Isolation of two cDNAs encoding zinc finger proteins which bind to the α_1-antitrypsin promoter and to the major histocompatibility complex class I enhancer

Catherine Mitchelmore*, Cinzia Traboni[1] and Riccardo Cortese+
European Molecular Biology Laboratory, Meyerhofstraße 1, 6900-Heidelberg, FRG and [1]Dipartmento Biopatologia, Policlinico Umberto I, Rome, Italy

Received August 15, 1990; Revised and Accepted September 27, 1990 EMBL accession nos X54249, X54250

ABSTRACT

Two partial cDNAs coding for DNA-binding proteins (AT-BP1 and AT-BP2) have been isolated. Both proteins, when prepared from λgt11 lysogens, bind to the B-domain of the α_1-antitrypsin promoter, an element which is important for the liver-specific expression of α_1-antitrypsin. Analysis of the cDNA sequences encoding these proteins reveals that both contain two zinc fingers of the Cys_2-His_2 type followed by a highly acidic stretch of 20 amino acids. AT-BP1 contains a second putative DNA-binding domain consisting of an 8-fold repeat of a SPKK (Ser-Pro-Lys/Arg-Lys/Arg) motif. Both proteins bind to the NF-κB recognition site in the MHC gene enhancer with significantly higher affinity than to the κ immunoglobulin gene enhancer, or to the B-domain of the α_1-antitrypsin gene promoter. Analysis of mRNA expression shows that AT-BP1 and AT-BP2 are expressed in all the tissues examined. While the physiological roles of AT-BP1 and AT-BP2 remain to be elucidated, their predicted amino acid sequence and their DNA-binding characteristics suggest a role as transcriptional regulators.

INTRODUCTION

Many DNA-binding proteins which act in *trans* to activate or repress transcription of particular genes have been recently described. Isolation of the cDNAs encoding such trans-acting factors reveals that their DNA-binding motifs tend to fall into one of several classes (reviewed in refs.1,2), including zinc finger proteins of the Cys_2-His_2 type, steroid hormone receptors containing Cys_4 type zinc finger motifs, the homeodomain family, the leucine zipper family and the helix-loop-helix family (3). In addition to the DNA-binding domains, several motifs important in trans-activation have been identified (2): these include acidic domains, glutamine-rich domains and proline-rich domains. These domains aid assembly of the transcription initiation complex near the transcription start site (4–6).

Tissue-specific gene transcription is achieved through the interaction of ubiquitous and cell type-specific trans-acting factors. In the liver, two proteins, LF-A1 and LF-B1, have been shown to be positively-acting transcription factors whose binding to the promoter of several liver-specific genes is essential for their tissue-restricted expression (7). Although these proteins were previously thought to be restricted to liver cells (8), a DNA-binding activity indistinguishable from LF-B1 is also detectable in kidney nuclear extracts (9). The cDNA encoding LF-B1 has been cloned (9,10) and the corresponding mRNA is present in similar amounts in liver and kidney, and at lower levels in intestine, spleen and thymus (9). The cloned LF-B1 gene encodes a protein which contains an unusually large homeodomain and binds to DNA as a dimer (11). Previous studies have shown that the binding of LF-B1 to its recognition site in the B-domain of the α_1-antitrypsin gene promoter is mutually exclusive with that of a protein called LF-B2 (8). The isolation of a cDNA clone encoding LF-B2 was therefore required to examine its possible function as a negative regulator of transcription.

A λgt11 expression library prepared from rat liver cDNA was screened (12) and a recombinant phage was isolated which bound to the B-domain (containing recognition sites for LF-B1 and LF-B2), but not to the B1-subdomain. The nucleotide sequence of this cDNA and a distinct cDNA isolated in a subsequent screening (with the original cDNA as probe) revealed a high sequence homology with a previously identified zinc finger protein called HIV-EP1 (13), MBP-1 (14), or PRDII-BF1 (15). In each case, the cDNAs were isolated by screening human cDNA λgt11 expression libraries with different probes containing recognition sites for the transcriptional activator NF-κB (16).

Here we show that AT-BP1 and AT-BP2 (for α_1-Anitrypsin Binding Protein) both contain zinc fingers and bind weakly to the α_1-antitrypsin B-domain, and with high affinity to the NF-κB recognition site of the MHC class I enhancer, but are distinct from LF-B2 or NF-κB. Whereas AT-BP2 is probably the rat homologue of PRDII-BF1, the factor AT-BP1 is a newly identified protein.

* To whom correspondence should be addressed

+ Present address: Instituto Ricerche di Biologia Molecolare, Via Pontina Km 30.6, Pomezia, Italy

MATERIALS AND METHODS

Screening λgt11 cDNA Libraries

A rat liver cDNA library in λgt11 (Clontech Laboratories, Inc.) was screened with five and four copies respectively of the A- and B-oligonucleotides derived from the α_1-antitrypsin promoter (AR-2 and AT75 oligonucleotides respectively, see below) according to the method of Singh et al (12).

For further screening of cDNA libraries, restriction fragments of the λgt11 cDNA inserts were gel-purified and radioactively labeled by the random hexamer priming method (17). Phage plaques from a second rat liver cDNA library (prepared by M. Frain) and a thyroid cDNA library (prepared by M. Price) were screened according to Frain et al. (10).

Subcloning and Sequencing of DNA

Recombinant phage DNA was prepared according to Manfioletti and Schneider (18). The EcoRI-digested DNA was subcloned into Bluescript and pUC plasmids, and further subclones were constructed in these vectors. Plasmid DNA was prepared from overnight cultures grown in the presence of antibiotic selection by alkali lysis (19). Both strands of the DNA were sequenced by the dideoxy chain termination method (20) using T7 Polymerase DNA sequencing kits from United States Biochemical Corporation and Pharmacia. Synthetic oligonucleotides were used for priming on the flanking sides of cDNA inserts or at internal sites. Difficult regions were sequenced with inosine or 7-deaza-guanosine mixes substituted for the guanosine mix. The sequencing reactions were run on 6% denaturing polyacrylamide gels.

Preparation of λgt11 Fusion Protein

Wild type and recombinant λgt11 lysogens were prepared in *E.coli* Y1089 cells as described by Huynh et al. (21). Lysogens were grown at 30°C to an OD_{600} of 0.5 and then heat-shocked by addition of 1/4 volume L-Broth at 80°C prior to transferring cultures to a shaking 44°C waterbath for 20 minutes. IPTG was added to 2mM final concentration and the cultures were allowed to grow at 37°C for a further 2 hours. Bacterial extracts were prepared according to Singh et al, 1988, except that the final supernatants were not dialyzed. Protein concentrations were determined using the Bio-Rad Bradford dye reagent. SDS-PAGE gels were prepared and stained with Coomassie blue according to Laemmli (22).

Gel Retardation Assays

1 pmol of single-stranded oligonucleotide was labelled with polynucleotide kinase (New England Biolabs) and [^{32}P]-γ-ATP (23) before annealing with 10 pmol of the (unlabelled) complementary strand. Unincorporated nucleotides were removed by passing the labelled probe over a Sephadex G50 column prepared in a 2ml disposable pipette. The oligonucleotides (top strand) used in these experiments are described below:

AT75	5′ AACTGGGGTGACCTTGGTTAATATTCACCAG 3′
B2	5′ AACTGGGGTGACCTTGGTTA 3′
MHC	5′ GATCCGGCTGGGGATTCCCCATCT 3′
HIV(2)	5′ GATCTAGGGACTTTCCGCTGGGGACTTTCCAG 3′
HIV(R)	5′ TTCCGCTGGGGACTTTCCAG 3′
κB	5′ GATCCAGAGGGGACTTTCCGAGAGG 3′
ApoB	5′ GGGCCCGGGAGGCGCCCTTTGGACCTTTTG 3′
AR-2	5′ GATCCCAGCCAGTGGACTTAGCCC 3′
MB	5′ TGAAGCAAAGTATTTAATTTGGTTAGTAATTAC TAAACAC 3′

Up to 8μg total protein in lysogen extracts was preincubated on ice in a 20μl reaction volume containing 20mM Tris (pH7.6), 8% Ficoll, 2.5mM KCl, 4mM spermidine, 1mM EDTA, 0.2mM DTT, 3μg poly(dI-dC), and 50ng sonicated salmon sperm DNA. Unlabelled double-stranded oligonucleotides were included in the preincubation mix in competition experiments. After 10 min, 20,000 cpm (Cerenkov) of labelled probe was added and incubation was continued for 30 min on ice. DNA-protein

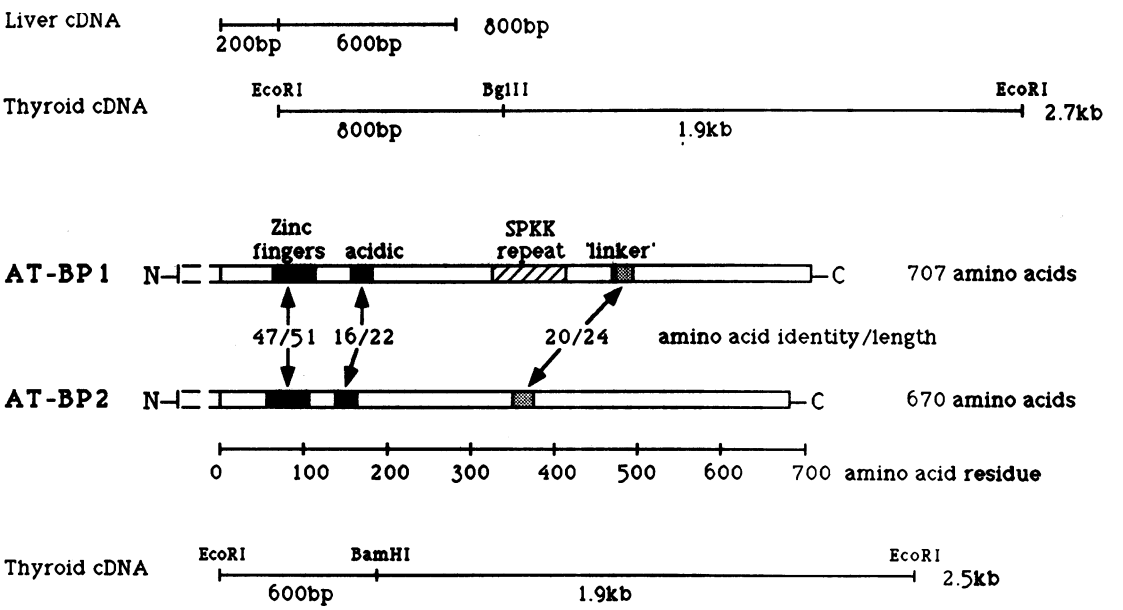

Figure 1. cDNAs and predicted protein structures of AT-BP1 and AT-BP2. The overlapping cDNAs encoding AT-BP1 isolated from a liver and thyroid cDNA library respectively are indicated at the top, and the cDNA encoding AT-BP2 isolated from a thyroid cDNA library at the bottom of the figure. The predicted protein structures are compared, and the amino acid identities between motifs in the two proteins are shown.

complexes were resolved from free DNA on a 6% polyacrylamide gel in 0.5×TBE. After the run the gel was dried onto DE52 paper and exposed to X-ray film.

Northern Analysis

Approximately 3μg of poly(A)-tailed RNA per rat tissue (kindly provided by M. De Felice) was size-fractionated on a 1% agarose 2.2M formaldehyde gel. A 0.24−9.5kb RNA ladder (Bethseda Research Laboratories) was run alongside. After electrophoresis, the gel was blotted onto a Hybond-N membrane (Amersham) by capillary transfer and the filter was UV-crosslinked. The RNA ladder lane was cut from the filter and stained with methylene blue (19). The filter was then hybridized with random primed restriction fragment probe as previously described (23).

The quantity and integrity of the loaded mRNA were subsequently checked by hybridization with a probe for the housekeeping GAPDH gene. The Amersham nonradioactive DNA labelling and detection kit was used precisely according to the instructions.

RESULTS

Isolation of cDNAs Encoding AT-BP1 and AT-BP2

A λgt11 expression library prepared from rat liver cDNA was screened with multimerised oligonucleotides corresponding to the A- and B-domains of the α_1-antitrypsin promoter (8), according to the protocol described by Singh et al. (12). Several positive clones were identified which bound to the multimerised B-oligonucleotide but not the A- or B1- oligonucleotides. Analysis of the cDNA inserts suggested that they were partial and encoded identical open reading frames of about 800bp which represented the DNA-binding portion of the β-galactosidase fusion protein. Attempts to isolate longer cDNAs from a second rat liver cDNA library were unsuccessful. We thus decided to screen a thyroid cDNA library since we had previously shown that an LF-B2-like DNA-binding activity was present in thyroid nuclear extracts (data not shown). Several cDNA clones were isolated which shared an identical 600bp overlap with the original clone and extended in the 3′ direction. The longest of these clones (2.7 kb) was sequenced and the extended sequence was designated AT-BP1 (for α_1-Antitrypsin Binding Protein). See Fig. 1.

A second class of positive clone was picked up in the thyroid cDNA library screening, by analysing phage plaques giving a weak hybridisation signal with the labelled probe. The longest of these, called AT-BP2, was sequenced and found to be highly homologous to AT-BP1, particularly in the DNA-binding domain.

Both AT-BP1 and AT-BP2 Encode Zinc-Finger Proteins

A comparison of the predicted amino acid sequences of AT-BP1 and AT-BP2 is shown in Fig. 2. Both proteins contain two zinc fingers of the C_2-H_2-type (see ref. 24 for a review) followed by a stretch of acidic residues. These motifs are highly conserved between the two proteins at the amino acid level, although third-base nucleotide substitutions are frequent. There is also a short stretch of 24 amino acids near the C-terminus which is highly conserved, see Figure 1. The overall identity in amino acid sequence is 40%, and for conserved residues is 59%. In AT-BP1, an 8-fold repeat of a Ser-Pro-Lys/Arg-Lys/Arg (SPKK) motif is also present. This motif is found in histone proteins (25) and may be important in DNA binding. The 3′ non-coding nucleotide sequences are noticeably A-T rich.

The DNA sequence of AT-BP2 has 82% identity with that of a previously identified human protein, PRDII-BF1 (13−15). At the amino acid level, there is 100% identity in the zinc finger domain and 76% identity overall, suggesting that AT-BP2 is the rat homologue of the human PRDII-BF1 gene.

Characterization of the DNA-Binding Activity of the Fusion Proteins

The β-galactosidase fusion proteins containing the DNA-binding domains of AT-BP1 and AT-BP2 (encoding the first 300 amino acids of each) were produced according to Singh et al., (12) and used in gel retardation assays to characterize their DNA binding specificity. Both fusion proteins bind to labelled probes corresponding to the α_1-antitrypsin B-domain (AT75) (Fig. 3) and the B2-domain (not shown). In each case, they form two specific DNA-protein complexes in gel retardation assays. The appearance of the faster migrating complex varies with extract preparation. Furthermore, the DNA-binding activity co-purifies with two proteins of approximately 100 and 140 kDa present in the recombinant lysogen extract of AT-BP1 and not that of wild type λgt11 (by purification over an AT75 affinity column; data not shown). The higher molecular weight species corresponds to the intact fusion protein, and the lower one probably represents a degradation product.

```
AT-BP1   FRERKLGGNVLKERGKGEIHG..DKDLGSKQTEPIRIKIFEGGYKSNEDYVYVRGRGRGKYI
AT-BP2   .........FSNKDDSEINSEQDKENSLIKNEPRRIKIFDGGYKSNEDYVYVRGRGRGKYI

         CEECGIRCKKPSMLKKHIRTHTDVRPYVCKLCNFAFKTKGNLTKHMKSKAHMKKCLELGVSM
         CEECGIRCKKPSMLKKHIRTHTDVRPYHCTYCNFSFKTKGNLTKHMKSKAHSKKCVDLGISV

         TSVDETETEEAENMEDLHKTSEKHSMSGISTDHQFSDAEESDGEDGDDNDEDDEDDDDFDDQG
         GLIDEQDTEESDEKQRF..GCER...SGF.......DLEESDGPDEDDND.NEEDDDDSQAES

         DLTPKTRSRSTSPQPPRFSSLPVNVGAVAHGVPSDSSLG..HSSLISYLVTLPSIQVTQLMTP
         GLSA.APSVTASPQHLPSRSGLQDPGSVEEDLRVSSCFSGVHTDPMDILPRALLTKMTVLSTV

         SDSCEDTQMTEYQRLFQSKSTDSEPDKDRLDIPSSMDEEAMLSSEPSSSPRDFSPSSYRSSPG
         QIDPNRTDLTAKAR........QNTGKDELELAPPVDT.............PISPEVTPRSPG

         YDSSPCRDNSPKRYLIPKGDLSPRRHLSPRRDLSPMRHLSPRKEAALRREMSQGDASPRRHLS
         HQMSV.....................HYSESDAL.........RSPAA.......GNTVASIQDS

         PRRPLSPGKDITTRRDLSPRRERRYMTTIRAPSPRRALYHNPPLPMGQYLQTEPIVLGPPNLR
         PSVGLPPA....TIAQLNPQPAARISSSV.............SPHPESQ.EPKQQITLQPS...

         RGLPQVPYFSLYGDQEGAYEHHGSSLFPEGPTDYVFSHLPLHSQQQVRAPIPMVPVGGIQMVH
         PGLP........................SPQTHLFSHLPLHSQQQSRTPYNMVPVGGIHVVP

         S.......LP..........PAISGLHPPPTLPLPT....EGSEEKKGAPGEALTQDPY.TLS
         AGLTYSTFVPIQAGPMQLTIPAVSVIHRTVGTPVDTITEVSGATNRPTGVAELSSVVPCIPIG

         RRH....EKQAPHVLQS..........SGLPSSPSSPRL....LMKQSTSEDSLNSTEREQE
         QIHVPGLQNLSPPALQSLSSLGMETVNIVGLANATVGPQVHPPGLALNAVGLQVLANAPAQSS

         ENIQTCTKAIASLRIA.................TEEAAL..............LGADQPTWV
         PATQTHIPGLQILNIALPTLIPSVGPVAVGTAGTPETATSNSKTVEPQTPAGQGHSAEPPQRL

         QESPQ.............KPLESAHISIRHFSGPE...PGQLCTSAAHPDLHDGEKDTFGTS
         PEGPQETPQTVPGPSVDHARPDGSSKMDTKKVASASHVLPGRSSAQAQPAPIPEALQKIVASG

         QTAVAHPTFYSKGSVDEKQVDFQSSKELSLSTEEGNEPSSEKNRLH*... (708 aa)
         PTSLPTDRPAPRPPVPHRQPIVHFSD...VSSDDDED......RLVIAT*  (671 aa)
```

Figure 2. Comparison of predicted amino acid sequences of AT-BP1 and AT-BP2. The nucleotide sequence data will appear in the EMBL nucleotide sequence database under the accession numbers X54249 (AT-BP1) and X54250 (AT-BP2). The amino acid sequences predicted from the longest reading frames of AT-BP1 and AT-BP2 are shown. Identical residues between the two proteins are indicated by lines, and conserved residues by dots between the two sequences. The cysteine and histidine amino acids located in the zinc fingers are in bold type, and the zinc fingers, acidic region, SPKK repeats, and conserved 'linker' region are underlined. The overall identity between the two proteins is 40% (59% conserved).

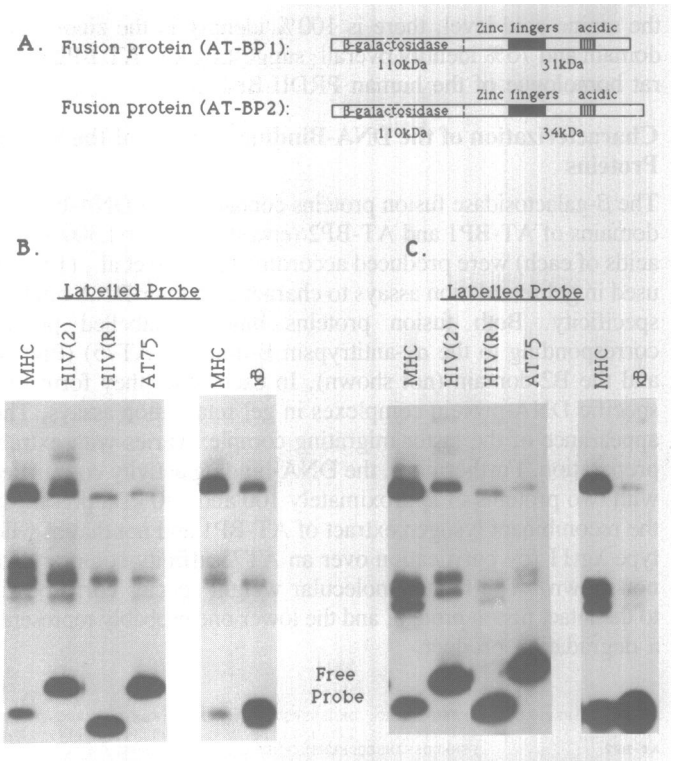

Figure 3. Gel retardation assays of the β-galactosidase fusion proteins of AT-BP1 and AT-BP2. The structure of the λgt11-encoded fusion proteins is shown in **A.** Extracts prepared from induced cultures of λgt11 lysogens were incubated with 20,000 cpm of the indicated probes (sequences given in methods section) for 30 min on ice, and resolved on a 6% polyacrylamide gel. The results for the AT-BP1 fusion protein are shown in **B.**, and for AT-BP2 fusion protein in **C.** Neither incubation of each fusion protein with AR-2 or MB oligonucleotides, nor incubation of the extract prepared from wild type λgt11 lysogen with any of the probes, resulted in formation of a retarded complex (data not shown).

Since the AT-BP2 sequence is highly homologous to the PRDII-BF1 factor shown to bind to the H-2^k enhancer (MHC) site (15), it was of interest to see whether the fusion proteins of AT-BP1 and AT-BP2 would also bind to the MHC or other sites known to be binding sites for NF-κB and H2-TF1 (Table 1; see also ref. 16 for a review). This was tested directly by gel retardation assays with labelled oligonucleotide probes corresponding to the MHC site, both HIV enhancer sites (HIV(2)), the right HIV enhancer site (HIV(R)) and the κ light chain enhancer site (κB). As shown in Fig. 3, the MHC probe is bound with much higher affinity by both fusion proteins than is the AT75 probe originally used in the isolation of the AT-BP1 protein. The affinity of binding decreases in the order MHC > HIV(2) > κB > HIV (R) = AT75.

In order to further analyze the binding specificity of each fusion protein and to determine if there is any quantitative difference in the affinity with which the two fusion proteins bind to the same probe, a competition analysis was performed (Fig. 4). Both proteins appear to have identical ranges of binding specificity as far as can be determined by this assay, which is perhaps not surprising as they share 47/51 amino acids in their zinc finger domains, and three of the divergent amino acids are found in the C-C loops not thought to contact DNA (26). However, it is important to note that the fusion proteins studied here contain only about 300 amino acids of each protein, and sequences outside this region may be important in modulation of DNA-binding characteristics.

Table 1. Binding sites for NF-κB and related proteins.

Gene Promoter/ Enhancer	Sequence
MHC1	TGGGGATTCCCCA
β2-microglobulin	AAGGGACTTTCCC
κ immunoglobulin	AGGGGACTTTCCG
SV40	TGGGGACTTTCCA
HIV (Left)	AAGGGACTTTCCG
HIV (Right)	TGGGGACTTTCCA
β1-interferon	GTGGGAAATTCCT
IL-2 receptor α	AGGGGAATCTCCC
α1-antitrypsin (proposed)	TGGGGATTCACCA
	14 nucleotides

Related sequence elements found in different gene promoter/enhancer elements which bind NF-κB and related DNA-binding proteins (see Baldwin et al., 1990). A sequence located in the AT75 probe shows a 12/13 match with the MHC element if a looping-out of 14 nucleotides is allowed (bottom of table).

The DNA-protein complexes formed with the MHC probe could be almost completely abolished by preincubation of the fusion proteins with 1 pmole of competitor MHC oligonucleotide. Approximately 20-fold more HIV(2) or κB oligonucleotide was required to produce the same level of competition. The oligonucleotides ApoB and AT75 were approximately 2-fold less efficient at competing for complex formation. These last two oligonucleotides have been shown to bind the transcription factors LF-A1 and LF-B1 (8); however other oligonucleotides which specifically bind these factors (AR-2 and MB respectively) do not compete for binding of the fusion proteins to an MHC probe when tested at the same concentration (data not shown). Since the fusion proteins bind with high affinity to NF-kB recognition sites, and in particular to the MHC site, we have examined the AT75 probe for a related DNA sequence. It is possible to propose a looping-out of 14 nucleotides in the AT75 probe which would create a 12/13 match with the MHC probe (Table 1). It is interesting to note that the HIV probe containing both NF-κB binding sites (HIV(2)) is bound much more strongly than when only the right site is used (HIV(R)); this may be due to co-operative binding or else a difference in binding affinity to each site.

AT-BP1 and AT-BP2 mRNAs are Ubiquitously Expressed

The tissue specificity and size of expressed mRNA was determined for each clone. Portions of the cDNAs that were outside the zinc finger-encoding region were used for hybridization with rat poly(A)-tailed RNA in Northern blot analysis (Fig. 5). The probe for AT-BP1 hybridizes to a band of 9.5kb, although other lower molecular weight species (or degradation products) could be detected on longer exposure. The probe for AT-BP2 hybridizes to a single band of 9.0kb and an additional band of 8.0kb which is only present in testes.

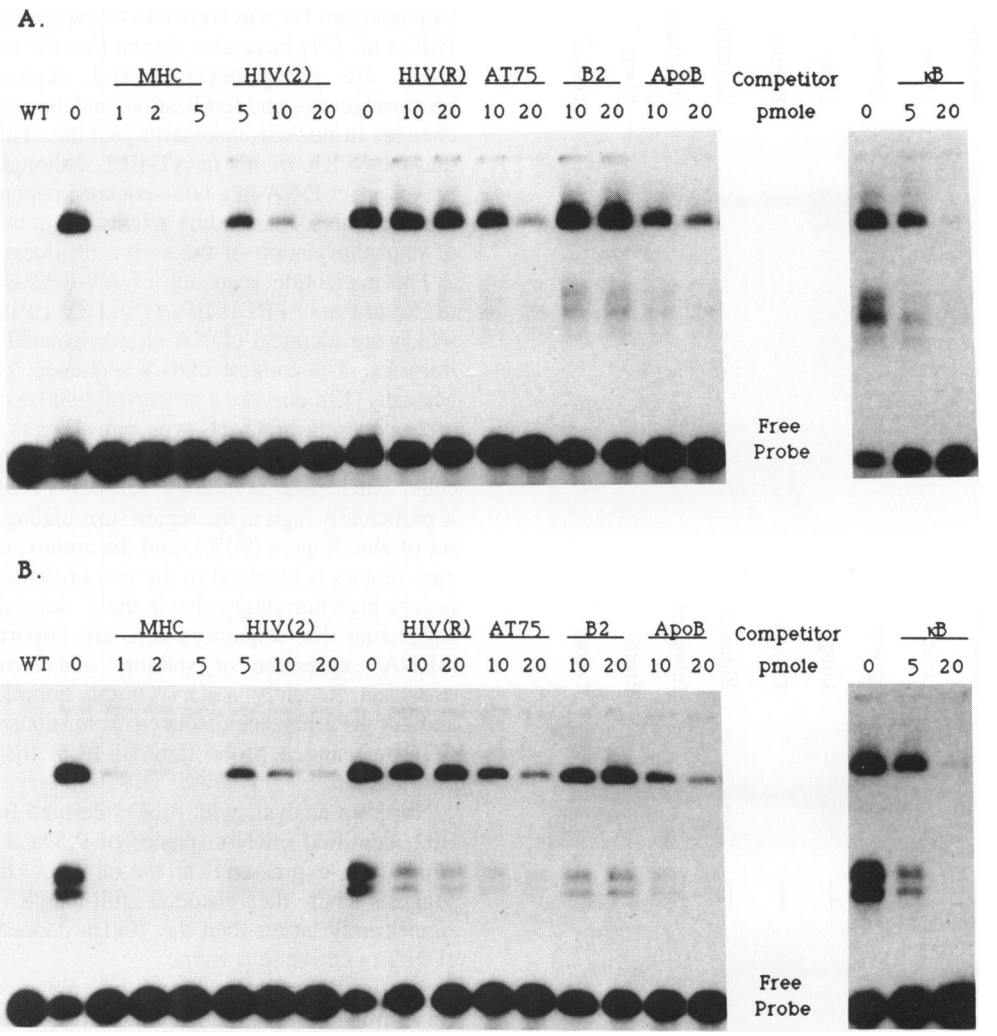

Figure 4. Competition analysis of the β-galactosidase fusion protein of AT-BP1 and AT-BP2. The structure and expression of the fusion proteins is the same as in Fig. 3. In this experiment, the indicated amounts of unlabelled competitor oligonucleotides were preincubated with the fusion protein of AT-BP1 (**A.**) or AT-BP2 (**B.**) for 10 min on ice before addition of 20,000 cpm of labelled MHC probe. The oligonucleotides, AR-2 and MB, when used at 20 pmole do not compete for complex formation (data not shown). In the first lane, extract prepared from wild type λgt11 lysogen was used in the assay.

The mRNAs encoding both proteins are expressed in all the tissues examined (Fig. 5). The levels in thyroid and in liver are low, taking into account the variation in RNA loading as detected by hybridization of the filter with the housekeeping GAPDH gene probe (Fig. 5). In brain tissue, AT-BP1 mRNA is more highly expressed than AT-BP2 mRNA.

DISCUSSION

The isolation and characterization of two partial cDNAs encoding DNA-binding proteins has been described. The proteins, AT-BP1 and AT-BP2 each contain two zinc fingers in tandem, followed by a stretch of acidic amino acids (Fig. 1). The zinc fingers are of the Cys_2-His_2 type (24) whose 3-dimensional structure has been determined (26). The consensus of the two repeats is *Tyr*-X-*Cys*-X_2-*Cys*-X_4-Lys-X_4-*Leu*-X_2-*His*-X_{3-5}-*His*, where the underlined residues are highly conserved in other zinc finger proteins (24). The acidic amino acid region (15/22 residues are aspartic or glutamic acid) may be involved in transcriptional activation (4). There is an additional region of high homology between the two proteins near the middle of the coding sequence (20/24 identical amino acids), but the significance of this is unclear; for example this region is not likely to form an α-helix because of the presence of three conserved proline residues. In the 3′ noncoding regions of both cDNAs, the corresponding mRNA sequence is very AU-rich, and contains one (AT-BP2) or two (AT-BP1) repeats of an AUUUA-motif previously shown to be important in destabilising the mRNAs of transiently expressed genes (27).

Several transcription factors are known to bind to the promoter of the α_1-antitrypsin gene: LF-A1, LF-B1 and LF-B2 (7,8). AT-BP1 and AT-BP2 are clearly distinct from LF-A1 and LF-B1. In addition, it is unlikely that these two proteins are LF-B2 since they bind to the B-domain sequence, containing both LF-B1 and LF-B2 binding sites, more strongly than to a sequence containing only the LF-B2 recognition site. The possibility remains that the full-length AT-BP1 or AT-BP2 protein might act as a negative regulator of α_1-antitrypsin expression by interfering with LF-B1 binding to the gene promoter.

Both AT-BP1 and AT-BP2 fusion proteins form complexes

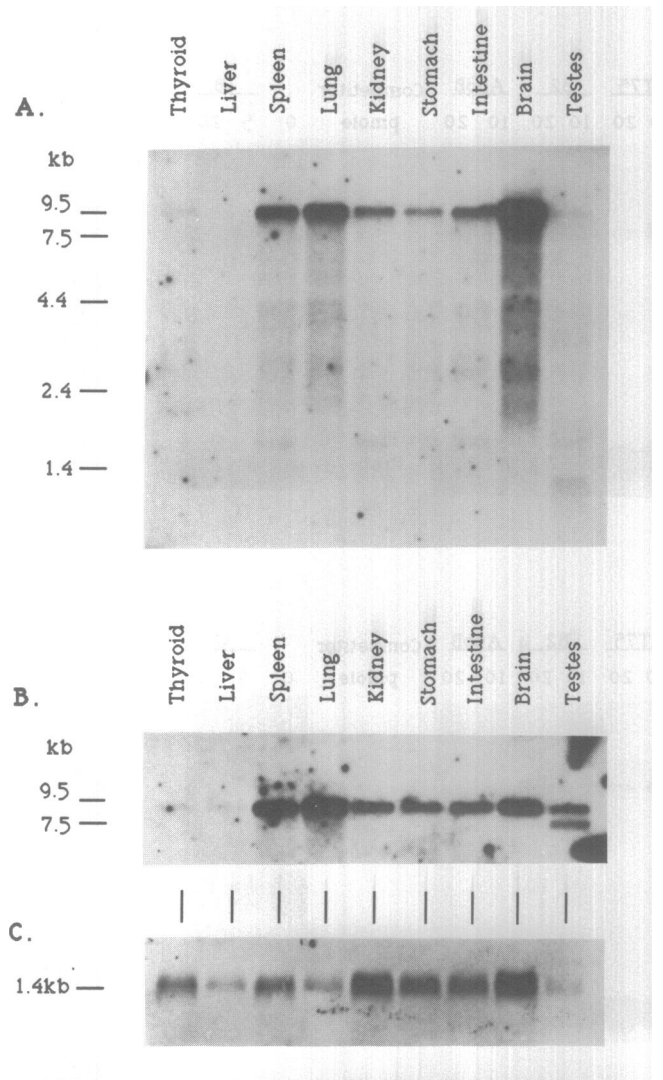

Figure 5. Northern blot analysis of poly(A)-tailed RNA (3μg/lane) from various rat tissues. The blot was hybridized with the [^{32}P]-labelled 1.9kb restriction fragments of **A.** AT-BP1 digested with BglII/EcoRI, or **B.** AT-BP2 digested with BamHI/EcoRI (Paonessa et al., 1988). In **C.**, the same blot was probed with the housekeeping GAPDH gene using the Amersham nonradioactive DNA labelling and detection kit. The migration of RNA size markers is indicated.

with the MHC probe which are competed 20–40-fold less efficiently by the other oligonucleotides tested as compared to competition with the MHC oligonucleotide. This site in the major histocompatibility complex class I promoter is important in gene activation and has previously been shown to be capable of binding at least two distinct transcription factors, H2TF1 and NF-κB (28). Since the factor H2TF1 also binds to the MHC site with higher affinity than to the immunoglobulin κ enhancer site, the possibility exists that either AT-BP1 or AT-BP2 might be identical to H2TF1. The DNA-binding subunit of NF-κB has recently been cloned and does not contain zinc fingers or any other previously recognised DNA-binding motif (29,30).

Although AT-BP1 is highly homologous to AT-BP2, it contains a second putative DNA-binding domain in addition to the zinc finger motif, consisting of an 8-fold repeat of Ser-Pro-Lys/Arg-Lys/Arg (SPKK). In sea urchin sperm-specific histones H1 and H2B, there are 4–6 SPKK repeats in the N-terminal region, which are spaced 0–3 residues apart. A 6-times repeat of SPKK isolated from H1 was shown to be capable of DNA binding (25). Hill et al. (31) have also shown that the serine residues in this motif are phosphorylated and dephosphorylated during spermatogenesis and fertilization, and these changes correlate with changes in nuclear chromatin packing. Thus one can speculate that the SPKK motifs in AT-BP1, although spaced irregularly, may contact DNA in a non-sequence dependent manner, and in addition, this DNA-binding activity might be regulated at the level of phosphorylation of the serine residues.

The nucleotide sequence of AT-BP2 is highly homologous (82%) to that of PRDII-BF1 (15), HIV-EP1 (13) and MBP-1 (14) which are identical cDNA clones isolated from human cDNA libraries. The longest cDNA sequence, reported by Fan and Maniatis (15), encodes a protein of 298kDa containing two widely separated sets of C_2-H_2-type zinc fingers. The corresponding mRNA is 9.5kb and can be induced by serum and virus in MG63 cells. Nucleotide homology between PRDII-BF1 and AT-BP2 is particularly high in the region surrounding the carboxy-terminal set of zinc fingers (90%), and the amino acid sequence of these zinc fingers is identical in the two proteins. Interestingly, there is very high homology also in the 3′ noncoding sequence (93%) suggesting that sequences here are important in regulation of mRNA expression or stability and have been retained in evolution. Recently, a cDNA highly homologous to PRDII-BF1 and AT-BP2 has been isolated from mouse lens epithelial cells (32) by using a probe derived from the αA-crystallin gene promoter.

Northern analysis with probes derived from AT-BP1 or AT-BP2 identified mRNA species of 9.5 and 9.0 kb respectively which were expressed in all the rat tissues that were tested. This suggests that the encoded full-length proteins might be significantly larger than the 70kDa deduced from their partial cDNA open reading frame.

Finally, the isolation of two homologous proteins containing very similar DNA-binding domains raises the possibility that they may be members of a gene family. However, by Southern blot analysis using restriction digests of DNA from human, mouse or rat genomes, Singh et al. (12) showed that the cDNA for MBP-1 was probably a single copy gene. This conclusion is surprising in light of our results, and more detailed analysis is required. It will be interesting to determine whether there are further members of this gene family.

ACKNOWLEDGMENTS

We are grateful to Iain Mattaj, Mattias Hentze, Dipak Ramji, Angus Lamond and Ben Blencowe for critical reading of the manuscript. We thank Dipak Ramji for useful discussions, and Monique Frain and Petra Stevenson for technical assistance. We are also grateful to P. Neuner and S. Weston for oligonucleotide synthesis and to the photography lab for preparation of photographs. We thank T. Maniatis and P. Sharp and their colleagues for sharing their unpublished data on PRDII-BF1 and MBP-1.

REFERENCES

1. Struhl,K. (1989) Trends Biochem. Sci. **14**, 137–140.
2. Mitchell,P.J., and Tijan,R. (1989) Science **245**, 371–378.
3. Murre,C., McCaw,P.S., and Baltimore,D. (1989) Cell **56**, 777–783.
4. Ptashne,M. (1988) Nature **335**, 683–689.
5. Lewin,B. (1990) Cell **61**, 1161–1164.
6. Ptashne,M., and Gann,A.A.F. (1990) Nature **346**, 329–331.

7. De Simone,V., Ciliberto,G., Hardon,E., Paonessa,G., Palla,F., Lundberg,L., and Cortese,R. (1987) EMBO J. **6**, 2759–2766.
8. Monaci,P., Nicosia,A., and Cortese,R. (1988) EMBO J. **7**, 2075–2087.
9. Baumhueter,S., Mendel,D.B., Conley,P.B., Kuo,C.J., Turk,C., Graves, M.K., Edwards,C.A., Courtois,G, and Crabtree,G.R. (1990) Genes Dev. **4**, 372–379.
10. Frain,M., Swart,G., Monaci,P., Nicosia,A., Stampfli,S., Frank,R., and Cortese,R. (1989) Cell **59**, 145–157.
11. Nicosia,A., Monaci,P., Tomei,L., De Francesco,R., Nuzzo,M., Stunnenberg,H., and Cortese,R. (1990) Cell **61**, 1225–1236.
12. Singh,H., LeBowitz,J.H., Baldwin,A.S.,Jr., and Sharp,P.A. (1988) Cell **52**, 415–423.
13. Maekawa,T., Sakura,H., Sudo,T., and Ishii,S. (1989) J. Biol. Chem. **264**, 14591–14593.
14. Baldwin,A.S.,Jr., LeClair,K.P., Singh,H., and Sharp,P.A. (1990) Mol. Cell. Biol. **10**, 1406–1414.
15. Fan,C-M., and Maniatis,T. (1990) Genes Dev. **4**, 29–42.
16. Lenardo,M.J., and Baltimore,D. (1989) Cell **58**, 227–229.
17. Feinberg,A.P., and Volgelstein,B. (1983) Anal. Biochem. **132**, 6–13.
18. Manfioletti,G., and Schneider,C. (1988) Nucl. Acids Res. **16**, 2873–2884.
19. Maniatis,T., Fritsch,E.F., and Sambrook,J. (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor University Press, Cold Spring Harbor.
20. Sanger,F., Nicklen,S., and Coulson,A.R. (1977) Proc. Natl. Acad. Sci. USA **74**, 5463–5467.
21. Huynh,T.V., Young,R.A., and Davis,R.W. (1988) In Glover,D.M. (ed.), DNA Cloning-A Practical Approach. IRL Press, Oxford, Vol.I, pp. 49–78.
22. Laemmli,U.K. (1970) Nature **227**, 680–685.
23. Paonessa,G., Gounari,F., Frank,R., and Cortese,R. (1988) EMBO J. **7**, 3115–3123.
24. Klug,A., and Rhodes,D. (1987) Trends Biochem. Sci. **12**, 464–469.
25. Suzuki,M. (1989) EMBO J. **8**, 797–804.
26. Lee,M.S., Gippert,G.P., Soman,K.V., Case,D.A., and Wright,P.E. (1989) Science **245**, 635–637.
27. Shaw,G., and Kamen,R. (1986) Cell **46**, 659–667.
28. Baldwin,A.S.,Jr., and Sharp,P.A. (1988) Proc. Natl. Acad. Sci. USA **85**, 723–727.
29. Kieran,M., Blank,V., Logeat,F., Vandekerckhove,J., Lottspeich,F., Le Bail,O., Urban,M.B., Kourilsky,P., Baeuerle,P.A., and Israël,A. (1990) Cell **62**, 1007–1018.
30. Ghosh,S., Gifford,A.M., Riviere,L.R., Tempst,P., Nolan,G.P., and Baltimore,D. (1990) Cell **62**, 1019–1029.
31. Hill,C.S., Packman,L.C., and Thomas,J.O. (1990) EMBO J. **9**, 805–813.
32. Nakamura,T., Donovan,D.M., Hamada,K., Sax,C. M., Norman,B., Flanagan,F.R., Ozato,K., Westphal,H., and Piatigorsky,J. (1990) Mol. Cell. Biol. **10**, 3700–3708.