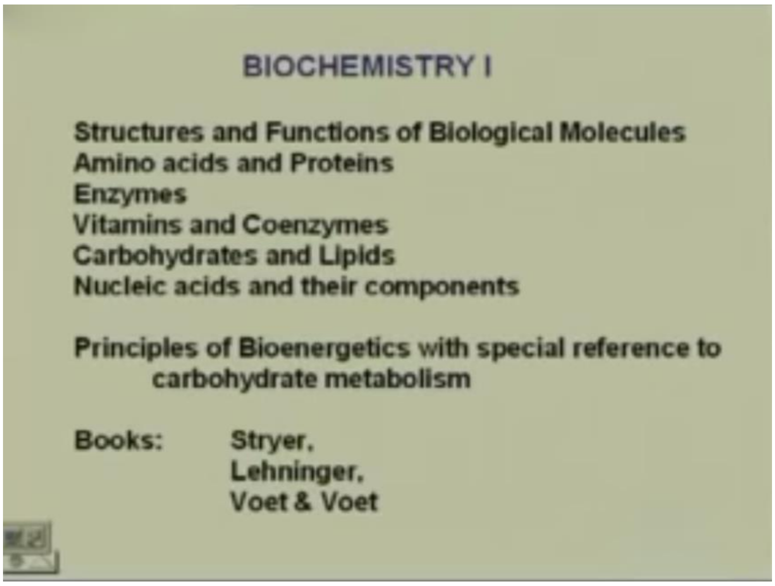
Biochemistry
Prof. S. Dasgupta
Department of Chemistry
Indian Institute of Technology – Kharagpur

Lecture -01
Amino Acids I

Welcome to the course biochemistry 1 conducted by Myself, Dr. S. Dasgupta, department of Chemistry, IIT Kharagpur. In this course, we will be studying certain aspects of biochemistry starting from the structures and functions of biomolecules right on to bioenergetics and metabolism.

(Refer Slide Time: 01:11)



BIOCHEMISTRY I

Structures and Functions of Biological Molecules
Amino acids and Proteins
Enzymes
Vitamins and Coenzymes
Carbohydrates and Lipids
Nucleic acids and their components

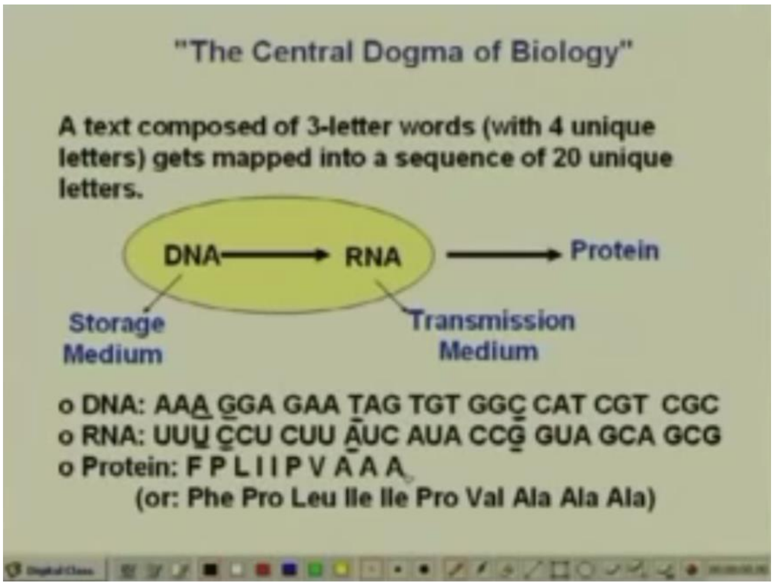
Principles of Bioenergetics with special reference to carbohydrate metabolism

Books: Stryer,
Lehninger,
Voet & Voet

The topics that we will be covering are structures and functions of biological molecules. In that, we will be considering amino acids and proteins, enzymes. Under enzymes, we will be considering the mechanisms of specific enzymes. Then, we will go on to vitamins and coenzymes, carbohydrates and lipids, nucleic acids and their components. All of the topics of this course will be covered in their entity as is relevant to this course.

We then, will be considering the principles of bioenergetics with special reference to carbohydrate metabolism. The books that we will be covering are common biochemistry books such as Stryer, Lehninger, Voet & Voet. Apart from this, there are many internet sites and websites that give us a lot of information on biology and biochemistry in general. So you could also look up those sites where there are specific course material available and specific tutorials and quiz materials also available for you to access.

(Refer Slide Time: 02:46)



"The Central Dogma of Biology"

A text composed of 3-letter words (with 4 unique letters) gets mapped into a sequence of 20 unique letters.

- DNA: AAA GGA GAA TAG TGT GGC CAT CGT CGC
- RNA: UUU CCU CUU AUC AUA CCG GUA GCA GCG
- Protein: F P L I I P V A A A
  (or: Phe Pro Leu Ile Ile Pro Val Ala Ala Ala)

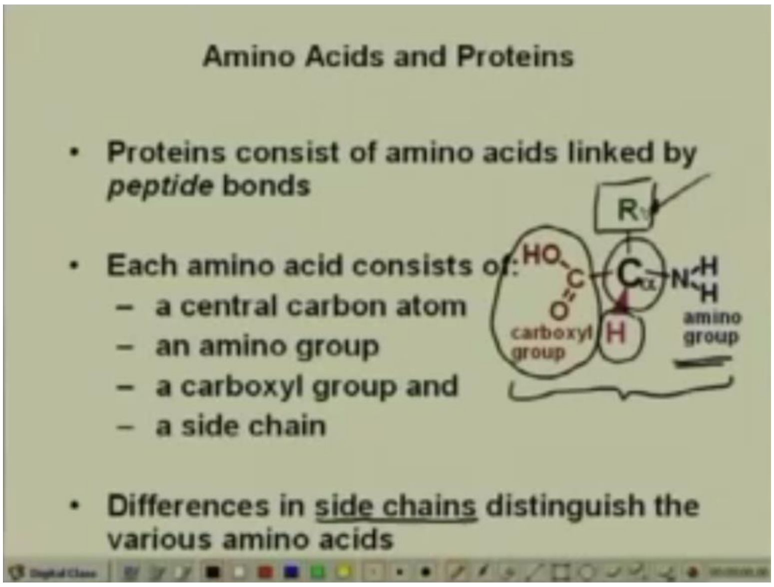
When we consider the central dogma of biology, the first thing that comes to mind is DNA. DNA is the storage medium. The central dogma of biology goes like this, DNA to RNA to protein. We have a text that is comprised of DNA, which is the 4 basis of DNA that are the storage medium. This is then transcripted to RNA, which is the transmission medium that also is comprised of 4 bases, one of them being a bit different.

The RNA is then translated to the protein. The alphabet of DNA, RNA and protein is slightly different. What is this alphabet? In DNA, we have 4 letters to the alphabet. The 4 letters are, as you can see here, A, G, T and C. These are the 4 letters that comprise the alphabet of DNA. If you look at the corresponding alphabet of RNA, you'll see that we have U, C, A and G.

When we go on to study the structure and the contents of nucleic acids the structures of each of these bases will be much clearer but for now, we have to know that the DNA and RNA are comprised of these letters, which actually represent nitrogenous bases. The protein alphabet is a bit different. The protein alphabet is sometimes represented as a 3-letter code which we will see in a moment or by the 1 letter code which is another representation of the very same 3-letter code.

The protein alphabet is comprised of 20 unique letters that tell us, what the protein sequence is. We will understand what an amino acid sequence is once we get into the details of what is the peptide bond and what is an amino acid. The first thing that we know or we try to understand here is the carbon atom.

(Refer Slide Time: 05:57)



Amino Acids and Proteins

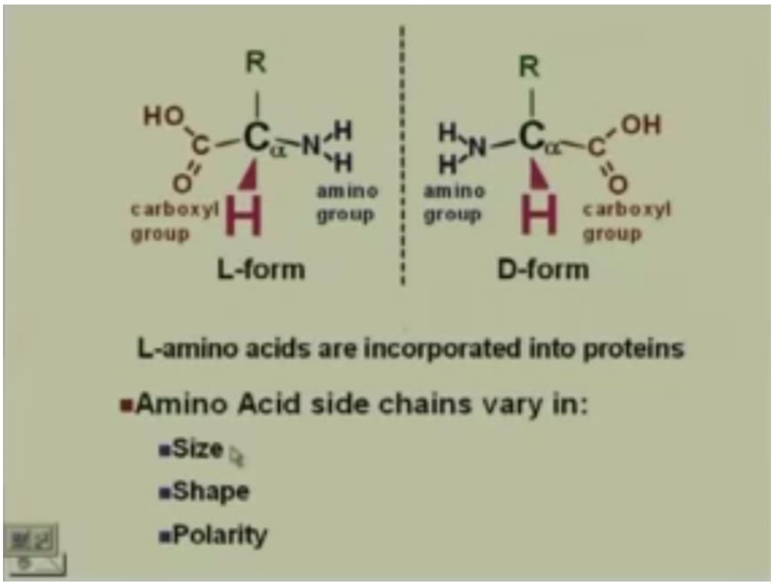
- Proteins consist of amino acids linked by *peptide* bonds
- Each amino acid consists of:
  - a central carbon atom
  - an amino group
  - a carboxyl group and
  - a side chain
- Differences in side chains distinguish the various amino acids

Now, if we consider what these proteins are actually made of, the proteins consist of amino acids that are linked by peptide bonds. We will understand this also in detail to see how a peptide bond is actually formed and how we can link these letters together into forming what would be like a sentence. Each amino acid consists of a central carbon atom, which you can see marked as C alpha here. So, we have the central carbon atom that is marked as C alpha.

To it, we have an amino group as NH2 group. We then have the carboxylic acid group, which is COOH. We also have what is called as an R group. Now this R group is what the side chain of the amino acid and we also have the hydrogen group, the hydrogen atom here. What is common to all amino acids is this part, because it has an amino group, it has an acid group, this is why it is called an amino acid. It also has an H, a hydrogen attached to it.

So, we see this central carbon is actually a cyral carbon, meaning that it is asymmetric, which again means that there are 4 different groups attached to the central carbon atom. And since all amino acids have a common set of groups here, in the amino group, the hydrogen atom and the carboxyl group, what differs is this side chain, which can be different atoms, different groups of atoms and this is what actually distinguishes the various amino acids. We are now going to consider what type of different R groups we can actually have.

(Refer Slide Time: 08:29)



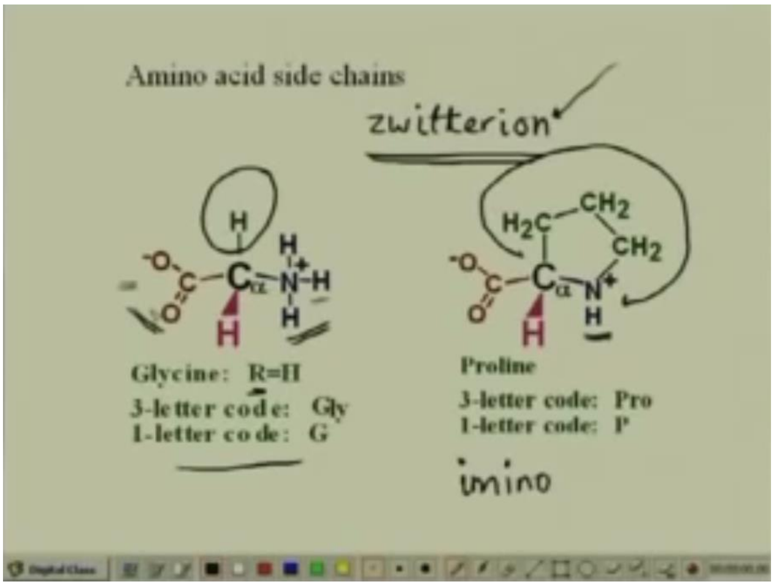
L-amino acids are incorporated into proteins

- Amino Acid side chains vary in:
  - Size
  - Shape
  - Polarity

Now, if we look at the different forms of the amino acids that could be incorporated into proteins, as we mentioned in the previous Slide, we have an amino group, we have a carboxyl group, we have a hydrogen atom and we also have an R group attached to it. Now, because of its cyrality it can have an L-form or D-form. Usually, L amino acids are incorporated into proteins.

Now, you understand, that these side chains, the R group can now differ in its size, it can differ in its shape, it can differ in its polarity. Do you understand each of these mean? But, there are 20 amino acids which have distinctive R groups with distinct properties of size, shape and polarity.

(Refer Slide Time: 09:28)



Amino acid side chains

Glycine: R=H
3-letter code: Gly
1-letter code: G

Proline
3-letter code: Pro
1-letter code: P

imino

We will consider the amino acids side chains by group in each case and what you have to remember is the 3 letter code of the amino acids along with the 1 letter code as well and obviously you have to remember what the side chain comprises. What we have listed here is glycine and proline. Glycine and proline are unique amino acids in ways that we see in a moment. Glycine is the simplest amino acid because the R group is just a hydrogen atom.

Now, this hydrogen atom makes the central carbon atom of glycine symmetric, because now it does not have 4 different groups attached to it. It has 2H atoms attached to it, which does not make it cyral anymore and this is the only such amino acid. So, if we look at the side chain of glycine, where the R group is written here and we know that this is where the R group is attached. This is R amino group and this is R carboxylic acid group.

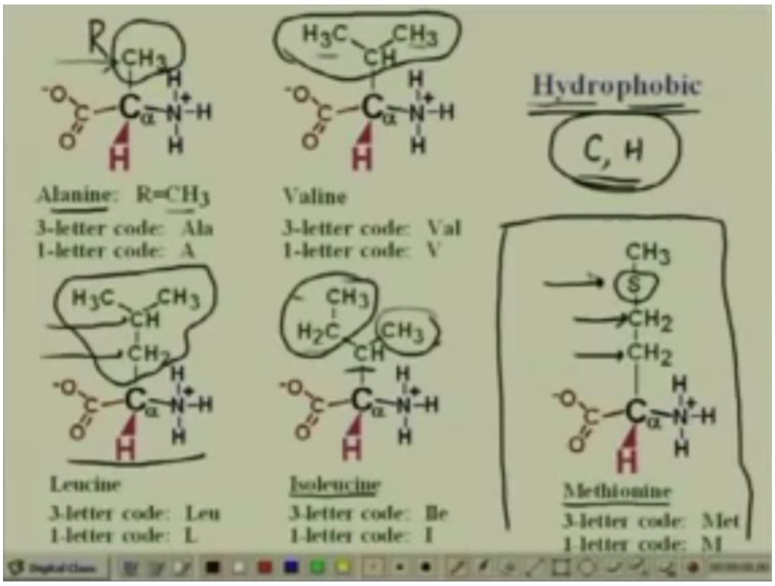
If you notice the way in which we have written the amino acid, we have written it with a NH3+ and a COO-. This is because, at physiological pH this OH is lost due to the PK value of carboxylic group, which we will be doing in the subsequent class, and this amino group is protonated which means that it has an additional hydrogen atom to it, making this nitrogen positively charged. This is what is called the Zwitterionic form of the amino acid.

It is written as the Zwitterionic form and is usually represented in this fashion. So, this is what we would call a zwitterion. It is represented in this fashion, because we would like to represent the amino acids as what they would be at the physiological pH. The next unique amino acid is Proline. Now, if you notice, it does not have a distinct R group attached to it, but the R group is actually linked up to what is called the amino group.

So, the side chain is actually CH2-CH2-CH2, but linked to the NH+, in this case. So, we have the C alpha with the hydrogen, with the carboxylic acid here, but instead of being an amino acid, proline is actually, what you would call an imino acid, because we do not have an amine group here, we have an imino. So we have an imino acid, where the CH2-CH2-CH2 bends on to itself to form, proline.

So, these are the 2 amino acids that are unique in their features. Glycine being just because it has the hydrogen atom there and it is acyral. Proline, it is an imino acid because the side chain bends back upon itself. And as I mentioned, when we represent the amino acids, we represent them in a zwitterionic form, where we write NH3+ and COO-, because this is how they would remain in or at the physiological normal solution.

(Refer Slide Time: 14:10)



Alanine: R=CH3
3-letter code: Ala
1-letter code: A

Valine
3-letter code: Val
1-letter code: V

Hydrophobic
C, H

Leucine
3-letter code: Leu
1-letter code: L

Isoleucine
3-letter code: Ile
1-letter code: I

Methionine
3-letter code: Met
1-letter code: M

If we go to the next group of amino acids, we consider now, hydrophobic amino acids. What are hydrophobic amino acids? Hydrophobic amino acids are those amino acids that are comprised mostly of carbon atoms and hydrogen atoms in their side chains. So, they would tend to be away from the solvent. Usually, the solvent being water or water based, they would be away from water. So, they would be hydrophobic, not liking to be in water.

Now, what are these specific side chains that we can have? The simplest one of these side chains is Alanine. Alanine, the 3 letter code is Ala and the one letter code is A. The side chain is the methyl group. So, this is what we would say is the R group and we you recognize again the zwitterionic representation of the amino acid. We then come to Valine. Valine is beta branched. We have CH-CH3-CH3. We then have Leucine with CH2-CH-CH3-CH3.

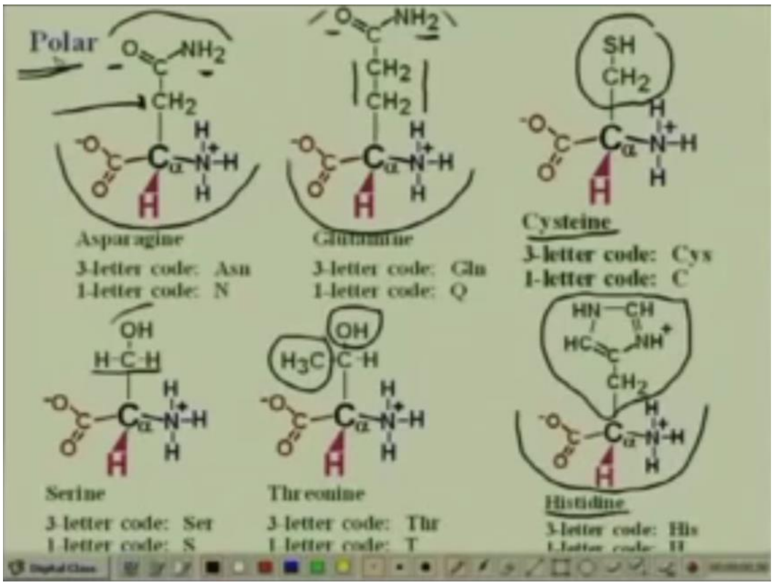
So, it is branched at the gamma atom. And the way these are represented is if this is the C alpha the next atom is the C beta, connected to the C beta are the C gamma atoms, which would be a unique representation of the amino acid Valine. If we look at Leucine, we would again have a unique representation considering that this is C alpha, the next one is the C beta, the next one is C gamma and attached to C gamma are 2 C deltas.

One is represented as Cd1 and other as Cd2. In the alanine side chain, we would just have a C beta carbon or Cb. So, these could actually be represented very clearly in a unique manner, where each amino acid this part being common, the side chains could be represented by the types of atoms that are attached to the C alpha. If we look at the isoleucine, we have CH-CH3-CH2-CH3. So, we have 2 beta carbon, one methyl group and one ethyl group attached to it. This is isoleucine.

Now, if you look at all the side chains that have been circled here, they are comprised of carbon and hydrogen only which make them hydrophobic in nature. Methionine can fall into this category as well, but it has a Sulphur atom and a methyl group attached to the Sulphur atom. So, we have the C alpha, we have C beta, C gamma and to the C gamma is attached this Sulphur atom and then we have a methyl group attached to the Sulphur.

Methionine, along with another amino acid, which we will consider in a moment, Cysteine are the 2 Sulphur containing amino acids and they could be grouped together in a group of their own or they could be considered in this group as well.

(Refer Slide Time: 18:36)



The next group that we will be considering are the polar amino acids. Now, what we mean by polar amino acids are those that have an oxygen or a nitrogen atom in their side chain and by virtue of having this oxygen and nitrogen in the side chain, these hetero atoms, they can participate in polar interactions not only amongst themselves but also within or with the solvent molecules.

So, they can participate in what we call, hydrogen bonding, which is extremely important in non-covalent interactions in proteins, which is what hold a protein folded together, the protein chain, the amino acid chain. We will see that in subsequent classes but the polar amino acids are the ones that are likely to interact with the solvent and in this interaction they can allow the oxygen and nitrogen atom that they are comprised of to interact with the solvent or within themselves to form a network and remain insolvent.

In contrast, the hydrophobic amino acids are unlikely to be on the surface of the protein. So, when we have a protein, which is actually a globular structure, which we will see in subsequent classes, we will see that there are certain side chains that prefer to be on the surface of the protein and there are certain side chains that prefer to be away from the solvent, which we have seen in the previous slide would be the once that are hydrophobic in nature.

Now, if you go back to look at the side chains that comprise this polar group of amino acids, what we have here is each of these have an oxygen or a nitrogen attached to it. We have, of course, the common part of each amino acid. In the asparagine side chain and the glutamine side chain, this being the common part again, we have amide groups. Amide groups are C double bond O NH2 groups.

So, this comprises the amide of asparagine, this comprises the amide of glutamine. The difference is that the glutamine chain is 1 carbon longer than the asparagine chain. So, what we have here is we have a CH2 that is the beta carbon attached to the alpha carbon followed by a gamma carbon that has attached to it an oxygen atom and a NH2. So, this amide group has a single beta carbon attached to the C alpha.

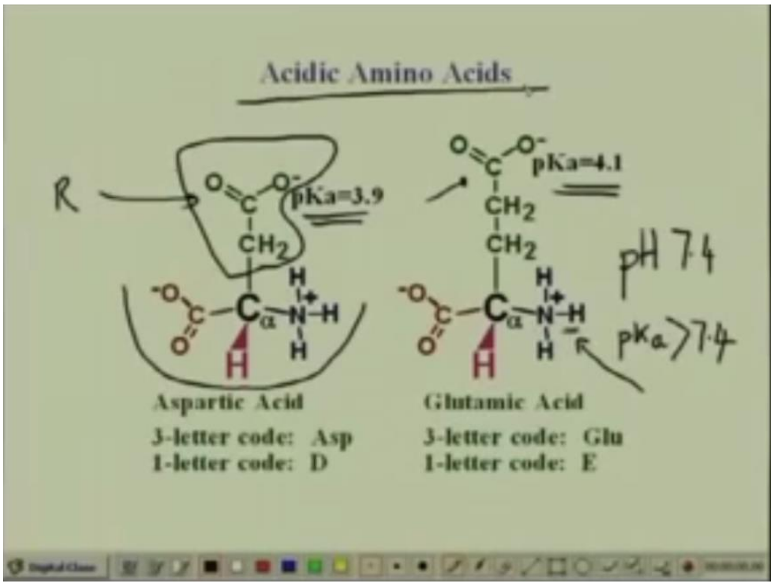
In glutamine, we have 2CH2 amides. We have a C alpha, a C beta, a C gamma, a C delta and to it is attached to C double bond O and the NH2 of the oxygen and the nitrogen. So, what can actually happen? Is this oxygen and nitrogen similarly in glutamine can participate in what is hydrogen bonding meaning that if we have a specific donor or an acceptor, then this could participate in hydrogen bonding not only with other amino acids but also with the solvents.

If we look at serine, serine is a small amino acid but a polar amino acid. The group that it has is CH2OH and it is this OH that can participate in hydrogen bonding. Threonine is the next amino acid. It has attached to the beta carbon, a CH3 and an OH. So, again it differs from serine. You see, how each of these are unique in their own way. The next amino acid is cysteine. Cysteine is another of the amino acids that has with it a Sulphur atom.

The other one that we saw on the previous slide was methionine that also had a Sulphur atom but to the Sulphur was attached a methyl group. Here, we have a hydrogen atom making this a thiol. So, we have a CH2-SH. Next is histidine. Histidine is a very important amino acid which will come across a lot when we consider the enzymes and enzyme mechanisms because of its specific polarity or specific properties of this side chain that is an imidazole group. So, again we have a common amino acid part here.

In histidine, we have 2 nitrogens in the side chain that is part of the imidazole ring. So, what you can see in this polar group of amino acids is all side chains that belong to this group contain what is called a heteroatom.

(Refer Slide Time: 24:54)



Acidic Amino Acids

pKa=3.9

pKa=4.1

Aspartic Acid
3-letter code: Asp
1-letter code: D

Glutamic Acid
3-letter code: Glu
1-letter code: E

The next group of side chains that we will be considering is acidic amino acids. We looked at asparagine and glutamine in our previous slide. What we found in asparagine was the C double bond NH2. Now, we know that an amide comes from a carboxylic acid. So, the asparagine comes from a specific carboxylic acid. Similarly, glutamine also comes from a carboxylic acid.

So, we group them into what are called acidic amino acids and we call the specific acids, aspartic acid, which gives rise to asparagine and glutamic acid, which gives rise to glutamine.

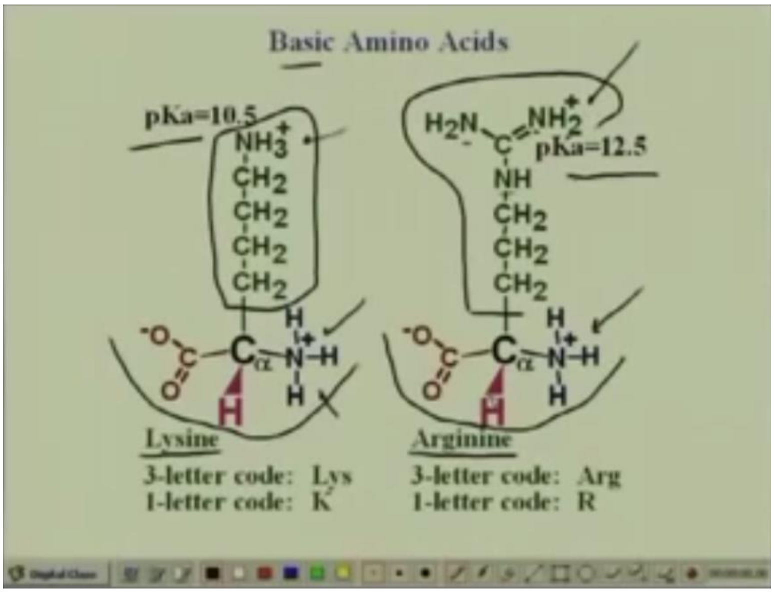
And now, what we have here is C double bond O, O minus. Now this is apart from the actual carboxylic acid that comprises the part that is common to all amino acids. This is part of the R group, the side chain. So, the side chain, in this case, also contains a C double bond O, O minus.

Similarly, the side chain in glutamic acid also has the C double bond O, O minus. But again, we have an additional CH2 in case of glutamic acid just like we had for glutamine. Now, what we have written here in addition is what is called a pKa value. We will learn more about this in our next class, but just for a preliminary information, if the pKa value is less than the pH of your solution, then your carboxylic acid is going to lose the proton.

Similarly, this has lost its proton but this has not. What does this mean, it means that the pKa value of this amino group is actually higher than the physiological pH, which is why it has still kept its proton attached to it. But, if you consider the physiological pH to be 7.4 it means that the pKa of this group is greater than 7.4 and we will see, how it is actually something close to between 9 and 10.

So, if we have the pKa value greater than 7.4, this going to remain protonated, but these carboxylic acid groups cannot remain protonated. These comprise what are known as acidic amino acids. If there are acidic amino acids, then it means that there are also basic amino acids.

(Refer Slide Time: 28:02)



Basic Amino Acids

Lysine
3-letter code: Lys
1-letter code: K

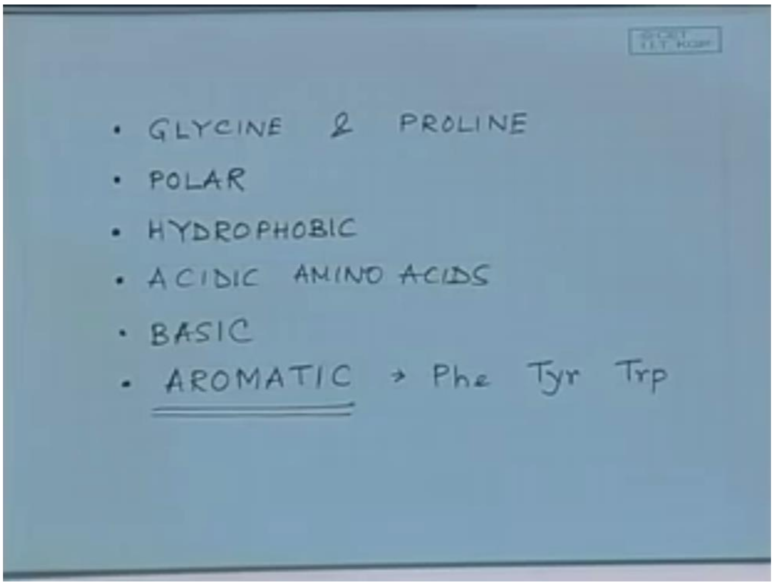
Arginine
3-letter code: Arg
1-letter code: R

So, what are these basic amino acids? These are Lysine and Arginine. Considering lysine and Arginine, let us look at the groups now. What are the side chains? This is the long side chain of lysine and this is the side chain of arginine. Now, if you look carefully at the side chains here, there are 2 pKa values written. What are these pKa values? These pKa values are actually greater than the physiological pH which is why they are still protonated.

So, we have protonated nitrogen because the physiological pH is 7.4 and we have not reached the pKa value, where this is going to lose its proton. The pKa value is where it is going to lose the proton, which as I said we will discuss in our subsequent classes. So, apart from the common part amino group that we have here, we have an additional amino group here because it is a basic amino acid.

In arginine, we have what is called guanidine group but that is a part of the side chain. It has 3 nitrogens here. And because of their properties especially the acidic and basic amino residues, these are the residues that prefer to be on the surface of the protein. So, if we look at the different structures of the amino acids that we have considered we have specific groupings.

(Refer Slide Time: 30:21)

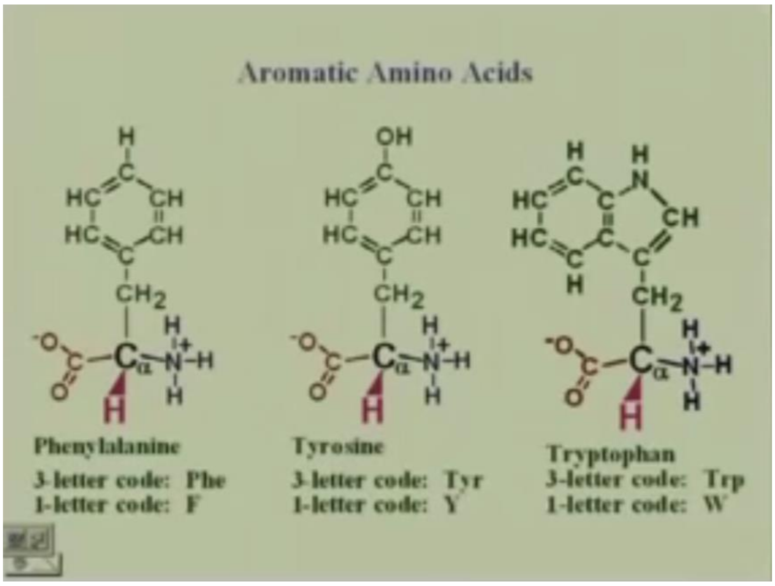


- GLYCINE & PROLINE
- POLAR
- HYDROPHOBIC
- ACIDIC AMINO ACIDS
- BASIC
- AROMATIC → Phe Tyr Trp

The groupings are the polar amino acids. So, we have in a group by itself glycine and proline because of the uniqueness in their properties. We have also other polar amino acids. We have hydrophobic amino acids. We have acidic amino acids. And we have basic amino acids. There is another group of amino acids called the aromatic amino acids. The aromatic amino acids are unique. As you can see, the name itself suggests their property.

They are aromatic in nature. Under this we have 3 amino acids. They are phenylalanine, tyrosine and tryptophan. Let us look at their side chains.

(Refer Slide Time: 31:51)



Aromatic Amino Acids

Phenylalanine
3-letter code: Phe
1-letter code: F

Tyrosine
3-letter code: Tyr
1-letter code: Y

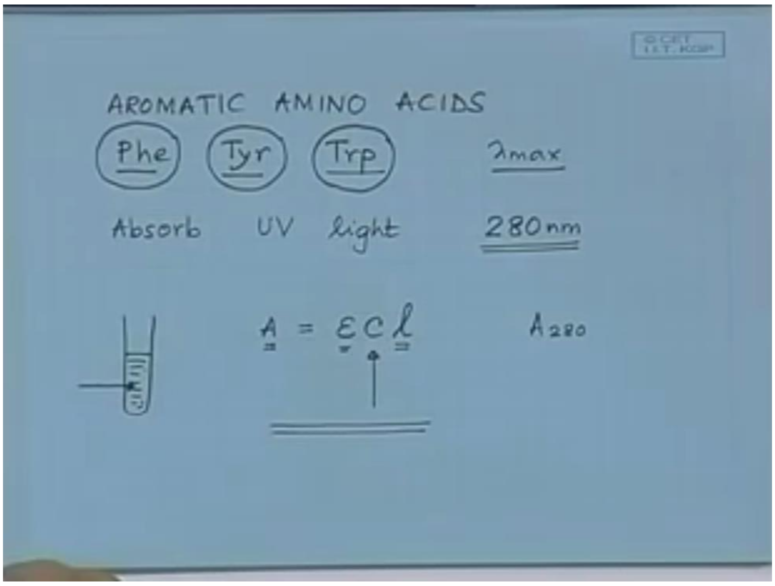
Tryptophan
3-letter code: Trp
1-letter code: W

These are the aromatic amino acids. This is phenylalanine. If you remember what alanine was, it was just a methyl group attached to the C alpha. In this case, 1 hydrogen has been replaced by a phenyl group. So, its name is phenylalanine. We, of course, have the common part of the amino acid here. The 3 letter code for phenylalanine is Phe and the 1 letter code is F. So, we have a phenyl group replacing one hydrogen of the alanine, phenylalanine.

So, this is aromatic in nature. We have tyrosine, which is similar to phenylalanine, the only difference being that this hydrogen replaced by an OH. So, tyrosine can actually also be involved in hydrogen bonding. In the grouping of amino acids, this could also therefore be put into a polar group. But, it is usually grouped under aromatic amino acids because of the phenyl ring here. So, we have a CH2 and we have a phenyl and an OH attached to this, which is called tyrosine.

In tryptophan, we have an indole ring attached to the CH2. This is a very bulky amino acid as you can see by the sheer size of it and it is quite rare in proteins in that it is not present to a very large extent in many proteins.

(Refer Slide Time: 34.01)



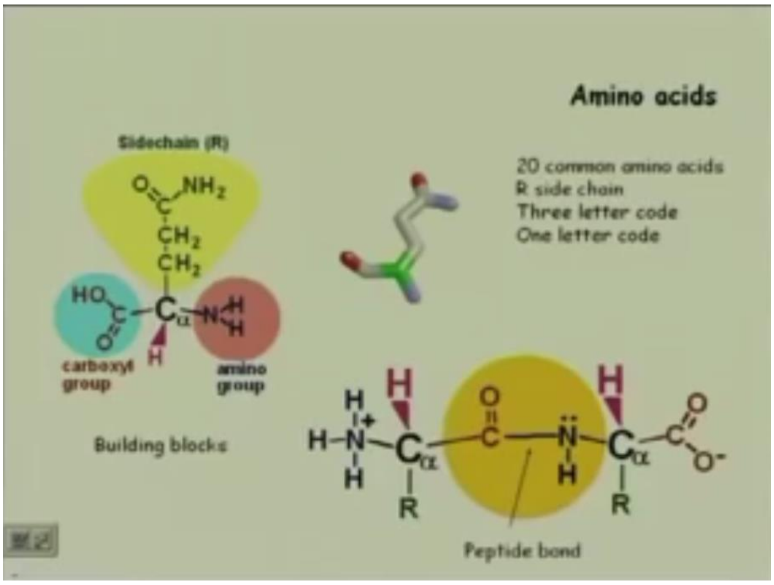
The unique properties of these aromatic amino acids which make the protein useful in an analytical way is all the aromatic amino acids that we have considered, the aromatic amino acids, which are, as I wrote previously, phenylalanine, tyrosine and tryptophan, each of these absorb UV light. They absorb ultra violet light. So, their presence in proteins can actually be utilized in this fashion.

What do we mean by that? They absorb UV light in the range, they have different lambda max values. But usually, we look at 280nm to identify a protein. So, if we have a solution that has a certain amount of protein in it, we can actually determine the amount of protein present in the solution, by a consideration of the number of phenylalanine,tyrosine and tryptophan that are present in the protein chain.

So, if we monitor or we find out the absorbents at 280nm, we know what is called the extinction of our protein and we know the length of the cell and we know the extinction coefficient of the protein. And we know what the absorbents is at 280nm, which is also represented as A280, we can determine the concentration of the protein. So, what we have is, the presence of these aromatic amino acids help us in determining whether our solution actually contains protein or not.

And we can also find out the content of the concentration of the protein in solution by virtue of their having phenylalanine, tyrosine and tryptophan. Of these, tryptophan has the highest extinction which means, that if we have a large number of tryptophan amino acids in the protein, you are going to have larger absorbents at 280nm. But the presence of the aromatic amino acids themselves will give the absorbents at 280nm which is how proteins are monitored in biochemistry laboratories.

(Refer Slide Time: 37:13)



Amino acids

20 common amino acids
R side chain
Three letter code
One letter code

The next thing that we are going to look into is the representation. When we consider the representation of amino acids as we have already seen, we have a carboxyl group, we have an amino group, we have a hydrogen which is common to all amino acids and we have a side chain R. Now, if you look at the side chain you recognize that this is an amide group. The 2 amide groups were asparagine and glutamine.

The glutamine side chain had 2CH2 groups. So, we have 2CH2 groups attached here. This is the stick representation, which you will look at in the next class where we asymmetric carbon in green and the others are grey, nitrogen in blue, oxygen in red and the other carbon atoms in grey. If we look at the linking of these amino acids, because when we are going to form proteins, these are actually the building blocks of protein.

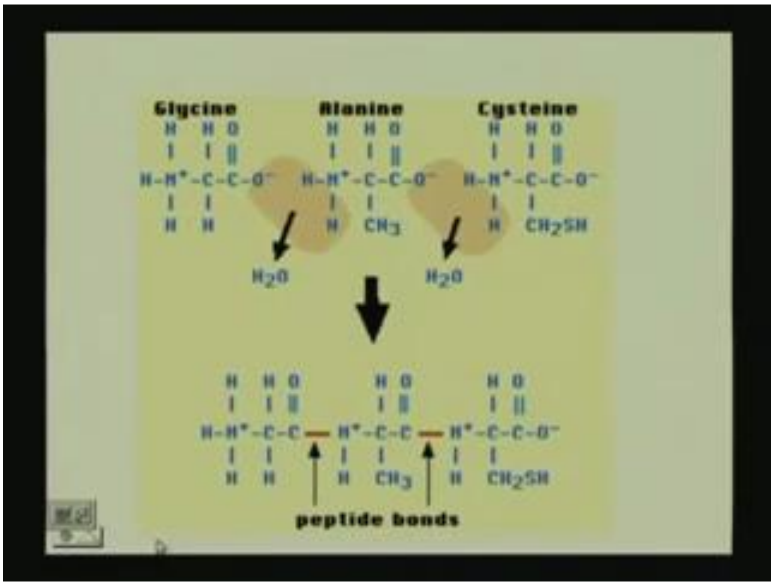
These building blocks have to be linked together. How are they linked together? They are linked together by what is called the peptide bond. If you look at the representation shown in left, it is not a zwitterionic representation because the proton of the carboxylic group is still attached to it and this is the NH2 group. In actual form, it would remain as NH3+ and COO-. But what we have here is two R groups.

Now, how do we have two R groups. In the first amino acid that we have on the left hand side, we have an amino terminal. This is called as dipeptide, because we have two amino acids linked by a peptide bond. In this peptide linkage, we have C double bond O and an NH. But, if we look at the original amino acid we are missing an OH from the carboxylic acid side and we are missing an H from the amino side.

What does that make? It makes H2O. So, when we are linking 2 amino acids by the elimination of H2O, we can form a peptide bond. We will look at into the features of peptide bonds once we consider the protein structure in general and the amino acids sequence. But, what we have to remember here is that when these amino acids are linked together on the left hand side you always have the N terminus and on the right hand side you always have the C terminus, because this the way the proteins are formed.

Here, this is the way they are synthesized. So, we have an amino terminal and we have a carboxylic acid terminal. And the first amino acid is the one that always has the NH3+ attached to it. And the last amino acid in a protein sequence or in a protein chain is the one that has the COO- attached to it. So, this is what we have a dipeptide linked by a peptide bond. There are certain features of the peptide bond unique to protein structure and peptide bonds that we will study in later classes.

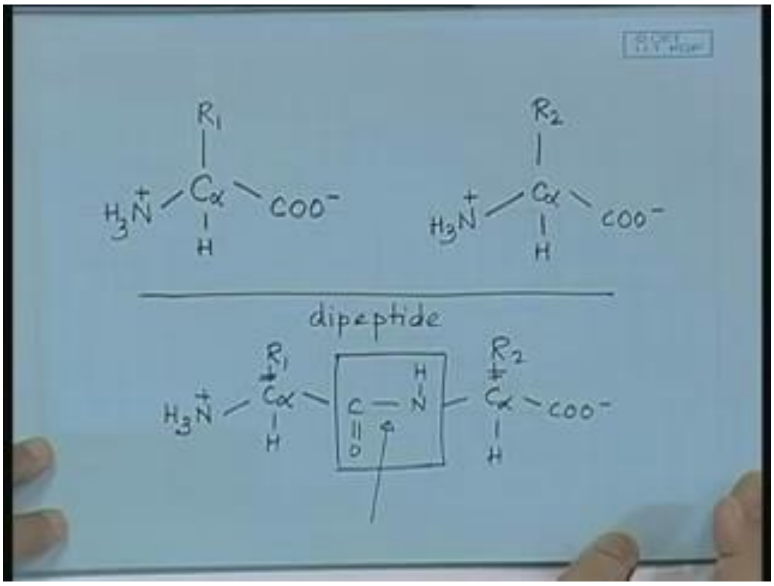
(Refer Slide Time: 41:11)



So, what do we essentially have. In this case, we have same Glycine. Now, what has happened to this glycine. In this case, we have a representation in this (()) (41: 23) form. The glycine is one, where we have, if you remember from our representation of the amino acids, we have a R group that is H. Here (Alanine), we have a R group that is CH3. In Cystine, we have a R group that is CH2SH.

So these are. For Glycine, we can actually we actually cannot distinguish, which is the R group because the hydrogen is present and the hydrogen is also the side chain. So, we have a NH3+, we have a COO and here, we are now forming and or eliminating water into forming a C double bond O and H. So, this is what has been formed. This should not be there. So, we have a C double bond O and H. A C double bond O and H. So, in the formation of our tri peptide, we have Glycine, Alanine, Cysteine.

(Refer Slide Time: 42:48)



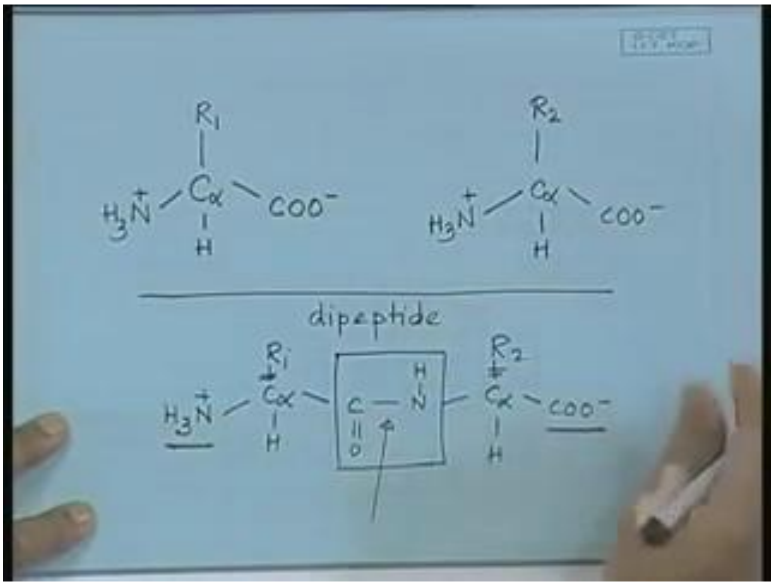
We can go on to form other peptide linkages. So, what we can have is, if we look at the basic structures of Amino Acids, we have a C Alpha, we have an NH3+, we have a COO-, we have an H and we have attached to this an R1 group. If you now link another amino acid, so you would have another NH3+ that would belong to our second amino acid, we have our COO- and we have our H and we have our R2.

So, now when we combine this two amino acids, to form what is called dipeptide, we would have linked these via a peptide bond and we would have our C double bond O, the NH coming from the second amino acid, the C alpha, the R2, the H and the COO-. So, this is what is our peptide linkage. This particular linkage is known as the peptide linkage. What do we have? We have linked now R1 and R2.

Now, in a representation of a protein, it is not very convenient to keep on writing all the atoms together. We know in the protein sequence, in the primary amino acids sequence, the proteins are the amino acids, the building blocks are linked together by the peptide bonds.

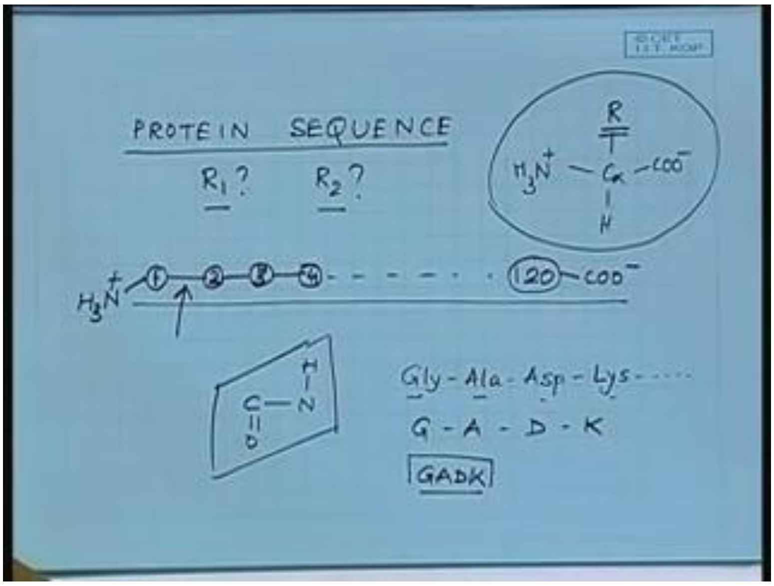
Now, since they are just linked by the peptide bonds, then it is not necessary to write what is common to all the amino acids, because these are certain features that we already know.

(Refer Slide Time: 45:17)



So, instead of writing in each case, the NH3+ or the COO-, we also know that the first amino acids is going to be linked with the NH3+ terminal and we know that the last one is going to be linked with the COO-. So, it is sufficient to write, instead of writing this in an elaborate fashion.

(Refer Slide Time: 45:32)



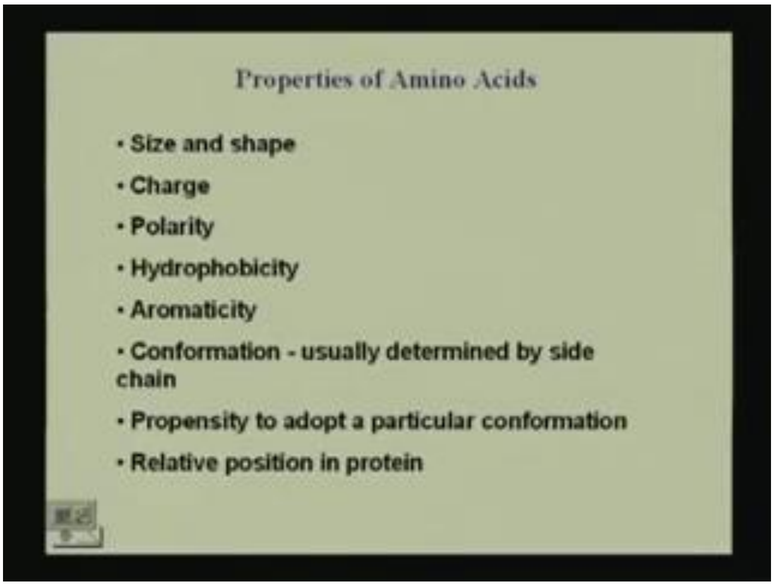
When we write a protein sequence as we will see later, to write what is called a protein sequence, all the information we actually need is what is R1? and what is R2? Because, we know that each amino acid looks the same. It is only the difference that we have in the R group. So, if we just know, what R1 and R2 are, when we have our protein sequence, we need to know how they are linked together. So, if this is amino acid 1, this is amino acid 2.

I know that this has to have NH3+ attached to it. 3, 4 and so on to say 120. I know this has to have COO- attached to it. And, I also know that these linkages are nothing, but peptide bonds. So, if this is the information that I know, I just need to know, what R1 is, what R2 is, what R3 is, what R4 is and so on and so forth. This is why we just write either the three letter code or the one letter code. In the three letter code, if this first one were the Glycine.

I would write Gly linked with Ala, Alanine linked with Acetic amino acid or aspartic acid linked with basic amino acid, Lysine and so on and so forth. Because, now, when I have glycine and Alanine and aspartic acid, I know, what the rest of the atoms are, because, I know the side chain of glycine, I know the side chain of Alanine and so on and so forth. If I write this in a one letter code, it would be G-A-D-K. So, if I just wrote GADK, you could write those structure of this tetropectide.

Similarly, when we consider a whole protein chain, in this case, say 120 amino acids, we could consider the whole protein chain. So, the difference is that we actually look at are in the properties of the amino acids side chains. What is so important?

(Refer Slide Time: 48:13)



Properties of Amino Acids

- Size and shape
- Charge
- Polarity
- Hydrophobicity
- Aromaticity
- Conformation - usually determined by side chain
- Propensity to adopt a particular conformation
- Relative position in protein

First is the size and shape of the amino acid that is extremely important in its accommodation. How is it going to be accommodated in the protein? We look at the charge on the protein. Is it acidic, is it basic? We look at the Polarity. Can it be involved in Hydrogen bonding? Is it a polar amino acid? Hydrophobicity. Where is this amino acid likely to be located? Is it going to be located in the center of the protein, because it likes to be away from the solvent or is it going to be likely to be on the surface of the protein?

But, we know that any hydrophobic amino acids would prefer to be in the core of the protein. Aromaticity. The aromatic amino acids that we consider Phenylalanine, tyrosine, tryptophane, these are important in imparting UV properties to amino acids, because these are the ones that absorb UV light and because of the absorption of UV lights, proteins can be detected in solution due to the presence of the aromatic amino acids.

And from what is known as the (()) (49:28), we can find out the concentration of the proteins provided, we know what the extinction coefficient is? The Conformation. Now, the Conformation, as we looked at, that is obviously usually determined by the side chain. We will see, since most of the side chains are linked by single bonds.

We will see how rotation about the side chains can actually bring about conformational changes to the amino acid orientations in the proteins and this change in conformation or the change in what is known as the (()) (50:05) angles will allow us to look at different properties of the amino acids, in the way, they interact with other amino acids. We also look at the propensity to adopt a particular conformation.

What does it mean? It means that if a protein were to have an amino acid that would likely form what is known as a helix or be part of the helix. That is what, we call the propensity. Is it likely to be in a helix, is it likely to be in a sheet? Now, these terms will be much more apparent, will be much more clearer, as we go to our subsequent classes. So, what we actually did learn today is the different types of the amino acids, the different groupings of the amino acids and the important properties of the side chains of the amino acids.

We consider that the central carbon atom, the asymmetric carbon atom, which is also known as the alpha carbon atom has linked to it, four different groups. These are hydrogen atom, an amino group, a carboxylic group and a side chain that is represented as R. Now, each of the R groups, there are twenty such common amino acids. The twenty common amino acids have twenty different R groups that differ largely in their properties.

We have different types of amino acids. The amino acids are, as I said, grouped into the type of R group or the type of side chain that they have attached to them. We have unique amino acids – Glycine and Proline. We have hydrophobic amino acids, we have polar amino acids, we have acetic and basic amino acids and we have aromatic amino acids. Some of these contain Sulphur, two of them methionine and cysteine.

There are others that have hetero atoms in them, oxygen and nitrogen. And there are others that have the side chains comprised entirely of carbon and hydrogen making them hydrophobic in nature. And lastly, we look at the overall properties that we can consider together the size, the shape, the charge, the polarity, the hydrophobicity and the aromaticity. And all of these will actually determine what the property of the protein is in general, because, we know that these amino acids are linked by peptide bonds.

In the linkage of the peptide bonds, we are bringing different types of amino acids together to form our protein sequence or amino acids sequence, which you will study in classes later on, where we will be looking at protein structures in detail. We looked at the peptide bond and we saw how the amino acids were linked together by the peptide bond and how we can actually represent the protein sequence, by just writing either the three letter code, one after the other or the one letter code, one after the other.

Because we know that the first amino acid is going to be the N terminus and the last amino acid is going to be the C terminus, which means the first amino acid is going to have the NH3+ attached to it. And the last amino acid is going to have COO- attached to it, making up the protein chain. We will learn in later classes, how the protein actually folds and how the hydrophobic amino acids tend to remain in the center of the protein. Thank you.