DNA to Protein Supplemental Page

Read the article before checking out this detailed overview of how proteins are made from DNA!

1. Central Dogma of Biology: DNA is transcribed into RNA, which is then translated into a protein
2. The process by which DNA becomes a protein occurs in a sequence of several different phases and modifications. Here is a brief overview:
    a. Transcription (DNA → mRNA) (occurs in the Nucleus)
        i. Initiation
        ii. Elongation
        iii. Termination
    b. Post-Transcriptional Processing (also occurs in the Nucleus)
        i. 5' Capping
        ii. 3' Polyadenylation
        iii. RNA Splicing
    c. Translation (mRNA → Protein) (Happens in the Ribosome)
        i. Includes its own Initiation-Elongation-Termination Steps
    d. Protein Folding and Post-Translational Modifications
        i. Folding/Post-Translational Modifications
3. **Step #1 → Initiation of Transcription**
    a. RNA Polymerase identifies and binds to a specific region of DNA called the *Promoter*, located **upstream** in the DNA sequence.
        i. **Upstream** = in the direction of the 5'
        ii. *Promoter =* The region that the RNA polymerase identifies to start the process of transcription.
    b. The Promoter contains a specific DNA sequence that the RNA polymerase actually attaches to. It varies between Prok. and Euk.
        i. In Prokaryotes, this is called the **Pribnow Box** (The exact sequence is TATAAT), just before the
        ii. In Eukaryotes, this is called the **TATA Box**. (sequence is exactly what it sounds like)
    c. Transcription Factors
        i. A complex variety of General Transcription Factors helps the RNA Polymerase bind to the Promoter

d. **Enhancer** → DNA regions that connect to the Promoter using a Protein Structure called the *Mediator Complex* that helps send regulatory signals to the RNA Polymerase II as it attaches to the Promoter
e. DNA begins unwinding into an **Open Complex** in order to create a transcription bubble, which exposes the *single-stranded template* for the base-pairs to start
f. **Abortive Initiation**: RNA Polymerase, still attached to the Promoter at this point, *pulls in downstream DNA* in order to gain the energy needed to break the *grip on the promoter*
    i. The enzyme may release smaller RNA Transcripts before it escapes the grip of the promoter
g. **Promoter Escape –** The enzyme moves away from the promoter to begin the elongation step.
    i. A *C-terminal domain* is attached to the RNA Polymerase, leading to a recruitment of the next set of proteins needed, and also releases initiation factors

4. **Step #2 → Elongation (Transcription)**
    a. **Elongation** occurs when the enzyme travels through the DNA strand to create a new RNA strand
        i. 3' to 5' direction
        ii. *Untwists the DNA Double helix* to form the transcription bubble as it does
    b. **Base Pairing and Nucleotide Addition:** Uracil replaces Thymine in mRNA when pairing with Adenine. Guanine and Cytosine continue to pair together
        i. *Free RNA Nucleotides* are identified by the RNA Polymerase and used in the creation of the new Nucleotide
    c. Phosphodiester bonds are used to hold together the RNA nucleotides
    d. **RNA Polymerase** contains a proofreading mechanism that allows it to correct any errors in gene expression
        i. RNA Molecule begins to pull away from the DNA template as the DNA transcription bubble closes up in order to keep less of the DNA exposed

5. **Step #3 → Termination of Transcription**
    a. In **Termination,** the RNA Polymerase Enzyme detaches from DNA and releases the RNA.
        i. In Bacteria, a protein called **RHO** is used to help destabilize the interactions between DNA and RNA

   ii. In Eukaryotes, the transcription stops at a specific DNA sequence *(AAUAAA)*
      1. A group of proteins will then cut the pre-mRNA strand away and **free it for further processing**

6. Step #4 → Post Transcriptional Processing

a. **Pre-mRNA** has been created and needs to undergo some changes before it can be translated into a protein
   i. *Various modifications* are made in order to turn the mRNA into something **mature** and ready for export
b. **5' Capping** → The 5 prime end is attached with a modified Guanine Nucleotide called a 7-methylguanylate.
   i. Recognized by the cell in order to leave the nucleus
   ii. Protects mRNA from degradation
   iii. Helps small ribosomal subunits identify the mRNA in translation
   iv. Interacts with spliceosome in later steps
c. **3' Poly-A Tail**
   i. 3' end is modified through the Poly-A Tail
      1. 50-250 Adenine Nucleotides are added to the 3' end
   ii. Helps keep the mRNA stable and protected
d. **RNA Splicing**
   i. A Spliceosome made of proteins is responsible for cutting on the non-coding regions of the DNA (*Introns*) while keeping exons.
      1. Spliceosome identifies specific sequences at the ends of introns and breaks it down, degrading it as exons are connected together
   ii. Some RNA Molecules, called **Ribozymes,** can do splicing completely on their own without a spliceosome.

7. Step #5 → Initiation of Translation

a. The mRNA *assembles at a specific "start" site* to establish the correct reading frame for protein synthesis
b. ***Ribosomes:*** Subunits remain separate until initiation begins
c. ***Initiator tRNA*** → An initiator tRNA carries the Methionine (AUG, first amino acid)
   i. Proteins coordinate the assembly of the initiation
d. ***Eukaryotic Initiation:***
   i. The eIF4F Complex recognizes the 5' Cap added to the mRNA earlier, which recruits Ribosomes to the process

ii. The *Ribosomal's smaller (40S) subunit* binds to the 5' end of the mRNA, forming what is called the **43S Complex** which begins searching through the mRNA for the start codon, AUG
iii. Initiation Factors lead to the large unit being recruited in order to start the elongation.

8. **Step #6 → Translation**
    a. ***Elongation Cycle*** begins as the codons are now read by the newly formed ribosome and the amino acids are being brought into the ribosome to synthesize proteins
        i. Codon Recognition → Specific Amino Acid enters the Ribosome carried by tRNA
        ii. Peptide Bond Formation → Peptide bonds are created between amino acids
        iii. Translocation → The Ribosome shifts further into the mRNA
9. **Step #7 → Post Translational Modifications**
    a. The now fully-created protein will continue to undergo even more changes after its synthesis is complete in the ribosome
    b. **Protein Folding:** The amino acid change folds into a *very specific three dimensional shape* specific to the protein, going into a secondary structure and eventually a tertiary structure. In larger proteins, quaternary structure can form too.
        i. **Chaperones:** While proteins fold spontaneously, others require molecular chaperones in order to fold
    c. Additional molecules may be added to the protein once it's done being transcribed
        i. *Acetylation, Methylation, Phosphorylation, Etc.*
        ii. Carbohydrates can be added to form Glycoproteins
    d. Unneeded proteins are marked for destruction by a polypeptide called Ubiquitin
    e. Modified Proteins travel to the Golgi Apparatus