Published online August 9, 2004

Nucleic Acids Research, 2004, Vol. 32, No. 14 **4145–4154**
doi:10.1093/nar/gkh745

Destabilization of tetraplex structures of the fragile X repeat sequence $(CGG)_n$ is mediated by homolog-conserved domains in three members of the hnRNP family

Samer Khateb, Pnina Weisman-Shomer, Inbal Hershco, Lawrence A. Loeb[1] and Michael Fry*

Unit of Biochemistry, Rappaport Faculty of Medicine, Technion—Israel Institute of Technology, P.O. Box 9649, Haifa 31096, Israel and [1]Department of Pathology, University of Washington School of Medicine, Box 357705, Seattle, WA 98195-7705, USA

Received June 16, 2004; Revised and Accepted July 17, 2004

ABSTRACT

Hairpin or tetrahelical structures formed by a $d(CGG)_n$ sequence in the *FMR1* gene are thought to promote expansion of the repeat tract. Subsequent to this expansion *FMR1* is silenced and fragile X syndrome ensues. The injurious effects of $d(CGG)_n$ secondary structures may potentially be countered by agents that act to decrease their stability. We showed previously that the hnRNP-related protein CBF-A destabilized G′2 bimolecular tetraplex structures of $d(CGG)_n$. Analysis of mutant proteins revealed that the CBF-A-conserved domains $RNP1_1$ and ATP/GTP binding box were sufficient and necessary for G′2 $d(CGG)_n$ disruption while the $RNP2_1$ motif inhibited the destabilization activity. Here, we report that a C-terminal fragment of CBF-A whose only remaining conserved domain was the ATP/GTP binding motif, disrupted G′2 $d(CGG)_n$ more selectively than wild-type CBF-A. Further, two additional members of the hnRNP family, hnRNP A2 and mutant hnRNP A1 effectively destabilized G′2 $d(CGG)_n$. Examination of mutant hnRNP A2 proteins revealed that, similar to CBF-A, their $RNP1_1$ element and ATP/GTP binding motif mediated G′2 $d(CGG)_n$ disruption, while the $RNP2_1$ element blocked their action. Similarly, the $RNP1_1$ and $RNP2_1$ domains of hnRNP A1 were, respectively, positive and negative mediators of G′2 $d(CGG)_n$ destabilization. Last, employing the same conserved motifs that mediated disruption of the DNA tetraplex G′2 $d(CGG)_n$, hnRNP A2 destabilized $r(CGG)_n$ RNA tetraplex.

INTRODUCTION

Strands of DNA or RNA that contain clusters of adjacent guanine residues are capable of forming tetrahelical structures termed quadruplexes or tetraplexes. At the core of these tetrahelices are cation-coordinated, Hoogsteen hydrogen-bonded tetrads of guanine residues [for reviews see (1,2)]. Three major types of tetraplex structures are distinguished by the stoichiometry and orientation of their nucleic acid strands; anti parallel G′4 monomolecular or G′2 bimolecular tetraplexes and parallel-stranded G4 four-molecular tetraplexes. These tetrahelices are further classified according to parameters such as the nucleotide sequence and size of their non-guanine spacer tracts, inclusion of bases other than guanine in the tetrad structure, the overall molecular geometry of the quadruplex, the glycoside torsion angles and the type of coordinating cation (1,2).

Numerous guanine-rich sequences of biological significance form tetraplex structures *in vitro* under physiologic-like conditions. Although the existence of such structures *in vivo* is still awaiting direct demonstration, it was argued that tetrahelices formed in genomic DNA undertake diverse biological tasks. For instance, folding of the telomere G-strand repeat sequence into monomolecular tetraplex structure was implicated in the regulation of the extension (3) or protection (4) of telomeric DNA. Also, formation of inter-chromosomal guanine tetrads was proposed to function in the pairing of meiotic chromosome (5). Further, tetraplex structures generated by runs of guanine residues in regulatory sequences were implicated in the transcriptional control of genes such as *c-myc* (6–8) or insulin (9–11).

A trinucleotide $d(CGG)_n$ repeat sequence at the transcribed but untranslated 5′ end of the *FMR1* gene readily forms hairpin (12–15) and tetraplex (16–18) structures. Blocking of DNA replication by these secondary structures (18–20) may instigate polymerase slippage which results in the expansion of this trinucleotide repeat. Stable hairpin or tetraplex structures of the expanded $d(CGG)_n$ sequence may in turn contribute to the obstruction of *FMR1* transcription and to the resulting absence of its protein product, FMRP, which is the direct cause of fragile X syndrome. In this context, agents that destabilize tetrahelical formations of $d(CGG)_n$ may eradicate the block to DNA replication and most importantly, may aid in restoring

*To whom correspondence should be addressed. Tel: +972 4 829 5328; Fax: +972 4 851 0735; Email: mickey@tx.technion.ac.il

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

Nucleic Acids Research, Vol. 32 No. 14 © Oxford University Press 2004; all rights reserved

FMR1 expression in fragile X cells by removing the impediment to *FMR1* transcription. Several proteins were found to unwind or destabilize tetraplex DNA. Yeast (21) and human (19,22–24) members of the RecQ family of DNA helicases efficiently unwind tetraplex structures of guanine-rich DNA sequences in reactions that require ATP hydrolysis. The RecQ protein Werner syndrome helicase/exonuclease (WRN) was specifically shown to unwind G′2 bimolecular tetraplex structures of $d(CGG)_n$ (19,23).

A heterogeneous nuclear ribonucleoprotein (hnRNP)-related protein, the CArG-box binding protein A (CBF-A, formerly also named qTBP42), destabilized G′2 $d(CGG)_n$ in an ATP-independent manner while it paradoxically stabilized G′2 tetraplexes of the d(TTAGGG) telomeric repeat sequence (25,26). CBF-A, which was originally identified as a muscle-specific transcriptional repressor (27), was more recently reported to be involved in the transcriptional and post-transcriptional regulation of the expression of several genes of diverse tissues (28–31). CBF-A possesses conserved protein motifs that hallmark members of the hnRNP family: $RNP2_1$ and $RNP1_1$ boxes that combine to form an RNA recognition motif 1 (RRM1), $RNP2_2$ and $RNP1_2$ motifs that pair into an RRM2 domain, and an ATP/GTP binding fold (27). Mutating each or a combination of these five conserved domains revealed that either the $RNP1_1$ motif or the ATP/GTP binding box was necessary and sufficient for the disruption of G′2 $d(CGG)_n$ whereas the $RNP2_1$ domain suppressed the tetraplex destabilization activity (26).

In the present paper, we report that additional members of the hnRNP family, mutant hnRNP A1 and hnRNP A2, were capable of destabilizing G′2 $d(CGG)_n$ and that the same homologous conserved domains acted in these proteins to mediate the destabilization activity. We also show that whereas CBF-A and hnRNP A2 disrupted a DNA tetraplex of the d(CGG) repeat sequence at similar efficiency, an RNA tetraplex $r(CGG)_n$ RNA was effectively destabilized by hnRNP A2 but not by CBF-A.

MATERIALS AND METHODS

Preparation of single- and double-stranded DNA and of DNA and RNA tetraplex structures

Synthetic DNA oligomers listed in Table 1 were obtained from Operon Technologies or Genosys and purified by denaturing gel electrophoresis in 8 M urea, 14% polyacrylamide (acrylamide/bisacrylamide, 19:1) (23). The RNA oligomer 3′-tail $r(CGG)_7$ (Table 1), which was a generous gift of the research department of Integrated DNA Technology, was purified by RNase-free high-performance liquid chromatography (HPLC). The DNA or RNA oligomers were 5′ end labeled with ^{32}P in bacteriophage T4 polynucleotide kinase-catalyzed reaction (32). Single-stranded oligomers were boiled for 3 min prior to their use. Double-stranded DNA was formed by the annealing of equimolar amounts of complementary DNA oligomers as we described (25). Bimolecular G′2 tetraplex structures of end-labeled 3′-tail $d(CGG)_7$ or 3′-tail $r(CGG)_7$ were generated and their stoichiometry was verified as detailed previously (33). Typically, being in equilibrium with their respective single strands, the tetraplex forms of the DNA or RNA repeat sequence constituted 50–65 or 60–85%, respectively, of the total amount of nucleic acid. A parallel-stranded, four-molecular G4 tetraplex structure of the IgG switch region sequence was prepared and characterized as we described in (25) and a G4 tetraplex structure of the 5′ E-box oligomer was formed according to Walsh and Gualberto (34).

Table 1. DNA and RNA oligonucleotides used in this study

Oligomer	Nucleotide sequence
3′-tail $d(CGG)_7$	5′-d(CGGCGGCGGCGGCGGCGGCGGCGTGGA-CTC)-3′
3′-tail $r(CGG)_7$	5′-r(CGGCGGCGGCGGCGGCGGCGGCGUGGA-CUC)-3′
5′- E-box	5′-d(TCAGGCAGCAGGTGTTGGGGGATCGA)-3′
TeR4	5′-d(TTAGGGTTAGGGTTAGGGTTAGGG)-3′
Anti-TeR4	5′-d(CCCTAACCCTAACCCTAACCCTAA)-3′
CArG box	5′-d(CTTTTACCTAATTAGGAAATGG)-3′
Anti-CArG box	5′-d(CCATTTCCTAATTAGGTAAAAG)-3′
IgG switch sequence	5′-d(TACAGGGGAGCTGGGGTAGA)-3′

Plasmids

A pGEX-A1 plasmid harboring mouse hnRNP A1 cDNA was kindly contributed by Dr Benoit Chabot (Universitè de Sherbrooke, Canada). A pGEX-2T construct containing human hnRNP A2 cDNA was the generous gift of Dr Ralph C. Nichols (Dartmouth School of Medicine, NH). Mouse CBF-A cDNA that was donated by Dr T. Miwa (Osaka University, Japan), was subcloned into pGEX-2T as we described in (26).

Generation of deletion, substitution or truncation mutations and expression of recombinant proteins

Previously described procedures (26) were employed to introduce deletion or substitution mutations into cDNA of hnRNP A1 or hnRNP A2 and to truncate CBF-A cDNA to form a C-terminal fragment. Wild-type or mutant plasmid DNA was electroporated (Eppendorf electroporator 2510) into *Escherichia coli XL-1-Blue* cells, the plasmid DNA was purified and the presence of a desired mutation was validated by direct nucleotide sequencing. Glutathione *S*-transferase (GST)-fused wild-type or mutant proteins were expressed and purified by slightly modified standard procedures as we described in (26), and ~95% purity of each recombinant protein was verified by SDS–PAGE. Recombinant hnRNP A2 or CBF-A proteins exhibited tetraplex DNA destabilization activity only after their GST tag was cleaved by thrombin protease digestion (26). In contrast, recombinant hnRNP-disrupted A1 tetraplex $d(CGG)_n$ at a similar efficiency with its GST tag fused or cleaved.

Assays of DNA or RNA destabilization and of DNA binding

Destabilization of 5′ ^{32}P-labeled double-stranded DNA, G4 tetraplex DNA or G′2 tetraplex DNA or RNA, was conducted at 33, 37 or 40°C for 10 or 15 min as specified, in reaction mixtures that contained in a final volume of 10 µl, specified amounts of protein and 150–300 fmol DNA or RNA substrate in buffer D (10 mM KCl, 0.5 mM DTT, 1 mM EDTA, 20% glycerol in 25 mM Tris–HCl buffer, pH 8.0). All the reagents that were used for the destabilization of tetraplex RNA were

prepared in RNase-free water and the assays were conducted in RNase-free reaction tubes. The tetraplex disruption reactions were terminated by rapid cooling of the mixtures to 4°C and addition of 3% SDS to a final concentration of 0.5%. Single-stranded products of the destabilization reaction were resolved from the intact tetraplex DNA or RNA substrates by electrophoresis at 4°C and 200–250 V in non-denaturing 10% polyacrylamide gel in 10 mM KCl, 0.5× TBE buffer (1.2 mM EDTA in 0.54 Tris–borate buffer, pH 8.3) until a bromophenol blue tracking dye migrated 7.0–7.5 cm into the gel. Gels were dried and the fractions of single-stranded DNA or RNA products, and of the remaining unwound tetraplex substrates were quantified by phosphorimaging analysis.

Binding of purified recombinant wild-type or mini CBF-A proteins to single-, double- or four-stranded DNA was conducted at 4°C for 20 min in reaction mixtures that contained in final volume of 10 µl, specified amounts of protein and 150–300 fmol [5′-^{32}P]DNA in buffer D. Protein–DNA complexes were resolved from free DNA by electrophoresis of the reaction mixtures at 4°C and 200–250 V in non-denaturing 10% polyacrylamide gel in 10 mM KCl, 0.5× TBE buffer. Free and protein-bound DNA was quantified by phosphorimaging analysis of the dried gels.

RESULTS

We have previously identified conserved domains in CBF-A that govern its ability to destabilize tetraplex d(CGG)$_n$. Our principal observations, as modeled in Figure 1, suggested that the suppressor RNP2$_1$ domain in wild-type CBF-A blocked either the RNP1$_1$ motif or the ATP/GTP binding fold (not represented in Figure 1). In the presence of the suppressor RNP2$_1$ domain, the one uninhibited motif remained free to mediate disruption of G′2 d(CGG)$_n$. Inactivation of either the RNP1$_1$ motif or the ATP/GTP binding box and the blocking of the remaining active domain by the RNP2$_1$ element, rendered the protein incapable of tetraplex d(CGG)$_n$ disruption (Figure 1). The presence of two active sites was reinforced by the finding that once the suppressor RNP2$_1$ domain was inactivated, destabilization activity was restored to mutant proteins that possessed either the RNP1$_1$ motif or the ATP/GTP binding box as their single active element (26).

In this study, we inquired whether members of the hnRNP family other than CBF-A were also capable of destabilizing DNA or RNA tetraplex structures of the trinucleotide repeat (CGG)$_n$. Consequently, we asked whether similar or different, conserved homologous domains of different hnRNPs mediate tetraplex disruption. To this end, we compared the relative tetrahelical d(CGG)$_n$ destabilizing capacity of wild-type and mutant CBF-A, hnRNP A1 and hnRNP A2 proteins. As seen in Figure 2, although these three proteins are of different length and they contain significant non-homologous stretches, the amino acid sequences of their four RNP boxes are highly preserved. Being somewhat less conserved, the ATP/GTP binding fold spans eight residues in hnRNP A1 and A2 but is only six residues long in CBF-A. In addition, the ATP/GTP binding box of hnRNP A2 differs by two or one residues from those of the hnRNP A1 or CBF-A proteins, respectively (Figure 2).

CBF-A and hnRNP A2 employ homolog-conserved domains to mediate destabilization of tetraplex d(CGG)$_n$

To examine whether recombinant hnRNP A2 possessed tetrahelical d(CGG)$_n$ destabilizing activity, it was incubated with terminally labeled bimolecular G′2 tetraplex form of 3′-tail

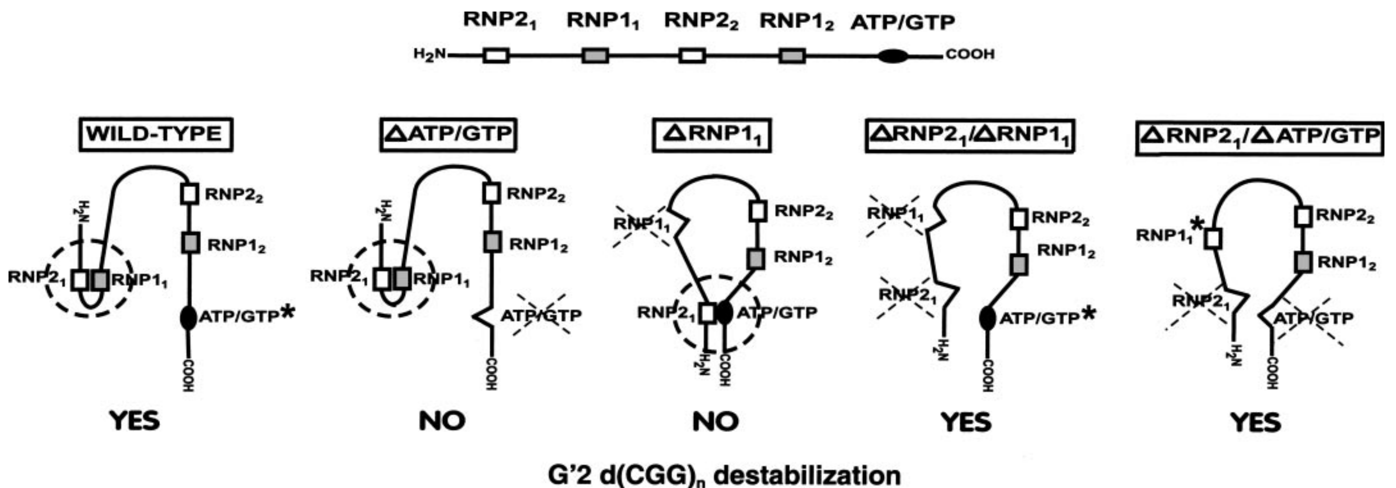

Figure 1. Modeling of the involvement of conserved motifs of CBF-A in the destabilization of tetraplex d(CGG)$_n$. The presented model is based on our previously published results (26). The five conserved motifs of wild-type CBF-A are schematically illustrated at the top of the diagram. Folded forms of wild-type and of five selected mutant proteins with their conserved motifs marked, are in the center and at the bottom are their previously determined respective capacities to destabilize G′2 bimolecular tetraplex d(CGG)$_n$ (26). Either the RNP1$_1$ element or the ATP/GTP binding domain that mediate tetraplex d(CGG)$_n$ disruption is blocked by the RNP2$_1$ motif. Coupled inhibitory RNP2$_1$ box and a suppressed RNP1$_1$ motif are marked within a dashed circle. This leaves an uninhibited ATP/GTP binding box, (denoted by an asterisk) to carry out tetraplex destabilization. In an alternative mode, (data not shown), RNP2$_1$ may inhibit the ATP/GTP domain, leaving a free RNP1$_1$ motif to mediate tetraplex disruption. Deletion of either the ATP/GTP binding box (ΔATP/GTP) or the RNP1$_1$ motif (ΔRNP1$_1$), and inhibition by RNP2$_1$ of the remaining active element, result in failure of these two mutant proteins to disrupt G′2 d(CGG)$_n$. However, deletion of the inhibitory RNP2$_1$ element from proteins that contain either an RNP1$_1$ motif or an ATP/GTP binding box (marked by an asterisk) as their single active element, enables the unblocked active motif to mediate G′2 d(CGG)$_n$ destabilization (26).

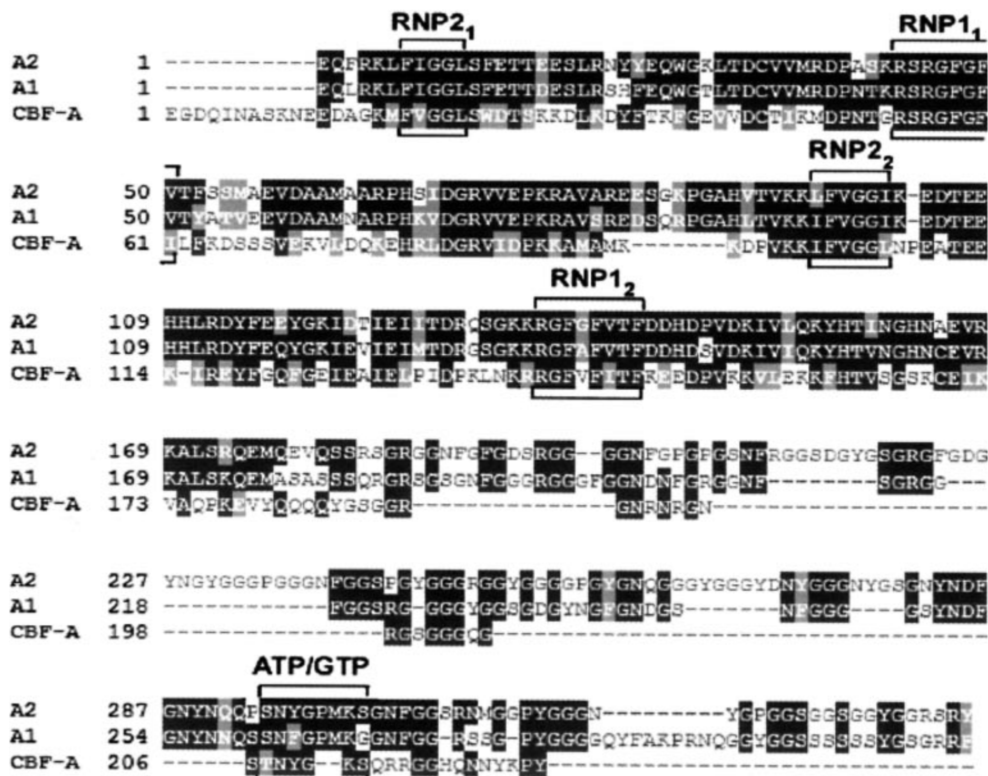

Figure 2. Homology of CBF-A, hnRNP A2 and hnRNP A1. Amino acid sequences of the three proteins were aligned by the ClustalW computer program and their homologous areas were marked by the Boxshade 3.21 program. Sequences of the four conserved RNP motifs and of the ATP/GTP binding domain are boxed.

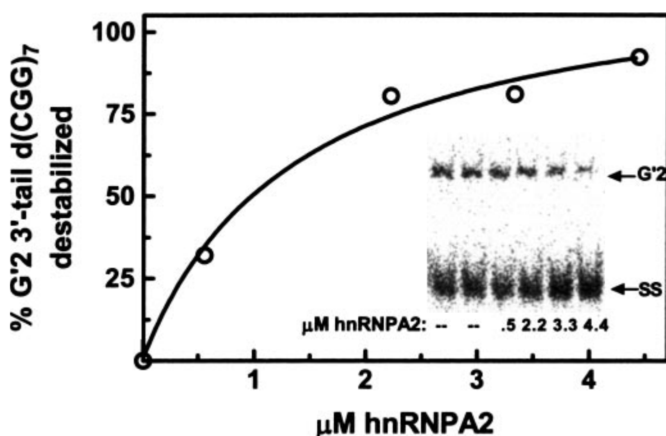

Figure 3. hnRNP A2 destabilizes a bimolecular tetraplex form of 3′-tail $d(CGG)_7$. Increasing amounts of recombinant hnRNP A2 that was purified and its GST was cleaved as described in Materials and Methods, were incubated at 37°C for 15 min with 200 fmol 5′-^{32}P G′2 3′-tail $d(CGG)_7$ under tetraplex DNA destabilization condition. Single-stranded products of the destabilization reaction were resolved from the remaining intact tetraplex DNA substrate by non-denaturing gel electrophoresis in 10% polyacrylamide as detailed in Materials and Methods. Shown in the inset is a phosphorimage of the electrophoretically separated G′2 and single-stranded forms of 3′-tail $d(CGG)_7$ and the results of their quantification are plotted.

$d(CGG)_7$ and single-stranded products of the reaction were resolved by non-denaturing gel electrophoresis. As seen in Figure 3, increasing amounts of hnRNP A2 progressively converted the tetraplex 3′-tail $d(CGG)_7$ substrate to its single-strand constituents. Hence, similarly to CBF-A (25,26), hnRNP A2 was also capable of disrupting a tetraplex structure of the $d(CGG)_n$ repeat sequence. To identify structural elements in hnRNP A2 that mediated the destabilization activity, we constructed and expressed in *E.coli* a series of recombinant hnRNP A2 mutants and the purified proteins were tested for their capacity to disrupt G′2 3′-tail $d(CGG)_7$. Results shown in Figure 4 demonstrated that the inactivating point mutations F[54]S or Y[269]F in the $RNP1_1$ motif or the ATP/GTP binding box, respectively, rendered hnRNP A2 incapable of G′2 3′-tail $d(CGG)_7$ disruption. These results suggested that either one of the two domains was required for the tetraplex $d(CGG)_n$ disruption activity. In parallel, however, when two amino acids were deleted from the $RNP2_1$ box of mutant proteins lacking an active $RNP1_1$ motif (Δ[14–15]/F[54]S) or an ATP/GTP binding domain (Δ[14–15]/Y[269]F), both double mutant proteins regained their tetraplex $d(CGG)_n$ destabilizing capability (Figure 4). Although higher amounts of the double mutant proteins relative to wild-type hnRNP A2 were necessary for maximum activity (Figure 4), inactivation of the $RNP2_1$ box clearly restored tetraplex $d(CGG)_n$ destabilization capacity to proteins that lacked either an $RNP1_1$ motif or an ATP/GTP binding domain. These results were interpreted as reflecting the blocking of either the $RNP1_1$ domain or the ATP/GTP binding box by the $RNP2_1$ motif of wild-type hnRNP A2. The restoration of tetraplex $d(CGG)_n$ disrupting activity by inactivation of the inhibitory $RNP2_1$ box in mutant proteins that possessed a single active motif, indicated that in the absence of inhibition by $RNP2_1$, either the $RNP1_1$ element or the ATP/GTP binding box alone

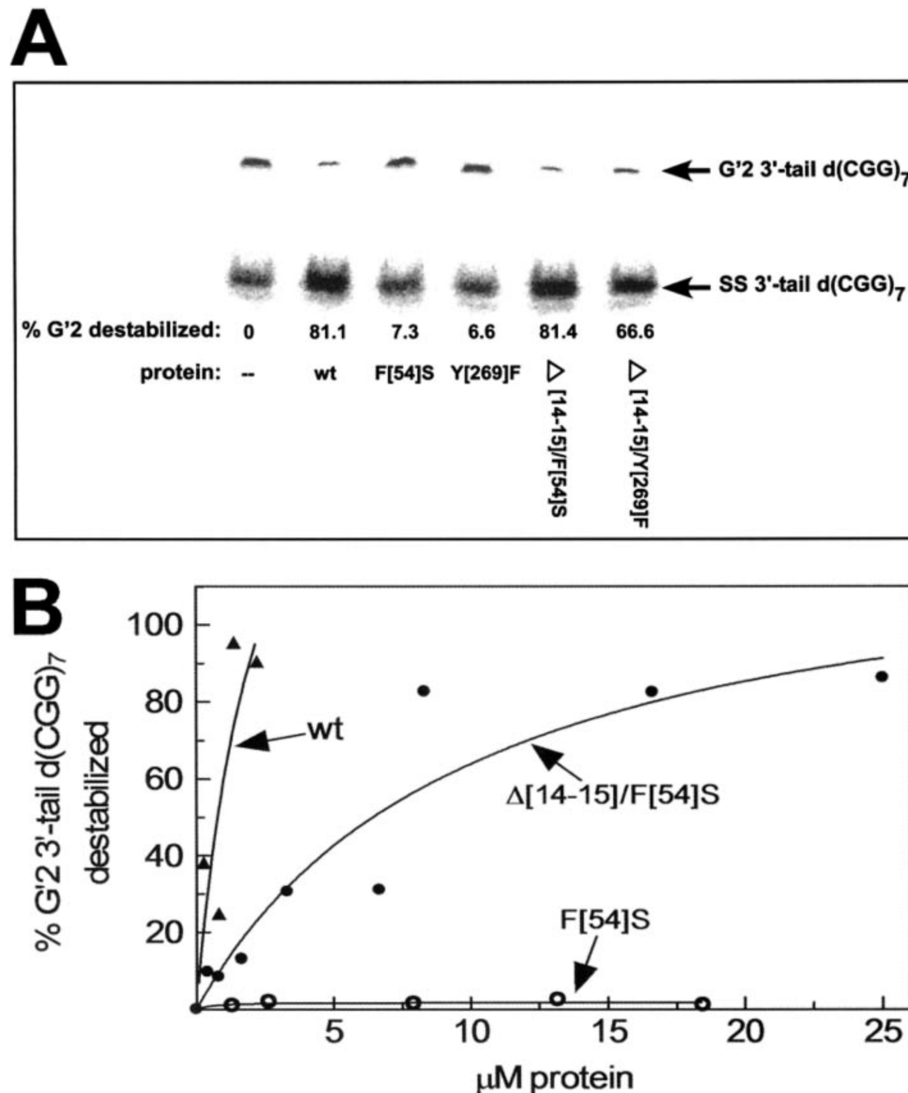

Figure 4. The $RNP1_1$ motif and the ATP/GTP binding box of hnRNP A2 act to destabilize G′2 3′-tail $d(CGG)_7$ and the $RNP2_1$ domain blocks their activity. (**A**) Phosphorimage of the destabilization of 5′-^{32}P G′2 3′-tail $d(CGG)_7$ by recombinant wild-type and mutant hnRNP A2 proteins. The tetraplex DNA substrate was incubated with each protein and single-stranded reaction products were resolved by gel electrophoresis as described in the legend to Figure 3. The amounts of proteins added were: 27.8 pmol wild-type hnRNP A2; 278 pmol F[54]S hnRNP A2 (mutated $RNP1_1$ motif); 139 pmol Y[269]F (mutated ATP/GTP box); and 278 pmol each of the respective double mutants with partially deleted [14–15] $RNP2_1$ motif: [14–15]/F[54]S and [14–15]/Y[269]F. Results of phosphorimaging quantification of the respective fraction of G′2 3′-tail $d(CGG)_7$ that each protein destabilized are indicated. (**B**) Destabilization of G′2 3′-tail $d(CGG)_7$ by wild-type hnRNP A2 and by mutant hnRNP A2 proteins with inactivated $RNP1_1$ box or of both $RNP1_1$ and $RNP2_1$ motifs. Increasing amounts of each protein were used to destabilize the tetraplex DNA substrate. Shown is a plot of phosphorimaging quantification of the accumulation of single-stranded products of the destabilization reaction.

sufficed for tetraplex $d(CGG)_n$ disruption. These results were qualitatively indistinguishable from those obtained with CBF-A [Figure 1 and (26)], and they strongly suggested that these two proteins utilized the same conserved motifs to disrupt tetrahelical forms of $d(CGG)_n$.

Inactivation of the $RNP2_1$ inhibitory motif enables hnRNP A1 to destabilize tetraplex $d(CGG)_n$

We next sought to corroborate the identification of $RNP1_1$ and $RNP2_1$ as the universal respective positive and negative mediators of tetraplex $d(CGG)_n$ destabilization. This was done by examining their roles in a third hnRNP ortholog, hnRNP A1. As seen in Figure 5, in contrast to CBF-A or hnRNP A2, recombinant wild-type hnRNP A1 did not exhibit G′2 3′-tail $d(CGG)_7$ disruption activity. We assumed that this failure to destabilize the tetraplex substrate was due to the degenerated sequence of the ATP/GTP binding motif of hnRNP A1 relative to hnRNP A2 or CBF-A (Figure 2). If this was the case, wild-type hnRNP A1 was not able to resolve G′2 $d(CGG)_n$ because its single active element $RNP1_1$ was blocked by the $RNP2_1$ inhibitory motif. To test this hypothesis, we constructed a [Δ18–22] hnRNP A1 mutant protein whose $RNP2_1$ domain was largely deleted. As shown in Figure 5, this mutant protein displayed robust G′2 3′-tail $d(CGG)_7$ destabilization activity. That this disruption activity was mediated by the $RNP1_1$ motif was demonstrated by the failure of the protein to resolve the tetraplex DNA substrate following the

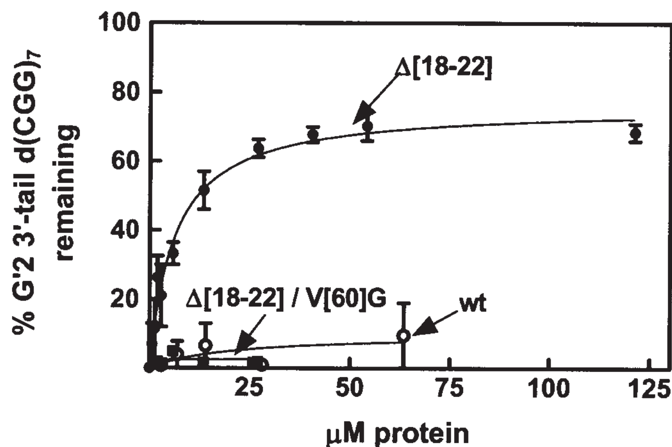

Figure 5. Inactivation of the $RNP2_1$ motif confers G′2 $d(CGG)_n$ destabilization capacity on hnRNP A1. Increasing amounts of recombinant wild-type or of the mutant hnRNP A1 proteins Δ[18–22] (deleted $RNP2_1$ box), or Δ[18–22]/V[60]G (deleted $RNP2_1$ box in combination with substituted $RNP1_1$ motif) were incubated with 200 fmol 5′-^{32}P G′2 3′-tail $d(CGG)_7$ under tetraplex $d(CGG)_n$ destabilization reaction conditions (see Materials and Methods). The single-stranded reaction products were resolved from the remaining intact tetraplex DNA substrate by non-denaturing gel electrophoresis. Shown is a plot of results of phosphorimaging quantification of the destabilization of G′2 3′-tail $d(CGG)_7$ by each protein.

introduction of the $RNP1_1$ inactivating mutation V[60]G into the [Δ18–22] hnRNP A1 mutant protein (Figure 5). These results indicated, therefore, that similar to CBF-A and hnRNP A2, the $RNP1_1$ and $RNP2_1$ motifs of hnRNP A1 acted, respectively, as positive and negative mediators of tetraplex $d(CGG)_n$ destabilization.

hnRNP A2, but not CBF-A, destabilizes $r(CGG)_7$ RNA tetraplex

Since both hnRNP A2 and CBF-A possess conserved elements typical of the RNA-interacting hnRNP family, we compared their capacities to interact with tetraplex forms of the RNA and DNA repeat sequences 3′-tail $r(CGG)_7$ and 3′-tail $d(CGG)_7$, respectively. Terminally labeled 3′-tail $r(CGG)_7$ or 3′-tail $d(CGG)_7$ were incubated for 10 min at their respective melting temperatures, 33 or 40°C, with increasing amounts of each protein. Quantification of the electrophoretically resolved substrates and products are shown in Figure 6. These data indicated that whereas hnRNP A2 efficiently destabilized tetraplex 3′-tail $r(CGG)_7$ (Figure 6A), CBF-A exerted an opposite effect by significantly increasing the thermal stability at 33°C of this RNA structure (Figure 6B). In clear contrast, the two proteins destabilized the DNA tetraplex G′2 3′-tail $d(CGG)_7$ at closely similar rates (Figure 6C). Hence, although hnRNP A2 and CBF-A destabilized G′2 3′-tail $d(CGG)_7$ at comparable efficiency, only hnRNP A2 was capable of disruption of its tetraplex RNA homologue G′2 3′-tail $r(CGG)_7$.

We next inquired whether or not hnRNP A2 utilized similar conserved domains to destabilize tetraplex forms of $d(CGG)_n$ or $r(CGG)_n$. Wild-type or mutant hnRNP A2 recombinant proteins were incubated at 33°C with terminally labeled tetraplex 3′-tail $r(CGG)_7$ and the extent of disruption of the tetrahelical RNA substrate was monitored following electrophoretic resolution of the reaction products in a non-denaturing gel. Results presented in Figure 7 indicated that introduction of the inactivating F[54]S or Y[269]F point mutations into the $RNP1_1$ box or the ATP/GTP binding domain, respectively, resulted in the loss of the tetraplex 3′-tail $r(CGG)_7$ disruption activity of hnRNP A2. However, these mutant proteins regained most of their tetraplex RNA destabilizing activity when their inhibitory $RNP2_1$ element was inactivated by deletion of two of its amino acid residues (Figure 7). These results were essentially indistinguishable from those obtained for the destabilization of G′2 3′-tail $d(CGG)_7$ by the same hnRNP A2 mutant proteins (Figure 4). It thus appeared that similar domains operate in hnRNP A2 to disrupt or inhibit the destabilization of DNA or RNA tetraplex forms of the $(CGG)_n$ repeat sequence.

An isolated ATP/GTP binding motif in a C-terminal fragment of CBF-A mediates selective destabilization of tetraplex $d(CGG)_n$

Our previous results (26) and the data shown above indicated that the ATP/GTP binding box was necessary and sufficient for tetraplex $d(CGG)_n$ destabilization by CBF-A or hnRNP A2. We inquired, therefore, whether a fragment of CBF-A devoid of any conserved motif except for the ATP/GTP binding box was capable of tetraplex $d(CGG)_n$ disruption activity. To this end, we constructed a [210–285] mini-gene that spanned the 75 C-terminal residues of CBF-A which included the ATP/GTP binding element but no other preserved domain. To specifically test whether the ATP/GTP binding box was responsible for the disruption of the tetraplex substrate by the CBF-A mini-protein, we prepared a mutant [210–285]/Δ260–266 mini-gene in which this motif was deleted (see Figure 8 for schemes of the CBF-A constructs). Wild-type, or each of the two mini CBF-A genes were expressed in *E.coli* and the respective proteins were purified and tested for their capacity to disrupt G′2 3′-tail $d(CGG)_7$. As is evident from results shown in Figure 8, the CBF-A [210–285] mini-protein effectively destabilized the G′2 3′-tail $d(CGG)_n$ tetraplex substrate. That this activity was mediated by the ATP/GTP binding domain was evident by the failure of the [210–285]/Δ260–266 mutant mini-protein to significantly disrupt G′2 3′-tail $d(CGG)_7$ (Figure 8).

In addition to its ability to destabilize tetraplex structures of $d(CGG)_n$, wild-type CBF-A protein was also shown to bind to, or to disrupt a wide variety of DNA sequences and structures (25,26,35,36). We thus compared the capacity of wild-type CBF-A and of its [210–285] mini-protein to interact with various DNA substrates. Data summarized in Table 2 indicated that wild-type CBF-A formed with varying efficacies complexes with single-stranded, duplex and tetraplex structures of diverse DNA sequences. In contrast, except for its measurable association with parallel G4 four-molecular tetraplexes of the 5′ E-box and the IgG switch sequences, the [210–285] CBF-A mini-protein did not detectably bind any of the tested DNA species. Results presented in Table 3 also showed that whereas wild-type CBF-A destabilized G′2 3′-tail $d(CGG)_7$ as well as double-stranded CArG box DNA, the [210–285] mini-protein selectively disrupted the tetraplex but not the duplex DNA substrate. All in all, these results demonstrated that an isolated ATP/GTP binding element is capable of mediating tetraplex $d(CGG)_n$ disruption activity in

the absence of any other conserved hnRNP motif. In contrast, the ability of CBF-A to interact with a variety of DNA sequences and structures was largely eliminated by the excision of a major N-terminal portion of this protein which included four of its five conserved motifs. Hence, relative to wild-type CBF-A, the [210–285] CBF-A mini-protein resolved tetraplex d(CGG)$_n$ with augmented selectivity.

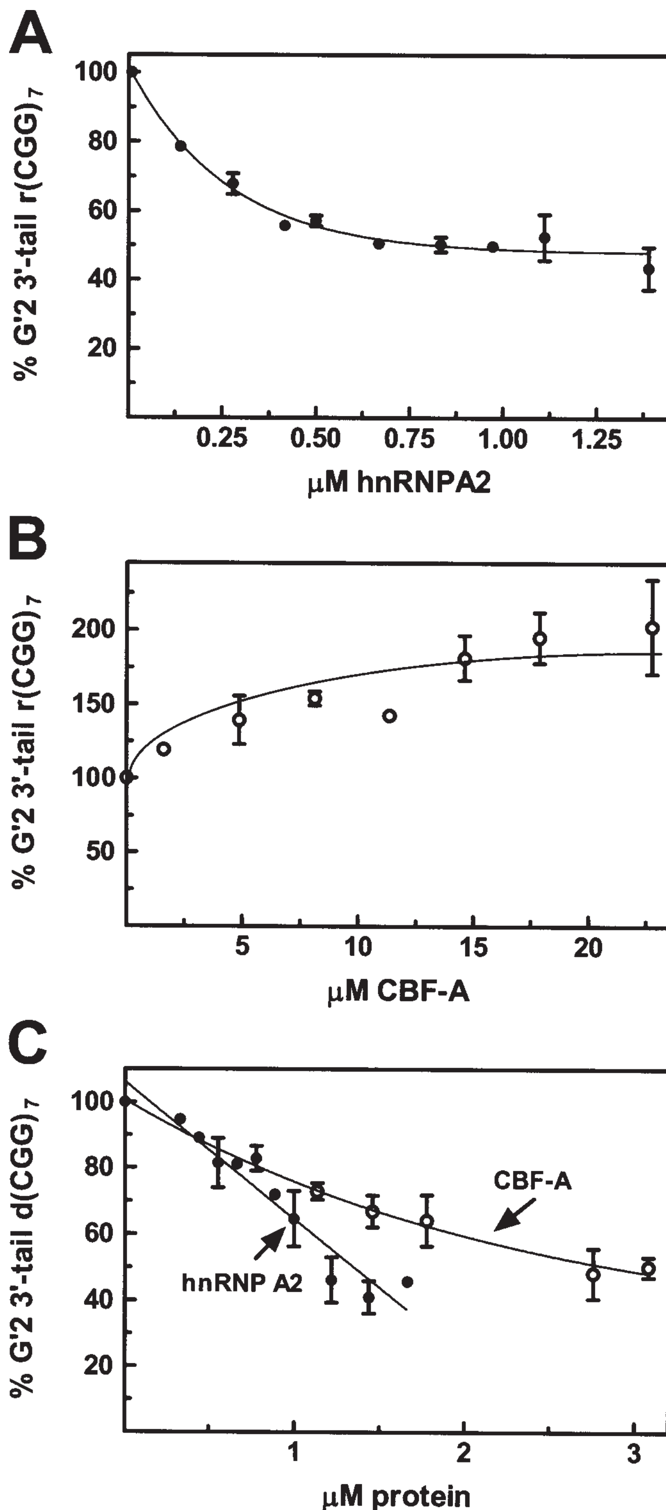

Figure 6. G′2 tetraplex form of the RNA repeat sequence 3′-tail r(CGG)$_7$ is disrupted by hnRNP A2 but stabilized by CBF-A. Increasing amounts of recombinant wild-type hnRNP A2 (**A**) or CBF-A (**B**) were incubated at 33°C for 10 min with 250 fmol 5′-^{32}P G′2 3′-tail r(CGG)$_7$ under RNA tetraplex destabilization reaction conditions (see Materials and Methods). In control experiments (**C**), hnRNP A2 or CBF-A were incubated at 40°C for 10 min with 200 fmol 5′-^{32}P G′2 3′-tail d(CGG)$_7$. The incubation temperatures, 33 or 40°C were determined as the respective melting temperatures, T_m, of the RNA and DNA tetrahelices. Following termination of the reactions, RNA or DNA single strands were separated from their tetraplex structures by electrophoresis through non-denaturing polyacrylamide gel and their respective proportions were quantified by phosphorimaging analysis as described in Materials and Methods. A value of 100% G′2 3′-tail r(CGG)$_7$ that was determined in samples that were incubated without protein at 33°C for 10 min represented 46.6% of the total RNA. This fraction of G′2 3′-tail r(CGG)$_7$ was progressively decreased (A) or increased (B) at 33°C in the presence of increasing amounts of hnRNP A2 or CBF-A, respectively. G′2 3′-tail d(CGG)$_7$ as determined in samples that were incubated without protein, represented 34.7% of the total DNA. This fraction of G′2 3′-tail d(CGG)$_7$ was decreased at 40°C to a similar extent in the presence of increasing amounts of either hnRNP A2 or CBF-A (C).

We next compared the extent of selective destabilization of G′2 3′-tail d(CGG)$_7$ by the [210–285] CBF-A mini-protein and native hnRNP proteins under more physiologic-like conditions. We found that the tetraplex d(CGG)$_n$ destabilization activities of wild-type CBF-A, hnRNP A2 and Δ[18–22] hnRNP A1 mutant protein were reduced by 9.0 to 55.0% in reaction mixtures that contained 1.0 mM ATP and 330-fold (w/w) excess of unlabeled double-stranded salmon sperm DNA. In contrast, the [210–285] CBF-A mini-protein maintained full G′2 3′-tail d(CGG)$_7$ destabilization activity in the presence of ATP and competing native DNA at these concentrations.

DISCUSSION

Hairpin (12–15) and tetraplex (16–18) structures formed by the d(CGG) trinucleotide repeat sequence were implicated in the blocking of DNA replication (18–20). Pausing of the DNA synthesis at these structures of d(CGG)$_n$ in *FMR1* may be responsible for the observed delayed replication of this gene in fragile X cells (37) and are a plausible primary cause for polymerase slippage and expansion of the repeat sequence. The transcription of *FMR1* may also be impeded by thermodynamically stable secondary structures of the expanded (CGG)$_n$ tract in either the non-transcribed DNA strand or in product mRNA molecules. Last, the expanded r(CGG)$_n$ tract in *FMR1* mRNA molecules that are synthesized in some fragile X individuals, (38) may also fold into structures that inhibit the translation of the FMRP protein product (39). In this context, agents that diminish the generation and the stability of hairpin and tetraplex structures of the (CGG) trinucleotide repeat in DNA or RNA should reduce the likelihood of its expansion in normal or fragile X premutation cells. As importantly, they may contribute to the restoration of the *FMR1* transcription and to the synthesis of FMRP in cells of individuals that already carry a full expansion mutation.

The cationic porphyrin 5,10,15,20-tetra(*N*-methyl-4-pyridyl)porphine (TMPyP4) (33) and the WRN protein (19,23) were shown to effectively destabilize G′2 dimeric tetraplex structures of d(CGG)$_n$. Likewise, the hnRNP-related

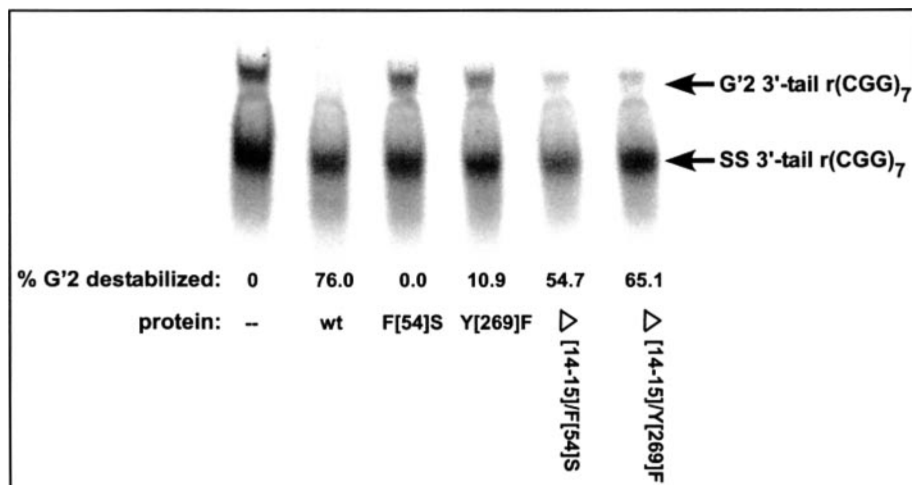

Figure 7. Destabilization of G′2 3′-tail r(CGG)$_7$ by hnRNP A2 is mediated by its RNP1$_1$ motif and the ATP/GTP binding box and is inhibited by the RNP2$_1$ domain. 250 fmol 5′-^{32}P G′2 3′-tail r(CGG)$_7$, were incubated at 33°C for 10 min under RNA destabilization reaction conditions with 19.5 or 39.0 pmol, respectively, of wild-type or each mutant hnRNP A2 protein and single-stranded reaction products were resolved by gel electrophoresis as described in the legend to Figure 6. Mutant proteins were as follows: F[54]S hnRNP A2 (mutated RNP1$_1$ motif); Y[269]F hnRNP A2 (mutated ATP/GTP box); and their respective double mutants with partially deleted [14–15] RNP2$_1$ motif: [14–15]/F[54]S or [14–15]/Y[269]F. Shown is a phosphorimage of gel-resolved reaction products and results of the quantification of the fraction of G′2 3′-tail r(CGG)$_7$ that each protein destabilized.

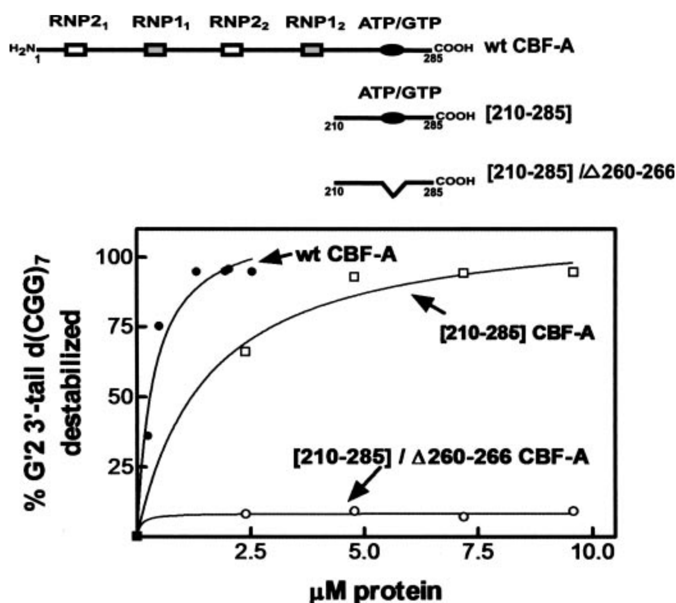

Figure 8. An ATP/GTP binding motif of a CBF-A C-terminal mini-protein mediates destabilization of G′2 3′-tail d(CGG)$_7$. Increasing amounts of recombinant wild-type CBF-A, its C-terminal [210–285] fragment or this fragment with a deleted ATP/GTP binding box [210–285]/Δ260–266, were incubated with 200 fmol 5′-^{32}P G′2 3′-tail d(CGG)$_7$ under tetraplex DNA destabilization conditions and the single-stranded reaction products were separated from the remaining intact tetraplex DNA substrate by non-denaturing gel electrophoresis. Full-length CBF-A and the two mini-proteins are drawn schematically at the top of the figure. Shown is a plot of results of phosphorimaging quantification of the destabilization of G′2 3′-tail d(CGG)$_7$ by each protein.

protein CBF-A was also found to disrupt bimolecular tetrahelical structures of d(CGG)$_n$ (25,26). Systematic analysis of the capacity of numerous CBF-A mutant proteins to destabilize tetraplex d(CGG)$_n$ indicated that two of the five conserved domains that CBF-A shares with other members of the hnRNP family, the RNP1$_1$ motif and the ATP/GTP binding box, were necessary and sufficient for the disruption of G′2 d(CGG)$_n$. At the same time, the RNP2$_1$ domain of CBF-A suppressed the tetraplex destabilization activity of either one of these two elements (26). In this work, we showed that another hnRNP ortholog, hnRNPA2, was also capable of destabilization of G′2 d(CGG)$_n$ (Figure 3). While this activity was abolished by the inactivation of either the RNP1$_1$ box or the ATP/GTP binding motif, the deletion of a significant portion of the RNP2$_1$ domain in each one of these two mutant proteins reestablished their ability to disrupt tetraplex d(CGG)$_n$ (Figure 4). Overall, these results indicated that the RNP1$_1$ motif or the ATP/GTP binding box of hnRNP A2 acted as positive mediators of G′2 d(CGG)$_n$ destabilization whereas the RNP2$_1$ element blocked this activity. In a similar vein, hnRNP A1 a third hnRNP ortholog, also utilized its RNP2$_1$ domain to suppress the tetraplex d(CGG)$_n$ activity that its RNP1$_1$ motif mediated (Figure 5). As these results were indistinguishable from those formerly obtained for CBF-A (Figure 1), we concluded that several members of the hnRNP family possessed an ability to disrupt tetrahelices of the d(CGG) trinucleotide repeat sequence and that homologous conserved domains were employed by the three ortholog proteins to execute this activity.

For any protein to be utilized as a dedicated tetraplex d(CGG)$_n$ destabilizing agent, its ability to interact with DNA sequences and structures other than the tetrahelical form of this trinucleotide repeat sequence should be diminished to a minimum. CBF-A which was originally identified by its preferential binding to the CArG box DNA sequence (27), was subsequently shown to associate with diverse DNA sequences and structures, most notably tetraplex structures of the telomeric repeat sequence d(TTAGGG)$_n$ (25,26). As the ATP/GTP binding box of CBF-A was found to be necessary and sufficient for G′2 d(CGG)$_n$ disruption, we constructed

Table 2. Binding of different DNA sequences and structures by wild-type and [210–280] CBF-A mini-protein

DNA ligand	Protein (pmol) required for binding 1.0 pmol DNA wt CBF-A	[210–285] CBF-A
Double-stranded E-box	333.0	>1000
G4 tetraplex 5′ E-box	1.0	1.0
Single-stranded anti TeR4	10.5	>1000
Double-stranded TeR4•anti TeR4	5.3	>1000
G′4 TeR4	112.5	>1000
Single-stranded CArG box	55.5	>1000
Single-stranded anti CArG box	16.6	>1000
Double-stranded CArG•anti CArG box	120.5	>1000
G4 tetraplex IgG switch sequence	1.1	385

The listed [5′-^{32}P]DNA ligands at 0.2–1.0 pmol each in a final reaction volume of 10 µl, were incubated for 20 min at 4°C with 0.05–20 pmol wild-type or 480 pmol [210–280] mini CBF-A recombinant proteins. Protein–DNA complexes were resolved by mobility shift electrophoresis and quantified by phosphorimaging analysis (see Materials and Methods). Undetectable binding of DNA to 480 pmol of the mini-protein is marked as >1000 pmol protein required to bind 1.0 pmol DNA ligand.

Table 3. Destabilization of different DNA sequences and structures by wild-type and [210–280] CBF-A mini-protein

DNA	% DNA destabilized wt CBF-A	[210–285] CBF-A
G′2 tetraplex 3′-tail $d(CGG)_7$	92.5	60.5
G4 tetraplex 5′ E-box	0.0	0.0
G4 tetraplex IgG switch sequence	0.0	0.0
Double-stranded CArG box	55.0	0.0

Terminally ^{32}P-labeled tetraplex or duplex DNA ligands at 0.2 pmol each were incubated for 15 min at 37°C with 30 pmol of wild-type CBF-A or 480 pmol [210–280] CBF-A mini-protein. Assay conditions for the destabilization of the DNA ligands, separation of the reaction products by non-denaturing gel electrophoresis and their quantification by phosphorimaging analysis were described in Materials and Methods.

a recombinant mini CBF-A protein whose single remaining conserved element was the ATP/GTP binding fold. This motif enabled the mini-protein to destabilize tetraplex $d(CGG)_n$, albeit at a reduced efficacy relative to the wild-type protein (Figure 8 and Table 3). It should be stressed, however, that neither the wild-type CBF-A protein, nor its C-terminal fragment required nucleotide triphosphates for their tetraplex $d(CGG)_n$ disruption activity. Most notable, the propensity of wild-type CBF-A to interact with diverse DNA sequences and structures was greatly diminished by the absence of the four conserved RNP motifs and of a major N-terminal portion of wild-type CBF-A (Tables 2 and 3). These findings provided an additional demonstration that the ATP/GTP box was necessary and sufficient for G′2 $d(CGG)_n$ destabilization. More interestingly, whereas the tetraplex destabilization activities of wild-type CBF-A, hnRNP A2 and mutant hnRNP A1 DNA were reduced to varying degrees in the presence of physiologic concentration of ATP and excess double-stranded DNA, the CBF-A mini-protein maintained full activity under these conditions (see Results). Hence, due to its higher selectivity, ectopically introduced CBF-A mini-protein should elicit more preferential disruption of secondary structures of the d(CGG) repeat in fragile X cells than wild-type protein.

Some individuals who carried a d(CGG) tract in excess of 200 repeats were reported to express normal, (39) or even elevated, (38,40,41) levels of *FMR1* mRNA. Yet, these fully affected persons failed to synthesize FMRP and in one case, their *FMR1* transcripts were found to be associated with a stalled 40S ribosomal subunit (39). It was proposed that secondary structures of the $r(CGG)_n$ sequence impeded ribosome advancement beyond the trinucleotide repeat in *FMR1* mRNA (39). Indeed, recent results indicated that $r(CGG)_n$ runs formed hairpin structures and folded into tetraplex formations (42). Thus, agents that resolve secondary structures of this RNA sequence are of potential utility. Data presented in Figures 6 and 7 showed that hnRNP A2 was able to proficiently disrupt a tetraplex formation of $r(CGG)_n$ and that this activity was mediated by the same conserved domains that executed destabilization of G′2 $d(CGG)_n$. Interestingly, CBF-A exerted an opposite effect by stabilizing this RNA tetraplex (Figure 6). It was proposed that some RNA interacting proteins may regulate the transport or translation of specific mRNA molecules by binding to and modulating the stability of their RNA tetraplex domains as suggested for the RNA interacting FMRP itself (43,44). In line with this notion, it is tempting to speculate that the opposite effect of hnRNP A2 and CBF-A on the stability of tetraplex forms of $r(CGG)_n$ might reflect their contrasting respective roles in the regulation of the transport of *FMR1* mRNA or its translation into FMRP.

ACKNOWLEDGEMENTS

This study was supported by grants to M.F. from the Conquer Fragile X Foundation, the US–Israel Binational Science Foundation, the Israel Science Foundation and the Fund for Promotion of Research in the Technion, and by a National Cancer Institute Program Project Grant CA77852 to L.A.L.

REFERENCES

1. Simonsson,T. (2001) G-quadruplex DNA structures—variations on a theme. *Biol. Chem.*, **382**, 621–628.
2. Arthanari,H. and Bolton,P.H. (2001) Functional and dysfunctional roles of quadruplex DNA in cells. *Chem. Biol.*, **8**, 221–230.
3. Henderson,E. (1995) Telomere DNA structure. In Blackburn,E.H. and Greider,C.W. (eds), *Telomeres*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp. 11–34.
4. Arthanari,H. and Bolton,P.H. (2003) Did quadruplex DNA play a role in the evolution of the eukaryotic linear chromosome? *Mini Rev. Med. Chem.*, **3**, 1–9.
5. Sen,D. and Gilbert,W. (1988) Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature*, **334**, 364–366.
6. Simonsson,T., Pecinka,P. and Kubista,M. (1998) DNA tetraplex formation in the control region of *c-myc*, *Nucleic Acids Res.*, **26**, 1167–1172.
7. Siddiqui-Jain,A., Grand,C.L., Bearss,D.J. and Hurley,L.H. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress *c-MYC* transcription. *Proc. Natl Acad. Sci. USA*, **99**, 11593–11598.
8. Grand,C.L., Powell,T.J., Nagle,R.B., Bearss,D.J., Tye,D., Gleason-Guzman,M. and Hurley,L.H. (2004) Mutations in the G-quadruplex silencer element and their relationship to *c-MYC* overexpression, NM23

repression, and therapeutic rescue. *Proc. Natl Acad. Sci. USA*, **101**, 6140–6145.
9. Hammond-Kosack,M.C., Kilpatrick,M.W. and Docherty,K. (1993) The human insulin gene-linked polymorphic region adopts a G-quartet structure in chromatin assembled *in vitro*. *J. Mol. Endocrinol*., **10**, 121–126
10. Nielsen,F.C. and Christiansen,J. (1995) Posttranscriptional regulation of insulin-like growth factor II mRNA. *Scand. J. Clin. Lab. Invest. Suppl*., **220**, 37–46.
11. Catasti,P., Chen,X., Moyzis,R.K., Bradbury,E.M. and Gupta,G. (1996) Structure-function correlations of the insulin-linked polymorphic region. *J. Mol. Biol*., **264**, 534–545.
12. Nadel,Y., Weisman-Shomer,P. and Fry,M. (1995) The fragile X syndrome single strand d(CGG)$_n$ nucleotide repeats readily fold back to form unimolecular hairpin structures. *J. Biol. Chem*., **270**, 28970–28977.
13. Gacy,A.M., Goellner,G., Juranic,N., Macura,S. and McMurray,C.T. (1995) Trinucleotide repeats that expand in human disease form hairpin structures *in vitro*. *Cell*, **81**, 533–540.
14. Chen,X., Mariappan,S.V., Catasti,P., Ratliff,R., Moyzis,R.K., Laayoun,A., Smith,S.S., Bradbury,E.M. and Gupta,G. (1995) Hairpins are formed by the single DNA strands of the fragile X triplet repeats: structure and biological implications. *Proc. Natl Acad. Sci. USA*, **92**, 5199–5203.
15. Mitas,M., Yu,A., Dill,J., Kamp,T.J., Chambers,E.J. and Haworth,I.S. (1995) Hairpin properties of single-stranded DNA containing a GC-rich triplet repeat: (CTG)$_{15}$. *Nucleic Acids Res*., **23**, 1050–1059.
16. Fry,M. and Loeb,L.A. (1994) The fragile X syndrome d(CGG)$_n$ nucleotide repeats form a stable tetrahelical structure. *Proc. Natl Acad. Sci USA*, **91**, 4950–4954.
17. Kettani,A., Kumar,R.A. and Patel,D.J. (1995) Solution structure of a DNA quadruplex containing the fragile X syndrome triplet repeat. *J. Mol. Biol*., **254**, 638–656.
18. Usdin,K. and Woodford,K.J. (1995) CGG repeats associated with DNA instability and chromosome fragility form structures that block DNA synthesis *in vitro*. *Nucleic Acids Res*., **23**, 4202–9.
19. Kamath-Loeb,A.S., Loeb,L.A., Johansson,E., Burgers,P.M. and Fry,M. (2001) Interactions between the Werner syndrome helicase and DNA polymerase delta specifically facilitate copying of tetraplex and hairpin structures of the d(CGG)$_n$ trinucleotide repeat sequence. *J. Biol. Chem*., **276**, 16439–16446.
20. Kang,S., Ohshima,K., Shimizu,M., Amirhaeri,S. and Wells,R.D. (1995) Pausing of DNA synthesis *in vitro* at specific loci in CTG and CGG triplet repeats from human hereditary disease genes. *J. Biol. Chem*., **270**, 27014–27021.
21. Sun,H., Bennett,R.J. and Maizels,N. (1999) The *Saccharomyces cerevisiae* Sgs1 helicase efficiently unwinds G-G paired DNAs. *Nucleic Acids Res*., **27**, 1978–1984.
22. Sun,H., Karow,J.K., Hickson,I.D. and Maizels,N. (1998) The Bloom's syndrome helicase unwinds G4 DNA. *J. Biol. Chem*., **273**, 27587–27592.
23. Fry,M. and Loeb,L.A. (1999) Human Werner syndrome DNA helicase unwinds tetrahelical structures of the fragile X syndrome repeat sequence d(CGG)$_n$. *J. Biol. Chem*., **274**, 12797–12802.
24. Mohaghegh,P., Karow,J.K., Brosh,R.M., Jr., Bohr,V.A. and Hickson,I.D. (2001) The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases. *Nucleic Acids Res*., **29**, 2843–2849.
25. Weisman-Shomer,P., Naot,Y. and Fry,M. (2000) Tetrahelical forms of the fragile X syndrome expanded sequence d(CGG)$_n$ are destabilized by two heterogeneous nuclear ribonucleoprotein-related telomeric DNA-binding proteins. *J. Biol. Chem*., **275**, 2231–2238.
26. Weisman-Shomer,P., Cohen,E. and Fry,M. (2002) Distinct domains in the CArG-box binding factor A destabilize tetraplex forms of the fragile X expanded sequence d(CGG)$_n$. *Nucleic Acids Res*., **30**, 3672–3681.
27. Kamada,S. and Miwa,T. (1992) A protein binding to CArG box motifs and to single-stranded DNA functions as a transcriptional repressor. *Gene*, **119**, 229–236.
28. Mikheev,A.M., Mikheev,S.A., Zhang,Y., Aebersold,R. and Zarbl,H. (2000) CArG binding factor A (CBF-A) is involved in transcriptional regulation of the rat Ha-ras promoter. *Nucleic Acids Res*., **28**, 3762–3770.
29. Bemark,M., Olsson,H., Heinegard,D. and Leanderson,T. (1998) Purification and characterization of a protein binding to the SP6 kappa promoter. A potential role for CArG-box binding factor-A in kappa transcription. *J. Biol. Chem*., **273**, 18881–18890.
30. Rushlow,W.J., Rajakumar,B., Flumerfelt,B.A., Naus,C.C. and Rajakumar,N. (2000) Changes in CArG-binding protein A expression levels following injection(s) of the D1-dopamine agonist SKF-82958 in the intact and 6-hydroxydopamine-lesioned rat. *Neuroscience*, **98**, 69–78.
31. Rushlow,W.J., Rajakumar,N., Flumerfelt,B.A. and Naus,C.C. (1999) Characterization of CArG-binding protein A initially identified by differential display. *Neuroscience*, **94**, 637–649.
32. Sambrook,J. and Russel,D.W. (2001) *Molecular Cloning: A Laboratory Manual*, *3rd edn*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
33. Weisman-Shomer,P., Cohen,E., Hershco,I., Khateb,S., Wolfovitz-Barchad,O., Hurley,L.H. and Fry,M. (2003) The cationic porphyrin TMPyP4 destabilizes the tetraplex form of the fragile X syndrome expanded sequence d(CGG)$_n$. *Nucleic Acids Res*., **31**, 3963–3970.
34. Walsh,K. and Gualberto,A. (1992) MyoD binds to the guanine tetrad nucleic acid structure. *J. Biol. Chem*., **267**, 13714–13718.
35. Sarig,G., Weisman-Shomer,P., Erlitzki,R. and Fry,M. (1997) Purification and characterization of qTBP42, a new single-stranded and quadruplex telomeric DNA-binding protein from rat hepatocytes. *J. Biol. Chem*., **272**, 4474–4482.
36. Sarig,G., Weisman-Shomer,P. and Fry,M. (1997) Telomeric and tetraplex DNA binding properties of qTBP42: a homologue of the CArG box binding protein CBF-A. *Biochem. Biophys. Res. Commun*., **237**, 617–623.
37. Hansen,R.S., Canfield,T.K., Lamb,M.M., Gartler,S.M. and Laird,C.D. (1993) Association of fragile X syndrome with delayed replication of the *FMR1* gene. *Cell*, **73**, 1403–1409.
38. 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.
39. Feng,Y., Zhang,F., Lokey,L.K., Chastain,J.L., Lakkis, L, Eberhart,D. and Warren,S.T. (1995) Translational suppression by trinucleotide repeat expansion at *FMR1*. *Science*, **268**, 731–734.
40. Tassone,F., Hagerman,R.J., Loesch,D.Z., Lachiewicz,A., Taylor,A.K. and Hagerman,P.J. (2000) Fragile X males with unmethylated, full mutation trinucleotide repeat expansions have elevated levels of FMR1 messenger RNA. *Am. J. Med. Genet*., **94**, 232–236.
41. 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.
42. Handa,V., Saha,T. and Usdin,K. (2003) The fragile X syndrome repeats form RNA hairpins that do not activate the interferon-inducible protein kinase, PKR, but are cut by Dicer. *Nucleic Acids Res*., **31**, 6243–6248.
43. 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.
44. 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.