2009

REGULATION OF LOW DENSITY LIPOPROTEIN RECEPTOR SPLICING EFFICIENCY

I-Fang Ling
University of Kentucky, iling2@uky.edu

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Recommended Citation
Ling, I-Fang, "REGULATION OF LOW DENSITY LIPOPROTEIN RECEPTOR SPLICING EFFICIENCY" (2009). University of Kentucky Doctoral Dissertations. 794.
https://uknowledge.uky.edu/gradschool_diss/794

This Dissertation is brought to you for free and open access by the Graduate School at UKnowledge. It has been accepted for inclusion in University of Kentucky Doctoral Dissertations by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.
ABSTRACT OF DISSERTATION

I-Fang Ling

The Graduate School
University of Kentucky
2009
ABSTRACT OF DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Medicine at the University of Kentucky

By

I-Fang Ling

Lexington, Kentucky

Director: Dr. Steven Estus, Professor of Physiology

Lexington, Kentucky

2009

Copyright © I-Fang Ling 2009
ABSTRACT OF DISSERTATION

REGULATION OF LOW DENSITY LIPOPROTEIN RECEPTOR SPlicing EFFICIENCY

Low density lipoprotein receptor (LDLR) is an apolipoprotein E (apoE) receptor and may play a role in Alzheimer’s disease (AD) development. A single nucleotide polymorphism (SNP), rs688, that has been identified to modulate the splicing efficiency of LDLR exon 12 and is associated with higher cholesterol and AD in some case-control populations. The exon 12 deleted mRNA is predicted to produce a soluble form of LDLR that fails to mediate apoE uptake. To gain additional insights, in this study, I seek to understand the regulation of LDLR splicing efficiency. To identify functional cis-elements within LDLR exon 12, I mutated several conserved putative exonic splicing enhancers (ESEs) to neutralize their affinity to serine/arginine-rich (SR) proteins. Transfection of wild type (WT) or mutant LDLR minigenes in HepG2 cells was performed, and splicing efficiency evaluated by quantitative RT-PCR. The results showed that two functional ESEs within exon 12, near rs688, are critical to LDLR splicing. To identify splicing factors that modulate exon 12 splicing, I co-transfected an LDLR minigene and vectors encoding different SR proteins and heterogeneous nuclear ribonucleoproteins (hnRNPs). After quantifying the splicing efficiency, I found that SRp20 and SRp38 increased exon 11-skipping. Moreover, ectopic expression of SRp38-2 and hnRNP G increased exon 11&12-skipping. Interestingly, the actions of hnRNP G did not require its RNA recognition motif (RRM). To further investigate the role of these splicing factors on LDLR splicing, I quantified the expression level of these splicing factors as well as LDLR splicing efficiency in human brain and liver. I found that SRp38 mRNA expression is associated with LDLR splicing efficiency. In conclusion, this study discovered that rs688 is located close to the two functional ESEs within LDLR exon 12, and revealed a role of SRp38 in LDLR splicing efficiency.

KEYWORD: LDLR, SNP, Splicing, ESE, SRp38
I-Fang Ling
Student’s Signature
November 20, 2009
Date
REGULATION OF LOW DENSITY LIPOPROTEIN RECEPTOR SPLICING EFFICIENCY

By

I-Fang Ling

Dr. Steven Estus
Director of Dissertation
Dr. Ok-Kyong Park-sarge
Director of Graduate Studies

November 20, 2009
Date
RULES FOR THE USE OF DISSERTATIONS

Unpublished dissertations submitted for the Doctor's degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with the permission of the author, and with the usual scholarly acknowledgments.

Extensive copying or publication of the dissertation in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this dissertation for use by its patrons is expected to secure the signature of each user.

<table>
<thead>
<tr>
<th>Name</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
REGULATION OF LOW DENSITY LIPOPROTEIN RECEPTOR SPlicing EFFICIENCY

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Medicine at the University of Kentucky

By
I-Fang Ling

Lexington, Kentucky

Director: Dr. Steven Estus, Professor of Physiology

Lexington, Kentucky

2009

Copyright © I-Fang Ling 2009
For my great teachers and lovely family

For every firm handshake that gives me the encouragement

For every warm hug that provides me the strength
ACKNOWLEDGMENTS

This dissertation could not be completed without the support and assistance of many people. First, I would like to thank my mentor, Dr. Steven Estus, for his wonderful patience, guidance, and encouragement. I still remember the first meeting I had with him. While I was just a student interested in rotating in his lab, he spent almost three hours explaining the projects to me with wisdom and passion. Over the last five years, every time that I struggled with my project, I always got excellent advice and generous support from him and the kind encouragement in the end of every meeting always gave me the instigation to accomplish the work. Dr. Estus not only taught me to be a good scientist, but a smart, genuine, and respectful one. Second, I would like to thank my committee members, Dr. Franca Cambi, Dr. Timothy McClintock, and Dr. Stefan Stamm for their scientific guidance and helpful insight. Also I thank Dr Wayne Cass for serving as the outside examiner for my dissertation. Third, I would like to extend the thanks to all the faculty and staff in the Department of Physiology. The department provides excellent training, as well as an inspiring environment for students to focus on research while having fun at the same time. Thank you for taking care of the students as a big family.

Dr. Estus guides my direction, but the members in the lab help to sharpen and vary my tools. I hope this dissertation has done their generous assistance justice. Thanks Ranga for initiating this interesting project. Thanks Haiyan for being like a big sister, giving me helpful advice. Thanks Karrie for always being a joy in the lab. Thanks Laura for caring for me as a good friend. Thanks Chris for sharing exciting thoughts. Special thanks to Jim who taught me to stand up for something I believe in. Thank you for all the discussions that help me to think more carefully and thoroughly.

Doing research is a long journey. Without the companionship of my friends, I may not be able to bear the loneliness. I am very blessed to be acquainted with many wonderful friends. The conversations we have shared, the roads we have driven, the food we have tasted, the blue water we have soaked our feet in, and the trails we have walked on, I will remember. Thank you for making my PhD life more colorful.
Last, but not least, I would like to thank my lovely family. My dad, Lun-Kun Ling, is a quiet person, but a firm pat from him on my shoulder gives me the most inspiring courage. My mom, Su-Jen Lu, felt reluctant to let me leave home to study abroad, but she had the most delighted smile on her face when I got the admission. My dear sister, I-Tien Ling, is always my best friend who reminds me when I am swollen-headed and comforts me when I am in sorrow. Thank you for being there for me. Thank you for being happier than myself when I achieve something. To make you all proud is my strongest motivation.
# TABLE OF CONTENTS

ACKNOWLEDGMENT ................................................................................................................................. iii

LIST OF TABLES ........................................................................................................................................... vii

LIST OF FIGURES ......................................................................................................................................... viii

CHAPTER ONE: INTRODUCTION .................................................................................................................. 1

Alzheimer’s Disease ...................................................................................................................................... 1
  Impact .......................................................................................................................................................... 1
  Pathogony .................................................................................................................................................. 1
  Genetics ...................................................................................................................................................... 3

APOE and AD .................................................................................................................................................. 5
  Cholesterol Metabolism ............................................................................................................................. 5
  Aβ Production .......................................................................................................................................... 6
  Aβ Clearance ............................................................................................................................................ 7
  Tau Phosphorylation ................................................................................................................................. 8

APOE Receptor: LDLR family ...................................................................................................................... 8

LDLR .............................................................................................................................................................. 10
  Gene Structure .......................................................................................................................................... 10
  Function .................................................................................................................................................... 11
  LDLR and Familial Hypercholesterolemia ................................................................................................. 12
  LDLR and AD .......................................................................................................................................... 12
  LDLR Mutations ...................................................................................................................................... 13

Regulation of Splicing .................................................................................................................................. 14

Specific Aims .................................................................................................................................................. 15

CHAPTER TWO: IDENTIFICATION OF FUNCTIONAL EXONIC SPLICING ENHANCER IN LDLR EXON 12 .......................................................... 18

Introduction .................................................................................................................................................. 18

Materials and Methods .............................................................................................................................. 20
  Cell culture ............................................................................................................................................... 20

Construction of LDLR minigenes .............................................................................................................. 20

Evaluation of LDR minigene splicing ........................................................................................................ 20

Mutagenesis of conserved putative ESE .................................................................................................. 21

Prediction of RNA secondary structure .................................................................................................. 22

Statistical analysis ...................................................................................................................................... 22

Result ............................................................................................................................................................ 22

Generation of LDLR exon 12 minigene ...................................................................................................... 22

Identification of conserved putative ESEs ................................................................................................. 23
LIST OF TABLES

Table 2.1 Conserved putative ESE regions in LDLR exon 12 ........................................ 33
Table 2.2 Mutation Primers .................................................................................................. 36
Table 3.1 Splicing factor expression in human brain and liver ........................................... 74
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>LDLR family members</td>
<td>16</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>Exon organization and protein domains of the human LDLR</td>
<td>17</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>Cloned LDLR exon 12 and flanking intronic sequences</td>
<td>30</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Rs688 effects upon LDLR exon 9-14 and exon 12 minigene</td>
<td>31</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>Identification of conserved putative ESE region</td>
<td>32</td>
</tr>
<tr>
<td>Figure 2.4</td>
<td>Mutagenesis of conserved putative ESEs</td>
<td>34</td>
</tr>
<tr>
<td>Figure 2.5</td>
<td>Effects of mutations of putative ESEs on LDLR exon 12 minigene splicing</td>
<td>37</td>
</tr>
<tr>
<td>Figure 2.6</td>
<td>Effects of putative ESE mutations on LDLR exon 9-14 minigene splicing</td>
<td>38</td>
</tr>
<tr>
<td>Figure 2.7</td>
<td>Predicted effects of mutations of putative ESEs on RNA secondary structure</td>
<td>39</td>
</tr>
<tr>
<td>Figure 2.8</td>
<td>c.1784G&gt;A has no effect on LDLR minigene splicing</td>
<td>40</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>Expression of transfected SR proteins</td>
<td>58</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>SR protein family effects on LDLR minigene splicing in vitro</td>
<td>59</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>HnRNP family member effects on LDLR minigene splicing in vitro</td>
<td>61</td>
</tr>
<tr>
<td>Figure 3.4</td>
<td>Splicing factors show dose-dependent effects on LDLR splicing</td>
<td>62</td>
</tr>
<tr>
<td>Figure 3.5</td>
<td>Effect of splicing factors on LDLR exon 12 minigene splicing</td>
<td>63</td>
</tr>
<tr>
<td>Figure 3.6</td>
<td>Effects of splicing factors on mutant LDLR minigene splicing</td>
<td>64</td>
</tr>
<tr>
<td>Figure 3.7</td>
<td>Evaluation of HepG2 cell viability after cantharidin treatment</td>
<td>65</td>
</tr>
<tr>
<td>Figure 3.8</td>
<td>Cantharidin effect on cell viability and LDLR minigene splicing</td>
<td>66</td>
</tr>
<tr>
<td>Figure 3.9</td>
<td>Effect of kinases and PP1 on LDLR minigene splicing</td>
<td>67</td>
</tr>
</tbody>
</table>
Figure 3.10 siRNA knockdown of ectopic SRp38 expression .............................. 68
Figure 3.11 siRNA decreased mRNA expression........................................... 70
Figure 3.12 Effect of siRNAs on endogenous SRp38 expression and LDLR minigene
splicing ........................................................................................................ 71
Figure 3.13 Effect of shRNAs on SRp38 expression and LDLR minigene splicing....... 72
Figure 3.14 Splicing factor expression in human brain and liver......................... 73
Figure 3.15 SRp38 and rs688 are associated with LDLR splicing efficiency .............. 75
Figure 4.1 Model of LDLR splicing regulation effect on cholesterol level and AD risk.....
......................................................................................................................... 83
CHAPTER ONE
INTRODUCTION

Alzheimer’s Disease

Impact
Alzheimer’s disease (AD) is the most common form of dementia which is characterized by loss of memory and other cognitive abilities. AD is an age-related disease such that the prevalence of AD increases with aging. While only 0.075% of population under age 65 have AD, 13% of population above age 65 are affected by AD in the United States. The number of people with the disease doubles for every 5-year age interval beyond age 65 (Alzheimer's 2009). Twenty-five to 30 million people worldwide are estimated to currently suffer from AD (Ferri, Prince et al. 2005; Minati, Edginton et al. 2009) with about 5.3 million Americans are living with AD (Hebert, Scherr et al. 2003; Alzheimer's 2009). In 2011, the first baby boomer will turn 65 years old. As the proportion of elders in the population increase, the number of people with AD is predicted to double at least by 2050 (Hebert, Scherr et al. 2003). Because of the demand of this disease, intensive caregiving is required for not only patients are affected by AD, but also their family members, friends, and caregivers. Additionally, people with AD are high user of healthcare services. The direct and indirect costs of Alzheimer's and other dementias to Medicare, Medicaid and businesses amount to more than $148 billion each year, much higher than costs for other Medicare beneficiaries (Alzheimer's 2009). An intervention that could delay the mean onset of AD by 5 years would reduce the expected prevalence by 50% after 40 years (Brookmeyer, Gray et al. 1998). Although the pathological mechanism behind the disease is still not completely understood, several genetic breakthroughs have been achieved to provide hints of the disease development and help understanding the biochemical pathology of the disease. The long-term goal is to prevent the onset or progression of the disease.

Pathology
Alzheimer’s disease was first described by a German neuropsychiatrist, Alois Alzheimer,
in 1906. The patient first had strange behavioral symptoms, including a loss of short-term memory. After autopsy, atrophy of the brain and dilatation of the ventricles were observed, and vast amyloid plaque and neurofibrillary tangles were stained around amygdala, hippocampus and neocortex. AD is now known as a neurodegenerative disease characterized by loss of neurons, extracellular amyloid plaques, and intracellular neurofibrillary tangles. A current AD neuropathological assessment known as the National Institute on Aging-Reagan Institute (NIA-Reagan) criteria is base on the demonstration of sufficient numbers of these hallmark lesions during autopsy (Jellinger and Bancher 1998; Newell, Hyman et al. 1999).

Amyloid plaques are extracellular insoluble fibrous aggregates of proteolytic products of the amyloid precursor protein (APP), an integral membrane glycoprotein. APP is abundantly expressed in neurons and has several roles in the development and function of the neurons (Mattson 1997; Young-Pearse, Bai et al. 2007; Young-Pearse, Chen et al. 2008). APP can be processed via non-amyloidogenic or amyloidogenic pathways. When undergoing non-amyloidogenic processing, APP is sequentially cleaved by α-secretase and γ-secretase. In contrast, when APP is sequentially cleaved by β-secretase and γ-secretase, amyloidogenic peptide is generated. Due to the slippage of γ-secretase at the cleavage site at the C terminal of amyloid, different species of amyloid-β peptide (Aβ) with variable length are produced. The major species is Aβ_{40} which is 40 amino acids long. The Aβ_{42}, about 10% of amyloid-β peptide, is the core molecule of amyloid plaques, but Aβ_{40} is usually colocalized with Aβ_{42} in a plaque (Jarrett, Berger et al. 1993). Whether the deposition of Aβ is the cause of neuronal degeneration is still a debate (Neve and Robakis 1998; Small and Duff 2008; Pimplikar 2009). Recently, soluble Aβ oligomers or dimers have been reported to be neurotoxic and impair memory (Lesne, Koh et al. 2006; Haass and Selkoe 2007; Shankar, Li et al. 2008). A lot of genetic data support the role of APP processing in AD development, leading to an Aβ-centric view of AD. A phase III clinical trial of a compound inhibiting Aβ aggregation has failed to prevent AD progression (Gauthier, Aisen et al. 2009). However, several clinical trials of vaccine targeting Aβ are ongoing (Rafii and Aisen 2009). These results would help to clarify the role of amyloid in AD.
Neurofibrillary tangles (NFTs) are intracellular aggregates of microtubule-associated protein tau (MAPT) that is abnormally hyperphosphorylated and arranged in a form of paired helical filaments (Grundke-Iqbal, Iqbal et al. 1986; Kosik, Joachim et al. 1986). NFT is not only seen in AD, but also in a class of neurodegenerative diseases called taupathies, including Pick’s disease, frontotemporal dementia, and Down syndrome. The number of NFTs directly correlates with the presence and the degree of dementia in AD (Alafuzoff, Iqbal et al. 1987; Arriagada, Growdon et al. 1992). Tau is highly expressed in neurons and stabilizes microtubules within cells (Drechsel, Hyman et al. 1992). Hyperphosphorylated tau fails to bind microtubules, causing microtubule dissociation and hence neurodegeneration (Brunden, Trojanowski et al. 2009). Whether NFTs contribute to the AD development or are just a secondary phenomenon has been disputed, since tau knockout mice develop a functional nervous system. However, mutations in MAPT gene that result in frontotemporal dementia with parkinsonism-17 (FTDP-17) have shown that tau can be pivotal in neurodegenerative process (Hong, Zhukareva et al. 1998). The link between Aβ and NFT is not clear, but the prevailing viewpoint is that the Aβ initiates a cellular cascade that results in tau aggregation (Hardy and Selkoe 2002).

**Genetics**

While the majority of AD cases occur after age 65, about 5% of AD cases develop before age 65, referred to as early-onset AD. Sometimes the patients are as young as age 30 (Hardy and Selkoe 2002).

Three genes have been linked to early-onset familial AD (FAD), including APP, presenilin 1 (PSEN1), and presenilin 2 (PSEN2). The first genetic cause of AD identified is a mutation of APP (Goate, Chartier-Harlin et al. 1991). Some rare missense mutations located close to the α-, β-, or γ-secretase cleavage site in APP change APP processing, increase Aβ production (Citron, Oltersdorf et al. 1992; Cai, Golde et al. 1993; Suzuki, Cheung et al. 1994), and thereby cause early-onset AD. Another example of mutation in APP resulting in AD is Down’s syndrome. APP is located in chromosome 21, and the Down’s syndrome is characterized by the presence of an extra copy of genetic material.
on the chromosome 21, either in whole (trisomy 21) or part. All patients with trisomy 21 develop AD at early age because they get an extra APP gene as well, resulting in over-production of Aβ. The causal relationship is supported by a case involving a patient with Down’s syndrome which resulted from partial translocation of chromosome 21 without APP and the patient did not develop AD (Prasher, Farrer et al. 1998). PSEN1 and PSEN2 are located in chromosomes 14 and 1 respectively. Their gene products are essential components of γ-secretase. Mutations on the two genes resulting in Aβ over-production are also genetic causes of familial AD (Schellenberg, Bird et al. 1992; Levy-Lahad, Wijsman et al. 1995; Duff, Eckman et al. 1996). Interestingly, all the three genes linked to early-onset AD are involved in APP processing. These genetic findings lead to “amyloid hypothesis” that accumulation of Aβ is the primary cause of AD.

Apolipoprotein E (APOE) has been linked to late-onset AD which accounts for about 95% of AD cases and occurs after age 65. APOE gene has three polymorphic alleles as ε2, ε3, and ε4. ε3 is the most (77%) and ε2 the least common allele (Mahley 1988). ε4 allele is a major genetic risk for late-onset AD (Strittmatter, Saunders et al. 1993). The ε4 allele frequency is about 15% in general population but increase to 40% in patients with AD. Individuals with one ε4 allele have three to four fold higher chance to develop AD than those without ε4 allele (Corder, Saunders et al. 1993). APOE is a constituent of many lipoproteins that transport cholesterol and lipids throughout the circulatory system and between cells. The blood-brain barrier (BBB) limits the passage of macromolecules and lipoproteins particles between the peripheral and the central nervous system. The nervous system is capable of de novo synthesis of lipid molecules, but can also redistribute lipoproteins locally for their lipid requirements (Pitas, Boyles et al. 1987; Poirier, Baccichet et al. 1993). In the brain, APOE is mainly produced by glial cells (Boyles, Pitas et al. 1985; Elshourbagy, Liao et al. 1985). Packed with lipoproteins, APOE is a ligand that is recognized by cell receptors that mediate the endocytosis of lipoprotein particles. APOE binds to a group of receptor known as the low density lipoprotein receptor (LDLR) family (Herz and Bock 2002; Herz and Chen 2006).

Several genes have been weakly linked to the late-onset AD, for example, LDLR-related
protein (LRP) (Kang, Saitoh et al. 1997; Beffert, Arguin et al. 1999) and $\alpha_2$-macroglobulin ($\alpha_2$M) (Blacker, Wilcox et al. 1998; Liao, Nitsch et al. 1998), but none has shown as solid association as APOE $\varepsilon$4 allele. The mechanism of APOE $\varepsilon$4 allele being pathogenic to AD is still not fully understood, but several pathways have been investigated.

**APOE and AD**

**Cholesterol Metabolism**
AD may be a result of dysfunction of cholesterol metabolism in brain. The brain contains 25% of all unesterified cholesterol in a body while accounting for only 2% of the body mass (Herz and Bock 2002). Cholesterol is an essential component of cell membranes and myelin sheaths, and is involved in synapse formation and plasticity (Pfrieger 2003). The primary function of APOE is to transport cholesterol and lipid molecules. Astrocyte-derived APOE-containing cholesterol is essential for synapse formation (Mauch, Nagler et al. 2001). Some reports shown that the use of statins, a cholesterol-lowering drug, decreases risk for AD, suggesting an association between cholesterol metabolism and AD risk (Jick, Zornberg et al. 2000; Wolozin, Kellman et al. 2000), although some following studies did not support the statement (Shobab, Hsiung et al. 2005). Whether the three APOE isoforms have functional differences regarding cholesterol metabolism is not clear, but some evidence has shown that APOE4 might be less efficient in transporting brain cholesterol (Rapp, Gmeiner et al. 2006) and promoting cholesterol efflux (Michikawa, Fan et al. 2000). The lower efficiency of cholesterol metabolism might cause synaptic and neuronal dysfunction in an A$\beta$-independent way.

Cholesterol may also influence AD development in an A$\beta$-dependent way. Both cholesterol and statins, cholesterol-lowering drugs, have been shown to modulate APP processing with lower cholesterol level resulting in lower A$\beta$ production (Simons, Keller et al. 1998; Simons, Schwarzler et al. 2002). The results may be related to that cholesterol is enriched in lipid rafts where APP processing by $\beta$- and $\gamma$-secretase is favored (Thinakaran and Koo 2008).
Several studies have shown that APOE expression increases after injury in the brain (Hayashi, Kamada et al. 2006; Xu, Bernardo et al. 2006). The injury-induced upregulation of APOE may help to redistribute lipids and strengthen APOE-mediated signaling for synaptic and neuronal repairs. Glia-derived lipoprotein particles protect central nervous system neurons from apoptosis by a receptor-mediated signaling pathway, and APOE3 shows a greater protective effect than APOE4 (Hayashi, Campenot et al. 2007). In addition, APOE3 increases neurite outgrowth while APOE4 either decreases outgrowth or has no effect (Nathan, Bellosta et al. 1994; Holtzman, Pitas et al. 1995). These results indicate that APOE4 is less effective than APOE3 in maintaining and repairing neuronal synapses. Possibly in aged brains, Aβ oligomers or misfolded tau or other neuronal stresses cause neuron damage, and APOE4 is less effective in repairing injured neurons, thereby resulting in late-onset AD.

**Aβ Production**

APOE and its receptors influence APP processing that produces Aβ, and may hence contribute to AD development. Reports demonstrate that an APOE receptor, LRP, interacts with APP either by direct binding extracellularly (Knauer, Orlando et al. 1996) or via adaptor protein such as FE65 which binds to both LRP and APP intracellularly (Trommsdorff, Borg et al. 1998; Pietrzik, Yoon et al. 2004). Because of the faster rate of LRP endocytosis, this interaction accelerates APP endocytic trafficking to acidic endosomes (Perez, Soriano et al. 1999) where β-secretase is abundantly present and active (Cole and Vassar 2007). Therefore, the increased distribution of APP in endosomes result in increased Aβ production (Ulery, Beers et al. 2000; Cam, Zerbinatti et al. 2005). Overexpression of a functional LRP minireceptor in neurons causes an age-dependent increase in soluble brain Aβ (Zerbinatti, Wozniak et al. 2004). APOE4 has been shown to result in more Aβ production than APOE3. In cultured neurons, ectopic expressions of APP and APOE4 increases Aβ production more than APOE3, and the difference is avoided by preincubating cells with receptor-associated protein (RAP), an LRP binding protein, indicating that the effect is LRP-dependent (Ye, Huang et al. 2005). In addition to LRP, three other APOE receptors have been shown to modulate APP processing. LRP1B reduces Aβ production because of its slow endocytosis rate (Cam,
Zerbinatti et al. 2004). APOER2 reduces Aβ production when F-spondin, an adaptor protein which bridges extracellular interaction between APP and APOER2, is present (Hoe, Wessner et al. 2005). Deletion of Sortilin-related receptor with A-type repeats (SORLA) in mice increases Aβ concentration, suggesting that SORLA promotes APP trafficking to Golgi compartment and hence reduces Aβ production (Andersen, Reiche et al. 2005).

**Aβ Clearance**

Amyloid plaques are a pathological feature of AD, and probably impaired Aβ clearance is a pathogenic event for AD. Aβ can be removed from the brain through two pathways: receptor-mediated clearance by cells or by drainage though the blood-brain barrier (BBB), and endopeptidase-mediated degradation.

Several reports shown that APOE is involved in Aβ clearance through APOE receptors. APOE and LRP and other LRP ligands have been identified in amyloid plaques (Rebeck, Reiter et al. 1993). APOE binds to Aβ via its c-terminal domain overlapping with lipid-binding region. Therefore, Aβ binding to APOE compromises the lipid binding function of APOE (Strittmatter, Weisgraber et al. 1993; Tamamizu-Kato, Cohen et al. 2008). Furthermore, Aβ binding to APOE influences APOE binding to APOE receptors. For example, the presence of Aβ increases binding and internalization of APOE-liposome to hippocampus neurons (Beffert, Aumont et al. 1998). Compared to APOE4, APOE3 has been shown to have a higher affinity for Aβ and to clear Aβ through receptors on the cell surface more efficiently, and hence overall result in lower Aβ accumulation (LaDu, Falduto et al. 1994; Holtzman, Bales et al. 1999; Holtzman, Bales et al. 2000). In addition to be cleared by cell endocytosis and lysed, Aβ can be removed from the brain by crossing the BBB in a receptor-dependent way. Recently, a report shown that at the BBB, Aβ binding to APOE4 redirects its clearance form LRP to very low-density lipoprotein receptor (VLDLR) which has s slower internalization rate. Therefore, VLDLR clears Aβ-APOE4 complex less efficiently than LRP and VLDLR clear Aβ-APOE2 or Aβ-APOE3 complexes at the BBB (Deane, Sagare et al. 2008).
Aβ can be processed directly by endopeptidase-mediated degradation. Several Aβ-degrading enzymes have been identified (Leissring 2008), but neprilysin (NEP) and insulin-degrading enzyme (IDE) are the two most studied Aβ protease that are expressed in both neurons and vascular cells (Bu 2009). Expression of IDE is reduced when APOE ε4 allele is present (Du, Chang et al. 2009), and APOE3 promotes Aβ degradation by NEP and IDE more efficiently than APOE4 (Jiang, Lee et al. 2008). These results indicate that APOE may also play a role in Aβ clearance by proteolytic degradation.

**Tau Phosphorylation**

Hyperphosphorylated tau is the main component of neurofibrillary tangles (NFTs) and is toxic to neurons, suggesting that dysregulation of tau may contribute to AD development as well. The presence of APOE ε4 allele is associated with higher content of NFTs in the neocortex (Nagy, Esiri et al. 1995; Ohm, Kirca et al. 1995). Overexpression of APOE4 in neurons but not in astrocytes increases tau phosphorylation in transgenic mice (Tesseur, Van Dorpe et al. 2000; Brecht, Harris et al. 2004). APOE normally expresses in astrocytes, not neurons, but APOE expression in neurons after injury has been reported (Aoki, Uchihara et al. 2003). Possibly in stressed AD brains, abnormal APOE expression in neurons increases tau phosphorylation (Bu 2009). The mechanism of how APOE influences tau phosphorylation is not clear. One hypothesis is that APOE isoforms may differentially regulate APOE receptor-mediated signaling pathway that thereby modulates the function of tau kinases and phosphatases, and causes tau phosphorylation differentially (Bu 2009). Another hypothesis is that the C-terminal-truncated fragments of APOE enter the cytosol and interact directly with tau and influence tau phosphorylation status (Brecht, Harris et al. 2004).

**APOE Receptors: LDLR family**

APOE binds to members of the low density lipoprotein receptor (LDLR) family. This family represents a group of modular type-I-transmembrane proteins that recognize and bind several ligands in the extracellular space. Once ligands bind, they initiate
endocytosis that transfers the ligand-receptor complex to the endosomal/lysosomal compartments within the cell. After releasing the cargo, the receptors recycle back to the cell surface, and the ligands are further processed or degraded. In addition to function as endocytosis receptors, several LDLR family members also play an important role in a variety of cellular signaling pathways (Jaeger and Pietrzik 2008).

The LDLR family includes LDLR, LRP (a.k.a. LRP1), LRP1B, megalin (a.k.a. gp330 or LRP2), multiple epidermal growth factor like domains 7 (MEGF7, a.k.a. LRP4), VLDLR, Apolipoprotein E receptor-2 (ApoER2, a.k.a. LRP8), LRP5, LRP6 and sortilin-related receptor with A-type repeats (SORLA, a.k.a. SORL1 and LR11). The members contains five characteristic domains: (a) ligand binding type cysteine-rich repeats, (b) epidermal growth factor type cysteine-rich repeats, (c) YWTD-containing β-propeller domain (Springer 1998; Jeon, Meng et al. 2001), (d) a single membrane spanning domain, and (e) a cytoplasmic tail. Except LRP5, LRP6 and SORLA, the cytoplasmic tails of all other members contain one or more ‘NPxY’ (Asp-Pro-any amino acid-Tyr) motifs which function in recruitment of adaptor proteins containing phosphotyrosine binding (PTB) domains. In contrast, the cytoplasmic tails of LRP5 and LRP6 contain five ‘PPPSP’ (Pro-Pro-Pro-Ser-Pro) motifs and SORLA contains one GGA-binding motif. The PPPSP motif is a target sites for glycogen synthase kinase 3 (GSK3) and essential for signaling pathway (Cseleynyi, Jernigan et al. 2008). The GGA-binding motif is a target site for GGA proteins (Golgi associated, gamma-adaptin ear containing, ADP-ribosylation factor binding protein), required for receptor trafficking. Another optional module is the O-linked sugar domain, which is found in some receptors immediately preceding the membrane-spanning segment (Herz and Bock 2002; Bu 2009) (Figure 1.1).

Although the LDLR family members may recognize distinct ligands, they share an ability to bind the receptor-associated protein (RAP). RAP functions as a chaperone for all the receptors in the endoplasmic reticulum where it facilitates proper folding (Bu, Geuze et al. 1995; Obermoeller-McCormick, Li et al. 2001) as well as prevents the premature interaction of ligands with receptors (Willnow, Armstrong et al. 1995; Willnow, Rohlmann et al. 1996).
LDLR and LRP are the two major APOE receptors in the brain. Deletion of the *Ldlr* gene in mice increases APOE levels in the brain parenchyma and cerebrospinal fluid (CSF) (Fryer, Demattos et al. 2005), and similarly, conditional deletion of the *Lrp1* gene in mouse forebrain neurons increases APOE levels (Liu, Zerbinatti et al. 2007). In contrast, overexpression of a functional LRPI minireceptor in the mouse brain decreases brain APOE levels (Zerbinatti, Wahrle et al. 2006). These results indicate that LDLR and LRP play an important role in APOE metabolism. However, they may regulate APOE and lipid metabolism differentially. While deletion of either *Lrp1* gene or *Ldlr* gene increases APOE levels, brain cholesterol levels are unchanged in *Ldlr*-knockout mice (Fryer, Demattos et al. 2005) but reduced in *Lrp1*-knockout mice (Liu, Zerbinatti et al. 2007). APOE–lipoprotein particles secreted by astrocytes have higher affinity for LDLR than for LRPI (Fryer, Demattos et al. 2005), whereas APOE-enriched lipoprotein particles and CSF-isolated HDL particles bind more prone to LRPI1 than to LDLR (Kowal, Herz et al. 1990; Fagan, Bu et al. 1996). In this study, I focus on LDLR because our lab has previously identified several LDLR splice variants that may be associated with AD (see below).

**LDLR**

**Gene Structure**

The *LDLR* gene (GeneID: 3949) is 45 kb and localized on chromosome 19p13.1-13.3. The gene consists of 18 exons, which correspond closely to the structural units of the protein. LDLR is a striking example of a mosaic protein encoded by a gene that was assembled by exon shuffling (Sudhof, Goldstein et al. 1985). The exon 1 encodes 168 bp upstream of the initial methionine codon and the signal sequence. Cysteine-rich repeats 1-2 and 6-7 of the ligand-binding domain are each encoded by individual exons. The other repeats 3-5 are all encoded by exon 4. The epidermal growth factor (EGF) repeats are each encoded by individual exons. Exons 9-13 encode the β-propeller domain. The O-linked sugar domain is encoded by exon 15. The transmembrane domain is interrupted by a single intron. Exon 17 encodes the C-terminal half of the transmembrane domain.
and a part of the cytoplasmic domain. The last exon encodes the remaining cytoplasmic domain and the 3′-untranslated region of the mRNA (Kim, Magoori et al. 1997) (Figure 1.2). The mRNA is 5.3 kb and encodes a protein of 860 amino acids (Yamamoto, Davis et al. 1984).

The LDLR gene promoter region contains regulatory elements consisting of three imperfect direct repeats of 16 bp. Repeat 1 (R1) and repeat 3 (R3) contain Sp1-binding sites that support basal transcriptional activity. Repeat 2 (R2) contains a sterol response element (SRE), a binding site for SRE-binding protein (SREBP). This transcription factor is synthesized in the endoplasmic reticulum as an inactive precursor, and is released to the nucleus by sterol-sensitive proteolysis (Sakai, Rawson et al. 1998) to upregulate transcription of LDLR and other genes in the cholesterol pathway (Sekar and Veldhuis 2004; Attie and Seidah 2005).

Function
LDLR plays a central role in cholesterol homeostasis. The LDLR is the cell surface receptor that regulates plasma cholesterol by mediating endocytosis of LDL, the major cholesterol transport protein in human plasma. LDL particles carry approximately 65-70% of plasma cholesterol in humans (Jeon and Blacklow 2005). Each LDL particle contains a highly-hydrophobic core consisting of esterified cholesterol molecules surrounded by a shell of phospholipids and unesterified cholesterol, as well as a single apolipoprotein B-100 (APOB-100). The LDLR also binds to lipoproteins that contain multiple copies of APOE, such as β-migrating forms of very low-density lipoprotein (β-VLDL), or certain intermediate and high-density lipoproteins (HDL) (Innerarity and Mahley 1978; Weisgraber, Innerarity et al. 1978).

LDLR binds to the APOB- or APOE-containing lipoproteins and initiates endocytosis via clathrin-coated pits, where the receptor molecules cluster on the cell surface (Anderson, Brown et al. 1977; Anderson, Goldstein et al. 1977). The complexes are then delivered to endosomes, where the lipoprotein particles are released because of the low-pH environment. The receptors are then recycled to the cell surface, whereas the lipoprotein
particles are processed upon fusion with lysosomes. The APOB or APOE protein is degraded to amino acids and the cholesterol esters are hydrolyzed to fatty acids and cholesterol. The free cholesterol inhibits expression of LDLR and 3-hydroxy-3-methyl-glutaryl coenzyme A (HMG-CoA) reductase, the rate-limiting step in cholesterol synthesis, forming a negative feedback regulation. Simultaneously, a reciprocal stimulation of cholesterol ester synthesis favors storage and utilization of excess cholesterol (Brown, Dana et al. 1975; Goldstein and Brown 2009).

**LDLR and Familial Hypercholesterolemia**

Familial hypercholesterolemia (FH) is a disease characterized by abnormal elevated concentration of plasma LDL and cholesterol, which lead to excess cholesterol deposition in tissues, resulting in accelerated atherosclerosis and increased risk of premature coronary heart disease (Soutar and Naoumova 2007). A series of papers by Drs. Goldstein and Brown demonstrated that FH results from defects in metabolism of LDL via the LDLR-mediated endocytosis, mostly caused by a loss-of-function mutation in the LDLR gene (Brown and Goldstein 1974; Goldstein and Brown 1974; Goldstein, Dana et al. 1975). FH is an autosomal dominant genetic disorder with a gene-dosage effect. While heterozygous FH is treatable with cholesterol-lowering medicine, homozygous FH shows weak response to medicine. Homozygous FH is very rare ($\leq 1/10^6$), whereas the frequency of heterozygous FH in most populations is about 1/500 and more than 1000 LDLR mutant alleles have been identified in FH subjects. These genetic variants have been reviewed (Hobbs, Brown et al. 1992) and listed online at [http://www.ucl.ac.uk/fh/](http://www.ucl.ac.uk/fh/) (Wilson, Gahan et al. 1998) and [http://www.umd.necker.fr/LDLR/research.html](http://www.umd.necker.fr/LDLR/research.html) (Villeger, Abifadel et al. 2002).

**LDLR and AD**

Several findings indicate that LDLR may be involved in AD development. First, the LDLR gene is located on chromosome 19p13.1-13.3, including a region that has been reported to link with late-onset AD (Wijsman, Daw et al. 2004). Second, LDLR is a main APOE receptor in the brain (Herz and Bock 2002). APOE influences Aβ production (Ye, Huang et al. 2005) or clearance in a receptor-dependent manner (Holtzman, Bales et al. 2002).
In addition, deletion of *Ldlr* gene in mice increases APOE levels in the brain (Fryer, Demattos et al. 2005), indicating that LDLR may affect APOE function even through a receptor-independent pathway. Since *APOE* allele variants are associated with AD (Strittmatter, Saunders et al. 1993), LDLR variants may contribute to AD as well. Third, LDLR is intimately linked with cholesterol homeostasis, and cholesterol metabolism has emerged as an AD modulator. Individuals with elevated levels of plasma cholesterol have higher susceptibility to AD (Jarvik, Wijisman et al. 1995). Reducing cholesterol level decreases Aβ production (Simons, Keller et al. 1998). Forth, LDLR-deficient mice show spatial learning deficits (Cao, Fukuchi et al. 2006), which is a pathological feature of AD patients (Swainson, Hodges et al. 2001). As such, LDLR is a considerable candidate to investigate for AD.

**LDLR Mutations**

Based on the phenotypic effect on the protein, LDLR mutations are divided into five classes (Hobbs, Russell et al. 1990; Jeon and Blacklow 2005). Class 1 mutations result in no protein synthesis due to disruptions of the promoter sequence, nonsense, frameshift, or splicing mutations (null alleles). Class 2 mutations disrupt transport of the LDLR from the endoplasmic reticulum to the Golgi apparatus due to mutations in the ligand-binding and EGF precursor regions (transport-defective alleles). Class 3 mutations are also primarily found in the ligand-binding and EGF precursor regions but interfere with cell surface binding of the receptor to LDL (binding-defective alleles). Class 4 mutations in the cytoplasmic and membrane-spanning domains inhibit the clustering of LDL receptors on the cell surface, so that the bound LDL particle is not internalized (internalization-defective alleles). Class 5 mutations in the epidermal growth factor precursor region prevent the proper release of LDL particles in the endosome and, as a result, the LDL receptor is not recycled to the cell surface (recycling-defective alleles.).

Some of the mutations result in LDLR splicing difference. For example, a mutation located in the end of LDLR intron 12 changes the consensus 3’ splice site, resulting in aberrant splicing (Nissen, Hansen et al. 1997). A study aimed to characterize LDLR mutations of FH in Spain families, and found 14 splicing mutations that affect splicing
donor or acceptor consensus sequences are expected to influence LDLR splicing (Mozas, Castillo et al. 2004). In addition to mutations located in splice sites, polymorphisms that affect splicing regulatory element sequences may change LDLR splicing efficiency.

**Regulation of Splicing**

Alternative splicing is a major mechanism for regulating gene expression and contributing to protein diversity. Estimates of the proportion of human genes that are alternatively spliced range from 35-60% (Mironov, Fickett et al. 1999; Modrek and Lee 2003). Alternative splicing is characterized by utilization of different splice sites. Splicing requires that small nuclear ribonucleoproteins (snRNPs) such as U1, U2, U4, U5 and U6, recognize a conserved 5’ splice site, branch point and 3’ splice site to assemble a spliceosome (Hastings and Krainer 2001). If these sequences vary from consensus sequence, non-snRNP splicing factors are required to achieve efficient splicing.

The two major non-snRNP splicing factor families are the serine/arginine-rich (SR) protein family and heterogeneous nuclear ribonucleoprotein (hnRNP) family. The SR proteins consist of one or two RNA recognition motifs (RRMs) and a carboxyl-terminal domain enriched with serine/arginine repeats (RS domain). The RRMs determine RNA binding specificity, whereas the RS domain mediates protein-protein interactions (Caceres and Kornblihtt 2002). Conventionally, SR proteins recognize exonic splicing enhancers (ESEs) and recruit splicing machinery close to the exon/intron boundary and therefore enhance splicing; a minority of SR proteins act to inhibit splicing (Lin and Fu 2007).

In contrast, hnRNPs were first described as a major group of chromatin-associated RNA-binding proteins (Krecic and Swanson 1999). These proteins consist of at least one RNA binding motif such as an RNA recognition motif (RRM), hnRNP K homology (Defesche, Schuurman et al.) domain or a arginine/glycine-rich (RGG) box as well as auxiliary domains for protein-protein interactions (He and Smith 2009). HnRNPs recognize ESEs or exonic splicing silencers (ESS) (Auclair, Busine et al.) to regulate splicing efficiency.
Additionally, intronic splicing enhancer (ISE) and intronic splicing silencer (Puissant and Houdebine) have also been identified, but the mechanism is not as well understood (reviewed in (Hastings and Krainer 2001; Black 2003)).

Specific Aims

Our lab previously showed that rs688 modulates LDLR exon 12 splicing efficiency (Zhu, Tucker et al. 2007; Zou, Gopalraj et al. 2008). Located in LDLR exon 12, rs688 is a single nucleotide polymorphism (SNP), with a major C and minor T alleles. In Caucasian populations, the minor allele carriers, i.e. C/T or T/T individuals, represent 60–65% of the population. The frequency of rs688T allele carriers in other races varies from 0-17% in African populations to 17-34% in Asian populations. Association studies revealed that rs688 associated with higher cholesterol in pre-menopausal women (Zhu, Tucker et al. 2007) and with Alzheimer’s disease in men (Zou, Gopalraj et al. 2008). I interpret the results as suggesting that enhancing LDLR splicing will lower cholesterol level and decrease AD risk. However, the mechanism of rs688 modulating LDLR exon 12 splicing and the regulation of LDLR splicing are still unclear. Therefore, I propose two specific aims.

1) To identify functional ESE(s) within LDLR exon 12.
Rs688 modulates LDLR splicing efficiency and the change from rs688C to rs688T allele is predicted to influence splicing regulatory elements affinity to splicing factors. Hence, I hypothesize that rs688 is located in or close to functional ESE(s) and thereby influence LDLR exon 12 splicing.

2) To identify splicing factors that modulate LDLR splicing efficiency.
To further investigate the regulation of LDLR splicing, I proposed to identify splicing factors modulating LDLR splicing that may help to understand the involvement of LDLR splicing physiologically and pathologically.
Figure 1.1 LDLR family members

Member of the LDLR family and their characteristic modules are shown. The furin cleavage sites in LRP1 and LRP1B are indicated by arrows. The four clusters of ligand-binding repeats in LRP1 are labeled (I–IV). Highlighted in blue are the two extra sequences in LRP1B (compared with LRP1).

Figure 1.2 Exon organization and protein domains of the human LDLR

The six functional domains of the three proteins are labeled in the lower portion of the figure. The ligand-binding repeats are numbered, and the EGF repeats are lettered A-C. 5′- and 3′-untranslated regions are indicated by solid lines. The positions at which introns interrupt the coding region are indicated by arrowheads. Exon numbers are shown between the arrowheads. Chr., chromosome; EGF, epidermal growth factor.

CHAPTER TWO
IDENTIFICATION OF FUNCTIONAL EXONIC SPLICING ENHANCER
IN LDLR EXON 12

Introduction

Alternative splicing is an important mechanism of gene expression regulation and contributes to protein diversity. Approximately, 70-80% of human genes are alternatively spliced (Johnson, Castle et al. 2003; Kampa, Cheng et al. 2004; Clark, Schweitzer et al. 2007). For efficient splicing, most introns require a conserved 5’ splice site and a branch point sequence followed by a polypyrimidine tract and a 3’splice site. Small nuclear ribonucleoproteins (snRNPs) recognize these regions and assemble spliceosome. Nevertheless, other non-snRNP splicing factors are involved in the process as well. For example, serine/arginine-rich (SR) proteins conventionally recognize exonic splicing enhancer (ESE) and recruit splicing machinery close to the exon/intron boundary, thereby enhancing splicing efficiency. The SR proteins consist of one or two RNA recognition motifs (RRMs) and a carboxyl-terminal domain enriched with serine/arginine repeats (RS domain). The RRM s determine RNA binding specificity, whereas the RS domain mediates protein-protein interactions (Caceres and Kornblihtt 2002). In contrast, heterogeneous nuclear ribonucleoproteins (hnRNPs) recognize exonic splicing silencer (Auclair, Busine et al.) and inhibit splicing. Additionally, intronic splicing enhancer (ISE) and intronic splicing silencer (Puissant and Houdebine) have also been identified, but the mechanism is not as well understood (reviewed in (Hastings and Krainer 2001; Black 2003)).

The interactions between some SR proteins and their specific ESE sequences have been identified in recent years. Cartegni et al. used a method called Systematic Evolution of Ligands by Exponential enrichment (SELEX) (Tuerk and Gold 1990) to identify ESE motifs specific for four SR proteins including SF2/ASF, SC35, SRp40, and SRp55, and developed an online ESE annotation tool called ESEfinder (http://rulai.cshl.edu/tools/ESE2/) (Cartegni, Wang et al. 2003). Fairbrother et al.
compared large sets of human gene sequence and identified 10 predicted ESE motifs which are enriched in human exons relative to introns and are significantly more frequent in exons with weak (non-conserved) splice sites than in exons with strong (conserved) splice sites. They developed an online tool called RESCUE-ESE (http://genes.mit.edu/burgelab/rescue-ese/) (Fairbrother, Yeh et al. 2002; Fairbrother, Yeo et al. 2004). Although there are other mechanisms of splicing regulation, ESEs appear most prevalent and these programs provide a reliable method to predict potentially functional ESEs of interested exons.

Changes in RNA sequence that disrupt functional regulatory cis-elements resulting in aberrant splicing can cause human diseases (reviewed in (Cooper, Wan et al. 2009; Tazi, Bakkour et al. 2009)). For example, mutations N279K and Delta280K, within exon 10 of MAPT (microtubule-associated protein tau) disrupt an ESE, causing exon 10 skipping and contributing to frontotemporal dementia with Parkinsonism linked to chromosome 17 (FTDP-17) (D'Souza and Schellenberg 2006). Another example is spinal muscular atrophy, which is caused by mutations within SMN1 (survival of motor neurons 1); although SMN2 differs from SMN1 by only a single base change, this base change disrupt an ESE in exon 7 of SMN2, resulting in exon 7 skipping, and therefore SMN2 cannot compensate for the loss of SMN1 (Lorson, Hahnen et al. 1999). Targeting the splicing regulation can be an alternative mean for human disease therapy.

Our lab previously identified a single nucleotide polymorphism (SNP), rs688, which modulates low-density lipoprotein receptor (LDLR) exon 12 splicing efficiency and is associated with higher cholesterol in pre-menopausal women (Zhu, Tucker et al. 2007) and with Alzheimer’s disease in men (Zou, Gopalraj et al. 2008). However, the mechanism and the regulation of exon 12 splicing are still unclear. We hypothesized that the rs688 can modulate LDLR splicing efficiency because it locates within or close to functional ESE(s) in LDLR exon 12. Here, to identify functional ESEs within LDLR exon 12, I performed a serious of mutations that neutralize affinity of predicted ESEs to associated SR proteins. I found two neighboring ESEs that flank rs688 and are critical to LDLR splicing.
Materials and Methods

Cell culture
HepG2 (human hepatocellular carcinoma) cells were cultured in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin at 37°C in a humidified 5% CO2-95% air atmosphere.

Construction of LDLR minigenes
An LDLR minigene containing exons 9-14 with rs688C or rs688T in a pcDNA3.1 backbone was previously described (Zhu, Tucker et al. 2007). I also created an additional LDLR minigene containing exon 12 in pSPL3b (the kind gift from Genzyme Corporation, MA). In addition to exon 12, this minigene included 219 bp of 5’ intronic LDLR sequence and 123 bp of 3’ intronic LDLR sequence. This minigene was generated by using PCR (Platinum Taq, Invitrogen) to amplify with primers EcoRI-Intron11F 5’GCGGAATTCTGAAGTTTTTCTGACCTGCA 3’ and BamHI-Intron12R 5’CGCGGATCCATAACTCAGGTCTAAGACCT3’. PCR fragments were cloned into pCR2.1 TOPO T/A cloning vector (Invitrogen). The TA clones and the splicing vector pSPL3b were digested with EcoRI/BamHI and, after gel-purification, the LDLR fragments were ligated into the vector. The pSPL3b backbone contains rabbit β-globin coding sequences separated by a portion of the intron of the HIV-tat gene (Buckler, Chang et al. 1991) (Burn, Connors et al. 1995). The integrity of the LDLR exon 12 minigenes was confirmed by sequencing.

Evaluation of LDLR minigene splicing
LDLR minigene splicing efficiency was evaluated by transfecting the clones into HepG2 cells by using FuGENE 6 as directed by the manufacturer (Roche Diagnostics, Basel, Switzerland). Briefly, 1.5 x 10^5 cells 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 RT–PCR as previously described (Zhu, Tucker 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′CATCGTGTTGGATCCTGTTCC3′ and, to obviate endogenous LDLR, a vector-specific antisense primer 5′GGGATAGGCTTACCTCGAA3′. 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. 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′TCTCAGTCACCTGGACAAC and 3′CCACACCAGCCACCTTCT3′, 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 30 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 and inefficiently spliced LDLR isoforms were quantified by fluorescence intensity. 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.

**Mutagenesis of conversed putative ESE**

Sequences of LDLR exon 12 across human, macaque, mouse, hamster, cow, wild boar, rat, and rabbit were analyzed by VISTA (http://genome.lbl.gov/vista/index.shtml) and ESEfinder (http://rulai.cshl.edu/tools/ESE2/) (Cartegni, Wang et al. 2003). If a putative ESE as defined by ESEfinder was conserved between more than five species, I considered it as 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 conserved putative ESEs to SR proteins, a series of primers were designed that
reduced the original binding affinity. To verify \textit{in silico} that these new sequences neutralized the putative ESE sequences, each was computationally evaluated by ESEfinder and ESE-RESCUE (http://genes.mit.edu/burgelab/rescue-ese/) (Fairbrother, Yeh et al. 2002). Also, to ensure that the mutations did not introduce a putative exonic splicing suppressor (Auclair, Busine et al.), the mutated sequences were analyzed by FAS-ESS (http://genes.mit.edu/fas-ess/) (Wang, Rolish et al. 2004). The primers listed in Table 2.2 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 contain the desired mutation, and extended by \textit{PfuUltra} high fidelity DNA polymerase. After temperature cycling, the product was treated with \textit{Dpn I} endonuclease which digests methylated and hemimethylated DNA, the parental DNA template. Then, the unmethylated vector DNA containing the desired mutations was 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.

\textbf{Prediction of RNA secondary structure}  
LDLR wild type and mutated sequences were analyzed by using NIPU software (http://www.bioinf.uni-freiburg.de/Software/NIPU/index.html) (Hiller, Zhang et al. 2007).

\textbf{Statistical analysis}  
The effects of rs688 alleles on LDLR minigene splicing efficiency were evaluated by a paired T-test because of inter-experiment variability. 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 16).

\textbf{Results}
**Generation of LDLR exon 12 minigene**

To identify sequences within LDLR exon 12 that modulate its splicing efficiency, I cloned 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 2.1). I planned to use this single-exon minigene in combination with site-directed mutagenesis to discern exon 12 sequences that modulate exon 12 splicing. To validate the single exon minigene, I compared its splicing relative to a previously reported LDLR exon 9-14 minigene (Figure 2.2A). This comparison was performed by transfecting minigenes containing each rs688 allele into HepG2 cells and quantifying splicing by reverse-transcriptase PCR (RT-PCR). The LDLR minigene containing only exon 12 showed a modest but significant effect for rs688 on splicing, i.e., rs688T increased the percentage of LDLR mRNA that lacked exon 12 (Delta 12) by 3.7 ± 0.6% (n = 3, p = 0.009; Figure 2.2B). For the LDLR exon 9-14 minigene, the rs688T allele increased the percentage of LDLR mRNA that lacked exon 12 by 12.6 ± 4.6% (n = 4, p < 0.01; Figure 2.2C). I noted that rs688 also influenced the percentage of LDLR that lacked both exons 11 and 12 (Delta 11+12), although to a smaller effect, i.e., the rs688T allele increased the percentage of LDLR mRNA lacking both exon 11 and 12 by 4.5 ± 1.9% (n = 4, p < 0.01; Figure 2.2C). In summary, the single exon minigene demonstrates rs688 effects that are qualitatively similar to those seen in the exon 9-14 minigene. Hence, I proceeded to use the single exon minigene to identify critical exonic elements that modulate exon 12 splicing.

**Identification of conserved putative ESEs**

Putative exon 12 ESEs that may modulate exon 12 splicing were identified in a two-step process. Exon 12 ESEs in eight species were first predicted by using ESEfinder. Then, these sequences were 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 2.3; Table 2.1). The location of rs688 is within two putative SRp40 binding sites, and the change from C to T allele is predicted to influence the binding affinity. To distinguish these two
putative SRp40 binding sites even though their sequences overlapped, I considered them as separate ESE regions and designated them as R5 and R6 respectively.

**Identification of functional ESEs**

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 2.4) by using site-directed mutagenesis with particular primers (Table 2.2). Wild type (WT) and each mutant (MT) minigene were then transfected into HepG2 cells and the splicing efficiency was evaluated by RT-PCR. Most of the mutations targeting conserved putative ESE regions did not alter exon 12 splicing efficiency. However, MT6 and MT7 significantly increased the percentage of Delta 12 from 18.2 ± 5.5% in the WT LDLR minigene to 64.3 ± 14.3% and 43.4 ± 13.1%, respectively (both \( p < 0.001 \); Figure 2.5). Hence, the sequences targeted in MT6 and MT7 appear critical for LDLR exon 12 splicing.

To discern whether exon 12 may be skipped because of a weak splice site donor, I also evaluated the effects of optimizing the 5’ splice site of intron 12 from GAGgtgtgg to a sequence more similar to the consensus splice donor site sequence of CAGgtgtragt (Mount 1982). Hence, I generated MT10 with the sequence CAGgtgtgg and MT11 with the sequence GAGgtgtagt (Figure 2.4). Analyzed by Alternative Splicing Database (ASD), these changes increased the donor site score strength from 6.67 in the WT minigene to 7.67 for MT10 and 10.46 for MT11. When I analyzed the effects of MT10 and MT11 on splicing efficiency, I found that MT10 had essentially no effect on exon 12 splicing, i.e., the percentage of Delta 12 is 14.1 ± 0.1% (\( p = 0.631 \); Figure 2.5). However, MT11 showed a clear trend towards decreased percentage of Delta 12, i.e., 2.6 ± 0.6% (\( p = 0.077 \); Figure 2.5). The trend associated with MT11 may reflect 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 minigene or applicable to a more physiologic context, I introduced the MT6 and MT7 into the LDLR exon 9-14 minigene. I also evaluated the effects of MT5 as a negative control. Splicing efficiency of the 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 ± 1.7%, compared to 18.3 ± 0.3% in the WT LDLR
minigene. In contrast, MT6 and MT7 significantly increased the percentage of delta 12 to
46.4 ± 4.8% and 46.4 ± 5.4%, respectively (both p < 0.001; Figure 2.6). 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.

**Evaluation of change in RNA secondary structure**

Although the results of these studies strongly suggest that regions 6 and 7 contain
elements critical for splicing regulation on the basis of ESEs, I also considered the
possibility that the mutations that were introduced may alter pre-mRNA secondary
structure. Therefore, I 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 diminish single-stranded
regions, including MT2, MT4 and MT9 (Figure 4). Interestingly, these mutations did not
alter LDLR minigene splicing. Conversely, MT6 and MT7 were not predicted to change
the pre-mRNA secondary structure and yet strongly decreased splicing efficiency. Hence,
the decreases in splicing efficiency observed in MT6 and MT7 were not because of
changes on RNA secondary structure but rather likely reflect the disruption of regulatory
motifs.

**Evaluation of c.1784G>A effect on LDLR minigene splicing**

In addition to rs688, I suspected that other mutations located in or close to the functional
ESE regions may also modulate LDLR exon 12 splicing. By searching the Universal
Mutation Database ([http://www.umd.necker.fr](http://www.umd.necker.fr)) (Varret, Rabes et al. 1998), I found an
LDLR human mutation, c.1784G>A, which has been reported in Korean patients with
familial hypercholesterolemia (Kim, Choi et al. 2004), and is located within the
regulatory region defined by MT7. This mutation changes the encoded amino acid from
arginine to glutamine, but is also predicted to decrease ESE affinity toward SR proteins
by ESEfinder (Figure 2.8A). To elucidate whether this mutation influences LDLR exon
12 splicing, I generated an LDLR exon 9-14 minigene containing the mutation, and
compared minigene splicing efficiency to wild type. The results revealed no significant change in splicing (Figure 2.8B), indicating that this mutation causes LDLR malfunction and familial hypercholesterolemia probably by amino acid substitution, not by aberrant splicing.

**Discussion**

In this study, I focused on identifying ESE regions that are functional to LDLR exon 12 splicing. By generating mutations that disrupt binding affinity of putative ESE regions to SR proteins, I identified two ESE regions defined by MT6 and MT7 as critical to LDLR exon 12 splicing. These two ESE regions are c.1769 - c.1772 and c.1784 - c.1787 respectively. They are separated by only 11 bp, and interestingly, rs688 at c.1773 is located within this 11 bp. Hence, this region appears critical for LDLR exon splicing efficiency.

While sequence-based prediction programs such as ESEfinder, RESCUE-ESE, FAS-ESS, and Splicing Rainbow provide useful information regarding investigating splicing regulation, several concerns about this approach have arisen from this study. First, ESE prediction has a lot of positive results, increasing extraneous work. By using ESEfinder, in the 140-bp human LDLR exon 12, twenty-five putative ESEs were predicted. This is probably because most binding sites for splicing factors are degenerate and overlapping. To refine the prediction, I compared sequences crossing eight species, and identified nine conserved putative ESE regions. Among them, minigene analyses confirmed two to be functional in splicing. Second, predicted binding protein may not be the functional protein. The region defined by MT6 is predicted to interact with SRp40 and the region defined by MT7 is predicted to bind SC35 or/and SF2/ASF by ESEfinder. However, my work that focused on identifying splicing factors which modulate LDLR splicing found that ectopic expression of SRp40 or SF2/ASF did not enhance LDLR splicing, but decreased, indicating that other splicing factors interact with the two regulatory elements. Additionally, the G-rich region between the two functional ESE regions matches binding site for hnRNP F and H, but ectopic expression of these two splicing factor did not
influence minigene splicing efficiency. The inability to actualize ESE prediction has been reported (Lastella, Resta et al. 2004; Auclair, Busine et al. 2006). These results revealed the limitation of the prediction programs. Splicing regulation is more complicated than protein-RNA recognition that protein-protein interaction is also critical to final splicing decision. While keep exploring binding sites for splicing factors, work also needs to be done to understand the crosstalk and the regulation of the splicing factors.

Another filter that can be used to refine an ESE prediction is the pre-mRNA secondary structure since it also contributes to exon recognition (Muro, Caputi et al. 1999; Hiller, Zhang et al. 2007). The majority of functional splicing regulatory elements are exposed in the loop of a hairpin RNA structure. When a well-defined ESE is placed in the stem of a RNA structure, the ESE may lose the ability to modulate splicing since the binding site is not exposed and therefore cannot be recognized by regulatory proteins. However, a functional regulatory element is not necessarily located in a single-stranded region. A few verified motifs are located in a double-stranded conformation. For example, an intronic splicing enhancer, GGGGATGGG, in growth hormone 1 (GH1) is predicted located in a double-stranded region by NIPU (Hiller, Zhang et al. 2007). In this study, the two ESE regions critical to LDLR exon 12 splicing seem to be in the latter category. By using NIPU, they are predicted to be located in double-stranded regions (Figure 2.7). The mechanism of how splicing factors recognize regulatory cis-elements that are located in double-stranded region is still unclear.

To understand the splicing regulation of LDLR exon 12 is important. First, rs688 is located close to the functional ESEs to modulate exon 12 splicing and associated with higher cholesterol and Alzheimer’s disease in some case-control populations (Zhu, Tucker et al. 2007; Zou, Gopalraj et al. 2008). Knowing the mechanism to improve LDLR splicing efficiency may provide insight for disease therapy or even prevent disease development. Second, optimizing the 5’ splice site of the LDLR intron 12 increased splicing efficiency close to 100% (Figure 2.5), indicating that the exon 12 is probably skipped because of weak splice donor site. Hence, the key of exon 12 splicing decision resides in the interaction between functional ESEs and the regulatory proteins. Although
the ESEs are predicted to bind SRp40 and SF2/ASF respectively, co-transfection of their expression vectors with LDLR minigene did not change the minigene splicing efficiency. Hence, more work is necessary to identify associated SR proteins.

In this study, by mutagenesis of putative conserved ESE regions, I identified two functional ESE regions that are important to LDLR exon 12 splicing, and that flank rs688. Hence, this region is critical for LDLR exon 12 splicing.
The LDLR exon 12 sequence is shown in bold upper case, and the flanking intronic sequences were shown in lower case. The red letter indicates the location of rs688 within exon 12, and the underlined sequences represent recognition sites for restriction enzymes, *EcoRI* and *BamHI*, respectively.

**Figure 2.1 Cloned LDLR exon 12 and flanking intronic sequences**
Figure 2.2 Rs688 effects upon LDLR exon 9-14 and exon 12 minigene

LDLR minigenes carrying rs688C or rs688T alleles were transfected into HepG2 cells and mRNA was isolated after 24 hours and converted to cDNA for analyses. PCR was performed with primers indicated as arrows in minigene schemes (A). When LDLR exon 12 minigene was transfected, rs688T allele significantly increased the percentage of delta...
12 (B, mean ± SD, n = 3, p = 0.009). When LDLR exon 9-14 minigene was transfected, the rs688T allele significantly increased the percentage of Delta 12 and Delta 11+12 (B, mean ± SD, n = 4, p < 0.01). Representative gel images and quantitative results were shown. Each line indicated a biological replicate.
Figure 2.3 Identification of conserved putative ESE regions

The putative ESEs within human LDLR exon 12 as discerned by ESEfinder are shown (A). Putative ESEs that were conserved in at least five of the species were considered putative ESE regions (B). These regions, R1 – R9, are denoted in A by underlining and in B by the shaded boxes. The position of rs688 is boxed. * represents a nucleotide that is conserved within the 8 species.
Table 2.1 Conserved putative ESE regions in LDLR exon 12

Putative ESEs crossing 8 species were predicted by using ESEfinder. Positions are shown as nucleotide numbers in LDLR exon 12. Conservation is shown as number of conserved species/total compared species (8), and followed by a list of species lacking the putative ESE.

<table>
<thead>
<tr>
<th>Conserved Putative ESE Region</th>
<th>Position of Putative ESE Region</th>
<th>Putative ESE Target</th>
<th>Position of Putative ESE</th>
<th>Putative ESE Sequence</th>
<th>Conservation</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>18-22</td>
<td>SC35</td>
<td>12-19</td>
<td>GGCCGCCT</td>
<td>7/8 - mouse</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SRp55</td>
<td>15-20</td>
<td>CGCCTC</td>
<td>6/8 - mouse, pig</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SC35</td>
<td>15-22</td>
<td>CGCCTCTA/CGTCTCTA</td>
<td>7/8 - pig</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SC35</td>
<td>18-25</td>
<td>CTCTACTG/CTTTACTG</td>
<td>5/8 - mouse, hamster, rat</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SRp40</td>
<td>20-26</td>
<td>CTATTGG/CTACTGG/TTACTGG</td>
<td>8/8</td>
</tr>
<tr>
<td>R2</td>
<td>31-35</td>
<td>SC35</td>
<td>30-37</td>
<td>GATTCCAA/GACTCCAA/GACTCTAA</td>
<td>8/8</td>
</tr>
<tr>
<td>R3</td>
<td>44-48</td>
<td>SRp40</td>
<td>41-47</td>
<td>CCACCTCC/CCACTCT/TCACCTCC</td>
<td>6/8 - bovine, rabbit</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SRp55</td>
<td>45-50</td>
<td>TCCATC</td>
<td>6/8 - mouse, hamster</td>
</tr>
<tr>
<td>R4</td>
<td>53-56</td>
<td>SC35</td>
<td>48-55</td>
<td>ATCTCCAG</td>
<td>6/8 - human, monkey</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SRp40</td>
<td>49-55</td>
<td>TCTCCAG/TCTCAAG</td>
<td>8/8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SRp55</td>
<td>54-59</td>
<td>AGCATC</td>
<td>7/8 - hamster</td>
</tr>
<tr>
<td>R5</td>
<td>62-63</td>
<td>SRp40</td>
<td>62-68</td>
<td>CGTCAAC/GTGCAAC</td>
<td>5/8 - hamster, bovine, rat</td>
</tr>
<tr>
<td>R6</td>
<td>64-67</td>
<td>SRp40</td>
<td>64-70</td>
<td>TCAATGG/TCAACGG</td>
<td>8/8</td>
</tr>
<tr>
<td>R7</td>
<td>79-82</td>
<td>SF2/ASF</td>
<td>75-81</td>
<td>AACCGGA</td>
<td>5/8 - mouse, hamster, rat</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SF2/ASF</td>
<td>78-84</td>
<td>CGGAAGA</td>
<td>6/8 - mouse, rat</td>
</tr>
<tr>
<td>R8</td>
<td>92-95</td>
<td>SF2/ASF</td>
<td>91-97</td>
<td>TGGAGGA</td>
<td>7/8 - hamster</td>
</tr>
<tr>
<td>R9</td>
<td>118-121</td>
<td>SRp40</td>
<td>113-119</td>
<td>TCACCC/CCACCCC</td>
<td>5/8 - bovine, pig, rabbit</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SF2/ASF</td>
<td>114-120</td>
<td>CACCCCT</td>
<td>8/8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SRp40</td>
<td>119-125</td>
<td>CTTCTCC</td>
<td>6/8 - bovine, rabbit</td>
</tr>
</tbody>
</table>
Figure 2.4 Mutagenesis of conserved putative ESEs

The mutations of conserved ESE regions are shown, with putative ESE regions in shaded boxes and mutation sequences in bold font on top (A). 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 (B); I 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 of LDLR intron 12.
<table>
<thead>
<tr>
<th>Mutation</th>
<th>Mutation Primer Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>MT1</td>
<td>CTCCTCAGTGGCCGCGgCGCtCTGGGTGACTCCAAAC</td>
</tr>
<tr>
<td>MT2</td>
<td>TGGCCGCTCTACTGGGTTG<strong>CAGTT</strong>AAAACCCTACTCCATCTC</td>
</tr>
<tr>
<td>MT3</td>
<td>GGGTTGACTCCAAACTTTCA<strong>TGCgC</strong>TCTCAAGCATCGATGTCAAC</td>
</tr>
<tr>
<td>MT4</td>
<td>CAAAATTTCACTCCATCTC<strong>GGCA</strong>ATCGATGTCAACGGGGG</td>
</tr>
<tr>
<td>MT5</td>
<td>ATCTCAAGCATCGAA<strong>AAT</strong>TAACCGGCGGCAACCG</td>
</tr>
<tr>
<td>MT6</td>
<td>TCTCAAGCATCGATGG<strong>GAGT</strong>CGGGCGGCAACCGGAAG</td>
</tr>
<tr>
<td>MT7</td>
<td>TGTCACGCGGGCAACC<strong>TAGT</strong>GACCATCTTGGAGGATG</td>
</tr>
<tr>
<td>MT8</td>
<td>AAGAGGCTGGCCACCT<strong>TAGC</strong>CTCCTTGGCCGCTTCTT</td>
</tr>
<tr>
<td>MT9</td>
<td>CCGGAAGACCATCTT<strong>CAGT</strong>GATGAAAAGAGGCTGGCCC</td>
</tr>
<tr>
<td>MT10</td>
<td>TTGGCCGTCCTTT<strong>CAG</strong>Ggtgtgcttacg</td>
</tr>
<tr>
<td>MT11</td>
<td>GGCCGTCTTTGAGg<strong>agt</strong>tcctagctagagatgc</td>
</tr>
<tr>
<td>c.1784</td>
<td>CAACGGGGGGAACCATCTTGG</td>
</tr>
</tbody>
</table>

Table 2.2 Mutation Primers

Mutation positions are underlined and in bold font.
Figure 2.5 Effects of mutations of putative ESEs on LDLR exon 12 minigene splicing

Wild type (WT) and mutant (MT) clones were transfected into HepG2 cells, and splicing efficiency was evaluated by RT-PCR. While MT1 to MT9 were targeted putative ESE regions, MT10 and MT11 were introduced to optimize 5’ splice site of intron 12. The effects of the ESE mutations on LDLR minigene splicing are shown as representative gel images (A) and quantitative results (B, mean ± SD, n = 3). MT11 showed a trend towards increasing splicing efficiency that the percentage of delta 12 is close to 0 (p = 0.077). While most ESE mutations didn’t change splicing efficiency, MT6 and MT7 significantly increased delta 12 (p < 0.001 and p = 0.003 respectively).
Figure 2.6 Effects of putative ESE mutations on LDLR exon 9-14 minigene splicing

The effects of the ESE mutations on LDLR exon 9-14 minigene splicing are shown as representative gel images (A) and quantitative results (B, mean ± SD, n = 3). The results were similar to the effect on the LDLR exon 12 minigene. MT6 and MT7 significantly decreased splicing efficiency (both p < 0.001).
Figure 2.7 Predicted effects of mutations of putative ESEs on RNA secondary structure

Wild type and mutation 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 single-strandedness. Mutated regions were highlighted in grey boxes.
Figure 2.8 c.1784G>A has no effect on LDLR minigene splicing

The mutation, c.1784G>A is predicted to change ESE affinity to SR proteins (A). Red boxes represent ESE affinity to SF2/ASF, blue box reflects ESE affinity to SC35, and green box represents ESE affinity to SRp40. The c.1784G>A had no effect on LDLR exon 9-14 minigene splicing. Quantitative results were shown (B, mean ± SD, n = 3).
CHAPTER THREE
IDENTIFICATION OF SPLICING FACTORS THAT MODULATE
LDLR SPICING

Introduction

To achieve efficient splicing, in addition to snRNPs recognizing consensus splicing sites, non-snRNP splicing factors binding to splicing regulatory elements are also required to assemble the spliceosome. The two major non-snRNP splicing factor families are the serine/arginine-rich (SR) protein family and heterogeneous nuclear ribonucleoprotein (hnRNP) family. The SR proteins consist of one or two RNA recognition motifs (RRMs) and a carboxyl-terminal domain enriched with serine/arginine repeats (RS domain). The RRMds determine RNA binding specificity, whereas the RS domain mediates protein-protein interactions (Caceres and Kornblihtt 2002). Conventionally, SR proteins recognize exonic splicing enhancers (ESEs) and recruit splicing machinery close to the exon/intron boundary and therefore enhance splicing; a minority of SR proteins act to inhibit splicing (Lin and Fu 2007).

In contrast, hnRNPs were first described as a major group of chromatin-associated RNA-binding proteins (Krecic and Swanson 1999). These proteins consist of at least one RNA binding motif such as an RNA recognition motif (RRM), hnRNP K homology (Defesche, Schuurman et al.) domain or an arginine/glycine-rich (RGG) box as well as auxiliary domains for protein-protein interactions (He and Smith 2009). HnRNPs recognize ESEs or exonic splicing silencers (Auclair, Busine et al.) to regulate splicing efficiency.

Some splicing factors are extensively regulated by phosphorylation. For example, phosphorylation of SF2/ASF modulates its cellular localization (Caceres, Screaton et al. 1998). In the cytosol, activated serine/threonine-protein kinase 1 (SRPK1) and SRPK2 phosphorylate SF2/ASF (Wang, Lin et al. 1998), and thereby trigger SF2/ASF to translocate to the nucleus, where it can be further phosphorylated by CDC-like kinase 1 (CLK1) and redistribute from nuclear speckles to splicing-active sites (Colwill, Pawson
et al. 1996). After splicing, SF2/ASF is dephosphorylated by protein phosphatase 1 (PP1), which facilitates its nuclear export (Stamm 2008). Similarly, cellular localization of hnRNP A1 is also regulated by phosphorylation, but through phosphorylation of its transport partner, F-peptide (Allemand, Guil et al. 2005). In addition, phosphorylation may influence activity of splicing factors. For instance, phosphorylation of SF2/ASF increases its binding to U1 70-kDa snRNP and enhance splicing (Xiao and Manley 1997).

Changes in RNA sequence that disrupt functional regulatory cis-acting elements can result in aberrant splicing and thereby cause human disease (reviewed in (Cooper, Wan et al. 2009; Tazi, Bakkour et al. 2009)). For example, mutations affecting RNA splicing are the most common cause of neurofibromatosis type 1 (NF1) (Ars, Serra et al. 2000) while mutations within and around exon 10 of microtubule-associated protein tau (MAPT) that disrupt exon 10 splicing regulation cause frontotemporal dementia with Parkinsonism linked to chromosome 17 (FTDP-17) (Liu and Gong 2008). These findings suggest that agents that alter splicing regulation can be a novel means of human disease therapy.

Our lab previously identified a single nucleotide polymorphism (SNP), rs688, which modulates low-density lipoprotein receptor (LDLR) exon 12 splicing efficiency and is associated with higher cholesterol in pre-menopausal women (Zhu, Tucker et al. 2007) and with Alzheimer’s disease in men (Zou, Gopalraj et al. 2008). I infer from the results that enhancing LDLR exon 12 splicing efficiency will lower cholesterol and reduce AD in relevant populations. However, the regulation of exon 12 splicing is still unclear. Here, I report that SRp20, SRp38 and hnRNP G decrease LDLR splicing in vitro. Moreover, SRp38 expression level is associated with LDLR splicing efficiency in vivo. I interpret these results as suggesting that SRp38 is a major modulator of LDLR splicing.

Materials and Methods

Cell culture
HepG2 (human hepatocellular carcinoma) cells were grown in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin at 37°C in a humidified 5% CO₂-95% air atmosphere.

**Vectors**

LDLR minigene containing exons 9-14 with rs688T or rs688C in pcDNA3.1, were previously described (Zhu, Tucker et al. 2007); the rs688T vector was derived from the rs688C by using site-directed mutagenesis (QuikChange Site-Directed Mutagenesis Kit, Stratagene) according to the manufacturer’s directions. Vectors encoding hnRNP F and hnRNP H, as well as their parent vector pcDNA4, were generous gifts from Dr. Paula Grabowski (University of Pittsburgh) (Han, Yeo et al. 2005). Vectors encoding the other splicing factors used here, as well as a pEGFP-C2 negative control vector encoding only EGFP were kindly provided by Dr. Stefan Stamm (University of Kentucky). These vectors, most of which encode the splicing factor as a fusion protein with EGFP, have been previously characterized (Screaton, Caceres et al. 1995; Novoyatleva, Heinrich et al. 2008; Heinrich, Zhang et al. 2009).

**Evaluation of LDLR minigene splicing efficiency in vitro**

Splicing factor effects on LDLR splicing efficiency were evaluated by co-transfecting vectors encoding each splicing factor with the LDLR exon 9-14 or exon 12 minigene (rs688C or rs688T) into HepG2 cells. Splicing factor effects were considered relative to LDLR minigene co-transfected with either pEGFP or “empty” pcDNA4 (similar results were obtained with each negative control). Vectors were transfected by using FuGene 6 transfection reagent as directed by the manufacturer (Roche Applied Sciences, Switzerland). Briefly, 1.5 x 10⁵ cells were seeded in a 6-well plate in 2 ml of medium without antibiotics one day before performing the transfection. The next day, 1 µg of LDLR minigene and 1 µg of splicing factor construct were mixed with 6 µl of FuGENE reagent in 94 µl of Opti-MEM (GIBCO) for 30 minutes and added to cell culture. Twenty-four hours later, transfection efficiency was evaluated by observing green fluorescence of EGFP-fused splicing factors by microscope (Nikon Diaphot 200, Japan). Then, total RNA was isolated (RNeasy Mini Kit, QIAGEN) and analyzed for LDLR
splicing patterns by reverse transcriptase-PCR (RT–PCR) as previously described (Zhu, Tucker et al. 2007). One µg of 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’CATCGTGGTGATCCTGTTC3’ and, to obviate endogenous LDLR, a vector-specific antisense primer 5’GGGATAGGCTTACCTTCGAA3’. 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. 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 30 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 (PAGE) 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 and inefficiently spliced LDLR isoforms were quantified by fluorescence intensity. For each sample, fluorescence values were corrected for background and normalized for length differences among amplicons. Sample splicing efficiency was then quantified as the amount of FL LDLR PCR product containing exons 10-14 divided by the total LDLR PCR product for that sample.

For experiments evaluating the dose-dependence of the splicing factors, 0.01, 0.1, or 1 µg of vectors encoding the splicing factor (total amount made up to 1 µg with “negative control” pcDNA4) was co-transfected with 1 µg of LDLR minigene containing rs688C. Twenty-four hours after transfection, total RNA was collected and splicing efficiency was evaluated by RT-PCR as described above.
Effect of splicing factors on mutant LDLR minigene splicing was evaluated by co-
transfecting 1 µg of vector encoding the splicing factor with 1 µg of wild type (WT) or
mutant (MT) LDLR minigenes. Twenty-four hours after transfection, total RNA was
collected and splicing efficiency was evaluated by RT-PCR as described above.

**Western blot**

Twenty-four hours after transfection, HepG2 cells were washed with 1 ml of room
temperature phosphate buffered saline (PBS) and lysed in 80 µl of RIPA butter (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. Lysates were pooled
from three wells for each sample and centrifuged at 10,000xg for 10 minutes at 4°C. Fifty
µl of protein extract was mixed with 10 µl of 6X SDS sample loading buffer containing
β-mercaptoethanol, boiled for 5 minutes and subjected to electrophoresis on a 12%
polyacrylamide gel. Proteins were transferred to nitrocellulose membranes (Bio-Rad).
The blots were then blocked with 5% non-fat powdered milk in PBS for 1 hour at room
temperature and probed overnight at 4°C with a mouse anti-GFP antibody (1:200
dilution; Cat. No. 11814460001, Roche Applied Science) or with a mouse anti-Actin
antibody (1:200 dilution; Cat. No. sc-8432, Santa Cruz Biotechnology) or with a rabbit
anti-SRp38 antibody (0.5 µg/ml; Cat. No. ARP41083_P050, Aviva Systems Biology).
After washing with 0.1% Tween-20 in PBS four times for 5 minutes each, the blots were
incubated with peroxidase-conjugated sheep anti-mouse antibody (1:1,000 dilution,
Jackson ImmunoResearch) or peroxidase-conjugate goat anti-rabbit antibody (1:50,000
dilution, Jackson ImmunoResearch) for 1 hour at room temperature. Bound peroxidase
was visualized by using a SuperSignal West Pico kit (Pierce) and a molecular imager
(ChemiDoc XRS System, Bio-Rad).

**Cell viability after cantharidin treatment**

HepG2 cells were grown with media containing 0, 0.7, 2, or 6 µM of cantharidin for 24
hours, and then stained with 5 µg/ml of calcein-AM (Molecular Probes, Invitrogen) and
10 µg/ml of propidium iodide (PI) (Molecular Probes, Invitrogen) for 20 min at 37°C.
The fluorescence of calcein-AM and PI was observed by microscopy (Nikon Diaphot 300, Japan). The number of viable and dead cells was counted in 4 different microscopic fields at 20X magnifications, and the ratio of viable cells to total cells was estimated.

**siRNA knockdown of SRp38**

Two siRNAs targeting SRp38 were predesigned by Integrated DNA Technologies (Varret, Rabes et al.) Inc. The sequences are

\[
5’-\text{ACAGUAGACCGACUGGAAGACCACG}3’ \text{ for siRNA 1 (HSC.RNAI.N054016.10.8.2)}
\]

\[
5’-\text{AGACUUGCGGCGUGAAUUUGGUCGT}3’ \text{ for siRNA 2 (HSC.RNAI.N054016.10.8.2)}.\]

The two siRNAs sequences correspond to SRp38 exons 5 and 2 respectively. Since SRp38 and SRp38-2 share the same exons 1-5, both SRp38 and SRp38-2 were targeted. A scrambled siRNA as a negative control (NC siRNA) and a positive control siRNA targeting hypoxanthine-guanine phosphoribosyltransferase (HPRT siRNA) were also purchased from IDT. One µg of vector encoding SRp38 or LDLR exon 9-14 minigene was co-transfected with respective concentration of siRNAs (0.1, 1, 2 nM) by using Lipofectamine 2000 transfection reagent as directed by the manufacturer (Invitrogen). Briefly, 4 x 10^5 cells were seeded in a 6-well plate in 2 ml of medium without antibiotics one day before performing the transfection. The next day, respective siRNA concentrations in 250 µl of Opti-MEM were prepared and mixed with 2 µg of LDLR minigene. Ten µl of Lipofectamine 2000 in 250 µl of Opti-MEM was prepared and added to the siRNA mixture. Solutions was incubated for 30 minutes and added to cell culture. Twenty-four hours later, total RNA was collected, and LDLR splicing efficiency was evaluated by RT-PCR as described above. SRp38 mRNA expression was quantified by real-time RT-qPCR. Total protein was also collected for SRp38 western blot.

**shRNA knockdown of SRp38**

Three shRNAs targeting SRp38 were purchased from Open Biosystems, Thermo Scientific (Cat. Number RHS4430-99290688, -98895307, -98911716). A non-silencing – GIPZ shRNA vector was also purchased as a negative control (NC shRNA). HepG2 cells were first transfected with 2 µg of respective shRNA vectors by using FuGene 6
transfection reagent as described above, and after two-days incubation, media was changed to selective media containing 3 μg/mL of puromycin. After selection for 6 days, cells were transfected with 2 μg of the LDLR exon 9-14 minigene by FuGene 6 reagent. Then, minigene splicing efficiency and SRp38 expression were quantified. Twenty-four hours later, total RNA or total protein was collected for analyses.

**Human tissues**

Human liver samples were obtained from the Brain and Tissue Bank for Developmental Disorders (Baltimore, MD) and have been previously described (Zhu, Tucker et al. 2007). The samples were from deceased individuals with an average age at death for women of 28 ± 9 years (mean ± SD, range of 15−44, n = 15) and for men of 27 ± 10 (range 13−46, n = 20). The average post-mortem interval (PMI) for women was 13 ± 5 hours (mean ± SD, range 4−19, n = 15) while the PMI for men was 10 ± 3 hours (range 3−14, n = 20).

Human anterior cingulate brain specimens were generously provided by the Sanders-Brown Alzheimer’s Disease Center Neuropathology Core and have also been described elsewhere (Zou, Gopalraj et al. 2008). The samples were from deceased individuals with an average age at death for females of 82 ± 7 years (mean ± SD, n = 31) and for males of 82 ± 8 (n = 28). The average postmortem interval (PMI) for females and males was 3.0 ± 0.8 h (mean ± SD, n = 31) and 3.1 ± 0.8 h (n = 28), respectively.

**Evaluation of LDLR splicing efficiency in vivo**

Total RNA was prepared from the human liver and brain specimens and converted to cDNA in 1 μg aliquots with random hexamers and reverse transcriptase (SuperScript III, Invitrogen) (Chomczynski and Sacchi 1987). The LDLR splicing efficiency in brain specimens was evaluated by RT-PCR as we previously described for the liver specimens (Zhu, Tucker et al. 2007). Briefly, LDLR exon 10 sense primer 5’CCTGGCCAGCAGCATGCGCTC3’ and exon 14 antisense primer 5’CATCGTGGTGATCCCTGTTTC3’ were used to PCR-amplify (Platinum Taq, Invitrogen) brain cDNAs corresponding to LDLR exons 10–14, as well as isoforms lacking exons 11 and/or 12. PCR profiles consisted of preincubation 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, e.g., 30 cycles. PCR products were separated by PAGE and quantified as described above.

**Real-time RT-qPCR for splicing factor expression in vivo**

The expression level of SRp20, SRp38 and hnRNP G in the liver and brain specimens was quantified by real-time RT-qPCR. The 20 μl RT-qPCR mixture containing approximately 20 ng of liver or brain cDNA, 1 μM of each primer and 1× SYBR Green PCR Master Mix (Ambion) was subjected to RT-qPCR (MJ Research PTC-200 with Chromo4 detector). PCR profiles consisted of preincubation at 95°C for 10 min, followed by 40 cycles of 95°C for 30 s, 60°C for 30 s, and 72°C for 30 s. Specificity of the reactions was evaluated by showing a single PCR product by gel electrophoresis and by performing a melting curve analysis after the PCR amplification. The copy numbers of PCR product in each sample were determined relative to standard curves that were amplified in parallel and were based upon standardized amounts of purified PCR products. The following mRNAs and primers were evaluated: SRp20: 5’CGGCTTTGTCTTTGTGAAT3’ and 5’TCACCATTCCAGTTCCAC3’; SRp38: 5’ATTTCACACTCGCCGTCCA3’ and 5’CCGTCACAAATCCACTTTC3’; hnRNP G: 5’GTGCAGTGGAATGGGAGGGA3’ and 5’CCATCATCTCCTTGGGACAA3’; ribosomal protein L13A (RPL): 5’CATCTCCTTCTCGGCATCA3’ and 5’AACCCTGTTGTCAATGCCCT3’; hypoxanthine-guanine phosphoribosyltransferase (HPRT): 5’GACCAGTCAACAGGGGACAT3’ and 5’AACACTTCTGCTGGTCTTTC3’. The geometric mean of RPL and HPRT was used to normalize splicing factor expression among the samples (Vandesompele, De Preter et al. 2002). All RT-qPCR assays were repeated twice.

**Statistical analysis**

The effect of splicing factor and phosphorylation-related protein ectopic expression on LDLR minigene splicing efficiency was analyzed by using analysis of variance (ANOVA) and a post-hoc Scheffe test (SPSS software, v 17, SPSS Inc., Chicago, IL).
The effect of cantharidin treatment on DLLR minigene splicing was also analyzed by ANOVA and a post-hot Scheffe test (SPSS). The correlation between LDLR splicing efficiency and splicing factor expression was analyzed by using a linear regression model (SPSS).

Results

Screening of SR protein family members for modulating LDLR splicing
To identify splicing factors that may modulate LDLR splicing efficiency, I quantified the effects of eleven widely studied SR protein family members on splicing of the LDLR exon 9-14 minigene (Zhu, Tucker et al. 2007). I co-transfected the LDLR minigenes containing either rs688C or rs688T alleles with vectors encoding candidate SR proteins into HepG2 cells. Since some of the splicing factors were encoded as EGFP fusion proteins, e.g., ASF, SRp38, and SRp38-2, I confirmed their ectopic expression by fluorescence microscopy (Figure 3.1). Splicing efficiency was quantified 24 hours later by RT-PCR. The proportion of full length (FL) LDLR transcript was consistently less with the rs688T allele than the rs688C allele (Figure 3.2), reproducing our earlier observation of a SNP-induced effect on splicing (Zhu, Tucker et al. 2007). None of the SR proteins showed a SNP-dependent effect. Rather, many of the SR proteins reduced LDLR minigene splicing efficiency, with SRp20, SRp38, and SRp38-2 showing the largest changes (Figure 3.2B). SRp20 reduced FL LDLR by specifically increasing an LDLR transcript that skipped exon 11, i.e., Delta 11 LDLR ($p < 0.001$) (Figure 3.2C). SRp38 acted similarly to increase Delta 11 (Figure 3.2C) while SRp38-2 acted primarily by increasing the LDLR isoform that lacked both exons 11 and 12 (Figure 3.2E). Interestingly, SRp38 and SRp38-2 are alternatively spliced isoforms from the same gene; SRp38 includes one RNA recognition motif (RRM) and three RS domains while SRp38-2 has the same RRM but only one RS domain (Komatsu, Kominami et al. 1999).

Screening of hnRNP family members for modulating LDLR splicing
HnRNPs are also critical splicing regulatory proteins. Therefore, I also screened ten well-characterized hnRNP family members for their effects on LDLR splicing. I found that
hnRNP G and hnRNP G-T showed the largest effects (Figure 3.3B), decreasing the inclusion of exons 11 and 12 in the final LDLR mRNA product, regardless of which rs688 allele was present ($p < 0.001$; Figure 3.3E). Since hnRNP G has been reported to influence splicing in an RRM-independent fashion (Heinrich, Zhang et al. 2009), I further evaluated the effects of a truncated hnRNP G form that lacks the RRM domain. The result was identical to hnRNP G (Figure 3.3B-E); I interpret these results as indicating that hnRNP G may modulate LDLR splicing by acting as a scaffold protein without binding to LDLR mRNA.

**Dose-dependent effect of splicing factors**

Overall, the *in vitro* screening identified SRp20, SRp38, SRp38-2, hnRNP G and hnRNP G-T as candidates for modulating LDLR splicing in human tissues. Since the expression of hnRNP G-T is restricted to testis (Elliott, Venables et al. 2000), I focused on the first four splicing factors in subsequent studies. To evaluate whether these splicing factors repressed LDLR minigene splicing in a dose-dependent manner, three doses (0.01, 0.1, and 1 µg) of the vectors encoding the splicing factors were co-transfected with 1 µg of LDLR minigene. Since three of the four splicing factors were encoded as EGFP fusion proteins, I confirmed that expression was indeed dose-dependent by using anti-GFP Western blots (Figure 3.4A). When I analyzed LDLR minigene splicing by RT-PCR, I found that as the dose of splicing factor increased, the splicing factor effects on LDLR splicing increased as well (Figure 3.4B-F). SRp20 and SRp38 acted mostly by increasing Delta 11, while SRp38-2 and hnRNP G increased Delta 11+12, consistent with the results described in Figures 3.2 and 3.3. Hence, these results confirm that these splicing factors modulate LDLR splicing in a dose-dependent manner.

**Effect of splicing factors on LDLR exon 12 minigene**

To evaluate whether these splicing factors also modulate splicing efficiency of LDLR exon 12 minigene, I co-transfected vectors encoding these splicing factors with the LDLR single exon minigene containing rs688C allele. When the vector encoding SRp20 was co-transfected, the expression of LDLR minigene transcripts was very low precluding their quantitation; the mechanism for this observation is unclear. In contrast,
SRp38, SRp38-2 and hnRNP G co-transfections did not decrease LDLR transcript expression, but all increased the percentage of Delta 12 LDLR (Figure 3.5B). Interestingly, when the LDLR exon 9-14 minigene was transfected, SRp38-2 showed stronger effect on decreasing minigene splicing than SRp38 (Figure 3.2 and 3.4); however, when the single exon minigene was transfected, SRp38 had bigger effect on decreasing exon 12 splicing (Figure 3.5). HnRNP G and hnRNP G lacking the RRM had similar effects on single exon minigene splicing, indicating that hnRNP G regulated exon 12 splicing without itself binding to LDLR pre-mRNA. I interpret the results as suggesting that these splicing factors may interact with regulatory cis-elements within exon 12 either directly or indirectly binding to the RNA.

**Effect of splicing factor on mutant LDLR minigenes**

As described in Chapter Two, I have identified two functional ESE regions defined by MT6 and MT7 that are critical to LDLR exon 12 splicing. To understand whether these splicing factors interact with the two functional ESE regions, I co-transfected vectors encoding relevant splicing factors with either wild type (WT) or mutant (MT) LDLR exon 9-14 minigene, and quantified splicing efficiency by RT-PCR. A vector encoding EGFP was co-transfected as a negative control for splicing factor ectopic expression. Compare to WT, when MTs were co-transfected with the vector encoding EGFP, the splicing efficiency decreased (Figure 3.6B). The MTs increased primarily the percentage of Delta 12 LDLR and some Delta 11+12 (Figure 3.6D-E). When MTs were co-transfected with relevant splicing factors, they decreased splicing efficiency distinctly; SRp20 and SRp38 still increased Delta 11 (Figure 3.6C), and SRp38-2 and hnRNP G increased Delta 11+12 (Figure 3.6E). If the splicing factors required binding to the two ESE regions to function, they would lose their effects on splicing when co-transfected with ESE mutants. Therefore, I interpreted the results as that these splicing factors didn’t interact with the two critical ESE regions within LDLR exon 12. Instead, they influence LDLR splicing by binding to other regulatory elements.

**SR protein phosphorylation and LDLR splicing**
Some splicing factors are regulated by phosphorylation. For example, SRp20 is modulated by phosphorylation (Saeki, Yasugi et al. 2005; Sen, Talukdar et al. 2009); SRp38 is activated by dephosphorylation by PP1 (Shi and Manley 2007). To gain more insight, I decided to evaluate whether phosphorylation status influences LDLR splicing. I first evaluated effect of cantharidin, an inhibitor of PP1 and PP2A, on LDLR splicing. Since cantharidin is known as a toxic compound, I checked HepG2 cell viability by using calcein AM and PI double staining after cantharidin treatment for 24 hours. At the highest tested cantharidin concentration of 6 μM, I didn’t observe significant cell death although most cells assumed a round morphology (Figure 3.7). The percentage of viable cells slightly decreased from 99.5 ± 0.5% to 96.7 ± 1.1% with 6 μM cantharidin treatment (Figure 3.8A). When I quantified LDLR exon 9-14 minigene splicing after cantharidin treatments, I found that 6 μM of cantharidin significantly decreased LDLR splicing efficiency (Figure 3.8B).

To further evaluate the effect of splicing-related kinase and phosphatase on LDLR splicing, I co-transfected the LDLR exon 9-14 minigenes containing either rs688C or rs688T alleles with vectors encoding splicing-related kinase and phosphatase into HepG2 cells. Consistent with the results observed from cantharidin treatment (Figure 3.8B), nuclear inhibitor of PP1 (NIPP1) decreased the splicing efficiency as well (Figure 3.9B). While SRPK1 had no significant effect on LDLR minigene splicing, CLK2 increased the percentage of the Delta 11+12 LDLR (Figure 3.9E). Surprisingly, ectopic expression of PP1 also decreased the splicing efficiency by increasing the Delta 11+12 LDLR (Figure 3.9E). I interpreted the results as suggesting that the LDLR splicing is modulated by multiple splicing proteins, and the phosphorylation status of these proteins may influence LDLR splicing.

**Knockdown of SRp38 expression**

To confirm that the splicing factors modulate LDLR splicing, I proceeded to evaluate changes in splicing efficiency after silencing expression of the splicing factor. I first used small interfering RNA (siRNA) to decrease SRp38 expression. When the siRNAs were co-transfected with a vector encoding SRp38, after 48 hours, I observed that siRNA 1 had
a clear effect on reducing the ectopic SRp38 expression (Figure 3.10). When I co-transfected the siRNAs with the LDLR exon 9-14 minigene, the siRNA 1 decreased endogenous SRp38 mRNA level to about 40% (Figure 3.11A). However, the siRNAs didn’t reduce endogenous SRp38 protein level successfully, and hence, no change in LDLR minigene splicing (Figure 3.12).

Considering that the difficulty of decreasing SRp38 protein level by siRNA may be due to stable protein expression, I then tried using short hairpin RNA (shRNA) to reduce SRp38 expression. NC shRNA or shRNAs targeting SRp38 (shRNA 1-3) were transfected into HepG2 cells and the transfected cells were selected by puromycin for 6 days. Then the LDLR exon 9-14 minigene was transfected, and SRp38 expression and minigene splicing were evaluated. The results were similar to what I observed while using siRNA to decrease SRp38. The shRNA 3 had the strongest effect on decreasing SRp38 mRNA expression to about 50% (Figure 3.12A). However, I did not detect a reduction on SRp38 proteins level or LDLR splicing efficiency (Figure 3.12B). I attribute that lack of change in LDLR splicing efficiency to the lack of siRNA or shRNA effect on SRp38 proteins levels.

**Correlation between splicing factor expression and LDLR splicing in vivo**

To investigate whether these splicing factors are associated with LDLR splicing efficiency *in vivo* in a dose-dependent fashion, I quantified SRp20, SRp38, and hnRNP G expression as well as LDLR splicing in 59 human brain and 35 human liver specimens. Expression of the splicing factors was quantified by RT-qPCR; SRp38 and SRp38-2 were amplified by common primers. LDLR splicing efficiency was evaluated by RT-PCR with primer sequences corresponding to LDLR exon 10 and exon 14. Although others had reported sex-dependent differences in SRp20 expression and our lab had seen a sex-dependent difference in LDLR splicing *in vivo* previously (Zhu, Tucker et al. 2007; Zou, Gopalraj et al. 2008), I did not discern any sex-dependent differences in the expression of SRp20, SRp38, or hnRNP G in brain or liver (Figure 3.13; Table 1). I proceeded to evaluate the correlation between splicing factor expression and LDLR splicing by using linear regression models that included each of the splicing factors as well as rs688 and
sex. I found that SRp38 expression correlated significantly with LDLR splicing in both the brain \((p < 0.001,\) observed power = 0.999) and liver \((p = 0.003,\) observed power = 0.890). In contrast, SRp20 and hnRNP G expression was not significantly associated with LDLR splicing. Interestingly, the inclusion of both SPp38 expression and rs688 genotype in the model of LDLR splicing tended to reduce the association of rs688 with LDLR splicing, i.e., the association of rs688 with splicing efficiency in the liver remained significant \((p = 0.040,\) observed power = 0.622) while the association of rs688 with splicing in the brain showed a strong trend \((p = 0.062,\) power = 0.544). Consistent with our finding that SRp38 acted similarly upon the rs688C and rs688T LDLR minigenes, the interaction between SRp38 expression and rs688 genotype was not significant \((p = 0.138)\). Overall, increased SRp38 expression is associated significantly with decreased efficiency of LDLR splicing while rs688 genotype is associated significantly with LDLR splicing in the liver and trends with LDLR splicing in the brain.

**Discussion**

In this study, I identified several splicing factors that influence LDLR minigene splicing in transfected cells. Two of the factors, SRp20 and SRp38, increased Delta 11 LDLR while two other factors, SRp38-2 and hnRNP G, increased Delta 11+12 LDLR. A comparison of splicing factor expression and LDLR splicing efficiency in human brain and liver specimens established that SRp38 expression correlated with LDLR splicing efficiency. Overall, I interpret our results as suggesting that SRp38 modulates LDLR splicing *in vivo*. Since the association between SRp38 and LDLR splicing is more robust than that between rs688 and LDLR splicing, and rs688 itself is associated with cholesterol and AD (Zhu, Tucker et al. 2007; Zou, Gopalraj et al. 2008), I propose that SRp38 may also emerge as a modulator of cholesterol homeostasis and AD risk.

The critical nature of LDLR to cholesterol homeostasis is well established in that the loss of a single LDLR allele causes familial hypercholesterolemia (Hobbs, Brown et al. 1992). In addition, LDLR is a receptor for apoE in the brain (Fryer, Demattos et al. 2005; Cao, Fukuchi et al. 2006). Since alleles of apoE are the major AD genetic risk factors, factors
that modulate LDLR expression or splicing represent potential modulators of AD risk. Consistent with this possibility, a previous work demonstrated that rs688, a SNP within LDLR exon 12, modulates exon 12 splicing, and is associated with AD in at least some case-control series (Zhu, Tucker et al. 2007; Zou, Gopalraj et al. 2008). The LDLR isoforms lacking exon 11 and/or 12 are predicted to encode a truncated LDLR that contains the ligand binding domain but lacks the transmembrane domain. Therefore, the truncated LDLR may act in a dominant negative fashion to interfere with normal LDLR function (Zhu, Tucker et al. 2007). Since these LDLR isoforms encode non-functional LDLR proteins, further work is necessary to elucidate the regulation of LDLR splicing.

Several observations support the possibility that SRp38 is involved in LDLR splicing. First, each of the two SRp38 isoforms modulated LDLR minigene splicing in vitro. Second, the SRp38 isoforms acted in a dose-dependent manner to alter LDLR splicing in vitro. Third, SRp38 overall expression correlated with LDLR splicing efficiency in the brain and liver in vivo. Although the minigene co-transfection studies evaluated SRp38 and SRp38-2 separately, my efforts to distinguish these SRp38 isoforms in the qPCR studies were unsuccessful. SRp38 is encoded by six exons and includes one RRM and three RS domains while SRp38-2 shares the same first five exons but uses an alternative exon 6. Hence, for my qPCR studies, I used primers which corresponded to exons 2 and 3 and amplified both SRp38 and SRp38-2. Overall, SRp38 and SRp38-2 share a common RRM domain and RS domain but differ in that SRp38 has two additional RS domains (Komatsu, Kominami et al. 1999). Since RRM domains are responsible for protein-RNA interaction, SRp38 and SRp38-2 may recognize the same RNA sequence but recruit different splicing factors. This may account for the actions of SRp38 and SRp38-2 in the in vitro minigene studies, i.e., SRp38 and SRp38-2 acted similarly to increase Delta 11+12 LDLR while only SRp38-2 also increased Delta 11 LDLR. Since I did not detect significant levels of Delta 11 LDLR in our in vivo studies, other factors may suppress Delta 11 LDLR in vivo.

To further confirm that SRp38 proteins are critical modulators of LDLR splicing, I tried using siRNA and shRNA to inhibit SRp38 expression and evaluated effects on LDLR
splicing. However, the RNA interference (RNAi) methods were only able to decrease SRp38 mRNA expression by 40-50% with no reduction in protein level. I attribute the observation of no change in LDLR splicing to the lack of effect on SRp38 protein level. The difficulty of decreasing SRp38 expression by RNAi is possibly because first, SR proteins seem to be very stable, and second, SRp38 has at least three pseudogenes (Clinton, Chansky et al. 2002). On the other hand, SRp38 has been reported regulated by phosphorylation. Dephosphorylated SRp38 is known as a general splicing repressor during M phase of cell cycle (Shin and Manley 2002) and during heat shock (Shin, Feng et al. 2004). Shi and Manley have shown that SRp38 is dephosphorylated by PP1 and 14-3-3 protects SRp38 from dephosphorylation (Shi and Manley 2007). Therefore, to block SRp38 activity through regulation of dephosphorylation and then evaluate the effect on LDLR splicing would be interesting.

Since rs688 previously has been found to modulate LDLR exon 11-12 splicing in a sex-dependent fashion, the positive SRp20 results in my screening were interesting because SRp20 expression has been reported to show sex-dependent expression differences. In particular, Antunes-Martins et. al. found that hippocampal SRp20 mRNA expression was higher in male mice than female mice both at basal state and following its upregulation by memory training, i.e., the spatial version of the Morris water maze and background contextual fear conditioning (Antunes-Martins, Mizuno et al. 2007). However, I did not detect differences in SRp20 expression between male and female humans. This may reflect that I compared expression in human anterior cingulate while Antunes-Martins et. al compared murine hippocampus. Characterization of murine SRp20 promoter revealed that SRp20 expression is probably regulated in a cell-specific fashion (Jumaa, Guenet et al. 1997). Therefore, it would be interesting to investigate SRp20 expression in human hippocampus.

Another interesting candidate regarding sex-dependent splicing is hnRNP G and its paralog, RBMY, which shares 57% sequence homology with hnRNP G. In humans, hnRNP G is encoded by the RBMX gene which is located on the X chromosome while RBMY is encoded by RBMY, an RBMX paralog on the Y chromosome (Ma, Inglis et al.
Hence, both males and females express RBMX while only males express RBMY, suggesting that RBMY could alter LDLR splicing in a sex-dependent fashion. Here I observed that hnRNP G and RBMY both decreased LDLR splicing with RBMY having a weaker effect (Figure 2). However, RBMY, like hnRNP G-T, is only expressed in testis where it is critical for spermatogenesis (Ma, Inglis et al. 1993) Therefore, the RBMY modulation of LDLR splicing in vitro may be physiologically relevant only in testis. In contrast, hnRNP G is ubiquitously expressed. I found that hnRNP G robustly decreased LDLR minigene splicing in an RRM-independent fashion, indicating that hnRNP G may interact with other splicing factor(s) and function as a scaffold protein to modulate LDLR splicing. Since HnRNP G has been reported to interact with Tra2-ß1, (Hofmann and Wirth 2002; Nasim, Chernova et al. 2003), I considered the possibility that Tra2ß1 could mediate hnRNP G effects. However, Tra2-ß1 overexpression did not cause a significant change in LDLR minigene splicing (Figure 1), suggesting that hnRNP G influenced LDLR splicing via other mechanisms in vitro.

In summary, I identified several splicing factors which modulate LDLR splicing in vitro with SRp20, SRp38, SRp38-2 and hnRNP G having the strongest effects. I then found that increased SRp38 mRNA expression correlated with decreased LDLR splicing efficiency in the human brain and liver. Overall, I interpret the results as suggesting that SRp38 may be critical in the regulation of LDLR splicing. Since other factors that modulate LDLR splicing, e.g., rs688, have been associated with cholesterol and Alzheimer’s disease (Zhu, Tucker et al. 2007; Zou, Gopalraj et al. 2008), I propose that SRp38 as a candidate for modulating cholesterol homeostasis and AD risk.

Copyright © I-Fang Ling 2009
Figure 3.1 Expression of transfected SR proteins

HepG2 cells co-transfected with vectors encoding EGFP-fused SR proteins were confirmed by observing green fluorescence under microscope at 20X magnification. Transfection efficiency of the SR proteins ranged from 20-30% by counting the green cells relative to total cells. Representative results from SRp38 and SRp38-2 transfection were shown.
Figure 3.2 SR protein family effects on LDLR minigene splicing in vitro

HepG2 cells were co-transfected with vectors encoding different SR proteins and LDLR rs688C or rs688T in an exon 9-14 minigene. The effects of the SR proteins and rs688 allele on LDLR minigene splicing are shown as representative gel images (A) and quantitative results (B-E, mean ± SD, n = 3, * and + reflect p < 0.01 when compared to rs688T and rs688C minigenes, respectively, co-transfected with the negative control pEGFP vector). The faint PCR products observed between FL and Delta 11, and between
Delta 12 and Delta 11+12 represent non-physiologic LDLR splice variants, i.e., FL LDLR lacking the first 74 bp of exon 14, and a Delta 13 LDLR isoform, respectively.
Figure 3.3 HnRNP family member effects on LDLR minigene splicing in vitro

HepG2 cells were co-transfected with vectors encoding different hnRNP family members and LDLR rs688C or rs688T-containing minigenes. The effects of the hnRNPs and rs688 allele on LDLR minigene splicing are shown as representative gel images (A) and quantitative results (B-E, mean ± SD, n = 3, * and + reflect p < 0.01 when compared to rs688T and rs688C minigenes co-transfected with the negative control pEGFP vector).
Figure 3.4 Splicing factors show dose-dependent effects on LDLR splicing

The indicated amounts of vectors encoding splicing factors were co-transfected with 1 µg of the vector encoding the rs688C allele LDLR minigene. The total amount of non-LDLR vector in the transfections was held constant at 1 µg by adding “negative control” pcDNA4 vector. The dose-dependent overexpression of SRp38, SRp38-2 and hnRNP G was confirmed by Western blots (A). Dose dependent effects on LDLR minigene splicing efficiency are shown as a representative image (B) and overall quantitation (C-F, mean ± SD, n = 3).
Figure 3.5 Effect of splicing factors on LDLR exon 12 minigene splicing

The LDLR exon 12 minigene containing rs688C allele was co-transfected with vectors encoding respective splicing factors. The effects of the splicing factors on LDLR single exon minigene splicing are shown as a representative gel image (A) and quantitative results (B, mean ± SD, n = 3, * reflects p < 0.01 when compared to rs688C minigenes co-transfected with the negative control pEGFP vector).
Figure 3.6 Effects of splicing factors on mutant LDLR minigene splicing

Equal amount of wild type (WT) or mutant (MT) LDLR exon 9-14 minigene was co-transfected with vectors encoding respective splicing factors into HepG2 cells. Splicing factor effects on LDLR minigene splicing efficiency are shown as a representative gel image (A) and overall quantitation (B-E, mean ± SD, n = 3).
After treatment of respective concentration of cantharidin for 24 hours, cell viability was evaluated by calcein-AM and PI double staining. Calcein-AM permeates cell membrane and after converted to calcein by esterase in a viable cell, it emits green fluorescence. Hence, calcein-AM stains viable cells. In contrast, PI is impermeable to cell membrane, and thus stains dead cells. PI binds to DNA and emits red fluorescence. Images shown were at 20X magnification.

Figure 3.7 Evaluation of HepG2 cell viability after cantharidin treatment
Figure 3.8 Cantharidin effect on cell viability and LDLR minigene splicing

Cell viability after cantharidin treatment was estimated by the ratio of viable cells to total cells, and quantitative results were shown (A). No significant increase in cell death was detected after 24-hours incubation with 6 μM cantharidin. Quantitative results of cantharidin effect on the LDLR splicing were shown (B, mean ± SD, n = 3, * and + reflect $p < 0.01$ when compared to rs688T and rs688C minigenes without cantharidin treatment).
Figure 3.9 Effect of kinases and PP1 on LDLR minigene splicing

HepG2 cells were co-transfected with vectors encoding kinases and PP1 and LDLR rs688C or rs688T-containing minigenes. The effects of phosphorylation-related protein ectopic expression on LDLR minigene splicing are shown as a representative gel image from C allele co-transfection (A) and quantitative results (B-E, mean ± SD, n = 3, * and + reflect p < 0.01 when compared to rs688T and rs688C minigenes co-transfected with the negative control pEGFP vector).
Figure 3.10 siRNA knockdown of ectopic SRp38 expression
A negative control scrambled siRNAs (NC siRNA) or 2 siRNAs targeting SRp38 were co-transfected with a vector encoding SRp38. Since the vector expresses an EGFP-fused SRp38, the expression of the ectopic SRp38 was evaluated by observing green fluorescence under microscopy. While siRNA 2 had no significant effect, siRNA 1 significantly decreased the green fluorescence, indicating reduction of the ectopic SRp38 expression. Images shown were at 20X magnification.
NC siRNA or 2 siRNA targeting SRp38 were transfected into HepG2 cells, as well as a positive control siRNA targeting HPRT (HPRT siRNA), and then RNA expression was quantified by real-time RT-qPCR. While 10 nM of HPRT siRNA efficiently decreased HPRT mRNA expression to 20% (A, mean ± SD, n = 3), the strongest effect on reducing SRp38 expression is about 40% by 30 nM of siRNA 1 (B, mean ± SD, n = 3). Similar knockdown result was observed when higher concentration of siRNA (100 nM) was transfected. SRp38 expression was normalized to the geometric mean of RPL and HPRT as described (Vandesompele, De Preter et al. 2002), but for the positive control siRNA targeting HPRT, the HPRT expression was only normalized to RPL expression.
Figure 3.12 Effect of siRNAs on endogenous SRp38 expression and LDLR minigene splicing

SRp38 protein expression was evaluated by Western blot after siRNA treatment, and no significant change was detected (A). Western blot for actin was also performed as an internal control. Protein ladder is labeled on the left of the representative image. The lower band of SRp38 is possibly dephosphorylated SRp38. Quantitative results for effect of the siRNAs on LDLR minigene splicing are shown (B, mean ± SD, n = 3).
Figure 3.13 Effect of shRNAs on SRp38 expression and LDLR minigene splicing
The effects of SRp38 shRNAs on SRp38 mRNA expression (A) and LDLR minigene splicing (B) are shown as quantitative results (mean ± SD, n = 3). While shRNA 3 showed the strongest effect on decreasing SRp38 mRNA expression to 50%, no significant change in LDLR minigene splicing.
Figure 3.14 Splicing factor expression in human brain and liver

Splicing factor expression was quantified by real-time RT-qPCR in 59 human brain (A) and 35 liver specimens (B). Expression was normalized to the geometric mean of RPL and HPRT as described (Vandesompele, De Preter et al. 2002). Each black dot reflects one specimen. The 59 brain specimens include 31 males and 28 females, and the liver specimens contain 20 males and 15 females.
<table>
<thead>
<tr>
<th></th>
<th>Human Brains</th>
<th></th>
<th>Human Livers</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Female</td>
<td>Male</td>
<td>P Value</td>
<td>Female</td>
</tr>
<tr>
<td>(n = 28)</td>
<td>(n = 31)</td>
<td></td>
<td>(n = 15)</td>
<td>(n = 20)</td>
</tr>
<tr>
<td><strong>SRp20</strong></td>
<td>0.883 ± 0.399</td>
<td>0.824 ± 0.438</td>
<td>0.59</td>
<td>3.552 ± 1.620</td>
</tr>
<tr>
<td><strong>SRp38</strong></td>
<td>0.958 ± 0.625</td>
<td>0.847 ± 0.538</td>
<td>0.47</td>
<td>6.264 ± 1.997</td>
</tr>
<tr>
<td><strong>hnRNP G</strong></td>
<td>0.539 ± 0.262</td>
<td>0.476 ± 0.241</td>
<td>0.35</td>
<td>1.795 ± 0.844</td>
</tr>
</tbody>
</table>

Table 3.1 Splicing factor expression in human brain and liver

These results reflect the mean ± SD for the expression of the indicated splicing factors. Expression was normalized to the geometric mean of RPL and HPRT as described (Vandesompele, De Preter et al. 2002).
Figure 3.15 SRp38 and rs688 are associated with LDLR splicing efficiency

The relationship between LDLR splicing efficiency, SRp38 expression and rs688 genotype are shown. As expression level of SRp38 increased, the splicing efficiency of LDLR decreased. (A) In brain, specimens included 16 rs688C/C, 18 rs688C/T, and 19 rs688T/T. The $r^2$ for the model is 0.309. (B) In liver, specimens included 8 rs688C/C, 16 rs688C/T and 15 rs688T/T. The $r^2$ for the model is 0.213. The solid lines represent fit lines and the dashed lines represent 95% confidence interval.
A genetic polymorphism located in a splicing regulatory element can influence gene splicing, resulting in aberrant protein expression or function and causing diseases. Understanding splicing regulation can lead to discovery of novel molecular mechanisms underlying disease development, and possibly new therapeutic targets. Our lab has previously identified a functional SNP, rs688, that modulates LDLR splicing efficiency, i.e., the minor T allele decreases LDLR exon 12 splicing efficiency compared to the major C allele (Zhu, Tucker et al. 2007). Furthermore, rs688 is associated with higher cholesterol in pre-menopausal women (Zhu, Tucker et al. 2007) and with Alzheimer’s disease in men (Zou, Gopalraj et al. 2008). However, the mechanism of rs688 modulating LDLR exon 12 splicing and the regulation of LDLR splicing are still unclear. To gain more insight, I aimed to investigate the regulation of LDLR splicing efficiency in this study: first, to identify functional ESEs within LDLR exon 12, and second, to identify splicing factors that modulate LDLR splicing efficiency.

Findings

To identify functional ESEs within LDLR exon 12, I first identified nine conserved putative ESE regions by using in silico methods, and then mutated these regions to neutralize their affinity to splicing factors. Two ESE regions separated by only 11 bp were confirmed as critical to LDLR exon 12 splicing, and interestingly, rs688 is located in between these regions. The results confirmed that rs688 is located close to important splicing regulatory elements, and hence, modulates LDLR splicing efficiency.

To identify splicing factors that modulate LDLR splicing efficiency, I screened splicing factors from both the SR protein family and the hnRNP family. Several splicing factors showed effects on LDLR minigene splicing with SRp20, SRp38, SRp38-2 and hnRNP G having the strongest effects. While SRp20, and SRp38 specifically increased the LDLR transcript lacking exon 11, SRp38-2 and hnRNP G increased the transcript lacking exons
11 and 12. All four splicing factors repressed LDLR splicing in a dose-dependent manner. Interestingly, hnRNP G did not require its RRM to affect LDLR splicing, indicating that hnRNP G may interact with other splicing factor that binds to LDLR pre-mRNA.

Co-transfection with vectors encoding kinases or PP1 revealed that LDLR splicing is also regulated by phosphorylation status. While SRPK1 had no significant effect on LDLR minigene splicing, CLK2 increased the percentage of the Delta 11+12 LDLR. Surprisingly, ectopic expression of PP1 also altered splicing efficiency by increasing the Delta 11+12 LDLR. I interpreted the results as suggesting that LDLR splicing is complex and regulated by multiple splicing factors. The phosphorylation status of these proteins affects LDLR splicing.

These splicing factors not only affected splicing efficiency of the LDLR exon 9-14 minigene, but also the exon 12 only minigene, indicating that they may interact with the functional ESE regions identified within LDLR exon 12. However, the effects of the ESE mutants and the splicing factors were distinct. I interpreted the results as that these splicing factors did not interact with the two critical ESE regions within LDLR exon 12. Instead, they influence LDLR splicing by binding to other regulatory elements.

While SRp38 RNAi treatments decreased SRp38 mRNA expression by 40-50% with no reduction in protein level and no changes in LDLR splicing, I identified a dose-dependent correlation between the SRp38 mRNA expression and LDLR splicing efficiency in human brain and liver. Increased SRp38 expression is associated significantly with decreased efficiency of LDLR splicing. SRp38 affects LDLR minigene splicing efficiency in vitro, and its mRNA expression is correlated with LDLR splicing in vivo. These results indicate that SRp38 may be critical to the regulation of LDLR splicing.

**Model**

Based on the findings, I propose a model that, 1) rs688 is located close to regulatory ESE
regions with the rs688T allele disrupting the binding of splicing factor to the ESE regions, resulting in decreased LDLR splicing efficiency, producing more truncated LDLR; 2) SRp38 modulates LDLR splicing efficiency, with increased SRp38 levels resulting in more truncated LDLR and hence increases in cholesterol level and AD risk. (Figure 4.1)

Several reasoning supports this model. First, LDLR has been linked to cholesterol homeostasis and AD. The relationship between LDLR and cholesterol in periphery has been well-established. Mutations impairing LDLR function result in defects in LDLR-mediated endocytosis, causing plasma LDL-cholesterol accumulation and developing familial hypercholesterolemia (Goldstein and Brown 1974; Goldstein, Dana et al. 1975; Hobbs, Brown et al. 1992). Individuals with elevated levels of plasma cholesterol have higher susceptibility to AD (Jarvik, Wijsman et al. 1995). In addition, in mouse model, LDLR deficiency increases brain APOE level (Fryer, Demattos et al. 2005). Since APOE genetic polymorphisms are highly associated with AD and APOE is involved in Aβ production (Ye, Huang et al. 2005) or clearance in a receptor-dependent manner (Holtzman, Bales et al. 1999; Holtzman, Bales et al. 2000), LDLR as an APOE receptor may play a role in AD as well. Second, rs688 that modulates LDLR splicing is associated with cholesterol in woman and AD in men (Zhu, Tucker et al. 2007; Zou, Gopalraj et al. 2008). The rs688T allele increases the proportion of LDLR that lacks exon 12. The exon-skipping causes a reading frame shift, leading to a premature stop codon in exon 13, and thereby a truncated LDLR without the transmembrane domain encoded by exons 16-17 is produced. This soluble form of LDLR can bind to ligands, but cannot process them. As a result, the rs688T allele is associated with higher cholesterol level and AD risk. Similarly, SRp38 modulates LDLR splicing efficiency as well. Ectopic expression of SRp38 increases the proportion of LDLR that lacks exons 11 and 12, resulting in a soluble form of LDLR. Since SRp38 expression shows a stronger correlation with LDLR splicing efficiency than rs688, it is highly possible that SRp38 also modulates cholesterol level and AD risk.

Discussion and Future Directions
In this study, I transfected the LDLR exon 9-14 and single exon 12 minigenes into HepG2 cells as *in vitro* models to evaluate LDLR splicing efficiency. Although the effects of rs688, splicing factors, and ESE mutations were consistent within these two minigene models, I did notice that the levels of effect were stronger for the LDLR exon 9-14 minigene. For example, the minigene splicing efficiency between rs688C and rs688T alleles for the exon 9-14 minigene was about 10% different, whereas it was only ~5% for the exon 12 minigene. In addition, the splicing pattern of the LDLR exon 9-14 minigene with four different transcripts, i.e., FL, Delta 11, Delta 12, and Delta 11+12, was similar to the one I observed when quantifying the LDLR splicing in human tissues. Several reports have shown that for the same gene, its minigenes with different length construction result in different splicing patterns (Kereszturi, Kiraly et al. 2009; Mowrer and Wolfe 2009). Here, the rs688 allele has a lesser effect in the shorter minigene maybe because of the absence of neighboring regulatory elements, leading to lack of crosstalk between splicing factors. Therefore, to create a physiologically relevant splicing system, it is important to ensure that a minigene produces a splicing pattern similar to the endogenous gene.

The first part of this study confirmed that rs688 is located close to functional ESE regions, and modulates LDLR exon 12 splicing efficiency. Therefore, I proposed that the sequences containing rs688C and rs688T allele have different affinity toward associated SR protein(s), and as a result, regulate LDLR exon 12 splicing. The two critical ESE regions are predicted to bind SRp40 and SF2/ASF respectively. However, when I screened the effects of splicing factors on LDLR minigene splicing, they did not enhance the LDLR splicing, but decreased, indicating that it is other splicing factors that interact with the two ESE regions. To identify the splicing factor(s) that interact with the two critical ESE regions may help to develop a strategy to enhance the LDLR exon 12 splicing, overcoming the allele effect. The associated protein(s) can be distinguished by comparing binding proteins between wild type and mutant LDLR RNAs and identified by Western blots and mass spectrometry.
The second part of this study identified SRp20, SRp38, SRp38-2 and hnRNP G as having the strongest effects on LDLR minigene splicing. The ectopic expression of SRp20 specifically increased the LDLR isoform lacking exon 11 whose expression was relatively low within the human tissues I examined, and I did not find that the SRp20 expression correlated with LDLR splicing \textit{in vivo}. However, a recent report has shown that when quantifying LDLR splicing in EBV-transformed lymphocytes from subjects, a fair amount of the Delta 11 LDLR was detected, and two mutations for FH at the 5’ splice site of the intron 11 increased the percentage of the Delta 11 and the Delta 11+12 LDLRs (Holla, Nakken et al. 2009). It would be interesting to investigate whether SRp20 expression is higher in lymphocytes and whether it is involved in the recognition of LDLR exon 11. In addition, SRp20 expression has been reported to show a sex-dependent difference in mice hippocampus, although I did not detect such a difference in human anterior cingulate. Characterization of murine SRp20 promoter revealed that SRp20 expression is probably regulated in a cell-specific fashion (Jumaa, Guenet et al. 1997). Hence, it would be interesting to examine SRp20 expression in human hippocampus.

SRp38 expression regulated the LDLR minigene splicing \textit{in vitro}, and also correlated with LDLR splicing efficiency \textit{in vivo}. I interpret the results as suggesting that SRp38 modulates LDLR splicing. To further confirm that, I tried using siRNA and shRNA to inhibit SRp38 expression and evaluated effects on LDLR splicing. Unfortunately, the RNAi methods were not able to reduce SRp38 protein level. SRp38 has been shown to be activated by dephosphorylation to repress splicing. It is dephosphorylated by PP1 and 14-3-3 keeps SRp38 from dephosphorylation (Shin and Manley 2002; Shin, Feng et al. 2004; Shi and Manley 2007). Therefore, to block SRp38 activity by inhibiting its dephosphorylation by PP1 or sequestration by 14-3-3 are alternative methods to investigate effect of SRp38 knockdown on LDLR splicing.

Since the association between SRp38 and LDLR splicing is more robust than that between rs688 and LDLR splicing, and rs688 itself is associated with cholesterol and AD (Zhu, Tucker et al. 2007; Zou, Gopalraj et al. 2008), I propose that SRp38 may also
emerge as a modulator of cholesterol homeostasis and AD risk. It would be interesting to investigate whether SRp38 levels correlate with cholesterol levels, and whether any factor, e.g. SNP, linked to SRp38 expression is also linked to AD risk.

The proposed model demonstrates the importance of splicing regulation of LDLR to cholesterol homeostasis and AD risk. Impaired LDLR splicing regulation increases cholesterol level and AD risk. Therefore, a modulator for LDLR splicing may play a role in cholesterol homeostasis and AD as well. An interesting possibility for a future direction would be to screen a chemical library for molecules that modulate LDLR exon 12 splicing. For this work, I could establish a cell-based model in the following fashion. First, I would generate an LDLR minigene with a reporter gene such as GFP and design the construct such that when Delta LDLR is produced, a stop codon is introduced before the GFP gene sequence, blocking GFP expression. The second step would be to generate a stable cell line that expresses the LDLR minigene and produces a splicing pattern similar to the endogenous gene. The third step would be to confirm the protein expression of FL and Delta LDLR by immunostaining or Western blot. The fourth step would be to incubate cells with or without a testing molecule and compare GFP expression. Any molecule increasing or decreasing fluorescence is a potential modulator for LDLR splicing. The cell-based model provides a fast screen for molecules that regulate LDLR splicing. For example, estrogen would be a good candidate to test since the association between rs688 and diseases is sex-dependent. This work could be done in collaboration with The National Institutes of Health Chemical Genomics Center (NCGC) to seek chemicals which enhance LDLR splicing and thereby lead to a drug discovery.

**Conclusion**

LDLR plays a central role in cholesterol homeostasis, and polymorphisms causing aberrant LDLR splicing have been linked to diseases such as FH and AD (Graham, McIlhatton et al. 2005; Zhu, Tucker et al. 2007; Zou, Gopalraj et al. 2008; Holla, Nakken et al. 2009). Understanding the splicing regulation can lead to discovery of novel molecular mechanisms underlying diseases, and possibly new therapeutic approaches.
However, the knowledge of LDLR splicing regulation is limited. In this study, I first identified two functional ESE regions with LDLR exon 12, and confirmed that rs688 is located close to the regulatory elements, explaining the underlying mechanism of rs688 modulating LDLR exon12 splicing. Then, I identified several splicing factors that regulate LDLR minigene splicing *in vitro*, and among them, SRp38. Lastly, I found that SRp38 expression is correlated with LDLR splicing efficiency in human brain and liver. I interpret the results as suggesting that SRp38 may be critical in the regulation of LDLR splicing, and hence may also be a modulator for cholesterol homeostasis and AD risk.
Figure 4.1 Model of LDLR splicing regulation effects on cholesterol level and AD risk

Rs688 modulates LDLR splicing. When the rs688C allele is present, more full length LDLR is produced and functions as a receptor for APOE-containing lipoproteins that transfer cholesterol in periphery and contribute to Aβ clearance in brain. However, when rs688T allele is present, more Delta 12 LDLR isoform is produced and is predicted to translate into soluble LDLR. The soluble LDLR still binds to APOE-containing lipoproteins and Aβ peptides, but does not mediate their internalization. As a result, plasma cholesterol levels and AD risk increase. Two splicing regulatory elements that flank rs688 are critical to LDLR splicing. The presence of rs688T allele reduces the affinity of these splicing regulatory elements to splicing factor(s), and therefore,
decreases LDLR splicing efficiency. Identifying the unknown splicing factor(s) to improve the interaction and enhance splicing can help to develop a strategy to overcome the allele effect as well as the poor 5’ splice donor sequence. SRp38 functions as a splicing repressor for LDLR splicing with higher SRp38 expression correlating with less efficient LDLR splicing and more delta LDLR produced. Since rs688 modulates LDLR splicing and is associated with cholesterol level and AD risk, SRp38 being a regulator for LDLR splicing can be a modulator for cholesterol homeostasis and AD risk as well.
This is a License Agreement between I-Fang Ling ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number 2302831178295
License date Nov 05, 2009
Licensed content publisher Nature Publishing Group
Licensed content publication Nature Reviews Neuroscience
Licensed content title Apolipoprotein E and its receptors in Alzheimer's disease: pathways, pathogenesis and therapy
Licensed content author Guojun Bu
Volume number
Issue number
Pages
Year of publication 2009
Portion used Figures / tables
Number of figures / tables 1
Requestor type Student
Type of Use Thesis / Dissertation
High-res requested No
Billing Type Invoice
Company I-Fang Ling
Billing Address 1608 University Ct. Apt. E311

Lexington, KY 40503
United States

Customer reference info
Total 0.00 USD

Terms and Conditions

Terms and Conditions for Permissions
Nature Publishing Group hereby grants you a non-exclusive license to reproduce this material for this purpose, and for no other use, subject to the conditions below:

1. NPG warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to Nature Publishing
Group and does not carry the copyright of another entity (as credited in the published version). If
the credit line on any part of the material you have requested indicates that it was reprinted or
adapted by NPG with permission from another source, then you should also seek permission
from that source to reuse the material.

2. Permission granted free of charge for material in print is also usually granted for any electronic
version of that work, provided that the material is incidental to the work as a whole and that the
electronic version is essentially equivalent to, or substitutes for, the print version. Where print
permission has been granted for a fee, separate permission must be obtained for any additional,
electronic re-use (unless, as in the case of a full paper, this has already been accounted for during
your initial request in the calculation of a print run). NB: In all cases, web-based use of full-text
articles must be authorized separately through the 'Use on a Web Site' option when requesting
permission.

3. Permission granted for a first edition does not apply to second and subsequent editions and for
editions in other languages (except for signatories to the STM Permissions Guidelines, or where
the first edition permission was granted for free).

4. Nature Publishing Group's permission must be acknowledged next to the figure, table or abstract
in print. In electronic form, this acknowledgement must be visible at the same time as the
figure/table/abstract, and must be hyperlinked to the journal's homepage.

5. The credit line should read:

Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference
citation), copyright (year of publication)

For AOP papers, the credit line should read:

Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME], advance online
publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

6. Adaptations of single figures do not require NPG approval. However, the adaptation should be
credited as follows:

Adapted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference
citation), copyright (year of publication)

7. Translations of 401 words up to a whole article require NPG approval. Please visit
http://www.macmillanmedicalcommunications.com for more information. Translations of up to a
400 words do not require NPG approval. The translation should be credited as follows:

Translated by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference
citation), copyright (year of publication).

We are certain that all parties will benefit from this agreement and wish you the best in the use of this
material. Thank you.

v1.1

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for
your reference. No payment is required.
If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 30 days of the license date. Payment should be in the form of a check or money order referencing your account number and this license number 2302831178295.

If you would prefer to pay for this license by credit card, please go to http://www.copyright.com/creditcard to download our credit card payment authorization form.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

If you find copyrighted material related to this license will not be used and wish to cancel, please contact us referencing this license number 2302831178295 and noting the reason for cancellation.

Questions? customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.
REFERENCES


Komatsu, M., E. Kominami, et al. (1999). "Cloning and characterization of two neural-
salient serine/arginine-rich (NSSR) proteins involved in the regulation of

Kosik, K. S., C. L. Joachim, et al. (1986). "Microtubule-associated protein tau (tau) is a
major antigenic component of paired helical filaments in Alzheimer disease." 

Kowal, R. C., J. Herz, et al. (1990). "Opposing effects of apolipoproteins E and C on
lipoprotein binding to low density lipoprotein receptor-related protein." J Biol


(Val1000Ile) polymorphism and Alzheimer's disease." Hum Mol Genet 7(12):
1953-6.


Neurodegener 3: 8.

apolipoprotein E and cholesterol metabolism through lipoprotein receptor LRP1." 


VITA

I-Fang Ling

Personal Information

Date of Birth: October 7th, 1979
Place of Birth: Taipei, Taiwan

Education

Fall 2004 – present PhD student, Department of Physiology, University of Kentucky, Lexington, KY

Fall 1998 – Spring 2002 Bachelor of Science, Department of Agricultural Chemistry, National Taiwan University, Taipei, Taiwan

Employment

Fall 2002 – Spring 2004 Laboratory technician, Center for Optoelectronic Biomedicine, College of Medicine, National Taiwan University

Fellowship

Fall 2007 – Spring 2008 Kentucky Opportunity Fellowship

Fall 2006 – Spring 2007 Kentucky Opportunity Fellowship

Abstract


**Publication**

Ling, I.F, and Estus, S. (2009) Role of SRp38 for low density lipoprotein receptor splicing. (Submitted Human Mutation)


