Published in final edited form as:

J Neurochem. 2010 November ; 115(3): 614–624. doi:10.1111/j.1471-4159.2010.06972.x.

Expression and regulation of a low density lipoprotein receptor exon 12 splice variant

I-Fang Ling, Rangaraj K. Gopalraj, James F. Simpson, and Steven Estus*

Department of Physiology, Sanders-Brown Center on Aging, University of Kentucky, Lexington, KY

Abstract

Since low density lipoprotein receptor (LDLR) contributes to cholesterol and amyloid beta homeostasis, insights into LDLR regulation may facilitate our understanding of cardiovascular disease and Alzheimer's disease (AD). Previously, we identified LDLR isoforms that lacked exon 12 or exons 11–12 and that are predicted to encode soluble, dominant negative, LDLR. Moreover, these isoforms were associated with rs688, an exon 12 polymorphism that was associated with LDL-cholesterol and AD risk. Here, we present evidence that while the truncated LDLR isoforms are translated *in vitro*, they represent < 0.1% of CSF proteins. Since these LDLR isoforms likely represent a loss of mRNA encoding functional LDLR, we then focused upon identifying intron-exon boundary and exonic splicing enhancer elements critical to splicing. Exon 12 inclusion is enhanced by altering the 5' splice site in intron 12 towards a consensus splice donor sequence, consistent with its being a weak 5' splice site. Additionally, of the nine evolutionarily conserved putative splicing enhancer regions within exon 12, two regions that flank rs688 were critical to exon 12 inclusion. Overall, these results suggest that LDLR splice variants represent a loss of mRNA encoding functional LDLR and provide insights into the regulatory elements critical for LDLR exon 12 splicing.

Keywords

Low density lipoprotein receptor; Alzheimer's Disease; RNA splicing; cholesterol; splicing regulatory element

Introduction

Variation in the efficiency of exon inclusion is emerging as a frequent mechanism underlying the actions of genetic polymorphisms that modulate disease risk. Efficient exon splicing requires consensus sequences at the 5' splice site, at the branch point and subsequent polypyrimidine tract and at the 3' splice site (Hertel 2008). When an RNA sequence varies from these consensus sequences, elements at other sites are required to achieve efficient splicing. Examples of these other sequences include exonic splicing enhancers (ESEs) which are conventionally recognized by members of the serine/arginine-rich (SR) protein family that enhance exon inclusion by recruiting splicing machinery to the exon/intron boundary. Splicing is also modulated by exonic splicing silencer (ESS) sequences, which bind heterogeneous nuclear ribonucleoproteins (hnRNPs) and typically inhibit splicing. Intronic splicing enhancer (ISE) and intronic splicing silencer (ISS)

*To whom correspondence should be addressed. Steven Estus, Phone: (859) 323-3985, ext. 264, Fax: (859) 323-2866, steve.estus@uky.edu, 800 S. Limestone St., Lexington, KY 40536-0230.

elements have also been identified although their mechanisms are not as well understood (reviewed in (Hastings & Krainer 2001, Black 2003)).

Changes in DNA sequence that disrupt functional RNA splicing regulatory elements can cause aberrant splicing and human disease (reviewed in (Cooper *et al.* 2009, Tazi *et al.* 2009)). For example, mutations that disrupt an ESE within exon 10 of microtubule-associated protein tau result in an increased frequency of exon 10 skipping that causes frontotemporal dementia with Parkinsonism linked to chromosome 17 (D'Souza & Schellenberg 2006). Empirically-based programs have been developed to identify ESE sequences *in silico*, e.g., ESEfinder (<http://rulai.cshl.edu/tools/ESE2/>) (Cartegni *et al.* 2003) and RESCUE-ESE (<http://genes.mit.edu/burgelab/rescue-ese/>) (Fairbrother *et al.* 2002, Fairbrother *et al.* 2004). Although there are other mechanisms of splicing regulation, ESEs are prevalent and these programs provide a useful method to predict functional ESEs.

Low density lipoprotein receptor (LDLR) plays a central role in cholesterol homeostasis (reviewed in (Hobbs *et al.* 1992)), is a major apolipoprotein E (APOE) receptor in brain (Cao *et al.* 2006, Fryer *et al.* 2005), and was recently implicated in amyloid- β (A β) homeostasis (Kim *et al.* 2009). Since cholesterol, APOE and A β may each contribute to AD, LDLR variants may modulate AD risk. Our lab previously identified a single nucleotide polymorphism (SNP), rs688, that modulates LDLR exon 12 splicing efficiency, i.e., the rs688T allele increases the proportion of LDLR isoforms lacking exon 12 or exons 11–12, compared to the major rs688C allele in minigene transfected cells. A similar effect was observed in human female liver and male brain tissue. Additionally, rs688T is associated with increased plasma LDL-cholesterol in pre-menopausal women (Zhu *et al.* 2007) and with increased AD risk in men (Zou *et al.* 2008), suggesting a connection between an LDLR SNP, splicing variants and disease status. The LDLR isoforms lacking exon 12 or exons 11–12 are predicted to produce truncated, soluble forms of LDLR (Zhu *et al.* 2007). However, whether these truncated LDLR proteins are translated and the critical factors necessary for exon 12 splicing are not known. Here, we present evidence that (i) these truncated LDLR proteins are translated within cells but are not detectable in CSF, (ii) the splice donor sequence in intron 12 contributes to exon 12 inefficient splicing, and (iii) two regions within exon 12 are critical for its inclusion within LDLR; these two splicing regulatory elements flank rs688.

Materials and Methods

Cell Culture

HepG2 (human hepatocellular carcinoma) and HEK293 (human embryonic kidney) cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum, 50 U/ml penicillin and 50 μ g/ml streptomycin at 37°C in a humidified 5% CO₂ - 95% air atmosphere.

Minigene Generation

Expression plasmids encoding the LDLR isoform lacking exon 12 (Delta 12 LDLR) or exons 11 and 12 (Delta 11+12 LDLR) were generated from HepG2 cell cDNA. LDLR exons 1–14 were amplified by using PCR (Platinum Taq, Invitrogen, Carlsbad, CA) with primers LDLR Exon1F 5'TAGGACACAGCAGGTCGTGA^{3'} and LDLR Exon14R 5'CAGGGAGGCACAGATACTGG^{3'}. PCR fragments were cloned into pcDNA3.1/V5-His-TOPO T/A cloning vector (Invitrogen). Clones lacking exon 12 or exons 11 and 12 were detected by PCR screening and clone integrity confirmed by sequencing.

An LDLR minigene containing exons 9–14 with rs688C or rs688T in a pcDNA3.1 backbone was described previously (Zhu *et al.* 2007). We also created an additional LDLR minigene

in pSPL3b (kind gift of Genzyme Corporation, Cambridge, MA) that contained LDLR exon 12 along with portions of its intronic flanking regions, i.e., 219 bp of 5' intronic sequence and 123 bp of 3' intronic sequence. This minigene was generated by using PCR to amplify with primers EcoRI-Intron11F $5'$ GCGGAATTCTGAAGTTTTTCTGACCTGCA $3'$ and BamHI-Intron12R $5'$ CGCGGATCCATAACTCAGGTCTAAGACCT $3'$. PCR fragments were cloned into pCR2.1 TOPO T/A cloning vector (Invitrogen). The clones and the splicing vector pSPL3b were digested with *EcoRI/BamHI* and, after gel-purification, the LDLR fragments were ligated into the pSPL3b vector which has been used extensively for splicing studies; the pSPL3b backbone contains rabbit β -globin exons as splice donor (SD) and splice acceptor (SA) exons separated by a portion of the intron of the HIV-tat gene (Buckler *et al.* 1991, Burn *et al.* 1995). Integrity of the LDLR exon 12 minigenes was confirmed by sequencing.

Antibody Generation and Purification

Rabbit antisera were raised against antigens corresponding to sequence near the carboxyl terminus of Delta-12 LDLR (GQSILDRYHQRSHFQC) or Delta 11+12 LDLR ((GC)-FGQISSTKPFVSPTASQV)); to enhance antigenicity, the peptides were conjugated to keyhole limpet hemocyanin through their cysteine residues (Bio-Synthesis Inc., Lewisville, TX). Antibodies that bound their respective peptide were purified from the antiserum by using Zymed's Rapid Affinity Purification (RAP) Kit as directed by the manufacturer (Invitrogen). Briefly, peptides were immobilized to a RAP coupling gel column via their cysteine residue. The column was loaded with antiserum, washed extensively with 10 mM PBS, and bound antibodies eluted with 10 mM glycine (pH 2.5). The purified antibody solution was pH neutralized and dialyzed against PBS.

Western Blot

HEK293 cells were transfected with the indicated LDLR expression plasmid by using FuGENE 6 as directed by the manufacturer (Roche Applied Sciences, Indianapolis, IN) and selected with G-418 till stably transfected. When 70% confluent, transfected cells or non-transfected control cells were washed with 5 ml of room temperature phosphate buffered saline (PBS) and lysed in 1 ml of RIPA buffer (50 mM Tris, pH 8.0, 150 mM NaCl, 1% NP-40, 0.5% deoxycholic acid, 0.1% SDS) containing 1X protease inhibitor cocktail (Roche Applied Science) for 30 minutes on ice with occasional rocking. A cell scraper was used to collect the cell lysate which was then subjected to centrifugation at 10,000 $\times$ g for 10 minutes at 4°C. Forty μ l of supernatant was mixed with 10 μ l of 5X SDS sample loading buffer containing β -mercaptoethanol, boiled for 5 minutes and subjected to polyacrylamide gel electrophoresis (PAGE) on a 7.5% or 4–20% gradient polyacrylamide gel (Bio-Rad). Proteins were transferred to nitrocellulose membranes (Bio-Rad). The blots were then incubated with Block Ace (AbD Serotec, Raleigh, NC) containing 0.1% Tween-20 for 1 hour at room temperature and probed overnight with chicken anti-LDLR antibody (1:1000 dilution, Chemicon, Temecula, CA), purified rabbit anti-Delta 11+12 LDLR antibody (1:1000 dilution) or rabbit anti-Delta 12 antibody (1:1000 dilution) at 4°C. After washing with 0.1% Tween-20 in Block Ace four times for 5 minutes each, the blots were incubated with peroxidase-conjugated donkey anti-chicken antibody (1:20,000 dilution, Jackson ImmunoResearch, West Grove, PA) or goat anti-rabbit antibody (1:10,000 dilution, Jackson ImmunoResearch) for 1 hour at room temperature. Bound peroxidase was visualized by using a SuperSignal West Pico kit (Pierce, Rockford, IL) and Hyperfilm (Amersham Biosciences, Piscataway, NJ). The blot with chicken anti-LDLR antibody was stripped with Re-Blot Plus Strong Solution (Chemicon) and reprobed with rabbit anti-LDLR antibody (1:200 dilution, Research Diagnostics, Concord, MA); bound antibody was then detected as described above.

To evaluate soluble LDLR in medium, 3×10^5 HepG2 cells per well were seeded in a 6-well plate in 2 ml of medium. The next day, medium was changed to DMEM supplemented with 10 $\mu\text{g/ml}$ insulin with or without 100 ng/ml 4 β -phorbol 12-myristate 13-acetate (PMA) and the cells returned to the incubator. After 4 hours, medium was collected and centrifuged at 10,000 $\times g$ for 10 minutes at 4°C. The supernatant was mixed with 200 μl of 0.15% deoxycholate, incubated at room temperature for 10 minutes, and the proteins were precipitated by adding 200 μl of 72% trichloroacetic acid and incubated on ice for 20 minutes. After centrifugation at 16,000 $\times g$ for 15 minutes at 4°C, the pellet was washed with 1 ml of acetone three times, air-dried, dissolved in 20 μl of ddH₂O, and subjected to PAGE analysis. Stably transfected HEK293 cells with Delta 11+12 or Delta 12 LDLR expression plasmid were grown in Opti-MEM for 48 hours. The conditioned medium was collected, supplemented with 1X protease inhibitor cocktail (Roche Applied Science) and concentrated by using iCON Concentrators 20K (Pierce). Forty μl of concentrated conditioned medium was subjected to PAGE analysis. Human CSF samples containing the indicated amounts of protein were analyzed in parallel. The sensitivity of this Western blot assay was evaluated by using known amounts of recombinant human LDLR (R&D Systems, Minneapolis, MN). Western blotting was performed as described above with chicken anti-LDLR antibody (1:1000 dilution, Chemicon). Where indicated, blots were stripped and reprobed with rabbit anti-Clusterin antibody (1:200 dilution, Santa Cruz Biotechnology, Santa Cruz, CA).

Evaluation of LDLR Minigene Splicing

LDLR minigene splicing efficiency was evaluated by transfecting the clones into HepG2 cells by using FuGENE 6 (Roche Applied Sciences). Briefly, 1.5×10^5 cells per well were seeded in a 6-well plate in 2 ml of medium without antibiotics one day before performing the transfection. The next day, 2 μg of LDLR minigene were mixed with 6 μl of FuGENE 6 reagent in 94 μl of Opti-MEM (Invitrogen) and added to each cell culture. Twenty-four hours after transfection, mRNA was isolated and analyzed for LDLR splicing patterns by reverse transcriptase-PCR (RT-PCR) as previously described (Zhu et al. 2007). RNA was converted to cDNA (SuperScript III, Invitrogen) and sequences corresponding to LDLR minigene splice products were PCR amplified (Platinum Taq, Invitrogen). For the LDLR exon 9–14 minigene, PCR was performed with an LDLR exon 10 sense primer 5'CATCGTGGTGGATCCTGTTC3' and, to obviate endogenous LDLR, a vector-specific antisense primer 5'GGGATAGGCTTACCTTCGAA3'. For the LDLR exon 12 minigene, PCR was performed with primers corresponding to the 5' and 3' β -globin exons in pSPL3b, i.e., 5'TCTCAGTCACCTGGACAACC3' and 5'CCACACCAGCCACCACCTTCT3', respectively. The PCR profiles consisted of initial denaturation at 94°C for 4 min, followed by cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 45 s, and final extension at 72°C for 7 min. The minimal number of PCR cycles necessary to discern products was performed, i.e. 23 cycles. PCR products were separated by polyacrylamide gel electrophoresis and visualized by SYBR-gold fluorescence on a fluorescence imager (Fuji FLA-2000). PCR product identities were determined by gel purification and direct sequencing (Davis Sequencing). The amount of full-length (FL) and inefficiently spliced LDLR isoforms was quantified by fluorescence intensity as described previously (Zou et al. 2008). Briefly, for each sample, fluorescence values were corrected for background and normalized for length differences among amplicons. The percentage of each isoform was calculated by the amount of the particular isoform divided by the total LDLR PCR product for that sample. Splicing efficiency refers to the percentage of the full-length transcript.

Mutagenesis of Converted Putative ESE

Sequences of LDLR exon 12 across human, macaque, mouse, hamster, cow, wild boar, rat, and rabbit were analyzed by ESEfinder (<http://rulai.cshl.edu/tools/ESE2/>) (Cartegni et al.

2003). If a putative ESE as defined by ESEfinder was conserved between more than five species, we considered it a conserved putative ESE. Conserved putative ESEs that overlapped one another were considered to be a conserved putative ESE region. To diminish the binding affinity of these conserved putative ESEs to SR proteins, mutagenesis primers were designed for each putative ESE that were predicted to reduce the original binding affinity (Table 1). To verify *in silico* that these new sequences neutralized the putative ESE sequences, each was computationally evaluated by ESEfinder and RESCUE-ESE (<http://genes.mit.edu/burgelab/rescue-ese/>) (Fairbrother et al. 2002). Also, to ensure that the mutations did not introduce a putative ESS, the mutated sequences were analyzed by FAS-ESS (<http://genes.mit.edu/fas-ess/>) (Wang *et al.* 2004). The primers listed in Table 1 were used to mutate the single exon LDLR minigene by using the QuikChange Multi Site-Directed Mutagenesis Kit (Stratagene) or the LDLR exon 9–14 minigene by using the QuikChange XL Site-Directed Mutagenesis Kit (Stratagene). Briefly, each LDLR minigene was denatured and annealed with either one or two synthetic oligonucleotide primers that contained the desired mutation, and extended by *PfuUltra* high fidelity DNA polymerase. After temperature cycling, the product was treated with *Dpn* I endonuclease and transformed into XL1-Blue supercompetent cells. LDLR mutations were confirmed by sequencing. The mutated LDLR minigenes were transfected into HepG2 cells and splicing efficiency evaluated by RT-PCR as described above.

Prediction of RNA Secondary Structure

LDLR wild type and mutated sequences were analyzed by using NIPU software with default parameters (<http://www.bioinf.uni-freiburg.de/Software/NIPU/index.html>) (Hiller *et al.* 2007).

Statistical Analysis

The effects of rs688 alleles on LDLR minigene splicing efficiency were evaluated by a paired T-test. The effects of the LDLR mutations on LDLR minigene splicing efficiency were analyzed by ANOVA with a *post-hoc* Fisher's Least Significant Difference (LSD) test (SPSS, version 17).

Results

Delta 12 and Delta 11+12 LDLR are predicted to encode soluble LDLR proteins because the loss of exon 12 or exons 11–12 produces a codon frameshift which results in novel carboxyl termini followed by premature stop codon for each protein (Zhu et al. 2007). To evaluate whether the proteins encoded by Delta 12 or Delta 11+12 LDLR are produced, we compared four antibodies for their sensitivity and specificity for detection of full-length LDLR as well as the truncated LDLR isoforms. The molecular weight of full-length LDLR is 95.4 kDa in the absence of glycosylation; following *N*- and *O*-linked glycosylation, full-length LDLR has a molecular weight of ~160 kDa. Delta 12 LDLR is predicted to be 64.1 kDa after removal of its leader sequence while Delta 11+12 LDLR is slightly smaller at 61.8 kDa. Both of these truncated forms lack the *O*-linked glycosylation site encoded within exon 15 but retain several potential *N*-linked glycosylation sites (Uniprot, <http://www.uniprot.org/uniprot/P01130>). A commercially available rabbit anti-LDLR antibody robustly detected full-length LDLR in HepG2 cell lysates and also detected truncated LDLR proteins in HEK293 cells stably transfected with Delta 12 or Delta 11+12 LDLR (Figure 1A); the bands marked with # in Figure 1A likely represent Delta 12 or Delta 11+12 LDLR proteins lacking *N*-glycosylation while those at ~75 kDa likely reflect these LDLR proteins with *N*-glycosylation. An LDLR antibody raised in chicken also labeled the ~75 kDa truncated LDLR proteins in transfected cells (Figure 1B). The identity of the proteins encoded by the truncated LDLR isoforms was further confirmed by using an

antibody raised against the novel carboxyl terminus of Delta 11+12 (Figure 1C) that labeled the 75 kDa protein in cells transfected with the Delta 11+12 LDLR expression plasmid but did not label cell lysates transfected with the Delta 12 LDLR expression plasmid (Figure 1C). Similarly, an antibody raised against the Delta 12 LDLR novel carboxyl terminus labeled a band at ~75 kDa in Delta 12 LDLR-transfected cells (Figure 1D) but did not label a protein of this size in the cells transfected with Delta 11+12 LDLR; this antibody also labeled an apparent artifactual band at 37 kDa. In summary, cellular proteins encoded by the LDLR isoforms lacking exon 12 or exons 11–12 are readily and specifically detected. Hence, these mRNAs are translated into the truncated LDLR proteins.

Since these truncated LDLR proteins lack the transmembrane domain and are predicted to be soluble proteins, we then used one of the commercial anti-LDLR antibodies to test whether soluble LDLR proteins are present in extracellular fluids. Others have reported that a soluble ~140 kDa LDLR protein resulting from cell surface LDLR proteolysis is increased in medium when cells are treated with PMA, a phorbol ester (Begg *et al.* 2004); the mechanism underlying the PMA-stimulated soluble LDLR is unclear but requires PKC and is inhibited by TNF- α protease inhibitor (Begg *et al.* 2004). Hence, we compared levels of this soluble LDLR protein to soluble LDLR arising from inefficient splicing. A ~140 kDa LDLR protein was clearly detected in concentrated conditioned medium from HepG2 cells and was increased when the cells were treated with PMA for 4 hours (Figure 2A). However, soluble ~75 kDa LDLR corresponding to Delta 12 or Delta 11+12 LDLR was not detected in the same medium (Figure 2A). Moreover, while we detected soluble ~75 kDa LDLR proteins in conditioned media of HEK293 cells transfected with Delta 11+12 or Delta 12 LDLR, similar LDLR proteins were not detected in ≥ 30 μ g of human CSF proteins (Figure 2B–C). To determine the sensitivity of this Western blot assay, we titrated decreasing amounts of recombinant human LDLR; we found that ≥ 30 ng of LDLR was readily detectable (Figure 2C). Lastly, we noted that these CSF samples contained intact proteins, as shown by their abundant amounts of soluble Clusterin (Figure 2D). We interpret these results as suggesting that since the CSF samples contained ≥ 30 μ g of protein and the assay detected ≥ 30 ng of LDLR, soluble LDLR in CSF represents $\leq 0.1\%$ of total proteins. Since APOE represents about 4% of CSF proteins (Kay *et al.* 2003), the amounts of soluble LDLR appear too low to be biologically meaningful. Hence, the Delta 12 and Delta 11+12 LDLR isoforms appear to alter LDLR function by reducing levels of mRNA encoding functional LDLR.

Since enhanced splicing of LDLR exon 12 could increase functional LDLR protein levels, and thereby be beneficial for cholesterol homeostasis and, perhaps, AD (Cao *et al.* 2006, Kim *et al.* 2009, Zou *et al.* 2008), we next investigated the RNA elements critical for exon 12 splicing. Previously, we found that the minor rs688T allele was associated with less exon 12 inclusion (Zhu *et al.* 2007, Zou *et al.* 2008). Here, we sought to evaluate the role of the 5' splice site and to identify functional ESEs by performing a series of mutagenesis and splicing studies. We began by cloning exon 12 and flanking intronic sequences into pSPL3b, a splicing vector wherein the LDLR sequence is flanked by HIV-tat intronic sequences and rabbit β -globin exons (Figure 3A top); we used this smaller minigene vector to facilitate mutagenesis. To validate this single exon minigene, we compared rs688 effects on splicing in this minigene relative to rs688 effects in the LDLR exon 9–14 minigene we previously described (Figure 3A bottom). This comparison was performed by transfecting each minigene containing each rs688 allele into HepG2 cells and quantifying splicing by RT-PCR. The LDLR minigene containing only exon 12 showed a modest but significant rs688 effect on splicing, i.e., rs688T increased the percentage of LDLR mRNA that lacked exon 12 by $3.7 \pm 0.6\%$ (mean $\pm$ SD, $n = 3$, $p = 0.009$; Figure 3B). For the LDLR exon 9–14 minigene, the rs688T allele increased the percentage of Delta 12 mRNA by $12.6 \pm 4.6\%$ ($n = 4$, $p < 0.01$; Figure 3C). Rs688 also influenced the percentage of the LDLR isoform that

lacked exons 11–12, i.e., the rs688T allele increased the percentage of Delta 11+12 LDLR by $4.5 \pm 1.9\%$ ($n = 4$, $p < 0.01$; Figure 3C). In summary, although these rs688-associated differences in splicing in the single exon minigene are less than those *in vivo* (Zhu et al. 2007, Zou et al. 2008), they demonstrate rs688 effects that are statistically significant and qualitatively similar to those seen in the exon 9–14 minigene and *in vivo*. Hence, we proceeded to use the single exon minigene to identify critical exonic elements that modulate exon 12 splicing.

Putative candidate exon 12 ESEs were identified in a two-step process. Exon 12 ESEs in eight species were first predicted by using ESEfinder. These sequences were then aligned by using Clustal, with sequences being considered to be conserved if they were consistent among at least five species. Conserved putative ESEs that overlapped one another were considered to be a conserved putative ESE region. Nine conserved putative ESE regions, R1–R9, were identified in exon 12 (Figure 4A–B). The location of rs688 is within two putative SRp40 binding sites (Figure 4A), and the conversion from C to T allele is predicted to influence SRp40 binding affinity. To distinguish these two putative SRp40 binding sites even though their sequences overlapped, we considered them as separate ESE regions designated as R5 and R6 respectively. To evaluate the function of each conserved region, the LDLR minigene was mutated to neutralize the putative affinity of each region to SR proteins (Figure 4C) by using site-directed mutagenesis with the primers indicated in Table 1.

We first considered the possibility that the mutations that were introduced may alter pre-mRNA secondary structure. Therefore, we evaluated whether the LDLR minigene mutations altered RNA secondary structures by using NIPU, which predicts RNA single-strandedness. Compared to the WT sequence, several mutations were predicted to markedly alter single-stranded regions, including MT2, MT4 and MT9 (Figure 5). To directly assess which conserved ESE are critical for splicing, wild type (WT) and each mutant (MT) minigene were transfected into HepG2 cells and splicing efficiency evaluated by RT-PCR. Most of the mutations targeting conserved putative ESE regions did not alter exon 12 splicing efficiency, including MT2, MT4 and MT9, suggesting that the possible effects of these mutations on RNA secondary structure do not modulate splicing. In contrast, MT6 and MT7, which were predicted to have minimal effects on RNA secondary structure (Figure 5), significantly increased the percentage of delta 12 from $18.2 \pm 5.5\%$ in the WT LDLR minigene to $64.3 \pm 14.3\%$ and $43.4 \pm 13.1\%$ respectively (both $p < 0.001$; Figure 6A–B). Hence, the sequences targeted in MT6 and MT7 appear critical for LDLR exon 12 splicing.

Previously, we and others showed that LDLR isoforms lacking exon 12 were major LDLR splice variants in human tissue (Zhu et al. 2007, Tveten *et al.* 2006). Our initial evaluation suggested that exon 12 may be skipped because of its weak splice site donor; as estimated by the Alternative Splicing Database (ASD) (Stamm *et al.* 2006), the 5' splice site in intron 12 is GAGgtgtgg which has a weak 6.67 score relative to the consensus splice donor sequence of CAGgtragt (Mount 1982), which has a score of 10.77. Hence, to evaluate whether optimizing this 5'splice site enhances LDLR exon12 splicing, we generated MT10 with the sequence CAGgtgtgg and MT11 with the sequence GAGgtgagt; analyzed by ASD, these changes increased the donor site score strength to 7.67 for MT10 and 10.46 for MT11. When we analyzed the effects of MT10 and MT11 on splicing efficiency, we found that MT10 had essentially no effect on exon 12 splicing, i.e., the percentage of delta 12 was $14.1 \pm 0.1\%$ ($p = 0.631$; Figure 6A–B). However, MT11 showed a clear trend towards increased exon 12 inclusion, i.e., the percentage of Delta 12 was $2.6 \pm 0.6\%$ ($p = 0.077$; Figure 6A–B). The trend associated with MT11 likely reflects the larger effect of MT11 on the donor site score, relative to MT10.

To examine whether the splicing effects of MT6 and MT7 were unique to the single exon LDLR minigene or were applicable to a more physiologic context, we introduced MT6 and MT7 into the LDLR exon 9–14 minigene. We also evaluated the effects of MT5 as a negative control. Splicing efficiency of these larger minigenes was evaluated by RT-PCR. Similar to the results observed in the LDLR exon 12 minigene, MT5 had no effect on splicing, i.e., the percentage of delta 12 was $21.1 \pm 1.7\%$, compared to $18.3 \pm 0.3\%$ in the WT LDLR (Figure 6C-D). In contrast, MT6 and MT7 significantly increased the percentage of Delta 12 to $46.4 \pm 4.8\%$ and $46.4 \pm 5.4\%$, respectively (both $p < 0.001$; Figure 6C-D). These results indicate that the regions 6 (c.1769 – c.1772) and 7 (c.1784 – c.1787) contain splicing regulatory elements that are important to exon 12 splicing.

Discussion

The primary findings of this study are several. First, the truncated LDLR proteins encoded by Delta 12 or Delta 11+12 LDLR are expressed within cells but represent $\leq 0.1\%$ of CSF proteins. Hence, the truncated LDLR proteins encoded by Delta 12 and Delta 11+12 LDLR appear to represent a loss of functional LDLR as opposed to a possible dominant negative, secreted form of the LDLR protein. Second, by optimizing the 5' splice site in intron 12 towards a consensus splice donor sequence, the proportion of exon 12-containing LDLR isoform is enhanced. Third, we present an approach to identify regions critical to exonic splicing. Forth, we apply this approach to an inefficiently spliced LDLR exon, and report that regions defined by MT6 and MT7 are critical to LDLR exon 12 splicing. These two ESE regions are c.1769 - c.1772 and c.1784 - c.1787 respectively. Separated by only 11 bp, these two regions flank rs688 at c.1773. Hence, this region appears critical for LDLR exon 12 splicing efficiency.

The role of soluble versions of LDLR family members is unclear. Previously, a splice variant of apoE receptor 2 (APOER2) that includes a furin cleavage site was shown to generate soluble APOER2 protein that acts as a dominant negative receptor by inhibiting Reelin signaling (Koch *et al.* 2002). Similarly, soluble forms of LDLR and LRP1 as a result of cell surface proteolytic cleavage have been reported, with the suggestion that these proteins also antagonize the actions of cell surface LDLR and LRP1 (Begg *et al.* 2004, Sagare *et al.* 2007). However, while we were able to detect soluble LDLR protein resulting from cell surface cleavage in concentrated cell media, we did not detect soluble LDLR protein corresponding to Delta 12 or Delta 11+12 LDLR. Similarly, soluble LDLR protein was not observed in human CSF with a detection limit of 30 ng. We interpret these results as suggesting that the truncated LDLR proteins encoded by Delta 12 or Delta 11+12 LDLR do not substantially accumulate in the extracellular space.

The approach that we have used to identify critical splicing elements within LDLR exon 12 may be generally applicable. The first step was to identify putative ESEs that were evolutionarily conserved. ESEfinder, RESCUE-ESE, and Splicing Rainbow (<http://www.ebi.ac.uk/asd-srv/wb.cgi>) provide useful information regarding possible splicing regulatory sites. Although the 140-bp human LDLR exon 12 was predicted initially by ESEfinder to contain 25 putative ESEs, the additional criterion of ESE conservation across five or more species focused attention upon nine regions that were enriched for conserved ESEs. The second step was a series of structure function studies that relied upon site-directed mutagenesis coupled with *in vitro* minigene splicing studies. This analysis confirmed two of the putative ESE regions as actually harboring functional RNA splicing regulatory elements. The possibility that ESEfinder and evolutionary conservation are yet likely to produce a false positive result was reinforced further by additional work, i.e., we suspected that other naturally occurring LDLR mutations located within the regions defined by MT6 and MT7 may also modulate LDLR exon 12 splicing. By searching the Universal

Mutation Database (<http://www.umd.necker.fr>) (Varret *et al.* 1998), we found a mutation, c.1784G>A, which has been reported in Korean patients with familial hypercholesterolemia (Kim *et al.* 2004) and is located in the regulatory region defined by MT7. This mutation changes the encoded amino acid from arginine to glutamine, but is also predicted by ESEfinder to block ESE affinity toward SR proteins, i.e., the presence of c.1784G>A neutralizes putative ESEs for SF2/ASF and SC35. To elucidate whether this mutation influences LDLR exon 12 splicing, we generated an LDLR exon 9–14 minigene containing the mutation; our results revealed no significant effect on splicing (I-FL and SE, unpublished observation), indicating that this mutation likely causes LDLR malfunction and hence familial hypercholesterolemia by amino acid substitution, not by aberrant splicing. In summation, although the approach used here does require elimination of multiple false positive candidates, functional exonic splicing elements are readily identifiable through the combination of *in silico* and *in vitro* analyses.

Pre-mRNA secondary structure analysis can also provide insight into ESE prediction because the binding of many splicing factors depends upon RNA secondary structure (Muro *et al.* 1999, Hiller *et al.* 2007). The majority of functional splicing regulatory elements are exposed in the loop of a hairpin RNA structure. When an ESE is located in the stem of a RNA structure, the ESE may lose its ability to modulate splicing since the binding site is not exposed and therefore cannot be recognized by regulatory proteins. However, a few functional motifs in double-stranded RNA regions have been verified, e.g., two *cis*-elements, ISE-2 and ISAR, form a double-stranded RNA structure to regulate fibroblast growth factor receptor-2 splicing (Muh *et al.* 2002). In this study, the two ESE regions critical to LDLR exon 12 splicing were predicted by NIPU to be located in double-stranded regions (Figure 6). The proteins that may interact with these regulatory *cis*-elements are unclear. One possibility is splicing factors such as MBNL1 which binds directly to the double-stranded region of a stem-loop structure to regulate splicing (Warf *et al.* 2009). Alternatively, splicing factors that bind to single stranded RNA sequence may inhibit the RNA from forming secondary structures. Interestingly, the regions defined by MT6 and MT7 are predicted to interact with SRp40, SC35 and SF2/ASF (Figure 4A) while the G-rich region between MT6 and MT7 matches the predicted binding site for hnRNP F and H. However, we have found that ectopic expression of SRp40, SF2/ASF, SC35, hnRNP F or hnRNP H does not influence the LDLR exon 9–14 minigene splicing efficiency (Ling and Estus 2010); similar inconsistencies between putative regulatory elements and their cognate factors have been reported elsewhere (Lastella *et al.* 2004, Auclair *et al.* 2006). Overall, we interpret these results as suggesting that while we have identified the RNA elements critical for LDLR exon 12 splicing, additional studies will be necessary to identify the splicing factors involved in this process.

In summary, we found that the truncated LDLR proteins encoded by LDLR isoforms lacking exon 12 or exons 11–12 are indeed translated but are not detectably present in extracellular fluids. Moreover, optimizing the 5' splice site in the LDLR intron 12 increased splicing efficiency to near 100%, suggesting that exon 12 is inefficiently included in LDLR because of its weak splice donor site. Lastly, we identified two regions that are critical to LDLR exon 12 splicing and that flank rs688. Hence, the key exon 12 splicing decision resides in the interaction between these functional RNA elements and their regulatory proteins, which will be the focus of future work. Since LDLR is critical to cholesterol and amyloid beta homeostasis, understanding LDLR exon 12 splicing efficiency may provide insights for therapeutic and/or preventive strategies when dealing with dyslipidemia, adverse cardiovascular events and/or AD (Zhu *et al.* 2007, Zou *et al.* 2008, Kim *et al.* 2009).

Acknowledgments

The authors gratefully acknowledge CSF supplied by the University of Kentucky AD Center, which is supported by P30AG028383, as well as NIH for grant support (R01AG026147 and P01AG030128).

References

- Auclair J, Busine MP, Navarro C, et al. Systematic mRNA analysis for the effect of MLH1 and MSH2 missense and silent mutations on aberrant splicing. *Hum Mutat.* 2006; 27:145–154. [PubMed: 16395668]
- Begg MJ, Sturrock ED, van der Westhuyzen DR. Soluble LDL-R are formed by cell surface cleavage in response to phorbol esters. *Eur J Biochem.* 2004; 271:524–533. [PubMed: 14728679]
- Black DL. Mechanisms of alternative pre-messenger RNA splicing. *Annual review of biochemistry.* 2003; 72:291–336.
- Buckler AJ, Chang DD, Graw SL, Brook JD, Haber DA, Sharp PA, Housman DE. Exon amplification: a strategy to isolate mammalian genes based on RNA splicing. *Proceedings of the National Academy of Sciences of the United States of America.* 1991; 88:4005–4009. [PubMed: 1850845]
- Burn TC, Connors TD, Klinger KW, Landes GM. Increased exon-trapping efficiency through modifications to the pSPL3 splicing vector. *Gene.* 1995; 161:183–187. [PubMed: 7665076]
- Cao D, Fukuchi K, Wan H, Kim H, Li L. Lack of LDL receptor aggravates learning deficits and amyloid deposits in Alzheimer transgenic mice. *Neurobiol Aging.* 2006; 27:1632–1643. [PubMed: 16236385]
- Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic acids research.* 2003; 31:3568–3571. [PubMed: 12824367]
- Cooper TA, Wan L, Dreyfuss G. RNA and disease. *Cell.* 2009; 136:777–793. [PubMed: 19239895]
- D'Souza I, Schellenberg GD. Arginine/serine-rich protein interaction domain-dependent modulation of a tau exon 10 splicing enhancer: altered interactions and mechanisms for functionally antagonistic FTDP-17 mutations Delta280K AND N279K. *J Biol Chem.* 2006; 281:2460–2469. [PubMed: 16308321]
- Fairbrother WG, Yeh RF, Sharp PA, Burge CB. Predictive identification of exonic splicing enhancers in human genes. *Science.* 2002; 297:1007–1013. [PubMed: 12114529]
- Fairbrother WG, Yeo GW, Yeh R, Goldstein P, Mawson M, Sharp PA, Burge CB. RESCUE-ESE identifies candidate exonic splicing enhancers in vertebrate exons. *Nucleic Acids Res.* 2004; 32:W187–190. [PubMed: 15215377]
- Fryer JD, Demattos RB, McCormick LM, et al. The low density lipoprotein receptor regulates the level of central nervous system human and murine apolipoprotein E but does not modify amyloid plaque pathology in PDAPP mice. *J Biol Chem.* 2005; 280:25754–25759. [PubMed: 15888448]
- Hastings ML, Krainer AR. Pre-mRNA splicing in the new millennium. *Curr Opin Cell Biol.* 2001; 13:302–309. [PubMed: 11343900]
- Hertel KJ. Combinatorial control of exon recognition. *J Biol Chem.* 2008; 283:1211–1215. [PubMed: 18024426]
- Hiller M, Zhang Z, Backofen R, Stamm S. Pre-mRNA secondary structures influence exon recognition. *PLoS Genet.* 2007; 3:e204. [PubMed: 18020710]
- Hobbs HH, Brown MS, Goldstein JL. Molecular genetics of the LDL receptor gene in familial hypercholesterolemia. *Hum Mutat.* 1992; 1:445–466. [PubMed: 1301956]
- Kay AD, Petzold A, Kerr M, Keir G, Thompson EJ, Nicoll JA. Cerebrospinal fluid apolipoprotein E concentration decreases after traumatic brain injury. *J Neurotrauma.* 2003; 20:243–250. [PubMed: 12820678]
- Kim J, Castellano JM, Jiang H, Basak JM, Parsadanian M, Pham V, Mason SM, Paul SM, Holtzman DM. Overexpression of Low-Density Lipoprotein Receptor in the Brain Markedly Inhibits Amyloid Deposition and Increases Extracellular Abeta Clearance. *Neuron.* 2009; 64:632–644. [PubMed: 20005821]

- Kim JH, Choi HK, Lee H, Park HY, Kim JW, Kim HJ, Lee ST. Novel and recurrent mutations of the LDL receptor gene in Korean patients with familial hypercholesterolemia. *Mol Cells*. 2004; 18:63–70. [PubMed: 15359125]
- Koch S, Strasser V, Hauser C, Fasching D, Brandes C, Bajari TM, Schneider WJ, Nimpf J. A secreted soluble form of ApoE receptor 2 acts as a dominant-negative receptor and inhibits Reelin signaling. *EMBO J*. 2002; 21:5996–6004. [PubMed: 12426372]
- Lastella P, Resta N, Miccolis I, Quagliarella A, Guanti G, Stella A. Site directed mutagenesis of hMLH1 exonic splicing enhancers does not correlate with splicing disruption. *J Med Genet*. 2004; 41:e72. [PubMed: 15173238]
- Mount SM. A catalogue of splice junction sequences. *Nucleic Acids Res*. 1982; 10:459–472. [PubMed: 7063411]
- Muh SJ, Hovhannisyanyan RH, Carstens RP. A Non-sequence-specific double-stranded RNA structural element regulates splicing of two mutually exclusive exons of fibroblast growth factor receptor 2 (FGFR2). *J Biol Chem*. 2002; 277:50143–50154. [PubMed: 12393912]
- Muro AF, Caputi M, Pariyarath R, Pagani F, Buratti E, Baralle FE. Regulation of fibronectin EDA exon alternative splicing: possible role of RNA secondary structure for enhancer display. *Mol Cell Biol*. 1999; 19:2657–2671. [PubMed: 10082532]
- Sagare A, Deane R, Bell RD, et al. Clearance of amyloid-beta by circulating lipoprotein receptors. *Nat Med*. 2007; 13:1029–1031. [PubMed: 17694066]
- Stamm S, Riethoven JJ, Le Texier V, Gopalakrishnan C, Kumanduri V, Tang Y, Barbosa-Morais NL, Thanaraj TA. ASD: a bioinformatics resource on alternative splicing. *Nucleic Acids Res*. 2006; 34:D46–55. [PubMed: 16381912]
- Tazi J, Bakkour N, Stamm S. Alternative splicing and disease. *Biochim Biophys Acta*. 2009; 1792:14–26. [PubMed: 18992329]
- Tveten K, Ranheim T, Berge KE, Leren TP, Kulseth MA. Analysis of alternatively spliced isoforms of human LDL receptor mRNA. *Clin Chim Acta*. 2006; 373:151–157. [PubMed: 16828075]
- Varret M, Rabes JP, Thiart R, et al. LDLR Database (second edition): new additions to the database and the software, and results of the first molecular analysis. *Nucleic Acids Res*. 1998; 26:248–252. [PubMed: 9399845]
- Wang Z, Rolish ME, Yeo G, Tung V, Mawson M, Burge CB. Systematic identification and analysis of exonic splicing silencers. *Cell*. 2004; 119:831–845. [PubMed: 15607979]
- Warf MB, Diegel JV, von Hippel PH, Berglund JA. The protein factors MBNL1 and U2AF65 bind alternative RNA structures to regulate splicing. *Proc Natl Acad Sci U S A*. 2009; 106:9203–9208. [PubMed: 19470458]
- Zhu H, Tucker HM, Gear KE, Simpson JF, Manning AK, Cupples LA, Estus S. A common polymorphism decreases low-density lipoprotein receptor exon 12 splicing efficiency and associates with increased cholesterol. *Hum Mol Genet*. 2007; 16:1765–1772. [PubMed: 17517690]
- Zou F, Gopalraj RK, Lok J, et al. Sex-dependent association of a common low-density lipoprotein receptor polymorphism with RNA splicing efficiency in the brain and Alzheimer's disease. *Hum Mol Genet*. 2008; 17:929–935. [PubMed: 18065781]

Figure 1. Expression of truncated LDLR proteins in stably transfected cells

Cells transfected with or without Delta 11+12 or Delta 12 LDLR were lysed and subjected to Western blot analysis with rabbit anti-LDLR antibody (A), chicken anti-LDLR antibody (B), rabbit anti-Delta 11+12 antibody (C), or rabbit anti-Delta 12 antibody (D). The label “Hep” refers to a HepG2 cell lysate, while “293” refers to HEK293 cell lysate, “293 D11+12” refers to HEK293 cells stably transfected with Delta 11+12 LDLR expression plasmid and “293 D12” refers to HEK293 cells stably transfected with Delta 12 LDLR expression plasmid. Bands marked with * reflect full-length LDLR, empty arrowhead likely reflects *N*-glycosylated Delta 11+12 and Delta 12 LDLR, and # likely represent these truncated LDLR proteins without *N*-linked glycosylation.

Figure 2. Examination of soluble LDLR proteins in extracellular fluids

LDLR proteins were detected by Western blot analysis with chicken anti-LDLR antibody. Cell lysate from HEK293 cells transfected with Delta 12 LDLR expression plasmid was used as a positive control for the Delta 12 LDLR protein while cell lysate from HepG2 cells was used to show the position of full-length LDLR protein. Soluble LDLR proteins were detected in HepG2 cell concentrated media; the proteins detected at ~140 kDa appear to be soluble LDLR resulting from cell surface cleavage based on size and their increased amount following PMA treatment. Similar findings were reported by Begg et al. who also observed two bands corresponding to LDLR (Begg et al. 2004). Soluble proteins corresponding to the Delta 12 or Delta 11+12 LDLR were not detected (A). Conditioned media from HEK293 cells transfected with Delta 11+12 or Delta 12 LDLR expression plasmid, and human CSF were also evaluated by the Western blot analysis (B). CSF samples from subjects 680 and 705 contained 38 and 70 μ g of protein per lane, respectively. Recombinant human LDLR was titrated to determine assay sensitivity (C). Decreasing amounts of LDLR were analyzed in parallel with CSF samples from subjects 601, 627 and 659, each of which contained 30 μ g of protein. The CSF portion of the blot was reprobbed with anti-Clusterin antibody to ensure protein loading (D).

Figure 3. Comparison of rs688 effects in LDLR exon 12 and exon 9-14 minigenes

LDLR minigenes carrying rs688C or rs688T alleles were transfected into HepG2 cells and after 24 hours, mRNA isolated and converted to cDNA for analyses. PCR was performed with primers indicated as arrows in minigene schemes (A). Representative gel images for LDLR exon 12 minigene transfection (B) and exon 9-14 minigene transfection (C) are shown.

Figure 4. Conserved exon 12 putative ESE regions and LDLR mutant sequences

The putative ESEs within human LDLR exon 12 as discerned by ESEfinder are shown (A). The X-axis shows LDLR exon 12 sequence, and the Y-axis indicates numerical scores for putative ESE strength. The dark gray boxes represent putative binding sites for SF2/ASF, while black boxes represent those for SC35, white boxes for SRp55, and light gray boxes for SRp40. Putative ESEs that were conserved in at least five of the evaluated species were considered putative ESE regions (B). These regions, R1 – R9, are denoted in A by underlining and in B by the shaded boxes. * represents a nucleotide that is conserved within the 8 species. The mutations of conserved ESE regions are also shown, with putative ESE regions in shaded boxes and mutated sequences in bold font on top (C). The position of rs688 is boxed. Upper case denotes exon 12 sequence while lower case denotes intronic sequence. Mutations 1–9 (MT1 to MT9) were introduced to neutralize affinity of the conserved putative ESE regions to SR proteins as predicted by ESEfinder; we ensured that a new ESE or ESS was not introduced by evaluating the mutant sequences with ESE-RESCUE and FAS-ESS as well. Mutations 10 and 11 (MT10 and MT11) were introduced to optimize the 5' splice site in LDLR intron 12.

Figure 5. Predicted effects of mutations of putative ESEs on RNA secondary structure
 Wild type and mutated sequences were analyzed by NIPU which computes PU score to predict RNA single-strandedness. X-axis represents the sequence of LDLR exon 12, and Y-axis represents PU value for each position. A higher PU value indicates higher probability of single-strandedness. Mutated regions were highlighted in grey boxes.

Figure 6. Effects of mutations of putative ESEs on LDLR splicing efficiency

Wild type (WT) and mutant (MT) clones were transfected into HepG2 cells, and splicing efficiency evaluated by RT-PCR. Representative gel images (A and C) and quantitative results (B and D; mean $\pm$ SD, $n = 3$) are shown. Mutations were introduced into LDLR exon 12 minigene (A-B). While MT1 to MT9 targeted putative ESE regions to reduce exon 12 inclusion, MT10 and MT11 were introduced to optimize the 5' splice site in intron 12 to enhance exon 12 inclusion. MT11 showed a trend towards increasing splicing efficiency ($p = 0.077$). While most ESE mutations did not change splicing efficiency, MT6 and MT7 significantly increased delta 12 ($p < 0.001$ and $p = 0.003$ respectively). When MT6 and MT7 were introduced into the LDLR exons 9–14 minigene (C-D), MT6 and MT7 significantly decreased splicing efficiency ($p < 0.001$).

Table 1**Mutagenesis Primers**

Mutation positions are underlined and in bold font. Upper case refers to exonic sequence while lower case refers to intronic sequence.

Mutation	Mutation Primer Sequence
MT1	CTCCTCAGTGGCCGC <u>GCGCT</u> CTGGGTTGACTCCAAAC
MT2	TGGCCGCCTCTACTGGGTTGC <u>CAGTT</u> AAACTTCACTCCATCTC
MT3	GGGTTGACTCCAAACTTCA <u>TGCGCT</u> CTCAAGCATCGATGTCAAC
MT4	CAAACCTTCACTCCATCTC <u>GGCA</u> ATCGATGTCAACGGGGG
MT5	ATCTCAAGCATCGA <u>AAT</u> CAACGGGGGCAACCG
MT6	TCTCAAGCATCGAT <u>GAGT</u> CGGGGGCAACCGGAAG
MT7	TGTCAACGGGGGCAACCT <u>TAGT</u> GACCATCTTGAGGATG
MT8	AAGAGGCTGGCCACCT <u>TAGC</u> CTCCTGGCCGTCTTTG
MT9	CCGGAAGACCATCTT <u>CAGT</u> GATGAAAAGAGGCTGGCCC
MT10	TTGGCCGTCTTT <u>CAG</u> gtgtggttacg
MT11	GGCCGTCTTTGAGgt <u>agt</u> cttacgtacgagatgc