

tRNA: Structure & Function

Major players in *protein synthesis*:
mRNA, tRNA and the ribosome

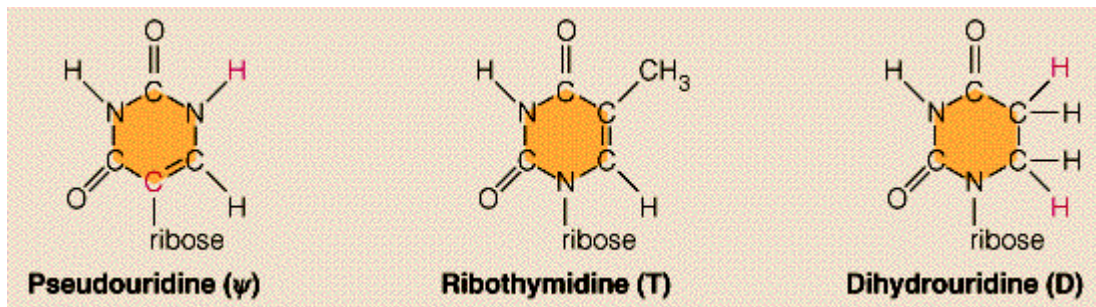
mRNA	Messenger RNA , a copy of DNA blueprint of the <i>gene</i> to be expressed.	Information
tRNA	Aminoacyl transfer RNA , also called <i>anticodon</i> or <i>adaptor</i> molecule. One or more tRNAs for each amino acid.	Supply
Ribosome	A very large complex of several rRNAs (ribosomal RNA) and many protein molecules. Total molecular weight over 2 million dalton.	Factory
Protein	Polypeptide chain with sequence dictated by the mRNA sequence. Also called the <i>gene product</i> .	Product

Estructura de los tRNA Primaria & Secundaria

-Todos los tRNA, cuando se escriben en una forma en la que existe el máximo apareamiento de bases intracatenarios mediante los pares permitidos A-U, G-C y G-U, forman una estructura en forma de hoja de trébol con cuatro brazos.

- Son cadenas sencillas que contienen entre 73 y 95 nucleótidos (25 Kd)

-Contienen muchas bases poco comunes, normalmente entre 7 y 15 por molécula. Estas modificaciones tienen un sentido funcional: **alquilación (metilación)** - impedir ciertos apareamientos, aumentar la hidrofobicidad para interactuar con las sintetasas u otras proteínas y para el plegamiento.

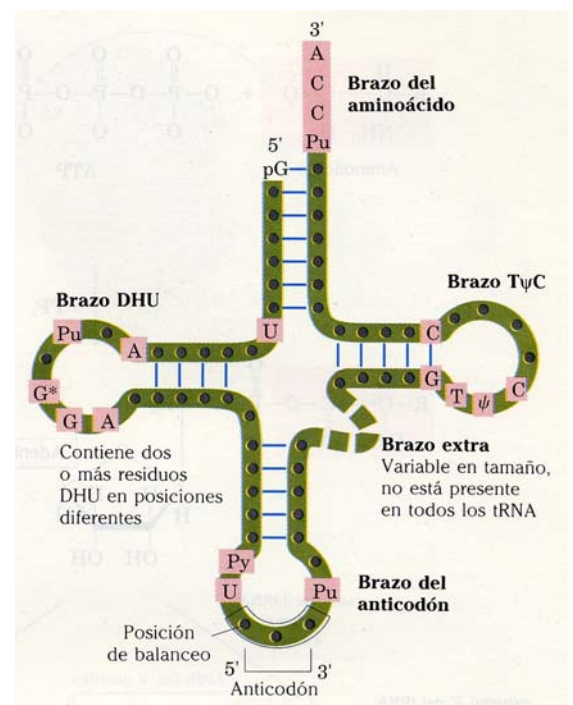
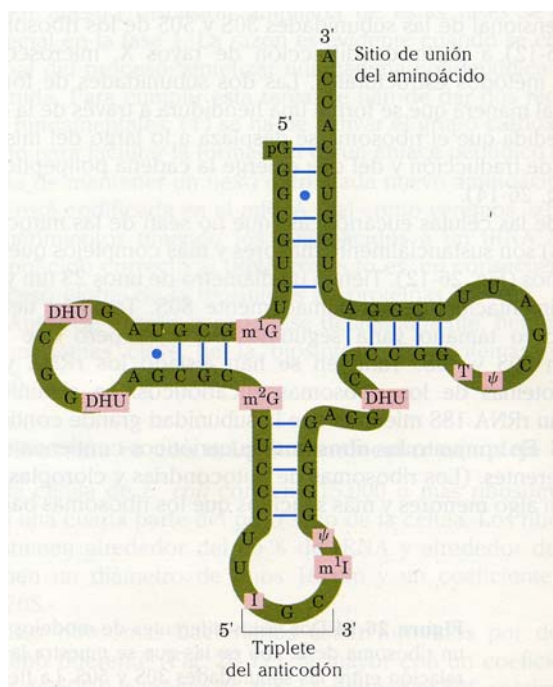


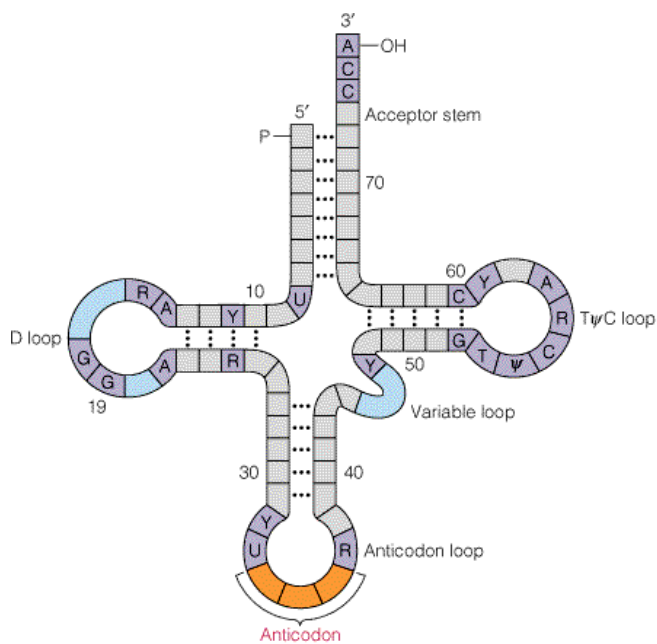
-El extremo 5' de los tRNA está fosforilado y generalmente suele ser pG

-La secuencia de bases del extremo 3' de las moléculas de tRNA maduras es CCA. El aminoácido activado está unido a un grupo hidróxilo 3' de la adenosina terminal

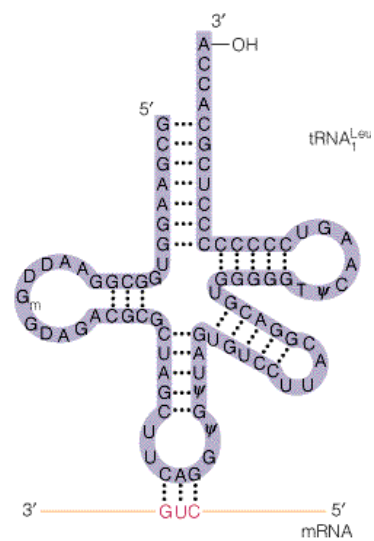
-Alrededor de la mitad de los nucleótidos en el tRNA tienen bases apareadas para formar dobles hélices. Cinco grupos de bases no están emparejadas: **la región terminal 3'-CCA**; **el lazo T ψ C, (ribotimidina, pseudouridina y citidina)**; **el brazo adicional**; **el lazo DHU (que contiene varios residuos dihidrouracilo)**, y **el lazo anticodón**.

-El lazo anticodón, consta de siete bases, con la secuencia siguiente.
pirimidina-pirimidina-XYZ-purina modificada-base variable

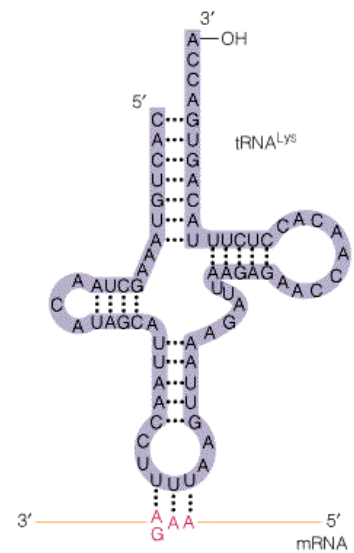




(a)



(b)



(c)

Estructura de los tRNA. (a) Estructura general de los tRNA. Las posiciones de las bases invariables o que rara vez varían se muestran en morado. Las regiones del bucle D y del bucle variable que pueden contener diferentes números de nucleótidos se muestran en azul. El anticodón se muestra en naranja. (b) un tRNA de leucina de E.coli. (c) Un tRNA mitocondrial de lisina humano. Códigos de las bases: Y = pirimidina, R = purina, ψ = pseudouridina, T = ribotimidina y D = dihidrouridina.

Estructura terciaria

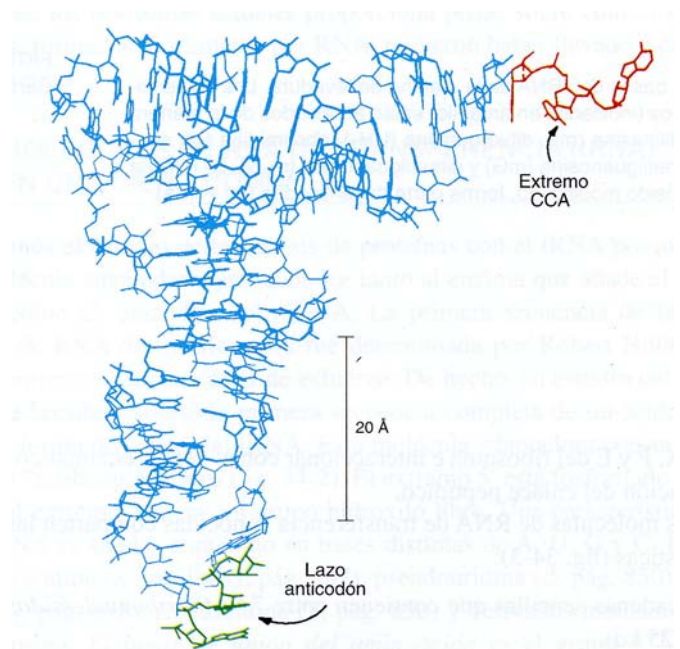
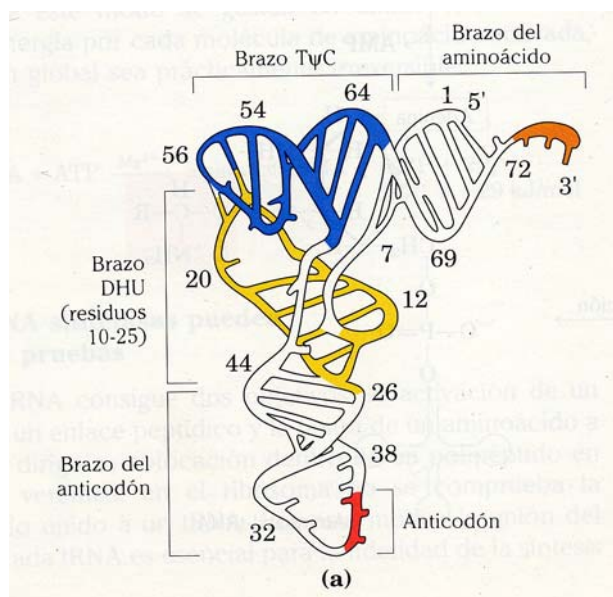
-La molécula tiene forma de L

-Hay dos segmentos de doble hélice del tipo de A-DNA. Cada una contiene alrededor de 10 bp.

-La mayoría de las bases de regiones no helicoidales participan en interacciones de puentes de hidrógeno poco frecuentes. Estas interacciones terciarias se establecen entre bases que normalmente no son complementarias. Además, el esqueleto de ribosa-fosfato interacciona con algunas bases e incluso con otras regiones del propio esqueleto.

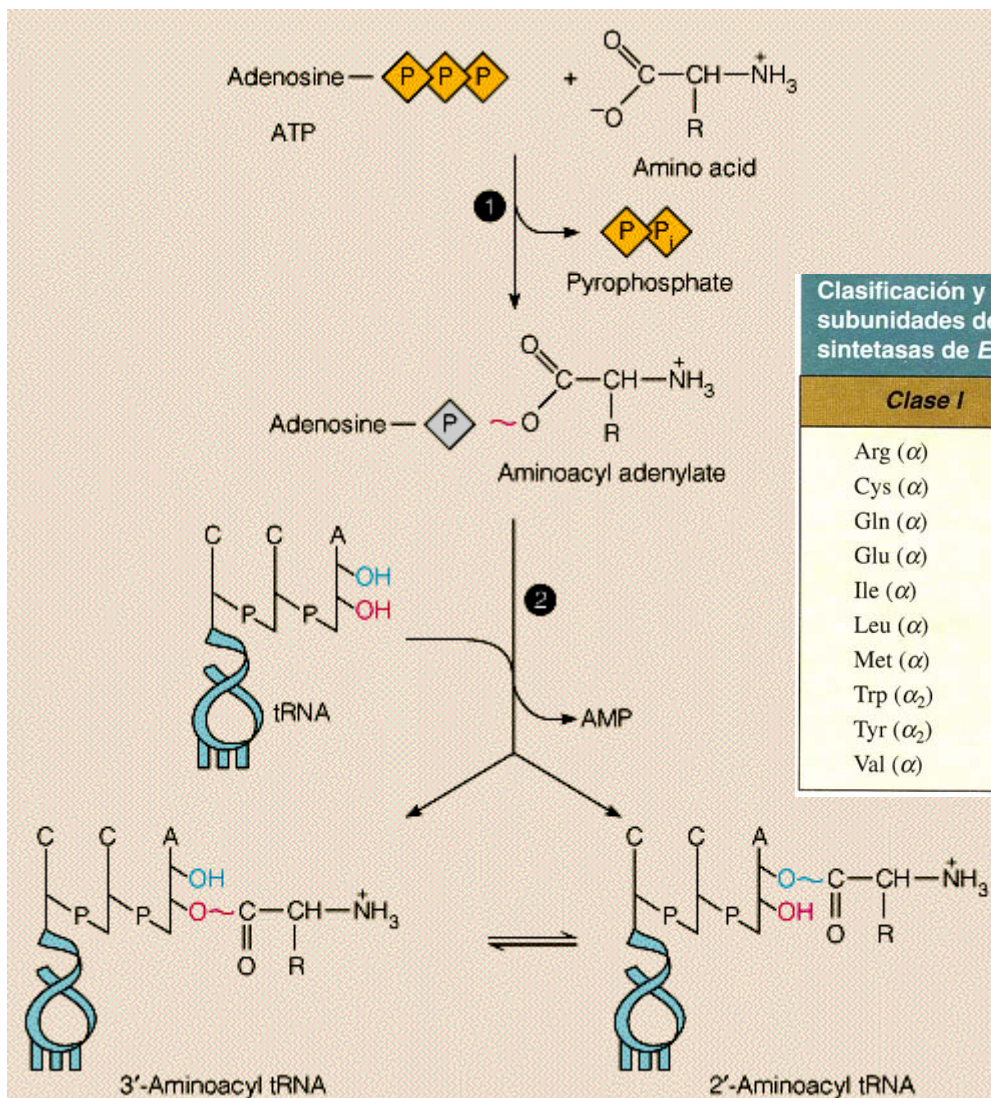
-El extremo CCA que contiene el lugar de unión del aminoácido está en un extremo de la L, mientras que el lazo anticodón está en el otro. Los lazos DHU y T ψ C forman el ángulo de la L.

-El extremo CCA y la región helicoidal adyacente no interaccionan fuertemente con el resto de la molécula. Esta parte de la molécula puede cambiar de conformación durante la activación de los aminoácidos y durante la síntesis de proteínas en el ribosoma.



tRNA Function: Synthetases

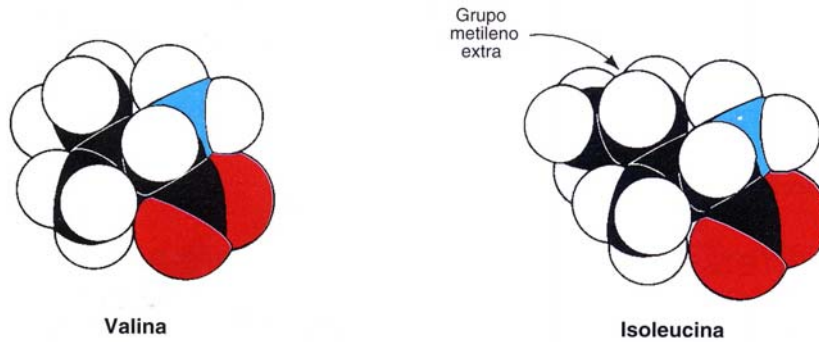
- Each tRNA is charged with the proper amino acid via a covalent ester bond at their 3' end by a family of enzymes called aminoacyl-tRNA synthetases. Each enzyme must recognize both the tRNA specific for an amino acid and the corresponding amino acid. This energy-consuming process is ATP-dependent and results in the cleavage of two high-energetic phosphate bonds.
- There are 20 different aminoacyl-tRNA synthetases, one for each amino acid. Despite the fact that they all carry out very similar tasks, they vary greatly in size.
- Since there are 61 amino acid codons, most tRNA synthetases must be able to recognize more than one type of tRNA (i.e. 6 codons for Arg). These tRNA are called cognate tRNAs for that particular synthetase. This mapping is achieved through so-called recognition domains on the tRNA.
- The recognition domain includes unique of sections of the acceptor stem and/or the anticodon



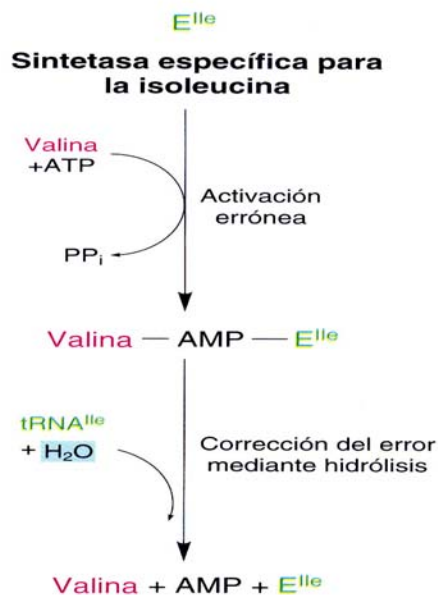
Clasificación y estructura de las subunidades de las aminoacyl-tRNA sintetasas de *E. coli*

Clase I	Clase II
Arg (α)	Ala (α_4)
Cys (α)	Asn (α_2)
Gln (α)	Asp (α_2)
Glu (α)	Gly ($\alpha_2\beta_2$)
Ile (α)	His (α_2)
Leu (α)	Lys (α_2)
Met (α)	Phe ($\alpha_2\beta_2$)
Trp (α_2)	Ser (α_2)
Tyr (α_2)	Pro (α_2)
Val (α)	Thr (α_2)

La función correctora de las aminoacil-tRNA sintetasas aumenta la fidelidad de la síntesis de proteínas.

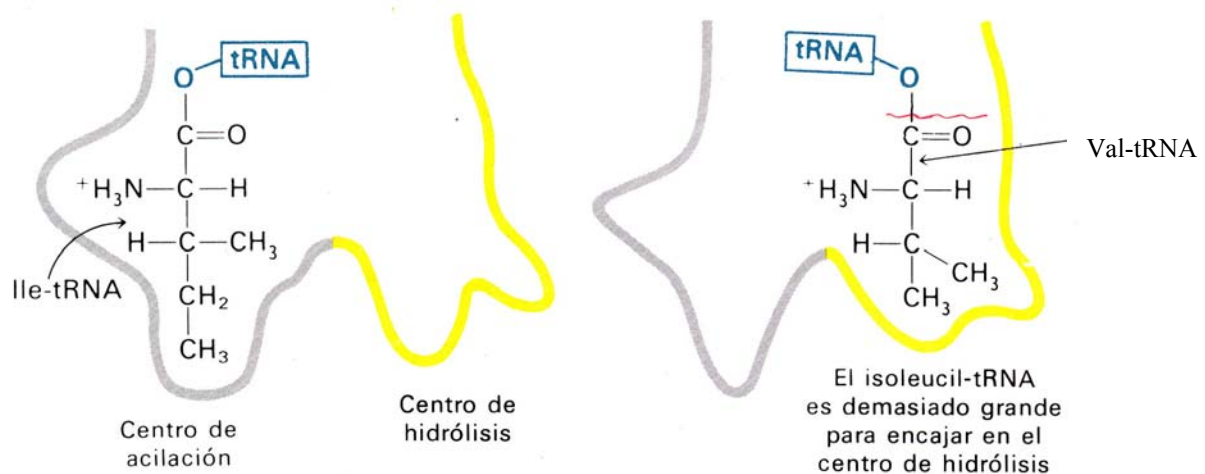


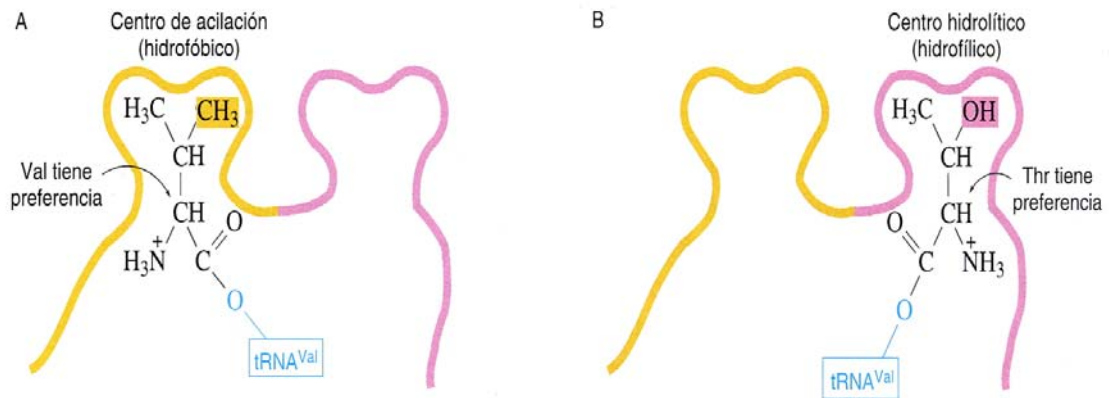
Modelos moleculares de valina e isoleucina. Se señala el grupo metilo extra que tiene la isoleucina. Las sintetasas propias de estos aminoácidos son capaces de diferenciarlos perfectamente.



Corrección por hidrólisis de un aminoacil-AMP incorrecto. La entrada del tRNA específico para la isoleucina induce la hidrólisis de valil-AMP.

Mecanismo de doble tamizado para rechazar los aminoácidos que son menores o mayores que el adecuado. El centro catalítico para la unión del aminoácido al fragmento AMP es el primer tamiz, el centro hidrolítico corrector es el segundo.





La valil-tRNA sintetasa rechaza a la treonina en dos etapas. (A) El centro de acilación hidrofóbico donde se transfiere el aminoácido activado al tRNA prefiere la valina a la treonina porque la valina es más hidrofóbica. (B) El centro de hidrólisis adyacente, por el contrario, es hidrofílico. Por lo tanto, la treonina unida por error al tRNA se hidrolizará preferentemente, ya que la energía de enlace del grupo hidroxilo de la treonina se utiliza para estabilizar el estado de transición.

Mecanismo propuesto para la formación de tirosil-AMP por acción de la tirosil-tRNA sintetasa. El átomo de fósforo α del ATP resulta atacado nucleofílicamente por el oxígeno del carboxilato de la tirosina. El estado de transición es estabilizado por puentes de hidrógeno entre el fosfato γ y las cadenas laterales de un residuo de treonina y otro de histidina. El pirofosfato unido al Mg^{2+} (no representado) abandona el estado de transición pentacovalente y se forma tirosil-AMP. Ado es la abreviatura de adenosina. [Según R.J. Leatherbarrow, A.R. Fersht y G. Winter. *Proc. Nat. Acad. Sci.* 82(1985):7841.]

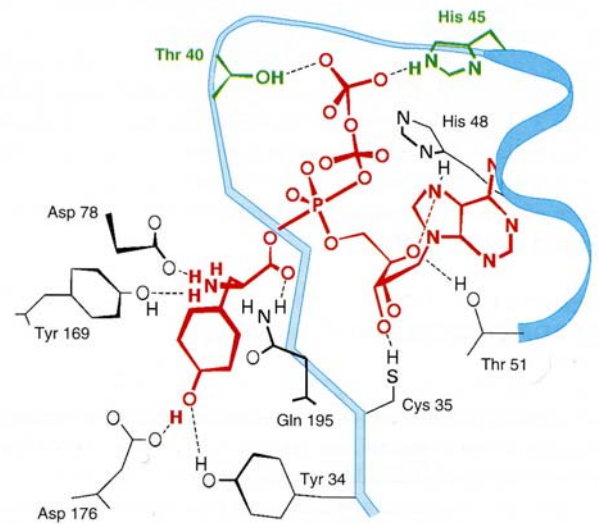
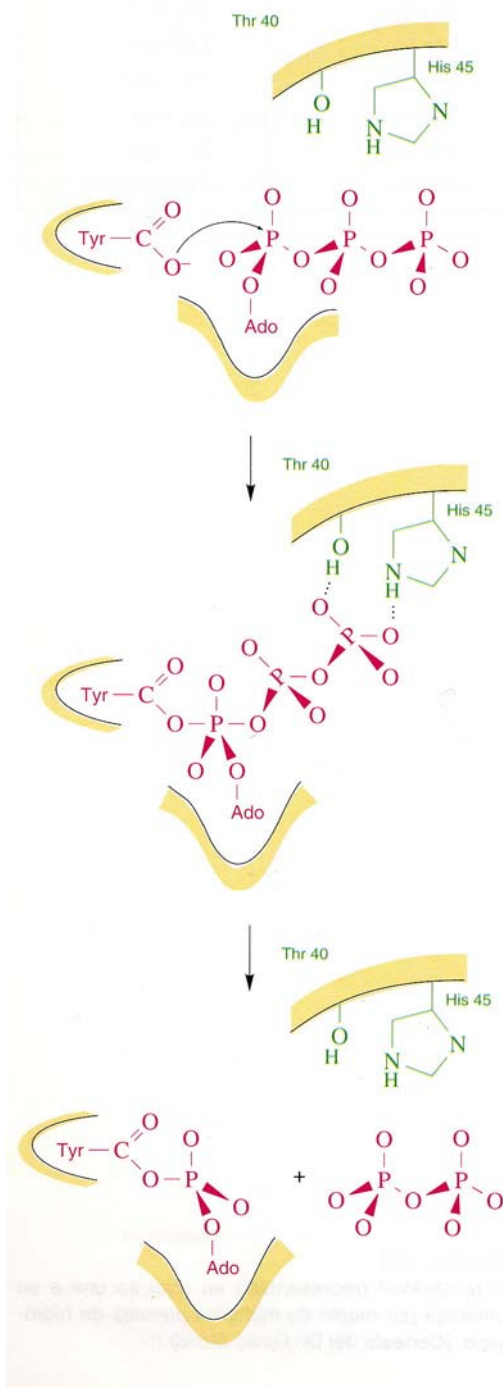
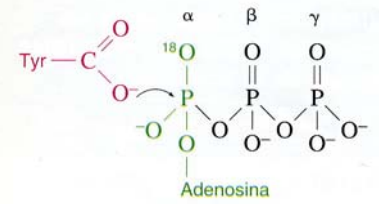
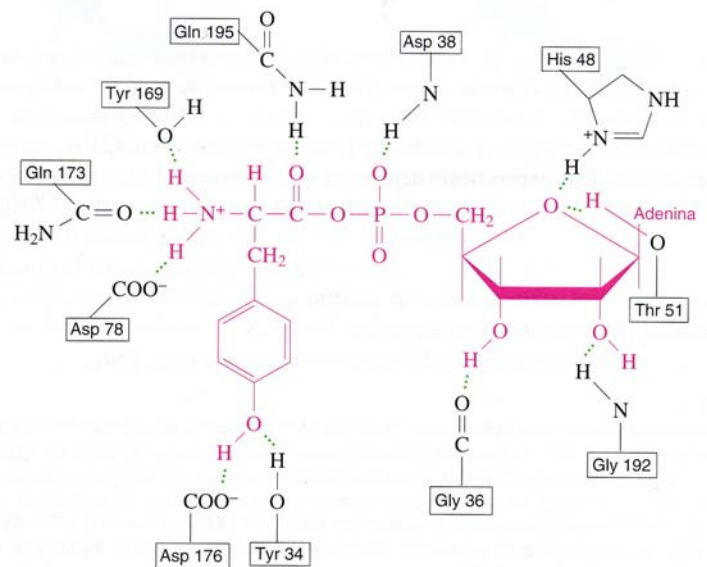
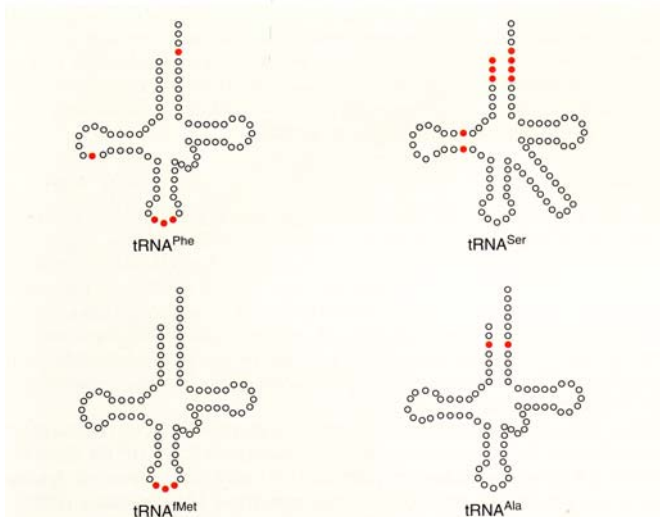


FIGURA 34-9

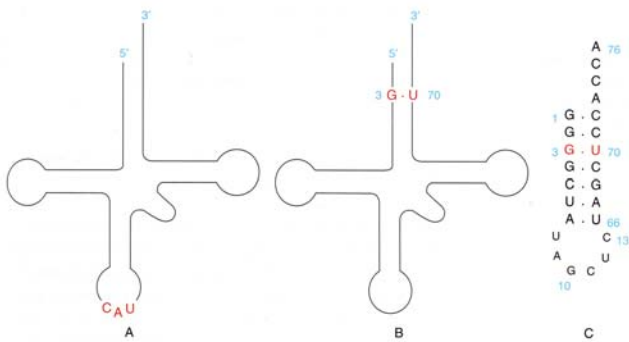
Estructura propuesta para el estado de transición durante la formación de tirosil-AMP a partir de tirosina y ATP (marcados en rojo). Las cadenas laterales de la treonina 40 y la histidina 45 (representadas en verde) juegan un papel crítico en la catálisis porque se unen selectivamente al fosfato γ en el estado de transición. [Según R.J. Leatherbarrow, A.R. Fersht y G. Winter. *Proc. Nat. Acad. Sci.* 82(1985):7841.]



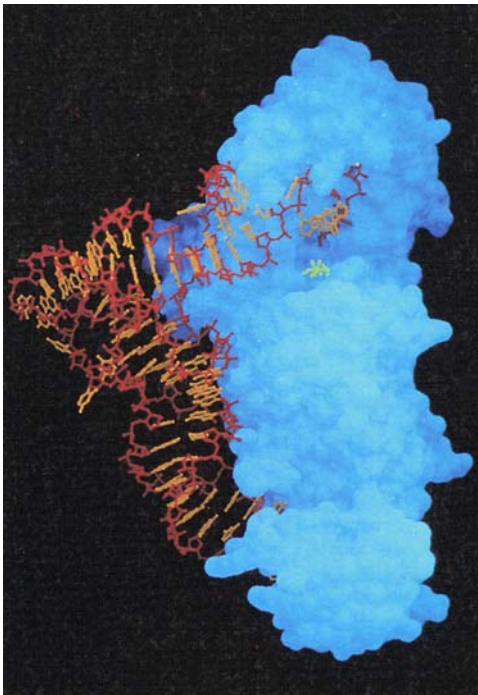
Las sintetasas reconocen el lazo anticodón y el tallo aceptor de las moléculas de RNA de transferencia.



Principales “elementos de identidad” en algunos tRNA. Los círculos rojos indican las posiciones en la hoja de trébol que se ha demostrado identifican el tRNA para su sintetasa relacionada

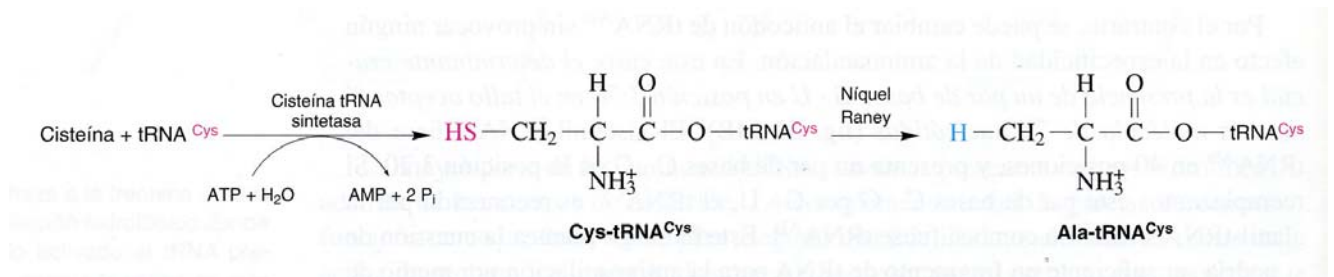


(A) La identidad del tRNA^{Val} cambia de Val a Met al mutar su anticodón de UAC a CAU. (B) La identidad del tRNA^{Cys} cambia de Cys a Ala al cambiar su par de bases 3:70 de C.G a G.U. (C) Una “microhélice” que contiene únicamente 24 de los 76 nucleótidos del tRNA^{Ala} es reconocida por su correspondiente sintetasa.



Modelo de la glutamil-tRNA sintetasa de E.coli acoplada con su tRNA y el ATP. El tRNA se representa mediante un modelo atómico detallado, la proteína mediante su superficie accesible al disolvente (azul). El ATP (en verde) y el tallo aceptor 3' del tRNA se ajustan en una hendidura profunda de la sintetasa. Esta hendidura acomodará también al aminoácido, Esta es una sintetasa monoérica de clase I.

El codón es reconocido por el anticodón del tRNA y no por el aminoácido activado.

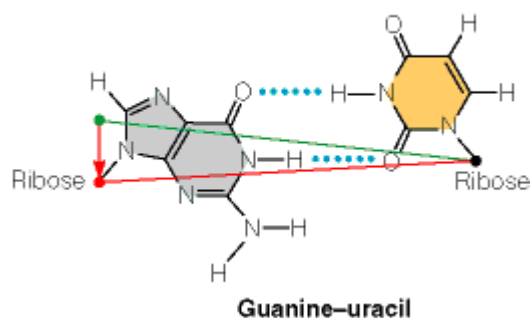
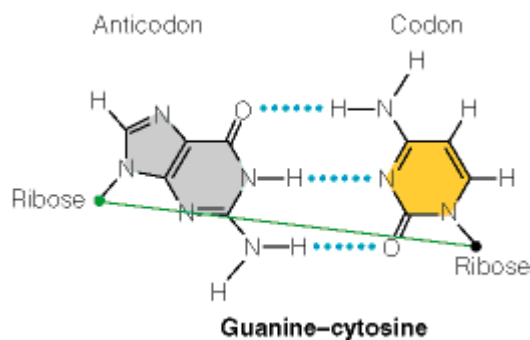


Algunas moléculas de tRNA reconocen más de un codón a causa del balanceo en el apareamiento de bases.

Hipótesis del balanceo (Crick)

- Las primeras dos bases de un codón del mRNA siempre forman pares de bases tipo Watson-Crick fuertes con los pares de bases del anticodón en el tRNA confiriéndole la parte más importante de la especificidad de código
- La primera base de algunos anticodones determina el número de codones leídos por un tRNA dado (ver tabla)
- Cuando un aminoácido es especificado por varios codones diferentes, los codones que difieren en cualquiera de las dos primeras bases requieren tRNA diferentes
- Para traducir los 61 codones se requiere un mínimo de 32 tRNA

El código es redundante. Varios codones pueden corresponder a un solo aminoácido, a veces mediante un balanceo en la posición 5' del anticodón.



Base at 5' Position in Anticodon		Base at 3' Position in Codon
G	pairs with	C or U
C	pairs with	G
A	pairs with	U
U	pairs with	A or G
I	pairs with	A, U, or C

Hipótesis del balanceo. Como ejemplo, presentamos la forma en que la base G del anticodón puede aparearse con C o con U de un codón. El movimiento ("balanceo") de la base en la posición 5' del anticodón es necesario para que exista esta capacidad (véase la flecha).

The Dangers of Wobble

Since a single tRNA can respond to more than one codon, one tRNA could *in theory* respond to codons for *two different* amino acids:

Example: AGC and AGU both code for Ser. Their anticodon is UCG using wobble rules. But anticodon UCA would also pair with AGU (Ser). This anticodon (UCA) would be converted post-transcriptionally to UCI (**Hey, that's us!**) which would recognize AGC and AGU as intended, but also AGA which codes for Arg -- not good, would result in an ambiguous Genetic Code!

One anticodon for serine is UCG:

Codon (5'->3')	AGC	AGU
Anticodon (3'->5')	UCG	UCG (wobble)
Amino acid	Ser	Ser

Now, in the following *hypothetical* example, anticodon UCA (*if it existed*) would also base-pair with codon AGU using wobble rules:

AGU	AGU	AGC	AGU	AGA
UCA =>	UCI (wobble)	UCI	UCI	UCI (wobble)
Ser	Ser	Ser	Ser	Arg

But anticodon UCA would be converted post-transcriptionally to UCI which is able to bind three codons (wobble rules), including one for Arg! For this reason the UCA/UCI anticodon does not exist! And the codon AGA for Arg is actually covered by the anticodon UCU (wobble) !

The Ribosome

Ribosomes can be found either free in the cytosol (cytoplasm) or attached to intracellular membranes.

Free ribosomes

- Found in the cytosol.
- May occur as a single ribosome or in groups known as polyribosomes or polysomes.
- Occur in greater number than bound ribosomes in cells that retain most of their manufactured protein.
- Responsible for proteins that go into solution in the cytoplasm or form important cytoplasmic structural or motile elements.

Bound ribosomes

- Found bound to the exterior of the endoplasmic reticulum (ER) constituting the rough ER.
 - Occur in greater number than free ribosomes in cells that secrete their manufactured proteins (e.g. pancreatic cells, producers of digestive enzymes).
 - Responsible for proteins that insert into membranes or are packaged into vesicles for storage in the cytoplasm or export to the cell exterior.
- About 15,000 ribosomes in a single E. coli cell, comprising 25 % of the dry cell mass.
- All ribosomes within one cell are identical.
 - All ribosomes are composed of two subunits (called small and large).
 - A substantial fraction of ribosomes are dissociated into free subunits in the cell.
 - Prokaryotic and eukaryotic ribosomes differ in composition.
 - Mitochondrial ribosomes resemble prokaryotic ribosomes.

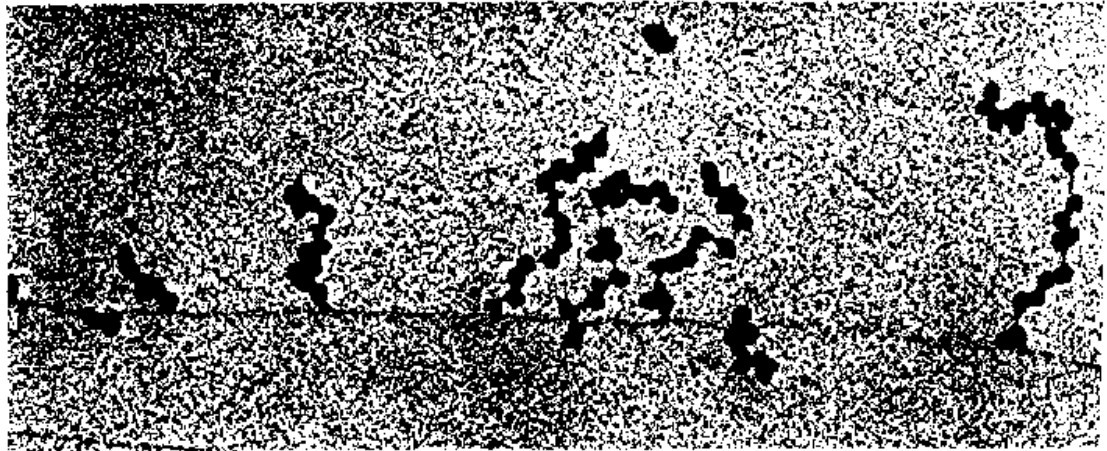
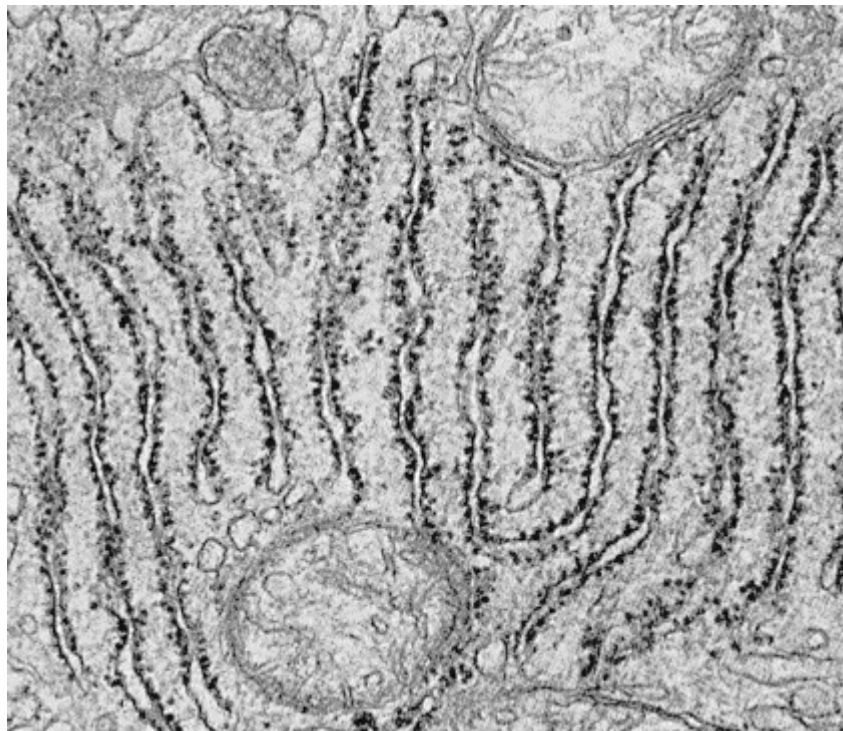


Figure 24-1

Electron micrograph of *E. coli* polysomes. Ribosomes are the dark structures connected by the faintly visible mRNA strand. A DNA strand connecting the polysomes from which the mRNA is being transcribed is visible as the diagonal line. (From O. L. Miller et al., Visualization of bacterial genes in action. *Science* 169: 392, 1970. Used with permission.)



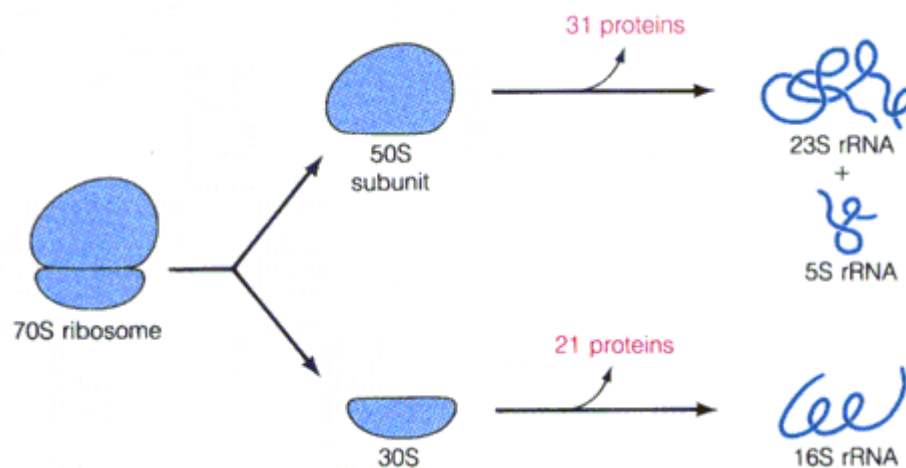
The composition of Ribosomes

Prokaryotic ribosomes

The sedimentation coefficient is measured in Svedberg units (S): the rate of sedimentation of a component in a centrifuge is related both to the molecular weight and the 3-D shape of the component.

RNA function and turnover in the cell

Type	Function	Different Kinds	Nucleotides	% of Synthesis	% of Total RNA	Stability
mRNA	Messenger	Thousands	500-6000	40-50	3	$T_{1/2} = 1-3$ min
rRNA	Ribosome	3 (23S, 16S, 5S)	2904, 1542, 120	50	90	Stable
tRNA	Adapter	~50	73-93	3	7	Stable

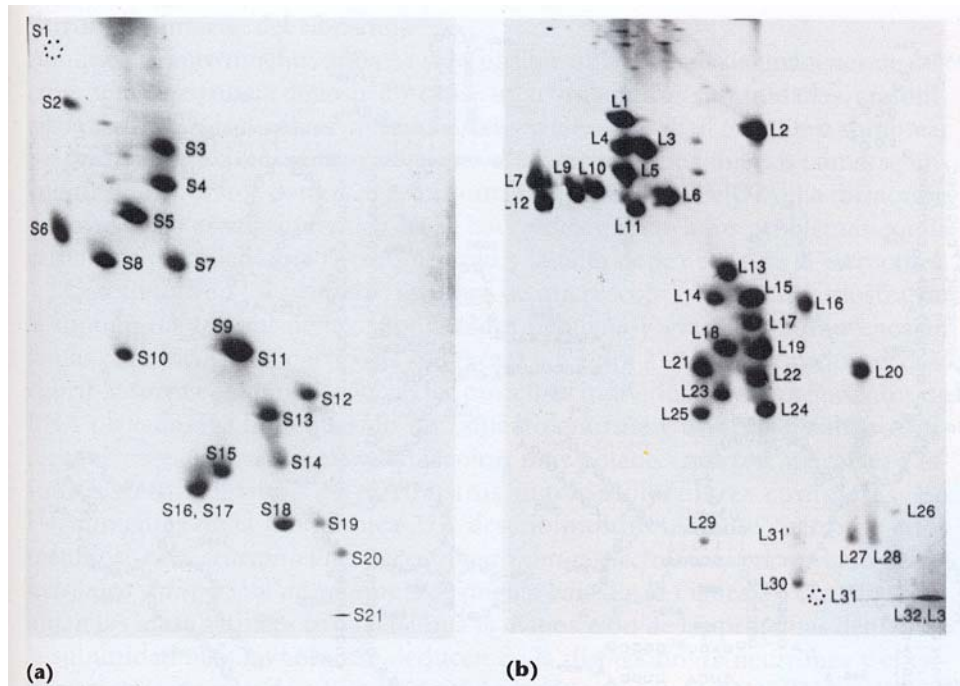


Componentes de los ribosomas de *E. coli*

	Ribosoma	Subunidad pequeña	Subunidad grande
Coefficiente de sedimentación	70S	30S	50S
Masa (kD)	2520	930	1590
RNA			
Mayor		16S, 1542 nucleótidos	23S, 2904 nucleótidos
Menor			5S, 120 nucleótidos
Masa del RNA (kD)	1664	560	1104
Proporción de la masa	66 %	60 %	70 %
Proteínas			
		21 polipéptidos	32 polipéptidos
Masa de las proteínas (kD)	857	370	487
Proporción de la masa	34 %	40 %	30 %

Fuente: Lewin, B., *Genes* (3.^a ed.), p. 145, Wiley (1987).

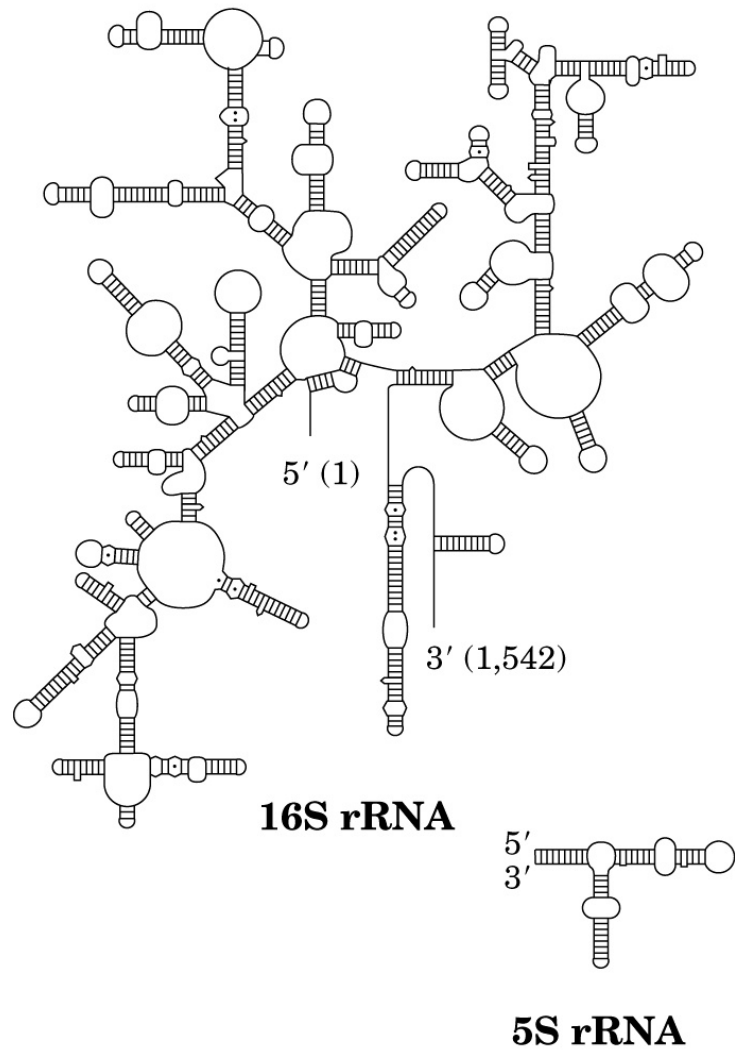
Separation of ribosomal protein by 2D gel electrophoresis (a) small (30S) subunit and (b) large (50S) subunit



Los RNAs Ribosómicos juegan un papel fundamental en la síntesis de proteínas

- La ruptura de un único enlace del rRNA de 16S por medio de la colicina E3, una nucleasa secretada por algunas bacterias, bloquea la síntesis de proteínas.
- La omisión de proteínas concretas durante la recostitución “in vitro” de las subunidades 30S provoca una disminución de la actividad del ribosoma en lugar de una pérdida total de su funcionalidad.
- Una secuencia del rRNA 16S selecciona el lugar de iniciación en el mRNA.
- La base oscilante del tRNA que ocupa el lugar P del ribosoma está apareada con una base localizada en una secuencia de 14 nucleótidos del rRNA 16S muy conservada. La presencia de esta secuencia idéntica en arqueobacterias, eubacterias y eucariotas indica que juega un papel crítico.
- El extremo aceptor 3' del tRNA interacciona con una región conservada del rRNA de 23S.
- Los ribosomas prácticamente desprovistos de proteínas aún pueden catalizar la formación de enlaces peptídicos. Por el contrario, el rRNA de 23S es esencial para la actividad peptidil-transferasa.
- La mayoría de los antibióticos que interfieren en la síntesis de proteínas lo hacen mediante interacciones específicas con el RNA ribosómico, no con las proteínas ribosómicas.

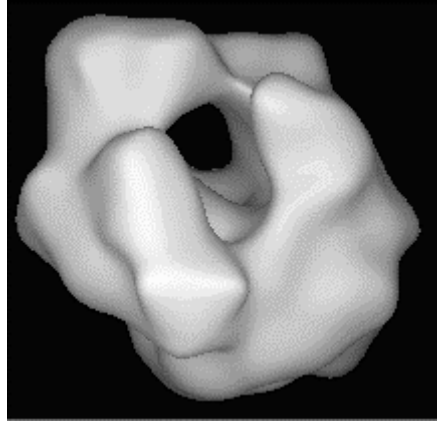
Proposed secondary structure of 16S rRNA



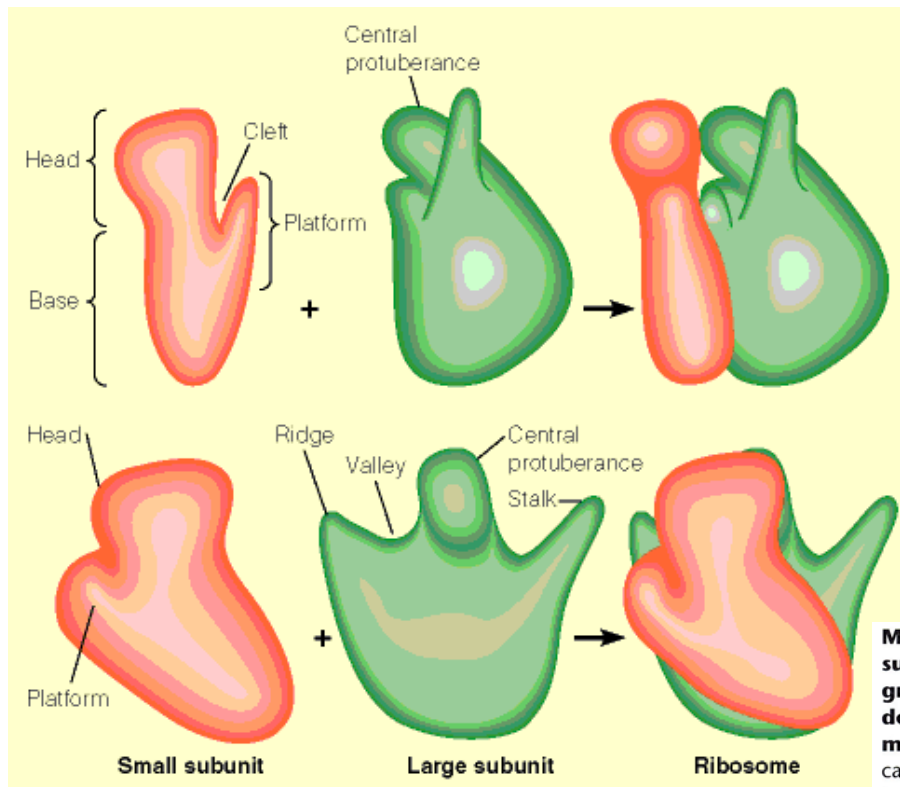
Structure of the Ribosome

3-dimensional structure of the 70S ribosome

the large subunit has a tunnel about 10 nm long and 2.5 nm in diameter. This tunnel is thought to be the channel that newly assembled polypeptide chains pass through on their way out of the ribosome.



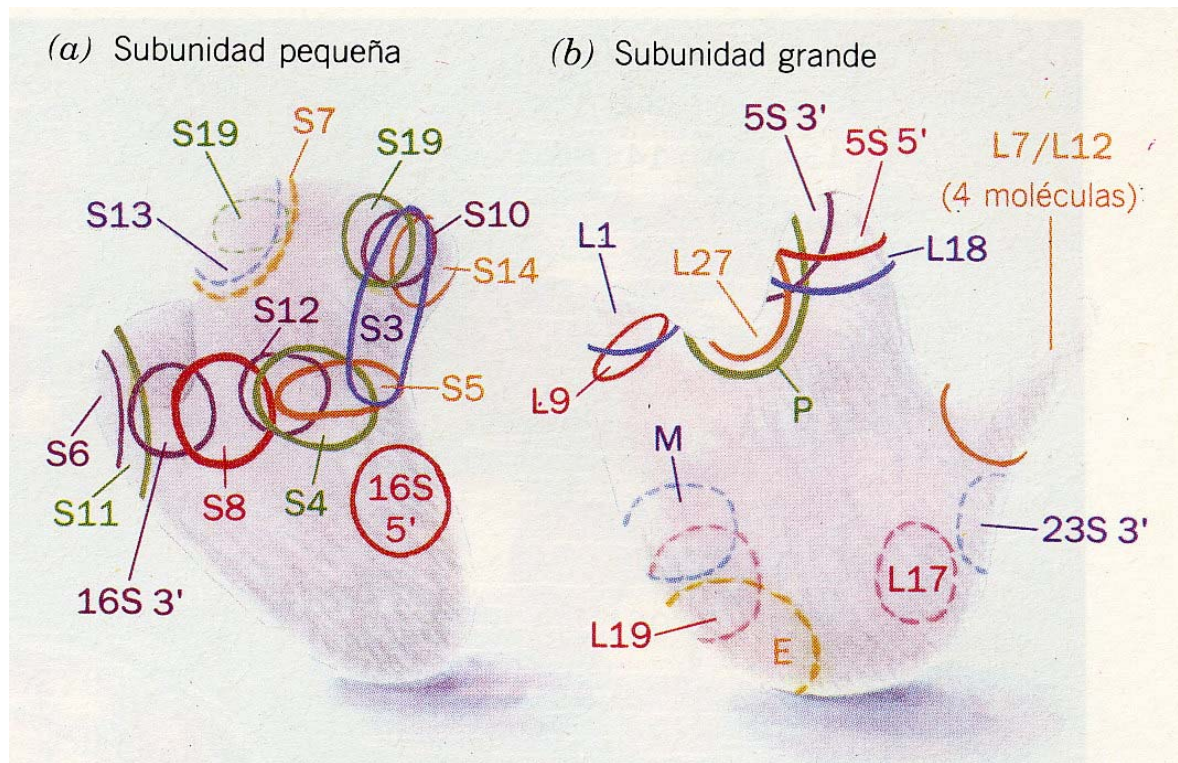
The two subunits interact very tightly and form the active site



Modelos de baja resolución de las subunidades ribosómicas pequeña y grande y su complejo 70S, según se deduce de los primeros estudios de microscopía electrónica. Las características más destacadas, que pueden también observarse en los estudios más recientes de mayor resolución, están indicadas por los nombres unidos.

Cortesía de M. Oaks, A. Scheirman, T. Atha, G. Shankweiler y J. A. Lake, *The Ribosome*, W. E. Hill, A. Dahlberg, R. A. Garret, P. B. Moore, D. Schlessinger y J. R. Warner, eds., p. 181 (Nueva York: American Society of Microbiology Press, 1990).

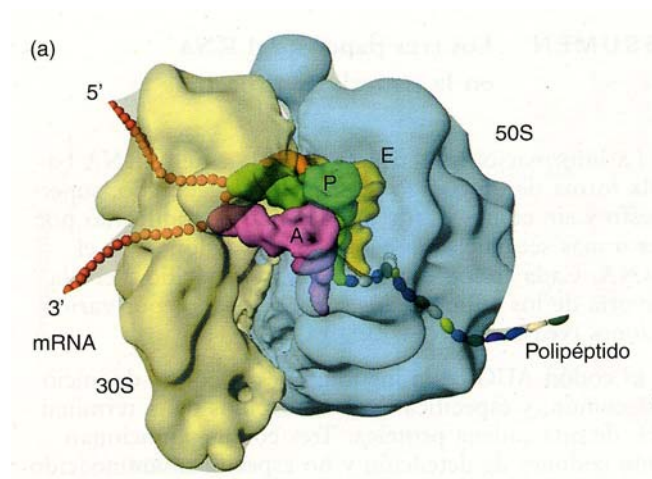
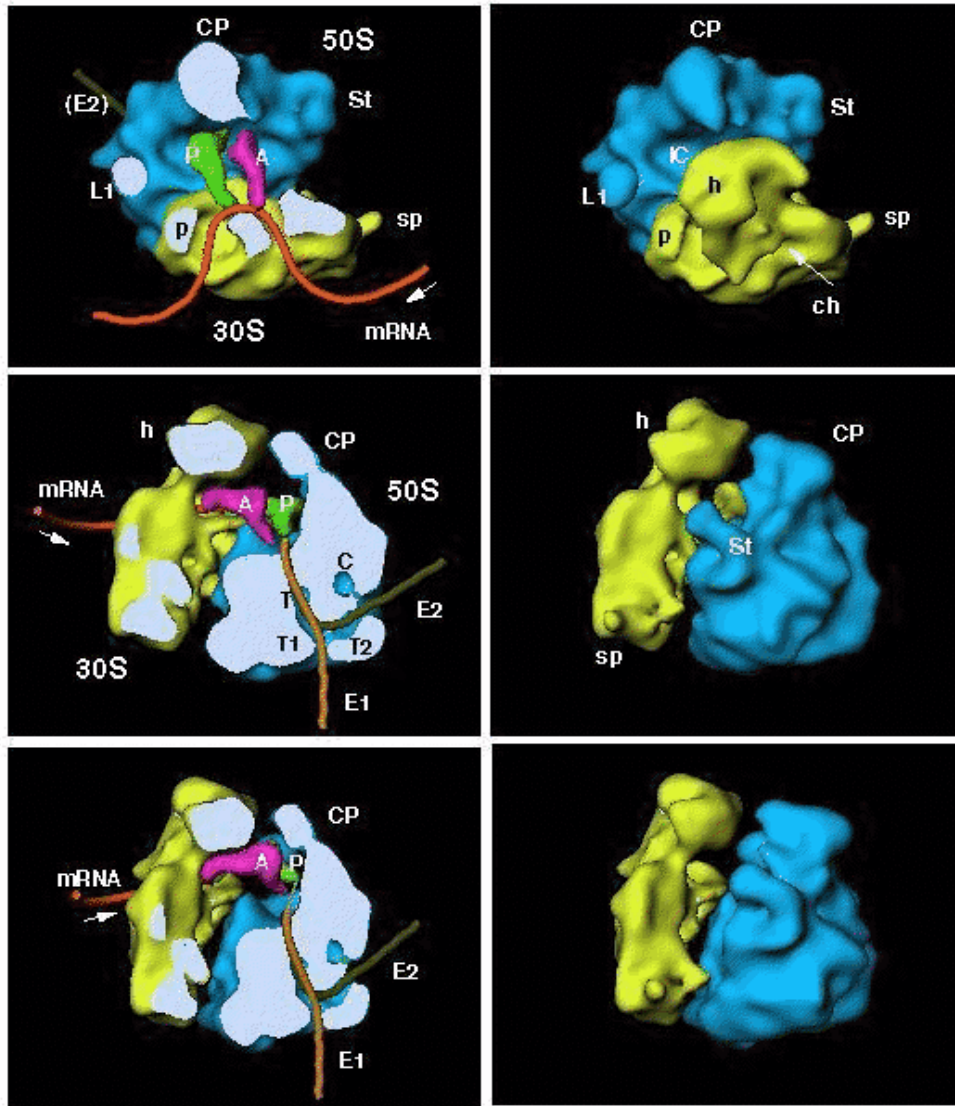
Las técnicas como el entrecruzamiento, la dispersión de neutrones, la inmunomicroscopía electrónica y la criomicroscopía electrónica han proporcionado un cuadro detallado de la estructura de algunas subunidades ribosómicas.

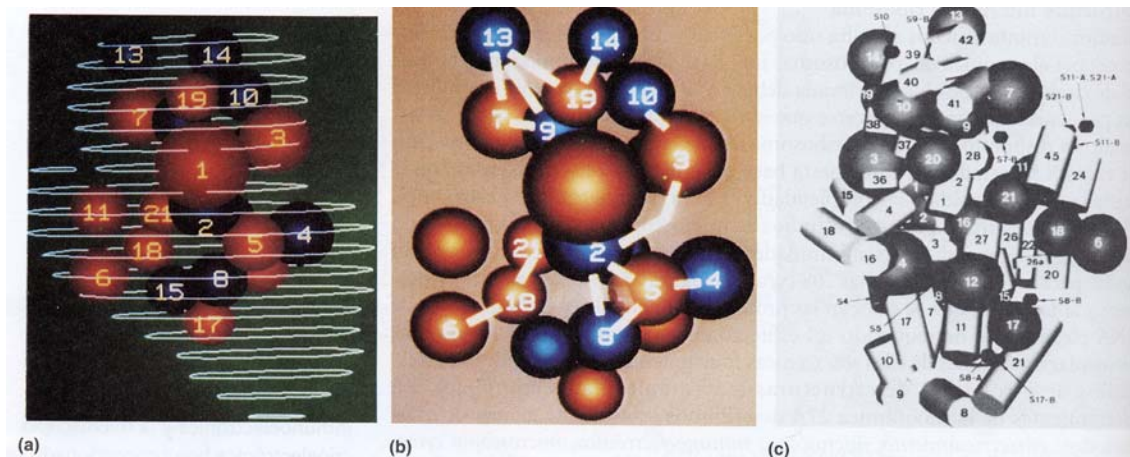


Mapa de las subunidades (a) pequeña y (b) grande del ribosoma de *E. coli*, indicando las localizaciones de algunas de las proteínas que las componen, determinadas por inmunomicroscopía electrónica. Los sitios que están señalados por líneas de puntos se encuentran al otro lado de la subunidad. Los símbolos 16S 3' y 16S 5' en la subunidad pequeña marcan los extremos del RNA de 16S. En la subunidad grande, la P indica el centro peptidil transferasa, la E marca el sitio donde el polipéptido naciente emerge del ribosoma (el final del túnel en la Fig. 30-24), la M especifica el sitio de anclaje a membrana del ribosoma, y 5S 3' y 5S 5' indican los extremos del RNA de 5S. [De Lake, J.A., *Annu. Rev. Biochem.* **54**, 512 (1985).]

Landmarks of the 30S subunit: h, head; p, platform; ch, channel presumed to be the conduit for mRNA; sp, spur.

Landmarks of the 50S subunit: CP, central protuberance; St, L7/L12 stalk; L1, L1 protein; IC, interface canyon; T, tunnel presumed to be the conduit for the nascent polypeptide chain; T1 and T2, lower tunnel segments, leading to alternative exit sites E1 and E2, respectively.





Disposición tridimensional de los componentes de la subunidad 30S.

(a) Disposición de las proteínas tal como se deduce de los experimentos de dispersión de neutrones. Las líneas de contorno muestran la forma global de la partícula. Las regiones que no ocupan las proteínas están rellenas de RNA. **(b)** Patrón de entrecruzamiento. Se reproduce la disposición de las proteínas de (a) con barras entre todos los pares de proteínas que están entrecruzados según han informado dos o más grupos de investigación independientes. Obsérvese que el patrón de entrecruzamiento concuerda bien con los datos de dispersión de neutrones. **(c)** Modelo completo, que muestra la situación del rRNA 16S (cilindros blancos) con relación a las proteínas, según se deduce de los estudios de entrecruzamiento RNA-proteína y RNA-RNA.

(a) Cortesía de P. Moore y V. Ramakrishnan de M. Capel et al., *Science* (1987) 238:1403-1406. © 1987 AAAS, reproducido con permiso; **(b)** reproducido con permiso de M. Capel et al., *Science* (1987) 238:1403-1406. © 1987 AAAS; **(c)** cortesía de E. Schüler y R. Brimacombe, *EMBO J.* (1988) 7:1509-1513. Con permiso de Oxford University Press.

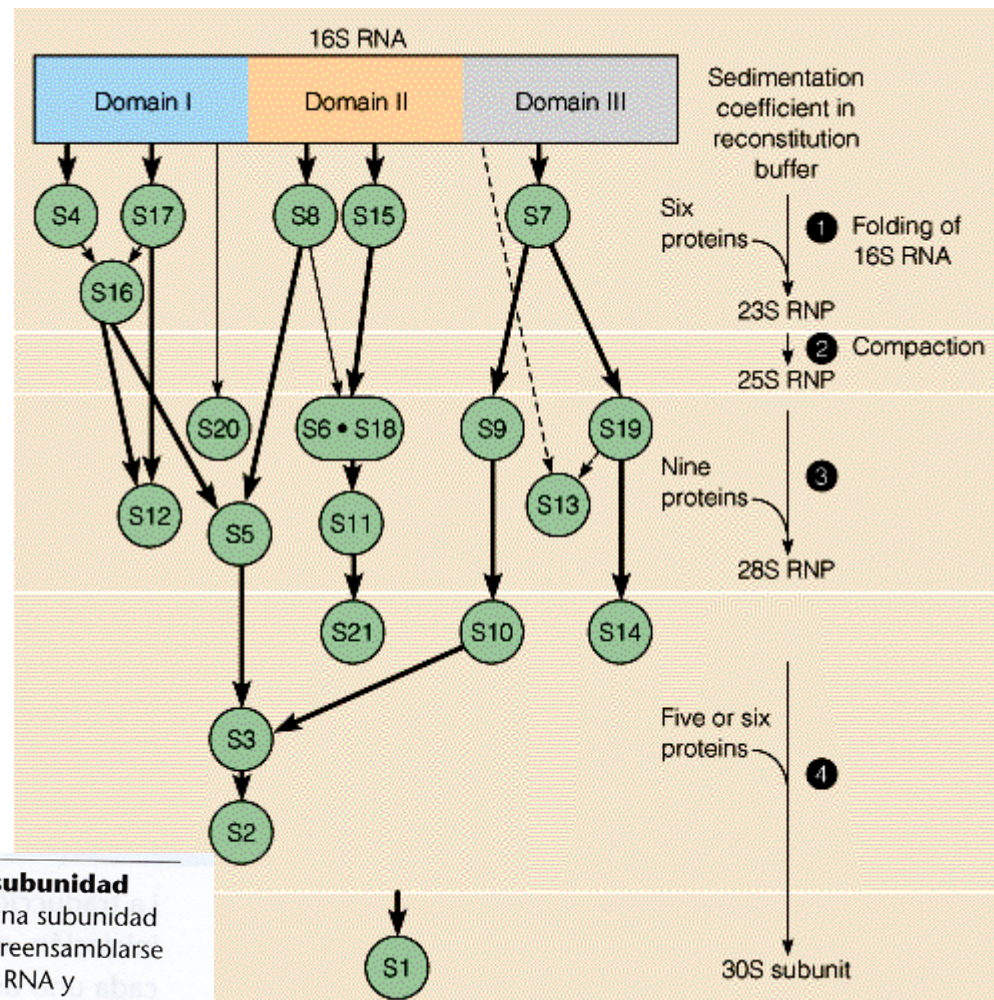
Estructura de la subunidad de 30S del ribosoma. (a) Diagrama de las posiciones relativas de las 21 proteínas de la subunidad de 30S del ribosoma, superpuesta con un perfil de su superficie. Si a la Parte (i) se la denomina visión frontal, las Partes (ii) y (iii) corresponden a vistas desde el lado izquierdo y desde abajo, respectivamente. Las proteínas reciben su numeración estándar de la Parte (i), en la cual S20 se encuentra directamente detrás de S3 (los colores diferentes de las esferas sólo pretenden facilitar la observación). Las distancias entre los pares de estas proteínas fueron determinadas por la difracción de neutrones de soluciones concentradas de subunidades de 30S reconstituidas, en las que las dos proteínas de interés estaban muy deuteradas, mientras que los otros componentes de la subunidad se encontraban protonados

normalmente (los deuterones desvían neutrones de forma bastante diferente a la de los protones). Estas medidas, aplicadas a muchos pares diferentes de proteínas, permitieron la elaboración de este mapa, en el cual el volumen de cada esfera es proporcional a la masa de la proteína correspondiente, y su posición marca el centro másico de la proteína. Compárese este mapa con la Fig. 30-31a. [Cortesía de Peter Moore, Yale University, y Malcolm Capel, Brookhaven National Laboratory]. (b) Localización de los segmentos bicatenarios del rRNA de 16S (cilindros) en relación a las proteínas de la subunidad de 30S (esferas) deducida a partir de estudios de entrecruzamiento RNA-proteínas. La perspectiva es la misma que en las Figs. 30-31a y 30-32(a,i). [De Schüler, D. y Brinacombe, R., *EMBO J.* 7, 1512 (1988).]

In 1968, Dr. M. Nomura, a professor here at UCI, was the first to show that the 30 S subunit could be reassembled from the individual components. He found that the order of addition of components was critical.



Masayasu Nomura

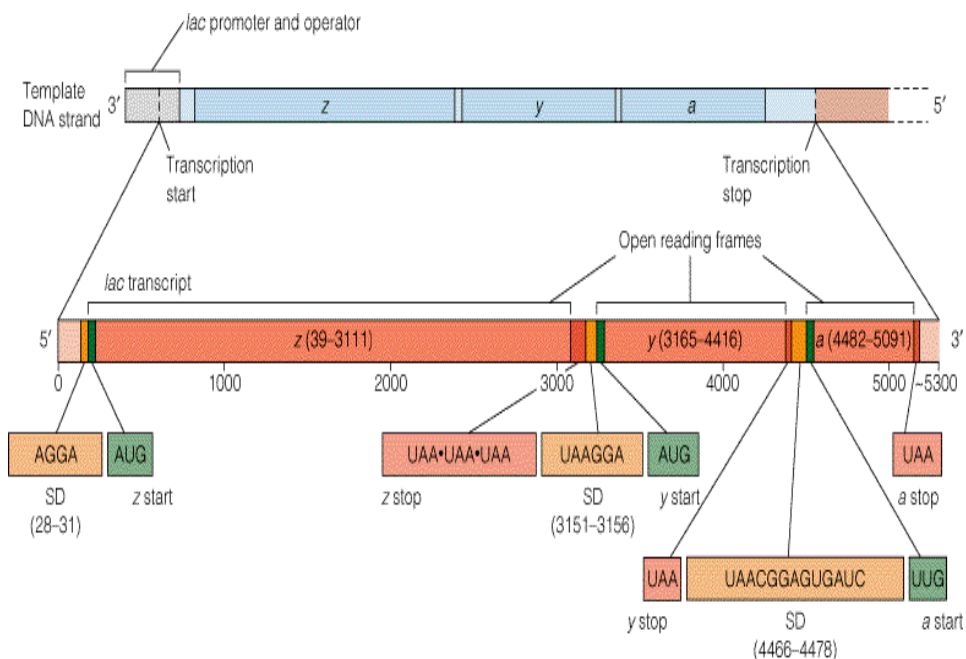
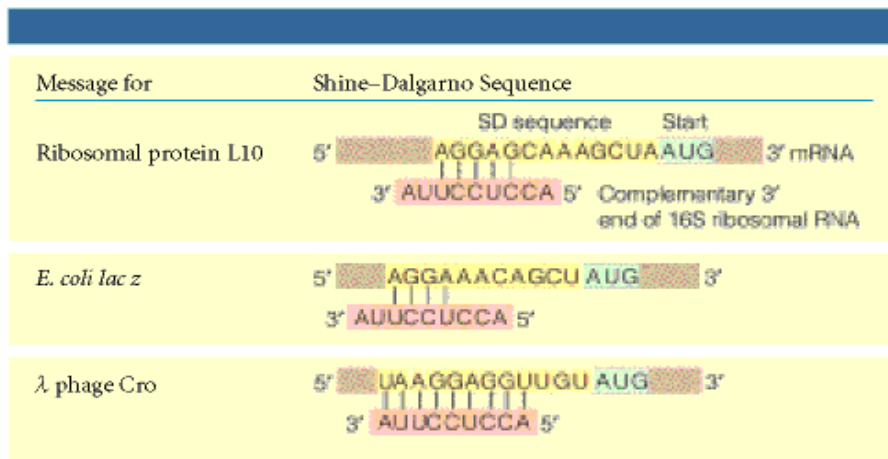
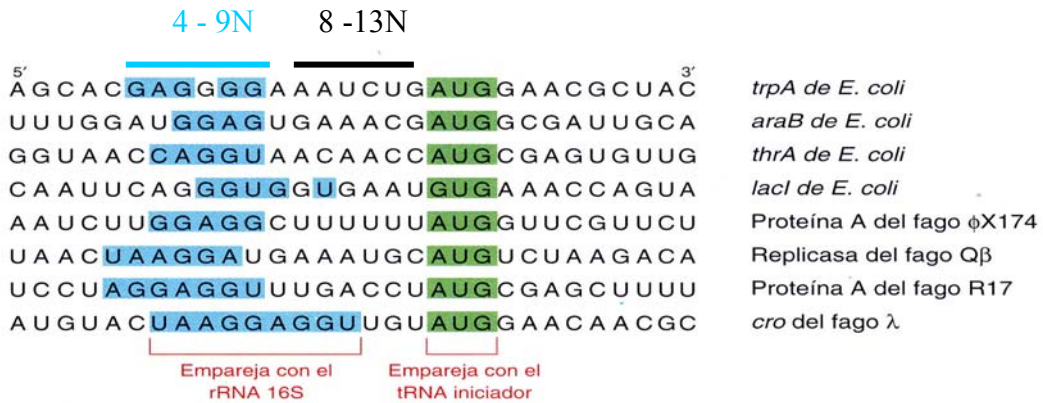


Mapa del ensamblaje de la subunidad

30S. El mapa muestra cómo una subunidad ribosómica 30S de *E. coli* puede reensamblarse en el tubo de ensayo a partir de RNA y proteínas (en verde). El orden de adición que se muestra en la figura es importante, puesto que algunas proteínas han de añadirse antes que otras para que el ajuste sea adecuado. Ciertas proteínas unidas inicialmente se asocian con cada uno de los dominios I-III del RNA que se indican en la Figura 27.15. Las flechas entre las proteínas indican la dependencia de la unión; cuanto más gruesa es la flecha, más esenciales son las proteínas previas. Así, por ejemplo, la unión de S3 en el paso 4 requiere que S5 y S10 estén ya unidas. La unión de S1 requiere una estructura ribosómica casi completa, pero resulta difícil establecer exactamente qué proteínas son esenciales. En cualquier caso, su unión completa el ensamblaje y se ha producido una subunidad 30S funcional. La ruta de la derecha muestra el desarrollo estructural de la subunidad. La abreviatura RNP indica partícula ribonucleoproteica.

Estructura de los mRNAs de los procariotas

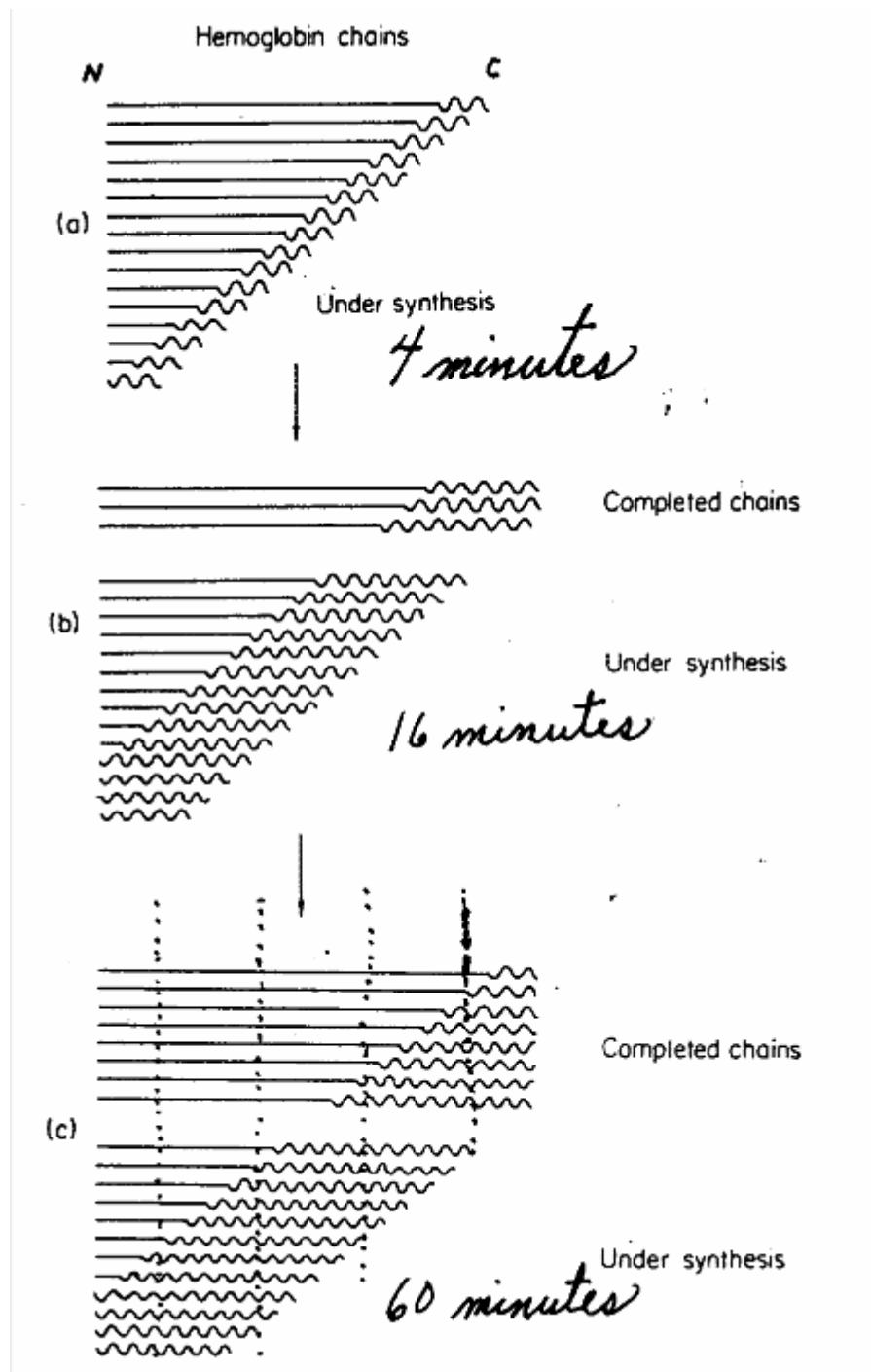
Las secuencias Shine-Dalgarno ayudan a alinearse a los ribosomas sobre los mRNA para comenzar la traducción de forma adecuada. Secuencias de la región de iniciación de la síntesis proteica en algunas moléculas de RNA bacterianas y víricas.



El mRNA del operón lac. El mRNA del operón lac de *E. coli* tiene unos 5300 nucleótidos de longitud y contiene los marcos de lectura abiertos para los genes lac z, y y a, flanqueados cada uno por secuencias de comienzo, de detención y de Shine-Dalgarno.

Polypeptide synthesis proceeds sequentially from N terminus to C terminus

Proved by the classic experiment of Dintzis in 1961. He synthesized hemoglobin in a test tube using a reticulocyte (red blood cell) system. He initiated protein synthesis and then added H³-labelled leucine. He isolated hemoglobin, partially labelled with incorporated H³-leucine, and created peptide fragments by digestion of hemoglobin with trypsin. Then he analysed peptide fragments to determine the relative amount of radioactivity. He obtained the following results:



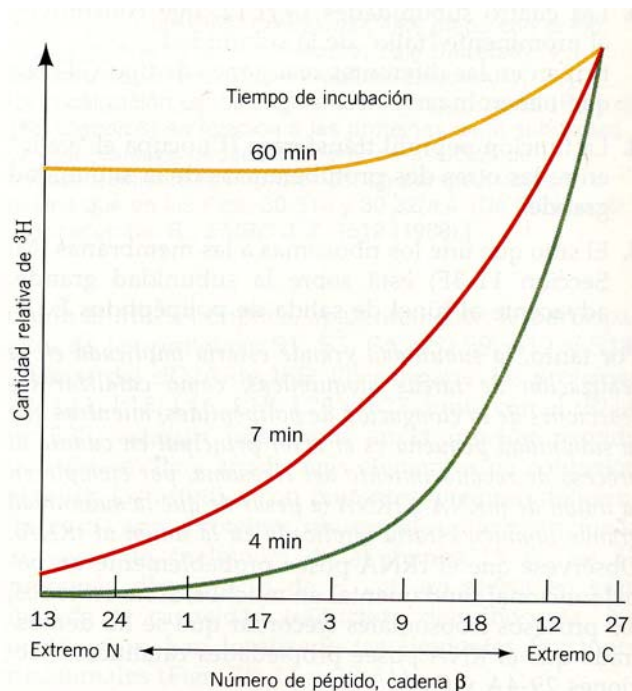


Figure 4.1.1: Incorporation of radioactive leucine into hemoglobin after different periods of incubation. There is a gradient from the C-terminus to the N-terminus of the protein, which implies that synthesis takes place sequentially from N-terminus to C-terminus. As the time of incubation is increased, the total amount of radioactivity is increased but the gradient remains. Results are expressed as the ratio of an H^3 label (tritium) in leucine to an internal control of C^{14} . Data of Dintzis (1961).

Ribosomes read mRNA in the 5' ---- 3' direction

To prove, one must know the Genetic Code. For example, **AAA** codes for Lys, **AAC** codes for Asn and **CAA** codes for Gln. In a cell-free protein synthesizing system, initiate protein synthesis by adding artificial mRNA, consisting of polyA with one C (cytosine) at the 3' end. Analyze the N and C termini of the peptide fragments.

Nucleotide Sequence	Protein Sequence	Direction of Synthesis
5'-AAAAAA(AAA)nAAC-3'	Lys-Lys-(Lys)n-Asn	5' -> 3'
3'-CAAAAA(AAA)nAAA-5'	Gln-Lys-(Lys)n-Lys	3' -> 5'