1-1-2015

The Role of Celf2 in the Signal Induced Alternative Splicing of Lef1 Exon 6 in T Cells

Sandya Ajith
University of Pennsylvania, ajith.sandya@gmail.com

The Role of Celf2 in the Signal Induced Alternative Splicing of Lef1 Exon 6 in T Cells

Abstract
Alternative splicing is the process by which an exon is preferentially included or excluded from an mRNA transcript. Recent global sequencing studies have shown that >95% of the transcriptome undergoes some form of alternative splicing. Such regulation often alters protein isoform expression, as is especially apparent in T cells of the immune system that change their expression of RNA and protein according to signaling cues. The focus of this thesis is on one alternative exon in the pre-mRNA of transcription factor LEF1 and its regulation by the splicing factor CELF2. LEF1 is crucial for T cell function as it upregulates the expression of TCRα. Upon signal induction in T-cells, CELF2 promotes the inclusion of exon 6 in LEF1 (LEF1-E6) in the final mRNA transcript. This increase in LEF-E6 inclusion generates an isoform of LEF1 that is preferentially active in promoting transcription of TCRα. CELF2 regulates LEF1-E6 inclusion upon stimulation by increasing its binding to two conserved elements (USE60 and DSE120) in the upstream and downstream introns flanking exon 6. My goal is to understand how the increase of binding of CELF2 to the USE60 and DSE120 upon stimulation results in an increase in LEF1-E6 inclusion. Using a combination of in vivo minigene assays, in vitro splicing assays and UV-crosslinking assays I correlate the binding of CELF2 to the function of the USE60 and DSE120. I show that the USE60 and DSE120 do not work synergistically to enhance inclusion but function antagonistic to each other. The USE60 is a repressor of splicing while the DSE120 is an enhancer. In order to achieve an increase in exon 6 inclusion only upon stimulation, CELF2 binding is highly regulated between the USE60 and DSE120. In unstimulated T cells, binding is biased towards the repressive USE60 and upon stimulation the increase in CELF2 binding happens purely on the activating DSE120. This bolus of CELF2 binding on the DSE120 upon stimulation leads to an increase in exon 6 inclusion. These studies reveal a model where binding of CELF2 to the DSE120 is inhibited in unstimulated cells and this inhibition is relieved upon stimulation.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Biochemistry & Molecular Biophysics

First Advisor
Kristen Lynch

Keywords
alternative splicing, CELF2, LEF1, T cells

Subject Categories
Biochemistry

THE ROLE OF CELF2 IN THE SIGNAL INDUCED ALTERNATIVE SPLICING OF LEF1 EXON 6 IN T CELLS

Sandya Ajith

A DISSERTATION

In

Biochemistry and Molecular Biophysics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2015

Supervisor of Dissertation

Signature_____

Kristen W. Lynch, Professor of Biochemistry and Biophysics

Graduate Group Chairperson

Signature_____

Kim A. Sharp, Associate Professor of Biochemistry and Biophysics

Dissertation Committee

Gregory D. Van Duyne, Professor of Biochemistry and Biophysics

James Shorter, Associate Professor of Biochemistry and Biophysics

Arjun Raj, Assistant Professor of Bioengineering

Russ P. Carstens, Associate Professor of Medicine

Yoseph Barash, Assistant Professor of Genetics

Tara Davis, Assistant Professor of Biochemistry and Molecular Biology, Drexel University College of Medicine

DEDICATION

I would like to dedicate this thesis to my late father, K.C Ajith. Thank you for teaching me to be eternally curious, believing that I could achieve anything and watching the discovery channel with me every Sunday morning when I was kid.

ACKNOWLEDGEMENTS

I'd like to thank my advisor Kristen Lynch for her consistent patience, wisdom and support through this journey. Thanks to her I know the importance of always celebrating the victories, no matter how small, even when the task ahead is daunting.

I'd like to thank my wonderful thesis committee for their guidance and invaluable wisdom through the progression of my thesis.

I'd like to thank Ben Black for hiring me out of undergrad and starting me down this path of discovery. Thank you for believing in my abilities and teaching me how to use Illustrator.

I'd like to thank current and past members of the Lynch Lab for making my time here unforgettable. Thank you for indulging my obsession with making crossword solving a group activity.

I'd like to thank Nicole Martinez for her unwavering friendship and support. Thank you for journeying with me to streets far and wide in search for new food trucks to try.

I'd like to thank Lee Solomon for being my lab husband. Thank you for motivating me to stay later in lab than I had originally planned and walking back home from lab with me in the pouring rain. Thank for all the conversations about *Farscape*, video games and *Twin Peaks* while we waited for experiments to finish. There is no Shepard without Vakarian.

I'd like to thank Bruce Berry for being my quarterback through this journey. You always calmed me down when I felt hopeless and gave me the best advice when I was in a bind.

I'd like to thank Samantha Falk for being my partner in crime. Between the blog, tennis, dream moods, coffee, quizzo and random chats during the day, you have been my source of creativity, stress relief and fun.

I'd like to thank my kindred spirit Nikolina Sekulic for reminding me of the bigger picture when things got too stressful.

I'd like to thank Chris Yarosh for being the younger brother I never had and for always getting me in trouble with the boss.

I'd like to thank Michael Mallory for being the best bench-mate a scientist could have. I hope to always strive to live by the motto we set.

I'd like thank Erin Wohler and the quizzo team for always making Mondays better.

I'd like to thank Sam Getchell and Kyle Harpole for always making me laugh and inspiring me with your awesome brains.

I'd like to thank Zach Smoak for being the Zach to my hobbles and Evan Smoak for always being able to put things in perspective for me.

I'd to thank my mother Amrita and my brother Sandeep for their constant support and encouraging words.

I'd like to thank my wife Stephanie for always believing in me even when I don't believe in myself. Thank you for knowing what I am capable of and challenging me to new heights. Thank you for your wonderful hugs after a hard day. I couldn't have done this without you.

ABSTRACT

THE ROLE OF CELF2 IN THE SIGNAL INDUCED ALTERNATIVE SPLICING OF LEF1 EXON 6 IN T CELLS

Sandya Ajith

Kristen Lynch

Alternative splicing is the process by which an exon is preferentially included or excluded from an mRNA transcript. Recent global sequencing studies have shown that >95% of the transcriptome undergoes some form of alternative splicing. Such regulation often alters protein isoform expression, as is especially apparent in T cells of the immune system that change their expression of RNA and protein according to signaling cues. The focus of this thesis is on one alternative exon in the pre-mRNA of transcription factor LEF1 and its regulation by the splicing factor CELF2. LEF1 is crucial for T cell function as it upregulates the expression of TCRα. Upon signal induction in T-cells, CELF2 promotes the inclusion of exon 6 in LEF1 (LEF1-E6) in the final mRNA transcript. This increase in LEF-E6 inclusion generates an isoform of LEF1 that is preferentially active in promoting transcription of TCRα. CELF2 regulates LEF1-E6 inclusion upon stimulation by increasing its binding to two conserved elements (USE60 and DSE120) in the upstream and downstream introns flanking exon 6. My goal is to understand how the increase of binding of CELF2 to the USE60 and DSE120 upon stimulation results in an increase in LEF1-E6 inclusion. Using a combination of *in vivo* minigene assays, *in vitro* splicing assays and UV-crosslinking assays I correlate the binding of CELF2 to the function of the USE60 and DSE120. I show that the USE60 and DSE120 do not work synergistically to enhance inclusion but function antagonistic to each other. The USE60 is a repressor of splicing while the DSE120 is an enhancer. In order to achieve an increase in exon 6 inclusion only upon stimulation, CELF2 binding is highly regulated between the USE60 and DSE120. In unstimulated T cells, binding is biased towards the repressive USE60 and upon

stimulation the increase in CELF2 binding happens purely on the activating DSE120. This bolus of CELF2 binding on the DSE120 upon stimulation leads to an increase in exon 6 inclusion. These studies reveal a model where binding of CELF2 to the DSE120 is inhibited in unstimulated cells and this inhibition is relieved upon stimulation.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Alternative splicing and spliceosome assembly

The central dogma of biology describes a linear progression of events that starts with DNA (genes) that is transcribed into messenger RNA (mRNA) that is translated into protein. This 1:1:1 association between the three states means that the proteomes of humans and the nematode *C. Elegans* should be relatively equal with ~20,000 protein for 20,000 RNA for ~20,000 protein-coding genes. However, global studies of RNA and protein have shown that diversity in both these realms goes well beyond the 1:1:1 ratio. It is this diversity that accounts for how although humans have a similar number of genes as worms - humans are much more biologically complex having greater proteomic and cellular diversity[3].

A major mechanism by which this diversity is achieved is at the level of RNA during a process called splicing. Splicing is a co-/post-transcriptional process by which the introns in a pre-mRNA transcript are excised and the exons, which contain the protein coding information, are ligated together to form a mature transcript. However, various versions of a transcript can be created by regulating not only intron removal but exon fate as well. During the process of splicing, through a system of regulatory sequences and proteins, certain exons can be preferentially included or excluded, leading to many isoforms of a protein from the same coding gene. This is called alternative splicing and global sequencing studies have shown that this process occurs in >95% of coding genes[4,5]

The ability to create and regulate the expression of functionally diverse isoforms makes alternative splicing a powerful tool used by the cell to dictate its internal and external environment in response to various stimuli. The mechanisms underlying how an exon is alternatively spliced in response to developmental signals are based on a network of regulatory proteins that control the splicing machinery. The chemistry of splicing is undertaken by the spliceosome, a RNA-protein macro-molecular machine whose final catalytic conformation is achieved on the pre-mRNA transcript via a step-wise assembly process.

The spliceosome is composed of 5 distinct RNAs - U1-2 and U4-6 - which associate with ~145 proteins to form ribonuclear protein complexes (snRNPs)[6,7]. These snRNPs interact with varying sequences on the intron and exon - particularly the 5' and 3' Splice Sites (ss), Poly-Pyrimidine Track (PPT) and the Branch Point Sequence (BPS). Besides the splice site consensus sequences, the BPS has a conserved adenine whose 2'OH performs the first nucleophilic attack and is therefore crucial to the catalytic activity of the final spliceosomal complex. The earliest step in the process is the formation of the E (early) complex. It involves the binding of U1 snRNP to the 5'ss, U2AF heterodimer (U2AF 35 and 65) to the PPT and 3'ss. This is then converted into the A complex by the ATP-dependent addition of U2 snRNP to the 3'ss. This is followed by the recruitment of the remaining snRNPs U4-U6 (tri-snRNP) to form the B complex. Finally, after a series of re-arrangements the splicing competent C-complex (catalytic complex) is formed with the release of U1 and U4 snRNP[6,8–10]. (Figure 1.1). The catalytic complex is now capable of enabling the nucleophilic attacks required for intron excision and lariat release.

Regulation of alternative splicing via *cis*- and *trans*-acting factors

It is important to note that the interactions that drive spliceosome assembly are largely weak, such that every step can be assisted or impeded by additional regulatory proteins. In all cases of alternative splicing that have been studied in detail, the regulation of exon fate involves *trans*-acting regulatory proteins interacting with *cis*-acting sequences on the pre-mRNA transcript as well as those in the splicing machinery. These synergistically lead either to promotion of spliceosome assembly and exon inclusion or to interference with the assembly process and exon exclusion[1,11,12]. These interventions by cis- and trans- acting elements can occur at various steps in the assembly process for example:

Splice Site recognition: The earliest step that can be regulated is the recognition of splice sites by the splicing machinery. Opportunities for regulation can arise from weak splice site strength (too divergent from the consensus sequence), steric blocking by other proteins or even RNA secondary structure[12–15].

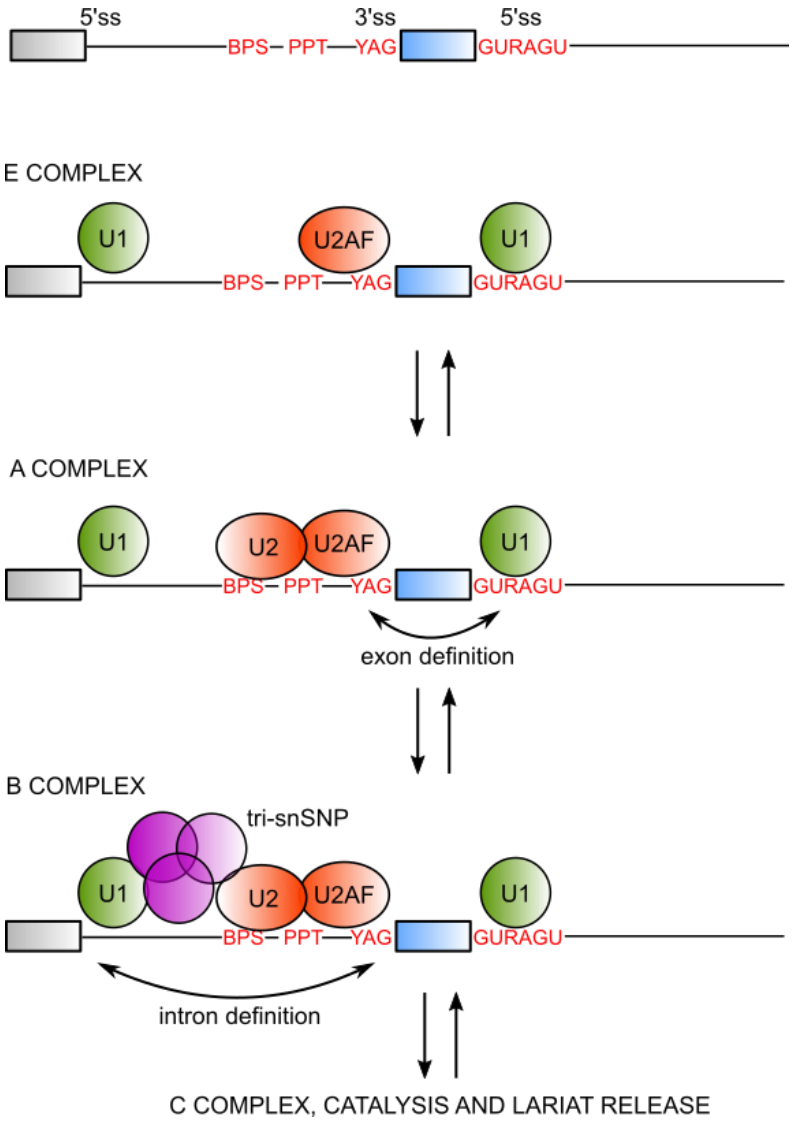


Figure 1.1: Assembly of the major spliceosome components during pre-mRNA splicing. (A) Schematic showing the location and identity of consensus sequences that guide spliceosome assembly. (Y=U or C, R-G or A, BPS=Branch Point Sequence, PPT = Poly-Pyrimidine Tract) (B) Schematic of spliceosome assembling on pre-mRNA. For simplicity, only U2AF, U1, U2 and tri-snRNP are shown.

Late spliceosome assembly: Regulation can also occur after splice site recognition and the ATP-dependent addition of U2 snRNP to the BPS. In order for C-complex formation, interactions between U1 and U2 have to occur across an intron, creating an “intron-defined” complex that is

required for catalysis. In higher eukaryotes however, studies have shown that the spliceosome first forms around the exon, creating an “exon-defined” complex[8,16]. Regulatory proteins can facilitate the conversion of exon- to intron- defined complex and cause inclusion and can also act to stabilize the exon-defined complex causing exon exclusion. Regulation can also occur after intron-definition is achieved by interfering with the recruitment of the tri-snRNP[12,17–19].

Considering the ubiquitous nature of alternative splicing and the various mechanisms by which spliceosome assembly can be regulated to achieve it, it comes as no surprise that a wide variety of *cis*-acting Sequence Regulatory Elements (SREs) and *trans*-acting regulatory proteins have been identified that typically bind to non-splice site sequences to control spliceosome assembly. SREs are varied in their lengths and depending on where they bind and the effect they have on an alternative exon, these *cis*-acting elements can be Exonic Splicing Enhancers or Inhibitors (ESEs or ESSs) or Intronic Splicing Enhancers or Inhibitors (ISEs or ISSs)[2,11,12,20].

The *trans*-acting regulatory proteins are mostly RNA Binding Proteins (RBPs) that bind to SREs and regulate alternative splicing events. Unlike SREs, RBPs cannot be broadly categorized into enhancers and repressors as their function is highly dependent on the context of the exon. A single RBP can positively or negatively influence hundreds of alternative splicing events in a cell. Conversely, the length of the SRE permitting, a single SRE can bind many RBPs that function in combination to influence exon fate.

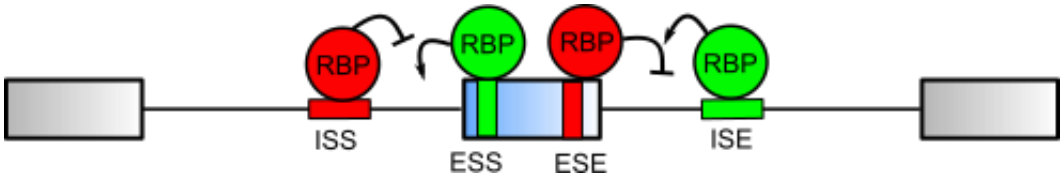


Figure 1.2: *cis* and *trans* regulatory elements regulate every individual alternative splicing event. Schematic showing *cis*-acting SREs - ESSs and ISSs inhibit exon inclusion (red boxes) and ESEs and ISEs activate exon inclusion (green boxes). *trans*-acting RBPs bind SREs and are activating or repressive to the spliceosome assembly process. (SRE = Splicing Regulatory Element, ESS = Exonic Splicing Silencer, ISS = Intronic Splicing Silencer, ESE = Exonic Splicing Enhancer, ISE = Intronic Splicing Enhancer. Partially adapted from Kornblihtt et al 2013.

There are approximately 50 RBPs that have been shown to directly interact with pre-mRNA and influence alternative splicing in mammals[3,21,22]. Each alternative splicing event is under the combinatorial control of many RBPs that bind SREs and act cooperatively or antagonistically to ultimately decide whether an exon is included or excluded from the final transcript[2,3,21,22] (Figure 1.2).

Alternative Splicing in the Immune System

The regulation of alternative splicing plays a crucial role in processes like the epithelial-mesenchymal transition, regulation of action potentials, heart development and, of importance to this thesis, the regulation of T-cell function in the immune system[21,23–25]. T-cells are lymphocytes that play a crucial role in adaptive immunity. Pre T cells develop from pluripotent stem cells in the red bone marrow that then travel to the thymus for maturation. The maturation process involves the expression of the T cell receptor (TCR) as well as either the CD4 or CD8 protein on its plasma membrane, termed CD4 or CD8 T cells. In the presence of a foreign antigen, TCR in conjunction with either the CD4 or CD8 protein bind antigen and trigger a signaling cascade that is the start of an immune response[26,27].

This signaling cascade leads to a large number of changes within the T cell, such as increased expression of TCR, increased proliferation and cytokine production and secretion to name a few. Therefore, the effectiveness of the immune system depends on its ability to orchestrate large changes in protein expression in response to antigen signaling. Several studies that investigated changes in alternative splicing in T cells using high-throughput RNA sequencing (RNA-seq) or microarrays have shown that one of the ways T cells respond to external signaling is through changes in alternative splicing[25,28–30]. In a 2012 study by the Lynch Lab to which I contributed, high-throughput RNA sequencing (RNA-seq) was used to highlight the alternative splicing networks involved in regulating the start of an immune response. The study used a Jurkat derived model T cell line[31], called JSL1, as well as primary human T cells to show that T cell stimulation caused significant changes in the isoform profile of 178 exons in 168 genes[25]. A handful of these genes

have been studied further and illustrate the various ways alternative splicing is used to regulate T cell activation.

The most studied alternative splicing event is the regulation of exons 4, 5, and 6 of the CD45 gene. CD45 is a transmembrane tyrosine phosphatase responsible for regulating antigen receptor signaling and lymphocyte development. Upon activation of T cells, exon 4, 5 and 6 are excluded, leading to the expression of an inactive form of the phosphatase. The repression of these exons is caused by ESSs present in all three exons, however mechanistic details are only available for exons 4 and 5. In the case of exon 4, this ESS binds to hnRNPL, hnRNPLL and PSF of which the latter two are responsible for signal induced exon exclusion. In the case of exon 5, SRSF1 binds an ESE within the exon but it's activating effect is displaced upon stimulation by the binding of hnRNP L and PSF to two flanking ESSs[17,32–35].

The regulation of CD45 exons 4 and 5 highlight the complexity involved in the coordination of SREs and the RBPs that bind to them. Besides CD45, mechanistic details are only available for the signal induced regulation of four other genes - CD3ζ exon 8 by SRSF1[36], Fas exon 6 by TIA-1[37], CD44 exon v5 by Sam68[38] and the focus of this thesis, LEF1 exon 6 by CELF2[39].

CELF2 is a regulator of alternative splicing

CELF2 is part of the CUG and ETR-3 Like Factor (CELF) family of proteins of which there are 6 members. All members of this family are characterized by three RNA Recognition Motifs (RRMs). RRM1 and 2 lie at the N terminus of the protein and RRM3 lies at the C terminus with a linker domain linking RRM 2 to 3. CELF2 shows high similarity (>90%) with CELF1 within their RRMs but diverge greatly in the linker domain[40,41]. Several studies have shown that the RRMs of CELF proteins bind UG-rich sequences. Structural studies of the RRMs of CELF1 show all three RRMs bind UGUU motifs. However, RRM1 and 2 have higher affinity for UG-rich RNA when linked together than when separate. Additionally, part of the linker domain was shown to greatly increase RRM3's affinity for RNA, showing that this region could be very important to function[42–47].

Figure 1.3: Domain structure of CELF2 protein. Schematic showing all annotated domains of CELF2. Red ovals corresponed to RNA Recognition Motifs (RRMs). Blue line corresponds to the Linker Domain. Numbers above the protein correspond to amino acid positions.

There are two alternative splicing events important to CELF2 structure and expression. CELF2 is involved in autoregulation of its own transcript by repressing exon 6 inclusion. If translated this would create a protein that is truncated at RRM2 as the skipping of exon 6 seems to cause a reading frame shift that introduces a premature stop codon and triggers the Nonsense Mediated Decay(NMD) pathway[48]. Additionally, there is evidence for the regulation of exon 14 that encodes for the beginning of RRM3. Molecular Dynamics coupled with NMR studies imply that the skipping of this exon would make RRM3 incapable of binding RNA. This isoform therefore has differential effects on alternative splicing as opposed to its full length protein. This isoform has been shown to be expressed at significant quantities only in the kidneys and liver[49].

Unsurprisingly for a RNA binding protein with three RRMs CELF2 plays large roles in alternative splicing and mRNA stability. CELF2's roles in mRNA stability are of particular interest as a target for disease therapeutics. CELF2 stabilizes the poly-glutamine extended Androgen Receptor (AR) mRNA in Spinal and Bulbar Muscular atrophy and targeted silencing of CELF2 successfully led to decay of the toxic AR mRNA[50]. CELF2 is of potential interest for cancer therapeutics as it hyper-stabilizes the anti-apoptotic factors COX2 and MCL1 mRNA in pancreatic and colon cancer cells, thereby inhibiting their translation and encouraging apoptosis[51–53].

In the case of alternative splicing, CELF2 has been shown to act as both an activator and repressor of exon inclusion. Besides LEF1-E6 there are 11 mechanistic studies of CELF2 regulating alternative splicing. An example of CELF2 activating exon inclusion is the regulation of Cardiac Troponin T (cTNT)'s exon 5 which is preferentially included in embryonic striated muscle but excluded in adult tissue. The isoforms created from the alternative splicing of exon 5 confer

different contractile properties to the muscle tissue. CELF2 ensures that exon 5 is included in embryonic tissue by binding to a UG-rich element downstream of exon 5 and acting across the exon to stabilize the binding of U2snRNP to the 3'ss and encourage exon definition[54–56].

An informative example of CELF2 repressing exon inclusion comes from exon 9 of the Cystic Fibrosis Transmembrane Regulator (CFTR) gene. CELF2 binds to a UG-rich sequence upstream of exon 9 and represses exon inclusion by displacing binding of constitutive splicing factor U2AF65 from the PPT. This repression was dependent on the rate of transcription as a slow rate of transcription allowed for greater CELF2 recruitment and binding upstream and greater exon skipping. A faster elongation rate presumably reduced the amount of CELF2 recruited upstream and thus, less displacement of U2AF65[57].

Based on all published studies of CELF2 regulating alternative splicing, it seems to have a positional dependence on how it influences exon inclusion. CELF2 binding upstream of the alternative exon as in CFTR exon 9, NMDAR1 exon 5 (N1), α actinin exon NM and CELF2's own exon 6 leads to exon skipping[48,57–59]. However, CELF2 binding downstream of an alternative exon as in cTNT exon 5 and NMDAR1 exon 21 (C1), encourages exon inclusion[54,55,59]. Global sequencing studies and MS2 tethering assays that correlate RBP binding to exon fate have revealed that a significant number of splicing factors including PTB, SRSF1,2,6,7,and10, TIA-1, Fox2α, FUS, hnRNPA1 and hnRNP F/H show evidence for positional dependent effects on exon fate[60–64]. Unpublished work from our lab that investigated CELF2 binding and alternative splicing regulation on a global scale in T cells confirms that CELF2 functions through a similar mechanism.

In the model T cell line JSL1 and in developing thymocytes, CELF2 expression increases upon signal induction through both an increase in transcription and mRNA stability. The increase in CELF2 expression is dependent on the NF-κB signaling pathway. The increase in CELF2 levels have wide effects in alternative splicing changes that happen during signal induction affecting a third of all splicing events that undergo a signal induced change[65]. One of the signal induced splicing events that requires CELF2 is the preferential increase of exon 6 of LEF1 upon T cell stimulation.

LEF1 exon 6 and its effect on TCRα expression

Lymphoid Enhancer-binding Factor 1 (LEF1) is a transcription factor involved in the regulation of a wide variety of cellular processes. More specifically, it regulates many developmental programs including that of the hair follicle, teeth, osteoblasts and mammary glands[66–69]. It has also been implicated in the progression of several cancer populations including gastrointestinal and pancreatic cancer[66,70]. It is characterized by a β-catenin binding domain called the Activation Domain (AD) at its N-terminus and a High Mobility Group (HMG) DNA Binding Domain (DBD) with a Nuclear Localization Signal (NLS) at its C-terminus. The N and C termini are separated by a Context Regulatory Domain that is encoded by alternative exon 6[70,71] (Figure 1.4).

In T cells, LEF1 is crucial for upregulating the expression of T-cell Receptor Alpha (TCRα). TCRα, along with TCRβ, is required to form a mature TCR which is crucial for its response to antigen binding and its maturation in the thymus. LEF1, through its CRD, forms protein-protein interactions in an enhanceosome complex that activates TCRα expression. LEF1 exon 6 (LEF1-E6) encodes part of the CRD and its exclusion from the final transcript creates an isoform of LEF1 that cannot upregulate TCRα (Figure 1.4 panel A).

Previous work done in the lab established that there is a preferential increase in inclusion of exon 6 during thymic development when immature T-cells transition from the Double negative (Dn) to double positive (Dp) state. This is recapitulated in the JSL1 cells when stimulated with the phorbol ester PMA. The preferential inclusion of exon 6 correlated with increased TCRα expression in both cases[39,72,73]. Additionally, Mallory et al was able to show that the enhancement of TCRα was a direct result of exon 6 splicing. The authors used a splice site morpholino to force exclusion of exon 6 and this resulted in a significant decrease in TCRα mRNA[39]. Therefore the mechanism by which LEF1 isoform choice is regulated by the alternative splicing of exon 6 has important implications for T-cell function and development (Figure 1.4).

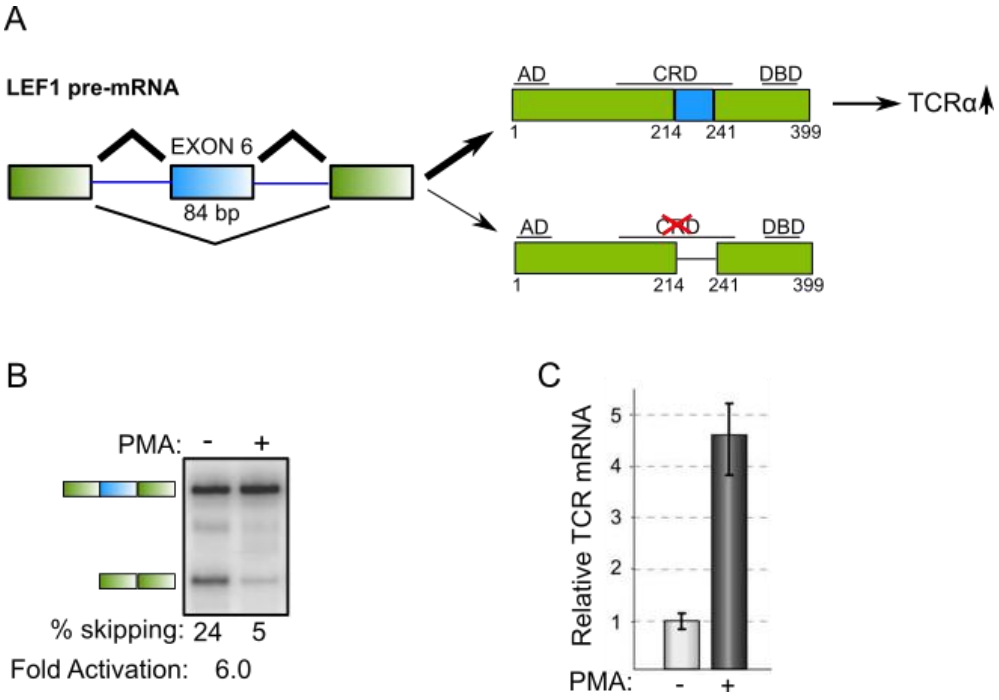


Figure 1.4: Regulation of LEF1-E6 inclusion upon PMA stimulation in JSL1 cells. (A) Schematic of the alternative splicing of 84 nucleotides (nt) of LEF1-E6 and the resulting protein isoforms. The Activation Domain (AD), Context Regulatory Domain (CRD) and DNA Binding Domain (DBD) are labeled on the isoforms. (B) Representative RT-PCR of endogenous LEF1 showing loss of skipping of LEF1-E6 upon PMA stimulation. Fold Activation (FA) is used as a measure of fold change of exon inclusion upon PMA stimulation. (C) Graph showing change in TCR-α mRNA upon PMA stimulation.

The regulation of LEF1-E6 by CELF2 in T cells

The same study by Mallory et al narrowed down the required SREs for regulation of LEF1-E6 inclusion upon T cell stimulation to two intronic elements upstream and downstream of the exon. Even replacing the exon (∆exon) did not affect enhancement of LEF1-E6 upon stimulation. The two intronic elements were called the Upstream Sequence Element, that is 60 nucleotides (nts) long (USE60), lies immediately upstream of exon 6 and includes the 3'ss and the Downstream Sequence Element that is 120 nts long (DSE120) and lies 31 nts downstream from exon 6. These sequences are highly conserved and rich in UG motifs, which are known CELF2 binding sites. CELF2 binds the USE60 and DSE120 and is functionally required for exon inclusion (Figure 1.5). Importantly, upon stimulation, the binding of CELF2 to these elements increases. The study showed that a 50% reduction in CELF2 protein by shRNA resulted in a 2-3 fold decrease in LEF1-

E6 inclusion and a corresponding decrease in TCRα mRNA expression[39]. In this thesis, I extend this study by Mallory et al by investigating the mechanism by which CELF2 interacts with the LEF1 pre-mRNA transcript to facilitate the preferential inclusion of variable exon 6 upon T cell stimulation.

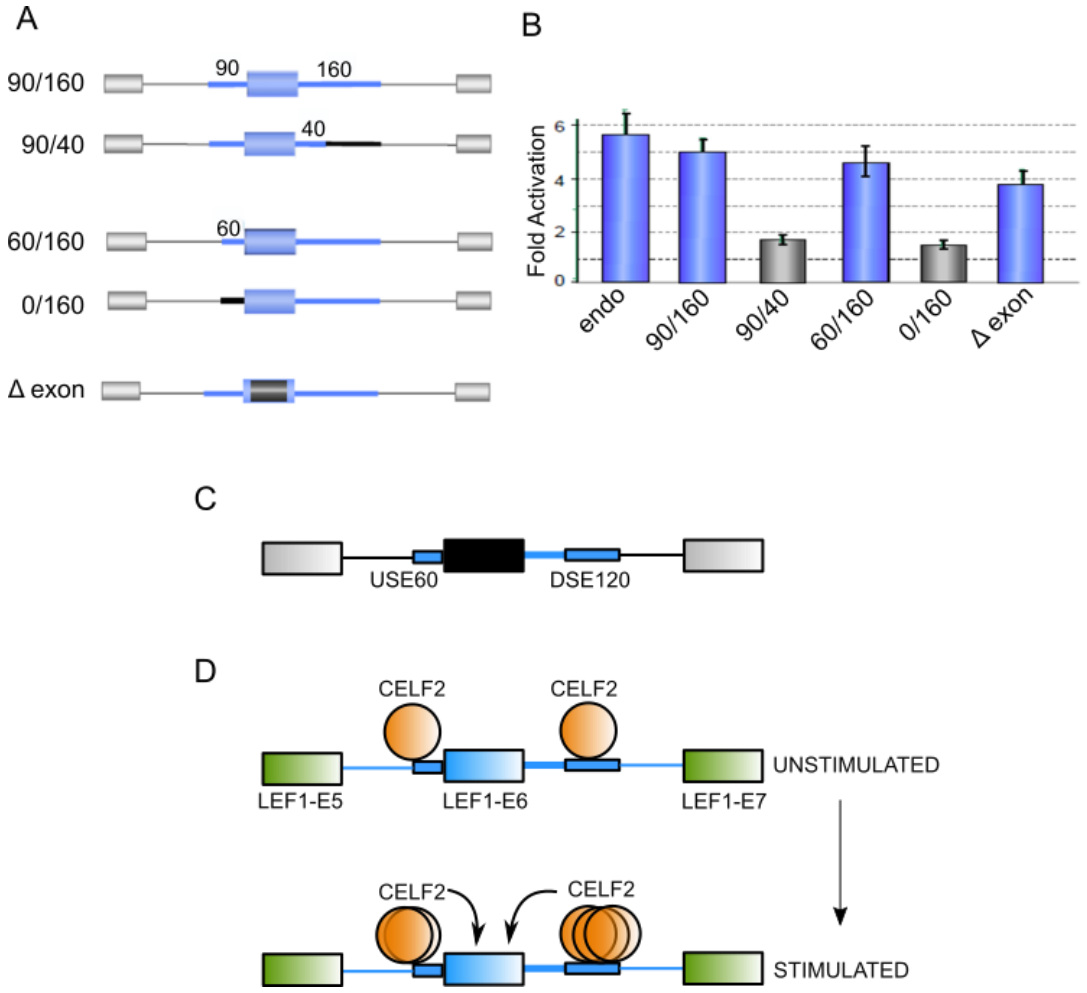


Figure 1.5: USE60, DSE120 and CELF2 are required for enhancement of LEF1-E6 inclusion upon T cell stimulation. (A) Schematic of minigenes used to discover SREs USE60 and DSE120 that are requried for LEF1-E6 regulation. Blue boxes and lines correspond to LEF1 exon 6 and flanking intronic sequences. Black boxes and lines refer to heterologous filler sequences. Grey boxes correspond to constitutive exons. Numbers above the lines are the approximate length of LEF1 intronic sequence (B) Graphical representation of Fold Activation for corresponding minigenes from (A). (endo = endogenous) (C) Schematic of Δexon minigene containing SREs, USE60 and DSE120. Smaller blue boxes refer to SREs. (D) Model for enhancement of LEF1-E6 inclusion upon T cell stimulation by the binding of CELF2 (based on data from Mallory et al 2011). Green boxes refer to flanking LEF1 exons. Orange circles correspond to CELF2 protein. Arrows towards exon correspond to activation of exon inclusion. Adapted from Mallory et al 2011.

CHAPTER 2: THE USE60 AND DSE120 ARE SUFFICIENT FOR SIGNAL INDUCED ENHANCEMENT OF LEF-E6 INCLUSION

Introduction:

Splicing is a complicated process that is under the influence of a large number of *cis* and *trans*-acting factors. Depending on the mechanism by which they exert their influence, splicing factors can either act on maintaining levels of inclusion in the unstimulated state (basal) levels or play a role in a signal induced change in exon inclusion. An informative example is the CD45 gene, whose exon 4 is regulated in a signal-dependent manner in T cells. Three splicing factors, hnRNP L, hnRNP LL and PSF act on exon 4 to repress exon inclusion but not all of them are involved in the signal induced repression. hnRNP L mediates basal levels of inclusion in unstimulated T cells by binding to an ESS in exon 4 and its effect remains unchanged upon stimulation. hnRNPLL and PSF, however, only bind the ESS in stimulated cells and cause further repression of CD45 exon 4[34,35,74].

In the case of CD45 exon 4, even though all three proteins bind to a single repressive element, the ARS, they have very distinct mechanisms of repression, with hnRNP L acting in both unstimulated and stimulated cell while hnRNP LL and PSF acting only in stimulated cells . Specifically, considering how many complex mechanisms can occur even on a single signal responsive element, it is imperative to distinguish between those that act on basal splicing and those that act on signal induced changes. Therefore, determining the minimum sequence requirements (*cis* factors) required for a signal-induced change in splicing is an important first step in reducing the complexity of the system. These minimum sequence requirements are crucial in being able to isolate the mechanism responsible for the signal induced change from the myriad of other mechanisms at play.

Previously Mallory et al. concluded that 2 intronic SREs, the USE60 and DSE120, were *required* for the signal induced enhancement of LEF1-E6 inclusion. The authors could not conclude that these regulatory sequences were *sufficient* for this regulation because the minimum construct used retained additional LEF1 sequences. In particular, the requirement of the sequence that connects the 5'ss to the DSE120, called the Downstream Connecting Sequence (DCS) for the

signal induced regulation of LEF1-E6 had not been tested *in vivo*. Here I describe the use of minigene assays in determining whether the DCS is required for enhancement of LEF1-E6 inclusion. Establishing the most minimal construct required for LEF1-E6 enhancement upon stimulation is imperative to being able to discover and understand the mechanism by which it occurs.

Results:

The DCS is a stretch of 31 nucleotides that extends from after the 5' splice site to the DSE120 (Figure 2.1). It is not as conserved as the DSE120, with the highest conservation found in the 8 nucleotides immediately after the 5'ss (Figure 2.1, nucleotides with asterisks above them). In order to determine whether the DCS is required for signal induced enhancement of LEF1-E6 inclusion, I created a minigene construct in which the DCS was replaced with a 38 nt heterologous sequence shown previously in the lab to have no effect on splicing regulation (ΔDCS).

A minigene is a simplified construct that contains the variable exon in question along with the relevant amount of flanking intronic sequence. The sequence of interest is amplified out of genomic DNA, inserted into an expression vector and placed under the control of a T7 promoter. The variable exon and relevant intronic sequences are flanked by two constitutive exons from a known and tested gene. These minigenes can be used for *in vivo* studies by transient transfection into a cell line of choice or for the construction of stable cell lines[75].

For this study, the alternative exon and intronic sequences are flanked by β-globin constitutive exons 1 and 2. The ΔDCS, 90/160, Δexon and 90/40 minigenes were transfected into the Jurkat-derived T-cell line called JSL1. The minigenes were tested under unstimulated and stimulated conditions. JSL1 cells were stimulated using Phorbol-12-myristate-13-acetate (PMA), which has been previously shown to mimic T-cell activation through the TCR[76]. After 72 hours of stimulation, RNA was extracted from unstimulated and stimulated cells and Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR) with radio-labeled primers was used to amplify the region in between and including the constitutive exons. The reactions were visualized on a denaturing

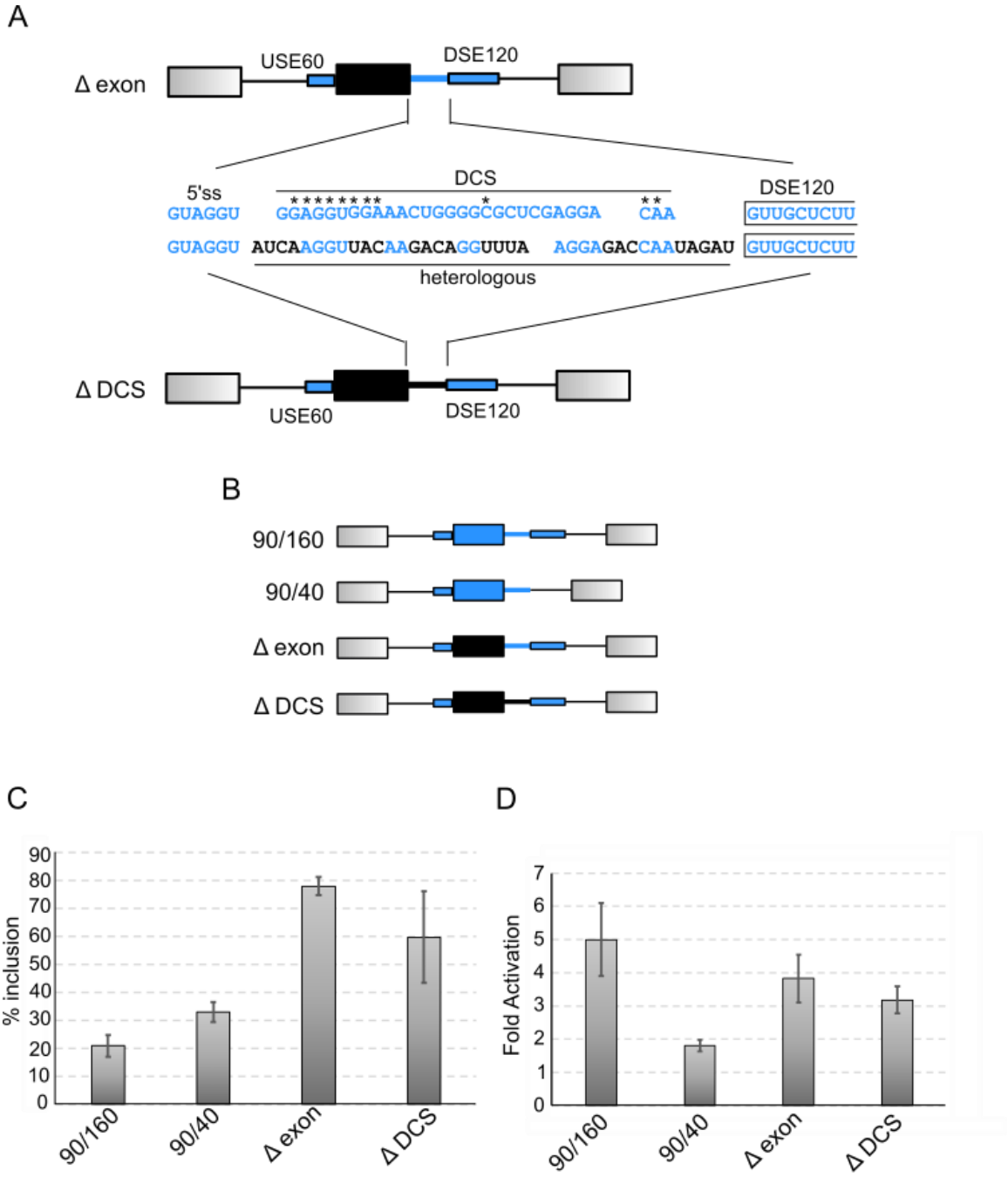


Figure 2.1: Downstream Connecting Sequence (DCS) is not required for enhancement of LEF1-E6 inclusion. (A) Schematic showing the sequence of the DCS from Δexon minigene compared to heterologous sequence from ΔDCS minigene. Blue coloured nucleotides indicate LEF1 sequences. Black coloured nucleotides indicate heterologous sequences. (B) Schematic of minigenes used to test the requirement of the DCS. (C) and (D) Graphical representation of % inclusion and Fold Activation respectively for corresponding minigenes from (B).

PAGE gel where the RT-PCR products are separated by size. The % exon included and % exon excluded for each condition was quantified by densitometry using a phosphor-imager. A measure called Fold Activation (FA) was used to accurately measure the amount that exon inclusion increases after stimulation. FA is calculated by using the formula: FA = $(\%exclusion/\%inclusion)_{unstimulated}$ / $(\%exclusion/\%inclusion)_{stimulated}$. This accounts for any variability in basal levels of inclusion that can skew the measurement. Based on results from Mallory et al and previous experience in the lab with LEF1 minigenes, a FA above 2.5 is considered a signal responsive minigene.

Figure 2.1 shows that the ΔDCS increased basal levels of inclusion relative to the 90/160 minigene, which implies the presence of an Intronic Splicing Silencer (ISS) in the DCS. However, ΔDCS did not significantly affect signal induced enhancement of exon 6 inclusion as the fold activation was comparable to the 90/160 and Δexon minigenes i.e well above the 2.5 fold. This suggests that the DCS influences basal levels of inclusion but is not responsible for the increase in exon inclusion upon PMA stimulation (Figure 2.1).

However, upon closer inspection of the heterologous sequence used in the ΔDCS, I noticed that it contained 4 of the 8 conserved residues of the DCS, in a similar position as the DCS (Figure2.2). The residues of interest were AGGT and therefore to ensure that those 4 residues did not play a part in LEF1-E6's signal induced regulation, I used a minigene construct that substituted the conserved G's in the DCS to C's (90/160 mut2). The 90/160mut2 minigene had higher levels of basal inclusion, similar to the ΔDCS, but did not have any effect on FA levels. The minigene data from the 90/160mut2 minigene taken together with the ΔDCS minigene confirm that the DCS is not required for the signal induced enhancement of LEF1-E6.

If the minimal requirement for regulation of LEF1-E6 is the USE60 and DSE120, it is important to confirm whether the location of these sequences is important to regulation or whether they serve as a recruiting tool to concentrate more CELF2 in the region of LEF1-E6. In order to test whether USE60 and DSE120 need to flank LEF1-E6 I created a minigene that moved the DSE120

upstream of the USE60 (DSE-I1). If the DSE120 was functioning just as a tool to recruit CELF2 to the region of exon 6, this displacement of the motif would not affect signal induced exon inclusion.

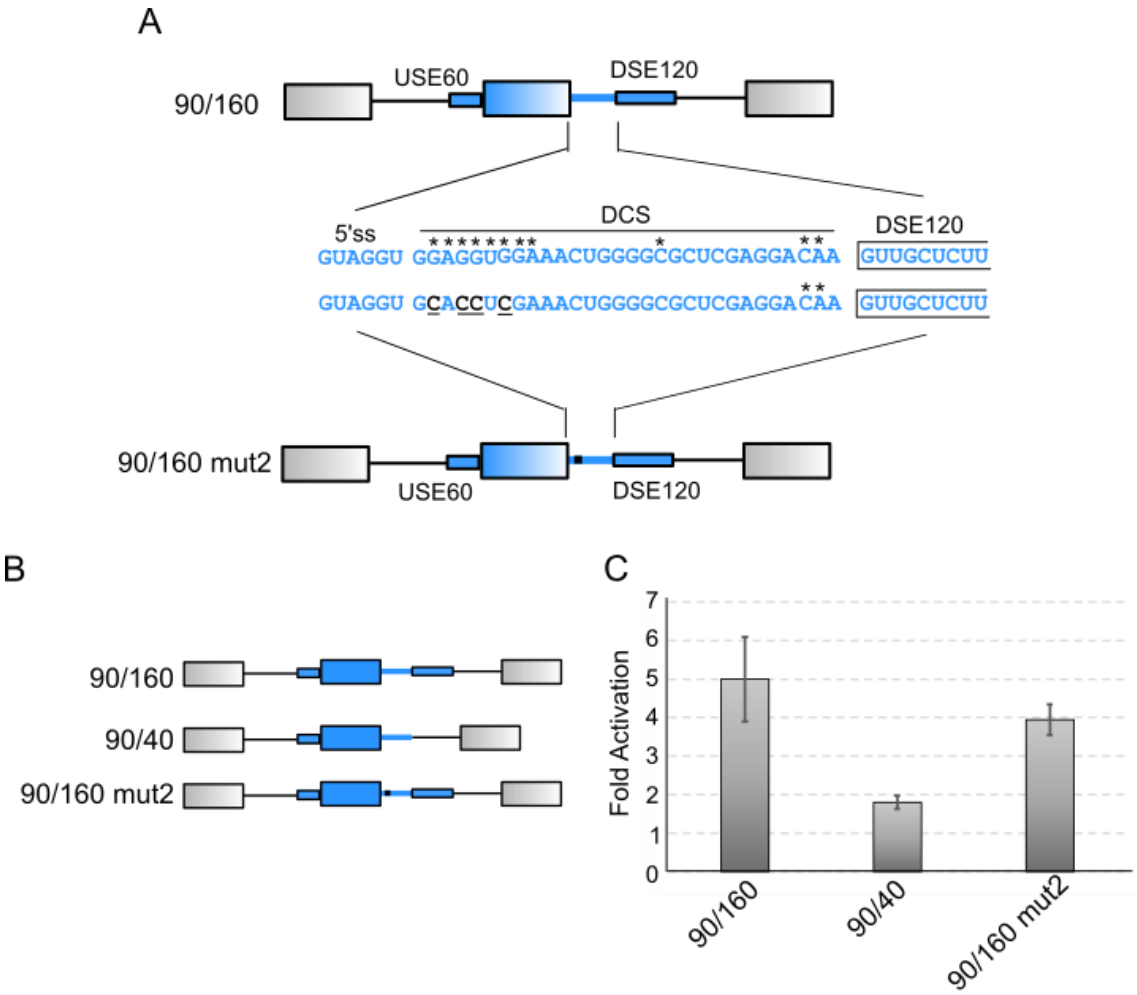


Figure 2.2: Conserved nucleotides in DCS are not required for enhancement of LEF1-E6 inclusion. (A) Schematic showing the sequence of the DCS from 90/160 minigene compared to that from 90/160 mut2 minigene showing G to C mutations. Blue coloured nucleotides indicate LEF1 sequences. Black and underlined nucleotides indicate mutations. (B) Schematic of minigenes used to test the requirement of the conserved G's in the DCS. (C) Graphical representation of Fold Activation for corresponding minigenes from (B).

.

However, if the binding of CELF2 downstream was crucial to the mechanism of exon 6 inclusion, there would be no signal dependent increase in inclusion in DSE-I1. Figure 2.3 shows that there is no significant increase in exon 6 inclusion upon PMA stimulation in DSE-I1 as the FA

remains close to that of the ΔDSE120 in DSE-I1 as opposed to the higher FA of the WT construct. This confirms that the location of the DSE downstream of the exon and hence binding of CELF2 downstream of the exon is required for this regulation.

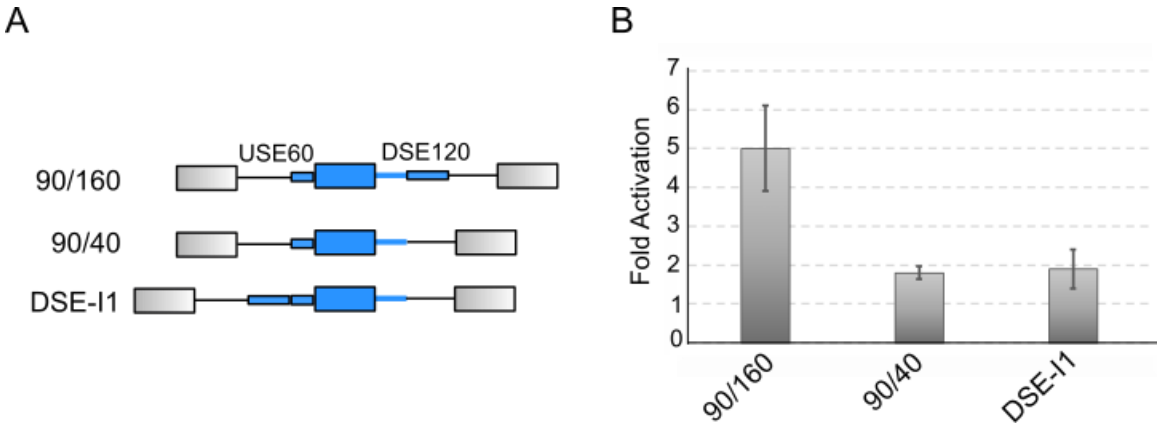


Figure 2.3: Location of the DSE120 downstream of LEF1-E6 is crucial for signal induced regulation. (A) Schematic of minigenes used to test the importance of the location of the DSE120 for signal induced regulation by LEF1. (B) Graphical representation of Fold Activation for corresponding minigenes from (A).

Discussion:

The minigene studies interrogating the requirement of the DCS have confirmed that the minimal sequence requirements for the signal induced enhancement of LEF1-E6 is the USE60 and DSE120. Additionally is it important to stress that the USE60 and DSE120 are *necessary and sufficient* for signal induced inclusion of LEF1-E6.

Having regulatory elements that bind the same protein flank an exon to cause inclusion has not been described in the literature before. The most studied example of splicing regulation by two *cis* elements in the flanking introns of an alternative exon is Poly-pyrimidine track binding protein's (PTB) repression of c-src's N1 exon[18,77,78]. PTB binds these two sequence elements to induce exclusion of the N1 exon in non-neuronal cells, where PTB expression is higher than in neuronal cells. Both these sequences are required for repression and studies have shown that an exon-defined spliceosomal complex is prevented from being converted to a splicing competent intron-defined complex[18,78].

However, the mechanism by which this is achieved is unclear, so is the reason for requirement of both upstream and downstream elements. It is known that PTB interacts with stem-loop 4 in the U1 snRNA but it is unclear which PTB molecule (ones bound upstream or downstream) is involved in this interaction or whether the interaction itself is required for repression. It is possible that this interaction with U1 inhibits intron-defined interactions from forming, but this is yet to be confirmed.

What is known about PTB's regulation is that binding downstream is required for PTB association upstream. This suggests a model where either one PTB molecule forms bridging interactions across the N1 exon or multiple PTB molecules interact with each other across the exon. This would result in the exon being looped out leading to disruptions in normal splicing interactions[77,79,80].

PTB's regulation of c-src's N1 exon is a good template by which to evaluate CELF2's regulation of LEF1-E6 alternative splicing. The interactions made by CELF2 with the splicing machinery are still unknown, however, just like PTB binding around the N1 exon, CELF2 binds SREs in the two flanking introns. The N1 exon is included in neurons where the expression of PTB is low and is excluded in non-neuronal cells where PTB levels are high[78–80]. An increase in PTB binding to the two intronic elements leads to an increase in exon *exclusion*. In LEF1-E6, an increase in CELF2 binding leads to an increase in exon *inclusion*. Mallory et al established that if the USE60 and DSE120 are radiolabeled and subjected to UV crosslinking in unstimulated and stimulated JSL1 nuclear extract, CELF2 is the only protein that increases binding upon stimulation. These experiments were done with the sequences in isolation, without the context of the rest of the RNA. It would be informative to know the pattern of CELF2 binding in the context of LEF1-E6 when both the USE60 and DSE120 are present. In the case of PTB, the downstream element is required to stabilize PTB binding upstream[79]. Is CELF2 similar to PTB wherein it requires both elements to bind to the RNA at all or do the USE60 and DSE120 function independently? This will be addressed in chapter 3 of this thesis.

One of the first steps towards deciphering the mechanism behind LEF1-E6 activation is to determine the function of the USE60 and DSE120. We know from PTB that the regulatory elements

ultimately lead to exclusion, however a more global study has shown that PTB binding downstream is more associated with inclusion of exon while upstream and exonic binding correlate with exclusion. The authors were also able to convert an exon that was activated by PTB binding downstream to being repressed by adding a PTB binding element upstream. The repression was most robust when there were twice the number of PTB sites downstream versus upstream[63]. Determining whether the USE60 and DSE120 are ISE's or ISS's would be crucial to interpreting downstream studies on the mechanism of LEF-E6 activation and its interaction with spliceosomal components. Chapter 4 of this thesis will focus on the functions of the USE60 and DSE120.

Another possibility is the requirement for another splicing factor that regulates CELF2 binding to the USE60 and DSE120. In the case of PTB, the exon defined complex that forms on the N1 exon in non-neuronal cells contains different proteins than the one that forms in neuronal cells where N1 is included. This could mean the involvement of other factors that clarify the role of the upstream and downstream elements[18]. Chapter 5 and 6 of this thesis addresses that possibility in CELF2's regulation of LEF1-E6.

CHAPTER 3: CELF2 BINDING TO THE USE60 AND DSE120 WITHIN THE CONTEXT OF LEF1 EXON 6

Introduction:

The enhancement of LEF1-E6 inclusion upon T cell stimulation is regulated by two *cis*-elements, one upstream of the exon called USE60 and one downstream of the exon called DSE 120. The splicing factor CELF2 interact with these two elements in unstimulated cells and upon stimulation there is an increase in CELF2 binding to these elements which leads to an increase in exon 6 inclusion. These experiments were done with the USE60 and DSE120 in isolation and outside the context of the exon they would be regulating. The sequence that would connect the USE60 and DSE120 could greatly influence the degrees to which CELF2 has access to these sequences. Therefore it is possible that within the context of the exon, the pattern of binding of CELF2 to the USE60 and DSE120 might be biased for one over the other.

Whether CELF2 binding between the USE60 and DSE120 is distributed equally or whether one sequence element is favoured over the other, can have large impacts on the influence CELF2 is having on the exon. There is a large body of literature confirming that the regulation of splicing by *trans*-factors is highly context dependent. Whether a splicing factor is a repressor or an activator can depend on whether it binds in the upstream intron, the exon or the downstream intron. Global studies that correlate protein binding with alternative exon fate have shown that the Rbfox family of proteins[81,82], TIA family of proteins[64], PTB[63], SRSF10[60], PUM2 and QKI[83] all show position-dependent effects on exon inclusion. Additionally, unpublished data from other members of the lab suggest that CELF2, also has positional effects on splicing. Since the crux of the mechanism behind the enhancement of LEF1-E6 upon stimulation lies in the differential binding of CELF2 to the USE60 and DSE120, accurately mapping its binding in the unstimulated and stimulated state in the context of LEF1-E6 is vitally important.

Results:

Previous work by Mallory et al showed via Ultra-Violet (UV) crosslinking studies in unstimulated and stimulated nuclear extract that CELF2 binds the USE60 and DSE120 in the unstimulated and stimulated state. UV-crosslinking is a standard procedure used to form stable

covalent crosslinks between nucleic acids and protein that are within a few angstroms of each other. Radio-labeled USE60 and DSE120 RNA is radio-labeled and incubated with JSL1 Nuclear Extract (NE) under splicing conditions and cross-linked with UV light (254nm). The RNA is then digested using RNases (T1 and A) and the proteins are separated on a denaturing SDS-PAGE gel. Proteins that bound to the RNA were detected by autoradiography as they are covalently linked to radio-labeled nucleotide. UV crosslinking is followed by Immunoprecipitation (IP) of CELF2 and control antibodies. Figure 3.1, adapted from Mallory et al, shows that concurrent with the enhancement of exon 6 inclusion upon stimulation, CELF2 binding to both sequences is enhanced upon stimulation[39] and the current hypothesis is that it is this increase in binding that leads to an increase in exon 6 inclusion. (Figure 3.1).

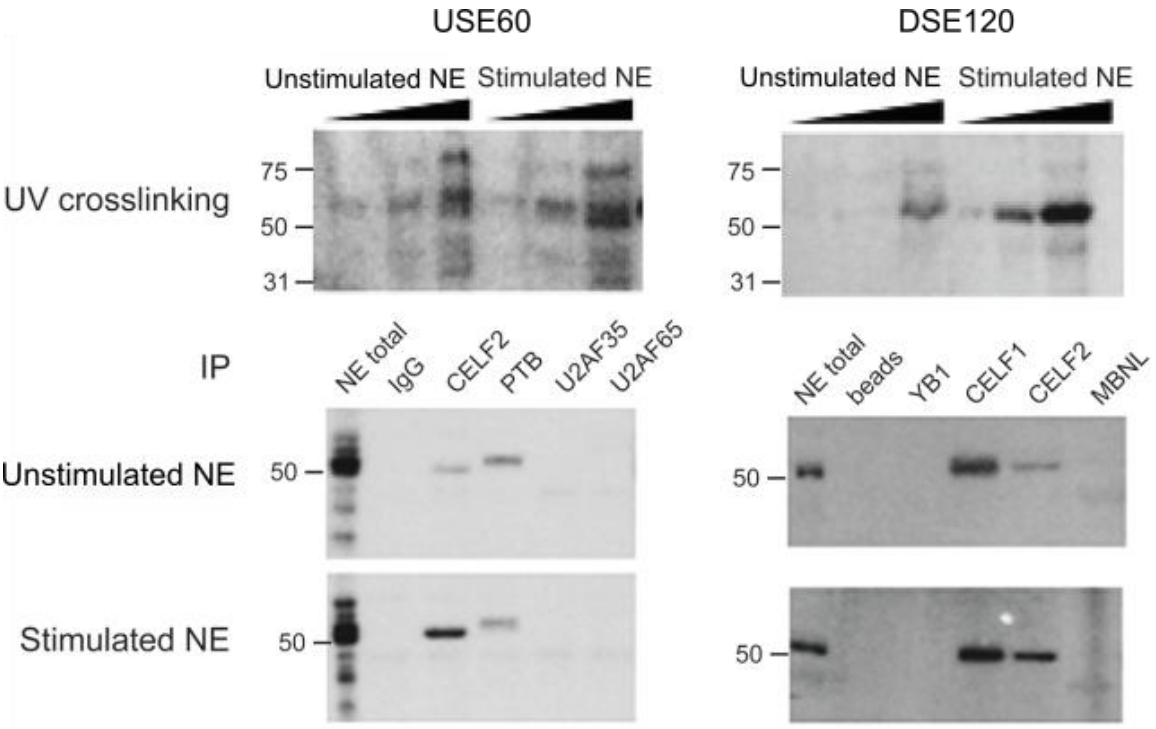


Figure 3.1: The USE60 and DSE120 bind CELF2 in a stimulation dependant manner. UV crosslinking was performed with body labeled USE60 and DSE120 in increasing amounts of Nuclear Extract (NE) denoted by black triangles above the lanes. Molecular weight markers are shown in kDa to the left of the gels. Immunoprecipitations (IP) of UV crosslinking reactions were done using antibodies after Rnase treatment. Adapted from Mallory et al.

To quantify how robust CELF2's interaction with these sequences are, I performed Electro-Mobility Shift Assays (EMSA) using bacterially expressed his-tagged CELF2 (his-CELF2) and radiolabeled USE60 and DSE120 (Figure 3.2 panel B). The radiolabeled RNA is incubated with his-CELF2 to allow complex formation. The complexes are visualized using autoradiography on a native gel. Figure 3.2 shows that the USE60 and DSE120, in isolation from each other and the exon, are potent binders of CELF2 and bind with an apparent Kd (dissociation constant) of 160nM.

For comparison, high affinity CELF2 binding sequences (2x and 4xUGUU), acquired from a Systematic Evolution of Ligands by Exponential Enrichment (SELEX) study by Faustino et al, was used[84]. The 2xUGUU and 4xUGUU have repeated instances of a sequence rich in UG motifs. The 2xUGUU sequences is approximately 60 nucleotides in length and serves as a length control for the USE60 while the 4xUGUU is approximately 120 nucleotides and serves as a length control for the DSE120. 2x has a total of 28 UG di-nucleotides and consequently 4x has a total of 56 UG di-nucleotides. This is substantially larger than the USE60 with 8 UG di-nucleotides and the DSE120 with 19 UG dinucleotides. Figure 3.2 shows that the 2xUGUU is a weak binder of his-CELF2, relative to the USE60, and does not saturate the RNA even at 1800nM of his-CELF2. The 4xUGUU however binds with an apparent Kd of 380nM and is closer in binding potency to the USE60 and DS120. The USE60 and DSE120 can therefore be categorized as high affinity CELF2 binding sequences. (Figure 3.2)

Considering the USE60 and DSE120 have equal affinity for CELF2 suggests a simplistic model where the increase in CELF2 binding upon stimulation occurs equally at both elements. However, this ignores the effect the LEF1-E6 regulatory landscape could have on protein binding, the least of which involves having both the USE60 and DSE120 present in the same substrate. In order to monitor CELF2 binding in the context of the LEF1 regulatory landscape, I repeated the UV crosslinking experiments with a wild-type (WT) construct that contained the USE60-exon6-DSC-DSE120 sequences (Figure 3.3). Figure 3.3 shows that many proteins bind the sequences in and around exon 6 however, the most predominant change is the band around 50kDa. In order to

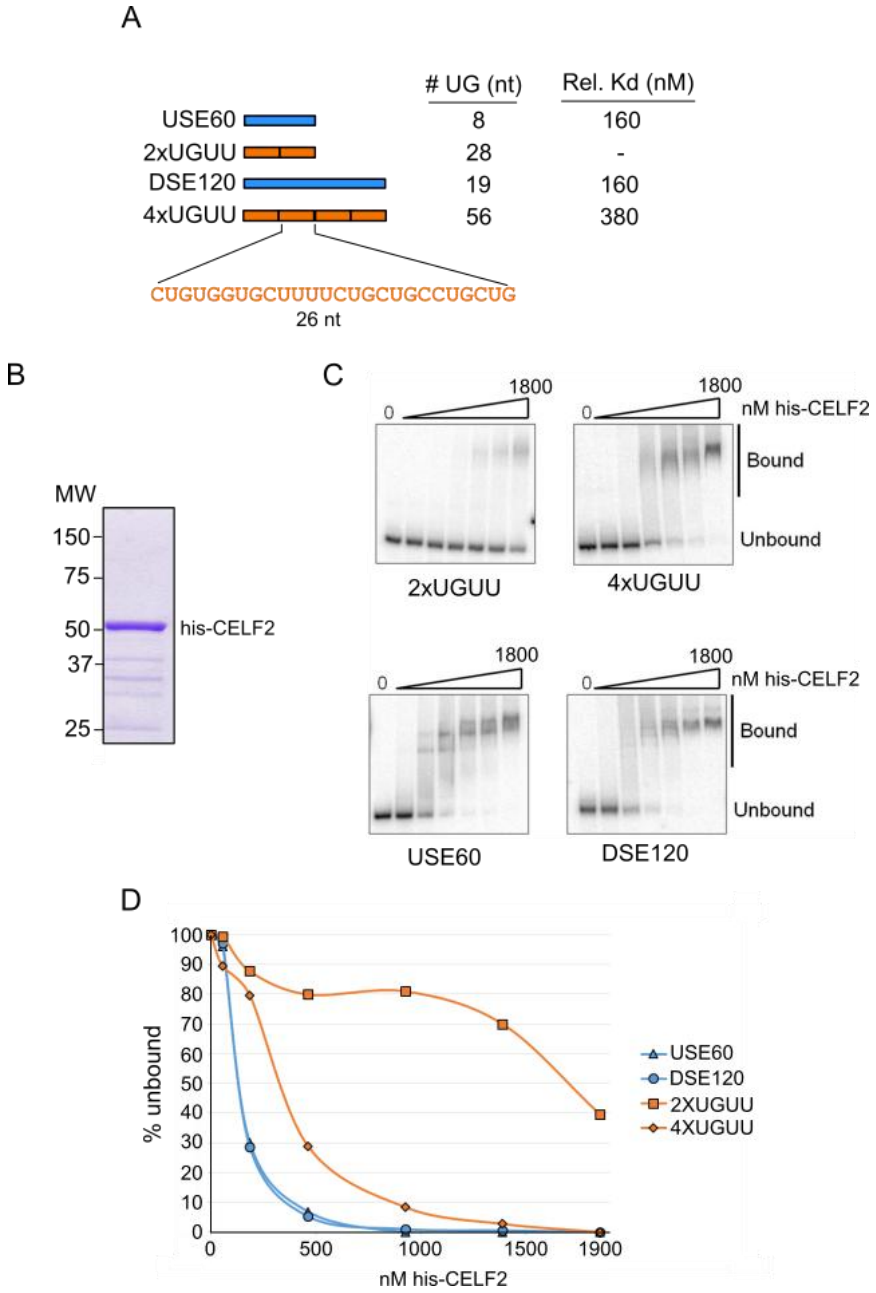


Figure 3.2: his-CELF2 binds USE60 and DSE120 with high affinity. (A) Schematic comparing the USE60 and DSE120 with high-affinity CELF2 binding sequences 2XUGUU and 4XUGUU. Numbers along side the table compare the number of UG di-nucleotides within each sequence and the relative Kds for his-CELF2 binding as obtained by Electro-Mobility Shift Assays (EMSAs). Long Blue box is DSE120, short blue box is USE60. One orange box corresponds to 1xUGUU which is 26 nucleotides (nt) in length. 2XUGUU and 4XUGUU are depicted in this figure. (B) Coomassie stained SDS-PAGE gel of purified his-CELF2. Molecular Weight (MW) migration markers in kilo-daltons (kDa) are on the left of the gel. (C) EMSAs with increasing amounts of his-CELF2 (in nano-molar (nM) concentration) with the USE60, DSE120, 2XUGUU and 4XUGUU visualized on native gels. Bound and Unbound RNA species are labeled to the right of the gels. (D) Graphical representation of %unbound from quantitation of bound and unbound species from (C).

confirm that the 50kDa band that increased upon stimulation was CELF2, I used a CELF2 antibody to IP CELF2 from the UV crosslinking reactions. Having confirmed that CELF2 binds the USE60-exon6-DSC-DSE120 sequence and that it's binding increases upon stimulation, I used constructs that had either the USE60 or the DSE120 replaced with heterologous sequence (alt60-exon6-DCS-DSE120 = ΔUSE60 and USE60-exon6-DCS-alt120 = ΔDSE120) to determine which regulatory element was being bound by CELF2 in each condition (Figure 3.3).

Figure 3.3 shows that in the unstimulated state ΔUSE60 is still capable of binding CELF2 but at slightly lower levels than WT (Figure 3.3 panel C). Importantly, the ΔUSE60 maintains the increase in CELF2 binding upon stimulation that is seen in the WT substrate. In unstimulated extract the ΔDSE120 also binds CELF2 close to WT levels. Strikingly however, the ΔDSE120 does not bind more CELF2 upon stimulation, especially when compared to WT and ΔUSE60. Although CELF2 is capable of binding both sequences, in the absence of the DSE120, CELF2 binds the USE60 but is incapable of any increased binding upon signal induction. However, in the absence of the USE60, CELF2 binds the DSE120 in unstimulated cells and can increase this binding upon stimulation. This suggests that in unstimulated cells, CELF2 binds to the USE60 and minimally to the DSE120 but the key signal-induced increase of CELF2 binding upon stimulation is localized to the DSE120.

Discussion:

The comparison of protein binding with and without the context of the exon highlights the complexity of RNA-protein interactions and the various factors that could influence them. Having more of the regulatory landscape of LEF2-E6 present during the experiment not only affects the pattern of binding of CELF2 but also is a powerful tool by which to accurately map where CELF2 binds in different cell states. Figure 3.3 convincingly shows that even though the USE60 and DSE120 are high affinity binders of CELF2 outside the context of LEF1-E6, within the context of the exon there is a preference for the DSE120 in stimulated cells. This is in contrast to the model from Mallory et al showing that there isn't an equal increase in CELF2 binding on both elements upon stimulation but a preferred increase on the DSE120 over the USE60 (Figure 3.3 panel D).

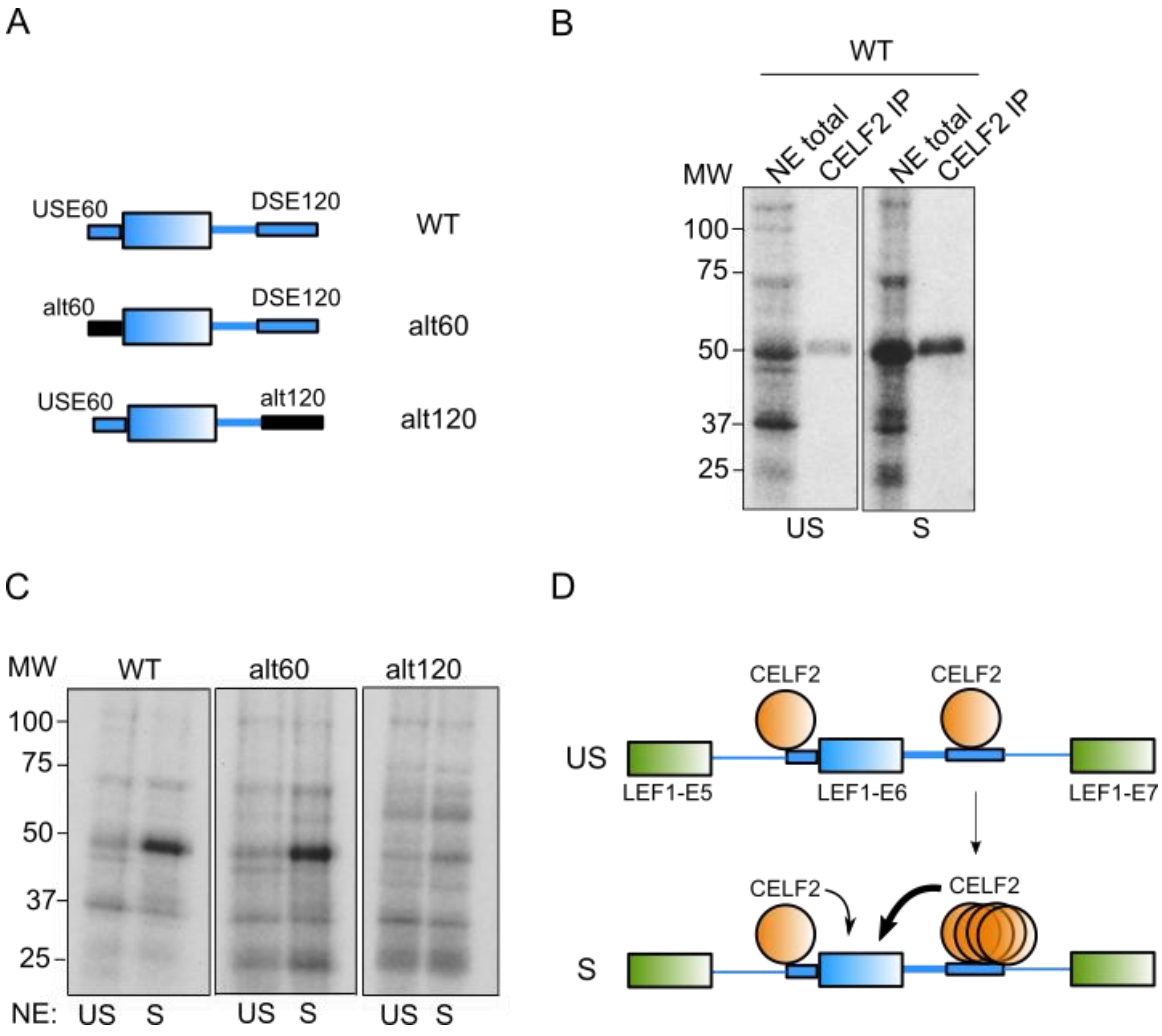


Figure 3.3: Increase of CELF2 binding upon stimulation occurs predominantly on the DSE120 not the USE60. (A) Schematic of LEF1 constructs used in UV crosslinking assay to map CELF2 binding on the USE60 and DSE120 in Unstimulated (US) and Stimulated (S) Nuclear Extract (NE). (B) UV crosslinking of the WT LEF1 construct in US and S NE coupled with IP for CELF2 confirming band at 50kDa that increases upon stimulation corresponds to CELF2. (C) Comparing UV crosslinking of WT, ΔUSE60 and ΔDSE120 in US and S NE showing requirement of DSE120 for increase of CELF2 binding upon simulation. (D) Model for LEF1-E6 regulation showing increased binding of CELF2 to the DSE120 in stimulated cell leading to activation of splicing. Arrow between the DSE120the site of increased CELF2 binding.

Additionally, it is important to note that within the context of LEF1-E6, CELF2 is capable of interacting with each element independent of the other as replacement of the USE60 did not interrupt CELF2's interaction with the DSE120 and vice versa. The fact that the USE60 and

DSE120 can bind independently is very unlike PTB's regulation of *csrc*'s N1 exon[79], where the loss of the downstream regulatory element results in a loss of PTB binding in the upstream element. In that case both sites are important to stabilize the interaction. That fact the USE60 and DSE120 are not both required to stabilize CELF2 binding is intriguing because, both elements are required to enhance LEF1-E6 inclusion upon stimulation[39]. In order to understand this mechanism, it is important to be able to correlate binding to function. Categorizing the USE60 and DSE120 as ISS's or ISE's, in the context of LEF1, would shed light on the significance of the change in binding upon stimulation.

CHAPTER 4: FUNCTION OF THE USE60 AND DSE120

Introduction:

One of the outstanding questions about the signal induced enhancement of LEF1-E6 is the specific functions of the USE60 and DSE120 i.e are they ISSs or ISEs. The fact that CELF2 binds preferentially downstream upon stimulation implies that the USE60 and DSE120 have differing functions, but it is also possible that they are both enhancing, with the DSE120 being a more robust activator of splicing. There aren't any instances in the literature where CELF2 binds on either side of an exon however there are many examples of CELF2 binding on either the upstream or downstream intron. Except for one instance involving one of the mutually exclusive exons SM and NM of α-actinin[59], every instance of CELF2 regulation of cassette exons that has been studied in molecular detail has shown that CELF2 binding upstream causes exon exclusion[57,58] while CELF2 binding downstream of an alternative exon causes inclusion[55,58,84].

To analyze the effects of the USE60 and DSE120 on exon splicing I used *in vitro* splicing assays with constructs that lacked one or both the elements. I then corroborated the *in vitro* splicing results with the exon inclusion levels obtained from the *in vivo* minigene assay. Clarifying what the functions of the USE60 and DSE120 are, in the context of LEF1, can help guide and interpret experiments that probe how CELF2 binding to these elements is regulated between unstimulated and stimulated cells. To correlate the functions of the USE60 and DSE120 with the regulation of CELF2 binding in the unstimulated and stimulated state, I used recombinant CELF2 expressed under each condition in RNA binding experiments. Together these experiments inform a model for how CELF2 regulates the enhancement of LEF1-E6 upon stimulation.

Results:

In order to determine whether the USE60 and DSE120 work in unison to cause exon inclusion or whether they follow the apparent positional dependent rules suggested by the literature, I used an *in vitro* splicing assay to monitor splicing in LEF1 RNA templates and compare the WT construct that contained both the USE60 and DSE120 (USE60/DSE120) with constructs in which the USE60 was deleted (Δ/DSE120) or the DSE120 was replaced (USE60/alt120). When compared

to the WT construct (USE60/DSE120) and a construct that lacks both elements(Δ/alt120), the RNAs with only the USE60 or DSE120 can shed some light on how the they function in the alternative splicing of LEF1-E6.

The *in vitro* splicing assay first involves the *in vitro* transcription of USE60/DSE120, Δ/DSE120, USE60/alt120 and Δ/alt120 RNA. The RNA is then incubated in nuclear extract to allow splicing to occur before using radio-labeled RT-PCR to amplify spliced and unspliced products. Splicing products are monitored on denaturing gels using autoradiography (Figure 4.1 panel A).

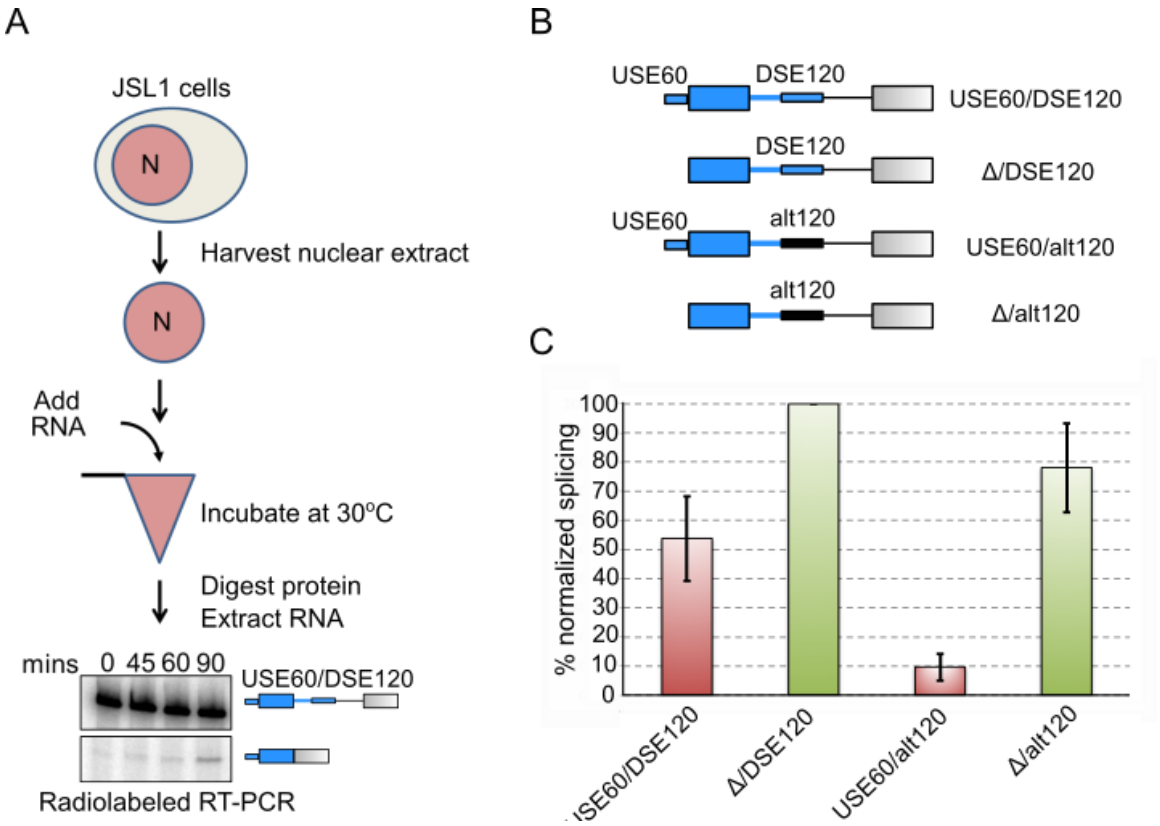


Figure 4.1: DSE120 activates splicing and the USE60 represses splicing *in vitro*. (A) Schematic showing the major experimental steps involved in conudcting *in vitro* splicng reactions. White circle is the cytoplasm of a cell and pink circle is the nucleus (N=nucleus). (B) Schematic showing two-exon constructs used in *in vitro* splicing reactions. (C) Graphical representation of results from *in vitro* splicing of constructs from (B). Splicing reactions were incubated for 90 minutes before analysis. Experiments were repeated atleast 3 times. Results were normalized to the % splicing of the Δ/DSE120 to account for variations in basal levels of splicing between experiments and facilitate comparison between them.

Figure 4.1 shows the results of the assay after incubation for 90 minutes in nuclear extract. Deleting the USE60 caused an increase in splicing of the downstream intron, showing that the USE60 represses splicing. However, replacing the DSE120 caused a decrease in splicing of the downstream intron, making the DSE120 an activator of splicing.

It is possible that the increase of splicing seen after the deletion of the USE60 is due to the absence of any sequence upstream of the exon. This could encourage intron definition on the downstream intron and lead to increased splicing. In order to ensure that the increase in splicing was not because of the lack of sequence upstream, I replaced the upstream sequence with heterologous sequence. Figure 4.2 shows that the alt60/DSE also yields an increase in splicing over the WT USE60/DSE120 construct, confirming the USE60 as a repressive sequence.

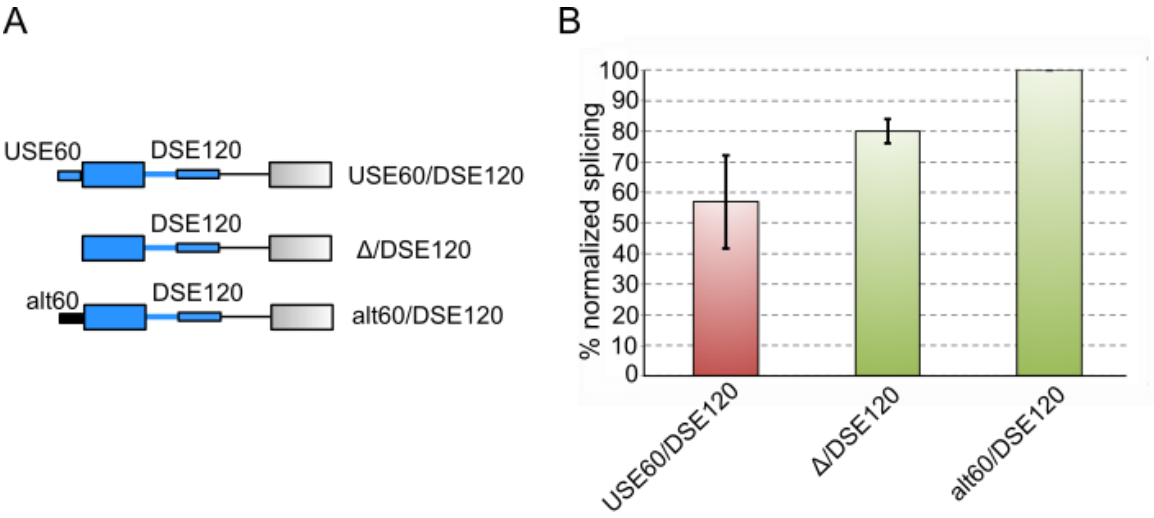


Figure 4.2: Deletion or replacement of the USE60 results in splicing activation. (A) Schematic showing two-exon constructs used in *in vitro* splicing reactions. (B) Graphical representation of results from *in vitro* splicing of constructs from (B). Splicing reactions were incubated for 90 minutes before analysis. Experiments were repeated atleast 3 times. Results were normalized to the % splicing of the alt60/DSE120 to account for variations in basal levels of splicing between experiments and facilitate comparison between them.

The roles of the USE60 as a splicing repressor and DSE120 as a splicing activator are corroborated by *in vivo* minigene results. LEF1-E6 is more included when the USE60 is replaced and is more excluded when the DSE120 is replaced (Figure 4.3). Considering that the signal

induced increase of binding of CELF2 is localized to the DSE120, the *in vitro* splicing results show that the reason this increase in binding results in an increase in LEF1-E6 inclusion is because the DSE120 is an activator of splicing. Since the USE60 is a splicing repressor, the effect of the USE60 dominates in unstimulated cells by the binding of CELF2 and upon stimulation this repression is relieved by increased CELF2 binding to the DSE120.

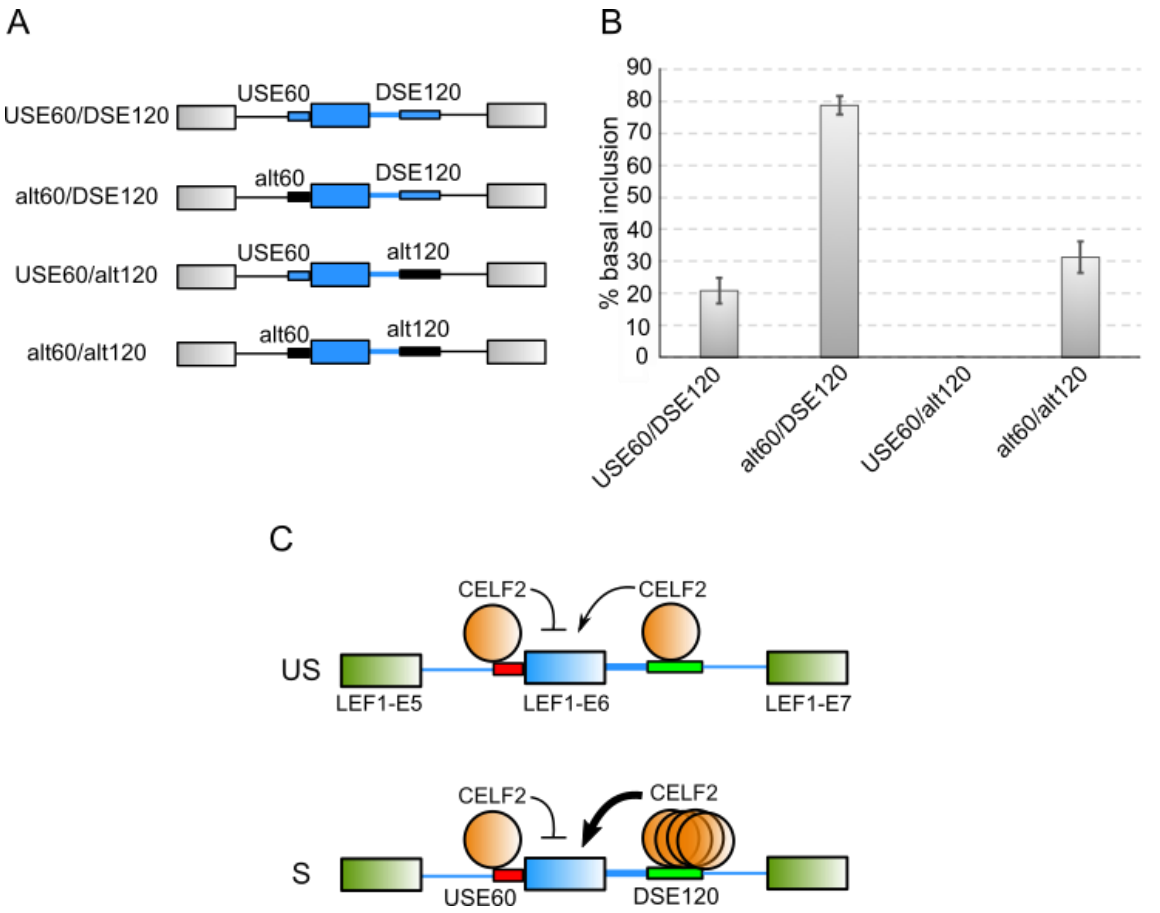


Figure 4.3: DSE120 activates splicing and the USE60 represses splicing *in vivo*. (A) Schematic of minigenes used to confirm *in vitro* splicing results from Figure 4.1 and 4.2 (B) Graphical representation of results from minigenes in (A) showing % basal inclusion of LEF1-E6. (C) Model for LEF1-E6 enhancement upon stimulation showing CELF2 binding ot the USE60 having a repressive and CELF2 binding to the DSE120 having an activating effect. Upon simulation, greater CELF2 binding to the activating DSE120 causes increase in LEF1-E6 inclusion. USE60 is coloured red indicating its role as an ISS and the DSE120 is coloured green indicating the DSE120s role as an ISE.

The differing roles that CELF2 plays upon binding to the USE60 versus the DSE120 is also apparent when CELF2 is significantly knocked down in unstimulated and stimulated JSL1 cells. Michael Mallory in the lab used a lenti-viral expression vector to cause very effective knockdown of CELF2 in both US and S JSL1 cells. In unstimulated cells, knockdown of CELF2 caused an increase in LEF1-E6 inclusion suggesting that CELF2 is functioning as a repressor in this condition. In stimulated cells, knockdown of CELF2 caused a decrease of LEF1-E6 inclusion suggesting that CELF2 serves as an activator in this condition. This is in line with the current model, wherein the effect of CELF2 binding the repressor USE60 is the dominant effect in unstimulated cells. Upon stimulation the effect of CELF2 binding the activator DSE120 is the dominant effect (Figure 4.4)

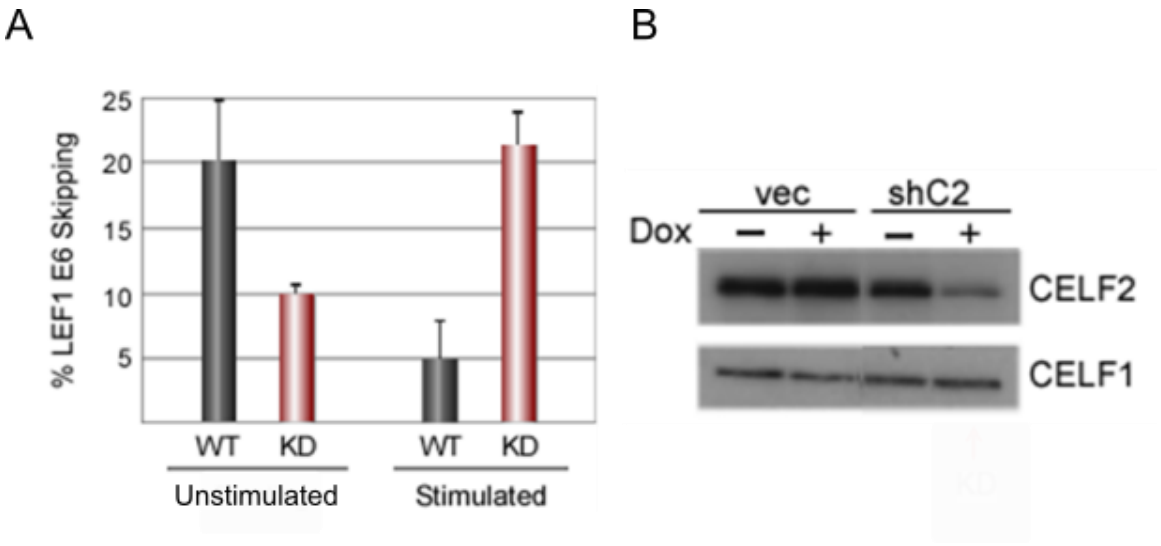


Figure 4.4: Knockdown of CELF2 in JSL1 cells has opposite effects on LEF1-E6 inclusion in unstimulated and stimulated cells. (A) Graphical representation of the effect of CELF2 knockdown on LEF1-E6 inclusion in unstimulated and stimulated cells. WT = Wild Type cells without lenti-viral knockdown of CELF2. KD = Cells with lenti-viral knockdown of CELF2. (B) Western blots of CELF1 and CELF2 on cell lysates with or without doxycycline (DOX) induced knockdown of CELF2 showing selective knockdown of CELF2 protein. Data and figure by Michael Mallory.

It is possible that the activating nature of the DSE120 is not solely due to the increased binding of CELF2 but due to some other protein interacting with the DSE120. In order to confirm that CELF2 binding downstream of LEF1-E6 leads to increased inclusion I created 4 minigenes that replaced the DSE120 with increasing amounts the high affinity SELEX sequences used in the EMSAs (1x-4xUGUU). The EMSAs have shown that CELF2 binding increases from the 2xUGUU

to the 4xUGUU sequences. Therefore, these constructs allow for precise control over how much CELF2 can bind downstream. LEF1-E6 inclusion can then be monitored under unstimulated and stimulated states (Figure 4.5).

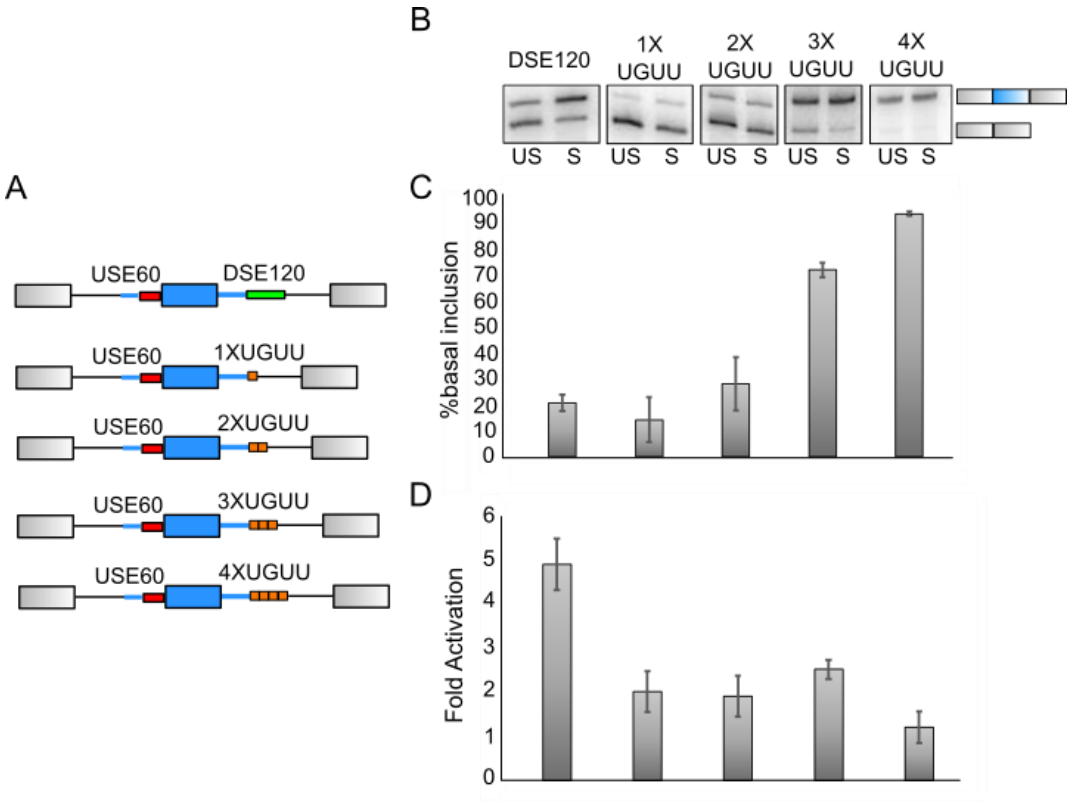


Figure 4.5: Systematically increasing CELF2 binding downstream of LEF1-E6 by step-wise increase of high affinity CELF2 binding sites increases basal inclusion levels but is not signal responsive. (A) Schematic of minigenes used to test effects of systematic increase of CELF2 binding downstream of LEF1-E6. (B) RT-PCR results visualized on a denaturing PAGE gel of splicing products of minigenes from (A) in Unstimulated (US) and Stimulated (S) cells. Schematics to the right of the gels indicate exon included and exon excluded products. (C) Graphical representation of results from (B) showing % basal inclusion of LEF1-E6 (US). Each bar corresponds to the gel above it. (D) Graphical representation of results from (B) showing Fold Activation (FA) of LEF1-E6 upon sitmulation. Each bar corresponds to the gel above it.

In *in vivo* minigene assays, if the amount of CELF2 binding downstream correlates with LEF1-6 inclusion, then inclusion should increase from the 1x to the 4xUGUU minigenes. If only a precise amount of CELF2 binding is required for inclusion, then inclusion levels should spike for

one of the UGUU minigenes and not the others. Finally, If CELF2 binding to the DSE120 is not the sole cause for LEF1-E6 inclusion, then there should be no correlation between inclusion levels and CELF2 binding. Figure 4.5 shows that increasing amounts of CELF2 binding downstream of LEF1-E6 results in increasing amounts of LEF1-E6 inclusion. Therefore, the signal induced enhancement of LEF1-E6 is solely due to the increase in CELF2 binding to the DSE120 downstream of exon 6.A striking result from Figure 4.5 however is that the 1x-4xUGUU sequences do not recapitulate the signal induced enhancement of exon 6 inclusion. The lack of signal responsiveness is not due to a lack of sensitivity caused by too much inclusion because the 1-3XUGUU minigenes are comfortably within the range to observe an increase. There is therefore something unique about the DSE120 in the context of LEF1-E6 that is capable of regulating the amount of CELF2 that binds to it unstimulated cells versus stimulated cells.

In order to probe whether there was a difference in the way CELF2 interacts with the USE60 and DSE120 in the unstimulated versus the stimulated state, I isolated protein from each condition. The first step towards isolating US and S-CELF2 was to stably express FLAG-tagged CELF2 in JSL1 cells. In order to obtain protein from both the unstimulated and stimulated conditions, I grew 30L of FLAG-CELF2 expressing JSL1 cells, and stimulated 15L with PMA. After 72 hours, both the unstimulated and stimulated cells were harvested. Since CELF2 is predominantly a nuclear protein[39], nuclear extract was separated from the harvest and subjected to a M2 FLAG affinity column to specifically pull out FLAG-tagged CELF2 (Figure 4.6).

To specifically probe how US and S-CELF2 interact with the USE60 and DSE120, I used an Electro-Mobility Shift Assay (EMSA). For this assay, the USE60 and DSE120 were radio-labelled and then incubated with increasing amounts of US-CELF2 or S-CELF2. The incubation allows RNA-protein complexes to form and then these radiolabeled RNA-protein complexes, are visualized in a non-denaturing acrylamide gel using autoradiography (Figure 4.6). Figure 4.6 shows that for the both the USE60 and DSE120, US-CELF binds in three distinct modes even at the highest CELF2 concentration. By sharp contrast, S-CELF2 binds very co-operatively to both the USE60 and DSE120, with the three species collapsing to one by the highest CELF2 concentration.

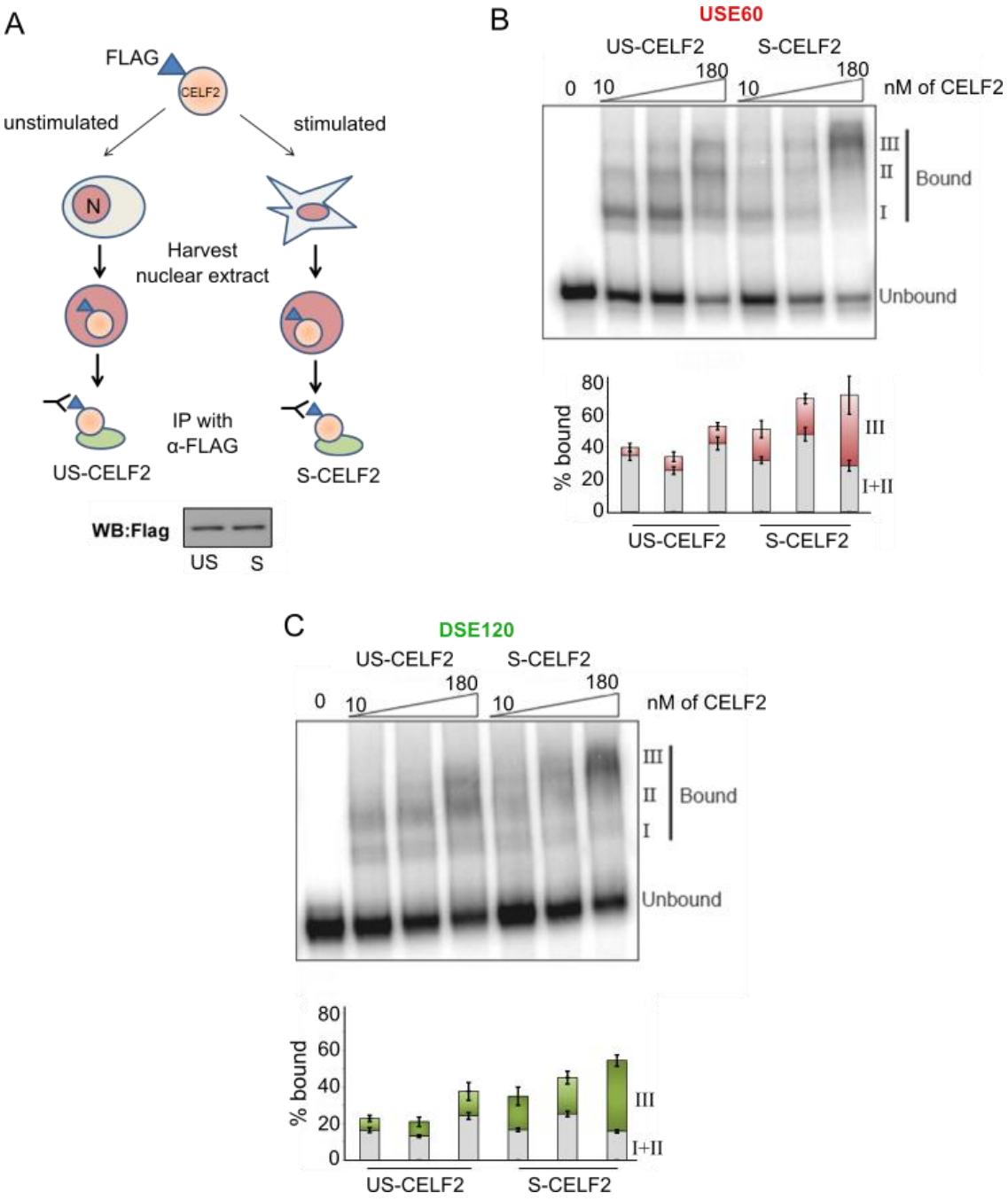


Figure 4.6: Differential binding of US-CELF2 and S-CELF2, purified in the absence of RNase, with USE60 and DSE120. (A) Schematic of strategy for FLAG-tagged CELF2 isolation from US and S JSL1 cells. Western blot (WB) using α-CELF2 antibody after SDS-PAGE on protein isolated is shown below the schematic. (B) EMSAs with increasing amounts of US-CELF2 and S-CELF2 interacting with USE60. Graph underneath the gel shows quatification of %bound species comapring bands III (red) compared to I+II (grey). Each bar corresponds to the lane above it. (C) EMSAs with increasing amounts of US-CELF2 and S-CELF2 interacting with DSE120. Graph underneath the gel shows quatification of %bound species comapring bands III (green) compared to I+II (grey). Each bar corresponds to the lane above it.

The EMSAs were repeated 3 times and the results quantified showing the cooperativity present in the binding of S-CELF2 to the RNA, that is absent from the US-CELF2 (Figure 4.6).

EMSAs were also repeated with US and S-CELF2 that was isolated in the presence of DNase and RNase. The removal of nucleic acids provides for a cleaner pull-down and several indirect associations that CELF2 made through RNA or DNA would be reduced (Figure 4.7). Figure 4.7 shows that the RNase treated US and S-CELF2 could not recapitulate the signal induced change in interaction observed from the non-RNase treated protein. However, since the yield from the pulldown was greatly improved, the EMSA was able to reach saturation and therefore apparent Kds were calculated for these interactions. The binding of US and S-CELF2 to the USE60 and DSE120 was compared to the high affinity CELF2 sequence 4xUGUU. Figure 4.7 shows that the US-CELF2 and S-CELF2 in the presence of RNase are very potent binders of the USE60 and DSE120 with a relative Kd of 15-20nM. This is a tighter interaction than the high affinity 4xUGUU sequence which they bind with a relative Kd of 65nM (data not shown). The interaction of US and S-CELF2 is non-cooperative and very similar to the binding of US-CELF2 in the absence of RNase to these constructs.

Discussion:

CELF2 binding upstream of an alternative exon being repressive to inclusion and downstream of an alternative exon being enhancing is a well-studied phenomenon in the literature[55,57–59,84]. cTNT exon 5[55] and NMDAR1 exon 21[58] have CELF2 regulatory elements downstream that contribute towards greater inclusion. NMDAR1 exon 5[58], CFTR exon 9[57], Tau exon 2[85] and CELF2's own exon 6 all have CELF2 regulatory elements upstream that contribute towards exon exclusion.

The literature therefore supports the model where CELF2 binding upstream is repressive and downstream is enhancing. The *in vitro* splicing results and the 1x-4xUGUU minigene results confirm that the DSE120 activates exon inclusion while the USE60 represses exon inclusion. Combining this with what was learnt about the pattern of CELF2 binding in unstimulated and stimulated cells, a clearer picture of the mechanism behind the enhancement of LEF1-E6 inclusion

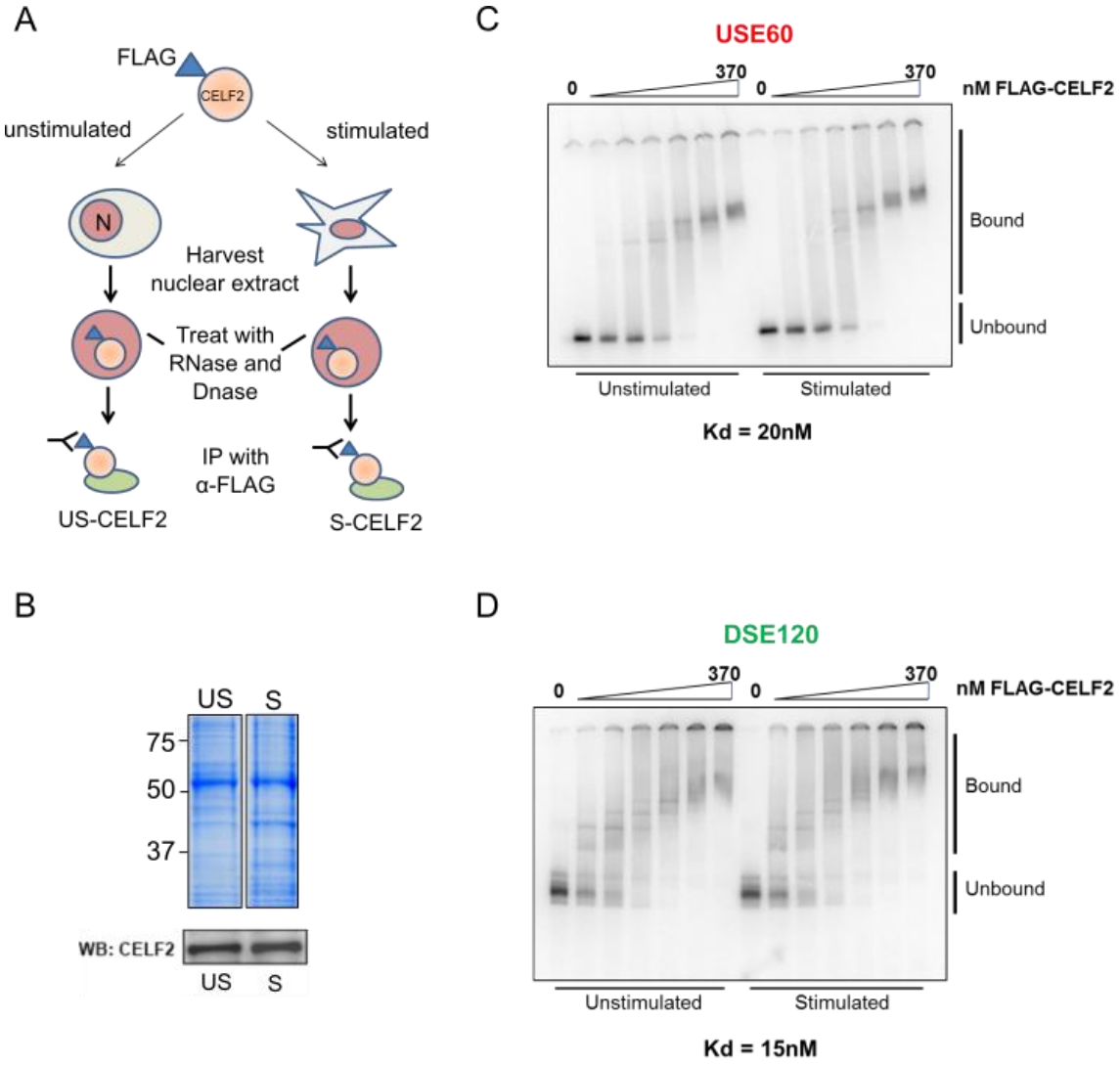


Figure 4.7: Differential binding of US-CELF2 and S-CELF2, purified in the presence of RNase, with USE60 and DSE120. (A) Schematic of strategy for FLAG-tagged CELF2 isolation from US and S JSL1 cells. (B) Coomassie stained poly-acrylamide gel with western blot (WB) using α-CELF2 antibody after SDS-PAGE on protein isolated using (A). (C) EMSAs visualied by autoradiography on native gels with increasing amounts of US-CELF2 and S-CELF2 interacting with USE60. Calculated Kd (dissociation constant) for the interaction is below the gel. (D) EMSAs with increasing amounts of US-CELF2 and S-CELF2 interacting with DSE120. Calculated Kd (dissociation constant) for the interaction is below the gel.

emerges. In unstimulated cells, the effects of CELF2 bound to the upstream repressor, USE60, are dominant. Upon stimulation, this repression is relieved by a bolus of CELF2 binding downstream to the activating DSE120. (Figure 4.3 panel C)

How is the binding of CELF2 so well regulated under these conditions? What is the mechanism behind the preference for the USE60 in unstimulated cells, and DSE120 in stimulated cells? One possibility is that there is an inherent difference in CELF2 itself between the two conditions. Perhaps CELF2's PTM landscape in unstimulated cells is inhibitory to binding the DSE120. Upon stimulation this PTM landscape changes now promoting CELF2 binding to the DSE120. A second possibility is that there are other splicing factors at play that actively keep CELF2 from binding the DSE120 in unstimulated cells. Upon stimulation either a down-regulation of this factor or the interference of another allows CELF2 to bind to the DSE120.

The EMSAs in the absence of RNase suggests that there is a difference in the way CELF2 pulldowns from US cells interact with the USE60 and DSE120 when compared to pulldowns from stimulated cells. This difference in interaction can be attributed to either an inherent difference in the CELF2 species between the two states or the presence of another splicing factor that was pulled down with CELF2. The loss of differential binding upon Rnase addition suggests that there was a factor present in the pulldown that lacked Rnase that was responsible for the change in interaction. This doesn't refute the necessity of a change in PTMs upon stimulation as it could be how CELF2 regulates its binding with this other unknown regulatory protein. The next two chapters of this thesis will discuss data that pertains to CELF2 PTMs in unstimulated and stimulated cells as well the possible involvement of other regulatory proteins in the regulation of LEF1-E6.

CHAPTER 5: POST-TRANSLATIONAL MODIFICATION OF CELF2 IN UNSTIMULATED AND STIMULATED T CELLS

Introduction

As seen in Chapter 4, one possible mechanism for how CELF2 binding the DSE120 increases from unstimulated to stimulated T cells is by a change in a post-translational modification (PTM). There are many examples of PTM's on splicing factors that modulate their ability to regulate splicing of their target exons. In some cases the modification directly affects the factors ability to bind its target mRNA[35,86–94]. SPF45's binding to *fas* exon 6 is regulated by 8 serine phosphorylations that are N-terminal to its RRM. Mutation of these serines to alanines resulted in greater binding to the *fas* substrate resulting in greater exon exclusion[94]. The kinase Chk1 phosphorylates serine 100 in RRM1 of the splicing factor HuR decreasing its affinity for the SIRT1 mRNA[91]. Similarly, PRMT1 methylates PSF and enhances its interaction with mRNA[86].

Modifications on splicing factors can also indirectly regulate splicing by regulating protein-protein interactions. A well-studied example of this from the Lynch Lab is the phosphorylation of PSF in unstimulated T cells. This modification causes it to be sequestered into a complex with the protein TRAP150. Upon stimulation, GSK3 activity is downregulated, which therefore reduces PSF phosphorylation. This leads to less PSF being sequestered by TRAP150, more being available to bind to its target mRNA CD45, and regulate splicing[35]. SMAR1's regulation of Sam68 is another great example of PTMs indirectly regulating alternative splicing. Acetylated Sam68 is required for the regulation of CD44 alternative exons. Smar1 in its unmodified state sequesters Sam68 into a complex with a deacetylase. Phosphorylation of SMAR1 leads to a disruption of protein-protein interaction with Sam68, allowing for the acetylation of Sam68 and its downstream regulation of CD44 alternative splicing[90]. PTMs can also indirectly regulate splicing by affecting the localization of splicing factors. The disruption of phosphorylation of SRSF1 by SRPK1 results in the translocation of SRSF1 out of the nucleus and thus making it unavailable to its target mRNA Rac1b exon3[92].

There is significant evidence in the literature that shows how PTMs can affect the function of the CELF family of proteins. CELF1 has two phosphorylations with described consequences in

the literature. In myoblasts, Ser28 of CELF1 can be phosphorylated by Akt which increases its interaction with Cyclin D1 mRNA[95]. Ser302 phosphorylation with cyclinD3 has also been shown to have effects on both CELF1's interaction with C/EBPβ mRNAs as well as the protein p21[95]. In the context of T cells, CELF1 is phosphorylated upon T cell activation which decreases its ability to bind GRE-containing mRNAs[96].

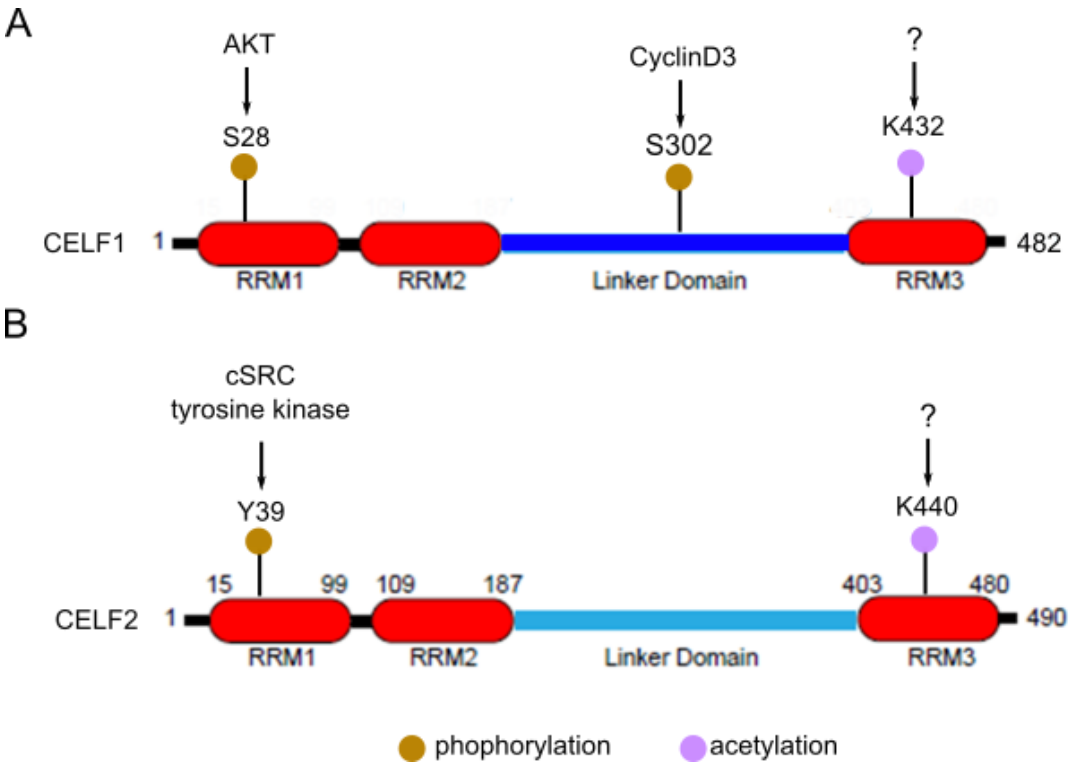


Figure 5.1: Post-translational modifications on CELF1 and CELF2 described in the literature. (A) Schematic of CELF1 showing annotated Post-Tranlational modifications (PTMs) and the modifying enzyme if known. Brown circles indicate phorylations and purple circles indicate acetylations. (B) Schematic of CELF1 showing annotated PTMs and the modifying enzyme if known. Brown circles indicate phorylations and purple circles indicate acetylations. Figure adapted from St. Louis 2013.

There is one well described instance of a PTM regulating CELF2 function. In rat smooth muscle cells, phosphorylation of Tyr39 in RRM1 of CELF2 by C-SRC tyrosine kinase enhances its interaction with COX2/PTGS2 mRNA[97]. The interaction stabilizes the mRNA preventing decay and is not a splicing regulatory event. However, it highlights the possibility of how PTMs on CELF2 can

modulate RNA binding. The only other modification that has been identified on CELF2 is the acetylation of Lys440[89,98]. It lies in RRM3 of CELF2 and therefore could have a role in modulating RNA binding as well (Figure 5.1).

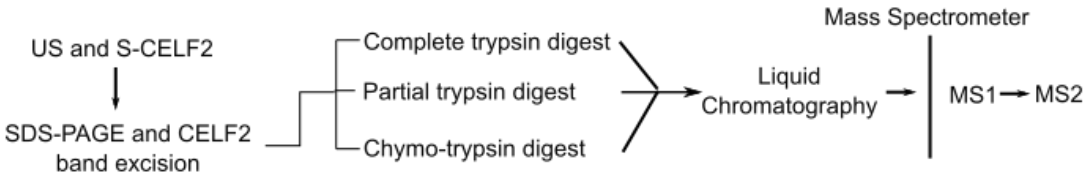
In order to better map CELF2's PTM landscape and ascertain whether T cell activation causes any changes in this landscape, I collaborated with the Wistar Institute's mass spectrometry core and used mass spectrometry to capture a global a view of CELF2's PTMs in the unstimulated and stimulated state.

Results:

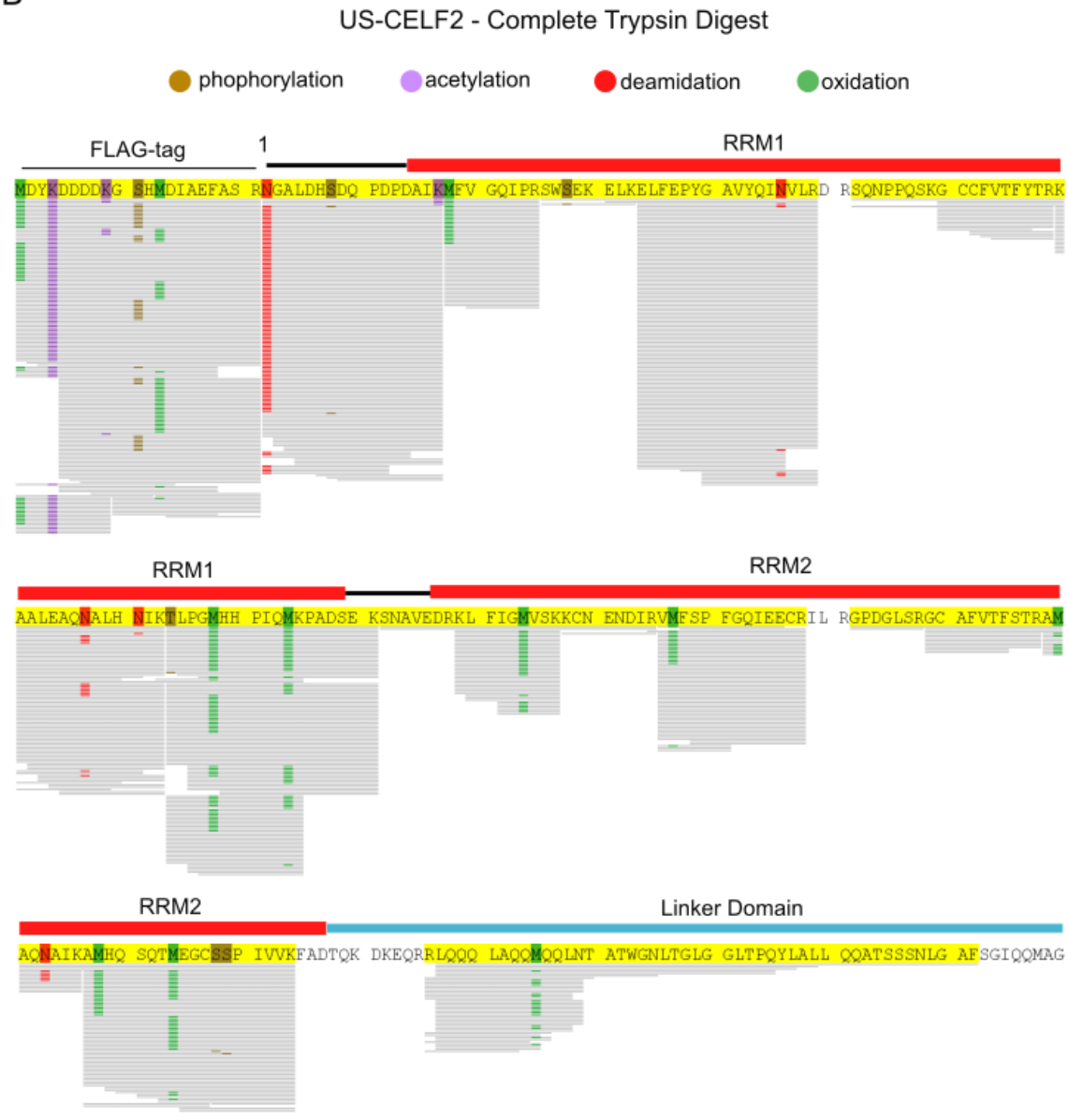
In order to investigate whether CELF2 had any significant changes in its PTM landscape upon stimulation, I used the FLAG-CELF2 protein isolated from US and S JSL1 cells according to the protocol outlined in Figure 4.7 panel A. Additionally, the purification was done in the presence of phosphatase and de-acetylase inhibitors in order to preserve as many of the native PTMs as possible (Figure 4.7 panel A)

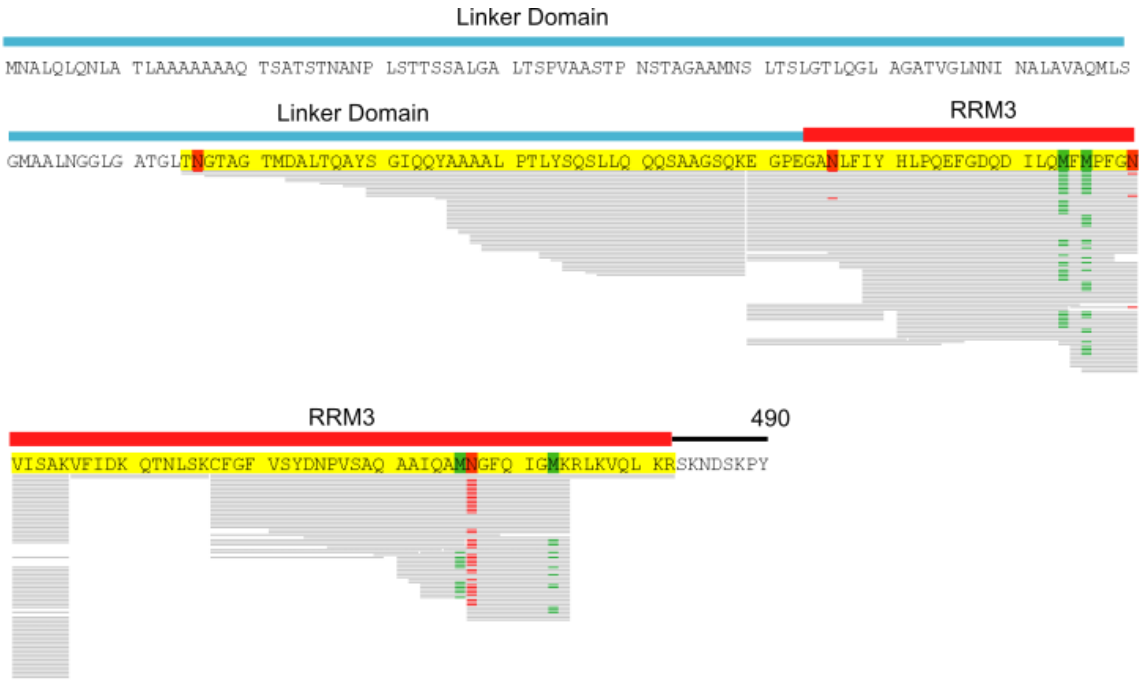
This protein was snap frozen and sent to the Mass spectrometry core at the Wistar Institute for further processing and analysis. At the institute, since the purification contained a significant portion of contaminating proteins (Figure 4.7), SDS-PAGE was performed and the band corresponding to CELF2 was specifically excised. This band was then subjected to three different proteolysis conditions to increase our chances of getting the best peptide coverage of the protein. The three conditions were complete Trypsin digest, partial trypsin digest and Chymo-trypsin digest. The resulting peptides were analyzed on their mass spectrometer. LC-MSMS was performed which collected both primary collision information (MS1) as well as secondary collision information(MSMS or MS2) in order to detect peptides with masses corresponding to any added PTMs (Figure 5.2, panel A). The institute mapped the peptides to CELF2's protein sequence and annotated any PTMs that were found.

A

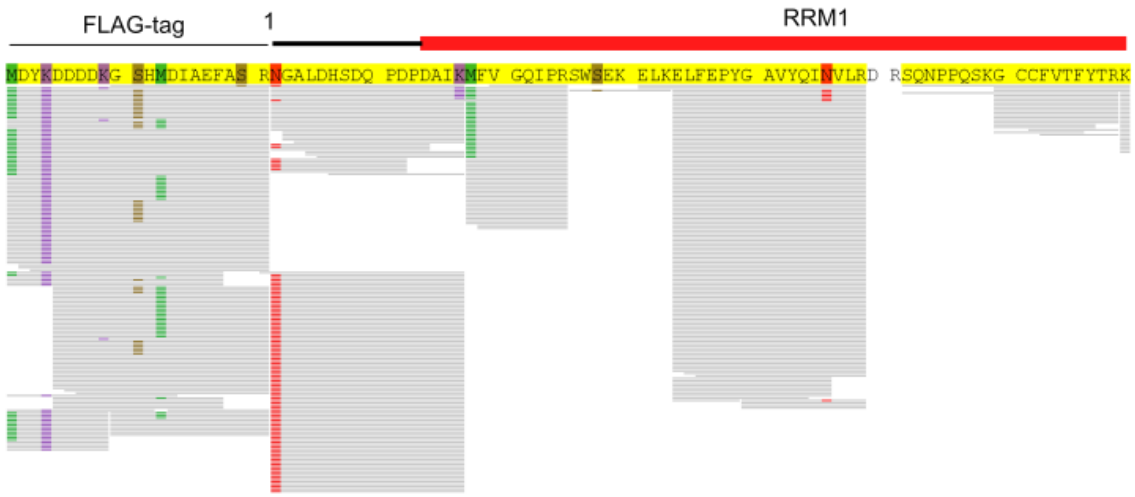

US and S-CELF2
SDS-PAGE and CELF2 band excision
Complete trypsin digest
Partial trypsin digest
Chymo-trypsin digest
Liquid Chromatography
Mass Spectrometer
MS1 → MS2

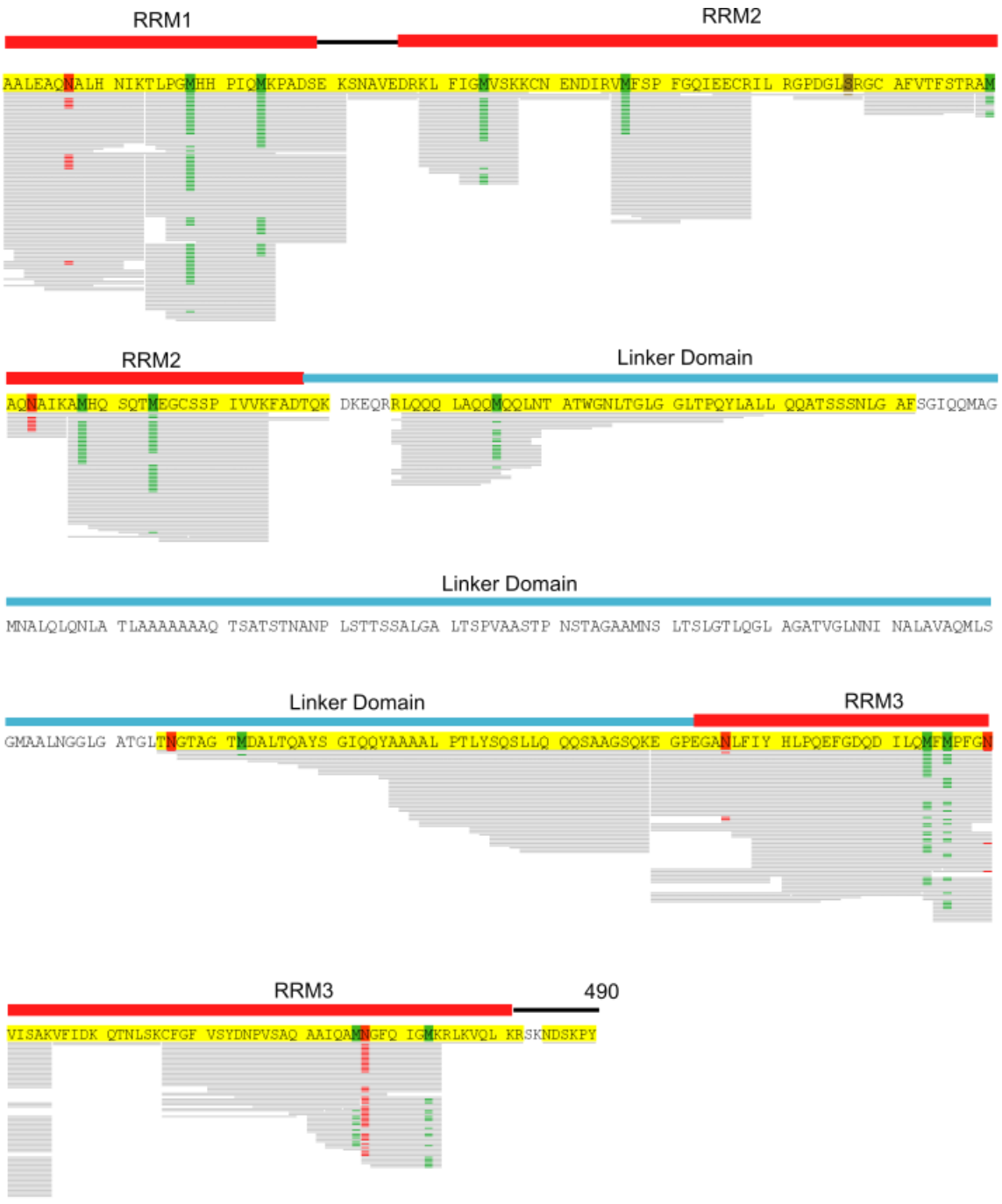
B


US-CELF2 - Complete Trypsin Digest
phophorylation
acetylation
deamidation
oxidation
FLAG-tag
1
RRM1
MDYKDDDDKG SHMDIAEFAS RNGALDHSDQ PDPDAIKMFV GQIPRSWSEK ELKELFEPYG AVYQINVLRD RSQNPPQSKG CCFVTFYTRK
RRM1
RRM2
AALEAQNALH NIKTLPGMHH PIQMKPADSE KSNAVEDRKL FIGMVSKKCN ENDIRVMFSP FGQIEECRIL RGPDGLSRGC AFVTFSTRAM
RRM2
Linker Domain
ACNAIKAMHQ SQTMEGCSSP IVVKFADTQK DKEQRRLQQQ LAQQMQQLNT ATWGNLTGLG GLTPQYLALL QQATSSSNLG AFSGIQQMAG

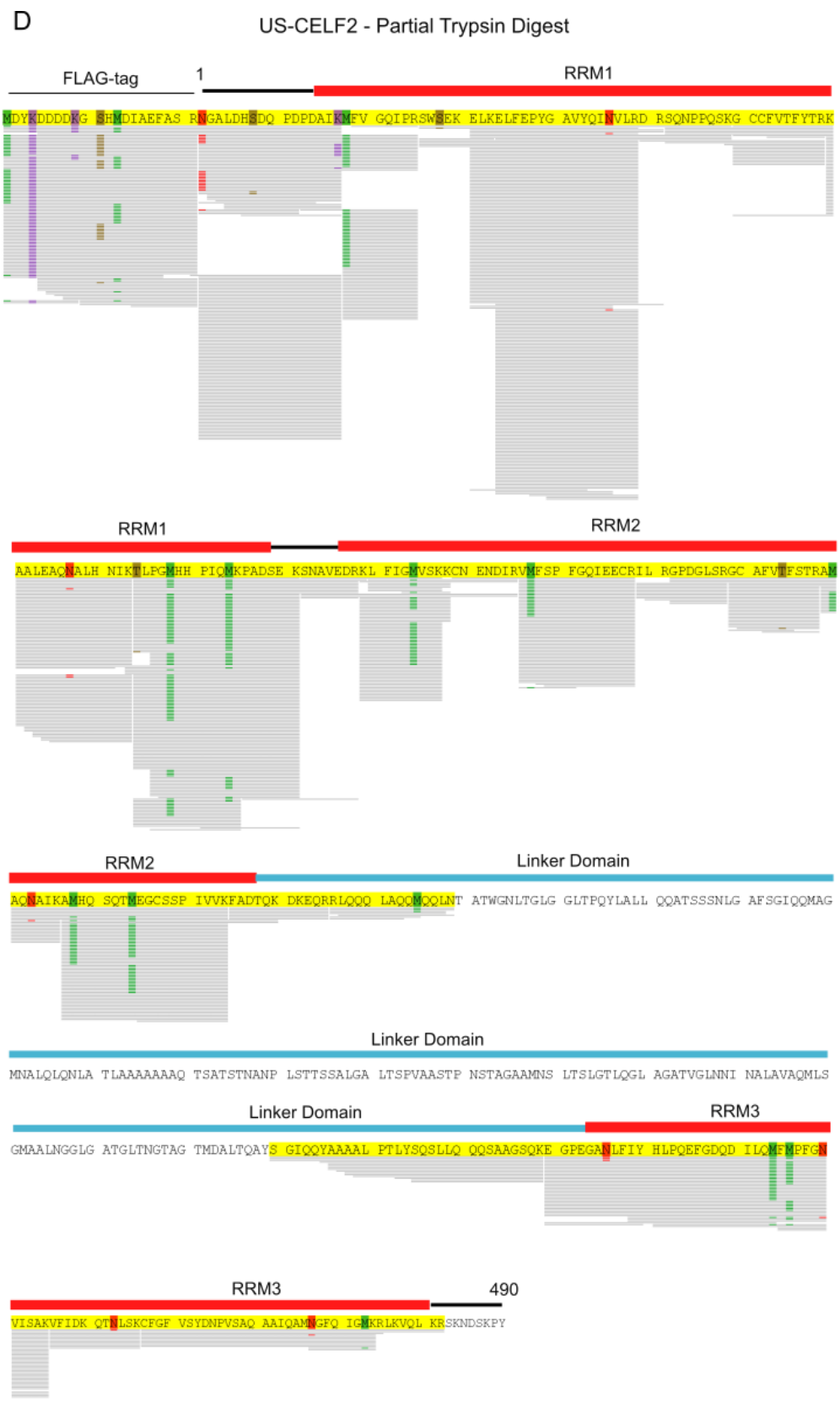

Linker Domain
MNALQLQNLA TLAAAAAAAQ TSATSTNANP LSTTSSALGA LTSPVAASTP NSTAGAAMNS LTSLGTLQGL AGATVGLNNI NALAVAQMLS
Linker Domain
RRM3
GMAALNGGLG ATGLTNGTAG TMDALTQAYS GIQQYAAAAL PTLYSQSLLQ QQSAAGSQKE GPEGANLFIY HLPQEFGDQD ILQMFMPFGN
RRM3
490
VISAKVFIDK QTNLSKCFGF VSYDNPVSAQ AAIQAMNGFQ IGMKRLKVQL KRSKNDSKPY

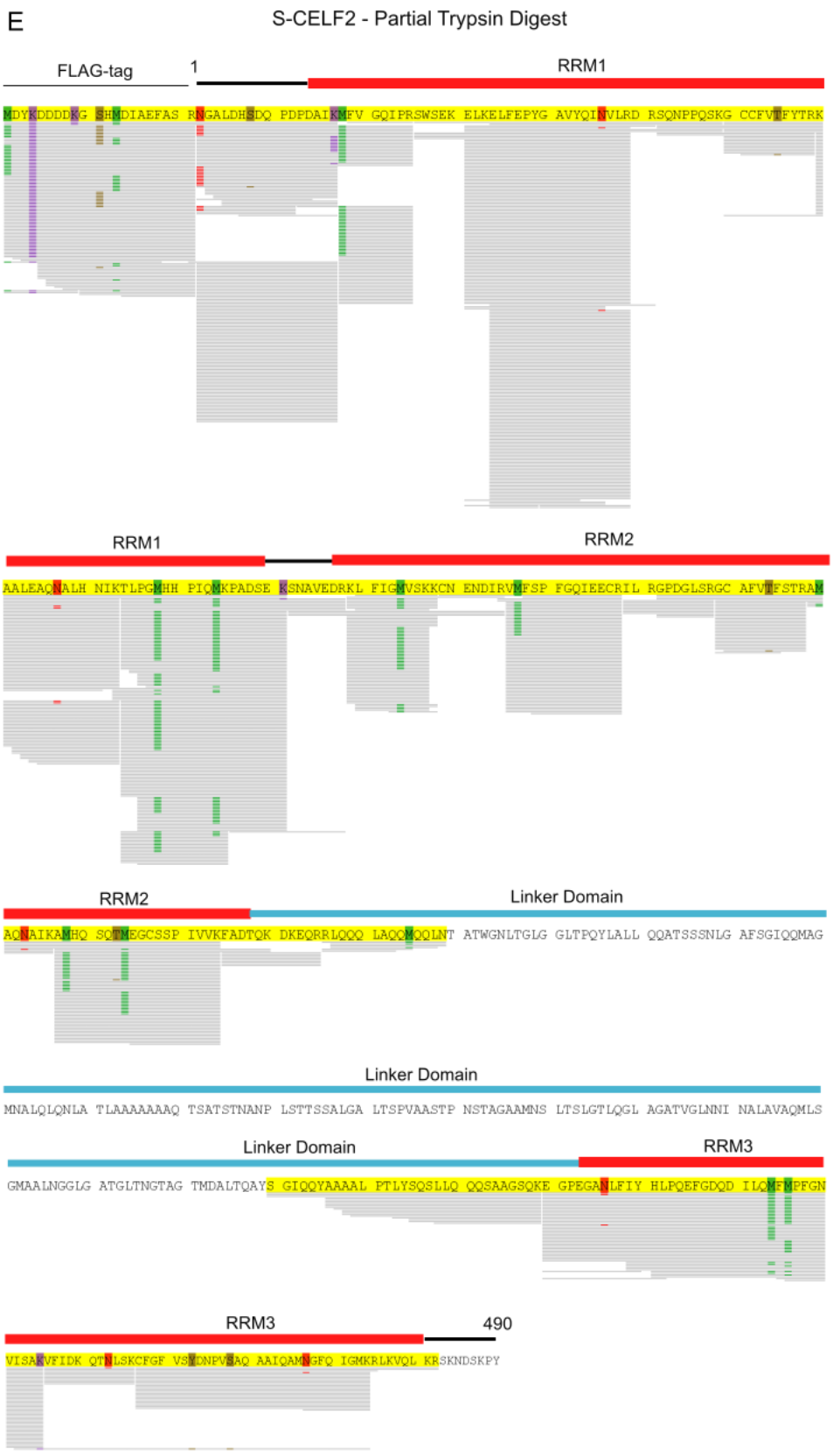
C


S-CELF2 - Complete Trypsin Digest
FLAG-tag
1
RRM1
MDYKDDDDKG SHMDIAEFAS RNGALDHSDQ PDPDAIKMFV GQIPRSWSEK ELKELFEPYG AVYQINVLRD RSQNPPQSKG CCFVTFYTRK


RRM1
RRM2
AALEAQNALH NIKTLPGMHH PIQMKPADSE KSNAVEDRKL FIGMVSKKCN ENDIRVMFSP FGQIEECRIL RGPDGLSRGC AFVTFSTRAM
RRM2
Linker Domain
AQNAIKAMHQ SQTMEGCSSP IVVKFADTQK DKEQRRLQQQ LAQQMQQLNT ATWGNLTGLG GLTPQYLALL QQATSSSNLG AFSGIQQMAG
Linker Domain
MNALQLQNLA TLAAAAAAAQ TSATSTNANP LSTTSSALGA LTSPVAASTP NSTAGAAMNS LTSLGTLQGL AGATVGLNNI NALAVAQMLS
Linker Domain
RRM3
GMAALNGGLG ATGLTNGTAG TMDALTQAYS GIQQYAAAAL PTLYSQSLLQ QQSAAGSQKE GPEGANLFIY HLPQEFGDQD ILQMFMPFGN
RRM3
490
VISAKVFIDK QTNLSKCFGF VSYDNPVSAQ AAIQAMNGFQ IGMKRLKVQL KRSKNDSKPY

D


US-CELF2 - Partial Trypsin Digest
FLAG-tag
1
RRM1
RRM1
RRM2
RRM2
Linker Domain
Linker Domain
Linker Domain
RRM3
RRM3
490


E
S-CELF2 - Partial Trypsin Digest
FLAG-tag
1
RRM1
RRM1
RRM2
RRM2
Linker Domain
Linker Domain
Linker Domain
RRM3
RRM3
490

F

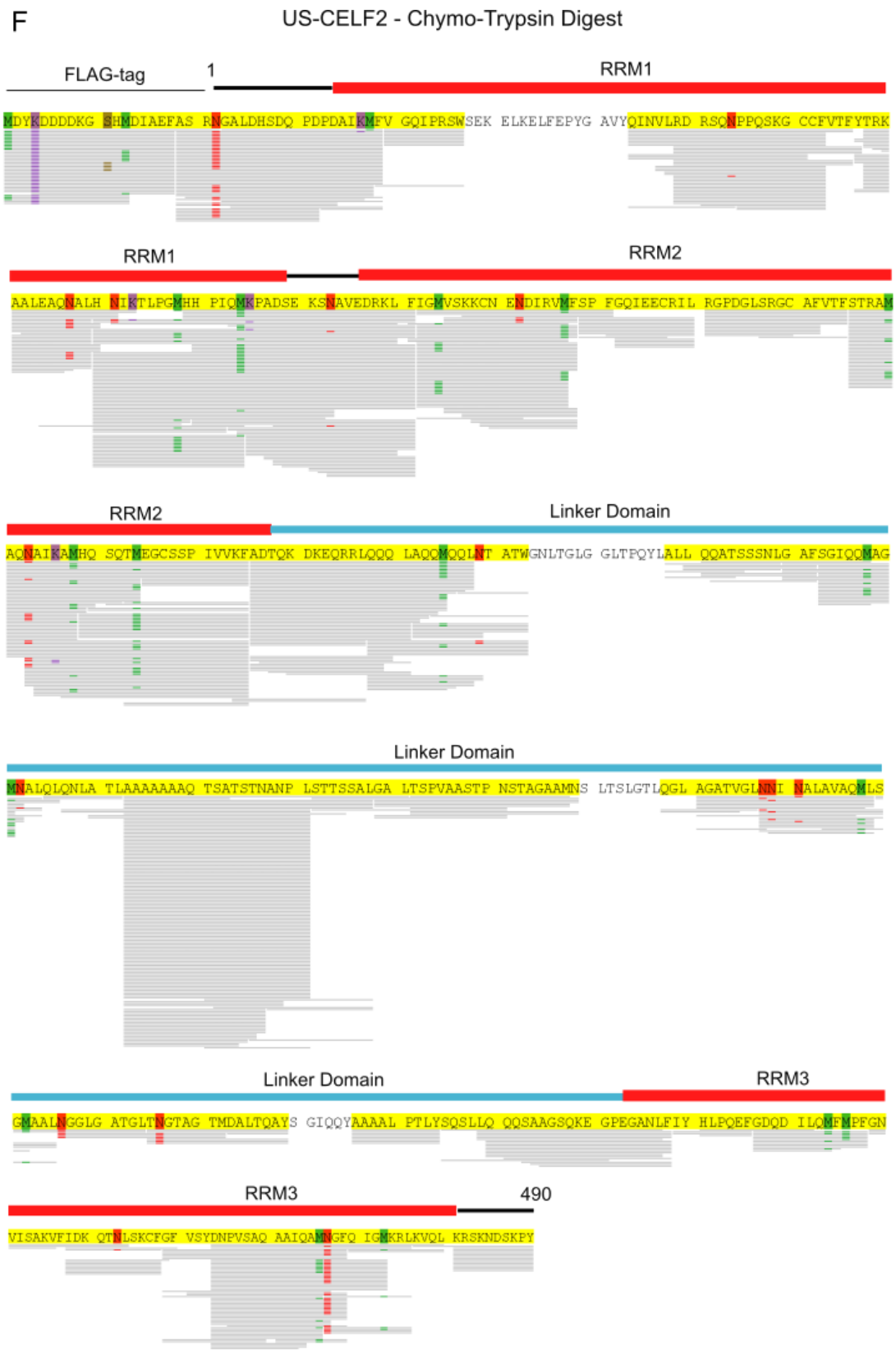

US-CELF2 - Chymo-Trypsin Digest
FLAG-tag
1
RRM1
RRM1
RRM2
RRM2
Linker Domain
Linker Domain
Linker Domain
RRM3
RRM3
490

Figure 5.2: Identification of PTMs on US and S-CELF2 using mass spectrometry. (A) Schematic showing major steps involved in obtaining peptides of US and S-CELF2 for mas spectrometry analysis. (B) - (G) Peptides (grey lines) identified by mass spectrometry for each digest condition (complete, partial and chymo-trypsin) for US and S-CELF2 with discovered PTMs marked in brown (phosphorylation), purple (acetylation), red (deamidation), green (oxidation). Peptides are overlayed with the protein sequence and domain they correspond to. Peptide maps courtesy the Wistar Institute's mass spectrometry core.

Figure 5.2 shows the peptides that were recovered from the mass spectrometry analyses of each digest condition, mapped against a schematic of the protein. As you can see, all three conditions provided great coverage of the three RRMs. The chymo-trypsin digest was the only condition that provided any peptides in the linker region of the protein, however there were no PTMs identified amongst them.

Analysis of the number of modified peptides compared to unmodified peptides in a small section of the N-terminus of US-CELF2 revealed that only a small percentage of the peptides have the modification (Figure 5.3). When I analyzed all the sites where PTMs were present, I found that to be the case in all instances. This suggests that only a small percentage of CELF2 was post-translationally modified. I did not consider deamidations or oxidations in this analysis. In order to isolate whether the small number of modified peptides was an accurate representation of the purified FLAG-CELF2 or a consequential loss in the mass spectrometer, I performed 2D gel electrophoresis on the purified unstimulated and stimulated protein. This procedure is used for the identification of the pH distribution of species of the same molecular weight. A protein without many PTMs will run close to its pI while modifications tend to cause a shift to a more acidic pH.

CELF2 has a pI of 8.6 and so any modified version of CELF2 would be expected to run at a more acidic pH. The 2-D gels were probed via western blot with anti-Flag and anti-CELF2 antibody to get a specific pH profile for CELF2. Figure 5.3 shows that the majority of the CELF2 protein runs ~8.5, which is close to the protein's pI. However, there is a small population of modified protein that runs at ~pH 2-3, very akin to the mass spectrometry results. As a control for the FLAG-CELF2 gel I also ran unstimulated and stimulated nuclear extract to check whether the purified protein was representative of the protein extracted from the cell. This also ensures that there weren't any drastic changes that resulted in a loss of PTMs during the FLAG purification process. As shown in Figure 5.3, CELF2 in nuclear extract also runs mostly around pH 8.6 with a small population running at a lower pH (Figure 5.3).

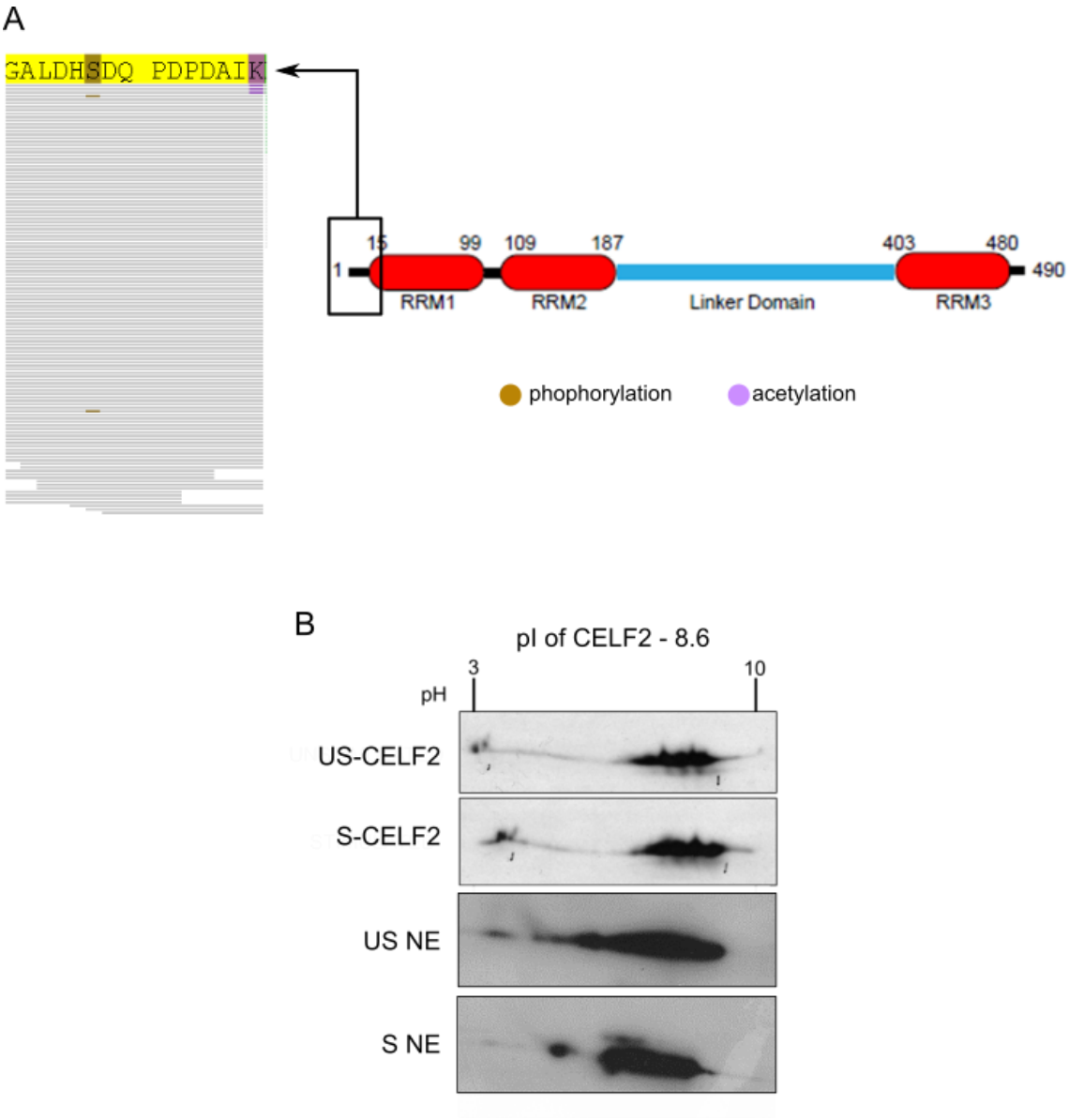
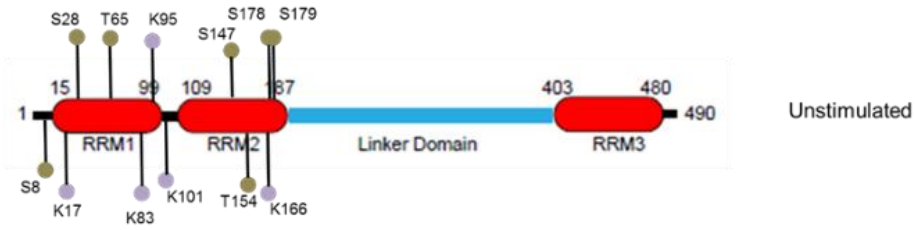


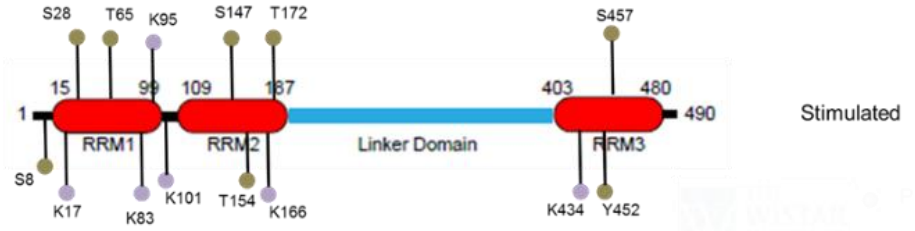
Figure 5.3: The population of modified FLAG-CELF2 isolated from US and S cells as well as endogenous CELF2 in US and S JSL1 nuclear extract is small. (A) Section of modified peptides from the N-terminus of FLAG-CELF2 showing paucity of modified peptides. Peptide map courtesy of the Wistar Institute mass spectrometry core. (B) 2-dimensional gel electrophoresis of US and S FLAG-CELF2 and US and S nuclear extract (NE). pH markers are above the gels.

This data confirmed that the mass spectrometry results were an accurate representation of the source material being provided. Focusing in on the small percentage of peptides that were modified, I mapped out where the modifications were on the protein (Figure 5.4). The modifications mostly lie in the RRMs with 5 on RRM1, 5 on RRM2 and 3 on RRM3. There was also 1 in the N-terminus of the protein and 1 in the sequence connecting RRM1 and RRM2. There were no PTMs found in the linker domain but this most likely because of the low peptide of coverage in this region (Figure 5.4).

A



B



C

Modification	Area (US)	Area (S)	US/S
S8	5.38×10^7	1.62×10^7	3.3
S28	1.29×10^8	1.21×10^8	1.1
K95	4.01×10^8	3.00×10^8	1.3
S147	2.17×10^9	3.01×10^9	0.7

Figure 5.4: Post-translational modifications of US and S-CELF2. (A) Schematic of US-CELF2 showing phosphorylation and acetylation sites discovered in Figure 5.2. (B) Schematic of S-CELF2 showing phosphorylation and acetylation sites discovered in Figure 5.2. (C) Calculation of fold change in S8, S28, K95 and S147 by diving the area under the mass peak in US compared to S-CELF2.

Since we were interested in changes in PTMs, I compared the intensity of the MS peptide peaks for identified PTMs in unstimulated and stimulated cells (Data from the Wistar institute). The intensity of the peaks can tell us whether there is more or less of that particular modification in each

condition. The modifications that had detectable MS1 peaks in both conditions were used for this analysis. I did not consider peaks that were completely absent in one condition versus the other as the reason for the absence is most likely due to a difficulty in detection in the mass spectrometer and not a physiological reality.

The fold change in intensities between US and S peaks was calculated by dividing the unstimulated peak area (Area (US)) by the stimulated peak area (Area (S)). Figure 5.4 panel C shows there were no significant changes in the intensity of the MS peaks between the two conditions.

Discussion:

Post-translational modifications on splicing factors can play a crucial part in the regulation of alternative splicing. In order to determine whether PTMs are important the function of CELF2, I collaborated with the Wistar Institute and used mass spectrometry to probe CELF2 PTM landscape in US and S cells. The small size of the population of modified peptides was a surprising find. It could indicate that CELF2 is not largely regulated via its PTMs in JSL1 cells. It is important to keep in mind that there could be a loss of PTMs after harvesting the cells. Since the 2D gel electrophoresis profile of unstimulated and stimulated nuclear extract looked very similar to the purified protein, if there was any loss of PTMs it was happening during the isolation of nuclear extract (Figure 5.3). However, the nuclear extract used for the 2D gel analysis is splicing competent and was used to recapitulate LEF1-E6 regulation in *in vitro* splicing experiments. This suggests that although less efficient, the levels of CELF2 and CELF2 PTMs are sufficient for splicing regulation. The small population size of modified CELF2 does not negate the possibility of PTMs being important for the regulation of LEF1-E6. There could be a mechanism by which the modified CELF2 binds to the USE60 and DSE120, either because of a greater affinity for these sequences that the unmodified protein or via interaction with other splicing factors.

Of the PTMs that were identified, 83% of them were newly annotated modifications, not mentioned in the literature to date. We were also able to pick up the two annotated modifications in our experiments adding more credence to our data. The number of modifications in the RRMs is

not surprising as most of the annotated and studied examples of PTM in splicing lie in the RRMs. In depth analysis of the RRM's PTMs however showed that there is no real change in their level upon T cell stimulation. The largest change observed is on S8 at a fold change of 3.2 (Figure 5.4). The lack of dramatic change in CELF2's PTMSs suggests that its preferential binding of the USE60 in unstimulated cells and DSE120 in stimulated cells is not caused by regulating PTMs on its RRMs.

An unfortunate drawback of this experiment is the lack of information in the linker domain of CELF2 that connects RRM2 and 3. A study by Singh et al showed that the linker domain is very important to CELF2's role as a regulator of splicing. The authors show that CELF2 is capable of activating exon inclusion via either RRM1 & RRM2 or via RRM3. However the caveat to both modes of regulation is that part of the linker domain is required – 70 residues adjacent to RRM1 & 2 and the 119 residues adjacent to RRM3. This suggests that this region is an important part CELF2's function as a regulator of splicing[99]. The linker domain has large segments that are rich in serines and threonines making them prime candidates for PTMs. Optimizing a proteolytic regime that allows for effective analysis of the linker region could shed light on why this region has been shown to be indispensable for CELF2 function as a splicing factor.

These experiments have shown that most of the CELF2 in JSL1 nuclear extract is unmodified. Of the small population that is modified, I discovered 12 previously undescribed PTMs on CELF2's RRMs with the help of the Wistar Institute's mass spectrometry core. This can be a guide for future structural and functional studies of CELF2. However, with the lack of information for the linker domain of CELF2, a definitive statement on whether there is a change in CELF2s PTM landscape cannot be made. Further efforts to study the PTMs of the linker domain will need to be made before this hypothesis is declared to not contribute to regulation of LEF1-E6.

CHAPTER 6: COMBINATORIAL CONTROL OF LEF1-E6 ALTERNATIVE SPLICING

Introduction:

Recent global studies of alternative splicing that probe the binding sites of splicing factors (called CLIP-seq or Cross Linking Immuno-Precipitation) are revealing an immense amount of overlap between various splicing factors and the alternative exons they bind around. A study that highlights this phenomenon was done by Huelga et al where they performed CLIP-seq on several hnRNP proteins (A1, A2/B1, F, M and U) and found that the majority alternative exons regulated by hnRNPs are regulated by 2 or more of the tested proteins. A similar study performed with the protein PTB and Quaking have shown that 25% of the exons regulated during myogenesis are under the control of both these proteins[100]. These global studies correlate with studies of individual splicing events, whose alternative exons are almost always are under the influence of multiple splicing factors.

One of the best examples of combinatorial splicing control comes from CD45 exon4, an alternative exon studied by the Lynch Lab. Here, exon 4 is repressed upon T cell stimulation and this repression is due to a ESS in exon 4 titled, quite simply, ESS1[34]. The lab has gone on to show that hnRNPL binds to this exon and causes repression by recruiting hnRNPA1 and together "holding" U1snRNP in conformation that makes spliceosome assembly more difficult[17,101]. This exon also relies on two other proteins, hnRNPLL and PSF that can enhance this repression upon T cell stimulation[33,35,74,102].

Other examples include the regulation of α-tropomyosin exon 2 by the splicing factors 9G8 and hnRNP F/H[103], DMD exon 39 by hnRNPA1, hnRNP A2/B1 and FUBP1[104], β-tropomyosin exon 6B by hnRNPA1, SC35 and ASF/SF2, FGFR1 exon IIIc by hnRNPF/H and Fox2[105] and HIV tev-specific exon 6D by SC35 and hnRNP H[106] to name a few. The CELF family of proteins can also regulate splicing in collaboration with other factors. The most prevalent family that collaborates with CELF is the MBNL family of factors. Together they have been shown to regulate cTNT exon 5, IR exon 11 and CLCN1 exon 7a[107,108]. More globally, these two families have been shown have antagonistic effects on exon fate in the developing skeletal and heart muscle[109]. CELF2 more

specifically also collaborates with PTB as in the case of cTNT exon 5[55] and α-actinin's NM and SM exons[59].

However there is less data on how CELF2 regulates signal induced alternative splicing and what other factors, if any, collaborate with CELF2 in a signal induced fashion. I have established so far that CELF2 represses exon 6 inclusion in unstimulated cells by predominantly binding upstream of the alternative exon. Upon T cell stimulation, CELF2 activates exon 6 inclusion by predominantly binding downstream of the alternative exon. Considering CELF2 ability to collaborate with other splicing factors, it is possible that where CELF2 binds around LEF1-E6 is regulated by a yet to be determined splicing factor. In order to identify the proteins that bind around LEF1-E6, I UV cross-linked radiolabeled USE60-exon6-DSE120 (WT) in US and S nuclear extract. Some of the bound proteins were identified by immunoprecipitation and their involvement in the signal induced enhancement of LEF1-E6 was assessed by knock down of the protein in JSL1 cells and monitoring levels of exon 6 inclusion.

Results:

UV crosslinking of proteins to radiolabeled RNA is a very powerful tool by which to narrow down which splicing factors bind around a region of interest. In this case, radiolabeled USE60-exon6-DCE-DSE120 RNA was incubated in US and S nuclear extract to allow proteins to bind. The RNA was then digested and the proteins analyzed on a SDS-PAGE gel. Proteins that were directly bound to the radiolabeled substrate appear upon phospho-imaging by virtue of being covalently linked to radio-labeled nucleotides. Figure 6.1 shows, 7 proteins, including CELF2 bind in and around LEF-E6 (figure 6.1 panel A).

In order to hone in on the proteins that just bind to the USE60 and DSE120, I performed the same assay with several deletion mutants of USE60-exon6-DCE-DSE120 (WT). The construct that replaces the exon with heterologous sequence, USE60-het-DCE-DSE120 (∆exon), didn't significantly change the binding pattern as compared to WT (Figure 6.1, panel B and C). Basal levels of CELF2 binding increase however suggesting that the exon harboured a splicing repressor. Signal induced increase of CELF2 however is unaffected. Upon replacement of the DCS in the

USE60-het-het-DSE120 construct (ΔDCS), we see a further increase in basal levels of CELF2 suggesting that the DCS also contained a basal splicing repressor while keeping signal induced increase in CELF2 binding intact (Figure 6.1, panel B and C).

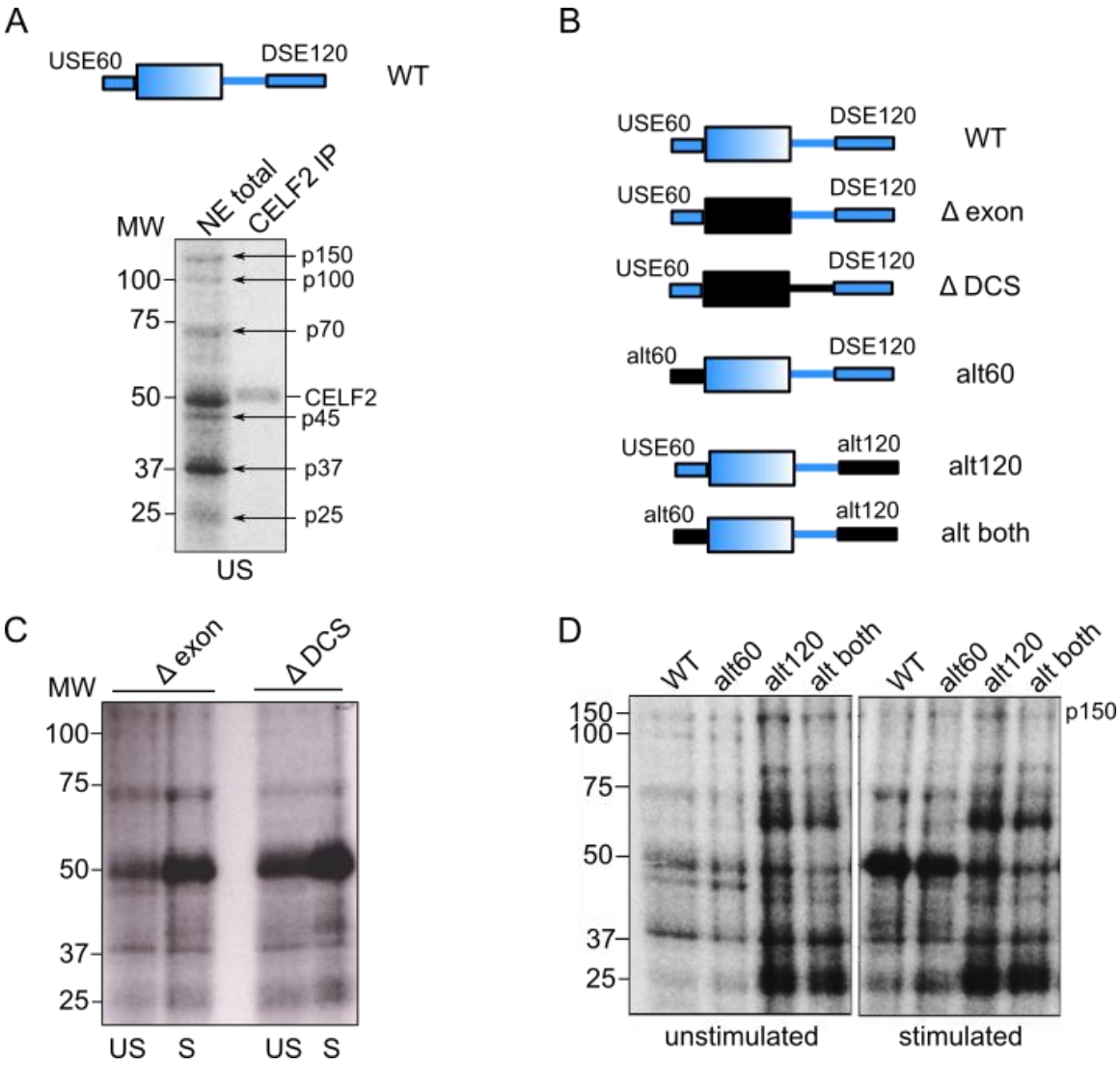


Figure 6.1: Proteins that bind the region of LEF1-E6. (A) UV crosslinking of the WT LEF1 construct in US and S NE coupled with IP for CELF2. Other prominent proteins that bind the RNA are labeled. (B) Schematic of LEF1 constructs used in UV crosslinking assay to identify proteins whose binding to the region around LEF1-E6 is sensitive to PMA stimulation or USE60 or DSE120 replacement. (C) UV crosslinking of body labeled Δexon and ΔDCS in US and S NE. (D) UV crosslinking of body labeled WT, alt60, alt120 and alt both in US and S NE. p150 is labeled as the only protein whose binding is insensitive to PMA stimulation or USE60 or DSE120 replacement.

In order to discover any protein that specifically bound to the DSE120 or the USE60, I also performed the UV crosslinking with USE60-exon6-DCE-het (alt120), het-exon6-DCE-DSE120 (alt60) and het-exon6-DCE-het (alt both). In the alt120 substrate (Figure 6.1, panel D), the binding of p100 is lost, suggesting that that protein binds specifically to the DSE120. p25, also increases upon replacement of the DSE120, suggesting there is element within that sequence that is inhibitory to p25 binding. Additionally, there is an extra protein that binds around 60 kDa. It is very possible that the heterologous sequence used to replace the DSE120 has binding sites of its own, which would explain the extra bands. However, as discussed in Chapter 3, the biggest change in the alt120 construct is the lack of increase in CELF2 binding upon stimulation.

The alt60 construct sees little to no change in any bound proteins highlighting the importance of the DSE120 in this regulation. Whereas, the alt both substrate looks almost identical to the alt120 substrate, also highlighting that most of the change that occur in binding occurs at the DSE120. One thing to note is that many of these protein bind substrate irrespective of what elements are replaced. It is most likely that these proteins are part of the spliceosomal machinery that are recruited to a functional 5' and 3' ss irrespective of the regulatory elements present. Besides p150, all the other proteins show sensitivity to stimulation and/or replacement of the USE60 or DSE120 and are potentially involved in the enhancement of exon 6 inclusion upon signal induction (Figure 6.1, panel D). I had previously confirmed that the 50kDa band was indeed CELF2 by IP with a CELF2 antibody. I have identified two more proteins, hnRNPH1 and hnRNP C, which bind around LEF1-E6.

Since the protein band that has the greatest change in binding around LEF1-E6 between unstimulated and stimulated nuclear extract is the one around 50kDa (Figure 6.1), I wanted to ensure that there weren't other proteins besides CELF2 that were contributing to the change. The best candidates that might overlap with CELF2 at 50 kDa are the hnRNP F/H family of proteins. This family consists of hnRNPH1, H2, H3 and hnRNP F. hnRNP H1 and H2 are the best characterized of the H proteins and are 96% identical. hnRNP F is 68% identical to hnRNPH1 but with 80% similarity in the region of RRM3[105]. These proteins are similar to CELF2 in that they also

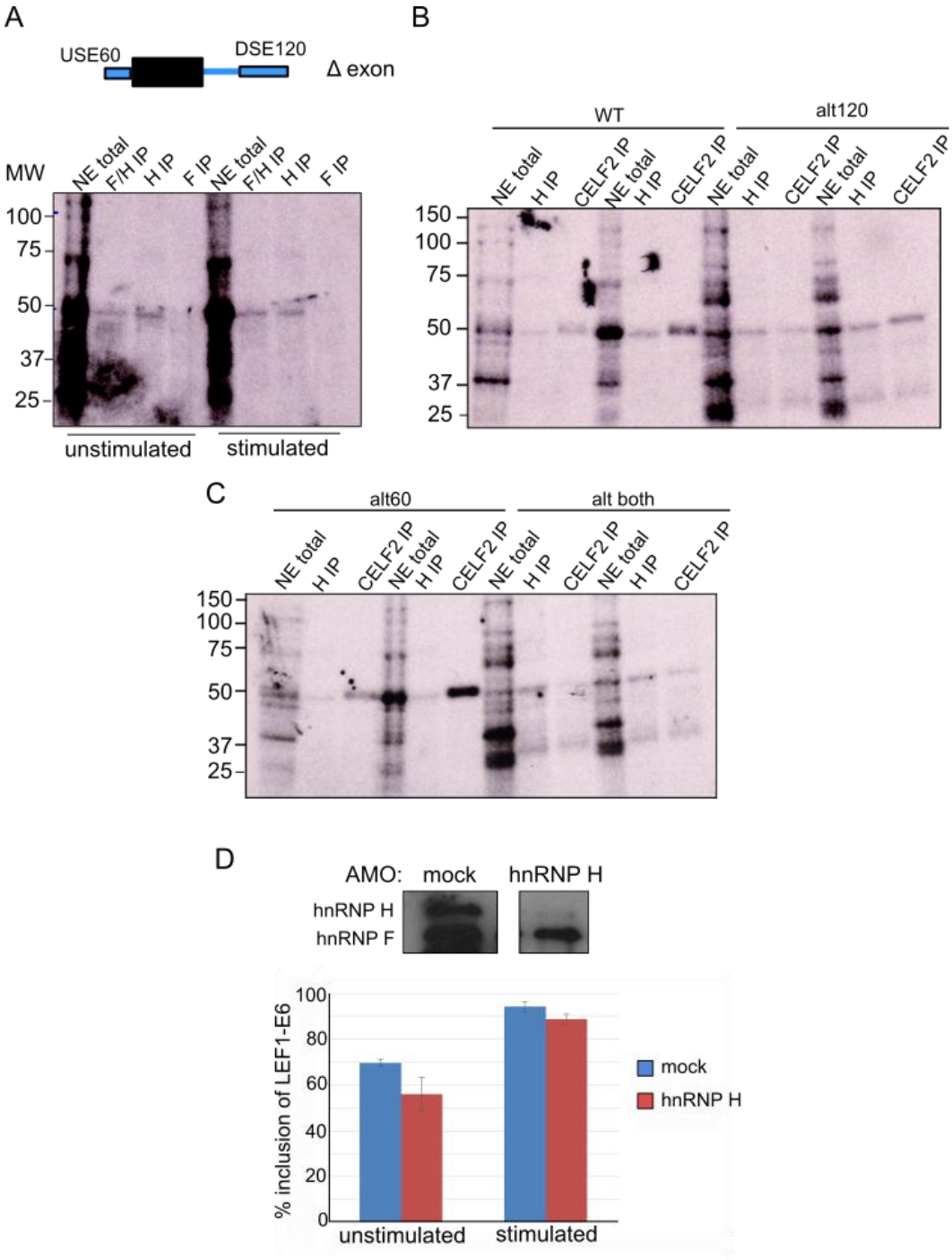


Figure 6.2: hnRNP H binds the region around LEF1-E6. (A) UV crosslinking of the Δexon LEF1 construct in US and S NE coupled with IP with hnRNP F/H, hnRNP H and hnRNP F antibodies. (B) UV crosslinking of WT and alt120 constructs in US and S NE coupled with IP with hnRNPH and CELF2 antibodies. (C) UV crosslinking of alt60 and alt both constructs in US and S NE coupled with IP with hnRNPH and CELF2 antibodies. (D) Western blot showing specific knockdown of hnRNP H in US and S JSL1 cells with graphical representation of RT-PCR results of LEF1-E6 inclusion upon hnRNP H knockdown. Knockdown of hnRNP H by Nicole Martinez.

contain 3 RRM domains with a linker domain connecting RRM2 and 3. They differ in that they have an additional glycine-rich domain at their C terminus. They are also similar to CELF2 in that their molecular weight is ~50kDa and therefore overlap with CELF2 on SDS-PAGE.

In order to determine if any members of the hnRNP F/H family bound the region around LEF1-E6, I used an antibody that recognized both hnRNP F and H. As shown in Figure 6.2 panel A, a band around 50kDa specifically precipitated with this antibody. In order to determine whether the band corresponded to the F or H proteins, I used a H family specific and a F specific antibody in the IP. Figure 6.2 panel A shows that the H antibody specifically pulled down the 50kDa band as opposed to the F antibody. I could therefore confirm that hnRNP H does bind around LEF1-E6. I also wanted to test whether hnRNP H binding changed upon stimulation or whether the change in binding seen in Figure 6.1 was purely due to CELF2. hnRNP H pulled down to approximately the same amounts in both US and S nuclear extract, confirming that the increase in binding was only due to CELF2 (Figure 6.2, panel A). In order to isolate whether H binds the USE60 or DSE120, I performed the IP with the alt60, alt120 and alt both substrates (Figure 6.2, panels B and C). Figure 6.2 panel C shows there is a slight decrease of binding in the alt60 construct, suggesting that the USE60 could be its preferred binding site. However, since the protein still binds in the alt both substrate, it seems that hnRNP H is capable of binding to a few sites around the exon

Although H binding doesn't change drastically upon stimulation, it is possible that its mode of binding changes such that it forms different interactions upon stimulation that allow for increased CELF2 binding to the DSE120. In order to determine whether hnRNP H played a role in signal induced enhancement of LEF1-E6 inclusion, Nicole Martinez in our lab knocked down this protein in US and S JSL1 cells. Analysis of LEF1-E6 inclusion showed that knockdown of H in US cells caused a small decrease in LEF1-E6 inclusion. In stimulated JSL1's knockdown of H had no effect on exon 6 inclusion (Figure 6.2 panel D). This implies the H serves as a modest activator of LEF1-E6 inclusion in US JSL1 cells. This however does not provide an explanation for how CELF2 binding is regulated at the DSE120 in US cells. If H played a part in restricting CELF2 binding to

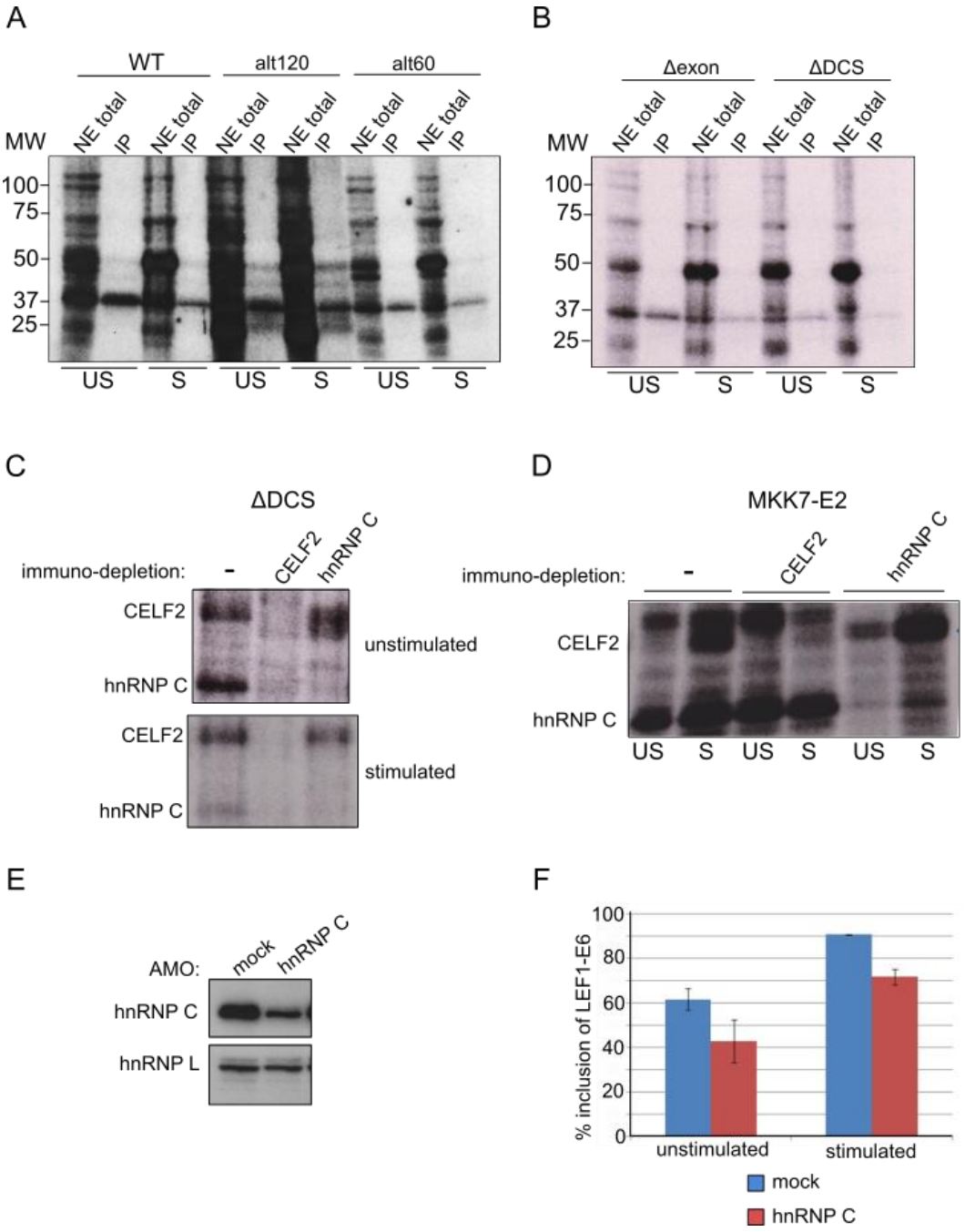


Figure 6.3: hnRNP C binding to the region around LEF1-E6 is depedant on binding of CELF2. (A) UV crosslinking of WT, alt120 and alt60 constructs in US and S NE coupled with IP with an hnRNP C antibody. (B) UV crosslinking of Δexon and ΔDCS constructs in US and S NE coupled with IP with an hnRNP C antibody. (C) UV crosslinking of ΔDCS in US and S NE immuno-depleted of hnRNP C or CELF2. (D) UV crosslinking of MKK-E2 in US and S NE immuno-depleted of hnRNP C or CELF2. MKK-E2 RNA from Nicole Martinez. hnRNP C depleted extract from Laura Agosto. CELF2 depleted extract from Sam Allom. (E) Western blot showing specific knockdown of hnRNP C in US and S JSL1 cells (F) Graphical representation of RT-PCR results of LEF1-E6 inclusion upon hnRNP C knockdown. hnRNP C knockdown done by Nicole Martinez.

the DSE120 in US cells, its knockdown should increase LEF1-E6 inclusion in this condition. Therefore, we can confirm that hnRNP H mildly contributes to the amount inclusion in US T cells but does not affect signal induced enhancement of LEF1-E6.

Another protein band who's binding changes upon stimulation is p37 which decrease upon stimulation as CELF2 binding levels increase. A prime candidate for this species is hnRNP C, a member of the hnRNP family of splicing factors with catalogued roles in mRNA splicing[110,111]. UV crosslinking the IP with an anti-hnRNP C antibody and radio-labeled WT, show a specific interaction between the two, confirming that hnRNPC binds in the LEF1-E6 region. Upon stimulation as CELF2 binding levels increase, hnRNPC levels decrease. This pattern is maintained in the alt60 substrate where there is also an increase in CELF2 upon stimulation and a corresponding decrease in hnRNP C. In the alt120 however, CELF2 binding is minimal and with the absence of the DSE120, there is no change in CELF2 levels upon stimulation. Figure 6.3 panel A shows that hnRNP C levels also remain unchanged in the alt120 substrate upon stimulation.

An important question to consider is whether the decrease in hnRNP C binding (with a concurrent increase of CELF2 binding to the DSE120) only possible in the *stimulated* state or can an increase of CELF2 binding to the DSE120 in the *unstimulated* state also elicit the same response from hnRNP C. Since both the ∆exon and the ∆DCS exhibit a greater amount of CELF2 binding to the DSE120 in the unstimulated state as compared to WT, I tested the IP of hnRNP C in these two substrates. In both the ∆exon and ∆DCS, the binding of hnRNP C in the unstimulated extract was significantly lower than in the WT substrate and reduced further upon stimulation (Figure 6.3, panel B). Therefore, the correlation between CELF2 and hnRNP C binding is not dependent on the state of the extract and only dependent on the amount of CELF2 bound to the DSE120.

To investigate whether a reduction in CELF2 can lead to increased binding of hnRNP C on the RNA, I repeated the UV crosslinking with the ∆DCS construct in extracts immuno-depleted of CELF2 and C (depletion done by Laura Agosto and Sam Allon respectively). When levels of CELF2 are reduced via immune-depletion, there is an almost complete loss of all protein binding, including hnRNP C, on the RNA. However, if levels of hnRNP C are also reduced via immune-depletion,

there is only a minimal effect on CELF2 binding (Figure 6.3 panel C). The fact that hnRNP C binding on the RNA is dependent on the presence of CELF2 suggests that CELF2 recruits hnRNP C to the RNA. This hypothesis is supported by the fact the upon hnRNP C depletion, CELF2 binding is largely unaffected (figure 6.3, panel C). This is in stark contrast to the behavior or CELF2 and hnRNP C in the region around MKK7 exon 2 (MKK&-E2), an alternatively spliced exon studied by Nicole Martinez in the lab. In the case of MKK7-E2, CELF2 and hnRNP C compete for binding as depletion of CELF2 leads to an increase in hnRNP C binding and vice versa (Figure 6.3, panel D).

To determine the effect hnRNP C has on signal induced enhancement of LEF1-E6, I used RNA from hnRNP C knockdowns in unstimulated and stimulated JSL1s (from Nicole Martinez in the lab) to probe the inclusion of LEF1-E6 via RT-PCR. LEF1-E6 inclusion is repressed in both cases suggesting that hnRNP C functions as an activator of LEF1-E6 splicing in both unstimulated and stimulated cells (Figure 6.3 panel E and F). A double knockdown of CELF2 and hnRNP C was done but the data was variable and inconclusive. This could be due to off target effects caused by the double knockdown.

I also performed UV crosslinking IPs on MBNL1, MBNL 2, PSF, HuR, Fox2 and hnRNP L, none of which were found to interact LEF1-E6 constructs.

Discussion:

The inclusion of LEF1-E6 in unstimulated and stimulated T cells is determined by the regulation of CELF2 binding to a repressive sequence (USE60) upstream of the exon and an activating sequence (DSE120) downstream of the exon. In unstimulated cells, CELF2 binding is biased towards the repressive USE60 while in stimulated cells, CELF2 binding is biased towards the activating DSE120. A possible mechanism for how CELF2 binding is regulated is by the involvement of a second protein that inhibits CELF2 binding downstream in the unstimulated state and this inhibition is relieved upon stimulation.

Seven proteins, including CELF2, bind LEF1 around the region of exon 6 (Figure 6.1). One of these proteins is hnRNP H, a splicing factor with a history of influencing alternative splicing[105,106,112–114]. hnRNP H has a very minimal effect on signal induced enhancement of LEF1

exon 6 inclusion as knockdown studies show it has a mild effect activating effect in unstimulated cells (Figure 6.2). This mild activation in unstimulated does not explain how CELF2 binding is restricted to the USE60 in this state. If H1 played a part in restricting CELF2 binding to the DSE120 in US cells, its knockdown should increase LEF1-E6 inclusion in this condition. Therefore, we can confirm that the hnRNP H family of proteins mildly contributes to the amount inclusion in US T cells but does not affect signal induced enhancement of LEF1-E6.

The other protein identified was p37 which was found to be hnRNP C. UV crosslinking of WT substrate in unstimulated nuclear extract shows that the levels of hnRNP C and CELF2 are relatively similar. However, upon stimulation, as CELF2 binding increases on the DSE120, hnRNP C binding decreases. Additionally, the decrease in hnRNP C binding is dependent on the increase of CELF2 binding to the DSE120 as is apparent when the DSE120 is replaced in the USE60-exon6-DCE-het substrate. Here, the lack of the DSE120 prevents an increase of CELF2 binding upon stimulation which also results in a lack of decrease in hnRNP C binding. This suggests that the reduction in hnRNP C binding is not what causes the concurrent increase in CELF2 binding to the DSE120 but the converse wherein hnRNP C is removed from the RNA by the increase in CELF2 binding to the DSE120 upon stimulation. Immuno-depletion of CELF2 shows that hnRNP C requires the binding of CELF2 to the USE60 or DSE120 in order to interact with the RNA. The converse is not true as the immuno-depletion of hnRNP C has a negligible effect on CELF2 binding (Figure 6.3).

One explanation for the requirement of CELF2 in hnRNP C's interaction with the RNA is that CELF2 forms a complex with hnRNPC and recruits it to the region around LEF1-E6. In unstimulated cells this CELF2-hnRNP C complex binds to the USE60. Upon stimulation, the increase of CELF2 binding downstream interferes with the upstream CELF2-hnRNPC interaction and releases hnRNP C from the transcript. It is important to note that hnRNP C is capable of binding the DSE120 in the absence of the USE60 (Figure 6.3 panel A) and that the decrease in binding due to CELF2 is independent of which elements it is bound to. This however does not explain the mechanism behind how CELF2 is biased towards the USE60 in unstimulated cells and what causes

the increase in binding to the DSE120 upon stimulation. The knockdown of hnRNP C shows that it is an activator of LEF1 splicing in both unstimulated and stimulated cells. If the presence of hnRNP C in unstimulated cells was responsible for the lack of CELF2 binding downstream, then it would serve as a repressor in unstimulated cells (Figure 6.3).

The dependence of hnRNP C binding on the presence of CELF2 is still very interesting to study as it highlights the various ways in which splicing factors co-operate or antagonize to interact with RNA. The one way dependence of hnRNP C on CELF2 in the case of LEF1 is particularly interesting considering data from Nicole Martinez in my lab, which shows that in the region around MKK7-E2, CELF2 and hnRNP C compete for binding. Depletion of CELF2 leads to an increase in hnRNP C binding and vice versa. It would be very interesting to determine the specific conditions that cause competition versus dependency (Figure 6.3 panel D).

Although hnRNP H and C bind to the region around LEF1-E6 neither is involved in the signal dependent enhancement of LEF1-E6 inclusion upon stimulation. Identification of p45, a protein whose binding also decreases upon increase CELF2 binding upon stimulation could help decipher this mechanism. Additionally, p100 could especially be important as its levels decrease upon stimulation and upon loss of DSE120. Considering it is likely p100 binds to the DSE120 (Figure 6.1 panel D), it could sterically block CELF2 from binding part of the DSE120 in unstimulated cells and this effect is reduced upon stimulation perhaps due to the down regulation of expression of this protein. Nonetheless, further studies into the identification of these proteins could shed light on the mechanism behind the signal induced enhancement of LEF1-E6 in T cells.

CHAPTER 7: CONCLUSIONS AND FUTURE DIRECTIONS

The region encoded by LEF1-E6 is crucial to the upregulation of T cell receptor α and is therefore preferentially included in the final LEF1 pre-mRNA transcript upon T cell stimulation. Previous work by Mallory et al identified two intronic *cis*-elements, the USE60 and DSE120, flank exon 6 and are required for this preferential inclusion upon stimulation. Mallory et al also discovered that the splicing factor CELF2 bound these two regulatory elements and the increase of CELF2 binding to these elements upon stimulation led to increased exon 6 inclusion

In Chapter 2 of my thesis I show that the USE60 and the DSE120 are required but also sufficient for enhancement of exon inclusion upon stimulation. Establishing the minimal sequence elements required for exon enhancement is crucial to accurately describing a mechanistic model for how exon 6 of LEF1 is preferentially included upon T cell stimulation. In Chapter 3 and Chapter 4, I correlate the binding of CELF2 to the function of the USE60 and DSE120. I show that the USE60 and DSE120 do not work synergistically to enhance inclusion but function antagonistic to each other. The USE60 is a repressor of splicing while the DSE120 is an enhancer. Consequently, I show that in order to achieve an increase in exon 6 inclusion only upon stimulation, CELF2 binding is highly regulated between the USE60 and DSE120. In unstimulated cells, binding is biased towards the USE60 and upon stimulation the increase in CELF2 binding happens purely on the DSE120, the splicing activator. This bolus of CELF2 binding on the DSE120 upon stimulation leads to an increase in exon 6 inclusion.

A more accurate analysis of this data is that in unstimulated cells, CELF2 binding on the DSE120 is inhibited, allowing the USE60 to have the dominant effect. Upon stimulation this inhibition is relieved, allowing more CELF2 binding to the DSE120, which overcomes the repressive effects of the USE60. This inhibition of CELF2 binding to the DSE120 in unstimulated cells is a key point of regulation in this system. *In vivo* mingene data combined with UV crosslinking data shows that both the exon and DCS regulate CELF2 binding to the DSE120 in unstimulated cells.

The minigene studies of the Δexon and ΔDCS show an increase in Exon 6 inclusion over WT in unstimulated cells. The UV crosslinking profiles of these constructs in unstimulated nuclear

extracts shows an increase in CELF2 binding as well. This suggests that some of the inhibition of CELF2 binding the DSE120 in unstimulated cells was relieved by the replacement of these sequences. However, both the ∆exon and ∆DCS constructs are still capable of a further increase of CELF2 binding on the DSE120 and therefore a further increase in exon 6 inclusion upon stimulation (Figure 6.1). Therefore, the signal responsive sequence that is responsible for inhibiting CELF2 binding to the DSE120 in unstimulated cell is contained within the DSE120. Creating minigenes that systematically replace/delete parts of the DSE120 using the ∆DCS as the template would isolate the specific part of the DSE120 that contributes to inhibiting CELF2 binding in unstimulated cells.

Having the sole point of regulation revolve around binding of CELF2 to the DSE120 however implies that no other sequence element would be required. In fact UV crosslinking analysis of the ∆USE60 shows that the increase of CELF2 binding to the DSE120 upon stimulation can still take place in the absence of the USE60. However, the minigene studies confirm that the USE60 is required for the signal induced enhancement of exon 6 inclusion. In the ∆USE60 minigene, despite the increase of binding on the DSE120 upon stimulation, there was no further increase in exon inclusion. The absence of the USE60 does increase basal levels of inclusion to 80% as it is a splicing repressor. This does not explain why there is no further increase upon inclusion because the ∆exon minigene also has a basal level of inclusion of 80% and is still capable of further increase in inclusion. Therefore, the regulation of CELF2 binding to the DSE120 is independent of the USE60. However, the mechanism by which it increases exon 6 inclusion is through the USE60. In other words, the increase of CELF2 binding to the DSE120 activates exon 6 inclusion by de-repressing the repression of the USE60.

This data leads to a model where binding of CELF2 to the DSE120 is inhibited in unstimulated cells and this inhibition is relieved upon stimulation. Once CELF2 is allowed to bind to DSE120 is doesn't function independently to overcome the effects of the USE60 but rather "activates" exon 6 inclusion by de-repressing the repression caused by the USE60. Therefore, when the USE60 is absent (as in the ∆USE60 minigene construct), the increase in CELF2 binding

to the DSE120 has no effect on inclusion as the target on which it acts is absent. To test this hypothesis in the future, the 1X-4X minigene constructs can be used. These minigenes are a convenient way to titrate the amount of CELF2 that binds downstream. In the presence of the USE60, each increase in high affinity binding sites and therefore each increase in CELF2 binding led to an increase in exon 6 inclusion. Replacing USE60 in these minigenes, can test whether the mechanism by which the increase in CELF2 binding downstream leads to increased exon inclusion, occurs through the USE60. If the activation of exon inclusion is through the de-repression of the USE60, there should be no increase in inclusion in any of the 1-4xUGUU minigenes.

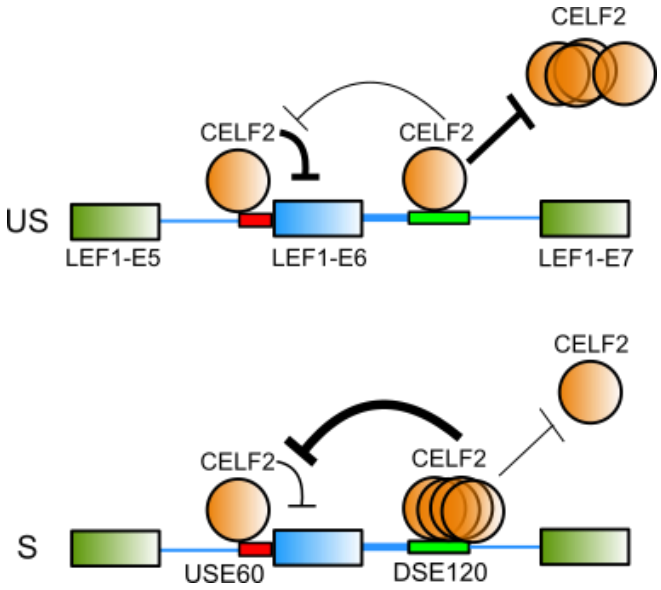


Figure 7.1: Model for the enhancement of LEF1-E6 inclusion upon T cell stimulation by the differential binding of CELF2 to the USE60 and DSE120.

In Chapter 5, I describe my efforts to determine whether the regulation of CELF2 binding to the DSE120 in unstimulated and stimulated cells is due to a change in its PTM landscape. In collaboration with the Wistar Institute's mass spectrometry core, we performed mass spectrometry studies of CELF2's PTM landscape and revealed 12 previously unannotated PTMs on CELF2 RRMs. None of these PTMs appeared to change after stimulation and therefore couldn't explain the differential binding of CELF2 on the DSE120. However, the linker domain escaped analysis under the current digest conditions. This domain has been shown to be crucial for the regulation of

alternative splicing by CELF2's RRMs[115]. It is possible that a PTM in the linker domain could be regulating CELF2 's interaction with RNA.

Evidence from Han et al has shown that CELF2 can regulate alternative splicing via either RRM1+RRM2 in tandem, or via RRM3[115]. The differential binding in unstimulated and stimulated cells could be due to the requirement for one set of RRMs versus the other. For future studies, the bacterial expression system for his-CELF2 can be used to create domain deletions of CELF2. These can be tested in EMSA to see if either set of RRMs has a particularly low affinity for the DSE120. These can also be used in in vitro splicing assays to test whether one set of RRMs is more efficient at activating splicing over the other. The EMSAs I have done with both his-CELF2 and the modified FLAG-CELF2 have shown that the PTM are not required for RNA interaction, and therefore determining whether there is a bias in RRM use for the DSE120 could shed some light on the mechanism by which CELF2 binding to the DSE120 is regulated.

Chapter 6 details my efforts into identifying other proteins that could also be playing a part in how LEF1-E6 is regulated. These proteins could interact with CELF2 differentially between resting and stimulated cells and thereby regulate when it binds the DSE120. These proteins could also interact with DSE120 differentially between unstimulated and stimulated cells and sterically hinder binding of DCSE120. I have identified two proteins, hnRNP H and hnRNP C, that bind around the region of LEF1-E6. However, there are a few other candidates that have yet to be identified. From the UV crosslinking of LEF1-E6 substrates in nuclear extract (Figure 6.1), p45 and p100 stand out as being the most sensitive to PMA stimulation and DSE120 replacement. Identification of these proteins via UV crosslinking IP should be pursued to assess their involvement in enhancement of LEF1-E6 upon signal induction.

Finally, the end result of the regulation of CELF2 binding is the manipulation of spliceosome assembly to either promote or inhibit exon inclusion. Initial results probing spliceosome involvement have shown that increasing splice site strength does not affect signal induced regulation suggesting that spliceosome assembly is unaffected at the point of splice site recognition. This is corroborated by preliminary results monitoring spliceosome assembly in WT and ΔDSE120 constructs. In both

cases, the formation of A complex (where splice site recognition takes place) was unhindered, even with the replacement of the DSE120. Future studies should include using the in vitro splicing conditions described in this thesis to optimize spliceosome assembly reactions to identify which stage in spliceosome assembly is being inhibited in the unstimulated state. Being able to determine the exact point in spliceosome assembly that is being affected by the lack of CELF2 binding downstream would provide valuable information behind the mechanisms of alternative splicing regulation by CELF2. Furthermore, the *in vitro* splicing reactions can be used to purify assembled spliceosomes from stimulated and unstimulated state to identify what specific interactions with the spliceosome are important to LEF1-E6 regulation.

Studies that have investigated CELF2's interactions with spliceosome components during alternative splicing regulation suggest that CELF2 activates exon inclusion by binding downstream of an alternative exon and stabilizing the association of U2 snRNP upstream[56]. Additionally, CELF2 represses exon inclusion by binding upstream of an alternative exon and displacing the association of U2AF65[57]. It is possible that these mechanism are at play in LEF1-E6 regulation. The current model states that CELF2 binding DSE120 encourages exon 6 inclusion by de-repressing the repressive effect of the USE60. The binding of U2AF65 at the PPT is crucial to U2 snRNP's association at the 3'ss and further assembly of the spliceosome. If the CELF2 that binds the USE60 inhibited or destabilized U2AF 65 binding, this would also effect binding of U2 snRNP and reduce exon 6 inclusion. Upon stimulation, greater CELF2 binding downstream to the DSE120 would stabilize U2 snRNP at the upstream 3'ss and therefore encourage exon inclusion. This would also explain the requirement of both the USE60 and DSE120. In the absence of the USE60, there would be no destabilization of U2AF65 and therefore no requirement for stabilization of U2snRNP.

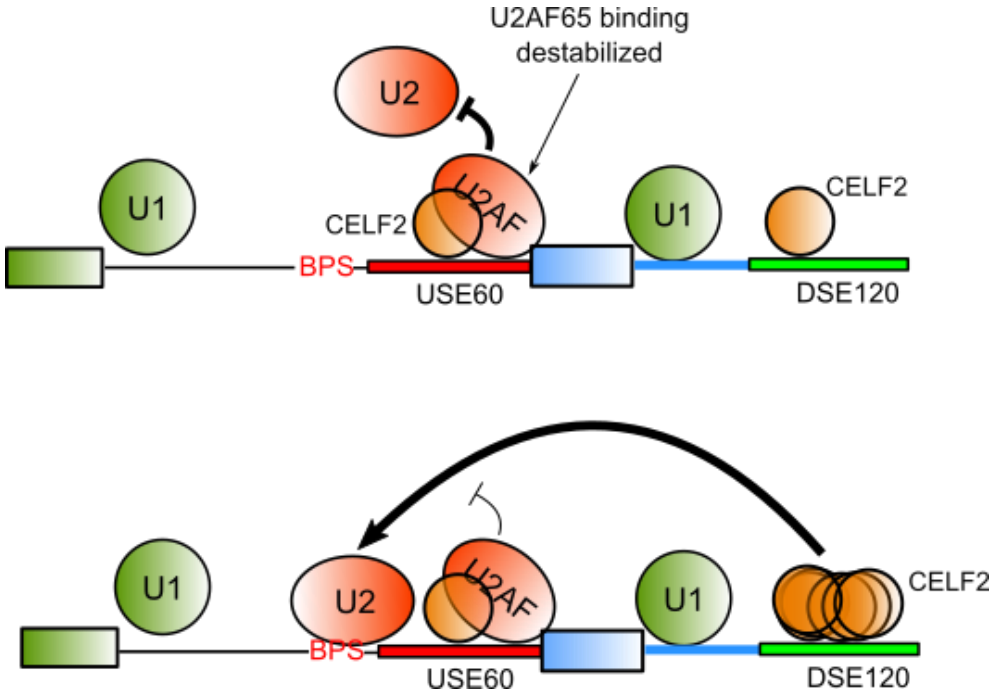


Figure 7.2: Possible model for CELF2's influence on the spliceosome during regulation of LEF1-E6 alternative splicing.

CHAPTER 8: MATERIALS AND METHODS

Minigenes and RNA

The 90/160, 0/160, 90/40, alt120 and Δexon minigenes were previously described in Mallory et al, 2011. Briefly, the minigenes consist of LEF1 exon 6 and surrounding intron flanked by intron and exon sequence from the human β–globin gene. alt-both was made using PCR by using the alt60 and alt120 sequences to replace the USE60 and DSE120 in the same construct. ΔDCS was made using PCR with primers to the DCS element Δexon and replacing it with____. The lab has used this sequence extensively in previous studies with no discernable effects on splicing. 1x-4x UGUU minigenes were made using synthetic oligos for the 1xUGUU sequences and with added restriction enzyme sites to allow oligomerization. These were cloned into the WT minigene to replace the DSE120.

Oligonucleotides encoding the USE60 and DSE120 were cloned directly downstream of a T7 polymerase promoter and served as templates for *in vitro* transcription. 1x-4x UGUU were amplified from minigenes with primers with a T7 tag attached using PCR and used as templets for *in vitro* transcription. The RNAs were transcribed with T7 polymerase (Promega) in the absence or presence of 32P-UTP to radioactively label probes. The USE60/DSE120, ΔUSE60/DSE120, alt60/DSE120, USE60/alt120 and ΔUSE60/alt120 were created by PCR with primers (with a T7 tag attached) to the respective minigene that contained the sequence. These were used as templates for *in vitro* transcription without radio-label.

Nuclear extract and recombinant proteins

Nuclear extract was purified from JSL1 cells using a standard protocol previously described in Lynch et al, 2001[116]. Recombinant his-tagged CELF2 was expressed in Rosetta (DE3) PLysS cells, a BL21 derivative of *E.Coli* and were purified using Nickel-NTA resin (GE Biosciences) according to company provided protocols. JSL1 cells stably expressing FLAG-CELF2 were grown under unstimulated or stimulated conditions and then lysed to prepare nuclear extract. Tagged proteins were purified from nuclear extract with EZ-View Red FLAG-conjugated resin (Sigma) in GFB300 (20 mM Tris-Cl, pH 7.5, 300 mM KCl, and 0.2 mM EDTA, pH 8.0). Following extensive washing in

GFB300, the proteins were eluted with 500 ng/ul of 3X Flag peptide (Sigma). Protease and Phosphotase inhibitor cocktails (Sigma) and de-acetylase inhibitor (Millipore) were used during the purification.

Cell culture

JSL1 cells (Lynch et al, 2000)[31] were cultured in RPMI+ 5% fetal calf serum at 37°C in 5% CO2. Sub-lines of JSL1 cells that stably express the minigenes described were created by transfecting 10 million cells with 10 ug of minigene plasmid by electroporation and grown under drug selection as described by Rothrock et al., 2003[117]. For splicing analysis, three independent clones of each minigene were either left untreated or treated with 20 ng/ml of PMA for 60 h, after which cells were harvested and total RNA extracted using RNABee (Tel-Test). Minigene derived spliced products were analyzed by RT PCR using vector-specific primers (see below).

RT-PCR

RT-PCR and analysis was carried out as previously described in detail by Rothrock et al. 2003. In brief, a low-cycle PCR protocol was used, such that the signal detected is linear with respect to input RNA. Minigenes were analyzed using the vector-specific primers ACT and GE3R (sequence published in Rothrock et al., 2003). Quantitation was done by densitometry using a Typhoon Phosphoimager (Amersham Biosciences).

Western blotting

Western blotting was carried out as previously described in Lynch et al, 2000. Antibodies for Western blots were as follows: anti-hnRNP L (4D11, Abcam), anti-CELF2 (HL1889, University of Florida Hybridoma Lab), anti-hnRNP C, (ab10294, Abcam), anti-hnRNP F/H (ab10689, Abcam), anti-hnRNP F (ab50982, Abcam), anti-hnRNP H (ab10374, Abcam).

***in vitro* splicing assay**

Unlabeled RNA substrates (10 nM) were incubated with 30% unstimulated JSL1 nuclear extract in a total volume of 12.5 µl under splicing conditions, which contains (final concentration):

12 mM Tris-HCl, pH7.5, 3.2 mM MgCl2, 4 mM ATP, 20 mM CP, 0.5 mM DTT, 0.125U

RNasin (Promega), 60 mM KCl, 0.1 mM EDTA, and 12% glycerol. Reactions were incubated for 90 min at 30°C; then the RNA was recovered from the reactions by proteinase K treatment, phenol-chloroform extraction and EtOH precipitation. The recovered RNA was analyzed by RT-PCR.

RNA electro-mobility shift assays (EMSA)

in vitro transcribed RNAs were gel-purified and adjusted to 104 cpm/ml specific activity. Each RNA was incubated with US or S FLAG-CELF2 or his-CELF2 in a total volume of 10ul under splicing conditions similar to that described for the in vitro splicing assays with the addition of 0.8 mg of BSA and 0.8mg. Reactions were incubated for 20 min at 30°C, after which heparin was added to a final concentration of 5 mg/ml and incubated for an additional 5 min at 30°C. Reactions were analyzed on a 4.5% native gel (Acrylamide/Bis 29:1 BioRad) and visualized by autoradiography.

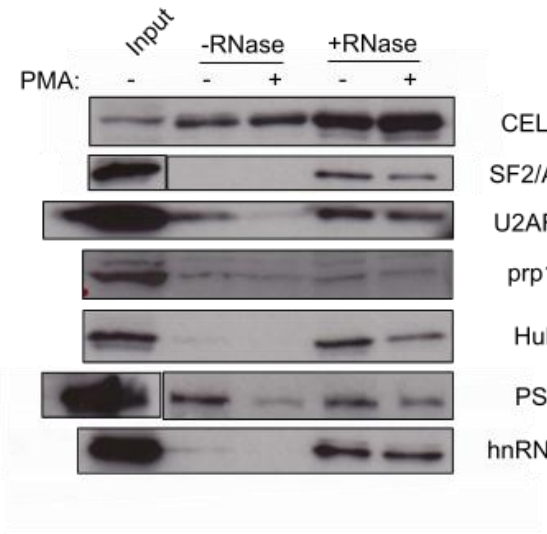
UV Crosslinking

Radiolabeled RNA was incubated in JSL1 nuclear extract under similar conditions described for the EMSAs. Reactions were incubated for 20 min at 30°C, crosslinked using UV light (254 nm) for 20 min on ice, and digested with 2 ug (final concentration) of RNase T1 and RNase A each for 20 min at 37°C. Reactions were analyzed under denaturing conditions on a 12% gel (Acrylamide/Bis 37.5:1, BioRad), and visualized by autoradiography.

APPENDICES

Appendix A: Analysis of CELF2's co-associated proteins

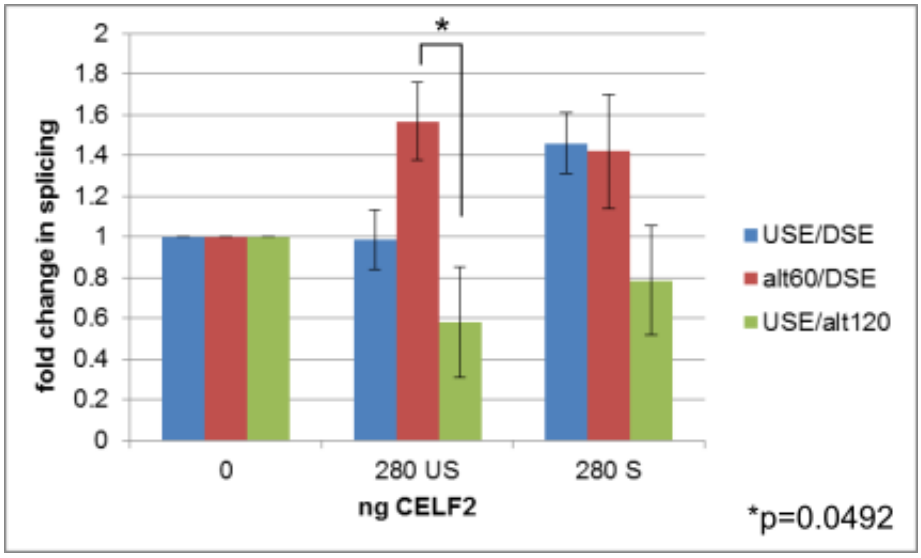
FLAG-CELF2 was IP'ed using it's FLAG-tag from US and S JSL1 cells in the absence of Rnase. The proteins that co-IPed with CELF2 were was analyzed using mass spectrometry in collaboration with the Wistar Institute. Some of these interactions were compared with proteins that co-IPed with FLAG-CELF2 in the presence of RNase to determine which of them were direct interactions and which of them required RNA. The figure below summarizes western blots that probe these interactions in the presence and absence of RNase in the US and S state. They also contain spectral counts from the mass spectrometry results.



	Unstimulated Spectra#	Stimulated Spectra#	S/US Spectra Ratio
CELF2	252	244	0.97
SF2/ASF	11	22	2.00
U2AF35	20	13	-1.54
prp19	21	21	1.00
HuR	16	27	1.69
PSF	136	77	-1.77
hnRNP C	17	17	1.00

Appendix B: *in vitro* splicing in the presence of US and S FLAG-CELF2

in vitro splicing assays as performed in Chapter 4 Figure 4.1 were repeated with increasing amount of US or S FLAG-CELF2 protein. Although the fold change in splicing upon addition of 280 ng of US or S FLAG-CELF2 was only significant between the alt60/DSE120 and USE60/alt120, the trends for all comparisons support the hypothesis US-CELF2 is repressive and S-CELF2 is activating. Whether this is due to change in PTMs or the influence of another protein is yet to be determined.



Appendix C: UV crosslinking in nuclear extract

1. Set up binding reaction:

   4 ul Nuclear Extract, protein or buffer (BC100 or BC300)
   1 ul BC400
   1 ul tRNA/BSA/BC100 mix (0.25ug/ul tRNA + 0.2 ug/ul BSA in BC100)
   0.4 ul 80mM MgCl2
   0.4 ul 25mM ATP
   0.4 ul 0.5M creatine phosphate (CP)
   1 ul H2O
   1 ul radiolabeled RNA

2. Incubate binding reaction at 30oC for 20 minutes
3. Pipet reactions into a 96-well U bottom plate on ice (keep top off plate)
4. Place handheld 254nm lamp overtop, cover in aluminium foil and crosslink for 20 minutes
5. Pipette reactions into new Eppendorf tubes that contain 20U RNase T1 + 20ug RNase A
6. Digest at 37^{o}C for 20 minutes
7. Add 13ul 2xSDS sample buffer and 1.3 ul 2-ME
8. Boil samples for 5 minutes and load on 10% SDS-PAGE gel (30 ml)

Appendix D: UV crosslinking in nuclear extract with immuno-precipitation

UV Crosslinking with Immuno-Precipitation (IP)

1. Set up binding reaction (use 2-4 reactions per IP):

   4 ul Nuclear Extract
   1 ul BC400
   1 ul tRNA/BSA/BC100 mix (0.25ug/ul tRNA + 0.2 ug/ul BSA in BC100)
   0.4 ul 80mM MgCl2
   0.4 ul 25mM ATP
   0.4 ul 0.5M creatine phosphate (CP)
   1 ul H2O
   1 ul radiolabeled RNA

2. Incubate binding reaction at 30^{o}C for 20 minutes
3. Pipet reactions into a 96-well U bottom plate on ice (keep top off plate)
4. Place handheld 254nm lamp overtop, cover in aluminium foil and crosslink for 20 minutes
5. Pipette reactions into new Eppendorf tubes that contain 20U RNase T1 + 20ug RNase A
6. Digest at 37^{o}C for 20 minutes
7. Pool 2-4 identical reactions in Eppendorf tube
8. Add 5-10 ul antibody
9. Bring volume to 400 ul with 1xRIPA buffer
10. Rotate overnight at 4^{o}C
11. Add 40 ul Protein-G sepharose beads pre-washed in 1xRIPA buffer
12. Rotate 1 hour at 4^{o}C
13. Spin down at 0.8g for 30 seconds and wash 3x with 1xRIPA buffer
14. After last wash, remove buffer and resuspend beads in
15. Add 13ul 2xSDS sample buffer and 1.3 ul 2-ME
16. Boil samples for 5 minutes and load on SDS-PAGE gel

1xRIPA
50mM TRIS pH 8
1% NP-40
0.5% sodium deoxycholate
150mM NaCl
0.1% SDS

Appendix E: PCR for making DNA templates for in vitro transcription

Forward Primers (FP), Reverse Primer (RP) and Plasmids:
in vitro splicing - 2-exon constructs:

Construct	FP	RP	Plasmid
USE60/DSE120	SA60	Adml BamHI R	Adml LEF1 S5
USE60/alt120	SA60	Adml BamHI R	Adml LEF1 S7m
alt60/DSE120	SA68	Adml BamHI R	Adml LEF1 S11
Δ/DSE120	SA58	Adml BamHI R	Adml LEF1 S5
Δ/alt120	SA58	Adml BamHI R	Adml LEF1 S7m
alt60/alt120	SA68	Adml BamHI R	Adml LEF1 43

EMSAs:

Construct	FP	RP	Plasmid
4xUGUU	SA34	SA35	LEF1 S32
DSE120	SA34	SA35	LEF1 S5

UV crosslinking

Construct	FP	RP	Plasmid
USE60/DSE120	SA44	SA65	LEF1 S5
USE60/alt120	SA44	SA55	LEF1 S7m
alt60/DSE120	SA69	SA65	LEF1 S11
alt60/alt120	SA69	SA55	LEF1 S43
Δexon	SA44	SA65	LEF1 S14
ΔDCS	SA44	SA65	LEF1 S45

Set up 3 reactions per construct.
1 x reaction:
30 ul H2O
5 ul 10xPfu buffer
10 ul 1mM dNTPs
1.25 ul 100ng/ul FP
1.25 ul 100ng/ul RP
1 ul 100ng/ul Plasmid (see above table)
1 ul Pfu
PCR Program:

1. 94°C 2 min
2. 94°C 45 sec
3. 45°C 45 sec
4. 72°C 45 sec
   repeat 2-4 for 34 cycles
5. 72°C 10 min
6. 4°C hold

Run all reactions on a 1% agarose gel. Cut out band corresponding to the right size and extract the DNA from the gel using Qiagen's gel extraction kit. Use 25 ul ddH_2O to elute DNA of column. You will usually need two columns per construct. Concentration of eluted DNA is between 50-80 ng/ul.

NOTE:

USE60 and 2xUGUU constructs were created by linearizing plasmid DNA (pcAT7 USE60 and pcAT7 2xRTB) with HINDIII.

Reaction:

10 ul plasmid
20 ul 10xBuffer
5 ul HINDIII
165 ul H_2O

Incubate for 2 hours at 37^{o}C. PCA extract and resuspend in 10ul ddH_2O.

Appendix F: *in vitro* transcription to create non-labeled or radio-labeled RNA for in vitro splicing, EMSAs or UV crosslinking assays

1. Thaw all stock reagents to room temperature before use.
2. For each transcription reaction refer to table below. Add DNA to tubes first. Create master mix of remaining ingredients (except P32-αUTP) and add appropriate amounts to tubes with DNA. Add exact amount of P32-αUTP at the end.

Stock Reagents	**non-labeled (in vitro splicing)**
Template DNA, PCR (~50-100ng/ul)	9.7 ul
5x transcription buffer (Promega)	5 ul
10xACG mix (8, 8, 2 mM)	2.5 ul
20mM UTP	1 ul
100mM DTT	2.5 ul
25mM CAP	0.8 ul
Rnasin (Promega)	1 ul
T7 RNA polymerase	2.5 ul
Total	25

Stock Reagents	**radio-labeled** (EMSA, UV crosslinking)
Template DNA, PCR (~50-100ng/ul) or 1000 ng (linearized)	5.5 ul or 1 ul
5x transcription buffer (Promega)	5 ul
10xACG mix (8, 8, 8 mM)	2.5 ul
20mM UTP	0.25 ul
100mM DTT	2.5 ul
P32-αUTP (3.3pmol/ul)	5 ul
Rnasin (Promega)	1 ul
T7 RNA polymerase	2.5 ul
autoclaved H2O	0.75 ul or 5.25 ul
Total	25

3. Incubate at 37^{o}C for 3-4 hours
4. Add 1ul RQ1 DNase and incubate at 37^{o}C for 15 minutes
5. Add 175 ul H2O to each reaction. Then add 200 ul PCA alcohol, shake vigourously and spin at max speed for 5 minutes
6. Transfer top layer carefully to a new eppendorf and ethanol precipitate by adding 1 ul glycogen, 500 ul 100% EtOH, 20 ul 3M NaOAc.
7. Freeze at -20^{o}C for 1 hour

8. Spin at max speed for 15 minutes
9. Wash pellet with 750 ul 75% EtOH
10. Re-suspend pellet in 8ul F dye and boil for 5 mins
11. Load samples onto 4% PAGE gel and run at 1200V until bottom dye reaches end of the gel but doesn't run off.
12. Cut out bands from gel and transfer to new Eppendorf
13. Use the tip of a pipette tip to break apart the gel piece
14. Add 400 ul elution buffer (0.1%SDS in TE buffer) and rotate overnight at room temperature
15. Briefly spin down the gel and transfer the supernatant into a new tube.
16. Add 400 ul PCA alcohol, shake vigourously and spin down at max speed for 5 mins
17. EtOH precipitate as in steps 6-9
18. Carefully remove last amounts of EtOH and resuspend pellet in 10 ul ddH2O.
19. Use 1 ul to measure concentration (non-labeled) or use in a scintillation counter (radio-labeled).
20. Dilute to: in vitro splicing = 1 ng/ul, UV crosslinking = $1x10^5$ cpm, EMSA = $1x10^4$ cpm

Appendix G: *in vitro* splicing without CELF2 protein

Per 13 ul reaction (for either radio-labeled or non-labeled RNA substrate):

0.5 µl 80mM $MgCl_2$
0.5 µl 25mM ATP
0.5 µl 0.5M CP
0.5 ul BC850
3 ul 13% PVA
3 ul BC100
4 µl Nuclear Extract
1 µl RNA substrate (final concentration: 8~10 fmol/ µl = 8~10 nM, 1ng/ul)

1. Mix everything except RNA (master mix) and then add the 12 µl of the master mix into the tube with 1 µl RNA substrate. (everything on ice).
2. Incubate at 30 °C for 90 minutes (or desired amount of time). Add 175 µl proteinase K treatment mix, flick tubes several times to mix well and incubate for 15 minutes at 30°C. (PK treatment per reaction = 100µl 2X PK buffer, 71µl H2O, 2.5µl proteinase K (20mg/ml stock), 1.5ul Glycogen)
3. Following PK treatment, PCA alcohol extract, ethanol precipitate and 70% ethanol wash.
4. Resuspend the pellet in 10 µl H2O. Store at -80 °C, or continue on with the RT-PCR analysis

2X PK buffer
20mM Tris pH7.5
200mM NaCl
25mM EDTA
2% SDS

RT-PCR assay for in vitro splicing reactions with constructs USE60/DSE120, alt60/DSE, USE60/alt120, Δ/DSE120, USE60/Δ:
Forward Primer: SA16 (sequence in LEF1 exon 6)
Reverse Primer: Adml BamHI R (sequence in Adml exon 2)
PCR program:

1. 94°C 2 min
2. 94°C 1 min
3. 70°C 1 min
4. 72°C 1 min
   repeat 2-4 for 25 cycle
5. 72°C 5 min
6. 4°C hold

Expected splice products: spliced RNA= 338nt; spliced RNA = 128nt

Electromobility Shift Assay (EMSA)

1. Cast 4.5% monomer gel and allow to polymerize for an hour
2. Set gel to pre-run for 20 minutes at 150V. In the meantime set up the binding reactions.

3. Binding conditions per 10 ul reaction:

1 ul	13% PVA
1 ul	tRNA/BSA mix in H_2O(0.25ug/ul yeast tRNA and 0.2ug/ul BSA)
0.9 ul	10mM DTT
0.4 ul	80mM $MgCl_2$
0.4 ul	25mM ATP
0.4 ul	0.5M CP
0.1 ul	RNasin
0.8 ul	ddH_2O
3 ul	BC300 or CELF2
1 ul	labeled RNA probe

4. Aliquot 8 ul of master mix to each reaction on ice
5. Add 3 ul his- or FLAG- CELF2 (in BC300)
6. Add I ul ^{32}P-αUTP labeled RNA probe (at 10^4 cpm/ul)
7. Incubate at 30^{o}C for 20 minutes
8. Transfer reactions to ice. Add 1 ul heparin (5 ug/ul) to each reaction
9. Load entire reaction on a 4.5% monomer gel and run at 150V for 1.5 hours
10. Transfer to whatman paper, dry gel and detect by autoradiography.

4.5% monomer gel

45 ml	ddH_2O
6 ml	5xTBE buffer
9 ml	30% acrylamide (29:1)
500 ul	20% APS
50 ul	TEMED

BC300

20mM TRIS pH 7.5
0.2mM EDTA
300 mM KCl
20% glycerol

Appendix H: FLAG-CELF2 purification with protease, phosphatase and de-acetylase inhibitors, RNase and DNase

1. Grow 30L of FLAG-CELF2 expressing JSL1 cells (15L for unstimulated and 15L for stimulated). Stimulate using 2xPMA concentration normally used for 6 well plates.
2. Harvest cells at 3000 rpm, 15 minutes, 4^{o}C.
3. Wash cells with 200ml cold PBS in 250ml conical tubes. Spin at 3000rpm, 15', 4^{o}C.
4. Assess Packed Cell Volume (PCV). Determine 5x PCV of Buffer A
5. Resuspend pellets in 5xPCV buffer A. Incubate on ice for 10'. Spin @ 4000rpm, 15', 4^{o}C
6. Discard supernatant. Resuspend pellets in 2xPCV of Buffer A. Dounce 10 strokes with pestle A.
7. Transfer dounce-ate to new SS34 open tubes and spin 3000rpm, 10', 4^{o}C
8. Remove supernatant (S100). Respin tubes at 16500rpm, 20', 4^{o}C
9. Add HALT Protease and phosphatase inhibitor cocktail (100x, Thermo-Fisher), sodium butyrate (de-acetylase inhibitor, Millipore), 15 ul DNase and 15 ul RNase to 0.7xPCV buffer C
10. Discard supernatant. Resuspend pellets in 0.7xPCV each. Dounce 10 strokes with Pestle A
11. Transfer to conical tube and rock at 4^{o}C for 30'. Transfer to SS34 tube and spin 16500 rpm, 30', 4^{o}C
12. In the meantime, equilibrate 750ul each (for US and S) FLAG beads in TBS. (3x, 8200g, 30'')
13. Remove and save supernatant from 11). Measure volume of sup and dilute 2x with Buffer F.
14. Add diluted nuclear extract to flag beads in conical tube and rotate at 4^{o}C for 1 hour.
15. Dilute 3x flag peptide to 300ug/ml in TBS (300mM KCl).
16. Spin in Tissue culture room centrifuge @ 4000rpm, 1min, 4^{o}C
17. Remove supernatant (unbound). Resuspend beads in TBS (300mM KCl), transfer to a micro-centrifuge tube. Rotate for 5mins and spin at 8200g, 30s. Repeat 3 times.
18. Elute, in batch, 4 times with 3x Flag peptide (300ug/ml). (Same as wash)
19. Re-spin all eluants at 13500rpm for 1 min to pellet residual beads. Transfer sup to new tubes.
20. Take 20ul total (10ul for each gel) and run gels for coomassie and westerm staining. Upon confirming fractions with CELF2 protein – Dialyze into BC300 Buffer. – 3 Hours.
21. Take another 10ul from each dialysis and run Coomassie and if enough volume, another 10 for a western.
22. While gel is running, freeze samples down (dry ice) in 11ul aliquots.

Appendix I: his-CELF2 expression and purification

1. Transform His-CELF2 into Rosetta pLysS cells. Plate out a couple of different dilutions.
2. Pick one colony for an O/N culture into 50ml of 2xYT with 1:1000 dilutions of Kanamycin and Chloramphenicol
3. Make a 1:100 dilution of above culture into 500ml 2xYT (with KAN and CHL). Grow cells till A600=0.8.
4. Induce with 1mM IPTG for 3 hours at 37^{o}C.
5. Add 250ul of 1M PMSF to each culture and centrifuge at 4000rpm for 10 mins.
6. Resuspend Pellet in 20ml Buffer A and freeze both at -80C.
7. Thaw pellet in Room Temperature water bath until a small amount of ice is remaining. Transfer pellet to ice to finish thawing
8. Sonicate using wand lab sonicator (3x) with 1 min cool down in between sonications
9. Centrifuge at 26000rpm for 1 hour. Separate supernatant from pellet. Take gel samples from both.
10. Equilibrate 50ul Ni-NTA beads (50% slurry, so 100ul total) in Buffer A. (resuspend 1ml buffer A, 1000xg for 10s – 3x)
11. Add 1ml of supernatant to beads and rotate at 4C for 30mins. Freeze the rest of the supernatant at -80 for scale up purification
12. Centrifuge at 1000xg for 10s. Remove Supernatant (Flowthrough)
13. Wash beads with Wash buffer (Buffer A +10mM additional Imidazole) 3x
14. Resuspend beads in Sample buffer
15. Run all samples on SDS-PAGE gel
16. Based on amount of protein bound to beads and left in flowthrough, adjust bead volume and scale up purification with the rest of the lysate from step 6.
17. For scale up - elute protein off of beads in Buffer A + 250-500mM Imidazole.

Buffer A:
50mM Na Phosphate
300mM NaCl
10mM Imidazole
10mM BME

REFERENCES

1. Lee, Y. & Rio, D. C. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annu. Rev. Biochem.* **84,** 291–323 (2015).
2. Kornblihtt, A. R. *et al.* Alternative splicing: a pivotal step between eukaryotic transcription and translation. *Nat. Rev. Mol. Cell Biol.* **14,** 153–165 (2013).
3. Nilsen, T. W. & Graveley, B. R. Expansion of the eukaryotic proteome by alternative splicing. *Nature* **463,** 457–463 (2010).
4. Blencowe, B. J. Alternative splicing: new insights from global analyses. *Cell* **126,** 37–47 (2006).
5. Wang, E. T. *et al.* Alternative isoform regulation in human tissue transcriptomes. *Nature* **456,** 470–476 (2008).
6. Wahl, M. C., Will, C. L. & Lührmann, R. The spliceosome: design principles of a dynamic RNP machine. *Cell* **136,** 701–718 (2009).
7. Zhou, Z., Licklider, L. J., Gygi, S. P. & Reed, R. Comprehensive proteomic analysis of the human spliceosome. *Nature* **419,** 182–185 (2002).
8. House, A. E. & Lynch, K. W. Regulation of Alternative Splicing: More than Just the ABCs. *J. Biol. Chem.* **283,** 1217–1221 (2008).
9. Konarska, M. M. & Sharp, P. A. Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. *Cell* **49,** 763–774 (1987).
10. Wassarman, D. A. & Steitz, J. A. Interactions of small nuclear RNA's with precursor messenger RNA during in vitro splicing. *Science* **257,** 1918–1925 (1992).
11. Fu, X.-D. & Ares, M. Context-dependent control of alternative splicing by RNA-binding proteins. *Nat. Rev. Genet.* **15,** 689–701 (2014).
12. Chen, M. & Manley, J. L. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell Biol.* (2009). doi:10.1038/nrm2777
13. Förch, P., Puig, O., Martínez, C., Séraphin, B. & Valcárcel, J. The splicing regulator TIA-1 interacts with U1-C to promote U1 snRNP recruitment to 5′ splice sites. *EMBO J.* **21,** 6882–6892 (2002).
14. Camats, M., Guil, S., Kokolo, M. & Bach-Elias, M. P68 RNA Helicase (DDX5) Alters Activity of Cis- and Trans-Acting Factors of the Alternative Splicing of H-Ras. *PLoS ONE* **3,** (2008).
15. Libri, D., Balvay, L. & Fiszman, M. Y. In vivo splicing of the beta tropomyosin pre-mRNA: a role for branch point and donor site competition. *Mol. Cell. Biol.* **12,** 3204–3215 (1992).
16. Hertel, K. J. Combinatorial control of exon recognition. *J. Biol. Chem.* **283,** 1211–1215 (2008).
17. Motta-Mena, L. B., Heyd, F. & Lynch, K. W. Context-dependent regulatory mechanism of the splicing factor hnRNP L. *Mol. Cell* **37,** 223 (2010).
18. Sharma, S., Kohlstaedt, L. A., Damianov, A., Rio, D. C. & Black, D. L. Polypyrimidine tract binding protein controls the transition from exon definition to an intron defined spliceosome. *Nat. Struct. 38 Mol. Biol.* **15,** 183–191 (2008).
19. Bonnal, S. *et al.* RBM5/Luca-15/H37 regulates Fas alternative splice site pairing after exon definition. *Mol. Cell* **32,** 81–95 (2008).
20. Matlin, A. J., Clark, F. & Smith, C. W. J. Understanding alternative splicing: towards a cellular code. *Nat. Rev. Mol. Cell Biol.* **6,** 386–398 (2005).

21. Kalsotra, A. & Cooper, T. A. Functional consequences of developmentally regulated alternative splicing. *Nat. Rev. Genet.* **12,** 715–729 (2011).
22. Gabut, M., Chaudhry, S. & Blencowe, B. J. SnapShot: The splicing regulatory machinery. *Cell* **133,** 192.e1 (2008).
23. Warzecha, C. C. *et al.* An ESRP-regulated splicing programme is abrogated during the epithelial-mesenchymal transition. *EMBO J.* **29,** 3286–3300 (2010).
24. Xie, J. & Black, D. L. A CaMK IV responsive RNA element mediates depolarization-induced alternative splicing of ion channels. *Nature* **410,** 936–939 (2001).
25. Martinez, N. M. *et al.* Alternative splicing networks regulated by signaling in human T cells. *RNA* **18,** 1029–1040 (2012).
26. Goldsby, R. A., Kindt, T. J., Osborne, B. A. & Kuby, J. *Immunology*. (W. H Freeman and Company).
27. Tortora, G. J. & Derrickson, B. in *Principles of Anatomy and Physiology* (Wiley).
28. Imbeault, M., Giguère, K., Ouellet, M. & Tremblay, M. J. Exon level transcriptomic profiling of HIV-1-infected CD4(+) T cells reveals virus-induced genes and host environment favorable for viral replication. *PLoS Pathog.* **8,** e1002861 (2012).
29. Wu, Z. *et al.* Memory T cell RNA rearrangement programmed by heterogeneous nuclear ribonucleoprotein hnRNPLL. *Immunity* **29,** 863–875 (2008).
30. Ip, J. Y. *et al.* Global analysis of alternative splicing during T-cell activation. *RNA N. Y. N* **13,** 563–572 (2007).
31. Lynch, K. W. & Weiss, A. A model system for activation-induced alternative splicing of CD45 pre-mRNA in T cells implicates protein kinase C and Ras. *Mol. Cell. Biol.* **20,** 70–80 (2000).
32. Tong, A., Nguyen, J. & Lynch, K. W. Differential expression of CD45 isoforms is controlled by the combined activity of basal and inducible splicing-regulatory elements in each of the variable exons. *J. Biol. Chem.* **280,** 38297–38304 (2005).
33. Melton, A. A., Jackson, J., Wang, J. & Lynch, K. W. Combinatorial Control of Signal-Induced Exon Repression by hnRNP L and PSF. *Mol. Cell. Biol.* **27,** 6972–6984 (2007).
34. Rothrock, C. R., House, A. E. & Lynch, K. W. HnRNP L represses exon splicing via a regulated exonic splicing silencer. *EMBO J.* **24,** 2792–2802 (2005).
35. Heyd, F. & Lynch, K. W. Phosphorylation-Dependent Regulation of PSF by GSK3 controls CD45 Alternative Splicing. *Mol. Cell* **40,** 126–137 (2010).
36. Moulton, V. R. & Tsokos, G. C. Alternative splicing factor/splicing factor 2 regulates the expression of the zeta subunit of the human T cell receptor-associated CD3 complex. *J. Biol. Chem.* **285,** 12490–12496 (2010).
37. Izquierdo, J. M. *et al.* Regulation of Fas alternative splicing by antagonistic effects of TIA-1 and PTB on exon definition. *Mol. Cell* **19,** 475–484 (2005).
38. Matter, N., Herrlich, P. & König, H. Signal-dependent regulation of splicing via phosphorylation of Sam68. *Nature* **420,** 691–695 (2002).
39. Mallory, M. J. *et al.* Signal- and Development-Dependent Alternative Splicing of LEF1 in T Cells Is Controlled by CELF2. *Mol. Cell. Biol.* **31,** 2184–2195 (2011).
40. St. Louis, I. V., Dickson, A. M., Bohjanen, P. R. & Wilusz, C. J. CELFish ways to modulate mRNA decay. *Biochim. Biophys. Acta* **1829,** 695–707 (2013).
41. Barreau, C., Paillard, L., Méreau, A. & Osborne, H. B. Mammalian CELF/Bruno-like RNA-binding proteins: molecular characteristics and biological functions. *Biochimie* **88,** 515–525 (2006).
42. Tsuda, K. *et al.* Structural basis for the sequence-specific RNA-recognition mechanism of human CUG-BP1 RRM3. *Nucleic Acids Res.* **37,** 5151–5166 (2009).

43. Teplova, M., Song, J., Gaw, H. Y., Teplov, A. & Patel, D. J. Structural Insights into RNA Recognition by the Alternate-Splicing Regulator CUG-Binding Protein 1. *Structure* **18,** 1364–1377 (2010).
44. Edwards, J. *et al.* Sequence determinants for the tandem recognition of UGU and CUG rich RNA elements by the two N—terminal RRMs of CELF1. *Nucleic Acids Res.* **39,** 8638–8650 (2011).
45. Edwards, J. M., Long, J., de Moor, C. H., Emsley, J. & Searle, M. S. Structural insights into the targeting of mRNA GU-rich elements by the three RRMs of CELF1. *Nucleic Acids Res.* **41,** 7153–7166 (2013).
46. Takahashi, N., Sasagawa, N., Suzuki, K. & Ishiura, S. The CUG-binding protein binds specifically to UG dinucleotide repeats in a yeast three-hybrid system. *Biochem. Biophys. Res. Commun.* **277,** 518–523 (2000).
47. Mori, D., Sasagawa, N., Kino, Y. & Ishiura, S. Quantitative analysis of CUG-BP1 binding to RNA repeats. *J. Biochem. (Tokyo)* **143,** 377–383 (2008).
48. Dembowski, J. A. & Grabowski, P. J. The CUGBP2 Splicing Factor Regulates an Ensemble of Branchpoints from Perimeter Binding Sites with Implications for Autoregulation. *PLoS Genet.* **5,** (2009).
49. Suzuki, H. *et al.* Alternative splicing produces structural and functional changes in CUGBP2. *BMC Biochem.* **13,** 6 (2012).
50. Miyazaki, Y. *et al.* Viral delivery of miR-196a ameliorates the SBMA phenotype via the silencing of CELF2. *Nat. Med.* **18,** 1136–1141 (2012).
51. Subramaniam, D. *et al.* RNA Binding Protein CUGBP2/CELF2 Mediates Curcumin-Induced Mitotic Catastrophe of Pancreatic Cancer Cells. *PLoS ONE* **6,** (2011).
52. Ramalingam, S. *et al.* Novel intestinal splice variants of RNA-binding protein CUGBP2: isoform-specific effects on mitotic catastrophe. *Am. J. Physiol. Gastrointest. Liver Physiol.* **294,** G971–G981 (2008).
53. Subramaniam, D. *et al.* Translation inhibition during cell cycle arrest and apoptosis: Mcl-1 is a novel target for RNA binding protein CUGBP2. *Am. J. Physiol. - Gastrointest. Liver Physiol.* **294,** G1025–G1032 (2008).
54. Ladd, A. N., Charlet-B., N. & Cooper, T. A. The CELF Family of RNA Binding Proteins Is Implicated in Cell-Specific and Developmentally Regulated Alternative Splicing. *Mol. Cell. Biol.* **21,** 1285–1296 (2001).
55. Charlet-B, N., Singh, G., Cooper, T. A. & Logan, P. Dynamic Antagonism between ETR-3 and PTB Regulates Cell Type-Specific Alternative Splicing. *Mol. Cell* **9,** 649–658 (2002).
56. Goo, Y.-H. & Cooper, T. A. CUGBP2 directly interacts with U2 17S snRNP components and promotes U2 snRNA binding to cardiac troponin T pre-mRNA. *Nucleic Acids Res.* **37,** 4275–4286 (2009).
57. Dujardin, G. *et al.* How Slow RNA Polymerase II Elongation Favors Alternative Exon Skipping. *Mol. Cell* **54,** 683–690 (2014).
58. Zhang, W., Liu, H., Han, K. & Grabowski, P. J. Region-specific alternative splicing in the nervous system: implications for regulation by the RNA-binding protein NAPOR. *RNA* **8,** 671–685 (2002).
59. Gromak, N., Matlin, A. J., Cooper, T. A. & Smith, C. W. J. Antagonistic regulation of α-actinin alternative splicing by CELF proteins and polypyrimidine tract binding protein. *RNA* **9,** 443–456 (2003).
60. Zhou, X. *et al.* Transcriptome analysis of alternative splicing events regulated by SRSF10 reveals position-dependent splicing modulation. *Nucleic Acids Res.* **42,** 4019–4030 (2014).

61. Erkelenz, S. *et al.* Position-dependent splicing activation and repression by SR and hnRNP proteins rely on common mechanisms. *RNA* **19,** 96–102 (2013).
62. Ishigaki, S. *et al.* Position-dependent FUS-RNA interactions regulate alternative splicing events and transcriptions. *Sci. Rep.* **2,** (2012).
63. Llorian, M. *et al.* Position-dependent alternative splicing activity revealed by global profiling of alternative splicing events regulated by PTB. *Nat. Struct. Mol. Biol.* **17,** 1114–1123 (2010).
64. Wang, Z. *et al.* iCLIP Predicts the Dual Splicing Effects of TIA-RNA Interactions. *PLoS Biol.* **8,** (2010).
65. Mallory, M. J. *et al.* Induced transcription and stability of CELF2 mRNA drives widespread alternative splicing during T-cell signaling. *Proc. Natl. Acad. Sci.* **112,** E2139–E2148 (2015).
66. Jesse, S., Koenig, A., Ellenrieder, V. & Menke, A. Lef-1 isoforms regulate different target genes and reduce cellular adhesion. *Int. J. Cancer* NA–NA (2009). doi:10.1002/ijc.24802
67. Kahler, R. A. *et al.* Lymphocyte enhancer-binding factor 1 (Lef1) inhibits terminal differentiation of osteoblasts. *J. Cell. Biochem.* **97,** 969–983 (2006).
68. Ishitani, T., Matsumoto, K., Chitnis, A. B. & Itoh, M. Nrarp functions to modulate neural-crest-cell differentiation by regulating LEF1 protein stability. *Nat. Cell Biol.* **7,** 1106–1112 (2005).
69. Mao, C. D. & Byers, S. W. Cell-context dependent TCF/LEF expression and function: alternative tales of repression, de-repression and activation potentials. *Crit. Rev. Eukaryot. Gene Expr.* **21,** (2011).
70. Arce, L., Yokoyama, N. N. & Waterman, M. L. Diversity of LEF/TCF action in development and disease. *Oncogene* **25,** 7492–7504 (2006).
71. Hovanes, K., Li, T. W. H. & Waterman, M. L. The human LEF-1 gene contains a promoter preferentially active in lymphocytes and encodes multiple isoforms derived from alternative splicing. *Nucleic Acids Res.* **28,** 1994–2003 (2000).
72. Giese, K., Kingsley, C., Kirshner, J. R. & Grosschedl, R. Assembly and function of a TCR alpha enhancer complex is dependent on LEF-1-induced DNA bending and multiple protein-protein interactions. *Genes Dev.* **9,** 995–1008 (1995).
73. Kuo, C. T. & Leiden, J. M. Transcriptional regulation of T lymphocyte development and function. *Annu. Rev. Immunol.* **17,** 149–187 (1999).
74. Topp, J. D., Jackson, J., Melton, A. A. & Lynch, K. W. A cell-based screen for splicing regulators identifies hnRNP LL as a distinct signal-induced repressor of CD45 variable exon 4. *RNA* **14,** 2038–2049 (2008).
75. Cooper, T. A. Use of minigene systems to dissect alternative splicing elements. *Methods San Diego Calif* **37,** 331–340 (2005).
76. Baier, G. The PKC gene module: molecular biosystematics to resolve its T cell functions. *Immunol. Rev.* **192,** 64–79 (2003).
77. Sharma, S., Falick, A. M. & Black, D. L. Polypyrimidine Tract Binding Protein Blocks the 5′ Splice Site Dependent Assembly of U2AF and the Prespliceosomal E Complex. *Mol. Cell* **19,** 485–496 (2005).
78. Sharma, S., Maris, C., Allain, F. H.-T. & Black, D. L. U1 snRNA Directly Interacts with Polypyrimidine Tract Binding Protein During Splicing Repression. *Mol. Cell* **41,** 579–588 (2011).
79. Chou, M.-Y., Underwood, J. G., Nikolic, J., Luu, M. H. & Black, D. L. Multisite RNA binding and release of polypyrimidine tract binding protein during the regulation of c-src neural-specific splicing. *Mol. Cell* **5,** 949–957 (2000).

80. Keppetipola, N., Sharma, S., Li, Q. & Black, D. L. Neuronal regulation of pre-mRNA splicing by polypyrimidine tract binding proteins, PTBP1 and PTBP2. *Crit. Rev. Biochem. Mol. Biol.* **47,** 360–378 (2012).
81. Weyn-Vanhentenryck, S. M. *et al.* HITS-CLIP and integrative modeling define the Rbfox splicing-regulatory network linked to brain development and autism. *Cell Rep.* **6,** 1139–1152 (2014).
82. Sun, S., Zhang, Z., Fregoso, O. & Krainer, A. R. Mechanisms of activation and repression by the alternative splicing factors RBFOX1/2. *RNA* **18,** 274–283 (2012).
83. Hafner, M. *et al.* Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141,** 129–141 (2010).
84. Faustino, N. A. & Cooper, T. A. Identification of Putative New Splicing Targets for ETR-3 Using Sequences Identified by Systematic Evolution of Ligands by Exponential Enrichment. *Mol. Cell. Biol.* **25,** 879–887 (2005).
85. Leroy, O. *et al.* ETR-3 represses Tau exons 2/3 inclusion, a splicing event abnormally enhanced in myotonic dystrophy type I. *J. Neurosci. Res.* **84,** 852–859 (2006).
86. Snijders, A. P. *et al.* Arginine methylation and citrullination of splicing factor proline- and glutamine-rich (SFPQ/PSF) regulates its association with mRNA. *RNA* **21,** 347–359 (2015).
87. Keshwani, M. M. *et al.* Conserved proline-directed phosphorylation regulates SR protein conformation and splicing function. *Biochem. J.* **466,** 311–322 (2015).
88. Vu, N. T. *et al.* hnRNP U Enhances Caspase-9 Splicing and Is Modulated by AKT-dependent Phosphorylation of hnRNP L. *J. Biol. Chem.* **288,** 8575–8584 (2013).
89. Choudhary, C. *et al.* Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325,** 834–840 (2009).
90. Nakka, K. K. *et al.* Nuclear matrix-associated protein SMAR1 regulates alternative splicing via HDAC6-mediated deacetylation of Sam68. *Proc. Natl. Acad. Sci.* **112,** E3374–E3383 (2015).
91. Abdelmohsen, K. *et al.* Phosphorylation of HuR by Chk2 Regulates SIRT1 Expression. *Mol. Cell* **25,** 543–557 (2007).
92. Gonçalves, V. *et al.* Phosphorylation of SRSF1 by SRPK1 regulates alternative splicing of tumor-related Rac1b in colorectal cells. *RNA* **20,** 474–482 (2014).
93. Bräuer, U., Zaharieva, E. & Soller, M. Regulation of ELAV/Hu RNA-binding proteins by phosphorylation: Figure 1. *Biochem. Soc. Trans.* **42,** 1147–1151 (2014).
94. Liu, Y. *et al.* Phosphorylation of the alternative mRNA splicing factor 45 (SPF45) by Clk1 regulates its splice site utilization, cell migration and invasion. *Nucleic Acids Res.* **41,** 4949–4962 (2013).
95. Salisbury, E. *et al.* Ectopic expression of cyclin D3 corrects differentiation of DM1 myoblasts through activation of RNA CUG-binding protein, CUGBP1. *Exp. Cell Res.* **314,** 2266–2278 (2008).
96. Beisang, D., Rattenbacher, B., Vlasova-St. Louis, I. A. & Bohjanen, P. R. Regulation of CUG-binding Protein 1 (CUGBP1) Binding to Target Transcripts upon T Cell Activation. *J. Biol. Chem.* **287,** 950–960 (2012).
97. Xu, K., Kitchen, C. M., Shu, H.-K. G. & Murphy, T. J. Platelet-derived Growth Factor-induced Stabilization of Cyclooxygenase 2 mRNA in Rat Smooth Muscle Cells Requires the c-Src Family of Protein-tyrosine Kinases. *J. Biol. Chem.* **282,** 32699–32709 (2007).
98. Gnad, F., Gunawardena, J. & Mann, M. PHOSIDA 2011: the posttranslational modification database. *Nucleic Acids Res.* **39,** D253–D260 (2011).

99. Singh, G., Charlet-B., N., Han, J. & Cooper, T. A. ETR-3 and CELF4 protein domains required for RNA binding and splicing activity in vivo. *Nucleic Acids Res.* **32,** 1232–1241 (2004).
100. Hall, M. P. *et al.* Quaking and PTB control overlapping splicing regulatory networks during muscle cell differentiation. *RNA* **19,** 627–638 (2013).
101. Chiou, N.-T., Shankarling, G. & Lynch, K. W. hnRNP L and hnRNP A1 induce extended U1 snRNA interactions with an exon to repress spliceosome assembly. *Mol. Cell* **49,** 972–982 (2013).
102. Yarosh, C. A. *et al.* TRAP150 interacts with the RNA-binding domain of PSF and antagonizes splicing of numerous PSF-target genes in T cells. *Nucleic Acids Res.* gkv816 (2015). doi:10.1093/nar/gkv816
103. Crawford, J. B. & Patton, J. G. Activation of α-Tropomyosin Exon 2 Is Regulated by the SR Protein 9G8 and Heterogeneous Nuclear Ribonucleoproteins H and F. *Mol. Cell. Biol.* **26,** 8791–8802 (2006).
104. Miro, J. *et al.* FUBP1: a new protagonist in splicing regulation of the DMD gene. *Nucleic Acids Res.* **43,** 2378–2389 (2015).
105. Mauger, D. M., Lin, C. & Garcia-Blanco, M. A. hnRNP H and hnRNP F Complex with Fox2 To Silence Fibroblast Growth Factor Receptor 2 Exon IIIc. *Mol. Cell. Biol.* **28,** 5403–5419 (2008).
106. Caputi, M. & Zahler, A. M. SR proteins and hnRNP H regulate the splicing of the HIV-1 tev-specific exon 6D. *EMBO J.* **21,** 845–855 (2002).
107. Kino, Y. *et al.* MBNL and CELF proteins regulate alternative splicing of the skeletal muscle chloride channel CLCN1. *Nucleic Acids Res.* **37,** 6477–6490 (2009).
108. Ho, T. H. *et al.* Muscleblind proteins regulate alternative splicing. *EMBO J.* **23,** 3103–3112 (2004).
109. Wang, E. T. *et al.* Antagonistic regulation of mRNA expression and splicing by CELF and MBNL proteins. *Genome Res.* gr–184390 (2015).
110. Izquierdo, J. M. Heterogeneous ribonucleoprotein C displays a repressor activity mediated by T-cell intracellular antigen-1-related/like protein to modulate Fas exon 6 splicing through a mechanism involving Hu antigen R. *Nucleic Acids Res.* **38,** 8001–8014 (2010).
111. Nasrin, F. *et al.* HnRNP C, YB-1 and hnRNP L coordinately enhance skipping of human MUSK exon 10 to generate a Wnt-insensitive MuSK isoform. *Sci. Rep.* **4,** 6841 (2014).
112. Chou, M.-Y., Rooke, N., Turck, C. W. & Black, D. L. hnRNP H is a component of a splicing enhancer complex that activates a c-src alternative exon in neuronal cells. *Mol. Cell. Biol.* **19,** 69–77 (1999).
113. Fogel, B. L. & McNally, M. T. A Cellular Protein, hnRNP H, Binds to the Negative Regulator of Splicing Element from Rous Sarcoma Virus. *J. Biol. Chem.* **275,** 32371–32378 (2000).
114. Mandler, M. D., Ku, L. & Feng, Y. A cytoplasmic quaking I isoform regulates the hnRNP F/H-dependent alternative splicing pathway in myelinating glia. *Nucleic Acids Res.* **42,** 7319–7329 (2014).
115. Han, J. & Cooper, T. A. Identification of CELF splicing activation and repression domains in vivo. *Nucleic Acids Res.* **33,** 2769–2780 (2005).
116. Lynch, K. W. & Weiss, A. A CD45 polymorphism associated with multiple sclerosis disrupts an exonic splicing silencer. *J. Biol. Chem.* **276,** 24341–24347 (2001).

117. Rothrock, C., Cannon, B., Hahm, B. & Lynch, K. W. A conserved signal-responsive sequence mediates activation-induced alternative splicing of CD45. *Mol. Cell* **12,** 1317–1324 (2003).