Investigating Alternative Splicing and Polyadenylation of the Interleukin 7 Receptor (IL7R) Exon 6: Implications for Multiple Sclerosis

by

Irina Evsyukova

Department of Biochemistry
Duke University

Date:_______________________

Approved:

________________________________________
Mariano A. Garcia-Blanco, Co-Supervisor

________________________________________
Michael D. Been, Co-Supervisor

________________________________________
G. Vann Bennett

________________________________________
Blanche Capel

________________________________________
Kenneth N. Kreuzer

________________________________________
Laura N. Rusche

Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biochemistry in the Graduate School of Duke University

2012
ABSTRACT

Investigating Alternative Splicing and Polyadenylation of the Interleukin 7 Receptor (IL7R) exon 6: Implications for Multiple Sclerosis

by

Irina Evsyukova

Department of Biochemistry
Duke University

Date: _______________________

Approved:

___________________________
Mariano A. Garcia-Blanco, Co-Supervisor

___________________________
Michael D. Been, Co-Supervisor

___________________________
G. Vann Bennett

___________________________
Blanche Capel

___________________________
Kenneth N. Kreuzer

___________________________
Laura N. Rusche

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biochemistry in the Graduate School of Duke University

2012
Abstract

Interleukin 7 receptor, IL7R, is expressed exclusively on cells of the lymphoid lineage and its expression is crucial for development and maintenance of T cells. While transcriptional regulation of IL7R expression has been widely studied, its posttranscriptional regulation has only recently been uncovered. Alternative splicing of IL7R exon 6, the only exon that encodes the transmembrane domain of the receptor, results in membrane-bound (exon 6 included) and soluble (exon 6 skipped) IL7R isoforms, respectively. Interestingly, the inclusion of exon 6 is affected by a single-nucleotide polymorphism associated with the risk of developing multiple sclerosis, a prototypic demyelinating disease of the central nervous system. Given the potential association of exon 6 inclusion with multiple sclerosis, we investigated the cis-acting elements and trans-acting factors that regulate exon 6 splicing.

We utilized mutagenesis of exon 6 and surrounding introns to identify multiple exonic and intronic cis-acting regulatory elements that impact inclusion of exon 6. At least two of these elements, one exonic splicing silencer and one exonic splicing enhancer, are located in the direct vicinity of the MS-associated SNP. We also uncovered a consensus polyadenylation signal (AAUAAA) in intron 6 of IL7R, 16 nucleotides downstream from the exon 6 5’ splice site, and showed that mutations to this site resulted in an increase in exon 6 inclusion. Additionally, we determined that the 5’ splice site of exon 6 is weak. We propose that this site may be responsible for exon 6 splicing regulation.
Using tobramycin RNA affinity chromatography followed by mass spectrometry, we identified \textit{trans}-acting protein factors that bind exon 6 and regulate its splicing. These experiments identified cleavage and polyadenylation specificity factor 1 (CPSF1) among protein binding candidates. siRNA-mediated knockdown of CPSF1 resulted in an increase in exon 6 inclusion, consistent with the results of mutations to the CPSF1 binding site. Correspondingly, CPSF1 depletion had no effect on a minigene with a mutation in the intronic polyadenylation site. Finally, 3’RACE and RT-PCR experiments on RNA from Jurkat cells suggested that the intronic AAUAAA site is utilized at low frequency by the polyadenylation machinery to produce a novel polyadenylated mRNA isoform. Together, our results suggest that competing pre-mRNA splicing and polyadenylation may regulate exon 6 inclusion and resultant levels of functional \textit{IL7R} produced. Since the intronic polyadenylated isoform of \textit{IL7R} is predicted to be translated into a membrane-bound protein product with a shortened, signal transduction-incompetent cytoplasmic tail, this may be relevant for both T cell biology and development of multiple sclerosis.
Dedication

This thesis is dedicated to the memory of my father, Dr. Yury M. Evsyukov (1931-2005). He did not live to see me join a lab for my graduate studies, choose my research project or complete my dissertation, but he still inspires me (and will inspire for the rest of my life) in all my endeavors.
Contents

Abstract .............................................................................................................................. iv

List of Tables .................................................................................................................... x

List of Figures ................................................................................................................... xi

List of abbreviations ........................................................................................................ xiii

Acknowledgements ......................................................................................................... xv

Chapter 1: Importance of posttranscriptional regulation of gene expression and its connections to T cell biology and disease ................................................................. 1

1.1 Posttranscriptional control of gene expression ............................................................. 1

1.2 Alternative splicing as a major regulator of gene expression ....................................... 2

1.2.1 An overview of pre-mRNA splicing ......................................................................... 2

1.2.2 Different mechanisms of alternative pre-mRNA splicing ........................................ 4

1.3 Alternative polyadenylation as an emerging regulator of gene expression .................. 5

1.3.1 An overview of pre-mRNA 3'-end processing .......................................................... 5

1.3.2 CPSF complex ............................................................................................................. 6

1.3.3 Interconnections between splicing and polyadenylation ....................................... 6

1.3.4 Alternative polyadenylation is widespread and important for regulation of gene expression ...................................................................................................................... 8

1.4 Alternative splicing and alternative polyadenylation in human disease ...................... 10

1.4.1 Alternative splicing of interleukin 7 receptor (IL7R) and multiple sclerosis .......... 11

1.5 Focus of this work ......................................................................................................... 15

Chapter 2: Identification of cis-acting regulatory elements that regulate splicing of IL7R exon 6 ......................................................................................................................... 25

2.1 Introduction .................................................................................................................. 25

2.2 Materials and Methods ............................................................................................... 28
2.3 Results.......................................................................................................................................................... 31

2.3.1 Multiple cis-acting regulatory elements in IL7R exon 6 regulate its inclusion ... 31

2.3.2 An intronic cis-acting element associated with 3’end formation and polyadenylation downregulates IL7R exon 6 inclusion .......................................................... 34

2.3.3 The 5’ splice site of IL7R exon 6 is weak .............................................................................................. 35

2.4 Discussion .................................................................................................................................................. 36

Chapter 3: Identification of cleavage and polyadenylation specificity factor 1 (CPSF1) as a regulator of IL7R exon 6 splicing ........................................................................................... 47

3.1 Introduction ................................................................................................................................................ 47

3.2 Materials and Methods ............................................................................................................................ 49

3.3 Results ...................................................................................................................................................... 56

3.3.1 Exon 6 splicing regulation is recapitulated in HeLa cell nuclear extracts in vitro .............................................. 56

3.3.2 RNA affinity chromatography identifies potential protein regulators of IL7R exon 6 splicing .......................................................... 57

3.3.3 CPSF1 regulates exon 6 splicing by promoting skipping of exon 6 ..................................................... 58

3.3.4 Analysis of potential intronic polyadenylated isoforms in IL7R ......................................................... 59

3.4 Discussion ................................................................................................................................................ 61

Chapter 4: Concluding discussion .................................................................................................................... 78

4.1 Posttranscriptional regulation of IL7R ....................................................................................................... 78

4.2 Potential role of different isoforms of IL7R in multiple sclerosis ............................................................ 80

4.3 Concluding remarks ................................................................................................................................. 82

Appendix ......................................................................................................................................................... 83

Table 3: Primers used in this study .................................................................................................................. 83

References ....................................................................................................................................................... 85
Biography..................................................................................................................................... 98
List of Tables

Table 1: Proteins identified by quantitative mass spectrometry to bind differentially to "C" and "sub(36-40)" RNAs ................................................................. 67

Table 2: Protein factors that do not show differential binding to "C" vs "sub(36-40)" templates as identified by mass spectrometry................................................................. 68
List of Figures

Figure 1.1: Mechanisms of alternative pre-mRNA splicing ......................................................... 18
Figure 1.2: Mammalian 3'-end processing machinery ............................................................... 19
Figure 1.3: Types of alternative polyadenylation ....................................................................... 20
Figure 1.4: Association analysis of SNPs from IL7R ................................................................. 21
Figure 1.5: Alternative splicing of IL7R exon 6 ....................................................................... 23
Figure 1.6: IL7R isoforms based on data from the Ensembl Genome Browser
(www.ensembl.org) ..................................................................................................................... 24
Figure 2.1 Minigenes that report on exon 6 inclusion ................................................................. 39
Figure 2.2: Mutagenesis across exon 6 ..................................................................................... 40
Figure 2.3: Deletion analysis of exon 6 inclusion ...................................................................... 41
Figure 2.4: Larger substitutions encompassing regulatory elements ESS1 and ESE2 ........... 42
Figure 2.5: Inclusion of exon 6 in Jurkat cells .......................................................................... 43
Figure 2.6: ESE2 works in heterologous context ..................................................................... 44
Figure 2.7: An intronic element associated with cleavage and polyadenylation influences
exon 6 inclusion .......................................................................................................................... 45
Figure 2.8: IL7R exon 6 5' splice site is weak ........................................................................... 46
Figure 3.1: Splicing regulation of IL7R exon 6 is recapitulated in HeLa cell nuclear extracts
in vitro ......................................................................................................................................... 65
Figure 3.2: Templates for tobramycin RNA affinity chromatography ..................................... 66
Figure 3.3: Western blot analysis of RNA affinity chromatography elution fractions .......... 71
Figure 3.4: Effect of siRNA-mediated CPSF1 knockdown on exon 6 inclusion of
minigenes sub(36-40) and sub(36-40)pAmut ........................................................................ 73
Figure 3.5: RT-PCR experiments to investigate mRNA isoforms of IL7R ......................... 75
Figure 3.6: 3' RACE experiments to investigate intronic polyadenylated isoforms of IL7R mRNA.
**List of abbreviations**

ATP adenosine triphosphate

bp base pair

BSA bovine serum albumin

CF cleavage factor

CNS central nervous system

CPSF cleavage and polyadenylation specificity factor

CstF cleavage stimulation factor

CTP cytidine triphosphate

DSE downstream sequence element

ESE exonic splicing enhancer

ESS exonic splicing silencer

GTP guanosine triphosphate

GWAS genome-wide association study

hnRNP heterogeneous nuclear ribonucleoprotein

IL7 interleukin 7

IL7R interleukin 7 receptor

IPA intronic polyadenylation

ISE intronic splicing enhancer

ISS intronic splicing silencer

miRNA microRNA

MS multiple sclerosis
nt nucleotide

PAS polyadenylation signal

PBMCs peripheral blood mononuclear cells

pre-mRNA pre-messenger RNA

RNA ribonucleic acid

RNAi RNA interference

RNP ribonucleoprotein

RT-PCR reverse transcription – polymerase chain reaction

siRNA small interfering RNA

SNP single-nucleotide polymorphism

snRNA small nuclear RNA

snRNP small nuclear ribonucleoprotein

SR proteins serine/arginine-rich proteins

TTP thymidine triphosphate

USE upstream sequence element

UTP uridine triphosphate

UTR untranslated region
Acknowledgements

My deepest gratitude is to my advisor, Dr. Mariano A. Garcia-Blanco, for giving me the opportunity to work in his lab, the freedom to select the projects that I wanted to pursue, and for his patience. It has not always been easy, but what I have learned during my years in the lab is invaluable. I could not be grateful enough for Mariano’s support outside of the lab as well, even in the most difficult times.

I would like to thank my committee members: Dr. Michael D. Been, Dr. G. Vann Bennett, Dr. Blanche Capel, Dr. Kenneth N. Kreuzer and Dr. Laura N. Rusche, for guiding me through my projects. Special thanks to Drs. Mike Been and Laura Rusche for their continuous support throughout the years.

Several of my labmates in the Garcia-Blanco laboratory have been instrumental in forming my project. Special thanks to Dr. David Mauger, my original lab mentor and friend, who not only taught me how to do lab work, but also made sure to plan my days in the lab hour-by-hour. I am indebted to Dr. Shelton Bradrick for all the lengthy discussions and excellent advice on my projects. My research would not have been possible without Shelton’s help. Many thanks to Dr. Jason Somarelli for his very positive attitude and willingness to help with pretty much anything.

I would like to thank Dr. Puneet Seth who created original constructs that this project heavily relies on; Dr. Will Thompson and the rest of the staff at Duke University Proteomics Core Facility who performed mass spectrometry experiments and analyzed the data; and Dr. Jeffrey Wilusz of Colorado State University who generously shared the last bit of CPSF1 antiserum that was in his possession. Dr. Klaus Hartmuth (Max Planck
Institute for Biophysical Chemistry, Göttingen, Germany) shared unpublished data on RNA affinity chromatography.

Last, but not least, I would like to thank my family – my Mom and my husband Michael especially – for dealing with me throughout these years. I would not have been able to succeed without your support and encouragement.
Chapter 1: Importance of posttranscriptional regulation of gene expression and its connections to T cell biology and disease

1.1 Posttranscriptional control of gene expression

Regulation of gene expression at the level of RNA represents a fundamental mechanism of generating transcript and protein diversity in eukaryotes. Essentially all steps of RNA processing can be regulated (Moore 2005):

- pre-mRNA capping;
- pre-mRNA splicing;
- polyadenylation;
- stability;
- editing;
- transport;
- localization and storage;
- mRNA translatability.

This complex regulation is achieved via the action of numerous RNA-binding proteins (that associate with RNAs to form dynamic ribonucleoprotein complexes (RNPs) that regulate every aspect of RNA metabolism) (Lukong et al., 2008) and various non-coding RNAs. The ultimate result of this regulation is flexibility with which the transcriptome responds to various stimuli (Moore 2005, Cooper et al., 2009). However, the high degree of complexity of posttranscriptional gene regulation also implies that
misregulation of any of the RNA or protein components would be deleterious and can result in disease.

**1.2 Alternative splicing as a major regulator of gene expression**

1.2.1 An overview of pre-mRNA splicing

Most metazoan protein-coding genes consist of short exons and much longer introns, which are retained and lost in mature mRNA, respectively. For mRNAs to be transported to the cytoplasm where they can direct protein synthesis, introns need to be efficiently and accurately removed. This is achieved via the process of RNA splicing (Tycowski et al., 2006).

Several layers of information, referred to as the “splicing code”, are included in the pre-mRNA sequence (Goldstrohm et al., 2001; Wang and Cooper 2007; Cooper et al., 2009). These include the consensus 5’ and 3’ splice sites, a branch point sequence upstream of the 3’ splice site, and a wide variety of intronic and exonic splicing elements (see below). The basic splicing machinery, known as the spliceosome, recognizes and binds to these elements and executes a highly complicated task of correctly defining an exon, ignoring anything that might resemble an exon, and joining appropriate exons together. Two cellular spliceosomes termed “major” and “minor”, work in parallel. The major spliceosome, a multi-megadalton ribonucleoprotein particle, consists of both RNAs (small nuclear RNAs, snRNAs: U1, U2, U4/U6, and U5) and a variety of protein factors (which together with the snRNAs form five small nuclear ribonucleoprotein complexes, or snRNPs) (Jurica and Moore 2003).
The spliceosome catalyzes exon joining and intron removal via two transesterification reactions (Jurica 2008; Smith et al., 2008; Staley and Woolford 2009; Wahl et al., 2009). The 5’ splice site is recognized first by the U1 snRNP, followed by the U2 snRNP binding to the branch point site, which results in the formation of the pre-splicing complex. Addition of the U4/U5/U6 tri-snRNP to the pre-splicing complex produces a fully assembled spliceosome, which undergoes conformational changes resulting in the first catalytic step of the splicing reaction, branch formation. After subsequent conformational changes and proper alignment of the two exons to be spliced, the second catalytic reaction – ligation of the exons – occurs. Importantly, spliceosome function is executed through dynamic RNA-RNA, RNA-protein and protein-protein interactions.

Mammalian splice and branch point sites are short sequences, which do not contain all the information necessary to direct correct exon definition. This information is provided by additional splicing regulatory elements that exist within both exons and introns. Based on their mode of action, these additional elements can be divided into at least four groups: exonic or intronic splicing enhancers (ESEs or ISEs) activate splicing from within exons or introns, respectively, while exonic or intronic splicing silencers (ESSs or ISSs) suppress splicing (Garcia-Blanco et al., 2004; Wang and Cooper 2007; Wang and Burge 2008; Cooper et al., 2009). Occasionally, the same element is able to activate and suppress splicing of two mutually exclusive exons (Carstens et al., 1998). All of the aforementioned cis-acting regulatory elements function through the binding of different trans-acting protein factors such as heterogeneous nuclear ribonucleoproteins,
hnRNPs, and SR (serine/arginine-rich) proteins and non-coding RNAs (Kishore and Stamm 2006; Khanna and Stamm 2010).

1.2.2 Different mechanisms of alternative pre-mRNA splicing

The layers of splicing regulation described above control both constitutive and alternative splicing of exons. Alternative splicing, which leads to the generation of multiple different transcripts from the same pre-mRNA, is a potent generator of protein diversity (Black 2003). Considering that multiple mRNAs can be generated from a single pre-mRNA via different combinations of splicing patterns, and that splicing can be regulated in a variety of ways (tissue-specific, developmental stage-specific, stimulus-specific) (Cooper et al., 2009), the resulting mRNA and protein diversity is truly enormous. According to some recent estimates, up to 94% of all human genes undergo alternative splicing (Wang et al., 2008), which means that this mechanism of post-transcriptional regulation is critical for virtually all cellular processes.

Alternative splicing proceeds through several different pathways (Figure 1.1): a) use of alternative 5’ splice sites; b) use of alternative 3’ splice sites; c) intron retention; d) a choice to include or skip an exon, and e) a choice between two mutually exclusive exons (Goldstrohm et al., 2001; Black 2003; Lynch 2004; Evsyukova et al., 2010). Together with other events that lead to transcript diversity (e.g., alternative polyadenylation which will be described in the following section) the five processes described in Figure 1 result in protein isoforms that can possess diverse or even opposite functions (Cooper et al., 2009).
1.3 Alternative polyadenylation as an emerging regulator of gene expression

1.3.1 An overview of pre-mRNA 3’-end processing

The majority of fully processed eukaryotic mRNAs possess a polyadenylated (poly(A)) tail, which consists of about 200 adenosine residues in higher eukaryotes and regulates proper mRNA translation, stability, and transport (Lutz and Moreira 2010). Formation of the pre-mRNA 3’ end proceeds through two tightly coupled reactions: (1) endonucleolytic cleavage at a specific site within the pre-mRNA, and (2) non-templated addition of adenosine residues. Similar to regulation of splicing, both of the 3’-end processing reactions are regulated by cis-acting elements within the pre-mRNA sequence and numerous trans-acting protein factors that bind these sequences (Colgan and Manley 1997; Zhao et al., 1999; Mandel et al., 2008; Lutz and Moreira 2010; Millevoi and Vagner, 2010; Campigli Di Giammartino et al., 2011).

Nearly all mammalian polyadenylation sites contain a core consensus polyadenylation signal (PAS), AAUAAA (or a close variant) about 10-35 nucleotides upstream from the actual cleavage site (as reviewed in Lutz, 2008; Lutz and Moreira, 2010; Proudfoot 2011) (Figure 1.2). Additionally, U/GU- rich sequences downstream from the cleavage site (downstream sequence element, DSE) as well as sequences directly upstream from PAS (upstream sequence element, USE) have been shown to enhance efficiency of polyadenylation (Proudfoot 2011 and references therein). Four multi-subunit protein complexes mediate 3’ end formation (Mandel et al, 2008; Campigli Di Giammartino et al., 2011): cleavage and polyadenylation specificity factor (CPSF), cleavage stimulation factor (CstF) and cleavage factors I and II (CFI and II). CPSF and
CstF bind to PAS and DSE, respectively, to initiate the processing reaction (Figure 1.2). Additional protein factors involved in 3’ end processing include poly(A) polymerase (PAP), poly(A)-binding protein (PABP), symplekin and others. Other auxiliary proteins that aid in the polyadenylation reaction have also been identified (as reviewed in Mandel et al., 2008).

1.3.2 CPSF complex

Mammalian cleavage and polyadenylation specificity complex (CPSF) contains five subunits: CPSF1 (CPSF160), CPSF2 (CSPF100), CPSF3 (CPSF73), CPSF4 (CPSF30) and hFip1 (Mandel et al., 2008; Figure 1.2), all of which are required for efficient cleavage and polyadenylation. CPSF1, the largest subunit, binds to the PAS AAUAAA directly (Keller et al., 1991; Murthy and Manley 1995). CPSF1 interacts with other subunits of CPSF as well as with other protein factors involved in polyadenylation, such as PAP and the CstF complex (Murthy and Manley 1992, 1995). The CPSF3 (CPSF73) component of the CPSF complex is thought to be the endoribonuclease mediating mRNA cleavage (Ryan et al., 2004; Dominski et al., 2005; Mandel et al., 2006). Somewhat less is known about functions of CPSF2 and hFip1 (for Factor interacting with Pap1p, human) is thought to bring PAP close to the polyadenylation site through protein-protein interactions (as reviewed in Mandel et al., 2008).

1.3.3 Interconnections between splicing and polyadenylation

The processes of splicing and polyadenylation are highly coupled and influence each other. In cell-free extracts, polyadenylation is stimulated by the presence of an
upstream intron in single-intron constructs (Niwa et al., 1990), and mutation to the AAUAAA polyadenylation signal inhibits splicing of an upstream intron (Niwa and Berget 1991). In a different set of \textit{in vitro} experiments, mutations to splicing regulatory elements (3’ splice site and branch point) reduced polyadenylation efficiency, while mutations to the AAUAAA and DSE sites affected splicing (Cooke et al., 1999). Together, these \textit{in vitro} experiments suggested that splicing and polyadenylation interact to recognize and define terminal exons. \textit{In vivo}, inhibition of 3’-end processing results in inhibition of splicing in single-intron pre-mRNAs, and CPSF is required for this inhibition (Li et al., 2001). Several splicing factors have been shown to regulate polyadenylation: (1) U2AF 65 interacts with PAP (Vagner et al., 2000) and CFI(m) (Millevoi et al., 2002, 2006) to promote 3’-end processing and couple polyadenylation and splicing; (2) U2AF 35 and 65 as well as hnRNP I bind USEs of several 3’-end polyadenylation elements (Danckwardt et al., 2007); (3) splicing co-activator SRm160 promotes 3’-end cleavage (McCracken et al., 2002) and (4) SRp20 regulates alternative polyadenylation of the calcitonin/calcitonin gene-related peptide mRNAs (Lou et al., 1998). Moreover, subunits of CPSF directly interact with components of U2 and U1 snRNPs (Kyburz et al., 2006; Lutz et al., 1996). Together with the role of U1 snRNP in regulating alternative intronic polyadenylation (see below), these data indicate a close interplay between splicing and polyadenylation in regulation of gene expression.
1.3.4 Alternative polyadenylation is widespread and important for regulation of gene expression

A large-scale analysis of mammalian polyadenylation has established that approximately 54% of human and 32% of mouse genes encode transcripts that are alternatively polyadenylated (Tian et al., 2005). Three major types of alternative polyadenylation have been described: (1) type I includes transcripts that contain only one poly(A) site in the 3’-most exon and thus produce just one polyadenylated product; (2) type II includes transcripts that contain more than one poly(A) signal in the common terminal exon and thus produce several polyadenylated isoforms, and (3) type III is intimately connected to alternative splicing and includes transcripts that contain poly(A) signals in upstream introns (IIIi) or exons (IIIe) and, similarly to type II, results in several polyadenylated isoforms (Figure 1.3) (Lutz 2008; Lutz and Moreira 2010). While type II alternative polyadenylation does not have a direct effect on the encoded protein product, the resultant protein levels may be different between polyadenylated isoforms due to differences in mRNA stability, translation efficiency and/or other RNA functions. Both IIIi and IIIe types of alternative polyadenylation would result in different protein isoforms and/or affect mRNA stability (Lutz and Moreira 2010). Interestingly, approximately 20% of human genes have been found to contain at least one intronic alternative polyadenylation site (Tian et al., 2007), resulting in different mRNA isoforms that are predicted to encode different proteins. The same study found that the presence of weak 5’ splice sites and intron size correlated with intronic polyadenylation events, suggesting a connection between alternative polyadenylation and alternative splicing. In
addition to the effect that alternative polyadenylation has on the regulation of mRNA processing by protein-binding factors, usage of alternative poly(A) signals is also predicted to affect regulation of transcripts by microRNAs (miRNAs) which are known to either inhibit mRNA translation or induce transcript degradation (Filipowicz et al. 2008). Since mRNA untranslated regions (UTRs) are most common locations for miRNAs binding, longer UTRs will have higher number of miRNA binding sites, so that modulation of the UTR lengths by alternative poly(A) usage would affect gene expression (Legendre et al., 2006; Sandberg et al., 2008; Ji et al., 2009; Mayr and Bartel 2009).

Importantly, several studies have found that alternative polyadenylation may be differentially regulated under different physiological conditions (as reviewed in Edwalds-Gilbert et al., 1997; Lutz and Moreira 2010). A prominent example of such regulation is usage of alternative polyadenylation in T cell activation. It has been shown that alternative polyadenylation of a critical transcription factor that plays a major role in inducing transcription during immune response, NF-AT (for “nuclear factor of activated T cells”), is regulated upon T cells activation (Chuvpilo et al, 1999) such that alternative poly(A) usage directs a switch in synthesis from a longer NF-AT isoform in naïve T cells to a shorter isoform in effector T cells. The authors speculated that different NF-AT isoforms regulate transcription of different classes of cytokine genes. Production of different isoforms of another protein important for immunoregulation, FKBP12 (for FK506-binding protein, 12 kDA), is also regulated by alternative polyadenylation in a manner dependent on T cell activation (Peattie et al., 1994). Moreover, a global analysis comparing resting and activated murine CD4+ T lymphocytes has found an increase in
the usage of proximal poly(A) sites and consequent 3’ UTR shortening upon T cell activation (Sandberg et al., 2008). Examples of regulated alternative polyadenylation in other cellular systems (e.g., mouse embryogenesis and gametogenesis, as well as cancer cell lines) (Ji et al., 2009, Mayr and Bartel 2009, Lutz and Moreira 2010 and references therein) also exist.

Lastly, activation of alternative intronic poly(A) sites (Figure 1.3, type IIIi) has recently been found to regulate production of truncated, secreted isoforms of 31 receptor tyrosine kinases (RTKs) (Vorlová et al., 2011). Consistent with studies in HeLa cells, these IPA events were shown to be regulated by the U1 snRNP in splicing-independent manner (Kaida et al. 2010; Vorlová et al., 2011). Moreover, production of soluble RTKs with dominant-negative functions was found to have significant effects on signaling through these receptors and implications for therapeutic intervention (Vorlová et al., 2011). Together, these data suggest that alternative polyadenylation is emerging as a widespread mechanism of regulating gene expression.

1.4 Alternative splicing and alternative polyadenylation in human disease

Given the complexity of regulation of both splicing and polyadenylation, it is not surprising that aberrant functioning of any of the components involved in this regulation may lead to disease. Mutations to any elements of the “splicing code” and polyadenylation signals and/or changes in expression levels and functioning of the protein trans-acting factors that regulate either of the processes can directly cause
disease, while genetic variation in these factors can predispose to certain conditions or modify their severity/progression.

The contribution of alternative splicing to disease has been widely studied. Several prominent examples of human diseases that may be influenced by splicing defects include cystic fibrosis, a number of muscular atrophies, retinitis pigmentosa, various autoimmune diseases and cancer (reviewed in Lynch 2004; Wang and Cooper 2007; Cooper et al., 2009; Evsyukova et al., 2010; Poulos et al., 2011). Mutations in the polyadenylation signal sequences or misregulation of the trans-acting protein factors that regulate polyadenylation have been shown to contribute to a number of diseases as well (as reviewed in Danckwardt et al., 2008), but studies on the role of alternative polyadenylation in disease have been limited (Chen 2006). Alternative polyadenylation has been linked to cancer (Betticher et al., 1995; Campigli Di Giammartino et al., 2011) and the autoimmune condition IPEX (immune dysfunction, polyendocrinopathy, enteropathy, X-linked), in which mutations leading to alternative poly(A) usage in transcripts encoding FOXP3 protein result in an unstable mRNA and malfunctioning of regulatory T cells (Bennett et al., 2001).

1.4.1 Alternative splicing of interleukin 7 receptor (IL7R) and multiple sclerosis

Autoimmunity, or the inappropriate reaction to self-antigens, underlies a number of diseases, some organ-specific and some systemic. Many (if not all) autoimmune diseases are thought to be complex, with both environmental and multiple genetic factors determining pathology. As is the case for any other disease in which aberrant splicing
plays a role, autoimmune diseases can result from mutations in both cis-acting regulatory elements and trans-acting factors.

Multiple sclerosis (MS) is a debilitating neurodegenerative disorder of the central nervous system (CNS) that is thought to be mediated by T-cell autoimmunity (Trapp and Nave 2008). Disease features include penetration of lymphocytes into the CNS, induction of local inflammation and immune responses, demyelination and axon damage (McFarland and Martin 2007; Trapp and Nave 2008; Zuvich et al., 2009). MS preferentially affects young adults and shows a strong gender preference with females being affected 2-3 times more frequently than males (Duquette 1998). The course of the disease varies greatly among individual patients, ranging from a mild, or even asymptomatic, disease to severe disability. MS has been divided into four clinical subtypes as a means of predicting prognosis and response to therapy: relapsing remitting (85% of diagnosed MS cases), primary progressive (10%), progressive relapsing (5%) and secondary progressive (develops in 50-80% of relapsing remitting patients) (Coyle 2000). More than 400,000 individuals are affected with MS in the US and up to 2.5 million worldwide (Coyle 2000; also available at www.nationalmssociety.org). Importantly, there is currently no treatment to significantly alter the long-term prognosis of the disease.

A disease of complex pathology, MS is caused by a combination of environmental (e.g., viral infection, latitude, vitamin D) (Kakalacheva and Lunemann 2011) and genetic factors. However, for many years the only genetic locus consistently linked to the disease has been the human leukocyte antigen locus, or HLA (Schmidt et al. 2007; Luckey et al. 2011). Yet, the association of HLA with MS explains less than 50%
of the genetic basis of the disease (Gregory et al., 2007). To identify additional MS susceptibility genes, multiple family-based and case-control studies have been performed on MS populations (reviewed in Fernald et al, 2005; Fugger et al. 2010; Baranzini 2011), including a recent large-scale genome-wide association study (GWAS) of over 9,000 MS patients (The International Multiple Sclerosis Genetics Consortium et al. 2011).

Interleukin 7 receptor (IL7R) has been identified in most of these studies as a susceptibility gene for multiple sclerosis. Specifically, a non-synonymous single nucleotide polymorphism (SNP) rs6897932 (T→C, T244I) within IL7R exon 6 shows the strongest association with MS in four independent family-based or case-control data sets (overall P=2.9x10⁻⁷) (Gregory et al. 2007; Lundmark et al 2007; the International Multiple Sclerosis Genetics Consortium 2007, 2009; Figure 1.4). The result for rs6897932 is due to overtransmission of the “C” risk allele to offspring affected with MS and is independent of the known HLA effect. Importantly, alternative splicing of IL7Rα transcripts was affected by this MS-associated SNP (Gregory et al., 2007).

The receptor for interleukin 7 (IL7) consists of two chains: the IL7Rα chain (CD127), which provides ligand specificity, and the common cytokine receptor γc-chain (CD132). IL7Rα is expressed almost exclusively on cells of the lymphoid lineage and is crucial for development and maintenance of T cells (as reviewed in Mazzhuccelli and Durum 2007). IL7Rα expression is tightly regulated, strictly based on whether or not these cells need to receive survival or, in some cases, differentiation signals. In addition to developing T cells in the thymus, naïve and memory T cells in the periphery require signaling through IL7Rα for their survival. Binding of IL7 to its receptor triggers a signaling cascade which, among other functions, ultimately results in transcriptional
induction of anti-apoptotic genes (e.g., Bcl-2) and re-distribution of cell-death proteins (e.g., Bax and Bad), thereby protecting T cells from apoptosis (Jiang et al. 2004 and references therein).

*IL7Rα* pre-mRNA consists of eight exons, with exon 6 coding for the transmembrane domain. Two isoforms of *IL7Rα* have been identified based on alternative splicing of exon 6: a membrane-bound isoform (*rIL7Rα*), in which exon 6 is included, and a soluble isoform (*sIL7Rα*), in which exon 6 is skipped (Figure 1.5). *sIL7R* lacks both the transmembrane and cytoplasmic domains of the receptor (Goodwin et al. 1990; Korte et al. 2000; reviewed in Evsyukova et al. 2010). Importantly, both isoforms are able to bind IL7 with high affinity (Goodwin et al. 1990; Rose et al. 2009).

Data from our laboratory has established that the MS-associated SNP affects splicing of the *IL7R* exon 6 by increasing its skipping 2-fold in cell culture (Gregory et al., 2007) and 4-5-fold when RNA from MS patients was analyzed (unpublished results). Additionally, it was reported that expression of *IL7Rα* mRNA was elevated in the cerebrospinal fluid of individuals with MS (Lundmark et al., 2007) and in whole blood samples from patients with two types of MS (primary progressive and relapsing remitting), and that increased expression of *sIL7Rα* correlated with a particular *IL7Rα* haplotype that was more common in the primary progressive MS patients (McKay et al., 2008). While the possibility that the effect of the MS-associated SNP is due to different properties of the transmembrane receptor (and, thus, aberrant signaling through the receptor), has not been excluded, these data suggest that reduced inclusion of exon 6 and the concomitant increase in *sIL7R* could be causally linked to immune dysfunction and MS. Characterizing exon 6 splicing control would shed light not only on the regulation of
*IL7R* expression and its role in T cell functioning, but may also be relevant to establishing a connection between *IL7R* splicing and MS.

### 1.5 Focus of this work

Different splicing isoforms of *IL7R* have been reported and have been implied to play a role in disease (Goodwin et al., 1990; Korte et al., 2000; Gregory et al., 2007; Rose et al., 2009). However, detailed characterization of these isoforms, either at the RNA or protein level, is lacking. Given the role that *IL7R* plays in T cell biology and the potential connection between *IL7R* splicing and MS, we aimed to investigate alternative splicing of *IL7R* exon 6 both at the level of *cis*-acting regulatory elements and *trans*-acting protein factors. Moreover, other mechanisms of posttranscriptional regulation of *IL7R* have been largely unexplored. At least 9 potential RNA isoforms are listed in the Ensemble Genome browser ([www.ensembl.org](http://www.ensembl.org), Figure 1.6), most of which are predicted to encode proteins, but the supporting data for many of the listed transcripts is limited or lacking. Characterization of these and other potential isoforms of *IL7R* would shed light not only on the mechanisms of *IL7R* regulation, but also on T cell biology and disease.

In Chapter 2, I will describe mutagenesis experiments that identified multiple potential *cis*-acting regulatory elements within *IL7R* exon 6. I will focus on the area surrounding the MS-associated SNP T→C and show that at least two regulatory elements, one exonic splicing enhancer and one exonic splicing silencer, are present around the SNP. I will provide data that confirm these elements are functional. The pattern of splicing regulation that was observed in DT and HeLa cells was also
recapitulated in Jurkat cells, a T cell line that naturally expresses IL7R. Additionally, I will use mutagenesis to demonstrate that IL7R exon 6 5’ splice site is inherently weak, which suggests that this splice site is most likely the regulated splice site that influences exon 6 inclusion. I will also show experiments that identified an intronic polyadenylation signal AAUAAA within IL7R intron 6, in the direct vicinity of exon 6 5’ splice site. Experiments in which this site was mutated resulted in increased exon 6 inclusion, suggesting that this site is important in exon 6 splicing regulation. I will conclude with a summary of all identified regulatory elements and future directions that will help determine each element’s role in exon 6 splicing regulation.

In Chapter 3, I will present findings of the RNA affinity chromatography/mass spectrometry experiments that were performed to identify trans-acting protein factors that bind IL7R exon 6 and surrounding intronic regions and regulate exon 6 inclusion. Components of the mammalian cleavage and polyadenylation specificity complex (CPSF) were identified by these experiments as IL7R pre-mRNA binding candidates. Given our finding of an intronic polyadenylation site in intron 6 of IL7R, we investigated if CPSF binding mediated the effect seen by the mutation of this site on exon 6 inclusion. I will show that CPSF1 depletion by RNAi resulted in increased exon 6 inclusion, consistent with mutations to the AAUAAA site. I will then describe identification of novel polyadenylated products that are produced at low frequency from the intronic AAUAAA site. I will conclude with future directions that aim at determining the potential role of these alternative polyadenylated isoforms in T cell biology and MS.

In Chapter 4, I will discuss implications of the effect that alternative splicing of exon 6 together with intronic polyadenylation and/or CPSF complex assembly on the
AAUAAA site in intron 6, would have on generation of different IL7R protein isoforms and their potential role in MS pathogenesis. I will conclude with proposing avenues for future research that build on the findings of this thesis and establish a connection between \textit{IL7R} posttranscriptional regulation and MS.
Figure 1.1: Mechanisms of alternative pre-mRNA splicing.

a) Alternative use of 5' splice sites; b) alternative use of 3' splice sites; c) intron retention; d) alternative use of a cassette exon; e) alternative use of two mutually exclusive exons. This figure is taken from Evsyukova et al., 2010.
Figure 1.2: Mammalian 3'-end processing machinery. Polyadenylation signal sequences as well as four multi-subunit protein complexes that bind them, are shown. This figure is taken from Mandel et al., 2008.
Figure 1.3: Types of alternative polyadenylation

Light blue boxes represent untranslated regions; dark blue boxes represent coding regions; lines represent introns. Type I alternatively polyadenylated transcripts have only one polyadenylation signal AAUAAA in their 3’-most exon; type II transcripts have more than one AAUAAA in their 3’-untranslated regions; type III refer to transcripts that contain alternative polyadenylation signals in the upstream introns (IIIi) or exons (IIIe). AUG represents translation start site; “Stop” represents translation stop codon. This figure is taken from Lutz and Moreira 2010.
Figure 1.4: Association analysis of SNPs from *IL7R*.

SNPs from *IL7R* in the US, European and combined family-based and case-control data sets. Lines represent introns; filled regions represent exons; unfilled regions represent untranslated regions. SNP rs6897932 represents the only significant SNP. This figures is taken from Gregory et al., 2007.
**A.**

![Diagram showing exon inclusion and skipping](image)

**B.**

**Nucleotide sequence**

<table>
<thead>
<tr>
<th>Exon Inclusion</th>
<th>Exon Skipping</th>
</tr>
</thead>
</table>
| Incl ...AATAGCTCAGGGAGATGGATCCTATCTTACTAAATCAGCATTTGAGTTTTTTCTCTGTCG | Skip ...AATAGCTCAGG--
| Incl TCTGGGCTCTTGCCCTGTGTTATGAAAGATTTAAAT | Skip ----------------------------------------|
| Incl TCCCCGATCATAAGAAGACTCTGGAACATCTTTGTAAGAAACCAAGAAA | Skip ------AATTTAAAT--

**Protein sequence**

<table>
<thead>
<tr>
<th>IL7R Protein</th>
<th>SIL7R Protein</th>
</tr>
</thead>
</table>
| rIL7R ...NSSGEMDILLIISILSSFSSVALLVILACVLWKKRIKIVWPSLPDHKTKLEHLCCKPRKLN... | sIL7R ...NSSGSLSYGPVIPVRRLWNIFVRNEDKISop

*Stop*
Figure 1.5: Alternative splicing of *IL7R* exon 6.

A. *Top*, genomic and pre-mRNA structure of the *IL7R* gene is shown. SNPs selected for genotyping in Gregory et al., 2007, are listed above. The MS-associated SNP, rs6897932, is highlighted. *Bottom*, alternative splicing of the *IL7R* pre-mRNA leads to the production of two isoforms: a membrane-bound (*rIL7R*, exon 6 included) or soluble (*sIL7R*, exon 6 skipped). The transmembrane isoform of the receptor consists of extracellular (EX), transmembrane (TM) and cytoplasmic (CYTO) domains. This figure is taken from Evsyukova et al., 2010. B. Skipping of exon 6 leads to a translation reading frame shift and translation termination due to a premature stop codon. *Top*, nucleotide sequence of a region of the *IL7R* mRNA encompassing exon 6 (green), exon 7 (blue) and parts of exons 5 and 8. An asterisk (*) denotes nucleotide and amino acid sequences that are the same between the *IL7R* isoforms that include (“Incl”) and skip (“Skip”) exon 6. Premature stop codon in exon 8 of the *IL7R* isoform that encodes *sIL7R* is highlighted in yellow. *Bottom*, amino acid sequence corresponding to the nucleotide sequence shown above. Transmembrane region is highlighted in red. Premature stop codon (“*Stop*”) in *sIL7R* is highlighted in yellow.
Figure 1.6: *IL7R* isoforms based on data from the Ensembl Genome Browser (www.ensembl.org).

A total of nine isoforms are shown, 7 of which are predicted to code for protein. Lines represent introns. Boxed regions represent exons. Unfilled boxes represent untranslated regions. Highlighted isoforms encode rIL7R (exon 6 included) and sIL7R (exon 6 skipped) protein products.
Chapter 2: Identification of *cis*-acting regulatory elements that regulate splicing of *IL7R* exon 6

2.1 Introduction

Bioinformatic analyses have determined that only about one half of the information required to correctly define an exon/intron boundary is contained within the “core” splicing motifs (3’ and 5’ splice site and branch point sequences) (Lim and Burge 2001). Thus, *cis*-acting regulatory elements within both exons and introns (exonic/intronic splicing enhancers, ESEs/ISEs and silencers, ESSs/ISSs) play a critical role in splicing regulation. Exonic splicing regulatory elements have been more extensively studied than intronic elements (Wang and Burge 2008), however, both classes are important in splicing regulation. ESEs are very abundant and it is likely that many if not all exons contain internal ESEs (Fairbrother et al., 2002; Wang and Burge 2008). ESEs usually function by recruiting different SR proteins. While ESSs may be somewhat less prevalent in exonic sequences (Zhang and Chasin 2004), they play a role at least as important as enhancers (Wang et al., 2004; Zhang and Chasin 2004) and act by recruiting hnRNP proteins. Both classes of regulatory elements consist of rather degenerate sequences, but comprehensive lists of ESE and ESS oligonucleotides have been generated to predict most common sequences (reviewed in Wang and Burge 2008). Importantly, these lists intend to predict outcomes of exonic mutations that frequently affect splicing (Cartegni et al., 2002; Fairbrother et al., 2002; Wang and Burge 2008).

A number of exonic regulatory elements have been discovered, including some in immunologically-relevant genes. CD45, a prototypic receptor tyrosine phosphatase that is
essential for regulation of signaling through the T-cell receptor, serves as a prominent example. Splicing of CD45 variable exon 4 is regulated by hnRNP L that binds to an ESS within this exon. This regulation occurs in a manner that is dependent on the activational state of the T cell (Lynch and Weiss 2001; House and Lynch 2006; Rothrock et al., 2003, 2005). Given that this and other trans-acting protein factors can be differentially expressed in certain tissues or physiological states of the cell (e.g., T cell activation), cis-acting elements are likely to exhibit their regulatory role in a tissue- (or cellular state)-specific manner.

*IL7R*, which is crucial for T cell development and maintenance, is expressed differentially based on the developmental stage of the T cell and availability of ligand (IL7). Several exons within the *IL7R* pre-mRNA are alternatively spliced (Goodwin et al., 1990; Pleiman et al., 1991; Korte et al., 2001; Rose et al., 2009; Figures 1.4 and 1.5); however, studies exploring any of these splicing events and/or the role that alternative isoforms play during different stages of T cell development, are limited. To test whether alternative splicing of exon 6 was differentially affected in transcripts from the “T” or the MS-associated “C” alleles, Gregory et al. analyzed exon 6 inclusion both *in vitro* and *in vivo* (Gregory et al., 2007). For the *in vitro* analysis, minigenes containing either “T” or “C” alleles of rs6897932, as well as parts of flanking introns, were created. When transfected into a variety of cell lines, transcripts from the minigenes containing the MS-associated “C” allele show an approximately two-fold increase in exon 6 skipping when compared with transcripts containing the “T” allele. Based on additional mutagenesis analysis (transversions to either “G” or “A” at the SNP position, as well as substitutions
around the SNP), the authors concluded that the disease-associated “C” allele affects exon 6 alternative splicing by augmenting the action of an exonic splicing silencer (ESS).

The *in vitro* results were supported by at least two lines of evidence *in vivo*: first, quantitative real-time PCR for allele-specific *IL7Rα* expression in peripheral blood mononuclear cells (PBMCs) from healthy controls showed a significantly lower expression of the exon 6-7 amplicon as observed for carriers of the “C” allele (Gregory et al., 2007). Second, semi-quantitative RT-PCR analysis of PBMCs from MS patients who are homozygous for the “C” allele showed a 4-5-fold increase in exon 6 skipping compared to patients who are homozygous for the “T” allele (Gregory et al., unpublished results). Additionally, expression of *IL7Rα* mRNA was elevated in the cerebrospinal fluid of individuals with MS in comparison with individuals with other neurological diseases. This suggests that changes in *IL7Rα* expression have pathophysiological significance. Lastly, McKay et al. established that expression of *sIL7Rα* mRNA was significantly elevated in whole blood samples from patients with two types of MS (primary progressive and relapsing remitting), and that increased expression of *sIL7Rα* correlated with a particular *IL7Rα* haplotype that was more common in the primary progressive MS patients. Taken together, these data suggest that: 1) there is a decrease in the ratio of transmembrane to soluble isoforms of *IL7Rα* in MS, and 2) aberrant alternative splicing of the *IL7Rα* exon 6 causes this change in isoform ratio. Given the importance of splicing regulation of exon 6 of *IL7R* and our interest in the effect that the MS-associated SNP has on alternative splicing of this exon, we employed mutagenesis to investigate *cis*-acting regulatory elements within this exon.
In this chapter, I present mutagenesis studies of exon 6 and surrounding intronic sequences that were performed in human cell lines. Several cis-acting regulatory elements that regulate inclusion of exon 6 were identified. We focused on two elements, one exonic splicing silencer and one exonic splicing enhancer, in the direct vicinity of the MS-associated SNP. We confirmed that these elements were functional and determined that exon 6 splicing regulation was similar in human T cells and non-T cell lines. We determined that the 5’ splice site of exon 6 is weak and is thus likely the regulated site responsible for exon 6 splicing regulation. We identified a consensus polyadenylation signal AAUAAA in intron 6 of *IL7R* 16 nucleotides downstream from exon 6 5’ splice site. Mutation to this site resulted in an increase in exon 6 inclusion, implying that this site is important for exon 6 splicing regulation. Together, these data provide the first comprehensive analysis of splicing regulatory elements within a critical transmembrane domain-coding exon within *IL7R*.

### 2.2 Materials and Methods

**Plasmids and cloning**

All substitution and deletion mutants were made in the context of the cloning vector pI-12 described previously (Carstens et al.; Gregory et al. 2007). For minigene schematics, see Figure 2.1. Briefly, mutant constructs were made by chimeric PCRs and purified with QIAquick PCR purification kit (Qiagen). Forward primer corresponding to the T7 promoter and reverse primer corresponding to the Sp6 promoter were used as common primers in chimeric PCRs. Sequences for common and mutation-specific
primers are listed in Table 3. All primers were synthesized by Integrated DNA Technologies (IDT, Coralville, IA). PCR chimeras and pI-12 vector were digested with XbaI and XhoI (NEB), purified with the QIAquick PCR purification kit, and ligated with 400 u T4 DNA Ligase (NEB). All PCR products were confirmed by sequencing by the Duke University Sequencing Facility.

**Cell culture and transfections**

Rat DT cells were cultured in low-glucose Dulbecco’s Modified Eagle’s Medium (DMEM, Invitrogen/Gibco) supplemented with 10% fetal bovine serum (Gemini) and antibiotics (penicillin/streptomycin, Invitrogen/Gibco). Human HeLa cells were cultured in high-glucose DMEM (Invitrogen/Gibco) with 10% fetal bovine serum and antibiotics. Human Jurkat cells were cultured in Roswell Park Memorial Institute Medium 1640 (RPMI1640) supplemented with 10% fetal bovine serum, 1 mM sodium pyruvate, 10 mM HEPES, glucose up to 4.5 g/l and antibiotics (all reagents were from Invitrogen/Gibco). Transient minigene transfections into DT and HeLa cell lines were performed with Lipofectamine™ 2000 (Invitrogen) according to the manufacturer’s instructions. A total of 25 ng of minigene DNA was transfected per well in 24-well format (24-well plates were from BD Biosciences). A total of 7.5x10^4 – 1x10^5 cells were seeded into each well 24 hours prior to transfection. Minigene transfections into Jurkat cells were performed with FuGENE 6 (Roche) according to the manufacturer’s instructions. Briefly, 2x10^6 cells were freshly passed and seeded into each well of a 6-well plate (BD Biosciences) immediately before transfection and 1 µg minigene DNA was used per well. Cells were harvested 48 hours post-transfection. All transfections were performed in triplicate.
**RNA isolation and semi-quantitative RT-PCR of transfected minigenes**

Total RNA was extracted with TRIZOL Reagent (Invitrogen) according to the manufacturer’s instructions and treated with TURBO™ DNase (Ambion/Life Technologies). A total of 2 µg of DNase-treated RNA was used to synthesize first-strand cDNA with 250 ng of random hexamer primers (Invitrogen) and Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV RT, Invitrogen) following the manufacturer’s protocol. PCRs were performed as follows: 1 ul of the RT reaction was mixed with 200 nM forward (T7) and 200 nM reverse (Sp6) primers and 100 nM dNTPs in 100 mM Tris-HCl, pH 8.3, 500 mM KCl, 20 mM MgCl$_2$ and 0.25 µl Taq polymerase was added. [$\alpha$-32P] dCTP (3000 Ci/mmol, 10 mCi/ml, PerkinElmer) was added to a final concentration of 0.1 µCi/µl. PCR products were loaded onto 5% non-denaturing polyacrylamide gels, electrophoresed for 3 hours at 120 V, dried using a gel dryer and exposed to Molecular Dynamics or Amersham Hyperfilm-MP phosphorimager screens. Quantifications were performed using ImageQuant software (Molecular Dynamics). Percent exon 6 inclusion was calculated according to the formula: % Inclusion = ((Inclusion product)/(Inclusion + Skipped product))x100%. Each calculation was the result of an average of triplicate samples. Error bars represent the standard deviation.
2.3 Results

2.3.1 Multiple cis-acting regulatory elements in IL7R exon 6 regulate its inclusion

To dissect cis-acting exonic elements that are important for splicing of IL7R exon 6, we created substitution mutants spanning the whole length of exon 6 (sub(1-4) through sub(91-94)) in the context of the splicing reporter minigene pI-12 described previously (Carstens et al. 1998; Gregory et al, 2007). The reporter contains full-length exon 6 surrounded by 614 and 573 nucleotides (nt) of IL7R introns 5 and 6 respectively, placed between two constitutively spliced adenoviral exons U (for “upstream”) and D (for “downstream”) (Figure 2.1; construction of the pI-12 splicing vector has been described in Carstens et al., 1998, and references therein). All of the substitutions were transversions and were introduced in the context of the MS-associated “C” allele at position 25 of the exon. The parental “C” minigene served as a positive control while minigenes with disruptions of either the 3’ or 5’ splice sites (3’ ss mut and 5’ ss mut) served as negative controls. Control and mutant minigenes were transfected into rat DT cells (transfection into HeLa cells was also performed and yielded identical results), and exon 6 inclusion was ascertained by semi-quantitative RT-PCR 48 hours post-transfection.

As shown in Figure 2.2, and consistent with previous studies from our laboratory, 50-55% of the transcripts containing the MS-associated “C” allele included exon 6 and, as expected, inclusion was minimal when either of the splice sites (3’ and 5’) was
mutated. Transfection with the five-nucleotide substitution minigenes had multiple effects on exon 6 inclusion, with some substitutions (sub(5-9), sub(10-14), sub(26-30), sub(36-40)) resulting in a significant decrease in exon 6 inclusion (less than 10% inclusion), while others (sub(20-24), sub(41-45), sub(51-55) through sub(61-65), sub(71-75)) resulted in an increase to between 70 and 80% inclusion. While substitutions sub(1-4) and sub(91-94) also showed profound effects on inclusion (less than 10% inclusion), their effects likely disrupt integrity of the splice sites and are, thus, not informative.

Importantly, three substitutions that impacted exon 6 inclusion are located in the vicinity of the MS-associated SNP. Two of these decreased inclusion revealing the presence of potential splicing enhancers, ESEs 1 and 2 (nts 5-14 and 26-40, respectively) and one increased inclusion, revealing the presence of a potential splicing silencer, ESS1 (nts 20-24). Given our interest in the MS-associated SNP, we first focused on these three exonic regulatory elements.

Since several of the five-nucleotide substitutions had a significant effect on exon 6 inclusion, it was important to determine if this effect was due to mutating existing regulatory elements rather than inadvertently introducing novel ones. To differentiate between these two possibilities, we introduced deletions of the same sequences that were substituted in Figure 2.2. As shown in Figure 2.3, introducing deletions confirmed the presence of the enhancer element ESE2 and silencer element ESS1. While deletion analysis of the ESE1 element failed to recapitulate the effects seen with sub(10-14), it revealed a region of decreased inclusion in line with sub(5-9). Deletion of nt 31-35, within what we called ESE2, resulted in no change to exon 6 inclusion, which may suggest that the regulatory sequence that we previously labeled “ESE2” is, in fact, two
different enhancer elements (Figure 2.3). Overall, deletion analyses confirmed the presence of two regulatory elements (ESS1 and ESE2) in the close vicinity of the MS-associated SNP.

To further test for the presence of ESS1 and ESE2, we introduced larger substitutions (Figure 2.4). As expected, while substitution of the sequences corresponding to ESS1 (sub(15-24)) resulted in almost 100% exon 6 inclusion, substitution of the sequences corresponding to ESE2 (sub(26-40)) resulted in a profound decrease in inclusion (less than 20% of transcripts included exon 6) (Figure 2.4). As an independent approach to determine whether cis-acting elements identified by our mutational analyses were indeed functional regulatory elements, we cloned ESE2 (nt 26-40) into heterologous context - a splicing reporter minigene that contains an alternatively spliced exon IIIc of the fibroblast growth factor receptor 2 (FGFR2) (Figure 2.6A). In DT cells, exon IIIc is not normally included. We reasoned that introducing a potential enhancer (ESE2) would result in exon IIIc being included in DT cells. As expected, transcripts containing IIIc-ESE2 chimeras included exon IIIc, while no inclusion was observed when a size-matched sequence that did not show any effect on exon 6 inclusion, was introduced into IIIc (Figure 2.6B). Additionally, introducing nt (1-14) of the IL7R exon 6 into FGFR2 exon IIIc resulted in inclusion of IIIc in DT cells, albeit to a less extent than with ESE2 (data not shown), suggesting the presence of another enhancer within exon 6, which most likely corresponds to nts 5-9 as hinted by our substitution and deletion analysis. These data served as an additional confirmation that our mutational analyses identified functional cis-acting elements that regulate inclusion of the IL7R exon 6. Together, deletion and substitution data suggest that at least two cis-acting regulatory
elements are present in the vicinity of the MS-associated SNP (ESS1 an ESE2). Importantly, previous data from our laboratory suggested that the MS-associated SNP augments an existing exonic splicing silencer within exon 6 (Gregory et al. 2007), and our current data suggest that this is likely ESS1.

While mutating exon 6 regulatory sequences showed the same effects on exon 6 inclusion in the context of rat DT and human HeLa cells (Figures 2.2-2.4, Gregory et al., 2007, and unpublished results), we aimed to test this regulation in T cells that naturally express IL7R. To this end, we used Jurkat cells, a T cell line. As shown in Figure 2.5, although absolute level of exon 6 inclusion was much higher overall, the effects of disrupting ESS1 (sub(20-24)) or ESE2 (sub(36-40)) resulted in increased and decreased exon 6 inclusion, respectively (compare Figures 2.2 and 2.5). We thus conclude that splicing regulation of exon 6 as seen in our experimental system (rat DT cells or human HeLa cells), recapitulated splicing regulation of exon 6 in T cells.

2.3.2 An intronic cis-acting element associated with 3’end formation and polyadenylation downregulates IL7R exon 6 inclusion

Examination of the sequences of exon 6 and introns 5 and 6 of IL7R revealed an AAUAAA polyadenylation signal 16 nt downstream from the 5’ splice site of exon 6 (Figure 2.7A). Therefore, we investigated whether intronic polyadenylation (IPA) contributes to the regulation of exon 6 splicing. The sequence AAUAAA serves as a binding site for cleavage and polyadenylation specificity factor 1 (CPSF1) (Keller et al., 1991; Murthy and Manley 1995) and mutations to this element have been shown to affect splicing (Niwa and Berget 1991; Cooke et al. 1999). We mutated the AAUAAA site
(pAmut in Figure 2.7A) within IL7R intron 6 and evaluated the effect of this mutation on exon 6 inclusion. Mutation of the consensus poly(A) signal in intron 6 of IL7R resulted in a significant increase in exon 6 inclusion, from 50% to about 80% in the context of the “C” minigene and from <10% to 30% in the context of the sub(36-40) minigene (Fig. 2.7B). These experiments suggested that the intronic poly(A) site affects splicing of exon 6 by suppressing its inclusion.

### 2.3.3 The 5’ splice site of IL7R exon 6 is weak

Mammalian splice sites commonly diverge from their consensus sequences (“weak” splice sites), and are thus prone to regulation by trans-acting factors. We examined exon 6 3’ and 5’ splice site sequences, ccag↓GG and AG↓gugacc, respectively (arrows designate exon/intron boundary, exons are in upper case, introns are in lower case). While exon 6 3’ splice site matches its consensus sequence (NPyAG↓PuN, where Py stands for pyrimidine and Pu for purine), exon 6 5’ splice site diverges from the consensus (AG↓GUPuAGU). To test whether this site was inherently weak, we introduced a mutation in which this splice site would match the consensus precisely (AG↓gugacc to AG↓gugagu) ([Figure 2.8](#)). The presence of the consensus 5’ splice site resulted in almost 100% inclusion of exon 6 in the context of both “C” and sub(36-40) minigenes. We thus concluded that exon 6 5’ splice site is inherently weak and will thus most likely represent the splice site that is regulated by the exonic splicing elements and trans-acting protein factors.
2.4 Discussion

Experiments presented in this chapter represent the first detailed characterization of \textit{cis}-acting elements that regulate splicing of exon 6 of \textit{IL7R}, which is the only exon that encodes the transmembrane domain of the receptor. Previous data from our laboratory have suggested that the MS-associated SNP affects splicing of this exon by increasing its skipping, and the same trend has also been observed in patients (Gregory et al., and unpublished results). We provided the first list of potential \textit{cis}-acting elements that are important for splicing of this exon.

Multiple potential elements have been identified. However, given our interest in the MS-associated SNP, we focused on the elements surrounding the SNP (the first 40 nt of exon 6). Our analysis indicated that other potential important regulatory elements are present within exon 6 (for example, the region between nt 51 and 65 shows an increase in inclusion, while sub(86-90) leads to a decrease in inclusion) which may be relevant for exon 6 splicing and will require further validation. It would be important to evaluate the effect that these additional elements have on exon 6 inclusion, as well as characterize the elements that we have identified (specifically, ESS1 and ESE2) further. ESS1 and ESE2 are located right downstream and upstream from the MS-associated SNP, respectively, but it would be interesting to define their exact boundaries by additional mutagenesis. The element that we called ESE1 will also require further study as while the effect that it has on exon 6 inclusion might be due to the disruption of the 3’ splice site, putting this element in heterologous context resulted in inclusion of an exon that is normally skipped in DT cells (data not shown). This suggests that an enhancer is preset in this region,
possibly immediately after the splice site (since substitutions and deletions of nt 5-9 resulted in a significant decrease in inclusion).

We have not attempted to systematically mutagenize intronic sequences of *IL7R* introns 5 and 6, however, we were able to identify an intronic element that corresponds to the consensus polyadenylation signal – AAUAAA – 16 nt downstream from exon 6 5’ splice site. Mutation to this site resulted in a significant increase in inclusion of this exon. We will address a possible mechanism via which this site is acting in the next chapter, but considering the effect that this site has on splicing of exon 6 and an interplay between splicing and polyadenylation sequences, it would be important to extend mutagenesis studies into this intron to characterize sequences that are potentially important for both splicing of exon 6 and intronic polyadenylation.

Lastly, we showed that exon 6 5’ splice site is inherently weak, and mutating this site “back” to the consensus sequence results in almost complete inclusion of exon 6. This result suggests that exon 6 5’ splice site is most likely the regulated splice site that is responsible for exon 6 splicing regulation. It would be interesting to determine whether the MS-associated SNP or any of the *cis*-acting regulatory elements that we identified in our mutagenesis studies act via regulating this splice site. Experiments with single-intron minigene constructs, in which only one splice site of *IL7R* exon 6 is functional (to exclude potential influence of the second splice site on splicing regulation), will be needed to address this issue. Considering other alternative splicing/alternative processing events that have been described or predicted for *IL7R* pre-mRNA, it would be interesting to establish which *cis*-acting sequences regulate those events as well, with the goal of
building a framework for detailed characterization of all of the *IL7R* alternatively processed isoforms.
Figure 2.1 Minigenes that report on exon 6 inclusion.

*Top,* minigene schematics; T7 and Sp6 represent sites for common PCR primers; the MS-associated SNP is highlighted with an asterisk; XbaI and XhoI sites represent restriction endonuclease sites used for cloning mutant minigenes. *Bottom,* sequences of the substitution (sub(1-4) through sub(91-94) and also sub(15-24) and sub(26-40)) and deletion (del(1-4) through del(36-40)) mutants used in transfection experiments. Exonic sequences are in upper-case letters; intronic sequences are in lower-case letters; the MS-associated SNP is shown with a capital letter and highlighted with an asterisk. Regions of the exonic/intronic sequences that were either substituted or deleted, are boxed.
Figure 2.2: Mutagenesis across exon 6.

Substitutions were introduced into exon 6 in the context of the MS-associated "C" allele. Control and mutant minigenes were transfected into cell lines and total RNA was analyzed for exon 6 inclusion 48 hours post-transfection. All transfections were done in triplicate. Error bars represent the standard deviation. The MS-associated SNP is shown with an arrow.
Deletions of the same sequences that were substituted in Figure 2.2, were introduced in the context of the MS-associated "C" allele. Mutant minigenes were transfected into cell lines, and exon 6 inclusion analyzed as in Figure 2.2. Transfections were performed in triplicate. Error bars represent the standard deviation. The MS-associated SNP is designated with an arrow. Only the first 40 nt of exon 6 (the area surrounding the SNP) were used for deletion analysis.
Figure 2.4: Larger substitutions encompassing regulatory elements ESS1 and ESE2.

Larger substitutions encompassing regulatory elements ESS1 and ESE2 (as identified in experiments shown in Figures 2.2 and 2.3) were introduced in the context of the MS-associated SNP. Experiments were done as in Figures 2.2 and 2.3.
Figure 2.5: Inclusion of exon 6 in Jurkat cells.

Control ("C") and mutant (sub(20-24) and sub(36-40)) minigenes were transfected into Jurkat cells (a T cell line that expresses IL7R naturally). Exon 6 inclusion was quantified as in Figures 2.2-2.4. Transfections were done in triplicate. Error bars represent the standard deviation.
Figure 2.6: ESE2 works in heterologous context.

A. ESE2 (nt 26-40 of IL7R exon 6) was introduced into minigenes containing FGFR2 exon IIIc surrounded by FGFR2 introns. Inclusion of IIIc interrupts EGFP open reading frame, while skipping of IIIc results in a functional EGFP product. In DT cells, IIIc is normally skipped. B. IIIc-ESE2 chimeric minigenes transfected into DT cells included exon IIIc, confirming that ESE2 functions as a splicing enhancer. The size of the IIIc inclusion product is 497 bp; the size of the skipped product is 352 bp. Products running at 586 and 441 bp correspond to products including and skipping IIIc, respectively, in which the downstream PCR primer annealed to a non-specific sequence in the vector.
Figure 2.7: An intronic element associated with cleavage and polyadenylation influences exon 6 inclusion.

A. Schematics of the \textit{IL7R} exon 6/intron 6 region. Consensus polyadenylation signal AAUAAA is present 16 nt downstream from the exon 6 5′ splice site, followed by the cleavage signal, “CA”, 19 nt downstream from the poly(A) site. Substitution mutation used to assess effect of this site on exon 6 inclusion is shown (pAmut). B. Mutation to the intronic AAUAAA signal results in increased inclusion of exon 6. Two minigenes, “C” and sub(36-40), were analyzed, either with (right) or without (left) the polyA signal mutation in intron 6.
Figure 2.8: *IL7R* exon 6 5' splice site is weak.

*IL7R* exon 6 5’ splice site (AG/gugacc, exon in upper case, intron in lower case) was mutated to match the consensus 5’ splice site sequence (AG/gugagu). The presence of the consensus 5’ splice site in the context of both “C” and sub(36-40) minigenes results in almost complete exon 6 inclusion.
Chapter 3: Identification of cleavage and polyadenylation specificity factor 1 (CPSF1) as a regulator of IL7R exon 6 splicing

3.1 Introduction

_Cis_-acting splicing regulatory elements exert their effects via _trans_-acting protein factors that bind these elements and regulate splicing. Expression of some of these protein factors is ubiquitous, while expression of others is regulated in a developmental stage-, tissue- or temporally-specific manner. Some of the most well-studied _trans_-acting protein factors include PTB (PTBP1, also known as hnRNP I) (Wagner and Garcia-Blanco 2001, 2002 and references therein; Amir-Ahmady et al., 2005; Sharma et al., 2005; Boutz et al., 2007; Xue et al., 2009; Tang et al., 2011), Fox family of proteins (Jin et al., 2003; Baraniak et al., 2006; Kuroyanagi 2009; Tang et al., 2009), hnRNP proteins H and F (Black 2003; Mauger et al., 2008 and references therein), Muscleblind family of proteins (Ho et al., 2004; Pascual et al., 2006; Lukong et al., 2008), Nova family of proteins (Jensen et al., 2000; Dredge and Darnell 2003; Dredge et al., 2005; Yano et al., 2010) as well as others. Protein factors that regulate splicing of immunologically-relevant genes have also been identified: regulation of the alternative splicing of CD45, a prototypic protein tyrosine phosphatase that is essential for T cell activation, by hnRNP L and LL, serves as a prominent example (Lynch 2004; Heyd and Lynch, 2010 and references therein; Motta-Mena et al., 2011; Yabas et al., 2011; Preußner et al., 2012).

Interestingly, pre-mRNAs for several genes that are important for the formation of myelin, and, thus, might be relevant to MS pathogenesis, are alternatively spliced and
have been studied at the level of cis-elements and trans-acting factors. Protein factors that regulate splicing of the major myelin component, myelin proteolipid protein (PLP) gene, have been identified (Wang et al., 2005, 2007, 2008, 2009, 2011). Another intriguing example of a trans-acting factor that regulates splicing of several myelin structural genes (including myelin basic protein, MBP, PLP and myelin-associated glycoprotein, MAG) is Quaking (QKI), a single hnRNP K homology (KH) RNA-binding domain protein. Although no data regarding the involvement of the human homolog of QKI in MS pathogenesis are available, mice possessing a quakingviable (Qkv) mutation have a twitching phenotype and show severe demyelination of the CNS, with reduced levels of QKI correlating with the severity of demyelination (Wu et al., 2002; Ryder and Williamson 2004; Zhao et al., 2006). The same studies found that the reduction in levels of QKI resulted in aberrant splicing patterns of the myelin structural genes.

As described in Chapter 2, multiple exonic and intronic cis-acting elements were identified in IL7R exon 6. However, nothing is known about protein factors that bind these elements and regulate exon 6 splicing. In addition to their importance as exon 6 splicing regulators, these factors are potential MS risk factors (via SNPs in their genes or due to differential expression levels). We employed tobramycin RNA affinity chromatography followed by mass spectrometry to identify trans-acting protein factors that bind exon 6 and regulate its splicing (Hartmuth et al. 2004; Ward et al. 2011, and also Hartmuth K., personal communication). Several RNA-binding proteins were identified as potential regulators of exon 6 splicing. Interestingly, we identified cleavage and polyadenylation specificity factor subunit 1 (CPSF1) as well as other components of the mammalian cleavage and polyadenylation machinery among potential regulators.
Given that splicing and polyadenylation interact, and that CPSF1 has been found to influence splicing decisions (see Introduction), we investigated its role in exon 6 splicing. In this chapter, I describe experiments that used siRNA-mediated knockdown of CPSF1. CPSF1 depletion resulted in an increase in exon 6 inclusion, consistent with the mutations introduced to the consensus CPSF1 binding site AAUAAA (as described in Chapter 2). Additionally, CPSF1 depletion had no effect on splicing of minigenes where the AAUAAA site was mutated, suggesting that CPSF1 binding to this site is needed to mediate effect on exon 6 inclusion. I also present RT-PCR and 3’ RACE experiments that aimed at identifying potential polyadenylated isoforms produced from the AAUAAA site in intron 6 and propose that this site is used to generate novel polyadenylated IL7R isoforms at low frequency. These data suggest that together with alternative splicing, alternative intronic polyadenylation may serve as an important regulator of IL7R expression.

3.2 Materials and Methods

In vitro transcription and in vitro splicing

DNA templates for in vitro transcription reactions were prepared by restriction endonuclease digestion as follows: “T”, “C” and substitution minigenes were digested with BbsI (NEB); DNA templates containing tobramycin aptamer were digested with DraIII. Digested templates were purified with the QIAquick PCR Purification Kit or the QIAquick Gel Extraction Kit (both from Qiagen). For small-scale in vitro transcriptions, 2 µg of digested DNA plasmid was mixed with 0.5 mM each of ATP, CTP, GTP and
12.5 nM UTP, 1X Transcription Buffer (40 mM Tris-HCl, pH 7.5, 6 mM MgCl$_2$, 2 mM spermidine-HCl, 5 mM NaCl), 20-40 units RNase Out (Invitrogen) and 40 units T7 RNA Polymerase Plus (Ambion/Life Technologies). [$\alpha$-32P] UTP (3000 Ci/m mole, 10 mCi/ml, PerkinElmer) was added to the final concentration of 2 nCi/µl). Reactions were incubated at 37°C for 90 minutes, and RNA was extracted with Phenol:Chloroform:Isoamyl Alcohol (25:24:1, Invitrogen) and precipitated with 0.4 M NH$_4$OAc and ethanol. One µl of each transcription reaction was taken out before and after phenol/chloroform extraction and used for scintillation counting on a Beckman LS-6500 liquid scintillation counter. Large-scale in vitro transcriptions (for the RNA templates used in tobramycin RNA affinity chromatography) were performed using Ambion’s MEGAScript T7 High Yield Transcription Kit and high-concentration T7 RNA Polymerase Plus (200 u/µl, Ambion/Life Technologies) following the manufacturer’s protocol. Reactions were incubated at 37°C for 6 hours, and RNA was extracted with RNeasy Mini Kit (Qiagen). In vitro transcripts that were uniformly labeled were run on denaturing 15% polyacrylamide gels (19:1) with 7 M Urea, to check for transcript integrity. Non-radioactively labeled transcripts (prepared by using the MEGAScript kit) were resolved on denaturing agarose gels to check for transcript integrity.

In vitro splicing reactions contained 100,000 counts per minute (cpm) of uniformly labeled RNA template, 1 mM ATP, 5 mM phosphocreatine, 64 mM KCl (total), 2 mM MgCl$_2$, 40 u RNase Out (Invitrogen) and 32% HeLa nuclear extract (v/v). HeLa nuclear extract was prepared according to Dignam et al. 1983. Reactions were incubated at 30°C for 90 minutes and treated with Proteinase K (Ambion/Life
Technologies) in 10 mM Tris-HCl, pH 7.5, 1% SDS, 0.15 mM NaCl, 10 mM EDTA, 25 ng/µl tRNA (Sigma) and 50 µg Proteinase K (Ambion/Life Technologies) for 15 minutes at 37°C, after which RNA was extracted with phenol:chloroform:isoamyl alcohol and precipitated with NH₄OAc and ethanol as described above. RNA was air-dried and resuspended in 10 µl of water and treated with 2 units of TURBO DNase (Ambion/Life Technologies) according to the manufacturer’s protocol. DNase reactions were incubated at 37°C for 30 minutes and extracted with phenol:chloroform:isoamyl alcohol and precipitated with NH₄OAc and ethanol as described above. RNA was resuspended in 10 µl of water, and all 10 µl were used in an RT-PCR reaction performed in the same manner as the RT-PCR for total RNA extracted from the transfected cell lines (see above). PCR primers were: forward primer, U, and reverse primer, D, which annealed to the sequences of the adenovirus upstream (U) and downstream (D) exons present in the pI-12 vector (for primer sequences, see Table 3).

**Tobramycin RNA affinity chromatography and mass spectrometry**

Tobramycin RNA affinity chromatography was performed according to Hartmuth et al., 2004, and personal communication of K. Hartmuth. A 2 ml packed bead volume of NHS-activated sepharose was washed 4 times in ice-cold 1 mM HCl and mixed with 1 ml 5 mM tobramycin in the coupling buffer (either 0.2 M NaHCO₃, 0.5 M NaCl, (pH 8.3) or 75 mM HEPES-KOH (pH 7.9) and 90% dimethylformamide) and incubated overnight with head-over-tail rotation at 4°C. The next day, beads were spun and blocked in either 0.2 M NaHCO₃, 0.1 M NaCl, 1 M ethanolamine (pH 8.0) or 0.2 M Tris-HCl, pH 8.4 for 2 hours at room temperature. The matrix was then washed two-three times with 9 ml of ice-
cold PBS (pH 8.0) and then 2 times with ice-cold PBS/NaN₃ and resuspended in 2 ml PBS/NaN₃. Matrices were stored in PBS/NaN₃ at 4°C for up to a month.

For chromatography experiments, 15-60 µl (packed bead volume) of matrix aliquots were blocked overnight in 20 mM Tris-HCl, pH 8.1, 1 mM CaCl₂, 1 mM MgCl₂, 0.2 mM DTT (together referred to as 1X binding buffer, or 1X BP), 300 mM KCl, 0.1 mg/ml tRNA, 0.5 mg/ml BSA and 0.01% NP-40 at 4°C. The next day, *in vitro* synthesized RNA templates were mixed with 1X BP, 145 mM KCl and 0.1 mg/ml tRNA, added to the tobramycin sepharose matrix, and incubated for 1 hour at 4°C with head-over-tail rotation. At the same time, appropriate amounts of HeLa nuclear extract were pre-cleared with an aliquot of the tobramycin-containing matrix. RNA immobilized on the matrix was washed 3 times with 1 ml of the wash buffer (1X BP, 145 mM KCl, 0.1 % NP-40), 20 mM Tris-HCl, pH 8.1, 1 mM CaCl₂, 1 mM MgCl₂, 0.2 mM DTT, 145 mM KCl, and conventional *in vitro* splicing reactions were assembled in the presence of 1 mM ATP, 5 mM phosphocreatine, 64 mM KCl (total), 2 mM MgCl₂ and 32% pre-cleared HeLa nuclear extract (v/v). Splicing mixes were added to the immobilized RNA, and reactions were incubated at 30°C for 5 minutes with head-over-tail rotation. Following incubation, reactions were immediately put on ice and then centrifuged at 250xg at 4°C to collect supernatant fraction, followed by three washes with the protein wash buffer (1X BP, 150 mM KCl, 0.1% NP-40) and three washes with either PBS or elution buffer without tobramycin (see below). All spins were done at 4°C. RNA-protein complexes were eluted with elution buffer (1X BP, 5 mM tobramycin, 145 mM KCl, 2 mM MgCl₂) for 20 minutes at 30°C. Supernatant, wash, and elution fractions were collected and used.
to precipitate protein as follows: an equal volume of phenol:chloroform:isoamyl alcohol was added, and, upon phase separation, the lower organic phase was transferred to a fresh tube. Four volumes of ice-cold acetone were added, and precipitation was allowed to occur overnight at -20°C. The next day, samples were spun at maximum speed at 4°C, washed two times in 4:1 acetone:water, air-dried for 15 minutes, resuspended in the SDS-gel loading buffer (100 mM Tris-HCl, pH 6.8, 4% SDS, 0.02 % (w/v) bromophenol blue, 20% (v/v) glycerol, and 200 mM DTT, added fresh before each use) and either loaded onto NuPAGE protein gels (Invitrogen) for silver staining or Western blotting analysis or stored at - 20°C. For mass spectrometry experiments, 1X protease inhibitors (Roche) were added to the elution buffer, and the eluates were stored at 4°C overnight until they could be submitted for the mass spectrometry analysis the following day.

Mass spectrometry analysis of the chromatography eluates was performed by the Duke University Proteomics Core Facility. Quantification of the label-free mass spectrometry experiments was performed with Rosetta Elucidator® software by Dr. Will Thompson of the Duke University Proteomics Core Facility.

**SiRNA duplexes and CPSF1 knockdown**

SiRNA duplexes were synthesized by Qiagen in a FlexiTube format and resuspended in water to a final concentration of 10 µM upon arrival. CPSF1 siRNA sequences were as follows: Hs_CPSF1_4, 5’-CCAGATGATCAGCGTCAAGAA-3’; Hs_CPSF1_5, 5’-AGGGCGGATCTTGATCATGGA-3’; Hs_CPSF1_6, 5’-CACGTGGAGTCTAAGGTGTAT-3’; and Hs_CPSF1_7, 5’-CCGAGCGTTCCACTTTGACAA-3’. AllStar Negative Control siRNA (Qiagen) was
used as a non-silencing siRNA control in all knockdown experiments. Transfections of siRNAs into HeLa cells were performed with Lipofectamine™ RNAiMax (Invitrogen) according to the manufacturer’s instructions. Briefly, 6 pmol of siRNA duplexes were diluted in 100 µl Opti-MEM I Medium (Invitrogen/Gibco) without serum in each well of a 24-well plate. Next, 0.8 µl of Lipofectamine™ RNAiMax was added to each well. SiRNA/transfection reagent mixtures were mixed gently and incubated for 20 minutes at room temperature. HeLa cells were diluted to obtain 3x10⁴ cells per 500 µl medium and 500 µl cells was added to each well to obtain a final siRNA concentration of 10 nM. At 48 hours post-transfection, cells were transfected with minigenes as described in Chapter 2. At 48 hours post-minigene transfection, cells were harvested with TRIZOL Reagent (Invitrogen) for RNA analysis or with 1X RIPA buffer (see below) for protein analysis. Transfection with all 4 CPSF1 siRNAs resulted in comparable decrease in CPSF1 levels as estimated by Western blotting (Figure 3.4B). However, transfections of IL7R minigenes into HeLa cells in which CPSF1 was depleted with siRNA Hs_CPSF1_5, did not have any effect on exon 6 inclusion and were excluded from analysis.

**Protein extraction, antibodies and Western blotting**

Total protein was harvested by scraping cells in 1X RIPA buffer (150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris-HCl, pH 7.5, 1X Protease Inhibitors (Roche)) and subjecting cells to freeze/thaw lysis (freezing at – 80°C followed by thawing at 37°C). Protein lysates were loaded directly onto NuPAGE® 4-12% Bis-Tris pre-cast gels (Life Technologies), transferred onto Protran nitrocellulose membranes (Whatman) and Western blot assays were performed according to standard
techniques. Anti-CPSF1 rabbit polyclonal antiserum was a generous gift from Dr. J. Wilusz, Colorado State University. Anti-actin mouse monoclonal antibody was from Santa Cruz Biotechnology.

3’RACE

3’RACE was performed according to conventional protocols. Briefly, reverse transcription was performed with total DNase-treated RNA from Jurkat cells, reverse Oligo(dT)-Anchor primer and M-MLV RT. A total of 1 µl of the RT reaction was used in the first PCR with the forward primer complementary to the IL7R exon 3 and PCR-Anchor as the reverse primer. Nested PCRs were performed with the following primer pairs: exon 4 forward and intron 6 reverse; exon 5 forward and intron 6 reverse; exon 3 forward and intron 5 reverse and exon 3 forward and intron 4 reverse (for primer sequences, see Table 3) under the following cycling conditions: initial denaturation for 5 minutes at 94°C, 30 cycles of denaturation at 95°C for 30 seconds, annealing at 51.5°C and extension at 70°C for 1 minute. Final extension was done for 7 minutes at 70°C. PCR products were resolved on 5% non-denaturing polyacrylamide/TBE gels. Gels were run at 120 V for approximately 3 hours, dried using a gel dryer and exposed to Molecular Dynamics Phosphorimager screens.
3.3 Results

3.3.1 Exon 6 splicing regulation is recapitulated in HeLa cell nuclear extracts *in vitro*

We decided to purify proteins that bound exon 6 and adjacent intronic regions from HeLa cell nuclear extracts, which were readily available in the laboratory. We first tested whether HeLa cell nuclear extracts contain the protein factors that are necessary to recapitulate exon 6 splicing regulation *in vitro*. To this end, we used RNAs that previously showed either significant increase in exon 6 inclusion (sub(20-24)) or a significant decrease in inclusion (sub(36-40)) in transfection experiments. Minigene templates were transcribed *in vitro* and transcripts were used in *in vitro* splicing reactions as outlined in Materials and Methods. The “C” RNA and the 3’ or 5’ splice site mutants served as positive and negative controls, respectively. As shown in Figure 3.1, although the overall efficiency of splicing was lower than in the cell culture system (a little over 10% of transcripts included exon 6 in the *in vitro* splicing reactions when the “C” control RNA was analyzed, compared to over 50% in the cell culture), the pattern of exon 6 splicing regulation *in vitro* reproduced the pattern observed in transfected cell lines. Specifically, *in vitro* splicing of the RNA transcripts with ESS1 disrupted, sub(20-24), resulted in a significant increase in exon 6 inclusion, while *in vitro* splicing of the RNA transcripts with ESE2 disrupted, sub(36-40), resulted in a significant decrease in inclusion (compare Figs. 1B to 3A). Importantly, the effect that the MS-associated “C” allele has on exon 6 inclusion as compared to the “T” allele, was reproduced in the *in vitro* system as well (Fig. 3A). We thus concluded that HeLa cell nuclear extracts contain
all the necessary proteins to recapitulate exon 6 splicing regulation in vitro and used these extracts in further chromatography experiments.

3.3.2 RNA affinity chromatography identifies potential protein regulators of IL7R exon 6 splicing

To identify trans-acting protein factors that bind exon 6 and/or adjacent intronic regions we employed tobramycin RNA affinity chromatography (Hartmuth et al. 2004; Ward et al. 2011, and also Hartmuth K., personal communication). Two sets of RNA templates were used for chromatography experiments (Figure 3.2). The first set consisted of two RNAs that contained full-length exon 6 flanked by 48 and 56 nucleotides from introns 5 and 6, respectively, one of which contained the MS-associated “C” allele at the SNP position in the exon, while the other minigene contained the five-nucleotide substitution sub(36-40). A second set of RNAs contained only the first 40 nucleotides of exon 6, again, either in the context of the MS-associated “C” allele or substitution sub(36-40). IL7R sequences in all templates were followed by streptavidin/tobramycin aptamer sequences described by Ward et al. (Ward et al. 2011). A “no RNA” control was routinely used in all chromatography experiments to control for the specificity of protein binding. Each chromatography experiment was performed in duplicate. RNA-protein complexes were eluted with excess tobramycin and duplicate elutions submitted to Duke University Proteomics Core Facility for analysis.

Mass spectrometry analysis identified a set of proteins (43 in the experiment with the “full-length” templates and 67 in the experiment with templates that contained just the first 40 nucleotides of exon 6) that were specific for the experimental templates (not
found in the “no RNA” control), the majority of which were known RNA-binding proteins (Tables 1 and 2). Rosetta Elucidator software was used to obtain quantitative data on protein factors that differentially bound the “C” vs sub(36-40) RNA templates. Several proteins were found to bind to the two RNAs differentially (Table 1), with some (Quaking, PTB) enriched on the sub(36-40) template, while others (HuR, MIF, MAP4, SF1) were enriched on the “C” RNA template.

Interestingly, among proteins that bound template RNA specifically yet didn’t show differential binding between the “C” and “sub(36-40)” templates, CPSF1 along with other components of the mammalian cleavage and polyadenylation machinery, were identified specifically with the full-length RNA templates and not with the templates that contained just the first 40 nt of exon 6 (Table 2). Western blot analysis on chromatography elutions confirmed binding of CPSF1, the major component of the mammalian cleavage and polyadenylation complex, to the full-length template and absence (or minimal) binding to the template that contained the first 40 nt of exon 6 (Figure 3.3).

3.3.3 CPSF1 regulates exon 6 splicing by promoting skipping of exon 6

Given that CPSF components were shown to bind in the vicinity of exon 6 and that CPSF-binding site AAUAAA was shown to downregulate exon 6 inclusion (Chapter 2, Figure 2.7) we sought to test whether CPSF components were involved in the splicing of this exon. To this end we depleted CPSF1 in HeLa cells using siRNAs and transfected minigenes to determine the effect of CPSF1 depletion on inclusion of exon 6. Inclusion
of exon 6 was increased over 2-fold upon depletion of CPSF1 with at least three different siRNAs, if the RNAs contained the AAUAAA site, while inclusion of exon 6 was not affected by CPSF knockdown in RNAs lacking AAUAAA (Figure 3.4). We thus concluded that CPSF1 downregulates IL7R exon 6 splicing and is likely to do this via the AAUAAA 16 nt downstream of the exon.

3.3.4 Analysis of potential intronic polyadenylated isoforms in IL7R

Two different and not mutually exclusive mechanisms could explain the effect of CPSF1 and the AAUAAA site on exon 6 inclusion: 1) binding of the cleavage and polyadenylation machinery results in proper cleavage and polyadenylation of the intronic mRNA, which is in line with recent reports (Kaida et al. 2010, Vorlová et al. 2011) or 2) assembly of cleavage and polyadenylation complexes on the consensus poly(A) sites in the absence of cleavage and polyadenylation interferes with spliceosome binding to the adjacent 5’ splice site. We sought to determine if the AAUAAA signal in intron 6 of IL7R is used to produce a new polyadenylated mRNA isoform.

To test if the consensus poly(A) site in intron 6 is used to cleave and polyadenylate IL7R pre-mRNA, we performed RT-PCR and 3’ RACE experiments on total RNA from Jurkat cells that naturally express IL7R (Figures 3.5 and 3.6). RT-PCR experiments with a specific PCR reverse primer that encompasses the AAUAAA site in intron 6 of IL7R revealed specific products when primers in the upstream IL7R exons 3, 4 and 5 were used (Figure 3.5). Several forward primers were used to ensure that the observed products are derived from processed mRNAs. Moreover, to ensure that the
observed products are not specific for *IL7R* intron 6 and can be observed with other intronic polyadenylated RNAs, we used a reverse primer that encompasses an intronic poly(A) site in intron 4 of *IL7R* and, as expected, detected a specific PCR product. PCRs with reverse primer in intron 5 of *IL7R*, where no intronic poly(A) site is present, were used as negative controls and, as expected, failed to produce any PCR products. These results suggested that mRNA species that are spliced upstream of exon 6 and at the same time contain at least the first 16 nt of *IL7R* intron 6, are present in Jurkat cells. However, these results do not differentiate between polyadenylated and non-polyadenylated RNA, and a possibility remained that the observed PCR products were derived from RNAs that have not yet been spliced downstream from exon 6.

To investigate polyadenylated species of the *IL7R* mRNA, we performed 3’RACE experiments with total RNA from Jurkat cells (Figure 3.6). Reverse transcription reactions were performed with a reverse oligo(dT)-anchor primer followed by a PCR with a forward primer in exon 3 of *IL7R* and the anchor primer as the reverse primer as outlined in Materials and Methods. Since we were not able to detect any specific products in the first PCR reaction, it was followed by a nested PCR with reverse primers in introns 6 and 4 (both introns contain intronic poly(A) signals) and 5 (no intronic poly(A) site), respectively, to enhance sensitivity. Similar experiments with nested PCRs were performed in one of the original studies describing the effect of IPA (Kaida et al., 2010).

Consistent with our RT-PCR results, nested PCR identified a product only when a poly(A) site was present within an intron, and no product was identified in a control PCR with a reverse primer in intron 5 (Figure 3.6). This strongly suggested that the intronic
poly(A) site within intron 6 of *IL7R* is used to produce an alternative polyadenylated isoform.

### 3.4 Discussion

Tobramycin RNA affinity chromatography is a powerful method to identify protein components of RNPs that combines high specificity of the tobramycin-tobramycin aptamer interaction and gentle method of elution that provides RNPs in their native state (Hartmuth et al., 2004). We have used this method to identify protein factor that bind *IL7R* exon 6 and regulate its inclusion.

Forty-eight proteins were identified in the chromatography/mass spectrometry experiments as specifically bound to the RNA templates that contained full-length exon 6 and short sequences from surrounding introns 5 and 6 (as opposed to the “no RNA” control), and 56 proteins were identified in the experiments that used RNA templates that contained just the first 40 nt of exon 6. Quantification of the results showed that while the majority of the proteins bound to both templates equally, several proteins exhibited differential binding (Table 1). Importantly, several proteins were found to be enriched on the “C” template (HuR, MIF, MAP4, SF1). It would be important to validate these proteins as factors that bind exon 6 and regulate its splicing, by siRNA-mediated approaches. Interestingly, a literature search has found that one of these factors, MIF, is associated with MS (Powell et al., 2005; Akcali et al., 2010). A connection between *IL7R* and/or its splicing and MIF remains to be established.
We were puzzled by the presence of CPSF1 as well as other components of the mammalian cleavage and polyadenylation machinery (namely, CPSF 2, hFip1 and Symplekin), in the tobramycin chromatography elution fractions. We validated the interaction between CPSF1 and its consensus binding site AAUAAA in intron 6 of \textit{IL7R} (Figure 3.3) and showed that both mutations to the AAUAAA site in intron 6 and knockdown of CPSF1 resulted in de-repression of exon 6 inclusion (Figures 2.7 and 3.4). Our data also suggested that the AAUAAA site in intron 6 may be used to produce polyadenylated \textit{IL7R} mRNA at low frequencies (Figures 3.5 and 3.6). It remains to be shown that 3’ RACE products observed are indeed derived from polyadenylated RNA, by using PCR-Anchor (see Materials and Methods) as a reverse primer in the PCR reactions and direct sequencing of the resulting products (experiments in progress).

Splicing and polyadenylation are intricately interconnected and influence one another. CPSF complex can both interact with components of the spliceosome (Kyburz et al., 2006; Lutz et al., 1996) and has been found to be required for splicing of single introns \textit{in vivo} (Li et al., 2001). Several splicing factors have been found to influence polyadenylation (for details, see Introduction). It was thus very plausible that by binding to its consensus AAUAAA site in intron 6 of \textit{IL7R}, CPSF1 was influencing splicing of exon 6.

Two potential mechanisms may explain the effect that CPSF1 has on exon 6 splicing. First, usage of the \textit{IL7R} intron 6 poly(A) site might be an example of an alternative intronic polyadenylation event that competes with splicing and is mediated by U1 snRNP in a splicing-independent manner (Kaida et al. 2010; Vorlová et al. 2011). Alternatively, the assembly of the CPSF complex on the intronic poly(A) site in the
absence of cleavage and polyadenylation may interfere with binding of U1 snRNP to the adjacent 5' splice site. The two mechanisms are not mutually exclusive. Interestingly, it has been established that the frequency of intronic polyadenylation events increases in the presence of weak upstream 5' splice site (Tian et al., 2005). Since we determined that exon 6 5' splice site is weak (Figure 2.8), it is possible that reduced binding of U1 snRNP to this site allows for the assembly of the CPSF complex on the intronic AAUAAA site.

Whatever the mechanism, assembly of the polyadenylation complex on the intronic poly(A) site in intron 6 is expected to contribute to the poor recognition of this exon by the splicing machinery. At the level of RNA, this would result in less full-length IL7R and, subsequently, less functional IL7R protein. If the poly(A) site in intron 6 is used to cleave and polyadenylate mRNA, the resulting mRNA can be translated and is predicted to produce a new transmembrane isoform with a truncated cytoplasmic domain. While still retaining its ligand-binding abilities, this isoform would not be able to initiate signaling and will thus have dominant-negative functions. It will be important to determine if this isoform exists at the protein level, and if its production is somehow dependent on the presence of the MS-associated SNP.

Moreover, while in general cleavage and polyadenylation factors are expressed constitutively (Proudfoot 2011), some data suggest that several cleavage and polyadenylation factors may be expressed at different levels depending on the developmental stage of the cell or its physiological state. For example, cleavage stimulation factor CstF-64 is expressed at limiting levels in pre-B cells, compared to mature B cells. This results in splicing of a specific hydrophobic C-terminus of the immunoglobulin antibody heavy chain pre-mRNA and production of a membrane-bound
isoform. In mature B cells, where levels of CstF-64 are elevated, an intronic polyadenylation signal is recognized instead, and a secreted protein isoform is produced (Takagaki et al. 1996; Takagaki and Manley 1998). Differential expression of the same cleavage stimulation factor CstF-64 regulates expression of the transcription factor NF-ATc in naïve versus effector T cells (Chuvpilo et al., 1999). Additionally, expression of a different cleavage and stimulation factor, CstF-77, has been found to be up-regulated in mouse lymphomas (Singh et al. 2009). Recent data also suggest that alternative polyadenylation plays a crucial role during global events such as cell proliferation, differentiation and development (as reviewed in Moreira and Lutz 2011), including T cell proliferation, during which polyadenylation at the alternative poly(A) sites results in the production of shorter mRNAs with different 3’ UTRs (Sandberg et al. 2008). It would be important to determine if any of the components of the cleavage and polyadenylation machinery are differentially expressed during stages of T cell development and/or T cell proliferation and differentiation, and whether differential expression of these factors correlates with changes in the production of the intronic polyadenylated isoform(s) of \textit{IL7R}. An intriguing (but completely unexplored) possibility would be that changes in the expression of either specific splicing factors that regulate \textit{IL7R} exon 6 inclusion or cleavage and polyadenylation factors that regulate intronic polyadenylation within intron 6 of \textit{IL7R} contribute to MS (or other pathologies) by either stimulating exon 6 skipping (and, thus, soluble \textit{IL7R} production) and/or by enhancing usage of the intronic poly(A) site (and, thus, potential expression of the truncated transmembrane isoform).
Figure 3.1: Splicing regulation of IL7R exon 6 is recapitulated in HeLa cell nuclear extracts in vitro.

Minigenes were transcribed in vitro and used in in vitro splicing reactions with HeLa cell nuclear extracts. 3' and 5" ss mut minigenes served as controls. Sub(36-40) resulted in a significant decrease in exon 6 inclusion in the cell culture system. Sub(20-24) resulted in an increase in inclusion. Both results were recapitulated in the nuclear extracts. Note the recapitulation of the effect that the MS-associated allele has on splicing (compare minigenes “C” and “T”). Each bar represents one in vitro splicing reactions. This experiment was repeated several times with identical results.
Figure 3.2: Templates for tobramycin RNA affinity chromatography.

*Top*, first set of templates contained full-length *IL7R* exon 6 (94 nt), 48 and 56 nt of introns 5 and 6, respectively, followed by tobramycin/streptavidin aptamer sequences. *Bottom*, second set of templates contained first 40 nt of *IL7R* exon 6 followed by the aptamer sequences. Both constructs were made in the context of either “C” allele or sub(36-40) (in the background of “C”). Templates were transcribed *in vitro* from the T7 promoter.
Table 1: Proteins identified by quantitative mass spectrometry to bind differentially to "C" and "sub(36-40)" RNAs

<table>
<thead>
<tr>
<th>Protein</th>
<th>Protein name</th>
<th>Fold change: C vs sub(36-40)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Identified in the experiment with the “full-length” templates</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QKI</td>
<td>Quaking</td>
<td>-2.9</td>
</tr>
<tr>
<td>PTBP1 (PTB)</td>
<td>Polypyrimidine tract binding protein 1</td>
<td>-2.2</td>
</tr>
<tr>
<td>ELAVL1 (HuR)</td>
<td>Embryonic lethal-abnormal vision-like protein 1</td>
<td>2.2</td>
</tr>
<tr>
<td><strong>Identified in the experiment with the templates that contained only the first 40 nt of exon 6</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIF</td>
<td>Macrophage inhibitory factor</td>
<td>2.0</td>
</tr>
<tr>
<td>MAP4</td>
<td>Microtubule-associated protein 4</td>
<td>1.8</td>
</tr>
<tr>
<td>SF1</td>
<td>Splicing factor 1</td>
<td>5.0</td>
</tr>
</tbody>
</table>
Table 2: Protein factors that do not show differential binding to "C" vs "sub(36-40)" templates as identified by mass spectrometry.

Components of the cleavage and polyadenylation machinery are highlighted.

<table>
<thead>
<tr>
<th>Protein name</th>
<th>Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identified in the experiment with “full-length” RNA templates</td>
<td></td>
</tr>
<tr>
<td>hnRNP Q</td>
<td>Heterogeneous nuclear ribonucleoprotein Q</td>
</tr>
<tr>
<td>DHX9</td>
<td>ATP-dependent RNA helicase A</td>
</tr>
<tr>
<td>ILF3</td>
<td>Interleukin enhancer-binding factor 3</td>
</tr>
<tr>
<td>ILF2</td>
<td>Interleukin enhancer-binding factor 2</td>
</tr>
<tr>
<td>hnRNP U</td>
<td>Heterogeneous nuclear ribonucleoprotein U</td>
</tr>
<tr>
<td>hnRNP R</td>
<td>Heterogeneous nuclear ribonucleoprotein R</td>
</tr>
<tr>
<td>hnRNP A/B</td>
<td>Heterogeneous nuclear ribonucleoprotein A/B</td>
</tr>
<tr>
<td>hnRNP A3</td>
<td>Heterogeneous nuclear ribonucleoprotein A3</td>
</tr>
<tr>
<td>TOP1</td>
<td>DNA Topoisomerase 1</td>
</tr>
<tr>
<td>hnRNP D-like</td>
<td>Heterogeneous nuclear ribonucleoprotein D-like</td>
</tr>
<tr>
<td>U2AF 65</td>
<td>Splicing factor U2AF 65 kDa subunit</td>
</tr>
<tr>
<td>E2AK2</td>
<td>Interferon-induced, dsRNA-activated protein kinase</td>
</tr>
<tr>
<td>HNRL2</td>
<td>Heterogeneous nuclear ribonucleoprotein U-like protein 2</td>
</tr>
<tr>
<td>CPSF1</td>
<td>Cleavage and polyadenylation stimulatory factor subunit 1</td>
</tr>
<tr>
<td>Sm D2</td>
<td>Small nuclear ribonucleoprotein Sm D2</td>
</tr>
<tr>
<td>DDX3X</td>
<td>ATP-dependent RNA helicase DDX3X</td>
</tr>
<tr>
<td>hnRNP L</td>
<td>Heterogeneous nuclear ribonucleoprotein L</td>
</tr>
<tr>
<td>Tra2b</td>
<td>Transformer-2 protein homolog beta</td>
</tr>
<tr>
<td>SF3B</td>
<td>Splicing factor 3B subunit 2</td>
</tr>
<tr>
<td>DDX5</td>
<td>Probable ATP-dependent RNA helicase</td>
</tr>
<tr>
<td>PABP2</td>
<td>Polyadenylate-binding protein 2</td>
</tr>
<tr>
<td>DDX21</td>
<td>Nucleolar RNA helicase 2</td>
</tr>
<tr>
<td>hnRNP H3</td>
<td>Heterogeneous nuclear ribonucleoprotein H3</td>
</tr>
<tr>
<td>L22</td>
<td>60S ribosomal protein L22</td>
</tr>
<tr>
<td>U2AF 35</td>
<td>Splicing factor U2AF 35 kDa subunit</td>
</tr>
<tr>
<td>hnRNP A0</td>
<td>Heterogeneous nuclear ribonucleoprotein A0</td>
</tr>
<tr>
<td>La</td>
<td>Lupus La protein</td>
</tr>
<tr>
<td>RL1D1</td>
<td>Ribosomal L1-domain containing protein 1</td>
</tr>
<tr>
<td>hFIP1</td>
<td>Pre-mRNA 3' end processing factor</td>
</tr>
<tr>
<td>TOE1</td>
<td>Target of EGR1 protein 1</td>
</tr>
<tr>
<td>SF3A</td>
<td>Splicing factor 3A subunit 1</td>
</tr>
<tr>
<td>PURB</td>
<td>Transcriptional activator protein Pur-beta</td>
</tr>
<tr>
<td>DHX36</td>
<td>Probable ATP-dependent RNA helicase DHX36</td>
</tr>
<tr>
<td>HNRL1</td>
<td>Heterogeneous nuclear ribonucleoprotein U-like protein 1</td>
</tr>
<tr>
<td>NEP1</td>
<td>Probable ribosome biogenesis protein NEP1</td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>Description</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>SRSF9</td>
<td>Ser/Arg-rich splicing factor 9</td>
</tr>
<tr>
<td>SF1</td>
<td>Splicing Factor 1</td>
</tr>
<tr>
<td>SF3B1</td>
<td>Splicing factor 3B subunit 1</td>
</tr>
<tr>
<td>G3BP1</td>
<td>Ras GTPase-activating protein-binding protein 1</td>
</tr>
<tr>
<td>NUCL</td>
<td>Nucleolin</td>
</tr>
<tr>
<td>hnRNPD0</td>
<td>Heterogeneous nuclear ribonucleoprotein D0</td>
</tr>
<tr>
<td>hnRNPQ</td>
<td>Heterogeneous nuclear ribonucleoprotein Q</td>
</tr>
<tr>
<td>YBOX1</td>
<td>Nuclease-sensitive element-binding protein 1</td>
</tr>
<tr>
<td>hnRNPA/B</td>
<td>Heterogeneous nuclear ribonucleoprotein A/B</td>
</tr>
<tr>
<td>RU17</td>
<td>U1 snRNP 70 kDa</td>
</tr>
<tr>
<td>NONO</td>
<td>Non-POU domain-containing octamer-binding protein</td>
</tr>
<tr>
<td>hnRNPA3</td>
<td>Heterogeneous nuclear ribonucleoprotein A3</td>
</tr>
<tr>
<td>H1.2</td>
<td>Histone H1.2</td>
</tr>
<tr>
<td>PABP2</td>
<td>Polyadenylate-binding protein 2</td>
</tr>
<tr>
<td>SRP 14kDa protein</td>
<td>Signal recognition particle 14kDa protein</td>
</tr>
<tr>
<td>PABP1</td>
<td>Polyadenylate-binding protein 1</td>
</tr>
<tr>
<td>ILF2</td>
<td>Interleukin enhancer-binding factor 2</td>
</tr>
<tr>
<td>ILF3</td>
<td>Interleukin enhancer-binding factor 3</td>
</tr>
<tr>
<td>hnRNP D-like</td>
<td>Heterogeneous nuclear ribonucleoprotein D-like</td>
</tr>
<tr>
<td>La</td>
<td>Lupus La protine</td>
</tr>
<tr>
<td>Sm D1</td>
<td>Small nuclear ribonucleoprotein Sm D1</td>
</tr>
<tr>
<td>PURB</td>
<td>Transcriptional activator protein Pur-beta</td>
</tr>
<tr>
<td>Sm D3</td>
<td>Small nuclear ribonucleoprotein Sm D3</td>
</tr>
<tr>
<td>Sm D2</td>
<td>Small nuclear ribonucleoprotein Sm D2</td>
</tr>
<tr>
<td>FUS</td>
<td>RNA-binding protein FUS</td>
</tr>
<tr>
<td>RUXF</td>
<td>Small nuclear ribonucleoprotein F</td>
</tr>
<tr>
<td>SRP 9 kDa protein</td>
<td>Signal recognition particle 9 kDa protein</td>
</tr>
<tr>
<td>RSMB</td>
<td>snRNP-associated proteins B and B’ and N</td>
</tr>
<tr>
<td>THO4</td>
<td>THO complex subunit 4</td>
</tr>
<tr>
<td>GBLP</td>
<td>Guanine nucleotide-binding protein subunit beta-2-like-1</td>
</tr>
<tr>
<td>H1.4</td>
<td>Histone H1.4</td>
</tr>
<tr>
<td>NEP1</td>
<td>rRNA small subunit methyltransferase NEP1</td>
</tr>
<tr>
<td>hnRNP G</td>
<td>Heterogeneous nuclear ribonucleoprotein G</td>
</tr>
<tr>
<td>TOP1</td>
<td>DNA Topoisomerase 1</td>
</tr>
<tr>
<td>DBPA</td>
<td>DNA-binding protein A</td>
</tr>
<tr>
<td>U1 snRNP A</td>
<td>U1 small nuclear ribonucleoprotein A</td>
</tr>
<tr>
<td>MAP4</td>
<td>Microtubule-associated protein 4</td>
</tr>
<tr>
<td>SF1</td>
<td>Splicing factor 1</td>
</tr>
<tr>
<td>snRNP E</td>
<td>Heterogeneous nuclear ribonucleoprotein E</td>
</tr>
<tr>
<td>RL6</td>
<td>60S Ribosomal protein L6</td>
</tr>
<tr>
<td>1433E</td>
<td>14-3-3 protein epsilon</td>
</tr>
<tr>
<td>SCAB1</td>
<td>Scaffold attachment factor B1</td>
</tr>
</tbody>
</table>

**Identified in the experiment with templates that contained first 40 nt of exon 6**
<table>
<thead>
<tr>
<th>Protein</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RL18</td>
<td>60S ribosomal protein L18</td>
</tr>
<tr>
<td>RSSA</td>
<td>40S Ribosomal protein SA</td>
</tr>
<tr>
<td>snRNP G and G-like</td>
<td>Small nuclear ribonucleoprotein G and G-like</td>
</tr>
<tr>
<td>SRSF3</td>
<td>Ser/Arg-rich splicing factor 3</td>
</tr>
<tr>
<td>PURA</td>
<td>Transcriptional activator protein Pur-alpha</td>
</tr>
<tr>
<td>RL31</td>
<td>60S Ribosomal Protein L31</td>
</tr>
<tr>
<td>H2B</td>
<td>Histone H2B different types</td>
</tr>
<tr>
<td>DDX39B</td>
<td>Spliceosome RNA helicase DDX39B</td>
</tr>
<tr>
<td>RBM3</td>
<td>Putative RNA-binding protein 3</td>
</tr>
<tr>
<td>RSS30</td>
<td>40S Ribosomal protein S30</td>
</tr>
<tr>
<td>CIRBP</td>
<td>Cold-inducible RNA-binding protein</td>
</tr>
<tr>
<td>PABP4</td>
<td>Polyadenylate-binding protein 4</td>
</tr>
<tr>
<td>hnRNP A0</td>
<td>Heterogeneous nuclear ribonucleoprotein A0</td>
</tr>
<tr>
<td>RSL29</td>
<td>60S Ribosomal Protein L29</td>
</tr>
<tr>
<td>HNRL2</td>
<td>Heterogeneous nuclear ribonucleoprotein U-like protein 2</td>
</tr>
<tr>
<td>1433G</td>
<td>14-3-3 protein gamma</td>
</tr>
<tr>
<td>ZN787</td>
<td>Zinc finger protein 787</td>
</tr>
<tr>
<td>RLL8</td>
<td>60S Ribosomal Protein L8</td>
</tr>
<tr>
<td>H1x</td>
<td>Histone H1x</td>
</tr>
<tr>
<td>hnRNP C1/C2</td>
<td>Heterogeneous nuclear ribonucleoprotein C1/C2</td>
</tr>
<tr>
<td>RLL30</td>
<td>60S Ribosomal Protein L30</td>
</tr>
<tr>
<td>PTMA</td>
<td>Prothymosin alpha</td>
</tr>
<tr>
<td>HnRNP U</td>
<td>Heterogeneous nuclear ribonucleoprotein U</td>
</tr>
<tr>
<td>RLL15</td>
<td>60S Ribosomal Protein L15</td>
</tr>
<tr>
<td>hnRNP R</td>
<td>Heterogeneous nuclear ribonucleoprotein R</td>
</tr>
<tr>
<td>RLL7</td>
<td>60S Ribosomal Protein L7</td>
</tr>
<tr>
<td>RSS3</td>
<td>40S Ribosomal protein S3</td>
</tr>
<tr>
<td>RSS20</td>
<td>40S Ribosomal Protein S20</td>
</tr>
</tbody>
</table>
Figure 3.3: Western blot analysis of RNA affinity chromatography elution fractions. Elutions for one set of duplicates are shown for each experiment. The blots were probed with anti-CPSF1 antiserum. CPSF1 molecular weight is 160 kDa.
Figure 3.4: Effect of siRNA-mediated CPSF1 knockdown on exon 6 inclusion of minigenes sub(36-40) and sub(36-40)pAmut

CPSF1 was depleted by three different siRNAs. A non-silencing siRNA served as control. siRNAs were transfected into HeLa cells, followed by minigene transfection in 48 hours. Minigene sub(36-40) is described in Figures 1.1 and 1.2. Minigene sub(36-40)pAmut contained a substitution mutation to the polyadenylation site AAUAAA in intron 6 (for mutation sequence, see Figure 2.7A). Exon 6 inclusion was measured by RT-PCR, and fold change difference between control siRNA-treated and CPSF1 siRNA-treated cells was calculated for each minigene. Levels of CPSF1 knockdown are shown below the graph for siRNAs Hs_CPSF1_4 and Hs_CPSF1_6. Knockdown of CPSF1 by siRNAs Hs_CPSF1_5 and HS_CPSF1_7 resulted in similar levels of CPSF1 depletion (data not shown).
Figure 3.5: RT-PCR experiments to investigate mRNA isoforms of IL7R.
RT step of the RT-PCR reaction was performed with random hexamer primers, either in the presence or absence of the RT enzyme. PCR product sizes and primer pairs are indicated above the gel. PCR products are highlighted with arrows. Note the absence of the correct-size product when primers in exon 3 (Exon 3 For) and intron 5 (Intron 5 Rev, no poly(A) site present) were used.
Exon 4 For/Intron 6 Rev (355 bp)
Exon 5 For/Intron 6 Rev (252 bp)
Exon 3 For/Intron 5 Rev (518 bp)
Exon 3 For/Intron 4 Rev (349 bp)

RT + - + - + - + -

527 bp
404 bp
307 bp
242 bp
Figure 3.6: 3' RACE experiments to investigate intronic polyadenylated isoforms of
*IL7R* mRNA.

RT reactions were performed with Oligo(dT)-Anchor primer and either in the
presence or absence of the RT enzyme. Primers used for nested PCR are shown
above the gel. PCR product sizes are indicated. PCR products are highlighted with
arrows. Note the absence of the product when primers in exon 3 (Exon 3 For) and
intron 5 (Intron 5 Rev, no poly(A) site) were used.
Chapter 4: Concluding discussion

4.1 Posttranscriptional regulation of \textit{IL7R}

Interleukin 7 receptor plays a critical role in T cell development and maintenance, with one of its major functions being protection of T cells from apoptosis (Mazzucchelli and Durum 2007). Underscoring its importance in T cell biology, aberrant expression of \textit{IL7R} has been implicated in several human diseases such as severe combined immunodeficiency (SCID) (Puel et al. 1998; Kalman et al. 2004) and acute lymphoblastic leukemia (Korte et al., 2000; Zhang et al. 2012), while genetic variation in \textit{IL7R} can predispose to such diseases as sarcoidosis (Heron et al. 2009), primary biliary cirrhosis (Mells et al., 2011), ulcerative colitis (Yamazaki et al., 2003; Anderson et al, 2011) and others, suggesting a larger role for this gene in the pathogenesis of autoimmunity.

\textit{IL7R} is expressed exclusively on the cells of the lymphoid lineage, and its expression is tightly regulated based on the developmental stage of the T cell and availability of the ligand (Mazzucchelli and Durum 2007). While transcriptional regulation of \textit{IL7R} expression has been widely studied (Pleiman et al., 1991; Franchimont et al., 2002; Xue et al., 2002; Park et al., 2004), posttranscriptional regulation mechanisms of \textit{IL7R} have only recently been uncovered. Several splicing isoforms have been reported for \textit{IL7R} at the RNA level (Goodwin et al., 1991; Pleiman et al., 1991; Korte et al., 2000; Gregory et al., 2007; Rose et al., 2009), but little attempts to characterize these splicing events in detail have been made. Moreover, to the best of our knowledge, other mechanisms of posttranscriptional regulation of \textit{IL7R} expression, such
as regulation by alternative polyadenylation, have not been explored. Considering the proposed role of the soluble IL7R isoform in MS and several reports implying the importance of this isoform in infection with the human immunodeficiency virus (HIV) (Koesters et al., 2006; Rose et al., 2009; Crawley et al., 2010), it would be important to characterize this and other potential IL7R isoforms and their contribution to T cell biology and disease.

We investigated splicing of IL7R exon 6 in detail, both at the level of cis- and trans-acting regulatory elements. Our experiments suggested the presence of a novel polyadenylated isoform arising from potential cleavage and polyadenylation in intron 6 of IL7R. As shown in Figure 1.5, several isoforms of IL7R have been predicted to exist, mainly derived from different splicing events. However, to our best knowledge, no alternative polyadenylation has been reported for this gene or other interleukin receptors. Bioinformatic studies have shown that up to 20% of human genes contain polyadenylation signals in their introns (Tian et al., 2005). Sequence analysis shows that poly(A) signals are present in several of the IL7R introns, and in introns 3, 4 and 6 these signals are in direct vicinity of the 5′ splice sites and are followed by a cleavage signal dinucleotide CA. If expressed, the protein product from the IPA in intron 6 would result in a truncated transmembrane isoform that would retain its ligand-binding abilities but would be deficient in signaling. Interestingly, if intronic polyadenylation occurs in introns 3 and 4, and if those novel mRNAs are expressed, the resultant protein products will be soluble. In the context of the suggested role of the soluble IL7R in MS pathogenesis, it would be important to investigate these potential polyadenylation events and determine if corresponding proteins are indeed produced. Alternative
polyadenylation of *IL7R* may thus constitute yet another layer of regulation of this receptor that contributes to the repertoire of *IL7R* isoforms and acts together with splicing to regulate its expression.

4.2 Potential role of different isoforms of *IL7R* in multiple sclerosis

Clearly, the causality of *IL7Rα* in MS pathogenesis has not been formally proven, and the exact role of the increased production of *sIL7R* or the effect that potential alternatively polyadenylated *IL7R* mRNAs have in MS has not been unraveled. One can envision, however, a number of plausible explanations that could connect aberrant splicing of exon 6, increased *sIL7R* production and MS. As discussed in Introduction, developing T cells in the thymus as well as naïve and memory T cells in the periphery require signaling through *IL7R* for survival and proliferation. *IL7R* ligand, IL7, is expressed constitutively, however, its amount is thought to be very low and just sufficient to maintain a finite number of T cells (as reviewed in Mazzucchelli and Durum, 2007). Therefore, lower expression of the transmembrane *IL7R*, due to increased skipping of exon 6, or expression of truncated transmembrane isoforms due to intronic polyadenylation, could lower IL7 mediated signaling below a critical threshold and compromise survival and proliferation of T cells (in this delicately balanced system, even a small decrease in functional *rIL7R* would be sufficient to generate an effect). Additionally, sequestration of the ligand by the increased amounts of the *sIL7R* isoform or by transmembrane isoforms with truncated cytoplasmic tails that are unable to transmit
the signal, could deprive T cells of their survival and proliferative signals. It must be noted, however, that the connection between this antagonistic mode of action for \( sIL7R \) and the development of MS, as proposed here, remains unproven.

Developing a hypothesis in which aberrant signaling through \( rIL7R \), and, thus, compromised T cell survival and proliferation, contribute to the etiology of MS is challenging for two reasons: (1) other unknown environmental and genetic factors may contribute to the development of the disease, and (2) the possible commensurate decrease in T cell populations is contrary to a disease model in which there is acute T cell proliferation and inflammatory response that results in axon demyelination. Two studies that identified an increase in CD8+ T cells from PBMC in MS patients (Liu et al., 2007; Haegele et al., 2007) also identified a relative decrease in CD8+ T cells in the CNS, in line with the assumption that a decrease in T cell populations is associated with the disease (Haegele et al., 2007). It is also possible that lower expression of rIL7R by the CD8+ T cells in MS patients renders these cells more susceptible to viral attack. Indeed, various viruses have been implicated in the development of MS (Grigoriadis and Hadjigeorgiou 2006; Lunemann and Munz 2007; Kakalacheva and Lunemann 2011), and in chronic viral infections a phenomenon known as CD8+ T cell exhaustion, which is characterized by poor function and viability of memory T cells, as well as low levels of rIL7R expression by these cells, has been described (day et al., 2006; Mazzucchelli and Durum 2007; Mueller and Ahmed 2009). Thus, it is possible that in individuals who genetically express lower levels of rIL7R, CD8+ T cell exhaustion occurs in response to chronic viral infection. Viral infection can subsequently induce autoimmunity resulting in demyelination and axon damage (Mueller and Ahmed 2009).
4.3 Concluding remarks

We have investigated posttranscriptional regulation of the clinically important interleukin 7 receptor (specifically, its exon 6). Our experiments revealed multiple exonic and intronic cis-acting regulatory elements, identified trans-acting protein factors that bind exon 6 and suggested that intronic polyadenylation might contribute to the repertoire of IL7R isoforms. Several important questions arise from our research and represent avenues for future work: first, what is the mechanism of CPSF1 repression of exon 6 inclusion? If poly(A) site in intron 6 is indeed used for cleavage and polyadenylation, are these novel transcripts translated, and if so, can their presence be detected in living cells? Secondly, what is the abundance of these new isoforms? Additionally, 3’ RACE and RT-PCR experiments have suggested that intronic polyadenylation events in other IL7R introns (e.g., intron 4) may also occur in cells. Given the importance of intronic polyadenylation for other classes of transmembrane receptors, it would be important to determine the contribution of IPA to IL7R expression. It would also be interesting to extend these studies into other interleukin receptors to determine if IPA influences their expression.

Lastly, experiments to evaluate the presence and abundance of these novel IL7R isoform(s) and/or expression levels of different polyadenylation factors in MS patients are needed. These studies will not only contribute to our understanding of IL7R expression, but also shed light on the role of this receptor in T cell biology and disease progression.
## Appendix

### Table 3: Primers used in this study

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>T7</td>
<td>5’-TAATACGACTCTATAGGGAGA-3’</td>
</tr>
<tr>
<td>Sp6</td>
<td>5’-ATTATTAGGTGACAATAGGAA-3’</td>
</tr>
<tr>
<td>U</td>
<td>5’-GCTGCGATCCACTAGTAAACG-3’</td>
</tr>
<tr>
<td>D</td>
<td>5’-GTTCCGAGGATGCATAGAGA-3’</td>
</tr>
<tr>
<td>IL7R exon 3</td>
<td>5’-GTGGAGGTAAAGTGCGATTTTC-3’</td>
</tr>
<tr>
<td>IL7R exon 4</td>
<td>5’-TAAATGCGATGCTAGTACCAGC-3’</td>
</tr>
<tr>
<td>IL7R intron 6</td>
<td>5’-TTTATTAGTTAGGAGAGGTCACTTTTTTTCC-3’</td>
</tr>
<tr>
<td>IL7R exon 5</td>
<td>5’-CAGGAGCAATGAGATGATTAAAGA-3’</td>
</tr>
<tr>
<td>IL7R intron 5</td>
<td>5’-CAGGAGCAATGAGATGATTAAAGA-3’</td>
</tr>
<tr>
<td>IL7R intron 4</td>
<td>5’-CATCCATTTAGGTTTTTTTATGATTGAGGAC-3’</td>
</tr>
<tr>
<td>Intron 6 For</td>
<td>5’-CCTTCTTCTTATTTTGGAGGAGATTT-3’</td>
</tr>
<tr>
<td>Intron 6 Rev</td>
<td>5’-CATACTCAGACTCTCTAAGGATGAGGAC-3’</td>
</tr>
<tr>
<td>Oligo(dT)-</td>
<td>5’-CAGCCTGAGTCTCATGACACACGGGACACCGAAGCT-3’</td>
</tr>
<tr>
<td>Anchor</td>
<td>5’-CAGCCTGAGTCTCATGACACACGGGACACCGAAGCT-3’</td>
</tr>
<tr>
<td>Sub(1-4) For</td>
<td>5’-CCTTCTTCTCTTATTTTACCTAGGATGATGC-3’</td>
</tr>
<tr>
<td>Sub(1-4) Rev</td>
<td>5’-CCTTCTTCTCTTATTTTACCTAGGATGATGC-3’</td>
</tr>
<tr>
<td>Sub(5-9) For</td>
<td>5’-CCTTCTTCTTTTATTTTACCTAGGATGATGC-3’</td>
</tr>
<tr>
<td>Sub(5-9) Rev</td>
<td>5’-CCTTCTTCTTTTATTTTACCTAGGATGATGC-3’</td>
</tr>
<tr>
<td>Sub(10-14) For</td>
<td>5’-CCTTCTTCTTTTATTTTACCTAGGATGATGC-3’</td>
</tr>
<tr>
<td>Sub(10-14) Rev</td>
<td>5’-CCTTCTTCTTTTATTTTACCTAGGATGATGC-3’</td>
</tr>
<tr>
<td>Sub(15-19) For</td>
<td>5’-CCTTCTTCTTTTATTTTACCTAGGATGATGC-3’</td>
</tr>
<tr>
<td>Sub(15-19) Rev</td>
<td>5’-CCTTCTTCTTTTATTTTACCTAGGATGATGC-3’</td>
</tr>
<tr>
<td>Sub(20-24) For</td>
<td>5’-CCTTCTTCTTTTATTTTACCTAGGATGATGC-3’</td>
</tr>
<tr>
<td>Sub(20-24) Rev</td>
<td>5’-CCTTCTTCTTTTATTTTACCTAGGATGATGC-3’</td>
</tr>
<tr>
<td>Sub(25-30) For</td>
<td>5’-CCTTCTTCTTTTATTTTACCTAGGATGATGC-3’</td>
</tr>
<tr>
<td>Sub(31-35) Rev</td>
<td>5’-CCTTCTTCTTTTATTTTACCTAGGATGATGC-3’</td>
</tr>
<tr>
<td>Sub(36-40) For</td>
<td>5’-CCTTCTTCTTTTATTTTACCTAGGATGATGC-3’</td>
</tr>
<tr>
<td>Sub(36-40) Rev</td>
<td>5’-CCTTCTTCTTTTATTTTACCTAGGATGATGC-3’</td>
</tr>
<tr>
<td>Del(1-4) For</td>
<td>5’-CCTTCTTCTTTTATTTTACCTAGGATGATGC-3’</td>
</tr>
<tr>
<td>Del(1-4) Rev</td>
<td>5’-CCTTCTTCTTTTATTTTACCTAGGATGATGC-3’</td>
</tr>
<tr>
<td>Del(5-9) For</td>
<td>5’-CCTTCTTCTTTTATTTTACCTAGGATGATGC-3’</td>
</tr>
<tr>
<td>Del(5-9) Rev</td>
<td>5’-CCTTCTTCTTTTATTTTACCTAGGATGATGC-3’</td>
</tr>
<tr>
<td>Del(10-14) For</td>
<td>5’-CCTTCTTCTTTTATTTTACCTAGGATGATGC-3’</td>
</tr>
<tr>
<td>Del(10-14) Rev</td>
<td>5’-CCTTCTTCTTTTATTTTACCTAGGATGATGC-3’</td>
</tr>
<tr>
<td>Del(15-19) For</td>
<td>5’-GAGATGGATCCTACTAACCAC-3’</td>
</tr>
<tr>
<td>Del(15-19) Rev</td>
<td>5’-CTGATGGTTAGTAGGATCCACTC-3’</td>
</tr>
<tr>
<td>Del(20-24) For</td>
<td>5’-TGGATCCTATCTTTCCATCGCATTT-3’</td>
</tr>
<tr>
<td>Del(20-24) Rev</td>
<td>5’-CAAAAAATGCTGATGGAAGATAGGATCCA-3’</td>
</tr>
<tr>
<td>Del(25-30) For</td>
<td>5’-CTATCTTACTAACGCATTGTTTTC-3’</td>
</tr>
<tr>
<td>Del(25-30) Rev</td>
<td>5’-GAAAAAAACTCAAAATGCGTTAGTAAAGATAG-3’</td>
</tr>
<tr>
<td>Del(31-35) For</td>
<td>5’-TTTACTAACCACATCTTTGAGTTTTTTTC-3’</td>
</tr>
<tr>
<td>Del(31-35) Rev</td>
<td>5’-GAGAAAAACTCAATGTGGTTAGTAAG-3’</td>
</tr>
<tr>
<td>Del(36-40) For</td>
<td>5’-CTAACCATCAGCATTTTTTTCTGTC-3’</td>
</tr>
<tr>
<td>Del(36-40) Rev</td>
<td>5’-GACAGAGAAAAATGCTGATGGTTAG-3’</td>
</tr>
<tr>
<td>For(26-40) transv</td>
<td>5’-GGATCCTATCTTTACTAACAACGACTACGGGGTCTTTTTTCTCTGTC-3’</td>
</tr>
<tr>
<td>Rev(26-40) transv</td>
<td>5’-CAGAGCGACAGAGAAAAAGACCCCGTAGTCGTGTTAGTAAGATAGGATCC-3’</td>
</tr>
<tr>
<td>3’ ss mut For</td>
<td>5’-TATTCTTGTTTCTGGGAGATGGAT-3’</td>
</tr>
<tr>
<td>3’ ss mut Rev</td>
<td>5’-ATCCATCTCCCGACGAAAGCAAGAATA-3’</td>
</tr>
<tr>
<td>5’ ss mut For</td>
<td>5’-TTATGGAAAAAACCAACCTTTCTTCAAC-3’</td>
</tr>
<tr>
<td>5’ ss mut Rev</td>
<td>5’-GTTGAGAAAGGTGTTTTTTTCCATAAA-3’</td>
</tr>
<tr>
<td>5’ ss consensus For</td>
<td>5’-GTTATGGAAAAAACGGTGTTTCTTCAACTAATAAG-3’</td>
</tr>
<tr>
<td>5’ ss consensus Rev</td>
<td>5’-CTTTATTAGTTGAAAGAACTCACCTTTTCTCCATAAC-3’</td>
</tr>
</tbody>
</table>
References


Crawley AM, Faucher S, Angel JB. 2010. Soluble IL-7Rα (sCD127) inhibits IL-7 activity and is increased in HIV infection. J Immunol 184: 4679-4687.


Lutz CS, Murthy KGK, Schek N, O’Connor JP, Manley JL and Alwine JC. 1999b Interaction between the U1 snRNP-A protein and the 160-kD subunit of cleavage-


Mueller SN, and Ahmed R. High antigen levels are the cause of T cell exhaustion during chronic viral infection. PNAS 2009; 106:8623-8628.


Biography

Irina Evsyukova was born to Yury and Natalia Evsyukova on November 13, 1982 in Moscow, Russia. She graduated from high school in 1999. Irina attended Lomonosov Moscow State University from 1999 to 2004 and earned a Bachelor’s of Science in Bioorganic Chemistry in 2003 and Master’s of Science in Bioorganic Chemistry/Biochemistry in 2004, graduating magna cum laude. Irina was accepted to Duke University in the Department of Biochemistry in 2004 and joined laboratory of Dr. Mariano A. Garcia-Blanco in May 2006. While at Duke, Irina presented her research at the International Congress on Autoimmunity in Ljubljana, Slovenia, Center for RNA Biology meetings, monthly meetings of the Department of Molecular Genetics and Microbiology and annual departmental retreats. Irina’s publications include:
