CHAPTER 4

Summary

Spermatogenesis is a complex process that occurs in successive mitotic, meiotic, and postmeiotic phases. Spermatogenesis involves remarkable biochemical and morphological changes that transform a relatively round diploid cell to a sickle shaped haploid spermatozoa. Spermatogenesis also witnesses extensive nucleoprotein changes: from somatic histones being prevalent in the diploid spermatogonia to the protamines condensing the nuclei in spermatozoa.

The DNA in a eukaryotic cell is packaged in the form of chromatin in the nucleus. The fundamental unit of chromatin is the nucleosome, which has 146 bp of DNA wrapped around an octamer of core histones. Further level of compaction is achieved with the help of linker histones (van Holde, 1988). During spermatogenesis, testis specific variants of both core histones and linker histones replace somatic histones, starting from the mid-pachytene stage. Linker histone H1t is expressed during the mid- and late- pachytene stages and constitutes about 50-55% of the total histone complement (Rao and Rao, 1987).

Globular domains of linker histones bind to nucleosomes through two DNA bindings sites: the primary site comprising of Lys 69, 73 and 85 and the secondary site comprising of Lys 40, 42 and 52. The sequences of histones H1t and H1d appeared to be largely conserved across the globular domain except for one substitution of a lysine in H1d to a glutamine in H1t. This substitution corresponded to the Lys 40 that is implicated in nucleosome binding. We also found that all reported histone H1ts harbored a glutamine at this position.

The present investigation is aimed at (a) understanding in detail the structural properties of linker histone H1t globular domain, particularly the change at the secondary site (b) examining the underlying differences in the DNA binding sites of histone H1 globular domain across species and (c) gene expression profiling of mouse spermatogenic cells in order to identify genes involved during various stages of spermatogenesis with a particular emphasis on identifying proteins involved in chromatin dynamics.

Towards the first objective, atomic models of the globular domains of both histones H1d and H1t were built by comparative modeling. The modeled globular domains were superimposed on the previously modeled chromatosome particle (Bharath *et al.*, 2003) and its interactions with the nucleosome were evaluated. While the gross features of the two models were similar to that of histone H5, it was found that the glutamine in H1t pointed away from the nucleosomal DNA at

the dyad axis and formed a H-bond with a methionine residue that is situated two residues before the glutamine.

In order to assess whether this change in the binding site would influence histone H1 function, binding studies of the histones H1d and H1t towards both four-way junction and reconstituted mononucleosome were undertaken. It was found that histone H1t indeed had a two-fold lower binding affninity towards both the substrates when compared with H1d. In order to conclusively prove that it is indeed the single amino acid change in the globular domain that is determining the observed differential binding affinity, mutants of both histones H1d and H1t were created, H1dK52Q and H1tQ54K respectively. On performing similar binding studies with the mutants, it was found that the H1d mutant, H1dK52Q, exhibited a binding affinity that was very similar to that of histone H1t and conversely, the H1t mutant, H1tQ54K, had binding properties very similar to that of histone H1d.

Differential binding affinities could also arise because of incorrect positioning of the histone globular domains on the nucleosome. Therefore, a chromatosome-stop assay was performed. Indeed both the proteins gave the 168 bp chromatosomal band indicating that they were positioned on the nucleosomes correctly. However, it was also observed that histone H1t gave a shorter pause at the chromatosome level, which could be explained based on its lower affinity to nucleosomes.

Having observed a two-fold lower *in vitro* binding affinity for histone H1t, the next obvious question that arises is to establish whether the differential binding of the histone proteins would bear an effect on any of its *in vivo* functions. One of the hallmark events of the pachytene stage in which histone H1t is expressed is recombination. Towards this, an *in vivo* MMS sensitivity assay (Downs *et al.*, 2003) in a heterologous *S. cerevisiae* model system was performed. Histone H1t, appeared to facilitate recombination more than histone H1d according better survival properties in the presence of MMS. Also, the mutant histone H1tQ54K reduced the recombination efficiency when compared to its wild-type. While this gave evidence of the contribution of the globular domain towards histone function, it also indicated that the C-terminus of histones also contribute towards histone function. Nevertheless, the weaker binding of H1t globular domain appears to influence its function and its role in H1 function cannot be ignored.

It would be illuminating to know which other histone H1 proteins are likely to show differential binding properties and how their globular domains would differ from that of histones H1d and H1t. Therefore, globular domains of all reported linker histones across species were compared by homology modeling. A non-redundant dataset containing 176 unique histone sequences was created and analyzed. A phylogenetic analysis showed the histone proteins to

cluster into seven major groups. Representatives from each of the sub-clusters were chosen and atomic models of all the representative histone proteins were built by homology modeling and superposed on the nucleosome to evaluate their interactions with DNA.

The analysis revealed all the histone globular domains to assume a structure that was very similar to the template chicken H5 protein with no major structural deviations. Histone proteins that had substitutions in both the primary and secondary binding sites were observed. However, it appeared as though all substitutions at the primary binding site, which were likely to weaken its DNA binding affinity, were relieved by compensatory mutations in the vicinity so as to maintain the primary site binding thereby indicating that the primary site binding is absolutely essential for globular domain anchoring on the protein. While the primary site binding was almost always conserved, histones with weaker secondary site, as observed with histone H1t, were observed. Thus, it can be hypothesized that the histone protein is tailored such that altered binding is achieved not by altering the primary site residues, but by substitution of the secondary site residues. Histone H1t, also appears to have achieved a reduced binding by having substitution at the secondary site.

Yeast histone Hho1, is unique in having two globular domains and no C-terminus equivalent. A qualitative model of the yeast chromatosome particle was also built. It was obvious both from the modeling studies and also from recent reports in literature (Ali *et al.*, 2004; Ali and Thomas, 2004) that it was the first globular domain (YGD1) that bound to nucleosome like other histone globular domains. The second globular domain (YGD2) was positioned on the nucleosome in a systematic fashion such that there was maximal interactions with the exiting DNA helix and the nucleosomal dyad axis and no steric clashes or short contacts with both YGD2 and DNA.

The nucleoprotein change during spermatogenesis involves the synergistic action of many other chromatin-associated proteins. Given the complex cellular interactions that are involved in spermatogenesis, the whole process requires not only knowledge of the function of individual genes, but also information of the overall gene expression patterns associated with germ cell differentiation. Thus, gene expression profiling of mouse spermatogenic cells using cDNA microarray was undertaken to identify the various genes involved in the process.

The cell types used for comparison were the diploid spermatogonial germ cells which reflect the pre-meiotic stage of spermatogenesis, meiotic tetraploid cells and the post-meiotic haploid cells. Highly enriched populations of tetraploid and haploid cells from mice testis were obtained by centrifugal elutriation. Ten day old mice testis was used as a source of diploid germ cells. In addition, the diploid germ cells were compared with hepatic cells, a diploid somatic cell.

RNA isolated from the different spermatogenic cells was hybridized against the commercially available mouse 15K cDNA array.

The reliability of the data was determined by comparing the expression profiles of the liver and the spermatogonial cells. These analyses confirmed a high concordance between the observed expression profiles and the genes already known to be either specific to, or highly expressed in liver, thus providing a high degree of confidence in the dataset. Many genes that were previously reported to be either required for spermatogenesis or found in the spermatocytes and spermatids, have been detected, in addition to a whole set of uncharacterized genes. There recently appeared a report on gene expression profiling in spermatogenic cells (Schultz *et al.*, 2003). The present study differs from the report in that a) the array technology is different: they have used oligonucleotide arrays, the gene set of which has very minimal overlap with the gene set spotted on the cDNA microarray that was used in this study, and b) they have used whole testis in a time series experiment and correlated gene expression patterns to the mitotic, meiotic and post-meiotic phases whereas in the present study elutriated highly enriched populations of tetraploid and haploid cells that directly reflect the meiotic and post-meiotic stages, were used for comparison.

Structural proteins involved in recombination and/or loosening the chromatin structure like the synaptonemal complex protein 3 and histone H2AZ, along with other assembly/remodeling proteins like CAF1A, BAF53A and Topoisomerase IIa were found to be up-regulated in the tetraploid cells concomitant with a decrease in proteins like nucleolin, which is relevant for nucleolar function. With the disassembly of the nucleolar structure, synthesis of a whole host of ribosomal proteins was also found to be down-regulated. Also, there appeared an increase in the transcription of genes involved in ubiquitin mediated protein degradation. Signaling proteins that are likely to be involved in meiosis were found to be highly expressed in the tetraploid cells. In addition, an elevated level of transcription was observed for several enzymes that are involved in glucose metabolism and oxidative stress in both tetraploid and haploid cells. Several cytoskeletal proteins that are likely to play a role in subsequent morphological changes during spermiogenesis were also found to be expressed in the tetraploid and/or the haploid cells. Thus, the analysis encompasses all main stages of spermatogenesis and the data obtained provides a fairly comprehensive insight into this complex process.