Influence of Friedreich Ataxia GAA Noncoding Repeat Expansions on Pre-mRNA Processing

Marco Baralle,[1,2] Tibor Pastor,[1,2] Erica Bussani,[1] and Franco Pagani[1,*]

The intronic GAA repeat expansion in the *frataxin* (*FXN*) gene causes the hereditary neurodegenerative disorder Friedreich ataxia. Although it is generally believed that GAA repeats block transcription elongation, direct proof in eukaryotic systems is lacking. We tested in hybrid minigenes the effect of GAA and TTC repeats on nascent transcription and pre-mRNA processing. Unexpectedly, disease-causing GAA_{100} repeats did not affect transcriptional elongation in a nuclear HeLa Run On assay, nor did they affect pre-mRNA transcript abundance. However, they did result in a complex defect in pre-mRNA processing. The insertion of GAA but not TTC repeats downstream of reporter exons resulted in their partial or complete exclusion from the mature mRNAs and in the generation of a variety of aberrant splicing products. This effect of GAA repeats was observed to be position and context dependent; their insertion at different distances from the reporter exons had a variable effect on splice-site selection. In addition, GAA repeats bind to a multitude of different splicing factors and induced the accumulation of an upstream pre-mRNA splicing intermediate, which is not turned over into mature mRNA. When embedded in the homologous frataxin minigene system, the GAA repeats did not affect the pre-mRNA transcript abundance but did significantly reduce the splicing efficiency of the first intron. These data indicate an association between GAA noncoding repeats and aberrant pre-mRNA processing because binding of transcribed GAA repeats to a multitude of *trans*-acting splicing factors can interfere with normal turnover of intronic RNA and thus lead to its degradation and a lower abundance of mature mRNA.

Introduction

Friedreich ataxia (FRDA [MIM 229300]) is the most prevalent inherited ataxia[1], with an estimated prevalence of 2–4/100,000. The neurodegenerative phenotype, associated with primary degeneration of dorsal root ganglia, is characterized by progressive sensory ataxia with onset before 25 years of age, areflexia, dysarthria, lower-limb areflexia, decreased vibration sense, muscle weakness in the legs, and extensor plantar response.[2,3] Diabetes, skeletal abnormalities, and hyperthrophic cardiomyopathy are nonneurological symptoms, and heart failure is a frequent cause of death. FRDA is an autosomal-recessive disorder, and most patients are homozygous for a noncoding expanded GAA triplet repeat sequence in the frataxin (FXN [MIM 606829] gene.[4] This gene is composed of seven exons spanning about 80 kb of genomic DNA, and the GAA repeats are located in the middle of an Alu sequence in the ~11-kb-long first intron. Normal alleles have only a small number of GAA trinucleotide repeats (usually 8–33), whereas expanded alleles contain more than 90 repeats.[1,5] An inverse correlation between the length of the GAA expansion and the age of onset and severity of the disease has been observed.[6,7] Interestingly, large expansions of GAA repeats are specific to mammals, and the human genome contains other genes with intronic expandable and potential pathogenic GAA triplets.[5,8] FRDA patients have a marked deficiency of frataxin mRNA[9] that eventually causes a significant reduction in the amount of frataxin protein, which is a highly conserved protein associated with the inner mitochondrial membrane and is involved in mitochondrial iron metabolism.[9]

Biochemical studies have shown that long GAA repeats can adopt in vitro non-B-DNA structure or sticky DNA[10–15], which can interfere with gene transcription.[11,12,14,16,17] However, the majority of studies have evaluated transcription by using phage or bacterial polymerases.[14,16,17] These prokaryotic systems cannot efficiently process the majority of eukaryotic pre-mRNAs; for example, transcripts derived from T7 (and also polIII) are not efficiently capped, spliced, and polyadenylated[18,19] because these polymerases lack the polII-specific carboxy-terminal domain, which provides a unique platform that connects pre-mRNA processing with transcription.[20–22] Indeed, indirect measurement of transcription efficiency in eukaryotic systems showed contradicting results with reduced transcription[11,12,14] or no effect.[14,23] GAA triplets have also been found to interfere with DNA replication[10,12,24] and to mediate position-effect variegation.[25] Recently, they have been also associated with a transcriptionally silent chromatin.[26,27]

Accurate mRNA biosynthesis requires both the classical splicing signals (the 5′ and 3′ splice sites, the branch-point and polypyrimidine sequences[31]) and a large number of highly degenerate intronic and exonic *cis*-acting regulatory elements.[28–30] The latter have been found to be a frequent target of mutations that cause aberrant splicing.[28,30,32] For example, through the specific binding of intronic repetitions to splicing factors, changes in the number of these repetitions are known to cause different human pathologies affecting pre-mRNA splicing. Transcribed GU, CA, and expanded CUG repeats are involved in cystic fibrosis (CF [MIM 219700]),[33] in the regulation of the endothelial nitric oxide synthase gene (eNOS

[1]International Centre for Genetic Engineering and Biotechnology, Padriciano 99, Trieste 34012, Italy
[2]These Authors contributed equally to this work.
*Correspondence: pagani@icgeb.org
DOI 10.1016/j.ajhg.2008.06.018.

[MIM 163729])[34], and in myotonic dystrophy (DM [MIM 160900]),[32] respectively. Interestingly, some intronic splicing regulatory sequences have been recently shown to facilitate the formation of pre-mRNA splicing intermediates acting selectively on the splicing efficiency of upstream or downstream introns.[35]

In this study, we have explored the effect of the GAA repeats on polII transcription and pre-mRNA processing by using hybrid minigenes as a model system. Our results indicate that GAA repeats can be efficiently transcribed and inserted in nascent pre-mRNAs, where they can bind to multiple splicing factors and affect the pre-mRNA processing in a position- and context-dependent manner. To explain the composite effect of the repeats, transcribed GAA triplets could form a "sticky RNA," which interferes with normal intronic pre-mRNA processing.

Material and Methods

Plasmid Construction

The pEDA minigene was previously described[36] and contains the genomic fibronectin EDA exon, its flanking introns, and part of the −1 and +1 exons embedded in the human α-globin gene under the control of the SV40 enhancer. To generate the pEDA Nco, pEDA Nde, and pEDA BclI, we individually modified the corresponding unique restriction sites in the downstream EDA intron to the KpnI site by inserting specific oligonucleotides. We inserted a 40 bp spacer derived from pBS in the HindIII site of the second α-globin exon to obtain pEDA BglII. For generating the pBRCA1 ex18 minigene, a BRCA1 exon 17, exon 18, and exon 19 cassette was amplified by PCR and cloned in the BstEII site in the third exon of the α-globin minigene.[37] A frataxin cassette, which included 180 nucleotides of the first exon, 1170 bases of the flanking intron, the last 786 bases of the intron, and 78 bases of the frataxin exon 2, was amplified in two fragments with Fra395Dir and Fra1750RevKpn, Fra10314Dir, and Fra11100RevXba, respectively (oligonucleotide sequences are provided in the Supplemental Data), and cloned in pCDNA3 to obtain the pFrx minigene. One hundred GAA or TTC repeats were obtained by PCR amplification with primers C27 dir and C27Rev on genomic DNA and inserted in the unique KpnI site sites of pEDA and pFrx constructs. The same number of repeats were amplified with specific primers and cloned in the unique BamHI and EcoRI sites of pCF ex9[38] and pBRA,[37] respectively. The human genomic frataxin sequence containing 217 GAA repeats was amplified from human FRDA transgenic YG22 mouse DNA[39] with primers GAA1724Dir and GAA2192Rev, digested with KpnI, and cloned in sense and antisense orientation in the pEDANco minigene to generate pEDANcoGAA217 and pEDANcoTTC217. Oligonucleotides with fifteen GAA or TTC repeats were inserted in pEDA Nde to generate pEDANde15TTC and pEDANde15GAA. All minigenes were verified by sequencing, and the correct size of repeat insertions was checked for every plasmid preparation.

Transfection, RT-PCR Analysis, Northern Blotting, and Nuclear Run On Assay

COS7 cells (2.5×10^5) or HeLa cells p100 (5×10^6) were grown in standard conditions and transfected with Effectene reagent with 1.5 μg or 3 μg of plasmid DNA, respectively.[40] RNA extraction, RT-PCR, and quantitation of the amplified products were done as previously described.[40] For the analysis of spliced forms, pEDA minigenes were amplified with primers α2-3 and EDA5, pBRCA1 exon 18 was amplified with BRC90 and glo800, pCF exon 9 was amplified with α2-3 and BRA2[38], and pFrx was amplified with Fra395Dir and Sp6. For nascent transcript analysis, total RNA was digested with DNase RNase free (Roche), purified with the RNeasy kit (QIAGEN), and subsequently amplified with specific primers. For the EDA minigene, we used the following: for u1, EDA1207D and EDA1251Rev; for d1, EDA+1F and EDA1940Rev; and for d2, EDA2764Dir and glo800. For the pFrx minigene, we used the following oligonucleotides: for U, Fra395Dir and Fra606Rev; for D, Fra10902Dir and Sp6; and for D1, Fra10754Dir and Fra11037Rev. For turnover of splicing intermediates, we amplified the transfected minigenes with α936 and EDA1251R to detect the upstream intermediates and with EDA1207D and glo395 to detect the downstream intermediates. The relative abundance of spliced and unspliced forms in the Frx minigene was analyzed by RT-PCR with primers specific for the flanking exons (Fra395Dir and Sp6) and/or for the intronic sequences (Fra606Rev and Fra10902Dir). To detect the spliced and U unspliced forms, we amplified the transfected pFrx minigenes with Fra395Dir, Fra606Rev, and Sp6 primers. Amplification with Fra395Dir, Fra10902Dir, and Sp6 primers allowed detection of the spliced and D unspliced forms. The amplified products were resolved on 2% agarose gels, and the relative efficiency of splicing was measured as the ratio between the intensity of the band originating from the mature spliced transcript and that originating from the unspliced pre-mRNAs.

For Northern analysis, total RNA (2 μg) was resolved on 1.2% agarose-formaldehyde gel, and blotted membranes were hybridized with α^{32}PdCTP-labeled α-globin cDNA or with the 270 bp EDA fibronectin alternative spliced exon.

For NRO analysis, nuclei from HeLa-transfected cells were isolated in hypertonic lysis buffer (10 mM Tris HCl [pH 7.5], 10 mM NaCl, 2.5 mM $MgCl_2$, 0.5% NP40, and 10% sucrose) and resuspended in transcription buffer (40 mM Tris HCl [pH 7.9], 500 mM KCl, 10 mM MgCl2, 40% glycerol, and 2 mM DTT) in the presence of ^{32}P pUTP and ^{32}P pGTP for 15 min at 30°C. After RNA extraction, the labeled nascent RNAs were isolated and hybridized to antisense single-strand M13 probes slot blotted on nylon membrane. Radioactive signals were quantitated on a Cyclon Imager. We obtained the A-to-F M13 probes spanning the pEDA minigene by cloning the different PCR-amplified sequences in M13 plasmid. The VA M13 probe for normalization of transfection efficiencies was a generous gift from N. Proudfoot. Each transfection was VA normalized, and data were expressed for each probe as a percentage relative to the values obtained with the pEDA construct.

Affinity Purification of RNA-Binding Proteins: Pull-Down Assay

Two synthetic $(GAA)_{10}$ and $(UUC)_{10}$ RNA oligonucleotides (Integrated DNA Technologies) were used as targets for pull-down assays. An amount of 10 μg of either target RNA oligo was placed in 400 μl of reaction mixture (100 mM NaAcetate [pH 5.0] and Sigma 5 mM sodium m-periodate), incubated for 1 hr in the dark at room temperature, precipitated with ethanol, and resuspended in 100 μl of 100 mM sodium acetate (pH 5.0). Approximately 400 μl of adipic acid dehydrazide agarose beads 50% slurry (Sigma) previously equilibrated with 100 mM sodium acetate

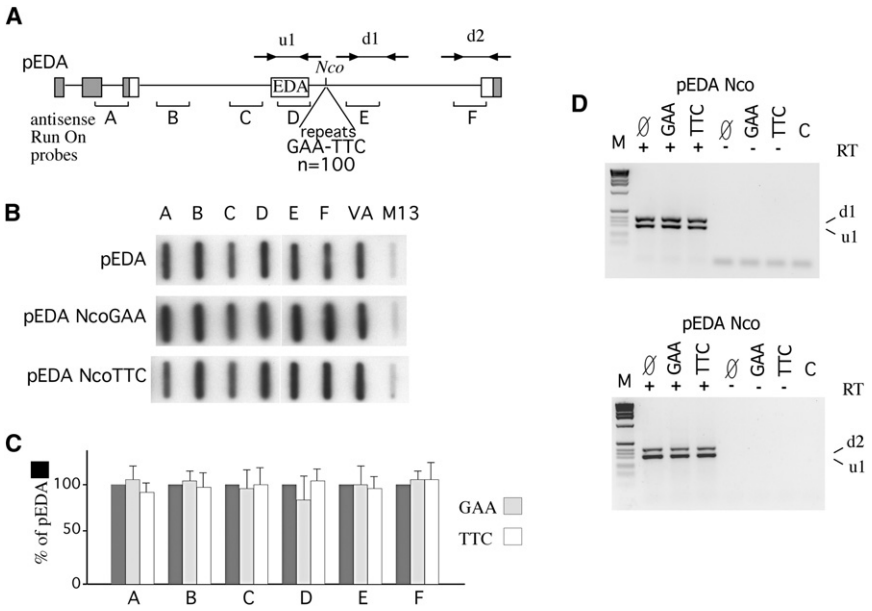

Figure 1. Expanded GAA- and TTC-Triplet Repeats Did Not Affect the Levels of Nascent Transcription

(A) Schematic representation of pEDA minigenes. Gray and white boxes correspond to α-globin and fibronectin exonic sequences, respectively; lines indicate intronic sequences. The position of the GAA and TTC repeats is indicated. The position of the u1, d1, and d2 nascent-transcript amplification products is indicated. The length and position of the M13 antisense Nuclear Run On probes relative to the minigenes are shown.

(B) Typical NRO analysis of HeLa cells. Cells were transfected with one of the indicated minigenes together with the cotransfection control VA construct. M13 corresponds to the empty M13 vector. Antisense single-strand M13 probes (A to F and VA) are indicated at the top.

(C) The graph shows the VA-normalized quantitation of three independent transfections relative to the values obtained with the pEDA (black columns) constructs over the corresponding probes. Data are expressed as means + SD. Light-gray and white columns show transcription levels obtained with pEDA NcoGAA and pEDA NcoTTC minigenes. All signals are corrected for the background hybridization shown in the M13-negative control.

(D) Analysis of pre-mRNA transcript abundance. The indicated minigenes were transfected in COS cells, and the resulting RNAs were coamplified with the u1-d1 (top) or u1-d2 (bottom) primer set and resolved on agarose gels. The pEDA without repeats is indicated with Ø. Controls without reverse transcriptase (RT) and untransfected cells (C) are indicated. M is the molecular 1 Kb marker.

(pH 5.0) was added to each periodate-treated RNA, and the mix was incubated for 12 hr at 4°C on a rotator.

The beads with the bound RNA were then washed three times with 1 ml of 2 M NaCl, equilibrated in 1× solution D (20 mM HEPES [pH 7.5], 100 mM KCl, 0.2 mM EDTA, 0.5 mM DTT, and 6% glycerol) and incubated, in a final volume of 500 μl, with 0.5 mg of HeLa cell nuclear extract (C4, Biotech), 10× solution D (200 mM HEPES [pH 7.5], 2 mM EDTA, 5 mM DTT, and 60% glycerol), 160 mM KCl, and Heparin (final concentration 2.5 μg/μl), for 30 min on a rotator at room temperature. The beads were then washed six times with 1 ml of 1× solution D before the addition of SDS sample buffer and loading on SDS-12% polyacrylamide gels. Proteins were visualized by Coomassie brilliant blue staining. Protein sequence analysis of the bands excised from the gel was performed with an electrospray ionization mass spectrometer (LCQ DECA XP-ThermoFinnigam). Protein bands were digested with trypsin, and the resulting peptides were extracted with water and 60% acetonitrile/1% trifluoroacetic acid. Fragments were then analyzed by mass spectrometry, and proteins were identified by analysis of the peptide MS/MS data with Turbo SEQUEST (ThermoFinnigam) and MASCOT (Matrix Science).

Immunoblots were performed onto PVDF membrane (Amersham Bioscience), which was blocked by incubation in 5% dried milk in 0.1 M PBS, 0.1% Tween-20, or for Mab 1H4, in 1× Western Blocking Reagent (Roche). Membranes were probed with primary antibodies raised against hnRNPA1 and hnRNPA2 (kindly provided by R. Klima, ICGEB, Italy; 1:1000), PTB (a kind gift from E. Buratti, ICGEB, Italy; 1:1000), ZNF 9 (kindly provided by B. Cardinali; 1:1000), PABPN1 (a kind gift from E. Wahle, Halle, Germany; 1:1000), ASF/SF2 (Zymed Laboratories; 1:1000), SR proteins (Map 1H4, Zymed Laboratories; 1:1000) and Tra2β (Abcam; 1:1000). Immunoreactive bands were detected with enhanced chemiluminiscence reagent (ECL; Amersham Biosciences).

Results

GAA Repeats Did Not Affect Transcription Elongation or Pre-mRNA Transcript Abundance in Transfected Cells

To test the effect of GAA repeats on PolII elongation and pre-mRNA transcript abundance, we used the alternatively spliced fibronectin pEDA minigene. This minigene has been extensively studied[36,41] and provides a good model for studying the process of cotranscriptional splicing and splicing regulatory sequences.[42–44] One hundred GAA or TTC repeats were introduced in the NcoI site of the pEDA construct (Figure 1A). This size expansion, found in affected individuals, has been shown to form abnormal DNA structures in vitro[13,14,45] and to inhibit prokaryotic polymerase transcription.[16] We performed a nuclear run on (NRO) assay to determine whether some differences in the transcription level could be observed between minigenes carrying the GAA or TCC expansion and the empty minigene. Analysis was performed with single-strand, antisense DNA probes that cover the minigene sequences upstream and downstream of the triplet repeats (Figure 1A). Cotransfection with the VAI-expressing plasmid was used as a control for transfection efficiency. NRO signals obtained for the GAA, TTC, or empty pEDA minigenes were comparable relative to the VAI control signal, resulting in no changes in nascent transcription (Figure 1B). Quantitation of three independent VAI normalized transfections confirmed that the insertion of either the GAA or TTC repeat did not affect pre-mRNA transcription levels (Figure 1C), as shown by the

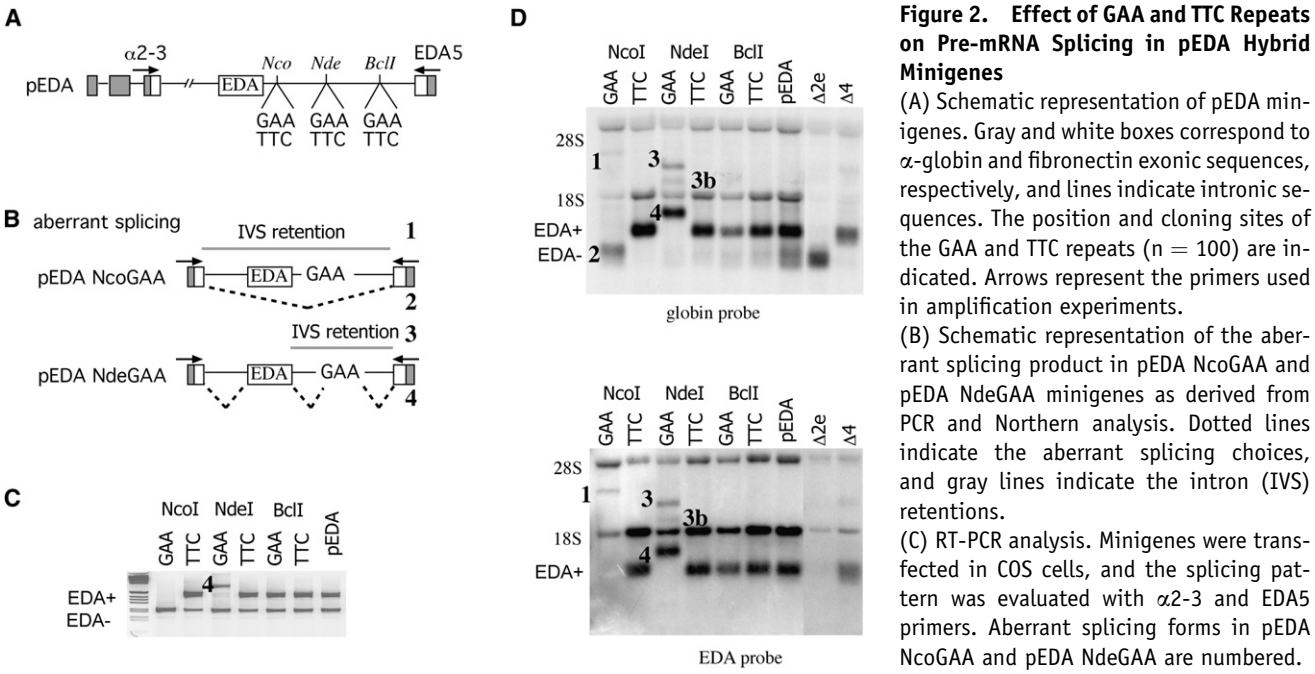

Figure 2. Effect of GAA and TTC Repeats on Pre-mRNA Splicing in pEDA Hybrid Minigenes

(A) Schematic representation of pEDA minigenes. Gray and white boxes correspond to α-globin and fibronectin exonic sequences, respectively, and lines indicate intronic sequences. The position and cloning sites of the GAA and TTC repeats (n = 100) are indicated. Arrows represent the primers used in amplification experiments.

(B) Schematic representation of the aberrant splicing product in pEDA NcoGAA and pEDA NdeGAA minigenes as derived from PCR and Northern analysis. Dotted lines indicate the aberrant splicing choices, and gray lines indicate the intron (IVS) retentions.

(C) RT-PCR analysis. Minigenes were transfected in COS cells, and the splicing pattern was evaluated with α2-3 and EDA5 primers. Aberrant splicing forms in pEDA NcoGAA and pEDA NdeGAA are numbered.

(D) Northern analysis. Minigenes were transfected in COS cells, and the resulting RNA was analyzed by Northern blotting with an α-globin (upper) or fibronectin EDA (lower) probes. Control Δ2e and Δ4 are EDA minigene mutants that showed complete exon skipping and inclusion, respectively.[36] Ribosomal RNAs and the two alternative spliced forms with and without the EDA exon are indicated. Aberrant splicing forms in pEDA NcoI GAA and pEDA NdeI GAA are numbered.

presence of similar amounts of transcribed RNA along the two minigenes and the empty pEDA. We also evaluated the relative abundance of nascent transcript before and after the triplet insertion in each of pEDA minigenes by RT-PCR. Pre-mRNA sequences, upstream (u1) and downstream (d1 or d2) of the GAA/TTC insertion, were coamplified (u1 plus d1 and u1 plus d2) from pEDA Nco, pEDA NcoGAA, and pEDA NcoTTC transfections. The ratio of two bands resulting from of u1-d1 and u1-d2 coamplifications showed similar patterns in the pEDA Nco, pEDA NcoGAA, and pEDA NcoTTC, indicating that the type of insertion has no effect on the relative amount of nascent transcripts before and after the insertions (Figure 1D). Similarly, an increased number of triplets (n = 217) inserted in the pEDA minigene did not affect the relative abundance of the nascent transcripts before and after the insertion (Figure S1). In summary, these results clearly demonstrate that the insertion of a disease-causing GAA repeat expansion in a mammalian reporter gene did not affect the polII transcription elongation or pre-mRNA transcript abundance.

GAA Repeats Affect Pre-mRNA Splicing in a Position-Dependent Manner in a Heterologous Hybrid Minigene System

The lack of any effect on transcription in the pEDA NcoGAA minigene was unexpected, and we decided to explore in more detail the effect of the repeats on the mRNA biosynthesis. The GAA or TTC repeats (n = 100) were cloned in different intronic positions along the entire pEDA construct, and the resulting minigenes were transfected in COS cells. The quality of the resulting RNAs was evaluated in parallel by two complementary approaches: by RT-PCR analysis with specific primers (Figure 2C) and by Northern analysis (Figure 2D). RT-PCR analysis can provide an easy way to analyze aberrant splicing, but as a result of the intrinsic characteristics of PCR, large-intron retention or peculiar aberrant splicing products can be missed by this technique. Interestingly, the insertion of the GAA repeats in the NcoI position (pEDA NcoGAA) induced two major events in comparison with the corresponding TTC minigene: complete exclusion of the EDA exon, as can be seen by RT-PCR and Northern analysis (Figures 2C and 2D), and retention of upstream and downstream introns. This latter event can be well appreciated only by Northern blot (aberrant band number 1, Figure 2D). To rule out potential interference with splicing regulatory elements at the cloning Nco site, which is in close proximity of the EDA exon, we inserted the repeats 400 bp downstream at the NdeI site. pEDA Nde GAA again showed an aberrant splicing product. In RT-PCR the normal EDA inclusion product disappeared and was replaced by a higher MW band (band 4 in Figure 2C). Sequencing of band 4 showed that it contained the EDA exon and a cryptic exon, which surprisingly included the entire GAA sequence (Figure 2B). In addition, Northern analysis showed two more bands (numbers 3 and 3b in Figure 2D). Their size and hybridization pattern indicates that band 3 corresponds to complete retention of the downstream intron (Figure 2B), whereas the faint 3b band probably corresponds to intron retention of the segment upstream or downstream of the GAA

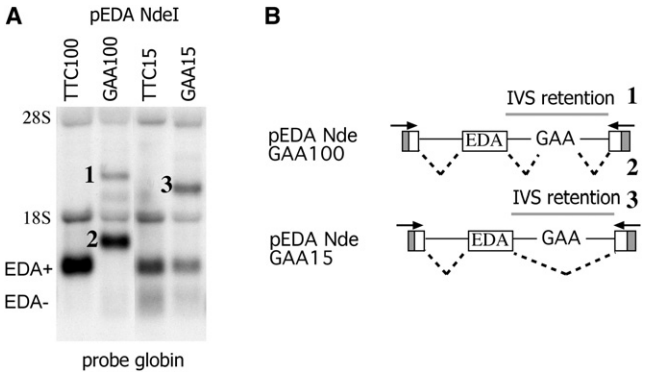

Figure 3. The Number of GAA Repeats Affects the Severity of the Aberrant Splicing Pattern

(A) Northern analysis of pEDA Nde minigenes with 100 or 15 GAA or TTC repeats. Minigenes were transfected in COS cells, and the resulting RNA was analyzed by Northern blotting with an α-globin probe. Ribosomal RNAs and the two alternative spliced forms with and without the EDA exon are indicated. Aberrant splicing forms in pEDA Nde GAA100 and pEDA NdeI GAA15 are numbered.

(B) Schematic representation of the splicing products observed in in pEDA Nde GAA100 and pEDA NdeI GAA15 as deduced from Northern analysis and the sequencing of aberrant products in RT-PCR analysis. Dotted lines indicate the aberrant splicing choices, and gray lines indicate the intron (IVS) retentions.

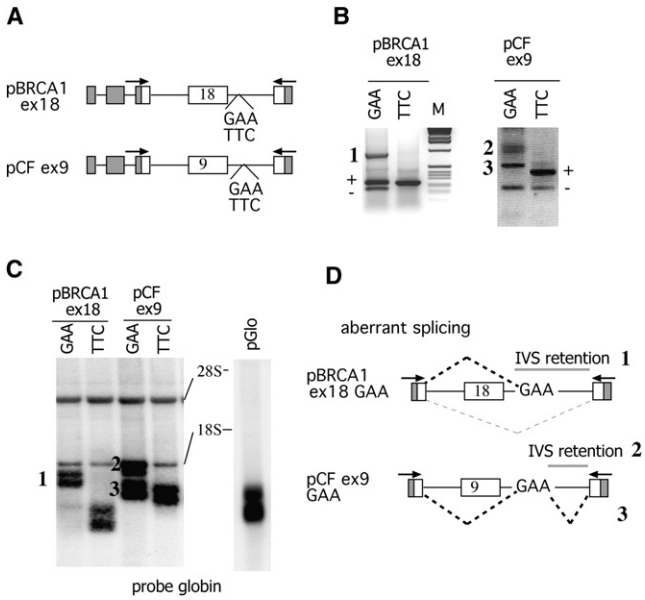

Figure 4. Effect of GAA-TTC Repeats in BRCA1 Exon 18 and CFTR Exon 9

(A) Schematic representation of minigenes containing the GAA or TTC (n = 100) repeats. Gray boxes correspond to α-globin sequence, and white boxes correspond to BRCA1 and CFTR exons. Lines indicate intronic sequences. Arrows represent the primers used in amplification experiments.

(B) RT-PCR analysis. Minigenes were transfected in COS cells, and the splicing pattern was evaluated with specific primers. Exon inclusion (+) and exclusion (-) as well as aberrant splicing (numbers) are indicated.

(C) Northern analysis. Minigenes were transfected in COS cells, and the resulting RNA was analyzed by Northern blotting with an α-globin probe. The basic pGlo minigene that contains the α-globin sequences with a short endogenous promoter showed a double band, which is due to alternative transcript initiation. Ribosomal RNAs are indicated. Aberrant splicing forms in pBRCA1ex18 GAA and pCFex9 GAA are numbered.

(D) Schematic representation of the aberrant splicing products observed in pBRCA1ex18 GAA and pCFex9 GAA. Dotted lines indicate the aberrant splicing choices, and gray lines indicate the intron (IVS) retentions.

insertion. Lastly, insertion of the repeats farther downstream had no effect on pre-mRNA processing (pEDA BclI constructs, Figure 2) suggesting that the number of GAA repeats affects splicing in a position-dependent manner.

The Number of Repeats Affects Abnormalities in Pre-mRNA Splicing

Because the number of GAA repeats is associated with disease severity in FRDA patients, we tested the effect of different repeat lengths on pre-mRNA splicing in hybrid minigenes. We created two additional pEDA minigenes by embedding 15 GAA or 15 TTC repeats within the NdeI restriction site and evaluated their splicing patterns. Compared to the construct with 100 TTC repeats, pEDA Nde$(TTC)_{15}$ had no effect on the normal splicing pattern (Figure 3A). On the contrary, pEDA Nde$(GAA)_{15}$ shifted the pre-mRNA processing toward normal splicing in comparison with pEDA Nde$(GAA)_{100}$, although this low number of repeats still has a pathological effect in the minigene in that it induces an aberrant splicing form with retention of the downstream intron (Figure 3, band 3).

GAA Repeats Affect Pre-mRNA Processing in Two Additional Heterologous Minigenes

To test whether the effect of the repeats on pre-mRNA processing was a peculiarity of the alternatively spliced fibronectin EDA minigene, we extended the study to two additional minigenes (Figure 4A). The GAA or TTC repeats were inserted downstream of the constitutively included BRCA1 exon 18 and downstream of the aberrant alternatively spliced CFTR exon 9 (Figure 4A). mRNAs were then evaluated by RT-PCR and Northern analysis. In the presence of the GAA repeats but not TTC, both BRCA1 exon 18 and CFTR exon 9 showed aberrant splicing products similar to those found in the EDA minigene. In particular, the pBRCA1 ex18GAA produces an aberrant band (number 1) that contains the GAA repeats followed by the downstream intronic sequences (Figure 4B). The pCF ex 9 GAA revealed a complex aberrant splicing pattern in which exon 9 was skipped, GAA was used as a cryptic exon, and downstream introns were retained (Figure 4). These different aberrant splicing products were also observed in Northern blots (Figure 4D) and are represented schematically in Figure 4E.

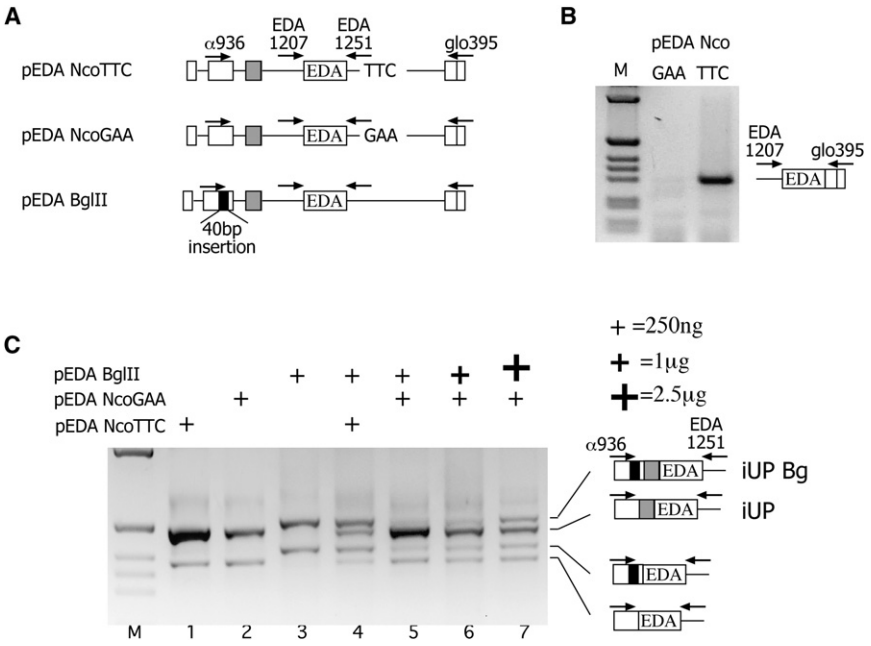

Figure 5. Turnover of Pre-mRNA Intermediates in pEDA Nco Minigenes

(A) Schematic representation of the hybrid minigenes used for evaluation of the splicing intermediates. The gray box corresponds to a hybrid α-globin fibronectin exon that is partially skipped in the upstream intermediates, and the black box is the 40 bp insertion in pEDA BglII. The indicated primer pairs α936–EDA1251R and EDA1207D-glo395 amplify, relative to the EDA exon, the upstream and downstream intermediates, respectively.

(B) Analysis of downstream intermediates. pEDA NcoGAA and pEDA NcoTTC minigenes were transfected in COS cells, and pre-mRNA was amplified with EDA1207D and glo395. The splicing product is shown, and its identity was verified by direct sequencing.

(C) Analysis and relative quantitation of upstream intermediates. Cotransfection experiments and RT-PCR amplification with α936 and EDA1251R are shown. Equal amounts of pEDA NcoGAA and pEDA NcoTTC (250 ng) were transfected in COS cells alone or in combination with different amounts of pEDA BglII (250 ng in lanes 3–5; 1 μg in lane 6; and 2 μg in lane 7). The major upstream intermediates originating from pEDA NcoGAA and pEDA NcoTTC (iUP) or from pEDA BglII (iUP Bg) are indicated.

Turnover of Pre-mRNA Splicing Intermediates in the pEDA Nco GAA Minigene

To investigate the GAA-dependent aberrant splicing mechanism more in detail, we evaluated the turnover of pre-mRNA splicing intermediates in pEDA NcoGAA and pEDA NcoTTC minigenes. Preferential activation of one pre-mRNA splicing intermediate (i.e., of an upstream or downstream intron) has been recently shown to regulate splice-site selection through differential binding of a neuronal specific splicing factor to intronic or exonic sequences.[35] In order to study the splicing intermediates in the pEDA Nco minigenes carrying the TTC or GAA repeats, we used two pairs of primers: the 1207D and glo395R pair and the α936 and EDA1251R pair. The first pair of primers detects intermediates containing the intron upstream of the EDA exon, whereas the second primer pair detects intermediates containing the intron downstream of the EDA exon. Amplification with EDA 1207D and glo395R showed a band corresponding to the removal of the TTC-containing intron of the downstream intermediate only in the pEDA TTC minigene (Figure 5B). The absence of any amplified band in pEDA NcoGAA is consistent with mature mRNA in the Northern analysis, in which this minigene showed complete exon skipping and the retention of the GAA-containing intron. On the contrary, amplification with α936 and EDA1251R showed two bands in both pEDANco GAA and pEDA NcoTTC plasmids (Figure 5C, lanes 1 and 2). The major upper band corresponds to the splicing intermediate in which the GAA- or TCC-containing intron is present and all the upstream introns are correctly removed (iUP) (Figure 5C, lanes 1 and 2, respectively), whereas the faint lower band was due to partial skipping of the −1 hybrid globin-fibronectin exon. In light of the fact that the iUP intermediate represents one of the two alternative pathways leading to exon inclusion (i.e., the intermediate in which the upstream intron is skipped), its presence in the GAA minigene was unexpected because the mature mRNA completely lacks the EDA (Figure 2). To quantify the relative amount of the iUP pre-mRNA, we set up a cotransfection experiment with the pEDABglII minigene. This pEDABglII internal-control minigene without repeats contains a 40-bp-long insertion in the second α-globin exon, and RT-PCR amplification results in a slightly higher intermediate, iUPBg (Figure 5C, lane 3). When equal amounts of the pEDA NcoGAA and pEDA NcoTTC were individually cotransfected with the same amount of pEDABglII (250 ng), the iUPBg intermediate was only evident in the pEDA NcoTTC lane, suggesting that the iUP produced by the pEDA NcoGAA is significantly reduced (Figure 3C, compare lanes 4 and 5). To confirm these data, we cotransfected pEDA NcoGAA with increasing amounts of pEDABglII (1 μg and 2.5 μg) (Figure 5, lanes 6 and 7). In light of the iUP intermediate band, cotransfection experiments showed that approximately ten times more pEDABglII plasmid is necessary to compete with the the pEDA NcoGAA than with the pEDA NcoTTC minigene (Figure 5C, lanes 5–7). Thus, the amount of iUP is significantly higher in the pEDA NcoGAA than in the pEDA NcoTTC. This indicates that the GAA repeats induce a block in the turnover of the upstream intermediate, which accumulates and is not processed into mature EDA+ mRNA, as shown by Northern blot analysis

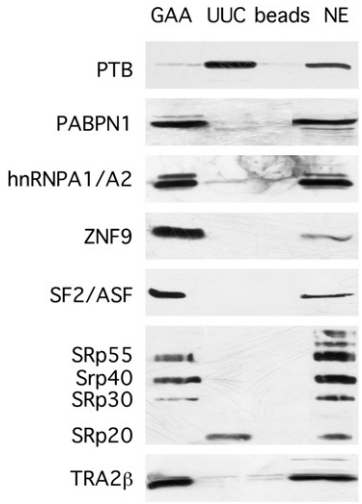

Figure 6. Binding Properties of the GAA and UUC RNAs
Immunoblot after pull-down analysis of $(GAA)_{10}$ and $(UUC)_{10}$ ribo-oligonucleotide sequences. Synthetic RNA oligonucleotides, $(GAA)_{10}$ and $(UUC)_{10}$, were used as targets for pull-down assays. The affinity-purified proteins pulled down by the RNAs were resolved on an SDS-PAGE and analyzed by immunoblotting with anti-PTB, anti-PABPN1, anti-hnRNPA1/A2, anti-ZNF9, anti-SF2/ASF, anti-SR proteins, and anti-Tra2β antibodies. The identity of the splicing factors is indicated. HeLa nuclear extract (NE) was used as a control.

(Figure 2D). One possibility is that the preferential accumulation of pre-mRNA splicing intermediates might occur through binding of *trans*-acting splicing factors in the intron.[35]

Binding Properties of GAA-Repeat Transcripts

In order to identify *trans*-acting factor(s) whose binding to GAA repeats could give us an explanation for the observed changes in the splicing pattern, we performed a pull-down analysis with synthetic RNA oligonucleotides containing either 10 GAA or 10 UUC trinucleotide repeats. We detected at least four protein bands that were enriched in the GAA-repeat fraction (data not shown). These protein bands of approximately 50, 35, and 18 kDa were excised, analyzed by mass spectrometry, and eventually identified as PABPN1 (Uniprot accession number Q86U42), a protein that is involved in nuclear polyadenylation, might contribute to mRNA export[46], and has an antiapoptotic role[47]; as hnRNPA1 and hnRNPA2 (Uniprot accession numbers P09651 and P22626, respectively), a well-known factor involved in splicing inhibition[31]; and as ZNF9 (Uniprot accession number P62633), whose function is yet to be understood.[48] The protein doublet associated with UUC oligonucleotide corresponds to the PTB (Uniprot accession number P26599), an important regulatory splicing protein, whose prevalent inhibitory effect on alternative splicing as a result of the recognition and binding to UC-rich sequences has already been well established.[49] The differential binding properties of these proteins were further confirmed by immunoblot analysis (Figure 6). We have also included in the analysis some members of SR protein family because these splicing factors have been previously reported to specifically interact with GAA-rich exonic-splicing regulatory elements.[41,50,51] Our results indicate that the GAA repeats represent the target for many different factors involved in pre-mRNA processing. More precisely, we observed their binding capacity not only to PABPN1, hnRNPA1, hnRNPA2, and ZNF9 proteins but also to several SR proteins (SF2/ASF, SRp30c [Uniprot accession number Q13242], SRp40 [Uniprot accession number Q13243], SRp55 [Uniprot accession number Q13247], and Tra2β [Uniprot accession number P62995]), yet notably not to SRp20 (Uniprot accession number P84103), which interacted with the UUC oligonucleotides (Figure 6). A similar association between GAA-rich RNA molecules and splicing factors (including several orthologous SR proteins and PAPBN1) has been detected also in *Xenopus* oocyte extracts.[52,53]

GAA Repeats Inhibit Splicing in the Frataxin Minigene

We tried to assess whether the observed effect of GAA repeats could be reproducible when the triplets were inserted in the frataxin minigene. The main characteristic of this minigene is that the GAA repeats are inserted in the first intron, which does not allow the potential skipping of the upstream exon. We made a frataxin minigene construct that carried exon 1 with 1170 bases of downstream intronic sequences and exon 2 with 786 bases of upstream intronic sequences under transcriptional control of the CMV promoter (Figure 7A). Either this empty minigene (Ø) or the $(GAA)_{100}$- and $(TTC)_{100}$-containing minigenes were transfected into COS cells. In order to evaluate the relative abundance of spliced and unspliced forms, we analyzed these constructs by RT-PCR with primers specific for the flanking exons and/or for the intronic sequences. We conducted the analysis with coamplification experiments to detect the relative efficiency of splicing measured as the ratio between the mature spliced transcript and the unspliced pre-mRNAs (see Material and Methods). Figure 7 clearly shows that the relative splicing efficiency of the frataxin minigene was significantly inhibited by the GAA repeats. In fact, the spliced-to-unspliced ratio is markedly reduced in the GAA minigene, and this is evident for both the upstream (S/U) and downstream (S/D) ratio. This effect was GAA-sequence specific given that the construct with the TTC triplets showed a splicing efficiency comparable to that of the empty one. Additional controls showed that there was no interference from contaminating DNA (lanes with minus RT) and that there was no background from endogenous Frataxin mRNA (data not shown).

In parallel, we also evaluated in the frataxin minigene the relative abundance of the nascent transcripts upstream and downstream of the triplet insertion (U and D1 in Figure 7A). As found for the EDA minigene (Figure 1D), the two bands resulting from U and D1 coamplifications showed a similar ratio independently of the type of insertion, indicating that the GAA repeats did not change the mRNA transcript

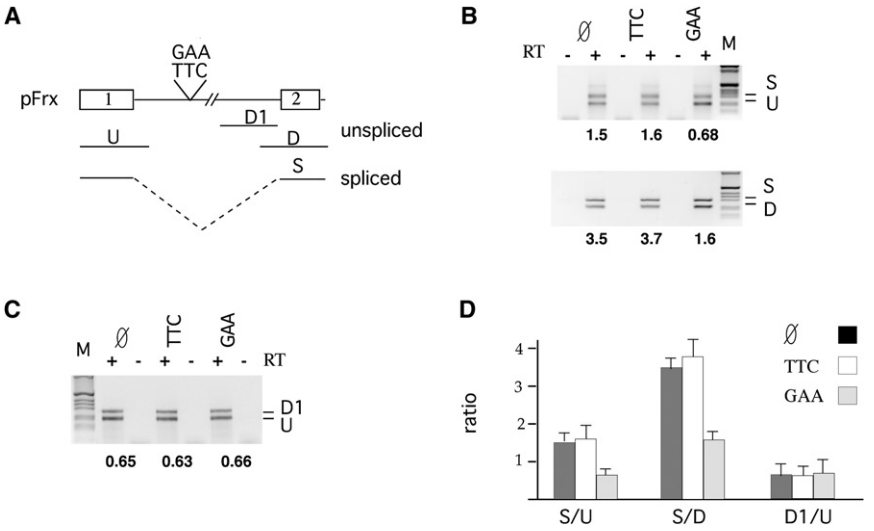

Figure 7. Effect of GAA-TTC Repeats on Pre-mRNA Splicing in the Homologous Frataxin Minigene

(A) Schematic representation of the frataxin minigene containing the GAA or TTC (n = 100) repeats and showing the unspliced (U, D, and D1) and spliced (S) amplification products.

(B) Analysis of splicing efficiency. The pFrx minigenes were transfected in COS cells, and RT-PCR analysis was performed so that the S- and U-unspliced transcripts (upper gel) or the S- and D-unspliced transcripts (lower gel) could be detected. The identity of the spliced and unspliced bands is indicated, and the number below each lane is the mean of the spliced versus unspliced ratio from four independent experiments. Controls without reverse transcriptase (RT) are indicated. M is the molecular 1 Kb marker.

(C) The pFrx minigenes were transfected in COS cells, and coamplification with U and D1 primer sets was performed so that pre-mRNA transcript abundance could be detected.

(D) The ratios between the various spliced (S) and unspliced (D and U) forms and between the unpliced D1 and U are expressed as means + SD of four independent transfection experiments.

abundance in the frataxin minigene (Figure 7C). On the other hand, analysis by RT-PCR and Northern blot did not reveal any significant effect on the mature mRNA (data not shown), indicating that the GAA repeats in their original context do not affect the quality of the final transcript. Thus, we conclude that the GAA-repeat-rich region can act as an intronic-splicing regulatory element that inhibits the splicing efficiency of the first *frataxin* intron.

Discussion

Despite extensive molecular investigations, the disease-causing mechanism of the GAA repeats in FRDA is not completely understood. GAA repeats have been associated with the formation of peculiar non-B DNA structure or sticky DNA, which is believed to interfere with nascent transcription. In this study, we unexpectedly found that the GAA-repeat expansions did not block transcription but induced a position- and context-dependent effect on pre-mRNA processing. Thus, inefficient pre-mRNA splicing rather than a direct transcriptional block is likely to be the cause of lower levels of mature frataxin mRNA in FRDA. The binding of transcribed GAA repeats to a multitude of *trans*-acting splicing factors might interfere with normal turnover of intronic RNA and thus lead to its degradation and a lower amount of mature mRNA.

Two complementary methodologies that were used for directly studying the transcription process, Nuclear Run On in HeLa cells and RT-PCR analysis on nascent transcripts, did not reveal any significant change in the polII transcript elongation and nascent transcript abundance along the minigenes (Figure 1 and Figure S1). The number of repeats we used has been previously reported to adopt a "sticky" DNA conformation in vitro, and this conformation is sufficient to inhibit T7 or Sp6 polymerase transcription.[14,16,17] By contrast, we show here that polII elongation is not influenced by the repeat-dependent "sticky" DNA. In yeast, disease-causing GAA repeats (n = 228) are also efficiently transcribed and do not induce the accumulation of truncated transcripts[23], suggesting that there are no significant differences in the ability of the yeast and mammalian transcription apparatuses to bypass GAA triplets. This is probably due to the common cotranscriptional pre-mRNA processing activity of the unique carboxy-terminal-domain structure of the large subunit of the polymerases.[20–22]

To understand the role of GAA-repeat expansions in pre-mRNA processing, we extensively evaluated the transcripts derived from several minigenes. As a first approach, we tested the effect of the GAA repeat in three artificial systems in which the GAA repeats are not normally found and subsequently analyzed the frataxin minigene that naturally carries the repeats. GAA repeats inserted in the three artificial splicing systems induced significant and complex modifications in the pre-mRNA splicing. Three major aberrant splicing events were observed to originate from the insertion of the GAA repeats. They consist of skipping of an upstream constitutive or alternative spliced exon, retention of intronic sequences, and usage of the GAA repeats as a cryptic exon. The type of aberrant splicing depends on the position used for insertion of the triplets and on the fact that in this context the repeats are not located in the first intron of the gene. Interestingly, the length of the repeat affects the severity of the splicing pattern (Figure 3), consistent with the variable phenotypic expression of the disease. On the other hand, the GAA repeats inserted in the frataxin minigene resulted in

a significant reduction in the splicing efficiency without affecting the abundance of the nascent transcript. The splicing inhibition was evident when both the upstream and downstream unspliced pre-mRNA was compared to the spliced form (Figure 6).

To clarify the mechanism of GAA-induced aberrant splicing, we studied the splicing intermediates generated during processing of pre-mRNA in one minigene in more detail. We observed that the pathological expansion induces a block in the turnover of one splicing intermediate that accumulates in the cell and is not further processed (Figure 5). This intermediate is probably highly unstable, not turned over into mature mRNA, and retained in the nucleus. The accumulation of the intermediate strongly suggests that pre-mRNA processing is blocked as a result of interference from multiple splicing factors bound to the nascent GAA-repeat transcript. Interestingly, the effect of the GAA repeats on pre-mRNA splicing intermediates is reminiscent of the recently described Nova-1 model of action on neuronal splicing regulation. In this case, the neuronal-specific Nova splicing factor regulates exon inclusion through a direct and asymmetric action on one of the flanking introns. It locally blocks or enhances the assembly of the spliceosome and facilitates the formation of pre-mRNA splicing intermediates.[35]

Our results support the notion that the expanded GAA repeats might serve as a binding platform for many different splicing factors, whose recruitment leads to the assembly of an inhibitory splicing complex. The presence of multiple *trans*-acting nuclear factors could affect the pre-mRNA processing in a context-dependent manner and lead to the production of aberrantly spliced forms. Thus, GAA repeats might be considered to be large intronic "sticky RNA" sequences within nascent transcripts. Bioinformatic[29] and experimental[41,51] evidence has shown that RNA sequences containing GAA are characteristic of many exonic splicing enhancers. The majority of splicing factors we have identified as binding to the GAA repeats are already recognized as regulators of normal or aberrant splicing. SR proteins are well-characterized activators of exon recognition and can also act as molecular retention signals of precursor mRNA.[52–54] It has been recently demonstrated in *Xenopus* oocyte extracts that GAA-rich enhancer exonic sequences bind the same splicing factors we report, namely the orthologous SR proteins and PABPN1, which actually contribute to the retention of pre-mRNA in the nucleus.[52,53] Because SR proteins facilitate exon recognition via splicing machinery and subsequent nuclear export, their inappropriate intronic positioning due to the GAA expansion might cause aberrant processing of pre-mRNA. Consequently, the turnover of splicing intermediates will not occur, and either generation of different aberrant splicing forms deriving from the heterologous minigenes or reduction in the splicing efficiency of the first intron in the frataxin minigene model system will take place. Future studies will address the mechanism by which reported GAA-binding factors can contribute to aberrant splicing and pre-mRNA nuclear retention.

The role of the context in determining the splicing outcome of the GAA repeats, as previously reported for the intronic CA dinucleotide in the eNOS gene, may explain the different phenotypic expression previously observed in mouse frataxin models. The nonconserved *frataxin* intron 1 of the mouse does not normally contain a GAA-repeat sequence, and in a knock-in mouse model, the insertion of 230 GAA repeats in the intron was not associated with any obvious pathological phenotype.[55] By contrast, the human genomic FRDA transgene with a similar GAA expansion at the correct intronic position exhibits progressive, although mild, FRDA-like pathology.[39] It would be interesting to compare the nascent transcripts and splicing efficiency in these two different mouse models and to characterize the contribution of the flanking intronic sequences.

Histone deacetylase (HDAC) inhibitors are under evaluation as potential therapeutic compounds in FRDA for their potential effect on *frataxin* transcription.[26] It is important to highlight that these drugs have been shown to be able to modify pre-mRNA splicing.[56–59] For example, in the neuromuscular disorder spinal muscular atrophy 1 (SMA1 [MIM 253300]), which results from the loss of function of the *survival of motor neuron 1* gene (*SMN1* [600354]), they restore exon 7 defective splicing of the paralogue *survival of motor neuron 2* gene (*SMN2* [601627]), which is not normally included in mRNA and contains a single synonymous mutation in exon 7. This treatment was shown to provide a sufficient amount of SMN proteins to compensate for the mutated *SMN1* gene.[56–58] The lack of a direct transcription interference of the GAA repeats combined with the changes in pre-mRNA turnover and splicing efficiency we have observed here suggests that the effect of HDAC inhibitors on pre-mRNA splicing rather than their effect on transcription might be responsible for the increased mRNA abundance. On the other hand, transfected minigenes might have a different chromatin structure and only partially mimic the recently described epigenetic changes observed in vivo in patients' cells.[26,27] The mechanism that leads to the formation of heterochromatin markers in vivo is unclear both in genes with repeat expansions and more generally in normal genes. Recent literature suggests the involvement of RNA processing in the generation of transcriptional silent chromatin[60,61], as well as a connection between chromatin modification and pre-mRNA splicing.[62] We can speculate that the GAA-mediated changes observed in the minigene transcript along with the bound splicing factors influence the epigenetic status of the *FXN* gene. Common pathways might connect the altered chromatin conformation found in patients' cells with changes in pre-mRNA processing, and the exploration of these links might provide new insight into the FRDA pathogenesis and lead to the development of new therapeutic strategies.

Supplemental Data

One figure and a table with the oligonucleotide sequences are available at http://www.ajhg.org/.

Acknowledgments

This work was supported by grant from the Associazione Italliana Ricerca sul Cancro and the Italian Cystic Fibrosis Foundation. We thank Elisa Goina for the pBRCA1 ex18 minigene, Cristiana Stuani for preparation of pEDA minigenes and technical assistance, Mark Pook for YG22 samples, and Rodolfo Garcia for helpful discussion.

Received: April 11, 2008
Revised: June 16, 2008
Accepted: June 18, 2008
Published online: July 3, 2008

Web Resources

The URL for data presented herein is as follows:

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/

References

1. Pandolfo, M. (2002). The molecular basis of Friedreich ataxia. Adv. Exp. Med. Biol. *516*, 99–118.
2. Harding, A.E. (1981). Friedreich's ataxia: A clinical and genetic study of 90 families with an analysis of early diagnostic criteria and intrafamilial clustering of clinical features. Brain *104*, 589–620.
3. Pandolfo, M., and Koenig, M. (1998). Friedreich's ataxia. In Genetic Instabilities and Hereditary Neuromuscular Diseases, R.D. Wells and S.T. Warren, eds. (San Diego, CA: Academic Press), pp. 373–398.
4. Campuzano, V., Montermini, L., Molto, M.D., Pianese, L., Cossee, M., Cavalcanti, F., Monros, E., Rodius, F., Duclos, F., Monticelli, A., et al. (1996). Friedreich's ataxia: Autosomal recessive disease caused by an intronic GAA triplet repeat expansion. Science *271*, 1423–1427.
5. Clark, R.M., Dalgliesh, G.L., Endres, D., Gomez, M., Taylor, J., and Bidichandani, S.I. (2004). Expansion of GAA triplet repeats in the human genome: Unique origin of the FRDA mutation at the center of an Alu. Genomics *83*, 373–383.
6. Durr, A., Cossee, M., Agid, Y., Campuzano, V., Mignard, C., Penet, C., Mandel, J.L., Brice, A., and Koenig, M. (1996). Clinical and genetic abnormalities in patients with Friedreich's ataxia. N. Engl. J. Med. *335*, 1169–1175.
7. Filla, A., De Michele, G., Cavalcanti, F., Pianese, L., Monticelli, A., Campanella, G., and Cocozza, S. (1996). The relationship between trinucleotide (GAA) repeat length and clinical features in Friedreich ataxia. Am. J. Hum. Genet. *59*, 554–560.
8. Clark, R.M., Bhaskar, S.S., Miyahara, M., Dalgliesh, G.L., and Bidichandani, S.I. (2006). Expansion of GAA trinucleotide repeats in mammals. Genomics *87*, 57–67.
9. Campuzano, V., Montermini, L., Lutz, Y., Cova, L., Hindelang, C., Jiralerspong, S., Trottier, Y., Kish, S.J., Faucheux, B., Trouillas, P., et al. (1997). Frataxin is reduced in Friedreich ataxia patients and is associated with mitochondrial membranes. Hum. Mol. Genet. *6*, 1771–1780.
10. Gacy, A.M., Goellner, G.M., Spiro, C., Chen, X., Gupta, G., Bradbury, E.M., Dyer, R.B., Mikesell, M.J., Yao, J.Z., Johnson, A.J., et al. (1998). GAA instability in Friedreich's ataxia shares a common, DNA-directed and intraallelic mechanism with other trinucleotide diseases. Mol. Cell *1*, 583–593.
11. Bidichandani, S.I., Ashizawa, T., and Patel, P.I. (1998). The GAA triplet-repeat expansion in Friedreich ataxia interferes with transcription and may be associated with an unusual DNA structure. Am. J. Hum. Genet. *62*, 111–121.
12. Ohshima, K., Montermini, L., Wells, R.D., and Pandolfo, M. (1998). Inhibitory effects of expanded GAA.TTC triplet repeats from intron I of the Friedreich ataxia gene on transcription and replication in vivo. J. Biol. Chem. *273*, 14588–14595.
13. Sakamoto, N., Chastain, P.D., Parniewski, P., Ohshima, K., Pandolfo, M., Griffith, J.D., and Wells, R.D. (1999). Sticky DNA: Self-association properties of long GAA.TTC repeats in R.R.Y triplex structures from Friedreich's ataxia. Mol. Cell *3*, 465–475.
14. Sakamoto, N., Ohshima, K., Montermini, L., Pandolfo, M., and Wells, R.D. (2001). Sticky DNA, a self-associated complex formed at long GAA*TTC repeats in intron 1 of the frataxin gene, inhibits transcription. J. Biol. Chem. *276*, 27171–27177.
15. Potaman, V.N., Oussatcheva, E.A., Lyubchenko, Y.L., Shlyakhtenko, L.S., Bidichandani, S.I., Ashizawa, T., and Sinden, R.R. (2004). Length-dependent structure formation in Friedreich ataxia (GAA)n*(TTC)n repeats at neutral pH. Nucleic Acids Res. *32*, 1224–1231.
16. Grabczyk, E., and Usdin, K. (2000). The GAA*TTC triplet repeat expanded in Friedreich's ataxia impedes transcription elongation by T7 RNA polymerase in a length and supercoil dependent manner. Nucleic Acids Res. *28*, 2815–2822.
17. Krasilnikova, M.M., Kireeva, M.L., Petrovic, V., Knijnikova, N., Kashlev, M., and Mirkin, S.M. (2007). Effects of Friedreich's ataxia (GAA)n*(TTC)n repeats on RNA synthesis and stability. Nucleic Acids Res. *35*, 1075–1084.
18. Mifflin, R.C., and Kellems, R.E. (1991). Coupled transcription-polyadenylation in a cell-free system. J. Biol. Chem. *266*, 19593–19598.
19. Sisodia, S.S., Sollner-Webb, B., and Cleveland, D.W. (1987). Specificity of RNA maturation pathways: RNAs transcribed by RNA polymerase III are not substrates for splicing or polyadenylation. Mol. Cell. Biol. *7*, 3602–3612.
20. McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M., and Bentley, D.L. (1997). The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. Nature *385*, 357–361.
21. Maniatis, T., and Reed, R. (2002). An extensive network of coupling among gene expression machines. Nature *416*, 499–506.
22. Proudfoot, N.J., Furger, A., and Dye, M.J. (2002). Integrating mRNA processing with transcription. Cell *108*, 501–512.
23. Krasilnikova, M.M., and Mirkin, S.M. (2004). Replication stalling at Friedreich's ataxia (GAA)n repeats in vivo. Mol. Cell. Biol. *24*, 2286–2295.
24. Pollard, L.M., Sharma, R., Gomez, M., Shah, S., Delatycki, M.B., Pianese, L., Monticelli, A., Keats, B.J., and Bidichandani, S.I. (2004). Replication-mediated instability of the GAA triplet repeat mutation in Friedreich ataxia. Nucleic Acids Res. *32*, 5962–5971.
25. Saveliev, A., Everett, C., Sharpe, T., Webster, Z., and Festenstein, R. (2003). DNA triplet repeats mediate heterochromatin-protein-1-sensitive variegated gene silencing. Nature *422*, 909–913.

26. Herman, D., Jenssen, K., Burnett, R., Soragni, E., Perlman, S.L., and Gottesfeld, J.M. (2006). Histone deacetylase inhibitors reverse gene silencing in Friedreich's ataxia. Nat. Chem. Biol. *2*, 551–558.
27. Greene, E., Mahishi, L., Entezam, A., Kumari, D., and Usdin, K. (2007). Repeat-induced epigenetic changes in intron 1 of the frataxin gene and its consequences in Friedreich ataxia. Nucleic Acids Res. *35*, 3383–3390.
28. Cartegni, L., Chew, S.L., and Krainer, A.R. (2002). Listening to silence and understanding nonsense: Exonic mutations that affect splicing. Nat. Rev. Genet. *3*, 285–298.
29. Fairbrother, W.G., Yeh, R.F., Sharp, P.A., and Burge, C.B. (2002). Predictive identification of exonic splicing enhancers in human genes. Science *297*, 1007–1013.
30. Pagani, F., and Baralle, F.E. (2004). Genomic variants in exons and introns: Identifying the splicing spoilers. Nat. Rev. Genet. *5*, 389–396.
31. Black, D.L. (2003). Mechanisms of alternative pre-messenger RNA splicing. Annu. Rev. Biochem. *72*, 291–336.
32. Faustino, N.A., and Cooper, T.A. (2003). Pre-mRNA splicing and human disease. Genes Dev. *17*, 419–437.
33. Buratti, E., Brindisi, A., Pagani, F., and Baralle, F.E. (2004). Nuclear factor TDP-43 binds to the polymorphic TG repeats in CFTR intron 8 and causes skipping of exon 9: A functional link with disease penetrance. Am. J. Hum. Genet. *74*, 1322–1325.
34. Hui, J., Stangl, K., Lane, W.S., and Bindereif, A. (2003). HnRNP L stimulates splicing of the eNOS gene by binding to variable-length CA repeats. Nat. Struct. Biol. *10*, 33–37.
35. Ule, J., Stefani, G., Mele, A., Ruggiu, M., Wang, X., Taneri, B., Gaasterland, T., Blencowe, B.J., and Darnell, R.B. (2006). An RNA map predicting Nova-dependent splicing regulation. Nature *444*, 580–586.
36. Muro, A.F., Caputi, M., Pariyarath, R., Pagani, F., Buratti, E., and Baralle, F.E. (1999). Regulation of fibronectin EDA exon alternative splicing: Possible role of RNA secondary structure for enhancer display. Mol. Cell. Biol. *19*, 2657–2671.
37. Goina, E., Skoko, N., and Pagani, F. (2008). Binding of DAZAP1 and hnRNPA1/A2 to an exonic splicing silencer in a natural BRCA1 exon 18 mutant. Mol. Cell. Biol. *28*, 3850–3860.
38. Pagani, F., Buratti, E., Stuani, C., and Baralle, F.E. (2003). Missense, nonsense, and neutral mutations define juxtaposed regulatory elements of splicing in cystic fibrosis transmembrane regulator exon 9. J. Biol. Chem. *278*, 26580–26588.
39. Al-Mahdawi, S., Pinto, R.M., Varshney, D., Lawrence, L., Lowrie, M.B., Hughes, S., Webster, Z., Blake, J., Cooper, J.M., King, R., et al. (2006). GAA repeat expansion mutation mouse models of Friedreich ataxia exhibit oxidative stress leading to progressive neuronal and cardiac pathology. Genomics *88*, 580–590.
40. Pagani, F., Stuani, C., Tzetis, M., Kanavakis, E., Efthymiadou, A., Doudounakis, S., Casals, T., and Baralle, F.E. (2003). New type of disease causing mutations: The example of the composite exonic regulatory elements of splicing in CFTR exon 12. Hum. Mol. Genet. *12*, 1111–1120.
41. Caputi, M., Casari, G., Guenzi, S., Tagliabue, R., Sidoli, A., Melo, C.A., and Baralle, F.E. (1994). A novel bipartite splicing enhancer modulates the differential processing of the human fibronectin EDA exon. Nucleic Acids Res. *22*, 1018–1022.
42. Cramer, P., Caceres, J.F., Cazalla, D., Kadener, S., Muro, A.F., Baralle, F.E., and Kornblihtt, A.R. (1999). Coupling of transcription with alternative splicing: RNA pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer. Mol. Cell *4*, 251–258.
43. de la Mata, M., Alonso, C.R., Kadener, S., Fededa, J.P., Blaustein, M., Pelisch, F., Cramer, P., Bentley, D., and Kornblihtt, A.R. (2003). A slow RNA polymerase II affects alternative splicing in vivo. Mol. Cell *12*, 525–532.
44. de la Mata, M., and Kornblihtt, A.R. (2006). RNA polymerase II C-terminal domain mediates regulation of alternative splicing by SRp20. Nat. Struct. Mol. Biol. *13*, 973–980.
45. Son, L.S., Bacolla, A., and Wells, R.D. (2006). Sticky DNA: In vivo formation in E. coli and in vitro association of long GAA*TTC tracts to generate two independent supercoiled domains. J. Mol. Biol. *360*, 267–284.
46. Calado, A., Kutay, U., Kuhn, U., Wahle, E., and Carmo-Fonseca, M. (2000). Deciphering the cellular pathway for transport of poly(A)-binding protein II. RNA *6*, 245–256.
47. Davies, J.E., Sarkar, S., and Rubinsztein, D.C. (2008). Wild-type PABPN1 is anti-apoptotic and reduces toxicity of the oculopharyngeal muscular dystrophy mutation. Hum. Mol. Genet. *17*, 1097–1108.
48. Pellizzoni, L., Lotti, F., Maras, B., and Pierandrei-Amaldi, P. (1997). Cellular nucleic acid binding protein binds a conserved region of the 5′ UTR of Xenopus laevis ribosomal protein mRNAs. J. Mol. Biol. *267*, 264–275.
49. Spellman, R., and Smith, C.W. (2006). Novel modes of splicing repression by PTB. Trends Biochem. Sci. *31*, 73–76.
50. Tacke, R., Tohyama, M., Ogawa, S., and Manley, J.L. (1998). Human Tra2 proteins are sequence-specific activators of pre-mRNA splicing. Cell *93*, 139–148.
51. Lavigueur, A., La Branche, H., Kornblihtt, A.R., and Chabot, B. (1993). A splicing enhancer in the human fibronectin alternate ED1 exon interacts with SR proteins and stimulates U2 snRNP binding. Genes Dev. *7*, 2405–2417.
52. Taniguchi, I., Masuyama, K., and Ohno, M. (2007). Role of purine-rich exonic splicing enhancers in nuclear retention of pre-mRNAs. Proc. Natl. Acad. Sci. USA *104*, 13684–13689.
53. Masuyama, K., Taniguchi, I., Okawa, K., and Ohno, M. (2007). Factors associated with a purine-rich exonic splicing enhancer sequence in Xenopus oocyte nucleus. Biochem. Biophys. Res. Commun. *359*, 580–585.
54. Masuyama, K., Taniguchi, I., Kataoka, N., and Ohno, M. (2004). SR proteins preferentially associate with mRNAs in the nucleus and facilitate their export to the cytoplasm. Genes Cells *9*, 959–965.
55. Miranda, C.J., Santos, M.M., Ohshima, K., Smith, J., Li, L., Bunting, M., Cossee, M., Koenig, M., Sequeiros, J., Kaplan, J., et al. (2002). Frataxin knockin mouse. FEBS Lett. *512*, 291–297.
56. Chang, J.G., Hsieh-Li, H.M., Jong, Y.J., Wang, N.M., Tsai, C.H., and Li, H. (2001). Treatment of spinal muscular atrophy by sodium butyrate. Proc. Natl. Acad. Sci. USA *98*, 9808–9813.
57. Brichta, L., Hofmann, Y., Hahnen, E., Siebzehnrubl, F.A., Raschke, H., Blumcke, I., Eyupoglu, I.Y., and Wirth, B. (2003). Valproic acid increases the SMN2 protein level: A well-known drug as a potential therapy for spinal muscular atrophy. Hum. Mol. Genet. *12*, 2481–2489.
58. Riessland, M., Brichta, L., Hahnen, E., and Wirth, B. (2006). The benzamide M344, a novel histone deacetylase inhibitor, significantly increases SMN2 RNA/protein levels in spinal muscular atrophy cells. Hum. Genet. *120*, 101–110.

59. Nissim-Rafinia, M., Aviram, M., Randell, S.H., Shushi, L., Ozeri, E., Chiba-Falek, O., Eidelman, O., Pollard, H.B., Yankaskas, J.R., and Kerem, B. (2004). Restoration of the cystic fibrosis transmembrane conductance regulator function by splicing modulation. EMBO Rep. *5*, 1071–1077.
60. Buhler, M., and Moazed, D. (2007). Transcription and RNAi in heterochromatic gene silencing. Nat. Struct. Mol. Biol. *14*, 1041–1048.
61. Grewal, S.I., and Elgin, S.C. (2007). Transcription and RNA interference in the formation of heterochromatin. Nature *447*, 399–406.
62. Sims, R.J. 3rd, Millhouse, S., Chen, C.F., Lewis, B.A., Erdjument-Bromage, H., Tempst, P., Manley, J.L., and Reinberg, D. (2007). Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. Mol. Cell *28*, 665–676.