Chapter for today: Chap. XIV

Major points for the day:

1. "The central dogma": DNA → RNA → Protein
2. The genetic code
3. The process of translation (synthesizing proteins from an RNA template)

Reprise

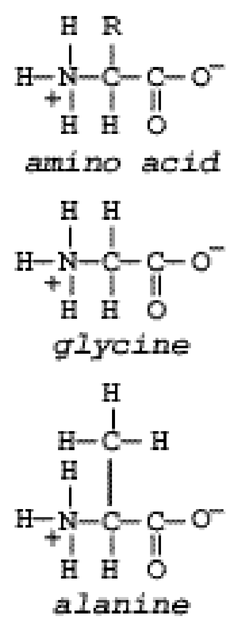
Last time I talked about the process by which information contained in the storage form of nucleic acid, DNA, is copied so that it can be used for useful work.

- The fact that DNA can be copied faithfully using the base pairing capacity of nucleotides along with the fact that it is chemically very stable make it an ideal hereditary molecule.
- Yet the stored information can not be used directly. Expressing the genetic instructions in DNA requires that the nucleotide sequence be copied into a form that can be used. That form is the second type of nucleic acid, ribonucleic acid or RNA. Copies of genes are termed messenger RNA or mRNA.
- Consistent with their different roles, DNA and mRNA have quite different fates in the cell. While DNA must remain unchanged for generations, the cell may need to express different parts of its genetic information depending on changes in internal or external factors. Therefore, mRNA is unstable; it is rapidly degraded by nucleases found in cells.

What was not discussed last time was how the mRNA is utilized to extract its genetic information. The vast majority of genes encode proteins and the process of producing proteins ("translation") is therefore centrally important to life.

The process of translation

In the early 1960s Francis Crick had hypothesized what the structure of a genetic code would be.



- While nucleic acids are polymers of four kinds of nucleotides, proteins are polymers of 20 kinds of amino acids
- An amino acid is a molecule in which a carbon atom is attached directly to three other chemical groups: an amino group (-NH_3) a carboxylic acid (-COOH) and a third variable group, called an R-group. The presence of the amino and acid groups gives the molecules their name.
- The R-group determines the identity of the amino acid and its special chemical properties. An amino acid with a single hydrogen as the R-group is called **glycine** while one with a methyl group (-CH_3) is **alanine**. R-groups can be polar, non-polar, acidic or basic. The chemical properties of the various R-groups confer on the protein its special structure and function.
- The problem was to understand what the relationship was between nucleotides and amino acids

Crick predicted that the genetic code would consist of code words, or "codons", of three nucleotides

- Given the four nucleotides (G, A, C or U) there are only 16 possible codons of two nucleotides (4^2), too few to specify the 20 amino acids

- But with 3 nucleotide codons there were 64 possible combinations (4^3). This number is easily sufficient to encode the 20 amino acids

Crick's predictions turned out to be correct. Much of the first half of the 1960s were spent in determining precisely how the information in mRNA is converted into proteins, and finding out what was the mechanism of translation.

The translational machinery

Transcription in eukaryotes occurs in the nucleus, where the DNA is located, while translation occurs in the cytoplasm

- The messenger RNAs (mRNAs) are transported out of the nucleus to the cytoplasm where they are translated on special molecular machines called ribosomes
- Ribosomes are made of a large number of proteins (ribosomal proteins) and several RNAs (ribosomal RNA or rRNA).
- Ribosomes are made of two distinct subunits termed the large and small subunit. These subunits have distinct roles in translation
- The small subunit interacts with the mRNA directly. This subunit has the role of making certain that the mRNA is decoded correctly
- In addition to its role in "reading" the mRNA, the ribosome also has the job of synthesizing proteins. This second role is the job of the large ribosomal subunit
- In the ribosome rRNAs perform most if not all of the catalysis. They both directly monitor reading of the mRNA and perform the chemistry required for synthesizing proteins

There are three kinds of RNAs involved in translation. You've already heard about mRNA and ribosomal RNA (rRNA). The first carries the genetic message and the second performs most of the catalysis of translation. But how does the ribosome decode mRNA codons into amino acids? The answer involves a third type of RNA called transfer RNA (tRNA).

- tRNAs are responsible for translating individual nucleotide "words" into specific amino acids
- To do this tRNAs have two "identities". First, they have a short nucleotide sequence that is complementary to the mRNA codon. This sequence is called the anticodon. On the ribosome, the tRNA anticodon must pair exactly with the mRNA codon. One job of the ribosome is to make certain that this interaction is perfect, and to reject any interaction that is less than perfect.
- The second identity is the ability to have a specific amino acid attached to it by a covalent bond. Special enzymes termed synthetases bind to tRNAs and attach an amino acid to the 3' end of the tRNA. This is another ester linkage, but this time between the carboxylic acid (COOH) of an amino acid and a hydroxyl (OH) at the 3' end of the tRNA

Determining the genetic code

A series of elegant experiments, principally done in the lab of Marshall Nirenberg at the NIH, allowed the complete determination of the genetic code by the mid 1960s

- The experiment made use of the role of the ribosome in pairing codons in mRNAs to anticodons in tRNAs
- In the first experiment, Nirenberg made a polymer of the nucleotide uridine monophosphate, or polyU, and bound it to ribosomes
- He then bound tRNAs to the ribosomes and attempted to find which of the classes of tRNAs bound to the mRNA

- The experiment showed that polyU specifically bound tRNAs that were carrying the amino acid phenylalanine (Phe), meaning that the codon UUU was interpreted by the ribosome to mean Phe. This was the first word of the genetic code ever identified.
- All the other codons were determined in the following years until we had a table showing the meaning of each of the 64 possible codons
- This was arguably the single most important set of scientific experiments of the 20th century since for the first time it opened up to view the way that the centrally important act of protein synthesis functioned

Structure of the genetic code

The genetic code is **redundant**, meaning it includes more codons than are necessary. There are an average of three codons for each amino acid

- The code consists of 61 codons specifying the 20 amino acids (sense codons), one of which also specifies initiation (AUG which codes for methionine), and three which specify termination (UAA, UAG and UGA)
- Almost all of these codon families consist of codons which differ only at the third position—GUU, GUC, GUA and GUG all specify valine; CCU, CCC, CCA and CCG all specify proline
- A gene consists of an AUG codon followed by a string of codons encoding amino acids (any of the 61 sense codons) and ending in a termination codon.

With a table of the genetic code, and the information about where to begin reading, we can determine the protein product of any gene

Function of the ribosome

The process of translation is occurs on the ribosome and is orchestrated by an assembly of tRNAs and protein factors in addition to the ribosome

To understand the process of translation you need to understand something about the structure of the ribosome.

The ribosome has two binding sites for tRNAs, called the decoding sites because the tRNA interacts with the mRNA to decode it while bound to them

- The aminoacyl-tRNA binds to the A site when it is recruited to the ribosome to decode the next codon during the process of extending a growing protein chain
- The second site, the P site, binds the peptidyl-tRNA, that is the tRNA to which the growing protein chain is attached
- When an aminoacyl-tRNA is accepted into the A site the ribosome immediately catalyzes the transfer of the growing protein to the from the 3’ end of the peptidyl-tRNA to the free amino group on the amino acid attached to the tRNA in the A site.

The decoding process requires protein factors called initiation or elongation factors

- Initiation factors regulate the process of translational initiation. Translation begins by a methionine tRNA binding to an AUG codon. The anticodon of the tRNA is complementary to the mRNA codon. Like DNA, the two nucleic acids have opposite polarity when paired (5’ to 3’).
- During initiation, only the small ribosomal subunit binds to the mRNA. After initiation the large subunit joins so that elongation can begin.

```
5’–AUG–3’ mRNA Codon
   |||
   UAC–5’ tRNA Anticodon
```

- Initiation occurs with the AUG codon bound at the ribosomal P site. The initiation factors are present to allow initiation only at correct AUG codons. They will reject any incorrect initiation events. After the correct initiation codon has been recognized the initiation factors dissociate from the complex and the large ribosomal subunit joins. The large subunit remember has the activity that actually polymerizes amino acids so protein synthesis can not begin until it joins.

Translation then continues with the addition of the second amino acid

- An elongation factor, EF-Tu, binds to all aminoacyl-tRNAs. EF-Tu deposits the tRNA into an empty A site. If the tRNA matches the codon displayed in the A site EF-Tu will dissociate, leaving the aminoacyl-tRNA behind. If it does not match then the EF-Tu-aminoacyl-tRNA complex dissociates
- As soon as the aminoacyl–tRNA binds to the A site, the ribosome catalyzes the transfer of the methionine to the free end of the amino acid in the A site, the amino end.
- After peptide transfer the ribosome moves the messenger RNA and tRNAs by three nucleotides, ejecting the methionyl–tRNA from the P site, and replacing it with the new peptidyl–tRNA (the two amino acid unit is now a peptide—a fragment of a protein). This process is catalyzed by a second elongation factor, EF-G

Translation consists of the repetition of the same steps over and over again:

- EF-Tu-aminoacyl-tRNA binds to the A site; EF-Tu dissociates
- The growing protein transfers to the amino acid in the A site
- EF-G promotes translocation of peptidyl-tRNA to the P site
- Then the process repeats again.

Each step of translation elongation occurs in less than a tenth of a second. An entire protein, which could consist of 500 amino acids, is synthesized in less than a minute.

The process of termination requires special termination factors that recognize the presence of a termination codon in the A site and catalyze peptide release, and ribosome dissociation

Gene expression must be regulated

Since the many cells in a body have very different roles, they also have distinct proteins that must be expressed to accomplish those roles

Even bacterial cells have different proteins that must be expressed in different physiological or environmental situations

Therefore, it is critical that the correct mRNAs be synthesized, and that they only be translated when their product is actually necessary

Next time we will consider how cells regulate the expression of genes