Cloning, Expression, and Characterization of the TATA-binding Protein (TBP) Promoter Binding Factor, a Transcription Activator of the Acanthamoeba TBP Gene*

Weibiao Huang and Erik Bateman†

From the Department of Microbiology and Molecular Genetics, Markey Center for Molecular Genetics, University of Vermont, Burlington, Vermont 05405

*This work was supported in part by National Eye Institute Grant EY 08706 (to E. B.), National Science Foundation, Vermont EPSCOR Grant RII8610679, and a grant from the Lucille P. Markey Charitable Trust to the Department of Microbiology and Molecular Genetics, University of Vermont. 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.

†To whom correspondence should be addressed. Tel.: 802-656-8608; Fax: 802-656-8749; E-mail: ebateman@moose.uvm.edu.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM/EMBL Data Bank with accession number(s) L46867.

TATA-binding protein (TBP) gene promoter binding factor (TPBF) is a transactivator which binds to the TBP promoter element (TPE) sequence of the Acanthamoeba TBP gene promoter and stimulates transcription in vitro. We have isolated a cDNA clone encoding TPBF. TPBF is a polypeptide of 327 amino acids with a calculated molecular mass of 37 kDa. The predicted amino acid sequence of TPBF shows no significant homology to other proteins. TPBF has two potential coiled-coil regions, a basic region, a proline-rich region, a histidine-rich N terminus, and a nuclear targeting sequence. The recombinant protein has an apparent molecular mass of 50 kDa, identical with that of TPBF purified from Acanthamoeba. Recombinant TPBF is able to bind DNA and activate transcription with the same specificity as natural TPBF. TPBF has two potential coiled-coil regions, a basic region, a proline-rich region, a histidine-rich N terminus, and a nuclear targeting sequence. The recombinant protein has an apparent molecular mass of 50 kDa, identical with that of TPBF purified from Acanthamoeba. Recombinant TPBF is able to bind DNA and activate transcription with the same specificity as natural Acanthamoeba TPBF, demonstrating the authenticity of the clone. Mobility shift assays of co-translated TPBF polypeptides and chemical cross-linking demonstrate that TPBF is tetrameric in solution and when bound to DNA. Analyses of TPBF mutants show that Coiled-coil II is essential for DNA binding, but Coiled-coil I and the basic region are also involved. TPBF is thus a novel DNA-binding protein with functional similarity to the tumor suppressor protein p53.

Accurate transcription initiation of all three classes of genes in eukaryotic cells requires stepwise assembly of several general transcription factors and the appropriate RNA polymerase on promoter DNA. The TATA-binding protein, TBP, is involved in transcription by all three RNA polymerases both in vitro and in vivo (1–5). TBP is complexed into SL1 (1), TFIID (6), and TFIIIB (7–10) for its function in RNA polymerase I, II, and III systems, respectively. These TBP-containing initiation factors are recruited to the different classes of promoters by specific protein-DNA (11, 12) and/or protein-protein interactions (1, 7–10, 13).

In the case of TATA-containing class II promoters, TFIIID, consisting of TBP and a large number of associated factors (TAFs) (5, 14–16), binds directly to DNA through specific interactions between TBP and the TATA box (11, 12) as the first step in formation of the initiation complex. This TFIIID-DNA complex then recruits other general transcription factors, such as TFIIA, TFIIE, TFIIF, and RNA polymerase II to form a complete initiation complex (17).

An additional class of transcription factors, known as sequence-specific transcription activators, is involved in efficient transcription by RNA polymerase II. These activators bind specifically to promoter sequences and modulate levels of expression of the selected genes, providing a regulatory strategy for eukaryotic cells to control development, differentiation, and their responses to extracellular stimuli. Evidence obtained in recent years suggests that sequence-specific activators stimulate transcription through direct or indirect (via coactivators) communication with the general transcription factors. Interactions between activators and general transcription factors TFIIA (18), TFIIA (19–21), TBP (22), TAFs (23, 24), TFIIF (25), and TFIIH (26) have been reported. However, the mechanism of transcription stimulation is not well understood.

Typical eukaryotic transcription activators are composed of discrete structural domains that have specific functions (27), for example, in multimerization, DNA binding, and transcription activation or repression. Different structural motifs involved in DNA binding and activation have been identified and used to classify transcription activators. A fully functional activator can be constructed by combination of functional domains from different activators (28).

The Acanthamoeba TBP gene promoter contains two major elements that are necessary for efficient transcription. The TBP box at −30 functions by binding TFIIID, which is necessary for basal transcription (29, 30). The TPE is a 23-base pair element centered around −90, which stimulates basal transcription up to 10-fold in vitro. The TPE binds a regulatory factor called TPBF (31), which was identified and purified previously in this laboratory (31, 32). TPBF is of interest because it regulates TBP gene transcription, but also because it is an apparently novel type of DNA-binding protein. Chemical interference assays demonstrated protein-DNA contacts on opposite faces of the DNA helix (32). This pattern, while reminiscent of the proposed model for p53 tetramer bound to DNA (33), is distinct from that produced by other factors (32). Although TPBF was found by gel filtration to be dimeric (32), the resolution was not sufficient to distinguish trimeric and tetrameric forms of the protein. Finally, TPBF is phosphorylated, and removal of phosphate increased DNA binding, suggesting that phosphorylation could play a regulatory function in vivo.
A Novel Tetrameric DNA-binding Protein

In order to determine the basis for these properties of TPBF and to permit further characterization of its role and mechanism in TBP gene expression, we isolated cDNA and genomic DNA clones encoding TPBF. Expression and analysis of cloned TPBF and mutant derivatives demonstrate that TPBF is a novel tetrameric DNA-binding protein. It contains a C-terminal coiled-coil domain necessary for tetramerization, as well as an apparently large central region involved in DNA binding. Other structural features of TPBF are discussed.

EXPERIMENTAL PROCEDURES

Purification of TPBF and Internal Peptide Sequencing—Purification of TPBF essentially followed the scheme described elsewhere (32). TPBF (5 mg) was loaded onto a 12% SDS-polyacrylamide gel (34), then blotted onto a polyvinylidene difluoride membrane (Applied Biosystems), and visualized by Ponceau S staining (35, 36). The TPBF band was cut out and used for tryptic digestion and internal amino acid sequencing by the Harvard Microchemistry Laboratory (36). Three peptides were sequenced (Fig. 1A). The amino acid sequences—AFQSNYR in peptide I and PYLTDAA in peptide II were used to design two sets of degenerate oligonucleotides for PCR amplification of the target gene.

Generating a TPBF-specific Probe by the Polymerase Chain Reaction—Two pairs of degenerate primers were synthesized corresponding to amino acid sequences—AFQSNYR and PYLTDAA: 1) GGC/G/CITCCAGCT(G/C)ACTACCG; 2) CGGTAGTT(G/C)GACTGGAAG(G/C/GC); 3) CCG/G/CITCCAGCT(G/C)ACG/GC/ATGC/AT/CIGC; 4) GCC/G/ATTCG/TGG/CTGG/GTGGAG/GTGC/GC.

Amplification of Acanthamoeba castellanii genomic DNA (37) by PCR was performed under the following cycle conditions: the first cycle at 94°C for 5 min, followed by 30 cycles of 94°C for 1 min, 48°C for 1 min, and 72°C for 2 min, and the last cycle at 72°C for 10 min. Several PCR products were generated using either primer combination (data not shown). The products were subcloned into the pSK(−) vector (Stratagene) and sequenced. One subclone, encoding the TPBF peptides, was obtained.

Isolation of a cDNA Clone and a Genomic Clone for TPBF—Using the subcloned PCR fragment as a probe, a A. castellanii cDNA library in λZAP (38) was screened. The cDNA library was plated out on 15-mm plates, each containing approximately 50,000 plaques. The plaques were blotted onto nitrocellulose filters (Schleicher & Schuell), and the DNA was denatured, neutralized, and immobilized (34). The filters were prehybridized at 65°C for 2 h in 500 mM NaHPO4, pH 7.2, 1% SDS, 1 mM EDTA, and denatured salmon sperm DNA (100 μg/ml), then hybridized with a 32P-labeled TPBF probe for 12 h at 65°C. Filters were then washed twice for 10 min at room temperature with 500 ml of 1 × SSC containing 0.1% SDS and once for 15 min at 65°C with 500 ml of 0.1 × SSC containing 0.1% SDS. Positive colonies were rescued as double-stranded plasmid DNA (39). A fragment of the cDNA clone of TPBF gene, an Acanthamoeba genomic DNA library, constructed in λEMBL3A (29), was screened using the cDNA-derived probe. One positive TPBF genomic clone was mapped to an 8-kilobase PstI fragment. The PstI fragment was subcloned into the pSK(−) vector and partially sequenced by primers derived from the cDNA sequence.

Northern blot Analysis—Total cellular RNA was isolated using guanidine isothiocyanate (38, 40), and mRNA was selected by two rounds of chromatography on oligo(dT)-cellulose (Life Technologies, Inc.). Samples of 10 μg of mRNA were subjected to electrophoresis on a 1% formaldehyde-agarose gel (34), transferred to a nitrocellulose filter, and hybridized with the 32P-labeled cDNA-derived probe as described above.

Southern blot Analysis—Acanthamoeba genomic DNA was prepared as described elsewhere (37). 2-μg aliquots were digested with BamHI, EcoRI, HindIII, PstI, MboI, HpalI, or MspI and separated by electrophoresis on a 0.8% agarose gel. Following depurination, the DNA was transferred to a nylon membrane (GeneScreen Plus, DuPont NEN) and probed with 32P-labeled cDNA (34).

Primer Extension of mRNA—2-μg of mRNA was dissolved in 10 μl of annealing buffer containing 20 μl Tris-HCl, pH 8.3, and 0.4 mM KCl and mixed with 50,000 to 100,000 cpm of labeled primer RT1 (GATGGTG-GCAGACATGGTGGTCTGT). The annealing reaction was incubated at 65°C for 10 min and then allowed to cool slowly to room temperature.

Electrophoretic Mobility Shift (EMS) Assay—The double-stranded 27-mer DNA fragment containing the Acanthamoeba TBP gene promoter sequence between −96 and −78 was used as the probe for EMS.
assays (32). For protein-DNA binding reactions, end-labeled DNA probe was incubated with 5 ng of purified native Acanthamoeba TPBF or purified recombinant TPBF or with 1–2-μl aliquots of TPBF synthesized in the in vitro transcription-translation reaction mixtures. The reaction conditions were as described previously (31). For the experiments shown in Fig. 4B and Fig. 8A, protein-DNA complexes were analyzed under the previously described conditions. For the experiments in Fig. 5 and Fig. 8B, the reactions were run on 6% native polyacrylamide gels at 10 V/cm for 5 h.

In Vitro Transcription Assay—Transcription with HeLa cell nuclear extracts was performed as described previously (31). A TBP promoter containing an intact TPBF binding site was used. 100 ng of purified recombinant full-length TPBF and mutant TPBF A254–296 were both assayed for their transcription activation. Transcription products were assayed by primer extension as described above.

Chemical Cross-linking of Proteins—Guanidinium thiocyanate (33) was used as a cross-linker. 200 ng of purified recombinant TPBF or ~20 ng of partially purified natural Acanthamoeba TPBF was incubated in DNA binding buffer without DTT and without DNA (32). Cross-linking was performed with a 25- to 50-fold molar excess of freshly prepared DTSSP solution or 0.001% glutaraldehyde. Reactions were incubated at room temperature for 30 min and quenched with 50 mM lysine. In some experiments, DTSSP cross-linked proteins were cleaved by incubation with 50 mM DTT at 37 °C for 30 min if necessary. Samples were resolved on 12% SDS-polyacrylamide gels. The gels were either stained with silver (34) or electroblotted for Western blotting analysis (see below).

Western Blotting—Proteins were transferred from a 12% SDS-polyacrylamide gel to a nitrocellulose filter in a Trans-Blot cell (Bio-Rad) using the conditions specified by the supplier. The membrane was blocked using 5% dry milk in phosphate-buffered saline containing 0.1% Nonidet P-40 after each incubation period. The membrane was washed extensively with phosphate-buffered saline containing 0.1% Nonidet P-40 for 30 min. It was then incubated with 1:200 diluted anti-TPBF sera for 30 min, followed by another 30-min incubation with 1:500 diluted horseradish peroxidase-linked anti-rabbit Ig (Amerham). The membrane was washed extensively with phosphate-buffered saline containing 0.1% Nonidet P-40 after each incubation period. Signals were detected using a Chemiluminescent Substrate Kit (Kirkgaard & Perry Laboratories).

RESULTS

Isolation of cDNA and Genomic Clones Encoding TPBF—Acanthamoeba TPBF was purified and subjected to internal sequencing as described under "Experimental Procedures." The amino acid sequences of three peptides were obtained (Fig. 1A) and used to design degenerate oligonucleotides for use in amplification of genomic DNA. PCR yielded a specific 200-bp fragment, which was subcloned and sequenced. 1 × 10⁶ plaques of an Acanthamoeba cDNA library were screened, and three positive cDNA clones were identified. All three clones contained sequences encoding peptides I, II, and III (Fig. 1A). Both strands of one of the cDNA clones were then completely sequenced. The 1060-bp cDNA contains an open reading frame of 327 amino acids with a predicted molecular mass of 37 kDa. The three clones had identical 5′ ends, but were not full-length based on the following observations: 1) there is a large discrepancy between the apparent size of natural Acanthamoeba TPBF (50–51 kDa) and the deduced size predicted from the cDNA; 2) the single open reading frame extends to the 5′ end of the cDNA; 3) primer extension of mRNA using a primer located 100 bp from the 5′ end of the cDNAs generated a product of 150 nucleotides (Fig. 2C and see below), which suggested that 50 bp was missing from the cDNA.

We used "rapid amplification of cDNA ends" (43) to obtain the missing 5′ end of the cDNA. In order to find a rapid amplification of cDNA ends primer, we isolated a genomic clone encoding TPBF and partially sequenced it. Two in-frame methionines were found in the 50-bp sequence preceding the 5′ end of the cDNA within the genomic copy of the TPBF gene. Using the primer starting from the first in-frame methionine in combination with primer RT2, we successfully obtained the missing part of the cDNA from mRNA by reverse transcription-PCR and reconstructed the full-length cDNA. The complete (both strands) sequence of the reconstructed TPBF cDNA is presented in Fig. 1A.

The complete cDNA comprises 1088 bp and contains an open reading frame of 327 amino acids with a predicted molecular mass of 37 kDa (Fig. 1A). There are several sequence motifs of potential importance to the function of TPBF as a transcription activator. First, it bears a putative nuclear localization signal

FIG. 1. Nucleotide sequence, predicted amino acid sequence, and schematic diagram of the TPBF gene. A, nucleotide and deduced amino acid sequences corresponding to the sequenced peptides. Open triangles show the two arrays of heptad repeats of hydrophobic amino acids. B, schematic presentation of TPBF indicating presumptive functional domains (see text). Relative positioning of the domains is shown by the scale above the gene. These data have been submitted to GenBank™ under accession number L46867.
that of TPBF protein in nuclear extracts, which contain low. The low abundance of the TPBF message is in contrast to amino acids 254 to 296 (bic residues in the sequence (indicated by therearetwosegmentscontainingheptadrepeatsofhydropho-
1), which appears in many transcription factors (44). Second,
254–296, residues 40 and 85 is proline-rich, which, by analogy to other factors, might be important in mediating transcription activation (47). Fourth, the N-terminal 24 amino acids of TPBF is unusually histidine-rich, containing 10 histidine residues. However, data base searches showed that TPBF lacks significant sequence homology to any other known genes (48). These features are considered further under “Discussion.”

Analysis of the Genomic Copy of the TPBF Gene and Its Transcript—Digestion of Acanthamoeba genomic DNA with BamHI, EcoRI, HindIII, or PstI followed by Southern blot analysis generated a single major band in each case (Fig. 2A), suggesting that the TPBF gene is unique within the Acanthamoeba genome. Multiple bands produced by digestion with MboI, Hpal, or Mspl for the existence of restriction sites within the gene.

Northern analysis (Fig. 2B) indicates that TPBF is transcribed into a single mRNA with a size of about 1,100 nucleotides. Primer extension of Acanthamoeba mRNA generated one single band of the expected size (Fig. 2C), indicating that the transcript begins about 30 bp downstream from an imperfect TATA box within the genomic copy of the TPBF gene. The TPBF transcript appears to be extremely rare. 10 μg of mRNA was required to obtain a clear signal in Northern blot analysis. Similarly, only 3 plaques from 1 x 10⁶ were obtained, suggesting that the level of TPBF expression in Acanthamoeba is very low. The low abundance of the TPBF message is in contrast to that of TPBF protein in nuclear extracts, which contain ~200 ng of TPBF/mg as judged by Western blotting (data not shown).

Expression of Full-length and Mutant TPBF in E. coli—In order to express full-length TPBF and a TPBF mutant lacking amino acids 254 to 296 (Δ254–296), we subcloned the corresponding cDNAs into the pET3a vector as detailed under “Experimental Procedures.” However, due to the toxicity of the proteins in E. coli, we were unable to maintain the expression plasmids in either E. coli BL21(DE3) or in BL21(DE3)pLysS (41). We therefore had to maintain the plasmids in E. coli LE392 and induce protein expression by infection with λCE6 (41) to provide T7 RNA polymerase. Both proteins were expressed at high levels in infected E. coli (Fig. 3, compare lanes 1 and 2, data not shown for Δ254–296).

The existence of histidine-rich sequences at the N terminus of full-length or deleted TPBF enabled us to purify them by Ni²⁺-affinity chromatography (Fig. 3, lane 5). The recombinant proteins were further purified by DEAE-cellulose chromatography, yielding a single major band (Fig. 3, lane 6).

The molecular mass of the full-length TPBF as measured by SDS-PAGE is 50 kDa, which differs significantly from the predicted mass of 37 kDa. However, the SDS gel mobility of recombinant TPBF perfectly matches that of natural Acanthamoeba TPBF purified from nuclei (Fig. 4A). The apparent discrepancy between SDS gel mobility and the predicted molecular mass of TPBF may be due to the abundance of positively charged amino acids in the protein. Natural TPBF migrates as a doublet on SDS gel due to phosphorylation (32). Surprisingly, recombinant TPBF showed the same mobility as the phosphorylated form of TPBF (Fig. 4A, lanes 1 and 2).

Recombinant TPBF Expressed in E. coli Binds the TPE Sequence and Transactivates the Acanthamoeba TBP Gene Promoter—The DNA binding activities of natural TPBF, purified recombinant TPBF, and the TPBF mutant Δ254–296 (shown in Fig. 4A) were compared by gel mobility shift assays. The complex between the recombinant TPBF and DNA had a mobility identical with that of the natural Acanthamoeba TPBF-DNA complex (Fig. 4B, lanes 1 and 2). The binding specificity was tested by competition with either a specific or mutant TPE (Fig. 4B, lanes 3 and 4). The results demonstrated a functional identity between recombinant and natural TPBF. Comparing the intensities of the bands, no significant difference in DNA binding activities was observed between recombinant and nat-
activation as stimulated by natural TPBF, as determined by titration with rTPBF. This may suggest that a significant portion of recombinant TPBF that is active in DNA binding is deficient in transactivation. As expected, the TPBF mutant Δ254–296 is unable to stimulate transcription (Fig. 4C, lane 3). To ensure that the observed transactivation was TPE-dependent, parallel experiments were done using a TBP promoter that lacks the TPE element. As expected, TPBF was not able to stimulate transcription in the absence of the TPE sequence (data not shown). We have also obtained similar results using Acanthamoeba extracts immunodepleted of TPBF (data not shown).

TPBF Forms a Tetramer When Bound to DNA—We previously reported that TPBF exists as a dimer or higher order oligomer (32). To definitively determine its oligomerization state, full-length TPBF and a truncated TPBF (Δ1–76) were synthesized individually or in combination using an in vitro transcription-translation system (49). The truncated protein was mutant Δ1–76 (Fig. 7A) which has DNA binding activity (Fig. 5, lane 6) and can be resolved easily from full-length TPBF in gel mobility shift assays. When individually synthesized proteins were assayed for TPE binding, TPBF formed one shifted protein-DNA band while the mutant generated one major band and a minor band. The minor band is caused by a partial length TPBF polypeptide produced during translation. The wild type and mutant bands had the expected difference in mobility (Fig. 5, lanes 1 and 2). When the polypeptides cotranslated at varying ratios while keeping the total protein amount approximately constant were assayed, five major bands were clearly visible. There is a gradual distribution from the largest to the smallest bands with increasing amounts of the mutant protein (Fig. 5, lanes 1–6). Two of the bands (indicated by open triangles) were seen in individually synthesized proteins. The other three bands (indicated by solid triangles) had intermediate mobilities distributed between the two outer bands. A faint band distributed in between the two lower major bands may be the product of association between TPBF and the shortened version of the mutant polypeptide.

The formation of five major different complexes by the cotranslated proteins strongly suggests that TPBF forms a tetramer when bound to DNA. The two outer bands correspond to homo-oligomeric complexes (L₄ and S₄), while the three inner bands correspond to hetero-oligomeric complexes (L₃S₁, L₂S₂, L₁S₃).
TPBF Exists as a Tetramer in Solution—In order to determine the oligomerization state of TPBF in solution, we performed cross-linking experiments with either DTSSP or glutaraldehyde, in the absence of DNA. Cross-linking of purified recombinant TPBF with either DTSSP or glutaraldehyde, followed by SDS-PAGE and silver-staining, produced two cross-linked bands with apparent molecular masses of about 160 kDa and 140 kDa (Fig. 6A). The size difference between these two bands indicates they are unlikely to be different oligomers. Most likely, the two bands were generated by cross-linking at different residues. Thus we believe that both bands correspond to one single form of oligomer. The size of the cross-linked bands, about 4 times that of the TPBF monomer, in combination with the above observation that TPBF binds DNA as a tetramer, suggests that TPBF is also tetrameric in the absence of DNA.

DTSSP cross-linking of partially purified Acanthamoeba TPBF detected by immunoblotting produced two cross-linked products apparently identical with those produced by recombinant protein (Fig. 6B), indicating that recombinant TPBF has the same structure as natural TPBF. This result also demonstrates that both cross-linked bands contain TPBF. Cross-linked bands were removable by treatment with 50 mM DTT, which cleaves the disulfide bond within DTSSP linking the monomers (Fig. 6B, lane 4).

Cross-linking also showed that mutant Δ254–296 is unable to form a tetramer in solution (data not shown). Loss of multimerization of TPBF mutant Δ254–296 is likely due to loss of the Coiled-coil II structure. Multimerization of TPBF is thus evidently necessary for binding to DNA (see also below).

Identification of Regions in TPBF Responsible for DNA Binding—To further delineate the coiled-coil region and investigate other regions required for specific DNA binding, we made a series of TPBF deletion mutants as detailed under “Experimental Procedures” and summarized in Fig. 7A. Mutant and wild type proteins were synthesized using coupled in vitro transcription-translation in the presence of [35S]methionine. Analysis of proteins by SDS-PAGE and autoradiography showed that a comparable amount of each mutant TPBF was made (Fig. 7B). The sizes of the mutant polypeptides were roughly proportional to the sizes of the mutated genes (compare A and B of Fig. 7). Some of the products migrated as doublets (Fig. 7B) because the translation system utilized either of two methionines located close to the N terminus (Fig. 1A). All these products were assayed by gel mobility shift for their abilities to bind the TPE element. Protein synthesized from the wild type TPBF construct generated a band that had exactly the same mobility as the band produced by natural Acanthamoeba TPBF (Fig. 8A, lanes 1 and 2). Deleting the first 20 amino acids, which are very histidine-rich, did not have an observable effect on DNA binding activity (Fig. 8A, lanes 2 and 3). Removal of amino acid residues 1 to 76 significantly reduced DNA binding activity (Fig. 8A, lanes 3 and 4), indicating that this region is involved in, but not essential for, DNA binding. Further deletion to residue 122 reduced DNA binding activity to near background level (Fig. 8A, lane 5, and 8B, lane 1). Deletion of residues 127–253 totally abolished DNA binding activity suggesting that an essential region had been removed (Fig. 8A, lane 6, and Fig. 8B, lane 2).

To examine the role of Coiled-coil II in DNA binding, we checked the internal deletion Δ254–296 in this system and again found it was inactive in DNA binding (Fig. 8A, lane 7, and Fig. 8B, lane 3). This deletion removes two heptads from Coiled-coil II, presumably preventing tetramerization. We also tested two C-terminal deletion mutants. Interestingly, removal of 7 amino acid residues from the C terminus increased the
A Novel Tetrameric DNA-binding Protein

DNA binding activity several fold (Fig. 8A, lanes 2 and 9). Removal of Coiled-coil II (mutant A278-327) resulted in the production of a polypeptide unable to bind DNA (Fig. 8A, lane 8, and Fig. 8B, lane 4). These results indicated that regions essential for DNA binding are distributed between amino acid residues 123 and 320. Presumably, the C-terminal Coiled-coil II drives tetramerization and is therefore necessary for DNA binding; while the regions that make DNA contact are located between amino acid residues 123 and 280.

**DISCUSSION**

We have isolated full-length cDNA encoding TPBF, an Acanthamoeba transcription activator, which regulates expression of the TBP gene (31, 32). To our knowledge, TPBF is the first regulatory protein isolated and cloned that controlsexpression of the TBP gene (31, 32). To our knowledge, TPBF is the first regulatory protein isolated and cloned that controlsexpression of the TBP gene (31, 32). To our knowledge, TPBF is the first regulatory protein isolated and cloned that controlsexpression of the TBP gene (31, 32).

TPBF shows the same DNA binding specificity and activity as natural TPBF. Fourth, recombinant TPBF is able to stimulate transcription in a TPE promoter-dependent fashion. Fifth, recombinant TPBF and natural TPBF bind avidly to a nickel affinity column due to the histidines present in the N terminus. Finally, antibody raised against recombinant TPBF recognizes natural TPBF in Acanthamoeba nuclear extracts. There are, however, some physical differences between recombinant and natural TPBF. For example, TPBF produced in E. coli, which is presumably not phosphorylated, comigrates during SDS-PAGE with the phosphorylated form of natural TPBF. It is thus possible that either recombinant TPBF has been modified or natural TPBF has additional unidentified modifications. Similarly, recombinant TPBF is somewhat less active in stimulating transcription than natural TPBF.

The predicted amino acid sequence of TPBF has no significant homologues in the public data bases (48), in accord with its unusual DNA binding properties. Many sequence-specific transcription factors have been grouped into several distinct families, such as basic helix-loop-helix, zinc finger, homeodomain, helix turn helix, or leucine zipper proteins (50). TPBF does not fall into any of these families, as judged by alignment between TPBF and the consensus sequences that characterize each family. However, the TPBF sequence contains several regions that suggest a function (Fig. 18). The N-terminal domain is remarkably histidine-rich, and these histidines can coordinate with chelated nickel. While we do not yet know whether TPBF requires metal for any of its activities, it conceivably contains Zn²⁺ or Ni²⁺, perhaps arranged in a configuration similar to the metal ions in urease (51). Metal coordination by histidines might stabilize a particular protein conformation analogous to zinc fingers or zinc clusters (50). While the histidines are not required for DNA binding based on mutagenesis, they could be involved in another function that our assays did not assess, for example, transcription activation. A similar possibility exists for the proline-rich region between amino acid residues 40 and 85. Proline-rich domains can function as activation regions in some transcription factors such as CTF (47). However, we have been unable to localize the transactivation domain directly.
A Novel Tetrameric DNA-binding Protein

since we have been unable to establish a homologous system to assay mutant TPBFs. We failed to recover activated transcription by simply adding back-purified recombinant or native TPBF into a nuclear extract in which TPBF has been seques-
tered by specific TPE DNA, or nickel affinity chromatography (data not shown). All these results suggest TPBF might employ a coactivator or adaptor to mediate its transactivation activity (18, 52). It will be of considerable interest to identify the trans-
activation domain of TPBF as well as its target.

Adjacent to and overlapping the proline-rich domain, there is a potential coiled-coil domain (Coiled-coil I) comprising a heptad repeat of hydrophobic residues. While this region could potentially contribute to tetramerization (see below), deletion mutagenesis suggests it is not necessary, but instead may have a stabilizing effect on the overall structure. However, because this region contains two proline residues which are likely to destabilize or prevent helix formation, the importance of Coiled-coil I is somewhat unclear.

In the central portion of TPBF, there is a putative nuclear localization signal and a region that is rich in basic residues. By analogy with leucine zipper proteins or basic helix-loop-helix proteins, it is possible that this latter region may be involved in DNA binding.

At the C terminus of TPBF there is an additional coiled-coil domain (Coiled-coil II), containing hydrophobic 4–3 repeats. Positions a and d (the positions within a heptad repeat are conventionally referred to as a, b, c, d, e, f, and g, see Ref. 53) of the heptads in the array named Coiled-coil II are almost perfectly hydrophobic. Although this region resembles a leucine zipper, its perfectly amphipathic hydrophobic character suggests that it is likely to form higher order oligomers, since perfect coiled-coil domains can form trimeric or tetrameric bundles (54). In accord with this prediction, direct chemical cross-linking and cotranslation experiments establish that TPBF exists as a tetramer both in solution and when bound to DNA.

Analyses of several TPBF deletion mutants supports the predictions made from inspection of its sequence. Although our analyses were constrained by an inability to express all mutants in E. coli, several important conclusions were reached. Chemical cross-linking and binding studies of mutant proteins demonstrate that Coiled-coil II is essential for tetramerization and therefore DNA binding. Coiled-coil II is the major, if not the only, region driving tetramerization of TPBF. Although Coiled-coil I is not essential for DNA binding (Fig. 6B), removal of the region greatly reduces DNA binding activity, suggesting its involvement in either stabilizing tetramer or tetramer-DNA complex. Other proteins with two or more separate coiled-coil domains have been reported (55, 56). The reovirus col attachment protein r1 has two coiled-coils, one is involved in the formation of a loose multimer while the other stabilizes the multimer (56). Our prediction of how the two coiled-coil functions in TPBF is similar to this model, i.e. Coiled-coil II mediates tetramerization of TPBF, and the tetramer is further stabilized by Coiled-coil I. This idea is supported by the preferential formation of homotetramer of wild type TPBF over a truncated version lacking Coiled-coil I (Fig. 5, lane 2). It is likely that the association of cotranslated polypeptides is not random. Instead, it favors the formation of tetramers containing Coiled-coil I, which provides additional stabilization. However, Coiled-coil I may also be involved in stabilizing tetramer-DNA interactions.

TPBF requires a relatively large domain for efficient, sequence-specific DNA binding. Unlike GCN4, for example, which only requires the C-terminal 56 amino acid residues for DNA binding (57), the region in TPBF required for efficient DNA binding spreads from 20 to 320. Since the mutant 11–122 had very weak DNA binding activity, the region from 20 to 122 is most likely involved in determining the binding efficiency but not specificity. The region involved in determining DNA sequence specificity is therefore contained within amino acids 123 to 281. We have preliminary evidence suggesting the basic region from 123 to 194 is in fact necessary for DNA binding. However, finer mapping needs to be done to further define the region necessary for sequence-specific DNA binding.

Previous studies showed that TPBF makes numerous symmetrical contacts with TPE (32). The overall pattern of the base and phosphate contacts, suggesting that TPBF contacts DNA symmetrically on opposite faces of the helix, is unique. This binding pattern is in keeping with the novel structural features of TPBF, especially its divergent coiled-coil region and the widely spread basic region. One of the unique features of TPBF is that it is tetrameric, which is probably determined by the arrangement of hydrophobic residues in positions a and d in Coiled-coil II (54). Tetramerization of TPBF can explain the protein-DNA contacts inferred from chemical interference assays. First, TPBF is able to occupy a large region of DNA since it contains a four-stranded helical bundle. Second, the region for tetramerization is located at the C terminus of TPBF allowing four DNA binding domains of TPBF, located more than 100 amino acid residues from the tetramerization domain, to reach relatively distal binding sites. It is possible that the tetramerized TPBF contacts the DNA helix perpendicularly from one side. Thus, its four DNA binding domains can make symmetrical contacts on opposite faces of the DNA helix. Interestingly, this DNA binding pattern resembles the way tumor suppressor p53 binds to its specific DNA recognition sequence (33). Although no obvious similarities exist between these two proteins in amino acid sequence or specific DNA binding sites, it is possible that they belong to a novel family of DNA-binding proteins that function in regulating expression of genes involved in growth and differentiation.

Acknowledgments—We thank Dr. Feng Liu for help in TPBF purification. We also thank Drs. David Pederson and Tom Orfeo for comments on the manuscript.

REFERENCES
1. Comai, L., Tanese, N., and Tjian, R. (1992) Cell 68, 965–976
2. Cormack, B. P., and Struhl, K. (1992) Cell 69, 685–696
3. Schultz, M. C., Reeder, R. H., and Hahn, S. (1992) Cell 69, 697–702
4. Yamashita, S., Hisatake, K., Doi, K., Roeder, R. G., Horikoshi, M., and Nakatani, Y. (1993) Science 261, 463–466
5. Hernandez, N. (1993) Genes & Dev. 7, 1291–1308
6. Pugh, B. F., and Tjian, R. (1992) J. Biol. Chem. 267, 679–682
7. Kavalov, C. A., Kozlowski, C. A. P., Pisano, M., Oudard, E. P., Colbert, T., Hahn, S., and Blanco, J. A. (1992) Cell 71, 1055–1064
8. Lobo, M. S., Tanaka, M., Sullivan, M. L., and Hernandez, N. (1992) Cell 71, 1029–1040
9. Taggart, A. K. P., Fisher, T. S., and Pugh, B. F. (1992) Cell 71, 1015–1028
10. White, R. J., and Jackson, S. P. (1992) Cell 71, 1041–1053
11. Sawadogo, M., and Roeder, R. G. (1993) EMBO J. 12, 165–175
12. Nakajima, N., Horikoshi, M., and Roeder, R. G. (1988) Mol. Cell Biol. 8, 4028–4040
13. Jantzen, H., Chow, A. M., King, D. S., and Tjian, R. (1992) Genes & Dev, 6, 1950–1963
14. Rigby, P. W. J. (1993) Cell 11, 1628–1640
15. Sharp, P. A. (1992) EMBO J. 11, 821–829
16. Greenblatt, J. (1991) Cell 66, 1067–1070
17. Buratowski, S. (1994) Cell 77, 1–3
18. Ge, H., and Roeder, R. (1994) Cell 78, 513–523
19. Liu, X. S., and Green, M. R. (1991) Cell 64, 971–981
20. Xu, X., Prorock, C., Ishikawa, H., Maldonado, E., Ito, Y., and Gelinas, C. (1993) Mol. Cell Biol. 13, 6733–6741
21. Kim, T. K., and Roeder, R. G. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 4170–4174
22. Stringer, K. F., Ingles, C. J., and Greenblatt, J. (1990) Nature 345, 783–786
23. Goodrich, J. A., Hoey, T., Thut, C., Admon, A., and Tjian, R. (1993) Cell 75, 519–530
24. Hoey, T., Weinzierl, R. O. J., Gill, G., Chen, J. L., Dynlacht, B. D., and Tjian, R. (1993) Cell 72, 247–260
25. Joliot, V., Demma, M., and Prywes, R. (1995) Nature 373, 632–635
26. Xiao, H., Pearson, A., Coulombe, B., Truant, R., Zhang, S., Regier, J. L., Trizeneberg, S. J., Reinberg, D., Flores, O., Ingles, C. J., and Greenblatt, J.

W. Huang and E. Bateman, unpublished data.
