The N-Terminus of the Fragile X Mental Retardation Protein Contains a Novel Domain Involved in Dimerization and RNA Binding†

S. Adinolfi,‡ A. Ramos,‡ S. R. Martin,‡ F. Dal Piaz,§ P. Pucci,|| B. Bardoni,⊥ J. L. Mandel,⊥ and A. Pastore*,‡

National Institute for Medical Research, The Ridgeway, NW7 1AA London, U.K., Centro Interdipartimentale per la Ricerca Biotecnologica, Universitá degli Studi di Bologna, Via Irnerio 42, Bologna, Italy, Universitá di Napoli, Via Cinthia, 80137 Napoli, Italy, and Institut de Génétique et de Biologie Moleculaire et Cellulaire, CNRS/INSERM/ULP, BP 10142, 1 rue Laurent Fries, 67404 Illkirch Cedex, France

Received May 28, 2003; Revised Manuscript Received July 7, 2003

ABSTRACT: Fragile X syndrome, the most common cause of inherited mental retardation, is caused by the absence of the fragile X mental retardation protein (FMRP). The emerging picture is that FMRP is involved in repression of translation through a complex network of protein−protein and protein−RNA interactions. Very little structural information is, however, available for FMRP that could help to understand its function. In particular, no structural studies are available about the N-terminus of the protein, a highly conserved region which is involved in several molecular interactions. Here, we explore systematically the ability of the FMRP N-terminus to form independently folded units (domains). We produced deletion mutants and tested their fold and functional properties by mutually complementary biophysical and biochemical techniques. On the basis of our data, we conclude that the N-terminus contains a domain, that we named NDF, comprising the first 134 amino acids. Most interestingly, NDF comprises two copies of a newly identified Agenet motif. NDF is thermally stable and has a high content of β structure. In addition to being able to bind to RNA and to recognize some of the FMRP interacting proteins, NDF forms stable dimers and is able to interact, although weakly, with the full-length protein. Our data provide conclusive evidence that NDF is a novel motif for protein−protein and protein−RNA interactions and contains a previously unidentified dimerization site.

Fragile X mental retardation syndrome, the most common cause of inherited mental retardation in humans, is caused by pathological expansion of a CGG trinucleotide repeat in the 5′ untranslated region (5′UTR)[1] of the *FMR1* gene (for recent reviews, see refs *1* and *2*). As a consequence of the expansion, the *FMR1* locus is usually silenced at the transcriptional level with consequent absence of the *FMR1* gene product, the FMR1 protein (FMRP), in affected individuals. Although ubiquitous, FMRP is most abundant in testes and brain, the organs most affected by the syndrome (*3*−*6*). FMRP is a cytoplasmic protein that is also found in distal dendrites, where its local concentration increases in response to neurotransmitter activation (*7*, *8*).

The cellular role of FMRP is still mostly unknown, despite the direct correlation between anomalies in the expression of the protein and the occurrence of fragile X syndrome. The current view is that FMRP is a negative regulator of translation in a manner critical for the development of neurons (*9*−*12*). To achieve this function, FMRP is an RNA-binding protein that forms an extended network of interactions with several nuclear and cytoplasmic protein partners (*12*−*16*). Several RNAs have been identified in vitro and in vivo as its potential partners (*9*, *17*−*20*). Among them, a small nontranslated RNA, BC1, has been shown to mediate the interaction of FMRP with specific mRNAs at synapses (*11*). FMRP is associated with polyribosomes as a component of a large mRNP complex which shuttles between polyribosomes and cytoplasmic granules (*12*, *21*, *22*). Direct interactions with the two close paralogues FXR1P and FXR2P and with the proteins NUFIP1, CYFIP1, and CYFIP2 have also been demonstrated (*23*−*26*). FMRP is able to heterodimerize in vitro with FXR1P and FXR2P (*26*, *27*) but seems mainly to homodimerize in vivo (*28*).

FMRP is a medium-size protein (632 residues long) (Figure 1). Its sequence comprises an N-terminal block of 444 residues (which spans up to the end of exon 14) with a high degree of conservation within the FMRP family (70−80% identity and 80−90% similarity) (*26*). This region contains two copies of the K-homology (KH) motif preceded by an N-terminus ca. 220 long. Conversely, no conservation is observed at the protein C-terminus, where an RGG box, a motif typically involved in nonspecific RNA recognition, is observed. Interestingly, many of the interactions so far

† The work was financed by a Human Frontier Grant (RGP0052/2001-B).

* To whom correspondence should be addressed.
‡ National Institute for Medical Research.
§ Universitá degli Studi di Bologna.
|| Universitá di Napoli.
⊥ Institut de Génétique et de Biologie Moleculaire et Cellulaire.

[1] Abbreviations: CD, circular dichroism; EM, electron microscopy; FMRP, fragile X mental retardation protein; GST, glutathione *S*-transferase; IPTG, isopropyl β-D-thiogalactopyranoside; KH motif, K-homology motif; NDF, N-terminal domain of FMRP; NMR, nuclear magnetic resonance; RNP, ribonucleoproteins; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; TEV, tobacco etch virus; TFA, trifluoroacetic acid; UTR, untranslated region; UV, ultraviolet.

10.1021/bi034909g CCC: $25.00 © 2003 American Chemical Society
Published on Web 08/12/2003

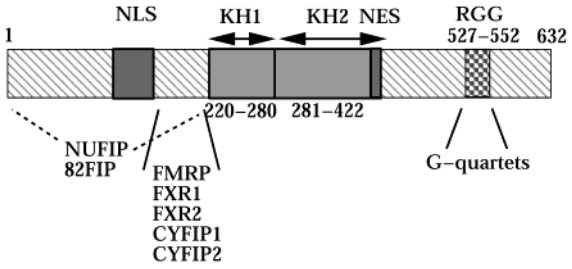

FIGURE 1: Schematic representation of FMRP showing the sequence motifs and their position along the sequence. NLS and NES stand for nuclear localization signal and nuclear export signal, respectively.

identified have been mapped to the N-terminus (*16*). This observation raises several important questions: How does the N-terminus recognize its partners? Is there any structural motif that is involved in the interactions? Do all the interactions compete for the same site by a subtle mechanism of mutual competition or are there distinct interacting sites? However, structural information is so far only available for the first KH module of FMRP (*29*), and very little is known about the structure of the FMRP N-terminus.

In this study, we describe the results of an extensive biochemical and biophysical characterization of the N-terminus of FMRP with the aim of understanding the role of this region in the molecular interactions known so far. We show that the N-terminus contains an independently folded domain localized in the first 134 residues. This region is able to bind to RNA as well as to homodimerize and to interact with other protein partners.

MATERIALS AND METHODS

Sample Preparation. The FMRP constructs used for this study are FMRP(1−217), FMRP(1−214), FMRP(1−180), FMRP(1−134), and FMRP(80−160), where the numbers in parentheses indicate the regions of the FMRP sequence. FMRP(1−217) was used for all of the biophysical studies, whereas, for comparison with previous data (*30*), FMRP(1−214) was used for the binding assays. All FMRP fragments were amplified by PCR with engineered *Nco*I on 5′ ends and *Not*I on 3′ ends from the complete cDNA. The constructs were cloned in a pET9-derived plasmid as fusion proteins either with a six histidine tag and/or glutathione *S*-transferase (GST). All clones were expressed in *Escherichia coli* BL21(DE3). The cells were grown in LB medium with either ampicillin (100 mg/L) or kanamycin (30 mg/L), induced for 3−4 h by addition of 0.5 mM isopropyl β-D-thiogalactopyranoside (IPTG) when the cultures reached an optical density of 0.8 at 600 nm. The cells were harvested (Beckman centrifuge) and resuspended in lysis buffer (20 mM Tris-HCl at pH 8, 200 mM NaCl, 0.2% IGEPAL CA-630, 2 mM β-mercaptoethanol, lysozyme from Sigma, and Dnase I and antiproteases from Boehringer Mannheim), sonicated (Branson sonifier, model 250/450) for 5 min at amplitude 5, and centrifuged at 18000 rpm for 40 min. The proteins were purified by affinity chromatography (using Ni-NTA gel or glutathione−Sepharose). The fusion protein or the His tag was removed, when necessary, by tobacco etch virus (TEV) protease. When necessary an additional affinity chromatography step of purification was performed to separate GST from the proteins of interest. One further step of purification was performed by gel filtration chromatography using a G75 16/60 column (Pharmacia). The purity of the recombinant proteins was checked by SDS−PAGE after each step of purification and by mass spectroscopy of the final products.

Limited Proteolysis. The proteins were incubated with trypsin, chymotrypsin, subtilisin, and endoprotease V8 using enzyme-to-substrate ratios (w/w) ranging from 1:100 to 1:250. The extent of each reaction was monitored by sampling the reaction mixture at different incubation times. Proteolytic fragments were fractionated by reverse-phase HPLC on a Phenomenex Jupiter C18 column (250 mm × 2.1 mm) with a linear gradient from 5% to 60% of acetonitrile containing 0.1% TFA over 60 min. Mass spectrometry identification of collected chromatographic fractions was performed by an API100 electron-spray mass spectrometer (Applied Biosystems) equipped with a quadrupolar analyzer.

Experiments To Probe the Aggregation State of the Constructs. FPLC gel filtration chromatography was performed using a G-200 10/30 High-Load column. The concentrations were 32, 7.1, and 35 μM for FMRP(1−217), FMRP(1−180), and FMRP(1−134), respectively.

Sedimentation equilibrium experiments were carried out using a Beckman XL-A analytical ultracentrifuge equipped with UV absorption optics. Protein concentrations of 20 μM were chosen so that the absorbance at 280 nm was in the 0.2−1.0 range (12 mm path length). The molar extinction coefficients of the three samples are 27790 cm^{-1} M^{-1} for FMRP(1−217) and FMRP(1−180) and 25100 cm^{-1} M^{-1} for FMRP(1−134). The measurements were carried out at 15 and 25 °C using speeds of 12000 and 16000 rpm for FMRP(1−217), 18000 and 22000 rpm for FMRP(1−180), and 16000 and 20000 rpm for FMRP(1−134). Data were recorded over an interval of 16 h. Each measurement was repeated after 8 h to ensure that equilibrium had been reached and that proteolysis was not occurring. In all data sets, the absorbance of the baseline was obtained by fitting the data at different speeds. A two-component self-association mechanism was assumed to allow for dimerization. The data were analyzed with the Origin XL-A/XL1 package (Beckman).

Light scattering experiments were performed on a luminescent spectrometer. Measurements were carried out at 20 °C on 200 μL samples using 50, 70, and 200 μM protein concentrations, in 20 mM Tris-HCl at pH 8, in a 1 cm path-length fluorescence cuvette with excitation and emission wavelengths of 780 nm. The protein solutions were filtered with a 0.02 μM cutoff filter prior to the measurements. UV spectra in the range 250−450 nm were also recorded to follow aggregation. Temperature-dependent measurements were performed in the range 20−70 °C using 11 μM samples. Once at 70 °C, reversibility was tested by lowering the temperature and recording a spectrum at 20 °C.

Experiments To Probe the Secondary and Tertiary Structure of the Constructs. Circular dichroism (CD) spectra were recorded on a Jasco J-715 spectropolarimeter equipped with a thermostatically controlled cell holder stabilized by circulating water from a Neslab RTE-110 water bath. Rectangular quartz cuvettes (Hellma) with 0.01 and 0.1 cm path lengths were used. Ten scans were averaged, and an appropriate buffer baseline was subtracted. The spectra were recorded in 20 mM Tris-HCl at pH 8 using protein concentrations of

6 and 100 μM. Temperature-dependent unfolding and refolding were followed in the 20−75 °C range by monitoring CD at 222 nm using a heating rate of 1 °C/min.

Nuclear magnetic resonance (NMR) experiments were performed on a Varian Unity 600 MHz spectrometer equipped with *z*-shielded gradient coils using 0.1−0.5 mM samples in 20 mM Tris-HCl at pH 8 to which 10% D_2O volume had been added. The experiments were performed at 27 °C.

Experiments To Probe the Functions of FMRP N-Terminal Fragments. Full-length GST-FMRP fusion protein (1 μg) produced in insect cells (*23*) was mixed with 3 μg of His-tagged FMRP(1−214), FMRP(1−180), and FMRP(1−134) produced in *E. coli* and with 20 μL of GST beads (Amersham). The assay was carried out in PBS at pH 7.4 and 0.5% Triton X-100 (binding buffer) at 4 °C overnight. As a negative control, the fragments were also tested with the GST beads alone. The beads were centrifuged at 2000 rpm for 2 min, washed four times with 600 μL of binding buffer, and finally resuspended in 10 μL of Laemmli sample buffer. The binding was probed by a murine monoclonal antibody raised against the FMRP N-terminus.

RNA binding to homopolymers was probed by following the assay described (*31*) with minor modifications. GST was used as a negative control. Five-tenth milliliter of 10 μg solutions of the purified proteins in the binding buffer (10 mM Tris-HCl at pH 7.4, 2.5 mM $MgCl_2$, 0.5% Triton, 150 or 500 mM NaCl) was individually mixed with agarose-bound poly(rA), poly(rU), poly(rG), and poly(rC) beads (purchased from either Sigma or Amersham) (50 μL of fully swollen matrix in the binding buffer). The mixture was left to incubate for 4 h at 4 °C. The beads were pelleted with a short spin in a microfuge and washed four times with 600 μL of binding buffer prior to resuspension in SDS−PAGE loading buffer. Bound proteins were eluted from nucleic acid by boiling and redissolved in an SDS−PAGE. The proteins were transferred to PDVF membrane (Amersham) and visualized by immunoblot. Filters were processed according to instructions from the manufacturer. Either a murine monoclonal antibody raised against the FMRP N-terminus or a monoclonal anti-6-His tag antibody was used for detection. The assay was repeated four times to check for reproducibility. To increase the sensitivity, the RNA-binding assay described here differs from that described in Adinolfi et al. (*30*) in the amount of RNA beads, the protein concentrations, and the incubation time (20 μL of RNA beads, 3 μg of protein concentration, and 20 min, respectively).

RESULTS

FMRP(1−217) Is Not the Minimal Unit Able To Retain a 3D Fold. The FMRP(1−217) construct contains the first seven exons of the FMR1 gene and directly precedes the first KH repeat. A preliminary study by CD of a slightly shorter construct [FMRP(1−214)] had shown that this region has some secondary structure and a high tendency to aggregate (*30*). Different biophysical techniques were used to characterize further the state of aggregation of FMRP(1−217). The molecular mass of FMRP(1−217) as calculated from analytical ultracentrifugation measurements is 34 kDa (at 20 μM protein concentration), which is between that expected for the monomeric (25 kDa) and the dimeric (50 kDa) species. Measurements were carried out at both low (10 mM) and high salt (150 mM) concentrations. A good fit to the data could be obtained only at low salt by assuming a monomer/dimer equilibrium with ca. 60% of the dimeric form (Figure 2a). This is consistent with a dissociation constant in the micromolar range. Studies at higher concentrations with this technique were prevented by the high molar

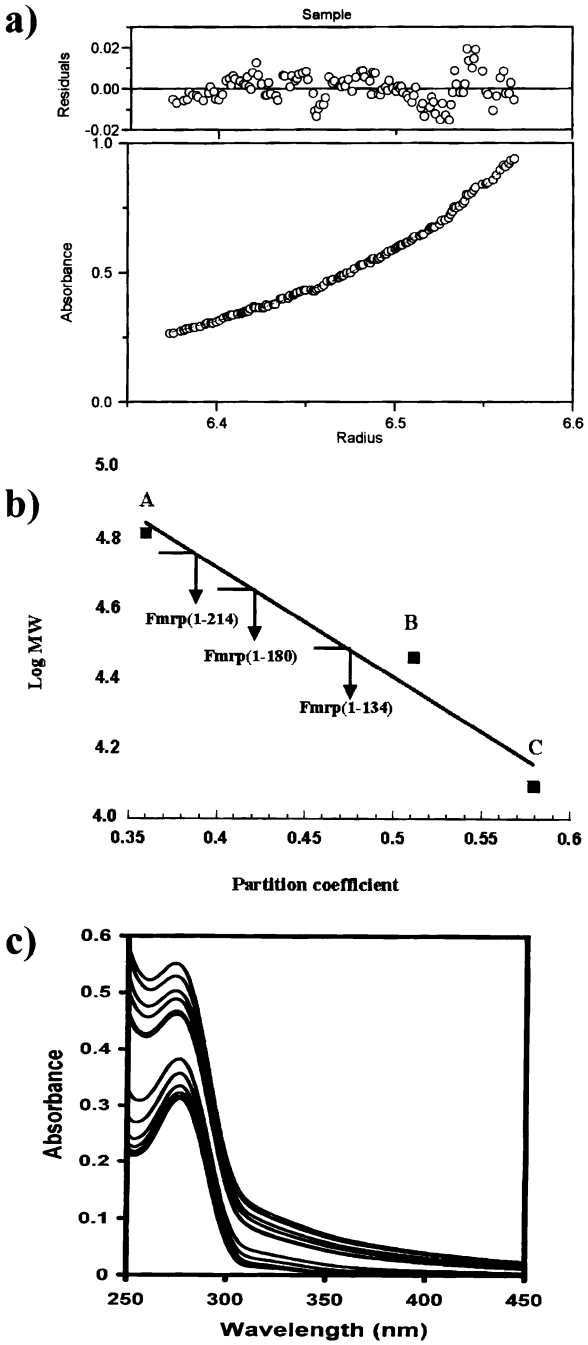

FIGURE 2: (a) Sedimentation equilibrium data of FMRP(1−217) at 16000 rpm and 20 °C. (b) Calibration curve for apparent molecular mass determination of FMRP(1−217), FMRP(1−180), and FMRP(1−134) in native condition by gel filtration chromatography. Bovine serum albumin, 67 kDa (A), carbonic anhydrase, 29 kDa (B), and cytochrome *c*, 12.4 kDa (C), were used as molecular mass standards. (c) UV spectra of 11 μM samples of FMRP(1−217) recorded at increasing temperatures in the range 20−80 °C.

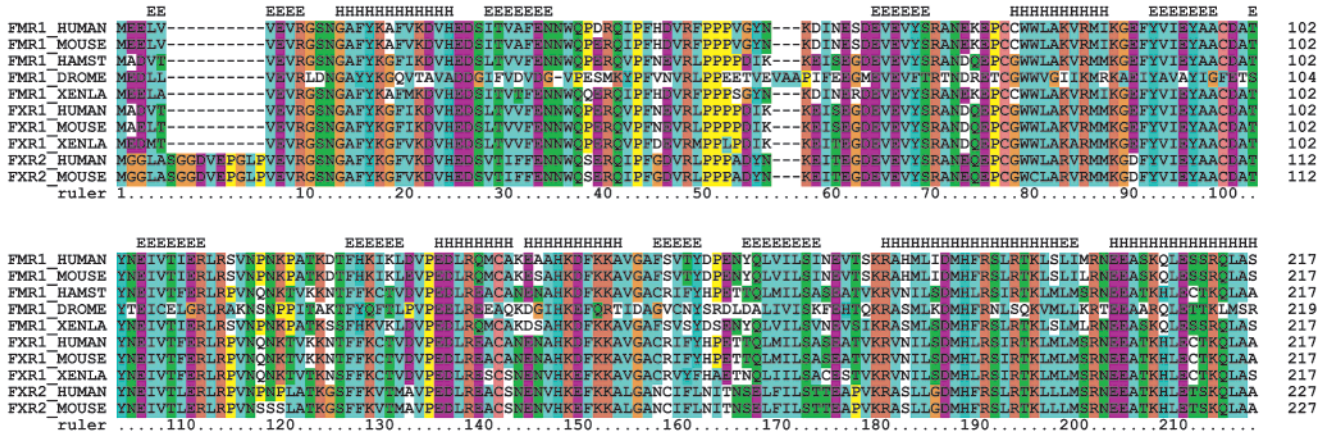

FIGURE 3: Multiple alignments of the N-terminal region of the FMR family, color-coded to highlight similarities (*39*). All G (orange) and P (yellow) residues are colored. Other coloring is by a conserved property in more than 40% (or more than 30% for K, R, and H) of a column according to the following convention: blue, hydrophobic; light blue, partially hydrophobic; red and pink, positive; purple, negative; green, hydrophilic. The sequence numbering (ruler) is displayed according to the FMR1_HUMAN sequence. Secondary structure predictions were made on the multiple aligned FMRP sequences using the PHD web server (40) and displayed in the first line.

extinction coefficient of FMRP(1−217) (27790 cm^{-1} M^{-1}). More qualitative techniques, such as analytical gel filtration and light scattering, had thus to be used.

At 30 μM protein concentration, the apparent molecular mass of FMRP(1−217) as estimated by analytical gel filtration chromatography was 56 kDa, a value compatible with that expected for a dimeric species (Figure 2b). No peak corresponding to the monomer was detected. Light scattering experiments were performed at two different concentrations (70 and 200 μM). When the FMRP(1−217) solutions were passed through a 0.02 μm cutoff filter before each measurement, the filter was readily blocked, preventing the measurements and suggesting the presence of high molecular mass species (data not shown). Similarly, ultraviolet (UV) spectra in the 250−450 nm range did not show a stable baseline, a phenomenon characteristic of polydisperse samples. An increase in light scattering was observed as the temperature was increased, indicating formation of large aggregates. The process was irreversible upon cooling (Figure 2c).

In summary, the N-terminus of FMRP contains secondary structure elements and exists in solution in a dimeric form. However, it has a strong tendency to further aggregate unspecifically at high concentrations.

FMRP(1−134) Forms a Specific Dimer. To find the smallest region able to retain a three-dimensional structure with no or little tendency to nonspecific aggregation, limited proteolysis was carried out on FMRP(1−217) (data not shown). Protease protection was observed for residues 80−160, suggesting the presence of a folded species in this region. A construct spanning residues from 80 to 160 was cloned and expressed but was insoluble and led to the formation of inclusion bodies even when different temperature conditions were tested.

We then explored systematically the properties of two fragments designed using both the exon−intron boundaries and the residue conservation: one spanning region 1−134 [FMRP(1−134)] and the second comprising residues 1−180 [FMRP(1−180)] (Figure 3). Both regions comprise either part or all of the proteolysis-protected region but terminate at the end of predicted secondary structure elements. In particular, a potential helix−loop−helix motif is predicted between amino acids 180−222. If this prediction were correct, this region could be involved in self-association of FMRP(1−217) since helix−loop−helix motifs are often responsible for dimerization (*32*). The two constructs were cloned and expressed, purified, and structurally characterized.

The state of aggregation of FMRP(1−180) and FMRP(1−134) was checked as for FMRP(1−217). Analytical ultracentrifugation measurements of 20 μM samples led to molecular masses of 44 and 33 kDa to be compared with the expected masses of the dimeric forms (45.6 and 31.4 kDa, respectively) (Figure 4a). Good curve fitting was achieved for both constructs by assuming a pure dimeric state, consistent with a dissociation constant in the nanomolar range. Accordingly, the apparent molecular masses of FMRP(1−180) (7.5 μM) and FMRP(1−134) (36 μM) as estimated by gel filtration were 43.9 and 33.4 kDa, respectively. Light scattering and UV spectra of FMRP(1−180) have a behavior similar to those of FMRP(1−217). FMRP(1−134), on the other hand, presented a stable baseline and no blockage of the filter, indicating a monodisperse sample with a calculated mass of 39 kDa (data not shown). These results show that the N-terminal 134 residues of FMRP are sufficient for homodimerization. C-Terminal extension of this region seems to promote nonspecific interactions that lead to aggregation.

To probe whether FMRP(1−134) is able to bind the full-length FMRP, interactions between the bacterially produced FMRP(1−134), FMRP(1−180), and FMRP(1−214) and the GST fusion full-length FMRP (*23*) were tested together in pull-down experiments. The three fragments were all able to interact with FMRP although with affinities decreasing with the length of the construct (Figure 4b).

FMRP(1−134) Folds into a Domain with Significant β Structure Content. The secondary and tertiary structures of the constructs were probed by CD and NMR. The far-UV CD spectra of FMRP(1−180) and FMRP(1−134) are shown in Figure 5a and compared with the spectra of FMRP(1−217). All three constructs have a minimum at 208 nm which is deeper in the two longer constructs. The spectra of FMRP(1−180) and FMRP(1−217) have very similar features, whereas that of FMRP(1−134) has a pronounced maximum at 230 nm, most likely arising from the contribution of aromatic side chains. An estimate of the secondary structure content by curve fitting suggests a minor helical content

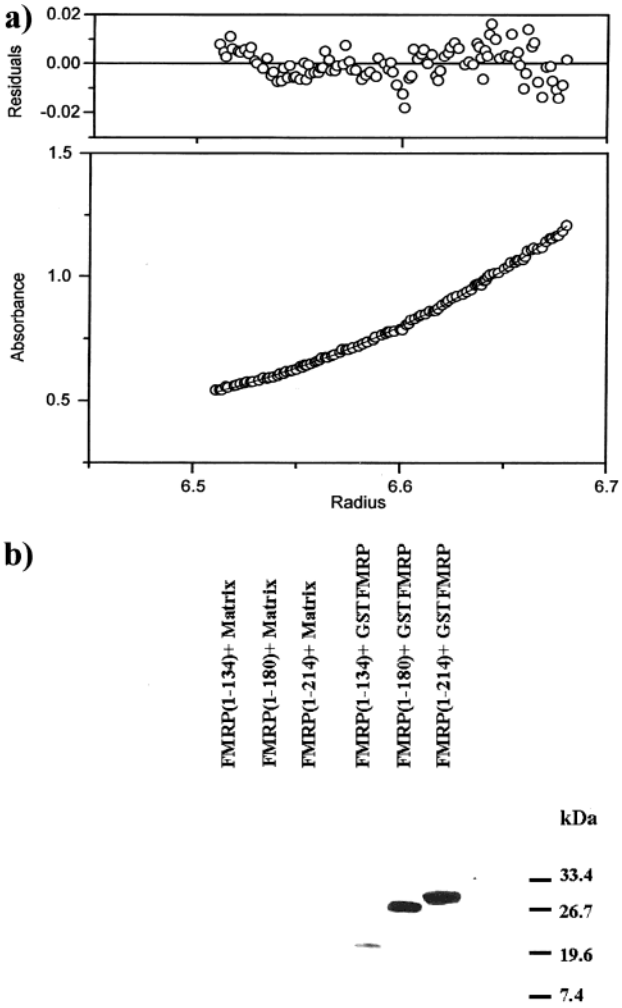

FIGURE 4: (a) Sedimentation equilibrium data of FMRP(1−134) at 16000 rpm and 25 °C. (b) GST pull-down assay to probe interaction of the fragments with full-length FMRP. GST fusion FMRP was mixed with 3 μg of FMRP(1−134), FMRP(1−180), and FMRP(1−214) (lanes 4−6, respectively). As a negative control, 3 μg of the fragments was mixed with glutathione−Sepharose (lanes 1−3). This assay was carried out in phosphate-buffered saline (PBS) (Pharmacia) and 0.5% Triton X-100.

(below or around 5%) with an approximately equal content of β and random structure (40−45%).

A 1D NMR ^{1}H spectrum of FMRP(1−134) was recorded and compared with those of the two longer constructs (Figure 5b). The spectrum of FMRP(1−134) shows a good resonance dispersion typical of a well-defined tertiary fold, as supported by the presence of resonances in the range from −1 to 0.8 ppm: these arise from aliphatic protons in proximity of aromatic rings which, when locked into a defined conformation, generate an additional local magnetic field. The 1D ^{1}H spectra of FMRP(1−180) and FMRP(1−217) are progressively broader, indicating a polydisperse state of these constructs at NMR concentrations (millimolar range). Resonances around 5 ppm in the spectrum of FMRP(1−134) [visible also in the spectrum of FMRP(1−180) but presumably too broad to be visible in FMRP(1−217)] are suggestive of an appreciable β content.

The thermal stability of the three constructs was determined by thermal denaturation followed by monitoring the ellipticity at 222 nm in the CD spectra. Two different concentrations (6 and 100 μM) were used for each construct. The melting point is comparable for FMRP(1−217) and FMRP(1−180) (57.6 ± 0.3 and 57.4 ± 0.2 °C, respectively) and lower for FMRP(1−134) (52.1 ± 0.2 °C). Interestingly, at low concentration, a conformational transition was observed at high temperature, with the signal becoming progressively more negative (Figure 5a). Comparison of the spectra at 25 °C and at 80 °C shows that the shallow minimum observed at 208 nm at room temperature shifts toward 215 nm and becomes deeper. At 80 °C the CD spectrum presents a single well-defined minimum at 215 nm, typical of a β-sheet conformation. When the same experiment was performed at 100 μM, the FMRP(1−217) solution turned opalescent around 40 °C, suggesting aggregate formation, which impaired complete data collection at this concentration. In all cases the process is cooperative but irreversible (Figure 5c). These results are consistent with a conformational transition occurring at high temperatures in which an intermolecular β conformation is stabilized, similar to that observed for β-amyloid formation (*33*). However, a preliminary check of the nature of the precipitate by electron microscopy (EM) showed only amorphous precipitation (data not shown).

Probing the RNA-Binding Properties of the Three Constructs. We have previously demonstrated that the FMRP N-terminus is able to recognize RNA homopolymers (*30*). We tested if also the shorter constructs retain RNA binding. Binding was tested using the four homopolymers immobilized on agarose beads. All constructs are able to bind the homopolymers when 150 mM NaCl concentration was used, although the signal is very weak for some homonucleotides (Figure 6). The constructs seem to have different specificity: FMRP(1−180) and FMRP(1−134) have the highest affinity for poly(rU) and the lowest for poly(rG). These results should be compared with those obtained for FMRP(1−214) [as described by Adinolfi et al. (*30*) and repeated here for consistency] which has clear preference for poly(rG) and poly(rU). At higher salt concentrations (500 mM NaCl), binding was retained, although generally weaker, by most of the constructs (data not shown). No signal was detected when the constructs were tested under the same conditions without the agarose beads and when His-tagged glutathione *S*-transferase (GST), a protein with no RNA-binding properties, was used as a negative control.

DISCUSSION

We have reported here a structural and functional characterization of the FMRP N-terminus, a region strongly conserved through evolution and known to be involved in several molecular interactions (*16*). We show that region 1−134 (which covers the first five exons) forms an independently folded unit, which we named NDF, for N-terminal domain of FMRP. While we were in the process of submitting the present paper, Ponting and co-workers reported detectable although low sequence homology between regions 3−43 and 62−108 of FMRP and the Tudor domain family, which also comprises the Chromo and MBT domains as well as a new plant-specific Agenet domain (*34*). This similarity would then place FMRP into a larger family of proteins that has evolved from a common ancestor with methyl substrate binding functions. The prediction is fully supported by our data since the expected fold for these regions would be predominantly β, as we observe by CD.

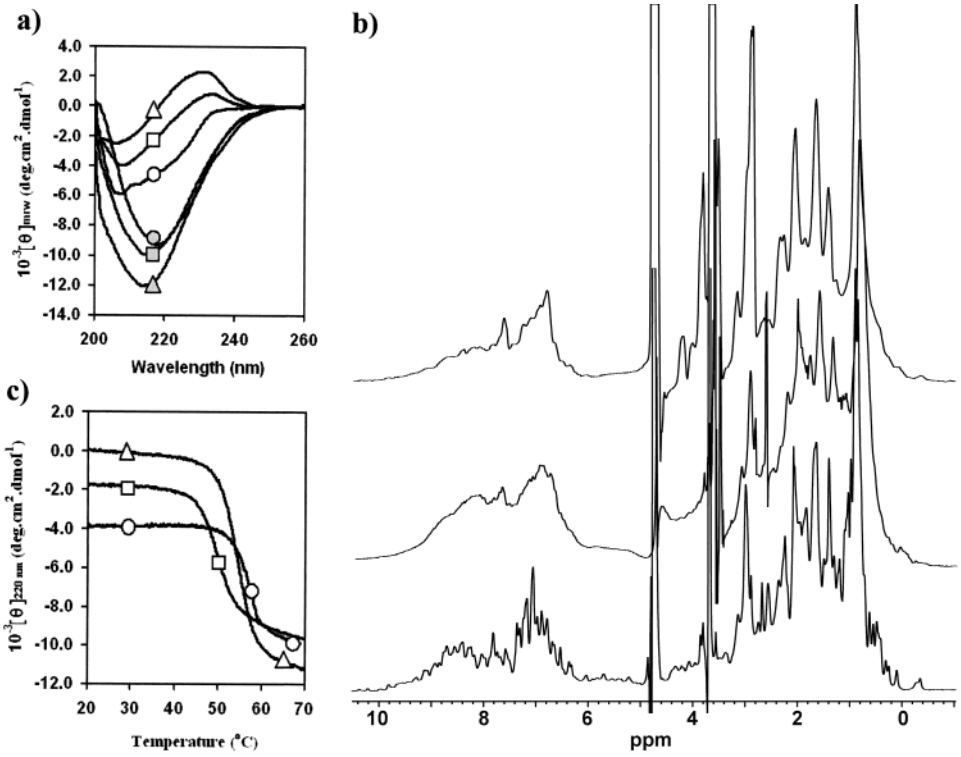

FIGURE 5: (a) Comparison of the far-UV CD spectra of the three constructs at 25 °C (open symbols) and 80 °C (gray symbols). The concentrations used to record these spectra are 100 μM. Triangles, squares, and circles are used for FMRP(1−134), FMRP(1−180), and FMRP(1−217), respectively. (b) Comparison of the 1D NMR spectra of FMRP(1−217) (top), FMRP(1−180) (middle), and FMRP(1−134) (bottom). The spectra were recorded at 25 °C and 600 MHz on samples around 0.2 mM concentration. The sharp peaks at 3.5 ppm are from the buffer. The peak at 4.7 ppm is residual water. (c) Thermal denaturation curves recorded for the three samples following the signal at 222 nm. The same convention as in panel a is used.

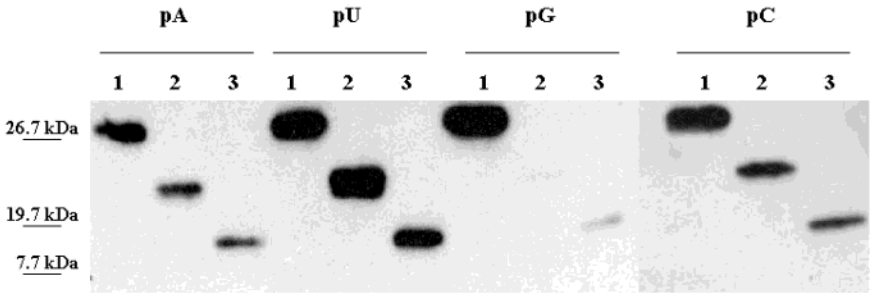

FIGURE 6: RNA-binding assays using homopolymers p(rA), p(rU), p(rG), and p(rC) immobilized on agarose beads. The constructs are reported as follows: FMRP(1−214) (lane 1), FMRP(1−180) (lane 2), and FMRP(1−134) (lane 3) performed at pH 7.4 and 150 mM NaCl as described in Materials and Methods.

Definition of a new structural domain in FMRP, in addition to the two KH motifs, is very important to revisit and interpret the functional data available. NDF is able to retain several of the functions mapped on the FMRP N-terminus. FMRP(1−134) recognizes RNA as it does the longer FMRP-(1−214) (*30*). Consistently, conserved positively charged residues which could contribute to RNA binding are present along the whole length of the N-terminus (Figure 3). However, the affinities to homopolymers progressively decrease with the length of the constructs, and the three proteins have different sequence specificities for the homopolymers. These findings suggest that the contribution of different binding sites distributed along the N-terminus is likely to be cooperative.

Many of the FMRP interacting proteins have been identified by two-hybrid screening using the protein N-terminus [construct FMRP(1−218)] as a bait. Further mapping was obtained by GST pull-down using deletion mutants designed, in the absence of structural indications, according to the exon boundaries. These do not, however, usually correlate with the domain definition especially in intracellular proteins, thus leading to the possibility that a negative result could be caused not by the absence of a region directly involved in the interaction but to loss of a region structurally essential. To demonstrate the power of a structural approach, a construct spanning NDF has been recently used to map the interacting region with NUFIP1 and the newly identified 82-FIP (35): NDF was shown to be sufficient to bind.

The folded NDF is also able to form stable and specific dimers and to bind, although weakly, full-length FMRP. A dimerization motif in FMRP(1−134) is somewhat unexpected since previous reports had excluded participation of the region N-terminal to exon 7 into dimerization and recognition of the close homologues FXR1 and FXR2 (*26*, *27*). However, the discrepancy can be explained by taking into account that the assays presented here were carried on

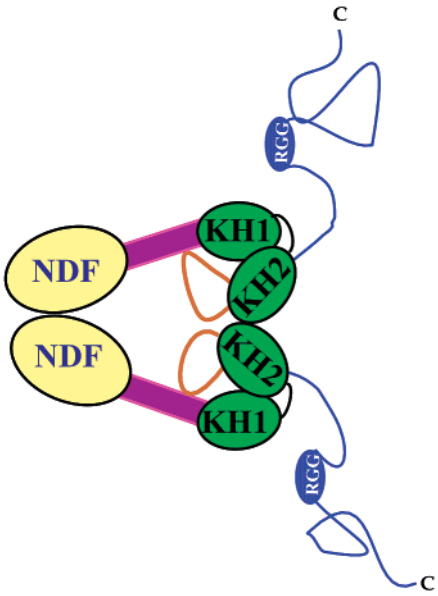

FIGURE 7: Hypothetical two-dimensional model of the architecture of the FMRP dimer. Ovals are used to indicate the so far identified globular domains (NDF, KH1, and KH2). The insertion between the second and third β strand observed in KH2 (*29*) is shown in orange. The tertiary fold of FMRP could bring the insertion to pack against the region 135−217 (violet bar), which links NDF to KH1. The nonconserved unstructured C-terminus with the RGG motif is indicated in blue. When available, we tried to take into account the information about the mutual position of the N- and C-termini in each module.

highly purified untagged proteins excluding the possibility of interferences. Also, the relatively low affinity by which FMRP(1−134) binds with the full-length FMRP makes it possible to miss this interaction site with methods less sensitive than the ones used in this study.

It remains possible that, together with NDF, other dimerization sites can cooperatively determine the dimer stability of the full-length protein. A putative coiled coil, identified in exon 7, has been proposed to be the main dimerization motif (*27*). Also, in our hands, region 181−217 has a strong tendency to form aggregates of hydrophobic nature. However, a closer analysis of the exon 7 sequence shows that this region has a clear helical periodicity but does not contain heptamotifs that could stabilize a coiled coil (*36*). Secondary structure predictions based on a multiple alignment would rather suggest that this region folds as a helix−loop−helix motif (Figure 3). Dimerization could occur through pairing two such motifs, leading to formation of a four-helix bundle. In the full-length protein, this region could pack either against other proteins or against distal regions of FMRP, thus masking additional exposed hydrophobic patches that cause further aggregation in our FMRP(1−217) construct. A plausible partner could be KH2, another region suggested to be implicated in dimerization. A mutant carrying an FMRP(I304N) point mutation in KH2 was shown to be unable to self-assemble and, consequently, to bind to polysomes and inhibit translation (*10*, *37*). However, because of the important structural role that the semiburied Ile304 plays in the fold of the KH domain (*38*), it is unlikely that this residue is directly involved in dimerization. The mutation should lead to disruption of the KH2 fold, thus altering the shape of the full-length protein and preventing dimerization. It is tempting to speculate that the tertiary structure of the full-length protein brings close together exon 7 and KH2, which could pack against each other (Figure 7). This hypothesis could explain the difficulties encountered in producing isolated recombinant KH2 in a folded form (*30*).

In conclusion, our results show that the region preceding the first KH domain contains a newly identified domain able to bind RNA and to function as a protein−protein interaction motif. Identification of another autonomously three-dimensional folding unit in FMRP that is able to retain some of the functions of the full-length protein opens new possibilities to investigating at atomic details the way FMRP recognize specific RNA targets and to understanding the essential role played by the N-terminus in molecular recognition. Further characterization of the structure of the NDF is currently in progress in our laboratory.

ACKNOWLEDGMENT

We are indebted to Steve Smerdon for help with the light scattering experiments and to David J. Thomas for critical reading of the manuscript.

REFERENCES

1. Crawford, D. C., Acuna, J. M., and Sherman, S. L. (2001) FMR1 and fragile X syndrome: Human genome epidemiology review, *Gen. Med. 3*, 359−371.
2. O'Donnell, W. T., and Warren, S. T. (2002) A decade of molecular studies of fragile X syndrome, *Annu. Rev. Neurosci. 25*, 315−338.
3. Abitbol, M., Menini, C., Delezoide, A. L., Rhyner, T., Vekemans, M., and Mallet, J. (1993) Nucleus basalis magnocellularis and hippocampus are the major sites of FMR1 expression in the human fetal brain, *Nat. Genet. 4*, 147−152.
4. Bachner, D., Steinbach, P., Wohrle, D., Just, W., Vogel, W., Hameister, A., Manca, A., and Poustka, A. (1993) Enhanced FMR1 expression in testis, *Nat. Genet. 4*, 115−116.
5. Devys, D., Lutz, Y., Rouyer, N., Bellocq, J. P., and Mandel, J. L. (1993) The FMR1 protein is cytoplasmic, most abundant in neuron and appears normal in carriers of a fragile X premutation, *Nat. Genet. 4*, 335−340.
6. Hinds, H. L., Ashley, C. T., Sutcliffe, J. S., Nelson, D. L., Warren, S. T., Housman, D. E., and Schalling, M. (1993) Tissue-specific expression of FMR1 provides evidence for a functional role in fragile X syndrome, *Nat. Genet. 3*, 36−43.
7. Weiler, I. J., Irwin, S. A., Klintsova, A. Y., Spenser, C. M., Brazelton, A. D., Miyashiro, K., Comery, T. A., Patel, B., Eberwine, J., and Greenough, W. T. (1997) Fragile X mental retardation protein is translated near synapses in response to neurotransmitter activation, *Proc. Natl. Acad. Sci. U.S.A. 94*, 5395−5400.
8. De Diego Otero, Y., Severijnen, L. A., van Cappellen, G., Schrier, M., Oostra, B., and Willemsen, R. (2002) Transport of fragile X mental retardation protein via granules in neurites of PC12 cells, *Mol. Cell. Biol. 22*, 8332−8341.
9. Brown, V., Small, K., Lakkis, L., Feng, Y., Gunter, C., Wilkinson, K. D., and Warren, S. T. (1998) Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome, *J. Biol. Chem. 273*, 15521−15527.
10. Laggerbauer, B., Ostareck D., Keidel, E. M., Ostareck-Lederer, A., and Fischer, U. (2001) Evidence that fragile X mental retardation protein is a negative regulator of translation, *Hum. Mol. Genet. 10*, 329−338.
11. Zalfa, F., Giorgi, M., Primerano, B., Moro, A., Di Penta, A., Reis, S., Oostra, B., and Bagni, C. (2003) The Fragile X Syndrome Protein FMRP associates with BC1 RNA and regulates the translation of specific mRNAs at synapses, *Cell 112*, 317−327.
12. Mazroui, R., Huot, M. E., Tremblay, S., Filion, C., Labelle, Y., and Khandjian, E. W. (2002) Trapping of messenger RNA by Fragile X Mental Retardation protein into cytoplasmic granules induces translation repression, *Hum. Mol. Genet. 11*, 3007−3017.
13. Ceman, S., Brown, V., and Warren, S. T. (1999) Isolation of an FMRP-associated messenger ribonucleoprotein particle and iden-

tification of nucleolin and the fragile X-related proteins as components of the complex, *Mol. Cell. Biol. 19*, 7925−7932.
14. Ceman, S., Nelson, R., and Warren, S. T. (2000) Identification of mouse YB1/p50 as a component of the FMRP-associated mRNP particle, *Biochem. Biophys. Res. Commun. 279*, 904−908.
15. Ohashi, S., Koike, K., Omori, A., Ichinose, S., Ohara, S., Kobayashi, S., Sato, T. A., and Anzai, K. (2002) Identification of mRNA/protein (mRNP) complexes containing Puralpha, mStaufen, fragile X protein, and myosin Va and their association with rough endoplasmic reticulum equipped with a kinesin motor, *J. Biol. Chem. 277*, 37804−37810.
16. Bardoni, B., and Mandel, J. L. (2002) Advances in understanding of fragile X pathogenesis and FMRP function, and in identification of X linked mental retardation genes, *Curr. Opin. Genet. Dev. 12*, 284−293.
17. Brown, V., Jin, P., Ceman, S., Darnell, J. C., O'Donnell, W. T., Tenenbaum, S. A., Jin, X., Feng, Y., Wilkinson, K. D., Keene, J. D., Darnell, R. B., and Warren, S. T. (2001) Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome, *Cell 107*, 477−487.
18. Schaeffer, C., Bardoni, B., Mandel, J. L., Ehresmann, B., Ehresmann, C., and Moine, H. (2001) The fragile X mental retardation protein binds specifically to its mRNA via a purine quartet motif, *EMBO J. 20*, 4803−4813.
19. Darnell, J. C., Jensen, K. B., Jin, P., Brown, V., Warren, S. T., and Darnell, R. B. (2001) Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function, *Cell 107*, 489−499.
20. Miyashiro, K. Y., Beckel-Mitchener, A., Purk, T. P., Becker, K. G., Barret, T., Liu, L., Carbonetto, S., Weiler, I. J., Greenough, W. T., and Eberwine, J. (2003) RNA cargoes associating with FMRP reveal deficits in cellular functioning in Fmr1 null mice, *Neuron 37*, 417−431.
21. Corbin, F., Bouillon, M., Fortin, A., Morin, S., Rousseau, F., and Khandjian, E. W. (1997) The fragile X mental retardation protein is associated with poly(rA)+ mRNA in actively translating polyribosomes, *Hum. Mol. Genet. 6*, 1465−1472.
22. Feng, Y., Gutekunst, C. A., Eberhart, D. E., Yi, H., Warren, S. T., and Hersch, S. M. (1997) Fragile X mental retardation protein: nulcear cytoplasmic shuttling and association with somatodendritic ribosomes, *J. Neurosci. 17*, 1539−1547.
23. Bardoni, B., Schenck, A., and Mandel, J. L. (1999) A novel RNA-binding nuclear protein that interacts with the fragile X mental retardation (FMR1) protein, *Hum. Mol. Genet. 8*, 2557−2566.
24. Schenck, A., Bardoni, B., Moro, A., Bagni, C., and Mandel, J. L. (2001) A highly conserved protein family interacting with the fragile X mental retardation protein (FMRP) and displaying selective interactions with FMRP-related proteins FXR1P and FXR2P, *Proc. Natl. Acad. Sci. U.S.A. 98*, 8844−8849.
25. Siomi, M. C., Siomi, H., Sauer, W. H., Srinivasan, S., Nussbaum, R. L., and Dreyfuss, G. (1995) FXR1, an autosomal homologue of the fragile X mental retardation gene, *EMBO J. 14*, 2401−2408.
26. Zhang, Y., O'Connors, J. P., Siomi, M. C., Srinivasan, S., Dutra, A., Nussbaum, R. L., and Dreyfuss, R. (1995) The fragile X mental retardation syndrome protein interacts with novel homologues FXR1 and FXR2, *EMBO J. 14*, 5358−5366.
27. Siomi, M. C., Zhang, Y., Siomi, H., and Dreyfuss, G. (1996) Specific sequences in the fragile X syndrome protein FMR1 and the FXR proteins mediate their binding to 60S ribosomal subunits and the interactions among them, *Mol. Cell. Biol. 16*, 3825−3832.
28. Tamanini, F., Van Unen, L., Bakker, C., Sacchi, N., Galjaard, H., Oostra, B., and Hoogeveen, A. T. (1999) Oligomerization properties of fragile-X mental-retardation protein (FMRP) and the fragile X related proteins FXR1P and FXR2P, *Biochem. J. 343*, 517−523.
29. Musco, G., Kharrat, A., Stier, S., Fraternali, F., Gibson, T. J., Nilges, M., and Pastore, A. (1997) The solution structure of the first KH domain of FMR1, the protein responsible for fragile X syndrome, *Nat. Struct. Biol. 4*, 712−716.
30. Adinolfi, S., Bagni, C., Musco, G., Gibson, T., Mazzarella, L., and Pastore, A. (1999) Dissecting FMR1, the protein responsible for fragile X syndrome, in its structural and functional domains, *RNA 5*, 1248−1258.
31. Swanson, M. S., and Dreyfuss, G. (1999) Preparation of heterogeneous nuclear ribonucleoprotein complexes, *Methods Mol. Biol. 118*, 299−308.
32. Robinson, K. A., and Lopes, J. M. (2000) Survey and Summary: *Saccharomyces cerevisiae* basic helix-loop-helix proteins regulate diverse biological processes, *Nucleic Acids Res. 28*, 1499−1505.
33. Dobson, C. M. (1999) Protein misfolding, evolution and disease, *Trends Biochem. Sci. 24*, 329−332.
34. Maurer-Stroh, S., Dickens, N. J., Hughes-Davies, L., Kouzarides, T., Eisenhaber, F., Ponting, C. P. (2003) The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains, *Trends Biochem. Sci. 28*, 69−74.
35. Bardoni, B., Castets, M., Huot, M. H., Schenck, A., Adinolfi, S., Corbin, F., Pastore, A., Khandjian, E. W., and Mandel, J. L. (2003) 82-FIP a novel FMRP (Fragile X Mental retardation Protein) interacting protein shows a cell cycle-dependent subcellar localization, *Hum. Mol. Genet*. (in press).
36. Yu, Y. (2002) Coiled-coils: stability, specificity, and drug delivery potential, *Adv. Drug Delivery Rev. 54*, 1113−1129.
37. Feng, Y., Absher, D., Eberhart, D. E., Brown, V., Malter, H. E., and Warren, S. T. (1997) FMRP associates with polyribosomes as an mRNP, and the I304N mutation of severe fragile X syndrome abolishes this association, *Mol. Cell 1*, 109−118.
38. Ramos, A., Hollingworth, D., and Pastore, A. (2003) The role of a clinically important mutation in the stability and RNA-binding properties of KH motifs, *RNA 9*, 293−298.
39. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin F., and Higgins, D. G. (1997) The CLUSTALX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, *Nucleic Acids Res. 25*, 4876−4882.
40. Rost, B., Sander, C., and Schneider, R. (1994) PHD-an automatic mail server for protein secondary structure prediction, *Comput. Appl. Biosci. 10*, 53−60.

BI034909G