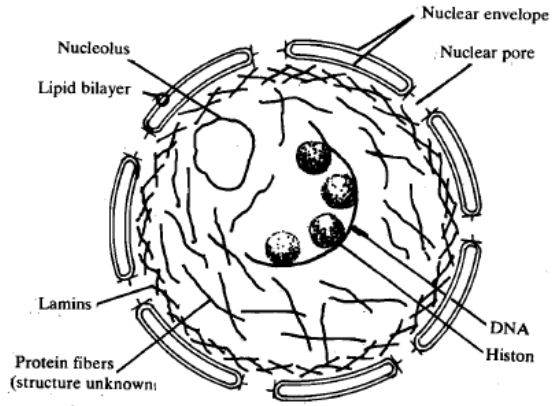
Structure of Nucleus: Nuclear envelope, Nuclear pore complex, Nucleolus; Chromatin: Euchromatin and Heterochromatin and packaging (nucleosome)

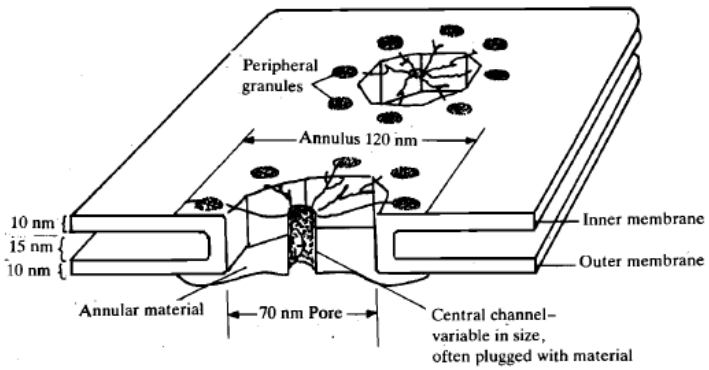
Dr Shuchismita Behera

Introduction: In eukaryotic cells—including those of animals, plants, protists, and fungi—the genomic DNA is compartmentalized within the nucleus, which is delimited by a double-membrane structure known as the nuclear envelope. This envelope delineates the boundary between the nucleoplasm and the cytoplasm and orchestrates nucleocytoplasmic transport, a highly regulated process essential for numerous cellular functions such as transcriptional regulation, signal transduction, and tumorigenesis. The compartmentalization afforded by the nucleus allows eukaryotic cells to spatially and temporally segregate transcription from translation, facilitating intricate gene regulatory mechanisms that are absent in prokaryotic systems.

Quantitative estimates suggest that a single human cell can translocate up to 80 megadaltons (MDa) of cargo per second across the nuclear envelope. This bidirectional transport of macromolecules is mediated by a sophisticated nuclear transport machinery, including the nuclear pore complex (NPC), nuclear envelope-associated proteins, and a cohort of soluble karyopherins and other transport receptors. Nucleus is composed of at least five distinct components: (i) a double membraneous nuclear envelope, which acts as barrier between nucleus and cytoplasm, (ii) a jelly like fluid nucleoplasm that fills the nuclear space, (iii) nucleolus (iv) chromatin fibres, and (v) nuclear matrix that provides skeletal framework to the nucleus (Fig. 1). Chromatin is dispersed throughout the nucleus except at the time of cell division when chromatin becomes condensed into discrete structures known as chromosomes

a





b

Fig 1 (a): The eucaryotic nucleus. A generalised diagram. The double-layered envelape, nucleolus, nuclear pores, and fibrous lamins are prominent nuclear structures. The DNA forms complexes with nonfibrous proteins called histones. Fig 1(b): Schematic diagram of the main elements present in a "typical" nuclear pore complex. Significant variations occur in the arrangement of the annular material (Source Egyankosh)

1. Nuclear membrane

The nuclear envelope is a physical barrier between nucleoplasm and cytoplasm. It is a phospholipid bilayer membrane which consists of two layers; inner and outer membrane separated by the perinuclear space. Under the electron micrographs, the outer nuclear membrane is seen to be continuous with the endoplasmic reticulum membrane (ER) and like the membrane of the rough endoplasmic reticulum (RER), the outer nuclear membrane contains cytoplasmic ribosomes over its outer surface that form the sites of protein synthesis (Figure 2). Since the ER is in continuation with the perinuclear space, both the perinuclear space and the outer nuclear membrane become the specialised regions of the endoplasmic reticulum. The inner membrane is attached to a meshwork of intermediate type V filaments called nuclear lamina. In mammals, three types of nuclear lamins line the inner surface of nuclear inner membrane, Lamin A (Mol. wt. 74 KDa), Lamin B (Mol. wt. 72 KDa) and Lamin C (Mol. wt. 62 KDa). Nuclear lamina is a dynamic structure, in mitosis, during mitotic spindle formation, these lamins undergo reversible disassembly and help in reformation of nuclear membrane in telophase. Nuclear lamina contributes to nuclear envelope stability and anchors critical proteins involved in chromatin condensation, DNA replication and gene transcription. Mutation of lamins or lamin associated inner nuclear membrane proteins can result in tissue specific diseases known as ‘nuclear envelopathies’. Mutation in lamin A and C lead to two muscular dystrophies known as Emery-Dreifuss muscular dystrophy and Limb-Girdle muscular dystrophy respectively.

Nucleus exports materials to cytoplasm including mRNA, tRNA and ribosome subunits, and imports enzymes, nucleotides, regulatory proteins etc from the cytoplasm. The nuclear membrane is perforated by nuclear pore complexes (NPC) which fuse the outer nuclear membrane and inner nuclear membrane together and creates channels for nucleocytoplasmic transport.

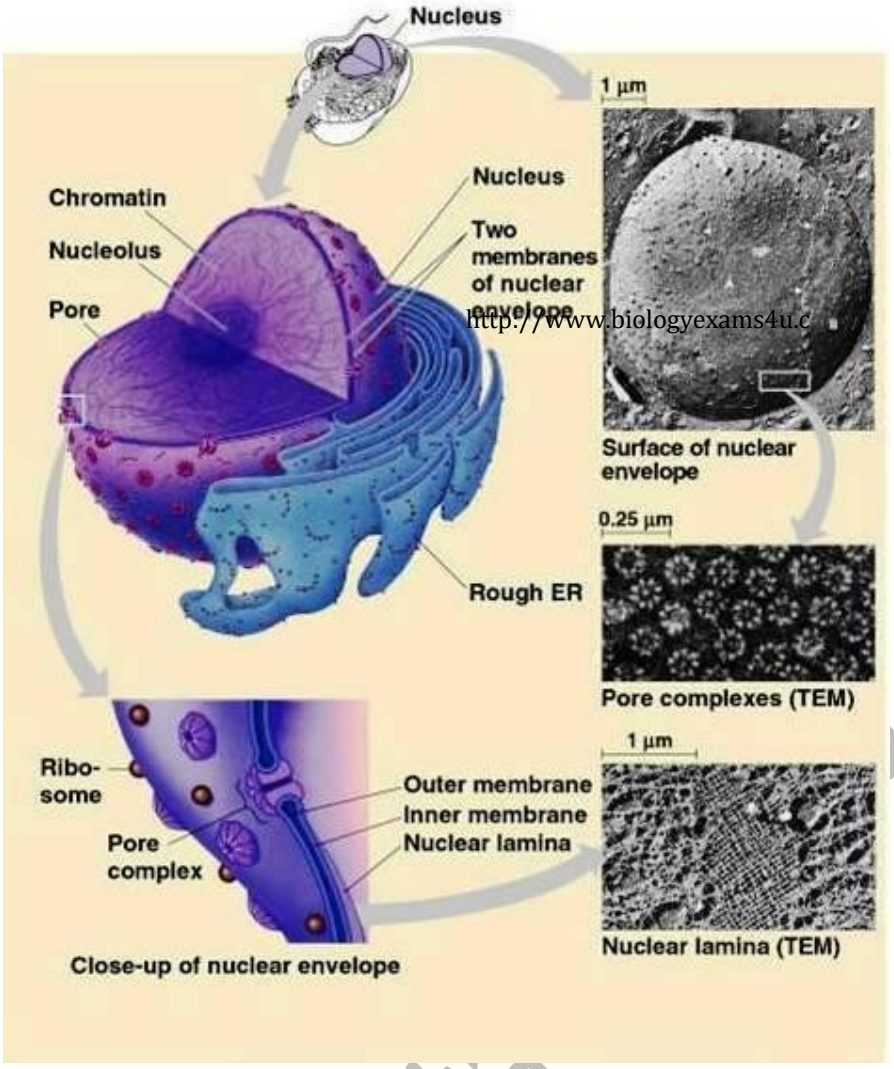


Figure 2: Diagram and pictures of Nuclear envelop and nuclear -pore complex (Source :epgpathshala)

2. Nuclear pore complex (NPC)

Nuclear pore complexes (NPC) are multiprotein suprastructures (50-112 MDa) that allows exchange of matter, energy and information between two major compartments of the cell, cytoplasm and the nucleus. The individual transport rates of NPCs are very high, about 1000 cargo transport complexes per second. All the histones and ribosomal proteins needed per cell cycle are transported through the NPC.

NPCs within nuclear envelope display variability in density and distribution between cell types. In majority of somatic cells, the nuclear pores are randomly distributed on the surface of nuclear membrane, however the NPCs are arranged in hexagonal symmetry in malphigian tubule cells of grass hoppers and in clusters in oocytes of *Xenopus laevis*.

In plants, protozoan and amphibian oocytes, NPCs account for 20% of total nuclear membrane area. In mammals, NPCs account for 5-15% of the total surface area of nuclear membrane. NPC density is determined by transcribeptional activity of the cells, more active cells have high number of NPCs. For e.g. non dividing RBCs and lymphocytes have NPC density of ~3/μm^2 whereas highly active cells like hepatocytes or brain cells have NPC density of 15-20μm^2.

About 3000-4000 NPCs are present on a mammalian cell nuclear envelope. While small sized molecules (less than 30 kDa) move freely through the nuclear pore, bigger molecules require nuclear localisation signal (NLS) and a soluble transport receptor to mediate transport through NPC. NLS is 4-8 amino acids long which are rich in positively charged amino acids like arginine and lysine and generally contain proline. The proteins with NLS are imported into nucleus by active transport.

2.1 Structural assembly of NPC

NPCs are elaborate and complicated protein assemblies. NPCs form a diffusion barrier for certain molecules and simultaneously remain highly permeable to other molecules. This remarkable selectivity is accomplished by the structural organisation of NPC. NPC morphology is conserved from yeast to humans. Using cry-electron microscopy and cry- electron tomography, NPC structure with resolution upto 6 nm has been delineated.

In vertebrates, NPC consists of ~600 individual proteins forming a structure ~125nm X ~150nm with a characteristic eight fold rotational symmetry. NPCs from various eukaryotic origins share a set of ~30 nucleoproteins (Nups). These nucleoporins (Nups) are often grouped as:

i) Transmembrane Nups, anchoring the NPC in the nuclear envelope.

ii) FG-Nups, (~1/3 of all Nups) contain phenylanine-glycine repeats (FG repeats)

iii) Structural Nups, (~1/2 of all Nups) forms a scaffold with transmembrane Nups and FG- Nups.

Nups localise on both sides of a symmetrical axis in the plane of the nuclear envelope, with the exception of some asymmetric nucleoporins. They occur in a copy number of eight or multiples of eight, explaining the octagonal symmetry of NPC. The donut shaped NPC is made up of three lattice like ring structures termed cytoplasmic ring (CR), inner ring complex (IRC) and nuclear ring (NR). The three rings are sandwiched between cytoplasmic extensions (extend ~35 nm into cytoplasm) and nuclear extensions which extend ~60 nm into nucleus and are linked by a distal ring to form a basket-like structure (Figure 3). Between cytoplasmic and nuclear rings is the 'central framework' which harbours the

central transport channel that accommodates all active transport across NPC.

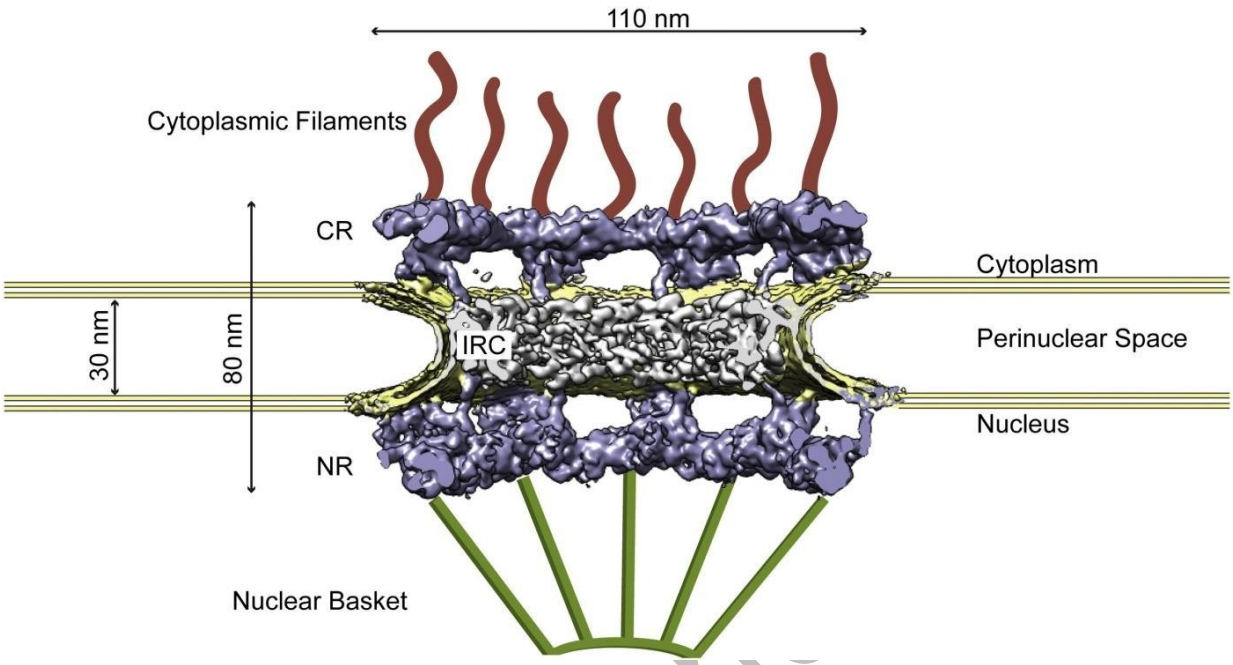


Figure 3: Overview of the nuclear pore complex (NPC). (Source: epgpathshala and *The Structure Inventory of the Nuclear Pore Complex. Thomas U. Schwartz.*)

3. Other Nuclear Components

3.1 Nucleoplasm

It is a jelly like fluid containing the soluble material of the nucleus. Various components of the nucleus can be isolated with dilute buffer or salt solutions such as proteins, RNA, RNA protein complexes and small molecules. Protein fraction of the nucleoplasm contains various enzymes involved in metabolic pathways, while other enzymes function in replication and transcription of DNA, and some proteins present are involved in the regulation of chromatin structure and function.

3.2 RNA and RNA protein complexes

Most of the RNA molecules in nucleoplasm are in association with proteins forming ribonucleoproteins of various sizes and types, for example some of these RNA molecules form **small nuclear RNAs,** which are involved in messenger RNA processing. These small RNAs never leave the nucleus and are destroyed inside the nucleus. The large RNA complexes pass through the nuclear pores to the cytoplasm

which act as a precursor of cytpplasmic ribosome, mRNA, or mRNA precursors. Low molecular weight substances such as coenzymes, nucleosides, precursors of nucleotides, metabolic intermediates and ions are also found in the nucleoplasm.

Nucleoside the nucleotide precursors are utilised in the synthesis of two nucleic acids, i.e. DNA and RNA.

4. **Role of Nuclear Envelope in Chromosome Organisation and in Transfer of Large Molecules**

Nuclear lamina, the fibrous protein network under the envelope, plays a major role in dissolution and reformation of the nuclear envelope during cell division. Chromatin fibres are attached to it at multiple Sites. The lamina1 protein bolds the condensed, inactive chromatin-at the peripheral region of the nucleus to keep it away from the unfolded, active chromatin.

Information regarding the transport of material through the nuclear pores is obtained by the knowledge of a typical mammalian cell which contains about 3,000 to 4,000 pore complexes in its nuclear envelopes.

The nuclear envelope like other cellular membranes is also selectively permeable and capable of passive diffusion, active transport and facilitated diffusion. Small membrane vesicles formed inside the nucleus fuse with the inner nuclear membrane discharging their contents to either the perinuclear space or to cytoplasm directly. The protein synthesised into the outer nuclear membrane having ribosomes on its surface is discharged into the perinuclear space and inro nucleus. Nuclear envelope is also involved in carrying out the function of oxidative phosphorylation and electron transport.

5. Nucleolus

Nucleolus (discovered by Fontana) is a non membrane bound *dynamic body* which disappears in the late prophase and reappears at nucleolar organizing region of chromosome in the telophase stage of cell division. Each nucleolus is produced by a **Nucleolar-Organizing Region** (NOR) presents on a nucleolar-organizing chromosome. All eukaryotic cells contain at least one such chromosome. Its number may be one or more (several hundred per nucleus), but 1 to 4 being the most common. The number of nucleoli per nucleus also differ. The yeast cell contains one relatively large nucleolus with respect to its nuclear volume. At the other extreme, *Xenopus oocytes* contain over 1000 nucleoli per nucleus.

It is a site of transcription of ribosomal RNA and assembly of ribosome. So, nucleolus has a high concentration of rRNA and proteins. Once the ribosomal subunits mature they come out from nucleus and released into the cytoplasm through nuclear pore.

Nucleolus structure : It consists of three major regions.

Fibrillar centers : Containing rRNA genes in the form of partly condensed chromatin.

Fibrillar component : Surrounds the fibrillar centers, which contains RNA molecules in the process of transcription.

Granular regions : Outermost regions having mature ribosomal precursor particles.

The chemical composition of the nucleolus is studied by direct chemical analysis and by cytochemical studies. Such studies have shown that nucleolus is composed of about 80% of proteins, and about 20% of a mixture of RNA and DNA.

Ribosomal subunit production:

The nucleolus is the most conspicuous structure within the interphase nucleus. It is responsible for synthesizing a large nascent precursor ribosomal RNA (pre-rRNA) that is 45S in mammals and the processing (cleavage and base modification) of this RNA to yield mature ribosomal RNA of 18S, 5.8S and 28S. The concomitant assembly of these RNAs with incoming ribosomal proteins generates small and large ribosomal subunits that then pass into the cytoplasm.

Nucleolar organizer:

Nucleolar organizers are secondary constrictions within the mitotic chromosomes of eukaryotes. The primary constriction occurs at the centromere. Historically, Heitz concluded that the number of secondary constrictions within certain chromosomes of *Drosophila funebris* is proportional to the number of nucleoli. He called these regions *sine acid thymonucleinico*. McClintock actually coined the term *nucleolar organizer* to describe the satellite segment at the tip of maize chromosome 6 that organizes the assembly of the nucleolus beginning in the telophase. As the name implies, the nucleolar organizer gives rise to the nucleolus in the telophase of mitosis. During the late telophase, RNA polymerase I begins to transcribe rRNA genes once again, and the newly synthesized rRNA likely nucleates the assembly of the nucleolus by recruiting prenucleolar bodies.

6. **CHROMOSOME**

A *chromosome* is an organized structure of DNA and protein that is found in the nucleus of eukaryotic cells. It is a single piece of coiled DNA containing coding and noncoding sequences. The word *chromosome* comes from the Greek word *chroma*, color and *soma*, body due to their property of being very strongly stained by particular dyes. A chromosomal DNA molecule contains three specific nucleotide sequences:

6.1 *Centromere*

The constricted region of linear chromosomes is known as the centromere. Although this constriction is called the centromere, it usually is not located exactly in the center of the chromosome and, in some cases, is located almost at the chromosome's end. The regions on either side of the centromere are referred to as the chromosome's arms. Centromeres help to keep chromosomes properly aligned during the complex process of cell division. The centromeres serve both as the sites of association of sister chromatids and as the attachment sites for microtubules of the mitotic spindle.

6.2 *Telomere*

Telomeres are repetitive stretches of DNA located at the ends of linear chromosomes. They protect the ends of chromosomes in a manner similar to the way the tips of shoelaces keep them from unraveling.

Origin of replication

The origin of replication (also called the *replication origin*) is a particular sequence in a chromosome at which replication is initiated.

6.3 **Chromosome number**

All eukaryotic cells have multiple linear chromosomes. Every cell maintains a characteristic number of chromosomes. Depending on the eukaryotic organism, the number of chromosomes typically varies from 2 to less than 50, but in rare instances are reach thousands. The majority of eukaryotic cells are diploid; that is, they contain two copies of each chromosome.

Morphologically chromosomes are classified according to the position of the centromere. If this is located centrally, the chromosome is metacentric, if terminal it is acrocentric, and in case the centromere is located in an intermediate position, the chromosome is sub-metacentic (Figure 4).

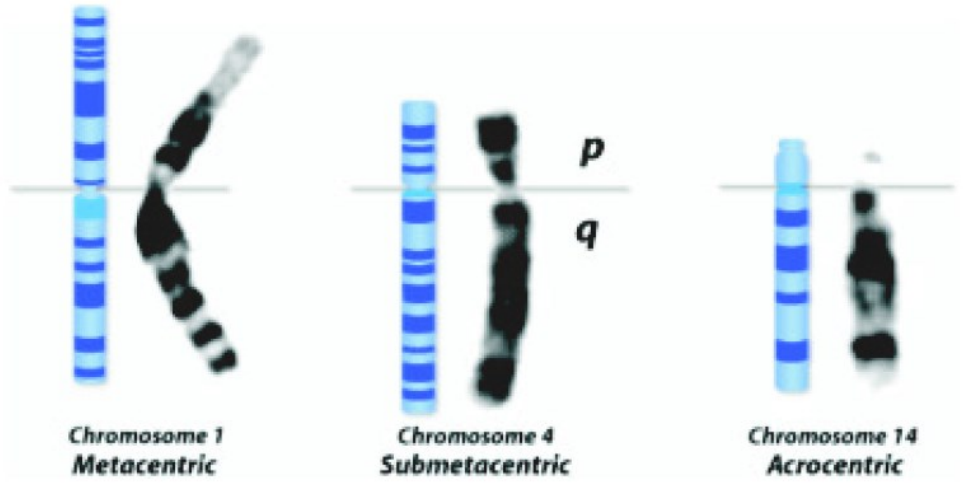


Fig. 4: Morphologically chromosomes are described as metacentric, submetacentric or acrocentric depending on the position of the centromere(Source: Egyankosh)

7. ORGANISATION OF DNA IN CHROMOSOMES

The prokaryotes such as bacteria and blue green algae lack a well-defined nucleus and possess a single double stranded circular DNA molecule that contains all the genes and is located in a special region of their cytoplasm. In addition to the chromosomes, most bacteria contain DNA of shorter chain lengths called **plasmids** which also form the closed loops. Eukaryotes have a complex organisation of membrane bounded nucleus with multiple linear chromosomes. Each chromosome contains a single double stranded helical DNA, and a set of five types of histone proteins.

The DNA molecule is incredibily long and has to fold itself in compact way so as to fit inside a cell which is only one millionth of an inch. Various proteins help in the compression of DNA molecule with out damaging it. A frame work of scaffold proteins guide the DNA molecule to compress. The DNA coils around the proteins known as "Histones" which give beaded appearance. The beaded structures are called "Nucleosomes" which get packed tightly during cell division so that the chromosomes become condensed and are visible as the cell division progresses specially at metaphase and telophase. A chromosome consists about 1/3 of DNA, 1/3 of histone protein, 1/3 of other DNA binding proteins and a small amount of RNA. The chromosome material is referred as "Chromatin" since it takes the color of the dye used while staining. The non-histone proteins are heterogenous and vary in different tissues in quantity and type. These include RNA and DNA polymerases and regulatory proteins. Histones are small basic proteins, they have a high content of the basic amino acids, arginine and lysine. Being basic, and having strong positive charge at the physiological pH, histones bind tightly to DNA, which is acidic due to the presence of phosphate group in it. There are five types of histones i.e. H2A, H2B, H3, H4 and H1. A chromatin contains almost equal number of H2A, H2B, H3 and **H4,** and about half the number of H1 molecules. Out of these, H2A, H2B, H3, and **H4,** are among the most conserved proteins known during evolution and are present in all eukaryotes. For example, the sequence of histone H3 from the rat differs from that of peas in only two amino acids, out of 102 total amino acids. Histone HI is not conserved between species and has tissue specific forms. It is loosely associated with chromatin and is involved in the maintenance of a higher order of folding of chromatin. **The complex containing two molecules each of H2A,H2B, H3 and H4 forms an** octamer **of proteins.**

Some of the properties of histones are given in following table

Histone	Molecular Weight	Amino Acid Composition	Species Variation	Number of Molecules per 200 Base Pairs of DNA
H_1	20,000	Lysine-rich	Wide	1
H_2A	13,700	Moderately lysine-rich	Fairly well-conserved	2
H_2B	13,700			2
H_3	15,700	Arginine-rich	Highly conserved	2
H_4	11,200			2

7.1 Nucleosomes and Chromatin Structure

7.1.1 Beads-on-a-string (Primary Structure): Electron microscopy of extended chromatin fibres reveals a "beads on a string", appearance, the 'beads' being the nucleosomes and the

'string' the linker DNA. Brief digestion with certain nucleases results in degradation of the linker DNA, liberating the nucleosomes as monomeric units. A nucleosome contains a histone octamer, DNA of 200 base pairs and one molecule of histone H1. The nucleosome is a flat disc shaped particle 11 nm in diameter and 5-7 nm in height. The DNA makes two complete turns around each histone octamer and these two turns are sealed off by an H1 molecule. Extensive nuclease digestion of a nucleosome removes the linker DNA and histone H1, yielding a nucleosome core particle. This particle comprises a disc-shaped histone octamer—consisting of two copies each of H2A, H2B, H3, and H4—associated with approximately 146 base pairs of DNA tightly wrapped around the histone core.

Histone H1 plays an important role in packaging the nucleosomes by further folding the nucleosome's fibre into higher levels of chromatin organisation. When nucleosomes are closely condensed forming **10nm** diameter filaments, the packing of DNA is about five to seven fold, i.e. five to seven times more compact than free DNA.

7.1.2 The 30 nm Fibre (Secondary Structure):

This is still a 1000-fold lower value than the packing ratio of DNA in metaphase, which is achieved by further folding the 10 nm diameter chromatin fibre into 30 nm diameter stage as seen in the electron micrograph. The 30 nm diameter chromatin fibre arises from the folding of the nucleosome chain into a helical solenoidal structure having about six nucleosomes per turn. The whole structure is stabilised by interactions between histone H1 molecules in neighbouring nucleosomes. The H1 molecules can interact with each other because they are located in the central hole of the solenoid. The nucleosome along with the linker H1 protein is also known as **chromatosome** (Figure 5). The DNA of a 30 nm solenoid has a packing ratio of about 40-fold (Figure 6).

More than one model had been proposed for the secondary structure of chromatin such as, solenoid, twisted-ribbon, cross-linker, and super bead models. However, the best analyzed and most acceptable ones are solenoid and zig-zag models. In the solenoid model, nucleosomes are arranged linearly in a solenoid-type helix with a bent linker DNA. For the solenoid to form an acidic patch domain on the surface of histone H2A and histone H2B, proteins of one nucleosome needs to interact with the N-terminal tails of histone H4 proteins in the nucleosomes present right next to it. Only then do all the nucleosomes interact to form a linear axis through the middle of the solenoid structure (Figure 6 and 7). The zig-zag arrangement also has certain properties which give it a unique structure. A relatively straight linker DNA connects the zig-zig arrangement of nucleosomes, unlike the bent linker in solenoid formation. In this case,nucleosomes that are not present right next to each other connect two stretches of DNA. In contrast, in the solenoid model, the connections are formed between adjacent nucleosomes.

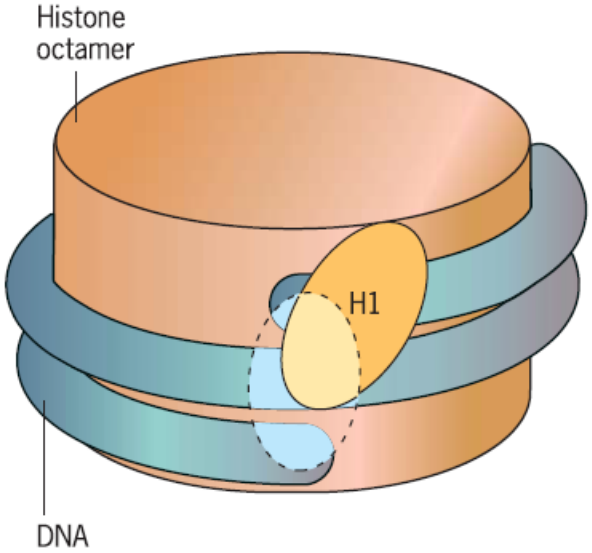


Figure 5 : Diagram showing the structure of nucleosome core particle and an associated histone H1 molecule (Source: Cell and Molecular Biology, Gerald Karp)

7.1.3 Higher Order (Tertiary Structures): A higher order structure is recognized as the tertiary structure when the genome is arranged in individual chromosomes. Essentially, chromosome depicts a tertiary arrangement. This is made due to interdigitation of independent DNA strands in a beads-on-a-string form (Figure 7 & 8). With the help of proteins that can change the structure of chromatin, one entire beads on-a-string DNA can move out of the compacted chromosome and be accessible for transcription factors to carry out transcription.

7.1.4 Looped Domains (Quaternary Structure): The next level of folding is the looped domains. It is thought that the loops of DNA are formed by special DNA-binding proteins that help specific regions of the 30 nm chromatin fibre, called nuclear matrix associated regions (NMARs) or scaffold/matrix associated regions (SMARs), to get attached with the nuclear matrix (Figure 7 & 9). The proteins that help the NMARs or SMARs to do so are called as nuclear matrix proteins. The loops vary in size quite widely, from 10 to 80 kb with an average of 40 kb. If we calculate, roughly one chromosome will contain about 2,600 looped domains. These domains allow the chromosomes to expose only a part of the DNA to transcription factors at a time. It extensively uses the chromatin remodelling proteins to help shift nucleosomes and the DNA such that, in small parts at a time, the entire actively transcribing genes on chromosome get exposed in the loops to all other molecules.

The genetic material in prokaryotes is packed into one or more double stranded DNA fibres that are folded into a compact structure known as **nucleoid** that contains about 80% DNA and rest as protein and RNA. This complex **is** attached at one side to the membrane inside the cell. The bacterial DNA is **supercoiled** and **super twisted** and is organised in about 45 loops radiating out from the central non-histone protein core that provides anchorage to DNA in the cell. Supercoiling and loop formation make the packaging of large DNA molecules into small structure to fit into the nucleus. The "naked" fibres of DNA can be seen when bacterial, viral or organelle chromosomes are examined under the electron microscope by staining the chromatin.

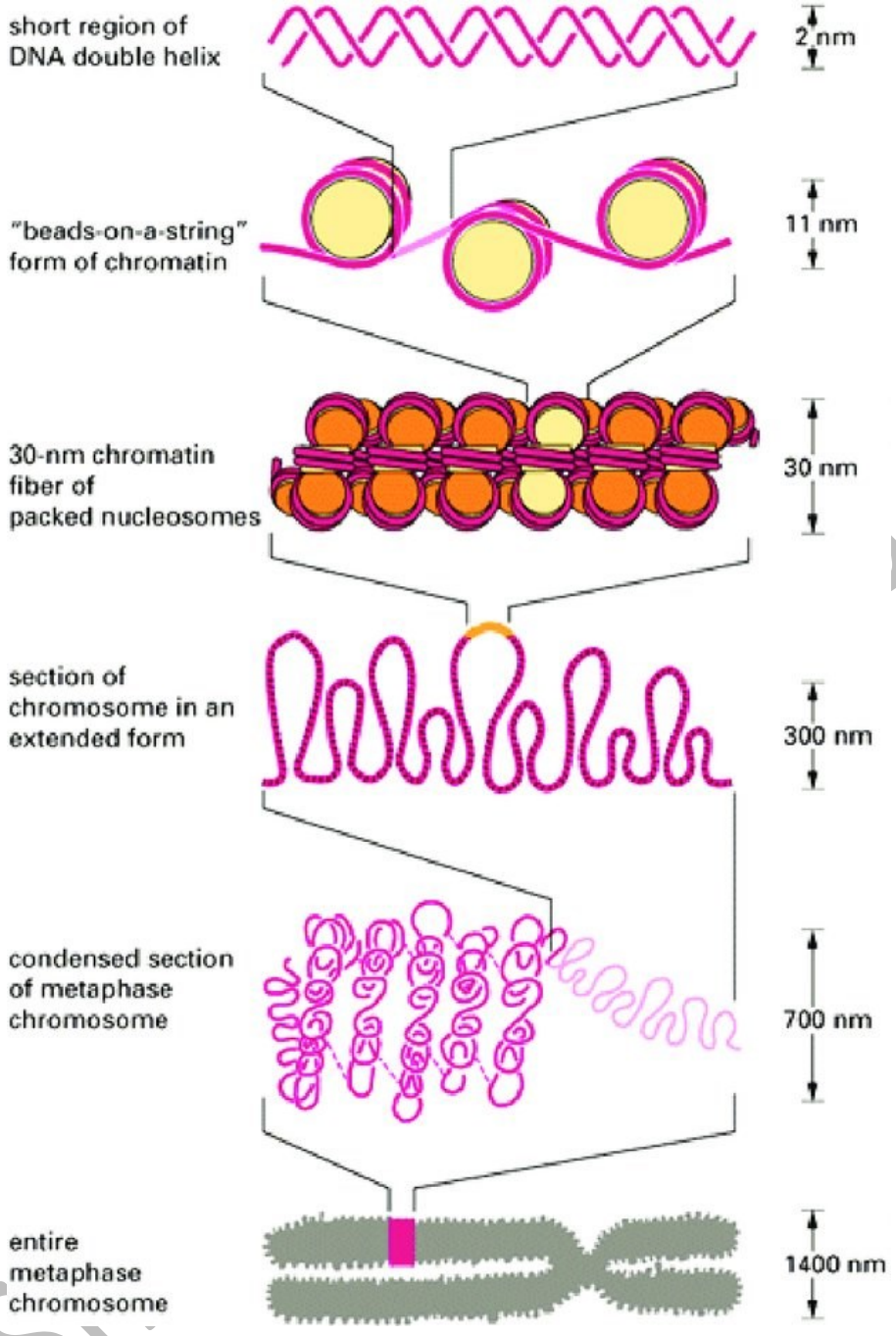


Figure 6: Illustration of the levels of organisation within the eukaryotic chromosome (Source: Nucleosome Positioning and Its Role in Gene Regulation in Yeast, Liu et al. 2018).

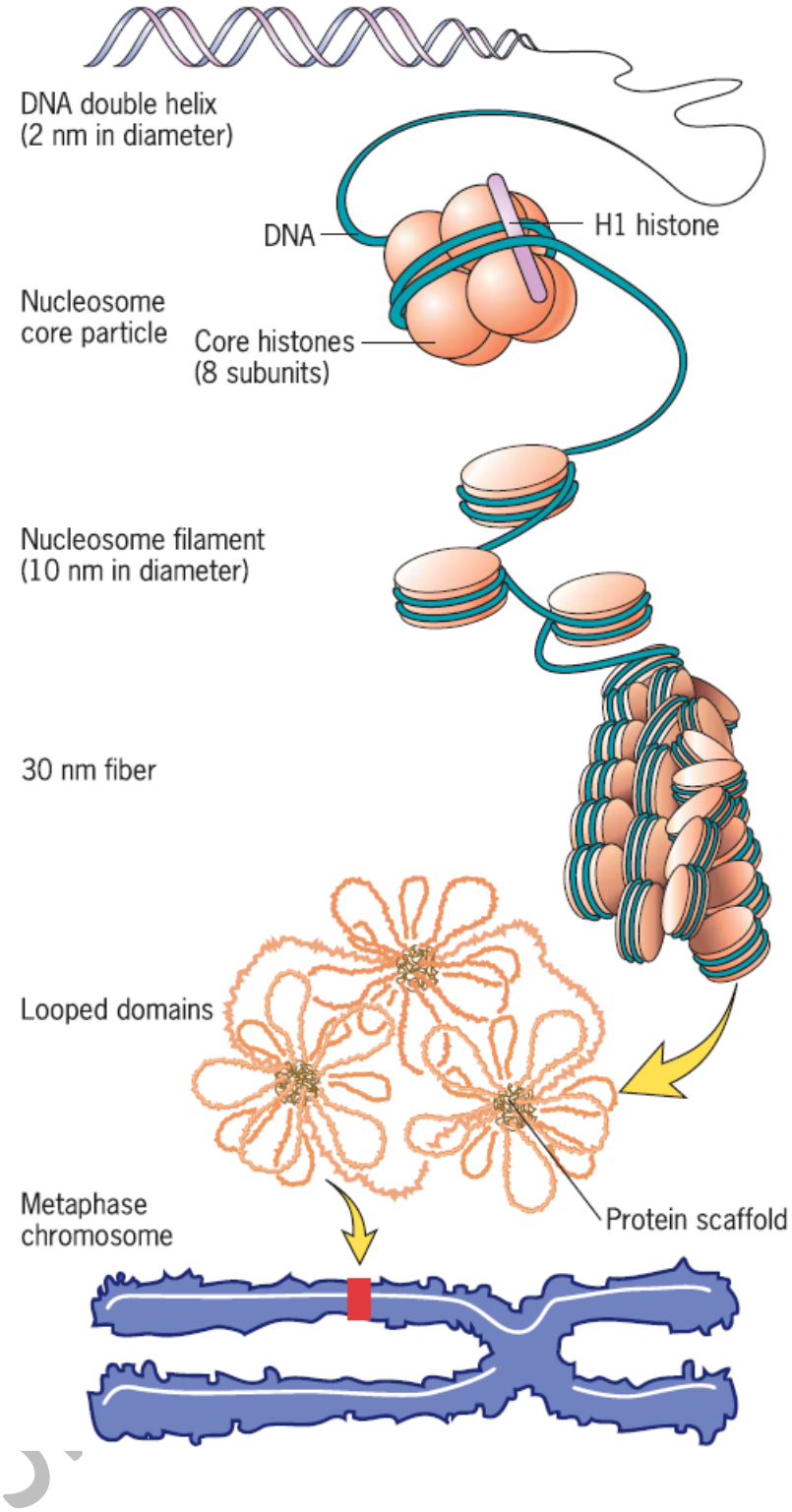


Figure 7 : Levels of organization of chromatin (Source: Cell and Molecular Biology, Gerald Karp)

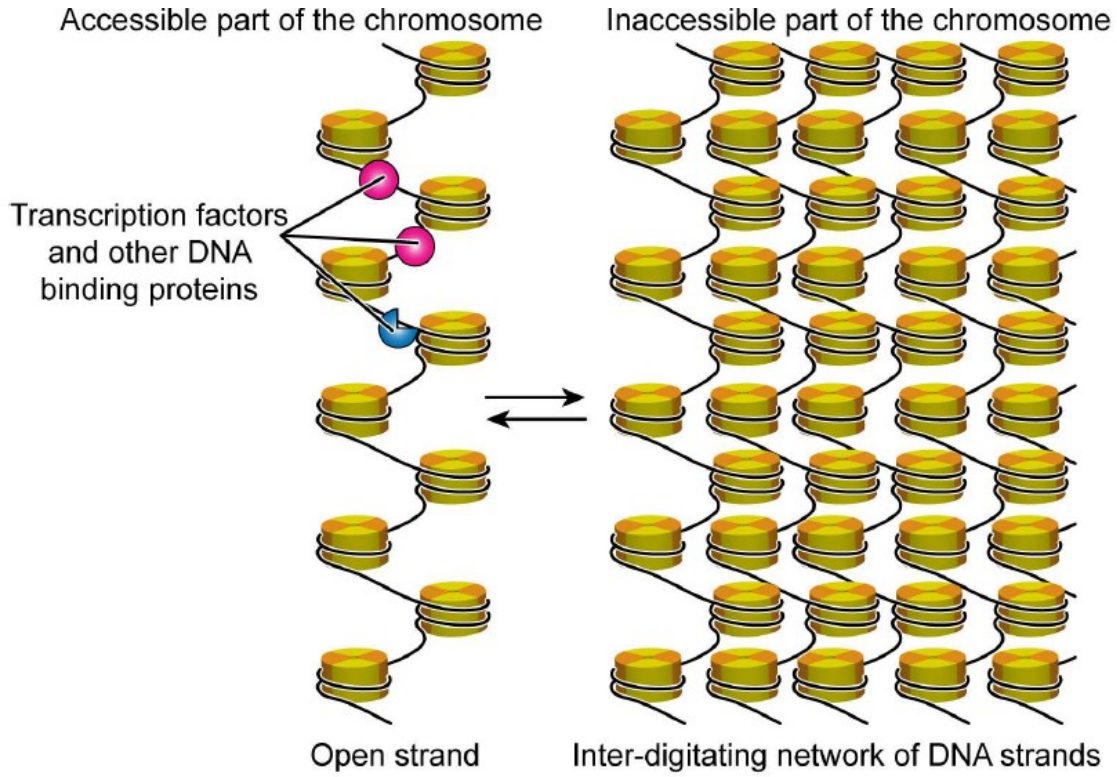


Fig. 8: The tertiary structure of the chromosome has parts which are inaccessible to transcription factors, however it also has the capability to quickly change into a more accessible form depending on the requirement. (Source: Egyankosh)

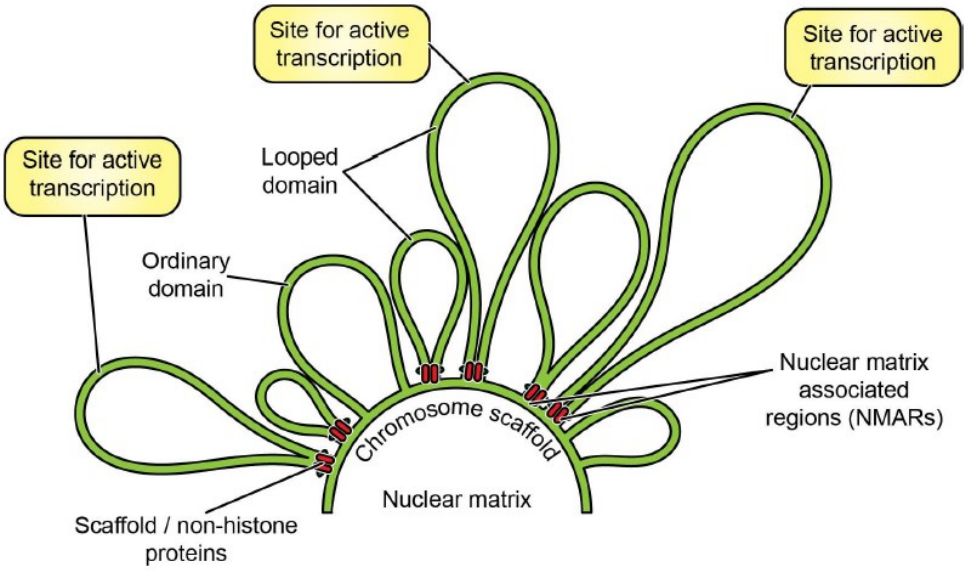


Figure 9: The Nuclear matrix associated regions (NAMRs) of the chromatin fibre tether to the nuclear matrix with the help of non-histone proteins. The loops can be loosened further and also be shifted with the help of proteins. (Source: Egyankosh)

8. Euchromatin and Heterochromatin

Euchromatin and heterochromatin are two distinct forms of chromatin, the complex of DNA and proteins that make up chromosomes within the nucleus of a eukaryotic cell. They differ in their structural organization, level of condensation, and transcriptional activity.

Euchromatin is a less condensed and more loosely packed form of chromatin. It appears as a diffuse and extended network during interphase. It is transcriptionally active, meaning that genes within euchromatic regions are more accessible and can be readily transcribed into RNA. This is because the DNA is less tightly wound around histone proteins, allowing regulatory proteins and transcription factors to access the DNA sequence. Euchromatin often contains genes that are actively involved in cellular processes. It is associated with the synthesis of RNA and the expression of functional proteins. It stains less intensely with DNA-specific dyes because of its less condensed structure.

Heterochromatin is a highly condensed and tightly packed form of chromatin. It appears as distinct, densely staining regions in the nucleus during interphase. It is generally transcriptionally inactive or less active compared to euchromatin. The tight packing of DNA makes it less accessible to the transcriptional machinery. Heterochromatin often contains repetitive DNA sequences, such as satellite DNA and transposons. It may also include genes that are temporarily silenced during certain stages of development or under specific cellular conditions. It stains more intensely with DNA-specific dyes due to its higher degree of condensation. It's important to note that euchromatin and heterochromatin are not fixed states but rather dynamic and can undergo changes in response to various cellular processes and signals. The transitions between euchromatin and heterochromatin are influenced by processes such as DNA methylation, histone modifications, and the binding of specific proteins. During cell division, euchromatin condenses further into heterochromatin to facilitate the segregation of chromosomes. In interphase, euchromatin becomes more accessible for gene expression. The regulation of these chromatin states is crucial for cellular functions, including gene expression, DNA replication, and repair.

Heterochromatin is divided into two types, which is called as constitutive and facultative. **Constitutive heterochromatin** is always inactive and condensed, and present in centromere. For example, repetitive DNA, centromeric DNA. It can affect the regions near itself. **Facultative heterochromatin** can exist in temporarily inactivated functional form. For example, female X chromosome in mammals. It is formed due to result of the genes that are silenced through RNA interference.