Supporting Information

Zu et al. 10.1073/pnas.1013343108

SI Materials and Methods

cDNA Constructs. The ATXN8 minigene (A8)(*KMQ_{EXP})-endo construct was generated by subcloning spinocerebellar ataxia type 8 (*SCA8*) cDNA into a pcDNA3.1 vector in the CAG direction. The *SCA8* locus containing the CAG repeat expansion was amplified by PCR from the bacterial artificial chromosome (BAC) transgene construct BAC-Exp (1)1 using the 5′ primer (5′-CGAACCAAGCTTATCCCAATTCCTTGGCTAGACCC-3′) containing an added HindIII restriction site and the 3′ primer (5′-ACCTGCTCTAGATAAATTCTTA AGTAAGAGATAA-GC-3′) containing an added XbaI restriction site. The HindIII/XbaI PCR product was cloned into the pcDNA3.1/*myc*-His A vector (Invitrogen) in the CAG orientation and placed under the control of the CMV promoter. The ATG start codon in the polyglutamine (polyGln) frame was mutated into AAG to remove the existing ORF and generate the A8(*KKQ_{EXP})-endo construct. To generate the A8(*KMQ_{EXP})-3Tf1, A8(*KKQ_{EXP})-3Tf1, A8(*KKQ_{EXP})-3Tf2, and A8(*KKQ_{EXP})-3Tf3 constructs, the HindIII/XbaI fragment was subcloned into pcDNA3.1/6Stops-3T vector. Stop codons between the 3′ end of the repeat and the tags were removed subsequently. In the resulting constructs, six stop codons (two for each frame) were placed prior to the 5′ end of the fragment, and each of three reading frames [polyGln, polyalanine (polyAla), and polyserine (polySer)] was tagged with myc-His, HA, and Flag epitopes, respectively. The ATT(CAG_{EXP})-3T construct was made by inserting the PCR fragment containing a pure CAG repeat into the pcDNA3.1/6Stops-3T vector. This construct contains a very limited sequence (5′-TAGAATT-CAG-3′) between the stop codon cassette and the CAG repeat tract. To remove the sequence between the last 5′ stop codon and the CAG repeat, the ATT(CAG_{EXP})-3T construct was digested with EcoRI, treated with mung bean nuclease, and ligated generating the TAG(CAG_{EXP})-3T construct, in which the last 5′ stop codon (TAG) is placed immediately upstream of CAG repeats, eliminating the existence of upstream alternative translation initiation. To generate the TAAG(CAG_{EXP})-3T construct, PCR was carried out using the 5′ primer (5′-AGTTAAGCTAGCTTAGCTAGGTAACTAAG-TAACTAGAATTAA-3′) and the 3′ primer (5′-TAGAAGGC-ACAGTCGAGGCTGATCA GCGGGTTT-3′). The PCR product was subcloned into the pcDNA3.1/6Stops-3T vector. To generate the TAGAG(CAG_{EXP})-3T construct, PCR was carried out using the 5′ primer (5′-AGTTAAGCTAGCTTAGCTA-GGTAACTAAGTAACTAGAATAGAGCA-3′) and the 3′ primer (5′-TAGAAGGCACAGTCGAGGCTGATCAGCGG-GTTT-3′). The resulting product was subcloned into the pcDNA3.1/6Stops-3T vector. The Huntingtin (HD)-3T, Huntingtin-like 2 (HDL2)-3T, spinocerebellar ataxia type 3 (SCA3)-3T, and myotonic dystrophy type 1 (DM1)-3T constructs were made by inserting the duplex primers containing 20 nt 5′ of the CAG repeats from HD, HDL2, SCA3, and DM1 into the *Eco*RI site of the ATT(CAG_{EXP})-3T construct. The extra nucleotides between the 5′ flanking sequence (HD, HDL2, SCA3, and DM1) and CAG repeats were removed by digesting with EcoRI and another restriction site on the duplex primers, followed by treatment with mung bean nuclease and DNA ligase.

The NheI/PmeI fragments of A8(*KMQ_{EXP})-3Tf1, HD-3T, HDL2-3T, SCA3-3T, and DM1-3T constructs containing six stop codons, expanded CAG repeats, and three tags were subcloned into the lentiviral vector, CSII.

The ATG-V5(CAG_{105})-3T construct was created by inserting an oligo (5′-GAATTATGGGTAAGCCTATCCCTAACCCTC-TCCTCGGTCTCGATTCTACGGGA -3′ and 5′-ATTCCCGT-AGAATCGAGACCGAGGAGAGGGTTAGGGATAGGCT-TACCCAT-3′) containing a V5 tag at the 5′ end of the ATT(CAG_{EXP})-3T construct. The QuickChange II XL Site-Directed Mutagenesis Kit (Stratagene) was used to change the ATG in front of the V5 tag to an ATC to generate the ATC-V5-(CAG_{105})-3T construct which contains no ORFs.

To generate the CAA_{EXP} constructs, a CAA repeat was amplified by PCR using the ACA_{13} and TTG_{15} primers. PCR products varied in size. A gel slice containing 200- to 550-bp fragments (67−183 repeats) was purified, and the resulting fragments were cloned into the pSC-A-amp/kan vector using Strataclone PCR Cloning Kit (Stratagene). Clones were sequenced, and desirable CAA repeats were excised and subcloned into pcDNA3.1/6Stops-3T. The resulting constructs were sequenced and CAA_{125}(−ATG), CAA_{90}(−ATG), and CAA_{38}(−ATG) constructs were obtained. Modified versions of these constructs containing an ATG in the polyGln frame [CAA_{125} (+ATG), CAA_{90}(+ATG), and CAA_{38}(+ATG)] were created using site-directed mutagenesis (Stratagene).

To generate CTG_{EXP}(Cys-myc/His), CTG_{EXP}(Ala-myc/His), and CTG_{EXP}(leucine-myc/His) constructs, a fragment of expanded CTG repeats was subcloned into pcDNA3.1/*myc*-His (A, B, and C respectively), and each of the three reading frames was C-terminally tagged. In the three resulting constructs, there is no ORF in each of three frames, and polyC, polyAla, and polyleucine are tagged individually in frame with an *myc*-His tag.

The 3′ flanking sequence of *DM1* in the CAG direction was amplified by PCR using 5′ primer (5′-CTCGAGGCTACAAG-GACCCTTCGAG-3′) and 3′ primer (5′-CCTGAACCCTAGA-ACTGTCTTCGACT-3′) and was cloned into a PCR cloning vector, pCR4-TOPO (Invitrogen). The XhoI/PmeI fragment of pCR4-DM1-3′ was subcloned downstream of CAG repeats of ATT(CAG_{EXP})-3T to generate the CAG-DM1-3′ construct containing expanded CAG repeats and the 3′ flanking sequence of *DM1*.

PCR-mediated mutagenesis was used to create several constructs in which the ATT or ATC alternative start codons were altered to ACT and ACC, respectively. All constructs were created using the 3′ primer (5′-TAGAAGGCACAGTCGA-GGCTGATCAG CGGGTTT-3′) and a unique 5′ primer. The ACT(CAG_{105})-3T primer (5′-AGTTAAGCTTAGCTAGGTA-ACTAAGTAACTAGAACTCAGCA-3′) was used to generate the ACT(CAG_{EXP})-3T construct from the ATT(CAG_{EXP})-3T template. The HDL2-3T:[ATT,ATC] construct was used as template to generate the HDL2-3T:[ATT,ACC], HDL2-3T:[ACT,ATC], and HDL2-3T:[ACT,ACC] constructs from the HDL2:[ATT,ACC] 5-1 (5′-AGTTAAGCTTAGCTAGGTAAC-TAAGTAACTAGAATTTCCTGCACAGAAACCACCTT-3′), HDL2:[ACT,ATC] 5-1 (5′-AGTTAAGCTTAGCTAGGTAA-CTAAGTA ACTAGAACTTCCT-3′), and HDL2:[ACT,ACC] 5-1 (5′-AGTTAAGCTTAGCTAGG TAACTAAGTAACTAGAA-CTTCCTGCACAGAAACCACCTT-3′) primers, respectively. Likewise, the SCA3:[ACT] construct was generated from SCA3 template and the SCA3:[ACT] 5-1 (5′-AGTTAAGCTTAGCTAGGTAAC-TAAGTAACTAGAAC TAACA-3′) primer. The HD: 5-1 primer (5′-AGTTAAGCTTAGCTAGGT AACTAAGTAACTAGAACT-TCGA-3′) was used along with the HD-3T:[ATT] template to generate the HD-3T:[ACT] construct.

The IntR(CAG_{EXP})-3T construct was made by inserting the duplex primer containing $(CAG)_{18}$CGCGCG into the EcoRI site of the ATT(CAG_{69})-3T construct. The extra nucleotides between $(CAG)_{18}$CGCGCG and $(CAG)_{69}$ were removed by

digestion with EcoRI, followed by treatment with mung bean nuclease and DNA ligase.

All PCR reactions to generate the above constructs were performed with Pfx polymerase (#11708-013; Invitrogen) to mitigate PCR-induced mutations. The above PCR products were digested with HindIII and PmeI, gel purified, and cloned into pcDNA3.1 containing the 6X STOP cassette. The integrity of all constructs was confirmed by sequencing.

Production of Polyclonal Antibodies. The polyclonal antibodies were generated by New England Peptide. The α-$SCA8_{GCA\text{-}Ala}$ antisera were raised against a synthetic peptide corresponding to the C terminus of the predicted polyAla frame of SCA8 in the CAG direction (VKPGFLT). The α-$DM1_{CAG\text{-}Gln}$ antisera were raised against a synthetic peptide corresponding to the C terminus of a predicted glutamine frame of DM1 in the CAG direction (SPAARGRARITGLEL).

Cell Culture and Transfection. HEK293T cells were cultured in DMEM supplemented with 10% fetal bovine serum and incubated at 37 °C in a humid atmosphere containing 5% CO_2. DNA transfections were performed using Lipofectamine 2000 Reagent (Invitrogen) according to the manufacturer's instructions.

Murine neuroblastoma (N2a) cells were cultured in unsupplemented DMEM and incubated at 37 °C in a humidified, 5% CO_2 incubator. DNA transfections were performed using TransIT-Express (Mirus) according to the manufacturer's instructions.

DM1 patient myoblasts with 50–70 CTG repeats, along with a normal control, were cultured in skeletal muscle growth medium (SGM; PromoCell) with Glutamax (Invitrogen), gentamicin (50 units/mL), decomplemented fetal calf serum, and the provided supplemental mix. Cells were grown to approximately 70% confluence on collagen-coated coverslips in six-well tissue-culture plates.

RNA Transfections. Plasmid DNA was linearized using PvuII. Transcription, capping, and polyadenylation was performed using 1 μg of DNA with the mScript mRNA Production System (Epicentre). Transfections were performed in six-well plates using 3 μg mRNA and 10 μL Lipofectamine 2000 (Invitrogen) per well. Cell lysates were collected 18–24 h posttransfection, and immunoblots were performed as described.

Immunofluorescence. The subcellular distribution of homopolymer proteins was assessed in transfected HEK293T cells by immunofluorescence. Cells were cultured on coverslips in six-well tissue-culture plates and transfected with plasmids the next day. Forty-eight hours posttransfection, cells were fixed in 4% paraformaldehyde in PBS for 30 min and permeabilized in 0.5% Triton X-100 in PBS for 10 min. The coverslips were blocked in 1% normal goat serum (NGS) in PBS for 30 min. After blocking, the cells were incubated for 1 h at 37 °C in blocking solution containing the rabbit anti-His (1:100), rat anti-HA (1:100), and mouse anti-Flag (1:200) primary antibodies. The coverslips were washed three times in PBS and incubated for 1 h at 37 °C in blocking solution containing secondary antibodies. Goat anti-rabbit conjugated to Cy3 (Jackson ImmunoResearch), goat anti-rat conjugated to Cys5 (Jackson ImmunoResearch), and goat anti-mouse conjugated to Alexa Fluor 488 (Invitrogen) were used at a dilution of 1:200.

DM1 patient myoblasts grown on coverslips were fixed in 4% paraformaldehyde for 30 min and blocked with 5% NGS for 1 h. Next, the cells were incubated with α-$DM1_{CAG\text{-}Gln)}$ (1:5,000) at 4° C overnight. Cells then were washed and incubated with goat anti-rabbit conjugated to Cy3 (Jackson ImmunoResearch) for1 h at room temperature, in darkness. Slides were washed three times for 5 min in 1× PBS, mounted with Vectashield HardSet mounting medium with DAPI (Vector Laboratories, Inc.) and coverslipped.

For mouse and human tissues, 9-μm cryosections were fixed in 4% paraformaldehyde for 15 min. Heat-induced epitope retrieval (HIER) was employed by steaming sections in citrate buffer, pH 6.0, at 90 °C for 20 min. HIER was used in all immunofluorescence tissue experiments except for $SCA8_{GCA\text{-}Ala}$ mouse and human experiments in which antigen retrieval was omitted altogether. A nonserum block (Biocare Medical) was applied to all tissues (except the SCA8 mouse tissue in which 10% NGS in a 0.3% Triton X-100 was used to block nonspecific immunoglobulin binding) and was allowed to incubate at room temperature for 1 h. The primary antibody (or antibodies, if double- or triple-labeled) was diluted either in a 1:5 solution of the nonserum block or in a solution of 5% NGS in PBS containing 0.3% Triton X-100 and was incubated at 4 °C overnight. Tissues then were incubated for 1 h in a 1:2,000 dilution of IgG–TRITC, in the dark at room temperature. If needed, a Sudan-black autofluorescence block was applied to the tissue for 1 h at room temperature in the dark (2). Staining was observed, and pictures were taken on an Olympus FluoView 1000 I×2 inverted confocal microscope. All mutant and control images were adjusted in unison, to the same specifications, and in a linear fashion, for intensity and contrast when deemed necessary.

Quantitative Imaging. Cells were counted using Image-Pro Plus 6.0 (Media Cybernetics) using an automatic dark-object count method. Images first were digitally annotated by a pathologist to separate coagulated blood from cardiac myocytes into distinct regions of interest. The images were split into grayscale versions of their blue color channel (nuclear stain color) and red color channel (positive marker color). The grayscale images were analyzed in each region of interest with Image-Pro Plus using a dark-object count method with automatic thresholding for nuclear intensity.

Immunohistochemistry. Myotonic dystrophy (DM) mutant and control mice were perfused in 10% formalin, and tissue was harvested and embedded in paraffin. Five-micrometer sections were deparaffinized in xylene and rehydrated through graded alcohol, incubated with 90% formic acid for 5 min, and washed with distilled water for 30 min. HIER was performed by steaming sections in citrate buffer, pH 6.0, at 90 °C for 20 min. To block nonspecific avidin-D/biotin binding, the Avidin-D/Biotin block (Vector Laboratories, Inc.) was used as described. To block nonspecific immunoglobulin binding, a nonserum block (Biocare Medical) was applied for 30 min. Primary 1C2 antibody was applied at a dilution of 1/12,000 in nonserum block (Biocare Medical) and incubated overnight at 4 °C. Biotinylated secondary α-mouse IgG purified in goat (Vector Laboratories, Inc.) was applied at a dilution of 1:200 for 30 min at room temperature. ABC reagent (Vector Laboratories, Inc.) was used for detection with Chromagen SG (Vector Laboratories, Inc.) for 10 min and counterstained with nuclear fast red.

Leukocyte cell pellets were isolated from peripheral blood of DM1 patients and controls. The cell pellets were fixed in 10% neutral buffered formalin for 30 min, washed, encapsulated in HistoGel (Richard-Allen), and placed in 70% ethanol. The pellets then underwent a short, 2-h cycle in the tissue processor and were embedded in paraffin blocks. Five-micrometer sections were cut, deparaffinized, and hydrated to water. HIER was employed with steam and Reveal Decloaker (Biocare Medical). A nonserum block (Biocare Medical) was applied for 30 min to prevent nonspecific immunoglobulin binding. The nonserum block 1:10 in PBS was used to dilute the α-$DM1_{CAG\text{-}Gln)}$ antibody to a concentration of 1:10,000. Slides were incubated overnight at 4 °C and washed three times for 5 min in PBS. The secondary antibody, DyLight 488-conjugated AffiniPure goat anti-rabbit (Jackson Immunoresearch) was applied and incubated for 2 h in the dark at room temperature at a concentration of 1:1,000. Slides were washed three times for 5 min in PBS, mounted with

Vectashield HardSet Mounting Medium with DAPI (Vector Laboratories, Inc.), and coverslipped. Staining was observed, and pictures were taken on an Olympus FluoView 1000 I×2 inverted confocal microscope. For consistency in Fig. S10*C*, Olympus FluoView software was used to reassign the green 488λ captured signal to red.

Cell Death Analysis. N2a cells were transfected with CAG_{105} and CAA_{90} (control) constructs with or without ATG initiation codons and were grown in DMEM without FBS. After 48 h, cells were stained with annexin-V to detect phosphatidylserine (an early marker of apoptosis) on the plasma membrane and were sorted by flow cytometry. For flow cytometric annexin-V analysis, floating cells were collected and combined with trypsinized, adherent cells in ice-cold PBS supplemented with 2% FBS. After washing, cells were resuspended in annexin-binding buffer (BD Biosciences), vortexed, and stained with annexin-V–APC (BD Biosciences) according to manufacturer's instructions. Cells were placed on ice and sorted immediately on a BD FACScalibur flow cytometer. Thirty thousand total events were collected. Two independent experiments, each done in triplicate, were performed, and data were combined and normalized to the ATT(CAA_{90}) average. Statistics were performed using a one-way ANOVA, $F(3, 20) = 0.07$, $P = 0.0003$, and post hoc P values were calculated with a one-tailed t-test.

Labeling and Immunoprecipitation of PolyGln, PolyAla, and PolySer Proteins with [^{3}H]-Amino Acids. HEK293T cells were cultured in DMEM supplemented with 10% fetal bovine serum and were transfected with CAG expansion constructs. Twenty-four hours posttransfection, the DMEM-based medium was replaced with the glutamine-, alanine-, and serine-free MEM medium (Invitrogen) supplemented with 10% fetal bovine serum. Then [^{3}H]-glutamine, [^{3}H]-alanine, or [^{3}H]-serine was added to the respective wells at 25 μCi/mL, and the cells were incubated for 16 h at 37 °C. Cells in culture plates were rinsed with PBS and lysed in RIPA buffer [150 mM NaC1, 1% sodium deoxycholate, 1% Triton X-100, 50 mM Tris-HC1 (pH 7.5), 1× protease inhibitors (Roche)] for 45 min on ice. The cell lysates were centrifuged at 16,000 × *g* for 15 min at 4 °C, and the supernatant was collected. To immunoprecipitate ^{3}H-labeled protein, 500 μg of tissue lysate was incubated with the desired antibody at 4 °C for 2 h and then with protein G-Sepharose at 4 °C overnight. Protein G-Sepharose was washed three times with RIPA buffer. Bound proteins were eluted from the beads with 1× SDS sample buffer, incubated at 90 °C for 10 min, and analyzed by protein gel electrophoresis.

Protein Labeling with [^{35}S]-Methionine. A T7-coupled transcription and translation kit (Promega) was used with [^{35}S]-methionine (MP Biomedical) for protein labeling. Labeled proteins were separated by PAGE in parallel on two separate gels. One gel was dried for 1.5 h at 70 °C and subsequently was used to generate a fluorograph. The other gel was transferred onto a nitrocellulose membrane and probed with the 1C2 monoclonal antibody specific to polyGln expansion tracts.

Preparation of *N*-Formyl-[^{35}S]Methionyl-tRNA$_i^{Met}$. Preparation of aminoacyl-tRNA synthetases and Met-tRNA$_i^{Met}$ transformylase from *Escherichia coli* MRE 600 (ATCC) was performed following Stanley's method (3). The reaction mixture included the *E. coli* MRE 600 fraction (1.5 mg/mL), 0.1 mM each of 19 unlabeled amino acids (minus methionine), L-[^{35}S]methionine (2 mCi/mL), 0.2 mM calcium leucovorin, unfractionated bovine liver tRNA (2 mg/mL), 10 mM ATP, 1 mM CTP, 15 mM $MgCl_2$, and 50 mM sodium cacodylate, pH 7.4. The aminoacylation reaction was carried out at 37 °C for 30 min. The reaction mixture was extracted with phenol, precipitated with ethanol, lyophilized, and suspended in water. *N*-formyl-[^{35}S]methionyl-tRNA$_i^{Met}$ was added at a concentration of 100 μg/mL for in vitro translation using a T7-coupled transcription and translation kit. The reaction products were analyzed as described above.

Immunoprecipitation. The protein concentration of tissue lysates was determined using the protein assay dye reagent (Bio-Rad). To immunoprecipitate polyGln protein, 500 μg of tissue lysate was incubated with rabbit polyclonal anti-His antibody at 4 °C for 2 h and then with protein G-Sepharose at 4 °C for overnight. Protein G-Sepharose was washed three times with RIPA buffer. Bound proteins were eluted from the beads with 1× SDS sample buffer, incubated at 90 °C for 10 min, and analyzed by immunoblotting.

Immunoblotting. Cells in each well of a six-well tissue-culture plate were rinsed with PBS and lysed in 300 μL RIPA buffer for 45 min on ice. DNA was sheared by passage through a 21-gauge needle. The cell lysates were centrifuged at 16,000 × *g* for 15 min at 4 °C, and the supernatant was collected. The protein concentration of the cell lysate was determined using the protein assay dye reagent (Bio-Rad). Twenty micrograms of protein were separated in a 4–12% or 10% NuPAGE Bis-Tris gel (Invitrogen) and transferred to a nitrocellulose membrane (Amersham). The membrane was blocked in 5% dry milk in PBS containing 0.05% Tween-20 and probed with the anti-His antibody (1:500) or 1C2 antibody (1:1,000) in blocking solution. After the membrane was incubated with anti-rabbit or anti-mouse HRP-conjugated secondary antibody (Amersham), bands were visualized by the ECL plus Western Blotting Detection System (Amersham).

The distinctive migration pattern of the polyAla may result from initiation at multiple sites and/or the formation of higher molecular weight protein complexes resistant to dissociation in SDS and 8 M urea. The polySer proteins might undergo beta elimination and addition from side chains of another polymer to create 3D-crosslinked polypeptides that would be extremely large (4).

MS. To immunoprecipitate polyAla protein for MS, transfected HEK293T cell lysate from five 150-mm dishes was incubated with mouse monoclonal antibody against the C-terminal HA tag at 4 °C for 2 h and then with protein G-Sepharose at 4 °C overnight. Protein G-Sepharose was washed three times with RIPA buffer. Bound proteins were eluted from the beads with 8 M urea. Samples were separated by parallel SDS-PAGE 4–15% Criterion Tris-HCl gels (Bio-RAD), one for MS and the other for immunoblotting. Protein bands of interest were excised manually after visualizion with Imperial Protein Stain (Thermo Scientific). Specified bands were cut out and subjected to in-gel trypsin digestion using standard methods (5), and extracted peptides were cleaned up further using "stage" tips (6). Mass analysis was performed using an LTQ-Orbitrap XL mass spectrometer (Thermo Scientific). Peptides derived from in-gel digestion were separated by reversed-phase chromatography with nanoHPLC. The gradient was 2–40% acetonitrile in H_2O containing 0.1% formic acid for more thn 60 min. Full MS scans were generated in the orbital trap at 60,000 resolution for 400 m/z. MS/MS scans were performed in a data-dependent manner using an inclusion list based on predicted tryptic peptides in the LTQ ion trap using collision induced dissociation (CID). Data were searched with SEQUEST V27 with semi-trypsin specificity, cysteine carbamidomethylation as a fixed modification, and N-terminal methionine oxidation as a variable modification. The search was performed against the combined database consisting of the National Center for Biotechnology Information human database V200906 and its reversed complement and an additional list of all possible proteins that could be initiated anywhere in the polyAla frame of the IntR (CAG_{EXP})-3T construct with or without an N-terminal methionine, which totaled >76,000 entries. Identified proteins were organized using Scaffold (www.proteomesoftware.com), and peptide

probabilities were calculated within this program using Peptide Prophet (7). The identification output was filtered using a precursor mass tolerance at 7 ppm.

In Vitro Translation. In vitro translation was performed using coupled reticulocyte lysate systems (Promega). Coupled transcription/translation reactions (50 μL) contained 50% lysate, 1 μL T7 RNA polymerase, 20 μM amino acid mixtures, 1 μL (40 units) RNasin ribonuclease inhibitor (Promega), and 1 μg plasmid DNA. Incubation was performed at 30 °C for 90 min. Ten percent of each reaction was analyzed by Western blotting.

Production and Purification of Lentiviral Vectors and Transduction of HEK293T Cells. HEK293T cells were plated on 150-mm tissue-culture dishes and transfected the following day when cells were 80–90% confluent. Thirty micrograms of the transducing vector, 20 μg of the packaging vector ΔNRF, and 10 μg of the VSV envelope pMD.D were cotransfected by calcium phosphate-mediated transfection. The medium was changed the next day, and conditioned media were collected 48 and 72 h after transfection. Conditional medium then was cleared by filtering through a 0.45-μm filter. The viral particles were concentrated by ultracentrifugation at 50,000 × *g* for 2 h. The pellet was resuspended in 20 μL of 1× HBSS and stored at −70 °C. HEK293T cells were seeded into each well of a six-well plate and transduced the next day. Transduced cells were analyzed by Western blotting after 5 d.

Injection of Mouse Brain with Lentiviral Vectors. Six-week-old FVB mice were anesthetized by intramuscular injection using a combination of ketamine and xylazine. Two microliters of lentiviral vectors [5×10^9 functional viral titer (TU)/mL] were injected into mouse striatum and cerebellum. The mouse was mounted in a stereotactic frame, and its head was shaved. A midline sagittal incision was made, and the cranium was exposed. For each injection site, a burr hole was drilled, and a Hamilton syringe was inserted to the depth described below the dura, plus an additional 0.5 mm. After 2 min, the syringe was retracted 0.5 mm to form a slight pocket in the parenchyma. After a pause of at least 2 min for pressure equalization, the injection was performed manually at an approximate rate of 0.5 μL/min. The syringe was left in place for an additional 3 min after injection and then was withdrawn over a period of 2 min or more. Once injections were complete, the scalp was sutured; the mouse was kept under a warming lamp until it recovered from the anesthesia and then was returned to standard housing. Animal care followed the guidelines set by the Institutional Animal Care and Use Committee at the University of Minnesota.

Polysome Profiling. Transfected HEK293T cells in 150-mm dishes were treated with cycloheximide (100 μg/mL) for 5 min and harvested by trypsinization. The cell pellet was resuspended in 375 μL of low-salt buffer [10 mM NaCl, 20 mM Tris (pH 7.5), 3 mM $MgCl_2$, 1 mM DTT, and 200 units RNAse inhibitor) and allowed to swell for 2 min. Then 125 μL lysis buffer (0.2 M sucrose, 1.2% Triton X-100 in low-salt buffer) was added, and the cells were homogenized using 15 strokes in a Dounce homogenizer using the tight-fitting pestle. Lysate was centrifuged at 16,000 × *g* for 1 min, and the nuclear pellet was removed. Cytoplasmic extract (1.5 mg measured at A_{260}) was layered onto a 5-mL, 0.5–1.5 M sucrose gradient and centrifuged at 200,000 × *g* in a Beckman SW50 rotor for 80 min at 4 °C. The gradients were fractionated using an ISCO density gradient fractionator monitoring absorbance at 254 nm. Ten fractions from each sample were collected into tubes containing 50 μL of 10% SDS.

Northern Analysis. The RNA from each fraction of the sucrose gradient was extracted using Tri-reagent (Sigma). For Northern blot analysis, an equal volume of the RNA from each fraction was separated on a glyoxal gel, blotted to a nylon membrane, and probed with a [^{32}P]ATP-labeled oligonucleotide (5′-TAGAAGGCACAGTCGAGGCTGATCAGCGGGTTTAAACTCAAT-3′) complementary to the 3′ end of the CAG-containing transcripts. Blots subsequently were probed with a [^{32}P]dATP-labeled GAPDH cDNA probe.

RT-PCR. For detection of CAG and CAA expansion transcripts, cells were transfected using Lipofectamine 2000 (Invitrogen) as described above. RNA and protein were harvested using Trizol (Invitrogen). Approximately 45 μg of RNA from each sample was resuspended in 50 μL diethylpyrocarbonate distilled water. The RNA sample was treated with an RNase-free DNase Set (Qiagen) and the RNeasy Plus Mini Kit (Qiagen) to remove DNA. A Superscript II Reverse Transcriptase System (Invitrogen) and the Myc Tag GSP Primer (5′-CAGATCCTCTTCTGAGATGAGTTTTTGTTC-3′) were used to reverse transcribe the RNA, and PCR was performed using the 336 forward (5′-ACCCAAGCTGGCTAGTTAAGC-3′) and 336 reverse (5′-TGTCGTCGTCGTCCTTGTAA-3′) primers at 95 °C for 2 min, then 35 cycles of 94 °C for 45 s, 59.5 °C for 30 s, 72 °C for 45 s, and a 6-min extension at 72 °C. Control reactions were performed using the β-actin forward (5′-TCGTGCGTGACATTAAGGAG-3′) and β-actin reverse (5′-GATCTTCATTGTGCTGGGTG-3′) primers. PCR conditions were 95 °C for 2 min, then 35 cycles of 94 °C for 45 s, 59.5 °C for 30 s, 72C for 45 s, followed by a 6-min final extension at 72 °C. PCR products were separated on a 1% agarose gel. For detection of CAG expansion transcripts in DM humans and mice, total RNA was extracted from frozen tissues with Trizol (Invitrogen) following incubation with lysis buffer and 0.5 mg/mL proteinase K, as well as precipitation and DNAse treatment. For strand-specific RT-PCR, a linker sequence (lk) was attached (5′-CGACTGGAGCACGAGGACACTGA-3′) to the 5′ end of primers specific for the antisense strand of dystrophia myotonica-protein kinase (DMPK): (*i*) 5′-CGCCTGCCAGTTCACAACCGCTCCGAGCGT-3′ or (*ii*) 5′-GACCATTTCTTTCTTTCGGCCAGGCTGAGGC-3′. Three micrograms of RNA were reverse-transcribed with Superscript III (Invitrogen) at 55 °C. PCR against the anti1B, antiN3, and antiA2 regions was carried out using the CTCF1b (5′-GCAGCATTCCCGGCTACAAGGACCCTTC-3′), antiN3 (5′-GAGCAGGGCGTCATGCACAAG-3′) and the antiA2 (5′-TAGGTGGGGACAGACAAT -3′) primers, respectively. The linker primer was used in all reactions. The PCR reactions were done using the following conditions: antiB1, 94 °C for 5 min, then 30 cycles of 94 °C for 30 s, 67 °C for 30 s, and 72 °C for 1 min followed by 10 min at 72 °C; antiN3, 94 °C for 5 min, then 30 cycles of 94 °C for 30 s, 63 °C for 30 s and 72 °C for 1 min followed by 10 min at 72 °C ; antiA2, 94 °C for 5 min, then 40 cycles of 94 °C for 30 s, 57 °C for 30 s, and 72°C for 1 min followed by 10 min at 72 °C. GAPDH was amplified using the GFw (5′-AGGTCGGTGTGAACGGATTTG-3′) and GRev (5′-TGTAGACCATGTAGTTGAGGTCA-3′) primers at 94 °C for 5 min and then 24 cycles of 94 °C for 30 s, 65C for 30 s, and 72 °C for 1 min followed by 10 min at 72 °C.

1. Moseley ML, et al. (2006) Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8. *Nat Genet* 38:758–769.
2. Schnell SA, Staines WA, Wessendorf MW (1999) Reduction of lipofuscin-like autofluorescence in fluorescently labeled tissue. *J Histochem Cytochem* 47:719–730.
3. Stanley WM, Jr., (1974) Specific aminoacylation of the methionine-specific tRNA's of eukaryotes. *Methods Enzymol* 29:530–547.
4. Tinette S, Feyereisen R, Robichon A (2007) Approach to systematic analysis of serine/threonine phosphoproteome using Beta elimination and subsequent side effects: Intramolecular linkage and/or racemisation. *J Cell Biochem* 100:875–882.

5. Shevchenko A, Wilm M, Vorm O, Mann M (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem* 68:850–858.
6. Rappsilber J, Ishihama Y, Mann M (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal Chem* 75:663–670.
7. Keller A, Nesvizhskii AI, Kolker E, Aebersold R (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* 74:5383–5392.

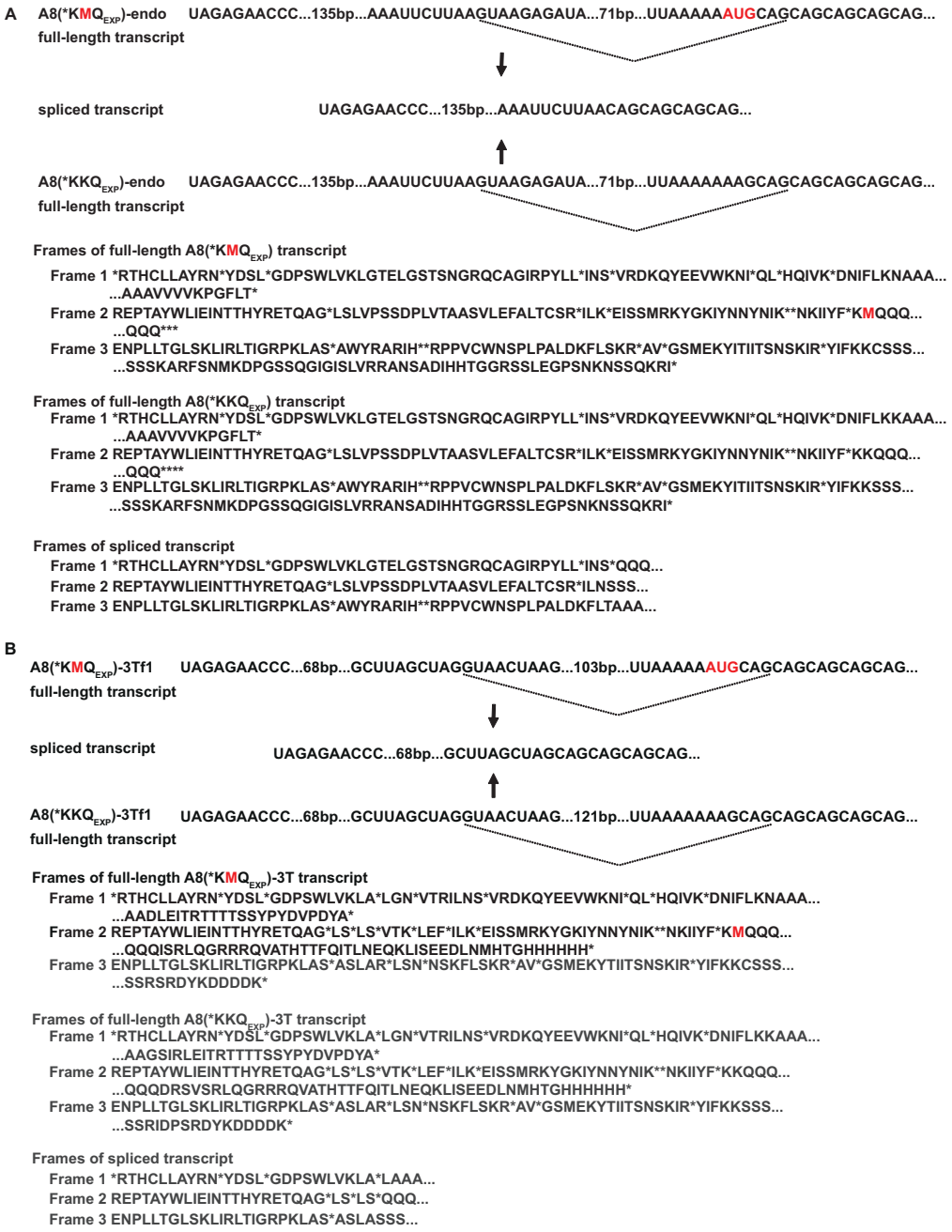


Fig. S1. Splicing of the SCA8 minigene in HEK293T cells. (*A*) Schematic representation of full-length and spliced transcripts of A8(*KMQ$_{EXP}$)-endo and A8 (*KKQ$_{EXP}$)-endo constructs showing 5′ splice sites located prior to CAG expansion and 3′ splice sites located within the CAG repeat region. The bottom panels show no ORF in any of three frames of full-length transcripts and spliced transcripts of both the A8(*KMQ$_{EXP}$)-endo and A8(*KKQ$_{EXP}$)-endo constructs. (*B*) Schematic representation of full-length and spliced transcripts of A8(*KMQ$_{EXP}$)-3Tf1 and A8(*KKQ$_{EXP}$)-3Tf1 constructs showing 5′ splice sites located in the 6X STOP cassette and 3′ splice sites located in the CAG repeat region. Bottom panels show no ORF in any of three frames of full-length transcripts and spliced transcripts of both the A8(*KMQ$_{EXP}$)-3Tf1 and A8(*KKQ$_{EXP}$)-3Tf1 constructs. Positions of the first 3′ stop codon in each frame are the same for the spliced and unspliced transcripts shown in *A* and *B*. RACE analysis performed on both the A8(KKQ$_{EXP}$)-endo and A8(KKQ$_{EXP}$)-3Tf1 transcripts confirmed that these transcripts lack AUG initiator codons.

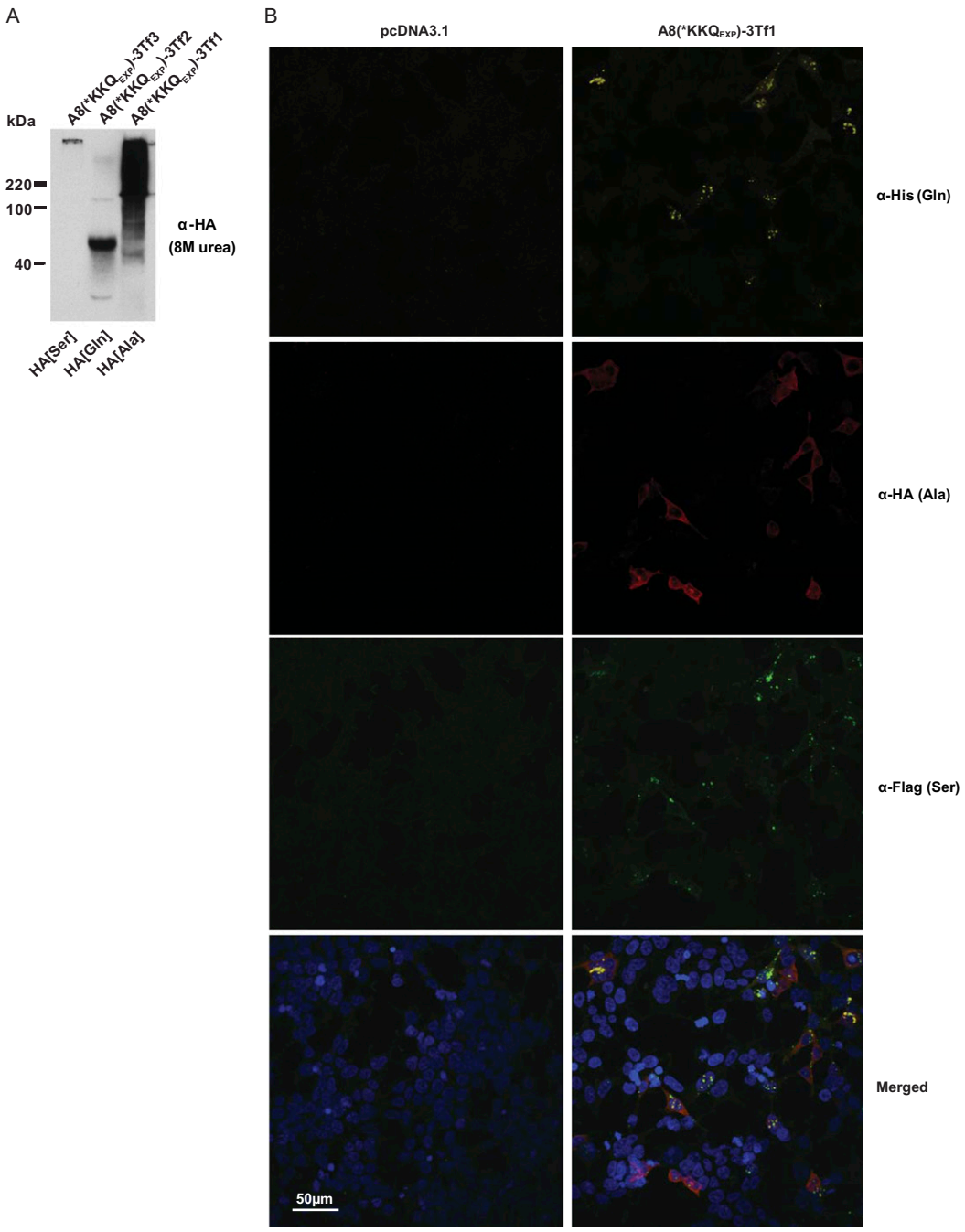


Fig. S2. (*A*) Immunoblot of HEK293T lysates after transfection with constructs in Fig. 1B and separation on a polyacrylamide gel in 8 M urea. (*B*) Immunofluorescence of tagged polyGln (α-His/cy3), polyAla (α-HA/cy5) and polySer (α-FLAG/FITC) proteins in transfected HEK293T cells shows RAN-translated proteins are expressed.

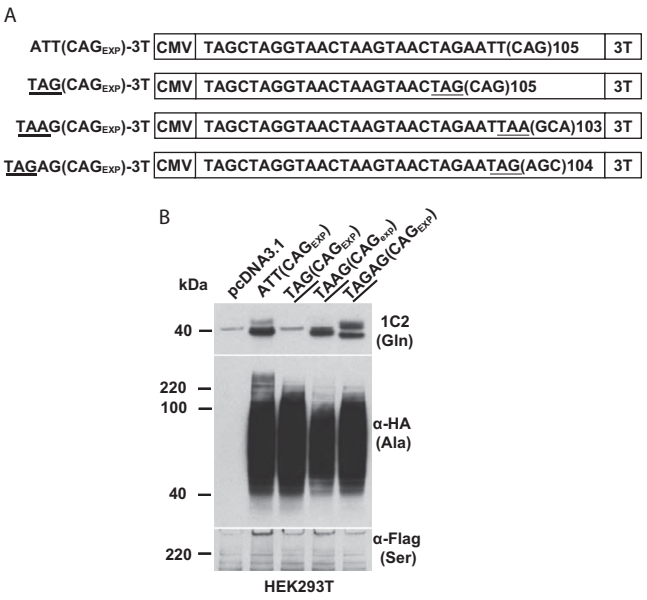
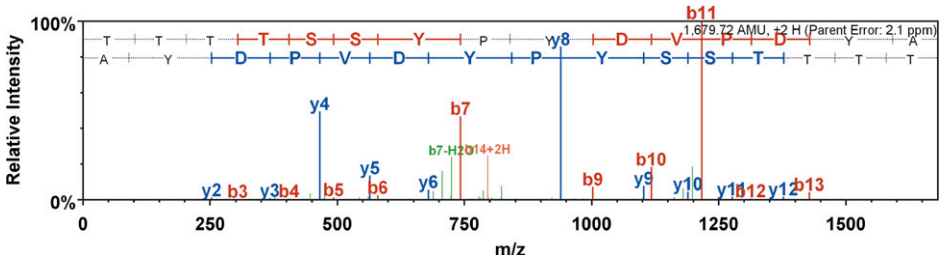


Fig. S3. Immunofluorescence of tagged polyGln (α-His/cy3), polyAla (α-HA/cy5), and polySer (α-FLAG/FITC) proteins in transfected HEK293T cells shows repeatassociated non-ATG (RAN)-translated proteins are expressed. (*A*) Schematic diagram showing constructs with and without stop codons immediately preceding repeats in the CAG(polyGln), GCA(polyAla), and AGC(polySer) frames. These constructs expressed full-length unspliced transcripts in transfected cells. All constructs contain 3′ epitope tags: myc-His (polyGln), HA (polyAla), and Flag (polySer). (*B*) Protein blots from transfected HEK293T cells probed with 1C2, α-HA, or α-FLAG antibodies.

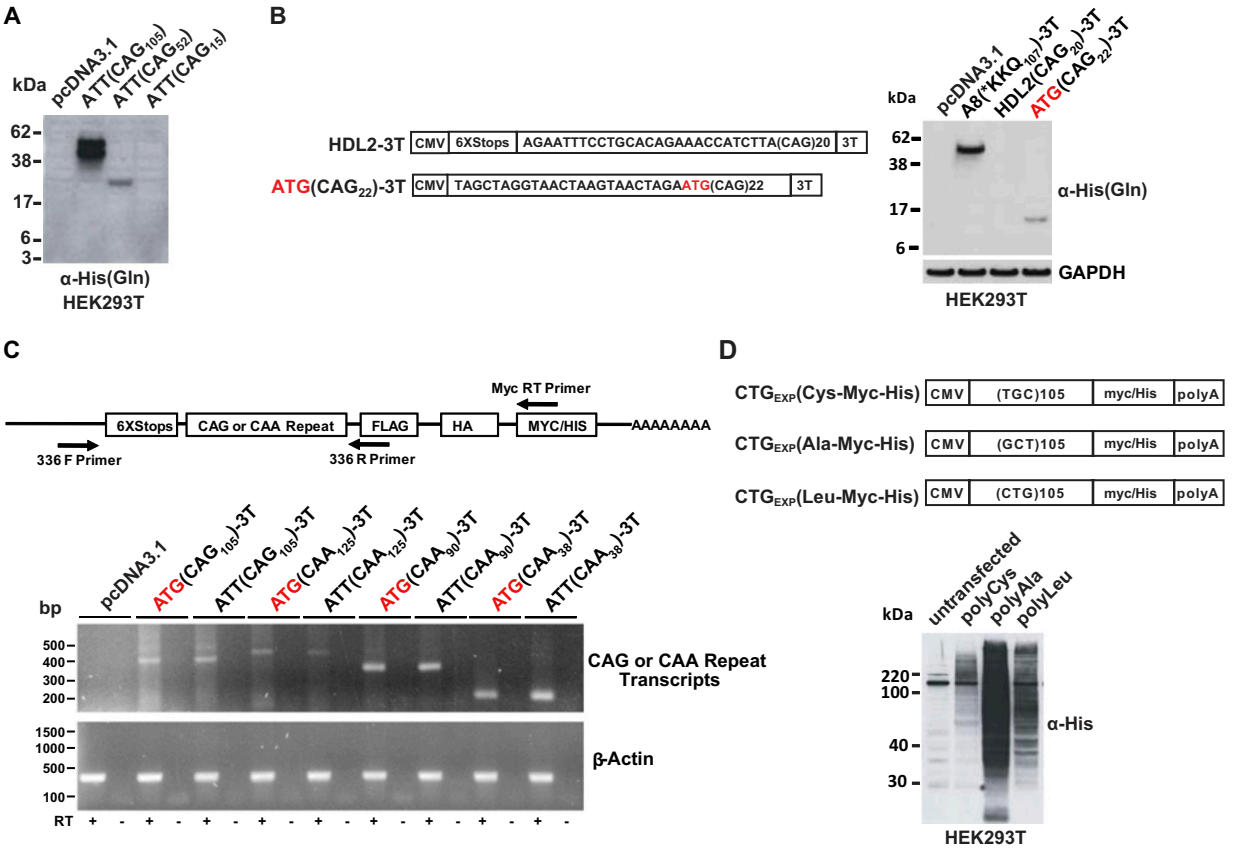


…	B Ions	B+2H	B-NH3	B-H2O	AA	Y Ions	Y+2H	Y-NH3	Y-H2O	…
1	102.1			84.0	T	1,680.7			1,662.7	15
2	203.1			185.1	T	1,579.7			1,561.7	14
3	304.2			286.1	T	1,478.6			1,460.6	13
4	405.2			387.2	T	1,377.6			1,359.6	12
5	492.2			474.2	S	1,276.5			1,258.5	11
6	579.3	290.1		561.3	S	1,189.5			1,171.5	10
7	742.3	371.7		724.3	Y	1,102.5			1,084.5	9
8	839.4	420.2		821.4	P	939.4			921.4	8
9	1,002.4	501.7		984.4	Y	842.4			824.3	7
10	1,117.5	559.2		1,099.5	D	679.3			661.3	6
11	1,216.5	608.8		1,198.5	V	564.3			546.3	5
12	1,313.6	657.3		1,295.6	P	465.2			447.2	4
13	1,428.6	714.8		1,410.6	D	368.1			350.1	3
14	1,591.7	796.3		1,573.7	Y	253.1				2
15	1,680.7	840.9		1,662.7	A	90.1				1

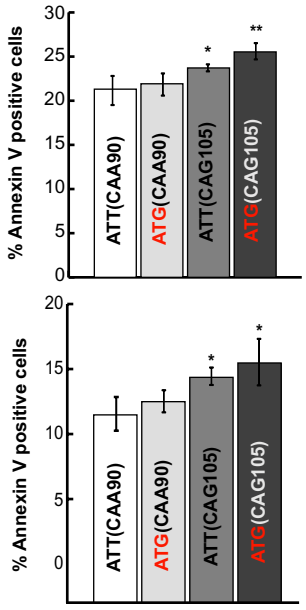
PolyA C-terminal peptide

TTTTSSYPYDVPDYA - Xcorr = 2.33, deltaCN = 0.39, Precursor chage +2, Precursor Error = 2.1 ppm

Fig. S4. Identification of C-terminal peptides of the polyAla protein by tandem MS. Representative identified spectrum of polyAla C-terminal peptide TTTTSSYPYDVPDYA. Matched b-ions are shown in red, and y-ions are shown in blue for the product ions of the associated precursor ion. Below each spectrum are fragmentation tables displaying matched product ions. The precursor ion was +2 charged with a mass error of 2.1 ppm. The SEQUEST Xcorr and ΔCN values were 2.33 and 0.39, respectively. Ten spectra with peptide probabilities at 95% were assigned to this peptide.



E



% Annexin Positive cells				
	ATT(CAA90)	ATG(CAA90)	ATT(CAG105)	ATG(CAG105)
plate 1	21.81	22.21	24.21	24.99
plate 2	22.64	23.01	23.59	26.61
plate 3	19.48	20.54	23.34	24.99
average	21.31	21.92	23.71333333	25.53
SD	1.638261273	1.260277747	0.447921124	0.935307436
p value		0.318077451	0.035186121	0.008961603

% Annexin Positive cells				
	ATT(CAA90)	ATG(CAA90)	ATT(CAG105)	ATG(CAG105)
Plate 1	10.32	12.51	14.57	13.65
Plate 2	12.8	11.65	13.64	15.57
Plate 3	11.3	13.31	14.87	17.18
average	11.4733333	12.49	14.36	15.46666667
SD	1.24905297	0.830180703	0.64132675	1.767267194
p value		0.152739725	0.011782951	0.016512077

Fig. S5. (*A*) Immunoblots of lysates after transfection with ATT(CAG)$_{EXP}$ with varying repeat lengths. (*B*) Immunoblot of polyGln protein from cells transfected with constructs of varying repeat lengths show RAN-translation of polyGln proteins occurs in the presence of 107 CAG repeats but not 20 CAG repeats. The ATG(CAG$_{22}$)-3T construct is a positive control. (*C*) Semiquantitative RT-PCR of CAG and CAA transcripts. (*Upper*) Schematic diagram depicting the RT-PCR strategy. The Myc-RT primer was used in a first-strand synthesis reaction, and 336 forward (336 F) and 336 reverse (336 R) primers were used for subsequent amplification of the repeat. (*Lower*) RT-PCR results for the CAG and CAA repeat constructs and β-actin control in the presence (+) or absence (–) of reverse transcriptase. (*D*) (*Upper*) CTG-containing constructs with myc/His tags in polycystine, polyAla, or polyleucine frames. (*Lower*) Corresponding immunoblots. (*E*) Summary of data for individual annexin-V experiments.

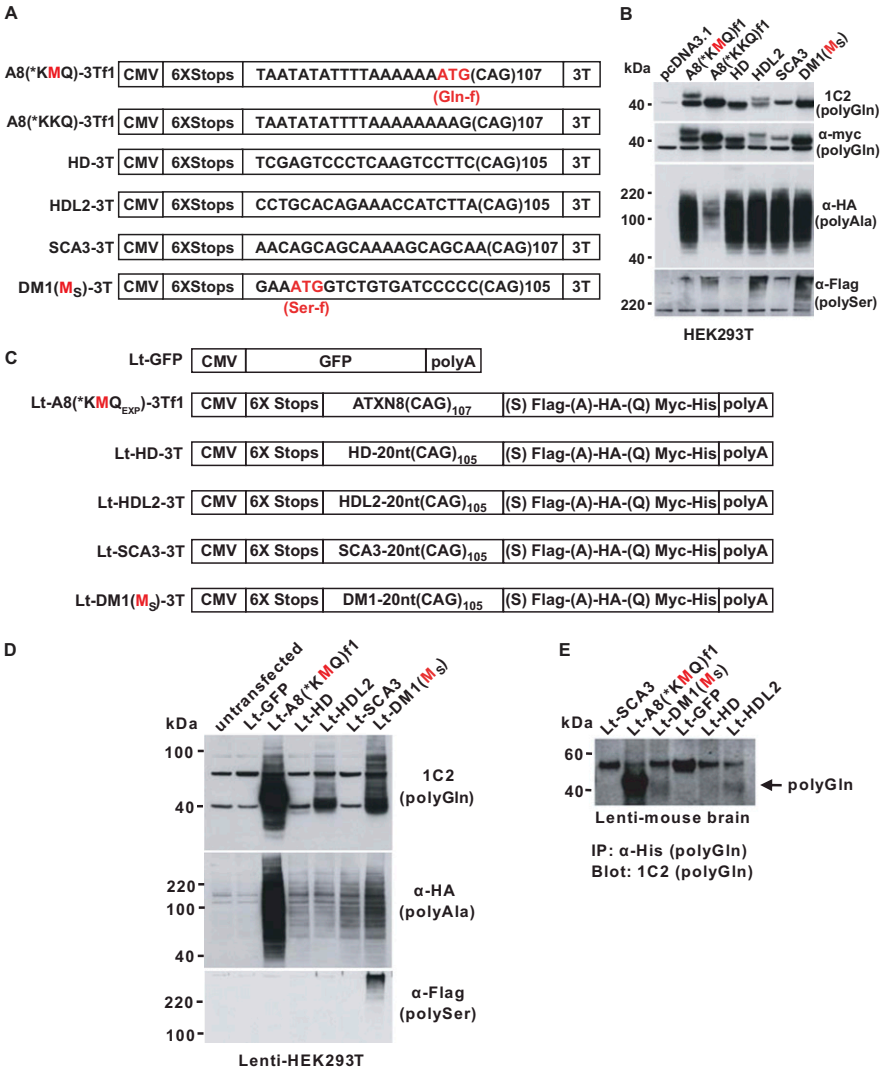


Fig. S6. Non-ATG translation in various sequence contexts. (*A*) Constructs with 20 bp of 5′ endogenous human flanking sequence upstream of repeat (CAG direction) at the HD, HDL2, DM1, and SCA3 loci. (*B*) Corresponding protein blots. (*C*) Schematic diagram showing triply tagged lentiviral constructs used for infection of HEK293T cells and mouse brains. All lentiviral constructs are in the CSII lentiviral vector. (*D*) Protein blots of HEK293T cells after lentiviral vector (Lt) infection with Lt-GFP, Lt-A8(*KMQ_{EXP})f1, Lt-HD, Lt-HDL2, Lt-SCA3, and Lt-DM1(M_S). Infected HEK293T cells show robust non-ATG translation of polyGln proteins for Lt-HDL2 and Lt-DM1. PolyAla but not polySer is expressed from all four constructs (Lt-HD, Lt-HDL2, Lt-SCA3, and Lt-DM1) without an ATG in the polyAla frame. (*E*) Protein blots of mouse cerebellar extracts after lentiviral vector infection and immunoprecipitation. The ∼40 kDa 1C2-positive protein was detected in cerebellar lysates injected with Lt-A8(*KMQ_{EXP}), Lt-HDL2, and Lt-DM1(M_S), but not Lt-HD, Lt-SCA3, or Lt-GFP. Two FVB animals were injected with each of these viruses, and 4 wk postinjection, polyGln-tagged protein was immunoprecipitated with anti-His antibody and probed with 1C2. As shown, polyGln-tagged protein was immunoprecipitated from tissue infected with the ATG-positive control virus Lt-A8(*KMQ_{EXP}) as well as from tissue infected with the Lt-DM1 and Lt-HDL2 lacking an ATG in the glutamine frame, although at a substantially lower level.

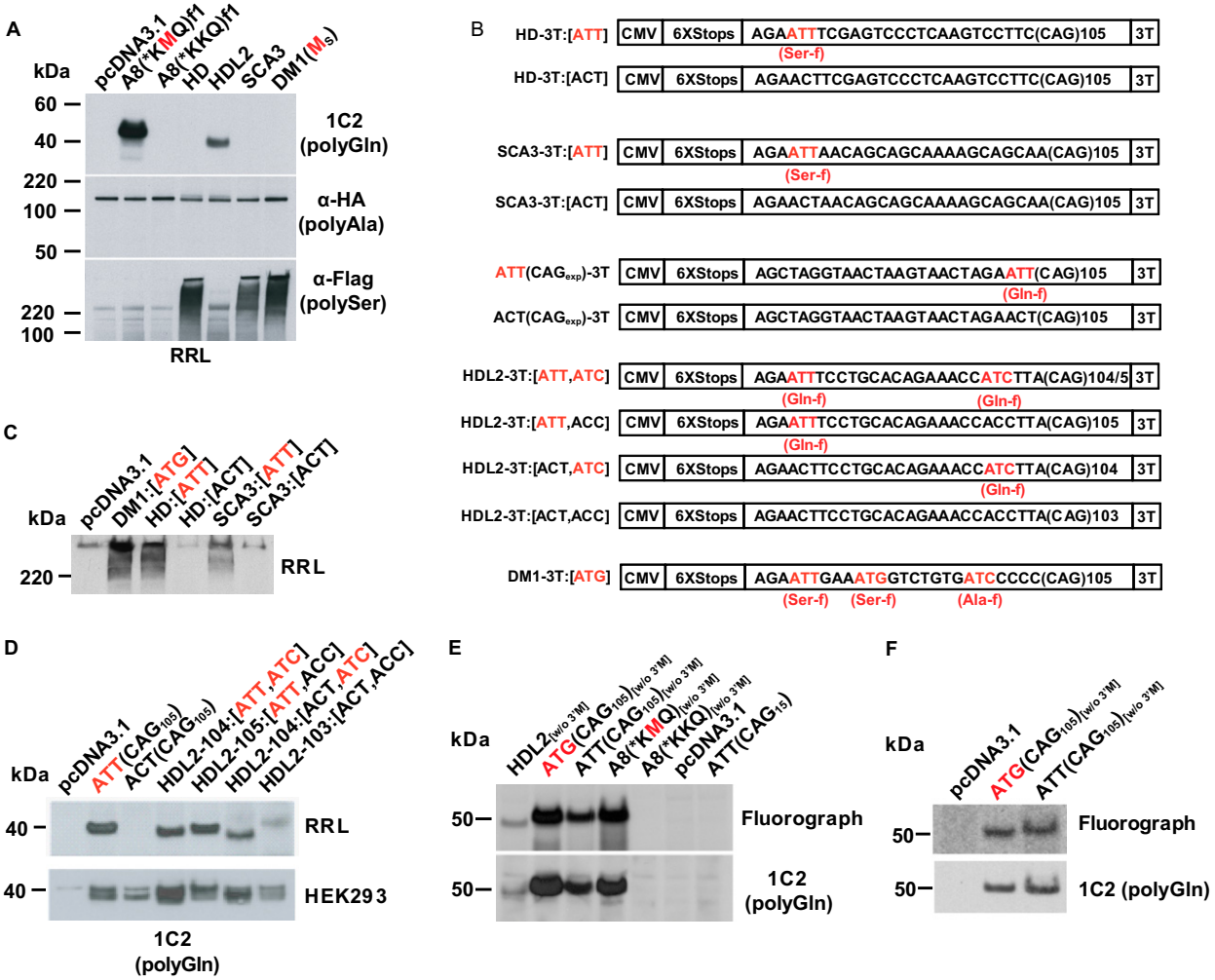


Fig. S7. RAN translation in cell-free rabbit reticulocyte lysates (RRLs) is less permissive and requires alternative start codons. (*A*) Protein blots after coupled in vitro transcription–translation of constructs in Fig. S6*A* using RRL. (*B*) Schematic diagrams of repeat constructs with and without ATT or ATC alternative start codons in the Gln (Gln-f) or Ser (Ser-f) frames, respectively. (*C*) Protein blots from RRL with DM1-3T, HD-3T, and SCA3-3T constructs with and without 5′ alternative start codons. (*D*) Protein blots of samples prepared using an in vitro RRL transcription/translation reaction (*Upper*) or from HEK293T cells (*Lower*) transfected with ATT(CAG_{105})-3T, ATC(CAG_{105})-3T, and HDL2-3T constructs. (*E*) (*Upper*) Fluorograph showing [^{35}S]-methionine incorporation. (*Lower*) Protein blot of the same in vitro translation products probed with the 1C2 antibody. (*F*) (*Upper*) Fluorograph showing in vitro translation in RRLs using [^{35}S]-labeled Met-$tRNA_i^{Met}$. (*Lower*) Protein blot of the same in vitro translation products probed with the 1C2 antibody.

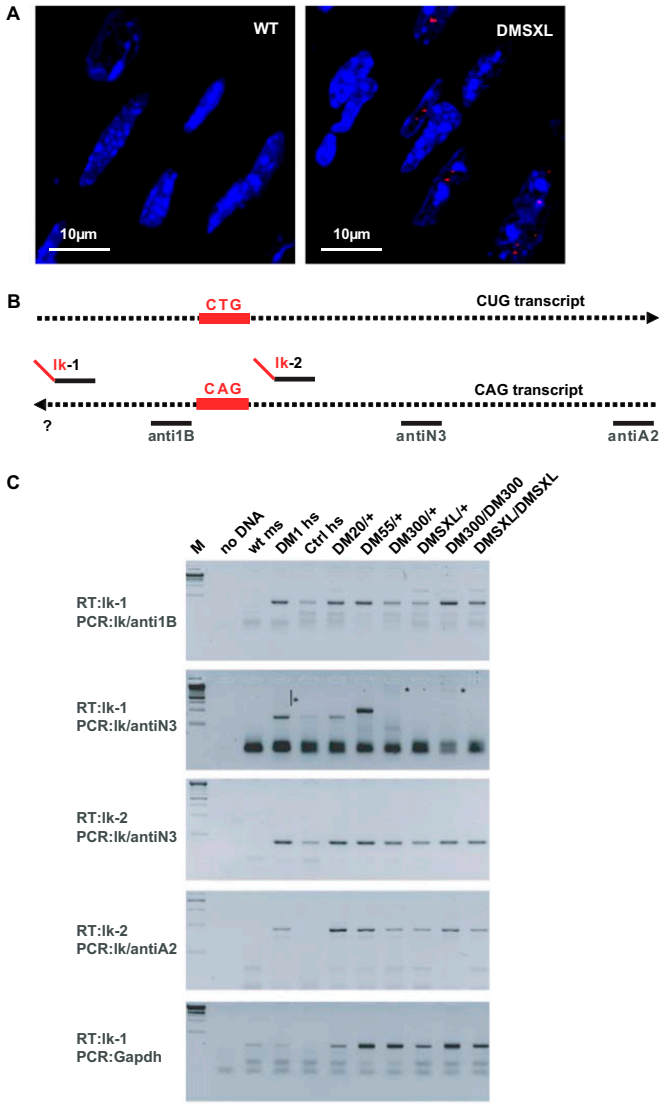


Fig. S8. CUG-containing ribonuclear inclusions and CAG DMPK antisense transcripts in DM mice. (*A*) FISH of CAG probe to detect CUG-containing RNA foci in cardiac sections from DMSXL and WT control animals. (*B*) Diagram showing the DMPK 3′ UTR and the location of antisense-specific primers for the CAG transcript. For strand-specific priming, a linker sequence (lk1 or lk-2) was attached to the DM1-specific primers for cDNA synthesis. PCR was performed using a primer complementary to the linker sequence and reverse primers anti1B, antiN3, or antiA2. The 3′ end of the DM1 CAG RNA is unknown. (*C*) Strand-specific RT-PCR of the human DMPK antisense strand in transgenic mice. Strand-specific reverse transcription and PCR were performed with RNA from a pool of 5-mo-old mouse hearts ($n = 3$) and with RNA from DM1 and control human heart samples. Various lines of transgenic mice have been assessed: DM20 mice with 20 CTGs, DM55 mice with 55 CTGs, DM300 mice with ~600 CTGs, and DMSXL mice with >1,000 CTGs. Ctrl hs heart, human control heart; DM1 hs heart, DM1 human heart; M, 250-bp DNA ladder; wt ms, wild-type mouse. Asterisks to the right of corresponding lanes indicate PCR products with large repeats that amplified with low efficiency. Primers used for DNA synthesis and for PCR are indicated on the left. Gapdh indicates PCR with primers for the mouse Gapdh cDNA that self-primed during reverse transcription. Note that these primers also amplified endogenous human GAPDH cDNA, at lower efficiency.

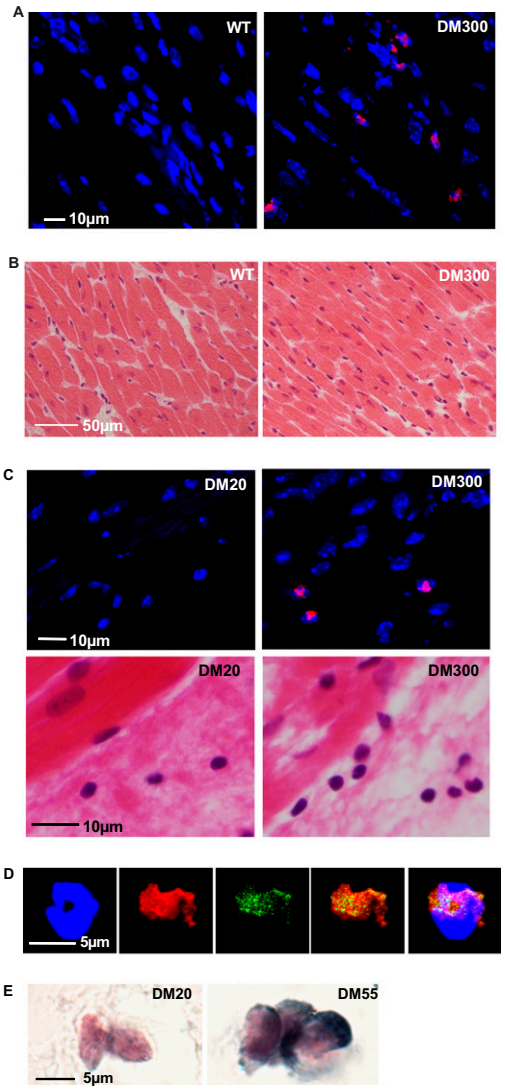


Fig. S9. RAN-translated polyGln protein in DM1 mouse cardiomyocytes and leukocytes and DM1 human leukocytes. (*A*) Immunofluorescence staining of cardiomyocytes using α-DM1$_{CAG\text{-}Gln}$ (cy3, red) in DM 300 (~600 CTG repeats) and WT control mice show rare (2.32 ± 2.04%) but reproducible staining in expansion mice but not control animals based on quantitative digital imaging of six representative fields. (*B*) H&E staining of cardiac tissue comparable to that used in *A* shows typical cardiac histology including large, boxy, centrally located myocyte nuclei in both DM300 and WT samples. (*C*) (*Upper*) Leukocytes in coagulated blood within heart chambers show frequent positive staining (10.69 ± 3.67%) with α-DM1$_{CAG\text{-}Gln}$ for DM300 mice but not DM20 control animals based on four representative fields. (*Lower*) Comparable (nonserial) H&E sections. (*D*) Colocalization of α-DM1$_{CAG\text{-}Gln}$ (cy3, red) with caspase-8 (Alexa Fluor 488, green) in mouse leukocytes colocalized with caspase-8, an early indication of polyGln-induced apoptosis (1). (*E*) Conditions for the α-DM1$_{CAG\text{-}Gln}$ antibody were optimized using frozen tissue. The 1C2 antibody does not adequately detect polyGln inclusions in frozen samples using available methods (2). Therefore we used the 1C2 antibody on fixed tissue to demonstrate independently that DM1 expansion animals express the DM1$_{CAG\text{-}Gln}$ protein. The 1C2 antibody previously has been shown to recognize expanded polyGln repeat tracts (>40) specifically (3). HRP-labeled 1C2-positive cytoplasmic stain (blue) is seen in leukocytes of the DM55 mouse but not of the DM20 control mouse.

1. U M, et al. (2001) Extended polyglutamine selectively interacts with caspase-8 and -10 in nuclear aggregates. *Cell Death Differ* 8:377–386.
2. Iuchi S, Hoffner G, Verbeke P, Djian P, Green H (2003) Oligomeric and polymeric aggregates formed by proteins containing expanded polyglutamine. *Proc Natl Acad Sci USA* 100: 2409–2414.
3. Trottier Y, et al. (1995) Polyglutamine expansion as a pathological epitope in Huntington's disease and four dominant cerebellar ataxias. *Nature* 378:403–406.

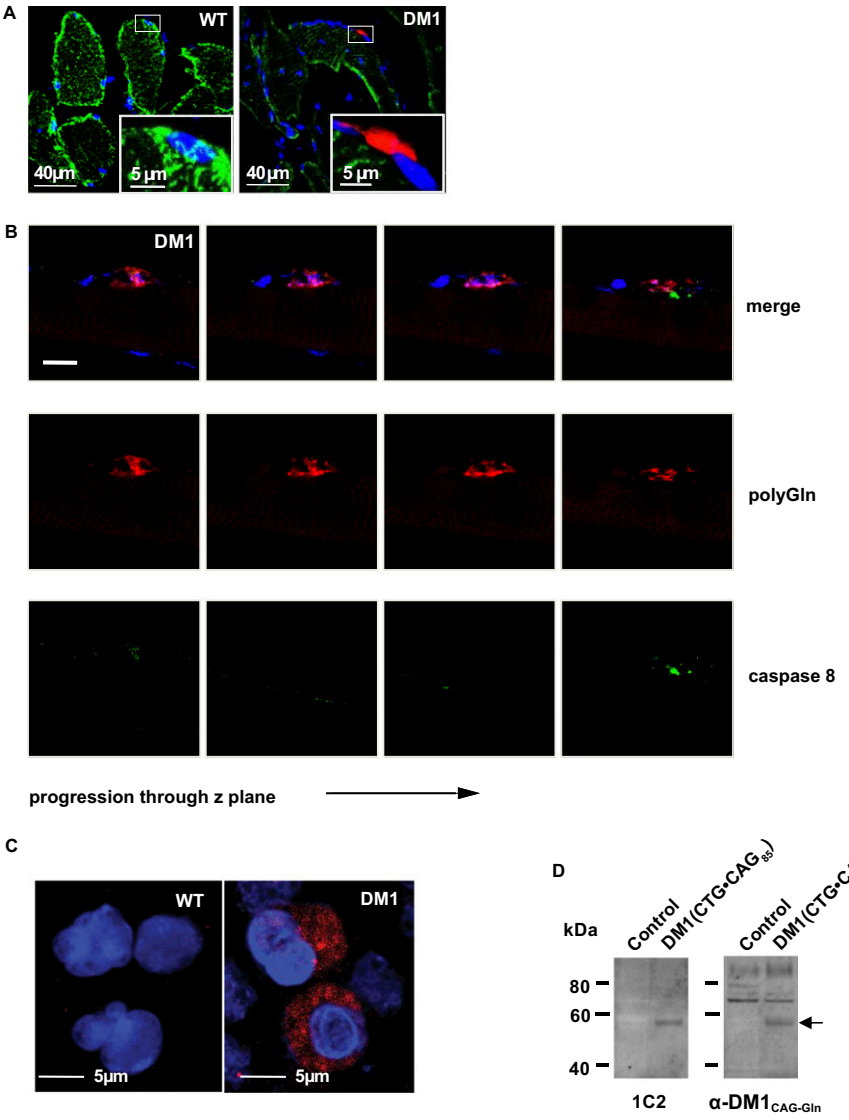


Fig. S10. DM1 polyGln protein in human tissue. (*A*) Immunofluorescence staining of skeletal muscle with α-$DM1_{CAG\text{-}Gln}$ (cy3, red) and α-desmin (green) antibodies. Infrequent but reproducible α-$DM1_{CAG\text{-}Gln}$ staining (cy3, red) was found in skeletal muscle from one DM1 autopsy case. Boxed regions are enlarged as insets. (*B*) DM1 human longitudinal skeletal muscle section showing coexpression of polyGln (red) and the early polyGln apoptotic marker caspase-8 (green). (*C*) Immunofluorescence α-$DM1_{CAG\text{-}Gln}$ (cy3, red) staining of human DM1 leukocytes from peripheral blood. (*D*) Protein blots of human peripheral blood from a patient with 85 CTG•CAGs probed with 1C2 and α-$DM1_{CAG\text{-}Gln}$ antibodies provide independent evidence that the $DM1_{CAG\text{-}Gln}$ expansion protein is expressed in blood.