USF binds to the APB α sequence in the promoter of the amyloid β -protein precursor gene

Alexander A. Vostrov, Wolfgang W. Quitschke, Frédérique Vidal¹, Alexander L. Schwarzman and Dmitry Goldgaber*

Department of Psychiatry and Behavioral Science, State University of New York at Stony Brook, Stony Brook, NY 11794-8101, USA and ¹Unité 273 de l'INSERM, Université de Nice–Sophia Antipolis, 06034 Nice, France

Received March 16, 1995; Revised and Accepted May 31, 1995

ABSTRACT

The APB α domain in the amyloid β -protein precursor (APP) promoter contains a nuclear factor binding domain with the core recognition sequence TCAGCTGAC. Proteins in nuclear extracts from brain and numerous cell lines bind to this domain and it contributes ~10–30% to the basal APP promoter activity. Included in this domain is the CANNTG motif, which is recognized by basic helix–loop–helix transcription factors. The same motif is also present in the CDEI element of the yeast centromere and in the adenovirus major late promoter (AdMLP). Here we present evidence based on thermostability, relative binding affinity, electrophoretic mobility and antibody recognition that the cellular proteins that bind to the APB α and CDEI motifs are USF. However, the relative binding affinity for the motifs is different. The affinity of USF for AdMLP is ~20-fold higher than for the APB α sequence and 5-fold higher than for the CDEI sequence. Mutational analysis suggested that the primary determinant for USF binding affinity resides within the octamer CAGCTGAC, which is composed of the E-box consensus sequence CANNTG followed by the dinucleotide AC. The human homolog of the mouse CDEI binding protein did not bind to either the CDEI sequence or APB α .

INTRODUCTION

Deposition of aggregated amyloid β -protein in the brain and cerebrovasculature are characteristic neuropathological features of Alzheimer's disease and Down's syndrome (1–5). Amyloid β -protein is derived from a larger transmembrane glycoprotein, the amyloid β -protein precursor (APP) (6–9). The APP gene is expressed in all major tissues, including brain (10–13). The level of APP gene transcript is increased in Down's syndrome and in certain areas of the brain in Alzheimer's disease (10–13). This suggests that in some cases overexpression of the APP gene could play a contributing role in the pathological processes leading to amyloid deposition.

The mechanism of APP gene expression has been the subject of extensive study (14–23). The APP promoter contains numerous putative binding sites for regulatory transcription factors (14,15,17,24). Functional analyses of 5' deletions of the APP promoter showed that 94 bp upstream from the transcriptional start site are sufficient for high levels of expression in numerous cell lines (17,18,22). This proximal APP promoter region contains two nuclear factor binding domains, designated APB α and APB β , that are conserved in the human, mouse and rat APP promoters (17,23). The factors that bind to these domains display sequence specificity, but have not yet been identified.

Binding site APB α contains the core recognition sequence TCAGCTGAC. Transverse block mutations across this sequence abolish nuclear factor binding and eliminate the functional activity mediated by the APB α binding site (17,23). The APB α domain includes the conserved motif CANNTG, which is recognized by the basic helix–loop–helix (bHLH)-containing family of DNA binding proteins (25). The APB α sequence also contains parts of the overlapping recognition sequences for transcription factors AP1 and AP4. However, the factors in nuclear extracts that bind to this element are distinct from both AP1 and AP4 (17). Furthermore, the core part of the APB α binding site bears a similarity to the GTCACATG motif. This octamer sequence was first characterized in the centromere DNA element I (CDEI) region of the yeast centromere (26).

The CDEI motif is also recognized by mouse nuclear proteins and microinjection of double-stranded oligonucleotides carrying this motif into fertilized mouse eggs resulted in detrimental effects on the early development of the embryos (27). A partial cDNA sequence encoding a candidate mouse CDEI binding protein has been identified (28). The protein, designated CDEBP, contains regions with extensive homology to APP. The human homolog of this protein was subsequently cloned by two groups of investigators and termed APPH (29) and APLP2 (30).

Bram and Kornberg (31) have pointed out the similarity of the CDEI motif to the sequence in the adenovirus major late promoter (AdMLP), which is recognized by the mammalian transcription factor USF (32). The recognition sequence for USF in AdMLP contains the central motif GCCACGTGAC. However, other binding sites for USF identified in a variety of cellular and viral promoters (33–38; for further references see 39) often differ in

* To whom correspondence should be addressed

Table 1. Examples of sequences that bind USF

adenovirus major late promoter [32]	CTGGC CACGTGAC CGC
rat γ -fibrinogen promoter [33]	GAC CCCGTGAC C
mouse metallothionein I promoter [34]	GCGGGG CGCGTGAC TAT
human β -globin LCR [35,36]	GCTGAC CACCTGAC TAAAC
<i>X. laevis</i> transcription factor IIIA promoter [37]	CAT CACGTGCT CCACTA
rat class I alcohol dehydrogenase promoter [38]	CTAGAT CACATGTG GGATC
amyloid β -protein precursor promoter [17]	CCGGAT CAGCTGAC TGCCT
CDEI [28]	TGTGC CATGTGAC TGAGAAGCT

The sequences are aligned with the 8 bp domain CANNTGAC, which is significant for USF binding affinity. The original references describing the binding of USF to the respective sequences are indicated. The sequences of the human APB α domain in the APP promoter and the CDEI sequences are included for comparison.

some nucleotides from the original USF binding site in AdMLP (Table 1).

Two thermostable proteins with molecular weights of 43 (USF43) and 44 kDa (USF44) which display USF activity were co-purified from HeLa cell nuclear extracts (40) and subsequently cloned from human and mouse cells (39,41,42). Both USF proteins are members of the bHLH protein family and bind to DNA as homodimers or heterodimers (42). Here we present evidence that the mammalian nuclear proteins that bind to both the APB α site in the APP promoter and the CDEI motif of the yeast centromere are USF.

MATERIALS AND METHODS

Plasmids

Plasmids dI2, containing human USF43 cDNA (41), and pM2-2, containing mouse USF44 cDNA (39), were kindly provided by Dr M. Sawadogo. A plasmid containing the open reading frame for the mouse CDEBP protein was described by Vidal *et al.* (28). Plasmid DNA was purified using the Quiagen Plasmid Maxi Kit according to the manufacturer's instructions.

cDNA library screening and cloning

The plasmid containing the main part of the open reading frame for the mouse CDEBP (28) was digested with the restriction enzyme *Nco*I. A 1.5 kb fragment was isolated and used for screening of a human brain cDNA library. Thirty three positive clones were isolated. To obtain the full-length cDNA, filters were rescreened with the 0.3 kb *Sph*I–*Dra*I restriction fragment from the same plasmid, which contains the 5'-region of the open reading frame. Sixteen of the previously isolated clones gave a positive signal for the 5'-region and their cDNA inserts were subcloned and sequenced.

Oligonucleotides

Oligonucleotides were synthesized by Bio-synthesis Inc. and gel purified prior to hybridization. The sequences of all double-stranded oligonucleotides used as probes or as competitors are shown in Table 2. Oligonucleotide sequences of APB α , CDEI and AdMLP were adapted exactly as described in the original publications (17,27,32). APM1–APM4 are mutant versions of APB α .

Table 2. Sequences of the oligonucleotides used in this study

AATTCCTGGC	C A C G T G A C	CGCAGCTGT	-AdMLP
TGTGC	C A T G T G A C	TGAGAAGCT	-CDEI
GGGCCGGAT	C A G C T G A C	TCGCCTGGCTCT	-APB α
GGGCCGGAT	C A <u>C</u> G T G A C	TCGCCTGGCTCT	-APM1
GGGCCGTGC	C A <u>I</u> G T G A C	TCGCCTGGCTCT	-APM2
GGGCCGGAT	C A <u>C</u> <u>A</u> T G A C	TCGCCTGGCTCT	-APM3
GGGCCGGAT	C A G C T G <u>C</u> A	TCGCCTGGCTCT	-APM4

The sequence domain with particular relevance for USF binding affinity is framed. The oligonucleotides APM1–APM4 are modifications of APB α . The nucleotides that differ from the wild-type APB α are underlined.

Nuclear extracts

Nuclear extracts from HeLa cells were prepared as described elsewhere (43). Nuclear extracts from rat and bovine brain were obtained by a modified procedure (17). The final concentration of protein in extracts varied from 1 to 4 mg/ml in a buffer containing 25 mM HEPES, pH 7.6, 40 mM KCl, 12.5 mM MgCl₂, 0.1 mM EDTA, 1 mM dithiothreitol, 0.5 mM NaS₂O₃, 0.1 mM phenylmethylsulfonyl fluoride and 10% glycerol. Extract preparations were aliquoted and stored at –80°C. Heat treatment of brain nuclear extract was performed at 100°C for 10 min. Nuclear extracts were then chilled on ice and the precipitated material was pelleted at 10 000 g. The supernatant was used for mobility shift electrophoresis.

Coupled *in vitro* transcription–translation

The coupled *in vitro* transcription–translation reaction was performed using the TNT Coupled Reticulocyte Lysate System (Promega) according to the manufacturer's instructions. Typically 2 μ g template plasmid DNA(s) were used per 25 μ l reaction mixture containing [³⁵S]methionine. Aliquots (5 μ l) of the reaction products were analyzed by electrophoresis in 10% SDS–polyacrylamide gels. Gels were fixed, dried and autoradiographed.

Electrophoretic mobility shift assay

Oligonucleotides were 5'-end-labeled with γ -³²P using T4 polynucleotide kinase. Labeled oligonucleotide (1 ng per binding reaction, 50 000–200 000 c.p.m.) was incubated in binding buffer [15 mM Tris–HCl, pH 7.5, 75 mM NaCl, 0.1 mM EDTA, 1 mM dithiothreitol, supplemented with 1 μ g poly(dI–dC)–poly(dI–dC)] for 1 h at room temperature either with 5 μ l nuclear extract containing 2–8 μ g protein or with 4 μ l *in vitro* translation products in a total reaction volume of 15 μ l. With the protein extract the final MgCl₂ concentration in the binding reaction was ~4 mM. The incubation mixture was loaded on 7% polyacrylamide gels containing 0.5 $\times$ Tris–borate–EDTA buffer and electrophoresed at 150 V constant voltage for 2–6 h. Gels were fixed in 10% acetic acid, 20% methanol and 5% glycerol, dried and autoradiographed. To quantitate the binding activity, areas containing binding complexes were excised from dried gels and the radioactivity was determined by liquid scintillation counting.

Supershift assays

Rabbit anti-human USF43 antibodies were a gift from Dr M. Sawadogo. These antibodies were raised against the C-terminal portion of the 43 kDa form of human USF, overproduced in *Escherichia coli* (42). After incubating oligonucleotides with bovine brain nuclear extract as described above, antibodies against USF43 were added to the reaction mixtures and incubated for an additional hour at 4°C. The products were then analyzed by mobility shift electrophoresis.

RESULTS

Proteins in nuclear extracts from HeLa cells, bovine brain and rat brain form similar binding complexes with AdMLP, CDEI and APB α motifs

The binding domain APB α contains the E-box consensus sequence and closely resembles the USF binding site of AdMLP (Table 2). It also bears a similarity to the CDEI-like sequence (Table 1). The APB α binding site was originally characterized by mobility shift electrophoresis in 1% agarose gels. This resulted in the identification of an apparently homogeneous binding complex (17,23). The specificity of this binding was demonstrated with nuclear extracts from a variety of cell lines and with mutations that abolished binding activity. To improve the resolution of the assay, we compared the mobility shift patterns in 6% acrylamide gels. The three end-labeled oligonucleotides containing the E-box consensus sequences of AdMLP, CDEI and human APB α (Table 1) were incubated with nuclear extracts from bovine brain, rat brain and HeLa cells. Mobility shift electrophoresis in acrylamide gels reveals multiple distinct binding complexes (Fig. 1). A major binding complex, designated C α , was observed with all nuclear extracts and oligonucleotides. This complex is surrounded by two minor complexes which migrate slightly above (C α 1) and below (C α 2) the major complex C α (Fig. 1). Additional binding complexes migrate below complex C α (Fig. 2A, arrowheads). However, most likely these bands represent degradation products, since their prominence varies between different extract preparations. In addition, the intensity of these bands was increased when the nuclear extract was frozen and thawed multiple times or left at room temperature for extended periods (not shown).

Since USF has been demonstrated to bind to the AdMLP sequence it is possible that this factor also binds to the CDEI-like and APB α sequences. It has been shown that the USF proteins retain their DNA binding ability after heat treatment (40). We therefore examined whether the binding proteins in bovine brain nuclear extract share this property with the USF proteins. The nuclear extract was heated at 100°C for 10 min and then used for mobility shift electrophoresis. We found that 50–70% of the binding activity remained in the extract after heat treatment and the mobility shift pattern did not change significantly. However, the complexes C α 1 and C α 2 became more pronounced after heat treatment, with a corresponding decrease in complex C α (Fig. 1A–C, compare lanes 3 and 4). Similar results were obtained with HeLa cell nuclear extract (not shown). This indicates that the nuclear binding proteins in the analyzed extracts are resistant to heat treatment, a property which they have in common with the USF proteins. The rearrangements of binding complexes C α , C α 1 and C α 2 is consistent with the observation that USF dimers dissociate at elevated temperature and randomly reassociate after cooling (40).

Figure 1. The ^{32}P 5'-end-labeled oligonucleotides (*) AdMLP (A), CDEI (B) and APB α (C) were incubated with nuclear extract from HeLa cells (lane 1), rat brain (lane 2) or bovine brain (lane 3). Binding of heat-treated bovine brain nuclear extract to the respective oligonucleotides is presented in lane 4. The positions of binding complexes C α , C α 1 and C α 2 are shown. The presence of additional complexes observed with all extracts are indicated by arrowheads in (A). The free oligonucleotide is not shown since it was routinely run off the gel in most assays.

Common nuclear binding proteins recognize AdMLP, CDEI and APB α sequences

USF binds to the core sequence GCCACGTGAC of the AdMLP (32). A number of promoter sequences that have been shown to bind USF display a certain degree of variability within this sequence (Table 2). The APB α sequence CAGCTGAC of the APP promoter and the CDEI-like sequence CATGTGAC display the same type of variability within the USF binding domain, suggesting that they may also be recognized by USF. This possibility was examined by performing mobility shift competition assays to determine the relative binding affinity of complex C α for the binding motifs of AdMLP, CDEI and APB α . In these experiments reference to complex C α also includes the surrounding complexes C α 1 and C α 2, which were not separated due to shorter electrophoresis times.

Radiolabeled oligonucleotides APB α , CDEI or AdMLP were mixed with increasing concentrations of unlabeled oligonucleotides and incubated with bovine brain nuclear extract (Fig. 2). The three oligonucleotides compete against each other for nuclear factor binding with different degrees of effectiveness. For example, a 2-fold molar excess of unlabeled AdMLP oligonucleotide competes for binding to labeled APB α in the same manner as a 50-fold molar excess of unlabeled APB α (Fig. 2C, lanes 4 and 5). Similarly, a 50-fold molar excess of unlabeled APB α reduces binding to labeled AdMLP to the same degree as a 2-fold excess of unlabeled AdMLP (Fig. 2A, lanes 4 and 5). The results from all competition experiments suggest that the binding affinity for the AdMLP sequence is ~20-fold higher than for the APB α sequence and 4–5-fold higher than for the CDEI-like sequence (Fig. 2A–C).

In order to estimate the importance of the CANNTGAC octamer motif for nuclear protein binding activity, mutations

Figure 2. The ^{32}P 5'-end-labeled oligonucleotides (*) AdLMP (A), CDEI (B) and APB α (C) were incubated with nuclear extract from bovine brain. Mobility shift electrophoresis was performed either in the absence of competitor (lane 1) or with a 2-, 10- or 50-fold molar excess (2X, 10X, 50X) of unlabeled competitors AdLMP (lanes 2-4), CDEI (lanes 5-7) and APB α (lanes 8-10). In addition, labeled oligonucleotide APB α (C) was also competed with unlabeled mutations APM1-APM3 at 2- and 10-fold molar excess (lanes 11-16) and mutation APM4 at 10- and 50-fold molar excess (lanes 17 and 18).

were introduced into the APB α sequence and used as competitors for binding to labeled wild-type APB α (Fig. 3C). The mutation APM1 differs from APB α in only two bases and it converts the wild-type APB α octamer sequence CAGCTGAC to CACGTGAT, which is the sequence found in AdMLP (Table 2). This replacement makes mutation APM1 as effective a competitor as AdMLP itself (Fig. 3C, compare lanes 11 and 12 with lanes 5 and 6). The mutation in APM2 reproduces the sequence of the binding domain in CDEI. This causes the binding of nuclear protein to APM2 to be as effective as the binding to CDEI (Fig. 2C, compare lanes 13 and 14 with 8 and 9). The binding affinity for mutation APM3 is in the same range as for CDEI (Fig. 2C, lanes 15 and 16). In mutation APM4 the AC dinucleotide that follows the CAGCTG hexamer was inverted to CA (Table 2). As a consequence, the binding affinity for APM4 is 5-10-fold lower than for the wild-type APB α sequence (Fig. 3C, lanes 17 and 18).

The results show that the affinity for formation of complex C α (including C α 1 and C α 2) is different for the three sequence motifs in AdLMP, CDEI and APB α . Nuclear proteins display the lowest binding affinity for the APB α sequence. However, the affinity for the APB α oligonucleotide can readily be increased 20-fold by converting the wild-type APB α domain to a sequence resembling

Figure 3. Schematic representation of alternatively spliced forms of APLP2 protein. Numbers indicate the amino acid position. The shaded and black boxes correspond to inserts of 54 and 12 amino acids respectively.

AdLMP (Table 2, APM1). Analysis of additional mutations (APM2-APM4) within the APB α domain suggests that the primary determinant for nuclear factor binding activity resides within the core sequence CANNTGAC (Table 2). This sequence consists of the core E-box domain CANNTG followed by the dinucleotide AC.

The human homolog to mouse CDEBP protein consists of alternatively spliced variants

Although the above experiments implicate USF as the primary candidate for formation of complex C α , binding of mouse CDEBP to CDEI-like sequences was demonstrated in experiments with a truncated fusion protein expressed in *E. coli* (28). We therefore examined this binding activity with the full-length protein of the human homolog synthesized *in vitro*. A human brain cDNA library was screened with a fragment containing the 5'-region of mouse CDEBP, yielding a total of 16 positive clones. One of the clones (C5) contained a 3.6 kb insert and appeared to be a full-length transcript encoding a 695 amino acid protein with a high degree of sequence similarity to murine CDEBP. The cloned sequence was found to be identical to that of APPH (29) and APLP2 (30), except for some minor differences in the 3' untranslated region. Here we refer to the cloned gene as APLP2. We also identified two clones, designated C3 and C7, which contain additional sequences within the open reading frame. C3 contains a 168 bp insert at position 927 (Fig. 3, amino acid 309), relative to the translational start site (+1), and C7 contains a 36 bp insert at position 1671 (Fig. 3, amino acid 557). The additional domain in C7 encodes an amino acid sequence with a high degree of similarity to the Kunitz-type protease inhibitor domain present in several alternatively spliced forms of APP. We assume that the APLP2 gene transcript is alternatively spliced and propose four possible different forms of encoded protein (Fig. 3). Such alternatively spliced forms were also reported by Sprecher *et al.* (29) (forms APLP2-2 and APLP2-4) and Wasco *et al.* (30) (forms APLP2-1, APLP2-2 and APLP2-3).

cDNA of form APLP2-1 encodes a protein that bears 92% similarity to the sequence of the mouse CDEBP protein (44). Fifty-seven of a total of 695 amino acids differ between the two sequences. Of these, 16 mismatches are clustered in the negatively charged protein domain (see descriptions of human APLP2 in 29,30), while others are evenly distributed along the sequence. Only 10 amino acids differ within the C-terminal 237 amino acid fragment, which was used for construction of the fusion protein to study CDEI binding activity (28). Specifically, only one Ser→Ala substitution was found in the putative HLH domain (28). Taking into account such a high degree of similarity, one could expect that

Figure 4. [35 S]Methionine-labeled proteins were synthesized by coupled *in vitro* transcription-translation and separated on 8% SDS-polyacrylamide gels. The USF43 and USF44 proteins were either synthesized separately (lanes 4 and 5) or together (lanes 1-3). In cases where the USF proteins were co-translated, the USF43:USF44 plasmid ratio was 1:3 (lane 1), 1:1 (lane 2) or 3:1 (lane 3). Translated APLP2 proteins are shown as indicated by brackets in lanes 7-10. Lane 6 shows a control reaction without added plasmid. The approximate migration of molecular weight markers (kDa) is shown in the center.

the mouse CDEBP and human APLP2 proteins share their putative CDEI binding properties.

From the three cDNA clones four recombinant plasmids were constructed that contained all four postulated alternatively spliced forms of APLP2. The full-length transcriptional units were cloned 5' to the T7 promoter and used in a coupled *in vitro* transcription-translation system.

In vitro translated USF proteins bind to AdMLP, CDEI and APB α sequences

In order to determine if either the USF or the APLP2 proteins contribute to formation of complex C α , these proteins were expressed in a coupled *in vitro* transcription-translation system. The translation reaction was performed in the presence of [35 S]methionine and the products of the transcription-translation reaction were analyzed by SDS-polyacrylamide gel electrophoresis (Fig. 4). USF43 and USF44 were either synthesized separately (Fig. 4, lanes 4 and 5) or together at plasmid ratios of 3:1, 1:1 and 1:3 (Fig. 4, lanes 1-3). The concentration ratios of the final translation products were largely proportional to the ratios of the plasmid concentrations in the reaction mixtures (Fig. 4, lanes 1-3). The translation products of the four alternatively spliced forms of APLP2 are shown in lanes 6-10 (Fig. 4).

The ability of these *in vitro* translated proteins to bind to AdMLP, CDEI and APB α sequences was examined by mobility shift electrophoresis. When labeled AdMLP, CDEI or APB α oligonucleotides were incubated with separately translated USF43 and USF44, two different binding complexes with slightly different mobilities were observed (Fig. 5A-C, lanes 4 and 5). These complexes represent homodimers of USF43 or USF44 that bind to the respective oligonucleotides. Similar binding of USF homodimers to the AdMLP sequence was described by Sirito *et al.* (39). When the co-translated products of USF43 and USF44 were incubated with the labeled oligonucleotides, three different binding complexes were observed. The intermediate band is the result of

Figure 5. *In vitro* translated proteins as depicted in Figure 4 were incubated with labeled oligonucleotides (*) AdMLP (A), CDEI (B) and APB α (C) and analyzed by mobility shift electrophoresis. The homodimer and heterodimer complexes of USF with their respective oligonucleotides are indicated (43, 44 and 43/44). The binding of *in vitro* co-translated USF proteins is shown in lanes 1-3 with USF43:44 ratios of 1:3, 1:1 and 3:1 respectively. The binding of separately translated USF43 and USF44 is shown in lanes 4 and 5 respectively. For comparison, the binding of proteins from bovine brain nuclear extract is displayed in lane 6. As a negative control, translation extract without added plasmid was incubated with labeled oligonucleotides (lane 7). No specific binding is observed with the *in vitro* translated APLP2 proteins (lanes 8-11). Faint bands, which are also present in the control lane, are indicated by brackets.

binding of USF43/44 heterodimers to the oligonucleotide (Fig. 5A-C, lanes 1-3). The ratio of the three complexes to each other is correlated with the ratio of USF43 to USF44 in the reaction mixture. When USF43 and USF44 were expressed at a ratio of 1:3, binding of USF44 homodimers and USF43/44 predominated, whereas homodimers of USF43 were barely detectable (Fig. 5A-C, lane 3). The opposite is true when USF43 and USF44 were expressed at a ratio of 3:1 (Fig. 5A-C, lane 1). When USF43 and USF44 were expressed at equimolar ratios, heterodimer binding predominated (Fig. 5A-C, lane 2). Binding of bovine brain nuclear extract is shown for comparison (Fig. 5A-C, lane 6). The mobility of the

Figure 6. USF43 and USF44 proteins were co-translated at a ratio of 1:1 and incubated with labeled APBα oligonucleotide (APBα*). Mobility shift electrophoresis was performed either in the absence of competitor (lane 1) or after addition of increasing molar excesses (2X, 10X, 50X) of unlabeled APBα (lanes 2–4), AdLMP (lanes 5–7), CDEI (lanes 8–10) or mutations APM1–APM4 (lanes 11–18). The positions of the complexes formed by the USF43/44 heterodimer and the respective homodimers are indicated (43/44, 43 and 44).

complex formed with bovine brain nuclear extract is very similar to the mobilities of the complexes formed by the *in vitro* translated USF proteins. However, no specific binding of any of the APLP2 proteins to either AdLMP, CDEI or APBα was observed. Faint bands indicated by brackets are probably non-specific, since they are also detected in the control experiment (Fig. 5A–C, lanes 7–11).

USF proteins bind to AdLMP, CDEI and APBα with the same relative affinities as the binding proteins in bovine brain nuclear extract

The competition experiments presented in Figure 2 suggested that the factors in crude nuclear extract display different binding affinities for the USF consensus sequences in the AdMLP, CDEI and APBα oligonucleotides. To determine the relative binding affinities of the *in vitro* translated USF, we performed mobility shift competition assays. The 5'-end-labeled APBα oligonucleotide was incubated with *in vitro* co-translated USF43 and USF44 proteins (ratio 1:1) (Fig. 6). Mobility shift electrophoresis resulted in separation of the three binding complexes formed by the USF43/44 heterodimer and the respective homodimers, as described in Figure 4. The binding of *in vitro* translated USF to APBα was competed with unlabeled APBα, AdLMP, CDEI and APBα mutations APM1–APM4 at the same molar ratios as described for the nuclear proteins in Figure 2C. The binding affinity of USF for the AdLMP sequence was ~20-fold higher than for the APBα sequence and 5-fold higher than for the CDEI sequence (Fig. 6). Similarly, the level of competition of mutations APM1–APM4 was in the same range as described for the bovine brain nuclear factors in Figure 2C. Furthermore, for each competitor the level of competition was the same for the USF homodimers and the heterodimer. Hence, all three forms of USF dimers display the same relative binding affinity for the analyzed consensus sequences.

APBα binding proteins from bovine brain nuclear extracts are recognized by antibodies against USF43

The experiments described above indicate that the properties of the USF proteins are most similar to the properties of the binding proteins present in crude nuclear extracts. To determine if the mobility shift complex that was formed between components of

Figure 7. Mobility supershift assay of complexes formed between APBα and USF43 (lanes 1 and 2), USF44 (lanes 3 and 4), USF43+USF44 (lanes 5 and 6) or bovine brain nuclear extract (lanes 7 and 8). Mobility shift electrophoresis was performed without (–) or with (+) addition of antibodies against USF43. The complexes formed between APBα and homodimers (43 and 44) or heterodimers (43/44) of USF43 and USF44 are indicated. The additional binding of antibody to the pre-formed APBα–protein complex results in a further reduction in electrophoretic mobility, observed as a supershift (S).

bovine nuclear extract and the APBα binding site indeed involved the USF proteins, a mobility supershift assay was performed. A pre-formed complex of *in vitro* translated USF43 with the 5'-end-labeled APBα oligonucleotide was incubated with antibodies against the C-terminal portion of human USF43. This antibody does not cross-react with USF44 (42). Mobility shift electrophoresis resulted in a complete supershift of the USF43 homodimer–APBα complex (Fig. 7A, lanes 1 and 2). In contrast, the USF44 homodimer–APBα complex was not recognized by the antibody and did not produce a supershift (Fig. 7, lanes 3 and 4). As described above, three binding complexes were observed with the co-translated USF43 and USF44 proteins and APBα (Fig. 5). A supershift was observed with the two binding complexes representing the bound USF43 homodimer and the USF43/44 heterodimer. The binding complex formed by the USF44 homodimer remained unaffected by the antibody supershift (Fig. 7, lanes 5 and 6). These experiments confirmed that the antibody specifically recognizes complexes which contain USF43.

Incubation of APBα with bovine brain nuclear extract resulted in formation of the major complex Cα (Fig. 6D). As described in Figure 1, this complex was surrounded by the minor complexes Cα1 and Cα2. Incubation of the binding complexes Cα, Cα1 and Cα2 with USF43 antibody resulted in a supershift of complexes Cα and Cα2. Only the minor complex Cα1 remained unaffected by the antibody against USF43 (Fig. 7, lanes 7 and 8).

In summary, the complexes formed by the *in vitro* translated USF proteins and the proteins from bovine brain nuclear extract appear to be similar in relative binding affinity, electrophoretic mobility and antibody recognition. This provides strong evidence that the major complex Cα represents binding of the USF43/44 heterodimer and the minor complexes Cα1 and Cα2 represent binding of small amounts of USF44 and USF43 homodimers respectively.

DISCUSSION

USF was originally identified as a cellular protein that interacts with and activates the nearly palindromic sequence GGCCACGT-

GACC in AdMLP (32). USF binding sites have been found in a variety of cellular and viral genes (Table 1). The high diversity of USF binding sites makes it difficult to deduce the precise consensus sequence. Many sites contain the palindromic core CACGTG hexamer sequence, while others contain versions of the less defined E-box CANNTG recognition motif for bHLH transcription factors (25). In addition, most E-box hexamers that bind USF are followed by an AC dinucleotide. USF binds with different affinity to different sequences. The affinity of USF for the binding site in the rat class I alcohol dehydrogenase gene promoter (Table 2) is 20 times less than for the AdMLP sequence (38). Nevertheless, USF has been shown to be important in mediating transcription from this promoter. Mutational analysis of the USF site in the human β -globin locus control region has been described (36). Some mutations within the E-box abolished USF binding, while others only decreased USF binding affinity. Mutation of 2 nt on the 5'-side of the E-box also affected USF binding. Both the APB α sequence of the APP promoter and the CDEI-like elements contain an E-box followed by an AC dinucleotide. The binding affinity of USF greatly depends on the sequence within the E-box, but, at least in the context of the APB α sequence, neighboring nucleotides are also important. This is illustrated by the observation that nuclear factor binding to the APB α domain is significantly reduced by mutation of nucleotides on the 5'-side of the E-box. (17). In a systematic study on the sequence requirements for USF binding it was determined that the optimal binding sequence is RYCAC₋₁G₊₁TGRY (45). This consensus sequence is best matched by the AdMLP binding site. Both the CDEI and APB α domains contain variations in the central (-1/+1) portion of the sequence, which may account for the differences in relative binding affinities (Table 1).

The CDEI element was originally identified as a sequence common to the centromeres of yeast chromosomes. The yeast gene encoding the 39.5 kDa CDEI binding protein has recently been cloned by three independent groups (54–56). The protein, designated CPF1, CBF1, or CP1, is required for chromosome stability and also participates in regulation of gene expression. Although CDEI-like sequences are not routinely found in the centromeres of higher eukaryotes, microinjection of oligonucleotides containing the CDEI motif into fertilized mouse eggs interferes with embryonic development (27). A truncated mouse protein of 511 amino acids, designated CDEBP, was subsequently cloned and its CDEI binding properties were demonstrated for a truncated fusion construct expressed in *E.coli* (28). We did not observe binding of full-length human APLP2 proteins (human homologs of mouse CDEBP) to the CDEI sequence. The reason for this is unclear. It is conceivable that despite the high degree of similarity between the CDEBP and APLP2 proteins, even subtle changes in amino acid composition have contributed to the loss of binding activity in the human proteins. Alternatively, the DNA binding activity of the CDEBP protein as reported by Vidal *et al.* (28) might be a property unique to the truncated fusion protein. However, the possibility cannot be excluded that some as yet unidentified alternatively spliced or post-translationally modified forms of APLP2 are capable of this binding. In a different study a human CDEI binding protein was described with a molecular weight range of 39–49 kDa (31). Incidentally, this molecular weight is more similar to USF than to CDEBP. Our data indicate that USF binds to the CDEI sequence. Hence, the detrimental effect of CDEI oligonucleotide injection on the early development of mouse embryos could be due to oligonucleotide competition with endogenous USF binding sites.

In order to determine if USF binds to APB α and CDEI-like sites, we performed electrophoretic mobility shift assays with oligonucleotides containing these elements. These assays were performed using both *in vitro* translated USF43 and USF44 proteins and nuclear extracts from rat brain, bovine brain and HeLa cells. Competition experiments revealed that both the *in vitro* translated USF proteins and proteins from nuclear extracts bind to the CDEI and APB α oligonucleotide sequences with the same respective 5- and 20-fold lower affinity than to the AdLMP sequence. To determine if the main APB α binding activity in brain nuclear extract is represented by USF, we performed supershift assays with a USF43-specific antibody. The results show that the main binding complex C α and the smaller complex C α 2 in bovine brain nuclear extract contain protein with specific USF43 immunoreactivity. However, the adjoining complex C α 1 was not recognized by the antibody against USF43. The same result was obtained with HeLa cell nuclear extract (data not shown).

It has been shown that the USF43 and USF44 proteins can bind to their recognition sites in homodimeric as well as in heterodimeric forms (42). We confirmed this observation by competing the binding of USF43 and USF44 to APB α with AdLMP, CDEI, APB α and APB α mutations APM1–APM4. In all cases the affinity for their respective binding sites was the same for the homodimeric and the heterodimeric forms. Furthermore, the main complex C α obtained with nuclear extract was flanked in close proximity by two minor complexes, C α 1 and C α 2. These minor complexes were not always readily distinguishable on the autoradiographs reproduced here. However, in experiments demonstrating the relative thermostability of the nuclear factor binding proteins, a rearrangement of the binding complexes was observed. After heat treatment the C α 1 and C α 2 complexes become more pronounced, with a concomitant reduction in the C α complex. This can best be explained by a dissociation of the USF43/44 heterodimers during heat treatment followed by a random reassociation in which the relative proportion of the homodimers increases.

It was previously shown that the proximal APP promoter comprising 94 bp upstream from the transcriptional start site (+1) contains two characterized nuclear factor binding sites, designated APB α (position -53 to -42) and APB β (position -93 to -82) (23). Proximal APP promoter activity under standard cell culture conditions is primarily mediated by APB β , which contributes 70–90% of promoter activity. A similar element has been reported in the rat APP promoter (22). The contribution of binding site APB α is lower and represents 10–30% of the total activity from the proximal human APP promoter. Mutations within the core recognition domain that eliminate factor binding also eliminate promoter activity from APB α (23). Consequently, the activity attributable to APB α is correlated with the ability of nuclear factor to bind to this domain.

Evidence presented here strongly suggests that the proteins in crude nuclear extracts that bind to APB α are homodimers and heterodimers of USF43 and USF44. However, the role of USF in the regulation of APP promoter activity remains elusive. Under standard cell culture conditions in a variety of cell lines a contribution of USF is not essential for high levels of expression from the APP promoter (23). However, the APB α binding site is completely conserved between the human, rat and mouse promoters (14,15,20,21), suggesting a possible functional role in promoter regulation. APP gene expression is regulated in a cell- and tissue-specific manner (11–13) and after exposure to external activating factors in responsive cell types (46–53,57). In addition, APP expression in

culture is dependent on cell density and developmental stage of the cells (47,53). The mechanisms by which most of these regulatory events are controlled have not yet been elucidated and it is possible that in some cases the binding of USF to APB α is an essential step in the regulation of APP gene expression.

ACKNOWLEDGEMENTS

We thank Dr Michèle Sawadogo for kindly providing plasmids containing the USF cDNAs and the antibody against USF43. This work was supported by National Institutes of Health grants NS30994 to WQ and AG09320 to DG.

REFERENCES

- Burger, P.C. and Vogel, F.S. (1973) *Am. J. Pathol.* **73**, 457–468.
- Ellis, W.G., McCulloch, J.R. and Corley, C.L. (1974) *Neurology* **24**, 102–106.
- Glennner, G.G. and Wong, C.W. (1984) *Biochem. Biophys. Res. Commun.* **120**, 885–890.
- Masters, C.L., Simms, G., Weinmann, N.A., Multhaup, G., McDonald, B.L. and Beyreuther, K. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4245–4249.
- Selkoe, D.J. (1991) *Neuron* **6**, 487–498.
- Goldgaber, D., Lerman, M.I., McBride, O.W., Saffiotti, U. and Gajdusek, D.C. (1987) *Science* **235**, 877–880.
- Kang, J., Lemaire, H.-G., Unterbeck, A., Salbaum, M.J., Masters, C.L., Grzeschik, K.H., Multhaup, G., Beyreuther, K. and Müller-Hill, B. (1987) *Nature* **325**, 733–736.
- Robakis, N.K., Wisniewski, H.M., Jenkins, E.C., Devine-Gage, E.A., Houck, G.E., Yao, X.L., Ramakrishna, N., Wolfe, G., Silverman, W.P. and Brown, W.T. (1987) *Lancet* **8529**, 384–385.
- Tanzi, R.E., Gusella, J.F., Watkins, P.C., Bruns, G.A.P., St George-Hyslop, P., Van Keuren, M.L., Patterson, D., Pagan, S., Kurnit, D.M. and Neve, R.L. (1987) *Science* **235**, 880–884.
- Cohen, M.L., Golde, T.E., Usiak, M.F., Younkin, L.H. and Younkin, S.G. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 1227–1231.
- Johnson, S.A., McNeill, T., Cordell, B. and Finch, C.E. (1990) *Science* **248**, 854–856.
- Neve, R.L., Finch, E.A. and Dawes, L.R. (1988) *Neuron* **1**, 669–677.
- Schmechel, D.E., Goldgaber, D., Burkhardt, D.S., Gilbert, J.R., Gajdusek, D.C. and Roses, A.D. (1988) *Alzheimer Dis. Associated Disorders* **2**, 96–111.
- Salbaum, J.M., Weidemann, A., Lemaire, H.-G., Masters, C.L. and Beyreuther, K. (1988) *EMBO J.* **7**, 2807–2813.
- LaFauci, G., Lahiri, D.K., Salton, S.R.J. and Robakis, N.K. (1989) *Biochem. Biophys. Res. Commun.* **159**, 297–304.
- Lahiri, D.K. and Robakis, N.K. (1991) *Mol. Brain Res.* **9**, 253–257.
- Quitschke, W.W. and Goldgaber, D. (1992) *J. Biol. Chem.* **267**, 17362–17368.
- Pollwein, P., Masters, C.L. and Beyreuther, K. (1992) *Nucleic Acids Res.* **20**, 63–68.
- Pollwein, P. (1993) *Biochem. Biophys. Res. Commun.* **190**, 637–647.
- Izumi, R., Yamada, T., Yoshikai, S.-i., Sasaki, H., Hattori, M. and Sakaki, Y. (1992) *Gene* **112**, 189–195.
- Chernak, J.M. (1993) *Gene* **133**, 255–260.
- Hoffman, P.W. and Chernak, J.M. (1994) *Biochem. Biophys. Res. Commun.* **201**, 610–617.
- Quitschke, W.W. (1994) *J. Biol. Chem.* **269**, 21229–21233.
- Beyreuther, K. and Masters, C.L. (1991) *Brain Res. Rev.* **16**, 86–88.
- Murre, C., McCaw, P.S. and Baltimore, D. (1988) *Cell* **56**, 777–783.
- Clarke, J. and Carbon, J. (1985) *Annu. Rev. Genet.* **19**, 29–56.
- Blangy, A., Léopold, P., Vidal, P., Rassoulzadegan, M. and Cuzin, F. (1991) *Nucleic Acids Res.* **19**, 7243–7250.
- Vidal, F., Blangy, A., Rassoulzadegan, M. and Cuzin, F. (1992) *Biochem. Biophys. Res. Commun.* **189**, 1336–1341.
- Sprecher, C.A., Grant, F.J., Grimm, G., O'Hara, P.J., Norris, F., Norris, K. and Foster, D.C. (1993) *Biochemistry* **32**, 4481–4486.
- Wasco, W., Gurubhagavatula, S., d'Paradis, M., Romano, D.M., Sisodia, S.S., Hyman, B.T., Neve, R.L. and Tanzi, R.E. (1993) *Nature Genet.* **5**, 95–100.
- Bram, R.J. and Kornberg, R.D. (1987) *Mol. Cell. Biol.* **7**, 403–409.
- Sawadogo, M. and Roeder, R.G. (1985) *Cell* **43**, 165–175.
- Chodosh, L.A., Cartew, R.W., Morgan, J.G., Crabtree, G.R. and Sharp, P.A. (1987) *Science* **238**, 684–688.
- Cartew, R.W., Chodosh, L.A. and Sharp, P.A. (1987) *Genes Dev.* **1**, 973–980.
- Ellis, J., Talbot, D., Dillon, N. and Grosveld, F. (1993) *EMBO J.* **12**, 127–134.
- Bresnick, E.H. and Felsenfeld, G. (1993) *J. Biol. Chem.* **268**, 18824–18834.
- Scotto, K.W., Kaulen, H. and Roeder, R.G. (1989) *Genes Dev.* **3**, 651–662.
- Potter, J.J., Cheneval, D., Dang, C.V., Resar, L.M.S., Mezey, E. and Yang, V.W. (1991) *J. Biol. Chem.* **266**, 15457–15463.
- Sirito, M., Lin, Q., Maity, T. and Sawadogo, M. (1994) *Nucleic Acids Res.* **22**, 427–433.
- Sawadogo, M., Van Dyke, M.W., Gregor, P.D. and Roeder, R.G. (1988) *J. Biol. Chem.* **263**, 11985–11983.
- Gregor, P.D., Sawadogo, M. and Roeder, R.G. (1990) *Genes Dev.* **4**, 1730–1740.
- Sirito, M., Walker, S., Lin, Q., Kozlowski, M.T., Klein, W.H. and Sawadogo, M. (1992) *Gene Expression* **2**, 231–240.
- Heberlein, U. and Tjian, R. (1988) *Nature* **331**, 410–415.
- Hanes, J., von der Kammer, H., Kristjansson, G.I. and Scheit, K.H. (1993) *Biochim. Biophys. Acta* **1216**, 154–156.
- Bendall, A.J. and Molloy, P.L. (1994) *Nucleic Acids Res.* **22**, 2801–2810.
- Donnelly, R.J., Friedhoff, A.J., Beer, B., Blume, A.J. and Vitek, M.J. (1990) *Cell. Mol. Neurobiol.* **10**, 485–495.
- Goldgaber, D., Harris, H.W., Hla, T., Maciag, T., Donnelly, R.J., Jacobsen, J.S., Vitek, M.P. and Gajdusek, D.C. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 7606–7610.
- Forloni, G., Demicheli, F., Giorgi, S., Bendotti, C. and Angeretti, N. (1992) *Mol. Brain Res.* **16**, 128–134.
- Quon, D., Catalano, R. and Cordell, B. (1990) *Biochem. Biophys. Res. Commun.* **167**, 96–102.
- Mönning, U., König, G., Prior, R., Mechler, H., Schreiter-Gasser, U., Masters, C.L. and Beyreuther, K. (1990) *FEBS Lett.* **277**, 261–266.
- Gray, C.W. and Patel, A.J. (1993) *Mol. Brain Res.* **19**, 251–256.
- Mobley, W.C., Neve, R.L., Prusiner, S.B. and McKinley, M.P. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 9811–9815.
- Adler, M.J., Coronel, C., Shelton, E., Seegmiller, J.E. and Dewji, N.N. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 16–20.
- Mellor, J., Jiang, W., Funk, M., Rathjen, J., Barnes, C.A., Hinz, T., Hegemann, J.H. and Philippsen, P. (1990) *EMBO J.* **9**, 4017–4026.
- Cai, M. and Davis, R.W. (1990) *Cell* **61**, 437–446.
- Baker, R.E. and Masison, D.C. (1990) *Mol. Cell. Biol.* **10**, 2458–2467.
- Trejo, J.A., Massamiri, T., Deng, T., Dewji, N.N., Bayney, R.M. and Brown, J.H. (1994) *J. Biol. Chem.* **269**, 21682–21690.