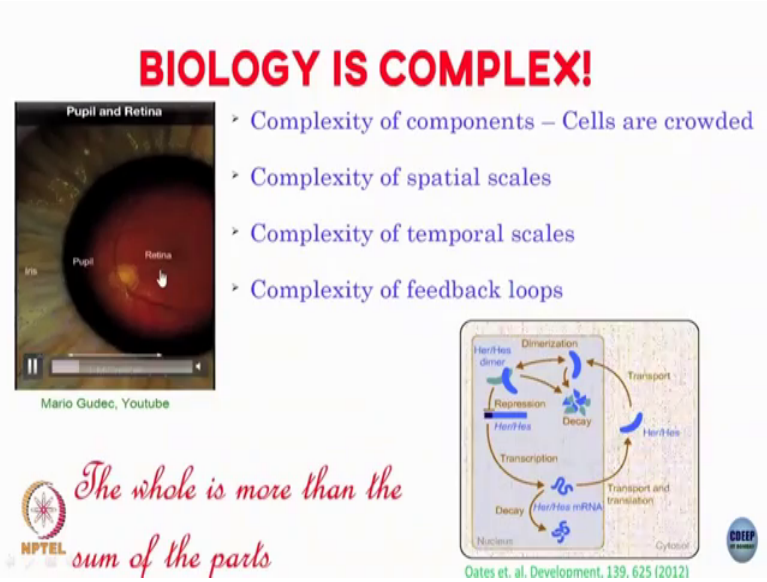
Physics of Biological Systems
Prof. Mithun Mitra
Department of Physics
Indian Institute of Technology, Bombay

Lecture – 04
Numbers and Sizes

(Refer Slide Time: 00:17)



So, the thing about biology which makes it challenging is that you know its sort of complicated and its complicated at various different levels. One is that there is a complexity of components. So, often we sort of when we look at textbook pictures of biological cells and stuff, we see very clean images and the reality is very far from that. Biology is not at all very clean cells are extremely crowded environments.

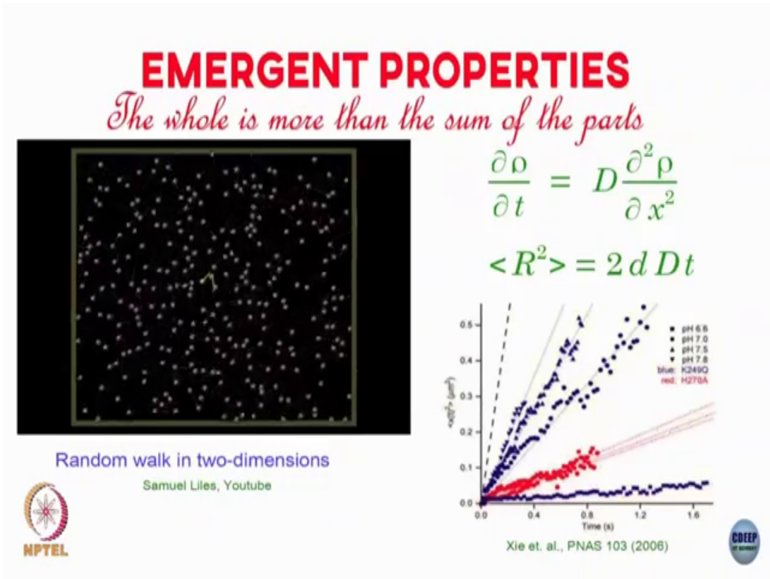
So, when we try to do modelling of such systems, we must take that crowding into effect and I will spend the reasonable time today trying to estimate how much this crowding what I mean by this crowding. So, there is a complexity of components, there is a complexity of spatial scales. So, it takes place over various scales which is what this sort of video tries to show starting from macroscopic features like you know the whole organism going up to the DNA and chromosomes and so on.

There is a complexity of temporal scales processes can take place in milliseconds seconds to going up to hours or if you think of evolution millions and years as well right. So, there is a whole range of temporal scales and depending on the process that you are interested in the sort of modeling that you would try to do would depend on whether something is taking place very quickly or very slowly so, what is the time scales that you are talking about.

Further there is a complexity of feedback loops everything talks to everything else. The DNA can produce a protein, DNA goes to RNA goes to protein that protein can come back and bind on the DNA and affect the expression pattern of the DNA itself. So, feedback loops are extremely common everywhere in biology so, this is one particular feedback loop. And the result that you sort of get out of all of this is sort of what is sum in this summed in this statement, but the whole is more than the sum of the parts right.

So, you can look at a very small component of the cell or of the organism and you try to model that, but then you have to build put everything back together in order to make sense of this whole process that you might be interested in which. And if might features what which we call emergent features which you may not have guessed by looking at the constituent systems themselves.

(Refer Slide Time: 02:35)



So, that is something that might be common to you as a physics audience which is what we call sort of emergent properties and one of the simplest examples one might think of is that of random walks. So, here is a cartoon of a random walk many many particles undergoing a simple random walk. If you look at the continuum limit of this what you get out of this is the diffusion equation and what; that means, is that starting from this very simple microscopic rules so, that an object goes for some time in one direction, it collides with another object moves in a new direction and so on.

So, this is very simple microscopic rules. Give rise to macroscopic properties which are somewhat universal. For example, in the case of a random walk the mean square displacement grows linearly with the time R square grows as I cannot point the pointer at the screen, but whatever R square grows as t right and this might be independent of the sort of microscopic details.

So, whole variety of processes would give rise to this macroscopic property where you have R square growing as t. So, often you will see examples of this sort of emergent features and as we were discussing last class sort of the ultimate puzzle would be life as an emergent property itself. It is not something we will do of course, but something good to keep in the background. So, we will try to look model these small constituent, yes.

Student: (Refer Time: 03:51) quality (Refer Time: 03:52).

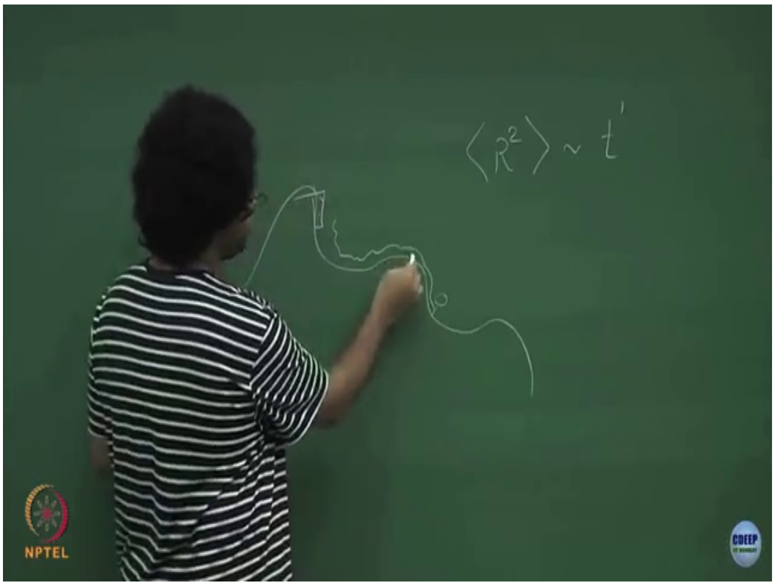
Ah.

Student: (Refer Time: 03:55).

So, D is the diffusion coefficient so, how fast it diffuses. Small d is the dimensionality of the random walk depending on whether you have a random walk in one dimensions, two dimensions or three dimensions you will get a different coefficient and t is the time. So, what is the end to end displacement in a time t and that follows this varies R square.

So, if I forget about the prefactors the main physics is contained in this R square going as t to the power of 1. And you so, this plot is actually for random walks of certain proteins on DNA. This is 2006 paper. These are this so, you have this DNA fragment.

(Refer Slide Time: 04:47)

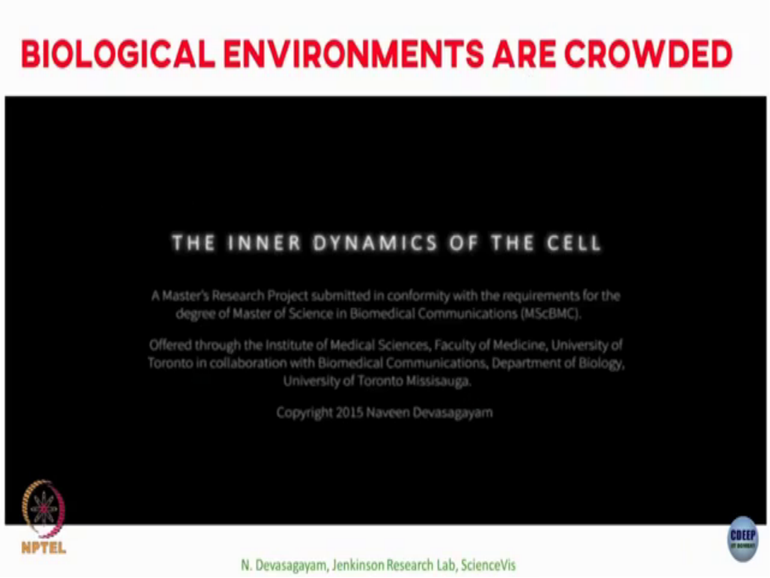


And you have proteins which bind on this DNA and then this sort of execute a random walk on this up on this DNA backbone in order to find a target site. So, this is in particular if i remember correctly this is some sort of a DNA repair protein that was being studied. So, it binds it may be there is some break in the DNA that it needs to repair. So, it needs to find the site that where it needs to go to perform its action.

So, it does that by doing a diffusion on this one dimensional backbone. So, if you look at the trajectories of this and you analyze that, you will see that it grows very linearly. So, R square grows this. So, this is x square which is the same as R square grows linearly with time the x axis is time and these different plots are at different pH conditions and so on.

So, simple things that we learn in sort of physics diffusion for example, diffusion; you would find them applicable very nicely applicable to various biological contexts all right.

(Refer Slide Time: 05:47)



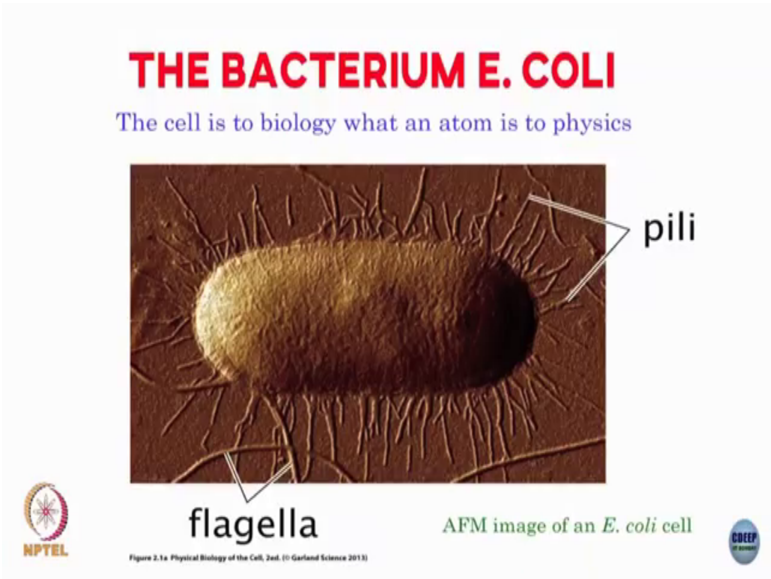
So, let us look at these things that we talked about crowding, spatial scales, temporal scales one by one. So, here is a nice video sorry here is a nice video which sort of tries to illustrate what we what I mean when I say that cells are crowded, let me right.

(Refer Slide Time: 0613)

So, this is what an artist reproduction of course, but realistic in that sense of what the interior of a cell would look like you have these; I do not know why the video quality is so bad. You have these long sort of objects which are microtubules, I will talk a little bit more about them, but and the small things are various proteins small ions which are crammed everywhere inside the cell right.

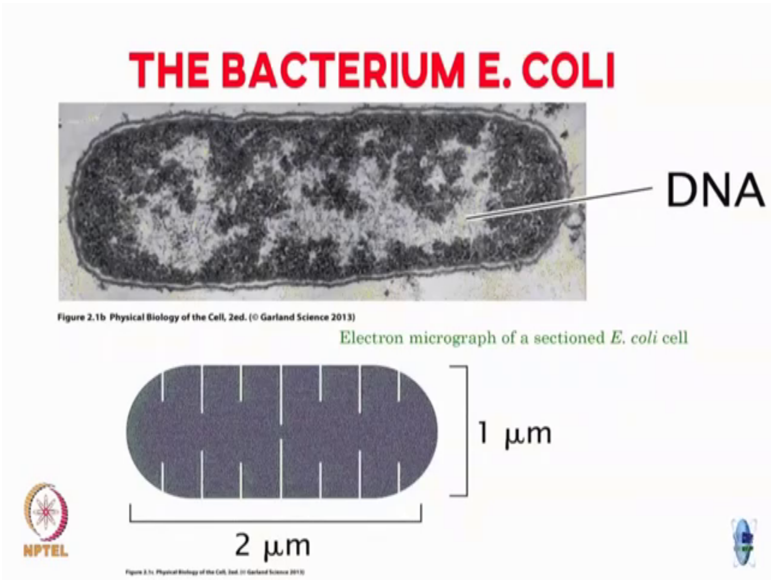
So, whenever we think of this canonical picture of a cell, I think of this I have a nucleus maybe and I have some gall g, I have some er. I have one or two things like this the reality looks something like this. If you were to transport a protein from one place to another place, you would have to do it through this extremely crowded environment. So, what we will try to do first off is try to estimate (Refer Time: 07:03) try to get an handle of the numbers of what I mean when I say the cell is extremely crowded.

(Refer Slide Time: 07:09)



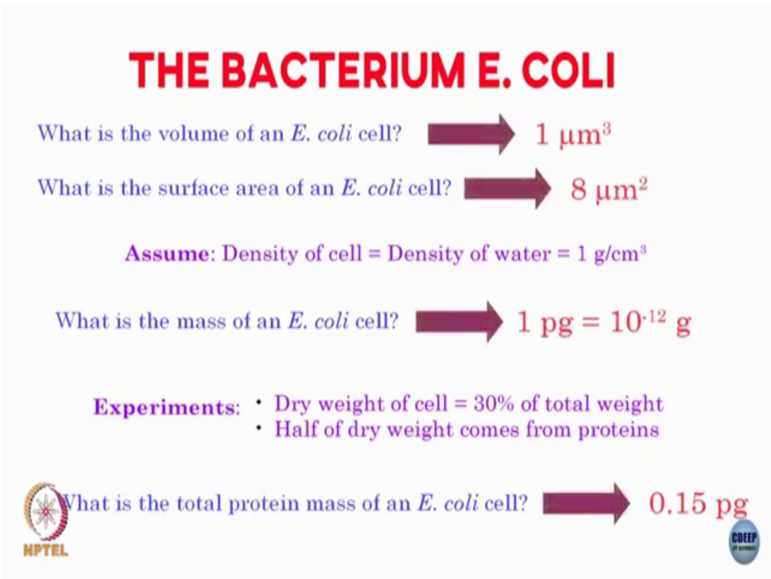
And to do that the system that we will choose is one of the canonical biological model organisms which is the E coli bacteria. So, here is a single E coli cell this is an AFM image, you will see this sort of small hair like things which are the pili or belie and then there will be this long sort of strands which are the flagella which the you which the bacteria are used to sort of move around. Again I will talk a little bit more about them.

(Refer Slide Time: 07:45)



So, if I want to get a sense of the scale of this image. So, this is the same E. coli cell an electron micrograph of that typically it would be around one micron on this side and two microns on that side ok. So, let us come back over here. So, here at least I can point. So, it is typically would be around one microns here and two microns over there right. So, this is like the standard ruler for biology of the E coli cell roughly is of the 1 micron in size all right.

(Refer Slide Time: 08:13)



So, simple the first thing that I ask is that what is there for the volume of an E coli cell right I will just say you can get a better answer, but I will just say it is one micron cube right roughly. Because I am talking about order of magnitudes, it might if I put that one and two you might get a slightly different number, but it is of the order of one micron cube. So, that is the volume of this E coli cell. What would be the surface area of this E coli cell roughly?

So, an E coli cell maybe I could say that depending on what if I want it to be a bit fancy, I could say that well it is like a cylinder and capped with two hemispherical caps right. This is roughly of the order of this is roughly two microns, this is roughly a micron. If I wanted to calculate the surface area what would the surface area be roughly, yes?

Student: (Refer Time: 09:16).

Something like that 5-6 microns square right. I think these numbers give me some 8 micron scale, it does not matter of that order all right. So, then let me assume that the density of the cell is same as the density of water. It is not really true, it is slightly higher than water, but water is a nice number 1 gram per cc.

So, I let me use that for the time being. So, then I can ask that what would be the mass of an E coli cell along with all of it is constituents. So, what would be the mass of this E coli cell? Given that it has a volume of 1 micron cubed and density of 1 grams per cc 1 nanogram. So, 10 to the power of 9 grams does that work out, anyone else?

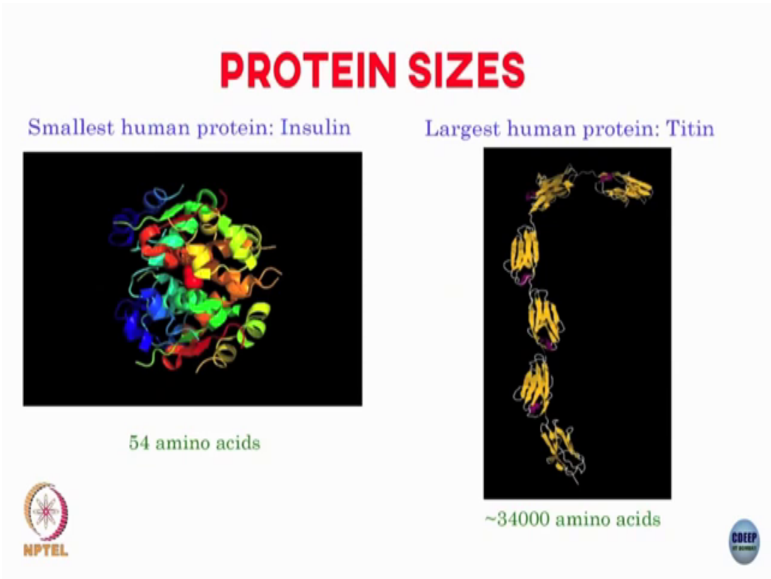
Student: Picogram.

Picogram I think that is the right thing. So, if you do it correctly it turns out to be 1 picogram, 10 to the power of minus 12 grams. So, it is very tiny as we would expect it is a single E coli cell it is 1 micron in size, it weighs 10 to the power of minus 12 grams ok. So, now, let me use an experimental result.

So, here is what experiment says that the dry weight of the cell is about 30 percent of the total weight and dry weight is everything proteins DNA all macromolecules inside the cell, the rest is water. The 70 percent is water, 30 percent is this dry weight and further experiments also tell that around roughly half of this dry weight comes from proteins. So, 15 percent of this total weight comes from proteins.

So, if I ask that what is the total protein mass of an E coli cell that would be 15 percent of this so, around 0.15 picograms. So, what I basically want to know is that here is one single E coli cell of 1 micron size, the mass of the total amount of proteins that it has is around 0.05 picograms ok. So, how many proteins is that ok? So, that is what I want to sort of estimate because I want to get it since of the crowding that is involved you want to get a sense of the numbers ok. To convert from this mass to sort of numbers I need to know typically about what are the typical protein masses and so on so ok.

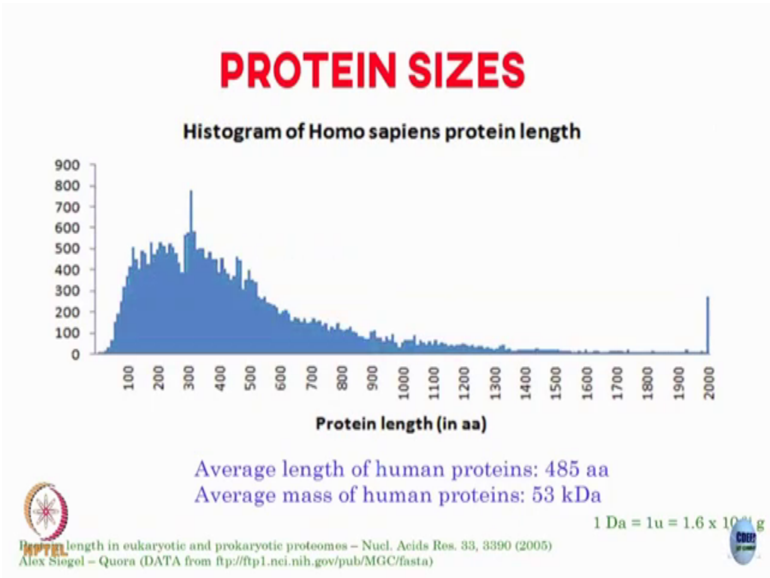
(Refer Slide Time: 11:43)



So, here is so, I digress from E coli for one slide just to show you. So, this is the smallest human protein, it is insulin which regulates your sugar and which we talked about last class. Insulin contains 54 amino acids. This is as far as they know this is the smallest human protein. Anyone knows which is the largest human protein? Any idea, no ok. So, the largest human protein is one that is found inside the muscle. It is called Titin and that has 34000 amino acids.

So, that is the range that we are talking about; 54 on one side for insulin and 34000 for titin. It is a multi unit protein titin. it helps you in extending and flexing your muscles. So, of course, we know what the proteins inside a human cell are and you can plot a histogram of these. So, here these are my two sort of extreme boundaries.

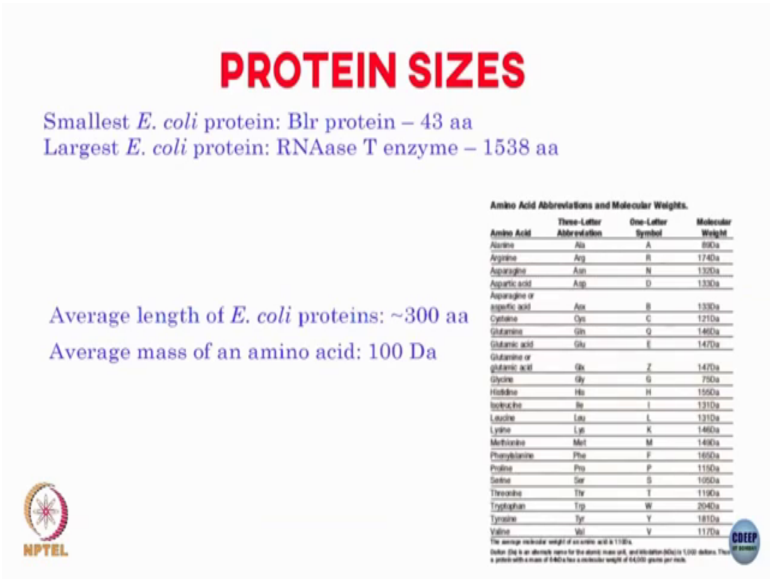
(Refer Slide Time: 12:41)



But you can plot a histogram. So, here is a histogram of homo sapiens protein lengths and here is the reference. So, here is the protein length in amino acids, it just goes up to 2000. So, you can imagine titin would be much to the right of this, but there are very few proteins that are. So, large most of the proteins reside in this range 200, 300,400, 5000 amino acids ok.

If you work out the numbers the average length of a human protein comes out to roughly around 485 amino acids and the average mass of a human protein comes out to around 53kilo Daltons and Dalton is like 1.6 into 10 to the power of minus 24 grams ok. So, it is a 1 ame so, that is for humans. So, you can do a similar analysis for E coli and I will expect you to do it as part of the assignment, I will upload tonight today.

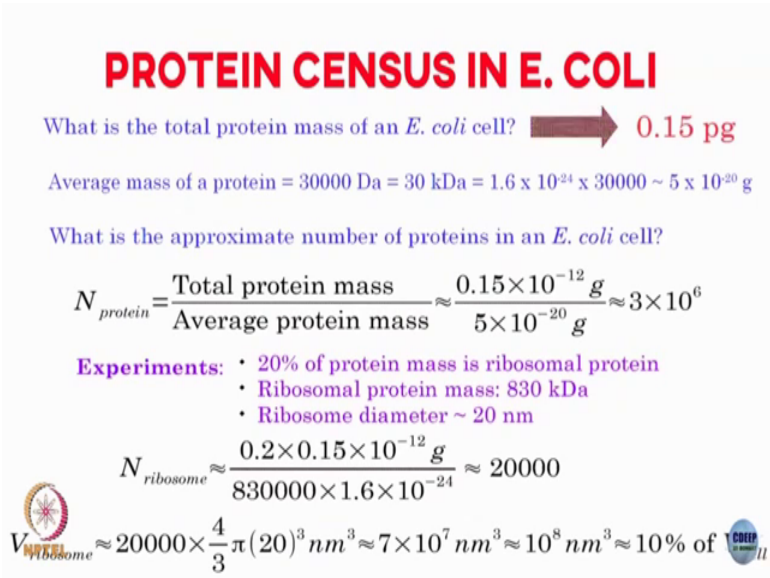
(Refer Slide Time: 13:41)



So, if you do it for E coli. The smallest E coli protein is 43 amino acids, the largest E coli protein is around 1538 amino acids. If you calculate the average length, that comes out around 300 roughly. So, it is not that different from a human. So, it is not that proteins get much larger as you grow the organism grows in complexity the average size is an E coli protein is 300 the average size of human protein is around 485 and you can further. So, I will take this number the average size.

Further I will take the average mass of any amino acids is around 1000 Daltons. This is again an approximation. So, these are the actual masses of the 20 different amino acids. So, you will see different numbers 89, 174, 132 roughly like that if you do a proper averaging, I think it comes to on 120 or so. So, again just for simplicity to make our life simple, let me say that the average mass is 100 Daltons.

(Refer Slide Time: 14:47)



So, these are the numbers that I you use that average length of a E coli protein is 300 amino acids. The average mass of an amino acid is around 100 Daltons ok, yes.

Student: (Refer Time: 14:51).

Why were the peaking at some points? So, that is not really the end, over here you could continue. So, this is like one peak like this. There are some proteins around that, but this 2000 is not the end of the human amino acids like I said it will go up to 34000. I have just shown up this plot just shows up to 2000 because the numbers become really really small after that.

So, it just so happens; I do not know why it happens it just so happens that there are some proteins around this molecular around this length of around 2000 minus also. But then there is a small peaks here also random fluctuations, I imagine ok.

So, an now armed with this information, I know that an E coli cell has a total protein mass of 0.15 picograms. I know that an average protein is 300 amino acids length and each amino acid is 100 Daltons which means that the average mass of a protein is around 30 kilo Daltons which is around 5 into 10 to the power of minus 20 gram.

So, now I could ask that therefore, what is the approximate number of proteins inside an E coli cell. So, what is that if I use these numbers quickly? 0.15 picograms and each protein is roughly 5 into 10 to the power of minus 20 grams. These are all rule of thumb estimates. We just want to get a sense of the numbers. So, what would be the number of proteins that he would get? Somebody said something?

Student: 300 million.

300 million right, you were saying something else million roughly like that 3 million; million is 6 zeroes right, million is zeroes so, 3 million roughly. So, 0.1 0.5 into 10 to the power of minus 12 grams and each protein is 5 into 10 to the power of minus 20. So, hopefully if I have done it correctly, it is around 3 million proteins. So, that immediately gives us a sort of sense of the crowding that is involved. You have the cell which is 1 micron small in that sense that contains 3 million proteins roughly.

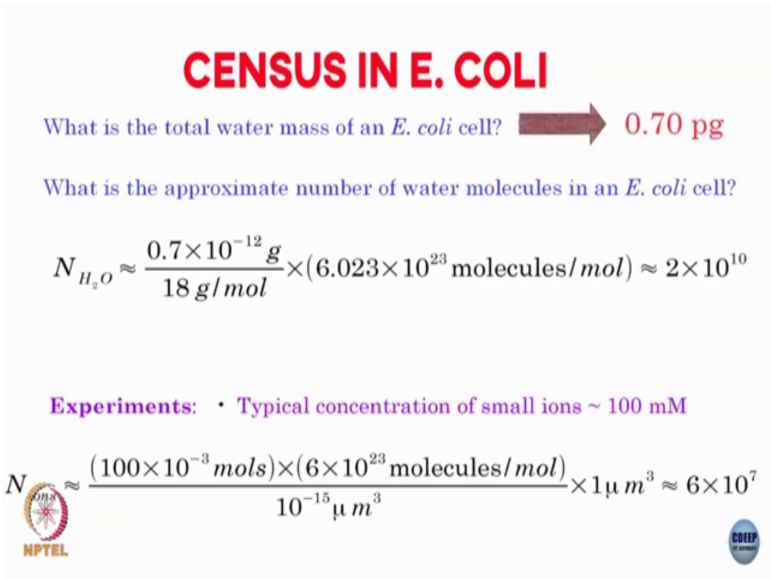
So, you can imagine that the correct picture would be to have you know my all this whole thing is going to be a soup of protein. You cannot take one step without hitting 10 other protein. We can do more estimates like this. So, for example, if I talk about ribosomes; remember we talked about ribosomes last class ribosomes bind on an RNA and it produces the protein right.

So, it is that big factory which binds on to RNA and produces the minor acid, joins together the minor acids to form the proteins. So, the protein is the ribosome and again from experiments, we know that around 20 percent of protein mass is ribosomal proteins and the ribosomal protein is roughly 830 kilo Daltons and it has a diameter of roughly 20 nanometers.

So, you then estimate the number of ribosomes that each equalise cell would contain a single E coli cell. So, this is the total number of all possible protein is combined, this is one single type of protein the ribosome. So, that single E coli cell would contain around 20000 ribosomes if you put in these numbers. You could also ask if you know the diameter of the ribosome, you could also ask that well what is the volume fraction of this E coli cell. Let us occupied by these ribosomal proteins and it turns out again if you put in these numbers. You assume I have assumed a sphere, you know these are all order of estimates you could do better.

So, if I have 20000 ribosomes, it sort of turns out that the ribosomal volume is around 10 to the power of 8 nanometer cubed. It is roughly around 10 percent of the E coli cell. It is only one single type of protein the ribosome admittedly it is a large rather large protein, but still ok. So, 3 million proteins in total 2000 ribosomes in particular.

(Refer Slide Time: 19:03)



All right, we can just talk a little bit about water. Let us remember we said the 30 percent of the cell is 30 percent of the dry weight is proteins and other stuff macromolecules basically remaining 70 percent was water. So, you can estimate. So, water is 0.7 picograms.

Therefore and then you can estimate what is the approximate number of water molecules. What is the approximate number of water molecules? What would be the range million, billion, trillion? What is the molecular weight of water?

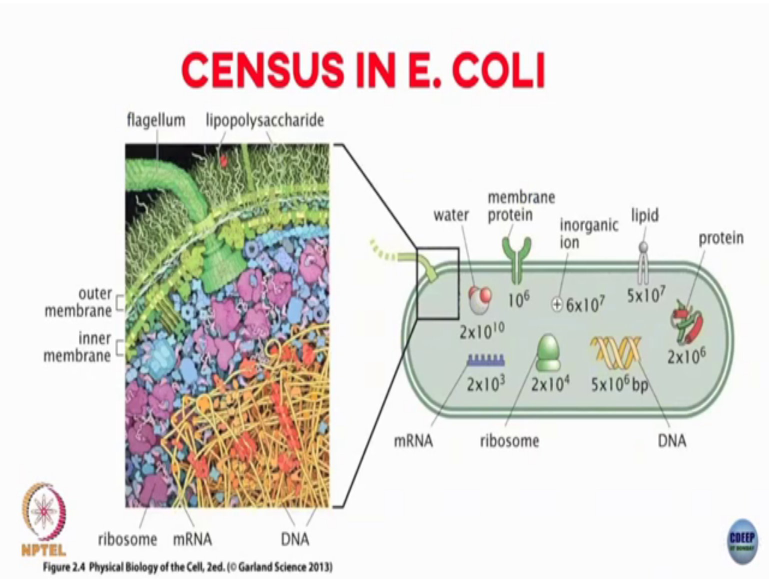
Student: 18.

18 grams per mole right, 18 grams per moles. You convert that 1 mole is 6 into 10 to the power of 23 number of molecules. So, if you put in all of this 18 grams per mole, 6 into 10

power 23.7 picograms. You get 2 into 10 to the power of 10 which is 20 billion; 20 billion water molecules, 6 million proteins, 20 billion water.

You could also estimate small ions like sodium, chlorine things like that. The typical concentrations of small ions around roughly 100 milli molar. Again if you put in that number and you calculate the number of small ions that comes out to around 60 million 6 million protein, 60 million small ions, 20 billion water molecules; all sitting inside that 1 micron cubed volume.

(Refer Slide Time: 20:37)



So, here is if we go to Rob Phillips, he has an estimate for all other different types of species RNA, DNA, lipids, proteins, you know organic ions which we calculated and so on. So, the inside of a cell actually looks well even this is much more dilute than what it actually is, but the point is that it is extremely crowded.

So, whenever you are trying to do physics of objects within this within an environment like this, you need to take into account this crowding. It will affect the physics if you want to go from one point to another through diffusion and you are living in such a crowded environment that will affect the timescales of processes that you are calculating. So, that is something to keep in mind always.

If you are not, you should at least realize that is an approximation that you have made and that is subject to changes if you take into account this crowding. It also leads to various other interesting effects like the depletion forces and so on which will again talk about as the course progresses.

(Refer Slide Time: 21:39)

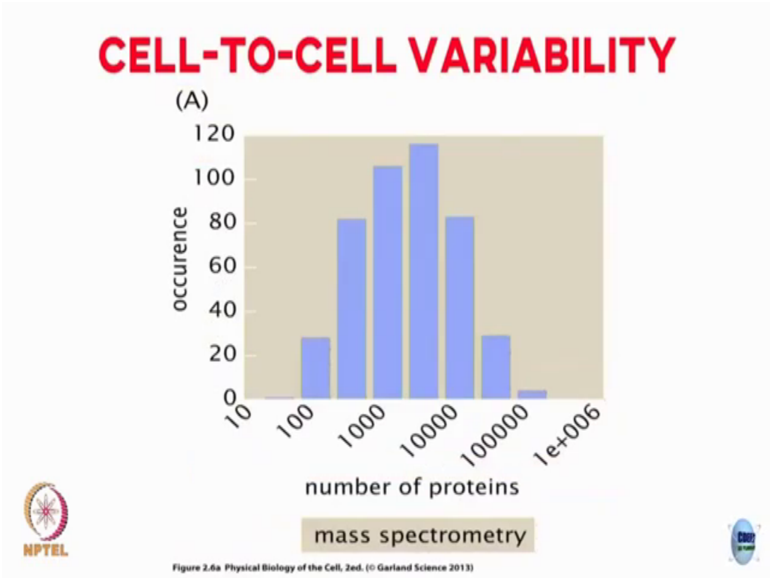
Table 2.1: Observed macromolecular census of an *E. coli* cell. (Data from F. C. Neidhardt et al., Physiology of the Bacterial Cell, Sinauer Associates, 1990 and M. Schaechter et al., Microbe, ASM Press, 2006.)

Substance	% of total dry weight	Number of molecules
Macromolecules		
Protein	55.0	2.4×10^6
RNA	20.4	
23S RNA	10.6	19,000
16S RNA	5.5	19,000
5S RNA	0.4	19,000
Transfer RNA (4S)	2.9	200,000
Messenger RNA	0.8	1,400
Phospholipid	9.1	22×10^6
Lipopolysaccharide (outer membrane)	3.4	1.2×10^6
DNA	3.1	2
Murein (cell wall)	2.5	1
Glycogen (sugar storage)	2.5	4,360
Total macromolecules	**96.1**	
Small molecules		
Metabolites, building blocks, etc.	2.9	
Inorganic ions	1.0	
Total small molecules	**3.9**	

Table 2.1 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

So, this is just that in numbers this is again from Rob Phillips a. It has various number of molecules of various species transfer RNA phospholipids this and that why did we not. So, here is another important point.

(Refer Slide Time: 21:53)

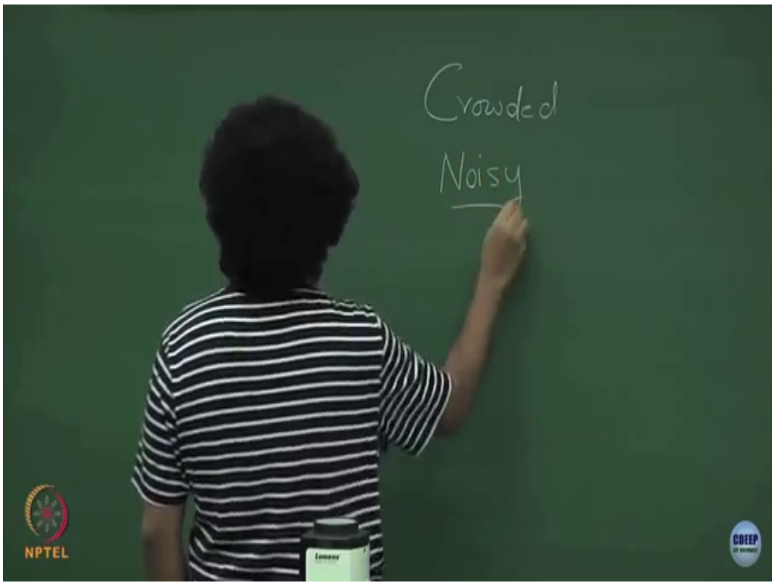


So, why did we just do a rule of thumb estimate? Because it does not make sense to do too accurate on estimate and that is because these numbers are constantly sort of. It is not that these numbers are sort of set in stone. There is a wide variability in these numbers. If you look at one cell versus another or even if you looked at a single cell at different points in it is life cycle and you estimate at the number of proteins.

So, here is a histogram of that you will see that many cells this is the maximum, but there are many cells which are very few proteins many cells if you have a large number of proteins ok.

So, there is a huge variability in the number of proteins and there is again a very generic feature biological systems are extremely noisy.

(Refer Slide Time: 22:45)



So, they are crowded, they are crowded, they are noisy.