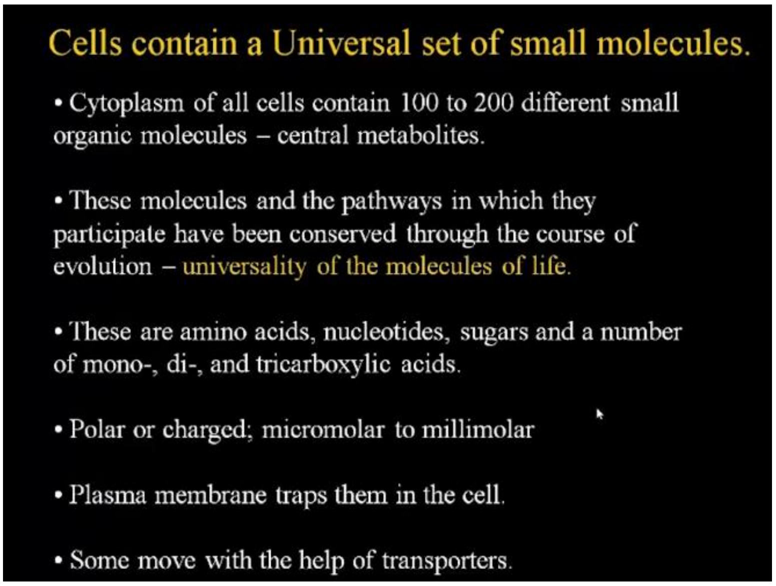
Introduction to Biomolecules
Prof. K. Subramaniam
Department of Biotechnology
Indian Institute of Technology - Madras

Lecture – 1
Introduction to Biomolecules - Part – 1/2

(Refer Slide Time: 00:14)



So to begin with you know among the subjects the main broad areas one needs to learn to understand biology to get a holistic view of biology there are a few subjects that are really the pillars of biology and one of them is biochemistry. The founding concepts of whatever specialized area you are going to be interested and you are going to pursuing either in your job or in research for all of them the foundations come from these 3 topics.

One is biochemistry, the other one is genetics and the third one is cell biology. So the concepts of these 3 subjects are really crucial to pursue any area of research or you know any kind of specialization in biology. So therefore you need to take additional effort and interest to learn the concepts that you are going to you know encounter in this course.

It is primarily because living systems be it how a healthy human body works or how a human embryo developed into a human adult or identifying the causes and treating human disease or to understand the natural world as it exists you know different organisms, biology is not just only the human body and human diseases that is only a small fraction of the whole thing that biology encompasses.

For all of that we need to learn the chemistry of it, primarily because organisms are made up of molecules and the molecules follow the laws of chemistry and physics and biochemistry therefore focuses on these molecules. What are these molecules, what do they do and how do they form the an organism's body and how it enables the organisms do different things, how do they reproduce and how organisms change over time giving rise to variations on which natural selection works.

For all of that we need to understand the logics of these molecules and that is what biochemistry deals with and the genetics gives you the principles of how this inheritance, you know these molecules how the information gets transferred from one generation to another generation and what are the guiding rules there or governing rules there and that is why genetics becomes another pillar in learning biology.

And the third is these molecules organize themselves into fundamental living units called cells. So therefore, there are cell specific characteristics in terms of their behavior, etc. and therefore there are set of cell biology specific concepts that you do not see in physics or chemistry and that is why the cell biology principles become important. So these 3 subjects are really crucial to understand everything else and that is why biochemistry is one of the core courses.

So viewed in this way with you are interested to understand biology and for that recognizing that biochemistry is an essential subject you will find the biochemistry interesting and you will be willing to invest more time and effort in learning and when you do that you will really find biochemistry very exciting. On the other hand if you look at biochemistry purely in for this is one of the core courses get away with it, I got to somehow manage to get a good grade out of this, then you may not find it very interesting.

So you will find it really hard, you know because it involves a lot of complicated concepts and you will find it hard. So therefore please orient yourself to the first way of looking at the subject. So now let us begin. So to begin with if you just look in an open minded way, think about living things and compare any living thing you need not take humans but take the simplest organisms and then you compare them with any of the machines that you use.

You will quickly realize how that organism is way more sophisticated than the most sophisticated machines or an instrument or a gadget that you use and similarly if you take what you can see with your eyes and compare with any camera you will quickly realize how sophisticated your eyes are and your ability to see things and this much of sophistication and another part I want to bring into the discussion is look at the diversity.

Organisms are starting from the SARS-CoV-2 virus that is making us not to meet in person and talk via google meet to a more complicated bacterium like E coli to plants, animals, birds, reptiles and all the creatures that are in the ocean and ourselves and the other mammals the diversity is enormous and this much diversity and the sophistication that I first explained all of that very surprisingly made possible by a couple of 100 molecules.

So you do not need to really worry about a whole lot of complex chemical structures here, so we have very, very few molecules that we need to learn. So if you look at the first bullet in the slide, cytoplasm meaning the cellular component in the main matrix within the plasma membrane so contain about 100 to 200 different small organic molecules and these are what we call as metabolites.

So the small organic molecules, so what are the big one I will introduce as we go, the small molecules are only less than 200 and these 200 we call as the central metabolites and these molecules and the pathway in which they participate meaning these molecules undergo a series of reactions and that series is called a pathway participate have been conserved, for example morning whatever you ate.

I am sure you ate something wheat based or a rice based and both are nothing but starch and starch is a polymer and if you cleave into monomer and the monomer is glucose. So glucose is one of these 200 molecules that we are learning about, small organic molecule and the way we get energy from glucose, the series of chemical reactions through which we get energy from glucose that series.

And the conversions the glucose undergoes to produce energy is the same in bacteria as well and that is what this bullet is telling you as it has been conserved. So glucose is glucose in E coli too, E coli being a bacterium and there what glucose is used the same way only you use glucose in your cell and this conservation we call as universality of molecules of life.

So if you think of defining life you will soon find it is not an easy question. You know people will say organisms, living things grow, many other things grow too, even mountains grow, you will think the organisms reproduce, there are crystals that reproduce. So there are lot of non-living things have a lot of things that we associate only with life. So at the end the way people think about life is the kind of molecules that make life are not found outside life.

The molecules that make life they are there in all living things doing very similar things that is what is this concept of universality of molecules of life. And what are these small molecules they are in the third bullet, these are amino acids, nucleotides, sugars and a number of mono, di and tricarboxylic acids. do not worry about these names for now, we will actually be understanding the structure of each one of them.

At that time these names will become intuitive, you will very readily understand why these are called with these names. So these are the main molecules amino acids, nucleotides, sugars and tricarboxylic acids. These are the four major classes and there are variations in the among these four and that is what we are primarily going to learn and these molecules may be polar like having ability to interact with water.

And they may be charged like they can dissociate releasing a proton or they can accept a proton so they can be positively or negatively charged. Some are hydrophobic too, you know they do not interact with water and they are present in different concentrations, you know they can be present in micromolar quantities to millimolar quantities. So the next thing is all of these are present within the boundary of plasma membrane.

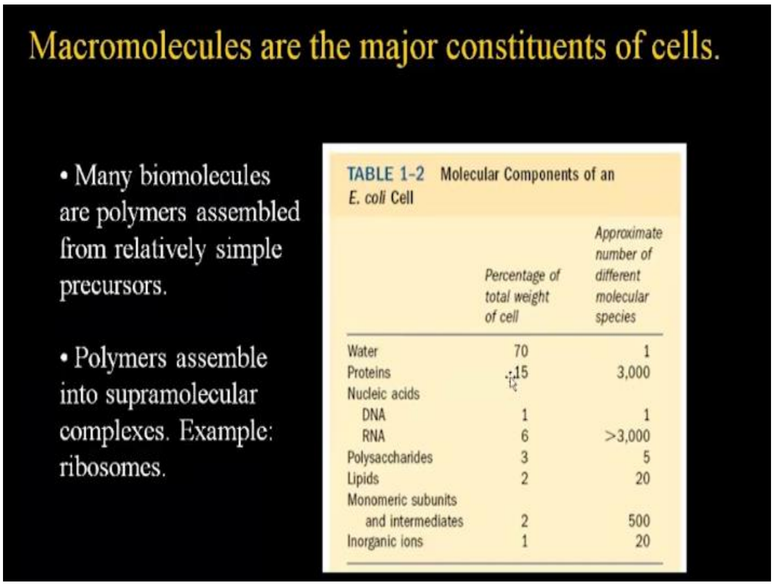
So plasma membrane itself is made up of these molecules, primarily a variation of these tricarboxylic acids. So the plasma membrane is hydrophobic, so most of these molecules are hydrophilic, so they are trapped within the plasma membrane and some of them help in transport across the plasma membrane and they are transporters.

(Refer Slide Time: 11:37)

So then you have some molecules that are produced only in some organisms not in all organisms and they are not part of this universality of the molecules of life. So they are like specialized ones, they are not like your core these are like electives and these are most often produced by plants. They are called secondary metabolites and to give you an idea you know the resin in the gum that you use, nicotine you know produced by the tobacco plant, caffeine produced by tea and coffee plants, quinine, morphines.

So these are secondary metabolites, they are not essential for existence but they produce them for additional sophistication of life. So I will talk about specific examples when we get there. All these molecules put together collectively present in an organism we call them a metabolome. All proteins that an organism produces you call proteome and all genes present in an organism you call as genome. So similarly, all the metabolites present you call as a metabolome.

(Refer Slide Time: 12:58)



	Percentage of total weight of cell	Approximate number of different molecular species
Water	70	1
Proteins	15	3,000
Nucleic acids		
DNA	1	1
RNA	6	>3,000
Polysaccharides	3	5
Lipids	2	20
Monomeric subunits and intermediates	2	500
Inorganic ions	1	20

So next let us look at their various of compositions like what is the relative abundance of each of this and when you look at it, you will find water is like 70% of the cellular molecules or water, proteins next highest that is 15. They are made up of amino acids and we will learn about that soon. Then you have nucleic acids, they come in 2 flavors, one is DNA, another one is RNA and there you see you have about 3000 different proteins.

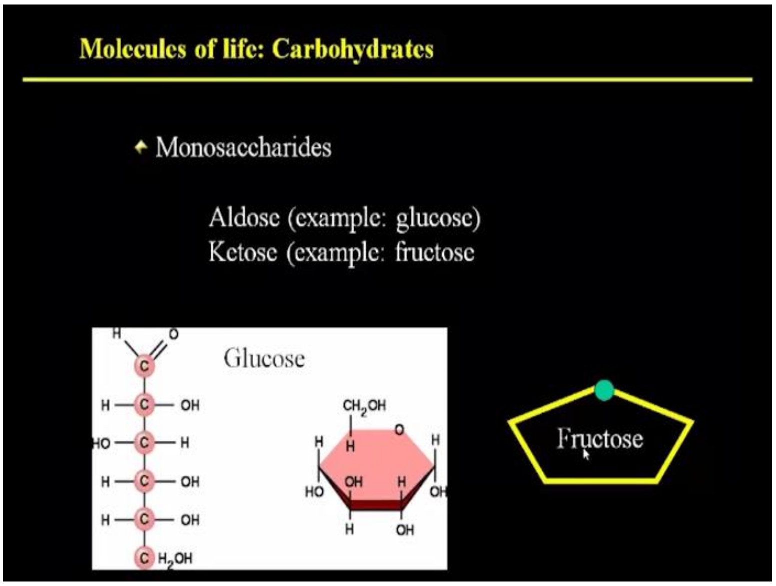
And similarly you have more than three 3000 different RNA. Polysaccharides only about 3% and you have 5 kinds, lipids 20 kinds but only 2%. Then you have the monomeric subunits of these molecules and the intermediates, they make up about 2% and there are only 500 types and then you have about 20 different inorganic ions making about 1% of the cellular weight and now from the metabolites we move to another important concept that is many biomolecules.

You know other than these small molecules that are present in the cell which is numbering 100 to 200. Majority of the molecules in body exist in the form of polymers. So several monomers join together in a repetitive fashion making polymers. For example amino acids that we saw in the previous two slides ago. So amino acids, these amino acids are monomers, when they join together they make a protein.

So most of the molecules are present as polymers so that is an important concept you need to remember and these polymers sometimes assemble into large structures more than the dimensions of normal molecules called supramolecular complexes. A good example familiar to all of us is ribosomes, you know the ribosomes are the cellular structures, non-membranebound structures involved in protein synthesis. So these are the structures that make proteins by reading the sequence present in the mRNA.

So the main point of this slide is that most molecules are present as polymers and then you saw in this table the relative abundance of the different classes of molecules and then the varying kinds like water there is only one H2O, but protein you have about 3000 different proteins and so on.

(Refer Slide Time: 15:54)



Okay so now let us get into first set of molecules. So the first molecules of life that we are going to learn about are carbohydrates. So before we get into the details of carbohydrate let me explain its importance then you may be motivated to listen more and to learn about carbohydrates. So carbohydrates are the primary energy source you know for cellular activities for the molecule that I mentioned few minutes ago is glucose.

So glucose is actually a central molecule in the energy production in all organisms. So by oxidizing glucose through cellular respiration, we obtain the energy required for all our work for the functions of our internal organs to the physical activity, for example to take a class like what I am doing right now. The energy required for all of it comes from glucose. So the food that we primarily eat; particularly in India more than 70% of what we eat are actually carbohydrates.

You think of like today when you are going to sit down for lunch do not pay attention to the varieties that are there in your plate but instead look at the relative abundance which one you are eating lot more than the others, then you will soon realize it is either derived from wheat or derived from rice and that is what you are primarily eating and both of them as I said earlier they are primarily starch and starch is a polymer made up of glucose.

So most of what we eat is carbohydrate. We will learn a lot more about carbohydrate, other functions of carbohydrate when we get to the carbohydrate chapter. So today's class is a preliminary introduction to the different molecules. So carbohydrates are the monomer we usually call as monosaccharides, saccharide you know stands for sugar in either Greek or Latin which I cannot recollect now.

Mono meaning the monomer, monosaccharide and glucose is a monosaccharide, a single monomer, and when you join two of them you call that as disaccharide and when you join multiple you call oligosaccharide. When you make a long chain you call polysaccharide and these monosaccharides exist in two different main forms. One as shown in this you know the molecular structure you have an aldehyde group CHC double bond O.

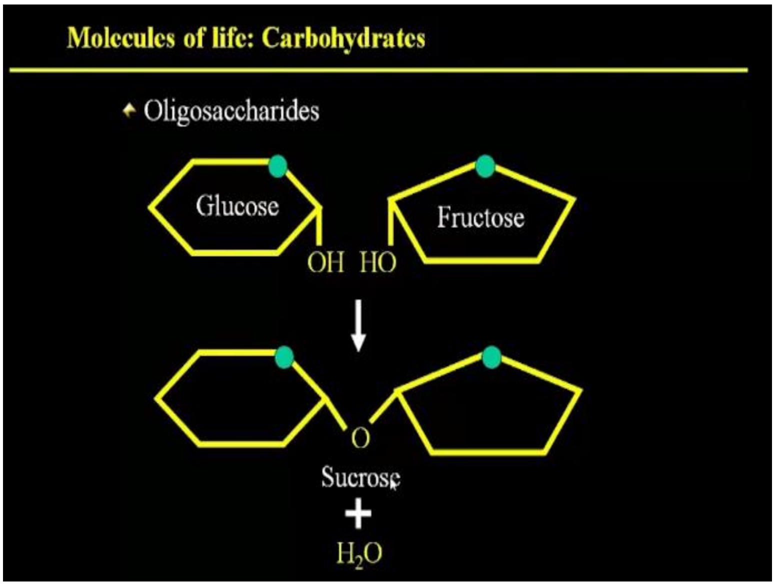
And when you have the aldehyde group present in the monosaccharide molecule you call that as an aldose. So glucose is an example. So later we will learn these molecules do not exist the way it is written here. So when you have the aldehyde group it is an aldose and glucoses and aldose therefore and these molecules do not exist like this. This aldehyde carbonyl group can interact with any of these alcoholic groups present in other carbons and they form a ring like structure.

So we will get into this ring structure again later, I needed to tell this now simply to go to the structure on the right side. So such a linked ring-like structure is what is shown here. So this is how they exist in solution or in our cell, in nature this is how they exist and some of the monosaccharides instead of having an aldehyde group they have a ketone group C double bond O alone and fructose is one such example.

When you join glucose and fructose, then you get what is called sucrose which is a disaccharide and that is the common sugar that we use. You know when you make coffee to sweeten it you add sugar that sugar is sucrose which is a disaccharide containing the monosaccharides, glucose and fructose, glucose is an aldose, fructose because it has a keto group we call that as ketose.

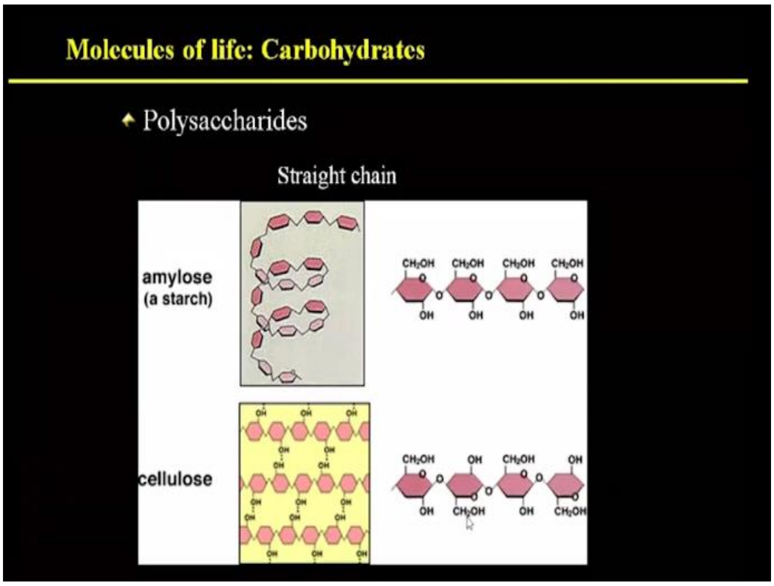
So how do they join? They join through these hydroxyl groups present on the side chains. You see this after linking these are the hydroxyl groups you see. So the carbon itself is numbered like this side where my cursor is pointing is C1 and this is 2, this 3, 4, 5 and 6, so this is the sixth carbon. So there are 6 and this is how the numbering goes on. So from this end we start and then this free CH2OH group attached, this one is the sixth one.

(Refer Slide Time: 21:23)



So now any of these OH groups can function as a functional group to join with another molecule and that is what is shown here OH of glucose. So this is the C1 O hydroxyl group alcoholic group and this is C4, 1, 2, 3, 4. So when they join together so it is like a dehydration leading to condensation, this bond we call glycosidic bond COC, so this is the glycosidic bond. So this is a dehydration based polymerization. So this is a glycosidic bond and here you see this between glucose and fructose as in sucrose disaccharide.

(Refer Slide Time: 22:10)



These can continue to link in very long chains as shown here, you know this is glucose structure, this is 1, 4 1,4 1,4 linkage a long chain continuously going that is part of starch, you know starch contains two kinds of chains, one is amylose and amylose is simply 1, 4 linkage. So we will be talking about one more term in terms of whether this OH is oriented this way or it is oriented above this ring and that leads to alpha and beta.

Do not worry about it for now, I will explain alpha and beta later when we go to carbohydrate chapter. For now you need to understand just this is 1 and this is 4 and this is 1, 4 linkage. So 1, 4 glucose continuous chain is amylose which is one of the components of starch. Then you can have branches, for example you have another OH available here, another OH available here that can link to other molecules.

So suppose let us say this particular glucose already having linked to another glucose now let us say its sixth carbon's hydroxyl group links to another glucose molecule then you got a branch here. Such branched structure leads to amylopectin or a stored carbohydrate in our liver primarily in animals called glycogen. So amylopectin and glycogen they are again glucose but they have a lot of branches that is all, otherwise amylose is good enough as a polysaccharide to remember.

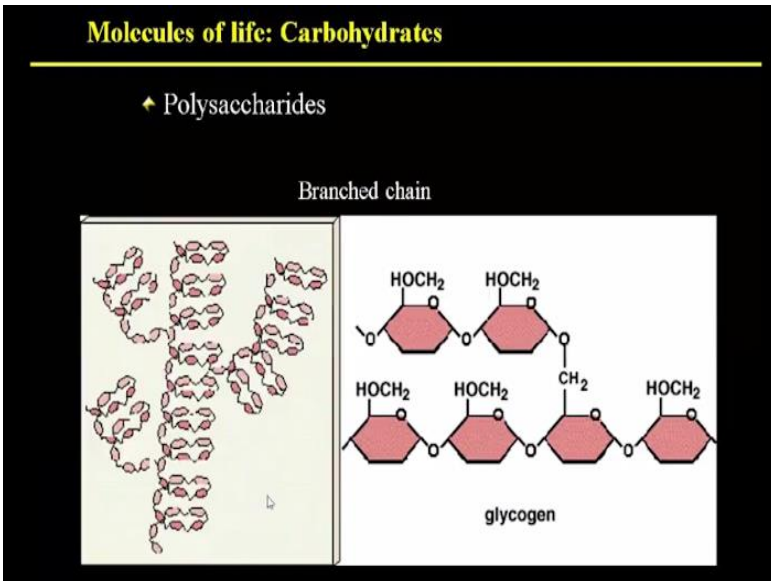
So this is starch, so this is our principal source of energy and carbohydrates are not just source of energy alone, they are also major structural components. If you take a tree or a tree branch, the main trunk of the tree or the leaf or the structural element for all of them is actually large number of chains of glucose but linked in this fashion, it is again 1,4 linkage but this OH is present in this glucose on above the ring the plane of the ring and this is what we call as the beta orientation.

Therefore starch we say alpha 1, 4 this we say beta 1, 4. So our body does not have a way of hydrolyzing this beta 1, 5 linkage and that is why we cannot eat the cellulose whereas our body can hydrolyze this alpha 1, 4 therefore you are eating starch. So cellulose is the most abundant biomolecule in the planet because you know you can easily see plants occupy almost every possible ecological niche and they are most abundant and now they are primarily made up of cellulose.

So cellulose there is providing a structural function. The tree trunk is able to stand straight and grow to more than 100 meters in some cases that is made possible by the structural strength provided by cellulose fibers and some of the organisms like microorganisms can cleave this beta 1, 4 linkage, for example fungus you would have seen fungus growing on dead wood lying particularly in rainy season that is because they can hydrolyze the beta 1, 4 linkage and make glucose out of cellulose.

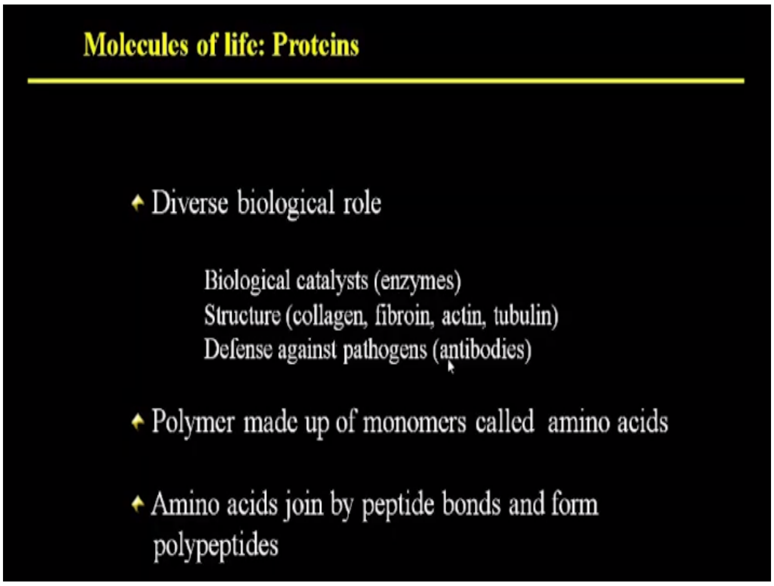
Some of the animals like for example cattle that we raise in our farms they eat grasses and leaves and the stems of some of the plants like you know sorghum or maize that is because these cattle harbor microorganisms in their stomach which produce enzymes that can cleave this beta 1, 4 linkage. So in this sense cellulose becomes food as well, not just a structure, alright.

(Refer Slide Time: 26:59)



So here is branch chain, example glycogen I just told you. So here you have here it is 6 and here it is 1. So here you have 1, 6 linkage branch in addition to alpha 1, 4 that is there and this is a cartoon representation of a large glycogen.

(Refer Slide Time: 27:28)



Alright so more about carbohydrates much later when we get to carbohydrate chapter, so now let us look at what is the next important molecule that we need to know. So they are the proteins. So what are proteins? Proteins are again polysaccharides. The monosaccharide repeating unit of polymer, sorry I said polysaccharide, they are polymers. So the repeating monomer in these polymers is amino acids.

So there we called monosaccharide, here we call amino acid. There the polymer we call polysaccharide, here we call that as polypeptide or simply protein. So two amino acids joined together you call dipeptide, dipeptide, tripeptide and so on. Then you have oligopeptide, then you have polypeptide, large number of amino acids linked together. So the first thing just like in carbohydrate here again we will direct our attention to what proteins do.

So all the chemical reactions, you know the series of reactions through which the metabolites undergo conversions which we called pathways in the first slide, the individual reactions of those pathways beat producing energy out of glucose or cleaving individual glucose molecules out of glycogen or starch or replicating the DNA or assembling microtubules based spindle for cell division, all of them are done by catalysts.

These are all chemical reactions that do not readily take place at the standard conditions of 37 degrees in aqueous medium and that concentrations present in our cytoplasm. They do not normally take place in the time duration of life, so they are all catalyzed and the catalyst for all these reactions are called enzymes and these enzymes are proteins. So now you understand the importance of proteins.

Without proteins nothing will happen in a living cell and these proteins are also surface molecules. For example the SARS-CoV-2 virus binds to our lung cells by a protein present on its surface called spike protein and that binds to a receptor present on our lung epithelial cells and they are structural components as well just like cellulose carbohydrate being a structural component in plants.

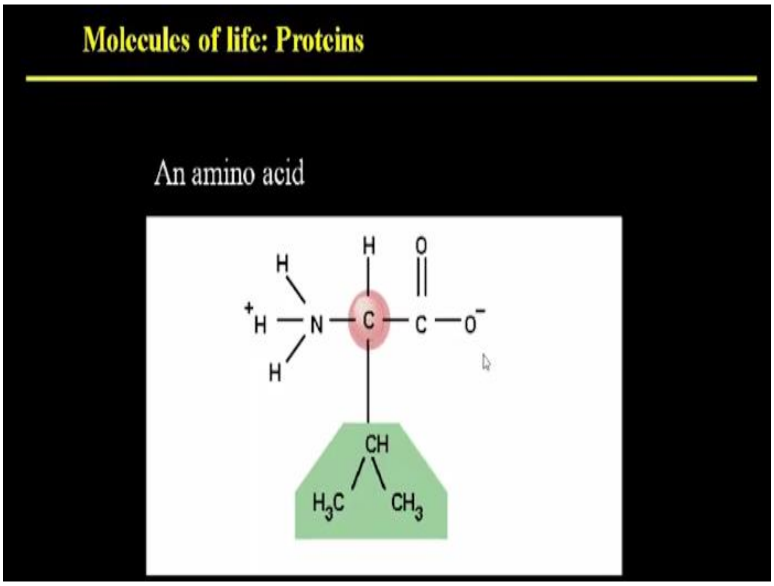
In animals, our cells are glued together using extracellular matrix say the material present among the cells, not inside the cell but in the space among the different cells and that extracellular matrix contains primarily collagen and fibroin another important molecule. For example in silk you have fibers, our hair keratin and within the cell we have structural scaffolds made up of actin and tubulin.

So these are structural molecules, again they are proteins and say some of our defense cells like the immune cells, B cells or T cells they produce molecules that fight against pathogens and those molecules themselves are proteins in nature. For example, the antibodies produced by our B cells they are made up of polypeptides. So therefore proteins do diverse roles, they are catalysts, they are structural molecules and they defend against pathogens.

They do many more, so we are not getting into all of them. For example, they are involved in cell signaling also. So we are looking at major things which are core into all you know organisms. So only those we are looking at to get an idea of the functional diversity. So I already told you that the monomers of these polymers are amino acids and amino acids are joined by similar dehydration based bonding as we saw in carbohydrates, but here we call them peptide bonds.

So there in carbohydrates we called it glycosidic bonds or glycosidic linkage, here the bond between two amino acids we call them as peptide bonds. So you have many of them in a polypeptide chain. So now we look at that part of proteins.

(Refer Slide Time: 32:40)

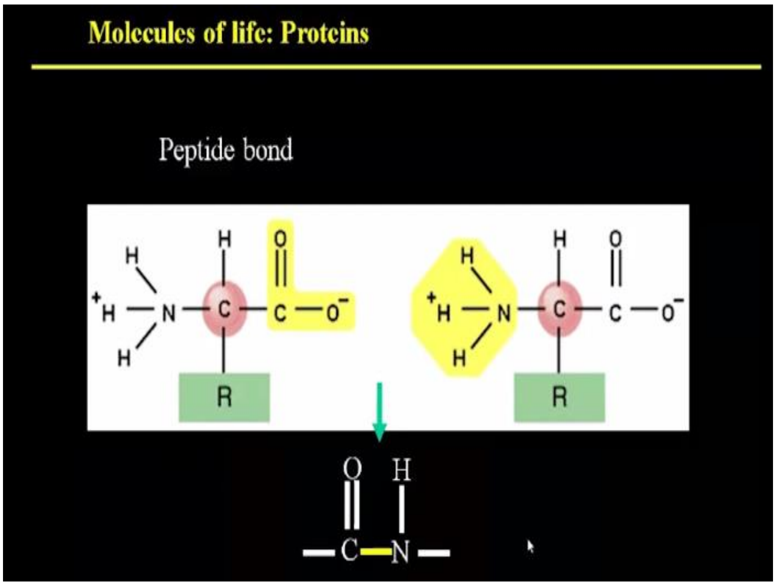


So this is a you know a typical amino acid. So in amino acid as the name says contains two important groups the functional groups, one is the carboxylic acid group. So shown here on the right, so that is the acid part of the name, and on the left you have the amino group so the amino part of the name. So these are amino acids. So an amino group and a carboxylic group linked to a carbon and carbon can form 4 bonds, the third one is a hydrogen.

These three are common to all amino acids. So, one amino acid differs from another amino acid only in its R group. So here you see a particular R group. In the most simplest amino acid, this R group is another hydrogen as seen in glycine, amino acid glycine has only a hydrogen atom bonded to the central carbon as the R group. So this carbon we call as the alpha carbon.

And as you see in this amino acid which has this R group has 4 different groups attached hydrogen, amino group, this R group and this carboxylic group. So this is a chiral carbon, center first generating stereoisomers. So, we will learn about that later, I just wanted to draw your attention to this asymmetry here, so the four groups are asymmetric attached to this carbon.

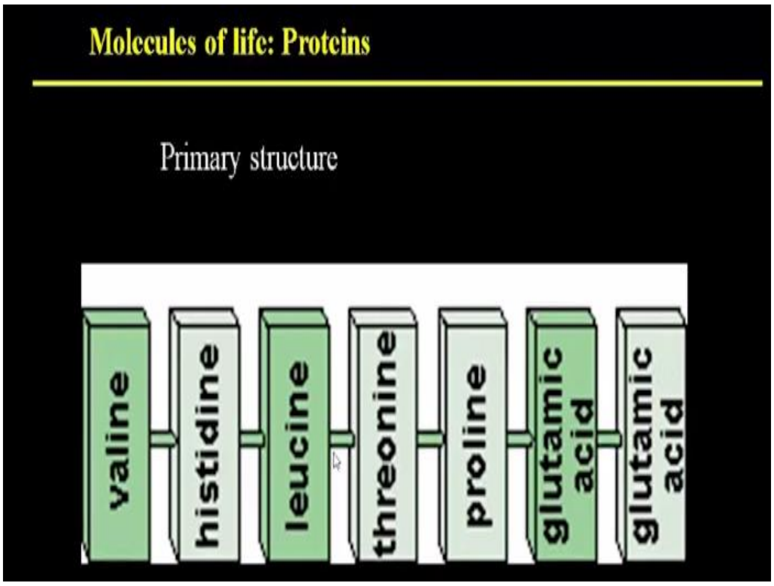
(Refer Slide Time: 34:38)



So now here you see two amino acids, the carboxylic acid of one and the amino group of another one again through a dehydration reaction can form this bond and this C-N bond is what we call as the peptide bond. So this C to the C alpha this we call as C-C alpha and this we call N-C alpha and this C-N is the peptide bond. So, these three are very important in determining the protein structure.

So again, we will learn about it later. So this is the peptide bond like glycosidic bond is to polysaccharides, peptide bond is for polypeptides and proteins.

(Refer Slide Time: 35:29)



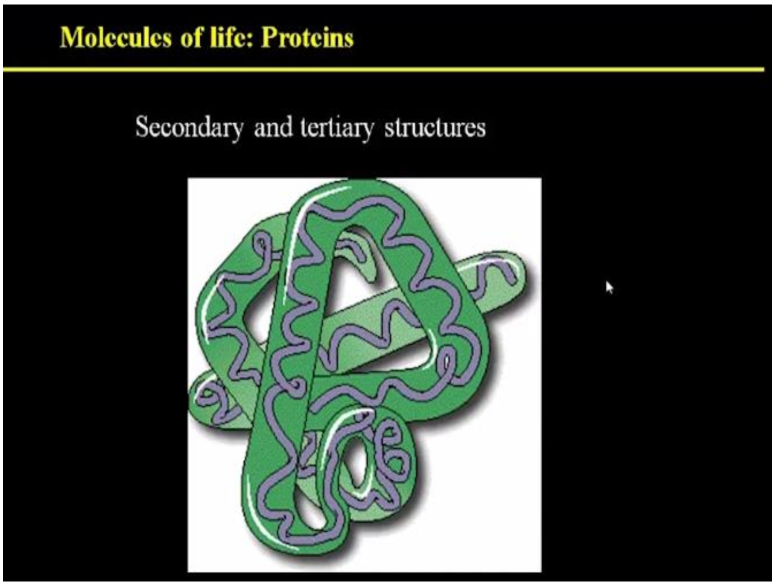
So this shows you the linkage. So there you saw the same kind of monosaccharide like glucose repeating multiple times in the polysaccharide, but here what you are seeing is it is not repetitive, sometimes they are repetitive but most of the time they are not repetitive, the different amino acids link together. For example, the valine you know that is what we saw here, so this is valine.

So valine linked to histidine linked to leucine linked to threonine and so on you have different amino acids. Here you see glutamic acid repeating twice, so they can repeat that it is just illustrate that point. So this long chain of amino acids linked via peptide bonds is what is a primary structure. So simply the amino acid sequence of a polypeptide we call as the primary structure. And these you know in carbohydrates we did not think about structure too much simply because they are repetitive and they are most often fibrous.

So therefore, we did not worry so much about it, but in terms of proteins for their functions the structure matters a lot and therefore we will briefly look at the structures. So depending on the sequence of the amino acids the side chains that come next to each other varies. For example the leucine side chain here will encounter histidine on one side, threonine on another side and threonine again one side leucine, another side proline.

So therefore the side chains that will come adjacent is determined by this primary sequence and these side chains can have attractive or repulsive interactions or could cause steric hindrance all that in determining the structure. So as a result, the primary structure determines a certain way of three-dimensional structure that is possible for a given sequence and therefore we focus on how this primary structure gives to specific additional higher ordered structures.

(Refer Slide Time: 37:56)

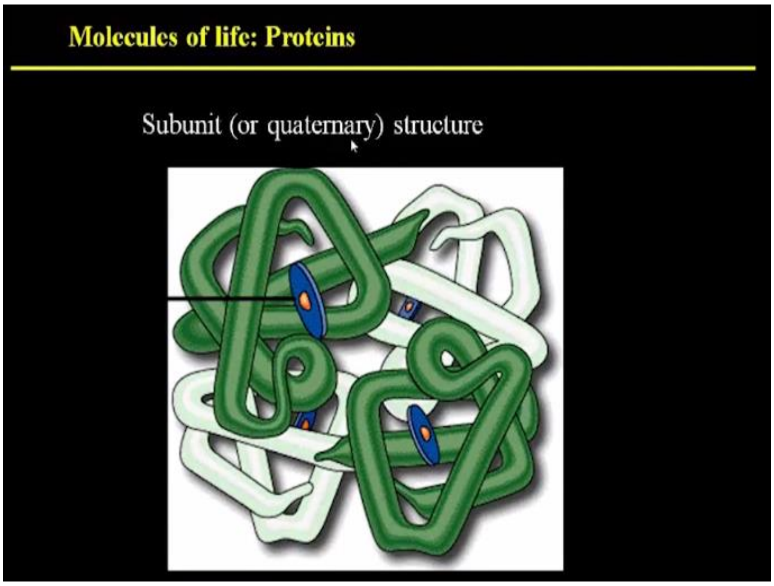


The first level of higher ordered structure we call as secondary structure. So secondary structure is actually in a long polymer of polypeptide like let us say you have 300 amino acids, let us say about 10 or 20 amino acids in a given part of this long 300 might form a local structure like for example as you see this purple ribbon like structure here which is twisted in a helical way.

So that sort of a local structure we call a secondary structure and this long 300 amino acid chain might further fold in an additional level of hierarchy in the structural organization to give you a three-dimensional structure and that is tertiary structure. The amino acid structure sequence itself is the primary structure and the local folding in a long chain of polypeptide we call a secondary structure.

And the overall folding of all these local folding into the final three-dimensional structure we call as tertiary structures. So most proteins have all these three structures, some have even additional ones where multiple polypeptides may interact among themselves.

(Refer Slide Time: 39:22)



A good example is hemoglobin present in our red blood cells that help in carrying oxygen from lungs to other tissues. So hemoglobin contains 4 polypeptides and they have additional groups that we do not worry for now. So they have a heme moiety which is not an amino acid based polypeptide moiety instead it is a different organic unconnected organic molecule that is embedded through interactions among the amino acids.

And this organic structure with the heme moiety coordinates an iron atom and so therefore this structure here. So proteins might contain like this non-proteinaceous moieties among the polypeptide chain. So hemoglobin has a structure like this, you know it has primary because of amino acid sequence, it has secondary and then it has tertiary and then multiple like four such polypeptides come together and that forms another hierarchy of structure which we call quaternary. So, these are the structures of proteins.