Cell Biology: Cellular Organization, Division. And Processes

Prof. Shikha Laloraya

Department of Biochemistry

Indian Institute of Science – Bangalore

Lecture 22

Chromatin organization

Hi, I am Shikha Laloraya, Professor of Biochemistry at IISc. Welcome to this lecture on chromatin organization in the ongoing course on Cell Biology: cellular organization division and processes.

(refer time: 00:41)

Today, we will discuss chromatin organization, that is, how is DNA packaged in the nucleus? The conundrum is that human cells have a lot of DNA that needs to fit inside the tiny nucleus. Human cells have a genome size of 3000 million base pairs, in 23 chromosomes. As you know, the somatic cells are diploid and so, there are 23 pairs of chromosomes, and the total extended length of DNA in a human cell is roughly about 2 meters. But this has to fit within a nucleus having a diameter of only 5 to 10 microns. So, how does this happen? To achieve this, the DNA of eukaryotic cells is bound to basic proteins termed histones that help package it to fit within the nucleus in a highly condensed state. The complex of DNA with histones and non-histone proteins in the nucleus is termed chromatin. There are four histone proteins that form an octamer around which the DNA is wrapped. So, in the chromatin in fact there is twice as much protein as there is DNA.

(refer time: 02:11)

Now this is a painting of the nucleus by David Goodsell and it shows actually the crowded environment within the nucleus. So, here you can see the DNA being replicated by DNA polymerase, the thing in purple here in the middle, and the parental DNA strand is shown over here and also we can see the two daughter strands in white, which are coming out on the other side. Also, you might note the chromatin fibers on either side of this assembly, that is on either side of the replisome. And these fibers as you can see have arrays of this ball-like organization, which, as I will explain to you, are nucleosomes. And of course, there are lot many other molecules present inside the nucleus. So, it is a very complicated milieu inside the nucleus.

(refer time: 03:26)

The most basic level of chromosome organization is the nucleosome. EM analysis of the interphase nuclei reveal a fiber with a diameter of about 30 nanometers; you might have seen this in your books. And when the chromatin is subjected to treatments that allow it to unfold partially then it appears as a thinner “beads on a string” structure that is about 10 nanometers in diameter. In this “beads on a string” structure, the string of course is DNA and each bead is a nucleosome core particle.

The nucleosome core particle consists of DNA, which is bound around the proteinaceous core which is formed of histone proteins. Each nucleosome core particle consists of 8 histone proteins. So, there are two molecules of each histone protein inside this histone octamer, which makes up the nucleosome. The histones present in the nucleosome are histone H2A, H2B, H3, and H4. Now double-stranded DNA that is 147 nucleotide pairs long is part of this core nucleosome. This histone octamer forms a core around which the double-stranded DNA is wrapped around 1.67 turns. So, it is wrapped not exactly two but a little less than two turns, 1.67 is the calculated number of turns, based on the length of 147 base pairs.

(refer time: 05:27)

Now here is a schematic representation of the assembly of core histones into the nucleosome. Histones here are shown H3, H4, H2A and H2B; they are relatively small proteins about 102 to 135 amino acids and they share a common structural motif, which is known as the histone fold. This is formed of three alpha helices, which you can see here are connected by loops. Each of the core histones also has an N-terminal tail that can undergo several forms of covalent post-translational modifications and these tails are highly flexible and they project outwards from the nucleosome. The H3-H4 dimer is formed, and the H2A-H2B dimer are formed from the histone “handshake” that is these two molecules, which are having this motif of the histone fold, they come together in this configuration, which is referred to as a histone handshake interaction. And then they all come together and form this histone octamer. So, here you can see the complete histone octamer with the DNA wound around it.

(refer time: 07:05)

Each nucleosome core particle is separated from the next one by linker DNA, which can vary in length from a few nucleotide pairs up to about 80. This organization of the chromatin fiber was deduced by digestion of chromatin with nucleases that cleave the DNA by cutting the DNA which is between the nucleosomes. So, in this experiment after micrococcal nucleus digestion, the linker DNA between the nucleosome core particles gets degraded by the enzyme .

The nuclease can actually easily cut the exposed linker DNA, but not the DNA which is wrapped tightly around the nucleosome core. Hence after the dissociation of this core particle into its protein components and the DNA, the length of the DNA and that was wrapped around the core can be deduced by gel electrophoresis and it was found that this protected length is 147 nucleotide pairs.

(refer time: 08:25)

So the histone proteins as I already mentioned, in the nucleosome core, there are four different types and there are two molecules of each. So, here is the structure of a nucleosome from the protein data bank and you can see that there are 8 histone proteins, they are represented as these tubes of different colours and that in fact follows their polypeptide backbone, and here the DNA is seen as these thinner tubes or lines and they are outside and they are circling around this octamer as you can see.

The tails of the 8 histones are seen extending outwards; you can see this short stretches projecting outward and in reality they are actually longer than what is shown here and only one of them is shown of a proper length. So you can imagine how this structure looks when all the tails are hanging out. In the nucleus these nucleosomes, they are arrayed by the millions along the DNA strands. So, this is only one of them but they are arrayed as you might have seen in the diagrams that I just showed you.

(refer time: 09:51)

Now here is another structure of the nucleosome showing the DNA in orange and the histone proteins are shown in blue, this is a side view. The histones are highly conserved proteins based on their sequence, that is they are nearly identical in all the eukaryotic organisms where they are present and even a slight modification in a histone protein sequence could be lethal or detrimental.

(refer time: 10:24)

The surface of the octamer, which is shown here on the left, has positively charged residues. As I already mentioned histones are basic proteins so that means they are rich in positively charged amino acids and this is represented with these bright blue nitrogen atoms. And these interact strongly with the negatively charged phosphate groups on the DNA which is shown here. So, on the right side you have the DNA shown and here you can see the yellow phosphorus and the red oxygen atoms.

So, they are quite exposed and so, there are nice associations between the negative charge on the DNA backbone and the positive charge on the surface of these histone molecules. So, this helps in the tight association between the DNA strand to the nucleosome core. And in order to form the nucleosome you need to realize that the DNA, which is normally a long straight molecule, it needs to be bent into these two tight circles. So, it also helps to have these stabilizing forces. Now this association or the interface between the DNA and the histones, it is quite extensive. There are 142 hydrogen bonds that are formed between the DNA and the histone core in each nucleosome. And nearly a half of these hydrogen bonds form between the amino acid backbone of the histones and the phosphodiester backbone of the DNA.

Numerous hydrophobic interactions and salt linkages are also holding the DNA and the protein complex together. As explained already the core histones are rich in lysines and arginines and their positive charges can neutralize the negatively charged DNA backbone. So, because there are so many different types of interactions holding these two molecules, sets of molecules, together, DNA of varying sequences can be bound on a histone octamer core because of these very many varied interactions.

(refer time: 13:05)

Now in this slide I have shown various levels of folding of chromatin. You can see each of the nucleosome core particle shown over here, looking a little different in this representation, this is a diagram, it is separated from the next one by this linker DNA, ok. And this linker DNA I already told you, it can vary in length from a few nucleotide pairs to up to 80. On an average the nucleosomes they repeat at intervals of about 200 nucleotide pairs, that is, the protected DNA of the nucleosome core plus the linker.

And this formation of nucleosomes, beads on a string organization, it packages the DNA molecule into a thread, which is about one third of its initial length. See here it is not at all packed in any way. So, after winding around the nucleosomes in this way it is packaged or shortened to one third of its initial length, and this is the first basic level of packing or packaging of chromatin.

This “beads on a string” 10 nanometer fiber is then further folded into a 30 nanometer fiber, and you can see this here and then this is further folded by formation of loops. Now in mitotic chromosomes these loops are formed and then the loops themselves are further folded to produce these condensed rod like metaphase chromosomes. This process also requires an important family of non-histone proteins such as the SMC, the Structural Maintenance of Chromosomes, protein complexes.

(refer time: 15:22)

Now how does the packaging into the 30 nanometer fiber occur? The most accepted model which is consistent with the available structural data is a sort of variation, which is known as the zigzag model. And the linker DNA that connects adjacent nucleosomes varies in size. So, these differences also introduce further local perturbations in this zigzag structure that is how the nucleosomes come together in the 30 nanometer fiber.

And then the presence of other DNA binding proteins and also the presence of certain DNA sequences in certain regions that are difficult to fold into nucleosomes, interrupts the 30 nanometer fiber with certain irregular features. So, it is not the smoothest sort of organization, there are slight variations along it. Now there is an additional histone known as histone H1 and this histone H1 is involved in this process of packaging of the 10 nanometer fiber into the 30 nanometer fiber.

The histone H1 is larger than the core histones and it is also less conserved. A single histone H1 molecule binds to each nucleosome and it contacts both the DNA as well as the protein. The histone H1 changes the path of the DNA as it exits from the nucleosome. And this change in the exit path in the DNA is important for compacting nucleosomal DNA to form this 30 nanometer fiber assembly.

(Video Starts: 17:31)

Now in this animation by Drew Berry and Etsuko Uno we will see how the DNA is wrapped around the histone proteins. Here you can see the double helical DNA assembly and in the background you can see nucleosomes- sort of a blurred view. So, here is a bunch of histone proteins, you can see the tails hanging out, the other half of this. So, now you have the octamer and the DNA wound around it and you can actually see the zigzag folding of the 10 nanometer fiber and you can see how it comes together to form the 30 nanometer fiber; this animation also further shows further folding of the 30 nanometer fiber. Now this is hypothetical of course, it shows some sort of spiral folding but in reality the situation may be a little more complicated because there are distinct loops which are formed, but nevertheless, there are different levels of folding, there are higher levels of folding and ultimately you get this compact mitotic chromosome, which is the most condensed form of chromatin, which is formed as cells enter mitosis. So, in metaphase you have the most compact organization of these DNA molecules.

(Video Ends: 19:03)

(refer time: 19:04)

Variations in histones can also arise from histone modifications. So, as I already mentioned, the amino terminal tails of the four core histones are not only highly conserved for each of them, they are also important for the functions of these histones in regulating chromatin organization. These tails are modified by covalent attachment of various groups to the residues present in these regions.

For example, there can be acetylation of lysines, methylation of lysines, phosphorylation of serine residues, present in these histone tails. Some of the histone modifications of the histone tails could occur just after their synthesis but before the assembly, but many of the modifications take place once the nucleosome has been assembled. These nucleosome modifications are made and removed by enzymes, which are present inside the nucleus. For example, the acetyl groups are added to the histone tails by histone acetyl transferases, or HATs, and these acetyl groups can also be removed by histone deacetylases or HDACs; a very well-known HDAC that many of you may have heard of is Sir2, it is a silent information regulator protein in yeast, which is conserved in other organisms. And this enzyme, it deacetylates H4K16 to initiate gene silencing in yeast.

(refer time: 20:58)

Shown here is a schematic representation of histone modifications; the residues in the N and C terminal tails of each of the histones are shown here. So, here are the tails for each of these histones, H2A, H2B, H3 and H4. So, in two of them you can see there is the N-terminal as well as the

C-terminal tail sticking out. The potential sites of modification are shown; you can see that these tails, they appear to be quite rich in lysines and arginines. And there are also several serine and threonine residues present in the histone tails. So, these are all sites of potential modifications and the modifications that occur on these tails are various combinations of methylation, acetylation, ubiquitylation, phosphorylation and even citrullination.

(refer time: 22:11)

Now the modifications of the histone tails have several important consequences. Of course, you know adding a new group can change the biochemical nature of that region of polypeptide and these modifications they can actually affect the stability of the 30 nanometer chromatin fiber and also the higher order structures that are formed. Histone acetylation tends to destabilize chromosome structure, chromatin structure, perhaps because adding the acetyl group removes the positive charge from the lysine thereby making it more difficult for the histones to neutralize the charges on the DNA as the chromatin is being compacted. The modified histone tails, they can also recruit specific proteins to a stretch of chromatin that has been modified in a particular way and depending on the precise tail modifications these additional proteins can either cause further compaction or they can facilitate access to the DNA of other factors.

And if we start considering the various possibilities of various combinations of modifications on the histone tails, then the number of distinct marks on each histone tail is very large because there are so many residues that can undergo different types of modifications, which may or may not be simultaneously present. It was proposed that through covalent modification of the histone tails, a given stretch of chromatin could convey a particular meaning or a message to the cell. And this was referred to as the "histone code hypothesis". According to this hypothesis, each combination of modifications on the histones would attract specific proteins that could then execute the appropriate functions. Now the histone tails are extended and accessible even when the chromatin is condensed; they are, as I mentioned, they are accessible, they are hanging out of the nucleosome core. So, they allow such modifications not only to occur, but also to be read by other proteins referred to as the reader complex that may then execute some appropriate biological function by recruiting other additional proteins.

(refer time: 25:07)

Another way to achieve variation in nucleosome composition is by incorporation of variant histones. Most eukaryotes have specialized variant core histones that are quite diverged from the main histones based on the amino acid sequence. For example, in sea urchins there are five histone H2A variants and each of them is expressed at a different time during its development. Nucleosomes that have incorporated these variant histones, may differ in stability from regular nucleosomes; I mean they may have different properties. And they may facilitate high rates of transcription or replication that occur during these early stages of development, in sea urchins. So, there are variants known for all the histones so far, except for histone H4. Another example of interesting histone variant is the centromere specific histone, CENP-A. CENP-A is a variant of the histone H3 that replaces the normal

H3 in the nucleosome, which is formed at the centromere of budding yeast. And additional proteins, proteins of the kinetochore, they bind to this nucleosome and help in attaching it to the microtubule.

(refer time: 26:50)

This is represented here, and I think we have discussed it a little bit earlier also. So, here is an example of a centromere specific histone CENP-A. So, you can see here is the centromeric DNA and in this region the nucleosomes appear to be different from all the other nucleosomes. And this is a representation showing that these nucleosomes are different because they have incorporated CENP-A instead of the regular histone H3. Now once this assembly is there, additional proteins bind to this nucleosome and then they help in attaching it to the microtubule. So, this variant helps in marking the centromeric region such that it can be recognized and attached to the microtubule.

(refer time: 27:46)

Higher levels of folding of chromatin is brought about by the formation of chromatin loops and loops are also seen in interphase cells. Now under the microscope, loops can readily be seen in lampbrush chromosomes in an amphibian oocyte. The homologous replicated chromosome pair in these cells form a highly extended structure containing a total of four replicated DNA molecules or chromatids.

And these are highly active in gene expression and they form extended chromatin loops, which are projecting from the chromosome axis that can also be readily seen by light microscopy. A given loop always has the same DNA sequence and most of the genes present in the DNA loops are being actively expressed. And the rest of the DNA, which is the majority of the DNA remains highly condensed in the chromomeres on the axis of these chromosomes and it is not expressed.

Another interesting case is of polytene chromosomes in secretory cells of fly larvae, example in the salivary gland cells of *Drosophila* larvae, multiple cycles of DNA synthesis occur without cell division and this is known as endo reduplication. And all the homologous chromosome copies are held side by side, that is in a paired configuration, creating a single large polytene chromosome, which can also be readily seen by microscopy.

These chromosomes are large and they have distinct alternating dark bands and light inter-bands and both of these regions have genes. This band-interband pattern, reflects different levels of gene expression and chromatin organization. So, the genes which are in the less compact inter-bands are being expressed more highly than those which are in the compact bands. During development, when ecdysone concentrations increase, the expression of some of the genes coding for proteins that the larvae requires for each moult or for the pupation, is induced.

And as this happens distinctive chromosome puffs are formed, and the old puffs disappear correlating with expression of new genes that become expressed and the old ones are being shut off. The puffs, they can be seen by microscopy and they arise from decondensation of a chromosome band and the puffs actually have chromatin, which is arranged again in a loop-like configuration.

(refer time: 31:10)

Additional levels of higher order chromosome organization are seen in mitotic chromosomes. So, shown here are the changes in chromosome organization starting with interphase here, going counter-clockwise, that occur as the cell enters mitosis and reaches metaphase where it is the stage where the chromosomes are most highly condensed. Formation of loops and their further folding might be important for this mitotic chromosome condensation.

Stay tuned for my next lecture on SMC proteins and chromosome organization, where we will discuss the mitotic chromosome organization. Thank you.