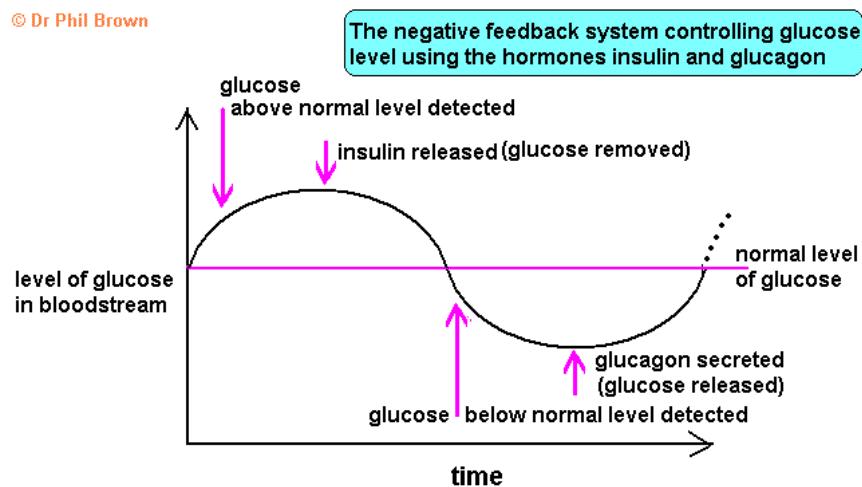


First lecture notes:

Welcome to the principles of biotechnology course. In this lecture, I will introduce you to this course. First, by giving an example of diabetes and need of insulin in the treatment of diabetes will be discussed. This case will highlight the need and application of Biotechnology. Then you will understand the basic definition of Biotechnology. You will also understand the difference between technology, engineering and science. This is to emphasize that biotechnology encompasses all these components to make commercial products. I will discuss the root, stem and fruits of biotechnology. These discussions will be reflected in the course content.

This video will help you to appreciate how insulin is made by biotechnology principles. Please go through these video and similar videos present on YouTube.

- <https://www.youtube.com/watch?v=DM4eMdrKDek>

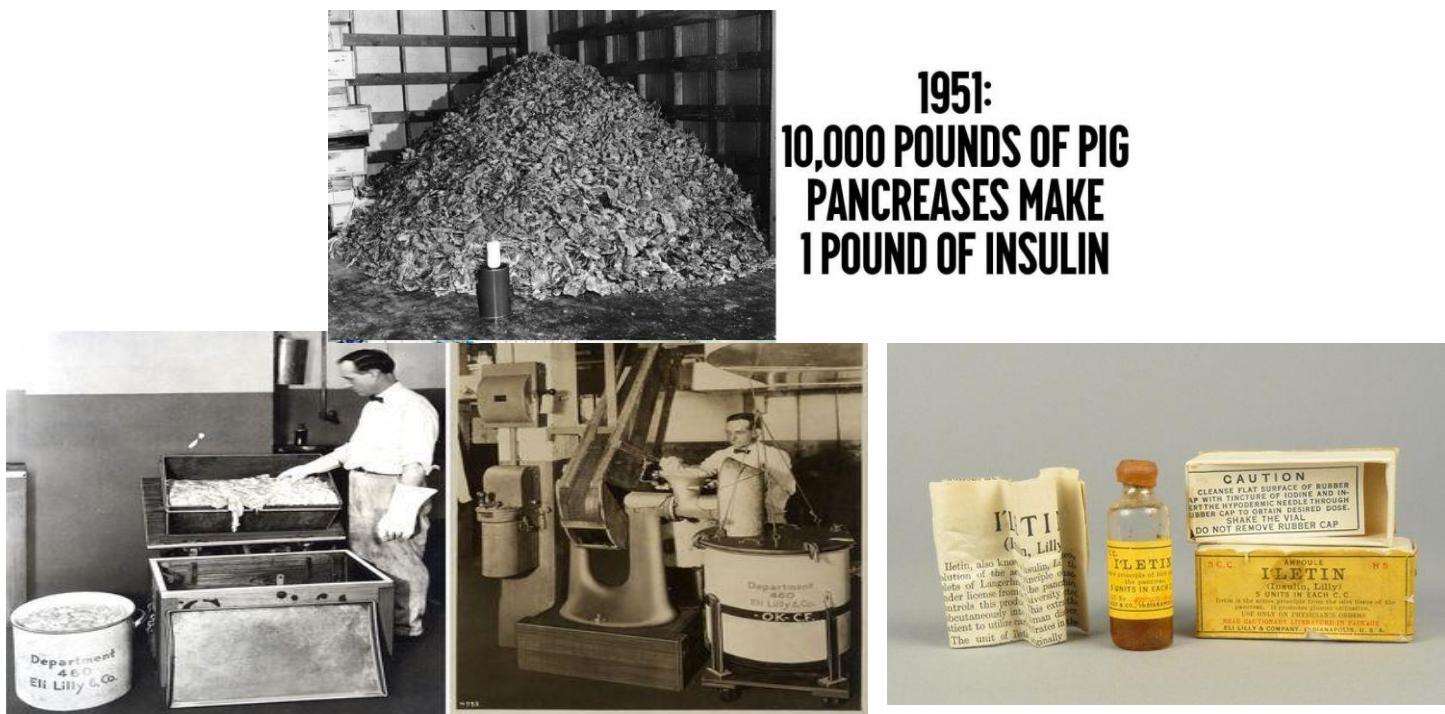


This representative graph presents the role of insulin to control blood glucose level. When we eat, glucose level increases in the blood, insulin helps it bring down the glucose level to a normal value. (you do not need to know the detail mechanism of this process, as these are part of biochemistry course/metabolism topic in BSBE department). If it goes below the normal level another molecule known as glucagon helps the releases of glucose from liver.

Since I am giving an example of insulin. I will stick myself to the main point. Imagine a case if insulin is not present in our body/ or it loses its function, or the cells which produce insulin are damaged/reduced (these topics will not be covered in detail, it is more of medicine topics). In summary one can say if insulin is not present, then our glucose levels will not be controlled and glucose level can rise in the blood. This condition can lead to one of the forms of a disease known as diabetes. This condition has short term or long-term health effects. Please read this topic, if you wish to know more about them (although it is not required for this class/or for the course)

It looks clear, the medicine for this condition can be insulin itself. Meaning in such patients, if one gives a therapy contain insulin from outside into the blood, the glucose level can be controlled and the patient disease condition can be managed.

Then the obvious question will arise. From where the insulin should come for injecting humans? Before the practice of Biotechnology, Insulin was isolated from animal sources such as pigs and cattle. To meet the demand of worldwide insulin, many animals were slaughtered (You need to remember that animals also use metabolic processes as present in humans and insulin is used by them for glucose control). You can see the figure given here representing some processes where insulin is extracted from pancreas (pancreas secret insulin) and eventually purified by multiple processes of purification. This animal insulin then used by doctors to inject humans for treating diabetes.



As you would have already realized the problems at least one is slaughtering of animals is unethical (now for research/application purposes worldwide, one has to take government permission with strong justification if animals to be used for medical research).

Second, there is some small variation in animal insulin and human insulin (Insulin a polypeptide, chain of amino acids). Therefore, in some instances, side reaction happens by injecting animal insulin to humans.

Now let's link the problem associated with the insulin obtained from animals to the development of biotechnology field which overcome these problems.

As I mentioned that biotechnology is an interdisciplinary area, where biological systems are used to make products. Discovery of biological processes and their controls (for example central dogma of life), cell biology, microbiology, biochemistry and many more fundamentals science discoveries have led to understand the insulin making process in the body. As you would have known that a gene is responsible in making a polypeptide and protein in a cell through transcription and translation processes. (These topics will be covered in future lecture). Some scientist starts fusing Genes of different organism, which led to the discovery of genetic engineering research. The gene which makes Insulin can be engineered in a bacterium to make insulin. That also in large amounts (please see the video, Large amounts can be made by understanding Biochemical engineering processes, There is a course in BSBE on Biochemical

engineering). We will discuss this more in detail in the course. The point is by doing this in short time, large amount of human insulin can be made in those biological systems which can be then purified. This whole thing comes in Biotechnology.

The question comes: what is biotechnology: The word is made up of two words, Bio and Technology, I can say in simple words technology made through biology. Technology refers to application of knowledge to practical purposes, commercialization, benefits to society.

Engineering refers to the design and manufacture of complex products. Science refers to the operation of general laws especially as obtained and tested through [scientific method](#). (These definitions are taken from English dictionaries)

Since biotechnology refers to technology made in biosystems (in cells, microbes), it is used for commercialization and benefits to the society. It means it is not just science and engineering it goes in the technology domain where products are made for societal usage. In other words, you can design and make a new car which can also fly (that is great for engineering new product), what whether this can be used commercially and at a price where people can make use of it, will come in technology domain. But all of them are important for Biotechnology, without knowing basic biology, one cannot practice genetic engineering, and putting both of them in practice with market requirement and sales cannot make a product for commercial usage. In our case Insulin.

Biotechnology in the context of science, engineering and technology can be represented in the form of a tree. Multiple fundamental concepts of biology and other disciplines are required and consists of roots of biotechnology. Genetic engineering tools and techniques make stem of the tree. These can only be put in practice if they are connected to roots. Products are represented as fruits of the tree. (in our example the fruit is insulin). In this course the contents are made in this way only. Please remember one course will not make you biotechnologist, but this course as the name suggest will teach you fundamental principles of biotechnology.

This is again putting the perspective of tree making biotechnology.

You will notice a small overlap with LIF101 course in background biology, however here it will be somewhat elaborate. This overlap was intentionally kept by the course committee of the institute, since most of you are from non-biology background. So, consider this overlap an advantage for connecting yourselves to basic principles of biology which are highly relevant to the principles of biotechnology. These concepts can be also found in advanced courses in BSBE. But they will be even further elaborated in those courses. Time, I will emphasize where and in which course you can take in future to understand such concepts in depth, if you will be interested to know about it. Please see the course content.

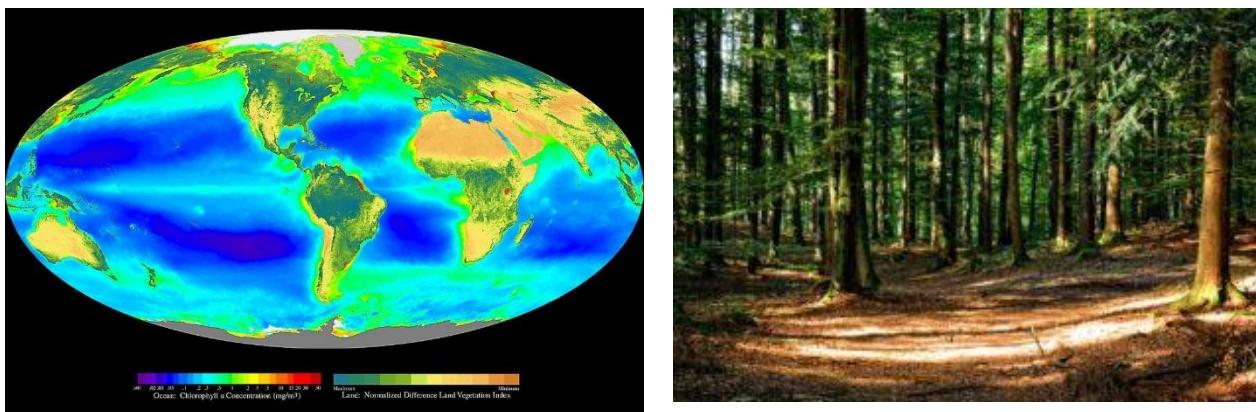
Lecture Notes 2: Understanding Biological Organization

Let's discuss about very basic school level concepts about biological organization. This is to set our tone towards other lectures as well as for principles of biotechnology course.

The Biosphere: If we visualize earth near enough, we see signs of life, green mosaics of forests, regions of land, ocean.

Ecosystem: All living things in a particular area, along with all the nonliving components of the environment with which life interacts, such as soil, atmospheric gases etc. Grasslands, forests, desert are some of the examples.

Communities: Living systems inhabiting in a particular ecosystem is called biological community, In a forest: trees, diverse animals, mushrooms, microorganisms etc.



Population: All the individuals of a species living within a particular area. In other word community can be also be defined by the set of populations that inhabit a particular area.

Organisms: Individual living things, each forest animal, or each tree can be called as an organism.

Organ and Organ systems: A leaf of a tree is called an organ, a body part that carries a special function is called an organ. Organs in human or plants are organised into an organ system, each team up for a larger function.

Tissues: While organs and organ systems can be seen by our eyes, but tissues are made up of group of cells that work together and performs a specialized function and can be visualized with the help of a microscope.

Cells: It is a fundamental unit of structure and function. Some organism like amoeba are single cell and carry all the function of life. Trillions of human cells make specialized tissue for function of our body.

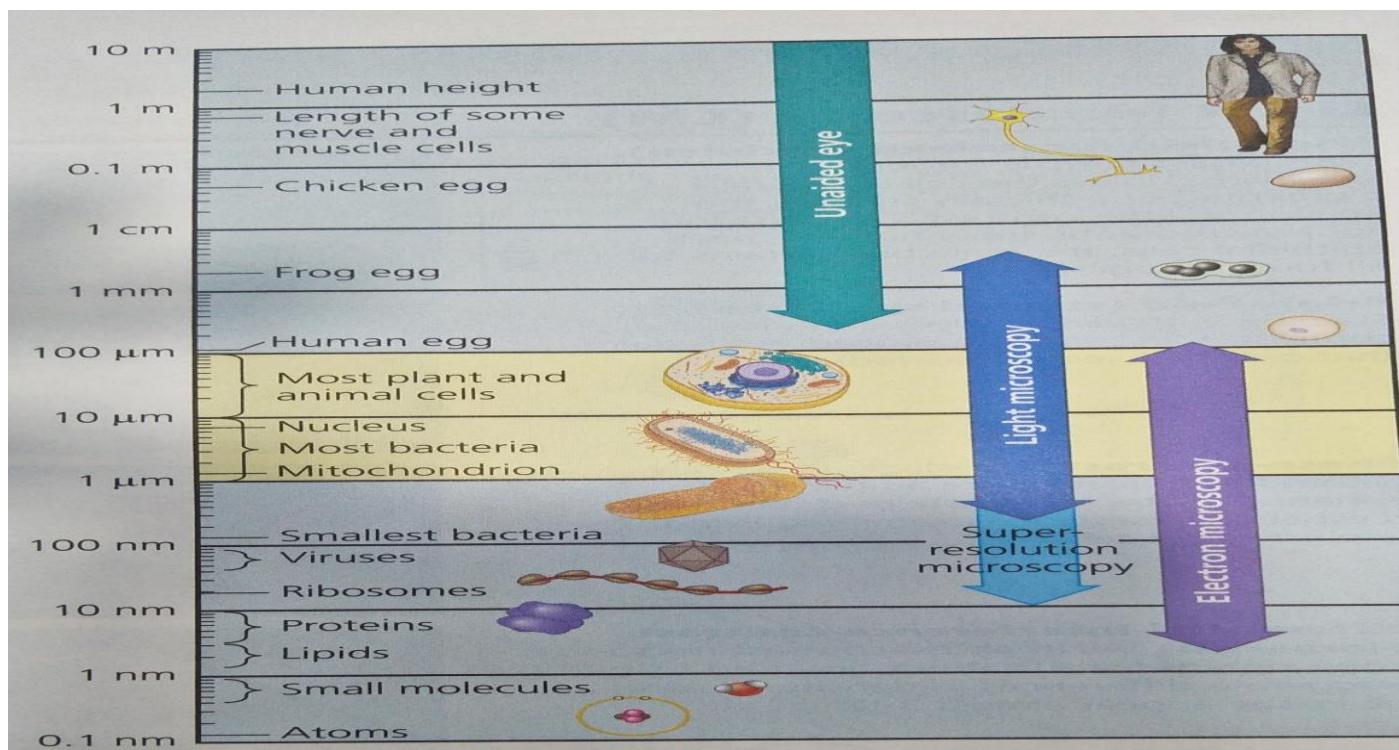
Organelles: Individual functional components present in the cells are known as organelles.

Molecules: Organelles are made up of molecules. A chemical structure made of two or more small chemical units.

One of the characteristics of living systems is the precise organization at different scales

The following figures represent, length scale of some of the living systems and components present in them. For example, atoms are in Angstrom size. Biomolecules can fit in nm size (1-10 nm). Between 10-100 nm comes various virus and small bacteria. Many organelles come between 1 um to 10 ums. Most of the cells can range from 10 -100 um size. Some cells are also lengthy and can fall between .1 m to 1 mm some nerve and muscle cells. Size more than 100 um can be seen by unaided eyes. After that light microscopy can be used up to 0,2 um. Below that high resolutions' methods are used. With these methods for deeper understanding of ultrastructure and high resolution, electron microscopy, x ray diffraction, cryoelectronic microscopy and NMR are used.

The details of these methods are taught in other courses of BSBE such as Analytical Methods and Structural biology, and Cell and Molecular biology courses.



Light Microscopy (LM)

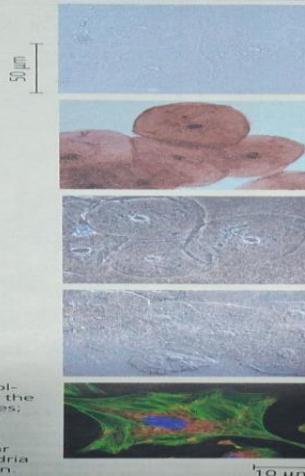
Brightfield (unstained specimen). Light passes directly through the specimen. Unless the cell is naturally pigmented, or artificially stained, the image has little contrast. (The first four light micrographs show human cheek epithelial cells; the scale bar pertains to all four micrographs.)

Brightfield (stained specimen). Staining with various dyes enhances contrast. Most staining procedures require that cells be fixed (preserved).

Phase-contrast. Variations in density within the specimen are amplified to enhance contrast in unstained cells, which is especially useful for examining living, unpigmented cells.

Differential-interference-contrast (Nomarski). As in phase-contrast microscopy, optical modifications are used to exaggerate differences in density, making the image appear almost 3-D.

Fluorescence. The locations of specific molecules in the cell can be revealed by labeling the molecules with fluorescent dyes or antibodies; some cells have molecules that fluoresce on their own. Fluorescent substances absorb ultraviolet radiation and emit visible light. In this fluorescently labeled uterine cell, nuclear material is blue, organelles called mitochondria are orange, and the cell's "skeleton" is green.



Confocal. The top image is a standard fluorescence micrograph of fluorescently labeled nervous tissue (nerve cells are green, support cells are orange, and regions of overlap are yellow). Below is a confocal image of the same tissue. Using a laser, this "optical sectioning" technique eliminates out-of-focus light from a thick sample, creating a single plane of fluorescence in the image. By capturing sharp images at many different planes, a 3-D reconstruction can be made. The standard image is blurry because out-of-focus light is not excluded.

Deconvolution. The top of this split image is a compilation of standard fluorescence micrographs through the depth of a white blood cell. Below is an image of the same cell reconstructed from many薄 images at different planes, each of which was processed using deconvolution software. This process digitally removes out-of-focus light and reassigns it to its source, creating a much sharper 3-D image.

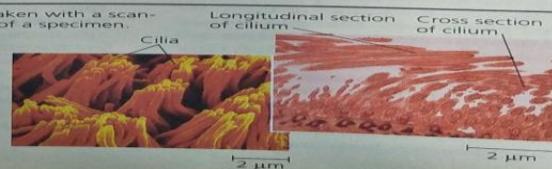
Super-resolution. On the top is a confocal image of part of a nerve cell, using a fluorescent label that binds to a molecule clustered in small sacs in the cell (vesicles) that are 40 nm in diameter. The image is blurry because 40 nm is below the 200-nm limit of resolution for standard light microscopy. Below is an image of the same part of the cell, seen using the new "super-resolution" technique. Sophisticated equipment is used to light up individual fluorescent molecules and record their positions. Combining information from many molecules in different planes "breaks" the limit of resolution, resulting in the sharp greenish-yellow dots seen here. (Each dot is a 40-nm vesicle.)



Electron Microscopy (EM)

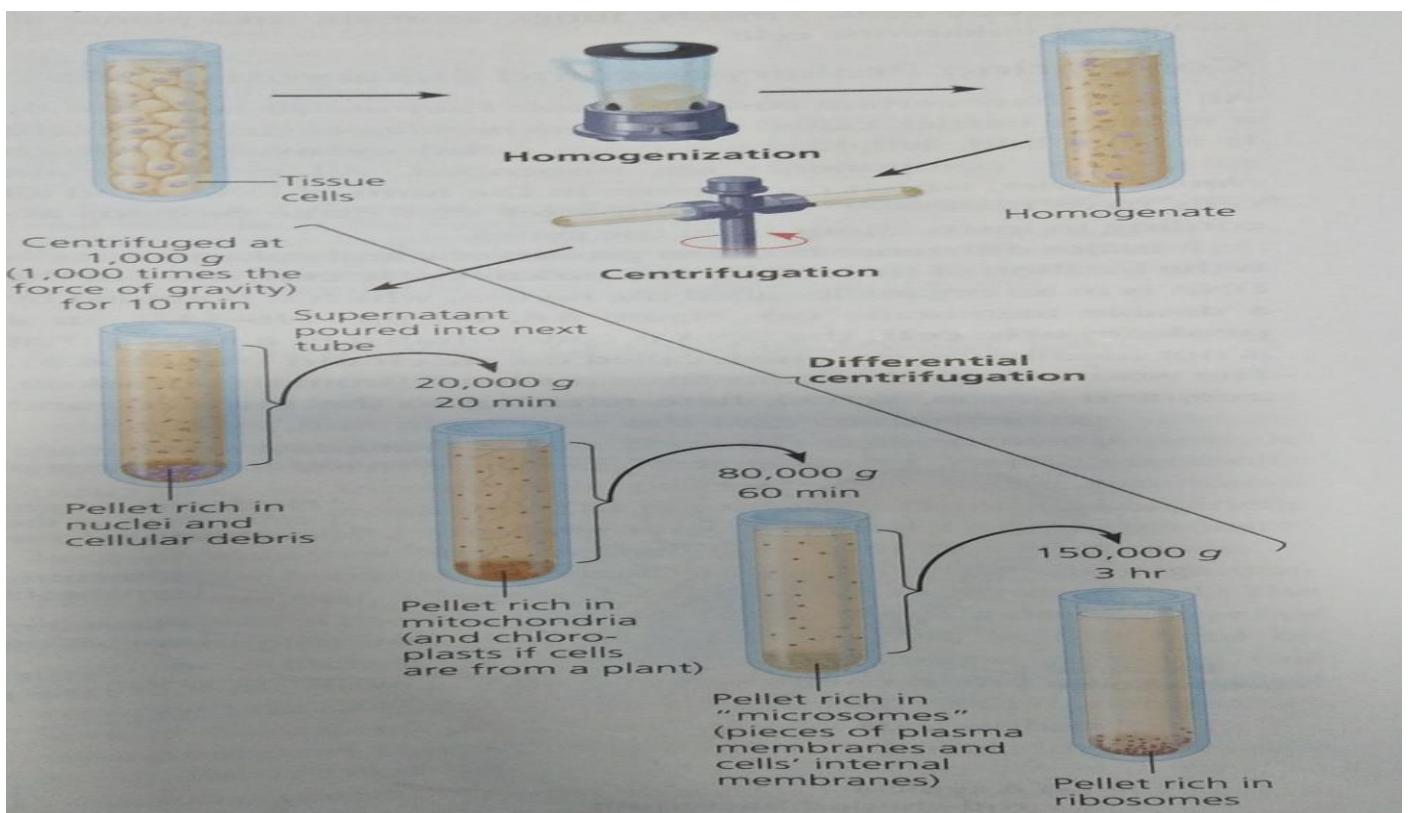
Scanning electron microscopy (SEM). Micrographs taken with a scanning electron microscope show a 3-D image of the surface of a specimen. This SEM shows the surface of a cell from a trachea (windpipe) covered with cilia. Beams of the electron beam move inhaled debris upward toward the throat. The SEM and TEM shown here have been artificially colorized. (Electron micrographs are black and white, but are often artificially colorized to highlight particular structures.)

Abbreviations used in this book:
LM = Light Micrograph
SEM = Scanning Electron Micrograph
TEM = Transmission Electron Micrograph



Transmission electron microscopy (TEM). A transmission electron microscope profiles a thin section of a specimen. Here we see a section through a tracheal cell, revealing its internal structure. In preparing the TEM, some cilia were cut along their lengths, creating longitudinal sections, while others were cut straight across, creating cross sections.

Most of the times to study different cell components cell fractional method is used. This is based on the size and density of the given component to be separated. In this method (please see diagram below), cells are homogenized in a blender to break them. The resulting mixture is known as homogenate is centrifuged. The supernatant (that is liquid part) is poured into another tube and centrifuged at a higher speed for a longer time. The process is repeated several times. This will lead to series of pellet (sedimented material at the bottom of the tube). These pallets will contain a cell component. This a very standard technique practiced in Biology.

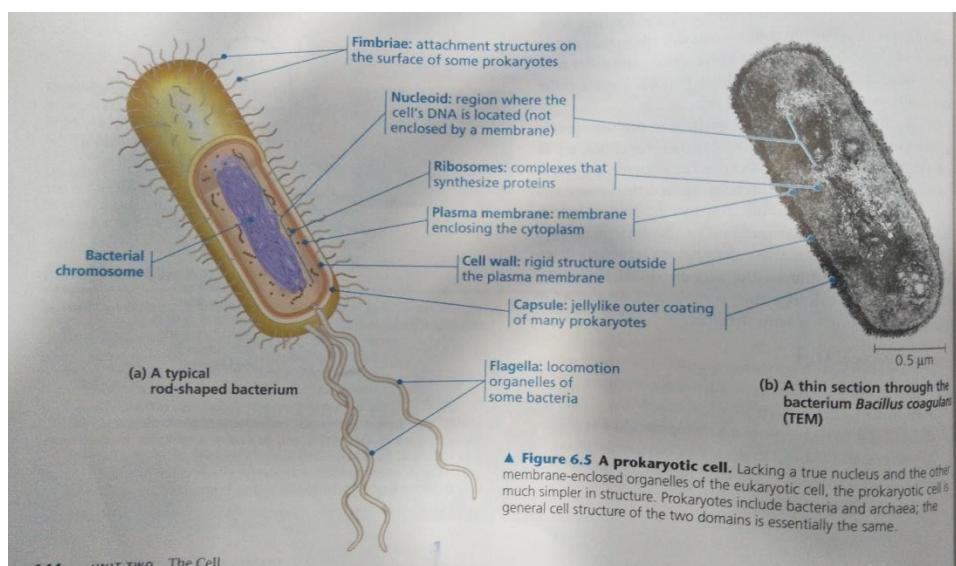


Lecture Notes: Functions of Cell Parts

Let's continue with the review of living systems. In the previous slide, I reviewed basic self-organization of components which makes living systems and their scales and visualization. In this topic, we will revisit the organelles functions. For principles of Biotechnology we do not need to go very details of each organelles. That, in future can be studied in a course of cell and molecular biology of BSBE Department. However, for this course you must know the name of each organelle and its function.

First you need to remember that cells can be divided into two types in living systems: Prokaryotic and another one is Eukaryotic. Prokaryotic cells are present in the domain of Life (Domain word comes from the classification of different organisms) of organisms Archaea and Bacteria. We will keep on bringing bacterial context in the whole course, so keep in mind that bacterial cells are prokaryotic. Eukaryotic cells are present in the domain Protists, Fungi, Animal and Plant cells. While we will keep on discussing bacteria, animal and plant cells; If you wish to know little more about Archaea, Protists, Fungi please google it and you will find a lot of reading material. Although it is not important for the course. They are subjects in themselves.

The prokaryotic cell is much simpler, yet complete all activities of life. We will keep on discussing about them at various stages of the course. They lack a true nucleus. The chromatin (DNA+proteins) is not surrounded/enclosed in the membrane. This is known as nucleoid. Prokaryotes also lack most of the membrane enclosed organelles. But contains protein synthesis organelle, ribosomes. Please look the diagram for other features of a typical prokaryote. See cell wall, capsule surrounding a prokaryotic cell. You also observe thread like structure known as fimbriae and long thread like structures Flagella. Note their function. Time to time these features will appear during the concepts covered in subsequent lectures.



By looking the diagram below, one can summarize the structure and function of cell parts of eukaryotic cell. You will be able to notice the difference between prokaryotic and eukaryotic cells. Please see and compare your self

1) Plasma membrane: It is made up of double layer of lipids, within which proteins are embedded. Some also attaches sugar groups. It serves as the external cell barrier, acts in the transport of substances into and out of cell, cell-cell communication.

2) Cytoplasm: Cellular region between nuclear and plasma membrane. It consists of fluid containing dissolved solutes, stored nutrients, secretory products and organelles.

3) Mitochondria: Rod like double membrane structures, the inner membrane is folded in cristae. The main site for adenosine triphosphate synthesis (ATP).

4) Ribosomes: Dense particles consisting of two subunits, each consists of ribosomal RNA and Proteins. They are free or bound to endoplasmic reticulum. Sites for protein synthesis. You will know about their function in detail in the topic translation.

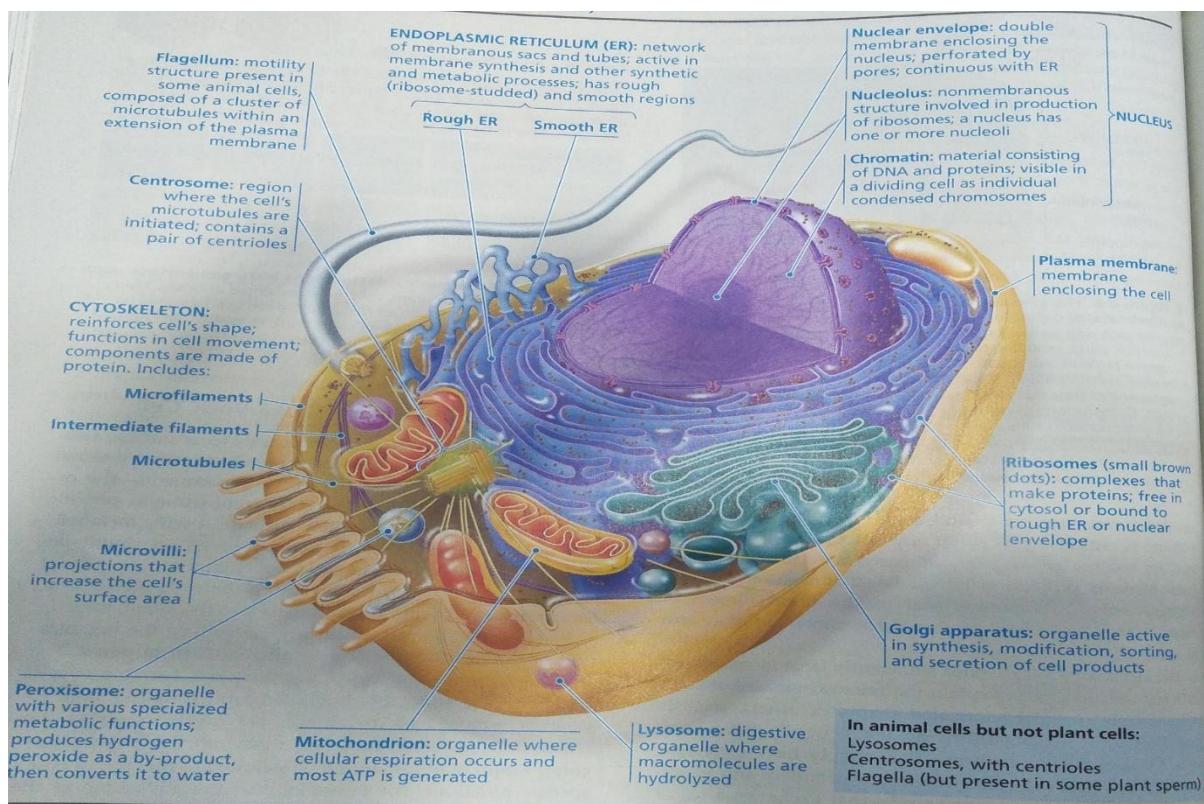
5) Rough endoplasmic reticulum: Membrane systems enclosing a cavity and externally bound with ribosomes due to which they appear rough. Protein modification, proteins are bound in vesicles of ER and transported to Golgi. We will talk more on one of the functions of this later.

6) Smooth endoplasmic reticulum: Membrane systems of sacs and tubules. They are sites of lipids and steroid synthesis and detoxification (liver).

7) Golgi Apparatus: A stack of smooth membrane sacks and associated vesicles. Packages, modifies and segregates proteins for secretion from cells.

8) Lysosomes: Membrane sacks containing hydrolases (digestive enzymes). These are sites for intracellular digestion.

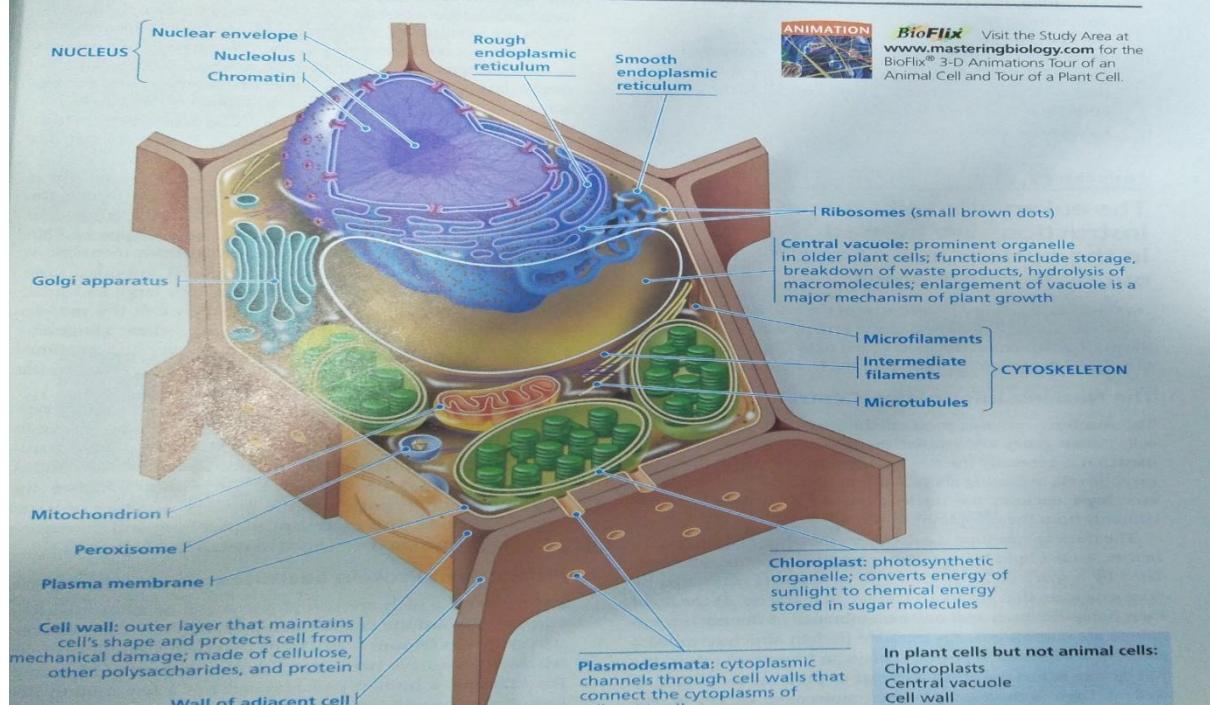
9) Nucleus: Largest organelle, surrounded by nuclear envelop, contains fluid: nucleoplasm, nucleoli and chromatin. It is the control center of the cell, responsible for transmitting genetic information and provides instructions for protein synthesis.



In plants, the following cell parts are important to mention. Others described in the previous slides hold true for plant cells as well.

- 1) Central Vacuole: Large membrane enclosed compartment. It is used to store ions, waste products, pigments and protective compounds.
- 2) Chloroplasts: Membrane enclosed organelles containing stacked structures called of chlorophyll containing membrane sacs called thylakoids surrounded by inner fluid (stroma). Chloroplasts are sites of photosynthesis.

Plant Cell (cutaway view of generalized cell)



I have restricted myself to these cell parts only as they will be frequently appearing in our discussion during the course. More focus will be given on Nucleus part that is chromatin where DNA, the genetic material/molecule of life is located, since genetic engineering/biotechnology involves DNA molecules.

Lecture Notes 4: The Nucleus: Center of Information

Let's elaborate on the structure and function of the nucleus. In the previous lecture, I introduced the nucleus function and its part. As we will move now towards DNA, the molecule of life. But before that, I wish to elaborate to some level about the nucleus in which DNA is present in the chromosomes.

The nuclear envelope is a double membrane structure. These membranes are made up of lipids. They are separated by about 20-40 nm. The nuclear envelope has perforations known as pores. These pores are about 100 nm in size. At the lip of the pores, the layers are continuous. Please see the close-up view of the pores in the diagram given below. Each pore is made up of a pore complex made up of protein molecules. This complex regulates the entry and exit of proteins, RNA and large complexes of macromolecules. Towards the inner surface, a netlike array of filaments is present that maintains the shape of the nucleus by mechanically supporting the nuclear envelope. This is known as Nuclear Lamina.

The nucleolus is a prominent structure present in the nucleus. It appears in the electron microscope as a dense granule. Here a Ribosomal RNA (rRNA) is synthesized from instructions in the DNA. Also, proteins are imported from the cytoplasm and assembled with rRNA into large and small units of Ribosomes. These subunits are exported out of the nucleus through the nuclear pores to the cytoplasm, where both units assemble into a ribosome. Details of these processes are not part of this course.

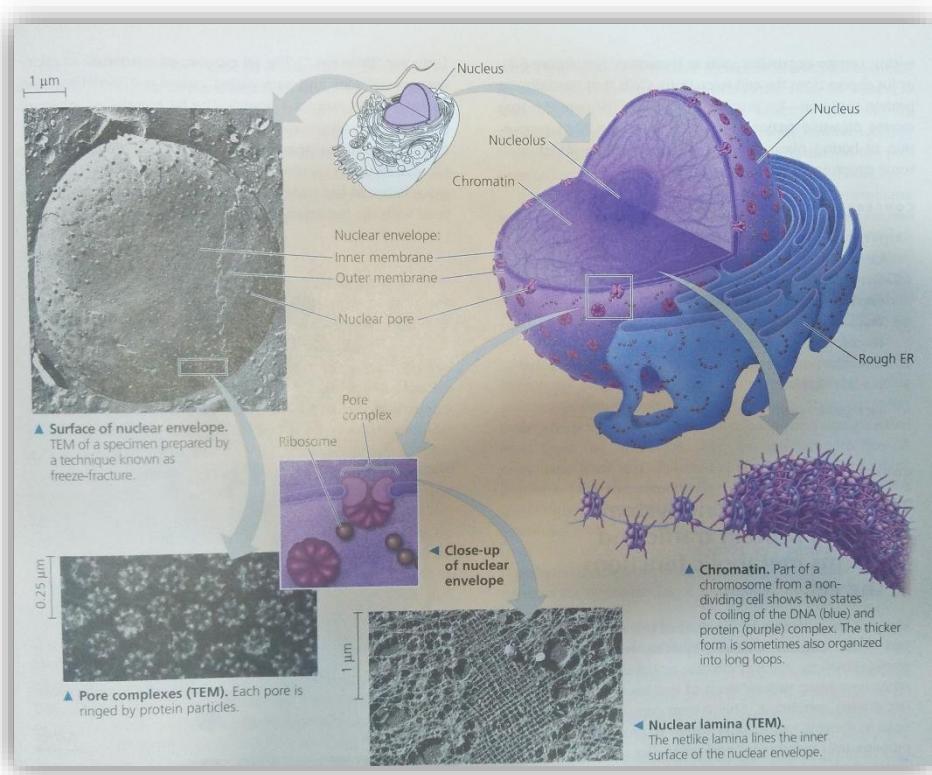


Figure 6.9 from Campbell biology page 149

As mentioned in the previous lecture, DNA and protein association makes chromatin and present in the nucleus. The organization of it in the Chromosomes is represented in the figure below.

Chromosome Organization: DNA is organized into discrete units called chromosomes, that carry genetic information. Each chromosome contains one long DNA molecule associated with proteins. Some of these proteins help to coil the DNA molecule of each chromosome, reducing its length to fit in the nucleus.

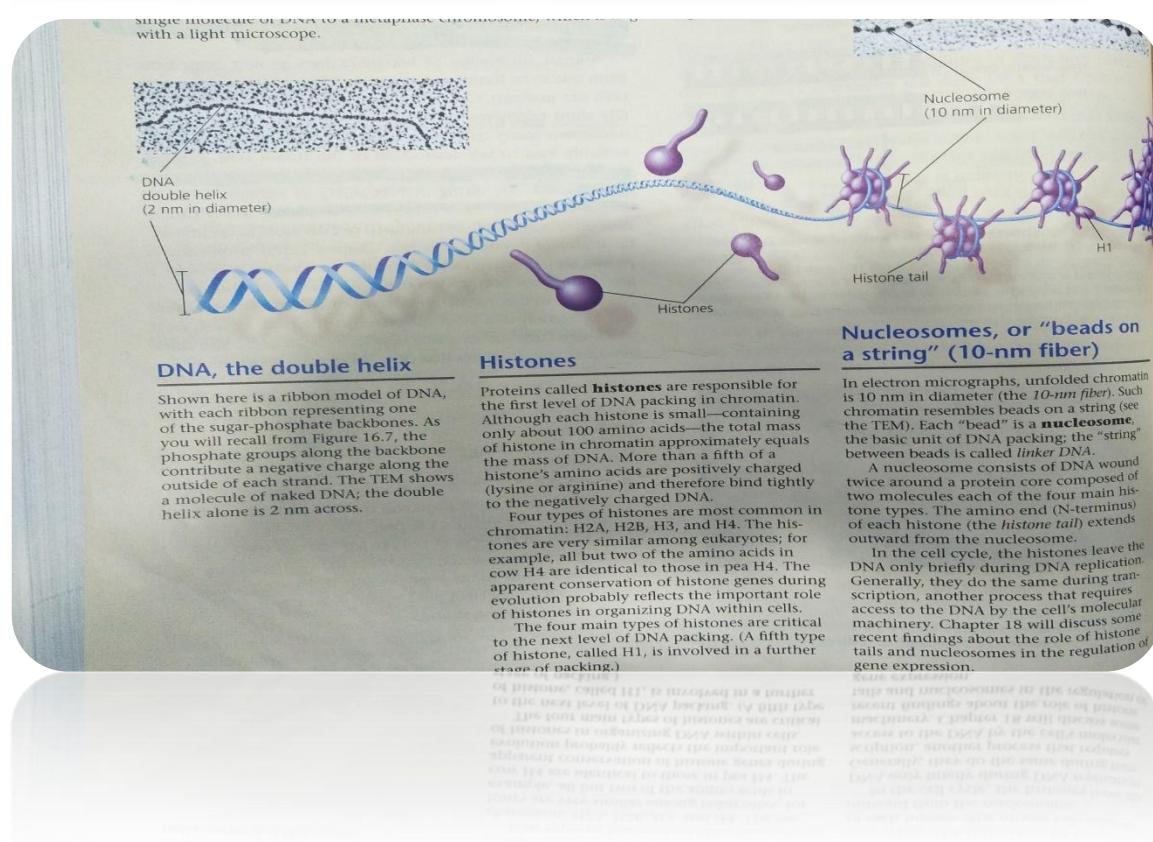
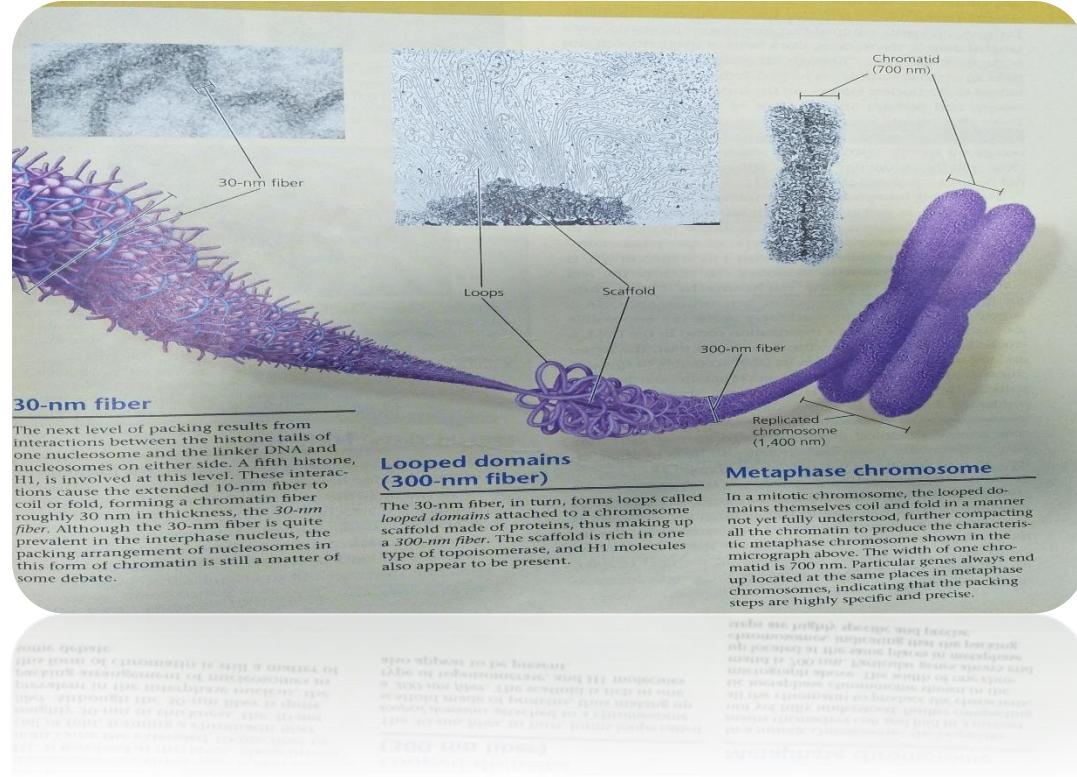
When the cell is not dividing. The DNA complexed with proteins (chromatin) looks like diffused mass (spread structures) in the micrographs (photo taken from a microscope), even though discrete chromosomes are present. As cells prepare to divide, the chromosomes coil, becoming thick enough and can be distinguished from one another. (details of this process are not part of this course). Each eukaryotic species has a characteristic number of

chromosomes. Humans have 46 chromosomes in the nucleus. Therefore, there will be 46 DNA molecules. Sex cells contain half of it. We will talk about cell division after replication of DNA molecule will be covered.

A chromosome is compacted and if viewed at a different scale, the organization of it can be understood as outlined below.

Bottom-up organization: DNA, proteins (histones), nucleosomes (10 nm), 30 nm structures, 300 nm fiber structures, 1400 nm chromosome

DNA and histone proteins make nucleosomes which can be compacted in 30 nm structures and then fibers appearing structures in the chromosomes. These assemblies are highly precise, the rules of which are not yet fully understood and there is a lot of scope in understanding it.



Lecture Note 5: Molecule of Life

In the previous lecture I talked about the Nucleus, its parts, and its function. DNA and proteins make chromatin.

We know now that DNA is the molecule which is the information centre of the cell and therefore control all the activities of the cell and thus biochemistry, physiology of an organism.

Mendel's suggested centuries ago about heritable factors and Morgan suggested the presence of Gene on the chromosome. These are parts of classical genetics and the chromosomal basis of inheritance. In this course, we will not discuss them, but a molecular basis of inheritance that relies on DNA transferring information. But before going to understand those details, this lecture will focus on how scientists come to know that DNA is the genetic material?

Around 1940; protein was the hot contender because of the heterogeneity of protein structure and specificity of their functions, which seems more relevant as well as genetic material. DNA perhaps looked to be simpler, whose physicochemical properties look more or less uniform and not much knowledge about it. Most of the work done to prove these facts was not animals or plants but bacteria and viruses.

In 1928, to discover Vaccine against pneumonia (infection of lungs), British medical officer Frederich Griffith was carrying some experiments. He was Studying *Streptococcus pneumoniae* (genus, species, scientific names are written in italics). He Had two strains (species can be further divided into strains having some differences, varieties). One was pathogenic (it causes pneumonia), smooth strain, S (symbol of strain), outside they have capsule which protects them from animal defence system. Another was non-pathogenic (harmless), rough strain, R (symbol of strain). Lacks capsule and are non-pathogenic.

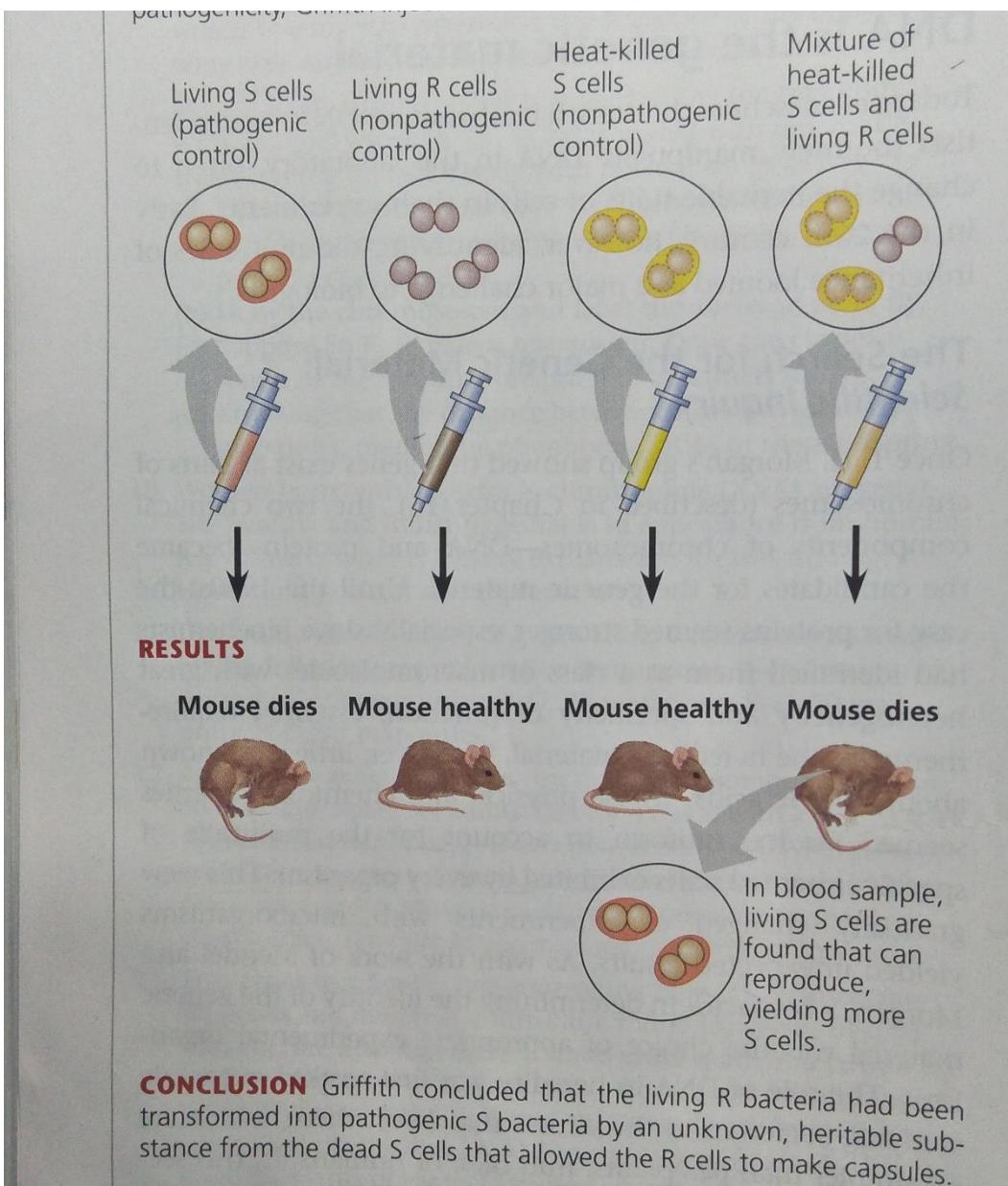
To test the trait for pathogenicity, Griffith took mice and injected these two strains as following procedure and he observed

1. Living S cells (pathogenic)> injected mice>pneumonia>mice die
2. Living R cells (non-pathogenic)> injected mice>no pneumonia> mice survive
3. Heat killed S cells (non-pathogenic)> injected mice>mice survived
4. Mixture of heat-killed S cells and live R cells> injected mice
 - What will be the outcome of the fourth experiment???
 - Mice died

Griffith was also isolating the bacteria from the mice after injection. He found in the fourth experiment that while he injected dead S bacteria in the mixture with R live bacteria, to his surprise he also found S bacteria. These S bacteria could further reproduce to S bacteria and they were pathogenic in nature.

Conclusion he made: SOME CHEMICAL COMPONENT OF THE DEAD PATHOGENIC CELLS (S) MADE A HERITABLE CHANGE TO R BACTERIA. Although he did not find the component.

Griffith called this phenomenon as TRANSFORMATION. (now a commonly used term when DNA is transferred to a living cell to change the genotype and phenotype of the cells).



Oswald Avery Effort to identify the Transforming substance

It took about 14 years for American bacteriologist (who study bacteria) to find the component. He focussed on three main candidates as transforming substances

- DNA
- RNA (another component well known at that time present in the cell)
- Proteins

He took heat-killed S bacteria (pathogenic), broken it up, and now extracted its cellular components. Meaning there are ways by which you could extract DNA fraction, RNA fraction, and Protein fraction of the cell.

He used enzymes which inactivate these three types of fractions as follows:

- RNA + RNase [this enzyme cleaves RNA molecules in small units so that it will not work as a biological molecule]
- DNA + DNase [this enzyme cleaves DNA molecule into nucleotides, so that it will not work for its function]
- Proteases [enzymes which digest proteins, so that they will be non-functional]

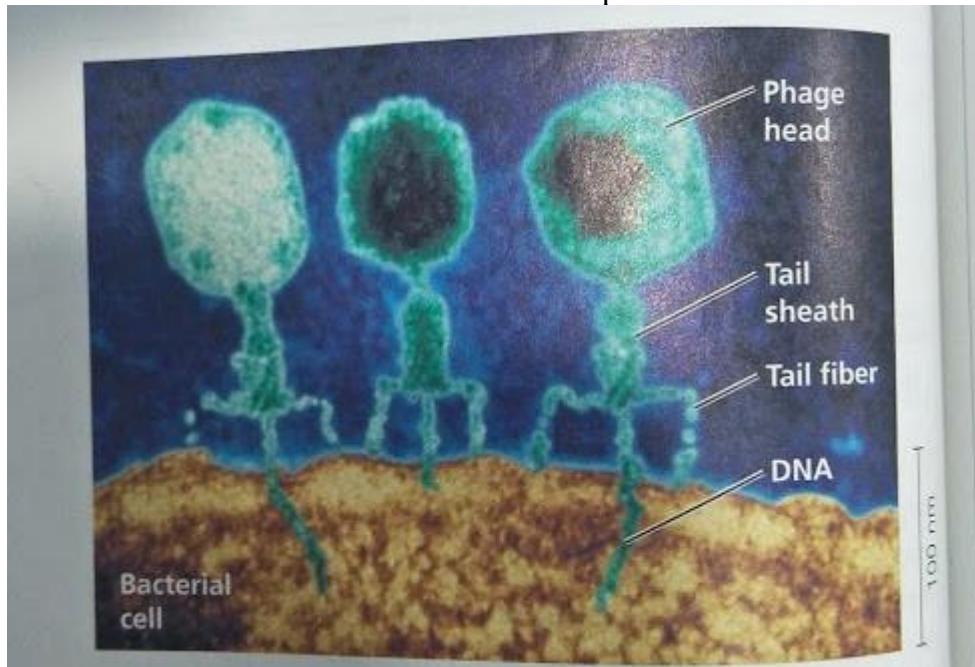
He did a series of experiments and combinations of them. What could be those combinations???? Think

And then each combination was mixed with R (non-pathogenic bacteria>mice injected> What was the outcome for transforming agent. It was only DNA molecule from S bacteria which could transform R Bacteria into Pathogenic S type bacteria, proving that DNA is the material who can transform and can be called as heritable material.

- However, still, people were not convinced for many years as very little was known about DNA and they were still believing protein as the genetic material.

Alfred Hershey and Martha chase Experiment: to prove DNA is the genetic material (1952)

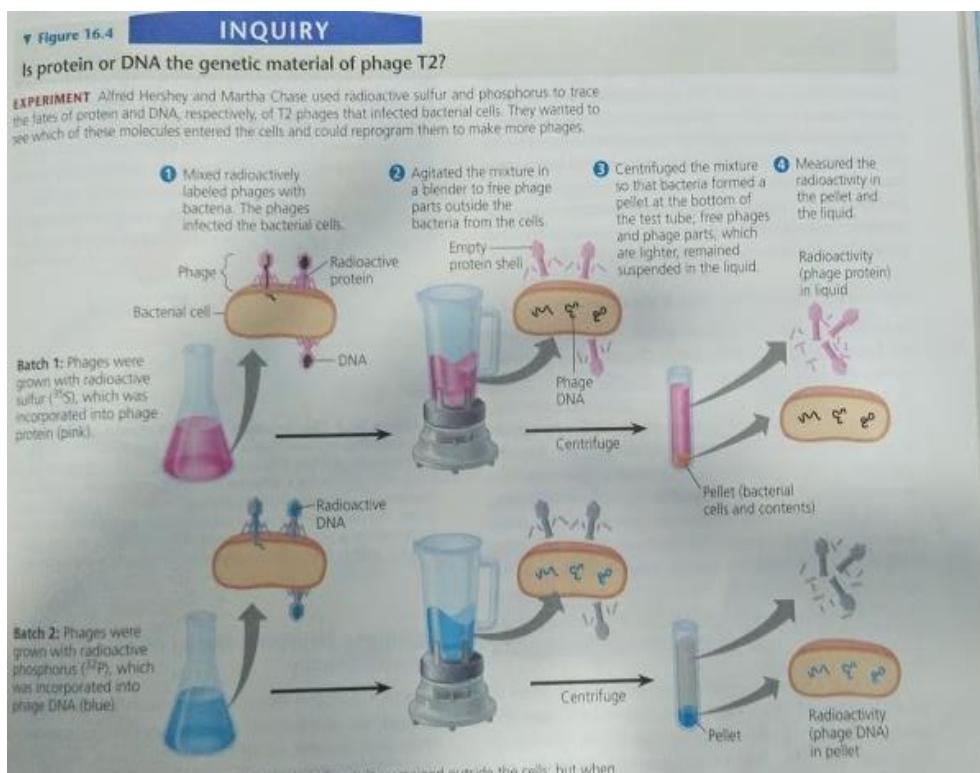
Phages have been widely used as tools by researchers in molecular biology. They took T2 (name of the phage): this is one of the phages which infects *Escherichia coli* (E. coli), which normally lives in the intestine and model organisms used by molecular biologists. It was known that T2 phage is made up of DNA and proteins. They also knew that E. coli after rupture can produce a lot of T2 phages. It was not known what was responsible for this?? DNA or protein??



- Phages have been widely used as tools by researchers in molecular biology. They took T2 (name of the phage): this is one of the phages which infects *Escherichia coli* (E. coli), which normally lives in the intestine and model organisms used by molecular biologists. It was known that T2 phage is made up of DNA and proteins. They also knew that E. coli after rupture can produce a lot of T2 phages
- It was not known what was responsible for this?? DNA or protein?? They devised an experiment to prove that only one of the components enters the E. coli cell as follows:
 - They used a radioactive isotope of sulfur to tag protein in one batch of T2 phage (Sulfur is present in amino acids, please find which amino acid has sulfur??). in another set of T2, A radioactive isotope of phosphorous to tag DNA was used. Phosphorous is only part of DNA and not protein and Sulfur is not part of DNA.
 - In one separate experiment T2 labelled with S* was used to infect E. coli
 - In another set of experiment T2 labelled with P* was used to infect E.co

Now they tested which part of the T2 has gone inside the bacteria? And therefore, wanted to know the ability to program its own synthesis in E. coli. They after infection with T2 of both kinds separately, they blended the mixture of phage and E. coli to free phage parts outside the bacteria.

- They centrifuged the mixture so that bacteria settle at the bottom of the test tube. It was expected that lighter parts of phage remain suspended in the supernatant.
- Now they measured the radioactivity in supernatant and pellet fractions.
- They found radioactivity of Sulfur representing proteins of phage in supernatant fraction and no activity in the pellet fraction. Suggesting protein part of T2 is not present with the Bacteria.
- In the T2 with P* infection, they found radioactivity was found in the pellet only and not in the supernatant
- This experiment confirmed that DNA has gone inside E. coli and not protein and thus acts as the genetic material which transformed information produces bacteriophages by utilizing Host machinery.

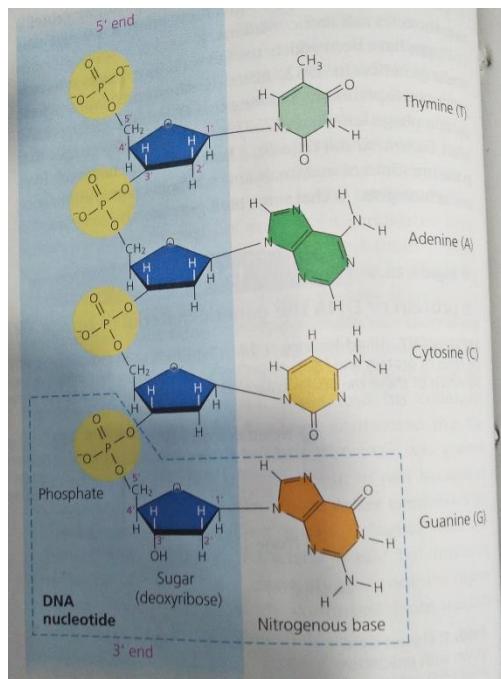


Lecture Note 6: Molecule of Life

In the previous lecture, I talked about the experiments which led to the discovery of DNA as the genetic material. There was another very important discovery by Chargaff's. In 1950 Chargaff's reported the base composition of DNA that varies from one species to another, for example, ~30% of human DNA is composed of Adenine, while E. Coli DNA is composed of ~25% Adenine. He also found that in the DNA of each species, he studied, The number of Adenines equals Thymine and Cytosine equals Uracil. For example, In Human DNA, A = T (both ~30%), G = C (both ~20%). Although Chargaff did not explain the meaning of these rules, but they became a part major part of the discovery of the Double helix by Watson and Crick. Read below.

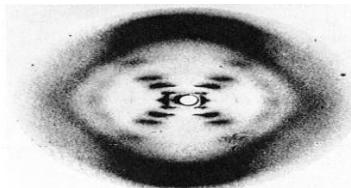
Now an obvious question among the scientist was that how the DNA looks like structurally, to explain its function.

By 1950, there was at least full awareness about the covalent bonding pattern of a nucleotide (Phosphate attachment to the sugar molecule and their attachment to the base (please see below figure). They were also aware of two types of purines (adenine, guanine) and two types of pyrimidine (thymine, cytosine) bases in it. Please watch each covalent bond carefully. One must read the book to grasp these bonds. Look at bases is attached to the sugar at C1' (^ prime), and phosphate is attached at both 3', 5' ends of the sugar and both sugars are attached by a phosphodiester bond. This bond is formed when two of the hydroxyl group of phosphoric acid reacts with hydroxyl groups on other molecules, here deoxyribose sugar. These make up the backbone of the DNA molecule.



Watson and Crick were able to build the structural model of DNA molecule on the basis of following data taken from other scientists.

They used the X-Ray diffraction data of DNA collected by Franklin. (X-ray diffraction is a technique in which one can study in the structural biology course of BSBE or in material sciences). The diffraction pattern tells you about the structural arrangement of molecule. By looking at such data, they could tell that the diffraction pattern belongs to a helical molecule, which is made up of two strands. They could also guess the diameter of the helix to about 2 nm and the distance between to bases of about 0.34 nm. They started building their model with sticks and balls.

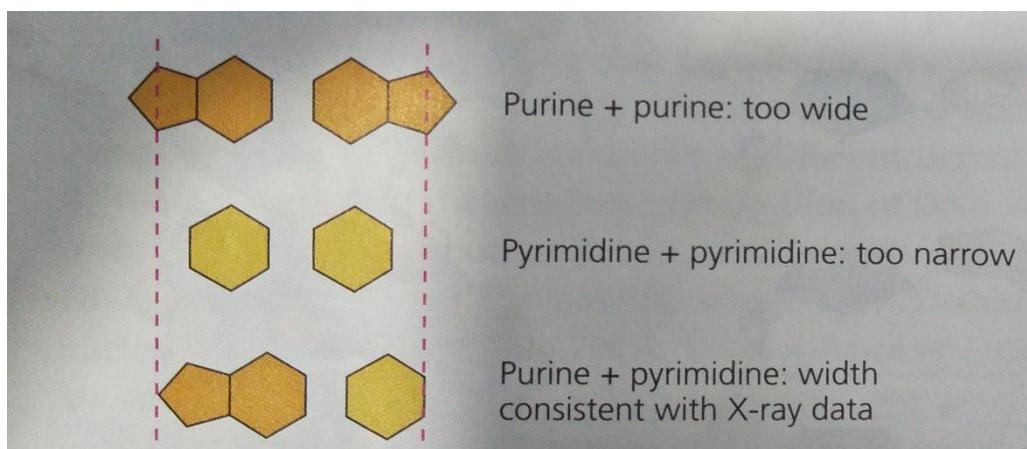


(a) Rosalind Franklin (b) Franklin's X-ray diffraction

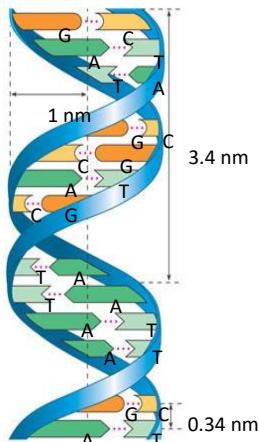
Photograph of DNA

Initially, they were putting in their structure, phosphate backbone inside the double helix, and bases outside. This was quite an unstable situation for the double-strand to hold. They again took a look in Franklin explanation where she had explained that the phosphate backbone should be outside and bases should inside. Watson and Crick realized it and found a very good explanation for it. That bases are hydrophobic, if they are put inside, they will prefer interior and away from water and also stabilized by stacking interactions. While phosphate being negatively charged will repel each other if put adjacent to each other and prefer water molecules to get stability if present away from each other. They did and corrected it.

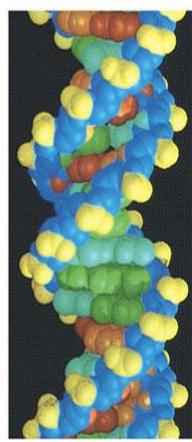
Now they put bases and realized Chargaff's rule, they also gave rationale of putting the bases in front of each other of two strands as the following figure. You can see that purine and pyrimidine pairs made a sense to fit in the width of the two-strand obtained from X-ray diffraction data.



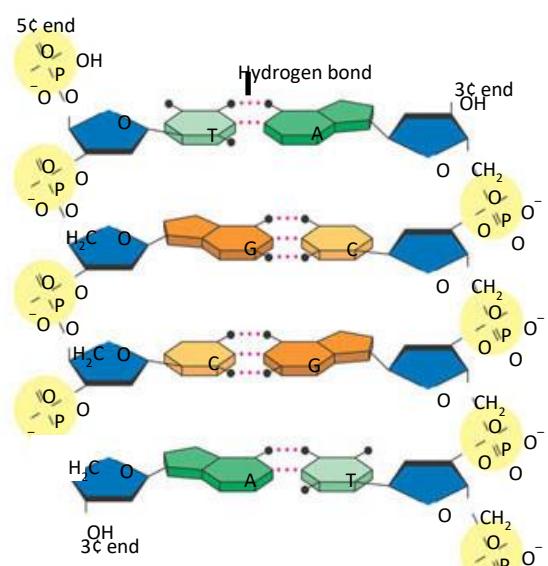
They also found that a possible hydrogen-bonding pattern will provide stability to two strands. AT will pair with two H bonds, and GC will pair with three H Bonds. They also found stability to two strands if they run in an antiparallel manner that parallel manner. Please note the ends of each strand by focusing on 5' ends of sugar and 3' end of sugar at the terminus. If in one stand it is 5' to 3', in another stand will be from 3' to 5'. This will become very important in the future as when the information of DNA will be duplicated and transmitted. (Replication and transcription). The final structure looks like:



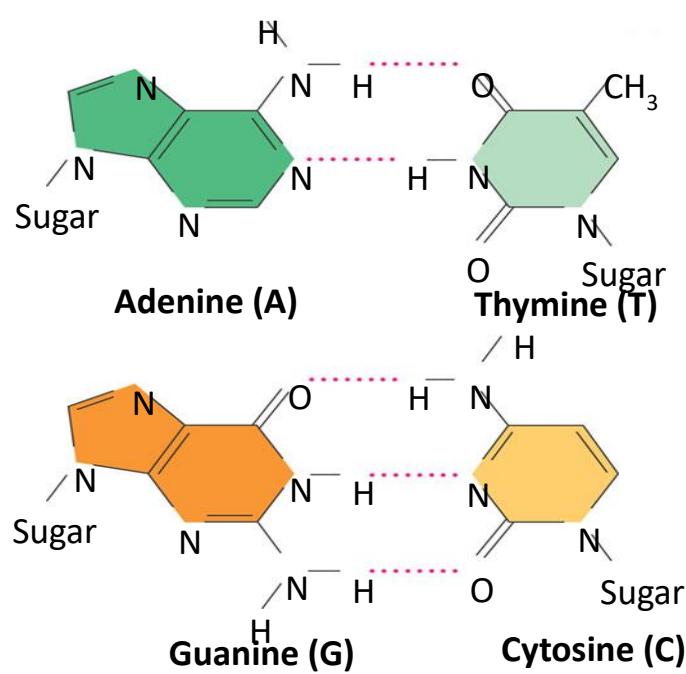
(a) Key features of DNA structure



(c) Space-filling model



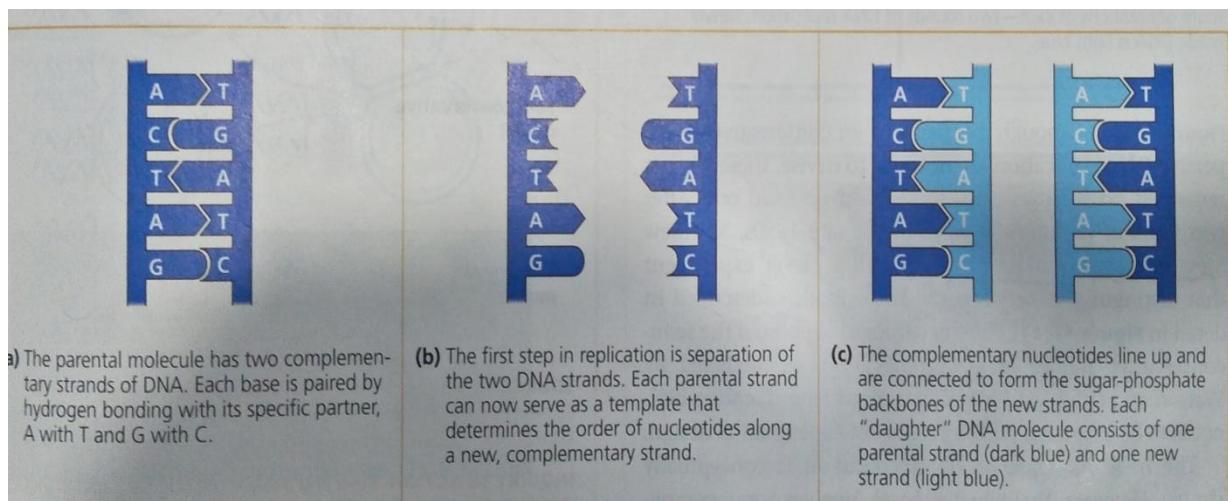
(b) Partial chemical structure



Lecture Note 7: Molecule of Life

In the previous lecture, we discussed DNA structure discovery by Watson and Crick. This discovery led towards better rational thinking about its function. The specific base pairing suggested that the linear sequence of nucleotides can be varied in countless ways and any length depending on a particular species. Moreover, the beauty of the DNA double-helical structure suggested its way of replication and this was recognized by Watson and Crick. They wrote, a statement in their research paper “ It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism.”

They also imagined that prior to duplication, the hydrogen bonds are broken and two chains unwind and separate. Now each chain makes a template for the formation onto itself of a new chain so that two pairs of DNA molecules can be created. A model for DNA replication proposed by Watson and Crick is represented below.

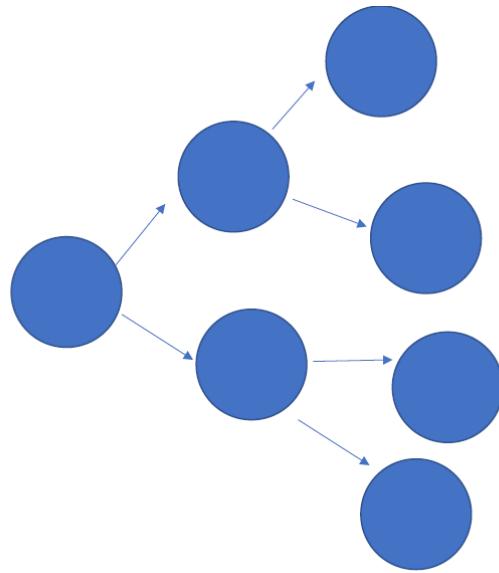
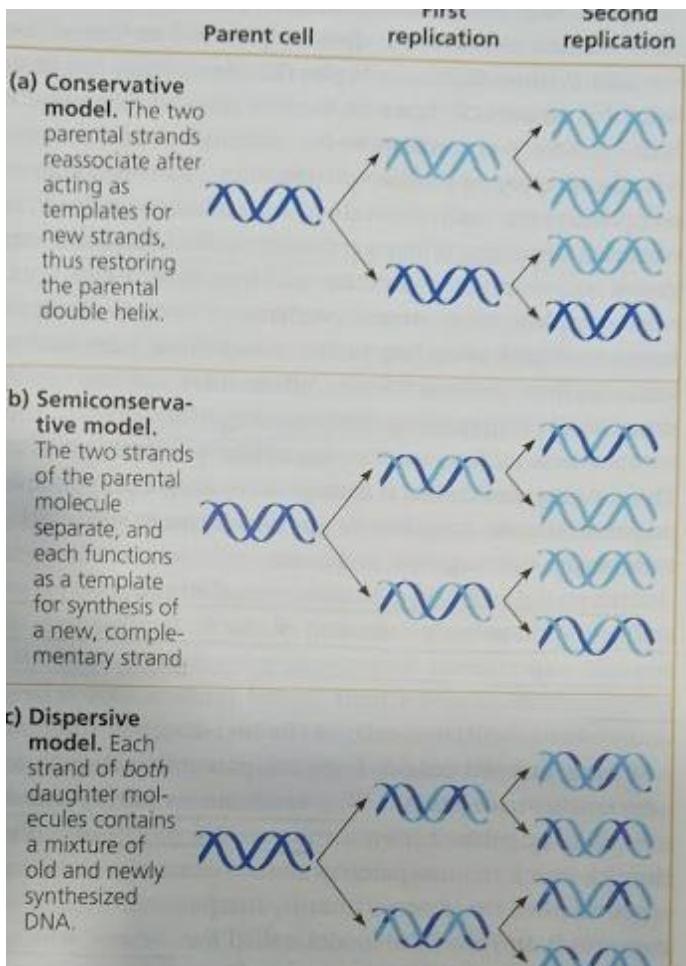


The question is why we are discussing this point. Recall that in the cell, DNA molecule is present in each chromosome. Cell division is the property of a cell which starts from the embryo stage itself. When cell divide, each cell should have same DNA molecule for functioning of the cell. Thus, the parent cell should duplicate the DNA and each copy will distribute to two cells, within the chromosome. That will happen only when DNA duplicates first as the molecule of life which is then followed by cell division.

Other than the model given by Watson and Crick for replication, which is known as semiconservative mode of replication. This means that the daughter DNA molecule will have one parent DNA strand with it as shown in the previous figure and the subsequent figure.

Conservative model: In this model DNA strand after duplication will reassociate in such a way that the parent strands will rewind with each other and daughter strands will reassociate with each other.

Dispersive model: That all four stands of DNA will have some old part and some new parts. Please see the next for these three models.



The proof of semiconservative replication model came from Meselson and Stahl in 1958. They worked on *E. coli* cells, whose life cycle and cell division were known, meaning in how many hrs one will expect a cell to divide, therefore suggesting that DNA will replicate. They fed *E. coli* with a medium (food of bacteria) in which nucleotide precursors were labelled with a heavy isotope of Nitrogen ^{15}N . This means in biomolecules including DNA during the synthesis inside bacteria, ^{15}N instead of ^{14}N nitrogen will be present in the nitrogenous bases. So, DNA will be labelled all over with ^{15}N . Please see DNA nitrogenous base structure where N is present (normally N should be ^{14}N). After that, they transferred the bacteria to a medium containing ^{14}N labelled precursor. They took bacterial cells as per their duplication cycle to extract DNA which replicated once. And then replicated again. They extracted DNA from such bacteria and centrifuged the DNA molecules to separate on the basis of Densities. In a way, they will have the following bacteria in a given medium in their experiments.

- 1) *E. coli* completely was grown in ^{15}N medium for many generations. Then bacteria were collected by centrifugation and DNA was isolated from them. It should contain ^{15}N labelled DNA.

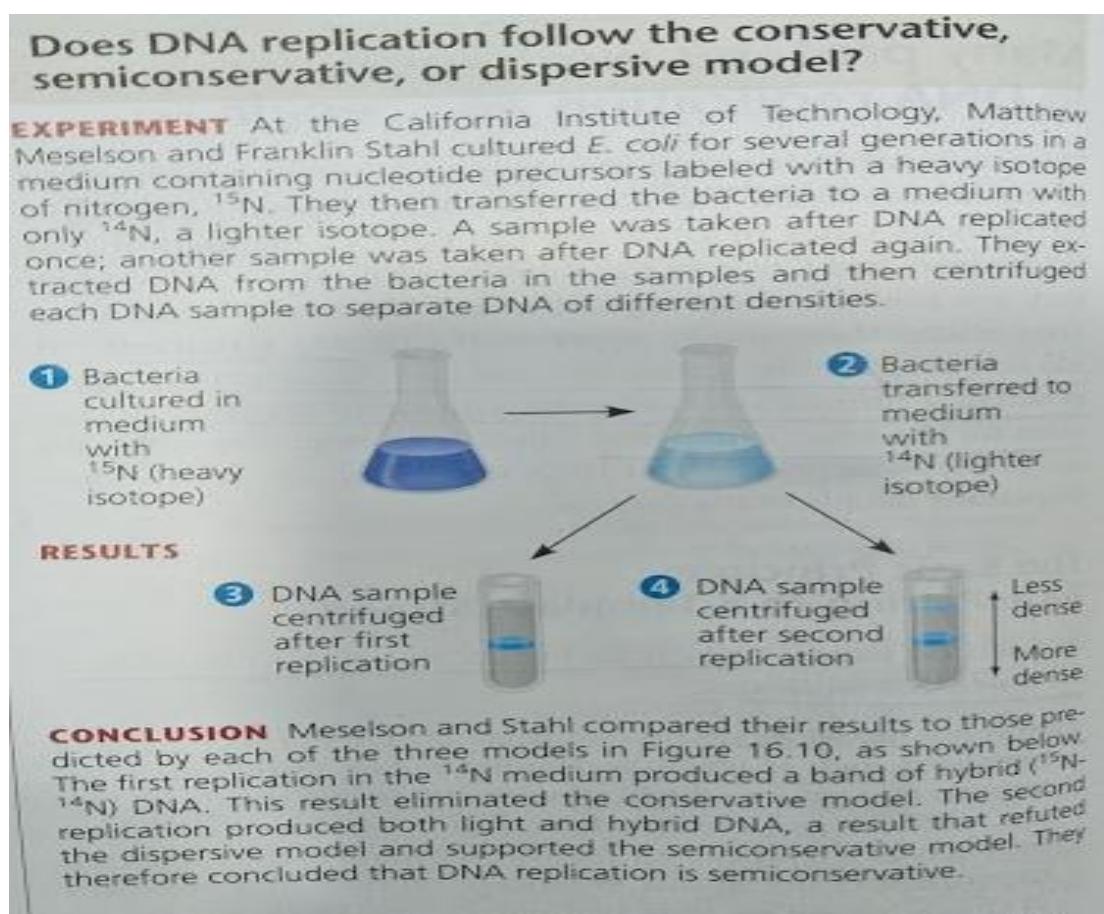
- 2) *E. coli* completely was grown in ^{14}N medium for many generations. Then Bacteria were collected by centrifugation and DNA was isolated from them. It should contain ^{14}N labelled DNA.

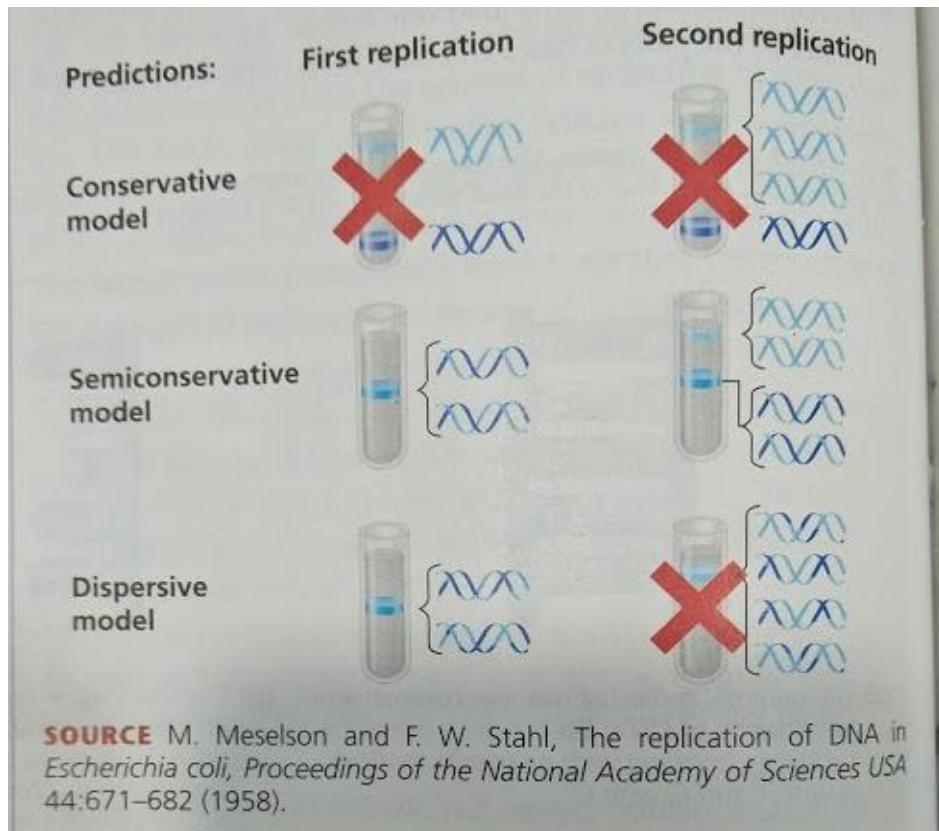
- 3) When the *E. coli* from the first condition bacteria was transferred to the ^{14}N medium and the bacteria was allowed to duplicate once. This means DNA would have duplicated one time. Meaning the parent cell first have seen ^{15}N medium and Now ^{14}N . Remember the synthesis of daughter strand on the parent DNA strand will add nitrogenous bases which were now ^{14}N labelled. The DNA extracted from these *E. coli* Cells would have original DNA (^{15}N) and daughter DNA which was (^{14}N) labelled.

- 4) When in the third condition, they let some *E. coli* grow and divide further for the second round in ^{14}N medium. Meaning one cell became two (scenario 3) and now they would become four. And the DNA of them was isolated. The expectation would be; two cells will have ^{15}N and ^{14}N DNA, while two cells will have ^{14}N and ^{14}N DNA strands.

For each extracted DNA from each condition, they centrifuged the sample on the basis of their densities and compare their height in the tube. The idea is dense DNA will move down more (^{15}N ^{15}N strands), as compared to the lighter isotope (^{14}N ^{14}N strands). And if there is a hybrid of two (^{15}N ^{14}N , it's position will be in between the two).

Now observe the figure below for Meselson and Stahl experiment. Also, find out yourself how the bands for the above four conditions will look like???



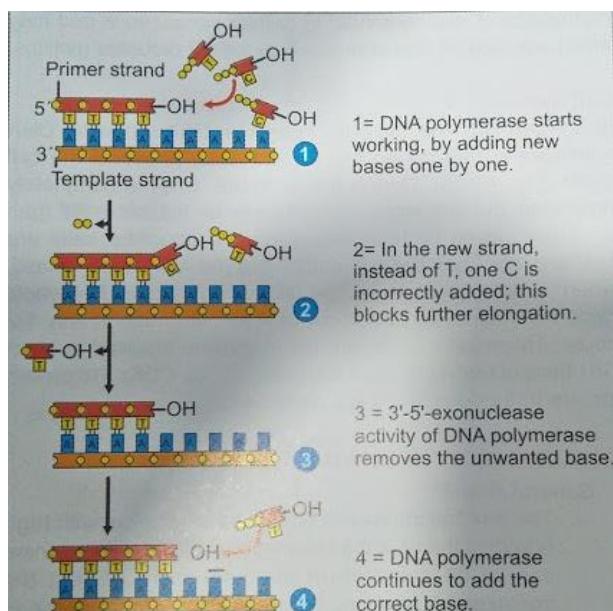


These experiments proved, semiconservative mode of replication and not dispersive or conservative mode of replication.

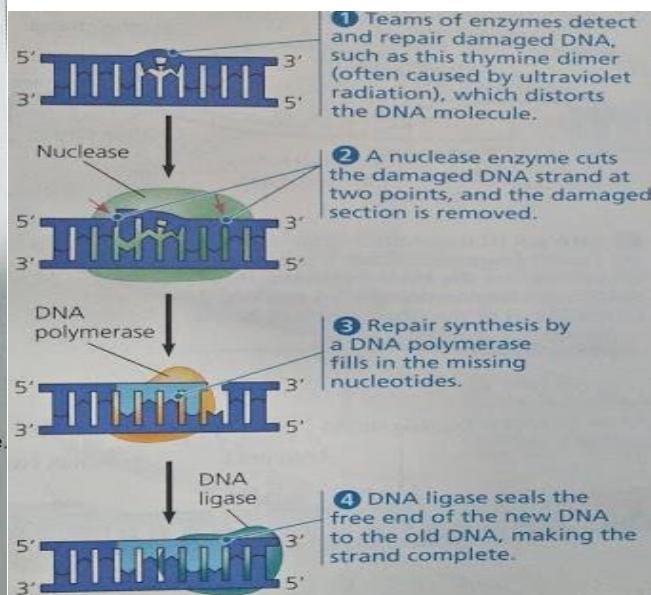
In the next note, we will discuss about the replication of the process inside the cell.

Lecture Note 9: Molecule of Life: Proof reading and repairing of DNA

Accuracy of DNA replication cannot be attributed to only complementary bases pairing. Although the errors in the completed DNA molecule is 1 in 10 billion nucleotides. Initial nucleotide pairing errors between template strand and daughter strand can be much larger. Meaning lets say the template strand has G, daughter stand should have C. However, in place of C due to errors, edition of other than C nucleotide can also be possible. Such errors could be 1 in 10^5 nucleotides. These errors are taken care of by DNA polymerase due to its proofreading activity. Upon finding an incorrectly base-paired it removes it and the right base is added. This activity is just like typing a wrong letter and deleting and correcting it.



Proof Reading of DNA Polymerase



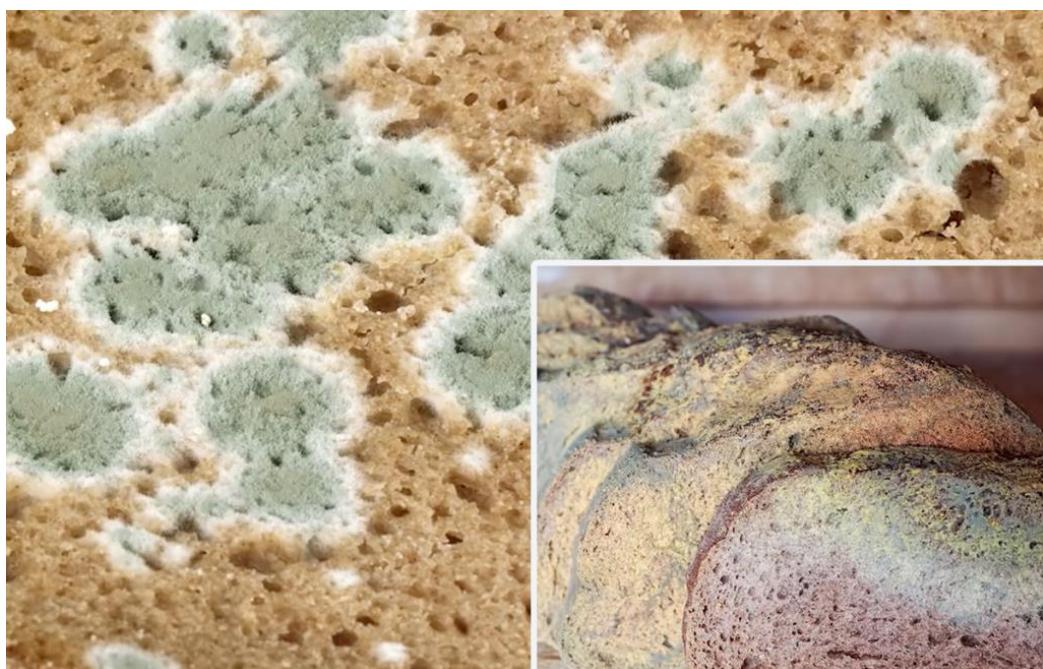
Nucleotide Excision Repair

Mismatched nucleotide sometimes even evade proofreading by DNA polymerase. In these cases, repair enzymes come into play and fix this error. Such errors such as altered nucleotides can also happen even after replication. In fact, such damages of DNA keeps on happening through chemicals and physical agents, such as cigarette smoke, Xray, etc. Cell continuously monitors DNA for such damages and try to fix them. Around 100-130 such repair enzymes are present in E.coli and Human cells.

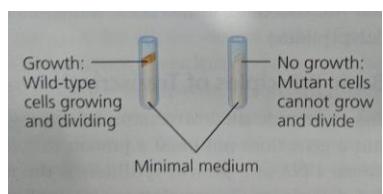
Faithfull replication of the DNA and repair of DNA damage are important for the functioning of the organism and for passing on completely accurate DNA to the next generation. Although the error rate significantly lowered, but if it slips through, there is a permanent change in the DNA, which is known as Mutation. If it happens in Germ cells, this passes to the next generation and effects a phenotypic character. Some mutations may be harmful, Some can be beneficial. In either case, the mutations are the source of variation on which natural selection operates during evaluation and ultimately responsible for the appearance of a new species. Over a long period of time, this balance between DNA replication, repair, and low mutations rate has over a long period of time allowed the evolution of the rich diversity of species on earth today.

Lecture Note 10: Molecule of Life: Flow of Information

The DNA inherited by an organism leads to a specific trait by dictating the synthesis of proteins and RNA molecules involved in protein synthesis. In other words, proteins are the link between gene (genotype) and a character/trait (phenotype). Gene expression is the process by which DNA directs the synthesis of proteins. It involves two stages: Transcription, Translation. Before going into details of these two processes, Let's find out how the fundamental relationship between genes and proteins was established? In 1902, British physician, Archibald Garrod was the first to suggest that genes dictate a phenotype through enzymes that catalyze specific reactions in the cell. Garrod was studying a disease known as Alkaptonuria. In this disease, urine is black as it contains a chemical alkapton, which darkens in the presence of air. Garrod speculated the most people will have enzymes to metabolize alkapton, whereas people with this disorder have inherited an inability to make this enzyme. He gave the name "inborn errors of metabolism for such conditions". Garrod was ahead of his time to speculate that Mendel's law may hold true in Humans. Research several decades after Garrod proved that genes dictate the production of enzymes A breakthrough in demonstrating the relationship between genes and enzymes came in 1920. 1920, Beadle and Tatum, worked on *Neurospora crassa* (a bread mold).

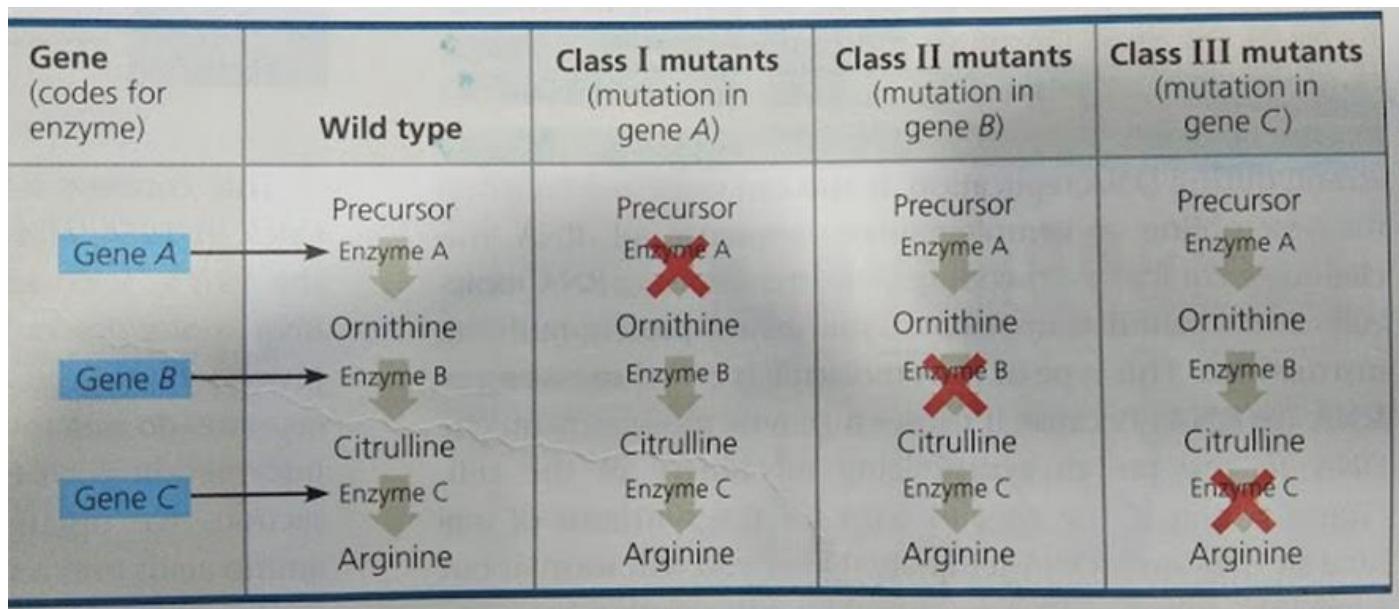


They bombarded X-rays on this mold. (It was known that X-ray can bring genetic changes). Remember that such changes can affect an organism. In this case, the effect of such changes was seen on the nutritional needs of *Neurospora crassa*. It was different than the wild type *Neurospora crassa* (wild type meaning which was not given X-rays). Remember that wild type mold has a modest nutritional requirement. It can grow on a simple solution of inorganic salts, glucose, vitamins, and a support medium such as agar. This is known as a **minimal medium**. The cells then utilize these nutritional components to synthesize essential, molecules in the cell, such as amino acids. However, the mutants (*Neurospora crassa* which was created by X-ray, cannot grow on the minimal medium). But they could grow when supplemented with 20 amino acids. This is known as a **complete medium**. This happened as the mutants were defective in making these amino acids in their body, but giving them in the medium supported their growth. One can say that Xray created metabolic defects in them. To characterize the metabolic defect, Beadle and Tatum, took samples of mutants grown on the complete medium and distributed them in a number of tubes. Each will contain minimal medium plus a single additional nutrient, such as an amino acid arginine. So suppose if the addition of an amino acid arginine brings back complete growth of the mold, one can conclude that the mold was mutated in the metabolic pathway of Arginine.

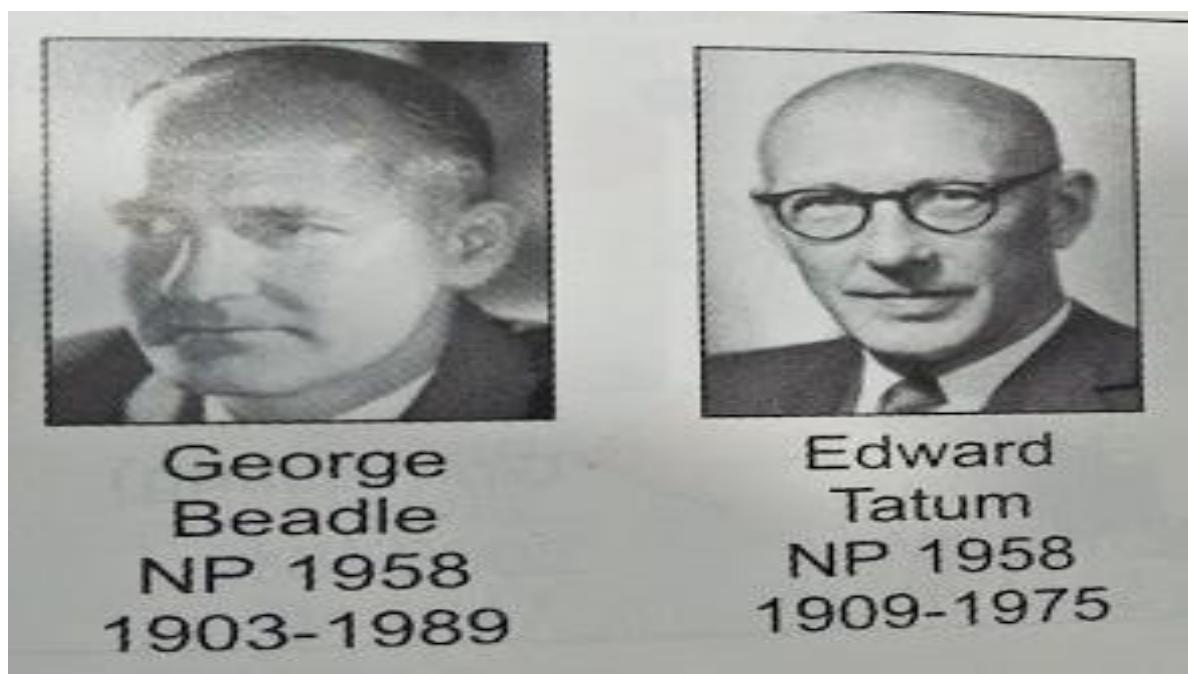


Classes of *Neurospora crassa*

	Wild type	Class I mutants	Class II mutants	Class III mutants
Condition				
Minimal medium (MM) (control)				
MM + ornithine				
MM + citrulline				
MM + arginine (control)				
Summary of results	Can grow with or without any supplements	Can grow on ornithine, citrulline, or arginine	Can grow only on citrulline or arginine	Require arginine to grow



Later on, a detailed analysis of the pathway and purification of enzymes and linkages of each enzyme to its catalytic function was established. This finding led to the support that one gene can be responsible for one enzyme synthesis. One gene-one enzyme hypothesis. Meaning that one gene will dictate the synthesis of one enzyme.



As researchers learned more about proteins, they made revisions to the one gene–one enzyme hypothesis. First of all, not all proteins are enzymes. Keratin, the structural protein of animal hair, and the hormone insulin are two examples of nonenzyme proteins. Because proteins that are not enzymes are nevertheless gene products, molecular biologists began to think in terms of one gene–one protein. However, many proteins are constructed from two or more different polypeptide chains, and each polypeptide is specified by its own gene. For example, hemoglobin, the oxygen-transporting protein of vertebrate red blood cells, contains two kinds of polypeptides, and thus two genes code for this protein (see Figure 5.20). Beadle and Tatum's idea was therefore restated as the *one gene–one polypeptide hypothesis*. Even this description is not entirely accurate, though. First, many eukaryotic genes can each code for a set of closely related polypeptides via a process called alternative splicing, which you will learn about later in this chapter. Second, quite a few genes code for RNA molecules that have important functions in cells.

- One gene on protein, One gene only polypeptide

Lecture Note 11: Molecule of Life: Basic Principles of Transcription and Translation

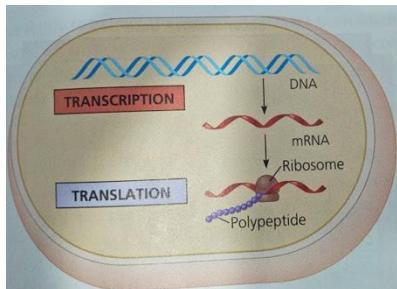
Genes provide instructions for making a specific protein. Gene does not build a protein directly. The bridge between the two is the RNA molecule, which is chemically similar to the DNA molecule, except that it contains ribose sugar instead of deoxyribose sugar and nitrogenous base uracil rather than thymine. Thus, DNA will have A, G, C or T, while RNA will have A, G, C, and U bases. RNA molecule usually is single-stranded. DNA and proteins are polymers that contain information written in two different chemical languages, nucleotides and amino acids. Two major stages are required getting from DNA to protein: Transcription and Translation.

Transcription: It is the synthesis of RNA using the information in the DNA. The two nucleic acids are written in different forms of the same language. The information is transcribed or rewritten from DNA to RNA. Just as the DNA strand provides a template for making a new complementary strand during DNA replication. It also serves as a template for assembling a complementary sequence of RNA nucleotides. This type of RNA is known as mRNA which takes the information to the protein-making machinery to synthesize protein.

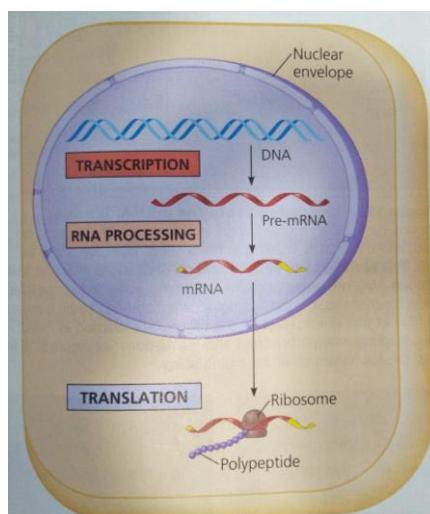
Translation: It is the synthesis of a polypeptide using the information in mRNA. During this stage, there is a change in the language. The cell must translate the nucleotide sequence of mRNA molecule into the amino acid sequence of a polypeptide. The sites of translation are ribosomes, that facilitate the linkage of amino acids into a polypeptide chain.

Both processes occur in all organisms meaning those which lacks membrane-bound nucleus (bacteria and archaea) and those have membrane-bound organelles. Although the basic mechanisms of transcription and translation are similar for bacteria and eukaryotes, there is an important difference in the flow of genetic information within the cells.

In a bacterial cell, since there is no membrane around the nucleus, mRNA made from DNA is immediately translated into a polypeptide.

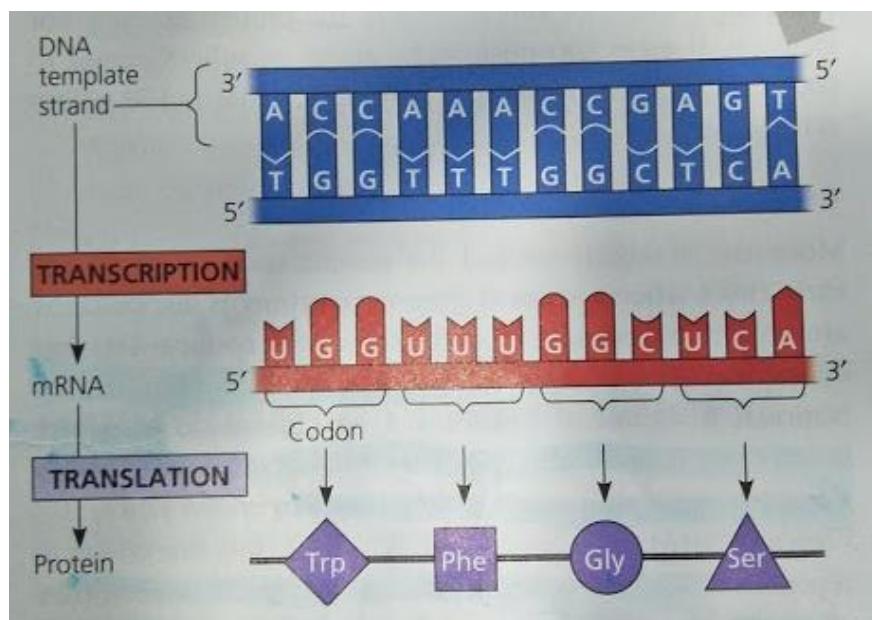


Due to compartmentalization, pre-mRNA is made and processed to mRNA, which leaves the through the nuclear membrane into the cytoplasm. Then it is translated into a polypeptide chain



Lecture Note 12: Molecule of Life: Genetic Code

Biologists recognize a problem that how four nucleotide bases in DNA will specify 20 amino acids in a protein. How many bases correspond to an amino acid? Each nucleotide translates to one amino acid, only four amino acid will be translated. If the code is two-letter, e. g AG, $4^2 = 16$ amino acids will be used. If the code is three letters, then 64 amino acids. During transcription, the gene determines the sequence of bases along the length of an mRNA molecule. This tripled code is known as genetic code.



1. Three nucleotides encode an amino acid. Proteins are built from a basic set of 20 amino acids, but there are only four bases. Simple calculations show that a minimum of three bases is required to encode at least 20 amino acids. Genetic experiments showed that *an amino acid is in fact encoded by a group of three bases, or codon.*

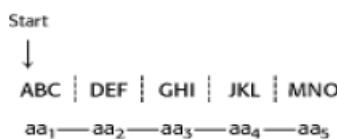
2. *The code is nonoverlapping.* Consider a base sequence ABCDEF. In an overlapping code, ABC specifies the first amino acid, BCD the next, CDE the next, and so on. In a nonoverlapping code, ABC designates the first amino acid, DEF the second, and so forth. Genetics experiments again established the code to be nonoverlapping.



3. The code has no punctuation. In principle, one base (denoted as Q) might serve as a "comma" between groups of three bases.

... QABCQDEFQGHIQJKLQ ...

This is not the case. Rather, the sequence of bases is read sequentially from a fixed starting point, without punctuation.



4. The genetic code is degenerate. Some amino acids are encoded by more than one codon, inasmuch as there are 64 possible base triplets and only 20 amino acids. In fact, 61 of the 64 possible triplets specify particular amino acids and 3 triplets (called stop codons) designate the termination of translation. Thus, *for most amino acids, there is more than one code word.*

		Second mRNA base					
		U	C	A	G		
U	UUU	Phe	UCU	UAU	UGU	Cys	
	UUC		UCC	UAC	UGC		
	UUU	Leu	UCA	UAA	Stop	UGA	Stop
	UUG		UCG	UAG	Stop	UGG	Trp
C	CUU		CCU	CAU	CGU		
	CUC	Leu	CCC	CAC	CGC		
	CUA		CCA	CAA	CGA	Arg	
	CUG		CCG	CAG	CGG		
A	AUU		ACU	AAU	AGU		
	AUC	Ile	ACC	AAC	AGC	Ser	
	AUA		ACA	AAA	AGA		
	AUG	Met or start	ACG	AAG	AGG	Arg	
G	GUU		GCU	GAU	GGU		
	GUC		GCC	GAC	GGC		
	GUA	Val	GCA	GAA	GGA	Gly	
	GUG		GCG	GAG	GGG		

The genetic code is nearly universal. Shared by organisms from the simplest bacteria to the most complex animals. No ambiguity and Redundancy/degenerate.

The stunning variety of living systems (Figure 1.7) belies a striking similarity. The common use of DNA and the genetic code by all organisms underlies one of the most powerful discoveries of the past century—namely, that *organisms are remarkably uniform at the molecular level*. All organisms are built from similar molecular components distinguishable by relatively minor variations. *This uniformity reveals that all organisms on Earth have arisen from a common ancestor.* A core of essential biochemical processes, common to all organisms, appeared early in the evolution of life. The diversity of life in the modern world has been generated by evolutionary processes acting on these core processes through millions or even billions of years. As we will see repeatedly, the generation of diversity has very often resulted from the adaptation of existing biochemical components to new roles rather than the development of fundamentally new biochemical technology. The striking uniformity of life at the molecular level affords the student of biochemistry a particularly clear view into the essence of biological processes that applies to all organisms from human beings to the simplest microorganisms.

What is the biological significance of the extensive degeneracy of the genetic code? If the code were not degenerate, 20 codons would designate amino acids and 44 would lead to chain termination. The probability of mutating to chain termination would therefore be much higher with a nondegenerate code. Chain-termination mutations usually lead to inactive proteins, whereas substitutions of one amino acid for another are usually rather harmless. Thus, *degeneracy minimizes the deleterious effects of mutations*. Degeneracy of the code may also be significant in permitting DNA base composition to vary over a wide range without altering the amino acid sequence of the proteins encoded by the DNA. The G + C content of bacterial DNA ranges from less than 30% to more than 70%. DNA molecules with quite different G + C contents could encode the same proteins if different synonyms of the genetic code were consistently used.

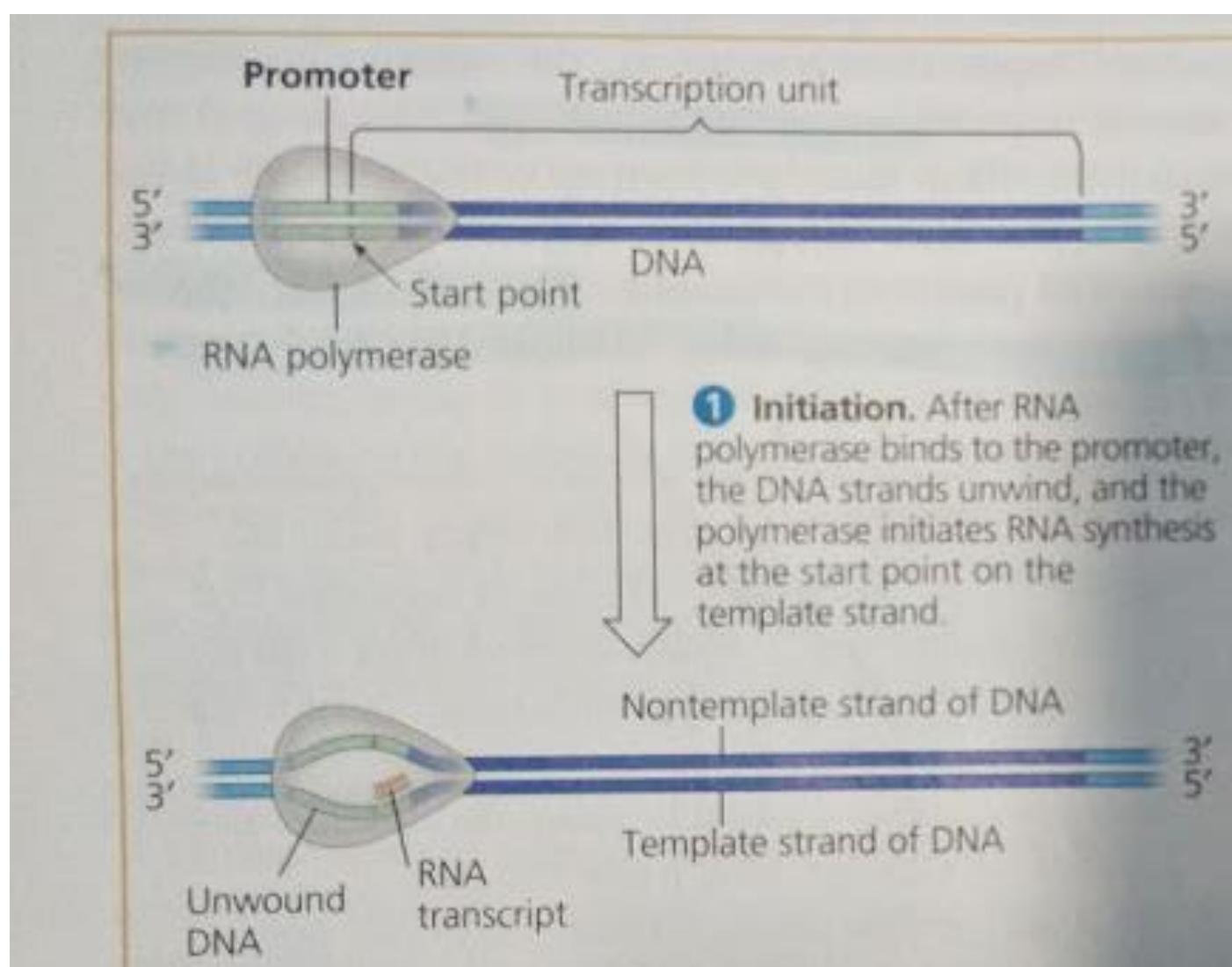


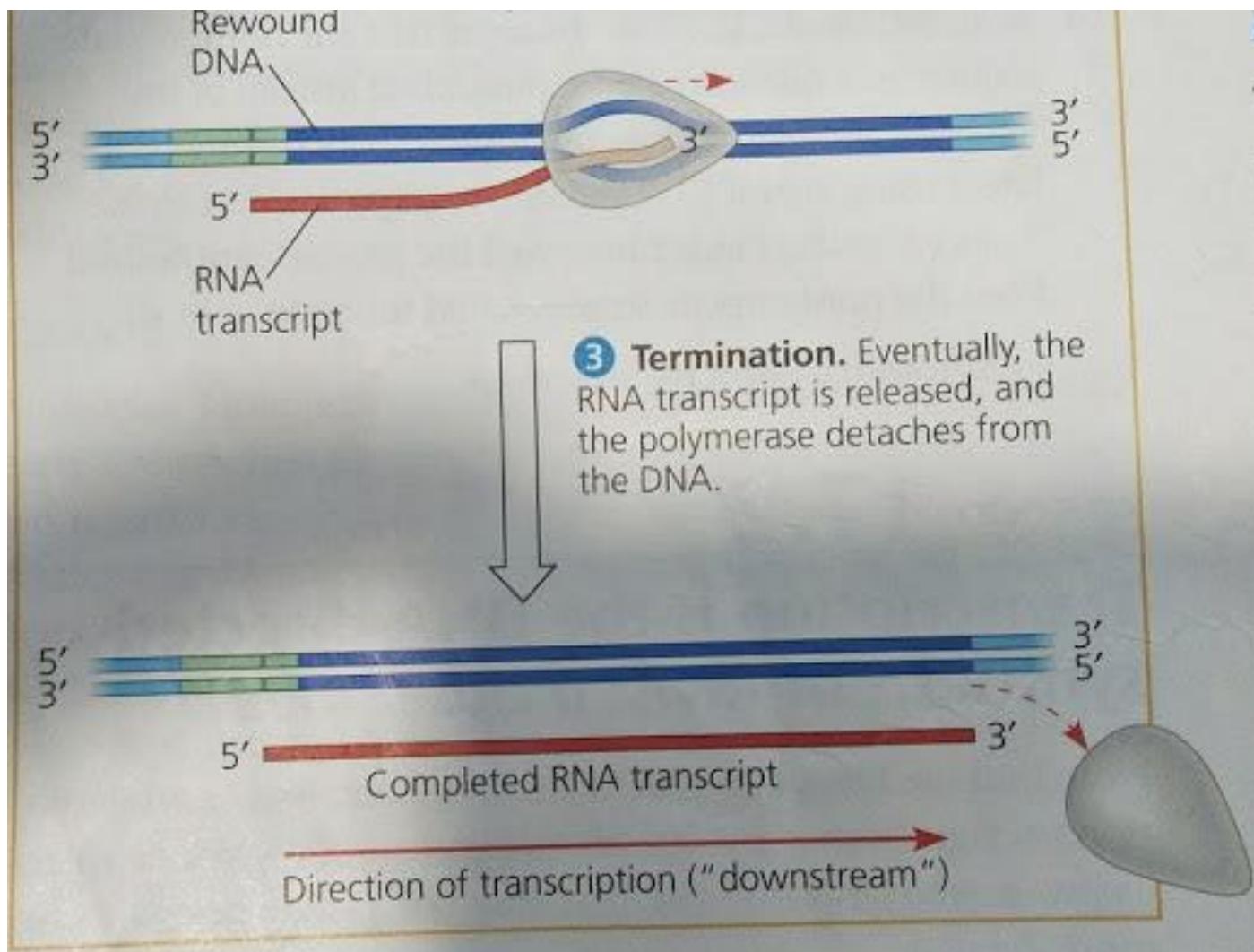
Lecture Note 13: Molecule of Life: Translation

Molecular components of Transcription: DNA, RNA polymerase enzyme: This enzyme separates the DNA strand, Joins together RNA nucleotides complementary to the DNA template. As the DNA polymerase, RNA polymerase is able to start a chain from scratch and therefore do not need a primer. But like DNA polymerase, RNA polymerase also works from 5' end to 3' end. The DNA sequence where RNA polymerase binds is known as Promoter and where it ends is known as Terminator. The direction of transcription is referred to as downstream and the other direction is known as upstream. For example, the promoter sequence can be said as upstream of the terminator sequence. The stretch of DNA that is transcribed into RNA is known as TRANSCRIPTIONAL Unit. In bacteria, RNA polymerase alone can bind to the promoter while in eukaryotes, there are other protein factors that plays role in the binding of RNA polymerase to the promoter. These together are known as the transcription initiation complex. The promoter sequence is known as the TATA box.

Transcription process: After RNA polymerase binds, the DNA strands unwind and the polymerase initiates the RNA synthesis at the start point on the template strand. The polymerase moves downwards, unwinding the DNA and elongating the RNA transcript 5'>>3 prime direction. The DNA strand reform a double helix.

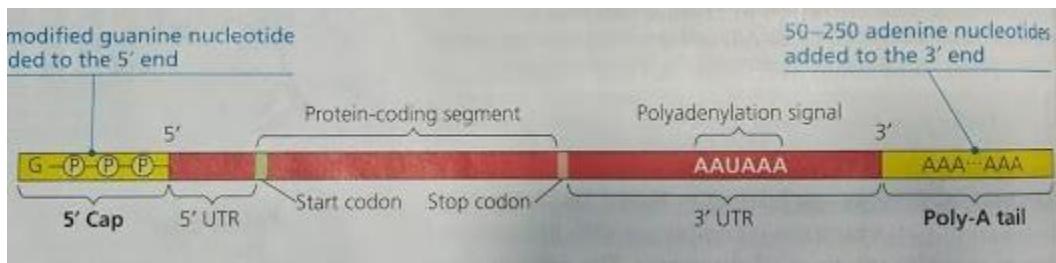
At the termination The RNA transcript then is released, and polymerase detaches from the DNA.



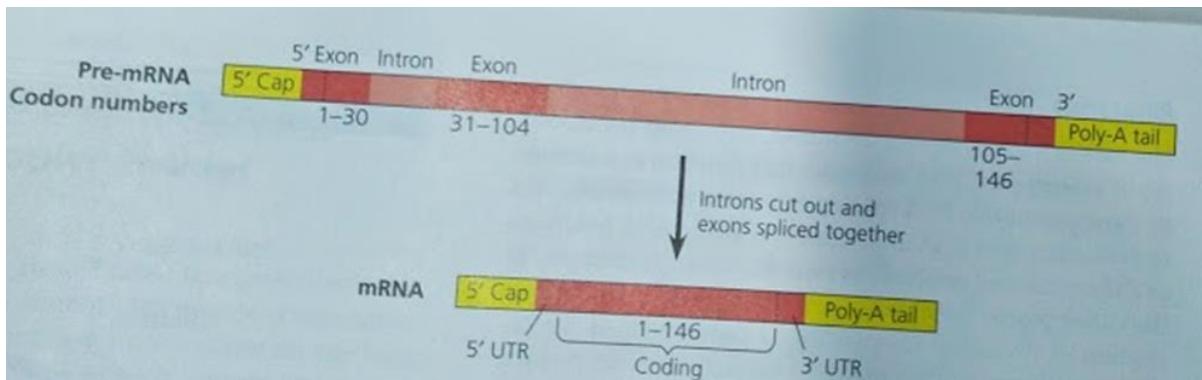


Lecture Note 14: Molecule of Life: Modification of RNA after transcription in Eukaryotes

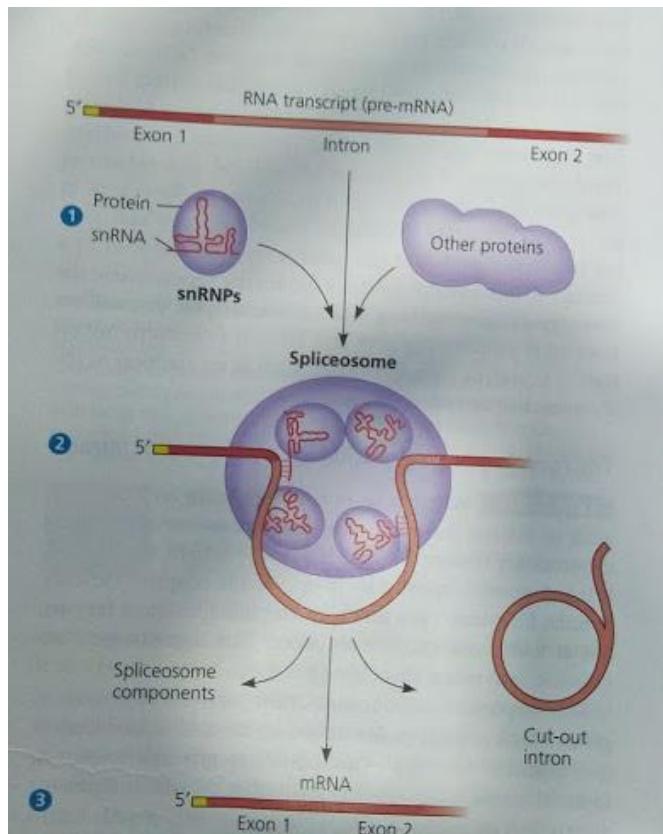
Eukaryotic cell modify its mRNA transcript (pre-mRNA) before it is translocated in the cytoplasm. First, a 5' cap of a modified Guanine is added. This is achieved after 20-40 nucleotides are added while during the elongation of RNA on the DNA template. Towards the 3' end a long polyA tail is added, after 3' untranslated region. The function of these modifications are 1) to help mature mRNA in transportation from the nucleus to the cytoplasm. 2) These modifications provide stability from the hydrolytic enzymes and 3) they help in translation, as the binding of 5' site to the ribosomes. Also, there are untranslated regions present in the mRNA. These regions do not get translated into the part of the protein, but helps in the recognition of ribosomes during translation.



Another important modification which happens in the pre-mRNA is the removal of some parts of it, by a process known as RNA splicing. This is similar to editing a video. They can be classified as coding and noncoding regions or exons and introns. Only exons/coding regions are part of the mature mRNA and make a polypeptide during translation. Introns are spliced out.



This is known as RNA splicing. This process is carried with the help of a spliceosome, which recognizes a small sequence present on the ends of the introns. snRNPs (small nuclear ribonucleoproteins) recognize these parts. They have protein and small nuclear RNA (snRNA). It is found that snRNA catalyses the removal of introns and joining of exons. These are also known as Ribozymes, RNA as an enzyme.

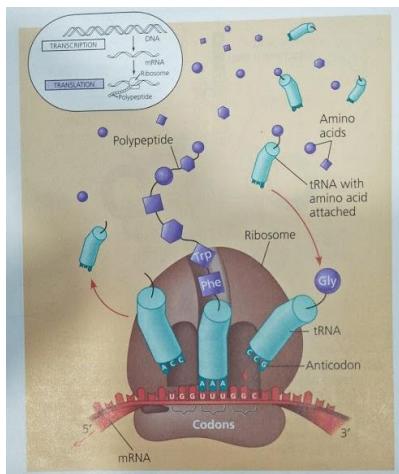


The functional importance of introns: Although the functions of many introns are a very active area of research. Some facts which are known about its importance are highlighted here.

Some of them participate in gene expression regulation. Another importance is that a gene can encode more than one polypeptides, by combining exons in different combinations during the splicing process. This is also known as exon shuffling. This helped scientists to explain that the number of proteins is more than the number of genes present in humans.

Lecture Note 15: Translation

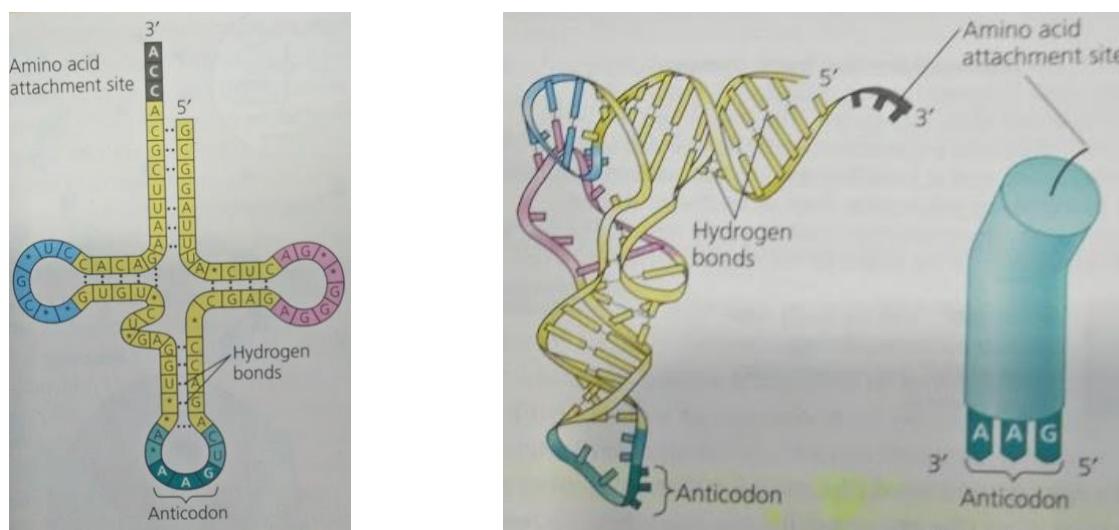
Basic concept of Translation: Flow of genetic information from mRNA to the synthesis of protein. In this process, a cell reads the genetic message, in the form of codons present on the mRNA. tRNA (transfer RNA) translates the message. It brings the amino acid from the pool of amino acid present in the cytoplasm to the ribosome where these amino acids are put in a polypeptide chain. Remember, the cell keeps all the amino acids and tRNA is in the cytoplasm. Ribosomes which are made up of proteins and RNA molecules, adds each amino acid brought by tRNA. Process is simple in understanding, but complex in molecular mechanism. We will study the generic view of the process, in this course.



Lecture Note 15: t-RNA structure and function

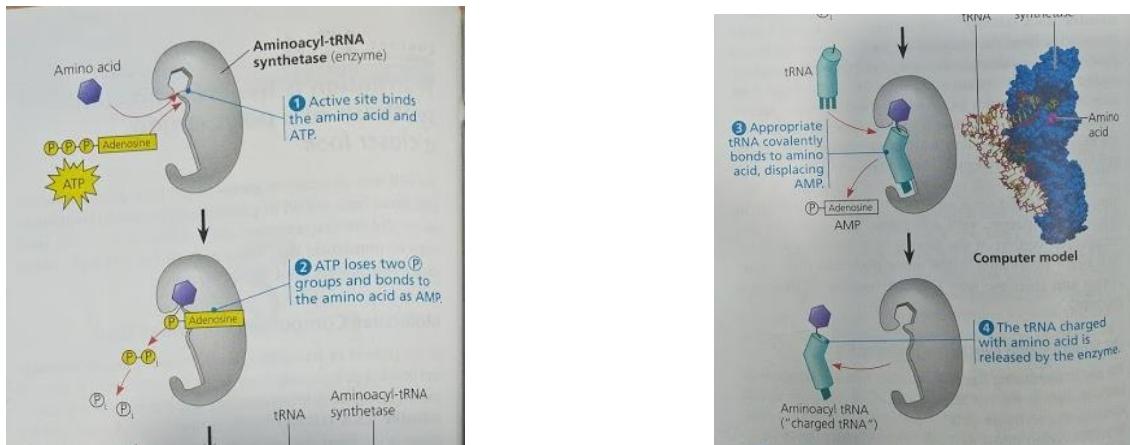
The structure and function of tRNA: tRNA molecule contains two important parts. One part which has complementary code (anticodon), reads mRNA codon. On the other end, it carries an amino acid. For example, consider, there is a codon on mRNA which has codon GGC, which is translated to glycine. The tRNA that base pairs with this codon will have anticodon CCG and carry glycine on it. And when this anticodon-codon will recognize each other, glycine will be added by the ribosome to the polypeptide. tRNA, therefore, works as a translator, it reads the code on mRNA with the help of a complementary code and interprets it as a protein word, in the form of amino acid. It means each tRNA molecule will have a different codon and the capability to bring a different amino acid. The question is from where the tRNA comes from. As mRNA is made from DNA, tRNA and rRNA (ribosomal RNA) is also made from the DNA template and present in the cytoplasm. Each tRNA molecule is used multiple times in the cell, picking up an amino acid and delivering it to the ribosome and again doing the same, repeatedly.

Structurally see below, the tRNA molecule is made up of a single RNA molecule, about 80 nucleotides long. It has complementary bases, which can pair with each other, and the molecule can fold to form a clover leaf-like structure. Both 2D and 3D version of it are shown below to appreciate both complementary base pairing in the chain as well as different regions of a tRNA molecule. Note an important point that Anticodon is written from 3' to 5' to match complementary codon written from 5' to 3'. For example ${}^5\text{UUC}{}^3$ codon on mRNA will be read by, ${}^3\text{AAG}{}^5$ anticodon on tRNA (see below figure). 3' end of tRNA attaches amino acid.

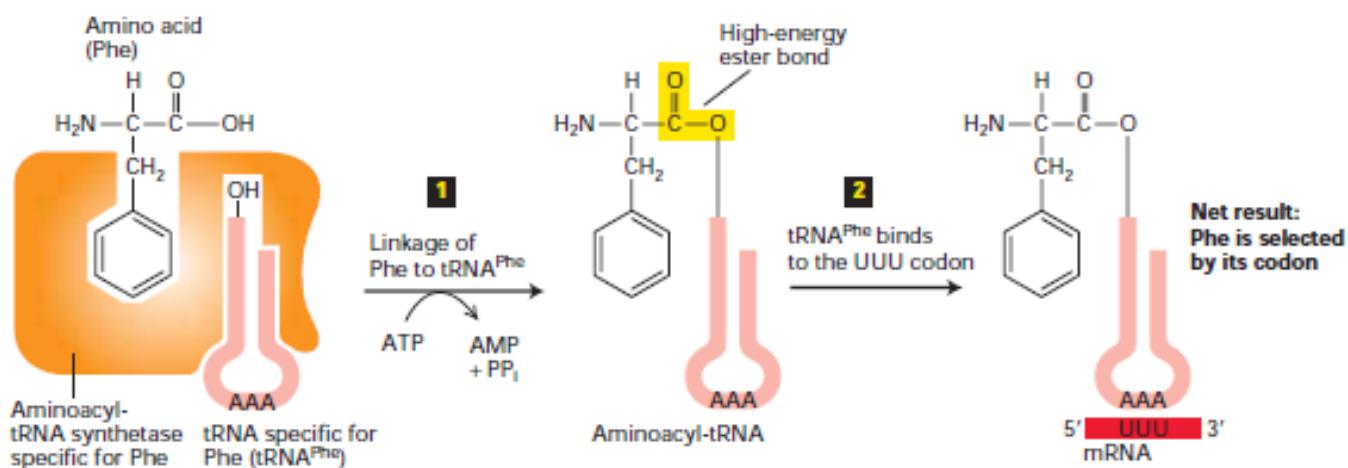


Lecture Note 17: Importance of two molecular events in translation

Molecules of tRNA are not all identical, each carries a specific amino acid on one end. Each has an anticodon on the other end. A specific enzyme called an aminoacyl-tRNA synthetase joins each amino acid to the correct tRNA molecule. There are 20 different synthetases. The active site binds the amino acid and ATP. ATP loses two P groups and joins amino acid as AMP. Appropriate tRNA covalently Bonds to amino Acid, displacing AMP. The activated amino acid is released by the enzyme, Aminoacyl tRNA (an “activated amino acid”).

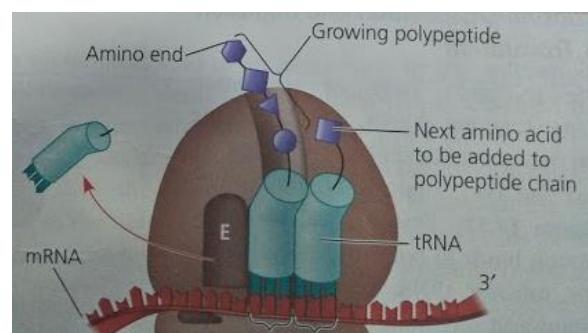
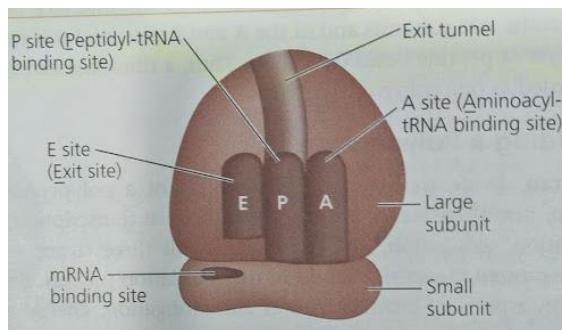


20 different synthetases, one for each amino acid. The pairing of tRNA anticodon with appropriate codon: If one tRNA existed for each mRNA, there will be 61 tRNA molecules. 45 tRNA molecules are present, some tRNA molecules are able to bind to more than one codon. Some rules of base pairing are relaxed for the third nucleotide of a codon and anticodon 3'-UCU-5' anticodon can base pair with mRNA codons 5'-AGT-3' or 5'-AGG-3' [both codes for Arginine] This flexibility is known as a wobble. This explains why synonymous codons vary most often at the third nucleotide base and not others.



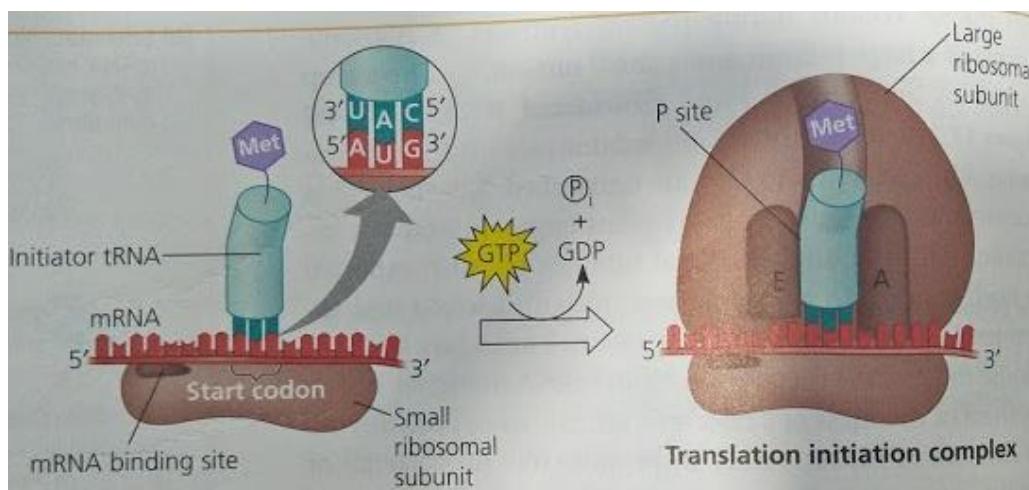
Lecture Note 17: Ribosome structure and function

The ribosomes are made of two subunits: larger and smaller. These units are constructed of proteins and RNA molecules named ribosomal RNA or rRNA. Both eukaryotic and prokaryotic ribosomes are roughly similar. Eukaryotic ribosomes are synthesized in the nucleolus, and the units are transported to the cytoplasm from the nucleus. The ribosome has three tRNA binding sites: The P site, The A site, and the E site, mRNA binding site. The P site (holds the tRNA carrying the growing polypeptide chain, The A site holds the tRNA carrying the next amino acid to be added to the next chain. The E site, discharge the tRNA. The ribosomes hold the tRNA and mRNA in close proximity.

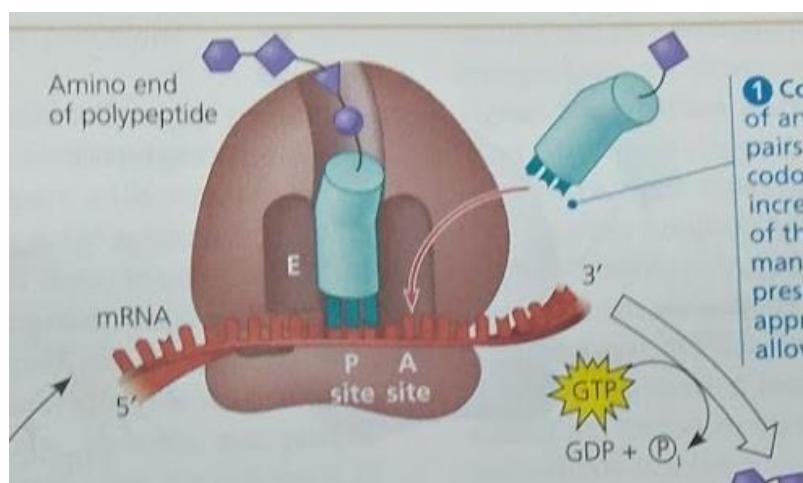


Lecture Note 19: Building a Polypeptide

The initiation stage of translation: Brings together mRNA, tRNA bearing the first amino acid of the polypeptide, and two subunits of a ribosome. A small ribosomal subunit binds to a molecule of mRNA. In a prokaryotic cell, the mRNA binding site on this subunit recognizes a specific nucleotide sequence on the mRNA just upstream of the start codon. An initiator tRNA, with the anticodon UAC, base-pairs with the start codon, AUG. This tRNA carries the amino acid methionine (Met). The arrival of a large ribosomal subunit completes the initiation complex. Proteins called initiation factors are required to bring all the translation components together. GTP provides the energy for the assembly. The initiator tRNA is in the P site; the A site is available to the tRNA bearing the next amino acid.

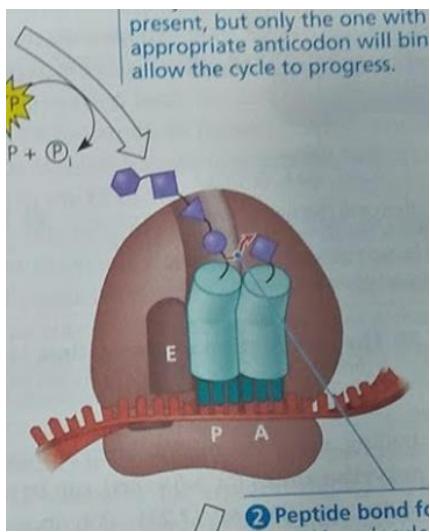


Elongation of Translation: In the elongation stage of translation, Amino acids are added one by one to the preceding amino acid. Codon recognition: The anticodon of an incoming aminoacyl tRNA base-pairs with the complementary mRNA codon in the A site. Hydrolysis of GTP increases the accuracy and efficiency of this step.

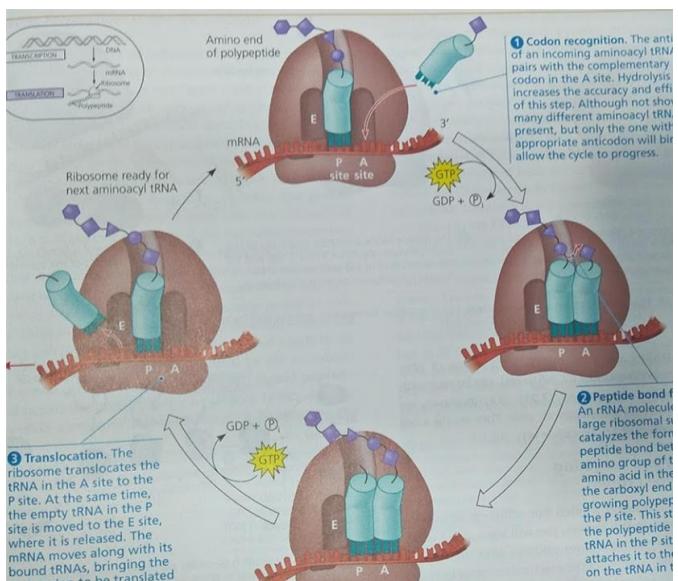


<https://www.ncbi.nlm.nih.gov/books/NBK21653/figure/A910/?report=objectonly>

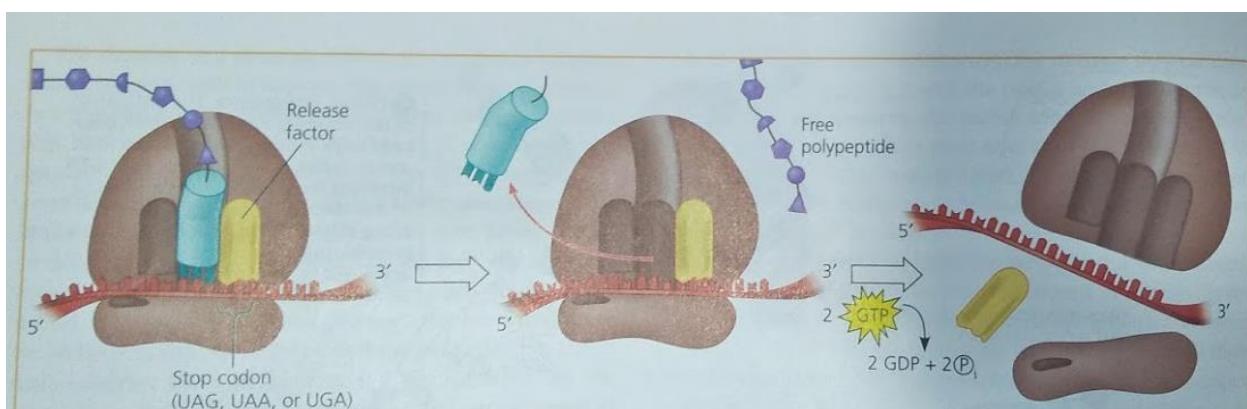
Peptide bond formation. An rRNA molecule of the large subunit catalyzes the formation of a peptide bond between the new amino acid in the A site and the carboxyl end of the growing polypeptide in the P site. This step attaches the polypeptide to the tRNA in the A site.



Translocation. The ribosome translocates the tRNA in the A site to the P site. The empty tRNA in the P site is moved to the E site, where it is released. The mRNA moves along with its bound tRNAs, bringing the next codon to be translated into the A site.

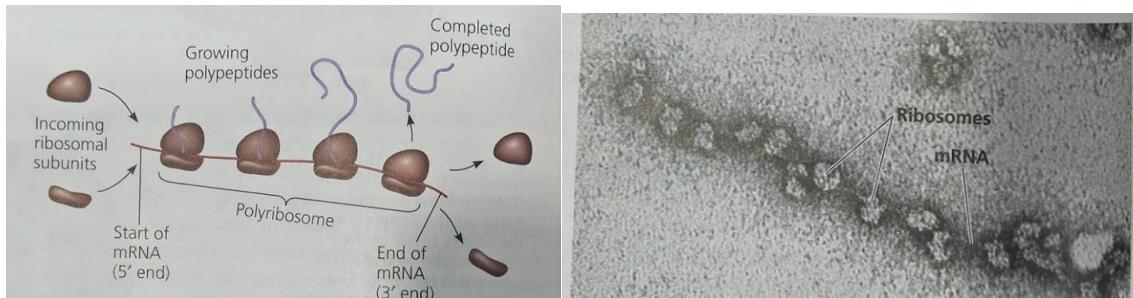


Termination of Translation: When a ribosome reaches a stop codon on mRNA, the A site of the ribosome accepts a protein called a release factor instead of tRNA. The release factor hydrolyzes the bond between the tRNA in the P site and the last amino acid of the polypeptide chain. The polypeptide is thus freed from the ribosome. The two ribosomal subunits and the other components of the assembly dissociate.

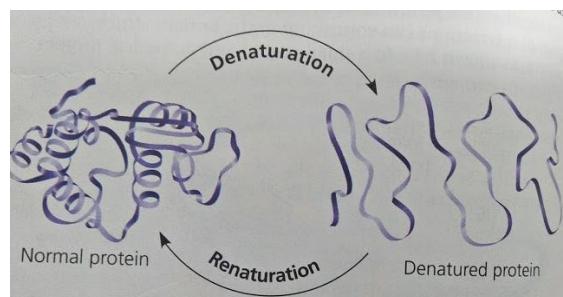


Lecture Note 20: Making a Functional protein

Polyribosomes: To make multiple copies of the polypeptide, a single mRNA can be translated by multiple ribosomes. For example, if the ribosome has left the start codon and crossed enough to give a space to another ribosome to bind to the start codon, the second ribosome can start synthesizing the polypeptide. Multiple ribosomes can bind to mRNA and keep on translating it into a polypeptide. These strings of ribosomes can be visualized with the help of electron microscope. It is present both in prokaryotes and eukaryotes.

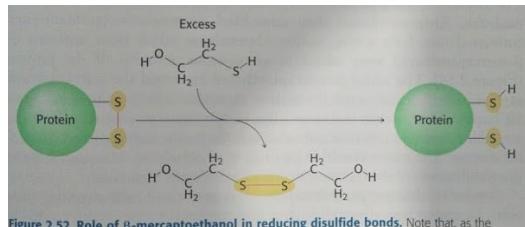
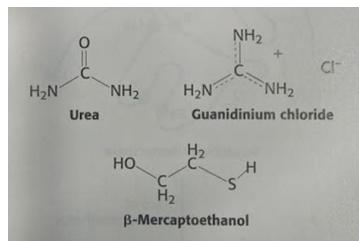
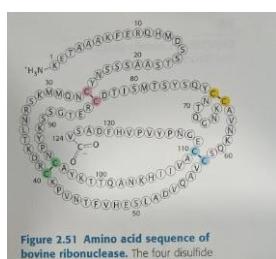


During its synthesis, a polypeptide starts coiling and folding spontaneously as a consequence of its amino acid sequence, the primary structure, forming a protein with secondary and tertiary structure. This leads to a functional protein. Gene determines the primary polypeptide chain sequence; this sequence determines the three-dimensional structure of the protein.



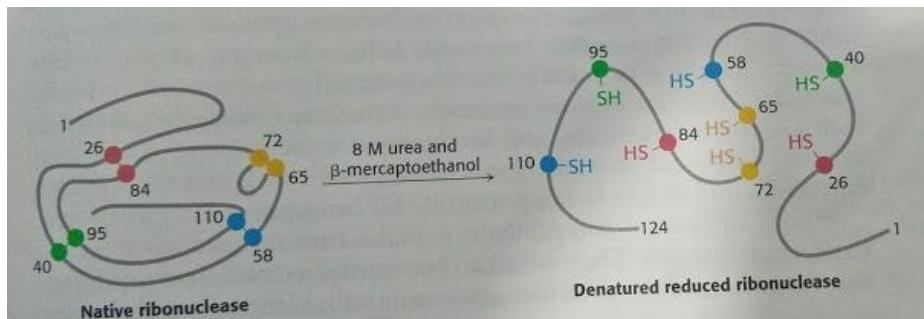
Amino acid sequence of a protein determines its three-dimensional structure.

Anfinsen's experiment: Enzyme ribonuclease, 124 amino acids long, four disulfide bonds. Destroy the structure of ribonuclease enzyme, and restore the structure. How this can be done? Urea, guanidinium hydrochloride disrupt non covalent bonds. They are also known as detergents. Disulfide bonds can be broken with β -mercaptoethanol.



When Ribonuclease was added in urea and Beta ME, Devoid of enzymatic function of ribonuclease, denatured. He freed denatured protein from urea and β -mercaptoethanol by dialysis and found slowly the protein retained its function, suggesting that the protein adopted the structure which gives it desired function. **Primary sequence determines the structure and function.** He got a different result, when the reduced ribonuclease was oxidized in the urea, and then dialyzed to check for the activity. Ribonuclease of this kind only show very little activity. Wrong pairing of disulfides in urea, 105 pairs might be possible. Only one type will be the one responsible for functional activity. Rest will be scrambled disulfide bridges. He further found that the scrambled ribonuclease, retained its activity in trace amount of β -mercaptoethanol in aqueous solution,

suggesting the rearrangement of disulfide pairing until the native structure regained in about 10 Hrs. The process was driven by driven by the decrease in free energy as the scrambled conformations were Converted into the stable, native conformation of the enzyme.





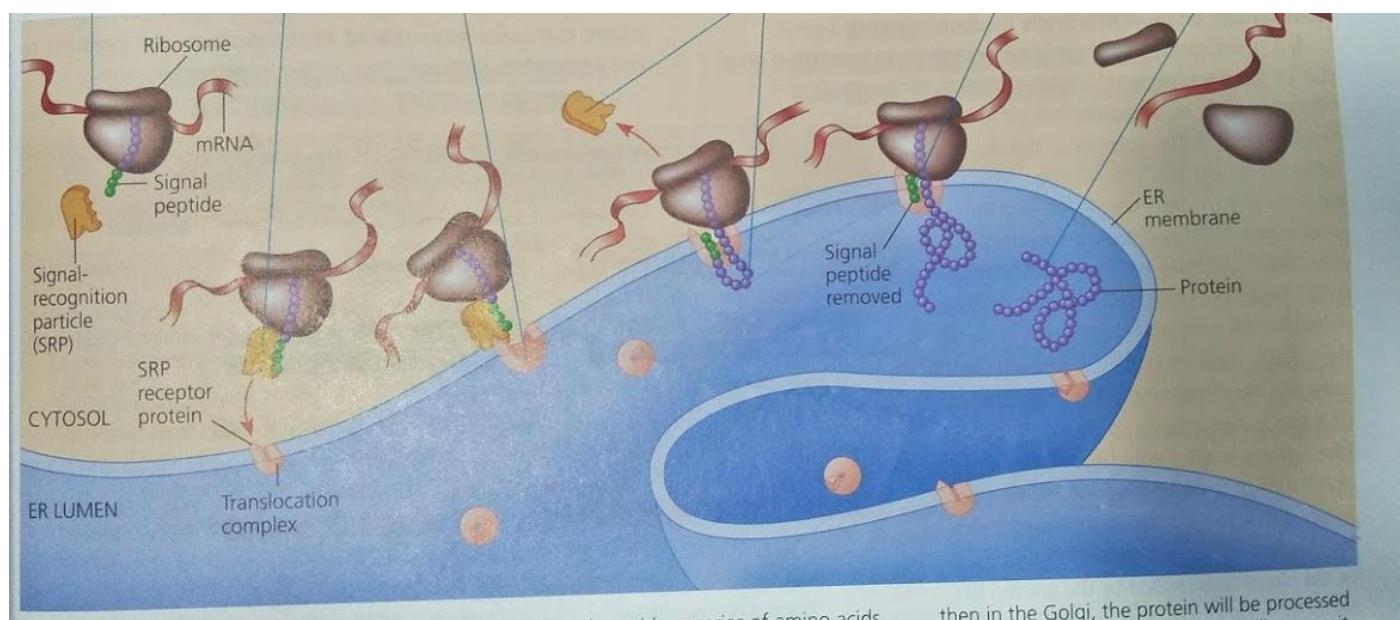
Lecture Note 21: Targeting Polypeptides to Specific Locations

Please recollect that two populations of ribosomes can be seen, one free in the cytoplasm and another population can be found bound to the endoplasmic reticulum. While free ribosomes make proteins, which stays in the cytoplasm, the bound ribosomes make proteins of endomembrane systems. For example, nuclear envelop, ER, Golgi apparatus, lysosomes, vacuoles, and plasma membranes, as well as proteins secreted from the cell, such as insulin. It should be noted that ribosomes are identical and switch their status from free from bound. What determines whether a ribosome is free or going to bind to the ER?

Polypeptide synthesis begins on a free ribosome in the cytosol. An SRP binds to the signal peptide, halting synthesis momentarily. The SRP binds to a receptor protein in the ER membrane. This receptor is part of a protein complex (a translocation complex) that has a membrane pore and a signal-cleaving enzyme.

The SRP leaves, and the polypeptide resumes growing, meanwhile translocating across the membrane. (The signal peptide stays attached to the membrane.). The signal-cleaving enzyme cuts off the signal peptide. The rest of the completed polypeptide leaves the ribosome and folds into its final conformation.

The polypeptide if needs to be secreted, it is released into solution within the ER lumen. If it belongs to the ER, it will remain bound to the ER.



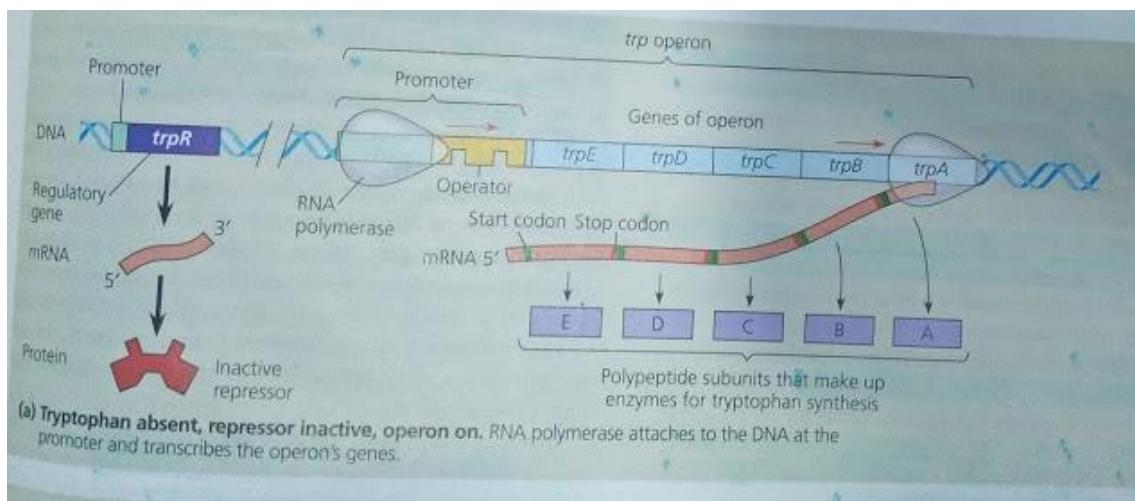
Other kinds of signal peptides are used to target polypeptides to mitochondria, chloroplasts, nucleus, and other organelles, that are not part of endomembrane systems. Translation of them is completed in the cytosol. And then the protein is imported into the organelle. A specific signal peptide works as a postal zip code to deliver the protein.

Lecture Note 22: Regulation of Gene Expression: Tryptophan Operon

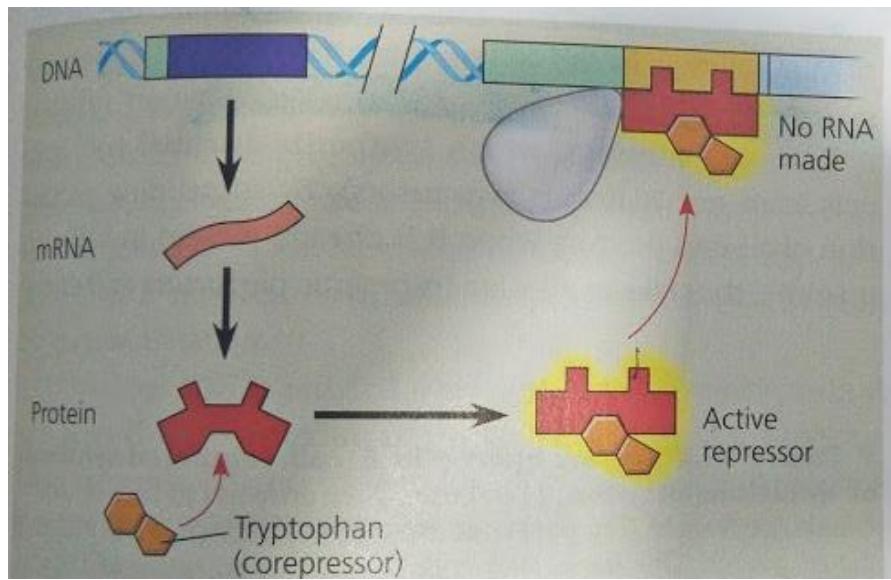
Consider that each cell type in an organism contains the same genome, thus the same number of genes. However, their expression, meaning making a product from them varies depending on the cell type, change of environment condition, etc. This happens both in eukaryotic and prokaryotic cells. All these expressions are regulated in the cell. We will learn basic principles underlining the regulation of gene expression in both prokaryotes and eukaryotes. Individual bacteria respond to environmental change by regulating their gene expression. *E. coli*, that lives in the human colon. Can tune its metabolism to the changing environment and food sources. For example, if the environment is lacking in the amino acid tryptophan, which is important for bacterial survival, the cell will respond by making the tryptophan from another compound. Later on, if the human host eats a tryptophan-rich meal, the bacterial cell stops making it. This is just one example of how bacteria respond to the expression of genes by tuning the metabolism. This metabolic control occurs on two levels. Adjusting the activity of metabolic enzymes already present. Regulating the genes encoding the metabolic enzymes.

Operons: The Basic Concept: 1961, Jacob and Monod. Genes are grouped (one transcriptional unit) for a cooperative function. They are under the control of one promoter and controlled by a single on-off switch, Operator (a segment of DNA). This is present between the promoter and the enzyme coding gene. The operator is worked upon by a repressor protein. The repressor binds to the operator and switches off the transcription.

Tryptophan absent, repressor inactive, operon on: When tryptophan is absent, the cell will make the tryptophan, with the help of five enzymes. These enzymes without a problem, as a repressor in this case will be inactive and not be able to bind to the operator sequence. This situation will not inhibit the RNA Polymerase to transcribe all the genes present in the trp operon. See the figure.



Tryptophan present, repressor active, operon off: when tryptophan is present, it binds to the inactive repressor, makes it active. This complex then binds to the operator, provides constraint to RNA polymerase to transcribe the operon genes. Thus, stop the synthesis of Tryptophan.

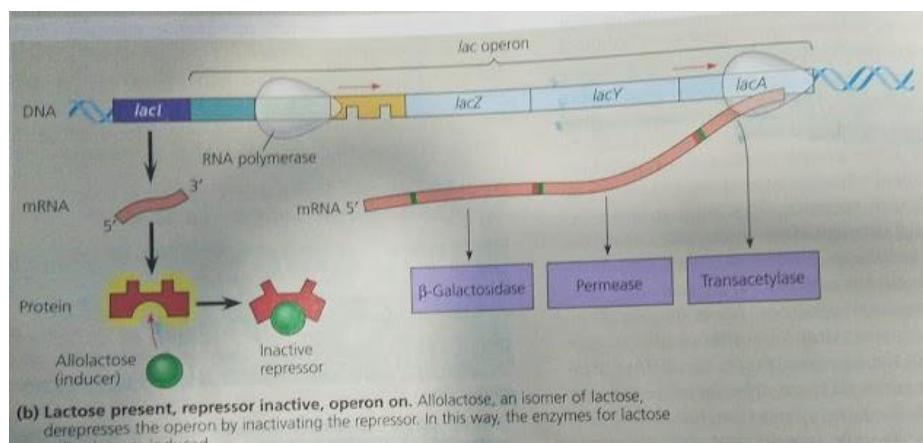


The trp operon is one of the examples of how gene expression can respond to changes in the cell's internal and external environment.

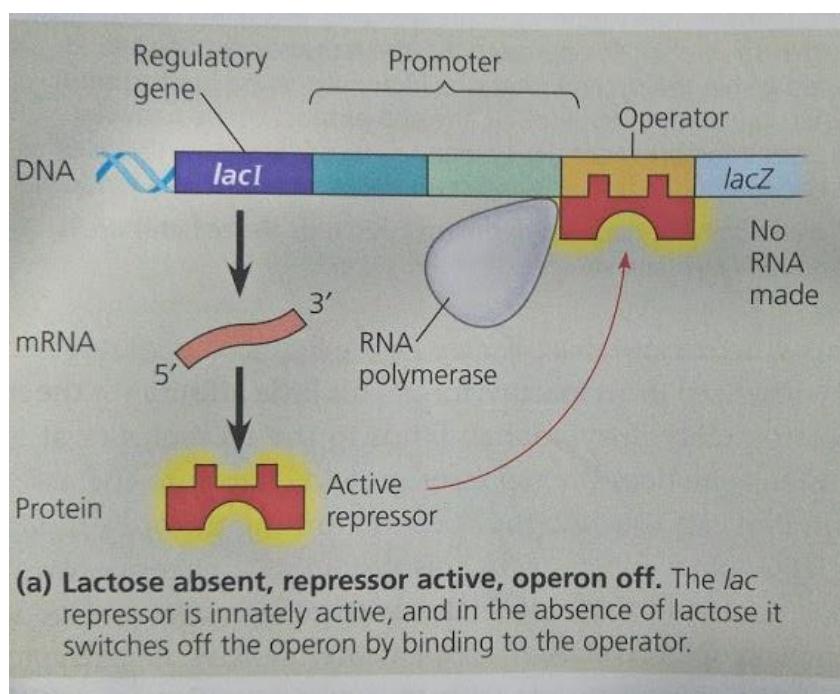
Lecture Note 23: Regulation of Gene Expression: Lactose, lac Operon

Jacob and Monod carried their pioneering work on lac operon. If the host drinks milk, *E. coli* can make use of lactose present in the milk. Lactose, a disaccharide is hydrolyzed by the bacteria into glucose and galactose inside the cell. This reaction is catalyzed by the enzyme β -galactosidase. A few molecules are present in the cell of *E. coli*. If lactose is added to bacteria, β -galactosidase amount in the cell increases a thousandfold within 15 minutes. The gene for β -galactosidase is part of the lac operon, where two more genes are present.

Lactose present, repressor inactive, operon on: Three genes are expressed into three enzymes. β -galactosidase helps in the breakdown of lactose, Permease helps the entry of lactose inside bacteria. Transacetylase. It also helps in metabolism, but its complete function is not yet clear. LacI gene makes lac repressor, which is active by itself. The inducer however can make it inactive. Here inducer is allolactose, an isomer of lactose, which is formed from lactose inside the cell in small amounts. This acts as an inducer by binding to lac repressor and inactivates it. This inhibits binding of repressor to the operator. Thus RNA polymerase can easily synthesize mRNA transcripts and makes three enzymes after translation, to use lactose by *E. coli* as a source of energy.



Lactose absent, repressor active, operon off: Once lactose is utilized, there is no need to express three enzymes, equilibrium switch and repress becomes active. It binds to the operator, hence shut down the synthesis of enzymes.

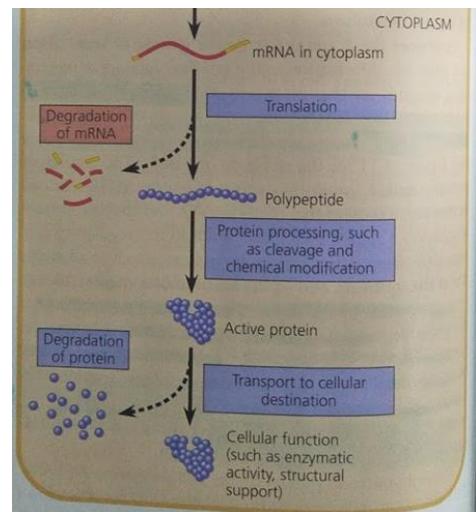
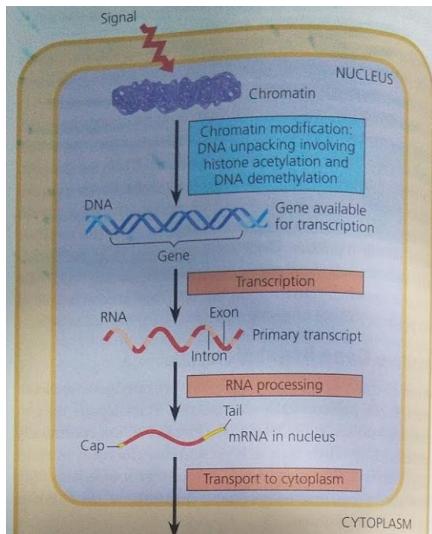


- Trp operon is said to be repressible operon, because its transcription is usually on, but can be inhibited /repressed by a specific molecule trp upon binding allosterically to repressor, regulator protein.
 - The enzymes for trp synthesis are known as repressible enzymes, and function in anabolic pathways, build the molecules.
- In the other hand, lactose is an inducible operon, as it is usually off, but can be stimulated /induced when a specific molecule (lactose/allolactose) interacts with the regulatory protein.
 - The enzymes of lac operon regulatory pathways are known as Inducible enzymes, and generally works in catabolic pathways break the molecules.
- Regulation of Both trp and lac operons is considered as negative control genes because the operon is switched off by the active form of repressor.

Lecture Note 24: Regulation of Gene Expression: Eukaryotic Cells

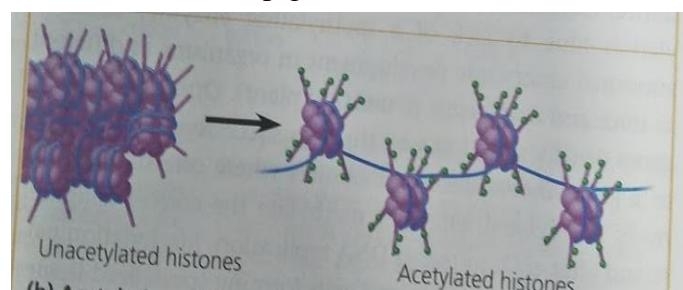
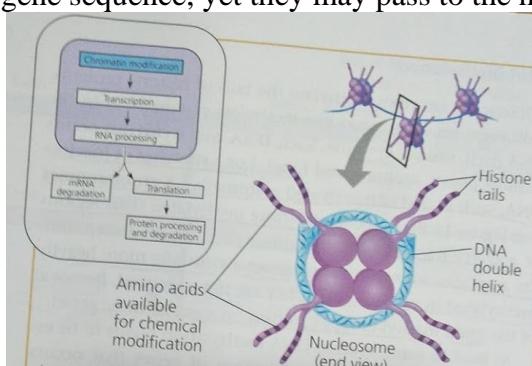
A typical human cell might express about 20% of its protein coding genes at any given time. Highly differentiated cells for example muscles and nerve cells, express even lesser fraction of genes. Almost all the cells have small genome, however the subset of genes expressed in the cells of each type. This is called differential gene expression, the expression of different genes by cells with the same genome.

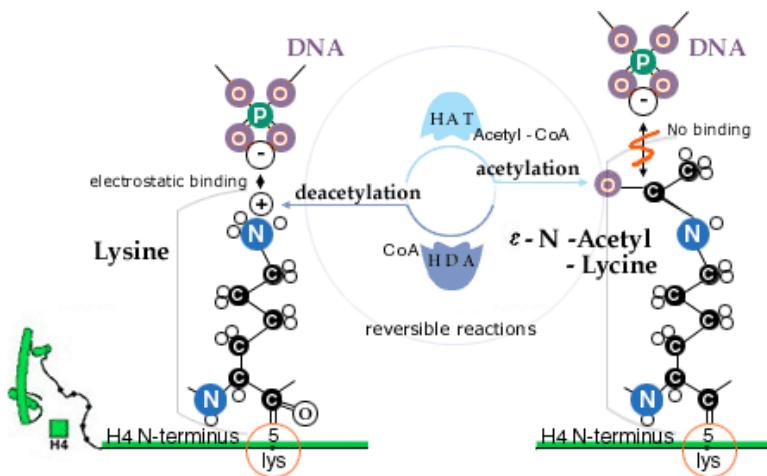
Stages in Gene Expression that can be Regulated in Eukaryotic Cells: It happens at the level of DNA, RNA, and protein. The expression of any given gene does not involve every stage shown. The figure is general to represent different stages where control can be executed. These are potential control points at which gene expression can be turned on, off, accelerated or slowed down.



Structural organization of Chromatin packs cell's DNA into compact form inside the nucleus and also help in regulating gene expression. Chemical modification of histone tails can affect the configuration of chromatin and thus gene expression. Histone acetylation seems to loosen chromatin structure and thereby enhance transcription. These nucleosome modifications are added and removed by enzymes that reside in the nucleus; for example, acetyl groups are added to the histone tails by histone acetyl transferases (HATs) and taken off by histone deacetylases (HDACs). Adding an acetyl group removes the positive charge from the lysine, thereby making it more difficult for histones to neutralize the charges on DNA as chromatin is compacted.

In addition, DNA Methylation and Epigenetic inheritance: the chemical modifications although do not change the gene sequence, yet they may pass to the next generation, it is known as epigenetic inheritance.

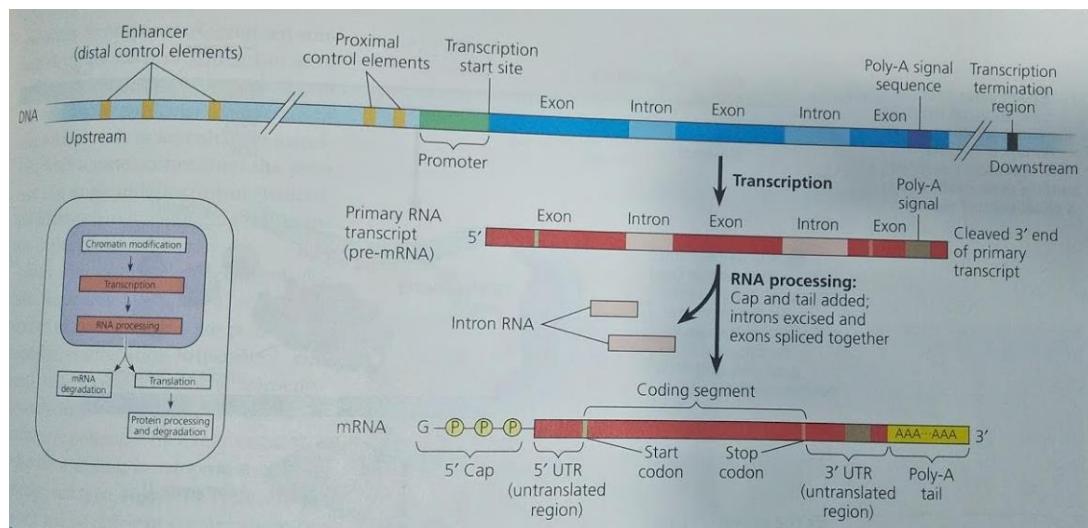




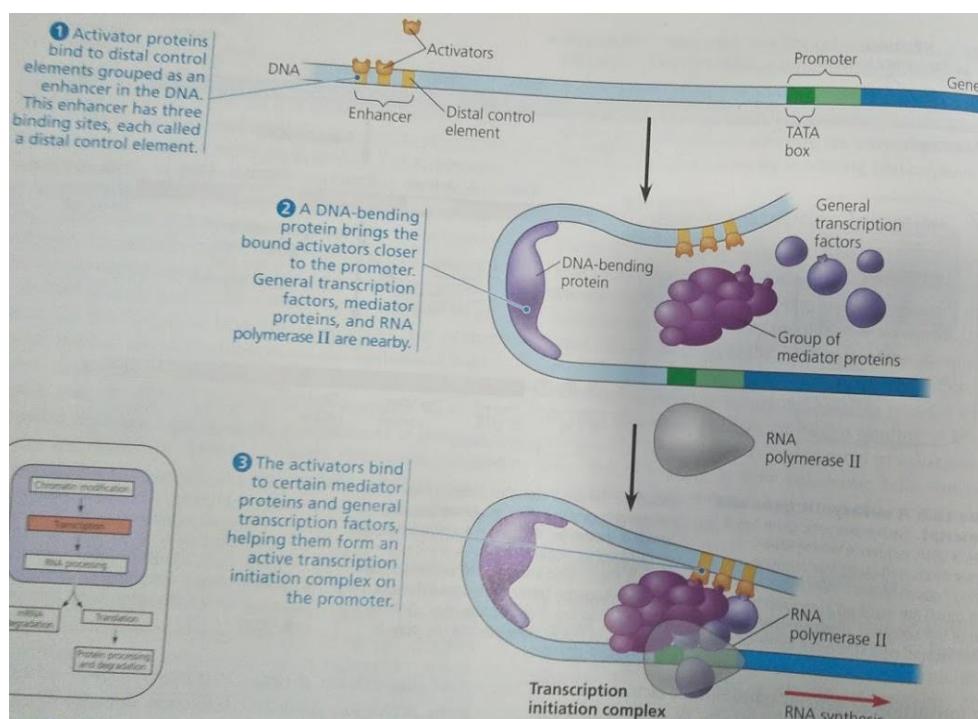
Lecture Note 25: Regulation of Gene Expression: Regulation of transcription

Another example of regulation of gene expression control is through transcription initiation: To understand this topic, recollect the architecture of a gene, which is also evident from the following figure.

Apart from the gene, there are associated control elements with most eukaryotic genes. They can be present near to the gene as well as far away from the gene. Thus, named as proximal control elements and distal control elements. They are able to bind to protein molecules (specific transcription factors), which in turn regulate the transcription. These specific transcription factors can be cell specific. While there are general transcription factors which can be part of every cell. While general transcription factors with RNA polymerase can produce low number of transcripts, but having a particular specific transcription factor and its interaction with general transcription factors and RNA polymerase can give rise to more expression of a given gene.



A model for the action of enhancers and transcription: Activator proteins bind to distal control elements grouped as an enhancer in the DNA. A DNA-bending protein brings the bound activators closer to the promoter. Other transcription factors, mediator proteins, and RNA polymerase are nearby. The activators bind to certain general transcription factors and mediator proteins, helping them form an active transcription initiation complex on the promoter.



A particular combination of control elements will be able to activate transcription only when the appropriate activator proteins are present.

▲ Figure 19.7 Cell type-specific transcription. Both liver cells and lens cells have the genes for making the proteins albumin and crystallin, but only liver cells make albumin (a blood protein) and only lens cells make crystallin (the main component of the lens of the eye). The specific transcription factors (activators and repressors) made in a particular type of cell determine which genes are expressed. In this example, the genes for albumin and crystallin are shown at the top, each with an enhancer made up of three different control elements. Although the enhancers for the two genes share one control element, each enhancer has a unique combination of elements. All the activators required for high-level expression of the albumin gene are present only in liver cells (a), whereas the activators needed for expression of the crystallin gene are present only in lens cells (b). For simplicity, we consider only the role of activators here, although the presence or absence of repressors may also influence transcription in certain cell types.

