ROLE OF DCPS IN MAMMALIAN RNA REGULATION AND HUMAN DISEASES

By

MI ZHOU

A dissertation submitted to the

Graduate School-New Brunswick

and

The Graduate School of Biomedical Sciences

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Cell and Development Biology

Written under the direction of

Dr. Megerditch Kiledjian

And approved by

_____

_____

_____

_____

New Brunswick, New Jersey

October, 2015

ABSTRACT OF THE DISSERTATION

Role of DcpS in Mammalian RNA Regulation and Human Diseases

By MI ZHOU

Dissertation Director

Dr. Megerditch Kiledjian

In eukaryotic cells, mRNA degradation plays an important role in the control of gene expression and is therefore highly regulated. The scavenger decapping enzyme DcpS is a multifunctional protein that plays a critical role in mRNA degradation.

We first sought to identify DcpS target genes in mammalian cells using a cell permeable DcpS inhibitor compound, RG3039, which was initially developed for therapeutic treatment of Spinal Muscular Atrophy (SMA). Microarray analysis following DcpS decapping inhibition by RG3039 revealed the steady state levels of 222 RNAs were altered. Of a subset selected for validation by qRT-PCR, two non-coding transcripts dependent on DcpS decapping activity, were identified and referred to as DcpS Responsive Noncoding Transcript (DRNT) 1 and 2 respectively. Only the increase in DRNT1 transcript was accompanied with an increase of its RNA stability and this increase was dependent on both DcpS and Xrn1. Our data indicate that DcpS is a transcript-restricted modulator of RNA stability in mammalian cells and the RG3039

quinazoline compound is pleotropic, influence gene expression in both an apparent DcpS dependent and independent manner.

A surprising development was uncovered in a collaborative study where two distinct mutations in the DcpS gene (c.636+1G>A, $DcpS^{Ins15}$ and c.947C>T, $DcpS^{T316M}$) were identified as the underlying cause of autosomal recessive intellectual disability within a consanguineous family. Both of the mutations were confirmed to disrupt DcpS decapping activity in vitro and/or in vivo, indicating that the decapping activity of DcpS is critical for normal neurological development. Consistent with a role for DcpS in neuronal cells, our studies with the $DcpS^{Ins15}$ variant uncovered a link between this variant DcpS and Spinal Muscular Atrophy (SMA). Exogenous expression of $DcpS^{Ins15}$ in SMA patient fibroblast cells increased SMN2 mRNA and corresponding SMN protein levels. Our findings suggest that strategies to shift wild type DcpS splicing patterns to partially yield the variant DcpS ^{Ins15} splicing pattern may be beneficial for SMA therapeutics.

ACKNOWLEGEMENT

I would like to thank my advisor, Dr. Mike Kiledjian for all of his guidance, support and advice during the past four years. I appreciate all his contributions of time, ideas, and funding to make my PhD experience productive and stimulating. I am so grateful for his guidance on both academic and life, including independent thinking ability, the writing and presentation skills, as well as personal and communication skills. His encouragement and unwavering support has sustained me through frustration and depression. Without his pushing me ahead, the completion of my PhD study would be impossible.

I would like to thank the members of my thesis committee, Dr. Lori Covey, Dr. Sam Gunderson, Dr. Paul Copeland for their advice and suggestions during my PhD study as well as sharing facilities and reagents and providing indispensable help to my research work.

I would like to thank our collaborators Dr John B. Vincent from University of Toronto (Canada) and Dr Rami Abou Jamra from Friedrich-Alexander University (Germany). They discovered the family with Intellectual Disability patients from remote Pakistan area and identified the critical mutations in DcpS by genomic sequencing analysis. Their genius work was the foundation for my project of DcpS in Intellectual Disability and SMA.

I would like to thank the current and past members of the lab, Xinfu Jiao, Ewa Grudzien, Xiaobin Luo, Huijuan Cui, Mangen Song, and Madel Durens, for their support, help and friendship in the past four years. I would especially like to thank Xinfu for his

technical support and generous contribution of time in helping me with research experiments throughout my PhD study.

Last and most importantly, I would like to thank my family for the support they provided me through my entire life. In particular, I must acknowledge my husband and also a former lab member, You Li, for his support, tolerance, patience and love in work and life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

General mRNA Degradation

Two opposing processes, nuclear transcription and cytoplasmic mRNA degradation, invokes the intricate patterns of gene expression in eukaryotic cells. In the nucleus, the nascent transcript is synthesized and chaperoned by cap-recognition factors that facilitate its processing and then subjected to quality control measures prior to its export. In the cytoplasm, the mRNA template is translated into protein but eventually meets its demise at the hands of the decay machinery.

Therefore, mRNA degradation plays an important role in the control of gene expression and is highly regulated. Following an initial deadenylation step to remove the poly(A) tail, the remaining mRNA can undergo exonucleolytic decay from either 3′ or the 5′ end. In the 5′ to 3′ decay pathway, the 5′ cap is firstly removed by Dcp2 or Nudt16 decapping enzymes, resulting in a 5′ end monophosphorylated RNA (Lykke-Andersen, 2002; Song et al, 2010; van Dijk et al, 2002; Wang et al, 2002). The uncapped product undergoes 5′ to 3′ exoribonuclease decay by Xrn1 (Hsu & Stevens, 1993). In the 3′ to 5′ pathway, the multi-subunit exosome complex degrades the deadenylated RNA from the 3′ end, leaving the residual m^7GpppN cap structure (Anderson & Parker, 1998; Liu et al, 2006). The resulting cap structure is a substrate of the scavenger decapping enzyme, DcpS to release m^7Gp and ppN products (Wang 2001, Liu, 2002). Factors involved in the two major mRNA degradation pathways are evolutionarily conserved from yeast to mammals.

5′ Decapping enzymes

Dcp2 is highly conserved in eukaryotes and is the best characterized 5′ decapping enzyme. It is a member of the Nudix family of proteins with a central Nudix domain, which catalyzes the decapping step. Dcp2 can hydrolyze both monomethyl (m^7G) and trimethyl (m2, 2, 7G) capped RNA with poor activity on unmethylated capped (G-cap) RNA (Cohen et al, 2005; Piccirillo et al, 2003; Steiger et al, 2003; Wang et al, 2002). Structural and biochemical analyses revealed that Dcp2 is an RNA binding protein which directly interacts with both the cap and RNA body to recognize its substrate (Deshmukh et al, 2008). The C-terminus of the NUDIX domain forms a conserved RNA binding channel that contributes to the substrate specificity of Dcp2 decapping and it requires a capped RNA substrate that is longer than ~25 nucleotides (Deshmukh et al, 2008; Li et al, 2009; Steiger et al, 2003). Dcp2 preferentially binds a 5′ terminal stem loop structure termed Dcp2 binding and decapping element (DBDE), which promotes recruitment of Dcp2 and subsequent decapping (Li et al, 2008). The DBDE consists of at least an 8 basepair long stem with an intervening loop positioned within the 5′terminal 10 nucleotides of an mRNA for optimal Dcp2-mediated decapping but is not restricted to a specific primary sequence.

Dcp2 is highly expressed in the mouse embryonic brain, heart, liver and kidney, while it can only be detected in the corresponding adult brain but not the latter three tissues (Song et al, 2010). This non-ubiquitous expression pattern indicates Dcp2 is tissue-specific and developmentally regulated. Genome-wide profiling of Dcp2 responsive mRNAs further revealed the selective function of Dcp2 in transcript decapping. Only about 200 mRNAs were elevated upon reduction of Dcp2, among which

a subset of mRNAs was involved in innate immunity indicating Dcp2 plays a role in host-defense of viral challenge (Li et al, 2012).

Similar to Dcp2, Nudt16 is also a member of the Nudix family of hydrolases and was initially identified as a 29 kD nuclear protein in Xenopus (X29) that selectively bound and decapped the U8 snoRNA in vitro (Ghosh et al, 2004). Nudt16 was also shown to be conserved in metazoans and function in nuclear snoRNA decapping (Taylor & Peculis, 2008). More recently, human Nudt16 was shown to be cytoplasmic and possess mRNA decapping activity on a subset of mRNAs in human cells including the Angiomotin-like 2 mRNA (Song et al, 2010).

The transcript specificity of Dcp2 and Nudt16 indicated that additional mRNA decapping enzymes are present in mammalian cells. Mammalian genomes contain 22 Nudix family proteins, two of which are Dcp2 and Nudt16. It was recently shown that mouse protein Nudt2, Nudt3, Nudt12, Nudt15, Nudt17, and Nudt19 had different degrees of decapping activity in vitro on both monomethylated and unmethylated capped RNAs, among which Nudt17 and Nudt19 were similar to Dcp2 and predominantly generated m^7GDP (Song et al, 2013). The other four Nudt proteins , Nudt2, Nudt3, Nudt12, and Nudt15, however generated both m^7GMP and m^7GDP from monomethylated capped RNAs (Song et al, 2013) Since these enzymes have thus far not been reported to modulate the stability of endogenous mRNAs, whether these Nudt proteins are true mRNA decapping enzyme in cells remains to be demonstrated.

Xrn1 exonuclease

The 5′-3′ exonuclease Xrn1 is a ~175 kDa protein which is evolutionarily conserved, with orthologs identified in all eukaryotes investigated from yeast to

mammalian cells. Its highly conserved N-terminal domain confers the exonuclease activity, which has a nuclear counterpart termed Xrn2 (or Rat1 in yeast) exonuclease and contributes to snoRNA and rRNA processing (Nagarajan et al, 2013). Xrn1 is proposed to degrade the mRNA body after decapping. Consistent with this, Xrn1 preferentially degrades 5′-monophosphorylated RNA in vitro, while capped or 5′-triphosphate RNAs are resistant to Xrn1 activity (Stevens & Poole, 1995). In yeast, lesions in the Xrn1 gene lead to accumulation of uncapped mRNA intermediates, highlighting its function in the control of mRNA degradation (Hsu & Stevens, 1993). Recent studies have also shown that XRN1 is important for degradation of a novel class of long noncoding RNAs (lncRNAs) in yeast. These are called XRN1-sensitive Unstable-Transcripts (XUTs) and are often antisense to open reading frames (van Dijk et al, 2011). XRN1 is enriched in cytoplasmic foci, together with other proteins required for 5′-3′ mRNA decay including decapping enzyme Dcp1/2 and deadenylase CCR4(Sheth & Parker, 2003). These foci are called mRNA-processing bodies (P-bodies).

Besides the bulk mRNA decay pathway, Xrn1 is also involved in mRNA surveillance. It has been extensively studied in the NMD pathways. In yeast, Xrn1 is shown to be recruited to polyribosomes to degrade PTC-containing mRNA after decapping (Hu et al, 2010). In Drosophila, the NMD pathway is initiated with endonucleolytic cleavage of the nonsense-containing mRNA by SMG6 (Huntzinger et al, 2008). The resulting 3′ cleavage product is degraded by XRN1. This mechanism is also conserved in human cells (Eberle et al, 2009). Similar to the NMD pathway, Xrn1 has also been demonstrated in RNA silencing pathway. 3′ decay intermediates generated by

RNA-induced silencing complex (RISC) accumulate in Xrn1-depleted cells (Orban & Izaurralde, 2005b)

Exosome exonuclease complex

The exosome complex is a multi-subunit protein complex capable of degrading various types of RNAs. It is found in both eukaryotic cells and archaea. Substrates of the exosome include messenger RNA, ribosomal RNA, and many species of small RNAs. This macromolecular complex has a central inactive core structure arranged in a ring consisting of six subunits to which other accessory components such as exonuclease subunit Rrp44 or Rrp6 can assemble on and form the active complex (Liu et al, 2006; Schmid & Jensen, 2008; Shen & Kiledjian, 2006). All core subunits are highly conserved [illegible] al (Allmang et al, 1999). In eukaryotic cells, the exosome complex is present in the cytoplasm, nucleus and the nucleolus. It interacts with different protein partners in different compartments to regulate specific RNA degradation in these cell compartments. The exosome has 3′-5′ exoribonucleolytic activity, and in eukaryotes also an endoribonucleolytic activity (Lebreton et al, 2008; Liu et al, 2006).

In the nucleus, exosome complex functions in the 3' processing of 5.8S rRNA precursors (the first identified function of the exosome), as well as many snoRNAs. It is also responsible for degradation of aberrant RNA precursors including pre-mRNAs, pre-tRNAs and pre-rRNAs (Houseley et al, 2006). In the cytoplasm, exosome complex is involved in the turn-over of mRNA molecules. It is the major player of the 3′-5′ degradation pathway of bulk mRNA decay and therefore plays a critical role in maintaining cellular hemostasis. The exosome is also important for various mRNA

surveillance pathways. It has been reported to be involved in the non-sense mediated decay (NMD) pathway (Mitchell & Tollervey, 2003; Takahashi et al, 2003), non-stop decay pathway which degrades mRNAs that lack a termination codon (van Hoof et al, 2002), and no-go decay pathway which targets mRNAs with stalled ribosomes (Doma & Parker, 2006). The 5′ fragment of mRNA cleaved by RISC complex is also degraded by the exosome complex. (Orban & Izaurralde, 2005a).

The Scavenger Decapping Enzyme DcpS

Scavenger decapping enzyme DcpS is a ~40 kDa protein and a member of the Histidine Triad (HIT) family of nucleotide hydrolases (Liu et al, 2002), which are characterized by the conserved HIT motif (Seraphin, 1992) contained in a larger HIT fold region. DcpS functions on cap structure or capped oligonucleotides shorter than 10 bases rather than capped RNA, which is different from Dcp2 (Liu et al, 2002). It hydrolyzes the triphosphate link of cap structure at the alpha phosphate position to generate m^7GMP and a nucleotide diphosphate (Liu et al, 2002). Mutations of conserved histidines in the HIT motif abolish decapping activity confirming the significance of this motif in catalysis (Liu et al, 2002). The N7-methyl moiety is essential for substrate specificity of DcpS, as it does not function on unmethylated cap (Liu et al, 2002).

The DcpS decapping enzyme was shown to co-purify with exosomal components in a 300 kDa complex (Wang & Kiledjian, 2001), implying the existence of a coupled 3′ to 5′ degradation and decapping pathway. Consistently, DcpS activity in cell extract is dependent on 3′ degradation of mRNA (Wang & Kiledjian, 2001). The proposed model is

that DcpS cleaves the cap structures resulting from exosome degradation of mRNAs. Although the direct interaction between DcpS and exosomal proteins has not been identified to date, the association of DcpS with the exosome suggests potential coordination of function between the exosome and DcpS in the regulation of mRNA decay.

In addition, Dcs1p, the homologue of DcpS in yeast, can influence 5' to 3' exoribonucleolytic activity and leads to the accumulation of stable uncapped mRNA in yeast strains disrupted for the DCS1 gene (Liu & Kiledjian, 2005). Dcs1p was subsequently identified as an obligate cofactor for the 5' to 3' exoribonuclease Xrn1 by an unknown function independent of its decapping catalytic activity (Sinturel et al, 2012). Moreover, the *C.elegans* Dcs-1 was shown to physically interact with Xrn-1 and promote specific microRNA degradation, also independent of its decapping activity (Bosse et al, 2013).

Despite the role of DcpS in cytoplasmic events, it is primarily a nuclear protein in S. pombe and human cells but localizes throughout the cytoplasm in *C.elegans* (Lall et al, 2005; Salehi et al, 2002; Shen et al, 2008). DcpSis a Crm1-dependent nucleocytoplasmic shuttling protein in mammalian cells (Shen et al, 2008). As a modulator of cap structure in cells, DcpS can potentially influence the function of other cap binding proteins including the nuclear cap-binding protein complex Cbp20 and Cbp80 or the cytoplasmic eIF4E cap-binding protein (Bail & Kiledjian, 2008). Reduction of DcpS protein levels results in accumulation of cap structure, which in turn sequesters the cap binding complex (CBC) and leads to a corresponding decrease in proximal intron splicing (Shen

et al, 2008). DcpS therefore appears to function at multiple levels in the regulation of gene expression.

Insights into the molecular mechanism of DcpS decapping were provided by structural analysis. The crystal structure of DcpS reveals a homodimer complex where each monomer possesses a distinct binding and activity site (Gu et al, 2004). The homodimer displays an asymmetric structure in the ligand-bound form and a dynamic state, where each monomer can alternate from an open to closed state. Hydrolysis of cap structure only occurs in the closed conformation (Gu et al, 2004). This mechanism is supported by kinetic studies that demonstrate the significance of dynamic conformational changes of the N-terminal domain for cap hydrolysis and confirm the mutually exclusive hydrolysis function between the two catalytic active sites (Liu et al, 2008b).

pre-mRNA Splicing and Regulation

Pre-mRNA splicing is one of the most fundamental processes in eukaryotic gene expression that generates mature mRNA for translation. Pre-mRNA splicing is carried out by the major spliceosome complex, which consists of five small nuclear ribonucleoprotein complexes (snRNPs), together with other factors. In eukaryotes, the major spliceosome is composed of the U1, U2, U4, U5, and U6 snRNPs involved in the vast majority of splicing. A small subset of introns are spliced by a minor spliceosome composed of U11, U12, U4atac, U5 and U6atac snRNPs (Black, 2003). Splicing takes place in a sequence of well-organized steps. In the first step, U1 snRNP binds to the 5' splice site which contains a critical GU dinucleotide to form the so-called early E complex (Ruby & Abelson, 1988), facilitated with two non-snRNP factors, nuclear cap-

binding complex (CBC) and ASF/SF2 (Harper & Manley, 1991; Izaurralde et al, 1994). In the second step, U2AF, a non-snRNP protein binds to the 3' splice site which contains polypyrimidine tract and a conserved terminal AG site at the end of intron (Ruskin et al, 1988). Next, the U2 snRNP base pairs with the branch A site which is an adenosine approximately 18 -35 nucleotides upstream of the 3' splice site. (Taggart et al, 2012; Zhuang et al, 1989). Finally, the recruitment of U4/U6/U5 tri-snRNP complex completes the B complex assembly that undergoes a conformational shift (Behrens & Luhrmann, 1991; Fortner et al, 1994). Upon this switch, the complex becomes the catalytically active C complex by bringing the 5' splice site and branch site adenosine within proximity (Chiara et al, 1996). Following the active complex formation, the branch site adenosine forms a new 2'-5' phosphodiester bond with the proximal 5' splice site using its 2' hydroxyl group to generate a lariat containing the 3' exon and free 5' exon. Then the 5' exon forms a new 3'-5' phosphodiester bond with the 3' exon using its 3' hydroxyl group and the intronic portion of the lariat is displaced (Kramer, 1996).

Although the basic steps of pre-mRNA splicing are well characterized, the fact that all exons are not constitutive and can be alternatively spliced adds complexity to the process. Alternative splicing joins a different combination of exons together that forms distinct protein-coding mRNAs from a single pre-mRNA. This combinatorial approach vastly expanded the information relative to the number of genes encoded. Hence, how a cell decides which exons to utilize requires a very delicate and precise regulation.

The regulation of splicing involves both *cis*-acting elements and *trans*-acting factors. *Cis* elements are specific sequences on the pre-mRNA and are categorized into groups as exon splicing enhancers (ESE), exon splicing silencers (ESS), intron splicing

enhancers (ISE) and intron splicing silencers (ISS) (Matlin et al, 2005). *Trans* factors are often RNA-binding proteins that bind to the cis element on pre-mRNA and influence the alternative splicing. Generally, the *trans* factors bind to enhancers to activate adjacent splice sites or deactivating adjacent silencer elements, while the factors bind to silencers to repress splice sites and enhancers (Matlin et al, 2005). For example, a large family of SR proteins usually bind to splicing enhancers to facilitate the recruitment of spliceosomal snRNPs (Graveley, 2000). It has been shown that SR proteins directly interact with U1 snRNP to recognize the 5′ splice site as well as the U2AF complex and U2 snRNP to recognize the 3′ splice site (Blencowe et al, 1999; Hastings & Krainer, 2001). On the other hand, members of the heterogeneous nuclear ribonucleoprotein (hnRNP) family are often binding to splicing silencers to inhibit splicing. For example, the polypyrimidine tract-binding protein (PTB) from hnRNP family represses the SM (smooth muscle) exon in the alpha-actinin gene, by binding to key sites in the polypyrimidine tract. This leads to co-operative occupation to additional downstream sites to displace an activator CELF protein which also binds to the same region but activate the SM splicing (Spellman et al, 2005). PTB also regulates self-splicing by repressing Exon 11 of PTB pre-mRNA splicing in an autoregulatory feedback loop (Spellman et al, 2005). Collectively, the final decision of splicing is generally accomplished through the combinatorial or competitive effects of both activating and inhibitory mechanisms.

It is estimated that a large fraction of human heritable diseases involve mutations which disrupt splicing (Cartegni et al, 2002). Among them, around 10% of the heritable disorders are caused by direct mutations of splice site sequences and about 25% of the

diseases are caused by mutations on splicing regulatory sequences outside of splicing sites (Cooper et al, 2009; Sterne-Weiler et al, 2011). As a result, further investigation in splicing regulation emphasized its importance in the detection and treatment of human diseases.

Translation

Cap-dependent Translation Initiation

Translation initiation is a crucial step in protein synthesis and a principal regulatory point in the control of translation. Translation initiation is rate limiting and involves assembly of elongation-competent 80S ribosomes on mRNA (Jackson et al, 2010). In eukaryotic cells, cap-dependent ribosomal recruitment involves the assembly of the eIF4F translation initiation complex to the 5′ m^7G cap structure of mRNA, and recruitment of the small ribosomal subunit-contained 43S complex to form the 48S complex (Gallie, 2002; Gingras et al, 1999). The eIF4F complex is composed of the cap-binding protein eIF4E, the scaffold protein eIF4G and the ATP-dependent RNA helicase eIF4A (Gallie, 1998). eIF4E binding to the 5′ cap is stabilized by an interaction with eIF4G and is further stabilized when cytoplasmic poly(A) binding protein (PABPC1) binds to eIF4G (Haghighat & Sonenberg, 1997; Imataka et al, 1998). eIF4G mediates the recruitment of the 43S complex to mRNA by an interaction with its component, eIF3 (Korneeva et al, 2000). The eIF4A helicase can promote binding of the 43S complex to mRNA by disrupting mRNA secondary structures (Seal et al, 1983). It is also required for scanning of the small subunit of the ribosome for the start-codon of mRNAs that contain

structured 5’ UTRs (Pestova et al, 2001). On the other hand, the eIF4E-binding protein (4EBP) family consists of three members, each of which can bind eIF4E at the interface recognized by eIF4G (Haghighat et al, 1995; Pause et al, 1994; Poulin et al, 1998). By doing so, 4EBP prevents eIF4F assembly and permits the secondary structure of the 5’ UTR to impede the scanning ribosome. 4EBP also responds to upstream signal molecules such as insulin, growth factors, and amino acids and operates downstream of the mTOR pathway (Choi et al, 2003). Moreover, 4EBP is a phosphoprotein wherein phosphorylation disrupts its interaction with eIF4E. Therefore translation initiation is sensitive to the complex assembly that recognizes its 5’ cap.

The 43S preinitiation complex comprised of 40S ribosomal subunit, eIF3 factor, methionine-initiator tRNA, and GTP contained eIF2 G protein (Ranu & London, 1979; Thach et al, 1966). This 43S ribosomal complex interacts with eIF4F complex at the 5′ end of mRNA and 5′ to 3′ scans in a linear manner until it encounters a start codon. In eukaryote, the Kozak sequence, $GCCGCCA^{-3}CCA^{+1}UGG^{+4}$, which contains the AUG start codon, plays a major role in translation initiation (Kozak, 1986). Base pairing between the AUG start codon and the methionine-initiator tRNA anticodon elicits hydrolysis of the GTP by eIF2 with the GTPase activating protein eIF5, releasing the eIF2-GDP complex from the initiation complex. The GTP-bound eIF5B factor facilitates 60S subunit joining the 48S initiation complex, displacing eIF3,eIF5 and eIF1 factors. The GTP hydrolysis through eIF5B G protein further releases eIF1A and GDP contained eIF5B, forming the 80S ribosome (Jackson et al, 2010).

In cap-dependent translation, the 5′ UTR of cellular mRNAs plays an important role in translation initiation and can be categorized into two groups: mRNAs that have

relatively short, unstructured 5´UTRs and mRNAs that have lengthy, highly structured 5´UTRs (De Benedetti & Harris, 1999; Koromilas et al, 1992). Translation initiation of mRNAs with long, structured 5' UTRs could provide a great challenge for the recruitment and maintenance of the initiation complex and ribosomal scanning compared to unstructured 5´ ends. mRNAs that contain unstructured 5' UTR usually have high translation efficiency and are less regulated at the initiation step compared to mRNAs with structured 5' ends (Pickering & Willis, 2005). Therefore structured 5' UTR-containing mRNAs are generally more sensitive to the regulation of cap-dependent translation initiation and the recruitment of the eIF4F complex and its associated helicase activity. As such, the eIF4F complex assembly plays a pivotal role in selectively controlling the translation of mRNAs with structured 5´ends.

Cap-independent translation initiation by IRES

Translation can also initiate internally within an mRNA independent of cap structure. Internal ribosomal entry site (IRES) elements are present in both viral and eukaryotic mRNAs, leading to the discovery of the cap independent translation process. An IRES element can be located either at the 5´ end or in the middle of an RNA. It usually forms a multi-pseudoknot structure that initiates translation independent of some or all of the eIF4F initiation factors (Lopez-Lastra et al, 2005; Merrick, 2004; Svitkin et al, 2005).

Viral IRES elements have been categorized into four groups based on the extent of cellular factors necessary for their function (Fraser & Doudna, 2007; Ho et al, 2000; Kieft, 2008). Group I IRES's are least dependent on initiation factors and include the Cricket paralysis virus, Plautia stali intestine virus, and Taura syndrome virus (Bushell &

Sarnow, 2002). For example, cricket paralysis virus-like viruse (CrPV) IRES can directly recruit the cellular 40S ribosome in the absence of any canonical initiation factors and operate independently of the methionine-initiator tRNA (Pestova & Hellen, 2003). Group II viruses usually operate independent of the scanning mechanism and consist of classical swine fever virus and hepatitis C virus, and porcine teschovirus 1 (Fraser and Doudna, 2007). For example, hepatitis C virus (HCV) IRES, require only eIF3, eIF2, and initiator Met-tRNA to assemble 80S ribosomes (Pestova et al, 1998). Group III viral IRES's use nearly most of the cellular initiation factors such as eIF4A, eIF4G, and IRES trans-activating factors (ITAFs) in addition to those of Group II and are found in encephalomyocarditis virus, foot-and-mouth-disease virus, and theiler's murine encephalomyelitis virus (Kolupaeva et al, 1998). The Group IV IRES's are most reliant on the cellular machineries and requires a supplemental source of extract along with ITAFs to recruit the ribosome and include poliovirus and rhinovirus (Kieft, 2008). Viral gene expression systems have evolved various means to compete against, bypass, or adopt the host's machinery at the translation level. For example, the EMCV IRES (poliovirus), requires all of the canonical initiation factors except for the cap binding protein eIF4E to recruit the ribosomal 43S for pre-initiation (Pestova et al, 1996).

Spinal Muscular Atrophy

SMN complex

Survival motor neuron (SMN) protein is a ubiquitously expressed protein localized in both the cytoplasm and nucleus of all cells (Liu & Dreyfuss, 1996). It is

concentrated in distinct nuclear bodies such as Cajal bodies which are involved in the assembly and modification of RNPs (Cioce & Lamond, 2005), and Gems which are nuclear structures containing high SMN concentration without snRNPs (Liu & Dreyfuss, 1996). Gems and Cajal bodies are shown to associate in the nucleus and the Cajal body marker coilin regulates its interaction with SMN and formation of Gems (Hebert et al, 2001).

SMN protein self-oligomerizes and interacts with a number of other core proteins (Gemin2-8) to from a macromolecular complex named SMN complex (Pellizzoni, 2007). The best-characterized SMN complex function is in snRNP biogenesis. In the cytoplasm, the SMN complex recruits seven Sm proteins and facilitates their assembly into a heptameric ring onto newly exported pre-snRNAs and ensuring the efficiency and specificity of the Sm core on to the correct snRNA targets (Golembe et al, 2005; Pellizzoni et al, 2002). After Sm core assembly, the SMN complex remains and potentially interacts with the trimethyl-guanosine synthase TGS1 which recognizes the m7G cap on the snRNA and transfers two additional methyl groups to form the m3G cap (Massenet et al, 2002; Mouaikel et al, 2003). The SMN complex also binds importin-beta and may facilitate the nuclear import of snRNPs (Massenet et al, 2002; Narayanan et al, 2002). However, additional roles for the SMN complex in snRNP biogenesis are remaining further investigation.

SMA

Spinal Muscular Atrophy (SMA) is a common autosomal recessive disorder that results in progressive loss of spinal anterior horn motor neurons. SMA is caused by reduced levels of SMN protein, the consequence of loss or mutation in the survival motor

neuron 1 gene (SMN1) (Lefebvre et al, 1995). Although SMN deficiency is reported to correlate with a decrease in the levels of spliceosomal snRNPs, these changes are predominantly tissue specific rather than uniform in all cells (Gabanella et al, 2007). Importantly, low levels of SMN protein disproportionally affect the minor splicing pathway that utilizes the U11, U12 snRNAs, rather than the classical U1, U2 snRNAs (Boulisfane et al, 2010). Recently, a subset of U12 intron-containing genes were identified to be SMN-dependent, where levels of mature mRNA from these genes were reduced upon reduction of SMN protein (Lotti et al, 2012).

A second, nearly identical gene to SMN1, termed SMN2, was found in the human genome. It shares the very similar promoter and differs by only two-nucleotides in the coding region relative to SMN1 (Monani et al, 1999a; Monani et al, 1999b). However, only the homozygous loss of SMN1, and not SMN2, results in SMA (Lefebvre et al, 1995). Therefore, all SMA patients still retain the SMN2 gene (Lefebvre et al, 1995). In contrast to the processing of the SMN1 pre-mRNA into functional mRNA, the majority of transcripts produced by the SMN2 locus lack exon 7 due to inappropriate splicing due to a single nucleotide substitution of a C→T within exon 7 (Lorson et al, 1999; Monani et al, 1999a). The absence of exon 7 leads to a defective, truncated, and unstable protein that is rapidly degraded (Burnett et al, 2009). The approximate 5 – 10 % of full length SMN transcript generated from the SMN2 gene is sufficient to sustain life through development and generally one year after birth (Lorson et al, 1999; Patrizi et al, 1999). In transgenic mice, severity of SMA was reported to be modulated by variable copies of the SMN2 gene, with increasing amelioration of the SMA phenotype with increasing copies of the SMN2 transgene (Monani et al, 2000). Eight copies of the human SMN2 transgene

were sufficient to cure SMA in mice (Monani et al, 2000). These findings suggest increasing expression of the endogenous SMN2 gene would be beneficial to SMA patients.

C5-quinazoline compounds in SMA therapeutic application

At present, no cure for SMA is available. One possible therapeutic approach is based on attempts at increasing the amount of SMN protein produced by the SMN2 genes through promoter activation, reduction of exon7 alternative splicing, or both (Cherry et al, 2012; Darras & Kang, 2007; Hastings et al, 2009; Hofmann & Wirth, 2002). Over the years, a number of SMN-inducing compounds have been identified using cultured fibroblasts derived from SMA patients, such as butyrate and valproic acid (Gilbert et al, 2001; Sumner et al, 2003). However, many of these compounds required nonphysiologically high concentrations (micromolar and millimolar) to increase SMN expression. Moreover, these compounds have extremely short half-lives *in vivo*, or have toxic side effects that make them inappropriate as therapeutic agents (Andreassi et al, 2001; Gilbert et al, 2001). At the meantime, a newly discovered compound that shifted the balance of SMN2 splicing toward the production of full-length SMN2 messenger RNA with high selectivity was reported to improve motor function and longevity in SMA Δ7 mice (Naryshkin et al, 2014). This orally available compound demonstrates a new therapeutic potential for SMA.

A recent study described C5-quinazoline as a promising compound that increased human SMN2 promoter driven expression of the bacterial β-lactamase reporter in a screen of a mouse motor neuron hybrid cell line NSC-34 (Jarecki et al, 2005). The initially identified compound was optimized by a focused medicinal chemistry effort to

develop a series of modified C5-quinazolines. One such compound, a piperidine 2,4-diaminoquinazoline known as D156844, was capable of inducing SMN2 promoter activity at low concentration (EC50=40nM) (Thurmond et al, 2008). D156844 was also reported to increase full-length/Δ7 SMN transcripts and protein levels in SMA patient fibroblast cultures, as well as the number of SMN-containing nuclear structures, termed Gems. However, D156844 treatment in SMA mouse models revealed a non-statistically significant increase of full-length SMN transcript or SMN protein in SMNΔ7 SMA mice ($SMN2^{+/+}$; $SMN\Delta7^{+/+}$; $mSmn^{-/-}$) (Butchbach et al, 2010). Nevertheless, a statistically significant increase in the lifespan of these SMA model animals was detected (Butchbach et al, 2010). In an effort to understand how D156844 may function, a proteomic screen was carried out. A protein array was screened with ^{125}I-lableled D156844, which resulted in the identification of DcpS as the cellular protein that it binds to and inhibits its decapping activity (Singh et al, 2008).

Further optimization generated the RG3039 derivative from D156844, which is also a potent DcpS decapping inhibitor (Gogliotti et al, 2013). Although the quinazoline compounds were capable of improving motor neuron function and extending survival of SMA model mice (Butchbach et al, 2010; Gogliotti et al, 2013; Van Meerbeke et al, 2013a), statistically significant increases in SMN2 mRNA or SMN protein were not evident. The lack of detectable increase in SMN mRNA and protein in treated animals confounds the mechanism by which RG3039 promotes survival of SMA mice.

Intellectual Disability

Neurological disorders represent a wide range of illnesses associated with varying clinical manifestations such as severity, age of onset, prognosis and treatment responses. Most neurological disorders lead to ill-health but rarely result in direct deaths. Disease pathogenesis of neurological disorders is generally characterized by a progressive decline in health function and thus an overall decrease in the quality of life for affected individuals. This is corroborated by high healthcare costs and long-term dependence on health and social care professionals throughout life. Most neurological disorders are induced predominantly by specific genetic factors that differ among varying classes of neurological diseases, some of which are caused by a single gene or locus, while others are caused by the interplay between genetic and environmental factors. As a result, genomic variation plays an important role in disease development and treatment option in neurological diseases.

Intellectual Disability (ID), or mental retardation, is a neurodevelopment disorder, which is characterized by impairment in conceptual, practical, social skills and an intelligence quotient (IQ) below 70 before the age of eighteen. The mild form of Intellectual Disability which is characterized by intelligence quotient (50-70), could be solely of genetic origin. However, severe forms of Intellectual Disability (IQ<50) with incidence of (0.4%), are thought to be caused by environmental or genetic factors (Ropers, 2010). Intellectual Disability is prevalent in 1-3% of the general population (Leonard & Wen, 2002; Roeleveld et al, 1997; Ropers, 2006).

Intellectual Disability is highly heterogeneous in nature and one of the important unsolved healthcare problems that impacts both affected individuals and their families.

For most Intellectual Disability cases, the underlying etiology is genetic in origin. It is reported that chromosomal aberration account for 15% of Intellectual Disability cases, while X-chromosomal mutations are the cause of approximately 10% of Intellectual Disability cases. In the remaining more than half of the cases, the etiology of Intellectual Disability remain unsolved and are likely caused by autosomal chromosomes mutations (Ropers & Hamel, 2005).

The most common Intellectual Disability caused by chromosomal aberration is Down syndrome (trisomy 21st), which can be easily detected by light microscopy. It is estimated that 1/750 to 1/800 babies are born with Down syndrome; this frequency is remarkably consistent in most parts of the world despite prenatal diagnosis (Besser et al, 2007; Collins et al, 2008).

X-chromosomal defects are the leading cause of intellectual disability in males and are referred to as X-linked Intellectual Disability. Fragile X-Syndrome is the most common case, which is diagnosed with the characteristic expanded trinucleotide repeat CGG within the first exon of the FMR1 gene (Kremer et al, 1991; Oberle et al, 1991; Yu et al, 1991). The expression of FMR1 is crucial for neuronal cells during fetal developmental (Devys et al, 1993; Pieretti et al, 1991). This expansion of repeat CGG causes hypermethylation of both the repeat and the promoter regions, leading to loss of FMR1 gene expression and subsequent aberrant neuronal development (Hansen et al, 1992; Pieretti et al, 1991; Vincent et al, 1991).

The frequency of autosomal dominant forms of Intellectual Disability is rare because most of the affected individuals do not reproduce. On the other hand, the molecular explanation of autosomal recessive intellectual disability is lagging because

most of the investigations are restricted by small family sizes and infrequent parental consanguinity. In this situation, even the possibility of a genetic causation will often not be considered (Najmabadi et al, 2007). As a result, very limited information is currently available concerning the role of autosomal recessive intellectual disability. It is postulated that the rate of autosomal recessive disorders is significantly higher in countries where consanguineous marriages are more frequent due to an increase in the probability of homozygosity of autosomal recessive mutations that may impact intellectual disability.

Materials and Methods

Plasmid constructs

The pcDNA3-Flag-DcpS^{H277N} plasmid was generated by replacing the wild type DcpS open reading frame from pcDNA3-Flag-DcpS (Liu et al, 2008b) with the DcpS cDNA containing the H277N mutation from pET-hDcpS-H277N (Liu et al, 2002) at the BamHI/XhoI restriction sites. The pTK-IREShyg-Flag-DcpS^{WT} and pTK-IREShyg-Flag-DcpS^{H277N} plasmids were constructed by inserting wild type and mutant human DcpS open reading frames amplified from pcDNA3-Flag-DcpS and pcDNA3-Flag-DcpS^{H277N} using primers 5' CACTATAGGCTAGCATGGACTACAAGGACGACG 3' and 5' TTATCTATGCGGCCGCCAGTGTGATGGATTCAG 3' containing the NheI and NotI restriction endonuclease sites respectively into the same restriction sites of the pTK-IREShyg-FLAG plasmid (Tan et al, 2014) in frame. The shRNA-resistant plasmid pTK-Flag-DcpS^{WT}shRR and pTK-Flag-DcpS^{H277N}shRR were generated by PCR mutagenesis using primers 5' TTCTCCAATGATATCTATAGTACGTATCACTTGTTCCCTCC 3' and 5' TGCTGTAGATATCGTTTGAAAATTGTAATTGGAGCTCAGGG 3' to alter the shRNA target sequence in pTK-IREShyg-Flag-DcpS^{WT} and pTK-IREShyg-Flag-DcpS^{H277N}.

pET-28a-DcpS^{Ins15} and pET-28a-DcpS^{T316M} plasmids was constructed as following. cDNAs encoding the various DcpS proteins were amplified from pDrive cloning plasmids which contained the full length the T316M mutation or the 15 aa insertion cDNA using primers 5' GAATTCGGATCCATGGCGGACGCAGCTCCTC 3' and 5' AAGCTTCTCGAGTCAGCTTTGCTGAGCCTCCTGC 3' containing the BamH1

and Xho1 restriction endonuclease sites respectively. The cDNAs were inserted into the corresponding restriction enzyme sites within pET28a.

pIRES-hyg- DcpS^{Ins15} plasmid was constructed by transferring DcpS^{Ins15} ORF cDNA from pET-28a-DcpS^{Ins15} into pIRES-hyp-Flag vector by BamH1 and Xho1digestion. The shRNA-resistant plasmid pIRES-hyg-DcpS^{Ins15}shRR was then generated by PCR mutagenesis using primers 5' TTCTCCAATGATATCTATAGTACG TATCACTTGTTCCCTCC 3' and 5' TGCTGTAGATATCGTTTGAAAATTGTAATT GGAGCTCAGGG 3' to alter the shRNA target sequence in pIRES-hyg-DcpS^{Ins15} plasmid.

pLJM-Flag-$^{DcpSWT/H277N/Ins15}$ plasmids was constructed inserting wild type and mutant human DcpS open reading frames amplified from plasmid pTK-Flag-DcpS^{WT}shRR, pTK-Flag-DcpS^{H277N}shRR and pIRES-hyg-DcpSIns15shRR using primers 5' GCTAGCGCTACCGGTCGCCACCATGGACTACAAGGACGACGATGACAAG 3' and 5' TCTAGATTCGAATTCTGGATCAGTTATCTATGCGGCCGC 3' containing the AgeI and EcoRI restriction endonuclease sites respectively into the same restriction sites of the pLJM1-GFP plasmid (Addgene).

All constructs were confirmed by sequencing.

His-tag protein purification

Recombinant proteins expressed from the pET vectors were generated in *E. coli* BL21(DE3) cells induced with 0.5 mM IPTG for 2 hours and cells harvested by a 20 minute centrifugation at 4000rpm at 4℃ and washed once with cold PBS. The resulting

pellet was resuspended in 1X binding buffer (5 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl, pH 7.9) with 1M urea, 0.5% TritonX-100 and protease inhibitors, lysed by sonication and the cell lysate isolated by centrifugation at 14000g for 20 minutes at 4℃. Purification of the recombinant protein was carried out using a charged nickel column according to the manufacturer (Novagen), except that 300 mM urea and 0.5% TritonX-100 was included in the binding buffer. The nickel column was then washed twice using wash buffer (60 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl, pH 7.9) with 800 mM urea and 0.5% TritonX-100 and once using wash buffer without urea and TritonX-100. Protein eluted from the nickel column was dialyzed against PBS and concentrated by Centricon centrifugal filter columns (Amicon) and stored at -80°C in 10% glycerol. Protein concentrations were determined by Bradford assay.

Cell culture and transfections

SH-SY5Y cells were grown in F/12 (Invitrogen) supplemented with 10% fetal bovine serum, penicillin–streptomycin and sodium pyruvate under 5% CO_2 at 37°C. SMA patient fibroblast cell lines GM03813, GM00232 and HEK293T cells were grown in DMEM (Invitrogen) and DcpS mutation patient lymphoblast cells PJMR5-3,4,7,8 cell were growth in RPMI1640 (Invitrogen) with the same complement of supplements as with SH-SY5Y cells. RG3039 was administered to the culture medium at a final concentration of 100nM and cells grown for 2 days before harvesting.

The 293T cells stably transformed with DcpS specific shRNA ($DcpS^{KD}$) was described in Shen et al. (2008). $DcpS^{KD}$ cells stably expressing an shRNA resistant wild

type (pTK-Flag-DcpS^{WT}shRR), catalytically inactive histidine 277 to asparagine mutant (pTK-Flag-DcpS^{MT}shRR) or the empty pTK-Flag vector were generated by transfecting the respective plasmids into DcpS^{KD} cells, with 3 μg/mL puromycin and 50 μg/mL hygromysin selection 48 hours post transfection. The 293T cells stably expressing inserted DcpS mutant (pTK-Flag-DcpS^{MT}shRR) were generated by transfecting the respective plasmids into293T cells with 50 μg/mL hygromysin selection 48 hours post transfection.

Lentiviral production and infection

All shRNA plasmids and packaging plasmids were purchased from Sigma. 293T cells were transfected by pLKO1-DcpS shRNA (Sigma # TRCN0000005571), pLKO.1-shXRN1 (Sigma # TRCN0000049676) or pLKO.1-shRRp41 (Sigma # TRCN0000051365), and pCMV-VSVG and psPAX2 with Lipofectamin2000 (Life technology) to generate viral particles. All protein expression lentiviral particles were produced in similar way. 293T cells were transfected by pLJM-GFP, pLJM-Flag-DcpS^{WT}, pLJM-Flag-^{MT} or pLJM-Flag-DcpS^{Ins15}, and pCMV-VSVG and psPAX2 with Lipofectamin2000 (Life technology). Culture medium was harvested two days post transfection and frozen at -80°C for storage (Bail et al, 2010).

All cells were infected with the viral particles in the presence of 8μg/ml hexadimethrime bromide, washed and fed with fresh medium 24 hours later. Cells were harvested for Western Blot analyses or real-time PCR analysis at indicated times post-infection.

RNA isolation, reverse transcription and Real time PCR

Cells were treated with actinomycin D (5 mg/L, Sigma) 48 hours post RG3039 treatment/ lentiviral infection, cells harvested and total RNA isolation carried out with TRIzol (Invitrogen) at the indicated time intervals. Total RNAs were treated with DNase (Promega) to degrade DNA prior to oligo(dT) (IDT)-directed reverse transcription with M-MLV-RT (Promega) according to the manufacturer's instructions. Values were quantified by real-time PCR using SYBR green PCR core reagent (Applied Biosystems) and the abundance of specific mRNAs were quantified using the standard-curve method according to the recommendations of the manufacturer. mRNA levels were normalized to the GAPDH mRNA and plotted against time. The primers used for real-time PCR are listed in Table I. mRNA half-life was calculated as reference (Tan et al, 2014).

Microarray

SH-SY5Y cells were treated for 48 hr with 100 nM RG3039, total RNA isolated with TRIzol (Invitrogen) according to the manufacturer and used for microarray analysis using the Affymetrix U133 Plus 2.0 human microarray at Expression Analysis (Durham, NC).

Western Blotting

Cells were sonicated in phosphate-buffered saline, and protein extract was resolved by 12.5% SDS-PAGE. Affinity-purified DcpS polyclonal antibody (1:200

dilution)(Shen et al, 2008), TRAF3(1:200 dilution, Santa Cruz), SMN(1:1000 dilution, BD), monoclonal anti-Flag antibody (1:5,000 dilution, Sigma), monoclonal anti-GAPDH antibody (1:1,000 dilution, Santa Cruz) were used for Western blot analysis and visualized using secondary antibodies coupled to horseradish peroxidase (Jackson ImmunoResearch) and chemiluminescence (ECL; GE Healthcare).

Immunofluorescence

To study the localization of $DcpS^{Ins15}$ protein, GM03813 cells were infected with lentivirus which express Flag-$DcpS^{Ins15}$. Flag tagged proteins were detected by indirect immunofluorescence with 1:200 dilution of anti-FLAG antibody (Sigma) and 1:200 dilution of FITC-conjugated anti-mouse secondary antibody. DcpS protein was detected with 1:20 affinity-purified DcpS polyclonal antibody and 1:200 dilution of Texas Red-conjugated anti-rabbit secondary antibody. Nuclei were stained for 3 min with DAPI (1 μg/ml).

RACE and DNA gel electrophoresis

5'RACE was conducted using 5' RACE System for Rapid Amplification of cDNA Ends, version 2.0 (Life Technologies). 3′ RACE was conducted with first strand cDNA synthesis using the adapter linked oligo-dT primer for reves-transcription. PCR products were resolved by 2% agarose gel electrophoresis and DNA fragments was purified using QIAquick Gel Extraction Kit (Qiagen) and cloned into pGEM-T Easy Vector (Promega).

Four colonies from each of the isoforms were picked and sequenced. Primers used in RACE and all semi-quantitative RT-PCR are listed in Table II.

Generation of labeled RNA and cap structures

Unlabeled, uncapped RNA corresponding to the pcDNA3 polylinker spanning the SP6 and T7 promoters was generated from SP6 RNA polymerase transcribed RNA, cap labeled and cap structure isolated as described (Liu et al, 2008a).

Electrophoretic mobility shift assays

Electrophoretic mobility shift assays of compound competition were carried out with *in vitro* P32 labeled cap structure (~2,000 cpm per reaction). Binding reactions were carried out in decapping buffer (10 mM Tris-HCl at pH 7.5, 100 mM KCl, 2 mM $MgCl_2$, 2 mM DTT, 0.5 mM $MnCl_2$) with 0.1μg of His-$DcpS^{MT}$ protein. RG3039, cold cap structure and D156676 compound was added to 2.5, 5, 10, 20 nM in a 20-μl total volume. Following a 30-min binding reaction at room temperature, the complexes were resolved on a 5% polyacrylamide gel (60:1 acrylamide:bis) in 0.5x Tris-Borate-EDTA buffer at 120V and exposed to PhosphorImager.

Electrophoretic mobility shift assays of $DcpS^{Ins15}$ and cap structure carried out with P32 labeled cap structure in the same buffer as above. 0.1, 0.5 and 1 μg His-$DcpS^{Ins15}$ protein was tested as well as 0.1μg of His-$DcpS^{MT}$ protein was used as positive control. Following a 30-min binding reaction at room temperature, the complexes were

resolved on a 5% polyacrylamide gel (60:1 acrylamide:bis) in 0.5x Tris-Borate-EDTA buffer at 120V and exposed to PhosphorImager.

In vitro decapping assays

Decapping assays were carried out with 10, 20, 40ng of the indicated recombinant proteins or 5μg total cell extract derived from patient lymphoblasts. Proteins or extract were incubated with the indicated labeled cap structures in decapping buffer (10 mM Tris-HCl at pH 7.5, 100 mM KCl, 2 mM $MgCl_2$, 2 mM DTT, 0.5 mM $MnCl_2$). Decapping reactions were carried out for 5min (for recombinant proteins) or 10, 20, 30 min (for cell extract) at 37°C and stopped by extracting once with phenol:chloroform (1:1). Decapping products were resolved by polyethyleneimine-cellulose TLC plates (Sigma-Aldrich) and developed in 0.45 M $(NH4)_2SO_4$ in a TLC chamber at room temperature. The TLC plates were air-dried and exposed to PhosphorImager for quantitation.

Dicistronic reporter assay

293T cells were treated with 100 nM RG3039 for two days prior to transfection with dicistronic construct. One day post-transfection with or without RG3039 presence, cells were harvested and assayed for Renilla and Firefly Luciferase activities using the Dual Luciferase Reporter Assay System (Promega). Cap-dependent translation efficiency was defined as the ratio of Renilla to Firefly Luciferases. Similar experiments were conducted in the 293T $DcpS^{KD}$, 293T Con^{KD}, 293T DcpS overexpression cell lines.

Generation of RNA in vitro

RNAs were in vitro transcribed from PCR-generated templates that contained T7 promoter sequences at the 5′ end. Cap- Renilla template was generated from pcDNA3.1 RL-hcvIRES-FL. Cap- Firefly template was generated from pcDNA3.1-FL. emcvIRES-Renilla template was generated from pRL. hcvIRES- Firefly template was generated from pcDNA3.1 RL-hcvIRES-FL. CrpvIRES- Firefly template was generated from pRL-CrPV-FL (Donovan & Copeland, 2010). Primers used in PCR of different reporter gene were listed in Table III. All PCR products were purified by gel electrophoresis before *in vitro* transcription. All RNAs were generated and purified using RiboMax kit (Promega) following the instruction. RNA and DNA concentration was determined by spectrophotometer at OD260.

In vitro translation

In vitro translation of firefly and renilla luciferase was conducted in Rabbit reticulocyte lysate kit (Promega) according to instruction. In each 50 uL reaction, RG3039 was added to 10uM final concentration. Lysate was incubated at 30 degree water buth and translation was stopped by ice-chilling at 10 min, 20 min and 30 min time point. Luciferase activity was detected using Dual Luciferase Reporter Assay System (Promega). The activity of control sample at 30 min was set to 100% and all other time points and RG3039 treated samples were normalized to it.

Table I. Real Time qRT-PCR Primers

PAQR8 F	AACGTCTGGACCCATTTACTG
PAQR8 R	CAGGTGAGGTAAGTGATTGAC
ATOH7 F	AAAGCTGTCCAAGTACGAGAC
ATOH7 R	CGAAGTGCTCACAGTGGAG
MAOB F	ACTCTGGACTGAATGTGGTTG
MAOB R	GATACGATTCTGGGTTGGTCC
RAB26 F	CGATTCAAGGATGGTGCTTTC
RAB26 R	ATGGGGTAACACTGCCGGAAC
DRNT1/HS370762 F	CACCTAGACTCATCACTTAGATCCACC
DRNT1/HS370762 R	GAGACCTGATGGCTACAACTGACA
DRNT2/BC011766 F	TGGAGAAGCGATGGATGACAGAGA
DRNT2/BC011766 R	GGTGAACGGACACAATTGCCAGAA
XRN1 F	AATTTGACAACTTGTACCTGG
XRN1 R	AGCCATAAAGAACACTTTCC
DcpS F	ATCCTGGAGAAGACGCC
DcpS R	CATCTCCCGTCTCTCGG
RRP41 F	TACATTGAGCAGGGCAACAC
RRP41 R	ATGGCTGCTTCGAAAGTCTG
GAPDH F	AGGTGAAGGTCGGAGTCAACGG
GAPDH R	CGTTCTCAGCCTTGACGGTGC
CLEC11A F	GAGGCCCTGATGCTGAAG
CLEC11A R	GGTTGGCGTTGCTTCCTC
MIB2 F	TGCAGTTCAACCACGAGAC
MIB2 R	TTCACTGTGTCAAGGTCGC
PIP4K2B F	TCGGGAGAGGTTTGGAATTG
PIP4K2B R	TGGCTGTCACTGTTGATGG
RNF165 F	CTGTCTATGCTGGAAGATGGAG
RNF165 R	TCCCAGTTGTGTCTCAATGTC
CSTF3 F	TCGTACAGAGGATCAGACCC
CSTF3 R	ATATCTCCCTTTTCTGCGAGC
DP1 F	CACCCACATCAGTCATCCAC
DP1 R	AGCACTTCTATGTCATCGTGG
Actin F	GAGAAAATCTGGCACCACACC
Actin R	CTGGATAGCAACGTACATGGC
DRNT1 upstream F	ACAGTCCAGTAATGTCACC
DRNT1 upstream R	CCCATTTGTGACATTCTAATG
DRNT2 upstream F	TCTCTTGGGGACCGG

DRNT2 upstream R	GTGCTCACTCTGTCGG
Dcps-WT F	AACAGCAGCTCGATGAC
DcpS-WT R	GATGGTCTCCCTTCATCC
DcpS total F	ATCCTGGAGAAGACGCC
DcpS total R	CATCTCCCGTCTCTCGG
Smn2 full length F	CATACTGGCTATTATATGGGTTTT
Smn2 full length R	TGCCAGCATTTCTCCTTAATTTAA
Smn total F	ACTGCAGCTTCCTTACAACAGTGG
Smn total R	CAATGGTAGCTGGGTAAATGCAACCG

Table II. Primers used in semi-quantitative RT-PCR

Anchor-oligo-dT	GCGAGCTCCGCGGCCGCGTTTTTTTTT
HS370762 5′race	CACCTAGACTCATCACTTAGATCCACC
HS370762 5′race nested	TGTCAGTTGTAGCCATCAGGTCTC
HS370762 3′race	AGATTTGAGGGACAACTGGG
HS370762 3′race nested	TGGGGAACTTTGTAGTTATATATGG
BC011766 5′race	GGTGAACGGACACAATTGCCAGAA
BC011766 5′race nested	TCTCTGTCATCCATCGCTTCTCCA
BC011766 3′race	TGGAGAAGCGATGGATGACAGAGA
BC011766 3′race nested	TTCTGGCAATTGTGTCCGTTCACC
HS370762 PF	CACCTAGACTCATCACTTAGATCCA
HS370762 PR1	AGGTTGGAAGACATTCAGCA
HS370762 PR2	CATTAGAATGTCACAAATGGG
HS370762 PR3	AATGCACAGTCCAGTAATGTCACC
HS370762 PR4	AATCGCTTGAACCCAGG
DcpS ex3 F	TTACTTTACCCCACCTGG
DcpS ex6 R	GATGGTCTCCCTTCATCC
DcpS ex2F	ATCCTGGAGAAGACGCC
DcpS ex3R	CATCTCCCGTCTCTCGG

Table III. Primers used for generating RNA template in vitro

T7-RL/FL F	TAGGCGTGTACGGTGGGAGGTC
T7-hcvIRES F	TAATACGACTCACTATAGGGATCCGGCGACACT CCGCCATGAG
T7-EMCV IRES F	TAGAGTACTTAATACGACTCACTATAGGC
T7-CrPV IRES F	TAATACGACTCACTATAGGTAACGTTACTGGCC GAA
A60-RL R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTTTTTTTTTGCGGAGTGTCGCC GGATCC
A60-EMCV-RL R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTTTTTTTTTAGAATTATTGTT CATTTTGAGAACTCGC
A60-IRES-FL R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTTTTTTTTTACAATTTGGAC TTTCCGCCC
A60-cap-FL R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTTTTTTTTTACACGGCGATC TTTCCGCCC

Chapter I DcpS is a Transcript Specific Modulator of RNA in Mammalian Cells

Summary

The scavenger decapping enzyme DcpS is a multifunctional protein initially identified by its property to hydrolyze the resulting cap structure following 3′ end mRNA decay. In *S. cerevisiae*, the DcpS homolog Dcs1 is an obligate cofactor for the 5'-3' exoribonuclease Xrn1 while the *C. elegans* homolog Dcs-1, facilitates Xrn1 mediated microRNA turnover. In both cases, this function is independent of the decapping activity. In mammals, the decapping function of DcpS facilitates first intron splicing likely by modulating nuclear cap structure concentrations. However, whether DcpS and its decapping activity can affect mRNA steady state or stability in mammalian cells remains unknown. We sought to determine DcpS target genes in mammalian cells using a cell permeable DcpS inhibitor compound, RG3039 initially developed for therapeutic treatment of spinal muscular atrophy. Global mRNA levels were examined following DcpS decapping inhibition with RG3039. The steady state levels of 222 RNAs were altered upon RG3039 treatment. Of a subset selected for validation, two transcripts that appear to be long non coding RNA HS370762 and BC011766, were dependent on DcpS and its scavenger decapping catalytic activity and referred to as noncoding DcpS responsive transcripts (ncDRT) 1 and 2 respectively. Interestingly, only the increase in DRNT1 transcript was accompanied with an increase of its RNA stability and this increase was dependent on both DcpS and Xrn1. Importantly, unlike in yeast where the DcpS homolog is an obligate cofactor for Xrn1, stability of additional Xrn1 dependent

RNAs were not altered by a reduction in DcpS levels. Collectively, our data demonstrate that DcpS in conjunction with Xrn1 can regulate RNA stability in a transcript selective manner in mammalian cells.

Introduction

In eukaryotic cells, mRNA degradation plays an important role in the control of gene expression and therefore is highly regulated. Following an initial deadenylation step to remove the poly(A) tail, the remaining mRNA can undergo exonucleolytic decay from either 3′ or the 5′ end. In the 5′ to 3′ decay pathway, the 5′ cap is initially removed by Dcp2 or Nudt16 decapping enzymes, resulting in a 5′ end monophosphorylated RNA (Lykke-Andersen, 2002; Song et al, 2010; van Dijk et al, 2002; Wang et al, 2002). The uncapped product undergoes 5′ to 3′ exoribonuclease decay by Xrn1 (Hsu & Stevens, 1993). Xrn1 is highly conserved in eukaryotes and functions in the decay of cytoplasmic mRNAs (Muhlrad et al, 1995), in the quality control of aberrant mRNAs (Conti & Izaurralde, 2005) and noncoding RNAs (Chernyakov et al, 2008; van Dijk et al, 2011). In the 3′ to 5′ pathway, the multi-subunit exosome complex degrades the deadenylated RNA from the 3′ end, leaving the residual m^7GpppN cap structure (Anderson & Parker, 1998; Liu et al, 2006). The resulting cap structure is a substrate of the scavenger decapping enzyme, DcpS to release m^7Gp and ppN products (Wang 2001, Liu, 2002). In yeast, the DcpS homolog Dcs1p, facilitates 5' to 3' exoribonucleolytic decay (Liu & Kiledjian, 2005) as an obligate cofactor for the Xrn1 exoribonuclease by an unknown mechanism independent of its decapping catalytic activity (Sinturel et al, 2012). Moreover, the *C.elegans* Dcs-1 was shown to physically interact with Xrn-1 and promote

specific microRNA degradation also independent of its decapping activity (Bosse et al, 2013).

DcpS hydrolyzes the triphosphate linkage of the cap structure utilizing an evolutionarily conserved Histidine Triad (HIT) motif (Liu et al, 2002) with the central histidine serving as the nucleophile for the hydrolase activity (Lima et al, 1997). Insights into the molecular mechanism of DcpS decapping were provided by structural analysis, revealing a homo-dimer complex where each monomer possesses a distinct binding and catalytic site (Gu et al, 2004). The homo-dimer displays an asymmetric structure in the ligand-bound form and a dynamic state, where each monomer can alternate from an open to closed state in a mutually exclusive manner (Liu et al, 2008b). Hydrolysis of cap structure only occurs in the closed conformation (Gu et al, 2004).

A potent DcpS decapping inhibitor was identified in efforts to isolate small molecules that increase expression of the survival of motor neuron 2 (SMN2) gene in spinal muscular atrophy (SMA) patient cells. Co-crystal structures of a C5 substituted quinazoline compound, D156844, revealed it binds to the DcpS cap binding active site and inhibits its decapping (Singh et al, 2008; Thurmond et al, 2008). Further optimization generated the RG3039 derivative, which is also a potent DcpS decapping inhibitor (Gogliotti et al, 2013). Although the quinazoline compounds were capable of improving motor neuron function and extending survival of SMA model mice (Butchbach et al, 2010; Gogliotti et al, 2013; Van Meerbeke et al, 2013b), statistically significant increases in SMN2 mRNA or SMN protein were not evident.

In this chapter, we took advantage of the DcpS decapping inhibition property of RG3039 to determine the impact on DcpS decapping on mRNA expression. Interestingly only 222 RNAs were detected to be altered upon RG3039 treatment and of a subset of genes validated, four were found to be responsive to RG3039 independent of DcpS and two were mediated through DcpS. Importantly, both the DcpS responsive transcripts are putative non-coding RNAs, with the stability of one RNA modulated by both DcpS and Xrn1. Our data indicate that DcpS is a transcript-restricted modulator of RNA stability in mammalian cells and the RG3039 quinazoline compound is more pleotropic than initially anticipated and can influence gene expression in both a DcpS dependent and independent manner.

Results

Steady-state levels of RNAs are altered in cells treated with RG3039

Despite its role in the last step of 3′ end mRNA decay pathway, DcpS also contributes to mRNA or miRNA turnover in *S. cerevisiae* and *C. elegans*, respectively. To determine whether the human DcpS decapping function influences the expression of mRNAs in mammalian cells, we utilized a DcpS decapping inhibitor to monitor changes in global mRNA levels following cellular treatment with RG3039. RG3039 is a cell permeable quinazoline derivative compound that is a potent and selective inhibitor of DcpS decapping (Gogliotti et al, 2013; Singh et al, 2008). RNA isolated from human SH-SY5Y retinoblastoma cells treated for two days with 100nM RG3039 were used to screen the Affymetrix U133 Plus 2.0 human microarray. RNA profiles from treated and untreated cells revealed levels of 222 RNA were altered more than 2 fold between the

two cell parameters with p values <0.05 (Table IV) with 171 increased and 51 decreased. Approximately 75% of the RNAs corresponded to mRNAs, while the remaining 25% mapped to uncharacterized transcripts consistent with non-coding RNA. A subset of 12 RNAs were randomly chosen for further validation by direct quantitative reverse transcription (qRT)-PCR. The steady state levels of six were found to reproducibly and significantly change greater than two fold upon RG3039 treatment. The levels of four RNAs, HS370762, BC011766, ATOH7 and RAB26, were elevated, while PAQR8 and MAOB were reduced compared to untreated cells (Figure 1A). Similar changes for the six RNAs were also detected in HEK293T cells treated with 100nM RG3039 (Figure 1B) demonstrating the generality of the results in different cell lines. All subsequent analyses were carried out in HEK293T cells

RNA steady state levels are influenced by DcpS

To determine whether changes in the levels of the six RNAs were mediated through DcpS, we utilized a HEK293T cell line constitutively expressing a DcpS directed shRNA ($DcpS^{KD}$) which reduced DcpS protein levels by 90% (Figure 2A) (Shen et al, 2008). Consistent with the outcome of RG3039 treated cells, steady state levels of both HS370762 and BC011766 increased in $DcpS^{KD}$ cells (Figure 2B). Complementation with shRNA-resistant exogenous wild type DcpS ($DcpS^{WT}$) (Figure 2A), but not a catalytically inactive DcpS mutant (Figure 2A) harboring an asparagine in the central histidine of the HIT motif ($DcpS^{MT}$), reversed the phenotype and reduced their RNA to levels comparable to control cells (Figure 2B). These data demonstrate that steady state accumulation of HS370762 and BC011766 RNAs are modulated by DcpS and moreover,

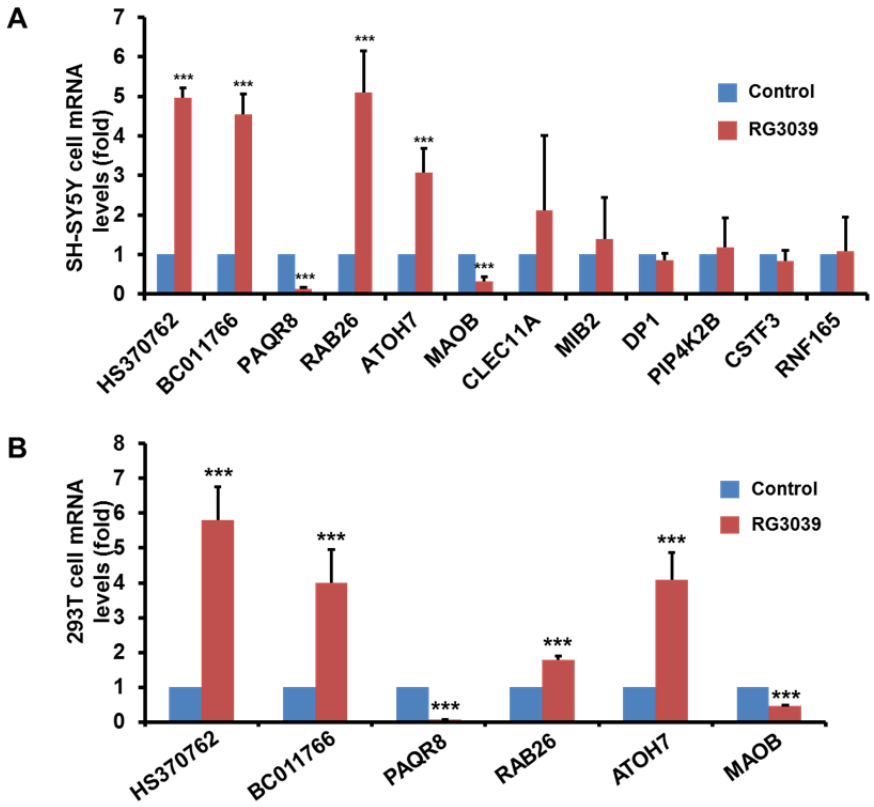

Figure 1. Validation of a subset of RG3039 target RNAs.

(A) Levels of twelve randomly chosen RNAs from the 222 RNAs that deviated by at least 2 fold in SH-SY5Y neuroblastoma cells treated for 48 hours with 100nM RG3039 in were tested by quantitative reverse-transcription (qRT)-PCR. RNA levels are presented relative to actin mRNA and derived from at least three independent experiments. Six of the RNAs were reproducibly altered greater than 2 fold and constitute RG3039 responsive RNAs. **(B)** Validated RG3039 target RNAs from SH-SY5Y cells were further confirmed in HEK293T cells to determine the consistency of RG3039 in different cell backgrounds. RNA levels are presented relative to the GAPDH mRNA and derived from at least three independent experiments. All six RNAs were similarly altered in the HEK293T cells as they were in SH-SY5Y cells. Error bars represent +/- SD. (*** represents $p < 0.001$ (Student's *t* test))

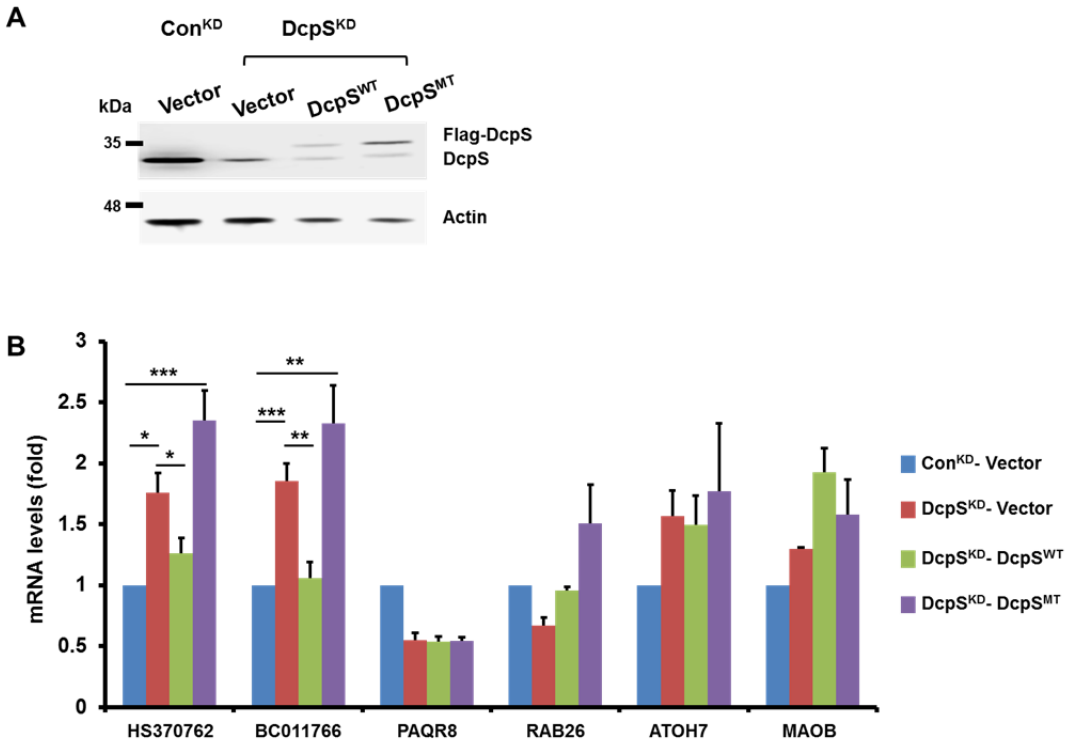

Figure 2. RG3039 target RNAs, HS370762 and BC011766 are regulated by DcpS.

(A) Endogenous and exogenously expressed DcpS protein levels were measured by Western Blot in HEK293T control knock down (Con^{KD}) or DcpS knock down ($DcpS^{KD}$) cells, complemented with shRNA-resistant $DcpS^{WT}$ of a catalytically inactive DcpS harboring a histidine 277 to asparagine ($DcpS^{MT}$) mutation as shown. Complementation with a control plasmid lacking the DcpS cDNA is denoted as "Vector". **(B)** RNA levels for the indicated RNAs were measured by qRT-PCR from cells in A. RNA levels are presented relative to the GAPDH mRNA and derived from three independent experiments. RNA levels for only HS370762 and BC011766 were responsive to depletion and complementation with DcpS. Error bars represent +/- SD. P values are denoted by asterisks; * represents $p < 0.05$, ** represents $p < 0.01$, *** represents $p < 0.001$ (Student's *t* test).

by the decapping activity of DcpS. In contrast, levels of the remaining four mRNAs were either consistent with RG3039 treatment but not complemented by DcpS, or inconsistent with RG3039 treatment suggesting they are not direct DcpS targets. We conclude DcpS decapping activity is important in HS370762 and BC011766 RNA expression, while the levels of ATOH7, RAB26, PAQR8 and MAOB mRNAs are responsive to RG3039 in a DcpS independent manner.

HS370762 and BC011766 transcripts are putative noncoding RNAs

HS370762 and BC011766 are uncharacterized transcripts. HS370762 maps to human EST AW665080 on a region of chromosome 9 containing multiple ESTs (Figure 3A) and BC011766 maps to a short transcript on human chromosome 3 (Figure 3D). Rapid amplification of cDNA ends (RACE) analysis was used to map the termini of both RNAs. The 3′ polyadenylated tail of HS370762 was identified with oligo-dT linked adaptor 3′ race and found to correspond to the minus strand. 5′ RACE extended the RNA to the 5′ end of the AW665080 EST. Multiple ESTs contiguous with the DNA map to this region indicating HS370762 is a nonspliced transcript. The 5′ end was further confirmed by direct reverse transcription and PCR (RT-PCR) amplification with a series of 5′ primers within, and upstream, of AW665080. As expected, a robust RT-PCR product resulted with the AW665080 specific primers, while primers upstream of AW665080 generated relatively less intense products (Figure 3B) suggesting the predominant transcript corresponded to the AW665080 EST. More significantly, qRT-PCR analysis confirmed RNA from AW665080 was responsive to RG3039 treatment,

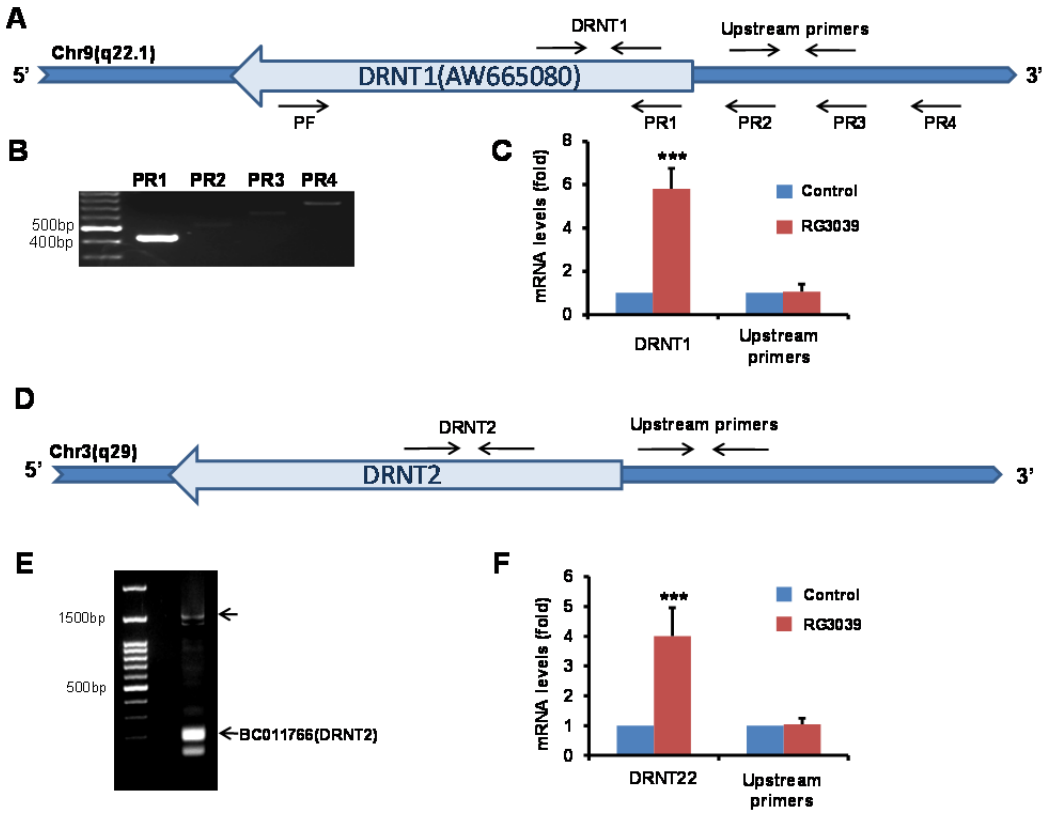

Figure 3. Mapping the genomic loci corresponding to HS370762 (DRNT1) and BC011766 (DRNT2) RNAs.

(A) Schematic of the genomic loci and direction on Chromosome 9 for HS370762 /DRNT1 is shown. Black arrows indicate position and orientation of primers used in (B) and (C). **(B)** PCR of oligo(dT) reverse transcribed total RNA from HEK293T cells using the PF forward primer and the four different reverse primers (PR1, PR2, PR3 and PR4) are shown. **(C)** Levels of HS370762/ DRNT1 RNA and potential transcript originating upstream of HS370762 quantitated by qRT-PCR from RNA derived from RG3039 treated or untreated HEK293T cells. **(D)** Schematic of the BC011766/DRNT2 genomic loci on chromosome 3. . Black arrows indicate positions of qRT-PCR primers used in F. (E) Agarose gel of 5´RACE products. **(F)** Quantitation of DRNT2 and potential longer isoform of BC011766DRNT2RNAs are presented relative to GAPDH mRNA and derived from three independent experiments.

while transcripts transcribed from regions upstream of this EST were not responsive to RG3039 (Figure 3C). The lack of an extensive open reading frame within this 400-nucleotide RNA suggests HS370762 is likely a noncoding RNA which we will refer to as DcpS-responsive noncoding transcript 1 (DRNT1).

A similar analysis was carried out to map the BC011766 RNA. The 3′ polyadenylated tail of BC011766 was identified and corresponded to the minus strand. 5′ RACE of BC011766 revealed the major transcript of approximately 500 nucleotides contiguous with the genomic sequence (Figure 3D and 3E). Unexpectedly the 5′ RACE also revealed a less intense and longer product that extended an additional 1100 nucleotides contiguous with the genomic sequence upstream from the 5′ end of BC011766. To verify whether this longer transcript was an isoform of BC011766, qRT-PCR primers that amplifying a region 150bp upstream of BC011766 confirmed the low abundance of this transcript and its lack of response to RG3039 treatment (Figure 3F), indicating that only levels of the 518-nucleotide BC011766 RNA are elevated by an inhibition of DcpS decapping. The lack of a predicted open reading frame in the BC011766 RNA indicates it is also a noncoding RNA and we will refer it as DcpS-responsive noncoding transcript 2 (DRNT2).

DRNT1 RNA stability is altered by RG3039.

We next determined whether changes in the steady state levels of the RG3039 target RNAs could be attributed to transcript stability. RNA half-lives for the six different RG3039 targets were determined following actinomycin D-directed

transcriptional silencing in RG3039 treated or untreated HEK293T cells. Consistent with the observed increase in DRNT1steady state levels, its stability increased >2 fold with RG3039 treatment (Figure 4) while the stability of DRNT2 was unchanged following RG3039 treatment. Interestingly, the PAQR8 mRNA was destabilized by RG3039 treatment. This result is consistent with the observed RG3039 mediated decrease in steady-state levels (Figure 2). Stability of the remaining mRNAs was refractory to RG3039 treatment indicating the changes in steady state mRNA levels for these mRNAs were not mediated through transcript stability. Collectively, our data reveal that RG3039 is pleotropic and has the capacity to impart transcript specific regulation of a subset of RNAs either through RNA stability (DRNT1 and PAQR8) or by a manner independent of RNA stability (DRNT2, ATOH7, RAB26 and MAOB).

DcpS modulates the stability of DRNT1.

DRNT1 RNA stability was next examined in $DcpS^{KD}$ cells (Figure 5A) to determine whether the RG3039-directed increase in DRNT1 stability was mediated through DcpS. The half-life of DRNT1 increased in $DcpS^{KD}$ cells relative to control cells ($T_{1/2}$ of 2.7 hrs and 1.7 hrs respectively), while statistically significant changes in half lives were not observed for PAQR8 and DRNT2 RNAs (Figure 5B). -Furthermore, lentiviral directed shRNA knockdown of Xrn1 (Figure 5A) but not Rrp41 component of the 3′–5′ exosome complex, stabilized DRNT1 (Figure 5B), indicating that DRNT1 is primarily degraded by Xrn1 in the 5′end decay pathway. Importantly, the stability of DRNT1 was comparable in cells with either singular reduction of DcpS or Xrn1 proteins

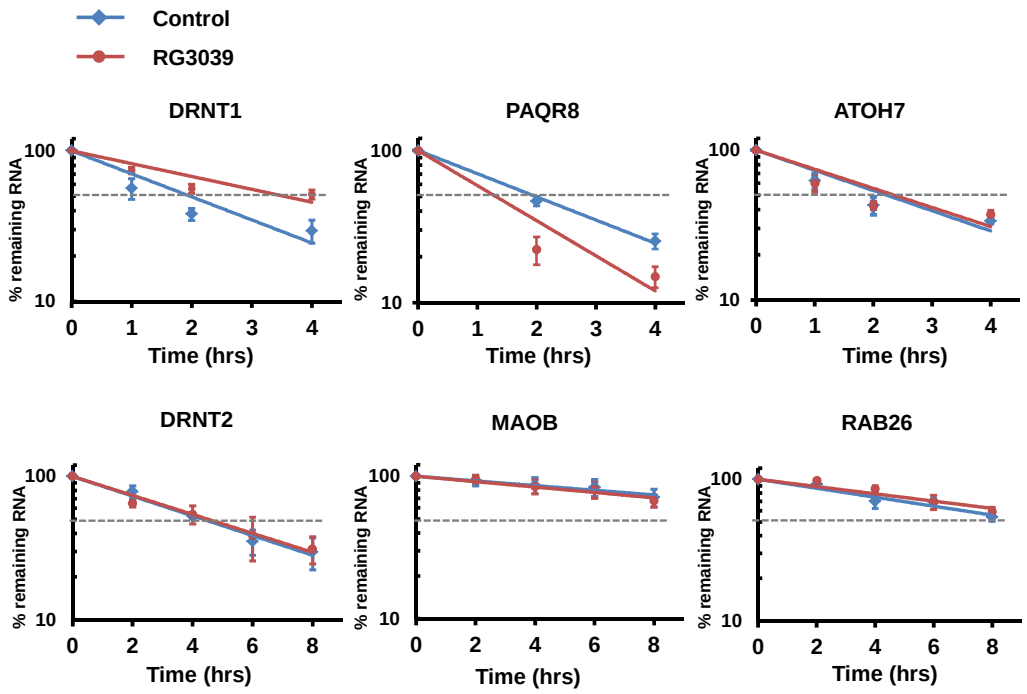

Figure 4. RG3039 influences the stability of HS370762 and PAQR8 RNAs.

Stability of the denoted RNAs was determined following actinomycin D-mediated transcriptional silencing in HEK293T cells treated or untreated with RG3039. Values normalized to levels of GAPDH mRNA are presented relative to their respective levels at time zero. Data were derived from at least three independent experiments. Dotted line represents 50%.

(Figure 5B) or both reduced simultaneously (Figure 5C and 5D) suggesting DcpS functions in the same pathway as, and likely with, Xrn1. In contrast, PAQR8 and DRNT2 RNAs where both stabilized with reduced Xrn1 levels but not DcpS (Figure 5B). These data demonstrate DcpS and Xrn1 do not coordinately regulate all Xrn1 responsive RNAs and the observed coordinate regulation of DRNT1 by DcpS and Xrn1 is transcript specific and dependent on DcpS catalytic activity.

RG3039 regulates translation of luciferase reporter

Based on our previously published data where the decapping function of DcpS facilitates first intron splicing by a mechanism postulated to be by modulating nuclear cap structure concentrations (Shen et al, 2008), we asked whether DcpS could influence mRNA translation. We hypothesized that reduction in DcpS levels may impact mRNA translation by altering cytoplasmic cap structure concentrations and in turn translation initiation. To investigate this possibility, we utilized a dicistronic luciferase reporter system in cells treated with RG3039 to inhibit DcpS decapping. The pcDNA3.1-RL-hcvIRES-FL reporter contains two open reading frames. The upstream cistron encodes Renilla Luciferase translated by a cap-dependent mechanism, and the downstream cistron encoding the Firefly Luciferase whose protein synthesis was driven through a hepatitis C viral Internal Ribosomal Entry Site (HCV IRES) that recruited the ribosome directly via its secondary structure (Figure 6A). Stably transformed HEK293T $DcpS^{KD}$ cells constitutively expressing either $DcpS^{WT}$ or catalytic inactive $DcpS^{MT}$ (Figure 6B) were pre-treated with 100nM RG3039 for two days and then transfected with the dicistronic luciferase plasmid as well as control untreated cells. Cells were retreated with

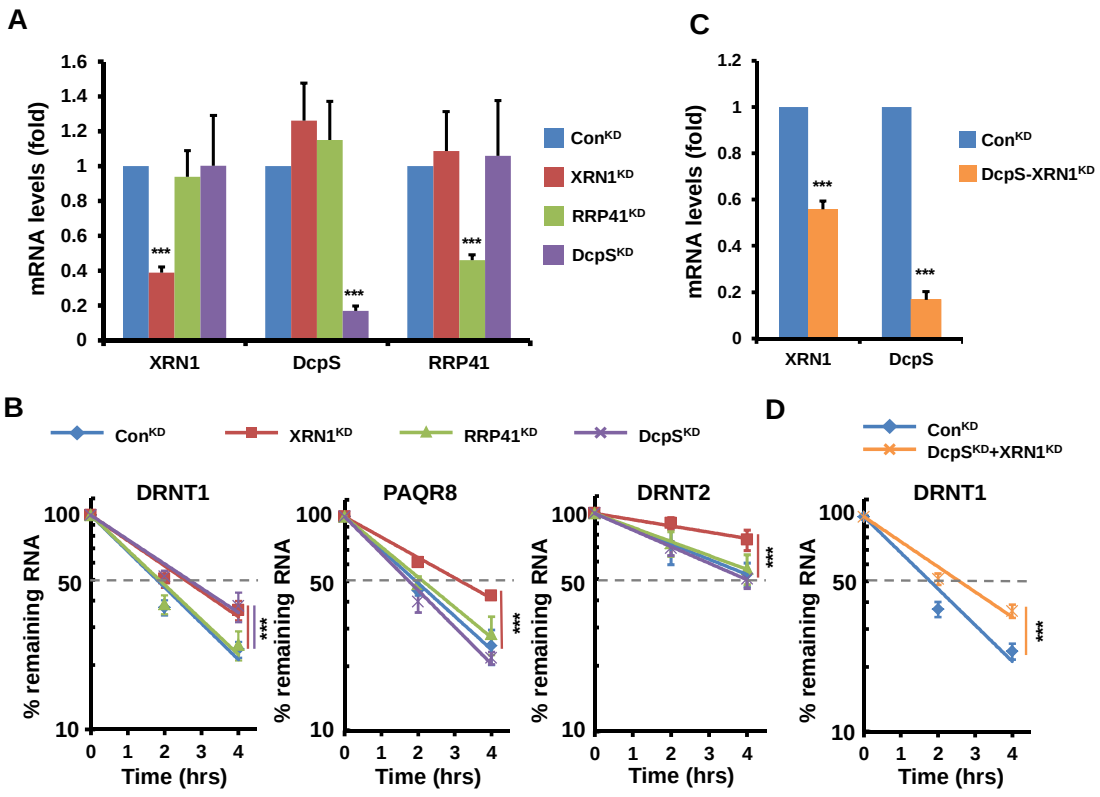

Figure 5. Stability of the DRNT1 RNA is mediated through both DcpS and Xrn1.

(A) DcpS, XRN1 and RRP41 mRNA levels were measured by qRT-PCR in HEK293T cells with control or the indicated mRNA knock down backgrounds. **(B)** Degradation of the indicated RNAs was determined by qRT-PCR in cells from A and as described in the legend to Figure 4. **(C)** Levels of DcpS and XRN1 mRNAs were measured by qRT-PCR in HEK293T cells with either a control or double DcpS and XRN1 knocking down. **(D)** Degradation of the DRNT1 mRNA was determined by qRT-PCR in cells from C and as described in the legend to Figure 4. All mRNA levels are presented relative to GAPDH mRNA and derived from three independent experiments, p-values from comparison of the decay rates are presented with asterisks. (*** represents $p < 0.001$ (Two-tailed extra sum-of-squares F test))

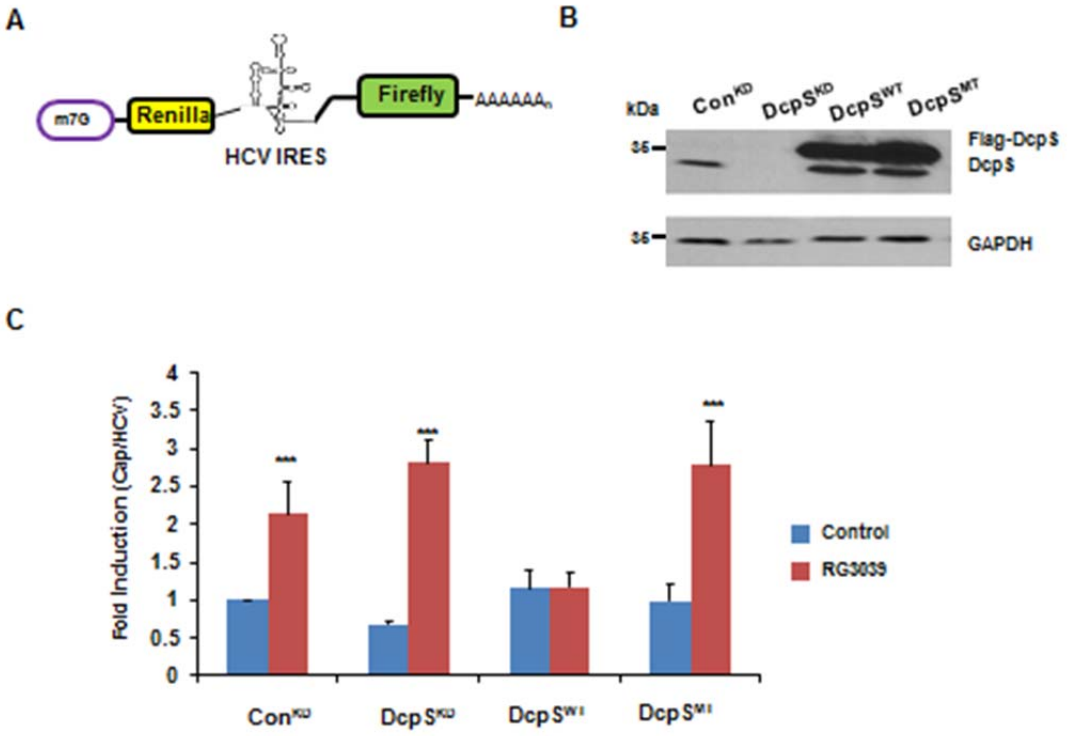

Figure 6. RG3039 alters the ratio of Cap dependent and Cap independent translation in cells.

(A) Schematic of the mRNA product of pRL-IRES-FL dicistronic luciferase plasmid is shown . **(B).** Western blot of DcpS knock down and overexpression cell lines. Specific antibody was used to detect DcpS. GAPDH was used as loading control. **(C)** All stable cells were pre-treated with RG3039 for two days and were transfected with the dicistronic reporter, then assayed for luciferase activity after 8 hours. The ratio of cap-dependent/cap- independent translation was calculated by dividing Renilla luciferase activity with Firefly luciferase activity. The ratio in control cell line without compound treatment was set as one. Three independent experiments were averaged and the error bars represent the standard deviation. P values are denoted by asterisks; *** represents $p < 0.001$ (Student's t test).

RG3039 following the transfections and cells were harvested 2 days post-transfection. The ratio of Renilla to Firefly luciferase activities were used to determine the extent of cap-dependent to cap-independent translation. RG3039 stimulated cap-dependent translation by two fold in control cells suggesting a role in translational regulation (Figure 6C). However, this affect was restricted to RG3039 and not a function of DcpS since a similar altered ratio of cap-dependent/cap-independent translation was not observed in the DcpS knockdown cells (Figure 6C). Importantly, complementation with wild type DcpS but not the catalytically inactive mutant DcpS reversed the stimulatory effect of RG3039 on cap dependent translation. These data indicate that RG3039 can augment cap dependent luciferase translation independent of DcpS, but overexpression of wild type DcpS likely sequesters RG3039 and prevents this function. This conclusion also predicts that the catalytically inactive DcpS is unable to bind RG3039 and thus unable to sequester RG3039.

Catalytic inactive DcpS mutant does not bind to RG3039

The results above indicate that RG3039 has a lower binding affinity to $DcpS^{MT}$ compared to $DcpS^{WT}$. To test this hypothesis, electrophoretic mobility shift assays (EMSA) with recombinant $DcpS^{MT}$ protein, which can bind but not hydrolyze cap structure, and ^{32}P labeled cap structure were carried out in the presence of RG3039 or unlabeled cap structure competitors. As expected, self-competition by cap structure efficiently competed the binding of ^{32}P-labeled cap structure (Figure 7), while RG3039 or a related quanzoline compound that is unable to bind DcpS (Singh et al, 2008) failed to

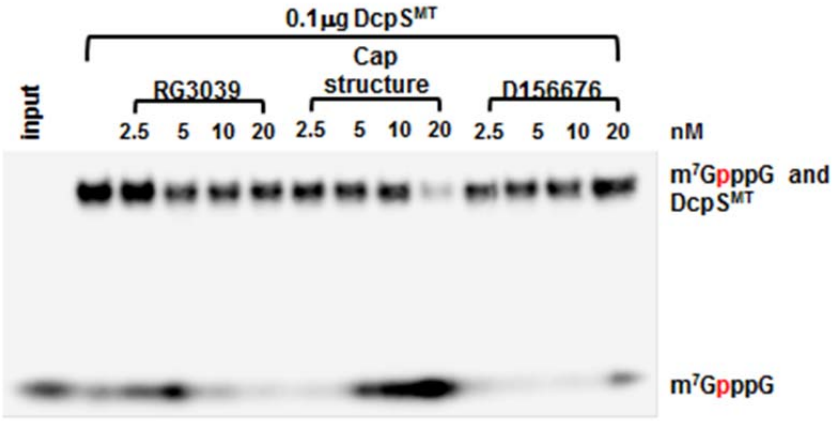

Figure 7. Catalytic inactive DcpS protein had lower affinity to RG3039 than cap structure.

The affinity of $DcpS^{MT}$ and RG3039 was tested by Eelectrophoretic Mobility Shift Assay. Recombinant $DcpS^{MT}$ and ^{32}P-alpha-GTP labeled cap were incubated in the presence of RG3039, cold cap structure and a related quinazoline control compound D156676 that does not bind DcpS. DcpS-cap complex was resolved by native PAGE. RG3039 did not compete with $DcpS^{MT}$-cap binding compared to cold cap structure.

compete labeled cap structure even at the highest concentration tested (Figure 7). Since RG3039 binds to $DcpS^{WT}$ with higher affinity than cap structure (Singh et al, 2008), these data indirectly demonstrate that $DcpS^{MT}$ has a reduced capacity to bind RG3039 compared to $DcpS^{WT}$. Collectively, our data indicate that, RG3039 binds to wild type DcpS more efficiently than catalytic mutant DcpS.

RG3039 specifically inhibited firefly luciferase translation independent of DcpS.

The luciferase assays above directly measure luciferase activity which is generally a reflection of the translation of each respective luciferase mRNA. To directly determine whether the relative increase in Renilla luciferase activity by RG3039 was a function of increased translation, we utilized a rabbit reticulocyte lysate (RRL) in vitro translation system. Before the translation experiment, we first verified that DcpS decapping activity exists in the reticulocyte lysate. As shown in Figure 8A, RRL contains DcpS decapping activity and this activity is efficiently competed with RG3039. Having determined the presence of DcpS decapping in RRL we next addressed whether translation of either luciferase mRNA is affected by RG3039 in vitro. Translation of two different luciferase mRNAs with either a 5′ cap, the HCV or EMCV IRES were monitored over time as indicated in Figure 8B. Unexpectedly, RG3039 specifically inhibited firefly luciferase mRNA translation by 2 fold regardless whether it utilized a cap or IRES for translation initiation (Figure 8B). In contrast, renilla luciferase mRNA translation was minimally affected by RG3039 whether it was capped or not (Figure 8B). RG3039 did not inhibit firefly luciferase protein activity because addition of RG3039 to in vitro translated luciferase protein did not affect luciferase activity (data not shown). Although we have

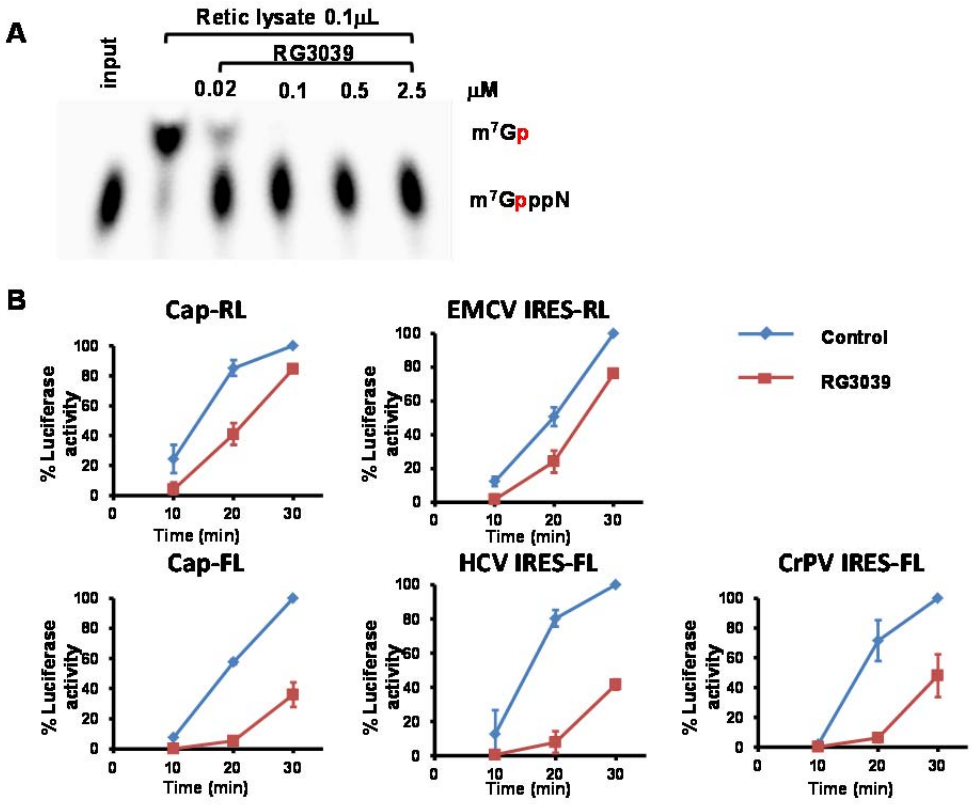

Figure 8. RG3039 inhibited firefly luciferase RNA translation in vitro.

(A) Rabbit reticulocyte lysate was tested for DcpS decapping activity in the presence of increasing RG3039 concentration, and products were resolved on TLC developed in 0.45M $NH_4(SO_4)_2$. Untreated lysate was used as a positive control. RG3039 demonstrated dose-dependent repression of DcpS activity in the rabbit reticulocyte lysate. **(B)** Capped renilla luciferase mRNA, capped firefly luciferase mRNA, EMCV IRES renilla luciferase mRNA, HCV IRES firefly luciferase mRNA and CrPV IRES firefly luciferase mRNA translation was monitored by luciferase activity at the indicated times in rabbit reticulocyte lysate with or without 20µM RG3039. Luciferase activity of untreated sample at 30min was set to 100%.

not yet ruled out the possibility that firefly luciferase activity was affected by incorporation of RG3039 during translation, these results revealed that the increased cap dependent/cap independent translation ratio we initially observed (Figure 7) was the consequence of inhibition of firefly mRNA translation rather than stimulation of cap dependent initiation. These data also indicate that RG3039 can function as a transcript specific mRNA translational inhibitor independent of DcpS. Further analyses are necessary to determine the mechanism by which RG3039 can selectively inhibit firefly luciferase mRNA translation.

Discussion

In this chapter, we demonstrate that the human DcpS protein and the DcpS decapping inhibitor, RG3039, both confer regulator roles on gene expression. Steady state levels of a small subset of RNAs changed upon treatment of either SH-SY5Y or HEK293T cells with RG3039. Of the six validated RNAs, increases in DRNT1 and DRNT2 were mediated through the inhibition of DcpS decapping since a similar increase was also observed in cells knocked down for DcpS that could be reversed with wild type but not a catalytically inactive DcpS protein (Figure 2). The function of RG3039 on the remaining four mRNAs was through a mechanism independent of DcpS (Figure 2). In the case of PAQR8, RG3039 promoted rapid decay of this mRNA while the stability of ATOH7, RAB26 and MAOB were unaffected suggesting an influence through mRNA stability and an mRNA stability independent mechanism. Our data further demonstrates a transcript specific role of DcpS in RNA decay where the stability of DRNT1, but not DRNT2 was increased in $DcpS^{KD}$ cells indicating DcpS promotes selective decay of

DRNT1. Furthermore, the comparable increased of DRNT1 stability in $DcpS^{KD}$, $Xrn1^{KD}$ and double $DcpS^{KD}$ and $Xrn1^{KD}$ cells suggests both proteins function within the same pathway and DcpS likely accentuates Xrn1 function similar to that reported in *S. cerevisiae* (Sinturel et al, 2012) and *C. elegans* (Bosse et al, 2013). Interestingly, unlike the scenario in *S. cerevisiae* and *C. elegans* where the decapping activity of the DcpS homolog was dispensable for its stimulation of Xrn1 decay, human DcpS requires a functional catalytic decapping activity. In addition, we also demonstrated a transcript specific role of RG3039 in the inhibition of firefly luciferase mRNA translation in a manner independent of DcpS. Collectively, our data demonstrate: 1) RG3039 can modulate the expression of a subset of transcripts by DcpS dependent and independent mechanisms; 2) despite functioning in the last step of mRNA 3′end decay, human DcpS can also modulate RNA stability in a transcript-specific manner and this modulation is dependent on its decapping activity and mediated through Xrn1 at least for the DRNT1 transcript.

RG3039 was initially identified and developed as a drug candidate for the treatment of SMA (Jarecki et al, 2005; Singh et al, 2008; Thurmond et al, 2008) and can increase the survival of mouse models of SMA (Butchbach et al, 2010; Gogliotti et al, 2013; Van Meerbeke et al, 2013b). Despite it being a potent inhibitor of DcpS decapping (Singh et al, 2008), the lack of detectable increase in SMN mRNA and protein in treated animals confounds the mechanism by which RG3039 promotes survival of SMA mice. Our results demonstrate that RG3039 can modulate RNA levels through its inhibition of DcpS as well as a mechanism independent of DcpS. Steady state levels of DRNT1 and DRNT2 increased with either RG3039 treatment (Figure 1) or knock down of DcpS

(Figure 2). The complementation of this phenotype with wild type but not a catalytically inactive DcpS demonstrates the significance of DcpS decapping for this phenotype. Surprisingly, modulation of four of the characterized RG3039 target RNAs (PAQR8, ATOH7, RAB26 and MAOB) did not appear to be mediated through DcpS since detected modification were not consistent in cells knocked down for DcpS or complemented by exogenous DcpS. Interestingly, the steady state increase in DRNT1 and decrease in PAQR8 could be attributed to a respective increase and decrease in RNA stability, while the stability of the remaining four RNAs was refractory to RG3039 treatment. How RG3039 can modulate RNA levels by a mechanism dependent and independent of transcript stability remains to be determined.

In an initial experiment, we found RG3039 treatment increased cap-dependent translation using a dicistronic luciferase reporter (Figure 6). Further investigation using in vitro translation system revealed that this phenomenon resulted from specific inhibition of firefly luciferase mRNA translation by the RG3039 compound (Figure 8). Therefore the effect of RG3039 on translation was not dependent on whether the translation was cap or IRES directed, but largely transcript specific. Nevertheless, the same effect was not observed in DcpS knockdown cells, indicating this function of RG3039 was not acting through DcpS (Figure 6). Conversely, the translational effect of RG3039 was accentuated in DcpS knockdown cells. Overexpression of wild type DcpS protein reversed the translational effect of RG3039, suggesting the excess DcpS protein functionally sequestered the compound. Overexpression of catalytic mutant DcpS had no impact on RG3039-directed translational effect, consistent with poor binding of the compound by the mutant DcpS (Figure 7). The transcript specific action of RG3039 independent of

DcpS further reinforces the pleotropic effect of the compound and may confound its potential clinical application.

At present, the mechanism by which DcpS confers stability is unclear, however, the stimulation of Xrn1 in humans is likely different than that observed in yeast and worms. We have been unable to detect a direct stable interaction between DcpS and Xrn1 by coimmunoprecipitations with or without crosslinking (data not shown) and currently do not know whether the stimulation of Xrn1 activity by DcpS is direct or indirect. The requirement of DcpS decapping activity and the lack of detectable interaction between DcpS and Xrn1 suggest a potential indirect mechanism. Whether accumulation of cap structure, as implicated in the efficiency of first intron splicing (Bail & Kiledjian, 2008; Shen et al, 2008) or possible cap structure byproducts generated in the absence of DcpS (Taverniti & Seraphin, 2014) may influence Xrn1 activity remains to be determined.

The elevation of DRNT2 levels in $DcpS^{KD}$ cells independent of changes in transcript stability strongly suggests a role for DcpS in non-decay regulation. Although this has not been tested directly, as the accumulation of an RNA is dictated by its synthesis and decay, an increase in steady state levels of DRNT2 in the absence of increased stability suggests a potential transcriptional event. Nevertheless, whether this would involve a direct role in transcription or an indirect consequence of modulating the stability of an mRNA or noncoding RNA that in turn impacts DRNT2 RNA levels remains to be determined. Collectively, despite its initial described function in the last step of the mRNA 3´end decay pathway, DcpS is a multifunctional protein involved in a greater network of regulated gene expression.

Of particular note, the two DcpS responsive RNAs were both lncRNAs. Interestingly, DcpS was recently implicated in a quality control mechanism that functions to clear aberrant U1 snRNAs (Shukla & Parker, 2014). However, we have been unable to detect a role for DcpS in the stability of wild type U1 snRNA (M.Z and M.K., unpublished observations), further underscoring the transcript specificity of DcpS mediated regulation of RNA stability. The mechanism underlying how DcpS can modulate the stability of specific substrates awaits further examination.

Chapter II Mutations of DcpS in Autosomal Recessive Intellectual Disability Indicate a Crucial Role for DcpS Decapping in Neurodevelopment

Summary

We report on the identification of mutations in DcpS involved in the mRNA decapping pathways in patients with intellectual disability. In a large family with multiple branches we identified a splice-site variant that results in the insertion of 15 amino acids (c.636+1G>A), and a missense variant (c.947C>T; p.Thr316Met) in DCPS in 3 individuals with intellectual disability with or without myopathy. DCPS hydrolyzes the resulting cap structure following decay of mRNA in the 3′ to 5′ mRNA degradation pathway. In vitro decapping assays showed an ablation of decapping function for both DcpS variants. Together, our finds establish a role for the scavenger decapping activity of DcpS, which is functional in the 3′-end mRNA decay pathway, in neuropathological disorders. Moreover, these studies further affirm an emerging theme linking proper mRNA metabolism to neurological defects.

Introduction

Intellectual disability, more commonly referred to as mental retardation, is a neurodevelopmental disorder that is characterized by impairment in conceptual, practical, social skills and an intelligence quotient (IQ) below 70 before the age of eighteen. For most Intellectual disability cases, the underlying etiology is genetic in origin. It is

reported that chromosomal aberration account for 15% of Intellectual disability cases, while X-chromosomal mutations are the cause for approximately 10% of intellectual disability cases (Ropers & Hamel, 2005). The rest of more than half of the cases the etiology of intellectual disability remain unsolved and are likely caused by autosomal chromosome mutations (Ropers & Hamel, 2005). The frequency of autosomal dominant forms of intellectual disability is rare because most of the affected individuals do not reproduce. On the other hand, the molecular explication of autosomal recessive intellectual disability is lagging behind because most of the investigation is hampered by small family sizes and infrequent parental consanguinity (Najmabadi et al, 2007). As a result, very limited information is currently available concerning the role of autosomal recessive intellectual disability.

DcpS is an essential enzyme that contributes to multiple steps of RNA regulation. In the 3′ to 5′ pathway mRNA decay pathway, the m^7GpppN cap structure that is generated through degradation of the mRNA by the multi-subunit exosome complex, is hydrolyzed by DcpS to release m^7Gp and ppN products (Anderson & Parker, 1998; Liu et al, 2002; Liu et al, 2006; Wang & Kiledjian, 2001). DcpS is also involved in the 5′ -3′ mRNA decay pathway through its mediation of the Xrn1 5′-3′ exonuclease which degrades the 5′ end monophosphorylated RNA after 5′ decapping (Bosse et al, 2013; Liu & Kiledjian, 2005; Sinturel et al, 2012). In addition, as a modulator of cap structure in cells, reduction of DcpS protein levels results in accumulation of cap structure, which in turn sequesters the cap binding complex (CBC) and leads to a corresponding decrease in proximal intron splicing (Shen et al, 2008). Thus, DcpS is expected to have numerous

effects on the expression levels of various pathways and disturb the cellular balance and function by regulating mRNA processing and degradation rates.

In this study, we report on mutations in DcpS in a consanguineous Pakistani family with autosomal recessive intellectual disability. We identified two different mutations in the DcpS gene (c.636+1G>A and c.947C>T), both of which lead to the loss of decapping activity. Our data indicate that loss of DcpS decapping function was responsible for the intellectual disability in patients and suggests that the decapping activity of DcpS is critical for normal neurodevelopment.

Results

DcpS mutations were identified in individuals with intellectual disability.

This four-generation pedigree of a consanguineous Pakistani family consisted of three affected (III-2, IV-3 and IV-7) individuals with intellectual disability including two females (IV-3 and IV-7) and one male (III-2) (Figure 9A). For all three affected individuals, motor development milestones were achieved late or not achieved at all and all have Intellectual Disability (ID), ranging from mild to severe (IQ level ranges 25-55). There was no previous history of ID in this family, and all the affected individuals were born from normal parents and therefore the mode of inheritance is therefore autosomal recessive.

In work carried out by our collaborators, two affected (III-2 and IV-3) and one normal individual (III-3) were selected for 500K SNP (NspI) genotyping (Affymetrix, USA). A single candidate region of 11.3 MB on chromosome 11 that is shared by

individuals III-2 and IV-3 was identified. Mutation analysis of the genes of 11q25.1-q25 locus with NGS data identified a splice donor site mutation in the DcpS gene at exon 4-intron 5 junction where c.636+1G>A (Figure 9B). Sequencing of the DcpS gene in individual IV-7 revealed her to be a compound heterozygote with one allele containing the c.636+1G>A mutation and a second *de novo* mutation in exon 6 of the second allele.

To determine the variant DcpS transcripts in patients, gene specific primers were used to amplify the DcpS cDNA from lymphocytes derived from individuals III-2 and IV-3. Sequence analysis and alignment revealed that an addition of 45 nucleotides were present in the DcpS transcript between exon4 and exon5, caused by the splice donor site mutation c.636+1G>A (Figure 9C). In silico translation of this mutant mRNA revealed the incorporation of 15 amino acids (IKVSGWHVLISGHPA) between amino acids 212-213 of the DcpS protein which was a consequence of an in frame 45 nucleotide insertion. The second point mutation c.947C>T is predicted to change threonine 316 to methionine (Thr316Met).

DcpS mutant proteins lose decapping activity

In order to gain further information on potential pathogenic effects of the identified variants, our collaborators performed molecular modeling using Modeler 9.9 (Sanchez & Sali, 2000). In the altered DcpS protein carrying a 15-residue insertion ($DcpS^{Ins15}$), the loop connecting the amino acids lysine (p.Lys207) and tyrosine (p.Tyr217) that build the active decapping site (Gu et al, 2004) is extended (Figure 10A

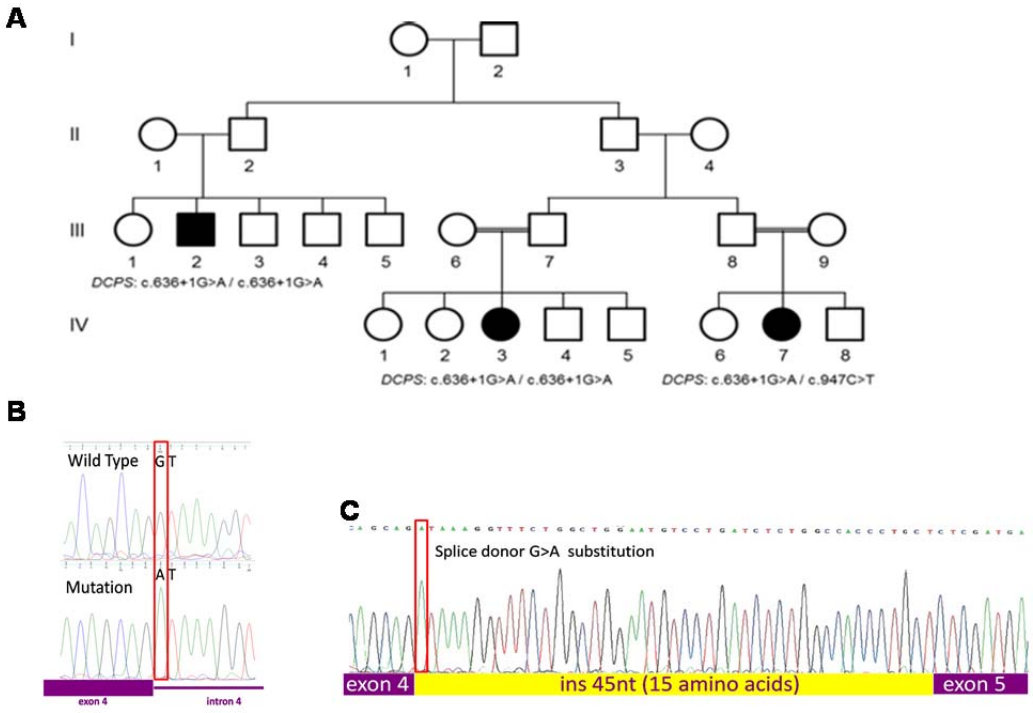

Figure 9. DcpS mutation was identified in a family with intellectual disability.

(A) Pedigree of Pakistani family. Individuals indicated with black shaded symbols are affected with apparent NS-ARID. **(B)** Sanger sequencing confirmation of the splice donor variant in DCPS, for wild type (unrelated control) and mutant (III-2) homozygotes. **(C).** RT-PCR followed by Sanger sequencing using lymphoblast cDNA from III-2 was used to confirm a change in splice donor usage at exon 4/intron 4 for the splice variant allele. For this variant we observed correct splicing between a cryptic splice site in intron 4 and the intron 4/exon 5 splice acceptor. The resulting mRNA carries an additional 45 nucleotides, and encodes an additional 15 amino acids. (Data provided by Dr Iltaf Ahmed)

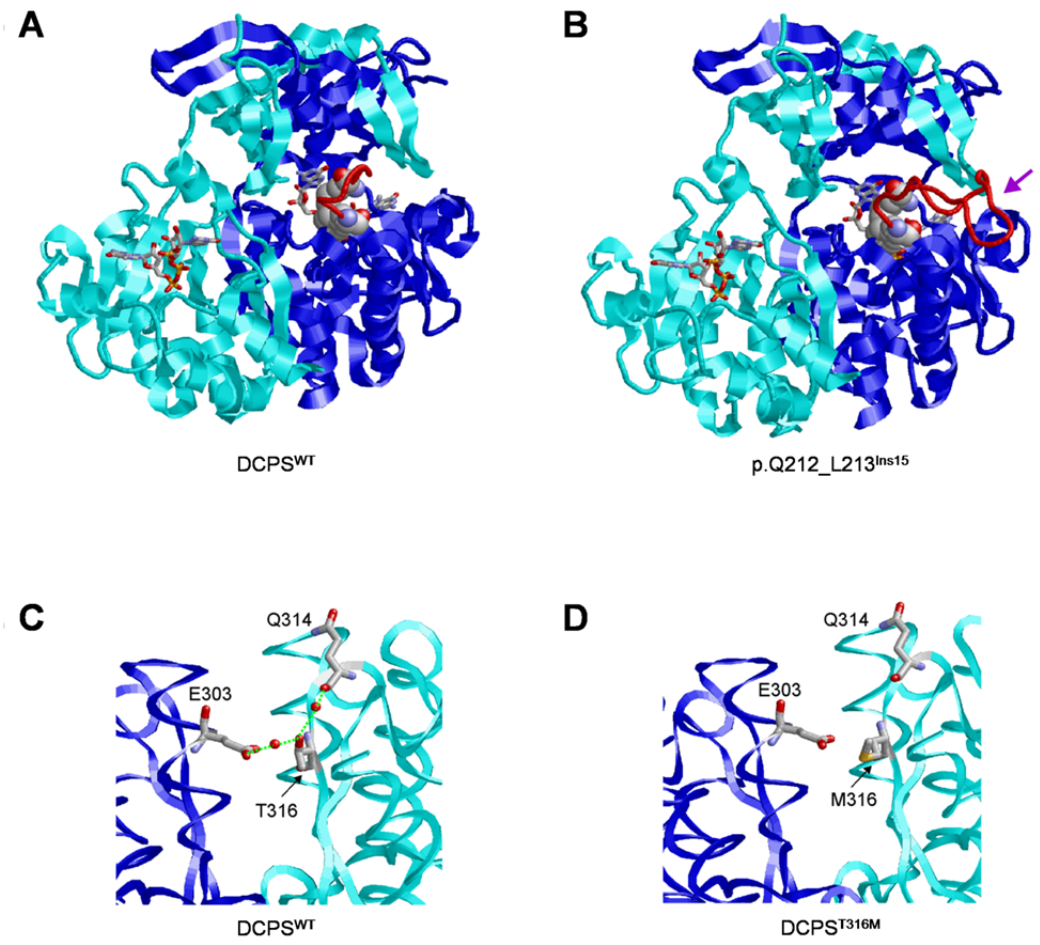

Figure 10. Molecular modeling of DcpS.

(A) Structure of dimeric DcpS bound to m7GpppG. The subunits are shown in cyan and blue and the ligand is shown in stick presentation and colored according to the atom type. Residues p.Lys207 and p.Tyr217 that are important for decapping activity are shown in space-filled presentation and their connection loop is highlighted in red. **(B)** In the mutant, which carries a 15-residue insertion in the loop connecting the active site lysine and tyrosine (loop shown in red; region of insertion marked by an arrow), the geometry of the active site becomes distorted and the long loop additionally interferes at the dimer interface. Therefore, the insertion is expected to reduce dimer stability and enzymatic activity. (Data provided by Dr Iltaf Ahmed)

and B). As a consequence, the active site is predicted to become distorted and consequently impact enzymatic activity. Molecular modeling of the homodimeric DcpS also predicted that substitution of methionine in place of threonine at position 316 would be detrimental to DcpS protein integrity. Thr316 normally forms a stabilizing water-mediated interaction with the glutamine at position p.314 and glutamate at position p.303 (Figure 10C and D). It is conceivable that this mutation could disrupt the stability of the DcpS homodimer formation and consequently decapping activity.

In order to test our hypothesis that these mutations would alter DcpS decapping activity, DcpS WT and mutant genes were cloned into the bacterial expression vector PET-28a and expressed in bacterial BL21 (DE3) cells (Figure 11A). Electrophoretic mobility shift assay revealed that $DcpS^{Ins15}$ protein failed to bind to cap structure due to the possible distorted active site (Figure 11B). Enzymatic activity of both recombinant mutant DcpS proteins was next analyzed. Significantly reduced decapping activity was detected by both mutants (the 15 amino acid insertion and Thr316Met), while robust decapping was detected by wild type DcpS. The wild type DcpS hydrolyzed almost all of the m7GpppG methylated cap structure to methylated guanosine monophosphate (m7Gp) at the lowest concentration (10ng) within the 5 minutes time period. The mutant with 15 amino acid insertion hydrolyzed only 3% at the highest concentration (40ng), whereas, 15% activity was shown by T316M mutant at its highest concentration (40ng) (Figure 11C). In summary, both DcpS mutant proteins were defective in scavenger decapping activity in vitro.

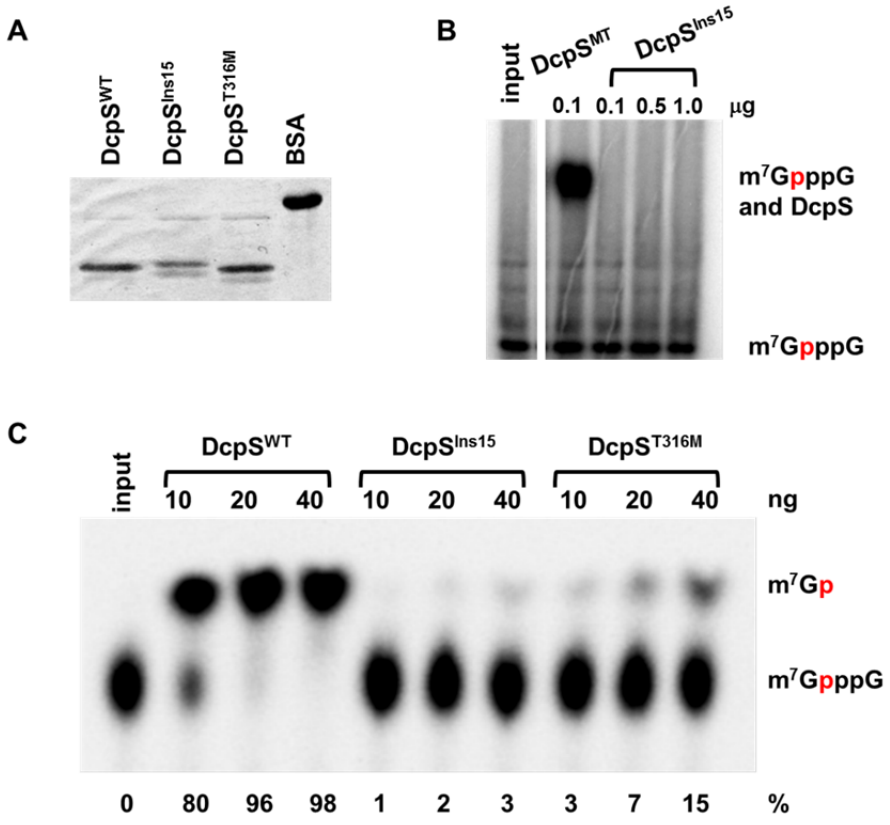

Figure 11. Mutant DcpS proteins did not bind to cap structure and lost decapping activity.

(A) Purified his-tag recombinant DcpS^{WT}, DcpS^{Ins15} and DcpS^{T316M} in pET28a (Novagen, EMD Biosciences, Danvers MA) vector expressed and purified from E. coli BL21(DE3) cells was examined on SDS-PAGE gel and stained with Coomassie Blue. **(B)** The affinity of DcpS^{Ins15} and cap was tested by electrophoretic mobility shift assay. Increasing amount of recombinant DcpS^{Ins15} and 32P- labeled cap structure were incubated to form protein-cap complexes. Recombinant catalytic inactive DcpS^{MT} protein was used as a positive control. (C) in vitro decapping assay of mutant DcpS proteins. Decapping was assayed using 10, 20, 30ng recombinant full length DcpS^{WT} or altered (DcpS^{Ins15} and DcpS^{T316M}) DcpS protein. Migration of the input cap structure (m^{7}GpppG) or m^{7}GMP product (m^{7}Gp) is designated, with the red "p" denoting the location of the ^{32}P. The percent decapping is indicated at the bottom of the panel.

Patient cell exhibits defect of DcpS decapping function.

DcpS activity was further investigated in whole cell extract obtained from Epstein–Barr virus transformed lymphocyte cells from two affected (III-2 and IV-3) and two heterozygous normal siblings (III-3 and IV-2). Cell extract was incubated at 37˚C with ^{32}P-labeled cap structure (32-P m7Gp*ppG) at the time intervals of 10, 20 and 30 minutes, respectively. DcpS enzymatic activity in cell extract of patients with the 15 amino acid insertion (III-2 and IV-3) were significantly reduced compared to heterozygous carriers (III-3 and IV-2) and normal control (rDcpS) (Figure 12). The mutant DcpS hydrolyzed 6-9% of the input m7GpppG methylated cap structure to m7Gp in 30 minutes whereas, DcpS of heterozygous carrier (III-3 and IV-2) and normal displayed 82-87% activity in the same time frame (Figure 12).

Mutant DcpS shows high mRNA level but low protein level

Using the lymphoblast cells from affected patients (III-2 and IV-3) and one unaffected carrier (III-3), we performed a series of cellular assays to characterize the c.636+1G>A DcpS mutation (DcpS^{Ins15}). To identify what are the DcpS transcripts present in the patients' cells, a primer set designed to distinguish between the DcpS^{WT} and DcpS^{Ins15} was used (Figure 13A). Regular RT-PCR performed on cDNA derived from homozygous patients (III-2 and IV-3) and heterozygous individual (III-3) cells revealed the presence of the major wild type isoform only in control and heterozygous cells but not in homozygous mutant cells (Figure 13A). Likewise, qRT Real Time PCR using primers detecting both DcpS^{WT} and DcpS^{Ins15} confirmed the similar total mRNA level of endogenous DcpS in both patient cells relative to control and heterozygous

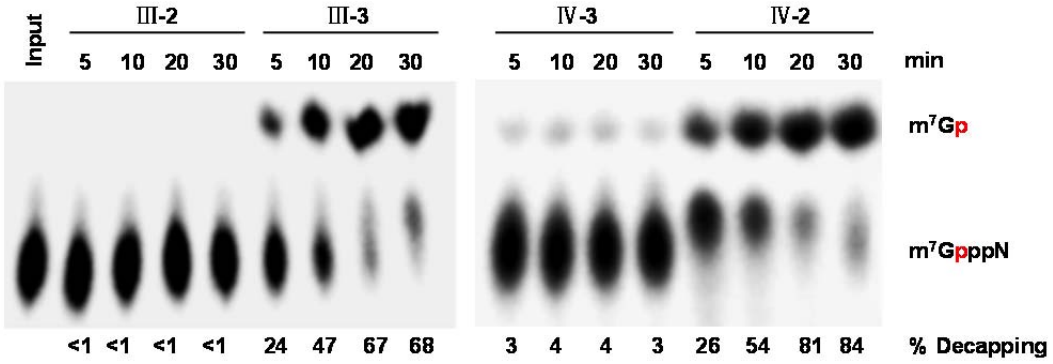

Figure 12. DcpS activity in lymphoblastoid cell lines.

Lymphoblast cell lines from homozygous affected (III-2 and IV-3) and heterozygous unaffected (III-3 and IV-2) family members were used for decapping assays. Decapping products and labeling are as in Figure 11

individual cells (Figure 13B). Using specific primers for $DcpS^{WT}$ also did not detect any $DcpS^{WT}$ mRNA generated in patient cells (Figure 13B) These data demonstrate the c.636+1G>A mutation results in an alteration of splicing that leads to exclusive generation of variant splice pattern in homozygous individuals while heterozygous individuals express both the WT and variant splice patterns.

We next assessed the consequences of the synthesis of the DcpS protein in patient cells. Consistent with the mRNA data, western blot analysis with an anti-DCPS antibody detected $DcpS^{WT}$ in control and heterozygous cells but not in homozygous mutant cells (Figure 13C). A slightly larger band of DcpS likely representing the 15 amino-acid in-frame insertion translated from mutation was detected at lower levels in both heterozygous cells and homozygous mutant cells but not in control cells (Figure 13C). These results suggest that the discovered splice site mutation completely abolished the default splicing and utilize the alternative splicing site 45 nt downstream instead. Furthermore, these data indicate that despite comparable mRNA levels for $DcpS^{Ins15}$ and $DcpS^{WT}$, considerably lower levels of $DcpS^{Ins15}$ protein are present in both homozygous and heterozygous cells.

$DcpS^{Ins15}$ localized in nucleus and hetero-dimerized with wild type DcpS

DcpS is a shuttling protein that is predominantly nuclear (Shen et al, 2008). We next determined whether $DcpS^{Ins15}$ is correctly localized in cells. Since the nuclear localization sequence (NLS) in DcpS is at the N- terminus of the protein and not disrupted by the c.636+1G>A mutation, it was anticipated to localize normally. We

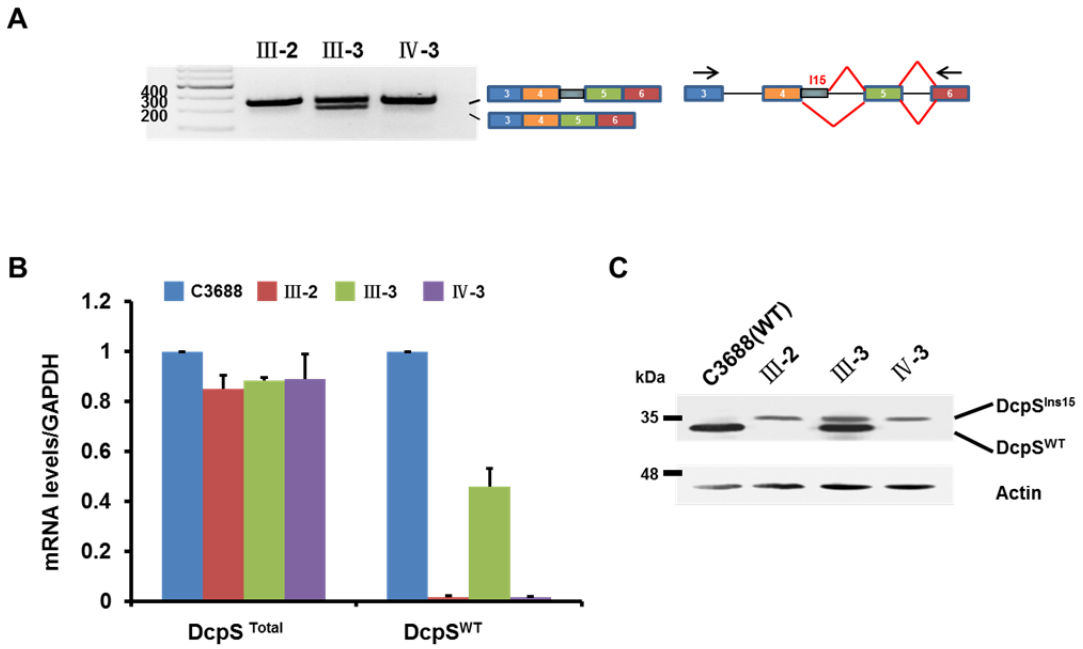

Figure 13. Characterize DcpS^{Ins15} mRNA and protein in patient cells.

(A) RT–PCR analysis for DCPS mRNA. RT–PCR across the DcpS splice site mutation (from exons 3 to 6) was performed using mRNA from patients (III-2 and IV-3) and an unaffected family member (III-3), and a schematic indicating the corresponding exon usage. **(B)** Total DcpS mRNA levels and DcpS^{WT} mRNA levels in patients (III-2 and IV-3) and unaffected family member (III-3) were tested by real time PCR using specific primers. Wild type lymphoblast cell line C3688 was used as a control and set to 1. RNA levels are presented relative to the GAPDH mRNA and derived from three independent experiments. **(C)** Western Blot of DcpS. Wild-type and variant DcpS proteins from the indicated lymphoblast cell lines were detected by western blotting. A wild-type lymphoblastoid cell line is shown to indicate the position of wild-type DcpS. Actin was used as loading control.

confirmed this by exogenous expression of an epitope tagged $DcpS^{Ins15}$ protein in GM03813 which are SMA patient fibroblast cells that express normal levels of DcpS. As shown in Figure 14A, after infecting lentiviral particles harboring an expression system for Flag-$DcpS^{Ins15}$, the majority of $DcpS^{Ins15}$ protein was detected in the nucleus with an anti-Flag tag antibody as expected (Figure 14A). Infection with viral particles lacking an exogenous expression cassette was used as the control and no Flag signal was detected (Figure 14A).

To test the possibility that the 15 amino-acid-insertion may interfere with the dimerization interface, we next determined whether a heterodimer between $DcpS^{Ins15}$ and $DcpS^{WT}$ could form. HEK293T cells stably transformed with Flag-$DcpS^{Ins15}$ cDNA were generated and $DcpS^{Ins15}$ was immunopurified on a Flag affinity column (Figure 14B). Western blot analyses with the α-DcpS antibody revealed coimmunoprecipitation of wild type DcpS in the fractions with Flag-$DcpS^{Ins15}$ in a dose-dependent manner (Figure 14B) demonstrating that DcpS and $DcpS^{Ins15}$ can heterodimerize in cells. Since a heterodimer of wild type DcpS with a catalytically inactive DcpS can generate partially active protein with one active site (Liu et al, 2008b), the $DcpS^{Ins15}$ containing heterodimer may contain partial activity.

Reduced levels of $DcpS^{Ins15}$ protein in cells is not a consequence of protein degradation or microRNA mediated silencing

Closer examination of figure 13 revealed a disconnection between the level of $DcpS^{Ins15}$ mRNA and protein. Although mutant $DcpS^{Ins15}$ mRNA levels in the ID patient

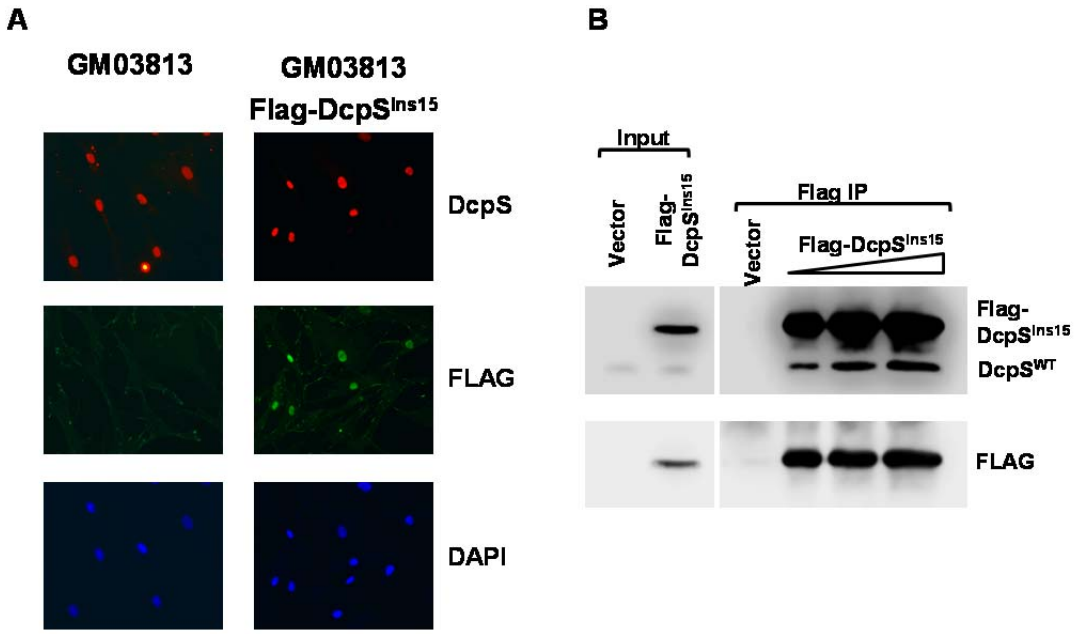

Figure 14. DcpS^{Ins15} localized in nucleus and was able to dimerize with wild type DcpS.

(A) GM03813 human fibroblast cells were infected with lentivirus which expressed Flag-DcpS^{Ins15} or empty vector. The red staining by DcpS antibody denotes localization of endogenous DcpS and the green staining by Flag antibody denoted the localization of DcpS^{Ins15}. The nuclei are designated by DAPI staining. **(B)** Anti-Flag immunoprecipitation was performed in HEK293T cell lines constitutively express Flag-tagged DcpS^{Ins15}. Western Blot using DcpS antibody indicated the DcpS^{WT} was co-purified with Flag-DcpS^{Ins15}. The lower band in anti-DcpS blot was not degradation product of Flag-DcpS^{Ins15} as it was not detected in anti-Flag blot. Whole cell extract was also included to designate the position of DcpS^{WT} and DcpS^{Ins15}.

cells were similar to DcpS levels in wild type lymphoblast cells (Figure 13B), $DcpS^{Ins15}$ protein levels were considerably lower (~25%) than wild type DcpS protein (Figure 13C). Our initial hypothesis was that the altered structure of $DcpS^{Ins15}$ resulting from the 15 amino acid insertion subjected the protein to rapid proteolysis. To test this hypothesis, $DcpS^{Ins15}$ protein levels were tested by western blot in patient lymphoblast cells treating with the proteasome inhibitor, MG132 to determine whether $DcpS^{Ins15}$ protein would accumulate in treated cells. However, $DcpS^{Ins15}$ didn't accumulate after 12 hours MG132 treatment at either 7.5 or 15 μM concentration (Figure 15A) while a known proteasome target protein, SMN, increased in MG132 treated samples (Figure 15A).

We next tested if $DcpS^{Ins15}$ was degraded by the lysosomal protein degradation pathway. Patient cells were cultured with chloroquine at 50μM overnight. Chloroquine neutralizes the pH in the lysosome and inhibits protein degradation in this pathway (Shintani & Klionsky, 2004). Detection of DcpS protein from chloroquine treated cells revealed that $DcpS^{Ins15}$ protein levels were not responsive to this treatment and did not appreciably accumulate. In contrast, level of TRAF3 protein, which is subjected to lysosomal degradation, increased indicating the chloroquine treatment conditions were sufficient to inhibit lysosomal degradation (Figure 15B). Thus, $DcpS^{Ins15}$ was not degraded by either of the two main protein degradation proteasome or lysosome pathways.

Since the low protein level of $DcpS^{Ins15}$ did not appear to be rapid protein degradation, defect in translation could be another explanation. We first hypothesized that the extra 45 nucleotides in $DcpS^{Ins15}$ could generate a novel microRNA target sites that could repress protein synthesis. To test this, we utilized an inducible Dicer

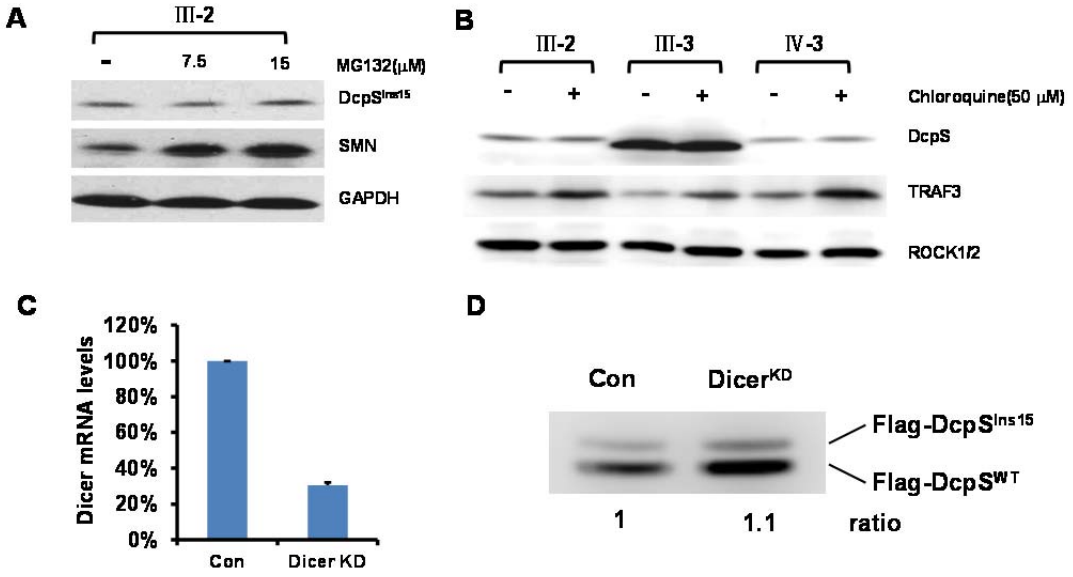

Figure 15. Low protein level of DcpS^{Ins15} was not caused by proteasomal/lysosomal protein degradation or microRNA directed translational inhibition.

(A) DcpS^{Ins15} protein levels in homozygote mutant patient III-2 were measured by Western Blot with treatment of proteasome inhibitor MG132 at 0, 7.5 and 15 μM for 12 hours. SMN protein level was detected as positive control for proteasome inhibition. GAPDH was used as loading control. **(B)** DcpS^{Ins15} protein levels in homozygote mutant patient III-2 and IV-3, and one heterozygote carrier (III-3) were measured by Western Blot following treatment with the lysosome inhibitor Chloroquine at 50μM overnight. TRAF3 protein levels were detected as a positive control and ROCK1/2 was used as a loading control. **(C)** Dicer mRNA levels were determined by qPCR in control and Dicer shRNA induced HEK293 cells. RNA levels are presented relative to the GAPDH mRNA and derived from at least two independent experiments. **(D)** Control and Dicer depleted HEK293 cells were co-transfected with constructs expressing Flag-DcpS^{Ins15} and Flag-DcpS^{WT}. Protein levels of DcpS^{Ins15} and DcpS^{WT} were determined by Western Blot. The ratio of DcpS^{Ins15} over DcpS^{WT} was indicated below.

knockdown 293 cell line (Schmitter et al, 2006). Dicer is an essential factor for microRNA expression and its depletion should ablate microRNA-mediated translational silencing. Flag-DcpS^{Ins15} plasmid was co-transfected with Flag-DcpS^{WT} into doxycycline induced cells and DcpS protein levels were determined by western blot 2 days after transfection. Dicer level was about 30% in the induced cells compared to non-treated cells as determined by real- time RT-PCR (Figure 15C). However DcpS^{Ins15} protein levels failed to accumulate in Dicer reduced cells after normalizing to wild type DcpS (Figure 15D). In summary, the reduced DcpS^{Ins15} protein level does not appear to be due to proteasomal/lysosomal degradation or microRNA-mediated translational repression. Possible translation repression mechanisms other than microRNA-mediated remain to be tested.

Discussion

In this chapter, we identified mutant DcpS genes in 3 individuals with intellectual disability from a large family with multiple branches. We identified a splice-site variant that results in the insertion of 15 amino acids (c.636+1G>A) and a missense variant (c.947C>T; p.Thr316Met) in the DcpS gene (Figure 9). *In silico* parameters and molecular modeling suggested deleterious effects of both variants (Figure 10) and functional analyses validated both mutations lead to abrogation of DcpS decapping function in vitro and in cells (Figure 11C and Figure 12). Using the lymphoblast cell lines generated from affected patients, we characterized the mutant DcpS with insertion of 15 amino acids (DcpS^{Ins15}) and analyzed its localization and capacity to form dimer with wild type DcpS (Figure 14). It is interesting that the transcription of this mutant DcpS

appeared normal and mRNA level of this isoform were similar to wild type DcpS in heterozygous cells (Figure 13B), while $DcpS^{Ins15}$ protein levels were significantly lower relative to wild type DcpS (Figure 13C). This phenomenon was not due to protein degradation by either proteasome or lysosome (Figure 15A and B), or microRNA induced translational repression (Figure 15D). Collectively, our data demonstrate that 1) homozygous mutations in DcpS that abolish its decapping function lead to intellectual disability in affected individuals; and 2) the insertion of 15 amino acids in DcpS results in reduced protein levels through a yet-to-be determined mechanism.

Although at present we only have identified disruptions in DcpS in one family, the presence of compound heterozygous variants in DcpS in an affected person supports a role for DcpS as the causative agent for intellectual disability in these individuals. Moreover, the identification of a second family, of Jordanian descent, with intellectual disability also caused by a homozygous mutation in DcpS leading to an in frame 7 amino acid insertion at amino acid K67 and nonfunctional variant protein (Ng et al, 2015b), further reinforces the significance of DcpS for neuronal function. The presence of obvious physical anomalies and neuromuscular phenotype in the Jordanian family suggests that a relatively wide clinical spectrum can be anticipated for different types and locations of mutations in DcpS.

In cells, a disruption of DcpS decapping activity is expected to lead to the accumulation of m7GpppN cap structure, which can sequester the nuclear cap binding protein and decrease the efficiency of first intron splicing (Shen et al, 2008) and potentially alter cytoplasmic mRNA translation by sequestering the cytoplasmic cap binding protein (Bail & Kiledjian, 2008). In addition, at least in *S. cerevisiae* and *C.*

elegans, DcpS orthologues contribute to Xrn1-mediated exonucleolytic decay (Bosse et al, 2013; Liu & Kiledjian, 2005; Sinturel et al, 2012). In human cells, DcpS also demonstrates transcript specific regulation of RNA stability coupled with Xrn1 as well as RNA decay independent regulation of RNA steady state levels (Zhou et al, 2015). Genomic mRNA and proteomic analysis of cells homozygous c.636+1G>A mutation that exclusively generate $DcpS^{Ins15}$ will enable an analysis of the underlying molecular defect attributed to expression of this variant protein.

The mutant $DcpS^{Ins15}$ exhibited much lower protein levels compared to wild type DcpS, despite similar mRNA levels of the two genes. This suggested the mutant DcpS had lower translational efficiency or faster protein turnover. Inhibition of the two primary protein degradation pathways mediated through the proteasome and lysosome, did not increase mutant DcpS protein levels, indicating the low level of this protein was not caused by impeded protein degradation. Alternatively, the insertion of 45nt could create sites for microRNA or some RNA-binding proteins, which could in turn slow down the progression of ribosomes and reduce translational efficiency. $DcpS^{Ins15}$ protein levels were unchanged in Dicer depleted cells, suggesting $DcpS^{Ins15}$ was not a microRNA target. Further work is needed to test whether the 45nt insertion recruited some protein factors to repress translation.

In an effort to generate a *DcpS* null mouse, we found that *DcpS* is an essential gene and its deletion is embryonic lethal (M. Kiledjian: unpublished observations), demonstrating a role for DcpS protein in embryogenesis. CRISPR screens of human genes also indicate the requirement of DcpS for cellular viability (Shalem et al, 2014), underlining the importance of this protein in fundamental cellular functions. The viability

of individuals with a decapping-compromised DcpS protein suggests that the decapping activity of DcpS is important for normal cognition, while an additional decapping-independent function of DcpS must also be required for normal development and cell survival. This additional function is yet to be determined. Further, we consider it unlikely that fully null homozygous mutations in DcpS will be identified in a clinical context, and that bi-allelic mutations would always be hypomorphic or with restricted effect on specific isoforms or functions of the protein.

In conclusion, we report the investigation of a family with apparent non-specific autosomal recessive intellectual disability leading to the identification of mutations in DcpS. At present the molecular consequence of defects in DcpS is unknown, yet it leads to a physiological consequence of adverse overall cognition in individuals harboring the mutant alleles. Future identification of the downstream target mRNAs or proteins altered in the cells of affected individuals will provide insights into the contribution of DcpS in gene expression and cognitive function.

Chapter III DcpS Insertion Variant Elevates Cellular SMN2 RNA and Protein Levels

Summary

Spinal Muscular Atrophy (SMA) is a common autosomal recessive disorder that results in progressive loss of spinal anterior horn motor neurons. It is a consequence of low levels of survival motor neuron (SMN) protein caused by the loss or mutation of the survival motor neuron 1 gene (SMN1). A duplicate gene, SMN2, usually remains intact and relieves the severity of SMA, though the majority of mRNA generated from SMN2 lack exon 7 due to inappropriate splicing resulting from a single nucleotide substitution of a C->T inside exon 7. Previously it was reported that a quinazoline derivative, RG3039, increased survival and function in a mouse model of SMA through inhibition of the decapping scavenger enzyme, DcpS. We now have evidence that a variant form of DcpS that utilizes a cryptic splice site initially identified as a mutation in non-syndromic autosomal recessive intellectual disability (ID) can significantly elevate SMN2 steady state mRNA and SMN protein in SMA patient cells. Most significantly, utilization of the cryptic splice site 45 nucleotides downstream of the endogenous splice site can be promoted by oligonucleotide directed splice switching technology. We designed and tested four oligonucleotides in an effort to alter the endogenous DcpS gene splicing pattern to that of the $DcpS^{Ins15}$. Shifting of the DcpS splicing pattern to the $DcpS^{I15}$ variant could be a potential therapeutic avenue to elevate full length SMN2 mRNA and SMN protein in SMA patient cells.

Introduction

Spinal Muscular Atrophy (SMA) is a common autosomal recessive disorder that results in progressive loss of spinal anterior horn motor neurons, lethal in its severe form, with no current cure. SMA is caused by reduced levels of SMN protein, the consequence of homozygous loss or mutation in the survival motor neuron 1 gene (SMN1) (Lefebvre et al, 1995). A second, nearly identical gene to SMN1, termed SMN2, was found in the human genome. It shares the same promoter and differs by only two-nucleotides in the coding region relative to SMN1 (Monani et al, 1999a; Monani et al, 1999b). The primarily difference is a single nucleotide substitution of a C→T inside exon 7 in SMN2 that disrupts splicing of exon 7 leading predominantly to exon 7 exclusion and the generation of nonfunctional truncated protein (Lorson et al., 1999; Monani et al., 1999). However, due to a low level of exon 7 inclusion from the SMN2 gene (Gavrilov et al., 1998; Lorson et al., 1999; Monani et al., 1999), an increase of SMN2 mRNA levels can lead to dosage dependent compensation of SMN1 lose and a decrease of SMA severity in humans (Chen et al., 1999; McAndrew et al., 1997; Wirth et al., 1999) as well as mice model systems (Hsieh-Li et al., 2000; Monani et al., 2000).

The primary strategies of current SMA therapeutics are focused on the importance of increasing SMN levels from three different approaches: 1) to increase SMN protein levels; 2), to increase SMN2 transcription levels; and 3) to promote the inclusion of exon 7 in SMN2 transcripts. Based on the first approach to express exogenous SMN, the feasibility of viral-based delivery of the SMN-coding region to restore normal SMN expression in SMA patient-derived cells was first demonstrated using adenoviral vectors (DiDonato et al, 2003). Later, a lentiviral vector based system was shown to retrogradely

increase SMN expression in motor neurons in an SMA mouse model (Azzouz et al, 2004). Recently, an adeno-associated viral vector (AAV9) based method, which can infect peripheral tissue as well as cross the blood–brain barrier, showed improvements in gene delivery efficiency and led to the rescue of a severe mouse model of SMA (Dominguez et al, 2011; Foust et al, 2010; Passini et al, 2010).

Major emphasis is also focused on an alternative approach to promote the inclusion of exon 7 in SMN2 transcripts through antisense oligonucleotide or small compound. The antisense oligonucleotide approach targets pre-mRNA processing, reducing exon 7 exclusion in SMN2 mRNA and increasing full-length SMN transcripts and protein levels. Many promising antisense oligonucleotide based methodology to treating SMA has proven successful both in vitro and in vivo using SMA mouse models (Cherry et al, 2013; Hua et al, 2010; Porensky et al, 2012; Rigo et al, 2014; Williams et al, 2009). Besides the antisense oligonucleotide, small molecule compounds that shift the balance of SMN2 splicing toward the production of full-length SMN2 are also showing therapeutic potential (Naryshkin et al, 2014). However, most strategies also have various obvious shortcomings and challenges in patients (Lunn & Stockwell, 2005; Wee et al, 2010). For example, oligonucleotide delivery requires a direct route of administration to the CNS, such as intrathecal or intracerebroventricular injections, and will require additional optimization for an efficient delivery to peripheral organs (Anderton et al, 2013).

A class of C5- substituted quinazoline compounds were identified that increased human SMN2 promoter driven expression of the bacterial β-lactamase reporter in a screen of a mouse motor neuron hybrid cell line NSC-34 (Jarecki et al, 2005). The

initially identified compound was optimized by a focused medicinal chemistry effort to develop a series of modified C5-quinazolines and shown to be an inhibitor of the scavenger decapping enzyme, DcpS (Singh et al, 2008; Thurmond et al, 2008). Although the quinazoline compounds were capable of improving motor neuron function and extending survival of SMA model mice (Butchbach et al, 2010; Gogliotti et al, 2013; Van Meerbeke et al, 2013a), statistically significant increases in SMN2 mRNA or SMN protein were not evident. The lack of detectable increase in SMN mRNA and protein in treated animals confounds the mechanism by which the quinazoline compounds promote survival of SMA mice. To add to this perplexing result, neither compound treatment nor DcpS knockdown in cell lines leads to statistically significant elevated SMN2 mRNA or protein (unpublished) suggesting that the compound likely functions through an SMN independent and possibly DcpS independent pathway. Therefore, the efficacy of current efforts focused on using DcpS inhibitor compounds in patients remains unclear.

In this chapter, we demonstrate that a variant form of DcpS, $DcpS^{Ins15}$, initially identified as a mutation in non-syndromic autosomal recessive intellectual disability (ID) (Chapter2) increases SMN2 mRNA expression. Lymphoblast cells from ID patients harboring a homozygous $DcpS^{Ins15}$ allele contain significant elevation of SMN2 mRNA but not in heterozygous siblings with normal intellect. More significantly, exogenous expression of $DcpS^{Ins15}$ into SMA patient fibroblasts lead to elevated SMN2 mRNA and SMN protein. The SMN2 increase is not observed when the cells are complemented with either wild type or a catalytically inactive DcpS. Of particularly high significance for potential SMA therapeutics, the DcpS pre-mRNA is amendable to antisense

oligonucleotide directed splice switching to promote generation of the DcpS^{Ins15} variant mRNA in SMA patient cells in a strategy to ameliorate the SMA phenotype in patients.

Results

SMN2 mRNA levels are elevated by expression of variant DcpS

The requirement of DcpS for normal neurological function (Ahmed et al, 2015b; Ng et al, 2015b) and the potential link of DcpS to SMN2 expression (Singh et al, 2008), prompted us to test whether SMN2 levels were altered in DcpS^{Ins15} homozygous cells. Quantitative reverse transcription (qRT)-PCR analysis of RNA derived from the homozygous patients with DcpS^{Ins15} mutation revealed a selective elevation of SMN2 in these cells. A 2 fold increase of full length SMN2 mRNA was detected in both DcpS^{Ins15} homozygous cell lines, but not in a heterozygous line relative to wild type lymphoblast cells (Figure 16A).

To determine whether the increase of full length SMN2 mRNA was due to loss of DcpS decapping activity, a 293T-DcpS^{KD} cell line which constitutively expresses DcpS-specific shRNA to reduce endogenous DcpS levels, was transfected with shRNA-resistant catalytically inactive mutant of DcpS (DcpS^{MT}), or DcpS^{Ins15}, and stable cell populations were selected. Similar to the DcpS^{Ins15} patient cells, expression of the DcpS^{Ins15} variant in DcpS^{KD} cells resulted in a >2 fold increase in full length SMN2 mRNA compared to control transfected cells (Figure 16B). Moreover, it was evident that the increase was not a result of a defect in DcpS decapping since expression of the DcpS^{MT} catalytically inactive mutant had no effect on SMN2 expression.

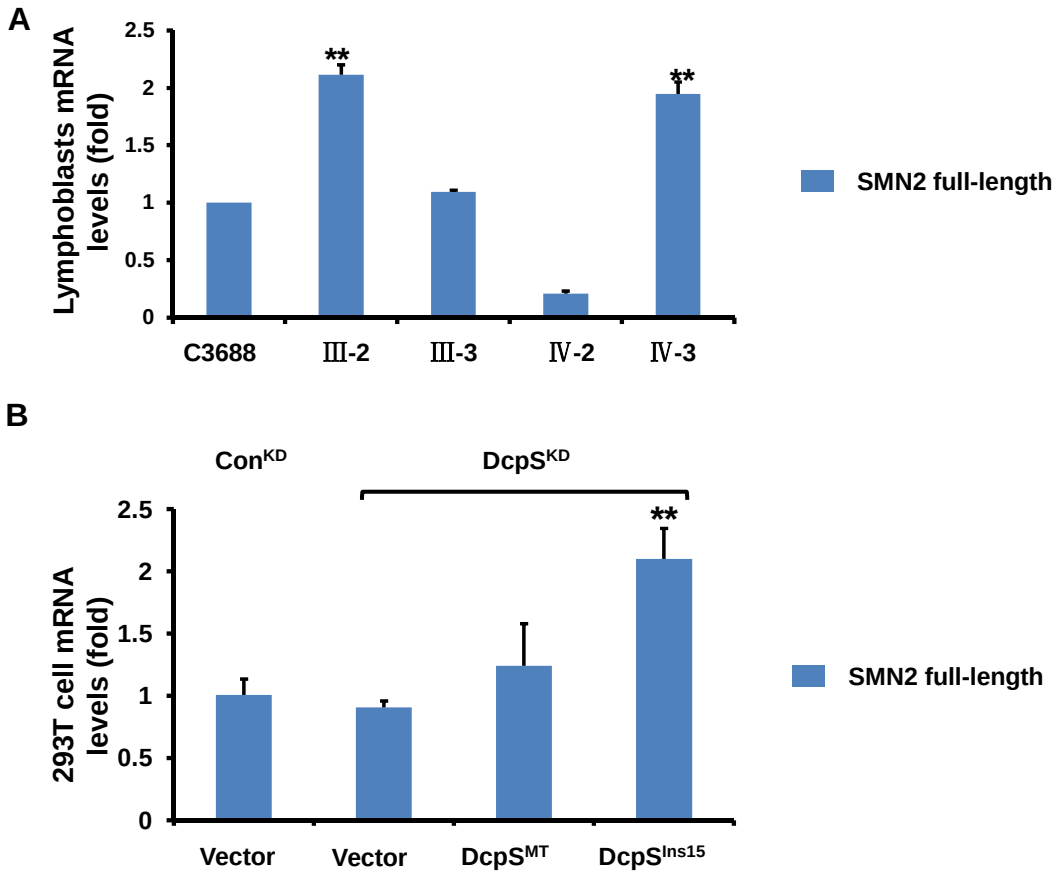

Figure 16. DcpS^{Ins15} correlated with increasing full length SMN2 mRNA.

(A) Full length SMN2 mRNA levels were detected by Real Time PCR in patients (III-2 and IV-3) and unaffected family member (III-3 and IV-2) as well as in WT lymphoblast cell line C3688, and normalized to C3688. **(B)** Full length SMN2 mRNA levels were detected by Real Time PCR in HEK293T control knock down (Con^{KD}) or DcpS knock down (DcpS^{KD}) cells, complemented with shRNA-resistant catalytically inactive DcpS^{MT}, or DcpS^{Ins15}. RNA levels are presented relative to the GAPDH mRNA and derived from three independent experiments. Error bars represent +/- SD. P values are denoted by asterisks; * represents $p < 0.05$, ** represents $p < 0.01$ (Student's *t* test).

These data demonstrate that the detected elevation of full length SMN2 in both the homozygous patient cells and in 293T cells is a consequence of $DcpS^{Ins15}$ expression. In addition, the enhanced level of elevation of SMN2 RNA observed with expression of $DcpS^{Ins15}$ compared to catalytic inactive DcpS indicates that $DcpS^{Ins15}$ is likely a gain of function variant coordinating an elevation of SMN2 in a manner independent of its decapping activity. Furthermore, the lack of detectable SMN2 mRNA increase in heterozygous lymphoblast cells indicated there may be a threshold level of $DcpS^{Ins15}$ variant that is necessary to increase SMN2 levels.

SMN RNA and protein levels were increased by $DcpS^{Ins15}$ in SMA patient cells

To determine whether expression of $DcpS^{Ins15}$ can elevate full length SMN2 mRNA in a clinically more relevant SMA type I patient-derived fibroblast cells (GM03813), we engineered a constitutive knockdown of endogenous DcpS with DcpS-specific shRNA. These cells were transfected with shRNA-resistant wild type DcpS ($DcpS^{WT}$), a catalytically inactive mutant of DcpS ($DcpS^{MT}$), or $DcpS^{Ins15}$. Similar to the 293T cells, expression of the $DcpS^{Ins15}$ variant into GM03813 cells resulted in a >2 fold increase of both full length and Δ7 SMN2 mRNA, compared to control transfected cells by qRT-PCR (Figure 17A). The increase of both the full length and Δ7 transcripts indicated the increase in full length mRNA is not a consequence of enhanced exon 7 splicing (Figure 17A).

In order to determine whether SMN protein levels could also increase as a consequence of SMN2 full length mRNA elevation, two SMA patient fibroblast cell lines

GM03813 and GM00232 expressing DcpS-specific shRNA were transfected with shRNA-resistant DcpS wild type ($DcpS^{WT}$), a catalytically inactive DcpS mutant ($DcpS^{MT}$), or $DcpS^{Ins15}$. Western Blot showed the protein level of SMN also increased about 2 fold in both GM03813 and GM00232 cells with $DcpS^{Ins15}$ expression (Figure 17B). Quantitation of the Western Blot was shown in Figure 17C. Collectively, our data indicated that $DcpS^{Ins15}$ specifically elevated SMN2 mRNA and protein levels independent of scavenger decapping activity.

Next, we determined whether $DcpS^{Ins15}$ could increase SMN2 without reducing endogenous DcpS. GM00232 SMA fibroblast cells were infected with increasing amount of lentivirus that expressed $DcpS^{Ins15}$. Lentiviruses expressing GFP or $DcpS^{MT}$ were used as controls. Western Blot analysis showed SMN protein levels increased about 2 fold when the expression level of $DcpS^{Ins15}$ reached levels comparable to endogenous wild type DcpS (Figure 18A). With low $DcpS^{Ins15}$ protein level which was similar to heterozygous individuals, there was also no SMN change (Figure 18A). Quantitation of the Western Blot showed that SMN protein level was elevated ~2 fold by expressing $DcpS^{Ins15}$ at levels that at least reached 1:1 ratio to endogenous DcpS (Figure 18B). Our data indicates a threshold level of $DcpS^{Ins15}$ is necessary in order to increase SMN.

Application of antisense morpholino oligonucleotides to alter DcpS pre-mRNA splicing in cells

The unexpected discovery of the $DcpS^{Ins15}$ alternative splice variant as an agent to increase full-length SMN2 mRNA and protein levels provides a potentially novel

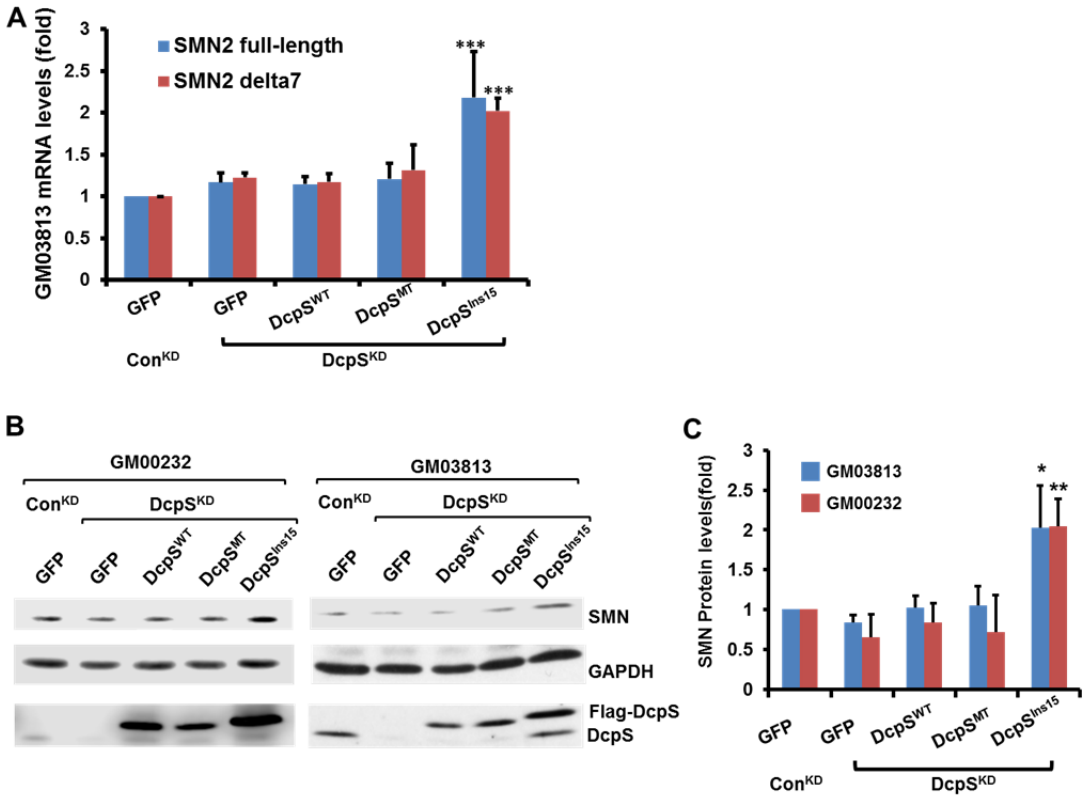

Figure 17. DcpS^{Ins15} increased full length SMN2 mRNA and SMN protein levels in SMA patient cells.

(A) GM03813 cells were infected with control (Con^{KD}) or DcpS shRNA expressing (DcpS^{KD}) lentiviruses, and complemented with shRNA-resistant DcpS^{WT}, catalytically inactive DcpS^{MT}, or DcpS^{Ins15}. Full length and exon7 truncated SMN2 mRNA levels were detected by Real Time PCR with transcript-specific primers. RNA levels are presented relative to the GAPDH mRNA and derived from three independent experiments. **(B)** SMN protein levels were detected by Western Blot in GM03813 and GM00232 cells from cells subjected to the same conditions as in (A). GAPDH was detected as loading control. **(C)** Quantitative summary of three independent Western Blot results as (B). Protein levels are normalized to GAPDH. Error bars represent +/- SD. P values are denoted by asterisks; * represents $p < 0.05$, ** represents $p < 0.01$ (Student's *t* test).

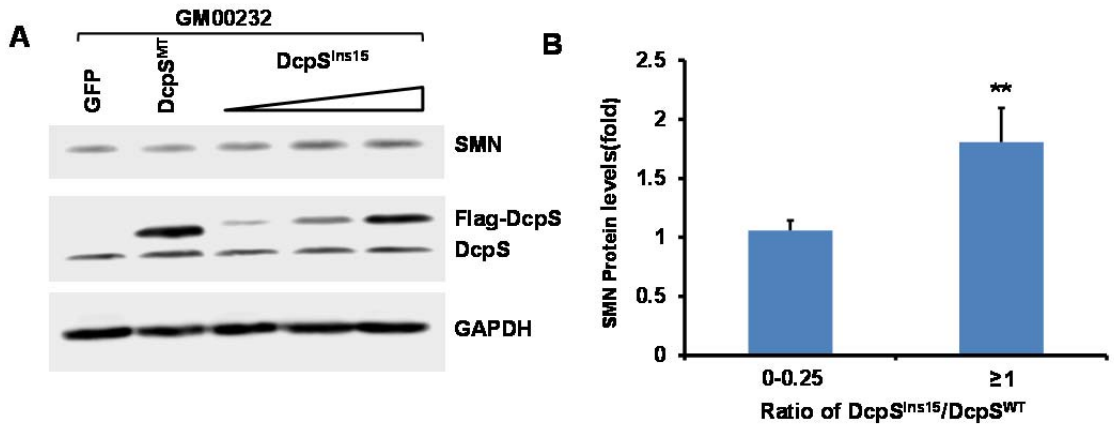

Figure 18. DcpS^{Ins15} increased SMN protein levels in a dose dependent manner.

(A) SMN protein levels were detected by Western Blot in GM00232 cells infected with lentiviruses expressing GFP, catalytically inactive DcpS^{MT}, or increasing amount of DcpS^{Ins15}. GAPDH was detected as loading control. **(B)** Quantitative summary of three independent Western Blot results as (A). Protein levels are normalized to GAPDH. Error bars represent +/- SD. P values are denoted by asterisks; ** represents $p < 0.01$ (Student's *t* test).

therapeutic avenue for SMA. The DcpS pre-mRNA is amendable by antisense splice switch oligonucleotide (SSO) technology to shift the endogenous DcpS splicing pattern to a predominant $DcpS^{Ins15}$ variant in SMA patients. With this approach, an antisense oligonucleotide complementary to the endogenous splice site is used to block the endogenous splicing pattern and shift it towards the $DcpS^{Ins15}$ splicing pattern in cells (Figure 18A). Our objective was to utilize SSO's to partially mask the endogenous 5′ splice site of DcpS intron 4 and redirect splicing to the cryptic splice site 45 nucleotides downstream to generate $DcpS^{Ins15}$ at levels sufficient to support increased full length SMN2 mRNA while retaining effective levels of wild type DcpS to maintain its normal function.

Four different vivo-morpholino antisense oligonucleotides targeting different positions of exon4 and intron4 junction in DcpS pre-mRNA were designed to directly switch splicing of DcpS in cells (Figure 19B). GM03813 cells were treated with a titration of antisense morpholino oligonucleotides containing a cell permeable guanidinium based dendramer delivery system (vivo-morpholino; GeneTools LLC) for 24 hours and splicing patterns were determined by RT- PCR. By using a pair of primers spanning exon3 to exon 6, three possible splicing patterns were shown in Figure 20A: wild type, insertion of the 45 nucleotides and a third product that skipped exon4 completely and spliced from exon 3 directly into exon 5. The addition of the vivo-morpholino oligonucleotide #1, #3 and #4 almost completely inhibited utilization of the endogenous exon 4 splice site at 1 μM concentration while oligonucleotide#2 didn't work as efficiently (Figure 20B). However all four oligonucleotides we tested not only

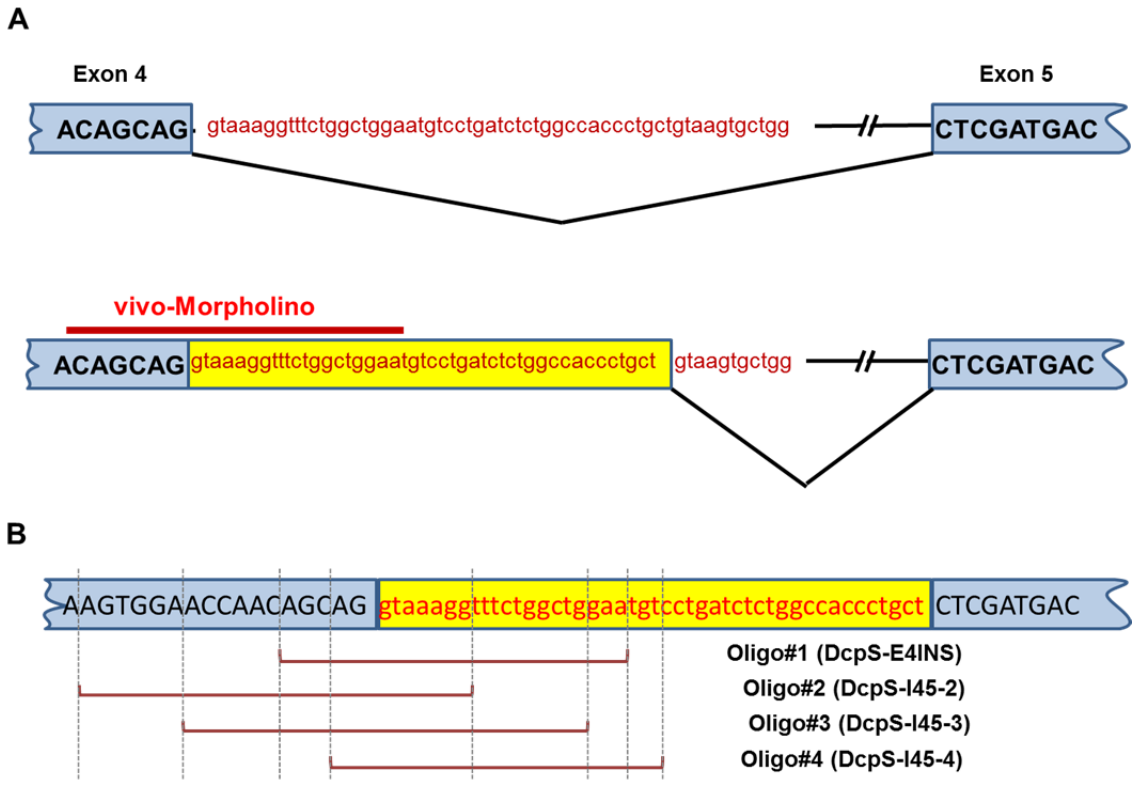

Figure 19. Design Morpholino anti-sense oligonucleotides that shift DcpS splicing.

(A) Schematic presentation of the splicing shift of DcpS by vivo-morpholino targeting DcpS exon4/intro4 junction. **(B)** Schematic presentation of the four designed vivo-morpholino oligonucleotides.

promoted utilization of the cryptic site to generate the $DcpS^{Ins15}$ splice variant, but also promoted skipping of exon 4 entirely (Figure 20B). This skipping effect was unexpected and is detrimental because it would generate a truncated DcpS transcript bypassing exon 4 and the insert.

After screening the splicing alteration efficiency of our four Morpholino oligonucleotides, we next checked whether $DcpS^{Ins15}$ protein was expressed in cells treated with the SSO's. GM03813 cells were treated with 0.5μM or 0.75μM Morpholino oligonucleotide#1 and 0.75μM or 1uM oligonucleotide #3 respectively and harvested after 6 days. Different DcpS transcripts were determined by RT-PCR using primers spanning exon3 to exon 6 and DcpS protein. A polyclonal antibody raised against the entire DcpS protein was use to enable detection of both $DcpS^{Ins15}$ and any putative truncated DcpS from the exon4-excluded transcript. However, no $DcpS^{Ins15}$ protein was detected in any of the treatment although there was $DcpS^{Ins15}$ mRNA expression (Figure 21). There was no truncated DcpS (lower band) detected either (Figure 21), which might be an indication that the truncated products were unstable. Interestingly, we observed high cell death rate after 6-day treatment with 1uM Morpholino oligonucleotide#1 and 1.25uM oligonucleotide #3, which reduced wild type DcpS transcripts more than 90% (Figure 20B). Thus we were unable to collect and test the longer incubation samples. In summary, although our findings demonstrate that the exon 4- intron 4 junction of DcpS is accessible to antisense oligonucleotide directed manipulation, at least the initial set of four antisense morpholino oligonucleotides tested do not sufficiently shift the endogenous DcpS pre-mRNA splicing to levels that generate detectable $DcpS^{Ins15}$ protein.

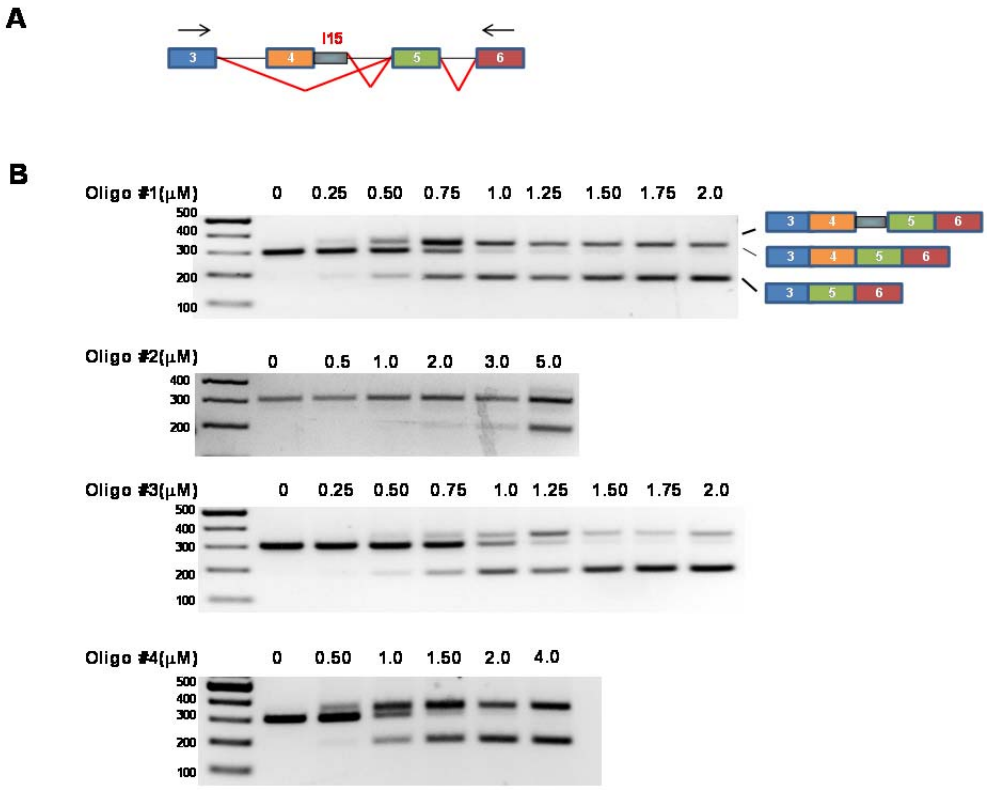

Figure 20. Morpholino oligonucleotide efficiencies were screening by RT-PCR.

(A) Schematic depicting the position of primers used for the following RT-PCR screens of possible splicing patterns caused by morpholino oligonucleotides. **(B)** Endogenous DcpS mRNAs with different splicing patterns were detected by RT-PCR in GM03813 cells after 2-day treatment of morpholino oligonucleotides #1-#4 at the indicated concentrations. Schematic represents the exon usage.

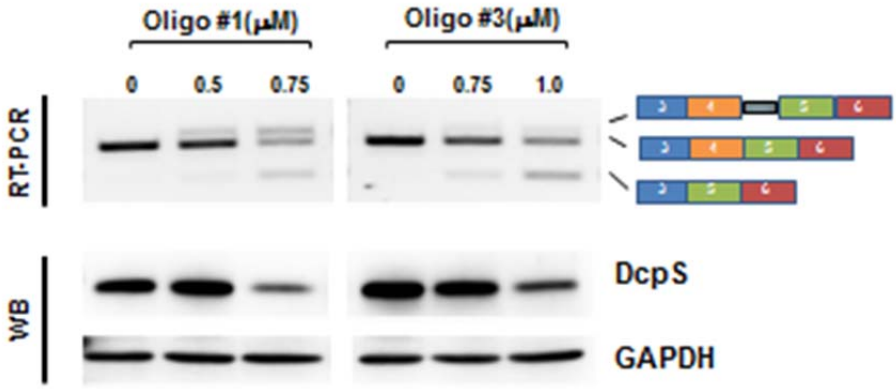

Figure 21. Morpholino oligonucleotides used are capable of changing endogenous DcpS transcript splicing but not the production of DcpS^{Ins15} protein.

GM03813 cells were harvested after 6-day treatment of morpholino oligonucleotide #1 and #3 at the indicated concentrations. Top: endogenous DcpS mRNAs with different splicing patterns were detected by RT-PCR with corresponding schematics indicating exon usage. Bottom: DcpS protein was detected by Western Blot using a DcpS polyclonal antibody. GAPDH protein was used as a loading control.

A panel of additional oligonucleotides that walk along the entire junction appears to be necessary to identify oligonucleotides that can efficiently shift the splicing pattern.

Discussion

In this chapter, we observed that full length SMN2 mRNA levels increased approximately 2 fold in $DcpS^{Ins15}$ homozygous patient lymphoblast cells compared to wild type and heterozygous lymphoblast cells (Figure 16A). We next confirmed that full length SMN2 mRNA levels also increased 2 fold in 293T $DcpS^{KD}$ cells expressing exogenous $DcpS^{Ins15}$ but not catalytic inactive $DcpS^{MT}$ (Figure 16B). A similar outcome was also detected in a the GM03813SMA patient fibroblast cell line where full length SMN2 mRNA levels also increased 2 fold in this cell line upon expression of exogenous $DcpS^{Ins15}$ (Figure 17A). The elevation of SMN2 RNA is accompanied with a corresponding increase of SMN protein levels (Figure 17B and 18). We showed that vivo-morpholino antisense oligonucleotides which targeted exon 4 and intron 4 junction in DcpS pre-mRNA were able to reduce the endogenous DcpS splicing pattern and partially shift the splicing to the $DcpS^{Ins15}$ isoform pattern in cells (Figure 20B). However, we also observed an unexpected alternative splicing whereby the entire exon 4 was skipped in morpholino treated cells (Figure 20B). Our studies establish a proof of principle for the use of vivo-morpholino technology to shift DcpS splicing, however, the current set of antisense oligonucleotides to not generate sufficient levels of variant mRNA to translate into detectable $DcpS^{Ins15}$ protein (Figure 21).

$DcpS^{Ins15}$ is likely a gain of function variant coordinating an elevation of SMN2 in a manner independent of its decapping activity. We performed all our initial experiments in DcpS knocked down cells to minimize any potential interference by endogenous DcpS since DcpS functions as a homodimer. However, we predicted that sufficient expression of $DcpS^{Ins15}$ should be able to elevate SMN2 even within the normal endogenous DcpS protein background. Indeed, in preliminary experiments we observed about 2 fold increase of SMN protein level by overexpressing $DcpS^{Ins15}$ up to at least 1:1 protein ratio to endogenous DcpS in wild type GM00232 cells (Figure 18). This is consistent with our data that heterozygous individuals did not show increased SMN2 mRNA levels. Therefore, to increase SMN2 mRNA required a certain threshold level of $DcpS^{Ins15}$ protein.

The application of vivo-morpholino antisense oligonucleotides to express endogenous $DcpS^{Ins15}$ was promising but still faces a few challenges. One problem is the unexpected exon 4 skipping caused by the morpholino oligonucleotides. While the morpholino oligonucleotides successfully inhibited the endogenous DcpS intron 4 splicing and promoted utilization of the $DcpS^{Ins15}$ splicing pattern, they also generated a splicing variant that skipped the entire exon 4. In most cases, this actually occurred more efficient than $DcpS^{Ins15}$ splicing. The binding of morpholino oligonucleotides at exon 4-intron 4 junction appear to disrupt both the endogenous and variant intron 4 splicing, favoring a bypass splicing pattern that omits exon 4 and splices exon 3 to exon 5. To eliminate this type of skipping, we will likely need to screen a series of oligonucleotides that span the exon 4-intron 4 junction at single nucleotide intervals. Another problem we faced was the low protein expression of $DcpS^{Ins15}$. Despite the significant expression of

$DcpS^{Ins15}$ splicing variant by morpholino oligonucleotides, we were unable to detect production of $DcpS^{Ins15}$ protein consistent with the low level of $DcpS^{Ins15}$ protein level observed in ID patient cells (Chapter 2). We have ruled out several possible regulatory pathways including proteasomal/lysosomal degradation of $DcpS^{Ins15}$ protein or translational repression by microRNAs (Chapter 2), but a clear understanding of why low levels of $DcpS^{Ins15}$ protein are generated is still lacking. As a therapeutic agent, a fine balance would have to be maintained between wild type and variant DcpS. Ideally we should express as much protein as possible to reach the threshold level. On the other hand, we would need to retain a certain level of wild type DcpS expression to maintain normal cell function otherwise it could cause neurological defects (Ahmed et al, 2015a; Ng et al, 2015a).

In addition to the therapeutic application of $DcpS^{Ins15}$ in SMA, the unknown mechanism beyond this specific regulation may also provide new perspectives in SMN2 gene regulation. Determining how $DcpS^{Ins15}$ upregulates SMN2 could provide novel insights into strategies to upregulate SMN2 expression independent of $DcpS^{Ins15}$ and avoid complications that could be associated with reduced wild type DcpS levels in cells.

Concluding remarks

Despite its initial described function in the last step of the mRNA 3´end decay pathway, DcpS is a multifunctional protein involved in a greater network of regulated gene expression. We identified two long non coding RNA, DRNT1 and DRNT2, as human DcpS responsive RNAs. DRNT1 RNA stability was regulated by DcpS in a transcript-specific manner. This modulation is dependent on its decapping activity and mediated through Xrn1. One the other hand, the regulation of DRNT2 steady state level by DcpS independent of transcript stability strongly suggests a role for DcpS in non-decay regulation.

Our studies have also uncovered a link of DcpS to human cognition. We identified mutant DcpS genes in intellectual disability patients from a Pakistan family within consanguineous marriages. Functional analyses validated that both of the two mutations ($DcpS^{Ins15}$ and $DcpS^{T316M}$) lead to abolishment of DcpS decapping function in vitro and in vivo. Moreover, a different homozygous mutation in DcpS from a second family with intellectual disability was identified independently by another group (Ng et al, 2015a). All these findings reinforce the significance of DcpS decapping activity for neuronal development. How DcpS decapping activity contributes to neurogenesis remains to be determined.

The fact that DcpS deletion is lethal in cells by CRISPR (Shalem et al, 2014) and embryonic lethal in mice (M. Kiledjian: unpublished observations), demonstrates the requirement of DcpS for cellular viability and embryogenesis. The viability of individuals with a decapping-defective DcpS suggests that the decapping of DcpS is only important

for normal cognition, while an additional decapping-independent function of DcpS must also be required for normal development and survival. This additional importance of DcpS in fundamental cellular functions has yet to be determined.

In a fortuitous development, we observed that lymphoblast cells from ID patients harboring a homozygous $DcpS^{Ins15}$ allele contain significant elevation of SMN2 mRNA uncovering a surprising connection between expressions of the $DcpS^{Ins15}$ variant with SMA. We showed that exogenous expression of $DcpS^{Ins15}$ into SMA patient fibroblasts led to elevation of SMN2 mRNA and SMN protein but not with either wild type or a catalytically inactive DcpS. We also demonstrated the feasibility of using vivo-morpholino antisense oligonucleotides, which targeted the exon 4 - intron 4 junction in the DcpS pre-mRNA as a tool to switch splicing of DcpS from wild type to $DcpS^{Ins15}$ isoform in cells. Our findings provide a new strategy in SMA therapeutic application, but the challenges such as low $DcpS^{Ins15}$ protein expression, nonproductive exon skipping as a consequence of the antisense morpholinos thus far used and the potential risk of intellectual disability, have yet to be overcome.

Finally, as a potential SMA drug, RG3039 is not a simple DcpS inhibitor and the extension of survival of SMA mice by RG3039 is likely independent of DcpS and SMN. We identified four mRNAs (PAQR8, ATOH7, RAB26 and MAOB) mediated by RG3039 independent of DcpS. Interestingly, the regulation of PAQR8 could be attributed to a decrease in mRNA stability, while the stability of the other three RNAs was unchanged upon RG3039 treatment. How RG3039 modulates RNA levels by a mechanism dependent and independent of transcript stability and independent of DcpS

needs to be further addressed to gain a more complete understanding of how this potential drug candidate may function in cells.

Table IV Microarray results

Affy Gene Symbol	Gene_Symbol	Unigene	Raw Est Fold Change	Individual Transcript p-value	Gene Descriptor
RFX4	RFX4	Hs.388827	11.647	0.0002	regulatory factor X, 4 (influences HLA class II expression)
---	PRO1621	---	11.37	0.0011	PRO1621 protein
---	KCNIP4	---	10.652	0.0004	Kv channel interacting protein 4
STARD10	STARD10	---	10.361	0.0038	START domain containing 10
---	LOC728245	---	8.947	0.0214	hypothetical protein LOC728245
LOC100129387	GABPB2	Hs.511316	8.289	0.0011	GA binding protein transcription factor, beta subunit 2
TRIML2	FLJ25801	---	8.034	0.0135	hypothetical protein FLJ25801
TRIM78P	TRIMP1	---	7.891	0.0023	tripartite motif-containing pseudogene 1
PROCA1	PROCA1	---	7.85	0.0022	proline-rich cyclin A1-interacting protein
ZPLD1	ZPLD1	---	7.59	0.0173	zona pellucida-like domain containing 1
MIR17HG	C13orf25	---	7.588	0.0101	chromosome 13 open reading frame 25
RAB26	RAB26	Hs.3797	7.31	0	RAB26, member RAS oncogene family
ABI3BP	ABI3BP	---	7.291	0.01	ABI gene family, member 3 (NESH) binding protein
---	EPHA5	Hs.175396	6.078	0.0082	EPH receptor A5
CD44	CD44	---	6.054	0.0251	CD44 molecule (Indian blood group)
RS1	RS1	---	5.999	0.0364	retinoschisis (X-linked, juvenile) 1
ANKRD9	ANKRD9	---	5.98	0.0012	Ankyrin repeat domain 9
TMEM144	TMEM144	Hs.59167	5.904	0.0057	transmembrane

		3			protein 144
---	PTK2	---	5.867	0.0003	PTK2 protein tyrosine kinase 2
TOB2	TOB2	---	5.835	0.0089	transducer of ERBB2, 2
---	ELF4	---	5.753	0.0105	E74-like factor 4 (ets domain transcription factor)
PIP4K2B	PIP4K2B	Hs.260603	5.66	0.0001	Phosphatidylinositol-4-phosphate 5-kinase, type II, beta
LOC283682	LOC283682	---	5.591	0.0381	hypothetical protein LOC283682
OPN5	OPN5	---	5.546	0.0308	opsin 5
LRMP	LRMP	---	5.402	0.0417	Lymphoid-restricted membrane protein
SLC17A9	C20orf59	---	5.28	0.0046	chromosome 20 open reading frame 59
SLC7A1	SLC7A1	---	5.194	0.0269	solute carrier family 7 (cationic amino acid transporter, y+ system), member 1
---	A2BP1	---	5.094	0.0006	Ataxin 2-binding protein 1
OR4D1	OR4D1	---	5.004	0.0026	olfactory receptor, family 4, subfamily D, member 1
SNCG	SNCG	---	4.988	0.0008	synuclein, gamma (breast cancer-specific protein 1)
FAM13C	FAM13C1	---	4.932	0.0275	family with sequence similarity 13, member C1
CHD6	CHD6	---	4.93	0.0199	chromodomain helicase DNA binding protein 6
LOC100288294	---	---	4.8	0.0248	Transcribed locus
MTTP	MTTP	---	4.794	0.0267	microsomal triglyceride transfer protein
JAKMIP3	C10orf39	Hs.106254	4.794	0.0008	chromosome 10 open reading frame 39
LTBP4	LTBP4	---	4.79	0.0201	Latent transforming growth factor beta binding protein 4

SLC2A12	SLC2A12	---	4.732	0.0264	solute carrier family 2 (facilitated glucose transporter), member 12
---	LRP4	---	4.627	0.0012	Low density lipoprotein receptor-related protein 4
SYT12	SYT12	---	4.61	0.002	Synaptotagmin XII
MYO7A	MYO7A	---	4.607	0.0439	myosin VIIA
UGT1A	UGT1A6	---	4.534	0.0455	UDP glucuronosyltransferase 1 family, polypeptide A9
C20orf96	C20orf96	---	4.496	0.0469	chromosome 20 open reading frame 96
HCG22	HCG22	---	4.488	0.0009	HLA complex group 22
ARHGEF12	ARHGEF12	---	4.476	0.0451	Rho guanine nucleotide exchange factor (GEF) 12
EBF3	EBF3	---	4.436	0.0284	early B-cell factor 3
IL1RAP	IL1RAP	---	4.344	0.0288	interleukin 1 receptor accessory protein
RAB26	RAB26	---	4.254	0.0002	RAB26, member RAS oncogene family
KCNJ3	KCNJ3	---	4.247	0.0066	potassium inwardly-rectifying channel, subfamily J, member 3
---	---	Hs.370762	4.155	0.0004	Similar to Laminin receptor 1
RANBP3L	RANBP3L	---	4.138	0.0153	RAN binding protein 3-like
CLEC11A	CLEC11A	Hs.512680	4.131	0.0019	C-type lectin domain family 11, member A
ZNF831	C20orf174	---	4.113	0.0136	chromosome 20 open reading frame 174
---	EHMT2	---	3.99	0.0117	Euchromatic histone-lysine N-methyltransferase 2
GRM1	GRM1	---	3.872	0.0396	glutamate receptor, metabotropic 1
TCF3	TCF3	---	3.863	0.0012	Transcription factor 3 (E2A immunoglobulin

					enhancer binding factors E12/E47)
---	CCND2	---	3.852	0.0202	Cyclin D2
VPS13B	VPS13B	---	3.851	0.0188	vacuolar protein sorting 13 homolog B (yeast)
---	GATA6	---	3.786	0.0356	GATA binding protein 6
ABCA1	ABCA1	---	3.692	0.0382	ATP-binding cassette, sub-family A (ABC1), member 1
---	HMGB2	Hs.434953	3.677	0.0003	High-mobility group box 2
LOC100272217	LOC158301	---	3.534	0.0021	Hypothetical protein LOC158301
C9orf45	C9orf45	---	3.529	0.028	chromosome 9 open reading frame 45 /// chromosome 9 open reading frame 45
---	CHRNA10	---	3.515	0.0444	Cholinergic receptor, nicotinic, alpha polypeptide 10
HIST2H2BE	HIST2H2BE	Hs.2178	3.417	0	histone cluster 2, H2be
---	RHBDD1	---	3.384	0.0141	Rhomboid domain containing 1
PTGIR	PTGIR	---	3.363	0	prostaglandin I2 (prostacyclin) receptor (IP)
LOC400960	LOC400960	Hs.516253	3.33	0.0003	hypothetical gene supported by BC040598
FAM45A	FAM45A	---	3.321	0.0447	family with sequence similarity 45, member A
FLJ39739	FLJ39739	---	3.311	0.021	FLJ39739 protein
LOC100272217	LOC158301	---	3.301	0.002	Hypothetical protein LOC158301
SIGLEC9	SIGLEC9	---	3.272	0.0437	sialic acid binding Ig-like lectin 9
BC011766	LOC728210	---	3.271	0.0006	Hypothetical protein LOC728210
CD79B	CD79B	---	3.256	0.0035	CD79b molecule, immunoglobulin-associated beta

RNF165	RNF165	Hs.501114	3.249	0	ring finger protein 165
LOC100288271	LOC730496	---	3.209	0.0188	hypothetical protein LOC730496
---	TJP2	---	3.18	0.029	Tight junction protein 2 (zona occludens 2)
GAL3ST2	GAL3ST2	---	3.16	0.0206	galactose-3-O-sulfotransferase 2
LOC253044	LOC253044	---	3.093	0.0056	hypothetical protein LOC253044
VWA5A	LOH11CR2A	---	3.069	0.0374	loss of heterozygosity, 11, chromosomal region 2, gene A
PNCK	PNCK	---	3.034	0.0215	pregnancy upregulated non-ubiquitously expressed CaM kinase
CRH	CRH	---	3.028	0.0214	corticotropin releasing hormone
A2M	A2M	---	3.01	0.0259	alpha-2-macroglobulin
LOC401588	LOC401588	---	2.982	0	hypothetical LOC401588
HIST1H1A	HIST1H1A	---	2.965	0.0257	histone cluster 1, H1a
PRDM16	PRDM16	---	2.959	0.0415	PR domain containing 16
ERF	ERF	---	2.955	0.0121	Ets2 repressor factor
FAM187B	FLJ25660	---	2.911	0.0283	hypothetical protein FLJ25660
ADAMTSL1	ADAMTSL1	---	2.883	0.0031	ADAMTS-like 1
C11orf71	C11orf71	Hs.91816	2.872	0	chromosome 11 open reading frame 71
CLEC11A	CLEC11A	---	2.871	0.0003	C-type lectin domain family 11, member A /// C-type lectin domain family 11, member A
TFDP1	TFDP1	Hs.79353	2.87	0	transcription factor Dp-1
TMPRSS2	TMPRSS2	---	2.847	0.0425	transmembrane protease, serine 2
---	X Kell blood group	Hs.282082	2.838	0.0001	X Kell blood group precursor-related

	precursor-r				family, member 7
LEKR1	FLJ16641	Hs.478048	2.773	0.0001	FLJ16641 protein
HIST1H2BB	HIST1H2BB	---	2.763	0.0419	histone cluster 1, H2bb
C20orf112	C20orf112	Hs.516977	2.748	0	chromosome 20 open reading frame 112
ZNF518B	KIAA1729	Hs.455089	2.736	0.0001	KIAA1729 protein
CDH16	CDH16	---	2.683	0.0091	cadherin 16, KSP-cadherin
HIST1H3B	HIST1H3B	---	2.656	0.0032	histone cluster 1, H3b
LOC644450	LOC644450	---	2.643	0.0108	hypothetical protein LOC644450
LOC729013	LOC729013	Hs.34068	2.627	0.0004	hypothetical protein LOC729013
---	MGC16733	---	2.622	0.0184	Hypothetical protein FLJ25037
CDKN2AIPNL	MGC13017	Hs.156506	2.62	0	similar to RIKEN cDNA A430101B06 gene
---	GENX-3414	---	2.613	0.0027	Genethonin 1
FAM3D	FAM3D	---	2.612	0.0496	family with sequence similarity 3, member D
---	LOC729776	Hs.645315	2.61	0	hypothetical protein LOC729776
LOC439911	LOC439911	---	2.609	0.0001	hypothetical gene supported by NM_194304
---	NSMCE2	---	2.608	0.043	Non-SMC element 2 homolog (MMS21, S. cerevisiae)
LOC645722	LOC645722	---	2.582	0.0014	hypothetical LOC645722
MIB2	MIB2	Hs.135805	2.564	0.0003	mindbomb homolog 2 (Drosophila)
---	C12orf32	Hs.198853	2.555	0.0003	Chromosome 12 open reading frame 32
HIST2H2AA3 /// HIST2H2AA4	HIST2H2AA3 /// HIST2H2AA4	---	2.532	0.0008	histone cluster 2, H2aa3 /// histone cluster 2, H2aa4
ST7OT1	ST7OT1	---	2.518	0.0021	ST7 overlapping transcript 1 (antisense

					non-coding RNA)
---	PBX1	---	2.515	0.0012	Pre-B-cell leukemia transcription factor 1
LOC646903	LOC646903	Hs.632559	2.513	0	hypothetical LOC646903
LPIN3	LPIN3	---	2.51	0.0283	lipin 3
HIST3H2A	HIST3H2A	---	2.507	0.0006	histone cluster 3, H2a
LOC730631	HIST2H4	Hs.55468	2.506	0	Histone 2, H4
C22orf45	ADORA2A	---	2.496	0.0068	Adenosine A2a receptor
TRIM54	TRIM54	---	2.493	0.0105	tripartite motif-containing 54
TRADD	TRADD	---	2.479	0.0265	TNFRSF1A-associated via death domain
---	LOC729705 /// LOC731763	---	2.465	0.001	hypothetical protein LOC729705 /// hypothetical protein LOC731763
FEZF2	ZNF312	---	2.454	0.0309	zinc finger protein 312
LOC729887	---	---	2.452	0.0024	Similar to Serine/threonine-protein kinase PLK1 (Polo-like kinase 1) (PLK-1) (Serine-threonine protein kinase 13) (STPK13)
C9orf123	C9orf123	Hs.7517	2.432	0.0002	chromosome 9 open reading frame 123
CNR1	CNR1	---	2.422	0.0027	Cannabinoid receptor 1 (brain)
IFNAR1	---	---	2.415	0.0009	Transcribed locus
ZNF836	FLJ16287	---	2.393	0.0072	FLJ16287 protein
MYH6 /// MYH7	MYH6 /// MYH7	---	2.392	0.0188	myosin, heavy chain 6, cardiac muscle, alpha (cardiomyopathy, hypertrophic 1) /// myosin, heavy chain 7, cardiac muscle, beta
BICD1	BICD1	---	2.353	0.0003	bicaudal D homolog 1 (Drosophila)
MRGPRX3	MRGPRX3	---	2.342	0.0359	MAS-related GPR,

					member X3
---	EIF2AK3	---	2.325	0.0107	Eukaryotic translation initiation factor 2-alpha kinase 3
CACNA1G	CACNA1G	---	2.309	0.0171	calcium channel, voltage-dependent, alpha 1G subunit
---	ARPC5L	---	2.301	0.0007	Actin related protein 2/3 complex, subunit 5-like
NPRL3	C16orf35	---	2.295	0.0164	Chromosome 16 open reading frame 35
---	FLJ10385	---	2.289	0.006	WD repeat domain 79
LOC342918	LOC342918	---	2.284	0.0099	hypothetical LOC342918
BHLHE23	BHLHB4	---	2.281	0.0033	basic helix-loop-helix domain containing, class B, 4
SRRM2	SRRM2	---	2.276	0.0201	Serine/arginine repetitive matrix 2
---	PRDM5	---	2.27	0.0031	PR domain containing 5
FAIM3	FAIM3	---	2.269	0.027	Fas apoptotic inhibitory molecule 3 /// Fas apoptotic inhibitory molecule 3
---	IMMT	---	2.254	0.0271	Inner membrane protein, mitochondrial (mitofilin)
PCDHAC1	PCDHAC1	---	2.248	0.0112	protocadherin alpha subfamily C, 1
MAPK8IP3	MAPK8IP3	---	2.236	0.0259	Mitogen-activated protein kinase 8 interacting protein 3
LOC284861	LOC284861	---	2.225	0.0129	hypothetical gene supported by BC039313
---	LOC440064	Hs.534870	2.225	0.0002	LOC440064
FLJ37453	FLJ37453	---	2.212	0.0003	hypothetical protein LOC645580
SQSTM1	SQSTM1	---	2.209	0.0089	sequestosome 1
COX18	FLJ38991	---	2.206	0.0034	Mitochondrial COX18

---	LOC727804	---	2.193	0.0023	Hypothetical protein LOC727804
GRIA3	GRIA3	---	2.191	0.0016	glutamate receptor, ionotrophic, AMPA 3
ZNF551	ZNF551	---	2.186	0.0176	Zinc finger protein 551
LOC645676	LOC645676	---	2.164	0.0002	hypothetical LOC645676
C22orf39	LOC128977	---	2.163	0.0002	hypothetical protein LOC128977
---	GCNT2	---	2.163	0.0149	Glucosaminyl (N-acetyl) transferase 2, I-branching enzyme (I blood group)
MST1P9	MSTP9	---	2.154	0.0039	macrophage stimulating, pseudogene 9
DPP7	DPP7	Hs.37916	2.143	0	Dipeptidyl-peptidase 7
C9orf122	C9orf122	Hs.632652	2.132	0.0001	chromosome 9 open reading frame 122
LOC400027	LOC400027	---	2.129	0.0023	hypothetical gene supported by BC047417
FAM200B	LOC285550	Hs.399980	2.128	0	hypothetical protein LOC285550
FLJ10038	FLJ10038	---	2.12	0.0005	hypothetical protein FLJ10038
LOH12CR2	LOH12CR2	---	2.115	0.0061	loss of heterozygosity, 12, chromosomal region 2
LILRB3	LILRB2 /// LILRB3	---	2.114	0.0425	leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 2 /// leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 3
C6orf226	LOC441150	---	2.089	0.0006	similar to RIKEN cDNA 2310039H08
KCNK12	KCNK12	---	2.083	0.0014	potassium channel, subfamily K, member

					12
---	STK35	---	2.065	0.0238	Serine/threonine kinase 35
---	C10orf95	---	2.057	0.0235	Chromosome 10 open reading frame 95
SH3GLB2	SH3GLB2	---	2.046	0.0486	SH3-domain GRB2-like endophilin B2 /// SH3-domain GRB2-like endophilin B2
KLHDC8A	FLJ10748	---	2.04	0.0488	Hypothetical protein FLJ10748
---	LOC401048	Hs.190365	2.016	0	hypothetical LOC401048
SERF2	SERF2	---	2.006	0.0004	Small EDRK-rich factor 2
PDLIM5	PDLIM5	---	-2.155	0.0276	PDZ and LIM domain 5
ABLIM2	ABLIM2	---	-2.263	0.0419	Actin binding LIM protein family, member 2
CDC25C	CDC25C	---	-2.281	0.0483	cell division cycle 25 homolog C (S. cerevisiae)
PHLDA1	PHLDA1	---	-2.29	0.0389	pleckstrin homology-like domain, family A, member 1
CDH15	CDH15	---	-2.3	0.0014	cadherin 15, M-cadherin (myotubule)
RHBDL3	RHBDL4	---	-2.393	0.0014	rhomboid, veinlet-like 4 (Drosophila)
ZNF821	LOC55565	---	-2.497	0.003	hypothetical protein LOC55565
---	PACSIN2	---	-2.508	0.0467	Protein kinase C and casein kinase substrate in neurons 2
SH3YL1	SH3YL1	---	-2.62	0.0008	SH3 domain containing, Ysc84-like 1 (S. cerevisiae)
PHLDB2	PHLDB2	---	-2.999	0.0303	pleckstrin homology-like domain, family B, member 2
LAT2	LAT2	---	-3.026	0.0116	linker for activation of T cells family, member 2
---	PAX8	---	-3.048	0.0129	Paired box gene 8

PDLIM5	PDLIM5	---	-3.061	0.0427	PDZ and LIM domain 5
CD99L2	CD99L2	---	-3.129	0.0244	CD99 antigen-like 2
C9orf156	C9orf156	---	-3.235	0.0257	chromosome 9 open reading frame 156
---	LOC730271	---	-3.29	0.0082	Hypothetical protein LOC730271
SNORA74A	SNORA74A	---	-3.358	0.0255	small nucleolar RNA, H/ACA box 74A
NPPC	NPPC	---	-3.377	0.0178	natriuretic peptide precursor C
HCG4P6	HCG4P6	---	-3.404	0.0403	HLA complex group 4 pseudogene 6
NCKIPSD	NCKIPSD	---	-3.424	0.0133	NCK interacting protein with SH3 domain
---	XPO1	---	-3.536	0.0462	Exportin 1 (CRM1 homolog, yeast)
IGHM	IGHM	---	-3.655	0.0219	immunoglobulin heavy constant mu
CNIH3	CNIH3	Hs.28659	-3.809	0	cornichon homolog 3 (Drosophila)
---	RGNEF	---	-3.823	0.0334	Rho-guanine nucleotide exchange factor
LOC100132080	PDGFA	---	-3.908	0.0282	Platelet-derived growth factor alpha polypeptide
C10orf85	C10orf85	---	-3.944	0.022	chromosome 10 open reading frame 85
MAOB	MAOB	Hs.46732	-3.945	0.0001	monoamine oxidase B
PAQR8	PAQR8	Hs.239388	-4.149	0.0001	progestin and adipoQ receptor family member VIII
COL13A1	COL13A1	---	-4.166	0.0345	collagen, type XIII, alpha 1
C10orf79	C10orf79	---	-4.205	0.0174	chromosome 10 open reading frame 79
---	IDH3A	---	-4.365	0.0106	Isocitrate dehydrogenase 3 (NAD+) alpha
NSMCE4A	C10orf86	---	-4.544	0.0306	Chromosome 10 open reading frame 86
OFCC1	OFCC1	---	-4.576	0.0388	orofacial cleft 1 candidate 1

AGR3	BCMP11	---	-4.642	0.0389	breast cancer membrane protein 11
PTH2R	PTHR2	Hs.159499	-4.703	0.0001	parathyroid hormone receptor 2
---	C21orf66	---	-4.887	0.0122	Chromosome 21 open reading frame 66
IFNA2	IFNA2	---	-5.113	0.0021	interferon, alpha 2
ZNF76	ZNF76	---	-5.187	0.0004	zinc finger protein 76 (expressed in testis)
---	USP33	---	-5.378	0.0122	ubiquitin specific peptidase 33
RIN2	RIN2	---	-5.488	0.0058	Ras and Rab interactor 2
TRAK1	TRAK1	---	-5.659	0.0058	trafficking protein, kinesin binding 1
---	CAPZA1	---	-5.681	0.0067	Capping protein (actin filament) muscle Z-line, alpha 1
LOC202181	LOC202181	---	-5.765	0.0112	Hypothetical protein LOC202181
KIAA0226	KIAA0226	---	-5.878	0.0068	KIAA0226
NCOR1	NCOR1	---	-5.947	0.0208	nuclear receptor co-repressor 1
---	HEG1	---	-5.992	0.0072	HEG homolog 1 (zebrafish)
FAM125B	FAM125B	---	-6.296	0.0021	family with sequence similarity 125, member B
FBXL21	FBXL21	---	-6.311	0.0087	F-box and leucine-rich repeat protein 21
PMS2L1	PMS2L1	---	-6.694	0.0003	postmeiotic segregation increased 2-like 1
FGF7	FGF7	---	-6.918	0.018	fibroblast growth factor 7 (keratinocyte growth factor)
RGS11	RGS11	---	-7.272	0.0084	regulator of G-protein signalling 11

Reference

Ahmed I, Buchert R, Zhou M, Jiao X, Mittal K, Sheikh TI, Scheller U, Vasli N, Rafiq MA, Brohi MQ, Mikhailov A, Ayaz M, Bhatti A, Sticht H, Nasr T, Carter M, Uebe S, Reis A, Ayub M, John P, Kiledjian M, Vincent JB, Jamra RJ (2015a) Mutations in DCPS and EDC3 in Autosomal Recessive Intellectual Disability Indicate a Crucial Role for mRNA Decapping in Neurodevelopment. *Hum Mol Genet*

Ahmed I, Buchert R, Zhou M, Jiao X, Mittal K, Sheikh TI, Scheller U, Vasli N, Rafiq MA, Brohi MQ, Mikhailov A, Ayaz M, Bhatti A, Sticht H, Nasr T, Carter MT, Uebe S, Reis A, Ayub M, John P, Kiledjian M, Vincent JB, Jamra RA (2015b) Mutations in DCPS and EDC3 in autosomal recessive intellectual disability indicate a crucial role for mRNA decapping in neurodevelopment. *Human molecular genetics*

Allmang C, Petfalski E, Podtelejnikov A, Mann M, Tollervey D, Mitchell P (1999) The yeast exosome and human PM-Scl are related complexes of 3' --> 5' exonucleases. *Genes Dev* **13:** 2148-2158

Anderson JSJ, Parker RP (1998) The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex. *EMBO J* **17:** 1497-1506

Anderton RS, Meloni BP, Mastaglia FL, Boulos S (2013) Spinal muscular atrophy and the antiapoptotic role of survival of motor neuron (SMN) protein. *Molecular neurobiology* **47:** 821-832

Andreassi C, Jarecki J, Zhou J, Coovert DD, Monani UR, Chen X, Whitney M, Pollok B, Zhang M, Androphy E, Burghes AH (2001) Aclarubicin treatment restores SMN levels to cells derived from type I spinal muscular atrophy patients. *Hum Mol Genet* **10:** 2841-2849

Azzouz M, Le T, Ralph GS, Walmsley L, Monani UR, Lee DC, Wilkes F, Mitrophanous KA, Kingsman SM, Burghes AH, Mazarakis ND (2004) Lentivector-mediated SMN replacement in a mouse model of spinal muscular atrophy. *J Clin Invest* **114:** 1726-1731

Bail S, Kiledjian M (2008) DcpS, a general modulator of cap-binding protein-dependent processes? *RNA Biol* **5:** 216-219

Bail S, Swerdel M, Liu H, Jiao X, Goff LA, Hart RP, Kiledjian M (2010) Differential regulation of microRNA stability. *RNA* **16:** 1032-1039

Behrens SE, Luhrmann R (1991) Immunoaffinity purification of a [U4/U6.U5] tri-snRNP from human cells. *Genes Dev* **5:** 1439-1452

Besser LM, Shin M, Kucik JE, Correa A (2007) Prevalence of down syndrome among children and adolescents in metropolitan Atlanta. *Birth Defects Res A* **79:** 765-774

Black DL (2003) Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* **72:** 291-336

Blencowe BJ, Bowman JA, McCracken S, Rosonina E (1999) SR-related proteins and the processing of messenger RNA precursors. *Biochemistry and cell biology = Biochimie et biologie cellulaire* **77:** 277-291

Bosse GD, Ruegger S, Ow MC, Vasquez-Rifo A, Rondeau EL, Ambros VR, Grosshans H, Simard MJ (2013) The decapping scavenger enzyme DCS-1 controls microRNA levels in Caenorhabditis elegans. *Mol Cell* **50:** 281-287

Boulisfane N, Choleza M, Rage F, Neel H, Soret J, Bordonne R (2010) Impaired minor tri-snRNP assembly generates differential splicing defects of U12-type introns in lymphoblasts derived from a type I SMA patient. *Hum Mol Genet* **20:** 641-648

Burnett BG, Munoz E, Tandon A, Kwon DY, Sumner CJ, Fischbeck KH (2009) Regulation of SMN protein stability. *Mol Cell Biol* **29:** 1107-1115

Bushell M, Sarnow P (2002) Hijacking the translation apparatus by RNA viruses. *J Cell Biol* **158:** 395-399

Butchbach ME, Singh J, Thorsteinsdottir M, Saieva L, Slominski E, Thurmond J, Andresson T, Zhang J, Edwards JD, Simard LR, Pellizzoni L, Jarecki J, Burghes AH, Gurney ME (2010) Effects of 2,4-diaminoquinazoline derivatives on SMN expression and phenotype in a mouse model for spinal muscular atrophy. *Hum Mol Genet* **19:** 454-467

Cartegni L, Chew SL, Krainer AR (2002) Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nature reviews Genetics* **3:** 285-298

Chernyakov I, Whipple JM, Kotelawala L, Grayhack EJ, Phizicky EM (2008) Degradation of several hypomodified mature tRNA species in Saccharomyces cerevisiae is mediated by Met22 and the 5'-3' exonucleases Rat1 and Xrn1. *Genes & development* **22:** 1369-1380

Cherry JJ, Evans MC, Ni J, Cuny GD, Glicksman MA, Androphy EJ (2012) Identification of novel compounds that increase SMN protein levels using an improved SMN2 reporter cell assay. *Journal of biomolecular screening* **17:** 481-495

Cherry JJ, Osman EY, Evans MC, Choi S, Xing X, Cuny GD, Glicksman MA, Lorson CL, Androphy EJ (2013) Enhancement of SMN protein levels in a mouse model of spinal muscular atrophy using novel drug-like compounds. *EMBO molecular medicine* **5:** 1035-1050

Chiara MD, Gozani O, Bennett M, Champion-Arnaud P, Palandjian L, Reed R (1996) Identification of proteins that interact with exon sequences, splice sites, and the branchpoint sequence during each stage of spliceosome assembly. *Mol Cell Biol* **16:** 3317-3326

Choi KM, McMahon LP, Lawrence JC, Jr. (2003) Two motifs in the translational repressor PHAS-I required for efficient phosphorylation by mammalian target of rapamycin and for recognition by raptor. *J Biol Chem* **278:** 19667-19673

Cioce M, Lamond AI (2005) Cajal bodies: a long history of discovery. *Annual review of cell and developmental biology* **21:** 105-131

Cohen LS, Mikhli C, Jiao X, Kiledjian M, Kunkel G, Davis RE (2005) Dcp2 Decaps m2,2,7GpppN-capped RNAs, and its activity is sequence and context dependent. *Mol Cell Biol* **25:** 8779-8791

Collins VR, Muggli EE, Riley M, Dip G, Palma S, Halliday JL (2008) Is down syndrome a disappearing birth defect? *J Pediatr* **152:** 20-24

Conti E, Izaurralde E (2005) Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. *Current opinion in cell biology* **17:** 316-325

Cooper TA, Wan L, Dreyfuss G (2009) RNA and disease. *Cell* **136:** 777-793

Darras BT, Kang PB (2007) Clinical trials in spinal muscular atrophy. *Current opinion in pediatrics* **19:** 675-679

De Benedetti A, Harris AL (1999) eIF4E expression in tumors: its possible role in progression of malignancies. *Int J Biochem Cell Biol* **31:** 59-72

Deshmukh MV, Jones BN, Quang-Dang DU, Flinders J, Floor SN, Kim C, Jemielity J, Kalek M, Darzynkiewicz E, Gross JD (2008) mRNA Decapping Is Promoted by an RNA-Binding Channel in Dcp2. *Mol Cell* **29:** 324-336

Devys D, Lutz Y, Rouyer N, Bellocq JP, Mandel JL (1993) The FMR-1 protein is cytoplasmic, most abundant in neurons and appears normal in carriers of a fragile X premutation. *Nature genetics* **4:** 335-340

DiDonato CJ, Parks RJ, Kothary R (2003) Development of a gene therapy strategy for the restoration of survival motor neuron protein expression: implications for spinal muscular atrophy therapy. *Human gene therapy* **14:** 179-188

Doma MK, Parker R (2006) Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* **440:** 561-564

Dominguez E, Marais T, Chatauret N, Benkhelifa-Ziyyat S, Duque S, Ravassard P, Carcenac R, Astord S, Pereira de Moura A, Voit T, Barkats M (2011) Intravenous scAAV9 delivery of a codon-optimized SMN1 sequence rescues SMA mice. *Hum Mol Genet* **20:** 681-693

Donovan J, Copeland PR (2010) The efficiency of selenocysteine incorporation is regulated by translation initiation factors. *J Mol Biol* **400:** 659-664

Eberle AB, Lykke-Andersen S, Muhlemann O, Jensen TH (2009) SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nature Structural & Molecular Biology* **16:** 49-55

Fortner DM, Troy RG, Brow DA (1994) A stem/loop in U6 RNA defines a conformational switch required for pre-mRNA splicing. *Genes Dev* **8:** 221-233

Foust KD, Wang X, McGovern VL, Braun L, Bevan AK, Haidet AM, Le TT, Morales PR, Rich MM, Burghes AH, Kaspar BK (2010) Rescue of the spinal muscular atrophy

phenotype in a mouse model by early postnatal delivery of SMN. *Nature biotechnology* **28:** 271-274

Fraser CS, Doudna JA (2007) Quantitative studies of ribosome conformational dynamics. *Q Rev Biophys* **40:** 163-189

Gabanella F, Butchbach ME, Saieva L, Carissimi C, Burghes AH, Pellizzoni L. (2007) Ribonucleoprotein assembly defects correlate with spinal muscular atrophy severity and preferentially affect a subset of spliceosomal snRNPs. *PLoS One*, Vol. 2, p. e921.

Gallie DR (1998) A tale of two termini: a functional interaction between the termini of an mRNA is a prerequisite for efficient translation initiation. *Gene* **216:** 1-11

Gallie DR (2002) Protein-protein interactions required during translation. *Plant molecular biology* **50:** 949-970

Ghosh T, Peterson B, Tomasevic N, Peculis BA (2004) Xenopus U8 snoRNA Binding Protein Is a Conserved Nuclear Decapping Enzyme. *Mol Cell* **13:** 817-828

Gilbert J, Baker SD, Bowling MK, Grochow L, Figg WD, Zabelina Y, Donehower RC, Carducci MA (2001) A phase I dose escalation and bioavailability study of oral sodium phenylbutyrate in patients with refractory solid tumor malignancies. *Clinical cancer research : an official journal of the American Association for Cancer Research* **7:** 2292-2300

Gingras AC, Raught B, Sonenberg N (1999) eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu Rev Biochem* **68:** 913-963

Gogliotti RG, Cardona H, Singh J, Bail S, Emery C, Kuntz N, Jorgensen M, Durens M, Xia B, Barlow C, Heier CR, Plasterer HL, Jacques V, Kiledjian M, Jarecki J, Rusche J, DiDonato CJ (2013) The DcpS inhibitor RG3039 improves survival, function and motor unit pathologies in two SMA mouse models. *Hum Mol Genet* **22:** 4084-4101

Golembe TJ, Yong J, Dreyfuss G (2005) Specific sequence features, recognized by the SMN complex, identify snRNAs and determine their fate as snRNPs. *Mol Cell Biol* **25:** 10989-11004

Graveley BR (2000) Sorting out the complexity of SR protein functions. *RNA* **6:** 1197-1211

Gu M, Fabrega C, Liu SW, Liu H, Kiledjian M, Lima CD (2004) Insights into the structure, mechanism, and regulation of scavenger mRNA decapping activity. *Mol Cell* **14:** 67-80

Haghighat A, Mader S, Pause A, Sonenberg N (1995) Repression of cap-dependent translation by 4E-binding protein 1: competition with p220 for binding to eukaryotic initiation factor-4E. *EMBO J* **14:** 5701-5709

Haghighat A, Sonenberg N (1997) eIF4G dramatically enhances the binding of eIF4E to the mRNA 5'-cap structure [published erratum appears in J Biol Chem 1997 Nov 14;272(46):29398]. *J Biol Chem* **272:** 21677-21680

Hansen RS, Gartler SM, Scott CR, Chen SH, Laird CD (1992) Methylation analysis of CGG sites in the CpG island of the human FMR1 gene. *Human molecular genetics* **1:** 571-578

Harper JE, Manley JL (1991) A novel protein factor is required for use of distal alternative 5' splice sites in vitro. *Mol Cell Biol* **11:** 5945-5953

Hastings ML, Berniac J, Liu YH, Abato P, Jodelka FM, Barthel L, Kumar S, Dudley C, Nelson M, Larson K, Edmonds J, Bowser T, Draper M, Higgins P, Krainer AR (2009) Tetracyclines that promote SMN2 exon 7 splicing as therapeutics for spinal muscular atrophy. *Science translational medicine* **1:** 5ra12

Hastings ML, Krainer AR (2001) Pre-mRNA splicing in the new millennium. *Current opinion in cell biology* **13:** 302-309

Hebert MD, Szymczyk PW, Shpargel KB, Matera AG (2001) Coilin forms the bridge between Cajal bodies and SMN, the spinal muscular atrophy protein. *Genes Dev* **15:** 2720-2729

Ho CK, Martins A, Shuman S (2000) A yeast-based genetic system for functional analysis of viral mRNA capping enzymes. *J Virol* **74:** 5486-5494

Hofmann Y, Wirth B (2002) hnRNP-G promotes exon 7 inclusion of survival motor neuron (SMN) via direct interaction with Htra2-beta1. *Hum Mol Genet* **11:** 2037-2049

Houseley J, LaCava J, Tollervey D (2006) RNA-quality control by the exosome. *Nat Rev Mol Cell Biol* **7:** 529-539

Hsu CL, Stevens A (1993) Yeast cells lacking 5'-->3' exoribonuclease 1 contain mRNA species that are poly(A) deficient and partially lack the 5' cap structure. *Mol Cell Biol* **13:** 4826-4835

Hu W, Petzold C, Coller J, Baker KE (2010) Nonsense-mediated mRNA decapping occurs on polyribosomes in Saccharomyces cerevisiae. *Nat Struct Mol Biol* **17:** 244-247

Hua Y, Sahashi K, Hung G, Rigo F, Passini MA, Bennett CF, Krainer AR (2010) Antisense correction of SMN2 splicing in the CNS rescues necrosis in a type III SMA mouse model. *Genes Dev* **24:** 1634-1644

Huntzinger E, Kashima I, Fauser M, Sauliere J, Izaurralde E (2008) SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. *Rna-a Publication of the Rna Society* **14:** 2609-2617

Imataka H, Gradi A, Sonenberg N (1998) A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. *EMBO J* **17:** 7480-7489

Izaurralde E, Lewis J, McGuigan C, Jankowska M, Darzynkiewicz E, Mattaj IW (1994) A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell* **78:** 657-668

Jackson RJ, Hellen CU, Pestova TV (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* **11:** 113-127

Jarecki J, Chen X, Bernardino A, Coovert DD, Whitney M, Burghes A, Stack J, Pollok BA (2005) Diverse small-molecule modulators of SMN expression found by high-throughput compound screening: early leads towards a therapeutic for spinal muscular atrophy. *Hum Mol Genet* **14:** 2003-2018

Kieft JS (2008) Viral IRES RNA structures and ribosome interactions. *Trends Biochem Sci* **33:** 274-283

Kolupaeva VG, Pestova TV, Hellen CU, Shatsky IN (1998) Translation eukaryotic initiation factor 4G recognizes a specific structural element within the internal ribosome entry site of encephalomyocarditis virus RNA. *J Biol Chem* **273:** 18599-18604

Korneeva NL, Lamphear BJ, Hennigan FL, Rhoads RE (2000) Mutually cooperative binding of eukaryotic translation initiation factor (eIF) 3 and eIF4A to human eIF4G-1. *J Biol Chem* **275:** 41369-41376

Koromilas AE, Lazaris-Karatzas A, Sonenberg N (1992) mRNAs containing extensive secondary structure in their 5' non-coding region translate efficiently in cells overexpressing initiation factor eIF-4E. *EMBO J* **11:** 4153-4158

Kozak M (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44:** 283-292

Kramer A (1996) The structure and function of proteins involved in mammalian pre-mRNA splicing. *Annu Rev Biochem* **65:** 367-409

Kremer EJ, Pritchard M, Lynch M, Yu S, Holman K, Baker E, Warren ST, Schlessinger D, Sutherland GR, Richards RI (1991) Mapping of DNA Instability at the Fragile-X to a Trinucleotide Repeat Sequence P(Ccg)N. *Science* **252:** 1711-1714

Lall S, Piano F, Davis RE (2005) Caenorhabditis elegans decapping proteins: localization and functional analysis of Dcp1, Dcp2, and DcpS during embryogenesis. *Mol Biol Cell* **16:** 5880-5890

Lebreton A, Tomecki R, Dziembowski A, Seraphin B (2008) Endonucleolytic RNA cleavage by a eukaryotic exosome. *Nature* **456:** 993-996

Lefebvre S, Burglen L, Reboullet S, Clermont O, Burlet P, Viollet L, Benichou B, Cruaud C, Millasseau P, Zeviani M, et al. (1995) Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* **80:** 155-165

Leonard H, Wen X (2002) The epidemiology of mental retardation: challenges and opportunities in the new millennium. *Mental retardation and developmental disabilities research reviews* **8:** 117-134

Li Y, Dai J, Song M, Fitzgerald-Bocarsly P, Kiledjian M (2012) Dcp2 Decapping Protein Modulates mRNA Stability of the Critical Interferon Regulatory Factor (IRF) IRF-7. *Mol Cell Biol* **32:** 1164-1172

Li Y, Ho ES, Gunderson SI, Kiledjian M (2009) Mutational analysis of a Dcp2-binding element reveals general enhancement of decapping by 5'-end stem-loop structures. *Nucleic Acids Res* **37:** 2227-2237

Li Y, Song MG, Kiledjian M (2008) Transcript-specific decapping and regulated stability by the human Dcp2 decapping protein. *Mol Cell Biol* **28:** 939-948

Lima CD, Klein MG, Hendrickson WA (1997) Structure-based analysis of catalysis and substrate definition in the HIT protein family. *Science* **278:** 286-290.

Liu H, Kiledjian M (2005) Scavenger decapping activity facilitates 5' to 3' mRNA decay. *Mol Cell Biol* **25:** 9764-9772

Liu H, Rodgers ND, Jiao X, Kiledjian M (2002) The scavenger mRNA decapping enzyme DcpS is a member of the HIT family of pyrophosphatases. *EMBO J* **21:** 4699-4708

Liu Q, Dreyfuss G (1996) A novel nuclear structure containing the survival of motor neurons protein. *EMBO J* **15:** 3555-3565

Liu Q, Greimann JC, Lima CD (2006) Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell* **127:** 1223-1237

Liu SW, Jiao X, Welch S, Kiledjian M (2008a) Analysis of mRNA decapping. *Methods Enzymol* **448:** 3-21

Liu SW, Rajagopal V, Patel SS, Kiledjian M (2008b) Mechanistic and kinetic analysis of the DcpS scavenger decapping enzyme. *J Biol Chem* **283:** 16427-16436

Lopez-Lastra M, Rivas A, Barria MI (2005) Protein synthesis in eukaryotes: the growing biological relevance of cap-independent translation initiation. *Biological research* **38:** 121-146

Lorson CL, Hahnen E, Androphy EJ, Wirth B (1999) A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proc Natl Acad Sci U S A* **96:** 6307-6311

Lotti F, Imlach WL, Saieva L, Beck ES, Hao le T, Li DK, Jiao W, Mentis GZ, Beattie CE, McCabe BD, Pellizzoni L (2012) An SMN-dependent U12 splicing event essential for motor circuit function. *Cell* **151:** 440-454

Lunn MR, Stockwell BR (2005) Chemical genetics and orphan genetic diseases. *Chem Biol* **12:** 1063-1073

Lykke-Andersen J (2002) Identification of a human decapping complex associated with hUpf proteins in nonsense-mediated decay. *Mol Cell Biol* **22:** 8114-8121

Massenet S, Pellizzoni L, Paushkin S, Mattaj IW, Dreyfuss G (2002) The SMN complex is associated with snRNPs throughout their cytoplasmic assembly pathway. *Mol Cell Biol* **22:** 6533-6541

Matlin AJ, Clark F, Smith CW (2005) Understanding alternative splicing: towards a cellular code. *Nat Rev Mol Cell Biol* **6:** 386-398

Merrick WC (2004) Cap-dependent and cap-independent translation in eukaryotic systems. *Gene* **332:** 1-11

Mitchell P, Tollervey D (2003) An NMD pathway in yeast involving accelerated deadenylation and exosome-mediated 3'-->5' degradation. *Mol Cell* **11:** 1405-1413

Monani UR, Lorson CL, Parsons DW, Prior TW, Androphy EJ, Burghes AH, McPherson JD (1999a) A single nucleotide difference that alters splicing patterns distinguishes the SMA gene SMN1 from the copy gene SMN2. *Hum Mol Genet* **8:** 1177-1183

Monani UR, McPherson JD, Burghes AH (1999b) Promoter analysis of the human centromeric and telomeric survival motor neuron genes (SMNC and SMNT). *Biochim Biophys Acta* **1445:** 330-336

Monani UR, Sendtner M, Coovert DD, Parsons DW, Andreassi C, Le TT, Jablonka S, Schrank B, Rossol W, Prior TW, Morris GE, Burghes AH (2000) The human centromeric survival motor neuron gene (SMN2) rescues embryonic lethality in Smn(-/-) mice and results in a mouse with spinal muscular atrophy. *Hum Mol Genet* **9:** 333-339

Mouaikel J, Narayanan U, Verheggen C, Matera AG, Bertrand E, Tazi J, Bordonne R (2003) Interaction between the small-nuclear-RNA cap hypermethylase and the spinal muscular atrophy protein, survival of motor neuron. *EMBO Rep* **4:** 616-622

Muhlrad D, Decker CJ, Parker R (1995) Turnover mechanisms of the stable yeast PGK1 mRNA. *Mol Cell Biol* **15:** 2145-2156

Nagarajan VK, Jones CI, Newbury SF, Green PJ (2013) XRN 5'-->3' exoribonucleases: structure, mechanisms and functions. *Biochim Biophys Acta* **1829:** 590-603

Najmabadi H, Motazacker MM, Garshasbi M, Kahrizi K, Tzschach A, Chen W, Behjati F, Hadavi V, Nieh SE, Abedini SS, Vazifehmand R, Firouzabadi SG, Jamali P, Falah M, Seifati SM, Gruters A, Lenzner S, Jensen LR, Ruschendorf F, Kuss AW, Ropers HH (2007) Homozygosity mapping in consanguineous families reveals extreme heterogeneity of non-syndromic autosomal recessive mental retardation and identifies 8 novel gene loci. *Human genetics* **121:** 43-48

Narayanan U, Ospina JK, Frey MR, Hebert MD, Matera AG (2002) SMN, the spinal muscular atrophy protein, forms a pre-import snRNP complex with snurportin1 and importin beta. *Hum Mol Genet* **11:** 1785-1795

Naryshkin NA, Weetall M, Dakka A, Narasimhan J, Zhao X, Feng Z, Ling KK, Karp GM, Qi H, Woll MG, Chen G, Zhang N, Gabbeta V, Vazirani P, Bhattacharyya A, Furia B, Risher N, Sheedy J, Kong R, Ma J, Turpoff A, Lee CS, Zhang X, Moon YC, Trifillis P, Welch EM, Colacino JM, Babiak J, Almstead NG, Peltz SW, Eng LA, Chen KS, Mull JL, Lynes MS, Rubin LL, Fontoura P, Santarelli L, Haehnke D, McCarthy KD, Schmucki R, Ebeling M, Sivaramakrishnan M, Ko CP, Paushkin SV, Ratni H, Gerlach I, Ghosh A, Metzger F (2014) Motor neuron disease. SMN2 splicing modifiers improve motor function and longevity in mice with spinal muscular atrophy. *Science* **345:** 688-693

Ng C, Shboul M, Taverniti V, Bonnard C, Ascia E, Nelson SF, Al-Raqad M, Altawalbeh S, Séraphin B, Reversade B (2015a) Loss of the Scavenger mRNA Decapping Enzyme DCPS Causes Syndromic Intellectual Disability with Neuromuscular Defects. *Hum Mol Genet*

Ng CK, Shboul M, Taverniti V, Bonnard C, Lee H, Eskin A, Nelson SF, Al-Raqad M, Altawalbeh S, Seraphin B, Reversade B (2015b) Loss of the scavenger mRNA decapping enzyme DCPS causes syndromic intellectual disability with neuromuscular defects. *Human molecular genetics*

Oberle I, Rousseau F, Heitz D, Kretz C, Devys D, Hanauer A, Boue J, Bertheas MF, Mandel JL (1991) Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. *Science* **252:** 1097-1102

Orban TI, Izaurralde E (2005a) Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. *Rna-a Publication of the Rna Society* **11:** 459-469

Orban TI, Izaurralde E (2005b) Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. *RNA* **11:** 459-469

Passini MA, Bu J, Roskelley EM, Richards AM, Sardi SP, O'Riordan CR, Klinger KW, Shihabuddin LS, Cheng SH (2010) CNS-targeted gene therapy improves survival and motor function in a mouse model of spinal muscular atrophy. *J Clin Invest* **120:** 1253-1264

Patrizi AL, Tiziano F, Zappata S, Donati MA, Neri G, Brahe C (1999) SMN protein analysis in fibroblast, amniocyte and CVS cultures from spinal muscular atrophy patients and its relevance for diagnosis. *European journal of human genetics : EJHG* **7:** 301-309

Pause A, Belsham GJ, Gingras AC, Donze O, Lin TA, Lawrence JC, Jr., Sonenberg N (1994) Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function. *Nature* **371:** 762-767

Pellizzoni L (2007) Chaperoning ribonucleoprotein biogenesis in health and disease. *EMBO Rep* **8:** 340-345

Pellizzoni L, Yong J, Dreyfuss G (2002) Essential role for the SMN complex in the specificity of snRNP assembly. *Science* **298:** 1775-1779

Pestova TV, Hellen CU (2003) Translation elongation after assembly of ribosomes on the Cricket paralysis virus internal ribosomal entry site without initiation factors or initiator tRNA. *Genes Dev* **17:** 181-186

Pestova TV, Hellen CU, Shatsky IN (1996) Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. *Mol Cell Biol* **16:** 6859-6869

Pestova TV, Kolupaeva VG, Lomakin IB, Pilipenko EV, Shatsky IN, Agol VI, Hellen CU (2001) Molecular mechanisms of translation initiation in eukaryotes. *Proc Natl Acad Sci U S A* **98:** 7029-7036

Pestova TV, Shatsky IN, Fletcher SP, Jackson RJ, Hellen CU (1998) A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal

translation initiation of hepatitis C and classical swine fever virus RNAs. *Genes Dev* **12:** 67-83

Piccirillo C, Khanna R, Kiledjian M (2003) Functional characterization of the mammalian mRNA decapping enzyme hDcp2. *Rna* **9:** 1138-1147

Pickering BM, Willis AE (2005) The implications of structured 5' untranslated regions on translation and disease. *Seminars in cell & developmental biology* **16:** 39-47

Pieretti M, Zhang FP, Fu YH, Warren ST, Oostra BA, Caskey CT, Nelson DL (1991) Absence of expression of the FMR-1 gene in fragile X syndrome. *Cell* **66:** 817-822

Porensky PN, Mitrpant C, McGovern VL, Bevan AK, Foust KD, Kaspar BK, Wilton SD, Burghes AH (2012) A single administration of morpholino antisense oligomer rescues spinal muscular atrophy in mouse. *Hum Mol Genet* **21:** 1625-1638

Poulin F, Gingras AC, Olsen H, Chevalier S, Sonenberg N (1998) 4E-BP3, a new member of the eukaryotic initiation factor 4E-binding protein family. *J Biol Chem* **273:** 14002-14007

Ranu RS, London IM (1979) Regulation of protein synthesis in rabbit reticulocyte lysates: additional initiation factor required for formation of ternary complex (eIF-2.GTP.Met-tRNAf) and demonstration of inhibitory effect of heme-regulated protein kinase. *Proc Natl Acad Sci U S A* **76:** 1079-1083

Rigo F, Chun SJ, Norris DA, Hung G, Lee S, Matson J, Fey RA, Gaus H, Hua Y, Grundy JS, Krainer AR, Henry SP, Bennett CF (2014) Pharmacology of a central nervous system delivered 2'-O-methoxyethyl-modified survival of motor neuron splicing oligonucleotide in mice and nonhuman primates. *J Pharmacol Exp Ther* **350:** 46-55

Roeleveld N, Zielhuis GA, Gabreels F (1997) The prevalence of mental retardation: a critical review of recent literature. *Developmental medicine and child neurology* **39:** 125-132

Ropers HH (2006) X-linked mental retardation: many genes for a complex disorder. *Current opinion in genetics & development* **16:** 260-269

Ropers HH (2010) Genetics of early onset cognitive impairment. *Annual review of genomics and human genetics* **11:** 161-187

Ropers HH, Hamel BC (2005) X-linked mental retardation. *Nature reviews Genetics* **6:** 46-57

Ruby SW, Abelson J (1988) An early hierarchic role of U1 small nuclear ribonucleoprotein in spliceosome assembly. *Science* **242:** 1028-1035

Ruskin B, Zamore PD, Green MR (1988) A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. *Cell* **52:** 207-219

Salehi Z, Geffers L, Vilela C, Birkenhager R, Ptushkina M, Berthelot K, Ferro M, Gaskell S, Hagan I, Stapley B, McCarthy JE (2002) A nuclear protein in Schizosaccharomyces pombe with homology to the human tumour suppressor Fhit has decapping activity. *Mol Microbiol* **46:** 49-62

Sanchez R, Sali A (2000) Comparative protein structure modeling. Introduction and practical examples with modeller. *Methods Mol Biol* **143:** 97-129

Schmid M, Jensen TH (2008) The exosome: a multipurpose RNA-decay machine. *Trends in biochemical sciences* **33:** 501-510

Schmitter D, Filkowski J, Sewer A, Pillai RS, Oakeley EJ, Zavolan M, Svoboda P, Filipowicz W (2006) Effects of Dicer and Argonaute down-regulation on mRNA levels in human HEK293 cells. *Nucleic Acids Res*

Seal SN, Schmidt A, Marcus A (1983) Eukaryotic initiation factor 4A is the component that interacts with ATP in protein chain initiation. *Proc Natl Acad Sci U S A* **80:** 6562-6565

Seraphin B (1992) The HIT protein family: a new family of proteins present in prokaryotes, yeast and mammals. *DNA Seq* **3:** 177-179

Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen TS, Heckl D, Ebert BL, Root DE, Doench JG, Zhang F (2014) Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343:** 84-87

Shen V, Kiledjian M (2006) A view to a kill: structure of the RNA exosome. *Cell* **127:** 1093-1095

Shen V, Liu H, Liu SW, Jiao X, Kiledjian M (2008) DcpS scavenger decapping enzyme can modulate pre-mRNA splicing. *RNA* **14:** 1132-1142

Sheth U, Parker R (2003) Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science* **300:** 805-808

Shintani T, Klionsky DJ (2004) Autophagy in health and disease: a double-edged sword. *Science* **306:** 990-995

Shukla S, Parker R (2014) Quality control of assembly-defective U1 snRNAs by decapping and 5'-to-3' exonucleolytic digestion. *Proc Natl Acad Sci U S A* **111:** E3277-3286

Singh J, Salcius M, Liu SW, Staker BL, Mishra R, Thurmond J, Michaud G, Mattoon DR, Printen J, Christensen J, Bjornsson JM, Pollok BA, Kiledjian M, Stewart L, Jarecki J, Gurney ME (2008) DcpS as a therapeutic target for spinal muscular atrophy. *ACS Chem Biol* **3:** 711-722

Sinturel F, Brechemier-Baey D, Kiledjian M, Condon C, Benard L (2012) Activation of 5'-3' exoribonuclease Xrn1 by cofactor Dcs1 is essential for mitochondrial function in yeast. *Proc Natl Acad Sci U S A* **109:** 8264-8269

Song MG, Bail S, Kiledjian M (2013) Multiple Nudix family proteins possess mRNA decapping activity. *RNA* **19:** 390-399

Song MG, Li Y, Kiledjian M (2010) Multiple mRNA decapping enzymes in mammalian cells. *Mol Cell* **40:** 423-432

Spellman R, Rideau A, Matlin A, Gooding C, Robinson F, McGlincy N, Grellscheid SN, Southby J, Wollerton M, Smith CW (2005) Regulation of alternative splicing by PTB and associated factors. *Biochem Soc Trans* **33:** 457-460

Steiger M, Carr-Schmid A, Schwartz DC, Kiledjian M, Parker R (2003) Analysis of recombinant yeast decapping enzyme. *RNA* **9:** 231-238

Sterne-Weiler T, Howard J, Mort M, Cooper DN, Sanford JR (2011) Loss of exon identity is a common mechanism of human inherited disease. *Genome Res* **21:** 1563-1571

Stevens A, Poole TL (1995) 5'-exonuclease-2 of Saccharomyces cerevisiae. Purification and features of ribonuclease activity with comparison to 5'-exonuclease-1. *J Biol Chem* **270:** 16063-16069

Sumner CJ, Huynh TN, Markowitz JA, Perhac JS, Hill B, Coovert DD, Schussler K, Chen X, Jarecki J, Burghes AH, Taylor JP, Fischbeck KH (2003) Valproic acid increases SMN levels in spinal muscular atrophy patient cells. *Annals of neurology* **54:** 647-654

Svitkin YV, Herdy B, Costa-Mattioli M, Gingras AC, Raught B, Sonenberg N (2005) Eukaryotic translation initiation factor 4E availability controls the switch between cap-dependent and internal ribosomal entry site-mediated translation. *Mol Cell Biol* **25:** 10556-10565

Taggart AJ, DeSimone AM, Shih JS, Filloux ME, Fairbrother WG (2012) Large-scale mapping of branchpoints in human pre-mRNA transcripts in vivo. *Nature structural & molecular biology* **19:** 719-721

Takahashi S, Araki Y, Sakuno T, Katada T (2003) Interaction between Ski7p and Upf1p is required for nonsense-mediated 3 '-to-5 ' mRNA decay in yeast. *Embo Journal* **22:** 3951-3959

Tan D, Zhou M, Kiledjian M, Tong L (2014) The ROQ domain of Roquin recognizes mRNA constitutive-decay element and double-stranded RNA. *Nat Struct Mol Biol* **21:** 679-685

Taverniti V, Seraphin B (2014) Elimination of cap structures generated by mRNA decay involves the new scavenger mRNA decapping enzyme Aph1/FHIT together with DcpS. *Nucleic Acids Res*

Taylor MJ, Peculis BA (2008) Evolutionary conservation supports ancient origin for Nudt16, a nuclear-localized, RNA-binding, RNA-decapping enzyme. *Nucleic Acids Res* **36:** 6021-6034

Thach RE, Dewey KF, Brown JC, Doty P (1966) Formylmethionine codon AUG as an initiator of polypeptide synthesis. *Science* **153:** 416-418

Thurmond J, Butchbach ME, Palomo M, Pease B, Rao M, Bedell L, Keyvan M, Pai G, Mishra R, Haraldsson M, Andresson T, Bragason G, Thosteinsdottir M, Bjornsson JM, Coovert DD, Burghes AH, Gurney ME, Singh J (2008) Synthesis and Biological

Evaluation of Novel 2,4-Diaminoquinazoline Derivatives as SMN2 Promoter Activators for the Potential Treatment of Spinal Muscular Atrophy. *J Med Chem* **51:** 449-469

van Dijk E, Cougot N, Meyer S, Babajko S, Wahle E, Seraphin B (2002) Human Dcp2: a catalytically active mRNA decapping enzyme located in specific cytoplasmic structures. *EMBO J* **21:** 6915-6924.

van Dijk EL, Chen CL, d'Aubenton-Carafa Y, Gourvennec S, Kwapisz M, Roche V, Bertrand C, Silvain M, Legoix-Ne P, Loeillet S, Nicolas A, Thermes C, Morillon A (2011) XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. *Nature* **475:** 114-117

van Hoof A, Frischmeyer PA, Dietz HC, Parker R (2002) Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science* **295:** 2262-2264.

Van Meerbeke JP, Gibbs RM, Plasterer HL, Miao W, Feng Z, Lin MY, Rucki AA, Wee CD, Xia B, Sharma S, Jacques V, Li DK, Pellizzoni L, Rusche JR, Ko CP, Sumner CJ (2013a) The DcpS inhibitor RG3039 improves motor function in SMA mice. *Hum Mol Genet* **22:** 4074-4083

Van Meerbeke JP, Gibbs RM, Plasterer HL, Miao W, Feng Z, Lin MY, Rucki AA, Wee CD, Xia B, Sharma S, Jacques V, Li DK, Pellizzoni L, Rusche JR, Ko CP, Sumner CJ (2013b) The DcpS inhibitor RG3039 improves motor function in SMA mice. *Hum Mol Genet*

Vincent A, Heitz D, Petit C, Kretz C, Oberle I, Mandel JL (1991) Abnormal pattern detected in fragile-X patients by pulsed-field gel electrophoresis. *Nature* **349:** 624-626

Wang Z, Jiao X, Carr-Schmid A, Kiledjian M (2002) The hDcp2 protein is a mammalian mRNA decapping enzyme. *Proc Natl Acad Sci U S A* **99:** 12663-12668.

Wang Z, Kiledjian M (2001) Functional Link between the Mammalian Exosome and mRNA Decapping. *Cell* **107:** 751-762.

Wee CD, Kong L, Sumner CJ (2010) The genetics of spinal muscular atrophies. *Curr Opin Neurol* **23:** 450-458

Williams JH, Schray RC, Patterson CA, Ayitey SO, Tallent MK, Lutz GJ (2009) Oligonucleotide-mediated survival of motor neuron protein expression in CNS improves phenotype in a mouse model of spinal muscular atrophy. *J Neurosci* **29:** 7633-7638

Yu S, Pritchard M, Kremer E, Lynch M, Nancarrow J, Baker E, Holman K, Mulley JC, Warren ST, Schlessinger D, et al. (1991) Fragile X genotype characterized by an unstable region of DNA. *Science* **252:** 1179-1181

Zhou M, Bail S, Plasterer HL, Rusche J, Kiledjian M (2015) DcpS is a transcript-specific modulator of RNA in mammalian cells. *RNA*

Zhuang YA, Goldstein AM, Weiner AM (1989) UACUAAC is the preferred branch site for mammalian mRNA splicing. *Proc Natl Acad Sci U S A* **86:** 2752-2756