



Modulation of prion protein expression through cryptic splice site manipulation

Received for publication, January 4, 2024, and in revised form, June 15, 2024. Published, Papers in Press, July 11, 2024.
<https://doi.org/10.1016/j.jbc.2024.107560>

Juliana E. Gentile^{1,2}, Taylor L. Corridon^{1,2}, Meredith A. Mortberg^{1,2}, Elston Neil D'Souza³, Nicola Whiffin^{3,4}, Eric Vallabh Minikel^{1,2}, and Sonia M. Vallabh^{1,2,*}

From the ¹McCance Center for Brain Health and Department of Neurology, Massachusetts General Hospital, Boston, Massachusetts, USA; ²Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA; ³Big Data Institute and Centre for Human Genetics, University of Oxford, Oxford, UK; ⁴Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA

Reviewed by members of the JBC Editorial Board. Edited by Elizabeth J. Coulson

Lowering expression of prion protein (PrP) is a well-validated therapeutic strategy in prion disease, but additional modalities are urgently needed. In other diseases, small molecules have proven capable of modulating pre-mRNA splicing, sometimes by forcing inclusion of cryptic exons that reduce gene expression. Here, we characterize a cryptic exon located in human *PRNP*'s sole intron and evaluate its potential to reduce PrP expression through incorporation into the 5' untranslated region. This exon is homologous to exon 2 in nonprimate species but contains a start codon that would yield an upstream open reading frame with a stop codon prior to a splice site if included in *PRNP* mRNA, potentially downregulating PrP expression through translational repression or nonsense-mediated decay. We establish a minigene transfection system and test a panel of splice site alterations, identifying mutants that reduce PrP expression by as much as 78%. Our findings nominate a new therapeutic target for lowering PrP.

Prion disease is a rapidly fatal neurodegenerative disease caused by the templated misfolding of the prion protein, PrP, encoded by the prion protein gene (*PRNP* in humans) (1). Prion disease naturally afflicts a range of mammals and has long been modeled in laboratory rodents, in which the full disease process can be induced. Both genetic (2) and pharmacological (3, 4) experiments in such models have demonstrated that reducing the amount of PrP in the brain is protective against prion disease, inspiring hope that a PrP-lowering therapy could be used to effectively treat, delay, and prevent disease in patients and individuals at risk (5). An RNase H1 antisense oligonucleotide (ASO) targeting *PRNP* RNA for degradation is now in preclinical development (3, 4, 6, 7), but additional therapeutic candidates are urgently needed.

Recently, the FDA-approved drug risdiplam (8–11) and clinical candidates kinetin and branaplam (12–16) have highlighted small molecule modulation of pre-mRNA splicing as another tool for therapeutic tuning of gene expression.

Branaplam causes incorporation of a piece of intronic sequence—variously called a nonannotated exon, cryptic exon, or poison exon—into mature *HTT* mRNA, causing a frameshift and nonsense-mediated decay (NMD) (15). Inspired by this work, we were led to inquire whether the architecture of *PRNP* would lend itself to disruption *via* splice site manipulation. *PRNP*'s coding sequence is located entirely within a single exon, precluding frameshift strategies. We hypothesized, however, that inclusion of a novel upstream open reading frame (uORF) in the *PRNP* 5'UTR could decrease PrP expression. It is known that uORFs can have dramatic effects on gene expression (17, 18) either through reduced abundance of ribosomes on the canonical ORF or possibly through NMD triggered by the presence of a stop codon prior to the final splice junction, though the latter mechanism is debated (19). The existence of Mendelian diseases caused by variants introducing uORFs (20), the evolutionary constraint of genetic variants that cause or extend uORFs in dosage-sensitive genes (21), as well as work with uORF-targeting ASOs (22) underscore the potential functional impact of uORFs.

Here, we identified a potential uORF within a cryptic exon located in *PRNP*'s sole intron, homologous to exon 2 in many nonprimate species. By genetically strengthening the splice sites surrounding the cryptic exon located in *PRNP*'s 5' UTR, we show that the mutations yielding the most robust inclusion of exon 2 reduced PrP expression by up to 78% in human cells. Certain other mutants reduced *PRNP* transcript levels and PrP protein expression without yielding cryptic exon inclusion detectable by qPCR, suggesting multiple mechanisms may be operative. These efforts nominate a novel strategy for lowering PrP.

Results

PRNP is a small gene of roughly 15 kilobases (kb) in humans (Fig. 1A). In all mammals, the entire coding sequence is contained in the final exon of the gene, while the 5'UTR is divided across exons; however, the number of exons differs. In mouse and most other preclinical species of interest, there are three constitutive exons (23, 24), with introns 1 and 2 dividing the 5'UTR (Fig. 1B). In Syrian hamsters, exon 2 is subject to

* For correspondence: Sonia M. Vallabh, svallabh@broadinstitute.org.

A cryptic exon to lower PrP

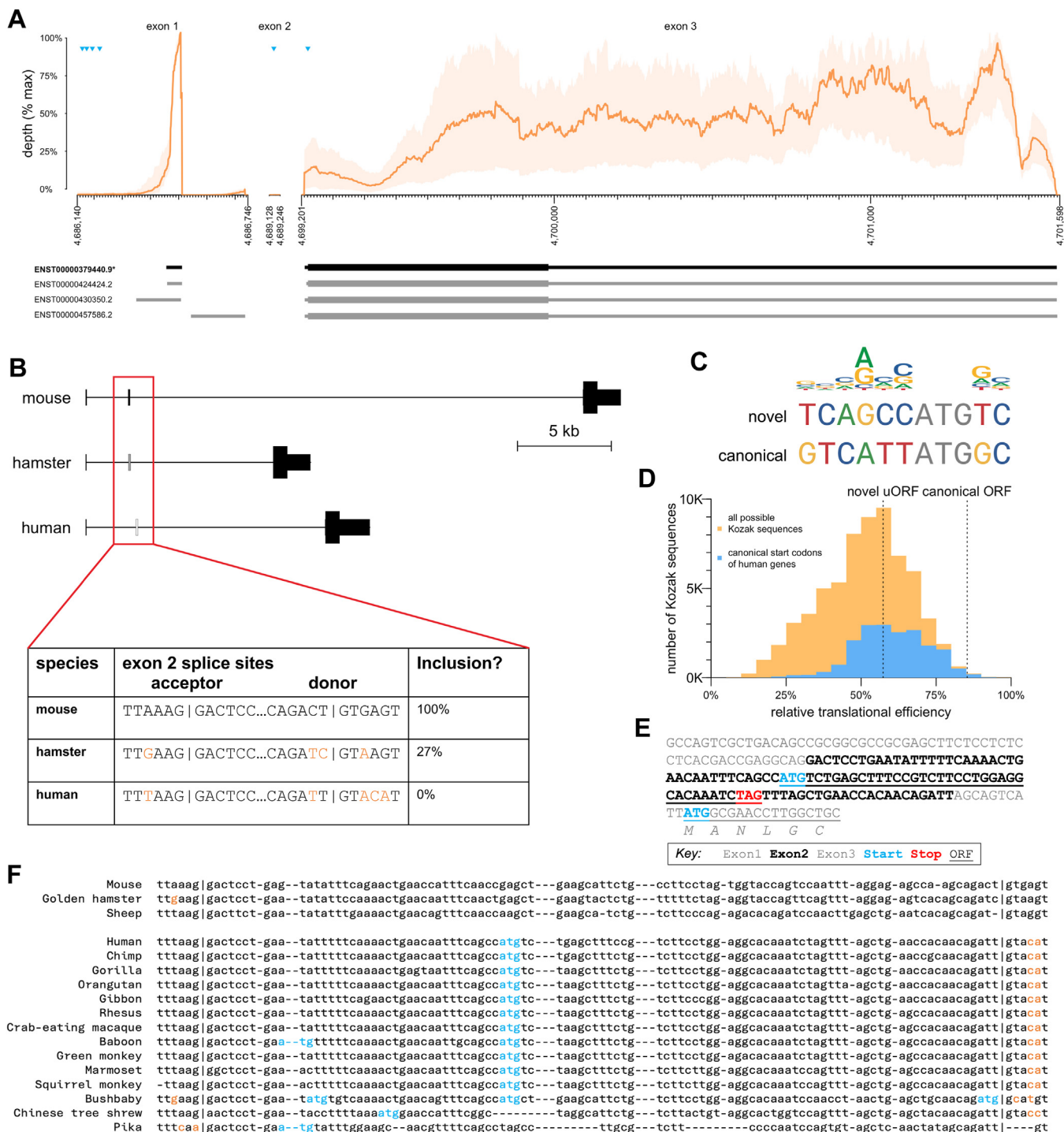


Figure 1. A cryptic exon in human PRNP. *A*, human PRNP transcript structure in human brain. Top panels show GTEX (30) v8 bulk RNA-seq coverage—mean (orange lines) and range (orange shaded area) across 13 brain regions. Coverage depth for exons 1 and 2 is normalized to the max for exon 1; depth for exon 3 is normalized to the max for exon 3. ATGs representing candidate upstream open reading frames and the canonical open reading frame are shown as blue triangles. Ensembl GRCh38.p14 annotated transcripts are shown below, canonical in black, alternatives in gray. *B*, comparison of orthologous exon 2 sequence in mouse, hamster, and human. Hamster inclusion percentage from ref (25). *C*, comparison of PRNP canonical and exon 2 novel ATG Kozak contexts with a sequence logo of human initiation sites (see Experimental procedures). *D*, relative strength of canonical and novel PRNP ORFs in context. Shown for comparison are histograms of translational efficiency of all 65,536 (4^3) possible Kozak contexts (yellow) and of all 18,784 actual human canonical ORF Kozak contexts (blue), expressed as a percentage of the translation of the most efficient Kozak context, TTCATCATGCA, according to data from Noderer *et al.* (28). *E*, annotated sequence of the PRNP 5'UTR if exon 2 were included. Frame is relative to the canonical ORF, and percentile indicates strength of the Kozak context as a percentile of all possible Kozak sequences, using rankings from ref (28). *F*, multiple alignment of PRNP exon 2 sequences known to be constitutively or variably included in mRNA from mouse (23), hamster (25), and sheep (24) versus all orthologous sequences in eutherian mammals that contain ATGs. ATGs are shown in blue and splice site variants absent from mouse, hamster, or sheep are shown in orange. A full alignment including all eutherian mammals is shown in Fig. S1. uORF, upstream open reading frame.

variable splicing and is included in ~27% of transcripts (25) (Fig. 1B). In humans and several closely related primate species, *PRNP* has only two annotated exons, the equivalent of exons 1 and 3 from other mammals; exon 2 remains as a cryptic exon within the sole intron (26). For clarity, herein we will refer to human *PRNP* exons 1, 2, and 3 and introns 1 and 2, even though the naturally occurring *PRNP* transcript contains only two exons and one intron.

Although essential splice sites—AG at the A-1 and A-2 and GT at the D + 1 and D + 2 positions—are conserved in human exon 2, we hypothesized that other nearby base pair substitutions may contribute to exclusion of this exon, particularly the loss of the G at the highly constrained D + 5 position (27) (Fig. 1B). Human *PRNP* exon 2 contains an ATG in a moderately strong Kozak context (Fig. 1C), estimated to yield 57% maximal translational efficiency, near the median of canonical ORFs of all other human protein-coding genes (28) (Fig. 1D). Human *PRNP* was previously reported (29) to already contain four uORFs in exon 1; however, RNA-seq data from human brain tissue (30) provide no support for transcription initiation beginning this far upstream: mean RNA-seq coverage at these uORFs is <0.5% of the peak coverage within exon 1 (Fig. 1A). Thus, if exon 2 were included, its ATG would yield a new, sole uORF upstream of *PRNP*'s canonical start codon (Fig. 1E) with the potential to downregulate PrP expression through its impact on ribosomal activity (17) (N. Whiffin, unpublished results). Its stop codon also occurs 22 bp prior to the exon two-thirds splice junction (Fig. 1E), creating a possible opportunity to trigger NMD (see Discussion). Alignment of *PRNP* exon 2 sequences across all available mammalian species (Figs. 1F and S1) reveals that exon 2 ATGs are present only in species with exon 2 splice site variants known or predicted to exclude exon 2 from mature mRNA, consistent with the possibility of exon 2 uORF having a strong negative effect on PrP expression. We thereby hypothesized that acting through either of these mechanisms, inclusion of exon 2 and thus the uORF of interest in *PRNP* mRNA, would reduce PrP expression.

To test this hypothesis, we first sought to generate a *PRNP* minigene system to support facile splice site manipulation, transfection, and screening in cell culture. A 4.8 kb minigene lacking most of intron 1 yielded no detectable PrP expression in HEK293 cells by Western blot (Fig. S2). A 6.5 kb construct retaining all of intron 1 and only the first and last 500 bp of intron 2 (Fig. 2A) expressed robustly and was used for all subsequent experiments. Codon optimization of exon 3 allowed for qPCR primer/probe pairs to discriminate minigene *PRNP* RNA from endogenous *PRNP* RNA (Fig. 2B).

Using this 6.5 kb minigene as a template, we designed a panel of splice site modifications that we hypothesized would strengthen exon 2 inclusion in the context of human *PRNP* (Fig. 3A). These included (1) installation of the consensus strongest (31) human splice donor and acceptor (“canonical splice site (ss”); six nucleotide changes required); (2) installation of the mouse *Prnp* exon 2 splice sites

(“mouse ss”; five changes required); (3) conversion of the donor +5 site from A to G, as this site shows the strongest nucleotide preference of any extended splice site position (27) (D + 5, A > G); (4) conversion of the acceptor -3 site from A to C, to assess whether this single change could mimic the effect of installing the consensus human splice site (A-3, A > C); and (5) conversion of the acceptor -4 site from T to A to assess whether this single change could mimic the effect of installing the mouse *Prnp* splice site (A-4, T > A).

Each mutant was separately transfected into HEK293 cells alongside the parent minigene construct and empty vector and GFP transfection controls and analyzed by qPCR. Each primer/probe set (Fig. 2B) was designed to amplify only if the targeted exons are adjacent. In keeping with these expectations, the empty vector and GFP controls yielded negligible signal for all primer pairs; trace amplification of exon 1 to 3 may reflect imperfect allele specificity, as only two bases differ from endogenous *PRNP* in the exon 3 codon-optimized primer. The parent minigene yielded PCR product for exon 1 to 3 but not for exon 1 to 2 or 2 to 3, reflecting the baseline exclusion of cryptic exon 2 in a human system.

All five splice site mutants appeared to reduce the amount of normally spliced *PRNP* RNA, as measured by the exon 1 to 3 primer pair, with the change being significant for four mutants (Fig. 3B). The two mutants yielding the greatest reduction—the canonical ss and mouse ss mutants—showed a corresponding increase in the presence exon 1 to 2 and 2 to 3 junctions (Fig. 3, C and D). For all other constructs, exon 2 remained undetectable, or nearly so, by these primer/probe sets. Note that the results for exons 1 to 2 and 2 to 3 are normalized to the highest value obtained for any mutant; 100% does not necessarily mean 100% exon 2 inclusion.

Immunoblots on cell lysates revealed apparent reductions in PrP for all mutants tested (Fig. 3, E and F). The canonical ss mutant yielded 33% and the mouse ss mutant 48% of the PrP expression level of the parent minigene (Fig. 3F). HEK293 cells express endogenous PrP, however, at ~15%, the level achieved by transfection of the parent minigene (Fig. 3, E and F); adjusting for this floor yielded residual PrP expression of 22% and 38% for the canonical ss and mouse ss mutants, respectively. Across all mutants, PrP levels tracked closely with exon 1 to 3 qPCR results, with significant reductions for mutants D + 5, A > G and A-3, A > C despite the lack of detectable exon 1 to 2 and 2 to 3 junctions (Fig. 3, C, D and F).

In order to dissect the mechanistic role of the exon 2 uORF in reducing PrP expression, we generated variants of the canonical ss and mouse ss mutants with the ATG mutated to CCC (Fig. S3) and transfected them into HEK293 cells (Fig. 4, A–D). On the absence of the ATG, exon 1 to 3 splicing was significantly increased for the mouse ss (Fig. 4A), a partial restoration toward the levels seen for the wildtype minigene. Exon 2 to 3 splicing was slightly, but not significantly, reduced for both canonical and mouse ss (Fig. 4B). Exon 1 to 2 splicing was not compared because the probe used throughout this study (Fig. 2B)

A cryptic exon to lower PrP

A Minigene construction



B qPCR Primer/probe binding sites

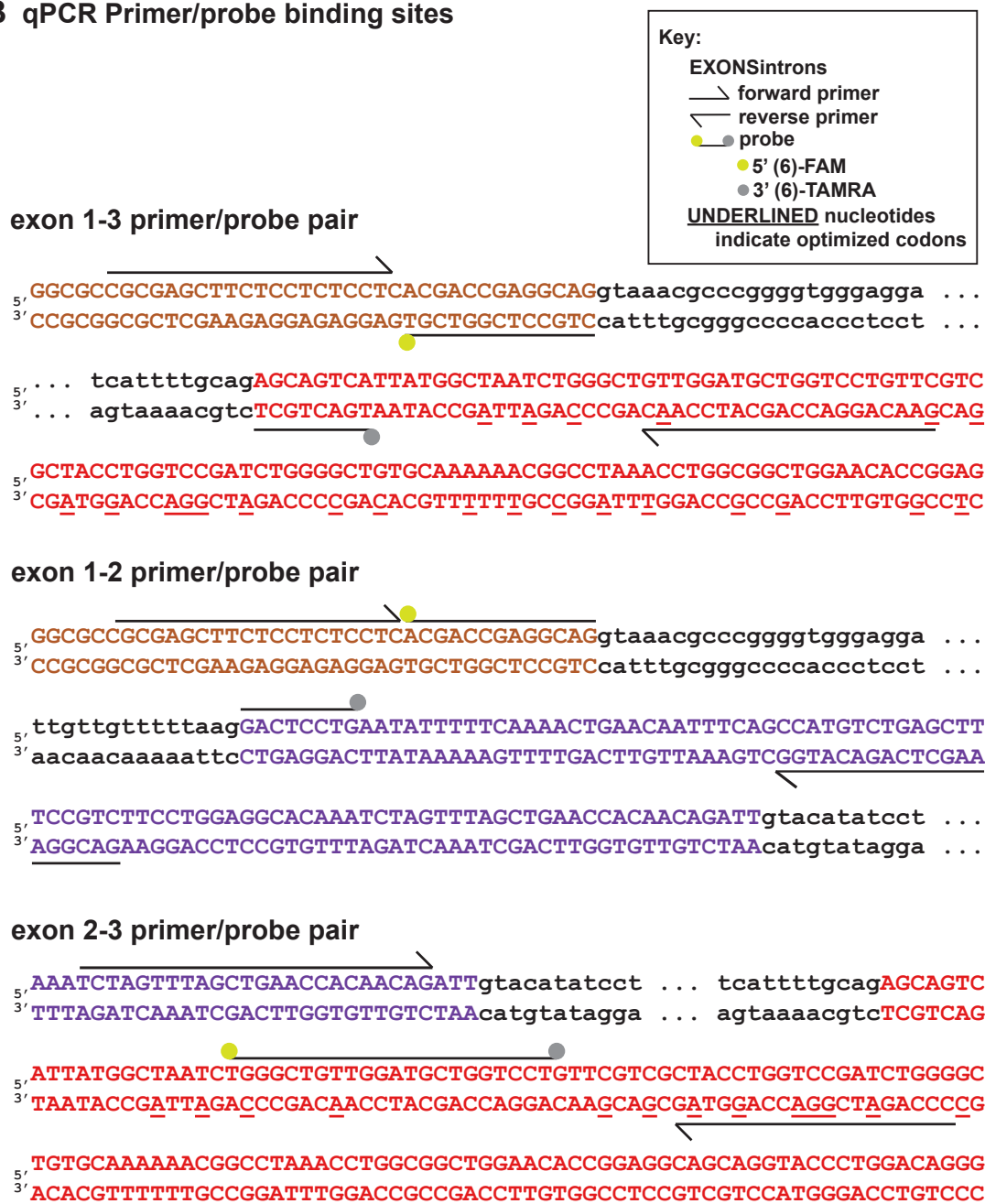


Figure 2. Design of PRNP minigene and primer pairs. A, diagram of minigene versus human reference sequence. Intronic sequence upstream of exon 2 and 500 bp on either end of the intronic sequence downstream of exon 2 are included. B, design of primer/probe pairs used to interrogate splicing of the minigene. Note that codon optimization in exon 3 (underlines) enables these pairs to discriminate the minigene from endogenous PRNP in HEK293 cells.

overlaps the ATG. For both the canonical and mouse ss, the ATG to CCC mutation partially but not entirely restored PrP protein levels (Fig. 4, C and D).

We also sought to confirm our findings in a neuronal cell line, using mouse N2a cells. The same pattern of splicing was confirmed for each mutant as observed in HEK293 cells

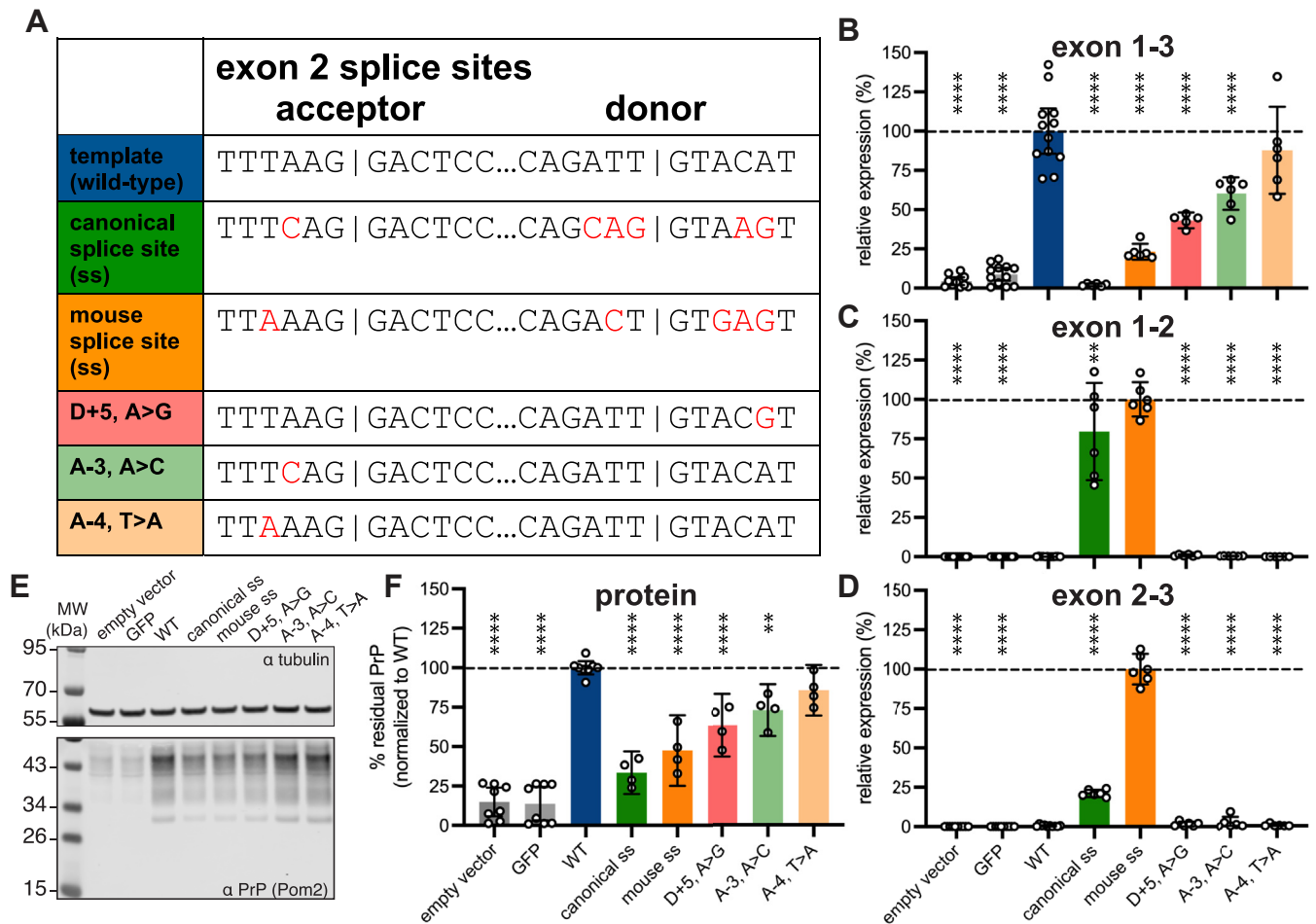


Figure 3. Inclusion of exon 2 lowers PrP expression. A, sequence variants of minigene tested in HEK293 cells. B–D, expression of (B) exon 1 to 3 (n = 5–12 biological replicates, meaning transfected wells/variant), (C) exon 1 to 2 (n = 6–12 transfected wells/variant) junctions in minigene mRNA for each variant transfected into HEK293 cells. Normalized to the template minigene for exons 1–3 and normalized to the highest-expressing variant for exons 1–2 and 2–3. Note that codon optimization in exon 3 enables discrimination from endogenous PRNP. E, immunoblot (POM2 primary antibody (49)) of PrP expression in HEK293 cells transfected with each variant. The molecular weight (MW) ladder marker at the top where the blot is cut off is 55 kDa. F, quantification of PrP expression from ≥ 4 immunoblots per construct, n = 4–8 transfected wells/variant. In all barplots, bar heights indicate means, and error bars indicate 95% confidence intervals. In B, C, D, and F, stars indicate statistical significance according to an ordinary one-way ANOVA and Dunnett's multiple comparison test with a single pooled variance. Overall ANOVA $p < 0.0001$. Individual tests $**p < 0.01$ and $***p < 0.0001$. Exact p values are provided in Tables S2 and S5. PrP, prion protein.

(Fig. 4, E–G). To distinguish minigene-expressed human PrP from endogenous mouse PrP, we used the human-specific 3F4 antibody. Although the magnitude of effect was more modest than observed in HEK293 cells, PrP expression was reduced for each mutant tested, with the canonical ss again providing the deepest reduction (Fig. 4, H and I).

Discussion

We find that splice site manipulation can modulate the level of PrP in human and mouse cell systems, reducing the levels of this disease-causing protein by 78% in the strongest condition tested. For the strongest mutants, which incorporated 5 to 6 nucleotide changes across the splice donor and acceptor sites, this reduction in protein level was observed in tandem with exon 2 inclusion at the mRNA level. Our data are consistent with a role for uORF-mediated translational repression; however, we cannot rule out that NMD may be at work, with the exon 1 to 2 and 2 to 3 qPCR simply picking up the small

fraction of exon 2-including mRNA that has not yet been degraded. NMD was long held to require 50 bp of distance between the stop codon and the splice donor (32) versus only 22 bp here, but data from protein-truncating variants in human tissues show this is not a hard-and-fast rule and that distance from the splice donor is but one of many imperfect predictors of NMD (33). Still, the evidence for NMD caused by uORFs in human genes is equivocal (19, 34). Abolition of the uORF by mutation of its initiation codon ATG to CCC increased exon 1 to 3 splicing for the mouse ss variant, which would be consistent with a role for NMD, but no increase in exon 1 to 3 splicing was detectable upon uORF abolition for the canonical ss. Further, uORF abolition only partially, but not completely, restored protein expression level, indicating that the uORF is not fully responsible for the reduction in expression. Moreover, the single point mutants tested here reduced PrP and normal exon 1 to 3 splicing without yielding any detectable exon 1 to 2 and 2 to 3 splicing. Thus, additional mechanisms not foreseen by our initial hypothesis could be

A cryptic exon to lower PrP

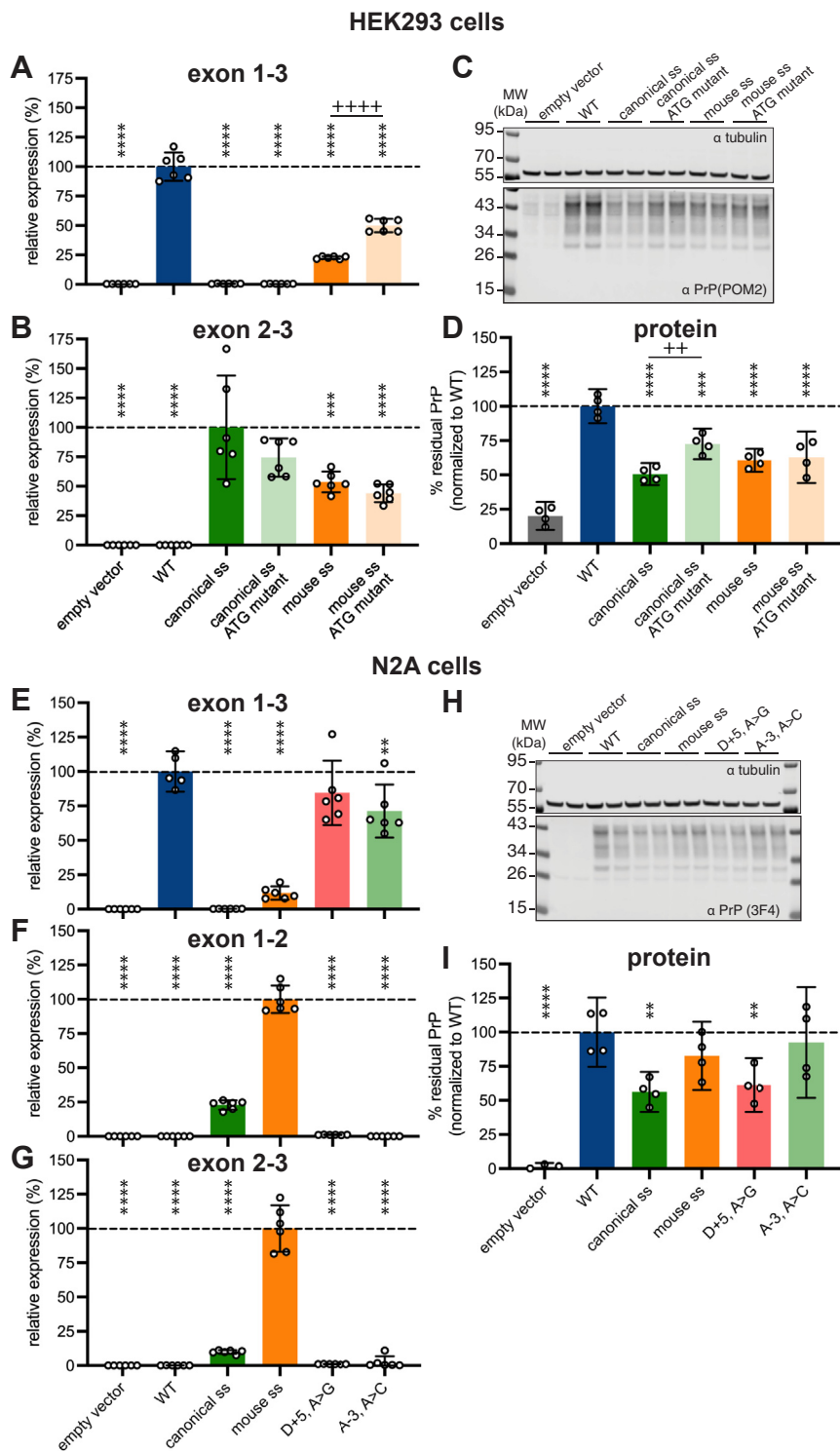


Figure 4. Contribution of uORF and replication in a neuronal cell line. A–D, HEK293 cells. As in Fig. 3, B and D, E, and F but with the ATG to CCC mutants included. Mutant sequences are provided in Fig. S3. (A) Exon 1 to 3 (n = 6 transfected wells/variant), (B) exon 2 to 3 (n = 6 transfected wells/variant), (C) representative POM2 immunoblot, (D) quantification of PrP expression from two immunoblots per construct, n = 4 transfected wells/variant. E–I, N2a cells. As in Fig. 3, B–F. E, exon 1 to 3 (n = 6 transfected wells/variant), (F) exon 1 to 2 (n = 6 transfected wells/variant), (G) exon 2 to 3 (n = 6 transfected wells/variant), (H) representative 3F4 immunoblot, (I) quantification of PrP expression from two immunoblots per construct, n = 4 transfected wells/variant. In all barplots, bar heights indicate means, and error bars indicate 95% confidence intervals. In A, B, D, E, F, G, and I, stars indicate statistical significance according to an ordinary one-way ANOVA and Dunnett’s multiple comparison test with a single pooled variance. Overall ANOVA $p < 0.0001$. Individual tests **, $p < 0.01$, ***, $p < 0.001$ and ****, $p < 0.0001$ relative to wildtype construct. ++ $p < 0.01$, +++ $p < 0.0001$ relative to corresponding variant without ATG to CCC mutation. Exact p values are provided in Tables S2 and S5. PrP, prion protein; uORF, upstream open reading frame.

operative: for example, these mutants might cause inclusion of exon 2 but also retention of a portion of intronic sequence, causing the exon 1 to 2 and 2 to 3 qPCR to not amplify, while resulting in a less stable transcript.

Our study has several limitations. The battery of splice manipulations that we tested was limited, leaving open the possibility that other splice site changes could yield more dramatic results. As our experiments were limited to human cell culture, *in vivo* relevance was not demonstrated. We also utilized a transfected minigene rather than editing of the endogenous *PRNP* gene in cells. Most importantly, the genetic engineering used to establish this proof of concept does not offer a direct path to therapeutic application.

In principle, several therapeutic modalities could be deployed to modulate *PRNP* splicing (35, 36). ASOs are a well-established modality capable of causing exon inclusion (37) but may be unlikely to be deployed toward this end: given the desired mechanism of reducing *PRNP* expression, RNase H1 ASOs are likely to yield greater target suppression than splice-modulating ASOs. Adenine base editors have been successfully deployed to disrupt splice sites (38, 39); however, the single point mutants identified here had relatively modest effects on PrP expression. Instead, small molecule modulation of *PRNP* splicing is the most enticing possibility suggested by our results. PrP-lowering small molecules could have desirable pharmacologic properties, particularly in terms of distribution to deep brain structures less well-reached by oligonucleotides (40). Attempts to discover small molecules to bind PrP have been unsuccessful (41), so splicing could offer a new mechanism for small molecule therapies in prion disease. Because *PRNP* does not share the preferred splice site motifs of any known splice-modulating small molecule series (10, 12, 14), discovery of a modulator would require a new screening effort. One challenge is that the mutant that reduced PrP expression most dramatically required six nucleotide changes, suggesting a high energy barrier to modulating this splicing event.

Despite these limitations, we are encouraged to discover a novel mechanism by which PrP expression can be influenced. PrP's role in prion disease is uniquely pivotal, as it serves as protein-only pathogen, amplification substrate, and mediator of neuronal neurotoxicity. The therapeutic benefit of PrP lowering has been shown across multiple prion strains (4), both through genetic reduction and by use of ASOs, and evidence for tolerability is provided by multiple nonhuman species as well as human genetics (42–47). Given this clarity, PrP and its precursors are disease targets worthy of ongoing creative angles of attack.

Experimental procedures

Kozak sequences

Files were retrieved from the Matched Annotation from NCBI, and EMBL-EBI (MANE) database (version 1.0) https://ftp.ncbi.nlm.nih.gov/refseq/MANE/MANE_human/release_1.0/cDNA Transcript sequences from the MANE. GRCh38.v1.0.refseq_rna.fna.gz file were then filtered to only MANE Select transcripts, and an 11-bp context surrounding

the CDS start were extracted, excluding transcripts where a 11 bp CDS context could not be retrieved, as in the case for a leaderless mRNA. To generate a sequence logo, these 11-bp sequences were then superimposed to align with each other and plotted using ggplot2 and the R package ggseqlogo using the bits method. To generate a histogram, the relative translational efficiencies of each sequence were taken from Noderer *et al.* (28) and normalized to the most efficient Kozak sequence.

Comparative genomic analyses

PRNP sequences, and multiple alignments thereof, were obtained from UCSC Genome Browser (48) (accessed September 6, 2023). Kozak sequence strength percentiles were obtained from the rank order among all possible Kozak sequences reported by Noderer *et al.* (28). GTE_x (30) RNA-seq coverage data were obtained from UCSC Table Browser (accessed November 14, 2023). Exon 1 and 2 in diagrams correspond to the canonical Ensembl transcript ENST00000379440.9.

Cell culture and transfections

HEK293 (ATCC cat no. CRL-1573) and N2a (ATCC cat no. CCL-131) cells were maintained in Dulbecco's modified Eagle's medium/F-12 (Gibco, cat no. 11320033) supplemented with 1% penicillin-streptomycin (Gibco, cat no. 15140163) and 10% FBS (Gibco, cat no. 16000044). Cells were periodically surveilled for *mycoplasma* (Lonza cat. no. LT07-418 and LT07-518) and were always negative. For transfection, cells were plated in a 12-well or 96-well plate for protein or RNA analysis, respectively, and were allowed to adhere for 18 h. Cells were then transfected using Lipofectamine 3000 transfection reagent (Invitrogen, cat no. L3000015) according to the manufacturer protocol. In short, lipofectamine 3000 reagent was diluted in Opti-MEM I reduced serum media (Gibco, cat no. 31985088) for a final mixture containing 3% lipofectamine 3000. In a separate tube, 1 μg (12-well plate) or 0.1 μg (96-well plate) DNA was mixed with 4% P3000 reagent in Opti-MEM. The two tubes were slowly mixed and then allowed to incubate at room temperature for 10 min before applying the mixture to the cell media. Transfection was incubated on cells for 48 h before lysing cells.

Plasmids

The PGK promoter (addgene 82579) was inserted into the backbone of the pcDNA3.1(+) plasmid (addgene V790-20), removing the CMV promoter, by digesting the vector with NruI and NheI. The minigene was synthesized with codon optimized exon 3 and then was ligated into this pcDNA3.1(+)-hPGK between NheI and EcoRI.

Western blot analysis

Following the 48 h transfection, cells were washed thoroughly with ice-cold PBS and then were lysed in 0.2% CHAPS containing cOmplete, Mini, and EDTA-free Protease Inhibitor Cocktail (Sigma, cat no. 4693159001). Protein concentration was determined using a DC protein assay kit (Bio-Rad, cat no.

A cryptic exon to lower PrP

5000112). NuPAGE 4 to 12%, Bis-Tris, mini protein gels (Invitrogen, cat no. NP0323BOX) were loaded with 10 μg total protein for each sample and run at 180 V in 1x MES buffer (Thermo, cat no. NP0002). Gels were transferred to PVDF membranes using an iBlot two device (iBlot 2 Transfer Stacks, PVDF, mini, Thermo, cat no. IB24002), 20 V, 7 min. Membrane was then cut right under 55 kDa band before blocking with LICOR TBS blocking buffer (LICOR, cat no. 927-60001), 1 h at room temperature. Primary antibodies were diluted in LICOR TBS blocking buffer + 0.2% Tween-20 (Teknova, cat no. T0710) and incubated at 4 °C overnight: α -Tubulin (Invitrogen, cat no. A11126), final 100 ng/ μl ; POM2 (Millipore, cat no. MABN2298), final 50 ng/ μl ; 3F4 (Sigma-Aldrich, cat no. MAB1562), final 0.1 ng/ μl ; and 6D11 (BioLegend, cat no. 808001), final 2 $\mu\text{g}/\mu\text{l}$. The specificity of POM2 has been characterized extensively elsewhere (49). Membranes were washed in 1x TBST and then incubated in secondary antibody (IRDye 800CW Goat anti-Mouse IgG, LICOR, cat no. 926-32210) diluted in LICOR TBS blocking buffer + 0.2% Tween-20 and incubated at room temperature for 1 h. Membranes were again washed with 1x TBST then scanned on a LICOR Odyssey CLx Infrared Imaging System. Blots were analyzed in Fiji (50). The same box was used for every PrP lane, with raw integrated density (the sum of the values of the pixels in the box) as the output value. Light contrast images were used for tubulin quantification and dark contrast images for PrP quantification. Background-subtracted PrP intensity was normalized to background-subtracted tubulin intensity. Raw uncropped gels are available in this study's online git repository.

qPCR

Following the 48 h transfection, cells were lysed using the Cells-to-CT 1-step Taqman Kit (Invitrogen, cat no. A25602) using the manufacturer protocol. In short, media were aspirated, each well was washed with 200 μl of ice cold 1x PBS, and then wash was completely aspirated. Room temperature DNase/lysis solution (0.5 μl : 50 μl) was added to the cells and then plate was put on a shaker for 5 min. Finally, 5 μl of room temperature stop solution was added to the cells and then plate was put back on the shaker for 2 min before moving the plate to ice. RT-PCR samples were prepared using Taqman 1-Step qRT-PCR master mix and Taqman gene expression assays for human *TBP* (Invitrogen, cat no. Hs00427620_m1) or mouse *Tbp* (Invitrogen, cat no. Mm00446971_m1). Custom primers and probes were ordered from Genscript to quantify the different splice variants (see Table S1 for sequences and Fig. 2 for alignment on the minigene sequence). Samples were run on a QuantStudio 7 Flex system (Applied Biosystems) using the following cycling conditions: reverse transcription 50 °C, 5 min; reverse transcription inactivation/initial denaturation 95 °C, 20 s; amplification 95 °C, 3 s, 60 °C, 30 s, 40 cycles. Each biological sample was run in duplicate, and the level of all targets were determined by $\Delta\Delta\text{Ct}$ whereby results were first normalized to the housekeeping gene *TBP* and then to

the wildtype template (exon 1–3) or the mouse ss (exon 1–2 and 2–3), depending on the primer pair used.

Experimental design and statistical analysis

All data were generated from at least three independent transfections, N are as indicated in figure legends. Throughout, all error bars in figures represent 95% confidence intervals. All data were compared with an ordinary one-way ANOVA and Dunnett's multiple comparison test, with a single pooled variance. *p* values less than 0.05 were considered nominally significant. In plots, ***p* < 0.01, ****p* < 0.001, and *****p* < 0.0001. Intron/exon diagrams were plotted in R, qPCR analysis was performed in Google Sheets, and barplots and statistical analyses were performed in GraphPad Prism.

Data availability

Source data for figures are provided in the form of supplementary tables. Raw data, Prism files, and R source code are available at https://github.com/ericminikel/cryptic_exon.

Supporting information—This article contains supporting information.

Author contributions—S. M. V., E. V. M., and J. E. G. conceptualization; S. M. V. and J. E. G. data curation; S. M. V., N. W., T. L. C., M. A. M., E. V. M., E. N. D. S., and J. E. G. formal analysis; S. M. V. and N. W. funding acquisition; S. M. V., M. A. M., E. N. D. S., J. E. G., and T. L. C. investigation; S. M. V. project administration; S. M. V. and N. W. supervision; S. M. V. writing—original draft; S. M. V., E. V. M., and J. E. G. writing—review & editing; N. W. and E. N. D. S. software; N. W., E. V. M., and J. E. G. visualization.

Funding and additional information—This study was supported by the Broad Institute (Chemical Biology and Therapeutic Science program funds) and National Institutes of Health (R01 NS125255). N. W. is supported by a Sir Henry Dale Fellowship jointly funded by Wellcome and the Royal Society (220134/Z/20/Z) and research grant funding from the Rosetrees Trust (PGL19-2/10025). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest—S. M. V. acknowledges speaking fees from Ultragenyx, Illumina, Biogen, and Eli Lilly, consulting fees from Invitae and Alnylam, and research support from Ionis, Gate, and Sangamo. E.V.M. acknowledges speaking fees from Eli Lilly, consulting fees from Deerfield and Alnylam, and research support from Ionis, Gate, and Sangamo.

Abbreviations—The abbreviations used are: ASO, antisense oligonucleotide; NMD, nonsense-mediated decay; PrP, prion protein; ss, splice site; uORF, upstream open reading frame.

References

1. Prusiner, S. B. (1998) Prions. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 13363–13383
2. Büeler, H., Raeber, A., Sailer, A., Fischer, M., Aguzzi, A., and Weissmann, C. (1994) High prion and PrPsc levels but delayed onset of disease in

- scrapie-inoculated mice heterozygous for a disrupted PrP gene. *Mol. Med.* **1**, 19–30
3. Raymond, G. J., Zhao, H. T., Race, B., Raymond, L. D., Williams, K., Swazey, E. E., *et al.* (2019) Antisense oligonucleotides extend survival of prion-infected mice. *JCI Insight* **5**, e131175
 4. Minikel, E. V., Zhao, H. T., Le, J., O'Moore, J., Pitstick, R., Graffam, S., *et al.* (2020) Prion protein lowering is a disease-modifying therapy across prion disease stages, strains and endpoints. *Nucleic Acids Res.* **48**, 10615–10631
 5. Vallabh, S. M., Minikel, E. V., Schreiber, S. L., and Lander, E. S. (2020) Towards a treatment for genetic prion disease: trials and biomarkers. *Lancet Neurol.* **19**, 361–368
 6. Mortberg, M. A., Zhao, H. T., Reidenbach, A. G., Gentile, J. E., Kuhn, E., O'Moore, J., *et al.* (2022) Regional variability and genotypic and pharmacodynamic effects on PrP concentration in the CNS. *JCI Insight* **7**, e156532
 7. Mortberg, M. A., Gentile, J. E., Nadaf, N. M., Vanderburg, C., Simmons, S., Dubinsky, D., *et al.* (2023) A single-cell map of antisense oligonucleotide activity in the brain. *Nucleic Acids Res.* **51**, 7109–7124
 8. Naryshkin, N. A., Weetall, M., Dakka, A., Narasimhan, J., Zhao, X., Feng, Z., *et al.* (2014) Motor neuron disease. SMN2 splicing modifiers improve motor function and longevity in mice with spinal muscular atrophy. *Science* **345**, 688–693
 9. Ratni, H., Ebeling, M., Baird, J., Bendels, S., Bylund, J., Chen, K. S., *et al.* (2018) Discovery of risdiplam, a Selective survival of motor neuron-2 (SMN2) gene splicing modifier for the treatment of spinal muscular atrophy (SMA). *J. Med. Chem.* **61**, 6501–6517
 10. Campagne, S., Boigner, S., Rüdiger, S., Moursy, A., Gillioz, L., Knörlein, A., *et al.* (2019) Structural basis of a small molecule targeting RNA for a specific splicing correction. *Nat. Chem. Biol.* **15**, 1191–1198
 11. Baranello, G., Darras, B. T., Day, J. W., Deconinck, N., Klein, A., Masson, R., *et al.* (2021) Risdiplam in type 1 spinal muscular atrophy. *N. Engl. J. Med.* **384**, 915–923
 12. Hims, M. M., Ibrahim, E. C., Leyne, M., Mull, J., Liu, L., Lazaro, C., *et al.* (2007) Therapeutic potential and mechanism of kinetin as a treatment for the human splicing disease familial dysautonomia. *J. Mol. Med.* **85**, 149–161
 13. Axelrod, F. B., Liebes, L., Gold-Von Simson, G., Mendoza, S., Mull, J., Leyne, M., *et al.* (2011) Kinetin improves IKBKAP mRNA splicing in patients with familial dysautonomia. *Pediatr. Res.* **70**, 480–483
 14. Palacino, J., Swalley, S. E., Song, C., Cheung, A. K., Shu, L., Zhang, X., *et al.* (2015) SMN2 splice modulators enhance U1-pre-mRNA association and rescue SMA mice. *Nat. Chem. Biol.* **11**, 511–517
 15. Krach, F., Stemick, J., Boerstler, T., Weiss, A., Lingos, I., Reischl, S., *et al.* (2022) An alternative splicing modulator decreases mutant HTT and improves the molecular fingerprint in Huntington's disease patient neurons. *Nat. Commun.* **13**, 6797
 16. [preprint] Krach, F., Boerstler, T., Reischl, S., Krumm, L., Regensburger, M., Winkler, J., *et al.* (2023) RNA splicing modulator induces peripheral neuropathy with increased neurofilament light chain levels via p53 signaling. *bioRxiv*. <https://doi.org/10.1101/2023.08.02.551640>
 17. Calvo, S. E., Pagliarini, D. J., and Mootha, V. K. (2009) Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 7507–7512
 18. Johnstone, T. G., Bazzini, A. A., and Giraldez, A. J. (2016) Upstream ORFs are prevalent translational repressors in vertebrates. *EMBO J.* **35**, 706–723
 19. Lee, D. S. M., Park, J., Kromer, A., Baras, A., Rader, D. J., Ritchie, M. D., *et al.* (2021) Disrupting upstream translation in mRNAs is associated with human disease. *Nat. Commun.* **12**, 1515
 20. Wright, C. F., Quaife, N. M., Ramos-Hernández, L., Danecek, P., Ferla, M. P., Samocha, K. E., *et al.* (2021) Non-coding region variants upstream of MEF2C cause severe developmental disorder through three distinct loss-of-function mechanisms. *Am. J. Hum. Genet.* **108**, 1083–1094
 21. Whiffin, N., Karczewski, K. J., Zhang, X., Chothani, S., Smith, M. J., Evans, D. G., *et al.* (2020) Characterising the loss-of-function impact of 5' untranslated region variants in 15,708 individuals. *Nat. Commun.* **11**, 2523
 22. Liang, X.-H., Shen, W., Sun, H., Migawa, M. T., Vickers, T. A., and Crooke, S. T. (2016) Translation efficiency of mRNAs is increased by antisense oligonucleotides targeting upstream open reading frames. *Nat. Biotechnol.* **34**, 875–880
 23. Westaway, D., Cooper, C., Turner, S., Da Costa, M., Carlson, G. A., and Prusiner, S. B. (1994) Structure and polymorphism of the mouse prion protein gene. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 6418–6422
 24. Westaway, D., Zuliani, V., Cooper, C. M., Da Costa, M., Neuman, S., Jenny, A. L., *et al.* (1994) Homozygosity for prion protein alleles encoding glutamine-171 renders sheep susceptible to natural scrapie. *Genes Dev.* **8**, 959–969
 25. Li, G., and Bolton, D. C. (1997) A novel hamster prion protein mRNA contains an extra exon: increased expression in scrapie. *Brain Res.* **751**, 265–274
 26. Lee, I. Y., Westaway, D., Smit, A. F., Wang, K., Seto, J., Chen, L., *et al.* (1998) Complete genomic sequence and analysis of the prion protein gene region from three mammalian species. *Genome Res.* **8**, 1022–1037
 27. Zhang, S., Samocha, K. E., Rivas, M. A., Karczewski, K. J., Daly, E., Schmandt, B., *et al.* (2018) Base-specific mutational intolerance near splice sites clarifies the role of nonessential splice nucleotides. *Genome Res.* **28**, 968–974
 28. Noderer, W. L., Flockhart, R. J., Bhaduri, A., Diaz de Arce, A. J., Zhang, J., Khavari, P. A., *et al.* (2014) Quantitative analysis of mammalian translation initiation sites by FACS-seq. *Mol. Syst. Biol.* **10**, 748
 29. Moreno, J. A., Radford, H., Peretti, D., Steinert, J. R., Verity, N., Martin, M. G., *et al.* (2012) Sustained translational repression by eIF2 α -P mediates prion neurodegeneration. *Nature* **485**, 507–511
 30. GTEx Consortium (2020) The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330
 31. Blakes, A. J. M., Wai, H. A., Davies, I., Moledina, H. E., Ruiz, A., Thomas, T., *et al.* (2022) A systematic analysis of splicing variants identifies new diagnoses in the 100,000 Genomes Project. *Genome Med.* **14**, 79
 32. Nagy, E., and Maquat, L. E. (1998) A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends Biochem. Sci.* **23**, 198–199
 33. Rivas, M. A., Pirinen, M., Conrad, D. F., Lek, M., Tsang, E. K., Karczewski, K. J., *et al.* (2015) Human genomics. Effect of predicted protein-truncating genetic variants on the human transcriptome. *Science* **348**, 666–669
 34. Aliouat, A., Hatin, I., Bertin, P., François, P., Stierlé, V., Namy, O., *et al.* (2020) Divergent effects of translation termination factor eRF3A and nonsense-mediated mRNA decay factor UPF1 on the expression of uORF carrying mRNAs and ribosome protein genes. *RNA Biol.* **17**, 227–239
 35. Rogalska, M. E., Vivori, C., and Valcárcel, J. (2023) Regulation of pre-mRNA splicing: roles in physiology and disease, and therapeutic prospects. *Nat. Rev. Genet.* **24**, 251–269
 36. Childs-Disney, J. L., Yang, X., Gibaut, Q. M. R., Tong, Y., Batey, R. T., and Disney, M. D. (2022) Targeting RNA structures with small molecules. *Nat. Rev. Drug Discov.* **21**, 736–762
 37. Bennett, C. F., Kordasiewicz, H. B., and Cleveland, D. W. (2021) Antisense drugs Make Sense for Neurological diseases. *Annu. Rev. Pharmacol. Toxicol.* **61**, 831–852
 38. Musunuru, K., Chadwick, A. C., Mizoguchi, T., Garcia, S. P., DeNizio, J. E., Reiss, C. W., *et al.* (2021) In vivo CRISPR base editing of PCSK9 durably lowers cholesterol in primates. *Nature* **593**, 429–434
 39. Davis, J. R., Wang, X., Witte, I. P., Huang, T. P., Levy, J. M., Raguram, A., *et al.* (2022) Efficient in vivo base editing via single adeno-associated viruses with size-optimized genomes encoding compact adenine base editors. *Nat. Biomed. Eng.* **6**, 1272–1283
 40. Jafar-Nejad, P., Powers, B., Soriano, A., Zhao, H., Norris, D. A., Matson, J., *et al.* (2021) The atlas of RNase H antisense oligonucleotide distribution and activity in the CNS of rodents and non-human primates following central administration. *Nucleic Acids Res.* **49**, 657–673
 41. Reidenbach, A. G., Mesleh, M. F., Casalena, D., Vallabh, S. M., Dahlin, J. L., Leed, A. J., *et al.* (2020) Multimodal small-molecule screening for human prion protein binders. *J. Biol. Chem.* **295**, 13516–13531
 42. Büeler, H., Fischer, M., Lang, Y., Bluethmann, H., Lipp, H. P., DeArmond, S. J., *et al.* (1992) Normal development and behaviour of mice lacking the neuronal cell-surface PrP protein. *Nature* **356**, 577–582

A cryptic exon to lower PrP

43. Bremer, J., Baumann, F., Tiberi, C., Wessig, C., Fischer, H., Schwarz, P., *et al.* (2010) Axonal prion protein is required for peripheral myelin maintenance. *Nat. Neurosci.* **13**, 310–318
44. Benestad, S. L., Austbø, L., Tranulis, M. A., Espenes, A., and Olsaker, I. (2012) Healthy goats naturally devoid of prion protein. *Vet. Res.* **43**, 87
45. Richt, J. A., Kasinathan, P., Hamir, A. N., Castilla, J., Sathiyaseelan, T., Vargas, F., *et al.* (2007) Production of cattle lacking prion protein. *Nat. Biotechnol.* **25**, 132–138
46. Minikel, E. V., Vallabh, S. M., Lek, M., Estrada, K., Samocha, K. E., Sathirapongsasuti, J. F., *et al.* (2016) Quantifying prion disease penetrance using large population control cohorts. *Sci. Transl. Med.* **8**, 322ra9
47. Minikel, E. V., Karczewski, K. J., Martin, H. C., Cummings, B. B., Whiffin, N., Rhodes, D., *et al.* (2020) Evaluating drug targets through human loss-of-function genetic variation. *Nature* **581**, 459–464
48. Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., *et al.* (2002) The human genome browser at UCSC. *Genome Res.* **12**, 996–1006
49. Polymenidou, M., Moos, R., Scott, M., Sigurdson, C., Shi, Y.-Z., Yajima, B., *et al.* (2008) The POM monoclonals: a comprehensive set of antibodies to non-overlapping prion protein epitopes. *PLoS One* **3**, e3872
50. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682

SUPPORTING INFORMATION

Modulation of prion protein expression through cryptic splice site manipulation

Juliana E. Gentile^{1,2}, Taylor L. Corridon^{1,2}, Meredith A. Mortberg^{1,2}, Elston Neil D'Souza³, Nicola Whiffin^{3,4}, Eric Vallabh Minikel^{1,2}, Sonia M. Vallabh^{1,2}

Supplementary tables 1-6 provided as a separate Excel file.

Mouse	ttaaag gactcct-gag--tatattcagaactgaaccatttcaaccagact--gaagcattctg--ccttcttag-tggtaccagttccaattt-aggag-agcca-agcagact gtaggt
Golden hamster	ttgaag gactcct-gaa--tatattcacaactgaacaatttcaactgagct--gaagtaactctg--ttttcttag-aggtaaccagttcagttt-aggag-agtcacagcagatc gtaagt
Sheep	ttaaag gactcct-gaa--tatattgaaactgaacatttcaaccaagct--gaagca-ctg--tcttccag-agacacagatccaacttgagctg-atacagcagat- gtaggt
Human	ttaaag gactcct-gaa--tatattcacaactgaacaatttcaaccagctc--tgagctttccg--tcttctgg-aggcacaactctagttt-agctg-aaccacaacagatt gtacat
Chimp	ttaaag gactcct-gaa--tatattcacaactgaacaatttcaaccagctc--tgagctttctg--tcttctgg-aggcacaactctagttt-agctg-aaccacaacagatt gtacat
Gorilla	ttaaag gactcct-gaa--tatattcacaactgagtaatttcaaccagctc--taagctttctg--tcttctgg-aggcacaactctagttt-agctg-aaccacaacagatt gtacat
Orangutan	ttaaag gactcct-gaa--tatattcacaactgaacaatttcaaccagctc--taagctttctg--tcttctgg-aggcacaactctagttt-agctg-aaccacaacagatt gtacat
Gibbon	ttaaag gactcct-gaa--tatattcagaactgaacaatttcaaccagctc--taagctttctg--tcttctgg-aggcacaactctagttt-agctg-aaccacaacagatt gtacat
Rhesus	ttaaag gactcct-gaa--tatattcacaactgaacaatttcaaccagctc--taagctttctg--tcttctgg-aggcacaactctagttt-agctg-aaccacaacagatt gtacat
Crab-eating macaque	ttaaag gactcct-gaa--tatattcacaactgaacaatttcaaccagctc--taagctttctg--tcttctgg-aggcacaactctagttt-agctg-aaccacaacagatt gtacat
Baboon	ttaaag gactcct-gaa--tgttttcacaactgaacaatttcaaccagctc--taagctttctg--tcttctgg-aggcacaactctagttt-agctg-aaccacaacagatt gtacat
Green monkey	ttaaag gactcct-gaa--tatattcacaactgaacaatttcaaccagctc--taagctttctg--tcttctgg-aggcacaactctagttt-agctg-aaccacaacagatt gtacat
Marmoset	ttaaag gactcct-gaa--acittttcacaactgaacaatttcaaccagctc--taagctttctg--tcttctgg-aggcacaactctagttt-agctg-aaccacaacagatt gtacat
Squirrel monkey	ttaaag gactcct-gaa--acittttcacaactgaacaatttcaaccagctc--taagctttctg--tcttctgg-aggcacaactctagttt-agctg-aaccacaacagatt gtacat
Bushbaby	ttgaag gactcct-gaa--tatgtgtcacaactgaacaatttcaaccagctc--gaagcattctg--tcttctgg-aggcacaactcaaacct-aactg-agctgcaacagatg gtaagt
Chinese tree shrew	ttaaag aactcct-gaa--taacctttaaatggaacatttccggc-----taggcattctg--tcttctgg-aggcacaactctagttt-agctg-aaccacaacagatt gtacat
Squirrel	tttaa gactcct-gaa--tatacctca-aactgaacaatttcaaccagctc--gaagcattctg--tgttcaag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Prairie vole	ttgaag gactcct-gaa--tatattcacaactgaacaatttcaaccagctc--gaagcattctg--ccttcttag-tggtacc-aggtaaccagctt-aggag-agtcacaacagatt gtacat
Chinese hamster	ttcaag gactcct-gaa--tatattcacaactgaacaatttcaaccagctc--gaagtaactctg--ccttcttag-agtaccagttcaagttt-aggag-agtcacaacagatt gtacat
Rat	ttaaag gactcct-gaa--tatattcacaactgaacaatttcaaccagctc--gaagcattctg--ccttcttag-cggtaccagttccagttt-aggag-agcca-agccagact gtaagt
Naked mole-rat	ttgaag ggcgcg-gaa--tatattcacaactgagtgatttcaaccagctc--aaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Guinea pig	ttgaag gaogcct-gagtatatatttcaaacctgaacaatttcaaccagctc--aaagcattctg--tcttcttag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Chinchilla	ttgaag gaogcct-gaa--atatttgaaacctgaacatttcaaccagctc--aaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Brush-tailed rat	ttgaag gactcct-gaa--tatattcacaactgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Rabbit	ttcaag gactcct-ga-----ccttggaaactgagcatttcaaccagctc--gaagcaactctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Pika	ttaaag gactcct-gaa--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Pig	tctaa gactcct-gaa--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Alpaca	ctaaag gactcct-gaa--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Bactrian camel	ttaaag gactcct-gaa--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Dolphin	ttaaag gactcct-gaa--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Killer whale	ttaaag gactcct-gaa--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Tibetan antelope	ttaaag gactcct-gaa--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Cow	ttaaag gactcct-gaa--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Domestic goat	ttaaag gactcct-gaa--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Horse	ttcaag gactcct-gaa--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
White rhinoceros	ttaaag gactcct-gaa--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Cat	ttaaag gactcct-gaa--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Dog	ttaaag gactcct-gaa--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Ferret	ttaaag gactcct-gag--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Panda	ttaaag gactcct-ggt--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Pacific walrus	ttaaag gactcct-gag--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Weddell seal	ttaaag gactcct-gag--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Black flying-fox	ttcaag gactcct-gca--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Megabat	ttcaag gactcct-gca--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
David's myotis (bat)	ttaaag gactcct-gaa--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Little brown bat	ttaaag gactcct-gaa--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Big brown bat	ttaaag gactcct-gaa--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Hedgehog	ttaaag gactcct-gaa--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Shrew	gttaaag aactcct-gaa--taaatgtgaacccaagcgtttccgccaagctc--aaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Star-nosed mole	ttaaag gactcct-gaa--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Elephant	ttcaag gactcct-gaa--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Cape elephant shrew	ttcaag gactcct-gaa--taaatgtgaacccaagcgtttccgccaagctc--aaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Manatee	ttcaag gactcct-gaa--tatatttgaaacctgaacaatttcaaccagctc--gaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Cape golden mole	ttcaag gactcct-gaa--tatatttcaaac-----caggtttcaaccagctc--aaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Tenrec	ttcaag gactcct-gaa--tatatttcaaac-----caggtttcaaccagctc--aaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat
Aardvark	ttcaag gactcct-gaa--tatatttcaaac-----caggtttcaaccagctc--aaagcattctg--tcttccag-aggcacaactctagttt-agctg-agtcacaacagatt gtacat

Figure S1. Multiple alignment of PRNP exon 2 orthologous sequence for all available eutherian mammals. As in Figure 1D, but including eutherian mammals without ATGs in exon 2. Lesser Egyptian jerboa is excluded because orthologous sequence was identified for only part of exon 2.

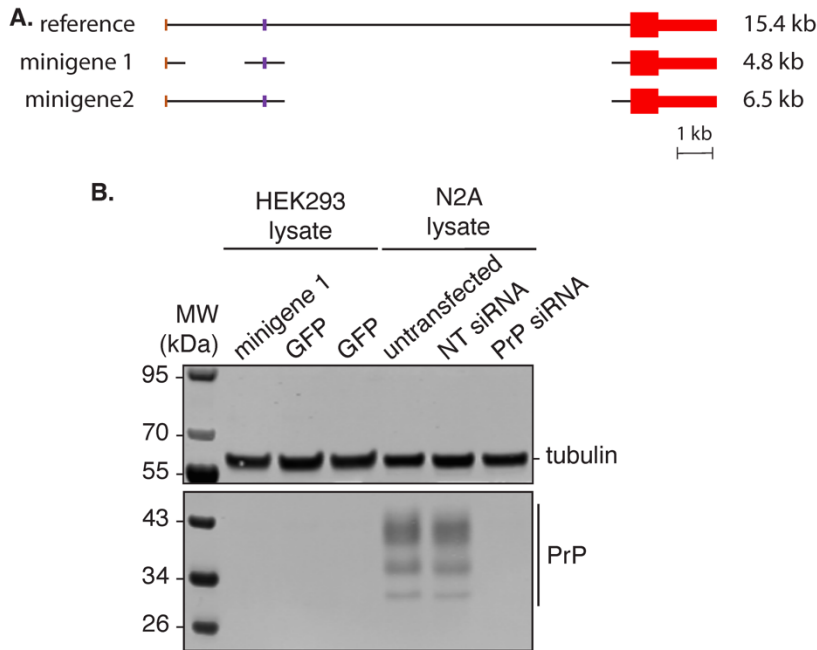


Figure S2. Alternative minigene construct tested in cells. **A)** Comparison of human reference sequence with an alternative “minigene 1” containing only 500 bp at either end of intron 1, and the “minigene 2” used throughout the main text of this manuscript. **B)** Immunoblot failing to detect any expression of minigene 1 in transfected HEK293 cells. Primary antibody: 6D11, see Methods.

	exon 2 splice sites	
	acceptor	donor
template (wild-type)	TTTAAG GACTCC...TCAGCCATGTCTGAG...CAGATT	GTACAT
canonical splice site (ss)	TTTCAG GACTCC...TCAGCCATGTCTGAG...CAGCAG	GTAAGT
canonical ss- ATG mutant	TTTCAG GACTCC...TCAGCCCCCTCTGAG...CAGCAG	GTAAGT
mouse splice site (ss)	TTAAAG GACTCC...TCAGCCATGTCTGAG...CAGACT	GTGAGT
mouse ss- ATG mutant	TTAAAG GACTCC...TCAGCCCCCTCTGAG...CAGACT	GTGAGT

Figure S3. Variants with the exon 2 uORF abolished. As in Figure 3A, but with additional mutants replacing the ATG with CCC.