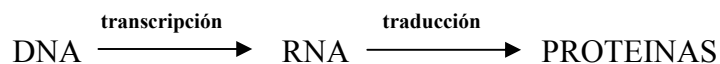
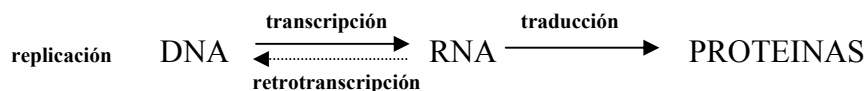


LA TRANSCRIPCIÓN

La información contenida en el DNA debe permitir la posterior síntesis de una proteína en la cual los distintos aminoácidos (AA) estén colocados en el orden dictado por su secuencia de bases, a pesar de que no hay ningún tipo de afinidad entre el DNA y los AA. Ya en 1953 se aceptaba que el **flujo de información genética** comienza a partir del DNA cromosómico. Este DNA sirve de molde para la síntesis del RNA (**transcripción**), que será el que determine la secuencia de AA de una proteína (**traducción**). Esta hipótesis constituye el **dogma central de la Biología**:



Este esquema puede completarse si incluimos los procesos en los que se copia la información contenida en el DNA (**replicación**) y los procesos que permiten la síntesis de DNA usando como molde al RNA (**retrotranscripción** o **transcripción inversa**), que utilizan los virus cuya información genética se almacena en forma de RNA:



En los organismos procariotas, los procesos de transcripción y traducción no están separados ni en el tiempo ni en el espacio: una molécula de RNA puede estar siendo traducida en una secuencia de proteína antes de que haya sido totalmente transcrita a partir del DNA (Figura 1). Por el contrario, **en los organismos eucariotas**, los procesos de transcripción y traducción están separados en el tiempo y en el espacio: la transcripción tiene lugar en el núcleo, y cuando el RNA llega al citoplasma podrá traducirse en una secuencia proteica (Figura 1).

TIPOS DE MOLECULAS DE RNA

La estructura química del RNA es muy parecida a la del DNA. Es una molécula larga, no ramificada, que contiene 4 tipos distintos de ribonucleótidos unidos por enlaces 3',5'-fosfodiéster. Hay tres características que lo diferencian del DNA. La primera es que el azúcar que contiene el RNA es **ribosa**, que únicamente se diferencia de la desoxirribosa del DNA por la presencia de un grupo OH en posición 2' del anillo de pentosa. La segunda diferencia radica en la composición de las bases nitrogenadas. El RNA no contiene timina, y en su lugar aparece el **uracilo**, de estructura química parecida. En tercer lugar, el RNA no forma una doble hélice como el DNA, sino que existe como una **hebra sencilla** de polirribonucleótido que puede presentar zonas de apareamientos locales dentro de la misma molécula.

En un principio se pensó que todo el RNA celular serviría como molde para la síntesis de proteínas. La cadena de RNA podría plegarse de forma que un determinado AA pudiera acceder a las cavidades formadas por la molécula de RNA, determinando así la secuencia de la proteína. Esta teoría resultó ser falsa fundamentalmente por el

hecho de que muchos AA no presentan ningún tipo de afinidad hacia las bases del RNA, y porque sería muy difícil seleccionar con precisión entre AA que tengan una estructura parecida. Se supuso, por tanto, que antes de incorporarse a las proteínas, los AA deberían unirse a moléculas adaptadoras capaces de unirse específicamente a las bases del RNA molde. Así, a medida que avanzaron los estudios sobre la traducción se puso de manifiesto que en la célula existen **tres clases de moléculas de RNA** mayoritarias: el RNA mensajero (**RNA_m**), el RNA ribosómico (**RNA_r**) y el RNA de transferencia (**RNA_t**). Todas ellas se sintetizan a partir de moléculas de DNA, en las regiones conocidas como genes de RNA. Las moléculas de RNA recién sintetizadas reciben el nombre de **transcrito primario**. En muchos casos, el transcrito primario no es la forma final de la molécula de RNA, sino que sufre una serie de modificaciones post-transcripcionales denominadas **procesamiento** o **maduración** del RNA (*RNA splicing*).

Aproximadamente el 85% del RNA celular se encuentra en los ribosomas, y como el número de ribosomas aumenta cuando la célula debe sintetizar proteínas, se pensó que **el RNA ribosomal** (RNA_r) serviría de molde para la síntesis proteica. El análisis posterior del RNA_r desechó esta hipótesis ya que se descubrió que las moléculas de RNA_r presentan muy poca variación tanto en su tamaño como en su composición de bases nitrogenadas. Si el RNA_r sirviese como molde para la síntesis de proteínas, su tamaño y su secuencia de bases variaría en función de la proteína que tuviese que codificar. El RNA_r es una molécula de hebra sencilla, y que presenta zonas donde hay apareamiento de bases complementarias, que forman estructuras en forma de horquilla. Estas estructuras se pliegan dando lugar a complejas estructuras secundarias que se conservan fielmente a lo largo de la evolución (Figura 2). El RNA_r está implicado en la unión a (1) otras proteínas, para formar el ribosoma, (2) al RNA_m que sirve de molde para la síntesis de proteínas, (3) al RNA_t, que es la molécula adaptadora, y (4) al RNA_r de la subunidad opuesta del ribosoma.

Zamecnik y Hoagland descubrieron que antes de incorporarse a las proteínas, los AA se unen a unas moléculas pequeñas de RNA al que llamaron **RNA de transferencia** (RNA_t). Este RNA constituye el 10% del RNA celular total. Cada AA se une a un RNA_t diferente. Las moléculas de RNA_t tienen entre 73 y 93 nucleótidos (peso molecular de 25-30 KDa). Una característica típica de los RNA_t es que contienen un alto contenido de bases atípicas (distintas de A, C, G y U). Estas bases atípicas a menudo contienen uno o más grupos metilo, que son añadidos enzimáticamente a las moléculas precursoras del RNA_t. La secuencia primaria de los RNA_t adquiere una conformación característica en la cual ciertos nucleótidos se aparean por complementariedad de bases, dando lugar a una estructura conocida con el nombre de estructura en hoja de trébol (Figura 3). En ella se distinguen **3 zonas** críticas: (1) el extremo 3' es siempre CCA, y es esta última A la que se une al extremo carboxilo del AA (2) una región especializada en la unión al ribosoma y (3) una región llamada anticodón, que es la que interacciona con la molécula de RNA molde. La estructura tridimensional del RNA_t se ha podido resolver, y muestra una estructura en forma de L, donde uno de los extremos contiene el AA y el otro el anticodón, separados por una distancia de unos 70 Å (Figura 3).

Los transcritos primarios de las moléculas de RNA_r y RNA_t, después de haber sido sintetizadas sufren una serie de modificaciones post-transcripcionales (**procesamiento del RNA**). Como resultado del procesamiento se ha observado que (1)

las moléculas terminan en un grupo 5'-monofosfato en lugar de 5'-trifosfato, (2) las moléculas de RNA_r y RNA_t son más pequeñas que sus transcritos primarios y (3) las moléculas de RNA_t contienen bases distintas a las normales (A, G, C, U), que no se encuentran en el transcrito primario.

El resto del RNA celular (<10%) es el que sirve de molde para la síntesis de proteínas. Como es el que transfiere la información genética del gen al ribosoma se le ha llamado **RNA mensajero** (RNA_m). Este RNA se une reversiblemente al ribosoma de forma que su secuencia va siendo traducida a proteína a medida que se desplaza a través de él. Su tamaño y su secuencia varían en función de la proteína que codifica. La molécula de RNA_m se traduce siempre a partir del extremo 5'. El RNA_m contiene también secuencias específicas para la unión al ribosoma, así como las señales para el inicio y final de la traducción. A veces la síntesis de una proteína no comienza en el extremo 5'-fosfato del RNA_m, sino a cierta distancia de éste (a veces hasta varios cientos de bases a partir del extremo). El tramo de RNA_m situado delante de las regiones codificantes y que no se traduce se llama **secuencia líder o secuencia conductora**. En algunos casos, la secuencia conductora contiene una región reguladora llamada **atenuador**, que determina la velocidad de la síntesis de proteínas.

El segmento del DNA que contiene la información para la síntesis de una cadena polipeptídica (incluyendo las señales para el inicio y el final de la traducción) se llama **cistrón**. Si una molécula de RNA_m codifica un único polipéptido se llama **RNA_m monocistrónico**. Si el RNA_m codifica varias proteínas diferentes, se llama **RNA_m policistrónico** (este tipo de RNA_m sólo se encuentra en procariotas). Una característica importante de los RNA_m de procariotas es que tienen una vida media corta, pues se degradan a los pocos minutos de haber sido sintetizados, y por lo tanto, la síntesis continuada de una proteína requiere a su vez la síntesis continua del RNA_m correspondiente.

A) LA TRANSCRIPCIÓN EN PROCARIOTAS

La transcripción (T) constituye el primer paso en la expresión y consiste en la síntesis de una molécula de RNA complementaria a una secuencia de DNA que actúa como molde para su síntesis. La molécula de RNA tendrá una secuencia idéntica a la hebra complementaria al molde de DNA. La hebra del DNA que se transcribe en una molécula de RNA se llama **hebra codificante, hebra con sentido o hebra (-)**, mientras que la hebra que no se transcribe se llama **hebra sin sentido o hebra (+)** (Figura 7). Este proceso está catalizado por un enzima llamado **RNA-polimerasa**.

El producto inmediato de la T se llama **transcrito primario**. Esta molécula de RNA es casi siempre inestable. En procariotas es rápidamente degradada (RNA_m) o modificada químicamente para dar los productos maduros (RNA_r o RNA_t).

Se define como **punto de partida de la T** al par de bases del DNA que corresponde al primer nucleótido incorporado al RNA y a menudo se encuentra dentro del promotor. Se le designa como posición +1. El punto de partida es en más del 90% de los casos una purina (A o G). Las secuencias anteriores al punto de partida se dice que están **corriente arriba** (*upstream*), y las que están después están **corriente abajo**

(*downstream*) (Figura 8). Las secuencias se escriben por convención de forma que la T se lleva a cabo de izquierda (*upstream*) a derecha (*downstream*), o sea, que el RNA_m se escribe en dirección 5'→3'.

1.- LA RNA-POLIMERASA BACTERIANA

La **RNA-polimerasa** mejor estudiada es la de *Escherichia coli*. Como en la mayoría de las bacterias, esta polimerasa sintetiza RNA_m, RNA_r y RNA_t. Una célula de *E. coli* contiene unas 7000 moléculas de polimerasa, de las cuales entre 2000 y 5000 pueden estar funcionando al mismo tiempo, según el estado de crecimiento de la célula. La RNA-polimerasa utiliza una molécula de **DNA como molde** para la formación de un polirribonucleótido a partir de ribonucleótidos trifosfato. La transcripción se inicia en unas secuencias especiales del DNA llamadas **promotores**.

La **formación del enlace fosfodiéster** requiere el ataque hidrofílico del grupo OH en posición 3' del último ribonucleótido de la cadena naciente de RNA sobre el grupo fosfato en posición α del nuevo ribonucleótido. En esta reacción se libera pirofosfato (Figura 4). La dirección de crecimiento del RNA es **5'→3'**. Esta reacción es rápida y a 37 °C es de unos 40 nucleótidos por segundo. A diferencia de las DNA polimerasas, la RNA polimerasa **no necesita de un cebador** para iniciar la síntesis y **carece de actividad exonucleasa 3'→5'** (función correctora).

La RNA polimerasa consta de **varias subunidades** (Figura 5). El enzima completo, también llamado **holoenzima** de *E. coli*, tiene un peso molecular de 465 KDa y contiene 4 subunidades distintas (α , β , β' y σ). La composición del holoenzima es $\alpha_2\beta\beta'\sigma$. **Las subunidades β y β'** forman el centro catalítico del enzima. **La subunidad σ** es necesaria para reconocer los promotores y **las subunidades α** están implicadas en (1) el ensamblaje correcto de todas las subunidades, (2) en el reconocimiento de los promotores y (3) en la interacción con los factores de regulación.

El holoenzima puede disociarse, formando por un lado **el núcleo del enzima** ($\alpha_2\beta\beta'$) y por otro **el factor sigma** (σ) (Figura 6). Sólomente el holoenzima puede iniciar la transcripción (T). Cuando la cadena de RNA naciente alcanza los 8-9 nucleótidos de longitud, el factor sigma se desprende, y la elongación continúa con el núcleo del enzima.

2.- ETAPAS DE LA TRANSCRIPCION

La Ttranscripción puede dividirse en cuatro fases tal y como se muestra en la Figura 9:

- **1.- Unión al molde de DNA:** La RNA-polimerasa se une al DNA de doble hebra. La zona del DNA a la que se une la polimerasa se llama **promotor**. Las dos hebras del DNA deben separarse para que una de ellas pueda servir de molde a la RNA polimerasa. Este desenrollamiento local se produce en el lugar donde se une la RNA-polimerasa al DNA.

- **2.- Iniciación:** Corresponde a la síntesis del primer enlace fosfodiéster del RNA. El enzima permanece unido al promotor mientras sintetiza un primer polirribonucleótido formado por 9 ribonucleótidos. En algunos casos este RNA de 9 bases puede desprenderse del DNA, con lo cual el proceso se interrumpe. Es lo que se llama una **iniciación abortiva**. Si esto ocurre, la síntesis de RNA debe comenzar de nuevo. Si se supera la fase de iniciación abortiva, la iniciación se da por terminada, el enzima se separa del promotor y comienza la siguiente fase.
- **3.- Elongación:** el enzima se desplaza a lo largo del DNA, al tiempo que la cadena de RNA naciente va creciendo. A medida que la polimerasa se desplaza, las dos hebras del DNA deben separarse para que una de ellas actúe como molde. En la zona desenrollada se forma un híbrido DNA-RNA entre el DNA molde y los ribonucleótidos recién añadidos por la polimerasa. Por detrás de la zona desenrollada, el DNA vuelve a adoptar su estructura de doble hélice.
- **4.- Terminación:** tiene lugar cuando la polimerasa llega a una zona del DNA denominada **terminador**. En este momento se interrumpe la formación de enlaces fosfodiéster y el complejo de T se disocia: el híbrido DNA-RNA se deshace, el transcrito primario de RNA y la RNA polimerasa se separan del DNA y el DNA vuelve a formar la doble hélice.

2.1.- UNION AL PROMOTOR

La unión de la RNA-polimerasa al promotor provoca la iniciación de la T en el punto de partida. La RNA-polimerasa continúa hasta encontrar una secuencia de terminación. Este proceso define la llamada **unidad de T**, que se extiende desde el promotor hasta el terminador (Figura 10). La unidad de T puede contener uno o varios genes. Cuando la unidad de T contiene un solo gen, originará un **RNA_m monocistrónico**, y si contiene varios genes, dará lugar a un **RNA_m policistrónico** (ya hemos dicho anteriormente que un cistrón es un segmento del DNA que contiene toda la información necesaria para sintetizar una proteína).

En bacterias, a menudo existe un único promotor para una serie de genes adyacentes. Este conjunto de genes que es controlado por un único promotor recibe el nombre de **operón** (Figura 11) y se transcribe como un único RNA_m (RNA_m policistrónico). Los genes de un operón a menudo codifican enzimas de **la misma ruta metabólica** y por lo tanto se controlan de forma coordinada: o se transcriben todos o no se transcribe ninguno. Cada gen del operón está representado en el RNA_m y su traducción en proteína se realiza de forma independiente. Cada cistrón contiene sus propios lugares de unión al ribosoma, su codón de iniciación y su codón de terminación y por lo tanto, la velocidad de traducción es distinta para cada uno.

A veces, los genes que codifican los enzimas necesarios para una determinada ruta metabólica no están adyacentes, sino que están localizados en distintas regiones del genoma. Este grupo de genes esparcidos por el genoma y cuya T está regulada de forma coordinada se llama **regulón**. Es el caso de los eucariotas y de los genes que codifican las proteínas necesarias para la síntesis de arginina (un AA).

El **reconocimiento del promotor** por parte del holoenzima de la RNA-

polimerasa comprende dos fases. En primer lugar, el holoenzima y el promotor forman un **complejo binario cerrado** y en segundo lugar, este complejo se convierte en un **complejo binario abierto** mediante la fusión de un pequeño fragmento del DNA implicado en la unión con el enzima. De esta forma se logra una unión fuerte entre el enzima y el DNA (Figura 12).

La RNA-polimerasa debe encontrar el promotor, de unas 60 pb en el contexto de un genoma que en *E. coli* puede ser de unas 4×10^6 pb. ¿Cómo es posible este reconocimiento?. El modelo más sencillo es suponer que la RNA-polimerasa se desplaza por **difusión al azar**. El holoenzima se une y se disocia del DNA muy rápidamente de las secuencias de que no contienen el promotor (secuencias de unión débil). Cuando por azar encuentra al promotor, se forma un complejo abierto en el que el enzima y el DNA se encuentran fuertemente unidos. En este modelo, el movimiento del enzima de un punto a otro está limitado por su velocidad de difusión. Para un promotor de 60 pb la constante de velocidad de asociación a la RNA-polimerasa sería $<10^{-8} \text{ M}^{-1}\text{s}^{-1}$. Sin embargo, el valor calculado *in vitro* es mayor que este límite teórico.

El proceso sería más rápido si el enzima permaneciese asociado al DNA. Así, para moverse de un lugar de unión débil a un promotor (1) el enzima va **intercambiando secuencias** de unión al DNA de manera indiscriminada hasta encontrar al promotor o (2) el enzima **se desplaza** a lo largo del DNA hasta encontrar el promotor. Es probable que el mecanismo *in vivo* sea una combinación de estos modelos (Figura 13).

CARACTERÍSTICAS DEL PROMOTOR

A partir del análisis de cientos de promotores distintos, se han encontrado en todos ellos **cuatro características comunes**. Estas características son fundamentales para el correcto funcionamiento de la RNA-polimerasa.

La primera característica común a todos los promotores estudiados es que en más del 90% de los casos, la base que está presente **en el punto de partida es una purina** (A o G).

La segunda característica común es una secuencia de 6 pb (TATAAT) cuyo centro se sitúa aproximadamente 10 nucleótidos corriente arriba del punto de partida (**secuencia -10**). Esta secuencia se llama también **caja de Pribnow** y presenta un alto grado de conservación.

La tercera característica común es otra secuencia de 6 pb (TTGACA) se halla centrada 35 nucleótidos corriente arriba del punto de partida (**secuencia -35**).

La cuarta característica común es la distancia que hay entre la secuencia -10 y la secuencia -35. Esta **secuencia espaciadora** tiene una longitud que oscila entre 16 y 18 nucleótidos. La secuencia de este fragmento no es importante, pero su longitud sí, por lo que se piensa que contribuye al correcto acoplamiento entre el promotor y la RNA-polimerasa.

La función de la secuencia -35 es la de reconocer a la RNA-polimerasa. La secuencia -10 está implicada en la formación del complejo abierto, ya que al ser rica en

pares A-T, la energía requerida para separar las dos hebras complementarias del DNA es menor. Por tanto, **las secuencias -35 y -10 del promotor influyen en el reconocimiento de la RNA-polimerasa y en el desenrollamiento de la doble hélice de DNA respectivamente** (Figura 15). Para que estas dos secuencias sean funcionales, es muy importante que estén separadas entre sí por una **secuencia espaciadora** que tenga una longitud de entre 16 y 18 nucleótidos.

2.2.- INICIACION

Una vez formado el complejo binario abierto mediante la fusión del DNA implicado en la unión con el holoenzima, se inicia la incorporación de los nucleótidos. Cuando se han unido los 8 ó 9 primeros nucleótidos, el enzima comienza a desplazarse a lo largo del DNA, dejando atrás al factor σ , formando **un complejo ternario DNA-polimerasa-RNA** (Figuras 9 y 12). A veces, el proceso de iniciación es interrumpido por la liberación del fragmento inicial de 8-9 nucleótidos de RNA (**iniciación abortiva**). En este caso, el enzima comienza a sintetizar de nuevo este fragmento hasta que se supera la fase de iniciación abortiva. En general, este proceso dura entre 1 y 2 segundos, de forma que la frecuencia máxima con que se inicia el proceso de T en un promotor es menor de un proceso por segundo.

2.3.- ELONGACION

Una vez superada la fase de iniciación abortiva, la RNA polimerasa abandona el promotor y comienza a desplazarse a lo largo del molde de DNA. Al tiempo que avanza el enzima se desplaza también la región de desenrollamiento local del DNA para exponer un nuevo segmento de DNA de cadena sencilla que sirva como molde. La longitud del segmento de DNA que se desenrolla es distinta en cada etapa del proceso y oscila entre 12 y 20 pb. La longitud del híbrido DNA-RNA es únicamente de unos 2-3 pb, y parece ser que también hay una zona del RNA recién sintetizado que permanece asociada a la polimerasa. A medida que la cadena de RNA se va alargando, se va disociando de la hebra molde de DNA y de la polimerasa, y se encuentra en forma de una molécula lineal de RNA. La zona del DNA que ya ha sido utilizada como molde por la polimerasa vuelve a adoptar la estructura de doble hélice.

2.4.- TERMINACION

La RNA-polimerasa se desplaza a lo largo del DNA sintetizando RNA hasta que encuentra **una secuencia de terminación**. En este punto, el enzima deja de añadir nucleótidos a la cadena de RNA naciente, libera el producto y se disocia del DNA molde. Los puentes de hidrógeno del híbrido DNA-RNA se rompen y se vuelve a formar la doble hélice de DNA (Figura 9).

Las secuencias de terminación en bacterias y en virus se han caracterizado con detalle. Estas secuencias no muestran ningún tipo de similitud más allá del punto donde se añadió la última base del RNA. Por tanto, **el mecanismo de terminación** debe residir en zonas de la secuencia que ya han sido transcritas por la polimerasa.

En *E. coli* se distinguen **dos tipos de terminación**. En un caso, el núcleo de la RNA-polimerasa puede terminar la transcripción *in vitro* sin la ayuda de otros factores. Este mecanismo se llama **no-dependiente de rho**, aunque es posible que intervengan otros factores que aún no conocemos. En el segundo caso, la **terminación depende del factor rho** tanto *in vitro* como *in vivo*.

2.4.1.- TERMINACION NO DEPENDIENTE DE Rho

Las secuencias de terminación no dependientes de rho **presentan dos características** estructurales comunes: (1) zonas ricas en las bases G y C que tienden a formar una estructura secundaria en forma de **horquilla** del pelo, y (2) una secuencia de aproximadamente **6 U** al final de ésta (Figura 16). Ambas características son esenciales para la terminación.

La horquilla se genera por el apareamiento entre secuencias convenientemente espaciadas, en las que abundan las bases G y C. Esta horquilla es de longitud variable y parece ser que provoca la terminación porque **ralentiza** o incluso detiene momentáneamente la síntesis de RNA por la RNA-polimerasa. Cuando la polimerasa se ha detenido durante un intervalo de unos 60 segundos, se desencadena la terminación. Es probablemente **la secuencia de 6 U** la que proporciona la señal para que **la RNA-polimerasa se disocie del molde de DNA** y para que el híbrido DNA-RNA se separe, ya que la unión entre el U del RNA y la A del DNA es particularmente débil (Figura 17). Este mecanismo es el que se produce en el operon *trp*. La importancia de la secuencia de U se ha visto confirmada en mutantes. Cuando esta secuencia está acortada o ha desaparecido, la RNA-polimerasa llega a la horquilla y se detiene, pero no se produce la terminación.

2.4.2.- TERMINACION DEPENDIENTE DE Rho

Hay casos en que no se termina correctamente la transcripción *in vitro*. A pesar de que la RNA-polimerasa se detiene en las secuencias de terminación, al cabo de un tiempo se reanuda la síntesis de RNA. Sin embargo, si a este sistema se le añade un extracto crudo celular, la terminación se produce normalmente. A partir de este extracto se ha podido purificar el factor responsable de la terminación, y se le ha llamado **factor rho**.

El factor rho es una proteína de unos 46 kDa. No se sabe cómo provoca la terminación, aunque parece ser que en este proceso, la proteína rho forma un **hexámero** que se une a una zona de entre 50 y 90 bases de la cadena de RNA recién sintetizado (Figura 18). Esta secuencia de RNA es especialmente rica en C y especialmente pobre en G. Una vez que el factor rho se ha unido al RNA se inicia su **actividad ATPasa**. La energía liberada por la hidrólisis del ATP es utilizada para disociar la RNA-polimerasa del molde de DNA y para romper los puentes de hidrógeno del híbrido DNA-RNA (En este caso no existe una secuencia de 6 U que pueda facilitar la separación del híbrido). Esta forma de terminación se encuentra frecuentemente en la transcripción del DNA vírico por RNA-polimerasas bacterianas.

3.- REGULACION DE LA TRANSCRIPCION

La transcripción constituye la primera etapa de la expresión génica. En la mayor parte de los casos, el control de la expresión génica tiene lugar en esta etapa. Para controlar la expresión de un gen, el primer paso (y en muchos casos, el único) consiste en **decidir si un gen determinado se transcribe o no**. ¿Quién decide qué promotor debe ser transcrito por la RNA-polimerasa?. Esta es la pregunta clave sobre la regulación de la expresión génica.

Los promotores son reconocidos por el holoenzima, de manera que todo promotor accesible será transcrito. La accesibilidad de un promotor puede controlarse mediante otras proteínas que actúen directamente sobre el promotor o sobre regiones próximas a él. Son las **proteínas reguladoras** las que deciden si un gen determinado estará o no en condiciones de ser transcrito, porque regulan el acceso de la RNA-polimerasa al promotor. Los genes que codifican las proteínas necesarias para el normal funcionamiento de la célula (metabolismo, crecimiento, división) se llaman **genes estructurales** y se transcriben de forma **constitutiva**, es decir, a una velocidad prácticamente constante. Los genes que codifican proteínas reguladoras se llaman **genes reguladores**, y su velocidad de T **depende de las condiciones ambientales**

En una bacteria, no todos los promotores son accesibles en un momento dado. A menudo **los nutrientes** presentes en el medio en que está creciendo la célula son los que determinan qué promotores son accesibles a la RNA-polimerasa y por tanto qué operones se expresan. Los enzimas que se sintetizan en respuesta a la presencia de sus sustratos en el medio son **enzimas inducidos**, y el proceso se **llama inducción enzimática**. Así, el promotor de un operón que codifica los enzimas del metabolismo de un determinado nutriente sólo será accesible cuando ese nutriente esté en el medio. Entre estos operones se encuentran aquellos implicados en el metabolismo de los azúcares: operón *lac*, operón *ara* y operón *gal*.

A veces, es **la falta de un nutriente** la que induce la síntesis de enzimas. Por ejemplo, muchas bacterias pueden sintetizar todos los aminoácidos (AA) necesarios para su crecimiento, pero los enzimas necesarios para esta síntesis sólo se encuentran en la célula cuando el AA está ausente. Un operón de este tipo es el operón *trp*, que codifica los enzimas necesarios para la síntesis del triptófano. El promotor de este operón sólo será accesible a la RNA polimerasa cuando no haya triptófano en el medio.

Las proteínas reguladoras que impiden la T de un gen se llaman **represores**. Se unen al DNA en el promotor o cerca de él e impiden el acceso de la RNA-polimerasa y la T del gen o del operón que reprimen. La zona del DNA a la que se unen se llama el **operador** y suele estar cerca del promotor, o incluso solapado con él. Los represores pueden unirse a moléculas pequeñas llamadas **efectores**, que determinan la afinidad del represor hacia el DNA. Si el efector disminuye esta afinidad se llama **inductor**, porque permite la T del DNA. Si, por el contrario, se aumenta la afinidad del represor al DNA, el efector se llama **correpresor**.

Encontramos un **ejemplo de inductor en el operón lac** (Figura 19). En este caso, la lactosa se une al represor del operón *lac*, de forma que éste se disocia del DNA permitiendo la T de los genes del metabolismo de la lactosa: **la lactosa es un inductor**. Si hay lactosa en el medio, se induce la síntesis de los enzimas que permiten su

degradación. Si no hay, se reprime la síntesis de estos enzimas, porque no son necesarios.

El operón *trp* nos proporciona un **ejemplo de correpresor** (Figura 20). Cuando no hay triptófano en el medio, el represor del operón *trp* no se une al DNA. En presencia de triptófano, éste se une al represor, aumentando su afinidad hacia el DNA e impidiendo la T del operón: **el triptófano es un correpresor**. O sea, que cuando no hay triptófano en el medio se permite la T de los enzimas que pueden sintetizarlo, mientras que en presencia de este aminoácido se reprime la T de estos enzimas, puesto que no hacen falta.

Hay proteínas reguladoras que activan la T del DNA, y reciben el nombre de **activadores** (Figura 21). Se unen al DNA a la altura del promotor o cerca de él y aumentan la frecuencia de unión entre la RNA-polimerasa y el promotor. Las zonas del DNA a las que se unen los activadores se llaman **lugares de activación**. Una de estas proteínas es AraC, necesaria para la T del operón *ara*.

B) LA TRANSCRIPCIÓN EN EUCARIOTAS

Las características básicas de la estructura del RNA_m y de la transcripción (T) en eucariotas son muy parecidas a las de los procariotas. Sin embargo hay ciertas **diferencias** notables (Figura 22):

1.- El DNA eucariótico está **asociado a las histonas**, formando la cromatina, cuya estructura presenta un alto grado de complejidad. Esta disposición hace que el DNA sea bastante inaccesible a la hora de iniciar la T, ya que se requieren cambios en el empaquetamiento de la cromatina. Este problema no existe en procariotas, ya que no presentan cromatina, y la T en un punto determinado del genoma puede empezar en cuestión de segundos.

2.- En eucariotas existen **tres tipos distintos de RNA-polimerasas**, cada una especializada en la síntesis de un tipo distinto de RNA. La RNA-polimerasa I sintetiza el RNA_r, la RNA-polimerasa II sintetiza RNA_m y la RNA-polimerasa III sintetiza RNA_t.

3.- El transcrito primario del RNA_m de eucariotas casi nunca es una molécula funcional, en el sentido de que no puede traducirse inmediatamente a proteína, como ocurre en procariotas. Este RNA_m precursor producido en el núcleo recibe el nombre de **RNA heterogéneo nuclear** (RNA_{hn}), debido a que su tamaño presenta grandes variaciones, a diferencia de los RNA_r y RNA_t, cuyo tamaño está bien definido. El RNA_{hn} sufre una serie de modificaciones bioquímicas antes de pasar al citoplasma en su forma definitiva. Este conjunto de modificaciones recibe el nombre de **procesamiento** (Figura 23). Después del procesamiento, tanto el extremo 5' como el extremo 3' del RNA_m están modificados. En el extremo 5' se encuentra una estructura compleja denominada **capucha**, y en el extremo 3' hay una cadena de longitud variable (hasta varias cientos de bases) de ácido poliadenílico (**poliA**). Además, durante el procesamiento se elimina una serie de secuencias intermedias que se llaman **intrones**.

Las secuencias restantes, que formarán parte del RNA_m que servirá de molde para la síntesis de proteínas se llaman **exones**. La molécula de RNA_m madura es unas diez veces menor que el transcrito primario del que procede.

4.- Como consecuencia de esta serie de modificaciones, el RNA_m de eucariotas tiene una **vida media larga**, a diferencia de los RNA_m de procariotas, que son degradados en cuestión de pocos minutos.

5.- Todas las moléculas de RNA_m de eucariotas son **monocistrónicas**, a diferencia del RNA_m de procariotas, que es policistrónico (Figura 24). Por tanto, su tamaño es menor: la longitud media de un RNA_m eucariótico es de 1-2 Kb.

LAS RNA-POLIMERASAS DE EUCARIOTAS

Todas las células eucariotas contienen tres RNA-polimerasas distintas y se llaman I, II y III (Figura 25). **La RNA-polimerasa I** se localiza en el nucleolo y sintetiza RNA_r. **La RNA-polimerasa III** funciona en el nucleoplasma y transcribe los genes de RNA_t y otros RNA pequeños como el RNA_r 5S y los RNA pequeños nucleares (RNA_{sn}) implicados en el procesamiento del RNA. **La RNA-polimerasa II** también se encuentra en el nucleoplasma y sintetiza principalmente RNA_{hn}, el precursor del RNA_m.

Las 3 RNA-polimerasas de eucariotas son proteínas grandes (> 500 KDa) y presentan una estructura más compleja que la de *E. coli*. Las más estudiadas son las RNA-polimerasas de *Saccharomyces cerevisiae* (la levadura de la cerveza). Están formadas por **2 subunidades grandes** y un número variable (entre 8 y 14) de **otras subunidades proteicas más pequeñas**. Las dos subunidades grandes tienen un peso molecular de unos 200 y unos 100 KDa, respectivamente. Estas dos subunidades grandes son muy parecidas en los tres tipos distintos de polimerasa, y además también se parecen a las subunidades β y β' de la RNA-polimerasa bacteriana, es decir probablemente representan el centro catalítico del enzima. En cuanto a las subunidades de menor tamaño, su función no se conoce todavía con detalle. Algunas de estas subunidades son muy parecidas entre las distintas polimerasas, pero otras son específicas de cada una (Figura 26).

Se han descrito **varios factores de transcripción** (FT) que son proteínas que ayudan a la polimerasa a localizar los lugares de iniciación y a comenzar la T. Se clasifican en dos tipos: los factores de transcripción generales y los específicos. Los **FT generales** se necesitan siempre, y ayudan en las etapas de (1) iniciación de la T y (2) elongación de la cadena de RNA. Los **FT específicos** sólo se necesitan para la T de determinados genes, y participan en la formación del **complejo de pre-iniciación**. Los complejos de pre-iniciación se forman cuando los FT específicos se unen al DNA en el promotor. De manera facilitan el reconocimiento y la unión de la RNA-polimerasa al promotor.

