ROLE OF HNRNP H IN A REVERSIBLE FINE-TUNING MECHANISM OF SPLICING RESPONSE UPON TRANSCRIPTIONAL PAUSING

by

Erkan Bayir

B.Sc. in Molecular Biology and Genetics, Bilkent University, 2009

Submitted to the Graduate Faculty of the

Kenneth P. Dietrich School of Arts and Sciences

in partial fulfillment of the requirements for the degree of

Master of Science

University of Pittsburgh

2016

UNIVERSITY OF PITTSBURGH

KENNETH P. DIETRICH SCHOOL OF ARTS AND SCIENCES

This thesis was presented

by

Erkan Bayir

It was defended on

July 7, 2016

and approved by

Dr. Craig L. Peebles, Professor, Department of Biological Sciences

Dr. Karen M. Arndt, Professor, Department of Biological Sciences

Dr. James M. Pipas, Professor, Department of Biological Sciences

Dr. Christine Milcarek, Professor, Department of Immunology, School of Medicine

Thesis Advisor: Dr. Paula Grabowski, Professor, Department of Biological Sciences

ROLE OF HNRNP H IN A REVERSIBLE FINE-TUNING MECHANISM OF SPLICING RESPONSE UPON TRANSCRIPTIONAL PAUSING

Erkan Bayir, M.Sc.

University of Pittsburgh, 2016

Alternative splicing is a common way of diversifying the proteome and regulating protein functions without increasing genome size. Chemotherapeutic agents, such as cisplatin, cause changes in the splicing of proliferation-associated exons, which contribute to the overall proliferative phenotype of cancer cells. Transcriptional pausing, which is seen in embryonic stem cells as a timing regulator of gene expression, also impacts splicing. The mechanisms by which transcriptional pausing and pause-release alter the functions of splicing machinery are largely unknown. Thus, I tested the effects of the transcription elongation inhibitor DRB (5,6-dichloro-1-β-D-ribofuranosylbenzimidazole) on the splicing of a set of inducible exons with a common UAGG-GGGG splicing code. These exons were previously identified through a genomewide search for the transcripts sensitive to hnRNP A1 and hnRNP H. I found that the endogenous human HNRNPH1 Exon 4 and its paralog HNRNPH3 Exon 3 had increased skipping upon DRB treatment, as did the pA^+ reporter-cloned GRIN1 CI, a brain-region-specific rat exon with this splicing code, previously studied in our lab. I used single and combined mutations to monitor the effect of DRB-induced transcriptional pausing and pause-release on splicing, since the shared motifs include two exonic UAGG motifs and a GGGG close to 5´ splice site. GGGG

mutants had increased skipping, compared to wild type, upon DRB treatment. UAGG-GGGG code is necessary for the response, as the triple mutant lost its ability to respond to DRB. These data point to a role of hnRNP H in response to transcription stress.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to thank my parents, for their endless support to my quest for science and learning, and for always being on my side in my life. I crossed the boundaries of my neighborhood, my hometown, my country, my continent and the ocean to do this work.

I would like to thank Dr. Paula Grabowski for scientific supervision and support at all parts of this project, and for her personal help. It would not be possible without her. Thanks very much to my committee members Dr. Craig Peebles, Dr. Karen Arndt, Dr. Jim Pipas, and Dr. Christine Milcarek for their guidance and help for shaping this work.

I am grateful to Pipas Lab, Arndt Lab and Saunders Lab for sharing some of the materials and help for the project. I would like to thank Dr. Kyounga Han, Dr. Ping An and Dr. Jill Dembowski for the previous work and tools they produced, partially forming the foundation of this study. I would like to thank my former student, lab mate and friend Chad Stein for being with me for three years.

Anneme ve babama …

To my mom and dad …

1.0 INTRODUCTION

Splicing is an important mechanism in the processing of eukaryotic pre-mRNAs, in which the non-protein-coding introns are removed and exons are spliced, thereby producing the mature mRNA. Splicing is controlled by different cis- and trans-acting elements, namely the enhancer and silencer RNA motifs acting in cis, and splicing factors acting in trans, to regulate the inclusion or skipping decision of an exon and defining the intron and exon boundaries.

Splicing factors are the proteins encoded by particular genes, and changes in their expression alter the splicing of a transcript under the control of these splicing factors. The stoichiometry of splicing factors in a cell can be different across tissues, and developmental stages, can differ before and after signal transduction, all of which result in multiple alternatively spliced isoforms of the same transcript (Stamm, 2002; Lipscombe, 2005; Boutz *et al.*, 2007; Zhang *et al.*, 2008).

Alternatively spliced isoforms of the same transcript give rise to different proteins, regulating the structure and function of proteins as well as diversifying the proteome without the need to increase the size of the genome. Since splicing factors are also encoded by genes, the transcripts of splicing factors are also subject to alternative splicing. The processes of transcribing and splicing a pre-mRNA transcript are spatially and temporally linked, suggesting a crosstalk or common regulation of both processes.

Splicing is regulated in response to a plethora of inputs, and this regulation provides a strategy for cells to adapt to environmental challenges and respond to different types of cellular stress.

1.1 SPLICING AND TRANSCRIPTION

Transcription of DNA into pre-mRNA by the RNA polymerase II enzyme has initiation, elongation and termination steps; which are all coupled to RNA processing. The coupling of transcription and RNA processing is especially important in splicing, a complex molecular mechanism of RNA modification in eukaryotic cells. Splicing and transcription are coupled, due to the C-terminal domain (CTD) of RNA polymerase II.

The phosphorylation of serines in the CTD is important for the movement of RNA polymerase II on the DNA (Phatnani and Greenleaf, 2006) and it also associates some of the splicing factors that are transferred from CTD onto the nascent RNA (David *et al.*, 2011; Hsin and Manley, 2012). The transcription-splicing coupling and the effect of this coupling on the splicing of several genes raised questions about whether there are proliferation-associated genes among the targets of this coupling, that might have an effect on proliferation, differentiation or the overall phenotype.

1.1.1 RNA Polymerase II C-terminal domain and coupling

A comprehensive proteomic analysis of immunopurified human RNAP II-coupled factors by mass spectrometry revealed that SR proteins and U1 snRNP are associated with

RNAP II, suggesting that SR family of splicing factors and U1 snRNP have a role in splicing-transcription coupling (Das *et al.*, 2007). The transcription initiation factor TFIIH was found to “specifically associate” with the spliceosome component U1 snRNP (Kwek *et al.*, 2002). The phosphorylation by TFIIH kinase CDK7 is also important due to its role in promoter-proximal pausing, and its inhibition by the specific inhibitor THZ1 causes the failure of capping and pausing (Nilson *et al.*, 2015). Many cancer types are responsive to THZ1 treatment, showing its potential as a cancer drug (Jeronimo *et al.*, 2016).

There are possible ways of crosstalk between transcription and splicing that are previously proposed, such as recruitment coupling (de la Mata *et al.*, 2006) and kinetic coupling (de la Mata *et al.*, 2003). Histone modifications, important in transcription, were found to be differentially associated with the splicing pattern of PTB-regulated exons, as another possible way of crosstalk between transcription and splicing (Luco *et al.*, 2010).

CDK12, which is capable of phosphorylating all serines in the RNA polymerase II CTD, was shown to have an affect on the alternative splicing of the SRSF1 gene (Liang *et al.*, 2015). Since the CDK12 is required for Ser2 phosphorylation in different species, and Ser7 phosphorylation stimulates CDK12 activity, it can constitute the link between the transcription and splicing (Jeronimo *et al.*, 2016).

Transcriptional stress can affect the splicing by severing the coupling between transcription and splicing. Human nuclear cofactor TCERG1 was suggested to be a possible coordinator of transcription and splicing (Montes *et al.*, 2012) and its mobility decreased upon DRB-induced transcription elongation inhibition (Sánchez-Hernández *et al.*, 2016). Splicing-transcription coupling appears as an important way of splicing control, causing changes in the splicing of many exons that are subject to regulation.

1.1.2 The role of cis-regulatory RNA elements in splicing

The splicing machinery needs to recognize the consensus sequences at the 5' and 3' splice sites and the branchpoint in order to define introns and exons. Most mammalian major-spliceosome introns start with GU at the 5' splice site, and end with an AG at the 3' splice site (Sheth *et al.*, 2006). Transcripts have intronic and exonic sequences, the cis-regulatory RNA elements, contributing to the selection of a particular splice site.

The enhancer and silencer sequences, depending on whether they are located within the intron or exon, are named as the intronic splicing enhancers (ISE), intronic splicing silencers (ISS), exonic splicing enhancers (ESE), and exonic splicing silencers (ESS). The copy number and the spatial arrangement of these short RNA elements is important, because the effect of regulatory RNA elements is dependent on the strength of the enhancer or silencer effect in a context-dependent way (Fu and Ares, 2014).

The combination of cis-regulatory RNA elements, or the splicing code, has a role in adjusting splicing patterns through interactions with the splicing factors (Wang *et al.*, 2008). When an RNA element is bound by trans-acting splicing factors, it has a role in the decision on whether an exon will be included or skipped.

1.1.3 Trans-acting splicing factors and their functions

Splicing factors are proteins that bind to RNA elements with some degeneracy via their RNA recognition motifs (RRM). The combined effect of cis-acting RNA elements and trans-acting splicing factors is important, because they act in an orchestrated way, not as one single cis-regulatory RNA element or one splicing factor causing the inclusion or

skipping of an exon. Therefore, the resulting pattern of splicing for a particular exon is an equilibrium that is subject to the composite effect of the splicing activators and splicing repressors, binding to the enhancer and silencer RNA elements.

Serine- and arginine-rich SR proteins are an important family of splicing factors, performing many functions. Spliceosome components U1 and U2 snRNPs are recruited to and stabilized at the splice sites by SR proteins (Feng *et al.*, 2008). Having an arginine- and serine-rich RS domain, SR proteins can come together with other RS-containing splicing factors and act as splicing enhancers (Chen and Manley, 2009). Their function depends on the sequence and location of the *cis* RNA motifs, and post-translational modifications like phosphorylation. Many SR proteins are phosphorylated and activated, and their dephosphorylation causes their nuclear export (Stamm, 2008).

Heterogeneous nuclear ribonucleoproteins (hnRNPs) are a family of splicing factors that have RNA recognition motifs (RRMs), and they often act as splicing repressors upon binding to silencer RNA sequences (Ladd and Cooper, 2002). Although many hnRNPs generally cause exon skipping, they sometimes favor an alternative cryptic or weak splice site (Matlin *et al.*, 2005). Splicing repressors can block the splicing factor binding to splicing enhancers, or block the splice sites by steric hindrance so that the spliceosome cannot be assembled (Chen and Manley, 2009). For instance, hnRNP A1 interferes with the U1 snRNP binding to the 5’ splice site (Eperon *et al.*, 2000). Close to the 3’ end of the CD45 exon 4, hnRNP A1 and U1 snRNP cause exon skipping in an hnRNP L-dependent manner (Chiou *et al.*, 2013). The effects of different splicing factors on a particular exon contribute to the final balance, leading to the inclusion versus skipping decision.

1.2 SPLICING-RELATED DISEASES: HEALTH IMPACTS

Pre-mRNA splicing is subject to regulation in response to different inputs, and the proteins that are made as a product of the aberrant splicing of certain transcripts can be detrimental to human health. Some of the aberrant splicing patterns were previously characterized in various diseases, such as neurodegenerative diseases and cancer.

There are cis-acting mutations on DNA that change the splice sites or create new splice sites, and there are trans-acting factors that can cause changes in splicing, which contribute to the development of diseases caused by aberrant splicing patterns (Ward and Cooper, 2010). The splicing change that causes a disease can be exploited to elucidate the splicing mechanism and develop a splicing therapy. If the aberrant splicing is reverted to the normal splicing pattern, the disease phenotype can be ameliorated.

There are different RNA-based therapeutic approaches for splicing-related diseases. The use of antisense oligonucleotides aims to revert the splicing from the splice isoform causing the disease to an isoform that rescues the phenotype (Sazani and Kole, 2003). RNA interference that specifically targets an aberrantly-spliced mRNA decreases the protein that contributes to disease, and small molecules that interfere with the splicing factors or RNA secondary structures can alter the aberrant splicing patterns associated with the disease (Cooper *et al.*, 2009; Scotti and Swanson, 2016).

1.2.1 Alternative splicing in cancer and early tumorigenesis

One of the most important diseases associated with aberrant splicing patterns is cancer. There are many cancer-related changes in splicing as a result of, or in the absence of

genomic mutations. Splice site mutations cause aberrant splicing in cancer (Venables, 2004). Splicing factors are subject to mutations that can cause cancer-related changes in splicing (Scotti and Swanson, 2016). The splicing of some genes is developmentally regulated, and they are preferentially re-expressed in cancer (David and Manley, 2010). Mutations close to splice sites can lead to aberrant splicing of tumor suppressor genes, promoting cancer (Supek *et al.*, 2014). Tumorigenesis-associated exons with similar splicing changes were identified in several human tumors (Danan-Gotthold *et al.*, 2015).

Cancer cells are not individual cells that are isolated from their environment; they rather co-evolve altogether upon exposure to different kinds of stresses. There are interactions between cancer cells, and the evolution of the tumor is a collective and programmed response performed by interacting cells (Lambert *et al.*, 2011). The external stress is beneficial to the tumor cell population as a whole, and possible changes in splicing caused by different kinds of stresses are exploited as a strategy for sustaining and expanding the tumor.

The link between cancer and splicing is extensively studied, and these studies aim to identify a set of cancer cell-specific splice isoforms along with significant changes in splicing associated with early tumorigenesis. The cancer-related changes in splicing are important in early tumorigenesis. Some of the causes of early tumorigenesis, such as genotoxic stress, possibly have a significant impact on splicing. The elucidation of alternative splicing in early tumorigenesis can be exploited to develop an RNA-based system of therapy to reprogram splicing and slow down cancer progression.

The pyruvate kinase M2 (PKM2) isoform, a product of alternative splicing, is subject to allosteric regulation by fructose-1,6-bisphosphate and phosphotyrosines, that causes the pyruvate kinase enzyme (encoded by PKM2) to be dimerized and diverting the pathway to the synthesis of biomolecules, which is needed by rapidly-dividing fetal and cancer cells (Christofk *et al.*, 2008). The glycolytic metabolism was proposed to be an adaptation of cancer cells during early tumorigenesis, in which blood vessels are not formed yet, and the oxygen is scarce (Cairns *et al*., 2011).

EGFR, CD44, NER, tyrosine hydroxylase, lactate dehydrogenase, cadherin-11, fibronectin, and Brn-3a were reported to have cancer-associated splice variants, and it was suggested that splicing changes might be widespread in human cancers (Xu and Lee, 2003). Bcl-x alternative splicing gives rise to a pro-apoptotic Bcl-x_S and an anti-apoptotic Bcl-x_L isoforms, and high Bcl-x_L expression was associated with many cancers as a possible way of cancer cells to escape apoptosis (Bauman and Kole, 2011). Cyclin D1, H-Ras, and VEGFA pre-mRNA transcripts are subject to alternative splicing with respect to apoptosis and cancer (David and Manley, 2010). Cancer-related splicing defects can be targeted to revert splicing by the use of RNA oligonucleotides.

An alternatively spliced variant of the gene encoding spleen tyrosine kinase was found to be frequently expressed in breast cancer tissues, and this splice variant was never found in normal mammary tissues (Wang *et al.*, 2003). SF3B1, encoding a core member of U2 snRNP complex, was observed to be mutated in 15% of 91 patients that have chronic lymphocytic leukemia, which shows that spliceosome components might also be subject to mutations in cancer (Wang *et al.*, 2011). More commonly observed splice variants can be used as markers for different types of cancers.

1.2.2 Neurological disorders and aberrant splicing patterns

There are several neurological disorders, which result from aberrant splicing in neuronal cells. For instance, a single nucleotide substitution in SMN2 (survival of motor neurons) gene, which encodes the SMN protein, causes spinal muscular atrophy disease due to exon skipping (Lefevbre *et al.*, 1995). SMN complex is important for snRNP biogenesis, and the abnormal splicing of SMN2 can affect splicing globally (Cooper *et al.*, 2009).

A deletion in the ATM gene intron 20 causes a cryptic exon to be included, and this splicing change causes ataxia-telengiectasia (Pagani *et al.*, 2002). Duchenne muscular dystrophy is also among the diseases partially caused by splicing defects in dystrophin gene, and a nonsense mutation resulting in partial skipping of exon 31 was identified to contribute to the disease (Disset *et al.*, 2006).

Cystic fibrosis transmembrane conductance regulator (CFTR) splicing defects caused by a tryptophan deletion were seen in about half of the cystic fibrosis patients in the US (Cooper *et al*., 2009; Scotti and Swanson, 2016). Neurological disease-related splicing changes provided the opportunity to elucidate the splicing mechanisms of these transcripts, and possible ways to exploit the splicing machinery to develop therapies.

1.3 ALTERNATIVE SPLICING, STRESS AND PAUSE-RELEASE

The splicing of many genes is altered in response to different kinds of cellular and metabolic stress such as oxidative stress, heat shock, and hypoxia (Biamonti and Caceres, 2009). Upon heat shock, various splicing factors are localized in nuclear

stress bodies, the active transcription sites that are suggested to act as rapid inducers of gene expression changes involving chromatin remodeling, transcription and splicing factor recruitment (Biamonti and Vourc'h, 2010). The phosphorylation of SR proteins is monitored upon heat and osmotic stress through cyclin dependent-like kinases Clk1/4 (Ninomiya *et al.*, 2011). Heat shock globally inhibits the splicing of posttranscriptionally spliced transcripts, whereas the transcripts including G-rich motifs in their introns are enriched among unaffected transcripts (Shalgi *et al.*, 2014).

Calcium (Ca^{2+}) signaling is particularly important for splicing in brain and neurons as well as with respect to the endoplasmic reticulum (ER) stress. An RNA element that is responsive to Ca^{2+}/calmodulin-dependent protein kinase (CaMK) is involved in the neuronal depolarization-induced splicing changes in GH_3 cells and neurons (Xie and Black, 2001; Xie *et al.*, 2005). HnRNP A1 is phosphorylated by MAPK and relocalized to stress granules upon osmotic stress, by which ionic concentration changes can affect splicing (Allemand *et al.*, 2005). Thus, cis-regulatory RNA elemens and trans-acting splicing factors can both be responsible for the Ca^{2+}-mediated regulation of splicing.

Genotoxic stress, especially by DNA damage inducing agents and irradiation, has an important role in splicing. Akt/PKB is particularly important in the response to genotoxic stress. ATM, ATR and DNA-PK activate Akt upon genotoxic stress (Viniegra *et al.*, 2004; Bozulic *et al.*, 2008). Activated Akt affects the SR splicing factors through SRPKs that phosphorylate SR proteins, and the EGF signaling regulates global splicing through Akt and SRPK (Zhou *et al.*, 2012).

UV-induced changes in splicing are widely studied, and they can affect exon inclusion or skipping via causing changes in hSlu7 subcellular localization (Shomron *et*

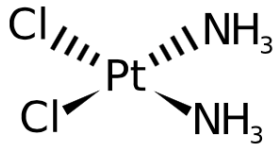
al., 2005), through RNA polymerase II CTD hyperphosphorylation (Muñoz *et al.*, 2009) or elevation of SRSF1 (Comiskey *et al.*, 2015).

Transcriptional stress can take place as a result of, or independent from genotoxic stress, and it also affects splicing. Transcriptional pausing and pause-release are likely to affect splicing due to the transcription-splicing coupling, or the splicing changes associated with transcriptional pausing and pause-release can be under the control of DNA damage response (DDR) pathways.

1.3.1 Alternative splicing upon genotoxic stress

Genotoxic stress can affect splicing by different ways; by post-translational modification of splicing factors, changing their cellular localization, transcriptional activation of splicing factors, or changing the rate of transcription (Thomas and Lieberman, 2013).

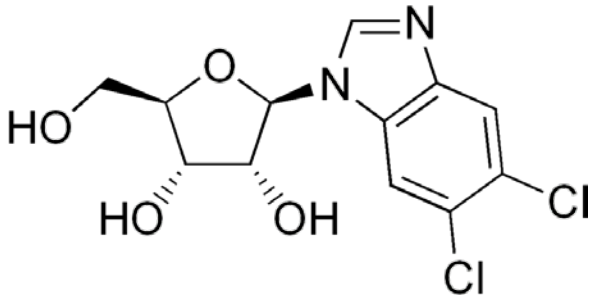
Cisplatin is a cytotoxic anticancer drug that targets DNA by forming adducts. It forms intrastrand (and to a lesser extent interstrand) DNA crosslinks between purines, and alters the 3D structure of the double helix by causing DNA bending (Jung *et al.*, 2007). Thus, cisplatin causes inhibition of replication and transcription.



cis-diamminedichloridoplatinum(II) (Cisplatin)

Cisplatin and doxorubicin, two well-known cancer chemotherapy drugs, cause a change in CDC25C C1/C5 splicing, to shift from C1 to C5 in MCF-7 cancer cells, due to the genotoxic effects of these drugs (Albert *et al.*, 2012). In a large-scale proteomic analysis of the proteins that are phosphorylated in response to DNA damage, many DNA damage response (DDR) signaling proteins and some of the splicing factors were found as targets of ATM and ATR kinase phosphorylation, and their phosphorylation by ATM and ATR can affect the splicing of their targets (Matsuoka *et al.*, 2007).

Cisplatin-induced DNA damage, which alters the double helix, is recognized by the cell and that initiates different DNA damage response pathways. The most important pathway that handles cisplatin-induced DNA damage is the “nucleotide excision repair pathway” (Basu and Krishnamurthy, 2010). Transcription-coupled repair is induced by the stalling of RNA polymerase II, and it is one of the ways that cells deal with cisplatin-induced genotoxic stress (Damsma *et al.*, 2007). The transcriptional elongation inhibitor, 5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole (DRB), damages the DNA and triggers homologous recombination repair (Stoimenov *et al.*, 2011). DRB-induced transcriptional pausing, similar to genotoxic stress, affects the splicing of many transcripts.



5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole (DRB)

One of the possible ways that cisplatin treatment affects splicing is the crosslinking of the RNA component of snRNPs. Aquated cisplatin was shown to cause the crosslinking of guanines on the opposite sides of an internal loop, which is derived from U2:U6 spliceosomal RNAs (Hostetter *et al.*, 2009).

There are a number of studies that focus on global proteomic changes in HeLa cells upon cisplatin treatment. The hnRNP H was found to be upregulated, whereas the hnRNP C was downregulated, in the nuclear extracts of cisplatin-treated HeLa cells (Wu *et al.*, 2010). In another study, the nuclear extracts of cisplatin-treated cells had a decrease in the hnRNP A1, hnRNP C1/C2 and hnRNP F levels, and an increase in the hnRNP K levels (Zhang *et al.*, 2012). The global or specific alterations of splicing as a result of cisplatin treatment can affect the proliferative phenotype of cells.

Genotoxic stress can affect splicing via transcriptional activation of splicing factors. The methylmethanesulfonate- and cyclophosphamide-induced genotoxic stress (caused by interstrand DNA crosslinks) were found to result in the upregulation of E2F1 and E2F1-dependent SRSF2 (SC35) upregulation, and it caused an increase in the proapoptotic spliced isoforms of caspase 9 and Bcl-x (Merdzhanova *et al.*, 2008). Thus, SRSF2 seems to be a candidate through which genotoxic stress can alter the splicing patterns of proliferation-associated transcripts.

Genotoxic stress-induced cotranscriptional exon skipping was observed in some transcripts, among which the genes encoding splicing factors are also present (Dutertre *et al.*, 2010; Solier *et al.*, 2010). Topoisomerase inhibitors, which inhibit topoisomerase I and cause the stalling of DNA replication and transcription, cause genotoxic stress and affect the alternative splicing. For instance, topoisomerase inhibitor camptothecin

caused the skipping of multiple exons in MDM2 transcripts independent of p53 (Dutertre *et al.*, 2010).

Another molecule, topoisomerase II inhibitor mitoxantrone, caused relocalization of hnRNP A1, SRSF1 and SRSF2 to nuclear stress granules, resulting in a change in CD44 splicing (Busa *et al.*, 2010). Topoisomerase 1 was recently found to be regulated through the CTD phosphorylation of RNA polymerase II during transcriptional pause release (Baranello *et al.*, 2016), which can be one of the links that connects the effects of genotoxic and transcriptional stress on alternative splicing.

1.3.2 Transcriptional inhibition and pause-release

Transcription is regulated in response to many different environmental or developmental inputs, to adapt to surrounding conditions. The time and scale of gene expression can be adjusted throughout development, in different stages and locations, under the effect of different signaling pathways. RNA polymerase II transcribes mRNA from genes, and it is one of the means through which the mRNA transcription is modulated.

Transcription has the initiation, elongation, and termination steps. In the initiation, RNA polymerase II binds to the promoter. In many genes, RNA polymerase II pauses at the promoter. The genes encoding the heat shock proteins (Hsp), which are important in the response to several types of cellular stresses, were among the first genes to be shown to have transcriptional pausing in *Drosophila* (Rasmussen *et al.*, 1993). Pausing was also observed in the proliferation-associated human *MYC* and *FOS* genes, which had RNA polymerase II pausing at promoter-proximal regions (Plet *et al.*, 1995).

Human and mouse embryonic stem cells have RNA polymerase II pausing in a significant portion of their genes, estimated to range from 30 to 90 per cent (Gilchrist *et al.*, 2010). Thirty per cent of all genes were found to have RNA polymerase II pausing in human embryonic stem cells (Guenther *et al.*, 2007), *Drosophila* and human primary cells, which seems to be the percentage of genes that show transcriptional pausing regardless of the species and cell type (Core *et al.*, 2008).

There are factors associated with RNA polymerase II promoter-proximal pausing that inhibit transcription elongation in vitro. DRB-sensitivity inducing factor (DSIF) and negative elongation factor (NELF), along with some additional factors have an effect on RNA polymerase II pausing (Narita *et al.*, 2003). NELF knockdown released the paused polymerase, supporting the role of NELF in transcriptional pausing (Muse *et al.*, 2007).

The pausing factors themselves are subject to phosphorylation by kinases, as a means of regulating the pausing of RNA polymerase II. The phosphorylation of NELF and DSIF by the positive transcription elongation factor b (P-TEFb) results in the release of NELF from the complex and turning DSIF into a factor favoring transcription elongation, making it a positive factor (Peterlin and Price, 2006).

Flavopiridol, which inhibits PTEF-b, prevents RNA polymerase II from processive elongation in vivo (Ni *et al.*, 2008). P-TEFb inhibition by flavopiridol blocked the pause-release in many genes in human embryonic stem cells (Rahl *et al.*, 2010), and it demonstrates the effect of these factors and their interactions.

There are some characteristics in common among the genes with transcriptional pause. Genes related to signal transduction pathways, development, proliferation, and

stress response were found to be enriched among the genes with RNA polymerase II pausing (Adelman and Lis, 2012).

The pluripotency genes in human embryonic stem cells are regulated through trancriptional pause release, and P-TEFb and CDK9 have a role in the activation of these genes in reprogramming, thereby causing changes in development (Liu *et al.*, 2014). The transcriptional pausing at the promoter-proximal regions of several genes is important in proliferation and stress response, because some of these factors themselves, such as the c-Myc, HSF, p53, and NF-κB, can contact P-TEFb and make it activate gene expression (McNamara *et al.*, 2016). Among these transcription factors, the role of c-Myc in the transcriptional pause release is evident in human embryonic stem cells, as c-Myc binds to P-TEFb, and the knockdown of c-Myc caused increased pausing in these cells (Rahl *et al.*, 2010).

This gives clues and raises questions about the possibility of the transcriptional pausing as a method of diversifying the transcriptome and proteome upon various environmental stresses. Genotoxic stress, which can impact transcriptional output, as well as the transcriptional stress itself, can affect RNA polymerase II pausing, which is likely to cause changes in alternative splicing.

1.3.3 Alternative splicing upon transcriptional stress

The coupling between transcription and splicing, and its effect on the splicing of several genes raised questions about whether there are alternatively spliced stress- or proliferation-associated transcripts that affect proliferation or the overall metabolic phenotype. The transcriptional stress induced by different means leads to similar global

outputs of alternative splicing, and the subset of transcripts that are sensitive to DRB also respond to UV exposure (Ip *et al.*, 2011).

Fibronectin EDI exon is among the well-studied exons that respond to the transcriptional stress in several cell types. DRB-induced transcription inhibition caused a 2-fold increase in EDI inclusion in Hep3B cells (Nogués *et al.*, 2002). The same exon responded similarly to low-dose UVC and cisplatin treatment in Hep3B cells, along with the increase in the proapoptotic spliced isoforms of two transcripts, Bcl-x and caspase 9 (Muñoz *et al.*, 2009). Different promoters may also have different effects on the splicing of EDI, and its inclusion is higher in wound healing and proliferating tissues.

P-TEFb can either be in its active form, that causes release of the paused RNA polymerase, or the inactive form that is bound by the 7SK noncoding RNA complex in the cell. The SR family members SRSF1 and SRSF2 were found to be part of the 7SK noncoding RNA complex and connect P-TEFb to RNA polymerase II (Ji *et al.*, 2013). Several hnRNP family members also bind the 7SK noncoding RNA, which increases the possibility of transcriptional stress and pausing to have an effect on the alternative splicing of the transcripts under the control of these splicing factors (Ji *et al.*, 2013).

Faster and slower transcriptional rates cause changes in the splicing of different genes. A subset of exons were documented to have increased inclusion upon slower RNA polymerase II transcription rates, and some of the exons are increasingly skipped upon transcriptional inhibition or in slower RNA polymerase II mutant strains (Fong *et al.*, 2014). The subset of exons with increased skipping upon DRB treatment were observed to be enriched in AT-rich sequences in their flanking introns, but there is also an enrichment of GU-rich sequences in the transcripts with higher exon inclusion upon

DRB treatment (Fong *et al.*, 2014). This model does not seem to be consistent, because there are exons with alternative splicing patterns that do not follow AT or GU enrichment. There are exons that do not respond to slow or fast transcription elongation rates. However, optimal transcription rate is needed for transcripts to have enough time for spliceosome assembly and splicing factor recruitment (Fong *et al.*, 2014).

Splicing is under tissue-specific and developmental regulation, and some of the slow-included or fast-included exons can have alternative splicing patterns that are correlated with the proliferative phenotype of the cell. The abundance and localization of splicing factors in the cell can also affect the splicing response to the changes in the rate of transcription.

Different signaling pathways that affect the proliferation of the cell also affect the rate of transcription. TNF-α and NF-κB were suggested to increase RNA polymerase II elongation at sites of transcriptional pausing among the target genes of this pathway, as a possible way of regulating the fate of their target transcripts (Danko *et al.,* 2013). The signaling pathways that affect transcription rates can be a way of adjusting the splicing, due to the spatial and temporal coupling between splicing and transcription in the cell.

The knockdown of the SR-family splicing factor SC35 was shown to cause abortive transcription elongation, providing a possible effect of splicing factors on the rate and processivity of transcription (Lin *et al.*, 2008). The effect of hnRNP A1/A2 knockdown and the effect of DRB-induced transcriptional inhibition were suggested to lead to similar results on gene transcription, due to a partial overlap in the repression of a similar set of transcripts (Lemieux *et al.*, 2015).

In the light of the findings related to genotoxic and transcriptional stress on the mRNA splicing, it is possible to speculate that the effect will be context- and sequence-dependent, and it can be controlled or regulated by different signaling pathways.

1.4 HNRNPH1 AND HNRNPH3: HNRNP H-ENCODING GENES

Heterogeneous nuclear ribonucleoproteins (hnRNPs) are a class of proteins that are mostly localized in the nucleus, some of which are also present in the cytoplasm. They are among the proteins with highest abundance in the cell, and there are many different types of hnRNPs in mammalian cells (Choi *et al.*, 1986). HnRNPs bind to RNA via RNA recognition motifs (RRMs), and this binding is sequence-specific, although the RNA motifs that are capable of binding to these RNA recognition motifs do not require a very strict sequence. The promiscuity of RRM and RNA motif binding is important, because one RNA motif can be bound by multiple hnRNPs or other splicing factors.

HnRNPs can have context-dependent effects on splicing, which vary due to the magnitude of the binding affinity and the location of the RNA motif with respect to the intron and exon. HnRNPs can usually act as splicing repressors by binding to silencer RNA motifs and preventing the spliceosome assembly, and they have additional functions such as protecting mRNAs from degradation, shuttling between nucleus and cytoplasm, and modulating transcript stability (Dreyfuss *et al.*, 1993). Thus, hnRNPs are worth investigating in detail, due to their several duties and effects on the transcriptome.

1.4.1 HnRNP H family of splicing factors and their target transcripts

HnRNP H belongs to a family of splicing factors that also contain the closely related but separate protein hnRNP F, both of which have ubiquitous expression in cells (Bent *et al.*, 1995). HnRNP H family consists of hnRNP H1, hnRNP H2 (H'), hnRNP H3 (2H9), and the hnRNP F (Caputi and Zahler, 2001). HnRNP H uses three glycine-rich highly-conserved quasi-RRMs (qRRMs) for binding RNA, and interacts with poly(G) stretches to affect splicing in a context-dependent manner (Dominguez *et al.*, 2006). HnRNP H was shown to bind to GGGA, and is also required for the interaction of U1 snRNP with the enhancer in HIV-1 tev-specific exon 6D (Caputi and Zahler, 2002).

HnRNP H family members are important in the coupling between transcription and splicing. For instance, hnRNP F interacts with the TATA-binding protein (TBP) and the C-terminal domain (CTD) of the RNA polymerase II (Yoshida *et al.*, 1999). The CTD can be used as a landing pad by the splicing factors to be transferred onto the newly emerging nascent transcript, in order to perform their functions in splicing.

HnRNP H is implicated in a number of diseases, especially in different cancers. HnRNP H is overexpressed in pancreatic, hepatocellular, and gastric carcinomas, with an increase in the cell nucleus in these cancer types (Honoré *et al.*, 2004). In gliomas, hnRNP H levels are increased, and this upregulation is involved in a shift to more malignant and aggressive phenotype (LeFave *et al.*, 2011). HnRNP H1 has an effect on the splicing of HER2 in breast cancer cells, and HER2 is a special gene that is present in HER2-positive breast cancer (Gautrey *et al.*, 2015).

HnRNP H and hnRNP F bind to G-rich sequence motifs for exerting their effect on splicing, which usually act as splicing repressors. HnRNP H2 and hnRNP F were previously shown to bind to similar sequences, albeit in different ways, to affect gene expression (Alkan *et al.*, 2006). HnRNP F was shown to bind only to the single-stranded G stretches, and this binding was proposed to obstruct the formation of G-quadruplex structures (Samatanga *et al.*, 2012). Poly(G) stretches, especially the GGGG motifs act as silencers, and they are bound by hnRNP H and hnRNP F (Sohail *et al.*, 2014). In a global analysis of splicing, the poly(G) stretches were found to favor exon skipping, and the binding of hnRNP H required a continuous G stretch (Rahman *et al.*, 2015).

The effect of hnRNP H on the splicing of proliferation-associated transcripts is important, due to its effect on the several important genes related to apoptosis. Bcl-x is one of the most well-known apoptosis-related genes, and hnRNP H causes the splicing of Bcl-x to favor the pro-apoptotic Bcl-x_S, rather than the anti-apoptotic Bcl-x_L, in HeLa cells (Garneau *et al.*, 2005). The 30-nucleotide G-rich stretch located in the Bcl-x was shown to be important in the response of Bcl-x splicing to the Protein Kinase C (PKC) inhibitor Ro, giving rise to a pro-apoptotic phenotype upon Ro treatment (Hai *et al.*, 2008). On the other hand, hnRNP H also has an anti-apoptotic effect on the MAP kinase pathway, by causing an increase in the A-Raf (Rauch *et al.*, 2010). This regulatory mechanism can be a fine-tuning response of multiple inputs, and it can affect the proliferation versus apoptosis decision of the cells upon different signals and environmental conditions.

1.4.2 Alternative splicing of HNRNPH1 and HNRNPH3 transcripts

HnRNP H proteins are encoded by HNRNPH1, HNRNPH2, and HNRNPH3 paralogous genes in human (Flicek *et al.*, 2012). Their mouse orthologs are Hnrnph1, Hnrnph2, and Hnrnph3 respectively. It is a very highly-conserved family, with orthologs in various organisms ranging from different fish species to fruitfly, dolphin to xenopus, sheep to cat. Their conservation is an obvious sign of crucial functions for their protein products, and the similarities of the alternative splicing patterns across orthologs and paralogs provide hints about possible gene duplication events in the evolution of this gene family.

HNRNPH1 and HNRNPH3 paralogous transcripts, which share easily-identifiable characteristics, have several different 5' and 3' untranslated regions (UTR) with varying lengths, suggesting tissue-specific regulation of their protein products depending on the needs of different types of cells. The arrangement of their exons has some similarities, encompassing a 139-base pair cassette exon (HNRNPH1 Exon 4 and HNRNPH3 Exon 3) with a high conservation and sequence identity (Grabowski, 2004).

The alternatively-spliced HNRNPH1 E4 and HNRNPH3 E3 cassettes correspond to parts of the RNA recognition motif domain in the hnRNP H1 and hnRNP H3 protein products, and they seem to be skipped in the non-coding transcripts, which undergo nonsense-mediated decay across several human and mouse orthologs. Interestingly, these alternatively-spliced cassette exons share a guanine stretch in the immediately downstream flanking intron (Grabowski, 2004), and this common putative *cis* regulatory RNA motif is bound by the hnRNP H protein (Stein, 2015). The direct binding between the transcript and its protein product suggests an autoregulatory loop, affecting the fate of the alternative splicing of its own transcript.

The presence of the GGGG motif in the immediately downstream flanking intron is a characteristic that was previously identified as part of a splicing code, found in a genomewide search for a common combination of exonic UAGG and intronic GGGG (Han *et al.*, 2005). This splicing code was initially identified in the NMDA R1 receptor-encoding GRIN1, a neuronal transcript explained in the next section.

1.5 NMDA R1 RECEPTOR AND THE GRIN1 GENE

1.5.1 The biology of NMDA R1 receptor and GRIN1 splicing

The NMDA (N-methyl-D-aspartate)-type receptors are glutamate-gated Ca^{2+} channels found in mammals, and they have many different functions in the strength of synapses, sensitivity of the postsynaptic membranes, and neuronal cell death in some instances (Aoki *et al.*, 1994). These receptors mediate the flow of calcium ions, and the calcium is involved in several signaling pathways in the cell, leading to synaptic plasticity within the context of memory and learning (Traynelis *et al.*, 2010).

NMDA receptors are shown to affect the synaptic connectivity, learning from the environmental experiences, action selection, and behavior (Lambot *et al.*, 2016). They were recently shown to promote the survival of neurons in cochlea (Zhang-Hooks *et al.*, 2016). The location-dependent effects of the NMDA receptors depend on where they are expressed in the brain.

NMDA receptors have multiple subunits, belonging to the R1, R2 or R3 classes of receptors, and each subunit has several alternatively-spliced isoforms (Cull-Candy *et al.*, 2001). The R1 receptor expression is low in several rat brain regions at birth, and its levels increase after three weeks (Luo *et al.*, 1996). Different NMDA receptor variants arise due to alternative splicing, and they have different functions as a result of diverse structures (Nakanishi *et al.*, 1992). Therefore, the alternative splicing of NMDA receptor subtypes contribute to the neuronal plasticity, diversity and functions in different regions of the brain. The alternative splicing can rapidly change upon changing environmental conditions, which is possibly a strategy for the brain to adapt during learning.

1.5.2 Alternative splicing of GRIN1 in different brain regions

GRIN1 has three alternatively spliced exons that give rise to eight isoforms, and these isoforms have different functions (Zukin and Bennett, 1995). The NI and CI cassette exons of the GRIN1 transcript, both of which are alternatively spliced, encode part of the extracellular and intracellular domains of the receptor respectively (Bradley *et al.*, 2006), and function in the sensitivity and the localization of the receptor (Ehlers *et al.*, 1995; Okabe *et al.*, 1999). Alternative splicing appears as a strategy to adjust the function and location of the receptor, adapting to different conditions.

The splicing of NI and CI are controlled by the brain region-specific CUGBP2 (or NAPOR), which was shown to cause CI inclusion and NI skipping in the forebrain, and with the exact opposite effect due to the absence of CUGBP2 in the hindbrain (Zhang *et al.*, 2002). The compartmentalization of the different versions of NMDA R1 receptor is important, as the extracellular domain partially encoded by NI plays a role in adjusting

the sensitivity of the receptor to different ions, and the CI-encoded intracellular domain is involved in the membrane localization and intracellular signaling (Cull-Candy *et al.*, 2001). The CI exon is particularly important, due to the presence of a GGGG motif that is present in the flanking intron immediately downstream of the exon (Han *et al.*, 2005).

The splicing of CI cassette exon is controlled by the splicing enhancer sequences for the binding of ASF/SF2, SC35, and SRp40 within the exon, and the intronic sites for hnRNP H and NAPOR/CUGBP2, with the presence of two exonic UAGG motifs and an intronic GGGG motif (Han *et al.*, 2005). Both the intronic and exonic RNA motifs, along with the splicing factors binding to these motifs, contribute to the regulation of the CI splicing, fine-tuning the inclusion and skipping. The GRIN1 splicing represents a good example of an orchestrated regulation by different splicing factors, adjusting the activity of the NMDA receptor in a tissue-specific and developmentally-specific manner (Black and Grabowski, 2003).

1.5.3 Alternative splicing of GRIN1 upon neuronal excitation

The splicing of CI is inducible by potassium chloride (KCl) excitation, giving rise to exon skipping upon depolarization, and the splicing is reversed to the steady-state levels upon the removal of KCl (An and Grabowski, 2007), suggesting an adjustable splicing regulation that responds to the voltage changes across the membrane and neuronal impulses. CI encodes an intracellular region of the NMDA R1 receptor; therefore the NMDA receptor agonists and intracellular signal transduction inhibitors (such as the Protein Kinase A) cause a decrease in the responsiveness of the alternative splicing to neuronal excitation, and the hnRNP A1 is the splicing factor that acts to increase the

sensitivity to neuronal depolarization (An and Grabowski, 2007). Therefore, it can be suggested that the alternative splicing of the CI can be influenced by the signaling pathways, contributing to or countering the effect of membrane voltage changes.

The splicing of GRIN1 CI cassette exon is responsive to KCl excitation of the rat cortical neurons, and the binding of hnRNP A1 to UAGG motifs was demonstrated to increase upon neuronal depolarization (An and Grabowski, 2007). Since the hnRNP A1-binding UAGG was part of the UAGG-GGGG splicing code influencing the CI splicing, the hnRNP H family members that bind to the GGGG motif are also of special interest. The binding of hnRNP H to the GGGG motif of its own transcript, taken together with the regulation of GRIN1 splicing by hnRNP A1 and hnRNP H, suggests a regulation of HNRNPH1 and HNRNPH3 splicing by their own protein products and different types of cellular stresses, possibly interfering with the binding of hnRNP H to the GGGG motif and the binding of hnRNP A1 to the UAGG motif.

1.6 THESIS GOAL AND OBJECTIVES

By describing the current knowledge about the alternative splicing of several exons with respect to cellular stresses and diseases, I provided the framework of this study, which aims to address different gaps and questions that are interrelated with each other. The questions involve the modulating and fine-tuning role of hnRNP H on splicing upon transcriptional stress and pause-release, the role of hnRNP H possibly as an adaptation mechanism for the cell to handle the changing conditions and give a splicing response to the intracellular and extracellular challenges that affect mRNA transcription. The

results shed light on the role of *cis* regulatory RNA elements and *trans*-acting splicing factors involved in the splicing response to stress.

How does hnRNP H act in response to the genotoxic and transcriptional stress, and is this alternative splicing regulation reversible upon transcriptional pause-release, possibly in relation to the tissue-specific and developmental regulation? Is the splicing of hnRNP H-encoding transcript subject to autoregulation? Can GRIN1, the previously-characterized transcript with an inducible splicing pattern upon depolarization, act as a model for the alternative splicing regulation upon transcriptional stress? Is the splicing response to transcriptional stress specifically regulated by a predicted splicing code?

In light of these questions, the specific goals of this study are as follows:

1- Determining the effect of genotoxic and transcriptional stress on the splicing of the previously predicted hnRNP H-encoding HNRNPH1 and HNRNPH3, sharing a UAGG-GGGG predicted splicing code

2- Establishing an RNA map to understand the cis-regulatory elements and trans-acting splicing factors on the splicing of conserved HNRNPH1 exon 4 and HNRNPH3 exon 3

3- Investigating the mechanism of alternative splicing upon transcriptional pausing and pause-release in the cloned rat GRIN1 CI reporter minigene and a set of mutants

Our main question is to understand the role of hnRNP H in the splicing response to genotoxic and transcriptional stress. The previous work on the HNRNPH1 exon 4 and HNRNPH3 exon 3 involved their fast response to the KCl excitation in neurons (An and Grabowski, 2007). Cancer cells are subject to many different types of stress including genotoxic and transcriptional stress, and splicing is possibly one of the ways to adapt to the rapidly changing environment. Transcriptional pausing is observed in many genes of

embryonic stem cells, hinting at a temporal regulation of splicing in development. I hypothesize that the hnRNP H has a role in fine-tuning splicing upon genotoxic and transcriptional stress, through the previously predicted UAGG-GGGG splicing code.

The previously-characterized combinatorial roles of hnRNP A1 and hnRNP H on the splicing of GRIN1 and the prediction of HNRNPH1 and HNRNPH3 in a genomewide search for the UAGG-GGGG splicing code (Han *et al.*, 2005) constitutes a foundation for a possible role of these splicing factors on the alternative splicing of the hnRNP H-encoding transcripts, which share the same splicing code. The different exons such as Bcl-x that are responsive to different cellular stresses, and the role of hnRNP H in the regulation of the exons (such as Bcl-x) that carries poly(G) stretches, suggest a role for the hnRNP H in the splicing of its own transcript and in the response to cellular stresses as a fast and reversible way of modulating and fine-tuning the stress response.

In order to analyze the splicing of HNRNPH1 exon 4 and HNRNPH3 exon 3, the percentages of exon inclusion upon genotoxic and transcriptional stress-inducer small molecules were determined. Bioinformatics tools were utilized to establish an RNA map of cis-regulatory elements. The reporter-cloned rat GRIN1 minigene, that shares the common UAGG-GGGG splicing code, was used as a model to understand the motifs that regulate the effect of transcriptional stress on splicing. The set of mutants, in comparison with the wild type, were transfected to observe the same treatments and their outcome. HeLa cells were used as a model system due to very high transfection efficiency, easy propagation and application of stress inducers, and wide use.

Understanding the role of the hnRNP H in the fine-tuning of alternative splicing upon transcriptional stress is one of the main steps required to elucidate a bigger cellular stress-response pathway, involving crucial steps of alternative splicing.

2.0 MATERIALS AND METHODS

2.1 PLASMID CONSTRUCTS AND PCR PRIMER DESIGN

All of the GRIN1 C1 cassette exon 5' splice site mutants used in this study were the derivatives of the wild type Exon 21 splicing reporter E21wt (Zhang et al., 2002) in the pBPSV pA^+ vector (Nasim *et al.*, 1990), as previously described (Han *et al.*, 2005). Rat GRIN1 transcript-203 sequence was obtained from Ensembl Genome Browser. GRIN1 E18/E20 fragment was previously cloned by Dr. Kyoung Han into the pBPSV pA^+ with XbaI and EcoRI (Han *et al.*, 2005). The pA^+ C1 cassette splicing reporters; wild-type C1wt0, UAGG mutant E17, UAGG-GGGG mutant T8, and the GGGG mutants 5m2 and 5m4 were the same as described (Han *et al.*, 2005; An and Grabowski 2007). Human HNRNPH1 transcript-001, human HNRNPH3 transcript-012, mouse Hnrnph3 transcript-001 sequences were obtained from Ensembl Genome Browser. HnRNP A1 expression plasmid was the derivative of the pcDNA4/HisMax, as described (Han *et al.*, 2005). The PKA expression plasmid was a gift from Dr. Jiuyong Xie (University of Manitoba).

The specific forward and reverse PCR primers that amplify the HNRNPH1 E3/E5, human HNRNPH3 E2/E4, mouse Hnrnph3 E2/E4 (Table 1) were designed within exons, BLAST-checked and optimized by using IDT Oligo Analyzer. H3E3 Forward primer was also used in the TAMRA-fluorolabeled form, and the PCR products were visualized with

the ImageGauge using the 540/575 nm (excitation/emission) wavelengths to confirm the splicing changes without gel background or other impurities interfering with the image.

The sequences of HNRNPH1 and HNRNPH3 primers are:

H1E4 Forward: 5’ TTGGGTGTTGAAGCATACTGG 3’

H1E4 Reverse: 5’ CATAAGCTTTCGTGGTGGATCA 3’

H3E3 Forward: 5’ AATGGTCCAAATGACGCTAG 3’

H3E3 Reverse: 5’ CCAGCAATCTTCTTGGTGG 3’

The specific forward and reverse PCR primers to amplify the GRIN1 C1 cassette exon were used, as described (Han *et al.*, 2005). Their sequences are:

NMDA3021: 5’ ATGCCCGTAGGAAGCAGATGC 3’

NMDA3255A: 5’ CGTCGCGGCAGCACTGTGTC 3’

The specific forward and reverse primers for the control PCR that amplify the beta actin (ACTB) E1/E3 and were designed within the exonic regions, BLAST-checked and optimized by using IDT Oligo Analyzer. Their sequences are:

ACTB Forward: 5’ AGCTTCTTTGCAGCTCCTTC 3’

ACTB Reverse: 5’ TCTTCTCCATGTCGTCCCAG 3’

2.2 SPLICING FACTOR PREDICTIONS AND THE RNA MAPS

Based on the sequences obtained from Ensembl genome browser, the Human Splicing Finder (HSF) online tool (http://www.umd.be/HSF3/) and the Exonic Splicing Enhancer Finder (ESEfinder) online tool (http://rulai.cshl.edu/tools/ESE/) were used to construct the RNA maps, including the scores for each predicted RNA motif represented by a

probability score to bind a particular splicing factor (Tables 2-8). Scores higher than 80 percent were highlighted as bold characters.

Rat GRIN1 CI cassette exon RNA maps, previously described (Han *et al.*, 2005), were updated and adapted according to the predictions from HSF and ESEfinder online tools. HNRNPH1 E4 and HNRNPH3 E3 RNA maps were prepared, according to the CLUSTALW multiple sequence alignment, provided online on EMBL-EBI CLUSTALW2 page: (http://www.ebi.ac.uk/Tools/msa/clustalw2/).

2.3 RNA PREPARATION FROM TISSUES AND CULTURED CELLS

Cryofrozen mouse brain and liver samples were gifts from Dr. James Pipas (University of Pittsburgh). The samples were thawed on ice, in order to minimize RNA degradation. Each tissue sample was weighed, using clean 15 mL tubes. 3 mL Trizol (Ambion Life Technologies) was added on the tissues, and the tissues were homogenized by dounce homogenizer. Total RNA was extracted from the homogenates via phenol-chloroform RNA extraction, according to the manufacturer's protocol.

After the transfection or treatment, 500 µl Trizol (Ambion Life Technologies) per well were used for extracting RNA from cultured HeLa CCL-2 or U2OS cells via phenol-chloroform total RNA extraction, according to the manufacturer's protocol.

2.4 CULTURING AND TRANSFECTION OF THE HELA CCL-2 CELLS

HeLa CCL-2 cells were grown in Dulbecco's Modified Eagle Medium DMEM (Cellgro) supplemented with heat-inactivated 10% (v/v) fetal bovine serum (FBS), inside a sterile 37°C incubator with 5% CO_2. 24 hours before the transfection or treatment, the cells were splitted to six-well plates to achieve 70% confluency.

For transient transfection, 500 ng pcDNA4-hnRNP A1, pcDNA4-PKA or pcDNA4 were used, along with 125 ng pA^+-C1wt0, pA^+-5m2 or pA^+-5m4, per well. The plasmids were mixed with 100 µl OptiMEM (Gibco) per well, and mixed to equal volume of OptiMEM that has 2 µl lipofectamine 2000 (Invitrogen) per well. The transfection mixture was incubated at room temperature for 20 minutes. The DMEM + 10% FBS on the cells was removed, and 1 mL OptiMEM was added before transfection.

For checking the transfection efficiency, the cells were seeded on six-well plates that have cover slips. The following day, the EYFP-A1 plasmid was also transfected, and the $EYFP^+$ cells on the cover slips were mounted on slides and they were observed under the microscope to estimate the percentage of transfection, which was determined as 70 percent, by checking the ratio of $EYFP^+$ nuclei to DAPI-stained cell population.

2.5 HNRNP A1 OVEREXPRESSION AND HNRNP A1/H KNOCKDOWN

In order to knock down the hnRNP A1 and hnRNP H, HeLa CCL-2 cells were seeded on a six-well plate 24 hours prior to the transfection. The confirmed siRNAs targeting

the hnRNP A1/A2 and hnRNP H, namely the siA1/A2 and siH, were transfected into the cells. HeLa CCL-2 samples overexpressing hnRNP A1 were prepared by Dr. Ping An.

2.6 GAMMA IR AND CISPLATIN-INDUCED DNA DAMAGE

The U2OS human osteosarcoma cells were grown in triplicates on the 60 mm plates overnight, and they were treated with 2.8-3 Gy of γ-radiation by exposing them to ^{137}Cs for 15 minutes. Following the γ-IR treatment, the cells were either harvested right after the irradiation (0h post-IR), or allowed to recover until the 8h, 18h and 24h time points. Double-stranded DNA damage was confirmed by Western Blot against phosphohistone Ser139 γ-H2AX (Figure 5).

The HeLa CCL-2 (ATCC) cells were splitted to six-well plates the day before the treatment, and grown overnight. They were treated with the 0-20-40 μM cisplatin, by preparing the mixture of 0.9% autoclave-sterilized NaCl with a main stock of 50 mM cisplatin (Sigma) and diluting the main stock in fresh DMEM + 10% FBS. The control cells were treated with the medium including equal volume of 0.9% NaCl. The treated and control cells were harvested after the 14-hour cisplatin treatment.

2.7 DRB TREATMENT AND TRANSCRIPTIONAL PAUSE-RELEASE

The HeLa CCL-2 cells, seeded on six-well plates 24 hours prior to treatment with or without lipofectamine transfection, were grown inside a sterile 37°C incubator in the presence of 5% CO_2, in DMEM+10% FBS. If the cells were not transfected, they were treated with 0-30-40 μM DRB. The DRB stock solution was 50 mM, which was prepared in sterile 95% ethanol. The dilutions were prepared from the 50 mM stock in the sterile DMEM+10% FBS, and this mixture was used to replace the growth medium of the cells.

If the cells were transfected and then treated with DRB, they were grown in six-well plates overnight, as previously explained in 2.4. The cells were transfected with the lipofectamine mixture, with the presence of the OptiMEM instead of DMEM+10% FBS. 4 hours after the transfection, the OptiMEM was removed and replaced by the DRB mix that was diluted in fresh sterile DMEM+10% FBS. The cells were incubated for 14 hours inside a sterile 37°C incubator in the presence of 5% CO_2.

For the transcriptional pause-release, the cells were washed once with the fresh DMEM+10% FBS following the 14-hour DRB treatment, and were incubated in the fresh DMEM+10% FBS for 30 minutes, 1 hour, and 2 hours. If the cells were harvested at the 0-hour time point, they were immediately harvested following the DMEM+10% FBS wash, without further incubating inside a sterile 37°C incubator.

2.8 LYSATE PREPARATION, SDS-PAGE AND WESTERN BLOTTING

Total cell lysates of HeLa CCL-2 cells were prepared from each well of the six-well plate and two wells with the same treatment were combined to obtain one 1.5 mL tube of total cell lysate. The cells were harvested in the +4°C cold room and on ice to minimize the protein degradation due to temperature increase. The cells were harvested by adding 75 µl ice-cold RIPA buffer (10 mM Tris-HCl pH 7.5, 100 mM NaCl, 1 mM EDTA, 0.5% sodium deoxycholate, 0.1% SDS, 1% Triton X-100) per well. The cells were scraped off the plate by the use of a rubber scraper and transferred into a 1.5 mL Eppendorf tube prechilled on ice. The duplicate wells having the same content were combined to obtain a final volume of 150 µl. The lysate was sonicated for 10 seconds on setting 3, and aliquotted into single-use 20 µl aliquots for flash-freeze in dry ice, and stored at -80°C.

The whole cell lysates were used for SDS-PAGE and Western blot analysis. The whole cell lysates, prepared as mentioned above, consisted of the cells scraped in RIPA buffer, sonicated and aliquoted. The concentrations of the whole cell lysates were not measured due to the presence of SDS. The SDS-PAGE was performed with the BioRad 12% precast acrylamide gels. The SDS-PAGE protein samples were prepared by cold centrifugation of the cell debris at +4°C at the 10,000 rpm speed for 5 minutes. After the spin, the supernatant was transferred into a prechilled ice-cold clean 1.5 mL eppendorf tube. 6X SDS loading buffer (1 M Tris-HCl pH 6.8, 30% glycerol, 10% SDS, 600 mM DTT, 0.012% bromophenol blue) was added to obtain a final concentration of 1X.

The 1X samples that include the SDS loading dye were boiled on the 100°C heat block for 10 minutes. The acrylamide gel was placed in BioRad MiniProtean acrylamide electrophoresis reservoir, and 15 µl of each 1X sample was loaded on the gel. The gel

was run in 700 mL of 1X SDS-PAGE Running Buffer (50 mM Tris-HCl, 400 mM glycine, 0.1% SDS) under 90V for 80 minutes, until the bromophenol blue dye band reached the bottom of the reservoir.

The gels were equilibrated in the 1X Transfer Buffer (25 mM Tris-HCl, 192 mM glycine, 20% methanol) for 15 minutes, and then sandwiched with the Immobilon-PSQ hydrophobic 0.2 µm PVDF membrane by wetting the membrane with absolute methanol before use. The sandwiched pad-gel-membrane-pad was completely soaked in the 1x Transfer Buffer, and the bubbles were removed with the help of a glass Pasteur pipette by rolling several times. The sandwiched pad-gel-membrane-pad was located into the color-coded plastic BioRad MiniProtean Tetra Cell cassette, inserted into the BioRad MiniProtean Tetra Cell transfer tank, and transferred under 120V for 75 minutes in the +4°C cold room in order to minimize overheating and potential hazards, according to the manufacturer's protocol.

For the Western blot analysis, the PVDF membrane with the transferred proteins was blocked in 5% non-fat milk powder dissolved in PBST (1X phosphate buffer saline [PBS], 0.1% Tween-20) for 1 hour at room temperature on the shaker while shaking at 60 rpm, followed by a 5-minute wash in 1X PBST. The PVDF membranes were marked with a small letter to label the type of antibody used and to recognize the surface that has the transferred proteins.

The membranes were incubated with the primary antibody for at least 1 hour (up to overnight, depending on the antibody) in the +4°C cold room, while rotating at 60 rpm. The primary antibodies were diluted in pre-chilled ice-cold 1X PBST + 1% BSA as follows: α-Beta Tubulin (1:200), α-hnRNP H (1:250), α-Phospho-Histone H2AX [Ser139]

(1:500), α-HA (1:250), α-ASF (1:500), α-SC35 (1:1,000). The membranes were washed four times, 10 minutes each, with 1X PBST at room temperature, to wash the unbound antibodies, following the primary antibody incubations. After the four washes, the PVDF membranes were incubated with the mouse, rabbit, or goat HRP-conjugated secondary antibodies, which were diluted 1:20,000 in the pre-chilled ice-cold 1X PBST + 1% BSA, while rotating in the +4°C cold room at 60 rpm for 1 hour. After the secondary antibody incubations, the membranes were washed four times, 10 minutes each, with 1X PBST at room temperature, to wash the unbound antibodies.

After the end of the washes, the membranes were treated with equal volumes of Perkin Elmer Western Lighting Chemiluminescence Plus Enhanced Luminol Reagent and Oxidizing Reagent for 3 minutes before developing the chemiluminescence signal images with the LAS-3000 FujiFilm Intelligent Dark Box by using 15-second increments of exposures for optimization, and the chemiluminescence images were then analyzed with the ImageReader, ImageGauge, and ImageJ softwares.

2.9 SPLICING ANALYSES: RT-PCR AND THE ELECTROPHORETIC BAND QUANTIFICATIONS

Total RNA was extracted from the HeLa CCL-2 cells, as described in 2.3, according to the manufacturer's protocol. After the phenol-chloroform RNA extraction, the total RNA concentration was measured by NanoDrop 2000, and diluted to 1 μg/μl in RNase-free sterile dH_2O. The total RNA was reverse transcribed with the Invitrogen SuperScript III Reverse Transcriptase enzyme in a reaction containing 2 μg of total RNA, 1 μl random

hexamer primers (Promega), 0.5 mM dNTPs, 10 mM DTT, 1X First Strand (FS) Buffer (50 mM Tris-HCl pH 8.3, 75 mM KCl, 3 mM $MgCl_2$) and 200 U Invitrogen SuperScript III Reverse Transcriptase in a final volume of 20 µl.

The primer-template mixture was denatured at 70°C for 10 minutes, annealed at room temperature for 10 minutes, and the reverse transcription reaction was carried out in the thermocycler at 50°C for 1 hour. Finally, the reverse transcription reaction was stopped by heat inactivation at 65°C for 15 minutes, according to the manufacturer's protocol.

Polymerase chain reaction (PCR) was performed at a volume of 10 µl each, that has 1 µl of the heat-inactivated reverse transcription reaction, 2 µl of 5X GoTaq Green Buffer (Mg^{2+} already optimized) to make a final concentration of 1X, 2.5 U of GoTaq Taq DNA polymerase (Promega), 0.2mM dNTPs, and 0.1 µM of each of the exon-specific forward and reverse primers that were diluted in nuclease-free dH_2O to a stock solution of 20 mM. The PCR steps and conditions were optimized as follows:

Initial denaturation at 94°C for 2 minutes; followed by 26 cycles of denaturation at 94°C for 30 seconds, annealing at 60°C for 30 seconds, and elongation at 72°C for 45 seconds (1 kb/min). Final elongation was performed at 72°C for 10 minutes, and the products were cooled to +4°C, flash frozen with dry ice and stored at -80°C.

PCR products were resolved on 1.8% agarose gel, prepared with 1.8 grams of ultrapure agarose, 100 ml 1X TBE (89 mM Tris Borate pH 8.3 + 2 mM Na_2EDTA) diluted from 10X stock with Nano-pure dH_2O, and 1 µl EtBr from a stock solution of 10 mg/ml. The agarose gels were run under 110V for 80 minutes, and the DNA ladder used in first

wells was 1 µl 1-kb Plus DNA ladder (Invitrogen). The gels were autoexposed with the FujiFilm LAS-3000 Intelligent Dark Box, analyzed with ImageReader and ImageGauge.

Electrophoretic band intensities were measured by ImageGauge, by selecting a rectangular region encompassing each well of the gel, subtracting the background from the peaks in the analysis curves of the software, in the Profile/MW setting. The graphed values show percent inclusion and skipping, and error bars show standard deviation.

The changes in splicing were calculated as a difference of the percentages of exon inclusion, which was denoted as the Ψ (PSI = Percent Spliced In) value, and it was calculated by subtracting the exon inclusion of the control group from the exon inclusion of the experimental group to obtain the ΔΨ (Wang and Burge, 2008). For the statistical analysis, the student's *t* test was used, with *p*-value < 0.05 at a confidence interval of 95%. Analyses were performed with Microsoft Office 2010 Excel in triplicates.

3.0 RESULTS

3.1 HNRNPH1 EXON 4 AND HNRNPH3 EXON 3 RESPOND SIMILARLY TO SINGLE-STRAND GENOTOXIC AND TRANSCRIPTIONAL STRESS

3.1.1 Introduction

HnRNP H family of splicing factors are encoded by the human HNRNPH1, HNRNPH2, and HNRNPH3 transcripts; two of which are alternatively spliced: HNRNPH1 and HNRNPH3. These transcripts were previously found in a genomewide search to have a common UAGG-GGGG splicing code with the brain region-specific GRIN1 CI cassette exon (Han *et al.*, 2005). Different proliferation-associated exons, such as the Fibronectin EDI exon, were previously studied for the effect of DRB-induced transcription elongation inhibition (Nogués *et al.*, 2002), UV- and cisplatin-induced genotoxic stress (Muñoz *et al.*, 2009) on alternative splicing. As a result of a previous study from our lab, Hnrnph3 Exon 3 splicing was observed to be inducible upon neuronal depolarization by the KCl excitation of rat primary cortical neurons (An and Grabowski, 2007), and we asked if the alternative splicing of this inducible exon is also subject to regulation upon genotoxic and transcriptional stress. I hypothesized that the previously-predicted exons sharing the common UAGG-GGGG splicing code would be regulated similarly upon genotoxic

and transcriptional stress. I designed PCR primers for assaying the splicing of several exons throughout the study (Figure 1) (Table 1), and analyzed the splicing upon specific type of stress or treatment of question.

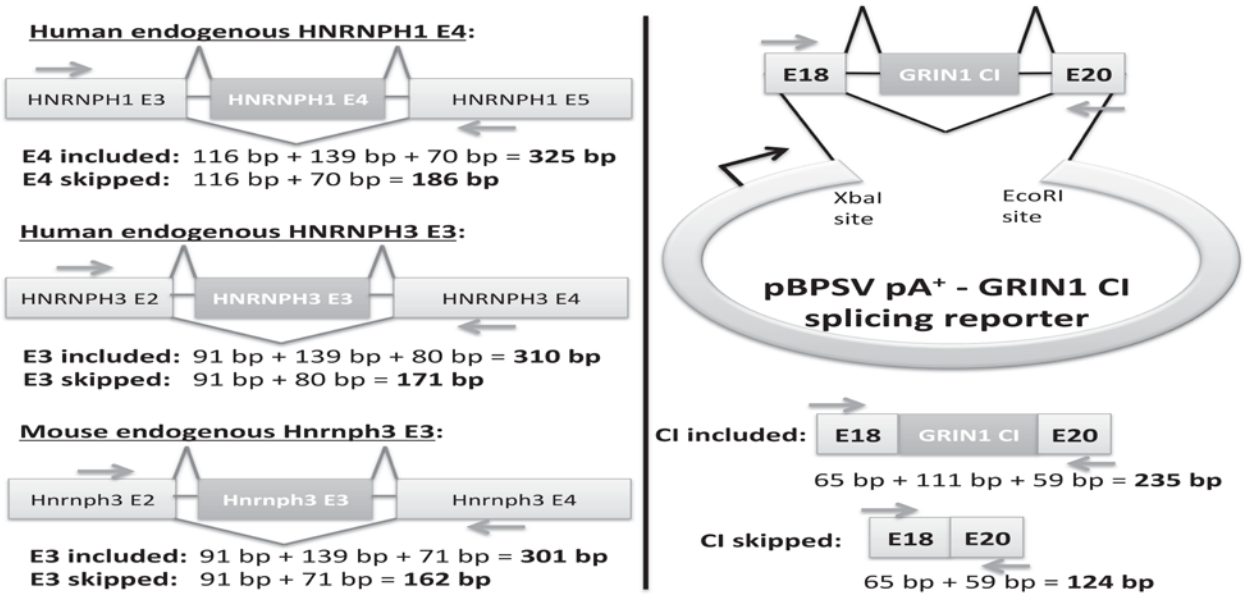


Figure 1: Schematic representation of the PCR design for splicing analyses. PCR primers complementary to the flanking exons of HNRNPH1 E4, HNRNPH3 E3, and GRIN1 CI were designed, based on the genomic DNA sequence data obtained from the Ensembl Genome Browser. Cassette exons are either included or skipped. The primer locations are scaled, compared to full-length exon size. Human HNRNPH1 E3+E4+E5, having the size of 144+139+179 bp, is partially amplified by exonic primers, resulting in a 325-bp (E4 included) or 186-bp (E4 skipped) product. Human HNRNPH3 or mouse Hnrnph3 E2+E3+E4, having the size of 135+139+185 bp, is partially amplified by the same primers that amplify both transcripts to produce the products with indicated sizes. GRIN1 E18/E20 fragment, cloned by Dr. Kyoung Han into the pBPSV pA$^+$ with XbaI and EcoRI, gives the indicated PCR products: 235-bp (CI included) or 124-bp (CI skipped).

Table 1: Primer sequences for the amplification of endogenous and reporter exons

Primer Name	Sequence
H1E4 Forward	5’ TTGGGTGTTGAAGCATACTGG 3’
H1E4 Reverse	5’ CATAAGCTTTCGTGGTGGATCA 3’
H3E3 Forward	5’ AATGGTCCAAATGACGCTAG 3’
H3E3 Reverse	5’ CCAGCAATCTTCTTGGTGG 3’
NMDA3021 (Fwd)	5’ ATGCCCGTAGGAAGCAGATGC 3’
NMDA3255A (Rev)	5’ CGTCGCGGCAGCACTGTGTC 3’
ACTB Forward	5’ AGCTTCTTTGCAGCTCCTTC 3’
ACTB Reverse	5’ TCTTCTCCATGTCGTCCCAG 3’

3.1.2 The alternative splicing of Hnrnph3 Exon 3 is subject to tissue-specific and developmental regulation in mouse

Tissues with different developmental needs or proliferative phenotypes have different splicing profiles, especially with relation to the proliferation-associated transcripts that contribute by giving rise to protein products that improve the adaptability of the cell to various types of stress. Since the hnRNP H splicing factor was shown to be differentially expressed in several types of cancer, and the transcript encoding hnRNP H3 was found to be inducible, I asked whether it has a proliferative phenotype-specific splicing in vivo.

Embryonic brain samples and adult cerebral cortex samples from various mice had higher inclusion of HNRNPH3 Exon 3, in comparison with the liver samples of adult mice, demonstrating the tissue specificity of the splicing of HNRNPH3 with respect to different proliferation rates and different developmental time points (Figure 2). The brain samples, belonging to a tissue with a much lower proliferation rate than the liver, has a distinct splicing of HNRNPH3, showing the inducibility at a tissue- and developmentally-specific level. After this result, I decided to check for the predictions for the regulatory RNA motifs in or flanking the HNRNPH3 Exon 3, to understand its splicing regulation.

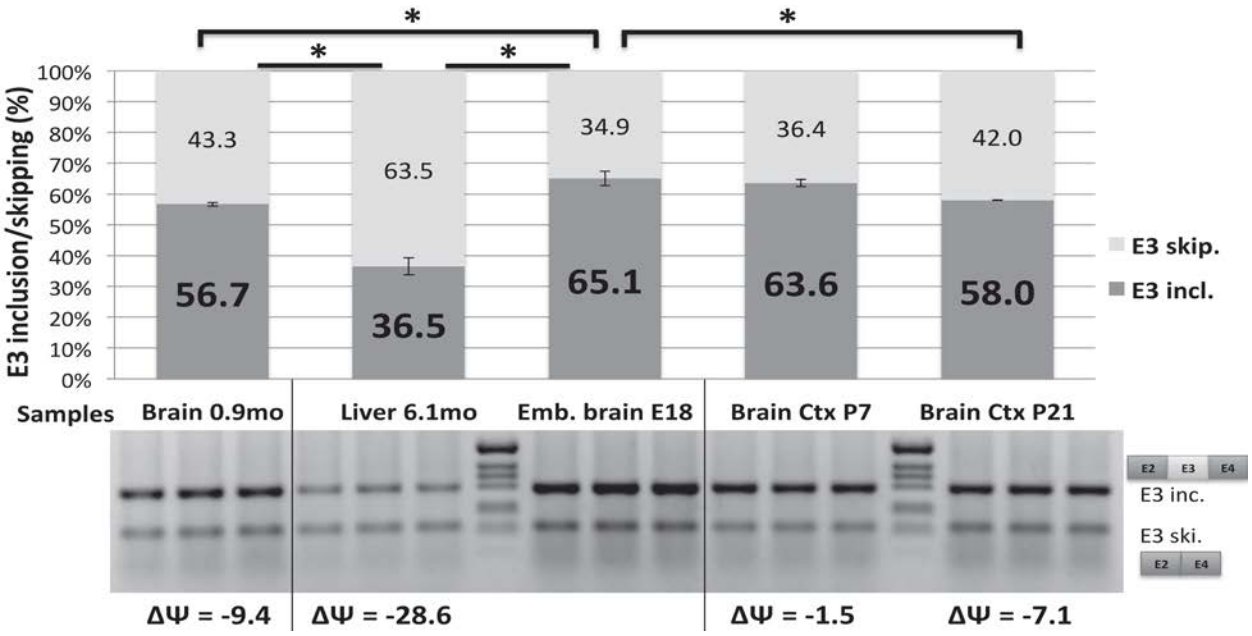


Figure 2: Embryonic and mature mouse brain samples have high inclusion of Hnrnph3 exon 3, in contrast to lower inclusion in liver. Mouse samples (n = 3) were used for total tissue RNA extraction and reverse transcription, from the indicated ages of embryonic and adult mouse brain and liver tissue samples. PCR was performed with the endogenous HNRNPH3 primers for 26 cycles. Electrophoretic band intensities were measured by ImageGauge software, selecting a rectangular region encompassing each well of the gel, subtracting the background from the peaks in the analysis curves of the software, in Profile/MW setting. The graphed values above show percent inclusion and skipping, and the error bars

show standard deviation. ΔΨ (difference in PSI = Percent Spliced In) values show the difference in percent inclusion, compared to brain sample at embryonic day 18. Unpaired, two-tailed student's *t* test was performed, and the changes in HNRNPH3 E3 inclusion were statistically significant, shown with asterisk. ($p < 0.05$)

Figure 3: CLUSTALW alignment of human HNRNPH1 Exon 4 and HNRNPH3 Exon 3 with their flanking introns. Highlighted section shows the 139-bp highly-conserved exon of the HNRNPH1 and HNRNPH3. At the left panel; H1 stands for the HNRNPH1 transcript, H3 stands for the HNRNPH3 transcript, H1E4 stands for HNRNPH1 Exon 4, and H3E3 stands for HNRNPH3 Exon 3. Asterisk (*) under the residues shows identical residues. 126 out of 139 bases are identical between HNRNPH1 Exon 4 and HNRNPH3 Exon 3, with 90.6% conservation. Approximately 50 bases upstream of both exons are highly conserved, with a disrupted stretch of 38 bases in the upstream intron. There is also conservation in the downstream, about a disrupted stretch of 30 bases in the downstream intron. The TAGG and GGGG motifs, which are part of the previously-predicted splicing code (Han *et al.*, 2005), are underlined.

3.1.3 HnRNP H is an auto-regulatory splicing factor, silencing the exon of its own transcript in HeLa cells

The HNRNPH3 E3 and its human paralog HNRNPH1 E4 share hnRNP A1- and hnRNP H-binding RNA motifs (Figure 3), triggering the question whether hnRNP H can regulate the splicing of its own transcript as a feedback loop mechanism. I used the HeLa CCL-2 samples that had stable hnRNP A1 knockdown, hnRNP H knockdown and hnRNP A1 overexpression to see their effects on the splicing of HNRNPH3 E3. I observed that the hnRNP H knockdown caused a strong increase in the inclusion of HNRNPH3 E3, while the hnRNP A1 overexpression had an anti-silencing effect (Figure 5). The hnRNP A1 knockdown did not have a statistically significant effect, due to the presence of hnRNP H. Therefore, hnRNP H acts as a strong silencer of HNRNPH3 Exon 3, whereas hnRNP A1 acts as an anti-silencer, on the splicing of the HNRNPH3 Exon 3 in HeLa cells.

3.1.4 HNRNPH1 E4 and HNRNPH3 E3 skipping increases upon cisplatin treatment, but not upon gamma-induced double strand breaks

The splicing of proliferation-associated transcripts is affected by the treatment of cancer cells with several cancer drugs, raising the question whether the splicing of our inducible HNRNPH3 Exon 3 responds to genotoxic stress. The human paralog of the HNRNPH3 Exon 3, the HNRNPH1 Exon 4, has very high sequence conservation, and I expected them to respond similarly to cisplatin-induced single-strand genotoxic stress, due to the presence of the putative hnRNP H- and hnRNP A1- binding motifs (Figure 4).

Table 2: Human HNRNPH1 Exon 4 and HNRNPH3 Exon 3 ESEfinder motif predictions

Seq: H1E4

GGTTGGAAATCGTGCCAAATGGGATAACATTGCCGGTGGACTTCCAGGGGAGGAGTACGG
GGGAGGCCTTCGTGCAGTTTGCTTCACAGGAAATAGCTGAAAAGGCTCTAAAGAAACACA
AGGAAAGAATAGGGCACAG

SRSF1		SRSF1 (IgM-BRCA1)		SRSF2		SRSF5		SRSF6	
threshold: 1.956		threshold: 1.867		threshold: 2.383		threshold: 2.67		threshold: 2.676	
Position*/Site/Score		Position*/Site/Score		Position*/Site/Score		Position*/Site/Score		Position*/Site/Score	
34 (-106) CGGTGGA	2.8407	34 (-106) CGGTGGA	3.4494	31 (-109) TGCCGGTG	2.589	42 (-98) TTCCAGG	3.6168	56 (-84) TACGGG	2.95
45 (-95) CAGGGGA	2.8302	45 (-95) CAGGGGA	3.092	39 (-101) GACTTCCA	4.0064	84 (-56) TCACAGG	6.3244	80 (-60) TGCTTC	3.8319
48 (-92) GGGAGGA	3.7142	48 (-92) GGGAGGA	2.2117	65 (-75) GGCCTTCG	3.319	98 (-42) TGAAAAG	3.366		
58 (-82) CGGGGGA	2.3829	58 (-82) CGGGGGA	3.3902	104 (-36) GGCTCTAA	3.0188	107 (-33) TCTAAAG	3.1403		
85 (-55) CACAGGA	5.985	85 (-55) CACAGGA	5.193	132 (-8) GGGCACAG	2.6486	116 (-24) ACACAAG	4.707		
117 (-23) CACAAGG	3.3386	117 (-23) CACAAGG	3.0512						

Seq: H3E3

GGTTGGAAATCGTGCCAAATGGGATAACATTGACGATGGACTACCAGGGGAGAAGCACAG
GGGAGGCCTTCGTGCAGTTTGCTTCAAAGGAGATAGCAGAAAATGCTCTGGGGAAACACA
AGGAAAGAATAGGGCACAG

SRSF1		SRSF1 (IgM-BRCA1)		SRSF2		SRSF5		SRSF6	
threshold: 1.956		threshold: 1.867		threshold: 2.383		threshold: 2.67		threshold: 2.676	
Position*/Site/Score		Position*/Site/Score		Position*/Site/Score		Position*/Site/Score		Position*/Site/Score	
30 (-110) TTGACGA	2.2699	34 (-106) CGATGGA	2.0656	39 (-101) GACTACCA	4.2932	41 (-99) CTACCAG	3.3509	80 (-60) TGCTTC	3.8319
45 (-95) CAGGGGA	2.8302	45 (-95) CAGGGGA	3.092	65 (-75) GGCCTTCG	3.319	55 (-85) GCACAGG	3.8201		
56 (-84) CACAGGG	5.2582	56 (-84) CACAGGG	4.7576	132 (-8) GGGCACAG	2.6486	84 (-56) TCAAAGG	4.7554		
58 (-82) CAGGGGA	2.8302	58 (-82) CAGGGGA	3.092			116 (-24) ACACAAG	4.707		
85 (-55) CAAAGGA	3.675	85 (-55) CAAAGGA	3.428						
117 (-23) CACAAGG	3.3386	106 (-34) CTCTGGG	2.6658						
		108 (-32) CTGGGGA	2.6607						
		117 (-23) CACAAGG	3.0512						

Table 3: Human HNRNPH1 exon 4 Human Splicing Finder enhancer motif predictions

Seq.	Linked SR Prot.	Enh. Motif	Value (0-100)	Seq.	Linked SR Prot.	Enh. Motif	Value (0-100)
43	SRp55	TGTATA	79.88	**548**	**SRp40**	**TCACAGG**	**100**
66	SC35	AGTTCCAA	75.29	**549**	**SF2/ASF (IgM-BRCA1)**	**CACAGGA**	**96.15**
78	**SRp40**	**TTAAAGC**	**83.65**	**549**	**SF2/ASF**	**CACAGGA**	**96.45**
137	**SRp40**	**TTAATGC**	**82.16**	**562**	**SRp40**	**TGAAAAG**	**82.28**
141	SRp55	TGCTTA	78.98	568	SC35	GGCTCTAA	78.98
144	**SRp40**	**TTAAAAC**	**80.24**	**571**	**SRp40**	**TCTAAAG**	**80.9**
157	SC35	GAATTGTA	77.14	**580**	**SRp40**	**ACACAAG**	**90.3**
161	SRp55	TGTATA	79.88	581	SF2/ASF (IgM-BRCA1)	CACAAGG	79.77
169	**SRp40**	**TCTCAAC**	**85.21**	**581**	**SF2/ASF**	**CACAAGG**	**81.07**
183	**SRp55**	**TGCATA**	**89.8**	596	SC35	GGGCACAG	76.64
216	**SRp40**	**TTAATGC**	**82.16**	**598**	**SRp40**	**GCACAGG**	**85.03**
220	**SRp55**	**TGCATC**	**92.17**	**599**	**SF2/ASF (IgM-BRCA1)**	**CACAGGT**	**94.38**

221	SC35	GCATCCTG	79.9	**599**	**SF2/ASF**	**CACAGGT**	**94.41**
230	SC35	GTTCAGCA	77.32	**627**	**SRp55**	**TATGTC**	**82.25**
250	SRp55	TATATA	75.97	639	SRp40	CTTATGG	78.8
258	**SC35**	**GTACCATG**	**81.99**	654	SRp55	TAAATC	74.56
294	**SRp40**	**ATAATGG**	**80.96**	658	SRp55	TCCATA	75.08
304	SRp55	TTCATA	75.08	**667**	**SRp40**	**TCTCTGC**	**87.13**
319	**SRp55**	**TGCATA**	**89.8**	671	SRp55	TGCTTA	78.98
325	SRp40	TTTAAAG	79.04	**674**	**SRp40**	**TTAAAAG**	**85.33**
341	SRp40	TGAAAGC	80.6	692	SC35	GTTTTGTA	78.06
345	SRp55	AGCATA	77.77	702	SF2/ASF (IgM-BRCA1)	CCTAGGT	73.23
353	SRp40	ATTCACC	78.62	**726**	**SC35**	**GATTTCCA**	**81.5**
358	SRp40	CCTAAAG	78.74	732	SF2/ASF (IgM-BRCA1)	CAAACTT	70.54
371	**SRp40**	**ATTCTGG**	**84.07**	734	SC35	AACTTGTG	75.35
382	SRp40	TTTAAAG	79.04	**738**	**SRp55**	**TGTGTC**	**86.15**
383	**SRp40**	**TTAAAGC**	**83.65**	**742**	**SRp40**	**TCAGTCC**	**81.26**
389	SRp40	CTATAAG	78.32	**745**	**SC35**	**GTCCCACG**	**84.27**
402	SC35	CATTTCTG	75.72	**746**	**SRp40**	**TCCCACG**	**83.77**
404	**SRp40**	**TTTCTGG**	**90.36**	**747**	**SF2/ASF (IgM-BRCA1)**	**CCCACGT**	**90.77**
410	**SC35**	**GGCTTGTG**	**88.32**	**747**	**SF2/ASF**	**CCCACGT**	**87.94**
430	SF2/ASF (IgM-BRCA1)	CCCCCCT	74.92	**753**	**SRp40**	**TTACACG**	**96.29**
431	SF2/ASF (IgM-BRCA1)	CCCCCTA	77.46	756	SF2/ASF (IgM-BRCA1)	CACGCAA	70.85
435	**SRp40**	**CTACTGG**	**94.49**	760	SF2/ASF (IgM-BRCA1)	CAAACTA	72.31
438	SF2/ASF (IgM-BRCA1)	CTGGGGT	74.92	760	SF2/ASF	CAAACTA	74.2
457	SC35	GGTTGAAG	75.54	774	SC35	GGTTTGAA	75.78
495	SC35	TGCCGGTG	76.34	**785**	**SC35**	**TGTCCCTA**	**81.13**
498	**SF2/ASF (IgM-BRCA1)**	**CGGTGGA**	**82.69**	786	SC35	GTCCCTAG	79.16
498	SF2/ASF	CGGTGGA	78.1	**787**	**SRp40**	**TCCCTAG**	**80.72**
503	**SC35**	**GACTTCCA**	**85**	801	SC35	GTCTGCTT	75.23
506	**SRp40**	**TTCCAGG**	**83.77**	**819**	**SRp55**	**TGAGTC**	**82.37**
509	**SF2/ASF (IgM-BRCA1)**	**CAGGGGA**	**80**	830	SC35	AACCCCAA	78
509	SF2/ASF	CAGGGGA	78.1	**838**	**SRp40**	**TCAATAG**	**85.69**
512	SF2/ASF (IgM-BRCA1)	GGGAGGA	73.15	851	SF2/ASF	GAGACTA	76.99
512	**SF2/ASF**	**GGGAGGA**	**83.17**	856	SRp55	TATGGC	77.06
520	SRp55	TACGGG	75.65	**871**	**SRp40**	**TTAATGC**	**82.16**
522	**SF2/ASF (IgM-BRCA1)**	**CGGGGGA**	**82.23**	884	SC35	GGCTTTAG	77.07
522	SF2/ASF	CGGGGGA	75.48	898	SRp40	TAACACC	79.52
529	**SC35**	**GGCCTTCG**	**80.82**	903	SF2/ASF (IgM-BRCA1)	CCCACCT	78.08
544	**SRp55**	**TGCTTC**	**81.35**	**927**	**SRp40**	**CCATAAG**	**80.18**

Table 4: Human HNRNPH1 exon 4 Human Splicing Finder silencer motif predictions

Seq.	Motif	Sil. motif	Value 0-100	Seq.	Motif	Sil. motif	Value 0-100
34	Motif 2 - [T/G]G[T/A]GGGG	TGTTGTAA	64.46	518	Motif 2 - [T/G]G[T/A]GGGG	AGTACGGG	61.28
61	Motif 1 - CTAGAGGT	CTTGAAGT	67.5	519	Motif 2 - [T/G]G[T/A]GGGG	GTACGGGG	61.98
66	Motif 3 - TCTCCCAA	AGTTCCAA	65.08	520	Motif 2 - [T/G]G[T/A]GGGG	TACGGGGG	70.71
67	Motif 3 - TCTCCCAA	GTTCCAAT	61.23	521	Motif 2 - [T/G]G[T/A]GGGG	ACGGGGGA	63.88
97	Motif 1 - CTAGAGGT	AATGAGGT	76.44	522	Motif 2 - [T/G]G[T/A]GGGG	CGGGGGAG	66.91
99	Motif 2 - [T/G]G[T/A]GGGG	TGAGGTAA	69.37	523	Motif 2 - [T/G]G[T/A]GGGG	GGGGGAG G	67.64
104	Motif 1 - CTAGAGGT	TAAAAGGT	72.32	524	Motif 1 - CTAGAGGT	GGGGAGGC	65.52
167	Motif 3 - TCTCCCAA	ATTCTCAA	66.85	524	Motif 2 - [T/G]G[T/A]GGGG	GGGGAGGC	67.64
186	Motif 1 - CTAGAGGT	ATAGATAT	63.52	526	Motif 2 - [T/G]G[T/A]GGGG	GGAGGCCT	67.1
206	Motif 2 - [T/G]G[T/A]GGGG	TGTTGAAT	64.46	546	Motif 3 - TCTCCCAA	CTTCACAG	65.99
257	Motif 3 - TCTCCCAA	TGTACCAT	66.85	548	Motif 1 - CTAGAGGT	TCACAGGA	61.63
261	Motif 1 - CTAGAGGT	CCATGGGT	64.15	555	Motif 1 - CTAGAGGT	AAATAGCT	66.19
261	Motif 2 - [T/G]G[T/A]GGGG	CCATGGGT	62.38	563	Motif 1 - CTAGAGGT	GAAAAGGC	66
262	Motif 2 - [T/G]G[T/A]GGGG	CATGGGTG	61.33	569	Motif 3 - TCTCCCAA	GCTCTAAA	63.5
266	Motif 2 - [T/G]G[T/A]GGGG	GGTGTGTT	68.12	572	Motif 1 - CTAGAGGT	CTAAAGAA	61.88
295	Motif 2 - [T/G]G[T/A]GGGG	TAATGGAA	61.41	591	Motif 1 - CTAGAGGT	GAATAGGG	64.09
328	Motif 1 - CTAGAGGT	AAAGAGTA	70.03	592	Motif 2 - [T/G]G[T/A]GGGG	AATAGGGC	61.28
338	Motif 2 - [T/G]G[T/A]GGGG	TGTTGAAA	64.46	593	Motif 1 - CTAGAGGT	ATAGGGCA	60.15
347	Motif 3 - TCTCCCAA	CATACCAT	61.93	593	Motif 2 - [T/G]G[T/A]GGGG	ATAGGGCA	63.35
355	Motif 3 - TCTCCCAA	TCACCTAA	72.75	598	Motif 1 - CTAGAGGT	GCACAGGT	70.29
359	Motif 1 - CTAGAGGT	CTAAAGTT	69.56	602	Motif 2 - [T/G]G[T/A]GGGG	AGGTGGGG	72.83
383	Motif 1 - CTAGAGGT	TTAAAGCT	64.5	**603**	**Motif 2 - [T/G]G[T/A]GGGG**	**GGTGGGG A**	**97.08**
392	Motif 1 - CTAGAGGT	TAAGAAGA	63.85	608	Motif 2 - [T/G]G[T/A]GGGG	GGATGGAT	75.24
405	Motif 2 - [T/G]G[T/A]GGGG	TTCTGGGC	62.75	612	Motif 2 - [T/G]G[T/A]GGGG	GGATGGTT	74.01
406	Motif 2 - [T/G]G[T/A]GGGG	TCTGGGCT	73.23	615	Motif 2 - [T/G]G[T/A]GGGG	TGGTTGGT	62.75
414	Motif 2 - [T/G]G[T/A]GGGG	TGTGATGT	66.35	616	Motif 2 - [T/G]G[T/A]GGGG	GGTTGGTT	77.06
416	Motif 2 - [T/G]G[T/A]GGGG	TGATGTTA	60.19	619	Motif 2 - [T/G]G[T/A]GGGG	TGGTTGGA	62.75
427	Motif 3 - TCTCCCAA	TTGCCCCC	62.97	620	Motif 2 - [T/G]G[T/A]GGGG	GGTTGGAT	78.29
428	Motif 3 - TCTCCCAA	TGCCCCCC	61.2	640	Motif 2 - [T/G]G[T/A]GGGG	TTATGGTA	60.19
429	Motif 3 - TCTCCCAA	GCCCCCCT	64.53	665	Motif 3 - TCTCCCAA	TCTCTCTG	74.95
430	Motif 3 - TCTCCCAA	CCCCCCTA	74.71	669	Motif 3 - TCTCCCAA	TCTGCTTA	61.48
431	Motif 3 - TCTCCCAA	CCCCCTAC	63.64	677	Motif 1 - CTAGAGGT	AAAGAAGA	66.66
436	Motif 2 - [T/G]G[T/A]GGGG	TACTGGGG	62.75	699	Motif 3 - TCTCCCAA	AGTCCTAG	61.37
437	Motif 2 - [T/G]G[T/A]GGGG	ACTGGGGT	76.43	710	Motif 1 - CTAGAGGT	ATTGATGT	60.67
449	Motif 3 - TCTCCCAA	TGTCCTTG	60.62	712	Motif 2 - [T/G]G[T/A]GGGG	TGATGTTT	60.19
452	Motif 2 - [T/G]G[T/A]GGGG	CCTTGGGT	65.43	718	Motif 3 - TCTCCCAA	TTTGCCAT	68.62

453	Motif 2 - [T/G]G[T/A]GGGG	CTTGGGTT	61.33	727	Motif 3 - TCTCCCAA	ATTTCCAA	66.85
457	Motif 2 - [T/G]G[T/A]GGGG	GGTTGAAG	61.39	728	Motif 3 - TCTCCCAA	TTTCCAAA	75.51
460	**Motif 2 - [T/G]G[T/A]GGGG**	**TGAAGGGT**	**81.96**	744	Motif 3 - TCTCCCAA	AGTCCCAC	78.98
461	Motif 1 - CTAGAGGT	GAAGGGTT	60.28	755	Motif 3 - TCTCCCAA	ACACGCAA	64.08
461	Motif 2 - [T/G]G[T/A]GGGG	GAAGGGTT	65.07	774	Motif 2 - [T/G]G[T/A]GGGG	GGTTTGAA	61.39
465	Motif 2 - [T/G]G[T/A]GGGG	GGTTGGAA	78.29	**785**	**Motif 3 - TCTCCCAA**	**TGTCCCTA**	**81.52**
475	Motif 3 - TCTCCCAA	CGTGCCAA	67.93	817	Motif 2 - [T/G]G[T/A]GGGG	TGTGAGTC	71.2
481	Motif 2 - [T/G]G[T/A]GGGG	AAATGGGA	65.42	829	Motif 3 - TCTCCCAA	TAACCCCA	65.38
482	Motif 2 - [T/G]G[T/A]GGGG	AATGGGAT	65.59	830	Motif 3 - TCTCCCAA	AACCCCAA	67.83
495	Motif 2 - [T/G]G[T/A]GGGG	TGCCGGTG	60.41	839	Motif 1 - CTAGAGGT	CAATAGAG	60.8
499	Motif 2 - [T/G]G[T/A]GGGG	GGTGGACT	70.15	841	Motif 1 - CTAGAGGT	ATAGAGTT	79.64
504	Motif 3 - TCTCCCAA	ACTTCCAG	75.7	845	Motif 2 - [T/G]G[T/A]GGGG	AGTTGAGA	68.47
508	Motif 1 - CTAGAGGT	CCAGGGGA	72.58	848	Motif 1 - CTAGAGGT	TGAGAGAC	68.18
508	Motif 2 - [T/G]G[T/A]GGGG	CCAGGGGA	70.33	848	Motif 2 - [T/G]G[T/A]GGGG	TGAGAGAC	69.37
510	Motif 2 - [T/G]G[T/A]GGGG	AGGGGAGG	63.88	900	Motif 3 - TCTCCCAA	ACACCCAC	77.98
511	Motif 1 - CTAGAGGT	GGGGAGGA	63.55	904	Motif 3 - TCTCCCAA	CCACCTAG	63.22
511	Motif 2 - [T/G]G[T/A]GGGG	GGGGAGGA	67.64	908	Motif 1 - CTAGAGGT	CTAGAATA	61.09
513	Motif 2 - [T/G]G[T/A]GGGG	GGAGGAGT	77.13				

Table 5: Human HNRNPH3 exon 3 Human Splicing Finder enhancer motif predictions

Seq.	Linked SR protein	Enhancer motif	Val. 0-100	Seq.	Linked SR protein	Enhancer motif	Val. 0-100
2	**SRp40**	**TTACACA**	**82.04**	**271**	**SF2/ASF (IgM-BRCA1)**	**CAGGGGA**	**80**
4	SRp40	ACACAGT	79.46	271	SF2/ASF	CAGGGGA	78.1
5	**SF2/ASF (IgM-BRCA1)**	**CACAGTG**	**80.92**	**278**	**SC35**	**GGCCTTCG**	**80.82**
5	SF2/ASF	CACAGTG	79.85	**293**	**SRp55**	**TGCTTC**	**81.35**
15	**SRp40**	**TTACACT**	**82.04**	**297**	**SRp40**	**TCAAAGG**	**90.6**
17	**SRp40**	**ACACTGG**	**92.22**	**298**	**SF2/ASF (IgM-BRCA1)**	**CAAAGGA**	**82.62**
18	**SF2/ASF (IgM-BRCA1)**	**CACTGGT**	**81.62**	**298**	**SF2/ASF**	**CAAAGGA**	**82.99**
18	**SF2/ASF**	**CACTGGT**	**80.14**	319	SF2/ASF (IgM-BRCA1)	CTCTGGG	76.77
22	**SC35**	**GGTCGCAA**	**81.99**	321	SF2/ASF (IgM-BRCA1)	CTGGGGA	76.69
25	SF2/ASF (IgM-BRCA1)	CGCAAAA	71.61	**329**	**SRp40**	**ACACAAG**	**90.3**
39	**SRp40**	**TTAATCC**	**80.3**	330	SF2/ASF (IgM-BRCA1)	CACAAGG	79.77
73	SC35	GTATTATG	75.91	**330**	**SF2/ASF**	**CACAAGG**	**81.07**
148	SF2/ASF (IgM-BRCA1)	CAGATGG	77.62	345	SC35	GGGCACAG	76.64
148	SF2/ASF	CAGATGG	79.62	**347**	**SRp40**	**GCACAGG**	**85.03**
160	**SRp40**	**TTTCAGC**	**86.77**	**348**	**SF2/ASF (IgM-BRCA1)**	**CACAGGT**	**94.38**
163	SF2/ASF (IgM-BRCA1)	CAGCCTT	78	**348**	**SF2/ASF**	**CACAGGT**	**94.41**
163	SF2/ASF	CAGCCTT	78.39	**371**	**SC35**	**GGATGGTG**	**81.32**
170	SC35	TAACACTG	76.09	403	SRp55	TGGGTC	76.48
206	SC35	GGTTAAAG	78.12	**405**	**SC35**	**GGTCACTA**	**90.84**
208	**SRp40**	**TTAAAGG**	**88.74**	**407**	**SRp40**	**TCACTAT**	**80.84**
243	SF2/ASF	TTGACGA	74.84	414	SRp55	TGCTTA	78.98
247	SF2/ASF (IgM-BRCA1)	CGATGGA	72.08	**443**	**SRp40**	**TTTCTGG**	**90.36**
252	**SC35**	**GACTACCA**	**86.79**	446	SF2/ASF (IgM-BRCA1)	CTGGGGT	74.92
254	**SRp40**	**CTACCAG**	**82.16**	**456**	**SRp40**	**TTAAAAG**	**85.33**
258	**SF2/ASF (IgM-BRCA1)**	**CAGGGGA**	**80**	492	SC35	AACTACAG	76.46
258	SF2/ASF	CAGGGGA	78.1	**494**	**SRp40**	**CTACAGA**	**81.74**
268	**SRp40**	**GCACAGG**	**85.03**	**495**	**SRp55**	**TACAGA**	**80.71**
269	**SF2/ASF (IgM-BRCA1)**	**CACAGGG**	**92.85**	**514**	**SRp40**	**TTAATGC**	**82.16**
269	**SF2/ASF**	**CACAGGG**	**92.2**	**520**	**SRp40**	**CTAATAG**	**81.68**

Table 6: Human HNRNPH3 exon 3 Human Splicing Finder silencer motif predictions

Seq.	Motif	Silencer motif	Value 0-100	Seq.	Motif	Silencer motif	Val. 0-100
6	Motif 1 - CTAGAGGT	ACAGTGTT	60.69	297	Motif 1 - CTAGAGGT	TCAAAGGA	61.63
22	Motif 3 - TCTCCCAA	GGTCGCAA	66.35	300	Motif 1 - CTAGAGGT	AAGGAGAT	63.7
63	Motif 3 - TCTCCCAA	TCTTTCAT	63.99	**320**	**Motif 2 - [T/G]G[T/A]GGGG**	**TCTGGGGA**	**83.25**
76	Motif 2 - [T/G]G[T/A]GGGG	TTATGGGG	72.25	340	Motif 1 - CTAGAGGT	GAATAGGG	64.09
77	**Motif 2 - [T/G]G[T/A]GGGG**	**TATGGGGT**	**83.25**	341	Motif 2 - [T/G]G[T/A]GGGG	AATAGGGC	61.28
79	Motif 2 - [T/G]G[T/A]GGGG	TGGGGTGA	70.71	342	Motif 1 - CTAGAGGT	ATAGGGCA	60.15
82	Motif 2 - [T/G]G[T/A]GGGG	GGTGATGG	63.28	342	Motif 2 - [T/G]G[T/A]GGGG	ATAGGGCA	63.35
84	**Motif 2 - [T/G]G[T/A]GGGG**	**TGATGGGA**	**89.15**	347	Motif 1 - CTAGAGGT	GCACAGGT	70.29
85	Motif 2 - [T/G]G[T/A]GGGG	GATGGGAA	69.35	351	Motif 2 - [T/G]G[T/A]GGGG	AGGTGGGG	72.83
147	Motif 1 - CTAGAGGT	TCAGATGG	61.69	**352**	**Motif 2 - [T/G]G[T/A]GGGG**	**GGTGGGGA**	**97.08**
149	**Motif 2 - [T/G]G[T/A]GGGG**	**AGATGGGA**	**82.32**	357	Motif 2 - [T/G]G[T/A]GGGG	GGATGGAG	75.24
150	Motif 2 - [T/G]G[T/A]GGGG	GATGGGAA	69.35	358	Motif 2 - [T/G]G[T/A]GGGG	GATGGAGA	63.28
177	Motif 3 - TCTCCCAA	GTTCCCCT	67.3	361	Motif 1 - CTAGAGGT	GGAGAGTT	74.09
178	Motif 3 - TCTCCCAA	TTCCCCTT	61.07	361	Motif 2 - [T/G]G[T/A]GGGG	GGAGAGTT	65.07
185	**Motif 2 - [T/G]G[T/A]GGGG**	**TGATGGGG**	**89.15**	365	Motif 2 - [T/G]G[T/A]GGGG	AGTTTGGG	68.47
186	**Motif 2 - [T/G]G[T/A]GGGG**	**GATGGGGT**	**80.18**	366	Motif 2 - [T/G]G[T/A]GGGG	GTTTGGGA	72.22
198	Motif 3 - TCTCCCAA	TGTCCTTG	60.62	367	Motif 2 - [T/G]G[T/A]GGGG	TTTGGGAT	72.42
201	Motif 2 - [T/G]G[T/A]GGGG	CCTTGGGT	65.43	371	Motif 2 - [T/G]G[T/A]GGGG	GGATGGTG	74.01
202	Motif 2 - [T/G]G[T/A]GGGG	CTTGGGTT	61.33	372	Motif 2 - [T/G]G[T/A]GGGG	GATGGTGT	63.28
208	Motif 1 - CTAGAGGT	TTAAAGGG	65.03	391	Motif 2 - [T/G]G[T/A]GGGG	TTTTGGGG	75.3
209	Motif 2 - [T/G]G[T/A]GGGG	TAAAGGGT	65.06	**392**	**Motif 2 - [T/G]G[T/A]GGGG**	**TTTGGGGG**	**83.25**
210	Motif 1 - CTAGAGGT	AAAGGGTT	62.91	393	Motif 2 - [T/G]G[T/A]GGGG	TTGGGGGT	70.71
210	Motif 2 - [T/G]G[T/A]GGGG	AAAGGGTT	61.32	394	Motif 2 - [T/G]G[T/A]GGGG	TGGGGGTT	75.56
214	Motif 2 - [T/G]G[T/A]GGGG	GGTTGGAA	78.29	398	Motif 2 - [T/G]G[T/A]GGGG	GGTTGTGG	72.22
224	Motif 3 - TCTCCCAA	CGTGCCAA	67.93	399	Motif 2 - [T/G]G[T/A]GGGG	GTTGTGGG	63.28
230	Motif 2 -	AAATGGGA	65.42	400	Motif 2 -	TTGTGGGT	62.75

	[T/G]G[T/A]GGGG				[T/G]G[T/A]GGGG		
231	Motif 2 - [T/G]G[T/A]GGGG	AATGGGAT	65.59	**401**	**Motif 2 - [T/G]G[T/A]GGGG**	**TGTGGGTC**	**88.1**
247	Motif 2 - [T/G]G[T/A]GGGG	CGATGGAC	68.45	419	Motif 2 - [T/G]G[T/A]GGGG	AAATGGGG	65.42
253	Motif 3 - TCTCCCAA	ACTACCAG	75.7	420	Motif 2 - [T/G]G[T/A]GGGG	AATGGGGG	76.43
257	Motif 1 - CTAGAGGT	CCAGGGGA	72.58	421	Motif 2 - [T/G]G[T/A]GGGG	ATGGGGGG	63.88
257	Motif 2 - [T/G]G[T/A]GGGG	CCAGGGGA	70.33	**422**	**Motif 2 - [T/G]G[T/A]GGGG**	**TGGGGGGG**	**87.61**
259	Motif 2 - [T/G]G[T/A]GGGG	AGGGGAGA	63.88	**423**	**Motif 2 - [T/G]G[T/A]GGGG**	**GGGGGGGT**	**84.54**
262	Motif 1 - CTAGAGGT	GGAGAAGC	64.21	424	Motif 2 - [T/G]G[T/A]GGGG	GGGGGGTA	72.48
262	Motif 2 - [T/G]G[T/A]GGGG	GGAGAAGC	60.23	444	Motif 2 - [T/G]G[T/A]GGGG	TTCTGGGG	62.75
268	Motif 1 - CTAGAGGT	GCACAGGG	61.88	**445**	**Motif 2 - [T/G]G[T/A]GGGG**	**TCTGGGGT**	**83.25**
270	Motif 1 - CTAGAGGT	ACAGGGGA	65.75	459	Motif 1 - CTAGAGGT	AAAGAATT	61.6
270	Motif 2 - [T/G]G[T/A]GGGG	ACAGGGGA	73.38	494	Motif 1 - CTAGAGGT	CTACAGAT	70.35
272	Motif 2 - [T/G]G[T/A]GGGG	AGGGGAGG	63.88	500	Motif 2 - [T/G]G[T/A]GGGG	ATTTGGGT	68.47
273	Motif 1 - CTAGAGGT	GGGGAGGC	65.52	501	Motif 2 - [T/G]G[T/A]GGGG	TTTGGGTG	71.2
273	Motif 2 - [T/G]G[T/A]GGGG	GGGGAGGC	67.64	503	Motif 2 - [T/G]G[T/A]GGGG	TGGGTGGG	70.71
275	Motif 2 - [T/G]G[T/A]GGGG	GGAGGCCT	67.1	504	Motif 2 - [T/G]G[T/A]GGGG	GGGTGGGG	76.58
294	Motif 3 - TCTCCCAA	GCTTCAAA	63.5	**505**	**Motif 2 - [T/G]G[T/A]GGGG**	**GGTGGGGA**	**97.08**

The scores for the predicted RNA motifs in Tables 2-6 were used to construct the map of regulatory RNA motifs in Figure H. I observed overlaps between an extended region of hnRNP H binding and ASF/SF2 binding within both the HNRNPH1 Exon 4 and the HNRNPH3 Exon 3 (Figure 4). I predicted that hnRNP H would have an effect on the splicing of HNRNPH3 Exon 3, to regulate the splicing of its own transcript.

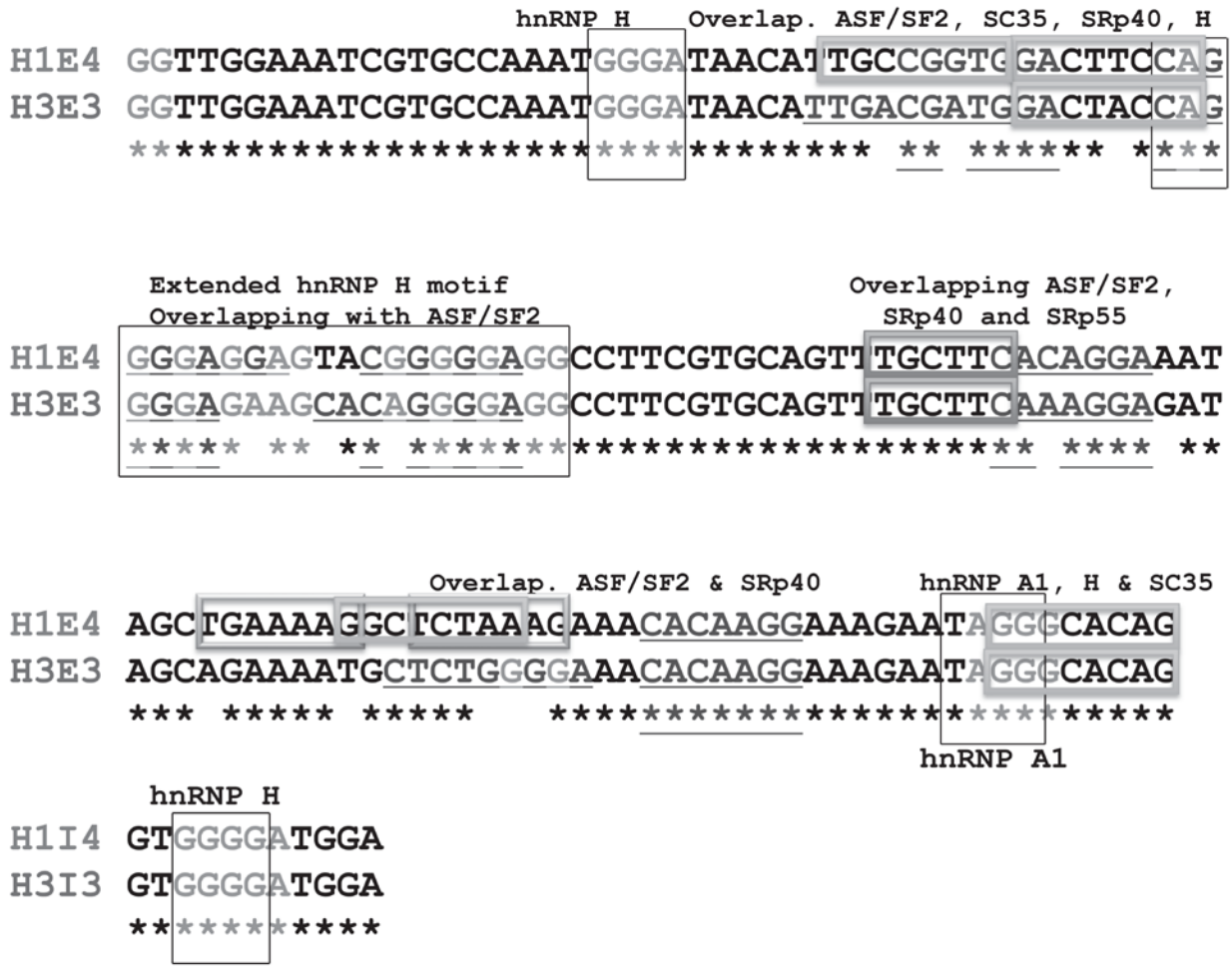


Figure 4: The map of regulatory RNA motifs for HNRNPH1 Exon 4 and HNRNPH3 Exon 3, based on the Human Splicing Finder and ESEfinder prediction scores. H1E4 stands for HNRNPH1 Exon 4; H3E3 stands for HNRNPH3 exon 3. H1I4 stands for HNRNPH1 Intron 4; H3I3 stands for HNRNPH3 Intron 3. Black thin boxed AGGG and GGGA sequences are predicted hnRNP H binding motifs. Underlined sequences are predicted ASF/SF2 binding motifs. Asterisk (*) shows exact sequence conservation. In the exonic section upstream of the 5' splice site, the TAGG is the putative hnRNP A1 motif. Downstream of the 5' splice site, the GGGG is the hnRNP H binding motif. The scores higher than 80 percent were considered, based on the values in Tables 2-6.

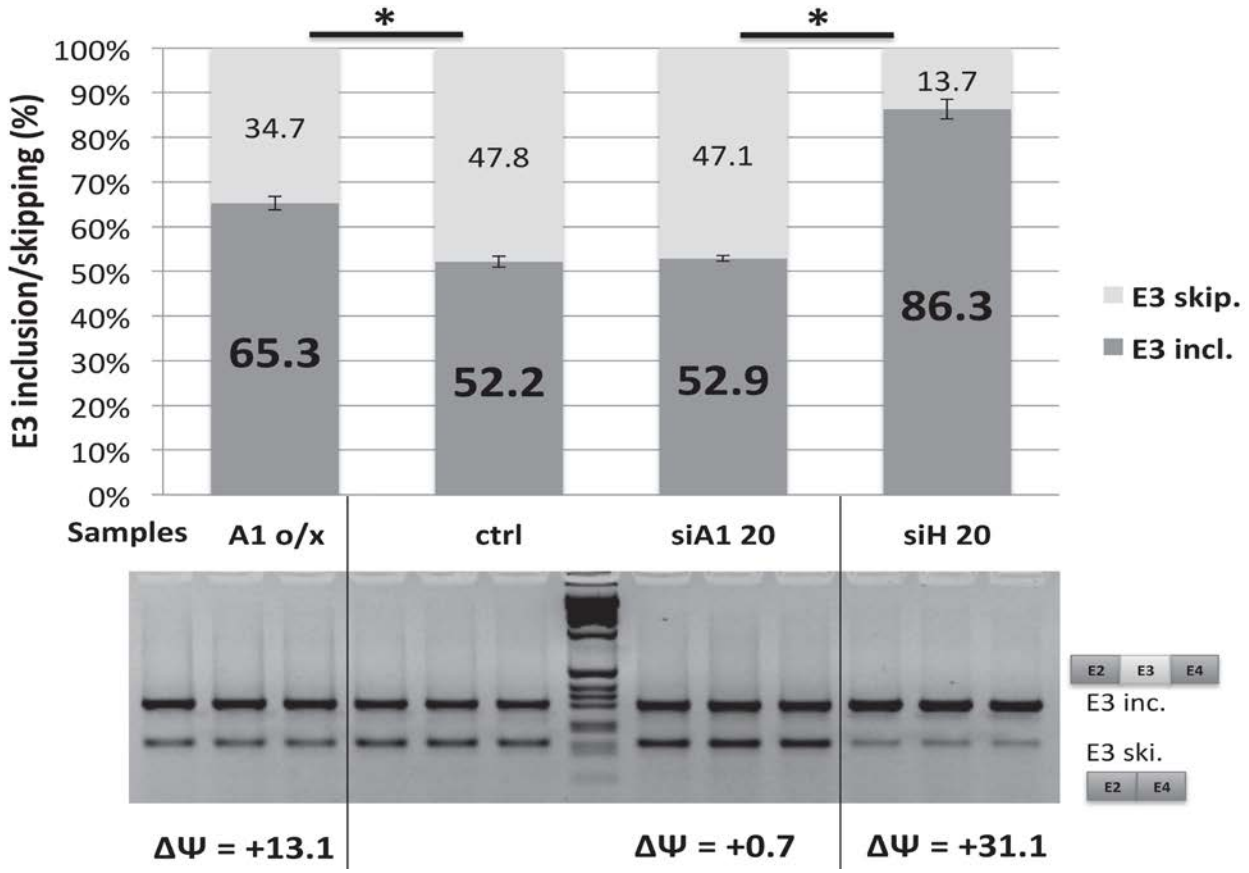


Figure 5: HnRNP H is a strong silencer, and hnRNP A1 acts as an anti-silencer, on the endogenous HNRNPH3 exon 3 in HeLa cells. HnRNP A1 and hnRNP H were knocked down (n = 3) in 6-well plates with 20 ng siRNA per well, compared to the non-targeting control siRNA. HnRNP A1 is overexpressed in HeLa cells as part of pcDNA3 vector. Total RNA was isolated and reverse transcribed. PCR was performed with the endogenous HNRNPH3 E2/E4 primers (26 cycles), and the products were run on 1.8% agarose gel. The values show percent inclusion and skipping, and the error bars show the standard deviation. ΔΨ (difference in PSI = Percent Spliced In) values show the difference in percentage of inclusion, compared to the HeLa control samples. Unpaired, two-tailed student's *t* test was performed, and the changes in the HNRNPH3 E3 inclusion were found to be statistically significant, shown with asterisk. ($p < 0.05$)

The inclusion of HNRNPH1 Exon 4 and HNRNPH3 Exon 3 decreased, and their skipping increased in a dose-dependent manner upon cisplatin treatment in HeLa cells (Figure 6). However, the gamma irradiation-induced double strand breaks, confirmed by the increase in the gamma-H2AX, did not cause a statistically significant change in the splicing of the HNRNPH3 Exon 3 in U2OS cells (Figure 7), possibly due to the different types of DNA damage and the diverse repair mechanisms acting differently in both cases. It is also possible that there can be cell type-specific and context-dependent effects, because the double-strand breaks can occur at random loci upon IR treatment.

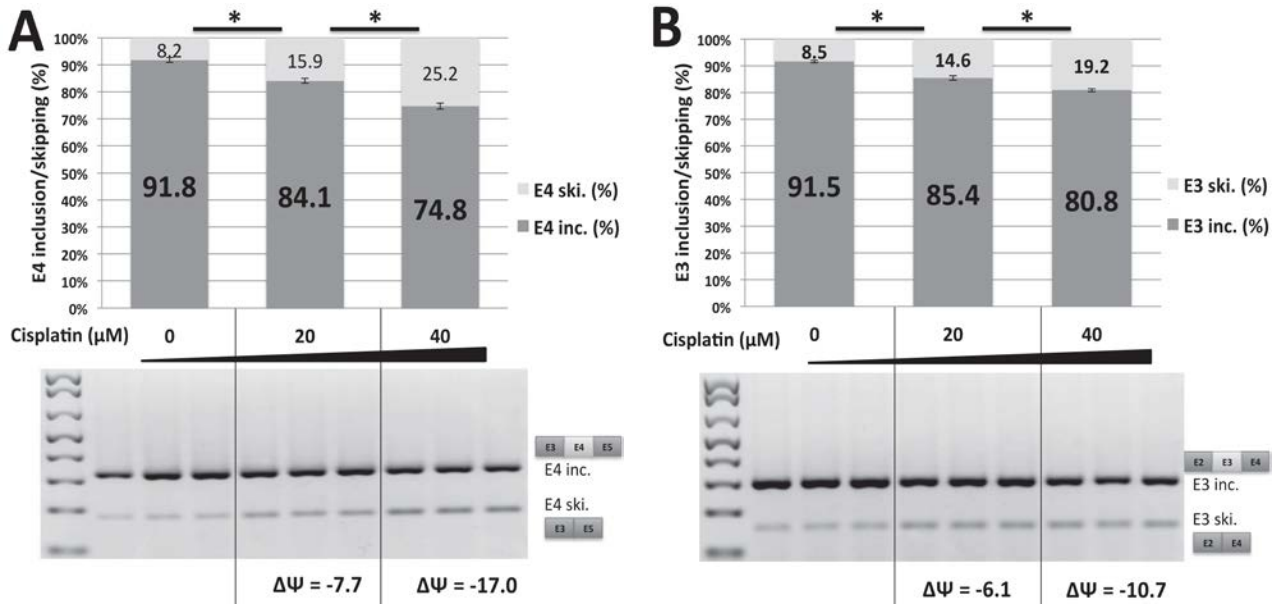


Figure 6: Cisplatin-induced DNA damage causes a dose-dependent increase in the skipping of the endogenous HNRNPH1 exon 4 and HNRNPH3 exon 3 in HeLa cells. HeLa cells (n = 3) were treated separately with 0-20-40 μM cisplatin for 14 hours. Total RNA was isolated and reverse transcribed. PCR was performed with endogenous **A)** HNRNPH1 E3/E5 and **B)** HNRNPH3 E2/E4 primers (26 cycles). The PCR products were run on 1.8% agarose gel. The electrophoretic bands were analyzed by the ImageGauge software. The values show percent inclusion and skipping, and the error bars show the standard deviation. ΔΨ (difference in PSI = Percent Spliced In) values show the difference in percentage

of inclusion, compared to the HeLa samples treated with 0 μM cisplatin. Unpaired two-tailed student's t test was performed, and the changes in the endogenous HNRNPH1 E4 and HNRNPH3 E3 inclusion were found to be statistically significant. ($p < 0.05$)

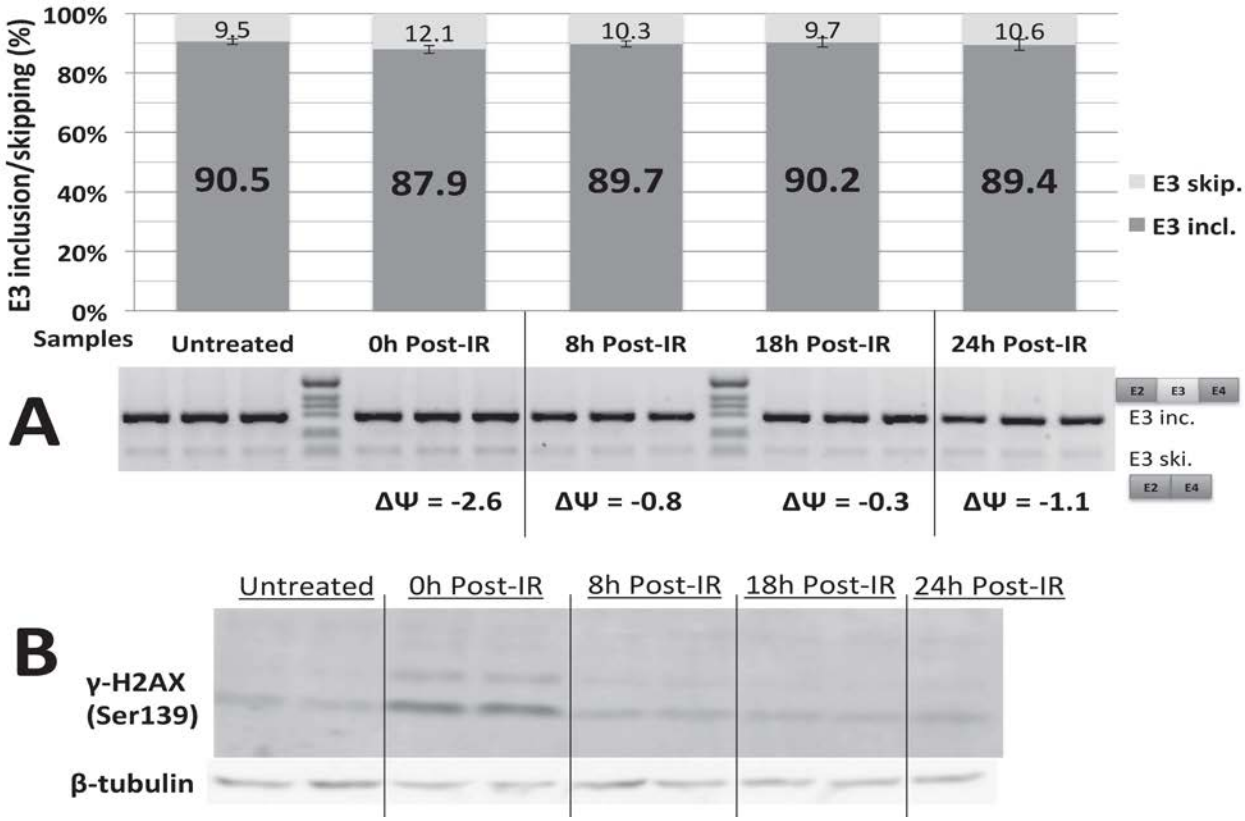


Figure 7: Gamma irradiation-induced double strand DNA breaks do not affect the splicing of the endogenous HNRNPH3 exon 3 in U2OS human osteosarcoma cells. The U2OS cells (n = 3) were grown on 60 mm plates overnight, and they were treated with 2.8-3 Gy of γ-radiation by exposing them to ^{137}Cs for 15 minutes. Following the γ-IR treatment, the U2OS cells were either harvested immediately (0h Post-IR), or allowed to recover until the indicated time points. **A)** Double strand DNA breaks induced by γ-IR do not cause a change in the splicing of HNRNPH3 Exon 3. PCR was performed with the endogenous HNRNPH3 E2/E4 primers (26 cycles), and the products were run on 1.8% agarose gel. The values show percent inclusion and skipping, and the error bars show the standard deviation. ΔΨ values show the difference in percentage of inclusion, compared to untreated U2OS cells. **B)** Double strand DNA breaks are confirmed by Western Blot against the phosphoserine 139 γ-H2AX ab. β-tubulin is the loading control.

3.1.5 DRB-induced transcription stress causes an alternative splicing response that is similar to cisplatin-induced genotoxic stress

The cisplatin-induced single-strand DNA damage can be repaired by the transcription-coupled repair (Damsma *et al.*, 2007), and can cause changes in splicing by changing the rate of transcription (Thomas and Lieberman, 2013). Therefore, I asked whether the transcriptional stress affects the splicing of HNRNPH1 and HNRNPH3 in HeLa cells.

The 5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) molecule is a known inhibitor of transcription elongation, and its effects on the splicing of several exons were previously studied. However, the regulatory RNA motifs that affect the splicing response to genotoxic and/or transcriptional stress remain unknown, and its effect on the splicing of HNRNPH1 and HNRNPH3 was not shown.

Poly(G) motifs and their location-dependent effects were previously studied, and the G-rich motifs adjacent to the 5' splice sites in vertebrate introns were observed to have enhancer effects, contributing to the selection of the splice site and the intron-exon boundaries, and affecting the strength or weakness of the splice site (McCullough and Berget, 1997; Carlo *et al.*, 2000; McCullough and Berget, 2000; Xiao *et al.*, 2009). In a genomewide search, an enrichment of GU-rich sequences was observed in a subset of exons with higher inclusion upon DRB treatment (Fong *et al.*, 2014). Since the 5' splice site GGGG motif was conserved across the human HNRNPH1 Exon 4 and HNRNPH3 Exon 3, I asked whether DRB would act to increase the inclusion of human endogenous HNRNPH1 Exon 4 and HNRNPH3 Exon 3.

The previously-published functions of the 5' splice site poly(G) sequences do not seem to fit the effect of DRB treatment that I observed on the splicing of HNRNPH1 and

HNRNPH3. I tested the effect of DRB-induced transcription elongation inhibition in HeLa cells, and I observed that there is a dose-dependent increase in the skipping of human endogenous HNRNPH1 Exon 4 and HNRNPH3 Exon 3 upon DRB treatment (Figure 8).

During the gel quantification, the background can be uneven or misleading. The fluoroconjugated primers allow the quantification of the specific PCR product without the agarose gel background EtBr. For this reason, I used TAMRA-conjugated HNRNPH3 Exon 3 primer to confirm the change in the inclusion and skipping. The Exon 3 skipping increased with a similar trend and bigger magnitude upon same doses of DRB-induced transcription elongation inhibition in HeLa cells (Figure 9).

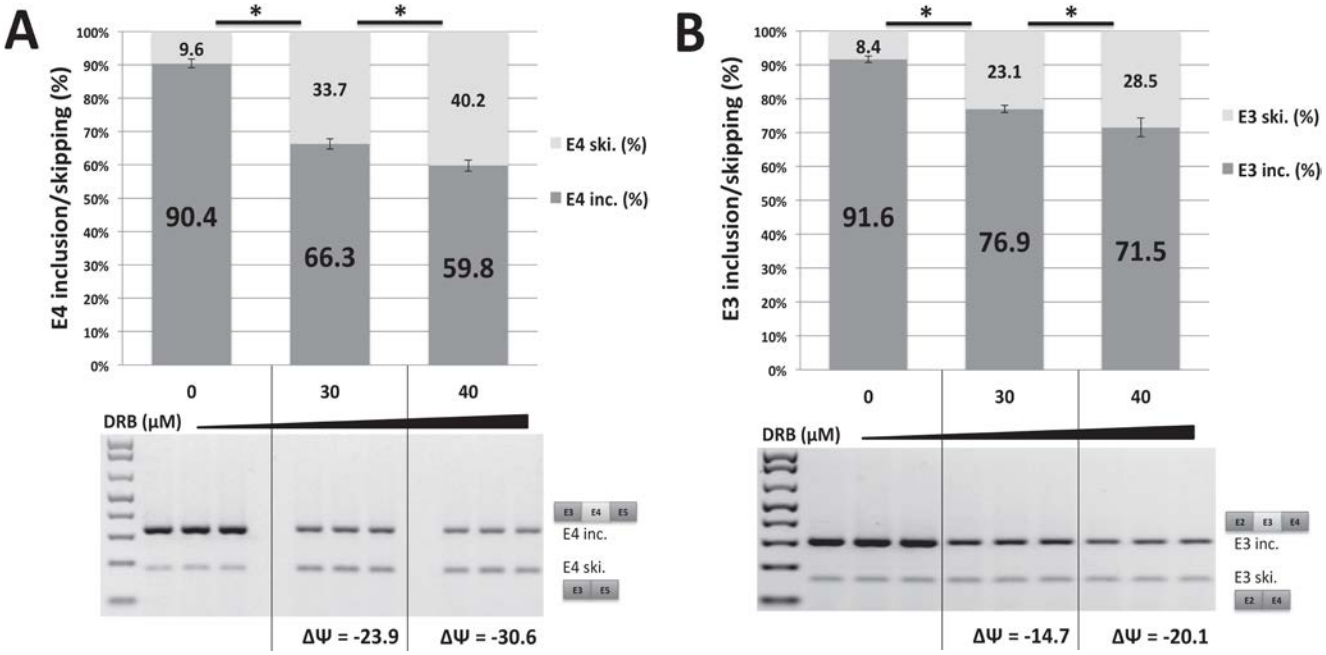


Figure 8: DRB-induced transcription elongation inhibition causes dose-dependent decrease in the inclusion of endogenous HNRNPH1 exon 4 and HNRNPH3 exon 3. A) HNRNPH1 Exon 4 (E4). **B)** HNRNPH3 Exon 3 (E3). HeLa cells (n = 3) were treated separately with the indicated doses of DRB for 24 hours. Total RNA was isolated and reverse transcribed. PCR was performed with HNRNPH1 E3/E5 and HNRNPH3 E2/E4 primers respectively (26 cycles), and the PCR products were run on 1.8% agarose

gel. Electrophoretic bands were analyzed by ImageGauge software. Numbers represent percentage of exon inclusion and skipping. ΔΨ (difference in PSI = Percent Spliced In) values show the difference in percentage of inclusion; compared to 0 μM DRB. Unpaired two-tailed student's t test was performed, and the changes in the endogenous HNRNPH1 E4 and HNRNPH3 E3 inclusion were found to be statistically significant. ($p < 0.05$)

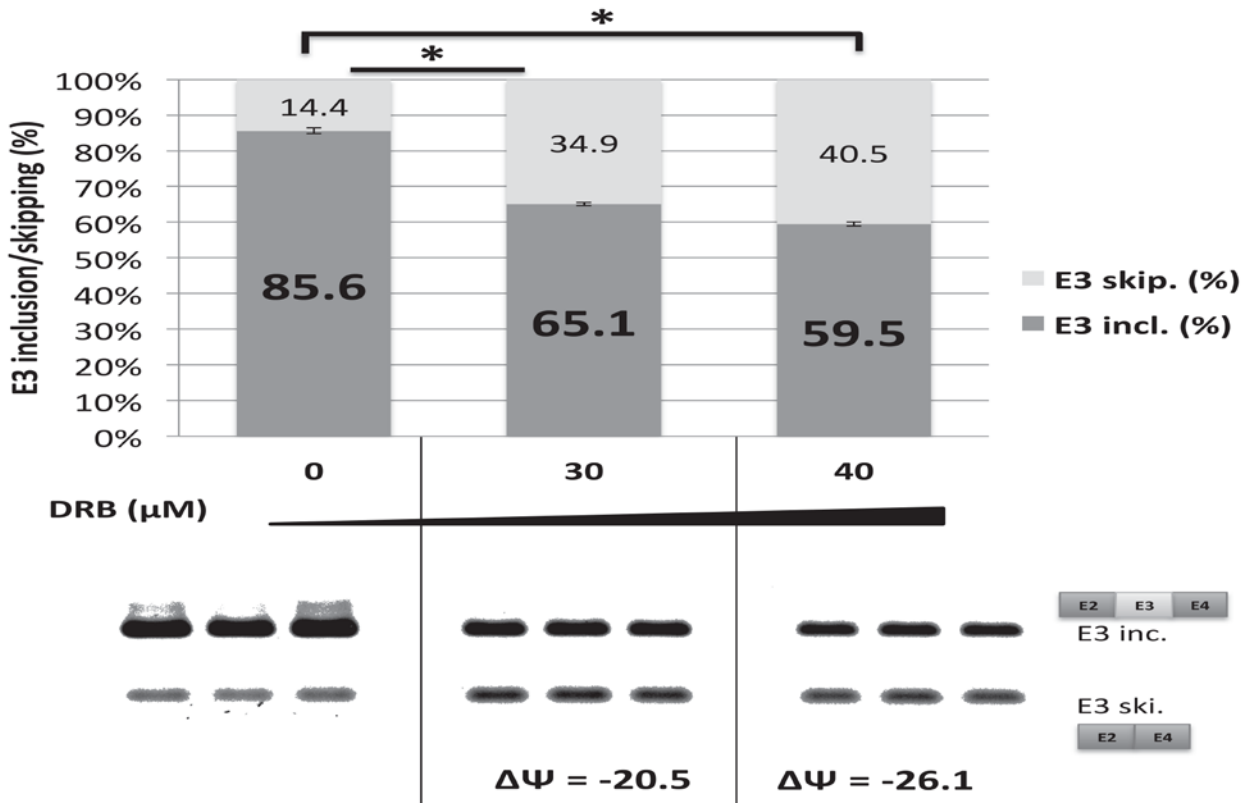


Figure 9: DRB-induced increase in the skipping of the endogenous HNRNPH3 exon 3 is confirmed by PCR with fluoroconjugated primers. HeLa cells (n = 3) were treated separately with the indicated doses of DRB for 24 hours. Total RNA was isolated and reverse transcribed. PCR was performed with the TAMRA-conjugated HNRNPH3 E2 forward and untagged E4 (*-E2/E4) primers for 26 cycles, and the PCR products were run on 1.8% agarose gel without Ethidium Bromide. The bands were visualized at 540/575 nm (excitation/emission), and the analysis was performed with the ImageGauge software. Numbers represent percentage of exon inclusion and skipping. ΔΨ (difference in PSI = Percent Spliced In) values show the difference in percentage of inclusion; compared to 0 μM DRB. Asterisks above the column charts show statistically significant difference. Unpaired, two-tailed student's t test was performed, and the exon inclusion changes were found to be statistically significant ($p < 0.05$).

3.2 THE HNRNP H-REGULATED GRIN1 CI CASSETTE EXON RESPONDS TO DRB AND TRANSCRIPTIONAL PAUSE-RELEASE

The GRIN1 CI exon, which is subject to tissue-specific and developmentally-specific regulation (Black and Grabowski, 2003) similar to HNRNPH3 Exon 3, is also under the regulation of hnRNP A1 and hnRNP H similar to that of HNRNPH3 Exon 3 (Han *et al.*, 2005). These exons were found together in the genomewide search for a UAGG-GGGG splicing code (Han *et al.*, 2005), and is inducible upon depolarization by KCl treatment (An and Grabowski, 2007), which raised the possibility of using the previously-cloned GRIN1 splicing reporter constructs and their GGGG mutants as a model system to test the effect of poly(G) stretches and the UAGG-GGGG splicing code in the mechanism of alternative splicing upon DRB-induced transcription stress.

The wild-type splicing reporter and the mutants of the wild-type CI construct are as described (Han *et al.*, 2005) (Figure 10). C1wt0 is the wild-type rat GRIN1 CI exon, cloned into the pBPSV pA^+ reporter, as described (Han *et al.*, 2005). The E17 mutant has two mutations in each UAGG motif, turning them into UAAU (Han *et al.*, 2005) (Figures 10-11). The T8 mutant is a triple mutant, destroying two exonic UAGG sites and the +6 GGGG close to the 5' splice site (Han *et al.*, 2005) (Figures 10-11). 5m2 and 5m4 are GGGG mutants, at the +6 after the 5' splice site downstream of the CI cassette exon (Han *et al.*, 2005) (Figures 10-11). The aim of using these UAGG and GGGG mutants is to find out whether the mutated motifs affect the DRB-induced splicing changes, if they cause increased skipping or inclusion, or whether they have any role.

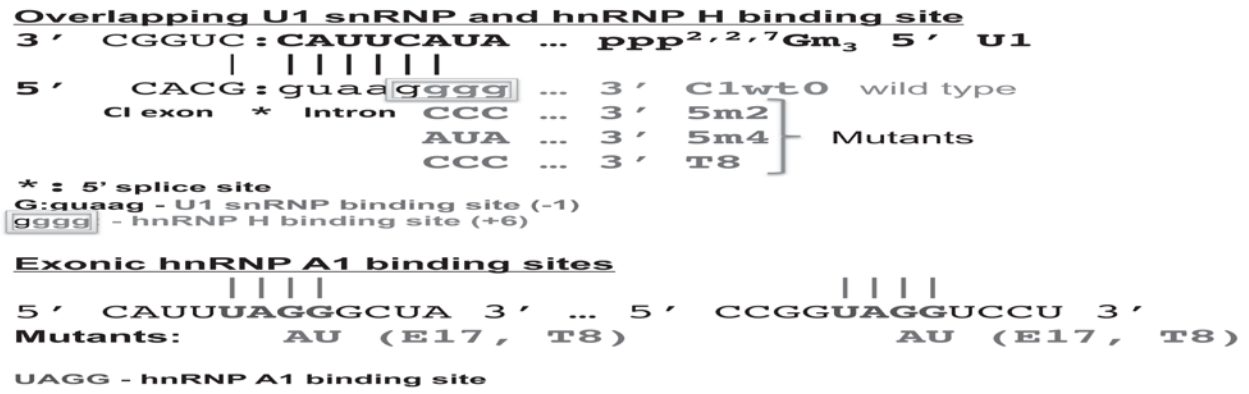


Figure 10: The schematic of the mutants, with respect to the overlapping binding sites for U1 snRNP and hnRNP H at the 5' splice site, and the exonic binding sites for hnRNP A1. The overlapping U1 snRNP and hnRNP H binding sites at the 5' splice site immediately downstream of the CI exon, and the splicing reporter mutant constructs used for the DRB-induced transcriptional pausing and pause-release experiment, along with the location of these binding sites with respect to the CI exon are shown. CIwt0 stands for the wild type pA^+ CI construct. 5m2 and 5m4 are three-nucleotide mutations at the +6 GGGG motif immediately downstream of the 5' splice site. E17 is a double mutant with 2 UAGG sites mutated into UAAU. Combining the mutated motif in E17 and the mutated motif in 5m2, T8 is a triple mutant with two UAGGs and one GGGG destroyed. Asterisk with a colon shows 5' splice site. Lowercase shows the intron, uppercase shows the exon. Constructs were cloned by Kyounga Han (Han *et al.*, 2005).

Table 7: GRIN1 CI exon Human Splicing Finder enhancer motif predictions

Seq	Linked SR protein	Enh. motif	Value 0-100 (variation)	Seq	Linked SR protein	Enh. motif	Value 0-100 (variation)
7	SRp40	agtctgg	**81.02 (+13.40 %)**	234	SC35	GCCTCCAG	79.84 (+19.20 %)
12	SC35	ggctgttg	**82.54 (+30.04 %)**	242	SRp40	CTTCAAG	**86.29 (+37.44 %)**
19	SC35	ggtccgtc	77.01 (+7.87 %)	248	SF2/ASF (IgM-BRCA1)	GAGACGT	70.77 (+0.88 %)
21	SRp55	tccgtc	**81.35 (+28.65 %)**	248	SF2/ASF	GAGACGT	**87.30 (+53.01 %)**
24	SC35	gtctgttg	78.55 (+14.03 %)	252	SF2/ASF (IgM-BRCA1)	CGTAGGT	**80.69 (+34.52 %)**
34	SF2/ASF (IgM-BRCA1)	cacagtg	**80.92 (+35.31 %)**	252	SF2/ASF	CGTAGGT	78.33 (+19.82 %)
34	SF2/ASF	cacagtg	79.85 (+25.42 %)	257	SC35	GTCCTCCA	**84.63 (+38.41 %)**
39	SRp55	tgaggc	77.19 (+12.74 %)	266	SRp40	AGACACG	**86.95 (+40.45 %)**
46	SF2/ASF (IgM-BRCA1)	ggcagga	76.08 (+18.87 %)	267	SF2/ASF (IgM-BRCA1)	GACACGg	72.15 (+5.57 %)

46	SF2/ASF	ggcagga	**84.62 (+43.10 %)**	267	SF2/ASF	GACACGg	**86.55 (+50.21 %)**
52	SRp40	agtcagg	**82.51 (+20.23 %)**	269	SF2/ASF (IgM-BRCA1)	CACGgta	71.00 (+1.66 %)
53	SF2/ASF	gtcaggg	76.47 (+12.92 %)	286	SC35	caccccag	79.04 (+16.00 %)
58	SC35	ggcctgag	**83.47 (+33.73 %)**	286	SF2/ASF (IgM-BRCA1)	cacccca	**81.92 (+38.70 %)**
61	SF2/ASF (IgM-BRCA1)	ctgagtg	74.69 (+14.18 %)	286	SF2/ASF	cacccca	79.27 (+23.26 %)
66	SRp40	tgaccgg	**84.67 (+30.07 %)**	288	SF2/ASF (IgM-BRCA1)	ccccagt	72.92 (+8.18 %)
70	SF2/ASF (IgM-BRCA1)	cggccag	76.92 (+21.74 %)	290	SRp40	ccagtcc	79.10 (+4.66 %)
83	SF2/ASF (IgM-BRCA1)	cagcggc	79.15 (+29.31 %)	293	SC35	gtcccgcg	**86.48 (+45.80 %)**
83	SF2/ASF	cagcggc	76.41 (+12.70 %)	295	SF2/ASF (IgM-BRCA1)	cccgcgt	77.54 (+23.83 %)
84	SRp55	agcggc	78.85 (+19.10 %)	297	SRp55	cgcgtc	**90.76 (+64.66 %)**
87	SC35	ggccccaa	**90.96 (+63.79 %)**	299	SC35	cgtctctg	79.35 (+17.23 %)
89	SRp40	ccccaag	**80.06 (+9.03 %)**	305	SRp40	tgactcc	**86.65 (+39.08 %)**
90	SF2/ASF (IgM-BRCA1)	cccaagg	74.54 (+13.66 %)	306	SC35	gactcctc	**82.48 (+29.79 %)**
91	SF2/ASF (IgM-BRCA1)	ccaagga	77.38 (+23.31 %)	311	SF2/ASF (IgM-BRCA1)	ctccctt	77.62 (+24.09 %)
107	SF2/ASF (IgM-BRCA1)	ggcaggt	74.31 (+12.87 %)	311	SF2/ASF	ctccctt	73.33 (+1.28 %)
107	SF2/ASF	ggcaggt	**82.59 (+35.55 %)**	318	SC35	gccccgtg	**81.93 (+27.58 %)**
113	SRp55	tgtggc	**80.97 (+27.19 %)**	319	SF2/ASF (IgM-BRCA1)	ccccgtg	72.54 (+6.88 %)
116	SC35	ggccgcca	**87.83 (+51.22 %)**	322	SRp55	cgtgtc	**80.84 (+26.70 %)**
122	SF2/ASF (IgM-BRCA1)	cacgctt	70.85 (+1.14 %)	328	SRp40	tgtctcc	**80.36 (+10.40 %)**
135	SRp40	ctaatcg	**83.23 (+23.51 %)**	329	SC35	gtctccct	76.77 (+6.88 %)
146	SRp55	tacata	**85.90 (+46.04 %)**	334	SRp40	cctctcc	**83.11 (+22.96 %)**
184	SF2/ASF (IgM-BRCA1)	AGCCCGA	72.77 (+7.66 %)	337	SF2/ASF (IgM-BRCA1)	ctccctt	77.62 (+24.09 %)
184	SF2/ASF	AGCCCGA	76.99 (+14.86 %)	337	SF2/ASF	ctccctt	73.33 (+1.28 %)
188	SC35	CGACCCTA	78.55 (+14.03 %)	348	SRp40	tcccacc	78.68 (+2.75 %)
189	SC35	GACCCTAA	78.18 (+12.55 %)	349	SF2/ASF (IgM-BRCA1)	cccacct	78.08 (+25.66 %)
215	SC35	GGCTATCA	78.43 (+13.53 %)	351	SC35	cacctctg	**80.64 (+22.40 %)**
227	SF2/ASF (IgM-BRCA1)	CACCCTG	79.38 (+30.09 %)	353	SRp40	cctctgc	**84.97 (+31.43 %)**
227	SF2/ASF	CACCCTG	77.64 (+17.23 %)	365	SRp40	cttcaac	**81.20 (+14.22 %)**
233	SC35	GGCCTCCA	**88.63 (+54.43 %)**				

Table 8: GRIN1 CI exon Human Splicing Finder silencer motif predictions

Seq.	Motif	Sil. motif	Value 0-100 (var.)	Seq.	Motif	Sil. motif	Value 0-100 (var.)
7	Motif 2 - [T/G]G[T/A]GGGG	agtctggc	61.28 (+3.20 %)	211	Motif 2 - [T/G]G[T/A]GGGG	TTAGGGCT	70.18 (+25.45 %)
15	Motif 2 - [T/G]G[T/A]GGGG	tgttggtc	80.14 (+50.35 %)	216	Motif 3 - TCTCCCAA	GCTATCAC	60.21 (+0.52 %)
27	Motif 2 - [T/G]G[T/A]GGGG	tgttgagc	75.30 (+38.24 %)	222	Motif 3 - TCTCCCAA	ACCTCCAC	60.79 (+1.99 %)
37	Motif 1 - CTAGAGGT	agtgaggc	68.14 (+20.36 %)	223	Motif 3 - TCTCCCAA	CCTCCACC	67.86 (+19.65 %)
37	Motif 2 - [T/G]G[T/A]GGGG	agtgaggc	76.43 (+41.06 %)	226	Motif 3 - TCTCCCAA	CCACCCTG	70.99 (+27.49 %)
39	Motif 2 - [T/G]G[T/A]GGGG	tgaggcag	69.37 (+23.43 %)	234	Motif 3 - TCTCCCAA	GCCTCCAG	61.64 (+4.10 %)
47	Motif 2 - [T/G]G[T/A]GGGG	gcaggagt	60.23 (+0.56 %)	235	Motif 3 - TCTCCCAA	CCTCCAGC	65.36 (+13.40 %)
48	Motif 1 - CTAGAGGT	caggagtc	63.23 (+8.08 %)	245	Motif 1 - CTAGAGGT	CAAGAGAC	79.62 (+49.05 %)
52	Motif 2 - [T/G]G[T/A]GGGG	agtcaggg	61.28 (+3.20 %)	247	Motif 1 - CTAGAGGT	AGAGACGT	73.34 (+33.36 %)
54	Motif 2 - [T/G]G[T/A]GGGG	tcagggcc	70.18 (+25.45 %)	252	Motif 2 - [T/G]G[T/A]GGGG	CGTAGGTC	63.08 (+7.71 %)
62	Motif 2 - [T/G]G[T/A]GGGG	tgagtgac	69.37 (+23.43 %)	258	Motif 3 - TCTCCCAA	TCCTCCAA	72.75 (+31.86 %)
66	Motif 2 - [T/G]G[T/A]GGGG	tgaccggc	65.06 (+12.65 %)	259	Motif 3 - TCTCCCAA	CCTCCAAA	**82.26 (+55.65 %)**
69	Motif 3 - TCTCCCAA	ccggccag	63.22 (+8.05 %)	262	Motif 1 - CTAGAGGT	CCAAAGAC	60.50 (+1.26 %)
78	Motif 3 - TCTCCCAA	cctggcag	61.36 (+3.40 %)	272	Motif 2 - [T/G]G[T/A]GGGG	Ggtaaggg	65.03 (+12.58 %)
82	Motif 1 - CTAGAGGT	gcagcggc	65.09 (+12.73 %)	273	Motif 2 - [T/G]G[T/A]GGGG	gtaagggg	61.98 (+4.96 %)
82	Motif 2 - [T/G]G[T/A]GGGG	gcagcggc	60.23 (+0.56 %)	274	Motif 1 - CTAGAGGT	taagggga	65.15 (+12.88 %)
87	Motif 3 - TCTCCCAA	ggccccaa	68.21 (+20.53 %)	274	Motif 2 - [T/G]G[T/A]GGGG	taagggga	**80.20 (+50.51 %)**
88	Motif 3 - TCTCCCAA	gccccaag	61.64 (+4.10 %)	279	Motif 2 - [T/G]G[T/A]GGGG	ggaagagc	61.98 (+4.96 %)
95	Motif 2 - [T/G]G[T/A]GGGG	ggagcagg	60.23 (+0.56 %)	280	Motif 1 - CTAGAGGT	gaagagca	71.99 (+29.97 %)
102	Motif 2 - [T/G]G[T/A]GGGG	gcaagggc	61.98 (+4.96 %)	285	Motif 3 - TCTCCCAA	gcacccca	71.43 (+28.57 %)
103	Motif 1 - CTAGAGGT	caagggca	65.85 (+14.62 %)	286	Motif 3 - TCTCCCAA	caccccag	66.96 (+17.41 %)
103	Motif 2 - [T/G]G[T/A]GGGG	caagggca	60.31 (+0.77 %)	292	Motif 3 - TCTCCCAA	agtcccgc	65.37 (+13.41 %)
108	Motif 2 - [T/G]G[T/A]GGGG	gcaggtgt	60.23 (+0.56 %)	301	Motif 3 - TCTCCCAA	tctctgac	67.60 (+18.99 %)
111	Motif 2 - [T/G]G[T/A]GGGG	ggtgtggc	**80.18 (+50.45 %)**	307	Motif 3 - TCTCCCAA	actcctcc	65.01 (+12.53 %)
113	Motif 2 - [T/G]G[T/A]GGGG	tgtggccg	73.23 (+33.07 %)	310	Motif 3 - TCTCCCAA	cctccctt	**83.14 (+57.85 %)**
117	Motif 3 - TCTCCCAA	gccgccac	62.06 (+5.16 %)	316	Motif 3 - TCTCCCAA	ttgccccg	62.55 (+6.37 %)
167	Motif 1 -	AAAGAG	70.09 (+25.22	330	Motif 3 - TCTCCCAA	tctccctc	92.56 (+81.39 %)

	CTAGAGGT	TG	%)				
169	Motif 1 - CTAGAGGT	AGAGTG GT	74.65 (+36.62 %)	332	Motif 3 - TCTCCCAA	tccctctc	60.05 (+0.11 %)
169	Motif 2 - [T/G]G[T/A]GGGG	AGAGTG GT	73.38 (+33.44 %)	334	Motif 3 - TCTCCCAA	cctctccc	67.86 (+19.65 %)
171	Motif 2 - [T/G]G[T/A]GGGG	AGTGGT AG	65.59 (+13.98 %)	336	Motif 3 - TCTCCCAA	tctccctt	**88.95 (+72.38 %)**
174	Motif 2 - [T/G]G[T/A]GGGG	GGTAGA GC	65.03 (+12.58 %)	341	Motif 3 - TCTCCCAA	cttaccat	62.81 (+7.02 %)
175	Motif 1 - CTAGAGGT	GTAGAG CA	73.12 (+32.80 %)	346	Motif 3 - TCTCCCAA	catcccac	**82.71 (+56.79 %)**
180	Motif 1 - CTAGAGGT	GCAGAG CC	71.74 (+29.35 %)	355	Motif 3 - TCTCCCAA	tctgcccc	73.67 (+34.18 %)
188	Motif 3 - TCTCCCAA	CGACCC TA	60.38 (+0.95 %)	356	Motif 3 - TCTCCCAA	ctgcccca	60.45 (+1.12 %)
210	Motif 2 - [T/G]G[T/A]GGGG	TTTAGG GC	68.11 (+20.27 %)	357	Motif 3 - TCTCCCAA	tgccccat	68.71 (+21.77 %)
211	Motif 1 - CTAGAGGT	TTAGGG CT	65.80 (+14.51 %)	364	Motif 3 - TCTCCCAA	tcttcaac	67.60 (+18.99 %)

Figure 11: The map of regulatory RNA motifs for GRIN1 CI exon and its flanking introns, based on the Human Splicing Finder prediction scores. The scores higher than 80 percent were considered, based on the values in Tables 7-8. TAGA and TAGG motifs are predicted for the binding of hnRNP A1. AGGG / GGGA / GGGG motifs are predicted for the binding of hnRNP H. There are overlaps between SRSF1, SRSF2 and SRSF5. The sites of the mutations (Han *et al.*, 2005) were shown with the thick box and arrows. The map was adapted from Kyounga Han, with the recent version of the HSF online splicing motif prediction tool.

3.2.1 The alternative splicing of HNRNPH1 E4 and hnRNP H-regulated GRIN1 CI exon is regulated similarly upon transcription stress

The structural similarities between HNRNPH1 Exon 4 and GRIN1 CI exon, and their similar splicing regulation by hnRNP A1 and hnRNP H caused me to ask whether their splicing regulation is similar upon transcriptional stress, and whether GRIN1 can be used as a model to investigate the role of UAGG and GGGG motifs in the alternative splicing regulation upon transcriptional stress. Since the brain-specific *GRIN1* is not expressed in HeLa cells, I transfected the CI cassette exon as part of a pA^+ splicing reporter construct into HeLa cells.

Upon DRB treatment, we observed a dose-dependent increase in the skipping of GRIN1 CI exon, similar to the HNRNPH1 Exon 4 splicing in HeLa cells (Figure 8). In light of the data showing that DRB-induced transcription elongation inhibition caused increased skipping of GRIN1 CI exon (Figure 12), I constructed an RNA map of GRIN1 CI (Figure 11) as an adaptation of the previously published work (Han *et al.*, 2005) in order to elucidate the splicing motifs that are responsible for the effect of DRB, based on the RNA motif predictions of Human Splicing Finder and ESEfinder algorithms. There were 2 exonic UAGG motifs and 2 exonic UAGA motifs identified as possible hnRNP A1-binding sites, as well as 1 exonic AGGG and 1 intronic GGGG motifs as possible hnRNP H-binding sites (Figure 11). The UAGG motifs were previously found by our lab to be bound by purified hnRNP H, albeit with a lower affinity than AGGG or GGGA motifs (Grabowski Lab, unpublished data).

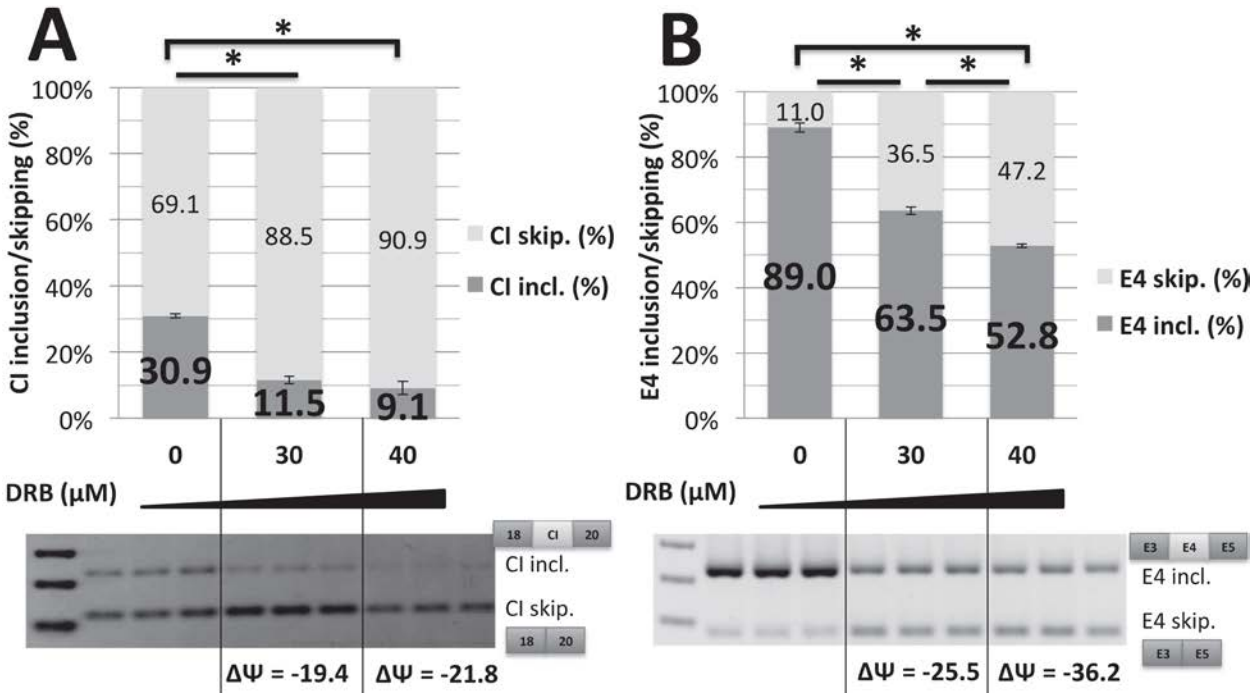


Figure 12: Endogenous HNRNPH1 exon 4 and pA+ reporter-cloned GRIN1 CI exon respond similarly to DRB-induced transcription elongation inhibition in the GRIN1 pA+ reporter-transfected HeLa cells. A) GRIN1 CI. **B)** HNRNPH1 E4. HeLa cells (n = 3) were grown and transfected separately with pA+-C1wt0 expressing wild-type GRIN1 CI cassette exon and its flanking introns. After removing the transfection medium, the indicated doses of DRB were added onto the cells and they were incubated with DRB for 14 hours. Total RNA was isolated and reverse transcribed. PCR was performed with CI and HNRNPH1 E3/E5 primers (26 cycles), and the products were run on 1.8% agarose gel. The electrophoretic bands were analyzed by ImageGauge software. Numbers represent percent inclusion and skipping. ΔΨ (difference in PSI = Percent Spliced In) values show the difference in percentage of inclusion; compared to 0 µM DRB. Asterisks above the column charts show statistically significant difference. Unpaired, two-tailed student's t test was performed, and the changes in the GRIN1 CI and HNRNPH1 E4 inclusion were found to be statistically significant. ($p < 0.05$)

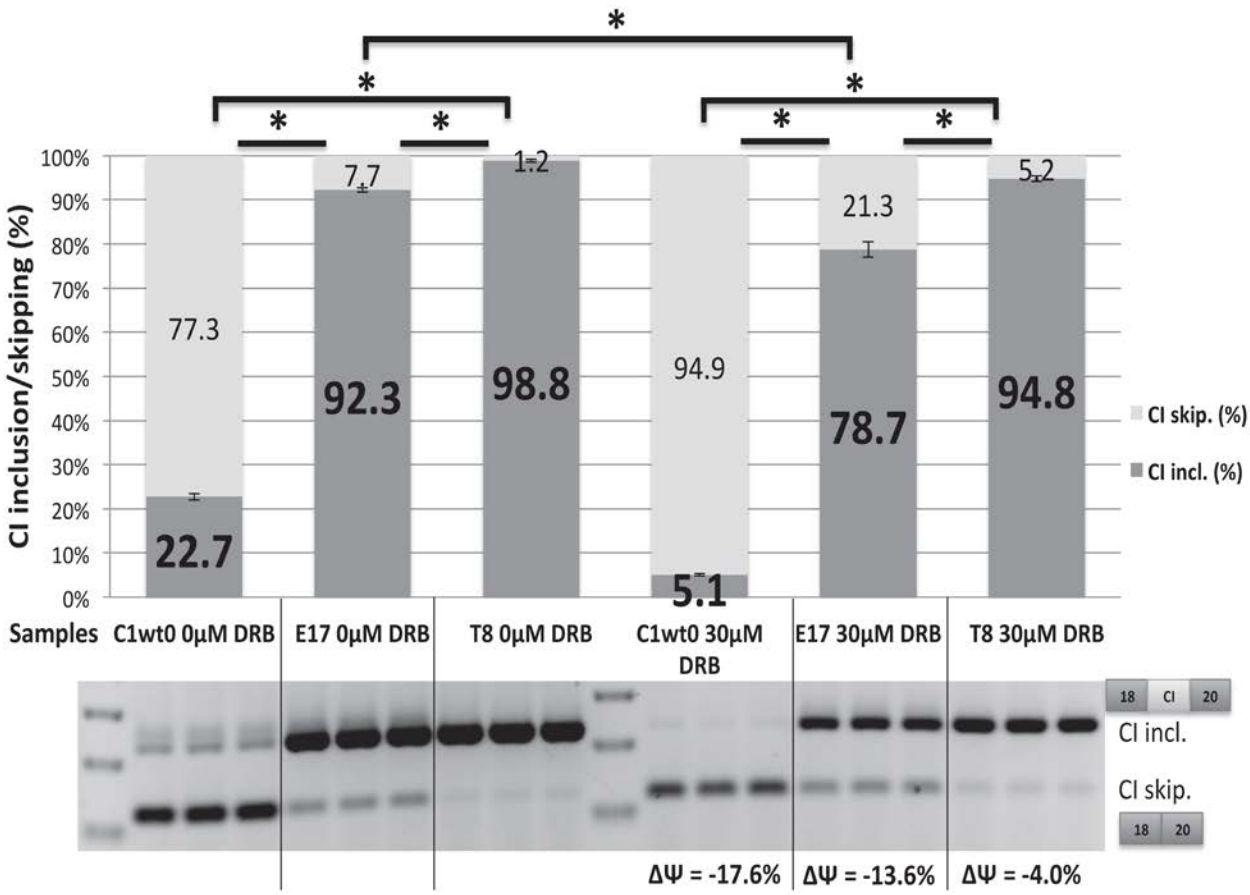


Figure 13: DRB causes an increase in the pA⁺ reporter-cloned GRIN1 CI skipping in wild type-, but the response is truncated in the UAGG mutant-transfected HeLa cells. 250 ng/well E17 (UAGG double mutant) and T8 (UAGGs + GGGG triple mutant) were transfected to HeLa cells (n = 3), versus wild-type C1wt0 to each well and the transfection medium was replaced 4 hours post-transfection by DMEM + 10% FBS with 0 µM versus 30 µM DRB. Total RNA was isolated and reverse transcribed. PCR was performed with CI primers (26 cycles), and the products were run on 1.8% agarose gel. The electrophoretic bands were analyzed by ImageGauge software. Numbers represent percent exon inclusion and skipping. Asterisks above the column chart show statistically significant difference. Unpaired, two-tailed student's t test was performed, and the changes in the GRIN1 CI inclusion were found to be statistically significant. ($p < 0.05$)

3.2.2 UAGG and GGGG motifs are necessary for the response to DRB in the pA$^+$ reporter-transfected HeLa cells

After observing the effect of DRB on the splicing of GRIN1 CI cassette exon, and by making a map of putative regulatory RNA motifs, I asked which of the RNA motifs were responsible for the splicing effect of DRB. I chose single and combined mutations of the UAGGs and GGGG to monitor the effects of DRB on splicing, since the shared RNA motifs include two exonic UAGG motifs and a GGGG close to 5´ splice site. I used E17, T8, 5m2, and 5m4 mutants, as described (Han *et al.*, 2005).

I transfected the wild type versus mutant pA$^+$ splicing reporters into HeLa cells, and afterwards I treated them with DRB for 14 hours. E17, the mutant with 2 UAGG motifs destroyed, had a truncated response to DRB, whereas the triple mutant T8 lacking 2 UAGGs and 1 GGGG was almost completely unresponsive to DRB treatment (Figure 13). Regardless of whether the wild type or the mutant splicing reporter were transfected, the splicing of endogenous HNRNPH1 Exon 4 responded similarly to DRB treatment, as an internal control (Figure 15).

These data showed that the two UAGG motifs as well as the GGGG motif are required for the splicing response to DRB-induced transcriptional stress. The mutated UAGG motifs cause an increase in CI inclusion in untreated control cells (Figure 13). GGGG motif causes a statistically-significant increase in exon inclusion in control cells without DRB treatment, showing that UAGG and GGGG motifs are acting as silencers for the CI exon (Figure 14).

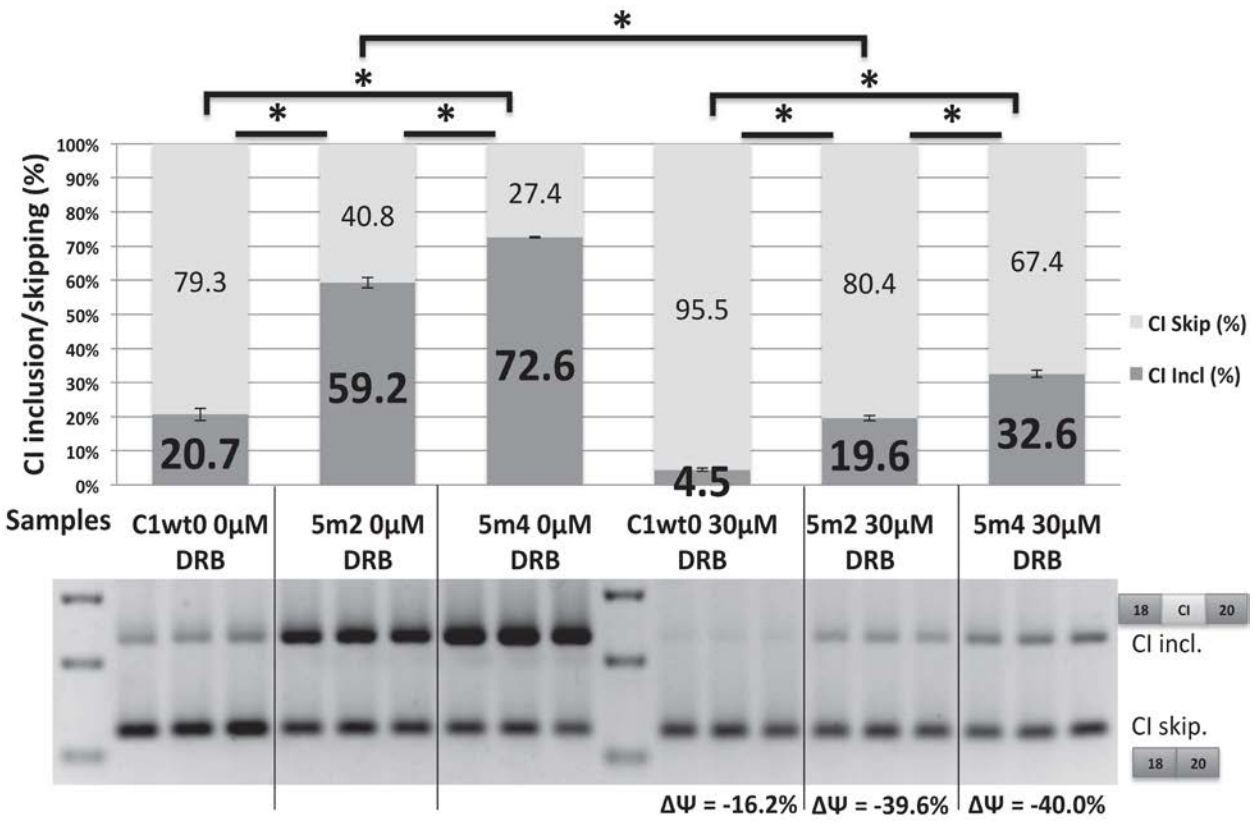


Figure 14: The intronic GGGG silencer motif immediately downstream of the 5' splice site inhibits the effect of DRB on GRIN1 cassette exon splicing. 250 ng of the wild-type C1wt0 and the GGGG mutants 5m2 and 5m4 were transfected into each well, and after 4 hours the transfection medium was replaced by the DMEM + 10% FBS including 0 μM or 30 μM DRB. Total RNA was isolated, reverse transcribed, and the PCR was performed with the primers recognizing the flanking exons of the GRIN1 CI cassette exon (26 cycles). PCR products were run on 1.8% agarose gel for 2 hours. The electrophoretic bands were analyzed by ImageGauge software. Numbers represent percentage of inclusion and skipping. ΔΨ (difference in PSI = Percent Spliced In) values show the difference in percentage of inclusion; compared to 0 μM DRB of each reporter transfection. Asterisks above the column chart show statistically significant difference. Unpaired, two-tailed student's t test was performed, and the changes in the GRIN1 CI inclusion were found to be statistically significant. ($p < 0.05$)

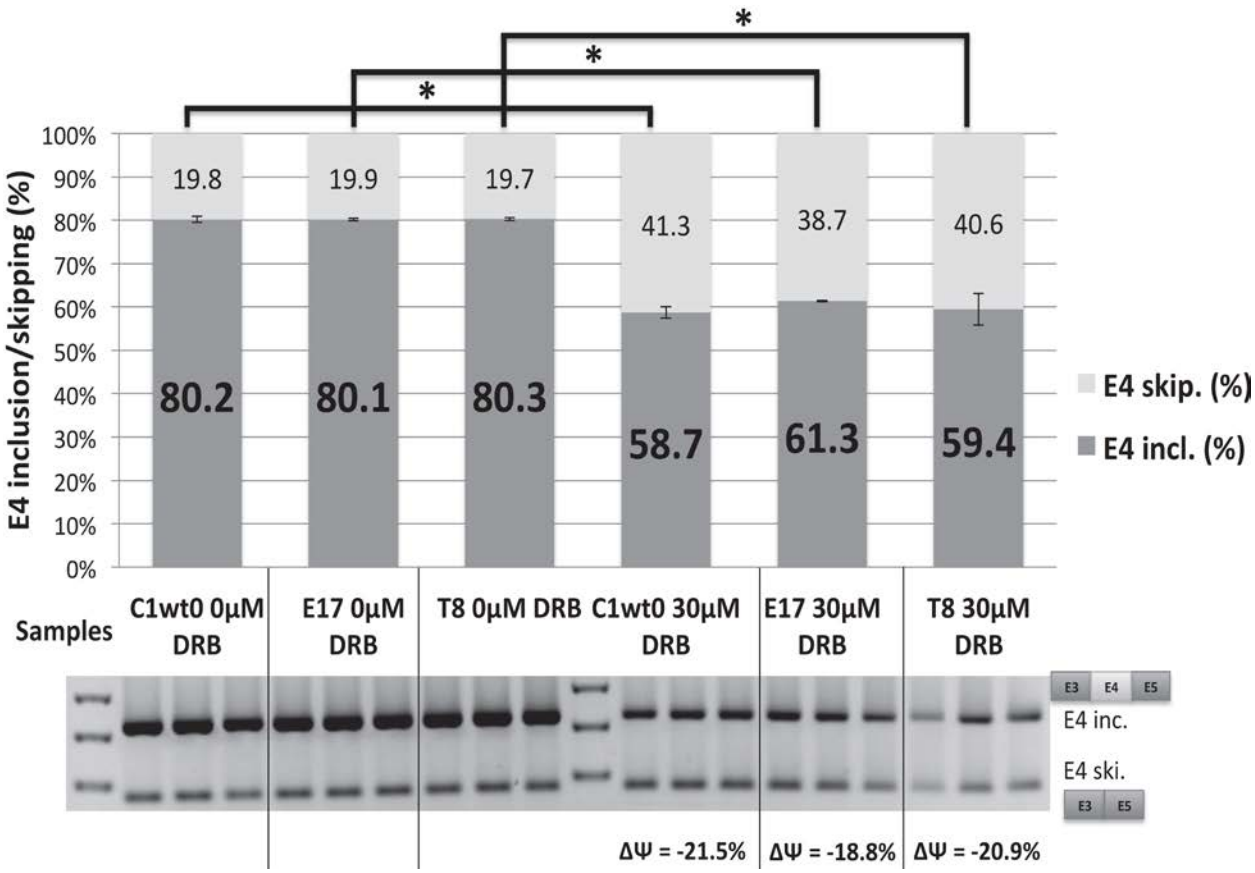


Figure 15: DRB treatment causes an increase in the endogenous HNRNPH1 exon 4 skipping in wild type and mutant GRIN1 pA+ reporter-transfected HeLa cells. 250 ng of the wild-type C1wt0 and the GGGG mutants 5m2 and 5m4 were transfected into each well, and after 4 hours the transfection medium was replaced by the DMEM + 10% FBS including 0 µM or 30 µM DRB. Total RNA was isolated, reverse transcribed, and the PCR was performed with HNRNPH1 E3/E5 primers (26 cycles). PCR products were run on 1.8% agarose gel for 2 hours. The electrophoretic bands were analyzed by ImageGauge software. Numbers represent percentage of inclusion and skipping. ΔΨ (difference in PSI = Percent Spliced In) values show the difference in percentage of inclusion; compared to 0 µM DRB of each reporter transfection. Asterisks above the column chart show statistically significant difference. Unpaired, two-tailed student's t test was performed, and the changes in the E4 inclusion were found to be statistically significant. ($p < 0.05$)

3.2.3 Mutated GGGG motif leads to a bigger splicing response to DRB in the pA$^+$ reporter-transfected HeLa cells

In the earlier studies, the 5' splice site sequences were found to be enhancers that determine intron-exon boundaries, strength of splice sites, and recruiting U1 snRNP to the 5' splice site (McCullough and Berget, 1997; Carlo *et al.*, 2000; McCullough and Berget, 2000). The role of GGGG is important for the regulation of GRIN1 splicing, as it was previously shown to be directly bound by hnRNP H (Han *et al.*, 2005). In addition, the GGGG motif of the HNRNPH3 transcript was shown to be bound by hnRNP H (Stein, 2015), which further supports its role in hnRNP H-mediated splicing. For this reason, I asked whether the mutations in GGGG have an effect over the DRB-induced splicing changes in GRIN1 CI exon inclusion. In the GGGG motif mutants (5m2 and 5m4), I observed a statistically significant increase in the effect of DRB, compared to that of the wild type (Figure 14). In untreated cells, GGGG it acts as a silencer, and the GGGG mutants have a big increase in GRIN1 CI exon inclusion (Figure 14). These data showed that the GGGG motif acts as an antagonist of the effect of DRB on the splicing of GRIN1 CI exon.

3.2.4 DRB-induced alternative splicing reverses to untreated levels upon pause-release in the pA$^+$ reporter-transfected HeLa cells

Transcriptional pausing is observed in a significant portion of genes in human and mouse embryonic stem cells (Guenther *et al.*, 2007; Gilchrist *et al.*, 2010). Pausing has a role in regulating the timing of gene expression, and the transcriptional pause release

is possibly important in the regulation of alternative splicing. In addition, DRB-induced increase in exon skipping is a phenomenon worth investigating in detail about whether it is reversible, or if it is a permanent change in splicing patterns. Therefore, I decided to treat the pA^+ reporter-transfected cells with DRB, and wash away the DRB, to check for the splicing upon transcriptional pause release.

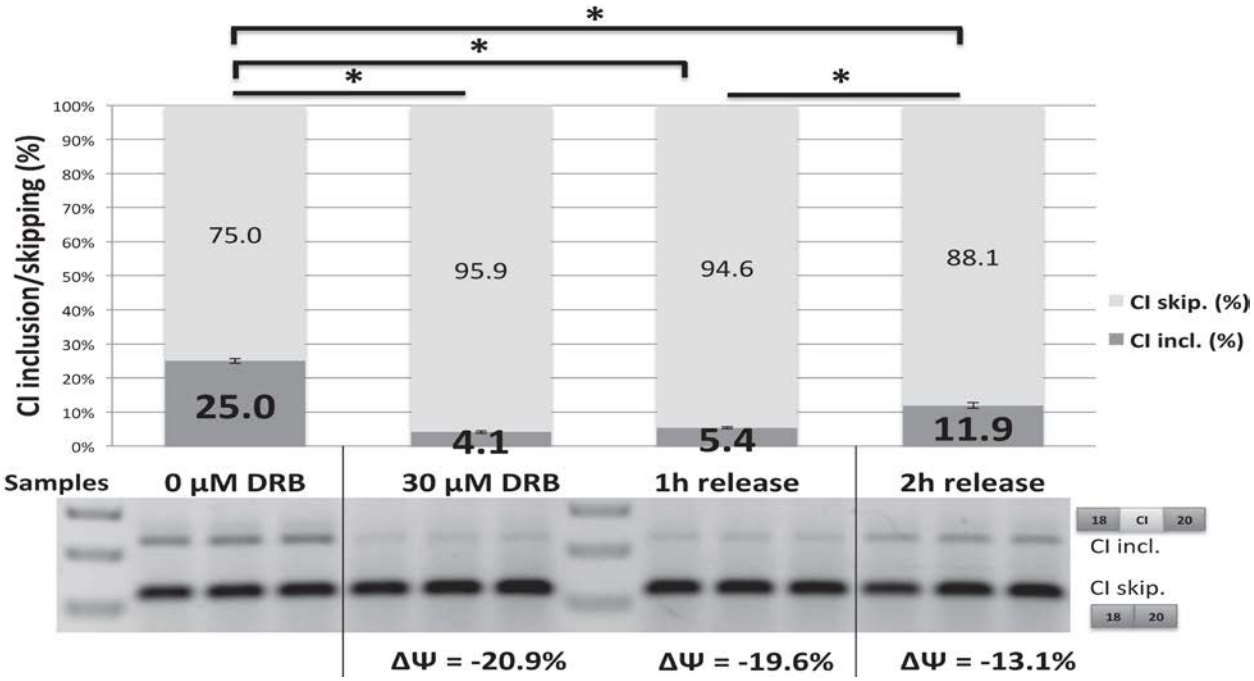


Figure 16: The silencing effect of DRB on the GRIN1 CI exon is reversible in the wild type pA^+ splicing reporter-transfected HeLa cells. 250 ng of the wild-type pA^+-C1wt0 was transfected into each well of HeLa cells, and after 4 hours the transfection medium was replaced by DMEM + 10% FBS including 0 µM or 30 µM DRB. Following the 14h DRB treatment, cells were either harvested (0 and 30 µM samples) or washed with DMEM + 10% FBS and incubated at 37°C for 30m and 1h respectively. Total RNA was isolated, reverse transcribed, and PCR was performed with the primers recognizing the flanking exons of the GRIN1 CI cassette exon (26 cycles). PCR products were run on 1.8% agarose gel for 2 hours. The electrophoretic bands were analyzed by ImageGauge software. Numbers represent percentage of inclusion and skipping. ΔΨ (difference in PSI = Percent Spliced In) values show the

difference in percentage of inclusion; compared to 0 μM DRB. Asterisks above the column chart show statistically significant difference. Unpaired, two-tailed student's t test was performed, and the changes in the GRIN1 CI inclusion were found to be statistically significant. ($p < 0.05$)

Upon washing away the DRB, the splicing change that happened over the course of 14-hour DRB treatment was reversed (Figures 16, 17). In the GGGG motif mutants, I observed a rapid reversal in which the inclusion increased within only 1 hour after washing away the DRB in 5m2 and 5m4 (Figure 17).

Also, the UAGG-GGGG splicing code was necessary for the splicing response to transcriptional pausing and pause release, because the E17 mutant had a diminished response to DRB and pause release, and T8 had no response at all (Figure 18). Beta actin-encoding transcript, ACTB, does not respond to the DRB treatment (Figure 19), showing the specificity of the effect of DRB, rather than a genomewide random perturbation of splicing.

These data point to the effect of DRB on the splicing of HNRNPH1, HNRNPH3 and GRIN1 through the action of hnRNP A1 and hnRNP H. In both cases, hnRNP H has a big effect on splicing. The combination of functional hnRNP A1 and hnRNP H motifs is necessary for the response to DRB, and GGGG is a strong silencer in CI. In HNRNPH3 exon 3, hnRNP H is a strong silencer and hnRNP A1 is an anti-silencer. The exonic UAGG - intronic GGGG code is present in HNRNPH1 exon 4 and HNRNPH3 exon 3, supporting the role of hnRNP A1 and hnRNP H in the splicing response to DRB.

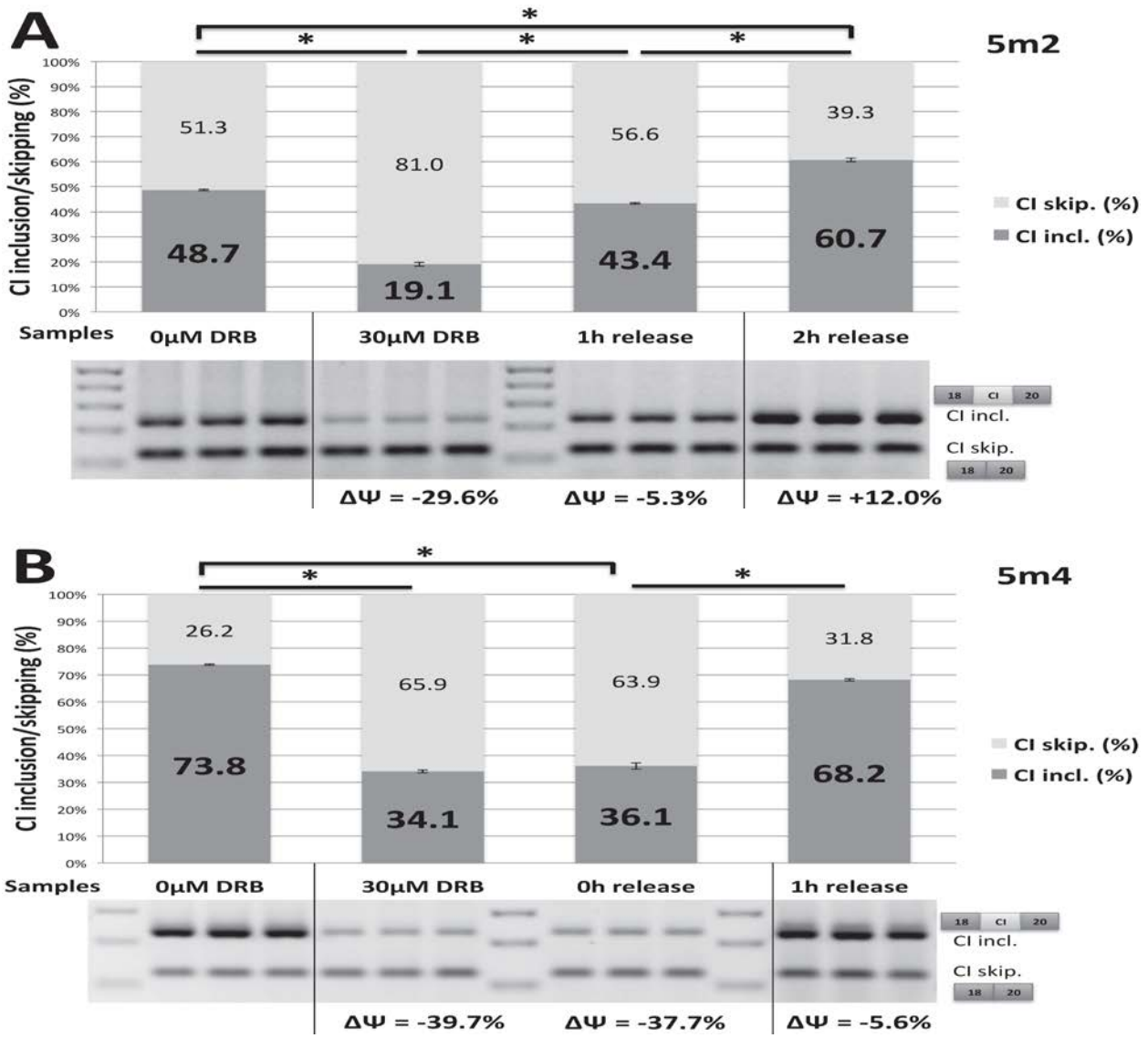


Figure 17: The silencing effect of DRB on GRIN1 CI exon is reversible, and it is faster in GGGG mutant than that of the wild type reporter-transfected HeLa cells. 250 ng of the **A)** pA^+-5m2 or **B)** 5m4 mutant was transfected into each well of HeLa cells, and after 4 hours the transfection medium was replaced by DMEM + 10% FBS including 0 μM or 30 μM DRB. Following the 14h DRB treatment, the wells were either harvested (0μM and 30μM samples) or washed with DMEM + 10% FBS and incubated at 37°C until the indicated pause-release time points. Total RNA was isolated, reverse transcribed, and PCR was performed with the primers recognizing the flanking exons of the GRIN1 CI cassette exon (26 cycles). PCR products were run on 1.8% agarose gel for 2 hours. The electrophoretic bands were analyzed by ImageGauge software. Numbers represent percentage of inclusion and skipping. ΔΨ

(difference in PSI = Percent Spliced In) values show the difference in percentage of inclusion; compared to 0 μM DRB. Asterisks above the column chart show statistically significant difference. Unpaired, two-tailed student's t test was performed, and the changes in the GRIN1 CI inclusion were found to be statistically significant. ($p < 0.05$)

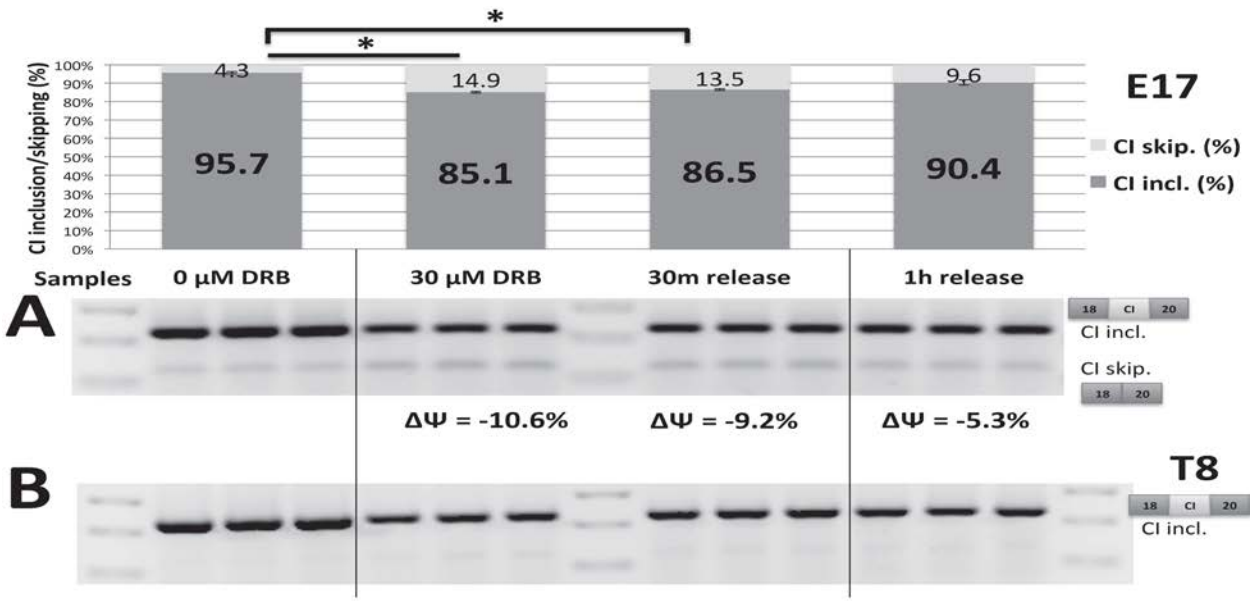


Figure 18: UAGG-GGGG splicing code is necessary for the splicing response to DRB-induced transcriptional pausing and pause-release. 250 ng of the **A)** pA^+-E17 double; or **B)** T8 triple mutant was transfected into each well of HeLa cells (n = 3), and after 4 hours the transfection medium was replaced by DMEM + 10% FBS including 0 - 30 μM DRB. Following the 14h DRB treatment, the wells were either harvested (0μM and 30μM samples) or washed with DMEM + 10% FBS and incubated at 37°C for 30m or 1h. Total RNA was isolated, reverse transcribed, and PCR was performed with the primers recognizing the flanking exons of the GRIN1 CI cassette exon (26 cycles). PCR products were run on 1.8% agarose gel for 2 hours. The electrophoretic bands were analyzed by ImageGauge software. Numbers represent percent inclusion and skipping. ΔΨ (difference in PSI = Percent Spliced In) values show the difference in percentage of inclusion; compared to 0 μM DRB. Asterisks above the column chart show statistically significant difference of the E17 mutant exon inclusion. Unpaired, two-tailed student's t test was performed, and the changes in the GRIN1 CI inclusion were statistically significant ($p < 0.05$). The CI skipping of the T8 mutant was not quantifiable.

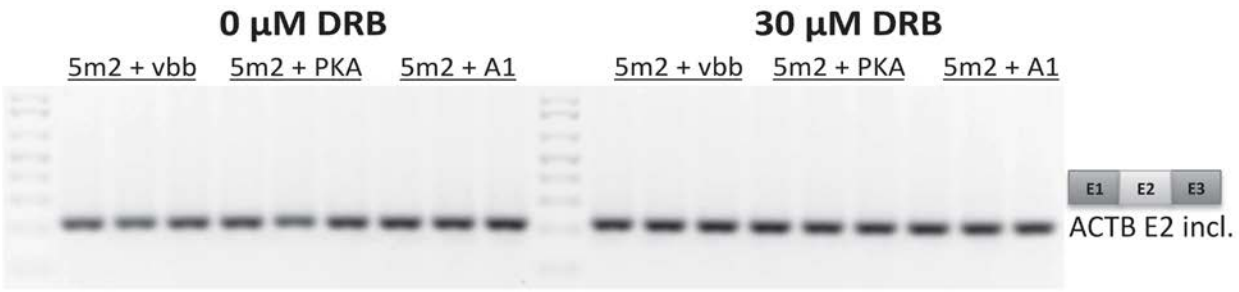


Figure 19: ACTB exon 2 does not respond to the DRB-induced transcription elongation inhibition in pA+ splicing reporter-transfected HeLa cells. 5m2 mutant was co-transfected with the pcDNA3.1/PKA/A1 plasmids into each well of HeLa cells, and after 4 hours the transfection medium was replaced by DMEM + 10% FBS including 0 µM or 30 µM DRB. Following the 14h DRB treatment, total RNA was isolated, reverse transcribed, and PCR was performed with the ACTB E1/E3 primers. PCR products were run on 1% agarose gel for 1 hour. The gel was visualized with ImageGauge. There is no ACTB exon 2 skipping upon DRB treatment or co-transfection with these plasmids.

4.0 DISCUSSION

RNA splicing was elucidated as a mechanism of diversifying the proteome of the cell, and as a fast response system to multiple environmental inputs. Its impact in human development and disease is a more recent phenomenon, and cancer is among the very important diseases directly tied to the aberrant splicing due to *cis*-acting mutations and changes in the splicing factor stoichiometry, localization, or activity in *trans*. The hnRNP H family is important in this framework, due to its overexpression in different cancers.

The main finding of this study was the splicing repressor effect of DRB-induced transcription on the exons sharing the UAGG-GGGG splicing code. Our main question was to understand the role of hnRNP H in the response to genotoxic and transcriptional stress. Cisplatin-induced single stranded DNA damage and DRB-induced transcriptional pausing had a similar effect on the splicing of HNRNPH1 exon 4 and HNRNPH3 exon 3, whereas gamma irradiation-induced double strand breaks did not have an effect. That is a sign of transcriptional pausing-induced splicing changes.

The effect of DRB-induced transcriptional pausing on splicing is not random, as the control exon (ACTB exon 2) did not respond to DRB. Pausing of transcription had a bigger effect when the 5' splice site GGGG motif was destroyed, and the triple mutant that lacks the two exonic UAGGs and the 5' splice site GGGG lost its ability to respond to DRB, meaning that the hnRNP A1-binding UAGG and hnRNP H-binding GGGG

motifs were necessary for the response to transcriptional pausing. Purified hnRNP H has the ability to bind to UAGG motifs, albeit with a lower affinity (Grabowski Lab, unpublished data), hinting about a possible competition by multiple splicing factors for the same RNA motif.

One of the novel findings of this study was the autoregulation of HNRNPH3 Exon 3 splicing by hnRNP H. This autoregulation, which leads to a transcript level regulation by alternative splicing, is a putative fine-tuning system that acts in the response to very diverse types of cellular stresses. The role of hnRNP A1 and hnRNP H on the splicing of GRIN1 transcript, and their responsiveness to stimuli were already previously studied in our lab (Han *et al.*, 2005; An and Grabowski, 2007). Transcriptional stress was known to induce changes in splicing, but the role of the hnRNP H-binding GGGG motifs in response to transcriptional pausing and pause-release, along with the necessity of the previously found UAGG-GGGG splicing code for the response to transcriptional stress and its reversibility are among the novel findings of this study.

The role of transcription stress on genomewide splicing alterations is a well-studied field of research, that shapes our understanding of how the splicing machinery responds to erroneous changes in the cell homeostasis, possibly being part of a bigger general stress response. Different types of cellular stress, such as the KCl excitation of neurons, flavopiridol- or DRB-induced transcription inhibition, and cisplatin-induced DNA damage may not seem to share a lot in common; but they somewhat converge to give rise to similar changes in the splicing of human HNRNPH1 and its paralog HNRNPH3. These data suggest a stress-mediated splicing response pathway, adjusting the splicing

of several transcripts for sparing some time for the cells to recover the damage caused by stress, and contribute to the overall phenotype upon homeostatic changes.

Previous studies by various research groups concluded that there are transcripts with higher exon inclusion upon transcription stress, such as the fibronectin EDI exon. These groups tried to explain the phenomenon with “recruitment coupling” (de la Mata *et al.*, 2006) by asserting that the CTD (C-terminal domain) tail of RNA polymerase II takes a role of docking pad for the RNA processing factors, transferred onto the newly emerging nascent transcript; and also with “kinetic coupling” (de la Mata *et al.*, 2003), in which these splicing factors have a longer temporal window period for performing their splicing regulation activity upon transcriptional pausing. However, these studies did not point to the role of a particular splicing factor in the regulation of splicing upon pausing or transcriptional stress. In addition, the transcriptional pause-release is known to be prevalent in stem cells (Guenther *et al.*, 2007; Core *et al.*, 2008; Gilchrist *et al.*, 2010) as a possible regulator of timing, but its effects on the splicing were still unknown.

In this study, I found that the intronic GGGG motif downstream of the 5’ splice site, which was previously shown to be directly bound by hnRNP H (Han *et al.*, 2005), has a silencer effect on splicing, and masks the effect of DRB on the splicing of GRIN1 CI exon. Mutations in the GGGG motif caused an increase in the splicing repressor effect of DRB. However, the mutations in the exonic hnRNP A1-binding UAGG motif seemed to counter the DRB effect. In the combinatorial mutants of the exonic UAGGs and the intronic GGGG, the response to DRB is completely lost. Both RNA motifs are required for the effect of DRB-induced transcriptional stress on splicing, suggesting a crosstalk between these motifs or a possible complex formation among the splicing

factors that bind to these motifs. The DRB-induced transcriptional stress and the putative crosstalk or complex formation between UAGG and GGGG upon this stress is still unknown, and it remains to be identified.

In the case of HNRNPH1, I observed that the feedback regulation by hnRNP H is another way of splicing response to stress. It is possible to predict that one role of this splicing can be the dose-dependent adjustment of the hnRNP H protein level in the cell, affecting the splicing of hundreds of transcripts in a concerted fashion. Since the DRB-induced skipping of the HNRNPH1 exon 4 leads to nonsense-mediated RNA decay, this splicing response acts as a switch that decides whether there will be more hnRNP H protein made in the cell, or less upon stress. It was previously shown that the hnRNP H overexpression was involved in a switch to more aggressive and metastatic phenotype in gliomas (LeFave *et al.*, 2011), therefore the nonsense-mediated decay of the hnRNP H transcript levels and fine-tuning of its splicing upon stress is important to keep the cell proliferation rate and phenotype on track. In order to address whether the splicing shift is a result of nonsense-mediated decay, UPF1 knockdown can be performed before the DRB-induced transcriptional stress.

Our analysis of different mutants revealed that the G stretch acts as a silencer. Two different mutants at this location had a bigger response to DRB treatment, showing that the binding of the hnRNP H protein to wild-type GGGG motif causes increased skipping of the HNRNPH1 Exon 4, albeit smaller. However, the increase in E4 skipping was much higher when hnRNP H fails to bind to the mutant GGGG motif. That motif might have an indirect role in optimizing the hnRNP H abundance in the cell, or act as a stress sensor, depending on the stoichiometry of hnRNP H protein. In order to validate

the effect of the hnRNP H, it can be knocked down prior to DRB-induced transcriptional stress. Since hnRNP H is a silencer of HNRNPH3 exon 3, and since I know that GGGG mutants had a much bigger effect on splicing, I would speculate that it can increase the DRB-induced skipping. In order to test whether there are other RNA motifs that act via hnRNP H binding, hnRNP H knockdown or overexpression can be combined with the GGGG mutants, to check the effect of DRB-induced transcriptional stress on splicing.

The 5’ splice site GGGG was previously published to have a role in the strength and functionality of the 5’ splice site, and act as an enhancer (McCullough and Berget, 1997; McCullough and Berget, 2000). However, I observed that it acts as a silencer for GRIN1 CI exon. This can be due to the interference of U1 snRNP, as it binds to a very close vicinity of the GGGG motif. U6 snRNP or other spliceosome components can also contribute to the effect of DRB-induced transcriptional stress in splicing.

The UAGG and GGGG mutations in GRIN1 CI exon were shown to be important in elucidating the effect of hnRNP H on the fine-tuning of splicing upon transcriptional stress. Since I observed that hnRNP H is a silencer acting on its own transcript, hnRNP H-encoding transcripts can be used in future experiments for constructing a splicing reporter, and performing site-directed mutagenesis on the exonic UAGG and intronic GGGG of HNRNPH1 exon 4 and HNRNPH3 exon 3. These mutants can be combined with transcriptional stress and the stoichiometric changes of hnRNP H by knockdown or overexpression, to test the changes in splicing.

The levels of the HNRNPH1 and the HNRNPH3 transcripts can be regulated via nonsense-mediated decay in the absence of the intact GGGG motif, and the HNRNPH2 can compensate the decreases in the levels of HNRNPH1 and HNRNPH3. The hnRNP

H2 is not subject to alternative splicing upon stress, making it a constantly available supply to perform the splicing regulation function of the hnRNP H family. The alternative splicing of hnRNP H1 and H3 leads to the shifts in the basal levels of hnRNP H proteins available in cells, contributing to the overall proliferative phenotype by acting upon the splicing of proliferation-associated transcripts, such as Bcl-x.

Transcriptional pausing and pause-release are directly tied to the SR family of splicing factors. SRSF1 and SRSF2 were found to be associated with specific gene promoters and RNA sequences, specifically with hnRNP H1 gene and its transcript (Ji *et al.*, 2013). In the RNA maps that I constructed, the SRSF1 and SRSF2 binding sites were at the very close vicinity of the hnRNP H and hnRNP A1 binding sites, hinting at a possible interaction or competition between these splicing factors. SRSF1 and SRSF2, as well as several hnRNPs, were found to be part of the 7SK non-coding RNA complex, acting as a link between the P-TEFb and the RNA polymerase II. SRSF2 has a specific role in RNA polymerase II CTD Ser2 phosphorylation and in CDK9-7SK association (Ji *et al.*, 2013). The role of SRSF2 in the transcriptional pausing and pause-release, along with its roles in splicing, suggests that the DRB-induced effects that I observed on HNRNPH1-H3 and GRIN1 splicing can also be related to the roles of SR proteins on transcription-splicing coupling. Transcriptional activation of HNRNPH1 by SRSF2, and the autoregulation of HNRNPH3 splicing that I observed in this study, made me think that SR proteins and hnRNPs can act in an orchestrated fashion to control the level of gene expression and the accuracy of splicing as part of the 7SK complex in a promoter- and sequence-dependent, cooperative manner.

Transcriptional pausing and its effect on splicing are also related to the 7SK RNA through the association of the 7SK complex with the hnRNP A1. Knocking down hnRNP A1/A2 resulted in a similar effect as DRB-induced inhibition of transcription elongation, which caused an increase in the 7SK-CDK9 interaction, changing the elongation rate (Lemieux *et al.*, 2015). However, the role of the hnRNP H and the mechanism through which these hnRNPs act remains unsolved. Our triple mutant T8 that lacks two hnRNP A1- and one hnRNP H-binding sites, along with the double mutant E17 lacking two hnRNP A1 sites, had a truncated response to DRB. This result supported the previous data about the hnRNP A1/A2 knockdown mimicking DRB. However, our hnRNP A1 knockdown did not affect HNRNPH3 splicing, possibly due to the presence of hnRNP H or hnRNP A2. In the case of GRIN1 CI exon, our GGGG mutants had a bigger response to DRB. HnRNP seemed to counter the effect of DRB, which remains unanswered.

The changes in the GRIN1 splicing upon transcriptional stress and pause-release are particularly important, due to the roles of NMDA R1 receptor in learning, memory, neuronal and synaptic plasticity. Embryonic stem cells or other types of quickly-dividing cells go through different cellular stresses, in part due to differential access to nutrients or oxygen, random or induced mutations, or the transcriptional pausing as a way of the timing regulation of gene expression. The regulation of GRIN1 splicing upon pausing or pause-release confers a survival advantage by rapidly changing the synapses and the availability of the NMDA R1 receptors on the surface. Cellular stresses can contribute to learning and memory through the alternative splicing changes in GRIN1, and this novel hnRNP H-regulated stress response has a potential to have further roles in the brain.

This study answered some of the questions about the role of hnRNP H and the hnRNP H-binding GGGG motif in relation to the genotoxic and transcriptional stress as well as the pause-release, but it also has given rise to new questions. What is the role of hnRNP H in the recruitment coupling, or kinetic coupling? Does hnRNP H recognize a broader range of degenerate targets to influence the stress response? What is the effect of the weak-affinity interactions between hnRNP H and the UAGGG motifs that are known to be bound by hnRNP A1? What is the effect of the competition between hnRNP A1 and hnRNP H on the splicing response to stress? What is the result of the autoregulation of splicing by hnRNP H? Is there a genomewide regulation of splicing of the hnRNP H targets upon transcriptional stress and pause-release? Answering these questions will enable a broader understanding of how hnRNP H utilizes different stress inputs, turns them into a splicing readout and keeps the cellular homeostasis stable.

In the future, a computational approach that will predict the targets of hnRNP H genomewide, and a transcriptomics approach to enrich the hnRNP H target transcripts and check their splicing in a global way would be interesting. By this way, the network of interactions or regulations as a target of hnRNP H in the presence versus absence of a transcriptional stress will be generated, and the stress-responsive targets of hnRNP H will be identified so as to elucidate the splicing cascade involved in the stress response.

There are also unanswered questions about the functional output of hnRNP H-mediated fine-tuning of splicing. What would be the resulting phenotype upon hnRNP H overexpression in primary cells? How do these cells respond to stress? If the response is masked via the presence of an intact GGGG motif and hnRNP H, how does the

knockdown of hnRNP H impact the transcriptional stress response of cancer cells? The combination of hnRNP H knockdown and DRB treatment can be applied.

The direct binding of hnRNP H to GGGG motifs is known, but in this study, the binding of hnRNP H was not assayed with respect to DRB treatment. How is the affinity of hnRNP H binding to GGGG affected by DRB treatment in vitro? Does DRB affect the in vitro splicing of HNRNPH1 or HNRNPH3? What kind of a splicing pattern will their GGGG mutants have upon DRB treatment? DRB treatment of the cells with HNRNPH1 and HNRNPH3 transcripts harboring GGGG mutations

The pause-release and its reversible effect on GRIN1 splicing was important, but the limits of reversibility is not known. To what extent can the cells accommodate the transcriptional pausing and pause-release? Some transcripts can stay at transcriptional pausing for a very long time in embryonic stem cells. If we did the transcriptional stress and pause-release experiments with mouse embryonic stem cells, would we be able to change the fate of the cell? Can we force the differentiation of a pluripotent stem cell into a specific lineage by just applying DRB-induced transcriptional stress and controlled pause-release? Can we change the splicing patterns of iPS cells (induced pluripotent stem cells)? Is the splicing-transcription coupling an important aspect of development and cell fate determination? Is there a way of splicing adaptation to continuous stress in extreme environments, such as exposure to low-dose chemotherapeutic agents, UV, or ionizing radiation? Can cancer cells evolve upon exposure to continuous stresses within a tumor microenvironment, by rapidly changing the splicing behavior?

On the whole, this study elucidated the role of hnRNP H in the autoregulation of its own splicing. In addition, I also determined the role of 5’ splice site-flanking GGGG

motif in transcriptional pausing and pause-release. Despite the remaining unknowns on the role of hnRNP H in the broad stress-sensing and splicing response upon stress, the *cis* RNA motif and the role of hnRNP H in *trans* for the response to transcriptional pause and pause-release were identified. The outcomes of this study can open the door for new experiments to understand the role of hnRNP H in the stress response pathways and stress-induced splicing regulation more in depth.

APPENDIX A

FLAVOPIRIDOL HAS A SMALL EFFECT ON THE SPLICING OF HNRNPH1 EXON 4

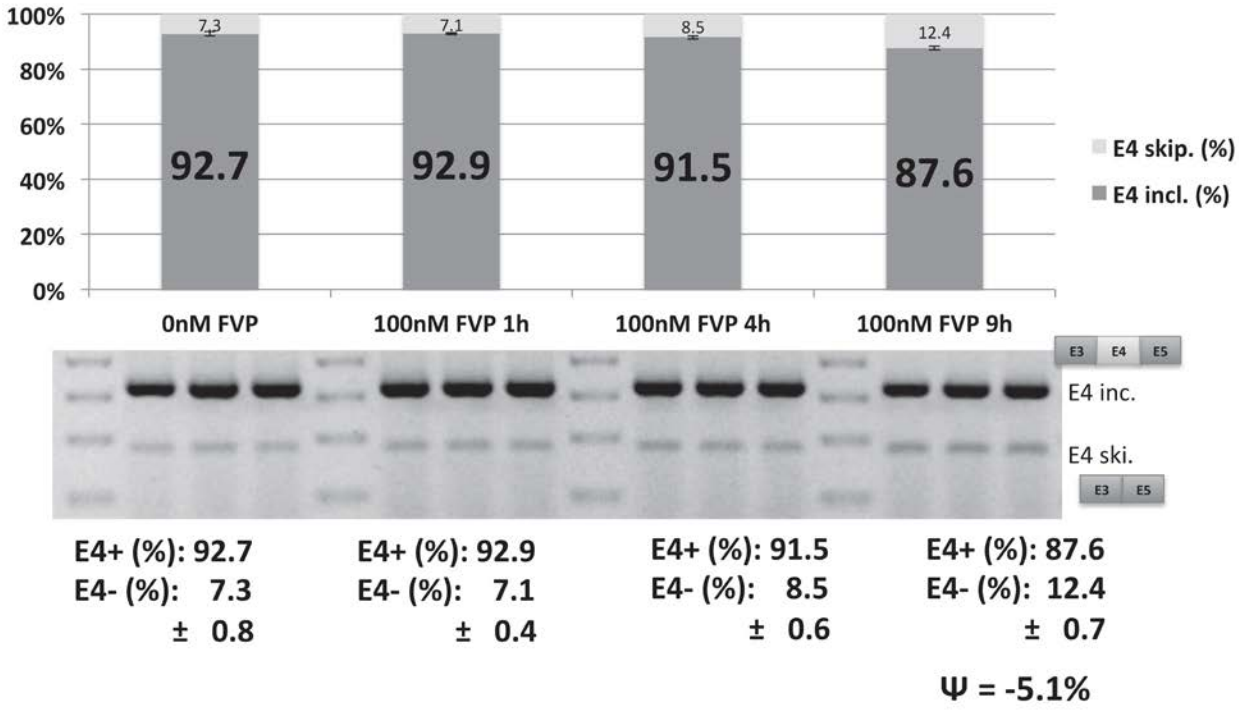


Figure 20: Flavopiridol has a small effect on the splicing of HNRNPH1 Exon 4. 100 nM FVP was added on the HeLa cells for the indicated time points, and the effect was very small. The dose and the duration of the treatment can be increased, or a very high dose can be applied for a short interval to see its effect on splicing.

APPENDIX B

PKA AND HNRNP A1 CAUSE A SMALL INCREASE IN UNTREATED HELA CELLS, WITH A SLIGHT ANTAGONIZING EFFECT ON DRB

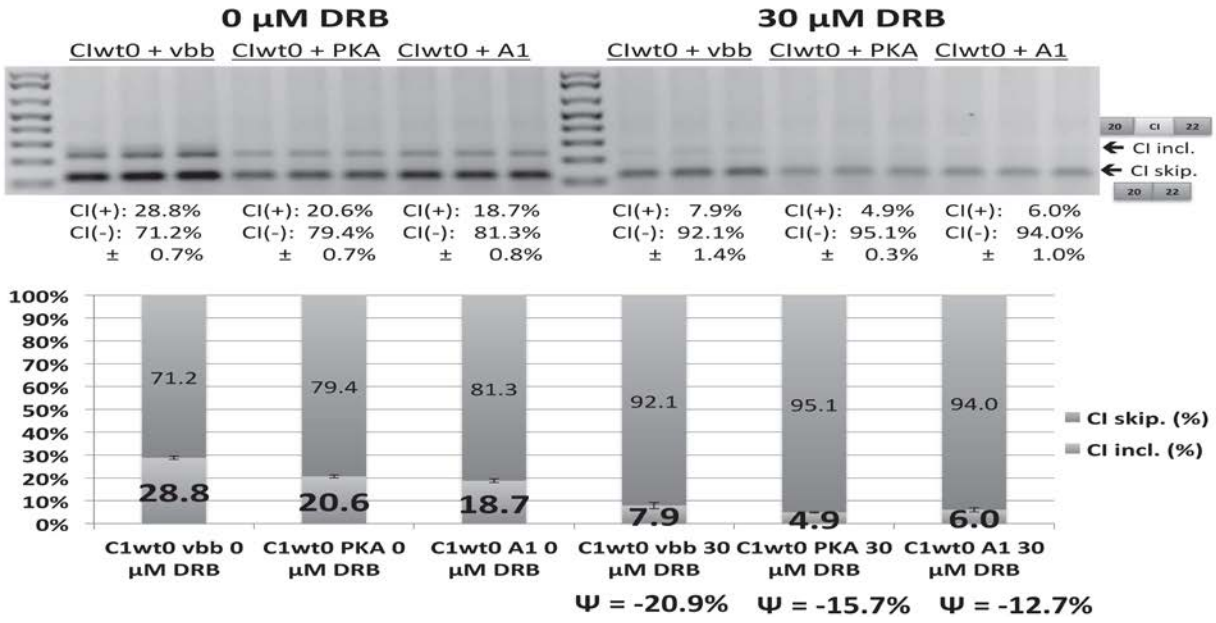


Figure 21: PKA and hnRNP A1 cause a small increase in untreated HeLa cells, with a slight antagonizing effect on DRB. The pcDNA3 vbb, pcDNA3.PKA and pcDNA3.hnRNP A1 plasmids were transfected to HeLa cells, and the cells were treated with 0 or 30 µM DRB. The PCR was performed with the primers for the GRIN1 CI amplification.

APPENDIX C

PKA AND HNRNP A1 CAUSE A SMALL INCREASE IN CI SKIPPING IN THE 5M2-TRANSFECTED HELA CELLS, WITH NO ANTAGONIZING EFFECT ON DRB-MEDIATED SILENCING

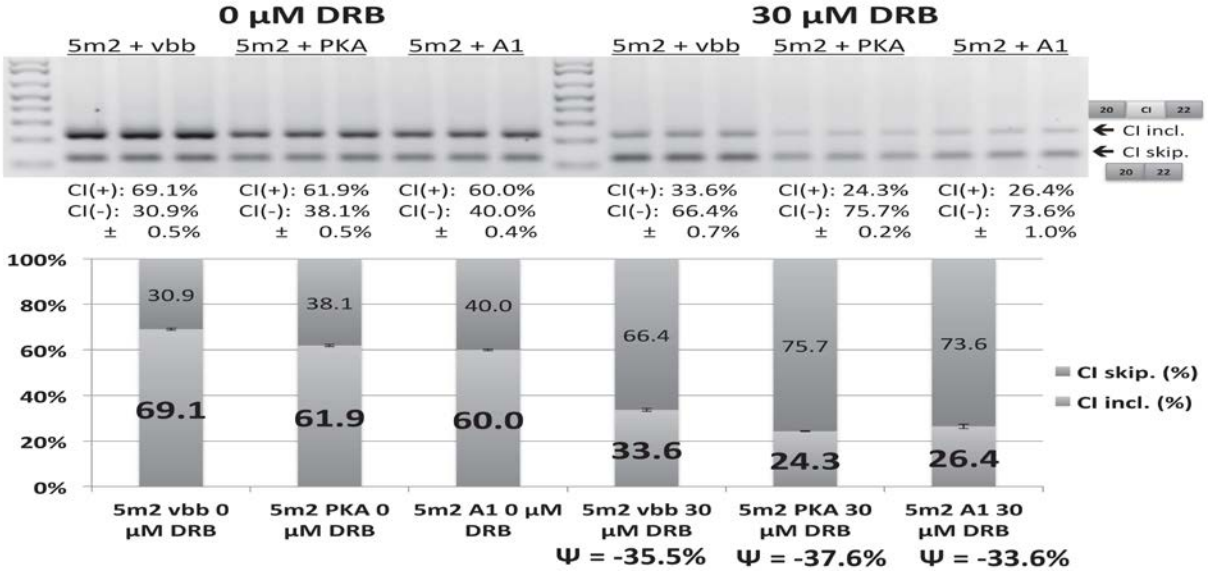


Figure 22: PKA and hnRNP A1 cause a small increase in CI skipping in the 5m2-transfected HeLa cells, with no antagonizing effect on DRB-mediated silencing. The pcDNA3 vbb, pcDNA3.PKA and pcDNA3.hnRNP A1 plasmids were transfected to HeLa cells, and the cells were treated with 0 or 30 µM DRB. The PCR was performed with the primers for the GRIN1 CI amplification.

APPENDIX D

PKA OVEREXPRESSION CAUSES A DELAY IN THE DRB PAUSE-RELEASE IN 5M2-TRANSFECTED HELA CELLS

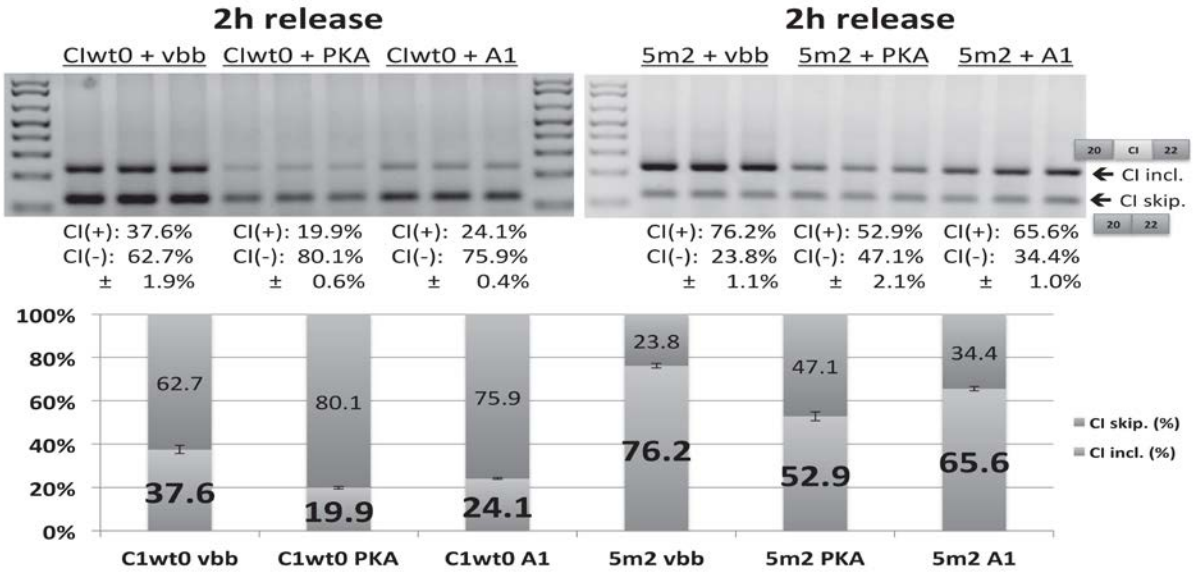


Figure 23: PKA overexpression causes a delay in the DRB pause-release in 5m2-transfected HeLa cells. The pcDNA3 vbb, pcDNA3.PKA and pcDNA3.hnRNP A1 plasmids were transfected to HeLa cells, and the cells were treated with 0 or 30 µM DRB. The PCR was performed with the primers for the GRIN1 CI amplification. After the 14h DRB treatment, the cell medium was replaced by DMEM+10% FBS lacking DRB.

BIBLIOGRAPHY

Adelman, K., & Lis, J. T. (2012). Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet, 13*(10), 720-731.

Albert, H., Battaglia, E., Monteiro, C., & Bagrel, D. (2012). Genotoxic stress modulates CDC25C phosphatase alternative splicing in human breast cancer cell lines. *Molecular Oncology, 6*(5), 542-552. doi:http://dx.doi.org/10.1016/j.molonc.2012.06.003

Alkan, Serkan A., Martincic, K., & Milcarek, C. (2006). The hnRNPs F and H2 bind to similar sequences to influence gene expression. *Biochemical Journal, 393*(1), 361-371. doi:10.1042/bj20050538

Allemand, E., Guil, S., Myers, M., Moscat, J., Cáceres, J. F., & Krainer, A. R. (2005). Regulation of heterogenous nuclear ribonucleoprotein A1 transport by phosphorylation in cells stressed by osmotic shock. *Proceedings of the National Academy of Sciences of the United States of America, 102*(10), 3605-3610. doi:10.1073/pnas.0409889102

An, P., & Grabowski, P. J. (2007). Exon Silencing by UAGG Motifs in Response to Neuronal Excitation. *PLoS Biol, 5*(2), e36. doi:10.1371/journal.pbio.0050036

Aoki, C., Venkatesan, C., Go, C., Mong, J., & Dawson, T. (1994). Cellular and subcellular localization of NMDA-R1 subunit immunoreactivity in the visual cortex of adult and neonatal rats. *The Journal of Neuroscience, 14*(9), 5202-5222.

Baranello, L., Wojtowicz, D., Cui, K., Devaiah, Ballachanda N., Chung, H.-J., Chan-Salis, Ka Y., . . . Levens, D. (2016). RNA Polymerase II Regulates Topoisomerase 1 Activity to Favor Efficient Transcription. *Cell, 165*(2), 357-371. doi:http://dx.doi.org/10.1016/j.cell.2016.02.036

Basu, A., & Krishnamurthy, S. (2010). Cellular Responses to Cisplatin-Induced DNA Damage. *Journal of Nucleic Acids, 2010*, 16. doi:10.4061/2010/201367

Bauman, J. A., & Kole, R. (2011). Modulation of RNA splicing as a potential treatment for cancer. *Bioengineered Bugs, 2*(3), 125-128. doi:10.4161/bbug.2.3.15165

Bent, H., Hanne, H. R., Henrik, V., Kurt, D., Xiangdang, L., Pavel, G., . . . Julio, E. C. (1995). Heterogeneous Nuclear Ribonucleoproteins H, H′, and F Are Members of a Ubiquitously Expressed Subfamily of Related but Distinct Proteins Encoded by Genes Mapping to Different Chromosomes. *Journal of Biological Chemistry, 270*(48), 28780-28789. doi:10.1074/jbc.270.48.28780

Biamonti, G., & Caceres, J. F. (2009). Cellular stress and RNA splicing. *Trends in Biochemical Sciences, 34*(3), 146-153. doi:http://dx.doi.org/10.1016/j.tibs.2008.11.004

Biamonti, G., & Vourc'h, C. (2010). Nuclear Stress Bodies. *Cold Spring Harbor Perspectives in Biology, 2*(6). doi:10.1101/cshperspect.a000695

Black, D. L., & Grabowski, P. J. (2003). Alternative Pre-mRNA Splicing and Neuronal Function. In P. Jeanteur (Ed.), *Regulation of Alternative Splicing* (pp. 187-216). Berlin, Heidelberg: Springer Berlin Heidelberg.

Boutz, P. L., Stoilov, P., Li, Q., Lin, C.-H., Chawla, G., Ostrow, K., . . . Black, D. L. (2007). A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons. *Genes & Development, 21*(13), 1636-1652. doi:10.1101/gad.1558107

Bozulic, L., Surucu, B., Hynx, D., & Hemmings, B. A. (2008). PKBα/Akt1 Acts Downstream of DNA-PK in the DNA Double-Strand Break Response and Promotes Survival. *Molecular Cell, 30*(2), 203-213. doi:http://dx.doi.org/10.1016/j.molcel.2008.02.024

Bradley, J., Carter, S. R., Rao, V. R., Wang, J., & Finkbeiner, S. (2006). Splice Variants of the NR1 Subunit Differentially Induce NMDA Receptor-Dependent Gene Expression. *The Journal of Neuroscience, 26*(4), 1065-1076. doi:10.1523/jneurosci.3347-05.2006

Busà, R., Geremia, R., & Sette, C. (2010). Genotoxic stress causes the accumulation of the splicing regulator Sam68 in nuclear foci of transcriptionally active chromatin. *Nucleic Acids Research, 38*(9), 3005-3018. doi:10.1093/nar/gkq004

Cairns, R. A., Harris, I. S., & Mak, T. W. (2011). Regulation of cancer cell metabolism. *Nat Rev Cancer, 11*(2), 85-95.

Caputi, M., & Zahler, A. M. (2001). Determination of the RNA Binding Specificity of the Heterogeneous Nuclear Ribonucleoprotein (hnRNP) H/H′/F/2H9 Family. *Journal of Biological Chemistry, 276*(47), 43850-43859. doi:10.1074/jbc.M102861200

Chen, M., & Manley, J. L. (2009). Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat Rev Mol Cell Biol, 10*(11), 741-754.

Choi, Y., Grabowski, P., Sharp, P., & Dreyfuss, G. (1986). Heterogeneous nuclear ribonucleoproteins: role in RNA splicing. *Science, 231*(4745), 1534-1539. doi:10.1126/science.3952495

Christofk, H. R., Vander Heiden, M. G., Harris, M. H., Ramanathan, A., Gerszten, R. E., Wei, R., . . . Cantley, L. C. (2008). The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. *Nature, 452*(7184), 230-233.doi:http://www.nature.com/nature/journal/v452/n7184/suppinfo/nature06734_S1.html

Comiskey, D. F., Jacob, A. G., Singh, R. K., Tapia-Santos, A. S., & Chandler, D. S. (2015). Splicing factor SRSF1 negatively regulates alternative splicing of MDM2 under damage. *Nucleic Acids Research*. doi:10.1093/nar/gkv223

Cooper, T. A., Wan, L., & Dreyfuss, G. (2009). RNA and Disease. *Cell, 136*(4), 777-793. doi:http://dx.doi.org/10.1016/j.cell.2009.02.011

Core, L. J., Waterfall, J. J., & Lis, J. T. (2008). Nascent RNA Sequencing Reveals Widespread Pausing and Divergent Initiation at Human Promoters. *Science, 322*(5909), 1845-1848. doi:10.1126/science.1162228

Cull-Candy, S., Brickley, S., & Farrant, M. (2001). NMDA receptor subunits: diversity, development and disease. *Current Opinion in Neurobiology, 11*(3), 327-335. doi:http://dx.doi.org/10.1016/S0959-4388(00)00215-4

Damsma, G. E., Alt, A., Brueckner, F., Carell, T., & Cramer, P. (2007). Mechanism of transcriptional stalling at cisplatin-damaged DNA. *Nat Struct Mol Biol, 14*(12), 1127-1133. doi:http://www.nature.com/nsmb/journal/v14/n12/suppinfo/nsmb1314_S1.html

Danko, Charles G., Hah, N., Luo, X., Martins, André L., Core, L., Lis, John T., . . . Kraus, W. L. (2013). Signaling Pathways Differentially Affect RNA Polymerase II Initiation, Pausing, and Elongation Rate in Cells. *Molecular Cell, 50*(2), 212-222. doi:http://dx.doi.org/10.1016/j.molcel.2013.02.015

Das, R., Yu, J., Zhang, Z., Gygi, M. P., Krainer, A. R., Gygi, S. P., & Reed, R. (2007). SR Proteins Function in Coupling RNAP II Transcription to Pre-mRNA Splicing. *Molecular Cell, 26*(6), 867-881. doi:http://dx.doi.org/10.1016/j.molcel.2007.05.036

David, C. J., Boyne, A. R., Millhouse, S. R., & Manley, J. L. (2011). The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65–Prp19 complex. *Genes & Development, 25*(9), 972-983. doi:10.1101/gad.2038011

David, C. J., & Manley, J. L. (2010). Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged. *Genes & Development, 24*(21), 2343-2364. doi:10.1101/gad.1973010

de la Mata, M., Alonso, C. R., Kadener, S., Fededa, J. P., Blaustein, M. a., Pelisch, F., . . . Kornblihtt, A. R. (2003). A Slow RNA Polymerase II Affects Alternative Splicing In Vivo. *Molecular Cell, 12*(2), 525-532. doi:http://dx.doi.org/10.1016/j.molcel.2003.08.001

Disset, A., Bourgeois, C. F., Benmalek, N., Claustres, M., Stevenin, J., & Tuffery-Giraud, S. (2006). An exon skipping-associated nonsense mutation in the dystrophin gene uncovers a complex interplay between multiple antagonistic splicing elements. *Human Molecular Genetics, 15*(6), 999-1013. doi:10.1093/hmg/ddl015

Dominguez, C., & Allain, F. H.-T. (2006). NMR structure of the three quasi RNA recognition motifs (qRRMs) of human hnRNP F and interaction studies with Bcl-x G-tract RNA: a novel mode of RNA recognition. *Nucleic Acids Research, 34*(13), 3634-3645. doi:10.1093/nar/gkl488

Dreyfuss, G., Matunis, M. J., Pinol-Roma, S., & Burd, C. G. (1993). hnRNP proteins and the biogenesis of mRNA. *Annual review of biochemistry, 62*(1), 289-321.

Dutertre, M., Sanchez, G., De Cian, M.-C., Barbier, J., Dardenne, E., Gratadou, L., . . . Auboeuf, D. (2010). Cotranscriptional exon skipping in the genotoxic stress response. *Nat Struct Mol Biol, 17*(11), 1358-1366. doi:http://www.nature.com/nsmb/journal/v17/n11/abs/nsmb.1912.html -supplementary-information

Ehlers, M., Tingley, W., & Huganir, R. (1995). Regulated subcellular distribution of the NR1 subunit of the NMDA receptor. *Science, 269*(5231), 1734-1737. doi:10.1126/science.7569904

Eperon, I. C., Makarova, O. V., Mayeda, A., Munroe, S. H., Cáceres, J. F., Hayward, D. G., & Krainer, A. R. (2000). Selection of Alternative 5′ Splice Sites: Role of U1 snRNP and Models for the Antagonistic Effects of SF2/ASF and hnRNP A1. *Molecular and Cellular Biology, 20*(22), 8303-8318. doi:10.1128/mcb.20.22.8303-8318.2000

Feng, Y., Chen, M., & Manley, J. L. (2008). Phosphorylation switches the general splicing repressor SRp38 to a sequence-specific activator. *Nat Struct Mol Biol, 15*(10), 1040-1048. doi:http://www.nature.com/nsmb/journal/v15/n10/suppinfo/nsmb.1485_S1.html

Flicek, P., Ahmed, I., Amode, M. R., Barrell, D., Beal, K., Brent, S., . . . Searle, S. M. J. (2012). Ensembl 2013. *Nucleic Acids Research*. doi:10.1093/nar/gks1236

Fong, N., Kim, H., Zhou, Y., Ji, X., Qiu, J., Saldi, T., . . . Bentley, D. L. (2014). Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. *Genes & Development, 28*(23), 2663-2676. doi:10.1101/gad.252106.114

Fu, X.-D., & Ares Jr, M. (2014). Context-dependent control of alternative splicing by RNA-binding proteins. *Nat Rev Genet, 15*(10), 689-701. doi:10.1038/nrg3778

Garneau, D., Revil, T., Fisette, J.-F., & Chabot, B. (2005). Heterogeneous Nuclear Ribonucleoprotein F/H Proteins Modulate the Alternative Splicing of the Apoptotic Mediator Bcl-x. *Journal of Biological Chemistry, 280*(24), 22641-22650. doi:10.1074/jbc.M501070200

Gautrey, H., Jackson, C., Dittrich, A.-L., Browell, D., Lennard, T., & Tyson-Capper, A. (2015). SRSF3 and hnRNP H1 regulate a splicing hotspot of HER2 in breast cancer cells. *RNA Biology, 12*(10), 1139-1151. doi:10.1080/15476286.2015.1076610

Gilchrist, D. A., Dos Santos, G., Fargo, D. C., Xie, B., Gao, Y., Li, L., & Adelman, K. (2010). Pausing of RNA Polymerase II Disrupts DNA-Specified Nucleosome Organization to Enable Precise Gene Regulation. *Cell, 143*(4), 540-551. doi:http://dx.doi.org/10.1016/j.cell.2010.10.004

Grabowski, P. J. (2004). A molecular code for splicing silencing: configurations of guanosine-rich motifs. *Biochemical Society Transactions, 32*(6), 924-927. doi:10.1042/bst0320924

Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R., & Young, R. A. (2007). A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells. *Cell, 130*(1), 77-88. doi:http://dx.doi.org/10.1016/j.cell.2007.05.042

Hai, Y., Cao, W., Liu, G., Hong, S.-P., Elela, S. A., Klinck, R., . . . Xie, J. (2008). A G-tract element in apoptotic agents-induced alternative splicing. *Nucleic Acids Research, 36*(10), 3320-3331. doi:10.1093/nar/gkn207

Han, K., Yeo, G., An, P., Burge, C. B., & Grabowski, P. J. (2005). A Combinatorial Code for Splicing Silencing: UAGG and GGGG Motifs. *PLoS Biol, 3*(5), e158. doi:10.1371/journal.pbio.0030158

Honoré, B., Baandrup, U., & Vorum, H. (2004). Heterogeneous nuclear ribonucleoproteins F and H/H′ show differential expression in normal and selected cancer tissues. *Experimental Cell Research, 294*(1), 199-209. doi:http://dx.doi.org/10.1016/j.yexcr.2003.11.011

Hostetter, A. A., Chapman, E. G., & DeRose, V. J. (2009). Rapid Cross-Linking of an RNA Internal Loop by the Anticancer Drug Cisplatin. *Journal of the American Chemical Society, 131*(26), 9250-9257. doi:10.1021/ja809637e

Hsin, J.-P., & Manley, J. L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes & Development, 26*(19), 2119-2137.doi:10.1101/gad.200303.112

Ip, J. Y., Schmidt, D., Pan, Q., Ramani, A. K., Fraser, A. G., Odom, D. T., & Blencowe, B. J. (2011). Global impact of RNA polymerase II elongation inhibition on alternative splicing regulation. *Genome Research, 21*(3), 390-401. doi:10.1101/gr.111070.110

Jeronimo, C., Collin, P., & Robert, F. (2016). The RNA Polymerase II CTD: The Increasing Complexity of a Low-Complexity Protein Domain. *Journal of Molecular Biology, 428*(12), 2607-2622. doi:http://dx.doi.org/10.1016/j.jmb.2016.02.006

Ji, X., Zhou, Y., Pandit, S., Huang, J., Li, H., Lin, Charles Y., . . . Fu, X.-D. (2013). SR Proteins Collaborate with 7SK and Promoter-Associated Nascent RNA to Release Paused Polymerase. *Cell, 153*(4), 855-868. doi:http://dx.doi.org/10.1016/j.cell.2013.04.028

Jung, Y., & Lippard, S. J. (2007). Direct Cellular Responses to Platinum-Induced DNA Damage. *Chemical Reviews, 107*(5), 1387-1407. doi:10.1021/cr068207j

Kwek, K. Y., Murphy, S., Furger, A., Thomas, B., O'Gorman, W., Kimura, H., . . . Akoulitchev, A. (2002). U1 snRNA associates with TFIIH and regulates transcriptional initiation. *Nat Struct Mol Biol, 9*(11), 800-805.

Ladd, A. N., & Cooper, T. A. (2002). Finding signals that regulate alternative splicing in the post-genomic era. *Genome Biol, 3*(11), 1-16.

Lambert, G., Estévez-Salmeron, L., Oh, S., Liao, D., Emerson, B. M., Tlsty, T. D., & Austin, R. H. (2011). An analogy between the evolution of drug resistance in bacterial communities and malignant tissues. *Nat Rev Cancer, 11*(5), 375-382.

Lambot, L., Chaves Rodriguez, E., Houtteman, D., Li, Y., Schiffmann, S. N., Gall, D., & de Kerchove d'Exaerde, A. (2016). Striatopallidal Neuron NMDA Receptors Control Synaptic Connectivity, Locomotor, and Goal-Directed Behaviors. *The Journal of Neuroscience, 36*(18), 4976-4992. doi:10.1523/jneurosci.2717-15.2016

LeFave, C. V., Squatrito, M., Vorlova, S., Rocco, G. L., Brennan, C. W., Holland, E. C., . . . Cartegni, L. (2011). Splicing factor hnRNPH drives an oncogenic splicing switch in gliomas. *The EMBO Journal, 30*(19), 4084-4097. doi:10.1038/emboj.2011.259

Lefebvre, S., Bürglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., . . . Melki, J. (1995). Identification and characterization of a spinal muscular atrophy-determining gene. *Cell, 80*(1), 155-165. doi:http://dx.doi.org/10.1016/0092-8674(95)90460-3

Lemieux, B., Blanchette, M., Monette, A., Mouland, A. J., Wellinger, R. J., & Chabot, B. (2015). A Function for the hnRNP A1/A2 Proteins in Transcription Elongation. *PLoS ONE, 10*(5), e0126654. doi:10.1371/journal.pone.0126654

Liang, K., Gao, X., Gilmore, J. M., Florens, L., Washburn, M. P., Smith, E., & Shilatifard, A. (2015). Characterization of Human Cyclin-Dependent Kinase 12 (CDK12) and CDK13 Complexes in C-Terminal Domain Phosphorylation, Gene Transcription, and RNA Processing. *Molecular and Cellular Biology, 35*(6), 928-938. doi:10.1128/mcb.01426-14

Lin, S., Coutinho-Mansfield, G., Wang, D., Pandit, S., & Fu, X.-D. (2008). The splicing factor SC35 has an active role in transcriptional elongation. *Nat Struct Mol Biol, 15*(8), 819-826. doi:http://www.nature.com/nsmb/journal/v15/n8/suppinfo/nsmb.1461_S1.html

Lipscombe, D. (2005). Neuronal proteins custom designed by alternative splicing. *Current Opinion in Neurobiology, 15*(3), 358-363. doi:http://dx.doi.org/10.1016/j.conb.2005.04.002

Liu, L., Xu, Y., He, M., Zhang, M., Cui, F., Lu, L., . . . Esteban, Miguel A. (2014). Transcriptional Pause Release Is a Rate-Limiting Step for Somatic Cell Reprogramming. *Cell Stem Cell, 15*(5), 574-588. doi:http://dx.doi.org/10.1016/j.stem.2014.09.018

Luco, R. F., Pan, Q., Tominaga, K., Blencowe, B. J., Pereira-Smith, O. M., & Misteli, T. (2010). Regulation of Alternative Splicing by Histone Modifications. *Science, 327*(5968), 996-1000. doi:10.1126/science.1184208

Luo, J., Bosy, T. Z., Wang, Y., Yasuda, R. P., & Wolfe, B. B. (1996). Ontogeny of NMDA R1 subunit protein expression in five regions of rat brain. *Developmental Brain Research, 92*(1), 10-17. doi:http://dx.doi.org/10.1016/0165-3806(95)00191-3

Matlin, A. J., Clark, F., & Smith, C. W. J. (2005). Understanding alternative splicing: towards a cellular code. *Nat Rev Mol Cell Biol, 6*(5), 386-398.

Matsuoka, S., Ballif, B. A., Smogorzewska, A., McDonald, E. R., Hurov, K. E., Luo, J., . . . Elledge, S. J. (2007). ATM and ATR Substrate Analysis Reveals Extensive Protein Networks Responsive to DNA Damage. *Science, 316*(5828), 1160-1166. doi:10.1126/science.1140321

McNamara, Ryan P., Reeder, Jonathan E., McMillan, Elizabeth A., Bacon, Curtis W., McCann, Jennifer L., & D'Orso, I. (2016). KAP1 Recruitment of the 7SK snRNP Complex to Promoters Enables Transcription Elongation by RNA Polymerase II. *Molecular Cell, 61*(1), 39-53. doi:http://dx.doi.org/10.1016/j.molcel.2015.11.004

Merdzhanova, G., Edmond, V., De Seranno, S., Van den Broeck, A., Corcos, L., Brambilla, C., . . . Eymin, B. (2008). E2F1 controls alternative splicing pattern of genes involved in apoptosis through upregulation of the splicing factor SC35. *Cell Death Differ, 15*(12), 1815-1823. doi:http://www.nature.com/cdd/journal/v15/n12/suppinfo/cdd2008135s1.html

Montes, M., Cloutier, A., Sánchez-Hernández, N., Michelle, L., Lemieux, B., Blanchette, M., . . . Suñé, C. (2012). TCERG1 Regulates Alternative Splicing of the Bcl-x Gene by Modulating the Rate of RNA Polymerase II Transcription. *Molecular and Cellular Biology, 32*(4), 751-762. doi:10.1128/mcb.06255-11

Muñoz, M. J., Santangelo, M. S. P., Paronetto, M. P., de la Mata, M., Pelisch, F., Boireau, S., . . . Kornblihtt, A. R. (2009). DNA Damage Regulates Alternative Splicing through Inhibition of RNA Polymerase II Elongation. *Cell, 137*(4), 708-720. doi:http://dx.doi.org/10.1016/j.cell.2009.03.010

Muse, G. W., Gilchrist, D. A., Nechaev, S., Shah, R., Parker, J. S., Grissom, S. F., . . . Adelman, K. (2007). RNA polymerase is poised for activation across the genome. *Nat Genet, 39*(12), 1507-1511. doi:http://www.nature.com/ng/journal/v39/n12/suppinfo/ng.2007.21_S1.html

Nakanishi, N., Axel, R., & Shneider, N. A. (1992). Alternative splicing generates functionally distinct N-methyl-D-aspartate receptors. *Proceedings of the National Academy of Sciences, 89*(18), 8552-8556. doi:10.1073/pnas.89.18.8552

Narita, T., Yamaguchi, Y., Yano, K., Sugimoto, S., Chanarat, S., Wada, T., . . . Handa, H. (2003). Human Transcription Elongation Factor NELF: Identification of Novel Subunits and Reconstitution of the Functionally Active Complex. *Molecular and Cellular Biology, 23*(6), 1863-1873. doi:10.1128/mcb.23.6.1863-1873.2003

Nasim, F. H., Spears, P. A., Hoffmann, H. M., Kuo, H. C., & Grabowski, P. J. (1990). A Sequential splicing mechanism promotes selection of an optimal exon by repositioning a downstream 5' splice site in preprotachykinin pre-mRNA. *Genes & Development, 4*(7), 1172-1184. doi:10.1101/gad.4.7.1172

Ni, Z., Saunders, A., Fuda, N. J., Yao, J., Suarez, J.-R., Webb, W. W., & Lis, J. T. (2008). P-TEFb Is Critical for the Maturation of RNA Polymerase II into Productive Elongation In Vivo. *Molecular and Cellular Biology, 28*(3), 1161-1170. doi:10.1128/mcb.01859-07

Nilson, Kyle A., Guo, J., Turek, Michael E., Brogie, John E., Delaney, E., Luse, Donal S., & Price, David H. (2015). THZ1 Reveals Roles for Cdk7 in Co-transcriptional Capping and Pausing. *Molecular Cell, 59*(4), 576-587. doi:http://dx.doi.org/10.1016/j.molcel.2015.06.032

Ninomiya, K., Kataoka, N., & Hagiwara, M. (2011). Stress-responsive maturation of Clk1/4 pre-mRNAs promotes phosphorylation of SR splicing factor. *The Journal of Cell Biology, 195*(1), 27-40. doi:10.1083/jcb.201107093

Nogués, G., Kadener, S., Cramer, P., Bentley, D., & Kornblihtt, A. R. (2002). Transcriptional Activators Differ in Their Abilities to Control Alternative Splicing. *Journal of Biological Chemistry, 277*(45), 43110-43114. doi:10.1074/jbc.M208418200

Okabe, S., Miwa, A., & Okado, H. (1999). Alternative Splicing of the C-Terminal Domain Regulates Cell Surface Expression of the NMDA Receptor NR1 Subunit. *The Journal of Neuroscience, 19*(18), 7781-7792.

Pagani, F., Buratti, E., Stuani, C., Bendix, R., Dork, T., & Baralle, F. E. (2002). A new type of mutation causes a splicing defect in ATM. *Nat Genet, 30*(4), 426-429.

Peterlin, B. M., & Price, D. H. (2006). Controlling the Elongation Phase of Transcription with P-TEFb. *Molecular Cell, 23*(3), 297-305. doi:http://dx.doi.org/10.1016/j.molcel.2006.06.014

Phatnani, H. P., & Greenleaf, A. L. (2006). Phosphorylation and functions of the RNA polymerase II CTD. *Genes & Development, 20*(21), 2922-2936. doi:10.1101/gad.1477006

Plet, A., Eick, D., & Blanchard, J. M. (1995). Elongation and premature termination of transcripts initiated from c-fos and c-myc promoters show dissimilar patterns. *Oncogene, 10*(2), 319-328.

Rahl, P. B., Lin, C. Y., Seila, A. C., Flynn, R. A., McCuine, S., Burge, C. B., . . . Young, R. A. (2010). c-Myc Regulates Transcriptional Pause Release. *Cell, 141*(3), 432-445. doi:http://dx.doi.org/10.1016/j.cell.2010.03.030

Rahman, M. A., Azuma, Y., Nasrin, F., Takeda, J.-i., Nazim, M., Ahsan, K. B., . . . Ohno, K. (2015). SRSF1 and hnRNP H antagonistically regulate splicing of COLQ exon 16 in a congenital myasthenic syndrome. *Scientific Reports, 5*, 13208. doi:10.1038/srep13208 http://www.nature.com/articles/srep13208 - supplementary-information

Rasmussen, E. B., & Lis, J. T. (1993). In vivo transcriptional pausing and cap formation on three Drosophila heat shock genes. *Proceedings of the National Academy of Sciences, 90*(17), 7923-7927.

Rauch, J., O'Neill, E., Mack, B., Matthias, C., Munz, M., Kolch, W., & Gires, O. (2010). Heterogeneous Nuclear Ribonucleoprotein H Blocks MST2-Mediated Apoptosis in Cancer Cells by Regulating a-raf Transcription. *Cancer Research, 70*(4), 1679-1688. doi:10.1158/0008-5472.can-09-2740

Samatanga, B., Dominguez, C., Jelesarov, I., & Allain, F. H.-T. (2012). The high kinetic stability of a G-quadruplex limits hnRNP F qRRM3 binding to G-tract RNA. *Nucleic Acids Research*. doi:10.1093/nar/gks1289

Sánchez-Hernández, N., Boireau, S., Schmidt, U., Muñoz-Cobo, J. P., Hernández-Munain, C., Bertrand, E., & Suñé, C. (2016). The in vivo dynamics of TCERG1, a factor that couples transcriptional elongation with splicing. *RNA, 22*(4), 571-582. doi:10.1261/rna.052795.115

Sazani, P., & Kole, R. Therapeutic potential of antisense oligonucleotides as modulators of alternative splicing. *The Journal of Clinical Investigation, 112*(4), 481-486. doi:10.1172/JCI19547

Scotti, M. M., & Swanson, M. S. (2016). RNA mis-splicing in disease. *Nat Rev Genet, 17*(1), 19-32. doi:10.1038/nrg.2015.3

Shalgi, R., Hurt, Jessica A., Lindquist, S., & Burge, Christopher B. (2014). Widespread Inhibition of Posttranscriptional Splicing Shapes the Cellular Transcriptome following Heat Shock. *Cell Reports, 7*(5), 1362-1370. doi:http://dx.doi.org/10.1016/j.celrep.2014.04.044

Sheth, N., Roca, X., Hastings, M. L., Roeder, T., Krainer, A. R., & Sachidanandam, R. (2006). Comprehensive splice-site analysis using comparative genomics. *Nucleic Acids Research, 34*(14), 3955-3967. doi:10.1093/nar/gkl556

Shomron, N., Alberstein, M., Reznik, M., & Ast, G. (2005). Stress alters the subcellular distribution of hSlu7 and thus modulates alternative splicing. *Journal of Cell Science, 118*(6), 1151-1159. doi:10.1242/jcs.01720

Sohail, M., Cao, W., Mahmood, N., Myschyshyn, M., Hong, S. P., & Xie, J. (2014). Evolutionarily emerged G tracts between the polypyrimidine tract and 3′ AG are splicing silencers enriched in genes involved in cancer. *BMC Genomics, 15*(1), 1-11. doi:10.1186/1471-2164-15-1143

Solier, S., Barb, J., Zeeberg, B. R., Varma, S., Ryan, M. C., Kohn, K. W., . . . Pommier, Y. (2010). Genome-wide Analysis of Novel Splice Variants Induced by Topoisomerase I Poisoning Shows Preferential Occurrence in Genes Encoding Splicing Factors. *Cancer Research, 70*(20), 8055-8065. doi:10.1158/0008-5472.can-10-2491

Stamm, S. (2002). Signals and their transduction pathways regulating alternative splicing: a new dimension of the human genome. *Human Molecular Genetics, 11*(20), 2409-2416. doi:10.1093/hmg/11.20.2409

Stamm, S. (2008). Regulation of Alternative Splicing by Reversible Protein Phosphorylation. *Journal of Biological Chemistry, 283*(3), 1223-1227. doi:10.1074/jbc.R700034200

Stein, C. (2015). *The role of hnRNP H in the splicing response to genotoxic stress.* University of Pittsburgh.

Stoimenov, I., Schultz, N., Gottipati, P., & Helleday, T. (2011). Transcription Inhibition by DRB Potentiates Recombinational Repair of UV Lesions in Mammalian Cells. *PLoS ONE, 6*(5), e19492. doi:10.1371/journal.pone.0019492

Thomas, M. P., & Lieberman, J. (2013). Live or let die: posttranscriptional gene regulation in cell stress and cell death. *Immunological Reviews, 253*(1), 237-252. doi:10.1111/imr.12052

Traynelis, S. F., Wollmuth, L. P., McBain, C. J., Menniti, F. S., Vance, K. M., Ogden, K. K., . . . Dingledine, R. (2010). Glutamate Receptor Ion Channels: Structure, Regulation, and Function. *Pharmacological Reviews, 62*(3), 405-496. doi:10.1124/pr.109.002451

Venables, J. P. (2004). Aberrant and Alternative Splicing in Cancer. *Cancer Research, 64*(21), 7647-7654. doi:10.1158/0008-5472.can-04-1910

Viniegra, J. G., Martínez, N., Modirassari, P., Losa, J. H., Cobo, C. P., Lobo, V. J. S.-A., . . . Rojas, J. M. (2004). Full activation of PKB/Akt in response to insulin or ionizing radiation is mediated through ATM. *Journal of Biological Chemistry*.

Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., . . . Burge, C. B. (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature, 456*(7221), 470-476. doi:http://www.nature.com/nature/journal/v456/n7221/suppinfo/nature07509_S1.html

Wang, L., Duke, L., Zhang, P. S., Arlinghaus, R. B., Symmans, W. F., Sahin, A., . . . Dai, J. L. (2003). Alternative Splicing Disrupts a Nuclear Localization Signal in Spleen Tyrosine Kinase That Is Required for Invasion Suppression in Breast Cancer. *Cancer Research, 63*(15), 4724-4730.

Wang , L., Lawrence , M. S., Wan , Y., Stojanov , P., Sougnez , C., Stevenson , K., . . . Wu , C. J. (2011). SF3B1 and Other Novel Cancer Genes in Chronic Lymphocytic Leukemia. *New England Journal of Medicine, 365*(26), 2497-2506. doi:doi:10.1056/NEJMoa1109016

Wang, Z., & Burge, C. B. (2008). Splicing regulation: From a parts list of regulatory elements to an integrated splicing code. *RNA, 14*(5), 802-813. doi:10.1261/rna.876308

Ward, A. J., & Cooper, T. A. (2010). The pathobiology of splicing. *The Journal of Pathology, 220*(2), 152-163. doi:10.1002/path.2649

Wu, W., Yan, C., Gan, T., Chen, Z., Lu, X., Duerksen-Hughes, P. J., . . . Yang, J. (2010). Nuclear proteome analysis of cisplatin-treated HeLa cells. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 691*(1–2), 1-8. doi:http://dx.doi.org/10.1016/j.mrfmmm.2010.06.002

Xie, J., & Black, D. L. (2001). A CaMK IV responsive RNA element mediates depolarization-induced alternative splicing of ion channels. *Nature, 410*(6831), 936-939. doi:http://www.nature.com/nature/journal/v410/n6831/suppinfo/410936a0_S1.html

Xie, J., Pabón, D., Jayo, A., Butta, N., & González-Manchón, C. (2005). Type I Glanzmann thrombasthenia caused by an apparently silent β 3 mutation that results in aberrant splicing and reduced Bbeta; 3 mRNA. *Thrombosis and Haemostasis, 93*(5), 897-903. doi:10.1160/TH04-09-0633

Xu, Q., & Lee, C. (2003). Discovery of novel splice forms and functional analysis of cancer- specific alternative splicing in human expressed sequences. *Nucleic Acids Research, 31*(19), 5635-5643. doi:10.1093/nar/gkg786

Yoshida, T., Makino, Y., & Tamura, T.-a. (1999). Association of the rat heterogeneous nuclear RNA-ribonucleoprotein F with TATA-binding protein1. *FEBS Letters, 457*(2), 251-254. doi:http://dx.doi.org/10.1016/S0014-5793(99)01048-0

Zhang, G., Sun, L., Lu, X., Chen, Z., Duerksen-Hughes, P. J., Hu, H., . . . Yang, J. (2012). Cisplatin treatment leads to changes in nuclear protein and microRNA expression. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis, 746*(1), 66-77. doi:http://dx.doi.org/10.1016/j.mrgentox.2012.03.004

Zhang, W., Liu, H., Han, K., & Grabowski, P. J. (2002). Region-specific alternative splicing in the nervous system: implications for regulation by the RNA-binding protein NAPOR. *RNA, 8*(5), 671-685.

Zhang, Z., Lotti, F., Dittmar, K., Younis, I., Wan, L., Kasim, M., & Dreyfuss, G. (2008). SMN Deficiency Causes Tissue-Specific Perturbations in the Repertoire of snRNAs and Widespread Defects in Splicing. *Cell, 133*(4), 585-600. doi:http://dx.doi.org/10.1016/j.cell.2008.03.031

Zhang-Hooks, Y., Agarwal, A., Mishina, M., & Bergles, Dwight E. (2016). NMDA Receptors Enhance Spontaneous Activity and Promote Neuronal Survival in the Developing Cochlea. *Neuron, 89*(2), 337-350. doi:http://dx.doi.org/10.1016/j.neuron.2015.12.016

Zhou, Z., Qiu, J., Liu, W., Zhou, Y., Plocinik, Ryan M., Li, H., . . . Fu, X.-D. (2012). The Akt-SRPK-SR Axis Constitutes a Major Pathway in Transducing EGF Signaling to Regulate Alternative Splicing in the Nucleus. *Molecular Cell, 47*(3), 422-433. doi:http://dx.doi.org/10.1016/j.molcel.2012.05.014

Zukin, R. S., & Bennett, M. V. L. (1995). Alternatively spliced isoforms of the NMDARI receptor subunit. *Trends in Neurosciences, 18*(7), 306-313. doi:http://dx.doi.org/10.1016/0166-2236(95)93920-S