750–755 *Nucleic Acids Research, 1999, Vol. 27, No. 3*

Identification of a mouse TBP-like protein (TLP) distantly related to the *Drosophila* TBP-related factor

Tetsuya Ohbayashi, Yasutaka Makino and Taka-aki Tamura*

Department of Biology, Faculty of Science, Chiba University and CREST Japan Science and Technology Corporation, 1-33 Yayoi-cho, Inage-ku, Chiba 263-8522, Japan

Received October 15, 1998; Revised and Accepted November 30, 1998

DDBJ/EMBL/GenBank accession no. AB017697

ABSTRACT

TATA-binding protein (TBP) is an essential factor for eukaryotic transcription. In this study, we demonstrated a mouse cDNA encoding a 21 kDa TBP-like protein (TLP). The TLP ORF, carrying 186 amino acids, covered the entire 180 amino acids of the C-terminal conserved domain of mouse TBP with 39% identity and 76% similarity. Northern blot analysis demonstrated that TLP mRNAs were expressed in various mammalian tissues ubiquitously and that their distribution pattern was analogous to that of TBP. By using anti-TLP antibody, we demonstrated the existence of TLP proteins in various mammalian cells and tissues. The *Drosophila* TBP-related factor (TRF) is a neurogenesis-related transcription factor that binds to the TATA-box and activates transcription. TLP did not bind to the TATA-box nor direct transcription initiation. Multiple amino acids critical for TBP function were deleted or substituted in TLP, while amino acids in *Drosophila* TRF much resembled those in TBP. Similarity between *Drosophila* TRF and mouse TLP was considerably lower (alignment score 35) than that between *Drosophila* TBP and mouse TBP (alignment score 88). Identity of nucleotide sequences between mouse and putative human TLPs (94%) was higher than that between TBPs (91%) in these two animals. Expression of TLP was nearly constant throughout the P19 differentiation process. Accordingly, we suggest that, even if higher eukaryotes generally contain multiple *tbp*-related genes, TLP is not a bona fide mammalian counterpart of *Drosophila* TRF.

INTRODUCTION

The TATA-binding protein (TBP) is one of the general transcription factors (GTFs) necessary for transcription initiation of eukaryotic genes. TBP is included in multiprotein complex GTFs such as SL1 (for rRNA genes), TFIID (for protein coding genes) and TFIIIB (tRNA, 5S rRNA and other small RNA genes) together with TBP-associated factors (TAFs) and is involved in all transcription systems (1,2). Accordingly, TBP is regarded as a universal transcription factor. In the case of RNA polymerase II (RNAPII)-driven genes, TBP or TFIID (plus TFIIA) binds to the TATA-box, then gathers other GTFs and RNAPII to form a functional preinitiation complex that is ready to initiate transcription (1). Moreover, TBP is capable of binding to a variety of cellular and viral proteins that modulate (activate or repress) transcription from a target gene (4–6). Accordingly, TBP is obviously a key molecule in gene regulation through communications with transcriptional regulators and basal transcription machineries. Hence, studies on TBP itself and its related factors are significant to elucidate the molecular mechanism of gene regulation.

TBP has been isolated from many kinds of eukaryotes and archaea (4,7–12). A polypeptide of TBP can be divided into two domains; an N-terminal species-specific domain and a C-terminal conserved domain (CCD). The CCD involves ~180 amino acids and is extraordinarily conserved among organisms (1,13). For example, CCD structures of yeast and human TBPs are 80% identical. The CCD is essential for TBP function, i.e. to bind to the TATA-box and to activate transcription from TATA-containing promoters. X-ray refraction studies (14,15) and mutation analyses (15,16) demonstrated that a strictly conserved ternary structure of the CCD consisting of multiple antiparallel β-sheets is critical for function. Thus, most point mutations and any short deletion within the CCD result in loss of most of the TBP function (16).

Initially, the eukaryotic cell was thought to contain only one kind of TBP. Indeed, the yeast (*Saccharomyces cerevisiae*) haploid genome contains a single *tbp* gene (17). In 1993, Crowley *et al*. (18) demonstrated one TBP-resembling protein in *Drosophila* and referred to it as TBP-related factor (TRF). TRF was originally isolated from the *shaker* mutant fly, which has a defect in its nervous function. Thereafter, TRF was demonstrated to have a function like TBP and to be involved in nervous system-specific gene expression (19). Since then, several investigators have been trying to identify a TBP-resembling protein in other higher eukaryotes.

In this communication, we describe the isolation of a mouse cDNA of a TBP-like protein designated TLP. We examined expression patterns and biochemical features of TLP and compared its structure with that of other TBP-resembling proteins. Eventually, we concluded that TLP is not a bona fide homolog of transcriptionally competent *Drosophila* TRF. We also discuss the copy number of *tbp*-related genes therein.

*To whom correspondence should be addressed. Tel: +81 43 290 2823; Fax: +81 43 290 2824; Email: btamura@nature.s.chiba-u.ac.jp

MATERIALS AND METHODS

Cloning of TLP cDNA

An EST database for mouse cDNAs was searched and two clones (AA79823 and W89738 corresponding to nucleotide positions 57–558 and 49–479 in TLP, respectively; DDBJ/EMBL/GenBank accession no. AB017697) were chosen as having a significant homology with mouse TBP. Two primers (5′-ATCTTCATCCTT-GTCCTCCAGCTTC and 5′-GCTCCCTCCAAAGCAATCTT-CCTTA) were constructed based on the EST clones and used for RT–PCR of mouse brain RNAs. The amplified DNA contained the TLP partial sequence from nucleotide 135 to 379. Using this DNA as a probe, we screened a mouse brain RNA-derived cDNA library in λZAPII and obtained one long cDNA (designated TLP). Sequence alignment was performed by use of the CLUSTAL W program (20).

Overexpression, purification and antibody for the recombinant TLP

The entire ORF linked with FLAG and oligo(histidine) tags at the N-terminus was inserted into an *Escherichia coli* expression vector (pET3a) (21). Recombinant TLP was expressed in *E.coli* BL21(DE3)*pLysS*. (21) Protein was induced by treatment of the bacterial culture with IPTG. Soluble recombinant protein was purified with Ni^{2+}–agarose (Qiagen) and anti-FLAG M2 affinity gel (Kodak) as recommended by the suppliers. Mouse TBP carrying a histidine tag was expressed and purified to near homogeneity as previously described (22,23). Rabbits were immunized with the purified TLP and polyclonal antibody was generated. The antibody was affinity purified by the use of TLP-immobilized HiTrap NHS-activated beads (Pharmacia).

Northern blotting

Membranes for multiple tissue northern blots on which various tissue-derived rat poly(A)$^+$ RNA had been blotted were purchased from Clontech. Each lane included 2 μg of poly(A)$^+$ RNA. We confirmed equal amounts of blotted RNAs by using a β-actin probe (data not shown). Probe DNA included the TLP sequence spanning nucleotides 135–379 (DDBJ/EMBL/GenBank accession no. AB017697). The Quick Hyb hybridization solution (Stratagene) contained 1×10^6 c.p.m./ml cDNA probe. Hybridization was carried out at 65°C for 1 h.

Extracts and western blotting

Cell extracts were prepared by the method described by Schreiber *et al*. (24). Rat liver nuclear extracts were prepared as previously described (25). The extracts were dialyzed against a buffer consisting of 25 mM HEPES–KOH (pH 7.9), 0.1 M KCl, 5 mM $MgCl_2$, 0.5 mM DTT, 0.1% NP-40, 10% glycerol. For western blotting, proteins were separated by 15% SDS–PAGE and electrophoretically transferred to PVDF membranes (Millipore). Proteins on the membrane were detected by the ECL protocol (Amersham).

Gel shift assay

Gel shift assay was performed as previously described (22) with a synthetic 23mer oligonucleotide probe containing the adenovirus major late promoter (AdML) TATA-box (5′-AGGGGGGCTA-TAAAAGGGGGTGG). In the binding reaction (20 μl), 1 ng end-labeled probe (labeled at 5′-end by kination) was incubated with 6 ng of each protein sample for 30 min at room temperature. Electrophoresis was conducted in a 5% polyacrylamide gel in a buffer consisting of 25 mM Tris-base (pH 8.3), 190 mM glycine and 5% glycerol, as previously reported (26). Specificity of the shifted bands was confirmed by use of wild-type and mutant competitor DNAs (data not shown).

Reconstituted *in vitro* transcription

Factors. The human TFIIB and TFIIE (27–29) expressed in *E.coli* as oligo(histidine)-tagged proteins were purified as described previously. The α/β and γ subunits of the human TFIIA were independently expressed in *E.coli* (30) and added into the reaction mixture. Human TFIIF was purified from Sf9 cells co-infected with recombinant baculoviruses encoding its RAP74 and RAP30 subunits (31). RNAPII was purified from calf thymus by the procedure of Hodo and Blatti (32).

In vitro transcription. Reconstituted *in vitro* transcription reaction was performed as previously described (33) with supercoiled DNA templates of the AdML and adenovirus E4 promoters followed by a G-free cassette (33,34). The final volume of one reaction was 20 μl in TX buffer [10 mM HEPES–KOH (pH 7.6), 25 mM KCl, 6 mM $MgCl_2$, 3% glycerol]. The protein concentration in each reaction was 45 ng for TBP, 50 ng for TFIIB, 120 ng for TFIIF and 0.2 μg for RNAPII. Supercoiled DNA template (400 ng) was used for each reaction. Transcription was performed for 45 min at 30°C after preincubation for 20 min at room temperature. Transcripts were resolved on a 5% sequencing gel.

RESULTS

Isolation of mouse TLP cDNA

To detect a gene resembling mouse *tbp*, we searched an EST database of mouse cDNA clones and found several possible clones. Among them, we selected two EST clones, submitted as AA79823 and W89738. Based on sequence data of these clones, we designed PCR primers and performed RT–PCR using mouse brain RNAs. We were able to amplify one DNA. By using this DNA as a probe, we finally obtained one long cDNA with 1293 nucleotides from a mouse brain cDNA library (DDBJ/EMBL/GenBank accession no. AB017697 in). One long ORF, which encoded a protein of 186 amino acids, was identified, whose calculated molecular mass was 20.8 kDa. Thus, we designated this cDNA as TLP (TBP-like protein) according to the various observations presented below. We aligned the amino acid sequences of TLP and mouse TBP (Fig. 1) and found that TLP exhibited a significant sequence similarity (76%) with the 180 amino acids of the CCD of mouse TBP, which is required and may be sufficient for TBP function. TLP covered the entire CCD when aligned with TBP and only two amino acids were deleted. From this significant structural similarity with the CCD, we denoted TLP as a TBP-resembling protein of mice.

Gene expression of TLP in mouse tissues

We examined the gene expression of TLP by northern blotting using RNAs from mouse, rat and human testes (Fig. 2A), since testes contained large amounts of TLP transcripts (below). RNAs from testes exhibited a prominent discrete signal with the same

```
TLP   1 MDADSDVALDILITNVVCVFRTRCHLNLRK
TBP 136 ------SGIVPQLQNIVSTVNLGCKLDLKT
              .:   : *:*....  *:*:*:.

TLP  31 IALEGANVIYK-RDVGKVLMKLRKPRITAT
TBP 160 IALRARNAEYNPKRFAAVIMRIREPRTTAL
        ***.. *. *: : .. *:*::*:** **

TLP  60 IWSSGKIICTGATSEEEAKFGARRLARSLQ
TBP 190 IFSSGKMVCTGAKSEEQSRLAARKYARVVQ
        *   **::****.***:::::**: ** :*

TLP  90 KLGFQVIFTDFKVVNVLAVCNMPFEIRLPE
TBP 220 KLGFPAKFLDFKIQNMVGSCDVKFPIRLEG
        **** . * ***: *::. *:: * ***

TLP 120 FTKNNRPHASYEPELHPAVCYRIKSLRATL
TBP 250 LVLTHQQFSSYEPELFPGLIYRMIKPRIVL
        :. .:: .:******.*.: **: . * .*

TLP 150 QIFSTGSITVTGPNVKA-VATAVEQIYPFV
TBP 280 LIFVSGKVVLTGAKVRAEIYEAFENIYPIL
         ** :*.:.:**.:*:* :  *.*:***::

TLP 179 FESRKEIL
TBP 310 KGFRKTT-
          **
```

Identity:39%
Similarity:76%

Figure 1. Amino acid sequences of mouse TLP and comparison between mouse TLP and TBP. Nucleotide sequence of mouse TLP has been submitted to the DDBJ, EMBL and GenBank databases with accession no. AB017697. The C-terminal conserved 180 amino acids of the mouse TBP from Ser136 were aligned with TLP by the CLUSTAL W program. The identical and similar amino acid residues between the two proteins are denoted by asterisks and dots, respectively. Single dot, low similarity; double dot, high similarity.

size (1.5 kb). The size of the northern signal (1.5 kb) roughly agreed with the length of the obtained cDNA (1.3 kb), if we consider the short poly(A) tail (Fig. 1). We thus conclude that mouse cells express TLP mRNA and that mammalian cells generally express structurally homologous transcripts.

We analyzed RNAs from various rat tissues and found that TLP mRNAs were expressed in all tissues tested (Fig. 2B), showing that *tlp* is a ubiquitous gene. Testes contained large amounts of the mRNAs (Fig. 2B, lane 8), whereas spleen, lung and liver contained only small amounts of them. The brain contained a moderate level of the mRNA (Fig. 2B, lane 2). Some tissues yielded non-specific signals with higher sizes (Fig. 2A, lane 1, and B). These bands may reflect intermediates of TLP mRNAs or cross-hybridization with unrelated transcripts. We further examined the gene expression of the mouse TBP (Fig. 2C). TBP mRNA, whose size (2.2 kb) is different from that of TLP, was ubiquitously distributed in all tested tissues, as expected. Interestingly, the distribution pattern of mouse TBP mRNAs was quite similar to that of TLP.

Identification of TLP protein

To investigate whether the *tlp* gene is expressed as a protein, we carried out western blotting for various mammalian tissues and cells using specific antibody. Anti-TLP antibody was generated in rabbits by immunization with purified TLP and antigen purified as described in Materials and Methods. Specificity of anti-TLP antibody for TLP and TBP was clearly demonstrated, because the antibody that recognized TLP did not detect even a 60-fold excess amount of mouse TBP (Fig. 3A). In SDS–PAGE, TLP was detected as a protein with an apparent molecular mass of 26 kDa. The anti-TLP antibody detected a protein with an apparent molecular size of 24 kDa in mouse NIH 3T3 cell (Fig. 3B, lane 1). According to the band of recombinant TLP (r-TLP, Fig. 3B, lane 7) carrying an N-terminal tag, this 24 kDa band was suggested to represent native TLP. The TLP-related bands appeared in double. It may be attributable to post-translational modification since the ratio of these two bands varied in each sample (Fig. 3B and C). We found considerable amounts of a band with an apparent molecular mass of 32 kDa (Fig. 3B and C, indicated by asterisks). This upper band may represent a non-specific protein having an epitope in common with TLP. The native TLP was observed as doublet in some lanes (Fig. 3B and C). These phenomena may reflect degradation or modification of the native TLP.

TLP was found both in the nucleus and cytoplasm of mouse cells (Fig. 3B, lanes 2 and 3, respectively). We further detected a TLP equivalent in rat brain whole cell extract and HeLa cell nuclear extracts (Fig. 3B, lanes 4 and 5, respectively). Highly purified nuclei did not contain the non-specific 32 kDa band (Fig. 3B, lane 6). We thus suggest that TLP is immunologically and structurally conserved in mammals. When we measured TLP in various rat tissues (Fig. 3C), we found it to be commonly present in various tissues. Eventually, we concluded that TLP indeed exists in mouse cells and is expressed in many mammalian tissues ubiquitously.

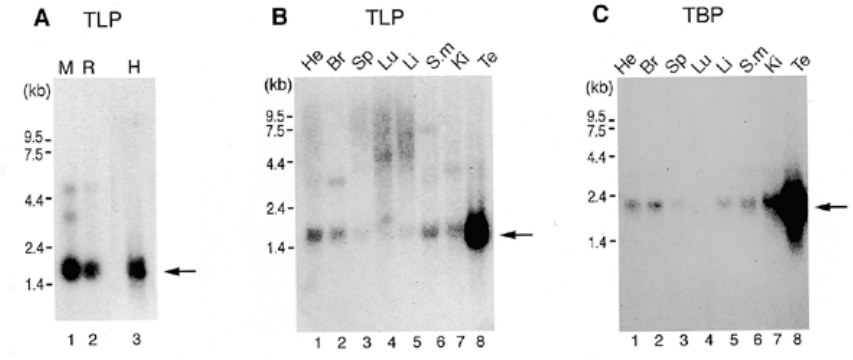

Figure 2. Expression of mRNAs of TLP and TBP in animal tissues. Gene expression of TLP and of TBP was determined by northern blotting using the TLP probe (from 135 to 379) and mouse TBP probe (from 466 to 1081), respectively. Arrows indicate specific signals. (**A**) Gene expression of TLP in various mammal testes. One microgram of poly(A)$^+$ RNA was analyzed. M, mouse; R, rat; H, human. Gene expression of TLP (**B**) and of TBP (**C**). Various kinds of rat RNAs on multiple tissue blots (Materials and Methods) were analyzed. He, heart; Br, brain; Sp, spleen; Lu, lung; S.m, skeletal muscle; Ki, kidney; Te, testis.

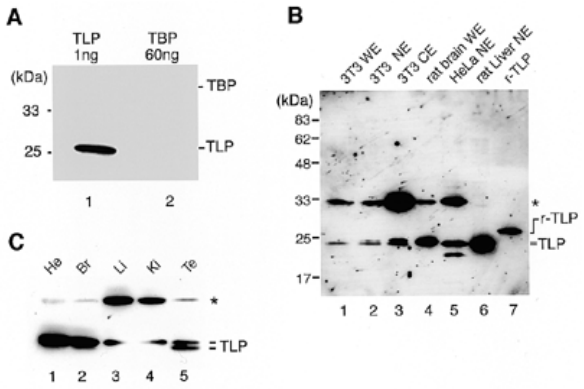

Figure 3. Detection of TLP protein by immunoblotting. Asterisks indicate unknown non-specific signal. (**A**) Specificity of the prepared anti-TLP antibody. One nanogram of recombinant FLAG-histidine-tagged TLP (lane 1) and 60 ng of mouse TBP (lane 2) were analyzed with the anti-TLP antibody. (**B**) TLP proteins in various mammalian cells and tissues. Whole cell extract (WE, 8 µg, lane 1), nuclear extract (NE, 5 µg, lane 2) and cytoplasmic extract (CE, 5 µg, lane 3) of mouse NIH 3T3 cells were used. Lane 4, WE of rat brain (5 µg); lane 5, NE of HeLa cells (20 µg); lane 6, NE of rat liver (20 µg); lane 7, 1 ng of the recombinant tagged TLP (r-TLP). Extracts were prepared by the method of Shreiber *et al*. (24) except for the rat liver NE (25). (**C**) TLP in rat tissues. Whole cell extracts (10 µg) from various rat tissues were analyzed. Lane 1, heart; lane 2, brain; lane 3, liver; lane 4, kidney; lane 5, testis. Abbreviations as in Figure 2.

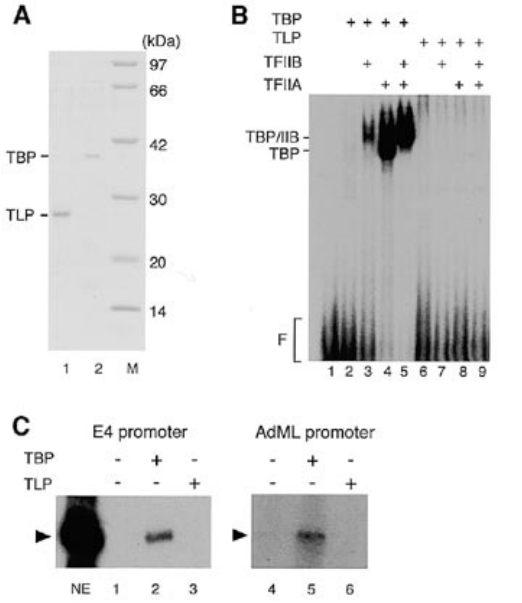

Figure 4. Function analysis of the mouse TLP. (**A**) Purity of the recombinant proteins used in this study. FLAG-histidine-tagged TLP (200 ng, lane 1) and oligo(histidine)-tagged mouse TBP (100 ng, lane 2) expressed in *E.coli* and purified on Ni–agarose were resolved by SDS–PAGE and stained with CBB. (**B**) Gel shift assay for TLP and TBP. TBP (10 ng), TLP (20 ng), TFIIA (10 ng) and TFIIB (10 ng) were used for each reaction as indicated together with the AdML TATA-box probe (1 ng). Positions of the free probe (F) and TBP- and TBP/TFIIB-specific shifted bands are shown. (**C**) Effects of TLP on transcriptional activation. The reconstitution transcription was performed as described in Materials and Methods using adenovirus E4 (E4) and major late (AdML) promoters. All lanes contained TFIIB, TFIIF and RNAPII. TBP (10 ng, lanes 2 and 5) and TLP (20 ng, lanes 3 and 6) were added to the reaction mixture as indicated. NE, rat liver nuclear extract (20 mg) prepared by the method of Tamura *et al*. (25).

Function assay for TLP

Functions of TBP are binding to TATA-box DNA and activation of basal transcription from a TATA-containing promoter. In the case of *Drosophila* TRF, it has essentially an equivalent activity as the authentic TBP (19). For use in biochemical assays, we purified TLP overexpressed in *E.coli* carrying FLAG and histidine tags to near homogeneity, avoiding any denaturing procedures (Materials and Methods; Fig. 4A). The recombinant mouse TBP was also purified. First, we carried out a gel shift assay using AdML TATA-box DNA as a probe (Fig. 4B). TBP stably bound to DNA in conjunction with TFIIA (Fig. 4B, lane 4), though TBP alone exhibited only a weak binding capacity (lane 2). TFIIB generated a supershift band (Fig. 4B, lane 3) and TFIIA increased that shifted band (lane 5). These results are coincident with previous observations. In contrast, TLP did not yield any detectable shift in the band of TATA-box DNA even if TFIIB and/or TFIIA were further added (Fig. 4B, lanes 6–9). These results suggest that TLP does not have an intrinsic DNA-binding ability like TBP and TRF.

Next, we carried out *in vitro* transcription by the reconstitution system in which recombinant TBP, TFIIB, TFIIF and purified RNAPII were included. E4 and AdML promoters were efficiently transcribed in this cocktail (Fig. 4C, lanes 2 and 5) and omission of TBP completely abolished the transcription (lanes 1 and 4). Hansen *et al*. (19) reported that the E4 promoter was efficiently transcribed by this type of reaction system when they employed *Drosophila* TRF instead of TBP. In this study, however, we could not find any transcriptional activation function of TLP for either promoter (Fig. 4C, lanes 3 and 6). Addition of TFIIA had no effect on the lack of transcriptional activation by TLP (data not shown). We suggest that TLP does not have a potential to stimulate transcription, at least in this reaction system.

DISCUSSION

TBP is one of the essential factors required for eukaryotic transcription initiation and is thus regarded as a key molecule for transcriptional regulation. The eukaryotic haploid genome was initially thought to contain a single *tbp*-related gene (in this communication, we use the term '*tbp*-related genes' to indicate authentic *tbp* plus *tbp*-like genes). However, in 1993, Crowley *et al*. (18) reported the existence of a TBP-related factor (TRF) in *Drosophila*. TRF resembles TBP (alignment scores: human, 60; *Drosophila*, 58) and was shown to bind to the TATA-box of the AdML promoter. Afterward, it was demonstrated to activate transcription of RNAPII genes (19). Since TRF was discovered in the mutant fly *shaker*, which has a defect in its nervous system, and was found to be abundantly expressed in the nervous system, this factor is believed to be responsible for nervous system-selective gene transcription (19).

Mammalian cells indeed contain a protein that resembles TBP

TBP-resembling factors have not been identified in higher animals so far. We initially tried to isolate a mouse counterpart of TRF. An EST database search for mouse cDNAs implied the existence of a *tbp*-resembling gene in mice. Based on the EST database, we finally obtained one *tbp*-like clone (DDBJ/EMBL/

nucleotide \ amino acid	mTLP	hTLP	dTRF	mTBP	hTBP	dTBP
mTLP		**100**	**35**	**39**	**39**	**35**
hTLP	94		**35**	**39**	**39**	**35**
dTRF	17	13		**60**	**60**	**58**
mTBP	45	26	51		**100**	**88**
hTBP	21	40	56	91		**88**
dTBP	35	27	63	72	72	

Figure 5. Comparison of nucleotide and amino acid sequences of TBP-resembling factors of various eukaryotes. Similarity among the C-terminal 180 amino acids of TBP, TRF and TLP of various organisms was assessed by the CLUSTAL W program and calculated aligned scores are presented. The aligned scores for amino acid and nucleotide sequences are shown in bold letters and normal letters, respectively. The nucleotide sequence for human TLP (hTLP) was deduced from multiple human EST cDNA clones (AA448144, AA412574, AA256080, AA298756, AA281228, AI075193 and AA02178). mTLP, mouse TLP; dTRF, *Drosophila* TRF; mTBP, mouse TBP; hTBP, human TBP; dTBP, *Drosophila* TBP.

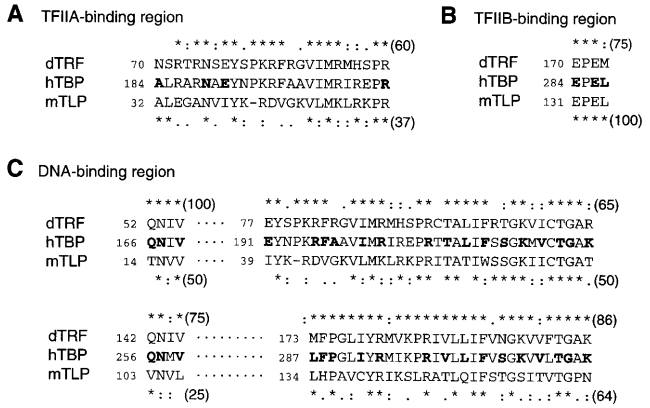

Figure 6. Structural comparison of the critical regions in TBP with TRF and TLP. Critical amino acid residues in human TBP (hTBP) identified by X-ray refraction studies were aligned with corresponding portions of *Drosophila* TRF (dTRF) and mouse TLP (mTLP). The TFIIA-binding region (**A**), TFIIB-binding region (**B**) and DNA-binding region (**C**) were analyzed. Critical amino acid residues in TBP are shown in bold. Identical and similar amino acid residues of TLP and TRF compared with those of TBP are denoted by asterisks and dots, respectively. Alignment scores for each region are presented in parentheses. Numbers presented at the left side of the sequences are amino acid positions in each protein.

GenBank accession no. AB017697). The encoded protein was designated TLP. We did not refer to it as 'mouse TRF' owing to its drastic divergence from TRF and its functional incompetency. Since TLP was highly homologous to mouse TBP (76% similarity) and it covered the entire 180 amino acids of the TBP CCD, we denoted TLP a TBP-resembling protein of mice.

Northern analysis of RNAs of mouse and other mammals detected a single discrete band of 1.5 kb (Fig. 2B) whose size was similar to that of the isolated cDNA (1.3 kb) that carried a short poly(A) tail (DDBJ/EMBL/GenBank accession no. AB017697), suggesting that the band represents TLP mRNA. Western blotting also demonstrated the presence of a polypeptide whose size was coincident with that of the recombinant TLP. This band was also observed in various mammalian cells and tissues (Fig. 3B and C). Moreover, the antibody was shown to be specific (Fig. 3A). Consequently, we suggest that the band represents native TLP protein. We found that TLP mRNA and its translation products were ubiquitously present in mammalian tissues. We detected human cDNAs analogous to mouse *tlp* in an EST database (data not shown). Wieczorek *et al.* (35) described a human TBP-like factor in their report, though they did not show any data for it. That TBP-like factor may be identical to human TLP. Consequently, we conclude that mammalian cells commonly have a TLP homolog and express its mRNA and translation product ubiquitously in many tissues.

Is TLP a mouse homolog of *Drosophila* TRF?

As described above, *Drosophila* has been demonstrated to harbor the *trf* gene as an alternative *tbp*-related gene in addition to the authentic *tbp* gene (19). Although we initially aimed at isolating a mouse *trf* (which was the reason why we screened a brain cDNA library), now we can claim that TLP is not a bona fide TRF homolog of mice, for the several reasons described below. We precisely compared the amino acid and nucleotide sequences of TBP and TBP-resembling proteins of *Drosophila*, mouse and *Homo sapiens* (Fig. 5). The human TLP (hTLP) is a hypothetical protein deduced from multiple human EST clones. We believe that human TLP truly exists because we isolated a partial human cDNA clone whose sequence completely matched the hypothetical hTLP (data not shown). We noticed that mouse TLP (mTLP) and *Drosophila* TRF (dTRF) exhibited an alignment score of only 35, whereas the identity of *Drosophila* TBP (dTBP) and mouse TBP (mTBP) was considerably higher (alignment score 88; Fig. 5). Moreover, nucleotide sequence similarity between hTLP and mTLP (alignment score 94) was higher than that between human TBP (hTBP) and mTBP (alignment score 91; Fig. 5). These data suggest that *tlp* homologs are much more conserved than *tbp* homologs. Consistent with this idea, the nucleotide sequence identity for the untranslated regions of mTLP and hTLP was considerably higher (alignment score 75) than that between mTBP and hTBP (alignment score 33; data not shown). Hence, it is phylogenetically implausible that *tlp* originated from *trf* or its ancestor gene.

To elucidate the structural relationship among TBP, TRF and TLP more clearly, we focused on three critical regions, including the TFIIA- and TFIIB- and DNA-binding regions within the CCD of TBP. TBP sequences were aligned with corresponding regions in TLP and TRF (Fig. 6). The *Drosophila* TRF and mTLP exhibited alignment scores of 60 and 37, respectively, to the TFIIA-binding region of TBP (Fig. 6A). The divergence of TLP from TBP for the DNA-binding region was thus reconfirmed to be higher than that of TRF. Alignment scores for four domains in the DNA-binding region between TRF and TBP were between 100 and 65, whereas those between TLP and TBP were between 64 and 25 (Fig. 6C). We examined sequence similarities between each repeating unit within a direct repeat (data not shown). The identity for the critical amino acid residues between the repeating unit for TRF (67%) was as high as that for TBP (83%). Surprisingly, identity between two repeating units of TRF (37%) was higher than those of TBP (25%). In contrast, the above two indexes for TLP were remarkably low (17% for the whole region and 28% for critical residues). These results are consistent with the facts that TRF binds to the TATA-box (18,19) but TLP does not (Fig. 4B). We conclude that TLP is not a bona fide mammalian homolog of *Drosophila* TRF.

TLP does not have a function shared by TBP and TRF

The above observations indicate that the critical sequences in TBP were highly conserved in TRF but not in TLP. This idea is consistent with the function assays in Figure 4B and C, showing that TLP had no apparent capability for DNA-binding and transcriptional activation. The TLP preparation we employed should not have been inactivated since we carefully prepared it as a soluble protein avoiding any denaturing conditions. Yamamoto *et al*. (16) described that any small deletions within the CCD abolished TBP function. TLP included two internal deletions when aligned with TBP (Fig. 1). One of them was located within the TFIIA-binding region (Fig. 6A) and the other was located within the DNA-binding region (Fig. 6C). Hence, it is mostly likely true that TLP has no intrinsic function such as does TBP. Interestingly, TRF and TLP exhibited weak (75) and perfect (100) alignment scores to the TFIIB-binding region of TBP (Fig. 6B). This fact implies that TLP may bind to TFIIB. From a functional point of view, we suggest that TLP is not a mouse counterpart of *Drosophila* TRF. Even though TLP diverged from an ancestor of TRF, the role of mammalian TLP is probably distinct from that of TRF. At the present time, we do not know the role of TLP. It may function in transcription without DNA-binding or with distinct DNA-binding ability. *Drosophila* TRF is expressed predominantly in the nervous system (18,19). According to the data of Figure 2B, we would emphasize that TLP may not have a major role in neurogenesis or in neural gene expression even if it participates in transcriptional regulation. Consistently, *tlp* gene expression in P19 cells was constant throughout the *in vitro* differentiation process induced by retinoic acid (data not shown).

How many *tbp*-related genes do higher eukaryotes have?

The above assumption brings us to the question of how many kinds of TBP-resembling proteins do the higher eukaryotes have? The genome of *S.cerevisiae* was proved to have only one *tbp* gene per haploid (17). However, the copy number for *tbp*-related genes in multicellular organisms is questionable. To our knowledge, earlier and present data indicate that *Drosophila*, mouse and higher plants contain at least two *tbp*-related genes. Plants (maize and *Arabidopsis*) have a couple of closely resembling *tbp* genes, both of which are functionally competent (11,12). Two kinds of *tbp*(-related) genes were detected in *Drosophila discoideum* (T.Kawata, personal communication). Since a human *tbp* cDNA was cloned in 1990 (10), no third *tbp*-related gene has been identified so far. Perhaps, no other *tbp*-related genes exist in higher eukaryotes. If this is the case, a single *tbp* ancestor gene was duplicated when multicellular organisms were generated. According to this 'two-TBP model', one copy of the *tbp* gene was strictly conserved through evolution, whereas the other copy evolved to *trf* in *Drosophila*, to *tlp* in mammals and diverged little in plants. If a *trf* homolog or a *tlp* homolog can be discovered in mammals or *Drosophila*, respectively, the alternative 'three-TBP model' would become more plausible. A database search for *tbp*-related sequences in the *Caenorhabditis elegans* genome uncovered two *tbp*-like DNA fragments distinct from the authentic *tbp* gene (data not shown).

ACKNOWLEDGEMENTS

The authors thank all colleagues in our laboratory for valuable discussion and encouragement through this study. We also thank Ms K.Osano and M.Kasahara for skillful technical assistance. A part of this study was supported by a grant from the program Grants-in-Aid for Scientific Research on Priority Areas of The Japanese Ministry of Education, Science, Sports and Culture.

REFERENCES

1 Hernandez,N. (1993) *Genes Dev.*, **7**, 1291–1308.
2 Sharp,P.A. (1992) *Cell*, **68**, 819–821.
3 Roeder,R.G. (1998) *Trends Biochem. Sci.*, **21**, 327–335.
4 Inostroza,J.A., Mermelstein,F.H., Ha,I., Lane,W.S. and Reinberg,D. (1992) *Cell*, **70**, 477–489.
5 Seto,E., Seto,E., Usheva,A, Zambetti,G.P., Momand,J., Horikoshi,N., Weinmann,R., Levine,A.J. and Shenk,T. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 12028–12032.
6 Ge,H. and Roeder,R.G. (1994) *Cell*, **78**, 513–523.
7 Horikoshi,M., Wang,C.K., Fujii,H., Cromlish,J.A., Weil,P.A. and Roeder,R.G. (1989) *Nature*, **341**, 299–303.
8 Muhich,M.L., Iida,C.T., Horikoshi,M., Roeder,R.G. and Parker,C.S. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 9148–9152.
9 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.
10 Kao,C.C., Lieberman,P.M., Schmidt,M.C., Zhou,Q., Pei,R. and Berk,A.J. (1990) *Science*, **248**, 1646–1650.
11 Gasch,A., Hoffmann,A., Horikoshi,M., Roeder,R.G. and Chua,N.H., (1990) *Nature*, **360**, 390–394.
12 Vogel,J.M., Roth,B., Cigan,M. and Freeling,M. (1993) *Plant Cell*, **5**, 1627–1638.
13 Hancock,J.M. (1993) *Nucleic Acids Res.*, **21**, 2823–2830.
14 Nikolov,D.V., Chen,H., Halay,E.D., Usheva,A.A., Hisatake,K., Lee,E.D., Roeder,R.G. and Burley,S.K., (1995) *Nature*, **377**, 119–128.
15 Bryant,G.O., Martel,L.S., Burley,S.K. and Berk,A.J., (1996) *Genes Dev.*, **10**, 2491–2504.
16 Yamamoto,T., Horikoshi,M., Wang,J., Hasegawa,S., Weil,P.A. and Roeder,R.G. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 2844–2848.
17 Goffeau,A., Aert,R., Agostini-Carbone,M.L., Ahmed,A., Aigle,M., Alberghina,L., Albermann,K., Albers,M., Aldea,M., Alexandraki,D. *et al.* (1997) *Nature*, **387** (suppl.), 1–105.
18 Crowley,T.E., Hoey,T., Liu,J.-K., Jan,Y.N., Jan,L.Y. and Tjian,R. (1993) *Nature*, **361**, 557–561.
19 Hansen,S.K., Takada,S., Jacobson,R.H., Lis,J.T. and Tjian,R. (1997) *Cell*, **91**, 71–83.
20 Thompson,J.D., Higgins,G. and Gibson,T.B. (1994) *Nucleic Acids Res.*, **22**, 4673–4680.
21 Studier,F.W., Rosengerg,A.H., Dunn,J.J. and Dubendorff,J.W. (1990) *Methods Enzymol.*, **185**, 60–89.
22 Kato,K., Makino,Y., Kishimoto,T., Yamauchi,J., Kato,S., Muramatsu,M. and Tamura,T. (1994) *Nucleic Acids Res.*, **22**, 1179–1185.
23 Hoffmann,A. and Roeder,R.G. (1991) *Nucleic Acids Res.*, **19**, 6337–6338.
24 Schreiber,E., Matthias,P., Muller,M.M. and Schaffner,W. (1989) *Nucleic Acids Res.*, **17**, 6419.
25 Tamura,T., Ohya,Y., Miura,M., Aoyama,A., Inoue,T. and Mikoshiba,K. (1989) *Technique*, **1**, 33–36.
26 Peterson,M.G., Tanese,N., Pugh,B.H. and Tjian,R. (1990) *Science*, **248**, 1625–1630.
27 Malik,S., Hisatake,K., Sumimoto,H., Horikoshi,M. and Roeder,R.G. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 9553–9557.
28 Hoffmann,A. and Roeder,R.G. (1991) *Nucleic Acids Res.*, **19**, 6337–6338.
29 Peterson,M.G., Inostroza,J., Maxon,M.E., Flores,O., Admon,A., Reingerg,D. and Tjian,R. (1991) *Nature*, **354**, 369–373.
30 Ozer,J., Moore,P.A., Bolden,A.H., Lee,A., Rosen,C.A. and Lieberman,P.M. (1994) *Genes Dev.*, **8**, 2324–2335.
31 Aso,T., Conaway,J.W. and Conaway,R.C. (1994) *J. Biol. Chem.*, **269**, 26575–26583.
32 Hodo,H.G. and Blatti,S.P. (1977) *Biochemistry*, **16**, 2334–2343.
33 Parvin,J.D., Shykind,B.M., Meyers,R.E., Kim,J. and Sharp,P.A. (1994) *J. Biol. Chem.*, **269**, 18414–18421.
34 Sawadogo,M. and Roeder,R.G. (1985) *Proc. Natl Acad. Sci. USA*, **82**, 4394–4398.
35 Wieczorek,E., Brand,M., Jacq,X. and Tora,L. (1998) *Nature*, **393**, 187–191.