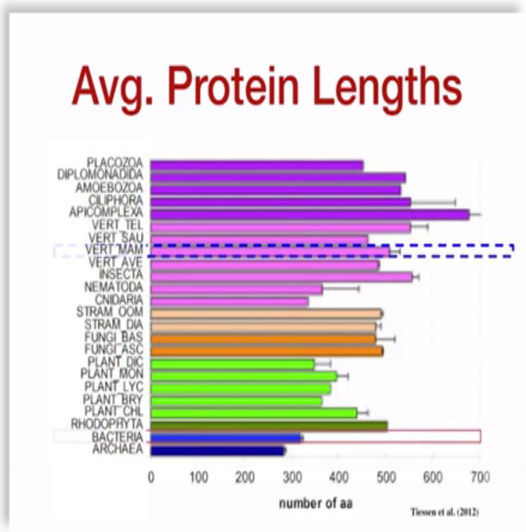
Cellular Biophysics
Professor Chaitanya Athale
Department of Biology
Indian Institute of Science Education and Research, Pune
Protein Abundance and Spacing

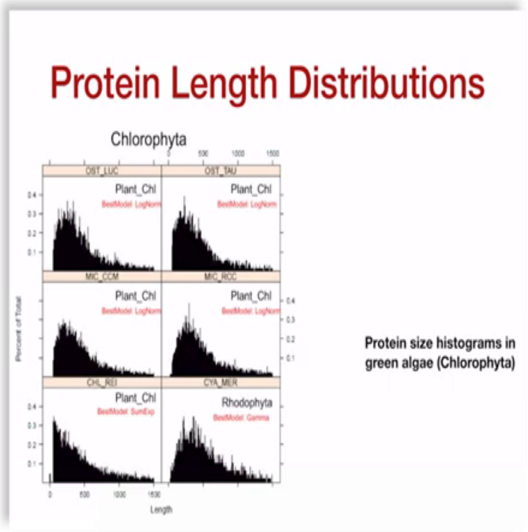
(Refer Slide Time: 0:17)



Hi! So far we have talked about DNA. Now we turn some attention to proteins. What you are looking at here is from a paper by Tiessen et al, the amino acid length on an average of proteins from multiple species, you see on the purple side Placozoa, diplomonadida, et cetera. Then you see vertebrates, blue dashed line indicates humans.

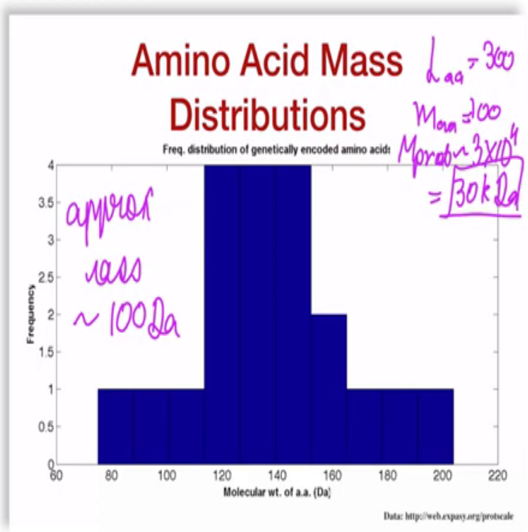
We stand at an average of about 500 amino acids per protein. Whereas bacteria are at 300 amino acids per protein. So, this means that knowing these statistics, by the way these statistics are publicly available on PubMed and the protein databases also. We can potentially also make sense of the protein abundance.

(Refer Slide Time: 1:30)



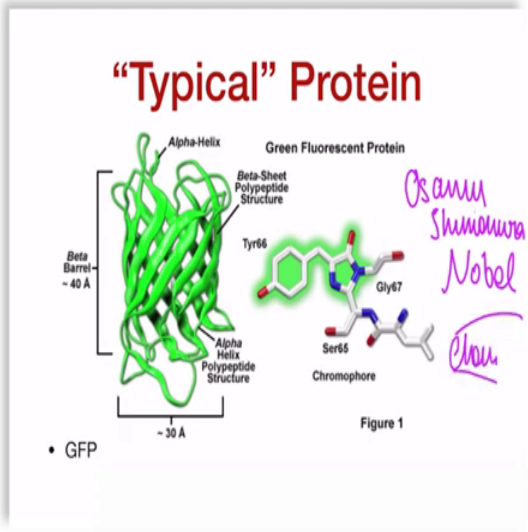
In a sort of similar spirit that I had shown earlier, which was missing the distribution of plant proteins, here you are looking at average length of plant proteins in terms of amino acids. And the average value again is somewhere around 200 to 300 amino acids for plant. The best model for this fit is a log normal, from green algae Chlorophyta.

(Refer Slide Time: 2:06)



So, the Expasy database protein scale also tells us about the molecular weight of amino acids, this is centred around 130 in terms of mass in Daltons, which means that we can allow ourselves to approximate mass of an amino acid has 100 Daltons. So, combined with the length amino acid as 300 and mass per amino acid as 100, we end up with the mass per protein on an average of 3 into 10^4 or 30 kilo Daltons.

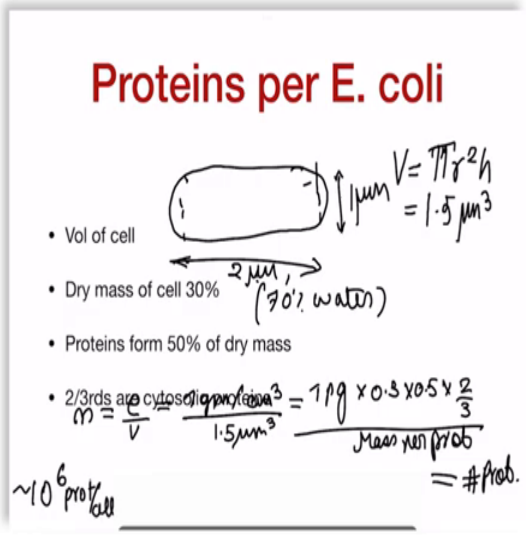
(Refer Slide Time: 3:20)



Now, I have made many approximations and simplifications here, but I will proceed to show you that these are quite useful. This is not a typical protein, that is why I put in inverted commas, but it is a protein you should know about.

The discovery of the protein itself is attributed to Osama Shimomura, who won a Nobel Prize for its discovery in chemistry. As you know there is no biology Nobel Prize, it is for medicine, you can claim many biologists won that and many biologists also won the Chemistry Noble Prize.

(Refer Slide Time: 4:06)



The proteins per *E.coli* is the question you ask. So one way to get at it is to ask - what is the volume of a cell. So we recall that he said that the cell is a rod of 1 micrometre or spherocylinder 1 micrometre height or diameter and length of 2 micrometre and if we ignore the semi-spheres here, then the volume is $\pi r^2 h$, which is…

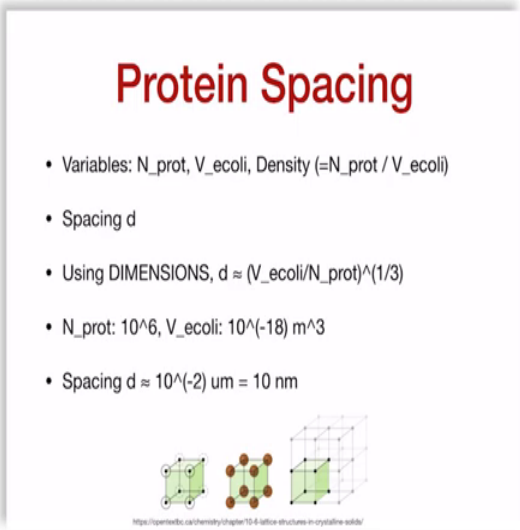
Again using my calculator I get something in the order of r^2 which is 0.5, so 0.5 squared is 0.25 micron squared into π is 0.7, 0.8 into height is 2 microns. So, I am left with 1.5 and this is micrometre cube. The dry mass of the cell is 30 percent, which means 70 is water which we have discussed earlier.

Proteins form 50 percent of the dry mass and two thirds of the protein is the cytosolic protein the remaining protein membrane proteins. So, the idea is that if you know the mass which means from the volume we get the mass of a cell as the density by volume, which is 1 gram per cm^3 by 1.5 $micrometre^3$ which is 10^3 kg per m^3 upon 1.5 into 10^{-18} meter cube this leaves us with.

So, you can use your calculator now. So the exact number is a very strange one it is 666 into, so 6.6 into 10 to the power, so I hope some of you have noticed that there is a mistake here. We are expecting 1 Picogram with the mass of the cell being the mass of the cell being only part of the picture, we want 1 multiplied by 0.3 to compensate for dry mass and by 0.5 for the protein mass and in addition by two thirds for the cytoplasmic.

All this divided by mass per protein will expectedly give us the number of proteins. You can do this calculation yourself and check if these numbers make sense because we expect 10^6 order magnitude proteins per E.coli.

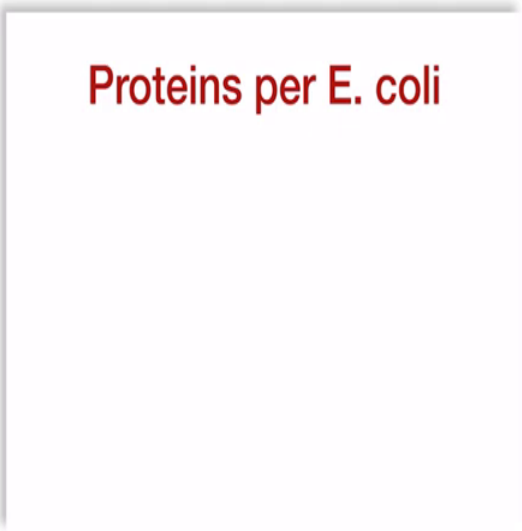
(Refer Slide Time: 8:20)

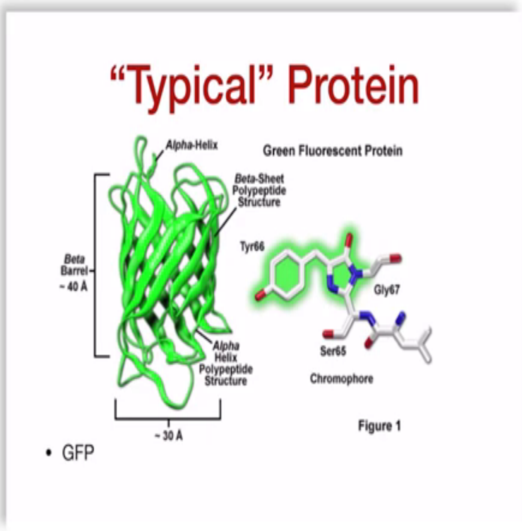


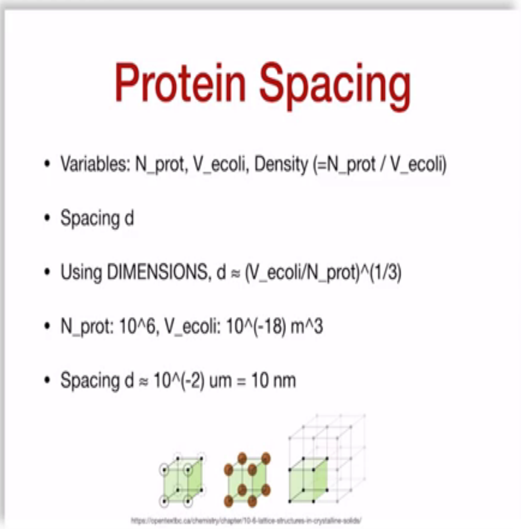
Now, if we ask the next question which is how many how *E. coli* is the cell with these proteins, then we are asking a question about protein spacing. The variables that we need are the number of proteins which you just identified as 10^6 the volume of an *E. coli* cell which is a femtoliter and the density which is the number of proteins per volume of *E. coli* that will give us the volume occupied by a single protein. But we want the spacing.

So, using dimensions and simple numerical literacy of what these mean, if you take the volume per number of protein and the cube root of it then we in fact have a distance in dimensions of L and units of meters or micrometres or nanometres of 10^{-18} m^3 because of the volume of 10 nanometre distance.

(Refer Slide Time: 9:33)





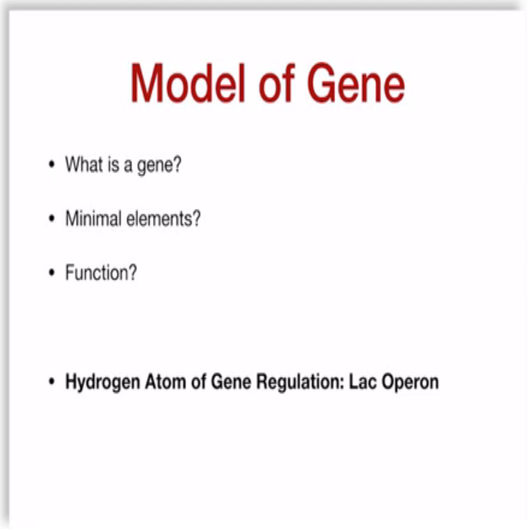


Now, it is important to bear in mind and that is why we got up this protein here that a single protein like a green fluorescent protein which is found in a jellyfish has a side length of 4 nanometres in terms of the height of the barrel and a diameter of 3 nanometres. So, roughly a few nanometres 3 to 4 nanometres that is what we argue is the size of a typical protein.

The spacing between the proteins is 10 nanometres assuming that each protein is only a point particle the protein to protein species is 10 nanometres, if however each protein has a diameter of 4 nanometres then it has a radius of 2 nanometres, you take out 2 from each side, so you are left with only 6 nanometers, this is roughly the size of another protein.

In other words spacing between proteins assuming only proteins populate the cell which, of course, is not true is that of 1 protein. In other words it is very hard to move around for proteins, because there are the additional structures like ribosomes, DNA, small ions and all of these serve to provide the crowded environment inside the cell.

(Refer Slide Time: 10:56)



So, I will now switch to a model gene, which can function as E. coli cell functions as a hydrogen atom of a cell as a hydrogen atom of a gene that gene is the Lac Operon.