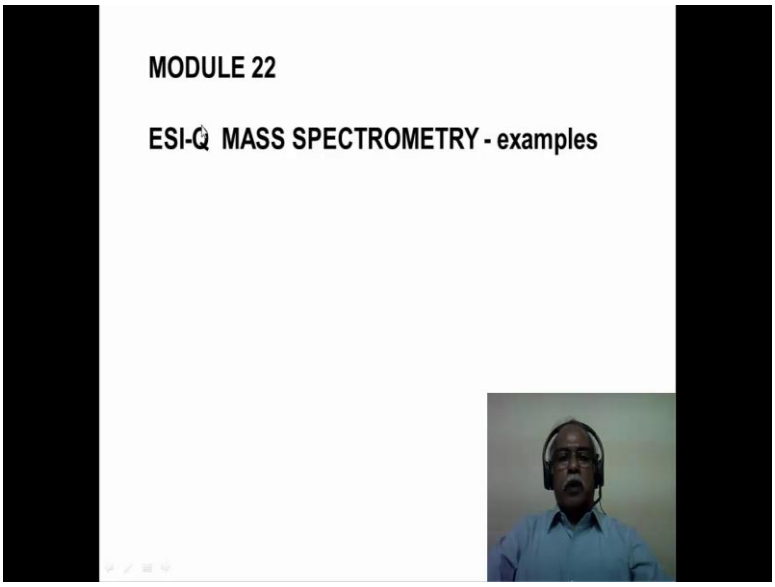
Application of Spectroscopic Methods in Molecular Structure Determination

Prof. S. Sankararaman
Department of Chemistry
Indian Institute of Technology, Madras

Lecture - 22
ESI and MALDI Mass Spectrometry

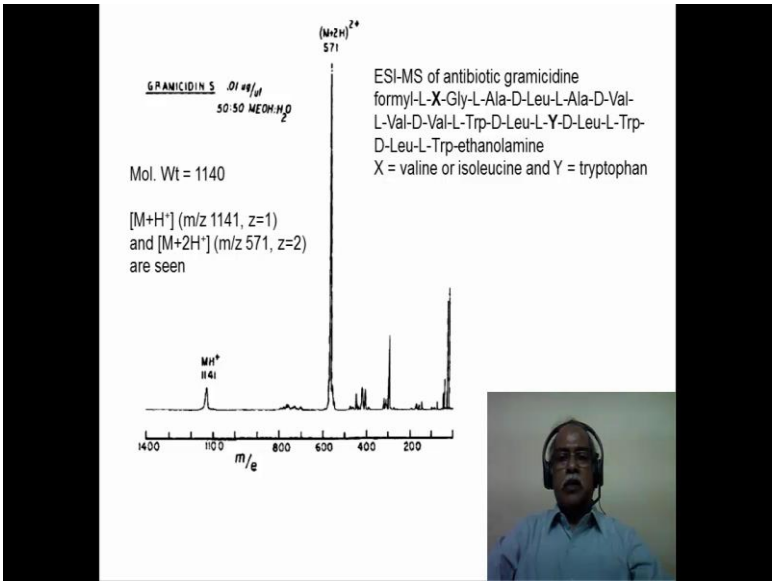
Hello, welcome to module 22 of the course on application of spectroscopic methods in molecular structure determination. We will continue with electrospray ionization mass spectrometry in this module also. Giving examples from bio molecular systems, where you have large proteins and peptides being analyzed by the mass spectrometry technique.

(Refer Slide Time: 00:31)



So, this is again, we are going to talk about examples of the electrospray ionization quadrupolar type of Mass Spectrometry in this particular module.

(Refer Slide Time: 00:40)

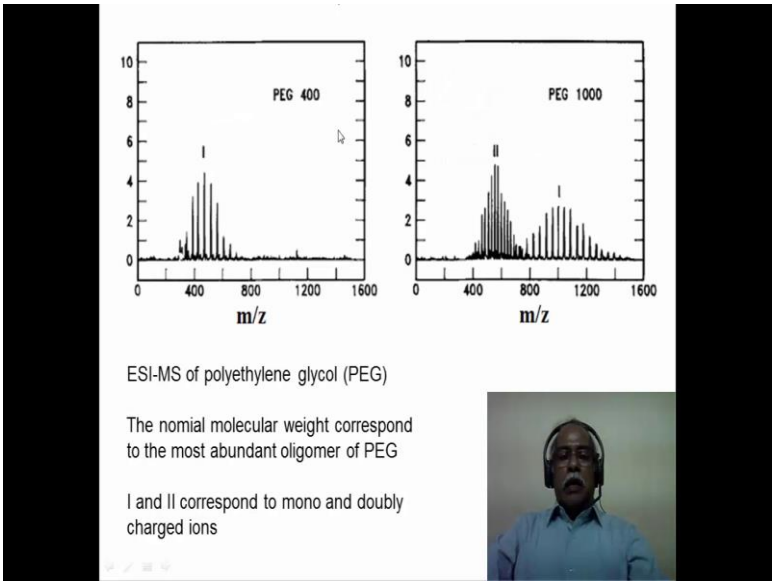


Let us start with an antibiotic called Gramicidine. Gramicidine has about 15 amino acid residues or so, which are given in the sequence. It is shown here in this particular slide. The molecular weight of this particular antibiotic is 1140. It is not a very large molecule, but the problem is that it is a non-volatile substance. And, that is why it could not be analyzed by electron impact ionization mass Spectrometry.

Secondly, it will undergo extensive fragmentation and you may not be able to see the molecular (Refer Time: 01:11) in the electron impact ionization mass Spectrometry. So, the ESI-MS was done using a 50-50 mixture of methanol water mixture. And, there are 2 types of ions that are produced under this condition. One is the molecular weight plus the proton, which corresponds to 1141. Then molecular weight plus 2 protons, which will be z is equal to 2. So, that would correspond to m by z value of 571. So, one sees the molecular ion plus the proton at 1141 as a small peak here. The doubly protonated system is the major peak. In particular conditions, this is a major ion that is produced.

So, from this we conclude that it is a molecular weight of 1140 which is the doubly charged ion, which corresponds to 571 in this particular case.

(Refer Slide Time: 02:02)



Now, polymers are usually not uni-dispersed system, they are poly-dispersed system. In other words, there is a range of molecular weights that are going to be present in synthetic polymers. The typical example is polyethylene glycol. Polyethylene glycol is commercially available. In different molecular weight, polyethylene glycols are available. They are called PEG. And, the number that is associated which is the average molecular weight of that particular polyethylene glycol. PEG four hundred means, it is the most abundant molecular ion, molecular weight. In this particular synthetic, polyethylene glycol is 400, molecular weight is one thousand is PEG 1000 and so on.

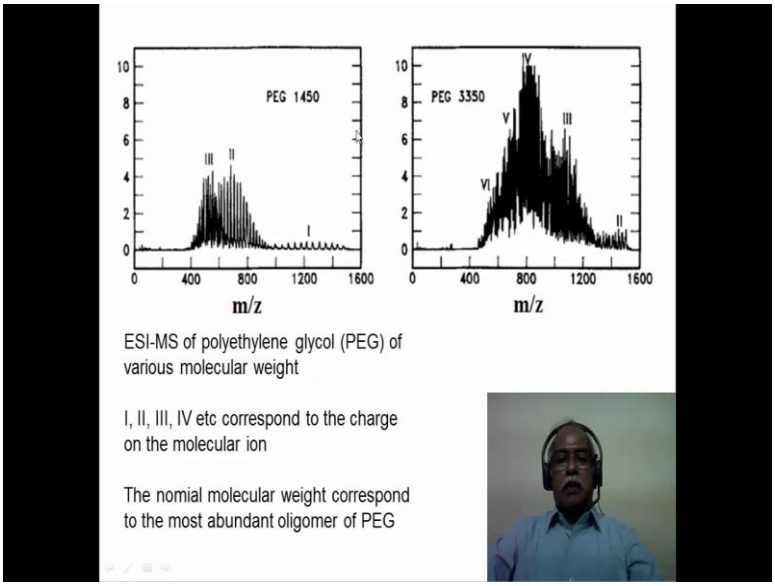
So, when you analyze the polyethylene glycol using electrospray ionization mass spectrometer, this is a kind of spectrum that one gets. Almost a Gaussian distribution of molecular weights; because it is a, polydispersity index is not equal to one. So, you have many molecular weight distributions that is present in this particular system. However, the nominal molecular weight corresponds to the most abundant oligomer in this particular system.

In other words, the most abundant oligomer in this particular system could be somewhere around 420 or 440 or something like this. That corresponds to the molecular

weight of about 400. So, that is why it is called polyethylene glycol; 400 in this particular case.

Polyethylene glycol 1000 is shown here. So, the most abundant nominal ion corresponds to roughly 1000 or so. And, this is a mono charged ion and this is doubly charged ion. In other words, it is protonated 2 times. So, the molecular weight is exactly halved, for example, in this particular case. So, you have this one and two, representing the mono and the doubly charged ions of the polyethylene glycol 1000.

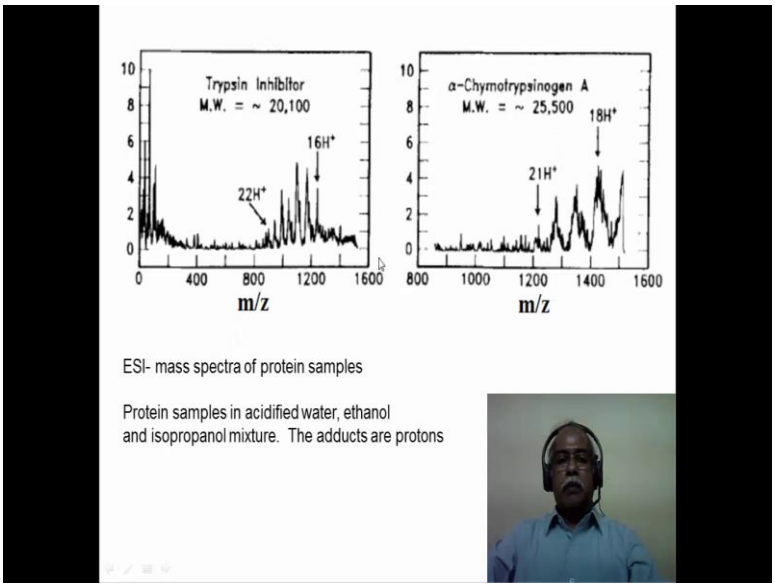
(Refer Slide Time: 03:52)



The high molecular weight polyethylene glycols, of course can be multiply protonated. This is polyethylene glycol 1450. The nominal molecular weight ion, molecular ion is shown here with the mono protonation corresponding to this. This multiple peaks that are shown here, essentially gives the distribution of molecular weights of this particular sample. And when it is doubly charged, for example, that becomes more intense. In other words, the mono charged species is not as abundant as the doubly charged and triply charged species in the case of high molecular weight polyethylene glycol. And, essentially you see the Gaussian kind of a distribution of molecular weights for this particular sample. And, the triply charged ion also shows the similar distribution pattern as far as the molecular weights are concerned.

And, when you go to even higher molecular weight of 3350, you will not only see the doubly triply charged, you have quadruply and 5 charge and 6 charges associated with the protonation of these polyethylene glycol chain. Polyethylene glycol is poly ether. So, it is easy to protonate the oxygen lone pairs in this type of samples. Lewis basic sites are plenty in nature in this kind of sample. So, depending upon the pH condition that one analyzes this molecules, one can go all the way up to 6 protonated species corresponding to m by z; where z is equal to 6 in this particular case.

(Refer Slide Time: 05:13)



Now, let us go into some proteins and polypeptide samples. Electrospray ionization of protein samples, essentially the protein samples were dissolved in acidified water, ethyl alcohol, isopropanol mixture, and the adducts are essentially proton. In other words, what you are looking at is multiply protonated protein molecule is what we are looking at.

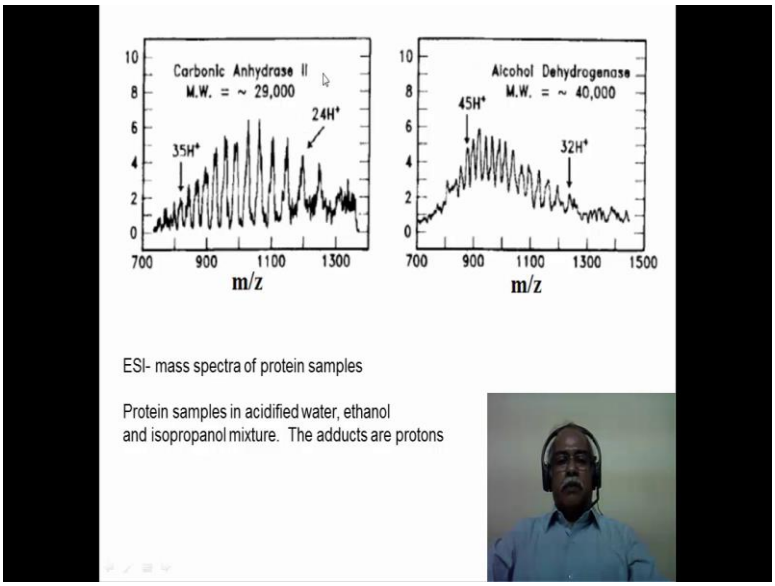
This is trypsin, which is an inhibitor; molecular weight is about 20,100 or so. What you are seeing is a molecular ion peak around 900, sorry, this is not 9 this is 1000 or so, and 1200 corresponding to 16 protonated species and 22 protonated species. In other words, in the m by z value, the z is equal to 22 in this particular case and z is equal to 16 in this

particular case. In between you have 17, 18, 19, 20, 21 and so on. These are the multiply charged ions of this particular inhibitor.

Let us take for example; this is 16, 17, 18, 19, 20. The 20 molecular weight corresponds to somewhere around 1000 or so. So, if m by z is z is 20, so it will essentially come to the molecular weight of 20,000 or so, in the case of the trypsin inhibitor.

Now, this is chymotrypsinogen, which is again a protein sample. This is also analyzed under the conditions of acidified water, ethanol and isopropanol alcohol mixture. Once again, you can see the multiply charged ions corresponding to 18 protonated species, 21 protonated species and so on. We will look in to the aspect of how we calculate the number of protons that are attached in a few minutes. But for the time being, assume that this is the 18 protonated species and this is 19, 20 and 21 protonated species. The molecular weight of this particular protein sample is 25,500. It is still not a very large molecular weight for a electrospray ionization mass spectrometry, nevertheless we are reaching to high molecule weight stages.

(Refer Slide Time: 07:07)

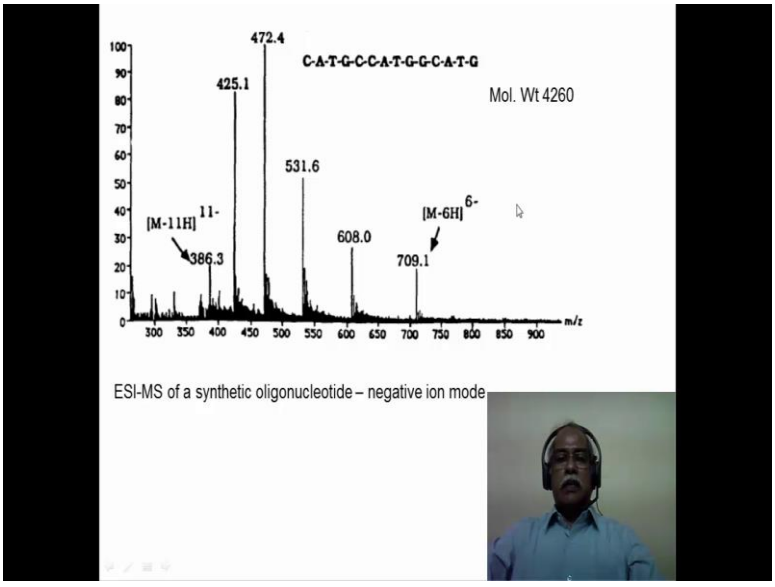


Carbonic anhydrase II is an enzyme. This is a protein sample. Molecular weight is 29,000 or so; approximately 29,000. One can determine the molecular weights by other

methodology also, other than the mass spectrometry. So, one can determine the molecular weight by mass spectrometry also, if one uses the electrospray ionization mass spectrometry. In this particular case, what you see is range of protonated species starting from 23 protonated species, 24, 25, all the way up to 35 protonated species corresponding to a molecular weight of 29,000 or so.

And, this is alcohol dehydrogenase. This is another enzyme used in the dehydrogenation of alcohol, for example, in the biological system. The molecular weight is 40,000 for this particular sample. So, you have a multi protonated species. There are many (Refer Time: 07:58) basic site present in this kind of enzymes in the form of nitrogen, carbonyl, oxygen and so on. So, protonation of multiply protonating the sample is not a difficult thing to do, as long as the PH can be adjusted to give certain number of protonation in the molecule. So, essentially, the 40,000 divided by z corresponding to 32 gives this particular peak. And, from this we get the information that the molecular weight is 40,000 for this particular sample of protein.

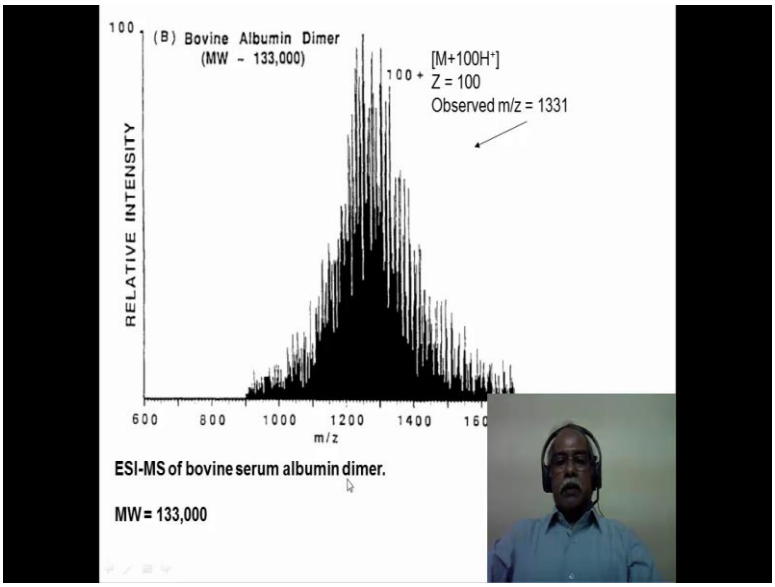
(Refer Slide Time: 08:27)



Here is a DNA sequence. The sequence is shown here very clearly. D N A, remember the skeleton of DNA has phosphate units which are negatively charged. So, it is advantageous to do the negative ion mode mass spectrometry in the case of DNA

samples; because DNAs are naturally negatively charged on the periphery of their because of the phosphate linkages that one has in DNA molecule. So, this is 6 minus charged species. Molecular weight is 4260. It is not a very large molecule, nevertheless one can see, for example all the way up to M minus eleven H. In other words, this is a; 11 negative charges are present in the skeleton of this particular polynucleotide. And, the molecular weight, if you divide the 4260 by 11, you essentially get around 386 or so, corresponding to the molecular weight of this particular oligonucleotide molecule.

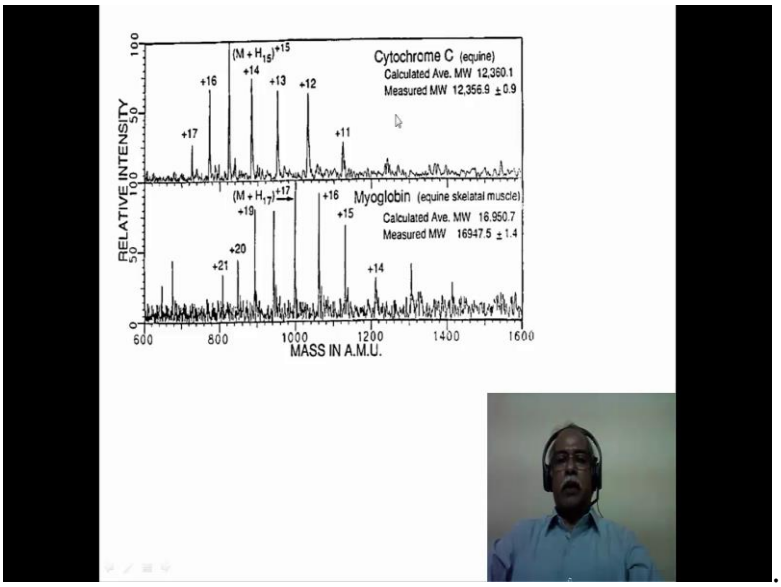
(Refer Slide Time: 09:26)



This is a Bovine albumin dimer. Molecular weight; look at this, they are reaching now hundred thousand molecular weight. 133,000 is the molecular weight of this particular protein. This is essentially the bovine serum albumin dimer. And, the molecular weight corresponds to this 133,000 or so.

The z value is 100. In other words, this is protonated 100 times at various sources, various sights. So, essentially the observed m by z is 1331. In other words, the Gaussian if you take the maximum abundant peak of this particular unit, this will correspond to 1331. 1331 times, 100 correspond to this particular molecular weight that is expected for this particular sample.

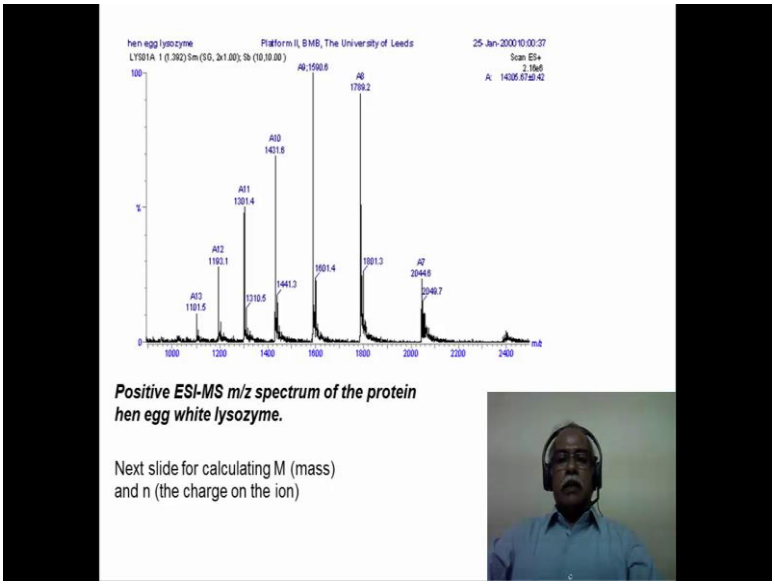
(Refer Slide Time: 10:15)



Now, Cytochrome C and myoglobin are 2 protein molecules. If the molecular weight of 12,360 and 16950 in this particular case, and here also multiply charged ions are seen. Charges starting from 14, 15, all the way up to 21 in the case of myoglobin.

And, in the case of cytochrome it is about plus 11 to plus 17 charges is what is seen here, corresponding to the molecular weights that are mentioned in the slide, in this particular case. The calculated molecular weight and the measured molecular weight, they are close to each other in terms of the accuracy with which you can measure the molecular weights using the electrospray ionization mass spectrometry.

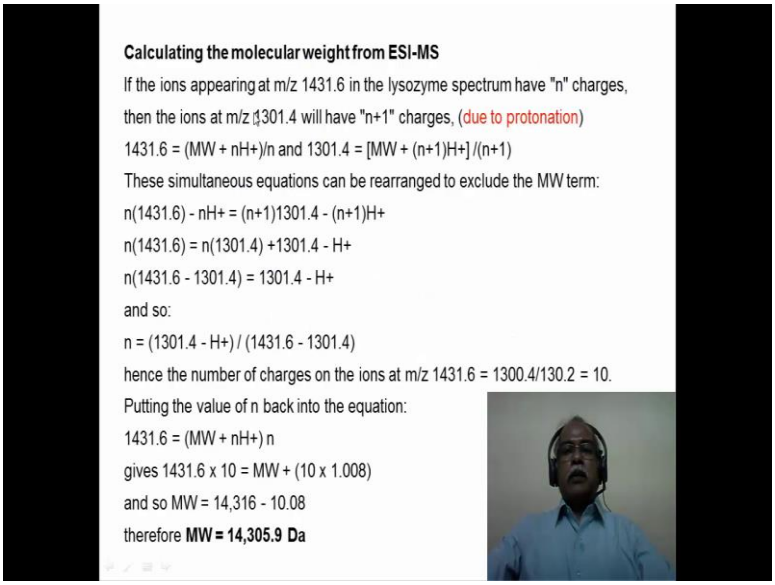
(Refer Slide Time: 10:56)



Now, let us take this example. This is a protein, which is isolated from the hen egg white lysozyme. This is the name of the enzyme. And this you what you see is, essentially different ion distribution in the form of a Gaussian distribution is what one sees. Let us take this particular ion at 1431 or so.

Let us take the example of the positive ion electrospray ionization mass spectrum of protein, which is isolated from the egg of the hen. This is the white; lysozyme is the name of the enzyme. This has a molecular weight of approximately 14,305 or so. That is the molecular weight of this compound. And, what you are seeing essentially is a charge of 10, charge of 11, 12, 13 and so on, starting from a charge of 7 in this particular case. In the next slide, we will see how the molecular weights are being calculated in this particular example.

(Refer Slide Time: 12:02)



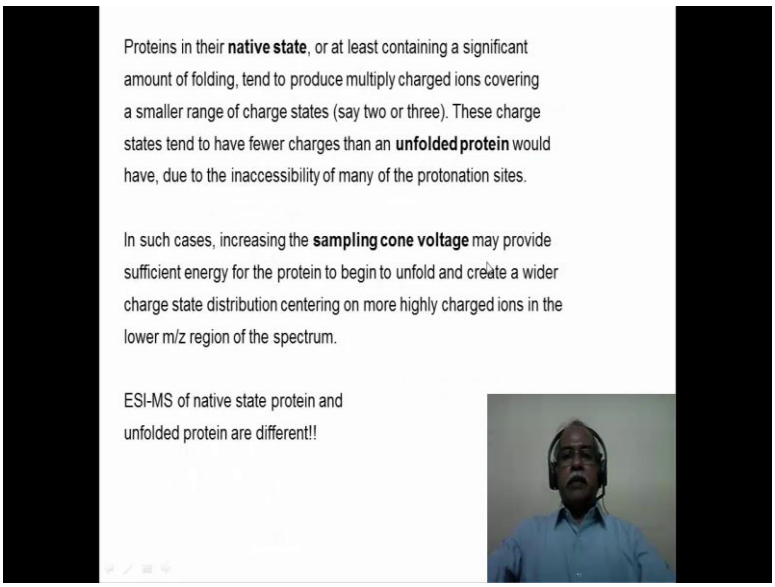
Now, let us arbitrarily assume that 1431.6 of the lysozyme spectrum has “n” number of charges in this particular ion. We do not know what is charge of this particular ion. Let us assume there are n charges there. The next ion at 1301 will have “n plus one” charge. This is because of the protonation; one protonation more than the 1436. So, one can write an expression; 1431.6 is equal to molecular weight times the n number of protons added, divided by the charge which is n. And, 1301 will be corresponding to the molecular weight plus the n plus one number of protons added, divided by n plus one which is the charge of this.

If we can solve these two simultaneous equations and rearranged that in terms of the molecular weight, excluding the molecular weight, for example; because the molecular weight will be same in both cases. So, that expression reduces to this particular one. And, if you can solve this expression to n, then n would correspond to essentially, 10. You go through this particular numerical analysis of this particular system. So, the number of charges on the ion at m by z of 1431 will be equal to roughly 10. Putting the value back into this equation here, 1431.6 is equal to molecular weight times 10 protons divided by 10.

So, taking that particular expression and solving for the molecular weight, one gets the molecular weight of 14,305. 14,305 is the expected molecular weight of this particular compound.

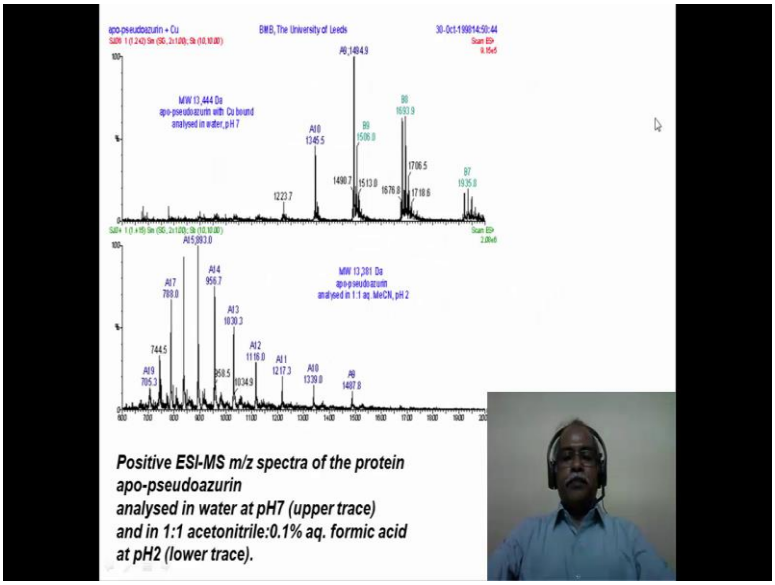
So, one need not actually know the charge of the particular ion, but what one needs to know is the adjacent ions are going to be 1 plus charge more or 1 minus charge less in terms of going on either side of the spectrum, starting from let us say the most abundant ion, in this particular case. So, it is fairly easy to calculate the number of charges and finally, arrive at the molecular weight of this particular enzyme easily.

(Refer Slide Time: 14:08)



Now, proteins in their native state has certain conformations, certain three-dimensional structures. It is not necessary that all the possible sites that can be protonated can be exposed under those conditions. Some of the sites can be really buried under the three-dimensional structure, which are inaccessible for protonation and so on. So, if for example, a native protein is denatured and if such protonation sites are exposed, then you have a wider distribution of charges, charged species in the case of denatured protein compared to a native protein. So, the electrospray ionization mass spectrum of a native protein and the unfolded protein can be very different in nature.

(Refer Slide Time: 14:48)



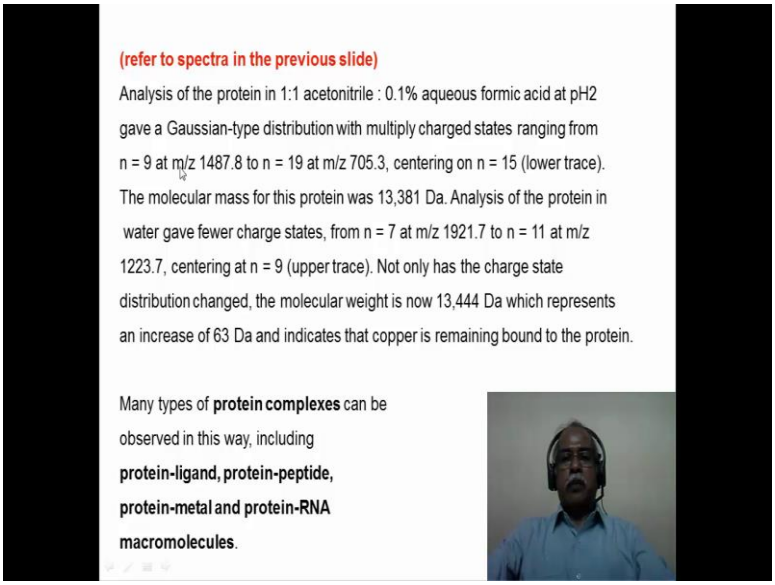
This is explained in this particular slide here. This is the; apo-pseudoazurin is the name of the protein in this particular case. The electron impact ionization mass spectrum of positive charge; electron impact ionization mass spectrum is what is given here. It is analyzed in two different pHs. One, the top trace is pH7 and the bottom trace is pH2. So, essentially you are changing the acidity of this system in water. So, in the case of pH7, which is a neutral pH, you have fewer charges on the upper protein. This is 7 charged, 8 charged, 9 charged, 10 charged species is what you are going to have in this particular case.

Whereas, at pH2 which is a much lower pH, more acidic system, you are starting from 9 charge all the way down to 19 charges are there in this particular species, that is having a molecular weight of 700; m by z value of 705.3 or so. So, not only that you have a molecular weight difference of about 13,444 in this particular case. Whereas, in this case 13,381. There is a 63 mass unit difference between these two.

So, let us assume that at the pH of two, the molecule is denatured. And as a result of denaturalization, multiple protonation sites are available with particular molecule. Whereas, in the native state at the pH of 7, not that many protonation sites are available. Not only that, there is also a copper that is bound in the system when it is in the natural

form. When it is denatured that copper is stripped off from the protein, and that is a reason you have molecular weight difference of 63, corresponding to the denaturalization and the removal of copper from the denatured protein is what you have seen here. And, this is explained here.

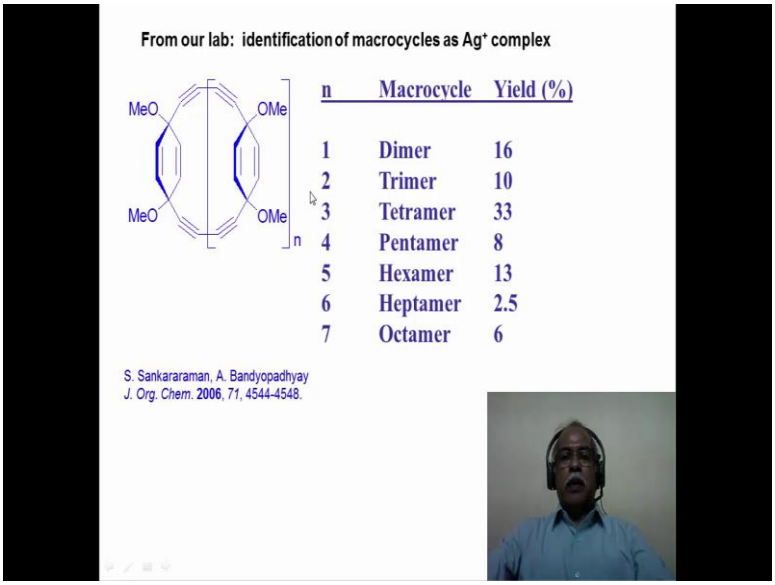
(Refer Slide Time: 16:33)



The analysis of the protein in a 1 is to 1 acetonitrile, aqueous formic acid at pH2 gave a Gaussian distribution of multiply charged ions. This is the Gaussian distribution we are talking about. And at a lower pH, for example, a fewer charged ions are being given here. Not only that, this is bound to copper and this is not bound to copper, which is evident from the difference in the molecular weight that is calculated for the top spectrum as well as from the bottom spectrum, which has a molecular weight of difference of 63 which corresponds to copper ion in this particular case.

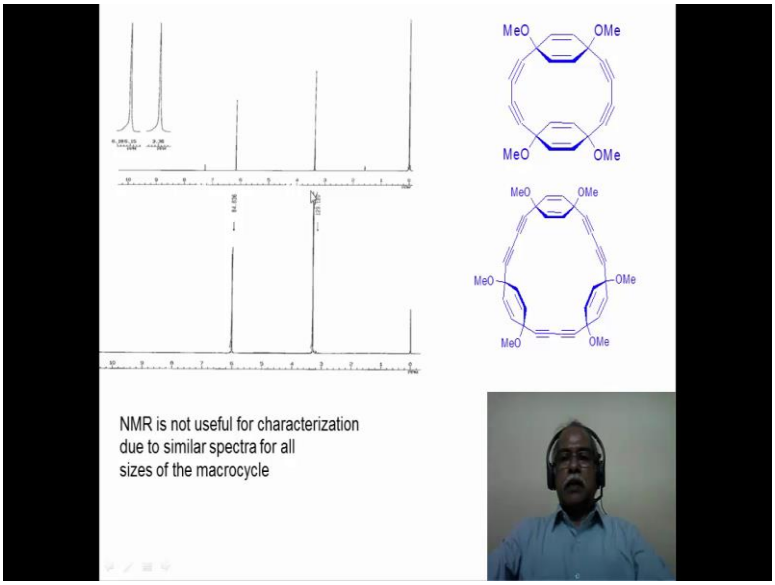
Many type of protein complexes can be analyzed and observed under electrospray ionization mass spectrometer, including protein-ligand, protein-peptide, protein-metal, protein-RNA type of macromolecular associated systems can be analyzed in the electrospray ionization mass spectrometer.

(Refer Slide Time: 17:35)



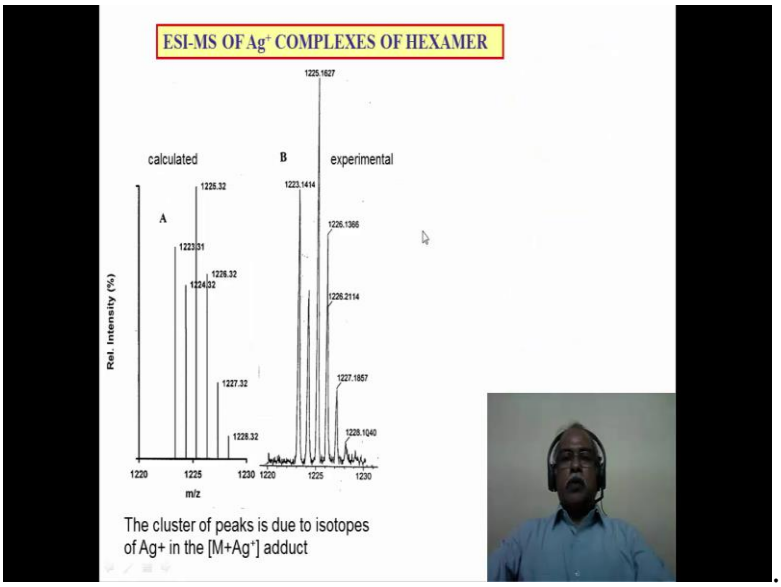
Now, let us give an example from our own research from our laboratory. Sometime back, we made lot of macro cycles of this type. So depending upon the value of n, you can have a dimer, trimer, all the way up to the octamer. And, these are the yields of this system. One of the problem with the characterization of this class of compounds is proton NMR spectrum is essentially useless; because of the highly symmetrical nature of this molecules, the cycle hexadiene protons give only one single line in the protein NMR spectrum and methoxy also give one single line in the proton NMR spectrum.

(Refer Slide Time: 17:54)



So, whether you take a dimer, trimer, tetramer or octamer, it always gives only one line for each of the ring protons as well as the methoxy protons in this molecule. And, not much of a chemical shift difference value also is available to distinguish some properly.

(Refer Slide Time: 18:26)



So, under this condition the electron impact ionization mass spectrometry using copper complex was very useful, sorry not copper, using silver complex are very useful. Why silver complex because silver is known to bind to the double bond and triple bonded system. So, we thought we will add silver triplet to the molecule, the solution of this molecule in acetonitrile and analyze the samples.

And, we were successful in doing that. So, silver ion has; silver as an element has an isotope 6 different isotopes are available for silver. So, essentially when silver coordinates to the; so, essentially when silver coordinates to this particular molecule, it forms a silver adduct. And, the silver adduct spectrum is what is shown here. This is a experimental spectrum and this is a calculated spectrum. One can calculate the expected isotope distribution based on the natural abundance of the various isotopes that are available, for example. Now, the molecular weight of this compound is shown here.

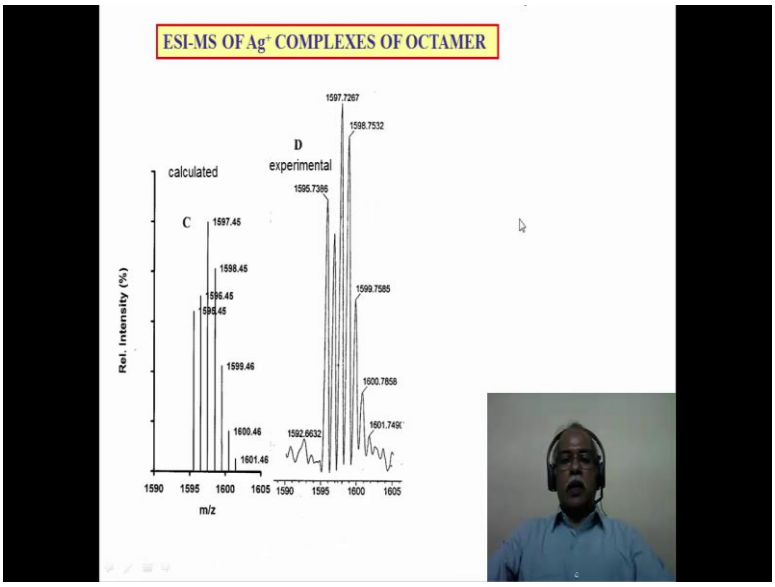
(Refer Slide Time: 19: 26)

n	Macro cycle	HRMS calc.for	HRMS found
2	Trimer	$C_{36}H_{30}O_6Ag$ 665.1093	665.1112
3	Tetramer	$C_{48}H_{40}O_8Ag$ 851.1774	851.1802
4	Pentamer	$C_{60}H_{50}O_{10}Ag$ 1037.2455	1037.2451
5	Hexamer	$C_{72}H_{60}O_{12}Ag$ 1223.3136	1223.3112
6	Heptamer	$C_{84}H_{70}O_{14}Ag$ 1409.3817	1409.3792
7	Octamer	$C_{96}H_{80}O_{16}Ag$ 1595.4497	1595.4481

This is the molecular weight plus the silver ion corresponding to 665 for the Trimer, 851 for the Tetramer and so on. And, this is the high resolution mass that is found for this molecule using the mass spectrometry. So, what you are seeing here is essentially the experimental spectrum showing the isotope distribution because of the silver isotopes that are present in nature. And, this is a calculated spectrum. You nearly have one to one

match of these two spectra in terms of the m by z values as well as the relative intensities of the ion peaks that you see in this spectrum. This is for the hexamer molecule.

(Refer Slide Time: 20:02)



The octamer molecule was also analyzed that way. You can see similarity in terms of the calculated spectrum and the experimental spectrum for the silver ion complex of this kind of class of molecule. So, essentially the entire range of polymer in this particular oligomer, in this particular cyclic oligomers, in this particular case was successfully analyzed. For example, using the electrospray ionization mass spectrometry, using silver ion to form an adduct with the substrate, so that it can be easily analyzed in this case.

So, hopefully we have seen large number of examples of peptides, proteins and bio molecular system with the molecular weight reaching nearly 100,000 or so. And, we saw the utility of the electrospray ionization mass spectrometer in analyzing such large macro molecular system. You have also seen some synthetic examples of polyethylene glycol and the macro cycle that were synthesized in our laboratory.

Thank you very much for your attention.