NF-Y, AP2, Nrf1 and Sp1 regulate the fragile X-related gene 2 (*FXR2*)

Lata MAHISHI and Karen USDIN¹

Gene Structure and Disease Section, NIDDK (National Institute of Diabetes and Digestive and Kidney Diseases), NIH (National Institutes of Health), Bethesda, MD 20892-0830, U.S.A.

Fragile X syndrome, the most common heritable form of mental retardation, is caused by silencing of the *FMR1* (fragile X mental retardation-1 gene). The protein product of this gene, FMRP (fragile X mental retardation protein), is thought to be involved in the translational regulation of mRNAs important for learning and memory. In mammals, there are two homologues of FMRP, namely *FXR1P* (fragile X-related protein 1) and *FXR2P*. Disruption of *Fxr2* in mice produces learning and memory deficits, and *Fmr1* and *Fxr2* double-knockout mice have exaggerated impairments in certain neurobehavioral phenotypes relative to the single gene knockouts. This has led to the suggestion that *FMR1* and *FXR2* functionally overlap and that increasing the expression of *FXR2P* may ameliorate the symptoms of an FMRP deficiency. Interestingly, the region upstream of the *FXR2* translation start site acts as a bidirectional promoter in rodents, driving transcription of an alternative transcript encoding the ABP (androgen-binding protein) [*aABP* (alternative ABP promoter)]. To understand the regulation of the human *FXR2* gene, we cloned the evolutionarily

conserved region upstream of the *FXR2* translation start site and showed that it also has bidirectional promoter activity in both neuronal and muscle cells as evidenced by luciferase reporter assay studies. Alignment of the human, mouse, rat, rabbit and dog promoters reveals several highly conserved transcription factor-binding sites. Gel electrophoretic mobility-shift assays, chromatin immunoprecipitation studies and co-transfection experiments with plasmids expressing these transcription factors or dominant-negative versions of these factors showed that NF-YA (nuclear transcription factor Y α), AP2 (activator protein 2), Nrf1 (nuclear respiratory factor/ α -Pal) and Sp1 (specificity protein 1) all bind to the *FXR2* promoter both *in vitro* and *in vivo* and positively regulate the *FXR2* promoter.

Key words: chromatin immunoprecipitation, fragile X mental retardation-1 gene (*FMR1*), fragile X-related gene 2 (*FXR2*), nuclear factor Y, promoter regulation, transcription factor.

INTRODUCTION

FXS (fragile X syndrome), the most common heritable form of mental retardation and the most common known cause of autism, belongs to a group of neurological diseases caused by expansion of a specific tandem repeat tract [1]. Expansion of a CCG·CCG repeat tract in the 5'-UTR (5'-untranslated region) of the *FMR1* (fragile X mental retardation-1 gene) gene to > 200 repeats leads to promoter heterochromatinization [2,3]. This, in combination with difficulties in the translation of any residual mRNA produced from the expanded allele, leads to reduced levels of the *FMR1* product, FMRP (fragile X mental retardation protein) [4]. FMRP is thought to be involved in normal synaptic function and thus is obviously crucial for normal brain function. However, in addition to symptoms related to the absence of FMRP in the CNS (central nervous system), individuals with FXS also show symptoms like macroorchidism, digestive difficulties and hypotonia that point to roles for FMRP in places outside of the CNS like testes, gut and muscle.

Two autosomal homologues of *FMR1*, *FXR1* (fragile X-related gene 1) and *FXR2*, have been identified [5]. FMRP shares 86 and 70% amino acid identity with the N-terminal regions of *FXR1P* (Fragile X-related protein 1) and *FXR2P* respectively [6]. Like FMRP, *FXR1P* and *FXR2P* are ubiquitously expressed during embryonic development and their tissue distribution overlaps with FMRP in adults [7]. They form homo- and hetero-multimers with one another [5]. The *Fxr2* knockout mice share some behavioural similarities with *Fmr1* knockout mice such as hyperactivity and impaired motor co-ordination and learning [6]. These neuro-

behavioral deficits are exaggerated in *Fxr2/Fmr1* double-knockout mice, indicating a co-operative contribution of *Fmr1* and *Fxr2* genes to these behaviours [8]. This has led to the suggestion that FMRP and *FXR2P* have overlapping functions and that *FXR2* may thus be capable of compensating, at least partially, for the loss of *FMR1* function in FXS. This is of interest because the poor translatability of mRNAs from FXS alleles limits approaches aimed at re-activating the *FMR1* gene [4,9,10]. The *FXR2* gene is also interesting in its own right, since its disruption in mice produces learning and memory impairments distinct from those seen in mice lacking FMRP [6].

While many factors important for the regulation of the *FMR1* gene have been identified [11–16], nothing is known about the regulation of the *FXR2* gene. Despite the conservation in their coding sequences and the similarity in their pattern of tissue expression, the promoters of these genes are not similar at the sequence level. Substantial sequence similarity exists in the non-coding regions of the *FXR2* genes in human and mouse [17] and this region overlaps with the *aABP* [alternative ABP (androgen-binding protein) promoter] that is used in rodents to transcribe the oppositely oriented *ABP* gene [18]. To characterize the *FXR2* promoter more fully, we compared the sequence of the human promoter with that of mouse, rat, rabbit and dog and identified a number of evolutionarily conserved regions. We demonstrated that the region upstream of the *FXR2* translation start site in humans also has bidirectional promoter activity in cells derived from both brain and muscle. We identified the major transcription factors that bind to the human *FXR2* promoter *in vitro* by gel shift assays and confirmed their *in vivo* association by ChIP

Abbreviations used: ABP, androgen-binding protein; *aABP*, alternative ABP promoter; AP2, activator protein 2; ChIP, chromatin immunoprecipitation; CNS, central nervous system; DMEM, Dulbecco's modified Eagle's medium; EMSA, electrophoretic mobility-shift assay; *FMR1*, fragile X mental retardation-1 gene; FMRP, fragile X mental retardation protein; *FXR1P*, fragile X-related protein 1; *FXR2*, fragile X-related gene 2; FXS, fragile X syndrome; NF-Y, nuclear transcription factor Y; Nrf1, nuclear respiratory factor/ α -Pal; Sp1, specificity protein 1; SV40, simian virus 40; UTR, untranslated region.

¹ To whom correspondence should be addressed (email ku@helix.nih.gov).

(chromatin immunoprecipitation) assay. The contribution of these transcription factors to promoter activity was examined using co-transfections of reporter constructs with plasmids expressing wild-type and dominant-negative versions of these transcription factors.

MATERIALS AND METHODS

Chemicals and reagents

[α - 32 P]dCTP (6000 Ci/mmol) and [γ - 32 P]GTP (6000 Ci/mmol) were purchased from MP Biomedicals (Irvine, CA, U.S.A.). The reporter gene vectors (pGL3-Basic, pGL3-Control and pRL-null) and the Dual Luciferase system were purchased from Promega (Madison, WI, U.S.A.). The transfection reagents FuGENE[®] 6 was purchased from Roche (Indianapolis, IN, U.S.A.) and Lipofectamine[™] 2000 and Cellfectin were purchased from Invitrogen (Carlsbad, CA, U.S.A.). The plasmid constructs Δ 4NF-YA13m29, dn-Nrf1, pSAP2B and pcDNA6-AP2DBD were generously provided by Dr R. Mantovani (Dipartimento di Scienze Biomolecolari e Biotecnologie, Università degli Studi di Milano, Milano, Italy), Dr R. Scarpulla (Department of Cell and Molecular Biology, Northwestern Medical School, Chicago, IL, U.S.A.), Dr M. Tainsky (Program in Molecular Biology and Genetics, Karmanos Cancer Institute and Wayne State University, Detroit, MI, U.S.A.) and Dr P. Kannan (MetroHealth Medical Center, Case Western Reserve University, Cleveland, OH, U.S.A.) respectively. Plasmids pPacSp4, pPacSp3 and pPacUSp3 that express Sp4, and the short and long isoforms of Sp3 respectively, were a gift from Professor G. Suske (Institut für Molekularbiologie und Tumorforschung, Philipps-Universität Marburg, Marburg, Germany). Plasmids pPac-Sp1 and pPac-Sp1DBD were a gift from Professor T. Johansen (Department of Biochemistry, University of Tromsø, Tromsø, Norway). The anti-Nrf1 (nuclear respiratory factor/ α -Pal) antibody was a gift from Dr R. Scarpulla. The anti-AP2 (activator protein 2), anti-NF-YA (nuclear transcription factor Y α) and anti-NF-YB antibodies and the rabbit and goat pre-immune sera were purchased from Santa Cruz Biotechnology (Santa Cruz, CA, U.S.A.). Anti-Sp1 (specificity protein 1) antibody was obtained from Upstate (Charlottesville, VA, U.S.A.) and Sp1 protein was purchased from Promega. Synthetic oligonucleotides were obtained from IDT (Coralville, IA, U.S.A.) and sequencing services were provided by MWG-Biotech (Highpoint, NC, U.S.A.). All other chemicals were purchased from Sigma (St. Louis, MO, U.S.A.) unless specifically noted otherwise.

Cloning of the human and rabbit *FXR2/aABP* promoter and construction of different plasmids

The human *FXR2* proximal promoter spanning nt -898 to +214 relative to the translation start site was PCR-amplified from human genomic DNA (negative strand of Chr17: 7458362-7459473, human genome May 2004 assembly) using forward primer 1 and reverse primer 2 (Table 1) (<http://genome.ucsc.edu/>) [19,20]. The PCR product was digested with SphI and PstI to generate a 706 bp fragment that spanned -742 to -37 bp with respect to the translation start site and was subcloned into the promoterless pGL3-Basic vector to generate p*FXR2*. Several 5' and 3' deletion fragments of the *FXR2/aABP* promoter were obtained by restriction digestion and religation. The two CCAAT-box mutants and the constructs containing deletions of the CGG tracts were generated by site-directed mutagenesis. These constructs were verified by DNA sequencing. The construction of p*FMR1* (aka pGL-FMR) has been described earlier [11].

Table 1 Oligonucleotides used

	5'-3' Sequence
PCR primers	
Forward 1 (-898 to -878 bp)	TTTTTGAGCTCCGGCGATGCATACGCGAAGAG
Reverse 2 (+185 to +204 bp)	TTTTTAAGCTTCGGGTTGGAGCAAGGGTCCA
Forward 3 (-657 to -638 bp)	CCCAATCCGATGCAGCGAGC
Reverse 4 (+46 to +64 bp)	TGGAGCCGCGACCTCGAC
Forward 5 (-690 to -656 bp)	TGGCCTGAGCGCTGTGTACTGTGCGGACGCCGCC
Reverse 6 (-655 to -621 bp)	GGGGAGGGAGGGCGCGGCTCGCTCAGTCCGATTG
Forward 7 (-655 to -621 bp)	CAATCGACTGAGCGAGCCCGGCCCTCCCTCCCC
Reverse 8 (-596 to -620 bp)	ACAACACTGAGAACCAACCAATCAG
Reverse 9 (-526 to -505 bp)	GTCGCTTGCTAGCGCTTGCAGCA
Forward 10 (-620 to -597 bp)	CTGATTGGTTGGTTCTCAGTGTGT
Reverse 11 (-318 to -298 bp)	TCCCGGAGCCCGCTGCTGCCG
Forward 12 (-305 to -288 bp)	TCCCGGAGCTGAGGCGAGCAGC
16cen-F	GTCTCTTCTTGTTTTAAAGCTGGG
16cen-R	TGAGCTCATTGAGACATTGG
EMSA competitor oligonucleotides (oligos)	
NF-Y oligo	CATTTTCTGATTGGTTAAAGTTG
Mut NF-Y oligo 1	ACGCCGCTTAATCGGACTGAG
Mut NF-Y oligo 2	CTCCCGCTGATTAATTGGTTC
AP2 oligo	GATCGAAGTACCGCCCGCGGCCCGT
Mut AP2 oligo 1	CCCCCTTTGGCTCCTCCTCCG
Mut AP2 oligo 2	CCGGCTCCTCTTTGGGGAGA
Nrf1 oligo	GATGACTGCGCGACGCGCGAGCTAGC
Mut Nrf1 oligo	CGGTGCGCCTTTGAGTAGGC
Sp1 oligo	ATTGATCGGGCGGGCGGAGC

The rabbit *FXR2* promoter region was PCR-amplified using the forward primer 3 and reverse primer 4 (Table 1) and cloned (GenBank[®] accession no. DQ322245) into the promoterless pGL3-basic vector. The *FXR2* promoter sequences for mouse (positive strand of Chr11: 69358267-69358940, mouse genome March 2005 assembly), rat (positive strand of Chr10: 5646328-56463961, rat genome June 2003 assembly) and dog (negative strand of chr5: 35496411-35497106, dog genome July 2004 assembly) were obtained from the public domain databases (<http://genome.ucsc.edu/>) [19,20]. The sequences were aligned using MultAlign software [21]. The transcription factor-binding sites were analysed using MatInspector [22], rVista [23] and TESS [24] software.

Transient transfections and reporter gene assays

Mouse C2C12 and human SK-N-MC cells were seeded in 24-well plates containing DMEM (Dulbecco's modified Eagle's medium; Invitrogen Life Technologies, Carlsbad, CA, U.S.A.) supplemented with 10% (v/v) fetal bovine serum. The initial plating density was 0.5×10^5 and 2×10^5 cells/well respectively. Co-transfections were done with various test constructs together with pRL-null, a plasmid expressing the *Renilla* luciferase gene to normalize for differences in the transfection efficiency. Transfections were carried out using FuGENE[™] 6 in C2C12 cells and using Lipofectamine[™] 2000 in SK-N-MC cells according to the supplier's instructions. *Drosophila* S2 cells were plated at a density of 4×10^5 /well in 400 μ l of serum-free Schneider medium (Invitrogen) and were grown at room temperature (20°C) overnight in 24-well plates. Cells were transfected using Cellfectin according to the manufacturer's instructions.

Luciferase activity was measured 36 h post-transfection using the Dual Luciferase Reporter assay system and a MicroLumat LB 96 P luminometer (Berthold Systems, Aliquippa, PA, U.S.A.). The luciferase activity from the reporter construct in each transfection was normalized to pRL-null.

The complete human nucleotide sequence is given with gaps in the alignment shown as dashes. The bases in the sequence of the remaining four species that are identical with the human sequence are shown as dots with only those bases that differ from the human sequence shown. The putative transcription factor-binding sites are marked by rectangles. The star symbol indicates that no sequence is available for the 5'-end of the rabbit *FXR2* promoter.

Nuclear extracts from the human neuronal SK-N-MC cell line were prepared using 'NE-PER nuclear and cytosolic extraction reagents' following the protocol provided by the manufacturer (Pierce, Rockford, IL, U.S.A.). The pFXR2 plasmid was digested with SstI and ApaI generating a 343 bp DNA fragment, fragment I (Figure 3A). This plasmid was also digested with ApaI and

HindIII to generate a 367 bp DNA fragment, fragment II (Figure 3A). Fragment I and fragment II were labelled using [α - 32 P]dCTP and Klenow fragment. Fragment III spanning -690 to -621 bp was PCR-amplified using pFXR2 as template together with forward primer 5 and reverse primer 6 (Table 1; Figure 3A). The fragment IV spanning -655 to -596 bp was PCR-amplified using pFXR2 as template with forward primer 7 and reverse primer 8 (Table 1; Figure 3A). Fragments III and IV were

labelled using [γ - 32 P]GTP and T4 polynucleotide kinase and purified using MicroSpin G-25 columns (Amersham Biosciences, Piscataway, NJ, U.S.A.). Binding was carried out at 20°C in 30 μ l of reaction buffer containing 25 mM Hepes (pH 7.5), 100 mM NaCl, 1 mM EDTA and 4 mM dithiothreitol with 6 fmol of labelled probe, 2 μ g of protein and 1 μ g of poly(dA-dT)·(dA-dT) in the presence of 1000-fold excess of specific and non-specific competitors. For antibody supershift assays, 10 μ g of antibody or normal rabbit/goat pre-immune serum was used in the reactions. The purified Sp1 protein binding reactions were carried out as per the supplier's instructions. Reactions were subjected to electrophoresis on 4 or 5% polyacrylamide gels (60:1, acrylamide/bisacrylamide) containing 1.6% (v/v) glycerol in 1× Tris/glycine/EDTA buffer (50 mM Tris, 380 mM glycine and 2.1 mM EDTA, pH 8.5).

ChIP

The Upstate Biotechnology (Charlottesville, VA, U.S.A.) ChIP assay kit was used according to the manufacturer's instructions with slight modifications. Briefly, 2×10^6 cells were cross-linked in 200 μ l of DMEM with 1% formaldehyde at 37°C for 5 min. The cells were washed in PBS and lysed in 200 μ l of a buffer containing 1% SDS, 10 mM EDTA and 50 mM Tris/HCl (pH 8.0). The chromatin from lysed cells was sonicated to lengths between 200 and 1000 bp and the cell debris was removed by centrifugation. The sonicated lysates were diluted to 2 ml using ChIP dilution buffer containing 0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris/HCl (pH 8.1) and 167 mM NaCl. The lysate was precleared, an aliquot of 100 μ l was removed for processing as 'Input', a control for the efficiency of the PCR, and the remaining 1.9 ml was incubated with the appropriate antibody or with pre-immune serum at 4°C overnight. The immunocomplexes were recovered by Protein A-agarose, and this 'Bound' material was washed and eluted in 1% SDS and 0.1 M NaHCO₃. The 'Input' sample and the 'Bound' samples were processed in parallel by reversing the cross-linking by incubation at 65°C for 4 h, deproteinizing with phenol/chloroform and ethanol precipitation in the presence of Pelletpaint (EMD Biosciences, San Diego, CA, U.S.A.). The 'Input' material was resuspended in 100 μ l of 1× TE (10 mM Tris/HCl, 1 mM EDTA, pH 8.0) and the 'Bound' material was resuspended in 30 μ l of the 1× TE. In each *FXR2* promoter and 16 centromere PCR reaction, 1 μ l of 'Input' DNA and 4 μ l of 'Bound' DNA was used. For the *FXR2* gene, the PCR primers used for Sp1 and NF-Y ChIP analysis were forward primer 1 and reverse primer 9 (Table 1); for Nrf1 ChIP analysis, forward primer 10 and reverse primer 11 were used, and for AP2 ChIP PCR, forward primer 12 and reverse primer 2 (Table 1) were used. The centromeric region of chromosome 16 was used as a negative control [25]. The h16CEN-F and h16CEN-R primers were used to amplify the 16 centromere region [25]. In each case roughly 250 times more of the starting material was used for PCR of the 'Bound' sample than the 'Input' sample. ChIP experiments were repeated three times with each antibody for the *FXR2* promoter and at least twice for the negative control.

RESULTS AND DISCUSSION

Sequence conservation in the *FXR2* promoter

The approx. 700 bp region immediately upstream of the *FXR2* translation start site shows substantial sequence conservation among various organisms (Figure 1). The sequence conservation decreases significantly upstream of this region (results

Figure 2 Comparative analysis of the human *FXR2/aABP* and *FMR1* promoters in C2C12 and SK-N-MC cells

Cells were transfected with the pFMR1, pFXR2 and paABP reporter constructs as described in the Materials and methods section. The luciferase activity was plotted as a percentage of the luciferase activity obtained for the pGL3-Control vector which contains a SV40 promoter and enhancer. The results shown represent the average of three independent experiments performed in duplicate.

not shown) due primarily to the presence of multiple order- or species-specific SINEs (short interspersed repeated elements). These elements have presumably been inserted into this region since these animals last shared a common ancestor before the mammalian radiation approx. 65 million years ago.

The human *FXR2/aABP* promoter region

The 706 bp region upstream of the human *FXR2* translation start site has a G + C content of approx. 70%. This is consistent with the observation that many other bidirectional promoters have a higher than average G + C content. Like many known bidirectional promoters [26], there are no canonical TATA box or initiator elements. Like many TATA-less promoters [27], the human *FXR2* promoter region has one CCAAT box (−656/−652 bp) and one inverted CCAAT box (−617/−613 bp). Along with two CCAAT boxes, the 706 bp human promoter region between nucleotide positions −742 to −37 contains various other potential transcription factor-binding sites (see Figure 3A) including two GC boxes (−595/−586 bp and −181/−172 bp), one Nrf1-binding site (−373/−362 bp) and two adjacent AP2-binding sites (−105/−76 bp). The human *FXR2* promoter also contains two stretches of (CGG·CCG)₄ repeats, one upstream of the *FXR2* transcription start site and the second in the 5'-UTR. These CGG·CCG repeats have been suggested to influence both transcription [28] and translation [29]. These sites are conserved in all five organisms (Figure 1).

Functional analysis of the human *FXR2* promoter

The activity of the 706 bp human DNA fragment upstream of the *FXR2* translation start site was examined using a luciferase reporter in transient transfection assays. For these experiments we used human SK-N-MC cells that are neuronal in origin and mouse C2C12 cells that are derived from skeletal muscle. These cells were chosen since while *FXR2* is ubiquitously expressed during fetal development, significant expression in adults is restricted to a few tissues that include brain and skeletal muscle [7]. The promoter was active in both directions in both cell lines (Figure 2) as it is in rat liver and brain where it drives production of both *FXR2* mRNA and an alternative transcript encoding the *ABP/SHBG* (sex hormone-binding globulin) [30,31]. The

Figure 3 Functional analysis of the human *FXR2/aABP* promoter

(A) Schematic diagram of the human *FXR2* promoter region. Fragments I–VI were used in the EMSA experiments described in Figure 4. (B, C) C2C12 (B) and SK-N-MC (C) cells were transfected with the various constructs shown on the left-hand side of the Figure and assayed for luciferase activity as described in the Materials and methods section. The luciferase activity was plotted as a percentage of the luciferase activity obtained for the full-length promoter construct. The results shown represent the average of three independent experiments performed in duplicate. Asterisks mark the location of point mutations in the constructs. Empty boxes mark the deletion of promoter construct.

Figure 4 EMSA of the *FXR2/aBP* promoter region using neuronal cell extracts

The probes used are described in the Material and methods section and are schematically shown in Figure 3(A). (A, B) With the exception of the lane marked '—', which contains the probe alone, all lanes contain the probe mixed with the nuclear extract. The lane marked '+' contains the probe, the nuclear extract and a non-specific competitor DNA as described in the Materials and methods section. The remaining lanes contain, in addition to the non-specific competitor DNA, the indicated specific competitors or antibodies. C-1, C-2 and C-3 indicate the DNA–protein complexes formed and 'SS' indicates the supershifted DNA–protein–antibody complex. The arrowheads indicate the non-specific DNA–protein complexes. In Figure 3(B) (lanes 9–13), lower exposure time is used for clarity of DNA–protein complex C3. Hence the DNA–protein complex C2 is not seen in these lanes. (C) With the exception of the lane marked '—', which contains the probe alone, all lanes contain the probe mixed with purified Sp1 protein. The lane marked '+' contains the probe, the pure Sp1 protein and a non-specific competitor DNA as described in the Materials and methods section. The remaining lanes contain, in addition to the non-specific competitor DNA, the consensus Sp1-binding oligonucleotide.

human *ABP* gene also has two promoters and multiple *ABP* transcripts have been reported in human liver [32,33], but, to date, there is no evidence of any *ABP* transcripts that initiate in this region in this organ and it remains to be seen if this region acts as an *ABP* promoter in human brain. However, the fact that both the human and rat *FXR2* promoters can function bidirectionally suggests that the ability to drive transcription in both directions was probably a feature of the promoter in the common ancestor of these organisms. Since many bidirectional promoters show co-ordinated expression of the two transcripts, the bidirectionality of the rodent *FXR2/ABP* gene may reflect the original co-regulation of these two genes. The insertion of multiple retrotransposable elements into this region since the rodent and primate lineages diverged may have disrupted the ability of the human promoter to function as an alternative promoter for the *ABP*

gene, necessitating the recruitment of a novel promoter element closer to the human *ABP* open reading frame.

The strength of the promoter acting in the *FXR2* direction was > 50-fold higher than the vector backbone in SK-N-MC cells (results not shown) and slightly lower than that in C2C12 cells. This corresponds to approx. 35 % and approx. 25 % of the activity of the moderately robust SV40 (simian virus 40) early promoter/enhancer in SK-N-MC and C2C12 cells respectively (Figure 2). This activity was comparable with that seen for the *FMR1* promoter under the same conditions (Figure 2). The activity of the promoter in the *aABP* direction was similar in C2C12 cells, but was significantly lower in SK-N-MC cells.

We generated a series of plasmid constructs that contained 5' or 3' deletions of this region or different point mutations and tested their activity in both SK-N-MC and C2C12 cells. Most of

the mutations showed the same general trend in both cell lines (Figures 3B and 3C). The one exception was a point mutation in one of the CCAAT boxes (paABP- Δ 2). This mutation had a negative effect on the aABP promoter in muscle cells but a positive effect in neuronal cells. Deletion of the conserved Nrf1-binding site had a particularly large impact on the *FXR2* promoter (pFXR2- Δ 2) in SK-N-MC cells (Figure 3C).

Identification of transcription factors binding to the *FXR2*/aABP promoter *in vitro*

To examine the transcription factors bound to the *FXR2*/aABP promoter, EMSAs were performed using SK-N-MC nuclear extracts and two different DNA fragments. The $-742/-404$ bp fragment I formed one DNA-protein complex with the nuclear extract from SK-N-MC cells (Figure 4A, lane 2). This complex was eliminated by an excess of either of two oligonucleotides, $-689/-621$ or $-655/-596$ bp (Figure 4A, lanes 6 and 7). The $-689/-621$ bp DNA fragment contains the CCAAT box and the $-655/-596$ bp fragment contains the inverted CCAAT box (Figure 3A). When used in EMSAs, the major protein-DNA complex formed using each of these two oligonucleotides was supershifted with antibody to NF-YB, a subunit of the trimeric NF-Y transcription factor (Figure 4A, lanes 12 and 18). An excess of unlabelled oligonucleotide containing the consensus CCAAT box eliminated the same DNA-protein complex (Figure 4A, lanes 10 and 16) and the complex was unaffected by an excess of oligonucleotide with the mutated CCAAT box (Figure 4A, lanes 11 and 17). Another DNA-protein complex was also eliminated by consensus CCAAT box (Figure 4A, lanes 10 and 16) and mutated CCAAT box oligonucleotides (Figure 4A, lanes 11 and 17). However, the same effect was seen with unrelated oligonucleotides, indicating that binding was not specific.

Fragment II ($-404/-37$) formed two major DNA-protein complexes (Figure 4B, lane 2). Antibody to AP2 eliminated complex 2 (Figure 4B, lane 4), while antibody to Nrf1 supershifted complex 3 (Figure 4B, lane 10). An excess of unlabelled oligonucleotide containing a consensus AP2-binding site and Nrf1-binding site eliminated complex 2 (Figure 4B, lane 3) and complex 3 respectively (Figure 4B, lane 9). Complexes 2 and 3 were unaffected by an excess of oligonucleotides with the mutated AP2 (Figure 4B, lane 7-8) and mutated Nrf1-binding site (Figure 4B, lane 13) respectively.

Purified Sp1 protein formed a DNA-protein complex with fragment I (Figure 4C, lane 2) and fragment II (Figure 4C, lane 5). Each of these complexes was competed by an excess of a consensus Sp1-binding oligonucleotide (Figure 4C, lanes 3 and 6).

Transcription factors binding to *FXR2*/aABP promoter *in vivo*

To determine if these four transcription factors are bound to the *FXR2*/aABP promoter *in vivo*, ChIP was performed using chromatin prepared from the neuronal cell line SK-N-MC. A region of the centromere of chromosome 16 was chosen as a negative control since it is known to be heterochromatinized and transcriptionally silent. While the amount of Nrf1, AP2 and NF-YA/NF-YB bound to the negative control amplicon was relatively low as expected, the amount of material immunoprecipitating with Sp1 was somewhat higher. Since Sp1 associates with histone deacetylases and can serve a repressive function, this binding may not be non-specific. Whatever the source of this binding, Sp1 binding to the *FXR2* promoter region was significantly higher than this as illustrated by the significant increase in the ratio of Bound to Input DNA amplified (Figure 5). Significant enrichment of the *FXR2* promoter region was also seen when antibodies to NF-YA/NF-YB or AP2 were used in the ChIP assay (Figure 5). A

Figure 5 *In vivo* association of NF-Y, Nrf1, AP2 and Sp1 with the *FXR2*/aABP promoter

(A) ChIP was carried out on neuronal cells using either pre-immune serum (PS) or antibody as indicated and described in the Materials and methods section. DNA isolated from immunoprecipitated material was amplified by PCR with primers (lane B: Bound). An aliquot taken from the chromatin sample prior to immunoprecipitation was also amplified as a control for the PCR (lane I: Input). The human chromosome 16 centromere region was used as a negative control. (B) ChIP experiments were repeated three times with each antibody for the *FXR2* promoter and at least twice for the negative control. Differences significant at $P < 0.05$ are indicated by asterisks.

much smaller enrichment of the *FXR2* promoter region was seen when the Nrf1 antibody was used (Figure 5).

The role of various transcription factors in regulation of the *FXR2*/aABP promoter

To test the role of NF-Y, Nrf1 and AP2 in the regulation of the *FXR2*/aABP promoter, we co-transfected our reporter constructs with dominant-negative versions of these transcription factors. In C2C12 cells both the *FXR2* and the aABP promoter activities were negatively affected by the dominant-negative forms of all three transcription factors (Figure 6A). This suggests that these factors act directly or indirectly to positively regulate transcription of both promoters in these cells. However, in SK-N-MC cells the dominant-negative version of NF-YA had little or no effect on the *FXR2* promoter but led to an increase in the activity of the

Figure 6 Effect of co-transfection of various transcription factors and dominant-negative versions of these factors on *FXR2* and *aABP* promoter activity

(A) *pFXR2* or *paABP* plasmid (200 ng) was transfected into C2C12 cells using FuGENE™ 6 with 600 ng of an empty cloning vector (–) or the same vector encoding a dominant-negative form of Nrf1, NF-YA and AP2. The luciferase activity was plotted as a percentage of the luciferase activity of the promoter construct with the empty vector. (B) The *pFXR2* or *paABP* plasmids (400 ng) were transfected into SK-N-MC cells using Lipofectamine™ 2000 with 600 ng of either an empty cloning vector (–) or the same vector carrying a dominant-negative form of Nrf1, NF-YA and AP2. The luciferase activity was plotted as a percentage of the luciferase activity of the promoter construct with the empty vector. (C) *pFXR2* or *paABP* plasmid (200 ng) was transfected into *Drosophila* S2 cells using Cellfectin with 250 ng of either the empty vector or the indicated Sp factor expression plasmids. The activity from the cells transfected with the full-length promoter construct and the backbone of the expression plasmid DNA was set to 1, and the luciferase activity obtained with the different Sp constructs was expressed relative to this value. The results shown represent the average of three independent experiments performed in duplicate.

aABP promoter. As in C2C12 cells the dominant-negative form of Nrf1 had a negative effect on the activity of both promoters.

To assess the role of Sp factors in *FXR2/aABP* promoter activity, we transfected the *FXR2* and *aABP* promoter plasmids into *Drosophila* S2 cells, which lack endogenous Sp factors. The expression of full-length Sp factors, Sp1 or Sp3 resulted in increased promoter activity in both orientations (Figure 6C).

There was no increase in *FXR2* or *aABP* promoter activity when co-transfections were done with plasmids expressing either the DNA-binding domains of Sp1 or Sp3 or full-length Sp4.

Nrf1 is highly expressed in brain [34] and it has strong homology to the *Drosophila* *ewg* (erect wing) gene that is required for normal central nervous development [35]. NF-Y and Sp1 are ubiquitous transcription factors consistent with the ubiquitous tissue distribution of FXR2P during embryonic development. AP2 is highly expressed in the CNS and is critical for normal neuronal development [36]. Our results thus suggest that despite the lack of extensive sequence conservation, the *FMR1* and *FXR2* promoters may be more similar than they appear: they are both TATA-less promoters that are regulated by Sp1, Nrf1 and AP2 [11–15] and have CGG · CCG repeats in 5'-UTR.

The knowledge of *FXR2* promoter regulation gained from these studies may help us to devise ways to up-regulate this gene and thus perhaps to ameliorate the effects of the FMRP deficit in FXS.

We thank Roland Owens (NIDDK, NIH) for his careful reading of this paper. This research was supported by the Intramural Research Program of the NIDDK, NIH.

REFERENCES

- Verkerk, A. J., Pieretti, M., Sutcliffe, J. S., Fu, Y. H., Kuhl, D. P., Pizzuti, A., Reiner, O., Richards, S., Victoria, M. F., Zhang, F. P. et al. (1991) Identification of a gene (*FMR-1*) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* **65**, 905–914
- Sutcliffe, J. S., Nelson, D. L., Zhang, F., Pieretti, M., Caskey, C. T., Saxe, D. and Warren, S. T. (1992) DNA methylation represses *FMR-1* transcription in fragile X syndrome. *Hum. Mol. Genet.* **1**, 397–400
- Coffee, B., Zhang, F., Ceman, S., Warren, S. T. and Reines, D. (2002) Histone modifications depict an aberrantly heterochromatinized *FMR1* gene in fragile X syndrome. *Am. J. Hum. Genet.* **71**, 923–932
- Feng, Y., Zhang, F., Lokey, L. K., Chastain, J. L., Lakkis, L., Eberhart, D. and Warren, S. T. (1995) Translational suppression by trinucleotide repeat expansion at *FMR1*. *Science* **268**, 731–734
- Zhang, Y., O'Connor, J. P., Siomi, M. C., Srinivasan, S., Dutra, A., Nussbaum, R. L. and Dreyfuss, G. (1995) The fragile X mental retardation syndrome protein interacts with novel homologs *FXR1* and *FXR2*. *EMBO J.* **14**, 5358–5366
- Bontekoe, C. J. M., McIlwain, K. L., Nieuwenhuizen, I. M., Yuva-Paylor, L. A., Nellis, A., Willemsen, R., Fang, Z., Kirkpatrick, L., Bakker, C. E., McAninch, R. et al. (2002) Knockout mouse model for *Fxr2*: a model for mental retardation. *Hum. Mol. Genet.* **11**, 487–498
- Bakker, C. E., de Diego Otero, Y., Bontekoe, C., Ragho, P., Luteijn, T., Hoogeveen, A. T., Oostra, B. A. and Willemsen, R. (2000) Immunocytochemical and biochemical characterization of FMRP, *FXR1P*, and *FXR2P* in the mouse. *Exp. Cell Res.* **258**, 162–170
- Spencer, C. M., Serysheva, E., Yuva-Paylor, L. A., Oostra, B. A., Nelson, D. L. and Paylor, R. (2006) Exaggerated behavioral phenotypes in *Fmr1/Fxr2* double knockout mice reveal a functional genetic interaction between Fragile X-related proteins. *Hum. Mol. Genet.* **15**, 1984–1994
- Tassone, F., Hagerman, R. J., Taylor, A. K., Gane, L. W., Godfrey, T. E. and Hagerman, P. J. (2000) Elevated levels of *FMR1* mRNA in carrier males: a new mechanism of involvement in the fragile-X syndrome. *Am. J. Hum. Genet.* **66**, 6–15
- Primerano, B., Tassone, F., Hagerman, R. J., Hagerman, P., Amaldi, F. and Bagni, C. (2002) Reduced *FMR1* mRNA translation efficiency in fragile X patients with premutations. *RNA* **8**, 1482–1488
- Kumari, D. and Usdin, K. (2001) Interaction of the transcription factors USF1, USF2, and alpha-Pal/Nrf-1 with the *FMR1* promoter – implications for Fragile X mental retardation syndrome. *J. Biol. Chem.* **276**, 4357–4364
- Kumari, D., Gabriellian, A., Wheeler, D. and Usdin, K. (2005) The roles of Sp1, Sp3, USF1/USF2 and NRF-1 in the regulation and three-dimensional structure of the Fragile X mental retardation gene promoter. *Biochem. J.* **386**, 297–303
- Smith, K. T., Coffee, B. and Reines, D. (2004) Occupancy and synergistic activation of the *FMR1* promoter by Nrf-1 and Sp1 *in vivo*. *Hum. Mol. Genet.* **13**, 1611–1621
- Carrillo, C., Cisneros, B. and Montanez, C. (1999) Sp1 and AP2 transcription factors are required for the human fragile mental retardation promoter activity in SK-N-SH neuronal cells. *Neurosci. Lett.* **276**, 149–152
- Lim, J. H., Booker, A. B., Luo, T., Williams, T., Furuta, Y., Lagutin, O., Oliver, G., Sargent, T. D. and Fallon, J. R. (2005) AP-2 alpha selectively regulates fragile X mental retardation-1 gene transcription during embryonic development. *Hum. Mol. Genet.* **14**, 2027–2034

- 16 Smith, K. T., Nicholls, R. D. and Reines, D. (2006) The gene encoding the fragile X RNA-binding protein is controlled by nuclear respiratory factor 2 and the CREB family of transcription factors. *Nucleic Acids Res.* **34**, 1205–1215
- 17 Kirkpatrick, L. L., McIlwain, K. A. and Nelson, D. L. (2001) Comparative genomic sequence analysis of the FXR gene family: FMR1, FXR1, and FXR2. *Genomics* **78**, 169–177
- 18 Joseph, D. R. (1998) The rat androgen-binding protein (ABP/SHBG) gene contains triplet repeats similar to unstable triplets: evidence that the ABP/SHBG and the fragile X-related 2 genes overlap. *Steroids* **63**, 2–4
- 19 Kent, W. J. (2002) BLAT – the BLAST-like alignment tool. *Genome Res.* **12**, 656–664
- 20 Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M. and Haussler, D. (2002) The human genome browser at UCSC. *Genome Res.* **12**, 996–1006
- 21 Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16**, 10881–10890
- 22 Quandt, K., Frech, K., Karas, H., Wingender, E. and Werner, T. (1995) MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res.* **23**, 4878–4884
- 23 Loots, G. G. and Ovcharenko, I. (2004) rVISTA 2.0: evolutionary analysis of transcription factor binding sites. *Nucleic Acids Res.* **32**, W217–W221
- 24 Schug, J. and Overton, G. C. (1997) TESS: Transcription Element Search Software on the WWW, <http://www.cbil.upenn.edu/tess>. Computational Biology and Informatics Laboratory, School of Medicine, University of Pennsylvania
- 25 Higashimoto, K., Urano, T., Sugiura, K., Yatsuki, H., Joh, K., Zhao, W., Iwakawa, M., Ohashi, H., Oshimura, M., Niikawa, N. et al. (2003) Loss of CpG methylation is strongly correlated with loss of histone H3 lysine 9 methylation at DMR-LIT1 in patients with Beckwith–Wiedemann syndrome. *Am. J. Hum. Genet.* **73**, 948–956
- 26 Trinklein, N. D., Aldred, S. F., Hartman, S. J., Schroeder, D. I., Otililar, R. P. and Myers, R. M. (2004) An abundance of bidirectional promoters in the human genome. *Genome Res.* **14**, 62–66
- 27 Mantovani, R. (1998) A survey of 178 NF-Y binding CCAAT boxes. *Nucleic Acids Res.* **26**, 1135–1143
- 28 Chandler, S. P., Kansagra, P. and Hirst, M. C. (2003) Fragile X (CGG)n repeats induce a transcriptional repression *in cis* upon a linked promoter: evidence for a chromatin mediated effect. *BMC Mol. Biol.* **4**, 3
- 29 Chen, L. S., Tassone, F., Sahota, P. and Hagerman, P. J. (2003) The (CGG)n repeat element within the 5' untranslated region of the FMR1 message provides both positive and negative *cis* effects on *in vivo* translation of a downstream reporter. *Hum. Mol. Genet.* **12**, 3067–3074
- 30 Sullivan, P. M., Wang, Y. M. and Joseph, D. R. (1993) Identification of an alternate promoter in the rat androgen-binding protein/sex hormone-binding globulin gene that regulates synthesis of a messenger RNA encoding a protein with altered function. *Mol. Endocrinol.* **7**, 702–715
- 31 Wang, Y. M., Bayliss, D. A., Millhorn, D. E., Petrusz, P. and Joseph, D. R. (1990) The androgen-binding protein gene is expressed in male and female rat brain. *Endocrinology* **127**, 3124–3130
- 32 Hammond, G. L., Underhill, D. A., Rykse, H. M. and Smith, C. L. (1989) The human sex hormone-binding globulin gene contains exons for androgen-binding protein and two other testicular messenger RNAs. *Mol. Endocrinol.* **3**, 1869–1876
- 33 Gershagen, S., Lundwall, A. and Fernlund, P. (1989) Characterization of the human sex hormone binding globulin (SHBG) gene and demonstration of two transcripts in both liver and testis. *Nucleic Acids Res.* **17**, 9245–9258
- 34 Gopalakrishnan, L. and Scarpulla, R. C. (1995) Structure, expression, and chromosomal assignment of the human gene encoding nuclear respiratory factor 1. *J. Biol. Chem.* **270**, 18019–18025
- 35 Efio, B. J., Chiorini, J. A. and Safer, B. (1994) A key transcription factor for eukaryotic initiation factor-2 alpha is strongly homologous to developmental transcription factors and may link metabolic genes to cellular growth and development. *J. Biol. Chem.* **269**, 18921–18930
- 36 Li, B. S., Kramer, P. R., Zhao, W., Ma, W., Stenger, D. A. and Zhang, L. (2000) Molecular cloning, expression, and characterization of rat homolog of human AP-2alpha that stimulates neuropeptide Y transcription activity in response to nerve growth factor. *Mol. Endocrinol.* **14**, 837–847

Received 17 May 2006/17 July 2006; accepted 4 August 2006

Published as BJ Immediate Publication 4 August 2006, doi:10.1042/BJ20060734