

SÍNTESE DE RNA – TRANSCRIÇÃO

Introdução e conceitos:

Transcrição é o processo pelo qual uma molécula de RNA é sintetizada a partir de um molde de DNA. Através da transcrição, são sintetizados todos os tipos de RNAs da célula, ou seja, o RNA mensageiro (mRNA), o RNA ribossômico (rRNA), o RNA transportador (tRNA) e outros RNAs menores. Todos os RNAs estão envolvidos ativamente na síntese protéica. O mRNA será usado para transferir a informação genética do DNA às proteínas, mas os demais RNAs sintetizados têm, por si, funções finais na célula, tanto estruturais como catalíticas.

É principalmente durante a transcrição que a célula exerce o controle da expressão gênica. Os genes não são transcritos indiscriminadamente, pois esse processo é regulado por proteínas. Na maioria dos casos, o principal ponto de regulação da atividade de um gene é a decisão de iniciar ou não a sua transcrição.

Existem muitas semelhanças entre a síntese de DNA e a de RNA. Entretanto, a função biológica de cada um desses processos é bastante diferente. A síntese de DNA deve ser precisa e uniforme. Estes conceitos se referem à necessidade da nova cópia de DNA ser exatamente uma réplica da original, e que esta síntese englobe toda a extensão da cadeia progenitora. Em contradição, a transcrição espelha o estado fisiológico da célula. Ela é extremamente variável para atender às suas necessidades num determinado momento. Tal variabilidade se reflete no fato de que somente um gene ou grupo de genes em particular é transcrito naquele instante fisiológico. Por outro lado, dependendo da necessidade da célula, o mesmo gene pode ser transcrito incontáveis vezes até que a demanda seja suprida.

Além da característica temporal da transcrição, este é um processo extremamente seletivo. Esta seletividade acontece em dois níveis: (1) somente partes do DNA são transcritas para produzir RNA. Por exemplo, nas células da maioria dos mamíferos, somente 1% de toda a seqüência do DNA é copiada para formar seqüências de RNA funcional e algumas partes do DNA podem nunca ser transcritas, e (2) somente uma porção ainda menor da seqüência de nucleotídeos do RNA sobrevive os passos de processamento pelo qual o RNA recém transcrito passa para se tornar maduro e funcional. Para resumir, a célula restringe a expressão de informação genética para a formação de produtos gênicos necessários naquele momento especificamente.

Mecanismos de transcrição:

A transcrição ocorre a partir da informação contida na seqüência de nucleotídeos de uma molécula de DNA fita dupla, sendo sempre no sentido 5' → 3'. Apenas uma das fitas do DNA, chamada de fita molde, é utilizada durante a síntese, que segue as mesmas regras de complementaridade e antiparalelismo, exceto pelo pareamento de uracil (U), ao invés de timina (T), com adenina (A). O RNA recém sintetizado (5' → 3') é, portanto, complementar à fita de DNA que serviu de molde (3' → 5') e idêntico a outra fita de DNA do duplex (5' → 3'). Por convenção, entretanto, a seqüência de nucleotídeos de um gene é sempre representada na orientação 5' → 3', ou seja, a fita que não serve de molde.

Em 1960 foi descoberta uma enzima capaz de, na presença de DNA fita dupla e dos ribonucleotídeos trifosfatados (ATP, CTP, GTP e UTP), sintetizar RNA. Essa enzima foi denominada RNA-polimerase (RNAP). A reação catalisada pelas RNAPs é mecanisticamente idêntica à reação catalisada pelas DNA-polimerases.

As RNAPs desempenham todas as atividades necessárias para a transcrição:

- 1) reconhecem e ligam-se a seqüências específicas de DNA;
- 2) desnaturam o DNA, expondo a seqüência de nucleotídeos a ser copiada;
- 3) mantêm as fitas de DNA separadas na região de síntese;
- 4) mantêm estável o duplex DNA:RNA na região de síntese;
- 5) restauram o DNA na região imediatamente posterior à da síntese; e
- 6) sozinhas ou com auxílio de proteínas específicas, terminam a síntese do RNA.

Existem, pelo menos, cinco RNAPs diferentes que estão envolvidas na síntese de RNAs específicos: RNAP I, RNAP II, RNAP III, RNAP mitocondrial (mtRNAP), e RNAP de cloroplastos. As RNAPs nucleares são abordadas a seguir:

A) *RNAP I, que sintetiza os rRNAs*

Essa polimerase catalisa a síntese de apenas um tipo de RNA, precursor dos rRNAs 18S, 5,8S e 28S. Entretanto esse é o RNA mais abundante na célula. Esses pré-RNAs são bastante longos, variando de 6.000 a 15.000 nucleotídeos. Para poder responder a essa demanda, a célula possui de 100 a 5.000 cópias dos genes para rRNA, uma ao lado da outra, e sua síntese ocorre em um ponto especializado no núcleo, o nucléolo. O complexo RNAP I é composto por 2 subunidades maiores (com 130 e 190 kDa) e por 4-10 subunidades menores.

B) *RNA polimerase II que sintetiza mRNA*

Todos os pré-mRNAs das células são sintetizados pela RNAP II, que, portanto, transcreve a maior parte dos RNA heterogêneos nucleares (hnRNA) precursores dos mRNAs. As RNAPs II possuem duas subunidades maiores de 215 e 139 kDa e vários componentes menores (de 6-8), de aproximadamente 50 kDa.

C) *RNAP III, que sintetiza os tRNA, rRNA 5S, e outros RNAs*

Essas polimerases catalisam a síntese de cerca de 10% do RNA da célula, do rRNA 5S, dos tRNAs, e de outro pequenos RNAs. A enzima é a mais complexa das RNAPs, com tamanho aproximado de 700 kDa, sendo formado por cerca de 14 subunidades. Assim como as outras RNAPs, 2 subunidades são maiores, com 160 e 128 kDa, e têm identidade com as correspondentes subunidades de RNAP II.

Embora consigam sintetizar RNA a partir de nucleotídeos livres, as RNA polimerases sempre precisam de proteínas acessórias, chamadas de fatores de transcrição, que auxiliem o início da transcrição. São estes fatores que identificam ao longo de um gene as seqüências específicas no DNA que indicam o local de início da transcrição, ou seja, a partir de qual base o gene deve ser copiado em RNA. O primeiro par de bases do segmento de DNA que será transcrito é chamado de sítio de início e recebe o valor +1. Seqüências anteriores ao sítio de início são chamadas *upstream* sendo que as bases progridem negativamente (-1, -2, -3...) a partir do +1 para trás. Seqüências localizadas após o sítio de início são chamadas *downstream*, e progridem positivamente (+2,+3, +4...).

As seqüências reguladoras da transcrição podem ser divididas em dois tipos: *promotores* e *elementos reforçadores (enhancers)*. Essas seqüências reguladoras nada mais são do que um padrão de bases encontrado em todos os genes, que tem uma função designada de acordo com os fatores/proteínas/enzimas que o reconhecem e se ligam a ele. Este conceito é como um “código enigmático e padrão” de bases que é entendido pelo agente que vai interpretá-lo. O agente específico capaz de decodificar este “código” se aproxima e sobre ele atua. Isto se dá de duas formas: ligando-se a ele e assim sinalizando a outros agentes o local de ação (*promotor*), ou ainda liga-se ao código, mas para modular a atividade de outros agentes

(*enhancer*). Por estarem presentes em todos os genes, podem ser chamadas de região consenso ou conservada, termos que também se aplicam para outras seqüências que não somente regulam, mas sinalizam .

Os *promotores* para transcrição no DNA são inicialmente reconhecidos por fatores de transcrição que, ligados ao DNA, interagem com outros fatores, formando um complexo ao qual as RNAPs se associam. Os promotores são, em geral, seqüências de 100 a 200 pb. Eles sempre estão próximos ao sítio de início da transcrição e possuem seqüências consenso, comuns a todos os *promotores*, que são reconhecidas pelos fatores de transcrição. Eles são responsáveis por uma transcrição em nível basal – muito baixo.

Os *enhancers* são seqüências pequenas de DNA que podem ocorrer na região 5' do gene, antes do *promotor* ou após o terminador, e ativam a expressão. Podem estar muito distantes do promotor (até milhares de pb) e ainda assim serem ativados. Alguns desses elementos de regulação dão a especificidade à transcrição, tanto temporal quanto tecido-específico.

O processo pode ser dividido em 3 partes: iniciação, alongamento e terminação. Essas etapas serão detalhadas para os genes transcritos pela RNAP II, genes que codificam proteínas.

Iniciação:

O primeiro passo da transcrição requer regiões reguladoras, acoplamento de fatores de transcrição, e fatores de RNAP II. Embora a organização das regiões reguladoras da transcrição por RNAP II variem para os diferentes genes e nos diferentes organismos, existem alguns componentes básicos:

A) elementos que constituem o *promotor*:

1 - *sítio de início da transcrição* representado pela primeira base a ser transcrita;

2 - *uma região que inclui uma seqüência conservada TATA Box* localizada na região -25, anterior, portanto, ao sítio de início da síntese. Este elemento determina onde a síntese deve ser iniciada e, normalmente, não afeta os níveis de expressão. É aqui que ocorre o reconhecimento inicial do gene a ser transcrito por parte dos fatores de transcrição;

3 - *uma região regulatória upstream denominada CAAT Box* (consenso GGNCAATCT), situada na região -60 a -80.

4 - *outras seqüências upstream*. Aqui e no CAAT Box ligam-se fatores que aumentam a eficiência de reconhecimento do promotor.

B) *enhancers*. Esses elementos aumentam a expressão do gene, independentemente da sua orientação, localização (a 5' do promotor ou após o sinal de término da transcrição) e distância do promotor.

O início da transcrição em genes que utilizam RNAP II é complexo e envolve uma cascata na qual vários fatores de transcrição vão se ligando ao DNA. Primeiramente há ligação de um fator ao TATA Box. Essa ligação sinaliza para que outros fatores se liguem entre si e a outros elementos do promotor. Quando todos os fatores já estiverem em seus devidos locais, o complexo estará pronto para receber a RNAP II e formar o chamado complexo de iniciação da transcrição. A RNAP II é responsável por abrir as fitas de DNA, e colocar a primeira base. Essa abertura das cadeias forma a “bolha de transcrição”, estrutura que é característica do processo. Para ativar a transcrição, outros fatores ligam-se ao *enhancer* e interagem com o complexo de iniciação da transcrição.

Alongamento:

É o aumento da cadeia de RNA para o qual outros fatores acessórios são necessários. Os fatores utilizados no complexo de iniciação da transcrição são liberados e a RNAP II sofre alterações na sua

conformação. A “bolha de transcrição” move -se sobre o molde de DNA, abrindo o DNA à frente e fechando atrás de si. Um duplex híbrido DNA-RNA forma-se dentro da bolha à medida que o RNA é sintetizado. Os erros na seqüência da cadeia produzidos nesta etapa são cuidadosamente corrigidos pela RNA polimerase. Uma peculiaridade irônica e não freqüente descrita sobre a meticulosidade dessa correção, é que a RNA polimerase ao demorar tanto tempo corrigindo seus erros pode até esquecer de continuar a transcrição.

Terminação:

Muito pouco se conhece sobre o término da transcrição. Sabe-se que todas as três RNAPs terminam a síntese em regiões consenso do DNA ricas em T. Essas regiões são “sinais” codificados pelo DNA, que indicam onde deve ser a extremidade 3' . Esses sinais são reconhecidos por enzimas que se ligam ao RNA, clivando-o para formar a extremidade 3' , ou seja, o sítio de clivagem precede esta região consenso. Após a clivagem, esses nucleotídeos que compunham a região consenso sinalizadora para a clivagem, são degradados no núcleo.

Particularidades da RNAP I e RNAPII:

Tanto a RNAP I quanto a RNAP II possuem mecanismos de transcrição semelhantes ao já descrito para RNAP II. Conhecem-se, porém, algumas propriedades particulares dessas duas enzimas.

Os *promotores* para RNAP I são geralmente compostos por 2 domínios de atividade: um, localizado entre -45 e +20 (domínio central ou *core*) e o outro, localizado entre -107 e -186 (domínio *upstream*). Essa organização é similar em todos os organismos, embora seqüências específicas desses *promotores* sejam variadas. A transcrição necessita de, pelo menos, 3 proteínas ou complexos protéicos além da RNAP I.

Já os *promotores* para RNAP III são bastante diferenciados daqueles até aqui descritos. Existe um tipo de promotor localizado a 5' de +1 que é mais parecido com os envolvidos com a RNAP II. Além desse *promotor*, existem outros 2 elementos que se localizam internamente, após o sítio de início de transcrição e dentro da região codificadora, sendo denominadas *regiões controladoras internas* (ICRs). Os dois tipos de ICRs são *promotores* de tRNA e de genes RNA 5S.

PROCESSAMENTO DE RNA

Introdução e conceitos:

Os diferentes RNAs sintetizados no processo de transcrição são chamados de transcritos primários, RNA precursor ou pré-RNA. Na maioria das vezes, esses transcritos primários não representam uma molécula de RNA madura, ou seja, aquela cuja seqüência e estrutura correspondem à forma final do RNA funcional. Nesses casos, esses transcritos primários precisam sofrer modificações, as quais chamamos de processamento de RNA. O processamento de RNA inclui alterações do tipo adição, deleção ou modificação de nucleotídeos, ou mesmo de regiões maiores do transcrito primário.

Processamento de mRNA:

A molécula de mRNA recém-sintetizada no núcleo precisa sofrer várias alterações bioquímicas. O objetivo é transformar-se no que é chamado de mRNA maduro ou mRNA processado, para só então, ser transportado ao citoplasma e lá ser traduzido. Esses precursores de mRNA são chamados de RNA nuclear heterogêneo (hnRNA – *heterogeneous nuclear RNA*). O processamento desses hnRNAs envolve modificações de nucleotídeos que acabam por aumentar a estabilidade do mRNA e convertê-lo em RNA maduro.

A) *Adição do cap* – A primeira modificação do hnRNA ocorre em sua extremidade 5', logo após o início do processo de transcrição. Um resíduo de guanina é ligado covalentemente ao primeiro nucleotídeo do hnRNA através de uma ligação fosfato 5'-5', com o resíduo de guanina estando na posição inversa a dos demais nucleotídeos. Essa estrutura é chamada de *cap*. O *cap* sofre, então, uma metilação na posição 7 da guanina, resultando no nucleotídeo 7-metilguanilato (m^7G). Todas essas reações são catalisadas por enzimas específicas de cada reação. O *cap* **protege** a extremidade 5' do transcrito da ação de exonucleases. Essas enzimas têm a propriedade de cortar ou clivar ácidos nucléicos, conseqüentemente, moléculas de mRNA sem *cap* são rapidamente degradadas. O *cap* também tem a função de **facilitar o splicing do RNA** e seu **transporte** do núcleo para o citoplasma. O *cap* tem também papel importante na **síntese protéica**.

B) *Adição da cauda de poli A (Poliadenilação)* – A maioria dos mRNAs possui uma seqüência de aproximadamente 200 resíduos de adenina na sua extremidade 3', que é chamada de cauda poli A. Essa cauda poli A não está codificada no DNA e não existe nos rRNAs e tRNAs. Ela é adicionada aos hnRNAs pela enzima poli A-polimerase. Um aspecto comum a todos pré-RNAs mensageiros é a presença da seqüência consenso AAUAAA, 11-30 nucleotídeos antes do sítio de poliadenilação (local onde deve ser adicionado a cauda poli A). Esta seqüência, chamada de seqüência sinal para poliadenilação, é reconhecida por um fator específico, que “marca” o local e permite que ocorra a clivagem por uma endonuclease. A adição da cauda poli A é feita na extremidade 3' OH gerada pela clivagem. Depois que o mRNA poliadenilado chega ao citoplasma, a cauda poli A vai diminuindo, com o decorrer do tempo, provavelmente devido à ação de nucleases. A cauda poli A é uma parte importante do processamento pois facilita o **transporte** do mRNA para o citoplasma e ao ser lentamente degradada, **estabiliza** a molécula no citoplasma. Existe também a possibilidade da cauda poli A estar relacionada com o **reconhecimento** da molécula de mRNA pelo complexo que posteriormente realizará a tradução.

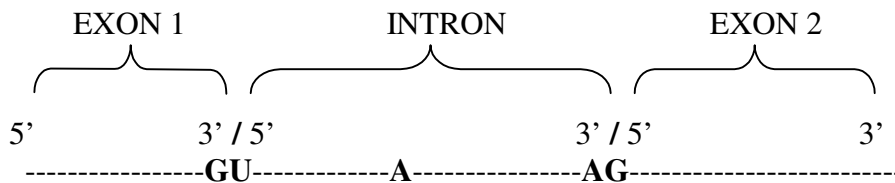
Há um grupo de mRNAs que é transportado ao citoplasma e que não possui cauda de poli A. São os mRNAs de histonas, as proteínas que participam do empacotamento do DNA. Mesmo sem poli A, a extremidade 3' do hnRNA é clivada como descrito anteriormente.

Em geral, a clivagem e a poliadenilação precedem a excisão de introns do hnRNA.

C) *Metilação* – Muitos mRNAs possuem o nitrogênio 6 dos resíduos de adenina metilado. Essa metilação ocorre apenas nos exons, antes da retirada dos introns do hnRNA. Provavelmente, o radical metil serve para proteger as porções do transcrito primário que precisam ser preservadas.

D) *Excisão de introns (splicing)* – A maioria dos transcritos primários de mRNA apresenta dois tipos de seqüências: os exons, porção codificadora do transcrito primário que estão presentes no RNA maduro; e os introns, seqüências que não estão presentes no RNA maduro e, portanto, não codificam nenhum aminoácido da proteína a ser sintetizada. Os introns precisam ser retirados para dar origem a um mRNA funcional (ou traduzível). Assim, após a adição do *cap*, da cauda poli A, e às vezes, a metilação de adeninas, o hnRNA precisa sofrer o processo de excisão dos introns e junção dos exons. Esse mecanismo é conhecido como *splicing* do RNA.

Para que o *splicing* ocorra corretamente, existem seqüências consenso sempre presentes localizadas nos introns dos mRNAs. Essas seqüências têm a finalidade de sinalizar o local onde deverá ser efetuado o *splicing* (chamado de sítio de *splice*) ou auxiliar na sua localização, e têm características importantes: 1) os primeiros e os últimos dois nucleotídeos da extremidade 5' e 3' dos introns são GU e AG, respectivamente; e 2) há uma adenina aproximadamente 18-20 nucleotídeos *upstream* da extremidade 3' do intron. Ambas têm um papel importante no processo de *splicing*. Além dessas seqüências consenso, há uma freqüência maior de algumas bases que estão próximas do sítio de *splicing*, tanto no exon como no intron, e há, também, uma tendência da região próxima à extremidade 3' do intron ser rica em pirimidinas (U e C). Sem essas seqüências consenso, ocorrem alterações no processo de *splicing*. Temos um mRNA no qual estão representados os sítios de *splicing* e a adenina no intron, esquematicamente:



O processo de *splicing* envolve a retirada de cada intron da molécula de hnRNA e união dos exons, liberando o intron. A remoção de introns envolve a quebra de uma ligação fosfodiéster na junção exon-intron, e a formação de outra, entre as extremidades dos exons.

Esse processo pode ocorrer de duas maneiras: mediado por “spliceossomos” e auto -*splicing*.

1- Mediado por “spliceossomos”

Os “spliceossomos” são constituídos pela associação de pequenas ribonucleoproteínas (snRNPs) e outras proteínas acessórias. As snRNPs, por sua vez, são formadas por proteínas associadas a pequenas moléculas de RNA nuclear (snRNAs), ricas em uracil. Algumas destas proteínas são comuns a todas as

snRNPs, outras são específicas. O “spliceossomo” é um complexo grande de aproximadamente 3×10^3 kDa, tendo quase o tamanho de um ribossomo.

Essa estrutura começa a ser montada assim que o intron é transcrito. Cada snRNP associa-se de diferentes maneiras e em diferentes locais do intron, para que o *splicing* ocorra. Este conceito explica porque o processo é tão preciso e específico. Essa associação pode ser auxiliada por pareamento de bases que ocorre entre as seqüências conservadas do intron e os snRNAs que fazem parte das snRNPs.

A primeira clivagem ocorre entre o primeiro exon e a extremidade 5' do intron. O radical OH que promove a clivagem vem da adenina conservada, que existe próximo à região 3' de todo intron. O radical 2' OH dessa adenina quebra a ligação fosfodiéster entre o exon e a extremidade 5' do intron, liberando o exon (primeira transesterificação). O radical 3' OH do exon liberado ataca, então, a ligação fosfodiéster que une a extremidade 3' do intron ao próximo exon, unindo os dois exons (segunda transesterificação). O intron é liberado na forma de laço. Não há alteração no número de ligações fosfodiéster durante o processo e nenhuma energia é necessária. Essas reações são decorrentes de uma cascata de eventos na qual os vários complexos snRNPs ligam-se às regiões pertinentes a serem clivadas. Não se sabe como as clivagens e ligações ocorrem, mas, provavelmente, os snRNAs têm um papel catalítico no processo.

Embora o “spliceossomo” seja montado ainda durante a síntese do RNA, geralmente o intron não é retirado até que a transcrição tenha acabado. O “spliceossomo” impede que o exon a 5' do intron afaste-se do exon a 3' do intron após a primeira clivagem. Após a remoção do intron, o complexo é desfeito. O “spliceossomo” é uma estrutura estável e grande o suficiente para não passar pelos poros nucleares. Talvez por isso, e também por se ligarem logo após a transcrição, os mRNAs não passem ao citoplasma antes de serem processados.

2 - *Auto-splicing*

A excisão de introns pode ocorrer devido à atividade catalítica do próprio RNA precursor, sendo esse processo conhecido como auto-splicing. Neste caso, as mesmas reações descritas anteriormente ocorrem, mas sem a ajuda de spliceossomos. Alguns genes utilizam uma guanina ao invés da adenina para a primeira reação de transesterificação.

Processamento de rRNA:

Existem quatro RNAs ribossômicos com diferentes coeficientes de sedimentação: 5S; 5,8S; 18S e 28S. São codificados por múltiplas cópias de genes, cujo número pode variar de 100-5000 por genoma haplóide, as quais estão localizadas uma ao lado da outra. Os rRNAs 18S, 5,8S e 28S fazem parte de uma unidade de transcrição, ou seja, a RNA-polimerase I produz um único RNA precursor que contém esses rRNAs separados um do outro por espaçadores internos. O transcrito primário apresenta, ainda, nas suas extremidades 5' e 3', espaçadores externos. O transcrito inicial é, então, alvo de clivagens por ribonucleoproteínas e modificações que resultam nos rRNAs maduros que vão constituir o ribossomo.

O processo de maturação ocorre em várias etapas de clivagem em uma ordem preferencial, mas não obrigatória, sendo a primeira delas a remoção dos espaçadores transcritos externos. Depois, há outras duas clivagens que liberam o RNA 18S maduro. As três próximas clivagens liberam o RNA 5,8S, e na maioria dos casos, o 28S maduro, os quais já permanecem ligados por pontes de hidrogênio devido à complementaridade de bases entre esses rRNAs.

O quarto tipo de rRNA, 5S, é transcrito separadamente. Os genes para RNA 5S estão localizados separados dos outros três e são transcritos pela RNA-polimerase III. Também se encontram em cópias múltiplas e separados por espaçadores. A transcrição desses genes não é coordenada com a dos outros RNAs e não se sabe por que eles estão localizados separadamente.

Após o processamento, os rRNAs participam na formação das subunidades do ribossomo e, só então deixam o núcleo.

O RNA ribossômico sofre, ainda, algumas modificações pós-transcricionais, entre elas, metilações. Muitas dessas metilações são extremamente conservadas e algumas delas têm uma função importante no início da síntese de proteínas. A modificação da base U para ψ (ψ) também é comum.

Processamento de tRNA:

Os genes de tRNA são encontrados em cópias múltiplas. Os transcritos primários precisam ser processados para originar o tRNA maduro. Isso envolve a ação de ribonucleases para gerar as extremidades 5' e 3' da molécula madura e a retirada do único intron. O mecanismo de retirada do intron de ~~p~~tRNA ocorre em duas etapas: na primeira, ocorre clivagem endonucleásica nas extremidades 5' e 3' do intron, liberando-o e dividindo o tRNA em duas metades (exons). As duas metades são mantidas juntas devido ao pareamento de bases do tRNA em regiões complementares, apesar de não estarem ainda ligadas covalentemente. Na segunda etapa ocorre a ligação dos dois exons para a formação do tRNA maduro. A ligação dos dois exons é feita por uma RNA-ligase.

Aparentemente não há seqüências consenso no intron, essenciais ao processo de *splicing*. As seqüências do exon e a conseqüente estrutura tridimensional do pré-tRNA é que são os elementos importantes para o reconhecimento e para a correta clivagem pelas endonucleases.

As moléculas de tRNA maduras contêm um grande número de nucleotídeos raros originados pela modificação pós-transcricional das bases comuns (A,U,C,G). A maioria das modificações envolve metilação e/ou tiolação e dependem de cisteína, metionina e treonina, que são as fontes dos grupos utilizados na alteração das bases. Outro processamento importante é a adição da seqüência CCA na extremidade 3' por uma enzima específica. Todos os tRNA têm essa seqüência.