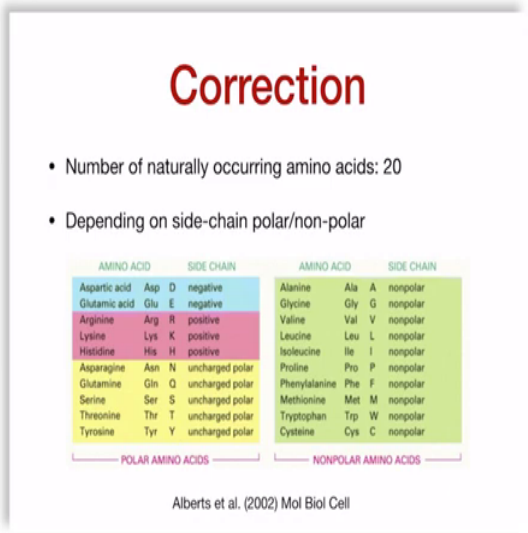
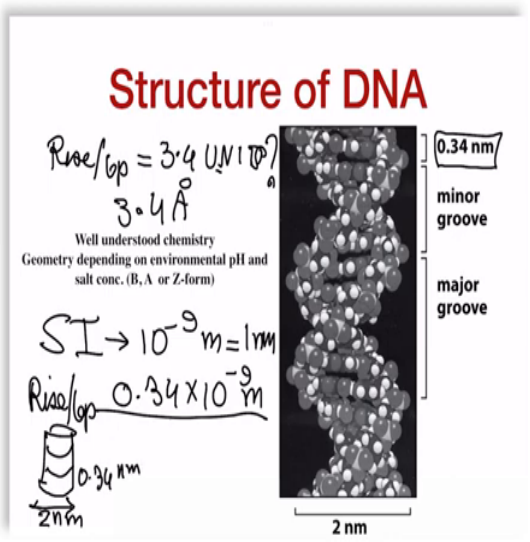
Cellular Biophysics
Professor Chaitanya Athale
Department of Biology
Indian Institute of Science Education and Research, Pune
Quantifying DNA and Chromatin

(Refer Slide Time: 0:17)



Hi, in the previous section I made a small error the number of naturally occurring amino acids are 20, this is a correction this depends on the side chain that we can classify them into two groups of polar or non-polar. This is a typical table you will find in most molecular biology textbooks and this one is from Bruce Albert's et al company. It is 20 natural occurring amino acids, not 24.

(Refer Slide Time: 0:50)

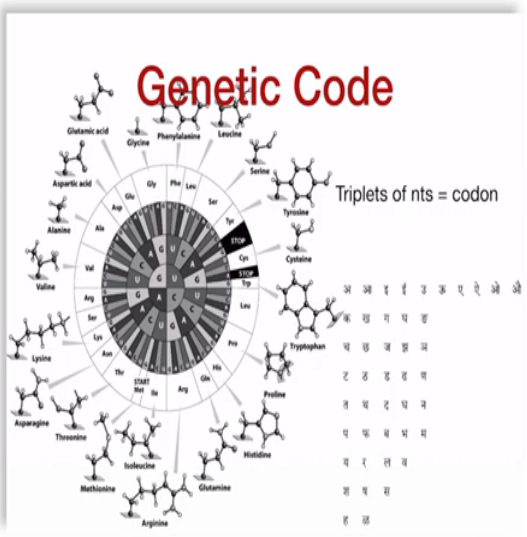


So, we move to another important biology, we talked about the size of Lipids, now we will talk about size of DNA and please remember that I keep insisting on nanometre because most biologists study the rise per base pair as 3.4 and then they do not remember the units. Since this is 0.34 nanometres, it should tell you that 3.4 refers to angstroms, but in SI units we are talking about 10^{-9} meters is 1 nanometer and 0.34 into 10^{-9} meters is the rise per base pair.

You might also be benefited by remembering that the diameter of the double helixs if you treat it as a simple cylinder is 2 nanometres. So, each brick is layered at 0.34 nanometres and the diameter is 200 meters, whichever radius is 1 nanometre depends on how you think about it, we are all familiar with the B DNA and that is the most commonly occurring DNA, which has a major anomaly depending on environment pH, salt concentration.

You may find B, A or Z forms. We are not going to worry so much about the less common forms and assume for our further calculations. Remember this is quantification and order of magnitude estimates biology by numbers. So, we are going to assume that it is B DNA largely.

(Refer Slide Time: 3:08)

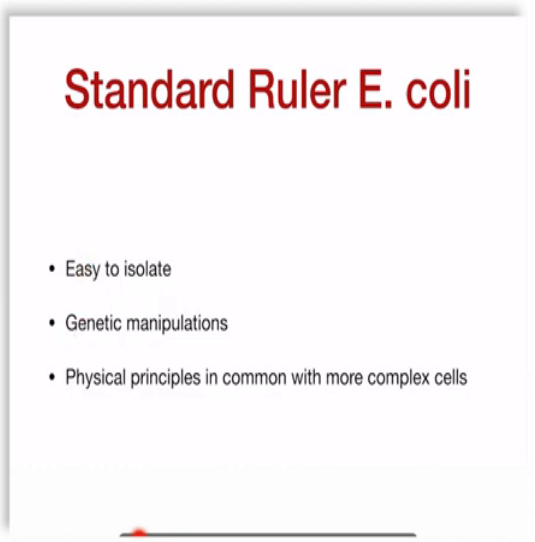


The amino acid genetic code is connected, the amino acids are connected to the genetic code through this beautiful representation in wheel. This tell us that as we go from inside outwards, for example, the start codon method is AUG and many of you studying molecular biology know this AUG. You can also AC and U, you can also find these producing isoleucine.

So, in a sense we know that a combination of triplets AAA, AAG will both result in lysine, this is the redundancy and some amino acids like GC and then UCAG are encoded by multiple such redundant codons. This codon to amino acids forms in some senses of the genetic code, it is the secret language of genes and proteins. So, triplets of nucleotides are a codon and they encode in a redundant manner for amino acids.

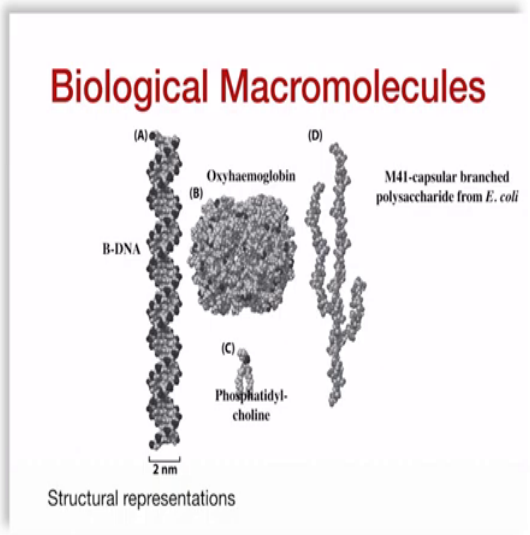
Some of them like UGA and UAA and UAG encode for stop. This is translational stop, there should be no more translation at and beyond this point. In a way it is like learning a language. The Barakhadi of Dev Nagri or Urdu or Tamil or Kanada or Telugu, or any language German, Spanish, Italian, Tagalog in Philippines, the point is that if you want to study molecular biology you need to understand at least the basis of the genetics and if you spend a lot of time in it you usually end up remembering it.

(Refer Slide Time: 5:41)



The standard ruler of E. coli which we said is a very well understood organism is also easy to isolate, possible to do genetic manipulations and physical principles are in common with more complex cells.

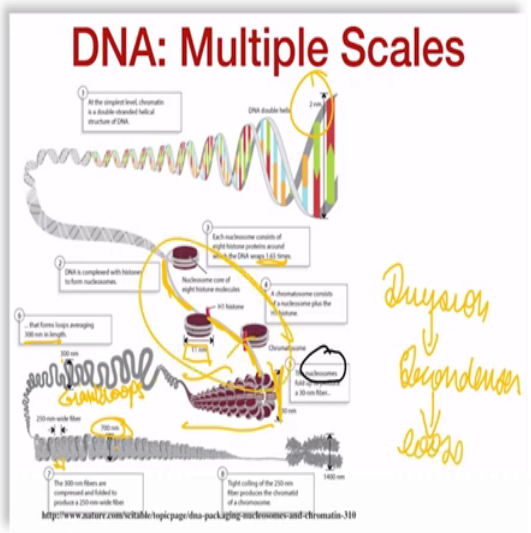
(Refer Slide Time: 5:56)



The biological macromolecules that make it up consists of these A, B, C, D and if I ask you to guess what these are it is probably easy for you to figure out that A is DNA, B it is a protein, it is a large protein, in fact, as you can probably see. It is Oxyhaemoglobin consists of the tetramer. C is a lipid phosphatidyl choline and D is one of those molecules for macromolecules whose biochemistry is even less understood than lipids.

Namely polysaccharide, this is a M41 capsular branched polysaccharide from *E. coli* the bacteria. These structural representations are important because they tell us the atomic positions, but our goal in this qualitative, this idea of quantification using rough approximations is that we use the approximate size scale to get an idea of how things relate to each other to then develop an intuition.

(Refer Slide Time: 7:06)



So, DNA in fact especially in eukaryotic cells is considered to be organized at multiple scales, so we are familiar with the 2 nanometre diameter and 0.34 nanometre rise per base pair, the helix B size, B form. And at its simplest level it is this that forms the chromosomes, even in human beings. But it is not available just as naked DNA or plain DNA. It is organized around nucleosomes.

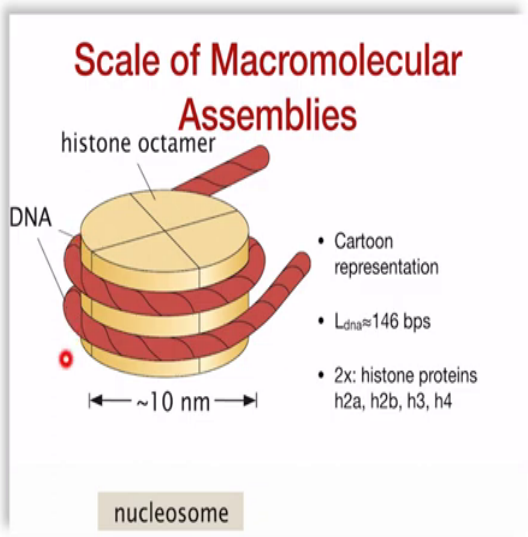
These nucleosomes form a core of 8 histone molecules around which this is wrapped in a one, two and a half. So, one is completed, two is completed and half. At the clip kind of molecule that acts on these histone octamer is the H1 histone or sometimes called the link of histone. The diameter of such a histone complex with DNA, the nucleosome as it is also sometimes called, is 11 nanometres.

It forms a sort of cylindrical geometry and is separated by a length of DNA, so it is 1.65 times or not, 2.65 times or 1.65 times. The linker DNA that joins each histone is thought to be a typically an average length. This then wraps into a more complex structure, which forms a 30 nanometre filament and it is called that because the diameter of that filament is 30 manometers.

That in turn forms loops averaging 300 manometers in length and these are the loops sometimes called giant loops, which then themselves are organized into 250 nanometre wide fibbers. This thickness of these fibbers with a diameter of 700 manometers. So, we went from 2 to 700 manometers. This is what the form that we can potentially find chromatin in.

Finally of 1.4 micron diameter for the two arms are called chromosome in mammalian cells and that is really astounding. There is such a high degree of compaction and arrangement and resolution of fine structure that it would be amazing how anything works because if it is so well organized at the time of division, at the time of mitosis, the chromatin de-condenses and loses most of this organization and becomes loose. We think it is either in this state or in 30 nanometre state and it reassembles again.

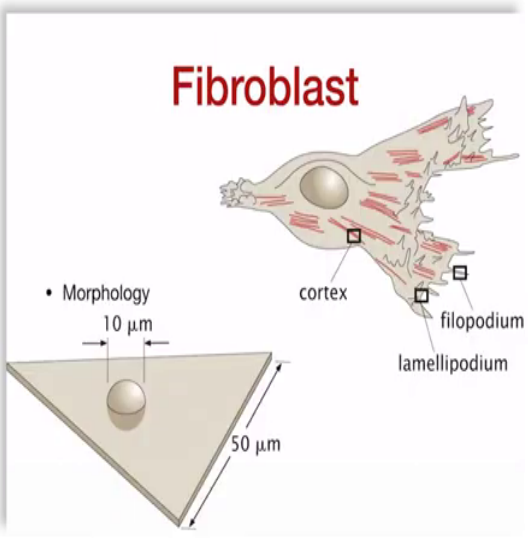
(Refer Slide Time: 10:27)



So, this cyclical assembly and reassembly is a very important point in cell biology. The linker DNA is of a 146 base pair, you could round off to 150 base pairs. There are two each copies of h2a, h2b, h3 and h4 that is the octamer 1, 2, 3, 4 into 2, 8 and it is linked at the exit points of the 1.65 rounds by a h1 histone diameter 10 manometers.

And if you assume that this is 2 + 2 nanometres, then it is 4 maybe about 6 nanometres in height. These numbers are useful for calculating simple things about the number of nucleosomes and the size of nucleosomal DNA as compared to plane DNA.

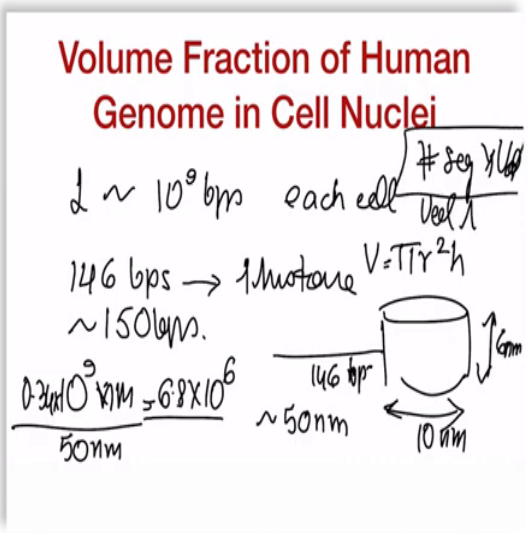
(Refer Slide Time: 11:22)



As far as cells go so far we have talked about the cartoon representation of E. coli, now let us talk about a cartoon representation of more complex cell, a larger cell this is the fibroblast. This fibroblast I said earlier it looks like a pitcher flies. It has cortical fibres which are made of actin that extend into these extensions called filopodia and lamellipodia.

They have a bulging nucleus in the center and it kind of forms almost a two-dimensional structure, so you can approximate it as a two-dimensional structure that has a bulging nucleus in the middle with a diameter of 10 microns and one side of this triangle is 50 microns. This means that we can calculate the approximate area and even volume of the nucleus as well as the cell.

(Refer Slide Time: 12:29)



So, now we can ask the question what volume fraction of the human genome forms the volume of the nucleus? How much of the cell nucleus is occupied by the genome? To do that we must know the approximate size of the human genome. So, you need to look this up but L human genome is and I remember the meaning of this sign, it is an approximation is 10^9 base pairs, each cell has this much DNA.

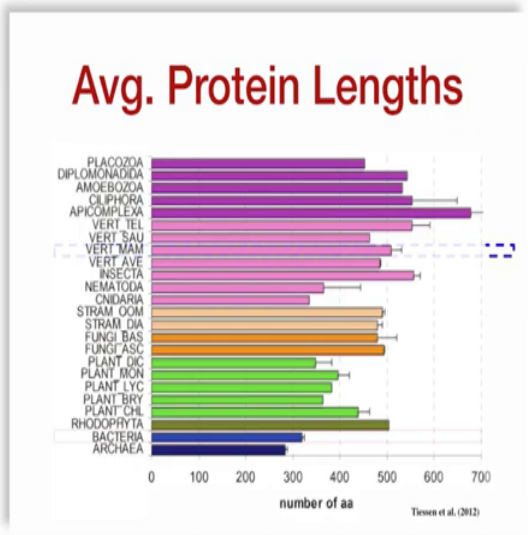
We said that 146 base pair is associated with one histone, so we call this 150 base pairs. The diameter of such a structure is around 10 nanometers. The length of the 146 base pair DNA is 146 x 0.34 nanometers which comes let us look at our calculators 146 x 0.34 is 49.64 so I am going to say 50 nanometers.

So, we have 10^9 base pairs into 0.34 nanometers divided by 150 nanometer segments, which gives us 6.8 x 10^6 segments. If we say that the volume of this is about is the volume of a cylinder which is,

$$vol_{cylinder} = \pi r^2 h$$

and the height is 6 nanometers then our genome volume is the number of segments into volume of the nucleosome and we want to know the ratio of this to the volume of the nucleus. I think you have understood the method by which we do this and I will expect you to solve this on your own.

(Refer Slide Time: 15:32)



We now come again after DNA to proteins.