Overview

DNA, abbreviated from deoxyribonucleic acids, is an organic chemical molecule with a complex molecular structure that is found in all unicellular and multicellular organisms including prokaryotes and eukaryotes, and in some viruses. It was not until 1869 when the chemical DNA was discovered but its role genetic inheritance came into picture only in 1943. After a decade in 1953, Watson and Crick deciphered the structure of DNA as a double-helix comprising of two spiral polymer strands wound each other and oriented in opposite directions (antiparallel). Each polymer strand is composed of monomer nucleotides that contain either of the four nitrogenous bases; purines (adenine, A and guanine, G) or pyrimidines (thymine, T and cytosine, C). The sequence of nucleotides encodes the genetic information that is the blueprint of life processes. Given the size of the mammalian genomes, it remains a mystery how the genetic material is structurally and functionally organized in a nucleus through processes that are evidently extremely efficient. The haploid human genome is approximately three billion base pairs long (around two meters long). Barring the mitochondrial DNA, the entire genetic material of a cell is packed in a nucleus only a few micrometres in diameter? This packaging of DNA in the nuclei is highly non-random and regulated in a cell-type and developmental stage-specific manner. In most eukaryotes, the condensed and packaged DNA molecules are visible as chromosomes.

The DNA-protein interactions play essential role in packaging and stabilizing the genome, preventing formation of unmanageable tangles [1], making the regulatory regions available for long-range interactions with target genes [2] and controlled exposure of functional domains to gene expression machineries. Over and above all these transmittable genetic information, there are some inheritable functions of genome that is not encoded by DNA sequences [3]; called epigenetic changes, featuring an “additional” genetic basis of inheritance. Functional aspects of genes (being active or suppressed) are affected by these epigenetic modifications without

any alteration in the DNA sequence. The epigenetic marks are established during developmental processes, and are affected by several factors including the environment, aging, diet and disease state. Normally epigenetic modifications are manifested as lineage-specific differentiation of progenitor cells to skin cells, hepatocytes, neurons etc. An atypical mode of epigenetic change could lead to diseases like cancer. Environmental cues or surroundings of a cell trigger epigenetic changes modulated by DNA methylation, histone post-translational modifications, histone variants, nucleosome positioning or by a noncoding RNA (ncRNA) [4]. Of these, post-translational modification of histones and their variations are the focus of this thesis.

Histones

Histones, found only in eukaryotes, are basic proteins with an overall positive charge because of the amino acids lysine and arginine. Thus, salt bridges and hydrogen bonds form with phosphate oxygen on DNA. They form tertiary structures also which provide energy in the form of electrostatic interactions that fold the DNA. There are five major families of histones: H1/H5, H2A, H2B, H3, and H4 [5,6]. Of these, two H2A-H2B dimers and a H3-H4 tetramer [7]; the core histones, form a bead-like structure around which 147 bp piece of DNA is wrapped [8,9] about 1.65 times. This unit comprising of histone octamer and 147 bp of DNA is known as a nucleosome. The histone H1, linker histone, attaches itself to entry and exit sites of DNA in each nucleosome, further packing the nucleosomes forming the 10 nm fibre or beads on a string architecture [10]. In this packaged form, the DNA in nucleosomes is protected from enzymatic cleavages, however, the linker DNA of around 50 bp that connect adjacent nucleosomes is not. Such a DNA-nucleosome complex is called chromatin. The chromatin can further take up higher-order conformation by compacting nucleosomes to form 30 nm fibre and so on to an entire mitotic chromosome of 1400 nm structure [11]. In addition to this, physical interactions of histones with DNA for packaging into nucleus, chemical modifications (post-