Interaction of the Transcription Factors USF1, USF2, and α-Pal/Nrf-1 with the FMR1 Promoter

IMPLICATIONS FOR FRAGILE X MENTAL RETARDATION SYNDROME*

Daman Kumari and Karen Usdin‡

From the Section on Genomic Structure and Function, Laboratory of Molecular and Cellular Biology, NIDDK, National Institutes of Health, Bethesda, Maryland 20892-0830

Hypermethylation of the FMR1 promoter reduces its transcriptional activity, resulting in the mental retardation and macroorchidism characteristic of Fragile X syndrome. How exactly methylation causes transcriptional silencing is not known but is relevant if current attempts to reactivate the gene are to be successful. Understanding the effect of methylation requires a better understanding of the factors responsible for FMR1 gene expression. To this end we have identified five evolutionarily conserved transcription factor binding sites in this promoter and shown that four of them are important for transcriptional activity in neuromuscular derived cells. We have also shown that USF1, USF2, and α-Pal/Nrf-1 are the major transcription factors that bind the promoter in brain and testis extracts and suggest that elevated levels of these factors account in part for elevated FMR1 expression in these organs. We also show that methylation abolishes α-Pal/Nrf-1 binding to the promoter and affects binding of USF1 and USF2 to a lesser degree. Methylation may therefore inhibit FMR1 transcription not only by recruiting histone deacetylases but also by blocking transcription factor binding. This suggests that for efficient reactivation of the FMR1 promoter, significant demethylation must occur and that current approaches to gene reactivation using histone deacetylase inhibitors alone may therefore have limited effect.

Fragile X syndrome is caused by the expansion of a CGG repeat in the 5’-untranslated region of the fragile X mental retardation (FMR1) gene (1, 2). This results in hypermethylation of the promoter and transcriptional silencing (3). The major symptoms of fragile X syndrome, mental retardation and macroorchidism, are consistent with the observation that high levels of FMR1 expression occurs in specific cells in brain and testis (4). The GC-rich human FMR1 promoter lacks a typical TATA-box and contains several potential Sp1 binding sites as well as an E-box and putative binding sites for the transcription factors α-Pal/Nrf-1,1 AP2, AGP/EBP, and Zeste (5). Four regions of protein binding in the unmethylated promoter of the FMR1 gene have been described by in vivo dimethyl sulfate footprinting analysis in human fibroblasts, peripheral lymphocytes, and lymphoblastoid cell lines (5, 6). These footprints correspond to the α-Pal/Nrf-1 site, 2 GC-boxes, and the E-box. However, the transcription factors that interact with these sites have not yet been identified. Moreover, whereas these footprints do reflect in vivo interactions, their relevance to the regulation of FMR1 transcription in brain and testis, the two major affected organs, is unknown.

Because FMR1 knockout mice demonstrate learning deficits and macroorchidism similar to those seen in fragile X patients (7), and mice show patterns of temporal and tissue-specific FMR1 expression similar to humans (8), many of the control elements important for FMR1 gene regulation are also likely to be evolutionarily conserved. To identify conserved promoter elements, we compared the sequence of the human FMR1 promoter with that from two other primates: Pan troglodytes (chimpanzee) and Macaca arctoides (stump-tailed macaque) as well as two more evolutionarily distant species: Mus domesticus (mouse) and Canis familiaris (dog). We also examined the activity of promoter mutations in transient expression assays using a neuronally derived cell line, PC12, and identified the major transcription factors that bind the FMR1 promoter in nuclear extracts of brain and testis. We also demonstrate that binding of one of these factors is abolished by methylation, and binding of the other two factors is also affected. This may have implications for therapeutic strategies aimed at reactivating the gene, since it indicates that methylation of the FMR1 promoter in Fragile X patients does not simply inhibit transcription via the formation of transcriptionally inactive chromatin, as suggested from in vivo dimethyl sulfate footprinting in lymphoblasts (6) and from the presence of deacetylated histones on the FMR1 promoter in individuals with Fragile X syndrome (9).

EXPERIMENTAL PROCEDURES

Isolation of the Mouse fmr1 Promoter—A 129/SVJ mouse (M. domesticus) BAC library was screened by polymerase chain reaction using two primers from exon 1 of the mouse FMR1 gene (Genome Systems, St. Louis, MI). The resultant clone was subcloned and mapped using standard procedures. Exon 1 of the mouse FMR1 gene was localized to a 3.1-kilobase EcoRI fragment that was subcloned into pZero (Invitrogen, Carlsbad, CA). This clone designated pEco3.3 was then sequenced using standard procedures. Sequence comparison was done using the GCG package (Wisconsin Package Version 10.0, Genetics Computer Group (GCG), Madison, WI). The mouse sequence was scanned against the GenBank database. The only significant matches were to the human cAMP-response element-binding protein; bp, base pair(s); EMSA, electrophoretic mobility shift assay.

Received for publication, October 20, 2000

Published, JBC Papers in Press, October 31, 2000, DOI 10.1074/jbc.M009629200

* 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.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank/EBI Data Bank with accession number(s) AF251347, AF251348, AF251349, and AF251350.

‡ To whom correspondence should be addressed: Bldg. 8, Rm. 302, National Institutes of Health, 8 Center Dr. MSC 6850, Bethesda, MD 20892–0830. Tel.: 301-496-2189; Fax: 301-402-0053; E-mail: ku@helix.nih.gov.

1 The abbreviations used are: Nrf, nuclear respiratory factor; CREB, cAMP-response element-binding protein; bp, base pair(s); EMSA, electrophoretic mobility shift assay.

This paper is available on line at http://www.jbc.org

Vol. 276, No. 6, Issue of February 9, pp. 4357–4364, 2001
Printed in U.S.A.
FMR1 5'-end (including GenBank27/EBI locus HUMFMR1S). The sequence was submitted to GenBank27/EBI (Accession Number: AF251347).

Amplification, Sequencing, and Computer Analysis of the Promoter Region from Different Species—Polymerase chain reaction amplification of FMR1 promoters that were cloned from the genomic DNA of chimpanzee (Pan troglodytes), macaque (Macaca fascicularis), and dog (Canis familiaris) was carried out using the ExpandTM HiFidelity polymerase chain reaction system (Roche Molecular Biochemicals) and the primers Fraxa f (5'-dAGC-CCCGAATCCTACACACAGCTTCTCA-3' from exon 1 and FRMUP2 (5'-dGCNTTCCGCCCTCACTCCAGAAG-3') homologous to the 5' end of the promoter that is conserved in mice and humans. The polymerase chain reaction product was directly sequenced using the Thermosequence radiolabeled terminator cycle sequencing kit (U. S. Biochemicals Corp.) according to the manufacturer's recommendations. The sequences were submitted to GenBank29 (accession numbers AF251349, AF251350, and AF251348 for chimpanzee, macaque, and dog, respectively).

Generation of Reporter Constructs—A single base insertion was introduced into the middle of the E-box site in p32.9 (10) using the QuickChange Mutagenesis™ protocol (Stratagene, La Jolla, CA) and the primer pair 5'-dGACACGTTGCTACATCGGTGTTCGTAGTGGTAC3' and 5'-dGTAACACTGAAACACGCTCAGTCAAGCGCTGTT-3'. The introduction of the single base change to the resultant plasmid (pUSFmut) was confirmed by sequencing. A 909-bp fragment from the original FMR1 promoter in p32.9 and a 910-bp fragment from the mutated promoter from pUSFmut containing 869 and 870 bases from the human FMR1 promoter, respectively, were cloned into KpnI-Ndel-digested plasmid pGL3-basic (Promega, Madison, WI), which contains the firefly Photinus pyralis luciferase reporter gene to make pGL-FMR and pGL-USFmut. The 909-bp wild-type FMR1 promoter was also cloned into HindIII-SpeI-digested pRL-null (Promega), which contains the sea pansy Renilla reniformis luciferase-coding sequence, to make the control plasmid pRL-FMR. Deletions in the promoter were created either by Exonuclease III and S1 nuclease treatment or by digestion with a combination of restriction enzymes followed by T4 DNA polymerase. Specifically plasmids pΔ-88/+244, pΔ-55/+244, and pΔ-447/+244 were generated by digesting pGL-FMR with BglII and filling in the recessed 3' ends with α-phosphothioate deoxyribonucleotides followed by digestion with Nhel, Exonuclease III, and S1 nuclease according to the method of the manufacturer. Plasmids pΔ-131/-123 were digested by pGL-FMR with BssHII followed by ligation, pΔ-131/-123dUSFmut was generated from pGL-USFmut in the same way. Plasmids pΔ-131/-123 were digested by pGL-FMR with BssHII for transcription and ligation. Plasmid pΔ-85/-9 was obtained by digestion of pGL-FMR with PmlI and treatment with Exonuclease III and S1 nuclease.

A 6-base deletion was made in the Initiator (Inr-like) element in the FMR1 promoter using the QuickChange Mutagenesis™ protocol and the primer pair 5'-dGGCGCGGGGTTCGCCACCAGCTCAGCTCC-3' and 5'-dGGATCTGAGGCGGCCGAAACTCCCTCAGGTG-3'. This resulted in plasmid pΔ-5'-5'. The deletion was confirmed by sequencing. Similarly, a four-base pair change was made in the two GC-boxes by QuickChange Mutagenesis™ protocol, and the mutation was confirmed by sequencing. Specifically, Sp1 site was mutated using primers 5'-dACCCTTGGACAGGAGACTGTCGCGGGAACCGTACG-3' and 5'-dGCTGCTGACCCGGCCAGCTTCTCTCTCTCAC-3' to get plasmid pSm1mut, and the Sp1-like site was mutated using primers 5'-dGCNTTCCGCCCTCACTCCAGAAG-3' and 5'-dGTAACACTGAAACACGCTCAGTCAAGCGCTGTT-3'. The mutations were confirmed by digestion with BsgHII and ligation to get pΔ-5'-131/-123dSm1mut and pΔ-131/-123dSm2mut, respectively.

Methylation of Reporter Constructs—Specific methylation of the E-box and α-Pal/NRf-1 sites were achieved by in vivo methylation with FmlI and BssHII, methylases, respectively. This was accomplished using the MBD vector for the methylation assay from the Methylation Assay Kit (Promega, Madison, WI) according to the manufacturer's protocol. The MBD was ligated into the XhoI site of the BssHIIHIM gene (pL3G39/BssHIM) was a gift of New England Biolabs, Beverly, MA. A similar clone was constructed for the FmlI methylase by subcloning the 3.5-kilobase EcoRI-Sau3A fragment from pEco72 m (a gift of MBI Fermentas Inc.) into pGL339 to generate pGL339/ FmlI. pGL-FMR was cotransformed into Enterococcus coli along with either pGL339/BssHIII or pGL339/PmlI. Methylation in all cases was confirmed by digestion with FmlI or BssHIII. Since the plasmids were not propagated in vivo methylation was accomplished with the low copy number pGL339 derivative, an unmethylated control was prepared by cotransforming pGL-FMR with pGL339.

Cell Culture, Transient Transfections, and Promoter Assays—Cells were grown in Dulbecco's modified Eagle's medium (Life Technologies, Inc.) supplemented with 5% fetal calf serum (Life Technologies Inc.), 10% heat-inactivated horse serum (Sigma), and 1% penicillin-streptomycin at 37 °C in 5% CO2, the culture medium was replaced 18–24 h before the transfection. Ten micrograms of test plasmid DNA together with 10 μg of the control plasmid pRL-FMR were introduced into ~106 cells by electroporation (300 V, 1180 capacitance; Cell-porator, Life Technologies). Transfected cells were plated in duplicate on 6-well plates. After 16 h, the culture medium was replaced with fresh medium. Cells were collected 42–44 h after transfection and assayed.

Transcription Factors, Competitor Oligonucleotides, and Antibodies—Oligonucleotides used in the electrophoretic mobility shift assay are listed in Table I. The consensus binding sites for the transcription factors are underlined. Double-stranded oligonucleotides containing the consensus binding sites for AP1, AP2, Sp1, OCT1, TFIIID, and CREB were obtained from Promega. The top and bottom strands of two variants of the α-Pal/NRf-1 site and that of consensus E-box and SREBP-1 sites were synthesized by Life Technologies. The complementary strands were annealed and used without further purification. Antibodies against USF1, USF2, Max, c-Myc, Sp1, Sp3, Sp4, and Egr-1 were from Santa Cruz Biotechnology, Inc. (Santa Cruz, CA). CREB antibody was from New England BioLabs. α-Pal/NRf-1 antibodies were a gift from Dr Brian Safer (NHBLI, NIH, Bethesda, MD).

Preparation of Nuclear Extracts—Nuclear extracts were prepared from the brain, testis, and liver of FVB/N mice (The Jackson Laboratory, Bar Harbor, ME) and from human lymphoblastoid cell lines (Coriell Cell Repositories, Camden, NJ) by a modification of the method of Dignam et al. (11). Briefly, the tissues and lymphoblastoid cells were washed with phosphate-buffered saline and resuspended in two volumes of ice-cold buffer A (10 mM HEPES, pH 7.9, 1.5 mM MgCl2, 10 mM NaCl, 0.5 mM dithiothreitol, 0.5 mM phenylmethylsulfonyl fluoride, 7 μg/ml caprin inhibitor II) containing 1 protease inhibitor mixture tablet (Complete™ mini, EDTA-free, Roche Molecular Biochemicals, Indianapolis, IN) per 10 ml of buffer A. The cells were lysed by 10 strokes with a Dounce homogenizer. The supernatant was removed after spinning at 25,000 × g for 20 min. The remaining pellets were quick-frozen in a dry ice/ethanol bath and stored in aliquots at −80 °C. Protein concentrations were determined using the Bio-Rad protein assay reagent. Nuclear extracts from PC12 cell lines were purchased from Promega.

Electrophoretic Mobility Shift Assays (EMSAs)—A 200-bp mouse FMR1 promoter fragment was amplified with primers mba1 (5'-dCTTCTTTGAGATCCTTTCCTGCGTC-3') and mba2 (5'-dGCTGCTGACCCGGCCGAAACTCCCTCAGGTG-3'). The amplified fragments were gel-purified and cloned into the HindIII site of pBS (SK+). For EMSA and DNase I footprinting experiments, a gel-purified SalI/SalI fragment containing the promoter was end-labeled at

---

2 Schug, J. and Overton, G. C. (1997) Technical Report CBIL-TR-1997-1001-v0.0, Computational Biology and Informatics Laboratory, School of Medicine, University of Pennsylvania.
The NoRI site using [32P]dCTP and Klenow DNA polymerase (Life Technologies). The labeled probe was extracted with phenol/chloroform and purified by passing through a Sephadex G-50 column. The promoter was methylated using SsoI methylase (New England BioLabs). Methylation was verified by digestion with the methylation-sensitive enzymes PvuII, BglI, and HpaII (New England BioLabs).

Binding was carried out at 30 °C or 4 °C in 30 μl of reaction buffer containing 25 mM HEPES, pH 7.5, 5 mM MgCl₂, 2 mM dithiothreitol, 100 mM NaCl, 0.25 μg of probe, 5 μg of protein, and 1 μg of poly[d(A-T)]·[d(A-T)] for 30 min. A 1000-fold excess of nonspecific DNA and transcription factor binding oligodeoxyribonucleotides were included in the reactions as nonspecific or specific competitors. For antibody supershift assays, 4 μg of antibody or BSA was incubated with the protein before the addition of the probe, and reactions were carried out for 40 min. The reactions were stopped by the addition of 30 μl of 2× gel loading buffer (200 mM Tris, pH 8.8, 10.5% glycerol, 0.002% bromphenol blue). Reactions were subjected to electrophoresis on a 4% polyacrylamide gel (60:1, acrylamide:bis) containing 1.6% glycerol in 1× Tris-glycine-EDTA buffer (50 mM Tris, 380 mM glycine, 2.1 mM EDTA, pH 8.5). Gels were dried and exposed to x-ray film.

**DNaese I Footprinting**—For DNaese I footprinting, a PstI/EcoN1 fragment of p32.9 was end-labeled at EcoN I site by [α-32P]dGTP and Klenow DNA polymerase (Life Technologies). Binding reactions were carried out with 1 ng of 3′ end-labeled probe and 20 μg of protein in a 100-μl reaction volume containing poly[d(A-T)]·[d(A-T)]. The reactions were treated with 0.25 units of DNase I in the presence of 10 mM MgCl₂ and 1 mM CaCl₂ for 10 min at 30 °C. The reactions were phenol-extracted and ethanol-precipitated with tRNA as carrier. The sample was dissolved in 100 μl of TE (10 mM Tris-HCl, pH 8.0, 1 mM Na₂EDTA) and butanol-precipitated. The pellet was washed with 70% ethanol, and resuspended in 10 mM Tris-HCl, pH 8.0, 1 mM EDTA, and 0.5% sodium dodecyl sulfate. The supernatant was collected, and the pellet was dissolved in 10 mM Tris-HCl, pH 8.0, 1 mM EDTA, and 0.5% sodium dodecyl sulfate. These were then treated with 0.25 units of DNase I in the presence of 10 mM MgCl₂, and 1 mM CaCl₂ for 10 min at 30 °C. The reactions were phenol-extracted and ethanol-precipitated with tRNA as carrier. The sample was dissolved in 100 μl of TE (10 mM Tris-HCl, pH 8.0, 1 mM Na₂EDTA) and butanol-precipitated. The pellet was washed with 70% ethanol, dried, and resuspended in 10 μl of formamide stop buffer (95% formamide, 20 mM EDTA, 0.05% bromphenol blue, 0.05% xylene cyanol FF). A 2-μl sample was electrophoresed on a 6% sequencing gel at 1600 V until the bromphenol blue dye reached the bottom of the gel. The gel was dried and exposed to x-ray film.

**Western Blot Analysis**—The nuclear extracts were subjected to electrophoresis on 10% polyacrylamide gels containing SDS and electroblotted to the NitroPure membrane (MSI, Westboro, MA) using standard procedures. Hybridization with primary antibody was carried out as specified by the supplier. Hybridization with the secondary antibody and signal detection were done using ECL Western blotting kit (Amer sham Pharmacia Biotech) as per the manufacturer's recommendation.

**RESULTS**

**Phylogenetic Footprinting of the 5′ End of FMR1 Gene Identifies Five Evolutionarily Conserved Regions**—A 466-bp fragment including 272 bp of sequence upstream of the transcription start site of the human FMR1 gene has previously been shown to contain all the elements necessary for the appropriate tissue-specific expression of the FMR1 gene in transgenic mice (12). To define the minimal promoter region more closely and to identify those transcription factor binding sites that are evolutionarily conserved and that may therefore be important for regulation of this gene, we compared the sequences of the 5′ end of the FMR1 gene that we obtained from a number of different mammals.

Fig. 1 shows a dot matrix comparison of the exon 1-containing region of the mouse fmr1 gene with the corresponding region of the human FMR1 sequence (GenBank™/EBI locus: HUMFMR1S). The homology between the two sequences is highest in the first exon, but islands of homology both 3′ and 5′ of exon 1 are also seen. The 5′ regions fall within the 272-bp previously defined minimal promoter fragment (12), the upstream border of which is marked by the black arrow in Fig. 1. No transcription factor binding sites were conserved between exon 1 and exon 6 of the α-Pal/NrF-1 site and 1313 bases of the start of transcription. This suggests that the minimal FMR1 promoter may only be 131 bp long. The conserved 3′ regions may represent regulatory elements in the first intron, but they have not been studied in any detail to date.

We used a primer derived from the conserved 5′ boundary of the promoter and a primer from exon 1 to amplify the FMR1 5′ region from chimpanzee, Stump-tailed macaque, and dog. The sequences were then aligned with the human and mouse promoter sequences to identify those regions that are conserved in all five species (Fig. 2). The four described previously in vivo protein binding sites, α-Pal/NrF-1, the 2 Sp1 or GC boxes, and the E-box (sometimes referred to as the c-Myc binding site) (5, 6) are conserved in all five species (shown in the dark gray boxes in Fig. 2). None of these sequences had a good TATA-box. However, they all contained a conserved motif close to the reported start of transcription (13) that resembles an Initiator (Inr) element. Such elements direct transcription initiation in certain TATA-less promoters (14).

The region containing the dimethyl sulfate hyperreactive G residue close to the reported transcriptional start site (5) is also well conserved (the complementary C is marked by an asterisk in Fig. 2). This region does not show a protein binding footprint in vivo, at least in cells with low FMR1 activity (5). This hyperreactivity is absent in individuals with fragile X syndrome and is therefore thought to reflect some DNA structure that is associated with transcription (5). The transcription factors Zeste (5), AP2 (15), and CREB (13) have all been suggested to be important for FMR1 regulation. However the binding sites for these factors were not conserved (shown in the open boxes Fig. 2).

**Mutational Analysis Confirms That the Evolutionarily Conserved Regions Are Important for Transcriptional Activity**—Mutated versions of the human FMR1 promoter were assayed in PC12 cells, a neurally derived cell line. Fig. 3 shows that single base insertion in the E-box (pGL-USFmut) or the deletion of 9 bases from the α-Pal/NrF-1 site (pA−131/−123) each reduced the expression of the reporter gene about 5-fold. Both mutations together (pA−131/−123/USFmut) reduced activity almost to that of a construct containing a deletion of almost all of the FMR1 sequence (pA−447/+244). Therefore the upstream
boundary of the minimal promoter region is probably close to the 5' end of the α-Pal/Nrf-1 site, as suggested by the phylogenetic footprinting. A deletion that included the α-Pal/Nrf-1 site and one of the GC-boxes (pΔ−151/−81) led to an increase in activity over that seen with the nine-base deletion in the α-Pal/Nrf-1 site alone (pΔ−131/−123). Similarly, a deletion that removes both the E-box sequence and its adjacent Sp1 site (pΔ−85/−9) produced a higher activity than the mutant with the single base insertion in the E-box. The activity of this construct was in fact higher than the full-length promoter (pGL-FMR). However, when a 4-base substitution mutation was made in either of the two GC boxes, the activity was reduced to 75 and 50%, respectively (pS1mut and pS2mut). This suggests that these two regions have a positive role in the control of gene expression. The increase in activity seen for deletion constructs pΔ−151/−81 and pΔ−85/−9 that lack these sites could be due to decreased distance between the α-Pal/Nrf-1 site and the transcription initiation site or to the deletion of additional sequences that contain negative regulators of FMR1 activity. Constructs containing both the α-Pal/Nrf-1 deletion and the point mutation in the first GC box (pΔ−131/−123/S1mut) had an activity similar to the α-Pal/Nrf-1 deletion alone, suggesting that there might be some interaction between factors that bind these two sites. However, a pΔ−131/−123:S2mut double mutant showed a slightly higher activity than the pΔ−131/−123 deletion by itself. Why this should be is not clear at this time but may be related to the uncovering of secondary transcription factor binding sites that can now be used. The pΔ−85/−9 construct lacks the TATA-like sequence upstream of the transcriptional start site. Since deletion of this sequence did not negatively affect promoter activity, it is possible that the TATA-like sequence is dispensable. However,
although the initiator-like element is evolutionarily conserved, its deletion had no effect on the promoter activity in PC12 cells (construct pΔ+5/+10 in Fig. 3). Since the region that includes the putative transcriptional start site is not conserved, it may be that transcription initiation in the FMR1 promoter occurs via a novel mechanism that involves as yet undefined signals.

Deletion of the region that included the CGG repeats in the 1st exon (pΔ+53/+244) produced a small increase in reporter gene activity. Since this sequence is located within the transcript, it is possible that this effect is mediated either at the level of transcription, mRNA stability, or translation. A negative effect of CGG repeats on translation has been suggested based on the observation that the levels of FMRF, the protein product of the FMR1 gene, are reduced in individuals with long premutation alleles (16). However, this effect is probably only significant at high repeat numbers. A 20-kDa protein has been shown to bind to CGG repeats and inhibit FMR1 transcription (17). Deletion of the CGG repeat tract may simply eliminate the effect of these or similar proteins.

USF1, USF2, and α-Pal/Nrf-1 Bind to the FMR1 Promoter—The 200-bp region of both the mouse and human FMR1 promoters, which contains all the evolutionarily conserved putative transcription factor binding sites (shown in bold in Fig. 2) was used as a probe for binding factors in nuclear extracts from mouse brain, testis, and liver and human lymphoblastoid and PC12 cell lines. EMSA using brain nuclear extracts and SDS-polyacrylamide gel electrophoresis and Western blotting of brain, testis, and liver nuclear extracts were carried out as described under “Experimental Procedures” using antibodies to c-Myc, CREB, USF1, and USF2 as indicated.

An excess of unlabeled oligonucleotide containing the E-box consensus sequence eliminated complexes I and III (Fig. 4, lane 5) as well as the minor shifted products. An oligonucleotide containing the consensus sequence for the α−Pal/Nrf-1 binding site eliminated complex II and III (Fig. 4, lane 9). Complexes II and III were also eliminated when poly[dI-dC] was present, presumably due to the GC richness of the α−Pal/Nrf-1 binding site (data not shown). Antibodies to the E-box-binding proteins USF1 or USF2 supershifted both complex I and III as well as the minor products, indicating that these complexes contained both USF1 and USF2 (Fig. 4, lanes 14 and 15). Antibodies to c-Myc or max, two other E-box binding proteins, had no effect on the gel mobility profiles (Fig. 4, lanes 12 and 13). This is not due to the absence of c-Myc in the extracts since Western blots with c-Myc antibody showed the presence of large amounts of full-length c-Myc protein (Fig. 4, inset). Antibodies to α−Pal/Nrf-1 shifted both complex II and complex III (Fig. 4, lane 20), indicating that α−Pal/Nrf-1 was involved in both of these complexes.

USF1 and USF2 were unable to bind the E-box variant with the single base insertion that was used in the promoter activity studies described above. Moreover, whereas an oligonucleotide containing the unmutated E-box was able to eliminate USF binding to the FMR1 promoter when present in a 10-fold excess, an oligonucleotide containing the mutated E-box did not eliminate binding completely even when a 1000-fold excess was used (data not shown).

Sp3 Binds a FMR1 Promoter Subfragment—Factor binding to the 2 GC boxes in the human promoter has been reported in fibroblasts and lymphoid cells (5, 6) and in vitro on a 71-bp promoter-containing fragment in extracts of the neuronally derived cell line SK-N-SH (15). However, GC-box-specific factors did not bind the full-length promoter in any of our extracts under a variety of reaction conditions; unlabeled GC-box oligonucleotides did not alter the EMSA profile (Fig. 4, lane 6), and antibodies to members of the Sp1 transcription factor family Sp1, Sp3, and Sp4 did not supershift complex I, II, or III (Fig. 4, lanes 16, 18, and 19). Antibodies to Egr-1, a related transcription factor that binds to some GC boxes, also had no effect (Fig. 4, lane 17). However, an 82-bp Sp1-Pml fragment that contained both the conserved GC-boxes but that lacked both the E-box and the α−Pal/Nrf-1 site did produce gel shifted products, one of which could be eliminated by competition with Sp1 oligonucleotide and was supershifted by Sp3 antibody (data not shown). It is possible that lack of Sp3 binding to the larger promoter fragment is due to binding of USF and/or α−Pal/Nrf-1. The in vitro binding of Sp factors in lymphoid cells and fibroblasts might reflect the high concentrations of Sp factors in these cells relative to brain, testis, and liver (19).

CREB and AP2 Do Not Bind the FMR1 Promoter—A CREB binding site that overlaps with the E-box sequence has also been suggested to be important for FMR1 regulation (13). However, the E-box consensus oligonucleotide that was able to abolish complex I lacked sequences corresponding to a CREB site (Table I). Moreover, oligonucleotides containing the con-
TABLE I

| Transcription factor | Sequence                                      |
|----------------------|-----------------------------------------------|
| α-Pal/Nrf-1 a        | 5′-d (GATAGCTGATCCCCAGGGATGATGACGAGATCCG) 3′ |
| α-Pal/Nrf-1 b        | 5′-d (GATAGCTGACGGAGGGCTGAGATGACGAGATCCG) 3′ |
| AP1                  | 5′-d (CTGTGAACTGAGGAGGGCTGAGATGACGAGATCCG) 3′ |
| AP2                  | 5′-d (AGGATATCCCTGAGGAGGGCTGAGATGACGAGATCCG) 3′ |
| CREB                 | 5′-d (ATCGATCCCTGAGGAGGGCTGAGATGACGAGATCCG) 3′ |
| E-box                | 5′-d (ATCGATCCCTGAGGAGGGCTGAGATGACGAGATCCG) 3′ |
| OCT1                 | 5′-d (ATCGATCCCTGAGGAGGGCTGAGATGACGAGATCCG) 3′ |
| Sp1                  | 5′-d (ATCGATCCCTGAGGAGGGCTGAGATGACGAGATCCG) 3′ |
| SREBP-1              | 5′-d (ATCGATCCCTGAGGAGGGCTGAGATGACGAGATCCG) 3′ |
| TFID                 | 5′-d (ATCGATCCCTGAGGAGGGCTGAGATGACGAGATCCG) 3′ |

The consensus sequence for CREB binding did not compete out any of the retarded bands (Fig. 4, lane 7). This is not due to the absence of CREB, because our extracts contained high levels of full-length CREB (Fig. 4, inset). Moreover, when a CREB consensus oligonucleotide is used as a probe with our extracts, a shifted product is seen (data not shown). We found no evidence for any other factor binding including AP2 (data not shown), which has also been reported to be involved in FMR1 regulation (15). The differences between our data and those previously reported might reflect differences between the source of the material used to prepare the nuclear extracts and other experimental conditions. However, the lack of evolutionary conservation of the CREB and AP2 binding sites and the failure of these factors to bind in brain and testis extracts suggests that these sites may be less important for FMR1 expression than previously thought.

A Similar Pattern of Protein Binding Is Seen in Extracts of Liver, Lymphoblasts, and PC12 Cells—A similar pattern of three shifted bands was observed with nuclear extracts from liver, human lymphoblastoid, and PC12 cell lines (Fig. 4, lanes 2 and 3, and data not shown). However, the two smaller of the three minor bands indicated by the bracket in the brain and liver extracts were not seen with the extracts prepared from the cell lines. These bands may be due to proteolysis during the preparation of the nuclear extract from the mouse tissues or specific USF variants that are only seen in certain cells. Complexes I, II, and III in liver extracts could be eliminated by competition with E-box and α-Pal/Nrf-1 oligonucleotides as in brain extracts (results not shown), suggesting that they are similar in composition to those seen in brain extracts.

Testis Extracts Have a Different Pattern of Protein Binding—Although a similar pattern of gel mobility shift was seen for brain, liver, lymphoblasts, and PC12 cells, the pattern of mobility shift in testis extracts was very different. Complex I was absent, and complex II and a large amount of heterodisperse complex, “complex III,” were seen together with a small amount of a novel band (complex IV) with faster mobility (Fig. 4, lane 1). Unlabeled α−Pal/Nrf-1 eliminates complex II and III, and anti-α−Pal/Nrf-1 antibody supershifts both complex II and III (data not shown). Unlabeled E-box oligonucleotide eliminates both complex III and complex IV, and both of these complexes are supershifted with antibodies to USF1 or USF2 (data not shown). Thus, complex II and III in testis are apparently similar to complex II and III in brain and liver, being comprised of α−Pal/Nrf-1, and α−Pal/Nrf-1 plus USF proteins, respectively. Complex IV contains USF1 and USF2 proteins, since it is supershifted by antibodies to both USF1 and USF2. However, since there are no unique testis-specific USF1 or USF2 polypeptides visible on Western blots of SDS-polyacrylamide gel electrophoresis gels of these extracts (Fig. 4, inset, and data not shown), the rapid mobility of USF proteins in complex IV may be due to either charge differences or to the association of USF with a different subset of proteins in this organ.

When the formation of complex III is prevented by the addition of α−Pal/Nrf-1 oligonucleotide, there is little if any increase in complex I in brain (Fig. 4, compare lanes 1 and 6) and liver or complex IV in testis (data not shown). No new bands corresponding to complex I are seen in testis extracts either (data not shown). In contrast, there is an increase in the amount of complex II seen in brain, testis, and liver extracts when excess E-box oligonucleotide is present (Fig. 4, compare lanes 4 and 5). One interpretation of these data is that the USF proteins in complex III do not bind the promoter effectively in the absence of α−Pal/Nrf-1, perhaps because they are complexed with other proteins or modified in some way and that in testis most of the USF is of this form.

Methylation of FMR1 Promoter Prevents α−Pal/Nrf-1 Binding—Methylation of the FMR1 promoter in fragile X patients (20) and in reporter constructs (21) greatly reduces the ability of the promoter to drive transcription. Methylation of the C residues in all the CpG residues in the promoter by Sss1 methylase eliminated complex II and III in gel shift assays done with the human FMR1 promoter and brain extracts but had only a small effect on the formation of complex I (Fig. 5, panel A). Consistent with this observation is the fact that methylation of three CpGs in the α−Pal/Nrf-1 site eliminated the DNase I footprint seen at the α−Pal/Nrf-1 site (Fig. 5, panel B, lane 4), whereas methylation of the single CpG in the E-box reduced but did not eliminate the DNase I protection at that site (Fig. 5, panel B, lane 6). Moreover, transfection experiments showed that methylation of the promoter at the E-box reduced promoter activity by about 20%, whereas methylation of the α−Pal/Nrf-1 site reduced activity by 55% (Fig. 5C). This suggests that binding of α−Pal/Nrf-1 is sensitive to methylation and that USF1/USF2 binding is also sensitive but somewhat less so than α−Pal/Nrf-1. The fact that a deletion in the α−Pal/Nrf-1 site which prevents α−Pal/Nrf-1 binding causes a 75% drop in promoter activity suggests that methylation of this site reduces the ability of α−Pal/Nrf-1 to drive transcription by more than 70%.

Methylation of the E-box also affects the footprint at the α−Pal site (Fig. 5B, lane 6). This suggests that the USF proteins bind to both the E-box and the α−Pal/Nrf-1 site and/or that there is some interaction between these factors. This is consistent with the observations that abolishing complex III by the addition of excess α−Pal/Nrf-1 oligodeoxyribonucleotide does not increase the amount of complex I that is seen and that mutations in the E-box or the α−Pal/Nrf-1 site both reduce promoter activity to 25% that of the wild-type promoter supports this idea. We are currently examining these possibilities in more detail.

DISCUSSION

We have identified five evolutionarily conserved sites in the promoter of the FMR1 gene (Fig. 2). These include the four in vivo protein binding sites found in fibroblasts and lymphoid cells (5, 6), namely the E-box, two GC-boxes, and a α−Pal/Nrf-1 binding site. The conservation of these sites among species as evolutionarily divergent as primates, canids, and rodents, which last shared a common ancestor 80–120 million years ago, indicates that they are important for regulation of this gene. That the E-box and the α−Pal/Nrf-1 sites together can account for almost all of the transcriptional activity of the FMR1 promoter in the neuronally derived PC12 cells supports this contention (Fig. 3). Removing the TATA-like sequence just upstream of the transcription start site does not abolish gene activity nor does a deletion of an Initiator-like element close to the reported transcription start site. Initiation of transcription...
and the gel-attached probe (lane 5) or a probe that has been completely methylated using the Dam methylase (lane 6). The location of the free probe and the gel-attached probe (lane 5) or a probe that has been completely methylated using the Dam methylase (lane 6). The location of the free probe and the gel-attached probe (lane 5) or a probe that has been completely methylated using the Dam methylase (lane 6). The location of the free probe and the gel-attached probe (lane 5) or a probe that has been completely methylated using the Dam methylase (lane 6).

of the FMR1 gene may thus involve a novel mechanism whose molecular details remain to be elucidated.

We have identified three major transcription factors that seem to positively affect transcription. They are USF1 and USF2. Failure of these factors to bind the methylated promoter would be an additional component. If this is true, it would have important implications for the development of therapies aimed at reactivating FMR1 transcription in patients with fragile X syndrome.

Acknowledgments—We thank A. V. Furano, N. Nossal, and R. Owens of the Molecular Laboratory of Cellular Biology, NIDDK and B. Safer of the Molecular Hematology Branch, NHLBI, National Institutes of Health, for thoughtful reading of this manuscript.

REFERENCES

1. Kremer, E. J., Pritchard, M., Lynch, M., Yu, S., Holman, K., Baker, E., Warren, S. T., Schlessinger, D., Sutherland, G. R., and Richards, R. I. (1991) Science 252, 1711–1714

2. Verkerk, A. J., Pieretti, M., Sutcliffe, J. S., Fu, Y. H., Kuhl, D. P., Polsky, A., Reiner, O., Richards, S., Victoria, M. F., and Zhang, F. P. (1991) Cell 65, 905–914

3. Pieretti, M., Zhang, F. P., Fu, Y. H., Warren, S. T., Oostra, B. A., Caskey, C. T., and Nelson, D. L. (1991) Cell 66, 817–822

4. Hinds, H. L., Ashley, C. T., Sutcliffe, J. S., Nelson, D. L., Warren, S. T., Housman, D. E., Scott, M., and Schalling, M. (1991) Nature 347, 36–43

5. Davein, R., Angers, P., Bally, N., Rose, T. M., Khandjian, W., and Rousseau, F. (1997) Hum. Mol. Genet. 6, 2051–2060

6. Schwenk, S., de Graaff, K., Deissler, H., Glaser, D., Wohrle, D., Hameister, H., and Poustka, A. (1993) Hum. Mol. Genet. 2, 2043–2050

7. Coffee, B., Zhang, F., Fearon, S. T., and Reines, D. (1999) Nat. Genet. 13, 36–43

8. Breathnach, R., and Chambon, P. (1981) Nature 292, 349–383

9. Verkerk, A. J., Pieretti, M., Sutcliffe, J. S., Fu, Y. H., Kuhl, D. P., Polsky, A., Reiner, O., Richards, S., Victoria, M. F., and Zhang, F. P. (1991) Cell 65, 905–914

10. Lavedan, C., Grabczyk, E., Usdin, K., and Nussbaum, R. L. (1998) Hum. Mol. Genet. 7, 359–366

11. Dignam, J. D., Lebovitz, R. M., and Roeder, R. G. (1983) Nucleic Acids Res. 11, 1475–1489

12. Heggner, G., Matsu, K., Gassmann, M., Schaffner, W., Luscher, B., Rulicke, T., and Uguzi, A. (1995) Hum. Mol. Genet. 4, 359–366

13. Hwu, W. L., Wang, T. R., and Lee, Y. M. (1997) DNA Cell Biol. 16, 449–453

14. Breathnach, R., and Chambon, P. (1981) Annu. Rev. Biochem. 50, 349–383

derived cell line. We have also shown that Sp3, a member of the Sp family of ubiquitously expressed zinc finger factors that bind GC-boxes, binds to the SpI-PalI mouse promoter subfragment containing the conserved GC-boxes. Deletions that include one or the other of these GC-boxes do decrease promoter activity (Fig. 3), suggesting some contribution of these regions to optimal gene activity.

In liver, the amount of USF protein that binds to the promoter in gel shift or DNase I footprinting assays is reduced relative to that in brain and testis, and very little α–Pal/Nrf-1 binding is seen at all (Fig. 4, and data not shown). The high level of USF1 and USF2 in adult brain (Fig. 4, and Ref. 25) and the high levels of α–Pal/Nrf-1 (26) and the α–Pal/Nrf-1-dependent USF variant in adult testis relative to other organs may account for their relatively high levels of FMR1 mRNA.

Methylation of the FMR1 promoter in individuals with fragile X syndrome is correlated with reduced transcription of this gene. In such individuals no transcription factor binding to the promoter in vitro is seen (5, 6). This may reflect the binding of proteins such as MEC2P2, that recognize methylated CpGs and bind the promoter in a sequence-nonspecific manner (27). One consequence of this binding is thought to be the recruitment of histone deacetylases to the region and subsequent formation of transcriptionally silent chromatin containing a large proportion of deacetylated histones (27). The FMR1 promoter in cells from fragile X patients is associated with deacetylated histones, and reactivation of the promoter using the demethylating agent 5-azacytidine increases transcription (9, 28) and the proportion of acetylated histones (9). However, treatment of cells from fragile X patients with the deacetylase inhibitor trichostatin A, which increases the level of acetylated histones directly, has little (29) or no effect on reactivation of the gene (9). This has led to the suggestion that there might be an additional component to the methylation-dependent silencing of this gene (9). We have shown that methylation significantly reduces α–Pal/Nrf-1 binding and also affects binding of USF1 and USF2. Failure of these factors to bind the methylated promoter would be a positive effect. If this is true, it would have important implications for the development of therapies aimed at reactivating FMR1 transcription in patients with fragile X syndrome.
15. Carrillo, C., Cisneros, B., and Montanez, C. (1999) Neurosci. Lett. 276, 149–152
16. Feng, Y., Zhang, F., Lokey, L. K., Chastain, J. L., Lakkis, L., Eberhart, D., and Warren, S. T. (1995) Science 268, 731–734
17. Muller-Hartmann, H., Deissler, H., Naumann, F., Schmitz, B., Schroer, J., and Doerfler, W. (2000) J. Biol. Chem. 275, 6447–6452
18. Deleted in proof
19. Suske, G. (1999) Gene 238, 291–300
20. Oberle, I., Rousseau, F., Heitz, D., Kretz, C., Devys, D., Hanauer, A., Boue, J., Bertheas, M. F., and Mandel, J. L. (1991) Science 252, 1097–1102
21. Hwu, W. L., Lee, Y. M., Lee, S. C., and Wang, T. R. (1993) Biochem. Biophys. Res. Commun. 193, 324–329
22. Xiao, Q., and Ojamaa, K. (1998) J. Mol. Cell. Cardiol. 30, 87–95
23. Johnson, A. D., and Owens, G. K. (1999) Am. J. Physiol. 276, C1420–C1431
24. Efiok, B. J., Chiorini, J. A., and Safer, B. (1994) J. Biol. Chem. 269, 18921–18930
25. Sirito, M., Lin, Q., Maity, T., and Sawadogo, M. (1994) Nucleic Acids Res. 22, 427–433
26. Escriva, H., Rodriguez-Pena, A., and Vallejo, C. G. (1999) Biochimie (Paris) 81, 965–971
27. Nan, X., Ng, H. H., Johnsen, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N., and Bird, A. (1998) Nature 393, 386–389
28. Chiurazzi, P., Pomponi, M. G., Willemse, R., Oostra, B. A., and Neri, G. (1998) Hum. Mol. Genet. 7, 109–113
29. Chiurazzi, P., Pomponi, M. G., Pietrobono, R., Bakker, C. E., Neri, G., and Oostra, B. A. (1999) Hum. Mol. Genet. 8, 2317–2325
