Functional and molecular characterization of the transcriptional regulatory region of *Tcp-10b*, a testes-expressed gene from the *t complex responder* locus

U. Kevin Ewulonu¹, T. J. Buratynski² and John C. Schimenti¹,*

¹The Jackson Laboratory, Bar Harbor, Maine 04609, USA
²Case Western Reserve Univ. School of Medicine, Cleveland, OH 44106, USA
*Author for correspondence

SUMMARY

Mouse *t* haplotypes contain several mutant alleles that disrupt spermatogenesis. Their phenotypes include sterility, reduced fertility and transmission ratio distortion (TRD). The substantial genetic analyses of these mutant alleles, coupled with intensive physical characterization of the *t* complex, provides a fertile ground for identifying and understanding genes essential to male gametogenesis. The *t complex responder* (*Tcr*) locus plays a central role in this process, interacting with other *t* haplotype-encoded genes to mediate TRD. A candidate responder gene, *Tcp-10b*, has been cloned and subjected to molecular characterization. Here, we define the transcriptional regulatory regions of this gene in transgenic mice. A 1.6 kb (but not 0.6 kb) DNA fragment upstream of the transcription start site contains all the regulatory signals for appropriate temporal and germ cell-specific expression of this gene. Two smaller fragments within this region bound specifically to nuclear factor(s) from germ cell protein extracts in gel shift assays. This work is a step towards understanding the mechanism of *Tcp-10b* regulated expression and may ultimately help reveal a common regulatory pathway shared by other similarly expressed spermatogenic genes.

Key words: *t* complex, spermatogenesis, transgenic mice, DNA-binding proteins

INTRODUCTION

*t* haplotypes are variant forms of the *t* complex, a 15cM stretch of DNA on mouse chromosome 17 (Silver, 1985). Most *t* haplotypes contain a recessive developmental lethal mutation and males heterozygous for two complementing *t* haplotypes are sterile. Despite these properties, *t* haplotypes are prevalent in wild mouse populations because of male transmission ratio distortion (TRD). This process causes male mice heterozygous for a *t* haplotype and a wild-type form of the *t* complex (+/t) to transmit the *t* chromosome to nearly all their offspring.

TRD occurs through the action of several trans-acting *t complex distorter* (*Tcd*) loci upon a *t complex responder* (*Tcr*) locus (Lyon, 1984, 1986). *Tcr* is the only gene in mammals that mediates the production of functionally inequivalent sperm. It has been mapped to a 40-155 kb region in the center of *t* haplotypes by molecular and genetic analyses of recombinant chromosomes called partial *t* haplotypes (Bullard et al., 1992; Fox et al., 1985; Rosen et al., 1990). Chromosome walking from a linked marker led to the identification of a candidate responder gene, *Tcp-10b*, which has been sequenced and molecularly characterized (Bullard and Schimenti, 1990, 1991; Cebra-Thomas et al., 1991; Schimenti et al., 1988). *Tcp-10b* is a member of a gene family (the *Tcp-10* genes), containing three or four copies in *t* haplotypes and two to four in wild-type chromosomes (Bullard and Schimenti, 1991). *Tcp-10b* transcripts are found only in male germ cells, from the spermatocyte stage onward (Schimenti et al., 1988).

Genes that are expressed in testis can be categorized into three groups based on their patterns of transcription (Wolgemuth and Watrin, 1991). (1) Those expressed exclusively during spermatogenesis. (2) Those expressed in several tissues, but which exhibit quantitative or qualitative differences in expression during male germ cell development. (3) Those that represent a testis-specific isotype of a more generally expressed gene.

In spite of the identical (or nearly identical) expression patterns of some genes within certain classes, cis-acting elements or transcription factors responsible for the appropriate regulation have yet to be identified. The lack of spermatogenic cell lines poses a major problem for the study of germ cell transcriptional regulation. Although in vitro testis transcription systems are being developed (Bunick et al., 1990; Van der Hoorn and Tarnasky, 1992), investigations into gene regulation depend upon transgenic mouse technology (Peschon et al., 1987; Robinson et al., 1989).
The present study was undertaken to define sequence elements in the 5' flanking region of Tcp-10b that are required for appropriate transcriptional activity. The results show that these elements are located within 1.6 kb upstream of the transcription start site and essential sequences have been localized to 1.0 kb. Transcripts of a chimeric Tcp-10b/LacZ transgene were detected from the spermatocyte through elongated spermatid stage. The RNA was translated immediately as determined by histochemical staining for LacZ. Two DNA fragments in the 1.6 kb flanking region were found to bind testicular protein in an electrophoretic mobility shift assay (EMSA). This work is a step towards the identification and characterization of transcription factors critical for meiosis and/or spermiogenesis, and may help elucidate the role of Tcp-10b in these processes.

**MATERIAL AND METHODS**

**DNA sequencing**

Various DNA fragments upstream of the Tcp-10b transcriptional start site were subcloned into the plasmid Bluescript (Stratagene) and sequencing of double-stranded templates was performed by the dideoxy-chain termination procedure (Sanger and Coulson, 1977) using Sequenase (US Biochemicals, Cleveland, Ohio) according to manufacturer's instructions. Both strands were sequenced, in some cases using internal synthetic oligonucleotides rather than the polylinker primers to initiate chain extension. DNA sequence analyses and homology searches were performed using MacVector (IBI) software package.

**Gene constructs**

Two constructs were used for generating transgenic mice, differing only in the length of Tcp-10b 5' flanking DNA (Fig. 1). p1.6Tcr-Lac-SV was constructed as follows: The 1.6 kb Nhel/EcoRI fragment shown in Fig. 1 was isolated from a genomic subclone by cleaving with EcoRI, Klenow filling, cleaving with Nhel and gel purifying. It was then ligated to Smal/Xhol digested pLacC plasmid (a kind gift of Jacques Peschen) to generate the intermediate plasmid p1.6Ter-Lac. pLacC contains the bacterial LacZ gene inserted into pUC18. The 0.85 kb SV40 fragment was generated by digesting the plasmid pCarbegal (Thummler et al., 1988) with Xhol and EcoRI. After Klenow filling, the fragment was ligated into a T4 Polynucleotide-blunted PstI site 3' to the LacZ gene in p1.6Tcr-Lac. p0.6Tcr-Lac-SV was generated by digesting the p1.6Tcr-Lac-SV with XmnI (see this site in Fig. 1) and HindIII, which removed the entire insert except for the 5'-most 1 kb of the Tcp-10b gene. This excised fragment was ligated into pBluescript cut with HindIII and HincII.

**Transgenic mice**

The inserts from each of the two plasmid constructs was excised by HindIII/KpnI digestion, purified on low melting point agarose gels, further purified with GeneClean (Bio101) and microinjected into 1-cell C57BL/6 × SJL F2 embryos. The injected embryos were transferred into pseudopregnant CF1 females while still at the 1-cell stage. Transgenics from the litter were identified by Southern blot hybridization of tail DNA, using a [32P]-labeled 2.5 kb PvuII LacZ fragment as probe. For the 1.6Tcr-Lac-SV construct, six founder transgenics were identified from eighteen mice born. Three independent lines were established from these and subsequently analyzed. For the 0.6Tcr-Lac-SV construct, four founder mice were identified from seventeen mice born, and three independent lines were also established and analyzed.

**Quantitation of transgene copy number**

Genomic DNA from each line was digested with SacI, yielding fragments corresponding to the tandem repeats, an internal fragment and two additional fragments corresponding to the junction of the transgene array and the host genome. Southern blots of the digested DNAs were probed to determine how many fold more intense the tandem array or internal bands were compared to the junction bands. The number of nucleotides detected in each fragment by the probe (a 1.95 kb PvuII-SacI lacZ fragment) was taken into account in order to normalize molecule copy number against band intensity. Quantitation was performed on a phosphorimager device (Molecular Dynamics).

**RNA isolation and northern blot analyses**

Total RNA was isolated from various tissues of male transgenic mice as described (Chomczynski and Sacchi, 1987). RNA electrophoresis, transfer onto Nytran membranes (Schleicher & Schuell) and hybridization was carried out by standard procedures (Maniatis et al., 1989). The probe used for transgene detection was the same as that used for the Southern blot hybridization, whereas that used for detecting endogenous Tcp-10 RNA was a cDNA fragment called Tcr16R/H.2 (Bullard and Schimenti, 1990).

**Staining for β-galactosidase**

Testes from six mature transgenic mice were dissected and used for to prepare a cell suspension from seminiferous epithelium by a modification (Murty and Schimenti, 1991) of the collagenase/trypsin method of Romrell (Romrell et al., 1976). The cells were resuspended in 50 ml sperm medium plus 0.2% bovine serum albumin and loaded onto a 1000 ml 2-4% BSA gradient prepared in a CELSEP chamber (Dupont) with a 10% BSA cushion at the bottom. 50 ml fractions were collected. Each was numbered starting from the bottom (high density BSA). The cells were pelleted, fixed, stained for LacZ activity and photographed as previously described (Murty and Schimenti, 1991).

![Fig. 1. Map of the Tcp-10b 5' flanking region and fragments used to generate transgenic mice. Selected restriction sites are shown: B, BglII; Bm, BamHI; H, HincII; N, Nhel; R, EcoRI; and X, XmnI. There are other HincII sites besides the one indicated. The transcriptional start site is indicated by the arrow (L. Snyder, L. Silver and J. Schimenti, unpublished observations). The DNA fragments used as promoter fragments in the transgenic constructs are shown as stippled lines above the map. The fragments that bind nuclear factor(s) are indicated, as is the position of a nanonucleotide resembling others present in the Pgk-2 and Zfp-35 genes.](image-url)
Gel shift assay
Nuclear extracts were prepared as described (Dignam et al., 1983) from liver, spleen and testicular cells enriched for spermatogonia, spermatocytes and spermatids of male mice. After ammonium sulfate precipitation for 30 minutes at 4°C and ultracentrifugation at 123,342 g for 25 minutes at 4°C, the pellet was resuspended in 1 ml of buffer D (20 mM Hepes-KOH, pH 7.9, 20% (v/v) glycerol, 100 mM KCl, 200 mM EDTA, 0.5 mM DTT, 0.5 mM PMSF) and dialyzed twice against buffer D over a 24 hour period. Samples were centrifuged for 10 minutes at 4°C to remove precipitated protein. 15 µg of each protein extract was incubated at 30°C for 30 minutes with 1 µg nonspecific DNA (lambda/BstElI) in a 15 µl volume containing 12% glycerol, 12 mM Hepes (pH 7.9), 60 mM KCl, 1 mM EDTA, 1 mM DTT and 1–5 µl of a 5' end labelled Tcp-10b fragment. The mixtures were loaded on a 4% non-denaturing polyacrylamide gel (30:1 acrylamide:bis) and electrophoresed at 125 volts for 2.5 hours. The gel was then dried and autoradiographed.

RESULTS
Nucleotide sequence of the Tcp-10b 5' flanking region
The nucleotide sequence of 1.6 kb of the Tcp-10b gene 5' flanking region is shown in Fig. 2. Similar to other testis-expressed genes, Tcp-10b does not possess canonical TATA or CCAAT boxes in the typical locations upstream of the transcriptional start site (Boer, 1987; Cunliffe et al., 1990; Galliot, 1989). Curiously, there are three closely spaced TATA sequences in the region between –104 and –86. Since they are further upstream than the typical 35 bases or so, their functional significance is questionable.

Transcriptional regulatory sequences defined in transgenic mice
To define the extent of 5' flanking sequence required for tissue-specific transcription, we established transgenic mice carrying the LacZ coding sequence flanked 5' by either 1.6 kb or 0.6 kb of Tcp-10b upstream sequence. Northern blot hybridization of total RNA isolated from various male transgenic tissues is shown in Fig. 3. A transcript of about 3.9 kb, exclusive to testis, was observed in all three transgenic lines carrying the 1.6 kb fragment. No hybridization was observed in the other tissues, in wild-type mice or in any of the three transgenics carrying the 0.6 kb upstream fragment. Two of the three 1.6Tcr-Lac-SV lines contained 2 copies of the transgene and the other had 3 copies. Two of the three 0.6Tcr-Lac-SV lines contained 3 copies of the transgene and the other had 4 copies. The results demonstrate that the 1.6 kb fragment contains the sequence information necessary for tissue-specific transcription of the transgene. The inability of the 0.6 kb fragment to direct

Fig. 2. Nucleotide sequence of the Tcp-10b 5' flanking region. Restriction enzyme sites relevant to making constructs are indicated. The multiple TATA sequences are underlined and the transcriptional start site is numbered +1. The nanonucleotide resembling those present in the Pgk-2 and Zfp-35 genes is double underlined.
expression in three independent lines indicates that essential transcriptional regulatory sequences of the Tcp-10b gene reside in the 1.0 kb region between the EcoRI and XmnI sites (Fig. 1).

To determine the stage of spermatogenesis at which the LacZ transcripts first appear, RNA was prepared from testes of transgenic mice ranging in age from 12.5 days old to adulthood (>60 days). As shown in Fig. 4 (top panel), LacZ transcripts first appear at day 14.5. This correlates with the appearance of pachytene spermatocytes (Bellve, 1979), which is in accordance with previously reported onset of Tcp-10 transcription (Schimenti et al., 1988). To confirm this correlation and compare transgene transcription to that of the endogenous Tcp-10 genes in these particular animals, the northern blot was stripped and rehybridized to a Tcp-10 cDNA probe (Fig. 4, bottom panel). The developmental expression patterns were identical.

**Assay for LacZ activity in testicular cells**

To identify the spermatogenic cell type(s) in which translation of the transgene mRNA occurs, LacZ activity in 1.6Tcr-Lac-SV mice was monitored as a function of neonatal development via X-gal staining. LacZ staining was observed in testicular cells from day 14.5 and older transgenics, but not in day 12.5 or younger animals. No staining was seen in adult 0.6Tcr-Lac-SV mice (Fig. 5D). To determine the cell types containing LacZ activity in the 1.6Tcr-Lac-SV mice, staining was performed on testicular cell samples fractionated by unit-gravity sedimentation. As shown in Fig. 5, staining was seen in spermatocytes (Fig. 5A), round spermatids (Fig. 5B) and elongating spermatids (Fig. 5C). We were unable to detect staining of mature spermatids. These results suggest that translation of the transgene mRNA occurs in all the cell types where it is transcribed. Alternatively, it is possible that lacZ activity produced in spermatocytes persists throughout spermiogenesis in cells that no longer translate the transgene RNA.

The appearance of LacZ activity in spermatocytes coincides with the onset of transcription in day 14.5 transgenic mice (see Fig. 4), demonstrating that this particular hybrid messenger RNA is not subject to post-transcriptional regulation. The temporal translation pattern of endogenous Tcp-10 has not been reported.

**Sequences in the Tcp-10b′ promoter that bind nuclear factor(s)**

A search for potential regulatory elements that bind specific nuclear factors was conducted using the electrophoretic mobility shift assay (EMSA). Various 5′ fragments corresponding to the Tcp-10b′ DNA present in the 1.6Tcr-Lac-SV construct were assayed. The results are shown in Fig. 6. Specific binding with testicular nuclear extracts from spermatocyte- and spermid-enriched fractions were observed with a 158 bp BglII-HincII and a 264 bp BamHI-HincII fragment (Figs 1, 6). The BglII-HincII fragment yielded three retarded bands (a, b, c), which could be eliminated by a 100-fold excess of the unlabeled fragment as competitor. The BamHI-HincII fragment yielded one retarded band, ‘a’, which could be removed by self-competition and another, ‘b’, which could not. An upstream, unlabeled DNA fragment could not eliminate the specific bands generated by either fragment (Fig. 6, lane 3a,b). At present, it is unknown whether the multiple retarded bands formed by the BglII-HincII fragment represent the binding of different proteins or the same factor(s) as monomeric and multimeric units. No binding was observed with the extracts prepared from either liver, spleen (Fig. 6, lanes 4a, 5a, 5b, 6b), or the spermatogonium-enriched fraction (data not shown), suggesting that the factor(s) binding to the above two fragments have a considerable level of tissue specificity.
DISCUSSION

Transcriptional control of spermatogenesis genes

Our understanding of gene regulation during spermatogenesis is relatively poor. This is due to the inability to maintain or differentiate spermatogenic cells in culture. Descriptive characterization of transcriptional patterns represents the bulk of existing data in this area. The compiled data reveal groups of similarly regulated genes (Wolgemuth and Watrin, 1991). These are activated at different times during spermatogenesis: some in spermatogonia, others in spermatocytes and others in the haploid stages. In addition, some genes are active both in testes and somatic cells, whereas others are exclusive to testis. The Tcp-10b gene is transcribed predominantly or exclusively in testicular germ cells and is first transcribed in spermatocytes. This pattern is shared by other genes, including tcte-1 (Sarvetnick et al., 1989), 177c3 (Rappold et al., 1987) and Pgk-2 (Gold et al., 1983).

More detailed investigations into spermatogenic gene regulation have utilized transgenic mice to characterize the promoter regions of testis-expressed genes. Appropriate temporal and spatial expression of the mouse protamine-1 (mP1), mouse protamine-2 (mP2) and human phosphoglycerate kinase-2 (PGK2) genes can be directed with a maximum of 425 bp (Peschon et al., 1989), 859 bp (Stewart et al., 1988) and 323 bp (Robinson et al., 1989), respectively, upstream of the transcriptional initiation site.

![Fig. 5. LacZ activity staining of fractionated transgenic testicular cells. Cel-Sep fractions of 1.6Tcr-Lac-SV mice enriched for (A) spermatocytes, (B) round spermatids and (C) elongating spermatids. (D) lacZ activity staining of unfractionated 0.6Tcr-Lac-SV cells. Arrowheads in C show an individual elongating spermatid. All animals were homozygous for the transgene insertions. Microphotographs were taken at 250× magnification.](image1)

![Fig. 6. Electrophoretic mobility shift assay. Nuclear protein was incubated with (left panel) a 158 bp BglII-HincII fragment (Fig. 1), or (right panel) a 264 bp BamHI-HincII fragment (Fig. 1). ‘+’ indicates the presence of extract or self/non-self competing DNA (all contain 1 µg of lambda DNA - see Materials and methods), while ‘−’ indicates the absence of same. Lanes 1-3 contain testicular nuclear extracts while lanes 4-5 (A) and 5-6 (B) were contain extracts from liver and spleen, respectively. Lane 3 in both panels contains 100 ng of a 255 bp EcoRI/RsaI fragment as control competitor. This fragment corresponds to the first 255 bases in Fig. 2. Lane 2 in both panels contains 100 ng of unlabelled probe fragment (specific competitor).](image2)
Genes sharing similar expression patterns may be activated by a common regulatory or signalling mechanism, possibly involving identical transcription factors. If this is the case, each of these genes would be expected to share common DNA-binding sites for such factors. The promoter studies in transgenic mice, coupled with sequence comparisons of promoter regions, has led to identification of potential regulatory sequences. Cunliffe et al. (1990) noted that the Zfp-35 and Pgk-2 genes, which have parallel transcription patterns in the testis (transcription of both begins in pachytene spermatocytes, although Zfp-35 is also expressed in somatic cells), have similar sequence elements in their 5' flanking regions. Zfp-35 has the sequence GGGTTTTCAT located at −226 to −215, which is identical at 10 of 12 positions to the sequence AGGTTTTTA-CAT located at −512 to −501 upstream of Pgk-2. The nonanucleotide GGGTTTTCAT resides upstream of Tcp-10b at positions −207 to −216, which differs by one nucleotide compared to either Zfp-35 or Pgk-2. It will be interesting to see whether other similarly expressed genes also contain this motif. Functional evaluation via in vitro mutagenesis will provide direct tests of such sequences.

Identification of male germ cell transcription factors

Analysis of transgenic mice carrying various lengths of the Tcp-10b gene 5' flanking region has enabled us to define a 1.0 kb segment containing essential transcriptional control element(s) for spermatocytes and more developed spermatogenic cells. Two promoter fragments bound to nuclear factor(s) in protein extracts from testis. Additional experiments will be required to determine whether binding of these proteins is required for effective transgene expression. The fact that the construct that was missing one of the DNA elements (0.6Tcr-Lac-SV) was incapable of transcription is consistent with this possibility.

Research into spermatogenic transcription factors is still in early stages. To date, no testis-specific transcription factors have been isolated. However, specific interactions of germ cell nuclear proteins with the RT7 and Pgk-2 promoters have also been identified (Gebara and McCarrey, 1992; Van der Hoorn and Tarnasky, 1992). As progress on the regulation of these and other testis-expressed genes advances, the potential existence of shared regulatory pathways may be identified. With regards to Tcp-10b, elucidation of its regulatory mechanisms will be important for understanding its potential involvement in TRD of t haplotypes.

We thank Peter Harte for providing the pC4βgal plasmid. This work was supported by grants from the Searle Scholars Program/The Chicago Community Trust to J. S., the Lucille Markay Charitable Trust to Case Western Reserve University and NIH grant R01 HD24374-02 to J. S. This work was performed in the Department of Genetics at Case Western Reserve University.

REFERENCES


mouse chromosome 17 gene encodes a testes-specific transcript with unusual properties. *Immunogenetics* 30, 34-41.


(Accepted 11 September 1992)