo relativamente lento em comparação com as restantes fases da síntese proteica.

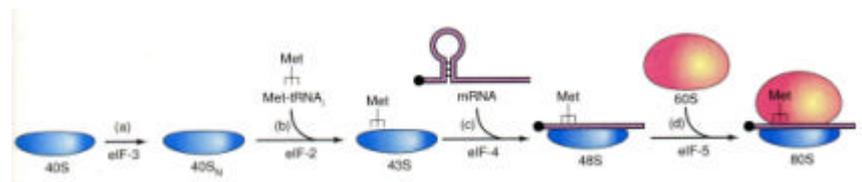


Figura 1.21 – Passos da fase de iniciação e respectivos co-factores envolvidos

Resumidamente pode dizer-se que nos eucariotas a iniciação começa com a ligação de GTP a um factor de iniciação denominado eIF-2 (eucariotic initiation factor 2). De seguida efectua-se a ligação de um N-formil-metionil-tRNA a este complexo. É o conjunto de factores assim formado que se liga então à subunidade 40S do ribossoma, a

qual com o auxílio de outros factores de iniciação reconhece então a extremidade 5' do mRNA (na qual se encontra a estrutura conhecida como CAP) por parte da subunidade 40S do ribossoma.

A subunidade 40S migra então no mRNA até encontrar um codon de iniciação. Neste ponto, liga-se a subunidade 60S, após a remoção de

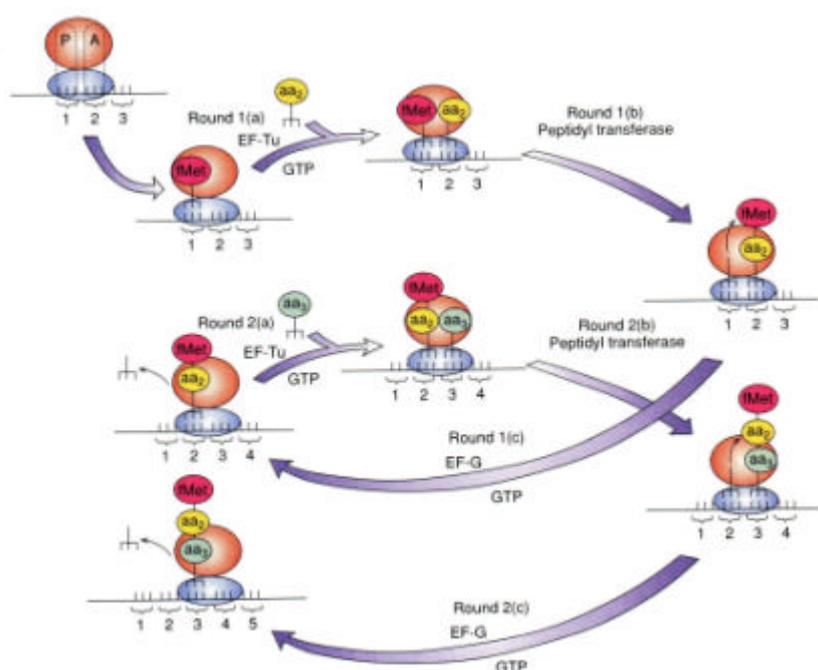


Figura 1.22 – Passos da fase de alongamento da tradução eucariótica

eIF-2 do complexo de iniciação.

Elongação: inclui todas as reacções desde a síntese da primeira ligação peptídica, até à adição do último aminoácido da cadeia polipeptídica.

Os aminoácidos são adicionados um a um, naquele que constitui o processo mais rápido da síntese proteica.

Assim que a subunidade 60S se liga ao complexo de iniciação, o ribossoma fica pronto a iniciar a elongação. Para tal necessita de aminoacil-tRNA, o qual entra o local A, num processo mediado pelo factor eEF-1 (eucariotic elongation factor 1). Assim que o aminoacil-tRNA se encontra correctamente posicionado no local A, a peptidil transferase (uma função da subunidade 60S) catalisa a formação da ligação peptídica entre os aminoácidos dos locais P e A.

O último passo na elongação é a translocação, processo

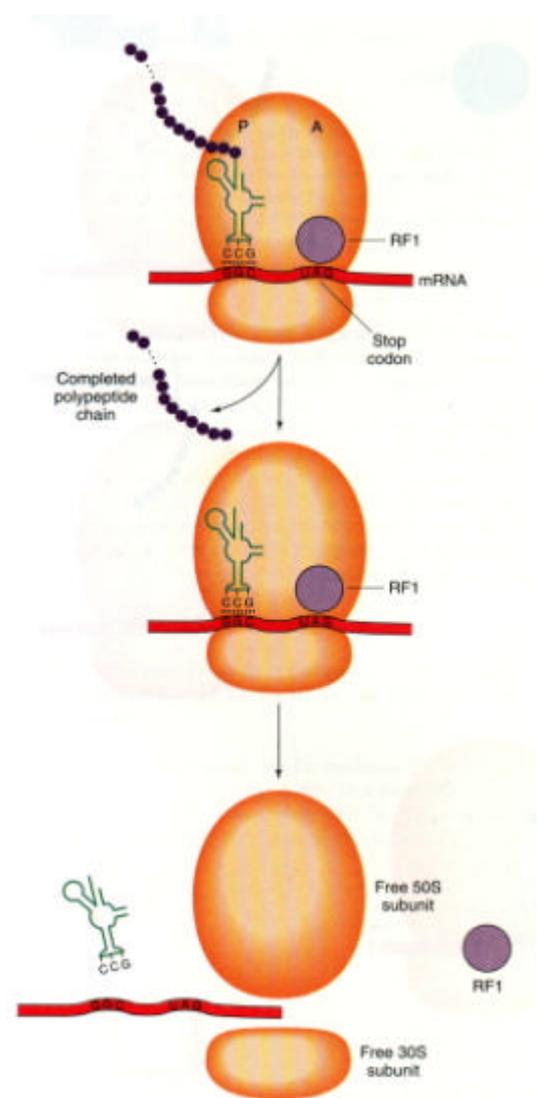


Figura 1.23 – Passos da fase de terminação da tradução eucariótica

em que o ribossoma avança três nucleótidos de forma concertada(e que requer o factor adicional eEF-2). Este processo envolve a libertação do tRNA do local P, a passagem do peptidil-tRNA do local A para o local P, e a exposição do próximo codon no local A agora vazio.

Terminação: inclui todos os passos necessários para a libertação da cadeia polipeptídica formada, bem como a dissociação do ribossoma do mRNA. Este é um processo lento, em comparação com o tempo necessário para adicionar um aminoácido na fase de elongação.

Dos 64 tripletos, apenas 61 codificam para aminoácidos, sendo os restantes três codons stop, ou de terminação. Qualquer destes três codons (UAG, UAA e UGA) é suficiente para terminar a síntese proteica. Aos codons de terminação não corresponde nenhum tRNA, sendo estes reconhecidos directamente pelo factor proteico eRF (eucariotic release factor).

A reacção de terminação envolve a libertação do polipéptido do último tRNA, a expulsão do tRNA do ribossoma, e a dissociação deste do mRNA.

A célula eucariótica é uma estrutura finamente organizada, cujas funções são efectuadas em locais celulares definidos. A síntese proteica não constitui excepção, podendo os polirribossomas ser classificados em 2 tipos (livres e ligados a membranas), aos quais corresponde a síntese de diferentes grupos de proteínas .

Os polirribossomas livres sintetizam proteínas que não interagem com membranas, enquanto os que se encontram associados às membranas sintetizam proteínas cuja futura localização depende da sua capacidade para se ligarem às membranas. Note-se no entanto que a denominação polirribossomas livres não significa que estes se encontrem livres em solução no citoplasma. Estes polirribossomas encontram-se associados ao citoesqueleto para o que provavelmente dependem do mRNA.

Os polirribossomas tendem a estar localizados perto de núcleos, nos locais de entrada do mRNA no citoplasma. A maior parte das proteínas sintetizadas são solúveis, e uma vez libertadas rapidamente difundem para longe do local de síntese. As proteínas que irão compor o citoesqueleto, tendem a integrar-se neste num local não muito distante do ponto de síntese.

As proteínas sintetizadas pelos ribossomas ligados a membranas têm vários destinos. Algumas são sequestradas em compartimentos celulares, outras são componentes membranares, e outras ainda são proteínas que se destinam a ser secretadas. Na maior parte dos casos das proteínas de membrana, a sua futura localização não depende da sequência da proteína madura, mas antes de uma sequência denominada “leader”, e que se localiza na zona terminal da cadeia polipeptídica nascente.

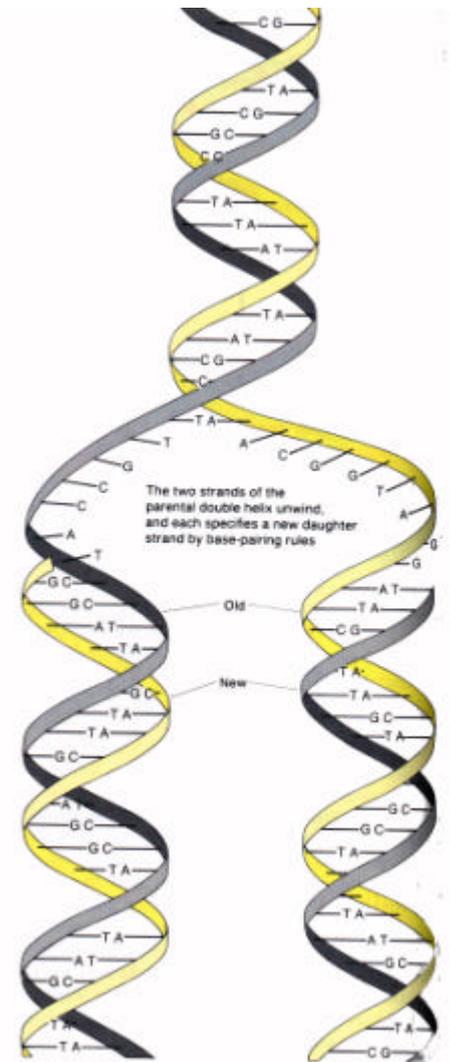


Figura 1.24 – A replicação é semiconservativa, tendo como base o emparelhamento de nucleótidos

Esta sequência, depois de ter determinado o destino da proteína será excisada do resto da proteína, originando a proteína madura.

1.1.2.3 Replicação

A descoberta da natureza de dupla hélice do DNA foi o ponto de partida para a compreensão do mecanismo de duplicação deste ácido nucleico, o qual é designado por replicação. Com efeito, o princípio básico do emparelhamento dos nucleótidos, em que assenta a estrutura do DNA é também o princípio essencial para a compreensão da replicação do DNA, já que esta assenta na noção de que cada cadeia de nucleótidos funciona

como padrão na síntese de uma nova cadeia (fig. 1.24). Assim, a replicação do DNA é semiconservativa, isto é, após a duplicação, cada molécula de DNA tem uma cadeia original, e uma cadeia sintetizada de novo (fig. 1.24).

Como a síntese se processa sempre no mesmo sentido ($5' \rightarrow 3'$) e as cadeias de DNA têm orientação invertida, numa das cadeias, a síntese é contínua, sendo na cadeia complementar descontínua, a partir de vários pontos de iniciação. Forma-se assim, em cada ponto de replicação uma estrutura em forma de Y ("Forks"). Como a replicação é bidireccional (excepto nos vírus de DNA linear), cada ponto de

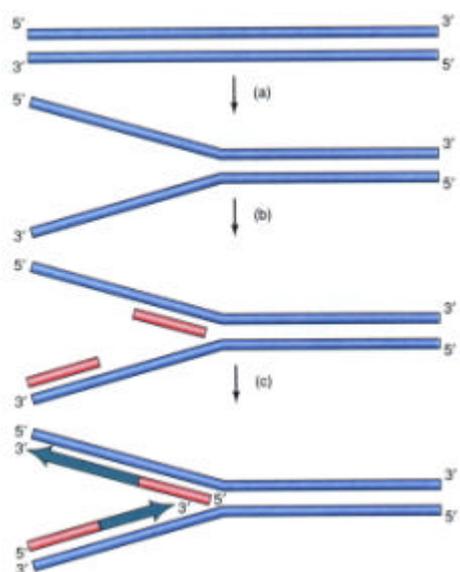


Figura 1.25 – A replicação é contínua numa cadeia (cadeia "leading") e descontínua na outra (Cadeia "lagging").

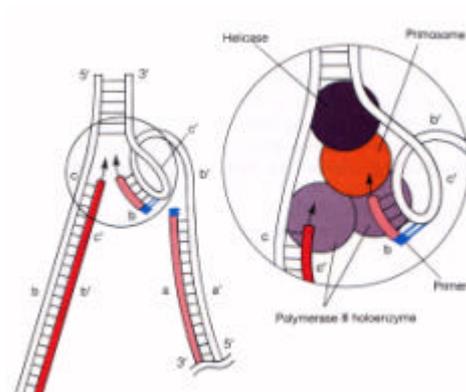


Figura 1.26 – Estrutura em loop que permite que uma polimerase sintetize as duas cadeias de DNA.

iniciação da replicação dá origem a dois “Forks” invertidos (—C D—). Para que a mesma polimerase possa sintetizar ambas as cadeias, forma-se uma estrutura local em ‘loop’ na cadeia de síntese descontínua (fig. 1.26).

Em bactérias de DNA circular, habitualmente apenas um ponto de iniciação é suficiente para replicar todo o genoma,

mas em eucariotas, cada cromossoma possui múltiplos pontos de iniciação, que cooperam para a rápida replicação do DNA. Estes pontos possuem uma sequência semelhante entre si. Conhecem-se 3 tipos principais destas sequências: a OriC da *E.coli* a sequência do vírus SV40, e a sequência de replicação autônoma em leveduras. A OriC (bem como a maior parte das sequências equivalentes em bactérias) é composta por repetições de sequências com 9 e 13 pares de bases (9-mers e 13-mers) ao longo de uma sequência com cerca de 240bp, sendo estes 9-mers e 13-mers locais de ligação da proteína DnaA, a qual inicia a replicação. A origem de replicação do SV40 possui cerca de 65bp. Os 17 cromossomas da levedura *S.cerevisiae* possuem mais de 400 sequências de iniciação da replicação. Cada uma destas sequências (denominadas ARS de *autonomously replicating sequence*) confere a um plasmídeo a capacidade de se dividir na levedura. Estudos mutagénicos efectuados ao ARS1 (cerca de 180bp) revelaram a existência de apenas uma região essencial, com cerca de 15bp, designada elemento A. Três outros elementos (B₁, B₂ e B₃) são

necessários para o funcionamento efectivo do ARS11. Um complexo de proteínas chamado ORC (“*origin recognition complex*”) liga-se ao ARS1 de modo dependente do ATP. Adicionalmente a estas sequências, uma sequência a elas adjacente, rica em Adenosinas e Timidinas facilita a abertura da dupla hélice.

A replicação eucariótica (aqui descrita para o SV40 – o modelo mais

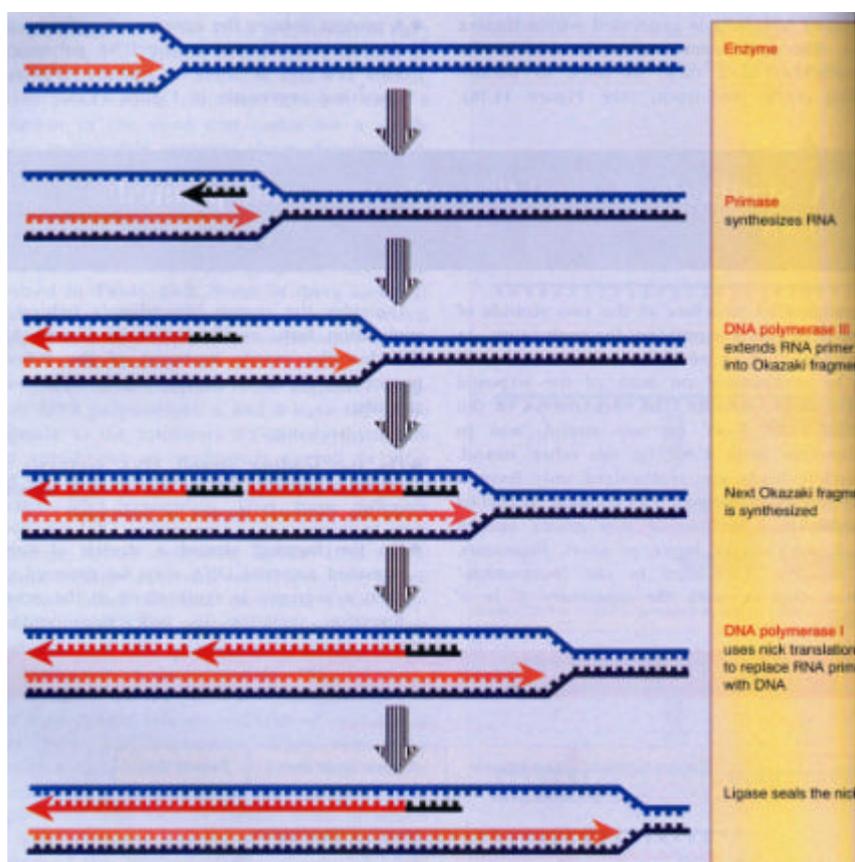


Figura 1.27 – A síntese descontínua da cadeia "lagging" exige primers de RNA sintetizados pela primase, formando-se temporariamente uma molécula mista RNA/DNA (fragmento de Okazaki).

bem conhecido) inicia-se com a ligação de uma proteína viral chamada T *antigen* ao DNA. Esta proteína multifuncional desenrola o DNA com a sua actividade de helicase, com a ajuda de ATP e um factor denominado RFA (“*replicating factor A*”), o qual é uma proteína que se liga a DNA de cadeia simples. A replicação inicia-se então bidirecionalmente a partir de primers de RNA produzidos por uma primase.

Duas polimerases distintas (α e δ) trabalham no “*fork*” eucariótico: a pol δ produz a cadeia “*leading*”, cuja síntese é contínua; a síntese da cadeia “*lagging*”, realizada de modo descontínuo é da responsabilidade da pol α , a qual se encontra fortemente associada a uma primase. Na realidade, a síntese da cadeia “*leading*” é na realidade o trabalho sequencial das polimerases α e δ . O complexo primase-pol α sintetiza

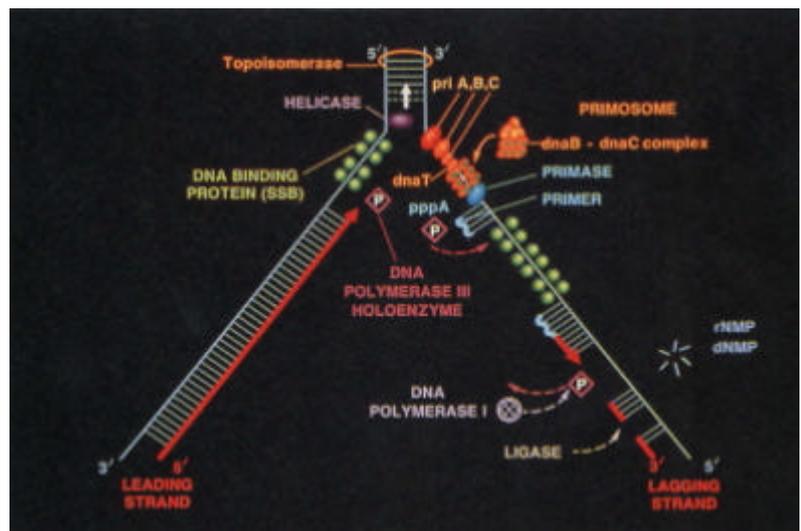


Figura 1.28 – Representação esquemática dos vários componentes que intervêm na replicação eucariótica.

o primer de RNA, e inicia a extensão do DNA, numa actividade

estimulada pelo factor de replicação C. A ligação de PCNA (*“proliferating cell nuclear antigen”*) liberta a pol α . A pol δ liga-se então ao PCNA que activa a polimerase a qual sintetiza então a cadeia *“leading”*. Finalmente enzimas chamadas topoisomerases executam o importante trabalho de libertar as tensões de torção introduzidas pelo desenrolar da hélice do DNA padrão.

No final de cadeias de DNA linear, existe o risco de em cada síntese, a cadeia *“lagging”* se tornar progressivamente mais curta, pois a síntese descontínua necessita sempre de um *“template”* maior que a sequência a amplificar. Para obviar a este risco, a célula possui nas extremidades do cromossoma estruturas especiais denominadas telómeros e compostas por repetições de sequências oligoméricas (ex. na levedura é: 5'-G₁₋₃T-3'). Os telómeros são sintetizados por um processo diferente do acima descrito por enzimas especializadas: as telomerases. Estas enzimas são transcriptases reversas modificadas, produzindo DNA a partir de um *“template”* de RNA que a elas se encontra associado, e que determina portanto a sequência do telómero.

Finalmente um outro aspecto da síntese do DNA deve ser alvo da nossa atenção: a reparação de erros. As polimerases, não são enzimas isentas de erros, e a incorporação de nucleótidos incorrectos por vezes ocorre. Vários mecanismos de reparação destes erros existem, e deles depende a transmissão do património genético, sem mutações. O primeiro mecanismo de reparação do DNA está incorporado nas próprias polimerases. Estas enzimas, não só sintetizam DNA (actividade polimerase 5'→3') como exercem uma actividade de verificação (*“proofreading”*) das bases recentemente adicionadas, removendo as bases incorrectas (actividade de exonuclease (3'→5')).

1.1.2.4 Mecanismos de Mutação do DNA

1.1.2.4.1 Erros na replicação do DNA

Quando, durante a replicação do DNA a polimerase incorpora nucleótidos errados, produz-se uma mutação. Na maior parte dos casos, tal não se verifica, apenas porque a maioria das polimerases, possuem não apenas actividade de síntese na direcção $5' \rightarrow 3'$, mas também actividade de *proofreading* na direcção $3' \rightarrow 5'$. Assim, se o nucleótido incorporado for o errado, a polimerase usa a sua actividade exonucleolítica, e volta a tentar.

Cada base do DNA existe em uma de várias formas possíveis, chamadas tautómeros. Trata-se de isómeros, em que a posição relativa e as ligações entre os átomos variam. Normalmente apenas a forma ceto aparece no DNA, enquanto as formas enol e imino são raras. No entanto, os tautómeros enol e imino têm a capacidade de emparelhar com bases diferentes das que a forma ceto emparelha, levando a erros na replicação. Assim, por exemplo a forma imino da citosina pode emparelhar com a adenina, a forma enol da timina pode emparelhar com a guanina, a forma imino da adenina pode emparelhar com a citosina e a forma enol da guanina pode emparelhar com a timina.

Uma outra forma de incorporação estável de bases erradas ocorre com a replicação de bases ionizadas, a qual é possivelmente mais frequente do que a mediada pelos tautómeros, e justifica os efeitos mutagêneos das radiações ionizantes.

Os emparelhamentos de purinas com purinas e pirimidinas com pirimidinas, ocasionam distorções na dupla hélice, já que forçam alterações na dimensão exterior da hélice. No entanto, mesmo este tipo de mutações (transversões) ocorrem, tendo inclusivamente sido observado por difracção de raios X estruturas de DNA albergando este tipo de emparelhamentos.

Também mutações frameshift podem originar-se por erros de replicação. este tipo de erros ocorre preferencialmente em zonas de repetição de nucleótidos (ex: AAAAAA → AAAAA; CTCTCT→CTCT), e pensa-se ser o resultado de um “deslizar” excessivo da polimerase, sem que este provoque uma desestabilização do emparelhamento imediatamente adjacente ao local de polimerização. Com o mesmo mecanismo se pode explicar mutações frameshift causadas por inserções, bastando para tal admitir que a cadeia que “desliza” é a cadeia nascente.

1.1.2.4.2 *Lesões espontâneas*

As duas mais frequentes lesões espontâneas causando mutações são a despurinação e desaminação. A mais comum é a **despurinação**, a qual envolve a interrupção da ligação glicosídica entre a base e a desoxirribose, com a conseqüente perda de uma base de guanina ou adenina da molécula de DNA. Uma célula de mamífero perde espontaneamente cerca de 10^4 purinas do seu DNA num período de 20 horas a 37°C. Assim se compreende a grande necessidade de mecanismos de reparação do DNA, para a manutenção do código genético, garantindo a viabilidade celular e a da progenia. Em certas condições, uma base pode ser adicionada ao nucleótido apurínico (como um último recurso de reparação), existindo neste processo 3 hipóteses em 4 de o resultado se traduzir numa mutação.

A **desaminação** de uma citosina origina um uracilo. Se a base assim alterada não for reparada, numa replicação posterior ela vai emparelhar com uma adenina, originando a substituição de um C por um G. Note-se que se o nucleótido onde ocorreu a desaminação possuir uma 5-metil-citosina, a base desaminada é uma 5-metil-uracilo, ou seja timina, a qual não é reconhecida pela maquinaria de recombinação, resultando numa mutação não reparável. Por este motivo as 5-metil-uracilo são “*hot spots*” de recombinação.

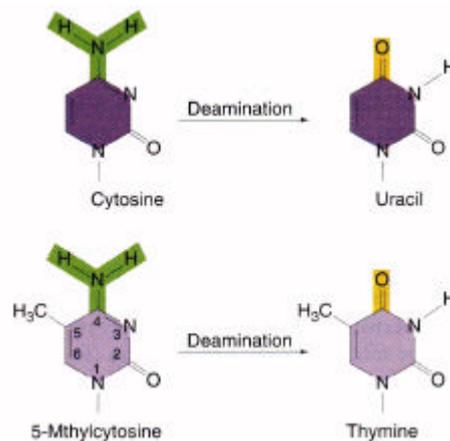


Figura 1.29 – Desaminação de bases nucleotídicas.

Uma forma menos frequente de mutação espontânea são os danos oxidativos. Espécies de oxigênio altamente reativas como os radicais superóxido (O_2^\bullet), peróxido de hidrogênio (H_2O_2) e o seu derivado radical hidróxilo (OH^\bullet) são produtos secundários do metabolismo aeróbico, e podem causar danos oxidativos no DNA ou seus precursores. Por exemplo a 7-hydrodeoxyguanosina ou 8-oxodG (GO) frequentemente emparelha com adenina, resultando num elevado número de mutações $G \rightarrow T$.

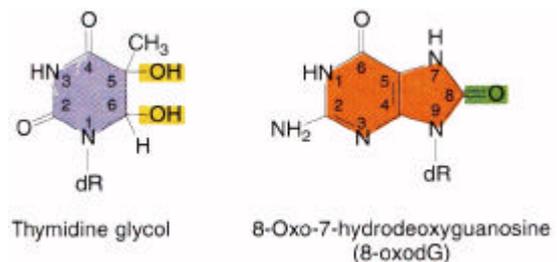


Figura 1.30 – Formação do 8-oxodG (GO)

1.1.2.4.3 **Recombinação genética**

1.1.2.4.3.1 Recombinação por crossing-over

O crossing-over é um mecanismo normal de troca de material genético entre dois cromossomas homólogos, durante a meiose. Trata-se de um fenómeno complexo em que habitualmente é conseguida uma extraordinária precisão, mas de onde podem surgir mutações quando algo de errado acontece. Existem hoje vários modelos de recombinação por crossing-over:

Recombinação homóloga Geral

A recombinação homóloga geral não envolve nenhuma especificidade em termos de sequência do DNA, para além de se verificar apenas entre moléculas de DNA homólogo. Esta recombinação envolve a formação de uma estrutura em que uma molécula de DNA (dupla hélice) emparelha parcialmente com outras homóloga, formando uma

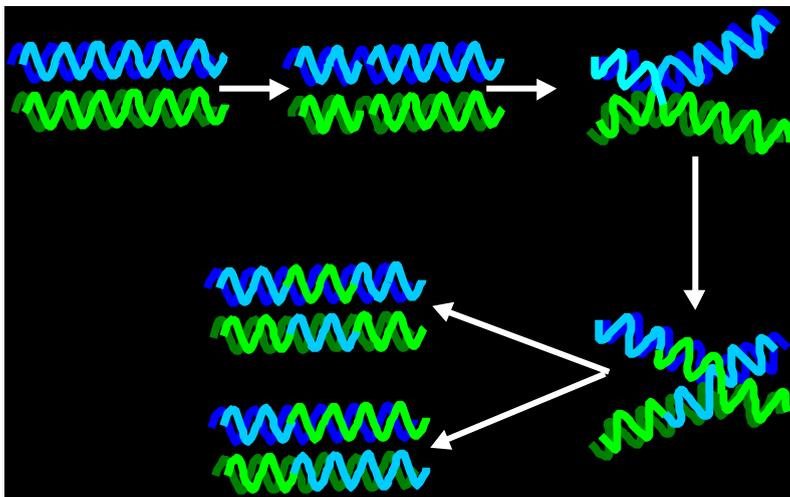


Figura 1.31 - Modelo de Holiday para a recombinação homóloga

estrutura em quiasma. Note-se que de cada cromossoma, apenas um dos dois cromatídeos (ou uma das duas moléculas de DNA de cadeia dupla) toma parte na recombinação. O local de corte e junção ocorre então no centro do quiasma, resultando dois cromossomas híbridos, mas intactos em termos do seu conteúdo genético.

O modelo mais plausível para explicar os passos necessários para a recombinação foi apresentado por Robin Holliday (fig.1.31) para explicar o fenómeno de conversão génica. Como se pode ver na fig. 1.31, o modelo de Holliday prevê que num dos passos iniciais se verifique um corte por uma endonuclease numa das cadeias de cada molécula de DNA, a que se segue a troca de extremidades de uma das cadeias de cada dupla hélice. As cadeias são então ligadas pela ligase formando-se cadeias híbridas. A este passo segue-se a migração do ponto de troca, terminando o processo com a resolução, a qual pode ocorrer por um de dois mecanismos (Fig. 1.31): 1) cortando e ligando mantendo as cadeias originais no resto do DNA ou; 2) cortando e ligando as cadeias novas, efectivamente trocando todo o material cromossómico deste o ponto de cross-over até ao telómero.

Este modelo, se bem que explique os resultados obtidos com fungos, não permite explicar todas as observações efectuadas com leveduras, o que levou Meselson & Radding a propor um mecanismo diferente. Este mecanismo, parte não de um corte de uma cadeia em cada molécula de DNA, mas apenas numa delas, seguindo-se a polimerização a partir de um dos terminais. Esta polimerização provoca o desemparelhamento de parte da cadeia em síntese, originando uma molécula de DNA parcialmente de cadeia simples, a qual pode então substituir a sua homóloga no outro cromossoma. Os últimos passos neste modelo consistem na ligação das cadeias com terminais, seguida da migração do ponto de junção, e resolução horizontal ou vertical. Note-se que este modelo explica a formação de produtos não passíveis de serem produzidos pelo modelo de Holliday, como é o caso de cromossomas irmãos que numa divisão possuem

trocas cromossômicas desiguais entre si. Um outro tipo de implicação deste modelo é a possibilidade de ocorrência de conversão de cromátídeos e de conversão de meios-cromátídeos. Estes fenômenos ocorrem quando se formam cromátídeos com cadeias de DNA não homólogas, isto é quando, por crossing-over, o produto final é uma molécula de DNA com cadeias pontualmente não emparelhadas (locais G:T, G:A, C:T, C:A). Quando isto sucede, os mecanismos de correção do DNA podem corrigir o mal-emparelhamento, do que resulta uma molécula de DNA convertida (conversão de cromátídeos), ou igual à molécula inicial. Pode ainda não ter ocorrido correção, e a molécula de DNA possuir o desemparelhamento local (conversão de meio-cromátídeo). Assim se obtêm as proporções menos frequentes, e não iguais no número de cromossomas resultantes de um crossing-over.

1.1.2.4.3.2 Recombinação somática

Trata-se de um processo enzimático complexo e altamente regulado, que permite a uma célula T ou B adquirir um gene específico para o TCR ou para uma imunoglobulina respectivamente.

O mecanismo de recombinação (ou rearranjo) somática (descrito em pormenor na secção 2.1) é habitualmente restrito às células T e B, e apenas aos genes do TCR e Ig. No entanto, a existência de sequências de recombinação ditas crípticas (sequências de recombinação imperfeitas) pode originar a que em casos raros, nestas células, os genes do TCR ou Ig sejam reorganizadas indevidamente com outros loci genéticos, do que resulta uma translocação intra ou intercromossômica com perturbação do programa genético. O programa genético da célula em que ocorrem estas anomalias pode ser perturbado de uma de duas formas: 1) a formação de uma proteína quimera pode desregular a função dos genes em questão; 2) a colocação de um gene habitualmente silencioso nos linfócitos debaixo da acção do enhancer do TCR ou das Ig, faz com que este seja expresso de forma

permanente. Exemplos de anomalias destes tipos, e suas consequências clínicas serão apresentadas na secção 4.2.

1.1.2.4.4 **Mutações mediadas por elementos de transposição**

Os elementos de transposição são elementos genéticos com a propriedade de serem capazes de se mobilizarem, e moverem-se para outro local do genoma, ou de criarem cópias de si próprios noutro local do genoma. Este conceito, criado por resultados experimentais produzidos por Marcus

Rhoades e Barbara McClintock em milho, foram inicialmente encarados com grande cepticismo. Sabe-se no entanto hoje que estes elementos são bastante representados em todo o mundo biológico, dos Procariotas aos Eucariotas superiores.

As sequências de inserção (fig.1.32) são os elementos de transposição mais simples, sendo constituídos por sequências de nucleótidos que se podem agrupar em famílias, de acordo com a sua sequência. Assim, a IS1 foi historicamente a primeira a ser descrita, possuindo cerca de 800bp, mas muitas outras foram entretanto descritas. Uma das características das IS é a presença de uma pequena sequência que aparece invertida em cada uma das extremidades (IR do inglês *inverted repeats*). Os *inverted repeats* possuem entre 16 nucleótidos e algumas dezenas, sendo o seu comprimento e sequência específico de cada IS.

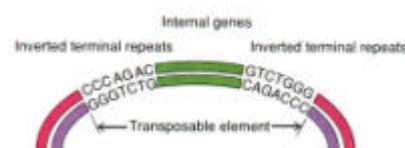


Figura 1.32 – Elemento de transposição simples ou sequência de inserção (IS do inglês *insertion sequences*).

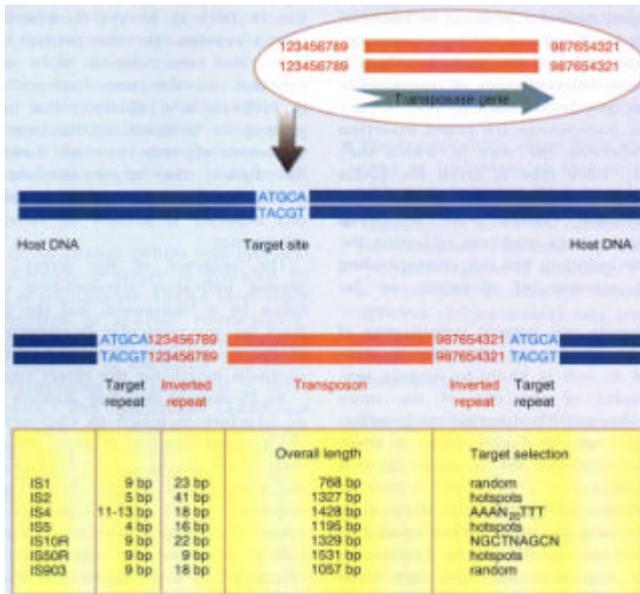


Figura 1.33 – características de alguns transposons e consequências da sua inserção

Um transposon (Tn1, Tn2, etc.) é um elemento de transposição mais complexo, contendo dois IS separados por uma sequência de DNA que pode conter um ou mais genes. Os dois IS nas extremidades são idênticos, e podem estar na mesma orientação, ou em orientações invertidas. Se estiverem na mesma orientação os IR presentes são os IR de cada um dos IS. Se estiverem em orientações invertidas, todo o IS funciona como um IR. Os transposons herdam dos IS que os constituem a capacidade de se deslocarem no genoma, mas com uma habilidade extra que é a de transportarem consigo o(s) gene(s) que os constituem. Na verdade, em bactérias, este é o mecanismo que muitas vezes se encontra por detrás da aquisição de resistência a antibióticos num único plasmídeo.

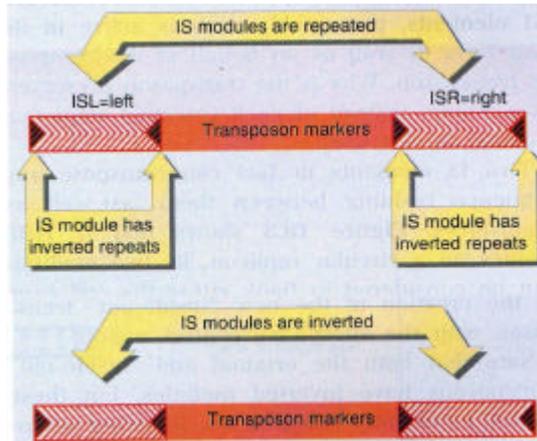


Figura 1.34 – Os transposons são elementos de transposição mais complexos. São constituídos por uma sequência central contendo genes ou marcadores, e sequências laterais que podem ser repetições directas ou invertidas de IS.

Os mecanismos de transposição podem ser conservativos ou replicativos. No primeiro ocorre a deslocação do elemento de transposição de um local para outro, sem que o numero de elementos tenha variado. No segundo caso, é deixada uma cópia do elemento de transposição no local original, isto é o elemento de transposição é replicado para um novo local.

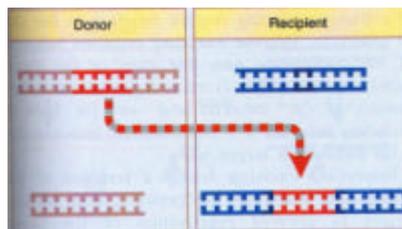


Figura 1.35 – Mecanismo de transposição conservativa

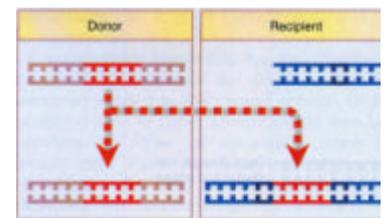


Figura 1.36 – Mecanismo de transposição replicativa

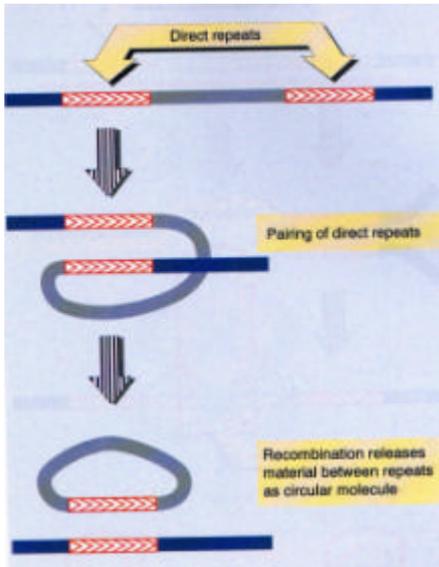


Figura 1.37 – Os transposons podem ser excisados por recombinação homóloga entre os direct repeats, deixando uma cópia dos repeats no local de inserção

O mecanismo de transposição replicativa é catalizado por uma transposase, e envolve um intermediário que possui todo o genoma do plasmídeo dador integrado no genoma do plasmídeo receptor, possuindo este intermediário duas cópias do elemento de transposição (uma em cada extremidade do genoma dador). Em seguida, ocorre como que um cross-over (mediado pela sequência do elemento de transposição), que é depois resolvida pela transposase (a qual pode ou não ser codificada nos elementos de transposição, pelo que a sua acção é efectuada em trans ou cis

respectivamente). Uma outra consequência da transposição é a ocorrência de uma duplicação do DNA receptor em cada extremidade da inserção. Esta duplicação resulta do facto de o corte em cada cadeia do DNA receptor ser efectuada em locais diferentes, e de a sequência de DNA de cadeia simples ser depois resolvida pela polimerase.

A presença de elementos de transposição origina uma alta frequência de deleções na sua vizinhança. Estas deleções ocorrem como eventos de transposição aberrantes e podem incluir ou não parte do elemento de transposição. A própria mobilidade dos elementos de transposição pode ser causa de mutações, já que quando um IS se posiciona no meio de um gene, produz uma interrupção deste. Se altera a sequência codificante, ou a separa do promotor, o resultado pode ser um gene não funcional, com funcionalidade alterada, ou com baixa ou nula expressão.

Os elementos de transposição acima descritos, são relativamente simples e pequenos, existindo nos procariontes. Nos Eucariotes foram também descritos elementos de transposição como os *Ty* (levedura), os *Copia like elements* (drosofila), os *FB elements* (drosofila), os *P elements* (drosofila), e os retrovírus.

Os *Ty elements* possuem uma sequência

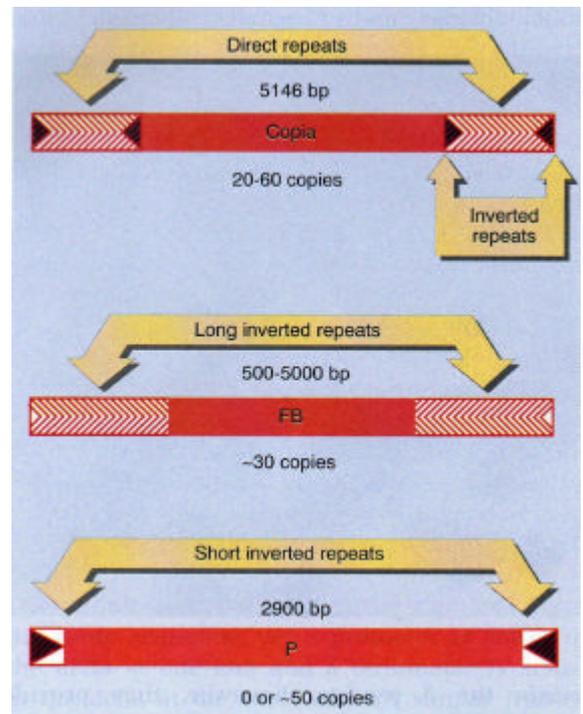


Figura 1.38 – Elementos de transposição dos eucariotas e suas características

altamente variável, mas terminada por uma sequência repetida (a sequência δ), a qual é também variável de elemento para elemento. Os Ty podem ter vários Kb de comprimento, e as sequências δ cerca de 300bp. Uma particularidade das sequências δ é que não se constituem em repetições invertidas, mas repetições directas.

Cerca de 10% do genoma da drosófila é composto por famílias de elementos de transposição disseminados, e que se movem como elementos discretos no genoma. Na drosófila foram caracterizados 3 tipos de elementos de transposição distintos: (*Copia like*, *Fold Back* ou *FB* e os *P elements*).

Os *Copia like* são compostos por 7 famílias diferentes, variando em tamanho de 5 a 8.5Kb. Membros de cada família são representados com 10-100 elementos no genoma da drosófila. Cada elemento possui uma repetição directa longa, e uma repetição invertida imperfeita e curta, sendo estruturalmente semelhante ao Ty da levedura.

Os elementos FB (*Fold Back*) variam em tamanho de algumas centenas de pares de bases a alguns Kb. Possuem entre si sequências semelhantes mas não iguais. Possuem terminais longos em repetições invertidas. Por vezes todo o elemento é constituído pelas repetições invertidas, com excepção de uma pequena sequência entre estas. O seu nome deriva do facto de em virtude de possuírem estas repetições tão longas estes elementos poderem hibridizar consigo próprios, como que dobrando-se sobre si mesmos. As propriedades observadas em elementos FB indicam que estes podem excisar-se a si próprios do genoma, provocando rearranjos cromossómicos com elevada frequência.

Os elementos P variam em tamanho de 0.5 a 2.9 Kb (os elementos mais pequenos derivam de deleções parciais num elemento P completo), tendo sempre uma sequência repetida invertida perfeita com 31 bp nas suas extremidades. Na região central dos elementos P

podem existir até 3 *open reading frames*, sugerindo que os elementos completos codificam para três proteínas.

Uma propriedade surpreendente dos elementos P é que se uma fêmea sem elementos P for cruzada com um macho com P+, a progenia apresenta mutagénese por inserção de elementos P (isto é os elementos P apresentam-se activos). Em contraponto, se a fêmea for P+ e o macho P-, a progenia não apresenta mutagénese por inserção os elementos P.

O modelo correntemente aceite para explicar as propriedades dos elementos P baseia-se na existência no interior do elemento de uma transposase e de produtos repressores dos elementos P. Se a fêmea possuir elementos P, então o óvulo possui proteínas repressoras dos elementos P, pelo que não ocorre transposição. Se no entanto a fêmea não possuir estes elementos, os elementos P do macho vão estar activos, inserindo-se no genoma, do que resulta mutagénese por inactivação de genes.

Os provírus (DNA de cadeia dupla inserida no genoma do hospedeiro) dos retrovírus podem ser considerados elementos de transposição já que podem efectivamente transportar-se de um lugar para outro no genoma da célula alvo. Com efeito, o genoma dos retrovírus tem algumas semelhanças com os elementos de transposição habituais. Em particular possuem um LTR (*long terminal repeat*) em cada extremidade semelhantes aos observados nos elementos Ty1 da drosófila.

A inserção dos retrovírus, tal como a dos elementos de transposição resulta na duplicação da zona de DNA onde se inserem. Uma outra forma de descrever esta similaridade entre os retrovírus e os elementos de transposição como os Ty1 e os *copia-like* é a de dizer que estes apresentam sequências que indicam serem reminiscências de retrovírus. Esta forma de ver o problema recebeu suporte experimental de experiências efectuadas por Jef Boeke and Gerald Fink, demonstrando a existência de um intermediário de RNA na inserção destas sequências. Os elementos de transposição que utilizam este

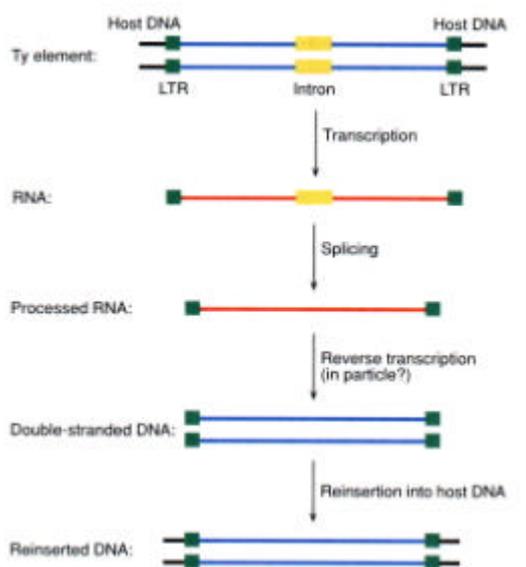


Figura 1.39 – Os retrotransposons e os retrovírus possuem semelhanças na forma de se integrar no genoma celular.

intermediário de RNA (ex. Ty1 e *copia-like*) são denominados retrotransposons, e são frequentes entre os eucariotas, sendo geralmente divididos em duas classes: Os retrotransposons virais (como os Ty1 e os *copia-like*) e os retrotransposons não virais (os mais

frequentes entre os mamíferos). Exemplos de retrotransposons não virais entre os mamíferos são os LINES (Long interspersed elements; 1-5Kb) e os SINES (short interspersed elements; ex. as sequências alu).

1.1.2.5 Mecanismos de reparação do DNA

As células eucarióticas desenvolveram uma série de mecanismos integrados de reparação de danos no DNA, os quais podem ser essencialmente de três tipos: mecanismos excisativos, não excisativos e mecanismos pós replicativos.

1.1.2.5.1 Mecanismos não excisativos

- **Evitar erros antes de surgirem:** alguns sistemas enzimáticos neutralizam compostos potencialmente perigosos antes de poderem reagir com o DNA. Um bom exemplo é o mecanismo de eliminação de radicais superóxido, catalizado pela superóxido dismutase e pela catalase ($2 \text{HO}\bullet \rightarrow \text{H}_2\text{O}_2$; $2 \text{H}_2\text{O}_2 \rightarrow \text{H}_2\text{O} + \text{O}_2$). Um outro exemplo é o do sistema catalizado pela proteína do gene *mut*, o qual previne a incorporação do 8-oxodT no DNA hidrolizando o trifosfato de 8-oxodT em monofosfato.
- **Reversão directa do dano:** A forma mais directa de reverter uma lesão é efectuar o seu processo inverso, restaurando assim a sequência alterada. No entanto, a reversão directa nem sempre é



Figura 1.40 – Esquema de acção da fotoliase que por acção da luz reverte a formação dos anéis de ciclobutano pirimidínico formados por acção da luz ultravioleta.

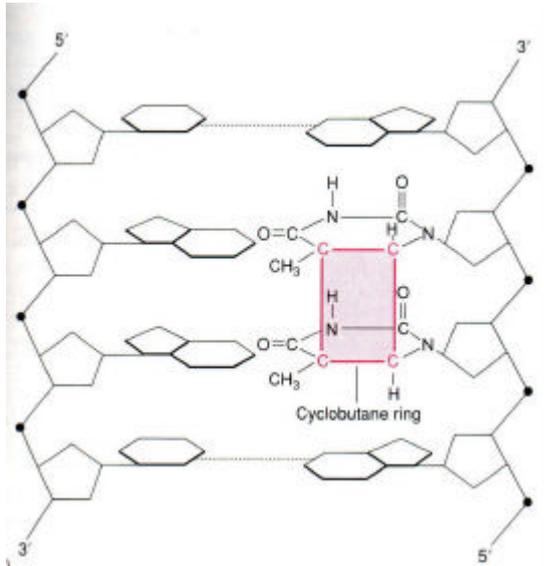


Figura 1.41 – dímero de ciclobutano pirimidínico formado por acção da luz ultravioleta.

possível, já que algumas reacções são no essencial irreversíveis.

Um dos casos em que a reversão é possível é o da fotodimerização por luz UV. O fotodímero de ciclobutano pirimidínico pode ser reparado por uma fotoliase (foi encontrada em eucariotas inferiores e bactérias, mas não em humanos).

Esta enzima liga-se ao dímero, catalisando a sua cisão em monómeros, na presença de luz visível de certos comprimentos de onda.

Um outro mecanismo de reversão directa é catalisado pelas alquiltransferases. Estas enzimas removem certos grupos alquilo adicionados às guaninas na posição O-6, por certos agentes químicos. Note-se no entanto que este é um processo enzimático

em que as alquiltransferases ficam inativadas, do que pode resultar uma saturação deste mecanismo em situações de grande alquilação do DNA.

1.1.2.5.2 Mecanismos excisivos

- Mecanismo de reparação geral por excisão:** Também denominado *nucleotide excision repair*, envolve a quebra da ligação fosfodiéster de ambos os lados da lesão, mas apenas na cadeia nucleotídica lesada. A sequência interrompida daí resultante é reparada por uma polimerase e uma ligase, tendo como padrão a cadeia nucleotídica complementar. Este mecanismo é no Homem bastante complexo, envolvendo pelo menos 17 proteínas

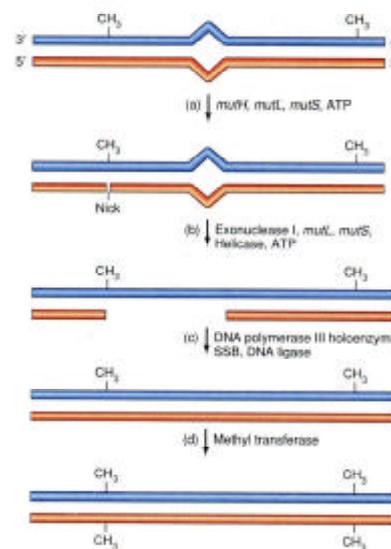


Figura 1.42 – Mecanismo geral de reparação por excisão.

- Mecanismo de reparação ligado à transcrição:** A transcrição e reparação estão também ligadas, o que demonstra o envolvimento de TFIIH (um factor de transcrição) na reparação, bem como o facto de nos eucariotas e nos procariotas, a cadeia transcrita é, nos genes activos, preferencialmente reparada, em desfavor da cadeia complementar.

- **Mecanismo excisativo específico:** Algumas lesões provocam danos geométricamente mais subtis, pelo que a sua detecção é mais difícil (não causam grandes distorções no DNA). Para estes casos, as células desenvolveram mecanismos específicos de detecção e reparação dos danos:

↳ **Reparação de DNA Glicosilase (*base excision repair*):** As DNA glicosilase não clivam as pontes fosfodiester, mas antes as ligações N-glicosilicas (base-pentose), gerando um nucleótido apurínico ou apirimidinico (locais AP). Esta lesão é posteriormente reparada pelo mecanismo de reparação endonucleolítica AP (ver abaixo). Existem numerosas glicosilases, como por exemplo Uracil-DNA glicosilase, a hipoxantina glicosilase (a hipoxantina é o resultado de uma desaminação da adenina), 3-metiladenina glicosilase, 3-metilguanina glicosilase, a 7 metilguanina glicosilase, etc.

↳ **Reparação endonucleolítica de AP:** Todas as células têm endonucleases que atacam os locais purínicos e pirimidínicos (AP endonucleases). A acção destas enzimas é vital já que como vimos a despurinação espontânea é muito elevada. Estas enzimas produzem quebras das ligações fosfodiester junto ao nucleótido apurínico. A este corte endonucleolítico segue-se um corte exonucleolítico, o preenchimento do vazio pela polimerase, e finalmente o

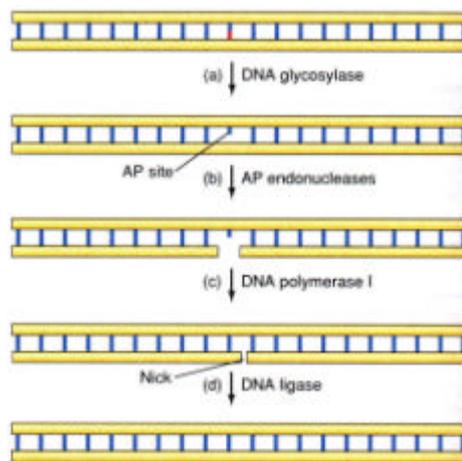


Figura 1.43 – Mecanismo de reparação endonucleolítica de locais apurínicos

restabelecimento de continuidade pela ligase.

- ☞ **O sistema GO:** Para prevenir lesões mediadas pela 8-oxodG, duas glicosilases codificadas pelos genes *mutM* e *mutY* trabalham em conjunto com o produto do gene *mutT* já referido. Quando surge um dano oxidativo produz lesões tipo GO, a glicosilase codificada pela *mutT* remove a lesão. No entanto, por vezes as lesões passam despercebidas, emparelhando com adenina. Neste caso, o produto do gene *mutY* remove a adenina levando à sua substituição pela citosina correcta por reparação de síntese, a que se segue a remoção do produto GO pelo produto *mutM*.

1.1.2.5.3 Mecanismos pós-replicativos

Alguns mecanismos de reparação são capazes de reconhecer os erros no DNA, mesmo após ou durante a sua replicação:

- **Reparação de *Mismatches*:** Este mecanismo detecta a existência de bases mal emparelhadas (*mismatches*) durante a replicação. Este processo envolve três passos essenciais: 1) a detecção da existência de um par de bases mal emparelhado; 2) o reconhecimento de que a base mal incorporada tem que ser a que existe na cadeia a ser sintetizada; 3) excisar a base incorrecta e reparar o corte. Assim, o passo crítico é o reconhecimento da cadeia sintetizada de novo, o que é conseguido recorrendo ao atraso na metilação relativamente à síntese de DNA. Assim, a cadeia de DNA não metilada é reconhecida como sendo a que deve ser processada.
- **Reparação por recombinação:** O gene *recA* envolvido no bypass

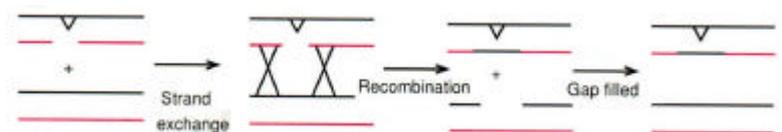


Figura 1.44 – Reparação por recombinação

SOS está também envolvido na reparação por recombinação. Neste mecanismo, o sistema de replicação pára junto a um dano, continuando depois da lesão, deixando uma quebra de alguns nucleótidos na cadeia sintetizada de novo. Neste

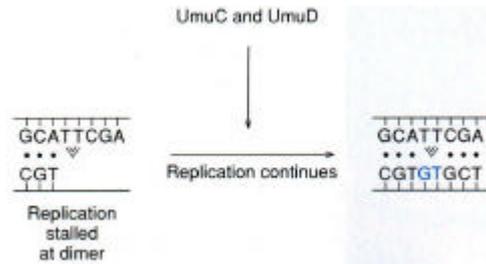


Figura 1.45 – Mecanismo de reparação SOS. este mecanismo apesar de permitir a continuação da replicação, ao introduzir aleatoriamente nucleótidos, pode produzir mutações.

Neste mecanismo de reparação, o DNA da zona de quebra é cortado da outra molécula de DNA homóloga, sendo esta depois reparada por síntese normal. Assim, este mecanismo de reparação difere do SOS, já que neste são adicionados nucleótidos de forma aleatória na zona de quebra (efectivamente criando mutações), enquanto na reparação por recombinação não são efectuadas nenhuma alteração ao programa genético da célula.

1.2 O ciclo celular

1.2.1 Divisão celular e anomalias genéticas

De acordo com a sua actividade reprodutora, a célula pode encontra-se em interfase (a fase entre duas divisões celulares), ou em fase M (a fase em que se encontra em mitose). A maior parte do ciclo celular passa-se portanto em interfase. Durante períodos específicos da interfase a célula integra estímulos e toma decisões centrais para a entrada na fase de mitose. De acordo com a actividade específica que se efectua, a

interfase subdivide-se em fases, com actividades e papéis específicos na divisão celular, e que culminam na mitose:

- **Fase G1** (de “GAP” 1; 6 a 12 horas). Não há qualquer síntese de ácidos nucleicos, mas uma intensa actividade de síntese proteica prepara a maquinaria necessária às fases

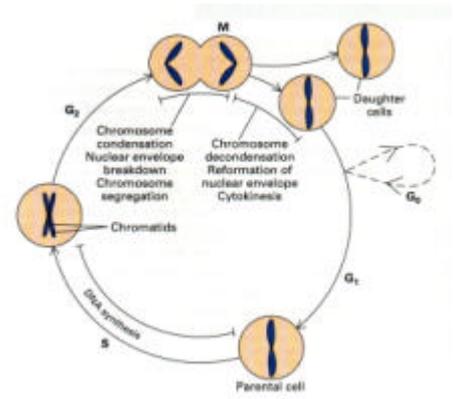


Figura 1.46 – Fases do ciclo celular

- subsequentes. A decisão de avançar para a duplicação do DNA ocorre durante a fase G1 (ponto de restrição).
- **Fase S**(6 a 8 horas). Inicia-se com a replicação do DNA, e continua até que todo o DNA se encontra duplicado (conteúdo em DNA 4n).
- **Fase G2** (de “GAP” 2; 3 a 4 horas). Decorre entre o fim da fase S e o início da mitose, durante o qual a célula organiza os seus 2 conjuntos completos de cromossomas diplóides. No final desta fase ocorre a decisão de avançar para a mitose
- Durante a **mitose** (1 hora), um conjunto diplóide de cromossomas é segregado para cada núcleo. Apenas nesta fase, os cromossomas são observáveis como entidades discretas, sendo a organização celular destruída, e organizado o fuso acromático entre os dois futuros núcleos. Toda a actividade sintética pára durante a mitose.

- **Fase G0.** Uma fase específica de células maduras, que pararam a divisão celular. No entanto alguns tipos de células, podem ser estimulados a deixar a fase G0 e entrar em G1.

1.2.2 Ciclinas e CDK

Durante todo o ciclo celular, existem pontos de decisão, os quais verificam se a célula se encontra já preparada para proceder para a fase seguinte. Alguns dos mais importantes são: O DNA já se encontra todo duplicado, e apenas duplicado? Existem erros na duplicação do DNA? A massa celular está já duplicada? Estes “checkpoints” existem sob a forma de factores activadores e inibidores. Destes, alguns encontram-se já identificados enquanto que a existência de outros se encontra apenas deduzida do comportamento observado em fusões de células em fases diferentes do ciclo celular. Nomeadamente o factor que determina a decisão de passar o ponto de restrição é totalmente desconhecido, enquanto que o factor que determina a entrada na fase S se encontra identificado como sendo uma proteína cinase, relacionada com a cinase que inicia a mitose. Esta última (MPF – M phase promoting Factor ou M phase Kinase), a primeira a ser identificada é a mais bem estudada, é um dímero composto por duas subunidades: uma a subunidade catalítica (p34) é activada no início da fase S, a outra (p45) é uma ciclina (acumula-se por síntese contínua durante a interfase), sendo destruída durante a mitose, o que é responsável pela inactivação da cinase da fase M, o que sinaliza a saída da mitose.

O estudo cristalográfico do dímero de MPF revelou que a ligação da ciclina provoca uma alteração conformacional na p34, a qual é essencial para a formação do centro catalítico. Uma MPF tem apenas um tipo de p34, mas pode conter uma de vários tipos de ciclinas. Basicamente dois tipos de ciclinas pouco relacionadas (A e B) podem fazer parte da MPF. A similaridade entre estas duas ciclinas centra-se

numa estrutura com cerca de 150 aminoácidos chamada a cyclin box. Em mamíferos existem ainda dois tipos de ciclinas B (B1 e B2).

A p34 encontra-se em níveis constantes ao longo de todo o ciclo celular, e como a síntese de ciclina é constante ao longo de toda a fase G1, a sua presença não constitui o sinal de activação da MPF. Na realidade, o dímero acumula-se numa forma inactiva, sendo a modificação da p34 que constitui o evento activador do dímero. A activação da p34 é efectuada por fosforilação/desfosforilação, já que para esta se encontrar activa é necessária a presença de grupos fosfatos em alguns resíduos e a sua ausência noutros. Como a actividade cinase da MPF é auto-catalítica, basta a activação de uma pequena porção, para que rapidamente toda a MPF disponível seja activada. Como a MPF não é uma fosfatase, e é necessária actividade de cinase e fosfatase para a activação da MPF, esta deve activar directamente a fosfatase que a activa, criando um circulo auto-catalítico completo.

O evento inactivador é constituído pela destruição proteolítica da ciclina, a qual ocorre durante a fase M.

Na realidade, o controlo da divisão celular é constituído por uma intrincada rede de reacções de cinase e fosfatase, que desencadeiam a activação/inactivação de uma série de factores, permitindo o desencadear temporalmente organizado das várias fases do ciclo celular. Toda esta série de reacções culminam na activação da MPF com a concomitante entrada na fase M. Dois modelos gerais podem explicar a actividade da MPF: 1) Pode ser um regulador central que fosforila proteínas alvo que por sua vez actuam para regular outras actividades, i.e., pode tratar-se de uma reacção clássica em cascata; 2) pode ser um regulador central, que activa ele próprio uma serie de substratos cruciais indispensáveis para realizar trabalhos regulatórios ou reorganizacionais envolvidos no ciclo celular.

Os substratos da MPF têm em comum uma estrutura proteica constituída por Ser-Pro flanqueada por resíduos básicos (habitualmente

na forma Ser-Pro-X-Lys). Substratos potenciais incluem a histona H1 (possivelmente necessária para condensar os cromossomas), lamininas (possivelmente necessárias para desorganizar o envelope nuclear), nucleolina (possivelmente envolvida na paragem da actividade ribossomal), bem como outras proteínas estruturais e enzimáticas. Dos modelos acima propostos, o mais provável é que a MPF actue directamente em todos (ou na maioria) destes substratos.

A proteína p34 foi pela primeira vez identificada no *Xenopus*, que pela dimensão do seu ovo (1mm de diâmetro) constitui um bom modelo para purificar proteínas envolvidas no ciclo celular. Posteriormente descobriu-se que o homólogo na levedura eram proteínas baptizadas como cdc2 (*S.Pombe*) e CDC28 (*S.Cerevisiae*). Dado o grande paralelo entre as proteínas de levedura e de mamíferos, o homólogo animal é habitualmente designado por cdc2 (o nome do exemplo mais bem caracterizado em levedura). Em todas as espécies, a proteína que emparelha com a cdc2/p34 para formar o dímero activo é uma ciclina, muito embora o nível de homologia entre espécies nas ciclinas seja bem menor que na subunidade catalítica (p34).

Na levedura o cdc2 é um regulador central quer da decisão de prosseguir de G1 para S quer de prosseguir de G2 para M. Em cada fase, cdc2 tem um par diferente no dímero: na mitose cdc2 emparelha com cdc13 formando uma cinase da fase M semelhante ao p34-ciclina B das células animais. Durante a fase G1, a forma activa de cdc2 tem um par diferente chamado cig2, e igualmente semelhante a ciclina-B. Note-se que as 2 formas de dímeros cdc2 podem coexistir na levedura, mas são diferentemente reguladas. Assim, a fase do ciclo celular pode também ser definida consoante o tipo de dímero activo em cada momento.

O produto de cdc25 é necessário para desfosforilar a cdc2 do dímero cdc2/cdc13. Apesar de não possuir um domínio clássico de fosfatase, a cdc25 é efectivamente uma fosfatase, que possivelmente tem como seu

alvo a tyr-15 da cdc2. Assim é provavelmente responsável pela desfosforilação que constitui o evento chave na activação da M phase cinase. Como mutantes de cdc25 não param a mitose se a duplicação de DNA não tiver sido conseguida, esta fosfatase é importante para assegurar que a fase S se completa antes de se iniciar a fase M.

Mutantes do gene wee1 permitem que a fase M se inicie sem que o crescimento necessário tenha ocorrido. Assim, este gene normalmente impede o início da fase M, se a massa celular crítica não tiver sido atingida. Wee1 codifica para uma cinase pouco usual. Pode fosforilar em serina/treoninas e tirosinas. Inibe cdc2 por fosforilação em tyr-15. Um outro gene mik1 tem efeito semelhante. Este controlo da cdc2 na transição G2/M parece ter sido preservado ao longo da evolução, já que genes homólogos são encontrados

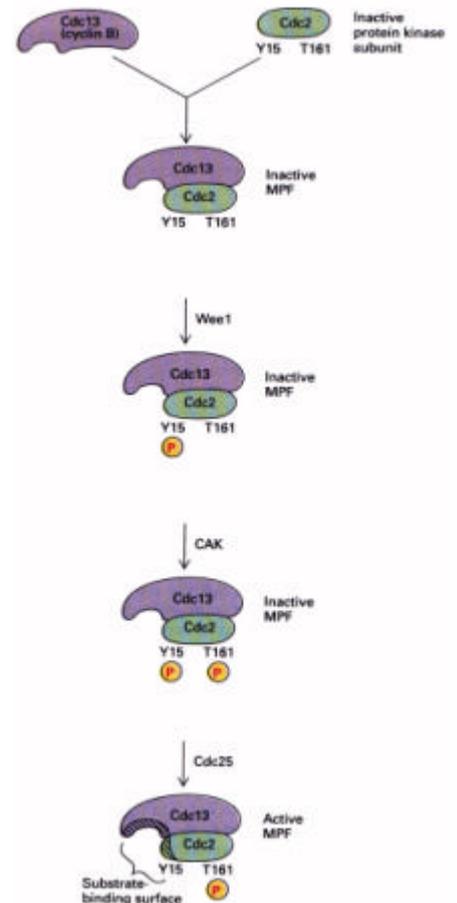


Figura 1.47 – via de activação da MPF por fosforilação da Y15 e T161 de Cdc2 respectivamente mediada pela Wee1 e CAK, e pela acção da Cdc25

em vários tipos de leveduras, em anfíbios e em mamíferos.

A activação da transição G1/S requer activação da *cdc2/cig2* (em *S. Pombe*), mas também a inactivação da *cdc2/cdc13* é indispensável. Com efeito, mutantes de *cdc13* não só não entram em fase M, como prosseguem por vários ciclos da fase S, sugerindo que a *cdc2/cdc13* é um inibidor da fase S. Assim, activação de *cdc2/cd13* durante a fase G2 impede a continuação da fase S, e estimula o início da fase M. A destruição da *cd13* no fim da fase M pára a fase M e permite um novo ciclo de fase S antes de se iniciar nova divisão. O funcionamento deste ciclo depende, provavelmente de *cdc18*. Esta proteína é activada após a passagem do ponto START, sendo necessária para entrar na fase S. Para *cdc18* ser activa é necessário que *cdc2/cdc13* esteja inactiva. A

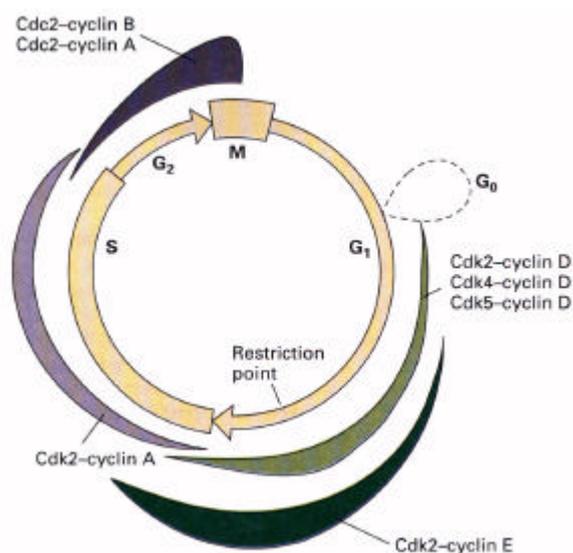


Figura 1.48 – Intervenção das várias ciclinas em diferentes fases do ciclo celular.

activação deste dímero na fase M inactiva a cdc18, impedindo novo ciclo de síntese de DNA. A actividade de cdc2/cdc13 é dependente do factor rum1. Sem este factor, as células entram em mitose prematuramente. Assim, rum1 deve ser um inibidor de cdc2/cdc13 (M phase kinase) é expresso entre G1 e G2 mantendo a M phase kinase inactiva (o que é particularmente importante em G1 antes da fosforilação em tyr-15), e reprimindo o nível de cdc13.

Desta imagem, dois conceitos gerais emergem para o controlo da divisão celular: 1) a presença de loops de feedback de controlo e; 2) a presença de duplicidade de efeitos, em que um factor activa a fase seguinte e reprime a anterior.

A passagem do ponto START está também dependente de cdc2. No entanto, a ciclina que se associa a cdc2 é diferente quando a cinase controla o ponto start e a transição G2M.

Os mecanismos envolvidos na regulação do ciclo celular em leveduras e em animais são semelhantes, ainda que o número de subunidades presentes em animais seja maior, havendo subunidades específicas para o controlo das transições G1/S e G2/M.

A comparação das subunidades envolvidas nos diferentes seres estudados estão descritas na tabela 1.4.

Como se pode ver, nos animais, a cinase dependente de ciclinas que regula a transição G2/M é única, á semelhança do que se passa com as leveduras. No entanto, existem várias ciclinas que com ela podem interactivar. O mesmo não se passa na transição G1/S, a qual é regulada por várias cinases e várias ciclinas.

Tabela 1.4 - comparação das subunidades envolvidas nas várias transições do ciclo celular em diferentes organismos

	Transição			
	G1/S		G2/M	
	subunidade catalítica	subunidade reguladora	subunidade catalítica	subunidade reguladora
<i>S.Pombe</i>	cdc2	cig 1,2	cdc2	cdc13
<i>S.cerevisae</i>	cdc28	CLN1-3	cdc28	CLB1-4
<i>Mamíferos</i>	cdk2(p33), cdk4	ciclinas A, D1, D2, D3, E	p34(cdc2)	ciclinas A, B1, B2

Cdk = cyclin dependent kinases

Uma outra diferença tem a ver com as ciclinas C e E, as quais acumulam durante a fase G1 como as restantes, mas não são destruídas na fase M.

Assim, como vimos, nos animais existem essencialmente 2 tipos de cinases : a cdk2 e a cdc2, respectivamente responsáveis pela transição G1/S e G2/M. Ainda que se trate de componentes diferentes, as estruturas regulatórias conhecidas são similares em ambas, pelo que se deduz que são similarmente regulados. Assim, a síntese de ciclinas na fase G1, vai progressivamente aumentando a sua concentração, até que se chega a um ponto em que se inicia a formação de dímeros. No entanto estes não são ainda activos, sendo necessárias reacções de fosforilação nos resíduos 15 e 161. A primeira fosforilação ocorre por acção da Wee1 cinase, enquanto a segunda é catalisada pela CAK (cdc2-activating kinase). O dímero activo promove então a entrada em mitose, activando concomitantemente por fosforilação a fosfatase cdc25. Esta por sua vez promove a remoção do grupo fosfato na tirosina 15 da cdc2, inactivando-a. A ciclina é por sua vez alvo de degradação proteolítica, resultando apenas a cdc2 fosforilada na thr-161, terminando assim a mitose. Não foram identificadas cinases que

actuem separadamente em thr-14, em alguns casos a mesma enzima actua na thr-14 e tyr-15.

A mitose ocorre como um processo temporal controlado. A mitose, como vimos, é iniciada pela activação da MPK. O progresso da mitose requer a degradação de ciclinas e outras proteínas. Por exemplo a separação dos cromossomas na anafase requer a actividade proteolítica, mas não directamente de ciclinas. Existem pelo menos 3 alvos proteolíticos: o primeiro evento é a degradação da ciclina A na metafase, seguindo-se a degradação de 2 alvos na anafase. Uma proteína desconhecida tem que ser degradada para a separação das cromátidas irmãs e a degradação de ciclina B é necessária para a inactivação da MPK. No final da mitose as fosforilações efectuadas pela MPK têm que ser revertidas.

A síntese de ciclinas D é activada pela acção de factores de crescimento. Estas ciclinas têm uma semi-vida muito curta, o que permite à célula responder a factores externos e à sua remoção. Estas ciclinas estão provavelmente envolvidas na reentrada na divisão celular de células em G₀.

O produto do gene de supressão tumoral RB (gene do retinoblastoma) é um substracto dos complexos cdk-ciclinas D, e exerce o seu efeito durante a fase G₁ que precede o ponto de restrição. O produto deste gene, na ausência de fosforilação liga-se ao regulador genético E2F, actuando como supressor de alguns genes, bloqueando a transição G₁/S. Quando por acção de cdk4,6-ciclina D1,2,3 é fosforilado, liberta a E2F, a qual passa a actuar como um activador genético, promovendo a entrada na fase S. RB é o alvo de várias vias que inibem o crescimento celular, pelo que pode ser uma forma importante de vários sinais manterem a célula em G₁ ou G₀. Alguns destes sinais (incluindo o TGFβ) actuam através de inibidores de cinases cdk (chamadas ckis). Exemplos de ckis são a p15/p16 (ink4), p21(c1p1/wAF1) e p27(kip1). A importância das ckis é realçada pelo facto de p16 ser também um

gene de supressão tumoral, o que indica que a p16 e possivelmente todas as outras ckis são necessárias para parar o crescimento celular. Na verdade, parece que a via das ckis até RB é uma via ventral para o bloqueio do crescimento celular, já que são conhecidos genes de supressão tumoral em todos os seus passos.

1.2.3 Apoptose

Durante a vida de um organismo multicelular, algumas células morrem num processo natural designado por Apoptose ou morte celular programada. Nos vertebrados, os exemplos mais visíveis de apoptose ocorrem no sistema imune e no sistema nervoso. O processo de morte é característico, envolvendo compactação celular, fragmentação membranar, condensação da cromatina e fragmentação do DNA. Este processo é activo, dependendo do RNA e da síntese proteica.

São conhecidas várias formas de activar a apoptose. Estas envolvem insultos moleculares (retirada de factores de crescimento, tratamento com glucocorticoides, irradiação γ), bem como estímulos específicos como os produzidos pelas células T citotóxicas ou activação de p53. Assim, a apoptose é importante na embriogénese e contenção do crescimento tecidual, bem como na resposta imune e na contenção de cancro.

Foram já descritas várias mutações recessivas associadas com o estímulo ou o bloqueio da apoptose. Alguns exemplos são: 1) a lpr no ratinho (esta mutação leva a deficiência em faz) que causa a proliferação excessiva levando a autoimunidade; 2) a gld (generalized lymphoproliferative disease) cuja mutação ocorre no gene que codifica para o receptor de fas. A proteína faz é um receptor membranar relacionado com o receptor para o TNF. Tanto o faz como o TNF são capazes de estimular a apoptose. Ao nível da porção transmembranar destes receptores existe uma sequência de 80 aminoácidos essencial para o envio do sinal apoptótico, pelo que é designado por “domínio da

morte”. Pouco se sabe sobre o mecanismo de sinalização do receptor faz ou TNF para o interior da célula. Contudo, foram identificadas algumas proteínas que interagem com o “domínio da morte”, as quais curiosamente também os possuem. Pensa-se assim que este domínio seja importante para promover a dimerização destas proteínas. Outros compostos importantes na via da apoptose são a família de proteases designada por ICE. O protótipo destas proteases é a “*IL2 β -converting enzyme*”: uma protease de cisteína que por proteólise converte o precursor da IL2 β na sua forma activa. O processo de apoptose desencadeado pelas ICE é inibido por crmA e pelo bcl-2.

Uma outra via apóptótica é desencadeada pelos linfócitos Tc, os quais matam as células alvo libertando na sua superfície grânulos contendo proteases de serina e outras proteínas líticas como a perforina, abrindo buracos na superfície das células alvo. As proteases de serina dos grânulos são chamadas granzimas, e na presença de perforina induzem morte por apoptose nas células alvo.

O bcl-2 inibe a apoptose. Foi originalmente descrito como um protooncogene activado em linfomas por translocação causando a sua hiper-expressão. O bcl-2 tem uma sequência de ancoragem membranar na zona c-terminal, tendo sido encontrado nas membranas externas de mitocôndria, núcleo e retículo-endoplasmático. Um outro gene semelhante ao bcl-2 nos mamíferos é o gene Bax. O produto do gene Bax pode dimerizar com bcl-2. O dímero bcl-2, bem como o dímero Bax bloqueiam a apoptose, mas o heterodímero (bcl-2/Bax) não o faz. Assim, a susceptibilidade de uma célula à apoptose é proporcional à proporção bcl-2/Bax.

1.3 Actividade Celular

1.3.1 Componentes dos sistemas de sinalização

Todos os sistemas de comunicação intercelular têm vários componentes. Tipicamente, uma molécula denominada ligando é libertada pela célula sinalizadora. Alguns ligandos são proteínas, enquanto outros são pequenas moléculas como péptidos, esteróides ou vitamina D. O ligando liga-se ao receptor, habitualmente uma proteína, na membrana celular ou no interior da célula alvo. Alguns tipos de complexo ligando-receptor são capazes de alterar a expressão genética

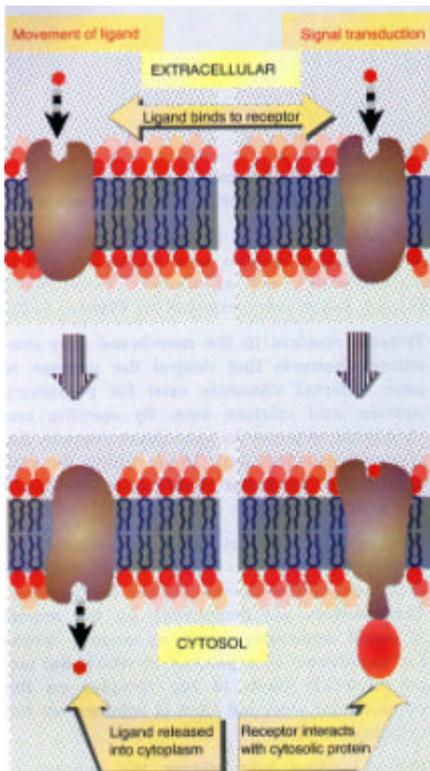


Figura 1.49 – Os sinais podem ser recebidos na membrana celular por um receptor o qual envia o sinal para o interior da célula. Em alguns casos a molécula sinalizadora é ela própria enviada para o citoplasma, enquanto noutros casos é o receptor que interage com componentes intracelulares para assegurar a passagem do sinal.

directamente, enquanto outros necessitam de enviar sinais através de uma cadeia de transdução, levando o sinal da membrana celular até ao núcleo (fig.1.49).

Alguns ligandos, chamados hormonas viajam a grandes distâncias na circulação sanguínea, antes de interagirem com a sua célula alvo. Estas moléculas podem actuar como sinais mestres para a actividade de órgãos diferentes, os quais podem assim responder de forma coordenada. Outros ligandos não actuam à distância, mas apenas nas células vizinhas das que os produziram.

1.3.2 Regulação da sinalização por modulação da conformação proteica

Interações de pequenas moléculas com as proteínas receptoras podem provocar intensas alterações conformacionais nestas. Por exemplo, alterações conformacionais estão na origem da regulação mediada por algumas proteínas cinases. A alteração da conformação coloca ou não amino-ácidos em posições favoráveis à fosforilação. Este mecanismo é extremamente eficiente, não só porque permite a rápida activação dos sinais, mas também porque permite revertê-los rápida e facilmente, reciclando os componentes de sinalização.

1.3.3 A Superfamília dos receptores das hormonas esteróides

Um grupo diverso de ligandos, incluindo várias hormonas esteróides, bem como outras pequenas moléculas como a hormona tiróide e a vitamina D, devido à sua estrutura não polar, são capazes de atravessar a membrana citoplasmática, actuando directamente no interior da célula alvo. Os receptores para estes ligandos chamam-se globalmente a superfamília dos receptores esteróides, e estão estrutural e funcionalmente relacionados. A ligação destes ligandos ao seu receptor, causa uma alteração conformacional, a qual provoca a sua libertação de uma proteína sequestradora. Nestas condições, o

complexo receptor-ligando desloca-se para o núcleo, onde em conjunto com outros reguladores da transcrição vão activar ou reprimir directamente a expressão génica, ao ligar-se a sequências reguladoras chamadas HRE (“*hormone response elements*”).

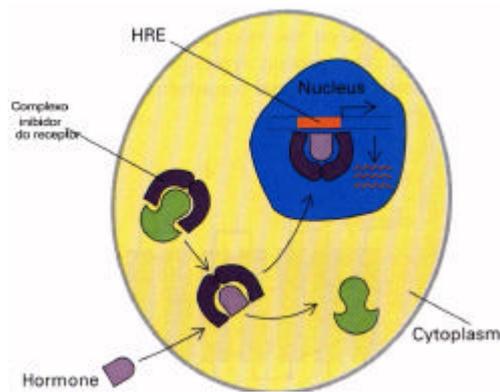


Figura 1.50 – As hormonas esteróides não possuem um receptor membranar, mas um receptor citoplasmático, já que são capazes de atravessar livremente a membrana plasmática.

1.3.4 Receptores transmembranares; vias de transdução de sinal

A maior parte dos ligandos são moléculas demasiado volumosas e/ou com carga, para poderem atravessar a membrana citoplasmática. Assim, os seus receptores são habitualmente proteínas de membrana, as quais medeiam a passagem de um sinal para o interior da célula. Os receptores transmembranares são assim compostos por 3 domínios principais: o extracelular, que se liga ao ligando; o transmembranar, que pode atravessar uma ou mais vezes a membrana; e o citoplasmático, que medeia a transmissão do sinal.

Muitos ligandos são dímeros, podendo assim ligar a mais do que um receptor. Desta forma, os ligandos colocam a porção citoplasmática de dois receptores fisicamente próximos, activando a via de sinalização destes receptores (Fig. 1.51). Alguns destes receptores são cinases, tendo por isso a habilidade de fosforilar certos resíduos em proteínas

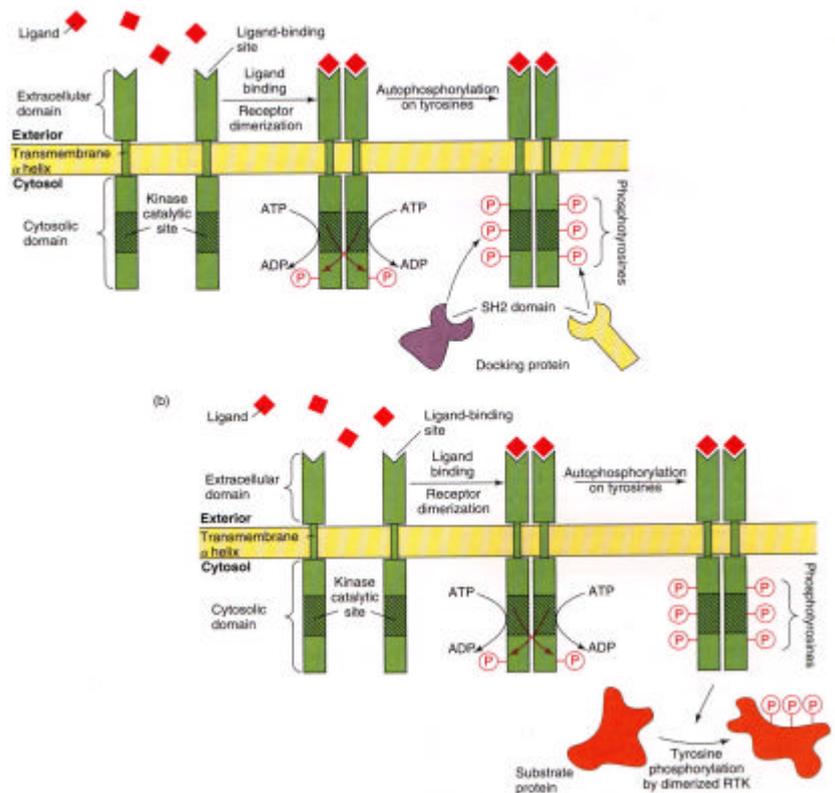


Figura 1.51 – Muitos receptores são dimerizados pelos seus ligandos, o que activa a zona citoplasmática, a qual pode a) ligar-se a proteínas cinases ou b) autofosforilar-se activando a sua capacidade de fosforilar substratos citoplasmáticos

alvo, enquanto outros são serinas/treoninas cinases, e outros não têm qualquer actividade enzimática.

Um dos mais bem conhecidos exemplos de receptores deste tipo são os receptores tirosina cinase, nomeadamente os receptores para os factores de crescimento. Estes receptores, após ligação do ligando

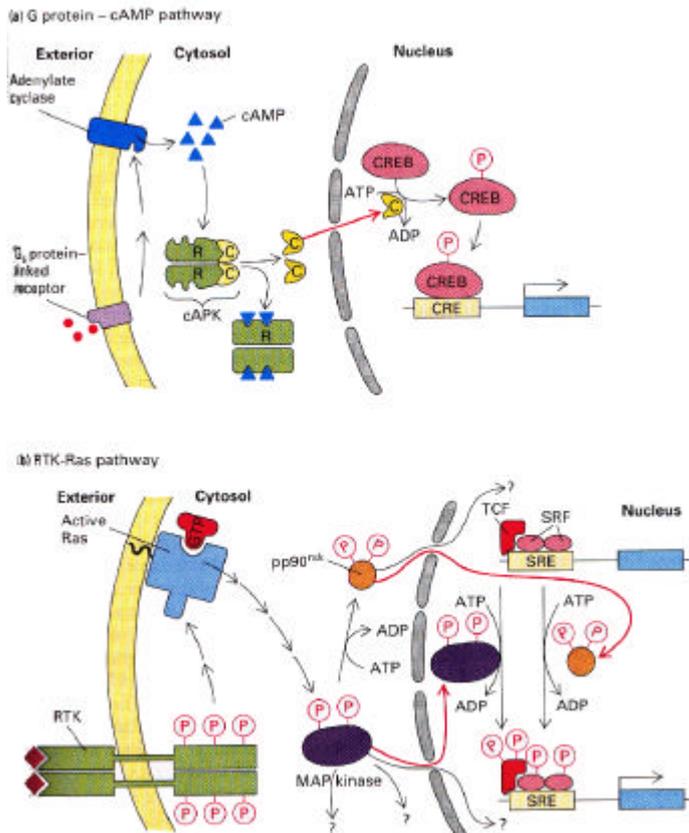


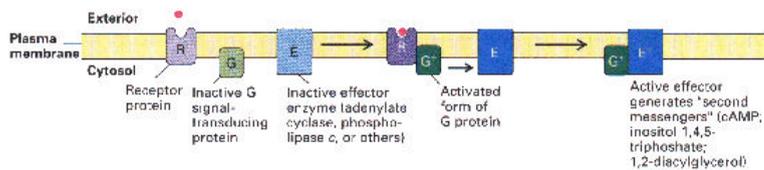
Figura 1.52 – Com frequência a propagação do sinal no interior da célula envolve a utilização de mensageiros secundários como a) cAMP ou b) proteínas G

dimerizam, originando a sua auto-fosforilação (fig. 1.51). Esta por sua vez activa a actividade de cinase, que vai fosforilar outras proteínas, iniciando uma cascata transdutora do sinal em que a fosforilação causa conformação proteica com conseqüente activação/inactivação proteica. Eventualmente a cascata de activação leva à fosforilação de factores de

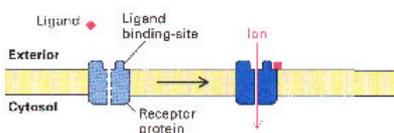
transcrição, com conseqüente alteração do estado de activação de um ou mais genes. Mas a auto-fosforilação não causa apenas a fosforilação de proteínas alvo.

Também a ligação a proteínas adaptadoras é estimulada pela auto-fosforilação. Interações entre o complexo e outras moléculas causam a propagação do sinal. (fig. 1.51) Com frequência a propagação deste sinal envolve proteínas G (fig. 1.52). As proteínas G são proteínas cuja vida é um ciclo entre a forma ligada a GDP (o estado inactivo) e a forma ligada a GTP (o estado activo); um exemplo particularmente importante de uma proteína G é a *ras*, a qual como veremos está envolvida na carcinogénese. A propagação do sinal das RTK leva à activação de uma proteína que se liga a proteínas-G inactivas, alterando a sua conformação, levando-a a ligar-se ao GTP. A proteína G assim activada liga-se então a uma proteína cinase citoplasmática, alterando a sua conformação, activando-a, o que a leva a fosforilar outras proteínas, incluindo proteínas cinases e factores de transcrição.

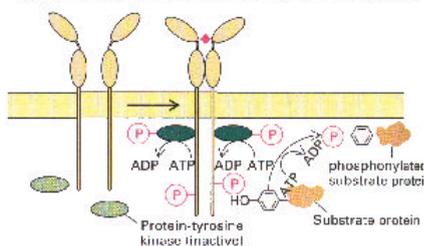
(a) G protein-linked receptors (epinephrine, glucagon, serotonin)



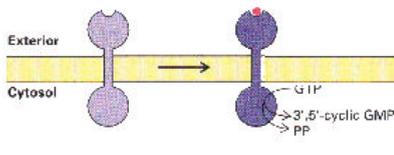
(b) Ion-channel receptors (acetylcholine)



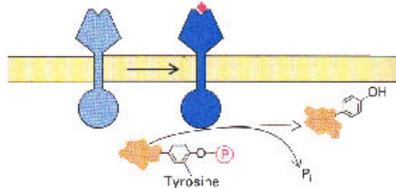
(c) Tyrosine kinase-linked receptors (erythropoietin, interferons)



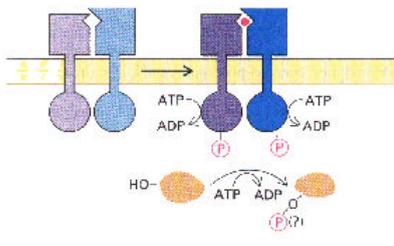
(d) Receptor guanylate cyclase (atrial natriuretic factor)



(e) Receptor tyrosine phosphatase (leukocyte CD45 protein)



(f) Receptor serine/threonine kinases (transforming growth factor β)



(g) Receptor tyrosine kinases (insulin, epidermal growth factor)

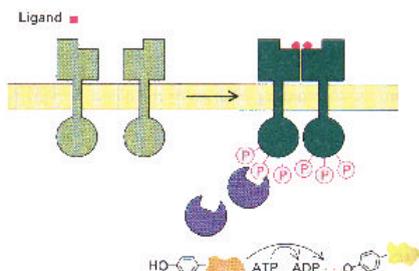


Figura 1.53 – Sumário dos vários tipos de receptores e vias de transdução de sinal.

