Lecture 1. Introduction to molecular biology. Part I.

Learning outcomes:

1. Describe the Chargaff, Griffith, Avery-MacLeod-McCarty, Hershey-Chase experiments and explain their significance.
2. Explain informational properties of macromolecules.
3. Explain the central dogma of molecular biology.
4. Briefly discuss the role of molecular biology in medicine.
5. Describe, identify and and draw the components of nucleosides and nucleotides.
6. Characterise and describe the chains of nucleic acids in DNA and RNA.

Many people believe that American biologist James Watson and English physicist Francis Crick discovered DNA in the 1950s. In reality, this is not the case. Rather, DNA was first identified in the late 1860s by Swiss chemist Friedrich Miescher. Then, in the decades following Miescher's discovery, other scientists--notably, Phoebus Levene and Erwin Chargaff--carried out a series of research efforts that revealed additional details about the DNA molecule, including its primary chemical components and the ways in which they joined with one another. Without the scientific foundation provided by these pioneers, Watson and Crick may never have reached

their groundbreaking conclusion of 1953: that the DNA molecule exists in the form of a three-dimensional double helix. Many people believe that American biologist James Watson and

English physicist Francis Crick discovered DNA in the 1950s. In reality, this is not the case. Rather, DNA was first identified in the late 1860s by Swiss chemist Friedrich Miescher. Then, in the decades following Miescher's discovery, other scientists--notably, Phoebus Levene and Erwin Chargaff--carried out a series of research efforts that revealed additional details about the DNA molecule, including its primary chemical components and the ways in which they joined with one another. Without the scientific foundation provided by these pioneers, Watson and Crick may never have reached their groundbreaking conclusion of 1953: that the DNA molecule exists in the form of a three-dimensional double helix.

Although few people realize it, 1869 was a landmark year in genetic research, because it was the year in which Swiss physiological chemist Friedrich Miescher first identified what he called "nuclein" inside the nuclei of human white blood cells. More than 50 years passed before

the significance of Miescher's discovery of nucleic acids was widely appreciated by the scientific community. Meanwhile, even as Miescher's name fell into obscurity by the twentieth century, other scientists continued to investigate the chemical nature of the molecule formerly known as nuclein. One of these other scientists was Russian biochemist Phoebus Levene. Levene is credited with many firsts. For instance, he was the first to discover the order of the three major components of a single nucleotide (phosphate-sugar-base); the first to discover the carbohydrate component of RNA (ribose); the first to discover the carbohydrate component of DNA (deoxyribose); and the first to correctly identify the way RNA and DNA molecules are put together.

During the early years of Levene's career, neither Levene nor any other scientist of the time knew how the individual nucleotide components of DNA were arranged in space; discovery of the sugar-phosphate backbone of the DNA molecule was still years away. The large number of molecular groups made available for binding by each nucleotide component meant that there were numerous alternate ways that the components could combine. Several scientists put forth suggestions for how this might occur, but it was Levene's "polynucleotide" model that proved to be the correct one.

Erwin Chargaff was one of a handful of scientists who expanded on Levene's work by uncovering additional details of the structure of DNA, thus further paving the way for Watson and Crick. Chargaff, an Austrian biochemist, had read the famous 1944 paper by Oswald Avery and his colleagues at Rockefeller University, which demonstrated that hereditary units, or genes, are composed of DNA. This paper had a profound impact on Chargaff, inspiring him to launch a

research program that revolved around the chemistry of nucleic acids. Of Avery's work, Chargaff (1971) wrote the following:

"This discovery, almost abruptly, appeared to foreshadow a chemistry of heredity and, moreover, made probable the nucleic acid character of the gene... Avery gave us the first text of a new language, or rather he showed us where to look for it. I resolved to search for this text."

As his first step in this search, Chargaff set out to see whether there were any differences in DNA among different species. After developing a new paper chromatography method for separating and identifying small amounts of organic material, Chargaff reached two major conclusions (Chargaff, 1950). First, he noted that the nucleotide composition of DNA varies among species. In other words, the same nucleotides do not repeat in the same order, as proposed

by Levene. Second, Chargaff concluded that almost all DNA--no matter what organism or tissue type it comes from--maintains certain properties, even as its composition varies. In particular, the amount of adenine (A) is usually similar to the amount of thymine (T), and the amount of guanine (G) usually approximates the amount of cytosine (C). In other words, the total amount of purines (A + G) and the total amount of pyrimidines (C + T) are usually nearly equal. (This second major conclusion is now known as "Chargaff's rule.") Chargaff's research was vital to the later work of Watson and Crick, but Chargaff himself could not imagine the explanation of these relationships--specifically, that A bound to T and C bound to G within the molecular structure of DNA

Chargaff's realization that A = T and C = G, combined with some crucially important X-ray crystallography work by English researchers Rosalind Franklin and Maurice Wilkins,

contributed to Watson and Crick's derivation of the three-dimensional, double-helical model for the structure of DNA. Watson and Crick's discovery was also made possible by recent advances in model building, or the assembly of possible three-dimensional structures based upon known molecular distances and bond angles, a technique advanced by American biochemist Linus Pauling. In fact, Watson and Crick were worried that they would be "scooped" by Pauling, who proposed a different model for the three-dimensional structure of DNA just months before they did. In the end, however, Pauling's prediction was incorrect.

Using cardboard cutouts representing the individual chemical components of the four bases and other nucleotide subunits, Watson and Crick shifted molecules around on their desktops, as though putting together a puzzle. They were misled for a while by an erroneous understanding of how the different elements in thymine and guanine (specifically, the carbon, nitrogen, hydrogen, and oxygen rings) were configured. Only upon the suggestion of American scientist Jerry Donohue did Watson decide to make new cardboard cutouts of the two bases, to see if perhaps a different atomic configuration would make a difference. It did. Not only did the complementary bases now fit together perfectly (i.e., A with T and C with G), with each pair held together by hydrogen bonds, but the structure also reflected Chargaff's rule

Although scientists have made some minor changes to the Watson and Crick model, or have elaborated upon it, since its inception in 1953, the model's four major features remain the same yet today. These features are as follows:

 DNA is a double-stranded helix, with the two strands connected by hydrogen bonds. A bases are always paired with Ts, and Cs are always paired with Gs, which is consistent with and accounts for Chargaff's rule.

 Most DNA double helices are right-handed; that is, if you were to hold your right hand out, with your thumb pointed up and your fingers curled around your thumb, your thumb would represent the axis of the helix and your fingers would represent the sugar-phosphate backbone. Only one type of DNA, called Z-DNA, is left-handed.

 The DNA double helix is anti-parallel, which means that the 5' end of one strand is paired with the 3' end of its complementary strand (and vice versa). As shown in Figure 4, nucleotides are linked to each other by their phosphate groups, which bind the 3' end of one sugar to the 5' end of the next sugar.

 Not only are the DNA base pairs connected via hydrogen bonding, but the outer edges of the nitrogen-containing bases are exposed and available for potential hydrogen bonding as well. These hydrogen bonds provide easy access to the DNA for other molecules, including the proteins that play vital roles in the replication and expression of DNA

Nucleic acid, naturally occurring chemical compound that is capable of being broken down to yield phosphoric acid, sugars, and a mixture of organic bases (purines and pyrimidines). Nucleic acids are the main information-carrying molecules of the cell, and, by directing the process of protein synthesis, they determine the inherited characteristics of every living thing. The two main classes of nucleic acids are deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). DNA is the master blueprint for life and constitutes the genetic material in all free-living organisms and most viruses. RNA is the genetic material of certain viruses, but it is also found in all living cells, where it plays an important role in certain processes such as the making of proteins.

Nucleic acids are polynucleotides—that is, long chainlike molecules composed of a series

of nearly identical building blocks called nucleotides. Each nucleotide consists of a nitrogen-containing aromatic base attached to a pentose (five-carbon) sugar, which is in turn attached to

a phosphate group. Each nucleic acid contains four of five possible nitrogen-containing bases: adenine (A), guanine (G), cytosine (C), thymine (T), and uracil (U). A and G

are categorized as purines, and C, T, and U are collectively called pyrimidines. All nucleic acids contain the bases A, C, and G; T, however, is found only in DNA, while U is found in RNA. The pentose sugar in DNA (2′-deoxyribose) differs from the sugar in RNA (ribose) by the absence of a hydroxyl group (—OH) on the 2′ carbon of the sugar ring. Without an attached phosphate group, the sugar attached to one of the bases is known as a nucleoside. The phosphate group connects successive sugar residues by bridging the 5′-hydroxyl group on one sugar to the 3′-hydroxyl group of the next sugar in the chain. These nucleoside linkages are called phosphodiester bonds and are the same in RNA and DNA.

DNA is a polymer of the four nucleotides A, C, G, and T, which are joined through a backbone of alternating phosphate and deoxyribose sugar residues. These nitrogen-containing bases occur in complementary pairs as determined by their ability to form hydrogen bonds between them. A always pairs with T through two hydrogen bonds, and G always pairs with C through three hydrogen bonds. The spans of A:T and G:C hydrogen-bonded pairs are nearly identical, allowing them to bridge the sugar-phosphate chains uniformly. This structure, along with the molecule's chemical stability, makes DNA the ideal genetic material. The bonding between complementary bases also provides a mechanism for the replication of DNA and the

transmission of genetic information. The double helical structure of normal DNA takes a right-handed form called the B-helix. The helix makes one complete turn approximately every 10 base

pairs. B-DNA has two principal grooves, a wide major groove and a narrow minor groove. Many proteins interact in the space of the major groove, where they make sequence-specific contacts with the bases. In addition, a few proteins are known to make contacts via the minor groove.

RNA is a single-stranded nucleic acid polymer of the four nucleotides A, C, G, and U joined through a backbone of alternating phosphate and ribose sugar residues. It is the first intermediate in converting the information from DNA into proteins essential for the working of a cell. Some RNAs also serve direct roles in cellular metabolism. RNA is made by copying the

base sequence of a section of double-stranded DNA, called a gene, into a piece of single-stranded nucleic acid. This process, called transcription

The Concept of Information Flow

The DNA macromolecule is an elegant structure that accommodates the need to store and transmit genetic information. A chromosomal DNA molecule is, in fact, two single chains of nucleotide subunits that are complementary to each other (Figure 1a). Thus there is present both the primary information, and the template necessary to make a new copy of that information within this chemically inert macromolecule. The redundancy helps to ensure that essential

information is not lost, because if one copy is damaged, the other can serve as a template for repair. The dangers inherent in transmitting genetic information are also reduced. If only a single chain were present, essential information might be lost either when the necessary template was assembled, or subsequently when the copies were produced.

Information transfer between DNA, RNA and protein macromolecules. (a) Complementary base pairing in a short segment of DNA. The four types of nucleotides are represented by the bars (solid or open, long or short), and each nucleotide can pair with only one complement. (b) Information transfer between DNA (drawn as in part a) and RNA (green). RNA can also be used to specify DNA by an analogous mechanism. (c) Information transfer from RNA to protein. The 20 types of amino acids (the four oval shapes represent four of these) are linked according to the sequence of nucleotides in the messenger RNA, which in turn was copied from the gene in part b. The next amino acid which will be added is shown with its adaptor transfer RNA (tRNA) attached. (d) The central dogma of molecular biology specifies the forbidden information transfers.

The central dogma of molecular biology predicts that a particular sequence of amino acids (a protein) cannot be used to specify or even alter a particular sequence of nucleotides (a gene). Instead, information flows from nucleic acids to proteins, in that an elaborate machinery exists to 'translate' the nucleic acid 'alphabet' to the amino acid 'alphabet' according to the rules of the genetic code. Cells exhibit no trace of a 'back-translation' machinery, and organisms can

transmit only their genes to their offspring. Even though the genetic material is not entirely constant, advantageous mutations do not arise in a directed manner. The predictions of the central dogma have withstood every challenge, and are likely to remain as the central organizing principles of molecular biology.

Modern molecular medicine encompasses the utilization of many molecular biological techniques in the analysis of disease, disease genes and disease gene function. The study of disease genes and their function in an unaffected individual has been possible by the development of recombinant DNA and cloning techniques. The basis of the term recombinant DNA refers to the recombining of different segments of DNA. Cloning refers to the process of preparing multiple copies of an individual type of recombinant DNA molecule. The classical mechanisms for producing recombinant molecules involves the insertion of exogenous fragments of DNA into either bacterially derived plasmid (circular double stranded autonomously replicating DNAs found in bacteria) vectors or bacteriophage (viruses that infect bacteria) based vectors.

The questions for self - control:

1. What is the subject of molecular biology? What are its connections with other sciences?
2. What was the history of molecular biology? What is the main carrier of genetic information, how it was discovered and studied?
3. What are the nucleosides, nucleotides and nucleic acids? What are their chemical structure and properties?

Recommended readings:

1. Russell P.J. iGenetics. A molecular approach. 3rd ed. 2009. Pearson. pp. 9-14.
2. Hartwell L. et al. Genetics. From genes to genomes. 4th ed. 2011. McGraw Hill. pp. 163-168.
3. Watson, James D. Molecular biology of the gene. 7th ed. Pearson. pp. 21-33.
4. John McMurry, et al. Fundamentals of General, Organic, and Biological Chemistry, 8th Edition. 2018. Pearson Education Limited. pp. 814-823.

Lecture 2. Introduction to molecular biology. Part II.

Learning outcomes:

1. Describe the three hypotheses of DNA replication.
2. Describe the Meselson-Stahl experiment and explain its significance.
3. Explain the molecular mechanism of semiconservative DNA replication.
4. Explain the role of main enzymes implicated in the replication process.
5. Explain proofreading mechanisms and error correction during DNA replication.

In 1953, after the double helix structure of DNA has just been discovered. One big question

concerned DNA replication. The structure of the DNA double helix provided a tantalizing hint about how copying might take place. It seemed likely that the two complementary strands of the helix might separate during replication, each serving as a template for the construction of a new, matching strand. There were three basic models for DNA replication that had been proposed by the scientific community after the discovery of DNA's structure

 Semi-conservative replication. In this model, the two strands of DNA unwind from each other, and each acts as a template for synthesis of a new, complementary strand. This results in two DNA molecules with one original strand and one new strand.

 Conservative replication. In this model, DNA replication results in one molecule that consists of both original DNA strands (identical to the original DNA molecule) and another molecule that consists of two new strands (with exactly the same sequences as the original molecule).

 Dispersive replication. In the dispersive model, DNA replication results in two DNA molecules that are mixtures, or “hybrids,” of parental and daughter DNA. In this model, each individual strand is a patchwork of original and new DNA.

The evidence that DNA replication was semi-conservative came from an elegant experiment completed by Matthew Meselson and Franklin Stahl. They labelled the parental DNA with a heavy isotope of nitrogen (15N) by growing bacteria in a growth medium that contained 15NH4Cl. They then grew the bacteria, in a medium that contained 14NH4Cl, in conditions such that any newly synthesised DNA would contain 14N. Since DNA replication is semi-conservative, after one round of DNA replication, each cell would have a DNA molecule that contains one ‘old’ parental strand labelled with 15N and one ‘new’ daughter strand labelled with 14N. This was shown by analysing the density of the DNA using density-gradient centrifugation. As predicted, they observed that the new daughter DNA molecule had a density consistent with the fact that it contained both 15N and 14N and that this daughter DNA contained one strand with 15N and another strand with 14N.

DNA replication is semiconservative, meaning that each strand in the DNA double helix acts as a template for the synthesis of a new, complementary strand.

DNA replication can be thought of in three stages; Initiation, Elongation, Termination

Initiation

DNA synthesis is initiated at particular points within the DNA strand known as ‘origins’, which

are specific coding regions. These origins are targeted by initiator proteins, which go on to recruit more proteins that help aid the replication process, forming a replication complex around the DNA origin. There are multiple origin sites, and when replication of DNA begins, these sites are referred to as replication forks.

Within the replication complex is the enzyme DNA Helicase, which unwinds the double helix and exposes each of the two strands, so that they can be used as a template for replication. It does this by hydrolysing the ATP used to form the bonds between the nucleobases, therefore breaking the bond holding the two strands together.

DNA Primase is another enzyme that is important in DNA replication. It synthesises a small RNA primer, which acts as a 'kick-starter' for DNA Polymerase. DNA Polymerase is the enzyme that is ultimately responsible for the creation and expansion of the new strands of DNA.

Elongation

Once the DNA Polymerase has attached to the original, unzipped two strands of DNA (i.e. the template strands), it is able to start synthesising the new DNA to match the templates. It is essential to note that DNA polymerase is only able to extend the primer by adding free nucleotides to the 3' end.

One of the templates is read in a 3' to 5' direction, which means that the new strand will be formed in a 5' to 3' direction. This newly formed strand is referred to as the Leading Strand. Along this strand, DNA Primase only needs to synthesise an RNA primer once, at the beginning, to initiate DNA Polymerase. This is because DNA Polymerase is able to extend the new DNA strand by reading the template 3' to 5', synthesising in a 5' to 3' direction as noted above.

However, the other template strand (the lagging strand) is antiparallel, and is therefore read in a 5' to 3' direction. Continuous DNA synthesis, as in the leading strand, would need to be in the 3' to 5' direction, which is impossible as we cannot add bases to the 5' end. Instead, as the helix unwinds, RNA primers are added to the newly exposed bases on the lagging strand and DNA synthesis occurs in fragments, but still in the 5' to 3' direction as before. These fragments are known as Okazaki fragments.

Termination

The process of expanding the new DNA strands continues until there is either no more DNA template left to replicate (i.e. at the end of the chromosome), or two replication forks meet and

subsequently terminate. The meeting of two replication forks is not regulated and happens randomly along the course of the chromosome.

Once DNA synthesis has finished, it is important that the newly synthesised strands are bound and stabilized. With regards to the lagging strand, two enzymes are needed to achieve this; RNAase H removes the RNA primer that is at the beginning of each Okazaki fragment, and DNA Ligase joins fragments together to create one complete strand.

The questions for self - control:

1. What were the three hypotheses of DNA replication and how the experiment of Meselson and Stahl revealed the real mechanism of this process?
2. What is the semiconservative mechanism of DNA replication and how this happens?

Recommended readings:

1. Alberts B. et al. Molecular biology of the cell / 6th ed. 2015. Garland Science, pp. 239-266.
2. Cooper G.M. The Cell: A Molecular Approach.
(https://www.ncbi.nlm.nih.gov/books/NBK9940/)

Lecture 3. Transcription of genetic information and mRNA processing

Learning outcomes:

1. Define the terms: transcription, promoter, enhancer, terminator.
2. Describe prokaryotic and eukaryotic RNA-polymerases' structure and functions.
3. Describe phases of transcription, explain the processes happening at each phase and their importance.
4. Explain the process, importance and difference of Rho-independent and Rho-dependent termination of transcription.
5. Explain the mechanism of polyadenylation, its importance.
6. Describe the structure of the cap fragment, its synthesis and functions.
7. Describe the splicing mechanism and its meaning.
8. Explain the effect of splicing on gene expression.

Transcription, the synthesis of RNA from DNA. Genetic information flows from DNA into protein, the substance that gives an organism its form. This flow of information occurs through the sequential processes of transcription (DNA to RNA) and translation (RNA to protein). Transcription occurs when there is a need for a particular gene product at a specific

time or in a specific tissue.

During transcription, only one strand of DNA is usually copied. This is called the template strand, and the RNA molecules produced are single-stranded messenger RNAs (mRNAs). The DNA strand that would correspond to the mRNA is called the coding or sense strand.
In eukaryotes (organisms that possess a nucleus) the initial product of transcription is called a pre-mRNA. Pre-mRNA is extensively edited through splicing before the mature mRNA is produced and ready for translation by the ribosome, the cellular organelle that serves as the site of protein synthesis. Transcription of any one gene takes place at the chromosomal location of that gene, which is a relatively short segment of the chromosome. The active transcription of a gene depends on the need for the activity of that particular gene in a specific cell or tissue or at a given time.
Small segments of DNA are transcribed into RNA by the enzyme RNA polymerase, which achieves this copying in a strictly controlled process. The first step is to recognize a specific sequence on DNA called a promoter that signifies the start of the gene. The two strands of DNA

become separated at this point, and RNA polymerase begins copying from a specific point on one strand of the DNA using a special type of sugar-containing nucleoside called ribonucleoside 5’-triphosphate to begin the growing chain. Additional ribonucleoside triphosphates are used as the substrate, and, by cleavage of their high-energy phosphate bond, ribonucleoside monophosphates are incorporated into the growing RNA chain. Each successive ribonucleotide is directed by the complementary base pairing rules of DNA. For example, a C (cytosine) in DNA directs the incorporation of a G (guanine) into RNA. Likewise, a G in DNA is copied into a C in RNA, a T (thymine) into an A (adenine), and an A into a U (uracil; RNA contains U in place of the T of DNA). Synthesis continues until a termination signal is reached, at which point the RNA polymerase drops off the DNA, and the RNA molecule is released.
Ahead of many genes in prokaryotes (organisms that lack a nucleus), there are signals called “operators” (see operons) where specialized proteins called repressors bind to the DNA just upstream of the start point of transcription and prevent access to the DNA by RNA polymerase. These repressor proteins thus prevent transcription of the gene by physically blocking the action of the RNA polymerase. Typically, repressors are released from their blocking action when they receive signals from other molecules in the cell indicating that the

gene needs to be expressed. Ahead of some prokaryotic genes are signals to which activator proteins bind to stimulate transcription.

Transcription in eukaryotes is more complicated than in prokaryotes. First, the RNA

polymerase of higher organisms is a more complicated enzyme than the relatively simple five-subunit enzyme of prokaryotes. In addition, there are many more accessory factors that help to

control the efficiency of the individual promoters. These accessory proteins are called transcription factors and typically respond to signals from within the cell that indicate whether transcription is required. In many human genes, several transcription factors may be needed before transcription can proceed efficiently. A transcription factor can cause either repression or activation of gene expression in eukaryotes.

The questions for self - control:

1. What is the transcription and which enzyme is necessary for this process? Describe the types, structure and functions of this enzyme.
2. What are the main stages of transcription? Describe each stage.
3. What is the structure of gene? Describe its regulatory and coding parts.
4. What is the processing of mRNA and which main stages it includes?

Recommended readings:

1. Alberts et al., pp. 333-362;
2. Russell, pp. 102-130;
3. Weaver, pp. 522-601.

Lecture 4. Translation of genetic information

Learning outcomes:

1. Explain the ribosome cycle and fidelity of translation.
2. Define the genetic code, tRNA, mRNA, codon, anticodon.
3. Describe the structure of tRNA and the mechanism of its charging.
4. Explain the scanning model of translation.
5. Explain the mechanism of translation and its phases.

6. Describe the structure of ribosomes and polysomes.

Translation, the synthesis of protein from RNA. Hereditary information is contained in the nucleotide sequence of DNA in a code. The coded information from DNA is copied faithfully during transcription into a form of RNA known as messenger RNA (mRNA), which is then translated into chains of amino acids. Amino acid chains are folded into helices, zigzags, and other shapes to form proteins and are sometimes associated with other amino acid chains. The specific amounts of amino acids in a protein and their sequence determine the protein's unique properties; for example, muscle protein and hair protein contain the same 20 amino acids, but the sequences of these amino acids in the two proteins are quite different. If the nucleotide sequence of mRNA is thought of as a written message, it can be said that this message is read by the translation apparatus in "words" of three nucleotides, starting at one end of the mRNA and proceeding along the length of the molecule. These three-letter words are called codons. Each codon stands for a specific amino acid, so if the message in mRNA is 900 nucleotides long, which corresponds to 300 codons, it will be translated into a chain of 300 amino acids.

Translation takes place on ribosomes—complex particles in the cell that contain RNA and protein. In prokaryotes (organisms that lack a nucleus) the ribosomes are loaded onto the mRNA while transcription is still ongoing. The mRNA sequence is read three bases at a time from its 5' end toward its 3' end, and one amino acid is added to the growing chain from its respective transfer RNA (tRNA), until the complete protein chain is assembled. Translation stops when the ribosome encounters a termination codon, normally UAG, UAA, or UGA (where U, A, and G represent the RNA bases uracil, adenine, and guanine, respectively). Special release factors associate with the ribosome in response to these codons, and the newly synthesized protein, tRNAs, and mRNA all dissociate. The ribosome then becomes available to interact with another mRNA molecule.

Any one mRNA is translated by several ribosomes along its length, each one at a different stage of translation. In eukaryotes (organisms that possess a nucleus) ribosomes that produce proteins to be used in the same cell are not associated with membranes. However, proteins that must be exported to another location in the organism are synthesized on ribosomes located on the outside of flattened membranous chambers called the endoplasmic reticulum (ER). A completed amino acid chain is extruded into the inner cavity of the ER. Subsequently, the ER transports the proteins via small vesicles to another cytoplasmic organelle called the Golgi apparatus, which in

turn buds off more vesicles that eventually fuse with the cell membrane. The protein is then released from the cell.

Translation has pretty much the same three parts, but they have fancier names: initiation, elongation, and termination.

Initiation ("beginning"): in this stage, the ribosome gets together with the mRNA and the first tRNA so translation can begin.

Elongation ("middle"): in this stage, amino acids are brought to the ribosome by tRNAs and linked together to form a chain.

Termination ("end"): in the last stage, the finished polypeptide is released to go and do its job in the cell.

The key components required for translation are mRNA, ribosomes, and transfer RNA (tRNA).

During translation, mRNA nucleotide bases are read as codons of three bases. Each 'codon' codes for a particular amino acid. Every tRNA molecule possesses an anticodon that is complementary to the mRNA codon, and at the opposite end lies the attached amino acid. tRNA molecules are therefore responsible for bringing amino acids to the ribosome in the correct order ready for polypeptide assembly

It is important to know that a single amino acid may be coded for by more than one codon. There are also specific codons that signal the start and the end of translation.

Aminoacyl-tRNA synthetases are enzymes that link amino acids to their corresponding tRNA molecules. The resulting complex is charged and is referred to as an aminoacyl-tRNA.

Initiation

For translation to begin, the start codon 5'AUG must be recognised. This is a codon specific to the amino acid methionine, which is nearly always the first amino acid in a polypeptide chain.

At the 5' cap of mRNA, the small 40s subunit of the ribosome binds. Subsequently, the larger 60s subunit binds to complete the initiation complex. The next step (elongation) can now commence.

Elongation

The ribosome has two tRNA binding sites; the P site which holds the peptide chain and the A site which accepts the tRNA.

While Methionine-tRNA occupies the P site, the aminoacyl-tRNA that is complementary to the next codon binds to the A site, using energy yielded from the hydrolysis of GTP. Methionine moves from the P site to the A site to bond to new amino acid there, and so the growth of the peptide has begun. The tRNA molecule in the P site no longer has an attached amino acid, and so leaves the ribosome.

The ribosome then translocates along the mRNA molecule to the next codon again using energy yielded from the hydrolysis of GTP. Now, the growing peptide lies at the P site and the A site is open for the binding of the next aminoacyl-tRNA, and the cycle continues. The polypeptide chain is built up in the direction from the N terminal (methionine) to the C terminal (the final amino acid).

Termination

One of the three stop codons enters the A site. No tRNA molecules bind to these codons so the peptide and tRNA in the P site become hydrolysed releasing the polypeptide into the cytoplasm.

The small and large subunits of the ribosome dissociate ready for the next round of translation.

The questions for self - control:

1. What is the genetic code and which properties it has?
2. What is biological translation? What are the main phases of translation? What happens at each phase?
3. What are the differences of translation process in eucariotes and procariotes?

Recommended readings:

1. Alberts et al., pp. 333-362;
2. Russell, pp. 102-130;
3. Weaver, pp. 522-601.