

Biomoléculas: ácidos nucleicos

Los ácidos nucleicos, **ácido desoxirribonucleico (ADN)** y **ácido ribonucleico (ARN)**, son los mensajeros químicos de la información genética de las células. En el ADN de las células está codificada la información que determina la naturaleza de la célula, controla el crecimiento y la división celular y dirige la biosíntesis de las enzimas y de otras proteínas requeridas para las funciones celulares.

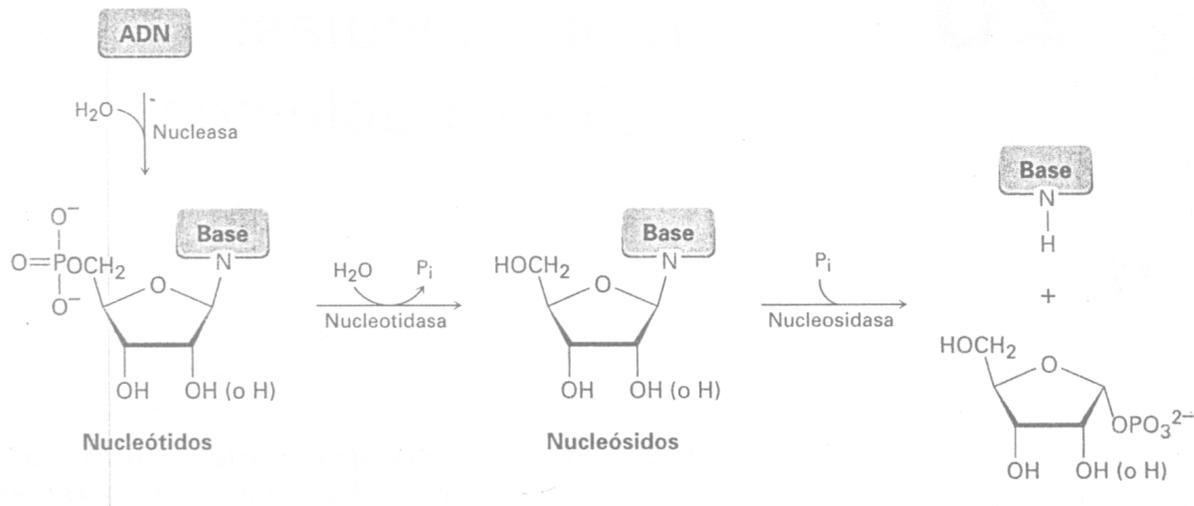
Además de los ácidos nucleicos, los derivados de ácidos nucleicos como el ATP están involucrados como agentes de fosforilación en muchas rutas bioquímicas, y varias coenzimas importantes, que incluyendo la NAD^+ , la FAD y la coenzima A, tienen componentes de ácidos nucleicos.

¿POR QUÉ ESTE CAPÍTULO?

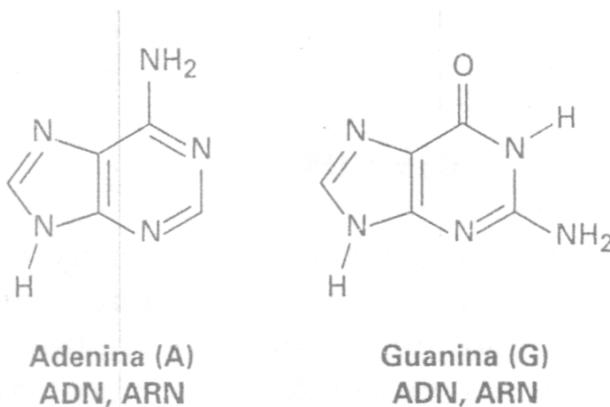
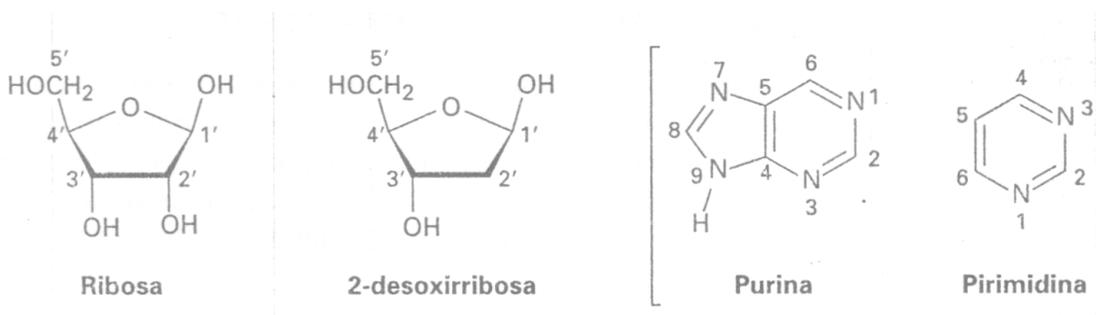
Los ácidos nucleicos son la última de las cuatro clases principales de biomoléculas que consideraremos. Se ha escrito y hablado demasiado acerca del ADN en los medios, que es probable que conozca las bases de la replicación y la transcripción del ADN. Por tanto, veremos brevemente los fundamentos y nos enfocaremos con más atención en los detalles químicos de la secuenciación y la síntesis del ADN.

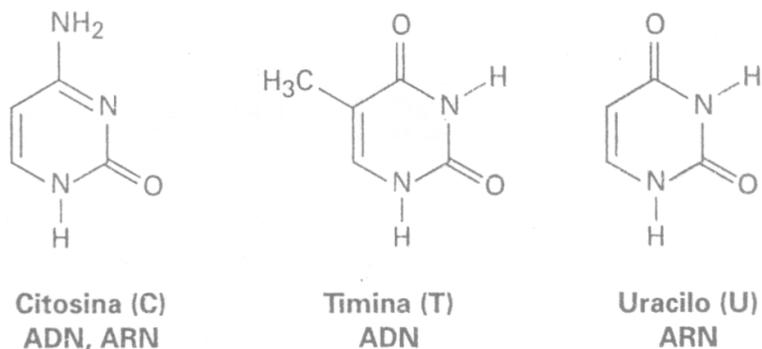
28.1 Nucleótidos y ácidos nucleicos

Al igual que las proteínas son biopolímeros conformados de aminoácidos, los ácidos nucleicos son biopolímeros conformados de nucleótidos unidos entre sí formando una cadena larga. Cada nucleótido está compuesto de un nucleósido unido a un grupo fosfato, y cada nucleósido está compuesto de un azúcar aldopentosa unida a través de su carbono anomérico al átomo de nitrógeno de una base heterocíclica de purina o de pirimidina.



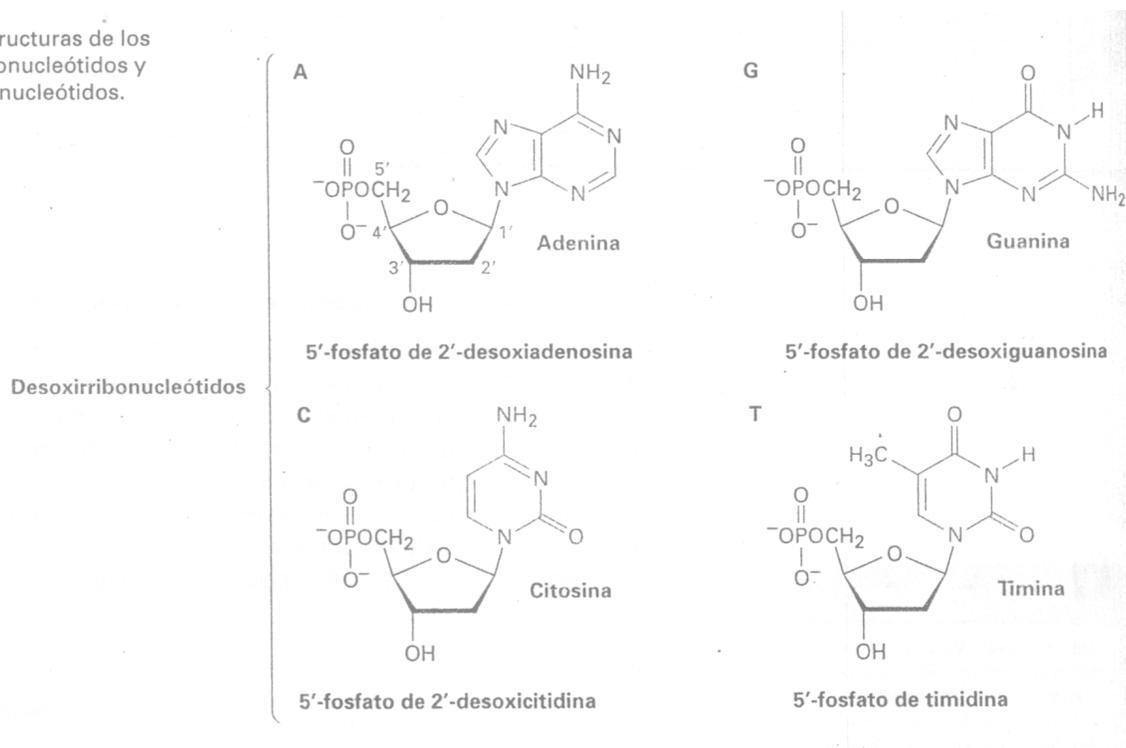
El componente azúcar en el ARN es la ribosa, y el azúcar en el ADN es la 2'-desoxirribosa. (El prefijo 2'-desoxi indica que falta el oxígeno de la posición 2' de la ribosa.) El ADN contiene cuatro amino bases distintas, dos purinas sustituidas (adenina y guanina) y dos pirimidinas sustituidas (citosina y timina). La adenina, la guanina y la citosina también se encuentran en el ARN, pero en el ARN se reemplaza la timina por una base pirimidina más estrechamente relacionada llamada uracilo.



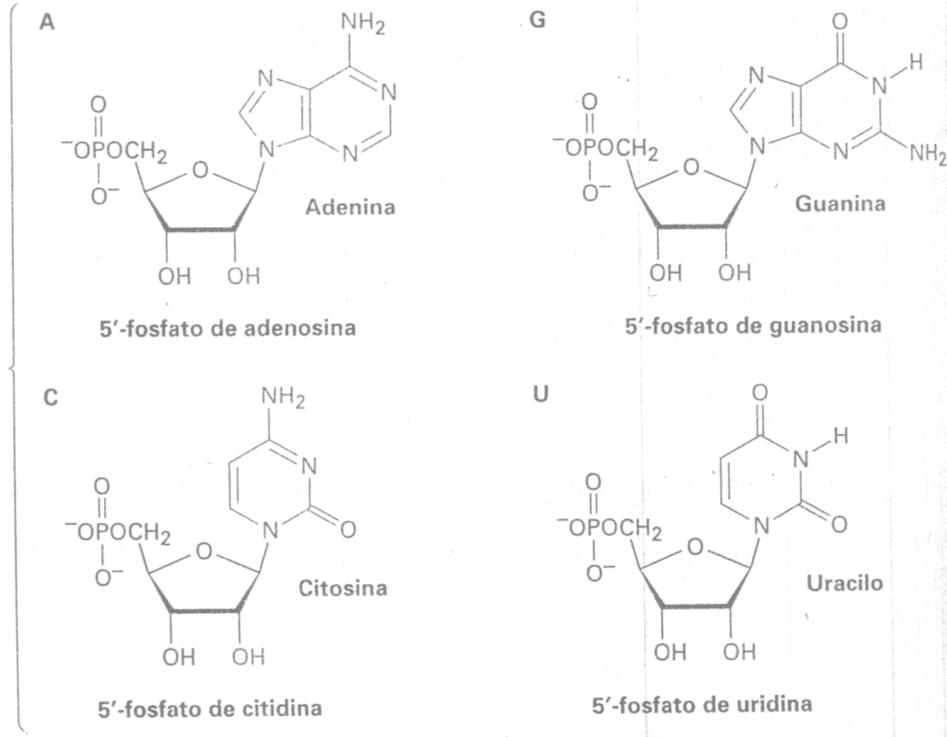


En la figura 28.1 se muestran las estructuras de los cuatro desoxirribonucleótidos y de los cuatro ribonucleótidos. Nótese que al nombrar y numerar los nucleótidos, a las posiciones en los azúcares se les da un superíndice prima para distinguirlas de las posiciones en la amino base. Por ejemplo, la posición 3 estaría en la base, mientras que la posición 3' estaría en el azúcar. Aunque son químicamente similares, el ADN y el ARN difieren de manera notable en su tamaño. Las moléculas de ADN son enormes, con masas moleculares de hasta varios miles de millones. Por el contrario, las moléculas de ARN son mucho más pequeñas, pues sólo contienen 60 nucleótidos y tienen masas moleculares tan bajas como 22 000.

Figura 28.1 Estructuras de los cuatro desoxirribonucleótidos y de los cuatro ribonucleótidos.



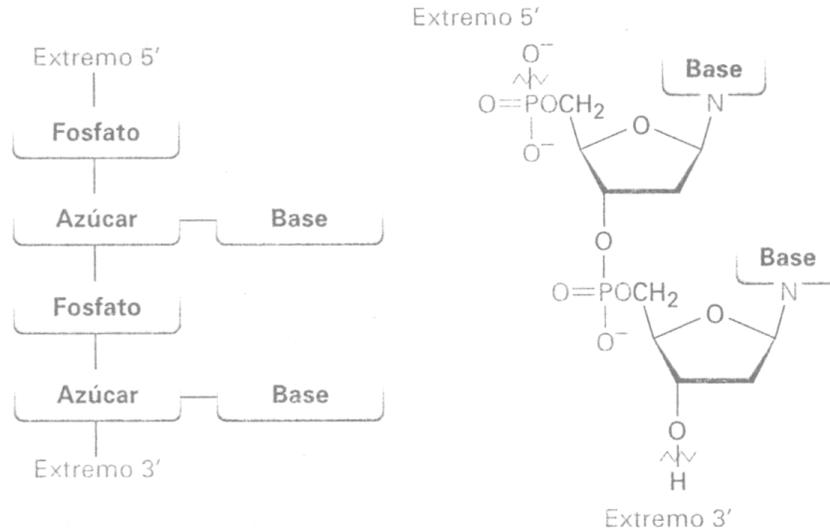
Ribonucleótidos



James Dewey Watson

James Dewey Watson (1928-) nació en Chicago, Illinois, y se matriculó en la Universidad de Chicago a los 15 años de edad. Recibió su doctorado en 1950 en la Universidad de Indiana y trabajó en la Universidad de Cambridge en Inglaterra de 1951 a 1953, donde él y Francis Crick dedujeron la estructura del ADN. Después de más de 20 años como profesor en la Universidad de Harvard, se cambió en 1976 al Laboratorio de Biología Cuantitativa en Cold Spring Harbor, Long Island, Nueva York. Compartió en 1962 el Premio Nobel de Medicina por su trabajo sobre los ácidos nucleicos.

Los nucleótidos están unidos entre sí en el ADN y en el ARN por enlaces fosfodiéster $[RO - (PO_2^-) - O]_{121}$ entre el fosfato, el grupo 5' hidroxilo en el nucleósido y el grupo 3'-hidroxilo en el otro nucleósido. Un extremo del polímero de ácido nucleico tiene un hidroxilo libre en C3' (el extremo 3') y el otro extremo tiene un fosfato en C5' (el extremo 5'). La secuencia de nucleótidos en una cadena se describe comenzando en el extremo 5' e identificando las bases en orden de ocurrencia, utilizando las abreviaciones G, C, A, T (o U para el ARN). Por lo tanto, una secuencia de ADN típica podría escribirse como TAGGCT.



Problema 28.1

Dibuje la estructura completa del dinucleótido AG del ADN.

Problema 28.2

Dibuje la estructura completa del dinucleótido UA del ARN.

28.2 Apareamiento de bases en el ADN: el modelo de Watson-Crick

Las muestras de ADN aisladas de tejidos diferentes de la misma especie tienen las mismas proporciones de bases heterocíclicas, pero con frecuencia las muestras de especies diferentes tienen proporciones de bases bastante distintas; por ejemplo, el ADN humano contiene casi 30% de adenina y de timina y casi 20% de guanina y de citosina. Sin embargo, la bacteria *Clostridium perfringens* contiene casi 37% de adenina y de timina y sólo 13% de guanina y de citosina. Nótese que en ambos ejemplos, las bases ocurren en pares. La adenina y la timina están presentes en cantidades iguales, como lo están la citosina y la guanina. ¿Por qué?

Francis Harry Compton Crick

Francis Harry Compton Crick (1916-2004) nació en Northampton, Inglaterra, y comenzó su carrera científica como físico. Seguida de una interrupción en sus estudios ocasionada por la Segunda Guerra Mundial, se cambió a biología y recibió su doctorado en 1954 por la Universidad de Cambridge. Permaneció en esta universidad como profesor. Compartió en 1962 el Premio Nobel de Medicina.

En 1953, James Watson y Francis Crick elaboraron su propuesta clásica para la estructura secundaria del ADN. De acuerdo con el modelo de Watson-Crick, el ADN bajo condiciones fisiológicas consiste en dos cadenas de polinucleótidos, que van en direcciones opuestas y que se enrollan entre sí en una doble hélice parecida a los pasamanos de una escalera de caracol. Las dos cadenas son complementarias en lugar de idénticas y se mantienen unidas por puente de hidrógeno entre pares específicos de bases, A con T y C con G. Esto es, siempre que se encuentra una base A en una cadena, se encuentra una base opuesta T en la otra cadena; cuando se encuentra una base C en una, se encuentra una G en la otra (figura 28.2). Por lo tanto, este apareamiento de bases complementario explica por qué siempre se encuentran en cantidades iguales A y T, al igual que G y C.

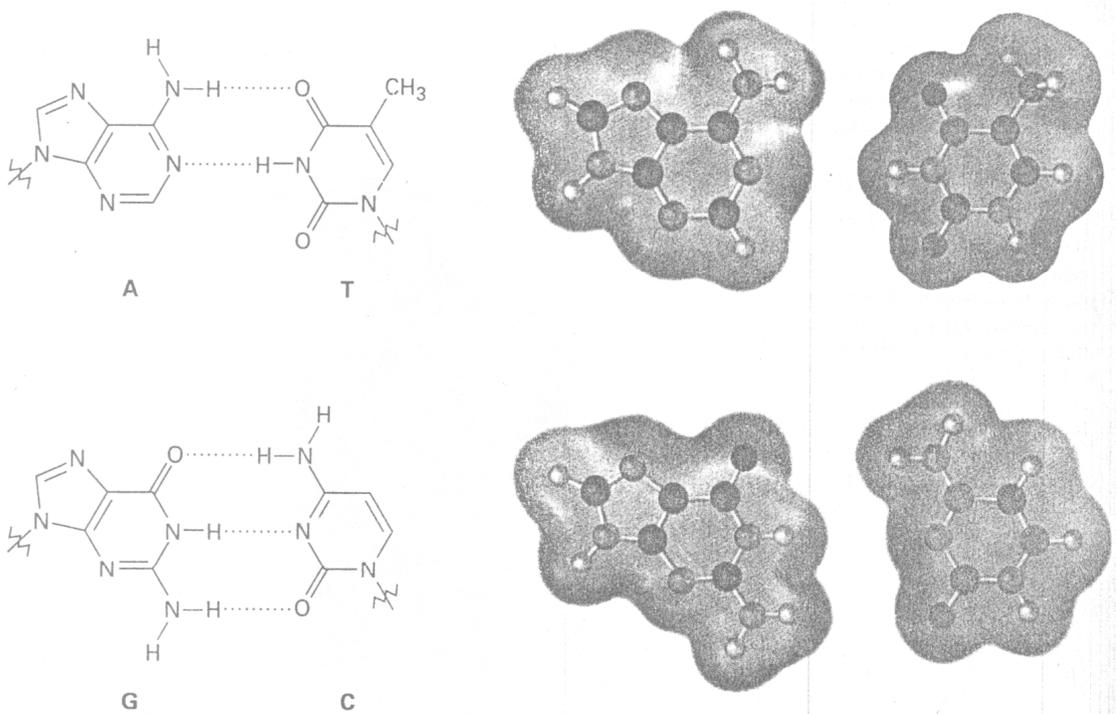


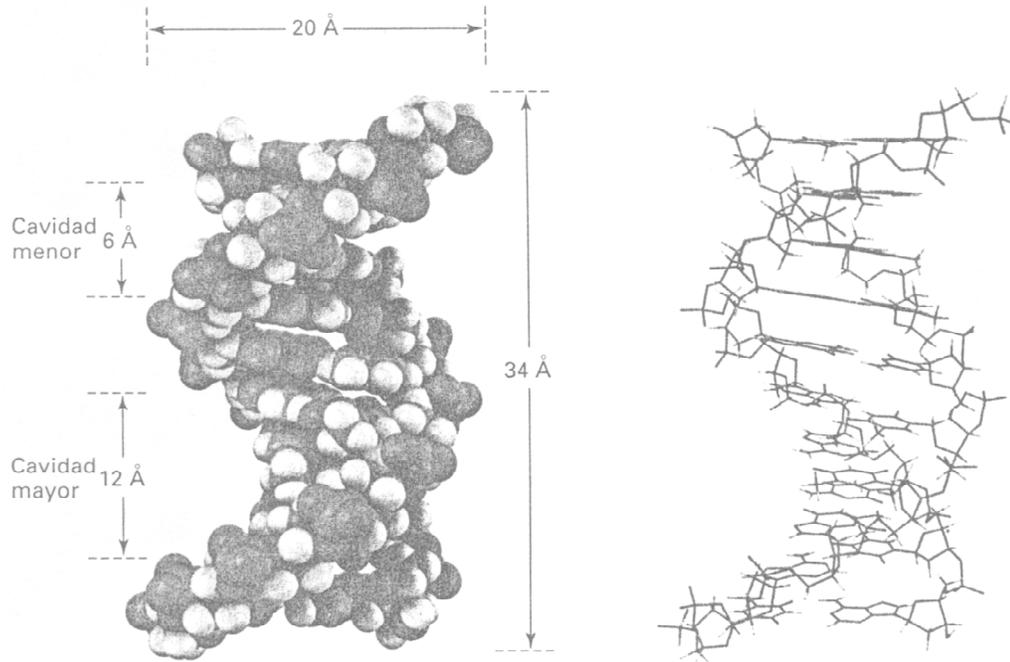
Figura 28.2 Puente de hidrógeno entre los pares de bases en la doble hélice del ADN. Los mapas de potencial electrostático muestran que las caras de las bases son relativamente neutras (verdes), mientras que los bordes tienen regiones positivas (azul) y negativas (rojo). El apareamiento de G con C y de A con T trae consigo regiones opuestamente cargadas.

En la figura 28.3 se muestra una vuelta completa de la doble hélice del ADN. La hélice es de 20 Å de ancho, hay 10 pares de bases por vuelta y cada vuelta es de 34 Å de longitud. En la figura 28.3 se puede ver que las dos cadenas de la doble hélice se enrollan de tal manera que resultan dos tipos de "cavidades", una *cavidad mayor* de 12 Å de ancho y una *cavidad menor* de 6 Å de ancho. La cavidad mayor es ligeramente más profunda que la cavidad menor, y ambas son alineadas por donadores y receptores de puente de hidrógeno. Como resultado, una variedad de moléculas aromáticas policíclicas planas, son capaces de deslizarse lateralmente, o *intercalarse*, entre las bases apiladas. Varios agentes cancerígenos y preventivos del cáncer funcionan interactuando con el ADN de esta manera.

La información genética de un organismo se almacena como una secuencia de desoxirribonucleótidos concatenados entre sí en la cadena de ADN. Para conservar esta información y pasarla a las generaciones futuras, debe existir un mecanismo de copiado del ADN. Para usar la información, debe existir un mecanismo que descodifique el mensaje del ADN e implemente las instrucciones que contiene.

Lo que Crick llamó el "dogma central de la genética molecular" dice que la función del ADN es almacenar información y pasarla al ARN. A su vez, la función del ARN es leer, descodificar y utilizar la información recibida del ADN para preparar proteínas, por lo tanto, tienen lugar tres procesos fundamentales.

Figura 28.3 Representación de una vuelta de la doble hélice en el ADN en modelos compactos y de marco de alambre. El esqueleto fosfato-azúcar va a lo largo del exterior de la hélice, y el puente de hidrógeno de las bases amino va de una a otra en el interior. Son visibles las cavidades mayor y menor.



Replicación

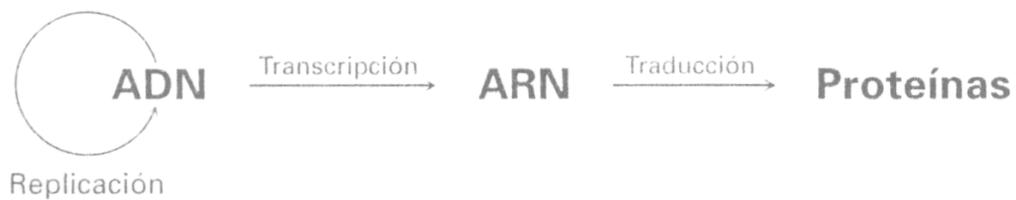
—proceso en el que se hacen copias idénticas del ADN de tal manera que la información puede conservarse y pasarse a la progenie

Transcripción

—el proceso por el que se leen y transportan los mensajes genéticos de los núcleos celulares a los ribosomas, donde ocurre la síntesis de proteínas

Traducción

—el proceso por el que se descodifican y utilizan los mensajes genéticos para sintetizar proteínas



EJEMPLO RESUELTO 28.1**Predicción de la secuencia de bases complementaria en el ADN de doble cadena**

¿Qué secuencia de bases en una cadena de ADN es complementaria a la secuencia TATGCAT en la otra cadena?

Estrategia Recuerde que A y G forman pares complementarios con T y C, respectivamente, y que van a través de la secuencia de reemplazamiento A por T, G por C, T por A y C por G. También recuerde que el extremo 5' está a la izquierda de la cadena original y que el extremo 3' está a la derecha.

Solución Original (5') TATGCAT (3')

Complemento (3') ATACGTA (5') o (5') ATGCATA (3')

Problema 28.3

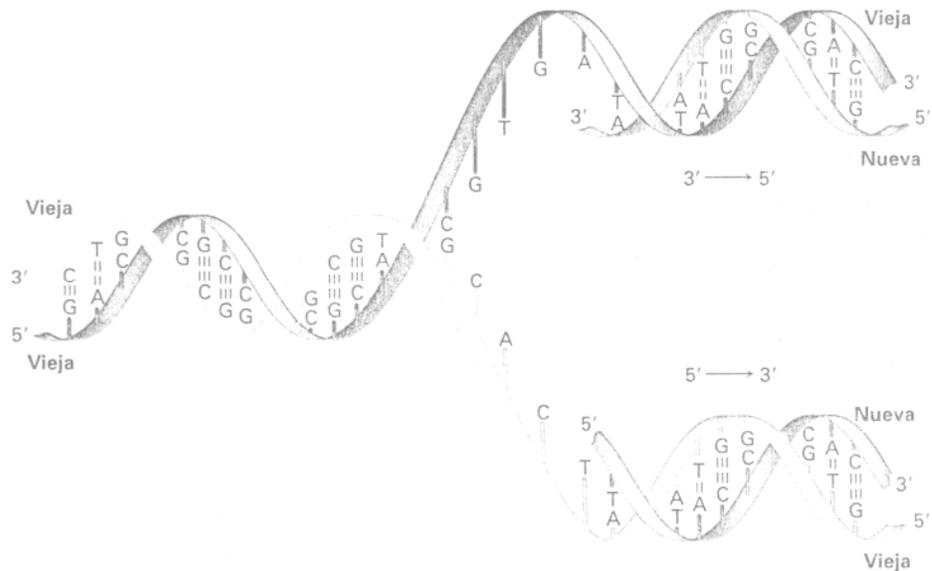
¿Qué secuencia de bases en una cadena de ADN es complementaria a la secuencia siguiente en la otra cadena?

(5') GGCTAATCCGT (3')

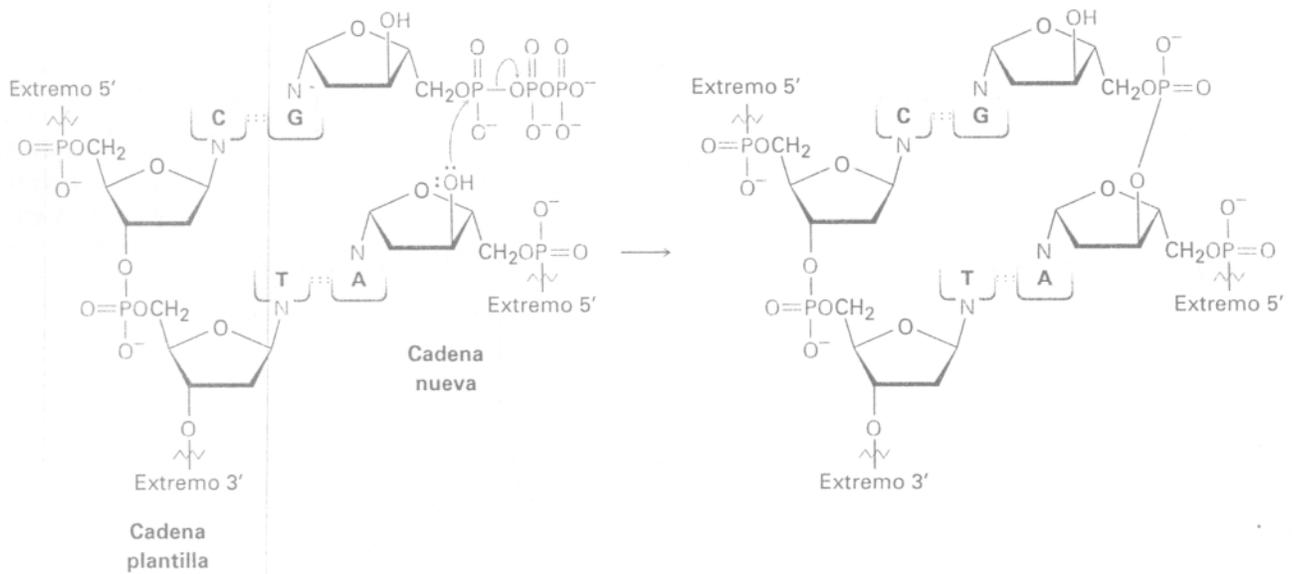
28.3 Replicación de ADN

La replicación del ADN es un proceso catalizado por enzimas que inicia con un desenrollamiento parcial de la doble hélice y el rompimiento de los puentes de hidrógeno entre las cadenas, causados por las enzimas llamadas *helicadas*. A medida que se separan las cadenas y se exponen las bases, se alinean nuevos nucleótidos en cada cadena de una manera complementaria, A a T y G a C, y empiezan a crecer dos cadenas nuevas. Cada cadena nueva es complementaria a su cadena plantilla antigua y se producen dos hélices de ADN idénticas (figura 28.4). Debido a que cada una de las moléculas de ADN nuevas contiene una cadena vieja y una cadena nueva, el proceso se describe como *replicación semiconservativa*.

Figura 28.4 Representación de la replicación semiconservativa de ADN. El ADN de la cadena doble original se desenrolla parcialmente, las bases quedan expuestas, los nucleótidos se alinean en cada cadena de una manera complementaria y empiezan a crecer dos cadenas nuevas. Ambas cadenas se sintetizan en la misma dirección 5' → 3', una de manera continua y una en fragmentos.



La adición de nucleótidos a la cadena en crecimiento sucede en la dirección 5' → 3' y es catalizada por la ADN polimerasa. El baso clave es la adición de un 5'-trifosfato de nucleósido al grupo 3'-hidroxilo libre de la cadena en crecimiento con la pérdida de un grupo saliente difosfato.



Debido a que las cadenas de ADN nuevas se sintetizan en la dirección 5' → 3', no pueden construirse exactamente de la misma manera. Una cadena nueva debe tener su extremo 3' cercano a un punto de desenredo (la *horquilla de replicación*), mientras que la otra cadena nueva tiene su extremo 5' cercano a la horquilla de replicación. Lo que sucede es que el complemento de la cadena original 5' → 3' se sintetiza de manera continua en una sola pieza para dar una copia recién sintetizada llamada *cadena conductora*, mientras que el complemento de la cadena original 5' → 3' se sintetiza de manera discontinua en piezas pequeñas llamadas *fragmentos de Okazaki* que se unen subsecuentemente a las ligasas del ADN para formar la *cadena retardada*.

La magnitud del proceso de replicación es asombrosa. El núcleo de todas las células humanas contiene 46 cromosomas (23 pares), cada uno de los cuales consiste en una molécula de ADN muy grande. A su vez, cada cromosoma está constituido de cientos de segmentos de ADN llamados *genes*, y se estima que la suma de todos los genes en una célula humana (el *genoma* humano) es de 2.9 miles de millones de pares de bases. A pesar del tamaño de estas moléculas enormes, su secuencia de bases se copia fielmente durante la replicación. El proceso de copiado sólo toma minutos, y ocurre un error casi en una de cada 10 a 100 mil millones de bases.

28.4 Transcripción del ADN

Como notó antes, el ARN es similar estructuralmente al ADN pero contiene ribosa en lugar de desoxirribosa y uracilo en lugar de timina. Existen tres tipos principales de ARN, cada uno de los cuales desempeña una función específica. Los tres son moléculas mucho más pequeñas que el ADN, y todas conservan una sola cadena en lugar de cadenas dobles.

- **ARN mensajero (ARNm)** transporta los mensajes genéticos del ADN a los ribosomas, pequeñas partículas granulares en el citoplasma de una célula donde tiene lugar la síntesis de proteínas.
- **ARN ribosómico (ARNr)** es un complejo con proteínas que provee la estructura física de los ribosomas.
- **ARN de transferencia (ARNt)** transporta los aminoácidos a los ribosomas, donde se unen para preparar proteínas.

La conversión de la información en el ADN en proteínas comienza en el núcleo de las células con la síntesis del ARNm por la transcripción del ADN. En las bacterias, el proceso comienza cuando se reconoce la ARN polimerasa y se une a una *secuencia promotora* en el ADN, típicamente consiste de alrededor de 40 pares de bases localizados hacia arriba (5') del sitio inicial de la transcripción. Dentro del promotor están

dos *secuencias consenso* hexaméricas, una localizada a 10 pares de bases hacia arriba del inicio y la segunda localizada a 35 pares de bases hacia arriba.

Después de la formación del complejo polimerasa-promotor, se desenrollan varias vueltas de la doble hélice del ADN, forman una "burbuja" y exponen aproximadamente 14 pares de bases de las dos cadenas. Se alinean los ribonucleótidos correspondientes a través de puentes de hidrógeno a sus bases complementarias en el ADN, la formación de enlaces ocurre en la dirección 5' → 3', la ARN polimerasa se mueve a lo largo de la cadena de ADN, y la molécula de ARN en crecimiento se desprende del ADN (figura 28.5). En cualquier momento, alrededor de 12 pares de bases en el ARN en crecimiento permanecen unidas con puentes de hidrógeno a la plantilla de ADN.

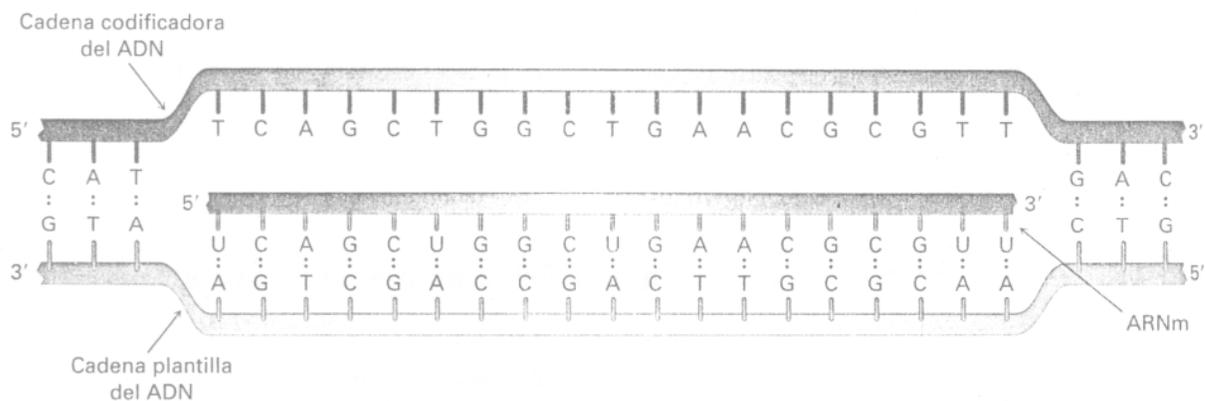


Figura 28.5 Biosíntesis del ARN utilizando un segmento del ADN como una plantilla.

A diferencia de lo que sucede en la replicación del ADN, donde se copian ambas cadenas, sólo una de las dos cadenas de ADN se transcribe en el ARNm. Con frecuencia a la cadena que contiene el gen se le llama cadena codificadora, o *cadena de sentido*, y a la cadena que transcribe se le llama cadena plantilla. Debido a que la cadena plantilla y la plantilla codificadora son complementarias, y a que la cadena plantilla y el ARN transcrito también son complementarios, *la molécula de ARN producida durante la transcripción es una copia de la cadena codificadora del ADN*. La única diferencia es que la molécula de ARN tiene una U donde la cadena codificadora del ADN tiene una T.

Otra parte de la representación en los vertebrados y en las plantas florales es que con frecuencia los genes no son segmentos continuos de la cadena de ADN. En lugar de eso, un gen comenzará en una sección pequeña del ADN llamada *exón*, después es interrumpida por una sección no codificadora llamada *intrón*, y retomará de nuevo hacia abajo de la cadena en otro *exón*. La molécula final de ARNm resulta sólo después de que las secciones no codificadoras se cortan y se juntan las piezas restantes entre sí

por medio de espliceosomas; por ejemplo, el gen para la fosfato triosa isomerasa en el maíz contiene nueve exones que son responsables de aproximadamente 80% de los pares de bases del ADN y ocho intrones que sólo son responsables de 20% de los pares de bases.

Problema 28.4

Muestre cómo el uracilo puede formar puentes de hidrógeno con la adenina.

Problema 28.5

¿Que secuencia de bases del ARN es complementaria a la secuencia de bases del ADN siguiente?

(5') GATTACCGTA (3')

Problema 28.6

¿A partir de qué secuencia de bases del ADN fue transcrita la secuencia de ARN siguiente?

(5') UUCGCAGAGU (3')

28.5 Traducción de ARN: biosíntesis de proteínas

La función celular principal del ARNm es dirigir la biosíntesis de los miles de diversos péptidos y proteínas requeridos por un organismo —quizá 100 000 en un ser humano—. El mecanismo de la biosíntesis de proteínas sucede en los ribosomas, pequeñas partículas granulares en el citoplasma de una célula que consisten de alrededor de 60% de ARN ribosómico y 40% de proteínas.

La secuencia de ribonucleótidos específica en el ARNm forma un mensaje que determina el orden en el que se unen los residuos de aminoácidos. Cada "palabra", o codón, a lo largo de la cadena de ARNm consiste de una secuencia de tres ribonucleótidos que es específica para un aminoácido dado; por ejemplo, la serie UUC en el ARNm es un codón que dirige la incorporación del aminoácido fenilalanina en la proteína en crecimiento. De los $4^3 = 64$ tripletes posibles de las cuatro bases en el ARN, 61 codifican aminoácidos y 3 codifican la terminación de la cadena. La tabla 28.1 muestra el significado de cada codón.

El mensaje insertado en el ARNm es leído por el ARN de transferencia (ARNt) en un proceso llamado traducción. Existen 61 ARNt distintos, uno para cada uno de los 61 codones que determinan un aminoácido. Un ARNt típico es una sola cadena y tiene aproximadamente la forma de una hoja de trébol, como muestra la figura 28.6 en la página 1111. Consiste en alrededor de 70 a 100 ribonucleótidos y está unido a un aminoácido específico por un enlace éster a través del hidroxilo 3' en la ribosa en el extremo 3' del ARNt. También cada ARNt contiene a la mitad de su hoja un segmento llamado anticodón, una secuencia de tres ribonucleótidos complementaria a la secuencia del codón; por ejemplo, la secuencia del codón UUC presente en el ARNm es leída por un ARNt unido a una fenilalanina que tiene la secuencia de bases del anticodón complementario GAA. [Recuerde que la secuencia de nucleótidos se escribe en la dirección 5' → 3', de tal manera que la secuencia en un anticodón debe ser la inversa, esto es, el complemento de (5')-UUC-(3') es (3')-AAG-(5'), el cual se escribe como (5')-GAA-(3').]

A medida que se lee cada codón sucesivo en el ARNm, diferentes ARNt traen los aminoácidos correctos a la posición para la transferencia mediada por enzima al péptido en crecimiento. Cuando se completa la síntesis de la proteína apropiada, un codón de "paro" señala el fin y la proteína es liberada del ribosoma. El proceso se ilustra en la figura 28.7.

Tabla 28.1 | Asignaciones de codones de tripletes de bases

Primera base (extremo 5')	Segunda base	Tercera base (extremo 3')			
		U	C	A	G
U	U	Fen	Fen	Leu	Leu
	C	Ser	Ser	Ser	Ser
	A	Tir	Tir	Paro	Paro
	G	Cis	Cis	Paro	Trp
C	U	Leu	Leu	Leu	Leu
	C	Pro	Pro	Pro	Pro
	A	His	His	Gln	Gln
	G	Arg	Arg	Arg	Arg
A	U	Ile	Ile	Ile	Met
	C	Tre	Tre	Tre	Tre
	A	Asn	Asn	Lis	Lis
	G	Ser	Ser	Arg	Arg
G	U	Val	Val	Val	Val
	C	Ala	Ala	Ala	Ala
	A	Asp	Asp	Glu	Glu
	G	Gli	Gli	Gli	Gli

EJEMPLO RESUELTO 28.2

Predicción de la secuencia de aminoácidos transcrita del ADN

¿Qué secuencia de aminoácidos se codifica por el siguiente segmento de una cadena codificadora de ADN?

(5') CTA-ACT-AGC-GGG-TCG-CCG (3')

Estrategia El ARNm producido durante la traducción es una copia de la cadena codificadora de ADN, con cada T reemplazada por U, por tanto, el ARNm tiene la secuencia

(5') CUA-ACU-AGC-GGG-UCG-CCG (3')

Cada conjunto de tres bases forma un codón, cuyo significado puede encontrarse en la tabla 28.1.

Solución Leu-Tre-Ser-Gli-Ser-Pro.

Figura 28.6 Estructura de una molécula de ARNt. El ARNt es una molécula con forma de hoja de trébol que contiene un triplete anticodón en una "hoja" y una unidad de aminoácido unida covalentemente a su extremo 3'. El ejemplo que se muestra es un ARNt de una levadura que codifica la fenilalanina. Los nucleótidos no identificados específicamente son análogos modificados químicamente de los cuatro nucleótidos comunes.

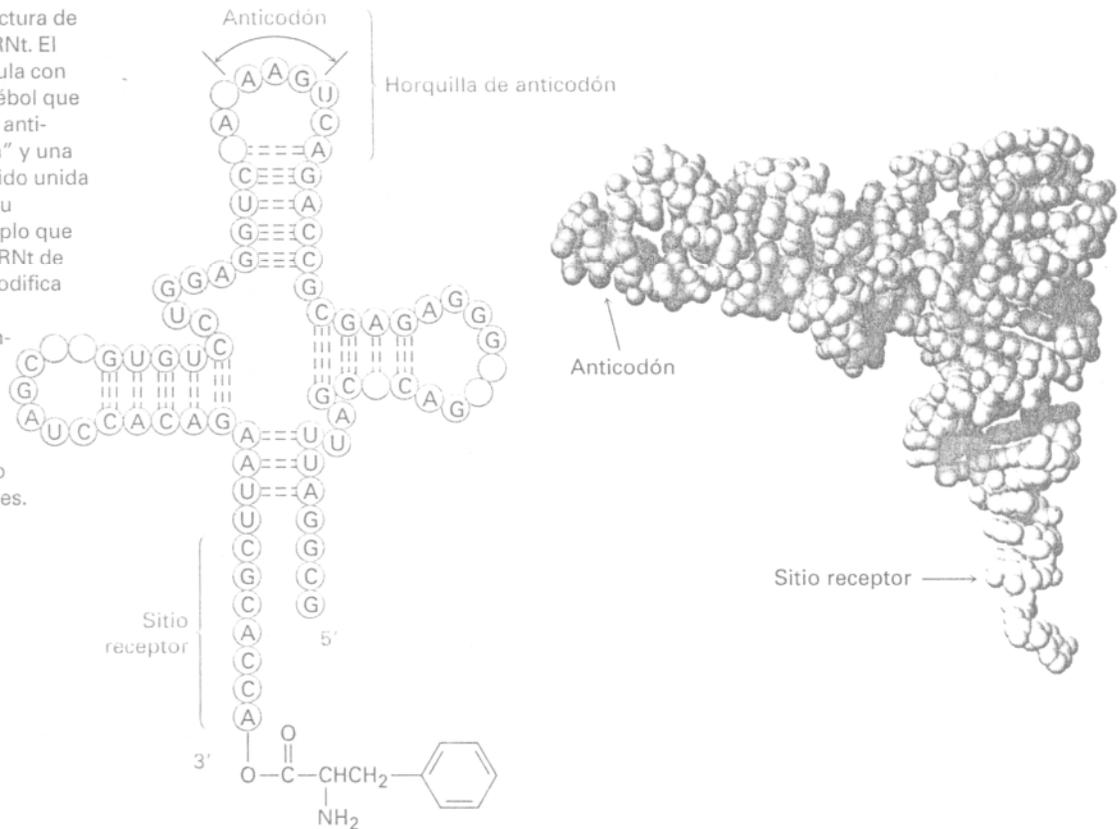
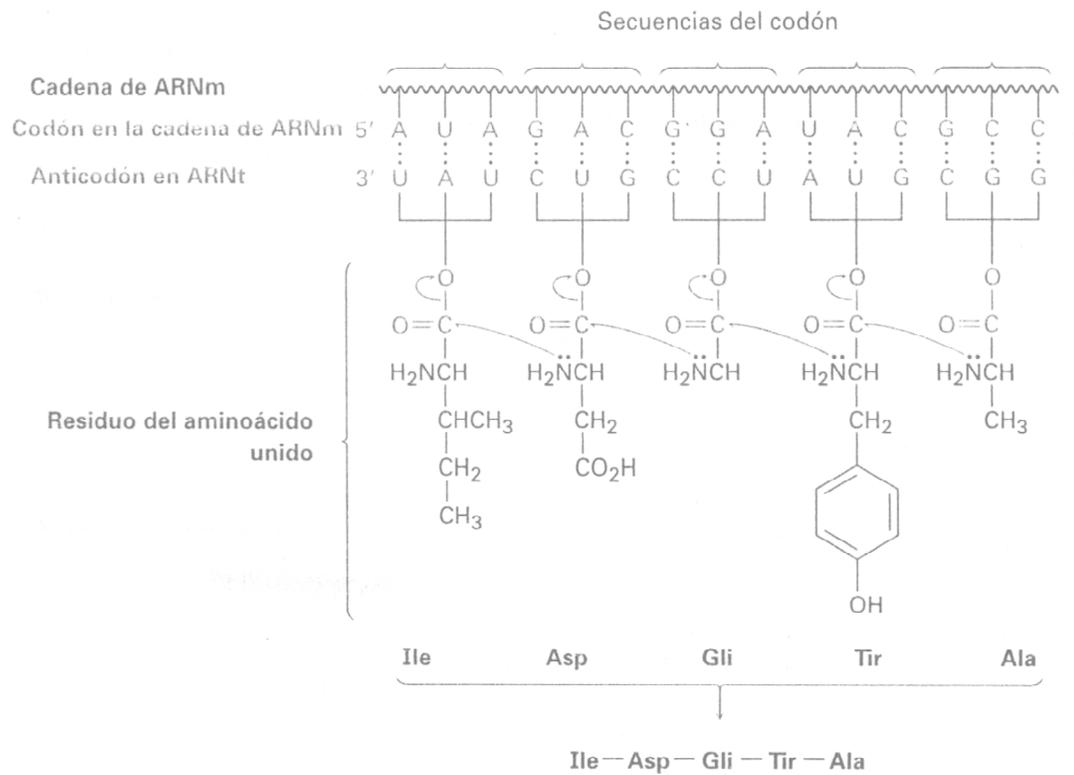


Figura 28.7 Representación de la biosíntesis de proteínas. Las secuencias de bases del codón en el ARNm son leídas por los ARNt que contienen las secuencias de bases del anticodón complementarias. Los ARN de transferencia ensamblan los aminoácidos apropiados en la posición para su incorporación en el péptido en crecimiento.



Problema 28.7

Liste las secuencias del codón para los siguientes aminoácidos:

- (a) Ala
- (b) Fen
- (c) Leu
- (d) Tir

Problema 28.8

Liste las secuencias del anticodón en los ARNt que transportan los aminoácidos mostrados en el problema 28.7.

Problema 28.9

¿Qué secuencia de aminoácidos se codifica por la secuencia de bases del ARN, siguiente?

CUU-AUG-GCU-UGG-CCC-UAA

Problema 28.10

¿Cuál es la secuencia de bases en la cadena de ADN original de la que se hizo la secuencia del ARNm en el problema 28.9?

28.6 Secuenciación de ADN

Una de las revoluciones científicas más grandes de la historia está ahora en curso en la biología molecular, ya que los científicos están aprendiendo cómo manipular y modificar la maquinaria genética de los organismos; sin embargo, ninguno de los avances extraordinarios de las dos décadas pasadas habría sido posible si no fuera por el descubrimiento en 1977 de los métodos para la secuenciación de las inmensas cadenas de ADN.

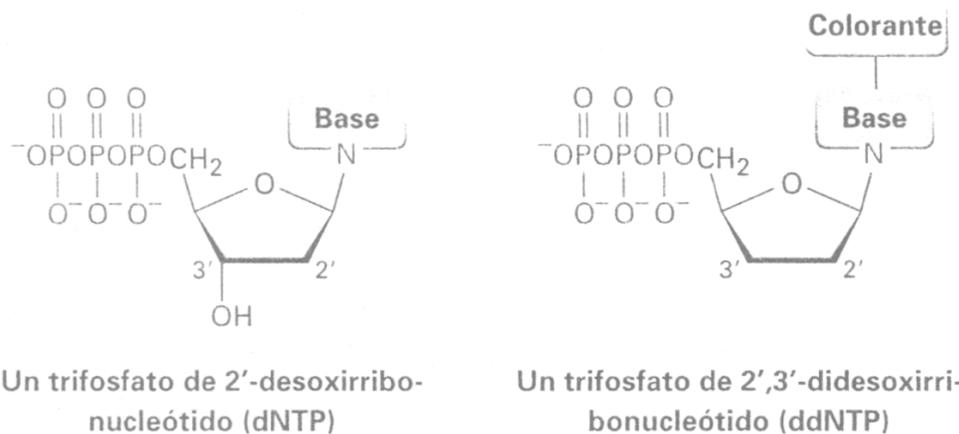
La primera etapa en la secuenciación de ADN es romper la cadena enorme en puntos conocidos para producir segmentos más pequeños y manejables, una tarea lograda por el uso de *endonucleasas de restricción*. Cada enzima de restricción diferente, de las que se conocen más de 3500 y aproximadamente 200 están disponibles comercialmente, rompe una molécula de ADN en un punto en la cadena donde ocurre una secuencia de bases específica; por ejemplo, la enzima de restricción Aitrl rompe entre G y C en la secuencia de cuatro bases AG-CT. Nótese que la secuencia es un *palíndromo*, lo que significa que la *secuencia* (5')- AGCT-(3') es la misma que su complemento (3')TCGA-(5') cuando ambas se leen en la misma dirección 5' → 3'. Lo mismo es verdadero para otras endonucleasas de restricción.

Si la molécula de ADN original se corta con otra enzima de restricción que tiene una especificidad diferente para la ruptura, se producen otros segmentos cuyas secuencias se traslapan parcialmente con aquéllas producidas por la primera

enzima. La secuenciación de todos los segmentos, seguida por la identificación de las regiones traslapadas, permite la secuenciación completa de ADN.

Están disponibles dos métodos de secuenciación de ADN. El *método Maxam Gilbert* utiliza técnicas químicas, mientras que el método didesoxi de Sanger utiliza reacciones enzimáticas. El método de Sanger es el que más se utiliza de los dos y fue el responsable de la secuenciación del genoma humano entero de 2.9 miles de millones de pares de bases. En los instrumentos comerciales de secuenciación, el método didesoxi comienza con una mezcla de los siguientes:

- El fragmento de restricción a ser secuenciado
- Un pequeño segmento de ADN llamado cebador (o molde), cuya secuencia es complementaria a la del extremo 3' del fragmento de restricción
- Los cuatro trifosfatos de 2'-desoxirribonucleótidos (dNTPs)
- Cantidades muy pequeñas de los cuatro trifosfatos de 2',3'-didesoxirribonucleótidos (ddNTPs), cada uno de los cuales se marca con un colorante fluorescente de un color diferente (un trifosfato de 2',3'-didesoxirribonucleótido es aquel en el que faltan los grupos 2' y 3' —OH de la ribosa).



Se adiciona la ADN polimerasa a la mezcla, y una cadena de ADN complementaria al fragmento de restricción empieza a crecer del extremo del ADN cebador. La mayor parte del tiempo sólo se incorporan en la cadena de crecimiento desoxirribonucleótidos normales debido a su concentración mucho mayor en la mezcla, pero a menudo se incorpora un didesoxirribonucleótido. Cuando esto sucede, se detiene la síntesis de ADN debido a que el extremo de la cadena ya no tiene un grupo 3'-hidroxilo para que adicione más nucleótidos.

Cuando se completa la reacción, el producto consiste en una mezcla de fragmentos de ADN de todas las longitudes posibles, cada uno terminado por uno de los cuatro dideoxirribonucleótidos marcados con un colorante. Esta mezcla de productos se separa de acuerdo con el tamaño de las piezas por electroforesis en gel (sección 26.2), y la identidad del dideoxirribonucleótido en cada pieza y, por lo tanto, la secuencia del fragmento de restricción, se identifica simplemente notando el color con que fluoresce el colorante fijado. La figura 28.8 muestra un resultado típico.

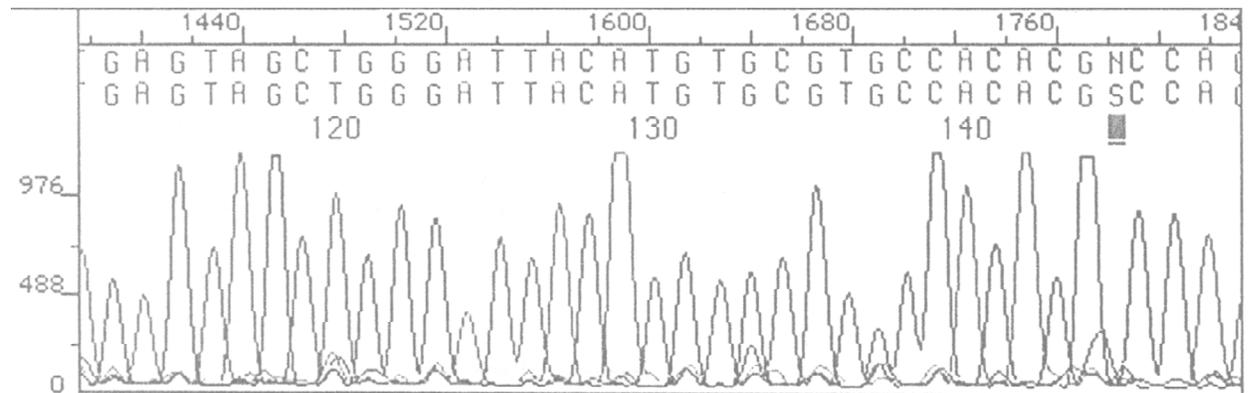


Figura 28.8 La secuencia de un fragmento de restricción determinada por el método dideoxi de Sanger puede leerse simplemente notando los colores del colorante fijado a cada uno de los varios nucleótidos terminales.

Es tan eficiente el método dideoxi automatizado que pueden secuenciarse con una exactitud de 98% secuencias de hasta 1100 nucleótidos de longitud, con un rendimiento de hasta 19 000 bases por hora. Después de una década de trabajo, se anunció en 2001 la información de la secuencia preliminar del genoma humano entero de 2.9 miles de millones de pares de bases. Extraordinariamente, nuestro genoma parece contener sólo alrededor de 30 000 genes, menos de un tercio del número predicho anteriormente y sólo el doble del número encontrado en el ascáride común.

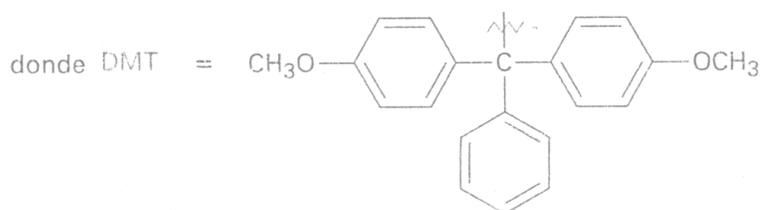
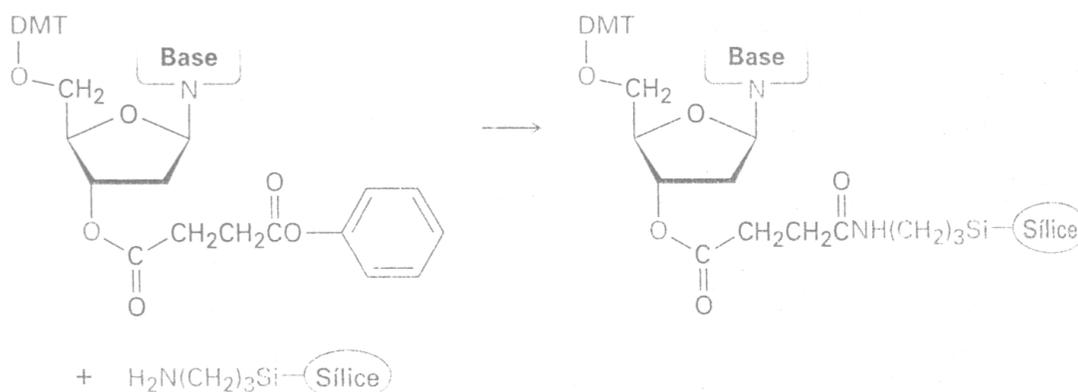
28.7 Síntesis de ADN

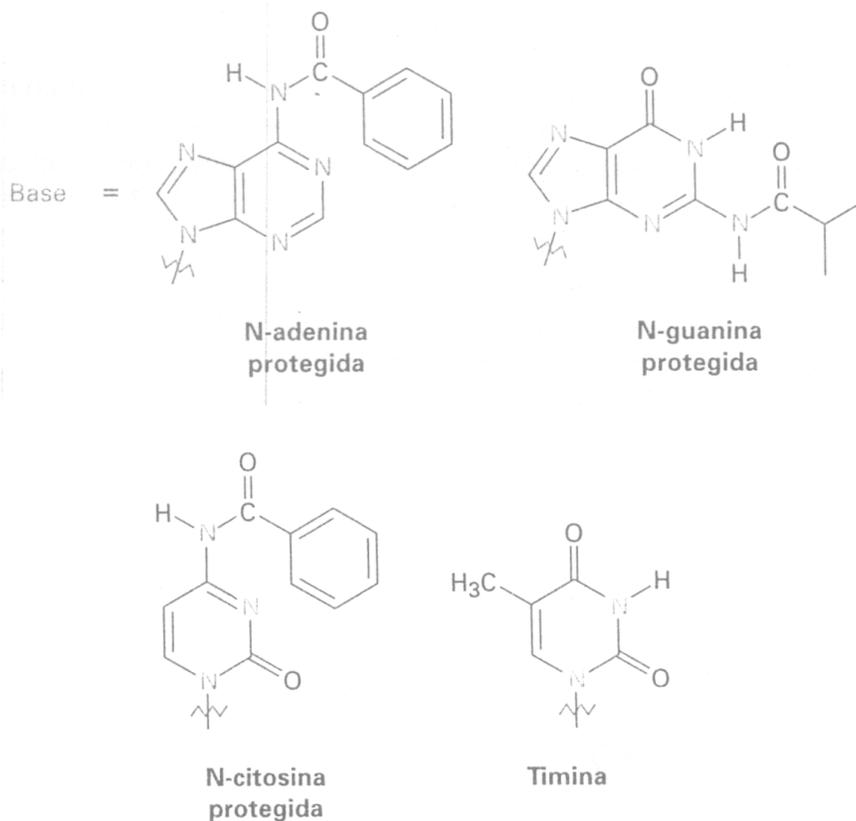
La revolución en curso en biología molecular ha traído consigo un incremento de la demanda de la síntesis química eficiente de segmentos cortos de ADN, llamados *oligonucleótidos*, o simplemente *oligos*. Los problemas de la síntesis de ADN son similares a los de la síntesis de proteínas (sección 26.7) pero son más difíciles debido a la complejidad de los monómeros nucleótidos. Cada nucleótido tiene múltiples sitios reactivos que deben protegerse y desprotegerse selectivamente en los tiempos apropiados, y el acoplamiento de

los cuatro nucleótidos debe realizarse en la secuencia apropiada. Sin embargo, están disponibles sintetizadores de ADN automatizados que permiten la síntesis rápida y confiable de segmentos de ADN de hasta 200 nucleótidos de longitud.

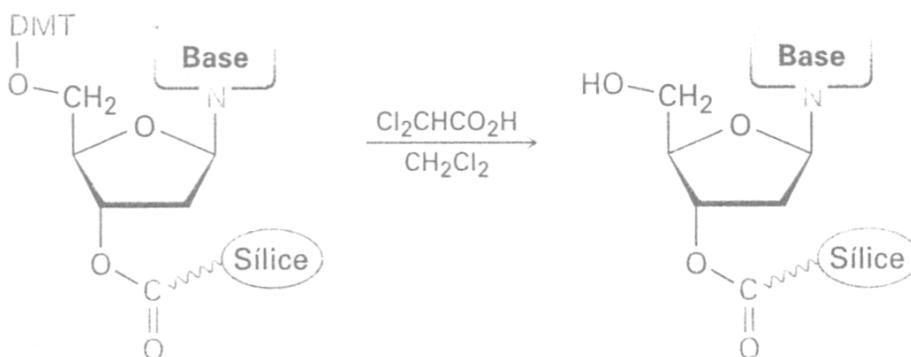
Los sintetizadores de ADN operan bajo un principio similar al del sintetizador en fase sólida de péptidos (sección 26.8). En esencia, un nucleótido protegido se une de modo covalente a un soporte sólido, y se adiciona de manera simultánea un nucleótido a la cadena en crecimiento por el uso de un reactivo de acoplamiento. Después de que se ha adicionado el nucleótido final, se eliminan todos los grupos protectores y se separa el ADN sintético del soporte sólido. Se necesitan cinco pasos:

Paso 1 El primer paso en la síntesis de ADN es fijar un desoxinucleósido protegido al soporte de sílice (SiO_2) por medio de un enlace éster al grupo —OH en 3' del desoxinucleósido. Deben protegerse el grupo —OH en 5' en el azúcar y los grupos —NH_2 libres en las bases heterocíclicas. Las bases adenina y citosina son protegidas por grupos benzoílo, la guanina es protegida por un grupo isobutirilo y la tiamina no requiere protección. El —OH en 5' de la desoxirribosa se protege como su éter *p*-dimetoxitritílico (DMT).

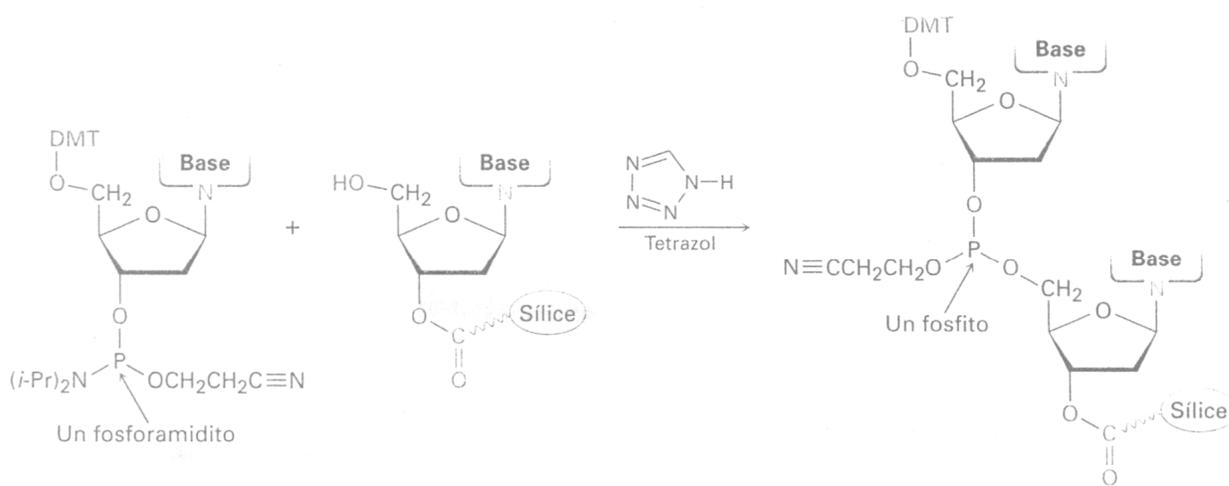




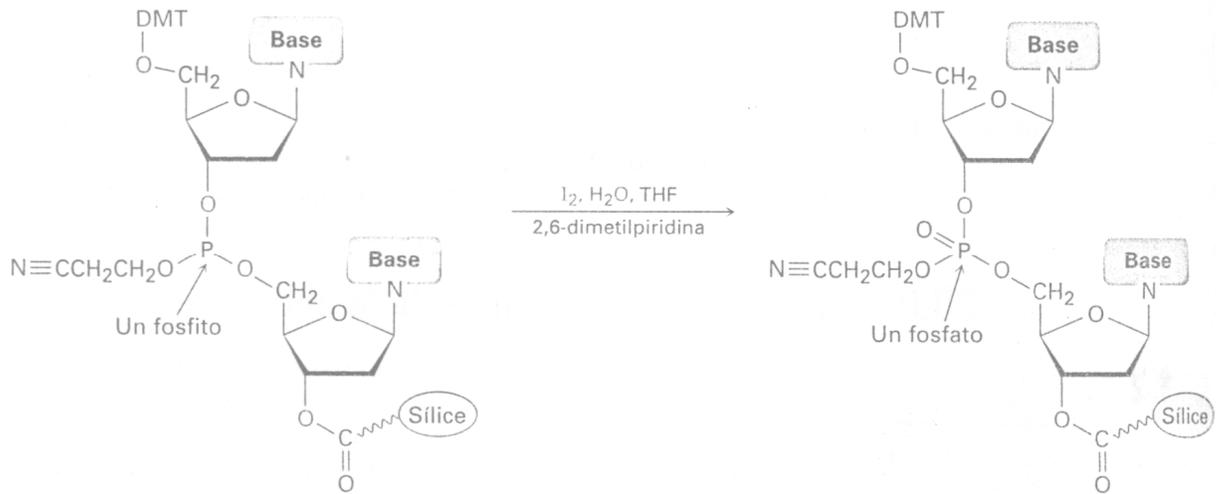
Paso 2 El segundo paso es la eliminación del grupo protector DMT por el tratamiento con ácido dicloroacético en CH_2Cl_2 . La reacción ocurre por un mecanismo $\text{S}_{\text{N}}1$ y procede rápidamente debido a la estabilidad del catión terciario dimetoxitritilo bencílico.



Paso 3 El tercer paso es el acoplamiento del desoxinucleósido unido al polímero con un desoxinucleósido protegido que contiene un grupo *fosforamidito* en su posición 3'. [Un fosforamidito tiene la estructura $R_2NP(OR)_2$.] La reacción de acoplamiento se lleva a cabo en acetonitrilo, un disolvente polar aprótico; ésta es catalizada por la amina heterocíclica tetrazol; y genera como producto un *fosfito*, $P(OR)_3$. Observe que uno de los átomos de oxígeno del fósforo está protegido por un grupo p-cianoetilo, $-OCH_2CH_2C \equiv N$. El paso de acoplamiento tiene lugar con un rendimiento superior a 99%.

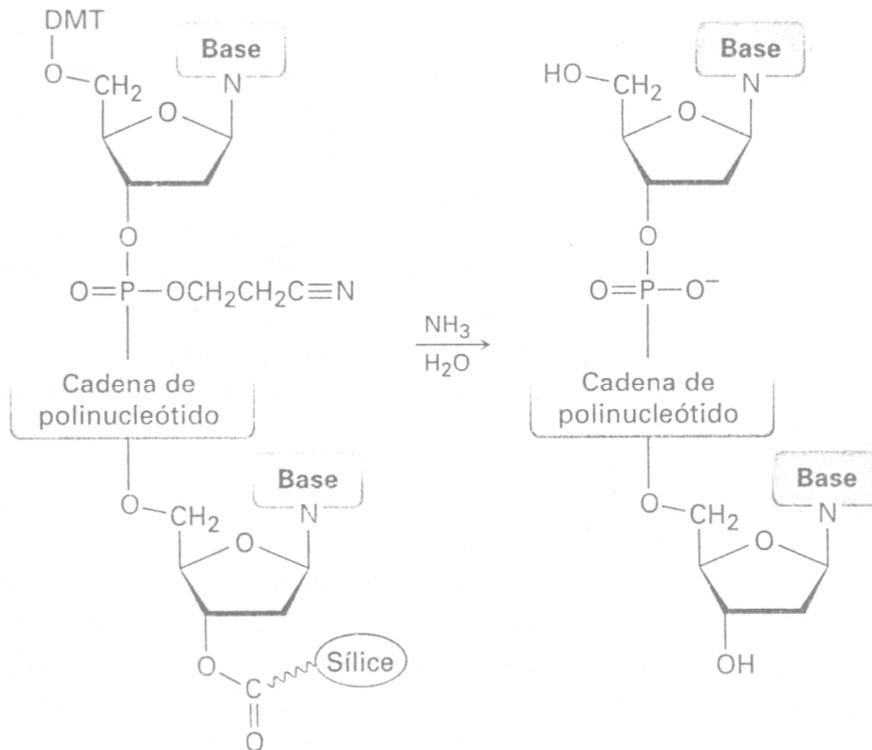


Paso 4 Con el acoplamiento realizado, se oxida el fosfito producido a un fosfato por el tratamiento con yodo en tetrahidrofurano acuoso en presencia de 2,6-dimetilpiridina. El ciclo (1) desprotección, (2) acoplamiento y (3) oxidación se repite hasta que se ha formado una cadena de oligonucleótido de la secuencia deseada.



Paso 5

El paso final es la eliminación de todos los grupos protectores y la ruptura del enlace éster que sostiene el ADN a la sílice. Todas estas reacciones se hacen al mismo tiempo por el tratamiento con NH₃ acuoso. La purificación por electroforesis produce el ADN sintético.



Problema 28.11

Los éteres *p*-dimetoxitritílicos (DMT) se rompen fácilmente por el tratamiento con un ácido suave. Muestre el mecanismo de la reacción de ruptura.

Problema 28.12

Proponga un mecanismo que explique la ruptura del grupo protector β -cianoetilo de los grupos fosfato en el tratamiento con amoníaco acuoso. (El acrilonitrilo, $\text{H}_2\text{C}=\text{CHCN}$, es un subproducto.) ¿Qué tipo de reacción está sucediendo?

28.8 La reacción en cadena de la polimerasa**Kary Banks Mullis**

Kary Banks Mullis (1944-) nació en la zona rural de Lenoir, Carolina del Norte; hizo sus estudios universitarios en el Tecnológico de Georgia; y recibió su doctorado por la Universidad de California, Berkeley, en 1973. De 1979 a 1986 trabajó en la corporación Cetus, donde realizó su trabajo sobre el desarrollo de la RCP. Desde 1988 siguió su propio ritmo como consultor y escritor independiente. Recibió en 1993 el Premio Nobel de Química.

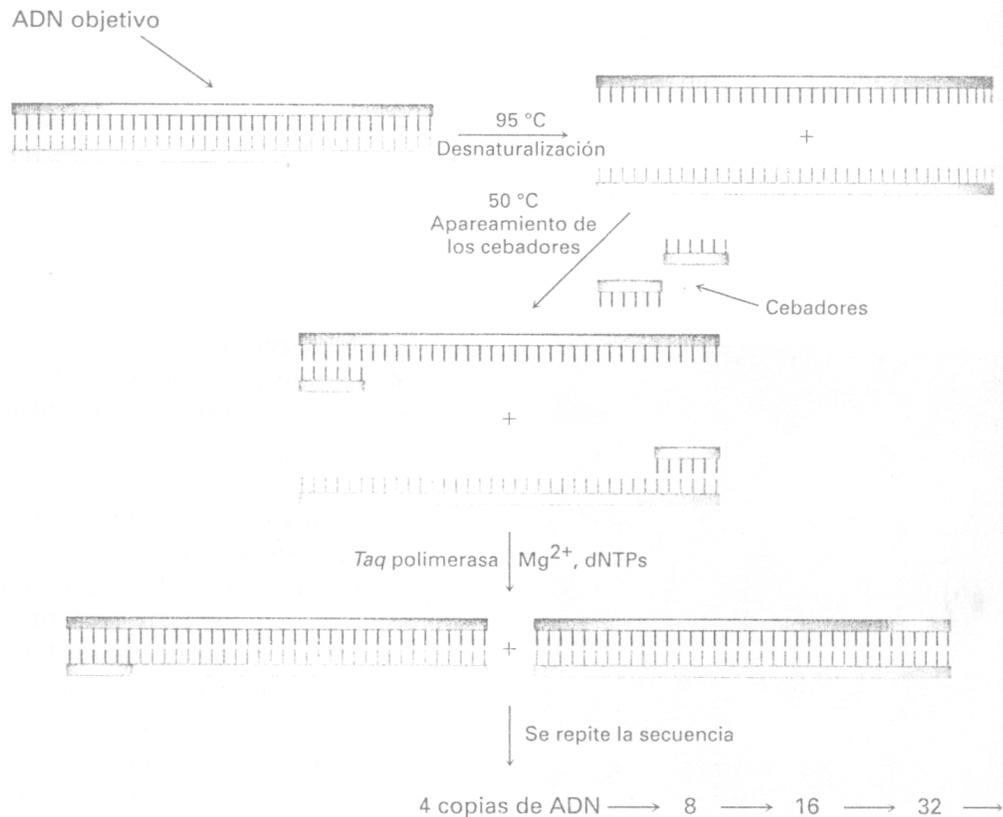
Con frecuencia sucede que sólo pueden obtenerse directamente a partir del ADN de un organismo pequeñas cantidades de una secuencia de genes, por lo que a veces se necesitan métodos que realicen la secuenciación para la obtención de cantidades mayores. La invención de la reacción en cadena de la polimerasa (RCP) por Kary Mullis en 1986 ha sido descrita que es a los genes lo que fue la invención de la imprenta de Gutenberg a la palabra escrita. Al igual que la imprenta produce múltiples copias de un libro, la RCP produce múltiples copias de una secuencia de ADN dada. Comenzado a partir de 1 *picogramo* de ADN con una longitud de cadena de 10,000 nucleótidos (1 pg = 10^{-12} g; alrededor de 100 000 moléculas), la RCP hace posible obtener varios microgramos (1 μg = 10^{-6} g; alrededor de 10^{11} moléculas) en sólo unas cuantas horas.

La clave de la reacción en cadena de la polimerasa es la *Taq* ADN polimerasa, una enzima termoestable aislada de la bacteria termófila *Thermus aquaticus* que se encontró en tu *t* manantial de agua caliente en el Parque Nacional de Yellowstone. La *Taq* polimerasa es capaz de tomar una sola cadena de ADN que tiene un segmento cebador corto de la cadena complementaria en un extremo y terminar de construir toda la cadena complementaria. Como ¡nuestra de manera esquemática la figura 28.9, el proceso global sucede en tres pasos. [Más recientemente, han llegado a estar disponibles enzimas ADN polimerasas termoestables mejoradas, que incluyen la *Vent* polimerasa y la *Pfu* polimerasa, ambas aisladas de las bacterias que crecen cerca de los respiraderos (chimeneas) geotérmicos en el fondo del océano. El grado de error de ambas enzimas es sustancialmente menor que el de la *Taq*].

- Paso 1** Se calienta el ADN de doble cadena a ser amplificado en presencia de la Taq polimerasa, ion Mg^{2+} , los cuatro monómeros de trifosfato de desoxinucleótido, (dNTPs) y un gran exceso de dos oligonucleótidos cebadores coitos de casi 20 bases cada uno. Cada cebador es complementario a la secuencia en el extremo de uno de los segmentos de ADN objetivos. A una temperatura de 95 °C, el ADN de cadena doble se desnaturaliza y se rompe espontáneamente en dos cadenas simples separadas.
- Paso 2** Se disminuye la temperatura entre 37 y 50 °C, lo que permite que los cebadores, debido a su concentración relativamente alta, se apareen formando nuevos puentes de hidrógeno con su secuencia complementaria en el extremo de cada cadena objetivo.
- Paso 3** Se eleva la temperatura a 72 °C y la Taq polimerasa cataliza la adición de más nucleótidos a las dos cadenas de ADN cebadoras. Cuando se finaliza la replicación de cada cadena, existen dos copias del ADN original. La repetición del ciclo desnaturalización-apareamiento-síntesis una segunda vez produce cuatro copias de ADN, la repetición de una tercera vez produce ocho copias, y así sucesivamente, en una serie exponencial.

Se ha automatizado la RCP, y pueden realizarse en una hora aproximadamente 30 ciclos, lo que resulta en un factor de amplificación teórico de 230 ($\sim 10^9$); sin embargo, en la práctica la eficiencia de cada ciclo es menor del 100% y habitualmente se logra una amplificación experimental de alrededor de 10^6 a 10^8 para 30 ciclos.

Figura 28.9 La reacción en cadena de la polimerasa. Los detalles se explican en el texto.



Enfocado a . . .



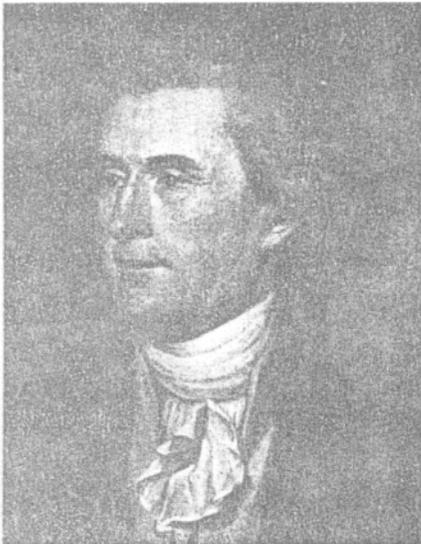
Identificación de ADN

La invención de la secuenciación de ADN ha afectado a la sociedad de varias maneras, algunas más dramáticas que las que se derivan del desarrollo de la *identificación de ADN*. La identificación de ADN surgió del descubrimiento en 1984 de que los genes humanos contienen secuencias repetitivas cortas de ADN no codificado, llamadas sitios de *repetición en tándem corto* (STR, por sus siglas en inglés). Además, los sitios STR son ligeramente distintos para cada individuo, excepto

en gemelos idénticos. Por la secuenciación de estos sitios puede obtenerse un patrón único para cada persona.

Quizás el uso más común y mejor publicitado de la identificación de ADN es el realizado por los laboratorios de criminalística para vincular sospechosos a la evidencia biológica —sangre, folículos del cabello, piel o semen— encontrada en la escena del crimen. Miles de juicios se han decidido con base en la evidencia de ADN.

Para su uso en procesos criminales, los laboratorios forenses en Estados Unidos han convenido en 13 sitios de STR centrales que son los más exactos para la identificación de un individuo. Con base en estos 13 sitios, se ha establecido un Sistema de Índice Combinado de ADN (CODIS, por sus siglas en inglés) que ha servido como un registro de delincuentes convictos. Cuando se obtiene una muestra de ADN de una escena del crimen, se somete a una ruptura con endonucleasas de restricción para cortar fragmentos que contengan el sitio STR, se amplifican los fragmentos utilizando la reacción de cadena de la polimerasa y se determinan las secuencias de los fragmentos.



Los historiadores se han preguntado por varios años si Thomas Jefferson procreó un hijo con Sally Hemings. La evidencia de la identificación de ADN obtenida en 1998 no es concluyente pero lo sugiere de manera firme.

Si coinciden el perfil de las secuencias de un individuo conocido y el perfil del ADN obtenido en una escena del crimen, la probabilidad de que el ADN pertenezca al mismo individuo es de aproximadamente de 82 mil millones a 1. En casos de paternidad, donde el ADN del padre y de la progenie están relacionados pero no son totalmente idénticos, la identidad del padre puede establecerse con una probabilidad de 100,000 a 1. Aun después de que han pasado varias generaciones, puede implicarse la paternidad por el análisis de ADN del cromosoma Y de los descendientes directos de la línea masculina. El más y mejor conocido de tales casos es el de Thomas Jefferson, quien pudo haber procreado un hijo con su esclava Sally Hemings. Aunque Jefferson no tiene descendientes de la línea masculina, los análisis de ADN de los descendientes de la línea masculina del tío paterno de Jefferson contuvo el mismo cromosoma Y que un descendiente de la línea masculina de Eston Hemings, el hijo más joven de Sally Hemings. Por tanto, es clara una mezcla de los dos genomas, aunque el responsable individual masculino para tal mezcla no puede identificarse de manera concluyente.

Entre sus varias otras aplicaciones, la identificación de ADN se utiliza ampliamente para el diagnóstico de los desórdenes genéticos, de manera prenatal y en los recién nacidos. La fibrosis cística, la hemofilia, la enfermedad de Huntington, la enfermedad de Tay-Sachs, la anemia de células falciformes y la talasemia, están entre las varias enfermedades que pueden detectarse, lo que permite el tratamiento temprano de un niño afectado. Además, al estudiar las identificaciones de ADN de los parientes con un historial de un desorden en particular, es posible identificar los patrones de ADN asociados con la enfermedad y quizás obtener indicios para la cura eventual. Además, el Departamento de Defensa de Estados Unidos solicita muestras de sangre y de saliva de todo el personal militar. Las muestras se almacenan y se extrae el ADN necesario para la identificación en caso de que surja una contingencia.