RNA (2000), *6*:79–87. Cambridge University Press. Printed in the USA.

Expanded CUG repeat RNAs form hairpins that activate the double-stranded RNA-dependent protein kinase PKR

BIN TIAN,[1] ROBERT J. WHITE,[2] TIANBING XIA,[3] STEPHEN WELLE,[4] DOUGLAS H. TURNER,[3] MICHAEL B. MATHEWS,[1] and CHARLES A. THORNTON[2]

[1]Department of Biochemistry and Molecular Biology, New Jersey Medical School, University of Medicine and Dentistry of New Jersey, Newark, New Jersey 17103, USA
[2]Department of Neurology, School of Medicine and Dentistry, University of Rochester, Rochester, New York 14642, USA
[3]Department of Chemistry, School of Medicine and Dentistry, University of Rochester, Rochester, New York 14642, USA
[4]Department of Medicine, School of Medicine and Dentistry, University of Rochester, Rochester, New York 14642, USA

ABSTRACT

Myotonic dystrophy is caused by an expanded CTG repeat in the 3′ untranslated region of the DM protein kinase (DMPK) gene. The expanded repeat triggers the nuclear retention of mutant DMPK transcripts, but the resulting underexpression of DMPK probably does not fully account for the severe phenotype. One proposed disease mechanism is that nuclear accumulation of expanded CUG repeats may interfere with nuclear function. Here we show by thermal melting and nuclease digestion studies that CUG repeats form highly stable hairpins. Furthermore, CUG repeats bind to the dsRNA-binding domain of PKR, the dsRNA-activated protein kinase. The threshold for binding to PKR is ~15 CUG repeats, and the affinity increases with longer repeat lengths. Finally, CUG repeats that are pathologically expanded can activate PKR in vitro. These results raise the possibility that the disease mechanism could be, in part, a gain of function by mutant DMPK transcripts that involves sequestration or activation of dsRNA binding proteins.

Keywords: CUG repeat; double-stranded RNA; dsRNA-binding motif; muscle; myotonic dystrophy; PKR

INTRODUCTION

Myotonic dystrophy (dystrophia myotonica, DM) is an autosomal-dominant, multisystem disease in which the phenotype is highly variable (Harper, 1989). The congenital form of DM is characterized by delayed muscle maturation, mental retardation, and respiratory distress. The adult-onset form is characterized by muscle wasting, neuropsychiatric impairment, heart block, and cataracts. The only disease manifestation in some late-onset cases is presenile cataract.

The genetic basis for DM is an expanded CTG repeat in the 3′ untranslated region of the DM protein kinase (DMPK) gene (Brook et al., 1992). The number of CTG repeats at this locus ranges from 5 to 37 repeats in the normal population, whereas individuals with DM have from 50 to more than 3,000 repeats. The severity of disease manifestations correlates with the length of the expanded repeat.

Reprint requests to: Charles A. Thornton, Box 673, URMC, Room 5-4306, 601 Elmwood Avenue, Rochester, New York 14642, USA; e-mail: charles_thornton@urmc.rochester.edu.

Transcripts from the mutant DMPK allele are retained in the nucleus (Taneja et al., 1995; Davis et al., 1997), resulting in a modest reduction of DMPK protein in skeletal muscle (Maeda et al., 1995). Partial loss of DMPK probably does not, however, provide a unitary explanation for the phenotype, because the development and maturation of skeletal muscle proceeds normally in mice that are homozygous for a DMPK-null allele (Jansen et al., 1996; Reddy et al., 1996). Although one strain of DMPK knockout mice developed a mild, late-onset myopathy, the phenotype was not clearly similar to DM (Reddy et al., 1996). Of note, point mutations in the DMPK gene have not been found in any kindreds with DM.

One proposed mechanism for the pathogenesis of DM is that nuclear accumulation of expanded CUG repeats is deleterious. A similar mechanism may be operating in spinocerebellar ataxia type 8 (SCA8), a dominantly inherited neurodegenerative disease that is also caused by the expansion of an untranslated CTG repeat (Koob et al., 1999). As there are no precedents in human genetics for transdominant RNA gain-of-function, mechanistic models are at an early stage of

development. Timchenko and colleagues (Timchenko et al., 1996; Philips et al., 1998) have proposed that expanded CUG repeats titrate a CUG-binding protein, CUGBP1, rendering it unavailable to perform its normal function as a splice enhancer.

To investigate the RNA gain-of-function model, we have studied the structural properties and protein interactions of CUG repeats in vitro. A panel of triplet repeat RNAs was synthesized by in vitro transcription. Optical melting and nuclease mapping studies indicate that CUG repeats form highly stable hairpins. Although conventional base pairing in the stem of the hairpin is interrupted by a periodic U•U mismatch, these transcripts are able to bind, and to activate PKR, the dsRNA-activated protein kinase. These observations suggest that interactions with dsRNA-binding proteins could play a role in the nuclear retention or toxicity of mutant DMPK transcripts.

RESULTS

An RNA folding algorithm (Walter et al., 1994; Mathews et al., 1999) predicts that CUG repeats form stable hairpins (Fig. 1A). G•C and C•G base pairs in the stem of the hairpin are separated by a periodic U•U mismatch. The most stable structure predicted for a CUG repeat of any length is the hairpin with a single loop and the longest possible stem, but multistem/loop structures (e.g., Fig. 1B) are also possible, as their predicted free energies are only slightly less favorable. CAG repeat RNAs are also predicted to form hairpins, whereas CUC repeat RNAs are not expected to form any stable secondary structure.

To test these predictions, UV absorbance melting experiments were conducted using a panel of triplet repeat RNAs synthesized by in vitro transcription. The plasmid templates for synthesizing CUG and CUC repeat transcripts were designed so that ATP could be omitted from the transcription reactions. This design minimizes the RNA-dependent polymerase activity of the T7 RNA polymerase (Cazenave & Uhlenbeck, 1994), and prevents the synthesis of antisense contaminants (Mellits et al., 1990) or hybrid sense/antisense run-off transcripts (Triana-Alonso et al., 1995). Similarly, CAG repeat transcripts were synthesized in the absence of UTP.

The normalized UV absorption melting curves for six $(CUG)_n$ transcripts ranging in size from 5 to 69 repeats all displayed a single sharp transition with 44–72% hyperchromicity (Fig. 2). The melting temperatures were ~75 °C for each of the $(CUG)_n$ transcripts, but the transitions were steeper, and the hyperchromicities larger, for longer repeats. The melting temperature of the transcript with 69 repeats was independent of RNA concentration over a 100-fold range of concentrations (not shown). By comparison, CAG repeat RNAs showed a broader transition with a lower melting point and a smaller hyperchromicity. As expected, the CUC repeat RNA showed no consistent transition.

The melting profiles of CUG repeat transcripts were consistent with a simple, unimolecular secondary structure, such as the predicted hairpin, or other, more complex, multistem/loop folded structures. To distinguish among these, RNAse T1, an endoribonuclease that cleaves preferentially after unpaired guanosines, was used to map the position of nuclease-sensitive loops.

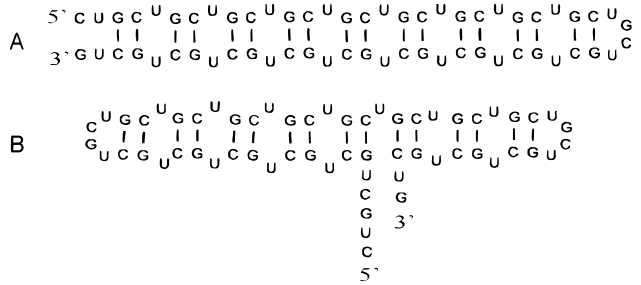

FIGURE 1. Folding algorithms predict a stable secondary structure for CUG repeats. **A**: A simple hairpin is the most stable secondary structure predicted for $(CUG)_{20}$. **B**: One example of an alternative multiloop structure.

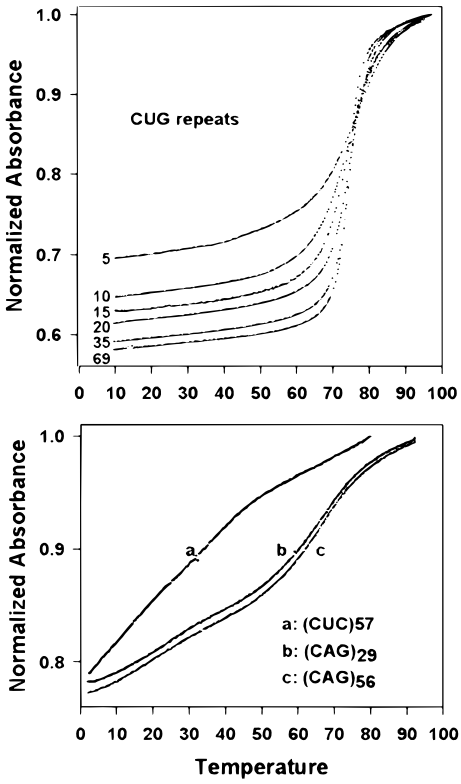

FIGURE 2. Thermal melting profiles of triplet repeat RNAs. Absorbance was normalized relative to the value at 80 °C (for CUC repeats) or 97 °C (all other transcripts). Transcripts have the following formulas: GGGCGG(UGC)$_n$U for CUG repeats, GGGAGG(AGC)$_n$ for CAG repeats, and 5′-GGCGCUGG(CCU)$_n$CCC for CUC repeats.

Digests of the 35-, 69-, and 140-repeat transcripts revealed a nuclease sensitive site near the midpoint of each transcript (Fig. 3). Thus, under the renaturation and digestion conditions used here, even the longest transcript (423 nt) formed a simple hairpin with an extended stem, rather than a series of smaller folds. These data confirm observations by Napierala and Krzyzosiak (1997), who used RNase and lead cleavage to map hairpin loops in transcripts with up to 49 CUG repeats, and extend the findings into the range of repeat sizes that cause DM. Expanded CUG repeats, however, are resistant to cleavage by RNAse III at a concentration that completely degrades authentic dsRNAs (Fig. 4A), confirming that the duplex structure is not due to a contaminating authentic dsRNA. Evidently, the U•U mismatches in the hairpin stem confer resistance to this dsRNA-modifying enzyme.

These results suggest that the dominant structural feature of an expanded CUG repeat is the duplex-like stem. To support this conclusion, we studied the interactions of CUG repeats with a dsRNA-binding protein. PKR was selected for these analyses because its binding activity is well characterized. Furthermore, an inventory of muscle mRNA using the serial analysis of gene expression (SAGE) method (Velculescu et al., 1995) indicated that PKR transcripts were more abundant in skeletal muscle than transcripts for other dsRNA-binding proteins (Table 1).

Interactions between triplet repeat RNAs and p20, the bacterially expressed RNA-binding domain of PKR, were shown in protection assays, where relatively high concentrations of p20 were able to suppress the digestion of $(CUG)_{69}$ by RNAse T1 (Fig. 4B). More detailed examination of the interaction was conducted by gel mobility shift analysis. Figure 5A shows that the $(CUG)_{69}$ transcript was retarded by p20 at a concentration of 50 μg/mL, whereas transcripts containing 10 or 20 repeats were unaffected. At this concentration, the $(CUG)_{35}$ transcripts were partially shifted to a slower mobility, confirming that p20 has a greater affinity for longer CUG repeat transcripts. More detailed analysis indicated that the threshold length for p20 interactions lies between 10 and 15 repeats: $(CUG)_{20}$ was shifted at high concentrations of p20, as was $(CUG)_{15}$ to a lesser extent, but $(CUG)_{10}$ and $(CUG)_5$ were not discernibly retarded (Fig. 5B). The behavior of transcripts containing 20–140 CUG repeats toward increasing p20 concentrations demonstrates the progressively increasing affinity of p20 for longer repeat transcripts (Fig. 5C). Furthermore, this experiment shows that increasing concentrations of p20 result in complexes of diminishing mobility, which suggests the binding of additional molecules of p20 to long CUG repeats. These findings are to be compared with the packing density of $\sim$11 bp for p20 binding to regular dsRNA duplexes (Manche et al., 1992). p20 also binds to CAG repeat RNAs to a limited extent, but with lower affinity than the $(CUG)_{69}$ and $(CUG)_{140}$ transcripts. RNAs containing 21 or 56 CAG repeats bound with an affinity similar to that of $(CUG)_{20}$ (data not shown). CUC repeat RNA did not bind to p20, as expected for an unstructured RNA. For $(CUG)_{140}$, the longest CUG repeat RNA tested, the binding affinity is bracketed by that of two naturally occurring viral RNAs that are inhibitors of PKR activation (Fig. 5D). The affinity is higher than that of HIV-1 TAR RNA and comparable to adenovirus-2 VA RNA_I ($K_d = 3.5 \times 10^{-7}$ M (Schmedt et al., 1995)), suggesting that the binding of CUG repeat RNA to p20 may be biologically significant.

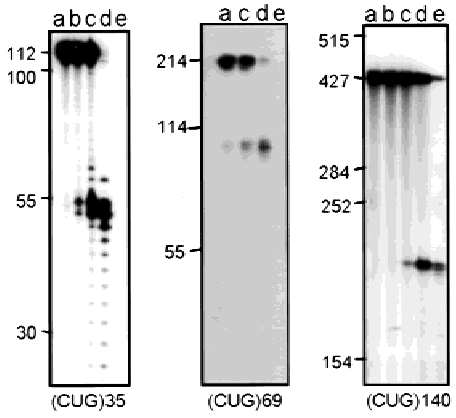

FIGURE 3. Nuclease mapping of CUG repeat RNAs. (CUG)35, 69, and 140 transcripts were digested with ribonuclease T1 (0 U for lane a, 0.01 U for lane b, 0.1 U for lane c, 1 U for lane d, and 10 U for lane e), and then loaded on denaturing gels. Markers indicate the size of the original CUG repeat transcripts or RNA standards synthesized from the polylinker of pBSK.

PKR is subject to both positive and negative regulation by RNA molecules (reviewed by Mathews & Shenk, 1991; Clemens, 1996). Extended perfect duplexes, such as those present in reovirus RNA, activate the kinase, whereas certain highly structured RNAs that are not perfectly duplexed, such as VA RNA_I, prevent activation. To determine the nature of the interactions with CUG repeat RNAs, in vitro kinase assays were conducted using PKR purified from interferon-treated human 293 cells. PKR was strongly activated by $(CUG)_{140}$ and $(CUG)_{69}$ as evidenced by autophosphorylation (Fig. 6) or by phosphorylation of its substrate, eIF2α (data not shown). Weak activation occurred with $(CUG)_{35}$ but not detectably with $(CUG)_{20}$ or $(CUG)_{15}$. Thus, PKR was activated by pathologically-expanded CUG repeats but not by repeat lengths in the normal range for the DM locus.

DISCUSSION

The optical melting studies indicate that CUG repeats form highly stable secondary structures with melting points of $\sim$75 °C, irrespective of their lengths. Based on optical melting studies of $(CTG)_{10}$ and $(CTG)_{30}$ ssDNA hairpins, which showed a similar constancy of melting

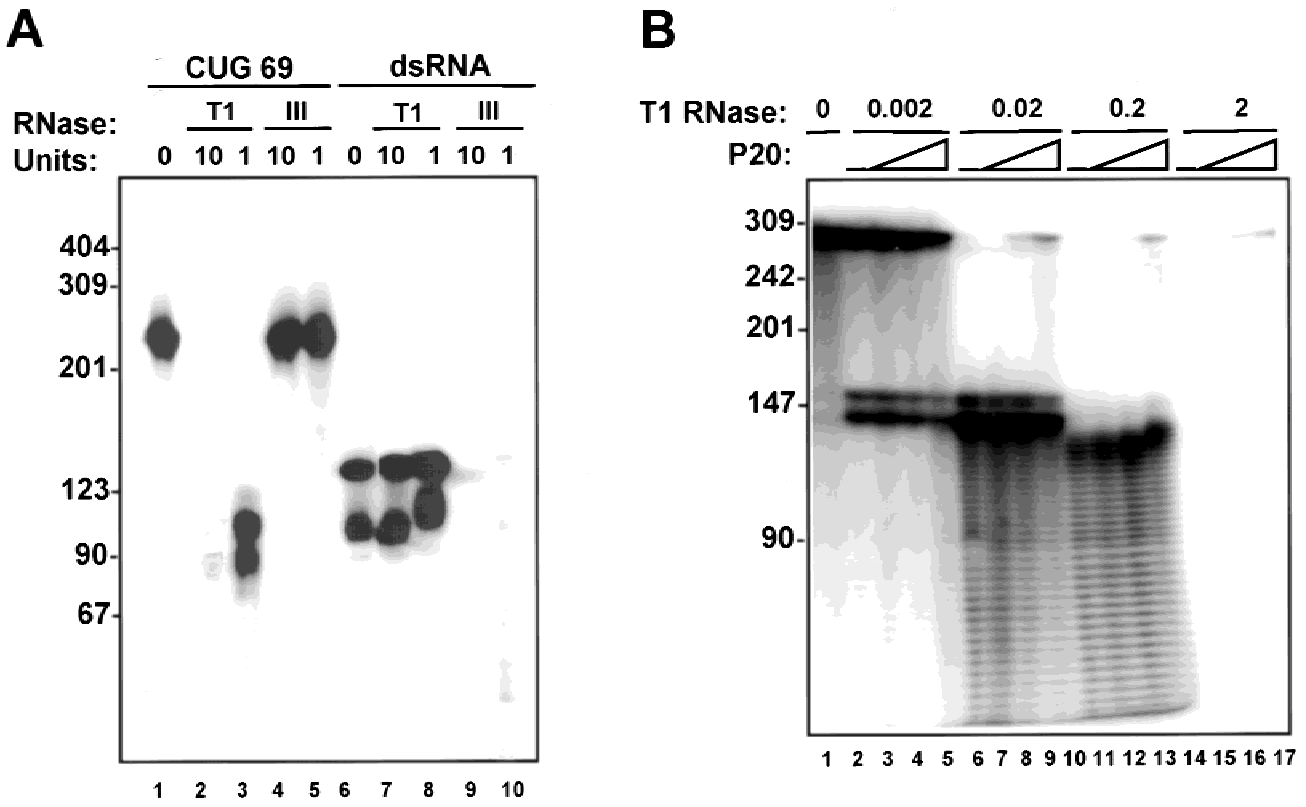

FIGURE 4. CUG repeat RNA is resistant to RNase III and partially protected by p20 from RNase T1 digestion. **A**: Double gel-purified $(CUG)_{69}$ and 145-bp dsRNA were incubated with T1 RNase or RNase III, then analyzed by 8 M urea/5% polyacrylamide gel electrophoresis. Lanes 1 and 6: incubated without RNase; lanes 2 and 7: 0.2 U/μL RNase T1; lanes 3 and 8: 0.02 U/μL RNase T1; lanes 4 and 9: 10 units RNase III; lanes 5 and 10, 1 U RNase III. **B**: $(CUG)_{100}$ was incubated with p20 then digested with T1 RNase at the concentrations shown (in units/μL). The p20 concentrations were: 0 μg/mL for lanes 1, 2, 6, 10, and 14; 1.2 μg/mL for lanes 3, 7, 11, and 15; 12 μg/mL for lanes 4, 8, 12, and 16; and 120 μg/mL for lanes 5, 9, 13, and 17.

temperature, it was proposed that the $(CTG)_{30}$ oligomer formed a series of smaller loops (Petruska et al., 1998). Our nuclease-mapping experiments, however, indicate that transcripts with up to 140 CUG repeats preferentially form a single-loop hairpin with an extended stem. The same result was obtained with transcripts that were refolded after chemical (Fig. 3) or thermal (Fig. 4) denaturation. The constancy in melting temperature for CUG repeats of different lengths is presumably due to the relatively small contribution of the loop to the total H° and S° for folding. These studies confirm the earlier nuclease-mapping studies of nonexpanded CUG repeats (Napierala & Krzyzosiak, 1997), and extend the findings into the range of repeat sizes that cause DM and SCA8. Our data suggest, therefore, that the transition from wild-type to disease allele ($>$50 repeats at the DM locus, and $\gtrsim$90 repeats at the SCA8 locus) is related to increasing length of the hairpin stem, rather than the adoption of a fundamentally different structure.

PKR binding and activation studies support the idea that CUG repeats form hairpins. PKR binds RNA via its

TABLE 1. Relative abundance of transcripts that encode dsRNA-binding proteins in normal skeletal muscle: number of occurrences in a database of 110,000 SAGE tags.

PKR	46
SON	6
ADAR1	3
ADAR2	2
RNA helicase A	1
PACT	1
K12h4.8 putative helicase	1
NF90/ILF3/MMP4/DRBP76	0
TRBP	0
2–5A synthetases	0
STAUFFEN	0

SAGE analysis (Velculescu et al, 1995) was carried out on pooled muscle mRNAs from 16 healthy adult subjects. The total counts for ADAR1, ADAR2, 2–5A synthetases, and TRBP include all isoforms deposited in GenBank (release #112). The tally could overestimate the abundance of a particular transcript, in the event that its SAGE tag is not unique (i.e., another transcript shares the same 14-nt sequence at the first *Nla*III restriction site upstream from the polyadenylation site). Note that transcripts that do not appear once in the database could still be significantly expressed in muscle (the 95% upper confidence bound for transcripts that were not detected is $\sim$4 per myonucleus, based on the Poisson distribution and an assumption of 150,000 transcripts per nucleus).

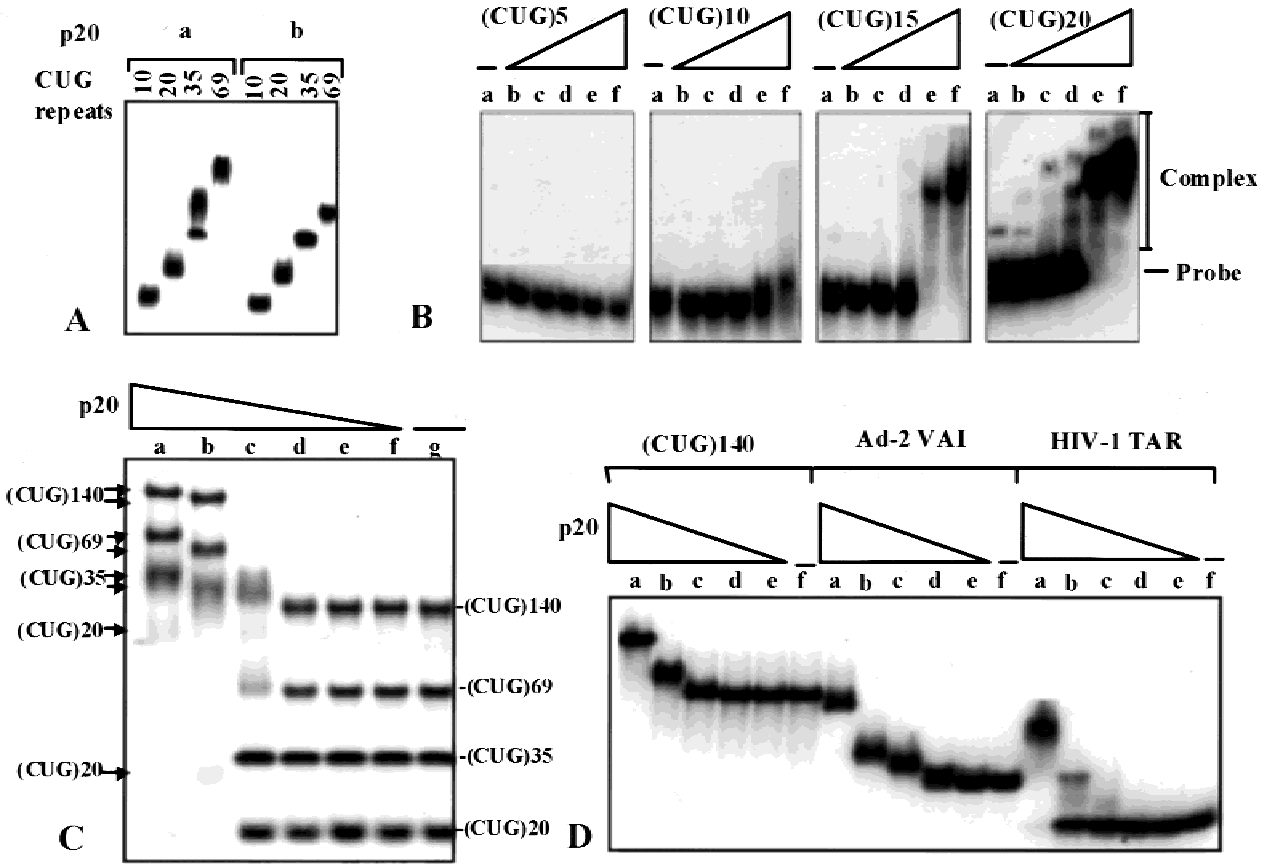

FIGURE 5. Mobility shift assays for purified triplet repeat RNAs using p20. **A**: Purified (CUG)10, 20, 35, and 69 repeat RNAs with same concentration (mg/mL) were incubated (a) with 0.05 mg/mL p20, or (b) without p20, and loaded on native gels. **B**: (CUG)5,10,15, and 20 repeat RNAs with same concentration (mol/mL) were incubated with different concentrations of p20 protein (a: 0 mg/mL, b: 0.01 mg/mL, c: 0.05 mg/mL, d: 0.1 mg/mL, e: 0.5 mg/mL, and f: 1 mg/mL), and loaded on native gels. **C**: (CUG)20,35,69, and 140 repeat RNAs were mixed together at the same concentration (mg/mL), then incubated with various concentrations of p20 protein (a: 0.5 mg/mL, b: 0.25 mg/mL, c: 0.05 mg/mL, d: 0.025 mg/mL, e: 0.005 mg/mL, f: 0.0025 mg/mL, and g: 0 mg/mL), and loaded on a native gel. **D**: (CUG)140, VA RNA_I from Adenovirus-2, and TAR RNA from HIV-, all purified by the same method and at same concentration, were incubated with different concentrations of p20 protein (a: 0.5 mg/mL, b: 0.1 mg/mL, c: 0.05 mg/mL, d: 0.01 mg/mL, e: 0.005 mg/mL, f: 0 mg/mL), and loaded on a native gel.

two dsRNA binding motifs (dsRBMs), both of which are required for efficient binding to dsRNA as well as to highly structured ssRNAs (Schmedt et al., 1995). The interactions between PKR and CUG repeat RNAs can be compared to their interactions with regular dsRNAs of known sizes. Although p20 can pack onto dsRNA as closely as ~11 bp (Manche et al., 1992; Schmedt et al., 1995), the threshold for binding is slightly higher, about 15–20 bp (Manche et al., 1992; Bevilacqua & Cech, 1996). Similarly, the minimum length of CUG repeat RNA that gives detectable binding to p20 in gel shift assays (Fig. 5) is 45 nt (15 repeats). Kinase activation appears to require PKR dimerization and intermolecular autophosphorylation (reviewed by Robertson et al., 1996). Accordingly, the minimum dsRNA chain length for activation is ~30 bp, and the efficiency of activation rises with increasing duplex size up to a maximum at ~85 bp (Manche et al., 1992). With CUG repeat RNA, the shortest polymers that activated PKR were 105 nt (35 repeats), and longer versions were increasingly effective (Fig. 6). These findings demonstrate a greater tolerance by PKR for mismatches in an RNA activator than has been observed previously. CUG hairpins with U•U mismatches at every third base pair are able to

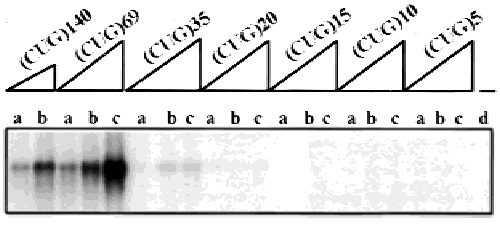

FIGURE 6. PKR activation by purified CUG repeat RNAs. Autoradiogram showing autophosphorylation of PKR after incubation with varying concentrations of the indicated RNAs (a: 0.2 μg/mL, b: 2 μg/mL, c: 20 μg/mL, d: no RNA).

activate PKR, whereas dsRNAs with guanosine•inosine mismatches at every eighth base pair, on average, were unable to activate in an earlier study (Minks et al., 1979).

PKR and RNase III are both dsRBM-containing enzymes that are specific for dsRNAs, although not to the exclusion of mismatched duplexes (Hunter et al., 1975; Robertson & Dunn, 1975; Minks et al., 1979). Strikingly, whereas CUG repeat hairpins bind and activate PKR, they are not cleaved by RNase III. It will be interesting to define the determinants responsible for this discrimination. One other imperfectly duplexed RNA has been shown to bind and activate PKR even though it is insensitive to RNase III (Robertson et al., 1996). Hepatitis delta RNA is highly structured but, like the CUG hairpin, it contains only short stretches of uninterrupted Watson–Crick base pairs (Circle et al., 1997).

The high stability of the CUG repeat hairpin suggests that expanded CUG repeats may form extended hairpins in vivo. Indeed, duplex formation may be the trigger for nuclear retention of mutant DMPK transcripts, as nuclear retention has been observed for dsRNA viral transcripts (Kumar & Carmichael, 1997), and some dsRNA-binding proteins are strongly associated with the nuclear matrix (Wreschner et al., 1985; Schroder et al., 1988). Furthermore, the lack of adenosine residues in CUG repeats may confer resistance to enzymes, such as the dsRNA-activated adenosine deaminases in combination with I-RNAse (Scadden & Smith, 1997), that target long duplex RNAs for rapid degradation. This unusual sequence characteristic may allow expanded CUG repeats to accumulate to higher levels in the nucleus than would be permitted for other duplex RNAs. In accord with this idea, in situ hybridization studies reveal that mutant DMPK transcripts accumulate in hundreds of foci per DM myonucleus (Davis et al., 1997), which is higher than the frequency of $\sim$8 transcripts per myonucleus that would be predicted from the normal abundance of DMPK mRNA (based on 11 occurrences of DMPK in a database of 110,000 muscle SAGE tags, an assumption that both alleles are transcriptionally active, and an estimate of 150,000 transcripts per myonucleus; S. Welle, C. Thornton, unpubl.)

Long duplex RNAs elicit potent biologic effects. In plants and invertebrates, dsRNAs provoke posttranscriptional gene silencing in a gene-specific, homology-dependent fashion. A different effect is observed in mammalian cells, where long duplex RNAs trigger a potent cellular stress response, independent of their sequence (reviewed by Haines et al., 1991; Jacobs & Langland, 1996). Natural activators of the dsRNA response include viral dsRNAs (genomes, replication intermediates, or products of bidirectional transcription). Many downstream effects of the dsRNA response, such as inhibition of translation initiation, accelerated mRNA and rRNA degradation, and apoptosis, can be understood as an attempt by host cells to prevent or slow viral replication. It is uncertain whether this response has functions unrelated to its antiviral effects, or whether it is activated by any endogenous transcripts. Although the conformation of expanded CUG repeats in vivo remains uncertain, we have recently observed that expression of an untranslated, expanded CTG repeat in transgenic mice can reproduce many of the muscle manifestations of DM, supporting the idea that the pathogenic effect of the DM mutation is mediated by mutant transcripts (C. Thornton, unpubl.)

Our results suggest that dsRNA-binding proteins are candidates for involvement in the nuclear retention or toxicity of expanded CUG repeats. PKR is an interesting candidate because: (1) it is significantly expressed in skeletal muscle (Table 1); (2) it may play a role in the regulation of muscle differentiation (Kronfeld-Kinar et al., 1999); (3) a significant fraction of PKR is localized in the nucleus (Jimenez-Garcia et al., 1993; Jeffrey et al., 1995); and (4) the length threshold for PKR activation by CUG repeats coincides with the threshold for causing DM by expanded CTG repeats (Fig. 6). Furthermore, activation of PKR inhibits protein synthesis, through its ability to phosphorylate the translation-initiation factor eIF2α (reviewed by Clemens, 1996; Mathews, 1996). It is noteworthy that the rate of muscle protein synthesis in vivo, derived from measurements of ^{13}C-leucine incorporation into muscle protein, is reduced in DM patients compared to healthy subjects (Griggs et al., 1990) or patients with other forms of muscular dystrophy (R. Tawil, pers. comm.).

Alternatively, the toxicity of expanded CUG repeats may involve sequestration, rather than activation, of dsRNA-binding proteins. Instability of the DM mutation in somatic cells can produce very large expansions of the CTG repeat. The expanded CUG repeats in muscle, heart, and brain range up to 30-fold longer than the longest transcript studied here, because they are generated from alleles with up to 13 kb of CTG repeats (Thornton et al., 1994). The extreme length of the expanded repeat and its focal accumulation in myonuclei (Taneja et al., 1995) may generate local concentrations that are high enough to act as a sink for binding proteins. The affinity of expanded CUG repeats for dsRNA-binding proteins appears adequate to support such interactions, as the binding affinity for PKR is similar to that of VA RNA$_I$, an adenovirus-2 transcript that inhibits PKR activation when it accumulates to high levels late in infection (reviewed by Mathews & Shenk, 1991). In view of the sequence nonselectivity that characterizes dsRNA interactions with dsRNA-binding proteins, and the presence in nuclear extracts of several proteins that bind preferentially to expanded CUG repeats (B. Tian & M.B. Mathews, data not shown), it is possible that several different proteins are susceptible to this effect. Efforts to identify other proteins that bind to expanded CUG repeats and to develop model systems to study gain of function by these transcripts are currently underway.

MATERIALS AND METHODS

Plasmid templates for synthesis of triplet repeat RNAs

Plasmid pBCE was constructed by using oligonucleotides to insert *Bsm*I and *Bam*HI restriction sites bracketed by two *Bsm*F1 sites at the *Bam*HI-*Eco*RI site of pBSKII. Oligonucleotides were used to insert a modified T7 RNA polymerase promoter (C substituted for A at the +4 position) at the *Bsm*F1-*Hin*dIII site of pBCE, creating plasmid pCNA. Derivatives of pCNA with 5, 10, 15, 20, 25, and 35 CTG repeats were made by cycles of $(CTG)_5$ linker addition at the *Bsm*F1 restriction site. A PCR product with a larger CTG repeat was obtained by amplification of the expanded repeat from a patient with mild DM as described (Thornton et al., 1994). The *Bsm*I and *Dpn*II fragment containing the expanded repeat was cloned at the *Bsm*I-*Bam*HI site of pBCE. The resulting plasmid, $pBCE_{74}$, contains an insert with 74 uninterrupted CTG repeats. *Bsm*F1 digestion of this plasmid releases a CTG repeat fragment. Nonpalindromic overhangs orient this fragment for ligation into concatamers of uninterrupted CTG repeats, or directional ligation into pCNA. This fragment and its dimer were inserted at the *Bsm*F1 site of pCNA to create plasmids $pCNA_{69}$ and $pCNA_{140}$. These plasmids contain 69 and 140 uninterrupted CTG repeats, respectively, fused directly to the T7 promoter. $pCNA_{100}$ is a subclone of $pCNA_{140}$ that underwent a deletion within the repeat tract.

Plasmid $pTQA_{52}$ with 52 CTC repeats was made by using oligonucleotides to insert an *Msc*I site fused to a modified T7 promoter (GGCGCT in the +1 to +6 positions) at the *Sac*I-*Pst*I site in pBSKII. Oligonucleotides 5′-$(CCT)_9$CC and 5′-$(GGA)_9$GG were used to synthesize expanded CTC repeat fragments by slipped strand replication in a thermal cycler as described (Ordway & Detloff, 1961). These PCR products were ligated into the *Msc*I site of pTQA. By a similar strategy, plasmids $pASU_{29}$ and $pASU_{56}$ were constructed with 29 and 56 uninterrupted CAG repeats fused to the wild-type T7 promoter.

The stability of triplet repeat plasmids was optimized by orienting the CTG repeat on the leading strand of DNA replication, cloning in strain HB101, growing cultures at 30 °C, and harvesting cultures before they reached stationary phase (Bowater et al., 1996). Despite these measures, plasmids with more than 50 CTG repeats showed minor instability, in that subclones often differed from the parent plasmid by ±1 to 5 repeats. All triplet repeat inserts were sequenced by dideoxy cycle sequencing.

RNA synthesis and purification

Plasmids in the pCNA, pASU, pTQA series were linearized at the final triplet repeat by digestion with *BsmF*1, *Bbs*I, or *Eco*47III, respectively. Linearized plasmids were extracted with phenol/chloroform and precipitated in ethanol. RNA for gel shift assays, PKR activation experiments, and nuclease digests in Figure 4 were synthesized using T7 RNA polymerase as described (Mellits et al., 1990), except that the nucleotide concentrations were modified. Nucleotides used for synthesis of CUG and CUC repeat RNA were 6 mM GTP and CTP, 12 μM UTP, and 1 mCi/mL [α-^{32}P]UTP. For reactions labeled with CTP, 6 mM GTP and UTP, 12 μM CTP, and 1 mCi/mL [α-^{32}P]CTP were used instead. ATP was omitted from transcription reactions for CUG and CUC repeats, and UTP was omitted from transcription reactions for CAG repeats. Transcription reactions were treated with 20 mg/mL DNaseI and 1 mM $CaCl_2$ for 30 min at 37 °C, extracted with phenol/chloroform, and precipitated with ethanol. The transcripts were purified sequentially on denaturing, followed by nondenaturing polyacrylamide gels as described (Mellits et al., 1990).

RNA for T1 digests in Figure 3 was synthesized with T7 RNA polymerase using unlabeled nucleotides, treated with shrimp alkaline phosphatase, labeled at the 5′ end with [γ^{-32}P]ATP using T4 polynucleotide kinase, and then purified on 40 cm 5 or 8% polyacrylamide/6 M urea gels. RNA was eluted from excised gel bands overnight at 4 °C in 0.2% SDS, 0.5 M ammonium acetate, pH 6, and then precipitated and washed in ethanol. For optical melting experiments, a fragment containing the triplet repeat and T7 promoter was first isolated from each plasmid. RNA was synthesized using T7 RNA polymerase, treated with DNAse I for 30 min at 37 °C, extracted with phenol/chloroform, passed through a Sepharose cartridge to remove unincorporated nucleotides, precipitated with ammonium acetate and ethanol, and then purified on 18 cm 6% polyacrylamide/6 M urea gels with UV shadowing. Purity of these transcripts was verified by end labeling with [γ^{-32}P]ATP, followed by analysis on sequencing gels.

Melting curves

Ethanol-precipitated RNA was resuspended in water and then adjusted to give a final concentration of 5 mM $MgCl_2$, 150 mM KCl, 10 mM PIPES, pH 7.0, and 0.5 mM EDTA. Absorbance versus temperature-melting curves were measured at 280 nm with a heating rate of 0.5 °C/min from 10 to 97 °C on a Gilford 250 spectrometer controlled with a Gilford 2527 thermoprogrammer. Melting temperatures were calculated as described (Serra et al., 1994).

Nuclease digests

For Figure 3, 5′ end-labeled RNA was dissolved in 10 mM Tris, pH 7.0, at a concentration of 1,500 cpm/μL. Nuclease digests were prepared on ice as 5-μL reactions containing 3 μL RNA and various concentrations of RNAse T1 in 12.5 mM Tris, pH 7.5, 5 mM $MgCl_2$, and 75 mM KCl. Digests were incubated at 37 °C for 45 min, and then loaded directly on 40-cm 5% polyacrylamide denaturing gels along with internally labeled RNA size standards. Gels were dried and subjected to autoradiography. For Figure 4, digestion with RNase III (kindly provided by H.D. Robertson, Cornell Medical School) was conducted in 30 mM Tris, pH 8.0, 100 mM NaCl, 5 mM $MgCl_2$, 0.1 mM EDTA, and 0.02 mM DTT; and digestion with RNase T1 was conducted in 10 mM Tris, pH 7.4, 150 mM NaCl, and 1 mM EDTA. Incubation was for 45 min at 37 °C. For p20 protection analysis, RNA was first incubated with p20 on ice for 20 min in the buffer used for mobility shift assays, and then incubated with RNase T1 for 45 min at 37 °C.

Frequency of transcripts that encode dsRNA-binding proteins

mRNA was isolated from normal adult human skeletal muscle biopsies, pooled, and analyzed by the SAGE technique (Velculescu et al., 1995) as described (Welle et al., 1999). Half of the tags (54,791) came from muscle of young men (21–42 years old, $n = 8$) and half (54,802) from muscle of older men (66–77 years old, $n = 8$). The number of occurrences for transcripts that encode each dsRNA binding protein was counted in the database of 110,000 SAGE tags.

Mobility shift assay

The *Escherichia coli*-expressed dsRNA binding domain of PKR, p20, was purified to apparent homogeneity and used in gel shift assays as described (Schmedt et al., 1995). p20 was incubated with in vitro-transcribed RNA and the other components indicated for 20 min at 4 °C and resolved on a 5%, 40 mM Tris-glycine polyacrylamide gel run at 4 °C.

Kinase assay

PKR from interferon-treated human 293 cells was partially purified to the Mono S stage as described (Kostura & Mathews, 1989). Kinase reactions (10 μL) were carried out with 0.5 μL PKR (about 5 ng) and 2.5 μCi of [γ^{32}P]ATP (ICN Radiochemicals) as described (Manche et al., 1992). Reactions were resolved on 10% polyacrylamide/SDS gels, fixed, dried, and subjected to autoradiography.

ACKNOWLEDGMENTS

This work was supported by the Muscular Dystrophy Association, grant AI34552 from the National Institutes of Health to MBM, a predoctoral fellowship 9810005T from the American Heart Association to BT, and the Wayne C. Gorrell Jr. Molecular Biology Laboratory (RJW and CAT). We thank Carolyn McClain for technical assistance. Charles Thornton is a Paul Beeson Jr. Physician Faculty Scholar of the American Federation for Aging Research.

Received July 14, 1999; returned for revision August 15, 1999; revised manuscript received October 22, 1999

REFERENCES

Bevilacqua PC, Cech TR. 1996. Minor-groove recognition of double-stranded RNA by the double-stranded RNA-binding domain from the RNA-activated protein kinase PKR. *Biochemistry 35*:9983–9994.

Bowater RP, Rosche WA, Jaworski A, Sinden RR, Wells RD. 1996. Relationship between *Escherichia coli* growth and deletions of CTG•CAG triplet repeats in plasmids. *J Mol Biol 264*:82–96.

Brook JD, McCurrach ME, Harley HG, Buckler AJ, Church D, Aburatani H, Hunter K, Stanton VP, Thirion JP, Hudson T, Sohn R, Zemelman B, Snell RG, Rundle SA, Crow S, Davies J, Shelbourne P, Buxton J, Jones C, Juvonen V, Johnson K, Harper PS, Shaw DJ, Housman DE. 1992. Molecular basis of myotonic dystrophy: Expansion of a trinucleotide (CTG) repeat at the 3′ end of a transcript encoding a protein kinase family member. *Cell 68*: 799–808.

Cazenave C, Uhlenbeck OC. 1994. RNA template-directed RNA synthesis by T7 RNA polymerase. *Proc Natl Acad Sci USA 91*:6972–6976.

Circle DA, Neel OD, Robertson HD, Clarke PA, Mathews MB. 1997. Surprising specificity of PKR binding to delta agent genomic RNA. *RNA 3*:438–448.

Clemens MJ. 1996. Protein kinases that phosphorylate eIF2 and eIF2B, and their role in eukaryotic cell translational control. In: Hershey JWB, Mathews MB, Sonenberg N, eds. *Translational control*. Cold Spring Harbor, New York: Cold Spring Harbor Press. pp 139–172.

Davis BM, McCurrach ME, Taneja KL, Singer RH, Housman DE. 1997. Expansion of a CUG trinucleotide repeat in the 3′ untranslated region of myotonic dystrophy protein kinase transcripts results in nuclear retention of transcripts. *Proc Natl Acad Sci USA 94*:7388–7393.

Griggs RC, Jozefowicz R, Kingston WJ, Nair KS, Herr BE, Halliday D. 1990. Mechanism of muscle wasting in myotonic dystrophy. *Ann Neurol 27*:505–512.

Haines DS, Strauss KI, Gillespie DH. 1991. Cellular response to double-stranded RNA. *J Cell Biochem 46*:9–20.

Harper PS. 1989. *Myotonic dystrophy*. London: W.B. Saunders Company.

Hunter T, Hunt T, Jackson RJ, Robertson HD. 1975. The characteristics of inhibition of protein synthesis by double-stranded ribonucleic acid in reticulocyte lysates. *J Biol Chem 250*:409–417.

Jacobs BL, Langland JO. 1996. When two strands are better than one: The mediators and modulators of the cellular responses to double-stranded RNA. *Virology 219*:339–349.

Jansen G, Groenen PJTA, Bachner D, Jap PH, Coerwinkel M, Oerlemans F, van den Broek W, Bohlsch B, Pette D, Plomp JJ, Molenaar PC, Nederhoff MGJ, vanEchteld CJ, Dekker J, Berns A, Hameister H, Wieringa B. 1996. Abnormal myotonic dystrophy protein kinase levels produce only mild myopathy in mice. *Nat Genet 13*:316–324.

Jeffrey IW, Kadereit S, Meurs EF, Metzger T, Bachmann M, Schwemmle M, Hovanessian AG, Clemens MJ. 1995. Nuclear localization of the interferon-inducible protein kinase PKR in human cells and transfected mouse cells. *Exp Cell Res 218*:17–27.

Jimenez-Garcia LF, Green SR, Mathews MB, Spector DL. 1993. Organization of the double-stranded RNA-activated protein kinase DAI and virus-associated VA RNAI in adenovirus-2-infected HeLa cells. *J Cell Sci 106*:11–22.

Koob MD, Moseley ML, Schut LJ, Benzow KA, Bird TD, Day JW, Ranum LP. 1999. An untranslated CTG expansion causes a novel form of spinocerebellar ataxia (SCA8). *Nat Genet 21*:379–384.

Kostura M, Mathews MB. 1989. Purification and activation of the double-stranded RNA-dependent eIF-2 kinase DAI. *Mol Cell Biol 9*:1576–1586.

Kronfeld-Kinar Y, Vilchik S, Hyman T, Leibkowicz F, Salzberg S. 1999. Involvement of PKR in the regulation of myogenesis. *Cell Growth Diff 10*:201–212.

Kumar M, Carmichael GG. 1997. Nuclear antisense RNA induces extensive adenosine modifications and nuclear retention of target transcripts. *Proc Natl Acad Sci USA 94*:3542–3547.

Maeda M, Taft CS, Bush EW, Holder E, Bailey WM, Neville H, Perryman, MB, Bies RD. 1995. Identification, tissue-specific expression, and subcellular localization of the 80- and 71-kDa forms of myotonic dystrophy kinase protein. *J Biol Chem 270*:20246–20249.

Manche L, Green SR, Schmedt C, Mathews MB. 1992. Interactions between double-stranded RNA regulators and the protein kinase DAI. *Mol Cell Biol 12*:5238–5248.

Mathews DH, Sabina J, Zuker M, Turner DH. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol 288*:911–940.

Mathews MB. 1996. Interactions between viruses and the cellular machinery for protein synthesis. In: Hershey JWB, Mathews MB, Sonenberg N, eds. *Translational control*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 505–508.

Mathews MB, Shenk T. 1991. Adenovirus virus-associated RNA and translation control. *J Virol 65*:5657–5662.

Mellits KH, Pe'ery T, Manche L, Robertson HD, Mathews MB. 1990. Removal of double-stranded contaminants from RNA transcripts: Synthesis of adenovirus VA RNAI from a T7 vector. *Nucleic Acids Res 18*:5401–5406.

Minks MA, West DK, Benvin S, Baglioni C. 1979. Structural requirements of double-stranded RNA for the activation of 2′,5′-oligo(A) polymerase and protein kinase of interferon-treated HeLa cells. *J Biol Chem 254*:10180–10183.

Napierala M, Krzyzosiak WJ. 1997. CUG repeats present in myotonin kinase RNA form metastable slippery hairpins. *J Biol Chem 272*:31079–31085.

Ordway JM, Detloff PJ. 1961. In vitro synthesis and cloning of long CAG repeats. *Biotechniques 21*:609–610.

Petruska J, Hartenstine MJ, Goodman MF. 1998. Analysis of strand slippage in DNA polymerase expansions of CAG/CTG triplet repeats associated with neurodegenerative disease. *J Biol Chem 273*:5204–5210.

Philips AV, Timchenko LT, Cooper TA. 1998. Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. *Science 280*:737–741.

Reddy S, Smith DBJ, Rich MM, Leferovich JM, Reilly P, Davis BM, Tran K, Rayburn H, Bronson R, Cros D, Balice-Gordon RJ, Housman D. 1996. Mice lacking the myotonic dystrophy protein kinase develop a late onset progressive myopathy. *Nat Genet 13*:325–334.

Robertson HD, Dunn JJ. 1975. Ribonucleic acid processing activity of *Escherichia coli* ribonuclease III. *J Biol Chem 250*:3050–3056.

Robertson HD, Manche L, Mathews MB. 1996. Paradoxical interactions between human delta hepatitis agent RNA and the cellular protein kinase PKR. *J Virol 70*:5611–5617.

Scadden AD, Smith CW. 1997. A ribonuclease specific for inosine-containing RNA: A potential role in antiviral defence? *EMBO J 16*:2140–2149.

Schmedt C, Green SR, Manche L, Taylor DR, Ma Y, Mathews MB. 1995. Functional characterization of the RNA-binding domain and motif of the double-stranded RNA-dependent protein kinase DAI (PKR). *J Mol Biol 249*:29–44.

Schroder HC, Wenger R, Rottmann M, Muller WE. 1988. Alteration of nuclear (2′-5′)oligoriboadenylate synthetase and nuclease activities preceding replication of human immunodeficiency virus in H9 cells. *Biol Chem Hoppe-Seyler 369*:985–995.

Serra MJ, Axenson TJ, Turner DH. 1994. A model for the stabilities of RNA hairpins based on a study of the sequence dependence of stability for hairpins of six nucleotides. *Biochemistry 33*:14289–14296.

Taneja KL, McCurrach M, Schalling M, Housman D, Singer RH. 1995. Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *J Cell Biol 128*:995–1002.

Thornton CA, Johnson K, Moxley RT. 1994. Myotonic dystrophy patients have larger CTG expansions in skeletal muscle than in leukocytes. *Ann Neurol 35*:104–107.

Timchenko LT, Miller JW, Timchenko NAXDD, Datar KV, Lin L, Roberts R, Caskey CT, Swanson MS. 1996. Identification of a (CUG)n triplet repeat RNA-binding protein and its expression in myotonic dystrophy. *Nucleic Acids Res 24*:4407–4414.

Triana-Alonso FJ, Dabrowski M, Wadzack J, Nierhaus KH. 1995. Self-coded 3′-extension of run-off transcripts produces aberrant products during in vitro transcription with T7 RNA polymerase. *J Biol Chem* 270:6298–6307.

Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. 1995. Serial analysis of gene expression. *Science 270*:484–487.

Walter AE, Turner DH, Kim J, Lyttle MH, Muller P, Mathews DH, Zuker. 1994. Coaxial stacking of helixes enhances binding of oligoribonucleotides and improves predictions of RNA folding. *Proc Natl Acad Sci USA 91*:9218–9222.

Welle S, Bhatt K, Thornton CA. 1999. Inventory of high-abundance mRNAs in skeletal muscle of normal men. *Genome Res 9*:506–513.

Wreschner DH, Nathanel T, Herzberg M. 1985. Double stranded RNA and the nuclear matrix: Implications for the 2–5A system. *Prog Clin Biol Res 202*:47–66.