

## TRASCIPCIÓN GENÉTICA

*Pablo J. Patiño, MD, MSc, Dr Sci.*

*Profesor Asociado, Grupo Inmunodeficiencias Primarias - Corporación Biogénesis  
Facultad de Medicina, Universidad de Antioquia*

Como se ha mencionado en capítulos anteriores, uno de los aspectos centrales del estudio de la biología molecular lo constituye el hecho de que la estructura y función de una célula son determinadas en gran medida por las proteínas que ella misma sintetiza. La información para la producción de estas proteínas está codificada en el ADN: pero, ¿cómo se convierte esta información en un polímero de aminoácidos? El dogma central de la biología molecular establece que el ADN con la información relevante se transcribe a un intermediario que es el RNA, el cual se traduce posteriormente a la proteína esperada (Figura 1).

¿Qué ventajas le confiere a la célula tener una molécula como el ADN para almacenar la información de las proteínas celulares?

- El ADN puede permanecer protegido en el núcleo, fuera de contacto con un medio químicamente poco amistoso, lo que limita la posibilidad de cambio (mutación) de la información.
- La información genética puede ser amplificada enormemente, pues se pueden generar muchas moléculas de un RNA determinado a partir de una sola copia de ADN.
- Se incrementan las posibilidades de regular la expresión de un gen, pues existen varios elementos que participan de este proceso en diferentes puntos, todos sometidos a un riguroso control.

Para que la información almacenada en el ADN, gracias a la combinación de cuatro nucleótidos, pueda ser convertida en proteínas debe, entonces, existir un código genético que permita la transformación en proteínas diferentes utilizando los 20 aminoácidos disponibles. Después de varios análisis se concluyó que cada aminoácido debería ser especificado por una combinación de al menos 3 nucleótidos, dicha combinación recibió el nombre de codón y esto constituye la base de lo que conocemos como código genético (Figura 2). De esta manera existen 64 posibles combinaciones diferentes de los cuatro nucleótidos para codificar 20 aminoácidos, por lo tanto la mayoría de los aminoácidos son codificados por más de un codón, esto es lo que se ha denominado degeneración o redundancia del código genético. Por ejemplo, algunos aminoácidos son codificados por cualquiera de seis codones distintos, además existen tres codones que no codifican para algún aminoácido.

El proceso de convertir el código genético almacenado en el ADN en una proteína no es un simple copiar el ADN en RNA, pues además de una serie de proteínas de reconocimiento y de enzimas, se requieren de al menos 3 tipos diferentes de RNA. El RNA mensajero (RNAm) que se transcribe a partir del ADN y es procesado hasta contener solo la información del gen que se traducirá en proteína. El RNA de transferencia (RNAt) que corresponde a

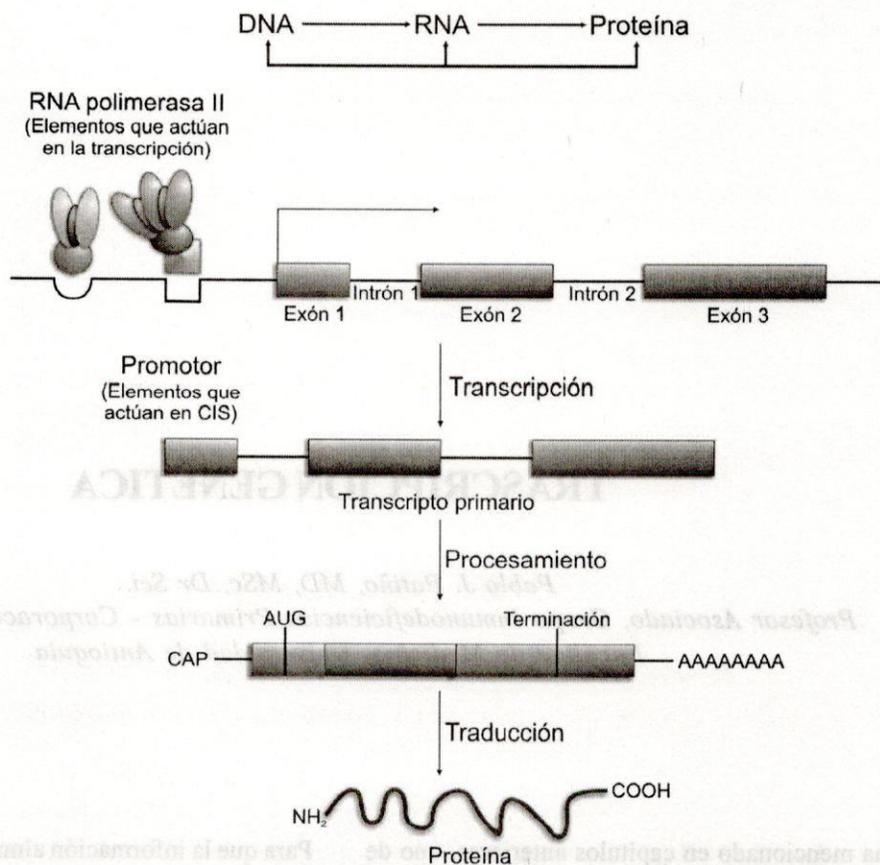


Figura 1. El dogma central de la biología molecular establece que la información para la estructura y función de los organismos se encuentra codificada en el ADN, la cual se transcribe a una molécula intermediaria que es el RNA, que posteriormente se traduce para generar secuencias de polipéptidos que dan origen a proteínas que cumplen una determinada función. En esta figura se esquematiza el proceso que ocurre en las células eucarióticas, en las cuales existe un mayor grado de complejidad que en las procariontas.

**Primera Posición      Segunda Posición      Tercera Posición**

	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Para	Para	A
	Leu	Ser	Para	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

Figura 2. El código genético se refiere a la información de los aminoácidos que constituyen las proteínas, la cual está codificada en una combinación de al menos 3 nucleótidos (tanto en el ADN como en el RNA), combinación que se denomina codón. Existen 64 posibles combinaciones diferentes de los cuatro nucleótidos que conforman los ácidos nucleico para codificar 20 aminoácidos, esto hace que una gran parte de los aminoácidos estén codificados por más de un codón. En esta figura se presentan los codones con base en la codificación del RNA.

diferentes moléculas de RNA que tienen la capacidad de transportar de manera específica los aminoácidos para la traducción de la proteína o sea que cada aminoácido tiene su propio RNAt. El RNA ribosomal (RNAr) el cual conforma los ribosomas, factorías donde ocurre la síntesis proteica y en las cuales este RNAr tiene un papel central en el proceso de polimerización de los aminoácidos. Además otras moléculas pequeñas de RNA son necesarias para el procesamiento del RNA y localización de proteínas en células eucariotas.

La transcripción es un fenómeno bastante similar al de la replicación del ADN, en el cual están involucrados diferentes factores y enzimas. Tal vez el más importante de estos factores es la RNA polimerasa, una holoenzima que está constituida por la aglomeración de factores diversos que permiten la síntesis del RNA a partir de un molde de ADN. Al igual que ocurre con la polimerización del ADN, el crecimiento del polímero de RNA es en sentido 5'→3'. Sin embargo, a diferencia de las ADN polimerasas, la RNA polimerasa no requiere de una secuencia iniciadora para comenzar la síntesis de RNA.

### Transcripción en procariotes

Al igual que ha ocurrido con el proceso de replicación del ADN, los mecanismos de la transcripción genética se han dilucidado inicialmente gracias a los estudios en el modelo de la bacteria *Escherichia coli* y luego se ha pasado a estudiar las características de este fenómeno en células eucariotas.

Una diferencia importante entre procariotes y eucariotes tiene que ver con la estructura de los genes y la forma como ellos se transcriben. Muchos de los genes procariotes son policistrónicos, lo cual significa que varios genes, cada uno de los cuales codifica para una proteína diferente (cistrón), se encuentran formando una unidad funcional que se transcribe como una sola molécula de RNAm que posteriormente se traduce en varias proteínas (Figura 3A). Por su parte, en las células eucarióticas cada gen constituye una sola unidad que da origen a un RNAm que se traduce posteriormente a una sola proteína (Figura 3B). Otra diferencia fundamental entre estos organismos y que se encuentra íntimamente relacionada con la transcripción, es el hecho que el RNAm de los procariotes se une rápidamente a los ribosomas (no existe envoltura nuclear) y por lo tanto se traduce a proteína aún antes de terminar la transcripción; mientras que el RNAm de los eucariotas debe sufrir una serie de modificaciones en el núcleo antes de ser exportado al citoplasma donde se une a los ribosomas para su traducción.

La RNA polimerasa de *E. coli* está formada por 4 subunidades diferentes que se denominan  $\alpha$ ,  $\beta$ ,  $\beta'$  y  $\sigma$  (Figura 4). Las 3 primeras subunidades constituyen el centro de la holoenzima y contienen la capacidad enzimática completa, por su parte la subunidad  $\sigma$  permite identificar el sitio correcto de inicio de la transcripción, por lo tanto esta subunidad es fundamental para la transcripción específica de los genes.

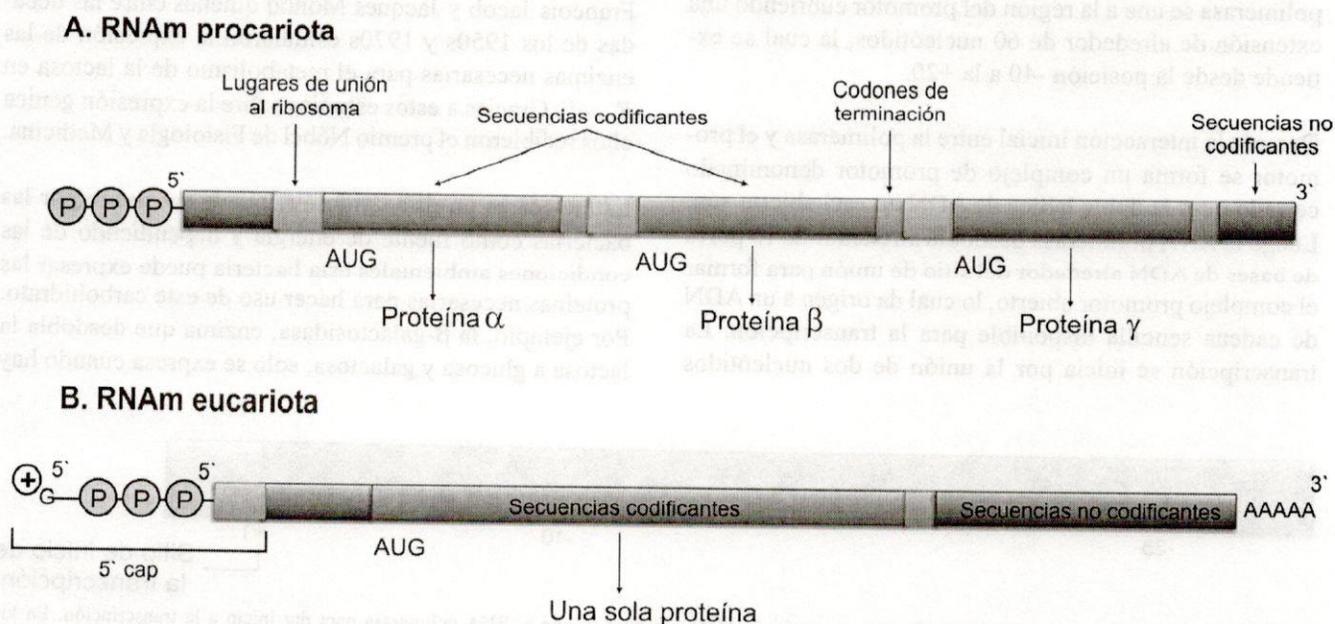


Figura 3. A. Una gran proporción de los genes en los organismos procariotes son policistrónicos: varios genes (cistrones), se encuentran formando una unidad funcional que se transcribe como una sola molécula de RNAm que posteriormente se traduce en proteínas independientes. B. En las células eucarióticas la estructura de los genes es diferente, cada gen constituye una unidad que origina un RNAm que luego se traduce a una sola proteína; además este RNAm sufre modificaciones postranscripcionales en sus extremos 5' y 3' que le confieren propiedades adicionales.

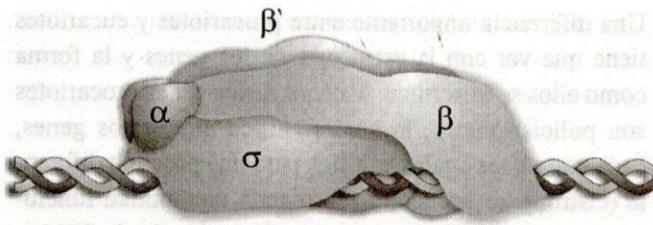


Figura 4. La RNA polimerasa de los organismos procariotes (*E. coli* en este caso) está formada por 4 subunidades diferentes que se denominan  $\alpha$ ,  $\beta$ ,  $\beta'$  y  $\sigma$ , cuya combinación constituye la holoenzima funcional.

Cada gen tiene una secuencia específica de ADN a la cual se une la RNA polimerasa para dar inicio a la transcripción, dichas secuencias se denominan promotores. El análisis de varios genes aislados de *E. coli* permitió identificar dos secuencias en la región 5' (corriente arriba) antes del sitio de inicio de la transcripción. Cada una de estas 2 secuencias tiene 6 nucleótidos y están ubicadas 10 y 35 nucleótidos antes del inicio de la transcripción, por lo que se conocen como los elementos -10 y -35, respectivamente (Figura 5). Aunque estas secuencias no son idénticas en todos los genes, ellas presentan suficiente similitud para establecer la presencia de secuencias consenso. La subunidad  $\sigma$  de la RNA polimerasa se une específicamente a las secuencias -35 y -10 del promotor y de esta manera se señala específicamente el inicio de la transcripción. Además, en algunos promotores bacterianos se encuentra una secuencia adicional corriente arriba a la secuencia -35 que permite la unión específica de la subunidad  $\alpha$  de la RNA polimerasa. En total, la RNA polimerasa se une a la región del promotor cubriendo una extensión de alrededor de 60 nucleótidos, la cual se extiende desde la posición -40 a la +20.

Durante la interacción inicial entre la polimerasa y el promotor se forma un complejo de promotor denominado cerrado pues la doble hélice de ADN no está abierta aún. Luego la RNA polimerasa desdobra alrededor de 15 pares de bases de ADN alrededor del sitio de unión para formar el complejo promotor abierto, lo cual da origen a un ADN de cadena sencilla disponible para la transcripción. La transcripción se inicia por la unión de dos nucleótidos

libres (Figura 6). Después de la polimerización de los primeros 10 nucleótidos, la subunidad  $\sigma$  se libera de la polimerasa permitiendo que la enzima se desplace a lo largo del molde de ADN para continuar la síntesis de RNA. A medida que la polimerasa se desplace la doble hélice de ADN se desdobra adelante y se aparea de nuevo detrás de ella, lo que permite mantener abierta una región de unas 17 pares de bases de ADN (Figura 7).

La síntesis de RNA continua hasta que la polimerasa encuentra una señal de terminación, la cual detiene la transcripción, el RNA se libera de la polimerasa y la enzima se disocia del ADN. La señal más común para terminar la replicación en los genes de *E. coli* consiste de una secuencia repetitiva simétrica invertida rica en GC seguida por cuatro nucleótidos de A (Figura 8). La transcripción de la secuencia simétrica invertida conduce a la síntesis de una secuencia de RNA que puede formar una estructura en forma de asa gracias al apareamiento de las bases complementarias dentro de la misma cadena. La formación de esta estructura en el RNA desestabiliza su interacción con el molde de ADN y termina la transcripción. En otros sistemas, tanto de procariotes como de eucariotes, la terminación de la transcripción depende de la unión de proteínas específicas a secuencias definidas de ADN.

#### Control de la transcripción en procariotes

La expresión génica es un fenómeno bastante complejo y su conocimiento inicial se logró gracias al trabajo de Francois Jacob y Jacques Monod quienes entre las décadas de los 1950s y 1970s estudiaron la expresión de las enzimas necesarias para el metabolismo de la lactosa en *E. coli*. Gracias a estos estudios sobre la expresión génica ellos recibieron el premio Nóbel de Fisiología y Medicina.

La lactosa es un disacárido que puede ser usado por las bacterias como fuente de energía y dependiendo de las condiciones ambientales esta bacteria puede expresar las proteínas necesarias para hacer uso de este carbohidrato. Por ejemplo, la  $\beta$ -galactosidasa, enzima que desdobra la lactosa a glucosa y galactosa, solo se expresa cuando hay

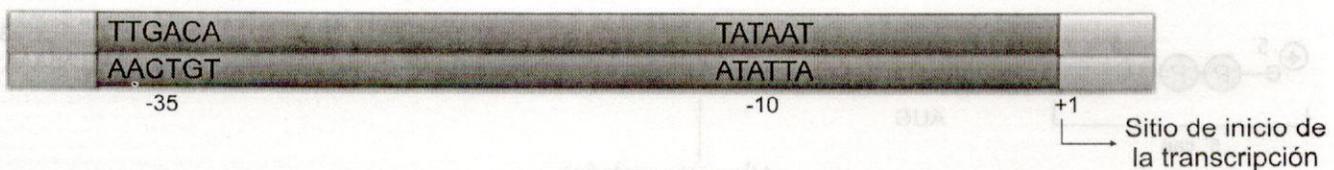


Figura 5. El promotor de un gen corresponde una secuencia de ADN a la cual se une la RNA polimerasa para dar inicio a la transcripción. En los procariotes se han identificado dos secuencias altamente conservadas (consenso) en la región 5' (corriente arriba) antes del sitio de inicio de la transcripción. Cada una de estas 2 secuencias tiene 6 nucleótidos y están ubicadas 10 y 35 nucleótidos antes del inicio de la transcripción, por lo que se conocen como los elementos -10 y -35, respectivamente. El elemento -10 se ha denominado caja TATA debido a que esta secuencia de nucleótidos es altamente conservada.

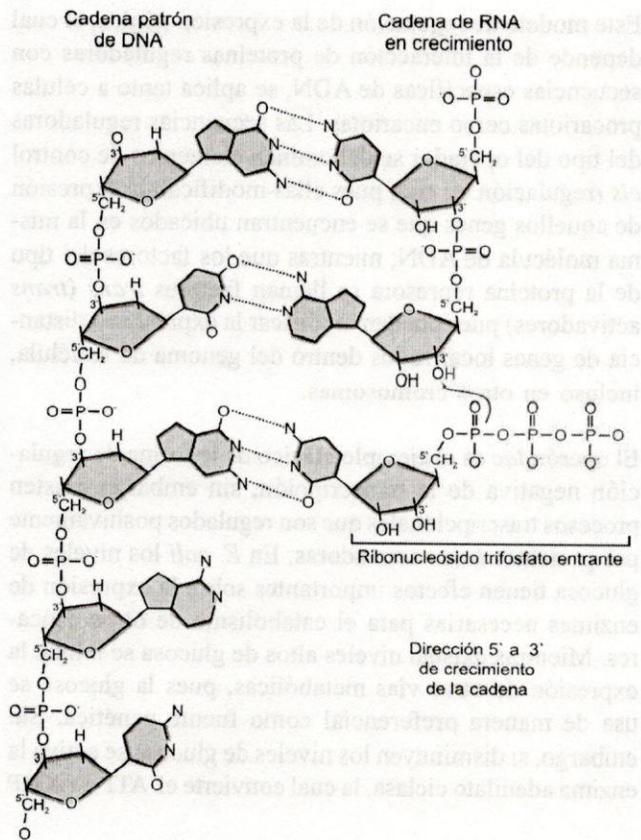


Figura 6. La transcripción se inicia por la unión de dos nucleótidos libres que realiza la RNA polimerasa en el sitio de inicio de la transcripción, los cuales son complementarios a la cadena de ADN usada como molde. Esta polimerización ocurre de manera similar a la del ADN se adicionan ribonucleótidos al grupo OH en el carbono 3' de la ribosa, o sea en dirección 5' → 3'.

lactosa presente para ser usada por la bacteria. Esto significa que la lactosa puede inducir la síntesis de las proteínas involucradas en su propio metabolismo. Las otras enzimas que se expresan junto con la  $\beta$ -galactosidasa son la permeasa de lactosa, la cual transporta lactosa a la célula y una transacetilasa cuya función no es conocida. Estos tres genes se transcriben en una sola molécula de RNA mensajero como un RNA policistrónico.

La expresión de los tres genes involucrados en el metabolismo de la lactosa depende de una unidad funcional denominada *operón*, en este caso se conforma el operón de lactosa (operón *lac*) (Figura 9). La transcripción del RNAm de este operón *lac* es controlada por una secuencia adyacente al sitio de inicio de la transcripción que se conoce como operador (*o*). Este operador consiste en una secuencia de ADN que pueden unir una proteína inhibidora de la transcripción que a su vez es codificada por el gen *i*, el cual se encuentra corriente arriba al operón. La proteína codificada por el gen *i* es por lo tanto un represor, que bloquea la transcripción cuando se une a la secuencia *o*. Cuando se adiciona lactosa al medio donde vive la bacteria, se induce la expresión del operón debido a que este disacárido es capaz de unirse al represor *i*, lo cual produce un cambio conformacional en esta proteína que le impida unirse al operador en el ADN. Este operador es una secuencia que contiene alrededor de 30 pb, que se ubica pocos nucleótidos corriente arriba del sitio de inicio de la transcripción.

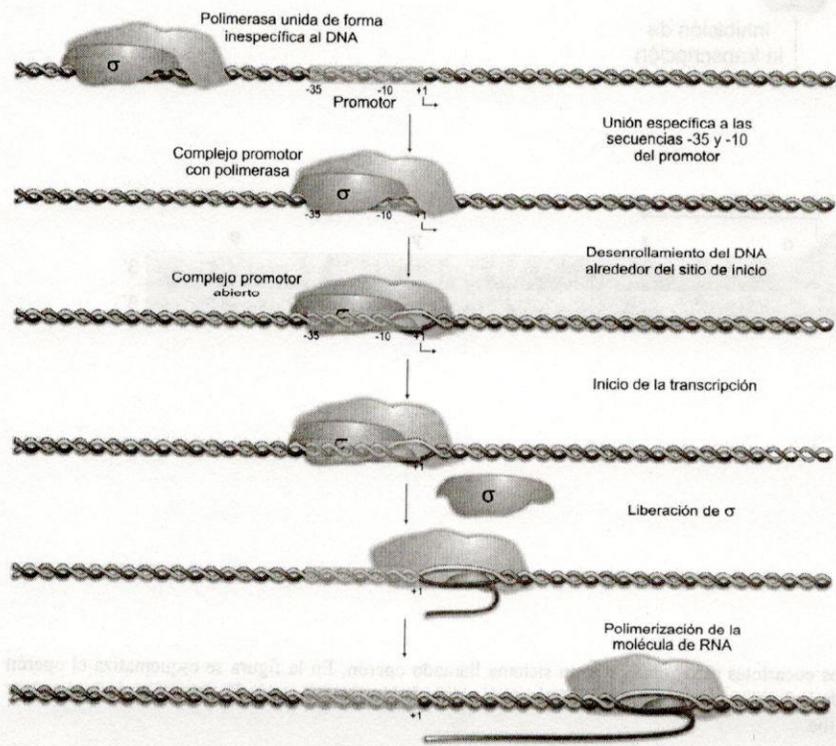


Figura 7. En esta figura se muestra el esquema general de la transcripción en procariotes. La RNA polimerasa reconoce la región promotora del gen e inicia la adición de nucleótidos en el sitio +1. Después de la polimerización de los primeros 10 nucleótidos, la subunidad  $\sigma$  se libera de la polimerasa permitiendo que la enzima se desplace a lo largo del molde de ADN para continuar la síntesis de RNA. A medida que la polimerasa se desplace la doble hélice de ADN se desdobra adelante y se aparea de nuevo detrás de ella, lo que permite mantener abierta una región de unas 17 pares de bases de ADN.

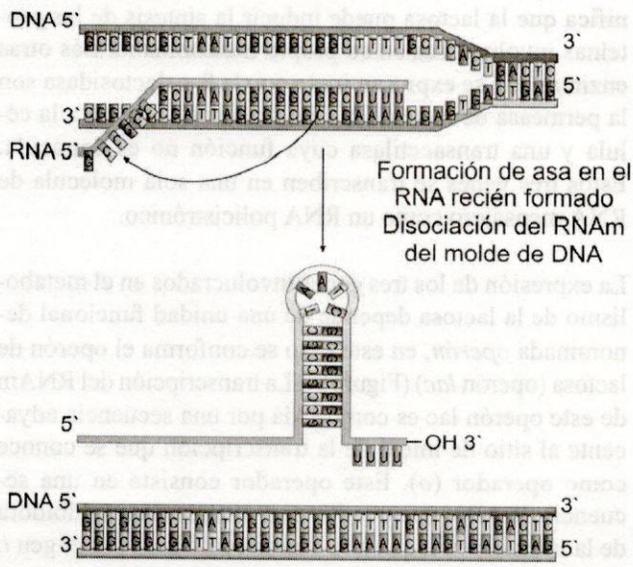


Figura 8. La señal más común para terminar la replicación en los genes procariotes consiste de una secuencia repetitiva simétrica invertida rica en GC; esta secuencia permite la formación de una estructura secundaria del RNA en forma de asa, la cual desestabiliza la RNA polimerasa haciendo que se desprenda del molde de ADN.

Este modelo de regulación de la expresión génica, el cual depende de la interacción de proteínas reguladoras con secuencias específicas de ADN, se aplica tanto a células procariotas como eucariotas. Las secuencias reguladoras del tipo del operador se denominan elementos de control *cis* (regulación en *cis*), pues ellas modifican la expresión de aquellos genes que se encuentran ubicados en la misma molécula de ADN; mientras que los factores del tipo de la proteína represora se llaman factores *trans* (*trans* activadores) pues pueden modificar la expresión a distancia de genes localizados dentro del genoma de la célula, incluso en otros cromosomas.

El operón *lac* es el ejemplo clásico de la forma de regulación negativa de la transcripción, sin embargo existen procesos transcripcionales que son regulados positivamente por proteínas *trans* activadoras. En *E. coli* los niveles de glucosa tienen efectos importantes sobre la expresión de enzimas necesarias para el catabolismo de otros azúcares. Mientras existan niveles altos de glucosa se inhibe la expresión de otras vías metabólicas, pues la glucosa se usa de manera preferencial como fuente energética. Sin embargo, si disminuyen los niveles de glucosa se activa la enzima adenilato ciclasa, la cual convierte el ATP en AMP

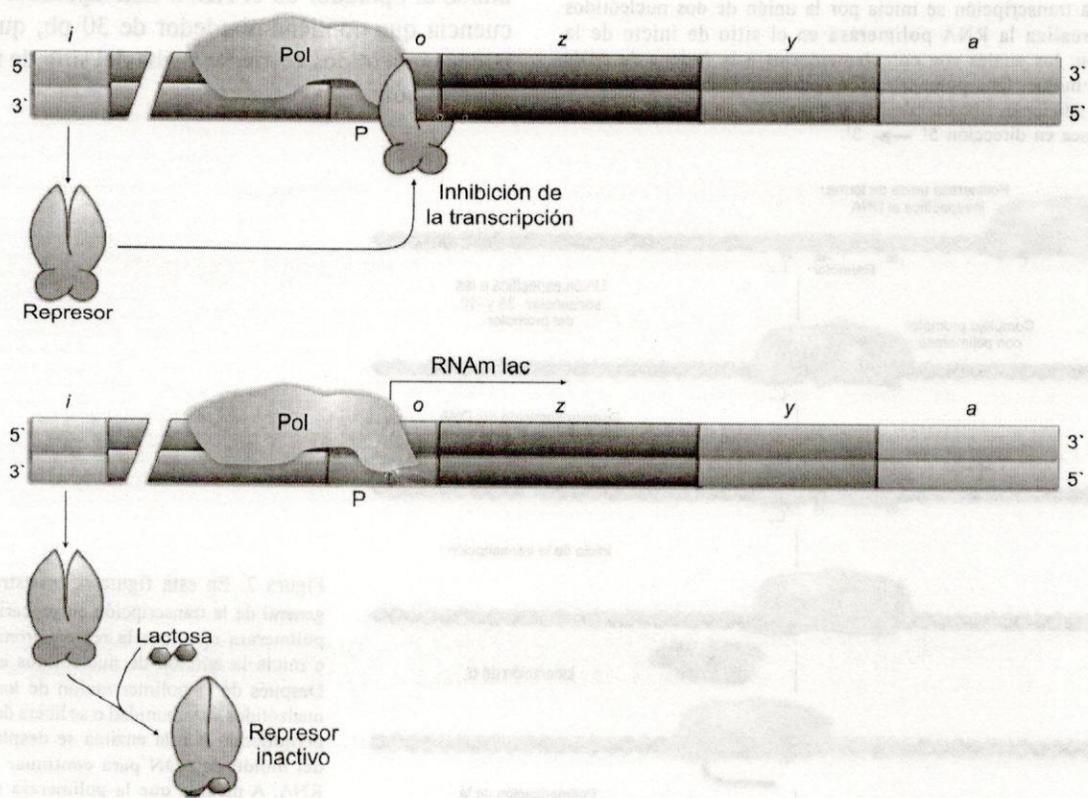


Figura 9. La expresión de los genes policistrónicos en los eucariotes es regulada por un sistema llamado operón. En la figura se esquematiza el operón lactosa (*lac*), el cual se encuentra inactivo en ausencia de lactosa lo que permite la síntesis de las proteínas necesarias para la obtención de este azúcar del medio. En presencia de la lactosa el operón se inhibe.

cíclico (AMPc); a su vez este AMPc puede unirse a una proteína reguladora de la transcripción conocida como proteína activadora de catabolito (CAP) (Figura 10). La unión de AMPc a CAP induce un cambio en su conformación que permite su interacción con una secuencia específica de ADN, la cual se encuentra unos 60 nucleótidos antes del sitio de inicio de la transcripción en el operón *lac*. Una vez unida al ADN, la CAP interactúa con la subunidad  $\alpha$  de la RNA polimerasa facilitando la unión de esta polimerasa al promotor, lo que finalmente activa la transcripción del operón *lac*.

Existe un tercer mecanismo de regulación de la expresión de genes bacterianos denominado atenuación transcripcional. Este mecanismo permite controlar la capacidad de la RNA polimerasa para continuar la elongación del RNA después de pasar por sitios específicos en el ADN. El mecanismo de atenuación se ha descrito en la regulación de los genes involucrados en el metabolismo del triptófano, los cuales constituyen el operón *trp* (Figura 11). Este operón se expresa solo cuando la bacteria no tiene una fuente externa de triptófano y se inhibe por medio de la unión de este aminoácido a un represor que bloquea la transcripción de este operón (regulación negativa clásica). Sin embargo, la atenuación de la transcripción introduce un mayor control de la transcripción de este operón. En la posición correspondiente al nucleótido 162 corriente abajo del sitio de inicio de la transcripción existe una secuencia

atenuadora que en presencia de niveles altos de triptófano induce la terminación de la transcripción. Este mecanismo de atenuación depende del hecho de que la traducción proteica bacteriana está acoplada a la transcripción (los ribosomas inician la traducción aún cuando el RNAm se está sintetizando). Cuando las células tienen niveles elevados de triptófano, la traducción del RNAm del operón *trp* procede a una velocidad normal de tal manera que los ribosomas van cerca del sitio de transcripción; esto permite que se forme una estructura en forma de asa en una región del RNAm cercana a la RNA pol que actúa como una señal de terminación de la transcripción y hace que se desprenda la polimerasa. Sin embargo, en presencia de niveles bajos de triptófano, los ribosomas se retrasan al inicio del RNAm, lo cual permite que una región diferente, mucho más cercana al extremo 5' del RNAm, forme la estructura en asa que ya no actúa como terminador de transcripción y por lo tanto ésta no es atenuada.

### Transcripción en eucariotes

Aunque la transcripción de genes ocurre por un mecanismo muy similar en todos los organismos, en las células eucariotas es mucho más complejo. A diferencia de los procariotes, organismos en los que todos los genes son transcritos por un solo tipo de RNA pol, en las células eucariotas existen varias isoformas de esta enzima las cuales transcriben diferentes clases de genes. Otra diferencia significativa consiste en que las RNA polimerasas de los eucariotes necesitan interactuar con varias proteínas para poder dar inicio a la transcripción.

En las células eucariotas existen tres tipos de RNA polimerasas (Tabla 1). La RNA pol I transcribe de manera

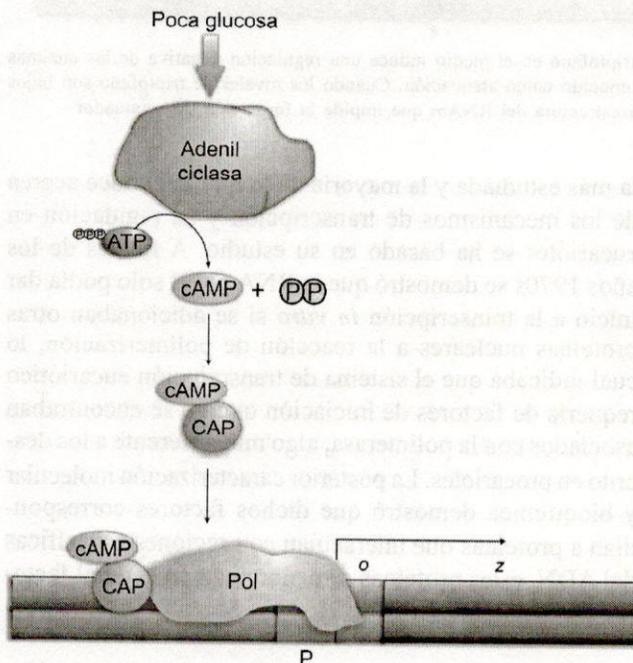


Figura 10. Control positivo del operón *lac*. Cuando disminuyen los niveles glucosa se activa la adenilato ciclasa y produce AMPc. Este a su vez activa a un factor activador de la transcripción específico para el operon *lac* (CAP).

Tipo de RNA sintetizado	RNA polimerasa
<u>Genes del núcleo</u>	
mRNA	II
tRNA	III
rRNA	
5.8S, 18S, 28S	I
5S	III
snRNA y scRNA	II y III
<u>Gen mitocondrial</u>	Mitocondrial
<u>Gen del cloroplasto</u>	Cloroplasto

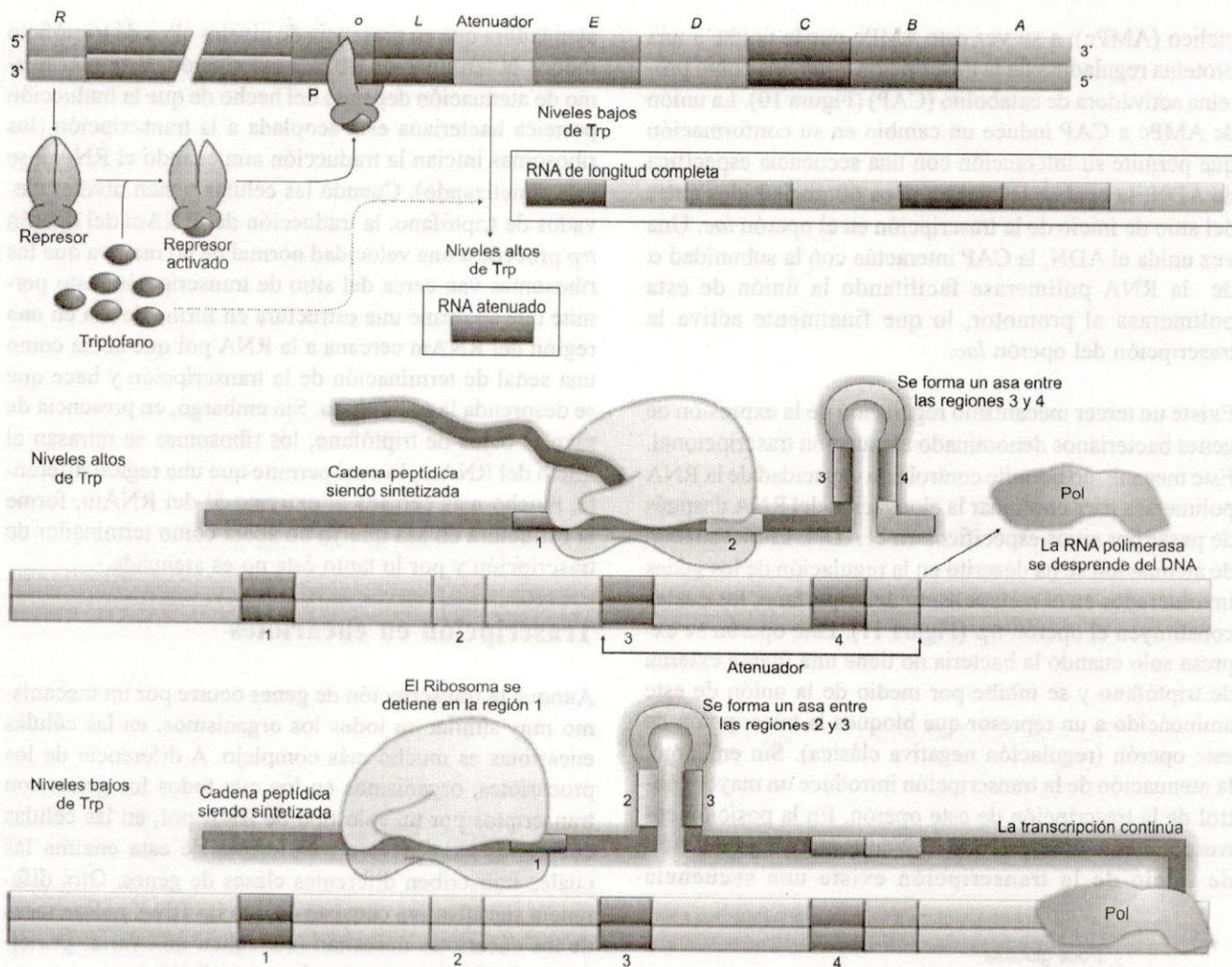


Figura 11. Control negativo del operón *trp* por atenuación. La presencia de triptofano en el medio induce una regulación negativa de las enzimas necesarias para la síntesis de este aminoácido, por medio de un mecanismo conocido como atenuación. Cuando los niveles de triptofano son bajos los ribosomas se demoran al inicio del RNAm, permitiendo que se forme una estructura del RNAm que impide la formación del atenuador

específica las especies más grandes del RNA ribosomal denominadas 28S, 18S y 5.8S. Por su parte la RNA pol II genera los RNAm de los genes que codifican proteínas. Mientras que la RNA pol III transcribe los genes de los RNAs de transferencia y la especie más pequeña del RNA ribosomal (5S). Existen otros RNAs pequeños que participan en el procesamiento del RNA y en el transporte de proteínas; estos RNAs se transcriben por las RNA pol II y III. Además en mitocondrias y cloroplastos se encuentran las RNA polimerasas propias de estas organelas, las cuales son similares a las RNA polimerasas bacterianas. Las tres RNA polimerasas nucleares son holoenzimas que constan de 8 a 14 distintas subunidades. Las 2 subunidades más grandes de estas RNA polimerasas son homólogas a las subunidades  $\beta$  y  $\beta'$  de la RNA pol bacteriana.

Puesto que la RNA pol II es la responsable de la transcripción de los genes que codifican las proteínas, ha sido

la más estudiada y la mayoría de lo que se conoce acerca de los mecanismos de transcripción y su regulación en eucariotes se ha basado en su estudio. A finales de los años 1970s se demostró que la RNA pol II solo podía dar inicio a la transcripción *in vitro* si se adicionaban otras proteínas nucleares a la reacción de polimerización, lo cual indicaba que el sistema de transcripción eucariótico requería de factores de iniciación que no se encontraban asociados con la polimerasa, algo muy diferente a los descritos en procariotes. La posterior caracterización molecular y bioquímica demostró que dichos factores correspondían a proteínas que interactúan con regiones específicas del ADN, estas proteínas se denominan en general factores de transcripción.

Se han descrito dos tipos de factores de transcripción. Los factores de transcripción generales tienen un papel central en la transcripción de todos los genes que se transcriben por las RNA polimerasas eucarióticas II y por

lo tanto hacen parte de la maquinaria básica de la transcripción, tanto en genes constitutivos como regulados específicamente. Por su parte, los factores de transcripción específicos se unen a secuencias de ADN que controlan la expresión de genes individuales por la RNA pol II y por lo tanto son responsables de regular la expresión génica de manera específica de tejidos o de estadio de diferenciación celular.

Muchos genes que se transcriben por la RNA pol II tienen en su promotor una secuencia de ADN similar a TATAA, la cual se ubica 25 a 30 nucleótidos corriente arriba del sitio de inicio de la transcripción (Figura 12). Esta secuencia es homóloga a la caja TATA que se encuentra en la posición nucleotídica -10 de los promotores bacterianos y es fundamental para el inicio de la transcripción; sin embargo, muchos otros genes no tienen esta secuencia consenso y la transcripción se inicia mediante el reconocimiento de otras secuencias.

El primer evento en la transcripción eucariótica es la unión de un factor general de transcripción denominado TFIID (factor D de transcripción para polimerasa II) a la caja TATAA (Figura 13). Este TFIID está constituido por varias subunidades peptídicas, las cuales incluyen a la pro-

teína unidora de TATA (TBP) (Figura 14) que se une a la secuencia TATAA y por otros 10 a 12 polipéptidos llamados factores asociados a TBP (TAFs). Luego de su unión a la cap TATA, la TBP se une a un segundo factor de transcripción general conocido como TFIIB, lo que conforma el complejo TBP-TFIIB sobre el promotor. A su vez el TFIIB actúa como un puente para la RNA polimerasa, la cual se une al complejo TBP-TFIIB en asociación con un tercer factor llamado TFIIF. Después de estas interacciones se requiere la unión de TFIIE y TFIIF a la polimerasa para poder dar inicio a la transcripción. El TFIIF tiene 2 funciones importantes: dos de sus subunidades son helicasas necesarias para desdoblarse el ADN en el sitio de inicio de transcripción, mientras que otra subunidad tiene actividad de proteína quinasa que fosforila ciertos aminoácidos presentes en el dominio carboxi terminal de la RNA polimerasa II; esta fosforilación facilita su disociación del complejo de iniciación y permite el desplazamiento de la enzima a lo largo del molde de ADN para poder sintetizar la cadena de RNA.

Como se había mencionado antes, en los promotores de muchos genes que son transcritos por la RNA pol II no existe una secuencia consenso de caja TATA, sino una secuencia iniciadora (*Inr*) que involucra el sitio de iniciación

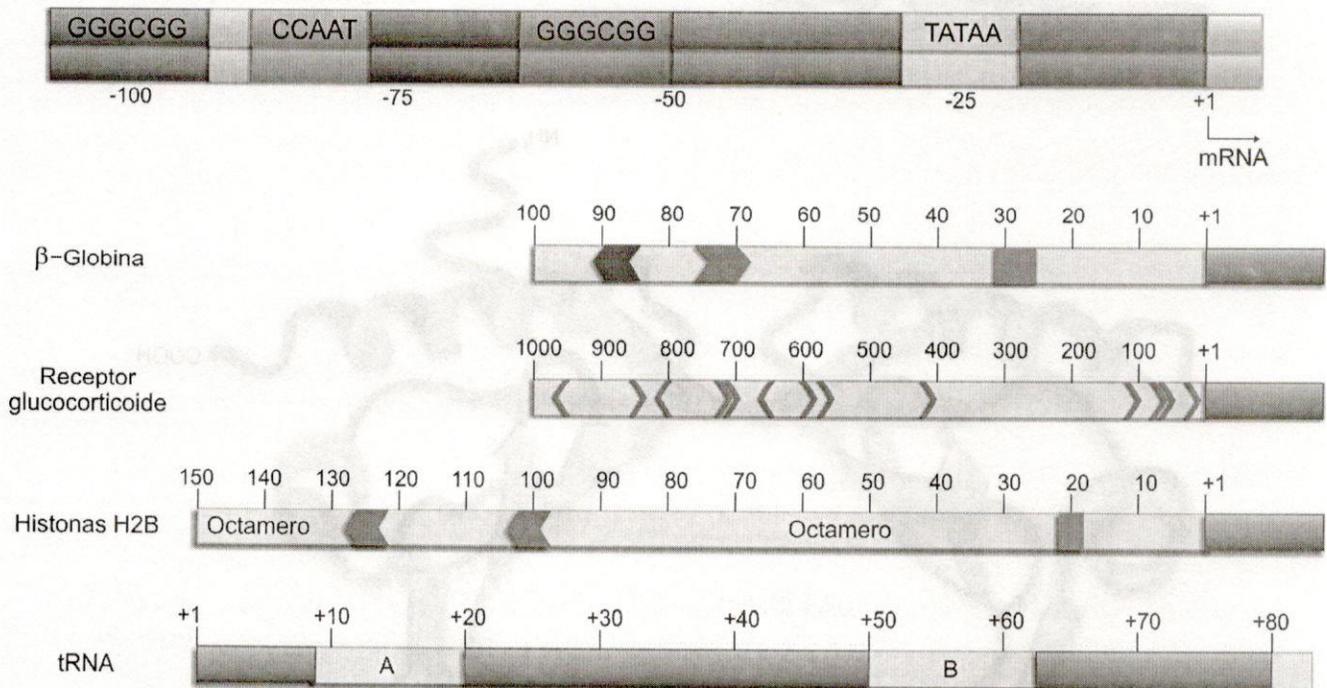
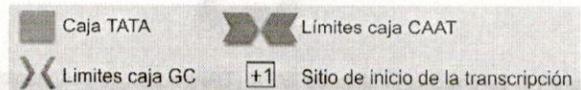


Figura 12. Promotor prototipo de un gen eucariótico. Para la transcripción adecuada de los genes en las células eucarióticas se requiere además de la caja TATA varios elementos o secuencias consenso en la región promotora, como por ejemplo las cajas GC y la caja CCAAT. En esta figura se esquematizan diferentes genes que poseen distintas secuencias reguladoras en su región promotora.



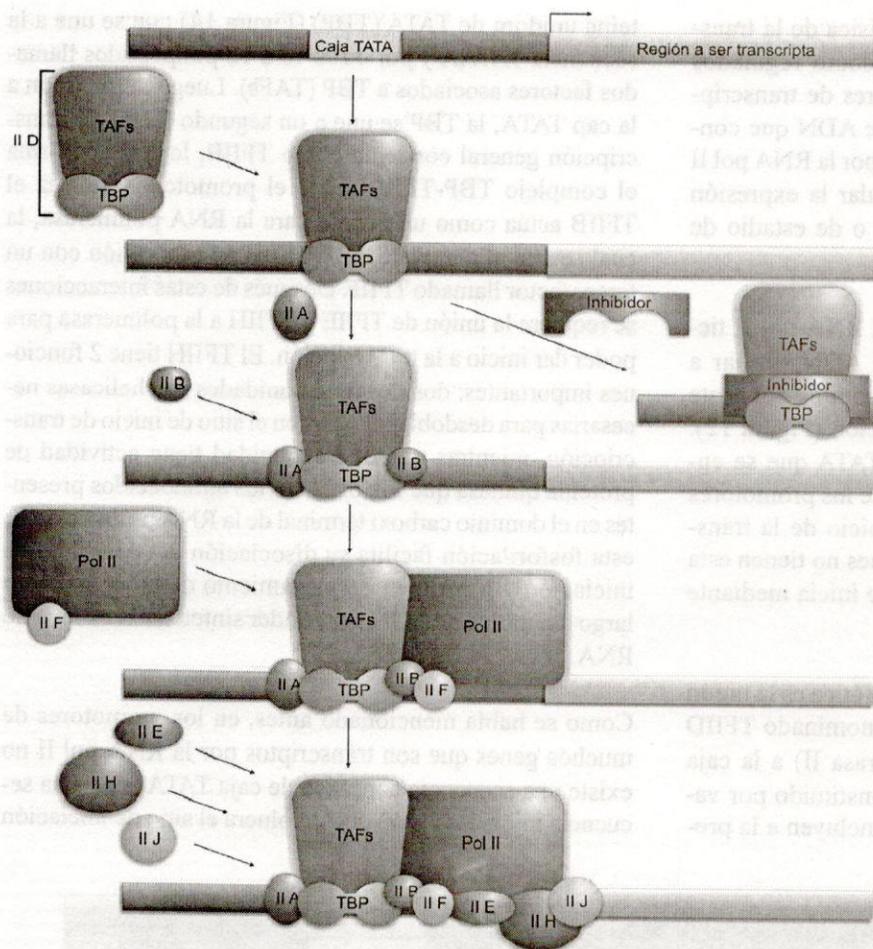


Figura 13. El primer evento en la transcripción eucariótica es la unión de un factor general de transcripción denominado TFIID (factor D de transcripción para polimerasa II) junto con la proteína unidora de TATA (TBP) a la caja TATA, lo cual desencadena el reclutamiento de otros polipéptidos llamados factores asociados a TBP (TAFs), todo lo cual en última instancia permite la unión de la RNA pol II a la región promotora y así iniciar la transcripción. Existen proteínas que intervienen en la interacción entre TFIID y TBP y de esta manera pueden inhibir el inicio de este proceso.

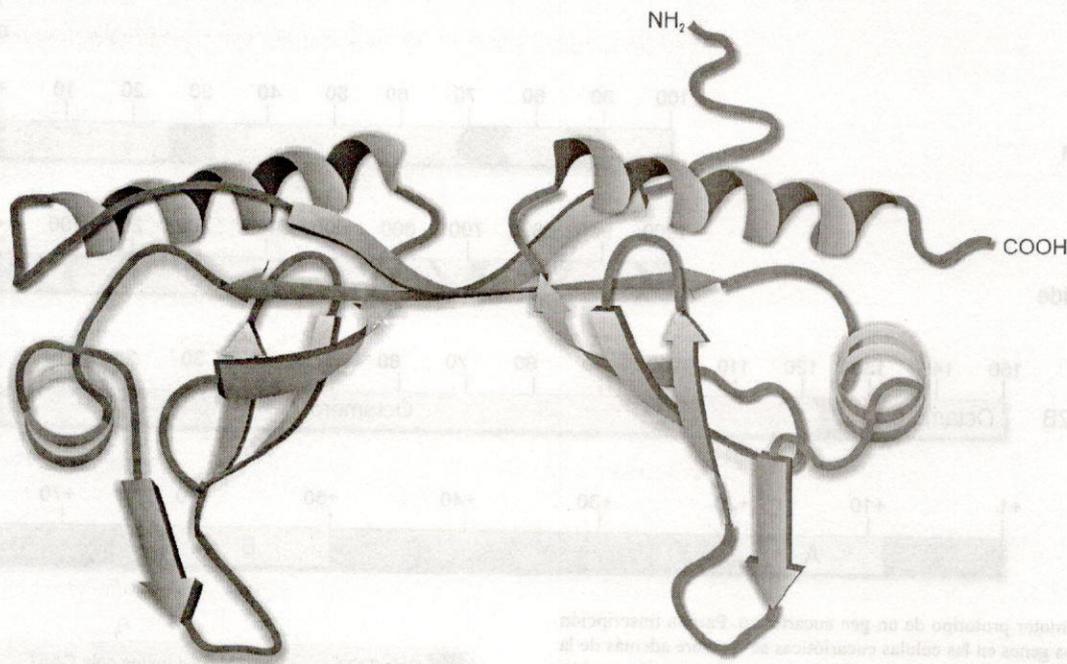


Figura 14. La proteína unidora de TATA (TBP) tiene una configuración tridimensional en forma de silla de montar que le permite unirse de manera específica a regiones específicas en el ADN.

de la transcripción. Sin embargo, el inicio de la transcripción en estos promotores también requiere de TFIID y TBP, así esta última proteína no tenga la secuencia TATA para interactuar. Parece ser que otras subunidades de TFIID se unen a la secuencia *Inr*, lo cual atrae TBP hacia el promotor permitiendo la interacción de TFIIB, RNA pol II y los demás factores de transcripción que se describieron antes. De esta manera TBP continúa teniendo un papel central en el inicio de la transcripción por la RNA pol II, aun en ausencia de una secuencia consenso TATA.

De igual manera y aunque las RNA polimerasas I y III de las células eucarióticas reconocen tipos diferentes de regiones promotoras, la proteína unidora de TATA también parece ser esencial para el inicio de la transcripción por estas dos RNA polimerasas.

Como se mencionó antes, la RNA pol I transcribe los RNA ribosomales. Estos transcritos son moléculas grandes de pre-RNAr (45S) que son luego procesados para generar los RNAr 28S, 18S y 5.8S. El promotor de los genes de RNAr, el cual incluye alrededor de 150 pares de bases antes del sitio de la transcripción, es reconocido por 2 factores de transcripción que interactúan entre sí y atraen la RNA pol I para formar el complejo de iniciación (Figura 15). La TBP hace parte de uno de los factores de transcripción, sin embargo ella no se une directamente a la secuencia del promotor porque no existe caja TATA en este tipo de promotores.

De otro lado, los promotores de los genes que se transcriben por la RNA pol III se encuentran dentro de la secuencia a ser transcrita (Figura 16), lo cual los diferencia de los promotores de las polimerasas I y II. La transcripción de los genes del RNAr 5S por parte de la RNA pol III dependen de la unión del TFIIA a secuencias específicas en el promotor de estos genes, después

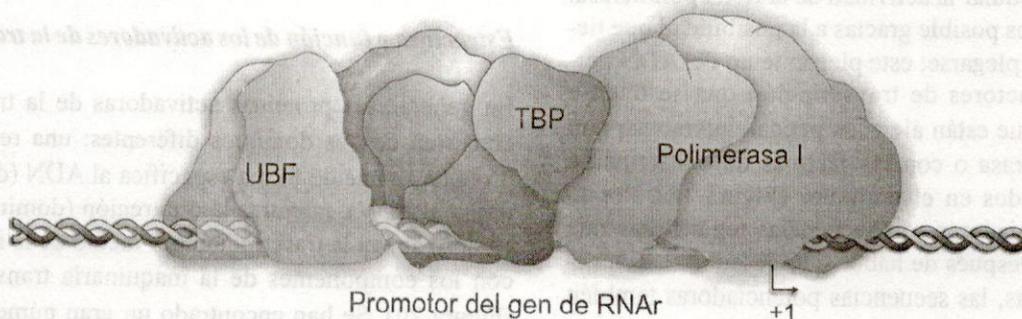


Figura 15. La región promotora de los genes que se transcriben por acción de la RNA pol I (la mayor parte de los RNA ribosomales) es reconocida por factores de transcripción (UBF y TBP) que interactúan entre sí y atraen la RNA pol I para formar el complejo de inicio de la transcripción. Aunque la TBP hace parte de uno de estos factores de transcripción no se une directamente a la secuencia del promotor porque no existe caja TATA en este tipo de promotores.

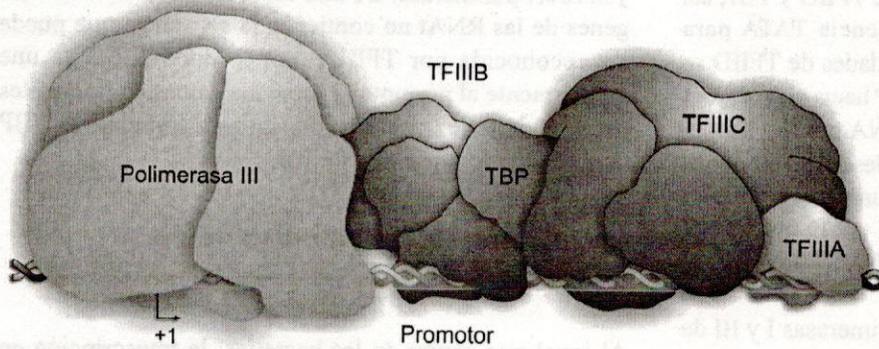
de lo cual se unen de manera secuencial TFIIC, TFIIB y la RNA polimerasa. De otro lado, los promotores de los genes de las RNAt no contienen la secuencia que puede ser reconocida por TFIIA, por lo que TFIIC se une directamente al promotor y atrae los demás componentes del complejo de transcripción. En este complejo la TBP está formando parte del TFIIB.

## Regulación de la transcripción en eucariotas

Al igual que ocurre en las bacterias, la transcripción en las células eucarióticas también es controlada por proteínas que se unen a secuencias específicas del ADN y regulan la actividad de la RNA polimerasa. Pero además de la regulación por proteínas unidoras de ADN, la expresión de los genes es modulada por el empaquetamiento del ADN en la cromatina y por la modificación que introduce la metilación del ADN, lo cual se discutirá más adelante.

Como se mencionó antes, los genes transcritos por la RNA pol II poseen dos elementos centrales en el promotor, la caja TATA y la secuencia *Inr*, que sirven para la unión específica de factores de transcripción generales. Además existen otras secuencias que actúan en *cis* (*cis* activadoras) que permiten la unión de factores reguladores que controlan la expresión de genes individuales. Estas secuencias reguladoras por lo general se encuentran ubicadas corriente arriba a la caja TATA. En muchos genes eucarióticos se han identificado dos secuencias reguladoras ubicadas dentro de las 100 pares de bases previas a la caja TATA, cuyas secuencias consenso corresponden a CCAAT y GGGCGG (esta última se ha denominado caja GC) e interactúan con proteínas específicas (Figura 12).

## Genes 5S rRNA



## Genes tRNA

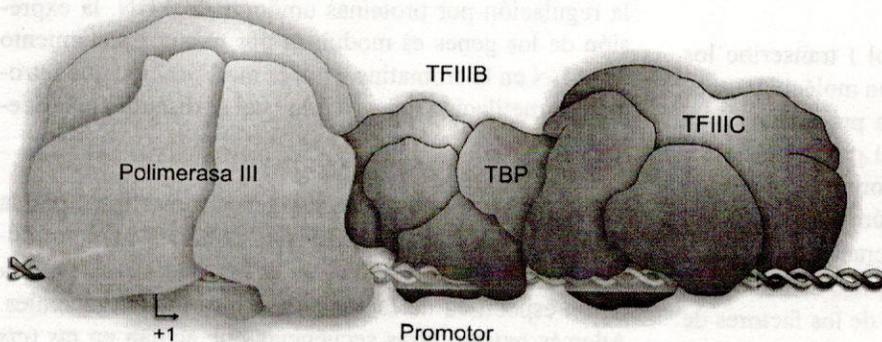


Figura 16. Los promotores de los genes que se transcriben por la RNA pol III se encuentran dentro de la secuencia a ser transcrita, lo cual los diferencia de los promotores de las polimerasas I y II.

De otro lado, muchos genes en las células eucarióticas son controlados por secuencias reguladoras muy retiradas del sitio de inicio de la transcripción, a veces tan lejanas como 10 kilobases, las cuales se denominan aumentadores o potenciadores. La actividad potenciadora de estas secuencias se mantiene cuando se ubican corriente arriba o corriente abajo del promotor o aun si la orientación de la secuencia está en uno u otro sentido (Figura 17). Los potenciadores, al igual que los promotores, regulan la transcripción de los genes por medio de la unión de factores de transcripción, los que a su vez se encargan de modular la actividad de la RNA polimerasa. Este fenómeno es posible gracias a la posibilidad que tiene el ADN para plegarse; este plegamiento del ADN permite que los factores de transcripción que se unen a potenciadores que están alejados puedan interactuar con la RNA polimerasa o con los factores de transcripción generales ubicados en el promotor (Figura 18). Por lo tanto los potenciadores son secuencias reguladoras que actúan en *cis*. Después de haber sido descubiertas en células eucarióticas, las secuencias potenciadoras también se identificaron en el ADN de los procariotes.

Una característica importante de los potenciadores es que con frecuencia contienen varias secuencias funcionales

que se unen a diferentes proteínas reguladoras de la transcripción y estas proteínas actúan en conjunto para regular la expresión genética. Por ejemplo, el potenciador del gen de la cadena pesada de las inmunoglobulinas contiene al menos 9 secuencias diferentes que sirven como sitios de unión para distintas proteínas reguladoras (Figura 19), muchos de los cuales funcionan como secuencias potenciadoras en otros genes; sin embargo, la combinación de estas 9 secuencias permite constituir un potenciador específico para la expresión de la cadena pesada de las inmunoglobulinas en los linfocitos B.

### **Estructura y función de los activadores de la transcripción.**

En general, las proteínas activadoras de la transcripción consisten de dos dominios diferentes: una región de la proteína se une de forma específica al ADN (dominio de unión al ADN), mientras la otra región (dominio de activación) activa la transcripción por medio de la interacción con los componentes de la maquinaria transcripcional (Figura 20). Se han encontrado un gran número de proteínas que actúan como factores de transcripción, las cuales hasta el momento se pueden agrupar hasta el momento en 4 grandes tipos estructurales:

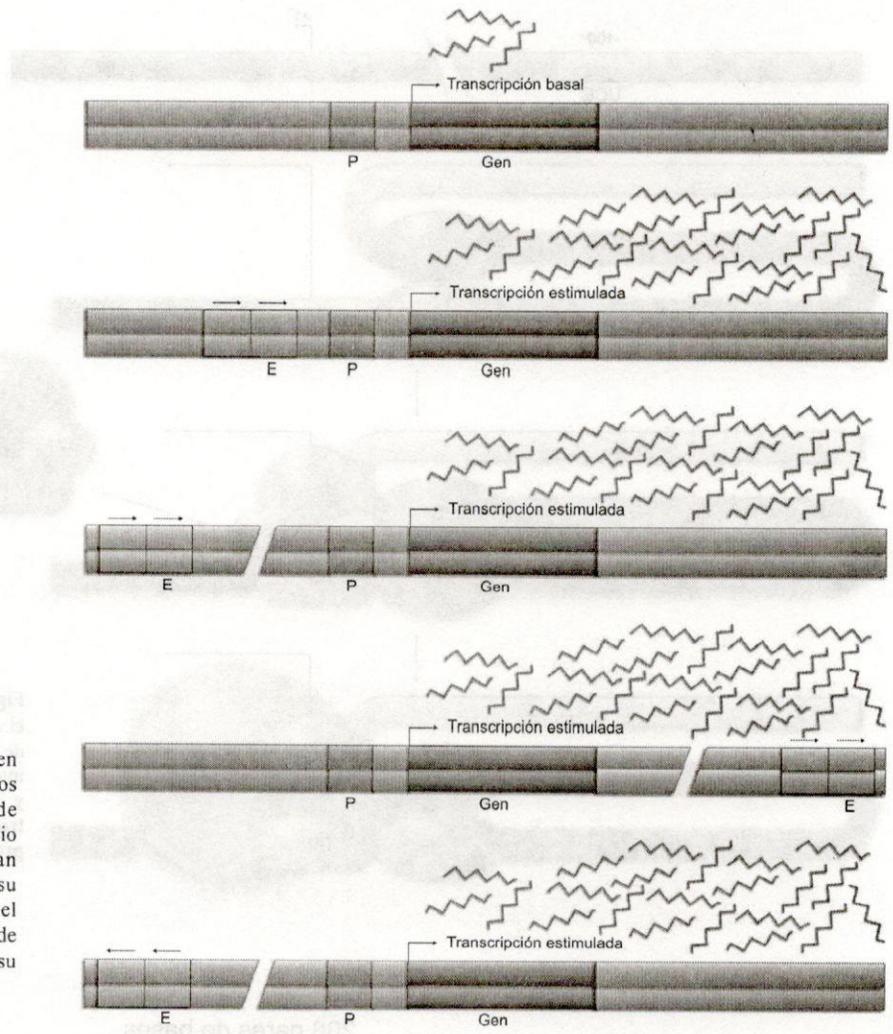


Figura 17. Potenciadores de la transcripción en eucariotes. La expresión de muchos genes en los eucariotes es controlada por la presencia de secuencias que están muy retiradas del sitio de inicio de la transcripción, las cuales se denominan potenciadores. Los potenciadores mantienen su actividad aunque se coloquen muy retirados del promotor o si colocan corriente abajo del sitio de inicio de la transcripción o incluso si se invierte su orientación en el ADN.

**Proteínas con dominios en dedos de zinc.** Estas son proteínas que contienen regiones de repeticiones de residuos de cisteína e histidina que se unen a iones de zinc y se pliegan para formar estructuras en forma de asas (dedos), las cuales se unen a secuencias específicas en el ADN (Figura 21). Ejemplos de factores de transcripción que contienen dedos de zinc son los receptores de hormonas esteroideas y la proteína de especificidad 1 (Sp 1).

**Proteínas con dominio hélice-giro-hélice.** En estas proteínas una hélice alfa hace la mayor parte del contacto con el ADN, mientras otras hélices se colocan sobre el complejo para estabilizar la interacción (Figura 22). Las proteínas con estos dominios se encontraron inicialmente en proteínas unidoras de ADN en procariotes. En eucariotas, un grupo importante de proteínas con este tipo de dominio lo constituyen las proteínas de homeodominio, las cuales tienen un papel central en la regulación de la expresión genética durante el desarrollo embrionario al interactuar con secuencias conservadas denominadas homeocajas, que se encuentran en los genes activos durante el desarrollo embrionario.

**Proteínas con cierre (cremallera) de leucina.** Estas son proteínas formadas por la dimerización de dos cadenas polipeptídicas idénticas (homodímeros), las cuales interactúan entre sí gracias a que en uno de los extremos de las hélices alfa presentan de cuatro a cinco residuos de leucina (a espacios de 7 aminoácidos), que extienden sus radicales hidrófobos hacia un lado de la región helicoidal permitiendo que se formen interacciones hidrófobas entre las leucinas de una cadena con las leucinas de la cadena opuesta, así formando una especie de cremallera (Figura 23). Después de la región rica en leucinas las hélices presentan aminoácidos cargados positivamente (lisina y arginina) que permiten la interacción con el ADN.

**Proteínas con dominio hélice-asa-hélice.** Estas proteínas presentan una estructura similar a las de cremallera de leucina, solo que la dimerización de las subunidades polipeptídicas ocurre por intermedio de dos hélices alfa que están separadas del resto de la cadena por una asa (Figura 24). Una característica de estas proteínas, al igual que las de cremallera de leucina, es que miembros dife-

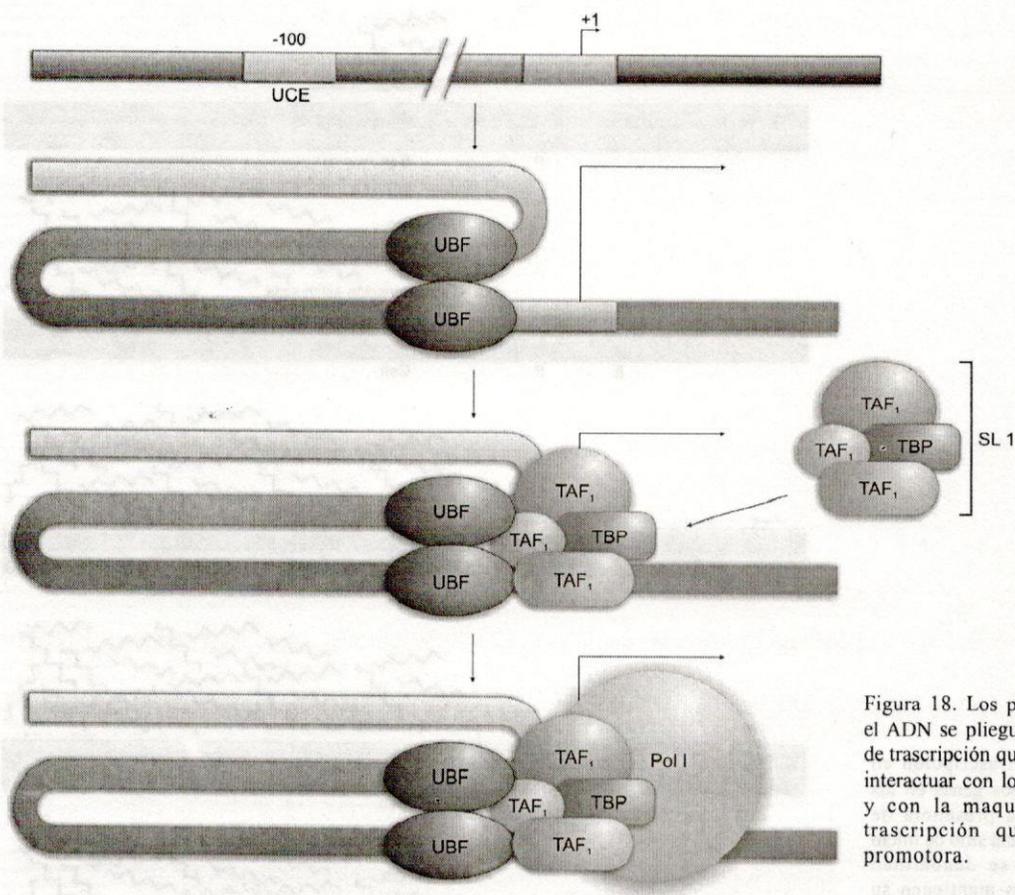


Figura 18. Los potenciadores actúan haciendo que el ADN se pliegue, lo cual permite que los factores de transcripción que se unen a estas secuencias puedan interactuar con los factores de transcripción generales y con la maquinaria encargada de iniciar la transcripción que se encuentra en la región promotora.

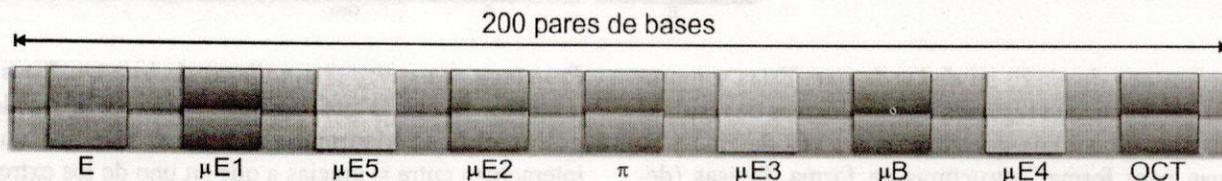


Figura 19. En esta figura se muestra un esquema del potenciador del gen de la cadena pesada de las inmunoglobulinas. Este potenciador ocupa alrededor de 200 pb y contiene nueve secuencias funcionales, las cuales en conjunto permiten la expresión específica de este gen en los linfocitos B.

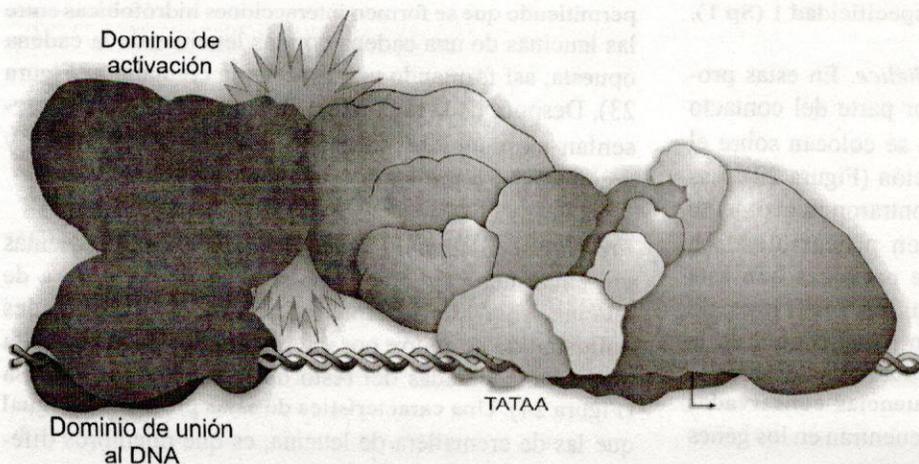


Figura 20. En general las proteínas que actúan como factores de transcripción contienen dos dominios diferentes: uno que reconoce específicamente e interactúa con el ADN (dominio de unión) y otro que interactúa con la maquinaria transcripcional (dominio de activación), el cual actúa como activador (en algunos casos como inhibidor) de la transcripción.

### Dedos de zinc

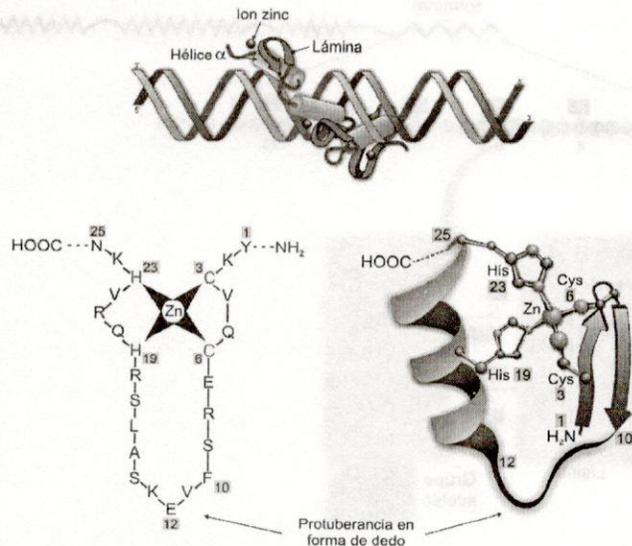


Figura 21. *Proteínas con dominios en dedos de zinc*. Estas son proteínas que contienen regiones de repeticiones de residuos de cisteína e histidina que se unen a iones de zinc y se pliegan para formar estructuras en forma de asas (dedos), las cuales se unen a secuencias específicas en el ADN

### Hélice - giro - hélice

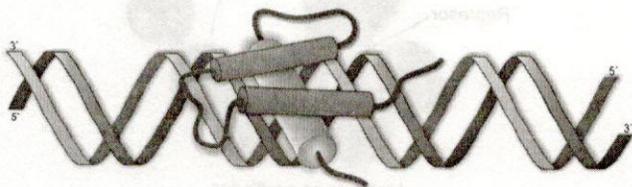


Figura 22. *Proteínas con dominio hélice-giro-hélice*. En estas proteínas una hélice alfa hace la mayor parte del contacto con el ADN, mientras otras hélices se colocan sobre el complejo para estabilizar la interacción.

### Cremallera de leucina

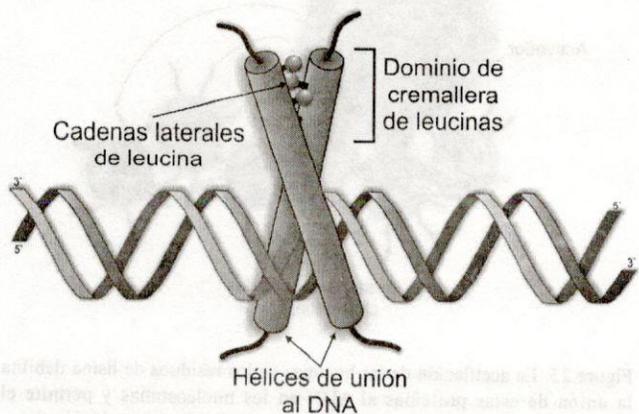


Figura 23. *Proteínas con cierre (cremallera) de leucina*. Estas son proteínas formadas por la dimerización de dos cadenas polipeptídicas idénticas (homodímeros), las cuales interactúan entre sí gracias a que en uno de los extremos de las hélices alfa presentan de cuatro a cinco residuos de leucina (a espacios de 7 aminoácidos) que extienden sus radicales hidrófobos hacia un lado de la región helicoidal permitiendo que se formen interacciones hidrófobas entre las leucinas de una cadena con las leucinas de la cadena opuesta, así formando una especie de cremallera.

### Hélice - asa - hélice

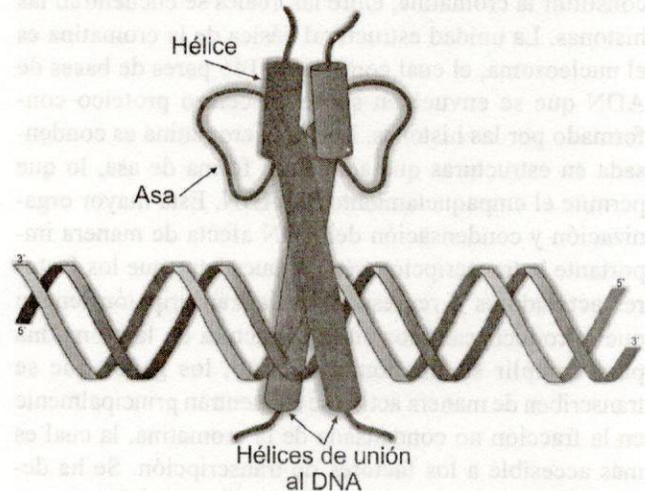


Figura 24. *Proteínas con dominio hélice-asa-hélice*. Estas proteínas presentan una estructura similar a las de cremallera de leucina, sin embargo, la dimerización de las subunidades polipeptídicas ocurre por intermedio de dos hélices alfa que están separadas del resto de la cadena por una asa.

rentes de estas familias de proteínas pueden dimerizar unos con otros y así dar una mayor diversidad de factores de transcripción que difieren en las secuencias que reconocen en el ADN.

Todas las proteínas que actúan como factores de transcripción poseen dominios de activación, los cuales parecen estimular la transcripción al interactuar con los factores de transcripción generales tales como TFIIB o TFIID, lo que a su vez facilita el ensamble del complejo de transcripción en el promotor.

Al igual ocurre en procariotes, en las células eucariotas también existen represores de la transcripción. En algunos casos los represores eucarióticos simplemente interfieren con la unión de factores de transcripción al ADN; por ejemplo, la unión del represor en una región cercana al sitio de inicio de la transcripción puede bloquear la unión de la RNA pol. Otros represores actúan por competición del sitio de unión de las secuencias reguladoras; en este caso, el represor tiene el mismo dominio de unión al ADN pero no posee el dominio de activación. Existen otros represores que inhiben la transcripción por medio de interacciones proteína-proteína, los cuales son llamados represores activos.

### Regulación de la transcripción por cambios en la estructura de la cromatina

Como se ha descrito en los capítulos previos, el ADN en las células eucariotas no se encuentra desnudo, sino que

presenta una estrecha unión a una serie de proteínas para constituir la cromatina, entre las cuales se encuentran las histonas. La unidad estructural básica de la cromatina es el nucleosoma, el cual consiste de 146 pares de bases de ADN que se envuelven sobre un centro proteico conformado por las histonas. Luego la cromatina es condensada en estructuras que adquieren forma de asa, lo que permite el empaquetamiento del ADN. Esta mayor organización y condensación del ADN afecta de manera importante la transcripción génica, pues hace que los factores activadores o represores de la transcripción tengan que introducir cambios en la estructura de la cromatina para cumplir su función. Por tanto, los genes que se transcriben de manera activa se encuentran principalmente en la fracción no condensada de la cromatina, la cual es más accesible a los factores de transcripción. Se ha demostrado que los genes que se transcriben activamente en varios organismos eucariotas se encuentran ubicados en fracciones no condensadas de la cromatina. Sin embargo, la decondensación de la cromatina no es suficiente para permitir que el ADN sea accesible a la maquinaria transcripcional, pues aún quedan las histonas y otros componentes básicos de la cromatina.

El hecho que el ADN se encuentre en una estrecha unión alrededor del centro del nucleosoma afecta de manera importante la posibilidad de los factores de transcripción para interactuar con el ADN y de la RNA polimerasa para transcribir a lo largo del molde de ADN. Para eliminar este efecto inhibitorio del nucleosoma ocurren varios fenómenos. Primero, se produce una acetilación de los residuos de lisina presentes en las histonas, lo cual reduce la carga positiva neta de estas proteínas y de esa manera debilita la unión de las histonas al ADN o altera su interacción con otras proteínas (Figura 25). De esta manera se posibilita el acceso de proteínas unidoras de ADN a la cromatina. En segundo término varias proteínas no histonas se unen al nucleosoma, entre las cuales se encuentran los factores de remodelación del cromosoma que facilitan la unión de los factores de transcripción a la cromatina al alterar la estructura del nucleosoma pero sin eliminar las histonas del ADN.

### Regulación de la transcripción por metilación del ADN

La metilación del ADN es un mecanismo que permite regular la transcripción por medio de modificación de la estructura de la cromatina. Los residuos de citosina en el ADN pueden ser modificados gracias a la adición de grupos metilo al carbono en la posición 5 de esta base (Figura 26). El ADN se metila en las citosinas que preceden a residuos de guanina (dinucleótidos CpG). La metilación

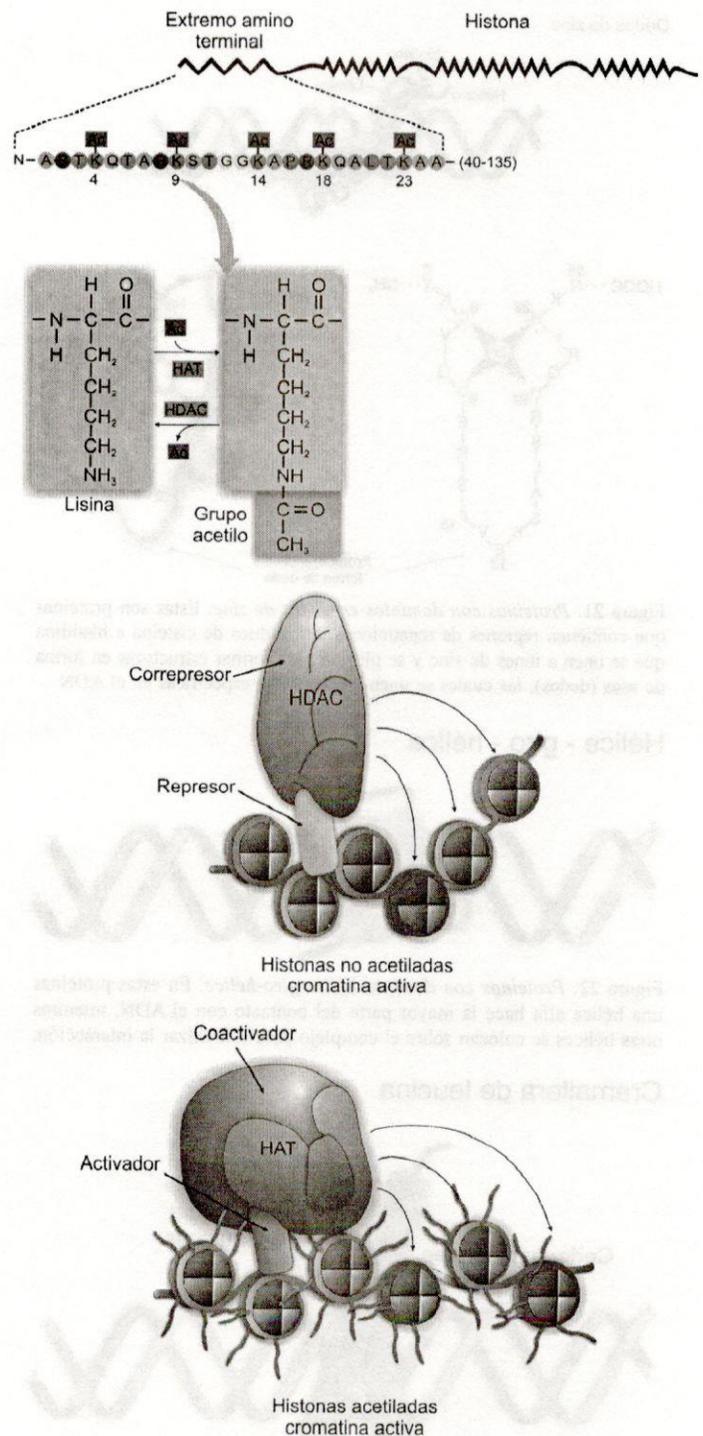


Figura 25. La acetilación de las histonas en los residuos de lisina debilita la unión de estas proteínas al ADN en los nucleosomas y permite el acceso de factores de transcripción y la maquinaria transcripcional.

de regiones con alta frecuencia de dinucleótidos CpG y cercanas a promotores conduce a una disminución de la actividad transcripcional; esto ocurre como consecuencia de la unión específica de la proteína MeCP2 al ADN metilado, la cual es un factor represor de la transcripción.

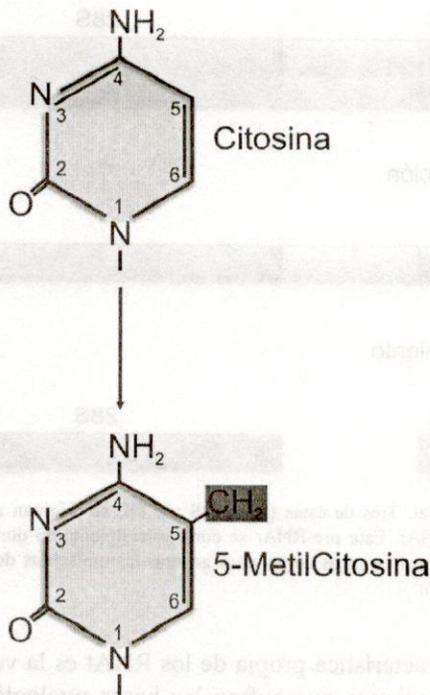


Figura 26. La metilación del ADN ocurre mediante la adición de grupos metilo en el carbono en la posición 5 de los residuos de citosina. El ADN se metila en las citosinas que preceden a residuos de guanina (dinucleótidos CpG). La metilación del ADN produce cambios en la función de esta molécula, particularmente se ha asociado a una disminución de la actividad transcripcional.

Aunque se ha discutido mucho acerca de la relevancia de la metilación del ADN en la regulación transcripcional, en las células de los vertebrados existe un papel claro de este mecanismo en el fenómeno conocido como marcado (“imprinting”) genómico. La mayoría de las veces las células diploides expresan tanto el alelo paterno como el materno de un gen, sin embargo se han descrito un pequeño número de genes eucariotas cuya expresión depende de si son heredados del padre o de la madre. Esto significa que en algunos casos solo se expresa el alelo paterno o materno, dependiendo de cual se encuentre transcripcionalmente inactivo. Esta inhibición de la transcripción ocurre como consecuencia de la metilación del ADN en dichos genes durante el desarrollo de las células germinales masculinas o femeninas. En las células somáticas del individuo que resulta de este apareamiento se continúa el patrón de metilación del ADN, lo cual permite que se mantenga la inhibición transcripcional de esos genes.

### Procesamiento del RNA

La mayor parte de las moléculas de RNA sufren un proceso de modificación antes de convertirse en moléculas funcionales. La excepción a este proceso son los RNAm bacterianos, pues éstos son usados como moldes para la

traducción a proteínas aún mientras el RNAm se está transcribiendo. En las células procariotas y eucariotas los transcritos iniciales de los RNAr y de los RNAt sufren una serie de etapas de procesamiento para producir RNAs maduros. Los transcritos primarios de los RNAm eucarióticos sufren las modificaciones más importantes, entre las cuales la eliminación de los intrones del RNA antes de salir del núcleo es uno de los eventos más importantes como se verá luego.

### Procesamiento de los RNAs ribosomal y de transferencia

El procesamiento de las moléculas de los RNAs ribosomal y de transferencia es en principio similar en procariotes y eucariotas. En los eucariotas existen 4 especies de RNAr, 3 de los cuales (28S, 18S y 5.8S) se originan a partir de la ruptura de una molécula precursora larga llamada pre-RNAr. Este pre-RNAr se corta inicialmente en dos fragmentos, uno que es el precursor del 18S y otro que contiene los RNAr 5.8S y 28S, luego estos sufren otros cortes para generar las moléculas definitivas y finalmente la molécula 5.8S se une a la 28S por medio de puentes de hidrógeno. En los procariotes los RNAr 23S, 16S y 5S son equivalentes a los RNAr de eucariotas y también se forman por el procesamiento de un transcrito inicial pre-RNAr (Figura 27).

Al igual que sucede con los RNAr, los RNAt de procariotes y eucariotas se forman a partir de moléculas precursoras más largas (pre-RNAt). El procesamiento de estos pre-RNAt depende del corte de regiones de RNA en los extremos 5' y 3' de la molécula. La eliminación del fragmento sobrante en el extremo 5' es dependiente de la acción de una enzima conocida como RNasa P. Esta RNasa P está constituida tanto por proteína como por RNA y ambos componentes son necesarios para su actividad catalítica. En la década de los 80s, el investigador Sidney Altman y sus colaboradores demostraron que la porción de RNA aislada de la RNasa P tenía la capacidad de catalizar el procesamiento del RNAt, por lo que se pudo demostrar que esta RNasa P constituía un tipo nuevo de moléculas con actividad catalítica que se denominaron ribozimas, de manera que la RNasa P se convirtió así en el prototipo de estas enzimas. La RNasa P corta el extremo 5' del pre-RNAt, mientras otra RNasa, ésta si proteica, corta el extremo 3' del pre-RNAt. Sin embargo, el procesamiento del extremo 3' requiere de la adición del trinucleótido CCA, el cual se encuentra en todos los RNAt y constituye el sitio de unión del aminoácido (Figura 28). En algunos RNAt este extremo CCA está codificado en sus respectivos genes, pero en otros este trinucleótido se adiciona durante su procesamiento por una enzima específica.

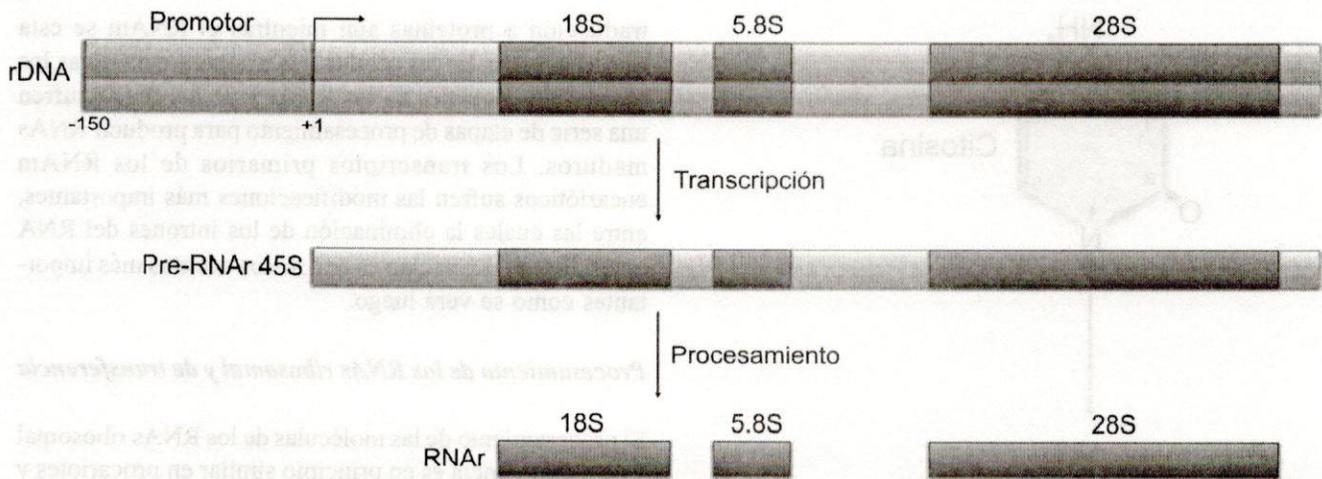


Figura 27. En los eucariotas existen cuatro moléculas que constituyen el RNA ribosomal. Tres de éstas (28S, 18S y 5.8S) se originan a partir de la transcripción de un gen que da origen a una molécula precursora larga llamada pre-RNAr. Este pre-RNAr se corta inicialmente en dos fragmentos, uno que es el precursor del 18S y otro que contiene los RNAr 5.8S y 28S, luego éstos sufren otros cortes para generar las moléculas definitivas. En la estructura final la molécula 5.8S se une a la 28S por medio de puentes de hidrógeno.

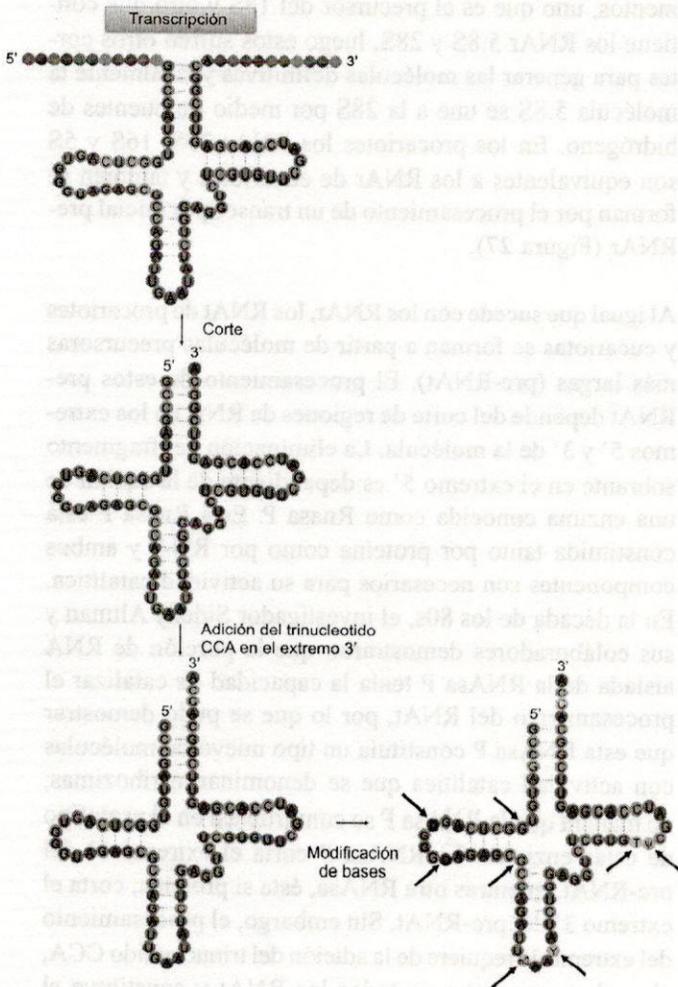


Figura 28. El RNA de transferencia sufre varios cambios después de su transcripción. El primero es un corte de secuencias en los extremos 5' y 3', luego ocurre la adición del trinucleótido CCA en el extremo 3', el cual constituye el sitio de unión del aminoácido, finalmente los RNAr sufren modificaciones químicas en ciertas bases nucleotídicas.

Otra característica propia de los RNAr es la variedad de modificaciones que sufren las bases nucleotídicas que componen su secuencia. Aproximadamente un 10% de las bases en los RNAr son modificados en posiciones específicas de la molécula (Figura 28). Modificaciones en las uridinas dan origen a nucleótidos tales como dihidouridina, ribotimidina y pseudouridina; mientras que la modificación de guaninas produce metilguanina e inosina (Figura 29). Aunque la función específica de estos nucleótidos modificados no se conoce, sí parecen ser importantes para la síntesis de proteínas al alterar las propiedades de apareamiento de bases de la molécula de RNAr.

### Procesamiento del RNAm

Puesto que en las bacterias el RNAm es traducido a proteína mientras aún se está transcribiendo, las moléculas de RNAm no sufren ninguna modificación importante después de su síntesis. Por su parte, los RNAm de las células eucarióticas deben ser transportados desde el núcleo hacia el citoplasma para su traducción; sin embargo, antes de salir del núcleo estas moléculas sufren modificaciones importantes en sus extremos 5' y 3', además de la eliminación de los intrones.

### Modificación en el extremo 5' ("capping")

Poco después de su síntesis, el RNAm es modificado en su extremo 5' por la adición de una molécula de 1-metilguanina, lo cual constituye la estructura conocida como capucha o casquete ("capping") del RNAm. Esta se forma por la unión de una molécula de GTP en una orientación inversa (5') al nucleótido ubicado en el extremo 5' del pre-RNAm, formando una unión 5'-5' (Figura

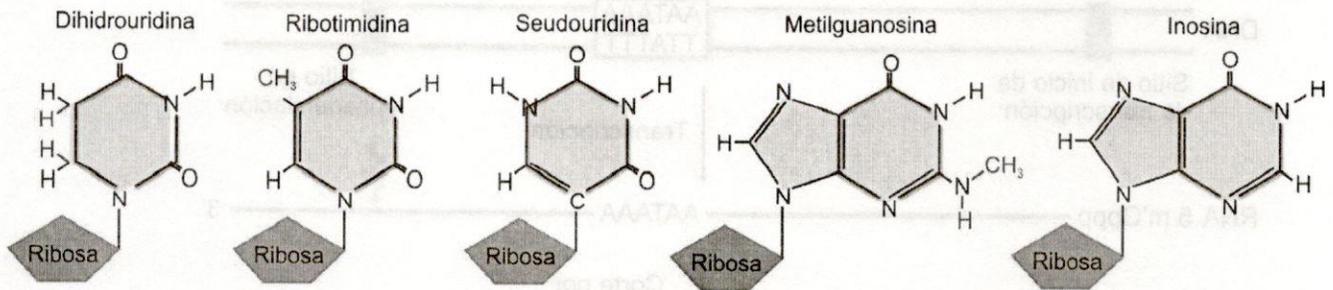


Figura 29. Las modificaciones en los RNAt ocurren esencialmente en las uridinas, las cuales dan origen a nucleótidos tales como dihidrouridina, ribotimidina y seudouridina, y en las guanosinas lo que da origen a metilguanosina e inosina.

30). Luego se adicionan grupos metilos a este residuo de guanina en la posición N-7 y a los grupos OH en la posición 2' de los dos primeros residuos de ribosa de esta cadena de RNAm. Esta modificación tiene un papel importante en la protección del RNAm frente a RNAsas y al mismo tiempo es necesaria para la ubicación del inicio de la traducción en el ribosoma, como se verá en el capítulo siguiente.

#### Modificación del extremo 3' (poliadenilación)

De otro lado, en el extremo 3' del pre-RNAm se presenta un fenómeno conocido como poliadenilación del RNAm, el cual consiste en la adición de una cola con alrededor de 200 nucleótidos de adenina. La adición de esta secuencia de adeninas depende de varias señales presentes en la secuencia del RNA. La señal más conservada que se ha encontrado en células de mamíferos es el hexanucleótido AAUAAA, el cual se encuentra entre 10 a 30 nucleótidos corriente arriba del sitio de poliadenilación. Además existe otro elemento importante para este procesamiento, una secuencia rica en G-U ubicada corriente abajo del sitio de poliadenilación. Estos elementos en el ADN son reconocidos por un complejo proteico que incluye una endonucleasa que corta la cadena de ADN y por la poli-A polimerasa que adiciona los nucleótidos de adenina en el sitio donde se ha cortado el RNA (Figura 31). Esta poliadenilación es una señal para la finalización de la transcripción, la que generalmente termina en algunos cientos de nucleótidos corriente abajo del sitio de poliadenilación. La poliadenilación es importante para una mayor estabilidad del RNAm y así asegurar su traducción a proteína.

#### Eliminación de intrones ("splicing")

Como se mencionó antes, las secuencias que codifican en los genes (exones) de la mayor parte de los organismos eucariotas se encuentran interrumpidas por regiones no codificadoras conocidas como intrones. Durante la transcripción, la RNA pol II realiza una copia de todas las

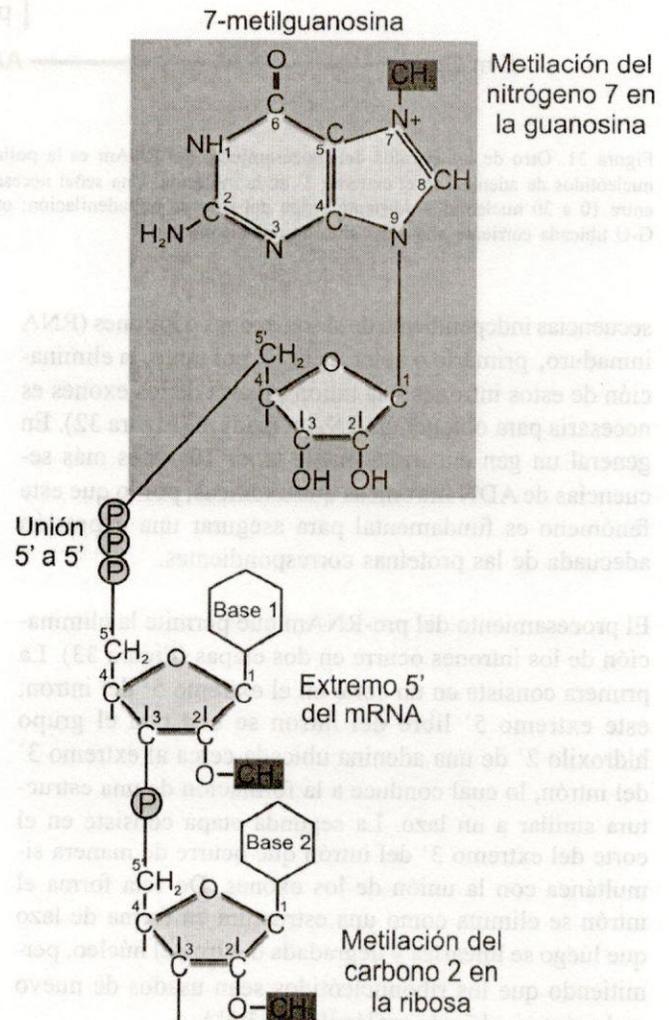


Figura 30. Una vez ocurre la síntesis de la molécula precursora del RNAm es modificada en su extremo 5' por la adición de 7-metilguanosina, lo cual constituye la capucha del RNAm. Esta se forma por la unión de una molécula de GTP en una orientación inversa (5') al nucleótido ubicado en el extremo 5' del pre-RNAm, formando una unión 5'-5'. Luego se adicionan grupos metilos a este residuo de guanina en la posición N-7 y a los grupos OH en la posición 2' de los dos primeros residuos de ribosa de esta cadena de RNAm.

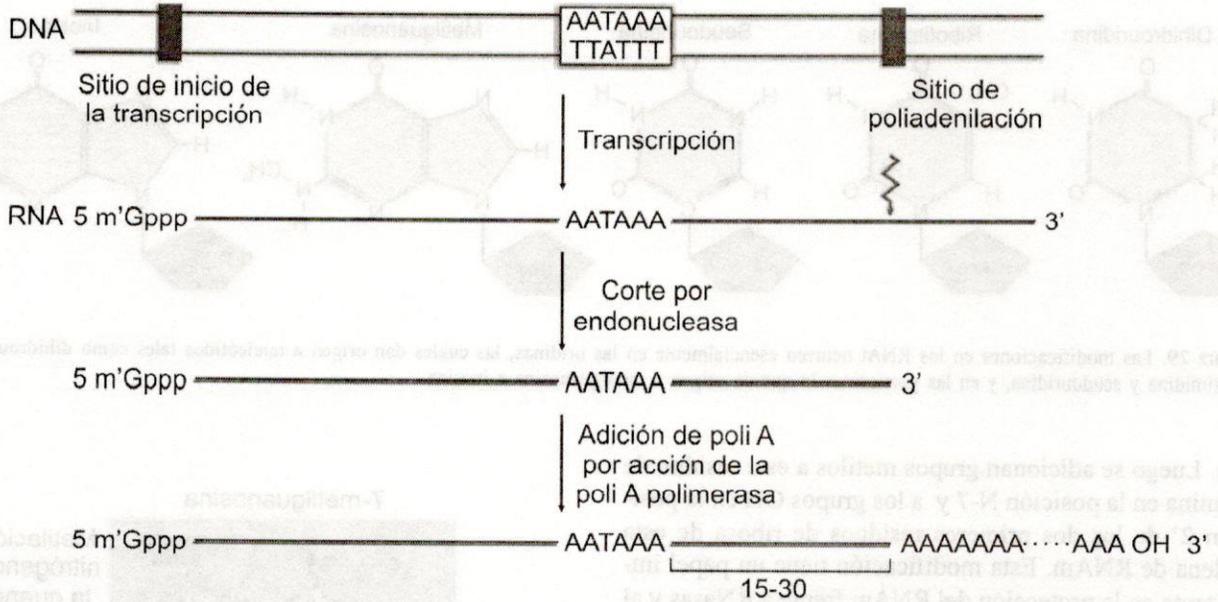


Figura 31. Otro de los eventos del procesamiento del RNAm es la poliadenilación que consiste en la adición de una cola con alrededor de 200 nucleótidos de adenina en el extremo 3' de la molécula. Una señal necesaria para esta adición es el hexanucleótido AAUAAA, el cual se encuentra entre 10 a 30 nucleótidos corriente arriba del sitio de poliadenilación; otro elemento importante para este procesamiento es una secuencia rica en G-U ubicada corriente abajo del sitio de poliadenilación.

secuencias independiente de si son exones o intrones (RNA inmaduro, primario o heterogéneo); por tanto, la eliminación de estos intrones y la unión precisa de los exones es necesaria para obtener un RNAm maduro (Figura 32). En general un gen eucariota puede tener 10 veces más secuencias de ADN intrónicas que exónicas, por lo que este fenómeno es fundamental para asegurar una expresión adecuada de las proteínas correspondientes.

El procesamiento del pre-RNAm que permite la eliminación de los intrones ocurre en dos etapas (Figura 33). La primera consiste en un corte en el extremo 5' del intrón; este extremo 5' libre del intrón se une con el grupo hidroxilo 2' de una adenina ubicada cerca al extremo 3' del intrón, lo cual conduce a la formación de una estructura similar a un lazo. La segunda etapa consiste en el corte del extremo 3' del intrón que ocurre de manera simultánea con la unión de los exones. De esta forma el intrón se elimina como una estructura en forma de lazo que luego se lineariza y degradada dentro del núcleo, permitiendo que los ribonucleótidos sean usados de nuevo en la síntesis de más moléculas de RNA.

Las reacciones de corte y empalme ("splicing") de intrones y exones ocurren gracias a la existencia de secuencias consenso altamente conservadas en el sitio de corte 5' del intrón, en el sitio de corte 3' y en el punto de ramificación del intrón (este último es donde se une el extremo del intrón 5' para formar el lazo) (Figura 34). Estas secuencias permiten el reconocimiento de los intrones por parte

de la maquinaria nuclear que realiza el proceso de corte y empalme. Estas reacciones tienen lugar en grandes complejos moleculares compuestos por proteínas y RNAs, denominados espliceosomas (de inglés "splice" corte y empalme). Los RNAs que hacen parte de estos espliceosomas son moléculas que tienen entre 50 y 200 nucleótidos, pertenecientes a los RNAs nucleares pequeños (snRNAs), los que a su vez se encuentran unidos con seis a diez moléculas proteicas para conformar las partículas de ribonucleoproteínas nucleares pequeñas (snRNPs). Se han descrito 5 snRNPs esenciales para la eliminación de los exones denominadas U1, U2, U4, U5 y U6.

La conformación del espliceosoma se inicia con la unión de la U1 al sitio 5' del intrón, lo cual se debe al apareamiento de bases entre la secuencia consenso en el extremo 5' del intrón y la secuencia complementaria presente en el extremo 5' del snRNA que hace parte de la snRNP U1 (Figura 35). Luego la snRNP U2 se une al punto de ramificación del intrón también por apareamiento de bases entre el snRNA U2 y esa región del intrón. En este momento un complejo preformado consistente de las snRNPs U4/U6 y U5 se incorpora al espliceosoma gracias a la unión de U5 a secuencias corriente arriba del sitio 5' del intrón (secuencia rica en pirimidinas). Luego U6 se disocia de U4 y desplaza a U1 del extremo 5' del intrón para permitir la formación del lazo. Entonces U5 se une a las secuencias en el sitio 3' del intrón lo que permite el corte del intrón y la unión de los exones. Además de permitir el reconocimiento de las regiones consenso para

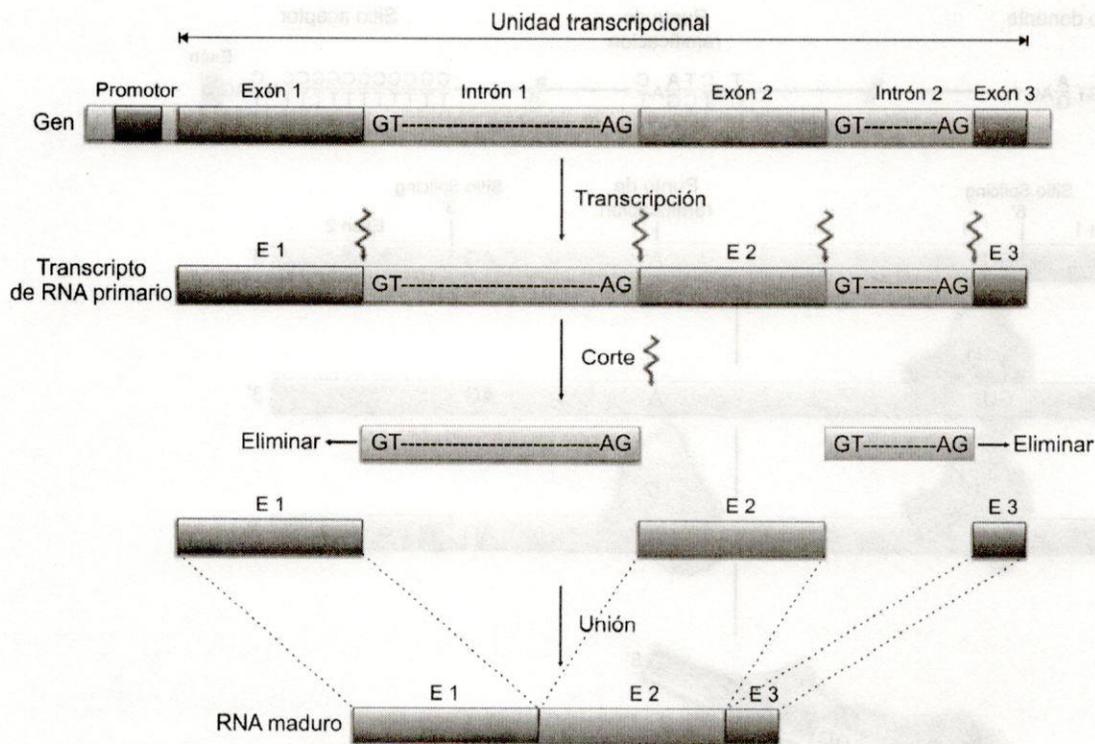


Figura 32. Durante la transcripción la RNA pol II realiza una copia de todas las secuencias que constituyen el gen, independiente de si son exones o intrones, lo cual origina un preRNA, RNA inmaduro, primario o heterogéneo. Posteriormente ocurre la eliminación de estos intrones y la unión de los exones entre sí para obtener un RNA maduro.

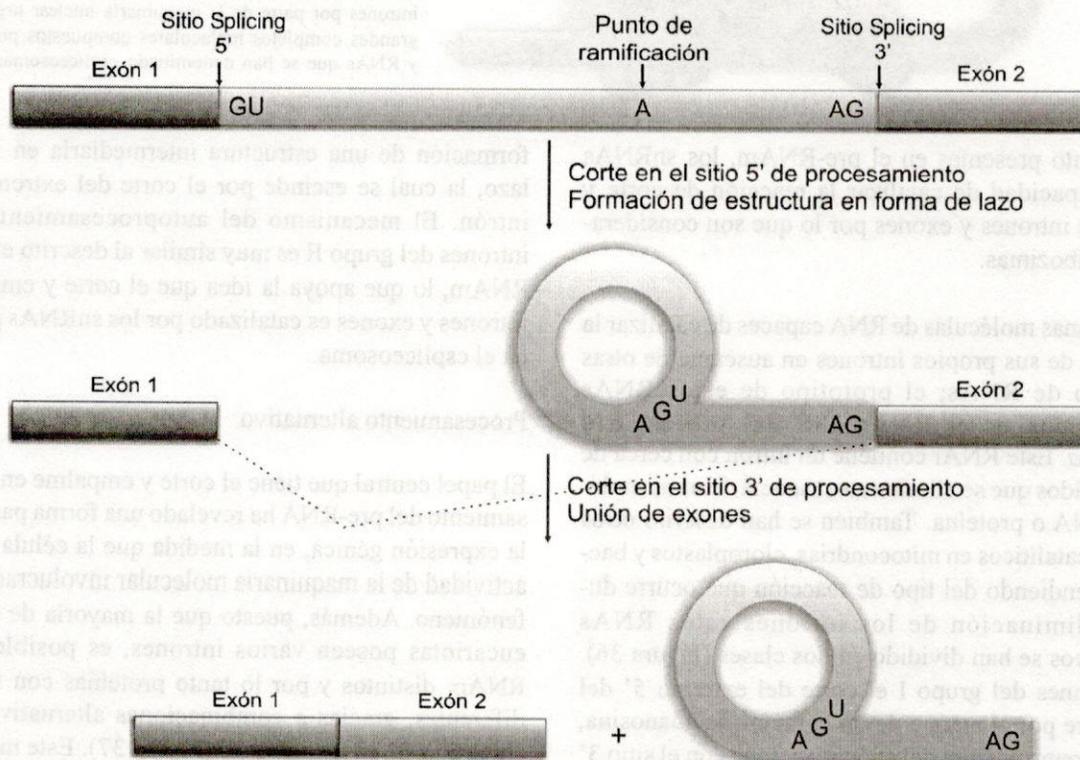


Figura 33. El procesamiento del pre-RNA que permite la eliminación de los intrones ocurre en dos etapas. La primera consiste en un corte en el extremo 5' del intrón; este extremo 5' libre del intrón se une con el grupo hidroxilo 2' de una adenina ubicada cerca al extremo 3' del intrón, lo cual conduce a la formación de una estructura similar a un lazo. La segunda etapa consiste en el corte del extremo 3' del intrón que ocurre de manera simultánea con la unión de los exones. La estructura en forma de lazo se lineariza y se degradada dentro del núcleo, y los ribonucleótidos son usados de nuevo en la síntesis de más moléculas de RNA.

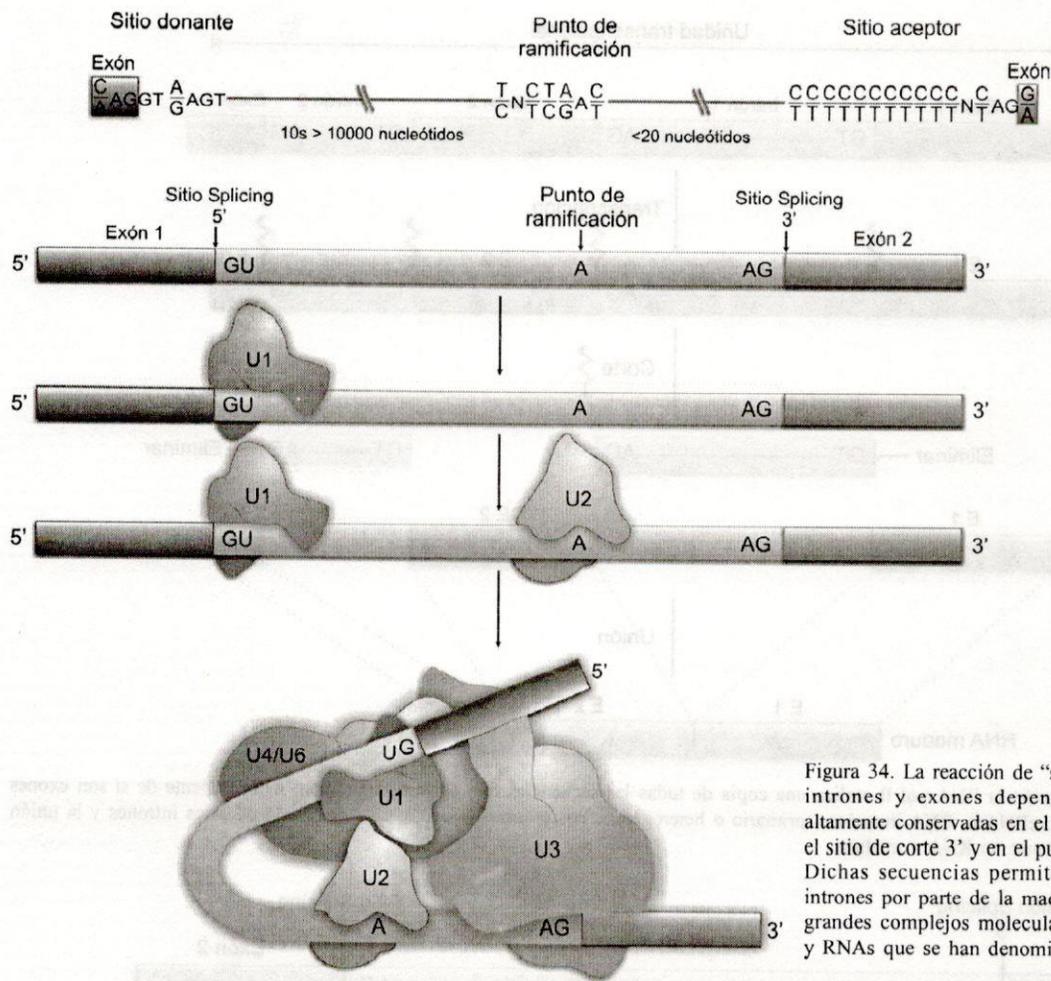


Figura 34. La reacción de "splicing" (corte y empalme) de intrones y exones depende de secuencias consenso altamente conservadas en el sitio de corte 5' del intrón, en el sitio de corte 3' y en el punto de ramificación del intrón. Dichas secuencias permiten el reconocimiento de los intrones por parte de la maquinaria nuclear organizada en grandes complejos moleculares compuestos por proteínas y RNAs que se han denominado espliceosomas.

procesamiento presentes en el pre-RNAm, los snRNAs tienen la capacidad de catalizar la reacción de corte y empalme de intrones y exones por lo que son considerados como ribozimas.

Existen algunas moléculas de RNA capaces de catalizar la eliminación de sus propios intrones en ausencia de otras proteínas o de RNAs; el prototipo de estos RNAs autocatalíticos es el RNAr 28S del protozoario *Tetrahymena*. Este RNAr contiene un intrón con cerca de 400 nucleótidos que se elimina en ausencia de otras moléculas de RNA o proteína. También se han descrito otros RNAs autocatalíticos en mitocondrias, cloroplastos y bacterias. Dependiendo del tipo de reacción que ocurre durante la eliminación de los intrones estos RNAs autocatalíticos se han dividido en dos clases (Figura 36). En los intrones del grupo I el corte del extremo 5' del intrón ocurre por el ataque de un cofactor de guanosa, luego el extremo 3' libre del exón reacciona con el sitio 3' del intron y el intron se escinde como una molécula lineal. En los del grupo II el corte del extremo 5' del intron se produce como consecuencia del ataque de un nucleótido de adenina ubicado dentro del intrón lo que conduce a la

formación de una estructura intermedia en forma de lazo, la cual se escinde por el corte del extremo 3' del intrón. El mecanismo del autoprocésamiento de los intrones del grupo II es muy similar al descrito en los pre-RNAm, lo que apoya la idea que el corte y empalme de intrones y exones es catalizado por los snRNAs presentes en el espliceosoma.

#### Procesamiento alternativo

El papel central que tiene el corte y empalme en el procesamiento del pre-RNA ha revelado una forma para regular la expresión génica, en la medida que la célula regula la actividad de la maquinaria molecular involucrada en este fenómeno. Además, puesto que la mayoría de los genes eucariotas poseen varios intrones, es posible generar RNAm distintos y por lo tanto proteínas con funciones diferentes, gracias a combinaciones alternativas de los sitios 5' y 3' en los intrones (Figura 37). Este mecanismo de procesamiento alternativo se ha descrito en muchos genes eucariotas, lo cual es importante para obtener una expresión génica específica de tejidos y una regulación del desarrollo embrionario.

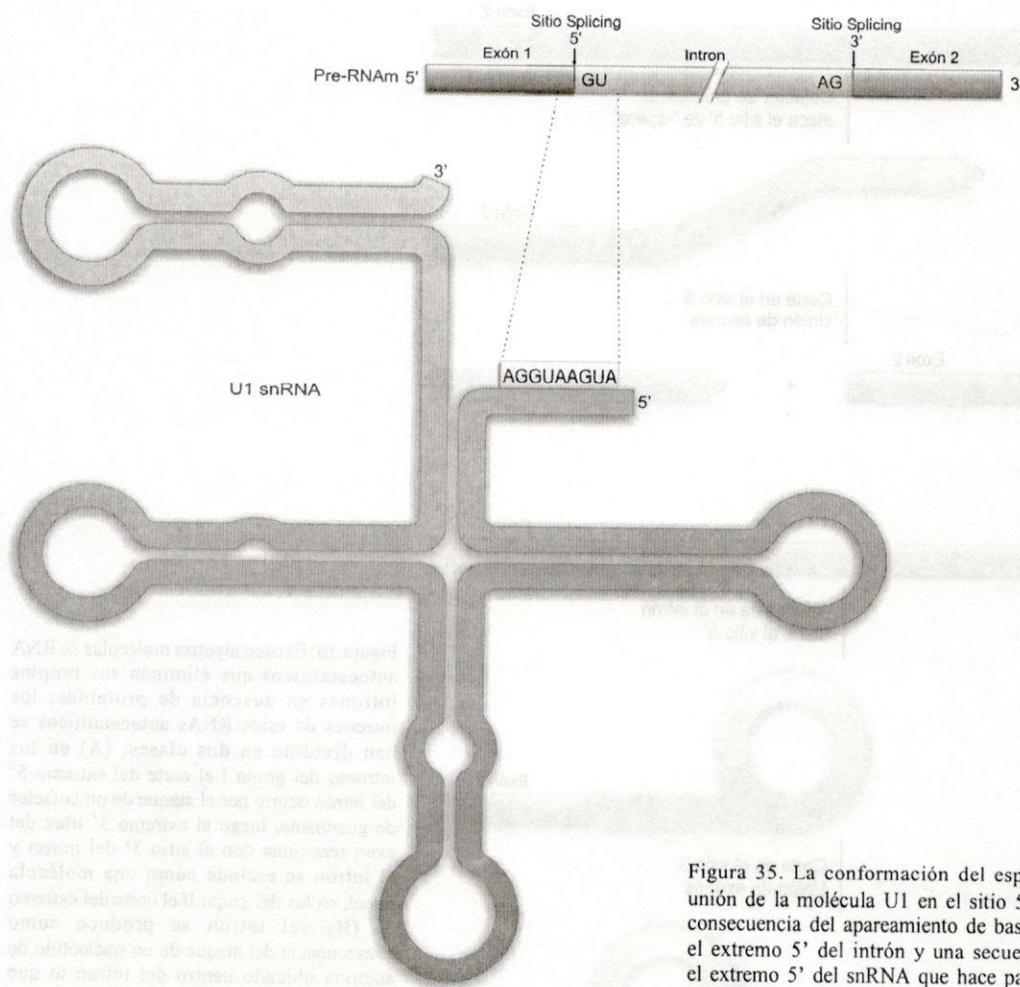


Figura 35. La conformación del espliceosoma se inicia gracias a la unión de la molécula U1 en el sitio 5' del intrón, lo cual ocurre como consecuencia del apareamiento de bases entre la secuencia consenso en el extremo 5' del intrón y una secuencia complementaria presente en el extremo 5' del snRNA que hace parte de la snRNP U1.

Un ejemplo interesante del procesamiento alternativo lo ofrecen algunos genes que codifican proteínas reguladoras de la transcripción. Como se mencionó antes, estas proteínas reguladoras generalmente contienen un dominio de unión al ADN y un dominio de activación. La mayoría de las veces estos dominios se encuentran codificados en exones separados, por lo tanto el procesamiento alternativo permite que el RNAm de estos genes se reorganice en combinaciones diferentes y a la final se puede obtener proteínas con función activadora (si contiene los exones para los diferentes dominios) o represora (si se elimina el exón con el dominio de activación) a partir del mismo gen (Figura 38).

Hasta el momento no se conoce exactamente como se regula el procesamiento alternativo en las células, sin embargo se han identificado proteínas que pueden contribuir a la selección de los sitios de corte de los intrones en el pre-RNAm. La expresión diferencial de estos factores proteicos puede ser fundamental para la regulación de la expresión génica durante el desarrollo y diferenciación, así como para ésta se mantenga específica en cada tejido.

## Edición del RNA

El fenómeno de edición del RNA tiene que ver con mecanismos de procesamiento que permiten modificar las secuencias codificadoras de proteínas, solo que dicha modificación ocurre cuando ya se han generado las moléculas de RNAm. Este mecanismo de procesamiento se evidenció inicialmente en los RNAm de mitocondrias de tripanosomas, en los cuales se agregan y eliminan residuos de uracilo en varios sitios a lo largo de la molécula. Después de esta descripción inicial se ha encontrado que la edición del RNAm puede ocurrir en los RNAm de mitocondrias de otros organismos, en los RNAm de cloroplastos de plantas y en los RNAm que se transcriben a partir de algunos genes nucleares en mamíferos.

La edición de estas moléculas de RNAm involucra cambios de un solo nucleótido, los cuales se producen como consecuencia de reacciones de modificación similares a las que ocurren durante el procesamiento del RNAt. Se ha demostrado que las reacciones de edición de RNAm

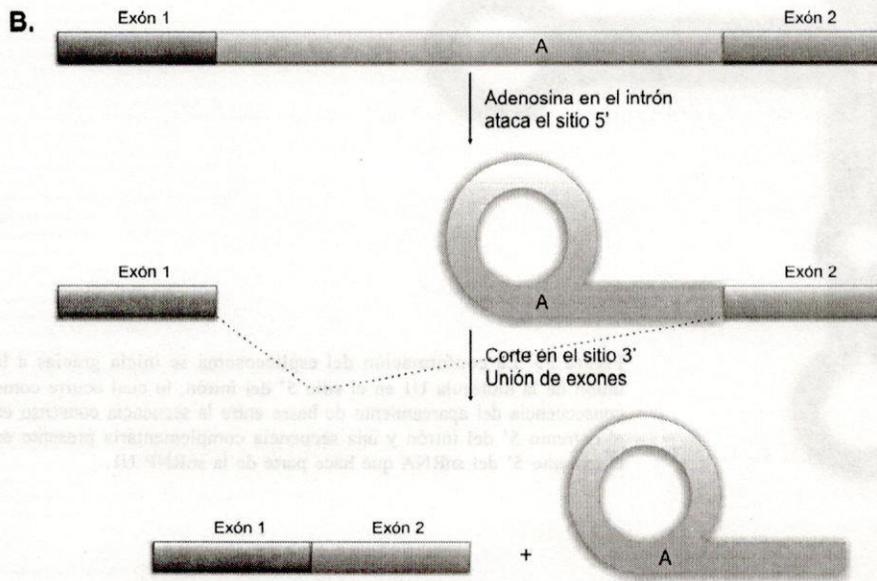
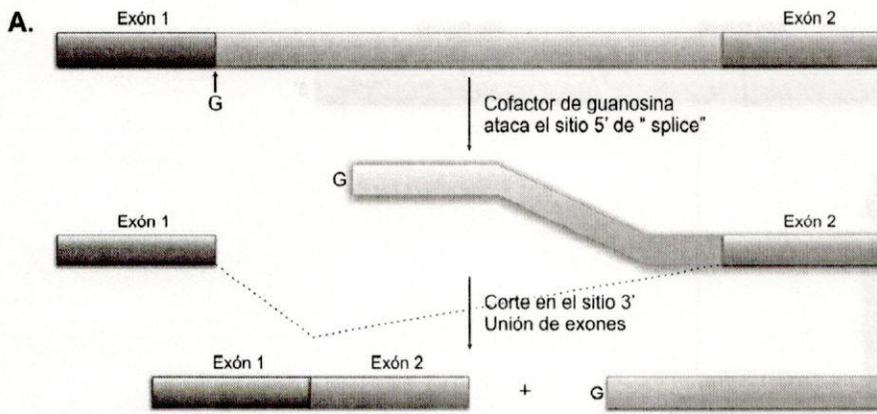


Figura 36. Existen algunas moléculas de RNA autocatalíticos que eliminan sus propios intrones en ausencia de proteínas; los intrones de estos RNAs autocatalíticos se han dividido en dos clases: (A) en los intrones del grupo I el corte del extremo 5' del intrón ocurre por el ataque de un cofactor de guanosina, luego el extremo 3' libre del exón reacciona con el sitio 3' del intron y el intrón se escinde como una molécula lineal; en los del grupo II el corte del extremo 5' (B) del intron se produce como consecuencia del ataque de un nucleótido de adenina ubicado dentro del intron lo que conduce a la formación de una estructura intermedia en forma de lazo, la cual se escinde por el corte del extremo 3' del intrón.

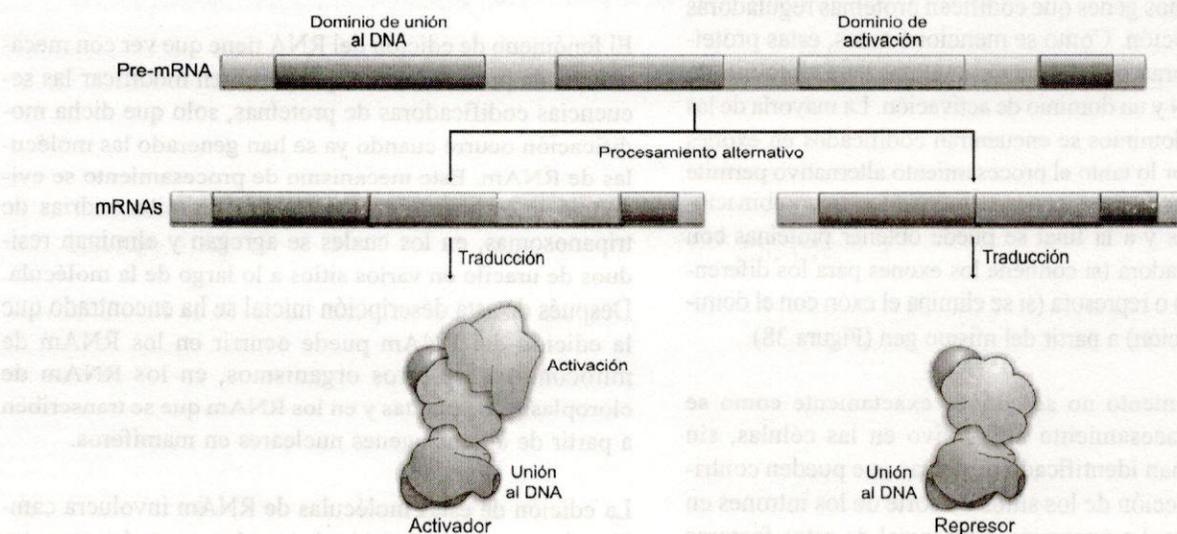


Figura 37. El procesamiento ("splicing") alternativo explica en gran medida la gran variedad de productos generados a partir de la limitada información genética de la secuencia que constituye un gen. El hecho que los genes eucariotas sean fragmentados por la presencia de intrones, permite que a partir de un solo gen se generen diferentes RNAm como consecuencia de la combinación alternativa entre los sitios 5' y 3' en los intrones.

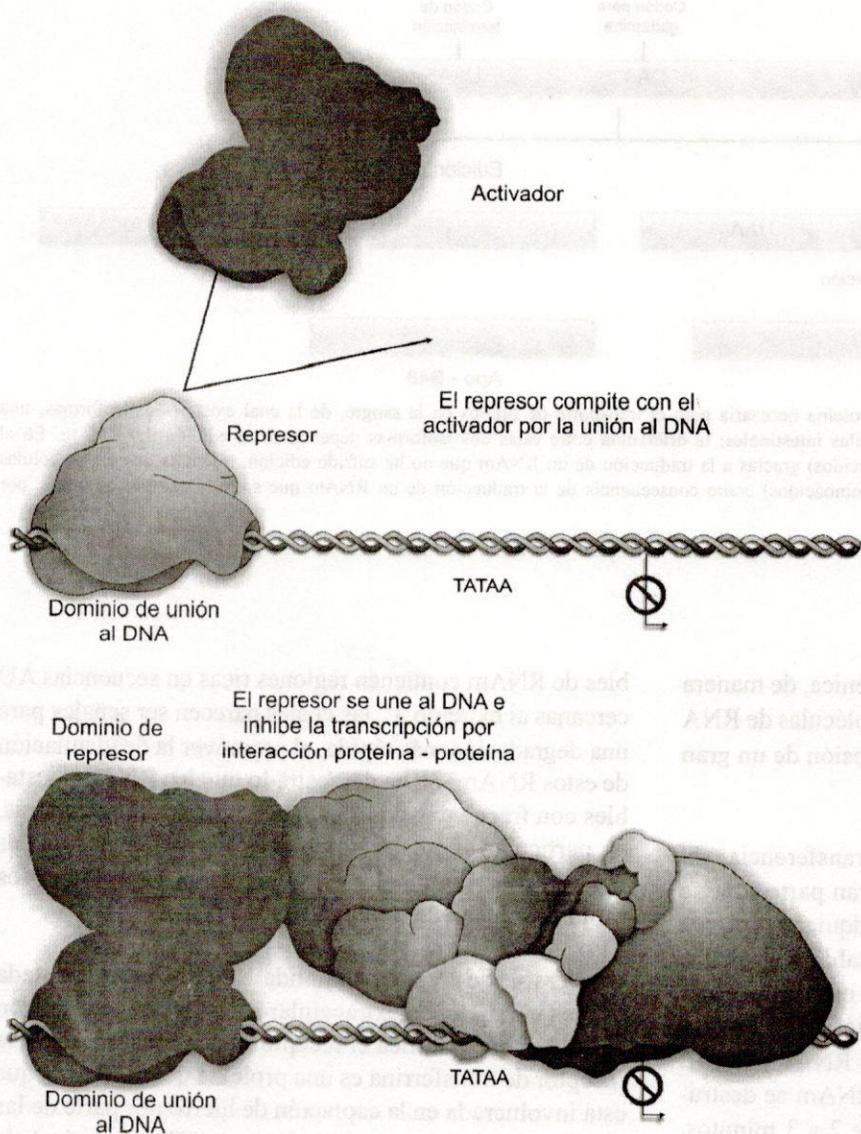


Figura 38. En esta figura se presenta un ejemplo de procesamiento alternativo, en el cual un RNAm se reorganiza en combinaciones diferentes de los exones, lo cual permite que los dominios que conforman la proteína den origen a una molécula con función activadora o represora a partir de una misma secuencia genética.

que tienen lugar en células mamíferas incluyen la desaminación de citosina a uridina y de adenosina a inosina.

En los humanos se ha descrito un caso claro de edición, el cual ocurre en el RNAm para la apolipoproteína B (Figura 39). Esta proteína se encarga del transporte de lípidos en el torrente circulatorio y es producida tanto por células del hígado como del intestino; sin embargo, estos dos tipos de células producen una isoforma de apolipoproteína B que depende de la edición del RNAm en cada una de ellas. En el hígado se sintetiza la Apo B100 (tiene 4536 aminoácidos) gracias a la traducción de un RNAm que no ha sufrido edición, mientras que en las células del intestino se sintetiza la Apo-B48 (tiene 2152 aminoácidos) como consecuencia de la traducción de un RNAm que ha sufrido el cambio de una C por una U. Esta desaminación cambia el codón que codifica para glutamina (CAA) por

un codón de terminación de la traducción (UAA), lo que conduce a la síntesis de una proteína mucho más corta. De esta manera un mecanismo de edición de RNAm específico de tejido da origen a la expresión de proteínas que son estructural y funcionalmente diferentes, pues la Apo-B100 de origen hepático transporta lípidos en circulación, mientras la Apo-B48 es necesaria para la absorción en el intestino de los lípidos presentes en la alimentación.

#### Degradación del RNA

Los niveles de RNA presentes en la célula van a ser determinantes para definir el grado de expresión de una proteína determinada y a su vez la función que dicha proteína cumple. Por tanto, el balance que existe entre la síntesis y la degradación de las moléculas de RNA es un factor que

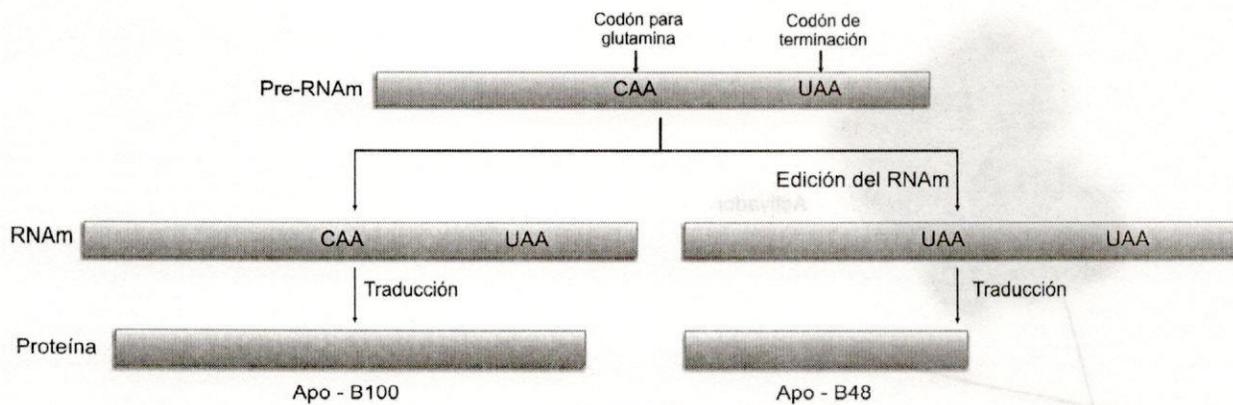


Figura 39. La apolipoproteína B humana es una proteína necesaria para el transporte de lípidos en la sangre, de la cual existen dos isoformas, una es producida por células del hígado y otra por células intestinales; la diferencia entre estas dos isoformas depende de la edición del RNAm. En el hígado se sintetiza la Apo B100 (tiene 4536 aminoácidos) gracias a la traducción de un RNAm que no ha sufrido edición, mientras que en las células del intestino se sintetiza la Apo-B48 (tiene 2152 aminoácidos) como consecuencia de la traducción de un RNAm que sufre el cambio de una C por una U gracias a un proceso de desaminación.

determina en gran medida la expresión génica, de manera que la regulación de la degradación de moléculas de RNA puede ser un punto de control de la expresión de un gran número de genes.

Las moléculas de RNA ribosomal y de transferencia son muy estables, lo cual tiene que ver en gran parte con las conformaciones estructurales que éstas adquieren una vez son sintetizadas. A su vez esta mayor estabilidad explica en gran medida la alta concentración de estas moléculas tanto en células eucarióticas como procarióticas, pues más del 90% de todo el RNA corresponde a RNAr y RNAt. Por el contrario, en los procariotes los RNAm se destruyen rápidamente, con una vida media de 2 a 3 minutos. Aunque esta mayor labilidad del RNA bacteriano se puede ver como una desventaja, no lo es tanto, pues permite que estos microorganismos cambien rápidamente los patrones de expresión génica ante las condiciones variables del ambiente donde se encuentren. De otro lado, en las células eucarióticas los RNAm son degradados a velocidades diferentes, lo cual permite que existan mecanismos adicionales para la regulación de la expresión génica en estas células.

La degradación de la mayoría de las moléculas de RNAm eucarióticas se inicia por medio del acortamiento de las colas poli-A; luego se elimina el casquete del extremo 5' de la molécula y posteriormente actúan las nucleasas en ambos extremos de la molécula para despolimerizar el RNA. La vida media de las moléculas de RNAm en las células mamíferas varía desde menos de 30 minutos hasta un poco más de 20 horas. Por lo general, las moléculas menos esta-

bles de RNAm contienen regiones ricas en secuencias AU cercanas al extremo 3', las cuales parecen ser señales para una degradación más rápida, al promover la deadenilación de estos RNAm. Se ha demostrado que los RNAm inestables con frecuencia codifican para proteínas reguladoras, en particular ciertos factores de transcripción, los cuales deben variar sus niveles en la célula en respuesta a los estímulos del medio ambiente celular.

Un ejemplo de como la estabilidad del RNAm es regulada por las condiciones extracelulares, es lo que sucede con el RNAm que codifica el receptor de la transferrina. Este receptor de transferrina es una proteína de membrana que está involucrada en la captación de hierro por parte de las células mamíferas, cuyo nivel de expresión depende de la disponibilidad de hierro para la célula. En presencia de cantidades adecuadas de hierro, el RNAm del receptor de transferrina se degrada rápidamente como consecuencia del corte específico por una nucleasa cerca al extremo 3' de la molécula. Sin embargo, cuando disminuye el suministro de hierro y los niveles celulares son bajos se aumenta el número de receptores de transferrina en la membrana plasmática como consecuencia de una mayor estabilidad del RNAm y por lo tanto una traducción de más moléculas de este receptor. Esta mayor estabilidad depende de la unión de una proteína a secuencias específicas cercanas al extremo 3' de la molécula de RNAm, conocida como elemento de respuesta al hierro (IRE), la cual la protege de la acción de la nucleasa. La unión de dicha proteína a la región IRE del RNAm es controlada por los niveles celulares de hierro, a medida que disminuye la concentración de hierro en la célula, la proteína se une al RNAm y lo protege de su degradación.