Journal of Cellular Biochemistry 54:219–230 (1994)

Primate Testicular Histone H1t Genes Are Highly Conserved and the Human H1t Gene Is Located on Chromosome 6

Daniel A. Koppel, Steven A. Wolfe, Leon A. Fogelfeld, Peggy S. Merchant, Leonard Prouty, and Sidney R. Grimes

Research Service (151), Veterans Administration Medical Center (D.A.K., S.A.W., L.A.F., P.S.M., S.R.G.) and Department of Biochemistry and Molecular Biology, Department of Medicine (L.A.F.) and Department of Pediatrics (L.P.) Louisiana State University Medical Center, Shreveport, Louisiana 71130-3932

Abstract The testis-specific histone H1t gene is known to be transcribed only in pachytene primary spermatocytes during spermatogenesis. Previous studies of the rat histone H1t gene revealed a unique promoter sequence element between the H1/GC box and the H1/CCAAT box. Proteins in crude nuclear extracts of rat testis bind specifically to this sequence element and a temporal correlation exists between the appearance of these DNA binding proteins and the onset of transcription. These discoveries led to a search for histone H1t genes in other mammalian species. The human and monkey histone H1t genes were amplified from genomic DNA using the polymerase chain reaction (PCR). The amplified genes were cloned and the genomic derived inserts were sequenced using linear PCR. Both proximal promoters contained the highly conserved H1/AC box, H1/CCAAT box, and H1/TATA box found in nongerminal H1 genes. Both promoters also contained the H1/GC box and the H1t/CCTAGG sequence element between the H1/GC box and H1/CCAAT box previously seen only in the H1t promoter. Specific amplification of the human H1t gene using template DNA samples from a NIGMS human/rodent somatic cell hybrid mapping panel has shown that the human histone H1t gene is located on chromosome 6. © 1994 Wiley-Liss, Inc.

Key words: spermatogenesis, sperm nuclear protein, DNA binding protein, chromatin, chromosomal protein, regulation of gene expression, tissue specific gene expression, testis specific histone

H1 is the largest and most variable of the histones. A primary function of H1 appears to be structural. By binding to linker DNA between nucleosomes, it is responsible for the higher-order packaging of chromatin in most eukaryotic cells [Cole, 1987]. Histone H1 has also been noted to be involved in the repression of transcription of some genes, possibly by the displacement of essential transcription factors [Zlatanova, 1990]. Vertebrates have seven known different subtypes of the histone H1 gene including histone H1° [Seyedin and Kistler, 1979; Lennox, 1984; Eick et al., 1984; Doenecke and Toenjes, 1986; Coles et al., 1987; Ohe et al., 1989; Albig et al., 1991]. Cross species comparisons have shown that the properties of most of the subtypes are conserved. The distribution of the subtypes on chromatin has been shown to be nonrandom, suggesting that they play a more important role than serving as passive chromatin structural elements [Breneman et al., 1993; Parseghian and Hamkalo, 1993; Schulze et al., 1993]. In higher eukaryotes, the number of H1 subtypes and the amount of each subtype varies from tissue to tissue and some subtypes are expressed in a tissue-specific manner [Lennox and Cohen, 1983; Maxson et al., 1983; Meistrich et al., 1985; Grimes, 1986]. The relative amounts of histone subtypes H1.1 and H1.2 (H1a and H1c, respectively; refer to Table II for H1 subtype nomenclature) are high in dividing cells, while subtypes H1.3 and H1° (H1e and H1°, respectively) are higher in nondividing cells [Lennox, 1984].

Subtypes H1.3, H1.4, H1.5, and H1° which constitute the largest amount of histone H1 in somatic tissues are much less abundant in germinal cells. During spermatogenesis somatic his-

Received August 5, 1993; accepted September 29, 1993.

Address reprint requests to Sidney R. Grimes, Research Service (151), Overton Brooks Veterans Administration Medical Center, 510 East Stoner Avenue, Shreveport, LA 71101-4295.

© 1994 Wiley-Liss, Inc.

tone subtypes initially make up about 50% of the H1 histone and drop off to 1% by the late pachytene stage. The somatic histones are first replaced by H1.1 and H1.2 in the spermatogonial stage where they are most abundant accounting for about 80% of H1. These two histones are then replaced by the testis-specific histone H1t in the primary spermatocyte stage where it constitutes about 60% of the H1 complement in late pachytene spermatocytes [Bucci et al., 1982; Meistrich et al., 1985]. Most of the histone H1t and core histones are then replaced with transition proteins TP1 and TP2 in the midspermatid stage of spermatogenesis [Grimes et al., 1977; Bucci, 1982; Grimes, 1986]. This dramatic nuclear protein transition occurs at a time when the spermatid nucleus begins to elongate, the chromatin undergoes an extreme degree of condensation, and the pattern of transcription changes dramatically [Meistrich et al., 1985].

Germinal histone H1 variants are found in most mammalian species and in most vertebrates. They are also present in other phyla including echinoderms [Hnilica and Johnson, 1970; Subirana, 1970] and arthropods [Kaye and McMaster-Kaye, 1974]. Translation of rat histone H1t occurs exclusively in pachytene primary spermatocytes and it is the most abundant H1 variant in these meiotic cells [Grimes, 1986]. It is presumed that the testis histone H1t may be essential for the unique chromatin structure and for the unique transcription pattern in primary spermatocytes; however these possibilities have not been adequately tested.

Histone H1t genes from both Sprague-Dawley rats and humans have been isolated and sequenced [Cole et al., 1986; Grimes et al., 1990; Drabent et al., 1991]. The two sequences are very similar, but they differ substantially from nongerminal H1 genes [Wolfe and Grimes, 1993]. Three highly conserved sequence elements, the H1/AC, H1/CCAAT, and TATA elements, are found in almost all nongerminal and testis H1 promoters (see Fig. 2) [Dalton and Wells, 1988; van Wijnen et al., 1988; Osley, 1991]. Only histone H1° lacks the H1/CCAAT element. Two of these, the H1/AC and H1/CCAAT elements, are specific to histone H1 promoters [Dalton and Wells, 1988; Gallinari et al., 1989]. We have described a unique sequence element between the H1/GC box and H1/CCAAT box seen only in the rat and human H1t promoters [Grimes et al., 1992a,b; Wolfe and Grimes, 1993].

In this study, we extended our analysis of the conserved sequence elements of histone H1t to other mammalian species. We applied a PCR procedure to amplify and sequence both human and monkey histone H1t genes. The human histone H1t gene was mapped using PCR to amplify the gene from template DNA samples isolated from a human/rodent somatic cell hybrid mapping panel.

MATERIALS AND METHODS

Reagents and Supplies

All oligonucleotide primers used were ordered from Bio-Synthesis, Inc., Lewisville, TX. Radiolabeled [γ-^{32}P] ATP (6,000 Ci/mmol) was purchased from New England Nuclear (Boston, MA). The Klenow fragment of DNA polymerase I, T4 DNA polymerase, and T4 DNA ligase were purchased from Boehringer Mannheim. T4 RNA ligase and T4 polynucleotide kinase were ordered from New England Biolabs. *AmpliTaq* DNA polymerase was ordered from Perkin Elmer Cetus. X-ray film was obtained from Eastman Kodak (X-OMAT XAR-5) and DuPont (Cronex 7).

PCR Primers

H1t specific primers were designed, with the aid of the *Oligo* software package from National Biosciences, Inc., by examining conserved sequences in both the rat and human H1t genes [Grimes et al., 1990; Drabent et al., 1991]. A total of nine primers were chosen for use in PCR amplification and sequencing. Two 20-mer primers identical to the rat sequence, one a 5′ primer consisting of the H1/AC box (−111 to −92 as numbered in Fig. 1B) and the other a 3′ primer containing the histone gene specific element that leads to a stem loop structure in mature mRNA (AAGTGGCTCTTAAAAGAGCC), were used initially to amplify the rat H1t gene from genomic DNA. For amplification of the full length monkey and human genes, the 3′ primer was replaced with a conserved 20-mer further downstream from the dyad symmetry element (744 to 763). Additionally, a 17-mer primer (CCGGTTC-CTTAAGCTT) and its complementary strand were both used for sequencing. The complement was also used for amplifying a fragment that extends from the H1/AC box to the *Hind* III site near the center of the coding region. Four additional primers were designed to complete the sequencing of both strands, (GTTATGTCTGAAA-CCGT) and (GCGGGAGAAAGGCTAAAGGA)

and their complements. These primers were based upon the monkey and the human sequences at (39 to 56) and (542 to 561), respectively.

Amplification of Histone H1t

Human placental (from one individual) and Rhesus Monkey liver genomic DNA samples (both purchased from Clontech Laboratories) were used as templates for amplification of the H1t gene. Other templates used are described in the Results section. The amplification was performed for 30 cycles, in 100 μl of solution containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, each dNTP at 200 μM, 100 pmol of each primer, 1 μg of the respective genomic DNA, and 2.5 units of *AmpliTaq* DNA polymerase. The samples were heated to 94°C for 5 min prior to the addition of *Taq* polymerase. Each cycle of the polymerase chain reaction consisted of a 1 min denaturation step at 94°C, a 1 min annealing step at 52–56°C, and a 2 min extension step at 72°C. The product was electrophoretically separated on a 1% low melting temperature (LMT) agarose gel and the expected sized band was excised. DNA was recovered from the agarose gel using the Elutip-d (Schleicher and Schuell) extraction procedure [Schmitt and Cohen, 1983] or by using a Spin-X (Costar) procedure. The DNA fragment was further amplified by running another 30 cycles using the same conditions as above except 100 ng of purified template DNA was used in the reaction. Protein was removed by extracting the final product with phenol/chloroform [1:1 (vol/vol)] and then the DNA was ethanol precipitated. At this point, the amplified product was either cloned into pUC19 (Bethesda Research Laboratories) or sequenced directly. To purify the samples for direct sequencing the amplified product was again electrophoresed on LMT agarose and purified as above.

Cloning and Screening

The exonuclease activity of T4 DNA polymerase (New England Biolabs) was used to remove any 3′ overhangs created by *Taq* polymerase [Clark, 1988]. Next, the PCR fragments were phosphorylated using T4 polynucleotide kinase and ATP. The amplified fragments were either ligated to *BamH* I linkers and then inserted into the *BamH* I site of pUC19 or blunt-end ligated into the *Sma* I site of pUC19. Competent *Escherichia coli* host strain JM83 was transformed with the ligated DNA. All white colonies potentially containing the recombinant plasmid were single colony isolated and grown in 5 mL L-broth cultures. Plasmids from these cultures were harvested using a mini-plasmid procedure [Birnboim and Doly, 1979]. Recombinant plasmids were identified by slot blotting on nitrocellulose (Schleicher and Schuell) [Meinkoth and Wahl, 1984] probing with a 30-mer [γ-^{32}P] ATP end-labeled oligonucleotide (−92 to −63) from the rat H1t proximal promoter [Grimes et al., 1990]. All positive recombinant plasmids were harvested from large scale cultures using alkaline lysis [Birnboim, 1983]. Plasmid DNA was purified on CsCl gradients containing ethidium bromide by centrifugation at 60,000 rpm for 4.5 h in a Beckman VTi65.2 vertical rotor. The supercoiled plasmid DNA was extracted with butanol to remove ethidium bromide, dialyzed against 10 mM Tris-HCl (pH 8.0), 1 mM EDTA, and quantified by reading A_{260} [Grimes et al., 1990].

Nucleotide Sequence Analysis

Cycle sequencing was conducted using a double stranded linear PCR method with the fmol™ kit developed by Promega. Approximately 10 pmol of each of the primers described above were end-labeled with [γ-^{32}P] ATP using 10 units of T4 polynucleotide kinase. Either 1 μg of plasmid DNA or 20 ng of amplified fragment were sequenced using PCR based upon the dideoxy chain termination protocol and using reagents provided in the fmol™ kit. The reaction was stopped with 4 μl of 95% formamide, 20 mM EDTA, 0.05% Bromophenol Blue, and 0.05% Xylene Cyanole solution. DNA fragments were separated by electrophoresis on a 6% polyacrylamide gel, containing 7 M urea and detected by autoradiography. The nucleotide sequences of both strands of the human and monkey histone H1t genes were determined by this method. Sequence data was analyzed with the IBI Pustell sequence analysis software package and with software available through Intelligenetics, Inc.

Chromosome Mapping

A National Institute of General Medical Sciences (NIGMS) human/rodent somatic cell hybrid mapping panel was used to map the human histone H1t gene. The panel consists of DNA samples isolated from 24 Human/rodent fusion cell lines. The rodent cell lines were derived from either mouse or Chinese hamster. Human cell line IMR91 was used as a positive control

and mouse line 3T6 and Chinese hamster line RJK88 were used as negative controls. All but two of the cell lines had retained only one human chromosome. Using human histone H1t specific primers, (−111 to −92) for the 5′ primer and (542 to 561) for the 3′ primer, 1 μg of DNA from each sample was amplified by 30 cycles of PCR. The products were then fractionated on a 1% agarose gel stained with ethidium bromide (see Fig. 3). The DNA was blotted to a nylon membrane and probed with a 57 bp histone H1t specific labeled probe (−89 to −33; see Fig. 2). To ensure that the amplified band was H1t, the remainder of the produce of interest was electrophoresed, excised, purified, and directly sequenced as explained above.

RESULTS

Cloning and Sequencing the H1t Gene

An initial experiment was conducted to determine whether the H1t gene could be successfully PCR amplified from genomic DNA. Two oligonucleotide primers used were designed based upon the known rat H1t gene sequence [Grimes et al., 1990]; one was identical to the rat H1/AC box and the other was complementary to the 3′ stem loop. PCR amplification of rat genomic DNA was conducted using these primers, the amplified DNA fragment was purified and cloned, and three different plasmid clones were sequenced. No discrepancies were found in comparing the coding region to the published rat sequence.

From a comparison of the published human and rat sequences, we designed new primers suitable for amplification of human and Rhesus monkey H1t genes from genomic DNA. Two new 3′ primers were used for amplification as discussed in the Materials and Methods section. Purified DNA fragments from four independent PCR amplifications of monkey genomic DNA were cloned into the *Sma* I site of pUC19 and the plasmid clones were screened. Two types of plasmid clones were isolated depending upon the 3′ primer used. One contained monkey DNA extending from the H1/AC box to the *Hind* III site within the coding region (Fig. 1A) and the other contained the full gene. In parallel experiments, four clones of the full length human histone H1t gene were isolated.

Plasmids containing the human and monkey histone H1t gene were sequenced using the sequencing strategy shown in Figure 1A. The cloned sequences of the human H1t gene (Fig. 1C) conflicted with the published sequence of the human H1t gene in three positions. Direct sequencing of the genomic DNA was used to verify that the discrepancies were not from PCR error. The first difference was in the leader region immediately following the TATA box. The sequence of the human gene cloned in our laboratory contained an additional G to give the sequence TATATAAG*G*CCCC, while the previously published sequence was TATATAAGCCCC [Drabent et al., 1991]. The two other mutations, transversions at site 1 and site 2 (Fig. 1C and Table I), were located in the coding region. The four clones of the human H1t (pDK 4, 7–9) gave two different sequences at the two sites and showed three apparent allelic combinations of the sequences (Table I). The fourth possible combination is found in the sequence reported by Drabent et al. [1991] also shown in Table I. At amino acid residue 13, the triplet sequences in our plasmid clones were either GTA which codes for valine or CTA which codes for leucine. At amino acid residue 178, the triplet sequences in our plasmid clones were either AAG which codes for lysine or CAG which codes for glutamine. In both cases one of our variations agreed with the previously published sequence of Drabent et al. [1991]. The variations were then confirmed by direct sequencing of the Clontech genomic DNA where both variations were found represented in the genomic DNA giving rise to all four possible human H1 sequences. In sequencing the DNA from the human/rodent hybrid cell line representing human chromosome 6 only a single sequence was derived. This sequence which is identical to the sequences in clones pDK7 and pDK8 is shown in Figure 1C with residue 13 reading GTA and residue 178 reading CAG. The remainder of the human histone H1t sequences were identical. An examination of the hydrophobicity plots of the human histone H1t amino acid sequences failed to show a significant effect caused by either of the two conservative substitutions. No differences were found in the sequences of four separate monkey H1t gene clones.

The total nucleotide sequence of the cloned monkey DNA fragment, including the PCR primers, was 874 bp (Fig. 1B) with an open reading frame encoding 207 amino acids. The 834 nucleotides of the monkey sequence between the primers are 93% homologous to the human H1t sequence (Fig. 1C). There are only 13 differences in the entire 5′ promoter region, and all of

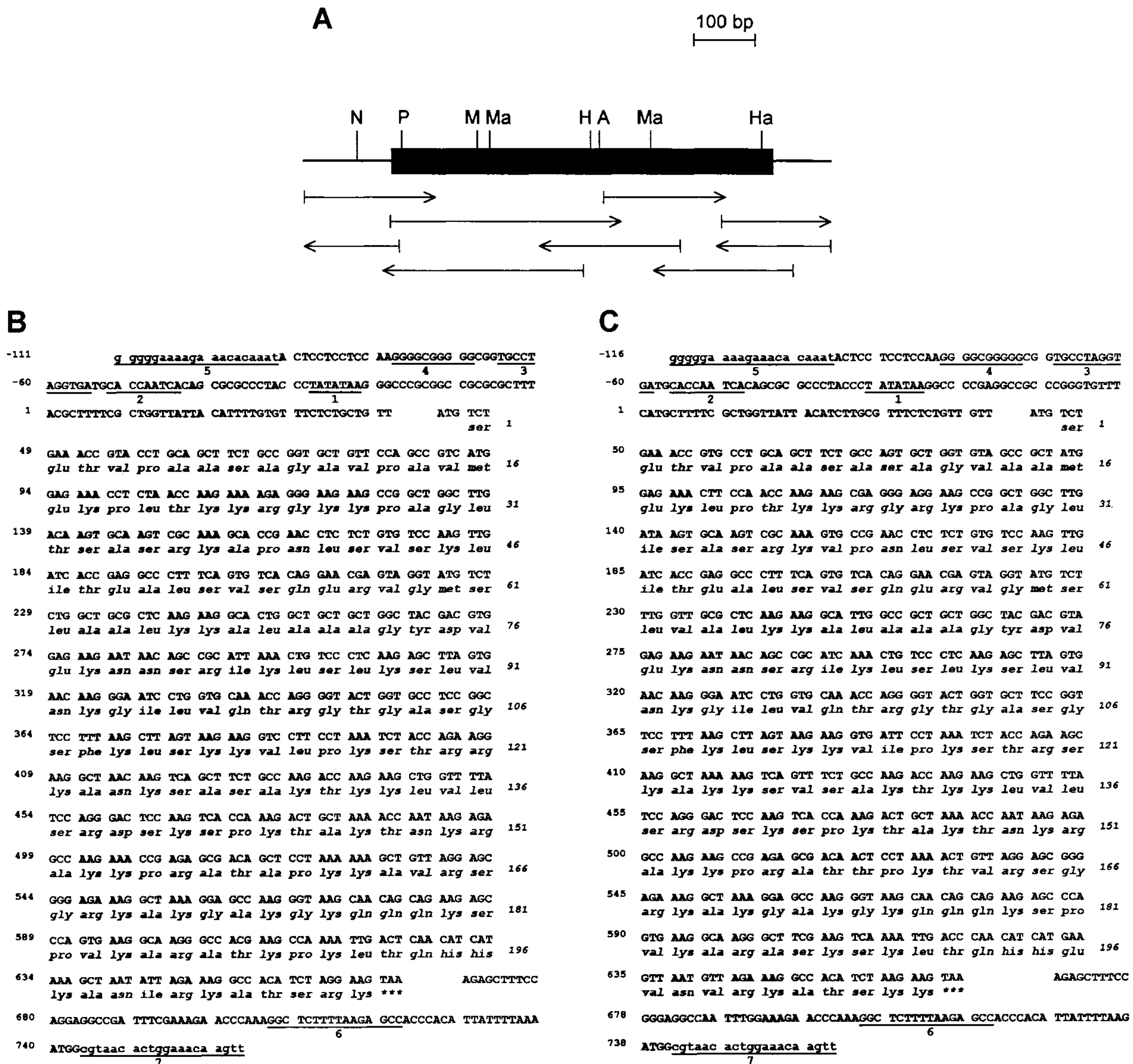

Fig. 1. Nucleotide sequencing strategy, partial restriction map, and nucleotide sequence of the Rhesus monkey and human testis-specific histone H1t genes. **A:** Nucleotide sequencing strategy and partial restriction map of the Rhesus monkey H1t gene. Arrows indicate the actual areas and direction sequenced with the respective primers. The restriction enzymes marked on the map include N, *Not* I; P, *Pst* I; M, *Mbo* I; Ma, *Mae* III; H, *Hind* III; A, *Ava* II; and Ha, *Hae* I. The coding region of the monkey H1t gene is marked as a black filled box. **B:** Nucleotide sequence of the monkey H1t gene together with the deduced amino acid sequence. The GenBank accession number for the monkey histone H1t sequence presented in this paper is M97756. The sequence shown includes the primers [lowercase underlined regions (5) and (7)], used to amplify the fragment during PCR. The oligonucleotide primer at (5) is derived from the rat H1t sequence and the primer at (7) is derived from the human H1t sequence (accession number M60094). Numbering on the left side refers to the nucleotide sequence relative to the transcription start site: +1. The italics numbering on the right refers to the amino acid residues. Underlined conserved sequence elements are 1) TATA box, 2) H1/CCAAT box, 3) H1t/CCTAGG element, 4) H1/GC box, and 6) H1 3′ dyad symmetry element. **C:** Nucleotide sequence of the human H1t gene together with the deduced amino acid sequence. The GenBank accession number for the human histone H1t sequence presented in this paper is M97755. The labels used are identical to those used for the monkey H1t sequence shown in B.

TABLE I. Sequence Differences Within the Coding Region of the Human H1t Gene*

H1t source	Area 1	Area 2
λEMBL3	GGT **CTA** GCC	CAA **CAG** CAG
pDK4	GGT **GTA** GCC	CAA **AAG** CAG
pDK7	GGT **GTA** GCC	CAA **CAG** CAG
pDK8	GGT **GTA** GCC	CAA **CAG** CAG
pDK9	GGT **CTA** GCC	CAA **AAG** CAG
Human genomic DNA	GGT $^{G}_{C}$TA GCC	CAA $^{C}_{A}$AG CAG
GM/NA10629	GGT **GTA** GCC	CAA **CAG** CAG

*A comparison of the nucleotide sequences of human H1t in the areas where differences were found. The sequence derived *from a* λEMBL3 clone was published by Drabent et al. [1991]. Plasmid clones pDK4, pDK7, pDK8, and pDK9 were derived from the cloning of the PCR product amplified from a Clontech human genomic template DNA sample. Human genomic DNA sequence was determined by direct sequencing of the Clontech DNA. The two nucleotides shown in Area 1 and Area 2 of the Human Genomic DNA template indicate that both nucleotides were represented in equal amounts at that site in direct sequencing. Template DNA from cell line GM/NA109629 gave only the single sequence shown. The sequence of H1t using this template was GGT GTA GCC in area one and CAA CAG CAG in area two. This was also the sequence of clones pDK7 and pDK8.

TABLE II. Cross Reference of H1 Subtype Nomenclature

A	B	C
H1.5	H1A	H1b
H1.4	H1B	H1d
H1.3	H1C	H1e
H1.2	H1D	H1c
H1.1	H1E	H1a
H1t	H1t	H1t
$H1^0$	$H1^0$	$H1^0$

A cross reference of three different nomenclatures used to refer to the different subtypes of H1 histones. Column A is the format used in this paper and also used by Albig et al. [1993]. Column B is a modified nomenclature used by Ohe et al. [1989]. Column C is used by Lennox [1984] and Seyedin and Kistler [1979] to describe subtypes of H1 histones found in rats and mice separated using 2D acid-urea and SDS gel electrophoresis.

these differences are located between the TATA element and the ATG start codon. The proximal promoters of the monkey, human, and rat histone H1t genes are very similar.

The human histone H1t promoter contains all of the consensus elements that are characteristic of cell-cycle-dependent H1 genes (Fig. 2 and Table II). The H1/AC box (AAACACA), which represents a portion of the 5′-primer used in PCR amplification, is located at position −106, relative to the presumed transcription start point (Fig. 1C). This element and the TATA box are found in all seven histone H1 promoters (Fig. 2). The H1/GC box (GGGGCGGGG) is located at −82, the H1/CCAAT box (CACCAATCA) is located at −56, and the TATA box (TATATAA) is located at −31. It is interesting that the H1/GC box is present only in the H1t and H1.1 histone genes, both of which are expressed in testis germinal cells (see Table II for H1 subtype nomenclature). The testis-specific H1t/CCTAGG sequence element (CCTAGG) [Grimes et al., 1992a,b] is found between the H1/GC box and the H1/CCAAT box at position −69 (Figs. 1B,C, 2). This element has previously been observed only in the promoter regions of the human and rat histone H1t genes [Grimes et al., 1990; Drabent et al., 1991]. More recently, the same sequence has been found in the mouse and rabbit histone H1t genes (manuscript in preparation).

The 3′ untranslated regions of the H1t genes are also conserved (Fig. 1B,C). In addition to the conserved histone specific dyad symmetry element (708 through 722) that leads to the formation of a stem-loop structure in mature mRNA, the downstream monkey sequence (723 through 743) is identical to the human sequence except in one position. The A at 739 in the monkey sequence is G in the human sequence. The conservation actually extends through the PCR primer region (nucleotide 745–763), since the primer hybridized successfully to the monkey genomic DNA. The 3′ PCR primer sequence was based solely upon the identity of the human and rat sequences in the region and it needs to be emphasized that we have not sequenced the actual monkey genomic sequence through the primer regions.

Human Histone H1t Gene Maps to Chromosome 6

The chromosome location of the human histone H1t gene was mapped by using a NIGMS human/rodent somatic cell hybrid mapping panel. In this panel, each sample was generated by the fusion of human cells and rodent cells (mouse or Chinese hamster). Ideally the cell lines will have lost all but one human chromosome. In this panel, all but two cell lines contain only one human chromosome. Genomic DNA from the human cell line IMR91, included with the panel, was used as a positive control, and DNA samples from mouse cell line 3T6 and Chinese hamster cell line RJK88 were used as

```
GAAACACAGATGGCGGCGGCGCAGCGCCATTCCGGGCCGGGAGCAGGCAGCCAGCAGCCCTGTCCTCACCGCGGTCCGCCCGCCGCCGCTAAATA   H1.0
 CAAACACAATTTGGGAGTCCAACGCGAGCGCGGCGGCCAGAGGGCGGTGGATTGGACGCTCCACCAATCACAGGGCAGCGCCGGCTTATATAA   H1.5
CAAACACAACTCGGGATCCGAGAGGACACTCTGCGGCTGCCAGCGAGGCGGGCTGGACAGCGCACCAATCACGGCGCACGTCCGCCCTATATAAA  H1.4
CAAACACAGCAGCGCGGTAGATACGAGGAGTCCTTTTCCAGCAGCGCCGCGCATGGAGCAAGGAACCAATCATCACTCAGCGTCTCTCTATATAAA H1.3
  CAAACACAACTTGGGAGCAGCGCAGCGGCTCAGAGCCTGCCAGCCAGGCGGGCGACCAGAGCACCAATCAGAGCGCGCCTGCGCTCTATATATA H1.2
                GAAACACATCGGCAGCGCGCAGGGCGCGGAGGGGCGGGACTGACGGGCACCAATCACGGCGCAGTCCCACCCTATAAATA H1.1
             CAAACACAAATACTCCTCCTCCAAGGGGCGGGGGCGGTGCCTAGGTGATGCACCAATCACAGCGC GCCCTACCCTATATAA   H1T
                1                        2           3          4                        5

57bp H1t Specific Probe - GCCCCTTCCCCAGGGGCGGGGGAGGCGCCTAGG GATGCACCAATCACAGCGC GCCCT
```

Fig. 2. Promoter regions of seven known human histone H1 genes. Nucleotide sequences of the seven known human histone H1 promoters are compared. The conserved sequence elements are underlined and numbered. All of the promoters except H1° show a high degree of conservation in the region extending from the H1/CCAAT box (4) through the TATA box (5). The H1° promoter is similar to the histone H5 gene found in species with nucleated erythrocytes [Doenecke and Toenjes, 1986]. Of particular interest is the unique nucleotide sequence of the human H1t gene between the H1/AC box (1) and H1/CCAAT box designated the H1t/CCTAGG element (3). The H1/GC box (2) is present only in the H1.1 (H1a, see Table II for nomenclature of H1 subtypes) and H1t genes. Below the promoter list, a 57 bp probe derived from rat histone H1t gene is shown [Cole et al., 1986; Grimes et al., 1990]. This probe was used in the Southern blot analysis for mapping of the human histone H1t gene (Fig. 3).

negative controls. The genomic DNA templates were each PCR amplified using primers and amplification conditions found to be specific for human histone H1t. The resultant DNA was then separated on a 0.7% agarose gel and visualized (Fig. 3A). Only amplified DNA samples from the human control and from the hybrid containing chromosome 6 were abundant and of the correct size. Minor nonspecific amplified bands of various sizes were seen with most of the other samples.

To confirm that the amplified band from the hybrid template containing chromosome 6 was H1t, the DNA was transferred from the agarose gel to a nylon membrane. The membrane was then probed with a histone H1t specific 57 bp probe (see Fig. 2 for the probe sequence). Autoradiography, shown in Figure 3B, revealed that only the DNA products in the human control lanes and in the lane containing hybrid sample number 6 hybridized to the probe. To further ensure that this band was indeed histone H1t, the remainder of the amplified product was purified and sequenced directly. The sequence was identical to that found for human H1t. This evidence demonstrates that human histone H1t is located on chromosome 6.

Comparison of Histone H1 Amino Acid Sequences

An alignment of the human, monkey, and rat histone H1t amino acid sequence [Cole et al., 1986; Grimes et al., 1990] is displayed in Figure 4. The figure shows that there is a high degree of conservation in amino acid sequence of histone H1t from different mammalian species. This is in contrast to the high degree of variability between the seven subtypes of H1 which have a lower degree of conservation than the core histones (Fig. 5). For example, the amino acid residues of human and monkey H1t are identical in 88% of the positions. The rat H1t sequence compared to the human and monkey sequences shows a 61% and 64% homology, respectively. In comparing the three protein domains of the testis-specific H1t gene (amino acids 1–37, 38–112, 113–207 in Fig. 4), the central domain is the most highly conserved with the monkey and the human sequences being 97% homologous. The N-terminal domains of the human and monkey are 78% homologous and the C-terminal domains are 86% homologous. Hydrophobicity plots of the human and monkey H1t histones were indistinguishable and the plot of the rat H1t histone was very similar to the human and monkey H1t plots.

In contrast, there was only a 47–49% homology of amino acid sequence found in comparing other human H1 proteins with the human H1t (Fig. 5). There is only a 46% homology between the rat histone H1t amino acid sequence and the mouse H1 (MH143–H1) sequence [Yang et al., 1987]. Furthermore, the rat testis histone H1t sequence and the nontestis rat H1d sequences are less than 50% homologous [Cole et al., 1990; Wells and Brown, 1991].

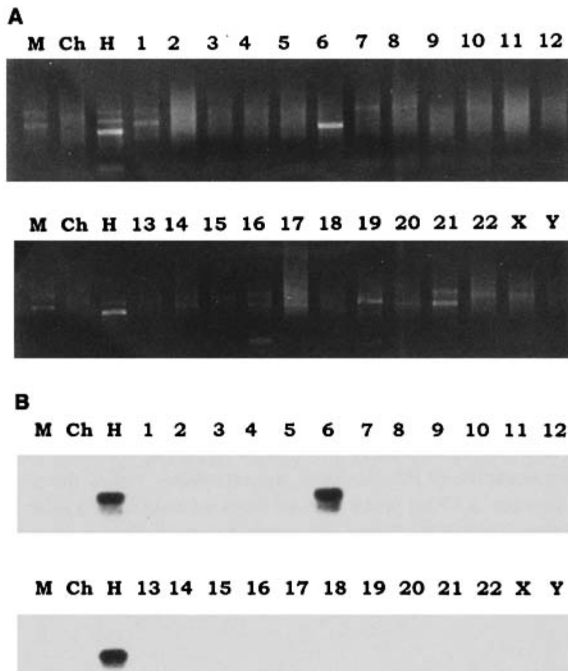

Fig. 3. Southern blot analysis of PCR products amplified from DNA templates derived from a NIGMS human/rodent somatic cell hybrid mapping panel. **A:** PCR amplified products of the NIGMS mapping panel were separated by electrophoresis on a 0.7% agarose gel. Parental cell lines from mouse (3T6), M; Chinese hamster (RJK88), Ch; and human (IMR91), H were the sources of template DNA for products run in lanes 1 through 3 in both upper and lower gels. In lanes 4 to 15 the human/rodent hybrids containing the individual human chromosomes are marked as 1 through 12 on the upper gel and 13 through Y on the lower gel. The primary band in the human lane is 672 bp. Neither the mouse nor the hamster samples have a band of this correct size. In the hybrid lanes there are some non-specific amplified bands but only chromosome 6 shows the H1t specific 672 bp band. **B:** The electrophoresed samples were transferred to a nylon membrane and probed with a 57 bp histone H1t specific probe (Fig. 2). The membrane was then exposed to X-ray film. The lane markings are the same as in the gel shown in A. Autoradiography shows that only the specific band in the human parental control lanes and the band in the hybrid cell line containing chromosome 6 bind to the histone H1t specific probe.

DISCUSSION

Currently, there are seven known subtypes of human histone H1 genes. Five of these are presumably cell cycle regulated histone H1 genes designated H1.1, H1.2, H1.3, H1.4, and H1.5 [Carozzi et al., 1984; Eick et al., 1984; Doenecke and Toenjes, 1986; Ohe et al., 1989; Albig et al., 1991; Drabent et al., 1991]. The sixth is the basally expressed H1° gene that is strikingly homologous to the histone H5 gene expressed in nucleated erythrocyte [Doenecke and Toenjes, 1986]. The seventh is the testis-specific histone H1t gene which is the focus of this paper.

In this study we used PCR to amplify and sequence primate (human and monkey) histone H1t genes. The human histone H1t gene sequence presented in this paper is identical to the previously published human histone H1t sequence [Drabent et al., 1991], except for three nucleotide differences. The first difference is an additional G in the H1t promoter sequence after the TATA box in all of our clones and in direct sequencing. The other differences are transversions at two sites within the coding region. Two possible sequence variations at the two sites make possible four sequences for the coding region. Three of these appear in our human H1t clone data shown in Table I. The fourth has been published by Drabent et al. [1991] and is also shown in Table I as the λEMBL clone.

The presence of three apparent allelic variants in our cloning from a single DNA source has at least three possible explanations. As the source was placental DNA, it is assumed that the DNA of the conceptus was contaminated with maternal decidua. The paternal contribution to the DNA source was then haploid, and the maternal diploid, as it included some of her somatic cells. Alternatively, there is an unlikely possibility that sequence variations could arise during the process of cloning. Potentially, paternal-maternal hybrid molecules of the two different DNA strands could be formed during PCR amplification. Following cloning and transfection into *E. coli,* DNA repair mechanisms may generate the transversions seen in Table I resulting in the additional sequences. This explanation seems unlikely due to the nonrandom nature of the mutations. Third, the differences might arise if H1t is a multi-copy gene, but the H1t gene has been reported to be single-copy gene [Cole et al., 1986]. It should be mentioned that the human/rodent fusion cell line GM/NA10629 used for mapping apparently has only one sequence of histone H1t gene represented on chromosome 6. Along these same lines, it is interesting that we have found only one sequence for the monkey histone H1t gene. The monkey H1t gene is remarkably similar to the human and rat histone H1t genes being more closely related to the human sequence.

The two amino acid differences in our H1t sequence compared to the published H1t sequence lie outside of the highly conserved central domain of the H1t (Figs. 1C, 4, 5). The

```
                                                                                   50
S E T V P A A S A S A G V A A M E K L P T K K R G R K P A G L I S A S R K V P N L S V S K L I T E A
: : : : : : : : : : G : V P : V : : : P L : : : : : K : : : : : T : : : : : A : : : : : : : : : : : :
: : : A : : : : S T L V P : P V : : P A : : R : : K : : G - - M A T A : : P R G F : : : : : : P : :

                                                                                   100
L S V S Q E R V G M S L V A L K K A L A A A G Y D V E K N N S R I K L S L K S L V N K G I L V Q T R
: : : : : : : : : : : : A : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : M : : : : A : : : : A : : : : : : : : : : : : : : : : : : : : : : A : : R : : : : : V : : : : K

                                                                                   150
G T G A S G S F K L S K K V I P K S T R S K A K K S V S A K T K K L V L S R D S K S P K T A K T N K
: : : : : : : : : : : : : : L : : : : : R : : N : : A : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : A A S G N D K G : G : : : A : : : A : : : G : : : A : R : : : S S : : - :

                                                                                   200
R A K K P R A T T P - K T V R S G R K A K G A K G K Q Q Q K S P V K A R A S K S K L T Q H H E V - -
: : : : : : : : A : K : A : : : : : : : : : : : : : : : : : : : : : : : : : T : P : : : : : : K A - -
V V : : : K : : P T - : G S G : R : : T : : : : : : L : : R : : : A : : : : : T N : N S G K S K M : M Q

- - N V R K A T S K K                        H1t human   (206)
- - : I : : : : : R :                        H1t monkey  (207)
K T D L : : : A G R :                        H1t rat     (207)
```

Fig. 4. A comparison of amino acid sequences of human, monkey, and rat histone H1t. The human H1t amino acid sequence is compared with the Rhesus monkey sequence, and Sprague-Dawley rat sequence respectively [Cole et al., 1986]. The top line is the human H1t sequence. The residues at the two sites where variations of amino acid sequences occurred in the human sequence are underlined (also see Table I). Arrows indicate the domain boundaries; amino acids 1–37, 38–112, and 113–207. For the other amino acids a colon signifies identity to the human sequence, whereas hyphens define gaps in each sequence necessary for sequence alignment. Numbers shown above and to the right refer to a general reference frame and the parentheses that follow each sequence indicate the protein's presumed length.

crystal structure of the central globular domain of histone H5, which is homologous to mammalian histone H1°, is strikingly similar to the DNA-binding domain of the catabolite gene activator protein CAP, providing a possible model for the binding of the globular domain to DNA [Ramakrishnan et al., 1993]. Therefore, it is relevant that the globular domain of the mammalian H1° is homologous to the globular domains of the other 6 human H1 histones including H1t (Fig. 5) implicating a possible similarity in DNA binding to this region.

Within the coding region, there is a much higher degree of conservation of encoded amino acids among the monkey, human, and rat H1t genes than there is among the seven different variants of H1 genes within any one species such as human. Furthermore, in comparing both the 5′ and 3′ noncoding regions of the H1t genes, we find that the human and monkey genes have all of the highly conserved sequence elements. These primate histone H1t genes also have a unique promoter element that has been found only in testis specific variants of the H1t gene. The data presented in this paper further support the hypothesis that the H1t/CCTAGG sequence within the histone H1t promoter is an essential sequence element that is possibly involved in regulating testis-specific transcription of the gene.

Previous studies indicated that there is high affinity testis-specific binding of nuclear proteins to the H1t/CCTAGG sequence element based upon competition assays using homologous and heterologous competitor DNA samples [Grimes et al., 1992a,b]. This binding protein potentially represents a tissue-specific transcription factor that positively regulates histone H1t expression in the testis [Grimes et al., 1992a,b]. Testis-specific factors must regulate transcription of many genes in pachytene primary spermatocytes and early spermatids. It is interesting that the appearance of this binding protein parallels the expression of the SRY gene in pachytene primary spermatocytes [Palmer et al., 1989; Su and Lau, 1993]. The SRY gene appears to be responsible for initiating testis determination and differentiation in man [Palmer et al., 1989; Su and Lau, 1993]. However, the SRY gene product is 23.9 kDa and the testis-specific H1t/CCTAGG binding factor appears to be about 70–80 kDa [Wolfe and Grimes, 1993]. The appearance of the H1t/CCTAGG binding activity also parallels the appearance of an acidic, histone binding protein designated nuclear autoantigenic sperm protein (NASP) [Lee and O'Rand, 1993] proposed to play a role in regulating some events in spermatogenesis. NASP mRNA is present only in primary spermatocytes reflecting the measured steady state level of H1t mRNA

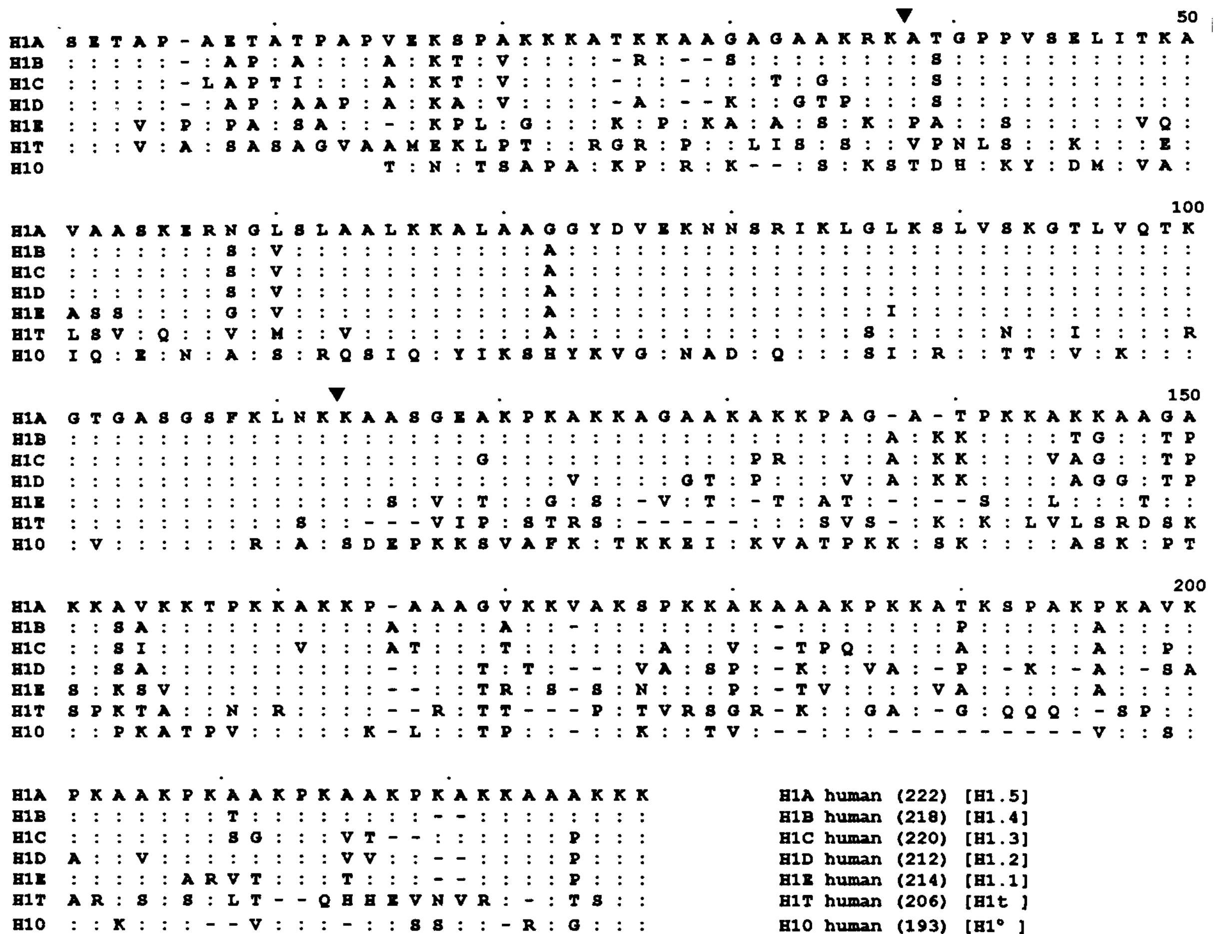

Fig. 5. Aligned amino acid sequences of the known human H1 histones. Amino acid sequences of the known human H1 histones are aligned. The H1.5 is from Ohe et al. [1989], the H1.4 and H1.3 are from Albig et al. [1991], the H1.1 and H1.2 are from Eick et al. [1989], and the H1° is from Doenecke and Toenjes [1986]. All sequences are compared to H1.5 which is the largest H1. Arrows indicate the domain boundaries; amino acids 1–37, 38–112, and 113–207. In the comparison, a colon signifies identity to the H1.5 sequence and hyphens are used to define gaps necessary for sequence alignment as in Figure 4. Numbers shown above and to the right refer to a general reference frame and the parentheses that follow each sequence indicate the protein's presumed length. The alignment was performed using the Needleman-Wunsch alignment of Genalign 5.4 by Intelligenetics.

and the presence of the H1/CCTAGG nuclear binding protein in primary spermatocytes.

Mapping the H1t gene was accomplished using PCR with template DNA samples from a human/rodent cell fusion mapping panel. Two primers were designed to amplify a portion of the H1t gene and thermal cycling parameters were optimized to preferentially amplify the human H1t gene in heterologous samples containing both human and rodent genomic DNA templates. Results from this experiment conclusively demonstrate that the human histone H1t gene maps to chromosome 6. To confirm that the product amplified from chromosome 6 was the histone H1t gene, direct PCR sequencing of genomic DNA template from the human/rodent fusion cell line GM/NA10629 was used.

The H1t gene is located on chromosome 6, but it is not known exactly where the gene is located on the chromosome. In situ experiments should further define the H1t locus in the human and the rat genome. Four human histone H1 gene loci have been identified thus far. Human histone H1 genes have been reported to be located on chromosomes 1, 6, 7, and 12 [Stephens, 1993; de Grouchy and Turleau, 1984], although a recent report suggests that all H1 genes except H1° are clustered on chromosome 6 [Albig et al., 1993]. Therefore, the histone H1 genes may be scattered, but there appears be at least two

separate histone H1 loci on chromosome 6. One of these is the H1t locus. Along these lines, the mouse testis-specific histone H1t gene is reported to be located on chromosome 13 [Owen et al., 1986]. It is interesting to note that the mouse H1t gene is linked to the bg and pe loci and the loci for the T-cell receptor gamma-chain gene (Tcrg), the prolactin gene (Prl), the Friend murine leukemia virus integration site 1 (Fim-1), the Murine Hanukuh Factor gene (Muhf/Ctla-3), and the dihdrofolate reductase gene (Dhfr) [Holcombe et al., 1991]. *Prl* and *Fim-1* have been mapped to human chromosome 6. The location of *Prl* is 6p22.2 to 6p21.3 and the location of *Fim-1* is 6p23 to 6p22.3. These data imply that the H1 locus found on chomosome 6 at 6p21 is the H1t locus. This result agrees with recently published data by Albig et al. [1993], indicating that human H1t histone gene maps to chromosome 6 at 6p21.1 to 6p22.2.

In summary, the comparison of the monkey and human gene sequences with the previously sequenced H1t genes confirmed the presence of the unique H1t/CCTAGG promoter element. Comparison of human histone H1t to the six other human H1 histones reveals important similarities and differences within the promoters of the H1 genes and within the three major domains of histone H1 coding region. This analysis has revealed highly conserved and presumably important regions of the promoter and coding region of the gene and will help direct future studies of regulation of expression and the functional and structural roles of the testis-specific histone H1t.

ACKNOWLEDGMENTS

This research was supported by the Veterans Administration and by an NIH grant 1R01HD29381-01A1 (SRG). We acknowledge the excellent technical assistance of Julie Wright.

REFERENCES

Albig W, Kardalinou E, Drabent B, Zimmer A, Doenecke D (1991): Isolation and characterization of two human H1 histone genes within clusters of core histone genes. Genomics 10:940–948.

Albig W, Drabent B, Kunz J, Kalff-Suske M, Grzeschik K-H, Doenecke D (1993): All known human H1 histone genes except the H1° gene are clustered on chromosome 6. Genomics 16:649–654.

Birnboim HC (1983): A rapid alkaline extraction method for the isolation of plasmid DNA. Methods Enzymol 100:243–255.

Birnboim HC, Doly J (1979): A rapid alkaline extraction procedure for screening recombinant Plasmid DNA. Nucleic Acids Res 7:1513–1523.

Breneman JW, Yau P, Teplitz RL, Bradbury EM (1993): A light microscope study of linker histone distribution in rat metaphase chromosomes and interphase nuclei. Exp Cell Res. 206:16–26.

Bucci LR, Brock WA, Meistrich ML (1982): Distribution and synthesis of histone 1 subfractions during spermatogenesis in the rat. Exp Cell Res 140:111–118.

Carozzi N, Marashi F, Plumb M, Zimmerman S, Zimmerman A, Coles LS, Wells JRE, Stein G, Stein J (1984): Clustering of human H1 and core histone genes. Science 224:1115–1117.

Clark JM (1988): Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. Nucleic Acids Res 16:9677–9686.

Cole KD, Kandala JC, Kistler WS (1986): Isolation of the gene for the testis-specific H1 histone variant H1t. J Biol Chem 261:7178–7183.

Cole KD, Kandala JC, Kremer E, Kistler WS (1990): Isolation of the genomic clone encoding the rat histone variant H1d. Gene 89:265–269.

Cole RD (1987): Microheterogeneity in H1 histones and its consequences. Int J Peptide Protein Res 30:433–449.

Coles LS, Robins AJ, Madley LK, Wells JRE (1987): Characterization of the chicken histone H1 gene complement. J Biol Chem 262:9656–9663.

Dalton S, Wells JRE (1988): A gene specific promoter element is required for optimal expression of the histone H1 gene in S-phase. EMBO J 7:49–56.

de Grouchy J, Turleau C (1984): "Clinical Atlas of Human Chromosomes. Second Edition." New York: John Wiley & Sons, p 465.

Doenecke D, Toenjes R (1986): Differential distribution of lysine and arginine residues in the closely related histones H1° and H5: Analysis of a human H1° gene. J Mol Biol 187:461–464.

Drabent B, Kardalinou E, Doenecke D (1991): Structure and expression of the human gene encoding testicular H1 histone (H1t). Gene 103:263–268.

Eick S, Nicolai M, Mumberg D, Doenecke D (1984): Human H1 histones: Conserved and varied sequence elements in two H1 subtype genes. Eur J Cell Biol 49:110–115.

Gallinari P, La Bella F, Heintz N (1989): Characterization and purification of H1TF2, a novel CCAAT-binding protein that interacts with histone H1 subtype-specific consensus element. Mol Cell Biol 9:1566–1575.

Grimes SR (1986): Nuclear proteins in spermatogenesis. Comp Biochem Physiol B 83:495–500.

Grimes SR, Meistrich ML, Platz RD, Hnilica LS (1977): Nuclear protein transitions in rat testis spermatids. Exp Cell Res 110:31–39.

Grimes SR, Wolfe SA, Anderson JA, Stein GS, Stein JL (1990): Structural and functional analysis of the rat testis-specific histone H1t gene. J Cell Biochem 44:1–17.

Grimes SR, Wolfe SA, Koppel DA (1992a): Tissue-specific binding of testis nuclear proteins to a sequence element within the promoter of the testis-specific histone H1t gene. Arch Biochem Biophysics 296:402–409.

Grimes SR, Wolfe SA, Koppel DA (1992b): Temporal correlation between the appearance of testis-specific DNA-binding proteins and the onset of transcription of the testis-specific histone H1t gene. Exp Cell Res 201:216–224.

Hnilica LS, Johnson AW (1970): Fractionation and analysis of nuclear proteins in sea urchin embryos. Exp Cell Res 63:261–270.

Holcombe RF, Stephenson DA, Zweidler A, Stewart RM, Chapman VM, Seidman JG (1991): Linkage of loci associated with two pigment mutations on mouse chromosome 13. Genet Res 58:41–50.

Lee YH, O'Rand MG (1993): Ultrastructural localization of a nuclear autoantigenic sperm protein in spermatogenic cells and spermatozoa. Anat Rec 236:442–448.

Lennox RW (1984): Differences in evolutionary stability among mammalian H1 subtypes; implications for the roles of H1 subtypes in chromatin. J Biol Chem 259:669–672.

Lennox RW, Cohen LH (1983): The histone H1 complements of dividing and non-dividing cells of the mouse. J Biol Chem 258:262–268.

Kaye JS, McMaster-Kaye R (1974): Histones of Spermatogenous cells in the house cricket. Chromosoma 46:397–419.

Maxson R, Cohn R, Kedes L, Mohun T (1983): Expression and organization of histone genes. Annu Rev Genet 17: 239–277.

Meistrich ML, Bucci LR, Trostle-Weige PK, Brock WA (1985): Histone variants in rat spermatogonia and primary spermatocytes. Dev Biol 112:230–240.

Meinkoth J, Wahl G (1984): Hybridization of nucleic acids immobilized on solid supports. Anal Biochem 138:267–283.

Ohe Y, Hayashi H, Iwai K (1989): Human spleen histone H1. Isolation and amino acid sequences of three minor variants, H1a, H1c, and H1d. J Biochem 106:844–857.

Osley MA (1991): The regulation of histone synthesis in the cell cycle. Annu Rev Biochem 60:827–861.

Owen FL, Taylor BA, Zweidler A, Seidman JG (1986): The murine γ-chain of the T cell receptor is closely linked to a spermatocyte specific histone gene and the beige coat color locus on chromosome 13. J Immunol 137:1044–1046.

Palmer MS, Sinclair AH, Berta AP, Ellis NA, Goodfellow PN, Abbas NE, Fellous M (1989): Genetic evidence that ZFY is not the testis-determining factor. Nature 342:937–939.

Parseghian MH, Hamkalo BA (1993): Fractionation of human histone H1 subtypes and generation of subtype-specific antibodies. FASEB J 7:243.

Ramakrishnan V, Finch JT, Graziano V, Lee PL, Sweet RM (1993): Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. Nature 362:219–223.

Schmitt JJ, Cohen BN (1983): Quantitative isolation of DNA restriction fragments from low-melting aragrose by Elutip-d affinity chromatography. Anal Biochem 133:462–464.

Schulze E, Trieschmann L, Schulze B, Schmidt ER, Pitzel S, Zechel K, Grossbach U (1993): Structural and functional differences between histone H1 sequence variants with differential intranuclear distribution. Proc Natl Acad Sci USA 90:2481–2485.

Seyedin SM, Kistler WS (1979): H1 histone sub-fractions of mammalian testes. I. Organ specificity of the rat. Biochemistry 18:1371–1375.

Stephens JC (1993): Cytogentic map of human genes. In O'Brien SJ (ed): "Genetic Maps Sixth Edition, Book 5: Human Maps." Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp 5.1–5.45.

Su H, Lau Y-FC (1993): Identification of the transcriptional unit, structural organization, and promoter sequence of the human sex-determining region Y (SRY) gene, using a reverse genetic approach. Am J Hum Genet 52:24–38.

Subirana JA (1970): Nuclear proteins from a somatic and a germinal tissue of the echinoderm *Holothuria tubulosa*. Exp Cell Res 63:253–260.

van Wijnen AJ, Wright KL, Massung RF, Gerretsen M, Stein JL, Stein GS (1988): Two target sites for protein binding in the promoter region of a cell cycle regulated human H1 histone gene. Nucleic Acids Res 16:571–592.

Wells D, Brown D (1991): Histone and histone gene compilation and alignment update. Nucleic Acids Res 19:2173–2188.

Wolfe SA, Grimes SR (1993): Histone H1t: A tissue-specific model used to study transcriptional control and nuclear function during cellular differentiation. J Cell Biochem, 53:156–160.

Yang Y-S, Brown DT, Wellman SE, Sittman DB (1987): Isolation and characterization of a mouse fully replication-dependent H1 gene within a genomic cluster of core histone genes. J Biol Chem 262:17118–17125.

Zlatanova J (1990): Histone H1 and the regulation of transcription of eukaryotic genes. TIBS 15:273–276.