THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 278, No. 8, Issue of February 21, pp. 6168–6174, 2003
Printed in U.S.A.

Fundamental Cellular Processes Do Not Require Vertebrate-specific Sequences within the TATA-binding Protein*

Received for publication, November 3, 2002
Published, JBC Papers in Press, December 5, 2002, DOI 10.1074/jbc.M211205200

Edward E. Schmidt‡§, Alla A. Bondareva‡, Jay R. Radke‡¶, and Mario R. Capecchi‖

From the ‡Department of Veterinary Molecular Biology, Marsh Laboratories, Montana State University, Bozeman, Montana 59717 and the ‖Howard Hughes Medical Institute, Eccles Institute of Human Genetics, University of Utah, Salt Lake City, Utah 84112-5331

The 180-amino acid core of the TATA-binding protein (TBPCORE) is conserved from Archae bacteria to man. Vertebrate TBPs contain, in addition, a large and highly conserved N-terminal region that is not found in other phyla. We have generated a line of mice in which the *tbp* allele is replaced with a version, $tbp^{\Delta N}$, which lacks 111 of 135 N-terminal amino acid residues. Most $tbp^{\Delta N/\Delta N}$ fetuses die in midgestation. To test whether a disruption of general cellular processes contributed to this fetal loss, primary fibroblast cultures were established from +/+, ΔN/+, and ΔN/ΔN fetuses. The cultures exhibited no genotype-dependent differences in proliferation or in expression of the proliferative markers dihydrofolate reductase (DHFR) mRNA (S phase-specific) and cdc25B mRNA (G_2-specific). The mutation had no effect on transcription initiation site fidelity by either RNA polymerase II (pol II) or pol III. Moreover, the mutation did not cause differences in levels of U6 RNA, a pol III-dependent component of the splicing machinery, in mRNA splicing efficiency, in expression of housekeeping genes from either TATA-containing or TATA-less promoters, or in global gene expression. Our results indicated that general eukaryotic cell functions are unaffected by deletion of these vertebrate-specific sequences from TBP. Thus, all activities of this polypeptide domain must either be compensated for by redundant activities or be restricted to situations that are not represented by primary fibroblasts.

Gene expression is a fundamental property of all living systems. Archae bacteria use only a single RNA pol[1] to transcribe mRNA, rRNA, and tRNA. Eukaryotes evolved three separate RNA pols with different specialties to perform these functions. TBP acts in promoter recognition for transcription initiation by the Archaea RNA pol as well as by all three eukaryotic RNA pols (1–3). In accordance with having such an important and conserved function, the 180-amino acid TBPCORE from Archaea and from all eukaryotes share both a high degree of amino acid similarity and very similar crystal structures (3–5).

The differences in complexity observed between Archaea and eukaryotes, or between single-celled and higher eukaryotes, are coincident with differences both in genomic complexity and in the complexity of the gene expression machinery. Similarly, although the general role of TBP in transcription initiation has been conserved, this function has become more complex. In Archaea TBP appears to work as a single-subunit entity (6); in eukaryotes TBP is at the core of obligate large multiprotein complexes (3). The complex SL1 functions in initiation by RNA pol I, TFIID is required for production of mRNA by RNA pol II, TFIIIB is used during transcription of tRNAs and some other small RNAs by RNA pol III. In mammals, TBP also interacts with SNAPc to direct production of small nuclear RNAs (snRNAs) by either pol II or pol III (3).

Attached to the TBPCORE, vertebrate species share a large N-terminal domain that differs from TBP domains in all other phyla (3, 7–11). Previous studies have implicated the vertebrate N terminus in general processes that are important for fundamental cellular activities. For example, biochemical analyses have demonstrated a role of the vertebrate N terminus in the function of the SNAPc complex (3, 12–14). Kinetic studies indicate that the N-terminal domain influences TATA-binding and DNA-bending by TBP (15). Another study suggests that the N terminus plays a role in determining rates of cell proliferation (16).

Because these fundamental cellular functions are common to most eukaryotes, including non-vertebrate metazoans, they evolved long before the vertebrate TBP N terminus.[2] Therefore, we hypothesized that this domain must have, in addition to any general functions, a vertebrate-specific role. To test this, we created a line of mice in which the wild type *tbp* allele was replaced with a version, entitled $tbp^{\Delta N}$, in which 111 of the 135 vertebrate-specific N-terminal amino acids were deleted (11). At gestational day (g.d.) 9.5, $tbp^{\Delta N/\Delta N}$ fetuses are generally normal (11). Over 90% of these mutant fetuses die between g.d. 10.5 and 12.5; loss results from a failure of their placenta to evade a maternal immune rejection response (11). Importantly, although less than 3% of the $tbp^{\Delta N/\Delta N}$ animals survive to weaning, these survivors are healthy and fertile (11).

Even though the defects in homozygous mutant mice are manifested as a highly specific and restricted phenotype, it is untested whether the mutation actually compromises general

* This work was supported by grants from the March of Dimes Foundation, National Institutes of Health, National Science Foundation, and USDA Animal Health Funds (to E. E. S.) and by awards from the Howard Hughes Medical Institute and the Mathers Charitable Foundation (to M. R. C.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Funded as a Basil O'Connor New Investigator of the March of Dimes Foundation, a Howard Hughes Medical Institute Fellow of the Life Sciences Research Foundation, and as an Investigator of the Montana Agricultural Experiment Station. To whom correspondence should be addressed. E-mail: eschmidt@montana.edu.

¶ Supported by a postdoctoral fellowship from the USDA and by an appointment in the laboratory of M. W. White.

[1] The abbreviations used are: pol, polymerase; TBP, TATA-binding protein; MEF, mouse embryo fibroblast; RT, reverse transcriptase; SAGE, serial analysis of gene expression; EST, expressed sequence tag; g.d., gestational day; DHFR, dihydrofolate reductase; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; TRF, TBP-related factor.

[2] A. A. Bondareva and E. E. Schmidt, submitted for publication.

cell functions in a sublethal manner. For example, one might hypothesize that all cells bearing this mutation would be physiologically compromised; however, only in the placenta are such defects critical to the function of the organ and the survival of the organism. To test this possibility, we analyzed fundamental properties of cells bearing this mutation. We show that $tbp^{\Delta N/+}$ and $tbp^{\Delta N/\Delta N}$ primary mouse embryo fibroblasts (MEFs) are normal in all measured general functions. Our data indicate that, if the TBP N terminus truly participates in any fundamental functions *in vivo*, these processes must be functionally redundant.

EXPERIMENTAL PROCEDURES

Mouse and MEF Production—The targeting and mouse production strategy for producing the TBP-ΔN mice has been reported elsewhere (11). The selectable *neo* cassette used for gene targeting was removed from the line by Cre/*lox*P, and for all samples in this article, the mutation had been back-crossed onto a C57Bl/6 background for ≥5 generations (11). For establishing primary fibroblast cell lines, whole uteri were harvested from females with timed pregnancies and were surface-sterilized with 70% ethanol. Deciduae were opened and embryonic surface tissue was harvested and dissociated with trypsin. Explanted cells were cultured in Dulbecco's modified Eagle's medium + 10% newborn calf serum + 1× penicillin/streptomycin/fungizone (BioWhittaker). Cultures were trypsinized and passed to fresh dishes when confluent, after 5–14 days. By this time under these conditions, fibroblasts had outgrown other surviving cell types to the point that only cells with fibroblastic morphology could be readily found under the microscope. These cultures were split in experiments.

RNase Protection, RT-PCR, and Primer Extension Assays—RNA harvests and RNase protection and primer extension assays were preformed as described previously (18, 19). Probes for DHFR (18), TFIIB, L7, GAPDH, Oct-1, and NF-Yb (19) have been described previously. For TBP-related factor-2 (TRF-2), a probe hybridizing to sequences between positions 705 and 833 of the mouse TRF-2 cDNA (20) was transcribed from a clone that we isolated by RT-PCR on mouse testis RNA.[3] A probe that hybridized to mouse U6 RNA sequences extending from the transcription initiation site to a position 92 bases downstream of the initiation site (21) was transcribed from a genomic clone that we isolated by PCR of mouse genomic DNA.[3] For β-actin, mouse genomic sequence was not available. A mouse genomic fragment,[3] extending from a position in exon 2 corresponding to base 89 of the mouse cDNA (22) to a position 51 base pairs into intron 2, was amplified from mouse genomic DNA using one primer designed from mouse exon 2, and a second primer designed from the genomic sequence of rat intron 2 (23), and this fragment was cloned into a plasmid vector. The β-actin probe hybridizes to 146 bases of the mouse pre-mRNA and 95 bases of the mouse mRNA (Fig. 3*B*). For TBP, a probe that was complementary to sequences extending from the *Bgl*II site in exon 2 (position 37 of the original published cDNA sequence) to position 168 in exon 3 (24) was transcribed from a subclone of the mouse TBP cDNA. This probe hybridized to 131 bases of wild-type TBP mRNA, and 81 bases of TBP-ΔN mRNA (Fig. 1*B*). For primer extensions, the following primers were used: U6 primer, 5′-ATC GAA TTC ACG AAT TTG CTG GTC ATC C-3′; β-actin primer, 5′-TGG GGT ACT TCA GGG TCA GGA TAC-3′.

For RT-PCR reactions, 2.5 μg of CsCl-purified DNA-free RNA (19) was used as template for oligo(dT)-primed reverse transcriptase reactions as described previously (25). PCR with primers spanning from exon 2 to exon 3 of the mouse β-actin gene confirmed that all cDNAs had uniform levels of β-actin cDNA and lacked detectable genomic DNA (data not shown). PCR was performed using 0.2% of each RT reaction per PCR reaction, as described in the figure legends, using mouse cdc25B-specific primers having the following sequences: 5′ primer, 5′-ATT CCA GCT CTG CCC AAG CTT TGG C-3′; 3′ primer, 5′-TCC ACA AAT CCG TCA TCT TCT TCA-3′. The PCR product spans from positions 626 to 821 of the mouse cdc25B mRNA (GenBank™ XM_123867; translation initiation is at position 607), which contains a centrally located splice junction for a 621-base pair intron at position 712 (GenBank™ NW_000178). Thus, genomic DNA gives a 816-base pair band that is clearly discernable from the 195-base pair cDNA signal (data not shown).

Serial Analysis of Gene Expression (SAGE)—SAGE libraries were prepared as described previously (26). Clones of individual plasmids containing SAGE tag concatenates were screened by PCR and clones containing inserts of >500 bp (roughly 30 SAGE tags) were arbitrarily chosen for sequence analysis. Data were ordered by SAGE2000 software (www.sagenet.org). Data in the article represent all tags that appeared ≥10 times in the combined libraries and represent all 46 genes present at an average of ≥1.5 parts per thousand. Summation of tag frequencies indicated that 77% of all tags sequenced were represented by these 46. The identities of mRNAs represented by individual tags was determined using the tag-to-gene mapper program on the NCBI SAGEmap homepage (www.ncbi.nlm.nih.gov/SAGE/SAGEtag.cgi).

RESULTS

mRNA and Protein Expression from the $tbp^{\Delta N}$ Allele—The $tbp^{\Delta N}$ allele retains all known transcriptional and post-transcriptional regulatory signals of the wild-type *tbp* allele (25, 27, 28), but it produces a version of TBP protein lacking 111 of 135 vertebrate-specific N-terminal amino acids (Fig. 1*A*, described in Ref. 11). Retention of the first 24 amino acids was predicted to preserve the *in vivo* stability of the mutant protein (11, 29). Thus, the allele is identical to the native *tbp* allele except that it encodes an epitope tag in place of amino acids 25–135 of the protein, and it contains a *lox*P-containing oligonucleotide insertion roughly 2.5-kb upstream of the gene (11).

To investigate the molecular consequences of removing the N terminus of TBP on basal cellular functions, we produced primary MEFs from g.d. 8.5–11.5 embryos. RNase protection assays using a probe that spanned the site of the ΔN mutation were used to compare levels of wild-type and mutant TBP mRNA accumulation in $tbp^{+/+}$, $tbp^{\Delta N/+}$, and $tbp^{\Delta N/DN}$ MEFs. Results showed that homozygous wild-type and homozygous mutant cells had similar levels of TBP or TBP-ΔN mRNA, respectively, and, in heterozygous cells, mRNA from each allele was equally represented (Fig. 1*B*). Western blots using either whole cell (Fig. 1*C*) or nuclear (data not shown) extracts indicated that MEFs of all three genotypes contained similar levels of total and nuclear TBP and TBP-ΔN protein; heterozygous MEFs exhibited similar amounts of both TBP and TBP-ΔN protein. Based on mRNA level, protein level, and subcellular distribution, TBP and TBP-ΔN proteins had no significant differences in synthesis/turnover rates or nuclear localization.

Previously, a mammalian *tbp*-gene family member, named TBP-like protein (TLP) (30) or TRF-2 (31, 32) was identified. The TRF-2$_{CORE}$ is far less closely related to the homologous region of mouse TBP (39% amino acid identity) than mouse TBP$_{CORE}$ is to yeast TBP$_{CORE}$ (81% identity; Fig. 2, Ref. 30). Different annotations of the human genome suggest that this may be the only other *tbp* family member in mammals (33) or that there may be one other more distantly related protein in the genome (34). We measured TRF-2 mRNA levels in adult mouse tissues and in MEFs of all three genotypes. Results showed that TRF-2 mRNA is expressed in wild-type MEFs (Fig. 2). TRF-2 was also expressed in all wild-type somatic mouse tissues examined at levels about 4-fold higher than those of TBP mRNA in each tissue and, consistent with previously published reports (32), TRF-2 mRNA was particularly enriched in testis. TRF-2 mRNA levels were unaffected by the TBP-ΔN mutation in both the MEFs and in adult tissues (Fig. 2). Because TRF-2 contains only a short N-terminal region fused to the core (30), we consider it highly unlikely that TRF-2 compensated for deletion of the N-terminal region of TBP in $tbp^{\Delta N/\Delta N}$ fibroblasts or mice.

Basal Properties of Cells Lacking the TBP N Terminus—Because we have shown that $tbp^{\Delta N/\Delta N}$ mutant animals occasionally survive to become healthy fertile adults (11), it is unlikely that large global defects result from the mutation. However, previous studies had implicated this domain in basal processes (see the Introduction). These studies led us to ponder

[3] E. E. Schmidt, unpublished data.

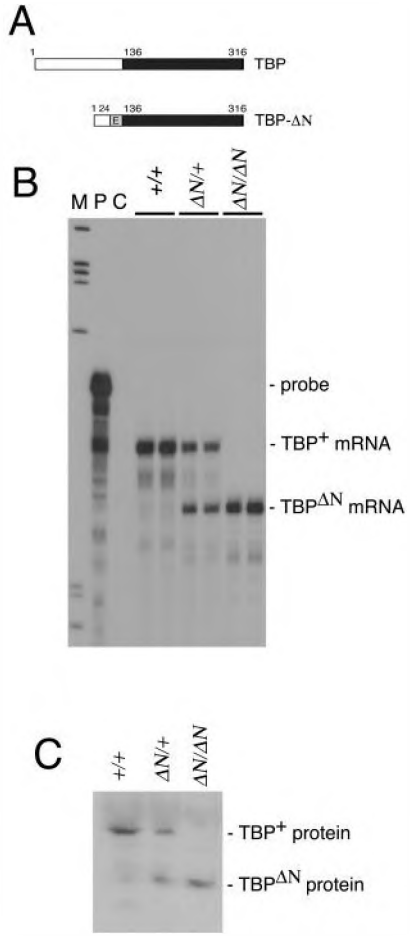

FIG. 1. **Primary structure and expression of TBP and TBP-ΔN in MEFs.** *A*, alignment of wild-type TBP and TBP-ΔN, showing the universal C terminus (*black box*) and the vertebrate-specific N terminus (*white box*). The mutant version retains amino acids 1–24 and 136–316 of the wild-type allele, and contains two copies of the 9-amino acid FLAG (Kodak) epitope tag (*gray box* labeled *E*) (11). *B*, expression of TBP mRNA in MEFs. Total RNA (10 μg) from two cultures each of $tbp^{+/+}$, $tbp^{\Delta N/+}$, and $tbp^{\Delta N/\Delta N}$ MEFs was analyzed by RNase protection using a probe that differentiates between wild-type and mutant TBP mRNA. All samples were supplemented to contain a total of 50 μg of RNA with yeast RNA. MEF genotypes are indicated at *top*. At *right* is indicated the positions of undigested probe and of the protected fragments for wild-type (TBP^+) and mutant ($TBP^{\Delta N}$) mRNA. Abbreviations: *M*, molecular size markers; *P*, 1:100 dilution of undigested probe; *C*, control containing probe hybridized to 50 μg of yeast RNA. *C*, Western blot of TBP and TBP-ΔN protein expression in MEFs. Whole cells of the indicated genotypes were harvested by scraping in 1% SDS, and lysates were sonicated and boiled. Nucleic acid content was estimated by reading the A_{260} on each lysate, and 20 μg of nucleic acid equivalents was boiled with an equal volume of Laemmli buffer for each determination (17). Western blots were probed with commercial antibody against the C-terminal domain of human TBP (BD PharMingen), which shows 100% amino acid identity with this region of mouse TBP and was visualized by ECL (Amersham Biosciences).

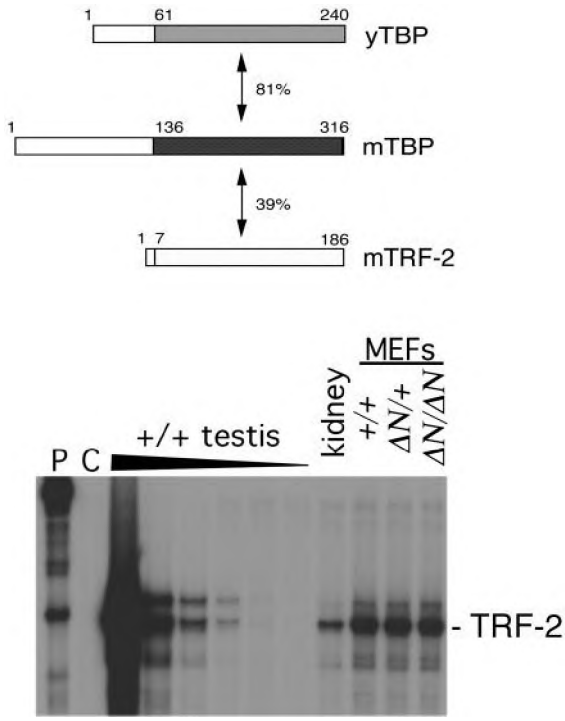

FIG. 2. **Mouse TRF-2 expression in MEFs.** At *top* is diagramed the primary structure and the C-terminal (*shaded regions*) amino acid conservation between yeast TBP (yTBP), mouse TBP (mTBP), and mTRF-2. *Below* is an RNase protection assay on dilutions of total RNA from wild-type mouse testis (10, 3.0, 1.0, 0.3, 0.1, and 0.03 μg from *left* to *right*), and on 20 μg each of total RNA from wild-type mouse kidney and from MEFs of the indicated genotypes. All samples were supplemented to 50 μg with yeast RNA. The TRF-2-specific RNase protection fragment is indicated at the *right*. Abbreviations as in the legend to Fig. 1.

whether, rather than only functioning in a particular situation (pregnancy), the N terminus functioned ubiquitously, but defects resulting from its deletion were sublethal in most cell types. Therefore, we tested whether defects in basal functions could be detected in cells bearing the TBP-ΔN mutation.

We first asked whether we could measure defects in SNAPc-dependent processes. In mammals, SNAPc is involved in transcription initiation on genes encoding snRNAs by pol III from TATA-containing promoters, and by pol II from TATA-less promoters (3). A classic example is its role in transcription of the human U2 gene by RNA pol II, and the U6 gene by pol III (35). RNase protection assays showed that U6 RNA levels in MEFs of the three genotypes were indistinguishable (Fig. 3*A*).

Splicing requires an elaborate ribonucleoprotein system that, in addition to U6 and U2, requires other snRNAs (36–39). As a potentially more global assay for whether any defects in snRNA expression resulted from the TBP-ΔN mutation, we examined the splicing efficiency of pre-mRNA encoding β-actin by RNase protection (Fig. 3*B*). Levels of correctly spliced cytoplasmic mRNA (95-base band) were unaffected by the mutation, and there was no evidence for increased nuclear accumulation of unspliced pre-mRNA (146-base band in lanes designated *Nu*). This indicated that splicing efficiency was similar between MEFs of the three genotypes. We concluded that deletion of TBP amino acids 25–135 was inconsequential for U6 RNA expression and mRNA splicing in the MEFs. Previous studies have shown that the first 55 amino acids of human TBP are sufficient for SNAPc functions *in vitro*, including U6 transcription (40). These data, in combination with our results, might suggest that SNAPc-dependent functions of the TBP N terminus reside entirely in the first 24 amino acids of the protein. Alternatively, it is possible that other compensatory interactions at the snRNA promoters stabilize SNAPc interactions in cells lacking the TBP N terminus (see "Discussion").

In one study, chicken DT40 cells that were heterozygous-null for TBP were shown to have lower proliferation rates than did the wild-type progenitor cell line, which correlates to reduced levels of mRNA and protein encoding cdc25B (16). Upon stably transfecting these cells with vectors that express wild-type TBP, proliferation rates match the progenitor cells. However, clones transfected with vectors that express a mutant version of TBP lacking all except the first 24 amino acids of the N

FIG. 3. **Expression of U6 RNA and splicing efficiency in MEFs.** *A*, U6 RNA expression. 5 μg of total RNA from MEFs of the indicated genotypes was supplemented to 50 μg with yeast RNA and assayed by RNase protection. The positions of undigested probe and U6 RNA are indicated at *left*. *B*, β-actin pre-mRNA splicing efficiency. At *top* is diagrammed the single-probe internally controlled RNase protection strategy for differentiating prespliced and spliced β-actin RNA. 5 μg each of cytoplasmic (*Cy*) and nuclear (*Nu*) RNA was supplemented to 50 μg with yeast RNA as above. The positions of undigested probe, unspliced pre-mRNA and spliced mRNA are indicated at the *left*. Abbreviations as in the legend to Fig. 1.

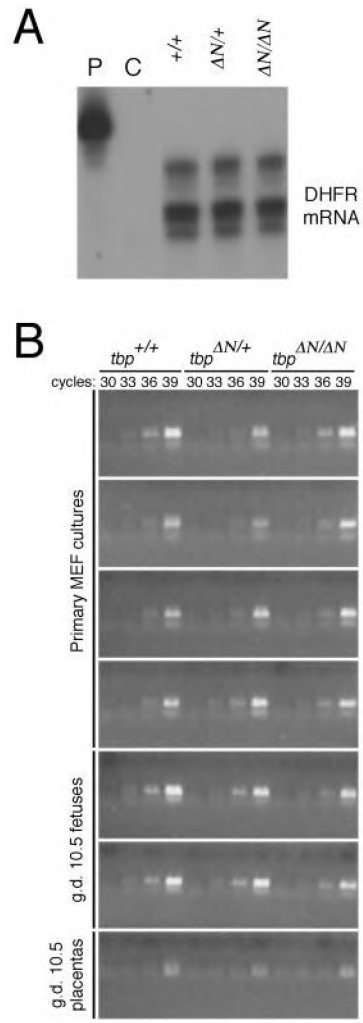

FIG. 4. **Expression of proliferative cell markers in MEFs and in tissues.** *A*, DHFR mRNA expression. 30 μg of total RNA from MEFs of the indicated genotypes was supplemented with yeast RNA to 50 μg and assayed by RNase protection using the DHFR-specific probe. *B*, cdc25B mRNA expression. Total RNA from four independent cultures each of wild-type ($tbp^{+/+}$), heterozygous ($tbp^{\Delta N/+}$), and homozygous mutant ($tbp^{\Delta N/\Delta N}$) MEFs, two individual g.d. 10.5 whole fetuses of each genotype, and one g.d. 10.5 whole placenta of each genotype (indicated at left; 21 independent tissue/cell samples shown) was assayed by RT-PCR. Reactions were terminated at the indicated PCR cycles (*top*) to generate an expression curve for each, showing that the reactions were not saturated. The PCR product spans a 621-base pair intron. Only the 195-base pair cDNA-derived sequence, not the 816-base genomic product, was detected, which confirmed that samples were not contaminated by genomic DNA. Control reactions using primers that spanned from exon 2 to exon 3 of mouse β-actin showed equivalent signals from all samples (not shown).

terminus exhibit an intermediate proliferation rate (16). These results suggest that the N terminus is required for full proliferative potential of the cells. Importantly, all of the cells in the study by Um *et al.* (16) contained one intact copy of the endogenous *tbp* gene including the entire N-terminal region, and thus, the anti-proliferative effect was manifested as a haploid-insufficiency.

We established MEF cultures from 139 embryos of heterozygous crosses, but we were unable to detect any significant correlation between genotypes and proliferation rates by either growth curves or thymidine labeling (data not shown). Even in comparing homozygous mutant cells, where both copies of *tbp* lacked the N terminus, and homozygous wild-type cells, proliferation rates were indistinguishable. As a potentially more sensitive molecular assay of proliferation, levels of the mRNA encoding DHFR, which is transcribed only during S-phase of the cell cycle (18, 41, 42), were measured. MEFs of all three genotypes exhibited similar levels of DHFR mRNA (Fig. 4*A*).

Previous studies have shown that cdc25B mRNA is expressed only during G_2 phase of the cell cycle (43, 44) and that expression of this mRNA may be particularly sensitive to mutations in the *tbp* gene (16). Therefore, as another molecular marker of proliferation rates, we compared relative cdc25B mRNA levels in $tbp^{+/+}$, $tbp^{\Delta N/+}$, and $tbp^{\Delta N/\Delta N}$ MEFs, as well as in g.d. 10.5 fetuses and placentas of all three genotypes. Results showed that cdc25B mRNA levels were similar between MEFs of all three genotypes, between fetuses of all three gen-

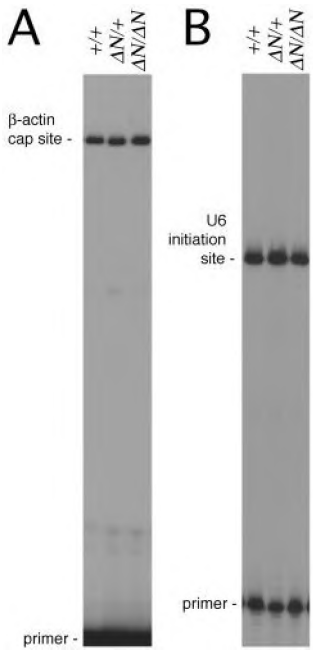

FIG. 5. **Transcription initiation site fidelity in MEFs.** Poly(A$^+$)-selected mRNA (*panel A*) or total RNA (*panel B*) was analyzed by primer-extension assays using either the β-actin primer (the β-actin gene is transcribed by pol II; *panel A*) or the U6 primer (the *u6* gene is transcribed by pol III; *panel B*). Positions of free primers and primer extension products are indicated at the *left* of each panel.

otypes, and between placentas of all three genotypes (Fig. 4*B*). The genotype-independent reduced level of cdc25B mRNA in placentas as compared with those in MEFs and fetuses likely represents a lower overall rate of proliferation in this organ. These data suggest that the TBP N terminus does not play a significant role, either as a dominant (*i.e.* haploid-insufficient) or recessive trait, in determining proliferation rates in MEFs.

Transcription and RNA Processing in tbp$^{\Delta N/\Delta N}$ and tbp$^{\Delta N/+}$ Cells—Because our study examines a targeted mutation in a fundamental component of the basal transcription machinery *in vivo*, we examined whether the mutation had measurable effects on initiation site fidelity or RNA accumulation. Transcription initiation site fidelity was assessed by primer extension assays on MEFs of all three genotypes. Results indicated that our mutation had no quantitative or qualitative effects on initiation site fidelity of RNAs transcribed by either pol II (β-actin mRNA) or by pol III (U6 RNA)(Fig. 5).

Although TBP, TRF-2, U6, β-actin, DHFR, and cdc25B mRNA levels were unaffected by the ΔN mutation (Figs. 1–5), we wished to determine whether there were effects on a broader range of mRNAs. Two additional highly abundant housekeeping mRNAs from genes with TATA-containing promoters, encoding ribosomal protein L7 and GAPDH, and three low abundance housekeeping mRNAs from TATA-less promoters, encoding TFIIB, Oct-1, and NF-Yb, were chosen for individual analysis (45–47). Expression of all of these mRNAs were unaffected by the TBP mutation (Fig. 6*A*). As an unbiased representation of global gene expression, we used SAGE to estimate the relative abundance of all mRNAs in a sample based on the frequency that unique 3′-end sequence tags are present in a cDNA library made from that mRNA sample (26). SAGE libraries were created from +/+ and ΔN/ΔN MEFs and we sequenced 6890 tags, giving resolution of all mRNAs present at ≥1.5 parts per thousand in each sample (77% of all of the mRNA in the cells; Fig. 6*B*). The SAGE data were externally corroborated because two of the tags, nos. 21 and 22, represented mRNAs that we had independently assayed by quantitative RNase protection and primer extension assays (β-actin, Fig. 3*B*, and GAPDH, Fig. 6*A*, respectively). None of the abundant tags showed a strong effect of the mutation (≥10-fold difference between cell types, *dark shaded zones* in Fig. 6*B*) and only three differed by 3-fold or more (*light shaded zones*). These three mRNAs may have been differentially regulated as a direct consequence of our mutation; however, they may also have simply represented stochastic differences between the cell lines that were independent of the TBP mutation. Further study will be required to distinguish these possibilities. Importantly, the data indicated that overall gene expression was not different in the two cell lines. We conclude that the TBP-ΔN mutation does not affect basal gene expression, and we infer that any gene expression defects arising from this mutation must be highly cell type- and/or gene-specific.

DISCUSSION

TBP is required by all eukaryotes for gene expression as well as for production of ribosomal RNAs and small RNAs. Consistent with this, the core of this protein is one of the most highly conserved proteins known. However, outside of this core, on the N-terminal side of the protein, different phylogenetic lineages exhibit vastly different embellishments (3). To test the function of the vertebrate-specific TBP N terminus, we developed a line of mice in which these sequences had been largely deleted (11). In this article, we show that this mutation had no measurable effects on basal functions in cultured primary fibroblasts.

Previous biochemical and cellular studies on the function of the vertebrate-specific TBP N terminus have shown that this polypeptide domain has strong effects on DNA binding by TBP (15), on how TBP interacts with SNAPc and with certain transcription factors (13, 14, 48–50), and on the function of TBP in transcribing genes encoding RNA components of the splicing machinery (14, 50). In addition, one study has shown that the TBP N terminus can affect rates of cell proliferation (16). By contrast, we show that MEFs lacking most of the TBP N terminus show no general defects in transcription, splicing, gene expression, or proliferation. Results presented here, indicating that removing most of the TBP N terminus has no basal physiological effects on fibroblasts, are corroborated by the observation that, even though few *tbp*$^{\Delta N/\Delta N}$ mice survive to weaning, survivors are healthy and fertile (11). Moreover, physiological processes that depend on proliferation are unaffected in survivors. For example, homozygous mutant fetuses grow and develop at normal rates in immune-compromised mothers (11), and homozygous mutant adults are normal size, they exhibit normal peripheral blood leukocyte counts, normal hair growth, and normal male fertility even after over a year of continuous mating (data not shown). What might be the reasons that our *in vivo* results on TBP-ΔN mutant mice do not show defects in any of the processes in which this domain has been implicated by *in vitro* studies?

One interesting property of living systems is that they are robust. In other words, they generally tolerate or rebound from perturbations of the system. Genetic studies and modeling experiments indicate that robustness can be accomplished by having parallel pathways that can independently accomplish the same task (51, 52). Conversely, fragments of biological systems, such as purified *in vitro* reactions, are generally optimized for a single reaction, and are relatively intolerant of perturbations. One might consider the possibility that the N terminus of TBP functions in a robust system. As examples, transcription factor Oct-1 may stabilize SNAPc on snRNA promoters (50) in the absence of the TBP N terminus, and the ability of the N terminus to affect DNA binding by the TBP$_{CORE}$ (15) may be partially redundant with activities of other proteins assembling at promoters. Elimination of the N terminus

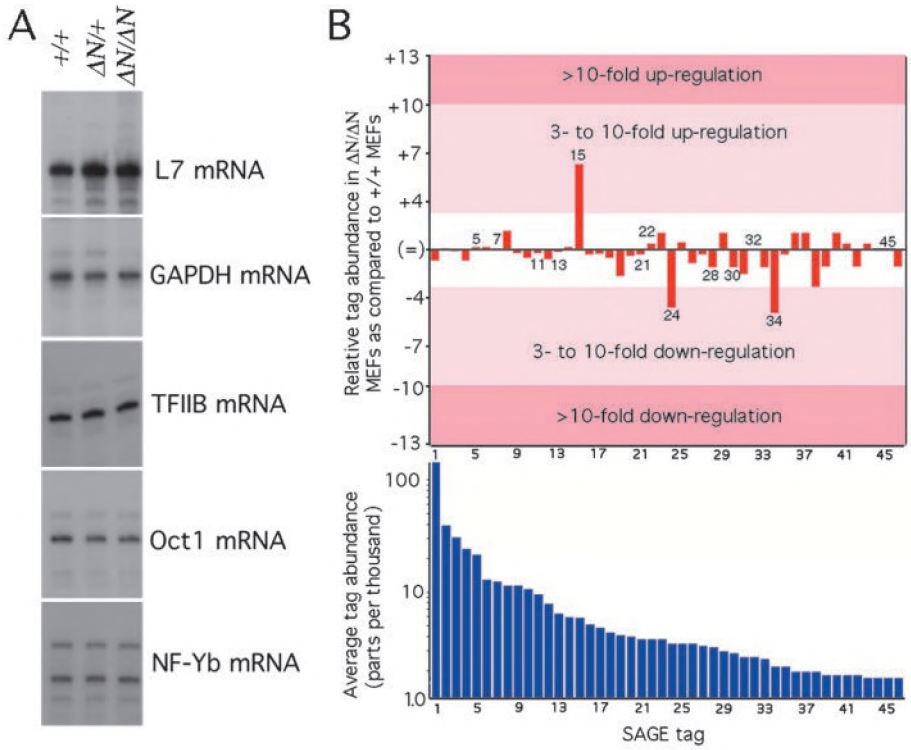

FIG. 6. **Global gene expression.** *A*, expression of selected housekeeping genes. RNase protection assays for the mRNAs indicated at the *right* were performed on total RNA from MEFs of the indicated genotypes as described in previous figures. *B*, expression of abundant mRNAs. SAGE was performed on libraries from +/+ and ΔN/ΔN MEFs at a resolution of 1.5 parts per 1000. Data are presented for the most abundant 46 tags. The *top panel* shows the relative expression profiles presented as fold-difference between the two cell lines, where up-regulation (positive values) are tags that are more abundant in the mutant than in wild-type MEFs, and down-regulation (negative values) are less abundant in the mutants. The *bottom panel* shows the abundance of each tag in both libraries combined, presented as parts per thousand. After exclusion of redundant di-tags, summation of all iterations of the 46 most abundant tags indicates that, of the 6890 tags in this analysis, 5872 (77%) were one of these 46. Tag frequencies for representative mRNAs, indicated by their numerical designation on the *bar graphs*, are as follows: tag 5, ribosomal S28 subunit; tag 7, ribosomal L23 subunit; tag 11, ferritin light chain mRNA; tag 13, ribosomal L36 subunit; tag 15, EST Mm.108736; tag 21, β-actin; tag 22, GAPDH; tag 24, EST Mm.154997; tag 28, ribosomal L28a subunit; tag 30, eTEF1α; tag 32, HMG box-2 protein mRNA; tag 34, EST Mm.194288; tag 45, cytochrome *c*.

may be compensated for by the other back-up proteins. One might envisage a model where the transcriptional assembly at a promoter forms a three-dimensional jigsaw puzzle that does not necessarily fall apart or lose activity as a result of removing this single piece. By such a model, cells lacking the TBP N terminus may be physiologically normal, but might also be more prone to failure in the event that the back-up system fails.

The previous study showing a role for the TBP N terminus in cell proliferation (16) is harder to reconcile with our data because that study, like ours, was performed in living cells. Differences in experimental design are likely responsible for the different results. For example, whereas both studies required several generations of *in vitro* culture to generate a line of cells bearing the respective targeted mutations, cells in our study were rejuvenated by turning them into viable mice, breeding the mutant animals, and then harvesting our experimental cells as primary cultures from the resultant fetuses. Moreover, all modified *tbp* alleles in our study were within the normal *tbp* locus, and none of our cells contained selectable marker genes or vector sequences that might have affected mutant cells (11). Conversely, event though the *tbp-null* allele in the study by Um *et al.* (16) was targeted into the *tbp* locus, it retained an expressed hygromycin resistance gene, and the rescuing expression cassettes were random insertions that brought with them vector sequences, including an expressed puromycin resistance gene. Our data strongly suggest that the N terminus of TBP can be deleted without measurably altering cell proliferation.

Vertebrate and non-vertebrate metazoans universally share the basal properties that have been attributed to the vertebrate TBP N terminus in previous studies (see above). Thus, if this domain indeed participates in such fundamental processes *in vivo*, they are likely accomplished equally well by the N termini of nematodes, dipterans, lepidopterans, and echinoderms, even though the N terminus in each group shows neither similarity to the vertebrate N terminus nor to each other (NCBI/GenBank™). The natural diversity in TBP N-terminal sequences between metazoan clades suggests that, in terms of accomplishing these general functions, this domain should exhibit large amino acid sequence plasticity. Conversely, we have found that natural selection has been particularly intolerant of mutations that would modify the amino acid sequence of TBP N terminus in vertebrates.[2] Thus, we suspected that, whether or not this domain participates in any fundamental cellular functions, the vertebrate TBP N terminus must have a truly vertebrate-specific activity, and it is this function that has been so highly conserved by natural selection in vertebrates. In designing the TBP-ΔN mutation, we hypothesized that this function would be involved in a specific gene regulation process. Consistent with this, we have found that mice homozygous for this mutation suffer from a highly specific defect that compromises a β2-microglobulin-dependent component of the interaction between the placenta of the developing fetus and the mother's immune system (11).

In conclusion, we report that removing the TBP N terminus is inconsequential for general physiological properties of mouse fibroblasts. As such, it is unlikely that this mutation, alone, compromises any vital basal cellular functions *in vivo*, leading us to dismiss the possibility that the mutation results in sub-

lethally sick cells. Rather, we posit that all activities of this polypeptide domain likely fall within two classes. Thus, all TBP N terminus-dependent functions must be either functionally redundant components of robust systems or they will be functionally restricted activities that are not manifested in primary embryonic fibroblast cultures.

Acknowledgments—We thank K. Thomas, C. Lenz, D. Taylor, J. Prigge, T. Frerck, N. Hobbs, L. Eng, K. Lustig, N. Meisner, T. Larson, D. Schwartzenberger, T. Tucker, A. Sealey, J. Kundert, and the members of the Schmidt and Capecchi laboratories for suggestions and contributions.

REFERENCES

1. Hausner, W., and Thomm, M. (2001) *J. Bacteriol.* **183,** 3025–3031
2. Takada, S., Lis, J. T., Zhou, S., and Tjian, R. (2000) *Cell* **101,** 459–469
3. Hernandez, N. (1993) *Genes Dev.* **7,** 1291–1308
4. Littlefield, O., Korkhin, Y., and Sigler, P. B. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 13668–13673
5. Thomm, M. (1996) *FEMS Microbiol. Rev.* **18,** 159–171
6. Qureshi, S. A., Bell, S. D., and Jackson, S. P. (1997) *EMBO J.* **16,** 2927–2936
7. Hashimoto, S., Fujita, H., Hasegawa, S., Roeder, R. G., and Horikoshi, M. (1992) *Nucleic Acids Res.* **20,** 3788
8. Nakashima, K., Nobuhisa, I., Deshimaru, M., Ogawa, T., Shimohigashi, Y., Fukumaki, Y., Hattori, M., Sakaki, Y., Hattori, S., and Ohno, M. (1995) *Gene (Amst.)* **152,** 209–213
9. Shimada, M., Ohbayashi, T., Ishida, M., Nakadai, T., Makino, Y., Aoki, T., Kawata, T., Suzuki, T., Matsuda, Y., and Tamura, T. (1999) *Nucleic Acids Res.* **27,** 3146–3152
10. Hoshiyama, D., Kuma, K., and Miyata, T. (2001) *Gene (Amst.)* **280,** 169–173
11. Hobbs, N. K., Bondareva, A. A., Barnett, S., Capecchi, M. R., and Schmidt, E. E. (2002) *Cell* **110,** 43–54
12. Lobo, S. M., Tanaka, M., Sullivan, M. L., and Hernandez, N. (1992) *Cell* **71,** 1029–1040
13. Henry, R. W., Sadowski, C. L., Kobayashi, R., and Hernandez, N. (1995) *Nature* **374,** 653–656
14. Mittal, V., and Hernandez, N. (1997) *Science* **275,** 1136–1140
15. Zhao, X., and Herr, W. (2002) *Cell* **108,** 615–627
16. Um, M., Yamauchi, J., Kato, S., and Manley, J. L. (2001) *Mol. Cell. Biol.* **21,** 2435–2448
17. Laemmli, U. K. (1970) *Nature* **227,** 680–685
18. Schmidt, E. E., and Merrill, G. F. (1989) *J. Biol. Chem.* **264,** 21247–21256
19. Schmidt, E. E., and Schibler, U. (1995) *J. Cell Biol.* **128,** 467–483
20. Ohbayashi, T., Kishimoto, T., Makino, Y., Shimada, M., Nakadai, T., Aoki, T., Kawata, T., Niwa, S., and Tamura, T. (1999) *Biochem. Biophys. Res. Commun.* **255,** 137–142
21. Das, G., Henning, D., Wright, D., and Reddy, R. (1988) *EMBO J.* **7,** 503–512
22. Tokunaga, K., Taniguchi, H., Yoda, K., Shimizu, M., and Sakiyama, S. (1986) *Nucleic Acids Res.* **14,** 2829
23. Nudel, U., Katcoff, D., Zakut, R., Shani, M., Carmon, Y., Finer, M., Czosnek, H., Ginsburg, I., and Yaffe, D. (1982) *Proc. Natl. Acad. Sci. U. S. A.* **79,** 2763–2767
24. Tamura, T., Sumita, K., Fujino, I., Aoyama, A., Horikoshi, M., Hoffmann, A., Roeder, R. G., Muramatsu, M., and Mikoshiba, K. (1991) *Nucleic Acids Res.* **19,** 3861–3865
25. Schmidt, E. E., and Schibler, U. (1997) *Dev. Biol.* **184,** 138–149
26. Velculescu, V. E., Zhang, L., Vogelstein, B., and Kinzler, K. W. (1995) *Science* **270,** 484–487
27. Schmidt, E. E., Ohbayashi, T., Makino, Y., Tamura, T., and Schibler, U. (1997) *J. Biol. Chem.* **272,** 5326–5334
28. Ohbayashi, T., Schmidt, E. E., Makino, Y., Kishimoto, T., Nabeshima, Y., Muramatsu, M., and Tamura, T. (1996) *Biochem. Biophys. Res. Commun.* **225,** 275–280
29. Varshavsky, A. (1997) *Genes Cells* **2,** 13–28
30. Ohbayashi, T., Makino, Y., and Tamura, T. (1999) *Nucleic Acids Res.* **27,** 750–755
31. Rabenstein, M. D., Zhou, S., Lis, J. T., and Tjian, R. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 4791–4796
32. Teichmann, M., Wang, Z., Martinez, E., Tjernberg, A., Zhang, D., Vollmer, F., Chait, B. T., and Roeder, R. G. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 13720–13725
33. Venter, J. C., Adams, M. D., Myers, E. W. *et al.* (2001) *Science* **291,** 1304–1351
34. Lander, E. S., Linton, L. M., Birren, B. *et al.* (2001) *Nature* **409,** 860–921
35. Reddy, R., Henning, D., Das, G., Harless, M., and Wright, D. (1987) *J. Biol. Chem.* **262,** 75–81
36. Sharp, P. A. (1988) *J. Am. Med. Assoc.* **260,** 3035–3041
37. Kramer, A. (1996) *Annu. Rev. Biochem.* **65,** 367–409
38. Reed, R. (2000) *Curr. Opin Cell Biol.* **12,** 340–345
39. Will, C. L., and Luhrmann, R. (2001) *Curr. Opin. Cell Biol.* **13,** 290–301
40. Mital, R., Kobayashi, R., and Hernandez, N. (1996) *Mol. Cell. Biol.* **16,** 7031–7042
41. Kaufman, R. J., and Sharp, P. A. (1983) *Mol. Cell. Biol.* **3,** 1598–1608
42. Farnham, P. J., and Schimke, R. T. (1986) *Mol. Cell. Biol.* **6,** 365–371
43. Hernandez, S., Hernandez, L., Bea, S., Cazorla, M., Fernandez, P. L., Nadal, A., Muntane, J., Mallofre, C., Montserrat, E., Cardesa, A., and Campo, E. (1998) *Cancer Res.* **58,** 1762–1767
44. Korner, K., Jerome, V., Schmidt, T., and Muller, R. (2001) *J. Biol. Chem.* **276,** 9662–9669
45. Li, X. Y., Hooft van Huijsduijnen, R., Mantovani, R., Benoist, C., and Mathis, D. (1992) *J. Biol. Chem.* **267,** 8984–8990
46. Suzuki, N., Peter, W., Ciesiolka, T., Gruss, P., and Scholer, H. R. (1993) *Nucleic Acids Res.* **21,** 245–252
47. Tsuboi, A., Conger, K., Garrett, K. P., Conaway, R. C., Conaway, J. W., and Arai, N. (1992) *Nucleic Acids Res.* **20,** 3250
48. Tansey, W. P., Ruppert, S., Tjian, R., and Herr, W. (1994) *Genes Dev.* **8,** 2756–2769
49. Tansey, W. P., and Herr, W. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 10550–10554
50. Mittal, V., Ma, B., and Hernandez, N. (1999) *Genes Dev.* **13,** 1807–1821
51. von Dassow, G., Meir, E., Munro, E. M., and Odell, G. M. (2000) *Nature* **406,** 188–192
52. Meir, E., von Dassow, G., Munro, E., and Odell, G. M. (2002) *Curr. Biol.* **12,** 778–786