

Marie A. Hillman · Jozef Gecz

Fragile XE-associated familial mental retardation protein 2 (FMR2) acts as a potent transcription activator

Received: June 19, 2000 / Accepted: January 26, 2001

Abstract Expansion of the FRAXE CCG repeat to a full mutation is associated with methylation and transcriptional silencing of the *FMR2* gene, and as a consequence, mild-to-borderline mental retardation. FMR2 is a member of a family of four proteins, AF4, LAF4, FMR2, and AF5q31. The proteins associated with this family localize to the cell nucleus. Various regions of FMR2, and each of the other members of the protein family, were cloned and analyzed for transcription activation in yeast and mammalian cells. In both yeast and mammalian cells, FMR2 showed strong transcription activation. AF4 activation potential was several-fold lower. Interestingly, isoforms of both FMR2 and LAF4 lacking exon 3 activated transcription better than the larger isoforms containing exon 3. Compared with the other members of the family, activation by FMR2 was the strongest. Our results show that FMR2 is a potent transcription activator and that its function is conserved. Elucidation of the function of the FMR2 protein as a transcription activator may place FMR2 within the molecular signalling pathways involved in nonspecific X-linked mental retardation (MRX).

Key words Alternative splicing · Transcription activation · FMR2 · AF4 · LAF4 · AF5q31 · XLMR

Introduction

FMR2, the gene associated with the *FRAXE* fragile site, was the first gene for a nonspecific familial X-linked mental retardation (XLMR) to be identified (Gecz et al. 1996; Gu

et al. 1996). The prevalence of *FRAXE* (mental retardation; MIM no. 309548) is estimated to be 1/50,000 to 1/100,000 live births (Mulley et al. 1995; Brown 1996) and it is the most prevalent form of nonspecific XLMR so far identified. *FRAXE* MR is a mild-to-borderline mental retardation (50 < IQ < 85) which manifests as learning difficulties and speech and behavioral problems. Several *FRAXE* full mutation males have been described with IQ within the normal range (Mulley et al. 1995; Murgia et al. 1996; Gecz et al. 1997b; Lo Nigro et al. 2000; for review, see Gecz 2000).

The *FMR2* gene is transcribed as a major 8.75-kb transcript (Gecz et al. 1997a). Shorter (*Ox19*; ~1.5kb; Chakrabarti et al. 1996) and longer (13.7kb; Gecz and Mulley 1999) specific isoforms were observed, the latter expressed predominantly in fetal brain. Additionally, alternative splicing of at least 5 of 20 internal exons was detected (Gecz et al. 1997a). The largest open reading frame of 3933bp is translated into a 1311 amino acid hydrophilic protein which is localized to the cell nucleus (Gecz et al. 1997a). Database searches have shown that the FMR2 protein is a member of a family of proteins including AF4 (Morrissey et al. 1993), LAF4 (Ma and Staudt 1996), and the recently characterized chromosome 5 gene *AF5q31* (Taki et al. 1999). Paralogous members of this *AF4/LAF4/FMR2/AF5q31* gene family show several similarities: size of transcript(s) at about 1.5, 8.0, and 13kb; alternative splicing and gene structure conservation (Gecz et al. 1997a; Nilson et al. 1997); protein size of ~1300 amino acids; and nuclear localization of the proteins (Ma and Staudt 1996; Gecz et al. 1997a). *AF4* was identified because of its involvement in a t(4;11) translocation, resulting in a fusion with the *ALL1* (*MLL*) gene, causing acute lymphoblastic leukemia (*ALL*) (Prasad et al. 1995). *LAF4* (Ma and Staudt 1996), identified from Raji Burkitt's lymphoma cells, is thought to have a role in lymphoid development. Preliminary studies on *AF4* (Prasad et al. 1995) and *LAF4* (Ma and Staudt 1996) show that these proteins have the potential to activate transcription of a reporter gene in an in-vivo transcription activation assay.

The normal function of the FMR2 protein is unknown. In the present study, we analyzed the transcription activa-

M.A. Hillman · J. Gecz (✉)

Centre for Medical Genetics, Department of Cytogenetics and Molecular Genetics, Women's and Children's Hospital, 72 King William Road, North Adelaide, SA 5006, Australia
Tel. +61-8-8204-7023; Fax +61-8-8204-7342
e-mail: jozef.gecz@adelaide.edu.au

J. Gecz

Department of Paediatrics, University of Adelaide, Adelaide, SA, Australia

tion potential of FMR2 and compared it with the other members of the gene family in yeast and mammalian cells. Our results show that, in this system, FMR2, AF4, and LAF4 are potent transcription activators. Interestingly, an alternative splicing product of FMR2 and LAF4 lacking exon 3 led to elevated transcription activation of the reporter gene in both yeast and mammalian cells.

Materials and methods

Plasmids

Effector plasmids were either pGBT9 or pM (Clontech, Palo Alto, CA, USA). These two vectors contained a GAL4 DNA binding domain (BD; amino acids 1–147), in pGBT9 flanked by the yeast *ADHI* promoter and *ADHI* transcription termination signal, and in pM flanked by the simian virus 40 (SV40) early promoter and polyA transcription termination signal. The pG5CAT (Clontech) reporter vector was used in cotransfections with pM into mammalian cells. pG5CAT contained the chloramphenicol acetyltransferase (*CAT*) gene downstream of 5 consensus GAL4 upstream activating sequences (UAS_{G17-mer(x5)}) and an adenovirus E1b minimal promoter. The β -galactosidase-expressing plasmid (pSV- β) was obtained from Promega (Madison, WI, USA) and used to normalize transfection efficiencies between experiments. All constructs were examined for expression of the GAL4 chimeric proteins by transfection into yeast or HeLa cells, and subsequent assays for transcription activation were carried out using β -galactosidase and *CAT* assays.

Design and preparation of expression constructs

Expression vectors for fusions between FMR2, LAF4, AF4, AF5q31, or ALL1 and the GAL4 DNA-BD were made by cloning reverse transcription (RT)-polymerase chain reac-

tion (PCR) amplified fragments of the corresponding gene into the multiple cloning site of either the pM or pGBT9 vectors. mRNA was isolated using Trizol reagent (GIBCO BRL, Grand Island, NY, USA) according to the manufacturer's protocol. Approximately 500ng of purified mRNA was random primed for first-strand cDNA synthesis, using SuperScript II RNase H⁻ reverse transcriptase (GIBCO BRL). Reverse transcription was carried out for 1h at 42°C in a thermal cycler (Hybaid, Ashford, Middlesex, UK). Negative controls, containing no reverse transcriptase, were run at the same time. Fragments were amplified from fetal brain (Gecz et al. 1997a) cDNA by PCR, using primers designed to contain a restriction enzyme site (see Table 1). Standard conditions for PCR were 30s at 94°C, 30s at 60°C, and 90s at 72°C for 35 cycles, in a final 50- μ l volume containing 100mM Tris-HCl, pH, 8.3; 50mM KCl; 1.5mM MgCl₂; 0.2mM of each dNTP; 50pmol of each primer; and 2U of *Taq* DNA polymerase (Boehringer Mannheim, Mannheim, Germany). PCR products were analyzed on a 1% agarose gel. The tested DNA fragments encoded the following amino acids: for FMR2 (FMR2) 481–541, (FMR2₂₋₁) 1–57, 61–364, and 392–541 (FMR2₂₋₅) 1–61, 348–364, 392–416, and 422–541; for LAF4 (LAF4) 426–484, (LAF4₂₋₁) 1–413, 424–450, and 462–485, (LAF4₂₋₅) 1–19 and 292–413, and 424–484; for AF4 (AF4) 464–523; and for AF5q31 (EST W26686) primers were designed to amplify the 460–500 paralogous amino acid region (Fig. 1a,b). After amplification, the PCR products were digested with restriction enzymes and cloned into the pM or pGBT9 vectors. An amount of 5–300ng of the digested PCR product and 50ng of the vector were ligated at 4°C overnight and transformed into XL1 Blue competent cells. The following day, positive clones were detected via colony PCR. The plasmid DNA was prepared using a QIAGEN Plasmid Mini Kit (Bothell, WA, USA).

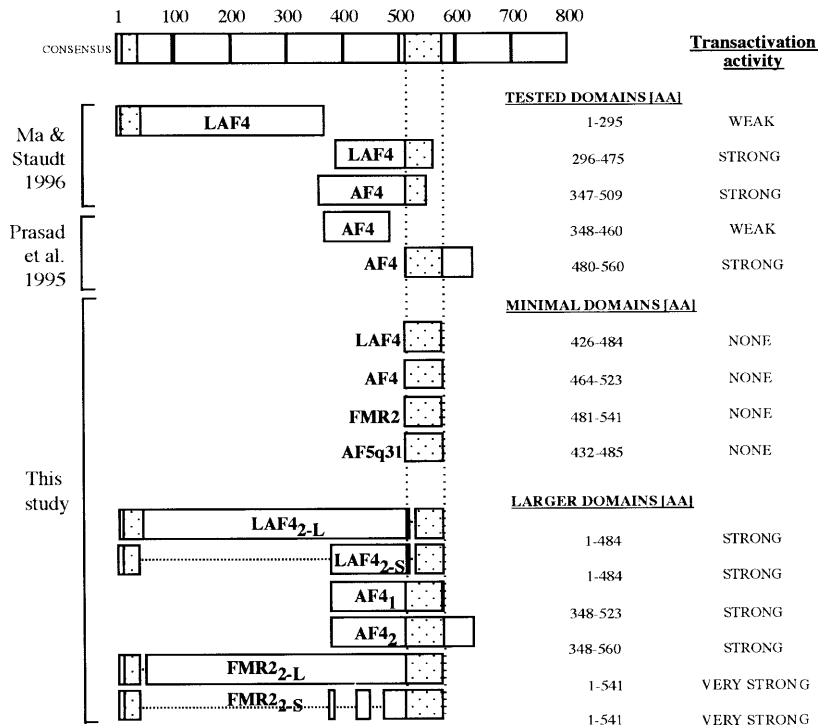
Positive constructs were sequenced using the Big Dye Terminator Cycle Sequencing Ready Reaction Kit (ABI, Foster City, CA, USA), to confirm that constructs had the correct insert and that the insert was in the correct reading frame.

Table 1. Oligonucleotide primers used for polymerase chain reaction (PCR)

Primer name	Sequence (5' to 3')	Size (bp)	RE site	AA
Primer 1	TCATCGGAAGAGAGTAG	17	*	*
LAF4 TAF	CCGGAATTCAGCGACTCAGAGAGCAGCTCC	30	<i>Eco</i> R1	426
LAF4 TAR	CCGGATCCGGATCAGAATAGGAGGCTTGTG	31	<i>Bam</i> H1	484
LAF4 TAF2	CCGGAATTCGACAGCTTCGACTTAGCCCTGCTC	33	<i>Eco</i> R1	1
LAF4 TAR3	CCGGTTCGACGGATCAGAATAGGAGGCTTGTG	31	<i>Sal</i> 1	484
AF4 TAF	CCGGAATTCCTCAGAAAGCACCAGTGACTCAGAC	33	<i>Eco</i> R1	464
AF4 TAR	CCGGGATCCGGCTGACTTTGGTCAGCCAGTTG	32	<i>Bam</i> H1	523
FMR2 TAF	CCGGAATTCAGCGAATCGGAGAGCAGCTCT	30	<i>Eco</i> R1	481
FMR2 TAR	CCGGGATCCGACTTGTTCTGGGATGTCAC	29	<i>Bam</i> H1	541
FMR2 TAF2	CCGCCCGGGAAGCCTTCCCAGTGATCCAAGC	31	<i>Sma</i> 1	348
FMR2 TAF3	CCGCCCGGGGATCTATTCGACTTTTTTCAGAGA	33	<i>Sma</i> 1	1
FMR2 TAR3	CCGGTTCGACAGACTTGTCTGGGATGTCAC	29	<i>Sal</i> 1	541
ALL1 TAF	CCGGAATTCCTGAAATCAGATTCAGACAATAAC	33	<i>Eco</i> R1	2829
ALL1 TAR	CCGGGATCCAAGACCCAATCCTTCACCAAG	30	<i>Bam</i> H1	2883

TAF, transcription activation forward primer; TAR, transcription activation reverse primer; AA, amino acid; RE, restriction enzyme

a



b

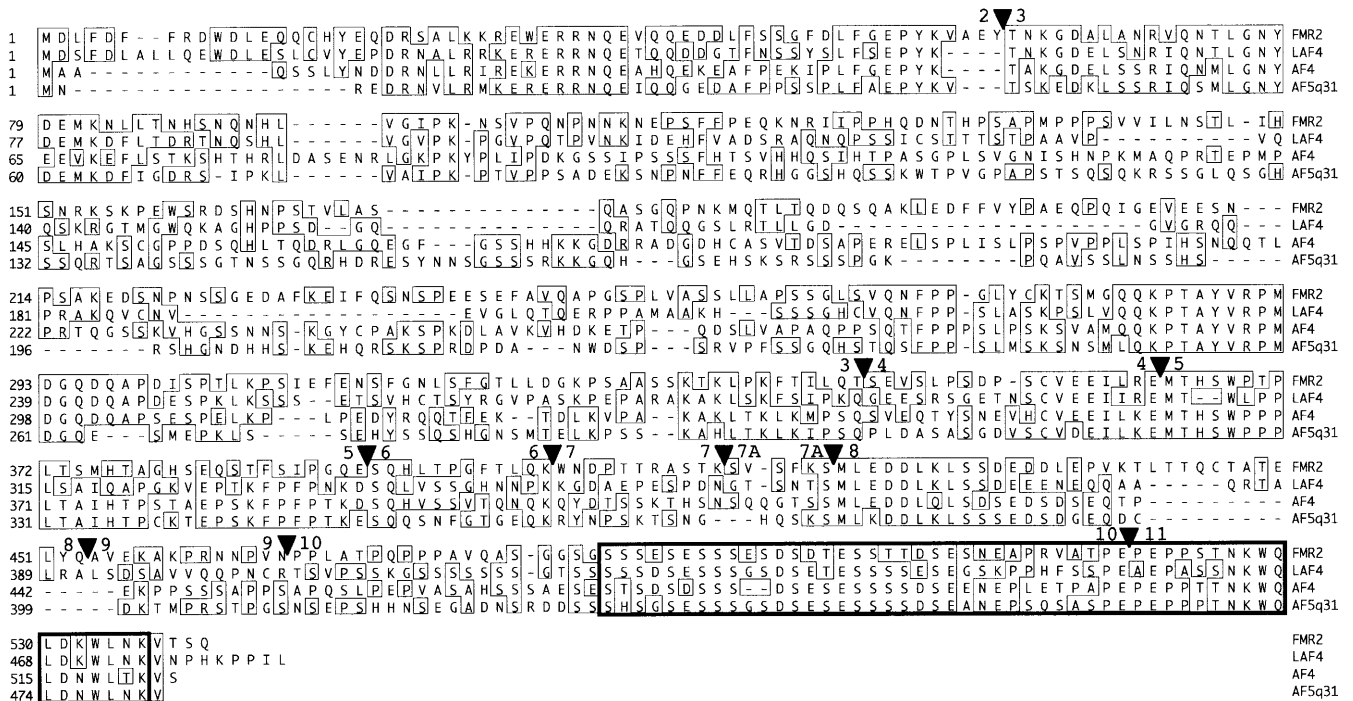


Fig. 1a,b. Construct design for activation experiments. **a** Schematic diagram of the constructs prepared and tested for transcription activation. *Dotted regions* indicate the presence of putative weak NH₂ terminal and strong central transcription activation domains. The COOH region of the proteins was not studied. Minimal domains for AF4/LAF4/FMR2/AF5q31 were predicted by the overlapping segments of AF4 and LAF4 previously shown to activate transcription. Amino acid

positions [AA] tested for transcription activation are indicated on the right of the diagram. Nomenclature for each construct tested is indicated. **b** Amino acid alignment of members of the AF4/LAF4/FMR2/AF5q31 protein family (NH₂ terminal region) demonstrating our predicted minimal transcription activation domain (*heavy box*). *Arrowheads* indicate exon/intron boundaries. *Boxed regions* are those that match the consensus exactly

Yeast cell transformations and β -galactosidase assays

An amount of 0.1 μ g of each experimental plasmid was transformed into competent *Saccharomyces cerevisiae* Y187 reporter strain, using the LiAc method protocol described in the TransAct Assay Kit (Clontech). *S. cerevisiae* Y187 contains an integrated *lacZ* reporter construct which is regulated by a wild-type GAL1 promoter. The GAL4 DNA binding domain has a high affinity for the GAL1 promoter. The recombinant pGBT9 derivative is transformed into Y187 where the fusion protein is expressed and targeted to the nucleus. Colonies containing the correct plasmid clone were detected via colony PCR, and then used to assay for transcription activation. If the protein of interest acted as an activator, then transcription of the *lacZ* gene would be initiated. This was detected by blue/white screening for β -galactosidase activity, using a colony lift filter assay, and further quantified using β -galactosidase assays of liquid cultures (Clontech).

For each transfection, the following controls were used: (1) pCL1 as a strong positive control, (2) pGBT9 + HA as a weak control, (3) the ALL1/pGBT9 construct containing the ALL1 activation domain, and (4) wild-type pM vector as a negative control. Transformations with all constructs were repeated three times.

Mammalian cell transfection experiments

A total of 5×10^5 HeLa cells in 100-mm culture dishes were transfected for 5h, using 14 μ g/ml of LipofectAMINE Reagent (GIBCO BRL) according to the manufacturer's specifications, with 2 μ g of each of the pM derivatives, 2 μ g of the pG5CAT reporter plasmid, and 0.4 μ g of the pSV- β plasmid, expressing β -galactosidase and used as an internal control to normalize transfection efficiency. Cellular extracts were prepared 48h post-transfection. CAT assays were performed as described in the Boehringer Mannheim CAT enzyme-linked immunosorbent assay (ELISA) Kit. The amount of extract used in the CAT assay was adjusted to 50 μ g/ μ l, based on the total amount of protein for each extract. Results of all CAT assays were normalized according to the β -galactosidase values.

For each transfection, the following controls were used: (1) the pM3-VP16 as a strong positive control containing the activation domain of the herpes virus, (2) the ALL1/pM construct encoding the ALL1 activation domain, and (3) wild-type pM vector as a negative control of transfection. Transfections with all constructs were repeated at least three times.

RT-PCR analysis

After transfection, cells were harvested and RNA was isolated from approximately 10^4 cells, using Trizol reagent (GIBCO BRL). To remove trace amounts of plasmid or genomic DNA, RNA was treated with 2U RNase-free DNase 1 for 30min at 37°C. DNase was removed from the sample by incubation with a DNase inactivation reagent

(Ambion, Austin, TX, USA) for 2min at room temperature. An amount of 1 μ g of purified total RNA was random primed for first-strand cDNA synthesis as described above. PCR amplifications using a GAL4-DNA-BD-specific forward primer (4.4: 5'-GACAGCATAGAATAAGTGCGAC-3') and exon 2 *FMR2* gene-specific reverse primer (*E2R1*: 5'-CGTCTTCTTGCTGGACTTCTTG-3') were carried out to ensure that recombinant plasmid-specific transcripts were produced. Standard conditions for PCR were 30s at 94°C, 30s at 55°C, and 30s at 72°C for 10 cycles, then 30s at 94°C, 30s at 60°C, and 30s at 72°C for 25 cycles, using the conditions described above. PCR products were analyzed on a 2% agarose gel.

Results

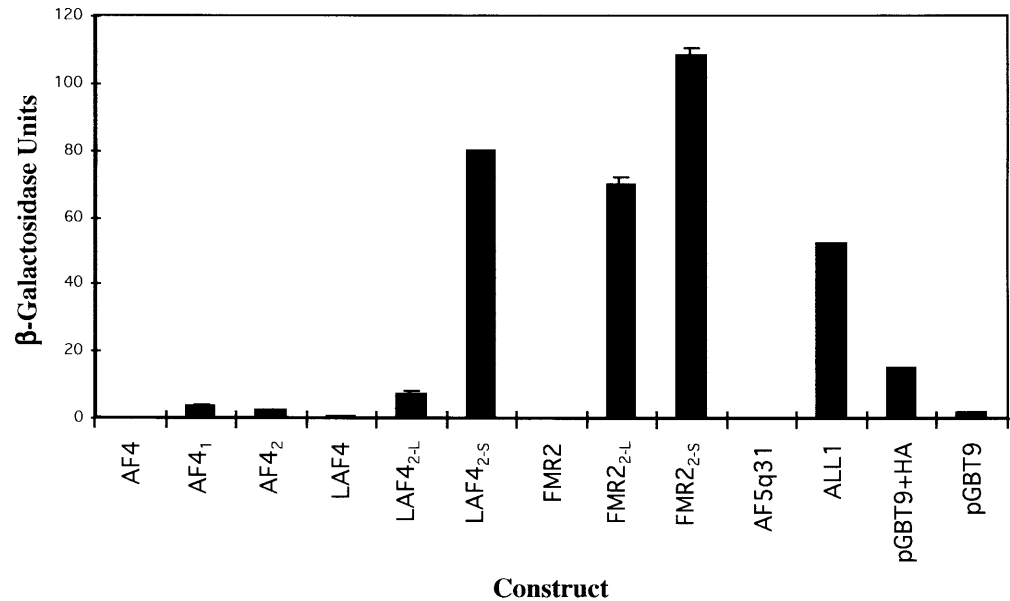
Definition of a minimal transactivation domain

Previous studies on AF4 and LAF4, two members of the *AF4/LAF4/FMR2/AF5q31* gene family (Prasad et al. 1995; Ma and Staudt 1996) indicated that these proteins may act as transcription activators. Initially, based on overlapping regions shown to activate transcription, small constructs containing the putative minimal activation domain were designed from all four members of the family: AF4 (AA464–523); LAF4 (AA426–484), FMR2 (AA481–541), and AF5q31 (Fig. 1a,b). However, assays using these “minimal constructs” in both yeast and mammalian cells showed no activation (see below). Subsequently, larger constructs, spanning most of the NH₂ terminus of the proteins, were designed, some of them identical to those previously tested (Prasad et al. 1995; Ma and Staudt 1996). Based on the AF4 and LAF4 activating regions, corresponding FMR2 protein constructs were designed. These comprised mostly the NH₂ terminus of the protein where extensive alternative splicing was observed (Gecz et al. 1997a). We did not prepare and analyze all possible FMR2 protein isoforms, focussing on the isoform missing the large exon 3 (~290 amino acids) and exon 5. By DNA sequencing we confirmed that the smaller isoform corresponded to FMR2 lacking exon 3, exon 5, and exon 7A (Gecz et al. 1997a). Our results showed that corresponding regions were also alternatively spliced in LAF4 (results not shown); thus, a construct lacking these regions in LAF4 was prepared as well. A schematic representation of all constructs used in the study is shown in Fig. 1a. The corresponding amino acid sequence is shown in Fig. 1b.

Transcription activation assays in *S. cerevisiae* Y187

All putative activation domains were initially tested in *S. cerevisiae*, using the β -galactosidase colony lift assay, and then using liquid o-nitrophenyl β -D-galactopyranoside (ONPG) assays. For all “minimal constructs” no activation was detected using either the colony lift assay (data not shown) or the more sensitive liquid ONPG assay (Fig. 2). Subsequently, the larger constructs were tested. Both

Fig. 2. Summary of the yeast *o*-nitrophenyl β -D-galactopyranoside (ONPG) activation assay. GAL4 DNA binding domain fusions in the pGBT9 vector were transformed into *Saccharomyces cerevisiae* Y187 cells. β -Galactosidase activity was assayed, and is shown as total β -galactosidase units for each sample. One unit of β -galactosidase is defined as the amount which hydrolyses 1 mmol of ONPG to *o*-nitrophenol and D-galactose per min. Error bars indicate SD



FMR2_{2-L} and FMR2_{2-S} augmented transcription of a reporter gene construct significantly, indicated by a rapid and strong color substrate reaction in both colony lift (data not shown) and liquid ONPG assays (Fig. 2). Interestingly, the shorter FMR2_{2-S} isoform, lacking exons 3, 5, and 7A, showed a much higher activation (40% higher, Fig. 2) than the larger exon 3-containing isoform FMR2_{2-L}. The level of activation of FMR2_{2-S} was higher than that of the control ALL1 construct (twofold), but less than that of the pCCL1 strong positive control (twofold; data not shown). The large FMR2_{2-L} construct activated at the level of the positive control ALL1 (Fig. 2). The LAF4_{2-L} and LAF4_{2-S} constructs with and without exon 3 showed potent activation, with the smaller (LAF4_{2-S}) isoform similarly showing higher (~tenfold) levels of transcription activation than the larger (LAF4_{2-L}; Fig. 2). AF4 was also shown to activate transcription (~tenfold less than ALL1), while AF5q31 did not. However, for this protein, only the AF5q31 minimal domain was tested, as the entire protein sequence was not available. In conclusion, the activation potential among the tested proteins was highest for FMR2.

Transcription activation assays in HeLa cells

All constructs used for transfection experiments were sequenced to confirm the open reading frame and to confirm that no mutations had occurred during the cloning process. Also, an internal control of protein expression (β -galactosidase expression plasmid pSV- β , see “Materials and methods”) was used in each transfection experiment. In order to confirm the in-vivo transcription from the recombinant pM vector constructs, total RNA was isolated from the transfected cells, DNaseI (to remove the recombinant plasmid DNA), reverse transcribed, and PCR amplified, using a combination of GAL4-DB- and FMR2-specific primers. An example of this experiment is shown in Fig. 3.

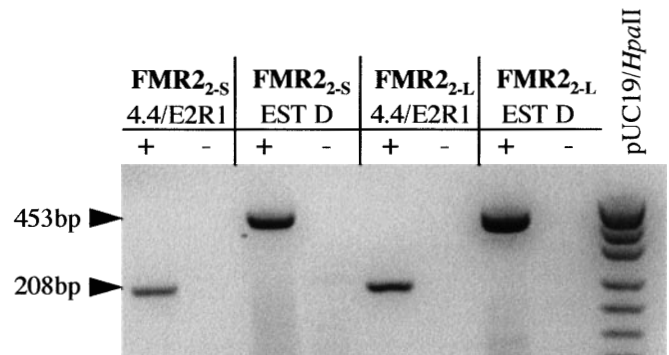


Fig. 3. Reverse transcription (RT)-polymerase chain reaction (PCR) analysis of the recombinant GAL4-DB/FMR2 RNA. As shown for the two FMR2 constructs, FMR2_{2-L} and FMR2_{2-S}, the recombinant RNA is produced from the transfected constructs. To control the quality of RNA, esterase D primers were used. The numbers above the lanes correspond to the oligonucleotides used (see Material and methods); + and - indicate lanes with and without the presence of reverse transcriptase in the cDNA synthesis reaction, as a marker, pUC/HpaII was used

Results of the HeLa assays (CAT ELISA) are shown in Fig. 4. All CAT activities were normalized to correct for any differences in transfection efficiency. The FMR2, AF5q31, LAF4, and AF4 “minimal” constructs did not exhibit any significant activation in HeLa cells, concordant with the results observed in yeast cells (Fig. 2). Both AF4 and LAF4 were seen to significantly augment transcription of the CAT reporter gene (Fig. 4). Consistent with the results of the yeast assay, the highest level of activation was observed for the short (FMR2_{2-S}) isoform. However, in this system, activation by FMR2_{2-S} was tenfold greater than that by LAF4_{2-S}, and about 40% higher with respect to the ALL1 control activation domain.

Fig. 4. Summary of the HeLa activation assay. GAL4 DNA binding domain fusions in the pM vector were cotransfected into HeLa cells along with a target promoter pG5CAT (5GAL4_{UAS}E1bCAT) and the β -galactosidase control vector (pSV- β). Chloramphenicol acetyltransferase (CAT) activity was measured 48h post-transfection, and activity shown as individual CAT values after normalization of results by assaying β -galactosidase levels as determined by pSV- β . Error bars indicate SD

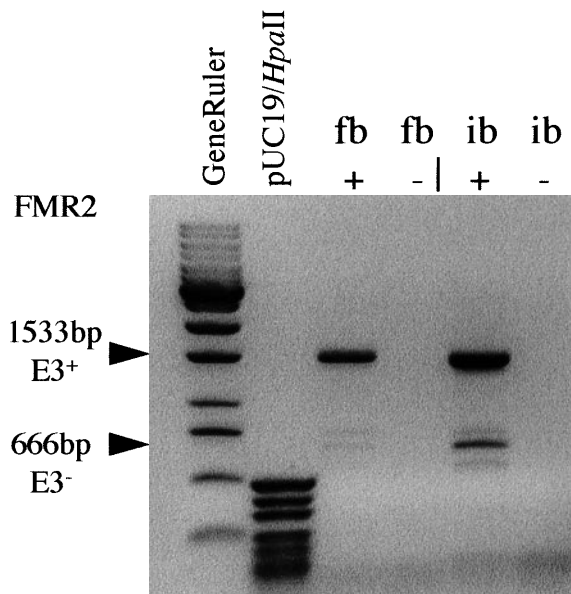
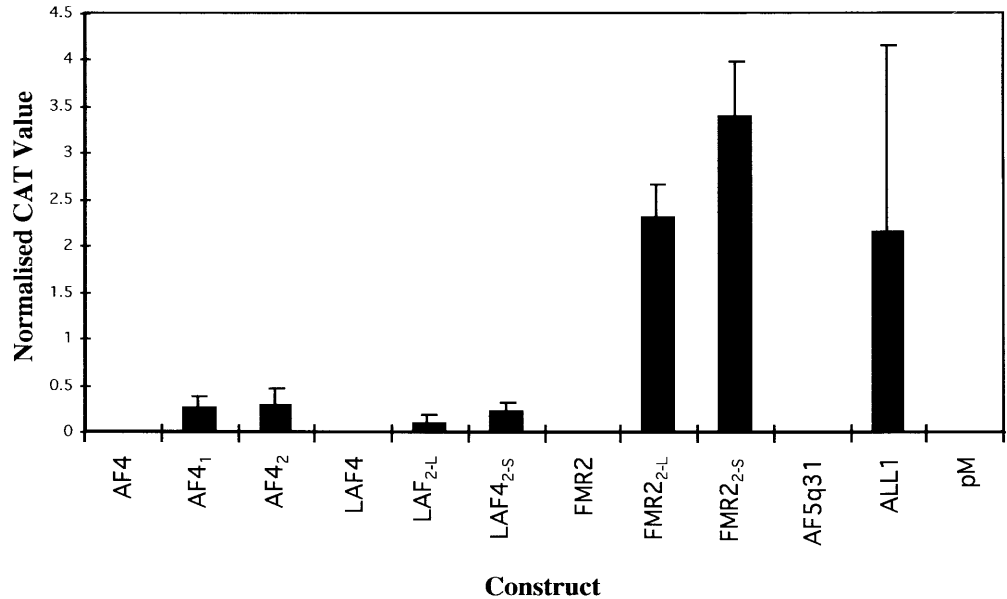


Fig. 5. RT-PCR analysis. Results for alternatively spliced *FMR2* isoforms showing variation in abundance between fetal and infant brain. PCR primers TAF3 and TAR3, covering the NH₂ terminus of the open reading frame (ORF) (*FMR2*_{2-S2-L}), were tested. The larger isoform (*FMR2*_{2-L}) was equally abundant in both tissues, but the smaller isoform (*FMR2*_{2-S}) showed markedly less abundance in infant brain. *fb* and *ib* indicate the source of the mRNA used (*fb*, fetal brain and *ib*, infant brain); + and - indicate lanes with and without the presence of reverse transcriptase in the cDNA synthesis reaction, as markers, GeneRuler (MBI) and pUC19/*Hpa*II (MBI) were used

Tissue specificity of *FMR2* exon 3 alternative splicing

It has been shown previously (Gecz et al. 1997a) that the 5' end of *FMR2* is alternatively spliced. To study the alternative splicing of the larger exon 3, fetal and infant brain mRNA was analyzed. Although the most abundant is the exon 3-containing isoform (*FMR2*_{2-L}), a significant difference in abundance of the exon 3-missing isoform (*FMR2*_{2-S})

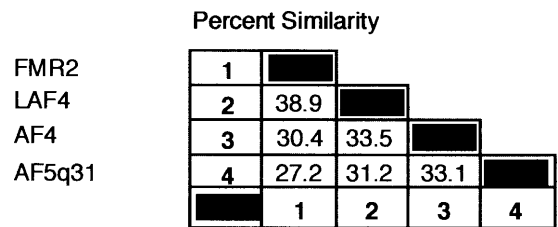


Fig. 6. Protein sequence distances among the members of the AF4/AF5q31/FMR2/LAF4 family. Only the NH₂ region of the corresponding proteins studied (as shown in Fig. 1b) have been analyzed. Percentages were calculated using ClustalW (DNA Star, Madison, WI, USA)

was detected (Fig. 5). Consistently, the level of the exon 3-missing isoform is higher in infant brain than in fetal brain. However, the overall expression of *FMR2* seems to be upregulated in infant brain.

Further analysis of the *FMR2* protein

To gain insight into the basis for different transcription activation potentials of the four proteins studied, corresponding paralogous regions (exon 1–exon 10) of these proteins were examined (Fig. 1b). Amino acid sequence comparison showed that conservation across this region (Fig. 6) is in the range of 27% (*FMR2* versus *AF5q31*) to 38.9% (*FMR2* versus *LAF4*). After removal of the sequence of the large exon 3 (287 amino acids for the *FMR2* protein), the overall conservation was slightly higher (40%–53%; results not shown). We speculated that differences of transcription activation potential observed, at least for the *FMR2* and *LAF4*, might be a consequence of different acidity and, thus, different charge of these proteins. Both the large *FMR2*_{2-L} and *LAF4*_{2-L} constructs tested have a relatively high negative charge, –28.29 and –7.3, respectively. The deduced amino acid sequence of the *FMR2*_{2-L} protein

shows that it has a primary length of 511 amino acids with an estimated mass of 56.4kDa and is rich in serine (68 residues; 13.3%) and proline (51 residues; 10.0%) amino acids. The larger LAF4_{2-L} isoform has a length of 460 amino acids, with an estimated mass of 50.3kDa and is rich in serine (64 residues; 13.9%) and proline (39 residues; 8.5%) amino acids. Removal of the amino acids of the alternatively spliced exon 3 and exons 5 and 7A results in an increase in the negative charge of LAF4 (-16.6), AF4 (from -15.0 to -23.8), and AF5q31 (from +0.37 to -15.13) proteins. For FMR2, the negative charge is slightly lower (-22.29). The amino acid sequence of the FMR2_{2-S} protein has a primary length of 222 residues with an estimated mass of 25.1kDa and is rich in serine (28 residues; 12.6%), glutamine (25 residues; 11.3%), and proline (19 residues; 8.6%) amino acids. Lastly, the deduced amino acid sequence of the LAF4_{2-S} protein shows that it has a primary length of 199 amino acids with an estimated mass of 21.6kDa and is rich in serine (38 residues; 19.1%), glutamine (22 residues; 11.1%), and proline (18 residues; 9.0%) amino acids. In conclusion, the increase in transcription activation potential of the protein constructs tested in this study (especially FMR2) correlated with the negative charge of that protein.

Discussion

FMR2 belongs to a family of four genes *AF4*, *LAF4*, *FMR2*, and *AF5q31*. Three of these genes were found to be associated with morbid phenotype; *AF4* (Prasad et al. 1995), *AF5q31* (Taki et al. 1999), and *FMR2* (Gecz et al. 1996; Gu et al. 1996). Two of them, *AF4* and *AF5q31*, were found to be involved in acute lymphoblastic leukemia (ALL), as rearrangement partners of the *ALL1* gene. An *LAF4*-associated phenotype is not yet known; however, based on its expression, it is thought to have a role in lymphoid development (Ma and Staudt 1996). *FMR2* is involved in a mild-to-borderline form of nonspecific X-linked mental retardation (Gecz et al. 1997b; Gecz 2000).

Our aims for this study were: (i) primarily to determine whether *FMR2* possessed a transcription activator function; (ii) to test whether this function was conserved; (iii) to assess levels of transcription activation by *FMR2* in comparison with that by other members of the gene family; and (iv) to compare this activity with that of other known transcription activators (e.g., *ALL1*, *VP16*).

We tested different regions of *FMR2* in comparison to the other member proteins for transcription activation in yeast and HeLa cells. Previous studies on both *AF4* (Prasad et al. 1995; Morrissey et al. 1997) and *LAF4* (Ma and Staudt 1996) have shown that these proteins can activate transcription. In our study, we show that *FMR2* does efficiently activate the *GAL4-E1b GAL1* and *CAT* reporter in both yeast and HeLa cells, respectively (Figs. 2,4). Our results also confirm previous studies on *LAF4* and *AF4* (Prasad et al. 1995; Ma and Staudt 1996). Regarding the *AF5q31* gene, only a partial protein sequence was available until recently

(Frestedt et al. 1996; Gecz et al. 1997a). Taki et al. (1999) cloned and characterized this gene, which they called *AF5q31*. The *AF5q31* gene encodes a 1163-amino acid protein, which is similar to *AF4*, *LAF4*, and *FMR2*. We did not find transcriptional activation for the small *AF5q31* construct, but a larger region was not tested. Based on our studies and conservation across the tested regions, we predict that the *AF5q31* protein might also act as a transcription activator protein. In conclusion, the members of the *AF4/LAF4/FMR2/AF5q31* family of proteins are able to activate transcription in yeast as well as mammalian cells.

Observation of alternative splicing within *FMR2* and *LAF4* and its role in transcription activation is interesting. Comparison of fetal and infant brain cDNAs showed that *FMR2* isoforms, lacking especially the large exon 3, had different abundance in either tissue. The smaller isoform is relatively more abundant in fetal brain than in infant brain. Smaller isoforms for either *FMR2* or *LAF4* showed consistently augmented levels of transcription activation. Thus, it would be of great interest to determine the spatial and temporal pattern of transcription of either of the *FMR2*_{2-L} *FMR2*_{2-S} isoforms. Differential transcription activation due to alternative splicing is not exclusive to *FMR2*; it has been documented for other proteins, including *SMAD2* (Yagi et al. 1999), *BRCA1* (Cui et al. 1998), and the orphan nuclear receptor *CAR* (Choi et al. 1997).

Alternative splicing of transcripts can result in proteins that have lost the ability to bind to a specific target sequence (Choi et al. 1997) or proteins that have increased ability to bind DNA (Yagi et al. 1999). Alternatively, splicing may lead to removal of an activation region (Cui et al. 1998) and subsequent aberrant transcription activation capacity. Our study has shown that removal of the large exon 3 in *FMR2* and *LAF4* acts to increase transcription activation. Removal of a repressive domain, in a similar mechanism to that of the MH2 domain in *SMAD2* allowing the MH1 domain to bind DNA (Yagi et al. 1999), may also be the case for *FMR2*. Conversely, physical changes in protein structure may effect activation levels. Travers (1998) discusses the idea that conformational changes in DNA and proteins are important for complex formations between transcription factors. This is interesting in light of our findings that alternative splicing results in varied transcription activation levels. It is also interesting to note that the larger product is more abundant than the smaller isoform. This variation in abundance may act as a regulatory mechanism for controlling transcriptional activation by the *FMR2* protein.

The function of transcription activation is to augment the basal levels of transcription of eukaryotic mRNA by RNA polymerase II. Several models for this action have been proposed (Mitchell and Tjian 1989; Triezenberg 1995; Kingston et al. 1996; Ptashne and Gann 1997). Studies of the herpes simplex *VP16* activation domain have shown that activation processes require biochemical interactions between activation region(s) and specific target protein(s) that depend most directly upon hydrophobic residues within helical segments of the activator (Sullivan et al. 1998). Paralogous regions of the *AF4/LAF4/FMR2/AF5q31*

family were analyzed in this respect, and these regions could therefore be correlated with their similar activation potential. The LAF4 and AF4 proteins have also been shown to have a nonspecific DNA binding potential (Prasad et al. 1995; Ma and Staudt 1996). Goodman and Mandel (1998) have suggested that transcription activators do not necessarily bind DNA directly. Rather, these transcription activators may act by modifying chromatin factors that repress gene expression, by linking DNA binding transcription activators to the basal transcription apparatus, or by altering components of the transcriptional machinery through covalent modifications. FMR2 and other members of the family may act by interacting with other proteins of the transcriptional machinery along these lines. The DNA binding capacity of FMR2 has not yet been tested. Computer analysis of FMR2 shows that the protein does not contain a recognized DNA binding domain motif, but does contain a helix-loop-helix dimerization domain signature (ETVELLRYA¹⁰⁹⁰⁻¹⁰⁹⁸), a domain that is known to interact with other proteins to form heterodimers (Dowell et al. 1992). The role of FMR2 (and the other members of the family) in transcription activation is also supported by the nuclear localization of the corresponding protein (Gecz et al. 1997a). Interestingly, the dispersed pattern of speckled nuclear localization for some FMR2-GFP constructs (data not shown) is very similar to the localization of other members of the transcription machinery (Jolly et al. 1999; Verschure et al. 1999; Berube et al. 2000).

Currently, at least eight genes have been identified as being involved in a form of nonspecific X-linked mental retardation: *FMR2* (Gecz et al. 1996; Gu et al. 1996); *GDII*, a Rab GDP-disassociation inhibitor (D'Adamo et al. 1998); *OPHNI*, a Rho-GTPase-activating protein (Billuart et al. 1998); *PAK3*, a p21-activated kinase (Allen et al. 1998); *RPS6KA3*, a growth factor-induced kinase (Merienne et al. 1999); *IL1RAPL*, an interleukin (IL)-1 receptor accessory-like protein (Carrie et al. 1999); *TM4SF2*, a member of the tetraspanin family of proteins (interaction with β 1-integrins; Zemni et al. 2000); and *ARHGEF6*, a guanine nucleotide exchange factor for Rho GTPase (Kutche et al. 2000). Only FMR2 has not yet been shown to participate in intracellular signalling (Chelly 1999; Gecz 2000; Gecz and Mulley 2000). Elucidation of a function for FMR2 in the activation of transcription may place FMR2 as one of the downstream effectors of the newly discovered *MRX* genes. The function of the FMR2 protein can be regulated in a manner similar to that of the CREB transcription factor, via phosphorylation by RSK2 (RPS6KA3; Merienne et al. 1999) or yet another kinase. A putative phosphorylation site was identified in FMR2 in the vicinity of the bipartite nuclear localization sequences (J. Gecz; unpublished results). The proposed role of FMR2 in transcription regulation is also attractive from the point of view of the mild nature of the FMR2-associated mental retardation phenotype. Moreover, although the function of FMR2 as a transcription regulator seems to be conserved (yeast activation study results; see above), most probably it is also, at least partially, redundant. This can be supported by the existence of several examples of individuals who do not have any

detectable FMR2 product and who are not affected (Lo Nigro et al. 2000; for review, see Gecz 2000).

FMR2-associated mental retardation is the most prevalent form of nonspecific mental retardation yet characterized (Gecz 2000). Although our study does not provide comprehensive evidence of FMR2 function, it points toward a pathway and molecular process in which FMR2 may take part. Investigation of proteins that interact with FMR2, and genes in which transcription is regulated by FMR2, will be the next step in the elucidation of its function and its role in mental retardation.

Acknowledgments The authors gratefully acknowledge Ingrid Lensink for advice on HeLa transfections and CAT assay technology. This work was supported by the National Health and Medical Research Council of Australia and the Women's and Children's Hospital Research Foundation.

References

- Allen KM, Gleeson JG, Bagrodia S, Partington MW, MacMillan JC, Cerione RA, Mulley JC, Walsh CA (1998) *PAK3* mutation in nonsyndromic X-linked mental retardation. *Nature Genet* 20:25-30
- Berube NG, Smeenk CA, Picketts DJ (2000) Cell cycle-dependent phosphorylation of the ATRX protein correlates with changes in nuclear matrix and chromatin association. *Hum Mol Genet* 9:539-547
- Billuart P, Bienvenu T, Ronce N, des Portes V, Vinet MC, Zemni R, Crollius HR, Carrie A, Fauchereau F, Cherry M, Briault S, Hamel B, Fryns JP, Beldjord C, Kahn A, Moraine C, Chelly J (1998) *Oligophrenin-1* encodes a rhoGAP protein involved in X-linked mental retardation. *Nature* 392:923-926
- Brown WT (1996) The *FRAXE* syndrome. Is it time for routine screening? *Am J Hum Genet* 58:903-905
- Carrie A, Jun L, Bienvenu T, Vinet MC, McDonell N, Couvert P, Zemni R, Cardona A, Van Buggenhout G, Frints S, Hamel B, Moraine C, Ropers HH, Strom T, Howell GR, Whittaker A, Ross MT, Kahn A, Fryns JP, Beldjord C, Marynen P, Chelly J (1999) A new member of the IL-1 receptor family highly expressed in hippocampus and involved in X-linked mental retardation. *Nature Genet* 23:25-31
- Chakrabarti L, Knight SJL, Flannery AV, Davies KE (1996) A candidate gene for mild mental handicap at the *FRAXE* fragile site. *Hum Mol Genet* 5:275-282
- Chelly J (1999) Breakthroughs in molecular and cellular mechanisms underlying X-linked mental retardation. *Hum Mol Genet* 8:1833-1838
- Choi HS, Chung M, Tzamelis I, Simha D, Lee YK, Seol W, Moore DD (1997) Differential transactivation by two isoforms of the orphan nuclear hormone receptor CAR. *J Biol Chem* 272:23565-23571
- Cui JQ, Wang H, Reddy ES, Rao VN (1998) Differential transcriptional activation by the N-terminal region of *BRCA1* splice variants *BRCA1a* and *BRCA1b*. *Oncol Rep* 5:585-589
- D'Adamo P, Menegon A, Lo Nigro C, Grasso M, Gulisano M, Tamanini F, Bienvenu T, Gedeon AK, Oostra B, Wu SK, Tandon A, Valtorta F, Balch WE, Chelly J, Toniolo D (1998) Mutations in *GDII* are responsible for X-linked non-specific mental retardation [published erratum appears in *Nat Genet* 1998 Jul;19:303]. *Nature Genet* 19:134-139
- Dowell SJ, Tsang JS, Mellor J (1992) The centromere and promoter factor 1 of yeast contains a dimerisation domain located carboxy-terminal to the bHLH domain. *Nucleic Acids Res* 20:4229-4236
- Frestedt JL, Hilden JM, Kersey JH (1996) *AF4/FEL*, a gene involved in infant leukemia: sequence variations, gene structure, and possible homology with a genomic sequence on 5q31. *DNA Cell Biol* 15:669-678
- Gecz J (2000) The *FMR2* gene, *FRAXE* and non-specific X-linked mental retardation: clinical and molecular aspects. *Ann Hum Genet* 64:95-106

- Gez J, Mulley JC (1999) Characterisation and expression of a large, 13.7kb *FMR2* isoform. *Eur J Hum Genet* 7:157–162
- Gez J, Mulley J (2000) Genes for cognitive function: developments on the X. *Genome Res* 10:157–163
- Gez J, Gedeon AK, Sutherland GR, Mulley JC (1996) Identification of the gene *FMR2*, associated with *FRAAXE* mental retardation. *Nature Genet* 13:105–108
- Gez J, Bielby S, Sutherland GR, Mulley JC (1997a) Gene structure and subcellular localization of *FMR2*, a member of a new family of putative transcription activators. *Genomics* 44:201–213
- Gez J, Oostra BA, Hockey A, Carbonell P, Turner G, Haan EA, Sutherland GR, Mulley JC (1997b) *FMR2* expression in families with *FRAAXE* mental retardation. *Hum Mol Genet* 6:435–441
- Goodman RH, Mandel G (1998) Activation and repression in the nervous system. *Curr Opin Neurobiol* 8:413–417
- Gu YH, Shen Y, Gibbs RA, Nelson DJ (1996) Identification of *FMR2*, a novel gene associated with the *FRAAXE* CCG repeat and CpG island. *Nature Genet* 13:109–113
- Jolly C, Voure'h C, Robert-Nicoud M, Morimoto RI (1999) Intron-independent association of splicing factors with active genes. *J Cell Biol* 145:1133–1143
- Kutche K, Yntema H, Brandt A, Jantke I, Northwang HG, Orth U, Boavida MG, David D, Chelly J, Fryns JP, Moraine C, Ropers HH, Hamle BCJ, Bokhoven H, Gal A (2000) Mutations in *ARHGEF6*, encoding guanine nucleotide exchange factor for Rho GTPases, in patients with X-linked mental retardation. *Nature Genet* 26:247–250
- Kingston RE, Bunker CA, Imbalzano AN (1996) Repression and activation by multiprotein complexes that alter chromatin structure. *Genes Dev* 10:905–920
- Lo Nigro C, Faravelli F, Cavani S, Perroni L, Novello P, Vitali M, Bricarelli FD, Grasso M (2000) *FRAAXE* mutation in a mentally retarded subject and in his phenotypically normal twin brother. *Eur J Hum Genet* 8:157–162
- Ma C, Staudt LM (1996) *LAF-4* encodes a lymphoid nuclear protein with transactivation potential that is homologous to *AF-4*, the gene fused to MLL in t(4;11) leukemias. *Blood* 87:734–745
- Merienne K, Jacquot S, Pannetier S, Zeniou M, Bankier A, Gez J, Mandel JL, Mulley J, Sassone-Corsi P, Hanauer A (1999) A missense mutation in *RPS6KA3* (*RSK2*) responsible for non-specific mental retardation. *Nature Genet* 22:13–14
- Mitchell PJ, Tjian R (1989) Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* 245:371–378
- Morrissey J, Tkachuk DC, Milatovich A, Francke U, Link M, Cleary ML (1993) A serine/proline-rich protein is fused to HRX in t(4;11) acute leukemias. *Blood* 81:1124–1131
- Morrissey JJ, Raney S, Cleary ML (1997) The FEL (*AF-4*) protein donates transcriptional activation sequences to HRX-FEL fusion proteins in leukemias containing T(4;11)(Q21;Q23) chromosomal translocations. *Leuk Res* 21:911–917
- Mulley JC, Yu S, Loesch DZ, Hay DA, Donnelly A, Gedeon AK, Carbonell P, Lopez I, Glover G, Gabarron I, Yu PWL, Baker E, Haan EA, Hockey A, Knight SJL, Davies KE, Richards RI, Sutherland GR (1995) *FRAAXE* and mental retardation. *J Med Genet* 32:162–169
- Murgia A, Polli R, Vinanzi C, Salis M, Drigo P, Artifoni L, Zacchello F (1996) Amplification of the Xq28 *FRAAXE* repeats—extreme phenotype variability. *Am J Med Genet* 64:441–444
- Nilson I, Reichel M, Ennas MG, Greim R, Knorr C, Siegler G, Greil J, Fey GH, Marschalek R (1997) Exon/intron structure of the human *AF-4* gene, a member of the *AF-4/LAF-4/FMR-2* gene family coding for a nuclear protein with structural alterations in acute leukaemia. *Br J Haematol* 98:157–169
- Prasad R, Yano T, Sorio C, Nakamura T, Rallapalli R, Gu Y, Leshkowitz D, Croce CM, Canaani E (1995) Domains with transcriptional regulatory activity within the ALL1 and AF4 proteins involved in acute leukemia. *Proc Natl Acad Sci USA* 92:12160–12164
- Ptashne M, Gann A (1997) Transcriptional activation by recruitment. *Nature* 386:569–577
- Sullivan SM, Horn PJ, Olson VA, Koop AH, Niu W, Ebright RH, Triezenberg SJ (1998) Mutational analysis of a transcriptional activation region of the VP16 protein of *Herpes simplex* virus. *Nucleic Acids Res* 26:4487–4496
- Taki T, Kano H, Taniwaki M, Sako M, Yanagisawa M, Hayashi Y (1999) *AF5q31*, a newly identified *AF4*-related gene, is fused to MLL in infant acute lymphoblastic leukemia with ins(5;11)(q31;q13q23). *Proc Natl Acad Sci USA* 96:14535–14540
- Travers A (1998) Transcription: activation by cooperating conformations. *Curr Biol* 8:616–618
- Triezenberg SJ (1995) Structure and function of transcriptional activation domains. *Curr Opin Genet Dev* 5:190–196
- Verschure PJ, van Der Kraan I, Manders EM, van Driel R (1999) Spatial relationship between transcription sites and chromosome territories. *J Cell Biol* 147:13–24
- Yagi K, Goto D, Hamamoto T, Takenoshita S, Kato M, Miyazono K (1999) Alternatively spliced variant of *SMAD2* lacking exon 3. Comparison with wild-type *SMAD2* and *SMAD3*. *J Biol Chem* 274:703–709
- Zemni R, Bienvenu T, Vinet MC, Sefiani A, Carrie A, Billuart P, McDonnell N, Couvert P, Francis F, Chafey P, Fauchereau F, Friocourt G, Portes V, Cardona A, Frints S, Meindl A, Brandau O, Ronce N, Moraine C, Bokhoven H, Ropers HH, Sudbrak R, Kahn A, Fryns JP, Beldjord C, Chelly J (2000) A new gene involved in X-linked mental retardation identified by analysis of an X;2 balanced translocation. *Nature Genet* 24:167–170