



University of Southern Denmark

## Antisense oligonucleotide rescue of deep-intronic variants activating pseudoexons in the 6-pyruvoyl-tetrahydropterin synthase gene

Martínez-Pizarro, Ainhoa; Leal, Fátima; Holm, Lise Lolle; Doktor, Thomas K; Petersen, Ulrika S S; Bueno, María; Thöny, Beat; Pérez, Belén; Andresen, Brage S; Desviat, Lourdes R

*Published in:*  
Nucleic Acid Therapeutics

*DOI:*  
10.1089/nat.2021.0066

*Publication date:*  
2022

*Document version:*  
Final published version

*Document license:*  
CC BY

### *Citation for pulished version (APA):*

Martínez-Pizarro, A., Leal, F., Holm, L. L., Doktor, T. K., Petersen, U. S. S., Bueno, M., Thöny, B., Pérez, B., Andresen, B. S., & Desviat, L. R. (2022). Antisense oligonucleotide rescue of deep-intronic variants activating pseudoexons in the 6-pyruvoyl-tetrahydropterin synthase gene. *Nucleic Acid Therapeutics*, 32(5), 378-390. <https://doi.org/10.1089/nat.2021.0066>

Go to publication entry in University of Southern Denmark's Research Portal

### **Terms of use**

This work is brought to you by the University of Southern Denmark.  
Unless otherwise specified it has been shared according to the terms for self-archiving.  
If no other license is stated, these terms apply:

- You may download this work for personal use only.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim.  
Please direct all enquiries to [puresupport@bib.sdu.dk](mailto:puresupport@bib.sdu.dk)

Open camera or QR reader and  
scan code to access this article  
and other resources online.



# Antisense Oligonucleotide Rescue of Deep-Intronic Variants Activating Pseudoexons in the 6-Pyruvoyl-Tetrahydropterin Synthase Gene

Ainhoa Martínez-Pizarro,<sup>1</sup> Fátima Leal,<sup>1</sup> Lise Lolle Holm,<sup>2</sup> Thomas K. Doktor,<sup>2</sup> Ulrika S.S. Petersen,<sup>2</sup> María Bueno,<sup>3</sup> Beat Thöny,<sup>4</sup> Belén Pérez,<sup>1</sup> Brage S. Andresen,<sup>2</sup> and Lourdes R. Desviat<sup>1</sup>

We report two new 6-pyruvoyl-tetrahydropterin synthase splicing variants identified through genomic sequencing and transcript analysis in a patient with tetrahydrobiopterin deficiency, presenting with hyperphenylalaninemia and monoamine neurotransmitter deficiency. Variant c.243 + 3A>G causes exon 4 skipping. The deep-intronic c.164–672C>T variant creates a potential 5' splice site that leads to the inclusion of four overlapping pseudoexons, corresponding to exonizations of an antisense short interspersed nuclear element *AluS<sub>q</sub>* repeat sequence. Two of the identified pseudoexons have been reported previously, activated by different deep-intronic variants, and were also detected at residual levels in control cells. Interestingly, the predominant pseudoexon is nearly identical to a disease causing activated pseudoexon in the *F8* gene, with the same 3' and 5' splice sites. Splice switching antisense oligonucleotides (SSOs) were designed to hybridize with splice sites and/or predicted binding sites for regulatory splice factors. Different SSOs corrected the aberrant pseudoexon inclusion, both in minigenes and in fibroblasts from patients carrying the new variant c.164–672C>T or the previously described c.164–716A>T. With SSO treatment PTPS protein was recovered, illustrating the therapeutic potential of the approach, for patients with different pseudoexon activating variants in the region. In addition, the natural presence of pseudoexons in the wild type context suggests the possibility of applying the antisense strategy in patients with hypomorphic *PTS* variants with the purpose of upregulating their expression to increase overall protein and activity.

**Keywords:** pseudoexons, splicing, antisense oligonucleotides, tetrahydrobiopterin

## Introduction

**T**HE PTPS (6-PYRUVOYL-TETRAHYDROPTERIN synthase; EC 4.2.3.12) enzyme, encoded by the *PTS* gene (OMIM \*612719), catalyzes the second step in tetrahydrobiopterin (BH<sub>4</sub>) synthesis from GTP in humans. BH<sub>4</sub> is the natural cofactor for phenylalanine, tyrosine and tryptophan hydroxylases, nitric oxide synthases, and glycerol ether monooxygenase [1]. PTPS deficient patients present with hyperphenylalanine-

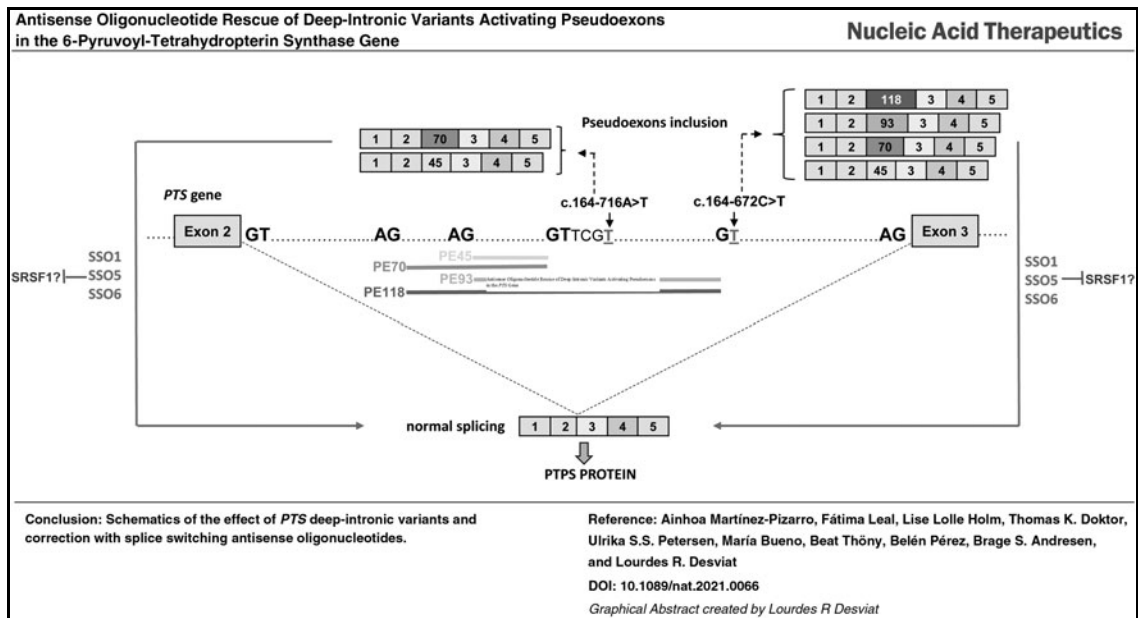
mia, and most of them present severe neurological manifestations, including truncal hypotonia, bradykinesia, oculogyric crises, hyperthermia, seizures, and impaired neurophysiological development. A subset of patients exhibits the less common atypical/peripheral form with minor or no changes in neurotransmitter levels. Treatment is based on oral BH<sub>4</sub> administration and supplementation with L-Dopa and 5-hydroxytryptophan, the precursors for catecholamines and serotonin, respectively, to normalize

<sup>1</sup>Centro de Biología Molecular Severo Ochoa UAM-CSIC, Centro de Diagnóstico de Enfermedades Moleculares (CEDEM), CIBERER, IdiPaz, Universidad Autónoma de Madrid, Madrid, Spain.

<sup>2</sup>Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Denmark.

<sup>3</sup>Congenital Metabolic Diseases Unit, Hospital Virgen del Rocío, Sevilla, Spain.

<sup>4</sup>Division of Metabolism, University Children's Hospital Zürich, Zürich, Switzerland.



monoamine neurotransmitters in the central nervous system [2]. Despite treatment, many patients have complications with neurological and developmental problems [3].

In this work we describe a novel deep intronic variant affecting splicing in the *PTS* gene. It is well known that splice defects significantly contribute to human disease [4,5]. Splice altering variants occur beyond disruption of the canonical splice sequences at exon-intron junctions [consensus 3' and 5' splice sites (ss), polypyrimidine tract, and branch point], which are easily identified. Intronic and exonic splice silencers and enhancers (ISSs, ESSs, ISEs, and ESEs, respectively), bound by regulatory splice factors, are less well conserved, and sequence variants that disrupt them can be misclassified as silent or missense, unless transcript analysis is performed, which is not usual in a diagnostic setting [6,7]. In addition, aberrant transcripts can result from deep intronic variants activating or creating potential splice sites, which are used in combination with other cryptic splice sites nearby resulting in the exonization of intronic regions known as pseudoexons [8,9].

Pseudoexons share characteristics marking them as high-risk sites in the human genome, where single nucleotide variants (SNVs) may cause disease by splicing activation [10,11]. Again, these splice defects and the responsible variants are not detected in routine genetic analysis focusing in exome sequencing. Recent studies have revealed the potential of high-throughput transcriptome analysis for increasing the diagnostic rate in heterogeneous Mendelian disorders, allowing identification of splice-altering variants in both exonic and deep-intronic regions [12–16].

RNA targeted therapies such as splice switching antisense oligonucleotides (SSOs) provide an avenue toward therapy for specific splicing defects [17]. SSOs are designed to modulate splicing by sterically blocking splice sites or splicing regulatory sequences in the pre-mRNA. Blocking of consensus splice sites or splice enhancers will result in exon skipping, a strategy that has been used to bypass a mutated exon and/or to restore the open reading frame disrupted through genomic deletions, as is the case for Duchenne muscular dystrophy (DMD) [17]. In contrast, blocking splice

silencers may instead promote exon inclusion, which is the therapeutic strategy used in spinal muscular atrophy (SMA) [18]. Therapeutic SSOs for DMD and SMA have been approved by the Food and Drug Administration (FDA)/European Medicines Agency (EMA), and several others are in clinical trials [19,20].

SSOs have also been used to prevent aberrant pseudoexon inclusion in the mRNA. Many pseudoexons are transposable sequence elements derived from short or long interspersed nuclear elements (SINEs or LINEs, respectively) [21], and their activation by point mutations has been described as pathogenic mechanism in many human disorders [8,22–24]. In neuromuscular diseases, deep intronic variants have been identified in 77 different genes [14]. In inherited retinal diseases, therapeutic potential of SSOs have been shown for deep-intronic pseudoexon activating variants in the *CEP290*, *USH2A*, *OPAI1*, and *ABCA4* genes, with an ongoing phase 1/2 clinical trial using SSOs to treat *CEP290*-associated Leber congenital amaurosis (NCT03140969) [23]. Recently, an SSO (Milasen) targeting an activated pseudoexon identified in a single patient with Batten's disease was developed and approved by FDA for clinical testing within 1 year, representing an extreme example of the use of SSOs as personalized medicines [25].

Previous studies by us and other groups reported the pathological activation by two different variants of overlapping pseudoexons in intron 2 of the *PTS* gene, which could be effectively reverted by use of a specific SSO [26]. We now report a novel deep intronic variant in the same region creating a 5' splice site and activating the same pseudoexons plus two others that use the new splice site. The patient carries on the other allele an also novel variant affecting the 5' splice site of exon 4. We have used transcript profiling in minigenes and in patients' fibroblasts to dissect the pathogenic mechanism for each variant and designed and tested a series of SSOs to block the insertion of the activated pseudoexons in the mature mRNA. Our results confirm SSO mediated pseudoexon skipping as a potential therapeutic approach for *PTPS* deficiency and point to the identification of high risk sites for the activation of *Alu*-derived pseudoexons.

## Materials and Methods

### Patient samples and genetic and biochemical analysis

Patient 1 was referred to the laboratory for biochemical and genetic analysis after giving a positive result in the neonatal screening for hyperphenylalaninemia. Biochemical analysis (pterin levels in blood and urine, and monoamine neurotransmitter levels in cerebrospinal fluid (CSF), measured by standard procedures [27]) indicated a clinical suspicion of *PTS* deficiency. Genetic analysis was performed in DNA isolated from blood samples, by use of Massive Parallel Sequencing using the Clinical-Exome Sequencing TruSight™ One Gene Panel (Illumina) and filtering for variants in all genes known to be responsible for hyperphenylalaninemia (*GCHI*, *PTS*, *PCBD1*, *QDPR*, *DNAJC12*, and *PAH*). *PTS* transcript analysis was performed in RNA extracted from fibroblast samples. Potentially pathogenic variants were confirmed by Sanger sequencing. Patient 2 has been previously reported [26]. Ethical approval for the present study was granted by the Institutional Ethics Committee (Universidad Autónoma de Madrid).

### Fibroblasts and cell lines

Patients' fibroblasts were referred to the laboratory for genetic analysis. Control fibroblasts CC2509 were obtained from Lonza. They were cultivated according to standard procedures and used with number of passes below 20. Briefly, cells were maintained in Minimum Essential Medium (MEM; Sigma Aldrich) supplemented with 1% (v/v) glutamine, 10% fetal bovine serum (FBS; Gibco), and 0.1% antibiotic mix (penicillin/streptomycin) under standard cell culture conditions [37°C, 95% relative humidity, 5% carbon dioxide (CO<sub>2</sub>)]. Cells were treated with puromycin (200 µg/mL) 5 h before harvesting for transcript analysis, to inhibit nonsense mediated mRNA decay (NMD).

For minigene analysis, human hepatoma cells Hep3B and SH-SY5Y neuroblastoma cells were used. Cells were grown in MEM for Hep3B and Dulbecco's modified Eagle's medium (DMEM; Gibco) for SH-SY5Y, both supplemented with 5% FBS, 1% glutamine, and 0.1% antibiotic mix under standard cell culture conditions (37°C, 95% relative humidity, 5% CO<sub>2</sub>).

### Minigenes

For *in vitro* evaluation of the splicing profile different minigene constructs were used. The previously described splicing reporter plasmid encompassing *PTS* exons 2–4 cloned in the pCR3.1 vector [26,28] was used as template for deep intron 2 variants c.164-672C>T and c.164-716A>T. The pSPL3 vector (Life Technologies) was used in the analysis of the c.243+3A>G variant. Using this vector, we cloned *PTS* exon 4 and ~150 bp of the flanking intronic regions, amplified from genomic DNA using specific primers (Forward: 5'-GCTTCCATGCTGAGGTC AAT-3' and Reverse: 5'-ACTATTCCCCAACCCACA-3'). Gene fragments were cloned into the pGEMT vector (No. A1360; Promega). The insert was excised with *EcoRI* and subsequently cloned into pSPL3. Variant minigenes containing the desired nucleotide changes were generated by site-directed mutagenesis with QuikChange Lightning Kit (Agilent Technologies, Santa Clara, CA, USA) using primers introducing the change and their reverse complement. Sanger sequencing confirmed the identity of the constructs.

### SSO design

The sequence of the region encompassing the pseudoexons was analyzed and SSOs designed following published guidelines [29,30]. Splice enhancer motifs were determined using ESE finder3.0. Pre-mRNA secondary structure was predicted using m-fold (<http://www.unafold.org/mfold/applications/rna-folding-form.php>), and RNAstructure software (<https://rna.urmc.rochester.edu/RNAstructureWeb/>) was used to predict secondary structure resulting from the interaction antisense oligonucleotide (AON)-target RNA. Specificity of the selected SSOs was verified using BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). SSOs were 20- or 25-mers with a T<sub>m</sub> above 60°C and a GC content between 50% and 60%. The exact sequences (Supplementary Table S1) and their location are depicted in Fig. 3A. A scrambled SSO was used as negative control. Given the abundance of *Alu* elements in the genome, we evaluated closely the predicted off-target binding sites, confirming that they are far from adjacent exons. All oligonucleotides were synthesized as full 2'-OMe phosphorothioates by Sigma or Eurogentec.

### Transfections and splicing analysis in minigenes

For minigene assays, Hep3B or SH-SY5Y cells were seeded in six-well plates at a density of  $4 \times 10^5$  or  $8 \times 10^5$ , respectively, in 2 mL MEM or DMEM 5% FBS and grown overnight. Cells were transfected using Lipofectamine 2000 (Invitrogen) with 2 µg per well of the corresponding plasmid minigene. For SSO treatment, cells were cotransfected with 2 µg of wild type or mutant minigenes and 50 nM of each SSO. Cells were harvested by trypsinization 48 h after transfection. Total RNA was isolated using TRIzol Reagent (ThermoFisher) and phenol-chloroform extraction. cDNA synthesis was performed using NZY First-Strand cDNA Synthesis Kit (NZYtech).

Splicing analysis was carried out by polymerase chain reaction (PCR) amplification with FastStart Taq DNA Polymerase (Roche) using specific primers SD6 (5'- TCTGAGT CACCTGGACAACC-3') and SA2 (5'- ATCTCAGTGGTATT TGTGAGC-3') for pSPL3 minigenes and PL3 (5'-GGGAGA CCAAGCTGGCTA-3') and PL4b (5'-AGTCGAGGCTGA TCAGCGG-3') for the pCR3.1 minigene. The amplification products were analyzed by 2% agarose gel electrophoresis. Pseudoexon bands were excised from the gel and purified using QIAquick Gel Extraction (Qiagen). The identity of each pseudoexon was confirmed by subcloning in pGEMT (Promega) and subsequent Sanger sequencing.

### Transfections and splicing analysis in fibroblasts

For SSO treatment, fibroblasts were seeded in six-well plates at a density of  $3 \times 10^5$  in 2 mL MEM 10% FBS and reverse transfected at the same time with 50 nM SSOs using Lipofectamine 2000. Cells were harvested by trypsinization after 48 h. RNA was isolated using TRIzol Reagent (Thermo Fisher) and phenol-chloroform extraction. Reverse transcription PCR (RT-PCR) analysis was performed using primers *PTS* cDNA F (5'-ATGAGCACGGAAGGTGGTG-3') and Exon 3-R (5'-CTCTCCATGTACTGTACCACAA-3') to amplify a fragment from exon 1 to exon 3. The amplification products were analyzed by 2.5% agarose gel electrophoresis and their identity and quantification confirmed by deep sequencing.

### Western blot analysis

Approximately  $4 \times 10^5$  fibroblasts were seeded per well in a six-well plate with 50 nM of each SSO, which were reverse transfected using Lipofectamine 2000. Ninety-six hours after transfection cells were harvested by trypsinization and lysed by freeze-thawing in a lysis buffer (10 mM Tris-HCl pH 7.5, 10% glycerol, 150 mM sodium chloride, and 0.1% Triton X-100) with protease inhibitor (Roche). Protein concentration was determined using Bradford reagent (Bio-Rad).

Fifteen micrograms of protein extract from control sample and 75 µg from patients' samples were electrophoresed on 12% SDS-polyacrylamide gels, transferred to a nitrocellulose membrane in an iBlot transfer device (Invitrogen) and blocked in Tris buffered saline-0.1% Tween supplemented with 5% nonfat dry milk. Membranes were incubated overnight at 4°C with the primary antibody (PTPS 1:500, sc-514628; Santa Cruz Biotechnologies) followed by a secondary horseradish peroxidase-conjugated anti-mouse IgG (1:2000, sc-514628; Cell Signaling). GAPDH was used as a loading control (using primary antibody at ab8245, 1:5000; Abcam). Proteins were detected using the Enhanced Chemiluminescence reagent (GE Healthcare). Images were scanned with a densitometer GS-900 (Bio-Rad) and quantified with Image Lab 5.2 (Bio-Rad).

### Deep sequencing

Amplicons were generated in a first PCR with fusion primers containing a specific sequence plus a common tag (CS1 and CS2):

- Forward primer: 5'-**ACACTGACGACATGGTTCTA CACTGTTTGGGAAATGCAACAA**-3'
- Reverse primer: 5'-**TACGGTAGCAGAGACTTGGT CTTCTCCATGTACTGTACCACAA**-3'

Libraries preparation and sequencing were carried out at Fundación Parque Científico de Madrid under protocols developed and optimized for next generation amplicon sequencing. Briefly, amplicons were diluted 1/50, and 1 µl was used as input for a second PCR of 10 cycles performed with a High-Fidelity DNA Polymerase in the presence of primers [5'-AATGATACGGCGACCACCGAGATCTACACTGACGACATGGTTCTACA-3' and 5'-CAAGCAGAAGACGGCA TACGAGAT-(10 nucleotides barcode)-TACGGTAGCAGAGACTTGGTCT-3'] of the Access Array Barcode Library for Illumina Sequencers (Fluidigm). The finally obtained amplicons were validated and quantified by Bioanalyzer, and equimolecular pool was purified by AMPure XP beads (Beckman Coulter). This pool was titrated by quantitative PCR using the "Kapa-SYBR FAST qPCR Kit for Light-Cycler480" and a reference standard for quantification. Amplicon pool was denatured before being seeded on a flow cell where clusters were formed and sequenced using a "MiSeq Reagent Kit v3," in a 2×300 pair-end sequencing run on a MiSeq sequencer.

### RNA pull down

Affinity purification of RNA-binding proteins was performed as previously described [31], using 3' end biotinylated RNA oligonucleotides (LGC Biosearch Technologies, Risskov, Denmark): PTS1 (5'-CCAGGUUCGAGCGAUUC UCCAGCCU-3'), PTS1MUT1 (5'-CCAGGUUCGUGCGA UUCUCCAGCCU-3'), PTS1MUT2 (5'-CCAGGUUCUAG

CUAUUCUCCAGCCU-3'), PTS1Rev (5'-AGGCUGGAG AAUCGCUCGAACCUGG-3'), PTS5 (5'-UUCUCCAGCC UCAGCCUCCGAGUA-3'), PTS5MUT (5'-UUCUCCAG CCUCAGCCUCUUGAGUA-3'), PTS5Rev (5'-UACUCGG GAGGCUGAGGCUGGAGAA-3'), PTS6 (5'-UCCCGAGU AGCUGAGAUUACAGGUG-3'), PTS6MUT (5'-UCUUG UGUAGCUUAUUAUUCUGGUG-3'), and PTS6Rev (5'-C ACCUGUAAUCUACAGCUACUCGGGA-3').

Five hundred picomole of each biotinylated RNA oligonucleotide was immobilized in 25 µL Dynabeads M-280 Strep-tavidin magnetic beads (Invitrogen) and incubated with 293 cell nuclear extract (Protein One). Blank beads were added as a control for background binding. Proteins were eluted in 25 µL protein sample buffer and analyzed by western blotting with antibodies against SRSF1 (32–4500 from Invitrogen), hnRNP A1 (R9778 from Sigma Aldrich), SRSF2 (04-1550 from Millipore), and SR proteins (33–9300 from Invitrogen).

### Surface plasmon resonance imaging

Surface plasmon resonance imaging (SPRi) was carried out as previously described [32]. Briefly, the biotinylated RNA-oligonucleotides described above were immobilized onto a G-strep sensorchip (SSENS) for 20 min. The following recombinant proteins were injected for 8 min, followed by dissociation for 4 min; 25–800 nM SRSF1 (GenScript) and 6.25–200 nM hnRNPA1 (ab123212; Abcam) with a continuous flow of SPR buffer (10 mM Tris-HCl, pH 7.9, 150 mM potassium chloride, 3.4 mM ethylenediaminetetraacetic acid, 0.005% Tween 20). Nuclear extract was used as a control of the oligos binding efficiency. Binding was fitted to a 1:1 kinetics model with Scrubber2 (v. 2.1.1.0; Biologics Inc.). For hnRNPA1 a biphasic 1:2 model was used in ClampXP (version 3.50; Biosensor Data Analysis).

### In silico analysis

The effect of the variants on splice site strengths and the presence of putative splicing regulatory elements were predicted using Human Splice Finder (HSF) (<http://www.umd.be/HSF/>) [33], MaxEntScan ([http://genes.mit.edu/burgelab/masent/Xmaxentscan\\_scoreseq.html](http://genes.mit.edu/burgelab/masent/Xmaxentscan_scoreseq.html)) [34], ESEFinder 3.0 software (<http://rulai.cshl.edu>) [35], and RBPmap (<http://rbpmap.technion.ac.il/>) [36].

The presence of repetitive sequences was analyzed using RepeatMasker (<http://www.repeatmasker.org/>).

### Processing of cross-linking immunoprecipitation datasets

Raw SRSF1 eCLIP reads [37] were downloaded from the ENCODE website and trimmed twice using cutadapt v1.15 [38]. Raw PAR-CLIP reads were downloaded from the GEO repository using the accession number GSE71095 [39] and trimmed using cutadapt. Raw HITS-CLIP reads were downloaded from the GEO repository using the accession numbers GSE131745 [40] and trimmed using cutadapt. Subsequently, reads were mapped with STAR v2.7.8a [41] to the human genome using GENCODE v39 hg38 annotations allowing for up to 100 mapping locations. Run parameters and further details are described in Supplementary Methods S1.

Visualization of CLIP reads

We used Gviz [42] to visualize the alignments along with GenomicFeatures [43] to parse the GENCODE GTF file with transcripts.

Results

Clinical, biochemical, and genetic analysis

Patient 1 was detected in newborn screening due to high Phe levels (226 μM) and later diagnosed as the severe form of *PTPS* deficiency based on urinary pterins (elevated neopterin, 22.9 mmol/mol creatinine and low biopterin, 0.3 mmol/mol creatinine) and monoamine neurotransmitter deficiency in CSF, corresponding to very low levels of homovanillic acid and of 5-hydroxyindoleacetic acid, (140 nM and 75 nM, respectively; for reference values see [44]). Prolactin, which is physiologically inhibited by dopamine, showed very high values (655–1590 μU/mL; reference values 127–637 μU/mL). Treatment with sapropterin (BH<sub>4</sub>), 5-OH Trp, and L-Dopa/carbidopa, adjusted to body weight, was implemented. Currently, the patient is 6 years old and developing normally, with no overt clinical symptoms.

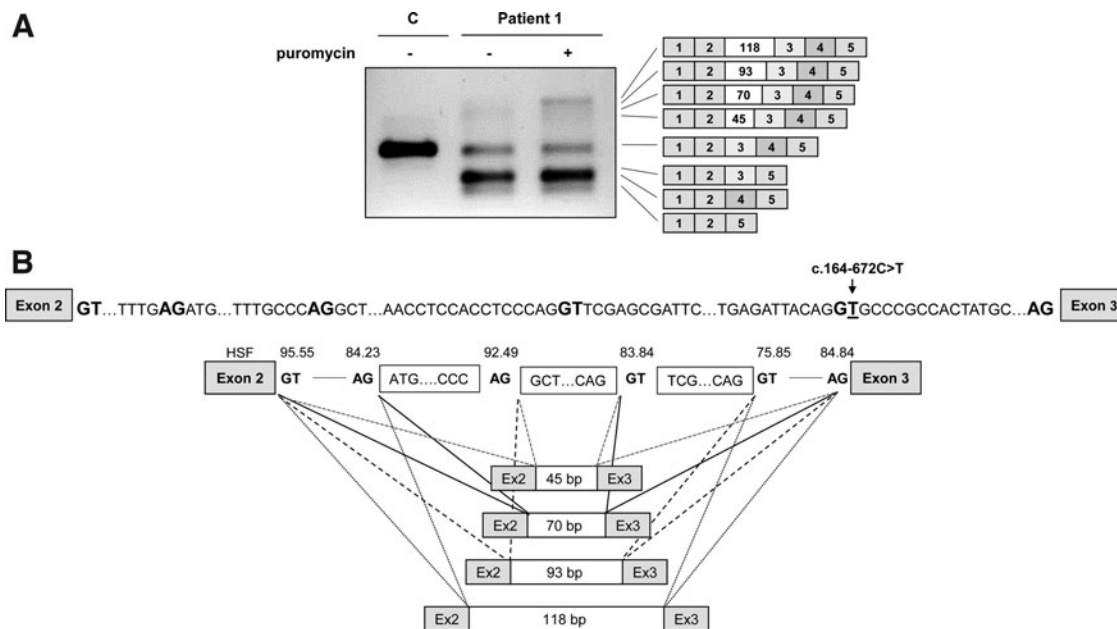
Genetic analysis by exome sequencing identified a single variant in the *PTS* gene: c.243+3A>G, affecting the 5'ss of exon 4 (decrease in HSF splicing score from 97.71 to 81.74 and from 8.7 to 4.4 in MaxEnt score). *PTPS* deficiency is inherited in autosomal recessive manner, so we sought to find the second pathogenic variant, which might reside within regions not covered by routine exome sequencing. We used Sanger sequencing to analyze a candidate region in intron 2 in which two different pathogenic variants were previously shown to activate a series of overlapping pseudoexons

[26,28]. This analysis revealed the presence of a novel variant c.164-672C>T that creates a potential 5'ss (HSF splicing score 75.85 and MaxEnt 5.8), downstream of the previously identified pseudoexons.

RT-PCR analysis in patient's fibroblasts revealed several aberrant bands of higher and lower molecular weight compared to the correctly spliced band, which was detected in low amounts (Fig. 1A). The lower sized bands were prominent in samples treated or not with puromycin, used as NMD inhibitor. After subcloning and sequencing we identified transcripts with exon 3 skipping [out of frame, p.(Val155Aspfs\*2), occurs naturally due to poor exon definition] [45], exon 4 skipping [in frame, p.(Ile63\_Glu81del)], and skipping of both exons 3 and 4 p.(Val55Glyfs\*8).

In puromycin treated samples faint high molecular weight bands were observed, which were identified by subcloning and sequencing analysis as insertion of several intronic sequences (pseudoexons) of 45, 70, 93, and 118 bp between exons 2 and 3. Pseudoexons of 45 and 70 bp (PE45 and PE70) have been previously reported to be activated by different variants in the region [26] and those of 93 and 118 bp (PE93 and PE118) correspond to the use of the newly created 5'ss by variant c.164-672C>T, in combination with upstream cryptic 3'ss (Fig. 1B). All correspond to exonizations of an antisense *AluS<sub>q</sub>* sequence, as previously described [28]. The newly created 5'ss corresponds to *Alu* position 156, present in naturally exonized *Alu*s [46,47].

Interestingly, by altering the conserved C at position 156 in the *Alu*y consensus sequence the *PTS* c.164-672C>T variant alters this to a 5'ss, identical to that used in the disease associated *F8* pseudoexon [24,48]. The *PTS* PE118 is nearly identical to the *F8* pseudoexon with the same 3'ss and 5'ss, consistent with the fact that this seems to be the most prominent



**FIG. 1.** *PTS* gene splicing profile in patient 1 fibroblasts. **(A)** RT-PCR analysis in C and patient 1 fibroblasts treated with (+) or without (-) puromycin. Shown on the *right* are the schematics of the identified bands confirmed by sequencing. **(B)** Sequence of the intron 2 region where the pseudoexons are located, showing the conserved GT and AG splice sites (in bold) used by the pseudoexons and the location of the novel variant. Shown below is the schematics of the identified pseudoexons along with the score of the splice sites according to HSF. C, control; HSF, Human Splice Finder; RT-PCR, reverse transcription polymerase chain reaction.

*PTS* PE isoform. Insertion of *PTS* PE45 and PE93 results in extra amino acid insertions (p.(Val55\_Val56ins15) and p.(Val55\_Val56ins31), respectively), while insertion of PE70 and PE118 results in frameshift and introduction of premature stop codons [p.(Val55Serfs\*33) and p.(Val55Aspfs\*37), respectively].

#### Functional analysis in minigenes of the novel variants

The individual effect on splicing of the novel variants was studied using minigenes transfected in Hep3B cells. Variant c.164-672C>T was analyzed in the previously described pCR3.1 minigene encompassing *PTS* exons 2–4 [26]. The mutant minigene produces, in addition to a normal sized, correctly spliced band, larger bands which after subcloning and sequencing were found to correspond to the insertion of PE70, PE93, and PE118, previously identified in patients' fibroblasts (Fig. 2A). Thus, this variant is responsible for the activation of the intronic pseudoexons, although some residual normal splicing is produced.

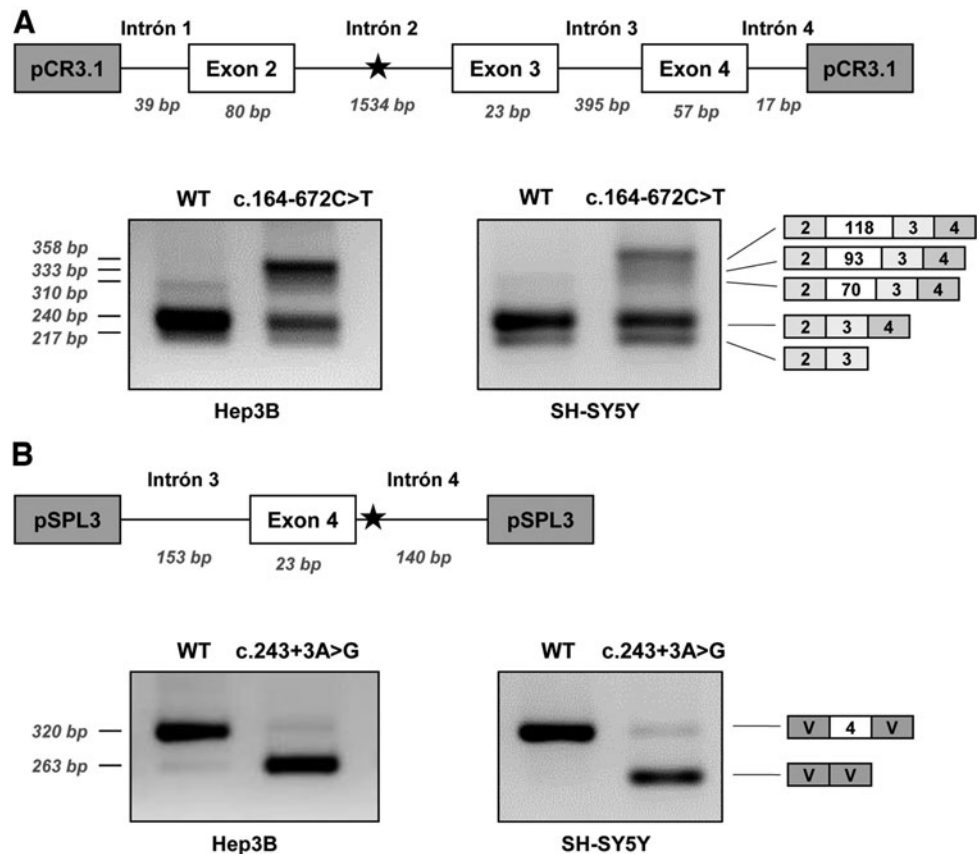
Variant c.243+3A>G was analyzed in a minigene with exon 4 and flanking intronic sequences cloned in the pSPL3 vector. As shown in Fig. 2B the c.243+3A>G variant causes exon 4 skipping, consistent with the predicted decrease in splice score of the mutant 5'ss and with the results obtained in fibroblasts. A faint band corresponding to the correctly spliced transcript is barely detectable, indicating that the normally spliced band observed in patient's fibroblasts (Fig. 1A) is probably mostly derived from the allele carrying the c.164-672C>T variant, which can be thus considered a "leaky" splicing variant.

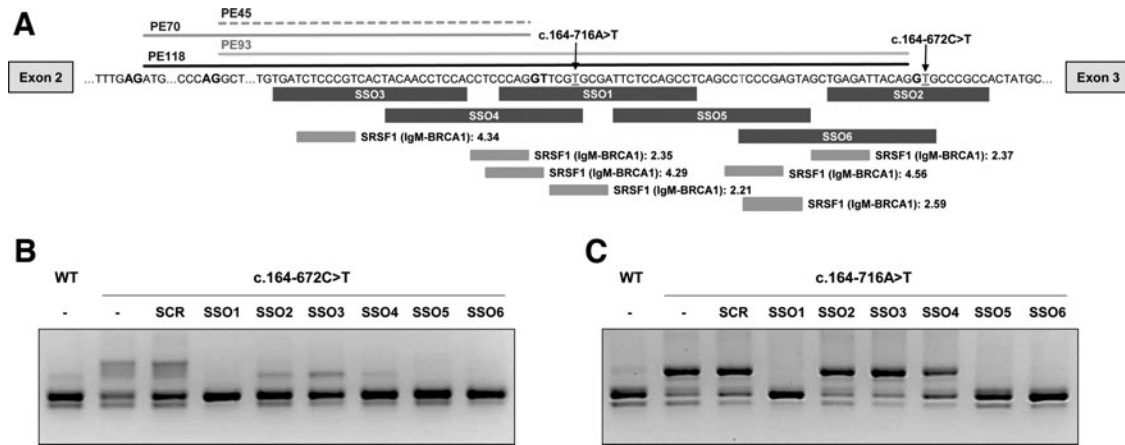
The recognition of the pseudoexons by the splicing machinery may vary depending on the tissue analyzed, due to tissue specific differences in splice factors, as evidenced in several studies [12,49]. Therefore, we subsequently analyzed the effect on splicing of both variants c.243+3A>G and c.164-672C>T in SH-SY5Y neuroblastoma cells, given the relevant role of the *PTS* gene in the central nervous system. The results showed a similar splicing pattern for both variants (Fig. 2A, B).

#### Pseudoexon exclusion with antisense oligonucleotides

Pathogenic pseudoexon inclusion resulting from deep intronic variants has the potential to be reverted using antisense therapy with SSOs. We designed six SSOs targeting the intronic region in which the different pseudoexons are located (Fig. 3A). A scrambled SSO was used as a nontarget control to discriminate between specific and unspecific effects. In this study we used full 2-OMe PS oligonucleotides, with SSO 1 having the wild-type sequence of the morpholino oligonucleotide (PTS-AMO3) used in a previous study to target PE45 and PE70 activated by variant c.164-716A>T (designated as c.164-712A>T in the previous study) [26]. SSO1 and SSO4 cover the 5'ss used by both these pseudoexons. SSO2 and SSO6 were designed for targeting the new 5'ss created by the novel variant c.164-672C>T. SSO3, which targets the TACAACCTC element previously suggested to be essential for PE45 inclusion [28]. SSO5 targets predicted potential binding sites for different SR splice factors.

**FIG. 2.** Minigene analysis of the novel splicing variants identified in patient 1. (A) Scheme of the pCR3.1 minigene used and RT-PCR analysis of WT and mutant (c.164-672C>T) minigenes in Hep3B and SH-SY5Y cells. (B) Scheme of the pSPL3 minigene used and RT-PCR analysis of WT and mutant (c.243+3A>G) minigenes in Hep3B and SH-SY5Y cells. The *star* indicates the location of the variants. WT, wild-type.





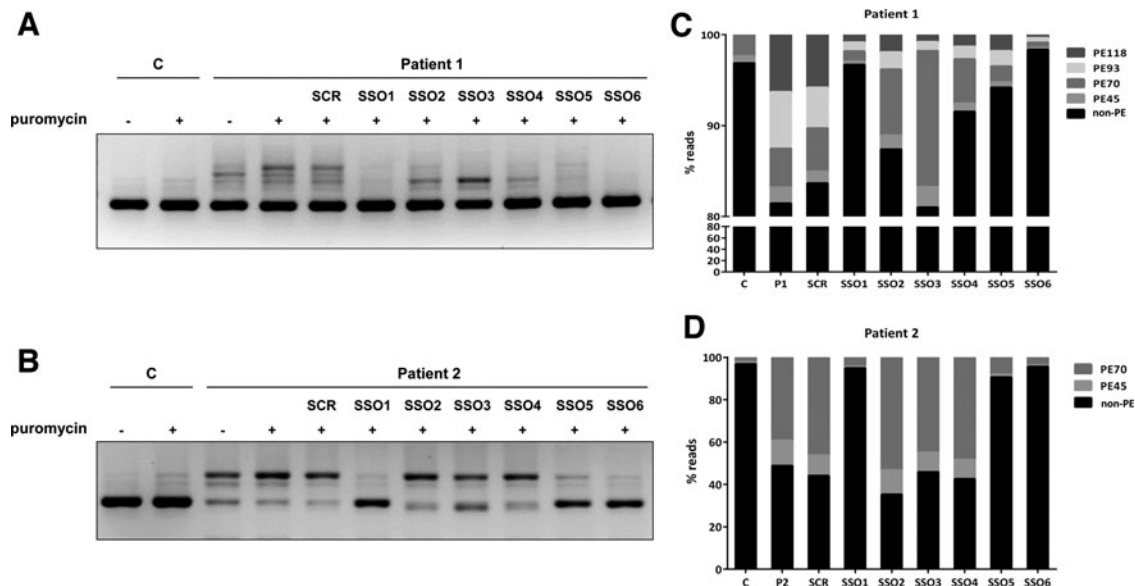
**FIG. 3.** SSO treatment in minigenes. (A) Sequence of the intron 2 region showing the identified pseudoexons, the conserved GT and AG splice sites used, the location of the c.163–716A>T and c.164–672C>T variants, and the region targeted by the designed SSOs (black boxes). Binding sites for SRSF1 predicted by ESE finder are shown in light grey boxes along with their scores. (B, C) RT-PCR analysis in WT and mutant minigenes untransfected (–) or transfected with 50 nM of each SSO. SCR, scrambled SSO; SSO, splice switching antisense oligonucleotide.

Minigenes carrying variant c.164-672C>T were co-transfected with SSOs at 50 nM, and RT-PCR analysis showed that SSOs 1, 5, and 6 completely prevented pseudoexon inclusion (Fig. 3B). For cells transfected with SSOs 2, 3, and 4 a faint band corresponding in size to PE70, which is also detectable in the wild-type untreated minigene, remained after treatment (Fig. 3B).

The same SSOs were also used with the mutant minigene carrying the previously studied variant c.164-716A>T that results in the inclusion of PE45 and PE70 [26]. Again, SSOs 1, 5, and 6 completely prevented pseudoexon inclusion, and in this case, SSO 2, 3, and 4 had no apparent effect (Fig. 3C).

We next transfected patient 1 fibroblasts, treated with puromycin, to inhibit NMD, with the six SSOs. For the sake of clarity, in the RT-PCR analysis we used a reverse primer hybridizing to exon 3, to avoid detection of the exon skipping transcripts resulting from the allele with variant c.243 + 3A>G. The results were concordant with what we observe with minigenes, as SSO 1, 5, and 6 resulted in almost complete disappearance of the higher molecular weight bands, corresponding to pseudoexon inclusions, while SSOs 2, 3, and 4a band presumably corresponding to PE70 remained (Fig. 4A).

The effect of the six SSOs was also analyzed in patient 2 fibroblasts, heterozygous for c.164-716A>T that causes insertion of PE45 and PE70 [26]. Again, SSOs 1, 5, and 6



**FIG. 4.** SSO treatment in patients' fibroblasts. RT-PCR analysis, using primers to amplify a fragment from exon 1 to exon 3, in control and patient 1 (A) or patient 2 (B) fibroblasts, transfected with 50 nM of each SSO. Puromycin was used as NMD inhibitor. (C, D) show the graphical representation of the quantification of each band corresponding to the transcripts without pseudoexons (non-PE) and to the insertion of each pseudoexon, expressed as percentage of reads obtained after deep sequencing the region in puromycin treated samples. NMD, nonsense mediated mRNA decay.



prevented pseudoexon inclusions and SSOs 2–4 showed no significant effect (Fig. 4B). In both patients 1 and 2 fibroblasts, a scrambled SSO had no effect on the splicing pattern.

Examining the hit SSOs we observe that SSO1 covers a sequence shared by all pseudoexons; however, SSOs 5 and 6 target sequences present only in the longer PE93 and PE118. To precisely define and quantify the effect of each of the SSOs on the inclusion of the different pseudoexons, we performed deep sequencing of the RT-PCR products from the transfected fibroblasts. The results are shown in Fig. 4C, D. In patient 1 cells, we observe similar levels (estimated by the number of reads) for PE70, PE93, and PE118, plus a low amount of PE45 (Fig. 4C). PE70 is predominant in patient 2 (Fig. 4D) and is also detectable in low amounts in control fibroblasts. SSO1 and 6 almost completely prevent all pseudoexon inclusion in both patients' fibroblasts, with SSO5 exhibiting a slightly lower effect. SSOs 2–4 prevalently hinder the inclusion of PE93 and PE118 and correspondingly have no effect in patient 2 cells (Fig. 4B, D).

#### *Analysis of splice factors binding to pseudoexon sequences*

SSO1 hybridizes to a region included in all pseudoexons, and its effect could be the result of blocking the binding of splice factors critical for their aberrant inclusion. The positive effect of SSOs 5 and 6 that lie downstream of the 5' splice site (for PE45 and PE70) or internal (PE93 and PE118) to the pseudoexons may also indicate the presence of splicing regulatory sites needed for pseudoexon recognition.

*In silico* analysis of the intronic sequences predict strong SRSF1 binding motifs in the region covered by hit SSOs 1, 5, and 6 (scores 4.3 and 4.6 according to ESEfinder) (Fig. 3A). To further investigate which splice factors bind to sequences targeted by SSOs 1, 5, and 6, we performed RNA affinity pull downs on biotinylated RNA oligonucleotides, containing the wild-type target sequences or with mutations that abolish the SRSF1 motifs (Supplementary Fig. S1). After incubation in HeLa cell nuclear extract, SR and hnRNPA1 proteins bound to each oligonucleotide were analyzed by SDS-PAGE and Western blotting. Using an anti-SR antibody there was no binding above background for SRSF2, SRSF3, SRSF4, SRSF5, or SRSF6 (data not shown). In addition, the pull downs with all RNA oligonucleotides showed little SRSF1 binding, although for PTS5 (antisense sequence to SSO5) we observe a stronger signal, which could be decreased by mutating the SRSF1 motif (Supplementary Fig. S1).

SPRi analysis of PTS5 shows binding of SRSF1 above background. When the potential SRSF1 binding site in PTS5 is disrupted by point mutations this binding is disrupted (Supplementary Fig. S2). SPRi analysis of PTS1 (sequence targeted by SSO1) and PTS6 (sequence targeted by SSO6) confirms the very low binding of SRSF1 observed in the pull-down experiments (Supplementary Fig. S2). Examination of the sequences using recombinant hnRNPA1 in the SPRi analysis indicated binding for all RNA oligonucleotides, confirming that they are all functional (Supplementary Fig. S2).

To provide additional evidence of SRSF1 binding to the targeted sequences, we examined publicly available SRSF1 CLIP datasets, including the K562 and HepG2 ENCODE eCLIP datasets [37], a PAR-CLIP dataset from HeLa cells

[39], and a HITS-CLIP dataset from MDA-LM2 cells [40]. Normally, CLIP datasets are analyzed by considering only the uniquely aligned reads, but this overlooks real binding sites in duplicated sequences. Because the *PTS* PE is derived from the highly repeated *AluS<sub>q</sub>* sequence [28], reads would therefore normally be filtered away. We therefore aligned CLIP reads allowing for up to 100 multiple mapping locations to allow for mapping of reads to duplicated regions, including *AluY* sequence elements. This resulted in reads aligning at or on top of the potential SRSF1 binding sites in *PTS* (Supplementary Fig. S3), providing further evidence that these binding sites may be functional.

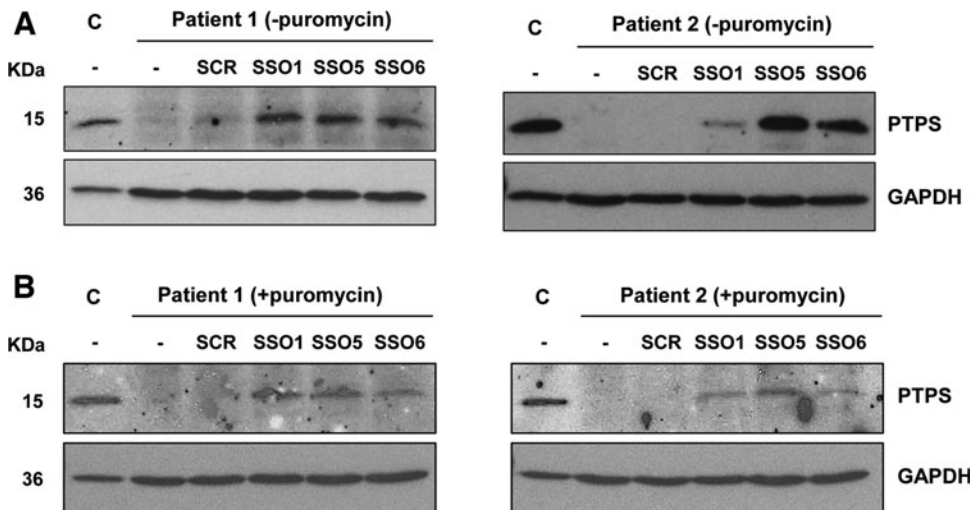
#### *PTPS protein recovery with SSO treatment*

SSOs 1, 5, and 6 corrected the mis-splicing defect in patients 1 and 2 fibroblasts carrying deep-intronic variants activating pseudoexons in intron 2. To illustrate the therapeutic potential of the approach, we analyzed PTPS protein levels in cells treated with the different SSOs or a scrambled SSO. As using an NMD inhibitor may favor correctly spliced transcripts, interfering with the interpretation of the results, we analyzed cells treated or not with puromycin. Both fibroblast cell lines are compound heterozygous, carrying mis-splicing variants in both alleles, although some residual normal splicing can be detected (Fig. 1A).

Western blot analysis revealed no detectable PTPS protein in untreated cells (Fig. 5). This confirms that in fibroblasts the resulting mis-spliced transcripts do not result in stable protein and that the protein translated from the residual normally spliced transcripts escapes detection limits with the antibody used. After transfection with the hit SSOs, PTPS protein is detected in both fibroblast cell lines, in cells treated or not with puromycin. Protein levels in patients' cells transfected with the different SSOs, estimated by densitometric analysis, ranged between 5% and 50%, relative to levels in control fibroblasts, depending on the SSO, treatment with or without puromycin and the cell line (Fig. 5).

## **Discussion**

Deep intronic variants that alter splicing are missed with standard exon-based diagnostic sequencing approaches. However, evidence has accumulated that such variants are an important class of disease-causing variants mainly due to the activation of pseudoexons [9,16]. Thus, it has been proposed that RNA analysis should be incorporated into the diagnostic workflow for patients with specific phenotypes and with no or incomplete genetic diagnosis [13]. In the present study routine clinical exome sequencing identified only one pathogenic *PTS* variant in patient 1 with a biochemical suspicion of PTPS deficiency. We subsequently used a candidate region approach, sequencing a deep-intronic region where previous studies had identified variants activating pseudoexons, followed by *PTS* transcript analysis in patients' fibroblasts. Interestingly, the analysis identified a novel intronic variant, c.164-672C>T, adding to the previously reported variants c.163+1695\_1751del57 and c.164-716A>T in the same region [26]. The novel variant results in the exonization of previously described PE45 and PE70, plus two longer overlapping PE93 and PE118. These two are defined by the newly created 5' splice site and the 3' splice site of the previously reported pseudoexons.



**FIG. 5.** PTPS protein analysis in patients' fibroblasts. Western blot analysis of the PTPS protein in untransfected (–) or transfected fibroblasts from patients 1 and 2, untreated (**A**) or treated (**B**) with puromycin. GAPDH was used as loading control. C, control fibroblasts; PTPS, 6-pyruvoyltetrahydropterin synthase.

Multiple pseudoexon inclusion due to a single intronic splice site coupled with different cryptic acceptor splice sites has already been described [50]. However, in this case, the novel variant also moderately increases the inclusion levels of upstream PE45 and PE70 compared to that observed in control fibroblasts, which could be due to the alleviation of some repressive mechanism, rather than to the occurrence of an activating event, as suggested in other cases [51]. In addition, some introns seem more prone to harbor disease-causing mutations than others, as observed in the *ABCA4* gene in which the clustering in certain introns of deep-intronic variants causing Stargardt disease has been reported [52]. In the present study, the pseudoexons correspond to transposable *Alu* elements, which may lead to the creation of new exons in the genome or to aberrant disease-causing transcripts due to single point variants [8,21].

A very interesting aspect of our results is the fact that the new *PTS* c.164-672C>T variant creates a 5' splice site identical to that used in the disease causing *Alu* derived *F8* pseudoexon that is activated by a deletion removing an inhibitory sequence in the 3' splice site [24], and this 5' splice site in *F8* is also activated when further strengthened by a point mutation altering the +3G>A also leading to disease [24,48]. Moreover, mutations in this sequence create the 5' splice site most frequently used in naturally exonized *Alus* [46,47]. The *PTS* PE118, the most prominent pseudoexon activated by the c.164-672C>T variant, is nearly identical to the *F8* PE, with exactly the same 3' splice site and 5' splice site. This underscores that similar mutations in the potential 5' splice site, as that here reported in *PTS* and previously in *F8*, in the numerous *AluY* sequence elements present in introns in other genes are likely to be a frequent cause of disease causing pseudoexon activation.

Similarly, mutations in the same positions (205 and 206) in the *Alu* sequence as the 5' splice site strengthened by the c.164-716A>T *PTS* variant have been reported to activate disease causing pseudoexons in *COL4A3* and *GUSB* genes [53,54], further supporting that these positions in the *Alu* sequences are high risk sites for disease causing pseudoexon activation.

Generally, most of the reported variants that activate pseudoexons either create or strengthen a preexisting sub-optimal 5' or 3' splice site [8,48], although in other pseudoexons, different types of activating variants have been identified

[10,11]. In the *PTS* intron 2 region, one variant creates a novel 5' splice site (c.164-672C>T), another strengthens a preexisting one (c.164-716A>T), and another brings closer a branch point sequence to the 3' splice site (c.163+1695\_1751del157) [26], revealing the spectrum of rare deleterious variants activating pseudoexons in this region.

Following the identification of the aberrant insertions, SSOs were designed with the aim of restoring normal splicing by forcing the skipping of the pathogenic pseudoexons. This strategy has been validated in minigene assays and in cultured patient-derived cells in a wide range of disorders and has efficiently shown the rescue of these loss-of-function variants [8,22,30]. Our results showed that both in minigenes and in the endogenous gene context using patients' fibroblasts, SSOs 1 and 6 (and to a slightly lesser extent, SSO5) were most effective for all *PTS* intron 2 pseudoexons, providing support for the potential of using a single SSO to treat patients with different pathogenic variants in the region. The relative amounts of each pseudoexon in the final transcript profile in patients' fibroblasts were measured using deep sequencing. This approach represents an efficient way to precisely quantify the efficacy of SSO-mediated splice correction for adjacent or overlapping pseudoexons.

The presence of SRSF1 motifs in the pseudoexon sequences indicated a possible mechanism of action of the hit SSOs, whereby they would block binding of this splice factor to the predicted sites, thus precluding pseudoexon inclusion, as previously described [55,56]. Taken together, our results using RNA pull downs and SPRi and analysis of CLIP data suggest that SSO5 could indeed function by blocking access to an SRSF1 ESE (CTCCCGAGTA) with two predicted overlapping binding sites (CTCCCGA—Score 4.56) and (CCGAGTA—Score 2.59) (Fig. 3A). When these sites are abolished by point mutations, binding decreases (Supplementary Figs. S1, S2). SSO6 covers most of the high score (4.56) SRSF1 motif; thus, it is plausible that it will block SRSF1 binding and cause exon skipping, although as the entire motif is not present, experimentally we observe low SRSF1 binding to the target sequence (Supplementary Figs. S1, S2).

SSO1 blocks the 5' splice site of PE45 and PE70, which may contribute to the exclusion of these pseudoexons. In addition, it is known that secondary RNA structure can affect pre-

mRNA splicing, regulating accessibility of cis-acting or trans-acting elements and the spatial distance in between [57]. In this case, SSO binding may also influence local RNA structure favoring pseudoexon exclusion.

SSO6 targets a region that encompasses the mutated nucleotide at position c.164-672C>T and, therefore, contains a mismatch toward the sequence in patient 2. Contrary to what has been reported in some studies [52,58], this single mismatch between the SSO and its target sequence apparently does not affect its splicing modulation capacity. Similarly, SSO1 with the wild-type sequence functions in patient 2 samples, despite the presence of a mismatch at the site of the c.164-716A>T variant. Of note, it can be envisaged that these same hit SSOs can be used for pseudoexon activating variants in similar *Alu*-derived activated pseudoexons, as the above mentioned are nearly identical in the *F8* gene. This implies for the first time, to our knowledge, that one single SSO could potentially be used for treating patients with different diseases.

Comparing the results obtained in a previous study [26], in which the phosphorodiamidate morpholino oligonucleotide (PMO) chemistry was used at concentrations 20–30  $\mu$ M, we observe similar splicing correction (pseudoexon exclusion) in patient 2 fibroblasts transfected with SSO1, same sequence as PTS-AMO3, but with 2-OMe-PS chemistry (see Fig. 3 in [26], compared to Fig. 4, this work). Both chemistries have proven to be effective *in vitro*, although *in vivo*, the neutral charge of PMOs improves tolerability as it results in lower binding to plasma proteins. Currently, 3 SSOs with PMO chemistry have received FDA approval for the treatment of Duchenne muscular dystrophy [59].

Protein analysis confirmed the recovery of PTPS protein levels after transfection of patient-derived fibroblasts with hit SSOs 1, 5, and 6. Both patients are compound heterozygous, carrying the pseudoexon activating variant in one allele. Although PTPS enzymatic activity could not be determined in this study, the levels of protein achieved here have been previously shown to correlate with a therapeutically relevant increase in enzymatic activity [26].

It must be noted that for rare pathogenic variants, occurring only in one or two families, developing a personalized genetic therapy represents a challenge, although the proof of concept of the feasibility of such an approach for a single patient with a deep intronic variant was recently reported [25] and more n-of-1 RNA therapy trials are being developed in academic or pharmaceutical settings [60,61]. Indeed, the recently founded n-LoRem Foundation aims to discover and develop antisense oligonucleotides for patients with unique mutations and to provide them for free as treatment for life [62]. Another important hurdle in preclinical development for SSOs targeting pseudoexons is the difficulty in generating adequate animal models, as intronic sequences are not conserved among species. However, with the advent of CRISPR/Cas gene editing technologies it is now feasible to easily generate humanized mouse models carrying the desired intronic pseudoexon sequences, although reproducibility of the splicing defect may depend on the whole genomic context and thus should be first confirmed *in vitro* (e.g., using minigenes or a similarly edited mouse cell line).

PE70 that corresponds to the “natural” predicted *Alu* sequence [28] is detectable in low abundance in the wild type context, as evidenced in minigene assays (Figs. 2, 3) and in control fibroblasts (Fig. 4). Residual pseudoexon inclusion

leading to a nonfunctional transcript in control individuals has been detected for several genes involved in disease [63,64]. It was initially considered an undesired unproductive splicing event. However, evidence has accumulated indicating that it is actually a well-represented mechanism of gene expression regulation [65]. During specific dynamic cellular processes, genes may be subjected to alternative splicing, resulting in such nonfunctional transcripts rapidly regulating final mRNA and protein levels [66]. Recently, this knowledge was exploited applying antisense technology to prevent nonproductive pseudoexon inclusion events, resulting in an increase in functional mRNA and protein levels in monogenic diseases characterized by haplo-insufficiency [67,68]. The authors detected by bioinformatic analysis more than 1,000 nonproductive splicing events in disease associated genes and using this approach, termed TANGO (targeted augmentation of nuclear gene output), confirmed an increase in productive mRNA and protein for different genes *in vitro* and *in vivo* [67].

The antisense approach targeting naturally occurring pseudoexon inclusion events could be applied to the *PTS* gene in the case of missense variants retaining residual activity. Over 100 variants have been reported in the *PTS* gene associated with PTPS deficiency, mostly missense variants [69], many of them affecting protein structure and stability [70–72]. It is also known that the milder peripheral forms of PTPS deficiency are associated to partial activity [73]. Future experiments to investigate the feasibility of the TANGO approach can be considered using fibroblasts derived from patients with such genotypes.

The present work adds to the current knowledge of the relatively high prevalence of variant induced pseudoexon inclusion as a disease mechanism in inherited metabolic disorders [74]. These variants are promising candidates for personalized therapies using SSOs, which, along with other RNA-therapies, have registered a boost in the last few years, with more than a dozen drugs approved for clinical use and many others currently in different stages of development. It can be anticipated that this type of drugs may shortly be in use for one or very few patients (personalized medicine) or to treat common diseases, once delivery and toxicity issues are solved [19].

## Acknowledgments

The excellent technical assistance of E. Montalvo and M. Alvarez and of A. Rassi from the Division of Clinical Chemistry and Biochemistry at the University Children’s Hospital Zürich is gratefully acknowledged, as well as the expert advice of R. Ramos from Fundación Parque Científico de Madrid.

## Author Disclosure Statement

No competing financial interests exist.

## Funding Information

This work was funded by grant PID2019-105344RB-I00/AEI/10.13039/501100011033 from Spanish Ministry of Science and Innovation, COST (European Cooperation in Science and Technology) Action CA17103, Fundación Isabel Gemio, Fundación La Caixa (LCF/PR/PR16/11110018) and grants to B.S.A. from the Danish Medical Research Council (FSS) (No. 9039-00281B) and the Novo Nordisk Foundation

(NNF19OC0058588). Centro de Biología Molecular Severo Ochoa receives an institutional grant from Fundación Ramón Areces.

**Supplementary Material**

- Supplementary Data
- Supplementary Figure S1
- Supplementary Figure S2
- Supplementary Figure S3
- Supplementary Table S1

**References**

1. Werner ER, N Blau, and B Thony. (2011). Tetrahydrobiopterin: biochemistry and pathophysiology. *Biochem J* 438:397–414.
2. Blau N, B Thony, GH Cotton, and K Hyland. (2001). Disorders of tetrahydrobiopterin and related biogenic amines. In: *The Metabolic and Molecular Basis of Inherited Disease*, Scriver C.R. *et al.* eds., 8th edn. McGraw-Hill, New York, pp. 1725–1776.
3. Manzoni F, E Salvatici, A Burlina, A Andrews, M Pasquali, and N Longo. (2020). Retrospective analysis of 19 patients with 6-pyruvoyl tetrahydropterin synthase deficiency: prolactin levels inversely correlate with growth. *Mol Genet Metab* 131:380–389.
4. Baralle D, and E Buratti. (2017). RNA splicing in human disease and in the clinic. *Clin Sci (Lond)* 131:355–368.
5. Scotti MM, and MS Swanson. (2016). RNA mis-splicing in disease. *Nat Rev Genet* 17:19–32.
6. Andresen BS, and AR Krainer. (2009). When the genetic code is not enough-how sequence variations can alter pre-mRNA splicing and cause (complex) disease. In: *Genetics of Human Complex Diseases*, Almasy, L. and Al-Chalabi, A. eds. Cold Spring Harbor Laboratory Press, New York, pp. 165–182.
7. Soemedi R, KJ Cygan, CL Rhine, J Wang, C Bulacan, J Yang, P Bayrak-Toydemir, J McDonald, and WG Fairbrother. (2017). Pathogenic variants that alter protein code often disrupt splicing. *Nat Genet* 49:848–855.
8. Dhir A, and E Buratti. (2010). Alternative splicing: role of pseudoexons in human disease and potential therapeutic strategies. *Febs J* 277:841–855.
9. Vaz-Drago R, N Custodio, and M Carmo-Fonseca. (2017). Deep intronic mutations, and human disease. *Hum Genet* 136:1093–1111.
10. Petersen USS, TK Doktor, and BS Andresen. (2022). Pseudoexon activation in disease by non-splice site deep intronic sequence variation - wild type pseudoexons constitute high-risk sites in the human genome. *Hum Mutat* 43: 103–127.
11. Keegan NP, SD Wilton, and S Fletcher. (2022). Analysis of pathogenic pseudoexons reveals novel mechanisms driving cryptic splicing. *Front Physiol* 12:806946.
12. Cummings BB, JL Marshall, T Tukiainen, M Lek, S Donkervoort, AR Foley, V Bolduc, LB Waddell, SA Sandaradura, *et al.* (2017). Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. *Sci Transl Med* 9:eaa15209.
13. Kremer LS, DM Bader, C Mertes, R Kopajtich, G Pichler, A Iuso, TB Haack, E Graf, T Schwarzmayr, *et al.* (2017). Genetic diagnosis of Mendelian disorders via RNA sequencing. *Nat Commun* 8:15824.
14. Gonorazky HD, S Naumenko, AK Ramani, V Nelakuditi, P Mashouri, P Wang, D Kao, K Ohri, S Viththiyapaskaran, *et al.* (2019). Expanding the boundaries of RNA sequencing as a diagnostic tool for rare Mendelian disease. *Am J Hum Genet* 104:1007.
15. Fresard L, C Smail, NM Ferraro, NA Teran, X Li, KS Smith, D Bonner, KD Kernohan, S Marwaha, *et al.* (2019). Identification of rare-disease genes using blood transcriptome sequencing and large control cohorts. *Nat Med* 25:911–919.
16. Truty R, K Ouyang, S Rojahn, S Garcia, A Colavin, B Hamlington, M Freivogel, RL Nussbaum, K Nykamp, *et al.* (2021). Spectrum of splicing variants in disease genes and the ability of RNA analysis to reduce uncertainty in clinical interpretation. *Am J Hum Genet* 108:696–708.
17. Crooke ST, JL Witztum, CF Bennett, and BF Baker. (2018). RNA-targeted therapeutics. *Cell Metab* 27:714–739.
18. Aartsma-Rus A. (2018). Genetic therapies for spinal muscular atrophy type 1. *Lancet Neurol* 17:111–112.
19. Hammond SM, A Aartsma-Rus, S Alves, SE Borgos, RAM Buijssen, RWJ Collin, G Covello, MA Denti, LR Desviat, *et al.* (2021). Delivery of oligonucleotide-based therapeutics: challenges and opportunities. *EMBO Mol Med* 13: e13243.
20. Crooke ST, XH Liang, BF Baker, and RM Crooke. (2021). Antisense technology: a review. *J Biol Chem* 296:100416.
21. Vorechovsky I. (2010). Transposable elements in disease-associated cryptic exons. *Hum Genet* 127:135–154.
22. Perez B, L Rodriguez-Pascau, L Vilageliu, D Grinberg, M Ugarte, and LR Desviat. (2010). Present and future of antisense therapy for splicing modulation in inherited metabolic disease. *J Inherit Metab Dis* 33:397–403.
23. Sangermano R, A Garanto, M Khan, EH Runhart, M Bauwens, NM Bax, van den LI Born, MI Khan, SS Cornelis, *et al.* (2019). Deep-intronic ABCA4 variants explain missing heritability in Stargardt disease and allow correction of splice defects by antisense oligonucleotides. *Genet Med* 21:1751–1760.
24. Jourdy Y, A Janin, M Fretigny, A Lienhart, C Negrier, D Bozon, and C Vinciguerra. (2018). Recurrent F8 intronic deletion found in mild hemophilia a causes Alu exonization. *Am J Hum Genet* 102:199–206.
25. Kim J, C Hu, C Moufawad El Achkar, LE Black, J Douville, A Larson, MK Pendergast, SF Goldkind, EA Lee, *et al.* (2019). Patient-customized oligonucleotide therapy for a rare genetic disease. *N Engl J Med* 381: 1644–1652.
26. Brasil S, HM Viecelli, D Meili, A Rassi, LR Desviat, B Perez, M Ugarte, and B Thony. (2011). Pseudoexon exclusion by antisense therapy in 6-pyruvoyl-tetrahydropterin synthase deficiency. *Hum Mutat* 32:1019–1027.
27. Bonafe L, B Thony, W Leimbacher, L Kierat, and N Blau. (2001). Diagnosis of dopa-responsive dystonia and other tetrahydrobiopterin disorders by the study of biopterin metabolism in fibroblasts. *Clin Chem* 47:477–485.
28. Meili D, J Kralovicova, J Zagalak, L Bonafe, L Fiori, N Blau, B Thony, and I Vorechovsky. (2009). Disease-causing mutations improving the branch site and polypyrimidine tract: pseudoexon activation of LINE-2 and antisense Alu lacking the poly(T)-tail. *Hum Mutat* 30:823–831.
29. Aartsma-Rus A, L van Vliet, M Hirschi, AA Janson, H Heemskerk, CL de Winter, S de Kimpe, JC van Deutekom, PA t Hoen, *et al.* (2009). Guidelines for antisense

- oligonucleotide design and insight into splice-modulating mechanisms. *Mol Ther* 17:548–553.
30. Albert S, A Garanto, R Sangermano, M Khan, NM Bax, CB Hoyng, J Zernant, W Lee, R Allikmets, *et al.* (2018). Identification and rescue of splice defects caused by two neighboring deep-intronic ABCA4 mutations underlying stargardt disease. *Am J Hum Genet* 102:517–527.
  31. Martínez-Pizarro A, M Dembic, B Perez, BS Andresen, and LR Desviat. (2018). Intronic PAH gene mutations cause a splicing defect by a novel mechanism involving U1snRNP binding downstream of the 5' splice site. *PLoS Genet* 14: e1007360.
  32. Holm LL, TK Doktor, MB Hansen, USS Petersen, and BS Andresen. (2022). Vulnerable exons, like ACADM exon 5, are highly dependent on maintaining a correct balance between splicing enhancers and silencers. *Hum Mutat* 43: 53–265.
  33. Desmet FO, D Hamroun, M Lalande, G Collod-Beroud, M Claustres, and C Beroud. (2009). Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res* 37:e67.
  34. Yeo G, and CB Burge. (2004). Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol* 11:377–394.
  35. Cartegni L, J Wang, Z Zhu, MQ Zhang, and AR Krainer. (2003). ESEfinder: a web resource to identify exonic splicing enhancers. *Nucleic Acids Res* 31:3568–3571.
  36. Paz I, I Kosti, M Ares, Jr., M Cline, and Y Mandel-Gutfreund. (2014). RBPmap: a web server for mapping binding sites of RNA-binding proteins. *Nucleic Acids Res* 42 (Web Server issue):W361–W367.
  37. Van Nostrand EL, GA Pratt, AA Shishkin, C Gelboin-Burkhart, MY Fang, B Sundararaman, SM Blue, TB Nguyen, C Surka, *et al.* (2016). Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat Methods* 13:508–514.
  38. Martin M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.J* 17: 10–12.
  39. Xiao W, S Adhikari, U Dahal, YS Chen, YJ Hao, BF Sun, HY Sun, A Li, XL Ping, *et al.* (2016). Nuclear m(6)A Reader YTHDC1 regulates mRNA splicing. *Mol Cell* 61: 507–519.
  40. Fish L, A Navickas, B Culbertson, Y Xu, HCB Nguyen, S Zhang, M Hochman, R Okimoto, BD Dill, *et al.* (2019). Nuclear TARBP2 drives oncogenic dysregulation of RNA splicing and decay. *Mol Cell* 75:967–981e9.
  41. Dobin A, CA Davis, F Schlesinger, J Drenkow, C Zaleski, S Jha, P Batut, M Chaisson, and TR Gingeras. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29:15–21.
  42. Hahne F, and R Ivanek. (2016). Visualizing genomic data using Gviz and bioconductor. *Methods Mol Biol* 1418:335–351.
  43. Lawrence M, W Huber, H Pages, P Aboyoun, M Carlson, R Gentleman, MT Morgan, and VJ Carey. (2013). Software for computing and annotating genomic ranges. *PLoS Comput Biol* 9:e1003118.
  44. Blau N, and Spronsen FJ. (2014). Disorders of phenylalanine and tetrahydrobiopterin metabolism. In: *Physician's guide to the diagnosis, treatment and follow-up of inherited metabolic diseases*, Blau N, Duran M, Gibson MK, Dionisi-Vici C, eds. Springer Berlin, Heidelberg, pp. 3–21.
  45. Liu TT, YH Chang, SH Chiang, YL Yang, WM Yu, and KJ Hsiao. (2001). Identification of three novel 6-pyruvoyl-tetrahydropterin synthase gene mutations (226C>T, IVS3 + 1G>A, 116–119delTGTT) in Chinese hyperphenylalaninemia caused by tetrahydrobiopterin synthesis deficiency. *Hum Mutat* 18:83.
  46. Sela N, B Mersch, N Gal-Mark, G Lev-Maor, A Hotz-Wagenblatt, and G Ast. (2007). Comparative analysis of transposed element insertion within human and mouse genomes reveals Alu's unique role in shaping the human transcriptome. *Genome Biol* 8:R127.
  47. Sorek R, G Lev-Maor, M Reznik, T Dagan, F Belinky, D Graur, and G Ast. (2004). Minimal conditions for exonization of intronic sequences: 5'' splice site formation in alu exons. *Mol Cell* 14:221–231.
  48. Dericquebourg A, Y Jourdy, M Fretigny, A Lienhart, S Claeysens, C Ternisien, P Boisseau, PS Rohrlisch, Negrier C, *et al.* (2020). Identification of new F8 deep intronic variations in patients with haemophilia A. *Haemophilia* 26: 847–854.
  49. Palhais B, M Dembic, R Sabaratnam, KS Nielsen, TK Doktor, GH Bruun, and BS Andresen. (2016). The prevalent deep intronic c. 639+919 G>A GLA mutation causes pseudoexon activation and Fabry disease by abolishing the binding of hnRNPA1 and hnRNP A2/B1 to a splicing silencer. *Mol Genet Metab* 119:258–269.
  50. Ikezawa M, I Nishino, Y Goto, T Miike, and I Nonaka. (1999). Newly recognized exons induced by a splicing abnormality from an intronic mutation of the dystrophin gene resulting in Duchenne muscular dystrophy. *Mutations in brief no. 213*. Online. *Hum Mutat* 13:170.
  51. Grodecka L, T Kovacova, M Kramarek, S Seneca, K Stouffs, C De Laet, F Majer, T Krsjakova, P Hujova, *et al.* (2017). Detailed molecular characterization of a novel IDS exonic mutation associated with multiple pseudoexon activation. *J Mol Med (Berl)* 95:299–309.
  52. Khan M, G Arno, A Fakin, DA Parfitt, PPA Dhooge, S Albert, NM Bax, L Duijkers, M Niblock, *et al.* (2020). Detailed phenotyping and therapeutic strategies for intronic ABCA4 variants in Stargardt disease. *Mol Ther Nucleic Acids* 21:412–427.
  53. Knebelmann B, L Forestier, L Drouot, S Quinones, C Chuet, F Benessy, J Saus, and C Antignac. (1995). Splice-mediated insertion of an Alu sequence in the COL4A3 mRNA causing autosomal recessive Alport syndrome. *Hum Mol Genet* 4:675–679.
  54. Vervoort R, R Gitzelmann, W Lissens, and I Liebaers. (1998). A mutation (IVS8+0.6kdbdelTC) creating a new donor splice site activates a cryptic exon in an Alu-element in intron 8 of the human beta-glucuronidase gene. *Hum Genet* 103:686–93.
  55. Palhais B, VS Praestegaard, R Sabaratnam, TK Doktor, S Lutz, P Burda, T Suormala, M Baumgartner, B Fowler, *et al.* (2015). Splice-shifting oligonucleotide (SSO) mediated blocking of an exonic splicing enhancer (ESE) created by the prevalent c.903+469T>C MTRR mutation corrects splicing and restores enzyme activity in patient cells. *Nucleic Acids Res* 43:4627–4639.
  56. Homolova K, P Zavadakova, TK Doktor, LD Schroeder, V Kozich, and BS Andresen. (2010). The deep intronic c.903+469T>C mutation in the MTRR gene creates an SF2/ASF binding exonic splicing enhancer, which leads to pseudoexon activation and causes the cbIE type of homocystinuria. *Hum Mutat* 31:437–444.

57. Buratti E, and FE Baralle. (2004). Influence of RNA secondary structure on the pre-mRNA splicing process. *Mol Cell Biol* 24:10505–10514.
58. Aartsma-Rus A, WE Kaman, M Bremmer-Bout, AA Janson, JT den Dunnen, GJ van Ommen, and JC van Deutekom. (2004) Comparative analysis of antisense oligonucleotide analogs for targeted DMD exon 46 skipping in muscle cells. *Gene Ther* 11:1391–1398.
59. Gupta A, JL Andresen, RS Manan, and R Langer. (2021). Nucleic acid delivery for therapeutic applications. *Adv Drug Deliv Rev* 178:113834.
60. Mullard A. (2020). N-of-1 drugs push biopharma frontiers. *Nat Rev Drug Discov* 19:151–153.
61. Kaiser J. (2021). RNA drugs custommade for rare diseases face wider test. *Science* 374:672.
62. Crooke ST. (2022). Addressing the needs of patients with ultra-rare mutations one patient at a time: the n-Lorem approach. *Nucleic Acid Ther* 32:95–100.
63. Broeders M, K Smits, B Goynuk, E Oussoren, H van den Hout, AJ Bergsma, AT van der Ploeg, and W Pijnappel. (2020). A generic assay to detect aberrant ARSB splicing and mRNA degradation for the molecular diagnosis of MPS VI. *Mol Ther Methods Clin Dev* 19:174–185.
64. Sakaguchi N, and M Suyama. (2021). In silico identification of pseudo-exon activation events in personal genome and transcriptome data. *RNA Biol* 18:382–390.
65. McGlincy NJ, and CW Smith. (2008). Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense? *Trends Biochem Sci* 33:385–393.
66. Mendell JT, NA Sharifi, JL Meyers, F Martinez-Murillo, and HC Dietz. (2004). Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nat Genet* 36:1073–1078.
67. Lim KH, Z Han, HY Jeon, J Kach, E Jing, S Weyn-Vanhentenryck, M Downs, A Corriero, R Oh, *et al.* (2020). Antisense oligonucleotide modulation of non-productive alternative splicing upregulates gene expression. *Nat Commun* 11:3501.
68. Han Z, C Chen, A Christiansen, S Ji, Q Lin, C Anumonwo, C Liu, SC Leiser, Meena, *et al.* (2020). Antisense oligonucleotides increase Scn1a expression and reduce seizures and SUDEP incidence in a mouse model of Dravet syndrome. *Sci Transl Med* 12.
69. Himmelreich N, N Blau, and B Thony. (2021). Molecular and metabolic bases of tetrahydrobiopterin (BH4) deficiencies. *Mol Genet Metab* 133:123–136.
70. Oppliger T, B Thony, H Nar, D Burgisser, R Huber, CW Heizmann, and N Blau. (1995). Structural and functional consequences of mutations in 6-pyruvoyltetrahydropterin synthase causing hyperphenylalaninemia in humans. Phosphorylation is a requirement for in vivo activity. *J Biol Chem* 270:29498–29506.
71. Oppliger T, B Thony, C Kluge, A Matasovic, CW Heizmann, A Ponzoni, M Spada, and N Blau. (1997). Identification of mutations causing 6-pyruvoyl-tetrahydropterin synthase deficiency in four Italian families. *Hum Mutat* 10:25–35.
72. Muniz JRC, NW Szeto, R Frise, WH Lee, XS Wang, B Thony, N Himmelreich, N Blau, KJ Hsiao, *et al.* (2019). Role of protein structure in variant annotation: structural insight of mutations causing 6-pyruvoyl-tetrahydropterin synthase deficiency. *Pathology* 51:274–280.
73. Niederwieser A, H Shintaku, W Leimbacher, HC Curtius, J Hyaneck, J Zeman, and W Endres. (1987). “Peripheral” tetrahydrobiopterin deficiency with hyperphenylalaninaemia due to incomplete 6-pyruvoyl tetrahydropterin synthase deficiency or heterozygosity. *Eur J Pediatr* 146:228–232.
74. Perez B, L Vilageliu, D Grinberg, and LR Desviat. (2014). Antisense mediated splicing modulation for inherited metabolic diseases: challenges for delivery. *Nucleic Acid Ther* 24:48–56.

Address correspondence to:

Lourdes R. Desviat, PhD  
 Centro de Biología Molecular Severo Ochoa  
 Nicolás Cabrera 1  
 Madrid 28049  
 Spain

E-mail: lruiz@cbm.csic.es

Received for publication August 24, 2021; accepted after revision March 14, 2022.