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

Daman KUMARI*, Andrei GABRIELIAN†, David WHEELER† and Karen USDIN*¹

*National Institute of Diabetes, Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0830, U.S.A., and †National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20892-0830, U.S.A.

Expansion of a CCG·CCG-repeat tract in the 5'-untranslated region of the FMR1 (Fragile X mental retardation 1) gene causes its aberrant transcription. This produces symptoms ranging from premature ovarian failure and Fragile X associated tremor and ataxia syndrome to FMR syndrome, depending on the size of the expansion. The promoter from normal alleles shows four protein-binding regions *in vivo*. We had previously shown that in mouse brain extracts two of these sites are bound by USF1/USF2 (upstream stimulatory factors 1 and 2) heterodimers and NRF-1 (nuclear respiratory factor-1). We also showed that these sites are involved in the positive regulation of FMR1 transcription in neuronally derived cells. In the present study, we show that Sp1 (specificity protein 1) and Sp3 are also strong positive regulators of FMR1 promoter activity. We also show that, like Sp1 and E-box-binding proteins such as USF1 and USF2, NRF-1 causes DNA

bending, in this case producing a bend of 57° towards the major groove. The combined effect of the four protein-induced bends on promoter geometry is the formation of a highly compact arch-like structure in which the 5' end of the promoter is brought in close proximity to the 3' end. We had previously shown that while point mutations in the GC-boxes decrease promoter activity, deletion of either one of them leads to an increase in promoter activity. We can reconcile these observations with the positive effect of Sp1 and Sp3 if protein-induced bending acts, at least in part, to bring together distally spaced factors important for transcription initiation.

Key words: DNA bending, Fragile X mental retardation 1 (FMR1), promoter regulation, specificity protein 1 (Sp1), upstream stimulatory factor 1 (USF1).

INTRODUCTION

CGG·CCG-repeat-expansion in the 5'-UTR (5'-untranslated region) of the FMR1 (Fragile X mental retardation 1) gene causes an increase in FMR1 transcription from alleles containing 60–200 repeats (premutation alleles) and a decrease in individuals with > 200 repeats (full mutation alleles). The basis of the increase seen in premutation carriers is not known. However, transcripts from these alleles are associated with premature ovarian failure [1–4] and a neurodegenerative disorder known as Fragile X associated tremor and ataxia syndrome [5–8]. A decrease in transcription seen in carriers of full mutation alleles is associated with aberrant promoter heterochromatinization and results in FMR [9,10]. Since the coding sequence of the gene remains intact, it may one day be possible to ameliorate the symptoms of these disorders by appropriately regulating transcription of the expanded alleles. However, many aspects of the regulation of the normal promoter are still unclear. The G + C-rich human FMR1 promoter, like the FMR1 promoters of other mammals, lacks a canonical TATA-box [11]. It does contain a number of evolutionarily conserved initiator-like elements (Figure 1). In addition there are also four evolutionarily conserved transcription factor binding sites: an E-box, two GC-boxes and an α -Pal/NRF-1 (hereafter referred to as NRF-1; where NRF-1 stands for nuclear respiratory factor-1) binding site (Figure 1). Footprinting data from normal human cells show binding to these four sites *in vivo* [12,13].

We previously identified three proteins from mouse brain extracts that bind to the full-length minimal promoter *in vitro* [11].

The 5'-end of the promoter is bound by NRF-1, whereas the E-box at the 3'-end is bound by homodimers and heterodimers of the transcription factors USF1 (upstream stimulatory factor 1) and USF2. Whereas a variety of related basic helix–loop–helix leucine zipper proteins are known to bind to E-boxes, no evidence was seen of binding by other proteins in extracts from a variety of different human cell lines and mouse tissues. All of the protein–DNA complexes seen in electrophoretic mobility-shift assays could be completely supershifted by antibodies to NRF-1, USF1 and USF2. Mutations in both the NRF-1 site and the E-box decreased promoter activity in transient transfection assays in neuronal cells consistent with factors binding to these sites being positive regulators of the FMR1 promoter. Recent co-transfection experiments in *Drosophila* cells confirm a role for NRF-1 in FMR1 activation [14]. However, no activation by USF1 or USF2 was seen in insect cells [14]. In fact, these proteins inhibited the transactivation of the promoter by other transcription factors, although the mechanism by which this occurs is unclear and does not seem to be mediated through the E-box [14]. While it is possible that other E-box binding proteins are involved, failure to see USF1/USF2 transactivation in insect cells does not necessarily preclude a role for these proteins in humans. It is known that USF1 and USF2 are relatively weak transactivators in transfection assays of some promoters [15–17], and USF activation of certain promoters can also depend on binding of cell-specific factors [18,19]. Chromatin immunoprecipitation experiments show that USF1 and USF2 do bind to the FMR1 promoter *in situ* [14], and *in vivo* footprinting data show no evidence of USF1/USF2 binding

Abbreviations used: DBD, DNA-binding domain; FMR1, Fragile X mental retardation 1; NRF-1, nuclear respiratory factor-1; Sp1, specificity protein 1; USF1, upstream stimulatory factor 1; UTR, untranslated region.

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

The four protein-bound sites in normal human lymphoblasts are indicated by the dark grey boxes. The three initiator-like sequences are shown by the light grey boxes. The three transcription start sites seen on normal and premutation alleles [41] are indicated by the arrows. Site I is used preferentially when the CGG · CCG-repeat in the 5'-UTR is within the normal range. As the repeat number increases, the usage of sites II and III increases [41]. The five mammals are *Homo sapiens*, *Pan troglodytes*, *Macaca arctoides*, *Canis familiaris* and *Mus domesticus*.

elsewhere on the promoter [13]. These data lend support to a model in which USF1/USF2 binding to the E-box contributes to FMR1 transcription activation in humans.

Whereas a small amount of protein binding to the GC-boxes could be detected with a promoter subfragment in our electrophoretic mobility-shift assay experiments, the significance of this binding was unclear, since point mutations and deletions in this region had seemingly paradoxical effects, and GC-box binding proteins can have both positive and negative effects on promoter activity [20–22]. To understand better the operation of this unusual promoter, we have extended our previous studies to examine the role of various GC-box binding factors on FMR1 promoter activity as well as the combined effects of transcription-factor binding on promoter architecture.

EXPERIMENTAL

DNA constructs

Mammalian two-hybrid constructs were made by PCR amplification of the coding region of USF1, USF2 and NRF-1. The coding sequence of NRF-1 was amplified from pGEM72f-Nrf-1 (a gift from Professor R. Scarpulla, Northwestern University Medical School, Chicago, IL, U.S.A. [23]). The coding sequence for human USF1 and mouse USF2 was amplified from plasmids pUSF1 and pUSF2 (Dr T. Kevin Howcroft; NIH, Bethesda, MD, U.S.A.). The PCR products were cloned into pM and pVP16 (BD Biosciences ClonTech, Palo Alto, CA, U.S.A.), so that in-frame fusions with the Gal4 DNA-binding domain and activation domain of VP16 were generated.

The E-box core sequence (underlined) and its flanking sequence in the FMR1 promoter (GTTGATCACGTGACGTGGT) and the NRF-1 core sequence (underlined) together with its flanking sequence in the FMR1 promoter (CAGCGCGCATGCGCGCGCT) were cloned into the *XbaI*–*SalI* sites of the vector pBend4 (provided by Dr S. Adhya; NIH, [24]). Probes 1 and 2 were generated by digestion with *MluI* and *EcoRV* respectively and labelled by using the Klenow subunit of DNA polymerase and [α - 32 P]dCTP to label the 3'- or 5'-end-labelled using T4 polynucleotide kinase and [γ - 32 P]GTP.

Phasing probes were constructed by using double-stranded oligonucleotide fragments containing three adjacent, in-phase (dA)₅ tracts and an NRF-1-binding site spaced 21, 26, 28 and

30 bp from the central deoxyadenosine in the second A-tract to generate NRF21, NRF26, NRF28 and NRF30 respectively. These oligonucleotides were cloned into the *Xba*I–*Sal*I sites of pBend4.

The sequences of the oligonucleotides are the following. NRF21: CTAGAAAGCGCGCATGCGCGGCAAAAACGGGCAAAAACGGGCAAAAACGGGTCGA. NRF26: CTAGAAGCGCGCATGCGCGTCGACGCAAAAACGGGCAAAAACGGGCAAAAACGGGTCGA. NRF28: CTAGAAGCGCGCATGCGCGTCGACACGCAAAAACGGGCAAAAACGGGCAAAAACGGGTCGA. NRF30: CTAGAAGCGCGCATGCGCGTCGACGTGACGCAAAAACGGGCAAAAACGGGCAAAAACGGGTCGA.

Each of these probes thus differed from the next one in the series by a fraction of a helical turn. Phasing probes were made from these vectors by digestion with *EcoRV* and end-labelling with T4 polynucleotide kinase and [γ - 32 P]GTP.

To replace the E-box with an intrinsically bent DNA, *Bgl*II and *Xba*I sites were introduced in the E-box by site-directed mutagenesis and intrinsically bent DNA was cloned into these sites in both orientations. The sequence of bent DNA was 5'-d(GAAAAAGCGCAAAAAGCGCAAAAAGCGCAAAAAGCGC)-3' and it results in a 72° bend, which is similar to the bend induced by USF binding [25]. To replace the NRF-1-binding site with the intrinsically bent DNA, the oligonucleotide 5'-d(GCAAAAAGCGCAAAAAGCGCAAAAAGC)-3' was cloned in the *Bss*HII site of plasmid pΔ - 131/- 123 [11]. This produces an intrinsic bend of 54°, which is similar to the 57° bend induced by NRF-1 binding (the present study).

Transient transfections in *Drosophila* S2 cells

Role of the Sp (specificity protein) family of transcription factors in the regulation of the FMR1 promoter was assayed by transient transfection in *Drosophila* S2 cells. S2 cells were used because they lack endogenous Sp factors. Cells (4×10^5) were plated in 400 μ l of serum-free Schneider medium (Invitrogen, Carlsbad, CA, U.S.A.) in each well of 24-well plates and were grown at room temperature (20°C) overnight. Cells were transfected using Cellfectin (Invitrogen) according to the manufacturer's instructions. The pGL-FMR reporter construct was used alongside derivatives of this plasmid, which contain mutations in one of the two GC-boxes [11]. Plasmids pPacSp4, pPacSp3 and pPacUSp3 that express Sp4, and the short and long isoform of Sp3 respectively, were a gift from Professor G. Suske (Institut für Molekularbiologie und Tumorforschung, Philipps-Universität Marburg, Germany). Plasmids pPac-Sp1 and pPac-Sp1DBD were a gift from Professor T. Johansen (University of Tromsø, Norway). Plasmid pPacControl was made by deletion of the Sp1 sequence from pPac-Sp1 by *Xho*I digestion and ligation.

Intrinsic and protein-induced DNA bending

The bending probes were subjected to electrophoresis on non-denaturing polyacrylamide gels (60:1, acrylamide/bisacrylamide) to check for intrinsic DNA bending. To assess protein-induced bending, protein binding was performed using mouse brain nuclear extracts and end-labelled promoter fragments as described previously [11] except that 4 mM dithiothreitol instead of 2 mM was used in the binding reactions. The DNA-protein complexes were subjected to electrophoresis on an 8 % gel (60:1, acrylamide/bisacrylamide).

Mammalian two-hybrid analysis

Protein–protein interaction between the USF proteins and NRF-1 was tested using a mammalian two-hybrid system (BD

Figure 2 The effect of Sp1 and Sp3 on FMR1 promoter activity in *Drosophila* S2 cells

(A) S2 cells were co-transfected with pGL-FMR1 [11] plasmid and either pPac-control or the specified pPac plasmids expressing the Sp factors as described in the Experimental section. The luciferase activity was assayed on extracts of these cells and normalized to the protein concentration of the extracts. (B) S2 cells were co-transfected with pPacSp1 or pPacUSP3 and pGL-FMR1, the wild-type promoter, or pGL-FMR1 derivatives containing mutations in one or other of the two GC-boxes.

Biosciences ClonTech), with different combinations of derivatives of the pM vector that contains the yeast Gal4 DNA-binding domain and VP16 derivatives containing herpes simplex virus 16 activation domain. The plasmid pG5Luc, which contains the luciferase gene driven by a promoter containing a number of Gal4-binding sites, was used as the reporter. The luciferase activity was assayed using the dual luciferase system as described earlier [11].

RESULTS

Sp1 and the long isoform of Sp3 activate transcription from the FMR1 promoter in *Drosophila* S2 cells

To assess the role of Sp family members on the activity of the FMR1 promoter, we performed a series of co-transfection experiments in *Drosophila* S2 cells. These cells were used since they do not contain endogenous Sp factors. When plasmids expressing the full-length Sp1 protein were co-transfected with an FMR1-promoter reporter construct (pGL-FMR), the activity of the FMR1 promoter was increased approx. 34-fold (Figure 2A). Our previous observation of weak Sp factor binding that was seen only on a promoter subfragment may reflect either suboptimal binding conditions used *in vitro* or the loss during preparation of the nuclear extracts, of post-transcriptional modifications or trans-acting factors that enhance Sp1/Sp3 binding. When a plasmid expressing the small isoform of human Sp3 protein (pPacSp3) was used, no difference in promoter activity was seen. However, co-transfection with the plasmid expressing the long isoform of Sp3 (pPacUSP3) gave an increase comparable with that seen with Sp1, suggesting that the long isoform of Sp3 also plays a role in the regulation of this promoter. A plasmid expressing Sp4 had little, if any, effect on promoter activity. We then compared the ability of pPacSp1 and pPacUSP3 to activate transcription of the wild-type FMR1 promoter or constructs containing mutations in each of the GC-boxes. Mutations in either of the GC-boxes significantly reduced the effect of both Sp1 and Sp3 (Figure 2B). These results demonstrate that Sp1 and Sp3 have a positive effect on the FMR1 promoter through a direct interaction with the GC-boxes.

Intrinsic and protein-induced bending of the FMR1 promoter

These data seem at odds with our previous data showing that deletions that include this region cause an increase in promoter ac-

Figure 3 Electrophoretic mobility-based analysis of intrinsic and protein-induced DNA bending

(A) Intrinsic DNA bending: (i) 100 bp overlapping fragments from the FMR1 promoter were amplified by PCR and run alongside *MspI* digested pBR322 DNA on a 6% native polyacrylamide gel; and (ii) sequence of fragment 5. The three transcription start sites in this fragment are indicated by the arrows and roman numerals. (B) Protein-induced DNA bending: (i) restriction map of the probes used in the experiments shown in (ii) and (iii), where the grey oblong indicates the location of the E-box or NRF-1-binding sites in the DNA probe; (ii) mobility of the DNA probes on a non-denaturing acrylamide gel; and (iii) mobility of DNA-protein complexes formed with mouse brain nuclear extracts on *MluI* or *EcoRV* digested DNA probes.

tivity [11]. Since it is known that Sp1 and E-box-binding proteins like USF1 and USF2 bend DNA [25,26], one way these observations could be reconciled if we assume that the FMR1 promoter architecture is important for optimal promoter activity, and both the deletions and the transcription-factor-induced bending act, at least in part, to bring distal regions of the promoter closer together.

To understand better the overall conformation of the FMR1 promoter, we used PCR to amplify overlapping fragments of 100 bp and ran these fragments alongside a standard DNA marker on 6% native polyacrylamide gels to look for any intrinsic migration anomalies (Figure 3A). Fragment 5 [Figure 3A(ii)], which contains all three transcription start sites, migrated slightly slower than the other fragments in the absence of nuclear extracts. Analysis of these fragments using the program BendoBase (available on request from D. W.) and a standard dinucleotide wedge model for sequence-directed DNA curvature [27,28] also suggests a small amount of intrinsic DNA curvature in this fragment.

DNA bending induced by the binding of NRF-1 was assessed by examining the mobility of DNA probes that contained the binding site either in the middle or at end of the fragment. Although this sort of analysis has been questioned in recent years, particularly for proteins containing a leucine zipper [29], recent analysis by Hardwidge et al. [30] support the general reliability of such assays for this group of proteins. The mobility of the probes in the absence of protein was indistinguishable [Figure 3B(ii)]. A difference in the mobility of DNA-protein complexes was seen, and this mobility varied depending on the position of the binding site on the probe fragment [Figure 3B(iii)]. This is consistent with NRF-1-induced DNA bending. Thus NRF-1 joins that large group of transcription factors known to bend DNA. The bending angle (α) for NRF-1 was calculated using the empirical relationship $\mu_M/\mu_E = \cos(\alpha/2)$, where μ_M is the mobility of the complex with the protein bound in the middle of the fragment and μ_E the mobility of the complex with the protein bound near the end of a DNA fragment [31]. The bending angle for the NRF-1-induced

Figure 4 Phasing analysis of NRF-1-induced DNA bend

(A) Diagrammatic representation of the principle of the phasing analysis. (B) Mobility of the NRF-1 phasing probes on a native polyacrylamide gel. M, markers. (C) Mobility of the NRF-1 phasing probes complexed with mouse brain nuclear extracts on a native polyacrylamide gel. The numbers in the probe names indicate the space between the NRF-1-binding site and the central deoxyadenosine in the middle adenosine tract.

bend was 57°. The bend angle for the DNA bend induced by binding to the E-box was also calculated and shown to be 72°, similar to what has been reported earlier for a variety of E-box-binding proteins, including USF1/USF2 [25].

To determine whether NRF-1 bending was oriented towards the major or the minor groove, phasing analysis was performed [32] with mouse brain nuclear extracts and DNA probes containing both an NRF-1 site and an intrinsic DNA bend. The intrinsic bend was produced by the insertion of three adjacent tracts of (dA-dT)₅ base-pairs, which generate minor-groove-directed bends of 18–22° per tract [25]. The probes were constructed so that the central nucleotide of the intrinsic bend was spaced 21, 26, 28 or 30 bp from the centre of the NRF-1-binding site. These probes thus have the two bending sites spaced between two and three complete turns of the helix apart. The DNA–protein complex with the fastest mobility is the one in which the two bends are closest to being out-of-phase. This creates a complex with the largest end-to-end distance and the smallest overall bend angle. The fastest mobility of the DNA–protein complex was observed with the fragment NRF21, which has 21 bases between the centre of the intrinsic bend and the centre of the NRF-1-binding site (Figure 4) or approx. two complete superhelical turns. Since the intrinsic bends are towards the minor groove, our data show that the net orientation of DNA bending induced by NRF-1 is towards the major groove.

Figure 5 Rasmol rendering of the modelled FMR1 promoter structure

The sequence of the modelled region is shown at the upper portion of the Figure. The transcription factor binding sites are shown within the boxed regions, with the bend angle and direction of the protein-induced bend shown above. The underlined bases indicate the bases on which the protein-induced bends were centred. The lower portion of the Figure shows a model of the promoter generated using Berkeley Enhanced Rasmol (<http://mc2.cchem.berkeley.edu/Rasmol/>). The FMR1 model is superimposed, for scale, on the structure of DNA from the nucleosome core particle of *Xenopus laevis*. The boxed regions correspond to the boxed regions in the upper panel. Numbers indicate the first and last base of the modelled sequence. Core particle DNA co-ordinates, representing 146 bp of DNA and making approx. 1.75 turns around the histone octamer core, were taken from PDB record 1AOI [33].

BendoBase was used to generate models of the 160 bp core FMR1 promoter region by superimposing protein-induced bend angles, centred on the underlined bases within the transcription-factor binding sites shown in Figure 5, on a structure calculated using a standard dinucleotide wedge model [27,28]. Protein-induced bends were simulated by adjusting the roll parameter used in the standard dinucleotide wedge model for bases at the bend centres. Using protein-induced bend geometries of 57° into the major groove for NRF-1, as determined in this study, 72° into the minor groove for E-box-binding proteins like USF [25] and 60° into the major groove for Sp1 [26], we predict the three-dimensional structure shown in dark grey in Figure 5. For comparison, the DNA trajectory of the nucleosome, taken from the X-ray crystallographic structure of the nucleosome core particle [33], is shown in light grey. At the point of closest approach, the 5' and 3' ends of the promoter are separated by a distance roughly equivalent to the diameter of nucleosome-bound DNA. The calculated structure is compact as a result of the combined effect of the four induced bends. An examination of the effects of all possible combinations of the four bends in the structures demonstrates that the removal of any one bend eliminates the planar curvature required for the compact structure and the close approach of the 5' and 3' ends.

DNA bending is not sufficient for transcription activation

The overall architecture of the promoter is thus consistent with the idea that deletions may mimic the effect of protein-induced bending in the FMR1 promoter by bringing together widely separated elements important for gene expression. To examine this issue in more detail, we first tested whether there was a direct effect

Figure 6 Interaction of USF proteins and NRF-1 *in vivo*

The luciferase activity resulting from the co-transfection of the two indicated test plasmids is expressed relative to the respective PM-protein and VP16 controls after normalization to the activity of a co-transfected pRLNull plasmid. Results are an average for three separate experiments, each performed in triplicate.

of DNA bending on the activation of the FMR1 promoter by replacing either the E-box sequence or the NRF-1-binding site with intrinsically bent DNA with the same bending angle. A precipitous fall in the activity of the promoter was seen when either the E-box site or the NRF-1 site was replaced with an intrinsically bent DNA fragment with the same bend angle. This was true when the intrinsic bend was into either the major or minor groove (results not shown). In addition, an expression construct containing just the DBD (DNA-binding domain) of Sp1 was also unable to activate transcription in *Drosophila* cells (Figure 2).

Analysis of transcription factor interactions

These data suggest that DNA bending alone is not sufficient for the activation of transcription from the FMR1 promoter. Because of the large size of the intrinsically bent fragment, which was used to replace the factor-binding sites, these findings do not preclude a role for protein-induced bending in facilitating interactions between distally spaced factors. Neither does the failure of the Sp1 DBD to activate transcription rule out a similar role for Sp1. While this work was in progress, a report was published showing that Sp1 and USF1/USF2 interact [15]. Other interactions involving the factors that bind the FMR1 promoter may also be important. For example, NRF-1 has already been shown to interact with Max, another E-box-binding protein that is closely related to USF1 and USF2 [34]. We therefore tested whether or not NRF-1 was capable of interacting with USF1/USF2 using a mammalian two-hybrid system. As a positive control, the interaction between USF1 and USF2 was tested using the same system, since while *in vivo* data are lacking, their heterodimerization *in vitro* and *ex vivo* is well established [35]. Using this assay, we were able to demonstrate directly their interaction *in vivo* (Figure 6). We also showed that, as is seen *ex vivo* [36], heterodimers interact *in vivo* much more efficiently than either USF1 or USF2 homodimers. Since this preferential association is not seen when these proteins are produced *in vitro* [36], it may reflect a specific post-translational modification that facilitates this interaction. Increased levels of the pG5Luc reporter gene were seen when a plasmid with the NRF-1 open reading frame fused to the VP16 activation domain was used in combination with a plasmid with the NRF-1 open reading frame fused to the yeast Gal4 DBD. This is consistent with *in vitro* observations of NRF-1 dimerization [34,37].

When the interaction of NRF-1 and USF1 or USF2, either separately or in combination, was examined, luciferase activities were not above the background value when NRF-1 was expressed as a fusion protein with the DBD, and only slightly above background when it was attached to the activation domain (Figure 6). There is thus probably little, if any, direct physical interaction between the NRF-1 and USF proteins in solution *in vivo*. This leaves open the possibility that bending facilitates the interaction of transcription factors with components of the general transcription machinery or with the as yet unidentified proteins that bind the CGG · CCG tract.

DISCUSSION

The data shown in the present study support a direct positive role for the GC-box binding proteins Sp1 and Sp3 in FMR1 promoter regulation. Our demonstration that NRF-1 binding induces a DNA bend of 57° that is directed towards the major groove suggests that, in combination with Sp1 and USF1/USF2, which were already known to cause DNA bending, the overall architecture of the FMR1 promoter might be quite compact. Since all the E-box-binding proteins studied to date bend DNA in a very similar fashion [25], a very similar compact architecture may well be seen even if other E-box-binding proteins are involved. Computer modelling predicts the combination of four protein-induced bends would produce a planar curvature capable of bringing the 5' and 3' ends of the promoter region into close proximity. Such a structure could provide a convenient DNA scaffold for the formation of a sizable protein assembly. This structure is consistent with the fact that Sp1 positively regulates the FMR1 promoter, whereas deletions that include this site lead to an increase in promoter activity: deletions may mimic the protein-induced bends in bringing factors at the different ends of the promoter into close proximity, thereby compensating for the loss of other transcription factor binding sites.

Sp1 is known to interact with both USF1/USF2 [15] and NRF-1 [14]. In addition, we have previously shown that constructs containing the FMR1 promoter and 5' UTR with a deletion of a region that includes the GC-box closest to the start of transcription and the E-box site have a promoter activity that is 5-fold higher than a construct with a point mutation in the E-box [11]. This suggests that interaction between factors upstream of this GC-box and those downstream of the E-box might be quite significant. The downstream factors may be components of the general transcription machinery or as yet unidentified factors that bind the CGG · CCG-repeat in the 5' UTR. The compact conformation of the FMR1 promoter may facilitate these interactions.

A role for Sp1 in the positive regulation of the FMR1 promoter is interesting, given the reports that the FMR1 promoter is responsive to cAMP [39], and our previous demonstration of a lack of CREB (cAMP-response-element-binding protein) binding, since Sp1 stimulates promoter activity by cAMP, independent of binding by CREB [40]. It has been shown that as the CGG · CCG-repeat length increases in the premutation range, upstream transcription start sites become more frequently used, so that with 82 or more repeats start site III shown in Figure 1 is used approx. 60% of the time [41]. This start site coincides with the E-box. How this affects the contribution of factors that bind the E-box to activity of the promoter is not yet known, but may point to an increased reliance on NRF-1 and the Sp factors in premutation carriers. In addition, while the binding of NRF-1 and USF1/USF2 to DNA is affected, at least to some degree, by CpG methylation [11], Sp1 binding is not [42,43]. Activation of the FMR1 promoter by Sp1 family members may thus play a significant role in the residual transcription seen in many individuals with Fragile X

syndrome who have fully methylated FMR1 alleles [44,45]. Sp1 and Sp3 can also have negative effects on transcription. This effect can be mediated through their interaction with proteins such as histone deacetylases [46], MECP2 [47] and MCAF (which also binds the repressor MBD1 when the promoter is methylated) [48]. Differences in the amount of Sp1/Sp3 or in the extent of the interaction of Sp1 and Sp3 with these proteins may also contribute to the variability in FMR1 expression that is seen in full mutation carriers.

REFERENCES

- Vianna-Morgante, A. M., Costa, S. S., Pares, A. S. and Verreschi, I. T. (1996) FRAXA premutation associated with premature ovarian failure. *Am. J. Med. Genet.* **64**, 373–375
- Murray, A., Webb, J., Grimley, S., Conway, G. and Jacobs, P. (1998) Studies of FRAXA and FRAXE in women with premature ovarian failure. *J. Med. Genet.* **35**, 637–640
- Macpherson, J., Murray, A., Webb, J. and Jacobs, P. (1999) Fragile X syndrome: of POF and premutations. *J. Med. Genet.* **36**, 171–172
- Allingham-Hawkins, D. J., Babul-Hirji, R., Chitayat, D., Holden, J. J., Yang, K. T., Lee, C., Hudson, R., Gorwill, H., Nolin, S. L., Glicksman, A. et al. (1999) Fragile X premutation is a significant risk factor for premature ovarian failure: the International Collaborative POF in Fragile X study—preliminary data. *Am. J. Med. Genet.* **83**, 322–325
- Jacquemont, S., Hagerman, R. J., Leehey, M., Grigsby, J., Zhang, L., Brunberg, J. A., Greco, C., Des Portes, V., Jardini, T., Levine, R. et al. (2003) Fragile X premutation tremor/ataxia syndrome: molecular, clinical, and neuroimaging correlates. *Am. J. Hum. Genet.* **72**, 869–878
- Leehey, M. A., Munhoz, R. P., Lang, A. E., Brunberg, J. A., Grigsby, J., Greco, C., Jacquemont, S., Tassone, F., Lozano, A. M., Hagerman, P. J. et al. (2003) The fragile X premutation presenting as essential tremor. *Arch. Neurol.* **60**, 117–121
- Brunberg, J. A., Jacquemont, S., Hagerman, R. J., Berry-Kravis, E. M., Grigsby, J., Leehey, M. A., Tassone, F., Brown, W. T., Greco, C. M. and Hagerman, P. J. (2002) Fragile X premutation carriers: characteristic MR imaging findings of adult male patients with progressive cerebellar and cognitive dysfunction. *Am. J. Neuroradiol.* **23**, 1757–1766
- Greco, C. M., Hagerman, R. J., Tassone, F., Chudley, A. E., Del Bigio, M. R., Jacquemont, S., Leehey, M. and Hagerman, P. J. (2002) Neuronal intranuclear inclusions in a new cerebellar tremor/ataxia syndrome among fragile X carriers. *Brain* **125**, 1760–1771
- Fu, Y. H., Kuhl, D. P., Pizzuti, A., Pieretti, M., Sutcliffe, J. S., Richards, S., Verkerk, A. J., Holden, J. J., Fenwick, Jr, R. G., Warren, S. T. et al. (1991) Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell (Cambridge, Mass.)* **67**, 1047–1058
- 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 (Cambridge, Mass.)* **65**, 905–914
- 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
- Drouin, R., Angers, M., Dallaire, N., Rose, T. M., Khandjian, W. and Rousseau, F. (1997) Structural and functional characterization of the human FMR1 promoter reveals similarities with the hnRNP-A2 promoter region. *Hum. Mol. Genet.* **6**, 2051–2060
- Schwemmler, S., de Graaff, E., Deissler, H., Glaser, D., Wohrle, D., Kennerknecht, I., Just, W., Oostra, B. A., Dorfner, W., Vogel, W. et al. (1997) Characterization of FMR1 promoter elements by *in vivo*-footprinting analysis. *Am. J. Hum. Genet.* **60**, 1354–1362
- 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
- Ge, Y., Jensen, T. L., Matherly, L. H. and Taub, J. W. (2003) Physical and functional interactions between USF and SP1 proteins regulate human deoxycytidine kinase promoter activity. *J. Biol. Chem.* **278**, 49901–49910
- Liu, M., Whetstone, J. R., Payton, S. G., Ge, Y., Flatley, R. M. and Matherly, L. H. (2004) Roles of USF, Ikaros, and Sp proteins in the transcriptional regulation of the human reduced folate carrier B promoter. *Biochem. J.* **383**, 249–257
- Pagliuca, A., Cannada-Bartoli, P. and Lania, L. (1998) A role for Sp and helix-loop-helix transcription factors in the regulation of the human Id4 gene promoter activity. *J. Biol. Chem.* **273**, 7668–7674
- Oyang, Y., Luo, X., Lu, T., Ismail, P. M., Krylov, D., Vinson, C. and Sawadogo, M. (1999) Cell-type-dependent activity of the ubiquitous transcription factor USF in cellular proliferation and transcriptional activation. *Mol. Cell. Biol.* **19**, 1508–1517
- Ismail, P. M., Lu, T. and Sawadogo, M. (1999) Loss of USF transcriptional activity in breast cancer cell lines. *Oncogene* **18**, 5582–5591
- Zaid, A., Hodny, Z., Li, R. and Nelson, B. D. (2001) Sp1 acts as a repressor of the human adenine nucleotide translocase-2 (ANT2) promoter. *Eur. J. Biochem.* **268**, 5497–5503
- Zhao, S., Venkatasubbarao, K., Li, S. and Freeman, J. W. (2003) Requirement of a specific Sp1 site for histone deacetylase-mediated repression of transforming growth factor beta Type II receptor expression in human pancreatic cancer cells. *Cancer Res.* **63**, 2624–2630
- Majello, B., De Luca, P. and Lania, L. (1997) Sp3 is a bifunctional transcription regulator with modular independent activation and repression domains. *J. Biol. Chem.* **272**, 4021–4026
- Herzig, R. P., Andersson, U. and Scarpulla, R. C. (2000) Dynein light chain interacts with NRF-1 and ERG, structurally and functionally related transcription factors from humans and drosophila. *J. Cell Sci.* **113**, 4263–4273
- Zwieb, C. and Adhya, S. (2001) Plasmid vectors for the analysis of protein-induced DNA bending. *Methods Mol. Biol.* **148**, 403–416
- Fisher, D. E., Parent, L. A. and Sharp, P. A. (1992) Myc/Max and other helix-loop-helix/leucine zipper proteins bend DNA toward the minor groove. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 11779–11783
- Sjottem, E., Andersen, C. and Johansen, T. (1997) Structural and functional analyses of DNA bending induced by Sp1 family transcription factors. *J. Mol. Biol.* **267**, 490–504
- Kabsch, W., Sander, C. and Trifonov, E. N. (1982) The ten helical twist angles of B-DNA. *Nucleic Acids Res.* **10**, 1097–1104
- Bolshoy, A., McNamara, P., Harrington, R. E. and Trifonov, E. N. (1991) Curved DNA without A-A: experimental estimation of all 16 DNA wedge angles. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 2312–2316
- McCormick, R. J., Badalian, T. and Fisher, D. E. (1996) The leucine zipper may induce electrophoretic mobility anomalies without DNA bending. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14434–14439
- Hardwidge, P. R., Kahn, J. D. and Maher, III, L. J. (2002) Dominant effect of protein charge rather than protein shape in apparent DNA bending by engineered bZIP domains. *Biochemistry* **41**, 8277–8288
- Thompson, J. F. and Landy, A. (1988) Empirical estimation of protein-induced DNA bending angles: applications to lambda site-specific recombination complexes. *Nucleic Acids Res.* **16**, 9687–9705
- Kerppola, T. K. and Curran, T. (1991) Fos-Jun heterodimers and Jun homodimers bend DNA in opposite orientations: implications for transcription factor cooperativity. *Cell (Cambridge, Mass.)* **66**, 317–326
- Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. and Richmond, T. J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature (London)* **389**, 251–260
- Shors, S. T., Efio, B. J., Harkin, S. J. and Safer, B. (1998) Formation of alpha-Pal/Max heterodimers synergistically activates the eIF2-alpha promoter. *J. Biol. Chem.* **273**, 34703–34709
- Sirito, M., Walker, S., Lin, Q., Kozlowski, M. T., Klein, W. H. and Sawadogo, M. (1992) Members of the USF family of helix-loop-helix proteins bind DNA as homo- as well as heterodimers. *Gene Expr.* **2**, 231–240
- Viollet, B., Lefrançois-Martinez, A. M., Henrion, A., Kahn, A., Raymondjean, M. and Martinez, A. (1996) Immunochemical characterization and transacting properties of upstream stimulatory factor isoforms. *J. Biol. Chem.* **271**, 1405–1415
- Gugneja, S. and Scarpulla, R. C. (1997) Serine phosphorylation within a concise amino-terminal domain in nuclear respiratory factor 1 enhances DNA binding. *J. Biol. Chem.* **272**, 18732–18739
- Reference deleted
- Hwu, W. L., Wang, T. R. and Lee, Y. M. (1997) FMR1 enhancer is regulated by cAMP through a cAMP-responsive element. *DNA Cell Biol.* **16**, 449–453
- Cote, F., Schussler, N., Boularand, S., Peirotes, A., Thevenot, E., Mallet, J. and Vojdani, G. (2002) Involvement of NF-Y and Sp1 in basal and cAMP-stimulated transcriptional activation of the tryptophan hydroxylase (TPH) gene in the pineal gland. *J. Neurochem.* **81**, 673–685
- Beilina, A., Tassone, F., Schwartz, P. H., Sahota, P. and Hagerman, P. J. (2004) Redistribution of transcription start sites within the FMR1 promoter region with expansion of the downstream CGG-repeat element. *Hum. Mol. Genet.* **13**, 543–549
- Harrington, M. A., Jones, P. A., Imagawa, M. and Karin, M. (1988) Cytosine methylation does not affect binding of transcription factor Sp1. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2066–2070
- Holler, M., Westin, G., Jiricny, J. and Schaffner, W. (1988) Sp1 transcription factor binds DNA and activates transcription even when the binding site is CpG methylated. *Genes Dev.* **2**, 1127–1135
- Tassone, F., Hagerman, R. J., Chamberlain, W. D. and Hagerman, P. J. (2000) Transcription of the FMR1 gene in individuals with fragile X syndrome. *Am. J. Med. Genet.* **97**, 195–203

-
- 45 Tassone, F., Hagerman, R. J., Taylor, A. K. and Hagerman, P. J. (2001) A majority of fragile X males with methylated, full mutation alleles have significant levels of FMR1 messenger RNA. *J. Med. Genet.* **38**, 453–456
- 46 Doetzlhofer, A., Rotheneder, H., Lagger, G., Koranda, M., Kurtev, V., Brosch, G., Wintersberger, E. and Seiser, C. (1999) Histone deacetylase 1 can repress transcription by binding to Sp1. *Mol. Cell. Biol.* **19**, 5504–5511
- 47 Kudo, S. (1998) Methyl-CpG-binding protein MeCP2 represses Sp1-activated transcription of the human leukosialin gene when the promoter is methylated. *Mol. Cell. Biol.* **18**, 5492–5499
- 48 Fujita, N., Watanabe, S., Ichimura, T., Ohkuma, Y., Chiba, T., Saya, H. and Nakao, M. (2003) MCAF mediates MBD1-dependent transcriptional repression. *Mol. Cell. Biol.* **23**, 2834–2843
-

Received 1 July 2004/1 October 2004; accepted 13 October 2004

Published as BJ Immediate Publication 13 October 2004, DOI 10.1042/BJ20041124