Fundamentals of Protein Chemistry
Prof. Swagata Dasgupta
Department of Chemistry
Indian Institute of Technology, Kharagpur

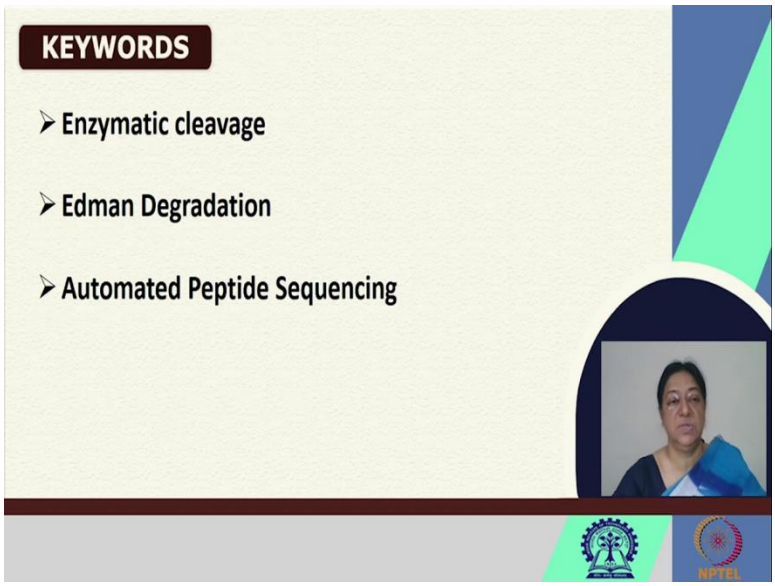
Module - 12
Special Topics in Protein Chemistry
Lecture - 57
Enzyme Action and Proteolytic Cleavage

In this lecture on our module on special topics in protein chemistry, we are going to be looking at enzyme action and proteolytic cleavage.

(Refer Slide Time: 00:24)

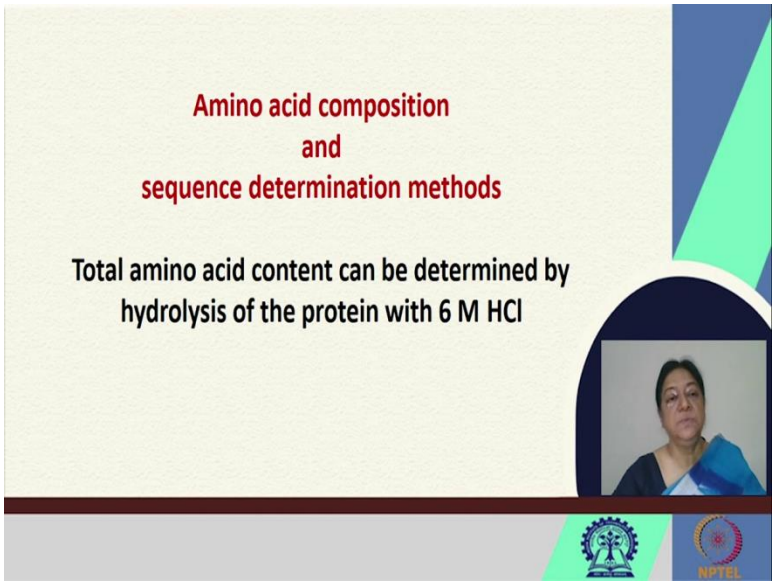
From this we will be looking at the studies of proteases, Edman degradation, what we mean by enzymatic cleavage and specific examples of sequence determination. In the previous modules of 6 and 7 we have looked at enzymes, enzyme classes, enzyme mechanisms, kinetics and inhibition in detail. However what we will look at here, is how to fragment a specific polypeptide chain and by that, determine the polypeptide sequence associated.

(Refer Slide Time: 01:00)



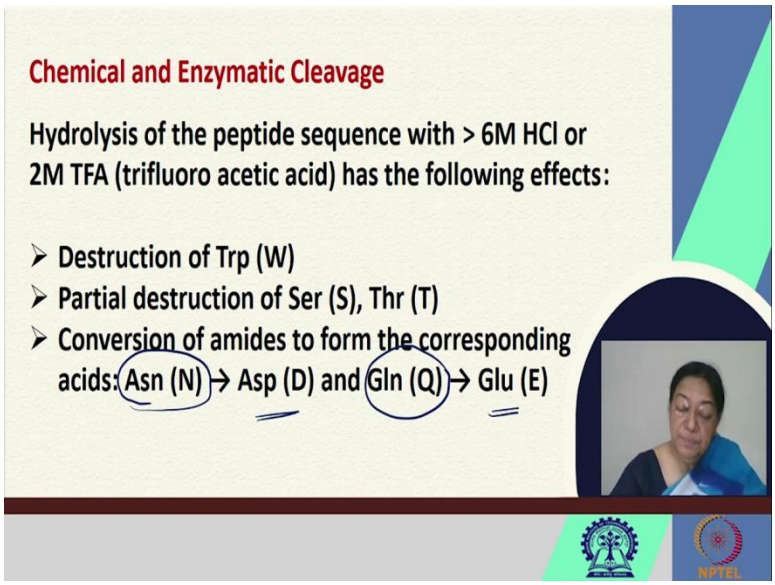
So we will be looking at enzymatic cleavage and also some degradation methods, into what we would want to do in automated peptide sequencing.

(Refer Slide Time: 01:12)



The amino acid composition and the sequence determination methods are important. The amino acid composition tells us how many of each amino acid types are present and the sequence determination methods tell us exactly in what sequence those amino acids are. So, if we want to determine the total amino acid content, this can be done by complete hydrolysis of the protein and identification of the specific amino acid residues present.

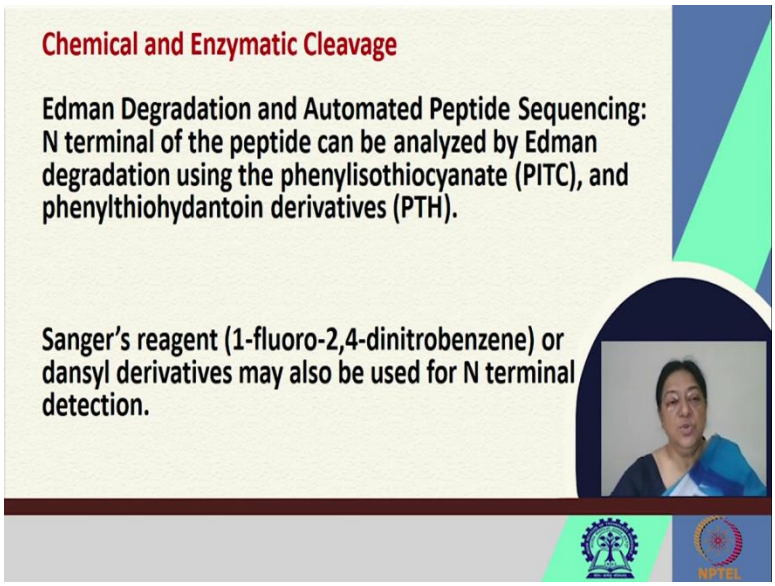
(Refer Slide Time: 01:53)



If we want to look at the methods of chemical and enzymatic cleavage, we can do hydrolysis of the peptide sequence with greater than 6 molar or greater than 2 molar TFA, trifluoroacetic acid and we have the following effects associated with this, which we also have to be aware of.

This could destruct our tryptophan residue altogether; this could result in the partial destruction of the serine and the threonine and because of the harsh methodology used, we could have a conversion of the amide and amino acid residues, that is asparagine and glutamine, to their respective acids. Based on this information, we would then have to construct our corresponding peptide.

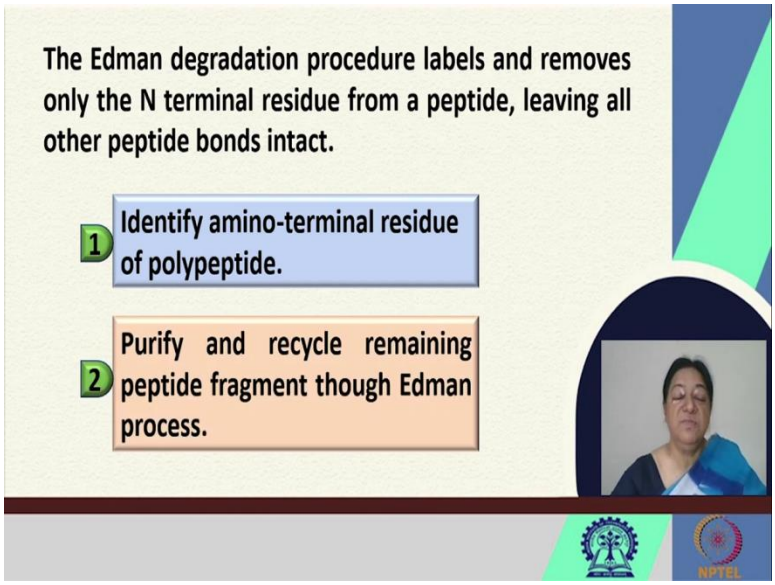
(Refer Slide Time: 02:49)



If we look at chemical and enzymatic cleavage, we can look at Edman degradation in automated peptide sequences. In this case, the N terminal of the peptide can be analyzed by Edman

degradation, using what is called the PITC and PTH derivatives and we have Sanger's reagent also, that is 1-fluoro-2, 4-dinitrobenzene or dansyl or dabsyl derivatives that also can be used for N terminal detection.

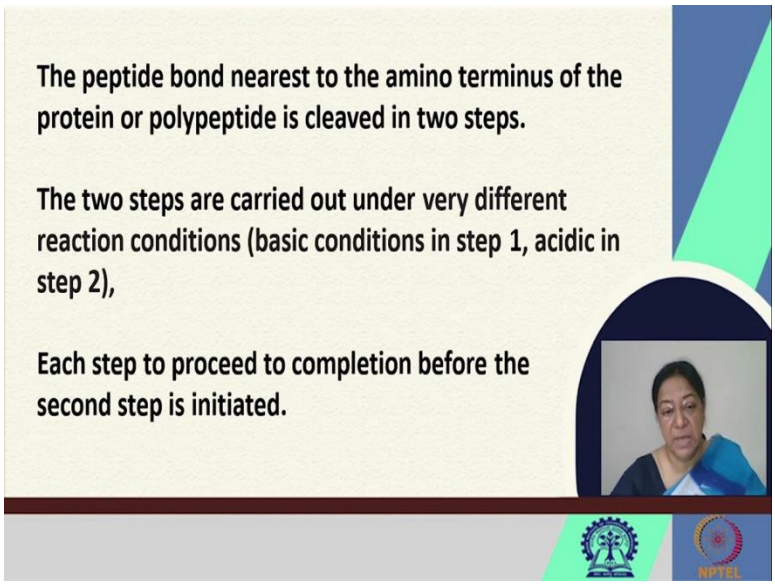
(Refer Slide Time: 03:24)



Without going into much detail of this, we will just have a brief overview of how these work. The Edman degradation procedure, it labels and removes only the N terminal residue from the peptide and leaves the other peptide bonding.

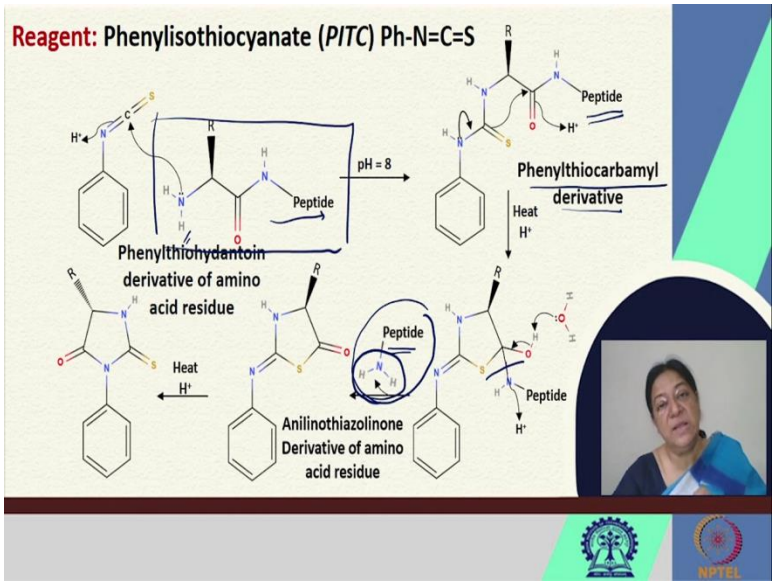
It can identify the N terminal residue of the polypeptide and the others can be purified and recycled for the Edman process and from that we can then determine the sequence of the peptide because each time we will have a new N terminal, as the previous N terminal has been removed by the Edman degradation procedure.

(Refer Slide Time: 03:59)



In this experiment the peptide bond nearest to the amino terminus of the protein, is cleaved in two steps. They are carried out under very different reaction conditions; 1 is basic and one is acidic and each step has to proceed to completion, before the 2nd step can be initiated.

(Refer Slide Time: 04:17)



So this [refer to slide] is the reagent that is used, PITC and here is our peptide and here is our N-terminus and the rest of the peptide is along this direction. We know that the N terminal residue is the first residue in the protein. The reagent is phenylisothiocyanate, we have a specific reaction under the basic conditions and the formation of an adduct, that is the phenylthiocarbamyl derivative.

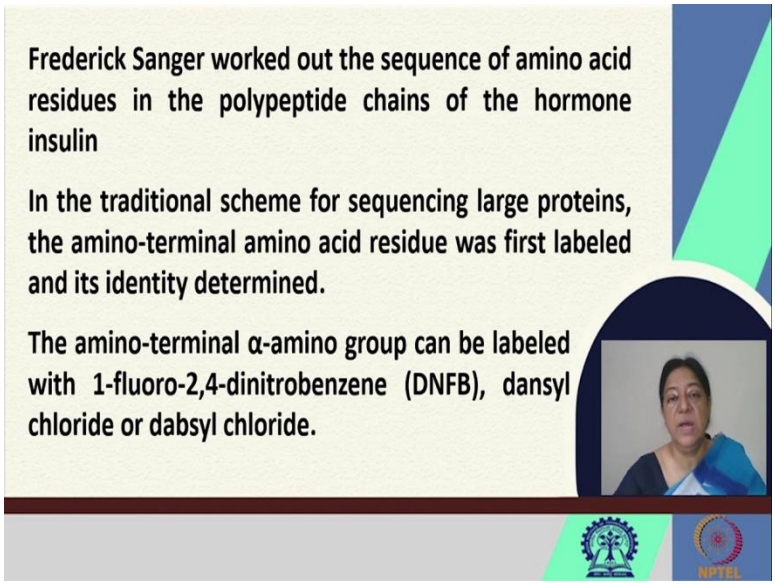
From that we have [refer to slide] a specific reaction that is going to give us under conditions of heat, a peptide that is attached to our PITC. So we have the PI derivative here, we have

hydrolysis and we have the connection in a manner that is going to be a resultant peptide, that is going to be cleaved from this specific residue attached to our phenylthiohydantoin derivative.

So what happens here, is initially we have our overall peptide given by this and we are now concerned with the specific amino acid residue, that is the first amino acid residue in this sequence. In the process, what happens is under these basic conditions the phenylthiocarbamyl derivatives is formed and once this is formed there is an adduct, with the peptide connected to it.

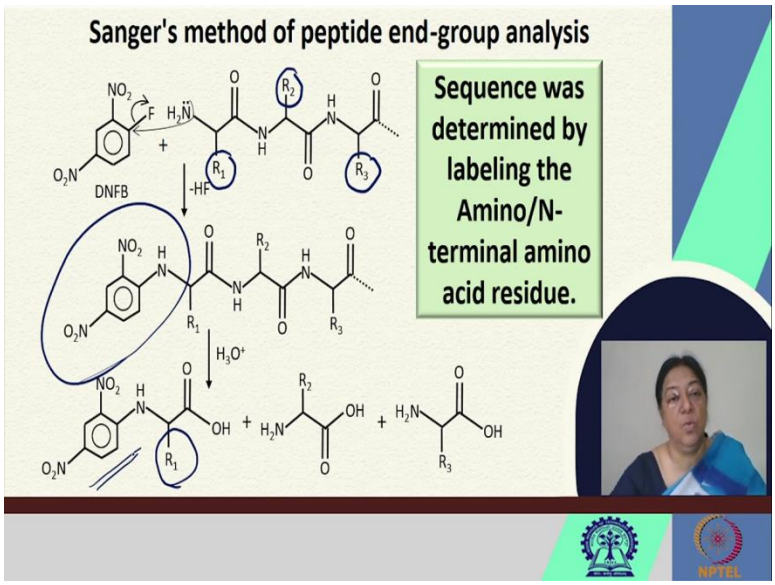
This is then hydrolyzed, where the peptide containing part with the new amino terminal is cleaved off and we have the derivative of the amino acid residue, that was the first one in the sequence. This cycle can go on and we can get the sequence, the N terminal or we can get amino acid number 1, then 2, then 3 and so on and so forth.

(Refer Slide Time: 06:38)



In the Sanger's method, there is a sequence of amino acid residues in the polypeptide chain. This was first worked out with insulin and here the amino terminal is also labeled and it is identity determined. This can be labeled with DNFB, dansyl chloride or dabsyl chloride which is going to give us the first amino acid in our sequence.

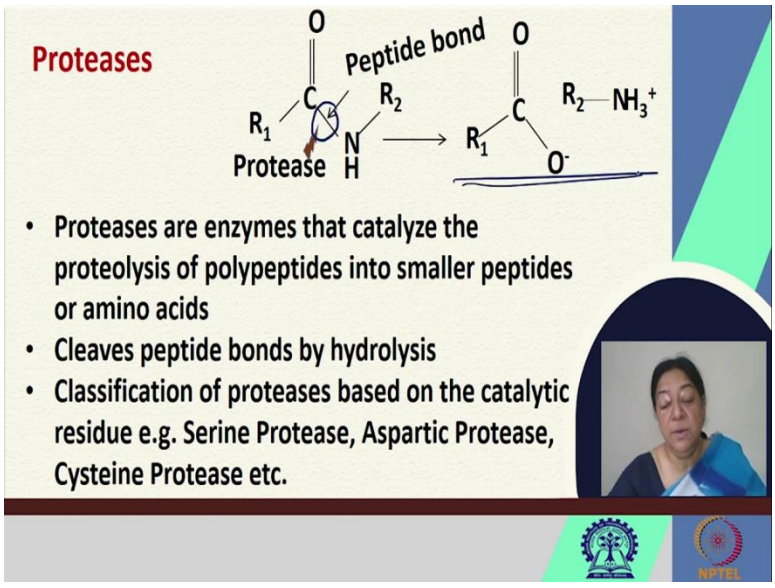
(Refer Slide Time: 07:02)



When we are looking at peptide end-group analysis, the sequence is determined by labeling the N-terminus of the residue. So here [refer to slide] is our peptide, where we have our specific amino acids given by their side chains; R_1, R_2, R_3. The reagent in this case is our DNFB, dinitrofluorobenzene and we have a specific reaction go on here, where we have this connected to our polypeptide chain now.

This will then be hydrolyzed. After the hydrolysis, we will have a connection with the first amino acid residue, with our derivative, that can be monitored and from this, we can determine what is R_1.

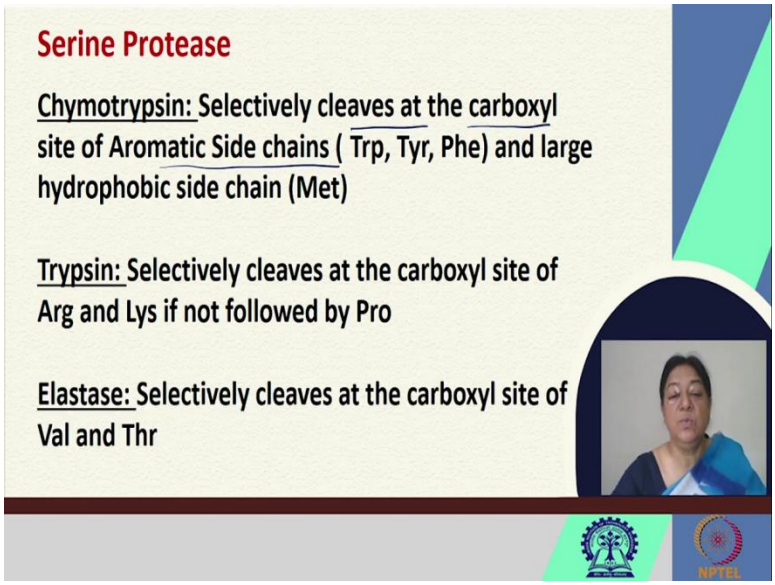
(Refer Slide Time: 07:53)



So what other methods are available? We have the proteases. Proteases is a topic that we covered when we looked at enzyme mechanisms, in terms of the workings of chymotrypsin and we know

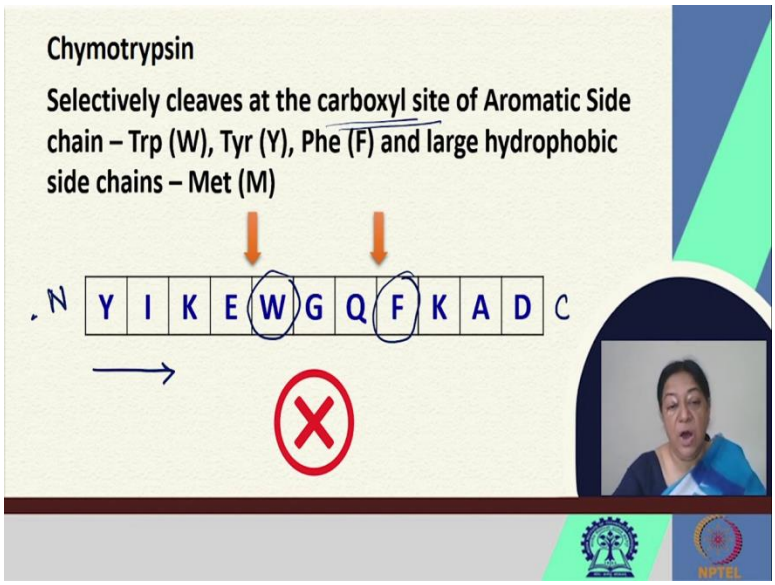
that this also cleaves peptide bonds by hydrolysis, in an enzymatic fashion and we have a classification based on the catalytic residue, like we looked at serine proteases, aspartic proteases or cysteine proteases, depending upon what is present in the active site. So the specific enzymatic cleavage that we now look at is the proteases acting on the peptide bond, resulting in the cleavage of the peptide bond.

(Refer Slide Time: 08:37)



This means that if we have a serine protease and we have our specific protein sequence, our polypeptide sequence, the chymotrypsin is going to fragment it in a fashion that we will selectively cleave at the carboxyl end of aromatic side chains. We have trypsin that will selectively cleave at the carboxyl site of arginine and lysine, provided it is not followed by a proline. Similarly, we can have elastase that would cleave at the carboxyl site of valine and threonine.

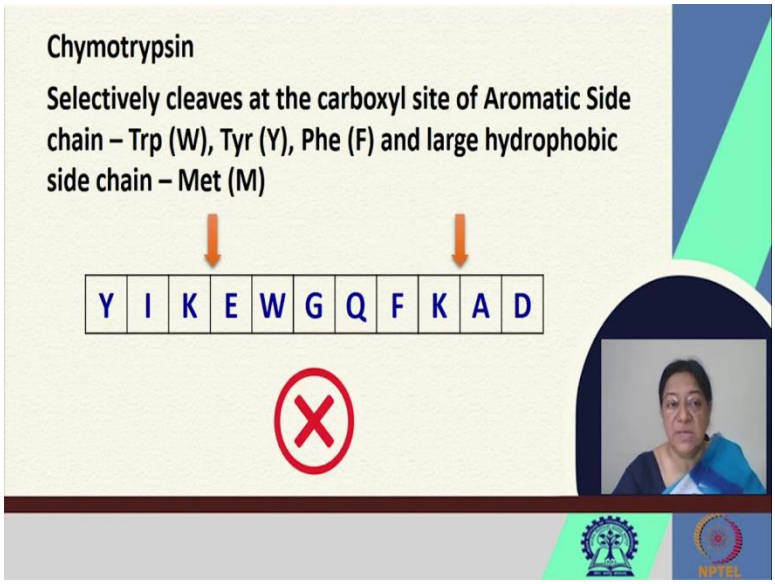
(Refer Slide Time: 09:11)



Let us look at how this action occurs. We have looked at the mechanistic details. We are just going to now look at how we can determine the overall polypeptide chain, out of the identity of a peptide in the specific order in which the amino acids are, by the usage of enzymatic cleavage and chemical cleavage, to give our peptide sequence. We know that this selectively cleaves at the carboxyl side of aromatic side chain.

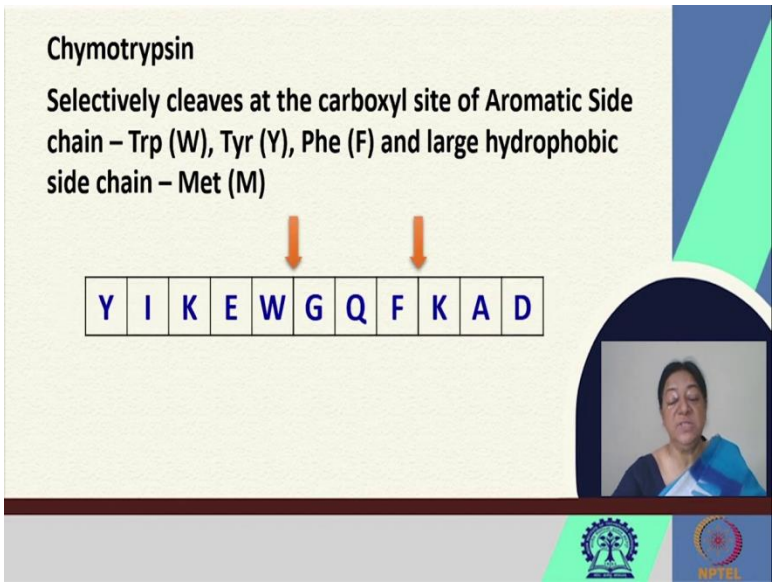
So given that this[refer to slide] is our peptide sequence. We now realize when we look at this, that we have two aromatic side chains here, but this cleaves at the carboxyl site. So, these arrows indicate the N site; because we know that our polypeptide sequence is in this fashion, so this is the N-terminus and this is the carboxyl terminus. So the N-terminus cleavage site is not correct.

(Refer Slide Time: 10:19)



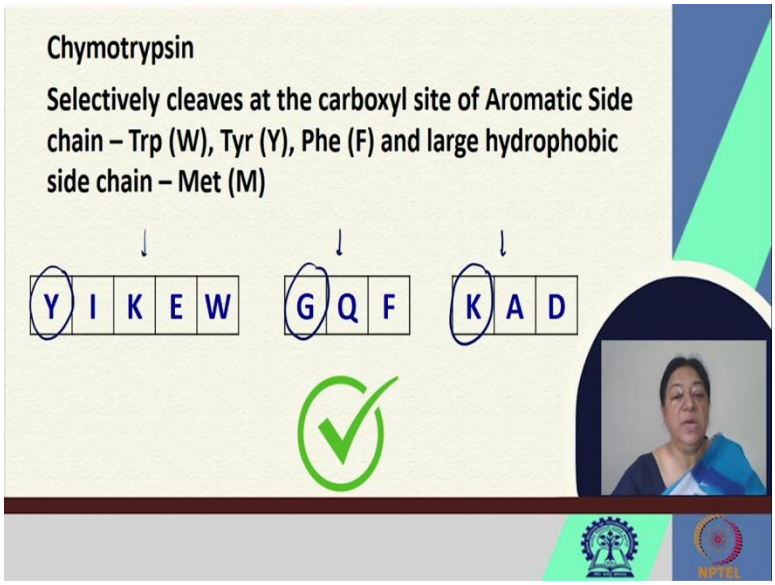
However when if we move them over again, we do not have the correct amino acids in this case; so still there will be no cleavage.

(Refer Slide Time: 10:28)



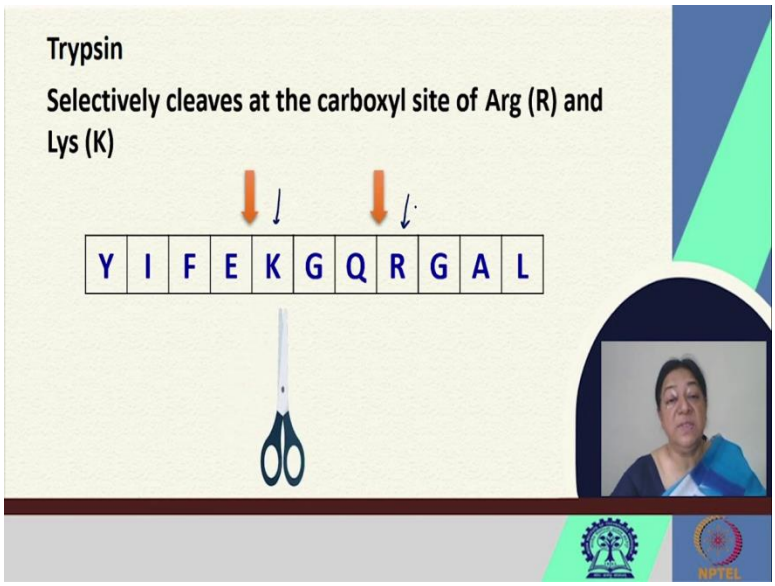
But if we move it to the carboxyl side of the aromatic side chains, we have two of them, the tryptophan and the phenylalanine; we will result in this cleavage.

(Refer Slide Time: 10:37)



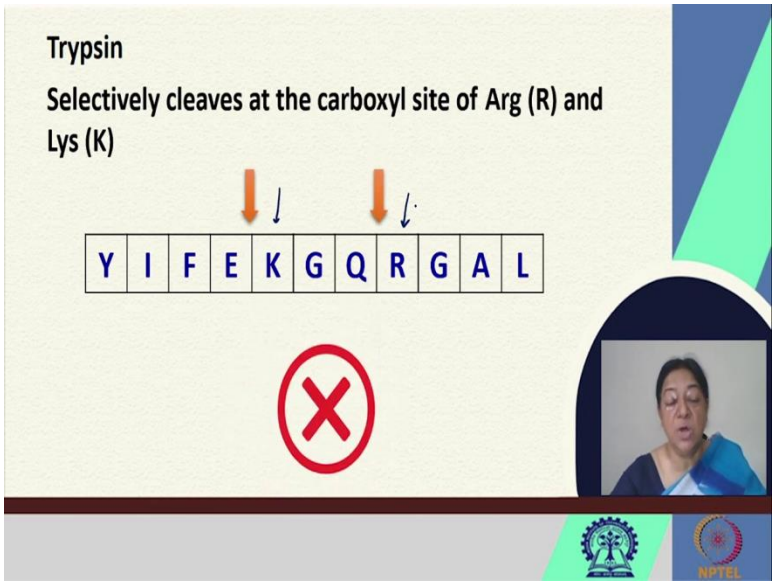
So then, we can now follow an Edman degradation or a Sanger method for the peptide fragments and determine the N-terminus of each of these fragments, into trying to figure out what the original chain was. This is a methodology where we can use chymotrypsin.

(Refer Slide Time: 11:02)



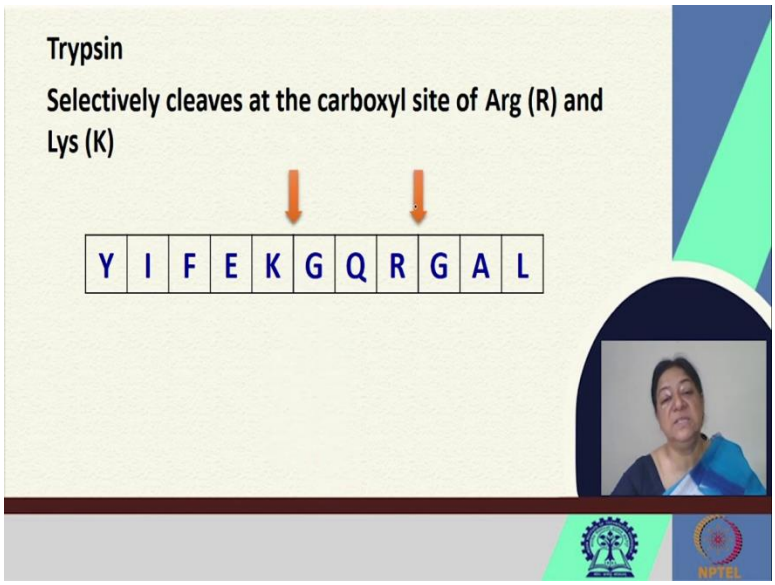
Similarly we can use trypsin. In the trypsin method, we find [refer to slide] out that this selectively cleaves at the carboxyl site of arginine and lysine. So in this case, if we try and identify where we have our lysine and arginine, then we would know that this will cleave at the carboxyl side.

(Refer Slide Time: 11:20)



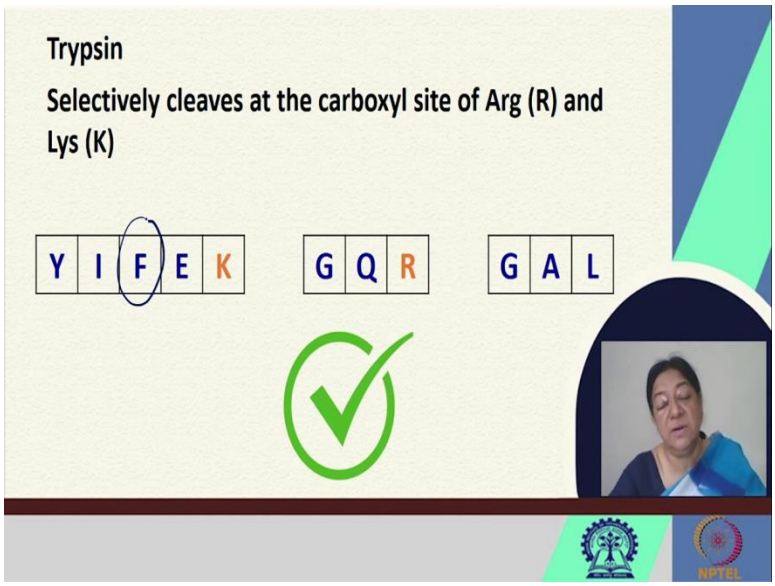
So this is not the carboxyl side, but it is the N terminal side of the specific amino acid.

(Refer Slide Time: 11:28)



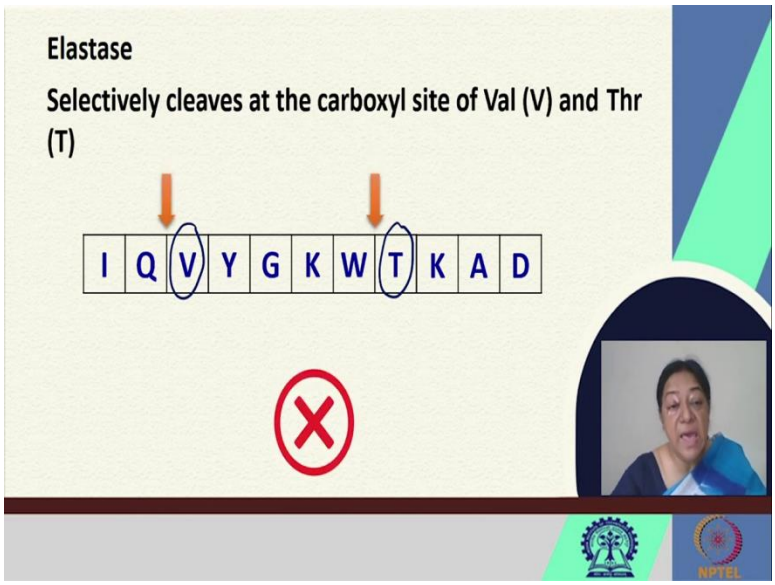
Thus we would have to look at the specific cleavage sites, giving us our specific fragments in this case.

(Refer Slide Time: 11:31)



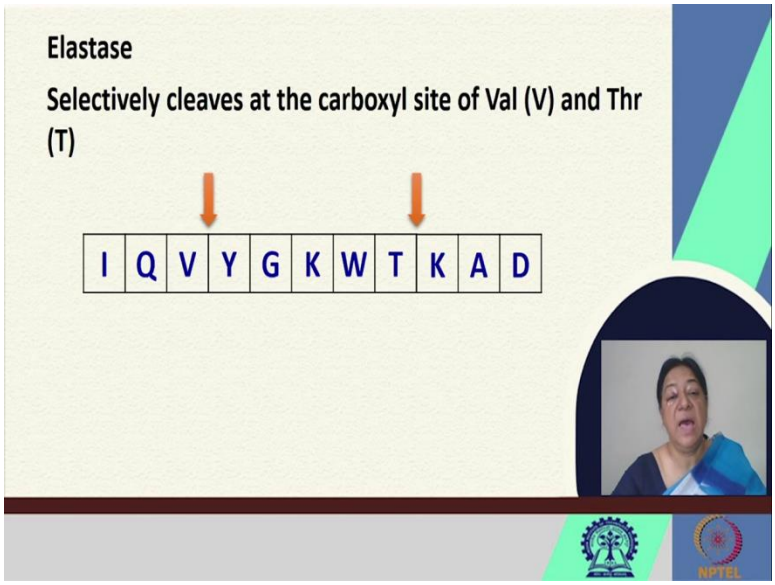
Again we can look [refer to slide] at the identity of the N-terminus or now that we have a phenylalanine, we can look at this specific fragment and look at another enzymatic cleavage mechanism.

(Refer Slide Time: 11:47)

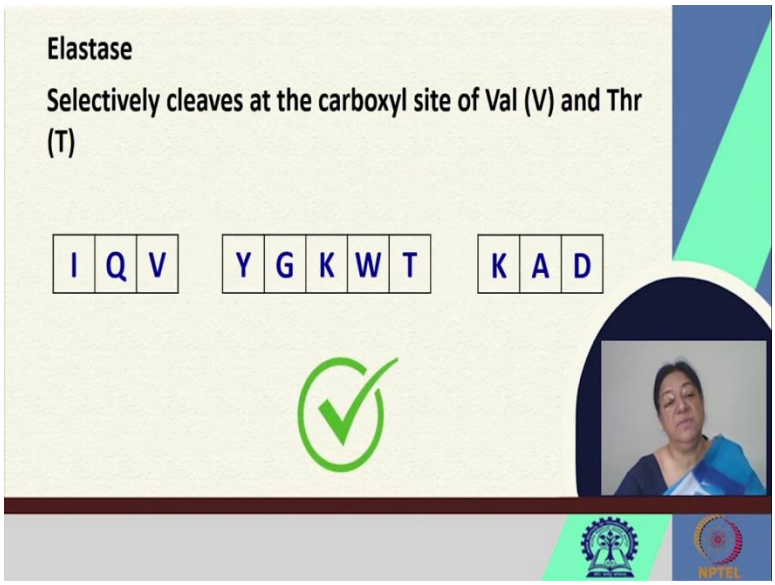


Similarly with elastase, it cleaves on the carboxyl side of valine and threonine. So we would have to know if we have a valine threonine present, then we can actually look at the specific C-terminus cleavage and find out that we have our cleavage.

(Refer Slide Time: 12:04)

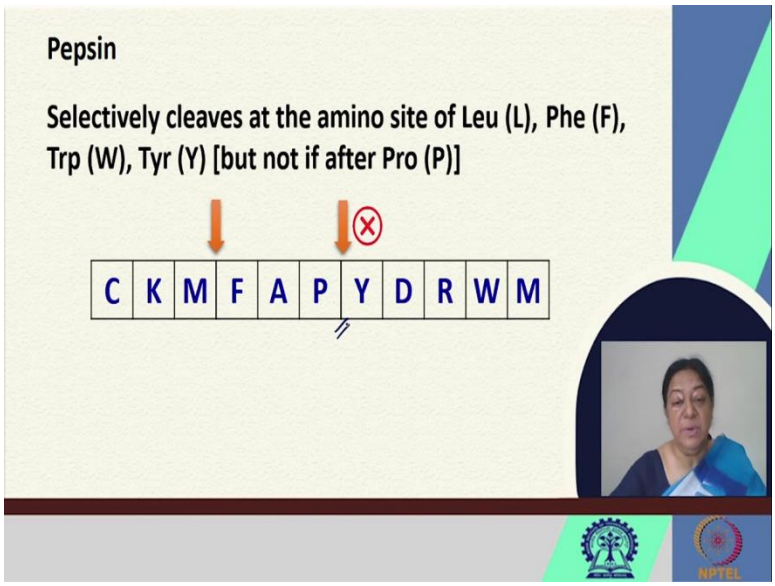


(Refer Slide Time: 12:07)



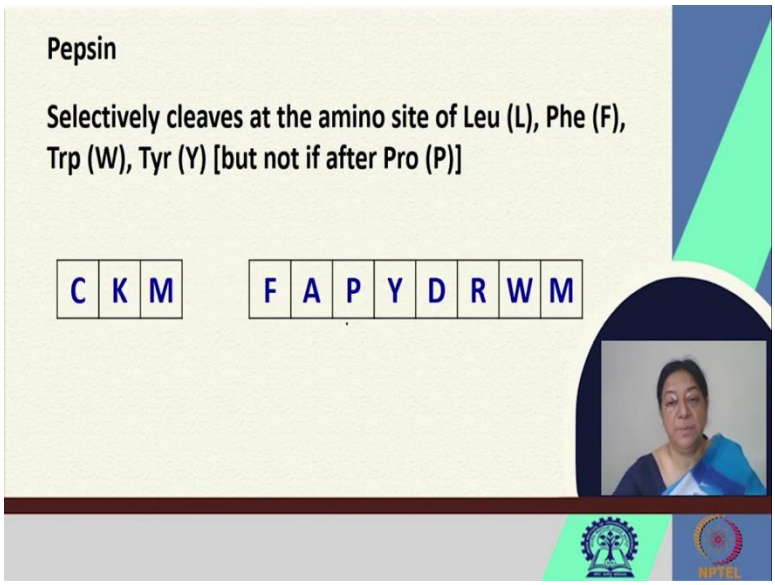
Now if we look at the same peptide and use these different enzymatic methods, we will then be able to piece together our results to find out the overall peptide sequence which we will see.

(Refer Slide Time: 12:20)



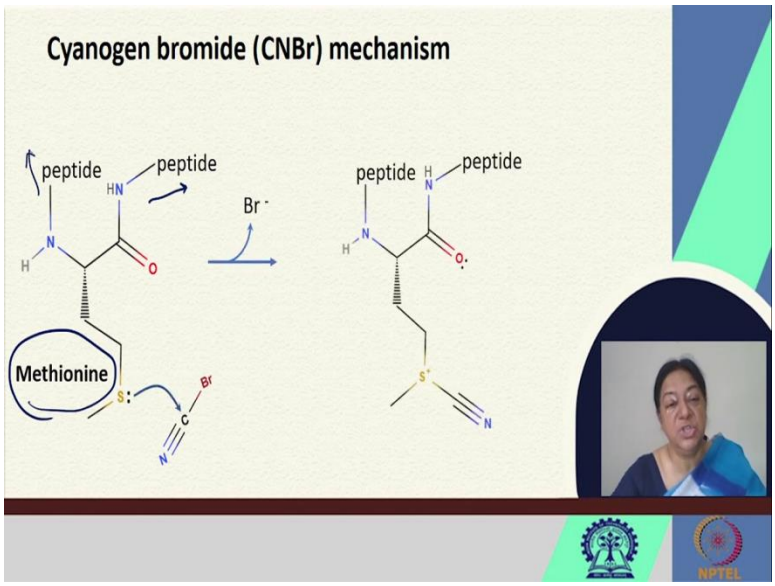
In the case of pepsin, this also selectively cleaves at the amino site of leucine, phenylalanine, tryptophan and tyrosine, provided it is not after a proline. If this is after a proline, there will be no cleavage here [refer to slide], however we will have these fragments formed.

(Refer Slide Time: 12:39)



So this is the methodology by which we can actually determine the ideas associated with the experiments.

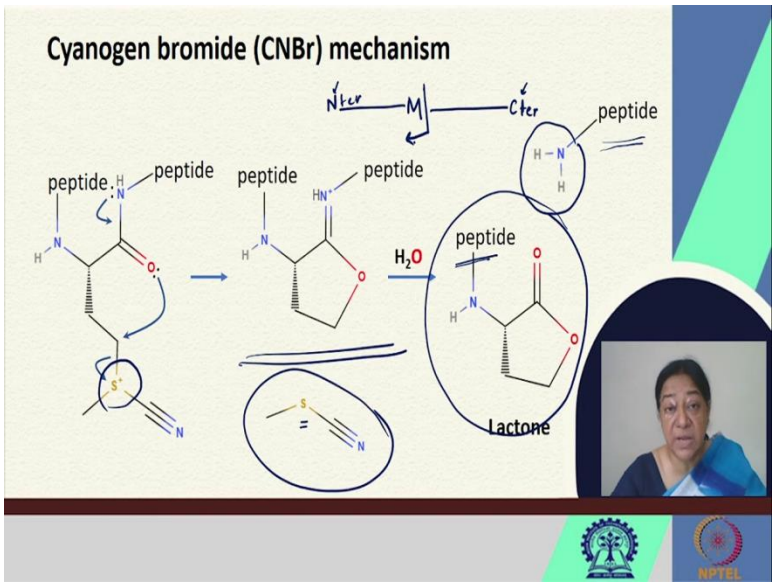
(Refer Slide Time: 12:45)



So we have another mechanism, this is a chemical cleavage. A chemical cleavage tells us this works on methionine, in the formation of what is called a lactose. Now we have our peptides, so this [refer to slide] is the N terminal side of our protein polypeptide chain and this is the C-terminus side of the polypeptide chain and we know that we have our specific peptide bond and this is the methionine side chain.

In this case, we have the CNBr, the chemical cleavage that is going to occur in this case, where we have a specific reaction occur in the elimination of the bromine ion and we have an adduct formation.

(Refer Slide Time: 13:32)

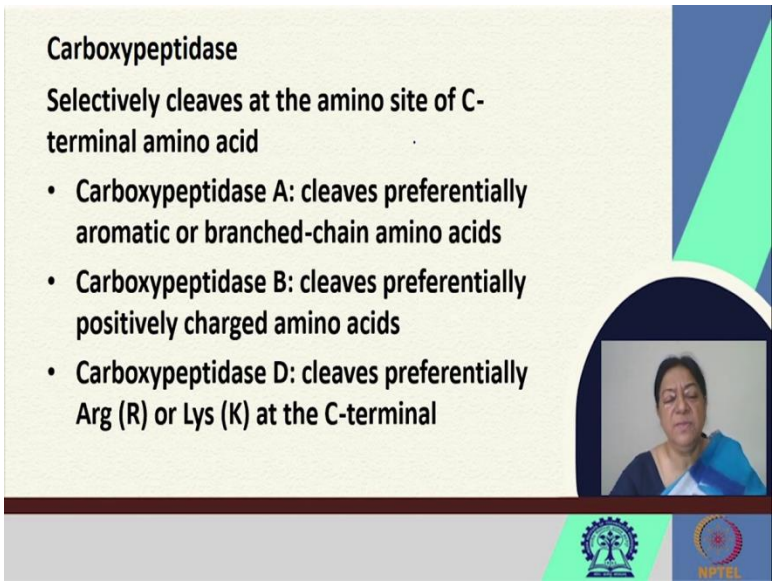


In this case when we have hydrolysis, we have a cleavage form here [refer to slide]. This is where we have the derivative from the methionine, the sulfur comes from the methionine in this case and with hydrolysis, we actually have a lactone formation with this fragment of the peptide and the rest of the fragment of the peptide, the polypeptide chain with the new amino terminus, has been cleaved.

CNBr will work on a methionine residue. So, this [refer to slide] is the N-terminus and this is the C-terminus of our polypeptide chain.

If there happens to be a methionine, we will have a lactone formation at this position and depending upon how many methionines we have, we will have lactone formations at each of these positions, resulting in specific number of fragments, depending upon how many methionines are present. Similarly for the enzyme case, we will be looking at the cleavage fragments from a tryptic digester, as it is called.

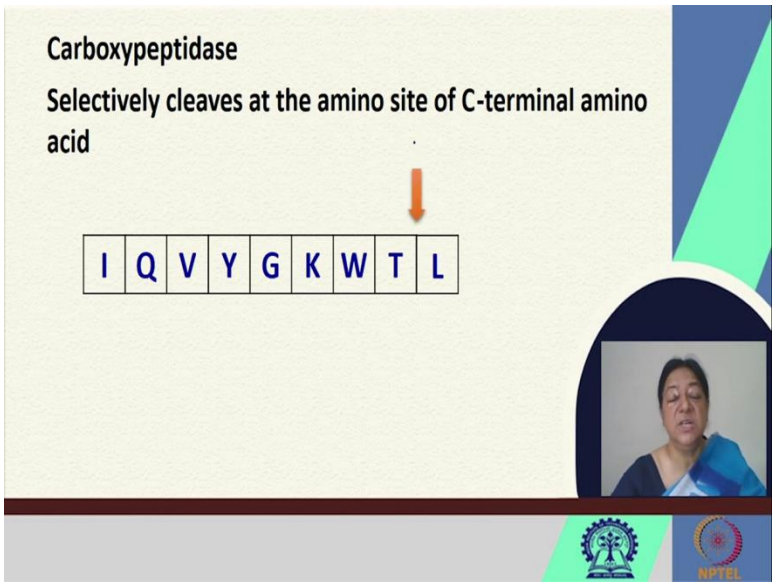
(Refer Slide Time: 14:57)



Carboxypeptidase works on another hand, where it cleaves at the amino site of the C-terminal amino acid. So we have an identification for the N-terminus using the Edman degradation method or the Sanger method. Now we want to know what is at the C-terminal amino acid.

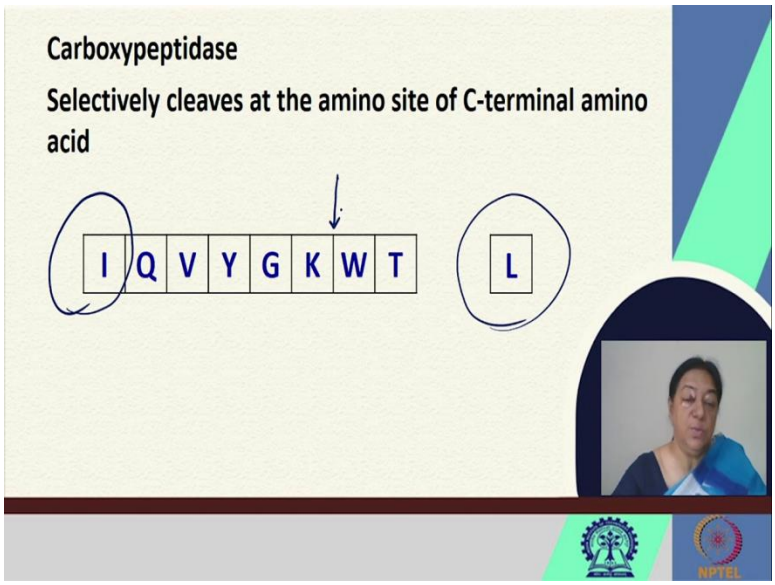
So, in this case we can use an enzymatic cleavage. For this there are several types. We have the carboxypeptidase A, B or D that have specificities because we have to remember that the enzymatic cleavage means we have to consider the substrate specificity, which is also important.

(Refer Slide Time: 15:38)



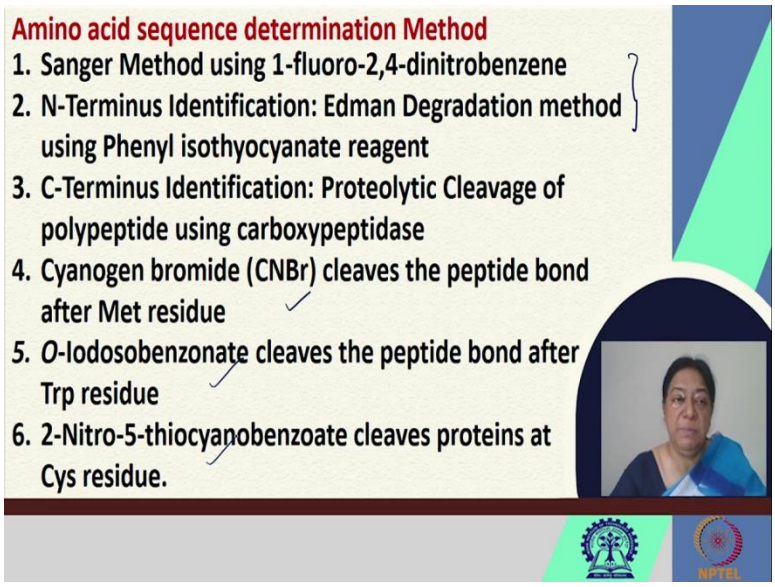
When we have a selective cleavage at the amino site of the C-terminal amino acid residue, it means it is just before the end.

(Refer Slide Time: 15:48)



So, this [refer to slide] is the amino site and we can identify our C-terminal amino acid. Similarly, if we use an Edman degradation method, we could find out the N-terminal amino acid for this specific polypeptide sequence. Then we can use other methodologies. For example if we use the trypsin, we would have a link break here and then based on this, we would be able to identify or work out what our peptide sequence is.

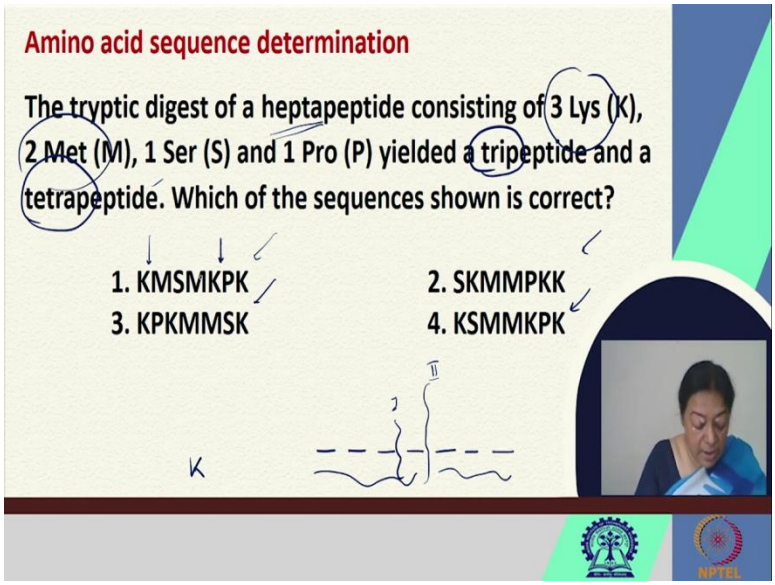
(Refer Slide Time: 16:18)



We have the amino acid sequence determination methods that have the Sanger method, the Edman method and the C-terminus method, that can be used with the carboxypeptides. So, we have one end that is the N-terminus that can be determined by these two methods.

We have the carboxypeptidase method that can be used to determine the C-terminus and the CNBr that will cleave up the peptide bond after the methionine residue and there are specific other chemical cleavage methods that can work on our polypeptide chain.

(Refer Slide Time: 16:50)



If we now look at a tryptic digest of a heptapeptide. What we mean by a heptapeptide, is that there are 7 amino acids and it has been found from say a complete hydrolysis of the polypeptide chain that there are 3 lysine residues, 2 methionine residues, 1 serine and 1 proline and obviously, there can be several sequences in which they can occur.

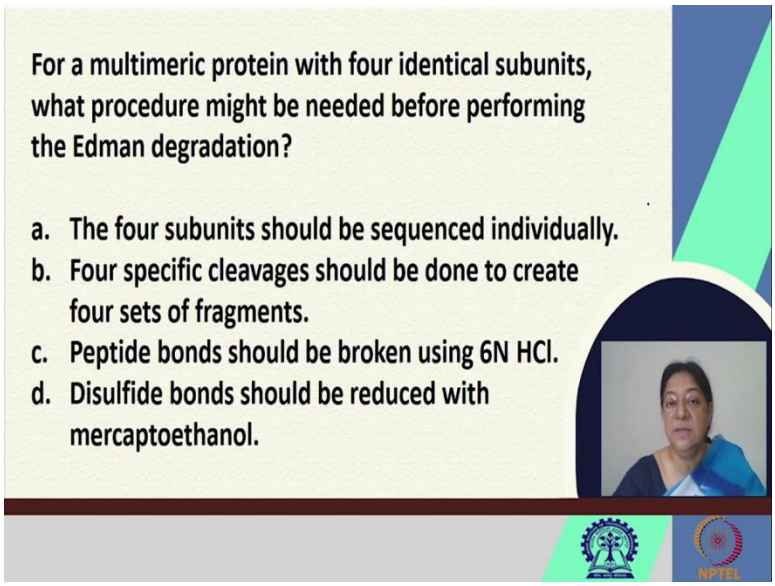
But the tryptic digest of the heptapeptide, consisting of this [refer to slide] combination of amino acid residues gave a tripeptide and a tetrapeptide, which meant that if we have a sequence where we have 1, 2, 3, 4, 5, 6, 7 we could have a cleavage here that is going to give us our tripeptide and tetrapeptide in one case. So this would be case 1. Or case 2 where we would have the cleavage here, giving us a tetrapeptide here and a tripeptide here. The identification is to where this specific amino acid is going to cleave and what we would expect.

We know that the trypsin will work after the lysine residues, so lysine is K. We could have a break here, a break here for this particular sequence in the identification of this peptide. If we look [refer to slide] at number 1, we can have a cleavage here, we can have a cleavage here; but none of this is giving us a tripeptide and a tetrapeptide.

In number 2, we could have a cleavage here and a cleavage here; but again we will have a single amino acid and a dipeptide. So we do not get a tripeptide or a tetrapeptide. In number 3, we have a K here, but we see that it is followed by a proline. This means that the cleavage will not occur here. We will have a cleavage here, which gives us a tripeptide and a tetrapeptide. In number 4 also, a cleavage here and a cleavage here does not give us the tripeptide and the tetrapeptide.

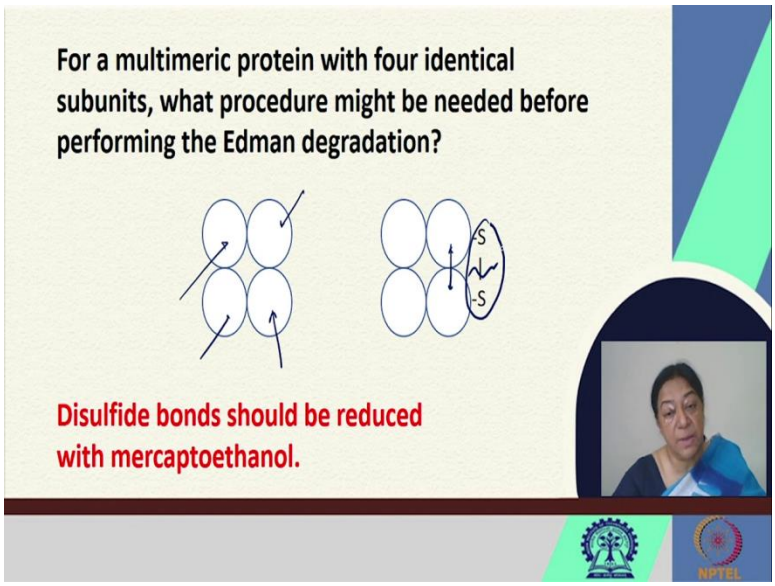
So this is where we have our answer, in which of the sequence shown is going to conform to our experiments here.

(Refer Slide Time: 19:19)



Now if we have another example, where we are looking at a multimeric protein with four identical subunits and we want to do an Edman degradation on it, what should we do? Now, is it necessary for the four subunits to be sequenced individually? Should the cleavage be done, will we have the four sets of fragments? Do we need to look at the broken peptide bonds or do we need to actually disrupt the disulfide bonds?

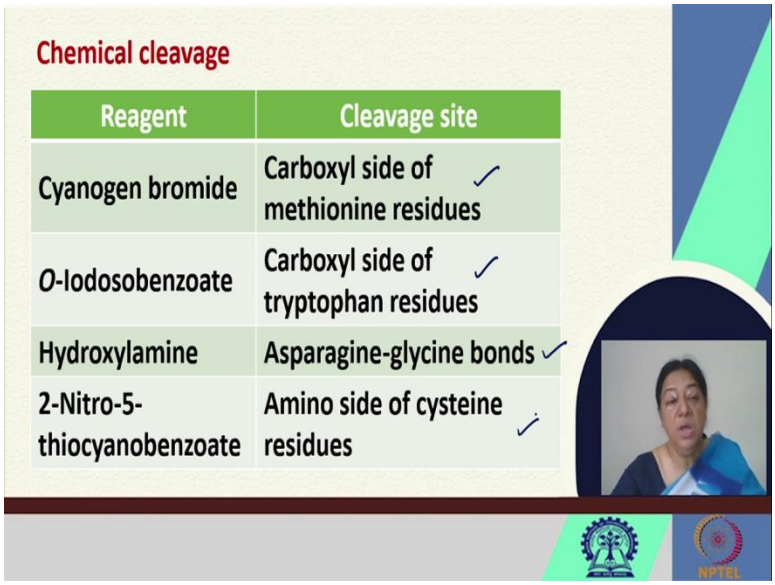
(Refer Slide Time: 19:54)



Since the four subunits are identical [refer to slide], it means that the four sequences are also identical. If we just take one and do the peptide fragmentation, the enzymatic cleavage, we will be able to identify the sequences for all of them. But we see that we have a connection here between these two subunits.

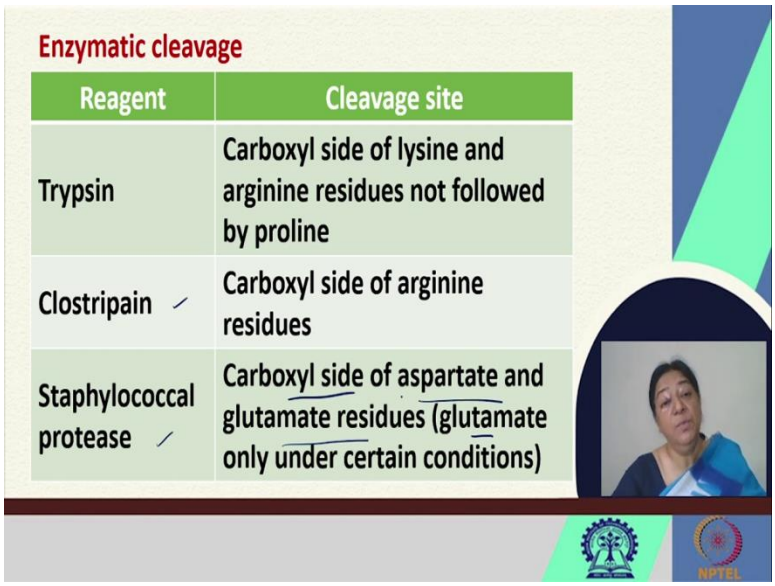
Given our methodology in the enzymatic cleavage, this disulfide linkage would remain. So, we would have a different peptide fragment for this particular set of subunits. In this case, what would be necessary, would be to break the disulfide bonds with β-mercaptoethanol and get exactly the subunits that can be sequenced in several methodologies.

(Refer Slide Time: 20:40)



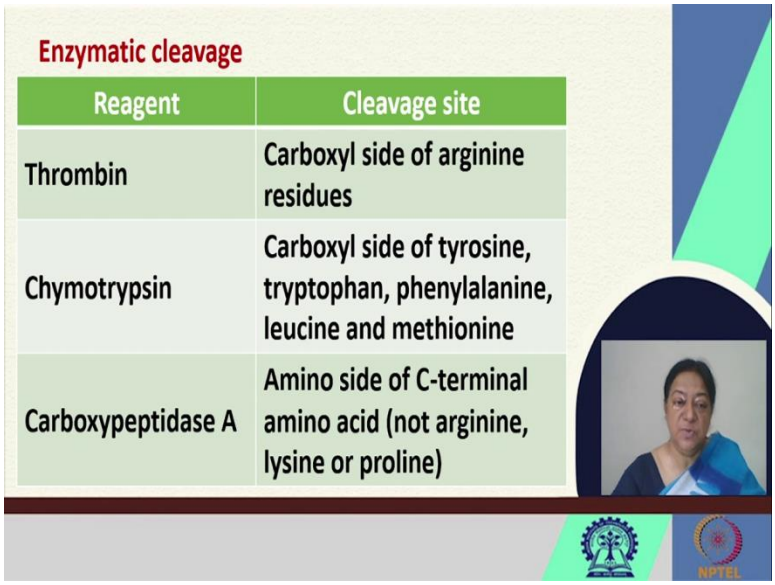
So whether we are using a chemical cleavage method with these different reagents, which are going to be used for specific ways in which we can identify the amino acid residues associated here or if we use enzymatic cleavage methods, where we can use trypsin.

(Refer Slide Time: 20:57)



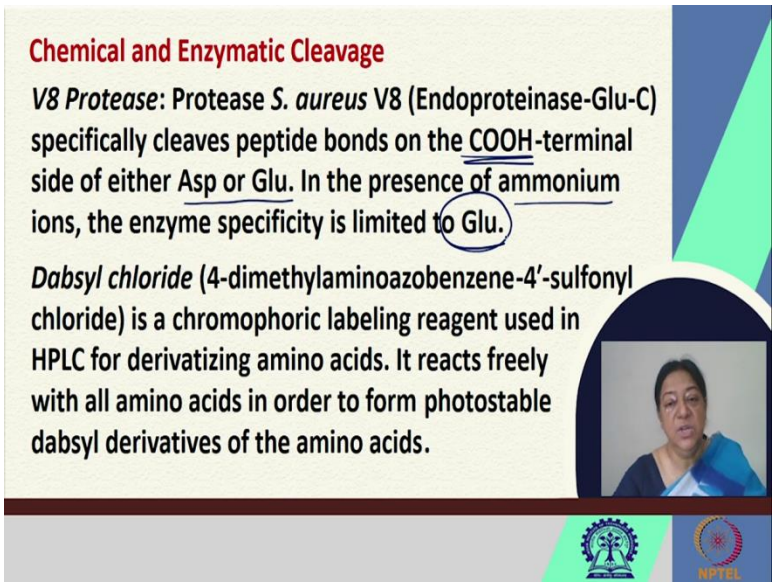
There are other methods; clostripain, staphylococcal protease or V8 protease, that cleaves on the carboxyl side of aspartate and glutamic acid and can be used to cleave only for glutamate under specific conditions, where we have ammonium ions.

(Refer Slide Time: 21:19)



Then we can also look at enzymatic cleavage for thrombin, for chymotrypsin, for carboxypeptidase and also for elastase and pepsin, there are many many enzymes we understand that can cleave these peptide bonds at specific locations. Given that there are enzymes, we understand their specificity associated with their mode of action and based on this, we can look at the identification of peptide sequences.

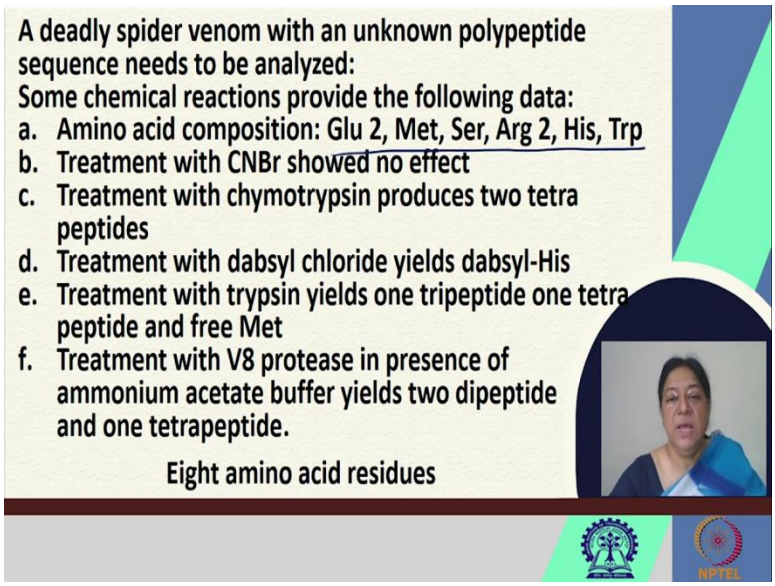
(Refer Slide Time: 21:44)



So here we have the V8 protease, which can cleave on the carboxyl side of the aspartic or glutamic acid and in the presence of ammonium ions, it is limited to glutamic acid.

Again we have another method of dabsyl chloride, this is again a labeling agent, chromophoric labeling agent, that is used in HPLC for derivatization of amino acid residues and it reacts freely with all amino acid residues and gives us photostable derivatives. So this can also be used to identify the N-terminus.

(Refer Slide Time: 22:28)



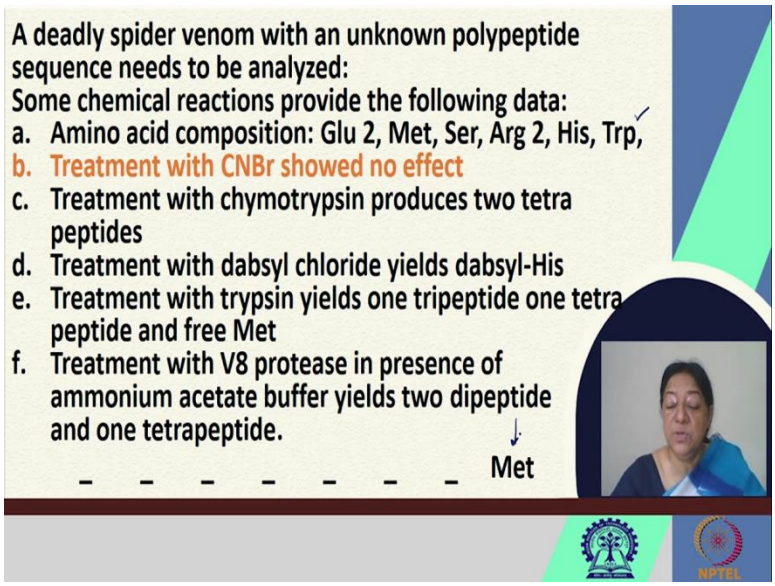
Let us look at how these experiments can actually be done. So, if we have a deadly spider venom that has an unknown polypeptide sequence that needs to be analyzed, we have to do some specific chemical reactions, specific enzymatic reactions and this is what is determined.

We have the amino acid composition. This is determined by a complete hydrolysis, where all the peptide bonds are broken and we can identify that this is what we get. It has an unknown polypeptide chain, for which we are going to determine the sequence. The treatment with CNBr showed no effect, the treatment with chymotrypsin gave two tetrapeptides, treatment with dabsyl chloride yields dabsyl histidine and treatment with trypsin gave 1 tripeptide, 1 tetrapeptide and free methionine.

Based on all these informations, we are now going to work on each of the points to find out what the overall sequence is. This gives us an idea of how to identify from the polypeptide sequence, from the specific enzymatic or chemical reactions that are done on the polypeptide, to determine what the sequence are.

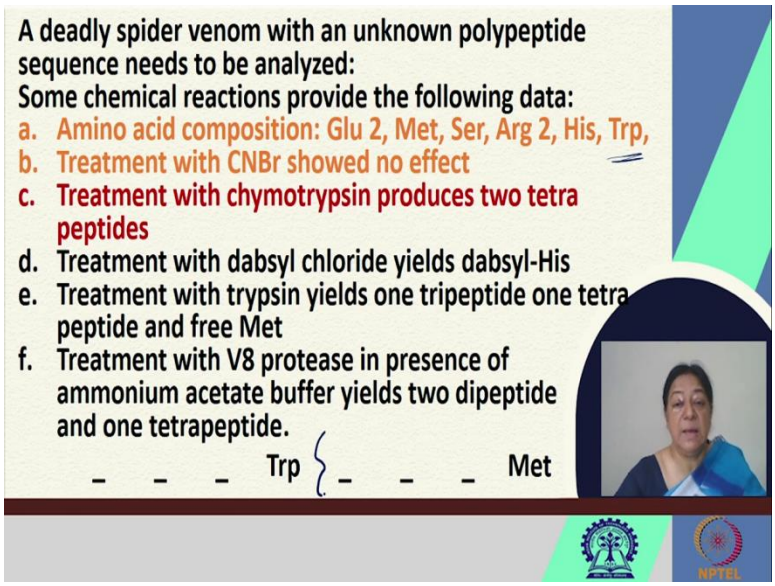
So we have 8 amino acid residues. From these 8 amino acid residues we want to identify what we have. The treatment with CNBr showed no effect. We understood from the previous slides or the previous part of the lecture that CNBr acts on methionine to form the lactone. Now, the lactone possibility is only when the methionine is within the protein or within the polypeptide chain.

(Refer Slide Time: 24:27)



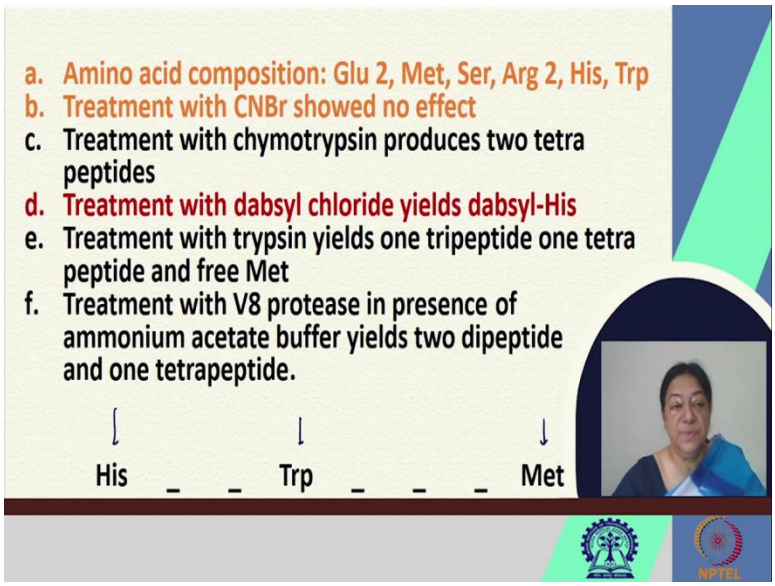
This means that when we have the treatment with CNBr showing no effect, it means that methionine has to be the C-terminus. We can also determine the C-terminus by using a carboxypeptidase. The next was treatment with chymotrypsin gave us 2 tetrapeptides. Now we have the amino acid composition that will tell us the chymotrypsin will work on the carboxyl site of aromatic amino acid residues.

(Refer Slide Time: 24:51)



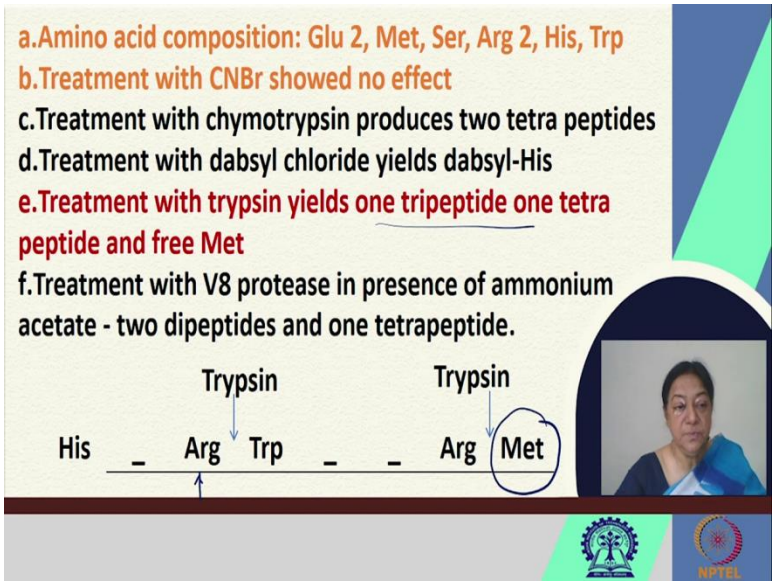
So, it gave 2 tetrapeptides and given that this is an octapeptide, we know that we can have the cleavage here [refer to slide]. That means that there are two tetrapeptides. The dabsyl chloride gave us dabsyl-His.

(Refer Slide Time: 25:11)



What does this mean? This means that this is giving us the first amino acid. So we have the histamine here [refer to slide], we have the tryptophan here and we have the methionine here. Now we have to identify the others.

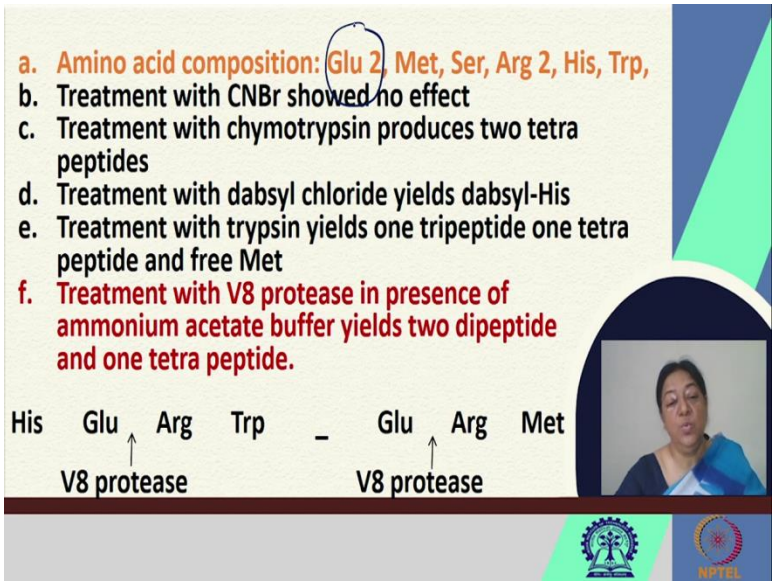
(Refer Slide Time: 25:23)



So we look [refer to slide] at treatment with trypsin. Trypsin will work after amino acid residues that are basic. We have arginine at two locations. The arginine here gave one tripeptide, one tetrapeptide. So the free methionine was here. So, this had to be preceded by an arginine and we had 1 tripeptide and 1 tetrapeptide.

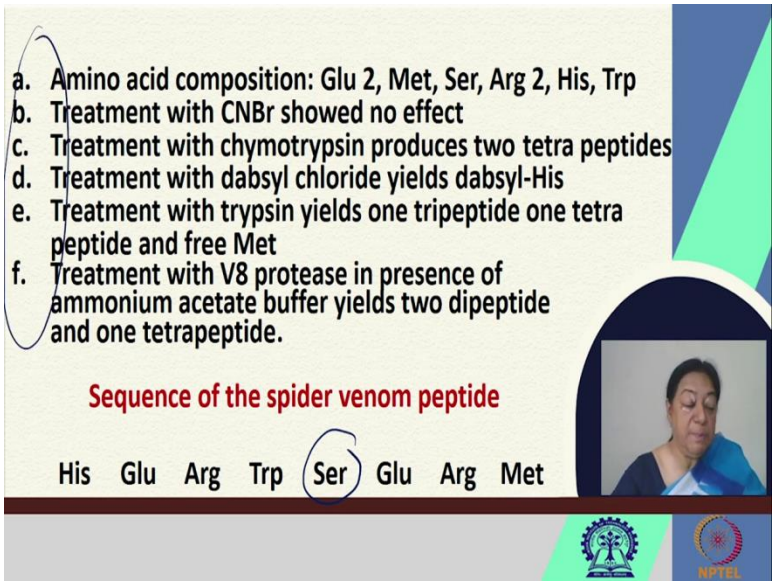
Given that, we could look at the location of the arginine based on the identification of the specific residues. So, given that the location of the arginine elsewhere would not give us the tripeptide, the tetrapeptide and the free met, this is the only location possible for the arginine residues. It's like a puzzle.

(Refer Slide Time: 26:16)



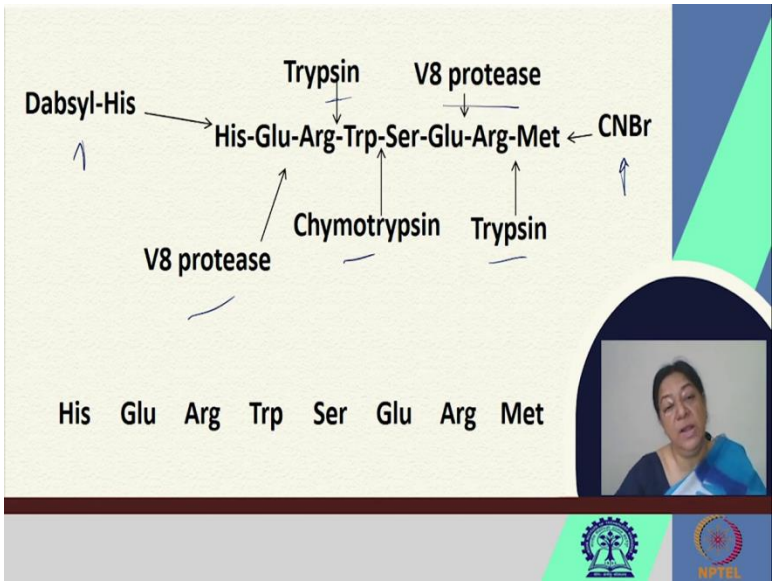
So we can identify them. The treatment of V8 protease gave 2 dipeptide and in the presence of ammonium; which means that it would work only on glutamic acid. In this case, we do not have an aspartic acid, so it does not matter. In this case, we have 2 dipeptides and 1 tetrapeptide. Based on that, we will work out the locations again and find out that the cleavage with the V8 protease,that will give us the rest of the sequence.

(Refer Slide Time: 26:47)



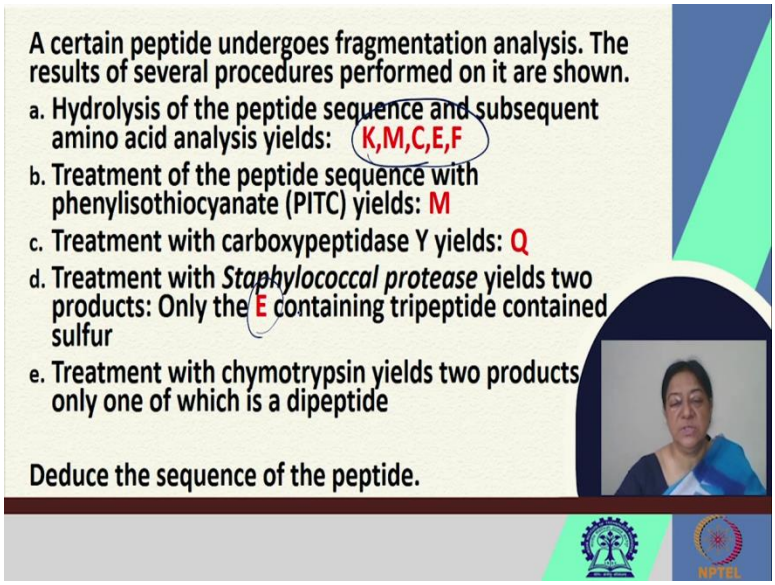
We have only 1 amino acid left, that is the serine. So we have [refer to slide] the sequence of the venom peptide, from all these specific chemical and enzymatic methods that are going to help us identify the sequence.

(Refer Slide Time: 27:03)



So this is the overall summary of what we looked at and how we identified the specific by looking at chemical methods of cleavage, by looking at enzymatic methods of cleavage, to determine the overall peptide sequence.

(Refer Slide Time: 27:21)

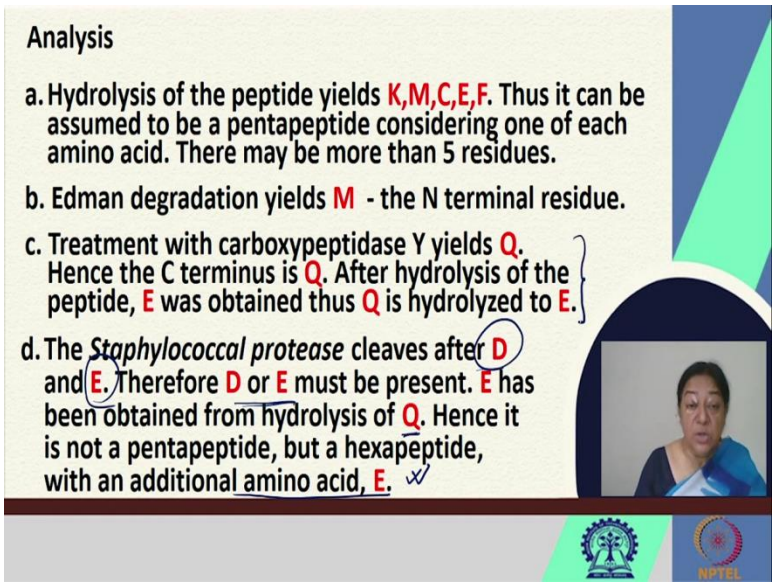


Let us look at just another example, a similar example where we have the hydrolysis of the peptide sequence and amino acid analysis, that gives us this set K, M, C, E, F of amino acid

residues and the fact that we have this, it may so happen that we have two of one of them, which does not necessarily mean that we have one each.

So, the treatment of the peptide sequence with PITC yields M. Treatment with carboxypeptidase gave Q and treatment with protease gave two products and the E containing product tripeptide had sulfur and treatment with chymotrypsin gave two products, one of which is a dipeptide. So again we want to find out what the sequence of the peptide is.

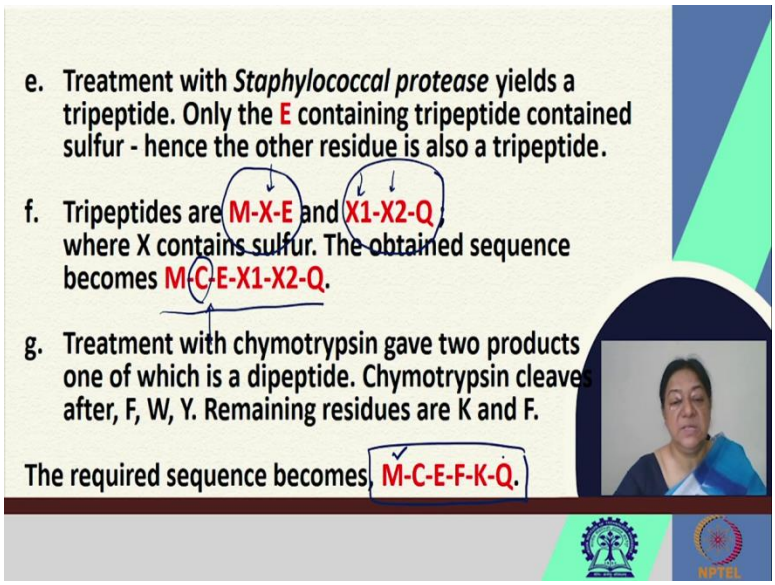
(Refer Slide Time: 28:14)



If we work in a similar analytic fashion, like we worked in the previous case. We have K, M, C, E, F which means that this is a pentapeptide, but could have more than five residues. But these are the specific identities of the amino acids present. The Edman degradation gave M, the N-terminal residue. The carboxypeptidase gave Q; that means, the C-terminus in Q and after hydrolysis of the peptide E was obtained, thus Q is hydrolyzed to E.

This means that there is another additional amino acid E. The staphylococcal protease cleaves after D and E. Therefore D or E has to be present. E has been obtained by hydrolysis of Q, which indicates that this is not a pentapeptide, but a hexapeptide with an additional E, because after hydrolysis we also got E.

(Refer Slide Time: 29:22)

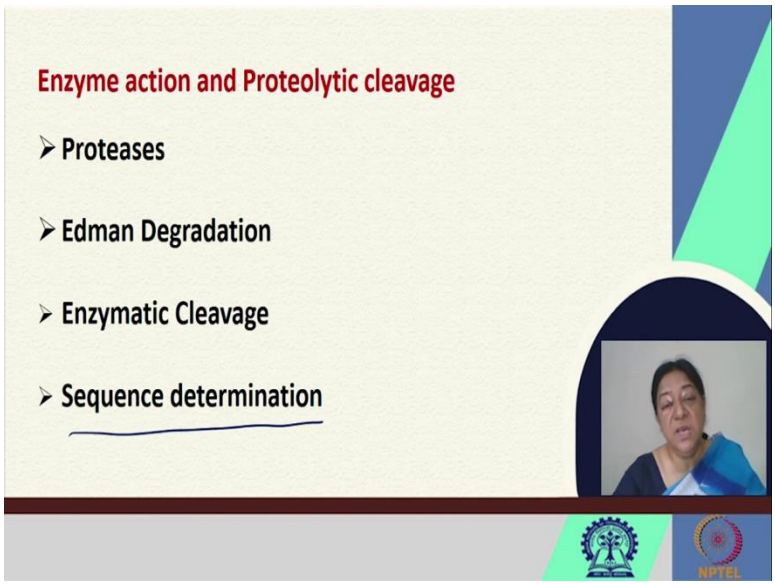


So working out on this mechanism, we have the E containing tripeptide that contain the sulfur. So, the other residue is also a tripeptide. We worked out the two different tripeptides, M-X-E and X1-X2-Q. M-C-E-X1-X2-Q is our hexapeptide and this was the one that contained the sulfur.

So, if we work it out with the different methodologies and the different analytical methods, it would be a good idea to look at the original problem and then try to work it out as to how each of our methodologies, in terms of the enzymatic cleavage and the chemical cleavage fit into an understanding of the overall sequence and how it is to be identified.

This can worked out to give us the original sequence, where we have the tripeptide M-X-E and X1-X2-Q. These are the ones that we are trying to identify, the locations of the specific amino acid residues and since one of the tripeptides that have the methionine or have the E or have the sulfur, indicated that X had to be cysteine. So, this gives us our identification. We get the N-terminus, we get the C-terminus and we get the specific fragments with the enzymatic cleavage.

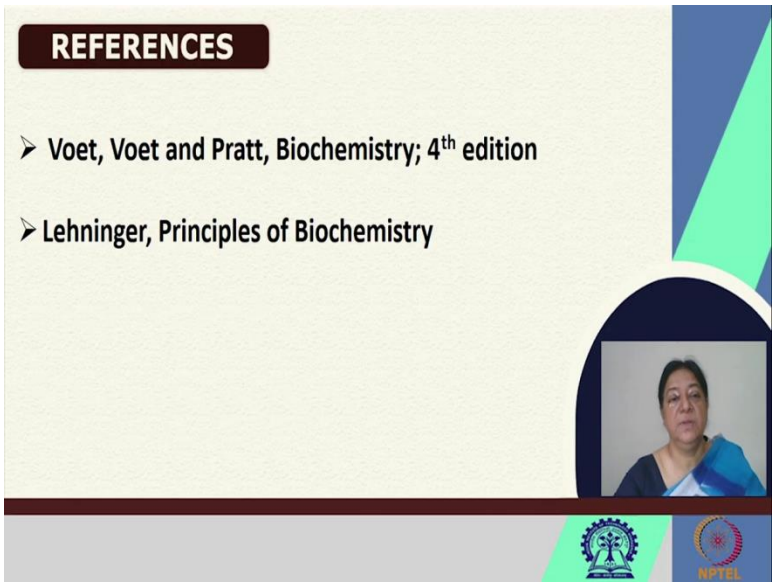
(Refer Slide Time: 30:47)



So from the enzymatic action and the proteolytic cleavage using proteases, Edman degradation, enzymatic cleavage, we can actually look for a sequence determination. There are metal methodologies now that have automated peptide sequencers, where you can actually give your peptide or your protein.

There would be a tryptic digest. From the mass of the tryptic digest and following the sequences, following specific chemical cleavage and enzymatic cleavage methods, we can actually determine the specific sequences of the peptides, pieced together by different methodologies and then find out our overall polypeptide sequence.

(Refer Slide Time: 31:33)



These [refer to slide] are the references.

Thank you.