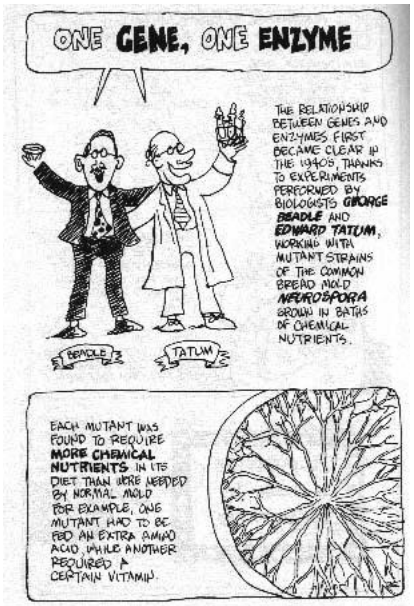
The Central Dogma: DNA--->RNA-->Protein

Chem 210

2/4/00

What do genes do?



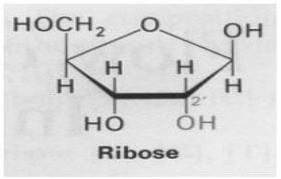
Genes must be related to the synthesis of proteins

- To use coding metaphor of Schrödinger, genes must "encode" information for the synthesis of proteins.
- Hypothesis: There must be a linear correspondence between structure of the gene and structure of the protein:
  - Sequence of the bases in a stretch DNA determine the sequence of amino acids in a protein
- Big questions:
  - How do you get from DNA to protein?
  - What is the "genetic code" which relates DNA sequence to protein sequence?

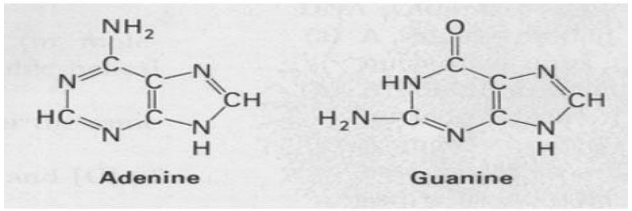
Protein synthesis also involves RNA

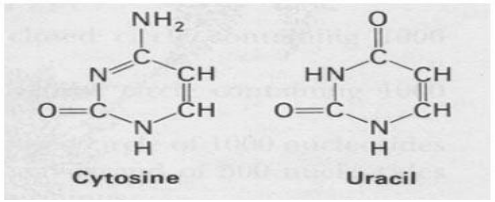
- Protein synthesis occurs at *ribosomes*, a complex found in all cells which is rich in RNA.
- Jacob and Monod in 1961 demonstrated that some RNA is synthesized right before protein synthesis, and this RNA is short lived.
  - Proposed that this RNA serves as an intermediate between DNA and protein synthesis
  - Called this "messenger RNA" (mRNA)

Key differences between RNA and DNA

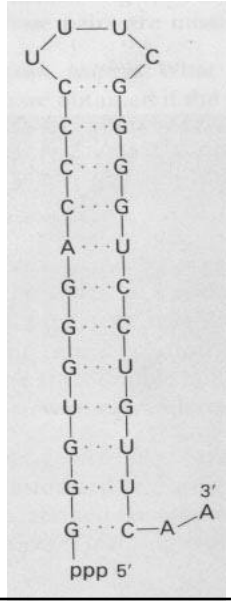


Ribose

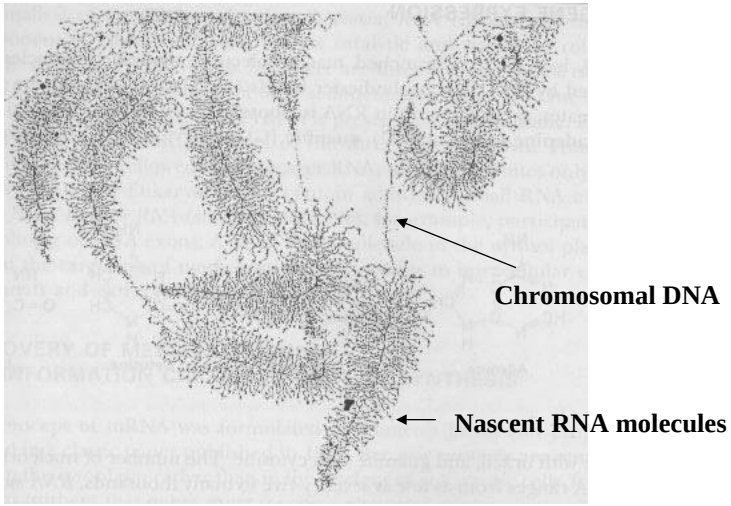




RNA is single stranded, but can base pair to itself



RNA is synthesized from a DNA template



All cellular RNA is synthesized by RNA Polymerases

- Requires a *template;*
  - Usually double stranded DNA (sometimes single stranded)
- All four *ribonucleotide triphosphates* (ATP, GTP,UTP, CTP) are required;
- *RNA polymerases* are enzymes which catalyze the initiation and elongation of RNA chains:

$(RNA)_{n\ residues}$ + ribonucleoside triphosphate ----> $(RNA)_{n+1\ residues}$ + PPi

- This process is called *transcription*

Mechanism of the chain elongation reaction

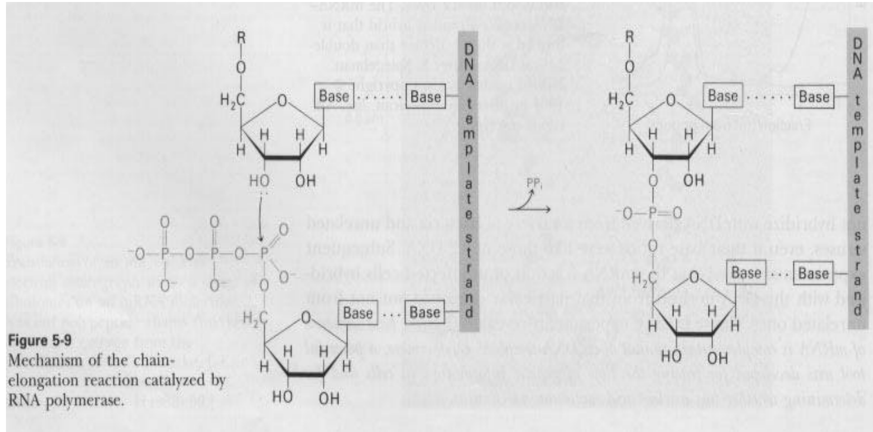
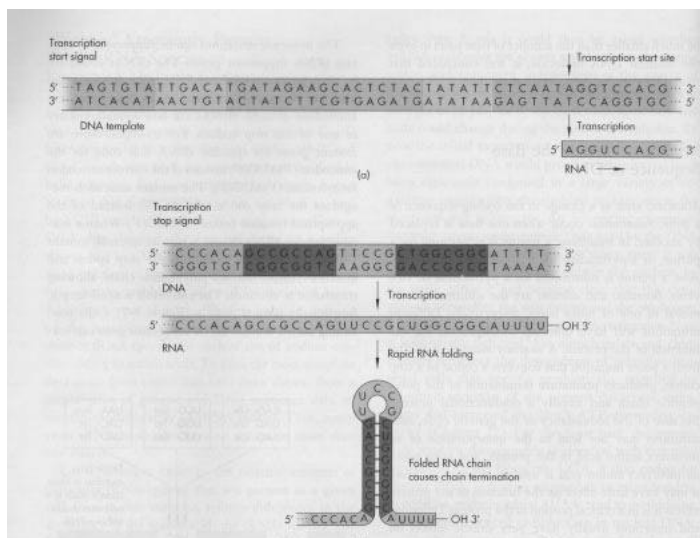


Figure 5-9
Mechanism of the chain-elongation reaction catalyzed by RNA polymerase.

- **The base sequence of the mRNA chain is the complement of that of the DNA template strand**

DNA contains start and stop signals for transcription



What is the structure of the genetic code?

- What coding ratio?
  - If each coding unit was only two nucleotides long, only 16 possible combinations (4 x 4 = 16); could code for 16 proteins maximum
  - So, coding ratio must be three or more
- Is the code overlapping or nonoverlapping?
- How is the correct group of amino acids read?
- Is the code *degenerate*?
  - Are there more than one code word for each amino acid?

Crick established the basic structure of the genetic code

- An amino acid is coded by a sequence of three bases--called a *codon.*
- The code is *nonoverlapping;* each group of three bases specifices only one amino acid.
- The sequence of bases is read sequentially from a fixed starting point.
- The code is degenerate--for many amino acids, there is more than one codon.

Elucidation of the genetic code

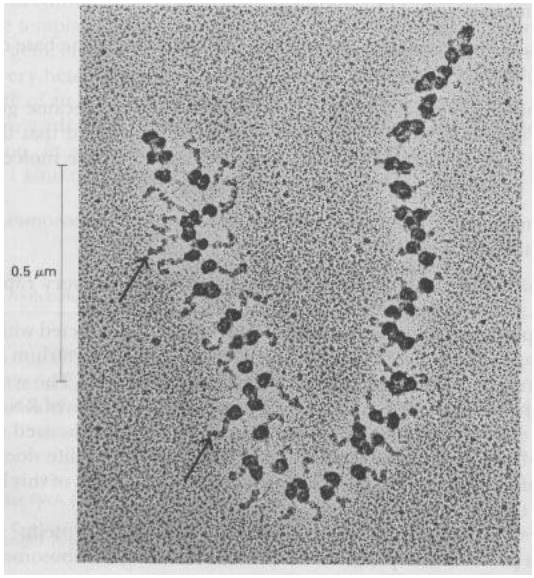
- Marshall Nirenberg showed that protein synthesis could occur in a cell free protein synthesizing system upon addition of a synthetic RNA molecule
- Adding the RNA poly(U) (UUUUU..) results in synthesis of a poly(phenylalanine) polypeptide;
  - "UUU" must code for phenylalanine
- This procedure was repeated for the other possible codon combinations

The Genetic code

TABLE 3-2
The Genetic Code

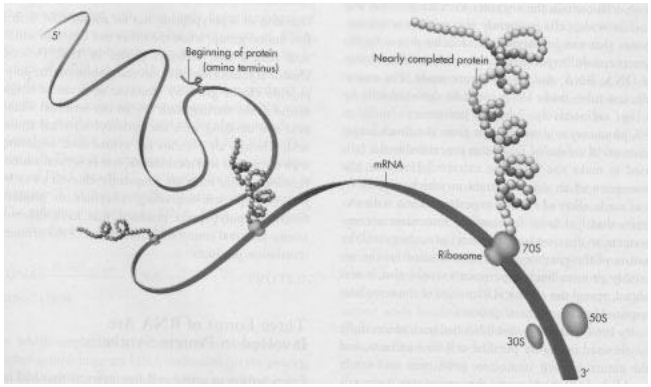
FIRST POSITION (5′ END)	SECOND POSITION U	C	A	G	THIRD POSITION (3′ END)
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Stop	Stop	A
	Leu	Ser	Stop	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

Protein synthesis occurs at ribosomes

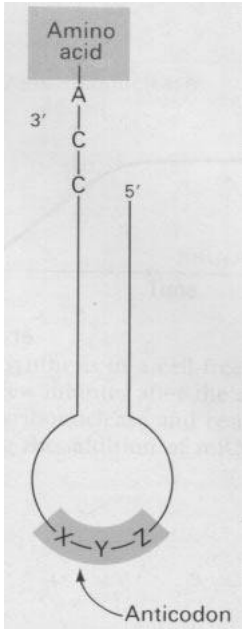


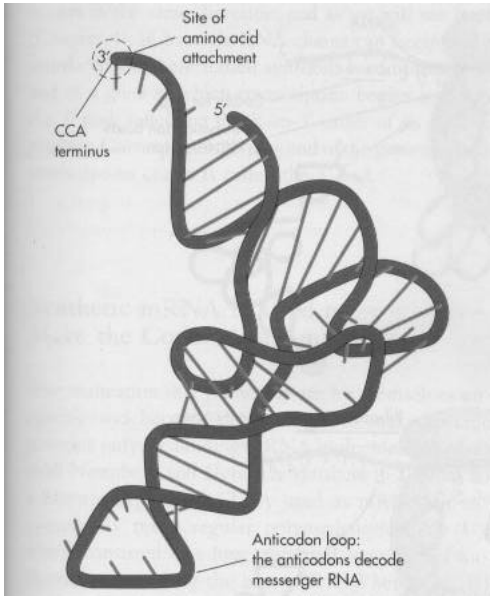
The molecular structure of ribosomes was recently elucidated

Mechanism of translation of mRNA

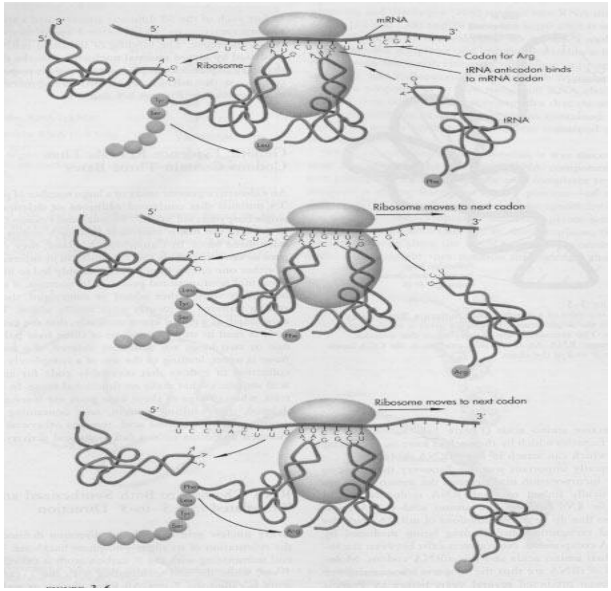


Transfer RNA are adapter molecules

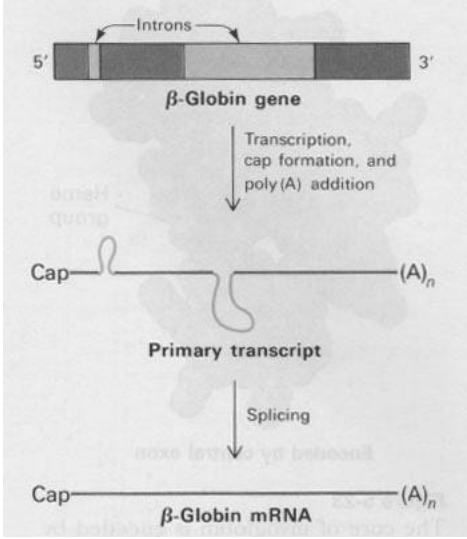




Mechanism of translation



Eukaryotic genes are divided into introns and exons



- *Introns* are intervening sequences in a gene which are transcribed into mRNA but do not code for amino acids.
- These sequences are spliced out of the mRNA before translation begins.
- The purpose of introns is unknown.

Mutations in DNA result in mutations in proteins

- *Nucleotide substitutions* may change a single amino acid.
- Nucleotide *deletions* or *insertions* can lead to large changes in protein sequence.
- Sequence of DNA can be used to understand the molecular basis of disease
- By modifying the sequence of DNA, new mutant proteins can be engineered with unique properties.