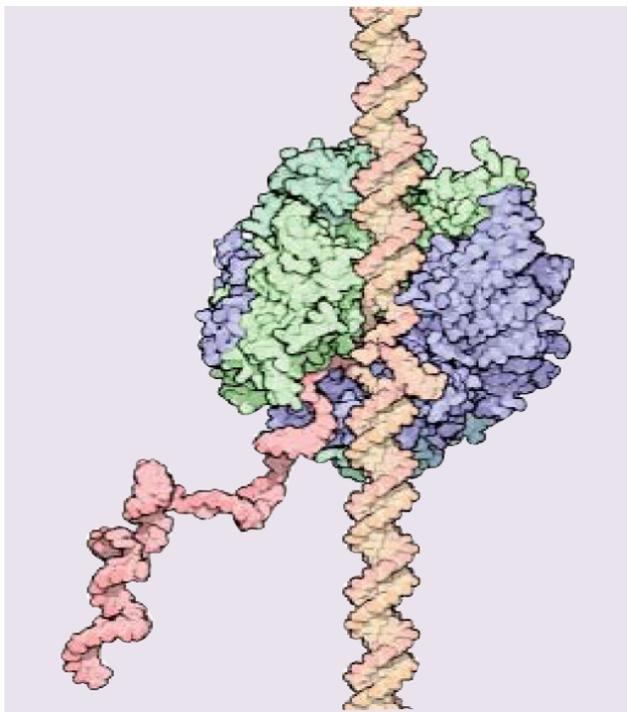


# Biología

## Unidad 6

### La base molecular de la variación genética



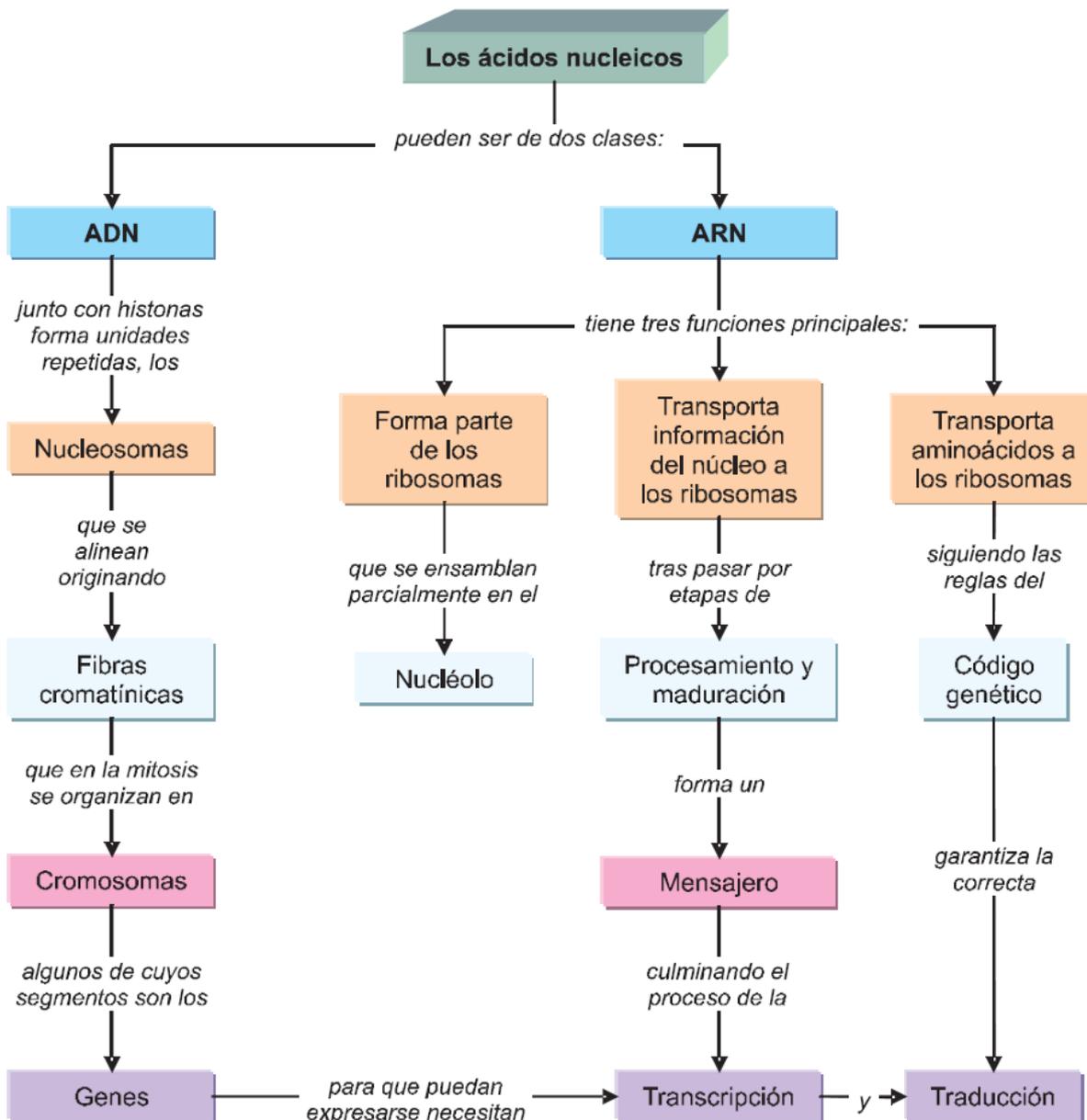
**Ilustración 6.1.** Detalle de la transcripción. La enzima ARN polimerasa, en verde y azul, se halla unida a una molécula de ADN. A la izquierda, molécula de ARNm en crecimiento (Fuente: <http://www.rcsb.org/pdb>).

En la Unidad 5 desarrollamos las leyes de la herencia pero quedaron, sin embargo, varias incógnitas por esclarecer: ¿De qué material están hechos los genes? ¿Cómo almacenan las “instrucciones” necesarias para construir una proteína? ¿Mediante qué mecanismos logran los genes, localizados en el interior del núcleo, comunicar a los ribosomas del citoplasma dichas instrucciones? ¿Cómo ejecutan los ribosomas las órdenes recibidas? ¿Y qué portentoso sistema regulador facilita que solo se sinteticen las proteínas adecuadas en el momento justo?

Responder a estas cuestiones requerirá cierto esfuerzo de abstracción, pues exige descender hasta el último peldaño en los dominios de la Biología: el mundo subcelular y molecular. Allí encontraremos gigantescas moléculas, como el ADN, que se retuercen, se pliegan y se estiran, nanomáquinas que cortan moléculas de ARN y “barajan” los trozos, moléculas cooperadoras y egoístas.

# Índice

<b>1. El portador de la información genética</b>	<b>232</b>
1.1. Estructura y organización del ADN	234
1.2. Organización del ADN en el núcleo de las células eucariotas	237
1.3. Organización del ADN en procariotas	239
Actividades	240
<b>2. La transcripción</b>	<b>241</b>
2.1. El “dogma” central de la Biología molecular	241
2.2. Redacción de mensajes genéticos	245
Actividades	253
<b>3. La traducción</b>	<b>254</b>
3.1. El código genético	254
3.2. La síntesis de proteínas	257
3.3. Regulación de la expresión génica	263
Actividades	266
Solucionario	268
Glosario	271
Bibliografía	272



Con el estudio de esta Unidad nos proponemos alcanzar los siguientes objetivos:

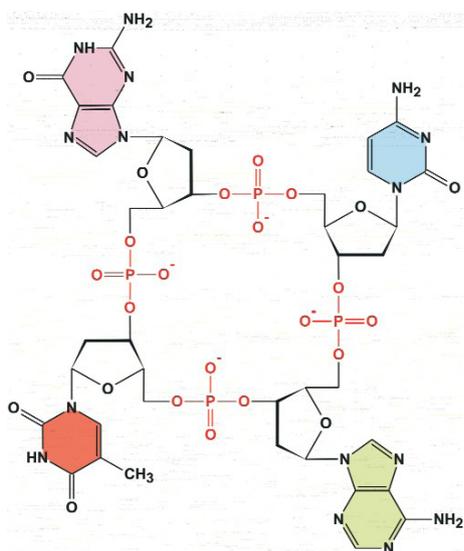
1. Explicar la estructura y la función de los distintos tipos de ácidos nucleicos, y su organización en eucariotas y procariotas.
2. Interpretar las experiencias que condujeron a averiguar la estructura del ADN y su papel como depositario de la información genética.
3. Relacionar el concepto de gen a nivel molecular con su definición en la genética clásica, y generalizar la expresión de la información genética como paradigma de la biología molecular.
4. Describir el proceso de transcripción, diferenciando entre procariotas y eucariotas.
5. Explicar el concepto y las características del código genético.
6. Describir el proceso de traducción, diferenciando entre procariotas y eucariotas.
7. Resolver ejercicios prácticos de transcripción, de aplicación del código genético y de elaboración e interpretación de esquemas de los procesos citados.

# 1. El portador de la información genética

En la Unidad 5 estudiamos la composición química de los ácidos nucleicos; ahora tenemos que responder a dos interrogantes: ¿dónde se localizan? y ¿cuál es su papel en las células?

La primera pregunta fue contestada por el citólogo y genetista sueco Torbjörn Oskar Caspersson (1910-1997). En 1936, Caspersson eliminó selectivamente el ARN de las células y fotografió la célula con luz ultravioleta, que los ácidos nucleicos absorben con intensidad; determinó así que el ADN —el único ácido nucleico que quedaba— se hallaba *exclusivamente* en los cromosomas. Experimentos similares mostraron que el ARN se localizaba principalmente en el citoplasma y en el nucléolo.

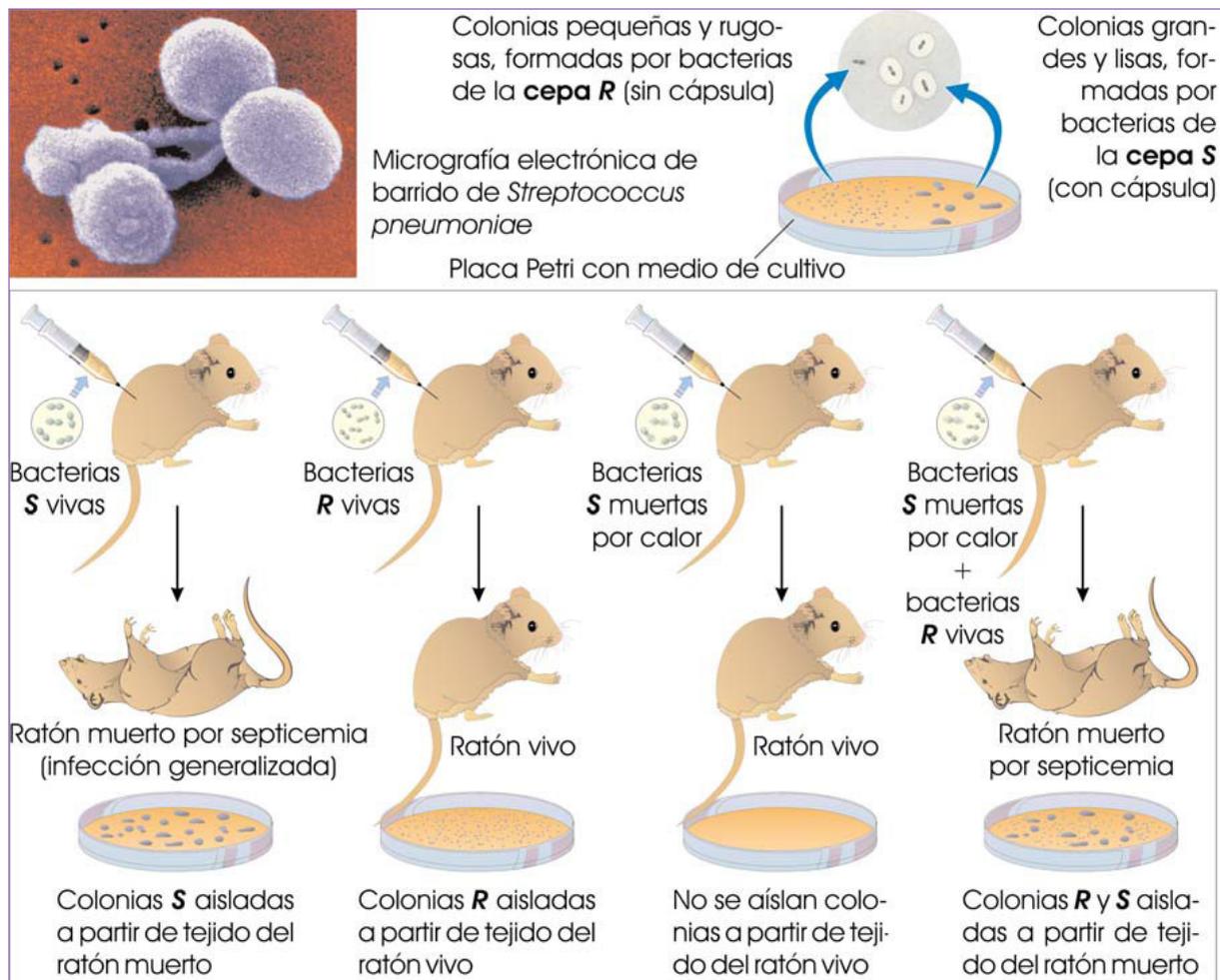
Ahora bien, si el ADN se encuentra en los cromosomas, ¿podría ser el material que forma los genes? La opinión general era que no, porque se asumía que el ADN era una molécula demasiado simple. Levene —el bioquímico ruso que, como vimos en la Unidad 4, estableció las diferencias químicas entre el ARN y el ADN— concluyó, a partir de determinaciones algo imprecisas, que los cuatro nucleósidos del ADN (dG, dC, dA y dT) se hallaban en igual proporción, y propuso que esta molécula consistiría en la repetición de un tetranucleótido lineal o cíclico [véase la *ilustración 6.2*]; ¿y qué información podría transmitir un mensaje en el que figurasen solo las letras GCAT GCAT GCAT... una y otra vez?



**Ilustración 6.2.** Esquema del hipotético tetranucleótido que Levene propuso (incorrectamente) como elemento estructural del ADN (Fuente: ASH).

Los genes, pues, no podían ser otra cosa que proteínas. El gen de la hemoglobina, por ejemplo, no sería más que una molécula de hemoglobina que permanecería “en reserva” dentro del núcleo; cada vez que se necesitara hemoglobina solo tendría que hacerse una “copia de trabajo” del prototipo. El ADN serviría simplemente para mantener unidos los genes —o sea, las proteínas— de cada cromosoma (gracias a las cargas negativas de los grupos fosforilo, que atraen a las proteínas).

Sin embargo, una serie de datos apuntaban insistentemente en otra dirección. Las proteínas que acompañaban al ADN en los cromosomas (las **protaminas** detectadas en los espermatozoides, mencionadas al comienzo de la Unidad 4, y las denominadas **histonas** en las demás células) resultaron poseer una estructura particularmente sencilla comparada con otras proteínas de la célula. Por el contrario, se advirtió que el peso molecular del ADN era muy superior al de las proteínas y, además, había el doble de ADN que de proteína en los cromosomas: demasiado para tratarse de un simple pegamento.



**Ilustración 6.3.** El experimento de Griffith mostró que al destruir por calor las bacterias que poseían la habilidad de sintetizar una cápsula, liberaban un “principio transformante” que podía inducir dicha capacidad en otras bacterias (Fuentes: <http://phil.cdc.gov/phil/home.asp> y ASH).

El rechazo definitivo al modelo proteínico del gen se debió a ciertos investigadores que trabajaban con neumococos (*Streptococcus pneumoniae*), las bacterias responsables de la neumonía humana. El oficial médico británico Frederick Griffith (1879-1941) había descrito en 1923 dos cepas de esta bacteria: una que desarrollaba una cápsula de polisacáridos que la protegía frente al sistema inmunitario de su huésped, por lo que era muy **virulenta**, y otra, mutante, incapaz de formar la cápsula y, por tanto, benigna. La primera se reproducía al cultivarla en un medio sólido formando **colonias** de aspecto liso, por lo que la llamó cepa **S** (del inglés *smooth*, “liso”); la segunda formaba colonias rugosas (*rough*), y la etiquetó con la letra **R**.

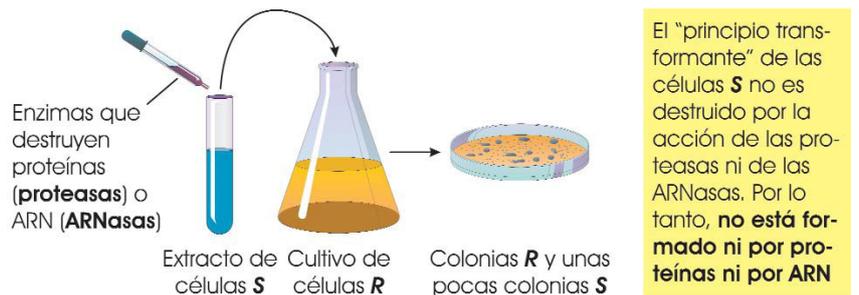
En 1928, Griffith informó de un resultado sorprendente. Había inyectado a un ratón una mezcla de bacterias S muertas por calor (por tanto, inocuas) y bacterias R vivas, y aisló de los tejidos del ratón bacterias S vivas [véase la ilustración 6.3]. Por supuesto, las bacterias S muertas no habían resucitado, pero las bacterias R habían adquirido la capacidad de producir la cápsula y se habían transformado en bacterias S. Las bacterias muertas, pues, deberían haber liberado el gen correspondiente (al que Griffith

llamó **principio transformante**), que se habría insertado en las bacterias R.

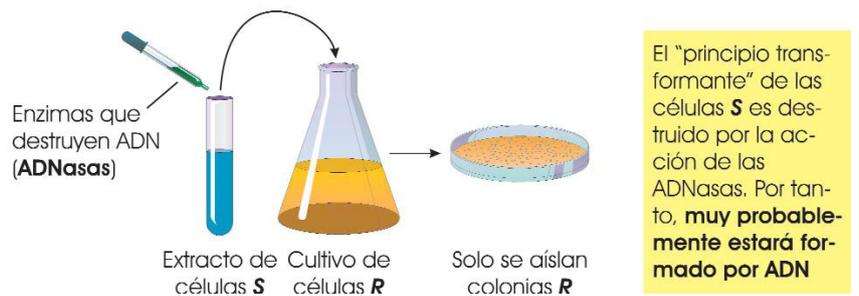
De ahí el enorme interés en averiguar la naturaleza química del “principio transformante”, tarea en la que se implicó el médico canadiense Oswald Theodore Avery (1877-1955). Durante una década trabajó con afán junto a dos colaboradores, el canadiense Colin Munro MacLeod (1909-1972) y el estadounidense Maclyn McCarty (1911-2005). En 1944, tras un proceso que fue esencialmente de eliminación [véase *la ilustración 6.4*], Avery, MacLeod y McCarty establecieron más allá de toda duda razonable que el **principio transformante —y, por tanto, el gen— era ADN**.



El “principio transformante” de las células **S** no es destruido por la acción del calor



El “principio transformante” de las células **S** no es destruido por la acción de las proteasas ni de las ARNasas. Por lo tanto, **no está formado ni por proteínas ni por ARN**



El “principio transformante” de las células **S** es destruido por la acción de las ADNasas. Por tanto, **muy probablemente estará formado por ADN**

**Ilustración 6.4.** Demostración de que el “principio transformante” es ADN (Fuente: *ASH*).

### 1.1. Estructura y organización del ADN

En 1950 casi todo el mundo estaba convencido de que el ADN era el portador de la información genética. En realidad, diversas investigaciones llevadas a cabo con virus mostraron que las proteínas asociadas al ADN eran casi superfluas, pues el ADN podía producir efectos genéticos por sí solo... justo lo contrario del modelo defendido por Levene. Era evidente que el ADN no podía consistir en unidades sencillas que se repetían, como se había supuesto. Urgía, pues, averiguar su auténtica estructura.

### La proporción de bases en el ADN

Una primera pista en esta línea la proporcionaron los estudios del bioquímico austriaco Erwin Chargaff (1905-2002) y sus colaboradores. En 1948 utilizaron una técnica similar a la cromatografía en papel [véase la ilustración 1.14] para medir la proporción exacta de los constituyentes del ADN. Pronto comprobaron que la suposición de Levene era errónea: los cuatro nucleósidos (a los que en lo sucesivo, y siempre que no de lugar a confusión, designaremos como G, C, A y T, sin la *d* delante) *no* se hallaban en cantidades iguales. De hecho, su proporción variaba de una especie a otra, pero en cada especie permanecía constante de un tejido a otro y no se modificaba con la edad o la nutrición.

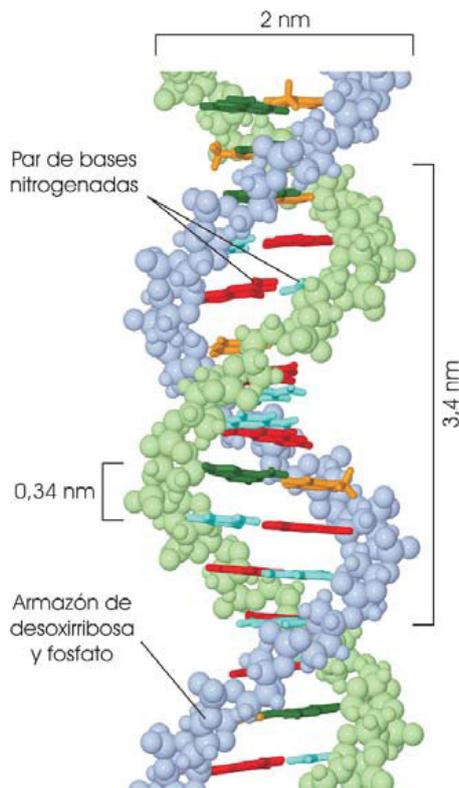
La tabla siguiente muestra el porcentaje de nucleósidos (y de las correspondientes bases nitrogenadas) en seis especies. Chargaff pudo percatarse de que, pese a la variación, se observaban ciertas regularidades. Así, la cantidad total de bases púricas (*Pur*) de cualquier muestra de ADN era la misma que la de bases pirimidínicas (*Pyr*); más aún, había tanta G como C, y tanta A como T. Estas relaciones se conocieron como **reglas de Chargaff**.

Organismo	Nucleósidos (% molar)				Relaciones		
	G	C	A	T	G/C	A/T	Pur/Pyr
<i>Homo sapiens</i>	19,9	19,8	30,9	29,4	1,01	1,05	1,03
Gallina	21,7	20,5	28,8	29,2	1,06	0,99	1,02
Erizo de mar	17,3	17,7	32,8	32,1	0,84	1,02	1,01
Levadura	18,7	17,1	31,3	32,9	1,09	0,95	1,00
Trigo	22,7	22,8	27,3	27,1	1,00	1,01	1,00
<i>Escherichia coli</i>	26,0	25,7	24,7	23,6	1,01	1,05	1,03

### El modelo de la doble hélice

Además de regularidades *numéricas*, también se apreciaron regularidades *estructurales* en la molécula de ADN: mediante **difracción de rayos X**, la fisicoquímica británica Rosalind Franklin (1920-1958) y el físico neozelandés Maurice Wilkins (1916-2004) detectaron una pauta que se repetía cada 3,4 nm, esto es, a intervalos bastante mayores que la distancia de un nucleótido a otro (0,34 nm). La conclusión lógica era que la molécula de ADN adoptaba la forma de una hélice; sus vueltas en espiral formarían las unidades repetidas que se apreciaban con los rayos X.

En 1953, el físico inglés Francis Harry Compton Crick (1916-2004) y el bioquímico estadounidense James Dewey Watson (n. 1928) reunieron estos y otros datos y con ellos elaboraron un modelo del ADN que marcaría un hito en la historia de la ciencia. Lo novedoso estribaba en representar al ADN no simplemente como una hélice, sino como una **doble hélice** [véase la ilustra-



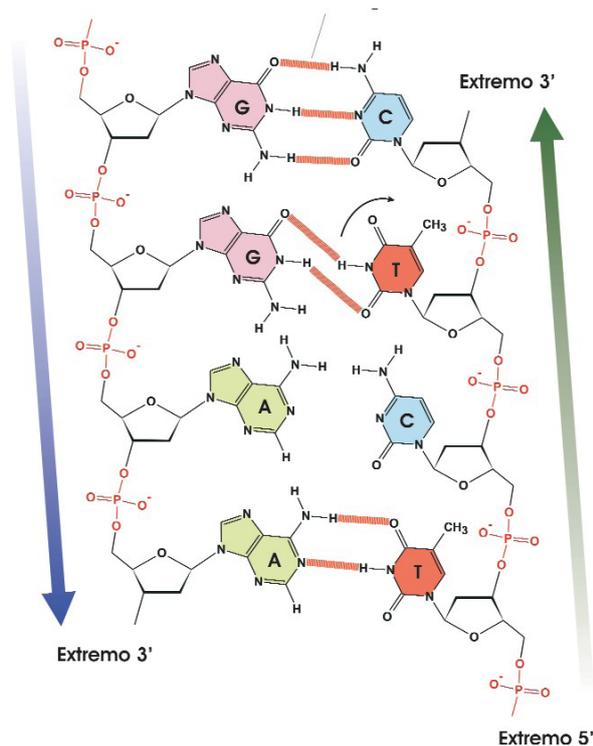
**Ilustración 6.5.** Modelo de una doble hélice de ADN postulado por Watson y Crick. La estructura se repite a intervalos de 3,4 nm, que corresponden a diez pares de bases (Fuente: ASH).

ción 6.5] en la que dos cadenas de polinucleótido se entrelazan una alrededor de la otra. Ambas hebras, cuyo exterior está dominado por un esqueleto donde alternan grupos fosforilo con restos de desoxirribosa, tienen orientación **antiparalela**: los enlaces 5'→ 3' van en sentidos opuestos. Desde el armazón azúcar-fosfato se proyectan hacia el interior bases nitrogenadas, que se unen a bases de la hebra opuesta mediante enlaces de hidrógeno. Los **pares de bases complementarias** así formados se apilan como los peldaños de una escalera de caracol; la presencia de miles de estos enlaces en una molécula de ADN contribuye en buena medida a su estabilidad.

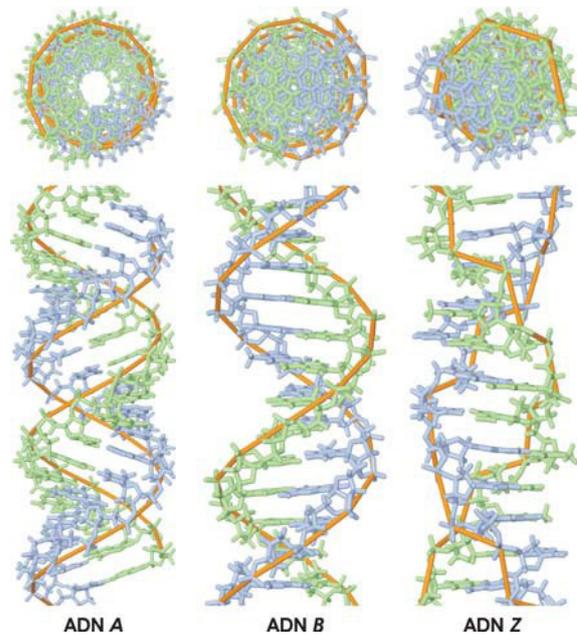
¿Qué pares de bases podrían ser complementarias? Dada la distancia entre ambas hebras, dos bases pirimidínicas (cada una con un anillo) se hallarían demasiado alejadas como para formar enlaces de hidrógeno. Por el contrario, dos bases púricas (con dos anillos cada una) se superpondrían. La única posibilidad es que se enfrenten una base púrica y otra pirimidínica, constituyendo así un “peldaño” de tres anillos. Como se explica en la ilustración 6.6, lo más habitual es que *la adenina de una hebra forme dos enlaces de hidrógeno con la timina de la otra, y que la guanina forme tres con la citosina*. Como consecuencia, cuanto más alta sea la proporción de pares G·C en el ADN mayor será la temperatura a la que habrá que calentarlo para romper los enlaces de hidrógeno y separar las hebras, proceso conocido como **desnaturalización** o “fusión” del ADN.

Una característica añadida de la estructura de la molécula de ADN hasta ahora descrita, conocida ya en tiempos de Watson y Crick como **forma B**, es que si se mira por encima de su eje se verá que las dos hebras en espiral se alejan del observador girando en sentido horario, como se aprecia en la ilustración 6.7; por convenio, se dice que la doble hélice *gira a la derecha*, o que el enrollamiento es dextrógiro. También lo es el de la llamada **forma A**, en la que los pares de bases están inclinados, mientras que el de la **forma Z** es *levógiro* (esto es, gira a la izquierda); ambas formas son bastante menos comunes que la B. Y más extraña aún es la existencia, en algunos virus, de moléculas de ADN formadas por una sola hebra.

Hasta el momento hemos distinguido dos niveles estructurales en el ADN: la **estructura primaria**, que concierne a la secuencia de nucleótidos, y la **estructura secundaria**, que corresponde a la ordenación regular y estable que adoptan los nucleótidos y que está representada, por ejemplo, en el modelo de doble hélice [véanse las ilustraciones 6.5 y 6.7]. Como seguidamente veremos, el ADN está organizado de forma muy compleja en el núcleo de las células eucariotas (**estructura terciaria**).



**Ilustración 6.6.** Entre A y C no se forman enlaces de hidrógeno, y entre G y T solo pueden formarse si la hélice sufre una distorsión local (flecha). Así, los únicos pares de bases que aparecen prácticamente siempre se corresponden con G•C y A•T (Fuente: ASH).



**Ilustración 6.7.** Algunas formas que puede presentar la doble hélice de ADN, vistas a lo largo del eje helicoidal (arriba) y lateralmente (abajo). En naranja se representa la dirección del “esqueleto” de cada molécula (Fuente: ASH).

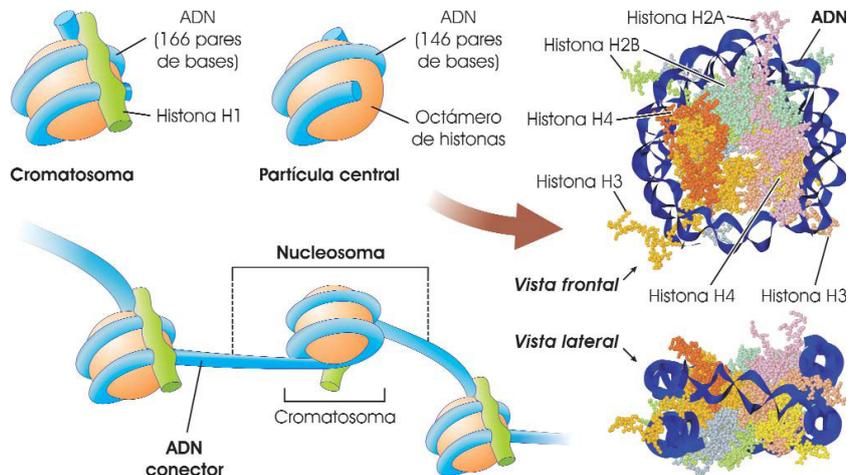
## 1.2. Organización del ADN en el núcleo de las células eucariotas

Que el ADN sea el principal componente de los cromosomas no significa que sea el único. Ya hemos mencionado anteriormente la presencia de proteínas llamadas **histonas**. Se trata, de hecho, de las principales proteínas de los cromosomas de las células eucariotas. Hay cinco tipos principales de histonas —designadas H1, H2A, H2B, H3 y H4—, todas ellas ricas en aminoácidos básicos, esto es, con carga positiva, lo que les permite interactuar con las cargas negativas de los grupos fosforilo del ADN.

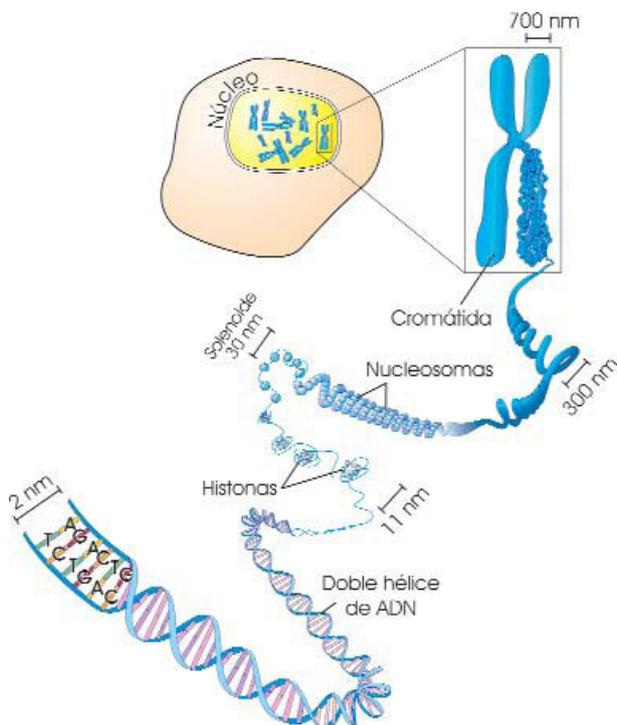
De tal interacción resultan los **nucleosomas**, las unidades básicas de la estructura de los cromosomas. El nucleosoma consta de [véase la ilustración 6.8]:

- Una **partícula central**, formada por dos copias de cada histona H2A, H2B, H3 y H4 (el **octámero**) y por un segmento de ADN de 146 pares de bases que describe 1,65 vueltas a su alrededor.
- El **ADN conector**, formado por dos “colas” que lo enlazan con el nucleosoma anterior y con el posterior. Incluye entre 10 y 80 pares de bases, dependiendo de la especie y el tipo de tejido.

Los primeros diez pares de nucleótidos del ADN conector contiguos a cada extremo de la partícula central se fijan a la histona H1 [véase la ilustración 6.8]. Se llama **cromatosoma** al conjunto formado por la partícula central, la histona H1 y el ADN al que se une; en total, supone  $146 + 10 + 10 = 166$  pares de bases (dos vueltas de ADN).



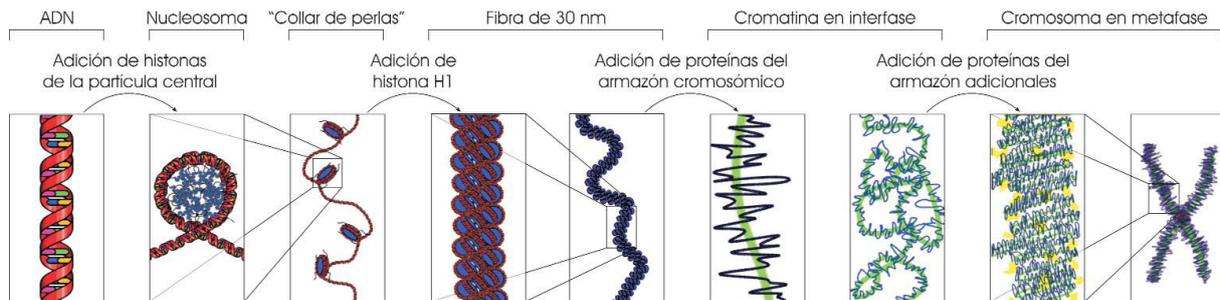
**Ilustración 6.8.** Izquierda: esquema de un nucleosoma y sus componentes. Derecha: modelo tridimensional de la partícula central. El ADN se representa en forma de cintas, y las histonas como bolas (Fuente: ASH).



**Ilustración 6.9.** Sucesivos niveles de compactación del ADN, que consiguen comprimir, por ejemplo, 8,3 cm de ADN en los 180 nm de longitud del cromosoma 1 humano (Fuente: <http://en.wikipedia.org/wiki>).

Los 146 pares de bases de la partícula central medirían, si se “estirara” la doble hélice, unos 50 nm; al arrollarlos en torno al octámero de histonas se consigue empaquetarlos en un espacio de unos 10 nm, lo que equivale a reducir su longitud a una quinta parte. Pero este no es más que el nivel inferior en la compactación del ADN, esquematizado en las ilustraciones 6.9 y 6.10. Así, las histonas H1 se unen, probablemente, para formar una especie de eje en torno al cual se enrolla una ristra de nucleosomas (unos seis por giro), formando el llamado **solenoid** o fibra de 30 nm de grosor. Este es el máximo nivel de plegamiento que alcanza la cromatina durante la **interfase**.

Sin embargo, cuando comienza la mitosis la fibra de 30 nm se pliega y empaqueta, formando cromátidas de unos 700 nm de grosor. No se conoce bien la forma en que esto ocurre, aunque en opinión de muchos se hallan implicadas un conjunto de proteínas no histónicas que componen una especie de **armazón cromosómico**, al que se ancla la fibra de 30 nm en algunos puntos dejando libres grandes bucles; dicho armazón se arrollaría en una hélice, cuyo posterior plegamiento formaría la cromátida.



**Ilustración 6.10.** Esquema detallado de la compactación de la cromatina y su asociación con las proteínas (wikipedia.org.Dominio público).

### 1.3. Organización del ADN en procariotas

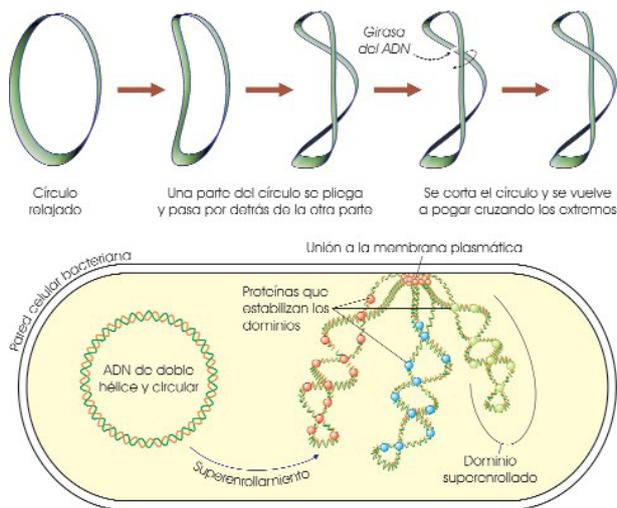
Además de no hallarse separado del citoplasma por una envoltura nuclear, el ADN de las células procarióticas se organiza de forma diferente a la propia de los eucariotas:

- En contraste con los múltiples cromosomas lineales hallados en las células eucariotas se observó, en las primeras bacterias estudiadas, que poseían una única molécula en doble hélice de ADN que, además, presentaba una notable peculiaridad: era circular, es decir, los dos extremos de cada una de sus hebras estaban covalentemente unidos.

Además de este **cromosoma bacteriano**, muchas bacterias contenían también **plásmidos**, pequeñas moléculas de ADN independientes del cromosoma bacteriano que estudiaremos con detalle en la Unidad 7, y que resultaron ser también circulares. No obstante, pese a que dichas observaciones iniciales siguen siendo válidas en términos generales, se han hallado bacterias con varios cromosomas circulares, y muchas tienen cromosomas y plásmidos *lineales*.

- El problema de compactar un ADN circular para que quepa en el interior de la célula procariota se soluciona de varios modos. En muchas arqueas tiene lugar gracias a histonas parecidas a las de los eucariotas, que organizan estructuras similares a nucleosomas. Pero en la mayoría de las bacterias se debe al **superenrollamiento**, esto es, al retorcimiento de una molécula de ADN constituida por dos cadenas que de por sí estaban arrolladas una alrededor de la otra: una enzima, la **girasa del ADN**, utiliza la energía de hidrólisis del ATP para formar superenrollamientos [véase la ilustración 6.11]. El cromosoma bacteriano, habitualmente unido a la membrana plasmática, se organiza con frecuencia no en un único superenrollamiento, sino en varios **dominios superenrollados** estabilizados, probablemente, por proteínas.

El ADN circular de las bacterias recuerda al que podemos hallar en dos orgánulos de las células eucarióticas: las mitocondrias y los cloroplastos. En su matriz y su estroma, respectivamente, se localizan varias copias de ADN circular que contienen genes



relacionados con algunas de las proteínas de dichos orgánulos. Ello da pie a hablar de una **herencia citoplásmica o extracromosómica**, que no sigue las leyes de Mendel y que, en los vertebrados, se transmite casi exclusivamente por vía materna.

**Ilustración 6.11.** Arriba: Introducción de superenrollamiento en un ADN circular por acción de la girasa del ADN, que “corta y pega”. Abajo: Modelo hipotético de organización del ADN circular de una bacteria, formando dominios estabilizados por la unión de ciertas proteínas (Fuente: ASH).

### Actividades

1. El modelo de la doble hélice de Watson y Crick, ¿justifica las leyes de Chargaff?
2. En el ADN extraído del saltamontes, el 29,3 % de las bases nitrogenadas corresponde a adenina. ¿Qué porcentajes esperarías hallar de las restantes bases?
3. En el ADN de un virus conocido como  $\Phi X174$  la proporción entre bases púricas y pirimidínicas es de 0,897. ¿Qué indica el dato acerca de la estructura de su ADN?
4. Según los datos de la tabla que aparece en el epígrafe 1.1, ¿cuál de las moléculas de ADN costará menos desnaturalizar por acción del calor? Razona la respuesta.
5. Un ADN lineal se desnaturaliza a 85 °C. Si ese mismo ADN fuera circular, ¿sería mayor o menor la temperatura a la que se separarían las dos hebras? Razónalo.
6. Una mutación en el gen ND1 del ADN mitocondrial humano ocasiona una afección visual conocida como **neuropatía óptica hereditaria de Leber (NOHL)**. Si Luís padece NOHL, pero Carmen no, ¿cómo serán sus hijos? ¿Y si fuera Carmen la afectada? Razona la respuesta.



### Recuerda

- El **ácido desoxirribonucleico (ADN)** constituye el material genético de los seres vivos y está formado por largos polímeros no ramificados de **nucleótidos** (estructura primaria).
- Las bases nitrogenadas de dos hebras **complementarias** interactúan mediante enlaces de hidrógeno (A-T, G-C), formándose una **doble hélice** (estructura secundaria).
- El ADN debe estar compactado para caber dentro de las células. Para ello se asocia con proteínas (histonas en los eucariotas) formando superenrollamientos o secuencias de **nucleosomas**, que a su vez se pliegan formando los **cromosomas**.

## 2. La transcripción

La determinación de la estructura del ADN no hizo más que enfatizar el problema principal: ¿cómo realiza su trabajo, esto es, de qué manera controla la síntesis de proteínas? Para lograrlo, la molécula de ADN tendría que “saber” cuál de los 21 aminoácidos debería situar en cada una de las diferentes posiciones de una proteína (formada quizá por miles de unidades). Cosa relativamente fácil si solo hubiese cuatro aminoácidos (llamémosles a, b, c y d): una especie de “lector” de ADN recorrería una hebra de la molécula y cada vez que encontrase, pongamos por caso, una G, incorporaría el aminoácido a en la correspondiente posición de la proteína; si a continuación leyese una C, agregaría el b; una nueva G daría lugar a otro aminoácido a, y así sucesivamente.

Pero resulta que hay 21 aminoácidos (22 en ciertas arqueas), y solo 4 bases nitrogenadas. ¿Cómo traducir, pues, un lenguaje construido con un alfabeto de 4 letras (el lenguaje del ADN) a un lenguaje con 21 letras (el lenguaje de las proteínas)?

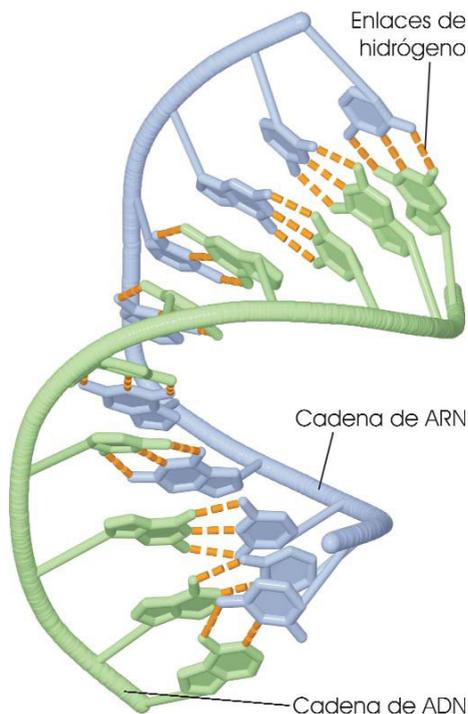
### 2.1. El “dogma” central de la Biología molecular

En 1954, el astrofísico ucraniano George Gamow (1904-1968) sugirió una solución inspirada en el código Morse. Al combinar solo dos signos, el punto (•) y el guión (-), los telegrafistas podían representar todos los símbolos del alfabeto; así, - • • era la “D”, y • • • significaba “S”. ¿No podrían asimismo las células combinar bases nitrogenadas (y, por tanto, nucleósidos) para “representar” aminoácidos en una especie de **código genético**?

Gamow aventuró que el ADN serviría de “molde” directo para el ensamblaje de los aminoácidos proteínicos: a lo largo del surco que hay entre las dos hebras de la doble hélice se moldearían cavidades delimitadas por distintas combinaciones de cuatro bases; en cada cavidad se insertaría un aminoácido específico, y cuando todos los aminoácidos se hubiesen alineado en el orden correcto, una enzima los uniría formando una cadena polipeptídica, que se desprendería del ADN. Para saber qué aminoácido “encaja” en una cavidad concreta bastaría conocer tres bases—un *triplete*—, ya que la cuarta quedaría automáticamente determinada por las reglas del apareamiento de bases. Según Gamow, el orden de las bases sería irrelevante (así, CAG y GAC definirían el mismo aminoácido), lo cual dejaría justo 20 combinaciones para los 20 aminoácidos conocidos por entonces.

### La implicación del ARN

Desafortunadamente, el modelo de Gamow no funcionó. Los huecos que quedaban entre las bases de la doble hélice no se diferenciaban lo suficiente como para seleccionar los aminoácidos. Además, según Gamow, los aminoácidos se ensamblaban directamente sobre el ADN, pero ya desde antes de 1940 se sabía que las proteínas se sintetizan en el citoplasma, físicamente separado del ADN por la envoltura nuclear en los eucariotas. ¿Cómo podía entonces transmitirse la información genética desde donde se almacena (el núcleo) hasta donde se utiliza (el citoplasma)?



**Ilustración 6.12.** Molécula híbrida de ADN-ARN (Fuente: ASH).

Precisamente, en el citoplasma se había hallado una molécula parecida al ADN: el ARN. Las investigaciones mostraron que las células de tejidos muy activos o en rápido crecimiento, como el hígado —donde se fabrica la mayor cantidad de proteínas—, eran justo las que tenían más ARN. Consideraciones de esta índole llevaron a sugerir en 1952 que el ADN no sería un molde *directo* para la síntesis de proteínas; pero sí indirecto, en el sentido de que actuaría de patrón para la síntesis de una cadena de ARN, la cual serviría a su vez de matriz para el ensamblaje de proteínas tras ser exportada al citoplasma.

Un dato a favor de tal hipótesis es que la desnaturalización del ADN no es un proceso irreversible; al contrario, si las hebras complementarias se mantienen a una temperatura suave durante algún tiempo se vuelven a formar enlaces de hidrógeno entre ellas. Este proceso inverso, denominado **renaturalización**, permitía también la formación de una doble hélice **híbrida**, con una hebra de ADN y otra de ARN en la que la timina había sido sustituida por el uracilo [véase la ilustración 6.12]. Era concebible, pues, que una molécula de ARN se sintetizara a base de fijarse ribonucleótidos a un molde de ADN.

### Las clases de ARN

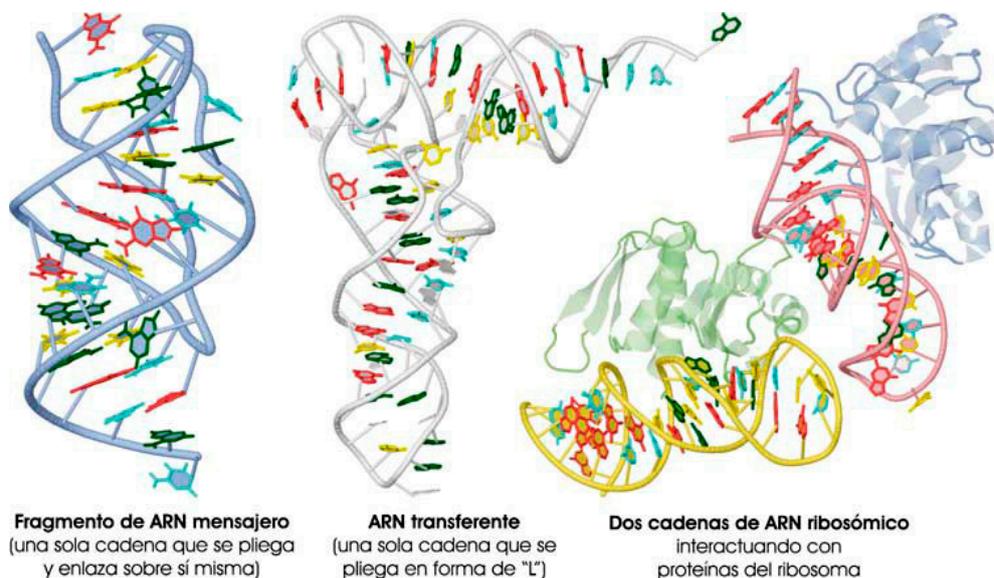
En 1953, el citólogo rumano George Emil Palade (1912-2008) había revelado la existencia de **ribosomas**, y había comprobado que el ARN abundaba especialmente en ellos. En 1955 se confirmó que era en los ribosomas donde ocurría la fabricación de proteínas, por lo que resultaba lógico suponer que el ARN de cada ribosoma —denominado **ARN ribosómico** y abreviado como **ARNr**— era una “copia” de un gen determinado, es decir, el molde para la síntesis de un polipéptido específico.

Esta hipótesis, que podía resumirse en la máxima “un gen, un ribosoma, una proteína”, permitía formular una predicción: supongamos que insertamos un nuevo gen en el ADN de una bacteria; entonces, la *expresión* del gen —la producción del polipéptido en él cifrado— no debería comenzar hasta que se hubiera “copiado” en forma de ARNr y éste se hubiera ensamblado

en ribosomas —proceso que comporta cierto tiempo—, y debería continuar incluso tras eliminar el gen, puesto que los ribosomas son estructuras estables y los moldes en forma de ARNr aún permanecerían.

Pero los experimentos mostraron que, una vez transferido a una bacteria, el gen empezaba a expresarse sin retardo aparente y, tras su destrucción, la síntesis de proteínas cesaba de inmediato. Era obvio que el proceso no se hallaba supeditado a la formación de moldes estables de ARNr. Además, cada ribosoma parecía servir para la síntesis de muchas proteínas diferentes, no de una concreta.

La única explicación posible fue sugerida en 1960 por los biólogos franceses François Jacob (n. 1920) y Jacques Lucien Monod (1910-1976), quienes postularon la existencia de un segundo tipo de ARN: el **ARN mensajero (ARNm)**, de corta vida, que llevaría la información copiada del ADN de los genes hasta los ribosomas; tras la fabricación de varias moléculas de proteína, el ARNm se disociaría de los ribosomas y se eliminaría, lo que dejaría a estos listos para recibir una nueva molécula de ARNm y producir otras proteínas. Antes de finales de 1961, varios equipos independientes demostraron la existencia del ARNm, confirmando así la hipótesis de Jacob y Monod.

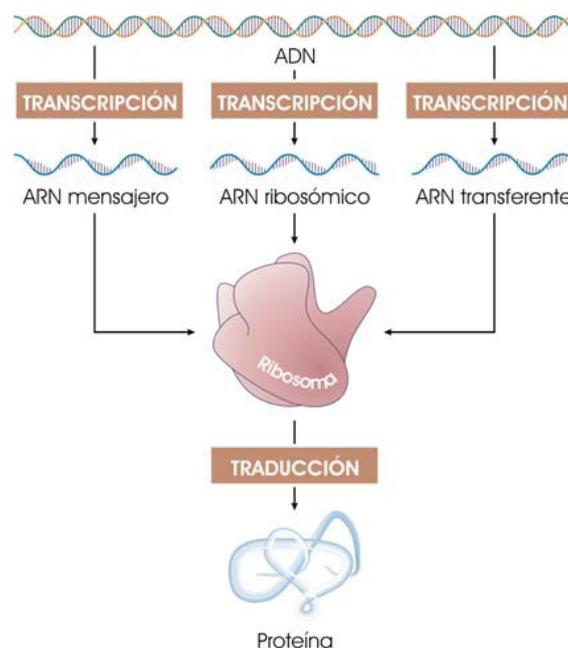


**Ilustración 6.13.** Clases de ARN más abundantes en las células eucarióticas. Se trata, en todos los casos, de moléculas monocatenarias (con una sola hebra), aunque en algunas zonas, denominadas horquillas, se forman enlaces de hidrógeno entre bases complementarias (Fuente: ASH).

Aun quedaba un escollo por sortear: nadie había podido establecer una relación física directa entre un ARNm y los aminoácidos que se han de concatenar utilizándolo como molde. Pero en 1954 Crick había anticipado ya una solución. Propuso que deberían existir unas moléculas, a las que llamó **adaptadores**, que se unirían por un lado a un aminoácido específico y por otro

lado a regiones concretas del ARNm, acarreando así el aminoácido hasta el lugar en el que se añadiría al polipéptido naciente. Experimentos posteriores confirmaron la predicción de Crick al detectar aminoácidos marcados radiactivamente asociados a pequeñas moléculas de ARN, hoy conocidas como **ARN transferente (ARNt)**.

Crick resumió estos hallazgos en el bautizado como **dogma central de la Biología molecular**, que comparaba la síntesis de proteínas con un proceso lingüístico. Un hispanohablante puede transcribir un texto de Tolstoi desde el alfabeto cirílico al latino (así, Война и мир se convertiría en *Voiná i mir*) y traducirlo del ruso al castellano (*Voiná i mir* se transformaría en “Guerra y paz”). Análogamente, una cadena de desoxirribonucleótidos de ADN puede experimentar su **transcripción** a una cadena de ribonucleótidos de ARNr, ARNt o ARNm, y un ribosoma puede ejecutar un “programa” capaz de “leer” el ARNm y culminar su **traducción** a la secuencia de aminoácidos de una proteína [véase la ilustración 6.14]. Pero mientras el texto en castellano se puede volver a traducir al ruso, para las proteínas no hay marcha atrás: una vez transferida la información a la proteína no puede ya salir de ella; esto es, no es posible usar la información almacenada en las proteínas para sintetizar ARN o ADN.



**Ilustración 6.14.** El dogma central de la Biología molecular tal y como fue formulado en la versión inicial de Crick, en 1958 (Fuente: ASH).

Más tarde Crick lamentaría haber empleado la palabra “dogma”, que tiene connotaciones religiosas (en el sentido de creencia que no se pone en duda) y se halla, pues, alejada de la buena práctica científica. Además, pronto se haría patente que su alegato “el ADN hace ARN, el ARN hace proteínas, y las proteínas nos hacen a nosotros” adolecía de excesiva simplificación, lo que le obligaría, como veremos, a introducir modificaciones. Es decir, contra el *dogma* empezarán a alzarse algunas *herejías*.

## 2.2. Redacción de mensajes genéticos

El “dogma” central obligó, una vez más, a redefinir el concepto de gen. La genética clásica utilizaba dicho término para aludir a una entidad un tanto abstracta, una “partícula” situada en algún punto del cromosoma y responsable de un rasgo del organismo. Pero para los biólogos moleculares el gen pasó a tener un significado físico concreto, el de un fragmento de una molécula<sup>1</sup>:

**Un gen consistiría, pues, en la secuencia completa de nucleótidos en la molécula de ADN que es necesaria para sintetizar una molécula de ARN.**

Dicho ARN —el denominado **transcrito primario**— puede experimentar modificaciones diversas (por ejemplo, su escisión en moléculas de menor tamaño). El resultado de las mismas constituye a veces el producto final del gen, como es el caso del ARNr o del ARNt; pero frecuentemente se trata solo un intermediario (el ARNm) que ha de ser **traducido** a un polipéptido funcional. Y, como adelantamos en la Unidad 4, en algunos virus los ARN no son los productos de los genes, sino *los propios genes*.

### Polimerasas de ARN

Ciñéndonos por el momento al núcleo de las células eucarióticas, la estructura de la mayoría de los genes sigue una pauta común. Todos incluyen dos **secuencias reguladoras** necesarias para la correcta síntesis del transcrito primario:

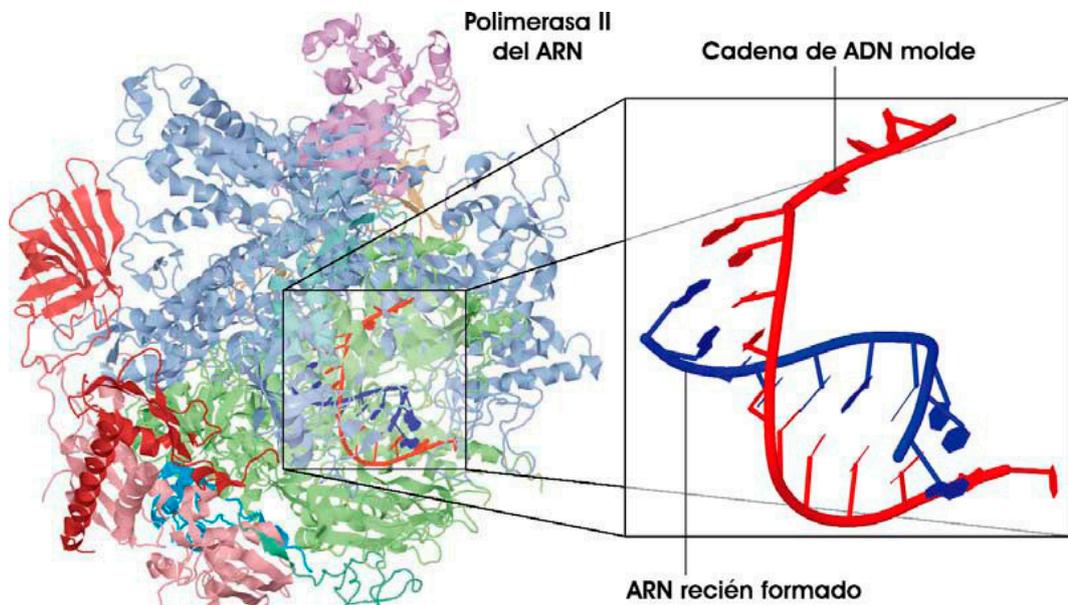
- El **promotor**: es la región del ADN a la que se une la **polimerasa de ARN dirigida por ADN**, la principal enzima responsable de la transcripción.
- El **terminador**: es una secuencia específica de nucleótidos que fuerza a la polimerasa a separarse del ADN y liberar el ARN recién formado.

Las polimerasas de ARN, o ARN polimerasas, son enzimas capaces de unirse a una doble hélice de ADN y utilizarla como molde o patrón para polimerizar ribonucleótidos y fabricar una molécula de ARN.

Por lo general solo se transcribe una de las dos hebras del ADN: la polimerasa avanza en el sentido 3' → 5' de dicha cadena de ADN, en tanto que la hebra de ARN crece en el sentido 5' → 3'. Es decir, el carbono 3' del último ribonucleósido forma un enlace fosfodiéster con el carbono 5' de un trifosfato de ribonucleósido (GTP, CTP, ATP o UTP), del que solo se retiene un grupo

<sup>1</sup> Parece difícil dar una definición de gen totalmente satisfactoria. Así, se conocen regiones de ADN que no se transcriben, pero que se consideran genes; por ejemplo, los llamados genes de recombinación (las regiones a las que se unen las enzimas responsables de la recombinación durante la meiosis), los genes de segregación (los sitios específicos para que las fibras del huso mitótico se adhieran a los cromosomas) o diversos fragmentos de ADN que interactúan con proteínas u hormonas.

fosforilo, el más próximo a la ribosa [véase la ilustración 6.16]; los otros dos se liberan como pirofosfato (PPi), cuya hidrólisis para dar dos fosfatos inorgánicos (Pi) suministra la energía que “impulsa” la polimerización. En cambio, el nucleósido del extremo 5’ —el aceptor inicial al que luego se agregarían los restantes nucleósidos— retiene los tres grupos fosforilo: casi siempre, en el extremo 5’ hay ATP o GTP.

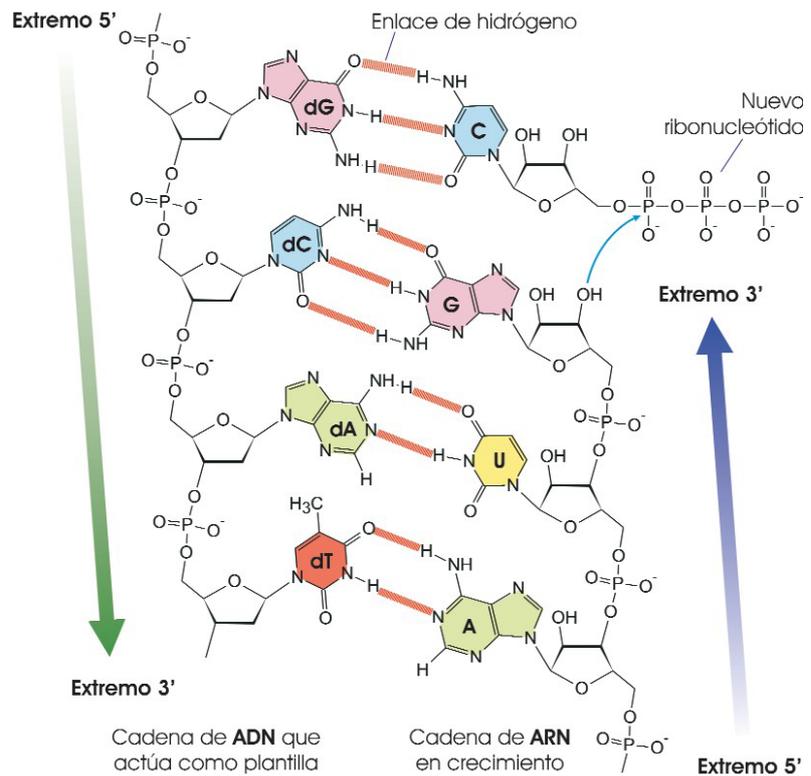


**Ilustración 6.15.** Modelo de la polimerasa de ARN II de *Saccharomyces cerevisiae* transcribiendo una hebra de ADN (Fuente: ASH).

¿Cómo “decide” la polimerasa cuál de los cuatro nucleótidos ha de añadir, en cada paso, al extremo 3’ del ARN en desarrollo? Las instrucciones a este respecto llegan del ADN por la vía de la complementariedad de bases. La enzima solo incorporará un ribonucleótido si su base nitrogenada se empareja con la del desoxirribonucleótido que tiene enfrente. De esta forma, la cadena de ARN que se ensambla será complementaria a una de las hebras de ADN (e idéntica en secuencia de bases nitrogenadas, por lo tanto, a la otra hebra, salvo en la sustitución de timina por uracilo).

Existen tres tipos de polimerasas en el núcleo de las células eucariotas:

- La **polimerasa I** se localiza en el nucléolo y es responsable de la síntesis del precursor de la mayoría de los ARNr.
- La **polimerasa II** interviene en la síntesis de los ARNm, de varios de los denominados *ARN nucleares pequeños* (ARNnp) y *ARN nucleolares pequeños* (ARNnop), así como de centenares de moléculas conocidas como *microARN* (miARN).
- La **polimerasa III** transcribe los genes que cifran todos los ARNt, un ARNr, uno de los ARNnp y una heterogénea colección de otros ARN de pequeño tamaño.



**Ilustración 6.16.** Alargamiento de una cadena de ARN por agregación de ribonucleótidos complementarios a una cadena de ADN (Fuente: ASH).

### Transcripción de la eucromatina

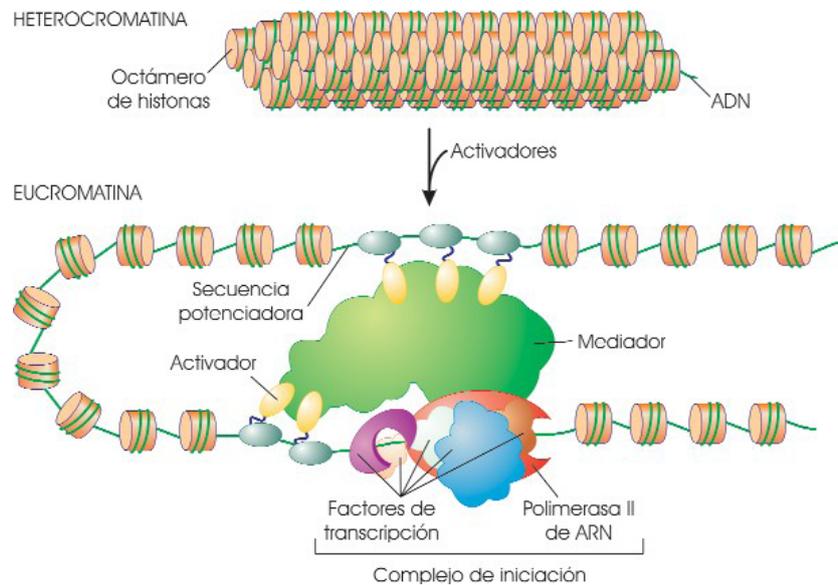
De todas las polimerasas de ARN de los eucariotas, la que se conoce mejor es, sin duda, la **polimerasa II**, un complejo capaz de sintetizar, de principio a fin, todos los ARNm necesarios para las diversas proteínas celulares. Pero, para poder ejercer su trabajo, la enzima debe superar un obstáculo: a menudo los genes que debe transcribir están “desactivados”.

En efecto, cada célula tiene que controlar estrictamente qué genes se deben expresar y cuáles no en un momento dado; las células embrionarias pueden, de este modo, diferenciarse en el transcurso del desarrollo, y los organismos se hallan capacitados para responder a los cambios en el medio.

Una forma muy extendida de mantener a los genes inactivos es empaquetar densamente el ADN en forma de **heterocromatina** (que, recordémoslo, es el estado de la cromatina que se tiñe intensamente con eosina), de modo que la polimerasa no tenga acceso a los mismos. Por esta razón, para que se transcriba un gen es necesario previamente que la cromatina se descondense y se transforme en **eucromatina** (que se tiñe débilmente con eosina), tarea de la que se ocupan —en respuesta a señales fisiológicas, como ciertas hormonas— los llamados **activadores de la transcripción**.

Dichos activadores son polipéptidos que se unen a regiones del ADN denominadas **secuencias potenciadoras**, muy alejadas del punto de inicio de la transcripción a veces [véase la ilustra-

ción 6.17]. Los activadores reclutan, a su vez, a otras proteínas genéricamente conocidas como **coactivadores**; entre ellas destacan ciertas enzimas capaces de añadir grupos acetilo a las histonas, disociándolas transitoriamente del ADN, o complejos de remodelación de la cromatina que desplazan a los nucleosomas, facilitando así el acceso de la maquinaria de transcripción al ADN.



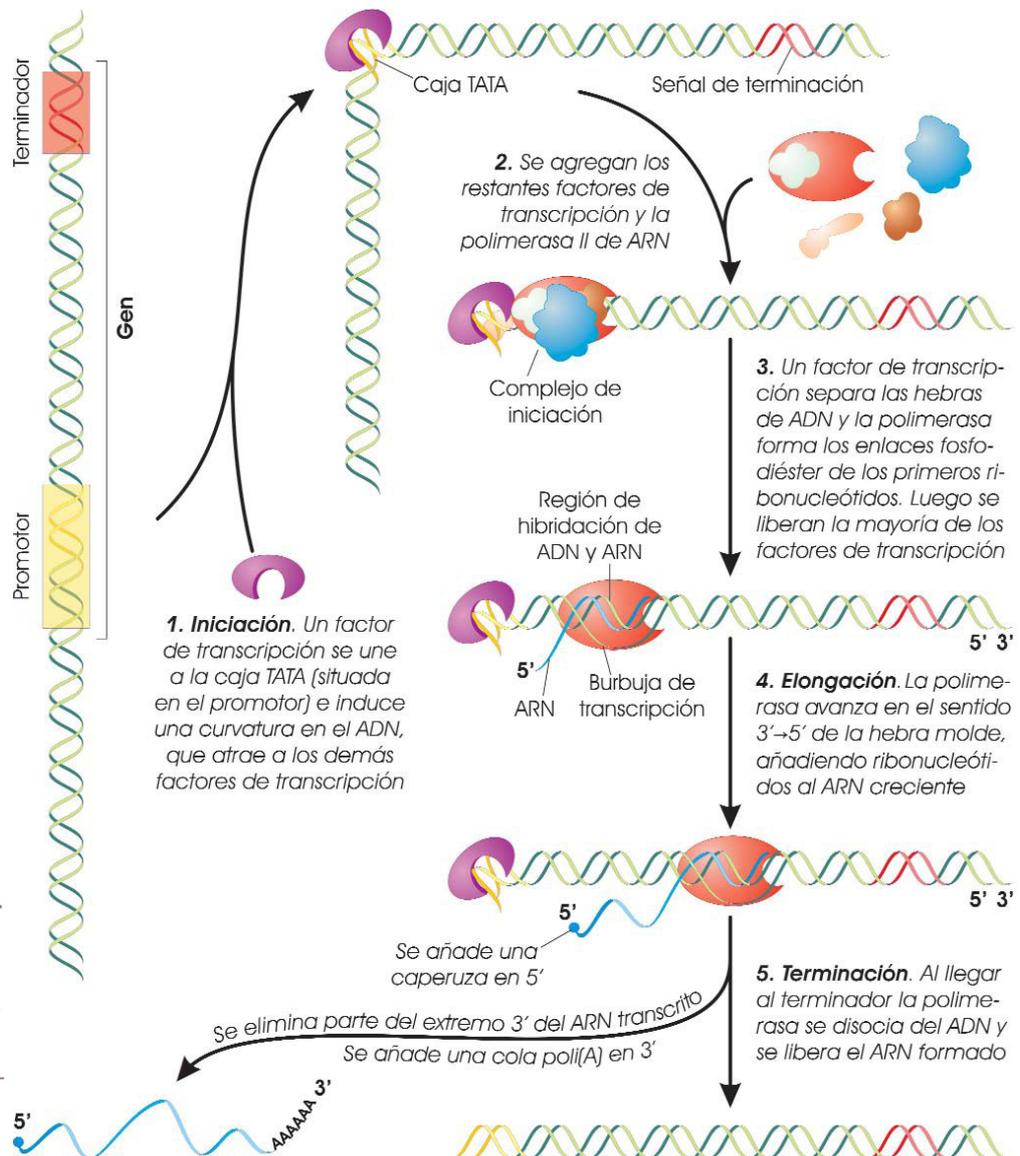
**Ilustración 6.17.** El mediador transmite las señales de regulación desde los activadores de la transcripción hasta el complejo de iniciación del proceso, formado por la polimerasa II y diversos factores de transcripción (Fuente: ASH).

Tras activarse los genes puede dar comienzo la transcripción, que por conveniencia se divide en tres etapas: iniciación, elongación y terminación [véase la ilustración 6.18].

**1. Iniciación.** La polimerasa II es incapaz por sí sola de localizar al promotor, unirse a él y comenzar a transcribir. Por tal razón, requiere la asistencia de una colección de proteínas conocidas como **factores de transcripción**. El proceso se inicia cuando uno de tales factores reconoce, en el interior del promotor, una corta secuencia de ADN que incluye el tetranucleótido dT-dA-dT-dA y se llama, por ello, caja **TATA**. La unión del factor de transcripción a la caja TATA causa una distorsión en el ADN que actúa como señal para atraer a los demás factores de transcripción y a la polimerasa II. Se ensambla así un agregado proteínico, el **complejo de iniciación**, que interactúa con los activadores gracias a otro complejo, el llamado **mediador** [véanse las ilustraciones 6.17 y 6.18].

Uno de los factores de transcripción puede desenrollar el ADN “cortándolo y pegándolo” (igual que la *girasa del ADN* de la ilustración 6.11), dejando libre una hebra que puede ser leída por la polimerasa II. Una vez la polimerasa ha engranzado los primeros nucleótidos del ARN sufre un cambio conformacional que la libera del complejo de iniciación, lo que marca el fin de esta primera etapa.

2. **Elongación**, o alargamiento de la cadena de ARN debido a la polimerasa II. En su transcurso —tras haberse transcrito entre 25 y 30 nucleótidos— se añade al extremo 5' de la nueva cadena de ARN la **caperuza**, consistente en un resto de trifosfato de guanosina con un grupo metilo en posición 7 ( $m^7Gppp$ ); sirve para distinguir al ARNm de otros tipos de ARN, para evitar su degradación por unas enzimas llamadas exonucleasas y para regular su exportación al citosol y su traducción.
3. **Terminación**. Tiene lugar cuando la polimerasa reconoce la secuencia de terminación dT-dT-dA-dT-dT- dT; la enzima continúa transcribiendo, pero ciertos complejos proteínicos asociados a la misma se unen a la secuencia A-A-U-A-A-A recién formada en el ARN y promueven la **escisión** de la molécula en dos fragmentos; el fragmento de ARN que estaba situado hacia el extremo 3' se degrada rápidamente, pero el otro fragmento queda protegido gracias a la **poliadenilación**, es decir, a la adición de una **cola poli(A)** integrada por entre 50 y 250 restos de adenosina [véase la ilustración 6.19] que también interviene en el transporte del ARN y en su traducción.



**Ilustración 6.18.** Secuencia de acontecimientos que conducen a la transcripción de un gen por la polimerasa II (fuente: ASH).

Parece razonable pensar que, tras la terminación, el ARNm está listo para exportarse al citoplasma y traducirse. Pero, como vamos a ver, la realidad es muy diferente.

### Corte y empalme (ajuste)

Los trabajos de Gamow y otros científicos en los años cincuenta del siglo XX habían permitido formular el llamado **principio de colinealidad**, según el cual existía una relación entre la ordenación lineal de los nucleótidos en el ADN y la ordenación lineal de los aminoácidos en las proteínas. La idea cobró fuerza tras el experimento llevado a cabo en 1964 por el genetista estadounidense Charles Yanofsky (n. 1925) y sus colaboradores. Examinaron distintas mutaciones en un mismo gen de la bacteria *Escherichia coli*, cada una de las cuales cambiaba un aminoácido diferente de una proteína; y hallaron que cuanto más alejada se hallaba la mutación del origen del gen tanto más alejado estaba el aminoácido alterado del origen de la cadena polipeptídica.

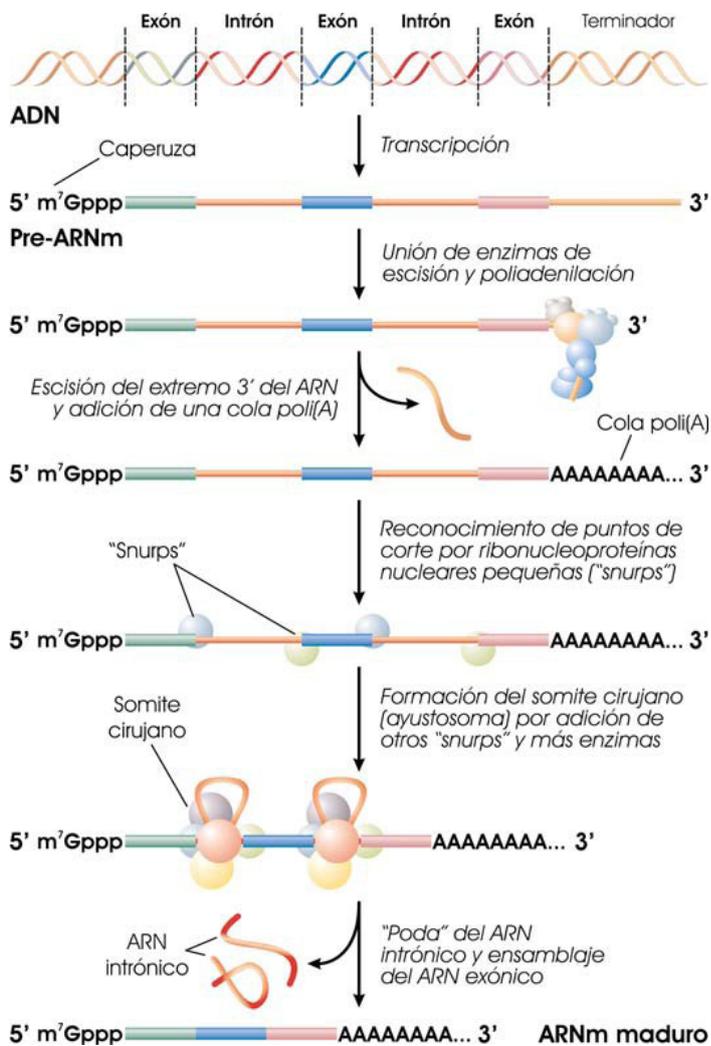


Ilustración 6.19. Procesamiento del ARNm: formación de la caperuza, escisión y poliadenilación y ajuste de exones (Fuente: ASH).

Por supuesto, *Escherichia coli* no es un eucariota, aunque nadie dudaba ni por un momento que también en las células eucarióticas se cumpliría el principio de colinealidad: lo contrario sería como imaginar que dos párrafos contiguos en el original ruso de *Guerra y paz* se situarían en páginas muy distanciadas en la versión española. Como decía Monod, “lo que vale para *Escherichia coli* vale también para el elefante”.

Pero una vez más lo impensable resultó ser lo correcto. Dos biólogos, el inglés Richard John Roberts (n. 1943) y el estadounidense Phillip Allen Sharp (n. 1944), descubrieron en 1977 que, tras su síntesis, el transcrito primario del adenovirus causante del resfriado común era cortado en trozos; algunos de esos trozos se eliminaban, y los restantes eran “empalmados”. Posteriores investigaciones mostraron que este proceso es habitual en células eucariotas y permitieron identificar la maquinaria responsable del mismo: designada a veces como **somite cirujano**<sup>2</sup> [véase la ilustración 6.19], está formada esencialmente por varias **ribonucleoproteínas nucleares pequeñas** —conocidas coloquialmente como “snurps”, por el acrónimo del término inglés **s**mall **n**uclear **r**ibonucleoproteins—, que no son

<sup>2</sup> El nombre original es *spliceosome*, derivado del inglés *splice* (“empalmar”). A veces se traduce directamente por “espliceosoma”, aunque es preferible decir **empalmosoma** o, mejor aún, **ayustosoma** (*ayuste* es una voz de origen náutico que significa “costura y unión de dos cabos”).

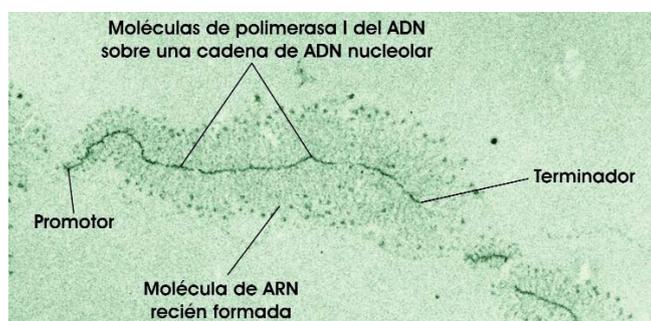
sino asociaciones de proteínas y moléculas de ARN relativamente cortas llamadas ARNnp, o **ARN nucleares pequeños**.

En definitiva, los genes eucariotas se hallan *fragmentados* (con la excepción de los genes de las histonas y algunos otros), de manera que las secuencias que finalmente aparecerán transcritas en el ARNm y, por tanto, se traducirán a proteínas —secuencias a las que se conoce como **exones**— se presentan intercaladas con **intrones**, o secuencias que no cifran proteínas. Los transcritos primarios de ARNm, o **preARNm**, serían como libros que contienen capítulos enteros carentes de significado, que hay que extirpar para obtener un relato coherente.

El descubrimiento de este proceso de **edición del ARN** en 1977 no acabó con las sorpresas. Tan solo tres años después quedó patente que la maquinaria celular puede “decidir” qué capítulos —o partes de los mismos— elimina y cuáles amnistía. En otras palabras, el preARNm puede ser editado de varias formas: los somites cirujanos pueden “atacar” en puntos de corte diferentes, según el tipo celular o la etapa de desarrollo del tejido. Hasta tres cuartas partes de los genes humanos se hallan sujetos a este proceso de **edición alternativa**, lo que explica la paradoja de que seamos capaces de fabricar más de 90 000 tipos distintos de proteínas con solo 20 000 o 25 000 genes... y pone en entredicho el famoso principio de “un gen, un polipéptido”.

### Síntesis y procesamiento de otros ARN

La eucromatina no solo contiene los genes que se transcriben a ARNm, sino que incluye también los responsables de la síntesis de **ARNt**, dirigida por la **polimerasa III**. El proceso es similar a la síntesis de ARNm, con la salvedad de que las regiones promotoras están ubicadas *dentro* de la secuencia que se transcribe. El preARNt experimenta también un proceso de maduración, que incluye la eliminación de parte del extremo 5', la adición de la secuencia C-C-A al extremo 3', la modificación de múltiples bases nitrogenadas y, en ocasiones, la eliminación de una secuencia intrónica.



**Ilustración 6.20.** Microfotografía electrónica de la transcripción de un preARNr (Fuente: <http://commons.wikimedia.org/wiki/>. Cortesía de Dr. Hans-Heinrich Trepte).

En cuanto a los **ARNr**, uno de ellos se sintetiza en el nucleoplasma gracias a la polimerasa III, y el resto lo hace en el nucléolo por intervención de la **polimerasa I**. En este último caso, los genes correspondientes se disponen uno a continuación del otro, de manera que se transcriben a la vez y originan un único preARNr. Seguidamente, esta larga cadena se escinde y modifica por acción de unas moléculas de ARN designadas como ARNnop (**ARN nucleolares pequeños**). Como consecuencia de la escisión se forman tres ARNr maduros que, tras asociarse con proteínas, originan las subunidades ribosómicas y se exportan al citoplasma. Merece la pena señalar que muchos ARNnop son en realidad fragmentos de ARN eliminados por ajuste de ARNm; es decir, están cifrados en los intrones de otros genes. Además, en el nucléolo se fabrican también otros ARN, como ciertos ARNnp.

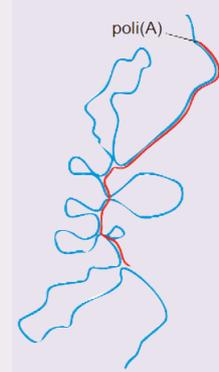
### Transcripción en bacterias y arqueas

En las bacterias, así como en las mitocondrias y cloroplastos, solo existe un tipo de polimerasa que sintetiza todos los ARN. Es más sencilla que las polimerasas eucariotas y no requiere factores de transcripción para unirse al promotor ni el desempaquetado previo del ADN. Además, los genes que codifican las proteínas de las bacterias no están fragmentados, y el ARNm no sufre procesos de maduración —algo comprensible si se tiene en cuenta que se sintetiza en el propio citoplasma, y es inmediatamente “colonizado” por ribosomas para traducirlo a proteínas antes incluso de que finalice su síntesis—. No obstante, se han hallado intrones en ciertos genes que codifican ARNr y ARNt bacterianos, así como en algunos genes mitocondriales de plantas y hongos.

Por último, cabe destacar que la transcripción en las arqueas se asemeja más a la eucariótica que a la bacteriana: poseen una sola polimerasa, pero es similar a la de los eucariotas, como también lo es la presencia del promotor o de factores de transcripción.

## Actividades

7. Si combináramos de dos en dos los cuatro nucleósidos, siendo relevante el orden (es decir, si GA no “fuese igual” que AG), ¿cuántos aminoácidos podríamos determinar? ¿Y si no importase el orden? Repite el razonamiento si los combináramos de tres en tres.
8. Teniendo en cuenta los resultados de la actividad 7, ¿existiría alguna posibilidad de un flujo de información inverso desde las proteínas al ADN? ¿Hay alguna relación entre el “dogma” central de la Biología molecular y la imposibilidad de heredar los caracteres adquiridos?
9. La secuencia 3'...dT-dA-dC-dC-dA-dC-dG-dT-dG-dG-dA-dC-dT-dG-dA...5' pertenece a una hebra del gen HBB de la hemoglobina. ¿Cómo será la hebra complementaria? ¿Qué secuencia tendrá el ARNm formado tras actuar la polimerasa II?
10. La ilustración adjunta esquematiza una microfotografía de un híbrido entre una cadena de ADN de un gen (azul) y el ARNm que codifica (en rojo). Señala los extremos 3' y 5' de cada uno. ¿Qué significado tienen los bucles observados?



## Recuerda

- ▶ Según el “dogma” central de la Biología molecular, el ADN se transcribe a ARN y este se traduce a proteínas, sin posibilidad de un flujo de información inverso. Hay varias clases de ARN:
  - **Mensajeros** (ARNm), que cifran la información precisa para sintetizar proteínas.
  - **Ribosómicos** (ARNr), que forman los ribosomas y catalizan la síntesis proteínica.
  - **Transferentes** (ARNt), que transportan aminoácidos hasta los ribosomas.
  - **ARN pequeños nucleares** (ARNnp) y **nucleolares** (ARNnop), que intervienen en el procesamiento y maduración del ARNm y ARNr, respectivamente.
- ▶ Las **polimerasas de ARN** sintetizan ARN (en sentido 5' → 3') tras unirse a un **promotor**. El transcrito primario no sufre modificación alguna en las bacterias, pero en los eucariotas experimenta un proceso de **edición**, como el ajuste de **exones**.

## 3. La traducción

En el interior de la célula eucariota existe un tráfico incesante a través de los complejos del poro de la envoltura nuclear. En el núcleo penetran todos los sillares necesarios para la síntesis de ARN, entre ellos los nucleótidos GTP, CTP, ATP y UTP y una dosis adicional de ATP que completará el suministro de energía del que carece el núcleo. Asimismo hay un notable flujo de proteínas hacia el interior: enzimas como las polimerasas de ARN, activadores y factores de transcripción, laminas nucleares, histonas y todas las proteínas que se han de ensamblar para formar los ribosomas.

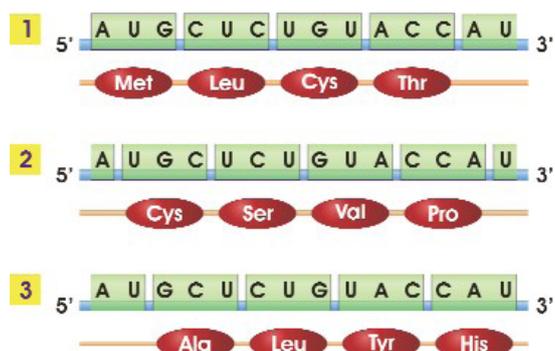
En sentido inverso, la misma vía da salida a moléculas como el AMP o el ADP, que retornan al citoplasma para recargarse, o a nucleótidos como el NAD<sup>+</sup>. Pero lo más llamativo es el vertido incesante de filamentos alargados, que atraviesan cada poro a la velocidad de 1  $\mu\text{m}$  por minuto; asociados con proteínas de transporte, se trata de hebras de ARN, ya maduras, que aportan los “planos” arquitectónicos (ARNm) o el material de construcción (ARNt) necesarios para la **síntesis de proteínas**. Y, por último, ocasionalmente se produce la salida de las subunidades de los ribosomas.

### 3.1. El código genético

Se ha comentado ya que la interpretación de los “planos” de ARNm por parte de los ribosomas es básicamente un proceso de traducción. Y como todo traductor, el ribosoma debe disponer de un diccionario. Recordemos [véase la actividad 7] que las entradas del mismo deben ser combinaciones de tres nucleótidos: las combinaciones binarias (con repetición) generarían 16 palabras, o sea, solo permitirían la incorporación de 16 aminoácidos a las proteínas. Crick propuso el término **codón** para designar a cada uno de los 64 *tripletes* de nucleótidos del ARNm y formuló una pregunta fundamental: ¿cómo averiguar en qué términos está redactado el diccionario?

#### El código “sin comas”

El propio Crick sugirió una respuesta que, en realidad, era una alternativa al fallido código de Gamow. Crick se dio cuenta de que una secuencia de ARNm podría traducirse siguiendo tres posibles **marcos de lectura** distintos, según se explica en la ilustración 6.21. ¿Cómo reconocerían las células cuál es la pauta correcta? ¿Acaso existirían “signos de puntuación” especiales (**comas**) que delimitasen los codones verdaderos evitando lecturas alternativas?



**Ilustración 6.21.** Un código sin comas podría dar lugar a distintos marcos de lectura: el mismo ARNm especificaría tres secuencias de aminoácidos, según dónde comenzase la lectura (Fuente: ASH).

El modelo de Crick eludía la necesidad de comas con un *diccionario de palabras sin sentido*, que limitaba el número de codones que cifran aminoácidos. Si, en el ejemplo de la ilustración 6.21, AUG<sup>3</sup> y CUC codificasen aminoácidos (Met y Leu, respectivamente), se evitaría que un cambio del marco de lectura originase un polipéptido incorrecto si UGC y GCU careciesen de sentido, esto es, si no definiesen ningún aminoácido —en la ilustración, en cambio, sí lo hacen—. Un cálculo sencillo mostraba que solo podrían existir, en tal caso, 20 codones con sentido... justo los correspondientes a los 20 aminoácidos identificados por entonces. Los 44 codones restantes serían **codones sin sentido**.

### El descifrado del código genético

El modelo de Crick era tan estilizado y coherente que se granjeó la aceptación casi universal... hasta el punto de ser calificado como “la más elegante teoría biológica que ha resultado ser incorrecta”. Efectivamente, la naturaleza parece haber optado por un esquema menos “perfecto”. Su escudriñamiento se remonta a 1955, cuando la bioquímica franco-rusa Marianne Grunberg-Manago (n. 1921) aisló a partir de extractos bacterianos una enzima llamada **polinucleótido fosforilasa** (PNPasa). La PNPasa cataliza la hidrólisis del ARN, pero también puede sintetizar ARN a partir de una mezcla de difosfatos de ribonucleósido que engarza al azar, sin necesidad de molde previo, lo cual posibilita la obtención *in vitro* de ARNm cuyas secuencias estén en cierto modo controladas.

Con auxilio de la PNPasa el bioquímico estadounidense Marshall Warren Nirenberg (n. 1927) y el alemán Johann Heinrich Matthaei (n. 1929) sintetizaron en 1961 un ARN formado únicamente por nucleótidos de uracilo —un poli(U)— y lo añadieron a un *sistema libre de células*: un tubo de ensayo que solo contenía ribosomas, enzimas, ARNt, aminoácidos y ATP como fuente de energía. Se originó un polipéptido formado por una sucesión de restos del aminoácido fenilalanina (Phe); puesto que el único codón posible en el poli(U) era UUU, estaba claro que, en el diccionario genético, UUU = Phe. Ensayos análogos demostraron que

<sup>3</sup> Al describir los codones se omiten los guiones entre nucleósidos que simbolizan enlaces fosfodiéster. Así, no se denotará A-U-G, sino AUG. Se asume que los codones se escriben siempre en sentido 5' → 3'.

el triplete CCC codifica para prolina (Pro) y AAA para la lisina (Lys); es decir, hasta los codones que según Crick presuntamente carecían de significado [véase *la actividad 11*] podían provocar la síntesis de proteínas. Más adelante, el biólogo de origen hindú Har Gobind Khorana (n. 1922) empleó ARNm sintéticos compuestos por nucleótidos alternantes [véanse *las actividades 13 y 14*]. Con este y otros tipos de métodos, en unos cuantos años se pudo desvelar la totalidad del código.

### Características del código genético

1. Las primeras investigaciones se llevaron a cabo en la bacteria *Escherichia coli*, pero la experimentación postrera condujo a la conclusión de que el código es **casi universal**: cada codón tiene el mismo significado para la mayoría de los organismos, tanto eucariotas como procariotas; si bien se conocen excepciones repartidas por una amplia gama de linajes, en particular en el sistema genético de las mitocondrias [véase *la ilustración 6.22*].
2. Aunque se trata de un código **sin superposición** (es decir, un nucleósido no puede formar parte de más de un codón o, lo que es lo mismo, dos codones sucesivos no comparten ningún nucleósido) y **sin comas**, no es necesario un diccionario de palabras sin sentido, como sugería Crick. Para evitar cambios en el marco de lectura se precisan tan solo un par de “signos de puntuación”:
  - Una **señal de inicio**, que suele corresponder al codón AUG (el cual, además, cifra el aminoácido metionina) y determina el marco de lectura correcto.
  - Una **señal de terminación** de la traducción o “punto final”, dada por tres codones sin sentido (en lugar de los cuarenta y cuatro de Crick) que no cifran aminoácidos. Habitualmente son UAA, UAG y UGA, señalizados mediante un stop en la ilustración 6.22.

Hay que advertir que el codón UGA no siempre actúa como codón de terminación; cuando una secuencia de nucleósidos especial llamada SECIS se halla presente en el ARNm, el codón UGA determina la selenocisteína (Sec), conocida como el *vigésimo primer aminoácido*. Este fenómeno se da en todos los organismos investigados hasta la fecha. Análogamente, en ciertas arqueas el codón UAG cifra el *vigésimo segundo aminoácido*, la pirrolisina (Pyl).

3. Como corolario del escaso número de codones sin sentido el código es **degenerado**, esto es, redundante: en su mayoría, los aminoácidos están representados por más de un codón. Cabe señalar que los codones *sinónimos* tienden a diferir en la base del extremo 3', siendo iguales las bases de las otras dos posiciones.

		Segundo nucleósido			
		U	C	A	G
Primer nucleósido (extremo 5')	U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys
		UUC Phe	UCC Ser	UAC Tyr	UGC Cys
		UUA Leu, <i>Stop</i> <sup>16</sup> <i>Inicio</i>	UCA Ser, <i>Stop</i> <sup>15</sup>	<b>UAA Stop, Gln<sup>7</sup></b>	<b>UGA Stop, Sec<sup>1</sup>, Trp<sup>2,3,4,5,6,8,11,14</sup>, Cys<sup>9</sup></b>
		UUG Leu <i>Inicio</i>	UCG Ser	<b>UAG Stop, Gln<sup>7,12</sup>, Leu<sup>13,15</sup>, Pyl<sup>17</sup></b>	UGG Trp
	C	CUU Leu, Thr <sup>3</sup>	CCU Pro	CAU His	CGU Arg
		CUC Leu, Thr <sup>3</sup>	CCC Pro	CAC His	CGC Arg
		CUA Leu, Thr <sup>3</sup>	CCA Pro	CAA Gln	CGA Arg
		CUG Leu, Thr <sup>3</sup> , Ser <sup>10</sup> <i>Inicio</i>	CCG Pro	CAG Gln	CGG Arg
	A	AUU Ile <i>Inicio</i>	ACU Thr	AAU Asn	AGU Ser
		AUC Ile <i>Inicio</i>	ACC Thr	AAC Asn	AGC Ser
		AUA Ile, Met <sup>2,3,6,11</sup> <i>Inicio</i>	ACA Thr	AAA Lys, Asn <sup>8,14</sup>	AGA Arg, <i>Stop</i> <sup>2</sup> , Ser <sup>6,8,14</sup> , Gly <sup>11</sup>
		AUG Met <i>Inicio</i>	ACG Thr	AAG Lys	AGG Arg, <i>Stop</i> <sup>2</sup> , Ser <sup>6,8,14</sup> , Gly <sup>11</sup>
	G	GUU Val	GCU Ala	GAU Asp	GGU Gly
		GUC Val	GCC Ala	GAC Asp	GGC Gly
		GUA Val	GCA Ala	GAA Glu	GGA Gly
		GUG Val <i>Inicio</i>	GCG Ala	GAG Glu	GGG Gly

<sup>1</sup> Todas las especies, cuando el elemento SECIS (secuencia especial de nucleósidos) está presente en el ARNm  
<sup>2</sup> Mitocondrias de vertebrados  
<sup>3</sup> Mitocondrias de *Saccharomyces cerevisiae* (levadura)  
<sup>4</sup> Mitocondrias de algunos mohos, protozoos y cnidarios  
<sup>5</sup> *Mycoplasma*, *Spiroplasma* (bacterias)  
<sup>6</sup> Mitocondrias de invertebrados  
<sup>7</sup> Protozoos ciliados, *Acetabularia* (planta marina)  
<sup>8</sup> Mitocondrias de equinodermos y algunos platelmintos  
<sup>9</sup> Ciliados euplótidos  
<sup>10</sup> Varias especies del género *Candida* (levaduras)  
<sup>11</sup> Mitocondrias de ascidias  
<sup>12</sup> *Blepharisma* (ciliado)  
<sup>13</sup> Mitocondrias de clorófitos (algas verdes)  
<sup>14</sup> Mitocondrias de tremátodos  
<sup>15</sup> Mitocondrias de *Scenedesmus obliquus* (alga verde)  
<sup>16</sup> Mitocondrias de *Thraustochytrium aureum* (cromista)  
<sup>17</sup> Arqueas metanógenas

**Ilustración 6.22.** El código genético se presenta en una tabla con 16 celdas, cada una de las cuales incluye los aminoácidos cifrados por los cuatro codones cuyo primero y segundo nucleósidos (indicados en los márgenes izquierdo y superior, respectivamente) son iguales. El código genético estándar aparece con caracteres en negrita; en caracteres más finos se indican las variantes registradas hasta la fecha (Fuente: ASH).

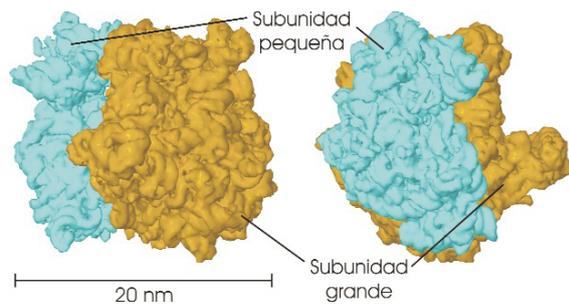
### 3.2. La síntesis de proteínas

Una vez descifrado el código genético, quedaba por aclarar el mecanismo de la traducción. Como sabemos, los codones de una molécula de ARNm no reconocen directamente los aminoácidos que especifican; es necesaria una compleja maquinaria en la que toman parte los ARNt, los ribosomas y una variada colección de enzimas.

#### Los ribosomas

Los responsables directos de la síntesis de proteínas son unas pequeñas partículas compactas, de forma algo oblonga y con unas dimensiones aproximadas de entre 15 y 25 nm: los **ribosomas**.

Los ribosomas carecen de membrana y solo se pueden observar con el microscopio electrónico. Habitualmente se asocian en hileras formadas por diez o más, llamadas **polisomas** (síncopa de *polirribosomas*). Existen dos clases de ribosomas: los que se hallan dispersos a millones por el citosol o adheridos al retículo endoplasmático de las células eucarióticas son **ribosomas 80S**, mientras que en el interior de las bacterias, mitocondrias y cloroplastos se hallan los **ribosomas 70S** (la “S” representa el *svedberg*, una unidad que indica la velocidad a la que sedimentan al centrifugarlos, y que se relaciona con su tamaño).



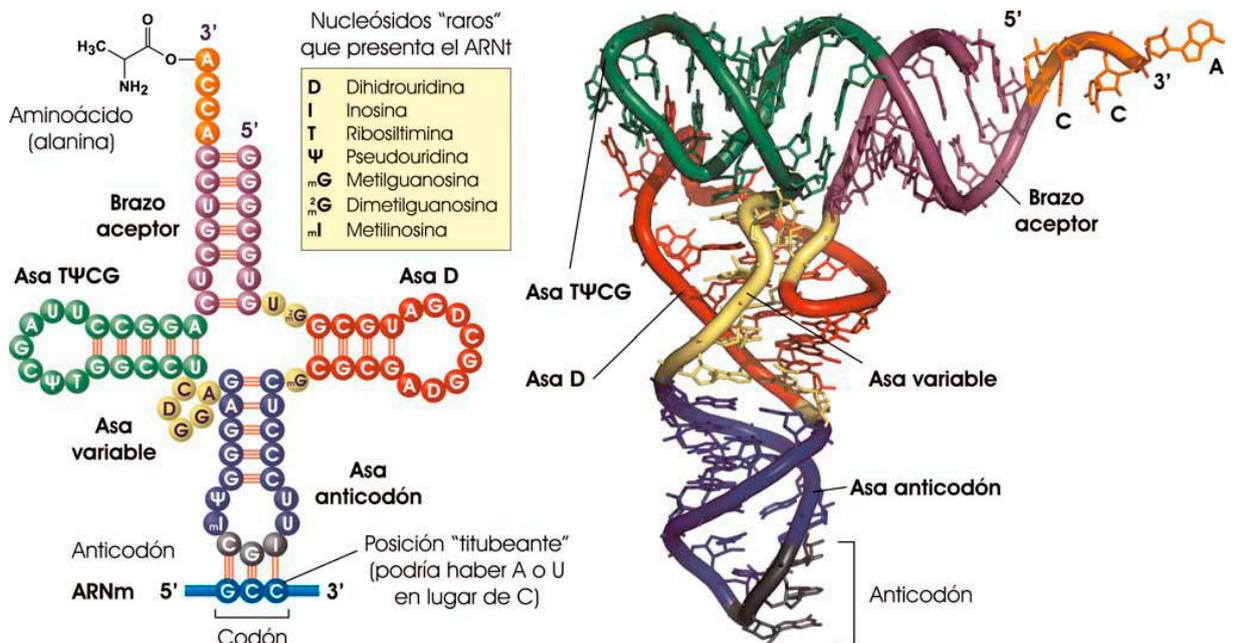
**Ilustración 6.23.** Representación esquemática de la superficie molecular de un ribosoma, mostrando las subunidades grande y pequeña (Fuente: ASH).

Cada ribosoma se compone de más de cincuenta proteínas y varias moléculas de un ácido nucleico, el ARN, organizadas en una subunidad grande y otra pequeña [véase la ilustración 6.23]. Los ribosomas reciben instrucciones del ARNm que indican con precisión qué aminoácidos tienen que engarzar y en qué orden; es decir, en estos orgánulos tiene lugar la traducción.

**La maquinaria de traducción**

El ARNt depende críticamente, para llevar a cabo su función, del plegamiento adecuado, debido a enlaces de hidrógeno. Se forma así una estructura compacta a modo de “L”. En uno de sus lados se halla el **anticodón** [véase la ilustración 6.24], que puede formar pares de bases con un codón complementario; por el otro lado —en el extremo 3’ de la molécula—, se une al aminoácido correspondiente al codón.

Si la complementariedad entre codones y anticodones fuese perfecta debería haber tantos tipos de ARNt como codones con sentido. Sin embargo, algunos ARNt solo requieren que se emparejen correctamente las dos primeras bases del codón, tolerando cierto **titubeo** (en inglés, *wobble*) en la tercera posición. Por ejemplo, el anticodón AAG (escrito en sentido 3’ → 5’) puede reconocer los codones UUC y UUU (ambos en sentido 5’ → 3’). A veces en la posición de titubeo del ARNt hay una guanosina desaminada, la **inosina**, capaz de aparearse con A, C y U, lo que facilita el reconocimiento de varios codones [véase la ilustración 6.24]. Esto explica por qué muchos codones sinónimos solo se



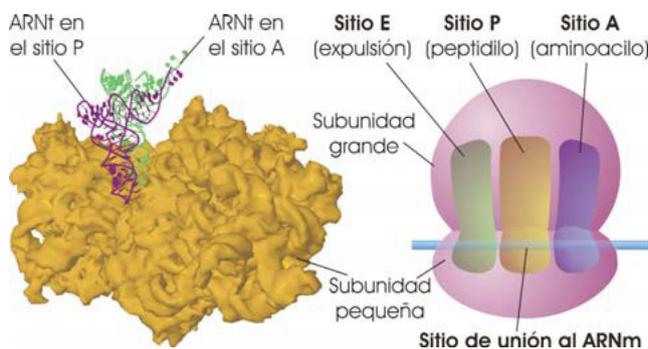
**Ilustración 6.24.** Izquierda: Estructura en “hoja de trébol” de un ARNt, un convenio usado para mostrar zonas donde se aparean bases complementarias de una misma cadena (brazos) y donde no lo hacen (asas); el nombre de dos de ellas remite a la presencia de bases modificadas químicamente cuando el ARNt ya se ha sintetizado. El brazo aceptor es el que se une al aminoácido, y el anticodón es la secuencia de tres nucleótidos que se enlaza con un codón. Derecha: Estructura real (tridimensional) de la molécula, en forma de “L” (Fuente: ASH y <http://en.wikipedia.org/wiki>).

distinguen en la tercera posición y, por lógica, debería permitir que hubiese *menos* tipos de ARNt que codones. Curiosamente, ocurre lo contrario: en el ser humano, por ejemplo, hay 497 genes de ARNt—sin incluir los mitocondriales—, aunque entre todos ellos solo hay representados 48 anticodones diferentes.

Aunque el ARNt actúa como el *adaptador* postulado por Crick, en realidad se necesita otro adaptador, la llamada **sintetasa del aminoacil-ARNt**, que acopla un aminoácido a su correspondiente ARNt (proceso llamado **activación del aminoácido**). Suele haber *una* de tales enzimas para *todos* los ARNt que se unen al mismo aminoácido:



Los aminoacil-ARNt así formados se unen a centros específicos de los ribosomas designados como sitios **A**, **P** y **E** [véase la ilustración 6.25]. En los sitios A y P se puede unir una molécula de ARNt solo si su anticodón está apareado con un codón complementario del ARNm (el cual también se une al ribosoma por un centro específico); ambos sitios se disponen de manera que las dos moléculas de ARNt se vean obligadas a emparejarse con



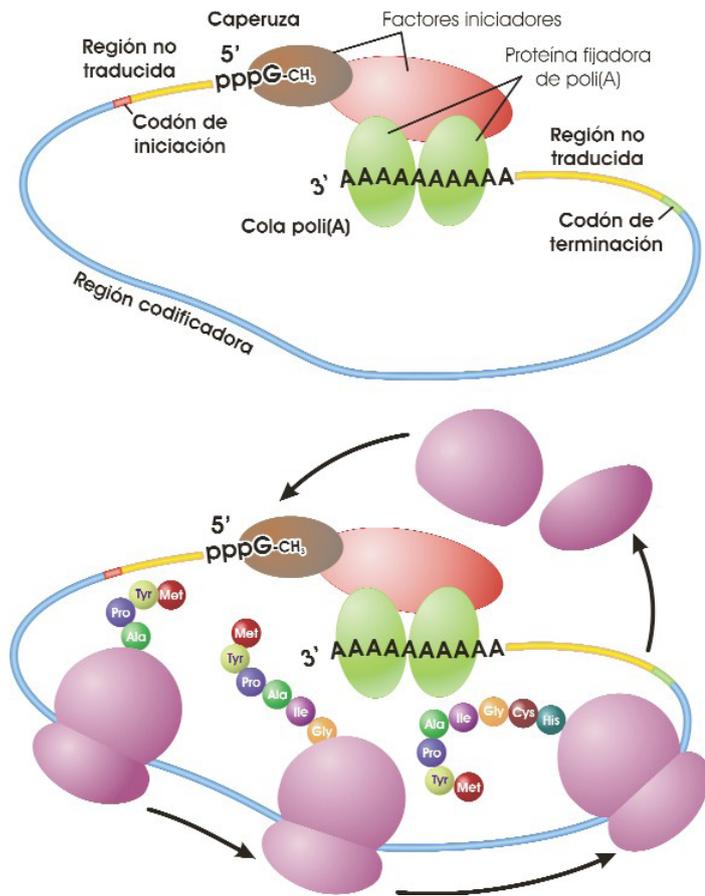
**Ilustración 6.25.** Modelo tridimensional (izquierda) y esquema simplificado (derecha) de las subunidades de un ribosoma y sus diferentes sitios de unión a los ARNt (Fuente: ASH).

codones adyacentes del ARNm, lo que mantiene el marco de lectura correcto. La determinación, en 2000, de la estructura completa del ribosoma—quizá el mayor logro de la biología molecular— permitió comprobar que los sitios A, P y E están formados básicamente por ARNr, y confirmó la sospecha previa de que son los ARNr, y no las proteínas, los responsables de la actividad ribosómica, incluida su habilidad para formar enlaces peptídicos (actividad enzimática conocida como **transferasa del peptidilo**). Estas moléculas de ARN con actividad catalítica se llaman **ribozimas**.

En una misma molécula de ARNm pueden coincidir varios ribosomas, formando **polisomas**, separados por distancias de 80 nucleótidos. Cada ribosoma se une al extremo 5' del ARNm, recorre la molécula generando la proteína (lo que lleva entre 20 y 120 segundos) y se desensambla en sus subunidades por el extremo 3'. El proceso se acelera gracias a proteínas que unen los extremos 5' y 3' del ARNm, haciendo que los polisomas sean circulares [véase la ilustración 6.26].

### Etapas en la síntesis de proteínas

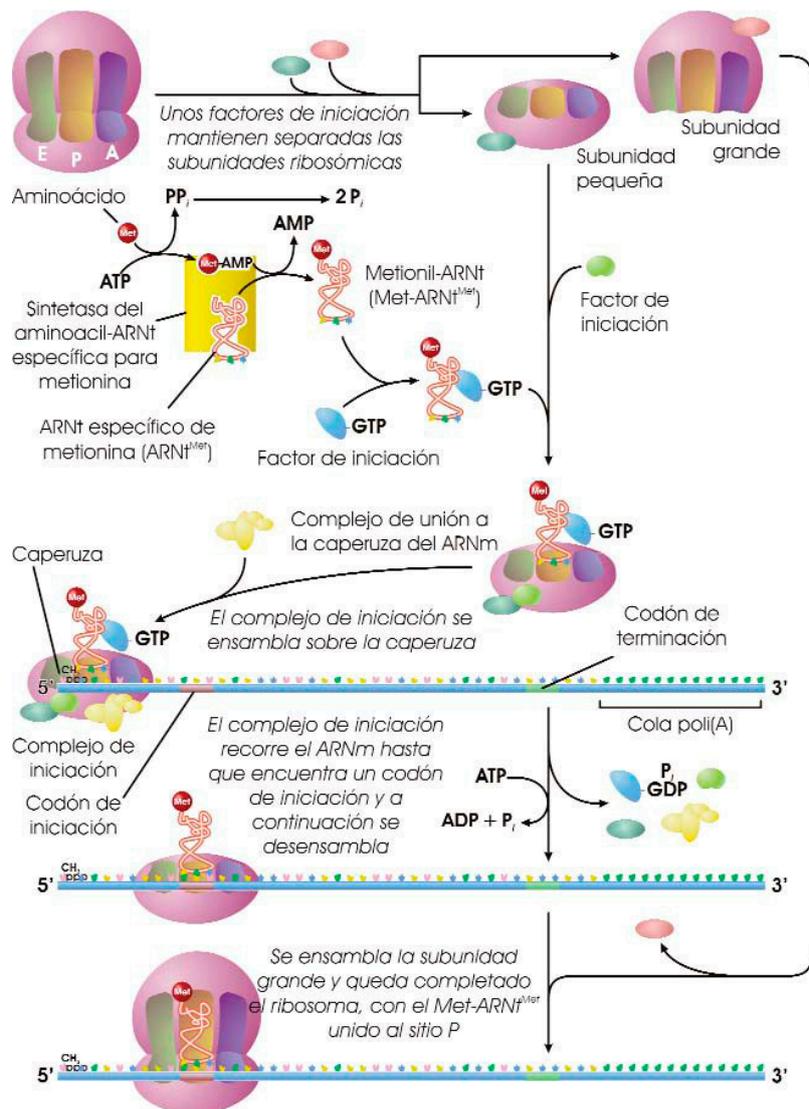
La traducción puede dividirse en una serie de etapas cuyos nombres coinciden con los de la transcripción, lo que quizá de lugar a confusiones:



**Ilustración 6.26.** Arriba: Estructura del ARNm de una célula eucariota, con los extremos 5' y 3' unidos. Abajo: Polisoma en el que las subunidades ribosómicas, una vez acabada la traducción, hallan fácilmente el extremo 5' del ARNm y pueden comenzar otra ronda rápidamente (Fuente: ASH).

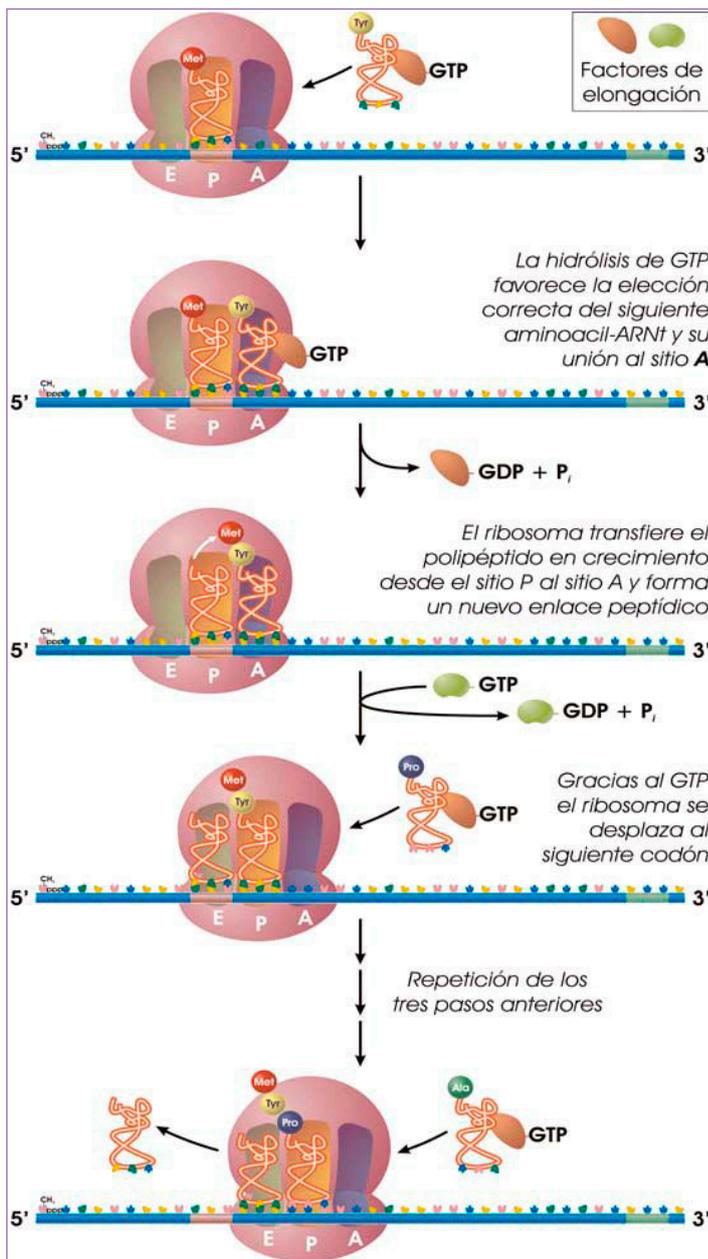
- 1. Iniciación.** La síntesis de proteínas en los eucariotas comienza cuando un ARNt especial llamado **ARNt iniciador**, que lleva *metionina*, se asocia con la subunidad menor del ribosoma y con proteínas denominadas **factores de iniciación** o **eIF**. El complejo formado se une a la caperuza del ARNm y recorre la molécula hasta encontrar un codón de inicio AUG; los eIF se disocian y permiten ensamblarse a la subunidad mayor, dejando al metionil-ARNt en el sitio P. El proceso requiere energía en forma de ATP y GTP [véase la ilustración 6.27].
- 2. Elongación.** Empieza cuando en el sitio P hay un polipéptido unido a un ARNt (peptidil-ARNt), o por lo menos un primer aminoácido. Tras llegar un nuevo aminoacil-ARNt al sitio A, la transferasa del peptidilo le traspasa el polipéptido del ARNt que hay en P, a la par que forma un enlace peptídico. Seguidamente, el ribosoma se desplaza (**translocación**) la distancia equivalente a un codón para expulsar el ARNt que ha perdido el polipéptido al sitio E (y, de ahí, al citosol), permitir el ingreso en el sitio P del peptidil-ARNt recién prolongado y dejar libre el sitio A para el siguiente aminoacil-ARNt (e iniciar un nuevo ciclo). En esta etapa también se requiere energía (GTP) y **factores de elongación** [véase la ilustración 6.28].
- 3. Terminación.** Cuando el sitio A ribosómico se “asoma” a un codón de terminación ya no se une a un nuevo aminoacil-ARNt; en su lugar lo hace a un **factor de liberación** proteínico que

tiene un sorprendente parecido con un ARNt. Esta unión fuerza a la transferasa del peptidilo a catalizar la adición de una molécula de agua al peptidil-ARNt en lugar de un aminoácido, liberando el extremo carboxilo de la cadena polipeptídica de su unión al ARNt. Otros factores de liberación promueven la disociación del ARNm, los ARNt y el ribosoma, procesos que requieren la hidrólisis de GTP [véase la ilustración 6.29].



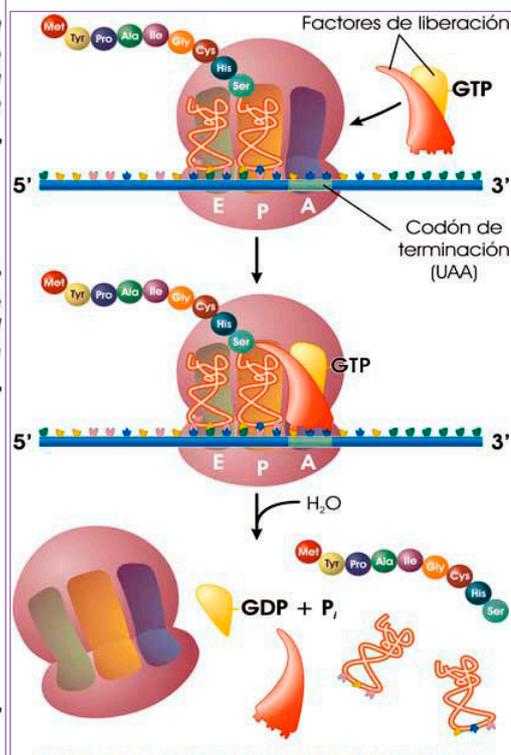
**Ilustración 6.27.** Activación del primer aminoácido que se incorpora a la cadena (*Met*) e iniciación de la traducción con el auxilio de factores proteínicos. La formación de un aminoacil-ARNt está favorecida por la hidrólisis del  $PP_i$  (pirofosfato) a  $P_i$ , que libera energía (Fuente: ASH).

El proceso de síntesis de proteínas no acaba aquí. Como vimos en la Unidad 3, para que una proteína sea útil debe plegarse adecuadamente, tarea en la que es asistida por las llamadas **chaperonas**. Posiblemente también habrá de asociarse a cofactores (coenzimas o iones metálicos) y a otras subunidades proteínicas, para constituir su estructura cuaternaria, o experimentar modificaciones covalentes, procesos que a menudo tienen lugar en el retículo endoplasmático rugoso y en el aparato de Golgi. Todo ello antes de ubicarse en la estructura de la que forman parte, o realizar la función que les corresponda.



**Ilustración 6.28** (izquierda). Elongación de una cadena polipeptídica gracias a la hidrólisis de GTP (que facilita la unión del aminoacil-ARNt al sitio A y la translocación del ribosoma) y a la actividad enzimática (transferasa) del ARNr (Fuente: ASH).

**Ilustración 6.29** (abajo). Terminación de la síntesis de proteínas y liberación del polipéptido formado (Fuente: ASH).



### Traducción en bacterias

La anterior descripción es válida en líneas generales para bacterias, mitocondrias y cloroplastos, aunque se aprecian diferencias. De entrada, los ribosomas de estas células y orgánulos (**70S**) son distintos de los que hay en el citoplasma eucariota (**80S**) y se disponen en polisomas no circulares. Puede también variar algún que otro detalle del código genético. Pero las diferencias más significativas atañen a la activación de aminoácidos y a la iniciación de la traducción:

- Muchas bacterias se las arreglan con menos de 20 sintetetasas del aminoacil-ARNt, lo que significa que una misma enzima puede colocar aminoácidos idénticos en ARNt diferentes, aunque solo uno de ellos posea el anticodón “correcto”; se precisa, pues, una segunda enzima que modifique químicamente el aminoácido “incorrecto” para adecuarlo al ARNt al que se

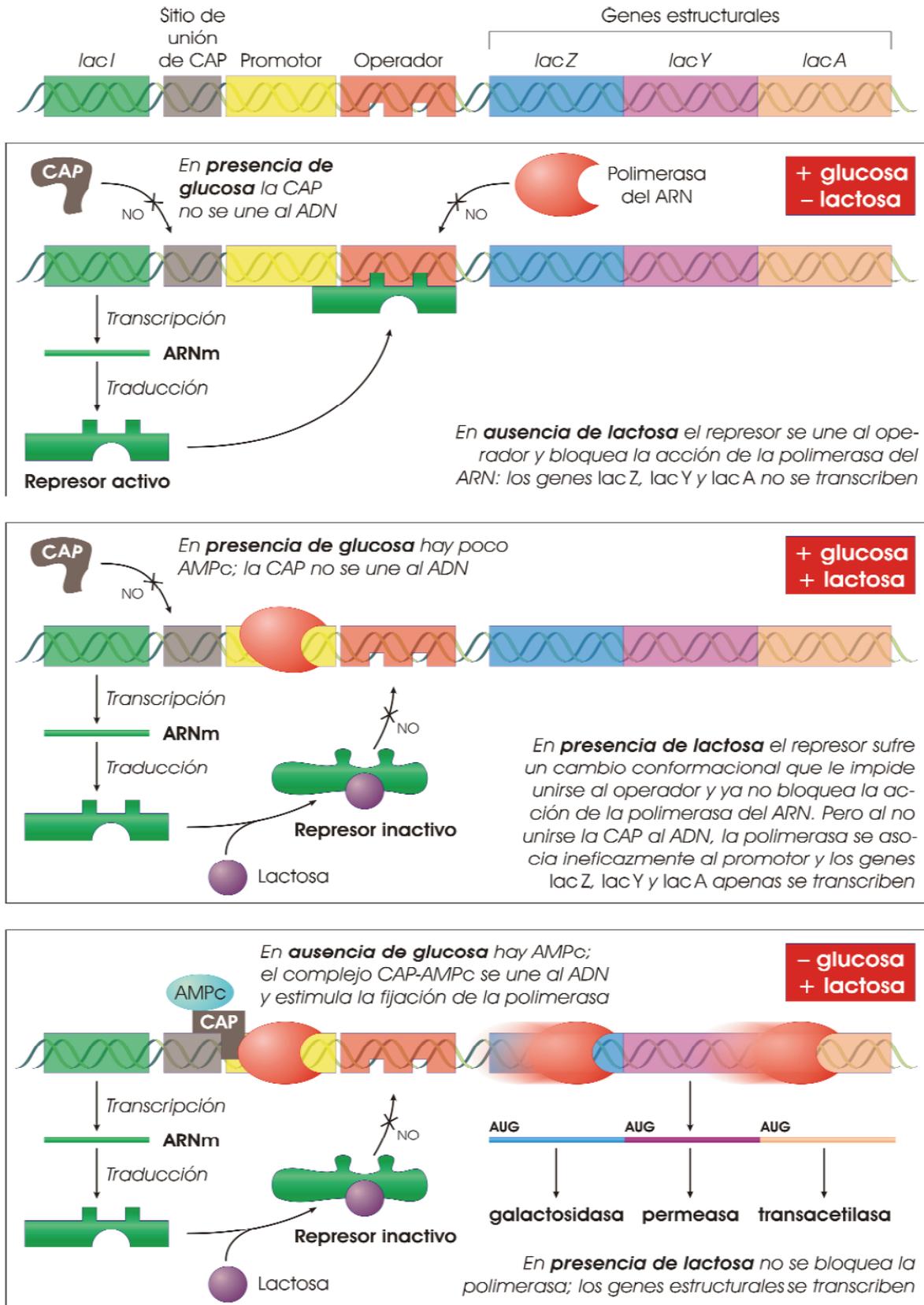
halla unido. Ese proceso afecta a la metionina unida al ARNt iniciador, a la que se añade un grupo formilo (–CHO) que la convierte en *formil-metionina* (el aminoácido que encabeza las proteínas bacterianas).

- También difiere el mecanismo para seleccionar el codón de inicio: en lugar de la caperuza eucariota, el ARNm bacteriano tiene un sitio de unión específico para el ribosoma, la *secuencia de Shine-Dalgarno*, que se ubica unos cuantos nucleótidos por delante del codón AUG y se une al ARNr de la subunidad menor del ribosoma.
- Por lo tanto, el ribosoma bacteriano podrá ensamblarse directamente a un codón de inicio emplazado en el interior de un ARNm, sin necesidad de recorrer esta molécula desde su extremo 5'. Así, *un solo* ARNm bacteriano puede incluir *varias* secuencias que se traduzcan a proteínas diferentes, flanqueada cada una de ellas por sus codones de iniciación y de terminación correspondientes. Este ARNm procederá, pues, de la transcripción conjunta de varios genes localizados uno junto al otro en el cromosoma bacteriano y que comparten un único promotor. Dicha agrupación de genes recibe el nombre de **operón**, porque *opera* como una unidad de transcripción, y los ARNm bacterianos se califican de **policistrónicos** —recuérdese que el cistrón es la unidad genética que cifra un polipéptido—. En cambio, un ARNm eucariota generalmente es **monocistrónico**, es decir, codifica una sola proteína.

### 3.3. Regulación de la expresión génica

El descubrimiento de los operones bacterianos proviene en realidad de estudios sobre un fenómeno que fue analizado con detalle por Jacob y Monod en 1959. Cuando una bacteria como *Escherichia coli* se cultiva en un medio en el que abunda la glucosa, apenas se detectan las enzimas necesarias para utilizar otros nutrientes, como la lactosa. Pero si se transfiere a un medio con lactosa y sin glucosa, al cabo de uno o dos minutos la célula comienza a sintetizar grandes cantidades de tres enzimas involucradas en el transporte y la hidrólisis de lactosa (para dar glucosa utilizable), codificadas por los genes lac Z, lac Y y lac A. Los tres genes forman parte del llamado operón lac.

Jacob y Monod descubrieron que, además de los genes que cifraban enzimas y su promotor colectivo —a los que denominaron genes estructurales—, en el operón lac se dan otro tipo de secuencias. Una de ellas, el gen regulador (*lacI*), codifica una proteína llamada represor; otra es una corta región interior del promotor bautizada como operador, que es reconocida por el represor. Cuando el represor está unido al operador se bloquea el acceso de la polimerasa del ARN al promotor, de modo que se impide la expresión de los genes estructurales. Ese bloqueo de la expresión génica está regulado de una forma que se podría calificar de ingeniosa [véase la ilustración 6.30]:



**Ilustración 6.30.** Características esenciales de la estructura del operón *lac* y de su control dual por los niveles de lactosa (control positivo o **inducción enzimática**) y de glucosa (control negativo o **represión enzimática**). Obsérvese que el transcrito de los genes *lac* tiene tres codones de inicio (AUG), de manera que codifica las tres enzimas necesarias para la utilización de la lactosa como fuente de carbono (Fuente:ASH).

- La **adición de lactosa** incrementa la concentración de un derivado suyo, la alolactosa, que se une a la proteína represora y la separa del ADN. La polimerasa puede ya unirse al promotor; de ahí que la lactosa actúe como **inductor** de la transcripción.
- La **adición de glucosa** reduce los niveles de un nucleótido, el AMP cíclico o AMPc [véase la ilustración 4.6]. La proteína denotada por **CAP** (siglas, en inglés, de “proteína activadora por catabolito”) no puede unirse entonces al AMPc, lo que la incapacita para asociarse al ADN. Al ser dicha asociación necesaria para fijar la polimerasa al promotor, la glucosa actuará como **inhibidor** de la transcripción.

La mayoría de los operones bacterianos se regulan mediante “interruptores genéticos” de este tipo; algunos de ellos serán sensibles a inductores (**regulación positiva**), otros a inhibidores (**regulación negativa**) y unos terceros, como el operón *lac*, a señales de ambos tipos. En los eucariotas, sin embargo, un mecanismo tan elegante y sencillo como este es inviable: son muchas las señales que han de participar en la regulación de un gen, según el tipo celular y la etapa del desarrollo, y no habría suficiente espacio cerca del promotor para incluir tan alto número de secuencias reguladoras.

Como se ha señalado páginas atrás, la regulación en eucariotas depende de mecanismos complejos que involucran el empaquetamiento del ADN en forma de cromatina, la presencia de activadores y represores de la transcripción —que etiquetan el ADN o las histonas (mediante, por ejemplo, grupos acetilo o metilo) para activar o silenciar genes— o los ARNnp, ARNnop y proteínas que controlan el procesamiento de los transcritos primarios. Merece la pena subrayar que quizá se haya sobrevalorado el protagonismo tradicionalmente atribuido a proteínas en todas estas actividades, en detrimento del ARN. Parte del ARN intrónico (como los ARNnop), e incluso del ARN exónico ensamblado, podría desempeñar una función reguladora operando a la manera de un código postal: dirigiendo moléculas de ARN hacia dianas receptoras en otras moléculas y reclutando proteínas para transformar las señales en acciones. Quizá, pues, sea finalmente el ARN el que encierre los secretos de la complejidad humana.

## Actividades

11. ¿Por qué en el código de Crick un codón formado por tres nucleótidos iguales (por ejemplo, AAA o CCC) no podría tener sentido?
12. ¿Cuántos codones formados por los nucleósidos G, G y U (en cualquier orden) podrían tener sentido en el código “sin comas” de Crick? Razónalo.
13. Cuando a un sistema libre de células se le agregó un ARNm en el que alternaban restos de A y C (...-A-C-A-C-A- C...) se obtuvo un polipéptido constituido por restos alternantes de treonina (Thr) e histidina (His). ¿Qué conclusión se puede sacar?
14. Un segundo experimento utilizó un ARNm a base de ...-A-A-C-A-A-C-A-A-C-A-A-C..., que se podía leer en tres marcos distintos, lo que originaba tres polipéptidos: uno con solo restos de asparagina (Asn), otro con treonina (Thr) y un tercero formado exclusivamente por glutamina (Gln). ¿Se puede establecer el significado de algún codón?
15. En un cultivo de células eucariotas animales se introduce un inhibidor de la síntesis de ribosomas de células procariontas, ¿podrán las células cultivadas sintetizar proteínas? Razona la respuesta.
16. ¿Qué secuencias de aminoácidos resultarían de la traducción de los siguientes ARNm?:
  - a. A-U-G-C-A-U-A-G-A-A-G-G-C-C-U-A-U-U-G-U-A.
  - b. C-A-U-G-U-U-U-C-U-U-A-A-A-G-G-U-C-G-U-U-C.
17. ¿Qué polipéptido codificará un fragmento de ADN como el de la actividad 9?
18. El ARNm de los organismos del planeta ficticio *Tatooine* incluye nucleótidos de tres tipos, y sus proteínas hasta 81 tipos de aminoácidos distintos. ¿De cuántos nucleótidos como mínimo estará formado cada codón? ¿Sería factible en tales circunstancias un mecanismo de traducción esencialmente igual al terrestre? Arguméntalo.
19. Explica qué ocurriría si una mutación eliminase 3 bases consecutivas del ADN.
20. ¿Por qué el orden de los nucleótidos en el ADN determina los caracteres genéticos y, en consecuencia, el fenotipo de los organismos?
21. ¿Cuánta energía consume una célula para crear un solo enlace peptídico mediante la traducción? ¿En qué etapas?
22. La mayor parte de los antibióticos utilizados en medicina actúan inhibiendo específicamente algún paso de la síntesis de proteínas de las células procariontas (bacterias). Investiga en qué etapas actúan los siguientes antibióticos: tetraciclinas, estreptomycin, eritromicina, rifamicina.
23. Los mutantes *lac I<sup>-</sup>* de *Escherichia coli* tienen un gen *lac I* defectuoso que no se traduce. ¿Cómo reaccionarán estas bacterias ante distintos niveles de lactosa?

24. ¿Qué ocurrirá con los mutantes *Oc* de *Escherichia coli*, a cuyo operador no reconoce el represor? ¿Y si la mutación afectase al promotor entero, inutilizándolo?
25. Jacob y Monod trabajaron con bacterias portadoras de dos copias del operón *lac*, una en el cromosoma bacteriano y otra en un plásmido. Describe qué ocurriría:
- Si una de las copias poseyese la mutación *lac I*<sup>-</sup> y la otra fuese “normal”.
  - Si una de las copias poseyese la mutación *Oc* y la otra fuese “normal”.



### Recuerda

- El **código genético** consta de tripletes de bases llamados **codones**. Cada codón cifra un solo aminoácido, pero cada aminoácido puede ser cifrado por varios codones (**degeneración**). Carece de **superposición** y de **comas** (los codones se leen sucesivamente, sin solaparse), pero posee señales de **inicio** y de **terminación**.
- La **traducción** de la información genética ocurre en los **ribosomas**, que se desplazan por el ARNm en dirección 5' → 3' leyendo cada codón sucesivo y reclutando el correspondiente aminoacil-ARNt. El proceso requiere energía (GTP), factores de iniciación, elongación y liberación, y la actividad **transferasa del peptidilo** del ARNr.
- Los genes de los procariotas son **policistrónicos**: varios de ellos funcionalmente relacionados se agrupan en un **operón**, transcribiéndose conjuntamente. La expresión de muchos operones se inhibe por acción de **represores** cifrados por **genes reguladores**; los represores se activan o se inactivan al unirse a ciertas moléculas.

# Solucionario

1. Sí: el hecho de que cada base púrica se aparee con una pirimidínica explica que la cantidad de ambas bases sea la misma ( $Pur = Pyr$ ); como cada adenina se une a una timina, la cantidad de A será igual a la de T y, análogamente, la cantidad de G será igual a la de C.
2. Si A = 29,3 %, entonces T = 29,3 %. Como A + T = 58,6 %, la diferencia  $100 - 58,6 = 41,4$  % corresponderá a G + C; como ambas deben ir en la misma proporción, G = C = 20,7 %.
3. Que está formado por una sola hebra, ya que en todo ADN de doble hélice  $Pur/Pyr = 1$ .
4. La del erizo de mar, ya que es la que tiene la menor proporción de pares G-C que, recordémoslo, se estabilizan mediante tres enlaces de hidrógeno; por tanto, a mayor número de enlaces G-C, más enlaces de hidrógeno habrá que romper para desnaturalizar el ADN.
5. Las dos hebras de un ADN circular no se pueden separar por acción del calor, aunque se rompan todos los enlaces de hidrógeno; para separarlas sería necesario que una de las hebras estuviese cortada, de manera que se pudiese “desenrollar” de la hebra intacta.
6. En la mayoría de los organismos pluricelulares, todo el ADN mitocondrial de un individuo lo hereda de su madre. Esto puede deberse a varias razones: a la simple diferencia de número (un óvulo contiene de 100 000 a 1 000 000 moléculas de ADN mitocondrial, mientras que un espermatozoide contiene solo entre 100 y 1000), a la degradación del ADN mitocondrial del espermatozoide en el óvulo fecundado y, en unas pocas especies, a la imposibilidad de que el ADN mitocondrial del espermatozoide entre en el óvulo.

En consecuencia, los hijos estarán afectados si es Carmen la que padece NOHL, pero no si la sufre Luís.

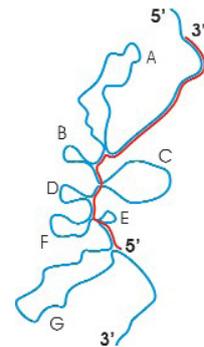
7. **Combinaciones de los 4 nucleósidos tomados de 2 en 2 (importando el orden):** GG, GC, GA, GT, CG, CC, CA, CT, AG, AC, AA, AT, TG, TC, TA, TT. En total,  $4 \times 2 = 16$  combinaciones.

**Combinaciones de los 4 nucleósidos tomados de 2 en 2 (sin importar el orden):** de la lista anterior se tachan las equivalentes y quedan 10 (GG, GC, GA, GT, CC, CA, CT, AA, AT, TT).

**Combinaciones de los 4 nucleósidos tomados de 3 en 3 (importando el orden):** GGG, GGC, GGA, GGT, GCG, GCC, GCA, GCT, GAG, GAC, GAA, GAT, GTG, GTC, GTA, GTT, CGG, CGC, CGA, CGT, CCG, CCC, CCA, CCT, CAG, CAC, CAA, CAT, CTG, CTC, CTA, CTT, AGG, AGC, AGA, AGT, ACG, ACC, ACA, ACT, AAG, AAC, AAA, AAT, ATG, ATC, ATA, ATT, TGG, TGC, TGA, TGT, TCG, TCC, TCA, TCT, TAG, TAC, TAA, TAT, TTG, TTC, TTA y TTT; en total,  $4^3 = 64$ .

**Combinaciones de los 4 nucleósidos tomados de 3 en 3 (sin importar el orden):** se tachan los que son equivalentes de la lista anterior y quedan en total 20 combinaciones (GGG, GGC, GGA, GGT, GCC, GCA, GCT, GAA, GAT, GTT, CCC, CCA, CCT, CAA, CAT, CTT, AAA, AAT, ATT y TTT).

- 8.** No. Dado que se necesitan combinaciones de tres nucleósidos (tripletes) para determinar los 21 aminoácidos, y que hay 64 posibles tripletes, es perfectamente posible que cada triplete se traduzca por un solo aminoácido. Pero en un flujo inverso cada aminoácido podría “traducirse” por varios tripletes; es decir, una proteína podría “leerse” de muchas maneras distintas. Puesto que las proteínas no pueden traducirse a ADN (que es lo que se hereda), las modificaciones en el organismo (por tanto, en sus proteínas) no pueden pasar a su descendencia.
- 9.** Secuencia complementaria: 5'...dA-dT-dG-dG-dT-dG-dC-dA-dC-dC-dT-dG-dA-dC-dT...3'. Secuencia del ARNm: 5'...A-U-G-G-U-G-C-A-C-C-U-G-A-C-U...3'.
- 10.** Los extremos 3' y 5' se señalan en la figura. El esquema muestra que hay zonas del ADN que se hibridan con el ARNm y otras que no (los siete bucles señalados con letras), que se corresponderán con intrones, transcritos pero posteriormente eliminados del ARNm.
- 11.** Supongamos que el codón AAA difra el aminoácido Lys. Como la secuencia de aminoácidos Lys-Lys es perfectamente posible, también lo será la secuencia de nucleótidos AAAAAA. Pero entonces sería factible un cambio del marco de lectura, ya que la secuencia que va desde el segundo al cuarto nucleótido, o desde el tercero al quinto, incorporaría al aminoácido Lys.
- 12.** Si GGU tiene sentido y se corresponde, por ejemplo, con el aminoácido Gly, sería posible la secuencia GGUGGU, correspondiente al dipéptido Gly-Gly. Si GUG y UGG tuvieran sentido se podría producir un cambio de marco de lectura. Por tanto, solo un codón podría tener sentido.
- 13.** Los únicos codones posibles son CAC y ACA. Uno de ellos corresponderá a Thr y el otro a His, pero como no sabemos si la lectura ha comenzado por CAC o por ACA —en principio, cualquier marco de lectura es posible—, no podemos concretar más las asignaciones.
- 14.** Hay tres codones posibles: AAC, ACA y CAA. Uno de ellos codificará Asn, otro Thr y el tercero Gln. Puesto que el codón ACA es común a este experimento y al de la actividad 13, debe codificar Thr (el aminoácido que aparecía en ambos experimentos). En consecuencia, CAC codificará His. AAC codificará, bien Asn, bien Gln, y CAA lo contrario.
- 15.** Los ribosomas de las células procariotas y las de las células eucariotas presentan una estructura y composición química diferentes; por tanto, aunque se añada un inhibidor de los ribosomas procariotas, los ribosomas eucariotas no se verán afectados ni tampoco, en principio, la síntesis proteínica. Sin embargo, las mitocondrias de las células eucariotas poseen ribosomas típicamente bacterianos, por lo que si el inhibidor pudiera introducirse en las mitocondrias, la síntesis de proteínas que tiene lugar en los citados orgánulos sí se vería afectada.
- 16.** a. Met – His – Arg – Arg – Pro – Ile – Val.  
b. His – Val – Ser.
- 17.** A partir del ARN formado por transcripción (solución de la actividad 9), y traduciéndolo en la dirección 5' → 3', se obtiene el polipéptido H<sub>2</sub>N-Met-Val-His-Leu-Thr-COOH.
- 18.** Con codones de 4 nucleótidos se podrán codificar hasta  $3^4 = 81$  aminoácidos, que son justamente los que hay en las proteínas tatooinianas. En tales circunstancias no habría degeneración del código genético ni codones de terminación, por lo que la última etapa de la traducción sería distinta a la terrestre. Quizá algún nucleótido se modifica químicamente después de la transcripción y actúa como señal de terminación, o puede que haya un codón sin sentido que codifica un aminoácido en ciertas condiciones, como ocurre con la selenocisteína.



**19.** Pueden darse dos posibilidades:

- Si las tres bases perteneciesen a un mismo codón, la proteína resultante tendría un aminoácido menos, sin que el resto de la secuencia variase.
- Si las tres bases perteneciesen a dos codones consecutivos, desaparecerían de la proteína los dos aminoácidos cifrados por ambos codones y, a cambio aparecería un nuevo aminoácido (el codificado por las tres bases no eliminadas de dichos codones). Por ejemplo, la secuencia A-U-G-C-A-U-A-**G-A-G**-G-U-C-C-U-A-U-U-G-U-A se traduce por el polipéptido Met – His – Arg – Gly – Pro – Ile – Val; si se eliminan los tres nucleósidos resaltados, el polipéptido resultante sería Met – His – Ser – Pro – Ile – Val.

Normalmente, tampoco en este caso se vería afectado el resto de la secuencia. Ahora bien, si el nuevo codón que aparece a cambio de los otros dos codones antiguos resultase ser un codón de terminación, de la cadena polipeptídica resultante desaparecerían todos los aminoácidos siguientes, por lo que podría quedar muy recortada.

**20.** Porque la secuencia de nucleótidos en la molécula de ADN es la que va a determinar la secuencia de aminoácidos en las proteínas y, en consecuencia, el fenotipo del organismo.**21.** Deben hidrolizarse cuatro enlaces fosfato de alta energía (dos ATP y dos GTP):

- Los dos primeros enlaces se hidrolizan durante la unión del aminoácido a su ARNt: la energía requerida para esta etapa procede de la hidrólisis de un ATP en AMP y P<sub>PPi</sub>; tras la hidrólisis del P<sub>PPi</sub> formado en dos P<sub>i</sub> y la reacción de AMP con otro ATP para dar dos ADP, todo ocurre como si se hubieran hidrolizado dos moléculas de ATP en ADP y P<sub>i</sub>.
- Un tercer enlace fosfato de alta energía (un GTP) se hidroliza cuando el aminoacil-ARNt se une al sitio A del ribosoma.
- El cuarto enlace (otro GTP) se hidroliza durante la translocación, es decir, cuando el ribosoma se desplaza al siguiente codón.

**22.** Las tetraciclinas inhiben la unión del aminoacil ARNt a la sede A del ribosoma; la estreptomina y la eritromicina inhiben la translocación y la rifamicina inhibe la transcripción (síntesis de ARNm).**23.** Al no producirse el represor, los genes lac nunca estarán inhibidos y siempre se expresarán, haya o no lactosa en el medio. Únicamente la presencia de glucosa reducirá en buena medida su expresión, ya que entonces habrá poco AMPc y la CAP no se unirá al ADN.**24.** Si el operador no se une al represor siempre se expresarán los genes. Si el promotor no “funciona”, la polimerasa del ARN no podrá comenzar la transcripción y nunca se expresarán.**25. a)** La copia normal produce represor, que se difunde por la célula y se une a ambos operadores lac, por lo que, en ausencia de lactosa, no se expresará ninguna de los dos copias.

- b)** Al haber un operador no funcional sus genes estructurales siempre se expresarán, aunque los de la copia normal no lo hagan en ausencia de lactosa.

# Glosario

**Colonia**

Grupo de microorganismos —generalmente bacterias o levaduras— procedentes de una única célula madre mediante mitosis sucesivas, que al crecer en un medio sólido forman masas con una forma, color, consistencia y brillo característicos de cada especie.

**Difracción de rayos X**

Técnica consistente en hacer pasar un haz de rayos X a través de un cristal de la sustancia sujeta a estudio; el haz se escinde en varias direcciones debido a la simetría de la agrupación de átomos y da lugar a un patrón de intensidades que revela la disposición relativa de los átomos en el cristal.

**Virulento**

Microorganismo que tiene carácter patogénico o nocivo; en otras palabras, que es capaz de causar una enfermedad.

# Bibliografía

ASIMOV, I.: Fotosíntesis. Barcelona, Plaza & Janés, 1992.

Uno de los más conocidos divulgadores científicos, y también afamado escritor de ciencia-ficción, nos muestra con estilo claro y ameno el proceso del que depende la vida. Aunque se trata de un libro antiguo (fue escrito en 1968), su principal atractivo es que, partiendo de preguntas casi triviales (¿por qué no se agotan la comida ni el oxígeno?), logra introducirnos en la comprensión de los esfuerzos de tantos científicos por desentrañar el mecanismo de la fotosíntesis.

CAIRNS-SMITH, A. G.: Siete pistas sobre el origen de la vida. Madrid, Alianza, 1990.

El autor, emulando a Sherlock Holmes, va buscando “pistas” entre los seres vivos actuales para intentar averiguar su origen, lo que le permite explorar la estructura y funcionamiento de las células desde una perspectiva sorprendente, prescindiendo de tecnicismos.

DE DUVE, C.: La célula viva (2 tomos). Barcelona, Prensa Científica, 1988.

En este libro, su autor, premio Nobel de Medicina, nos introduce en un maravilloso viaje por el interior de una célula eucariótica viva, reduciéndonos con la imaginación al tamaño de bacterias y permitiéndonos nadar a nuestro gusto por su interior. Combina magistralmente la amenidad y el rigor científico, y constituye la mejor forma de adentrarse en los contenidos de la asignatura.

MAYNARD SMITH, J., Y SZATHMÁRY, E.: Ocho hitos de la evolución. Barcelona, Tusquets, 2001.

Obra que recorre de forma panorámica la evolución de los seres vivos, desde el origen de la vida hasta la aparición del lenguaje, jalonándola de una serie de “transiciones principales” (la aparición de las células, el surgimiento del sexo, la emergencia de la pluricelularidad...). Dirigido a un público no especializado, hace hincapié en los principales problemas que deben resolver los biólogos y trata muchos de los aspectos de la asignatura desde una perspectiva evolutiva.

SCHRÖDINGER, E.: ¿Qué es la vida? Barcelona, Tusquets, 1983.

Es uno de los textos más influyentes en la historia de la Biología, escrito en 1944 por uno de los físicos más prestigiosos. Schrödinger, presentándose a sí mismo como un “físico ingenuo”, intenta dilucidar —con prosa clara y argumentos persuasivos— los problemas de la herencia y la organización celular; parte únicamente de consideraciones físicas y predice la estructura de los genes antes del descubrimiento de la doble hélice.

SOL, C. Y OTROS: Selectividad Biología: pruebas de 2006. Madrid, Anaya, 2007.

Es un libro bastante económico en el que se plantean y se resuelven las cuestiones formuladas en pruebas de acceso a la Universidad de toda España.

TEIXIDÓ, F.: Biología. Schaum. Madrid, McGraw-Hill/Interamericana, 2005.

Adaptado al currículo vigente de segundo curso de bachillerato, en cada uno de sus capítulos se resumen de forma concisa los principales conceptos de Biología, se aportan instrucciones y consejos para no cometer errores en los exámenes y se proponen y resuelven multitud de ejercicios y problemas. Útil para preparar las pruebas de acceso a la Universidad.

VOGEL, G. Y ANGERMANN, H.: Atlas de biología. Barcelona, Omega, 1987.

Se trata de un libro que conserva plena vigencia en la presentación de los contenidos básicos de la Biología de forma esquemática y asociada siempre a ilustraciones claras y detalladas.

WATSON, J.: La doble hélice. Barcelona, Salvat, 1987.

Best-seller internacional desde su publicación, en 1968, narra de forma autobiográfica los acontecimientos que desembocaron en el descubrimiento de la estructura del ADN. Constituye una interesante descripción del modo en que trabajan los científicos, de sus anhelos y sus mezquindades; en suma, un relato de la naturaleza del éxito.

## Aviso legal

---

El contenido de esta unidad es adaptación del existente en el libro de Biología para 2º de Bachillerato a distancia (NIP0: 660-09-096-2).

Adaptación: César Martínez Martínez  
Asesor Técnico Docente Biología y Geología. CIDEAD, 2016.

La utilización de recursos de terceros se ha realizado respetando las licencias de distribución que son de aplicación, acogiéndonos igualmente a los artículos 32.3 y 32.4 de la Ley 21/2014 por la que se modifica el Texto Refundido de la Ley de Propiedad Intelectual. Si en algún momento existiera en los materiales algún elemento cuya utilización y difusión no estuviera permitida en los términos que aquí se hace, es debido a un error, omisión o cambio de licencia original.

Si el usuario detectara algún elemento en esta situación podrá comunicarlo al CIDEAD para que tal circunstancia sea corregida de manera inmediata.

En estos materiales se facilitan enlaces a páginas externas sobre las que el CIDEAD no tiene control alguno, y respecto de las cuales declinamos toda responsabilidad.



DIRECCIÓN GENERAL DE  
FORMACIÓN PROFESIONAL

