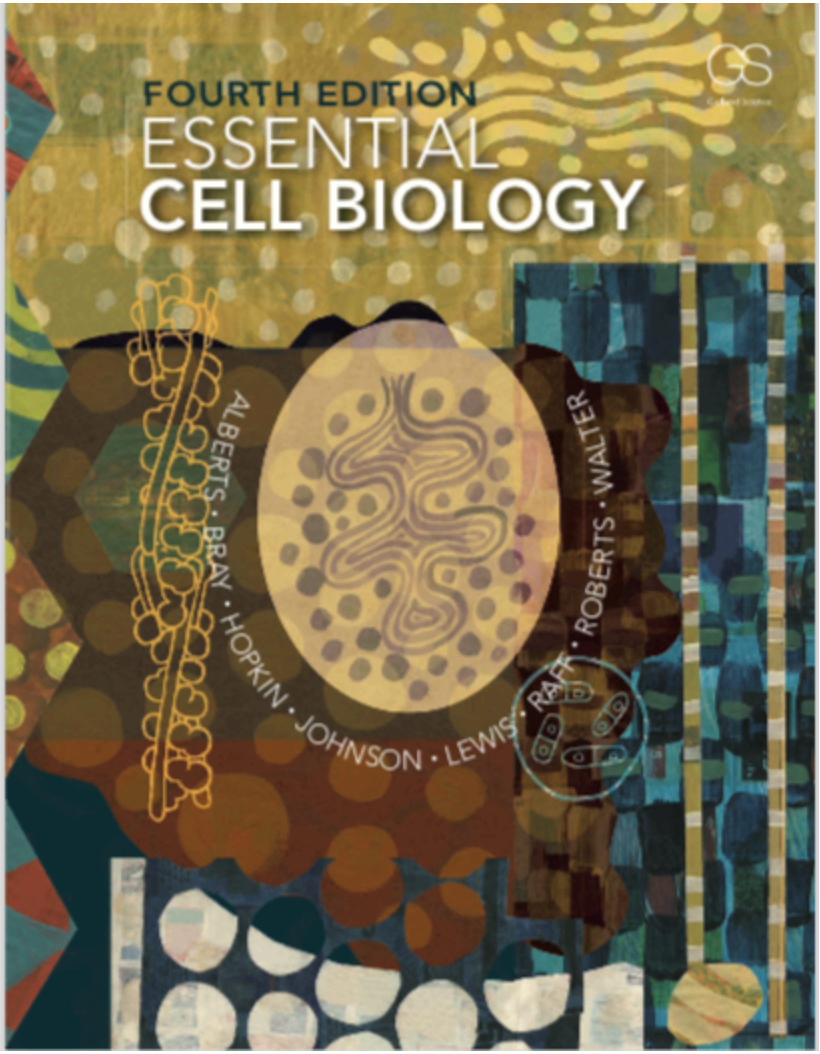
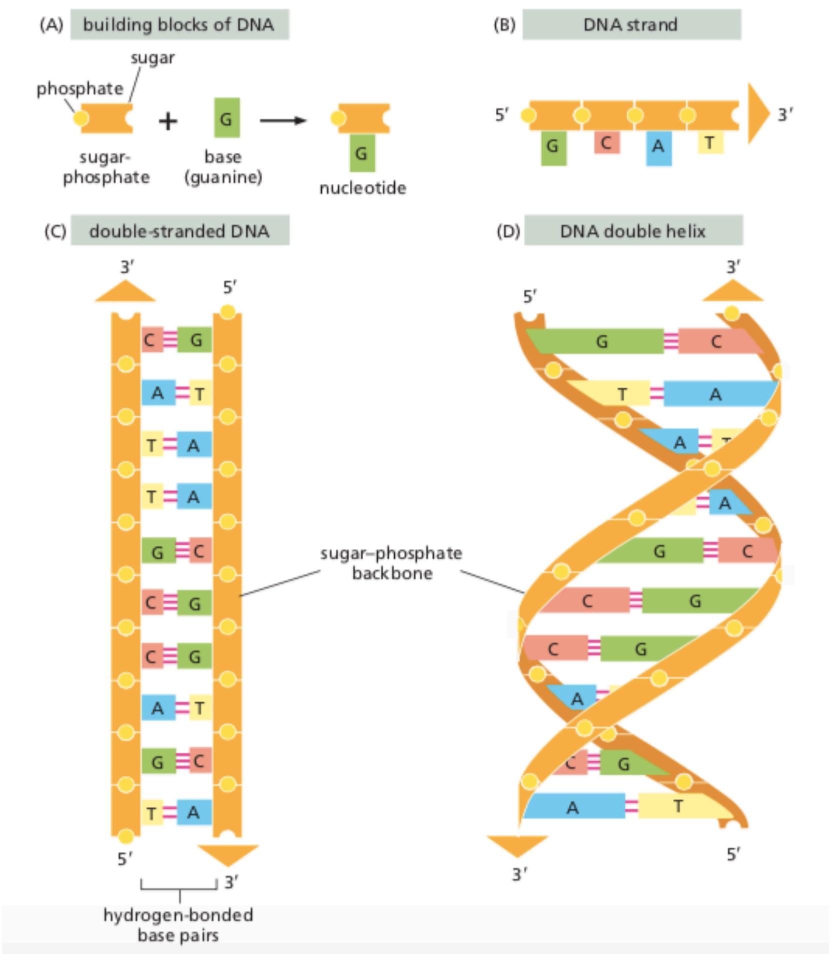
Sobre expresión génica, tomado del libro Essential Cell Biology de Alberts et al.


FOURTH EDITION
ESSENTIAL
CELL BIOLOGY
ALBERTS · BRAY · HOPKIN · JOHNSON · LEWIS · RAFF · ROBERTS · WALTER
GS

Algunas cosas que ya vimos

- Life depends on the stable storage and inheritance of genetic information.
- Genetic information is carried by very long DNA molecules and is encoded in the linear sequence of four nucleotides: A, T, G, and C.
- Each molecule of DNA is a double helix composed of a pair of antiparallel, complementary DNA strands, which are held together by hydrogen bonds between G-C and A-T base pairs.

La forma en que las subunidades de nucleótidos están unidas le da a una hebra de ADN una polaridad química. Si imaginamos que cada nucleótido tiene una perilla (el fosfato) y un orificio, cada hebra, formada por perillas entrelazadas con orificios, tendrá todas sus subunidades alineadas en la misma orientación. Además, los dos extremos de la hebra se pueden distinguir fácilmente, ya que uno tendrá un orificio (el hidroxilo de 3 ') y el otro una perilla (el fosfato de 5'). Esta polaridad en una hebra de ADN se indica refiriéndose a un extremo como el extremo 3 'y al otro como el extremo 5'. Esta convención se basa en los detalles del enlace químico entre las subunidades de nucleótidos.


(A) building blocks of DNA
sugar
phosphate
+
G
sugar-phosphate
base (guanine)
nucleotide
(B) DNA strand
5′
3′
G
C
A
T
(C) double-stranded DNA
3′
5′
C G
A T
T A
T A
G C
C G
C G
A T
G C
T A
5′
3′
hydrogen-bonded base pairs
sugar–phosphate backbone
(D) DNA double helix
5′
3′
G C
T A
A T
A
G C
C G
C G
A
C G
A T
3′
5′

- The genetic material of a eukaryotic cell is contained in a set of chromosomes, each formed from a single, enormously long DNA molecule that contains many genes.

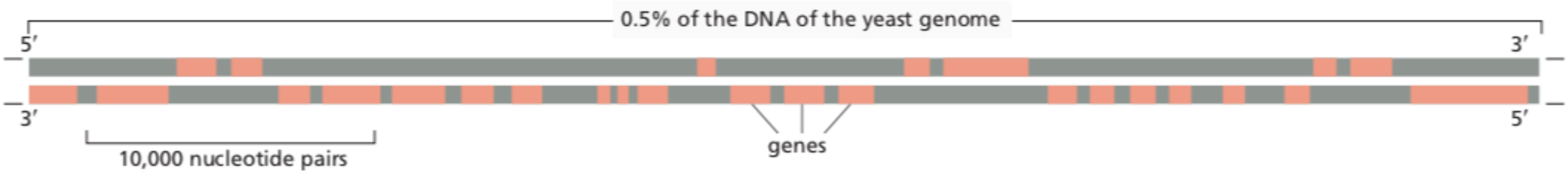
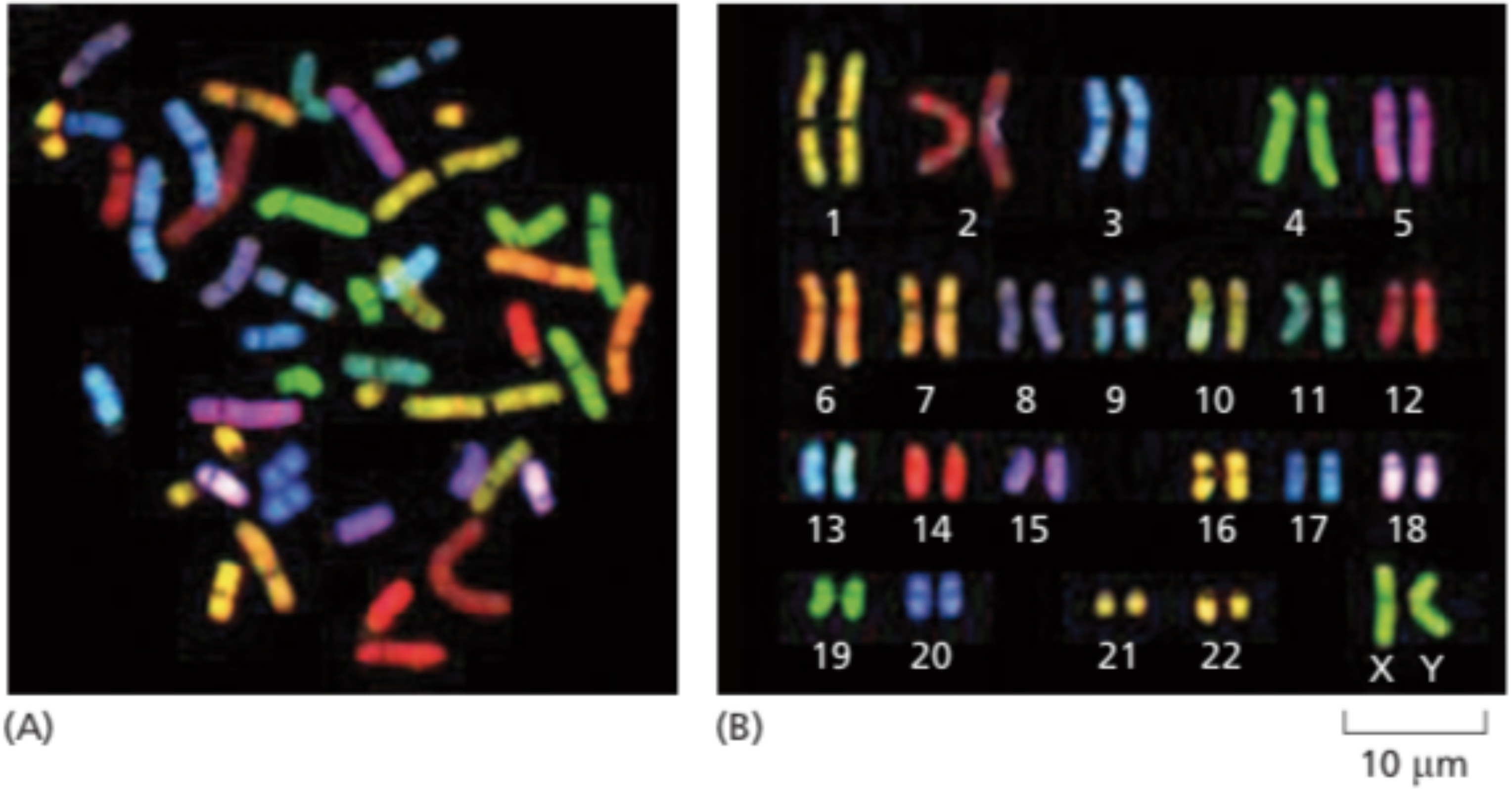


Figure 5–12 Genes are arranged along chromosomes. This figure shows a small region of the DNA double helix in one chromosome from the budding yeast *S. cerevisiae*. The *S. cerevisiae* genome contains about 12 million nucleotide pairs and 6600 genes—spread across 16 chromosomes. Note that, in each gene, only one of the two DNA strands actually encodes the information to make an RNA molecule, and this can be either strand, as indicated by the *light red* bars. However, a gene is generally denoted to contain both the "coding strand" and its complement, as in Figure 5–9. The high density of genes is characteristic of *S. cerevisiae*.



Each human chromosome can be "painted" a different color to allow its unambiguous identification. The chromosomes shown here were isolated from a cell undergoing nuclear division (mitosis) and are therefore in a highly compact (condensed) state. Chromosome painting is carried out by exposing the chromosomes to a collection of human DNA molecules that have been coupled to a combination of fluorescent dyes. For example, DNA molecules derived from chromosome 1 are labeled with one specific dye combination, those from chromosome 2 with another, and so on. Because the labeled DNA can form base pairs (hybridize) only to its chromosome of origin, each chromosome is differently colored. For such experiments, the chromosomes are treated so that the individual strands of the double-helical DNA molecules partly separate to enable base-pairing with the labeled, single-stranded DNA, while keeping the chromosome structure relatively intact. (a) Micrograph shows the array of chromosomes as they originally spilled from the lysed cell. (B) the same chromosomes have been artificially lined up in order. In this so-called *karyotype,* the homologous chromosomes are numbered and arranged in pairs; the presence of a Y chromosome reveals that these chromosomes came from a male. (From e. Schröck et al., *Science* 273:494– 497, 1996. With permission from the AAAS.)

¿Qué son los genes?

Son los segmentos a lo largo del ADN que son transcriptos en ARN.

When a gene is expressed, part of its nucleotide sequence is transcribed into RNA molecules, many of which are translated into protein.

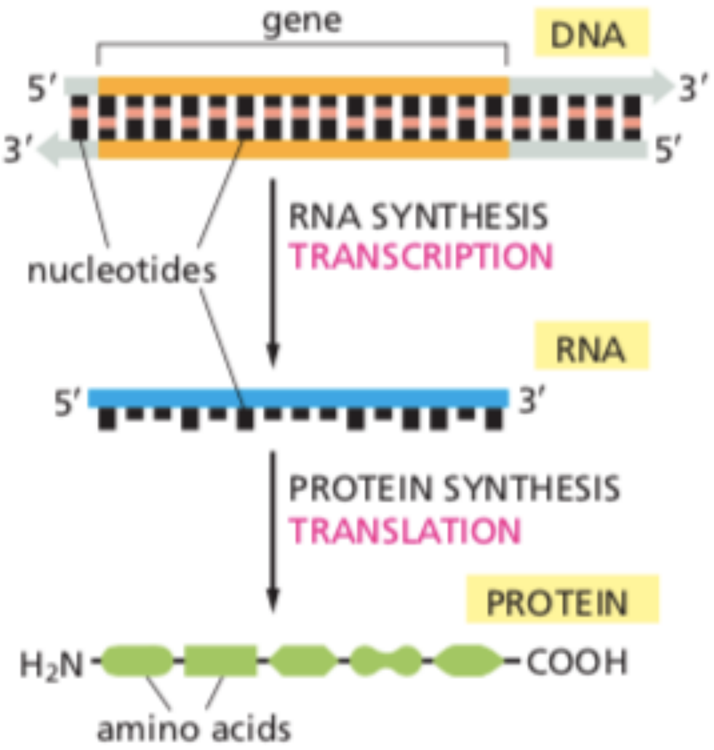
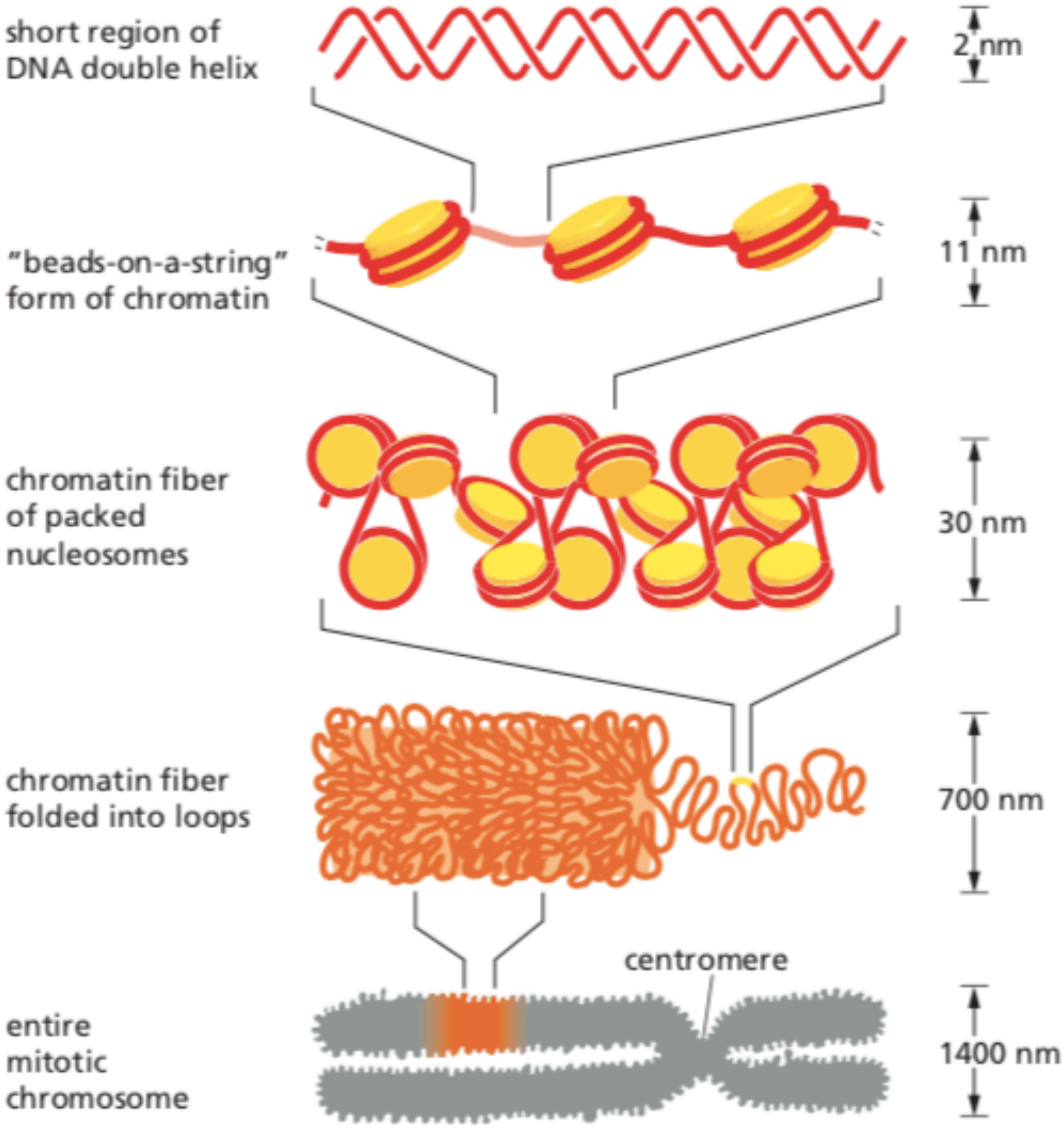

gene
DNA
5′
3′
3′
5′
nucleotides
RNA SYNTHESIS
TRANSCRIPTION
RNA
5′
3′
PROTEIN SYNTHESIS
TRANSLATION
PROTEIN
H_2N
COOH
amino acids

Figure 7–1 Genetic information directs the synthesis of proteins. The flow of genetic information from DNA to RNA (transcription) and from RNA to protein (translation) occurs in all living cells. It was Francis Crick who dubbed this flow of information "the central dogma." The segments of DNA that are transcribed into RNA are called genes.

- In eukaryotic chromosomes, the DNA is tightly folded by binding to a set of histone and nonhistone proteins. This complex of DNA and protein is called chromatin.
- Histones pack the DNA into a repeating array of DNA–protein particles called nucleosomes, which further fold up into even more compact chromatin structures.


short region of DNA double helix
2 nm
"beads-on-a-string" form of chromatin
11 nm
chromatin fiber of packed nucleosomes
30 nm
chromatin fiber folded into loops
700 nm
centromere
entire mitotic chromosome
1400 nm

NET RESULT: EACH DNA MOLECULE HAS BEEN PACKAGED INTO A MITOTIC CHROMOSOME THAT IS 10,000-FOLD SHORTER THAN ITS FULLY EXTENDED LENGTH

- A cell can regulate its chromatin structure—temporarily decondensing or condensing particular regions of its chromosomes—using chromatin-remodeling complexes and enzymes that covalently modify histone tails in various ways.
- The loosening of chromatin to a more decondensed state allows proteins involved in gene expression, DNA replication, and DNA repair to gain access to the necessary DNA sequences.

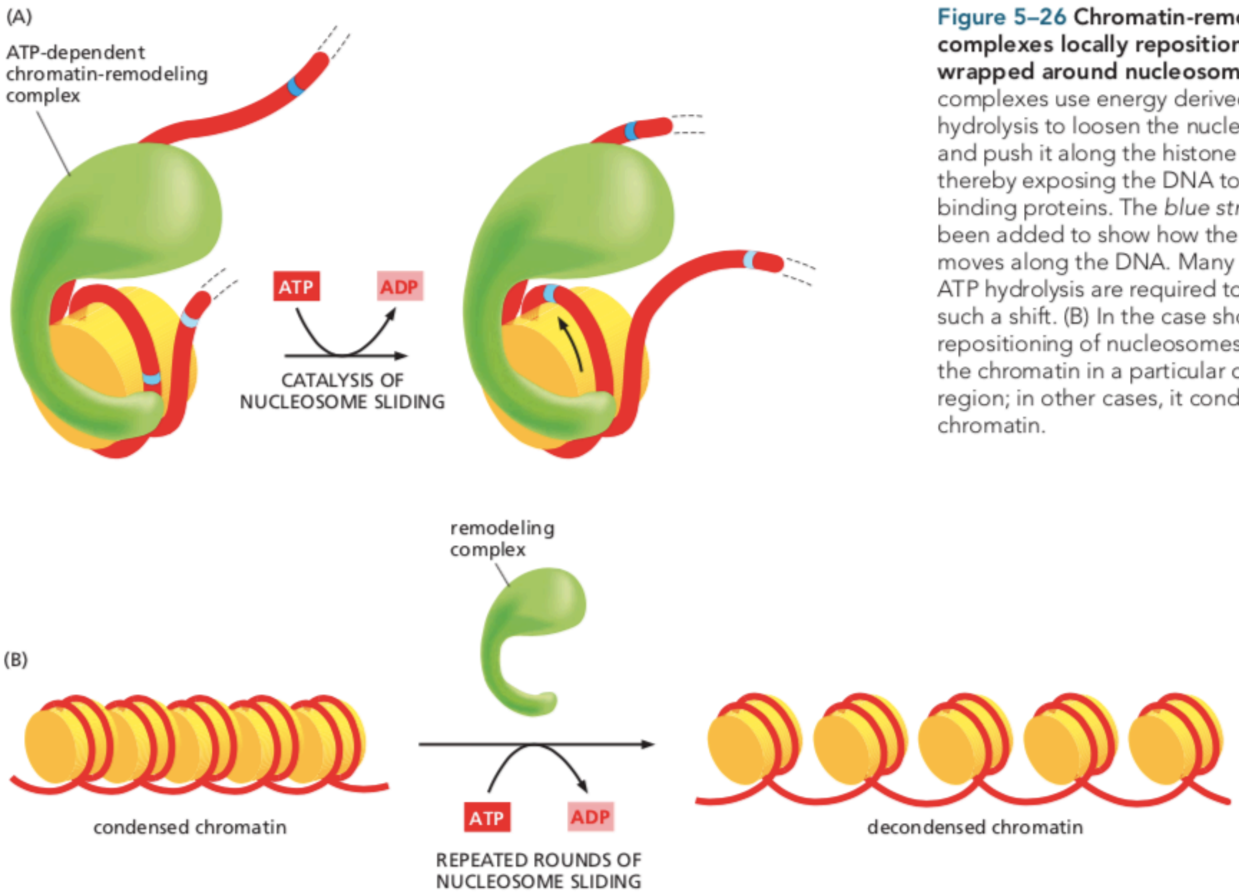

(A)
ATP-dependent chromatin-remodeling complex
ATP
ADP
CATALYSIS OF NUCLEOSOME SLIDING
remodeling complex
(B)
condensed chromatin
ATP
ADP
REPEATED ROUNDS OF NUCLEOSOME SLIDING
decondensed chromatin

Figure 5–26 Chromatin-remodeling complexes locally reposition the DNA wrapped around nucleosomes. (A) The complexes use energy derived from ATP hydrolysis to loosen the nucleosomal DNA and push it along the histone octamer, thereby exposing the DNA to other DNA-binding proteins. The *blue stripes* have been added to show how the nucleosome moves along the DNA. Many cycles of ATP hydrolysis are required to produce such a shift. (B) In the case shown, the repositioning of nucleosomes decondenses the chromatin in a particular chromosomal region; in other cases, it condenses the chromatin.

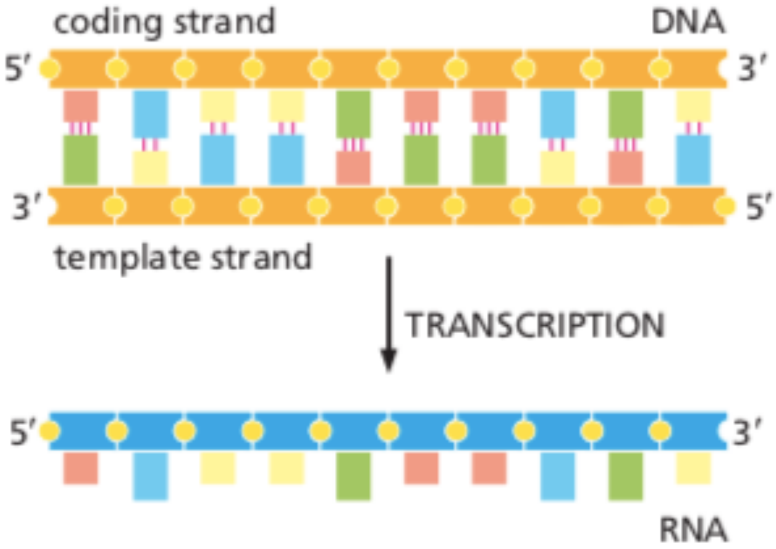
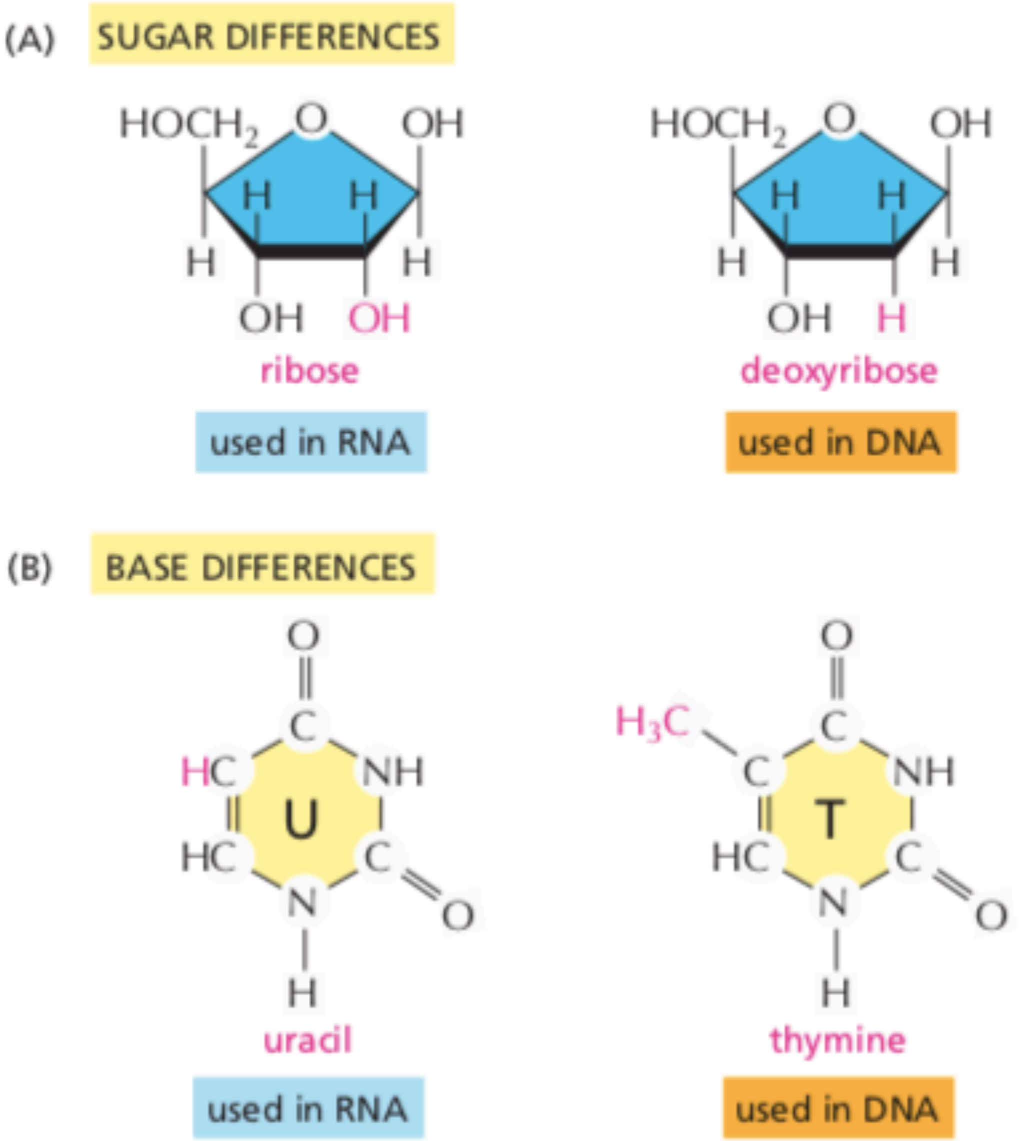

coding strand
DNA
5′
3′
3′
5′
template strand
TRANSCRIPTION
5′
3′
RNA

Figure 7–6 Transcription of a gene produces an RNA complementary to one strand of DNA. The transcribed strand of the gene, the *bottom* strand in this example, is called the template strand. The nontemplate strand of the gene (here, shown at the *top*) is sometimes called the *coding strand* because its sequence is equivalent to the RNA product, as shown. Which DNA strand serves as the template varies, depending on the gene, as we discuss later. By convention, an RNA molecule is always depicted with its 5′ end—the first part to be synthesized—to the left.

Además de que las moléculas de ARN están formadas por una sola hebra, hay otras diferencias químicas entre el ARN y el ADN.


(A)
SUGAR DIFFERENCES
HOCH2
O
OH
H
H
H
H
OH
OH
ribose
used in RNA
HOCH2
O
OH
H
H
H
H
OH
H
deoxyribose
used in DNA
(B)
BASE DIFFERENCES
O
C
HC
NH
U
HC
C
N
O
H
uracil
used in RNA
H3C
O
C
C
NH
T
HC
C
N
O
H
thymine
used in DNA

El ARN es de una sola hebra en general pero también se pueden pegar entre sí secuencias complementarias presentes en la misma cadena.
Esta habilidad para plegarse de distintos modos en estructuras tridimensionales compleja
le permite al ARN cumplir otras funciones además de ser un “reservorio” de información.

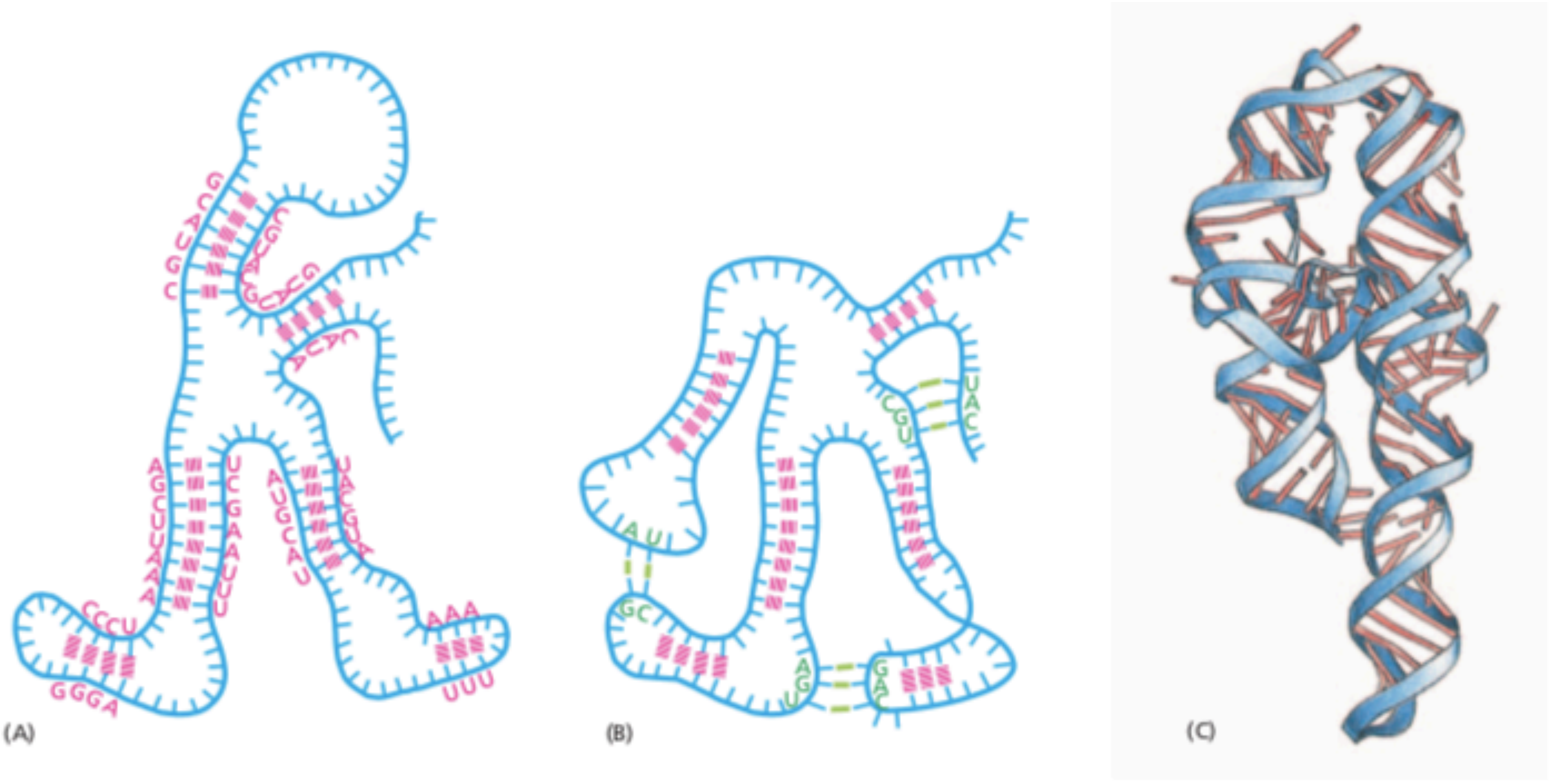


Figure 7–5 RNA molecules can form intramolecular base pairs and fold into specific structures. RNA is single-stranded, but it often contains short stretches of nucleotides that can base-pair with complementary sequences found elsewhere on the same molecule. These interactions—along with some “nonconventional base-pair interactions (e.g., A-G)—allow an RNA molecule to fold into a three-dimensional structure that is determined by its sequence of nucleotides. (A) A diagram of a hypothetical, folded RNA structure showing only conventional (G-C and A-U) base-pair interactions. (B) Incorporating nonconventional base-pair interactions (*green*) changes the structure of the hypothetical RNA shown in (A). (C) Structure of an actual RNA molecule that is involved in RNA splicing. This RNA contains a considerable amount of double-helical structure. The sugar–phosphate backbone is *blue* and the bases are *red*; the conventional base-pair interactions are indicated by red “rungs” that are continuous, and nonconventional base pairs are indicated by broken red rungs. For an additional view of RNA structure, see **Movie 7.1**.

Hay distintos tipos de ARN en las células. El que codifica para las proteínas es el mensajero.

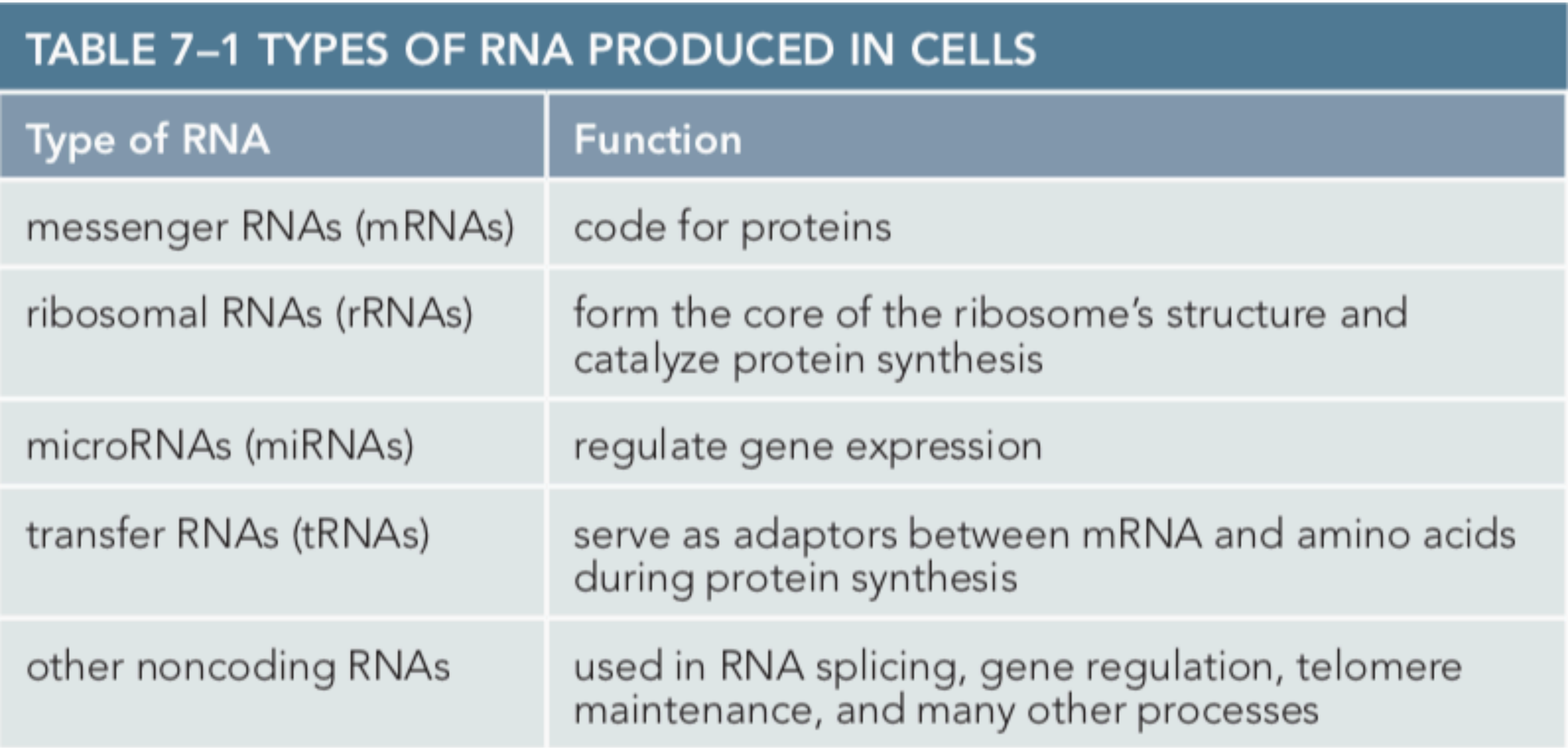
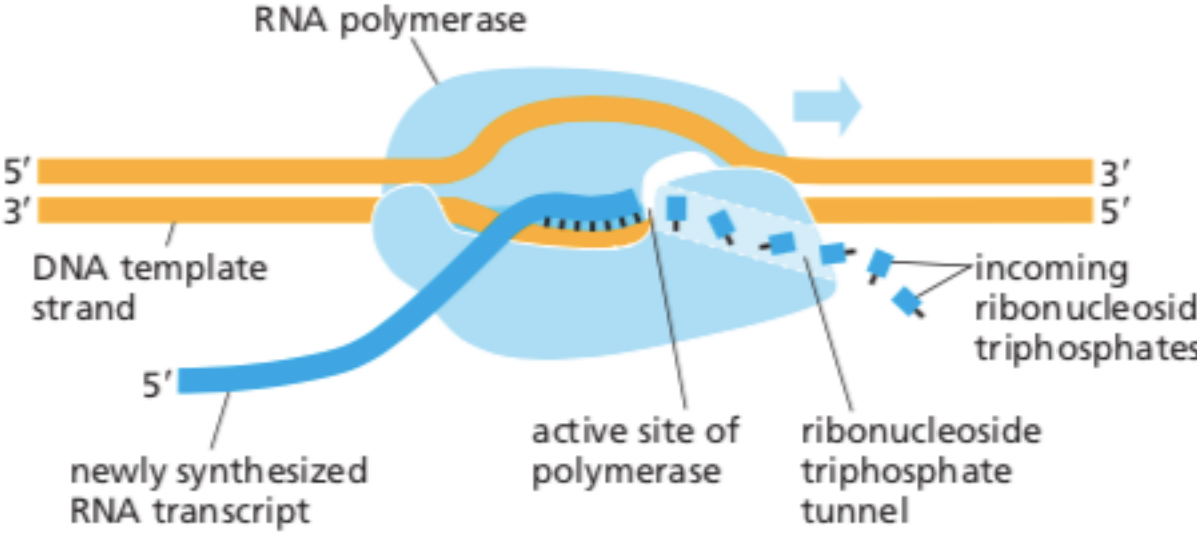


TABLE 7–1 TYPES OF RNA PRODUCED IN CELLS

Type of RNA	Function
messenger RNAs (mRNAs)	code for proteins
ribosomal RNAs (rRNAs)	form the core of the ribosome's structure and catalyze protein synthesis
microRNAs (miRNAs)	regulate gene expression
transfer RNAs (tRNAs)	serve as adaptors between mRNA and amino acids during protein synthesis
other noncoding RNAs	used in RNA splicing, gene regulation, telomere maintenance, and many other processes

¿Cómo ocurre la transcripción?

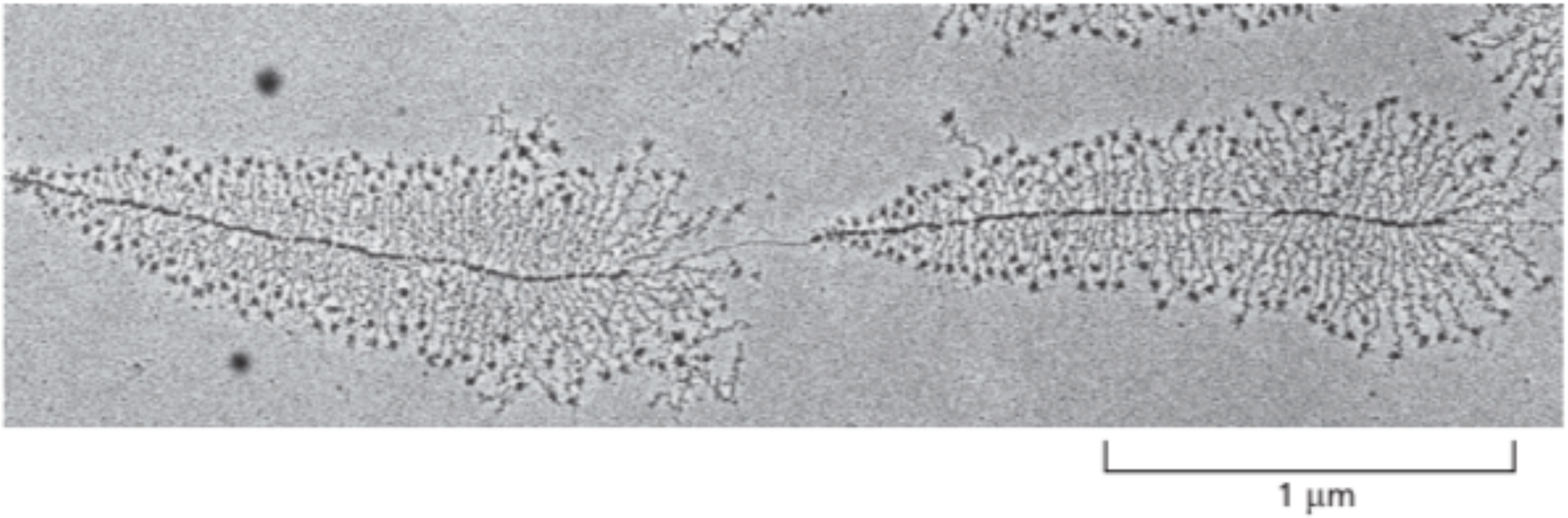
Interviene la ARN polimerasa que es una enzima.


RNA polymerase
5′
3′
3′
5′
DNA template strand
incoming ribonucleosid
triphosphates
5′
newly synthesized RNA transcript
active site of polymerase
ribonucleoside triphosphate tunnel

Las ARN polimerasas catalizan la formación de ciertos enlaces que permiten unir nucleótidos entre sí, armando así el esqueleto de azúcar-fosfato del ARN. Para ello utiliza una de las hebras de la doble hélice como molde. La transcripción del ARN es de cadena simple y complementaria a la cadena molde. La polimerasa se mueve paso a paso a lo largo del ADN (en la dirección 3’ a 5’), desenrollando la hélice de ADN justo delante para exponer una nueva región de la hebra molde. A medida que la polimerasa se mueve, desplaza el ARN recién formado, permitiendo que las dos hebras de ADN detrás de ella se re-enrollen. Se forma así, de forma transitoria, una región corta de hélice híbrida ADN/ARN (aproximadamente nueve nucleótidos de longitud) que se mueve a lo largo del ADN con la polimerasa.
La cadena de ARN en crecimiento se extiende de a un nucleótido a la vez en la dirección 5′ a 3 ′. Los trifosfatos entrantes (ATP, CTP, UTP y GTP) proporcionan la energía necesaria para impulsar la reacción.

La liberación casi inmediata de la hebra de ARN del ADN a medida que se sintetiza significa que se pueden hacer muchas copias de ARN a partir del mismo gen en un tiempo relativamente corto; la síntesis del siguiente ARN suele iniciarse antes de que se complete el primer ARN.

Un gen de tamaño mediano, ~ 1500 pares de nucleótidos, requiere aproximadamente 50 segundos para que una molécula de ARN polimerasa lo transcriba. En un momento dado, podría haber docenas de polimerasas acelerando a lo largo de este único tramo de ADN permitiendo que se sinteticen más de 1000 transcripciones en una hora.



Many molecules of RNA polymerase simultaneously transcribing two adjacent ribosomal genes on a single DNA molecule. Molecules of RNA polymerase are barely visible as a series of tiny dots along the spine of the DNA molecule; each polymerase has an RNA transcript (a short, fine thread) radiating from it.

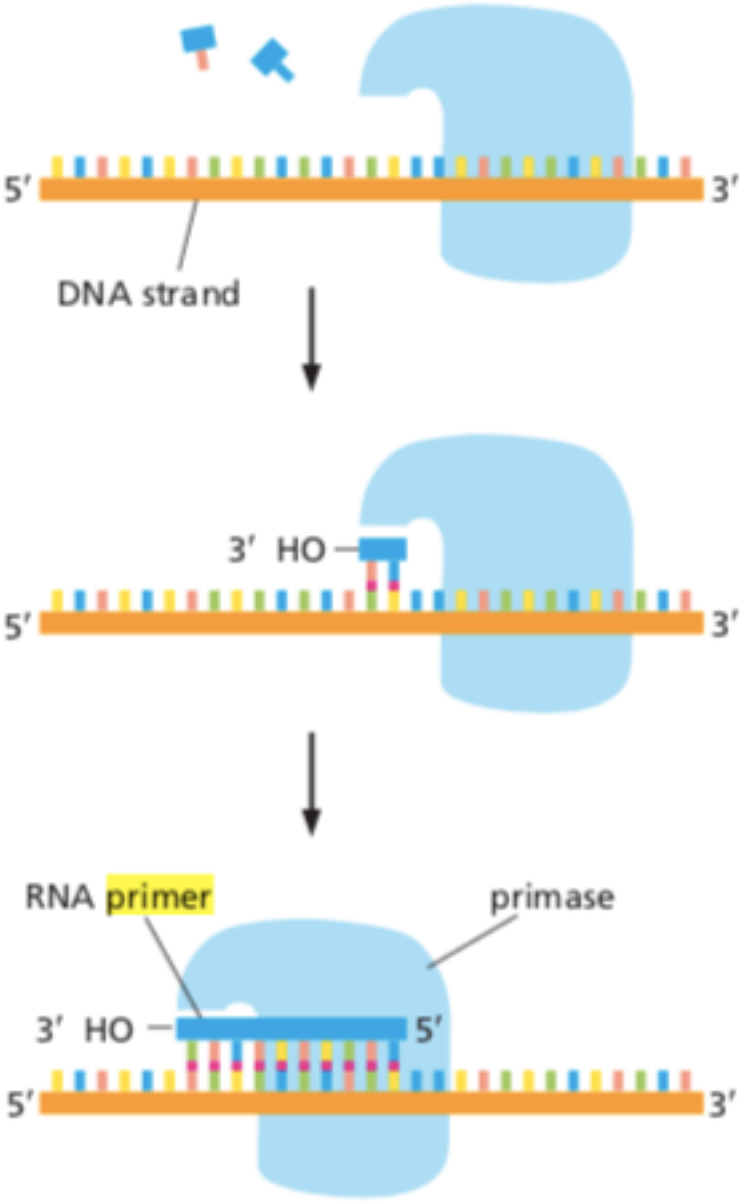
A diferencia de la ADN polimerasa involucrada en la replicación del ADN, las ARN polimerasas pueden iniciar una cadena de ARN sin un “primer”. Esta diferencia probablemente evolucionó porque la transcripción no necesita ser tan precisa como la replicación del ADN.

A diferencia del ADN, el ARN no se utiliza como forma de almacenamiento permanente de información genética en las células, por lo que los errores en las transcripciones de ARN tienen consecuencias relativamente menores para una célula.

Las ARN polimerasas cometen aproximadamente un error por cada 10^4 nucleótidos copiados en ARN, mientras que la ADN polimerasa comete solo un error por cada 10^7 nucleótidos copiados.

Para que la replicación del ADN ocurra, el modo en que funciona la ADN polimerasa requiere que el extremo 3′ esté correctamente apareado antes de que se puedan agregar más nucleótidos a la cadena de ADN en crecimiento. Entonces, ¿cómo puede la polimerasa comenzar una cadena de ADN completamente nueva? Para iniciar el proceso, se necesita una enzima diferente, una que pueda comenzar una nueva cadena de polinucleótidos simplemente uniendo dos nucleótidos sin la necesidad de que haya un extremo correctamente apareado. Esta enzima produce una cadena corta de ARN (ácido ribonucleico), utilizando la hebra de ADN como plantilla. Esta pequeña longitud de ARN, de aproximadamente 10 nucleótidos de longitud, está apareada con la cadena de la plantilla y proporciona un extremo 3′ con pares de bases como punto de partida para la ADN polimerasa. Por lo tanto, sirve como "primer" para la síntesis de ADN, y la enzima que sintetiza el cebador de ARN se conoce como primasa.

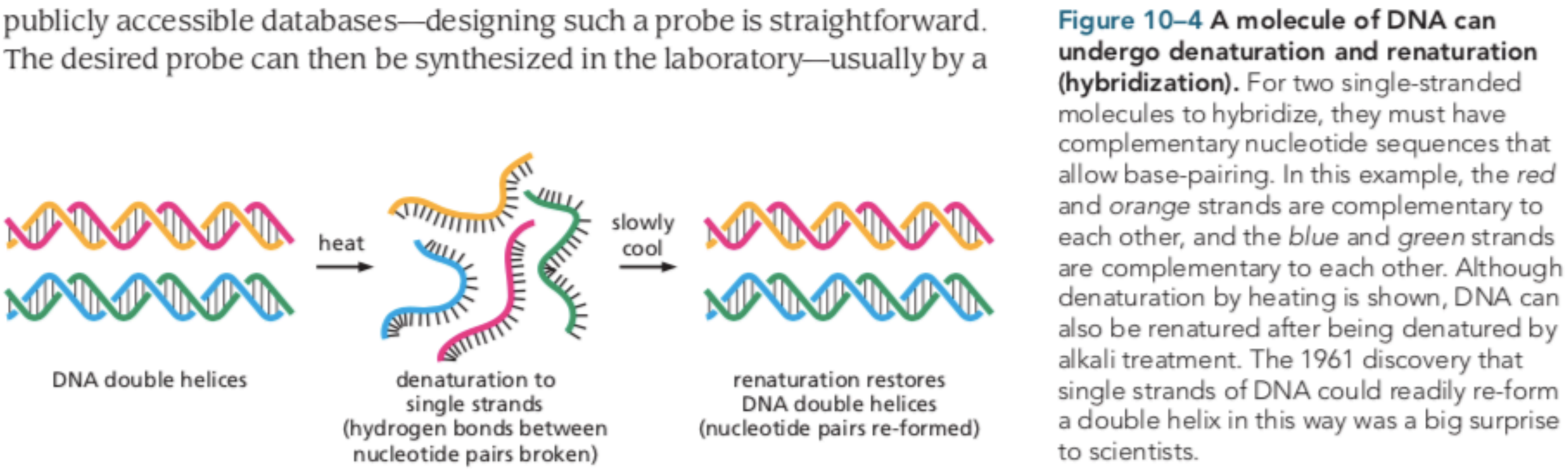
Comentario "al costado".



Seguimos por “el costado”.

La hibridación permite detectar secuencias específicas.

Under normal conditions, the two strands of a DNA double helix are held together by hydrogen bonds between the complementary base pairs. But these relatively weak, noncovalent bonds can be fairly easily broken. Such *DNA denaturation* will release the two strands from each other, but does not break the covalent bonds that link together the nucleotides within each strand. Perhaps the simplest way to achieve this separation involves heating the DNA to around 90oC. When the conditions are reversed—by slowly lowering the temperature—the complementary strands will readily come back together to re-form a double helix. This hybridization, or *DNA renaturation*, is driven by the re-formation of the hydrogen bonds between complementary base pairs.



publicly accessible databases—designing such a probe is straightforward. The desired probe can then be synthesized in the laboratory—usually by a

heat
slowly cool
DNA double helices
denaturation to single strands (hydrogen bonds between nucleotide pairs broken)
renaturation restores DNA double helices (nucleotide pairs re-formed)

Figure 10–4 A molecule of DNA can undergo denaturation and renaturation (hybridization). For two single-stranded molecules to hybridize, they must have complementary nucleotide sequences that allow base-pairing. In this example, the *red* and *orange* strands are complementary to each other, and the *blue* and *green* strands are complementary to each other. Although denaturation by heating is shown, DNA can also be renatured after being denatured by alkali treatment. The 1961 discovery that single strands of DNA could readily re-form a double helix in this way was a big surprise to scientists.

Seguimos por “el costado”.

La hibridación permite detectar secuencias específicas.

This fundamental capacity of a single-stranded nucleic acid molecule, either DNA or RNA, to form a double helix with a single-stranded molecule of a complementary sequence provides a very powerful and sensitive technique for detecting specific nucleotide sequences in both DNA and RNA. Today, one simply designs a short, single-stranded *DNA probe* that is complementary to the nucleotide sequence of interest. Because the nucleotide sequences of so many genomes are known—and are stored in publicly accessible databases—designing such a probe is straightforward. The desired probe can then be synthesized in the laboratory—usually by a commercial organization or a centralized academic facility. Such probes carry a fluorescent or radioactive label to facilitate detection of the nucleotide sequence to which they bind.

Once a suitable probe has been obtained, it can be used in a variety of situations to search for nucleic acids with a complementary sequence.

Clonado de ADN usando PCR

PCR is an iterative process in which the cycle of DNA amplification is repeated dozens of times. At the start of each cycle, the two strands of the double- stranded DNA template are separated and a unique primer is annealed to each. DNA polymerase is then allowed to replicate each strand independently. In subsequent cycles, all the newly synthesized DNA molecules produced by the polymerase serve as templates for the next round of replication. Through this iterative amplification process, many copies of the original sequence can be made—billions after about 20 to 30 cycles. Each cycle takes only about five minutes, and automation of the whole procedure enables cell-free cloning of a DNA fragment in a few hours. The original template for PCR can be either DNA or RNA.

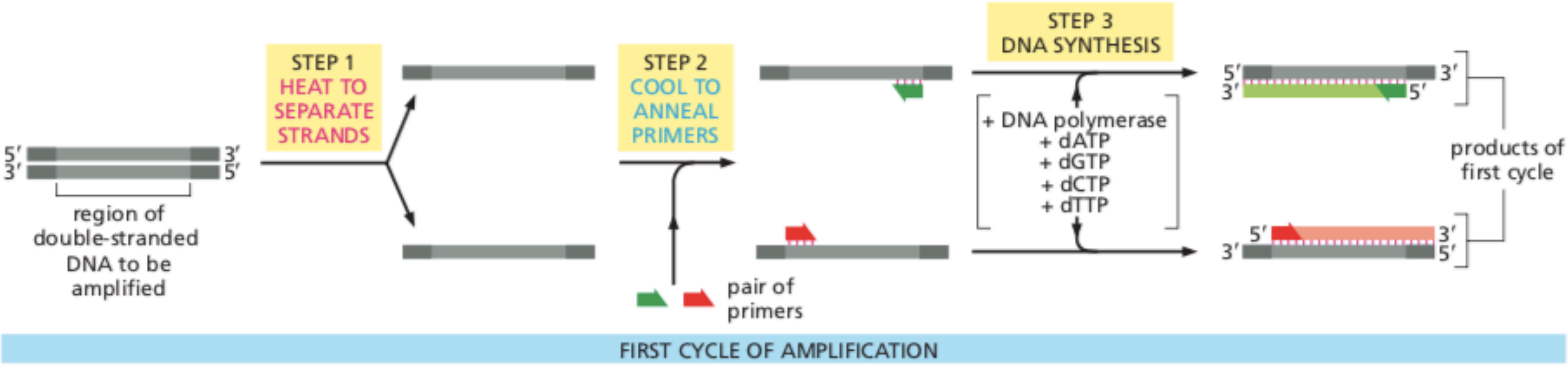

STEP 1
HEAT TO SEPARATE STRANDS
STEP 2
COOL TO ANNEAL PRIMERS
STEP 3
DNA SYNTHESIS
5′
3′
3′
5′
region of double-stranded DNA to be amplified
+ DNA polymerase
+ dATP
+ dGTP
+ dCTP
+ dTTP
pair of primers
products of first cycle
FIRST CYCLE OF AMPLIFICATION

Figure 10–14 A pair of PCR primers directs the amplification of a desired segment of DNA in a test tube. Each cycle of PCR includes three steps: (1) The double-stranded DNA is heated briefly to separate the two strands. (2) The DNA is exposed to a large excess of a pair of specific primers—designed to bracket the region of DNA to be amplified—and the sample is cooled to allow the primers to hybridize to complementary sequences in the two DNA strands. (3) This mixture is incubated with DNA polymerase and the four deoxyribonucleoside triphosphates so that DNA can be synthesized, starting from the two primers. The cycle can then be repeated by reheating the sample to separate the newly synthesized DNA strands (see Figure 10–15).

The technique depends on the use of a special DNA polymerase isolated from a thermophilic bacterium; this polymerase is stable at much higher temperatures than eukaryotic DNA polymerases, so it is not denatured by the heat treatment shown in step 1. The enzyme therefore does not have to be added again after each cycle.

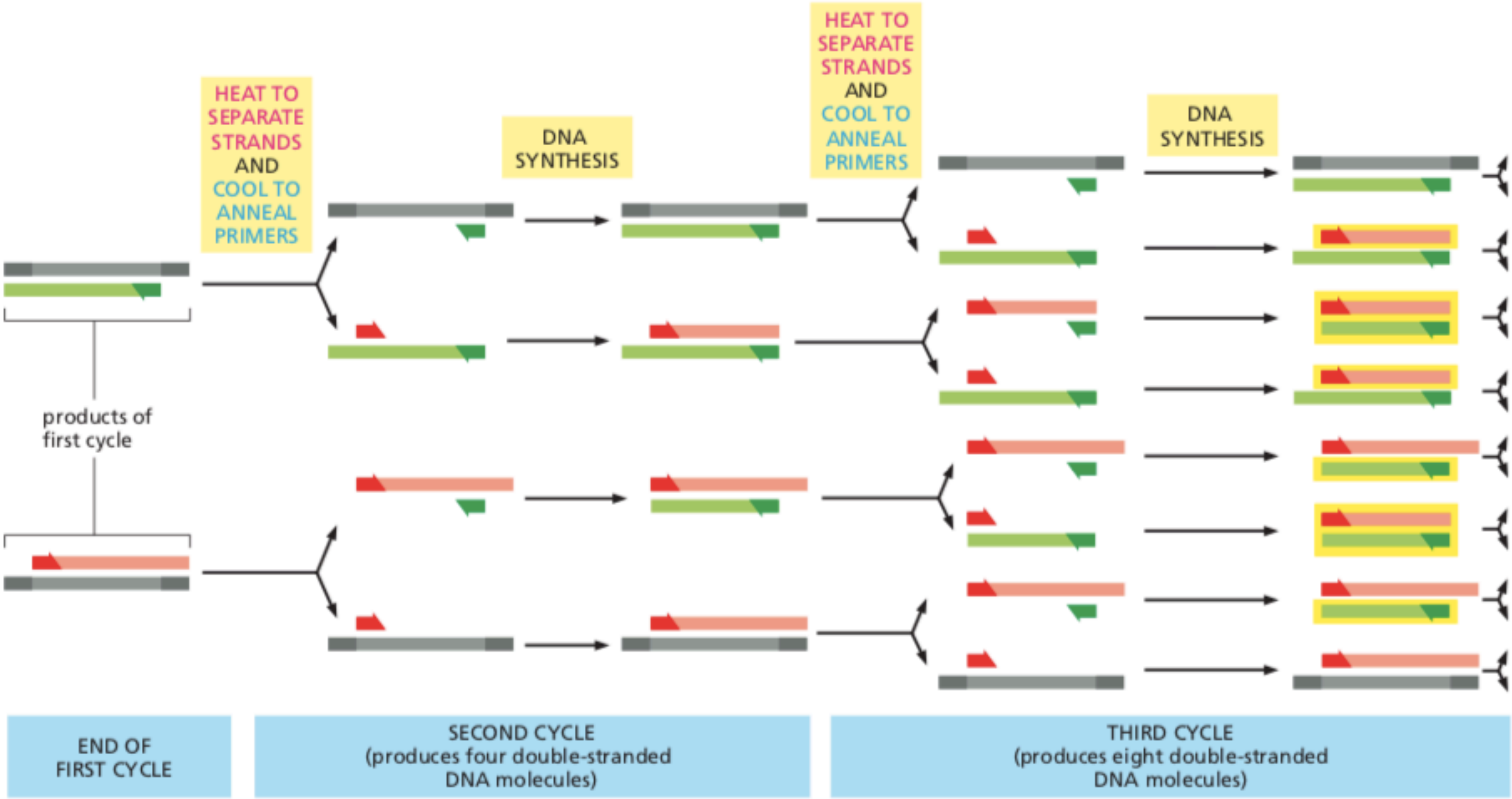
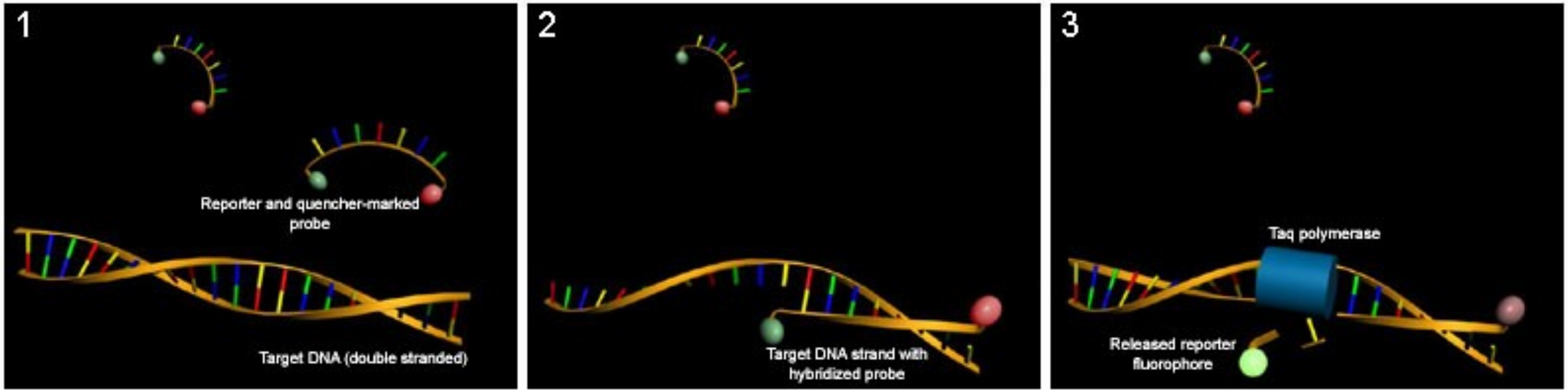

HEAT TO SEPARATE STRANDS AND COOL TO ANNEAL PRIMERS
DNA SYNTHESIS
HEAT TO SEPARATE STRANDS AND COOL TO ANNEAL PRIMERS
DNA SYNTHESIS
products of first cycle
END OF FIRST CYCLE
SECOND CYCLE (produces four double-stranded DNA molecules)
THIRD CYCLE (produces eight double-stranded DNA molecules)

Figure 10–15 PCR uses repeated rounds of strand separation, hybridization, and synthesis to amplify DNA. As the procedure outlined in Figure 10–14 is repeated, all the newly synthesized fragments serve as templates in their turn. Because the polymerase and the primers remain in the sample after the first cycle, PCR involves simply heating and then cooling the same sample, in the same test tube, again and again. Each cycle doubles the amount of DNA synthesized in the previous cycle, so that within a few cycles, the predominant DNA is identical to the sequence bracketed by and including the two primers in the original template. In the example illustrated here, three cycles of reaction produce 16 DNA chains, 8 of which (boxed in *yellow*) correspond exactly to one or the other strand of the original bracketed sequence. After four more cycles, 240 of the 256 DNA chains will correspond exactly to the original sequence, and after several more cycles, essentially all of the DNA strands will be this length. The whole procedure is shown in Movie 10.1.


1
Reporter and quencher-marked probe
Target DNA (double stranded)
2
Target DNA strand with hybridized probe
3
Taq polymerase
Released reporter fluorophore

CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=337932

(1) In intact probes, reporter fluorescence is quenched. (2) Probes and the complementary DNA strand are hybridized and reporter fluorescence is still quenched. (3) During PCR, the probe is degraded by the Taq polymerase and the fluorescent reporter released.

Fluorescence reporter probes detect only the DNA containing the sequence complementary to the probe. Using different-coloured labels, fluorescent probes can be used in multiplex assays for monitoring several target sequences in the same tube.
There is one method that relies on a DNA-based probe with a fluorescent reporter at one end and a quencher of fluorescence at the opposite end of the probe. The close proximity of the reporter to the quencher prevents detection of its fluorescence; breakdown of the probe by the 5' to 3' exonuclease activity of the polymerase breaks the reporter-quencher proximity and thus allows unquenched emission of fluorescence, which can be detected after excitation with a laser. An increase in the product targeted by the reporter probe at each PCR cycle therefore causes a proportional increase in fluorescence due to the breakdown of the probe and release of the reporter.

Another method uses non-specific fluorescent dyes that intercalate with any double-stranded DNA.

How is the presence of the virus detected in the samples?

By Madprime - Own work, CC0

The viral RNA is detected using PCR

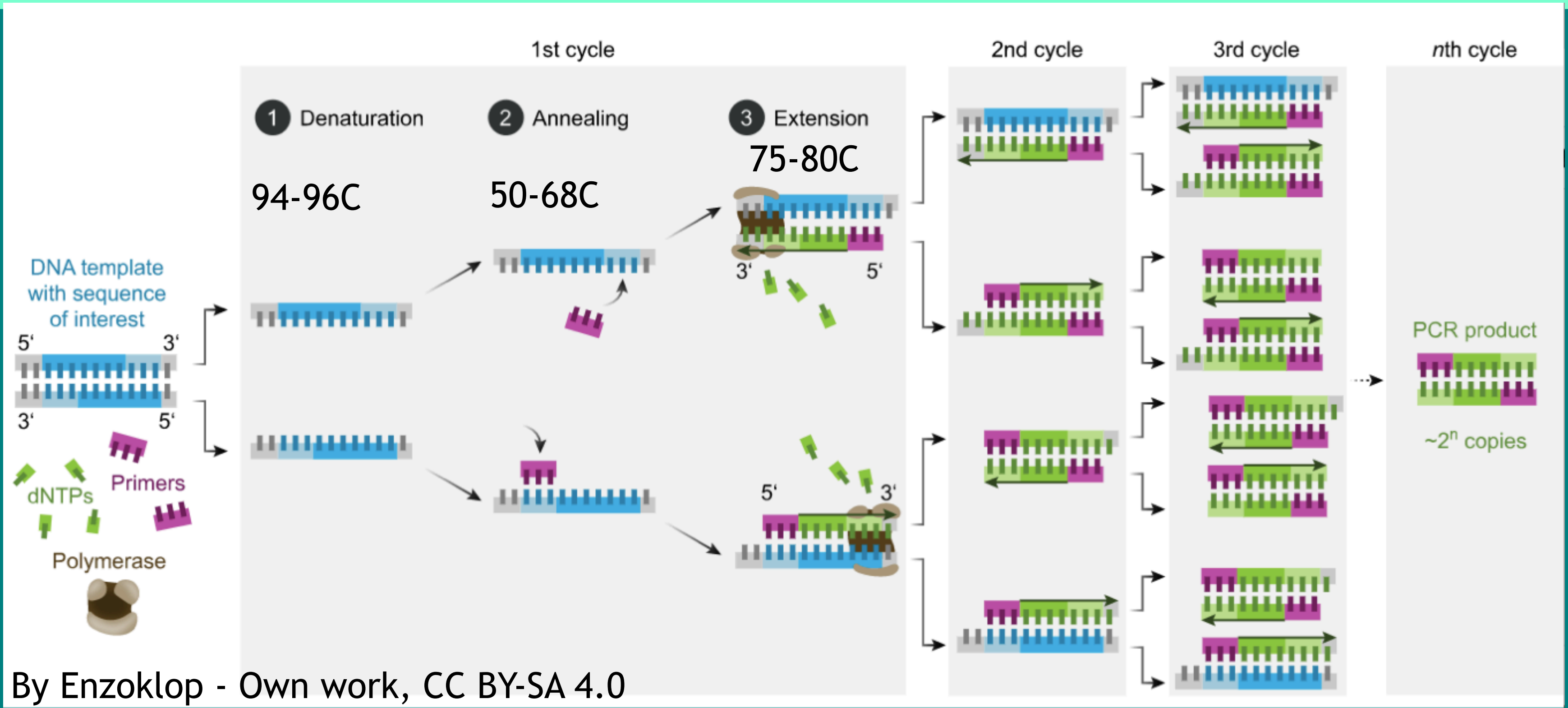
Transverse transcription of RNA into DNA.

DNA amplification in thermal cycles with DNA polymerase that “builds” new DNA strands.

By Karl Mumm - Own work, CC BY-SA 3.0

It does it by elongating a strand of a double-stranded region: PCR needs primers that are complementary to the ends of each of the strands of the DNA target

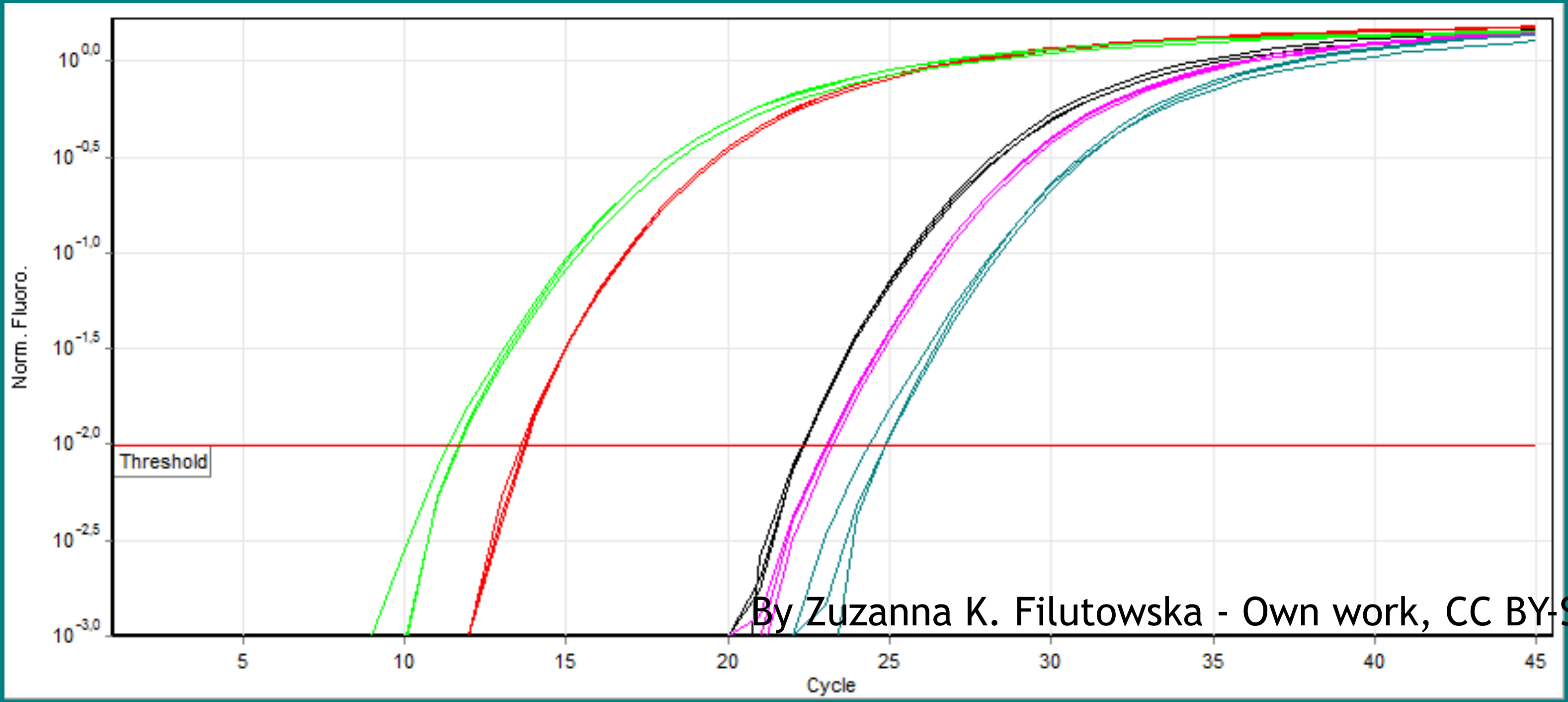
The process requires multiple (n) cycles. Final number of copies of target 2^n


1st cycle
2nd cycle
3rd cycle
nth cycle
① Denaturation
② Annealing
③ Extension
94-96C
50-68C
75-80C
DNA template with sequence of interest
5'
3'
3'
5'
dNTPs
Primers
Polymerase
3'
5'
5'
3'
PCR product
~2ⁿ copies

By Enzoklop - Own work, CC BY-SA 4.0

How do we “see” the copies of the target? Fluorescence

Real Time PCR:
monitors the amplification of the target in real time, not at the end.



By Zuzanna K. Filutowska - Own work, CC BY-SA 3.0

The detection can be done with non-specific fluorescent dyes that intercalate with any double-stranded DNA or with sequence-specific DNA probes.

The target is assumed to be in the sample if the observed fluorescence crosses a threshold. A calibration is necessary to quantify the amount of target (more or less target if threshold reached after fewer or more cycles).

In Situ Hybridization can reveal When and Where a gene is expressed

To see where a particular RNA is made, investigators use a technique called *in situ* hybridization (from the Latin *in situ*, "in place"), which allows a specific nucleic acid sequence— either DNA or RNA—to be visualized in its normal location.

In situ hybridization uses single-stranded DNA or RNA probes, labeled with either fluorescent dyes or radioactive isotopes, to detect complementary nucleic acid sequences within a tissue, a cell or even an isolated chromosome. *In situ* hybridization is frequently used to study the expression patterns of a particular gene or collection of genes in an adult or developing tissue.

tein-coding gene to drive the expression of some type of

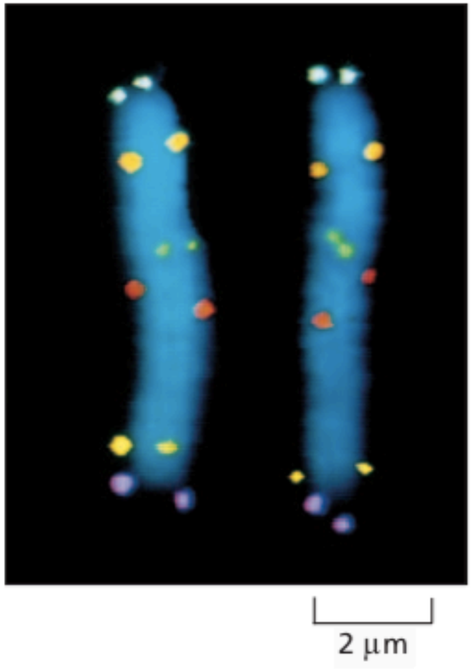
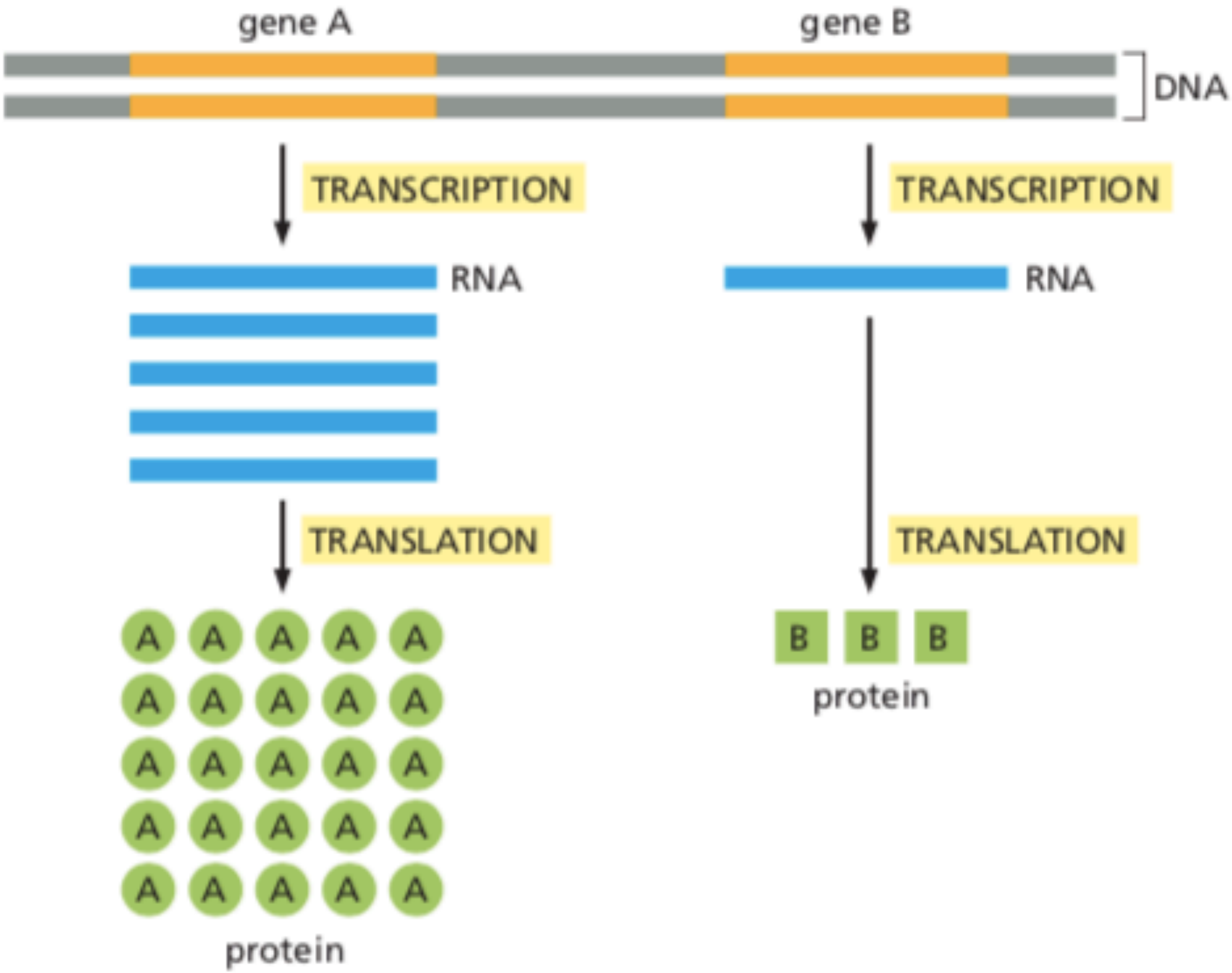


Figure 10–29 ***In situ*** **hybridization can be used to locate genes on isolated chromosomes.** Here, six different DNA probes have been used to mark the locations of their respective nucleotide sequences on human Chromosome 5 isolated from a mitotic cell in metaphase (see Figure 5–16 and Panel 18–1, pp. 622–623). The DNA probes have been labeled with different chemical groups and are detected using fluorescent antibodies specific for those groups. Both the maternal and paternal copies of Chromosome 5 are shown, aligned side by side. Each probe produces two dots on each chromosome because chromosomes undergoing mitosis have already replicated their DNA; therefore, each chromosome contains two identical DNA helices. The technique employed here is nicknamed FISH, for *fluorescence* in situ *hybridization*. (Courtesy of David C. Ward.)

Volvemos al ARN y a la transcripción.

Dos genes distintos pueden transcribirse a tasas muy distintas entre sí. Esto hace que las tasas a las que se producen la proteínas codificadas por los genes sean distintas entre sí también.


gene A
gene B
DNA
TRANSCRIPTION
TRANSCRIPTION
RNA
RNA
TRANSLATION
TRANSLATION
A A A A A
A A A A A
A A A A A
A A A A A
A A A A A
protein
B B B
protein

La expresión génica es un proceso complejo mediante el cual las células dirigen selectivamente la síntesis de los muchos miles de proteínas y ARN codificados en su genoma. Pero, ¿cómo coordinan y controlan las células un proceso tan intrincado, y cómo especifica una célula individual cuál de sus genes expresar? Esta decisión es un problema especialmente importante para los animales porque, a medida que se desarrollan, sus células se vuelven altamente especializadas, produciendo finalmente una serie de células musculares, nerviosas y sanguíneas, junto con los cientos de otros tipos de células que se observan en el adulto. Tal diferenciación celular surge porque las células producen y acumulan diferentes conjuntos de moléculas de ARN y proteínas: es decir, expresan diferentes genes.

El inicio de la transcripción es un proceso especialmente crítico porque es el punto principal en el que la célula selecciona qué proteínas o ARN se van a producir. Para comenzar la transcripción, la ARN polimerasa debe poder reconocer el inicio de un gen y unirse firmemente al ADN en este sitio. La forma en que las ARN polimerasas reconocen el sitio de inicio de la transcripción de un gen difiere algo entre bacterias y eucariotas.

Cuando una ARN polimerasa choca aleatoriamente con una molécula de ADN, la enzima se adhiere débilmente a la doble hélice y luego se desliza rápidamente a lo largo de su longitud. La ARN polimerasa se adhiere firmemente solo después de haber encontrado una región genética llamada **promotor**, que contiene una secuencia específica de nucleótidos que se encuentra inmediatamente aguas arriba del punto de partida para la síntesis de ARN. Una vez unida fuertemente a esta secuencia, la ARN polimerasa abre la doble hélice inmediatamente delante del promotor para exponer los nucleótidos en cada hebra de un tramo corto de ADN. Una de las dos hebras de ADN expuestas actúa como plantilla para el apareamiento de bases complementarias con los ribonucleósidos trifosfatos entrantes, dos de los cuales son unidos por la polimerasa para comenzar la síntesis de la cadena de ARN. El alargamiento de la cadena continúa hasta que la enzima encuentra una segunda señal en el ADN, el sitio de stop, donde la polimerasa se detiene y libera tanto la plantilla de ADN como la transcripción de ARN recién creada. Esta secuencia de terminación está contenida en el gen y se transcribe en el extremo $3'$ del ARN recién creado. Debido a que la polimerasa debe unirse fuertemente antes de que pueda comenzar la transcripción, un segmento de ADN se transcribirá solo si está precedido por un promotor. Esto asegura que las porciones de una molécula de ADN que contienen un gen se transcriban en ARN.

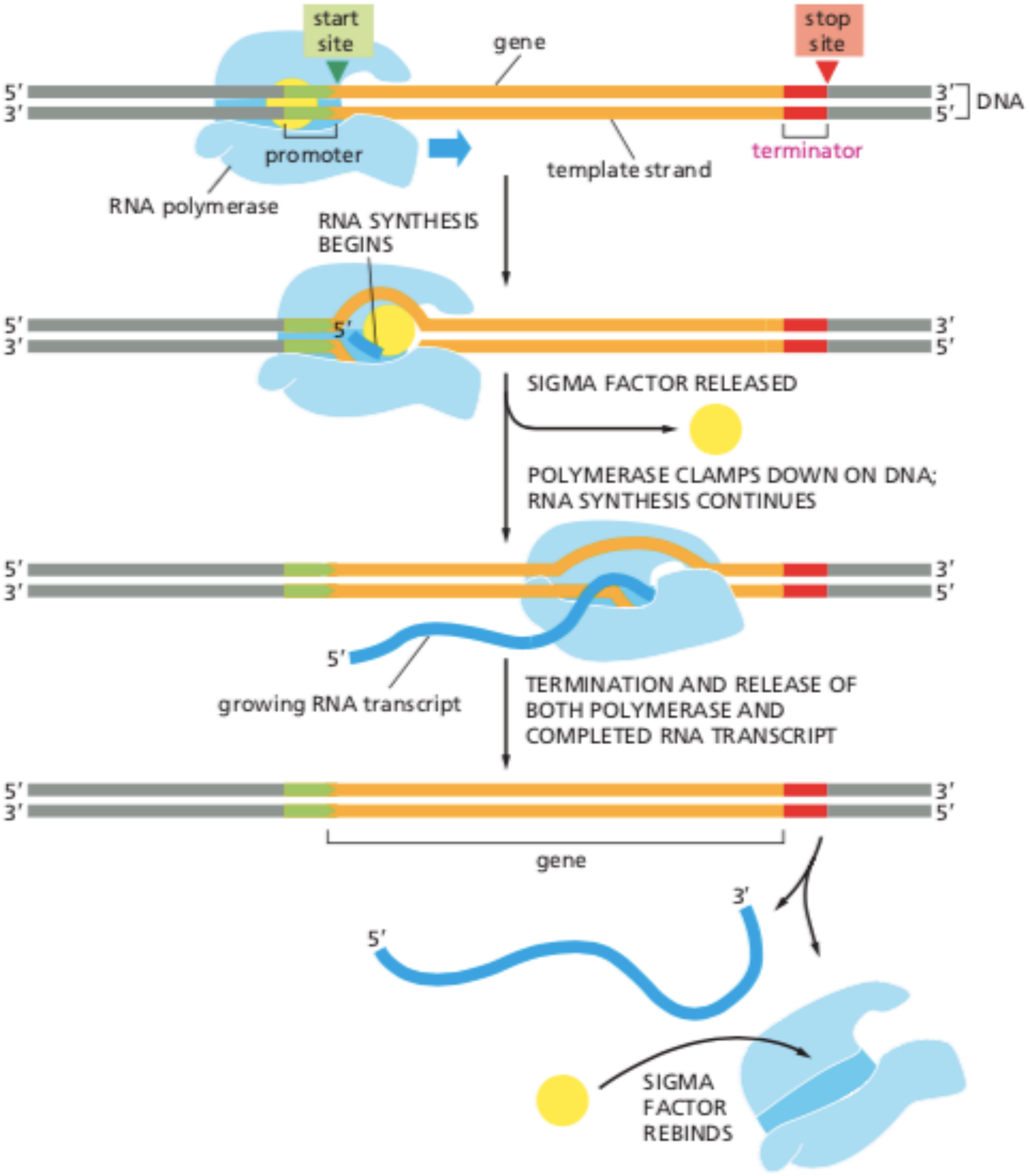

start site
gene
stop site
5′
3′
3′
5′
DNA
promoter
template strand
terminator
RNA polymerase
RNA SYNTHESIS BEGINS
5′
5′
3′
3′
5′
SIGMA FACTOR RELEASED
POLYMERASE CLAMPS DOWN ON DNA; RNA SYNTHESIS CONTINUES
5′
3′
3′
5′
5′
growing RNA transcript
TERMINATION AND RELEASE OF BOTH POLYMERASE AND COMPLETED RNA TRANSCRIPT
5′
3′
3′
5′
gene
3′
5′
SIGMA FACTOR REBINDS

Figure 7–9 Signals in the nucleotide sequence of a gene tell bacterial RNA polymerase where to start and stop transcription. Bacterial RNA polymerase (*light blue*) contains a subunit called sigma factor (*yellow*) that recognizes the promoter of a gene (*green*). Once transcription has begun, sigma factor is released, and the polymerase moves forward and continues synthesizing the RNA. Chain elongation continues until the polymerase encounters a sequence in the gene called the terminator (*red*). There the enzyme halts and releases both the DNA template and the newly made RNA transcript. The polymerase then reassociates with a free sigma factor and searches for another promoter to begin the process again.

En las bacterias, es una subunidad de la ARN polimerasa, el factor sigma (σ), el principal responsable de reconocer la secuencia promotora en el ADN. Pero, ¿cómo puede este factor “ver” al promotor si los pares de bases en cuestión están situados en el interior de la doble hélice del ADN? Resulta que cada base presenta características únicas en el exterior de la doble hélice, lo que permite que el factor sigma encuentre la secuencia promotora sin tener que separar las hebras de ADN entrelazadas.

Ejemplos de secuencias que corresponden a promotores y sitios de STOP en bacterias

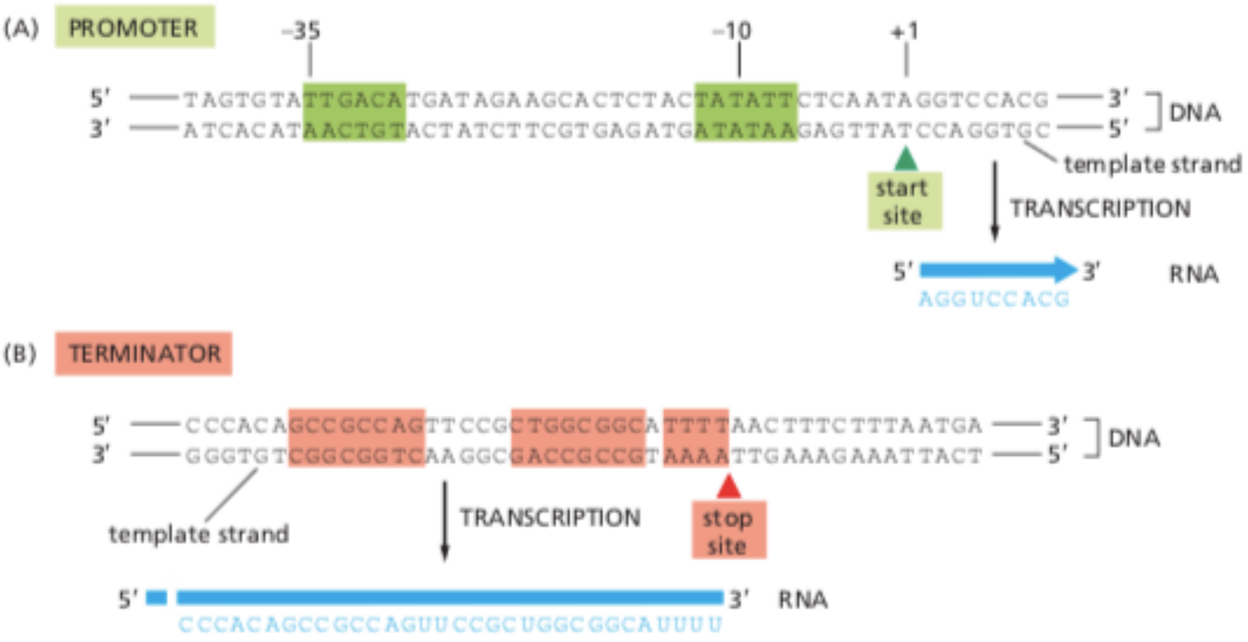

(A)
PROMOTER
−35
−10
+1
5′ TAGTGTATTGACATGATAGAAGCACTCTACTATATTCTCAATAGGTCCACG 3′
3′ ATCACATAACTGTACTATCTTCGTGAGATGATATAAGAGTTATCCAGGTGC 5′
DNA
template strand
start site
TRANSCRIPTION
5′ 3′ RNA
AGGUCCACG
(B)
TERMINATOR
5′ CCCACAGCCGCCAGTTCCGCTGGCGGCATTTTAACTTTCTTTAATGA 3′
3′ GGGTGTCGGCGGTCAAGGCGACCGCCGTAAAATTGAAAGAAATTACT 5′
DNA
template strand
TRANSCRIPTION
stop site
5′ 3′ RNA
CCCACAGCCGCCAGUUCCGCUGGCGGCAUUUU

Figure 7–10 Bacterial promoters and terminators have specific nucleotide sequences that are recognized by RNA polymerase. (A) The *green*-shaded regions represent the nucleotide sequences that specify a promoter. The numbers above the DNA indicate the positions of nucleotides counting from the first nucleotide transcribed, which is designated +1. The polarity of the promoter orients the polymerase and determines which DNA strand is transcribed. All bacterial promoters contain DNA sequences at −10 and −35 that closely resemble those shown here. (B) The *red*-shaded regions represent sequences in the gene that signal the RNA polymerase to terminate transcription. Note that the regions transcribed into RNA contain the terminator but not the promoter nucleotide sequences. By convention, the sequence of a gene is that of the non-template strand, as this strand has the same sequence as the transcribed RNA (with T substituting for U).

El siguiente problema al que se enfrenta una ARN polimerasa es determinar cuál de las dos cadenas de ADN utilizar como plantilla para la transcripción: cada cadena tiene una secuencia de nucleótidos diferente y produciría una transcripción de ARN diferente. El secreto radica en la estructura del propio promotor. Cada promotor tiene una cierta polaridad: contiene dos secuencias de nucleótidos diferentes aguas arriba del sitio de inicio de la transcripción que posicionan la ARN polimerasa, asegurando que se una al promotor en una sola orientación.Debido a que la polimerasa solo puede sintetizar ARN en la dirección 5 'a 3' una vez que la enzima está unida, debe usar la hebra de ADN orientada en la dirección 3 'a 5' como plantilla.

Células eucariotas

En estas células hay distintos tipos de ARN polimerasas

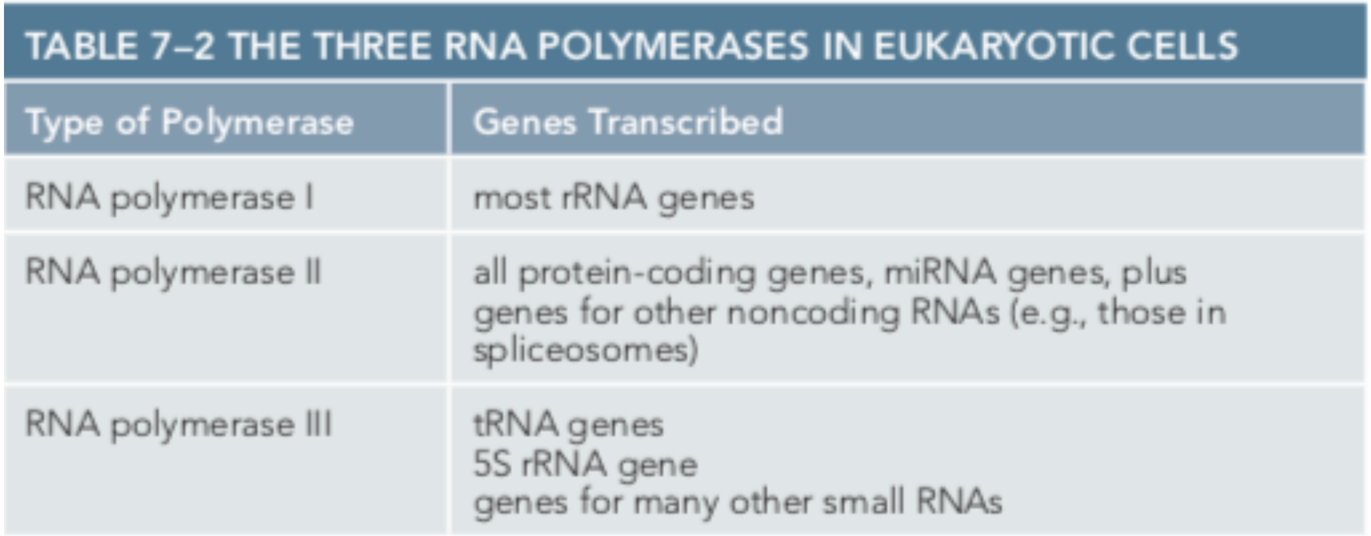


TABLE 7–2 THE THREE RNA POLYMERASES IN EUKARYOTIC CELLS

Type of Polymerase	Genes Transcribed
RNA polymerase I	most rRNA genes
RNA polymerase II	all protein-coding genes, miRNA genes, plus genes for other noncoding RNAs (e.g., those in spliceosomes)
RNA polymerase III	tRNA genes 5S rRNA gene genes for many other small RNAs

Una segunda diferencia respecto de las bacterias es que, mientras que la ARN polimerasa bacteriana (junto con su subunidad sigma) es capaz de iniciar la transcripción por sí misma, las ARN polimerasas eucariotas requieren la ayuda de un gran conjunto de proteínas accesorias. Entre ellas, las principales son los **factores de transcripción** generales, que deben ensamblarse en cada promotor, junto con la polimerasa, antes de que la polimerasa pueda comenzar la transcripción.

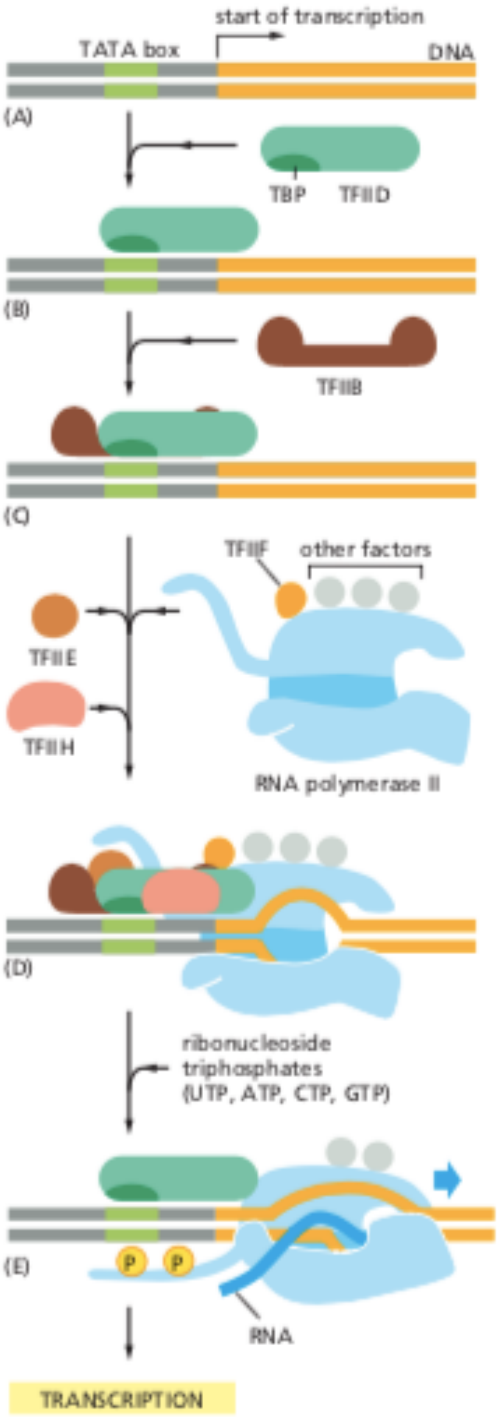

start of transcription
TATA box
DNA
(A)
TBP
TFIID
(B)
TFIIB
(C)
TFIIF
other factors
TFIIE
TFIIH
RNA polymerase II
(D)
ribonucleoside
triphosphates
(UTP, ATP, CTP, GTP)
(E)
P
P
RNA
TRANSCRIPTION

Figure 7–12 To begin transcription, eukaryotic RNA polymerase II requires a set of general transcription factors. These transcription factors are called TFIIB, TFIID, and so on. (A) Many eukaryotic promoters contain a DNA sequence called the TATA box. (B) The TATA box is recognized by a subunit of the general transcription factor TFIID, called the TATA-binding protein (TBP). For simplicity, the DNA distortion produced by the binding of the TBP (see Figure 7–13) is not shown. (C) The binding of TFIID enables the adjacent binding of TFIIB. (D) The rest of the general transcription factors, as well as the RNA polymerase itself, assemble at the promoter. (E) TFIIH then pries apart the double helix at the transcription start point, using the energy of ATP hydrolysis, which exposes the template strand of the gene (not shown). TFIIH also phosphorylates RNA polymerase II, releasing the polymerase from most of the general transcription factors, so it can begin transcription. The site of phosphorylation is a long polypeptide "tail" that extends from the polymerase.

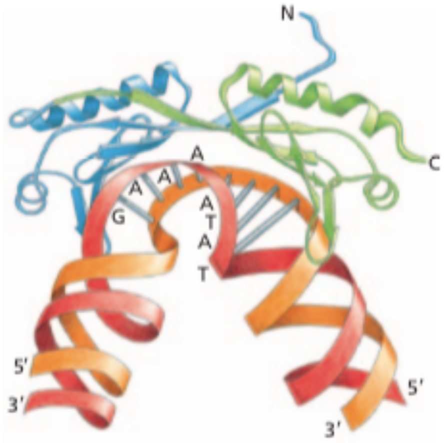

N
C
A
A
A
A
G
T
A
T
5′
3′
5′
3′

Figure 7–13 TATA-binding protein (TBP) binds to the TATA box (indicated by letters) and bends the DNA double helix. The unique distortion of DNA caused by TBP, which is a subunit of TFIID (see Figure 7–12), helps attract the other general transcription factors. TBP is a single polypeptide chain that is folded into two very similar domains (*blue* and *green*). The protein sits atop the DNA double helix like a saddle on a bucking horse (**Movie 7.4**). (Adapted from J.L. Kim et al., *Nature* 365:520–527, 1993. With permission from Macmillan Publishers Ltd.)

Una tercera característica distintiva de la transcripción en eucariotas es que los mecanismos que controlan su inicio son mucho más elaborados que los de los procariotas. En las bacterias, los genes tienden a estar muy cerca unos de otros en el ADN, con sólo tramos muy cortos de ADN no transcrito entre ellos. Pero en plantas y animales, incluidos los humanos, los genes individuales se distribuyen a lo largo del ADN, con tramos de hasta 100.000 pares de nucleótidos entre un gen y el siguiente. Esta arquitectura permite que un solo gen sea controlado por una gran variedad de secuencias reguladoras de ADN esparcidas a lo largo del ADN, y permite a los eucariotas participar en formas más complejas de regulación transcripcional que las bacterias.

Por último, pero no menos importante, el inicio de la transcripción eucariota debe tener en cuenta el empaquetado del ADN en nucleosomas y formas más compactas de estructura de cromatina.

Aunque el principio que rige la transcripción del ADN en ARN es el mismo en todos los organismos, los pasos que deben darse para producir las proteínas correspondientes difiere mucho entre bacterias y eucariotas. El ADN bacteriano se encuentra expuesto directamente al citoplasma, que contiene los ribosomas en los que tiene lugar la síntesis de proteínas. A medida que comienza a sintetizarse una molécula de ARNm en una bacteria, los ribosomas se unen inmediatamente al extremo 5’ libre del transcrito de ARN y comienzan a traducirlo en proteína.

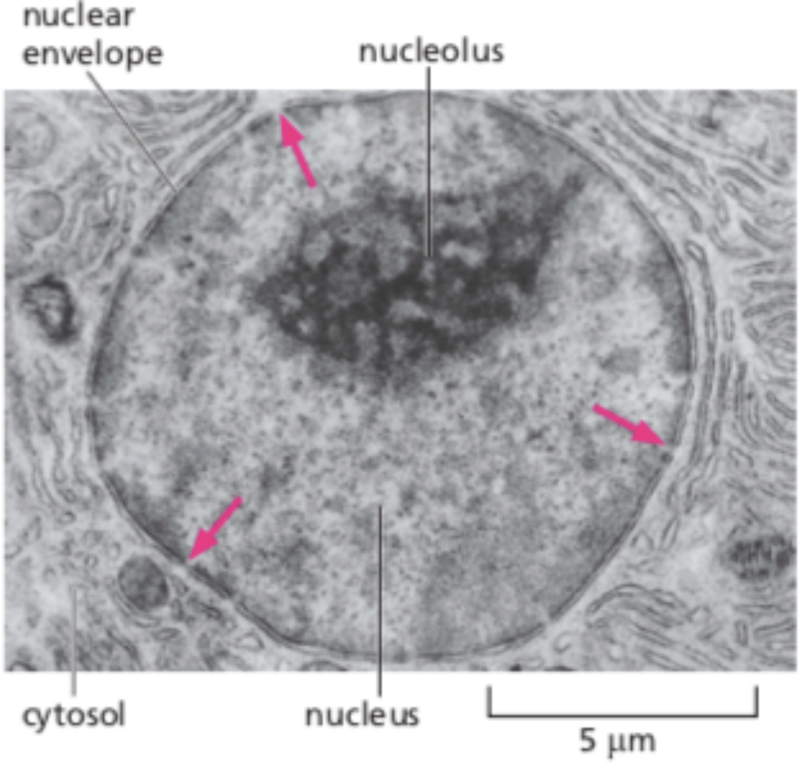


Figure 7–14 Before they can be translated, mRNA molecules made in the nucleus must be exported to the cytosol via pores in the nuclear envelope (*red arrows*). Shown here is a section of a liver cell nucleus. The nucleolus is where ribosomal RNAs are synthesized and combined with proteins to form ribosomes, which are then exported to the cytoplasm. (From D.W. Fawcett, A Textbook of Histology, 11th ed. Philadelphia: Saunders, 1986. With permission from Elsevier.)

En las células eucariotas el ADN está encerrado dentro del núcleo. Allí tiene lugar la transcripción, pero la síntesis de proteínas tiene lugar en los ribosomas del citoplasma. Por tanto, antes de que un ARNm eucariota pueda traducirse en proteína, debe ser transportado fuera del núcleo a través de pequeños poros en la envoltura nuclear. Antes de que pueda exportarse al citosol, un ARN eucariota debe pasar por varios pasos de procesamiento del ARN. Estos pasos tienen lugar a medida que se sintetiza el ARN. Las enzimas responsables del procesamiento del ARN viajan sobre la cola fosforilada de la ARN polimerasa II eucariota cuando sintetiza una molécula de ARN (figura 7-12) y procesan la transcripción cuando emerge de la polimerasa (figura 7-15).

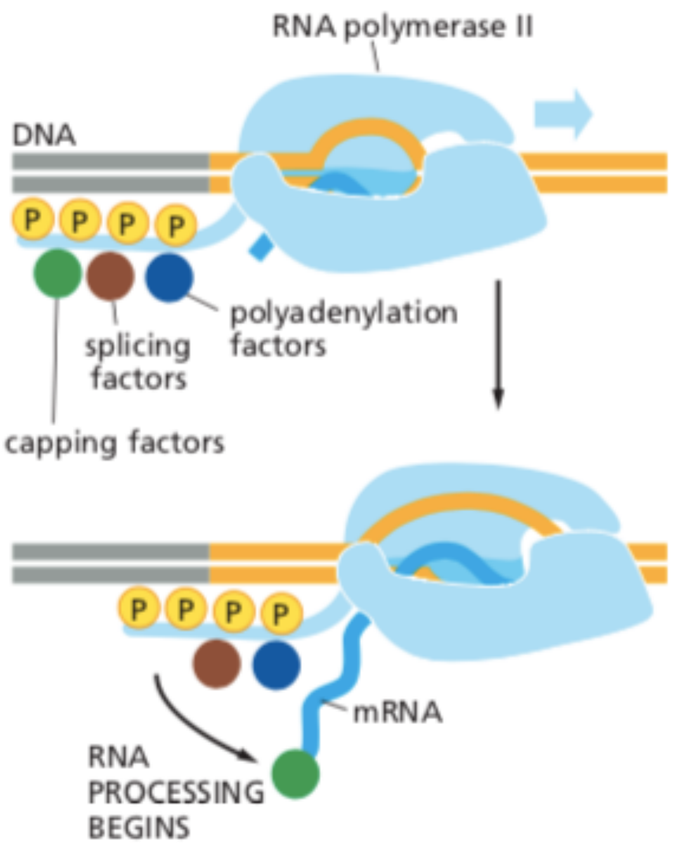


Figure 7–15 Phosphorylation of the tail of RNA polymerase II allows RNA-processing proteins to assemble there. Note that the phosphates shown here are in addition to the ones required for transcription initiation (see Figure 7–12). Capping, polyadenylation, and splicing are all modifications that occur during RNA processing in the nucleus.

Antes de que el ARNm pueda exportarse al citosol, un ARN eucariota debe pasar por varios pasos de procesamiento del ARN. Estos pasos tienen lugar a medida que se sintetiza el ARN. Las enzimas responsables del procesamiento del ARN viajan sobre la cola fosforilada de la ARN polimerasa II eucariota cuando sintetiza una molécula de ARN y procesan la transcripción cuando emerge de la polimerasa.

Distintos tipos de ARN se procesan de diferentes maneras. Dos pasos de procesamiento, “capping” y poliadenilación, ocurren solo en las transcripciones de ARN destinadas a convertirse en moléculas de ARNm. Estas modificaciones aumentan la estabilidad de la molécula de ARNm, facilitan su exportación al citoplasma y, en general, marcan la molécula de ARN como ARNm. También son utilizados por la maquinaria de síntesis de proteínas para asegurarse de que ambos extremos del ARNm estén presentes y que, por lo tanto, el mensaje esté completo antes de que comience la síntesis de proteínas.

En las eucariotas, los genes que codifican proteínas son interrumpidos por secuencias no codificantes llamadas intrones

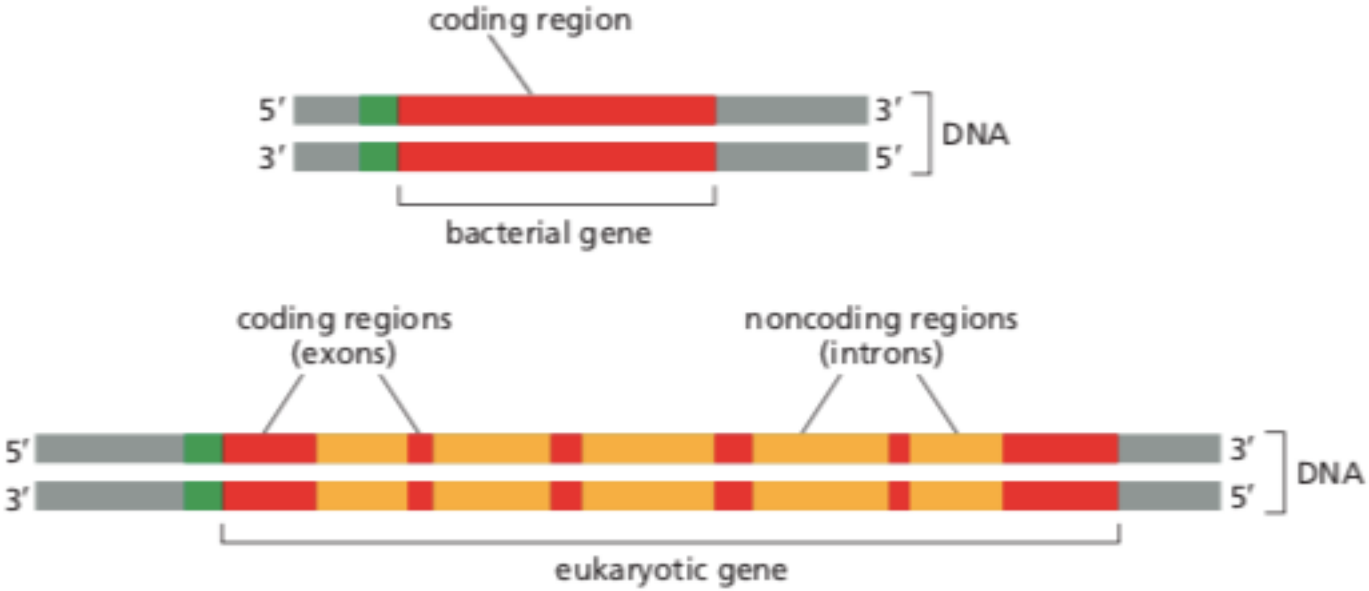
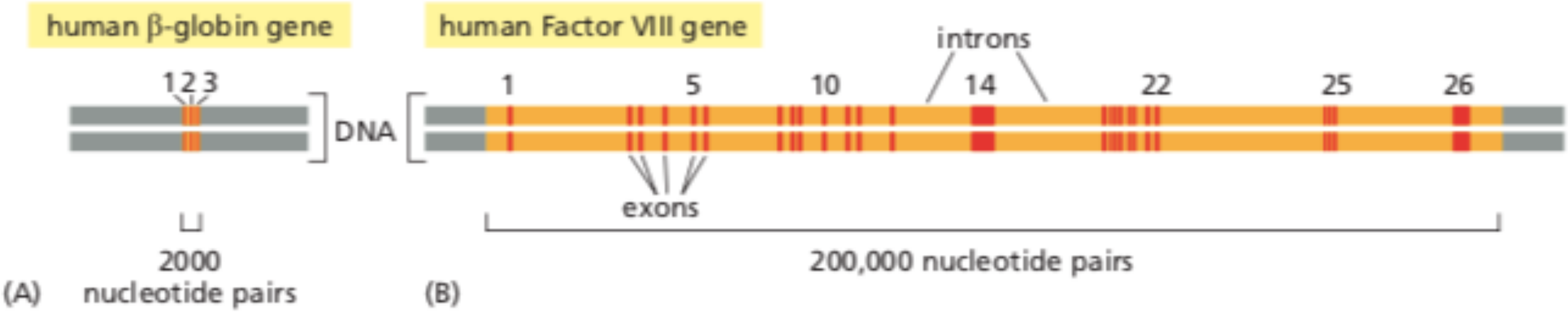

coding region
5′
3′
3′
5′
DNA
bacterial gene
coding regions
(exons)
noncoding regions
(introns)
5′
3′
3′
5′
DNA
eukaryotic gene

Figure 7–17 **Eukaryotic and bacterial genes are organized differently.** A bacterial gene consists of a single stretch of uninterrupted nucleotide sequence that encodes the amino acid sequence of a protein (or more than one protein). In contrast, the protein-coding sequences of most eukaryotic genes (*exons*) are interrupted by noncoding sequences (*introns*). Promoters for transcription are indicated in *green*.

Los fragmentos dispersos de la secuencia codificante, denominados exones, suelen ser más cortos que los intrones y, a menudo, representan sólo una pequeña fracción de la longitud total del gen. Los intrones varían en longitud desde un solo nucleótido hasta más de 10,000. Algunos genes eucariotas que codifican proteínas carecen de intrones o tienen pocos. La mayoría tiene muchos. (los términos "exón" e "intrón" se aplican tanto al ADN como a las secuencias de ARN correspondientes).


human β-globin gene
1 2 3
DNA
2000 nucleotide pairs
(A)
human Factor VIII gene
introns
1
5
10
14
22
25
26
exons
200,000 nucleotide pairs
(B)

Para producir un ARNm en una célula eucariota tanto los intrones como los exones se transcriben en ARN. Después del “capping”, y mientras la ARN polimerasa II continúa transcribiendo el gen, comienza el proceso de “splicing” (corte y empalme) del ARN, en el que los intrones se eliminan del ARN y los exones se unen. Luego se modifican los extremos 5 'y 3’ y el ARN es una molécula de ARNm funcional que puede traducirse.

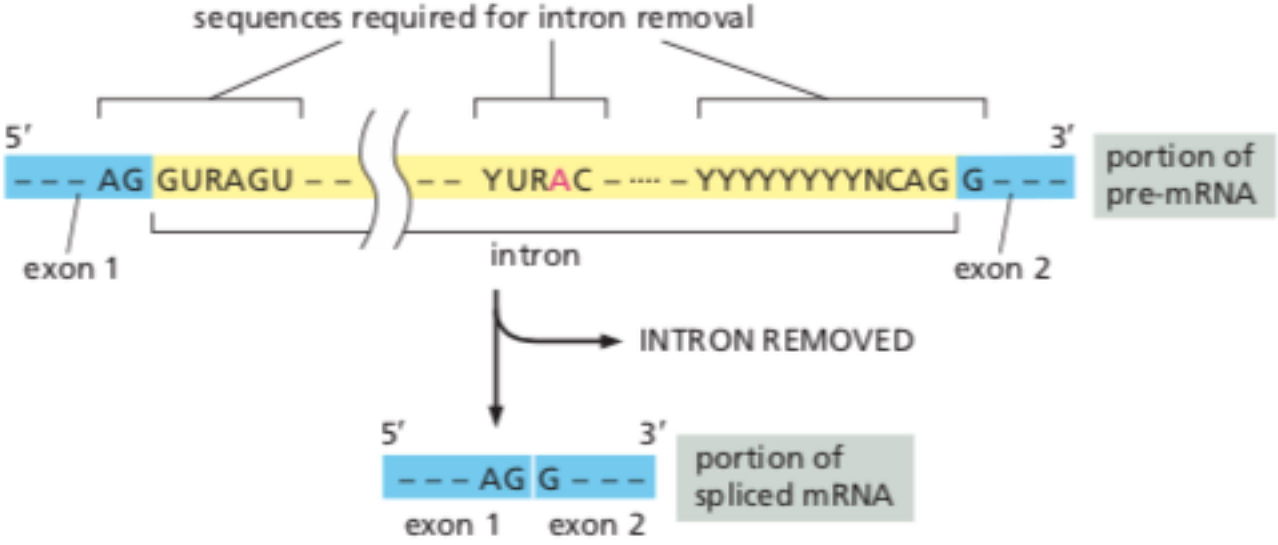

sequences required for intron removal
5′
3′
– – – AG GURAGU – –
– – YURAC – ····· – YYYYYYYYNCAG G – – –
portion of pre-mRNA
exon 1
intron
exon 2
INTRON REMOVED
5′
3′
– – – AG G – – –
portion of spliced mRNA
exon 1
exon 2

Figure 7–19 Special nucleotide sequences in a pre-mRNA transcript signal the beginning and the end of an intron. Only the nucleotide sequences shown are required to remove an intron; the other positions in an intron can be occupied by any nucleotide. The special sequences are recognized primarily by small nuclear ribonucleoproteins (snRNPs), which direct the cleavage of the RNA at the intron–exon borders and catalyze the covalent linkage of the exon sequences. Here, in addition to the standard symbols for nucleotides (A, C, G, U), R stands for either A or G; Y stands for either C or U; N stands for any nucleotide. The A shown in *red* forms the branch point of the lariat produced in the splicing reaction shown in Figure 7–20. The distances along the RNA between the three splicing sequences are highly variable; however, the distance between the branch point and the 5′ splice junction is typically much longer than that between the 3′ splice junction and the branch point (see Figure 7–20). The splicing sequences shown are from humans; similar sequences direct RNA splicing in other eukaryotes.

El splicing del ARN es llevado a cabo, principalmente, por moléculas de ARN llamadas ARN nucleares pequeños (ARNnn).

El splicing del ARN proporciona otra ventaja a los eucariotas, una que probablemente haya sido profundamente importante en la historia evolutiva temprana de los genes. Se cree que la estructura intrón-exón de los genes aceleró la aparición de proteínas nuevas y útiles: las proteínas nuevas parecen haber surgido de la mezcla y ensamblado de diferentes exones de genes preexistentes, muy parecido al montaje de un nuevo tipo de máquina a partir de un kit de componentes funcionales preexistentes. De hecho, muchas proteínas de las células actuales se asemejan a un mosaico compuesto de un conjunto común de fragmentos de proteínas, llamados dominios de proteínas.

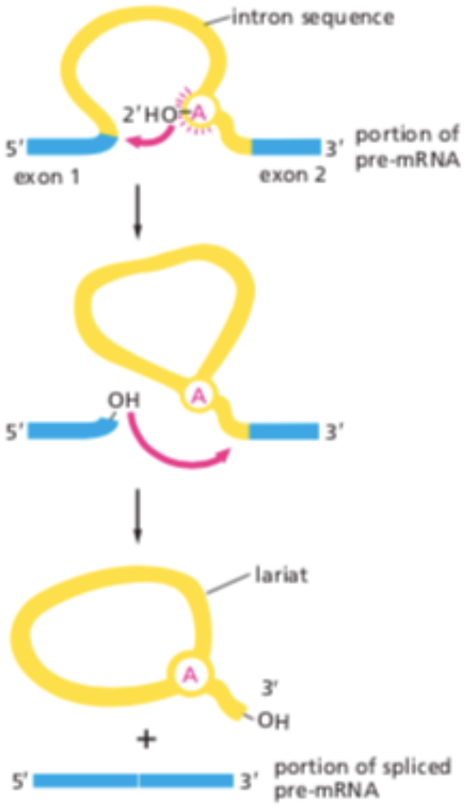


Figure 7–20 An intron in a pre-mRNA molecule forms a branched structure during RNA splicing. In the first step, the branch point adenine (*red* A) in the intron sequence attacks the 5′ splice site and cuts the sugar–phosphate backbone of the RNA at this point (this is the same A highlighted in *red* in Figure 7–19). In this process, the cut 5′ end of the intron becomes covalently linked to the 2′-OH group of the ribose of the A nucleotide to form a branched structure. The free 3′-OH end of the exon sequence then reacts with the start of the next exon sequence, joining the two exons together into a continuous coding sequence and releasing the intron in the form of a lariat structure, which is eventually degraded in the nucleus.

La existencia de exones e intrones tiene otros beneficios importantes. Entre otros, el “splicing” alternativo que permite producir proteínas distintas a partir del mismo gen. Se cree que alrededor del 95% de los genes humanos son sometidos a splicing alternativo. Esto permite a las eucariotas aumentar el ya enorme potencial de codificación de sus genomas.

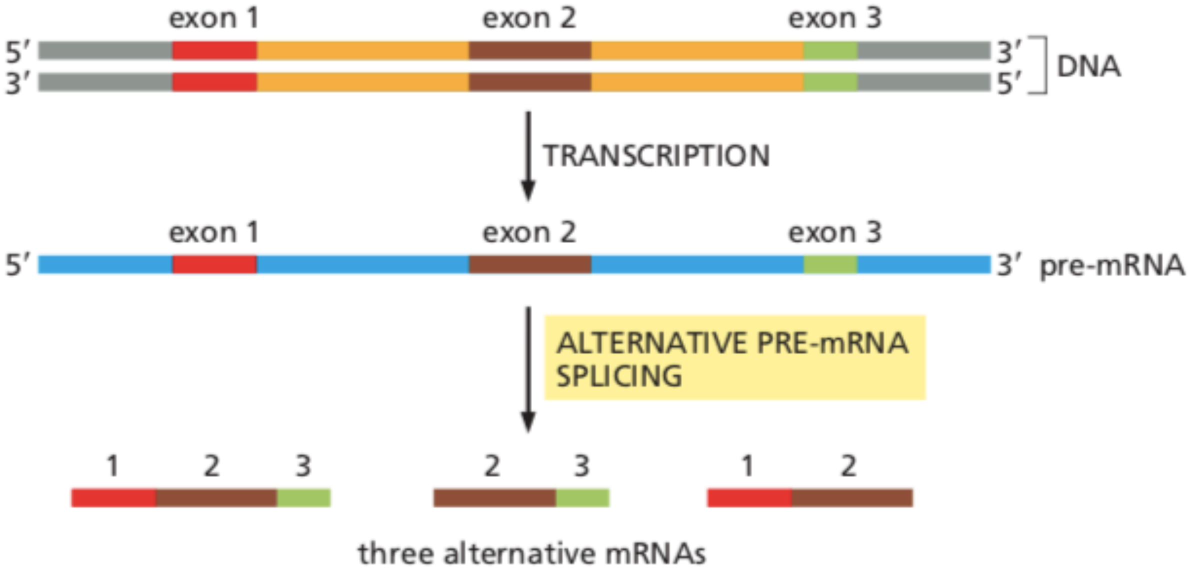

exon 1
exon 2
exon 3
5′
3′
3′
5′
DNA
TRANSCRIPTION
exon 1
exon 2
exon 3
5′
3′
pre-mRNA
ALTERNATIVE PRE-mRNA SPLICING
1 2 3
2 3
1 2
three alternative mRNAs

Figure 7–22 Some pre-mRNAs undergo alternative RNA splicing to produce various mRNAs and proteins from the same gene. Whereas all exons are present in a pre-mRNA, some exons can be excluded from the final mRNA molecule. In this example, three of four possible mRNAs are produced. The 5′ caps and poly-A tails on the mRNAs are not shown.

Camino a la traducción

El transporte de ARNm desde el núcleo al citosol, donde los ARNm se traducen en proteína, es altamente selectivo: solo se exportan ARNm procesados correctamente. Este transporte selectivo está mediado por complejos de poros nucleares, que conectan el nucleoplasma con el citosol y actúan como puertas que controlan qué macromoléculas pueden entrar o salir del núcleo. Para estar "lista para la exportación", una molécula de ARNm debe estar unida a un conjunto apropiado de proteínas, cada una de las cuales reconoce diferentes partes de una molécula de ARNm madura. Este conjunto determina si una molécula de ARNm abandonará el núcleo. Los "ARN de desecho" que quedan en el núcleo se degradan allí y sus componentes básicos de nucleótidos se reutilizan para la transcripción.

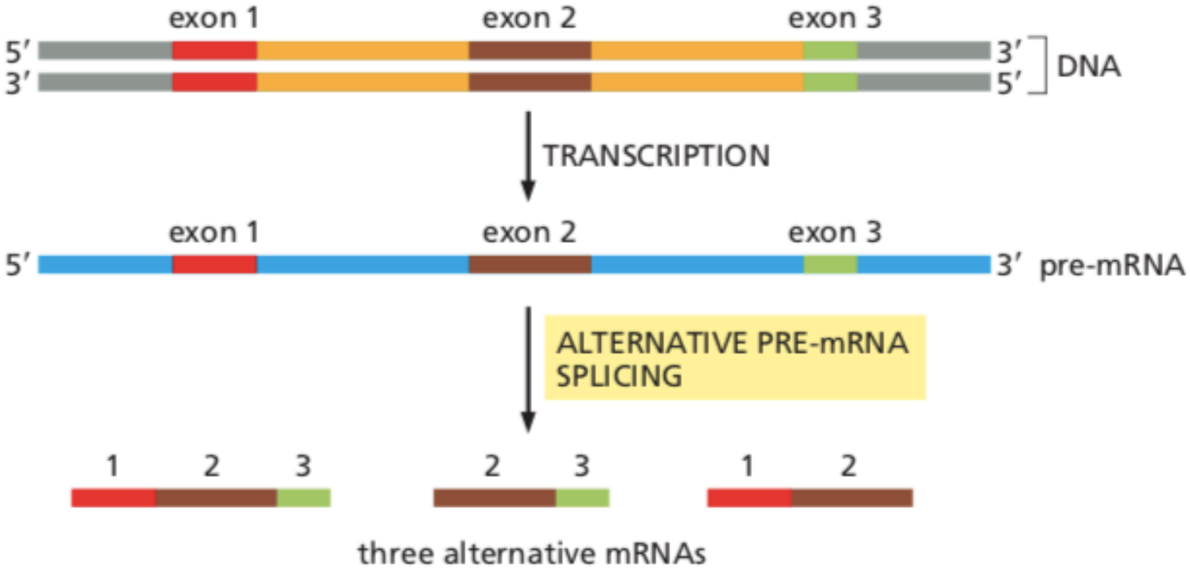
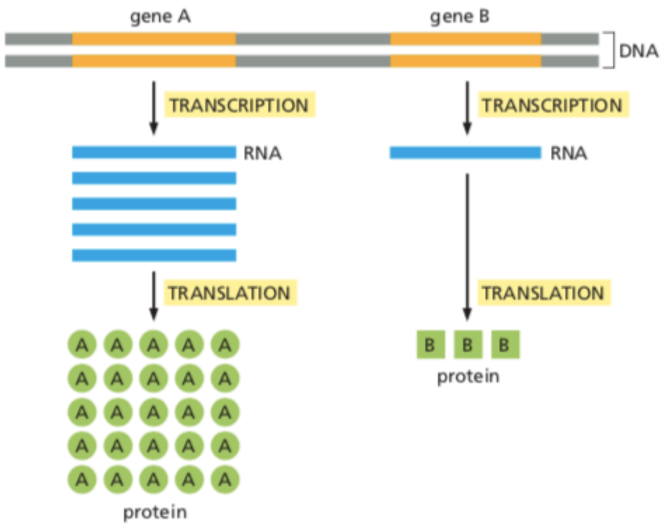

exon 1
exon 2
exon 3
5′
3′
3′
5′
DNA
TRANSCRIPTION
exon 1
exon 2
exon 3
5′
3′ pre-mRNA
ALTERNATIVE PRE-mRNA SPLICING
1 2 3
2 3
1 2
three alternative mRNAs

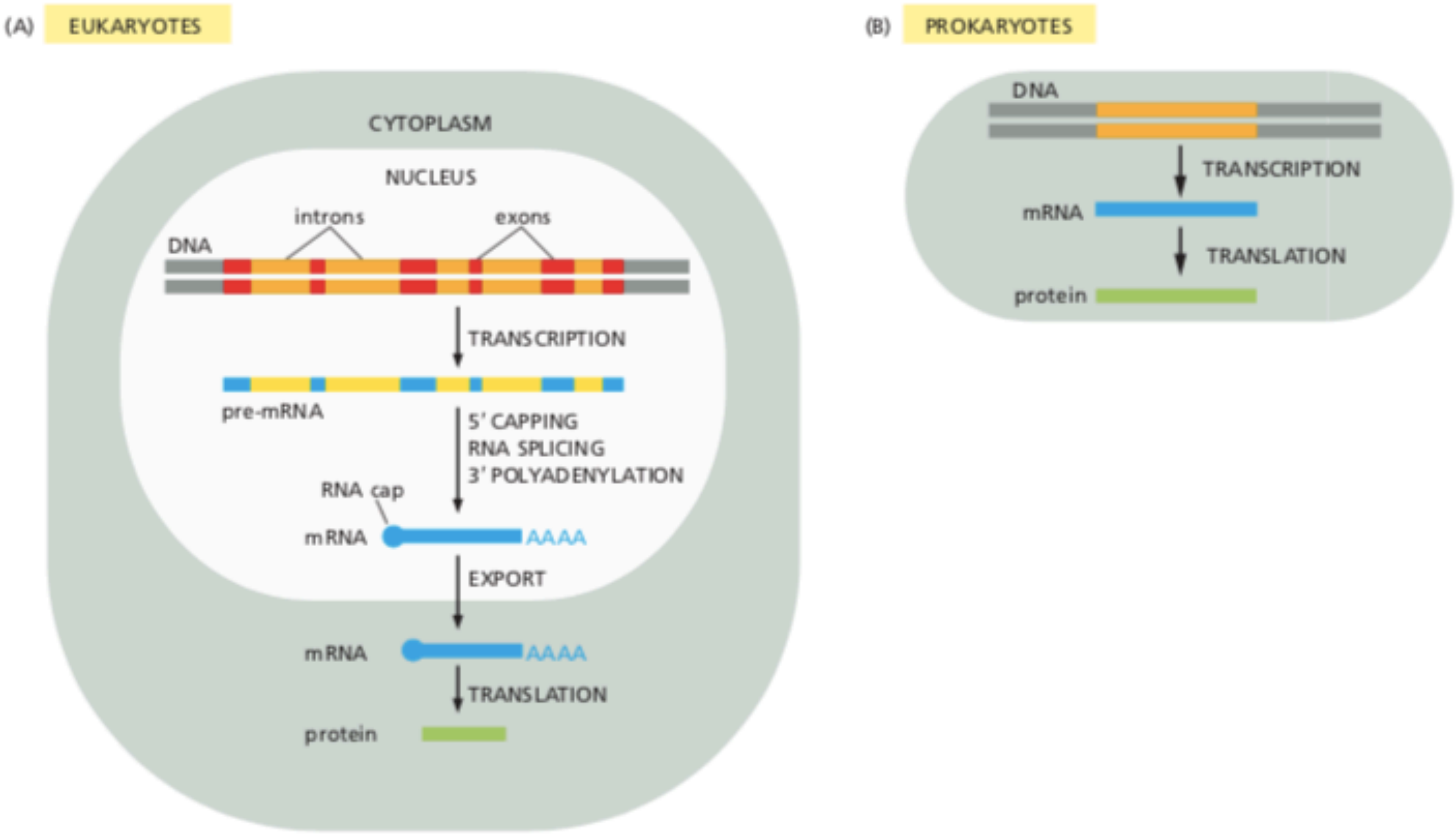
Figure 7–22 Some pre-mRNAs undergo alternative RNA splicing to produce various mRNAs and proteins from the same gene. Whereas all exons are present in a pre-mRNA, some exons can be excluded from the final mRNA molecule. In this example, three of four possible mRNAs are produced. The 5′ caps and poly-A tails on the mRNAs are not shown.

Degradación de ARNm

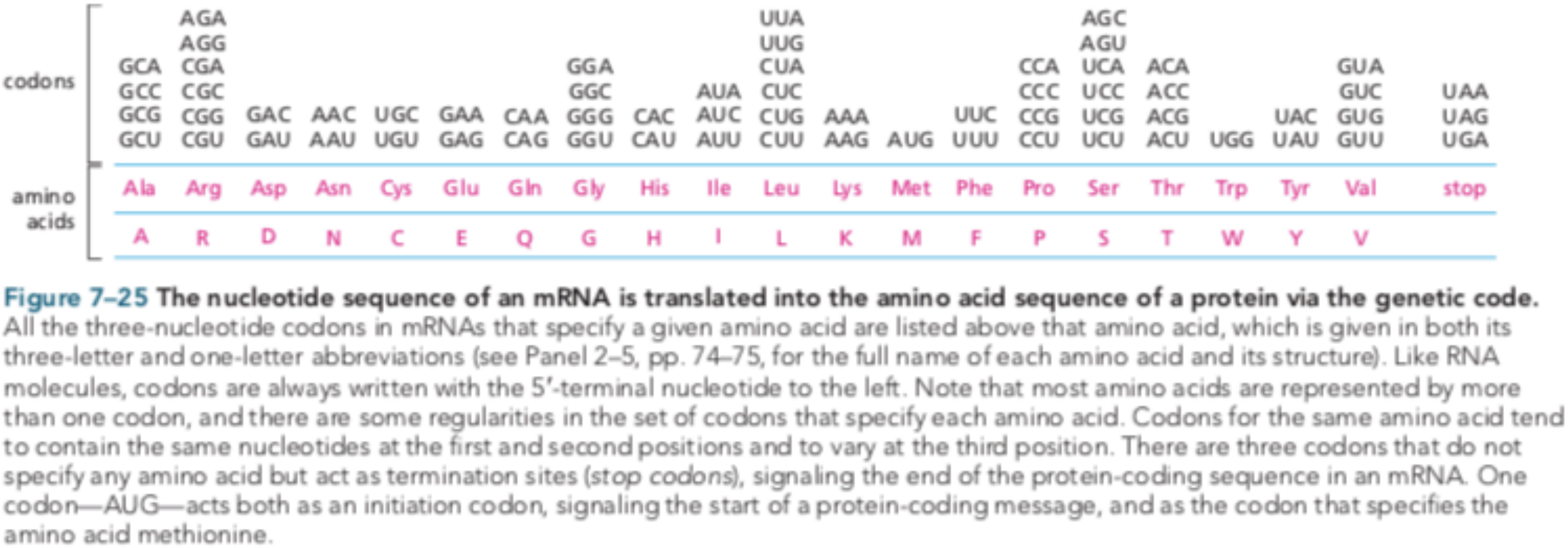
Debido a que una sola molécula de ARNm puede traducirse en proteína muchas veces, el tiempo que una molécula de ARNm madura persiste en la célula afecta la cantidad de proteína que produce. Cada molécula de ARNm es finalmente degradada en nucleótidos por las ribonucleasas (ARNas) presentes en el citosol.

La vida útil de las moléculas de ARNm difiere considerablemente, dependiendo de la secuencia de nucleótidos del ARNm y del tipo de célula. En las bacterias, la mayoría de los ARNm se degradan rápidamente y tienen una vida útil típica de unos 3 minutos. Los ARNm de las células eucariotas suelen persistir más tiempo: algunos, como los que codifican la β-globina, tienen una vida útil de más de 10 horas, mientras que otros tienen una vida útil de menos de 30 minutos.


gene A
gene B
DNA
TRANSCRIPTION
TRANSCRIPTION
RNA
RNA
TRANSLATION
TRANSLATION
A
B
protein
protein


(A)
EUKARYOTES
CYTOPLASM
NUCLEUS
introns
exons
DNA
TRANSCRIPTION
pre-mRNA
5' CAPPING
RNA SPLICING
3' POLYADENYLATION
RNA cap
mRNA
AAAA
EXPORT
mRNA
AAAA
TRANSLATION
protein
(B)
PROKARYOTES
DNA
TRANSCRIPTION
mRNA
TRANSLATION
protein

Código genético



codons	GCA GCC GCG GCU	AGA AGG CGA CGC CGG CGU	GAC GAU	AAC AAU	UGC UGU	GAA GAG	CAA CAG	GGA GGC GGG GGU	CAC CAU	AUA AUC AUU	UUA UUG CUA CUC CUG CUU	AAA AAG	AUG	UUC UUU	CCA CCC CCG CCU	AGC AGU UCA UCC UCG UCU	ACA ACC ACG ACU	UGG	UAC UAU	GUA GUC GUG GUU	UAA UAG UGA
amino acids	Ala	Arg	Asp	Asn	Cys	Glu	Gln	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val	stop
	A	R	D	N	C	E	Q	G	H	I	L	K	M	F	P	S	T	W	Y	V	

Figure 7–25 The nucleotide sequence of an mRNA is translated into the amino acid sequence of a protein via the genetic code. All the three-nucleotide codons in mRNAs that specify a given amino acid are listed above that amino acid, which is given in both its three-letter and one-letter abbreviations (see Panel 2–5, pp. 74–75, for the full name of each amino acid and its structure). Like RNA molecules, codons are always written with the 5′-terminal nucleotide to the left. Note that most amino acids are represented by more than one codon, and there are some regularities in the set of codons that specify each amino acid. Codons for the same amino acid tend to contain the same nucleotides at the first and second positions and to vary at the third position. There are three codons that do not specify any amino acid but act as termination sites (*stop codons*), signaling the end of the protein-coding sequence in an mRNA. One codon—AUG—acts both as an initiation codon, signaling the start of a protein-coding message, and as the codon that specifies the amino acid methionine.

La conversión de la información del ARN en proteína representa una **traducción** de la información a otro idioma que utiliza diferentes símbolos. Debido a que solo hay 4 nucleótidos diferentes en el ARNm pero 20 tipos diferentes de aminoácidos en una proteína, esta traducción no puede explicarse por una correspondencia directa uno a uno entre un nucleótido en el ARN y un aminoácido en la proteína.
Esto es distinto de la transcripción que significa copiar la información usando el mismo lenguaje (al pasar del ADN al ARN es como si un mensaje escrito a mano pasara a estar escrito “a máquina”).

En principio, una secuencia de ARNm se puede traducir usando tres marcos de lectura diferentes, dependiendo de dónde comience el proceso de decodificación. Sin embargo, solo uno de los tres posibles marcos de lectura en un ARNm especifica la proteína correcta.

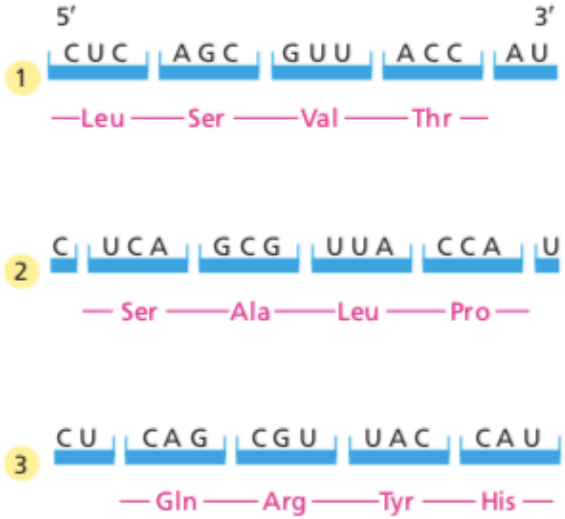
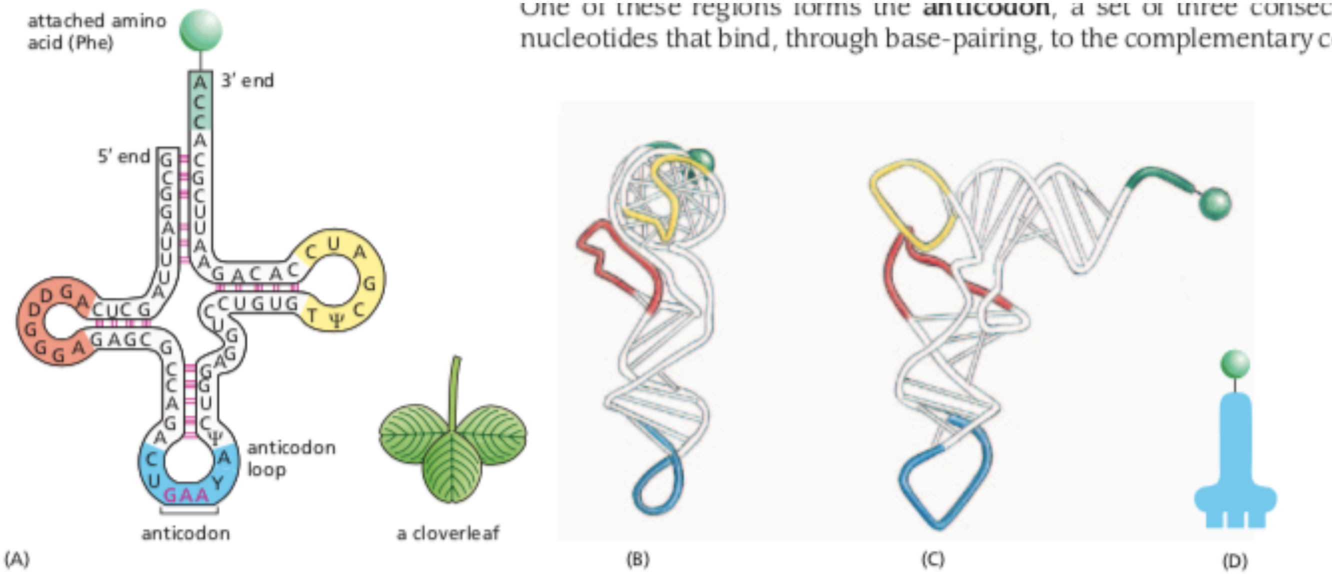

5′
3′
1
CUC AGC GUU ACC AU
—Leu—Ser—Val—Thr—
2
C UCA GCG UUA CCA U
—Ser—Ala—Leu—Pro—
3
CU CAG CGU UAC CAU
—Gln—Arg—Tyr—His—

Figure 7–26 In principle, an mRNA molecule can be translated in three possible reading frames. In the process of translating a nucleotide sequence (*blue*) into an amino acid sequence (*red*), the sequence of nucleotides in an mRNA molecule is read from the 5′ to the 3′ end in sequential sets of three nucleotides. In principle, therefore, the same mRNA sequence can specify three completely different amino acid sequences, depending on where translation begins—that is, on the reading frame used. In reality, however, only one of these reading frames encodes the actual message and is therefore used in translation, as we discuss later.

Los codones de una molécula de ARNm no reconocen directamente los aminoácidos que especifican: el grupo de tres nucleótidos no se une directamente al aminoácido. Más bien, la traducción de ARNm en proteína depende de moléculas adaptadoras que pueden reconocer y unirse a un codón en un sitio de su superficie y a un aminoácido en otro sitio. Estos adaptadores consisten en un conjunto de pequeñas moléculas de ARN conocidas como ARN de transferencia (ARNt), cada una de aproximadamente 80 nucleótidos de longitud.

One of these regions forms the **anticodon**, a set of three consecutive nucleotides that bind, through base-pairing, to the complementary codon



5′ GCGGAUUUAGCUCAGDDGGGAGAGCGCCAGACUGAAYAΨCUGGAGGUCCUGUGTΨCGAUCCACAGAAUUCGCACCA 3′
(E) anticodon

Figure 7–29 tRNA molecules are molecular adaptors, linking amino acids to codons. In this series of diagrams, the same tRNA molecule—in this case, a tRNA specific for the amino acid phenylalanine (Phe)—is depicted in various ways. (A) The conventional "cloverleaf" structure shows the complementary base-pairing (*red lines*) that creates the double-helical regions of the molecule. The anticodon loop (*blue*) contains the sequence of three nucleotides (*red letters*) that base-pairs with a codon in mRNA. The amino acid matching the codon–anticodon pair is attached at the 3′ end of the tRNA. tRNAs contain some unusual bases, which are produced by chemical modification after the tRNA has been synthesized. The bases denoted Ψ (for pseudouridine) and D (for dihydrouridine) are derived from uracil. (B and C) Views of the actual L-shaped molecule, based on X-ray diffraction analysis. These two images are rotated 90° with respect to each other. (D) Schematic representation of tRNA, emphasizing the anticodon, that will be used in subsequent figures. (E) The linear nucleotide sequence of the tRNA molecule, color-coded to match A, B, and C.

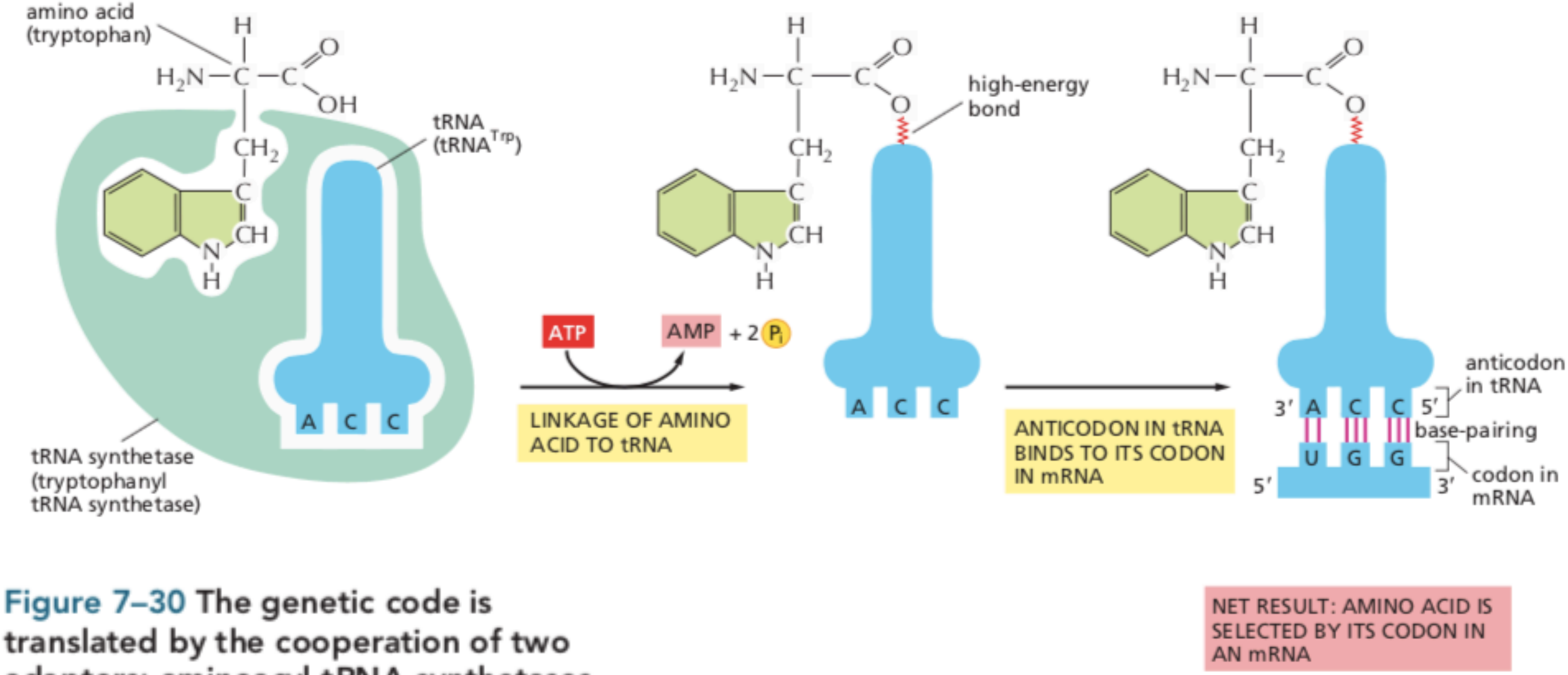

amino acid (tryptophan)
tRNA ($tRNA^{Trp}$)
tRNA synthetase (tryptophanyl tRNA synthetase)
A C C
ATP
AMP + 2 P_i
LINKAGE OF AMINO ACID TO tRNA
high-energy bond
ANTICODON IN tRNA BINDS TO ITS CODON IN mRNA
anticodon in tRNA
3′ A C C 5′
base-pairing
5′ U G G 3′
codon in mRNA
NET RESULT: AMINO ACID IS SELECTED BY ITS CODON IN AN mRNA

Figure 7–30 The genetic code is translated by the cooperation of two adaptors: aminoacyl-tRNA synthetases and tRNAs. Each synthetase couples a particular amino acid to its corresponding tRNAs, a process called charging. The anticodon on the charged tRNA molecule then forms base pairs with the appropriate codon on the mRNA. An error in either the charging step or the binding of the charged tRNA to its codon will cause the wrong amino acid to be incorporated into a protein chain. In the sequence of events shown, the amino acid tryptophan (Trp) is selected by the codon UGG on the mRNA.

Dos regiones de nucleótidos desapareados situadas en cada extremo de la molécula de ARNt son cruciales para la función de los ARNt en la síntesis de proteínas. Una forma el anticodón, un conjunto de tres nucleótidos consecutivos que se unen al codón complementario en una molécula de ARNm. La otra es una región corta de una sola hebra donde el aminoácido que coincide con el codón se une covalentemente al ARNt.

La traducción precisa y rápida de ARNm en proteína requiere una máquina molecular que pueda moverse a lo largo del ARNm, capturar moléculas de ARNt complementarias, mantener los ARNt en posición y luego enlazar covalentemente los aminoácidos que transportan para formar una cadena polipeptídica. Tanto en procariotas como en eucariotas, la máquina que hace el trabajo es el **ribosoma**, un gran complejo formado por docenas de pequeñas proteínas (las proteínas ribosomales) y varias moléculas de ARN llamadas ARN ribosómico (ARNr). Una célula eucariota típica contiene millones de ribosomas en su citoplasma.

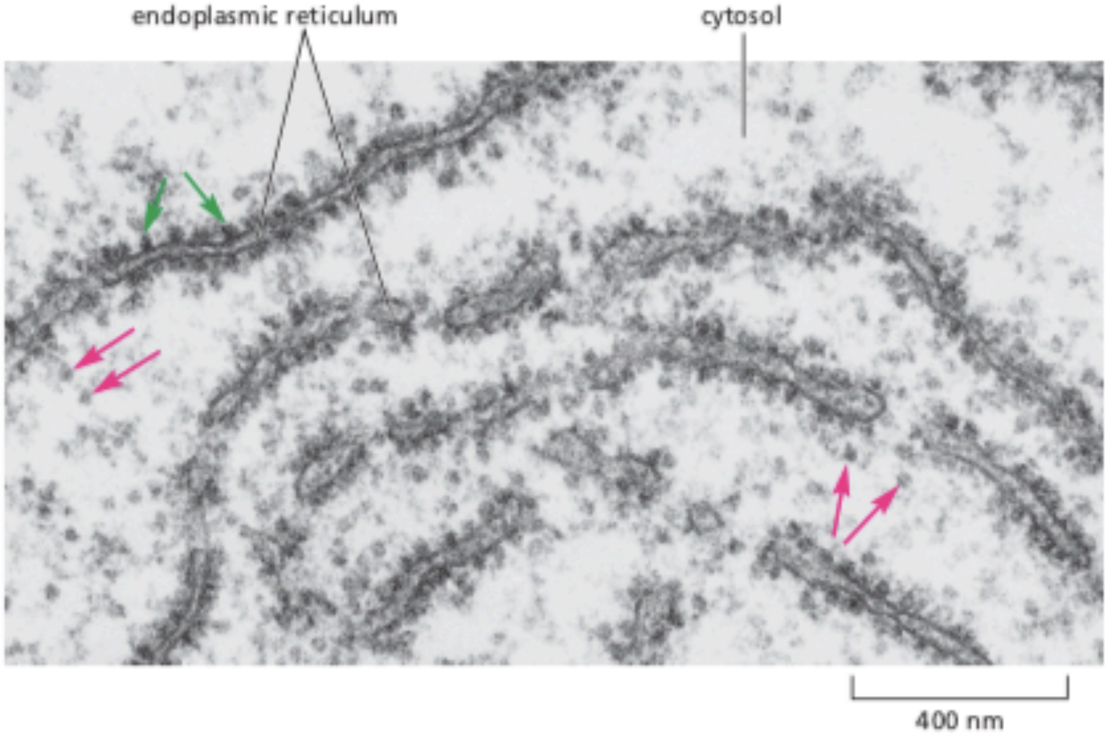


Figure 7–31 Ribosomes are located in the cytoplasm of eukaryotic cells. This electron micrograph shows a thin section of a small region of cytoplasm. The ribosomes appear as small gray blobs. Some are free in the cytosol (*red arrows*); others are attached to membranes of the endoplasmic reticulum (*green arrows*). (Courtesy of George Palade.)

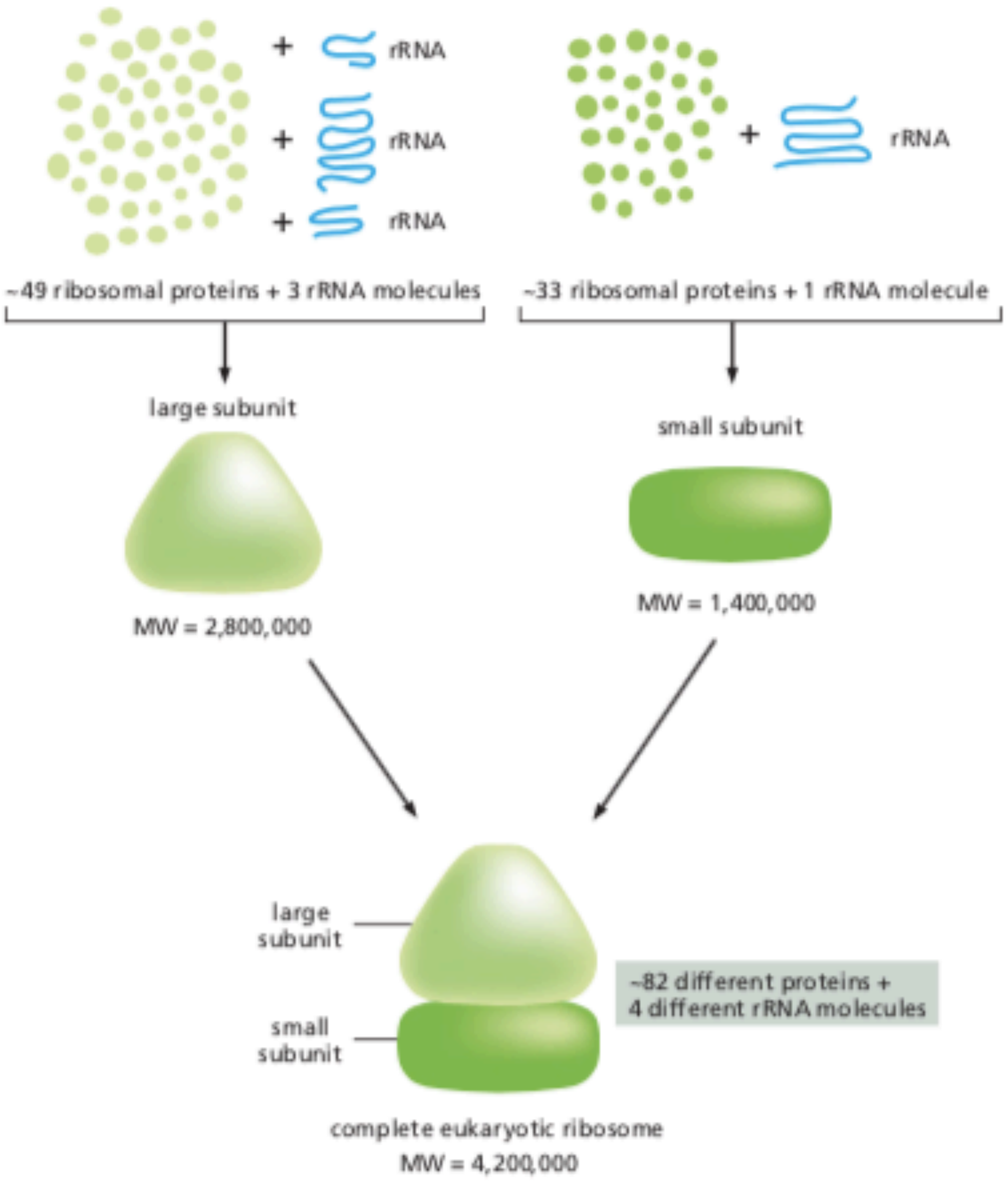
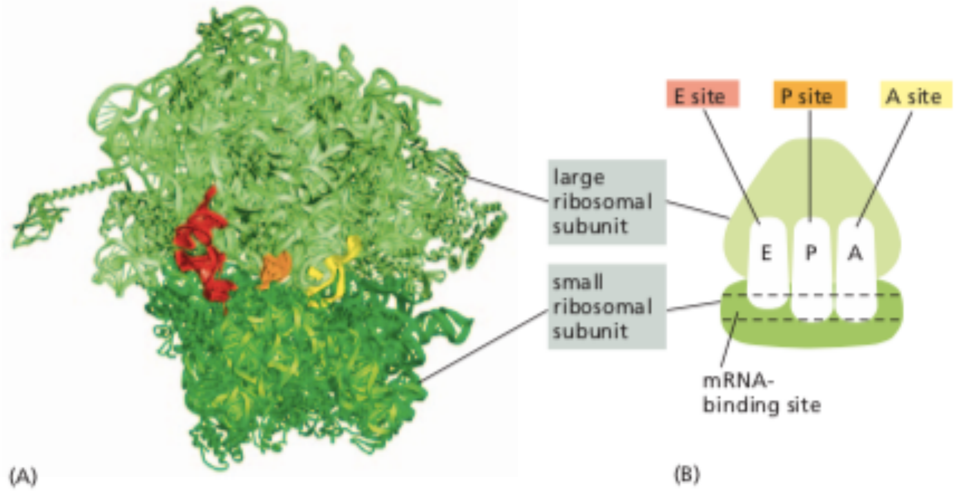

rRNA
rRNA
rRNA
rRNA
~49 ribosomal proteins + 3 rRNA molecules
~33 ribosomal proteins + 1 rRNA molecule
large subunit
small subunit
MW = 2,800,000
MW = 1,400,000
large subunit
small subunit
~82 different proteins + 4 different rRNA molecules
complete eukaryotic ribosome
MW = 4,200,000

Figure 7–32 The eukaryotic ribosome is a large complex of four rRNAs and more than 80 small proteins. Prokaryotic ribosomes are very similar: both are formed from a large and small subunit, which only come together after the small subunit has bound an mRNA. Although ribosomal proteins greatly outnumber rRNAs, the RNAs account for most of the mass of the ribosome and give it its overall shape and structure.


E site
P site
A site
large ribosomal subunit
small ribosomal subunit
E
P
A
mRNA-binding site
(A)
(B)

Además de un sitio de unión para una molécula de ARNm, cada ribosoma contiene tres sitios de unión para moléculas de ARNt, llamados sitio A, sitio P y sitio E. Para agregar un aminoácido a una cadena de péptidos en crecimiento, el ARNt cargado apropiado ingresa al sitio A mediante el apareamiento de bases con el codón complementario en la molécula de ARNm. Luego, su aminoácido se une a la cadena peptídica sostenida por el ARNt en el sitio P vecino. A continuación, la subunidad ribosómica grande se desplaza hacia adelante, moviendo el ARNt gastado al sitio E antes de expulsarlo. Este ciclo de reacciones se repite cada vez que se agrega un aminoácido a la cadena polipeptídica, y la nueva proteína crece desde su extremo amino hasta su extremo carboxilo hasta que se encuentra un codón de terminación en el ARNm.

Acá se ve describe el proceso

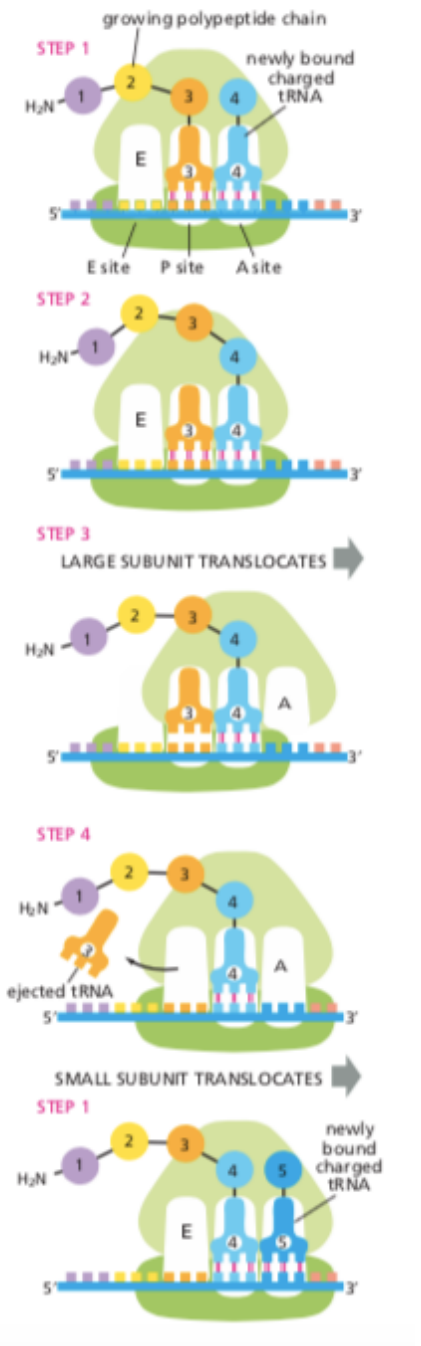
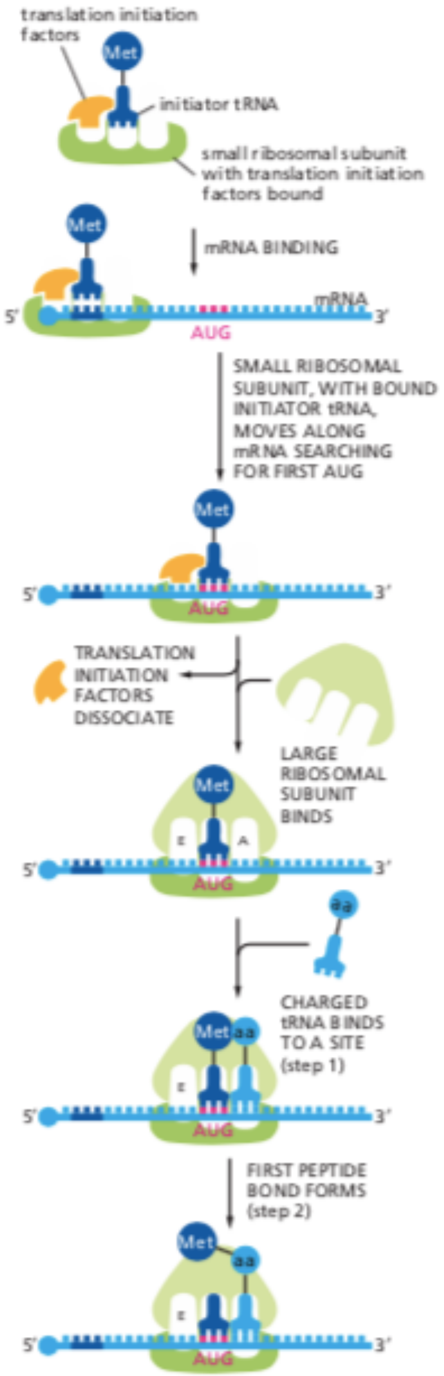

growing polypeptide chain
STEP 1
newly bound charged tRNA
H_2N
1
2
3
4
E
5'
3'
E site
P site
A site
STEP 2
STEP 3
LARGE SUBUNIT TRANSLOCATES
A
STEP 4
ejected tRNA
SMALL SUBUNIT TRANSLOCATES
STEP 1
newly bound charged tRNA
5

Figure 7–34 Translation takes place in a four-step cycle. This cycle is repeated over and over during the synthesis of a protein. In *step 1*, a charged tRNA carrying the next amino acid to be added to the polypeptide chain binds to the vacant A site on the ribosome by forming base pairs with the mRNA codon that is exposed there. Because only the appropriate tRNA molecules can base-pair with each codon, this codon determines the specific amino acid added. The A and P sites are sufficiently close together that their two tRNA molecules are forced to form base pairs with codons that are contiguous, with no stray bases in between. This positioning of the tRNAs ensures that the correct reading frame will be preserved throughout the synthesis of the protein. In *step 2*, the carboxyl end of the polypeptide chain (amino acid 3 in step 1) is uncoupled from the tRNA at the P site and joined by a peptide bond to the free amino group of the amino acid linked to the tRNA at the A site. This reaction is catalyzed by an enzymatic site in the large subunit. In *step 3*, a shift of the large subunit relative to the small subunit moves the two tRNAs into the E and P sites of the large subunit. In *step 4*, the small subunit moves exactly three nucleotides along the mRNA molecule, bringing it back to its original position relative to the large subunit. This movement ejects the spent tRNA and resets the ribosome with an empty A site so that the next charged tRNA molecule can bind (**Movie 7.8**).

As indicated, the mRNA is translated in the 5′-to-3′ direction, and the N-terminal end of a protein is made first, with each cycle adding one amino acid to the C-terminus of the polypeptide chain. To watch the translation cycle in atomic detail, see **Movie 7.9**.


translation initiation factors
Met
initiator tRNA
small ribosomal subunit with translation initiation factors bound
mRNA BINDING
mRNA
5′
3′
AUG
SMALL RIBOSOMAL SUBUNIT, WITH BOUND INITIATOR tRNA, MOVES ALONG mRNA SEARCHING FOR FIRST AUG
TRANSLATION INITIATION FACTORS DISSOCIATE
LARGE RIBOSOMAL SUBUNIT BINDS
E
A
aa
CHARGED tRNA BINDS TO A SITE (step 1)
FIRST PEPTIDE BOND FORMS (step 2)

En las células el inicio de la traducción requiere una señal específica para iniciar la traducción. El sitio en el que comienza la síntesis de proteínas en un ARNm es crucial, porque establece el marco de lectura para toda la longitud del mensaje. Un error de un nucleótido en cualquier sentido en esta etapa hará que todos los codones subsiguientes del ARNm sean mal interpretados, lo que dará como resultado una proteína no funcional. La tasa de iniciación determina la tasa a la que se sintetiza la proteína a partir del ARNm.

La traducción de un ARNm comienza con el codón AUG y se requiere un ARNt cargado especial para iniciar la traducción. Este ARNt iniciador siempre lleva el aminoácido metionina. Todas las proteínas recién creadas tienen metionina como aminoácido en el extremo que se sintetiza primero. La metionina suele eliminarse posteriormente mediante una proteasa específica.

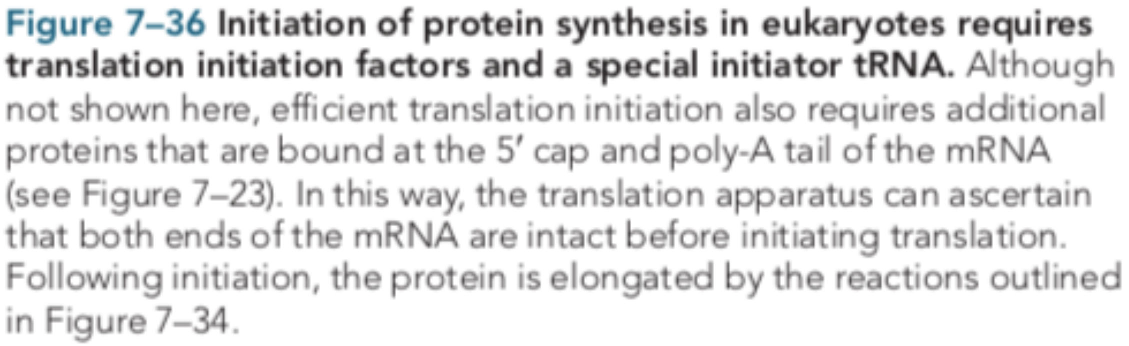
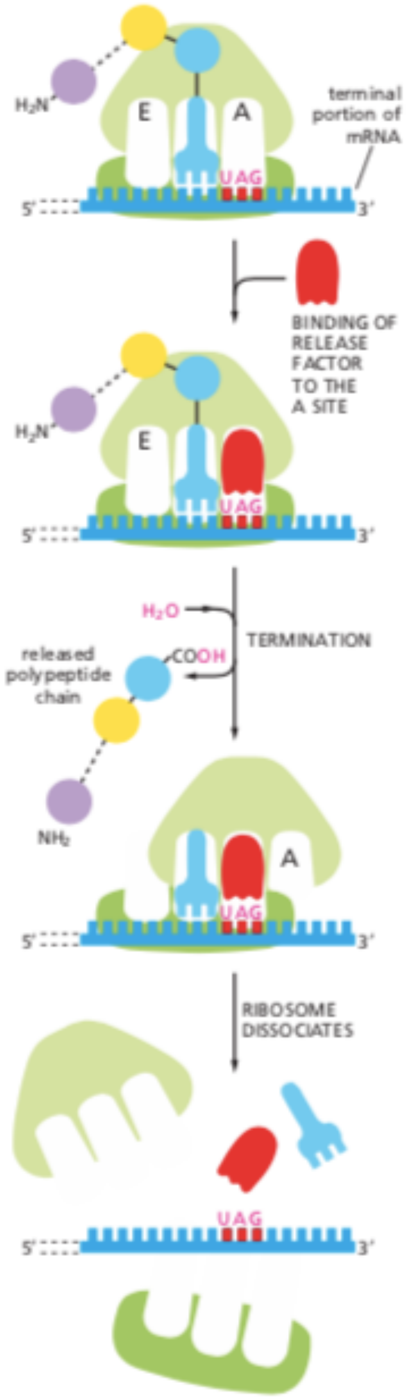


Figure 7–36 Initiation of protein synthesis in eukaryotes requires translation initiation factors and a special initiator tRNA. Although not shown here, efficient translation initiation also requires additional proteins that are bound at the 5′ cap and poly-A tail of the mRNA (see Figure 7–23). In this way, the translation apparatus can ascertain that both ends of the mRNA are intact before initiating translation. Following initiation, the protein is elongated by the reactions outlined in Figure 7–34.

El final de la traducción se indica por la presencia de uno de varios codones, llamados codones de terminación, en el ARNm.


H₂N
E
A
terminal portion of mRNA
UAG
5'
3'
BINDING OF RELEASE FACTOR TO THE A SITE
H₂O
TERMINATION
released polypeptide chain
COOH
NH₂
RIBOSOME DISSOCIATES

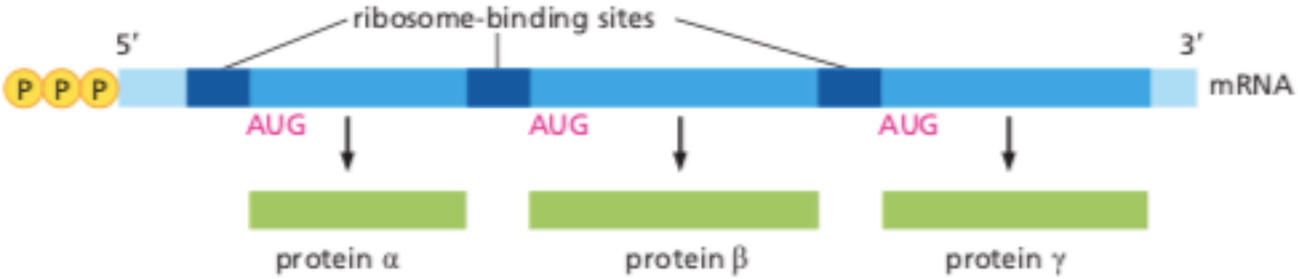


Figure 7–37 A single prokaryotic mRNA molecule can encode several different proteins. In prokaryotes, genes directing the different steps in a process are often organized into clusters (operons) that are transcribed together into a single mRNA. A prokaryotic mRNA does not have the same sort of 5′ cap as a eukaryotic mRNA, but instead has a triphosphate at its 5′ end. Prokaryotic ribosomes initiate translation at ribosome-binding sites (*dark blue*), which can be located in the interior of an mRNA molecule. This feature enables prokaryotes to synthesize different proteins from a single mRNA molecule, with each protein made by a different ribosome.

El mecanismo para seleccionar un codón de inicio es diferente en las bacterias. Los ARNm bacterianos no tienen “caps” 5’ para indicarle al ribosoma dónde comenzar la traducción. Contienen secuencias específicas de unión a ribosomas, de hasta seis nucleótidos de longitud, que se corriente arriba de los codones AUG en los que debe comenzar la traducción. Los ribosomas procarióticos pueden unirse directamente a un codón de inicio en el interior de un ARNm, siempre que un sitio de unión al ribosoma lo anteceda por varios nucleótidos. A menudo los mRNA procarióticos codifican varias proteínas diferentes, cada una de las cuales se traduce a partir de la misma molécula de mRNA. Por el contrario, un ARNm eucariota generalmente lleva la información de una sola proteína.

Degradación de proteínas para controlar su cantidad

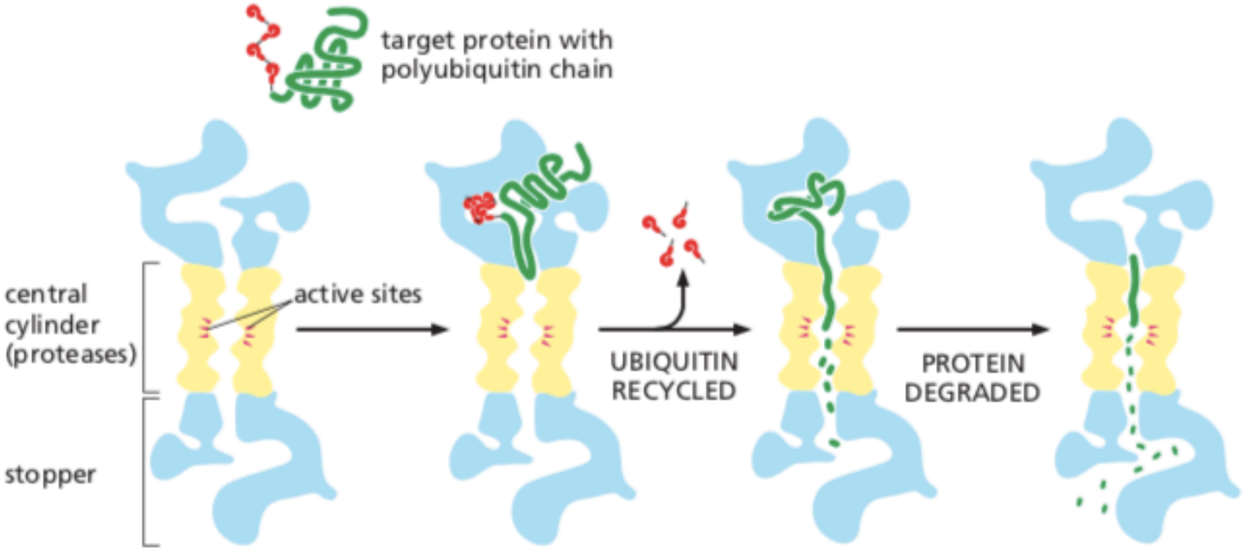


Figure 7–41 Proteins marked by a polyubiquitin chain are degraded by the proteasome. Proteins in the stopper of a proteasome (*blue*) recognize target proteins marked by a specific type of polyubiquitin chain. The stopper then unfolds the target protein and threads it into the proteasome's central cylinder (*yellow*), which is lined with proteases that chop the protein to pieces.

Las células poseen vías especializadas que descomponen enzimáticamente las proteínas en sus aminoácidos constituyentes (un proceso denominado proteólisis). Las enzimas que degradan las proteínas, primero en péptidos cortos y finalmente en aminoácidos individuales, se conocen colectivamente como proteasas. Las proteasas actúan cortando (hidrolizando) los enlaces peptídicos entre los aminoácidos. Una función de las vías proteolíticas es degradar rápidamente aquellas proteínas cuya vida útil debe ser corta. Otra es reconocer y eliminar las proteínas dañadas o mal dobladas.

En las células eucariotas, las proteínas son degradadas por máquinas proteicas llamadas proteasomas, presentes tanto en el citosol como en el núcleo. En eucariotas, los proteasomas actúan principalmente sobre proteínas que han sido marcadas para su destrucción por la unión covalente de una pequeña proteína llamada ubiquitina. Esta unión es producida por enzimas especializadas.

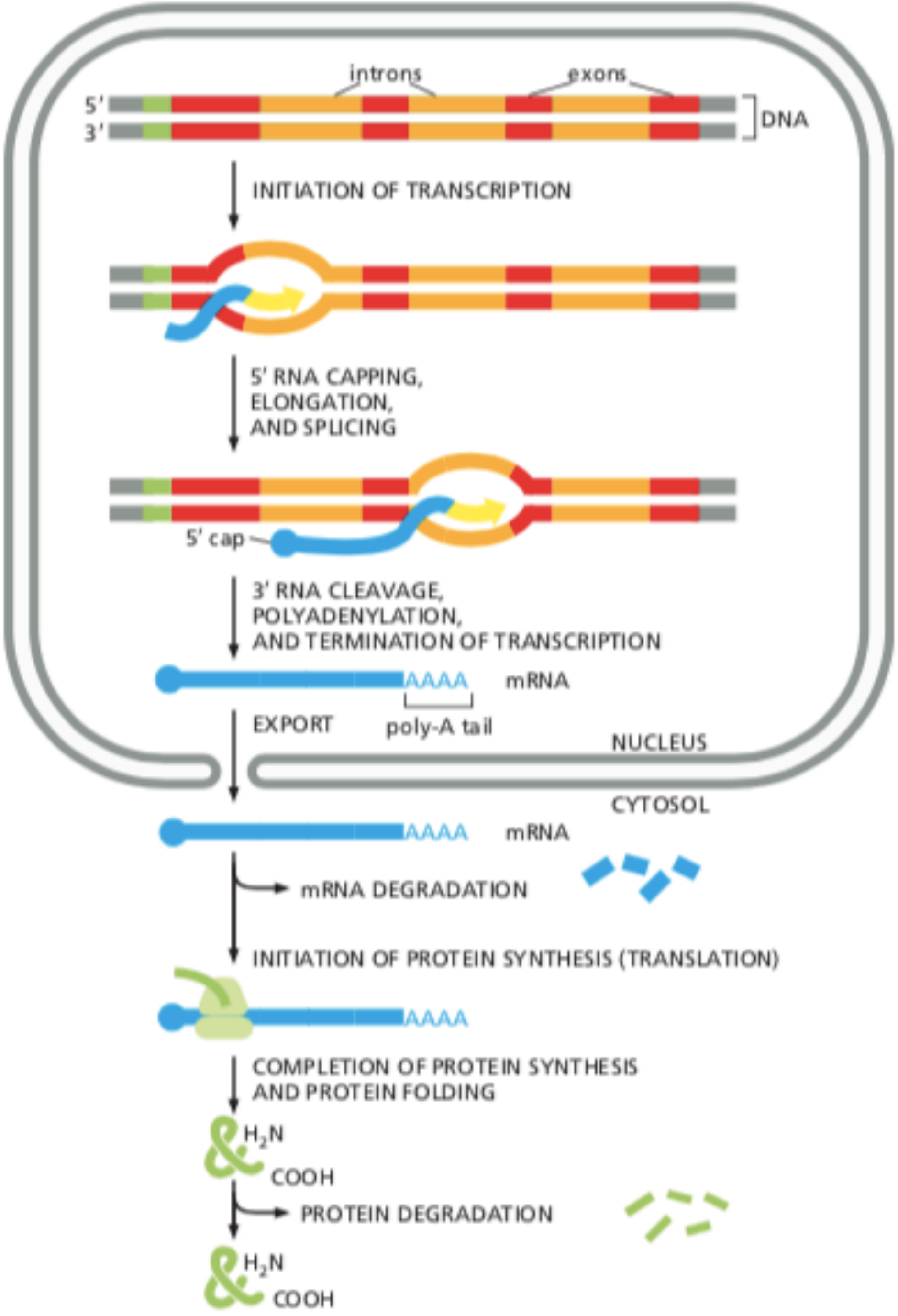

introns
exons
5'
3'
DNA
INITIATION OF TRANSCRIPTION
5' RNA CAPPING,
ELONGATION,
AND SPLICING
5' cap
3' RNA CLEAVAGE,
POLYADENYLATION,
AND TERMINATION OF TRANSCRIPTION
AAAA
mRNA
EXPORT
poly-A tail
NUCLEUS
CYTOSOL
AAAA
mRNA
mRNA DEGRADATION
INITIATION OF PROTEIN SYNTHESIS (TRANSLATION)
AAAA
COMPLETION OF PROTEIN SYNTHESIS
AND PROTEIN FOLDING
H2N
COOH
PROTEIN DEGRADATION
H2N
COOH

Figure 7–42 Protein production in a eukaryotic cell requires many steps. The final concentration of each protein depends on the rate of each step depicted. Even after an mRNA and its corresponding protein have been produced, their concentrations can be regulated by degradation. Although not shown here, the activity of the protein can also be regulated by other post-translational modifications or the binding of small molecules (see Figure 7–43).

Expresión génica

El ADN de un organismo codifica todas las moléculas de ARN y proteínas que se necesitan para producir sus células. Sin embargo, una descripción completa de la secuencia de ADN de un organismo, ya sean los pocos millones de nucleótidos de una bacteria o los pocos miles de millones de nucleótidos en cada célula humana, no nos permite reconstruir ese organismo.

El trabajo coordinado de esta maquinaria está directamente relacionado con la **expresión génica**. Incluso la bacteria unicelular más simple puede usar sus genes de manera selectiva, por ejemplo, activando y desactivando genes para producir las enzimas necesarias para digerir cualquier fuente de alimento disponible. En plantas y animales multicelulares la expresión génica está bajo un control mucho más elaborado.

En el transcurso del desarrollo embrionario, un óvulo fertilizado da lugar a muchos tipos de células que difieren dramáticamente tanto en estructura como en función. Las diferencias entre una célula nerviosa que procesa información y un glóbulo blanco que combate las infecciones, por ejemplo, son tan extremas que es difícil imaginar que las dos células contengan el mismo ADN. Ahora sabemos que casi todas las células de un organismo multicelular contienen el mismo genoma. En cambio, la diferenciación celular se logra mediante cambios en la expresión génica.

La expresión génica es un proceso complejo mediante el cual las células dirigen selectivamente la síntesis de los muchos miles de proteínas y ARN codificados en su genoma. Pero, ¿cómo coordinan y controlan las células un proceso tan intrincado, y cómo especifica una célula individual cuál de sus genes expresar? Esta decisión es un problema especialmente importante para los animales porque, a medida que se desarrollan, sus células se vuelven altamente especializadas, produciendo finalmente una serie de células musculares, nerviosas y sanguíneas, junto con los cientos de otros tipos de células que se observan en el adulto. Tal diferenciación celular surge porque las células producen y acumulan diferentes conjuntos de moléculas de ARN y proteínas: es decir, expresan diferentes genes.

Diferentes tipos de células producen diferentes conjuntos de proteínas

Una célula puede cambiar la expresión de sus genes en respuesta a señales externas.

La expresión génica se puede regular en varios pasos, desde el DNA hasta el RNA y hasta las proteínas.

Así, una célula puede controlar las proteínas que contiene:

(1) controlando cuándo y con qué frecuencia se transcribe un gen determinado,
(2) controlando cómo se empalma o procesa una transcripción de ARN,
(3) seleccionando qué ARNms se exportan del núcleo al citosol,
(4) regulando la rapidez con la que se degradan ciertas moléculas de ARNm,
(5) Seleccionando qué ARNms se traducen en proteína por los ribosomas,
(6) Regulando la rapidez con la que se destruyen proteínas específicas después de su fabricación

Además, la actividad de las proteínas individuales se puede regular adicionalmente de diversas formas.

How Transcriptional Switches Work 265

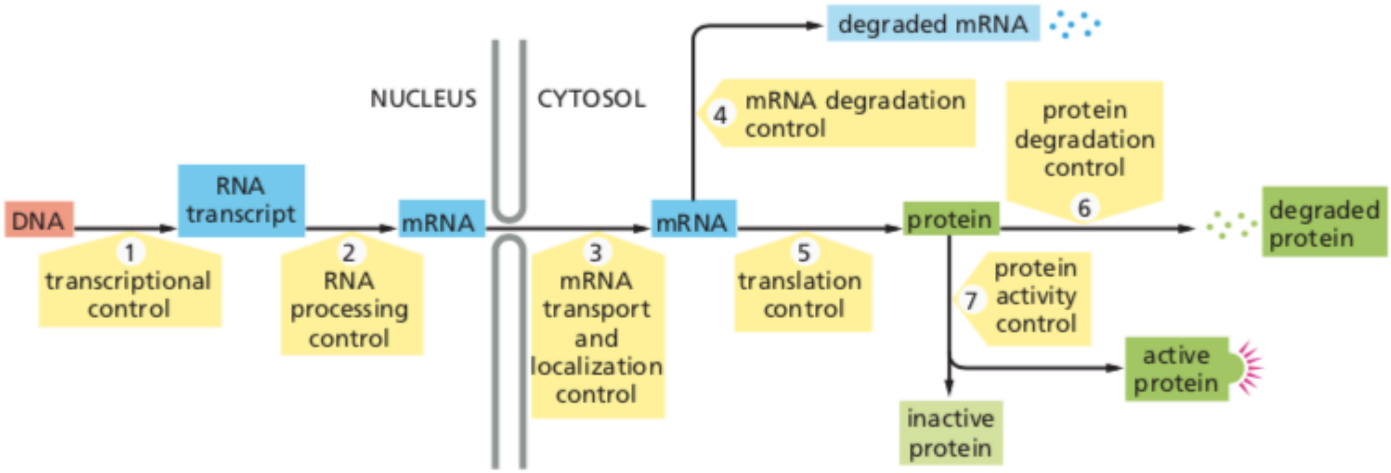

degraded mRNA
NUCLEUS
CYTOSOL
4 mRNA degradation control
protein degradation control
6
RNA transcript
DNA
mRNA
mRNA
protein
degraded protein
1 transcriptional control
2 RNA processing control
3 mRNA transport and localization control
5 translation control
7 protein activity control
active protein
inactive protein

Figure 8–3 Gene expression in eukaryotic cells can be controlled at various steps. Examples of regulation at each of these steps are known, although for most genes the main site of control is step 1—transcription of a DNA sequence into RNA.

Para la mayoría de los genes, el control de la transcripción (paso número 1 en la Figura) es el más usado. Esto tiene sentido porque solo el control transcripcional puede asegurarse de que no se sintetizan intermediarios innecesarios.

“Switches” transcripcionales

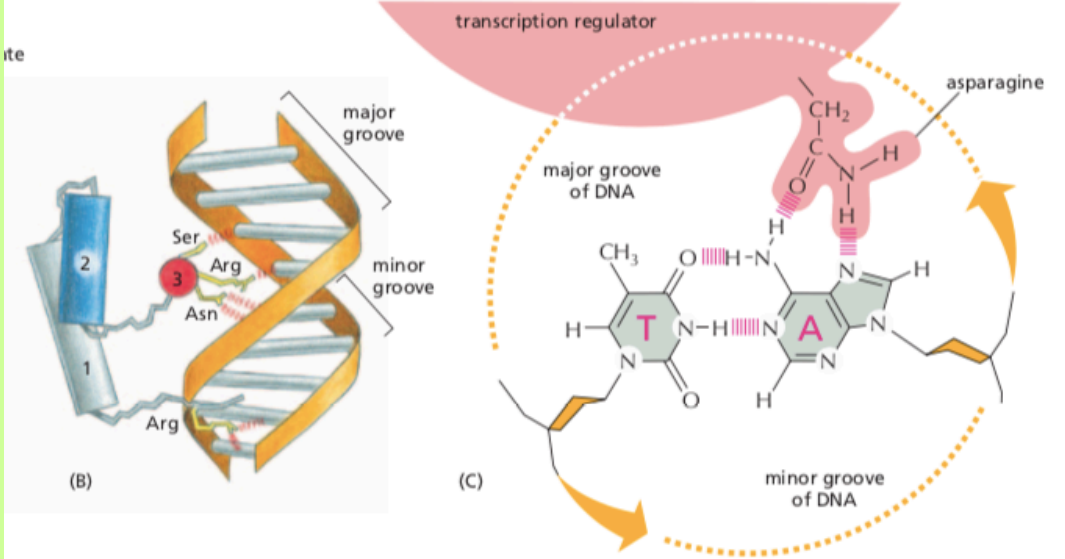
Antes vimos que la región promotora de un gen se une a la enzima ARN polimerasa y orienta correctamente la enzima para comenzar su tarea. Los promotores de genes bacterianos y eucariotas incluyen un sitio de inicio de la transcripción, donde comienza la síntesis de ARN, más una secuencia de aproximadamente 50 pares de nucleótidos que se extiende corriente arriba desde el sitio de inicio. Esta región contiene sitios que son necesarios para que la ARN polimerasa reconozca al promotor, aunque no se unen directamente a la ARN polimerasa. Son sitios de reconocimiento para proteínas que se asocian con la polimerasa activa: factor sigma en bacterias o factores de transcripción generales en eucariotas.

Además del promotor, casi todos los genes, ya sean bacterianos o eucarióticos, tienen secuencias de ADN reguladoras que se utilizan para activar o desactivar el gen. Algunas secuencias de ADN reguladoras son tan cortas como 10 pares de nucleótidos y actúan como interruptores simples que responden a una sola señal. Estos simples interruptores reguladores predominan en las bacterias. Otras secuencias de ADN reguladoras, especialmente las de eucariotas, son muy largas (más de 10.000 pares de nucleótidos) y actúan como microprocesadores moleculares, integrando la información proveniente de una variedad de señales para determinar cuándo se inicia la transcripción.

Para tener algún efecto, las secuencias reguladoras deben ser reconocidas por proteínas llamadas reguladores de la transcripción. Es la unión de un regulador de la transcripción a una secuencia de ADN reguladora que actúa como el interruptor para controlar la transcripción.

La bacteria más simple produce varios cientos de reguladores de transcripción diferentes, cada uno de los cuales reconoce una secuencia de ADN diferente y, por lo tanto, regula un conjunto distinto de genes. Los seres humanos producen muchos más, varios miles, lo que indica la importancia y complejidad de esta forma de regulación genética en el desarrollo y la función de un organismo complejo.

Las proteínas que reconocen una secuencia de nucleótidos específica lo hacen porque la superficie de la proteína se ajusta firmemente algunas “cosas” presentes en la superficie de la doble hélice del ADN cuyas características dependen de la secuencia de nucleótidos. En la mayoría de los casos, la proteína se inserta en el surco principal de la hélice del ADN y hace una serie de contactos moleculares con los pares de nucleótidos dentro del surco. Aunque cada contacto individual es débil, los 10 a 20 contactos que se forman típicamente en la interfaz proteína-ADN se combinan para asegurar que la interacción sea altamente específica y muy fuerte, una de las más estrictas y específicas conocidas en biología.



A transcription regulator interacts with the major groove of a DNA double helix. (A) This regulator
DNA via three α helices, shown as numbered cylinders, which allow the protein to fit into the major
form tight associations with the base pairs in a short stretch of DNA. This particular structural motif,
meodomain, is found in many eukaryotic DNA-binding proteins (Movie 8.1). (B) Most of the contacts with
ases are made by helix 3 (*red*), which is shown here end-on. The protein interacts with the edges of the
s without disrupting the hydrogen bonds that hold the base pairs together. (C) An asparagine residue
3 forms two hydrogen bonds with the adenine in an A-T base pair. The view is end-on looking down the
le helix, and the protein contacts the base pair from the major groove side. For simplicity, only one amino
contact is shown; in reality, transcription regulators form hydrogen bonds (as shown here), ionic bonds,
hobic interactions with individual bases in the major groove. Typically, the protein–DNA interface would
0–20 such contacts, each involving a different amino acid and each contributing to the overall strength of
–DNA interaction.

Ejemplo de “switch” transcripción en respuesta a cambios en el entorno

El genoma de la bacteria E. coli consta de una única molécula de ADN circular de ~ $4{,}6 \times 10^6$ pares de nucleótidos. Este ADN codifica aproximadamente 4300 proteínas. Las bacterias regulan la expresión de muchos de sus genes según las fuentes de alimentos disponibles.

En E. coli, cinco genes codifican enzimas que fabrican el aminoácido triptófano. Estos genes están dispuestos en un grupo en el cromosoma y se transcriben a partir de un solo promotor como una molécula de ARNm larga; estos grupos transcritos coordinadamente se denominan operones. Aunque los operones son comunes en las bacterias, son raros en los eucariotas, donde los genes se transcriben y regulan individualmente.

Cuando las concentraciones de triptófano son bajas, el operón se transcribe; el ARNm resultante se traduce para producir un conjunto completo de enzimas biosintéticas, que trabajan en conjunto para sintetizar triptófano. Cuando el triptófano es abundante, el aminoácido se importa a la célula y detiene la producción de las enzimas

terium can respond very rapidly to a rise in tryptophan concentration.

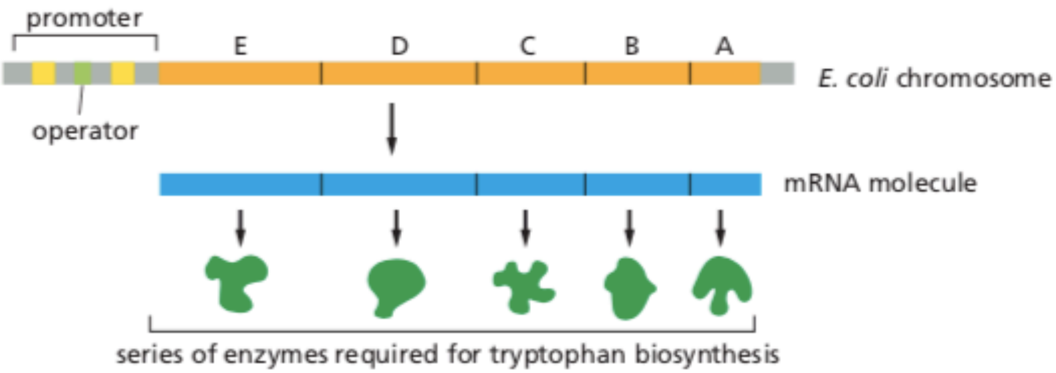


Figure 8–6 A cluster of bacterial genes can be transcribed from a single promoter. Each of these five genes encodes a different enzyme; all of the enzymes are needed to synthesize the amino acid tryptophan. The genes are transcribed as a single mRNA molecule, a feature that allows their expression to be coordinated. Clusters of genes transcribed as a single mRNA molecule are common in bacteria. Each of these clusters is called an *operon* because its expression is controlled by a regulatory DNA sequence called the *operator* (*green*), situated within the promoter. The *yellow* blocks in the promoter represent DNA sequences that bind RNA polymerase.

Así como la transcripción puede ser “apagada”, también puede ser “encendida"

Las proteínas activadoras a menudo tienen que interactuar con una segunda molécula para poder unirse al ADN. Por ejemplo, la proteína activadora bacteriana CAP tiene que unirse al AMP cíclico (cAMP) antes de poder unirse al ADN. Los genes activados por CAP se activan en respuesta a un aumento en la concentración de cAMP intracelular, que aumenta cuando la glucosa, la fuente de carbono preferida de la bacteria, ya no está disponible; como resultado, CAP impulsa la producción de enzimas que permiten a la bacteria digerir otros azúcares.

En algunos casos un mismo promotor está regulado por dos reguladores de la transcripción distintos.

Ejemplo muy estudiado: Operón LAC en E. coli.

El opera LAC está controlado por el represor Lac y por el activador CAP que acabamos de ver. El operón Lac codifica las proteínas necesarias para importar y digerir el disacárido lactosa. En ausencia de glucosa, la bacteria produce cAMP, que activa CAP para activar genes que permiten que la célula utilice fuentes alternativas de carbono, incluida la lactosa. El represor Lac apaga el operón en ausencia de lactosa. Esta disposición permite que la región de control del operón Lac integre dos señales diferentes, de modo que el operón se exprese en gran medida solo cuando se cumplen dos condiciones: la glucosa debe estar ausente y la lactosa debe estar presente.

Este circuito genético se comporta de forma muy similar a un interruptor que realiza una operación lógica en una computadora: la transcripción se activa solo si la glucosa está ausente y la lactosa está presente.

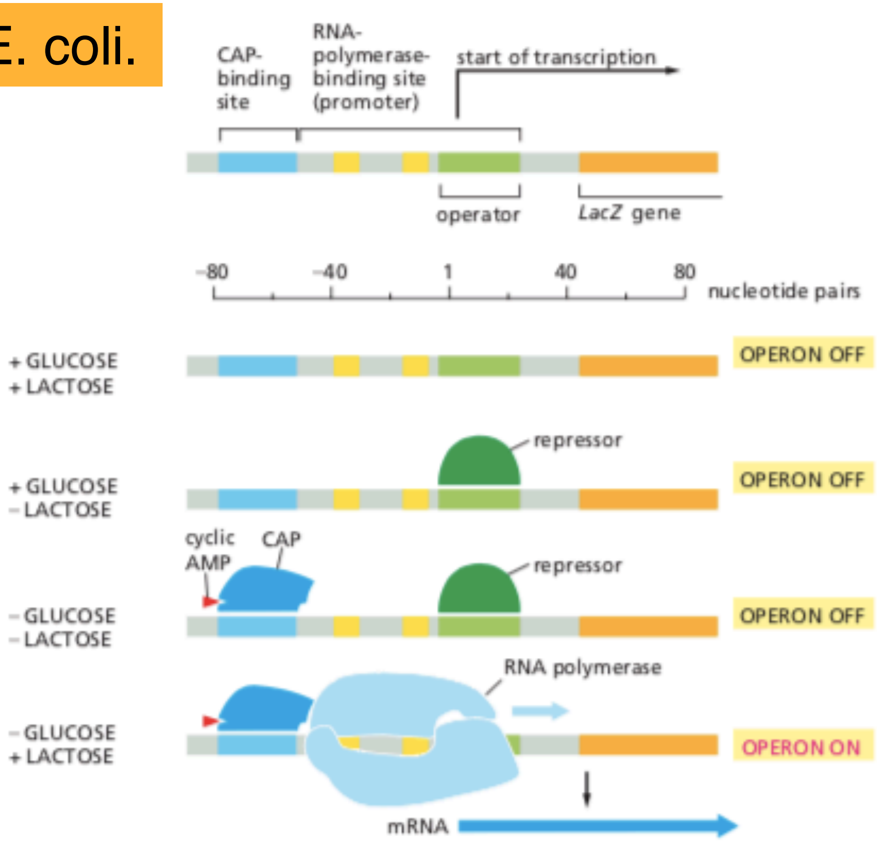

CAP-binding site
RNA-polymerase-binding site (promoter)
start of transcription
operator
LacZ gene
−80
−40
1
40
80
nucleotide pairs
+ GLUCOSE
+ LACTOSE
OPERON OFF
repressor
+ GLUCOSE
− LACTOSE
OPERON OFF
cyclic AMP
CAP
repressor
− GLUCOSE
− LACTOSE
OPERON OFF
RNA polymerase
− GLUCOSE
+ LACTOSE
OPERON ON
mRNA

Figure 8–9 The *Lac* operon is controlled by two transcription regulators, the Lac repressor and CAP. When lactose is absent, the Lac repressor binds to the *Lac* operator and shuts off expression of the operon. Addition of lactose increases the intracellular concentration of a related compound, allolactose; allolactose binds to the Lac repressor, causing it to undergo a conformational change that releases its grip on the operator DNA (not shown). When glucose is absent, cyclic AMP (*red* triangle) is produced by the cell, and CAP binds to DNA. *LacZ*, the first gene of the operon, encodes the enzyme β-galactosidase, which breaks down lactose to galactose and glucose.

Las células eucariotas pueden controlar la expresión génica a distancia.

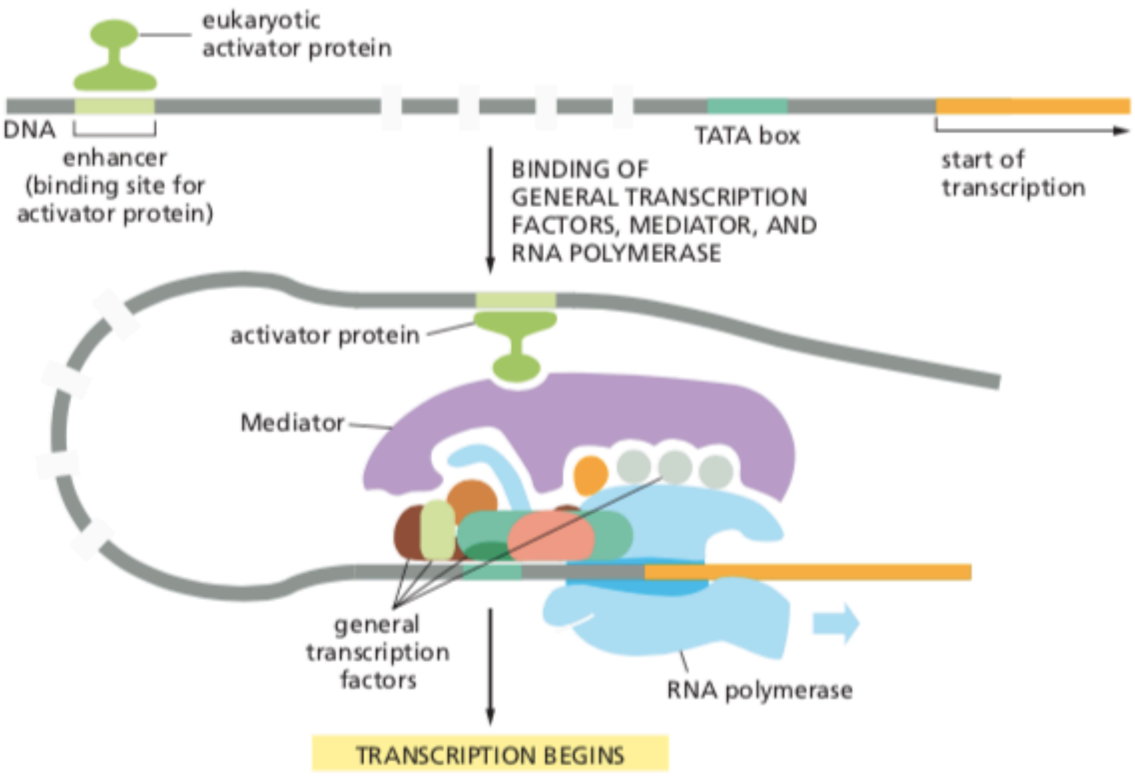

eukaryotic activator protein
DNA
enhancer (binding site for activator protein)
TATA box
start of transcription
BINDING OF GENERAL TRANSCRIPTION FACTORS, MEDIATOR, AND RNA POLYMERASE
activator protein
Mediator
general transcription factors
RNA polymerase
TRANSCRIPTION BEGINS

Figure 8–10 In eukaryotes, gene activation can occur at a distance. An activator protein bound to a distant enhancer attracts RNA polymerase and general transcription factors to the promoter. Looping of the intervening DNA permits contact between the activator and the transcription initiation complex bound to the promoter. In the case shown here, a large protein complex called Mediator serves as a go-between. The broken stretch of DNA signifies that the length of DNA between the enhancer and the start of transcription varies, sometimes reaching tens of thousands of nucleotide pairs in length. The TATA box is a DNA recognition sequence for the first general transcription factor that binds to the promoter (see Figure 7–12).

Los reguladores de la transcripción en eucariotas ayudan a iniciar la transcripción reclutando proteínas modificadoras de la cromatina.

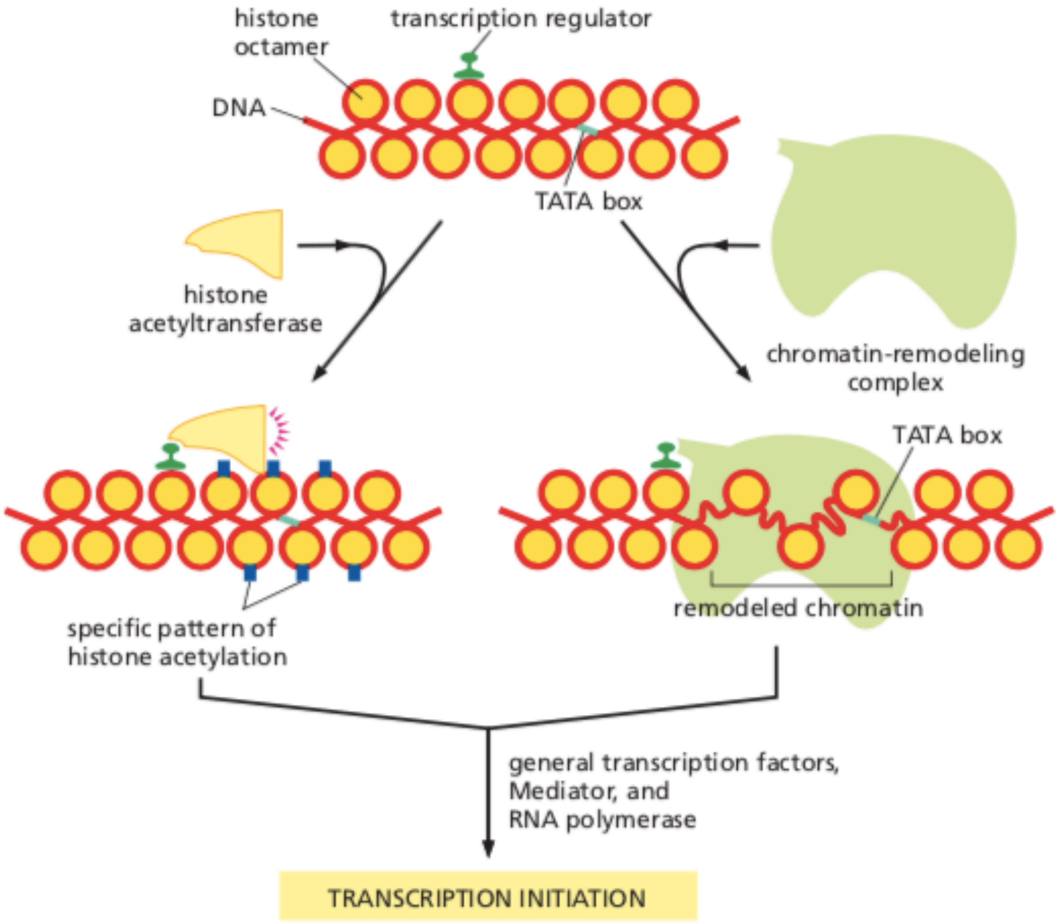

histone octamer
transcription regulator
DNA
TATA box
histone acetyltransferase
chromatin-remodeling complex
TATA box
specific pattern of histone acetylation
remodeled chromatin
general transcription factors, Mediator, and RNA polymerase
TRANSCRIPTION INITIATION

Figure 8–11 Eukaryotic transcriptional activators can recruit chromatin-modifying proteins to help initiate gene transcription. On the right, chromatin-remodeling complexes render the DNA packaged in chromatin more accessible to other proteins in the cell, including those required for transcription initiation; notice, for example, the increased exposure of the TATA box. On the left, the recruitment of histone-modifying enzymes such as histone acetyltransferases adds acetyl groups to specific histones, which can then serve as binding sites for proteins that stimulate transcription initiation (not shown).

Los mecanismos que crean tipos especializados de células

Además de los mecanismos para activar y desactivar genes en respuesta a señales en el entorno, las células de los organismos multicelulares han desarrollado esta capacidad en un grado extremo para formar arreglos organizados de células diferenciadas.

En particular, una vez que una célula en un organismo multicelular se compromete a diferenciarse en un tipo celular específico, la elección del destino generalmente se mantiene a través de divisiones celulares posteriores. Esto significa que la célula debe recordar los cambios en la expresión génica, que a menudo son provocados por una señal transitoria.

Este fenómeno de la memoria celular es un requisito previo para la creación de tejidos organizados y para el mantenimiento de tipos de células diferenciadas de manera estable.

Por el contrario, los cambios más simples en la expresión génica tanto en eucariotas como en bacterias suelen ser solo transitorios; el represor de triptófano, por ejemplo, desconecta el operón de triptófano en bacterias solo en presencia de triptófano. Tan pronto como se elimina el aminoácido del medio, los genes se vuelven a activar y los descendientes de la célula no recordarán que sus antepasados hayan estado expuestos al triptófano.

Como los reguladores de la transcripción en eucariotas pueden controlar el inicio de la transcripción a “distancia” a lo largo de la cadena, las secuencias de nucleótidos que controlan la expresión de un gen pueden extenderse a lo largo de tramos largos de ADN. Si bien muchos activadores y represores actúan individualmente, la mayoría de los reguladores de la transcripción eucariotas funcionan como parte de un “comité” de proteínas reguladoras.
La forma en que los grupos de reguladores de la transcripción trabajan juntos para determinar la expresión de un solo gen se llama control combinatorio.

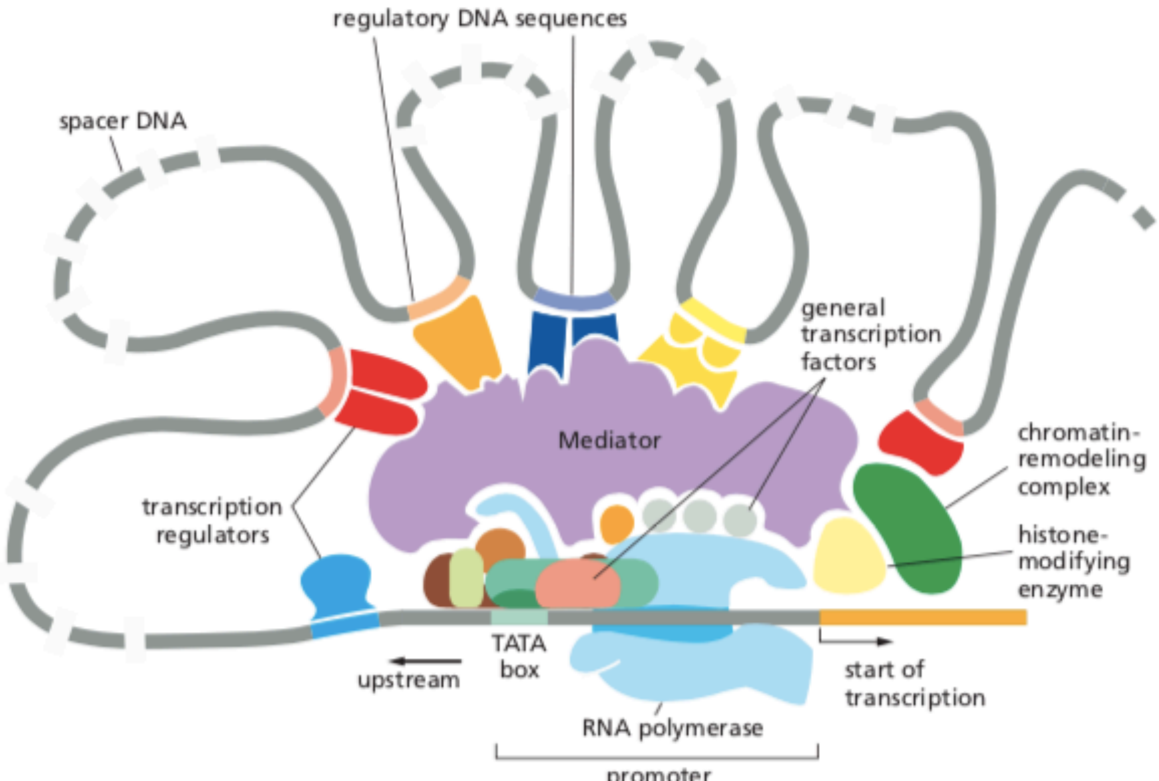

regulatory DNA sequences
spacer DNA
general transcription factors
Mediator
transcription regulators
chromatin-remodeling complex
histone-modifying enzyme
TATA box
upstream
start of transcription
RNA polymerase
promoter

Figure 8–12 Transcription regulators work together as a “committee” to control the expression of a eukaryotic gene. Whereas the general transcription factors that assemble at the promoter are the same for all genes transcribed by RNA polymerase (see Figure 7–12), the transcription regulators and the locations of their DNA binding sites relative to the promoters are different for different genes. These regulators, along with chromatin-modifying proteins, are assembled at the promoter by the Mediator. The effects of multiple transcription regulators combine to determine the final rate of transcription initiation.

La historia de Eve

Even-skipped (Eve, para abreviar) es un gen cuya expresión juega un papel importante en el desarrollo del embrión de Drosophila. Eve no se expresa de manera uniforme a lo largo del embrión: la distribución de la proteína muestra un "pattern" con una llamativa serie de siete franjas ordenadas, cada una de las cuales ocupa una posición muy precisa a lo largo del embrión.

Estas siete franjas corresponden a siete de los catorce segmentos que definen el plan corporal de la mosca: tres para la cabeza, tres para el tórax y ocho para el abdomen.
Este patrón nunca varía: Eve se encuentra en los mismos lugares en cada embrión de Drosophila.

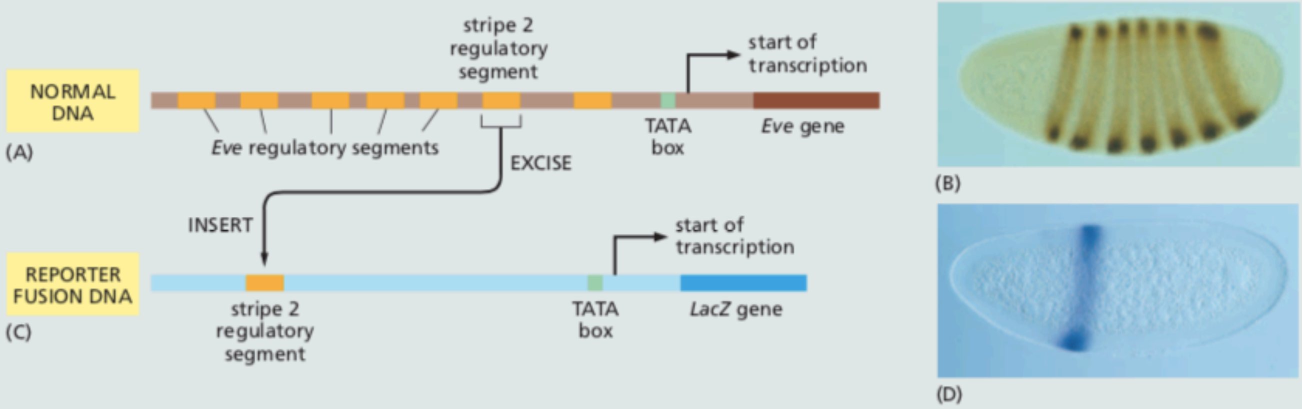

NORMAL DNA
(A)
stripe 2 regulatory segment
start of transcription
Eve regulatory segments
EXCISE
TATA box
Eve gene
INSERT
REPORTER FUSION DNA
(C)
start of transcription
stripe 2 regulatory segment
TATA box
LacZ gene
(B)
(D)

Figure 8–13 An experimental approach that involves the use of a reporter gene reveals the modular construction of the *Eve* gene regulatory region. (A) Expression of the *Eve* gene is controlled by a series of regulatory segments (*orange*) that direct the production of Eve protein in stripes along the embryo. (B) Embryos stained with antibodies to the Eve protein show the seven characteristic stripes of Eve expression. (C) In the laboratory, the regulatory segment that directs the formation of stripe 2 can be excised from the DNA shown in part A and inserted upstream of the *E. coli LacZ* gene, which encodes the enzyme β-galactosidase (see Figure 8–9). (D) When the engineered DNA containing the stripe 2 regulatory segment is introduced into the genome of a fly, the resulting embryo expresses β-galactosidase precisely in the position of the second *Eve* stripe. Enzyme activity is assayed by the addition of X-gal, a modified sugar that when cleaved by β-galactosidase generates an insoluble blue product. (B and D, courtesy of Stephen Small and Michael Levine.)

¿Cómo estudiar la regulación de la expresión?

Una forma de localizar una secuencia de ADN reguladora, y estudiar cómo funciona, es eliminar un fragmento de ADN de la región aguas arriba de un gen de interés e insertar ese ADN aguas arriba de un gen reportero, uno que codifica una proteína cuya actividad o presencia se pueda monitorear. Si el fragmento de ADN contiene una secuencia reguladora, impulsará la expresión del gen reportero. Cuando este fragmento de ADN se introduce posteriormente en una célula u organismo, el gen reportero se expresará en las mismas células y tejidos que normalmente expresan el gen del que se derivó la secuencia reguladora.

Al extraer varios segmentos de las secuencias de ADN aguas arriba de Eve y acoplarlos a un gen indicador, se encontró que la expresión del gen está controlada por una serie de siete módulos reguladores, cada uno de los cuales especifica una sola franja de expresión de Eve. Ver Figura.

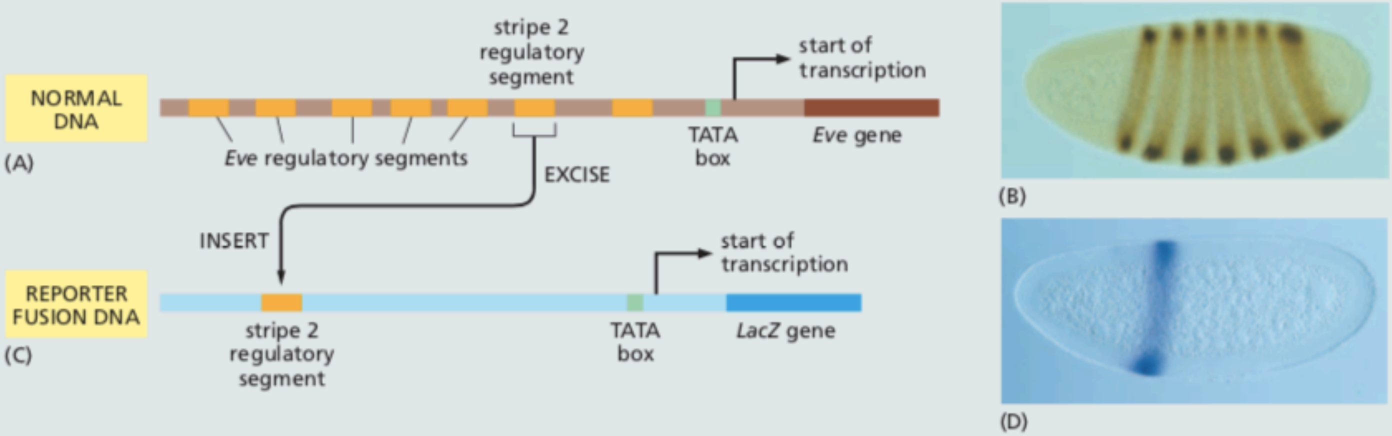


Figure 8–13 An experimental approach that involves the use of a reporter gene reveals the modular construction of the *Eve* gene regulatory region. (A) Expression of the *Eve* gene is controlled by a series of regulatory segments (*orange*) that direct the production of Eve protein in stripes along the embryo. (B) Embryos stained with antibodies to the Eve protein show the seven characteristic stripes of Eve expression. (C) In the laboratory, the regulatory segment that directs the formation of stripe 2 can be excised from the DNA shown in part A and inserted upstream of the *E. coli LacZ* gene, which encodes the enzyme β-galactosidase (see Figure 8–9). (D) When the engineered DNA containing the stripe 2 regulatory segment is introduced into the genome of a fly, the resulting embryo expresses β-galactosidase precisely in the position of the second *Eve* stripe. Enzyme activity is assayed by the addition of X-gal, a modified sugar that when cleaved by β-galactosidase generates an insoluble blue product. (B and D, courtesy of Stephen Small and Michael Levine.)

¿Cómo dirige cada uno de estos siete segmentos reguladores la formación de una sola franja en una posición específica?

Se encontró que cada segmento contiene una combinación única de secuencias reguladoras que se unen a diferentes combinaciones de reguladores de la transcripción. Estos reguladores, como el propio Eve, están distribuidos en patrones únicos dentro del embrión: algunos hacia la cabeza, otros hacia la parte posterior, algunos en el medio.

El segmento regulador que define la banda 2, e.g., contiene secuencias de ADN reguladoras para 4 reguladores de la transcripción: 2 que activan la transcripción de Eve y dos que la reprimen (figura 8-14). En la estrecha banda de tejido que constituye la banda 2, da la casualidad de que las proteínas represoras no están presentes, por lo que se expresa el gen Eve; en las bandas de tejido a ambos lados de la raya, los represores mantienen callada a Eve. Y así se forma una raya.

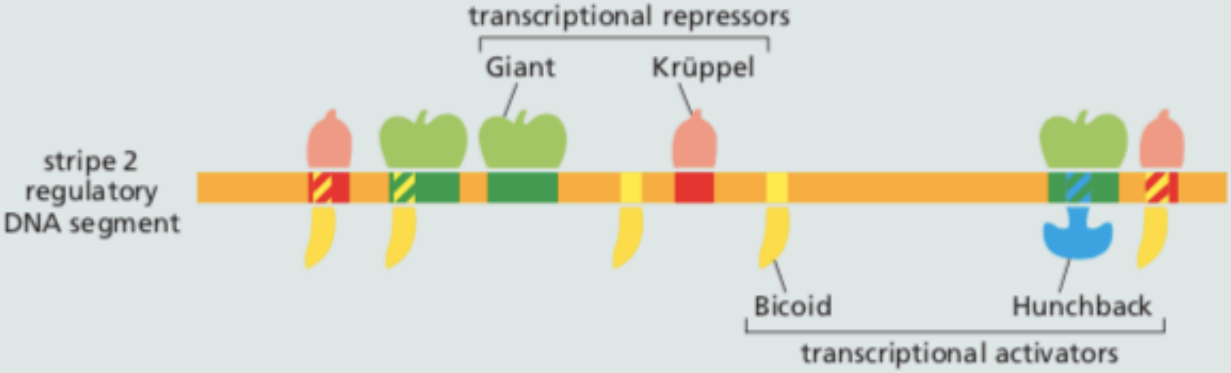

transcriptional repressors
Giant
Krüppel
stripe 2 regulatory DNA segment
Bicoid
Hunchback
transcriptional activators

Figure 8–14 The regulatory segment that specifies Eve stripe 2 contains binding sites for four different transcription regulators. All four regulators are responsible for the proper expression of *Eve* in stripe 2. Flies that are deficient in the two activators, called Bicoid and Hunchback, fail to form stripe 2 efficiently; in flies deficient in either of the two repressors, called Giant and Krüppel, stripe 2 expands and covers an abnormally broad region of the embryo. As indicated in the diagram, in some cases the binding sites for the transcription regulators overlap, and the proteins compete for binding to the DNA. For example, the binding of Bicoid and Krüppel to the site at the far right is thought to be mutually exclusive. The regulatory segment is 480 base pairs in length.

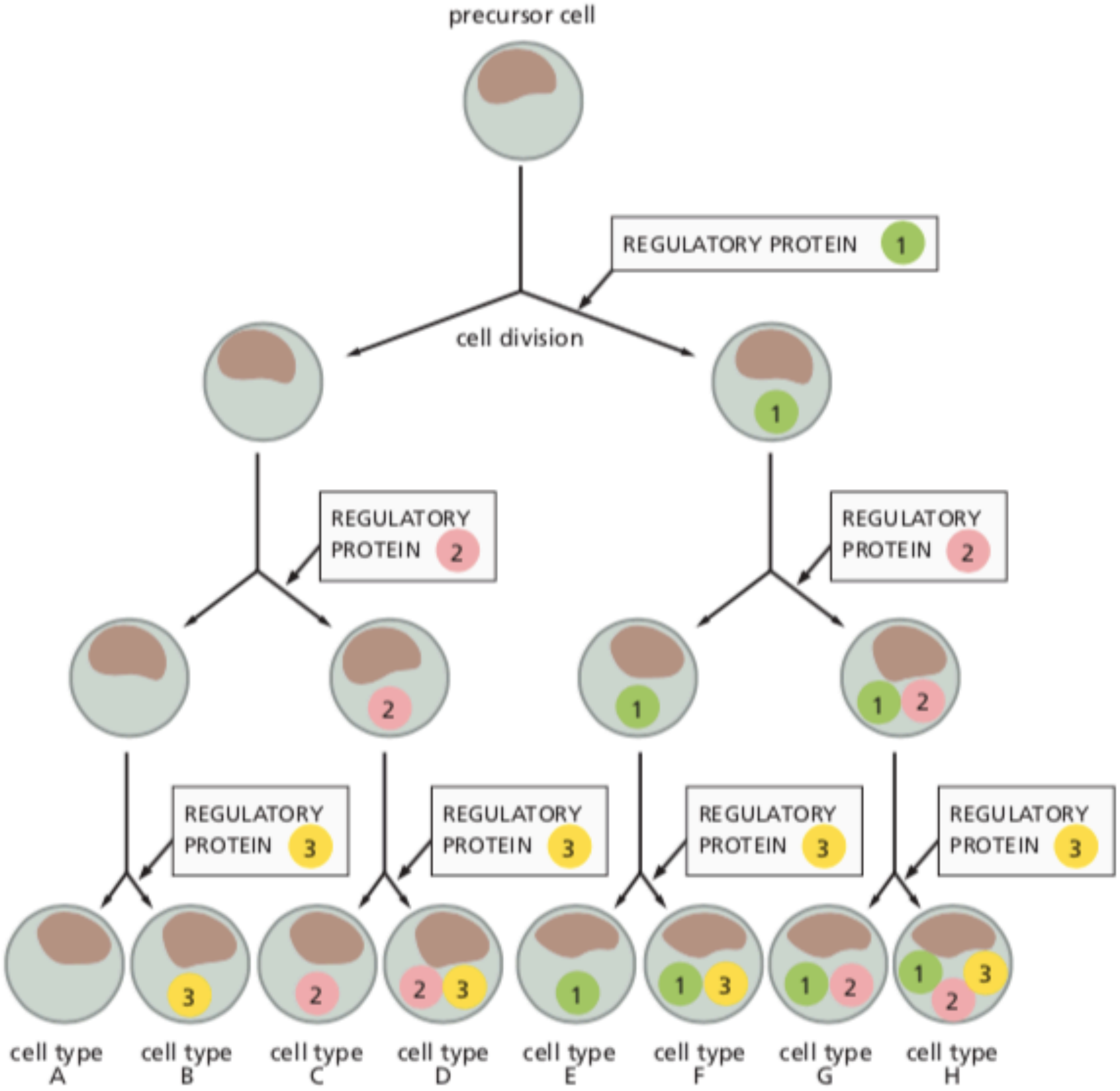

precursor cell
REGULATORY PROTEIN 1
cell division
1
REGULATORY PROTEIN 2
REGULATORY PROTEIN 2
2
1
1 2
REGULATORY PROTEIN 3
REGULATORY PROTEIN 3
REGULATORY PROTEIN 3
REGULATORY PROTEIN 3
3
2
2 3
1
1 3
1 2
1 3 2
cell type A
cell type B
cell type C
cell type D
cell type E
cell type F
cell type G
cell type H

Figure 8–17 Combinations of a few transcription regulators can generate many cell types during development. In this simple scheme, a “decision” to make a new transcription regulator (shown as a *numbered* circle) is made after each cell division. Repetition of this simple rule can generate eight cell types (A through H), using only three transcription regulators. Each of these hypothetical cell types would then express many different genes, as dictated by the combination of transcription regulators that each cell type produces.

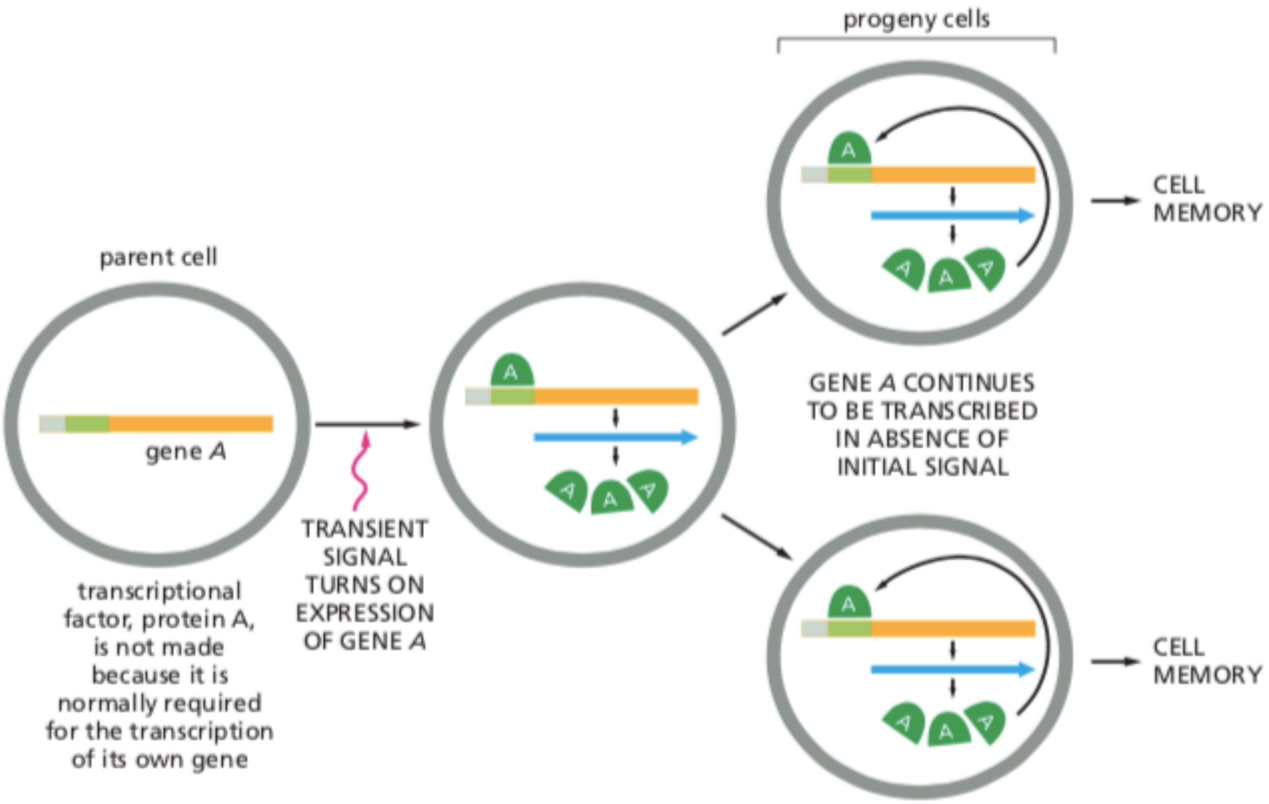

progeny cells
parent cell
A
A
A
A
gene *A*
A
A
A
A
A
A
A
A
A
transcriptional factor, protein A, is not made because it is normally required for the transcription of its own gene
TRANSIENT SIGNAL TURNS ON EXPRESSION OF GENE *A*
GENE *A* CONTINUES TO BE TRANSCRIBED IN ABSENCE OF INITIAL SIGNAL
CELL MEMORY
CELL MEMORY

Figure 8–20 A positive feedback loop can create cell memory. Protein A is a master transcription regulator that activates the transcription of its own gene—as well as other cell-type-specific genes (not shown). All of the descendants of the original cell will therefore "remember" that the progenitor cell had experienced a transient signal that initiated the production of protein A.

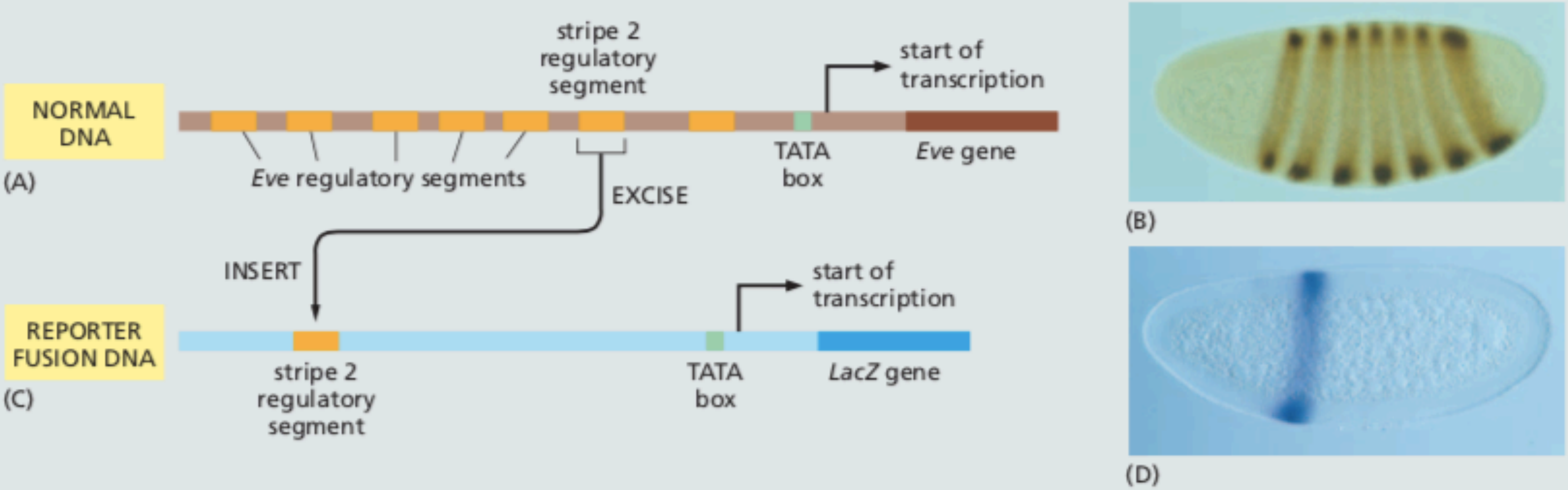

stripe 2
regulatory
segment
start of
transcription
NORMAL
DNA
(A)
Eve regulatory segments
TATA
box
Eve gene
EXCISE
INSERT
start of
transcription
REPORTER
FUSION DNA
(C)
stripe 2
regulatory
segment
TATA
box
LacZ gene
(B)
(D)

Figure 8–13 An experimental approach that involves the use of a reporter gene reveals the modular construction of the *Eve* gene regulatory region. (A) Expression of the *Eve* gene is controlled by a series of regulatory segments (*orange*) that direct the production of Eve protein in stripes along the embryo. (B) Embryos stained with antibodies to the Eve protein show the seven characteristic stripes of Eve expression. (C) In the laboratory, the regulatory segment that directs the formation of stripe 2 can be excised from the DNA shown in part A and inserted upstream of the *E. coli LacZ* gene, which encodes the enzyme β-galactosidase (see Figure 8–9). (D) When the engineered DNA containing the stripe 2 regulatory segment is introduced into the genome of a fly, the resulting embryo expresses β-galactosidase precisely in the position of the second *Eve* stripe. Enzyme activity is assayed by the addition of X-gal, a modified sugar that when cleaved by β-galactosidase generates an insoluble blue product. (B and D, courtesy of Stephen Small and Michael Levine.)

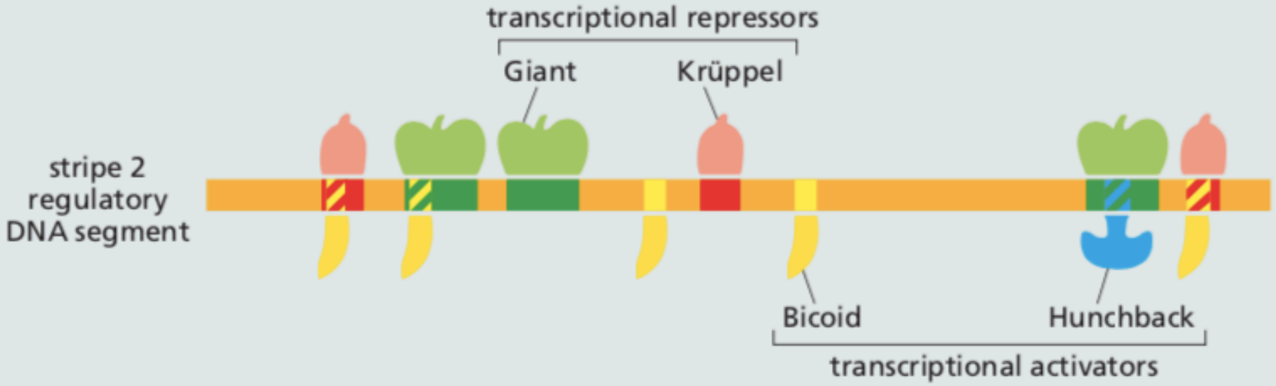

transcriptional repressors
Giant
Krüppel
stripe 2 regulatory DNA segment
Bicoid
Hunchback
transcriptional activators

Figure 8–14 The regulatory segment that specifies Eve stripe 2 contains binding sites for four different transcription regulators. All four regulators are responsible for the proper expression of *Eve* in stripe 2. Flies that are deficient in the two activators, called Bicoid and Hunchback, fail to form stripe 2 efficiently; in flies deficient in either of the two repressors, called Giant and Krüppel, stripe 2 expands and covers an abnormally broad region of the embryo. As indicated in the diagram, in some cases the binding sites for the transcription regulators overlap, and the proteins compete for binding to the DNA. For example, the binding of Bicoid and Krüppel to the site at the far right is thought to be mutually exclusive. The regulatory segment is 480 base pairs in length.

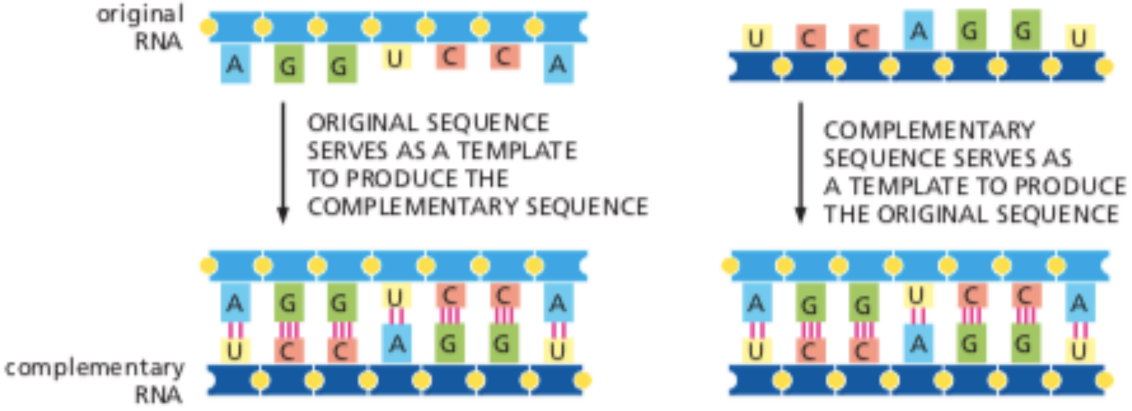

original RNA
A G G U C C A
U C C A G G U
ORIGINAL SEQUENCE SERVES AS A TEMPLATE TO PRODUCE THE COMPLEMENTARY SEQUENCE
COMPLEMENTARY SEQUENCE SERVES AS A TEMPLATE TO PRODUCE THE ORIGINAL SEQUENCE
A G G U C C A
U C C A G G U
complementary RNA

Figure 7–45 An RNA molecule can in principle guide the formation of an exact copy of itself. In the first step, the original RNA molecule acts as a template to form an RNA molecule of complementary sequence. In the second step, this complementary RNA molecule itself acts as a template to form an RNA molecule of the original sequence. Since each template molecule can produce many copies of the complementary strand, these reactions can result in the amplification of the original sequence.

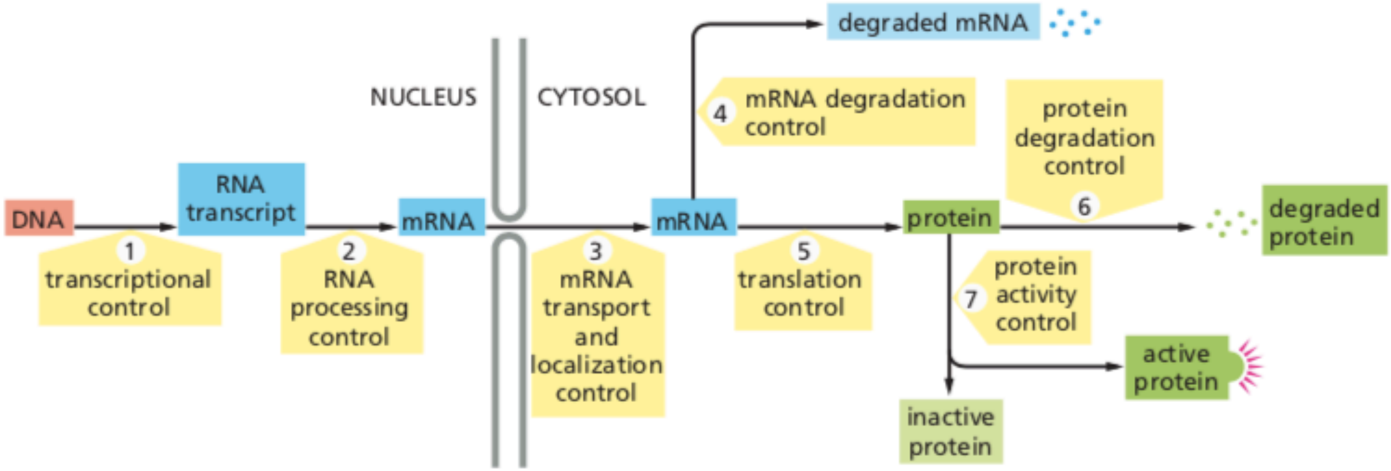
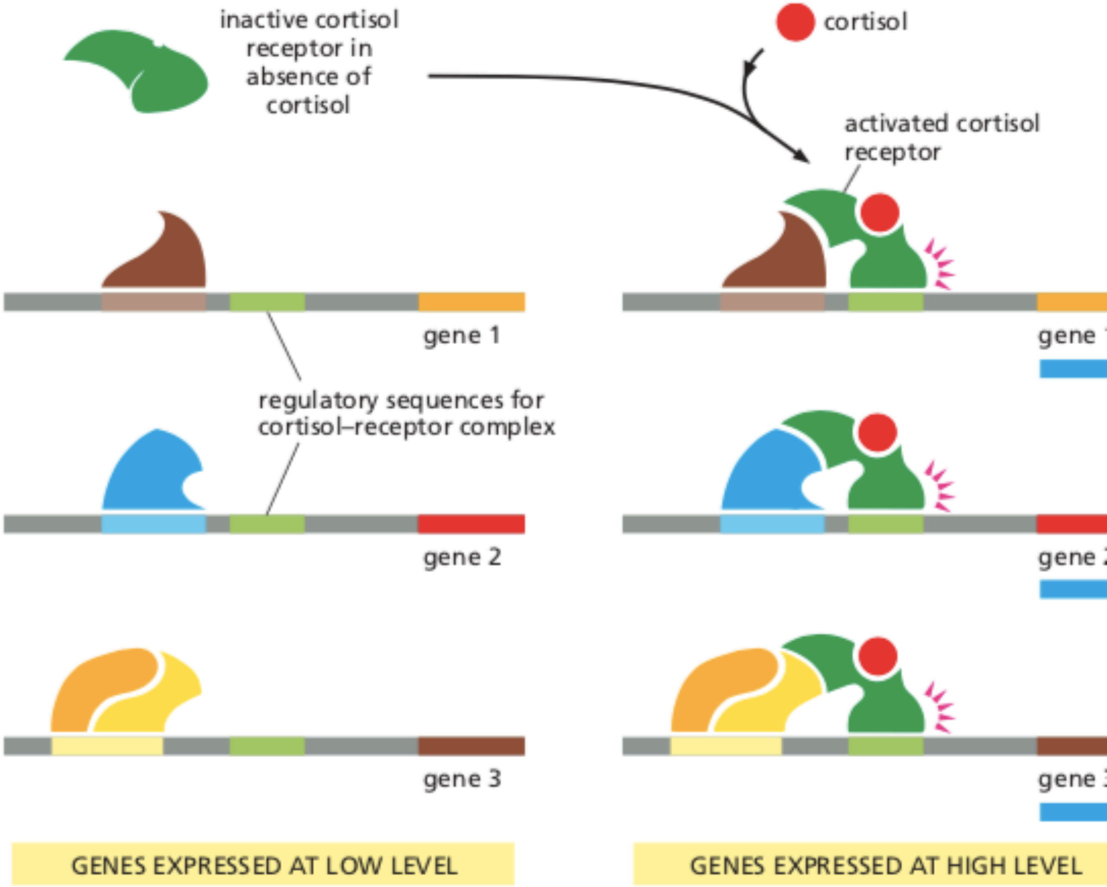

DNA
RNA transcript
mRNA
NUCLEUS
CYTOSOL
mRNA
degraded mRNA
protein
degraded protein
1 transcriptional control
2 RNA processing control
3 mRNA transport and localization control
4 mRNA degradation control
5 translation control
6 protein degradation control
7 protein activity control
active protein
inactive protein

Figure 8–3 Gene expression in eukaryotic cells can be controlled at various steps. Examples of regulation at each of these steps are known, although for most genes the main site of control is step 1—transcription of a DNA sequence into RNA.

Figure 8–15 A single transcription regulator can coordinate the expression of many different genes. The action of the cortisol receptor is illustrated. On the left is a series of genes, each of which has a different gene activator protein bound to its respective regulatory DNA sequences. However, these bound proteins are not sufficient on their own to activate transcription efficiently. On the right is shown the effect of adding an additional transcription regulator—the cortisol–receptor complex—that can bind to the same regulatory DNA sequence in each gene. The activated cortisol receptor completes the combination of transcription regulators required for efficient initiation of transcription, and the genes are now switched on as a set.


inactive cortisol receptor in absence of cortisol
cortisol
activated cortisol receptor
gene 1
regulatory sequences for cortisol–receptor complex
gene 2
gene 3
GENES EXPRESSED AT LOW LEVEL
gene 1
gene 2
gene 3
GENES EXPRESSED AT HIGH LEVEL

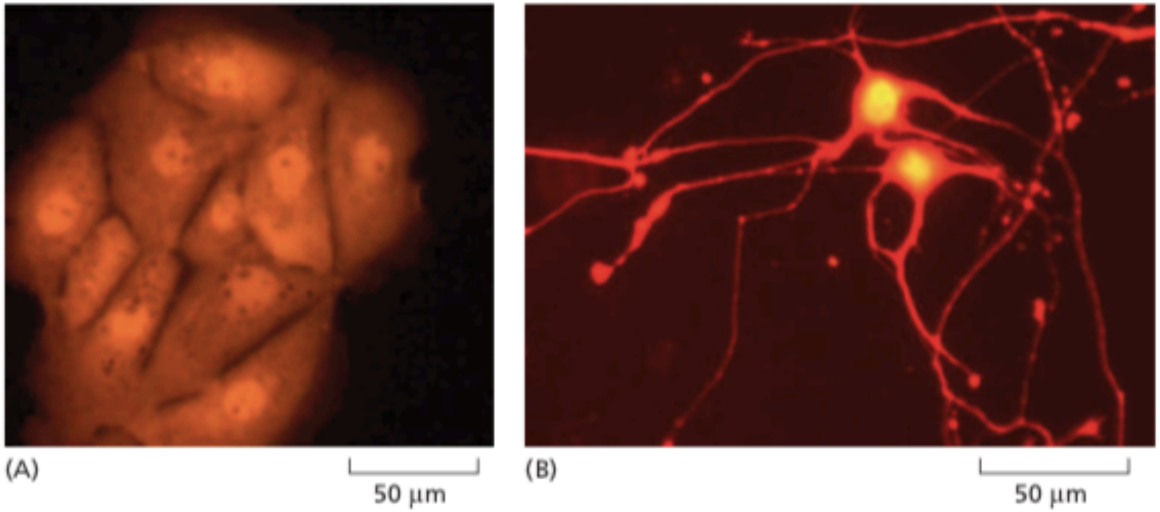


Figure 8–16 A small number of transcription regulators can conve one differentiated cell type direct another. In this experiment, liver cel in culture (A) were converted into ne cells (B) via the artificial introduction nerve-specific transcription regulato cells are labeled with a fluorescent c (From S. Marro et al., *Cell Stem Cell* 378, 2011. With permission from Els

In Situ Hybridization can reveal When and Where a gene is expressed

To see where a particular RNA is made, investigators use a technique called *in situ* hybridization (from the Latin *in situ*, “in place”), which allows a specific nucleic acid sequence— either DNA or RNA—to be visualized in its normal location.

In situ hybridization uses single-stranded DNA or RNA probes, labeled with either fluorescent dyes or radioactive isotopes, to detect complementary nucleic acid sequences within a tissue, a cell or even an isolated chromosome (**Figure 10–29**). The latter application is used in the clinic to determine, for example, whether fetuses carry abnormal chromosomes.

In situ hybridization is frequently used to study the expression patterns of a particular gene or collection of genes in an adult or developing tissue. In one particularly ambitious project, neuroscientists are using the method to assemble a three-dimensional map of all the genes expressed in both the mouse and human brain (**Figure 10–30**). Knowing where and when a gene is expressed can provide important clues about its function.

reporter genes allow specific proteins to be tracked in living cells

For a gene that encodes a protein, the location of the protein within the cell, tissue, or organism yields clues to the gene’s function. Traditionally, the most effective way to visualize a protein within a cell or tissue involved using a labeled antibody. That approach requires the generation of an antibody that specifically recognizes the protein of interest—a process that can be time-consuming and has no guarantee of success.

An alternative approach is to use the regulatory DNA sequences of the protein-coding gene to drive the expression of some type of reporter gene, one that encodes a protein that can be easily monitored by its fluorescence or enzymatic activity. A recombinant gene of this type usually mimics the expression of the gene of interest, producing the

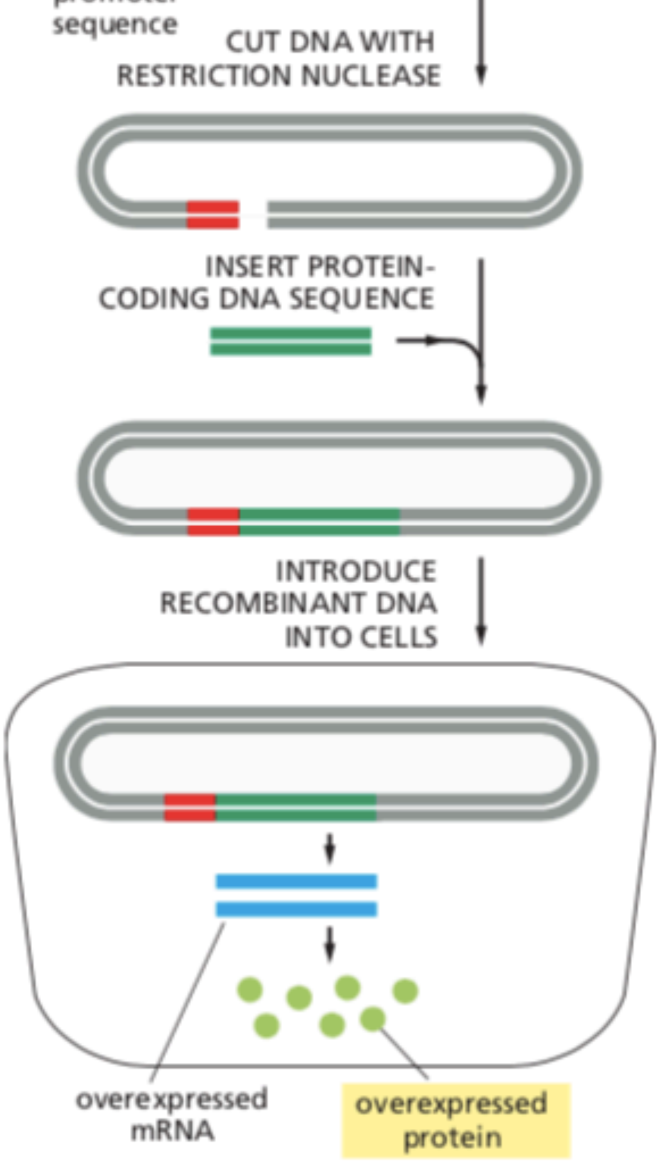

sequence
CUT DNA WITH RESTRICTION NUCLEASE
INSERT PROTEIN-CODING DNA SEQUENCE
INTRODUCE RECOMBINANT DNA INTO CELLS
overexpressed mRNA
overexpressed protein

Figure 10–38 Large amounts of a protein can be produced from a protein-coding DNA sequence inserted into an expression vector and introduced into cells. here, a plasmid vector has been engineered to contain a highly active promoter, which causes unusually large

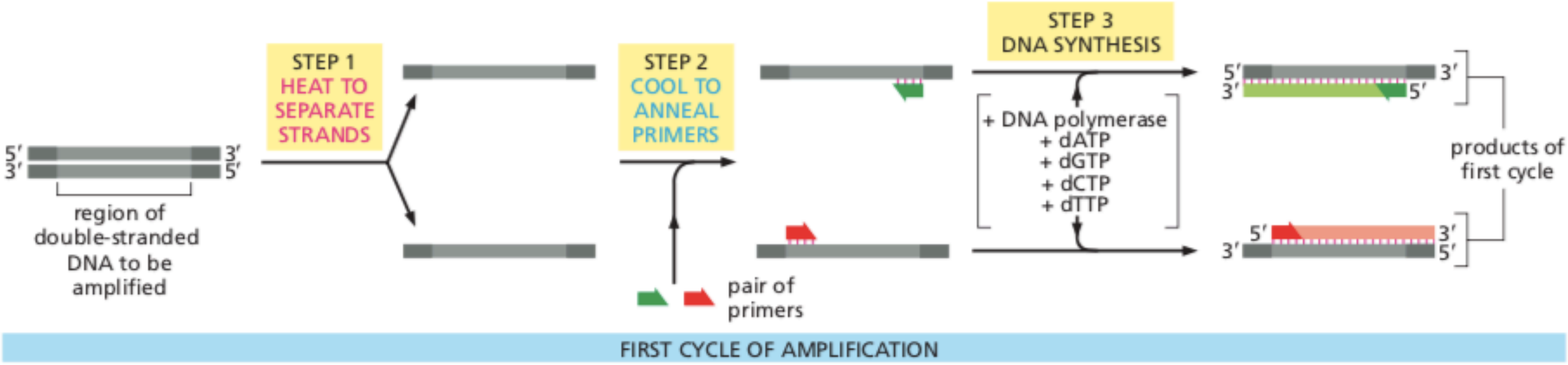

5′
3′
3′
5′
region of double-stranded DNA to be amplified
STEP 1
HEAT TO SEPARATE STRANDS
STEP 2
COOL TO ANNEAL PRIMERS
pair of primers
STEP 3
DNA SYNTHESIS
+ DNA polymerase
+ dATP
+ dGTP
+ dCTP
+ dTTP
5′
3′
3′
5′
5′
3′
3′
5′
products of first cycle
FIRST CYCLE OF AMPLIFICATION

Figure 10–14 A pair of PCR primers directs the amplification of a desired segment of DNA in a test tube. Each cycle of PCR includes three steps: (1) The double-stranded DNA is heated briefly to separate the two strands. (2) The DNA is exposed to a large excess of a pair of specific primers—designed to bracket the region of DNA to be amplified—and the sample is cooled to allow the primers to hybridize to complementary sequences in the two DNA strands. (3) This mixture is incubated with DNA polymerase and the four deoxyribonucleoside triphosphates so that DNA can be synthesized, starting from the two primers. The cycle can then be repeated by reheating the sample to separate the newly synthesized DNA strands (see Figure 10–15).

The technique depends on the use of a special DNA polymerase isolated from a thermophilic bacterium; this polymerase is stable at much higher temperatures than eukaryotic DNA polymerases, so it is not denatured by the heat treatment shown in step 1. The enzyme therefore does not have to be added again after each cycle.

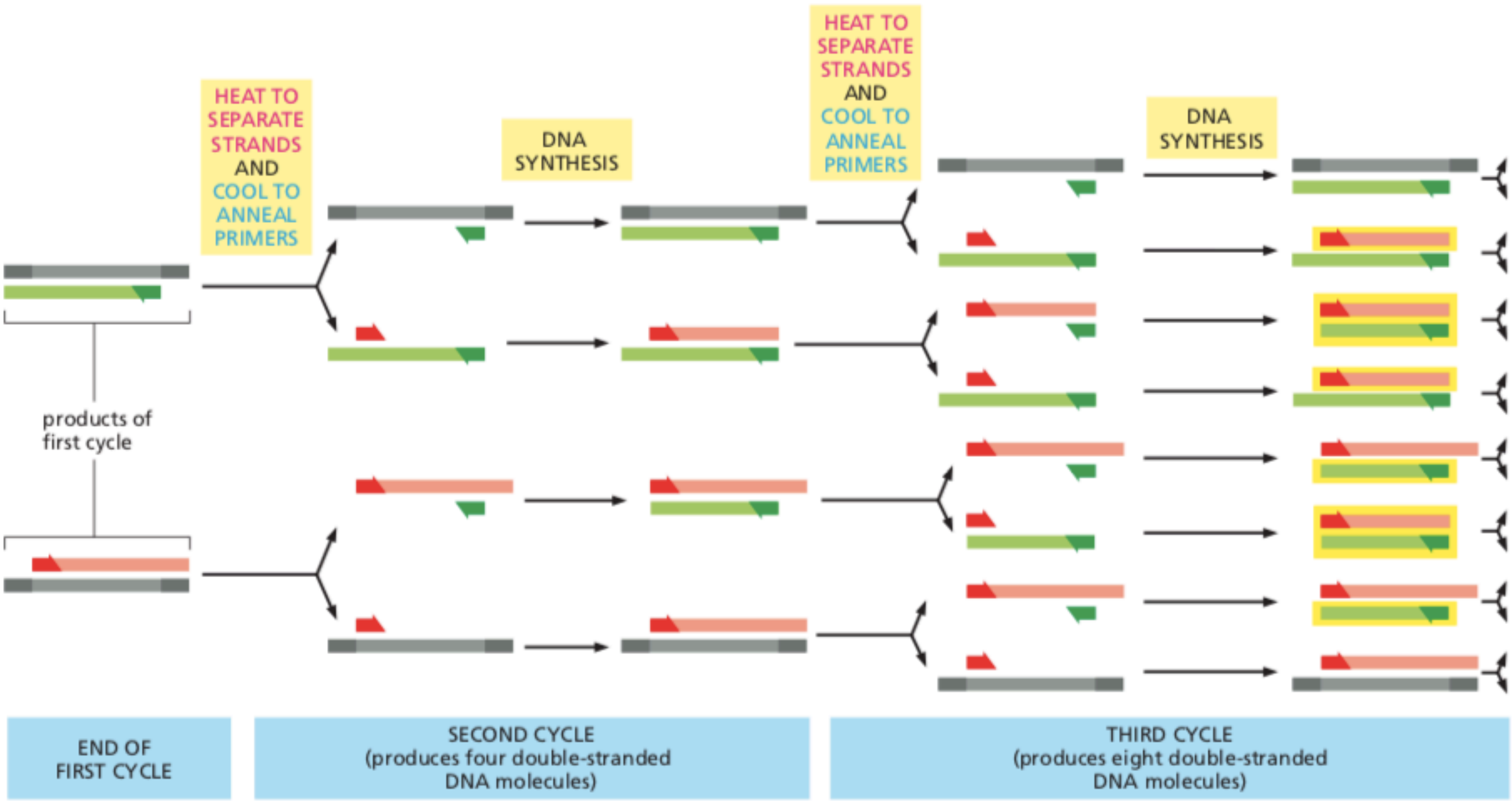

HEAT TO SEPARATE STRANDS AND COOL TO ANNEAL PRIMERS
DNA SYNTHESIS
HEAT TO SEPARATE STRANDS AND COOL TO ANNEAL PRIMERS
DNA SYNTHESIS
products of first cycle
END OF FIRST CYCLE
SECOND CYCLE (produces four double-stranded DNA molecules)
THIRD CYCLE (produces eight double-stranded DNA molecules)

Figure 10–15 PCR uses repeated rounds of strand separation, hybridization, and synthesis to amplify DNA. As the procedure outlined in Figure 10–14 is repeated, all the newly synthesized fragments serve as templates in their turn. Because the polymerase and the primers remain in the sample after the first cycle, PCR involves simply heating and then cooling the same sample, in the same test tube, again and again. Each cycle doubles the amount of DNA synthesized in the previous cycle, so that within a few cycles, the predominant DNA is identical to the sequence bracketed by and including the two primers in the original template. In the example illustrated here, three cycles of reaction produce 16 DNA chains, 8 of which (boxed in *yellow*) correspond exactly to one or the other strand of the original bracketed sequence. After four more cycles, 240 of the 256 DNA chains will correspond exactly to the original sequence, and after several more cycles, essentially all of the DNA strands will be this length. The whole procedure is shown in Movie 10.1.

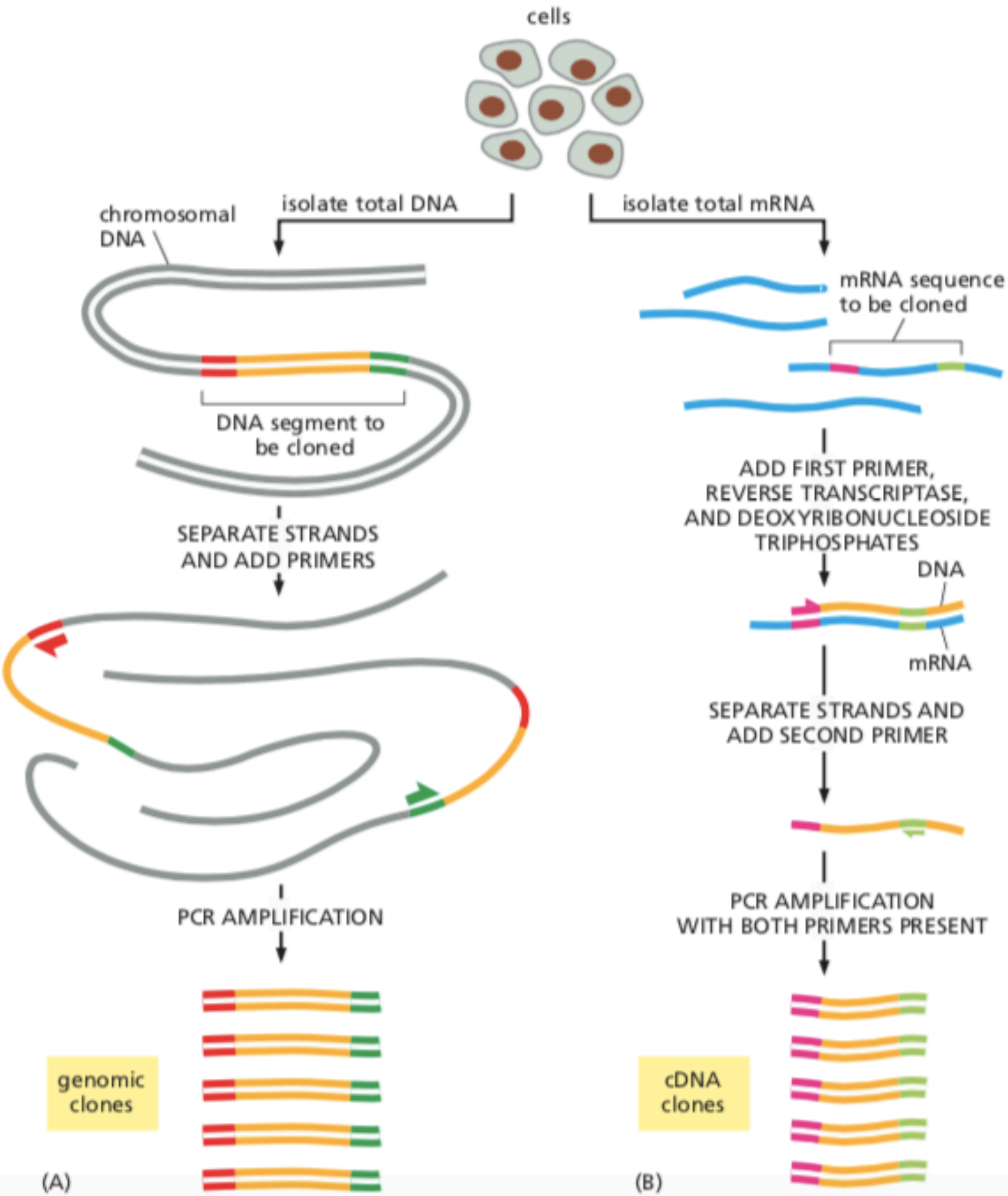

cells
isolate total DNA
isolate total mRNA
chromosomal DNA
DNA segment to be cloned
SEPARATE STRANDS AND ADD PRIMERS
PCR AMPLIFICATION
genomic clones
(A)
mRNA sequence to be cloned
ADD FIRST PRIMER, REVERSE TRANSCRIPTASE, AND DEOXYRIBONUCLEOSIDE TRIPHOSPHATES
DNA
mRNA
SEPARATE STRANDS AND ADD SECOND PRIMER
PCR AMPLIFICATION WITH BOTH PRIMERS PRESENT
cDNA clones
(B)

Figure 10–16 PCR can be used to obtain either genomic or cDNA clones. (A) To use PCR to clone a segment of chromosomal DNA, total DNA is first purified from cells. PCR primers that flank the stretch of DNA to be cloned are added, and many cycles of PCR are completed (see Figure 10–15). Because only the DNA between (and including) the primers is amplified, PCR provides a way to obtain selectively any short stretch of chromosomal DNA in an effectively pure form. (B) To use PCR to obtain a cDNA clone of a gene, total mRNA is first purified from cells. The first primer is added to the population of mRNAs, and reverse transcriptase is used to make a DNA strand complementary to the specific RNA sequence of interest. The second primer is then added, and the DNA molecule is amplified through many cycles of PCR.

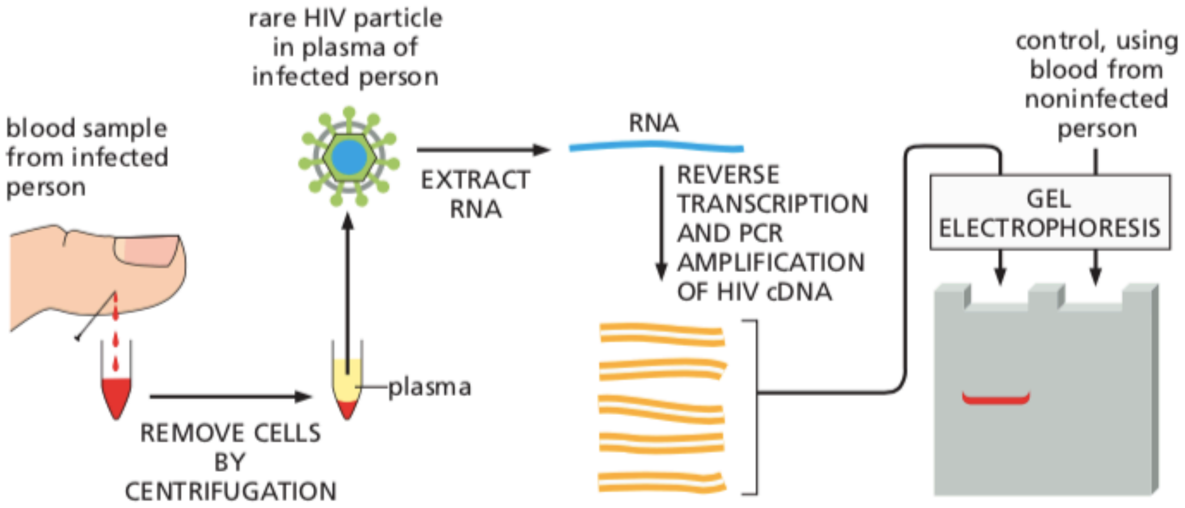

rare HIV particle in plasma of infected person
blood sample from infected person
EXTRACT RNA
RNA
REVERSE TRANSCRIPTION AND PCR AMPLIFICATION OF HIV cDNA
control, using blood from noninfected person
GEL ELECTROPHORESIS
plasma
REMOVE CELLS BY CENTRIFUGATION

Figure 10–17 PCR can be used to detect the presence of a viral genome in a sample of blood. Because of its ability to amplify enormously the signal from every single molecule of nucleic acid, PCR is an extraordinarily sensitive method for detecting trace amounts of virus in a sample of blood or tissue without the need to purify the virus. For HIV, the virus that causes AIDS, the genome is a single-stranded molecule of RNA, as illustrated here. In addition to HIV, many other viruses that infect humans are now detected in this way.

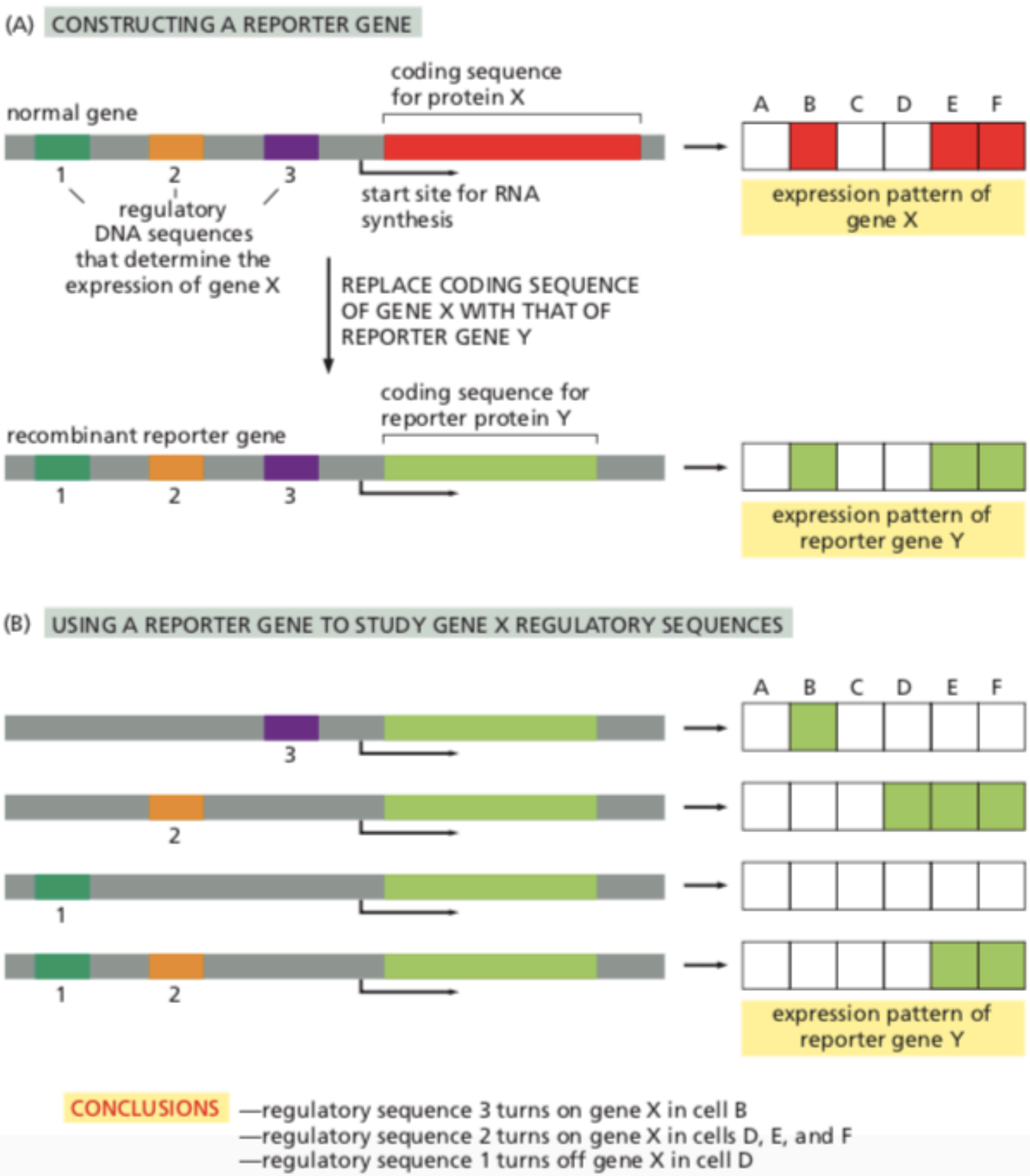

(A) CONSTRUCTING A REPORTER GENE
normal gene
coding sequence for protein X
1
2
3
regulatory DNA sequences that determine the expression of gene X
start site for RNA synthesis
A B C D E F
expression pattern of gene X
REPLACE CODING SEQUENCE OF GENE X WITH THAT OF REPORTER GENE Y
recombinant reporter gene
coding sequence for reporter protein Y
1
2
3
expression pattern of reporter gene Y
(B) USING A REPORTER GENE TO STUDY GENE X REGULATORY SEQUENCES
A B C D E F
3
2
1
1
2
expression pattern of reporter gene Y
CONCLUSIONS
—regulatory sequence 3 turns on gene X in cell B
—regulatory sequence 2 turns on gene X in cells D, E, and F
—regulatory sequence 1 turns off gene X in cell D

Figure 10–31 Reporter genes can be used to determine the pattern of a gene's expression. (A) Suppose the goal is to find out which cell types (A–F) express protein X, but it is difficult to detect the protein directly—with antibodies, for example. Using recombinant DNA techniques, the coding sequence for protein X can be replaced with the coding sequence for reporter protein Y, which can be easily monitored visually; two commonly used reporter proteins are the enzyme β-galactosidase (see Figure 8–13C) and green fluorescent protein (GFP, see Figure 10–32). The expression of the reporter protein Y will now be controlled by the regulatory sequences (here labeled 1, 2, and 3) that control the expression of the normal protein X. (B) To determine which regulatory sequences normally control expression of gene X in particular cell types, reporters with various combinations of the regulatory regions associated with gene X can be constructed. These recombinant DNA molecules are then tested for expression after their introduction into the different cell types.