

## Topic 13: ORGANIZATION OF DNA INTO GENES AND THE SYNTHESIS OF PROTEINS (lecture 21-22)

### OBJECTIVES:

1. Understand the nature of the genetic code including such issues as word size, redundancy, codon bias and reading frame.
2. **You will not be required to memorize the code.** However, you should have a general idea as to which amino acids have the most redundant number of codons and which the least. You will also be asked to translate an RNA message using a codon dictionary.
3. Know the fundamental differences between RNA and DNA.
4. Review the process of transcription. Understand the roles of the promoter site, initiation site, TATA box, RNA polymerase II and transcription factors.
5. Describe the structure of typical eukaryotic genes and how this structure necessitates considerable modification of newly synthesized mRNA. What additional modifications take place on the mRNA and what functional roles do these modifications play?

**gene-** text definition is “a discrete unit of hereditary information consisting of a specific nucleotide sequence.

### Two fundamental kinds of genes:

1. **structural genes-** carry information for the synthesis of proteins; also code for specific kinds of RNA's like ribosomal RNA's and transfer RNA's
2. **regulatory genes-** genes that control the expression (activity) of other genes.

Overview of protein synthesis: two fundamental processes taking place.

1. **transcription-** the information in the DNA is transcribed (literally written) into another form; namely RNA is synthesized (a specific kind of RNA known as messenger RNA or **mRNA**). RNA differs from DNA (fig. 5.29) in that it (a) is generally single stranded, (b) it has the pyrimidine base uracil instead of thymidine and (c) the sugar is ribose.
2. **translation-** the information carried in the RNA molecule is translated into a sequence of amino acids thereby forming a protein.

fig. 17.2- comparison of bacterial vs. eukaryotic; in eukaryotic cells the mRNA is modified prior to leaving the nucleus; translation takes place in the cytoplasm.

**The Genetic Code-** recall that DNA molecules are polymers of four different nucleotides (A,T, G and C) strung together in a single strand; the second strand consists of a sequence of nucleotides that are complementary to the nucleotides on the

first strand. How can four different letters produce the information needed to make proteins which consist of unique sequences of 20 different amino acids?

fig. 17.3- in the 1960's it was discovered that amino acids are coded for by three nucleotides; on the mRNA molecule these nucleotide triplets are known as codons. Thus, each "word" consists of the unique sequence of three nucleotides. During translation, this information is recognized and translated into amino acids.

fig. 17.4- the code; as shown in RNA

reads 5'  $\longrightarrow$  3'

1. the code is redundant for many amino acids; that is, more than one codon codes for a particular amino acid- 6 for LEU and ARG; 4 for VAL, SER, PRO, THR and ALA  
different groups of organisms show "preference" for usage of particular codons; this is known as **codon bias**
2. AUG codes for "start" or MET
3. UAA, UAG and UGA code for "stop"

4. **reading frame**- where you begin to read the code is important

starting with A : 5' ACAGCCUUGGUG3' Thr Ala Leu Val

starting with C : 5' ACAGCCUUGGUG3' Gln Pro Trp

5. **near universality**- the above code is common to most living organisms including prokaryotes, fungi, plants and animals; this universality has been used as evidence for common descent of all organisms on earth. However, there are exceptions. For instance, for genes in mitochondria the RNA codon UGA is Trp not stop, AUA is Met not Ile and AGA and AGG are "stop" not Arg! Other differences are found in chloroplasts and in ciliated protozoans like *Paramecium*.

### Transcriptional formation of mRNA and mRNA processing.

fig. 17.7- initiation of transcription; synthesis goes from the 3'→5' end of DNA so the synthesized mRNA proceeds from 5'→3' direction

mRNA            5'----- $\longrightarrow$  3'  
DNA            3'----- 5'

1. **RNA polymerase II**- enzyme that catalyzes the synthesis of mRNA
2. **promoter site** - region of the DNA where the RNA polymerase initially binds and mRNA synthesis begins; consists of several regions  
(a) **initiation site**- where mRNA synthesis begins

(b) “upstream” region in eukaryotes has a sequence TATA known as the **TATA box**; RNA polymerase II looks for this sequence; associated with this region are proteins called **transcription factors**; polymerase cannot bind without these proteins being present.

3. RNA polymerase II binds to region at initiation site to commence transcription

fig. 17.6- transcription;

1. RNA polymerase II causes the DNA duplex to separate and causes the helix to untwist
2. complementary RNA nucleotides interact with corresponding nucleotides on the DNA strand
3. the polymerase catalyzes the formation of mRNA by adding nucleotides to the 3' end so the polymerase synthesizes mRNA in the 5'→3' direction but reads DNA from the 3'→5'
4. elongation- process by which the mRNA chain is lengthened; text says that this occurs at a rate of 60 nucleotides/sec
5. for a particular stretch of DNA, more than one molecule of RNA polymerase II can be making mRNA at the same time.
6. when the polymerase reaches the termination sequence (AATAA in eukaryotes) synthesis ceases and the newly formed mRNA is released

Fate of the mRNA- in prokaryotes the mRNA goes directly to translation (protein synthesis); not so in eukaryotes because (1) some regions of the newly formed mRNA must be excised and (2) the 3' and 5' ends must be modified to facilitate transport out of the nucleus as well as to make the message more stable.

(1) eukaryotic gene structure- genes actually consists of bundles of regions called **introns** and **exons** (fig. 17.9)

introns- regions that must be excised from message; do not contain structural information

exons- regions that code for specific stretches of amino acids

exons often code for regions (or **domains**) of the protein which correspond to important structural and/or functional components; see Fig. 17.11.

(2) fig. 17.8- ends are modified in two ways

(a) GTP is attached to 5' end to make a cap which protects the RNA from degradative enzymes as well as enhances translation

(b) the 3' end is modified by the addition of 30-100 adenine nucleotides forming a poly A “tail”; this tail enhances mRNA stability and also facilitates transport out of the nucleus

Fig. 17.10- when mRNA is synthesized the newly formed RNA consists of (a) 5' leader sequence [ often called the 5' untranslated region], (b) exons, (c) introns and (d) 3' trailer sequence [often called the 3' untranslated region]. Special protein/RNA complexes known as **spliceosomes** cut out the introns and the remaining exons plus leader and trailer sequences are reattached to form the message.

Note: molecular biologists take advantage of this poly A tail to isolate mRNA from other RNA's. They use a poly T sequence attached to a solid surface. The mRNA binds to the poly T, the other RNA's do not and are washed off. The mRNA can then be released by changing the chemical conditions yielding relatively pure mRNA!