4. DISCUSSION

4.1 HSS cDNA is a novel human cDNA with several characteristic repeats

4.2 Amino acid sequence of HSS has no homology to any known protein sequence in databanks

4.3 Expression of HSS transcript is restricted to human testis

4.4 Cell specific expression of HSS transcript is within the round spermatids of human spermatogenesis

4.5 Conclusion
4. DISCUSSION

The process of spermatogenesis represents a unique cellular differentiation pathway. During spermatogenesis, the germ cells undergo a complex process of differentiation and morphological restructuring, resulting in the formation of mature spermatozoa. Reports already exist indicative of the coordinate expression of genes in a stage specific manner accompanying the changes associated with spermatogenesis, however such reports are few and need further investigations.

It has been shown that the total synthesis of RNA in germ cells changes as spermatogenesis goes on [Hecht, 1989]. Studies on the promoter region of the protamine 1 [Peschon et al., 1989], the protamine 2 [Bunick et al., 1990a], the phosphoglycerate kinase-2 [Robinson et al., 1989] and the testis specific H2B histone [Hwang et al., 1990] genes have revealed DNA sequence motifs controlling their transcriptional activity during spermatogenesis. It has also been shown that translational regulatory mechanisms are important to ensure the correct temporal appearance of proteins within the haploid germ cell types [Braun et al., 1989; Kwon and Hecht, 1993]. Therefore, though limited, the available literature confirms that a variety of well timed gene regulatory mechanisms must be operating to secure the accuracy of spermatogenesis, and isolation of battery of such genes would give insights into the nature of molecular signals that govern spermatogenesis.

4.1 HSS cDNA is a novel human cDNA with several characteristic repeats

In order to isolate testis specific cDNA clones, one of the most widely used approach is the use of specific polyclonal antibodies against sperm antigen(s) for screening testis cDNA expression library [O'Rand and Romrell, 1981; Wright et al., 1990; Liu et al., 1996]. Taking advantage of this attribute,
we have used this approach to identify a cDNA clone (designated as HSS) from a human testis cDNA expression library.

Homology studies conducted on HSS cDNA revealed no significant homology to any known sequence in the nucleotide (GenBank, EMBL-NDB) databases and various other non redundant nucleotide databases of the BLAST network. This makes HSS cDNA a novel human cDNA sequence with respect to its specificity of expression restricted only to the testis. We have added a new gene involved specifically in the spermatogenetic process. For the uniqueness of this novel gene, HSS cDNA has been awarded an Accession Number X91879 from EMBL, Heidelberg, Germany.

Other than the regular controlling motifs of eukaryotic full length cDNA viz. start codon, stop codon, mRNA degradation consensus sequence, polyadenylation consensus signal and eukaryotic ribosome binding site, several characteristic repeats were identified within the HSS cDNA. There were five direct repeats. Two direct repeats DR IV and DR V were present in 3' UTR. DR III was present in the ORF and was in frame and hence repetitive even at the amino acid level. DR I and DR II were present in the ORF but were not in frame. It was also observed that HSS cDNA nucleotide sequence had three mirror repeats MR I and MR II were within the ORF, whereas MR III at position 2229 was in the 3' untranslated region (UTR). Despite their presence within the HSS sequence, no direct function can be attributed at this point to these repeats with respect to the spermatogenetic process.

Direct repeats have been identified in cDNA of testis specific genes of SP-10 [Wright et al., 1990] and proacrosin gene [Baba et al., 1989]. In case of SP-10 all the three direct repeats were within the ORF whereas in proacrosin a unique repeat of CCCCCA was present at the 3' end of the cDNA within the ORF. In both cases no specific function has been attributed to these repeats. Hence in case of repeats within HSS cDNA we can only hypothesize their roles in translational control by aiding in secondary structure formation of the
HSS mRNA based on evidences of translational delay of some spermatogenesis specific genes [Kleene et al., 1984; Heidaran et al., 1988; Yelick et al., 1989;]. Recently, Lee et al. [1995] have demonstrated the need for translational delay during spermatogenesis by creating transgenic mice wherein premature translation of protamine 1 mRNA causes precocious condensation of the sperm nucleus and arrests spermatid differentiation.

Analysis of the nucleotide sequence revealed the presence of three palindromic repeats at key locations. Palindromic sequences or inverted repeats, have in the past been identified in genomic sequences near the putative controlling regions of genes, or at the origins of DNA replication which could provide a molecular switch for controlling transcription or replication [Sinden, 1994]. There are reports of gene products (mRNA) within testis specific genes, having conserved sequences at the 5' or 3' UTR which act as transcriptional or translational controls.

It has been reported earlier that post meiotic gene expression is a characteristic feature of many spermatogenic specific genes [Kleene, 1996]. Most of these genes are transcribed immediately after meiosis (round spermatid) and their translation is masked until a later stage wherein protein expression is observed only at the elongated spermatid stage which in itself is transcriptionally inactive [Davies and Willison, 1993]. Since transcription ceases during mid spermiogenesis in mammals, many of the spermatid and spermatozoon specific proteins are encoded by mRNAs that are stored as ribonucleoproteins (mRNPs). There is growing evidence that the 5' conserved sequences of mRNAs play important roles in modulating mRNA translation.

Presence of the unusual palindromic motifs at key locations within HSS cDNA makes this feature new and unique to a new family of testis specific genes. The presence of palindromic sequence (P I) in the 5' UTR of HSS cDNA could suggest that this might be involved in translational control. In case of testis specific isoenzyme superoxide dismutase (TscD) it has been shown that a 65
kDa cytoplasmic protein binds specifically to the 5' UTR (absent in somatic isoforms) repressing the translation. This aids in fine tuning of this enzyme within the maturing germ cells [Gu and Hecht, 1996].

Analysis of HSS cDNA nucleotide sequence revealed the presence of a palindromic sequence (P II) towards the 3' end of the cDNA, at position 2115 which is 26 bases upstream of the putative stop codon TAG and is located within the ORF. Palindromic sequence P III is present at position 2475, present 37 bases upstream of the poly A⁺ tail and is located within the 3' untranslated region.

There are enough evidences that certain conserved sequences at 3' untranslated regions (UTRs) of mRNAs are involved in modulating mRNA translation. [Hentze et al., 1987; Jackson and Standart, 1990; Kwon and Hecht, 1991]. Protamine 1 and protamine 2 genes of mouse are the two best studied examples of spermatogenetic specific genes for translational regulation by protein-UTR interactions. Translational delay has been observed in both protamine 1 and protamine 2 transcripts. In case of the protamine 2 gene of mouse it has been observed that the translational product is delayed for 3 to 7 days [Kleene et al., 1984]. Studies on protamine 1 using transgenic mice have shown that replacing its 3' UTR with 3' UTR of human growth hormone did not result in any delay in the translational product and translation occurred in both round and elongated spermatid [Braun et al., 1989]. Recently a specific binding of a 47 kDa protein to a conserved 8 nucleotide motif (GCCATCAC) in the 3' UTR of rat transition protein-2 mRNA has been shown to be involved in translational control mechanism [Schlicker et al., 1997].

All these studies confirm the specific function of the sequences present in the 5' or 3' UTR of testis specific genes as translational control elements. Preliminary gel shift assays have been conducted in our laboratory using oligonucleotides containing these palindromic repeats within HSS cDNA. Using proteins from human testis cytosol and nucleus extracts, a strong
binding of specific proteins to these palindromic repeats has been demonstrated (data not shown). As we did not find any homology within HSS cDNA at the nucleotide level in the various databanks for such unusual motifs, we suggest that the unique palindromic repeats within HSS cDNA could be novel controlling elements in a new family of testis specific genes with specific functions during spermatogenesis.

**4.2 Amino acid sequence of HSS has no homology to any known protein sequence in the databanks**

The amino acid sequence of HSS did not show any significant homology to other known proteins in the databases of EMBL-PDB, Swiss Prot and other non redundant protein databases of the BLAST network. Analysis of HSS amino acid sequence revealed several characteristic features viz., leucine zipper motif (LZ) with six leucine repeats, four putative N-linked glycosylation sites, twenty one putative kinase sites and six putative myristoylation sites. LZ motifs, identified as a hypothetical structure common to a new class of DNA binding proteins [Landschultz et al., 1988], have been identified in several nuclear transcription factors viz. GCN4 in yeast [Struhl, 1987; Ellenberger et al., 1992], C/EBP (rat liver) [Landschultz et al., 1989] and in proto-oncogenes like jun and fos [Vogt et al., 1987; Turner and Tjian, 1989]. The classical LZ motif present in DNA binding proteins have leucine repeats with an upstream basic domain called the cluster-spacer-cluster. These domains have putative role in recognizing specific sequences on the DNA, while the LZ motif aids in dimerization, a prerequisite for optimal targeting to the DNA [Busch and Sassone-corsi, 1990]. The LZ motif of HSS protein does not have any upstream basic domain which differed from the classical LZ motif in DNA binding proteins. Therefore the presence of LZ motif in HSS protein can be attributed a different function other than that of DNA binding.

Based on the hydropathic index and prediction of membrane spanning segments by the method of Klein et al., [1985], HSS protein was classified as
peripheral and not integral. Analysis of the HSS amino acid sequence for eukaryotic secretory signals or transmembrane helices did not reveal any sequence/signals that could help the transport of HSS protein to the sperm surface. Absence of such features within HSS amino acid/protein lead us to search the literature for novel mechanisms that might exist for proteins lacking signal peptides or transmembrane domains to be transported to the cell surface. Literature survey showed several proteins with LZ motifs that were not DNA binding proteins but membrane proteins viz. in the voltage gated K⁺ channels [McCormack et al., 1989], glucose transporters of vertebrate cells [White and Weber, 1989] and the fusion (F) glycoproteins of several paramyxoviruses [Buckland and Wild, 1989]. The function of LZ motifs in all the three cases was to aid in dimerization of individual monomers to form a functional dimer. In case of the F glycoproteins of paramyxoviruses, studies on the influenza haemagglutinin and the glycoprotein of vesicular stomatitis virus, it has been shown that after the synthesis of these glycoproteins in the endoplasmic reticulum, they must oligomerise in order to be transported to the cell surface via the golgi complex [Buckland and Wild, 1989]. These investigators have also shown that a failure in dimerization results in loss of transportation of these glycoproteins to the surface.

Polyclonal antisera used to identify reactive proteins expressed from HSS cDNA in an expression library, localized antigenic determinants on the sperm acrosomal compartment by immunofluorescence experiments conducted on human sperm and immunohistochemical studies on human testis sections (data not shown). Also flow cytometric analysis of these antibodies distinctly revealed that they bind to the surface of live human sperm (binding was lost on disrupting the acrosomal membrane).

Based on the results of HSS antibody localization on human sperm surface and the studies on F glycoproteins we can suggest a similar mechanism of LZ dimerization and transport of the HSS protein to the surface
of the sperm acrosomal compartment (a derivative of the golgi complex). The proposed mechanism for the leucine zipper motif in the HSS protein has been supported by recent studies done on the major histocompatibility complex (MHC) class II molecules which are membrane-anchored heterodimers (DR\(\alpha\) and DR\(\beta\)) that present peptides on the surface of antigen presenting cells to T cells. The hydrophobic transmembrane regions of DR\(\alpha\) and DR\(\beta\) are known to facilitate assembly of the heterodimers, and truncation of one or both transmembrane regions results in loss of heterodimer formation. Replacement of this hydrophobic transmembrane region with a LZ motif (from fos and jun transcription factors) restored the heterodimer formation and secretion of HLA DR2 onto the membrane surface [Kalandadze et al., 1996].

4.3 Expression of HSS transcript is restricted to human testis

Spermatogenesis involves a series of developmental changes of the male germ cell. Hence it is not too surprising that the action of a large number of germ cell isoforms of housekeeping genes are expressed during spermatogenesis. These isoforms are expressed only in the testis but due to their partial similarities, Northern blot analysis using testis specific full length cDNAs identifies signals in all the tissues expressing the somatic isoforms. As in case of hexokinase, being an important enzyme of the glycolytic pathway its expression is seen in almost all the major tissues. On the other hand the different testis specific isoforms of human hexokinase cDNA are specifically expressed only in the testis. Analysis of such cDNA reveal a unique testis specific sequence at the 5' end of all the three isoforms, and they lack the porin binding domain (PBD) seen in all the somatic isoforms [Welch et al., 1993; Mori et al., 1996]. Hence at the nucleotide level very little homology is seen except in the regions coding for the active site of the enzyme. Testis specific genes and testis specific isoforms of somatic genes have also shown the presence of multiple transcripts within testis with no expression in other
tissues. Splicing is one of the most common mechanism for such a phenomenon. Northern blot of different human tissues revealed that only one transcript of SP-10 (a human sperm specific gene) was present within the human testis [Herr et al., 1990] but later on it was demonstrated by RT-PCR on human polyA+ mRNA that there were eleven spliced forms of SP-10 cDNA within the human testis [Freemerman et al., 1995]. By quantitative RT-PCR experiments it was concluded that the biggest transcript identified by Northern blot was the most abundant (72%) among all the eleven spliced forms. Similar observations of spliced forms of testis specific genes have been reported in the case of c-kit receptor mRNA [Sorrentino et al., 1991; Rossi et al., 1992]. A mechanism other than splicing has been suggested for multiple transcripts in case of testis enhanced gene transcript (TEGT) wherein a mechanism of differential usage of polyadenylation sites have been described [Walter et al., 1995].

Northern blot analysis of HSS gene showed a single transcript of approximately 3 kb only in the testis and not in other human tissue analysed. Absence of any other smaller transcripts does not eliminate the presence of other forms as Northern hybridization is not as sensitive a technique as RT-PCR (used to identify other forms). Considering a single transcript present only in testis and not in any other tissue, it seems that this gene does not belong to the pool of testis specific genes that are isoforms of somatic/housekeeping genes. Hence HSS gene is suggested to be testis specific.

4.4 Cell specific expression of HSS transcript is within the round spermatids of human spermatogenesis

The remarkable arrangement of the developing germ cells (stages) in various cross sections of the seminiferous tubules described for rodents (rat and mouse) is apparently lacking in human since these associations are haphazard and difficult to identify. Yet a deeper observation of human testis
sections revealed certain constant cellular associations between the different
generations of the developing germ cells [Clermont, 1963] and six different
stages were classified (I-VI). These stages succeed each other in time in a
cyclic manner to result in "the cycle of the seminiferous epithelium in man".
Spermiogenesis on the other hand has been divided into 12 steps with three
characteristic phases viz., the golgi phase (3 steps), the cap phase (4 steps)
and acrosomic and maturation phase (5 steps). It is during these steps that
there is maximum reorganization of sperm specific macromolecules along with
control on transcription and translation of various testis/sperm specific genes.

Post-meiotic gene expression has been a characteristic feature of most
testis/spermatogenic specific genes. Hence initiation of transcription in most of
these genes begins immediately after meiosis [Erickson, 1990], though their
translation varies with respect to time. The mRNAs synthesized after meiosis
are stored as translationally inactive ribonucleoproteins (RNPs) and are
activated at the required time by different mechanisms not yet fully understood
[Kleene, 1993]. In order to determine this temporal relationship between the
synthesis of mRNA transcript and the appearance of protein, it is important to
localize the specific stages of germ cells that initialize the synthesis of these
mRNAs.

In situ hybridization is an important approach for detecting specific
mRNAs within individual germ cell types in testis sections. In particular, the
ability to localize and monitor changes in the levels of these germ cell specific
mRNAs has helped in partial understanding of temporal transcription of some
germ cell specific genes like SP-10 in humans [Kurth et al., 1993] and germ
cell specific isoymes like glyceraldehyde 3-phosphate dehydrogenase in
mice [Mori et al., 1992]. In situ hybridization studies on human testis sections
using riboprobes synthesized from the 5' and 3' end of HSS cDNA clearly
identify HSS transcripts primarily on the round spermatid of stages I, II and III
of the human seminiferous cycle and not on any other cell type. Hence HSS
cDNA is not only testis specific but also germ cell specific. The paucity of signal representing HSS mRNA in spermatogonia and spermatocyte suggest that the message is post meiotically transcribed.

During mammalian spermatogenesis the developing male germ cell undergoes massive structural changes during the haploid phase as the round spermatid transforms into the species specific spermatozoon [Hecht, 1990]. This is mainly due to the temporal expression of many male germ cell proteins such as the transition proteins (TP1), protamines (1 and 2), outer dense fibre protein (ODF) and the fibrous sheath protein [Heidaran et al., 1988; Morales et al., 1994; Carrera et al., 1994 and Hecht, 1995], occurring at various periods after the cessation of transcription in post meiotic germ cells, suggesting that their expression is translationally regulated. To accommodate this cessation of transcription, the mRNAs encoding some nuclear proteins and structural proteins are synthesized in the round spermatid stage which are not translated until the time of their requirement. In contrast to these, mRNAs of several testis specific isozymes and others such as the acrosomal protein SP-10 mRNA are expressed simultaneously with their transcription. Thus, the utilization of functional mRNA in the cytoplasm of eukaryotic cells can be regulated by controlling the stability of individual mRNAs or by altering their ability to bind ribosomes and be translated.

HSS protein expression studies done on human testes sections using immunohistochemical techniques showed that the translation of HSS is initiated at the elongated spermatid stage, during spermatogenesis [data not shown]. Combining the transcription and translation data of HSS it is clear that though HSS is transcribed during the round spermatid stage, whereas the translation of HSS is masked upto the elongated spermatid stage, due to mechanisms not yet identified.
4.5 Conclusion

HSS cDNA was isolated from human testis cDNA expression library and was characterized as 2.54 kb cDNA clone. The nucleotide sequence revealed the presence of 5 direct repeats, 3 mirror repeats and 3 inverted repeats. The significance of these repeats present in the sequence is not known at present. Hydropathy plot indicated that the translated protein is predicted to have no transmembrane domain. The amino acid sequence demonstrated the presence of LZ motifs. The function of the LZ motifs may be attributed in HSS protein dimerization and transport to the surface of the human sperm acrosomal compartment. Prediction of membrane spanning segments revealed that the protein is classified as peripheral and not as integral. Northern blot analysis of eight different human tissues has shown that HSS gene is expressed specifically in the testis and detects a single band of 3 kb. The in situ RNA hybridization studies done using riboprobes synthesized from both 5' and 3' end of the ORF revealed the expression of this gene only on round spermatid and not in any other cell type. Moreover its uniqueness is that it is a novel sequence as there is no homology either at nucleotide or amino acid level for which HSS DNA has been awarded an Accession Number X91879 by EMBL.