PMD patient mutations reveal a long-distance intronic interaction that regulates \textit{PLP1/DM20} alternative splicing

Jennifer R. Taube\textsuperscript{1}, Karen Sperle\textsuperscript{1}, Linda Banser\textsuperscript{1}, Pavel Seeman\textsuperscript{2}, Barbra Charina V. Cavan\textsuperscript{3}, James Y. Garbern\textsuperscript{4,†} and Grace M. Hobson\textsuperscript{1,5,6,∗}

\textsuperscript{1}Nemours Biomedical Research, Alfred I. duPont Hospital for Children, Wilmington, DE 19803, USA, \textsuperscript{2}Department of Child Neurology, DNA Laboratory, 2nd School of Medicine, Charles University and University Hospital Motol, 150 06 Prague 5, Czech Republic, \textsuperscript{3}Department of Pediatrics, Cebu Institute of Medicine, 6000 Cebu City, Philippines, \textsuperscript{4}Department of Neurology, University of Rochester Medical Center, Rochester, NY 14642, USA, \textsuperscript{5}Department of Biological Sciences, University of Delaware, Newark, DE 19716, USA and \textsuperscript{6}Department of Pediatrics, Jefferson Medical College, Thomas Jefferson University, Philadelphia, PA 19107, USA

Received January 31, 2014; Revised April 28, 2014; Accepted May 27, 2014

Alternative splicing of the proteolipid protein 1 gene (\textit{PLP1}) produces two forms, \textit{PLP1} and \textit{DM20}, due to alternative use of 5′ splice sites with the same acceptor site in intron 3. The \textit{PLP1} form predominates in central nervous system RNA. Mutations that reduce the ratio of \textit{PLP1} to \textit{DM20}, whether mutant or normal protein is formed, result in the X-linked leukodystrophy Pelizaeus-Merzbacher disease (PMD). We investigated the ability of sequences throughout \textit{PLP1} intron 3 to regulate alternative splicing using a splicing minigene construct transfected into the oligodendrocyte cell line, Oli-neu. Our data reveal that the alternative splice of \textit{PLP1} is regulated by a long-distance interaction between two highly conserved elements that are separated by 581 bases within the 1071-base intron 3. Further, our data suggest that a base-pairing secondary structure forms between these two elements, and we demonstrate that mutations of either element designed to destabilize the secondary structure decreased the \textit{PLP1/DM20} ratio, while swap mutations designed to restore the structure brought the \textit{PLP1/DM20} ratio to near normal levels. Sequence analysis of intron 3 in families with clinical symptoms of PMD who did not have coding-region mutations revealed mutations that segregated with disease in three families. We showed that these patient mutations, which potentially destabilize the secondary structure, also reduced the \textit{PLP1/DM20} ratio. This is the first report of patient mutations causing disease by disruption of a long-distance intronic interaction controlling alternative splicing. This finding has important implications for molecular diagnostics of PMD.

INTRODUCTION

One of the major challenges facing molecular diagnostics professionals today is determining whether a particular DNA variant is pathogenic. When the variant is in a coding region and results in an amino acid change, or it is in a splice donor or acceptor site, there are computer algorithms to help predict pathogenicity (1–6). However, when the variants are in other DNA regions, it can be difficult to predict the outcomes for patients (7–9), and the main avenue currently available is to determine a potential mechanism for the variant of unknown significance and design an assay to test it. For example, variants in transcriptional enhancer regulatory regions, those in potential alternative splicing enhancer or silencer regions and those within introns can be tested in transfection experiments and transgenic animals, and by protein factor binding assays to help predict pathogenicity. Although interesting disease mechanisms have been identified for variants in introns, including activation of

†Deceased 10 November 2011.

∗To whom correspondence should be addressed at: Nemours Biomedical Research, Alfred I. duPont Hospital for Children, RC1-235, 1701 Rockland Road, Wilmington, DE 19803, USA. Tel: +1 3026516829; Fax: +1 3026516767; Email: ghobson@nemours.org

© The Author 2014. Published by Oxford University Press. All rights reserved.
For Permissions, please email: journals.permissions@oup.com
cryptic splice sites and interruption of secondary structure at splice sites (7,8,10), the current practice in many molecular diagnostics laboratories is to sequence only coding, and splice donor and acceptor sequences, and sometimes the promoter regions of genes.

In the 13 years that our Molecular Diagnostics Laboratory at Nemours/Alfred I. duPont Hospital for Children has been testing for mutations of the proteolipid protein 1 gene (PLP1) in the X-linked leukodystrophy Pelizaeus-Merzbacher disease (PMD), many patients have tested negative, in large part due to difficulties inherent in clinical diagnosis of PMD because of the wide disability spectrum and the overlap of symptoms with other leukodystrophies. The disability spectrum in PMD ranges from patients with mild developmental delay who are able to walk and speak into their teen years to patients with severe developmental delay who never walk and never develop speech. Results of magnetic resonance imaging (MRI) testing of PMD patients show a diffuse pattern of hypomyelination and clinical signs include nystagmus, hypotonia progressing to spasticity, ataxia, titubation, tremors, athetotic movements and differing degrees of cognitive impairment (11,12). These MRI results and clinical signs are shared by patients with Pelizaeus-Merzbacher-like disease (MIM#608803) caused by autosomal recessive mutations of the gap junction protein, gamma 2 gene (GJC2), and they overlap with the MRI results and clinical signs of patients who have other leukodystrophies, such as hypomyelination with congenital cataract caused by mutations of the family with sequence similarity 16, member A gene (FAM126A, MIM# 610531), 4H syndrome caused by mutations of the polymerase (RNA) III (DNA-directed) polypeptide A and polypeptide B genes (POLR3A, MIM#614258; POLR3B, MIM#614366), and Allan-Herndon-Dudley syndrome caused by mutations of the solute carrier family 16, member 2 gene (SLC16A2, also called MCT8, MIM#300095). Testing for other leukodystrophies may be considered when molecular diagnostic testing for PMD is negative (11). However, because the families of some patients who test negative for PMD show strong X-linked recessive inheritance and we and others have shown that the alternative splice of PLP1 is important for normal function (13–16), we postulated that PMD-causing mutations might be found in regulatory regions for alternative splicing of PLP1 that were not being sequenced by our molecular diagnostics laboratory.

The normal splicing pattern of the PLP1 gene product includes two alternatively spliced transcripts, PLP1 and DM20, due to use of two alternative splice donor sites for exon 3A and 3B joining the same acceptor site for exon 4 (Fig. 1A). The PLP1 transcript, which is the predominant form expressed in myelin-producing oligodendrocytes of the central nervous system (CNS), is 105 bases longer than DM20 and gives rise to a protein that is 35 amino acids larger. The alternatively spliced forms are developmentally regulated, as DM20 is the predominant form prior to onset of myelination, but the PLP1 isoform becomes the predominant form when transcription of the PLP1 dramatically increases during myelination (17,18).

We and others have described the effects of patient mutations at the PLP1 and DM20 splice donor sites and their common acceptor site, as well as patient mutations within exonic and intronic enhancers of PLP1 alternative splicing (13–16,19–21). Patient mutations in exonic enhancers were detected in routine sequencing conducted by our diagnostics laboratory. The previously described intronic enhancer was detected because it is close to the splice donor site for the PLP1 form and within the region amplified and sequenced using the primers selected for sequencing exon 3 (13,14).

To determine potential mechanisms that could be disrupted by patient mutations in regulatory regions for alternative splicing of PLP1, we extended our splicing analyses to the whole of intron 3, a 1071-base intron. Here, we report that intron 3 has regions throughout its length that affect alternative splicing. A particularly strong regulatory element was identified at 323 to 308 bases upstream of the splice acceptor. We showed that this element participates in a long-distance interaction with another element that is 151–166 bases from the splice donor site of intron 3. Our data suggest that these two elements in the pre-mRNA participate in a double-stranded intra-intronic secondary structure that controls PLP1 alternative splicing. In addition, we report that patients in three families with no coding mutations of PLP1 whose clinical signs were highly suggestive of PMD had mutations in one of these elements. These mutations, which were predicted to destabilize the secondary structure, reduced the PLP1/DM20 ratio. Our data strongly suggest that a secondary structure within intron 3 is important for controlling regulation of the PLP1 alternative splice and that patient mutations can cause PMD by disrupting this structure. We therefore advise molecular diagnostics laboratories to include intron 3 when sequencing PLP1 for PMD diagnostics. This is the first report of patient mutations that cause disease by disrupting a long-distance intronic interaction controlling alternative splicing.

RESULTS

Deletions throughout intron 3 of PLP1 greatly reduced the ratio of PLP1 to DM20 alternatively spliced products

PLP1 is subject to alternative splicing, by use of either of two possible donor sites for exon 3 (Fig. 1A). We postulated that intron 3 of PLP1 might contain sequences important for regulation of the alternative splice and that PMD patient mutations could occur in these regulatory regions for several reasons. Intron splice enhancers (ISEs) and silencers (ISSs) are known to regulate alternative splicing of other genes (22). Also, we and others have identified splice site mutations that affect alternative splicing of PLP1 and cause PMD, some of which do not cause amino acid changes (13,14,16,21), and we have identified a patient mutation that disrupts an enhancer of PLP1/DM20 alternative splicing (13,23). We examined the conservation of sequences in intron 3 because conserved regions can indicate conserved function, such as a binding site for splicing regulatory factors (24–26). While intron 3 of PLP1 is 78% identical between mouse and human, the other introns of human PLP1 range from 47 to 64% identity to mouse (alignments not shown), suggesting that sequences in intron 3 may regulate PLP1/DM20 alternative splicing throughout the mammalian lineage.

To determine whether regulatory regions for the PLP1/DM20 alternative splice are limited to the immediate vicinity of the splice site or are located throughout intron 3, we created six overlapping deletions of 210 bp each in the minigene splicing
Three patient mutations and other base substitutions in intron 3 reduced the PLP1/DM20 ratio

In parallel to our direct testing of the intron 3 mutant constructs, we conducted a molecular analysis for potential splice-regulatory mutations in intron 3 of PLP1 in candidate families with mild classical PMD and suggested X-linked inheritance but no mutations in coding regions of PLP1 or GJC2 and no duplication of the PLP1 locus. The clinical presentation summary is in Table 1, and brain MRI is in Supplementary Material, Figure S1A. In families A and C, we examined the CA microsatellite marker in intron 3 (30), and PLP1 was implicated as the possible disease-causing gene (data not shown). On further analysis of these families and in family B, we identified sequence variants in intron 3 of PLP1 that were not previously described: family A, c.454-322G > A; family B, c.454-314T > A; family C, c.454-314T > G. In each family, the sequence variant segregated with disease, implicating the variants as the disease-causing mutations (Fig. 2).

To confirm that the mutation c.454-322G > A altered splicing of PLP1 in patient cells, we characterized the PLP1/DM20 ratio in cultured skin fibroblasts from the patient (II.5) in family A. RT-PCR of RNA from patient fibroblasts and a normal fibroblast control demonstrated a reduction of PLP1 message in the fibroblasts from the affected individual II.5 (Supplementary Material, Fig. S1B). The levels of PLP1 expressed in the cells of the patient were 6–9% of normal. Skin fibroblasts were not available from patients in families B and C.

The three patient mutations are within the 210-base E region, removal of which reduced the PLP1/DM20 alternative splice ratio (Fig. 1D). To further narrow down potential sequences portions of PLP1 intron 3 either deleted or replaced with 210 bases of unrelated sequence. Letters to the left were used to name the deletion and replacement constructs. (D) Quantification of PLP1/DM20 mRNA in Oli-neu cells transfected with deletion constructs. Standard deviation for normal was ± 0.00. (E) Quantification of PLP1/DM20 mRNA in Oli-neu cells transfected with replacement constructs. Standard deviation for normal was ± 0.01. Graphs in (D) and (E) represent the ratio of PLP1 product to the DM20 product normalized to the value for the normal construct (mean of replicates ± SD). All constructs were tested in triplicate. In cases where constructs were tested in repeated experiments, the mean of the ratio, standard deviations and t-tests are presented. Mutated constructs that had an expression ratio significantly different from the normal construct, as measured by Student’s t-test, are indicated: ***P < 0.001; **P < 0.01.
<table>
<thead>
<tr>
<th>Family member</th>
<th>Family A, USA II.5</th>
<th>Family B, Philippines II.6</th>
<th>Family C, Czech Republic III.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pregnancy</td>
<td>Full term, no complications</td>
<td>Normal</td>
<td>No complications</td>
</tr>
<tr>
<td>Early signs of disease</td>
<td>Limited leg movements at 3 months, rotary nystagmus at 6 months, unable to sit up at 9 months</td>
<td>Delayed early motor milestones, nystagmus at 6 months</td>
<td>Delayed early motor milestones, chin tremor, hypotonia at 3 months</td>
</tr>
<tr>
<td>Speech</td>
<td>Single words at 6 months, dysarthric but understandable as adult</td>
<td>Poor speech pronunciation and volume as adult</td>
<td>Nystagmus and hypotonia at 3 months</td>
</tr>
<tr>
<td>Walking</td>
<td>Began at 2.5 years, gait ataxic and increasingly spastic, crutches at age 10 years, wheelchair at approximately 13 years</td>
<td>Able to walk independently, spasticity increased from age 4, wheelchair use in adolescence</td>
<td>Able to walk independently, cerebellar ataxic gait</td>
</tr>
<tr>
<td>Ocular findings</td>
<td>Extraocular movements conjugate, end-gaze nystagmus, acuity of 20/30 as adult</td>
<td>Vision affected by optic nerve damage as adult</td>
<td>Mild nystagmus improved continuously to normal</td>
</tr>
<tr>
<td>Limb reflexes</td>
<td>Abnormally brisk</td>
<td>Hyperreflexia</td>
<td>Hyperreflexia patellar (L2-4), enlarged reflex zone clonus at Achilles tendon (L2S2)</td>
</tr>
<tr>
<td>Upper extremities</td>
<td>Mildly dystonic as adult, movements bradykinetic and mildly ataxic</td>
<td>Worsening paralysis as adult</td>
<td>Ataxic, brisk reflexes C5-7, pyramidal sign, Juster positive, intention tremor, finger to nose dysmetria/ hypometria, mild tremor distally, pyramidal signs—Finger to nose dysmetria/ hypermetria, mild tremor distally, pyramidal signs—Finger to nose dysmetria/ hypermetria, mild tremor distally, pyramidal signs—</td>
</tr>
<tr>
<td>Lower extremities</td>
<td>Severe spastic paraplegia as adult, no voluntary leg movement, bilateral Babinski signs</td>
<td>Spasticity, bilateral Babinski, ankle clonus</td>
<td>Brisk reflexes C5-7, increased reflexes patellar (L2-L4) with enlarged reflex zone, no ankle clonus, but pyramidal reflexes L2-S2, Babinski and Rossolimo positive—Bilaterally</td>
</tr>
</tbody>
</table>

Continued
that influence alternative splicing in the region surrounding the patient mutations, we introduced a series of seven 10-base deletions into the splicing minigene construct (deletions 1–7; Fig. 3A) and tested them by transfection into Oli-neu cells. Deletion 5, which included position c.454-314 where the mutations in families B and C are located, caused a highly significant reduction in the PLP1/DM20 ratio to 0.02 (Fig. 3B), similar to that of the deletion of the entire E region, which was 0.07 (Fig. 1D). Unexpectedly, deletion 4, which included position c.454-322 where the mutation in family A is located, did not cause reduction of the PLP1/DM20 ratio. However, we noted sequence similarity between bases in regions 3 and 4 (Fig. 3C). If the sequence of region 4 must be next to region 5 for splicing to occur, then we may have compensated for the function of the sequence in region 4 by moving the similar sequence in region 3 next to region 5.

To test the hypothesis that sequences in regions 4 and 5 are responsible for maintaining the normal PLP1 isoform levels, a series of mutant constructs was made throughout this region (Fig. 4A) and tested by transfection into Oli-neu cells (Fig. 4B). In designing these constructs, special focus was placed on the c.454-322 and c.454-314 positions, as these were the sites of the patient mutations in our three families. Results indicated that the c.454-322G>A mutation identified in family A reduced the PLP1/DM20 ratio to 0.13 of normal (Fig. 4B), with similar results for non-patient mutations at the same position, c.454-322G>T and c.454-322G>C, at 0.12 of normal. The c.454-314T>A and c.454-314T>G mutations in families B and C, respectively, reduced the PLP1/DM20 ratio to 0.13 and 0.15 of normal, respectively. All of the two-base substitution mutations across regions 4 and 5 resulted in a significant decrease in PLP1/DM20 ratio suggesting that regions 4 and 5 were important for alternative splicing of PLP1, although the two-base substitution c.454-318_-317GA>TG reduced the PLP1/DM20 ratio to 0.26, the single-base mutation at position c.454-315 had a much less severe effect with a level of 0.90 of normal, suggesting the base c.454-316 to be more important than base c.454-315 for splicing function.

Bioinformatics analyses suggest that PLP1/DM20 alternative splicing is dependent on formation of a double-stranded RNA secondary structure between long-distance interacting elements

Mutations in introns can cause aberrant alternative splicing by creating or destroying splicing enhancer or silencer motifs that are binding sites for trans-acting factors that influence splice-site selection by interacting with the splicing machinery. For example, a mutation deep within an intron of the fibrinogen beta gene (FGB) creates an ESE that affects alternative splicing by activating the inclusion of a cryptic exon (31). We performed bioinformatics analyses to identify predicted modifications in potential splicing factor binding sites in regions 4 and 5 due to our introduced mutations and to determine whether the changes could explain our transfection data produced from mutant constructs (Figs 3, 4, and 5, Supplementary Material, Table S1). Modification of predicted factor-binding sites did not correlate consistently with the transfection data.
example, the c.454-314T>A mutation in family B and the c.454-314T>G mutation in family C had no effect on predicted ESE sites, but there was a significant reduction in the PLP1/DM20 ratio when the mutant constructs were tested by transfection. Additionally, two constructs, c.454-309T>C and c.454-308_-307CA>GG, were both predicted to reduce the strength of a SRp40 site, but there was no change in the PLP1/DM20 ratio for the c.454-309T>C construct, while the c.454-308_-307 CA>GG mutation reduced the PLP1/DM20 ratio to 0.37 of normal. Our data did not suggest a clear connection between the predicted site modifications and our transfection data, so we sought an alternative explanation for how regions 4 and 5 could affect alternative splicing.

Alternative splicing can be influenced by pre-mRNA secondary structures that form around splice sites (32–34). To search for potential secondary structure involvement in the splicing of intron 3, we analyzed the 1071-base intron 3 of PLP1 by using the mfold program with standard parameters to represent the stability of the most likely secondary structures (35). The results included 32 potential structures for the intron. The initial delta G values for the first three structures were −353.4, −352.9 and −351.8 kcal/mol. In comparison, another PLP1 intron of similar size, intron 6 with 1127 bases, had delta G-values averaging −338.3 kcal/mol for the first three structures, indicating that intron 3 structures had a lower free energy and were more likely to be stable than intron 6 structures. In viewing the third structure created by mfold for intron 3, we noted that bases c.454-323 to c.454-308, regions 4 and 5 (Fig. 3) that included the three patient mutations, formed a double-stranded RNA structure with bases c.453+151 to c.453+166 (Fig. 5A). The position of the secondary structure within the intron (Fig. 5B) and the bases involved (Fig. 5C) are shown. We will refer to bases c.454-323_-308 as long-distance interaction site 3′ (LDIS-3′) and to c.453+151+166 as long-distance interaction site 5′ (LDIS-5′).

The mfold data led to the hypothesis that formation of a double-stranded intra-intronic RNA structure between LDIS-5′ and LDIS-3′ functions to promote PLP1 splicing and mutations that reduce the stability of the predicted secondary structure would lead to reductions in the PLP1/DM20 ratio. We first tested constructs in which LDIS-5′ bases were mutated to the corresponding bases in LDIS-3′. We also tested the reciprocal constructs in which bases in LDIS-3′ were mutated to corresponding bases in LDIS-5′. In these constructs, base-pairing could not occur between identical bases so the double-stranded RNA structure would not be formed, and we predicted that the PLP1/DM20 ratio would be reduced. Further, we tested mutations in which we swapped bases between the two regions. In these constructs, base-pairing would occur, and we predicted that the normal PLP1/DM20 ratio would be reestablished.

When we mutated bases in LDIS-5′ so that base-pairing with LDIS-3′ would be disrupted, three of the four constructs (+153_+157AGAA>TTCCT, +162_+165GTGC>CACG and +156_+157AA>T) functioned as predicted by our hypothesis, reducing the relative PLP1/DM20 ratio to no more than 0.20 (Fig. 5D, black bars). The remaining construct, c.453+165C>G, had a PLP1/DM20 ratio reduced to 0.69 of normal. Since the mutation of c.454-322G (in family A), which is predicted to be the complement of the base at c.453+165C in the intra-intronic structure, had reduced the PLP1/DM20 ratio to 0.12 of normal, we expected a similar reduction in the ratio for the c.453+165C>G base switch. However, we found a plausible explanation for this unexpected result that fit with our hypothesis by further investigating the secondary structure as follows. Using mfold and a 55-base normal RNA test sequence comprising 20 bases from the LDIS-5′, we found the normal test sequence to have an initial delta G value of −17.6 kcal/mol (Supplementary
The three LDIS-5′ mutation constructs, which had reduced PLP1/DM20 ratios (+153_157AAG AA TTCTT, +162_165GTGC CACG and +156_157AA TT) had increased initial delta G values averaging 211.6 kcal/mol (Supplementary Material, Table S1), indicating a less-stable intra-intronic RNA structure. Likewise, in the secondary structure most likely to form with the c.453+165C G base change, the guanine base mutation does not participate in base-pairing; however, there is a shift in base-pairing so that the guanine at position c.454-322 now base-pairs with the cytosine at position c.453+166, so an adenine-uracil base-pair is added (Supplementary Material, Fig. S2B). The resulting initial delta G of 216.4 kcal/mol is relatively closer to the initial delta G of 217.6 kcal/mol for the normal test sequence indicating a partially stable structure, which correlates well with the intermediate transfection result for the c.453+165C G mutation, a PLP1/DM20 ratio of 0.69 of normal, further supporting our hypothesis.

When we separately mutated additional bases in LDIS-3′ so that base-pairing with the LDIS-5′ would be disrupted, all four constructs functioned as predicted, reducing the relative PLP1/DM20 ratio to 0.12 of normal or less (Fig. 5D, white bars with thin outlines). Finally, when we made ‘double mutants’ by swapping complementary bases between LDIS-5′ and LDIS-3′ of the proposed RNA structure, which was predicted to recreate the secondary stem structure depicted in Figure 5C, the relative PLP1/DM20 ratio was restored to nearly normal levels (ranging from 0.68 to 1.08), as anticipated (Fig. 5D, white bars with thick outlines). Together, these data demonstrate a long-distance interaction between LDIS-5′ and LDIS-3′ and support our hypothesis that formation of the intra-intronic secondary structure between the two sites acts to facilitate PLP1 splicing.

We further reasoned that if LDIS-5′ and LDIS-3′ are functionally important for PLP1/DM20 alternative splicing, the sequences and secondary structure would be conserved across mammalian species because the alternative splice of PLP1/DM20 exon 3 is found throughout the mammalian lineage. When we performed bioinformatics analysis of conservation of the sequences at LDIS-5′ across 25 mammalian species, we found that all 16 bases are identical in each of the 25 mammalian

---

**Figure 3.** The 10-base deletion that included base c.454-314, which is mutated in two families with PMD (deletion 5), caused a significant reduction in the PLP1/DM20 ratio, while the deletion that included base c.454-322 (deletion 4), which is mutated in a third family, did not. (A) The location of a series of seven 10-base deletions in a 70-base region of intron 3 that includes the three PMD patient mutations is depicted. The bar labeled E refers to the E region shown in Figure 1. Predicted SF2/ASF sites are indicated by thick solid bars above the normal sequence. Predicted SRp40 sites are indicated by thin solid bars below the normal sequence. (B) Quantification of PLP1/DM20 mRNA in Oli-neu cells transfected with the 10-base deletion constructs. Standard deviation for the normal construct was ±0.06. Methods of data collection and statistical testing are identical to those used in experiments for Figure 1. †††P < 0.001; ††P < 0.01; †P < 0.05; n.s., no significant difference. (C) Diagram comparing the 10 bases removed in deletions 3 and 4. The PMD family mutation at base c.454-322 is boxed. Seven of 10 bases are the same in these two segments, so deletion 4 effectively brings these seven bases into the same position as normal relative to the bases at region 5.
PLP1 genes examined (Fig. 6A). The conservation of LDIS-3′ was also very high, with only two sites of variability among species, bases c.454-318 and c.454-309 (Fig. 6B). As shown in Figure 5C, base c.454-318 is within the secondary structure, but not paired, so any base would be tolerated at this position. Base c.454-309 is a uracil in most species, but it is a cytosine in some. Our data showed that a cytosine at this position had no significant effect on the PLP1/DM20 ratio (Fig. 4B). Thus, the base-pairing in all 25 species is nearly identical (Fig. 6C).

Outside of the bases in LDIS-5′ and LDIS-3′, the conservation falls off. When we performed bioinformatics analyses of the secondary structures formed by intron 3 sequences across the 25 mammalian species, we found the LDIS-5′ and LDIS-3′ base-pairing structure in the computed foldings for all of them. It is in the third-most energetically favorable structure for humans, but in 13 of the 25 species it is found in the most energetically favorable structure. These data showing conservation of LDIS-5′ and LDIS-3′ sequences and secondary structure across mammalian species argue that the mechanism regulating the alternative splice of PLP1 is functionally conserved.

The strong evidence for the importance of the secondary structure in PLP1/DM20 alternative splicing prompted us to further examine some of the results in the experiments shown in Figures 3 and 4 by using mfold. The region 5 deletion in LDIS-3′ that had reduced the PLP1/DM20 ratio dramatically (Fig. 3B) also had a higher-than-normal initial delta G value of −6.3 kcal/mol, due to removal of seven double bonds in the structure (Supplementary Material, Fig. S2C). However, the
deletion of region 4 in LDIS-3 had failed to reduce the PLP1/DM20 ratio despite containing base c.454-322 that caused reduction of the PLP1/DM20 ratio when mutated in family A. We previously hypothesized that the function of the sequence in region 4 may have been compensated by moving the similar sequence in region 3 next to region 5 (Fig. 3C). Using mfold analysis, we saw that because of the similar sequence next to the deleted region, deletion of region 4 resulted in a predicted secondary structure with an initial delta G value of -351.8 kcal/mol. The boxed-in gray region shows one of the secondary structures that can form between bases c.453 + 151_-166 within region A (termed LDIS-5’) and c.454-323_-308 within region E (termed LDIS-3’), where the three patient mutations are located. The structure shown does not depict the formation of the splicing lariat between the donor and branch-point of the acceptor site. (B) Linear depiction of intron 3 of the PLP1 pre-mRNA showing the locations of LDIS-5’ and LDIS-3. Regions A and E are marked by bars. (C) The secondary structure between LDIS-5’ (on top) and LDIS-3’ (on the bottom) is shown at the sequence level. These complementary regions are separated by 581 bases. Vertical lines indicate strength of hydrogen bonding: black and thick, triple hydrogen bond between guanine and cytosine; dark gray and intermediate thickness, double hydrogen bond between adenine and uracil; light gray and thin, single hydrogen bond between guanine and uracil. (D) Quantification of the PLP1/DM20 ratio in Oli-neu cells transfected with constructs containing mutations that affect the predicted secondary structure between LDIS-5’ and LDIS-3’. On the X-axis, the nomenclature for the mutations is based on the DNA sequence. Bases are numbered as shown in (C). The solid bars are constructs with mutations in LDIS-5’, c.453+151_-166. The adjacent open bars with thin borders are constructs in which the mutation is in LDIS-3’, c.454-323_-308. Each pair of solid bars and open thin-bordered bars disrupts the same base pairing(s) in the predicted intra-intronic structure shown in (C). The open bars with thick borders are constructs in which there are two complementary mutations, reestablishing the predicted base-pairing structure. The standard deviation for the normal construct was ±0.02. Methods of data collection and graphic representation are identical to those used in Figure 1.
constructs, we found a strong linear relationship (Supplementary Material, Fig. S3A). All the constructs listed in Supplementary Material, Table S1 were included in this analysis. Structures that had an initial \( \Delta G \) value more positive than \(-14\) kcal/mol consistently had \( PLP1/DM20 \) ratios below 0.25 of normal \( \left[ \log_{10}(-0.6) \right] \). Additionally, this plot suggests that a threshold \( \Delta G \) value of \(-16.4\) kcal/mol or lower indicates a construct with sufficient base-pairing interactions for formation of an RNA structure adequate for near-normal levels of \( PLP1/DM20 \) splicing.

We used our model of \( PLP1 \) intron 3 secondary structure and the mfold program to predict which bases are essential for the formation of the intra-intronic structure and therefore which mutations would result in an altered \( PLP1/DM20 \) ratio. We individually substituted each base in the structure with all possible mutations and recorded the free energy values (data not shown). For example, a change of c.453+150A to G, U or C resulted in a delta \( G \) value of \(-17.5\) kcal/mol in mfold and was predicted not to change the \( PLP1/DM20 \) ratio. However, the change of c.453+151G to another base increased the delta \( G \) value to an average of \(-14.1\) kcal/mol and was predicted to be likely to reduce the \( PLP1/DM20 \) ratio. Only the central non-pairing bases c.453+158 and c.454-315 are unessential for the formation of the secondary structure (Supplementary Material, Fig. S3B). Any single base change at this location is predicted to be unlikely to alter the \( PLP1/DM20 \) ratio, and indeed, we found that c.453-315T>G had a \( PLP1/DM20 \) ratio at 0.90 of normal levels (Fig. 4B). These data are consistent with our hypothesis that the formation of secondary structure in intron 3 is required for normal regulation of the \( PLP1/DM20 \) alternative splice.

**DISCUSSION**

Mutations that dysregulate \( PLP1/DM20 \) alternative splicing by changing invariant sequences at splice donor or splice acceptor sites have been relatively simple to classify as likely causative mutations for PMD (13–16,21). Further, mutations in both ESEs and ISEs near the donor and acceptor sites are likely causative mutations for PMD by loss of binding of splicing factors to the mutant sequences (14,15,20,23,27–29). Because of the importance of the \( PLP1/DM20 \) alternative splice to expression of the gene and its sensitivity to mutation, we postulated that regulation of \( PLP1/DM20 \) alternative splicing could be controlled by elements acting from a distance into intron 3 of \( PLP1 \) and that there could be PMD patient mutations in these elements that are not typically sequenced by molecular diagnostics laboratories, but that are responsible for pathogenicity in some of the patients in whom \( PLP1 \) mutations have not been found. In this work, we identified regulatory sequences throughout intron 3 and a strong regulatory region, \( LDIS-3' \), \(~70\%) of the way into the intron that affects the \( PLP1/DM20 \) ratio. Bioinformatics analysis suggested that this region was capable of participating in a secondary structure with a site \(~15\%) of the way into the intron, \( LDIS-5' \). In domain swap experiments, we demonstrated that mutations introduced into either strand of the double-stranded RNA helix disrupted the secondary structure and decreased the alternatively spliced ratio of \( PLP1 \) to \( DM20 \), while reciprocal mutations that would stabilize the double-stranded RNA secondary structure restored the \( PLP1 \) to \( DM20 \) ratio to normal levels. In addition, we sequenced intron 3 in its entirety in patients with clinical diagnoses of PMD and strong suggestion of X-linked inheritance, who had tested negative for \( PLP1 \) duplication and DNA sequence mutations in molecular diagnostics laboratories. We report patients in three families with mutations in \( LDIS-3' \) deep within intron 3 that reduced the \( PLP1/DM20 \) ratio in transfection experiments. Our data show that \( LDIS-5' \) and \( LDIS-3' \) interact to regulate the \( PLP1/DM20 \) alternative splice and strongly suggest that the interaction between them is the formation of secondary structure within intron 3. As diagrammed in Figure 7, the secondary structure between \( LDIS-5' \) and \( LDIS-3' \) favors the \( PLP1 \) form, while disruption of the secondary structure favors the \( DM20 \) form. Our data also show that the patient mutations can cause PMD by disrupting the interaction, most likely by disrupting secondary structure. Our findings advocate for molecular diagnostics laboratories to include intron 3 when sequencing \( PLP1 \) in PMD patients and further suggest that introns involved in alternative splicing of other disease genes should be checked when disease-causing mutations cannot be found in coding sequences.

Often, genes have multiple sites that are recognized by prediction programs as splice donor sites. According to the oriented scanning mechanism (36), which is the preferred mechanism underlying splice donor site selection, the 5’-most splice donor site is used most often because it is the first to be encountered by the splicing machinery. In this regard, the 5’-most site of
the PLP1/DM20 pre-mRNA is DM20, and it seems to be the default splice site, as it is the site that is used in all places where PLP1 is expressed except for the adult CNS (e.g. the peripheral nervous system, heart and early in development of the CNS) (17,37,38). When PLP1 expression is upregulated in the oligodendrocytes of the CNS during development, there is a switch to use of the PLP1 splice donor site 105 bases downstream of the DM20 site. Little is known about the mechanism governing that shift, which results in alternative splicing of PLP1 pre-mRNA. The previously described ESEs and ISEs in PLP1 may play a role. We performed bioinformatics analyses to determine further factor-binding sites in a region of intron 3 that affects the PLP1/DM20 ratio and compared the results with our transfection data, but we did not clearly identify functional factor-binding sites in the LDIS-3′ region (Supplementary Material, Table S1). However, it is well known that these bioinformatics prediction analyses can give false results, identifying sites that are actually not present and missing sites that are present. We did not try further to validate predicted sites, and we did not search for binding sites that could have been missed. Splicing factors may play a role in alternative splicing of PLP1 by binding the secondary structure or the single-stranded RNA and interacting with the splicing machinery at the PLP1 and/or DM20 splice site. Alternatively, the secondary structure may play a role in promoting the use of the PLP1 splice donor site by bringing it into proximity with the splice acceptor site at exon 4, with destabilization of this structure shifting splicing back toward the default DM20 site.

Although the interaction of splicing factors with regulatory sequences in RNA is currently thought to play a larger role in splicing regulation than secondary structure, there are several reports of regulation by secondary structure in the literature, and various mechanisms have been proposed (39,40). Stem-loop structures are predicted to play roles in alternative splicing of genes such as human growth hormone (GH1), survival of motor neuron 2 (SMN2), microtubule-associated protein tau (MAPT) and troponin T, type 2 (TNNT2) (41–44). In some cases, specific splicing factors are known to play a role by binding to the sequences involved in the structure. For example, in the case of MAPT, an RNA helicase has been demonstrated to interact with a stem-loop structure and induce splicing of tau exon 10 (45). In the case of cTNA, the splicing factors muscle-blind (MBNL1) and U2AF65 regulate alternative splicing by binding double- and single-stranded RNA, respectively (44). In the ENAH gene, the secondary structure functions as a bridge that positions a distal RBFOX splicing factor binding site close to alternative exon 11a, increasing inclusion of the exon (46). In other cases, secondary structures are found to overlap or be in close proximity to splice donor or acceptor sites, facilitating binding of splicing factors (34,47). In the case of Lepidoptera 14-3-3 zeta pre-mRNAs, inclusion of mutually exclusive exons is largely regulated by RNA structure but is adjusted by splicing factor recognition (48). The mechanism by which the secondary structure in intron 3 of PLP1 regulates splicing is currently unknown. Further studies will be needed to determine whether the structure itself regulates the alternative splice or whether splicing factors that bind the structure or positions near the structure are involved. While there is a report of the importance of a long-distance interaction in regulation of alternative splicing of SMN2 that is involved in the genetic disease...
spinal muscular atrophy (49), to our knowledge, our article is the first to report disease-causing mutations that dysregulate alternative splicing by disruption of a long-distance interaction within an intron, probably by destabilization of secondary structure.

MATERIALS AND METHODS

Preparation of minigene splicing constructs for transfection

An artificial minigene construct containing the PLP1 genomic sequences from base c.123 within exon 2 to c.468 of exon 4 inserted into the neomycin gene under control of the Rous sarcoma virus (RSV) promoter has been described (Fig. 1B) (13). Two bases in the original minigene construct that were found to differ from the reference sequence in intron 2 (at c.192-147A>G previously reported and at c.191+218_219insT) were corrected to match the reference genome sequence by site-directed mutagenesis using the QuikChange Multi Site-directed Mutagenesis Kit (Agilent Technologies, Inc., Santa Clara, CA, USA) with primers 12cor1 and 12cor2 (Supplementary Material, Table S2) according to the manufacturer’s instructions (13). Resulting plasmids were confirmed to be identical to the reference sequence (Z73964) by sequence analysis. The corrected minigene construct, designated pRSVneoPLP1-NRL, was mutated as described below.

Six constructs containing overlapping 210-base deletions of intron 3 were created. A SalI site was positioned at the 5’ end of the regions targeted for deletion, and an XhoI site was placed at the 3’ end using QuikChange Multi Site-Directed mutagenesis and mutagenesis primers [Integrated DNA Technologies (IDT), Coralville, IA, USA] (Supplementary Material, Table S2). Each plasmid was digested with SalI and XhoI to remove an intron portion and re-ligated. The constructs were designed so that the natural donor, branch-point and acceptor sites were not modified. Additionally, six constructs were made in which a replacement sequence of 210 bases was used to replace the deleted sequence. The replacement was a randomly generated sequence, from which any potential exonic splicing enhancers (ESEs) or splice sites, identified by bioinformatics analysis, were removed (Supplementary Material, Table S3). The double-stranded replacement sequence was created by hybridizing 11 overlapping oligonucleotide primers (listed in Supplementary Material, Table S3), ligating, cloning into the pPCR-Script vector, then digesting with SalI and XhoI. The resulting SalI–XhoI fragment was ligated into each of the SalI–XhoI-digested deletion constructs.

Constructs containing 10-base deletions, or substitutions of one to five bases within intron 3, were made with the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies, Inc.). The sequence of the forward mutagenic oligonucleotide primer (IDT and Eurofins MWG Operon, Huntsville, AL, USA) used to make each construct are indicated (10-base deletions, Supplementary Material, Table S4; 1–5-base substitutions, Supplementary Material, Table S5). The complementary primers used to make each construct are not shown.

Cell culture and transfections

Oli-neu cells were maintained in DMEM/N1/1% horse serum growth medium containing 2 mM l-glutamine (Invitrogen, Foster City, CA, USA), 10 μg/ml biotin (B4639, Sigma-Aldrich, St Louis, MO, USA), 2.5 μg/ml insulin (I5500, Sigma-Aldrich), 5 μg/ml apo-transferrin (T2036, Sigma-Aldrich), 100 μM putrescine dihydrochloride (P7505, Sigma-Aldrich), 20 μM progesterone (P0130, Sigma-Aldrich) and 0.03 μM sodium selenite (SS261, Sigma-Aldrich) on lysine-coated flasks (BD Biosciences, San Jose, CA, USA) (50). Transfections of normal and mutant constructs into Oli-neu cells were performed in triplicate. Plasmid DNA was prepared using EndoFree Plasmid Maxi Kit (Qiagen, Valencia, CA, USA). Briefly, 0.6 × 10^6 cells were plated in each well of a lysine-coated six-well plate on day 1. On day 4, transfections were performed with plasmid DNA (0.5 μg) and Lipofectamine 2000 (Invitrogen) according to the manufacturer’s protocol. On day 5, the medium was removed from the cells, and in-well lysis was performed. Adhering cells were dislodged with a rubber policeman and collected, and the resulting suspension was vortexed and stored or immediately processed for RNA isolation.

RNA isolation and reverse transcription-polymerase chain reaction (RT-PCR)

RNA was isolated from Oli-neu cells using the RNeasy Mini kit (Qiagen) and QiShredders (Qiagen) according to the manufacturer’s instructions for purification of total RNA from animal cells. Residual DNA was removed using Turbo DNA-free treatment (Life Technologies/Ambion, Austin, TX, USA). RNA was quantified with a NanoDrop ND-1000 spectrophotometer (Thermo Scientific, Waltham, MA, USA). The cDNAs were produced by first denaturing RNA at 65°C for 5 min, then performing reverse transcription in a reaction mixture containing 300 ng of DNase-treated RNA, 4 U Omniscript Reverse Transcriptase (Qiagen), 1× Buffer RT (Qiagen), 0.5 mM each dNTP (Qiagen), 1 μM Oligo (dT) 15 primer (Promega, Madison, WI, USA) and 10 U rRNasin (Promega) which was incubated for 60 min at 37°C.

PCR reactions were performed on the cDNA with Multiplex PCR kit (Qiagen) using 2 μl of the reverse transcription reaction in a 25 μl volume containing 0.2 μM of each primer, and 12.5 μl of Multiplex Master Mix (Qiagen). The reverse primer of each pair was labeled with 6-FAM. The reaction conditions were 15 min at 95°C; 25 cycles of 30 s at 94°C, 90 s at 60°C and 90 s at 72°C; and 10 min at 72°C. Primers specific to the minigene construct, neo-ex2F/neox-ex4R (Supplementary Material, Table S6) (IDT and Eurofins MWG Operon, were designed to overlap the neo-exon2 and exon4-neo junctions (Fig. 1B). These primers were then unable to amplify an endogenous neo-cassette in the Oli-neu cells or the endogenous mouse Plp1 (50). Amplification with primer pair neo-ex2F/neox-ex4R-6-FAM resulted in a 262 bp DM20 product and a 367 bp PLP1 product (including exon 3B). Dilutions of the multiplex ranging from 1:3 to 1:200 were subjected to capillary electrophoresis on a 3130xL Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) along with a GS500 LIZ size standard. Data were analyzed using Peak Scanner Software v. 1.0 (Applied Biosystems). The ratio of PLP1 to DM20 was calculated by dividing the PLP1 signal by the DM20 signal resulting from the PCR with the neo-ex2F/neox-ex4R primers.

Human subjects and genetic testing

The study of human subjects was approved by the Institutional Review Board at Alfred I. duPont Hospital for Children,
Wilmington, DE, and the Ethics Committee of the University Hospital Motol, Prague, Czech Republic, and informed consent was obtained as appropriate. DNA testing for PMD was performed in the Molecular Diagnostics Laboratory at the Alfred I. duPont Hospital for Children (families A and B) and the DNA Laboratory of the Department of Child Neurology, 2nd School of Medicine, Charles University and University Hospital Motol, Prague (family C). Table 1 summarizes the clinical presentations for patients with intron 3 mutations from families A, B, and C. Neurological examinations were performed by authors J.G. for family A, B.C. for family B and P.S. for family C. The clinical presentation in all three families is mild classical PMD.

The PLP1 copy number change (duplication) and coding-region mutations, and GJC2 (formerly GLA12) mutations for families A and B were ruled out by diagnostic testing in the Molecular Diagnostics Laboratory at the Alfred I. duPont Hospital for Children and for family C in the DNA Laboratory of Child Neurology, 2nd School of Medicine, Charles University and University Hospital Motol, Prague. Linkage analysis for PLP1 was completed by PCR amplification of a microsatellite marker consisting of dinucleotide CA repeats in intron 1 with primers CA-PLP1 and CA-PLP2 (30) or CA-PLP4 (Supplementary Material, Table S6). Mutations in PLP1 were detected or confirmed by DNA sequence analysis in the Molecular Diagnostics Laboratory at the Alfred I. duPont Hospital for Children or in the DNA Laboratory of the Department of Child Neurology, 2nd School of Medicine, Charles University and University Hospital Motol, Prague. Patient mutations have been submitted to the LOVD X-chromosome gene database (http://grenada.lumc.nl/LOVD2/MR/home.php, last accessed on 5 June 2014).

Skin fibroblast cultures
Skin fibroblast cultures were obtained by tissue explant as previously described (51). RNA isolation and cDNA production from cultured skin fibroblasts was performed as described for Oli-neu cells earlier in this section. For analysis of PLP1 splicing in fibroblasts, the primer pairs were hGAPDH F/R (14), PLPF4/5R2 and PLPF3/5R2 (Supplementary Material, Table S6). Mutations in PLP1 were confirmed by DNA sequence analysis in the Molecular Diagnostics Laboratory at the Alfred I. duPont Hospital for Children or in the DNA Laboratory of the Department of Child Neurology, 2nd School of Medicine, Charles University and University Hospital Motol, Prague. Patient mutations have been submitted to the LOVD X-chromosome gene database (http://grenada.lumc.nl/LOVD2/MR/home.php, last accessed on 5 June 2014).

Bioinformatics analyses
In silico searches for cis-acting splicing elements were performed using RESCUE-ESE (http://genes.mit.edu/burgelab/rescue-eselast accessed on 5 June 2014), ESEfinder 2.0 (http://rulai.cshl.edu/tools/ESE2/last accessed on 5 June 2014), ESEfinder 3.0 (http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder. cg?process=home, last accessed on 5 June 2014) and Automated Splice Site and Exon Definition Analysis (http://splice.uwo.ca/last accessed on 5 June 2014) (252–55). Donor and acceptor sites were investigated using Analyzer Splice Tool (http://ibis.tau.ac.il/ssat/SpliceSiteFrame.htm, last accessed on 5 June 2014) (56). Analysis of the RNA structure within intron 3 from PLP1 genes of 25 mammalian species (Supplementary Material, Table S7) was investigated with the mfold web server (http://mfold.rit.albany.edu/?q=mfold, last accessed on 5 June 2014) using default settings (35). Note that while all mutations were described at the DNA level, predicted secondary structure was described at the RNA level, substituting uracil for thymidine.

Intron 3 sequences from the PLP1 genes across 25 mammalian species (Supplementary Material, Table S7) were formatted into FASTA with EMBoss Script and compared using CLUSTAL W and CLUSTAL W PROF from SDSC Biology WorkBench (http://workbench.sdsc.edu, last accessed on 5 June 2014), and sequence logos were developed using WebLogo (http://weblogo.berkeley.edu, last accessed on 5 June 2014) (57–60).

SUPPLEMENTARY MATERIAL
Supplementary Material is available at HMG online.

ACKNOWLEDGEMENTS
We are thankful to the three families for their participation in our studies. We thank Selena Driscoll, Kaitlin McLean and Elisabet V. Eppes for their assistance in preparing and testing constructs, Angelique Davis-Williams for her diagnostic expertise, and Jacqueline Trotter for the gift of Oli-neu cells. We also thank the Department of Family Medicine of Cebu Velez General Hospital–Cebu Institute of Medicine and their interns for their assistance in the data and sample collections for family B. We acknowledge the Nemours Biomolecular Core and the Nemours Molecular Diagnostics Laboratory.

Conflict of Interest statement: G.M.H. receives pay and the Nemours Biomedical Research department derives revenue for diagnostic testing from the Molecular Diagnostics Laboratory, Alfred I. duPont Hospital for Children.

FUNDING
This work was supported by National Institutes of Health (P20GM103464 and R01NS058978 to G.M.H.); the Pelizaeus-Merzbecher Disease Foundation (to G.M.H.); Internal Grant Agency of Ministry of Health of the Czech Republic (NT14348 to P.S.); and the Project for Conceptual Development of Research Organization (00064203 to P.S.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the granting agencies and foundations.

REFERENCES


