3852–3863 *Nucleic Acids Research, 2011, Vol. 39, No. 9* *Published online 18 January 2011*
doi:10.1093/nar/gkq1323

Mutant CAG repeats of Huntingtin transcript fold into hairpins, form nuclear foci and are targets for RNA interference

Mateusz de Mezer, Marzena Wojciechowska, Marek Napierala, Krzysztof Sobczak and Wlodzimierz J. Krzyzosiak*

Laboratory of Cancer Genetics, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland

Received August 20, 2010; Revised December 12, 2010; Accepted December 13, 2010

ABSTRACT

The CAG repeat expansions that occur in translated regions of specific genes can cause human genetic disorders known as polyglutamine (poly-Q)-triggered diseases. Huntington's disease and spinobulbar muscular atrophy (SBMA) are examples of these diseases in which underlying mutations are localized near other trinucleotide repeats in the huntingtin (*HTT*) and androgen receptor (*AR*) genes, respectively. Mutant proteins that contain expanded polyglutamine tracts are well-known triggers of pathogenesis in poly-Q diseases, but a toxic role for mutant transcripts has also been proposed. To gain insight into the structural features of complex triplet repeats of HTT and AR transcripts, we determined their structures *in vitro* and showed the contribution of neighboring repeats to CAG repeat hairpin formation. We also demonstrated that the expanded transcript is retained in the nucleus of human HD fibroblasts and is colocalized with the MBNL1 protein. This suggests that the CAG repeats in the HTT mRNA adopt ds-like RNA conformations *in vivo*. The intracellular structure of the CAG repeat region of mutant HTT transcripts was not sufficiently stable to be protected from cleavage by an siRNA targeting the repeats and the silencing efficiency was higher for the mutant transcript than for its normal counterpart.

INTRODUCTION

Huntington's disease (HD) is a progressive neurodegenerative disorder caused by an autosomal dominant mutation in the huntingtin gene (*HTT*). This is the most prevalent disorder among trinucleotide repeat expansion diseases (TREDs), which are caused by CAG repeat expansions in translated regions of a single causative gene (1–3). Healthy individuals harbor fewer than 35 CAG repeats in the *HTT* locus, which results in the production of the cytoplasmic huntingtin protein. However, HD patients have 40 or more repeated CAG units, which gives rise to the mutant polyglutamine (poly-Q) protein that has pathogenic properties. Another poly-Q-dependent TRED is spinal and bulbar muscular atrophy (SBMA), also known as Kennedy's disease, which is triggered by exonic CAG expansions in the androgen receptor gene (*AR*). In SBMA patients the number of CAG repeats range from 38 to 62 in the *AR* locus, whereas healthy individuals have fewer than 36 repeats (4).

Long polyglutamine tracts in CAG expansion-dependent diseases function as 'sinks' that sequester other proteins (5). Early *in vitro* evidence demonstrated the accumulation of insoluble poly-Q protein inclusions in the neurons of patients with these diseases and this was considered a cause of neuronal damage (6). However, other reports have shown that such inclusions may be beneficial rather than deleterious to cells, serving as deposits of dysfunctional proteins that prevent these proteins from engaging in pathogenic interactions (7,8).

*To whom correspondence should be addressed. Tel: +48 61 8528503; Fax: +48 61 8520532; Email: wlodkrzy@ibch.poznan.pl
Present Addresses:
Mateusz de Mezer, Laboratory of Functional Genomics, Institute of Plant Genetics, Polish Academy of Sciences, Strzeszynska 34, 60-479 Poznan, Poland.
Marek Napierala, University of Texas M. D. Anderson Cancer Center, Department of Molecular Carcinogenesis and Center for Cancer Epigenetics, 1515 Holcombe Boulevard, Houston, TX 77030, USA.
Krzysztof Sobczak, Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Umultowska 89, 61-614 Poznan, Poland.

© The Author(s) 2011. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.5), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Another group of TREDs is mediated by the expression of mutant expansions in the non-coding region of a gene. Transcription of these repeats appears to be both necessary and sufficient to cause disease (9,10). Myotonic dystrophy type1 (DM1) and type 2 (DM2) as well as fragile X associated tremor ataxia syndrome (FXTAS) belong to this toxic RNA gain-of-function disorder family (11–14). The toxicity of the pathogenic RNA is driven by its propensity to fold into long hairpin structures, which triggers abnormal interaction with dsRNA-binding proteins (double stranded RNA-binding proteins) and results in the loss of their normal function in RNA metabolism. Interestingly, it has been postulated that mutant transcripts may also be 'toxic' to the cell in poly-Q-dependent diseases (15,16).

Over the past decade, our group has studied the structures of triplet repeat regions from a number of transcripts involved in TREDs (17–21). We determined that triplet repeats composed of CNG units adopt hairpin conformations *in vitro* that are either 'slipped', i.e. they show several 'in register' variants of alignment or else are 'frozen' in a single alignment imposed by the interacting flanking sequences (17,18). There are also 'split' hairpins formed by the interrupted repeats of CAG or CGG (19–21). Interestingly, expansion-prone CAG repeats in the *HTT* and *AR* loci are located in close proximity to other triplet repeats. In the case of *HTT* locus the CAG tract is adjacent to the polymorphic $(CCG)_n$ while in the *AR* locus expandable $(CTG)_3(CAG)_n$ complex repeat is in close proximity to the monomorphic $(CAG)_6$ stretch (3,4). Predicted interactions of short triplet repeat tracts with long polymorphic CAG repeats in the HTT and AR transcripts and their influence on the repeat hairpin formation have never been studied.

In this study, we analyzed the *in vitro* RNA structure of the two peculiar triplet repeat regions present in the HTT and AR transcripts. We further performed *in vivo* analysis of mutant HTT transcripts in human HD cells and found that RNAs containing CAG repeat expansions were accumulated in ribonuclear foci that colocalized with the MBNL1 (muscleblind-like 1) splicing factor. Despite the presence of typical toxic RNA features in HD cells driven by expanded CAG hairpin structures, a CAG repeat-targeting siRNA exhibited a greater silencing of the mutant HTT transcript than of the normal transcript.

MATERIALS AND METHODS

Preparation of RNA fragments of HTT and AR transcripts

The $(CAG)_{10}(CCG)_{10}$ and $(CAG)_{20}$ oligoribonucleotides were purchased from Metabion (Martinsried, Germany). Fragments of natural HTT and AR mRNAs were obtained by *in vitro* transcription of cDNA fragments obtained by PCR amplification. These HTT and AR fragments were selected based on structure prediction for RNAs differing in the length of the repeat flanking sequences using Mfold program (22,23). Results of the molecular modeling demonstrated that the structures formed by the repeats and their nearest flanks are autonomous and not influenced by long-range interactions. The template for HTT transcripts containing both the CAG and the CCG repeat tracts was amplified with Had1 T7 forward and Hd2 reverse primers, and the template for the AR transcripts was obtained with AR1 T7 forward and AR2 reverse primers (Table 1). Both forward primers contained the T7 RNA polymerase promoter *in vitro* transcription was carried out with T7 RNA polymerase (Epicenter) as described previously (24). RNA products were purified by denaturing PAGE and were 5′-end-labeled with [γ-^{32}P] ATP using T4 polynucleotide kinase (Epicenter).

Table 1. Oligonucleotides used in this study

Oligonucleotide name	Sequence 5′→3′
Had1 T7	TAATACGACTCACTATAGGACGCCTGGA AAGCTGATGA
Hd2	CGGCTGAGGAAGCTGAGGA
AR1 T7	TAATACGACTCACTATAGGCGCGAAGT GATCCAGAA
AR2	CTTGGGGAGAACCATCCTCT
GAPDH	F: GAAGGTGAAGGTCGGAGTC R: GAAGATGGTGATGGGATTTC
HTT	F: CCCTGGAAAAGCTGATGAAG R: CACGGTCTTTCTTGGTAGCTG
ATXN3	F: GGAAGAGACGAGAAGCCTAC; R: TCACCTAGATCACTCCCAAGT
ATXN7	F: GAGCGGAAAGAATGTCGG R: CGTCCTTCCCAGGAAGTTTG
ATN1	F: AGGCTGAGGCTACCTGTGAA R: CAGAGCCTGTGGGAAGTCTC
5′-ETS	F: GTCCCCTCGTCTCTCCTCTC R: ACAGCGAGGGCTGTCTGC
HTT pre-mRNA	F: CTGCACCGACCGTGAGTT R: GGTTGCTGGGTCACTCTGTC
siHD 5′	ACUUGAGGGACUCGAAGGCCU[a]
siCUG7/CAG7	CUGCUGCUGCUGCUGCUGCUGC[a]
siHD sep	UGGCGGCUGUGGCUGCUGCdTdT[a]
siHD 3′	UGAGGAAGCUGAGGAGGCGGC[a]

[a]Ribonucleotide sequences of antisense strand of siRNA.
dT, deoxyribotymidine.

RNA structure probing

Prior to structure probing reactions, gel purified 5′-end-labeled RNAs were subjected to denaturation and renaturation in a solution containing 20 mM Tris–HCl (pH 7.2), 80 mM NaCl and 2 mM $MgCl_2$ by heating at 90°C for 1 min and slowly cooling to 37°C. Incomplete RNA digestion was initiated by mixing 5 µl of RNA sample (20–60 nM; 50 000 cpm) and 5 µl of probe solution containing one of the four nucleases (S1, T1, T2 or V1). For the S1 nuclease reaction, 1 mM $ZnCl_2$ was present. In single-stranded RNA structures, T1 and T2 RNases exclusively cleave after G and A residues, respectively, whereas S1 nuclease cleaves after all residues (25,26) with some preference for Gs (27,28). In addition,

ribonuclease V1 was used, which is specific for double-stranded regions of RNA (25) and exhibits a preference for U (28,29). All nuclease reactions were performed at 37°C for 10 min and were terminated by adding an equal volume of stopping solution (7.5 M urea and 20 mM EDTA with dyes) and freezing. Nuclease digestion products were separated by short (12% gel), medium (10%) and long (6%) denaturing electrophoresis, and these were run in parallel with products of alkaline hydrolysis (1-nt ladder) and limited T1 nuclease digestion (G-specific ladder) of the transcript to visualize RNA cleavage occurring in different sections of the studied molecules at high resolution. The denaturing gel contained 7.5 M urea, 90 mM Tris–borate buffer and 2 mM EDTA. Products of the RNA structure probing were visualized by autoradiography and Phosphorimaging (Typhoon, Molecular Dynamics). The intensity of cleavages induced by the different probes is proportional to the strength of radioactive signals from cleavage products which were measured and categorized into three groups: strong, medium and weak.

Cell culture and siRNA transfection

Human fibroblast cell lines from a normal control patient (GM07492) as well as heterozygous HD (GM04208–21/44 CAG, GM04281–17/68 CAG, GM03864–15/46 CAG and GM09197–21/151 CAG), DRPLA (GM13717–15/65 CAG), SCA3 (GM06153–18/69 CAG), SCA7 (GM03561–8/62 CAG) and DM1 patients (GM03987–13/500 CTG) were purchased from Coriell Cell Repositories (NJ, USA). Fibroblast cultures were maintained in MEM medium (Lonza) containing 10% FBS, 1× antibiotic antimycotic solution (Sigma) and 1× MEM non-essential amino acid solution (Sigma). For RNA fluorescence *in situ* hybridization (FISH) and immunofluorescence (IF), cells were grown on cover slips pre-coated with 0.05% gelatin in 1× PBS. Prior to transfection, siRNA duplexes were prepared as follows: 100 µM of single-stranded siRNAs (Table 1) (Metabion) were prepared at 20 µM in annealing buffer (Metabion). Then, siRNAs were denatured for 1 min at 90°C followed by slow cooling at room temperature for 40 min. The siRNAs were then diluted in serum and antibiotic-free Opti-MEM (Invitrogen) and mixed with the transfection reagent, Lipofectamine 2000 (Invitrogen). The mixture was added to the cells grown in a 6-well plate and incubated at 37°C for 4 h. Fluorescent siRNAs (BLOCK-iT Florescent Oligo, Invitrogen) were used to monitor transfection efficiency.

Total RNA isolation and semi-quantitative RT–PCR

Total RNA was extracted from cells with TRIzol (Invitrogen). Equal amounts of RNA (1 µg) were reverse-transcribed with random primers (Promega) using SuperScript II Reverse Transcriptase (Invitrogen) according to the manufacturer's protocol. Oligonucleotides and PCR conditions used for RT–PCR of human HTT, ATXN3, ATXN7 and ATN1 cDNA are shown in Tables 1 and 2, respectively. Each sample used was assayed in three independent RT reactions and normalized to *GAPDH* expression levels. PCR products were visualized by 1.3% agarose gel electrophoresis and quantified using Gel-Pro 3.1.

Table 2. RT–PCR conditions applied for the indicated genes

Gene name	RT–PCR conditions (number of cycles; annealing temperature)
GAPDH	18; 60°C
HTT	30; 58°C
ATXN3	26; 58°C
ATXN7	28; 45°C
ATN1	27; 55°C
5′-ETS	19; 59°C
HTT (pre-mRNA)	29; 63°C

Subcellular fractionation and RNA extraction

Fibroblasts were harvested by trypsinization and collected in 15 ml tubes on ice, washed three times with cold PBS and pelleted at 500*g* for 3 min. Cell pellets were then split. One-fourth of the pellet was used for total RNA isolation with TRIzol (Invitrogen), whereas the remaining part was subjected to fractionation into the nuclear and cytoplasmic fractions with PARIS ™ Kit (Ambion). RNA was extracted from the subcellular compartments according to the manufacturer's guidelines. In general, 300–500 ng of RNA was used for cDNA preparation with SuperScript II Reverse Transcriptase (Invitrogen) and random hexamer primers (Promega). To control for genomic DNA contamination mock reactions were performed that lacked reverse transcriptase. Genomic DNA contamination was removed with DNase I treatment before RNA denaturation. The RNA samples were incubated with RNase-free DNase for 15 min at 25°C followed by DNase heat inactivation at 65°C for 10 min in the presence of 5 mM EDTA. Oligonucleotides used and PCR conditions for RT–PCR are shown in Tables 1 and 2, respectively. The depletion of nuclear RNA in cytoplasmic fraction was determined by RT–PCR with primers for HTT pre-mRNA and for 5′-external transcribed spacer (5′-ETS) of pre-rRNA. The 5′-ETS is cleaved prior to export of rRNA from the nucleus (30). PCR products were resolved in 1.3% agarose gel in TBE, stained with ethidium bromide and quantified with Gel-Pro 3.1. HTT mRNA levels in total RNA isolated from one-fourth of the cell pellet and in the cytoplasmic fraction were normalized relative to the levels of GAPDH mRNA whereas nuclear HTT mRNA level was normalized relative to the 5′-ETS.

RNA FISH, IF and microscopy

For RNA FISH and IF cells were fixed in 4% PFA/PBS at 4°C and washed three times in PBS, followed by permeabilization in 2% pre-chilled acetone/PBS for 5 min. Pre-hybridization was performed in 30% formamide and 2× SSC for 10 min, followed by hybridization in buffer containing 30% formamide, 2× SSC, 0.02% BSA, 66 µg/ml yeast tRNA, 10% dextran sulfate, 2 mM vanadyl ribonucleoside complex and 2 ng/µl DNA/LNA

probes $(CTG)_6$-CA or $(CAG)_6$-CA. The probes were labeled at the 5′-end with Cy3 and modified at positions 2, 5, 8, 13, 16 and 19 with LNA. Post-hybridization washing was done in 30% formamide and 2× SSC at 45°C for 30 min followed by 1× SSC at 37°C for the next 30 min. Slides were either mounted in Vectashield medium (Vector Laboratories, INC. Burlingame, CA, USA) with DAPI or were used for subsequent IF. In control experiments, prior to RNA FISH cells were treated for 1.5 h at 37°C with 0.01 μg/ml RNase A (Sigma). For RNA FISH in combination with IF, the slides were treated as described above, but DAPI staining was omitted. The slides were incubated in blocking buffer (5% non-fat milk, 10% normal serum and 0.3% Triton X-100 in PBS) for 1 h at room temperature, and the reaction with rabbit anti-MBNL1 primary antibody (A2764) (a gift from Dr Thornton) was performed overnight at 4°C. After the reaction, the slides were washed in PBS three times for 5 min each and then incubated with a secondary anti-rabbit antibody conjugated to Cy2 (Jackson ImmunoResearch Laboratories, Inc.). Following PBS washes, the slides were mounted in Vectashield medium with DAPI. All FISH/IF images were acquired on a Zeiss Axiovert 200M microscope equipped with a cooled AxioCam HRc camera and processed with LSM 510 software.

RESULTS

Selection of transcripts for structural analysis of HTT and AR triplet repeats

In the *HTT* locus, the polymorphic CCG tract is 12 bp downstream of the expansion-prone $(CAG)_n$. The *AR* locus is composed of $(CTG)_3(CAG)_n$ with a monomorphic $(CAG)_6$ tract 18 bp downstream (Figure 1A and C). This study aimed to determine the structure of HTT and AR mRNA fragments within the triplet repeat regions and the few dozen neighboring nucleotides. For the RNA structure analysis we selected transcripts corresponding to the most frequent normal alleles of the *HTT* and *AR* genes and mutant allele's characteristic of HD and Kennedy's disease. The selection of normal triplet repeat variants was based on the results of our prior genotyping studies (31,32) and the data extracted for *HTT* and *AR* genes are shown in Figure 1B and D.

CAG and CCG repeats interact with each other *in vitro* in HTT transcripts

First, we aimed to determine whether the two neighboring repeat tracts in the *HTT* locus interact with each other and what kind of structure is formed by the selected sequence flanking the repeats. To do this, we first analyzed a transcript containing 10 CAG and 10 CCG repeats (HTT-10-10) in the context of its natural flanking sequence (Figure 2A, left and middle sections). The results of RNA structure probing experiments showed that repeated sequences of the transcript form an autonomous structure in which the CAG and CCG repeats interact with each other over nearly their entire length (Figure 2A, right section). Only a large 13-nt terminal

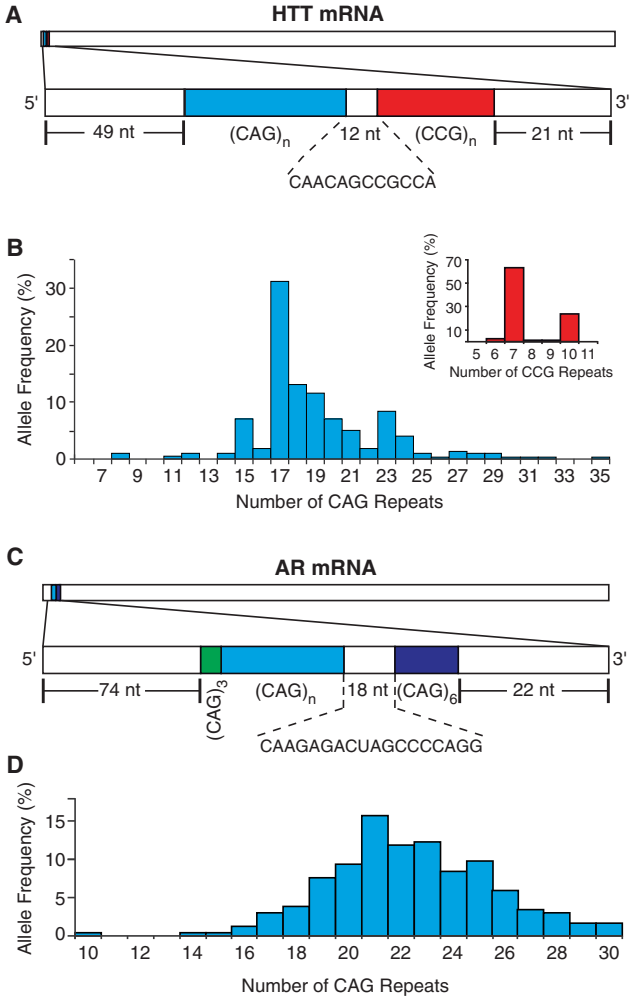

Figure 1. Characteristics of repeat regions from HTT and AR mRNAs. **(A** and **C)** Schematic representation of HTT and AR mRNA fragments. White bars, specific flanking sequences; blue bars, polymorphic CAG repeats; red bar, polymorphic CCG repeat region; dark blue bar, monomorphic $(CAG)_6$ tract; green bar, monomorphic $(CUG)_3$ sequence. Genotyping of CAG and CCG repeat tracts in the *HTT* gene **(B)** and the polymorphic CAG tract from the *AR* gene **(D)**.

loop (loop a) formed mainly by the specific sequence separating the two repeat tracts was well recognized by T1, T2 and S1 single strand-specific nucleases. The stem structure that was formed by the two different types of repeats, the CAG at its 5′ side and the CCG at 3′ side, was composed of repeated C-G and G-C base pairs and A:C mismatches.

To gain better insight into the structure and stability of hairpins composed of the interacting repeats in the HTT transcript, we prepared RNAs composed of CAG and CCG units only. Transcripts harboring $(CAG)_{10}(CCG)_{10}$ and $(CAG)_{20}$ were analyzed. Both RNAs were cleaved in the center by single strand-specific nucleases and the number of cleavages observed in that region was consistent with the coexistence of two slipped variants (Supplementary Figure S1A). To determine differences in the structural stability of the investigated transcripts, we performed UV-monitored structure melting experiments. The obtained results confirmed that the

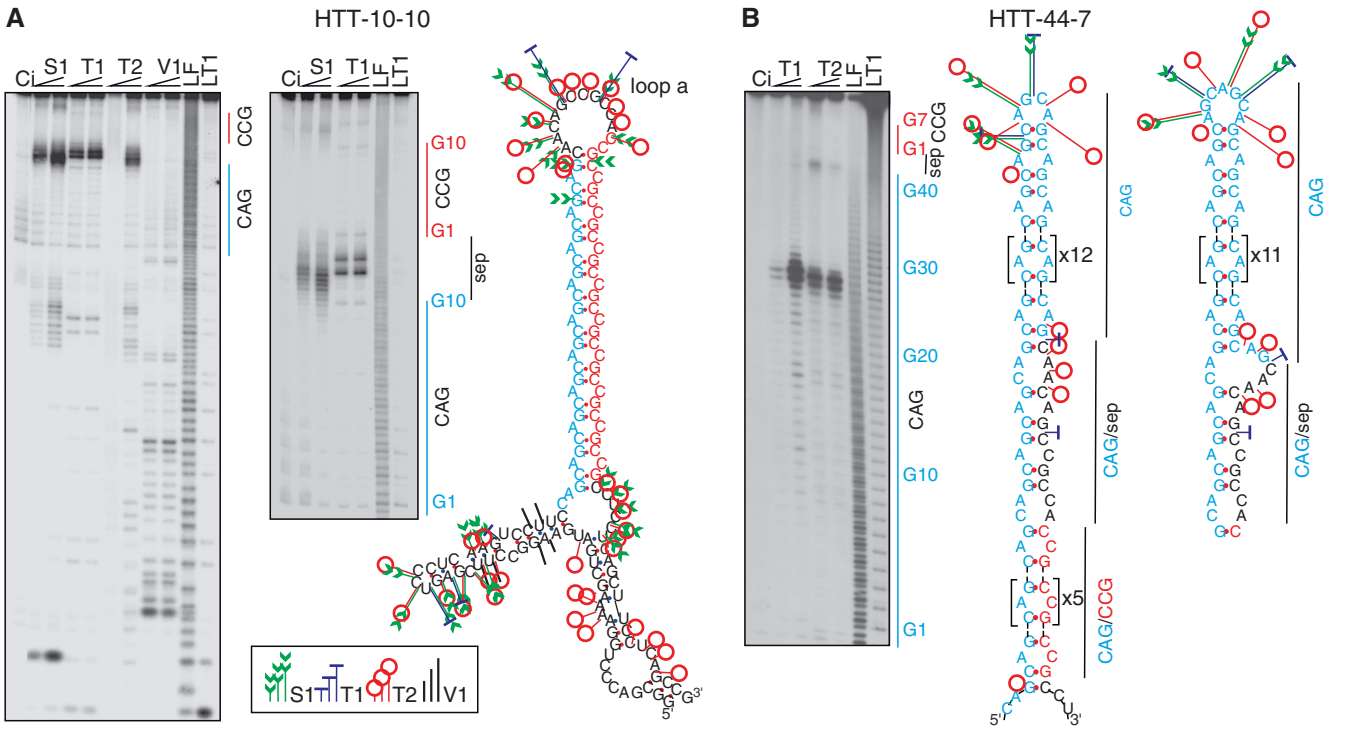

Figure 2. *In vitro* structure of CAG and CCG repeat tracts from HTT mRNA. **(A)** Structure analysis of the HTT-10-10 mRNA fragment. The left autoradiogram (short run) represents the pattern of RNA fragments generated by the nucleases: S1 (2.5 u/μl and 5 u/μl), T1 (0.1 u/μl and 0.15 u/μl), T2 (0.1 u/μl and 0.2 u/μl) and V1 (0.1 u/μl and 0.2 u/μl). The middle autoradiogram (medium run) depicts cleavage products generated by S1 nuclease and T1 RNase in the repeat region. Positions of selected G residues from the repeated sequences (G1 shows the G residue of the first triplet repeat) and the 12-nt repeat separation region (sep) are indicated. The right section shows the secondary structure formed by the HTT mRNA fragment. **(B)** Secondary structure model of the trinucleotide repeat region in the HTT mRNA fragments containing expanded CAG repeats (HTT-44-7). The autoradiogram shows cleavages generated by RNase T1 (0.1 u/μl and 0.15 u/μl) and T2 (0.1 u/μl and 0.2 u/μl) in the repeat region. The middle section presents a model of the secondary structure of the trinucleotide repeat region from HTT-44-7 whereas the right section shows an alternative model of the secondary structure in the terminal loop and CAG/sep region. Structural motifs that are repeated several times are surrounded by brackets. LF and LT1 indicate the 1-nt and G-specific ladders, respectively; Ci, incubation control in which RNA was incubated in the absence of nucleases. Sites and intensity of cleavages induced by the different probes are indicated (inset legend).

$(CAG)_{10}(CCG)_{10}$ transcript formed a significantly more stable structure than the $(CAG)_{20}$ transcript, with the folding free energy change $\Delta G^{37°C} = -15.38 \pm 0.43$ kcal/mol and $\Delta G^{37°C} = -5.8 \pm 0.39$ kcal/mol and melting temperatures of 77.5 and 62.4°C, respectively (Supplementary Figure S1B). This difference in stability is likely to be related to the nature of base mismatches occurring periodically in the stems of hairpins in both transcripts.

Variations in the number of CAG repeats in HTT modulate RNA hairpin architecture

To determine the consequences of normal lengthening and pathogenic expansions of the CAG repeats on RNA structure, we analyzed normal HTT transcripts containing $(CCG)_7$ and 17 or 24 CAG repeats (Supplementary Figure S2) as well as the mutant transcript harboring 44 CAG repeats (Figure 2B). The overall structural architecture of the triplet repeat region in these four transcripts was very similar. In all analyzed RNAs, a large part of the stem structure was formed by the interacting $(CAG)_7$ and $(CCG)_7$ repeats, which is similar to that described above for the HTT-10-10 repeat (Figure 2A). The cleavage pattern detected in the terminal loop region and CAG/sep region of transcripts containing CAG repeat tracts longer than CCG repeat tracts is consistent with co-existence of two structural variants of this part of HTT hairpin (Figure 2B and Supplementary Figure S2). This conclusion is supported by the results of structural analysis of 3′-end labeled HTT transcripts (data not shown). The ability to form this kind of slipped structure is typical for trinucleotide repeats (24). The cleavage pattern generated by nucleases within the repeat flanking sequences of all compared HTT transcripts was almost identical (Figure 2A). Therefore, the structure formed by the two repeat tracts was indeed autonomous and was not influenced by the structure of the flanking sequence. The only significant difference between the repeat regions containing long normal repeats and the expanded mutant repeats was the length of the hairpin stem composed of pure CAG repeats.

The CAG repeat hairpin is stabilized by the neighboring short $(CUG)_3$ in the AR transcript

Next, we aimed to determine the influence on hairpin structure formation of two short repeated tracts, $(CUG)_3$

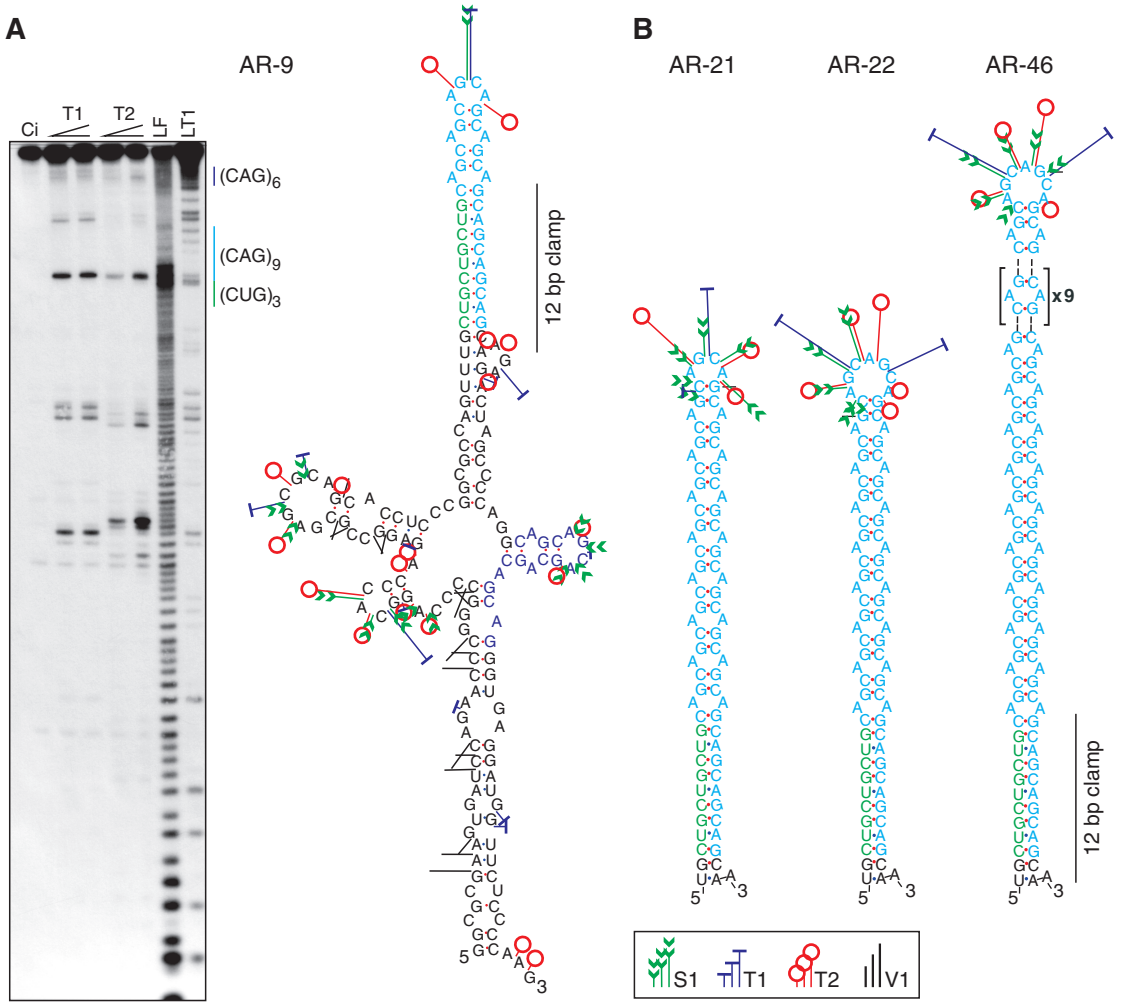

Figure 3. *In vitro* structure analysis of repeat region from AR mRNA fragments. **(A)** The autoradiogram (right section) depicts RNA fragments generated by T1 (0.1 u/µl and 0.15 u/µl) and T2 (0.1 u/µl and 0.2 u/µl) enzymatic probes in the AR-9 transcript (containing 9 CAG repeats). Other symbols are as in the legend to Figure 2. The right section shows the secondary structure formed by the AR-9 transcript. **(B)** Secondary structures of the trinucleotide repeat region from AR transcripts harboring different numbers of CAG repeats (from left to right, 21, 22 and 46 CAG repeats). The hairpin structure stabilizing interaction between CUG and CAG repeats is indicated by a 12-bp clamp.

and $(CAG)_6$, located in close proximity to the polymorphic CAG repeats subject to expansion in the *AR* locus. For RNA structure probing experiments, we selected three normal *AR* gene variants containing 9, 21 and 22 CAG repeats (AR-9, AR-21 and AR-22) as well as one mutant transcript with an expanded $(CAG)_{46}$ tract (AR-46). Based on nuclease digestion, it appeared that in all transcripts, three CUG triplets interact with the 3′-end of the $(CAG)_n$ sequence, forming a perfect 12-bp helical region together with a few nucleotides of repeat flanking sequence (Figure 3). The remaining polymorphic CAG repeats fold into an RNA hairpin structure, which contains either a 4- or a 7-nt terminal loop (Figure 3B, left and middle sections). Due to stabilizing interactions between CUG and CAG repeats in both normal and mutant AR transcripts, the slipped nature of the hairpin that is characteristic of triplet repeat sequences was not observed. The only difference between analyzed mRNA fragments with increasing numbers of CAG repeats was the length of the CAG hairpin structure (Figure 3B). In each AR transcript tested, the specific flanking sequence formed identical structures, and the short $(CAG)_6$ sequence located 18 nt downstream from the polymorphic CAG repeats folded into autonomous short hairpins (Figure 3A).

In summary, structural analysis of the complex triplet repeat regions of HTT and AR transcripts revealed that both CAG-harboring RNAs form *in vitro* hairpin structures that are stabilized by $(CCG)_n$ or $(CUG)_3$ motifs adjacent to the polymorphic CAG. Due to the similarity in structural conformation and contribution of the neighboring repeats to the hairpin formation in HTT and AR transcripts, we selected human cells expressing an HTT mutant transcript for further *in vivo* studies. We utilized these cells to examine the colocalization of mutant CAG RNA with the dsRNA-binding protein MBNL1 and to test the susceptibility to RNA interference (RNAi) to characterize the *in vivo* structure.

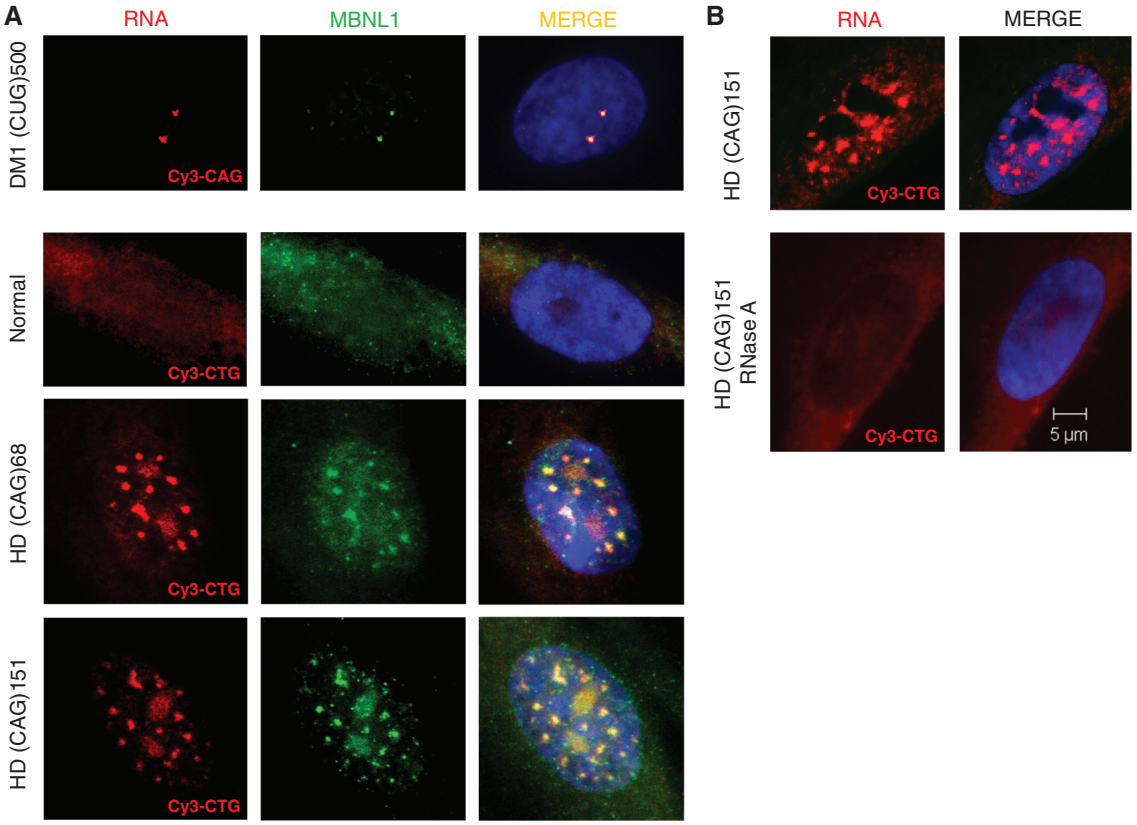

Figure 4. Mutant CAG RNA repeats form nuclear aggregates in human HD cells and colocalize with MBNL1. **(A)** Representative images of RNA FISH and IF performed on cultured human fibroblasts. HD cells expressing 68 and 151 CAG repeats from the *HTT* gene show nuclear accumulation of CAG RNA aggregates, as detected with a CTG probe. As a positive control, human DM1 fibroblasts were included in the FISH and IF assays to detect ribonuclear CUG inclusions with a CAG probe and to show their colocalization with MBNL1 protein. In normal cells, MBNL1 is detected over the entire cell, whereas the presence of CAG RNA foci in HD cells causes colocalization with the splicing factor as depicted also in DM1 cells. CUG and CAG RNA foci are displayed in red, MBNL1 is shown in green and nuclear DNA was stained with DAPI (blue). **(B)** HD cells expressing 151 CAG repeats were treated with RNase A followed by FISH with the CTG probe. The CAG RNA nuclear aggregates were sensitive to RNase A treatment. CAG RNA foci (red), DAPI nuclear stain (blue).

Expanded CAG repeat RNAs are accumulated in HD nuclei and colocalize with MBNL1 protein

To determine whether mutant HTT transcripts are retained in the nuclei of cultured human HD cells, we performed FISH and tested for presence of CAG repeat aggregates. Fibroblast cells expressing expanded 68 and 151 CAG repeats from the *HTT* locus were used for this experiment. As a positive control for FISH we selected DM1 cells in which nuclear deposition of mutant CUG^{exp} transcripts has been previously shown (15). Figure 4 depicts that numerous nuclear RNA inclusions composed of CAG repeats were detected in the analyzed HD cells. We confirmed the presence of RNA in the nuclear inclusions with RNase A treatment (Figure 4B). Additionally, treatment with DNase did not affect the pattern of foci (data not shown). IF was performed to determine whether nuclear CAG aggregates in HD cells colocalize with the MBNL1 protein. Analysis of MBNL1 intracellular localization in HD cells revealed that the dsRNA-binding protein colocalized with nuclear CAG repeat inclusions as shown for CUG^{exp} foci in DM1 cells (Figure 4A). In contrast, in normal fibroblasts the protein was distributed throughout the cell.

Targeting CAG repeats by RNA interference in human HD cells

To determine whether the structure of the repeat region in the HTT transcript is susceptible to RNA interference, we utilized several siRNA duplexes (Table 1). Three patient derived HD fibroblast cell lines containing 15/46, 21/44 and 17/68 CAG repeats in the *HTT* gene were transfected with siRNAs, and the silencing effect was assessed for each *HTT* allele 24 h post-transfection by RT–PCR amplification of the repeat region (Figure 5A). Two siRNAs (siHD sep and siCUG7/CAG7) targeting the hairpin and two (siHD 5′ and siHD 3′) targeting flanking sequences (Figure 5B) were highly effective in silencing HTT transcripts. However, only the siCUG7/CAG7 siRNA showed some allele discrimination, cleaving the mutant transcripts more efficiently. This result implies that the structure formed by the CAG repeats in the analyzed HD cells was not stable enough to form a barrier for RNA-induced silencing complex (RISC) activity. In contrast to the four active siRNAs, the CCG repeat targeting RNA duplex did not trigger any considerable silencing effect, which may be explained by its high stability and inability to activate RISC (data not shown).

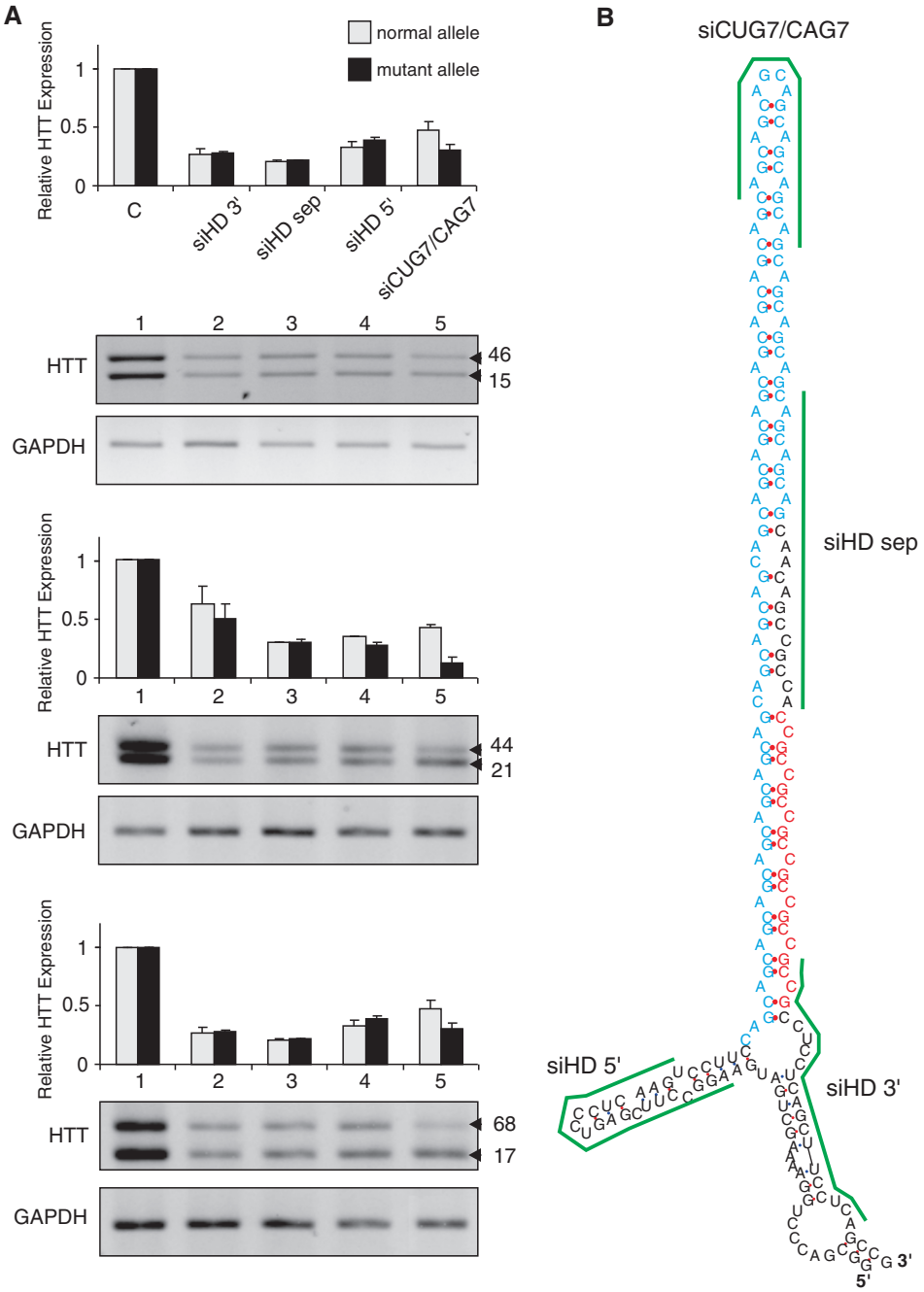

Figure 5. Silencing effects of siRNAs on HTT mRNA expression levels. **(A)** Results of quantitative RT–PCR showing HTT mRNA expression levels in three human HD cell lines harboring 15/46, 21/44 and 17/68 CAG repeats (CAG repeat sizes from normal and mutant alleles are shown on the side). The cells were either treated with transfection reagent only (lane 1) or treated with one of four indicated siRNAs used at 20 nM (lanes 2–5). Bar graphs represent mean values from three independent experiments normalized to GAPDH expression (SDs are shown as error bars). **(B)** Potential localization of target sequences (green lines) for the indicated siRNAs on a structural model of the HTT transcript. For siCUG7/CAG7 only one possible binding site is shown.

Next, to determine the CAG repeat tract cleavage sites by the repeat targeting siRNA, we performed 5′-RACE analysis. As controls for this experiment, we used siRNAs targeting non-repeating regions of the HTT transcript (Supplementary Figure S3). As expected, the HTT-specific siRNAs gave rise to single discrete amplification products with lengths corresponding to sites of AGO2-mediated cleavage. In contrast, multiple cleavage sites were observed in the CAG repeat region for HD cells treated with the CUG7/CAG7 siRNA, indicating that the repeated sequence was recognized by RISC at numerous positions (Supplementary Figure S3).

In addition to HD cells, we analyzed SCA3, SCA7 and DRPLA fibroblast cell lines to determine whether similar allele discrimination by the CAG repeat targeting duplex siCUG7/CAG7 could be achieved in cellular models of

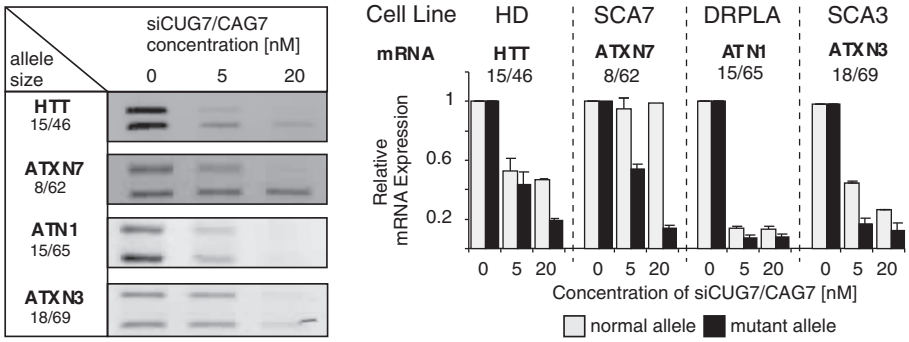

Figure 6. Effect of CUG7/CAG7 siRNA on the silencing of HTT, ATXN7, ATXN3 and ATN1 transcripts. Quantitative RT–PCR analysis of four transcripts containing CAG repeats in human HD, SCA7, SCA3 and DRPLA fibroblasts treated with two different concentrations of CUG7/CAG7 siRNA (left panel). Upper bands represent mutant and lower bands represent normal mRNA variants. Bar graphs (right panel) show relative expression levels of the indicated mRNA variants calculated in three independent experiments and normalized to GAPDH expression (error bars represent SDs).

other polyglutamine diseases (Figure 6). In all CAG expansion-containing cell lines, the siRNA duplex used at two different concentrations was effective in silencing targeted genes. The efficiency of silencing was higher for mutant than for normal alleles. This result may be a consequence of the presence of more potential RISC binding sites in mutant alleles.

Downregulation of huntingtin transcript by siCUG7/CAG7 duplex predominates in the cytoplasm

In order to determine whether the observed siRNA silencing of the HTT mRNA took place only in the cytoplasm or also in the nucleus of HD cells, we used fractionated samples of RNA. Fibroblasts transfected with 20 nM siCUG7/CAG7 duplex were partitioned into the nuclear and cytoplasmic fractions and each of them was individually analyzed for the RNAi effect. The depletion of nuclear RNA from the cytoplasmic fraction was demonstrated by minute amounts of the HTT pre-mRNA and the 5′-ETS fragment of pre-rRNA in the cytoplasmic fraction (Figure 7). Mature HTT transcripts were detected in both the nuclear and cytoplasmic RNA isolated from HD cells expressing either 21/44 or 17/68 CAG repeats. Upon siCUG7/CAG7 treatment, we observed a decreased level of transcripts from both HTT alleles and the silencing was significantly more pronounced in the cytoplasm than in the nucleus. In cells expressing 21/44 CAG repeat RNAs the relative HTT mRNA levels decreased to 0.16 ± 0.005 in the cytoplasm and to 0.85 ± 0.03 in the nuclear fraction. Similar correlation was found in fibroblasts containing 17/68 CAG repeats in HTT locus where the relative HTT mRNA levels in the cytoplasm diminished to 0.4 ± 0.06 while in the nucleus only to 0.88 ± 0.08. Analysis of HTT pre-mRNA expression confirmed the low level of nuclear silencing with siCUG7/CAG7 in both the analyzed cell lines (Figure 7). Therefore, utilization of siCUG7/CAG7 duplex targeting HTT transcript caused its efficient silencing which was significantly more pronounced in the cytoplasmic fraction of RNA. It is also possible that the minor HTT silencing which we observed in the nuclear fraction resulted from contamination of the RNA with traces of the cytoplasmic fraction rather than being an effect of the nuclear RNAi.

DISCUSSION

Triplet repeat regions from transcripts of genes implicated in trinucleotide repeat expansion diseases form hairpin structures with different architectures. These conformations resemble dsRNA and are thought to trigger the pathogenesis of toxic RNA-dependent disorders such as DM1 but have also been suggested to play a role in poly-Q-triggered diseases (14,16). In this study, we provided *in vitro* and *in vivo* data showing that the structural features of mutant CAG repeat transcripts from *HTT* and *AR* genes may predispose them as potential pathogenic factors in polyglutamine diseases. Herein, we (i) determined the structural conformations of HTT and AR mRNA fragments within regions flanking polymorphic CAG repeats and other triplet repeats of neighboring sequences, (ii) showed that HTT transcripts harboring CAG expansions form intranuclear foci that colocalize with MBNL1 and (iii) found that hairpin structures of CAG repeat mutant transcripts are not strong enough to resist RISC attack.

According to the results of RNA *in vitro* structure analysis, the CAG repeats from the AR transcript form a 'frozen' type of hairpin stabilized by a short $(CUG)_3$ tract adjacent to the polymorphic $(CAG)_n$. This is in contrast to previous reports showing that specific flanking sequences of SCA1 and SCA6 repeats in the ATXN1 and CACNA1A transcripts stabilize the hairpin (18,19). The only difference between normal and mutant AR transcripts was the length of the hairpin formed by the pure CAG repeats. On the other hand, the majority of HTT repeat variants have a tripartite modular structure composed of interacting CAG and CCG repeats at its base, a central module formed by the interacting CAG repeats and the specific sequence, and a terminal section formed by the fold-back structure composed of pure CAG repeats. The stability of the base section of the hairpin structure composed of interacting CAG and CCG repeat tracts is much higher than other hairpin parts, which is

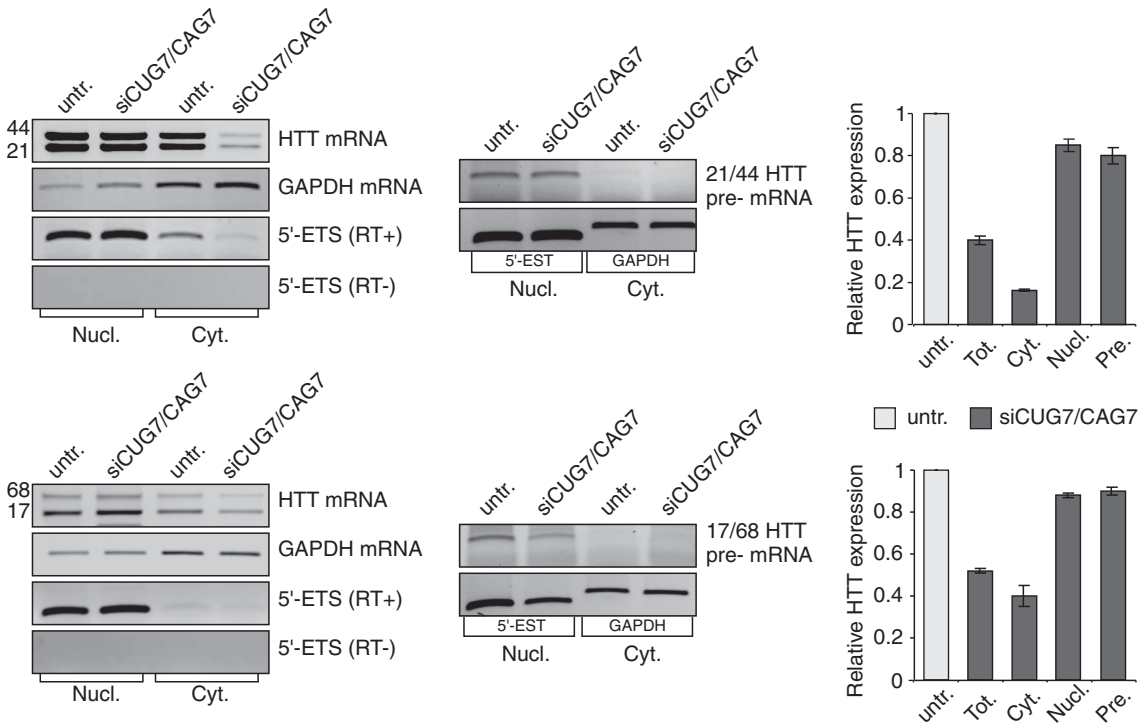

Figure 7. siRNA targeting CAG repeats acts predominantly in the cytoplasm. RT–PCR results for the nuclear and cytoplasmic RNA isolated from HD fibroblasts untransfected and siRNA transfected. The experiments were carried out in triplicate and representative gels are shown. Quantitative results of RT–PCR shown in the bar diagrams represent the relative levels of HTT RNA expression (±SD) normalized to HTT RNA level in untransfected HD cells and either to GAPDH expression (for total and cytoplasmic RNA) or to the 5′-ETS (for the nuclear RNA). As shown, HTT RNA silencing was significantly stronger in the cytoplasm than in the nucleus. Low level of nuclear silencing was confirmed by analysis of HTT pre-mRNA. (RT+) indicates the presence and (RT–) the absence of reverse transcriptase in the reaction; Untr., untransfected; siRNA, siCUG7/CAG7; Nucl., nuclear; Cyt., cytoplasmic; Tot., total; Pre., pre-mRNA.

driven by the properties of the reiterated A:C mismatches. In consequence, the upper and central modules of HTT repeat hairpin may exist in the alternative structural variants resulting from slippage of the CAG repeats. The most variable section of the HTT repeat hairpin is the part containing pure CAG repeats. In transcripts from the *HTT* mutant alleles the majority of the CAG repeat region forms the terminal section of the hairpin structure, similar to the previously reported CAG hairpin architecture formed by mutant transcripts from other poly-Q disease genes (18–20).

Evidence for involvement of abnormally long poly-Q tracts in the pathogenesis of polyglutamine diseases is well documented (5). However, the possible contribution of mutant CAG RNA in poly-Q-triggered degeneration has also been suggested. Ho *et al*. (15) described the formation of intranuclear CAG RNA aggregates in COSM6 cells transfected with a plasmid encoding long interrupted CAG repeats. Despite the presence of sequestered MBNL1 in the foci, their toxic effect was milder in comparison to CUG foci, and they did not considerably alter the splicing of the two tested pre-mRNAs, cardiac troponin T (TNNT2) and insulin receptor (INSR) (15). This raises the question of whether the presence of interruptions in the CAG was responsible for the diminished toxicity of the mutant RNA. More evidence for involvement of RNA in poly-Q-triggered degeneration was provided by a study conducted using SCA3 *Drosophila* as a model (16). Transcripts containing CAG expansions were retained in the nuclei of fly eyes, but their toxicity was mitigated by the presence of interruptions altering the CAG repeat sequence. Therefore, the purity of long expanded CAG tracts which allow the formation of long hairpin structures within the repeats seems to be the most significant determinant of pathogenesis triggered by mutant RNA. Moreover, long CAG repeats present in the 3′-UTR conferred neuronal degeneration in a transgenic *Drosophila* model (16). Our study showed that in cultured human HD fibroblast cells, endogenous HTT transcripts harboring uninterrupted CAG expansions are sequestered in the nuclei where they form RNA foci. The conformations of the repeated RNA sequences have been revealed *in situ* by their colocalization with the MBNL1 protein. Therefore, we provide evidence that expanded CAG repeats from HTT transcripts form hairpin structures *in vitro* and that due to their structural conformation, these RNAs are being deposited *in vivo* in the nuclei of human cells. However, it remains unclear whether the nuclear CAG repeat aggregates represent a stable fraction of toxic RNA that does not leave the nuclear compartment, as shown for CUG RNA foci, or whether they are efficiently exported to the cytoplasm after temporary participation in the formation of foci. If CAG foci represent the soluble fraction of a short half-life

RNA in the nucleus, this could affect their toxic capacity. For CUG RNA foci, it was shown that nuclear but not cytoplasmic localization triggers pathogenesis in DM1 cells (33). Therefore, efficient export of toxic RNA to the cytoplasm may rescue the cell from the harmful effects of the RNA. If this is true for mutant CAG repeat RNA, which is indeed translated into toxic protein its involvement in the pathogenesis of poly-Q-triggered diseases may be diminished but not abrogated.

Sequence specific targeting of transcripts by RNAi was shown to be inefficient for heavily structured targets such as those embedded in the stem portions of stable hairpins (34–36). Due to a lack of RNA helicase activity, RISC cannot unwind such firm RNA segments (37–40). Therefore, RISC may be considered a programmable endogenous probe for stable structure detection within cellular transcripts. With this rationale, we tested an siRNA duplex against the CAG repeat region to evaluate its activity toward structures formed within the HTT transcript. Our experimental results showed that structures formed by normal and mutant CAG repeat regions are susceptible to RNAi what was also observed by other authors (41,42). Unexpectedly, we observed a moderate preference for silencing the mutant HTT transcript with siRNA targeting the CAG repeat region. Similar results were obtained for SCA3, SCA7 and DRPLA, which contain expanded CAG repeats. These observations demonstrate that although structures formed by the CAG repeats *in vivo* are stable enough to form nuclear foci that bind MBNL1 and trigger alternative splicing of several genes (Mykowska A. *et al*., submitted for publication), these structures are not strong enough to resist attack by RISC. Periodic occurrence of A:A mismatches is a structural factor responsible for the lower stability of the CAG repeat hairpins compared to hairpins with a fully base-paired stem. The recently resolved crystal structure of the CAG repeat duplex shows the nature of the A:A interaction with atomic resolution (43). The A:A pairs fit well within a regular RNA-A helix, and the only interaction between the paired adenine residues is the unusual and weak C2–H2–N1 hydrogen bond. Adenine residues adopt different positions within the helix depending on which residues play roles as H-bond donors or acceptors. Therefore, the decreased stability and increased flexibility of CAG repeat hairpins may explain their susceptibility to interactions with RISC.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Dr Charles A. Thornton for rabbit MBNL1 antibody A2764.

FUNDING

This work was supported by funding from the Sixth Research Framework Program of the European Union, Project RIGHT [LSHB-CT-2004-005276], the Ministry of Science and Higher Education grant [N N401 097536], and the European ERDF in PO IG grant [POIG.01.03.01-30-098/08].

Conflict of interest statement. None declared.

REFERENCES

1. Bates,G.P. (2005) History of genetic disease: the molecular genetics of Huntington disease - a history. *Nat. Rev. Genet*., **6**, 766–773.
2. Martin,J.B. and Gusella,J.F. (1986) Huntington's disease. Pathogenesis and management. *N. Engl. J. Med.*, **315**, 1267–1276.
3. The Huntington's Disease Collaborative Research Group. (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell*, **72**, 971–983.
4. La Spada,A.R., Wilson,E.M., Lubahn,D.B., Harding,A.E. and Fischbeck,K.H. (1991) Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature*, **352**, 77–79.
5. Gatchel,J.R. and Zoghbi,H.Y. (2005) Diseases of unstable repeat expansion: mechanisms and common principles. *Nat. Rev. Genet*., **6**, 743–755.
6. DiFiglia,M., Sapp,E., Chase,K.O., Davies,S.W., Bates,G.P., Vonsattel,J.P. and Aronin,N. (1997) Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. *Science*, **277**, 1990–1993.
7. Rangone,H., Humbert,S. and Saudou,F. (2004) Huntington's disease: how does huntingtin, an anti-apoptotic protein, become toxic? *Pathol. Biol*., **52**, 338–342.
8. Saudou,F., Finkbeiner,S., Devys,D. and Greenberg,M.E. (1998) Huntingtin acts in the nucleus to induce apoptosis but death does not correlate with the formation of intranuclear inclusions. *Cell*, **95**, 55–66.
9. Amack,J.D. and Mahadevan,M.S. (2001) The myotonic dystrophy expanded CUG repeat tract is necessary but not sufficient to disrupt C2C12 myoblast differentiation. *Hum. Mol. Genet*., **10**, 1879–1887.
10. Mankodi,A., Logigian,E., Callahan,L., McClain,C., White,R., Henderson,D., Krym,M. and Thornton,C.A. (2000) Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. *Science*, **289**, 1769–1773.
11. Brook,J.D., McCurrach,M.E., Harley,H.G., Buckler,A.J., Church,D., Aburatani,H., Hunter,K., Stanton,V.P., Thirion,J.P., Hudson,T. *et al*. (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*, **69**, 385.
12. Liquori,C.L., Ricker,K., Moseley,M.L., Jacobsen,J.F., Kress,W., Naylor,S.L., Day,J.W. and Ranum,L.P. (2001) Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science*, **293**, 864–867.
13. Sellier,C., Rau,F., Liu,Y., Tassone,F., Hukema,R.K., Gattoni,R., Schneider,A., Richard,S., Willemsen,R., Elliott,D.J. *et al*. (2010) Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients. *EMBO J*., **29**, 1248–1261.
14. Orr,H.T. (2004) RNA gains a new function: a mediator of neurodegeneration. *Trends Neurosci*., **27**, 233–234.
15. Ho,T.H., Savkur,R.S., Poulos,M.G., Mancini,M.A., Swanson,M.S. and Cooper,T.A. (2005) Colocalization of muscleblind with RNA foci is separable from mis-regulation of alternative splicing in myotonic dystrophy. *J. Cell Sci*., **118**, 2923–2933.
16. Li,L.B., Yu,Z., Teng,X. and Bonini,N.M. (2008) RNA toxicity is a component of ataxin-3 degeneration in Drosophila. *Nature*, **453**, 1107–1111.
17. Napierala,M. and Krzyzosiak,W.J. (1997) CUG repeats present in myotonin kinase RNA form metastable "slippery" hairpins. *J. Biol. Chem*., **272**, 31079–31085.

18. Michlewski,G. and Krzyzosiak,W.J. (2004) Molecular architecture of CAG repeats in human disease related transcripts. *J. Mol. Biol.*, **340**, 665–679.
19. Sobczak,K. and Krzyzosiak,W.J. (2004) Imperfect CAG repeats form diverse structures in SCA1 transcripts. *J. Biol. Chem.*, **279**, 41563–41572.
20. Sobczak,K. and Krzyzosiak,W.J. (2005) CAG repeats containing CAA interruptions form branched hairpin structures in spinocerebellar ataxia type 2 transcripts. *J. Biol. Chem.*, **280**, 3898–3910.
21. Napierala,M., Michalowski,D., de Mezer,M. and Krzyzosiak,W.J. (2005) Facile FMR1 mRNA structure regulation by interruptions in CGG repeats. *Nucleic Acids Res.*, **33**, 451–463.
22. Mathews,D.H., Sabina,J., Zuker,M. and Turner,D.H. (1999) Expanded Sequence Dependence of Thermodynamic Parameters Improves Prediction of RNA Secondary Structure. *J. Mol. Biol.*, **288**, 911–940.
23. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
24. Sobczak,K., de Mezer,M., Michlewski,G., Krol,J. and Krzyzosiak,W.J. (2003) RNA structure of trinucleotide repeats associated with human neurological diseases. *Nucleic Acids Res.*, **31**, 5469–5482.
25. Ehresmann,C., Baudin,F., Mougel,M., Romby,P., Ebel,J.P. and Ehresmann,B. (1987) Probing the structure of RNAs in solution. *Nucleic Acids Res.*, **15**, 9109–9128.
26. Knapp,G. (1989) Enzymatic approaches to probing of RNA secondary and tertiary structure. *Methods Enzymol.*, **180**, 192–212.
27. Sobczak,K., Michlewski,G., de Mezer,M., Kierzek,E., Krol,J., Olejniczak,M., Kierzek,R. and Krzyzosiak,W.J. (2010) Structural diversity of triplet repeat RNAs. *J. Biol. Chem.*, **285**, 12755–12764.
28. Sobczak,K., Michlewski,G., de Mezer,M., Krol,J. and Krzyzosiak,W.J. (2010) Trinucleotide repeat system for sequence specificity analysis of RNA structure probing reagents. *Anal. Biochem.*, **402**, 40–46.
29. Krol,J., Sobczak,K., Wilczynska,U., Drath,M., Jasinska,A., Kaczynska,D. and Krzyzosiak,W.J. (2004) Structural features of microRNA (miRNA) precursors and their relevance to miRNA biogenesis and small interfering RNA/short hairpin RNA design. *J. Biol. Chem.*, **279**, 42230–42239.
30. Renalier,M.H., Mazan,S., Joseph,N., Michot,B. and Bachellerie,J.P. (1989) Structure of the 5′-external transcribed spacer of the human ribosomal RNA gene. *FEBS Lett.*, **249**, 279–284.
31. Rozanska,M., Sobczak,K., Jasinska,A., Napierala,M., Kaczynska,D., Czerny,A., Koziel,M., Kozlowski,P., Olejniczak,M. and Krzyzosiak,W.J. (2007) CAG and CTG repeat polymorphism in exons of human genes shows distinct features at the expandable loci. *Hum. Mutat.*, **28**, 451–458.
32. Sobczak,K. and Krzyzosiak,W.J. (2004) Patterns of CAG repeat interruptions in SCA1 and SCA2 genes in relation to repeat instability. *Hum. Mutat.*, **24**, 236–247.
33. Dansithong,W., Wolf,C.M., Sarkar,P., Paul,S., Chiang,A., Holt,I., Morris,G.E., Branco,D., Sherwood,M.C., Comai,L. *et al.* (2008) Cytoplasmic CUG RNA foci are insufficient to elicit key DM1 features. *PLoS One*, **3**, e3968.
34. Rana,T.M. (2007) Illuminating the silence: understanding the structure and function of small RNAs. *Nat. Rev. Mol. Cell Biol.*, **8**, 23–36.
35. Schubert,S., Grunweller,A., Erdmann,V.A. and Kurreck,J. (2005) Local RNA target structure influences siRNA efficacy: systematic analysis of intentionally designed binding regions. *J. Mol. Biol.*, **348**, 883–893.
36. Kertesz,M., Lovino,N., Unnerstall,U., Gaul,U. and Segal,E. (2007) The role of site accessibility in microRNA target recognition. *Nat. Genet.*, **39**, 1278–1284.
37. Ameres,S.L., Martinez,J. and Schroeder,R. (2007) Molecular basis for target RNA recognition and cleavage by human RISC. *Cell*, **130**, 101–112.
38. Shao,Y., Chan,C.Y., Maliyekkel,A., Lawrence,C.E., Roninson,I.B. and Ding,Y. (2007) Effect of target secondary structure on RNAi efficiency. *RNA*, **13**, 1631–1640.
39. Tafer,H., Ameres,S.L., Obernosterer,G., Gebeshuber,C.A., Schroeder,R., Martinez,J. and Hofacker,I.L. (2008) The impact of target site accessibility on the design of effective siRNAs. *Nat. Biotechnol.*, **26**, 578–583.
40. Lu,Z.J. and Mathews,D.H. (2008) Efficient siRNA selection using hybridization thermodynamics. *Nucleic Acids Res.*, **36**, 640–647.
41. Miller,V.M., Xia,H., Marrs,G.L., Gouvion,C.M., Lee,G., Davidson,B.L. and Paulson,H.L. (2003) Allele-specific silencing of dominant disease genes. *Proc. Natl Acad. Sci. USA*, **100**, 7195–7200.
42. Hu,J., Matsui,M., Gagnon,K.T., Schwartz,J.C., Gabillet,S., Arar,K., Wu,J., Bezprozvanny,I. and Corey,D.R. (2009) Allele-specific silencing of mutant huntingtin and ataxin-3 genes by targeting expanded CAG repeats in mRNAs. *Nat. Biotechnol.*, **27**, 478–484.
43. Kiliszek,A., Kierzek,R., Krzyzosiak,W.J. and Rypniewski,W. (2010) Atomic resolution structure of CAG RNA repeats: structural insights and implications for the trinucleotide repeat expansion diseases. *Nucleic Acids Res.*, **38**, 8370–8376. doi: 10.1093/nar/gkq700.