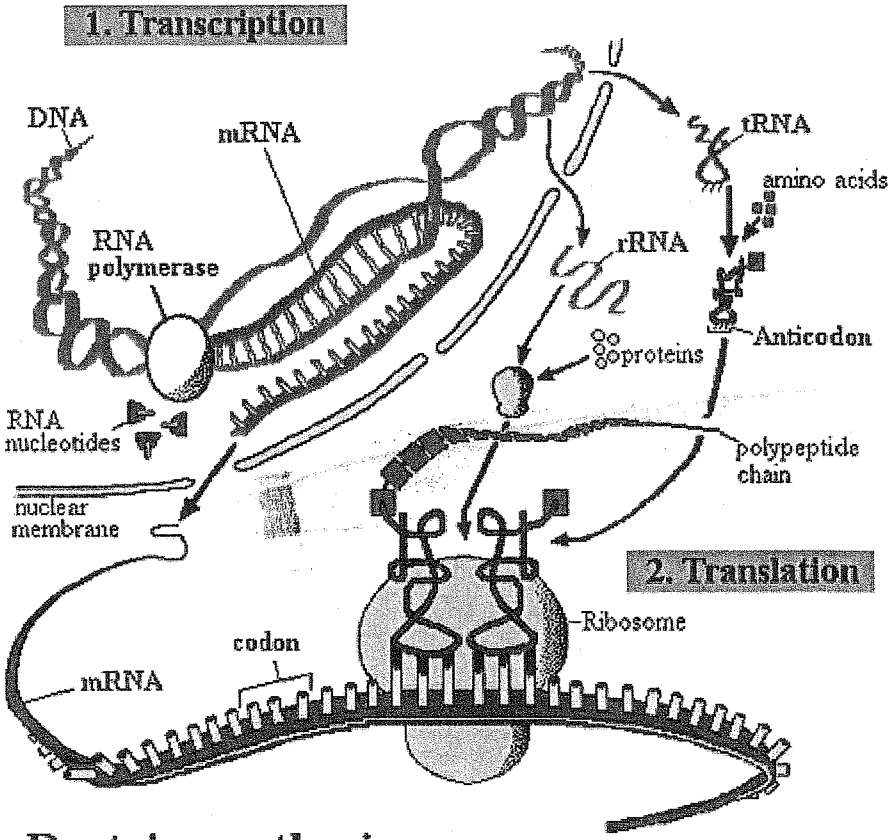
Protein Synthesis



Protein synthesis is the process by which DNA encodes for the production of amino acids and proteins. This process can be divided into 2 parts:

Transcription:

Before the synthesis of a protein begins, the corresponding RNA molecule is produced by RNA transcription. One strand of the DNA double helix is used as a template by the RNA polymerase to synthesize a messenger RNA (mRNA). This mRNA migrates from the nucleus to the cytoplasm. During this step, mRNA goes through different types of maturation including one called splicing when the non-coding sequences are eliminated. The coding mRNA sequence can be described as a unit of 3 nucleotides called a codon.

Translation:

The ribosome binds to the mRNA at the start codon (AUG) that is recognized only by the initiator tRNA. The ribosome proceeds to the elongation phase of protein synthesis. During this sate, complexes, composed of an amino acid linked to tRNA sequentially bind to the appropriate codon in MRNA by forming complementary base pairs with the tRNA anticodon. The ribosome moves from codon to codon along the mRNA. Amino acids are added one by one, translated into polypeptide sequences dictated by DNA and represented by mRNA. At the end a release factor binds to the stop codon, terminating translation and releasing the complete polypeptide from the ribosome.

One specific amino acid can correspond to more than one codon.

Chapter 10
RNA and Protein Synthesis

Ribonucleic acid (RNA) is very similar to DNA except for these differences:

- RNA is *single stranded* (instead of a double helix)
- RNA contains the sugar *ribose* instead of the sugar deoxyribose
- RNA *uses Uracil (U) as a nitrogenous base* instead of thymine (T) so nucleotides in RNA pair up as A-U
- RNA bases can pair up, even though the RNA molecule is single stranded, because *RNA has a secondary structure and can fold up and base pair with itself where complementary.* These double-stranded portions of RNA are *called* **hairpins**.
- RNA *comes in a variety of types*

RNA molecules are important for the production of proteins. And, just after DNA is replicated, the complementary strands head out to produce proteins.

You know that the DNA harbors the genes that code for what proteins will be produced in your body. These are called ***structural genes***. But the code buried in segments of DNA is no what initiates protein production. First, the DNA must be *"rewritten"* (called **transcription**) into a strand of RNA, and the RNA carries – (therefore called messenger RNA) – the information out of the cell's nucleus to the ribosomes. At a ribosome, the original message is translated, and then the appropriate protein can be produced. Protein synthesis is initiated on ribosomes that exist free in the cytoplasm. The ribosomes that are attached to the endoplasmic reticulum (ER) make proteins to be secreted or transported to other organelles.

Types of RNA:

1. Pre-RNA: Precursor RNA, which is any newly transcribed RNA that has to be processed before it can be functional. Includes pre-mRNA (precursor to mRNA) and precursors of other types of RNA.
2. snRNA: Small nuclear RNA that take part in converting pre-RNA into mature RNA (In eukaryotes snRNA is found **splicesosomes**, which are granules that look like small ribosomes that never leave the nucleus).
3. mRNA: This is messenger RNA. It is a molecule that carries the genetic code for polypetides. mRNA attaches to a ribosome, where the code is then translated into a sequence of amino acids in a polypeptide.
4. rRNA: This is ribosomal RNA. rRNA molecules make up more than ½ of ribosome and this is where amino acids are joined to form polypeptides.
5. tRNA: This is transfer RNA. Each tRNA is specific to one kind of amino acid. tRNA carries the amino acid to mRNA at a ribosome and fits into the right place on the growing polypeptide.

Transcription: Rewriting the DNA's Message

Remember that the transcription of DNA into RNA begins with **RNA polymerase** (an enzyme) that *binds to a sequence of nucleotides on the DNA*. When an original strand of *DNA is unzipped during DNA replication*, its nucleotide bases are used as templates for the production of new complementary strands of DNA. Those *new complementary strands are used during transcription as the templates* (called **copy DNA -cDNA)** from which a strand of RNA is produced. The type of RNA molecule produced from the transcribed message is called messenger RNA (*mRNA*) because it then *carries the DNA's message outside of the nucleus for processing*. (When RNA polymerase reaches the termination signal on the DNA, it leaves the DNA, and the RNA detaches.)

The *complementary strand of DNA contains certain sequences (such as T-A-T-A- and C-C-A-A-T, called TATA or CAT boxes*) that exist on the strand just before the spot where transcription should start. Transcription begins at a region on the mRNA that is untranslated. This region, aptly called the 5' – untranslated region (UTR), serves as a signal to the ribosomes to begin translation at the appropriate ***codon* (a group of 3 bases)** that reads A-U-G. The nucleotides that follow those indicators are "read" and transcribed into the corresponding bases that are appropriate for an RNA molecule. (*Mutations* that *force the code to be read from the wrong place are called **frameshift mutations***.) So, A is the template for the addition of U. T is the template for the addition of an A, C is the template for the addition of a G, and G is the template for the addition of a C. An original DNA template message of 5'-T-A-C-G-A--G-C-A-3' becomes a complementary *DNA message* of 3'-**T-A-C-G-A-C-G-T**-5'. Then, that message serves as a template to create the following *mRNA sequence*: **A-U-G-C-U-G-C-A.**

As transcription occurs down the template, certain areas are read and transcribed. A large *precursor* is made called heterogeneous nuclear RNA (*hnRNA*), which has both intervening sequences (called *introns*) and expressed sequences (called *exons*). The introns are removed, and the exons are joined by a process known as splicing. (Splicing takes place only in eukaryotes, not prokaryotes). When all the nucleotides have been read, exons have been spliced together, and the messages have been transcribed, the end of the material to be transcribed is indicated by a termination codon. The **termination codon** – either T-A-A, T-A-G, or T-G-A – is *like a stop sign* along the strand. Transcription of a message completely ends at the 3'UTR (untranslated region), which can be quite a long sequence.

Transcription of mRNA is similar to the replication of DNA except that uracil (U) pairs with Adenine of the cDNA rather than thymine.

Processing the RNA

First step in processing RNA: Once the DNA message is transcribed, and the pre-mRNA is produced, the pre-mRNA gets a cap and a tail. The cap and tail help to stabilize the pre-mRNA molecule as it transports the important messages. (Cap to the 5'end and a poly-A tail to the 3" end of a pre-mRNA.) The nucleotides that did not code

for anything – introns – are deleted out, and the important nucleotides – the exons – are joined together (Refer to Figure 1). Think of when you use a word processing program, and you highlight a sentence in the middle of a paragraph and select the "Cut" command. The sentences before and after the one you deleted then join up to make the paragraph flow smoothly. Once the cutting and splicing is done, the mRNA is mature enough to leave the nucleus and head out into the cytoplasm on its own with the important genetic information (Refer to Figure 2). You wouldn't trust an unlicensed 15 year old to drive a moving van filled with you most treasured belongings, now would you? Maturity is important even at the cellular level.

Second step in processing RNA: splicing of pieces out of middle of the pre-RNA. Eukaryotic genes contain strings of nucleotides that do not code for amino acids that are interspersed with coding sections. (Noncoding called *introns*, separate coding sections called *exons*.)

In eukaryotes **spliceosomes** (granules that look like small ribosomes, but never leave the nucleus) splice pre-RNA in the nucleus helping it to convert to mature RNA.

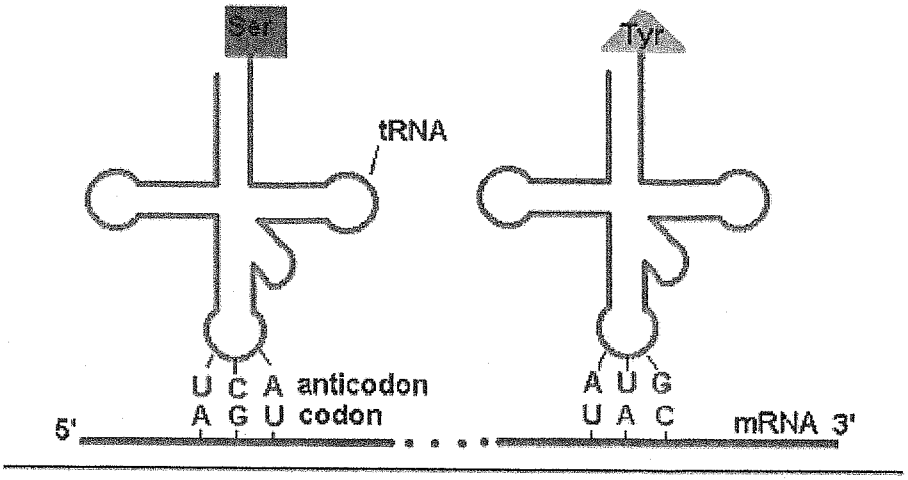
Translation: Putting the Code into the Right Language

Once a *mature mRNA* leaves the nucleus of a cell, it *heads for a ribosome in the cell's cytoplasm (*some head for a ribosome attached to the endoplasmic reticulum). As the strand of mRNA, which is carrying the ever-so-precious genetic information, slides through the two parts of a ribosome, the code it is carrying is read 3 bases at a time (remember this is called a *codon*).

Genes reside on a strand of DNA, and genes direct the production of amino acids, which are then put together to form a protein. **Translation** is the part of the process where the information from the gene encoded in RNA molecules is used to create the amino acids and then the protein (refer to Figure 2). To perform this amazing feat, however, the ribosome must be able to take the nucleotide bases and equate it with codons that specify amino acids. Each amino acid is represented by certain codons. The "language" that bridges the gap between gene and amino acid is the *genetic code*. (**Remember:** *in the genetic code, certain codons act as initiating codons and terminating codons* and the code *"degenerates"* meaning that most amino acids are specified by more than one codon.)

The genetic code was deciphered within the last 25 years or so and earned the men (*Nirenberg and Khorana*) who cracked it a pair of Noble Prizes. The genetic code uses the four nucleotide bases – thymine, cytosine, adenine, and guanine – to "spell" out what amino acid should be produced. But, codons usually are written as the RNA sequence, so uracil (U) is used in place of thymine. And because each codon is "spelled" by 3 bases, the genetic code contains 64 different codons (4^3) (refer to Figure 3). Amazingly, every organism uses the same genetic code. Each organism produces a different number of amino acids, however. Humans use only 20 amino acids, so several codons specify one amino acid.

Genetic Code



2nd base in codon

1st base in codon	U	C	A	G	3rd base in codon
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	STOP	STOP	A
	Leu	Ser	STOP	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

Amino acids specified by each codon sequence on mRNA. Key for the above table:

Ala: Alanine	Cys: Cysteine	Asp: Aspartic acid	Glu: Glutamic acid
Phe: Phenylalanine	Gly: Glycine	His: Histidine	Ile: Isoleucine
Lys: Lysine	Leu: Leucine	Met: Methionine	Asn: Asparagine
Pro: Proline	Gln: Glutamine	Arg: Arginine	Ser: Serine
Thr: Threonine	Val: Valine	Trp: Tryptophane	Tyr: Tyrosisne

DNA transfers information to MRNA in the form of a code defined by a sequence of nucleotides bases. During protein synthesis, ribosomes move along the mRNA molecule and "read" its sequence 3 nucleotides at a time (codon) from 5' end to the 3" end. Each amino acid is specified by the mRNA's codon, and then pairs with a sequence of 3 complimentary nucleotides carried by a particular tRNA (anticodon).

Since RNA is constructed from 4 types of nucleotides, there are 64 possible triplet sequences or codons (4 x 4 x 4). Three of these possible codons specify the termination of the polypeptide chant. They are called "stop codons". That leaves 61 codons to specify only 20 different amino acids. Therefore, most of the amino acids are represented by more than one codon.

Crick and others tested the triplet (3) code theory. It was hypothesized 4 nucleotides (arranged in all possible arrangements of two) would have to be arranged into triplets to code for 20 amino acids. They found that if the genetic code were composed of doublet codons (4^2) instead of the triplet codons a maximum of 16 codes could be formed, not enough to specify 20 different amino acids.

The mRNA strand moves through the ribosome, bringing codon after codon into position so that it can be read. As the mRNA codon is read (AGU), a transfer RNA (tRNA) molecule brings the appropriate anticodon (UCA), amino acid to the mRNA-ribosome. (Remember: an anitcodon is a triplet of nucleotides that can be base-paired with the appropriate mRNA codon) *Amino acids are then joined together by peptide bonds to form proteins*. Following translation, protein molecules may need to be modified a bit before they become functional. But, once the protein is created, it is soon capable of being expressed within the organism.

Suppose that you are getting too much sun (or unnatural ultraviolet rays). To protect yourself, the cells on the tip of your nose want to darken a bit. If the original template strand of DNA contained the gene for depositing melanin (skin pigment) on the tip of your nose, the complementary strand of DNA created during DNA replication also contained that information. Then, when the bases were transcribed for mRNA during the process of transcription, the mRNA was handed that genetic information. The mRNA carried the info to the ribosome, and the codons on the mRNA were read during translation.

During protein synthesis, the tRNA brought in the amino acids for melanin (in the appropriate order, of course), and peptide bonds joined the amino acids together to form the protein melanin. The genes directed that the melanin be deposited on the tip f your nose, so lo and behold, you develop a freckle in a very obvious spot. The protein was created and is now expressed, in an attempt to protect you from damaging UV rays, whether you like it or not.

Ribosomal RNA and Ribosomes

Remember: DNA has 2 strands, each of which has sequences of nitrogenous bases that form the genetic code. The genetic code, which is derived from the nucleotide bases in the genes on strands of DNA, is "interpreted" and then a ribonucleic acid (RNA) molecule called messenger RNA (mRNA) is produced from the DNA template. The mRNA uses the information from the genetic code to create amino acids – building blocks of protein – in the cell. The amino acids are then taken by transfer RNA (tRNA) to an organelle called a ribosome, where the final proteins are made.

Each nucleus has a round mass inside it called a nucleolus (in eukaryotes). The nucleolus produces the third type of RNA molecules – that is, *ribosomal RNA (rRNA)*. This type of RNA *helps to make ribosomes, which get transferred from the nucleus to the cytoplasm*

to help in making proteins. A *functional ribosome* will *contain rRNA and protein and hold mRNA and tRNA in place for protein synthesis.*

Protein Synthesis

Remember: The "events" of protein synthesis (in other words the order of their occurrence) is as follows:

- Transcription
- mRNA
- ribosomes
- tRNA
- peptide bond formation

Protein synthesis occurs on ribosomes and involves:

- Initiation
- Elongation
- Termination

Initiation: *mRNA binds with a small ribosomal subunit (in cytoplasm)* and the 1st codon (AUG) is in position to initiate the process. It will base-pair with the anticodon of a tRNA (that has methionine, which will become the first amino acid in the polypeptide). (In other words, *the mRNA codon at the A site bonds a tRNA with the complementary anticodon* to make the first amino acid in the polypeptide chain.)

Ribosome has 2 sites where tRNA can bind. First, the **P** (peptidyl) site, where the initiation codon, AUG (located on the mRNA) is positioned. Second, the **A** (aminoacyl) **site**, where the second amino acid is located.

Elongation: three steps are involved here. First, the mRNA codon at the A site will bind a tRNA with a complementary anticodon. The amino acid that is carried by the tRNA will become the next amino acid in the polypeptide chain.

Second, an amino acid at the *P site is attached to one at the A site*. Peptide bond formation leaves the tRNA at the P site empty and the A site holds the growing peptide. The tRNA this is empty leaves the ribosome and will pick up another molecule of its amino acid.

Third, a molecule of the nucleotide GTP provides the energy needed to *mRNA (with tRNA attached to it) so it can move from the A site to the P site.* This translocation brings the next mRNA codon to the ribosme A site. Then the elongation cycle repeats with the anticodon of the appropriate tRNA binding to the next codon, which brings the next amino acid into position at the A site.

Termination: protein synthesis stops when the ribosome reaches a *STOP* codon on mRNA. A releasing factor (a special protein) binds to the *Stop* codon and pushes the mRNA off the ribosome, and the ribosomal unites separate.

Post-translation: Most *polypeptides* must be *trimmed, finished and folded* after the ribosome releases them. They can also *join with other polypeptides, forming a complete protein.*

Remember: Each mRNA molecule usually has many ribosomes attached to it and its transcribing its message as they move along. One mRNA with many ribosomes attached to it forms a cluster called a **polysome**.

According to the current theories of protein synthesis *the genetic information for protein synthesis is carried from DNA to the ribosomes by RNA.*

Archibald Garrod was the 1st scientist to suggest that genes control the synthesis of enzymes.

Crick described the *central dogma* of molecular biology: The genetic code for a protein is transcribed into RNA, which is then translated into the sequence of amino acids in the protein.

DNA $\xrightarrow{\text{Transcription}}$ RNA $\xrightarrow{\text{Translation}}$ Protein

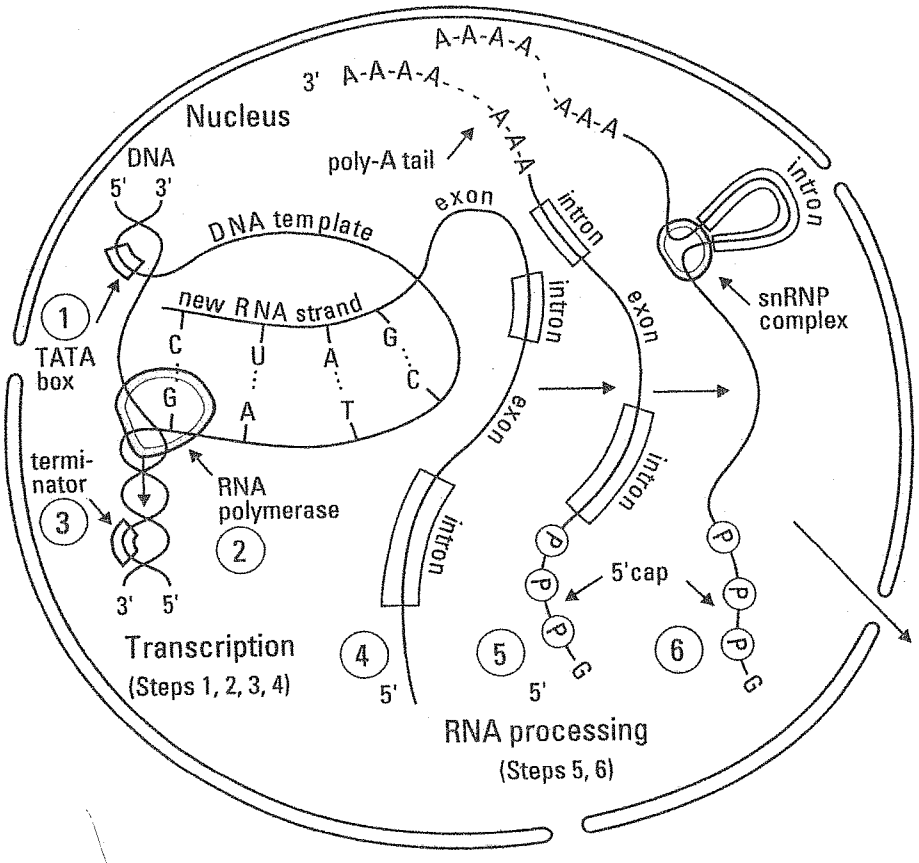


Fig. 1

Transcribing the DNA message and processing the mRNA: the process of transcription occurs in the nucleus of a cell

Protein Synthesis

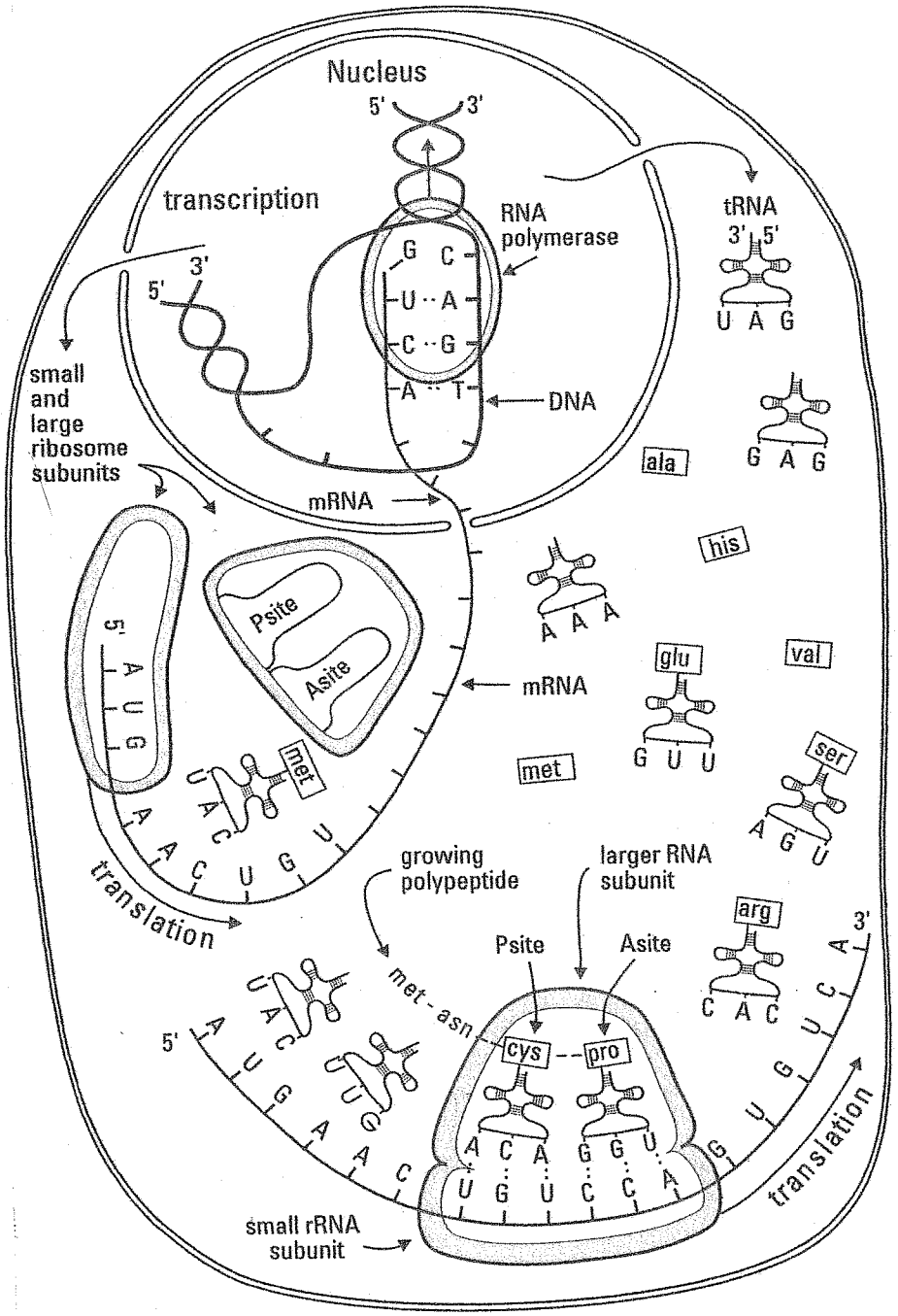


Fig. 2

The mRNA moves out of the nucleus and into the cytoplasm after transcription and RNA processing. Then, translation can take place.

Genetic Code

First Letter ↓	Second Letter U	Second Letter C	Second Letter A	Second Letter G	Third Letter ↓
U	phenylalanine	serine	tyrosine	cysteine	U
	phenylalanine	serine	tyrosine	cysteine	C
	leucine	serine	STOP	STOP	A
	leucine	serine	STOP	tryptophan	G
C	leucine	proline	histidine	arginine	U
	leucine	proline	histidine	arginine	C
	leucine	proline	glutamine	arginine	A
	leucine	proline	glutamine	arginine	G
A	isoleucine	threonine	asparagine	serine	U
	isoleucine	threonine	asparagine	serine	C
	isoleucine	threonine	lysine	arginine	A
	methionine & START	threonine	lysine	arginine	G
G	valine	alanine	aspartate	glycine	U
	valine	alanine	aspartate	glycine	C
	valine	alanine	glutamate	glycine	A
	valine	alanine	glutamate	glycine	G

Fig. 3

One letter from the vertical column on the left, one letter from the horizontal column, and one letter from the vertical column on the right are the codon

The amino acids that the codon codes for are given.

Note: the same amino acids are coded for by more than 1 codon

EX: UUU codes for phenylalenine, so does UUC.

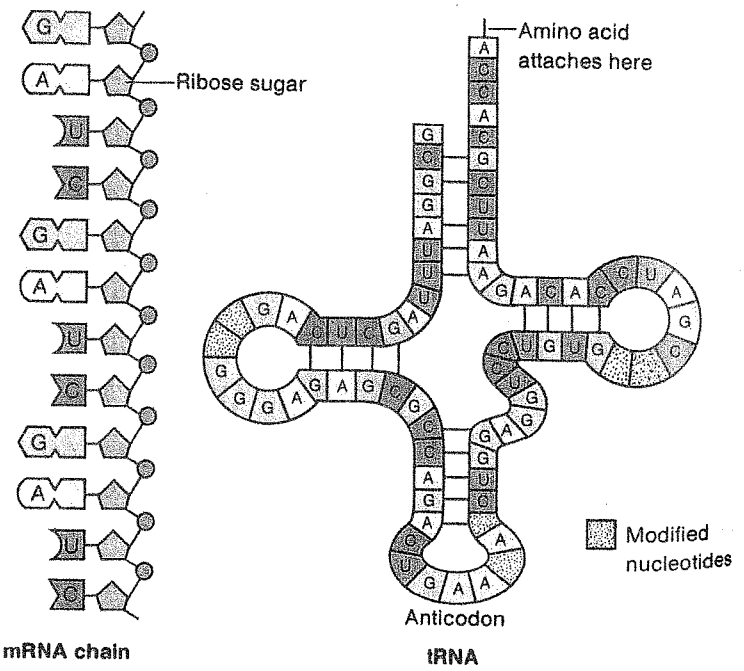
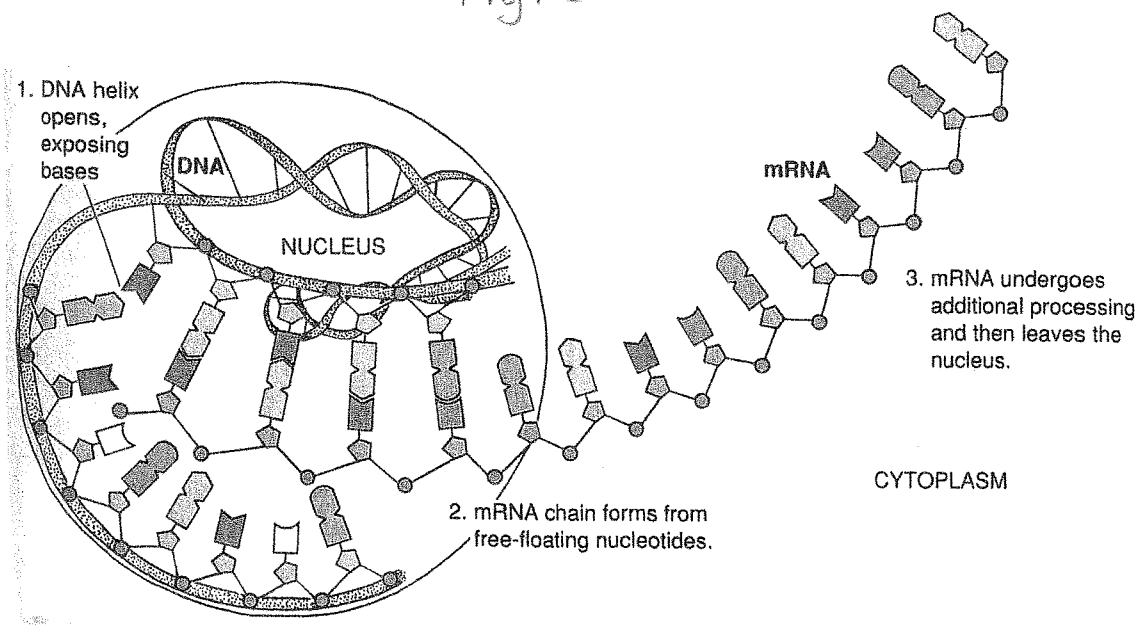


Fig. 4

Molecules of mRNA are made of single strands of uncoiled nucleotides. (left)

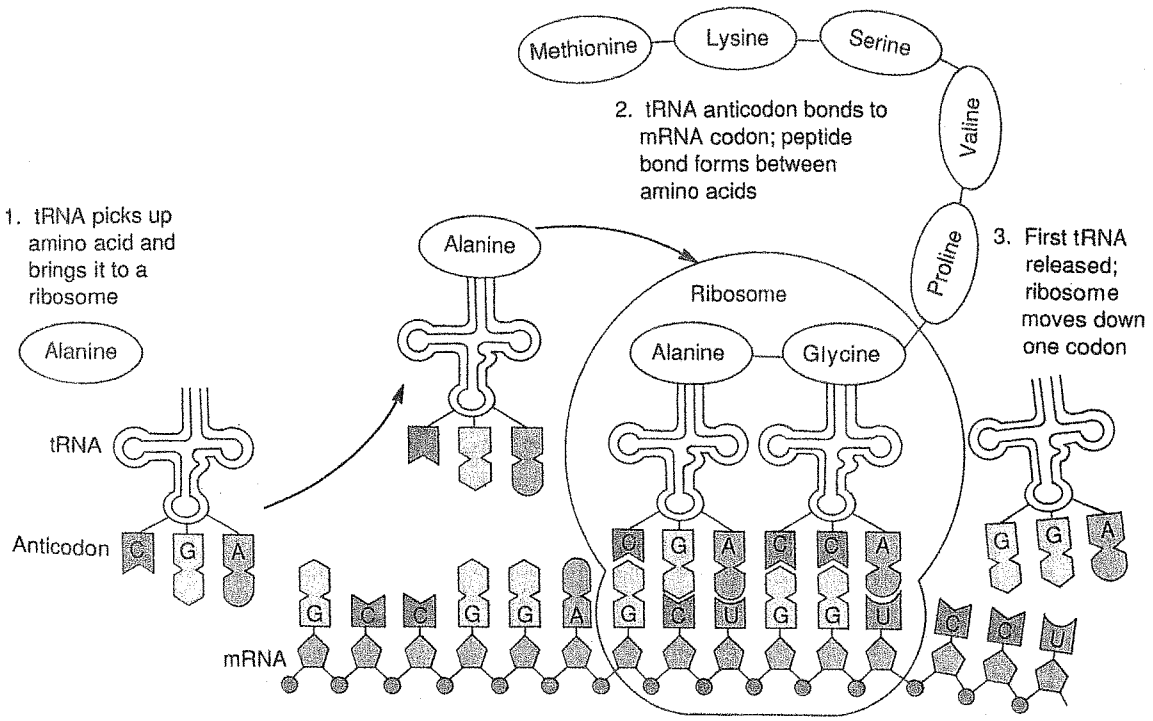
Cloverleaf shape of tRNA results from base pairing @ various pts. along the single strand.

Fig. 5



Transcription occurs in the nucleus. A completed molecule of mRNA enters the cytoplasm through a nuclear pore. It travels through the cytoplasm until it reaches a ribosome.

Fig. 6



Translation (building protein from mRNA)

1. A molecule of tRNA w/ a methionine amino acid (a.a.) attached moves toward a ribosome. There, the anticodon of the tRNA bonds to the initiation codon of the mRNA on the ribosome.
2. Another tRNA-amino acid complex bonds to the mRNA next to the first tRNA. 2 a.a. are now positioned side by side. Enzymes catalyze the formation of a peptide bond between them.
3. The first tRNA is released to pick up another methionine molecule. A chain of 2 a.a. is now attached to the second tRNA. The ribosome moves down one codon on the mRNA molecule.
4. tRNA molecules bring more a.a. to the ribosome and bonding process is repeated.
5. Protein synthesis stops when the ribosome reaches a termination codon on the mRNA molecule. The protein chain is then released from ribosome for use.

Once protein is completed, mRNA falls off the ribosome, mRNA breaks down into individual nucleotides that return to nucleus - unless cell needs another molecu of that protein the same mRNA can be read by another ribosome.

CONCEPT MAPPING