Defining the role of RNA-binding proteins and their regulation by arginine methylation

by

Jingwen Song

Department of Medicine

Division of Experimental Medicine

McGill University, Montréal, Québec, Canada

September 2016

A Thesis Submitted to the Faculty of Graduate Studies and Research in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy.

ABSTRACT

RNA-binding proteins maintain RNA metabolism homeostasis by fulfilling their functional roles in the regulation of basic cellular processes. Bioinformatic searches demonstrate that a large number of RNA-binding proteins contain the glycine-arginine rich motif (GAR motif or RRG/RG motif), which is a major target of arginine methylation. Arginine methylation is a fundamental post-translational modification catalyzed by protein arginine methyltransferases (PRMTs). Although there are >1000 RGG/RG motif-containing proteins, the versatile roles of arginine methylation in the regulation of RNA metabolism still remain to be elucidated. The goal of my thesis was to identify novel spliced targets of a well-established RGG/RG motif-containing protein, Sam68, and to characterize the novel function of a newly identified arginine methylated protein named Aven.

The first part of my thesis focuses on Sam68, a member of the STAR (Signal Transduction and Activation of RNA) family of RNA-binding proteins. It not only serves as a hub protein in signaling transduction, but also plays a role in RNA metabolism, such as in pre-mRNA alternative splicing and mRNA translation. Sam68 is required for cells to undergo proliferation, differentiation, and apoptosis. Herein, we report that Sam68 contributes to adipogenesis by regulating alternative splicing of the ribosomal protein S6 kinase beta-1 (S6K1), thereby preventing the production of a small S6K1 isoform (S6K1-iso2). S6K1-iso2 produces a nuclear protein called p31, which we have shown is a potent adipogenesis suppressor. This is consistent with previous findings showing that Sam68 null mice are lean and protected from dietary induced obesity. Mechanistically, we show that Sam68 binds to the intron 6 of S6K1 pre-mRNA, and competes with the function of serine/arginine-rich splicing factor 1 (SRSF1) in the splicing of S6K1. These findings define a new splicing target of Sam68 in adipogenesis and cancer.

In the second part of my studies, I investigated the role of Aven, a cell death regulator containing an N-terminal RGG/RG motif. We demonstrated the methylation of the RGG/RG motif both *in vitro* and *in vivo*. The Aven RGG/RG motif mediates the binding to G-quadruplex (G4)-containing RNAs and their polysomal associations. Aven bound G4

motifs within the open reading frames (ORFs) of mixed lineage leukemia (MLL) family proteins MLL1 and MLL4 mRNAs to promote their translation. These observations suggest a new role for arginine methylation in the regulation of mRNA translation. These finding have implications for T-cell acute lymphoblastic leukemia, where Aven's expression is known to be up-regulated. Furthermore, we identify BRISC (The BRCC36 isopeptidase-containing complex) as an Aven interacting complex. We show that Aven stabilizes the BRISC complex in polysomes to control polysomal Lys63 (K63) ubiquitination homeostasis. Under oxidative stress, Aven dissociates Abro1 from the BRISC complex thereby decreasing the deubiquitinating activity of BRCC36, leading to transient accumulation of K63 ubiquitin in polysomes. This increase in K63 influences mRNA translation of a subset of stress-responsive genes, providing a survival advantage to cells in response to oxidative stress. These findings define a novel survival role for Aven and provide a new link between K63 ubiquitination and mRNA translation.

In conclusion, my thesis identifies new signaling pathways of two important RNA-binding proteins, Sam68 and Aven. I show how Sam68 is required for adipogenesis to regulate cellular metabolism, and I define several new modes of cell survival processes regulated by Aven. My work has implications for obesity, diabetes and cancer.

SOMMAIRE

Les protéines liant l'ARN sont importantes dans la regulation des processus cellulaires. La méthylation des protéines s'est révélée être une modification post-traductionnelle importante. Bien que les analyses bioinformatiques aient permis de repérer plus de 1000 protéines contenant des motifs RGG/RG, le rôle que jouent les arginine méthyltransférases (PRMTs) et leurs cibles dans ces processus n'a pas encore été élucidé. Par conséquent, il est d'une grande importance de caractériser de nouvelles cibles et de nouvelles fonctions pour les méthyltransférases. Mes projets avaient pour but de déterminer une nouvelle fonction pour Sam68, une protéine contenant des motifs RGG/RG bien établie, et de caractériser les fonctions d'une cible de la méthylation de l'arginine récemment établie : la protéine Aven.

La première partie de ma thèse est centrée sur la protéine Sam68, un membre de la famille des protéines activatrices de la traduction du signal de l'ARN (dites STAR pour signal transduction activator of RNA). Cette protéine joue pas qu'un rôle central dans la traduction des signaux, mais aussi dans le métabolisme de l'ARN, y compris la transcription, l'épissage alternatif et le transport de l'ARN. En se couplant à ses cibles d'ARN, Sam68 participe à des processus physiologiques importants comme l'apoptose, la prolifération et la différentiation. Dans les présents travaux, nous rapportons que Sam68 participe à l'adipogenèse en régulant l'épissage alternatif de la protéine S6K1, ce qui empêche la production d'une petite isoforme de S6K1 (S6K1-iso2), qui, comme nous l'avons montré, est un inhibiteur puissant de l'adipogenèse. Ces résultats sont conséquents avec des résultats antérieurs montrant que les souris nulles Sam68 peuvent exprimer un phénotype mince. Sur le plan des mécanismes, nous montrons que Sam68 se lie à l'intron 6 du pré-ARNm de S6K1 en compétition avec le facteur SRSF1 pour la régulation de l'épissage de S6K1. Ces résultats permettent de caractériser une nouvelle cible de la protéine Sam68 et de fournir une description détaillée de son mode d'action dans la régulation de l'épissage alternatif.

Dans la deuxième partie de mes travaux, nous avons étudié le rôle de la protéine Aven. Aven est un régulateur de la mort cellulaire qui, dans sa portion N-terminale, contient un motif RGG/RG dont la fonction n'est pas encore connue. Nous avons montré in vivo et in

vitro que le motif RGG/RG est méthylé par la protéine PRMT1. Nous avons observé que le motif RGG/RG est responsable de la liaison d'Aven aux ARN adoptant une structure G-quadruplex et de son association avec les polysomes. En particulier, Aven se lie aux phases ouvertes de lecture des ARNm MLL1 et MLL4 de protéines de la famille MLL (pour mixed lineage leukemia) et stimule leur traduction en recrutant une hélicase de l'ARN appelée DHX36. Par conséquent, Aven et PRMT1 contribuent à la prolifération des cellules leucémiques. Ces résultats laissent entrevoir un nouveau mode d'action pour la méthylation de l'arginine dans la régulation de la traduction de l'ARNm. Nous avons également étudié le rôle de la protéine Aven sous conditions de stress. Nous avons montré qu'Aven est une composante régulatrice du complexe BRISC (pour BRCC36 isopeptidase containing complex), et qu'elle stabilise le complexe dans les polysomes afin de maintenir l'homéostasie de l'ubiquitination polysomale liée à la protéine K63 (Lys63). Sous des conditions de stress, Aven se dissocie du complexe BRISC, ce qui diminue l'activité « désubiquitinante » de celui-ci et mène à l'accumulation transitoire de K63 dans les fractions de polysomes. L'augmentation anormale des taux de K63 dans les polysomes perturbe la traduction de l'ARNm d'un sous-groupe de gènes réagissant au stress. De fait, nous avons découvert la présence d'un rapport entre l'ubiquitination liée à la protéine K63 et la traduction de l'ARNm.

En somme, l'ensemble de mes travaux caractérise plusieurs substrats de protéines méthyltransférases et montre leurs rôles dans l'épissage alternatif et la traduction de l'ARNm, tant sous des conditions normales que sous des conditions de stress. Ces travaux pourraient avoir de l'importance pour plusieurs types de maladies, dont le cancer.

PREFACE

As chapters of this thesis, I will include the text and figures of three original research manuscripts and a review article that have been published, submitted or in preparation for publication. Each of these chapters (Chapters 2, 3, 4 and 5) contains its own summary, introduction, materials and methods, results, discussion, and references sections. A general introduction and literature review will be presented in Chapter 1, whereas a final discussion will be included in Chapter 6.

Papers included in this Thesis:

Chapter 2 **Song**, **J.**, Perreault, JP., Topisirovic, I., and Richard, S. (2016). RNA G-quadruplexes and their potential regulatory roles in translation (Translation. 4 (2): e1244034. doi:10.1080/21690731.2016.1244031)

Chapter 3 **Song J.**, and Richard S. (2015). Sam68 regulates S6K1 alternative splicing during adipogenesis. Molecular & Cellular Biology. 35(11):1926-39. doi: 10.1128/MCB.01488-14

Chapter 4 Thandapani P.*, **Song J.***, Gandin V., Cai, Y., Rouleau S.G., Garant J. M., Boisvert F., Yu, Z. Perreault J., Topisirovic, I., Richard S. (2015). Aven recognition of RNA G-quadruplexes regulates translation of the Mixed Lineage Leukemia proto-oncogenes. Elife. doi: 10.7554/eLife.06234 **(*authors contributed equally)**

Chapter 5 **Song**, **J.**, Yu, Z., Cai, Y., Topisirovic, I., and Richard, S. (2016).Polysome-associated Aven/BRISC complex protects against oxidative stress-induced cell death by selectively modulating K63-linked ubiquitination and mRNA translation (manuscript in preparation for submission)

Contribution of Authors:

The candidate performed most of the research presented in this thesis and wrote all of the included manuscripts with support and assistance from Dr. Stéphane Richard. The contribution of other authors to this work is described below:

In Chapter 2, Dr. Jean-Pierre Perreault and Dr. Ivan Topisirovic contributed to the writing and critical reading of the manuscript.

In Chapter 4, Palaniraja Thandapani performed western blots, co-immunoprecipitation and methylation assay (Figure 4.1B-C&E-F, Figure 4.3, Figure 4.4, Supplementary Figure 4.1-4.3A&B), in vitro RNA binding assay (Figure 4.2A-D, Figure 4.5C-F), proliferation assay (Figure 4.6), and luciferase assay (Figure 4.10). I performed the co-immunoprecipitation and western blots (Figure 4.1D), CLIP assay (Figure 4.5G-H), polysome fractionation experiments, RNA extraction, Real-time RT-qPCR (Supplementary Figure 4.3C, Figure 4.7, Figure 4.8, Figure 4.9), and data quantification (Supplementary Figure 4.3D-H). Dr. Valentina Gandin and Yutian Cai provided expertise in polysome fractionation and help with data analysis. Samuel G Rouleau performed the in-line assays (Figure 4.5I). Jean Michel Garant performed genome-wide searches to identify mRNAs with G4 sequences in ORFs (Supplementary Figure 4.4). Dr. Zhenbao Yu performed the peptide binding assay (Figure 4.2E). Dr. François Michel Boisvert performed the SILAC analysis. Dr. Jean-Pierre Perreault provided G-quadruplex expertise. Dr. Ivan Topisirovic provided expertise in mRNA translation and data analysis.

In Chapter 5, Dr. Zhenbao Yu performed the western blot in Figure 5.1A and contributed to experimental design and critical reading of the manuscript. Yutian Cai assisted with polysome fractionation. Dr. Ivan Topisirovic provided expertise in mRNA translation, data analysis and critical reading of the manuscript.

All studies in this thesis were conducted under the supervision of Dr. Stéphane Richard.

ACKNOWLEDGEMENTS

I would first like to express my gratitude to my supervisor, Dr. Stéphane Richard, for giving me the opportunity to work on the brilliant projects and for teaching me how to be a successful scientist. I have been extremely lucky to have a supervisor who cared about my work, and who responded to my questions and queries promptly.

Thanks to all the past and present members of the Richard laboratory for the deep friendship and for that you make my life here remarkable. In particular, I wish to thank Dr. Zhenbao Yu for valuable suggestions in my projects. Also I would like to express my gratitude to my committee members, Dr. Stéphanie Lehoux, Dr. Rongtuan Lin, Dr. Ivan Topisirovic and Dr. Vahab Soleimani for their direction, dedication and invaluable advice regarding my project. And I would also like to express my deep appreciation to Dr. Ivan Topisirovic, for his insightful advice and willingness to share his knowledge through exciting collaborations.

I would like to thank all the people I encountered at the Lady Davis Institute. I would like to acknowledge the McGill MICRTP program for their financial support for my Ph.D, and Lady Davis Institute as well as Department of Experimental Medicine for the various travel awards.

Lastly, I would like to convey my gratitude to all my family members, especially my parents, for your endless love and constant moral support. I am also grateful to everyone who befriended me in Montreal, without whose friendship and love I never could pursue and complete these years of life and study.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

53BP1	P53 binding protein-1
ABRAXAS	Coiled-Coil Domain-Containing Protein 98
Abro1	Abraxas Brother Protein 1
aDMA/ADMA	Asymmetrically dimethylated arginine
Akt	Protein kinase B
ALL	Acute Lymphoblastic Leukemia
ALS	Amyotrophic lateral sclerosis
AMSH	Associated Molecule With The SH3 Domain Of STAM
AMSH-LP	AMSH-like protease
APC/C	Anaphase-promoting complex/cyclosome
Apaf-1	Apoptotic protease activating factor 1
AR	Androgen receptor
ASK1	Apoptosis signal-regulating kinase 1
ASH2L	ASH2 like histone lysine methyltransferase complex subunit
ATL-D	Ataxia-Telangiectasia like disorder
ATM	Ataxia Telangiectasia Mutated
ATR	ATM kinase related protein
BAD	BCL-2 antagonist of cell death
BAZ1A	Bromodomain Adjacent To Zinc Finger Domain 1A
Bcl	B-cell lymphoma
BRCA1	Breast cancer 1
BRCC36/BRCC3	BRCA1/BRCA2-containing complex, subunit 3
BRD4	Bromodomain Containing 4
BRE	Brain And Reproductive Organ-Expressed
BRISC	BRCC36 isopeptidase containing complex
BRWD3	Bromodomain and WD repeat-containing protein 3
BRG1	Brahma-related gene-1
BLIMP1	B-lymphocyte-induced maturation protein 1
Btk	Bruton Tyrosine Kinase
CA150	Tcerg1 transcription elongation regulator 1
CARM1	Co-activator methyltransferase
CBP	CREB binding protein
cDNA	Complementary DNA
CDK7	Cyclin-Dependent Kinase 7
CHD1	chromo-ATPase/helicase-DNA-binding
Chk2	Checkpoint kinase 2
COPR5	Coordinator of PRMT5 and differentiation stimulator
CPSF	Cleavage/polyadenylation specificity factor
cIAP	Cellular inhibitor of apoptosis
CIRBP	Cold-inducible Rna-binding protein

CLIP	Cross-linking immunoprecipitation
CTCFL	CCCTC-Binding Factor Like
CtsF	Cleavage stimulation factor
DART	Drosophila arginine methyltransferase
DDAH	Dimethylarginine dimethylaminohydrolase
DDR	DNA damage repair
DNA	Deoxyribonucleic acid
DNA-PK	DNA-dependent protein kinase
DNMT1	DNA methyltransferase I
DHX36	DEAH-Box Helicase 36
DSB	Double strand break
DUB	Deubiquitinating enzyme
E2F1	E2F Transcription Factor 1
EBNA	Epstein-Barr nuclear antigen
EEN	SH3 Domain Containing GRB2 Like 1
EGF	Epidermal growth factor
eIF	Eukaryotic initiation factor
ERα	Estrogen receptor
Erk	Extracellular signal-regulated kinase
ES Cells	Embryonic stem cells
ESCRT	Endosomal sorting complexes required for transport
ETO	RUNX1 Translocation Partner 1
EJC	Exon junction complex
EWS	Ewing sarcoma
FEN1	Flap endonuclease I
FMRP	Fragile mental retardation protein
FOXO1	Forkhead box O1
FTD	Frontotemporal dementia
FUS/TLS	Fused in Sarcoma/Translocated in Sarcoma
FXR	The Farnesoid X receptor
GAR	Glycine- and arginine-rich
GAP	GTPase activating protein
GFP	Green fluorescent protein
GLD-1	GermLine development Defective-1
GMP	Guanosine 5'-monophosphate oxidoreductase
GRP33	Glycine rich protein 33
Grb-2	Growth Factor Receptor Bound Protein 2
GRAP	GRB2-related adapter protein
GTP	Guanosine-5'-triphosphate
HECT	Homologous to the E6AP carboxyl terminus
HITS-CLIP	High-throughput sequencing of RNA isolated by crosslinking immunoprecipitation
hnRNP	Heterogeneous nuclear ribonucleoprotein
HOW	Held out wings

HP1β	Heterochromatin protein 1
HR	Homologous recombination repair
HRE	Hexanucleotide repeat expansion
HRMT	HnRNP Arginine Methyltransferase
hBRM	Human Brahma
IFMAR1	Interferon Alpha And Beta Receptor Subunit 1
IgG	Immunoglobulin
IP	Immunoprecipitation
IKK	IkB kinase
JAMM	Jab1/Mov34/Mpr1 Pad1 N-terminal+ (MPN+)
kb	Kilobases
kDa	Kilo Dalton
KH	HnRNP K homology
LC/MS/MS	Liquid chromatography coupled with tandem mass spectrometry
Lck	Lymphocyte-specific protein tyrosine kinase
LSm	Sm-like protein
MAPK	Mitogen-activated protein kinases
MBP	Myelin basic protein
MBNL1	Muscleblind-like protein 1
Merit40	Mediator Of RAP80 Interactions And Targeting Subunit Of 40 KDa
MLL	Mixed lineage leukemia
MMA	Monomethylated arginine
MRE11	Meiotic Recombination Mutant 11
MRN	MRE11/RAD50/NBS1
mRNA	Messenger RNA
mTOR	Mammalian target of rapamycin
MS	Mass spectrometry
MSH2-6	MutS protein homolog2-6
NBS	Nijmegen breakage syndrome
NCK	Non-catalytic region of tyrosine kinase adaptor protein
Npl3	Nuclear protein localization 3
NSD1	Nuclear receptor binding SET domain protein 1
NURD	Nucleosome remodeling deacetylase
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
ORC1	Origin recognition complex 1
ORF	Open reading frame
OTU	Ovarian tumor deubiquitinase
PABP	Poly-A binding protein
PAD	Peptidylarginine deiminase
PAF1	Polymerase associated factor
PAP	PolyA polymerase

PCAF	p300/CREB-binding protein associated factor
PCNA	Proliferating cell nuclear antigen
PGC-1α	Peroxisome proliferator-activated receptor-gamma coactivator
PGM	Proline Glycine motif
pICln	Chlorine channel regulator
PIKK	Phosphatidylinositol-3-kinase like protein kinase family
PI3K	Phosphatidylinositol 3-kinase
PML	Progressive multifocal leukoencephalopathy
POH1 (PSMD14)	26S proteasome non-ATPase regulatory subunit 14
PPARγ	Peroxisome proliferator-activated receptor gamma
PRC2	Polycomb repressive complex 2
PRMT	Protein arginine methyltransferase
PTEF	The positive transcription elongation factor
PTM	Post translational modification
PTK	Protein-tyrosine kinase 6
PKC	Protein kinase C
PLCγ-1	phospholipase C γ1
RNA PolII	RNA polymerase II
RAD50	Radiation sensitive gene 50
Rap80	BRCA1-A complex subunit
RBP	RNA-binding protein
RBBP	Retinoblastoma Binding Protein
RNA	Ribonucleic acid
RNAi	RNA interference
RNF	Ring Finger Protein
RNP	Ribonucleoprotein
RIGI	Retinoic Acid-Inducible Gene I Protein
RING	Really interesting new gene
RUNX1	Runt-related transcription factor 1
S6K1	Ribosomal protein S6 kinase beta-1
Sam68	Src-associated in mitosis of 68kDa
SAP49	Splicing factor 3B subunit 4
SCF	Skp, Cullin and F-box containing complex
SELEX	Systematic evolution of ligands by exponential enrichment
SERBP1	Plasminogen activator inhibitor 1 RNA-binding protein
SF1	The steroidogenic factor 1
SHANK1	SH3 and multiple Ankyrin repeat domain protein 1
SHMT	Serine hydroxymethyltransferase
SLM	Sam68-Like Mammalian Protein 1
SILAC	Stable isotope labelling in cell culture
siRNA	Small interfering RNA

Sm	Small nuclear ribonucleoprotein
SMA	Spinal muscular atrophy
SMN	Survival of motor neuron
SND1	Staphylococcal nuclease domain-containing protein 1
snRNP	Small nuclear ribonucleoparticle
SRSF1	Serine/Arginine-Rich Splicing Factor 1
SPF30	Survival of motor neuron-related-splicing factor 30
STAR	Signal Transduction and Activation of RNA
SWI/SNF	SWItch/Sucrose Non-Fermentable
TAK	TGF-β–Activated Kinase
TAP	Nuclear RNA export factor 1
T/WCL	Total/whole cell lysate
TDRD3	Tudor domain containing protein 3
TNF	Tumor necrosis factor
TRAF	TNF Receptor Associated Factor
Ub	Ubiquitin
UCH	Ubiquitin C-terminal hydrolases
UV	Ultravoilet
U1C	U1 small nuclear ribonucleoprotein C
U2AF	U2 snRNP auxiliary factor
USP	Ubiquitin-specific protease
WDR5	WD repeat-containing proteim 5
YY1	Yin Yang 1
YT521	YTH Domain Containing 1

Chapter 1

Introduction and Literature Review

1.1 Protein arginine methylation

Post-translational modifications (PTM) are covalent enzymatic reactions on proteins that occur after they have been synthesized. These modifications add diversity to the standard amino acids by introducing small functional moieties such as amide, phosphate, acetate and methyl groups (Khoury, Baliban et al. 2011). These modifications are often crucial for regulating protein function, and they have led to the fields of signal transduction and epigenetics. Types of PTMs include serine/threonine/tyrosine phosphorylation; N- and O-linked protein glycosylation; lysine ubiquitination/sumoylation/neddylation; lysine acetylation; lysine/arginine methylation, and proline isomerization (Yang and Seto 2008; Larsen, Sylvestersen et al. 2016).

Arginine is a positively charged amino acid, known to mediate hydrogen bonding and amino-aromatic interactions. Arginine can be methylated in 3 ways in eukaryotes (Bedford and Richard 2005): ω-N^G-monomethylarginine (MMA), ω-N^G,N^G-asymmetric dimethylarginine (aDMA), and ω-N'^G,N^G-symmetric dimethylarginine (sDMA). These modifications are carried out by protein arginine methyltransferases (PRMTs). Methylation does not change the cationic charge, rather it changes the hydrogen bonding and increases hydrophobicity (Fuhrmann and Thompson 2016).

1.1.1 PRMTs

In mammalian cells, there are nine PRMTs referred to as PRMT1 through 9 (Yang and Bedford 2013). They are grouped into type I, type II, and type III enzymes. Type I and II enzymes catalyze the formation of MMA as a biosynthetic intermediate, while PRMT7, a type III enzyme, generates MMA as its final product (Feng, Maity et al. 2013). Type I enzymes (PRMT1, PRMT2, PRMT3, PRMT4 (CARM1), PRMT6 and PRMT8) produce

aDMA, whereas type II enzymes (PRMT 5 and PRMT9) catalyze the formation of sDMA (Yang and Bedford 2013) (Figure 1.1). PRMTs are ubiquitously expressed and regulate many cellular processes (Figure 1.2). Here I will discuss the biochemical properties of the major type I, type II and type III PRMTs. See section 1.1.4 for the biological roles of PRMTs.

PRMT1 is a major methyltransferase catalyzing the formation of >85% methylarginine in mammalian cells (Tang, Frankel et al. 2000). PRMT1 is highly conserved in all eukaryotes and the dimerization of PRMT1 is crucial for AdoMet binding (Zhang and Cheng 2003). The active sites of PRMT1 are essential for its enzymatic activity and are required for its product specificity (Gui, Gathiaka et al. 2014). Intriguingly, loss of PRMT1 increases the global levels of MMA and sDMA, detected by type-specific antibodies, inferring a dynamic interplay between PRMT1 and other PRMT-catalyzed arginine methylation types (Dhar, Vemulapalli et al. 2013). The enzymatic activity of PRMT1 is also regulated by oxidation and alternative splicing, which generates different PRMT1 isoforms with unique substrate specificities (Goulet, Gauvin et al. 2007; Morales, Nitzel et al. 2015).

PRMT1 has hundreds of substrates with major groups of proteins being histones and RNA-binding proteins (Boisvert, Rhie et al. 2005). On histones, PRMT1 is associated with transcriptional activation, since it specifically methylates histone H4 arginine 3 (H4R3), a transcriptional activation mark (Wang, Huang et al. 2001). A proteome-wide analysis searching for arginine methylation sites in large numbers of human proteins revealed that depletion of PRMT1 or PRMT4 increases the RNA-binding function of hnRNPUL1(heterogeneous nuclear ribonucleoprotein U-like protein1), an RNA transport ribonucleoprotein (Larsen, Sylvestersen et al. 2016). Additionally, PRMT1 also methylates R296 of the oncoprotein Ash2L, which is a component of the mammalian histone H3K4 methyltransferase complex (Butler, Zurita-Lopez et al. 2011). Moreover, PRMT1 asymmetrically methylates R30 of RelA, thus preventing the binding of RelA to DNA, and thereby repressing the NF-κB target genes in response to TNFα (Reintjes, Fuchs et al. 2016).

PRMT1 is crucial in mammals, since *prmt1*-knockout mice die at E6.5 (Pawlak, Scherer et al. 2000). *PRMT1*-conditional knockout mice demonstrate that PRMT1 is required for cell viability. PRMT1 plays a key role in DNA damage, checkpoint defects, chromosomal

instability, and the DDR (DNA damage response) pathway (Yu, Chen et al. 2009). Mice with central nervous system (CNS)-specific knockout of PRMT1 exhibit postnatal growth retardation, reduced cell number in white matter tracts, and prominent defects in myelination (Hashimoto, Murata et al. 2016). The B cell-specific PRMT1 knockout mouse model reveals that PRMT1 is required for lymphocyte function and antibody production against the type 2 T cell-independent antigen NP-Ficoll, but not against the T cell-dependent antigen NP-OVA (Hata, Yanase et al. 2016).

PRMT5 is the main type II enzyme in mammals. It is complexed with methylosome protein 50 (MEP50)/WDR77/p44 and the crystal structure of this hetero-octameric complex reveals that the MEP50-PRMT5 interaction plays a key role in PRMT5 substrate recognition and its affinity for substrates (Antonysamy, Bonday et al. 2012; Wang, Fuhrmann et al. 2014). This core structure interacts with different protein complexes to fulfill different functions in transcription, DNA damage, pluripotency and pre-mRNA splicing (Greenblatt, Liu et al. 2016). For instance, it catalyzes the two repressive histone marks H3R8me2s and H4R3me2s, by interacting with COPR5, a component of the SWI/SNF complex (Pal, Vishwanath et al. 2004; Lacroix, El Messaoudi et al. 2008). Additionally, the PRMT5-MEP50 complex associates with either pICln or RioK1 in distinct complexes to methylate specific substrates (Guderian, Peter et al. 2011). pICln facilitates the methylation of Sm proteins by PRMT5, and inhibits the spontaneous assembly of Sm proteins to the U snRNA, thereby playing a role in snRNP assembly (Meister, Eggert et al. 2001; Neuenkirchen, Englbrecht et al. 2015).

Prmt5-null mice exhibit early embryonic lethality between days 3.5 and 6.5, and knockdown of PRMT5 in cells leads to proliferation defects (Wang, Pal et al. 2008). Conditional knockout of PRMT5 in early primordial germ cells (PGCs) results in male and female sterility, along with the up-regulation of LINE1 and IAP transposons and activation of DDR (Kim, Gunesdogan et al. 2014). Loss of PRMT5 in blastocysts reveals that PRMT5 is required for the maintenance of pluripotency (Tee, Pardo et al. 2010). However, in human cells, microarray analysis demonstrated that PRMT5 regulates proliferation in the self-renewing state rather than pluripotency (Gkountela, Li et al. 2014). Conditional knockout of PRMT5 in mouse adult bone marrow exhibits reduction of hematopoiesis progenitor cells

(HPCs), with impaired cytokine signaling and increased p53 expression, suggesting that PRMT5 also plays a role in lymphocyte development (Liu, Cheng et al. 2015).

PRMT7 contains two catalytic domains, and both are required for its activity (Miranda, Miranda et al. 2004). It is a type III enzyme, which preferentially monomethylates arginines within an RxR sequence (Feng, Maity et al. 2013). It was also shown that depletion of PRMT7 leads to enhanced cellular resistance to DNA damaging agents, indicating that PRMT7 is required for the DDR (Karkhanis, Wang et al. 2012). A structural analysis illustrated that Glu181 residue in the Glu-Xaa8-Glu (double E)-loop and Gln329 residue in the Thr-His-Trp (THW) loop of PRMT7 are critical for its enzymatic activity (Debler, Jain et al. 2016; Jain, Warmack et al. 2016).

Patients have been identified with recessive mutations in PRMT7, and they are mainly females who display pseudohypoparathyroidism (Akawi, McRae et al. 2015). In mice, the lack of PRMT7 leads to obesity and muscle regeneration defects. PRMT7 has been shown to epigenetically regulate the expression of the CDK inhibitor p21, leading to premature senescence and exhaustion of the satellite stem cell pool (Blanc, Vogel et al. 2016). *Prmt7*-null mice have decreased energy expenditure with reduced expression of genes involved in muscle oxidative metabolism, including PGC-1α expression (Jeong, Lee et al. 2016). The B cell-specific knockout of PRMT7 mouse model also reveals that PRMT7 is required for germinal center formation by regulating transcription of Bcl6, Prdm1 and Irf4 (Ying, Mei et al. 2015).

1.1.2 Methylarginine readers

The Tudor domain is the only known protein domain to interact with methylated arginines. Among more than 30 Tudor domain-containing proteins (TDRDs) in the mammalian genome (Yang and Bedford 2013), at least ten bind methyllysine residues and eight bind methylarginine residues (Gayatri and Bedford 2014). Various TDRDs are involved in the regulation of splicing (SMN and SPF30), and gene expression (TDRD3 and SND1), as well as the small RNA silencing pathway (TDRD1/6/9) (Gayatri and Bedford 2014). TDRDs are preferentially recruited to these methylated regions near transcription start sites (TSS) to

modulate transcription (Yang and Bedford 2013). Specifically, TDRD3 is recruited to the H4R3me2a and H3R17me2a sites methylated by PRMT1 and PRMT4 respectively, providing the first evidence that TDRDs regulate transcription (Yang, Lu et al. 2010).

1.1.3 Protein arginine demethylation

To date, the existence of arginine demethylases (RDMs) is still controversial, although mounting evidence suggests that arginine methylation is a dynamic modification (Yang and Bedford 2013). The first putative arginine demethylase, Jumonji C (JmjC) domain-containing protein 6 (JMJD6), was subsequently shown to be a lysine hydroxylase (Mantri, Krojer et al. 2010). Recently, a biochemical study was carried out, showing that a subset of JmjC lysine demethylases (KDMs) also function as RDMs on histone and non-histone substrates in purified form (Walport, Hopkinson et al. 2016).

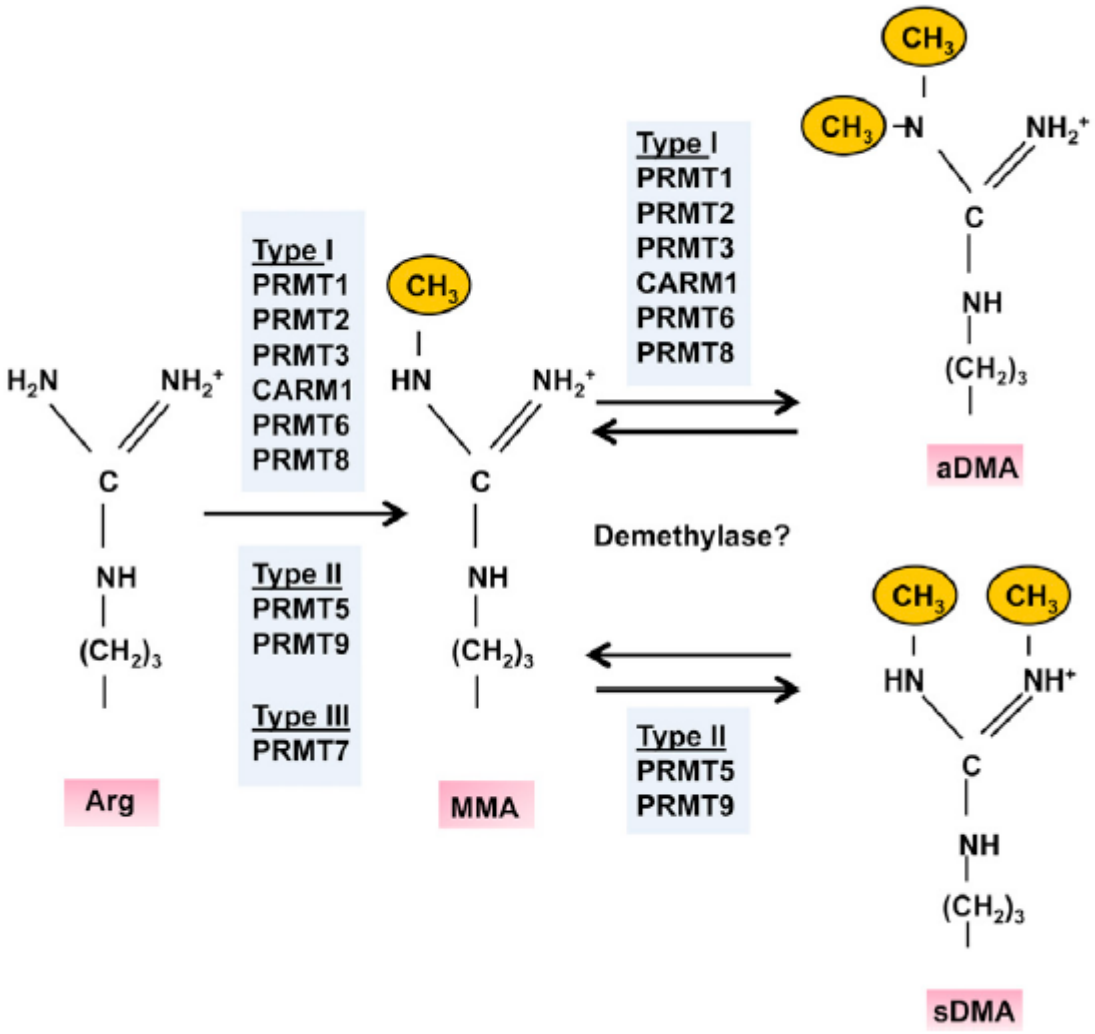


Figure 1.1 Mammalian arginine methylation by different types of PRMTs.
(adapted from (Thandapani, O'Connor et al. 2013)).

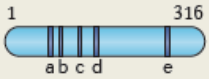
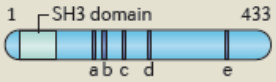
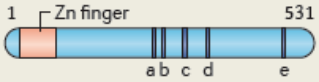
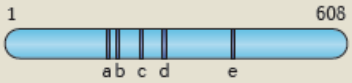
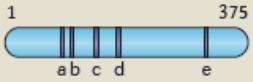
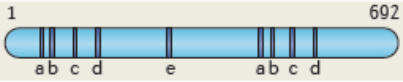
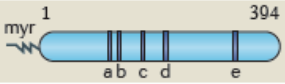
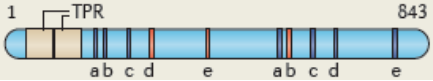
PRMT (locus)	Domain structures of human enzymes*	Function	Family	Primary substrates‡
PRMT1 (19q13.3)		Transcription activation, signal transduction, RNA splicing and DNA repair	Type I	H4R3, MRE11, 53BP1 and SAM68
PRMT2 (21q22.3)		Transcription regulation	Type I	H3R8
PRMT3 (11p15.1)		Ribosomal homeostasis	Type I	RPS2 and p53
CARM1 (19p13.2)		Transcription activation, RNA splicing, cell cycle progression and DNA repair	Type I	H3R17, AIB1, p300, CBP and RNA Pol II CTD
PRMT5 (14q11.2)		Transcription repression, signal transduction and piRNA pathway	Type II	H3R8, H4R3, E2F1, p53, EGFR and CRAF
PRMT6 (1p13.3)		Transcription regulation	Type I	H3R2 and H2AR29
PRMT7 (16q22.1)		Male germline gene imprinting	Type II and type III	H4R3, H2AR3 and H3R2
PRMT8 (12p13.3)		Brain-specific function	Type I	Unknown
PRMT9 (4q31.23)		Unknown	Not classified	Unknown

Figure 1.2 Domain structures of mammalian PRMTs and their major functions. (adapted from(Yang and Bedford 2013)).

1.1.4 The cellular processes regulated by arginine methylation

As a fundamental post-translational modification, arginine methylation is involved in a variety of cellular functions, such as RNA processing, signal transduction, transcription, and DNA repair (Bedford and Richard 2005).

1.1.4.1 Transcriptional coactivators/corepressors

PRMTs generally modulate transcription, as they are recruited to the promoter regions where they methylate histones, as well as transcription factors, co-regulators, and RNA polymerase II (RNA polII), acting as transcriptional coactivators or corepressors.

Transcriptional coactivator

PRMT1 and PRMT4 are known to function in transcription as coactivators. PRMT1 mainly deposits the H4R3me2a mark associated with active transcription, recruiting other transcriptional coactivators including p300/CBP and potentiating histone acetyltransferases for binding of transcription factors (Bedford and Clarke 2009). This was also confirmed in mice, where H4R3me2a and histone acetylation were shown to control the transcription of Cdk5 and CaMKII in the nucleus accumbens (NAc) of mouse brains after cocaine treatment (Li, Zhu et al. 2015). PRMT4 is another transcriptional coactivator that specifically methylates the transcription active sites H3R17, H3R26 and H3R42 (Yang and Bedford 2013). It has been shown that the C-terminal domain (CTD) of RNA polII is methylated by PRMT4, providing the docking site for TDRD3. When this methylation is prevented, small nuclear and nucleolar RNAs are deregulated (Sims, Rojas et al. 2011). Histone H3R17me2 mark generated by PRMT4 recruits the PAF1 complex (Paf1C), which plays critical roles in RNA polymerase II transcription elongation (Wu and Xu 2012). Additionally, mounting evidence demonstrates that PRMT1 and PRMT4 recruit various transcription factors such as p53, YY1, NF-κB, PPARγ, RUNX1, and E2F1 to specific promoters where they activate transcription (Yadav, Cheng et al. 2008; Zhao, Jankovic et al. 2008; Yost, Korboukh et al. 2011). Using mass spectrometry and a methyl-arginine-specific antibody, Zhao et al. (2008) discovered two PRMT1-methylated arginines in RUNX1, which interact with the transcriptional corepressor SIN3A. PRMT1-dependent methylation abrogates the association between RUNX1 and

SIN3A, leading to transcription activation (Zhao, Jankovic et al. 2008).

Transcriptional corepressor

PRMT5, PRMT6 and PRMT7 mainly function as transcriptional corepressors. PRMT5 symmetrically dimethylates H4R3 in transcriptionally repressed loci, and the H4R3me2s mark recruits DNMT3a (DNA methyltransferase 3A), which further enhances gene repression by methylating adjacent DNA. This implies that histone methylation is associated with epigenetic silencing (Zhao, Rank et al. 2009). Furthermore, PRMT5 acts as a general transcriptional repressor by functioning with repressor complexes and transcription factors, including BRG1, hBRM, Blimp1, and Snail (Bedford and Clarke 2009). For instance, PRMT5 is a component of the Snail-silencing complex, and as such, interacts with the transcriptional corepressor Ajuba, to repress E-cadherin transcription (Chiang and Ayyanathan 2013). PRMT6 has been shown to methylate histones at H3R2 and H2AR29 with transcriptional repression activity (Waldmann, Izzo et al. 2011). PRMT6 dimethylates H3R2 in mammalian cells and blocks the recruitment of transcriptional coactivator WDR5 to the methylated H3K4 site, thereby acting as a transcriptional repressor (Guccione, Bassi et al. 2007; Neault, Mallette et al. 2012). The *in vivo* substrates of PRMT7 remain unknown. However, PRMT7 has been shown to deposit H4R3me2s, and to play a role in DNA methylation of the imprinting control region (ICR). Mechanistically, CTCFL, which binds ICR via its zinc finger region, interacts with the PRMT7 C-terminus. The CTCFL-PRMT7 complex methylates H4R3 and recruits Dnmt3a and Dnmt3b for DNA methylation (Jelinic, Stehle et al. 2006).

1.1.4.2 DNA repair

Genotoxic agents damage DNA and thus lead to serious lesions in the genome. To counteract DNA damage, mammalian cells utilize versatile signaling machinery known as DNA damage response (DDR). Various proteins known as PIKKs (phosphatidylinositol-3-kinase like protein kinase family members), such as ATM (ataxia telangiectasia mutated), ATR (ataxia telangiectasia and Rad3-related), and DNA-PK (DNA-dependent protein kinase), are involved in DDR (Ciccia and Elledge 2010; Polo and Jackson 2011). The major DNA repair pathways are homologous recombination repair (HR),

non-homologous end joining (NHEJ), mismatch repair (MMR), nucleotide excision repair (NER) and base excision repair (BER) (Natarajan and Palitti 2008; Curtin 2012). In HR, the mammalian MRN complex containing MRE11, Rad50 and NBS1 plays a critical role. MRE11 is methylated by PRMT1 in its glycine-arginine-rich (GAR) motif both *in vivo* and *in vitro*, and this methylation recruits MRE11 translocation to the double-strand break (DSB) foci (Boisvert, Dery et al. 2005; Boisvert, Hendzel et al. 2005). *Mre11* (RK/RK) mice display cell cycle checkpoint defects and chromosome instability with impairment in ATR/CHK1 signaling (Yu, Vogel et al. 2012). 53BP1 harboring a Tudor domain is associated with chromatin via H4K20me2, playing a role in NHEJ (Botuyan, Lee et al. 2006). 53BP1 itself can also be methylated by PRMT1, and methylase inhibitors block 53BP1 recruitment at DSBs (Boisvert, Rhie et al. 2005; Suchankova, Legartova et al. 2014). As a player in BER, DNA polymerase β (Polβ) is methylated by PRMT1 and PRMT6 with non-redundant roles. Polβ contains an N-terminal lyase domain and a C-terminal domain with polymerase nucleotidyl-transferase activity. Methylation at R83 and R152 by PRMT6 increases its polymerase activity, while methylation at R137 by PRMT1 blocks its interaction with PCNA (Proliferating cell nuclear antigen) for downstream functions (El-Andaloussi, Valovka et al. 2006; El-Andaloussi, Valovka et al. 2007; Morettin, Baldwin et al. 2015).

1.1.4.3 Cytoplasmic localization

Biochemical and proteomic analyses have demonstrated that Sam68 and FGF2 mainly localize in the nucleus, and that this localization is mediated by arginine methylation (Bedford and Richard 2005; Sylvestersen, Horn et al. 2014). It was shown that hypomethylated Sam68 localizes in the cytoplasm, indicating that arginine methylation is crucial for its nuclear function (Cote, Boisvert et al. 2003). Additionally, in response to cellular stresses such as heat shock, oxidative stress, energy deprivation and glucose starvation, cells form large, complex ribonucleoprotein particles called stress granules (SGs) (Anderson and Kedersha 2008). Various mRNAs, most of which encode oncogenes, caspases and cystatins, are sequestered and prevented from ongoing translation (Lopez de Silanes, Galban et al. 2005). Many of the proteins associated with stress granules are methylated by PRMTs. It has been recently demonstrated that PRMT1 and PRMT5 methylate distinct

arginine residues of the SG nucleating protein Ras-GAP (GTPase activating protein) SH3 binding protein 1 (G3BP1), which impedes SG assembly (Tsai, Gayatri et al. 2016). Some methylarginine effector molecules are also found in stress granules to promote SG assembly, including SMN (survival motor neuron protein) and TDRD3 (Pek, Anand et al. 2012).

1.1.4.4 RNA metabolism

The majority of PRMT substrates are associated with RNAs, since the majority of RNA-binding proteins, including hnRNPs (A1, A2, K, R and U), harbor RGG/RG motifs. Therefore, arginine methylation has been shown to be involved in RNA metabolism, including in mRNA splicing, transport, translation and turnover (Pahlich, Bschir et al. 2005).

PRMT5 is known to be a putative splicing regulator by virtue of its function in assembly of snRNPs (small nuclear ribonucleoproteins). Spinal muscular atrophy (SMA) is caused by loss-of-function mutation of *SMN* gene, making it an important cause of infant death (Lefebvre, Burglen et al. 1995). One well-studied mutation in the Tudor domain, E134K, impedes SMN association with methylated arginines in RGG/RG motif-containing proteins (Cote and Richard 2005; Workman, Kolb et al. 2012). During spliceosome assembly, PRMT5 methylates arginines in the RG-rich clusters of Sm family proteins, which are recognized and bound by the SMN Tudor domain (Brahms, Raymackers et al. 2000). Disruption of arginine methylation by PRMT5 blocks the SMN-pICln-Sm interaction, and impedes the Sm proteins from forming snRNPs (Cote and Richard 2005; Neuenkirchen, Englbrecht et al. 2015). As a result, the reduced arginine methylation of Sm proteins in PRMT5-depleted cells leads to aberrant constitutive splicing and the alternative splicing of mRNAs with weak 5’ donor sites (Bezzi, Teo et al. 2013; Koh, Bezzi et al. 2015). Moreover, arginine methylation of RNApolII by PRMT5 recruits SMN and senataxin to resolve R-loops (RNA-DNA hybrids), playing a role in the transcription termination pathway (Zhao, Gish et al. 2016). Other PRMTs are also implicated in splicing regulation. PRMT4 has been demonstrated to be involved in splicing by promoting exon skipping. It asymmetrically dimethylates a variety of splicing factors, including the transcription factor CA150, spliceosome-associated protein SAP49, SmB and U1C (Cheng, Cote et al. 2007; Greenblatt, Liu et al. 2016). PRMT1 regulates alternative splicing of genes important for megakaryopoiesis by modulating the expression of the

splicing regulator RBM15. PRMT1 methylates RBM15 at R578, leading to its ubiquitination and degradation (Zhang, Tran et al. 2015). Arginine methylation of the spliceosome-associated protein SAP145 by PRMT9 generates a docking site for SMN binding, implying that PRMT9 plays a key role in U2 snRNP maturation (Yang, Hadjikyriacou et al. 2015).

Arginine methylation also regulates mRNA translation. SNBP/ZNF9 (nucleic acid binding protein) is a key player in promoting cap-dependent and cap-independent translation, and arginine methylation by PRMT1 on R25 and R27 impedes its RNA binding ability (Wei, Hu et al. 2014). Moreover, arginine methylation of the translation repressor Scd6 promotes its binding to the eukaryotic translation initiation factor eIF4G1, thereby further repressing translation initiation (Poornima, Shah et al. 2016).

1.1.4.5 Signal transduction

Although arginine methylation appears to be a stable mark, mounting evidence suggests that PRMTs (mainly PRMT1 and PRMT8) play a role in various signaling pathways. PRMT8, which contains a myristoylation site, associates with the plasma membrane and contributes to signal transduction. PRMT8 also has a proline-rich domain that interacts with SH3 domains, implying its role in signal transduction (Lee, Sayegh et al. 2005; Sayegh, Webb et al. 2007). *Prmt8*-knockout mice reveal that PRMT8 functions as a phosphatidylcholine-hydrolyzing phospholipase D, which is required for normal neurological function (Kim, Park et al. 2015).

Furthermore, PRMT1 binds the cytoplasmic region of the type I interferon receptor, revealing its role in signal transduction in immune responses (Altschuler, Wook et al. 1999; Bedford and Richard 2005; Ghildiyal and Sen 2015). PRMT1 has also been shown to participate in insulin/EGFR-triggered signaling, estrogen signaling and glucose metabolism (Iwasaki and Yada 2007; Le Romancer, Treilleux et al. 2008; Lv, Chen et al. 2015). Specifically, PRMT1 methylates R198 and R200 of the extracellular domain of EGFR to facilitate its binding to EGF, resulting in strengthened dimerization and enhanced signal transduction (Liao, Hsu et al. 2015). It is worth mentioning that arginine methylation interferes with phosphorylation when the methylation site is in proximity to the phosphorylation site (Yamagata, Daitoku et al. 2008). For instance, methylation by nuclear

PRMT1 superimposes the phosphorylation site of FOXO1, implying the role of PRMT1 in FOXO1-induced apoptosis (Altan, Yokobori et al. 2016).

1.1.5 PRMTs and Cancer

Numerous reports have implied that PRMTs are misregulated in cancer, and thus inhibitors of PRMTs make them promising therapeutic targets (Yang and Bedford 2013; Greenblatt, Liu et al. 2016).

PRMT5 is commonly overexpressed in cancers, including in solid tumors, lymphomas and leukemias. The 5-methylthioadenosine phosphorylase (MTAP) gene is frequently deleted in human cancers. Recently, biochemical studies and large-scale screens revealed that the viability of MTAP-depleted cells is reduced upon PRMT5 depletion (Mavrakis, McDonald et al. 2016). MTA, a substrate of MTAP, is a potent and specific inhibitor of PRMT5, and reduces PRMT5 methylation activity in MTAP-depleted cells, creating a set of targets in cancer (Kryukov, Wilson et al. 2016; Marjon, Cameron et al. 2016). It has been reported that the MYC oncogene directly upregulates transcription of PRMT5 and other core snRNP particle assembly genes during lymphomagenesis, thereby promoting effective pre-mRNA splicing, cell survival and proliferation (Koh, Bezzi et al. 2015). PRMT5 and KLF4 (Kruppel-like factor 4) are overexpressed in breast tumors, where PRMT5 methylates KLF4 and impedes its ubiquitination. This leads to KLF4 accumulation and tumor initiation and progression (Hu, Gur et al. 2015). Furthermore, PRMT5 is crucial for Epstein-Barr virus (EBV)-induced B cell transformation, and it regulates the epigenetic-repressive marks (Alinari, Mahasenan et al. 2015). PRMT5 is also a marker of malignancy in glioma cells and is highly expressed in glioblastoma multiforme (GBM). Depletion of PRMT5 impairs the ERK signaling pathway, resulting in impeded tumor growth (Han, Li et al. 2014). Another study using liquid chromatography - tandem mass spectrometry (LC-MS/MS) analysis showed that PRMT5 methylates and positively regulates MYCN expression levels, thereby playing a major role in neuroblastomas tumorigenesis (Park, Szemes et al. 2015). In human melanoma cells, loss of PRMT5 leads to decreased expression of the melanocyte-lineage specific oncogene MITF (microphthalmia-associated transcription factor), and increased

expression of cell cycle regulator p27, indicating that PRMT5 regulates growth of melanoma cells (Nicholas, Yang et al. 2013).

PRMT1 is another fundamental tumorigenesis regulator implicated in leukemia, glioblastoma, and lung, liver and breast cancers. In acute leukemia cells, PRMT1 and KDM4C, an H3K9 demethylase, act synergistically to mediate the epigenetic reprogramming of MLL fusions, contributing to their transforming ability (Cheung, Fung et al. 2016). Additionally, AE9a is a splicing isoform of AML1-ETO, which can cause leukemia in mice. Shia et al. (2012) identified that PRMT1 interacts with and weakly methylates AE9a at the promoter regions of AE9a-activated genes, which leads to leukemia progression (Shia, Okumura et al. 2012). It also has been observed that PRMT1 methylates the pro-oncoprotein SRSF1, and together they contribute to leukemogenesis in pediatric acute lymphoblastic leukemia patients (Zou, Zhang et al. 2012). PRMT1 is also a potential biomarker of gliomas, inasmuch as the level of PRMT1 histone substrate, H4R3, is higher in glioma cells as compared to normal brain tissue (Wang, Tan et al. 2012). PRMT1 mediates tumor metastasis by promoting epithelial-mesenchymal transition (EMT) in lung cancer. Avasarala et al. (2015) identified that Twist1, a transcription factor that represses E-cadherin, is methylated by PRMT1 and this methylation mark is important for active E-cadherin repression (Avasarala, Van Scoyk et al. 2015). Accordingly, PRMT1 also promotes EMT in breast cancer cells and in mouse models, via the transcriptional regulation of ZEB1 (Gao, Zhao et al. 2016). Moreover, PRMT1 is crucial for hepatocellular carcinoma (HCC) progression, since miR-503 targeting the 3'-UTR of PRMT1 remarkably impairs the invasion and migration of HCC cells (Li, Liu et al. 2015).

PRMT4 is also proven to be an oncogene mainly in lung and breast cancer. PRMT1 and PRMT4 were reported to be drastically overexpressed in 60 NSCLC (non-small cell lung carcinoma) patients, and PRMT4 level was correlated with the degree of tumor differentiation (Elakoum, Gauchotte et al. 2014). Estrogen receptor α (ERα) expression in breast cancer generally promotes a more differentiated phenotype than that observed in ERα-negative tumors. PRMT4 is a potential biomarker of breast cancer, especially in the ER-positive luminal-like subgroup, and its expression is correlated with poor prognosis

(Habashy, Rakha et al. 2013). Mechanistically, PRMT4 acts as a coactivator of ERα which simultaneously blocks cell proliferation and induces differentiation (Al-Dhaheri, Wu et al. 2011). By using the Zinc-Finger Nuclease technology to knock out PRMT4 in several breast cancer cell lines, Wang et al. (2014) identified the SWI/SNF core subunit BAF155 as a PRMT4 arginine methylation substrate, which plays a role in breast cancer cell migration and metastasis (Wang, Zhao et al. 2014). Intriguingly, MED12 (RNA polymerase II mediator complex subunit 12) has been demonstrated as another methylation substrate of PRMT4, and cells defective in MED12 are associated with poor prognosis after chemotherapy in breast cancer patients (Wang, Zeng et al. 2015).

1.2 Substrates of arginine methylation

1.2.1 RGG/RG motif

The first evidence of arginine methylation on RGG motifs is from 1985, when Lischwe and colleagues showed that dimethylated arginines are present in the C-terminus of nucleolin, implying that RGG sequences are substrates of PRMTs (Lischwe, Cook et al. 1985). Subsequently, it was demonstrated that RGG sequences of hnRNP U bind homopolymeric RNA. Tri-RGG sequences were shown to be conserved amongst different species (Kiledjian and Dreyfuss 1992). Aromatic residues were found in between the RGGs, contributing to the hydrophobic stacking of RNA molecules. In addition, a bioinformatic analysis confirmed that RGG sequences are evolutionally conserved, with ~70 occurrences within ~15 genes in the human genome (Corley and Gready 2008). This was recently redefined, when Larsen et al. (2016) identified 3300 human proteins harboring 8030 arginine methylation sites in HEK293 cells, by performing high-resolution mass spectrometry (Larsen, Sylvestersen et al. 2016). Moreover, Richard group has evaluated the functional enrichments in proteins examined by progressively relaxed motif definition (Thandapani, O'Connor et al. 2013). In this study, multiple groups of proteins containing tri-RGG, di-RGG, tri-RG and di-RG motifs were identified (Thandapani, O'Connor et al. 2013). There are 31 protein isoforms harboring tri-RGG motifs (RGG(X_{0-4})RGG(X_{0-4})RGG), including the RNA-binding proteins hnRNP A1,

hnRNPU, nucleolin, and fibrillarin, the transcription factors FET (FUS/EWS/TAF15), the translational repressor LSM14A and the chromatin binding protein SERBP1, which is localized in the stress granule and methylated by PRMT1 (Lee, Hsieh et al. 2012; Lee, Wei et al. 2014). There are 88 di-RGG motif-containing proteins harboring RGG(X_{0-4})RGG sequences. Some of the well-known members are the RNA-binding proteins Sam68, hnRNPK, Aven, as well as FMRP, which is methylated by various PRMTs (PRMT1, PRMT3 and PRMT4) (Wei, Mundade et al. 2014). Moreover, it was identified that a large number of proteins contain RG motifs rather than RGG motifs. There are 314 proteins including MRE11, DHX9, 53BP1, SHANK1 and TDRD3, which contain RG(X_{0-4})RG(X_{0-4})RG sequences, and they were shown to be involved in RNA biogenesis, DNA damage signaling and mRNA translation. DHX9 is a known PRMT1 substrate, and also can be found in the interactome of PRMT1 spliced isoforms, as was determined by a SILAC-based quantitative affinity purification mass spectrometry analysis (Baldwin, Bejide et al. 2015). Finally, there are more than 1,700 proteins with a di-RG motif (RG(X_{0-4})RG). Some notable proteins are the splicing factor SRSF1, the ubiquitin ligase RBBP6, and the transcription factor E2F-1, which is methylated in a competitive fashion by PRMT1 and PRMT5 (Zheng, Moehlenbrink et al. 2013).

Several RGG/RG motif containing proteins interact with G-quadruplex RNA, and this association has been shown to increase G-quartets RNA stability (Taha, Nouri et al. 2014; Anderson, Chopra et al. 2016). Indeed, GO (gene ontology) analysis revealed that most of the RGG/RG motif-containing proteins are RNA-binding proteins, although some of them have no other RNA binding domains (Thandapani, O'Connor et al. 2013). Accordingly, Larsen et al. (2016) identified that somatic mutations are prevalent at the methylated arginines of RNA-binding proteins and that arginine methylation differentially regulates the RNA-binding ability of these RGG/RG-containing proteins, such as SRSF2 and hnRNPUL1 (Larsen, Sylvestersen et al. 2016). RGG/RG motifs also mediate protein-protein interactions, as methylated RGG/RG motifs provide docking sites for Tudor domain-containing proteins (Thandapani, O'Connor et al. 2013). Herein, I discuss the biochemical properties and biological functions of two RGG/RG motif-containing proteins: Sam68 (section 1.2.2) and

Aven (section 2.7.4).

1.2.2 An RGG/RG motif-containing protein: Sam68

1.2.2.1 Structure of Sam68

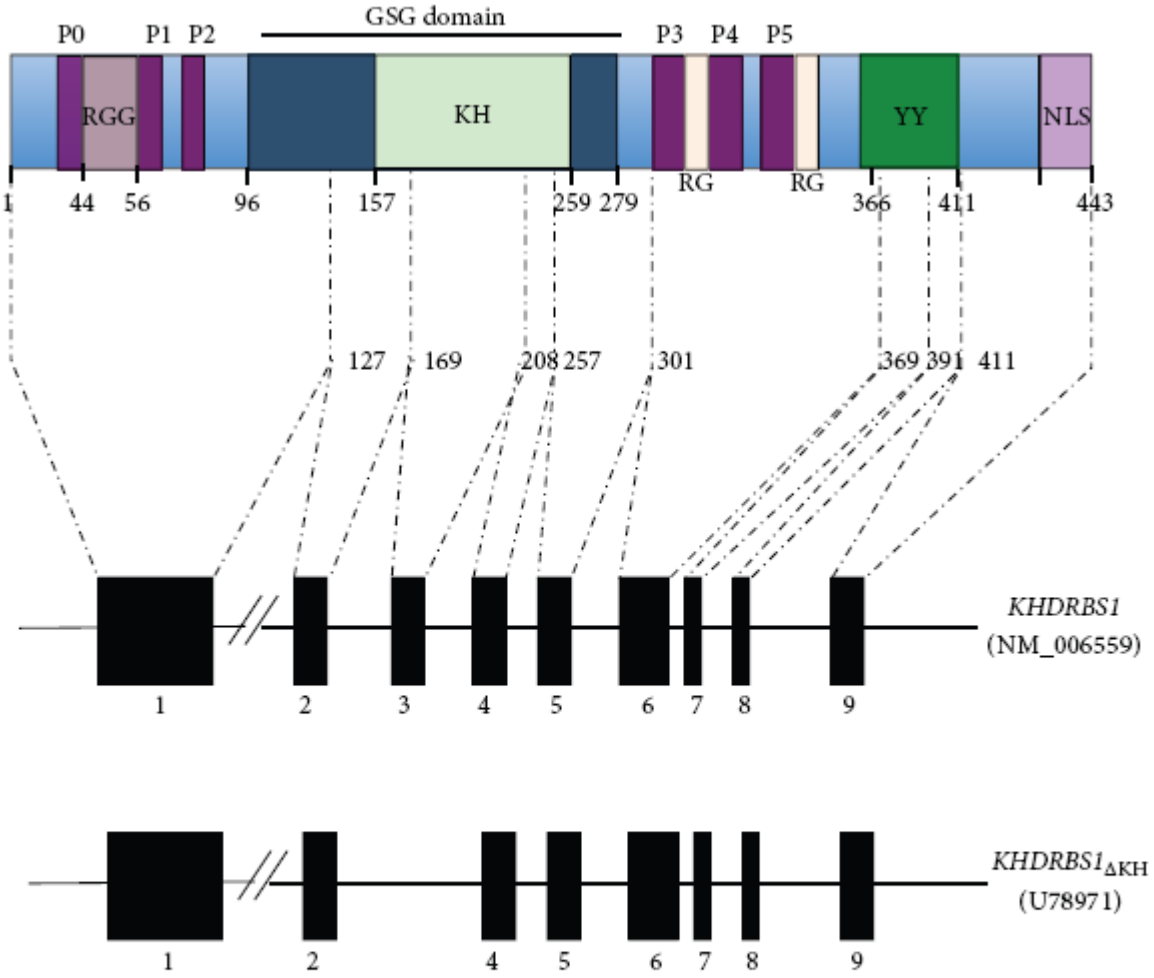
A well-known RGG/RG motif-containing protein, Sam68, is the Src-associated substrate during mitosis of 68 kDa belonging to the KH (heteronuclear ribonucleoprotein particle K homology) domain family of RNA-binding proteins. Sam68 was initially discovered as a cell cycle-regulated phosphorylation substrate of Src. Sam68 is transcribed from the *KHDRBS1* gene and contains several functional domains including the previously termed GSG (GRP33, Sam68, GLD-1) domain, which harbors a KH domain, N- and C-terminal regions of the KH, a YY (Tyrosine-rich) domain, an RGG/RG motif and a nuclear localization signal (NLS) (Figure 1.3) (Frisone, Pradella et al. 2015). Dimer formation allows the KH domain to bind bipartite A/U-rich RNA sequences (particularly UAAA or UUUA) with high affinity, as shown by SELEX (systematic evolution of ligands by exponential enrichment) (Lin, Taylor et al. 1997; Galarneau and Richard 2009) and CLIP (crosslinking and immunoprecipitation)-seq (our unpublished data). Recently, the atomic resolution structure of Sam68 bound to RNA was established, revealing that this unique dimerization is crucial for Sam68 biological functions (Feracci, Foot et al. 2016). Additionally, Sam68 contains several proline-rich domains that interact with SH3 domain- and WW domain-containing proteins (Lukong and Richard 2003). Recently, it was reported that SH3 inhibitors disrupt the interaction between Sam68 and hOSF (human osteoclast-stimulating factor), which is an intracellular protein produced by osteoclasts. The same group also demonstrated that two key proline residues (Pro427 and Pro430) are required for this binding (Han, Liu et al. 2016). The C-terminal tyrosine-rich domain, which is phosphorylated by tyrosine kinases, exerts a negative impact on Sam68 RNA-binding activity, possibly via the dissociation of homodimers (Meyer, Tripsianes et al. 2010). It has been demonstrated that the RNA-binding activity of Sam68 is modulated by its tyrosine phosphorylation levels, and by its association with SH3-containing proteins (Lukong and Richard 2003; Feracci, Foot et al. 2014) (Figure 1.3). Thus, Sam68 and other KH domain-containing proteins are referred to as signal transduction and activation of

RNA (STAR) family proteins (Vernet and Artzt 1997). These other proteins include GRP33, GLD-1, QKI (Quaking), SLM-1, SLM-2, HOW, KEP1, Sam50 and SF1 (Sanchez-Jimenez and Sanchez-Margalet 2013). Sam68 can be acetylated by histone acetyltransferases at lysine residues in its GSG domain, positively regulating its RNA-binding ability (Babic, Jakymiw et al. 2004). It has been shown that SMAR1 (Scaffold/matrix-associated region-binding protein 1) in complex with HDAC6 (histone deacetylase 6) maintains Sam68 in a deacetylated state, conferring a tumor suppressive effect in breast tumor cells (Nakka, Chaudhary et al. 2015). Moreover, Sam68 can be methylated by PRMT1 at its RGG/RG motif *in vivo* and *in vitro*, and this is required for Sam68-mediated RNA export. The arginines within the Sam68 RGG/RG motif undergo monomethylation and asymmetrical dimethylation, as shown by biochemical studies and hmSILAC (heavy methyl stable isotope labeling by amino acids in cell culture) with high-resolution mass spectrometry-based proteomics (Cote, Boisvert et al. 2003; Bremang, Cuomo et al. 2013). Lastly, SUMOylation is another post-translational modification occurring on Sam68, which results in repressed cyclin D1 expression, G_1 arrest and cell proliferation defects (Babic, Cherry et al. 2006; Filosa, Barabino et al. 2013).

1.2.2.2 Function of Sam68

Role of Sam68 in signal transduction

Although Sam68 was identified as a Src substrate during mitosis (Fumagalli, Totty et al. 1994) , its role during mitosis remains undefined. Sam68 is the target of various tyrosine kinases, including Erk1/2 or Erk5, Fyn, Lck, Tec, Jak3, Btk and the insulin receptor (Sanchez-Jimenez and Sanchez-Margalet 2013). Tyrosine-phosphorylated Sam68 associates with SH2 domain-containing proteins, including Src family kinases, Brk (Btk6), PLCγ-1, Ras-GAP and Tec family kinases. The proline-rich domain of Sam68 provides the binding site of SH3 domain-containing proteins, such as p85 PI3K (phosphatidylinositol 3-kinase), PLCγ-1, Grb-2, GRAP (Grb2-related adaptor protein) and Nck (Sanchez-Jimenez and Sanchez-Margalet 2013). Thus, Sam68 acts as a scaffold protein to mediate the signal transduction in response to the various stimuli from the outer membrane. It has been demonstrated that Sam68 is involved in the TNFα pathway, the insulin receptor pathway,



GSG: GRP33/SAM68/GLD-1 domain
KH: heterogeneous nuclear ribonucleoprotein particle K (hnRNP K) homology domain
P0–P5: proline-rich motifs
RGG and RG: arginine/glycine/glycine and arginine/glycine boxes
YY: tyrosine-rich domain
NLS: nuclear localization signal

Figure 1.3 Genomic structures of Sam68 and its variant.
Sam68 contains an RGG domain, a GSG (GRP33/Sam68/GLD-1) domain of 200 amino acids, and regulatory domains containing motifs for protein-protein interactions. The KH domain is 70-100 amino acids long, containing the conserved N- and C- terminal flanking sequences. An NLS (nuclear localization signal) indicates that it predominantly localizes in the nucleus (Frisone, Pradella et al. 2015).

the EGF-mediated pathway, Met signaling and NF-κB signaling, which all lead to different biological functions of Sam68 (Sanchez-Margalet, Gonzalez-Yanes et al. 2003; Lukong, Larocque et al. 2005; Kunkel and Wang 2011; Locatelli, Lofgren et al. 2012; Fu, Sun et al. 2016). Interestingly, Sam68 also acts as a convergent hub of the PI3K and p38MAPK pathways (Venigalla and Turner 2012). Recently, it has been demonstrated that Sam68 plays a significant role in TLR2 (Toll-like receptor) and TLR3 signaling, as well as in the downstream activation of NF-κB p65 (Tomalka, de Jesus et al. 2016).

Regarding the biological functions of Sam68, it mostly acts as a proto-oncogene in an RNA-dependent or -independent manner. Tyrosine phosphorylation of Sam68 at its C-terminus facilitates Sam68 nuclear localization, and thus promotes cell cycle progression (Lukong, Larocque et al. 2005). Elevated Sam68 expression levels have been observed in various human cancers including breast cancer, prostate cancer, ovarian cancer, colorectal cancer, lung cancer and lymphoma (Frisone, Pradella et al. 2015). Increased levels of tyrosine-phosphorylated Sam68 are frequently observed in **prostate** cancer cells, with increased proliferation and survival ability against chemotherapeutic agents (Busa, Paronetto et al. 2007). Sam68 expression and nuclear localization are related to poor survival of patients with **colorectal** cancer (Liao, Liu et al. 2013). Recently, upregulated Sam68 was observed in human and mouse colon cancer, with elevated PAR (polymers of ADP-ribose) production and NF-κB-mediated anti-apoptotic transcription (Fu, Sun et al. 2016). Similarly, high expression levels of Sam68 correlate with lymph node metastasis and poor prognosis of **NSCLC** (Non-small cell lung cancer) patients (Zhang, Xu et al. 2014). Sam68 is responsible for cell cycle progression and cell proliferation, as well as cell adhesion-mediated drug resistance via the Akt pathway in non-Hodgkin **lymphoma** (NHL) (Wu, Xu et al. 2015). Overexpression of Sam68 is related to increased proliferation, cell cycle progression and lymph node metastasis of epithelial **ovarian** cancer (EOC) tumor tissues (Dong, Che et al. 2016). It has also been shown recently that depletion of Sam68 results in MAPK/PI3K pathway defect and impaired proliferation, upon insulin and leptin treatment in **breast** adenocarcinoma cell lines (Perez-Perez, Sanchez-Jimenez et al. 2016). Moreover, Sam68 is also involved in T-cell acute lymphoblastic **leukemia** (T-ALL) by participating in the Akt/mTOR signaling pathway in

T-ALL cell lines (Wang, Li et al. 2016).

Although they have a normal lifespan and development (Richard, Torabi et al. 2005), Sam68$^{-/-}$ mice show impaired fertility, delayed onset of mammary tumorigenesis and metastasis, basal motor coordination failure, fewer excitatory synapses in the hippocampus, maintenance of bone mass with aging, lean phenotype with protection from obesity and increased white adipose tissue browning, as well as hypersensitivity to genomotoxity caused by DNA-damaging agents (Lukong and Richard 2008; Richard, Vogel et al. 2008; Bianchi, Barbagallo et al. 2010; Huot, Vogel et al. 2012; Klein, Younts et al. 2013; Zhou, Cheng et al. 2015; Fu, Sun et al. 2016).

Role of Sam68 in transcription

Sam68 links signaling pathways with gene transcriptional regulation by directly associating with transcription factors. Independent of its RNA-binding ability, Sam68 interacts with the transcription cofactor CBP and represses CBP-dependent gene transcription by competing with other transcription co-regulators of CBP (Hong, Resnick et al. 2002). Notably, Sam68 interacts with the androgen receptor (AR) and enhances its transcriptional activity in prostate cancer cells (Rajan, Gaughan et al. 2008). It has also been proposed that Sam68 interacts with the splicing factor, SRSF1 (ASF/SF2), and this association might enhance the transcriptional repression of cyclin D1 (Bielli, Busa et al. 2011). Sam68 regulates cyclin D1 expression, playing a role in proliferation of astrocytes in the central nervous system (Chen, Liu et al. 2016). Fu et al. (2013) reported that Sam68 is recruited to the promoter region of CD25 in normal and transformed human T cells, implying that Sam68 is a transcriptional regulator of NF-κB (Fu, Sun et al. 2013). Additionally, Sam68 acts as a transcriptional coactivator of the p53 tumor suppressor in response to DNA damage, since the depletion of Sam68 results in the defects of cell cycle arrest and apoptosis mediated by p53 (Li and Richard 2016) .

Role of Sam68 in RNP complex formation

In the nucleus, Sam68 was shown to participate in pre-mRNA processing by interacting with various RNPs, such as hnRNPK, hnRNPA1, hnRNPG, and FAST (Paronetto, Achsel et al. 2007; Simarro, Mauger et al. 2007; Ulke-Lemee, Trinkle-Mulcahy et al. 2007).

Neurological defects have been associated with alterations in these hnRNPs, although the possible role of Sam68 in these neurological defects remains to be elucidated (Sanchez-Jimenez and Sanchez-Margalet 2013). A recent study demonstrated that Sam68 RNPs containing G3BP1, FMRP and Sam68 are observed in spreading initiation centers (SIC) of newly adhering cells, and this localization is mediated by its association with mRNAs (Bergeman, Caillier et al. 2016). In cancer cells, Sam68 is associated with subnuclear organelles called nuclear bodies (SNBs). SNBs contain splicing factors, signaling components, and nucleic acids, which are modulated in response to transcription inhibitors, growth signals, and mitosis (Chen, Boisvert et al. 1999; Huot, Vogel et al. 2009). Various proteins including PTK6, SLM-1/2, hnRNPA2/B1, hnRNP G, YT521, DBC1 and hnRNP L are associated with Sam68 in these SNBs (Hartmann, Nayler et al. 1999; Derry, Richard et al. 2000; Rajan, Dalgliesh et al. 2009; Close, East et al. 2012). Interestingly, Mannen et al. (2016) discovered that Sam68 SNBs are composed of two distinct RNase-sensitive substructures, and that hnRNPL acts as the adaptor bridging these two substructures (Mannen, Yamashita et al. 2016). Moreover, these Sam68-containing-SNBs have been observed to regulate alternative splicing in response to genotoxic stress, in order to overcome the induced damage (Busa, Geremia et al. 2010). In the cytoplasm, Sam68 can also be found in different cellular structures, mainly in stress granules (SGs). These SGs contain translation initiation components, including phosphorylated translation initiation factor (eIF2α), T-cell intracellular antigen (TIA-1), and Ras-GAP binding protein (G3BP1) and melanoma differentiation-associated gene 5 (MDA5) (Anderson and Kedersha 2009; Finnen, Pangka et al. 2012; Langereis, Feng et al. 2013). It has been shown recently that the KH domain of Sam68 is essential for its localization to SGs, and Sam68 is not a constitutive component of SG formation during EV71 (enterovirus 71) infection (Zhang, Chen et al. 2016).

Role of Sam68 in alternative splicing

Sam68 was originally shown to bind to the splicing factor YT521-B to regulate the splicing site selection, providing the first evidence that signaling mediators regulate pre-mRNA splicing (Hartmann, Nayler et al. 1999). Subsequently, increasing evidence suggests that Sam68 binds to intronic elements of pre-mRNAs, spliceosome-associated

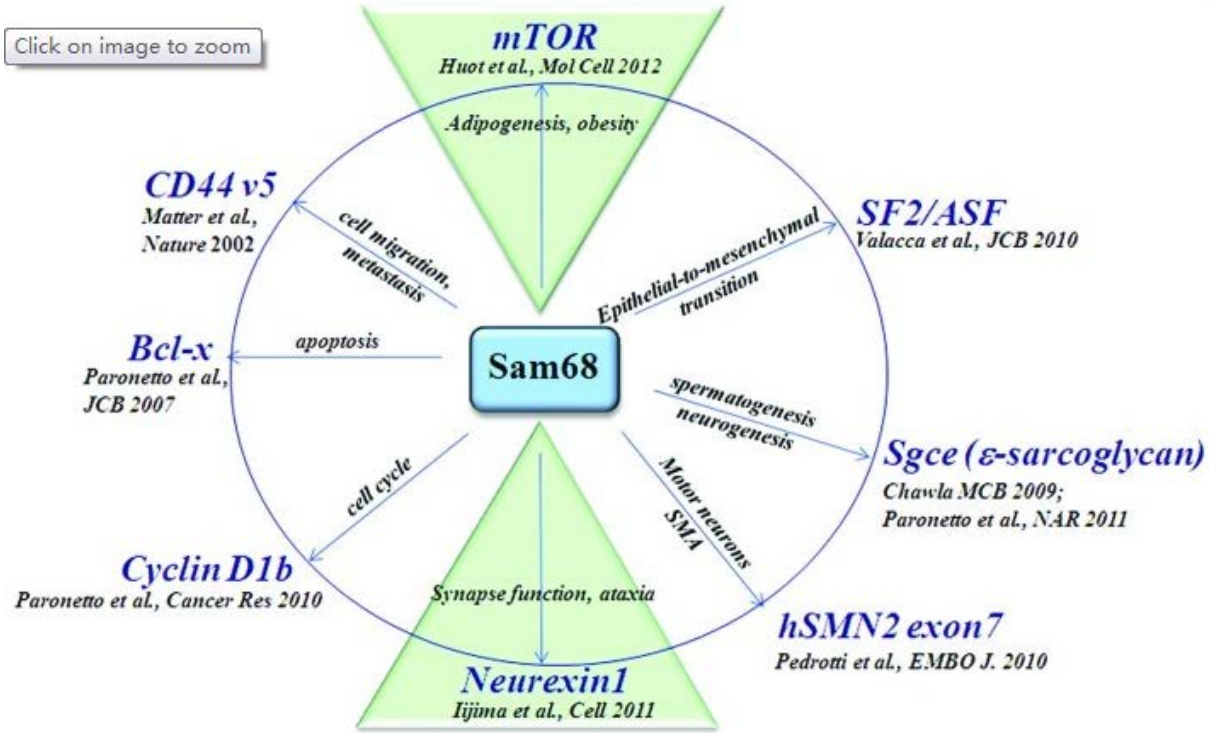


Figure 1.4 Schematic representation denoting the multiple roles of Sam68 in alternative splicing (Vogel and Richard 2012).

proteins and splicing factors (Venables, Vernet et al. 1999; Bedford, Frankel et al. 2000; Vogel and Richard 2012; Naro and Sette 2013) (Figure 1.4).

Firstly, I discuss the roles of Sam68 in alternative splicing linked with tumorigenesis, as Sam68 is a potential biomarker of various cancers (Rajan, Gaughan et al. 2008; Zhang, Li et al. 2009; Song, Wang et al. 2010; Li, Yu et al. 2012; Liao, Liu et al. 2013). Inclusion of exon v5 of CD44 pre-mRNA is observed in tumor progression, and is stimulated by the Ras signaling pathway. Sam68 was shown to bind to exon v5 sequences and promote the inclusion of v5, when Sam68 is phosphorylated by Erk in mouse T-lymphoma cells (Matter, Herrlich et al. 2002; Prochazka, Tesarik et al. 2014). More reports show that Sam68 associates with other splicing factors such as SRm160, Brm, and SND1, to coordinately regulate CD44 splicing (Batsche, Yaniv et al. 2006; Cheng and Sharp 2006; Cappellari, Bielli et al. 2014). Moreover, Sam68 affects splicing machinery in a signaling-dependent manner. U2AF (U2 snRNP auxiliary factor) participates in the key step in spliceosome assembly. The pre-mRNA occupancy of U2AF is repressed when Sam68 is phosphorylated by Erk in lymphoma cells and *in vivo*, resulting in the dysregulation of spliceosome assembly (Tisserant and Konig 2008). Additionally, Sam68 plays a role in EMT by modulating the expression of a short isoform of SF2/ASF (SRSF1), which is subjected to non-sense mediated decay (NMD) (Valacca, Bonomi et al. 2010). In prostate cancer cells, Sam68 binds to Cyclin D1 pre-mRNA and favors the splicing of D1b isoform, which is predictive of poor outcomes in patients, thereby increasing the risk of prostate cancer (Paronetto, Cappellari et al. 2010). This regulation of cyclin D1 splicing by Sam68 is also observed in breast cancer cell lines (Wu, Luo et al. 2014). Moreover, in coordination with hnRNP A1, Sam68 regulates Bcl-xl alternative splicing in HEK293 cells. Overexpression of Sam68 increases the production of the pro-apoptotic Bcl-x short isoform, which is counteracted by the transcription factor FBI-1 (Bielli, Busa et al. 2014). Phosphorylation of Sam68 at tyrosine residues switches its function from pro-apoptotic to anti-apoptotic (Paronetto, Achsel et al. 2007).

In addition to tumorigenesis, Sam68 is required for spermatogenesis, neurogenesis and adipogenesis by regulating the alternative splicing of key genes. In differentiating germ line cells, Sam68 binds to transcriptionally active chromatin sites, interacts with phosphorylated

RNA polymerase II, and promotes the alternative splicing of mRNA targets, such as murine *Sgce* (Paronetto, Messina et al. 2011). Dysregulation of Sam68 has been observed in neurogenerative diseases. Neurexins (NRXs) are the family of synaptic receptors, and their distinct alternative spliced isoforms exhibit different functions. Sam68 regulates neurogenesis by modulating alternative splicing of neurexin-1, and its activity is regulated by neuronal depolarization via calcium/calmodulin-dependent kinase IV signaling (Iijima, Wu et al. 2011). In wild-type neurons, this exon-skipping event leads to NRX protein variants with different ligand specificities (Aoto, Martinelli et al. 2013). Cerebellar Sam68$^{-/-}$ neurons fail to increase exon skipping at the alternatively spliced segment 4 of neurexin-1 upon depolarization (Iijima, Wu et al. 2011). The Fragile X-associated Tremor/Ataxia Syndrome (FXTAS) is an RNA gain-of-function disorder derived from the expansion of CGG repeats. The CGG repeats recruit intranuclear RNA aggregates containing Sam68, hnRNP G and MBNL1 (muscle blind-like 1) (Cieply and Carstens 2015). Consequently, Sam68 is sequestered by CGG repeats and its splicing function is therefore lost (Sellier, Rau et al. 2010). In another disorder termed spinal muscular atrophy (SMA), Sam68 modulates SMN splicing by repressing exon 7 inclusion, favoring the recruitment of hnRNPA1 (Pedrotti, Bielli et al. 2010; Pedrotti and Sette 2010) and interfering with U2AF65 at the 3’ splice site of exon7 (Pagliarini, Pelosi et al. 2015). Finally, Sam68 regulates the alternative splicing of mTOR during adipogenesis. Depletion of Sam68 results in mTOR intron 5 retention and decreased wild-type mTOR protein level. Thus, by modulating the alternative splicing of mTOR, Sam68 regulates the abundance of mTOR signaling pathway (Huot, Vogel et al. 2012).

1.3 RNA processing

Messenger RNA is extensively processed before being exported to the cytoplasm. After transcription begins, it undergoes 5’ end capping, splicing, and 3’ end processing, which largely determines the fate of an mRNA and this process is called RNA processing. RNA processing involves both RNA modification and the loading of various protein factors, such as RNA-binding proteins (RBPs). RBPs are categorized by their RNA-binding domains,

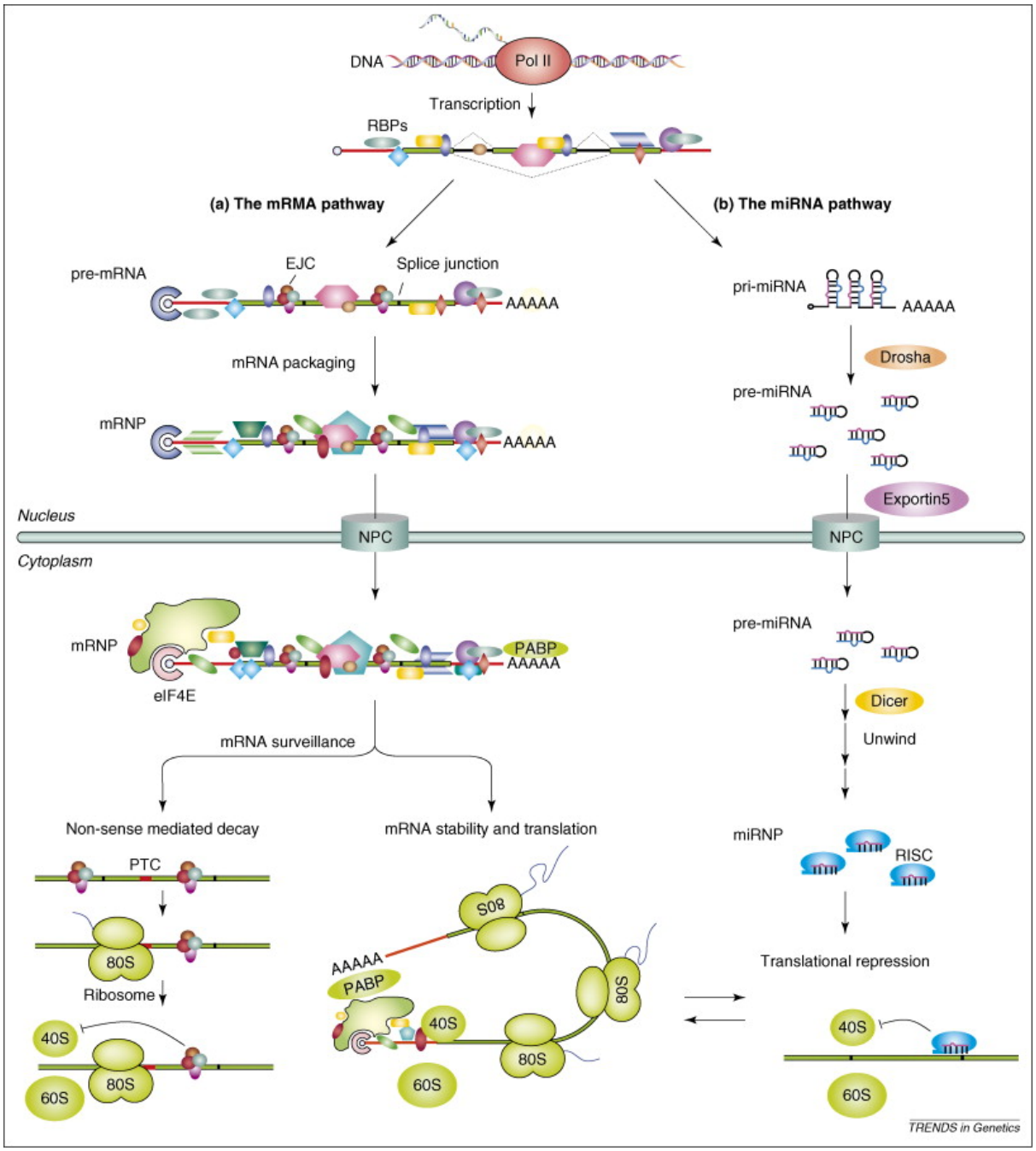


Figure 1.5 RNA processing in eukaryotes.

RNA processing is divided into nuclear processing and cytoplasmic processing. The functions of RNA-binding proteins and accessory proteins are shown. The miRNA (microRNA) processing is also modulated by specific RNA-binding proteins. RBP, RNA-binding protein. EJC, exon junction complex. PTC, premature termination codon. NPC, nuclear core complex (Lukong, Chang et al. 2008).

including the K homology (KH) domain, RNA recognition motif (RRM), the DEAD motif, the double-stranded RNA binding motif (DSRM) as well as the zinc-finger domain (Lukong, Chang et al. 2008; Calabretta and Richard 2015). Some RBPs directly participates in RNA processing, while other RBPs can be found in the complexes called EJC (exon junction complex) for splicing regulation and THO/TREX (transcription export complex) during transcription elongation (Hocine, Singer et al. 2010) (Figure 1.5). In the following, the main steps in RNA processing regulated by RBPs will be discussed.

1.3.1 Transcription

The RNA polymerase II (RNAPII) C terminal domain (CTD) contains heptapeptide repeats (YSPTSPS), which are phosphorylated during transcription on Y1, T4 and all the three serine residues (Allison, Moyle et al. 1985; Corden 1990; Egloff, Dienstbier et al. 2012). Thus, dynamic phosphorylation and dephosphorylation during the transcription cycle allows CTD to recruit various factors to regulate transcription, RNA processing and chromatin modification (Hsin, Sheth et al. 2011; Sims, Rojas et al. 2011). Generally, serine 2 (S2) and serine 5 (S5) predominate on mRNA encoding genes. S5 is phosphorylated by the general transcription factor TFIIH to play a role in transcription initiation, by recruiting mRNA-capping enzymes (Ghosh, Shuman et al. 2011). As the RNAPII escapes from the promoter region, the phosphorylation on S5 is replaced by phosphorylation on S2, catalyzed by p-TEFb, BRD4 and CDK12/CDK13 (Bartkowiak, Liu et al. 2010; Devaiah, Lewis et al. 2012; Zhou, Li et al. 2012). S2 phosphorylation is crucial for transcription termination and 3' end cleavage/polyadenylation (CPA), since termination and cleavage factors harbor CTD interaction domains (CIDs) which bind the phosphorylated CTD (Davidson, Muniz et al. 2014).

1.3.2 5'capping

5' end capping is a key determinant of RNA processing, stability, localization and translation efficiency, adding to the complexity of "epitranscriptomic" gene regulation (Topisirovic, Svitkin et al. 2011; Jaschke, Hofer et al. 2016). This process requires three

enzymes: RNA triphosphatase, guanylyltransferase, and 7-methyltransferase (Shuman 2001). After the PolII has transcribed 25-30 nucleotides of the RNA, the RNA triphosphatase firstly acts on the RNA terminus to remove the γ–phosphate. Subsequently, the guanylyltransferase converts GMP to GTP to form GpppN to be methylated. Finally, the 7-methyltransferase is required to methylate the guanine at the N7 position (Hocine, Singer et al. 2010). Alternatively, recent studies highlight various non-canonical capping structures (Shuman 2015). For instance, a *de novo* transcription initiation mechanism was discovered by Bird et al. (2016), who observed that the non-canonical initiating nucleotide (NCIN) caps, including NAD+, NADH, and dpCoA caps, are added by RNA PII *in vitro* and *in vivo*, a process which may occur in all organisms (Bird, Zhang et al. 2016).

5’ capping is tightly linked with the transcription process, via the coupling between the capping enzymes and phosphorylated CTD of RNAPII (Bentley 2015). Recently, a study incorporating mass spectrometry, cryo-electron microscopy, and protein crosslinking revealed the molecular basis of such coupling (Martinez-Rucobo, Kohler et al. 2015). This study showed that capping enzymes associate with RNAPII and scan to the end of the RNA, and bind to the unprocessed 5’-triphosphate end. Subsequently, CTD serine 5 dephosphorylation during transcription is linked with the release of capping enzymes, which occurs before the nascent transcripts reaching 500 nucleotides long (Zorio and Bentley 2004; Bentley 2015). Finally, it has been demonstrated that the pyrophosphohydrolase DXO (Dom3Z) possesses the decapping activity, which is implicated in a quality control mechanism in mammalian cells (Jiao, Chang et al. 2013).

1.3.3 pre-mRNA splicing

1.3.3.1 Mechanism of pre-mRNA splicing

RNA splicing is the process by which noncoding intervening sequences, namely introns, are removed from pre-mRNAs (Kornblihtt, Schor et al. 2013). Lately, large numbers of genome-wide studies in different organisms have shown that, the vast majority of transcripts undergo alternative splicing, providing an important source of diversity in gene expression (Barbosa-Morais, Irimia et al. 2012; Marquez, Brown et al. 2012; Mele, Ferreira et al. 2015).

There are two main steps in splicing: Firstly, the 2'-OH of the RNA nucleotide at the branch point attacks the last nucleotide of the intron, forming a lariat as the intermediate. Secondly, the 3'-OH of the exon's RNA nucleotide attacks the last nucleotide of the intron, joining two adjacent exons and releasing the lariat. Intron identification is determined by the presence of "GU" at the 5' splice site, "A" at branch point and "AG" at 3' splice site (Carpenter, Ricci et al. 2014).

Splicing is mainly carried out by spliceosome consisting of U1, U2, U4, U5, and U6 snRNPs (small nuclear RNPs), and is facilitated by various regulatory proteins (Stark and Luhrmann 2006; Irimia and Roy 2014). In the splicing process, U1 firstly binds to the 5' splice site via complementary base pairing, while U2 binds the intron branch point facilitated by U2AF (U2 auxiliary factor). Subsequently, a snRNP containing U4, U5 and U6 then facilitates the spliceosome assembly. U4 then disassociates allowing U6 to replace U1 at the 5' splice site, after which U6 interacts with U2 to link the branch point with the 5' splice site. The 5' end of the intron is then cleaved to form a lariat. Finally, U5 brings the 3' end of the upstream exon and the 5' end of downstream exon into close proximity, followed by a second cleavage reaction to join the two exons together (Douglas and Wood 2011; Cook-Andersen and Wilkinson 2015)(Figure1.6A).

1.3.3.2 Regulation of pre-mRNA splicing

Spliceosome may select different splice sites, leading to long/short isoform production or intron retention. These choices are made depending on the "strength" of the splice sites, which are regulated by the *cis*-acting elements (enhancer/silencer elements) and the presence of *trans*-acting factors. These cis-elements are much more variable than splice sites and are influenced by different pathways and stimuli, allowing the fine tuning of pre-mRNA splicing (Fu and Ares 2014). A greater understanding of splicing regulation has been achieved by various high-throughput approaches, including microarrays with splicing profiling, iCLIP, PAR-CLIP and HITS-CLIP (Hafner, Landthaler et al. 2010; Konig, Zarnack et al. 2010; Ray, Kazan et al. 2013; Weyn-Vanhentenryck, Mele et al. 2014). These analyses elucidated that global splicing is tightly modulated by snRNPs (Yoshida, Sanada et al. 2011; Xiao, Tang et al. 2012; Munding, Shiue et al. 2013).

Acting as *trans*-acting regulatory factors, RBPs play essential roles in RNA splicing, by binding enhancer and silencer elements (Figure 1.7) to dictate which splice sites are used. Serine/Arginine-rich (SR) and hnRNP proteins are accessory proteins for splicing. Large scale sequencing analysis illustrated that exonic splicing enhancers (ESEs) are bound by SR proteins, while exonic splicing silencers (ESSs) are bound by hnRNPs; other splicing factors bind to intronic splicing enhancers (ISEs) and silencers (ISSs) (Anko, Muller-McNicoll et al. 2012; Pandit, Zhou et al. 2013). By binding to adjacent elements, RBPs either act cooperatively or competitively to regulate splicing. For example, PTB and QKI overlap extensively between splicing regulatory networks, whereas hnRNPC competes with U2AF65 on Alu-associated 3' splice sites (Hall, Nagel et al. 2013; Zarnack, Konig et al. 2013). Recently, an enhanced CLIP technique has been developed to robustly address RBP binding sequences with improved specificity (Van Nostrand, Pratt et al. 2016). Moreover, to address some disadvantages of CLIP, namely the need for relatively large amounts of material and highly specific antibodies, a novel technique called TRIBE (targets of RNA-binding proteins identified by editing) has been characterized (McMahon, Rahman et al. 2016). To identify the cell-specific binding targets of an RBP, TRIBE couples the RBP of interest to the catalytic domain of the Drosophila RNA-editing enzyme ADAR (Double-stranded RNA-specific adenosine deaminase) (Vogel, Schneider et al. 2014), and expresses this fusion protein *in vivo*.

Ultimately, through alternative splicing, the cells utilize different combination of exons to make different proteins from a single gene, increasing the diversity of cellular functions.

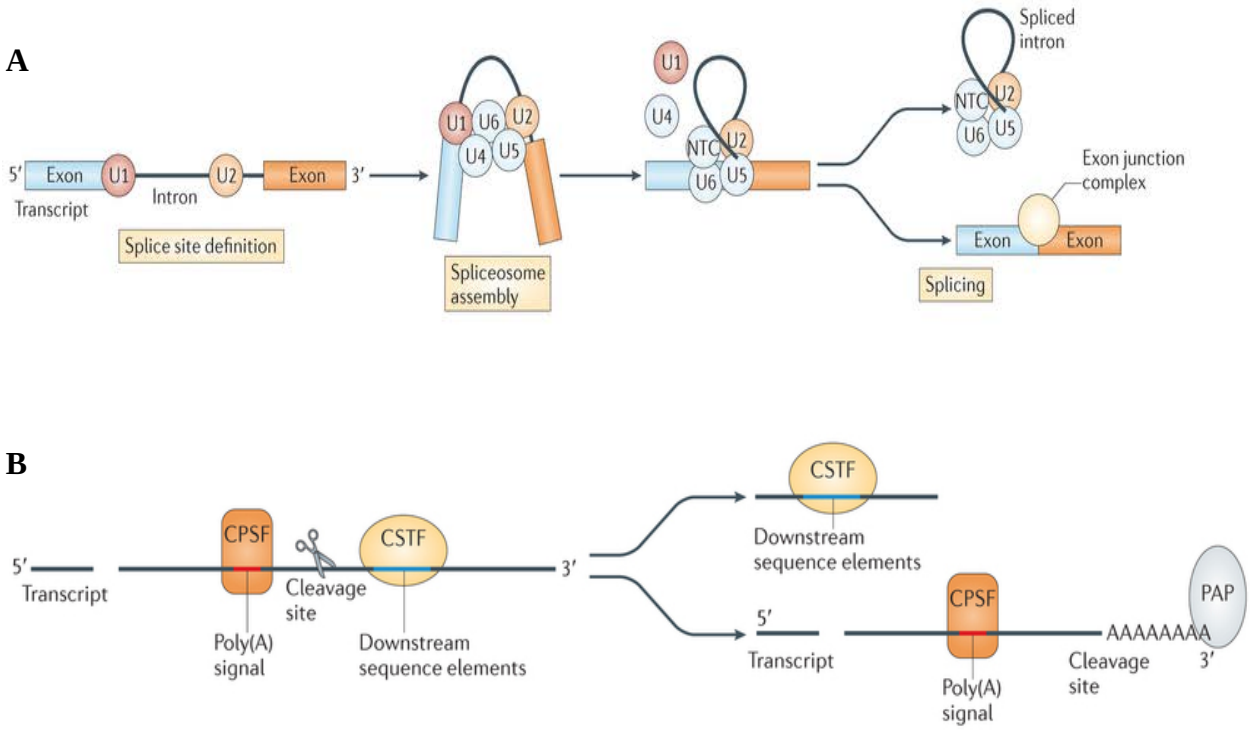


Figure 1.6 Pre-mRNA processing into mature mRNAs: intron splicing and polyadenylation.

A) Schematic representation of splicing mechanism by spliceosome. NTC, the NineTeen Complex of proteins associated with spliceosome during spliceosome activation and intron removal. B) Mechanism of transcript cleavage and polyadenylation (Carpenter, Ricci et al. 2014).

1.3.4 3’ polyadenylation

3’ end polyadenylation is the final step where endonucleolytic cleavage occurs 10-30 nucleotides downstream of an AAUAAA sequence, and polyA is added at the 3’ end by polyA polymerase (PAP) (Figure 1.6B). The polyA tail is generated for the stability of mRNA and to ensure translation efficiency (Proudfoot 2004). However, this view has been recently challenged by Park et al. (2016), who demonstrated that the length of polyA tail and translation efficiency is correlated only to a threshold of ~20 nucleotides during cell cycle, as shown by TAIL-seq and ribosome profiling (Park, Yi et al. 2016). Multiple proteins participate in 3’ end polyadenylation, such as CPSF (cleavage/polyadenylation specificity factor), CtsF (cleavage stimulation factor) and two cleavage factors (CFIm and CFIIm) (Shi, Di Giammartino et al. 2009; Proudfoot 2011). CPSF and CtsF are highly conserved in human and are responsible for cleavage and polyadenylation. Interestingly, these 3’ end formation factors also play a role in alternative splicing (Misra and Green 2016). CPSF recognizes the polyA signal and recruits PAP, while CtsF recognizes U/GU-rich sequences and is directly involved in polyadenylation (Figure 1.6B). Cleavage by CFIm and CFIIm is coupled to polyA addition, in a process which also requires PABP (PolyA binding protein) and various other factors. PABP facilitates PAP to catalyze the polyA addition, modulates polyA length, and directly stimulates translation termination as shown recently (Ivanov, Mikhailova et al. 2016).

Similarly to alternative splicing, RNA transcripts can be alternatively polyadenylated via the recognition of alternative polyA signals, generating different isoforms or transcripts with distinct stabilities or localizations. Alternative polyadenylation may be tissue-specific and coupled with alternative splicing, and thus it may contribute to different diseases (Erson-Bensan 2016). Indeed, large numbers of genes contain various potential polyadenylation and cleavage sites, and different polyA site selections make it possible to achieve an additional layer of gene expression complexity (Erson-Bensan and Can 2016). Although the detailed mechanism is still under debate, it has been demonstrated that the U1 snRNP controls the polyadenylation site selection and thus can modulate the expression of genes involved in tumorigenesis (Kaida, Berg et al. 2010; Berg, Singh et al. 2012). Studies

also have concluded that alternative polyadenylation is closely linked with tumorigenesis (Masamha, Xia et al. 2014; Xia, Donehower et al. 2014).

1.3.5 RNA export and degradation

At the last step of RNA processing, the mature mRNA is exported from the nucleus to the cytoplasm through the NPC (nuclear pore complex), facilitated by the TAP/NXF1-dependent pathway and the THO/TREX complex, as well as the EJC (Hocine, Singer et al. 2010; Okamura, Inose et al. 2015). EJC deposited on exon-exon junctions is a crucial player in pre-mRNA processing. As a dynamic structure consisting of both RBPs and RNAs, EJC forms a RNP harboring at least 10 proteins (Lukong, Chang et al. 2008; Zhang and Sachs 2015). EJC modulates RNA non-sense mediated decay and enhances translation initiation (Figure 1.5) (Hocine, Singer et al. 2010). The presence of EJC downstream of a stop codon triggers the degradation of the mRNA (Muhlemann 2016). Thus, the EJC increases the efficiency of normal mRNA translation, while facilitates the degradation of the aberrant mRNAs (Stalder and Muhlemann 2008; Celik, Kervestin et al. 2015). It also has been shown that SR proteins and EJC function together for efficient nuclear export, linking pre-RNA splicing to mature RNA export (Singh, Kucukural et al. 2012; Muller-McNicoll and Neugebauer 2013).

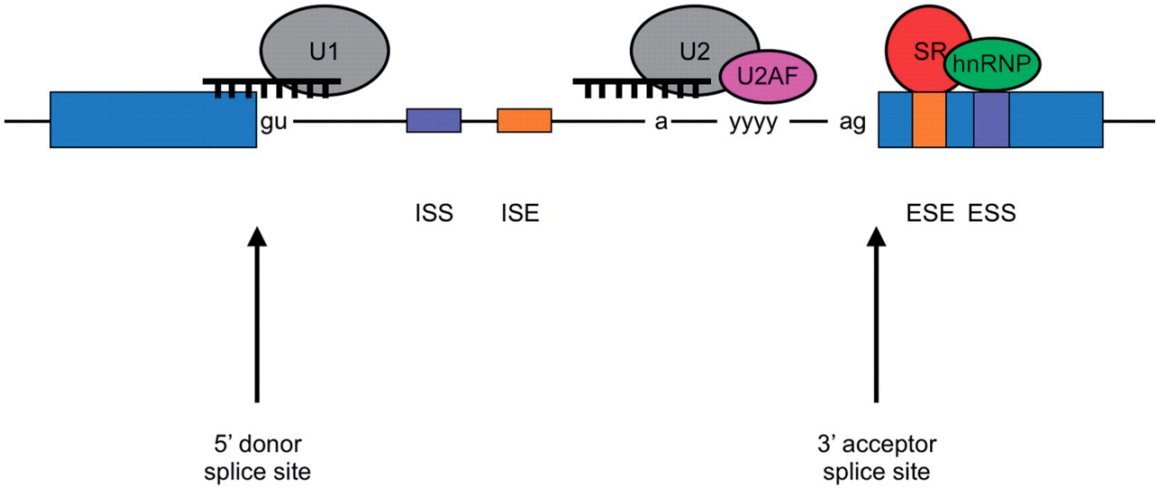


Figure 1.7 *Cis*-acting elements and *trans*-acting factors control alternative splicing. ISS, intron splicing silencer; ISE, intron splicing enhancer; ESE, exon splicing enhancer; ESS, exon splicing silencer (Douglas and Wood 2011).

1.4 Protein ubiquitination

1.4.1 Ubiquitination process

Ubiquitination is a post-translational modification in which ubiquitin is attached to a lysine residue of the substrate. Ubiquitination was initially known to be associated with protein degradation, and was more recently shown to function in signal transduction as well, by mediating protein-protein interactions to affect their subcellular localization (Komander and Rape 2012). Ubiquitin (Ub) is an 8.5 kDa small protein which requires 3 enzymes to be attached to a protein: the E1 ubiquitin-activating enzyme, E2 ubiquitin-conjugating enzyme, and E3 ubiquitin ligase. There are different types of E3 ligases, including the family of homologous to the E6AP carboxyl terminus domain (HECT), the group of really interesting new gene (RING), the RING-related E3s, as well as members of the U-box family (Metzger, Hristova et al. 2012). Ubiquitination consists of three steps. Firstly, the C-terminus of Ub forms a thioester bond with the cysteine of E1 in an ATP-dependent manner. Next, Ub is transferred to E2 via the catalytic cysteine. Lastly, E3 binds the Ub-E2 complex and/or the substrate to transfer Ub to the substrate's lysine. In this process, HECT E3s catalyze a thioester intermediate formation between the Ub and E3, whereas the vast majority of E3s, namely RING family E3s, mediate the direct transfer of Ub from E2 to the substrate (Metzger, Hristova et al. 2012) (Figure 1.8A). Thus, the substrate is monoubiquitinated via an isopeptide bond. Substrates can be ubiquitinated on multiple lysines, termed multiubiquitination; otherwise specific E2/E3 combinations are able to catalyze multiple cycles of ubiquitination on one lysine, forming polyubiquitination (Komander and Rape 2012). Different ubiquitination patterns target proteins for various fates: monoubiquitination generally regulates DNA repair and gene expression; K48 polyubiquitination mostly targets proteins for proteasomal degradation; while others such as K63, K29 and K11 polyubiquitinations preferentially modulate endocytosis, kinase activation, signaling transduction and translation (Figure 1.8B) (Di Lello and Hymowitz 2016; Hu and Hochstrasser 2016).

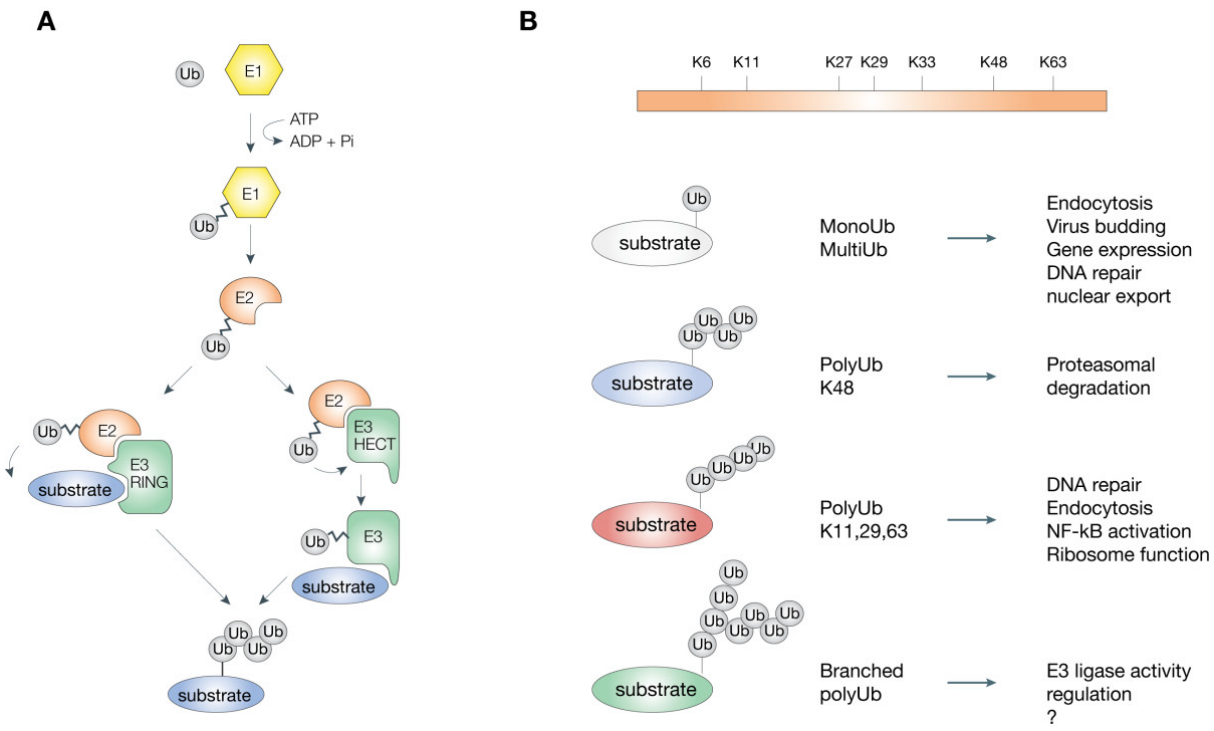


Figure 1.8 The ubiquitination process and functions.

A) Schematic representation of the ubiquitination process. Ubiquitin-activating (E1), ubiquitin-conjugating (E2) and ubiquitin-protein ligase (E3) enzymes are depicted. B) Schematic representation of the different ubiquitination with various linkages (K6, K11, K27, K29, K33, K48 and K63) and their functions (Woelk, Sigismund et al. 2007).

1.4.2 Deubiquitination

Deubiquitinating enzymes (DUBs) catalyze the removal of Ub chains. By participating in the quality control step of ubiquitination, DUBs prevent ubiquitin ligases from being constitutively on, and remove the Ub-chain from substrates (McDowell and Philpott 2013).

There are various types of DUBs. Human cells contain approximately 55 USPs, 14 OTUs (ovarian tumor DUBs), 10 JAMM family DUBs, 4 UCHs (ubiquitin C-terminal hydrolases) and 4 Josephin domain DUBs (Clague, Barsukov et al. 2013). Some of the DUBs are housekeeping enzymes (USP14, UCH37 and RPN11) which protect ubiquitin from degradation, in order to keep enough free Ub for chain-reassembly (Pfoh, Lacdao et al. 2015). A comprehensive analysis reported that most USPs are active against all ubiquitin linkages, showing that most of the USP family members are non-specific to the ubiquitin code but specific to their substrates (Komander and Rape 2012). Other families of DUBs are linkage-specific. For instance, JAMM family DUBs including AMSH-LP, AMSH, BRCC36, and POH1 (PSMD14), are often K63-specific (Komander and Rape 2012). These DUBs are important in cellular functions, and the dysregulation of DUBs in tumorigenesis has been intensively studied (Bhattacharya and Ghosh 2014; Pal, Young et al. 2014).

BRCC36 is one of the components of BRCC36 isopeptidase containing complex (BRISC), which comprises RAP80, ABRAXAS (CCDC98), BRE (BRCC45), MERIT40, and BRCA1 (Feng, Wang et al. 2010). This complex is localized in the nucleus where it plays a role in DDR. The E3s RNF8 and RNF 168 generate Ub-rich chains on histones, providing docking sites to recruit DNA repair enzymes including BRCA1, Rap80 and other checkpoint molecules (Al-Hakim, Escribano-Diaz et al. 2010). Subsequently, Rap80 targets the BRISC complex to the DSB loci, contributing to the tight control of BRCA1-dependent DSB repair (Sobhian, Shao et al. 2007; Panier and Durocher 2013). It has been recently shown that BRCC36 plays a role in 5-fluorouracil induced DDR in colorectal cancer cells, which is dependent on B7-H3 expression (Sun, Zhang et al. 2016). BRCC36 can also associate with Abro1 (KIAA0157) instead of ABRAXAS to form a cytosolic complex, whose function is poorly defined. It has been reported that serine hydroxymethyltransferase (SHMT) directs BRISC activity at K63 ubiquitin chains conjugated to the type I interferon receptor 1

(IFNAR1), which mediates type I interferon signaling (Zheng, Gupta et al. 2013).

1.4.3 Cellular functions of ubiquitination

1.4.3.1 Proteasomal degradation

Intracellular degradation of soluble proteins is generally mediated by the 26S proteasome. K48 is the most abundant linkage in all organisms, and K48 ubiquitination level increases dramatically when proteasomes are inhibited (Kim, Bennett et al. 2011). It has been reported recently that multiple ubiquitinated K48 lysines are more accessible for proteasomal degradation than K48-polyubiquitin chains, suggesting that proteasomal degradation efficiency is largely dependent on the number of K48 chains rather than the length (Lu, Lee et al. 2015). Additionally, homotypic K11 polyubiquitin chains do not tightly bind to proteasomal receptors, whereas heterotypic K11/K48-polyubiquitin chains bind to the proteasome and stimulate protein degradation during mitosis (Grice and Nathan 2016). In mitosis, APC/C (anaphase promoting complex/cyclosome) E3 ligase catalyzes ubiquitin chain formation to direct the degradation of cell-cycle related proteins (Min and Lindon 2012; Meyer and Rape 2014).

1.4.3.2 Lysosomal degradation

The degradation of plasma membrane-associated proteins occurs in lysosomes, a process which is catalyzed by K63 ubiquitination. Ubiquitination is initiated at the membrane, whereby the addition of Ubs targets the substrate to the endocytosis. Accordingly, deubiquitination by DUBs facilitates substrate recycling back to the plasma membrane (Clague, Liu et al. 2012). A mass spectrometry-based targeted proteomics study revealed that activated EGFR is linked via a K63 ubiquitinaion chain, which is required for its endocytosis and degradation (Huang, Zeng et al. 2013). Moreover, ESCRT (endosomal sorting complexes required for transport) complex recognizes the surface of endocytosis vesicles or endosomal membrane lipids, acting as an effector to facilitate proteolysis. The ubiquitinated membrane substrates are bound by the ESCRT complex with a modest preference for binding to K63 linkage, implying the role of K63 ubiquitination in proteolysis (Grice and Nathan 2016).

1.4.3.3 Protein-protein interaction

Ubiquitination also participates in signal transduction by modulating protein-protein interactions, protein localization and substrate activity. Monoubiquitination has been shown to be sufficient for recruiting protein binding partners in response to DNA damage. Monoubiquitinated PCNA recruits the Y family of DNA polymerase, rescuing stalled replication forks from collapsing. Subsequently, USP1 deubiquitinates PCNA to turn off the recruitment signal, allowing the replication machinery back to normal (Huang, Nijman et al. 2006; Liang, Dexheimer et al. 2014).

K63 linkage ubiquitination also regulates protein-protein interactions. K63 modification increases the affinity of the yeast splicing factor Prp3 for a U5 snRNP component (Prp8), thereby stabilizing U4/U6/U5 snRNPs (Song, Werner et al. 2010). Additionally, K63 modification on the ribosomal protein L28 stabilizes the polysomes and promotes translation (Spence, Gali et al. 2000). Building upon on this, it was recently shown that K63 accumulation in polysomes facilitates mRNA translation and cellular survival in response to oxidative stress in yeast (Silva, Finley et al. 2015). Furthermore, K63 ubiquitination plays a significant role in DNA repair. It was recently proposed that RNF8 ubiquitylates the H1-type linker histones in response to DNA damage, providing the initial binding platform for RNF168, which subsequently ubiquitylates H2A at K13/K15. This indicates that a dynamic histone code is formed to repair DNA damage by RNF8 and RNF168, the respective K63 ubiquitination writer and reader (Thorslund, Ripplinger et al. 2015).

1.4.3.4 Protein localization and activity

Activation of NF-κB in response to external stimuli, such as TNFα, is dependent on K63 ubiquitination. TNFα triggers the formation of a K63 chain on TRAF6 and K11/K63 chain on cIAP1 (Xu, Skaug et al. 2009), while cIAPs inhibits non-canonical NF-κB activation by stimulating K48-linked proteasomal degradation (Nachbur and Silke 2016). K63 also mediates the activation of TAK1/IKK kinase and the antiviral protein RIG-I (Oshiumi, Miyashita et al. 2013). The K63-modified RIG-I is able to dimerize, after which it is able to facilitate downstream signaling events (Hou, Sun et al. 2011).

Ubiquitination also plays a role in protein localization. Upon the monoubiquitination of lysine residues at its C-terminus, p53 is exported out of the nucleus, a process which is

reversed by the deubiquitinating activity of USP10. Since the nuclear export signal (NLS) of p53 is in the proximity of its C-terminus, it is proposed that the monoubiquitination of p53 affects its localization by modulating the accessibility of its NLS to the export machinery (Ayroldi, Petrillo et al. 2015; Wei, Yang et al. 2015). Under stress conditions, monoubiquitinated p53 is modified into a polyubiquitinated state with the help of MDM2 (Murine Double Minute 2), while deubiquitination of p53 translocates it to the outer mitochondrial membrane, where it plays a significant role in apoptosis (Dai, Luo et al. 2016).

Thus, by affecting the intra- or inter- molecular binding events, ubiquitination is involved in a variety of cellular processes.

Objective and Hypothesis

RGG/RG motifs are preferred substrates of arginine methyltransferases, PRMTs. The goal of my work was to identify the functions of two RGG/RG motif containing-proteins: Aven and Sam68. For Sam68 project, Richard lab has reported that Sam68 regulates alternative splicing of mTOR and loss of Sam68 leads to the lean phenotype of mice. However, full-length mTOR could not fully rescue the adipogenesis defect in Sam68-depleted preadipocytes (Huot, Vogel et al. 2012). My objective was to examine other mechanisms that contribute to lean phenotype of Sam68-deficient mice. Thus I hypothesize that Sam68 regulates the alternative splicing of other mediators in mTOR signaling, whose dysregulation leads to adipogenesis defect.

For Aven project, Richard lab and other groups have identified Aven contains an uncharacterized RGG/RG motif at its N-terminus (Thandapani, O'Connor et al. 2013). Studies have reported that Aven is a survival factor, which is overexpressed in several acute leukemias (Paydas, Tanriverdi et al. 2003; Choi, Hwang et al. 2006). Our objective was to determine the function of RGG/RG motif of Aven in acute leukemic cells. Herein, we hypothesize that Aven regulates leukemogenesis, and this regulation is mediated by its RGG/RG motif.

Chapter 2

RNA G-quadruplexes and their potential regulatory roles in translation

2.1 Preface

G-quadruplex is prevalent in DNA and RNA. In this review, we focus on RNA G-quadruplex in 5’ UTR, 3’ UTR and open reading frames and summarize recent studies revealing the functions of G4s in mRNA translation.

2.2 Abstract

DNA guanine (G)-rich four-stranded helical nucleic acid structures called G-quadruplexes (G4), have been extensively studied during the last decades. However, emerging evidence reveals that 5'- and 3'-untranslated regions (5'- and 3'-UTRs) as well as open reading frames (ORFs) contain putative RNA G-quadruplexes. These stable secondary structures play key roles in telomere homeostasis and RNA metabolism including pre-mRNA splicing, polyadenylation, mRNA targeting and translation. Interestingly, multiple RNA-binding proteins such as nucleolin, FMRP, DHX36, and Aven were identified to bind RNA G-quadruplexes. Moreover, accumulating reports suggest that RNA G-quadruplexes regulate translation in cap-dependent and -independent manner. Herein, we discuss potential roles of RNA G-quadruplexes and associated *trans*-acting factors in the regulation of mRNA translation.

2.3 Structure of G-quadruplexes

In 1910, Bang demonstrated that guanylic acid forms gels at high concentration, providing the first evidence that guanine (G)-rich sequences may form higher-order structures (Bang 1910). Fifty years later, Gellert et al. (1962) reported that guanylic acid has the ability to form tetrameric structures by self-association (Gellert, Lipsett et al. 1962). These non-canonical structures were identified in conserved DNA sequences of telomeres and shown to form so-called G-quadruplexes (G4) structures *in vitro* (Lipps, Gruissem et al. 1982; Sundquist and Klug 1989). Since then, increasing evidence shows that both DNA and RNA containing spaced guanine repeats form G4 structures (Sundquist and Heaphy 1993; Tuesuwan, Kern et al. 2008). G-quadruplexes are folded in G-quartets that are square planar arrangements formed via Hoogsten pairing of adjacent guanines (Bochman, Paeschke et al. 2012; Bugaut and Balasubramanian 2012; Millevoi, Moine et al. 2012; Murat and Balasubramanian 2014; Rhodes and Lipps 2015; Mendoza, Bourdoncle et al. 2016). The G-quartets lay on the top of each other to form four-stranded structures which are stabilized by a cation positioned in the middle of the tetrads with preference for potassium (Figure 2.1). The ability of potential G4 sequences to form G-quadruplexes is therefore influenced by the nature of the central cation, the number of stacking G-quartets, the length of the sequences connecting the strands (at least in DNA G-quadruplexes), the direction of the strands, and the presence of an alternative Watson-Crick pair-based stable structure (Mukundan and Phan 2013; Beaudoin, Jodoin et al. 2014).

2.4 DNA G-quadruplexes

Bioinformatic search using a script that identifies “regular” G-quadruplex (i.e. $G_xN_{1\text{-}7}G_xN_{1\text{-}7}G_xN_{1\text{-}7}G_x$, where G stands for guanine, N can be any nucleotide (A, G, C, U) and $x \geqslant 3$) has shown that the human genome harbors ~376,000 potential G-quadruplex sequences (pG4) (Huppert and Balasubramanian 2005; Todd, Johnston et al. 2005). With the development of high resolution sequencing-based methods > 716,000 G4 structures were identified, where ~451,000 were not predicted by computational methods (Rodriguez and

Miller 2014; Chambers, Marsico et al. 2015). These sequences are highly conserved in mammals, with limited conservation in evolutionarily lower organisms (Konig, Evans et al. 2010; Frees, Menendez et al. 2014). Genome-wide searches using G-quadruplex-specific probes and structure-based pull-down strategies have revealed that the pG4 sequences are not randomly distributed, being enriched at telomeres, promoter regions and replication origins (Lam, Beraldi et al. 2013; Rhodes and Lipps 2015). The highest abundance of pG4 is in human telomeres, where the G4 formation protects the end of chromosomes by inhibiting telomerase activity (Healy 1995; Todd, Johnston et al. 2005; Maizels and Gray 2013). Abnormal telomerase overexpression has been observed in > 85% of cancers. Thus, extensive investigations were elicited on emerging anti-cancer therapies using small molecules to stabilize telomeric G-quadruplexes as a means to inhibit telomerase (Wang, Liu et al. 2011; Shalaby, Fiaschetti et al. 2013; Crees, Girard et al. 2014; Moye, Porter et al. 2015). Moye et al. (2015) characterized the stable human telomeric G-quadruplexes and demonstrated that these G4s are able to extend into parallel, intermolecular conformations, aligning with the intrinsic RNA moiety of the human telomerase RNA (hTR). They also showed that telomerase colocalizes with a subset of telomeric G4 structures *in vivo* (Moye, Porter et al. 2015). Additionally, structural and computational analysis revealed that pG4 sequences are found in ~40% of gene promoter regions, mostly acting as transcriptional repressors (Balasubramanian, Hurley et al. 2011; Maizels and Gray 2013; Valton, Hassan-Zadeh et al. 2014). It has been shown that pG4 sequences occur with high frequency in the promoter regions of genes encoding oncogenes such as *c-Myc, c-Kit, KRAS, PDGF-A* (platelet derived growth factor-A), and *hTERT* (human telomerase reverse transcriptase), while tumor suppressor genes correlate with low pG4 abundance in their promoter regions (Brooks, Kendrick et al. 2010; Millevoi, Moine et al. 2012; Morgan, Batra et al. 2016). Recently, Onel et al. (2016) showed that *Bcl-2* forms G-quadruplexes in its promoter by nuclear magnetic resonance (NMR) spectroscopy and dimethylsulfate (DMS) footprinting assays, and these G4s were shown to inhibit transcription by promoter-driven luciferase assay (Onel, Carver et al. 2016). Another study also demonstrated that the human tyrosine hydroxylase (*hTH*) gene harbors a G4 structure in the 3' proximal promoter region, acting as a necessary element for

transcriptional regulation. Since hTH is linked to several neurological and psychiatric disorders such as Parkinson's disease and Schizophrenia (Farhath, Thompson et al. 2015, these findings suggest that promoter G4 sequences may be linked with these diseases. Furthermore, emerging evidence suggests that ~90% of DNA replication origins contain pG4 sequences (Cayrou, Coulombe et al. 2011; Besnard, Babled et al. 2012; Bochman, Paeschke et al. 2012; Cayrou, Coulombe et al. 2012). Recently, it was discovered by performing nascent strand sequencing that G4 sequences may also position nucleosomes at a subset of human replication origins (Foulk, Urban et al. 2015). Moreover, it was reported that Potential DNA:RNA Hybrid G-Quadruplex Sequences (PHQS) are present in > 97% of human genes, and these PHQS may modulate transcription (Zheng, Xiao et al. 2013). Additional studies show that helicases, which are the molecular motors to unwind DNA and RNA, are involved in the active resolution of G4s. The best-characterized DNA G4 helicase are Pif1, RecQ, FANCJ, DDX11, BLM (Bloom syndrome protein) and WRN (Werner syndrome protein) (Mendoza, Bourdoncle et al. 2016). Known RNA-DNA G4 helicases are DHX9 and RHAU (DHX36) (Chakraborty and Grosse 2011; Chen, Murat et al. 2015).

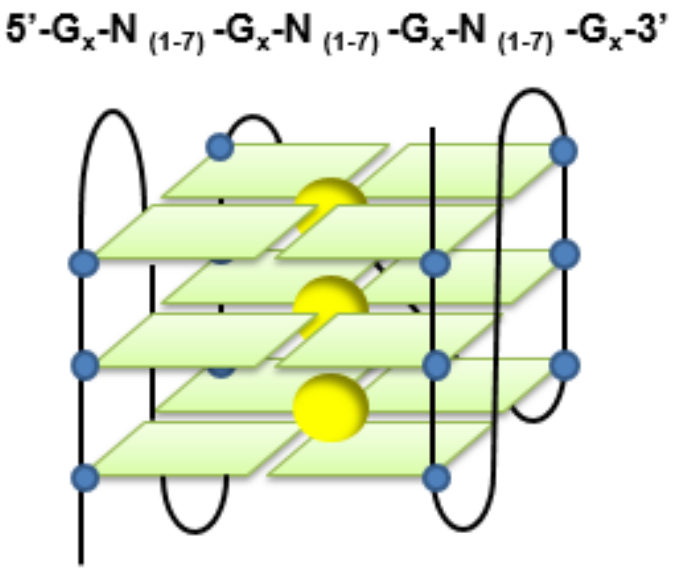


Figure 2.1 Structure of G-quadruplexes.
Guanine-rich sequences fold into G-quadruplex structures (G4 structures), composed of planar G-quartets. This representation is a G-quadruplex parallel structure that could be observed in RNA as well as DNA molecules. The guanines are represented as blue, while cations are yellow.

2.5 RNA G-quadruplexes

G-quadruplexes also form in RNA and are more stable than their DNA counterparts (Bugaut and Balasubramanian 2012). RNA G-quadruplexes almost exclusively adopt a parallel conformation in which the four strands all have the same directionality. The 2'-hydroxyl group of the ribose locks the RNA in an anti-conformation, which favors the parallel topology. Consequently, RNA G-quadruplexes have less topological diversity than DNA G-quadruplexes.

In eukaryotes, RNA G4 structures are enriched at telomeres and within specific protein encoding transcripts i.e. mRNAs. In telomeres, the G-rich TERRA (telomeric repeat-containing RNAs, or TelRNA) RNA which is transcribed from the human C-rich telomeric DNA, was shown to form G4 structures *in vitro* (Xu, Kaminaga et al. 2008; Takahama, Takada et al. 2013) and *in cellulo* (Xu, Suzuki et al. 2010). The high resolution of the G4 structure in TERRA revealed that the 2'-hydoxyl group provides intramolecular hydrogen bonding within the parallel-stranded structures, and this is important for ligand targeting and higher-order arrangement (Martadinata and Phan 2013). The human telomerase RNA forms G4 structures at its 5' end in the presence of potassium, as visualized using gel electrophoresis and UV, CD, and NMR spectroscopy (Gros, Guedin et al. 2008; Martadinata and Phan 2014).

RNA G-quadruplexes are frequently present in mRNAs. Kumari et al. (2007) demonstrated that among all the genes in the human transcriptome available at that time, ~3,000 of the 5'-UTRs were identified to possess at least one pG4 sequence (Kumari, Bugaut et al. 2007). This was revised, when Beaudoin and Perreault (2010) identified 9,979 5'-UTRs to contain at least one pG4 sequence, amongst the 124,315 transcripts from the human UTRfull dataset based on a "regular" definition of G-quadruplex (Beaudoin and Perreault 2010). Subsequently, it was discovered that 1,453 human pG4 sequences possess two short distal loops of 1 nucleotide in length and a long central loop of up to 70 nucleotides long in the 5'-UTRs, which significantly expands the number of pG4s in the transcriptome (Jodoin, Bauer et al. 2014). Moreover, a bioinformatics search for pG4 sequences in mRNA-coding regions was performed, and it reveals ~1600 pG4s in human ORFs (Thandapani, Song et al.

2015). A bioinformatic search for “regular” G-quadruplex in 3’-UTRs of the human transcriptome, detected 8,903 pG4 sequences showing that the enrichment for G-quadruplex is not limited to the 5’-UTRs (Beaudoin and Perreault, 2013). Additionally, other bioinformatic analyses, in-line probing and luciferase reporter assays revealed the existence of pG4 sequences in both the 5’- and 3’-UTR of transcripts, indicating G4 structures function in the post-transcription regulation *in cellulo*, and neighbouring C-rich sequences that affect G4 folding (Huppert, Bugaut et al. 2008; Beaudoin, Jodoin et al. 2014; Jodoin, Bauer et al. 2014). Recently, the existence of G-quadruplex formation in RNA was further confirmed in human cells by using stabilizing ligands that specifically trap RNA G-quadruplexes (Biffi, Di Antonio et al. 2014). The authors were able to visualize G4 structures in the cytoplasm of human cells using G-quadruplex-specific antibodies.

While recent reviews have broadly covered the G4 structures in DNA and RNA molecules as well as their functions in telomere maintenance and RNA metabolism in physiology and pathology conditions (Millevoi, Moine et al. 2012; Murat and Balasubramanian 2014; Rhodes and Lipps 2015), we will discuss in detail the emerging areas of RNA G4 structures in mRNA translation and known *trans*-acting factors that bind these structures.

2.6 G-quadruplexes in the control of mRNA translation

mRNA translation is one of the most fundamental processes in RNA metabolism, and its regulation is tightly controlled. Protein synthesis is composed of four main steps: translation initiation, elongation, and termination as well as ribosome recycling (Hershey, Sonenberg et al. 2012). For most eukaryotic mRNAs, translation initiation involves the association of the 7-methylguanosine cap with the cap-binding complex called eukaryotic initiation factor 4F (eIF4F, Figure 2.2A) (Hinnebusch 2014; Hinnebusch, Ivanov et al. 2016). eIF4F complex contains the cap-binding protein eIF4E, the scaffold protein eIF4G that bridges interaction between eIF4F and multifactor complex (MFC), thus allowing recruitment of the mRNA to the ribosome, and the ATP-dependent DEAD box RNA helicase eIF4A required to unwind

secondary structures in the 5’-UTRs (Sonenberg and Hinnebusch 2009; Jackson, Hellen et al. 2010). The association between the mRNA and eIF4F is the first step of translation, followed by the recruitment of 43S initiation complex, composed of the 40S ribosomal subunit, the eukaryotic initiation factors eIF3, eIF1, eIF1A and eIF5, as well as the ternary complex containing methionine-loaded tRNA, eIF2 and GTP (Figure 2.2A) (Holcik and Sonenberg 2005; Hinnebusch 2014). The 43S complex recognizes the initiation codon, where it is joined by a 60S ribosomal subunit to form the 80S ribosome (Sonenberg and Hinnebusch 2009; Jackson, Hellen et al. 2010). Excessive secondary structures in 5’-UTRs impede mRNA translation in a cap-dependent manner in eukaryotes (Pelletier and Sonenberg 1985; Koromilas, Lazaris-Karatzas et al. 1992). The 3’-UTRs also participate in translational regulation, where the added poly (A) tail is bound by the poly(A)-binding protein (PABP) and eIF4G, resulting in the circularization of mRNAs and enhanced overall initiation rate (Sonenberg and Hinnebusch 2009). An alternative mode of translation is driven by IRES (Internal Ribosome Entry Sites), and it occurs in a cap-independent mode (Komar and Hatzoglou 2011). Independent of the presence or integrity of several canonical initiation factors (especially eIF4E), IRES directly recruits ribosomes, bypassing the requirement for the 5’cap and eIF4E. Efficient IRES-driven translation is facilitated by the IRES *trans*-acting factors (ITAFs, Figure 2.2B) (Hellen and Sarnow 2001; Holcik and Sonenberg 2005). Collectively, secondary structure in 5’-UTR is thought to have a major impact on translation efficiency (Hinnebusch, Ivanov et al. 2016).

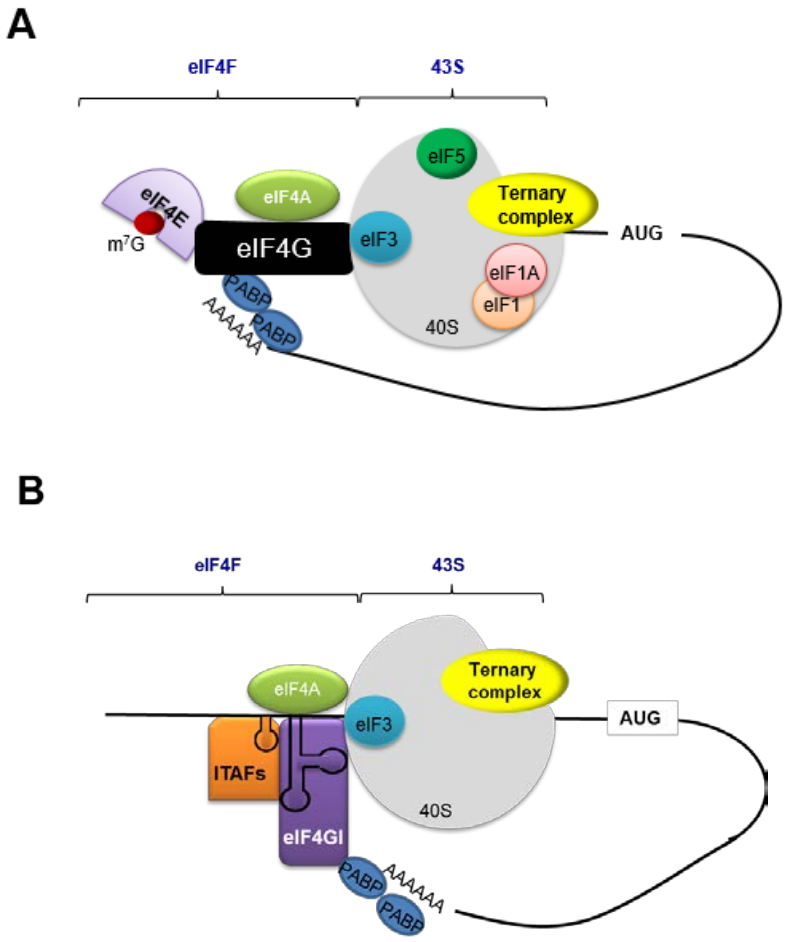


Figure 2.2 Cap-dependent and Cap–independent translation initiation.
A) In cap-dependent translation, eIF4E binds to the 5'-m^{7}GpppN of the mRNA (m^{7}G). The capped 5'-end is associated with 43S complex by a bridging protein called eIF4G. eIF4G is also bound to eIF4A, the RNA helicase that unwinds 5' secondary structures. PABP binds the poly (A) tail and brings the 5'-end and 3'-end of the mRNA together through the interaction with eIF4G. eIF3, eIF5, eIF1/eIF1A and ternary complex are shown as represented.
B) ITAFs and eIG4GI (also known as p97/DAP5/NAT1, purple) facilitates IRES cap-independent translation (Holcik and Sonenberg 2005).

2.6.1 G-quadruplexes in 5’-UTR and translational control

Kumari et al. (2007) reported that pG4 sequences in the human *NRAS* mRNA are conserved in different organisms (Kumari, Bugaut et al. 2007). They documented the formation of G4 structures *in vitro* by circular dichroism (CD) spectroscopy and UV-melting experiments, while luciferase reporter assays revealed that the RNA G4 in 5’-UTR of *NRAS* inhibits translation by ~80% in rabbit reticulocyte lysates (Kumari, Bugaut et al. 2007). Moreover, it was established that the human ZIC-1 mRNA forms a 27 nucleotide G4 structure within its 5’-UTR and represses protein production by ~80% in HeLa cells using the dual-luciferase plasmid based assay (Arora, Dutkiewicz et al. 2008). The presence of G4 structures in 5’-UTR of various human mRNAs and multiple strategies such as bioinformatic analyses, mutagenesis and reporter gene-based expression assays showed that G4s in 5’-UTRs correlate with translational repression of various mRNAs including *MT3-MMP* (Morris and Basu 2009), *ERS1*(Balkwill, Derecka et al. 2009), *BCL-2* (Shahid, Bugaut et al. 2010), *TRF2* (Gomez, Guedin et al. 2010), *ADAM10* (Lammich, Kamp et al. 2011; Dai, Liu et al. 2015) and *TGFβ2* (Agarwala, Pandey et al. 2013) (Figure 2.3). Moreover, in-depth analysis using CD spectroscopy and in-line probing, identified several 5’-UTRs that harbor pG4 sequences including *EBAG9*, *AASDHPPT*, *FZD2*, *BARHL1*, *NCAM2*, and *THRA* (Beaudoin and Perreault 2010) (Figure 2.3). Most of these genes are involved in transcriptional regulation, protein modification, G-protein-mediated signaling, cation transport and developmental processes. The C-to-A substitution, known to destabilize G4 formation, was able to rescue the repressed translation of all but one gene (Beaudoin and Perreault 2010). Recently, G4 structures with longer central loops (> 7 nucleotides) in the *HIRA*, *TOM112* and *APC* 5’-UTR were also shown to have the ability to repress translation when tested by luciferase reporter assays (Jodoin, Bauer et al. 2014) (Figure 2.3). Similar conclusions were reached in the study where the “irregular” G4 structures were discovered in the *H2AFY* and *AKIRIN* 5’-UTR (Rouleau, Beaudoin et al. 2015) (Figure 2.3). It was also shown that antisense oligonucleotides can be used to inhibit or promote the formation of RNA G4 structures (Rouleau, Beaudoin et al. 2015). Additionally, by using ribosome footprinting on a transcriptome-wide scale, Wolfe et al. (2014) reported that the 12-nucleotide guanine quartet

motifs that can form G4 structures in 5’-UTRs rendering mRNAs exceptionally sensitive to eIF4A. As a key factor in cap-dependent translation initiation, eIF4A plays a role in scanning the 5’-UTR of the mRNAs for start codons (Bhat, Robichaud et al. 2015). The natural compounds silvestrol, hippuristanol and pateamine A are implicated to have anticancer activity (Bordeleau, Robert et al. 2008; Schatz, Oricchio et al. 2011). By using silvestrol in murine T-ALL models and primary human T-ALL samples, Wolf et al. observed that eIF4A promotes the T-cell acute lymphoblastic leukemia development and maintenance by unwinding the G4 structures in oncogenes, superenhancers-associated transcription factors and epigenetic regulators including *MYC, NOTCH1, MYB, CDK6, MDM2, CCND3, ETS1,* and *BCL-2* (Wolfe, Singh et al. 2014) (Figure 2.3, 4B). It was, however, suggested that motifs other than 5’-UTR G4 structures may be required to render mRNA translation sensitive to eIF4A (Rubio, Weisburd et al. 2014). To this end, mRNAs with long, but not short 5’-UTRs, appear to exhibit eIF4A-sensitivity, thereby suggesting that the length of 5’-UTR may also determine eIF4A requirement (Sinvani, Haimov et al. 2015; Gandin, Masvidal et al. 2016).

G4 structures in 5’-UTRs also influence cap-independent, IRES-driven-translation. The IRES within the 5’-UTR of the *FGF2* mRNA forms a G4 structure affecting cap-independent translation (Bonnal, Schaeffer et al. 2003). Deletion analysis in human liver adenocarcinoma cells showed that the pG4 sequences are sufficient to facilitate IRES activity (Baird, Turcotte et al. 2006). Another example was shown by Morris et al., (2010) who reported that the *hVEGF* (human vascular endothelial growth factor) mRNA forms a G4 structure essential for

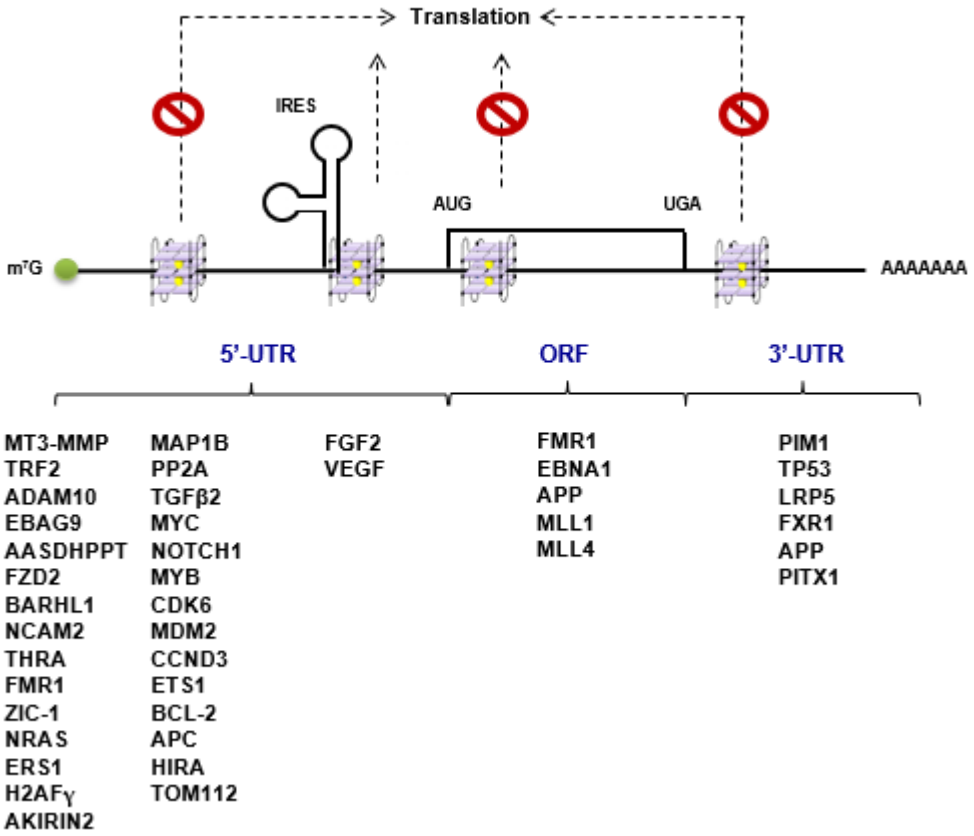


Figure 2.3 Possible roles of G-quadruplexes in mRNA translation and mRNAs that harbor G4 structures.

G-quadruplexes in 5'-UTRs, ORF and 3'-UTRs mainly repress cap-dependent translation, whereas G-quadruplexes in 5'-UTR near IRESs likely enhance the IRES-mediated translation. The genes harboring G4 structures in 5'-UTRs, ORF and 3'-UTRs are listed below.

IRES-mediated translation (Morris, Negishi et al. 2010). Interestingly, it was also shown that the stabilization of the G4 structure leads to inhibition of IRES-mediated translation of VEGF-A (Cammas, Dubrac et al. 2015). These findings show that G4 structures may influence IRES-mediated cap-independent translation, although the mechanism on how this is achieved is unclear (Jackson 2013).

2.6.2 G-quadruplexes in open reading frames and translational control

The role of G4 structures in translational control has been focused mainly on G4 sequences in the 5'-UTRs. However, ORFs also contain G4 sequences and these sequences frequently encode low complexity amino acid sequences, amino acid repeats or short motifs (Thandapani, Song et al. 2015). While G-quadruplex in 5'-UTR decreases protein expression by inhibiting ribosome scanning processes, G4 structures in ORFs likely contributes to other translation-related processes, such as elongation (Endoh, Kawasaki et al. 2013), ribosomal frameshift (Giedroc and Cornish 2009), no-go mRNA decay (Doma and Parker 2006; Harigaya and Parker 2010) and translational folding of newly synthesized proteins (Komar 2009; Zhang, Hubalewska et al. 2009; O'Brien, Vendruscolo et al. 2012). Originally, it was demonstrated that the Herpes virus thymidine kinase ORF contains pG4 sequences, leading to the expression of full-length thymidine kinase (Horsburgh, Kollmus et al. 1996). Subsequently, the ORFs containing G-quadruplexes were found in *FMR1* and *APP* (amyloid precursor protein) mRNAs (Schaeffer, Bardoni et al. 2001; Westmark and Malter 2007). Murat et al. (2014) revealed that the mRNA encoding Epstein-Barr virus-encoded nuclear antigen 1 (EBNA1) forms a G4 in its ORF and using antisense oligonucleotides to G4 sequences to destabilize G4 formation, they observed that the structure impairs translation elongation (Murat, Zhong et al. 2014) (Figure 2.3). Recently, genome-wide bioinformatic analysis showed that two mixed lineage leukemia (MLL) proto-oncogenes *KMT2A* (lysine methyltransferase 2A or mixed lineage leukemia, *MLL1)* and *KMT2B (MLL4)* harbor pG4 sequences in their ORFs ~200 nucleotides downstream of the start codon (Thandapani, Song et al. 2015) (Figure 2.3). By performing *in-line* probing analysis with G-A mutagenesis and the dual luciferase reporter assay, it was confirmed that the MLL mRNAs form *bona-fide* G4

structures that block translation by >75% (Thandapani, Song et al. 2015). Endoh & Sugimoto (2016) observed that the positioning of the G4 in ORFs, but not in 5'-UTR has a dramatic impact on translational efficiency (Endoh and Sugimoto 2016). These results suggest that by acting 'roadblocks', G-quadruplexes in ORFs may play a significant role in protein synthesis by inhibiting ribosomal progression during elongation.

2.6.3 G-quadruplexes in 3'-UTR and translational control

Polyadenylation is an essential process in post-transcriptional regulation. Following transcription, the mRNA is subsequently cleaved and a poly (A) tail added downstream of a canonical polyadenylation signal (AAUAAA) (Colgan and Manley 1997; Millevoi and Vagner 2010). G-quadruplexes are located in the 3'-UTR of some mRNAs and their presence affects translational output. The 3'-UTR of the proto-oncogene *PIM1* harbors a conserved pG4 sequence (Figure 2.3) and Arora & Suess (2011) showed using reporter assays that this pG4 sequence inhibits translation (Arora and Suess 2011). A pG4 structure in *TP53* 3'-UTR maintains the 3'-end processing under DNA damage and the G4 formation is critical for p53 protein expression contributing to p53-induced apoptosis (Decorsiere, Cayrel et al. 2011). More recently, a study combining *in silico, in vitro* and *in cellulo* approach, demonstrated that *LRP5* and *FXR1* mRNAs form G4 structures in their 3'-UTR, affecting the ratio of short/long isoforms produced (Beaudoin and Perreault 2013) (Figure 2.3). Moreover, via a bioinformatic approach and CD spectrophotometry, Crenshaw et al. (2015) identified a candidate G-quadruplex in the 3'-UTR of *APP* mRNA, whose overexpression leads to Alzheimer's disease. The authors also showed that this G-quadruplex inhibits APP protein expression by dual luciferase reporter assay (Crenshaw, Leung et al. 2015) (Figure 2.3).

In addition to translational control, G-quadruplexes in the 3'-UTR also play a role in alternative splicing, polyadenylation and mRNA targeting. For example, G4 structures in the 3'-UTR of PIM1 mRNA favor translational repression (Arora and Suess 2011), while those in the 3'-UTR of *TP53* and *hTERT* mRNAs regulate alternative splicing (Gomez, Lemarteleur et al. 2004; Marcel, Tran et al. 2011). The G-quadruplex formed in the intron 6 of hTERT mRNA acts as a splicing silencer to control the splicing efficiency, since a potent

G-quadruplex-stabilizing agent impaired the splicing machinery of hTERT (Gomez, Lemarteleur et al. 2004). Marcel et al. (2011) reported that the G-quadruplex in intron 3 of TP53 mRNA modulates the splicing of intron 2, shown by site-directed mutagenesis and fluorescent protein-reporter splicing assay (Marcel, Tran et al. 2011). Additionally, G4 structures within the *LRP5* and *FXR1* 3'-UTRs increase the efficiency of alternative polyadenylation site selection, leading to the expression of shorter transcripts (Beaudoin and Perreault 2013). Finally, G4 structures in 3'-UTR of two dendritic mRNAs *PSD95* and *CaMKIIa* facilitate their localization in neurites (Subramanian, Rage et al. 2011).

2.7 G-quadruplexes and *trans*-acting binding proteins

Many cellular RNAs associate with RNA-binding proteins to form ribonucleoprotein (RNP) complexes (Lukong, Chang et al. 2008; Calabretta and Richard 2015). A large number of RNA-binding proteins contain RGG/RG motif (Thandapani, O'Connor et al. 2013) or the related RGG/YGG motif (Castello, Fischer et al. 2012). The RGG/RG motifs of nucleolin and FMRP have been shown to associate with RNA G4 forming sequences (Hanakahi, Sun et al. 1999; Darnell, Jensen et al. 2001). Some RNA G-quadruplex binding proteins are required for the unwinding of G4 structures during translation progression (Sissi, Gatto et al. 2011). The *trans*-acting factors that bind G4 sequences and their links with diseases are discussed below.

2.7.1 Nucleolin

Nucleolin is a 100kDa nucleolar phosphoprotein that contains an RGG/RG motif (Hanakahi, Sun et al. 1999). Nucleolin interacts with both DNA and RNA G-rich sequences and plays a role in translational repression of specific mRNAs (Abdelmohsen, Tominaga et al. 2011; Abdelmohsen and Gorospe 2012; Thandapani, O'Connor et al. 2013). *C9orf72*, a disease-related gene, harbors the expansion of a $(GGGGCC)_n(GGCCCC)_n$ repeats within the first intron, leading to amyotrophic lateral sclerosis and frontotemporal dementia (DeJesus-Hernandez, Mackenzie et al. 2011; Turner, Hardiman et al. 2013). It has been

reported that the DNA/RNA hexanucleotide repeat expansion (HRE), $(GGGGCC)_n$, forms extremely stable G-quadruplex structures and this may play a role in *C9orf72* gene activity, protein binding as well as translation into pathologic dipeptides (Reddy, Zamiri et al. 2013; Haeusler, Donnelly et al. 2014). It was also elucidated that the DNA of *C9orf72* HRE forms antiparallel- and parallel-G-quadruplexes, while the HRE RNA sequence only adopts the parallel-stranded G4s. Nucleolin preferentially binds the HRE G-quadruplexes, which is mislocalized in patient cells carrying the *C9orf72* mutation (Haeusler, Donnelly et al. 2014). These studies show DNA/RNA G4 sequences bound by nucleolin as a determinants of repeat-associated neurodegenerative diseases.

2.7.2 FMRP

Fragile X syndrome (FXS) is caused by expansion of $(CGG)_n$ repeats in the *FMR1* gene, and subsequent hypermethylation of the *FMR1* gene promoter, leading to the loss of FMRP expression (Darnell and Klann 2013; Colak, Zaninovic et al. 2014). FXS is recognized to be the most-frequent heritable syndrome of mental insufficiency (Richter, Bassell et al. 2015). This implicates the role of RNA G-quadruplexes in neuronal function (Richter, Bassell et al. 2015). In addition to binding to the G4s, the RGG/RG motif of FMRP also modulates its association with polysomes (Blackwell, Zhang et al. 2010), being consistent with its role in translational regulation (Corbin, Bouillon et al. 1997; Richter, Bassell et al. 2015). Originally, it was demonstrated that 432 mRNAs were co-immunoprecipitated with the FMRP RNP complex from mouse brain, nearly 70% of which contain a G4 sequence. Dramatic changes in polysome-association of these mRNAs were observed in the absence of FMRP, indicating that translational dysregulation of these mRNAs may underpin FXS (Brown, Jin et al. 2001). Recently, this was redefined by large-scale CLIP (crosslinking immunoprecipitation) studies, where 34 218 consensus FMRP binding sites in 3 703 genes were identified (Anderson, Chopra et al. 2016). It was reported the identification of a G4 sequence in the FMR1 mRNA that provides the docking site for its encoded fragile X mental retardation protein (FMRP). Luciferase reporter assay demonstrated that the binding of FMRP to a G4 sequence in 5’-UTR of a reporter gene strongly represses its translation initiation in vitro (Schaeffer,

Bardoni et al. 2001). Intriguingly, the biophysical characterization was carried out directly showing the binding between the FMR1 G4 structures and several FMRP isoforms, which regulates different FMRP isoform production involved in a feedback mechanism (Blice-Baum and Mihailescu 2014). *MAP1B*, *PP2A* and *Shank1* genes that are essential for neural development have been shown to harbor one or more G-quadruplexes in their 5'-UTRs and/or 3'-UTRs, whereby the absence of FMRP increases production of the corresponding proteins, indicating that FMRP plays pivotal roles in neonatal brain development (Lu, Wang et al. 2004; Castets, Schaeffer et al. 2005; Zhang, Gaetano et al. 2014) (Figure 2.3). Furthermore, to explore the functions of FMRP in translational control in neurons, Napoli et al. (2008) found that FMRP inhibits cap-dependent translation by recruiting and/or stabilizing CYFIP1 in synaptoneurosomes. As a novel eIF4E binding protein, CYFIP1 binds eIF4E independently of other factors to impede cap-dependent translation of many mRNAs (Figure 2.4A). FMRP is dephosphorylated in response to synaptic activation, resulting in the release from mRNAs encoding synaptic proteins and their derepression (Napoli, Mercaldo et al. 2008; Simone, Fratta et al. 2015). In contrast, Bechara et al. (2009) reported a novel role of FMRP in translation activation. They observed that FMRP binds *SOD1* mRNA through a motif called SoSLIP (*SOD1* mRNA Stem Loops Interacting with FMRP), which competes with FMRP binding to G4 structures favoring translation (Bechara, Didiot et al. 2009). The solution structure of FMRP in complex with the *in vitro* selected G-rich RNA sequence, sc1 RNA, revealed that arginines within the RGG/RG motif of FMRP are positioned in the major groove of the G4 structure (Phan, Kuryavyi et al. 2011). Recently, X-ray crystallography analysis demonstrated that an RGG peptide of human FMRP was able to bind to the selected G-rich RNA *in vitro* (Vasilyev, Polonskaia et al. 2015). The RGG peptide was shown to stabilize G-quartets and facilitate G4 formation. It was also revealed that the specific binding of these RNAs with FMRP likely also involves the hydrogen binding with RNA duplexes, shown by mutagenesis and footprinting (Vasilyev, Polonskaia et al. 2015). High throughput sequencing of RNAs collected by HITS-CLIP (crosslinking immunoprecipitation) identified that FMRP also interacts with ORFs and stalls ribosomes on mRNA encoding presynaptic and postsynaptic proteins implicated in autism spectrum disorders (ASD) (Figure 2.4A)

(Darnell, Van Driesche et al. 2011). Moreover, FMRP was shown to repress translation by direct binding to the L5 protein on the 80S ribosome (Chen, Sharma et al. 2014), and this activity is depended on the integrity of the RGG/RG motif (Taha, Nouri et al. 2014).

2.7.3 RHAU (DHX36)

RHAU, the RNA helicase associated with AU-rich elements, also known as DHX36, is a member of ATP-dependent DEAH-box RNA helicase family (Vaughn, Creacy et al. 2005; Booy, McRae et al. 2015). DHX36 has been shown to be the predominant G-quadruplex resolving helicase *in celullo*, with high activity in unwinding RNA G4 structures (Chalupnikova, Lattmann et al. 2008; Creacy, Routh et al. 2008; Booy, Meier et al. 2012). The solution structure of DHX36 recognizing a G-quadruplex was resolved and identified a three-anchor-point electrostatic interaction (Heddi, Cheong et al. 2015). It was also elucidated that DHX36 uses a local, non-processive mechanism to unwind G4 structures, mimicking the DEAD-box RNA helicase eIF4A (Chen, Murat et al. 2015). By applying genome-wide analysis, DXH36 was shown to bind ~106 RNAs, the majority harboring pG4 sequences including the human telomerase RNA (Lattmann, Giri et al. 2010; Lattmann, Stadler et al. 2011). Sexton et al. (2011) also discovered the association between DHX36 and *TERC* pG4 sequences in the HEK293 cells (Sexton and Collins 2011). Another study confirmed that the interaction between DHX36 and the *TERC* G-quadruplex disrupts the formation of P1 helix, a structure which defines template boundary for reverse transcription. DHX36 was sufficient to unwind the quadruplexes and promote the formation of a stable P1 helix, and DHX36 depletion led to a reduction in average telomere length (Booy, Meier et al. 2012; Booy, McRae et al. 2015). Besides the function on telomere maintenance, DHX36 regulates translation, although the precise mechanism has not yet been elucidated. By performing a RNA-co-immunoprecipitation screen, Booy et al. (2014) discovered that DHX36 interacts with G-quadruplexes in 3'-UTR of *PITX1* mRNA *in cellulo*, and depletion of DHX36 results in increased PITX1 expression, a transcription factor with roles in development and cancer (Booy, Howard et al. 2014). Finally, it was also demonstrated that the DHX36-mediated regulation of PITX1 involves miRNA regulatory components (Booy, Howard et al. 2014).

Moreover, the impact of Aven on mRNA translation regulation is dependent on DHX36 (see below).

2.7.4 Aven

Aven is an anti-apoptotic protein first shown to interact and stabilize the pro-survival factor Bcl-xl, while inhibiting Apaf-1, thereby stimulating survival in particular when the cells are exposed to stress (Chau, Cheng et al. 2000). Aven is overexpressed in acute myeloid and acute lymphoblastic leukemia where its expression has been associated with poor prognosis (Paydas, Tanriverdi et al. 2003; Choi, Hwang et al. 2006). Recently, Aven was shown to bind RNAs with G4 structures via its RGG/RG motif (Thandapani, Song et al. 2015). The RGG/RG motif of Aven is arginine methylated by PRMT1 and this promotes association with the methylarginine interactors SMN and TDRD3. The arginine methylation and binding to SMN and TDRD3 is required for association of Aven with polysomes. *In vitro* binding and photocrosslinking immunoprecipitation assays revealed that Aven associates with *KMT2A* and *KMT2B* mRNAs. Interestingly, the G4 sequences bound by Aven were in the coding regions regulating their mRNA translation (Figure 2.4C). Mechanistically, Aven recruits DHX36 onto the polysomes likely to facilitate unwinding of the G4 structures during translation. Depletion of Aven/DHX36 inhibits *KMT2A/KMT2B* translation, whereas luciferase assays with G-to-A mutagenesis revealed that both Aven and DHX36 are required to rescue translation in the presence of G4 motifs (Thandapani, Song et al. 2015). Thus, Aven represents a G4 interacting protein that recruits DHX36 to unwind G4 structures containing mRNAs during translational elongation. These findings foster our understanding that RNA G4 binding proteins can play a key role in modulating mRNA translation.

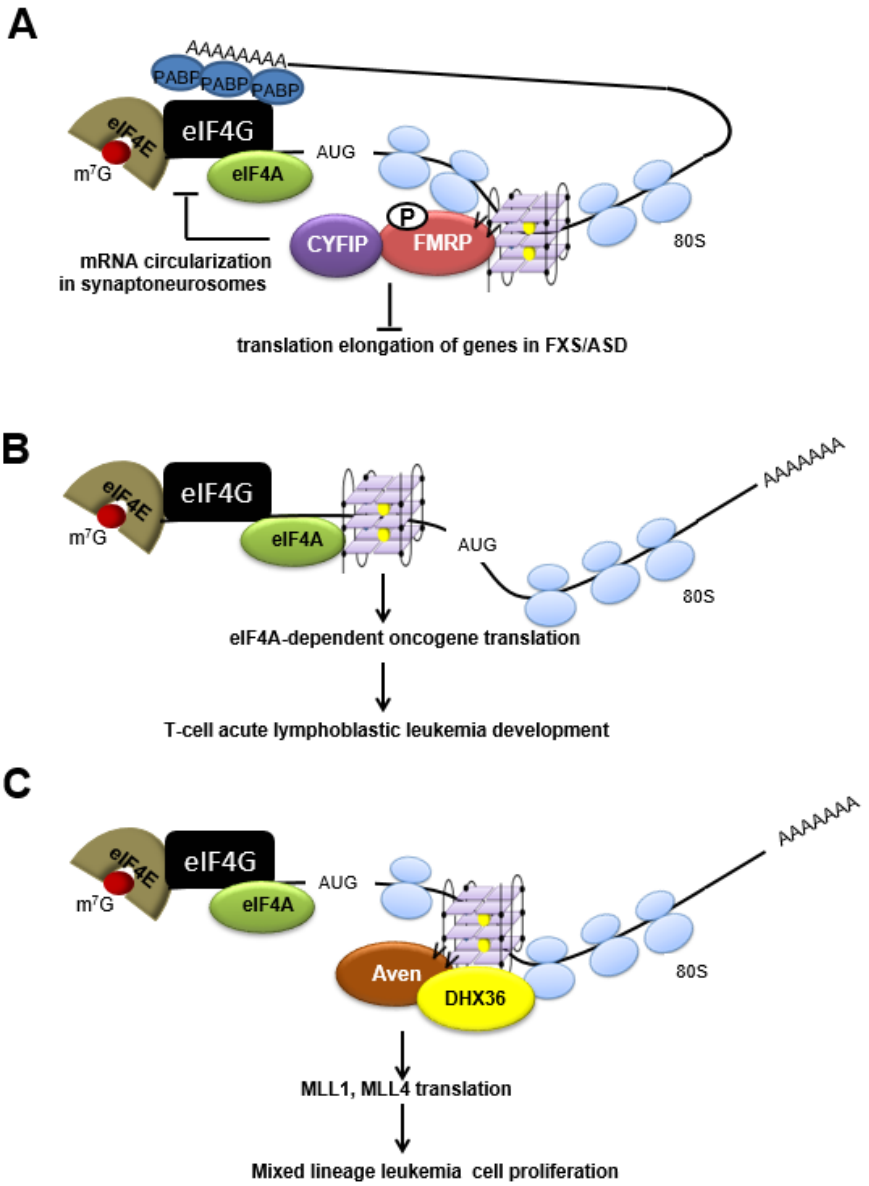


Figure 2.4 Schematic illustration of the functions of RNA-binding proteins that bind RNA G4 structures in mRNAs.

A) Phosphorylated FMRP binds ORFs of mRNAs and inhibits translation. It stalls ribosomes in the elongation stage, resulting in the repressed translation of transcripts related to FXS/ASD. It recruits the co-factor CYFIP in synaptoneurosomes. By interacting with CYFIP, FMRP prevents the complex assembly between eIF4E, eIF4G and PABP, thereby inhibiting translation initiation. RGG/RG motif is denoted as ⋎⋎.

B) eIF4A unwinds the G4 structures in 5'-UTR of many key transcription factors and oncogenes, thereby contributing to the T cell-acute lymphoblastic leukemia development.

C) Methylated Aven binds G4 structures in the ORFs of MLL1 and MLL4 mRNAs in an RGG/RG motif (denoted as ⋎⋎) dependent manner. Aven also recruits DHX36 onto the polysomes that may facilitate unwinding of G4 structures. Thus Aven favors the translation of oncogenic proteins to increase leukemic cell proliferation.

2.8 Conclusions and perspectives

Increasing evidence shows that G-quadruplexes play essential roles in RNA metabolism. Advances in computational analyses and genome-wide sequencing have facilitated the characterization of RNA quadruplexes and their implication in translational control. Further advancements are required to confirm their occurrence *in vivo*. Moreover, the *trans*-acting factors that bind G4 structures are key to understanding the role of G-quadruplexes in the regulation of mRNA fate. Nucleolin, FMRP, DHX36 and Aven represent the tip of the iceberg of *trans*-acting factors that recognize G4 RNA structures. Large scale analyses including HITS-CLIP-seq, SHAPE-seq, and the high-throughput RNA affinity profiling (HiTS-RAP) hold great promise for knowledge advancement as to the role of G4 structures and interacting proteins in *vivo* (Deigan, Li et al. 2009; Xu, Bolduc et al. 2012; Moore, Zhang et al. 2014; Siegfried, Busan et al. 2014; Tome, Ozer et al. 2014). It is possible that multiple RNA-binding proteins recognize one G-quadruplex, and there could be functional interplay. The binding of different RNA-binding proteins are likely to: 1) stabilize the G4 structures, 2) unwind the G4s by recruiting helicases, 3) act as chaperones to transport G4 containing mRNAs, and 4) serve as scaffold proteins to recruit other proteins or RNAs.

A growing number of disease-related genes are regulated by G4 structures and their RNA-binding proteins. This includes mRNAs encoding the tumor suppressor *TP53* (Decorsiere, Cayrel et al. 2011), oncogene *NRAS* (Kumari, Bugaut et al. 2007; Biffi, Di Antonio et al. 2014), oncogenes *KMT2A/KMT2B* (Thandapani, Song et al. 2015), anti-apoptotic *Bcl-2* (Shahid, Bugaut et al. 2010; Wolfe, Singh et al. 2014), *FMR1* (Didiot, Tian et al. 2008; Blice-Baum and Mihailescu 2014), and the telomerase *hTERT* (Gomez, Lemarteleur et al. 2004; Bidzinska, Cimino-Reale et al. 2013). This indicates that by affecting mRNA translation, splicing, and polyadenylation, G4 structures may be implicated in numerous human diseases, especially cancer. Thus, targeting G4 structures with synthesized small molecules is attractive to modify oncogene expression. Extensive studies have evaluated some ‘drug-like’ molecules including the pyridostatins (PDS) (Bugaut, Rodriguez et al. 2010), cationic porphyrins and derivatives (Faudale, Cogoi et al. 2012; Morris, Wingate et al. 2012; Huang, Zhu et al. 2014) and bisquinolinium compounds (Gomez, Guedin et al.

2010). Short antisense oligonucleotides affect RNA G4 folding and translation regulation of specific mRNAs (Rouleau, Beaudoin et al. 2015) and may eventually be of therapeutic potential. A new polyaromatic molecule, RGB-1, was recently shown to specifically stabilize RNA G-quadruplexes, but not DNA G-quadruplexes or other RNA structures. RGB-1 inhibits translation in mammalian cells and decreases NRAS expression, providing a new tool to understand G4 structures and therapeutic applications (Katsuda, Sato et al. 2016). Thus, a better understanding of the various contributions of G4 structures in the human transcriptome may provide important insights into strategies targeting G4s in molecular medicine.

2.9 References

Abdelmohsen, K. and M. Gorospe (2012). "RNA-binding protein nucleolin in disease." RNA Biol 9(6): 799-808.

Abdelmohsen, K., K. Tominaga, et al. (2011). "Enhanced translation by Nucleolin via G-rich elements in coding and non-coding regions of target mRNAs." Nucleic Acids Res 39(19): 8513-8530.

Agarwala, P., S. Pandey, et al. (2013). "The G-quadruplex augments translation in the 5' untranslated region of transforming growth factor beta2." Biochemistry 52(9): 1528-1538.

Anderson, B. R., P. Chopra, et al. (2016). "Identification of consensus binding sites clarifies FMRP binding determinants." Nucleic Acids Res.

Arora, A., M. Dutkiewicz, et al. (2008). "Inhibition of translation in living eukaryotic cells by an RNA G-quadruplex motif." RNA 14(7): 1290-1296.

Arora, A. and B. Suess (2011). "An RNA G-quadruplex in the 3' UTR of the proto-oncogene PIM1 represses translation." RNA Biol 8(5): 802-805.

Baird, S. D., M. Turcotte, et al. (2006). "Searching for IRES." RNA 12(10): 1755-1785.

Balasubramanian, S., L. H. Hurley, et al. (2011). "Targeting G-quadruplexes in gene promoters: a novel anticancer strategy?" Nat Rev Drug Discov 10(4): 261-275.

Balkwill, G. D., K. Derecka, et al. (2009). "Repression of translation of human estrogen receptor alpha by G-quadruplex formation." Biochemistry 48(48): 11487-11495.

Bang, I. (1910). "Untersuchungen u¨ ber die Guanylsau¨ re.
Biochemische Zeitschrift." 26: 293-311.

Beaudoin, J. D., R. Jodoin, et al. (2014). "New scoring system to identify RNA G-quadruplex folding." Nucleic Acids Res 42(2): 1209-1223.

Beaudoin, J. D. and J. P. Perreault (2010). "5'-UTR G-quadruplex structures acting as translational repressors." Nucleic Acids Res 38(20): 7022-7036.

Beaudoin, J. D. and J. P. Perreault (2013). "Exploring mRNA 3'-UTR G-quadruplexes: evidence of roles in both alternative polyadenylation and mRNA shortening." Nucleic Acids Res 41(11): 5898-5911.

Bechara, E. G., M. C. Didiot, et al. (2009). "A novel function for fragile X mental retardation protein in translational activation." PLoS Biol 7(1): e16.

Besnard, E., A. Babled, et al. (2012). "Unraveling cell type-specific and reprogrammable human replication origin signatures associated with G-quadruplex consensus motifs." Nat Struct Mol Biol 19(8): 837-844.

Bhat, M., N. Robichaud, et al. (2015). "Targeting the translation machinery in cancer." Nat Rev Drug Discov 14(4): 261-278.

Bidzinska, J., G. Cimino-Reale, et al. (2013). "G-quadruplex structures in the human genome as novel therapeutic targets." Molecules 18(10): 12368-12395.

Biffi, G., M. Di Antonio, et al. (2014). "Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells." Nat Chem 6(1): 75-80.

Blackwell, E., X. Zhang, et al. (2010). "Arginines of the RGG box regulate FMRP association with polyribosomes and mRNA." Hum Mol Genet 19(7): 1314-1323.

Blice-Baum, A. C. and M. R. Mihailescu (2014). "Biophysical characterization of

G-quadruplex forming FMR1 mRNA and of its interactions with different fragile X mental retardation protein isoforms." RNA 20(1): 103-114.
Bochman, M. L., K. Paeschke, et al. (2012). "DNA secondary structures: stability and function of G-quadruplex structures." Nat Rev Genet 13(11): 770-780.
Bonnal, S., C. Schaeffer, et al. (2003). "A single internal ribosome entry site containing a G quartet RNA structure drives fibroblast growth factor 2 gene expression at four alternative translation initiation codons." J Biol Chem 278(41): 39330-39336.
Booy, E. P., R. Howard, et al. (2014). "The RNA helicase RHAU (DHX36) suppresses expression of the transcription factor PITX1." Nucleic Acids Res 42(5): 3346-3361.
Booy, E. P., E. K. McRae, et al. (2015). "Biochemical characterization of G4 quadruplex telomerase RNA unwinding by the RNA helicase RHAU." Methods Mol Biol 1259: 125-135.
Booy, E. P., M. Meier, et al. (2012). "The RNA helicase RHAU (DHX36) unwinds a G4-quadruplex in human telomerase RNA and promotes the formation of the P1 helix template boundary." Nucleic Acids Res 40(9): 4110-4124.
Brooks, T. A., S. Kendrick, et al. (2010). "Making sense of G-quadruplex and i-motif functions in oncogene promoters." FEBS J 277(17): 3459-3469.
Brown, V., P. Jin, et al. (2001). "Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome." Cell 107(4): 477-487.
Bugaut, A. and S. Balasubramanian (2012). "5'-UTR RNA G-quadruplexes: translation regulation and targeting." Nucleic Acids Res 40(11): 4727-4741.
Bugaut, A., R. Rodriguez, et al. (2010). "Small molecule-mediated inhibition of translation by targeting a native RNA G-quadruplex." Org Biomol Chem 8(12): 2771-2776.
Calabretta, S. and S. Richard (2015). "Emerging Roles of Disordered Sequences in RNA-Binding Proteins." Trends Biochem Sci 40(11): 662-672.
Cammas, A., A. Dubrac, et al. (2015). "Stabilization of the G-quadruplex at the VEGF IRES represses cap-independent translation." RNA Biol 12(3): 320-329.
Castello, A., B. Fischer, et al. (2012). "Insights into RNA biology from an atlas of mammalian mRNA-binding proteins." Cell 149(6): 1393-1406.
Castets, M., C. Schaeffer, et al. (2005). "FMRP interferes with the Rac1 pathway and controls actin cytoskeleton dynamics in murine fibroblasts." Hum Mol Genet 14(6): 835-844.
Cayrou, C., P. Coulombe, et al. (2012). "New insights into replication origin characteristics in metazoans." Cell Cycle 11(4): 658-667.
Cayrou, C., P. Coulombe, et al. (2011). "Genome-scale analysis of metazoan replication origins reveals their organization in specific but flexible sites defined by conserved features." Genome Res 21(9): 1438-1449.
Chakraborty, P. and F. Grosse (2011). "Human DHX9 helicase preferentially unwinds RNA-containing displacement loops (R-loops) and G-quadruplexes." DNA Repair (Amst) 10(6): 654-665.
Chalupnikova, K., S. Lattmann, et al. (2008). "Recruitment of the RNA helicase RHAU to stress granules via a unique RNA-binding domain." J Biol Chem 283(50): 35186-35198.
Chambers, V. S., G. Marsico, et al. (2015). "High-throughput sequencing of DNA G-quadruplex structures in the human genome." Nat Biotechnol 33(8): 877-881.
Chau, B. N., E. H. Cheng, et al. (2000). "Aven, a novel inhibitor of caspase activation, binds Bcl-xL and Apaf-1." Mol Cell 6(1): 31-40.

Chen, E., M. R. Sharma, et al. (2014). "Fragile X mental retardation protein regulates translation by binding directly to the ribosome." Mol Cell 54(3): 407-417.
Chen, M. C., P. Murat, et al. (2015). "Insights into the mechanism of a G-quadruplex-unwinding DEAH-box helicase." Nucleic Acids Res 43(4): 2223-2231.
Choi, J., Y. K. Hwang, et al. (2006). "Aven overexpression: association with poor prognosis in childhood acute lymphoblastic leukemia." Leuk Res 30(8): 1019-1025.
Colak, D., N. Zaninovic, et al. (2014). "Promoter-bound trinucleotide repeat mRNA drives epigenetic silencing in fragile X syndrome." Science 343(6174): 1002-1005.
Colgan, D. F. and J. L. Manley (1997). "Mechanism and regulation of mRNA polyadenylation." Genes Dev 11(21): 2755-2766.
Corbin, F., M. Bouillon, et al. (1997). "The fragile X mental retardation protein is associated with poly(A)+ mRNA in actively translating polyribosomes." Hum Mol Genet 6(9): 1465-1472.
Creacy, S. D., E. D. Routh, et al. (2008). "G4 resolvase 1 binds both DNA and RNA tetramolecular quadruplex with high affinity and is the major source of tetramolecular quadruplex G4-DNA and G4-RNA resolving activity in HeLa cell lysates." J Biol Chem 283(50): 34626-34634.
Crees, Z., J. Girard, et al. (2014). "Oligonucleotides and G-quadruplex stabilizers: targeting telomeres and telomerase in cancer therapy." Curr Pharm Des 20(41): 6422-6437.
Crenshaw, E., B. P. Leung, et al. (2015). "Amyloid Precursor Protein Translation Is Regulated by a 3'UTR Guanine Quadruplex." PLoS One 10(11): e0143160.
Dai, J., Z. Q. Liu, et al. (2015). "Discovery of Small Molecules for Up-Regulating the Translation of Antiamyloidogenic Secretase, a Disintegrin and Metalloproteinase 10 (ADAM10), by Binding to the G-Quadruplex-Forming Sequence in the 5' Untranslated Region (UTR) of Its mRNA." J Med Chem 58(9): 3875-3891.
Darnell, J. C., K. B. Jensen, et al. (2001). "Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function." Cell 107(4): 489-499.
Darnell, J. C. and E. Klann (2013). "The translation of translational control by FMRP: therapeutic targets for FXS." Nat Neurosci 16(11): 1530-1536.
Darnell, J. C., S. J. Van Driesche, et al. (2011). "FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism." Cell 146(2): 247-261.
Decorsiere, A., A. Cayrel, et al. (2011). "Essential role for the interaction between hnRNP H/F and a G quadruplex in maintaining p53 pre-mRNA 3'-end processing and function during DNA damage." Genes Dev 25(3): 220-225.
Deigan, K. E., T. W. Li, et al. (2009). "Accurate SHAPE-directed RNA structure determination." Proc Natl Acad Sci U S A 106(1): 97-102.
DeJesus-Hernandez, M., I. R. Mackenzie, et al. (2011). "Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS." Neuron 72(2): 245-256.
Didiot, M. C., Z. Tian, et al. (2008). "The G-quartet containing FMRP binding site in FMR1 mRNA is a potent exonic splicing enhancer." Nucleic Acids Res 36(15): 4902-4912.
Doma, M. K. and R. Parker (2006). "Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation." Nature 440(7083): 561-564.
Endoh, T., Y. Kawasaki, et al. (2013). "Translational halt during elongation caused by

G-quadruplex formed by mRNA." Methods 64(1): 73-78.
Endoh, T. and N. Sugimoto (2016). "Mechanical insights into ribosomal progression overcoming RNA G-quadruplex from periodical translation suppression in cells." Sci Rep 6: 22719.
Faudale, M., S. Cogoi, et al. (2012). "Photoactivated cationic alkyl-substituted porphyrin binding to g4-RNA in the 5'-UTR of KRAS oncogene represses translation." Chem Commun (Camb) 48(6): 874-876.
Foulk, M. S., J. M. Urban, et al. (2015). "Characterizing and controlling intrinsic biases of lambda exonuclease in nascent strand sequencing reveals phasing between nucleosomes and G-quadruplex motifs around a subset of human replication origins." Genome Res 25(5): 725-735.
Frees, S., C. Menendez, et al. (2014). "QGRS-Conserve: a computational method for discovering evolutionarily conserved G-quadruplex motifs." Hum Genomics 8: 8.
Gandin, V., L. Masvidal, et al. (2016). "nanoCAGE reveals 5' UTR features that define specific modes of translation of functionally related MTOR-sensitive mRNAs." Genome Res 26(5): 636-648.
Gellert, M., M. N. Lipsett, et al. (1962). "Helix formation by guanylic acid." Proc Natl Acad Sci U S A 48: 2013-2018.
Giedroc, D. P. and P. V. Cornish (2009). "Frameshifting RNA pseudoknots: structure and mechanism." Virus Res 139(2): 193-208.
Gomez, D., A. Guedin, et al. (2010). "A G-quadruplex structure within the 5'-UTR of TRF2 mRNA represses translation in human cells." Nucleic Acids Res 38(20): 7187-7198.
Gomez, D., T. Lemarteleur, et al. (2004). "Telomerase downregulation induced by the G-quadruplex ligand 12459 in A549 cells is mediated by hTERT RNA alternative splicing." Nucleic Acids Res 32(1): 371-379.
Gros, J., A. Guedin, et al. (2008). "G-Quadruplex formation interferes with P1 helix formation in the RNA component of telomerase hTERC." Chembiochem 9(13): 2075-2079.
Haeusler, A. R., C. J. Donnelly, et al. (2014). "C9orf72 nucleotide repeat structures initiate molecular cascades of disease." Nature 507(7491): 195-200.
Hanakahi, L. A., H. Sun, et al. (1999). "High affinity interactions of nucleolin with G-G-paired rDNA." J Biol Chem 274(22): 15908-15912.
Harigaya, Y. and R. Parker (2010). "No-go decay: a quality control mechanism for RNA in translation." Wiley Interdiscip Rev RNA 1(1): 132-141.
Healy, K. C. (1995). "Telomere dynamics and telomerase activation in tumor progression: prospects for prognosis and therapy." Oncol Res 7(3-4): 121-130.
Heddi, B., V. V. Cheong, et al. (2015). "Insights into G-quadruplex specific recognition by the DEAH-box helicase RHAU: Solution structure of a peptide-quadruplex complex." Proc Natl Acad Sci U S A 112(31): 9608-9613.
Hellen, C. U. and P. Sarnow (2001). "Internal ribosome entry sites in eukaryotic mRNA molecules." Genes Dev 15(13): 1593-1612.
Hershey, J. W., N. Sonenberg, et al. (2012). "Principles of translational control: an overview." Cold Spring Harb Perspect Biol 4(12).
Hinnebusch, A. G. (2014). "The scanning mechanism of eukaryotic translation initiation." Annu Rev Biochem 83: 779-812.

Hinnebusch, A. G., I. P. Ivanov, et al. (2016). "Translational control by 5'-untranslated regions of eukaryotic mRNAs." Science 352(6292): 1413-1416.
Holcik, M. and N. Sonenberg (2005). "Translational control in stress and apoptosis." Nat Rev Mol Cell Biol 6(4): 318-327.
Horsburgh, B. C., H. Kollmus, et al. (1996). "Translational recoding induced by G-rich mRNA sequences that form unusual structures." Cell 86(6): 949-959.
Huang, X. X., L. N. Zhu, et al. (2014). "Two cationic porphyrin isomers showing different multimeric G-quadruplex recognition specificity against monomeric G-quadruplexes." Nucleic Acids Res 42(13): 8719-8731.
Huppert, J. L. and S. Balasubramanian (2005). "Prevalence of quadruplexes in the human genome." Nucleic Acids Res 33(9): 2908-2916.
Huppert, J. L., A. Bugaut, et al. (2008). "G-quadruplexes: the beginning and end of UTRs." Nucleic Acids Res 36(19): 6260-6268.
Jackson, R. J. (2013). "The current status of vertebrate cellular mRNA IRESs." Cold Spring Harb Perspect Biol 5(2).
Jackson, R. J., C. U. Hellen, et al. (2010). "The mechanism of eukaryotic translation initiation and principles of its regulation." Nat Rev Mol Cell Biol 11(2): 113-127.
Jodoin, R., L. Bauer, et al. (2014). "The folding of 5'-UTR human G-quadruplexes possessing a long central loop." RNA 20(7): 1129-1141.
Katsuda, Y., S. Sato, et al. (2016). "A Small Molecule That Represses Translation of G-Quadruplex-Containing mRNA." J Am Chem Soc 138(29): 9037-9040.
Komar, A. A. (2009). "A pause for thought along the co-translational folding pathway." Trends Biochem Sci 34(1): 16-24.
Komar, A. A. and M. Hatzoglou (2011). "Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states." Cell Cycle 10(2): 229-240.
Konig, S. L., A. C. Evans, et al. (2010). "Seven essential questions on G-quadruplexes." Biomol Concepts 1(2): 197-213.
Koromilas, A. E., A. Lazaris-Karatzas, et al. (1992). "mRNAs containing extensive secondary structure in their 5' non-coding region translate efficiently in cells overexpressing initiation factor eIF-4E." EMBO J 11(11): 4153-4158.
Kumari, S., A. Bugaut, et al. (2007). "An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation." Nat Chem Biol 3(4): 218-221.
Lam, E. Y., D. Beraldi, et al. (2013). "G-quadruplex structures are stable and detectable in human genomic DNA." Nat Commun 4: 1796.
Lammich, S., F. Kamp, et al. (2011). "Translational repression of the disintegrin and metalloprotease ADAM10 by a stable G-quadruplex secondary structure in its 5'-untranslated region." J Biol Chem 286(52): 45063-45072.
Lattmann, S., B. Giri, et al. (2010). "Role of the amino terminal RHAU-specific motif in the recognition and resolution of guanine quadruplex-RNA by the DEAH-box RNA helicase RHAU." Nucleic Acids Res 38(18): 6219-6233.
Lattmann, S., M. B. Stadler, et al. (2011). "The DEAH-box RNA helicase RHAU binds an intramolecular RNA G-quadruplex in TERC and associates with telomerase holoenzyme." Nucleic Acids Res 39(21): 9390-9404.
Lipps, H. J., W. Gruissem, et al. (1982). "Higher order DNA structure in macronuclear

chromatin of the hypotrichous ciliate Oxytricha nova." Proc Natl Acad Sci U S A 79(8): 2495-2499.
Lu, R., H. Wang, et al. (2004). "The fragile X protein controls microtubule-associated protein 1B translation and microtubule stability in brain neuron development." Proc Natl Acad Sci U S A 101(42): 15201-15206.
Lukong, K. E., K. W. Chang, et al. (2008). "RNA-binding proteins in human genetic disease." Trends Genet 24(8): 416-425.
Maizels, N. and L. T. Gray (2013). "The G4 genome." PLoS Genet 9(4): e1003468.
Marcel, V., P. L. Tran, et al. (2011). "G-quadruplex structures in TP53 intron 3: role in alternative splicing and in production of p53 mRNA isoforms." Carcinogenesis 32(3): 271-278.
Martadinata, H. and A. T. Phan (2013). "Structure of human telomeric RNA (TERRA): stacking of two G-quadruplex blocks in K(+) solution." Biochemistry 52(13): 2176-2183.
Martadinata, H. and A. T. Phan (2014). "Formation of a stacked dimeric G-quadruplex containing bulges by the 5'-terminal region of human telomerase RNA (hTERC)." Biochemistry 53(10): 1595-1600.
Mendoza, O., A. Bourdoncle, et al. (2016). "G-quadruplexes and helicases." Nucleic Acids Res 44(5): 1989-2006.
Millevoi, S., H. Moine, et al. (2012). "G-quadruplexes in RNA biology." Wiley Interdiscip Rev RNA 3(4): 495-507.
Millevoi, S. and S. Vagner (2010). "Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation." Nucleic Acids Res 38(9): 2757-2774.
Moore, M. J., C. Zhang, et al. (2014). "Mapping Argonaute and conventional RNA-binding protein interactions with RNA at single-nucleotide resolution using HITS-CLIP and CIMS analysis." Nat Protoc 9(2): 263-293.
Morgan, R. K., H. Batra, et al. (2016). "Identification and characterization of a new G-quadruplex forming region within the kRAS promoter as a transcriptional regulator." Biochim Biophys Acta 1859(2): 235-245.
Morris, M. J. and S. Basu (2009). "An unusually stable G-quadruplex within the 5'-UTR of the MT3 matrix metalloproteinase mRNA represses translation in eukaryotic cells." Biochemistry 48(23): 5313-5319.
Morris, M. J., Y. Negishi, et al. (2010). "An RNA G-quadruplex is essential for cap-independent translation initiation in human VEGF IRES." J Am Chem Soc 132(50): 17831-17839.
Morris, M. J., K. L. Wingate, et al. (2012). "The porphyrin TmPyP4 unfolds the extremely stable G-quadruplex in MT3-MMP mRNA and alleviates its repressive effect to enhance translation in eukaryotic cells." Nucleic Acids Res 40(9): 4137-4145.
Moye, A. L., K. C. Porter, et al. (2015). "Telomeric G-quadruplexes are a substrate and site of localization for human telomerase." Nat Commun 6: 7643.
Mukundan, V. T. and A. T. Phan (2013). "Bulges in G-quadruplexes: broadening the definition of G-quadruplex-forming sequences." J Am Chem Soc 135(13): 5017-5028.
Murat, P. and S. Balasubramanian (2014). "Existence and consequences of G-quadruplex structures in DNA." Curr Opin Genet Dev 25: 22-29.
Murat, P., J. Zhong, et al. (2014). "G-quadruplexes regulate Epstein-Barr virus-encoded

nuclear antigen 1 mRNA translation." Nat Chem Biol 10(5): 358-364.
Napoli, I., V. Mercaldo, et al. (2008). "The fragile X syndrome protein represses activity-dependent translation through CYFIP1, a new 4E-BP." Cell 134(6): 1042-1054.
O'Brien, E. P., M. Vendruscolo, et al. (2012). "Prediction of variable translation rate effects on cotranslational protein folding." Nat Commun 3: 868.
Onel, B., M. Carver, et al. (2016). "A New G-Quadruplex with Hairpin Loop Immediately Upstream of the Human BCL2 P1 Promoter Modulates Transcription." J Am Chem Soc 138(8): 2563-2570.
Paydas, S., K. Tanriverdi, et al. (2003). "Survivin and aven: two distinct antiapoptotic signals in acute leukemias." Ann Oncol 14(7): 1045-1050.
Pelletier, J. and N. Sonenberg (1985). "Insertion mutagenesis to increase secondary structure within the 5' noncoding region of a eukaryotic mRNA reduces translational efficiency." Cell 40(3): 515-526.
Phan, A. T., V. Kuryavyi, et al. (2011). "Structure-function studies of FMRP RGG peptide recognition of an RNA duplex-quadruplex junction." Nat Struct Mol Biol 18(7): 796-804.
Reddy, K., B. Zamiri, et al. (2013). "The disease-associated r(GGGGCC)n repeat from the C9orf72 gene forms tract length-dependent uni- and multimolecular RNA G-quadruplex structures." J Biol Chem 288(14): 9860-9866.
Rhodes, D. and H. J. Lipps (2015). "G-quadruplexes and their regulatory roles in biology." Nucleic Acids Res 43(18): 8627-8637.
Richter, J. D., G. J. Bassell, et al. (2015). "Dysregulation and restoration of translational homeostasis in fragile X syndrome." Nat Rev Neurosci 16(10): 595-605.
Rodriguez, R. and K. M. Miller (2014). "Unravelling the genomic targets of small molecules using high-throughput sequencing." Nat Rev Genet 15(12): 783-796.
Rouleau, S. G., J. D. Beaudoin, et al. (2015). "Small antisense oligonucleotides against G-quadruplexes: specific mRNA translational switches." Nucleic Acids Res 43(1): 595-606.
Rubio, C. A., B. Weisburd, et al. (2014). "Transcriptome-wide characterization of the eIF4A signature highlights plasticity in translation regulation." Genome Biol 15(10): 476.
Schaeffer, C., B. Bardoni, et al. (2001). "The fragile X mental retardation protein binds specifically to its mRNA via a purine quartet motif." EMBO J 20(17): 4803-4813.
Sexton, A. N. and K. Collins (2011). "The 5' guanosine tracts of human telomerase RNA are recognized by the G-quadruplex binding domain of the RNA helicase DHX36 and function to increase RNA accumulation." Mol Cell Biol 31(4): 736-743.
Shahid, R., A. Bugaut, et al. (2010). "The BCL-2 5' untranslated region contains an RNA G-quadruplex-forming motif that modulates protein expression." Biochemistry 49(38): 8300-8306.
Shalaby, T., G. Fiaschetti, et al. (2013). "G-quadruplexes as potential therapeutic targets for embryonal tumors." Molecules 18(10): 12500-12537.
Siegfried, N. A., S. Busan, et al. (2014). "RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP)." Nat Methods 11(9): 959-965.
Simone, R., P. Fratta, et al. (2015). "G-quadruplexes: Emerging roles in neurodegenerative diseases and the non-coding transcriptome." FEBS Lett 589(14): 1653-1668.
Sinvani, H., O. Haimov, et al. (2015). "Translational tolerance of mitochondrial genes to metabolic energy stress involves TISU and eIF1-eIF4GI cooperation in start codon selection."

Cell Metab 21(3): 479-492.
Sissi, C., B. Gatto, et al. (2011). "The evolving world of protein-G-quadruplex recognition: a medicinal chemist's perspective." Biochimie 93(8): 1219-1230.
Sonenberg, N. and A. G. Hinnebusch (2009). "Regulation of translation initiation in eukaryotes: mechanisms and biological targets." Cell 136(4): 731-745.
Subramanian, M., F. Rage, et al. (2011). "G-quadruplex RNA structure as a signal for neurite mRNA targeting." EMBO Rep 12(7): 697-704.
Sundquist, W. I. and S. Heaphy (1993). "Evidence for interstrand quadruplex formation in the dimerization of human immunodeficiency virus 1 genomic RNA." Proc Natl Acad Sci U S A 90(8): 3393-3397.
Sundquist, W. I. and A. Klug (1989). "Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops." Nature 342(6251): 825-829.
Taha, M. S., K. Nouri, et al. (2014). "Subcellular fractionation and localization studies reveal a direct interaction of the fragile X mental retardation protein (FMRP) with nucleolin." PLoS One 9(3): e91465.
Takahama, K., A. Takada, et al. (2013). "Regulation of telomere length by G-quadruplex telomere DNA- and TERRA-binding protein TLS/FUS." Chem Biol 20(3): 341-350.
Thandapani, P., T. R. O'Connor, et al. (2013). "Defining the RGG/RG motif." Mol Cell 50(5): 613-623.
Thandapani, P., J. Song, et al. (2015). "Aven recognition of RNA G-quadruplexes regulates translation of the mixed lineage leukemia protooncogenes." Elife 4.
Todd, A. K., M. Johnston, et al. (2005). "Highly prevalent putative quadruplex sequence motifs in human DNA." Nucleic Acids Res 33(9): 2901-2907.
Tome, J. M., A. Ozer, et al. (2014). "Comprehensive analysis of RNA-protein interactions by high-throughput sequencing-RNA affinity profiling." Nat Methods 11(6): 683-688.
Tuesuwan, B., J. T. Kern, et al. (2008). "Simian virus 40 large T-antigen G-quadruplex DNA helicase inhibition by G-quadruplex DNA-interactive agents." Biochemistry 47(7): 1896-1909.
Turner, M. R., O. Hardiman, et al. (2013). "Controversies and priorities in amyotrophic lateral sclerosis." Lancet Neurol 12(3): 310-322.
Valton, A. L., V. Hassan-Zadeh, et al. (2014). "G4 motifs affect origin positioning and efficiency in two vertebrate replicators." EMBO J 33(7): 732-746.
Vasilyev, N., A. Polonskaia, et al. (2015). "Crystal structure reveals specific recognition of a G-quadruplex RNA by a beta-turn in the RGG motif of FMRP." Proc Natl Acad Sci U S A 112(39): E5391-5400.
Vaughn, J. P., S. D. Creacy, et al. (2005). "The DEXH protein product of the DHX36 gene is the major source of tetramolecular quadruplex G4-DNA resolving activity in HeLa cell lysates." J Biol Chem 280(46): 38117-38120.
Wang, Q., J. Q. Liu, et al. (2011). "G-quadruplex formation at the 3' end of telomere DNA inhibits its extension by telomerase, polymerase and unwinding by helicase." Nucleic Acids Res 39(14): 6229-6237.
Westmark, C. J. and J. S. Malter (2007). "FMRP mediates mGluR5-dependent translation of amyloid precursor protein." PLoS Biol 5(3): e52.
Wolfe, A. L., K. Singh, et al. (2014). "RNA G-quadruplexes cause eIF4A-dependent

oncogene translation in cancer." Nature 513(7516): 65-70.
Xu, W., F. Bolduc, et al. (2012). "The use of a combination of computer-assisted structure prediction and SHAPE probing to elucidate the secondary structures of five viroids." Mol Plant Pathol 13(7): 666-676.
Xu, Y., K. Kaminaga, et al. (2008). "Human telomeric RNA in G-quadruplex structure." Nucleic Acids Symp Ser (Oxf)(52): 175-176.
Xu, Y., Y. Suzuki, et al. (2010). "Telomeric repeat-containing RNA structure in living cells." Proc Natl Acad Sci U S A 107(33): 14579-14584.
Zhang, G., M. Hubalewska, et al. (2009). "Transient ribosomal attenuation coordinates protein synthesis and co-translational folding." Nat Struct Mol Biol 16(3): 274-280.
Zhang, Y., C. M. Gaetano, et al. (2014). "FMRP interacts with G-quadruplex structures in the 3'-UTR of its dendritic target Shank1 mRNA." RNA Biol 11(11): 1364-1374.
Zheng, K. W., S. Xiao, et al. (2013). "Co-transcriptional formation of DNA:RNA hybrid G-quadruplex and potential function as constitutional cis element for transcription control." Nucleic Acids Res 41(10): 5533-5541.

Chapter 3

Sam68 regulates S6K1 alternative splicing during adipogenesis

3.1 Preface

Previously, Richard lab identified over 1000 human proteins with RGG/RG motifs by genome-wide searches. Sam68, the STAR family RNA-binding protein, is one of the proteins that contain RGG/RG motif. It has been demonstrated that Sam68 is methylated by PRMT1 on its RGG/RG motif, and Sam68 deficient mice exhibit the lean phenotype. Herein, I further characterized the detailed mechanism of how Sam68 regulates adipogenesis.

3.2 Abstract

The requirement for alternative splicing during adipogenesis is poorly understood. The Sam68 RNA-binding protein is a known regulator of alternative splicing and Sam68 mice deficient exhibit adipogenesis defects due to defective mTOR signaling. Sam68 deficient pre-adipocytes were monitored for alternative splicing imbalances in players of the mTOR signaling pathway. Herein, we report that Sam68 regulates isoform expression of ribosomal S6 kinase (Rps6kb1). Sam68-deficient adipocytes express Rps6kb1-002 and its encoded p31S6K1 protein, in contrast to wild type adipocytes that do not express this isoform. Sam68 binds an RNA sequence encoded by Rps6kb1 intron 6 and prevents serine/arginine-rich splicing factor 1 (SRSF1)-mediated alternative splicing of Rps6kb1-002, as assessed by crosslinking and immunoprecipitation (CLIP) and minigene assays. Depletion of p31S6K1 with siRNAs partially restored adipogenesis of Sam68-deficient pre-adipocytes. The ectopic expression of p31S6K1 in wild type 3T3-L1 cells resulted in adipogenesis differentiation defects, showing that p31S6K1 is an inhibitor of adipogenesis. Our findings indicate that Sam68 is required to prevent the expression of p31S6K1 in adipocytes for adipogenesis to occur.

3.3 Introduction

Src-associated substrate during mitosis of 68kDa (Sam68) is an RNA-binding protein that belongs to the evolutionary STAR (signal transduction activator of RNA) family (Richard 2010; Bielli, Busà et al. 2011). Sam68 is a sequence specific RNA-binding protein that binds repeats of U(U/A)AA sequences in single stranded RNA (Lin, Taylor et al. 1997; Galarneau and Richard 2009). The binding of Sam68 near alternative splice junctions in pre-mRNAs has been shown to regulate splice site selection and regulate the usage of alternative exons (Bielli, Busà et al. 2011). Sam68 promotes the inclusion of CD44 variable exon 5 (v5) and interaction of Sam68 with SND1 (staphylococcal nuclease domain 1) enhances v5 inclusion (Matter, Herrlich et al. 2002; Cappellari, Bielli et al. 2014). The alternative splicing of Bcl-x is regulated by Sam68 and its interaction with hnRNPA1 and FBI-1, affecting pro-survival and apoptotic pathways (Paronetto, Achsel et al. 2007; Bielli, Busà et al. 2014). Sam68 regulates the epithelial-to-mesenchymal transition by decreasing the presence of an alternative serine/arginine-rich splicing factor 1 (*Srsf1*) transcript degraded by nonsense-mediated mRNA decay (Valacca, Bonomi et al. 2010). Sam68 has been shown to regulate alternative splicing of mRNAs during neurogenesis (Chawla, Lin et al. 2009) and in cerebellar neurons (Iijima, Wu et al. 2011). Stimulation of cerebellar neurons using the glutamate receptor agonist kainic acid was dramatically attenuated without Sam68 indicating that Sam68 is required for activity-dependent alternative splicing of Nrxn1 *in vivo* (Iijima, Wu et al. 2011).

The role of Sam68 in alternative splicing has implications for spinal muscular atrophy and FXTAS (fragile X-associated tremor/ ataxia syndrome). Sam68 promotes the skipping of exon 7 leading to a non-functional SMN2 protein and it was shown that the inhibition of Sam68 enhanced exon 7 inclusion in endogenous SMN2, increases SMN levels in SMA patient cells (Pedrotti, Bielli et al. 2010). Expanded CGG repeats in the 5'-UTR of the FMR1 gene causes FXTAS and Sam68 associates with these repeats in RNA aggregates, blocking Sam68 from fulfilling its splicing functions (Sellier, Rau et al. 2010). The inhibition of Sam68 phosphorylation prevents Sam68 from aggregating with RNA, suggesting it may be a therapeutic option for FXTAS patients (Sellier, Rau et al. 2010).

Sam68 null mice have revealed numerous unexpected physiological roles for Sam68. Male Sam68-/- mice are infertile with defects in spermatogenesis, a process where Sam68 has been shown to regulate alternative splicing (Paronetto, Messina et al. 2011) and the polysomal recruitment of specific mRNAs in germline cells (Paronetto, Messina et al. 2009). Ablation of Sam68 leads to increased energy expenditure, decreased number of early adipocyte progenitors, and defective adipogenic differentiation, resulting in mice having a lean phenotype protected against dietary induced obesity (Huot, Vogel et al. 2012). The lack of Sam68 results in mTOR (mammalian target of rapamycin) intron 5 retention and the production of a short transcript (named mTORi5) leading to reduced mTOR protein levels resulting in defects in insulin-stimulated S6 and Akt phosphorylation (Huot, Vogel et al. 2012).

mTOR signaling plays a major role in the regulation of mRNA translation, cell growth, metabolism, and autophagy (Dann, Selvaraj et al. 2007; Laplante and Sabatini 2012; Shimobayashi and Hall 2014). The TSC complex (tuberous sclerosis 1 and 2 heterodimer) is a GTPase activating protein (GAP) on the Ras-like protein Rheb, which activates the mTOR complex 1 (mTORC1) (Inoki, Zhu et al. 2003; Saucedo, Gao et al. 2003; Tee, Manning et al. 2003), and PRAS40 (proline rich Akt substrate of 40kDa) is an inhibitory mediator of mTORC1 signaling. The phosphorylation and inhibition of TSC complex and PRAS40 by the upstream kinase Akt (serine/threonine protein kinase B) activates mTORC1 signaling (Thedieck, Polak et al. 2007; Vander Haar, Lee et al. 2007; Wang, Harris et al. 2007). Activated mTOR signaling results in phosphorylated 4EBP1 (initiation factor 4E-binding protein1) and S6K1 (S6 kinase 1) (Dann, Selvaraj et al. 2007; Sonenberg and Hinnebusch 2009; Shimobayashi and Hall 2014). Active S6K1 phosphorylates the 40S ribosomal protein S6, thereby facilitating mRNA translation, while phosphorylated 4EBP1 promotes the release of eIF4E (eukaryotic translation initiation factor 4E) and initiates translation (Sonenberg and Hinnebusch 2009).

In the present manuscript, we identify Sam68 as an RNA-binding protein that prevents the production of the alternative short isoform of Rps6kb1, encoding p31S6K1, in mouse pre-adipocytes and white adipose tissue (WAT). The binding of Sam68 to an Rps6kb1

intronic RNA sequence counteracted the alternative splicing effects of the SR protein, SRSF1. Expression of p31S6K1 in pre-adipocytes inhibited differentiation, while the depletion of p31S6K1 in Sam68-deficient pre-adipocytes partially restored the adipogenic differentiation defects in a p70S6K1-independent manner. Our findings show that Sam68 is a regulator of Rps6kb1 alternative splicing during adipogenesis.

3.4 Materials and Methods

3.4.1 Alternative splicing assessment and real-time PCR

Total RNA was isolated using the TRIzol® reagent according to the manufacturer's instructions (Invitrogen). Four micrograms of RNA was incubated at 65°C for 5 min and then 42°C for 1 h with 100 pmoles oligo(dT) primer and 100 U of M-MLV reverse transcriptase (catalog no. M1701, Promega) according to the manufacturer's protocol. cDNAs were then amplified by PCR. Endogenous *Rps6kb1* and *Rps6kb1-002* were amplified with the common forward primer 5´-GCA ATG ATA GTG AGG AAT GCT AAG -3´ located in exon 5. The reverse primer for *Rps6kb1* was 5´-GCT GTG TCT TCC ATG AAT ATT CC-3´ located in exon 6 and for *Rps6kb1-002* the reverse primer was 5´-GAA TAG GAG GGC AGA TCC CAT CC -3´ located in exon 6b.

For the real-time RT-PCR, mouse *Rps6kb1* was amplified with 5´-CGT GGA GTC TGC GGC G-3´ located in exon 1 and 5´-CAT ATG GTC CAA CTC CCC CA -3´ located in exon 2, mouse *Rps6kb1-002* was amplified with 5´-TAT GCC TTT CAG ACC GGA GG-3´ located in exon 5 and 5´- ACC TCC CTA AGA CTG CAC CT-3´ located in exon 6b, 18S rRNA was amplified with 5´- GTA ACC CGT TGA ACC CCA TT -3´ and 5´- CCA TCC AAT CGG TAG TAG CG-3´, C/EBPα was amplified with 5´- CGC AAG AGC CGA GAT AAA GC-3´ and 5´- GCG GTC ATT GTC ACT GGT CA-3´, GLUT4 was amplified with 5´-TCG TGG CCA TAT TTG GCT TTG TGG-3´ and 5´- TAA GGA CCC ATA GCA TCC GCA ACA -3´, PPARγ was amplified with 5´-GAA CGT GAA GCC CAT CGA GGA C -3´ and 5´-CTG GAG CAC CTT GGC GAA CA -3´ , as previously described (33), and GAPDH was amplified with 5´- AGC CAC ATC GCT CAG ACA C-3´ and 5´- GCC CAA TAC GAC

CAA ATC C-3′. Sam68 was amplified with 5′-GTG GAG ACC CCA AAT ATG CCC A-3′ and 5′-AAA CTG CTC CTG ACA GAT ATC A-3′. Moreover, primers for mouse GAPDH, Sam68, C/EBPα and PPARγ were purchased from QIAGEN (Valencia, CA). Real-time RT-PCR (RT-qPCR) was performed on a 7500 Fast Real-Time PCR System (Applied Biosystems, Foster City, CA) using SyBR Green PCR Mastermix (QIAGEN, Valencia, CA).

3.4.2 Plasmid constructions

The GFP-Sam68 expression vector encoding an N-terminal green fluorescent protein (GFP) was described previously (Chen, Boisvert et al. 1999). The GFP-SRSF1 expression vector was obtained from Addgene (catalog no. 17990, Cambridge, MA). Gene *Rps6kb1* exon 6, intron 6 and exon 7 were amplified from mouse genomic DNA by PCR using the forward primer 5′- GGG GGA TCC GGA GGA GAA CTA TTT ATG CAG TTA -3′ containing a *Bam*HI site, and the reverse primer 5′- GGG CTC GAG CTT GGT GAT TAA GCA TGA TGT TCT-3′ containing an *Xho*I site. The DNA fragment was then subcloned in the corresponding site of pcDNA3.1 containing a FLAG epitope tag. The mutation of the Sam68 binding site (SBS) in intron 6 of the minigene was performed as follows: primers were used in a two-step PCR reaction forward primer 5′- ATG ATT CAT GTA ATT CCA AGC AAA ACC ACC TT-3′ and reverse primer 5′- AAG GTG GTT TTG CTT GGA ATT ACA TGA ATC AT-3′. The plasmids encoding full-length p31S6K1 were purchased from IDT and subcloned in pcDNA3.1. An expression vector encoding p31S6K1 was kindly provided by Rotem Karni (Hebrew University-Hadassah Medical School). The common forward primer for RT-PCR and RT-qPCR detection of the *Rps6kb1* minigene transcripts was F1: 5′- GAT TAC AAG GAT GAC GAC GAT AAG-3′. The reverse primers for RT-PCR detection were R1: 5′- AGG ATG GAG GGT GTG TCC TAG AGG-3′ and R2: 5′- CTT GGT GAT TAA GCA TGA TGT TCT-3′. The reverse primers for RT-qPCR detection were R3: 5′- CAA TTC AAG GAA ATT CTG CAG TG-3′ and R4: 5′- GCC ATG GAG ATT TCA GCC AAG -3′.

3.4.3 Synthetic RNA oligonucleotides

The RNA oligonucleotides with 3´ biotin tags were synthesized and purchased from IDT. The sequences of these oligonucleotides are Rps6kb1-SBS: 5´- CAU GAU UCA UGU AAU UAA AAG CAA AAC CAC CUU C-3´ biotin; Rps6kb1-SBSmut: 5´- CAU GAU UCA UGU AAU UCC AAG CAA AAC CAC CUU C-3´ biotin.

3.4.4 Pre-adipocyte differentiation and white adipose tissue (WAT)

Sam68-deficient 3T3-L1 cells were generated using pRetrosuper harboring an shRNA targeting Sam68 (Sam68sh) and pRetrosuper 3T3-L1 cells were used as control, as described previously (Huot, Vogel et al. 2012). Pre-adipocyte 3T3-L1 adipogenic differentiation was performed as described (Sun, Ma et al. 2005). The cells were fixed with 3% formaldehyde and 0.025% glutaraldehyde and incubated with Oil-Red-O solution (Sigma-Aldrich, St. Louis, Mo). Cell extracts were prepared and analyzed, as described previously (Huot, Vogel et al. 2012). Antibodies for Sam68 (Millipore), p70 S6K (BD Transduction Laboratories, Cell Signaling), GFP (Roche), SRSF1 (Santa Cruz), FLAG M2, β-actin and β-tubulin (Sigma) were purchased.

Generation of stable 3T3-L1 clones overexpressing p31S6K1 were generated as follows: Cells were transfected with either pcDNA3.1 and pcDNA3.1 Flag-p31 plasmid constructs. After 48 h post transfection, G418 was added in the medium and individual clones selected several weeks later. The expression level of p31S6K1/Flag was assessed by immunoblotting.

3.4.5 RNA interference and transfection

Human KHDRBS1 (Sam68) siGENOME SMARTpool siRNA (catalog no. M-020019-00-0010) and mouse KHDRBS1 (Sam68) siGENOME SMARTpool siRNA (catalog no. M-065115-01-0010) were ordered from Dharmacon (Thermo Scientific). Human SRSF1 siGENOME SMARTpool siRNA (catalog no. M-018672-00-0005) and mouse SRSF1 siGENOME SMARTpool siRNA (catalog no. M-040886-01-0005) were ordered from Dharmacon (Thermo Scientific). Mouse RPS6KB1 (p70/p31) siGENOME set of 4

siRNA (catalog no. MQ-040893-02-0002) was ordered from Dharmacon (Thermo Scientific). The following siRNAs were also ordered from Dharmacon (Thermo Scientific): mouse p31 siRNA-A sense sequence : 5'- GCU CUU CAC UGC AGA AUU UUU-3' and antisense sequence: 5'- AAA UUC UGC AGU GAA GAG CUU-3', mouse p31 siRNA-B sense sequence: 5'- ACA CAG AAG CUG CAU UUA AUU-3' and antisense sequence: 5'- UUA AAU GCA GCU UCU GUG UUU-3', siGFP siRNA sense sequence: 5'-AAC ACU UGU CAC UAC UUU CUC UU-3' and antisense sequence: 5'-GAG AAA GUA GUG ACA AGU GUU UU-3'.

For siRNA transfections, typically cells were plated in 6-well plates and transfected with 100 nM siRNA using Lipofectamine™ RNAiMAX (siRNA), as recommended by the manufacturer (Invitrogen). HEK293 plated in 6-well plates were transfected using Lipofectamine™ 2000. Each well received a total of 5 μg with GFP-SRSF1 (0, 0.25, 1 and 2 μg), GFP-Sam68 (1 μg) and 2 μg of the indicated minigene and empty vector was used to compensate.

3.4.6 RNA binding assays

3T3-L1 cells were lysed in 1 mL of cell lysis buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 1% Triton X-100, 40 units/ml RNaseOUT and supplemented with Roche Complete Mini, EDTA-free protease inhibitor) and incubated for 15 min at 4°C. Lysates were cleared by centrifugation and 5 μL of 100 μM biotinylated RNA was added to the lysates and incubated at 4°C for 60 min with constant end-over-end mixing with Streptavidin Sepharose beads. The beads were washed 3 times with lysis buffer and once with 1X PBS. Protein samples were analyzed on SDS-polyacrylamide gels and transferred to nitrocellulose membrane for immunoblotting.

3.4.7 Immunoprecipitation

Transfected HEK293 cells were lysed in buffer containing 1% Triton X-100, 150 mM NaCl, 20mM Tris (pH 7.5) and proteinase inhibitor (Roche) for 15 min on ice. Total cell lysates were clarified by centrifugation for 10 min at 10,000 x g at 4^{o} C. The lysates were incubated

with indicated antibodies 1 h or overnight at 4˚C and then 20 µl of 50% protein A/G slurry was added. The mixture was incubated for 30 min at 4˚C. The Protein A/G Sepharose beads were washed 3 times with lysis buffer and once with 1X PBS. The samples were boiled and subjected to standard Western blot analysis.

3.4.8 UV-crosslinking and immunoprecipitation (CLIP)

3T3-L1 cells (pRetrosuper and Sam68sh cells) were treated with 4-thiouridine to a final concentration of 100 µM directly to the cell culture medium 8 h prior to crosslinking. The cells were washed with ice-cold PBS and irradiated with 0.15 J/cm^2 of 365nm UV light at 4°C. The cells were collected by centrifugation at 514 g for 1 min at 4°C. The cell pellets were resuspended in CLIP lysis buffer supplemented with protease inhibitors (Roche) and 0.5 U/µl RNasin (Promega) and sonicated twice with 10s bursts (Huppertz, Attig et al. 2014). The lysates were added with 10 µl of 1:250 dilution of RNase I (Life Technologies) and 2 µl Turbo DNase (Life Technologies) shaking at 37°C for 3 min. The lysates were then cleared and immunoprecipitated with 2 µg anti-Sam68 or anti-SRSF1 antibody and control mouse/ rabbit IgGs (Santa Cruz). Proteinase K Buffer (containing 1.2 mg/ml Proteinase K) was added to the immunoprecipitates and incubated for 20 min at 37°C. RNA was isolated through TRIzol® reagent and subjected to RT-qPCR. The reverse primers below were used for the reverse transcription reaction. qPCR was performed with *Rps6kb1* (intron 6 SBS, forward) 5´-GAT TCA GGT CAT GAT TCA TG -3´ and *Rps6kb1* (intron 6 SBS, reverse) 5´-CAG TGG GAA GGT GGT TTT GC-3´; *Rps6kb1* (exon 6, SRSF1 site, forward) 5´-GAG GAG AAC TAT TTA TGC AG -3´ and *Rps6kb1* (exon 6, SRSF1 site, reverse) 5´-GAA TAT TCC CTC TCT TTC TAA-3´; *Rps6kb1* (exon 7, forward) 5´- TTT ACT TGG CTG AAA TCT CC-3´ and *Rps6kb1* (exon 7, reverse) 5´- CTT GGT GAT TAA GCA TGA TG-3´.

3.5 Results

3.5.1 Sam68 regulates the alternative splicing of *Rps6kb1* in pre-adipocytes and mouse white adipose tissue (WAT)

Sam68-deficient pre-adipocytes are unable to differentiate into mature adipocytes (Huot, Vogel et al. 2012). We reported that Sam68-deficient pre-adipocytes have decreased mTOR expression, as they increase the production of a short $mTOR_{i5}$ isoform rather than synthesizing the full-length mTOR mRNA (Huot, Vogel et al. 2012). The Sam68-deficient pre-adipocyte defect is partially rescued by the ectopic expression of the full-length mTOR expression, suggesting that there may be other splicing events regulated by Sam68 in the mTOR signaling pathway. To identify these alternative splicing events that contribute to the differentiation defects of Sam68-deficient pre-adipocytes, we monitored the presence of spliced isoforms in the mTOR signaling pathway. Using the ENSEMBL database, we identified the existence of spliced isoforms for the murine *Rps6kb1, TSC1, Rheb, Akt1,* and *Deptor* genes, but not for *IRS1*, *TSC2*, *4EBP1 or eIF4E* (data not shown). Amongst the candidate isoforms tested, we observed that the mRNA levels of isoform *Rps6kb1-002* were dramatically increased in Sam68-deficient cells (Sam68sh) compared to control pRetrosuper 3T3-L1 cells (Figure 3.1A, B). We also noted a slight to moderate up-regulation of isoforms *TSC1-003*, *Rheb-003, Akt1-003,* and *Rps6kb1-005* in Sam68sh 3T3-L1 cells and we did not observe significant fluctuations with the following isoforms in Sam68sh 3T3-L1 cells: *TSC1-006*, *Rheb-002*, *Akt1s1-201*, and *Deptor-002* (data not shown).

We next examined the levels of isoform *Rps6kb1-002* in white adipose tissue (WAT) of wild type and Sam68 null mice. The level of *Rps6kb1-002* was more abundant in white adipose tissue of Sam68 deficient mice compared with littermate control mice (Figure 3.1C). The increase in *Rps6kb1-002/ Rps6kb1* mRNA ratio was also confirmed by RT-qPCR using primers specific for each isoform in Sam68-deficient pre-adipocytes and WAT isolated from Sam68 null mice (Figure 3.1D). Thus the loss of Sam68 promotes the production of splicing variant *Rps6kb1-002*.

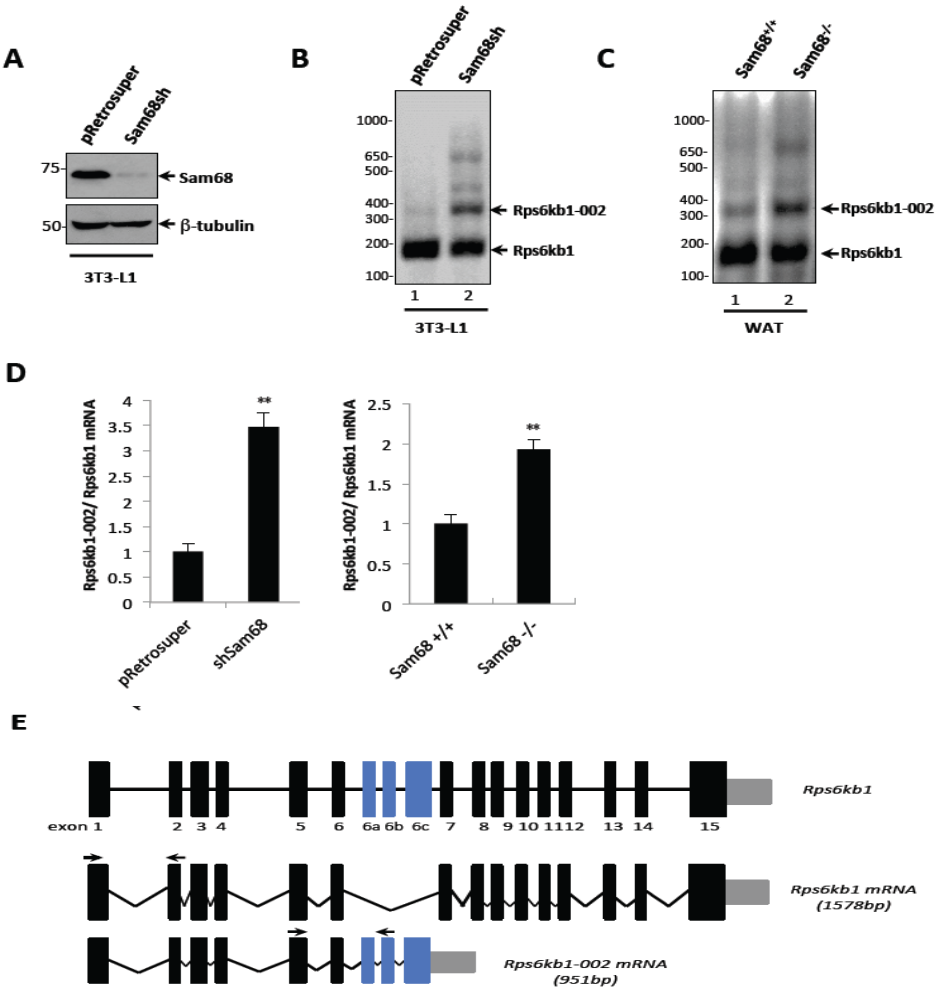


Figure 3.1 Sam68 regulates the alternative splicing of *Rps6kb1* in mouse pre-adipocytes and WAT.

A, Mouse pre-adipocytes (3T3-L1 cells) were transfected with pRetrosuper empty plasmid (pRetrosuper) and pRetrosuper containing the Sam68sh sequences. These cells were selected for analysis by immunoblotting using anti-Sam68 and β-tubulin antibodies.

B, C Total RNA from undifferentiated pRetrosuper and Sam68sh 3T3-L1 cells and mouse WAT was isolated and analyzed using a three primer RT-PCR strategy with a common forward primer in exon 5 and reverse primers in exon 6 and 6b. The DNA markers are shown on the left in base pairs. Schematic representation of the spliced variants is shown in *panel E*.

D, Total RNA from undifferentiated pRetrosuper and Sam68sh 3T3-L1 cells and mouse WAT was isolated and subjected to RT-qPCR. The presence of the *Rps6kb1-002* expressed as a ratio of total *Rps6kb1* transcripts. Error bars represent ± standard deviation of the mean ** $p < 0.01$.

E, Schematic representation of *Rps6kb1* gene, the wild type isoform *Rps6kb1* and alternative spliced isoform *Rps6kb1-002*. Constitutive exons are shown as black boxes, and alternative exons are shown in blue. Introns are shown as horizontal lines, and splicing events are indicated by angled lines. The 3'-UTRs are shown as grey boxes. The filled arrowheads denote the primers utilized to define the alternative splice isoforms by RT-qPCR of *panel D*.

3.5.2 Sam68-deficiency increases the expression of p31S6K1

Rps6kb1 encodes p85/p70 S6K1 and the inclusion of alternative exons 6a, 6b and 6c leads to the generation of the *Rps6kb1-002* isoform (Figure 3.1E). The *Rps6kb1* transcript generates 2 proteins due to alternative mRNA translation start sites resulting in p70S6K1 and p85S6K1, whereas the shorter *Rps6kb1-002* isoform harbors only the first 6 exons with alternative exons 6a, 6b, 6c and its alternative splicing was shown to be positively regulated by SRSF1 (Karni, de Stanchina et al. 2007). The presence of a stop codon in exon 6c generates a truncated protein of 31 kDa, termed p31 or p31S6K1 that expresses the S6K1 N-terminal domain followed by a truncated kinase domain. The increase of *Rps6kb1-002* mRNA was reflected at the protein level, since we observed the presence of p31S6K1, as well as p70S6K1 in Sam68-deficient pre-adipocytes by immunoblotting with an N-terminal S6K1 antibody (BD Transduction Lab Inc., Figure 3.2A). We observed similar results in mouse WAT isolated from Sam68$^{-/-}$ mice using a different anti-S6K1 antibody (Cell Signaling Inc.) that detects p31S6K1 in addition to p70S6K1 and p85S6K1 (Figure 3.2B). As SRSF1 is a known regulator of p31S6K1 (Karni, de Stanchina et al. 2007), and Sam68 has been shown to regulate the alternative splicing of *Srsf1* associated with non-sense mediated decay (Valacca, Bonomi et al. 2010), we performed immunoblotting to examine SRSF1 levels in the absence of Sam68. The depletion of Sam68 in pre-adipocytes or Sam68-deficient WAT did not affect the SRSF1 protein levels (Figure 3.2A, 3.2B). These findings show that Sam68-deficient pre-adipocytes have increased p31S6K1 expression using two different N-terminal S6K1 antibodies with little to no effect on the global expression of p70S6K1 and p85S6K1.

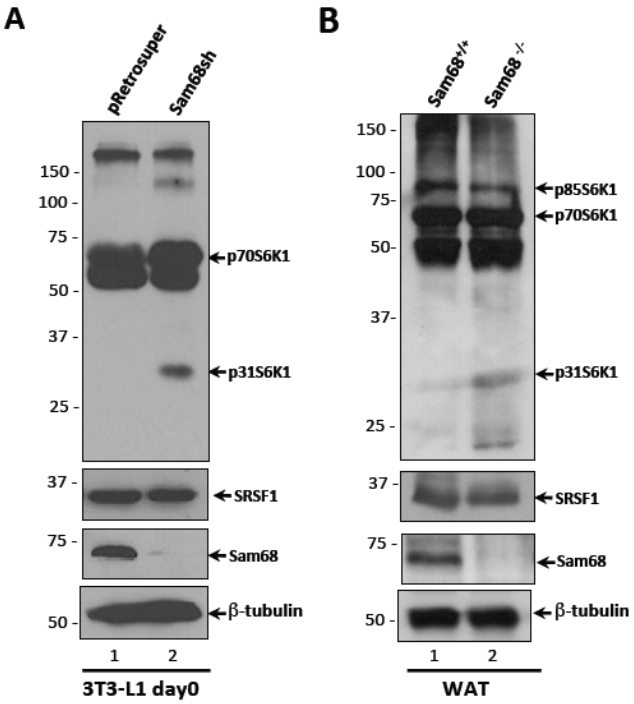


Figure 3.2 Sam68 regulates the expression of p31S6K1, but not SRSF1 in pre-adipocytes and WAT.

A, Protein extracts from pRetrosuper and Sam68sh 3T3-L1 cells,

B, or from mouse WAT, were immunoblotted with anti-S6K1 antibodies (BD Transduction Labs for panel A and from Cell Signaling Inc. for panel B. Anti-Sam68 and anti-β-tubulin antibodies were used to monitor the levels of Sam68 and β-tubulin, respectively. The molecular mass markers are shown on the left in kDa.

3.5.3 Sam68 binds an RNA element within intron 6 that diminishes SRSF1 binding to *Rps6kb1* exon 6

Sam68 binds RNA with U(U/A)AA motifs with high affinity (Lin, Taylor et al. 1997; Galarneau and Richard 2009). Sam68 binding sites (SBS) often reside near splice sites within pre-mRNAs, acting either as splice enhancers or silencers to regulate neighboring splice site usage (Bielli, Busà et al. 2011). Since the absence of Sam68 increases the splicing of *Rps6k1-002*, we searched intron 6 for repeats of the U(U/A)AA motif. The *Rps6kb1* gene sequence is shown in Figure 3.3A with the *Rpsk6kb1-002* alternative exons in blue. We identified a putative Sam68 binding site within the *Rpsk6kb1* intron 6 with an encoded sequence of 5'-UAAUUAAA-3', termed SBS, 68 nucleotides downstream of a putative SRSF1 binding site in exon 6 (Figure 3.3A). To determine whether Sam68 binds the SBS sequence, we synthesized a biotinylated RNA of the SBS, as well as biotinylated control RNA (SBSmut) that has the 5'-UAAUUAAA-3' replaced with 5'-UAAUUCCA-3' (Figure 3. 3B). The RNA oligonucleotides harboring either wild type or mutated SBS were used to perform affinity pull-down assays. Cell lysates from undifferentiated wild type 3T3-L1 cells were incubated with the biotinylated RNAs and complexes purified with streptavidin Sepharose beads. The bound proteins were separated by SDS-PAGE and immunoblotted for Sam68. Wild type SBS bound Sam68 with a much higher affinity than the mutated SBS, as assessed by increasing the concentration of salt in the wash buffer (Figure 3.3B). These findings show that Sam68 associates *in vitro* with RNA sequences within intron 6 of the *Rps6kb1* pre-mRNA.

We next examined whether endogenous Sam68 bound the intron 6 SBS *in vivo* using UV-crosslinking and immunoprecipitation (CLIP) with a dilution of RNase I (1:250) that digests RNAs into fragments of 50 to 300 nucleotides in length (Huppertz, Attig et al. 2014). Pre-adipocytes (pRetrosuper and Sam68sh cells) were prepared for CLIP, as described in ‘Materials and Methods’ and immunoprecipitated with either control immunoglobulin G (IgG), anti-Sam68 antibodies or anti-SRSF1 antibodies. The precise binding site was mapped by using primers indicated in Figure 3.4A. Anti-Sam68 immunoprecipitations compared to IgG, enriched ~25-fold for the *Rps6kb1* intron 6 region spanning the SBS, but not an RNA

region spanning *Rps6kb1* exon 7 in pRetrosuper 3T3L1 cells (Figure 3.4B). In contrast, there was no RNA enrichment detected in anti-Sam68 immunoprecipitations in Sam68sh 3T3-L1 cells, as expected (Figure 3.4B). This finding suggests that Sam68 associates *in vivo* with the SBS site of the *Rps6kb1* intron 6. CLIP with anti-SRSF1 antibodies revealed a modest 2-fold enrichment of *Rps6kb1* exon 6 fragment encompassing the putative SRSF1 binding site over control in pRetrosuper 3T3L1 cells (Figure 3.4C). Interestingly, cells depleted of Sam68 contained a ~ 6-fold increase in SRSF1 at this site (Figure 3.4C), suggesting that Sam68 occupancy at the SBS prevents SRSF1 binding to exon 6.

Sam68 and SRSF1 do not have overlapping RNA binding sites in *Rps6kb1* but they may mediate interact by protein-protein interactions. Mass spectrometry analysis using endogenous Sam68 in HeLa cells indeed immunopurifies SRSF1 (Huot, Vogel et al. 2009). We further confirmed the interaction using the ectopic expression of both GFP-Sam68 and GFP-SRSF1 in HEK293 cells. The cells were lysed and immunoprecipitated with anti-SRSF1 antibodies or control IgG. The bound proteins were separated by SDS-PAGE and immunoblotted with anti-Sam68 or anti-SRSF1 antibodies. GFP-Sam68 and endogenous Sam68 co-immunoprecipitated with SRSF1 (Figure 3.4D, upper panel). These data confirm that Sam68 binds the SBS sequence and influences SRSF1 binding to its RNA element likely by protein-protein interaction.

3.5.4 Minigene assays indicate that Sam68 suppresses the alternative splicing of *Rps6kb1-002* by binding *Rps6kb1* intron 6 SBS

A splicing minigene was constructed with a cytomegalovirus (CMV) promoter driving the expression of a 1.6kb genomic fragment encompassing *Rps6kb1* exon 6, intron 6, and exon 7 (Figure 3.5A). The minigene transcription start site was located in the plasmid upstream of *Rps6kb1* exon 6 followed by the sequence of a FLAG epitope tag. A forward primer complementary to the FLAG cDNA sequence (F1) and reverse primers in exons 6c (R1), exon 6a (R3) and 7 (R2 and R4) that recognize fragments corresponding to *Rps6kb1*-002 and *Rps6kb1*, respectively. pRetrosuper and Sam68sh 3T3-L1 cells were transfected with either pcDNA3.1, the wild type *Rps6kb1* or the SBS mutated (SBSmut) minigene. Total RNA was

isolated 48 h after transfection, treated with RQ1 DNase and monitored for *Rps6kb1*-002 and *Rps6kb1* transcripts using 3-primer RT-PCR. There was little expression of the *Rps6kb1*-002 fragment from the wild type minigene (Figure 3.5B, lane 2), which increased significantly in minigene SBSmut transfected pRetrosuper and Sam68sh cells (Figure 3.5B, lanes 3 and 6). Quantification using RT-qPCR revealed that deletion of the SBS increased the *Rps6kb1*-002/ *Rps6kb1* mRNA ratio to ~6-fold in pRetrosuper 3T3-L1 (Figure 3.5C). Sam68sh cells harbored an *Rps6kb1*-002/ *Rps6kb1* mRNA ratio of about 8-fold that increased to ~40-fold with the transfection the SBSmut minigene (Figure 3.5C). These observations suggest that Sam68 represses *Rps6kb1-002* when bound to the SBS in *Rps6kb1* intron 6. The additive effect observed with the SBSmut minigene in Sam68sh 3T3 cells, suggests that the intron 6 SBS bound by Sam68 (Figure 3.5B, 3.5C) is a regulatory site, but there are likely other Sam68 binding sites in the *Rps6kb1* minigene that could contribute to the suppression of isoform *Rps6kb1-002*.

We next overexpressed Sam68 to examine its influence on the wild type and SBSmut *Rps6kb1* minigenes. The expression of GFP-Sam68 in HEK293 cells was confirmed by immunoblotting (Figure 3.5D). The *Rps6kb1-002/Rps6kb1* mRNA ratio was higher in pcDNA3.1 transfected HEK293 cells with the SBSmut *Rps6kb1* minigene compared to the wild type minigene by RT-qPCR (Figure 3.5E), consistent with data in 3T3-L1 cells (Figure 3.5B, 3.5C). Expression of GFP-Sam68 completely quenched the expression of *Rps6kb1-002* and the ratio of *Rps6kb1-002/Rps6kb1* mRNA was below 1.0 (Figure 3.5E). These findings suggest that elevated levels of Sam68 associates with SRSF1 and prevents it from up-regulating *Rps6kb1-002*.

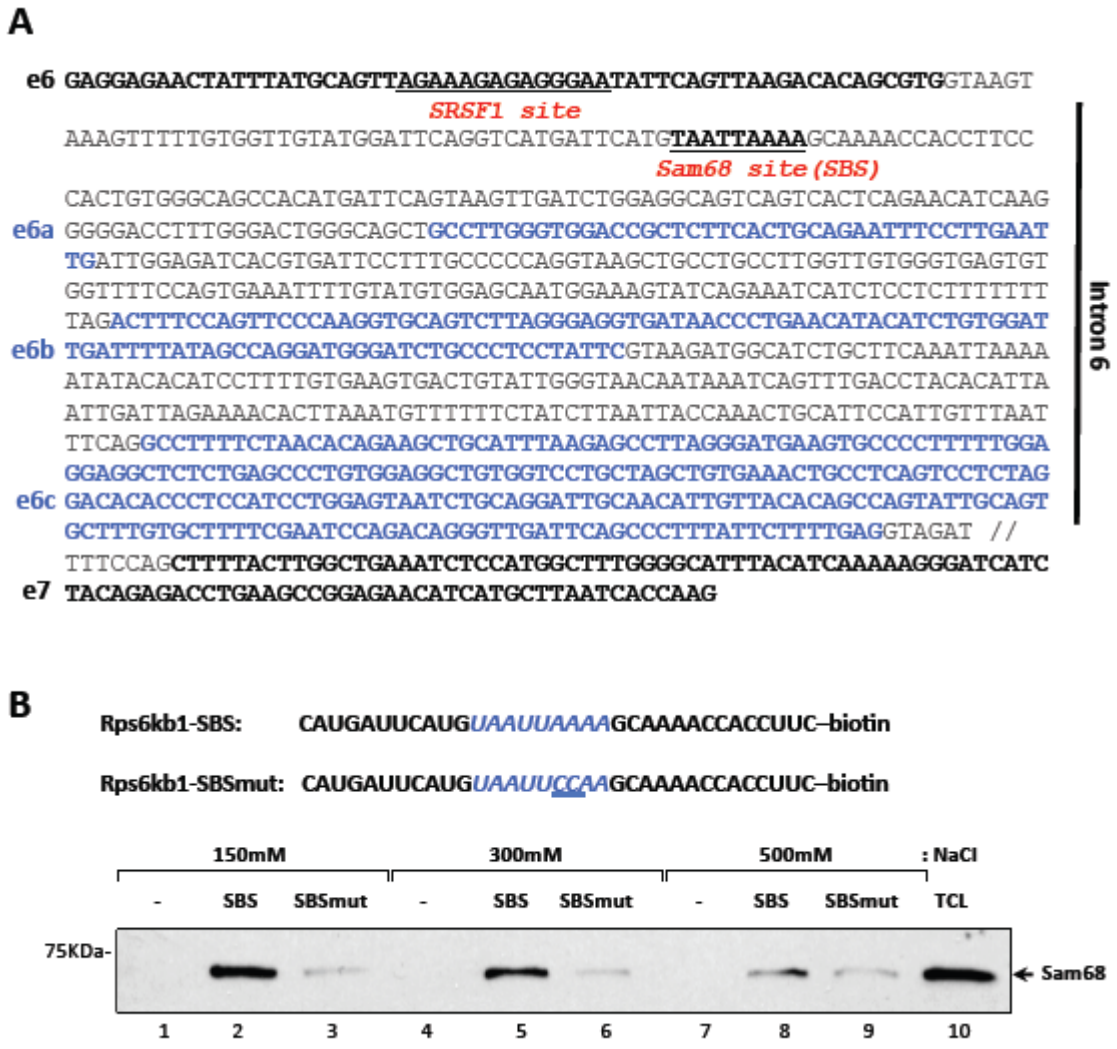


Figure 3.3 Sam68 associates *in vitro* with RNA elements in *Rps6kb1* intron 6.

A, Sequence spanning mouse *Rps6kb1* exon6, intron 6 which contains the three alternative exons (6a, 6b, 6c) and exon 7 taken from the ENSEMBL browser. Underlined sequences represent the Sam68 binding site (SBS) within intron 6 and the SRSF1 binding site in exon 6.

B, The sequence of intron 6 and SBS, as well as the mutated version (SBSmut) of the synthetic RNAs generated. Affinity pull-down assays were performed with the biotinylated RNA and Streptavidin beads using 3T3-L1 cells and immunoblotted with Sam68 antibodies. TCL represents total cell lysate. The molecular mass markers are shown on the left in kDa

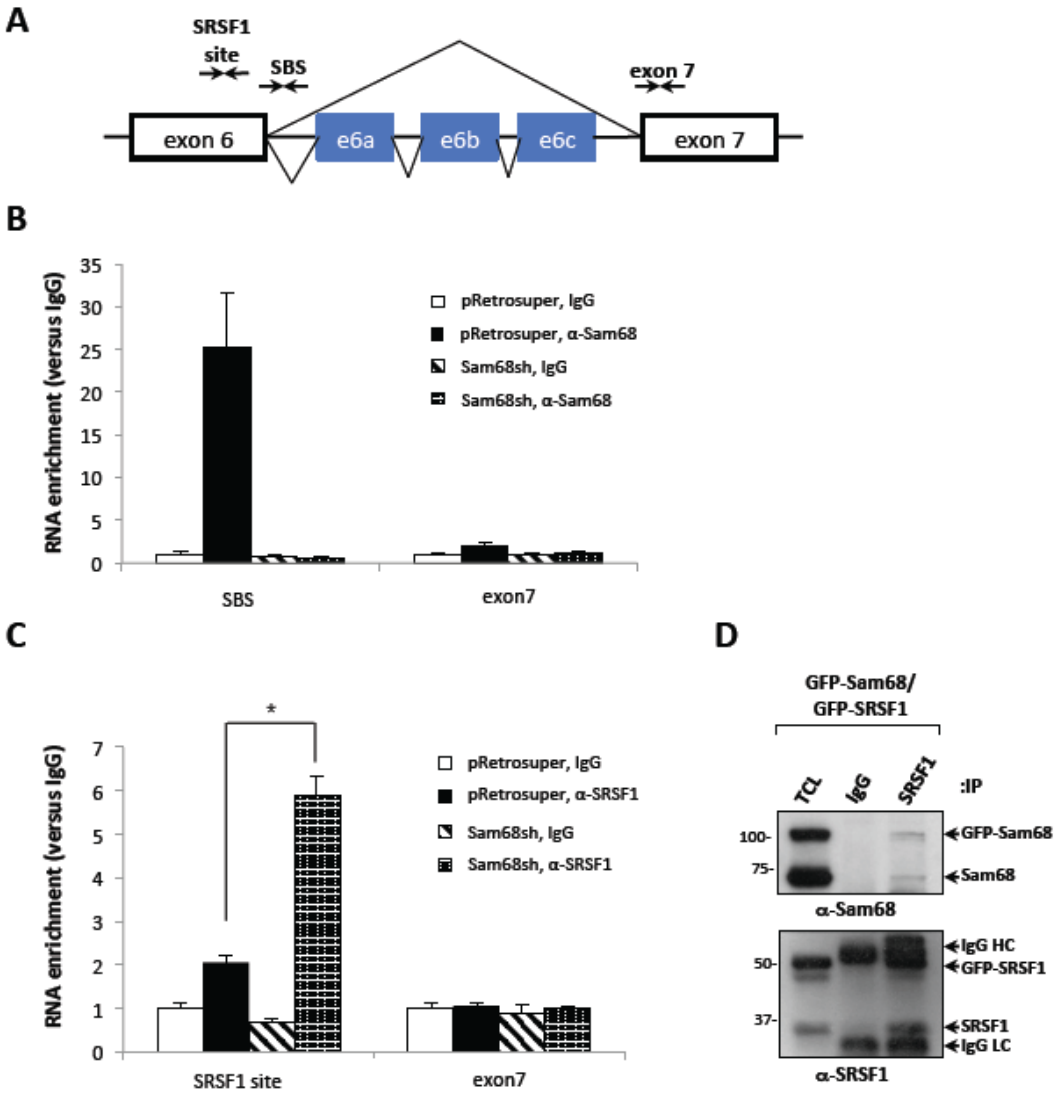


Figure 3.4 Sam68 binds an intronic SBS and prevents the binding of SRSF1 to its consensus site in *Rps6kb1* exon 6.

A, Schematic of the genomic architecture of *Rps6kb1* spanning exons 6 to 7 with three alternative exons located in intron 6. Arrows depict the primer pairs used in RT-qPCR to detect the RNA-bound Sam68 binding site (SBS), the exon 6 SRSF1 binding site, and exon 7 as a negative control.

B,C, CLIP assays were performed using anti-Sam68 antibodies, and anti-SRSF1 antibodies or control IgGs. Bound RNA was analyzed in triplicate by RT-qPCR with the primers shown in panel ***A***. The levels of bound RNA of the SBS, SRSF1 binding site and exon 7 in immunoprecipitates were normalized to the levels of the total RNA in the input. Mean values are expressed as fold enrichment. Error bars represent ± standard deviation of the mean. * $p < 0.05$.

D, HEK293 cells were transfected with expression vectors encoding GFP-Sam68 and GFP-SRSF1. After 48 h, the cells were lysed and subjected to immunoprecipitation (IP) and immunoblotting. The migration of GFP-Sam68, Sam68, GFP-SRSR1, SRSF1 and the heavy chain (HC) and light chains (LC) of IgG is indicated. TCL represents the total cell lysate. The molecular mass markers are shown on the left in kDa.

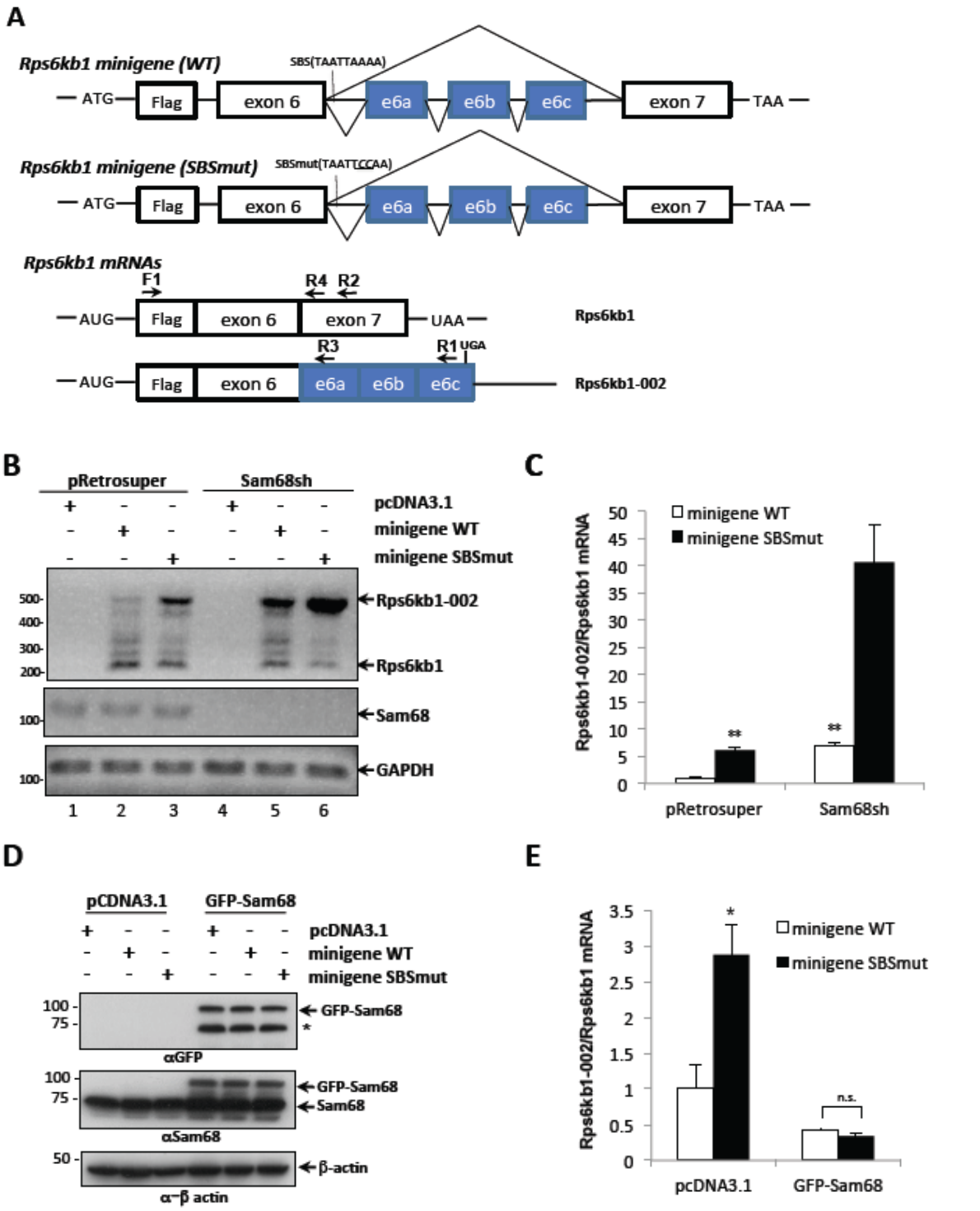


Figure 3.5 *Rps6kb1* minigene assay defines intron 6 as the minimal requirement.
A, Mouse genomic fragment encompassing *Rps6kb1* exons 6 to 7 was cloned in pcDNA3.1 with an N-terminal FLAG epitope tag. The SBS is shown in the proximity of the 5' splice site within intron 6. The same mouse genomic fragment containing a 'CC' replacing the 'AA' in the Sam68 binding site was also constructed in pcDNA3.1 with an N-terminal FLAG epitope tag. The splicing of wild type *Rps6kb1* and SBSmut *Rps6kb1* minigene, leads to two different transcripts. Arrows indicate the primers used in the minigene assay.
B, C pRetrosuper and Sam68sh 3T3-L1 cells were transfected either with pcDNA3.1 alone,

Rps6kb1 (WT) or *Rps6kb1* (SBSmut) minigenes. Total RNA was isolated, digested with RQ1 DNase and the levels of *Rps6kb1-002* and *Rps6kb1* transcript assessed by semi-quantitative RT-PCR using primers F1 with R1 and R2 (*panel B*), or by RT-qPCR using F1 with R3 and R4 (*panel C*). The mRNA levels of Sam68 and GAPDH were also assessed (*panel B*). Error bars represent ± standard deviation of the mean. ** $p < 0.01$.***D,*** HEK293 cells were co-transfected with either pcDNA3.1 or GFP-Sam68 alone with *Rps6kb1* minigene or *Rps6kb1* minigene SBSmut. The cells were harvested after 48 h and the cellular extracts were immunoblotted with anti-GFP, anti-Sam68 and anti-β-actin antibodies. The asterisk denotes a non-specific protein recognized with anti-GFP antibodies. ***E,*** The total RNA from the cells indicated in ***D*** was isolated, treated with RQ1 DNase, and the mRNA levels of *Rps6kb1* and *Rps6kb1-002* were assessed by RT-qPCR. The mRNA expression level of *Rps6kb1-002* was normalized to *Rps6kb1* levels. Error bars represent ± standard deviation of the mean. * $p < 0.05$ and n.s. denotes not significant.

3.5.5 Sam68 counteracts the positive effects of SRSF1 for *Rps6kb1-002* expression

We next confirmed that SRSF1 was responsible for regulating the p31S6K1 levels in Sam68sh 3T3-L1 cells. Depletion of SRSF1 using siRNA reduced the levels of p31S6K1 (Figure 3.6A) and decreased the *Rps6kb1*-002/ *Rps6kb1* mRNA ratio by RT-qPCR in Sam68sh, but not pRetrosuper 3T3-L1 cells (Figure 3.6B). These findings confirm that SRSF1 is required for the production of *Rps6kb1*-002 in adipocytes.

We next examined the influence of SRSF1 and Sam68 on the *Rps6kb1* minigene in HEK293 cells, a cell type that is easily transfected. Firstly, SRSF1-depleted cells by siRNA were generated (Figure 3.7A). HEK293 cells treated with siGFP contained an *Rps6kb1*-002/ *Rps6kb1* transcript ratio that was similar in SRSF1-depleted cells using the wild type minigene (Figure 3.7B, lanes 2 and 5). The SBSmut minigene transfected in siGFP cells displayed an *Rps6kb1*-002/ *Rps6kb1* transcript ratio close to 8.0 (Figure 3.7B, lane 3), suggesting that uncoupling Sam68 binding to the SBS stimulates *Rps6kb1*-002. The induction of *Rps6kb1*-002 observed with the SBS mutation was SRSF1-dependent (Figure 3.7B, lane 6). These findings show that SRSF1 is the positive factor required to promote the expression of *Rps6kb1*-002. We next ectopically expressed SRSF1 and Sam68 in HEK293 cells (Figure 3.7C). The transfection of GFP-tagged SRSF1 increased the appearance of *Rps6kb1*-002 in a dose-dependent manner (Figure 3.7D) and this increase was attenuated with the overexpression of GFP-Sam68 (Figure 3.7D), consistent with findings of Figure 3.5D and 3.5E. These data suggest that Sam68 counteracts the positive effects of SRSF1 for expression of the *Rps6kb1*-002 isoform.

3.5.6 The ectopic expression of p31S6K1 suppresses adipogenesis

The direct role of the S6K1 short isoform in adipogenesis was investigated. Pre-adipocytes were stably transfected with pcDNA3.1 (control) or an expression vector encoding FLAG-p31S6K1. Two stable clones (FLAG-p31 #3 and #10) ectopically expressing p31S6K1 were selected (Figure 3.8A) and the expression was comparable to levels observed in Sam68-deficient cells (data not shown). p31S6K1 ectopically expressing cells were

monitored for lipid accumulation following differentiation for 4 days. We observed a notable decrease of lipid accumulation in both p31S6K1 overexpressing 3T3-L1 cell lines, compared to control cells, as measured by Oil-Red-O staining (Figure 3.8B). We subsequently examined the expression pattern of adipose-specific markers at differentiation day 0 and day 4 (Figure 3.8C). The mRNA levels of PPARγ, C/EBPα and GLUT4 were increased dramatically in control cells (pcDNA3.1) upon differentiation, as expected, while in clones #3 and #10, the expression of these markers was largely absent after differentiation day 4, compared to the control cells (Figure 3.8C). These results show that p31S6K1 is a repressor of adipogenesis.

3.5.7 Depletion of p31S6K1 in Sam68-deficient pre-adipocytes partially rescues the adipogenesis defect

We examined whether the elevated level of p31S6K1 is a contributing factor for the adipogenesis defect of Sam68-deficient 3T3-L1 cells. To test this possibility, we decreased the expression of p31S6K1 using siRNA specific to this isoform. pRetrosuper and Sam68sh 3T3-L1 cells were transfected with either siGFP or two different p31S6K1 siRNAs, abbreviated sip31-A and sip31-B. The elevated expression of p31S6K1 in Sam68sh 3T3-L1 cells was depleted in sip31-A and sip31-B transfected cells (Figure 3.9A). We monitored cell morphology, expression of key adipocyte markers and lipid accumulation in these cells. Sam68sh 3T3-L1 cells exhibited reduced adipogenesis compared to pRetrosuper cells, as previously reported and the presence of transfected siGFP had no influence on adipogenesis (Huot, Vogel et al. 2012). The transfection of p31S6K1 siRNA increased lipid accumulation in Sam68-deficient cells, suggesting that the loss of p31S6K1 expression partially rescued the adipogenesis defect observed in Sam68sh cells (Figure 3.9B, compare siGFP with sip31-A and sip31-B). We also examined the expression pattern of adipose-specific markers at differentiation day 0 and 4 of 3T3-L1 cells. The mRNA levels of PPARγ, C/EBPα and GLUT4 were increased in pRetrosuper cells upon differentiation, as expected (Figure 3.9C). Strikingly, the attenuated expression of differentiation markers (PPARγ, C/EBPα and GLUT4) in Sam68sh cells was partially de-repressed with the depletion of p31S6K1 (Figure 3.9C).

These results indicate that expression of p31S6K1 represses adipogenesis and is a contributing factor for the observed defects in Sam68-deficient pre-adipocytes.

To exclude the possibility that elevated levels of p31S6K1 inhibit adipogenesis by decreasing the levels of the p70S6K1 isoform, we abrogated the total S6K1 expression using siRNAs. pRetrosuper 3T3-L1 cells were transfected with either siGFP or siRNA targeting both the p70/p31 S6K1 isoforms. Deletion of total S6K1 in pRetrosuper 3T3-L1 cells had no influence on adipogenesis, as cells differentiated normally and expressed high levels PPARγ, C/EBPα and GLUT4 at day 4 of differentiation (Figure 3.10B, 3.10C). Since p31S6K1 is absent in pRetrosuper cells, these findings demonstrate that deletion of p70S6K1 alone does not affect the differentiation of 3T3-L1 cell, as reported previously (Carnevalli, Masuda et al. 2010). Next we examined whether Sam68sh cells were partially rescued with siRNAs targeting p70/p31S6K1 and indeed this was the case (Figure 3.10B, 3.10C). The partial de-repression we observed was similar to decreasing the levels of p31S6K1 alone (compare with Figure 3.9). These observations indicate that p31S6K1 inhibits adipogenesis independent of the p70S6K1 isoform.

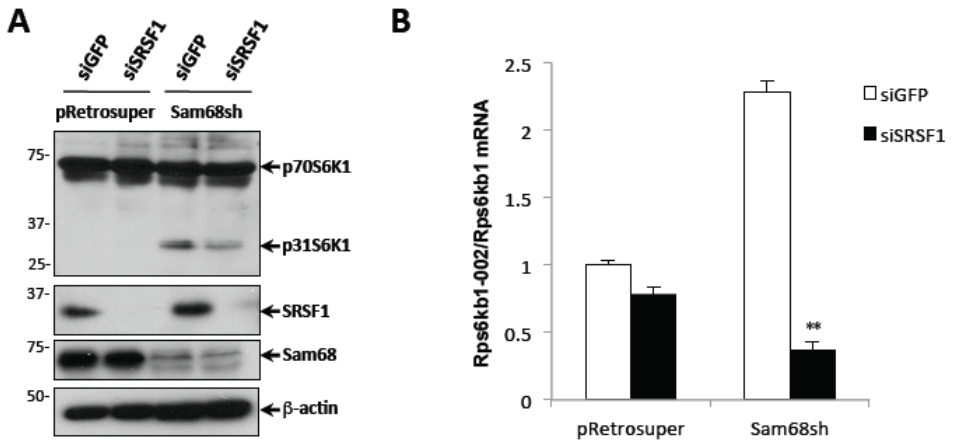


Figure 3.6 The presence of the *Rps6kb1-002* isoform in Sam68-depleted pre-adipocytes requires SRSF1.

A, pRetrosuper or Sam68sh 3T3-L1 cells were transfected with siGFP or siSRSF1. The protein extracts were prepared 48 h after and immunoblotted with the indicated antibodies. β-actin is shown as the loading control. The molecular mass markers are shown on the left in kDa.

B, The total RNA was isolated from pRetrosuper or Sam68sh 3T3-L1 cells transfected with siGFP or siSRSF1. The mRNA levels of Rps6kb1 and Rps6kb1-002 were assessed by RT-qPCR. The mRNA expression level of *Rps6kb1-002* was normalized to total *Rps6kb1*. Error bars represent ± standard deviation of the mean. ** $p < 0.01$.

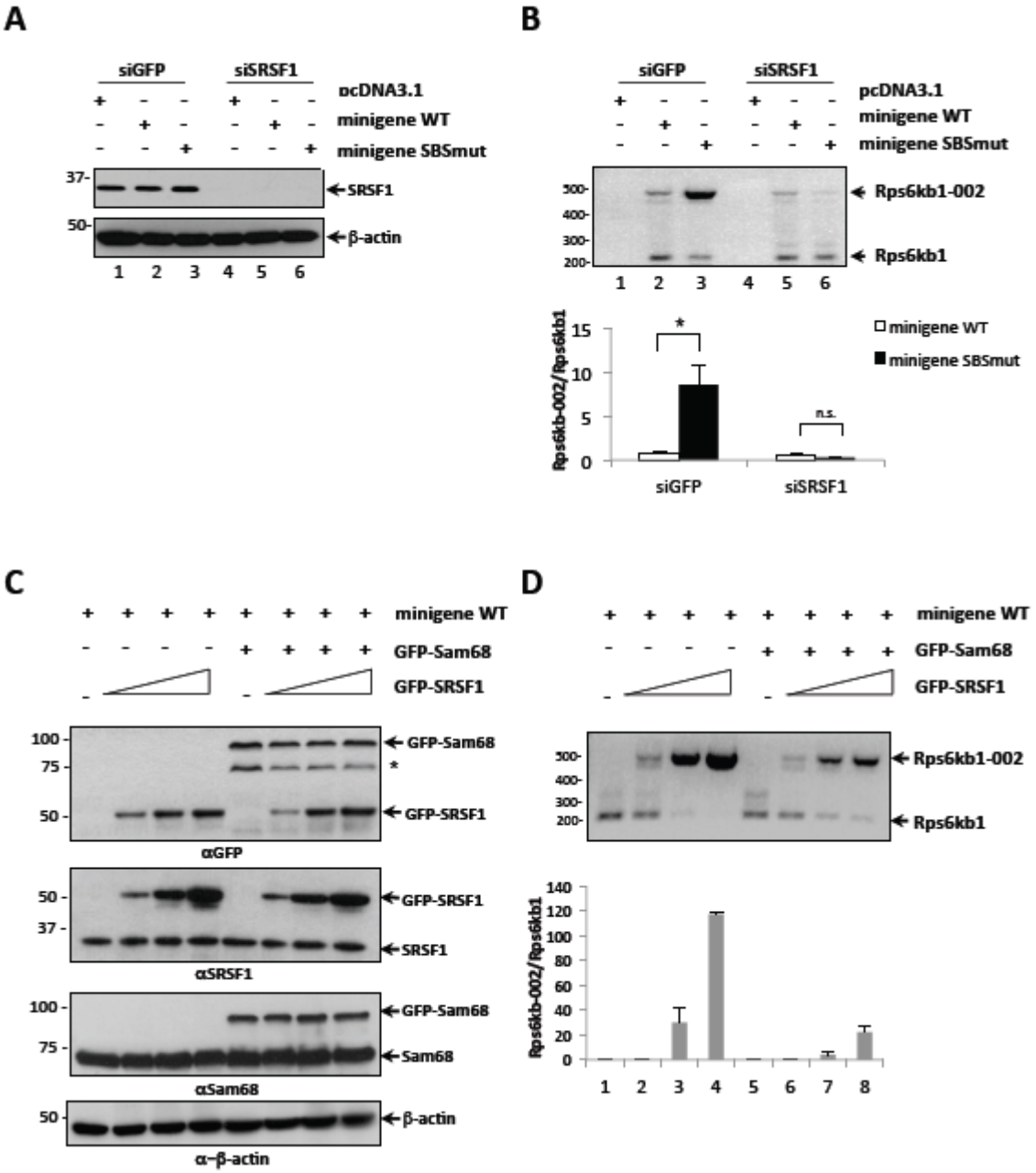


Figure 3.7 Sam68 competes with SRSF1 for the positive regulation of *Rps6kb1* splicing.

A, HEK293 cells were transfected with siRNAs targeting either GFP or human SRSF1. After 24 h, cells were transfected either with pcDNA3.1, *Rps6kb1* minigene or *Rps6kb1* minigene SBSmut plasmid. The cells were harvested after 48 h. Protein extracts were immunoblotted with the indicated antibodies. β-actin is shown as the loading control and * denotes an unknown protein. The molecular mass markers are shown on the left in kDa. ***B,*** Total RNA was isolated, treated with RQ1 DNase, and the mRNA levels of Rps6kb1 and Rps6kb1-002 were assessed by RT-PCR using the primer pairs indicated in panel ***A.*** Densitometry analysis was performed from 2 independent experiments and fold inclusion of *Rps6kb1-002* was normalized to the level of *Rps6kb1*. Error bars represent ± standard deviation of the mean. * $p < 0.05$. n.s., denotes not significant.

C, HEK293 cells were co-transfected with either *Rps6kb1* minigene plasmid alone, or with

GFP-Sam68, or with increasing amount of GFP-SRSF1. The cells were harvested after 48 h. The protein extracts were immunoblotted with the indicated antibodies. β-actin is shown as the loading control.
D, The total RNA was isolated, treated with RQ1 DNase, and the mRNA levels of *Rps6kb1* and *Rps6kb1-002* were assessed by RT-PCR. Densitometry analysis was performed from 2 independent experiments and fold inclusion of *Rps6kb1-002* was normalized to the level of *Rps6kb1*. Error bars represent ± standard deviation of the mean.

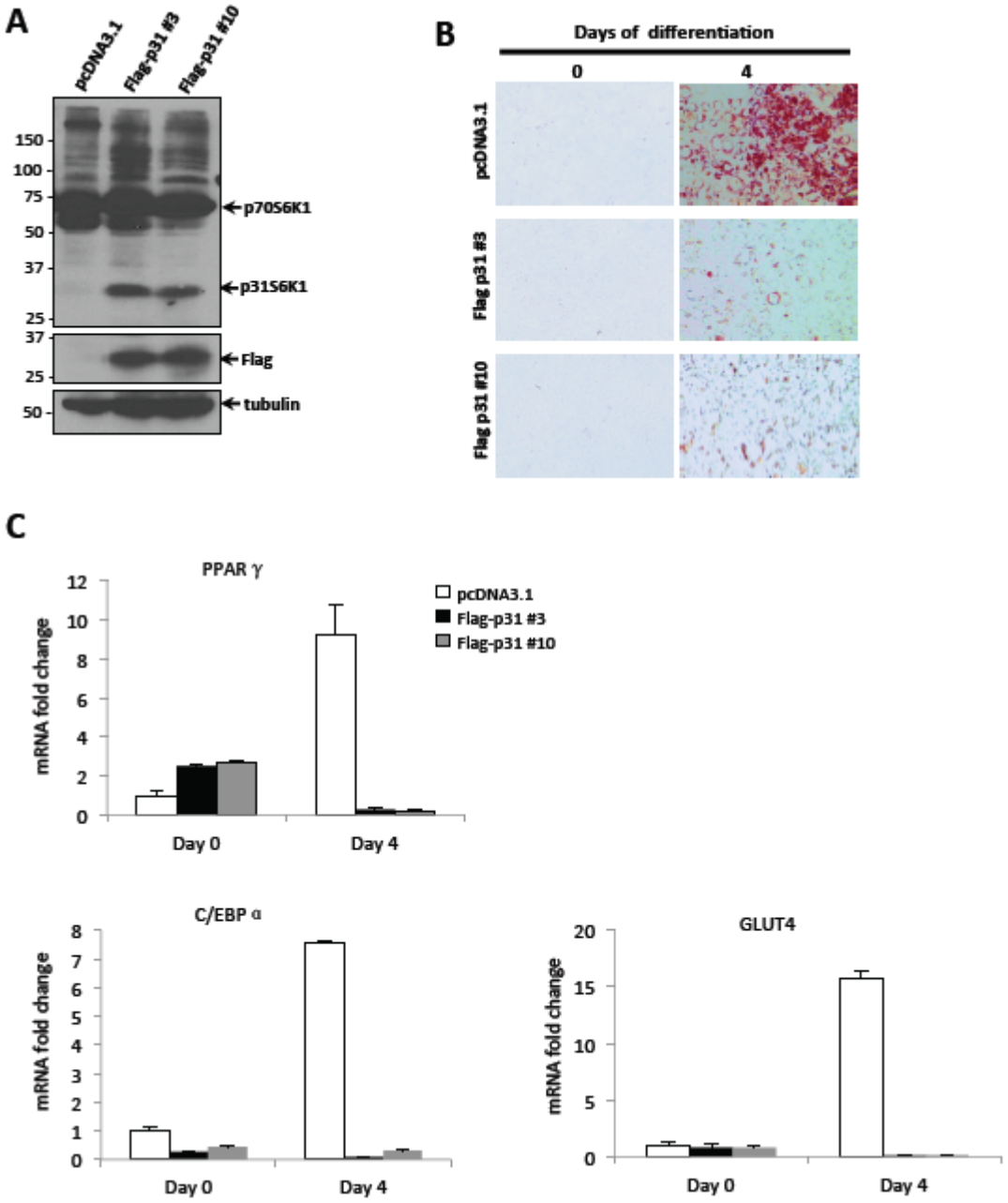


Figure 3.8 Ectopic expression of p31S6K1 suppresses adipogenesis.
A, 3T3-L1 cells were stably transfected with pcDNA3.1 or FLAG-p31S6K1. A polyclonal population of pcDNA3.1 and two individual clones (FLAG-p31 #3 and FLAG-p31 #10) were selected for analysis by immunoblotting using indicated antibodies. β-Tubulin is shown as the loading control. The molecular mass markers are shown on the left in kDa.
B, The cells indicated in panel ***A*** were induced to differentiate for 4 days. Adipocyte differentiation was assessed by Oil-Red-O staining.
C, The mRNA levels of PPARγ, C/EBPα and GLUT4 normalized to 18S rRNA were assessed by RT-qPCR. The data are expressed as relative values from differentiation day 0 and 4. Error bars represent ± standard deviation of the mean.

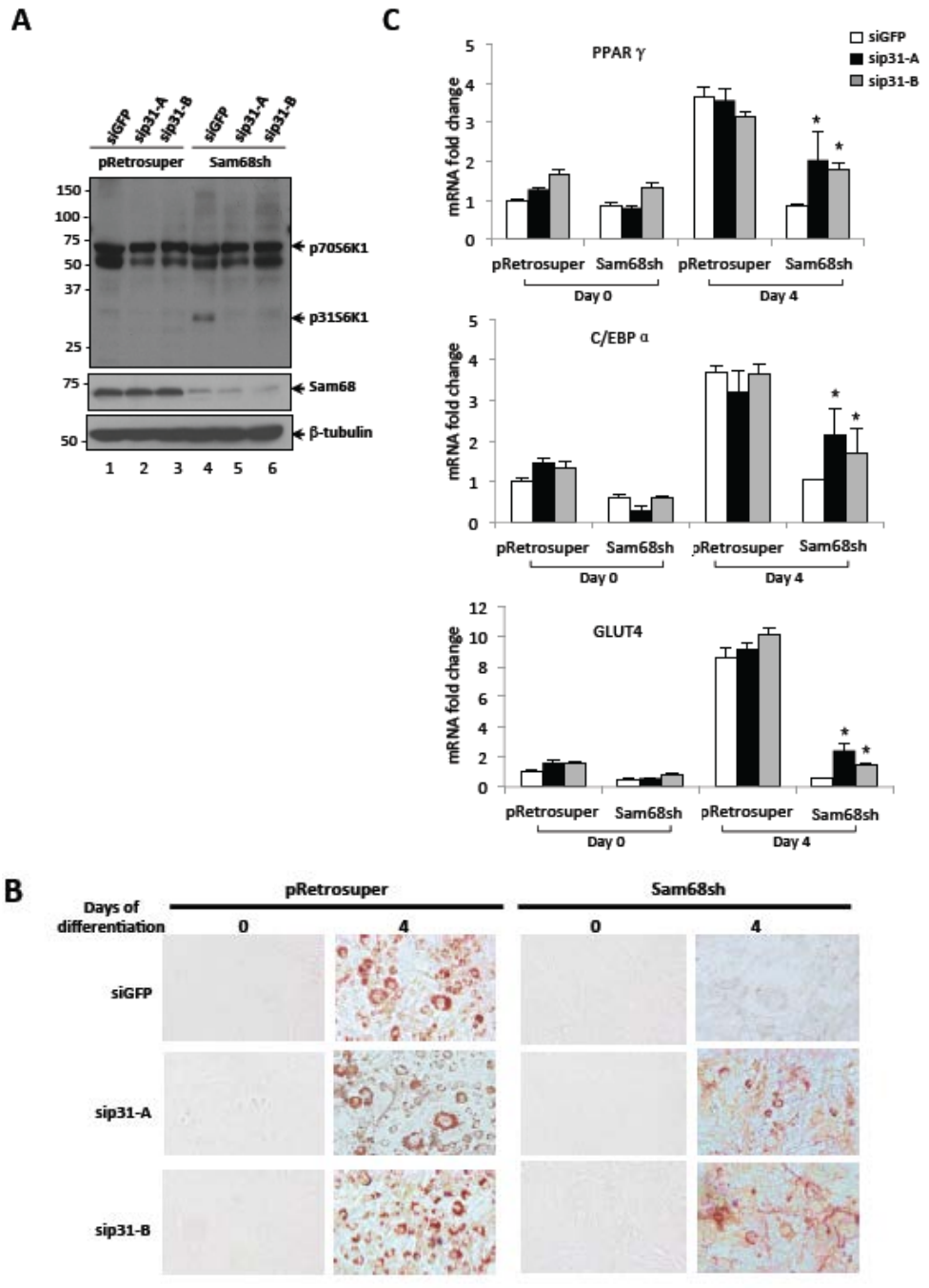


Figure 3.9 The expression of p31S6K1 contributes to the adipogenesis defects of Sam68-deficient mouse pre-adipocytes.

A, pRetrosuper or Sam68sh 3T3-L1 cells were transfected with siGFP (control), sip31-A and sip31-B. The protein extracts were immunoblotted with the indicated antibodies. β-Tubulin is shown as the loading control.

B, The cells indicated in panel ***A*** were induced to differentiate for 4 days. Adipocyte differentiation was assessed by Oil-Red-O staining.

C, The mRNA levels of PPARγ, C/EBPα and GLUT4 normalized to GAPDH were assessed by RT-qPCR. The data are expressed as relative values from differentiation day 0 and 4. Error bars represent ± standard deviation of the mean. * $p < 0.05$.

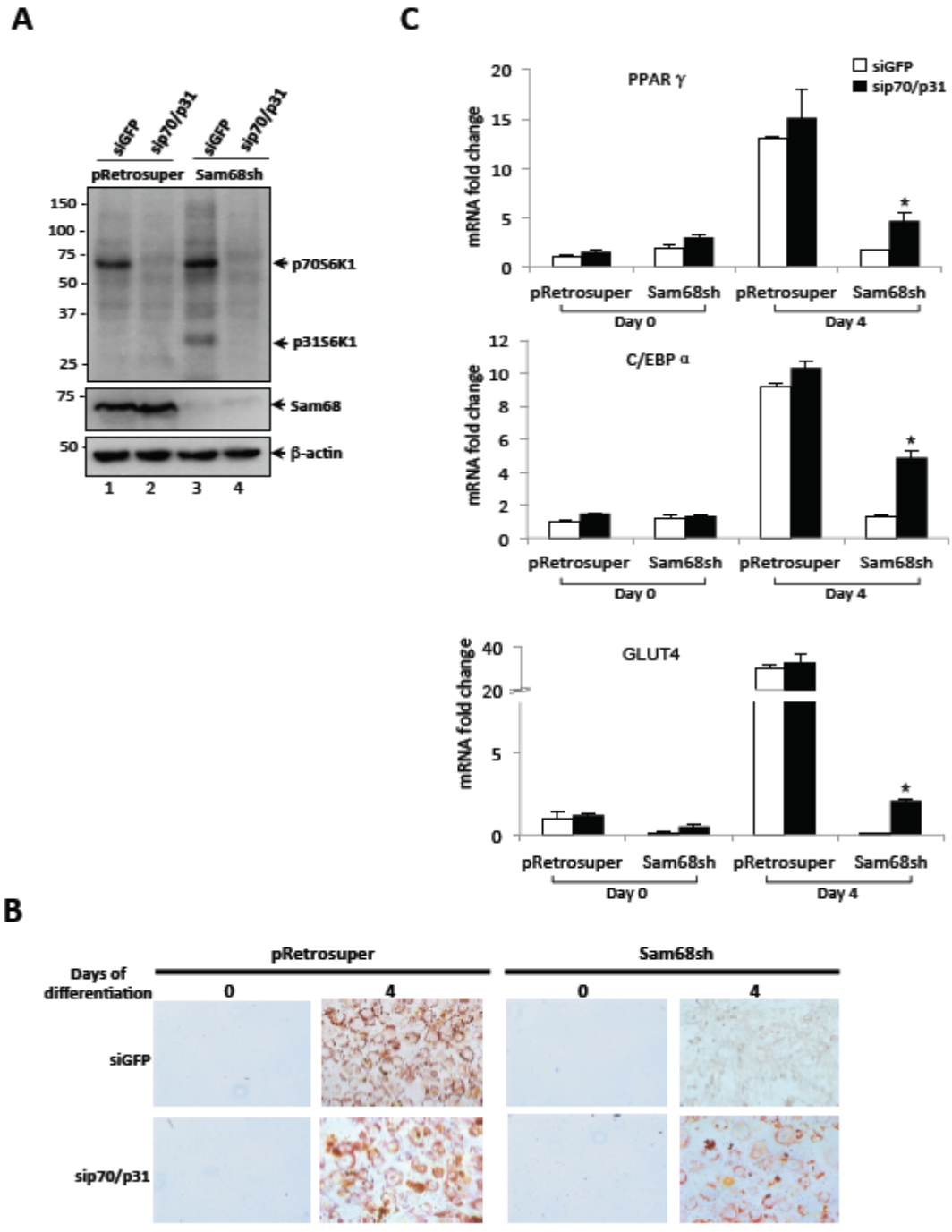


Figure 3.10 p31S6K1 contributes to the adipogenesis defects independently of p70S6K1 isoform.

A, pRetrosuper or Sam68sh 3T3-L1 cells were transfected with siGFP (control) and sip70/p31. The protein extracts were immunoblotted with the indicated antibodies. β-actin is shown as the loading control. The molecular mass markers are shown on the left in kDa.

B, The cells indicated in panel ***A*** were induced to differentiate for 4 days. Adipocyte differentiation was assessed by Oil-Red-O staining.

C, The mRNA levels of PPARγ, C/EBPα and GLUT4 normalized to 18S rRNA were assessed by RT-qPCR. The data are expressed as relative values from differentiation day 0 and 4. Error bars represent ± standard deviation of the mean. * $p < 0.05$.

3.6 Discussion

Alternative splicing leads to the generation of key isoforms required for cellular differentiation and proliferation (Fu and Ares 2014). In the present manuscript, we report that the Sam68 RNA-binding protein exerts a suppressive effect on the alternative splicing of *ribosomal S6 kinase* (*Rps6kb1*) in adipocytes. Consequently, Sam68-depleted 3T3-L1 pre-adipocytes harbor elevated levels of the short isoform 2 of S6K1 (*Rps6kb1-002*) and its encoded protein p31S6K1. Mechanistically, Sam68 binds an RNA element (Sam68 binding site, SBS) in intron 6 near the 5' splice site and prevents the usage of alternative exons 6a, 6b and 6c that generate *Rps6kb1-002* by counteracting the positive effects of the serine/arginine-rich splicing factor 1 (SRSF1) that binds within exon 6. The ectopic expression of p31S6K1 in wild type pre-adipocytes inhibited adipogenesis and the depletion of p31S6K1 using two separate siRNAs, partially restored the adipogenesis defects in Sam68-deficient pre-adipocytes. These findings demonstrate that the expression of Sam68 in adipocytes is required to prevent the expression of the short isoform 2 of S6K1, a potent suppressor of adipogenesis.

We identified an A/U-rich intronic sequence bound by Sam68, 46 nucleotides downstream of the 5' splice site of *Rps6kb1* exon 6. By associating with this intronic Sam68 binding site, Sam68 promotes the skipping of *Rps6kb1* exons 6a, 6b and 6c located in intron 6, thus preventing the expression of *Rps6kb1-002*. Sam68 is an established regulator of alternative splicing and it is known to function by directly associating with A/U-rich elements near 5' splice sites (Matter, Herrlich et al. 2002; Chawla, Lin et al. 2009; Pedrotti, Bielli et al. 2010; Huot, Vogel et al. 2012). The *Rps6kb1* intron 6 UAAUUAAA bipartite sequence is recognized by Sam68 with relative high affinity. Mutation of the SBS to UAAUUCCA diminished Sam68 association with this RNA sequence. By performing CLIP assays, we confirmed that Sam68 localizes directly at the SBS *in vivo,* as anti-Sam68 immunopreciptitations enriched the SBS RNA sequence 25-fold over control immunoprecipitations (Figure 3.4B). The presence of Sam68 at its SBS in the *Rps6kb1* minigene assay suppressed the effects of SRSF1 on the production of *Rps6kb1-002.* SRSF1 is a known positive regulator of *Rps6kb1-002* (Karni, de Stanchina et al. 2007), but its *Rps6kb1*

binding site(s) and how it regulates the production of *Rps6kb1*-002 remains unknown. We show that SRSF1 displayed reduced binding to its *Rps6kb1* exon 6 GAAAGAGAGGGAA site in the presence of Sam68 by CLIP assays. Additionally, we show that Sam68 regulates the binding of SRSF1 to its RNA binding site by directly interacting with SRSF1. In addition, several possible mechanisms have been proposed to explain how RNA-binding proteins suppress neighboring SR proteins (Witten and Ule 2011; Das and Krainer 2014; Fu and Ares 2014). Sam68 could compete directly with SRSF1 for the splicing machinery for intron 6 definition. Sam68 may also alter the neighboring RNA secondary structure and/or the rate of transcription, as proposed by Muchardt and coworkers (Batsche, Yaniv et al. 2006). The rate of transcription may influence the binding of SRSF1, thus influencing exon selection (Witten and Ule 2011; Das and Krainer 2014; Fu and Ares 2014).

We show that *Rps6kb1-002* and its encoded protein, p31S6K1, are present in Sam68-depleted pre-adipocytes and mature 3T3-L1 cells, as well as in mouse white adipose tissue of Sam68 null mice. Sam68 protein expression increases ~3-fold during the differentiation of 3T3-L1 pre-adipocytes to mature adipocytes (days 2 to 8) (Huot, Vogel et al. 2012), which is consistent with its role in preventing the expression of *Rps6kb1-002*. Using an anti-S6K1 antibody that recognizes all the isoforms sharing the common N-terminus, we observed that Sam68 deficiency leads to increased p31S6K1 expression, without apparent reduction in p70S6K1 and p85S6K1 expression. The latter is probably due to the fact that p70/p85 S6K1 are considerably more abundant than p31S6K1 (Karni, de Stanchina et al. 2007). The cellular role of p31S6K1 is unknown, however, it does have oncogenic properties. The expression of p31S6K1 is sufficient to induce transformation in NIH 3T3 cells (Karni, de Stanchina et al. 2007). Unlike mice that only express one short isoform, humans generate two short isoforms (h6A and h6C) of S6K1 and their expression is elevated in breast cancer cell lines (Ben-Hur, Denichenko et al. 2013). Depletion of these isoforms in breast cancer cell lines decreases their proliferation (Ben-Hur, Denichenko et al. 2013). These short isoforms lack kinase activity because their kinase domains are truncated, however, they retain the ability to bind mTORC1, as they contain the Raptor binding motif (Schalm and Blenis 2002). p31S6K1 has been shown to associate with mTOR and increase its activity (Ben-Hur,

Denichenko et al. 2013), however, this activity in Sam68-deficient pre-adipocytes is difficult to assess, since these cells have reduced mTOR levels (Huot, Vogel et al. 2012). Indeed Sam68-deficient pre-adipocytes exhibit decreased phosphorylation of rpS6 and AKT during adipogenesis (Huot, Vogel et al. 2012). p31S6K1 has been shown to also be nuclear, unlike p85S6K1 and p70S6K1 (Rosner and Hengstschläger 2011), therefore, it may also fulfill other functions.

S6K1$^{-/-}$ mice have decreased adipose tissue mass, increased energy expenditure, and are resistant to dietary-induced obesity (Um, Frigerio et al. 2004). S6K1 participates in the up-regulation of transcription factors during the commitment phase of adipogenesis (Carnevalli, Masuda et al. 2010). Adipocytes normally express p70/p85S6K1, but not p31S6K1 (Figure 3.3). The fact that p31S6K1 blocks differentiation, suggests that it has independent properties with p70/p85S6K1.

Sam68 likely regulates many alternative spliced events that contribute to the observed lean phenotype of Sam68-deficient mice (Huot, Vogel et al. 2012). Sam68 regulates the splicing of Bcl-$_X$ (Paronetto, Achsel et al. 2007), as well as mTOR, tripeptidyl peptidase II (Tpp2) and Tubby (Tub) (Huot, Vogel et al. 2012). Sam68 has also been shown to regulate the alternative splicing of the *Srsf1* transcript in colon cancer cells to influence epithelial-to-mesenchymal transition (Valacca, Bonomi et al. 2010). Indeed, we also detected an increase in *Srsf1* transcript in 3T3-L1 cells (data not shown), but this did not affect SRSF1 protein levels. Depletion of both p70S6K1 and p31S6K1 rescued the adipogenesis of Sam68 deficient cells to a similar extent, indicating that the negative effects of p31S6K1 in adipogenesis are independent of p70S6K1. Therefore, p31S6K1 is yet another contributor to the Sam68-deficiency induced adipogenesis defects observed.

In conclusion, we report that the Sam68 regulates alternative splicing of multiple mediators in mTOR signaling. We show that the mouse *Rps6kb1* gene is a Sam68 RNA spliced target, as it bound an intron 6 RNA element to regulate alternative exon usage by antagonizing the effect of SRSF1. We also show that the short isoform of *Rps6kb1*, namely p31S6K1, is a potent repressor of adipogenesis and its presence in Sam68 deficient pre-adipocytes contributes to the adipogenesis defects observed in these cells.

3.7 Acknowledgments

We thank Zhenbao Yu and Gillian Vogel for helpful discussions and for critically reading the manuscript. The work was supported by a grant from the Canadian Institute of Health Canada (MOP-123531) to S. R.

3.8 References

1. Bielli P, Busà R, Paronetto MP, Sette C. 2011. The RNA-binding protein Sam68 is a multifunctional player in human cancer. Endocr Relat Cancer 18:R91-102.
2. Richard S. 2010. Reaching for the stars: Linking RNA-binding proteins to diseases. Adv Exp Med Biol 693:142-157.
3. Galarneau A, Richard S. 2009. The STAR RNA-binding proteins GLD-1, QKI, SAM68 and SLM-2 bind bipartite RNA motifs. BMC Mol Biol 10:47.
4. Lin Q, Taylor SJ, Shalloway D. 1997. Specificity and determinants of Sam68 RNA binding. Implications for the biological function of K homology domains. J Biol Chem 272:27274-27280.
5. Matter N, Herrlich P, Konig H. 2002. Signal-dependent regulation of splicing via phosphorylation of Sam68. Nature 420:691-695.
6. Cappellari M, Bielli P, Paronetto MP, Ciccosanti F, Fimia GM, Saarikettu J, Silvennoinen O, Sette C. 2014. The transcriptional co-activator SND1 is a novel regulator of alternative splicing in prostate cancer cells. Oncogene 33:3794-3802.
7. Bielli P, Busà R, Di Stasi SM, Munoz MJ, Botti F, Kornblihtt AR, Sette C. 2014. The transcription factor FBI-1 inhibits SAM68-mediated BCL-X alternative splicing and apoptosis. EMBO Rep 15:419-427.
8. Paronetto MP, Achsel T, Massiello A, Chalfant CE, Sette C. 2007. The RNA-binding protein Sam68 modulates the alternative splicing of Bcl-x. J. Cell Biol. 176:929-939.
9. Valacca C, Bonomi S, Buratti E, Pedrotti S, Baralle FE, Sette C, Ghigna C, Biamonti G. 2010. Sam68 regulates EMT through alternative splicing-activated nonsense-mediated mRNA decay of the SF2/ASF proto-oncogene. J Cell Biol 191:87-99.
10. Chawla G, Lin CH, Han A, Shiue L, Ares MJ, Black DL. 2009. Sam68 regulates a set of alternatively spliced exons during neurogenesis. Mol Cell Biol 29:201-213.
11. Iijima T, Wu K, Witte H, Hanno-Iijima Y, Glatter T, Richard S, Scheiffele P. 2011. SAM68 Regulates Neuronal Activity-Dependent Alternative Splicing of Neurexin-1. Cell 147:1601-1614.
12. Pedrotti S, Bielli P, Paronetto MP, Ciccosanti F, Fimia GM, Stamm S, Manley JL, Sette C. 2010. The splicing regulator Sam68 binds to a novel exonic splicing silencer and functions in SMN2 alternative splicing in spinal muscular atrophy. EMBO J. 29:1235-1247.
13. Sellier C, Rau F, Liu Y, Tassone F, Hukema RK, Gattoni R, Schneider A, Richard S,

Willemsen R, Elliott DJ, Hagerman PJ, Charlet-Berguerand N. 2010. Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients. EMBO J. 29:1248-1261.
14. Paronetto MP, Messina V, Barchi M, Geremia R, Richard S, Sette C. 2011. Sam68 marks the transcriptionally active stages of spermatogenesis and modulates alternative splicing in male germ cells. Nucleic acids research Feb 25:(Swiercz, Person et al.).
15. Paronetto MP, Messina V, Bianchi E, Barchi M, Vogel G, Moretti C, Palombi F, Stefanini M, Geremia R, Richard S, Sette C. 2009. Sam68 regulates translation of target mRNAs in male germ cells, necessary for mouse spermatogenesis. J Cell Biol 185:235-249.
16. Huot ME, Vogel G, Zabarauskas A, Ngo CT, Coulombe-Huntington J, Majewski J, Richard S. 2012. The Sam68 STAR RNA-binding protein regulates mTOR alternative splicing during adipogenesis. Mol Cell 46:187-199.
17. Laplante M, Sabatini DM. 2012. mTOR signaling in growth control and disease. Cell 149:274-293.
18. Dann SG, Selvaraj A, Thomas G. 2007. mTOR Complex1-S6K1 signaling: at the crossroads of obesity, diabetes and cancer. Trends Mol Med 13:252-259.
19. Shimobayashi M, Hall M. 2014. Making new contacts: the mTOR network in metabolism and signalling crosstalk. Nat Rev Mol Cell Biol 15:155-162.
20. Inoki K, Zhu T, Guan KL. 2003. TSC2 mediates cellular energy response to control cell growth and survival. Cell 115:577-590.
21. Saucedo LJ, Gao X, Chiarelli DA, Li L, Pan D, Edgar BA. 2003. Rheb promotes cell growth as a component of the insulin/TOR signalling network. Nat Cell Biol 5:566-571.
22. Tee AR, Manning BD, Roux PP, Cantley LC, Blenis J. 2003. Tuberous sclerosis complex gene products, Tuberin and Hamartin, control mTOR signaling by acting as a GTPase-activating protein complex toward Rheb. Curr Biol 13:1259-1268.
23. Wang L, Harris TE, Roth RA, Lawrence JCJ. 2007. PRAS40 regulates mTORC1 kinase activity by functioning as a direct inhibitor of substrate binding. J Biol Chem 282:20036-20044.
24. Vander Haar E, Lee SI, Bandhakavi S, Griffin TJ, Kim DH. 2007. Insulin signalling to mTOR mediated by the Akt/PKB substrate PRAS40. Nat Cell Biol 9:316-323.
25. Thedieck K, Polak P, Kim ML, Molle KD, Cohen A, Jenö P, Arrieumerlou C, Hall M. 2007. PRAS40 and PRR5-like protein are new mTOR interactors that regulate apoptosis. PLoS One 2:e1217.
26. Sonenberg N, Hinnebusch AG. 2009. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. 136:731-745.
27. Chen T, Boisvert FM, Bazett-Jones DP, Richard S. 1999. A role for the GSG domain in localizing Sam68 to novel nuclear structures in cancer cell lines. Mol Biol Cell 10:3015-3033.
28. Sun Y, Ma YC, Huang J, Chen KY, McGarrigle DK, Huang XY. 2005. Requirement of SRC-family tyrosine kinases in fat accumulation. Biochemistry 44:14455-14462.
29. Huppertz I, Attig J, D'Ambrogio A, Easton LE, Sibley CR, Sugimoto Y, Tajnik M, König J, Ule J. 2014. iCLIP: protein-RNA interactions at nucleotide resolution. Methods 65:274-287.
30. Karni R, de Stanchina E, Lowe SW, Sinha R, Mu D, Krainer AR. 2007. The gene

encoding the splicing factor SF2/ASF is a proto-oncogene. Nat Struct Mol Biol. 14:185-193.
31. Huot ME, Vogel G, Richard S. 2009. Identification of a Sam68 ribonucleoprotein complex regulated by epidermal growth factor. J Biol Chem (Swiercz, Person et al.).
32. Carnevalli LS, Masuda K, Frigerio F, Le Bacquer O, Um SH, Gandin V, Topisirovic I, Sonenberg N, Thomas G, Kozma SC. 2010. S6K1 plays a critical role in early adipocyte differentiation. Dev Cell 18:763-774.
33. Fu XD, Ares MJ. 2014. Context-dependent control of alternative splicing by RNA-binding proteins. Nat Rev Cancer 15:689-701.
34. Witten JT, Ule J. 2011. Understanding splicing regulation through RNA splicing maps. Trends Genet 27:89-97.
35. Das S, Krainer AR. 2014. Emerging functions of SRSF1, splicing factor and oncoprotein, in RNA metabolism and cancer. Mol Cancer Res 12:1195-1204.
36. Batsche E, Yaniv M, Muchardt C. 2006. The human SWI/SNF subunit Brm is a regulator of alternative splicing. Nat Struct Mol Biol 13:22-29.
37. Ben-Hur V, Denichenko P, Siegfried Z, Maimon A, Krainer A, Davidson B, Karni R. 2013. S6K1 alternative splicing modulates its oncogenic activity and regulates mTORC1. Cell Rep 3:103-115.
38. Schalm SS, Blenis J. 2002. Identification of a conserved motif required for mTOR signaling. Curr Biol 12:632-639.
39. Rosner M, Hengstschläger M. 2011. Nucleocytoplasmic localization of p70 S6K1, but not of its isoforms p85 and p31, is regulated by TSC2/mTOR. Oncogene 30:4509-4522.
40. Um SH, Frigerio F, Watanabe M, Picard F, Joaquin M, Sticker M, Fumagalli S, Allegrini PR, Kozma SC, Auwerx J, Thomas G. 2004. Absence of S6K1 protects against age- and diet-induced obesity while enhancing insulin sensitivity. Nature 431:200-205.

Chapter 4

Aven recognition of RNA G-quadruplexes regulates translation of the Mixed Lineage Leukemia proto-oncogenes

4.1 Preface

As another RGG/RG motif-containing protein discovered by previous genome-wide searches, Aven is an anti-apoptotic protein and is known to be overexpressed in acute leukemias. However, the role of RGG/RG motif of Aven has not been characterized. Herein we identified that Aven plays a role in cell survival and leukemogenesis, via the association between its RGG/RG motif and the G-quadruplexes of specific mRNAs, including MLL1 and MLL4.

4.2 Abstract

G-quadruplexes (G4) are extremely stable secondary structures forming stacks of guanine tetrads. DNA G4 structures have been extensively studied, however, less is known about G4 motifs in mRNAs, especially in their coding sequences. Herein, we show that Aven stimulates the mRNA translation of the mixed lineage leukemia (*MLL*) proto-oncogene in an arginine methylation-dependent manner. The Aven RGG/RG motif bound G4 structures within the coding regions of the MLL1 and MLL4 mRNAs increasing their polysomal association and translation, resulting in the induction of transcription of leukemic genes. The DHX36 RNA helicase associated with the Aven complex and was required for optimal translation of G4 mRNAs. Depletion of Aven led to a decrease in synthesis of MLL1 and MLL4 proteins resulting in reduced proliferation of leukemic cells. These findings identify an Aven-centered complex that stimulates the translation of G4 harboring mRNAs, thereby promoting survival of leukemic cells.

4.3 Introduction

RNA-binding proteins (RBPs) coordinate many steps of RNA metabolism ranging from splicing, RNA processing, RNA transport, mRNA translation and RNA degradation (Glisovic, Bachorik et al. 2008). RBPs associate with specific RNA motifs and/or secondary structures within coding, untranslated regions and non-coding RNAs in functional units called ribonucleoprotein (RNP) complexes (Mitchell and Parker 2014). Defects in RBPs have been associated with many complex diseases ranging from neurological disorders to cancer (Lukong, Chang et al. 2008; Cooper, Wan et al. 2009; Castello, Fischer et al. 2013; Ramaswami, Taylor et al. 2013).

RBPs are predominantly defined by the presence of RNA binding domains within their sequences (Chen and Varani 2013). Recently, several 'interactome capture' strategies have been performed to identify RBPs genome-wide. In addition to identifying the known RBPs, these approaches have identified numerous mammalian proteins that do not possess a canonical RNA binding domain (Baltz, Munschauer et al. 2012; Castello, Fischer et al. 2012; Kwon, Yi et al. 2013). Interestingly, RBPs that harbor repeated sequences including YGG and RGG motifs were identified (Castello, Fischer et al. 2012). The RGG/RG motif is enriched in proteins associated with RNA and is a known RNA binding interface (reviewed in (Thandapani, O'Connor et al. 2013)). The RGG/RG motif, also called RGG box, was shown to be an RNA recognition motif (Kiledjian and Dreyfuss 1992). Subsequently, the RGG/RG motifs of Nucleolin, FMRP, FUS and EWS were also shown to bind guanine-rich sequences that are potential G-quadruplexes (Darnell, Jensen et al. 2001; Takahama, Kino et al. 2011; Takahama, Takada et al. 2013; Haeusler, Donnelly et al. 2014). RGG/RG motifs also mediate protein-protein interactions. Notably, the RGG/RG motif of yeast Scd6 mediates interactions with eIF-4G which leads to stress granule formation and inhibition of cap-dependent translation (Rajyaguru, She et al. 2012). Despite these recent advances, little is known about the role of RGG/RG motifs that bind both RNA and proteins.

G-quadruplexes (G4) are planar structures of stacks of guanine tetrads stabilized by monovalent potassium or sodium ions. G-quadruplexes have been shown to regulate DNA replication, DNA repair, gene expression and telomeres (Bates, Mergny et al. 2007). Less is

known about G4 structures found in RNA. There are > 1500 potential G4s (PG4s) in 5′-UTR of mRNAs alone (Beaudoin and Perreault 2010), but not all PG4s form stable G-quadruplexes, which are influenced by the numbers of G-quartets, the possibility of bulge formation, the length of the loops and the presence of alternative Watson-Crick base pair-based stable structure (Burge, Parkinson et al. 2006; Mukundan and Phan 2013; Jodoin, Bauer et al. 2014). Some PG4s in the 5′-UTR of mRNAs form *bona fide* G-quadruplexes and inhibit cap-dependent translation (Kumari, Bugaut et al. 2007; Beaudoin and Perreault 2010; Bugaut and Balasubramanian 2012). Recently, inhibitors of DEAD box RNA helicase eIF4A, or eIF4A1 depletion have been shown to selectively inhibit translation of mRNAs with G-quadruplexes in their 5'UTR (Wolfe, Singh et al. 2014; Modelska, Turro et al. 2015). However, presence of G-quadruplexes in 5'UTRs does not appear to be sufficient to render translation of mRNAs sensitive to changes in eIF4A activity (Rubio, Weisburd et al. 2014). In addition to the incomplete understanding of the role of 5'UTR G-quadruplexes in translation control, little is known about how G4 structures in open reading frames (ORFs) affect translation.

Arginine residues within RGG/RG motifs are preferred substrates for methylation by protein arginine methyltransferases (PRMTs) (Thandapani, O'Connor et al. 2013). Arginine methylation is known to regulate many cellular processes including signal transduction, transcription, pre-mRNA splicing, and DNA repair (Bedford and Richard 2005; Bedford and Clarke 2009; Xu, Wang et al. 2013). PRMT1 generates > 85% of asymmetric dimethylarginines found in cells with preference for RGG/RG motif containing proteins (Bedford and Clarke 2009). PRMT1 is known for its nuclear roles in regulating gene expression and DNA damage (Strahl, Briggs et al. 2001; Wang, Huang et al. 2001; An, Kim et al. 2004; Boisvert, Déry et al. 2005), however less is known about its cytoplasmic roles. PRMT1-deficient mice die at E6.5 and the absolute removal of PRMT1 in mouse embryo fibroblasts (MEFs) leads to cell death (Pawlak, Scherer et al. 2000; Yu, Chen et al. 2009). To identify other biological processes regulated by arginine methylation, we performed a bioinformatics approach to identify proteins harboring RGG/RG motifs and one such protein we identified was Aven (Thandapani, O'Connor et al. 2013).

Aven is a predominantly cytoplasmic protein required for cell survival and it has been shown to function as an apoptotic inhibitor by interaction with and stabilizing the pro-survival protein Bcl-x_L, as well as inhibiting the function of Apaf-1 (Chau, Cheng et al. 2000). It was proposed that the proteolytic cleavage of Aven by Cathepsin D is required for its anti-apoptotic activity (Melzer, Fernández et al. 2012). Furthermore, Aven is required for ataxia telangiectasia mutated (ATM) activation in *Xenopus* oocytes and HeLa cells (Guo, Yamada et al. 2008) and ataxia telangiectasia related (ATR) activation following DNA damage in osteosarcoma cells (Baranski, Booij et al. 2015). High Aven expression correlates with poor survival in metastatic patients with osteosarcomas (Baranski, Booij et al. 2015). The elevated Aven expression is also frequently observed in acute myeloid leukemia (AML) and acute lymphoblastic leukemia (T-ALL) and is associated with poor prognosis (Paydas, Tanriverdi et al. 2003; Choi, Hwang et al. 2006). A transgenic mouse model with T cell-specific overexpression of Aven showed that its expression enhanced T-cell lymphomagenesis in the absence of p53 (Eismann, Melzer et al. 2013). The mechanism by which Aven promotes hematological malignancies is yet to be understood.

Herein, we report that the methylation of the RGG/RG motif of Aven functions in the translational control of mRNAs harboring G4 structures in their ORFs. The association of Aven with polysomes was dependent on the arginine methylation of its RGG/RG motif and on the methyl-dependent interactions with the Tudor domains of SMN and TDRD3, previously shown to be associated with polysomes (Goulet, Boisvenue et al. 2008; Sanchez, Dury et al. 2013). We identify Aven to be an RNA-binding protein, as its RGG/RG motif bound G4 motifs in the ORFs of mRNAs encoding the mixed lineage leukemia (MLL) family proteins MLL1 and MLL4. The RGG/RG motif of Aven also associated with the G4 RNA helicase, DHX36, and this helicase was required for optimal translation of Aven regulated mRNAs. Furthermore, Aven deficient T-ALL cell lines had reduced MLL1 and MLL4 protein levels, but not mRNA levels, which were paralleled by proliferation defects. These findings define a hitherto unknown mechanism of action for arginine methylation in regulating translation of a subset of mRNAs including those encoding pivotal leukemogenic transcriptional regulators MLL1 and MLL4.

4.4 Materials and Methods

4.4.1 Cells, Reagents and Antibodies

HEK293T, U2OS, MOLT-4 and CCRF-CEM were from the American Type Culture Collection (Manassas, VA). *PRMT1*$^{FL/-;CreERT}$ mouse embryo fibroblasts (MEFs) were described previously (Yu, Chen et al. 2009). Protein A-Sepharose, 4-hydroxytamoxifen (OHT), anti-FLAG (M2) antibody-coupled agarose beads, mouse anti-FLAG (M2), anti-Myc, anti-MLL4 (WH0009757M2) and anti-α-tubulin were purchased from Sigma-Aldrich (St. Louis, MO). Mouse anti-Aven (ab77014, Abcam) was used for immunoprecipitations and rabbit anti-Aven (ProSci 2413, ProScience) was used for immunoblotting. Mouse anti-rpS6 was from Santa Cruz Biotechnology (Santa Cruz, CA). Rabbit anti-PRMT1 and ASYM25b antibodies were purchased from Millipore (Billerica, MA). Anti-SMN was from Transduction Laboratories (Lexingtong, KY). Anti-TDRD3 was a kind gift from Mark Bedford (Smithville, TX). Anti-MLL1 antibodies (A300-086A) were from Bethyl Laboratories (Montgomery, TX) and anti-DHX36 (ab70269) was from Abcam.

4.4.2 DNA Constructs

The full-length human FLAG-Aven and FLAG-AvenΔRGG lacking amino acids (1-73) were cloned in pcDNA3.1. The full-length human Aven was also cloned in pcDNA3.1 with 5 epitope tags of Myc at the N-terminus between the *Bam*HI and *Xho*I sites. GST Aven RGG/RG including amino acids (1-73) was cloned in pGEX5x-1 plasmid. FLAG-AvenR-K was assembled using G-blocks purchased from IDT. Aven arginines R5, R8, R11, R13, R14, R17, R19, R24, R28, R37, R50, R51, R53, R55, R57, R60, R63 and R66 were replaced with lysines. pGL3-MLL4 and pGL3-MLL4 G4 mutant was generated by cloning MLL4 WT nucleotide sequence (262-318) (5´-CGG GTC CAG CGG GGC CGG GGA CGG GGT CGG GGC CGG GGC TGG GGC CCG AGT CGA GGC-3´) or (5´-GCG TCC AGC GCG CCC GCG CCC GCG CCC GCG CCC GCG CCT GGG CCC CGA GTC GAG GC-3´) in fusion with the pGL3 basic plasmid.

4.4.3 PAR-CLIP (photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation)

HEK293T cells expressing either pcDNA3.1, FLAG-Aven and FLAG-AvenΔRGG were incubated with 100 mM 4-thiouridine (4SU) for 8 h prior to crosslinking. The cells were washed with ice-cold PBS and irradiated with 0.15 J/cm^2 of 365nm UV light at 4°C. The cells were collected by centrifugation at 514 g for 1 min at 4°C. The cell pellets were resuspended in lysis buffer supplemented with protease inhibitors (Roche) and 0.5 U/ml RNasin (Promega) (Hafner, Landthaler et al. 2010). 10 µl of 1:250 dilution of RNase I (Life Technologies) and 2 µl Turbo DNase (Life Technologies) was added to the lysate while shaking at 37°C for 3 min. The lysates were then cleared and immunoprecipitated with 25µl anti-FLAG M2 affinity beads (Sigma). The beads were washed twice with high salt buffer, twice with the lysis buffer and incubated with proteinase K buffer (containing 1.2 mg/ml Proteinase K) for 20 min at 37°C. RNA then was isolated through TRIzol® reagent and subjected to RT-qPCR. For endogenous Aven, UV crosslinked lysates were processed as described above expect they were incubated with rabbit IgG or anti-Aven (Proscience).

The reverse primers were used for the reverse transcription reaction: For MLL1 G4 structure the RT reverse primer is 5´-GAG GAG GCT GCT GAG GCG GC-3´; for MLL1 negative control region (>300bp downstream of G4) the RT reverse primer is 5´-TCT TCT TGA TCT TAT CTC CA-3´; for MLL4 G4 structure the RT reverse primer is 5'- CTC TCC TCC TCC GGC ACG CAG C-3'; for MLL4 negative control region (>300bp downstream of G4) the RT reverse primer is 5´- GAT TGT CAC AGC TGC TTC TGC-3´ qPCR was performed with MLL1 (G4 structure, forward) 5´- GGC GGG AAG CAG CGG GGC TG-3´ and MLL1 (G4 structure, reverse) 5´-CTG AGG CGG CGG CCG CTC CC-3´; MLL1 (negative control, forward) 5´- CAT CTG TGT TTT CCC CTC TA -3´ and MLL1 (negative control, reverse) 5´-CTT ATC TCC AGA TTT GGT CT-3´; MLL4 (G4 structure, forward) 5´- GCG CCG GCT CCG CCG CCT GT-3´ and MLL4 (G4 structure, reverse) 5´- GCA CGC AGC CTC GAC TCG GG -3´; MLL4 (negative control, forward) 5´- TGC AGG AGG AAG CAG CAA GC-3´ and MLL4 (negative control, reverse) 5´- CTG CTT CTG CCA CCA CTA CT-3´.

4.4.4 Polysome Profiling

Polysome profiling has been performed as described in detail (Gandin, Sikström et al. 2014). Briefly, HEK293T cells in 150mm plates were transfected with the indicated expression plasmids using Lipofectamine 2000. Approximately 70% confluent cells were treated with 100 µg/ml cycloheximide for 5 min to 'freeze' mRNA translation. The cells were washed twice with ice cold-PBS and lysed in hypotonic lysis buffer containing 50 mM Tris-HCl (pH 7.5), 2.5 mM $MgCl_2$, 1.5 mM KCl, 100 µg/ml cycloheximide, 2 mM DTT, 0.5% Triton X-100, and 0.5% sodium deoxycholate. The lysates were spun at 13,000 rpm for 10 min at 4ºC and layered onto a 10% to 50% sucrose gradient as previously described (Gandin, Gutierrez et al. 2013). The gradients were formed using a SW40 rotor (Beckman) at 36,000 rpm for 2 h at 4°C. One ml fractions were collected by upward displacement with 60% sucrose and absorbance was continuously recorded at 254 nm using ISCO fractionator (Teledyne, ISCO). Collected fractions were precipitated with 10% TCA, separated by SDS-PAGE and proteins visualized by immunoblotting. For RNA analysis, 800 µl TRIZOL was added to the 1 ml fractions and RNA was isolated using standard procedures. Isolated RNA was quantified using RT-qPCR. The cDNA samples were serially diluted and the efficiency and *Cq* values were used to generate a standard curve (Piques, Schulze et al. 2009). One standard curve was generated for each primer pair. All standard curves had R^2 value higher than 0.99, with a slope between -3.58 and -3.10. Each data point for each fraction was plotted against the standard curve to calculate the percentage of input.

4.4.5 RNA Binding Assays

Biotinylated RNAs were purchased from IdT. The RNAs were dissolved in binding buffer (10 mM Tris-acetate, pH 7.7, 200 mM potassium acetate, 5 mM magnesium acetate), heated to 75°C for 10 min and allowed to renature at 21°C for 5 min (Phan, Kuryavyi et al. 2011). For the RNA binding assays 100 nM final concentration of biotinylated RNA was incubated with cellular lysates containing 2 mg/ml heparin for 1 h on ice. Then 25 µl of 50% Streptavidin agarose slurry was added and incubated at 4ºC for 30 min with constant end-over-end mixing.

The beads were then washed 2x with cell lysis buffer with increasing salt concentration and once with PBS. The samples were then boiled with 25 µl of 2x Laemmli buffer, resolved by SDS-PAGE and visualized by immunoblotting.

4.4.6 Peptide RNA binding assay

Fluorescein labelled RNA were dissolved in binding buffer (10 mM Tris-acetate, pH 7.7, 200 mM potassium acetate, 5 mM magnesium acetate), heated to 75°C for 10 min and allowed to renature at 21°C for 5 min (Phan, Kuryavyi et al. 2011). The biotinylated peptides (20 pmol in 50 µl ddH_20) were allowed to bind the streptavidin coated high capacity binding plates (Pierce # 15503) overnight at 4ºC or 2 h at room temperature. The peptides were removed and the plates washed four times with binding buffer. Different concentrations of the fluoresceinated RNA (IDT) was allowed to bind the peptides for 1 h at room temperature. The unbound RNAs were removed and the plates washed four times with binding buffer. Fluorescence was measured at 521 nm on a Synergy H4 instrument (BioTek). The peptides were purchased from Epicypher Inc and their sequence are: Aven50-74 biotin-RRGRGRGRGFRGARGGRGGGGAPRG, termed DiRGG; Aven50-74(Me2a) biotin-RRGRGRGRGFRGAR(Me2a)GGR(Me2a)GGGGAPRG, termed DiRGGme; Aven2-26: biotin- QAERGARGGRGRRPGRGRPGGDRHS, termed TriRG; Aven2-26(Me2a): biotin-QAER(Me2a)GAR(Me2a)GGR(Me2a)GRRPGR(Me2a)GRPGGDRHS, termed TriRGme.

4.4.7 RT-qPCR primers

Gene	Primer	Sequence (5'->3')
MLL1	Forward	GAGGACCCCGGATTAAACAT
	Reverse	GGAGCAAGAGGTTCAGCATC
MLL4	Forward	CAGACCCGGCAGACAGATGAG
	Reverse	AGATGTTACGTAGTCAAGGCACA
rpS6	Forward	AATGGAAGGGTTATGTGGTCCG
	Reverse	CCCCTTACTCAGTAGCAGGC

HOXA9	Forward	TCAAAAGGATAGCGCTGCCA
	Reverse	TGCATTACCAGAGAGCCGTG
MEIS1	Forward	ACCGTTTGCGACTTGGTACT
	Reverse	TGCTCACAACCAGACAGCTC
Actin	Forward	ACCACACCTTCTACAATGAGC
	Reverse	GATAGCACAGCCTGGATAGC
HOXA1	Forward	ACCAAGAAGCCTGTCGCTC
	Reverse	ACTTTCCCTGTTTTGGGAGGG
HOXC10	Forward	ACCACAGGAAATTGGCTGAC
	Reverse	GATCCGATTCTCTCGGTTCA
Aven	Forward	CTCTGCCTCCGACTCAAC
	Reverse	CCTTGCCATCATCAGTTCTC
HOXA7	Forward	AAGCTGCCGGACAACAAATC
	Reverse	GAAGCCCCCGCCGTATATTT
GAPDH	Forward	AATCCCATCACCATCTTCCA
	Reverse	TGAGTCCTTCCACGATACCA

4.4.8 Generation of stable clones

CCRF-CEM and MOLT-4 cells maintained in RPMI with 10% FBS were transduced with lentiviruses harboring either shRNA targeting AVEN (5'CCG GGA GAA TGA TGA ACA GGG AAA TCT CGA ATT TCC CTG TTC ATC ATT CTC TTT TTT G-3'), PRMT1 (5'-CGG GTG TTC CAG TAT CTC TGA TTA CTC GAG TAA TCA GAG ATA CTG GAA CAC TTT TTG-3') or control shRNA in the vector pLKO.1. The lentiviruses were generated in HEK293T cells following recommended manufacturer’s protocol with modifications in transfection as follows (9 μg psPAX2 -; 4 μg vsv-g; 9 μg shRNA) per 10 cm plate. The shRNAs were purchased from the shRNA library from Dharmacon. Post-infection, bulk populations of stably infected cells were selected with 2 μg/ml puromycin.

4.4.9 siRNA Transfections

Small interfering RNAs (siRNAs; Dharmacon Inc.) were transfected using Lipofectamine RNAi MAX (Invitrogen) as per the manufacturer's protocol. The final concentration of the siRNA was 40nM and the cells were lysed 72 h post-transfection. The siRNA target sequence for PRMT1 was 5′-CGU CAA AGC CAA CAA GUU A-3′. The siRNA target sequences for Aven were siAven 5′-GAG GAG AAA GAA UGG GAU AUU-3′. For SMN, TDRD3 and DHX36 siRNAs, SMARTpools were purchased from Dharmacon Inc.

4.4.10 Immunoprecipitations and Immunoblotting

PRMT1 $^{FL/-;CreERT}$ MEFs or HEK293T cells were transfected using Lipofectamine 2000 (Invitrogen) as per manufacturer's instructions. After 24 to 48 h, the cells were lysed with cell lysis buffer (20 mM Tris pH 7.4, 150mM NaCl, 1 mM EDTA, 1mM EGTA, 1% Triton X-100). For immunoprecipitations, cell lysates were incubated with the primary antibody for 1 h on ice. Then 25 µl of 50% protein A-Sepharose slurry was added and incubated at 4ºC for 45 min with constant end-over-end mixing. The beads were then washed twice with cell lysis buffer and once with 1x PBS. The samples were then boiled with 25 µl of 2x Laemmli buffer, resolved by SDS-PAGE, transferred to nitrocellulose membranes and the proteins visualized by immunoblotting.

4.4.11 Recombinant GST Pull-Down Assays

U2OS cells transfected with FLAG-Aven full length or FLAG-AvenΔRGG were lysed in lysis buffer 48 h after transfection. Cell lysates were prepared and incubated for 1 h at 4˚C with 20 µl of 50% slurry of the purified GST-Tudor proteins bound to the glutathione agarose. 1 µg of GST protein was used for each pull-down. Following the incubation, the beads were washed three times with lysis buffer and the proteins eluted in 1x Laemmli buffer. The bound proteins were analyzed by SDS-PAGE and visualized by immunoblotting.

4.4.12 In-Line Probing

In line probing assays were performed, as previously described (Beaudoin, Jodoin et al. 2013). Briefly, trace amounts of labelled RNA (50 000 cpm), were heated at 70°C for 5 min and then slow-cooled to room temperature over 1 h in buffer containing 100 mM Tris–HCl (pH 7.5) and 100 mM LiCl or KCl of 10 µl. Following this incubation, the final volume of each sample was adjusted to 20 µl such that the final concentrations were 50 mM Tris–HCl (pH 7.5), 20 mM $MgCl_2$, and 100 mM LiCl or KCl. The reactions were then incubated for 40 h at room temperature, ethanol-precipitated and the RNAs dissolved in formamide dye loading buffer (95% formamide, 10 mM EDTA, and 0.025% bromophenol blue). The radioactivity of the in-line probing samples were measured, and equal amounts in terms of desintegrations per minute of all conditions of each candidate were fractionated on denaturing (8 M urea) 10 % polyacrylamide gels. The SAFA software was used to quantify each band. The intensity of the band incubated in KCl was then divided by the intensity of the corresponding band incubated with LiCl. G4 formation is confirmed when this value exceed 2 (Beaudoin and Perreault 2010; Beaudoin and Perreault 2013). Histograms show the mean result and standard deviation of two separate experiments, that is two different RNA transcription, labeling, and in line probing.

The sequences used were for wild type MLL1 5′-GGC CGC GGC GGC GGC GGC GGG AAG CAG CGG GGC UGG GGU UCC AGG GGG AGC GGC CGC CGC CUC-3′ and for the G/A mutant 5′- GGC CGC GGC GGC GGC GGC GAG AAG CAG CGA AGC UGA AGU UCC AGA AAG AGC GGC CGC CGC CUC -3′. The sequences used were for wild type MLL4 5′-GGC CCG CGG GUC CAG CGG GGC CGG GGA CGG GGU CGG GGC CGG GGC UGG GGC CCG AGU CGA GGC UG-3′ and for the G/A mutant 5′-GGC CCG CGG GUC CAG CGA AGC CGA AGA CGA AGU CGA AGC CGA AGC UGA AGC CCG AGU CGA GGC UG-3′.

4.4.13 Potential G-quadruplex-forming sequences

The following RNA sequences were purchased from IDT.

sc1 (G4 WT) 5'- GCUGC gg UGU gg AA gg AGU gg UC gg GUUGCGCAGCG-biotin-3';

sc1 (G4m) 5'- GCUGC aa UGU gg AA aa AGU gg UC gg GUUGCGCAGCG-biotin-3';

MLL1 (G4; 220-258) 5'- CGGC ggg AAGCAGC gggg CT gggg TTCCA gggg GAGCGG biotin-3';

MLL1 (G4m) 5'- CGGC aaa AAGCAGC gggg CT aaaa TTCCA gggg GAGCGG-biotin-3';

MLL4 (G4; 267-310) 5'-CCAGC gggg CC gggg AC gggg UC gggg CC gggg CU gggg CCCGA-biotin-3';

MLL4 (G4m) 5'-CCAGC *aaaa* CC gggg AC gggg UC *aaaa* CC gggg CU gggg CCCGA-biotin-3'.

4.4.14 Immunofluorescence

U2OS cells grown on coverslips were washed twice with 1x PBS and fixed with 4% paraformaldehyde for 10 min at room temperature. The cells were then washed twice with 1x PBS and permeabilized with 0.5% Triton X-100 for 10 min. Permeabilization was followed by three washes with 1x PBS and cells were blocked with 10% FBS in PBS and labeled with primary antibody diluted in PBS containing 5% FBS. After three washes, the cells were labeled with Alexa Fluor 540 conjugated secondary antibody. DNA was counterstained with 4.6-diamidino-2-phenylindole (DAPI). The coverslips were washed 3 times with 1x PBS and mounted on slides using Immuno-Mount (Thermo Scientific). Images were captured using a Zeiss M1 microscope with epifluorescence optics.

4.4.15 Mass spectrometry and SILAC

Myc-Aven expressing HEK293T cells were lyzed and immunoprecipitated with anti-Myc epitope tagged antibodies. The immunoprecipitated proteins were resolved by SDS-PAGE, visualized by Coomassie Blue (SimplyBlue Safestain, Life Technologies), excised from the gel and digested with trypsin and subjected to LC/MS/MS analysis as previously described (Boisvert, Ahmad et al. 2012). U2OS cells were cultured in DMEM (Dulbeccos's Modified Eagle Medium) depleted of arginine and lysine, as described previously (Boisvert, Ahmad et al. 2012). DMEM was supplemented with 10% of dialyzed FBS. Arginine and lysine were

substituted with light (Arg0, Lys0), medium (Arg6, Lys4) or heavy (Arg10, Lys8) amino acids (Cambridge Isotope Laboratories, Inc.). The cells were cultured in the labeled medium for 6 passages for metabolic incorporation of the labeled amino acids. The light, medium and heavy labeled cells were transfected with empty vector pcDNA3.1, FLAG-Aven and FLAG-AvenΔRGG, respectively. Cell lysates were immunoprecipitated with anti-FLAG agarose beads and the immunoprecipitated complexes were mixed prior to mass spectrometry analysis.

Trypsin digested peptides were separated using a Dionex Ultimate 3000 nanoHPLC system. 10 µl of sample (volume) in 1% (vol/vol) formic acid was loaded with a constant flow of 4 µl/min onto a trap column (100 µm, 2 cm, PepMap C18, Dionex Corporation). After trap enrichment peptides were eluted off onto a PepMap C18 nano column (75 µm x 50 cm, Dionex Corporation) with a linear gradient of 5-35% solvent B (90% acetonitrile with 0.1% formic acid) over 240 minutes with a constant flow of 200 nl/min. The HPLC system was coupled to a QExactive mass spectrometer (Thermo Fisher Scientific Inc) via a EasySpray source. The spray voltage was set to 2.0 kV and the temperature of the column was set to 40 ºC. Full scan MS survey spectra (m/z 350-1600) in profile mode were acquired in the Orbitrap with a resolution of 70,000 after accumulation of 1,000,000 ions. The ten most intense peptide ions from the preview scan in the Orbitrap were fragmented by collision induced dissociation (normalised collision energy 35% and resolution of 17,500) after the accumulation of 50,000 ions. Maximal filling times were 250 ms for the full scans and 60 ms for the MS/MS scans. Precursor ion charge state screening was enabled and all unassigned charge states as well as singly, 7 and 8 charged species were rejected. The dynamic exclusion list was restricted to a maximum of 500 entries with a maximum retention period of 40 s and a relative mass window of 10 ppm. The lock mass option was enabled for survey scans to improve mass accuracy. Data were acquired using the Xcalibur software.

Data were processed, searched and quantified using the MaxQuant software package version 1.4.1.2 as described previously (Cox and Mann 2008) employing the Human Uniprot database (16/07/2013). The settings used for the MaxQuant analysis were: 2 miscleavage

were allowed; fixed modification was N-ethylmaleimide on cysteine; enzymes were Trypsin (K/R not before P); variable modifications included in the analysis were methionine oxidation and protein N-terminal acetylation. A mass tolerance of 7 ppm was used for precursor ions and a tolerance of 20 ppm was used for fragment ions. A maximum false positive rate of 1% was allowed for both peptide and protein identification.

4.4.16 Methylation Assays

Glutathione-S-transferase (GST)-AvenRGG/RG was incubated with GST-tagged PRMT1 with 0.55µCi of (methyl-^{3}H)S-adenosyl-L-methionine in the presence of 25mM Tris-HCl (pH 7.5) for 1 h at 37ºC in a final volume of 25 µl. Reactions were stopped by adding 25µl of 2x Laemmli buffer followed by boiling for 10 min. The proteins were separated by SDS-PAGE, stained with Coomassie Blue and the destained gel was soaked in EN^{3}HANCE (Perkin Elmer Life Sciences), as per manufacturer's instructions and visualized by fluorography.

4.4.17 Generating Aven$^{-/-}$ cells using CRISPR/Cas9

HEK293T were cotransfected with a pEGFP-C1 (Clontech, Palo Alto, CA), a Cas9 expression vector (Addgene, Cambridge, MA) and expression plasmids encoding the following gRNAs; 5'-GGG GCC AGC GCG CCG GTA AGA GG-3' and 5'-GCA GCG GCG GTA GCC AGA GGC GG-3' targeting *Aven* exon 1. The gRNAs expression plasmids were synthesized by IDT (Coralville, IA), as described (Mali, Yang et al. 2013). Single cells expressing GFP were sorted using fluorescence activated cell sorting (FACS) several days after transfection and individual clones were expanded and screened by genomic PCR and by immunoblotting.

4.5 Results

4.5.1 Aven RGG/RG motif is methylated by PRMT1

Aven harbors an N-terminal RGG/RG motif (Thandapani, O'Connor et al. 2013), a nuclear export sequence (NES) (Esmaili, Johnson et al. 2010) and a predicted BH3 domain (Hawley, Chen et al. 2012) (Figure 4.1A). To define the function of the Aven RGG/RG motif, we initially investigated whether the motif was methylated by protein arginine methyltransferase 1 (PRMT1). An *in vitro* methylation assay was performed using a glutathione-S-transferase (GST)-Aven RGG/RG fusion protein incubated with recombinant GST-PRMT1 in the presence of (*methyl*-^{3}H)-S-adenosyl-L-methionine. The proteins were separated by SDS-PAGE, stained with Coomassie Blue to visualize loading, and the methylated proteins were observed by fluorography. The GST-Aven RGG/RG fusion protein migrated as a doublet and was methylated by PRMT1, while GST alone was not methylated (Figure 4.1B, lanes 5, 6).

We next examined whether Aven and PRMT1 associated *in vivo*. HEK293T cells were transfected with Myc-epitope tagged Aven or empty vector (pcDNA3.1) and cell extracts were immunoprecipitated with anti-Myc antibodies. The immunoprecipitated proteins were separated by SDS-PAGE and immunoblotted with anti-PRMT1 antibodies (Figure 4.1C). PRMT1 was present in anti-Myc immunoprecipitates of the Myc-Aven transfected cells, but not in the empty vector transfected cells (Figure 4.1C, lanes 3, 4). Immunoblotting with anti-Myc confirmed the presence of Myc-Aven (Figure 4.1C, lanes 5-8). To examine whether Aven is a substrate of PRMT1 *in vivo*, we depleted HEK293T cells of PRMT1 using siRNA. The cells were also transfected with empty plasmid (pcDNA3.1) or an expression vector encoding 5 tags of the Myc epitope linked to Aven (Myc-Aven). Cellular lysates were immunoprecipitated with immunoglobulin (IgG) control or anti-Myc antibodies, resolved by SDS-PAGE and immunoblotted with ASYM25b, an asymmetric dimethylarginine-specific antibody. We observed that Myc-Aven was arginine methylated in *PRMT1* proficient, but not in PRMT1-deficient cells (Figure 4.1D, compare lanes 3 and 4). An anti-Myc immunoblot confirmed the immunoprecipitations. Immunoblots of total cell lysates (TCL) confirmed the

myc-Aven expression and the PRMT1 knockdown and rps6 was used as a loading control (Figure 4.1D).

A similar experiment was performed using conditional *PRMT1*$^{FL/-;CreERT}$ mouse embryo fibroblasts (MEFs) (Yu, Chen et al. 2009). Ablation of PRMT1 was achieved by treating the cells with 4-hydroxytamoxifen (OHT) for 6 days. An expression vector encoding Myc-Aven was transfected into control (*PRMT1*$^{FL/-;CreERT}$; -OHT) and PRMT1-deficient (*PRMT1* $^{FL/-;CreERT}$; +OHT) MEFs. Myc-Aven was arginine methylated in *PRMT1* proficient, but not in PRMT1-deficient cells (Figure 4.1E, compare lanes 1 and 2). We next proceeded to show that endogenous Aven is arginine methylated. Conditional *PRMT1*$^{FL/-;CreERT}$ were used for Aven immunoprecipitations and indeed endogenous Aven was asymmetrically dimethylated by PRMT1 (Figure 4.1F). Endogenous Aven migrates at ~50 kDa (Figure 4.1F), while myc-Aven was generated with multiple Myc tags to avoid overlap with heavy chain of IgG during immunoprecipitation and migrated at ~70 kDa (Figure 4.1C-E). To identify the arginines within the RGG/RG motif that are methylated, immunopurified Myc-Aven was subjected to mass spectrometry analysis. We identified the dimethylation of Aven R37, R63 and R66 (Supplementary Figure 4.1). Moreover, we also identified the dimethylation of R8, R11, R50, as well as the monomethylation of R5, R8, R28 and R37. Arginines R5, R8, R11, R37, R50 and R63 are also conserved in murines (Supplementary Figure 4.1). Taken together, these findings demonstrate that the Aven RGG/RG motif is methylated by PRMT1 on conserved arginines.

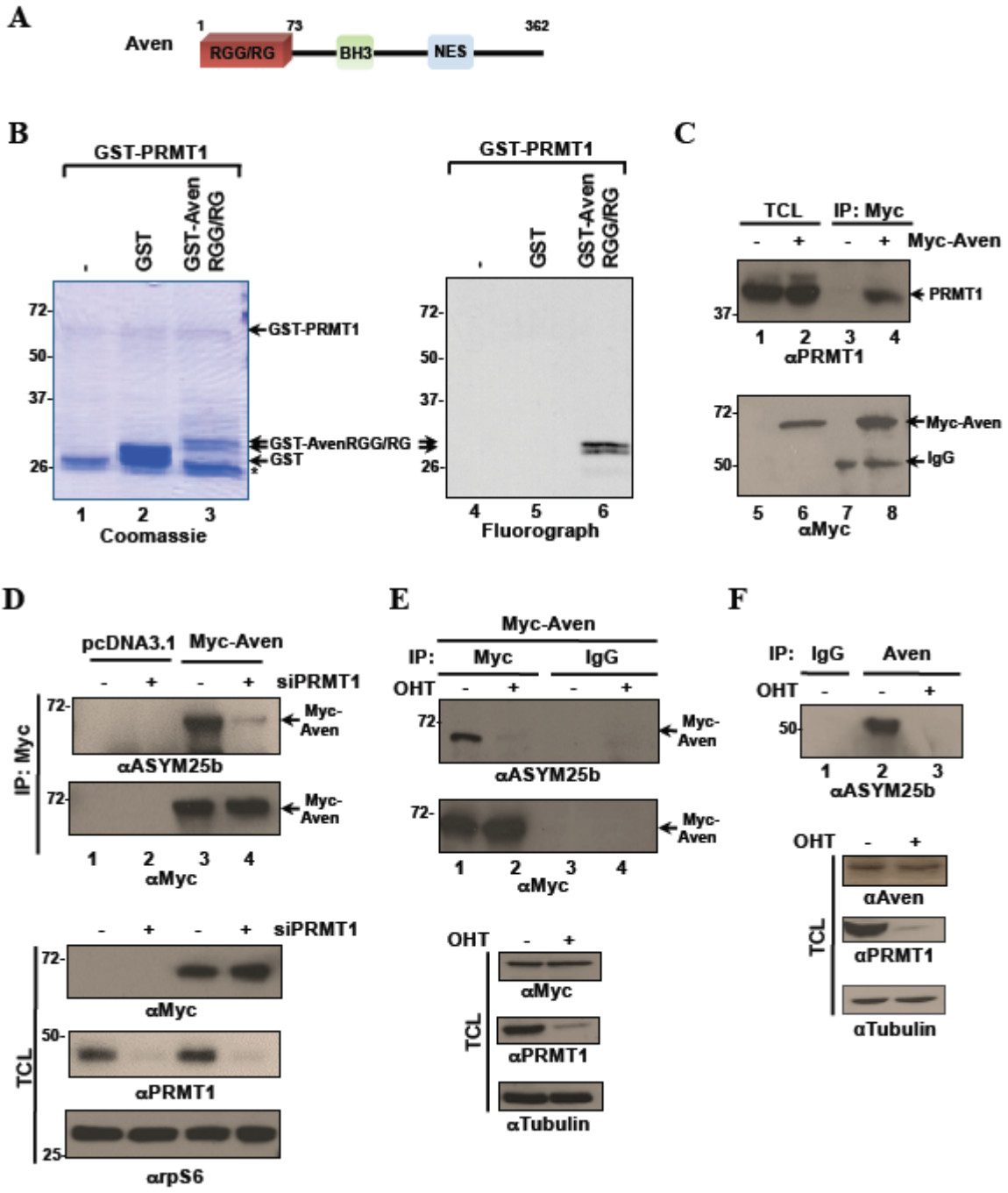


Figure 4.1 Aven is a substrate of PRMT1.

A) Schematic diagram of Aven with its RGG/RG motif, putative BH3 domain and NES.

B) *In vitro* methylation assay with GST-PRMT1 and GST-AvenRGG/RG with (^{3}H)-S-adenosyl-L-methionine as the methyl donor (n > 4). Proteins were resolved by SDS-PAGE, stained with Coomassie Blue (left) and analyzed by fluorography (right). The migration of the molecular mass markers is shown on the left in kDa and the migration of the GST-PRMT1, GST, GST-AvenRGG/RG proteins is indicated with arrows. The asterisk (*) denotes degraded proteins from the GST-PRMT1 preparation.

C) HEK293T cells were transfected with empty vector pcDNA3.1 (lanes 1 and 3) or transfected with Myc-Aven-pcDNA3.1 (lanes 2 and 4) were lysed and immunoprecipitated

(IP) with anti-Myc antibodies (lanes 3 and 4). The bound proteins were resolved by SDS-PAGE and immunoblotted with anti-PRMT1 (lanes 3 and 4) and -Myc antibodies (7 and 8). TCL denotes input total cell lysates and IgG represents the heavy chain of immunoglobulin G.

D) HEK293T cells were cotransfected with siGFP (- PRMT1) or siPRMT1 (+ PRMT1) and pcDNA3.1 or Myc-Aven plasmids. After 48 h, the cells were lysed and immunoprecipitated (IP) with anti-Myc antibodies, the proteins were resolved by SDS-PAGE and immunoblotted with anti-ASYM25b, anti-PRMT1, and anti-Myc antibodies. rpS6 levels were obtained by using anti-rpS6 antibodies and show equivalent loading.

E) *PRMT1*$^{FL/-;CreERT}$ MEFs treated with 4-hydroxytamoxifen (OHT) for 6 days or left untreated were transfected with Myc-Aven followed by anti-Myc antibody immunoprecipitations and the methylation monitored by immunoblotting with ASYM25b (lanes 1-4) or anti-Myc antibodies (lanes 5-8). TCL were immunoblotted with anti-PRMT1, anti-Aven and anti-tubulin antibodies as indicated.

F) *PRMT1*$^{FL/-;CreERT}$ MEFs treated with OHT for 6 days or left untreated were lysed and immunoprecipitated with anti-Aven (ab77014) or IgG antibodies. Immunoprecipitates were blotted with ASYM25b. TCL were immunoblotted with anti-PRMT1, anti-Aven and anti-tubulin antibodies as indicated.

4.5.2 RGG/RG motif of Aven binds G4 sequences

To define the role of the Aven RGG/RG motif, we generated a mutant that lacks the motif by deleting amino acids 1 to 73. AvenΔRGG was not recognized by ASYM25b, confirming that all the methylarginines reside in the N-terminus of Aven (Figure 4.2A). AvenΔRGG was able to activate ATM, comparably to wild type Aven and deletion of the RGG/RG motif did not interfere with its ability to oligomerize (Supplementary Figure 4.2). Aven is predominantly localized in the cytoplasm with some weak nuclear staining (Supplementary Figure 4.2), as reported previously (Chau, Cheng et al. 2000) and FLAG-AvenΔRGG had the same cellular localization pattern as wild type Aven (Supplementary Figure 4.2).

RGG/RG motifs are enriched amongst RNA-binding proteins and they possess inherent RNA binding activity (reviewed in (Thandapani, O'Connor et al. 2013). A high affinity RNA sequence that forms a G4, termed *sc1*, binds RGG/RG sequences (Phan, Kuryavyi et al. 2011). To test whether the RGG/RG motif of Aven binds G4 sequences, we performed binding assays using *sc1*. Biotinylated RNA sequences of *sc1* were generated, heated, and slowly cooled in the presence of K^+ to favor formation of the G4 RNA structure. The RNA was used in an affinity 'pull-down' assay with HEK293T cell lysates. The presence of Aven after different washes of sodium chloride was monitored by immunoblotting following separation of the bound proteins by SDS-PAGE. Aven bound the *sc1* wild type G4 RNA, but not a mutant *sc1* RNA sequence that is predicted not to form the G4 structure (Figure 4.2B).

To determine whether the Aven binding to the G4 RNA structures was mediated by the RGG/RG motif, binding assays were performed in HEK293T cells expressing FLAG-Aven and the FLAG-AvenΔRGG. FLAG-Aven, but not FLAG-AvenΔRGG, bound the *sc1* G4 structure indicating that the RGG/RG motif is necessary for binding (Figure 4.2C). We subsequently investigated whether methylation of the RGG/RG motif influences binding to the *sc1* G4 structure. Hypomethylated Aven was obtained by depleting HEK293T cells of PRMT1 with siRNAs. Aven bound equally well to biotinylated *sc1* G4 RNA from wild type or PRMT1 depleted cells (Figure 4.2D). A biotinylated Aven RGG/RG peptide denoted as TriRG was synthesized with or without asymmetric dimethylarginines and RNA binding was measured using fluorescently-labeled RNA. RGG/RG motif whether harboring arginine or

asymmetric dimethylarginines bound a fluoresceinated *sc1* G4 RNA, induced to fold into a G4 structure prior to binding, with same relative affinities with a Kd of ~80 to 90 nM (Figure 4.2E). Another set of peptides spanning the DiRGG motif of Aven with and without asymmetric dimethylarginines also bound the fluoresceinated *sc1* G4 RNA with lower affinity than the TriRG peptides (Kd of ~175 to 200 nM, Supplementary Figure 4.2), suggesting that several RGG/RG motifs in Aven are able to bind RNA. These finding show that Aven is an RNA-binding protein that interacts with G4 RNA sequences via its RGG/RG motifs independent of arginine methylation.

4.5.3 RGG/RG motif of Aven interacts with TDRD3 and SMN in a methyl-dependent manner

Arginine methylation is known to regulate protein-protein interactions with Tudor domain-containing proteins (Chen, Nott et al. 2011). Methylated RGG/RG motifs are known to interact with certain Tudor domain-containing proteins, such as TDRD3, SMN and SPF30 (Selenko, Sprangers et al. 2001; Cote and Richard 2005). Since arginine methylation did not influence the ability of Aven to bind RNA, we next investigated whether Aven interacts with TDRD3, SMN and SPF30 in a methyl-dependent manner. U2OS cellular lysates expressing FLAG-Aven or FLAG-AvenΔRGG were incubated with the GST-TDRD3, -SMN and -SPF30 bound to Sepharose beads. The presence of bound Aven was detected by SDS-PAGE followed with anti-FLAG immunoblotting. FLAG-Aven, but not FLAG-AvenΔRGG, interacted with the GST Tudor domains of TDRD3 and SMN, suggesting that TDRD3 and SMN Tudor domains interact with the Aven RGG/RG motif (Figure 4.3A). GST-SPF30 Tudor domain had a weak interaction with FLAG-Aven, but not with FLAG-AvenΔRGG (Figure 4.3A). To verify whether Aven interacts with TDRD3 and SMN *in vivo*, co-immunoprecipitations were performed. Endogenous TDRD3 and SMN co-immunoprecipitated with FLAG-Aven, but not with FLAG-AvenΔRGG (Figure 4.3B, C). To verify whether FLAG-Aven interacts with TDRD3 and SMN in an arginine methylation dependent manner, co-immunoprecipitations were performed in PRMT1-depleted and control HEK293T cells. Indeed cells depleted of PRMT1 showed reduced interaction between

FLAG-Aven and SMN and TDRD3, as compared to the control (Figure 4.3D). To confirm the interaction of endogenous Aven with SMN and TDRD3, conditional *PRMT1*$^{FL/-;CreERT}$ MEFs were treated with OHT or a vehicle and the cellular lysates were immunoprecipitated with anti-Aven antibodies. Immunoprecipitates were resolved by SDS-PAGE and immunoblotted with anti-SMN and anti-TDRD3 antibodies. The ability of SMN and TDRD3 to co-immunoprecipitate with endogenous Aven was lost in PRMT1-deficient cells (+ OHT; Figure 4.3E). These findings confirm that methylation of the RGG/RG motif is required for interaction between the Aven/ TDRD3 and Aven/ SMN complexes.

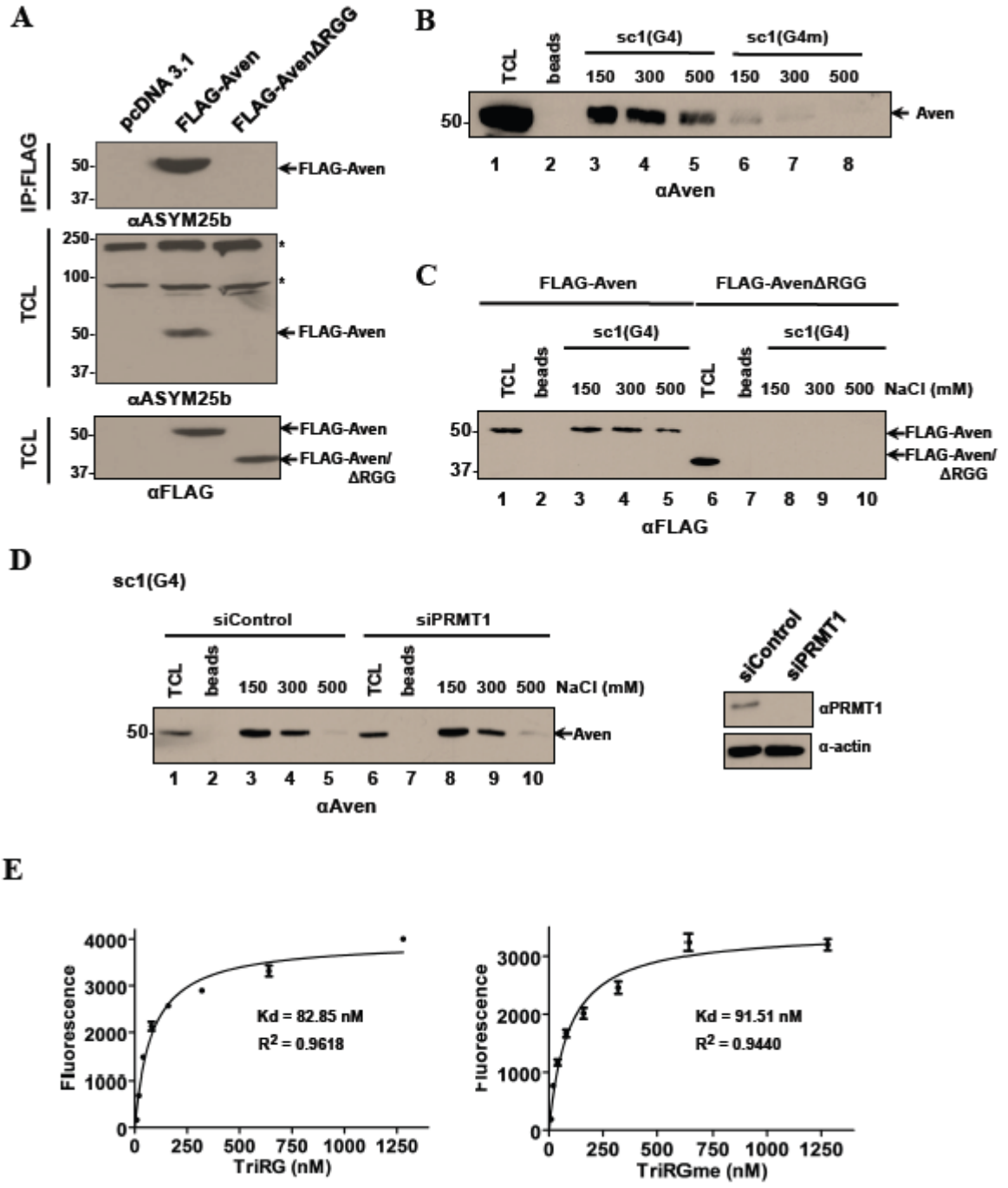


Figure 4.2 Aven binds G4 RNA sequences in an arginine methylation independent manner.

A) U2OS cells transfected with pcDNA3.1, or expression vectors encoding FLAG-Aven or FLAG-AvenΔRGG were immunoprecipitated with anti-FLAG antibodies and immunoblotted with ASYM25b or anti-FLAG antibodies as indicated. The molecular mass markers are shown on the left in kDa and the migration of FLAG-Aven and FLAG-AvenΔRGG is shown. The asterisks (*) denote unknown arginine methylated proteins.

B) Biotinylated *sc1* G4 or (G4m) bound to Streptavidin were incubated with HEK293T cell

lysates. The bound proteins were washed with increasing NaCl as indicated and visualized by SDS-PAGE followed by immunoblotting with anti-Aven antibodies.
C) Biotinylated *sc1* G4 or (G4m) bound to Streptavidin were incubated with cellular lysates expressing FLAG-Aven and FLAG-AvenΔRGG and detected as in panel A.
D) Biotinylated *sc1* G4 RNA bound to Streptavidin beads was used to pull-down Aven from PRMT1-depleted HEK293T cells. Aven binding was performed as in panel A. PRMT1 depletion was confirmed by immunoblotting.
E) Biotinylated methylated and unmethylated Aven TriRG peptides were pre-bound on Streptavidin plates and were incubated with fluorescein-labeled *sc1* G4 RNA. The bound RNA was quantified by measuring absorbance at 515nm. The experiment was performed in triplicate.

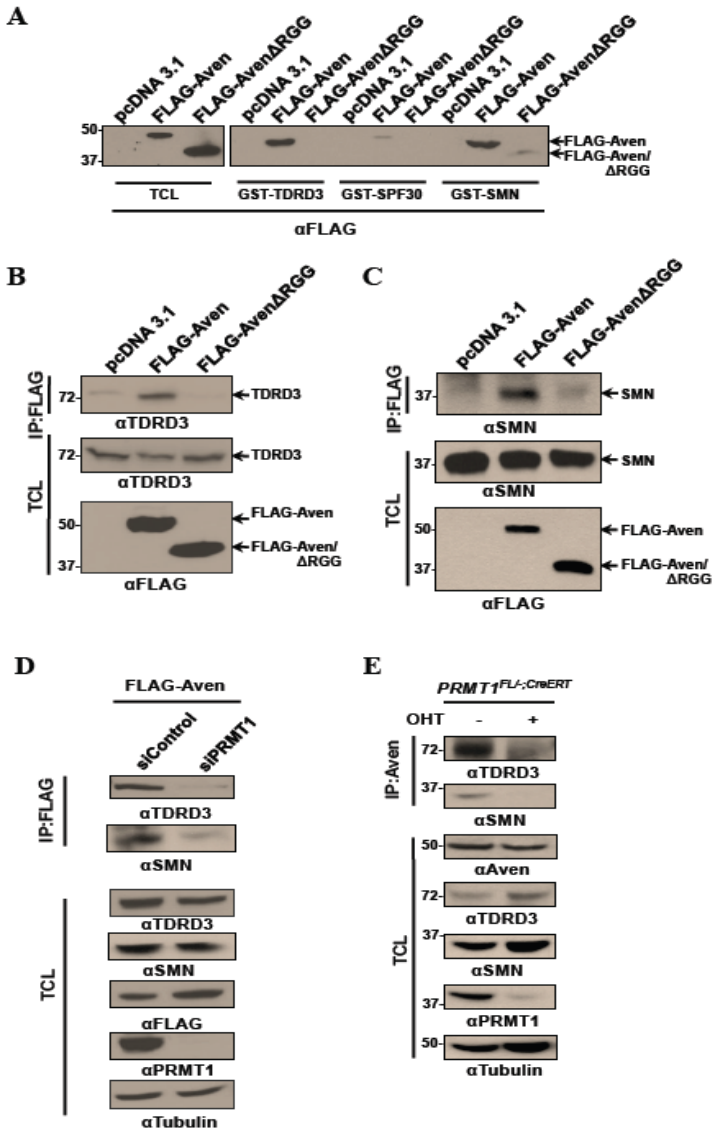


Figure 4.3 Tudor domains of TDRD3 and SMN recognize methylated Aven.

A) Recombinant Tudor domains of TDRD3, SPF30 and SMN were fused GST and used in 'pull-down' assays with HEK293T lysates expressing pcDNA3.1 (control), FLAG-Aven or FLAG-AvenΔRGG. The bound proteins were separated by SDS-PAGE and immunoblotted with anti-FLAG antibodies.

B, C) Lysates from HEK293T lysates expressing pcDNA3.1 (control), FLAG-Aven or FLAG-AvenΔRGG were immunoprecipitated with anti-FLAG antibodies. Co-immunoprecipitation of endogenous TDRD3 and SMN was detected by immunoblotting.

D) Aven interaction with TDRD3 and SMN was reduced in cells deficient for PRMT1 using siRNAs. FLAG-Aven was co-expressed with either siControl or siPRMT1 in U2OS cells. Anti-FLAG antibody immunoprecipitations were performed and the presence of endogenous TDRD3 and SMN monitored by immunoblotting following separation by SDS-PAGE.

E) *PRMT1*$^{FL/-;CreERT}$ MEFs treated with OHT for 6 days or left untreated were lysed and immunoprecipitated with anti-Aven antibodies. Co-immunoprecipitation of endogenous TDRD3 and SMN was detected by immunoblotting (upper panels).

4.5.4 Methylated RGG/RG motif regulates association of Aven with polysomes.

To identify the interactome of the RGG/RG motif of Aven, we used a stable isotope labeling by amino acids in cell culture (SILAC) approach to quantify protein complexes differentially associated with Aven and AvenΔRGG (Blagoev, Kratchmarova et al. 2003). U2OS cells transfected with pcDNA3.1, FLAG-Aven or FLAG-AvenΔRGG were light (L), medium (M) or heavy (H) SILAC labeled, respectively and immunoprecipitated with anti-FLAG antibodies. The data was expressed as fold-enrichment of Aven over control (M/L) and AvenΔRGG over control (H/L). Aven was identified and quantified with high M/L and H/L ratios, while PRMT1 was only found with a high ratio M/L showing as expected that it associates with its RGG-containing substrates. Of the 146 proteins enriched with Aven, but not AvenΔRGG, ~23% were ribosomal proteins and ~10% were RNA-binding proteins. These data suggested that Aven, but not AvenΔRGG, associates with ribosomal proteins and/or ribonucleoprotein complexes, thereby implying that Aven associates with ribosomes in an RGG/RG motif dependent manner. To address this question, we first assessed whether endogenous Aven and PRMT1 associate with ribosomes by sedimenting cytoplasmic extracts on 5-50% sucrose gradients by ultracentrifugation (Figure 4.4A). Subsequently, sucrose gradients were fractionated to separate cytoplasmic mRNPs, ribosomal subunits, monosomes and polysomes (Gandin, Sikström et al. 2014) and the amount of Aven, PRMT1, rpS6 (ribosomal protein S6) and β-tubulin in each fraction was determined by immunoblotting. Aven and PRMT1 co-sedimented with the heavier polysomal fractions with rpS6, while β-tubulin was restricted to the lighter fractions of the sucrose gradient corresponding to cytoplasmic mRNP fractions (Figure 4.4B, fractions 11 to 16). To investigate whether the RGG/RG motif and its methylation regulates the recruitment of Aven to polysomes, we performed polysomal fractionation in HEK293T cells transfected with FLAG-Aven or FLAG-AvenΔRGG (Supplementary Figure 4.3). FLAG-Aven cosedimented with the heavier polysomal fractions (Figure 4.4C). In contrast, FLAG-AvenΔRGG was largely restricted to lighter fractions representing free-ribosomal subunits (Figure 4.4D). These findings demonstrated that RGG/RG motif of Aven is required for its recruitment to polysomes. To further investigate the role of RGG/RG motif methylation in polysomal localization of Aven,

the FLAG-Aven was co-transfected with siPRMT1, which reduced PRMT1 expression by 2.7-fold (Supplementary Figure 4.3). Similar to what has been reported in fission yeast (Bachand and Silver 2004), PRMT1 depletion did not have a major effect on the monosome/polysome ratio, thus indicating that PRMT1 does not exert major impact on global protein synthesis (Supplementary Figure 4.3). Nonetheless, PRMT1 depletion shifted FLAG-Aven into the lighter fractions as compared to a control (Figure 4.4E, Supplementary Figure 4.3). We confirmed that Aven does not co-sediment with mRNPs other than polysomes by showing that puromycin, an aminonucleoside antibiotic that dissociates polysomes (Blobel and Sabatini 1971), leads to redistribution of both FLAG-Aven and rpS6 towards the lighter fractions corresponding to free ribosomal subunits, monosomes, and cytoplasmic mRNPs (Figure 4F). Taken together, our findings demonstrate that the arginine methylation of the Aven RGG/RG motif by PRMT1 is required for the association of Aven with polysomes.

TDRD3 and SMN are known to be polysome-bound (Goulet, Boisvenue et al. 2008; Sanchez, Dury et al. 2013). Thus to determine their requirement for polysomal localization of Aven, FLAG-Aven was transfected in HEK293T cells depleted of SMN or TDRD3 or both using siRNAs. The depletion of TDRD3 (6-fold) and SMN (2-fold) was confirmed by immunoblotting (compare Figure 4.4G and 4.4H). FLAG-Aven cosedimented in the polysome fractions with endogenous TDRD3, SMN and rpS6 in siCTRL (Figures 4.4G, fractions 11 to 16). In contrast, depletion of both TDRD3 and SMN, but not single TDRD3 or SMN depletion (data not shown) reduced polysomal association of Aven as compared to the control (Figure 4.4H). These findings suggest that TDRD3 and SMN are required for the recruitment of Aven to polysomes, whereby these proteins likely play a redundant role.

4.5.5 Aven RGG/RG motif binds the G4 sequences of MLL1 and MLL4

Next, we reasoned that since RGG/RG motifs are encoded by G-rich sequences (codons: Gly GGN; Arg CGN or AGA/G), it is likely that certain mRNAs encoding RGG/RG motif-containing proteins may occasionally comprise a PG4 sequences (Figure 4.5A). To identify these RGG/RG encoding sequences and to identify all PG4 sequences in the coding sequences, we performed a bioinformatic search for PG4 sequences in mRNA coding regions

(G_x-N_{1-7}-G_x-N_{1-7}-G_x-N_{1-7}-Gx where x $\geq$3 and N is any of the nucleotides (A, C, G or U)), and it revealed ~1600 PG4 in human ORFs. We also provide a cG/cC score where > 2 predicts higher G4 formation over Watson-Crick base pairing with neigbhoring sequences (Beaudoin, Jodoin et al. 2014). In addition, we also use RNAfold v2.1.7 which is a new scoring system to identify RNA G4-folding (Lorenz, Bernhart et al. 2013). RNAfold v2.1.7 was less efficient in predicting G4 formation than the cG/cC ratio (Beaudoin, Jodoin et al. 2014).

We observed a preference for amino acids G>>R/A/P>E/L consistent with the frequency of guanine in each codon. Some PG4 sequences identified encoded RG-rich sequences, as well as glycine-rich sequences. As Aven is a known survival protein and its depletion decreases the proliferation of leukemic cells (Eismann, Melzer et al. 2013), we searched for mRNAs harboring PG4 sequences involved in leukemic cell survival. We identified human KMT2A, KMT2B and KMT2D, the family of Mixed Lineage Leukemia (MLL) histone methyltransferases known to be mutated and/or part of fusion proteins, as the result of chromosomal translocations in leukemia (Liedtke and Cleary 2009; Smith, Lin et al. 2011). KMT2A (MLL1) has a PG4 sequence (nucleotide 223-253) that encodes a glycine-rich sequence (AGSSGAGVPGG, Figure 4.5B; cG/cC score of 2.975, KMT2B (MLL4) has 3 PG4 sequences (nucleotides 262-287; 3139-3166; 6274-6301) that encode an RGG/RG motif (RVQRGRGRG, Figure 5B; cG/cC score of 2.796), a glycine-rich (RGAGAGGPRE; cG/cC score of 1.887), an alanine-glycine-rich (RAGVLGAAGD; cG/cC score of 2.468) sequences and KMT2D has a PG4 from nucleotide 961 to 986 that encodes RVCRACGAG (cG/cC score of 2.029). We examined whether Aven associated with the conserved PG4 RNA sequences of KMT2A (MLL1) and KMT2B (MLL4) near the initiator ATG (i.e. 223-253 and 262-287). Indeed, Aven bound either MLL1 or MLL4 PG4 RNA sequences in the RGG/RG motif-dependent manner, but not to those harboring guanine to adenine mutants (Figure 4.5C, D, E, F). We next examined whether FLAG-Aven binds the endogenous MLL1 and MLL4 PG4 sequences *in vivo* using PAR-CLIP (photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation) (Hafner, Landthaler et al. 2010) with a dilution of RNase I that digests RNAs into fragments of 50 to 300 nucleotides in length (Huppertz, Attig et al. 2014). HEK293T transfected with pcDNA3.1, FLAG-Aven or FLAG-AvenΔRGG were

prepared for CLIP, as described in ‘Materials and Methods’ and immunoprecipitated with anti-FLAG antibodies (Figure 4.5G). Anti-FLAG immunoprecipitations of FLAG-Aven expressing cells, but not pcDNA3.1 or FLAG-AvenΔRGG transfected cells, enriched ~30-fold for the MLL4 G4 sequence, whereas an RNA region 300 nucleotides downstream was not enriched (noG4; Figure 4.5G, right panel). MLL1 G4 sequence was also enriched (~5-fold) in FLAG-Aven immunoprecipitations, but not a region without G4 motifs (noG4) (Figure 4.5G, left panel). We investigated whether endogenous Aven could associate with the PG4s of MLL1 and MLL4 mRNAs. PAR-CLIP assays confirmed that MLL1 and MLL4 PG4 sequences associated with immunoprecipitated endogenous Aven, but not immunoglobulin G control albeit with a lower affinity (~4-6 fold, Figure 4.5H) than with overexpressing FLAG-Aven (Figure 4.5G). Ultraviolet light crosslinking at 365nM was required for this association (Figure 4.5H). These findings demonstrate that Aven is associated *in vivo* with the PG4s of MLL1 and MLL4.

We next performed in-line probing experiments, to determine whether the MLL PG4 sequences formed *bona fide* G4 structures. This assay compares the cleavage pattern in two conditions: in presence of $K^{+,}$ which support G4 formation, and in presence of Li^{+}, which does not. G4 folding leads to an increased cleavage for the nucleotides within the loop regions, since they bulge out (Beaudoin and Perreault 2010; Beaudoin and Perreault 2013). Using MLL1 and MLL4 PG4, we observed increased cleavage in the predicted loops of the formed G4, when incubated in K^{+} compared to Li^{+}. However, no such difference was observed for a mutant RNA, where guanines were replaced with adenines (Figure 4.5I), confirming that the PG4 of MLL1 and MLL4 form G4 RNA structures.

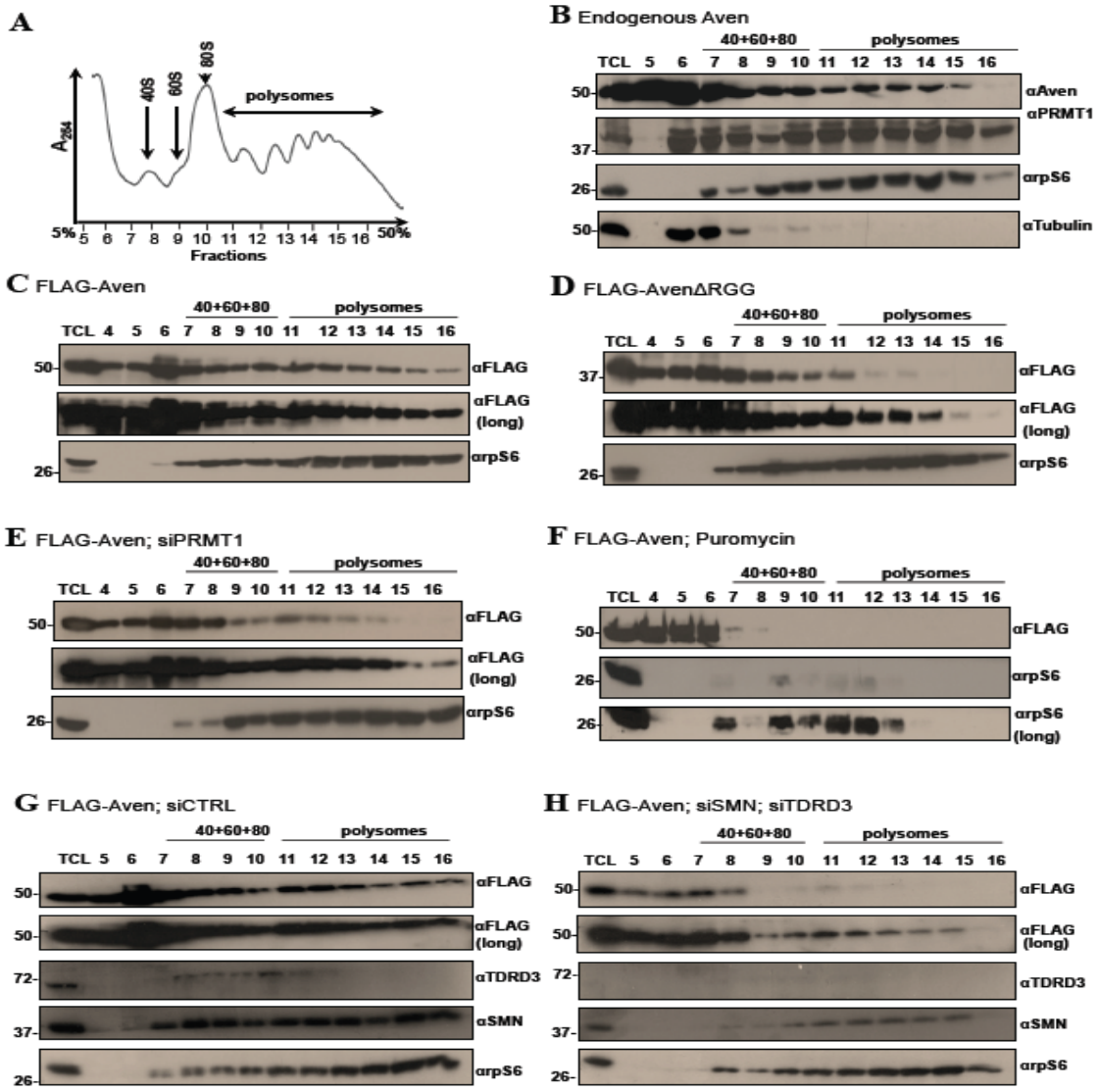
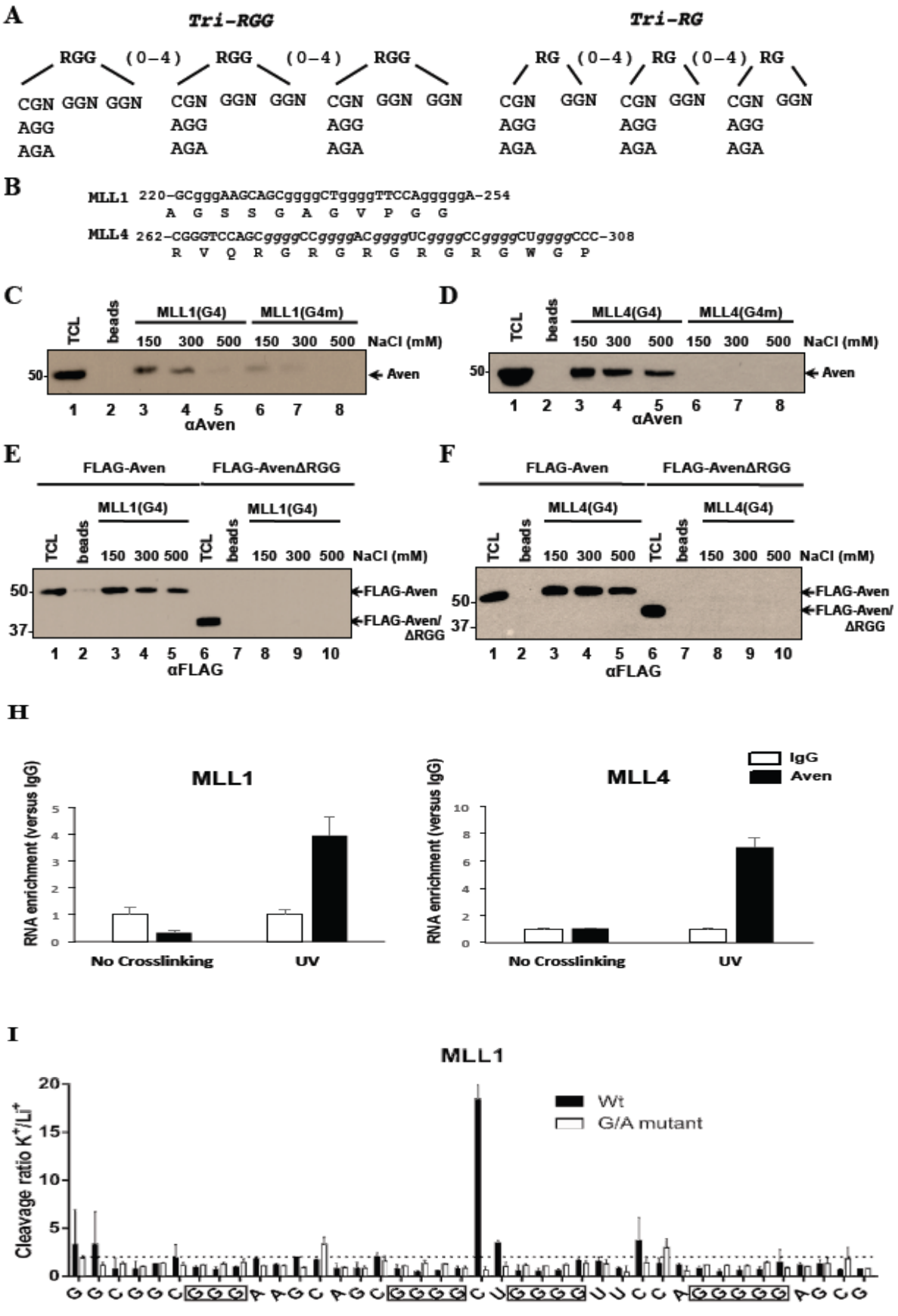


Figure 4.4 Methylation of Aven and its association with TDRD3 and SMN is required for polysomal localization.

A) Cytoplasmic extracts from HEK293T cells were sedimented by centrifugation on a 5 to 50% sucrose gradient. Polysome profiles were obtained by continuous monitoring of UV absorbance at 254nm. 40S, 60S and 80S indicate the positions of the respective ribosomal subunits and the monomer on the gradient.

B) The distribution of endogenous Aven and PRMT1 across the gradient of *panel A* was monitored by immunoblotting. Ribosomal protein rpS6 was used as a loading control, whereas □-tubulin served as a cytoplasmic marker.

C-H) The distribution of FLAG-Aven or FLAG-AvenΔRGG across the gradient was monitored by immunoblotting as well as FLAG-Aven in siControl, siPRMT1, siSMNsiTDRD3 or with puromycin treatement. Both short (5 s, panels C-H) and long exposures (30 s, panels C-F) are shown. rpS6 was used as a loading control. The exposure time was determined using a standard curve with increasing amounts of lysates expressing FLAG-Aven immunoblotted with anti-FLAG antibodies for various times. Each polysomal profile experiment was performed independently at least 3 times.


A
Tri-RGG
RGG (0–4) RGG (0–4) RGG
CGN GGN GGN CGN GGN GGN CGN GGN GGN
AGG AGG AGG
AGA AGA AGA
Tri-RG
RG (0–4) RG (0–4) RG
CGN GGN CGN GGN CGN GGN
AGG AGG AGG
AGA AGA AGA
B
MLL1 220–GCgggAAGCAGCggggCTggggTTCCAgggggA–254
A G S S G A G V P G G
MLL4 262–CGGGTCCAGCggggCCggggACgggguCgggggCCgggguCgggggCCC–308
R V Q R G R G R G R G R G R G W G P
C
TCL
beads
MLL1(G4)
MLL1(G4m)
150 300 500 150 300 500 NaCl (mM)
50
Aven
1 2 3 4 5 6 7 8
αAven
D
TCL
beads
MLL4(G4)
MLL4(G4m)
150 300 500 150 300 500 NaCl (mM)
50
Aven
1 2 3 4 5 6 7 8
αAven
E
FLAG-Aven
FLAG-AvenΔRGG
MLL1(G4)
MLL1(G4)
TCL beads 150 300 500 TCL beads 150 300 500 NaCl (mM)
50
37
FLAG-Aven
FLAG-Aven/ ΔRGG
1 2 3 4 5 6 7 8 9 10
αFLAG
F
FLAG-Aven
FLAG-AvenΔRGG
MLL4(G4)
MLL4(G4)
TCL beads 150 300 500 TCL beads 150 300 500 NaCl (mM)
50
37
FLAG-Aven
FLAG-Aven/ ΔRGG
1 2 3 4 5 6 7 8 9 10
αFLAG
H
MLL1
RNA enrichment (versus IgG)
No Crosslinking
UV
MLL4
IgG
Aven
RNA enrichment (versus IgG)
No Crosslinking
UV
I
MLL1
Cleavage ratio K+/Li+
Wt
G/A mutant
GGCGGCGGGAAGCAGCGGGGCUGGGGGUUCCCAGGGGGAGCG

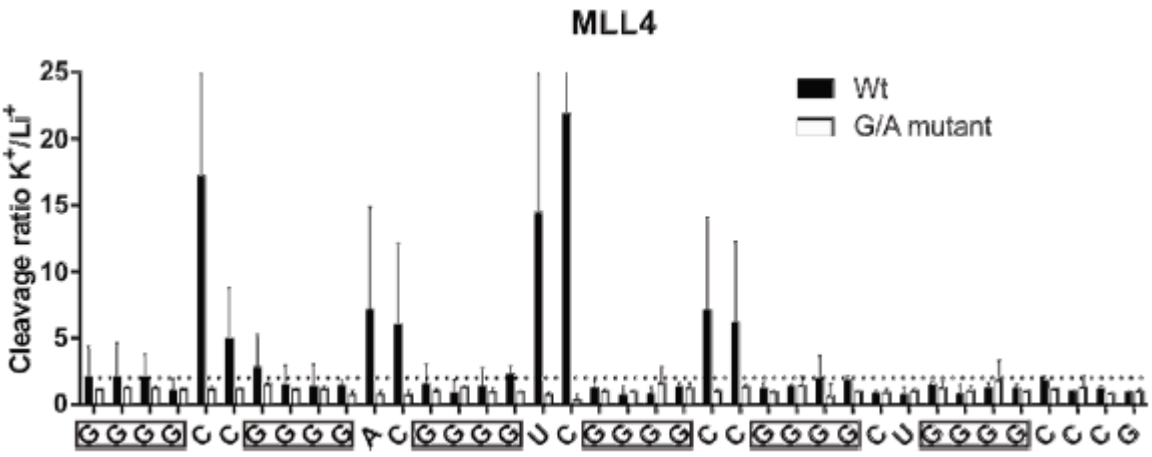


Figure 4.5 Aven RGG/RG motif binds G4 RNA structures of MLL1 and MLL4.

A, B) RNA sequences of the RGG/RG motifs and the PG4 motifs of MLL1 and MLL4.

C, D) Biotinylated MLL1 G4 or a mutant sequence (G4m), biotinylated MLL4 G4 or a mutant sequence (G4m) bound to Streptavidin beads were incubated with HEK293T cell lysates. The bound proteins were washed with increasing concentrations of NaCl and visualized by SDS-PAGE followed by immunoblotting with anti-Aven antibodies.

E, F) HEK293T cells expressing FLAG-Aven and FLAG-AvenΔRGG were processed as in *panel A, B* except the bound proteins were visualized by immunoblotting with anti-FLAG antibodies.

G) PAR-CLIP assays were performed using anti-FLAG antibodies. The bound RNA was analyzed in triplicate from two biological replicates by RT-qPCR with the primers spanning the PG4 sequence or a sequence ~300 nucleotides downstream. The levels of bound RNA in immunoprecipitates were normalized to the levels of the total RNA in the input. Mean values are expressed as fold enrichment over pcDNA3.1. Error bars represent ± SEM.

H) PAR-CLIP assays were performed on HEK293T cells using anti-Aven antibodies. The bound RNA was analyzed in triplicates by real-time RT-PCR with the primers spanning the PG4 sequence, as indicated in panel G. The level of bound RNA in immunoprecipitates was normalized to the levels of the total RNA in the input. Mean values are expressed as fold enrichment over IgG. Error bars represent ± SEM, n=2.

I) In-line probing of MLL1 and MLL4 PG4. The nucleotide sequence of the MLLs PG4 is shown below, the boxed guanines represent the predicted G-tracks. K^+/Li^+ ratios of the band intensities of the MLLs G4 (black) and G/A-mutant (white) for each nucleotide are shown. Error bars represent ± standard deviation, n = 2. The dashed line represents a 2-fold change, an arbitrary set threshold that indicates G4 formation when exceeded.

4.5.6 Aven regulates the protein, but not the mRNA, levels of MLL1 and MLL4 in T-ALL cells

Aven is required for the proliferation of T-ALL cells (Eismann, Melzer et al. 2013). We generated Aven-deficient T-ALL cells, MOLT4 and CCRF-CEM using a lentivirus that expresses an shRNA against Aven and we achieved >80% knockdown (Figure 4.6A, B). We next monitored the levels of MLL1 and MLL4 protein by immunoblotting and their mRNAs by RT-qPCR. Both MLL1 and MLL4 protein levels were reduced in Aven-deficient MOLT4 and CCRF-CEM cells (Figure 4.6A, B), whereas the levels of corresponding mRNAs remained unchanged (Figure 4.6C, D). Next we investigated whether the reduced protein levels of MLL1 and MLL4 in Aven depleted cells are associated with the reduced expression of *HOX* genes, which are well-established transcriptional targets of MLLs (Krivtsov and Armstrong 2007). Aven-deficient cells, exhibited reduced expression of several key *HOX* genes such as *HOXA9*, *HOXA7*, *HOXA1* and *MEIS1* (Figure 4.6E, F). These findings suggest that Aven regulates the translation of MLL1 and MLL4 mRNAs thereby leading to an increase in MLL1 and MLL4 protein levels and an increase in the transcription of leukemic genes.

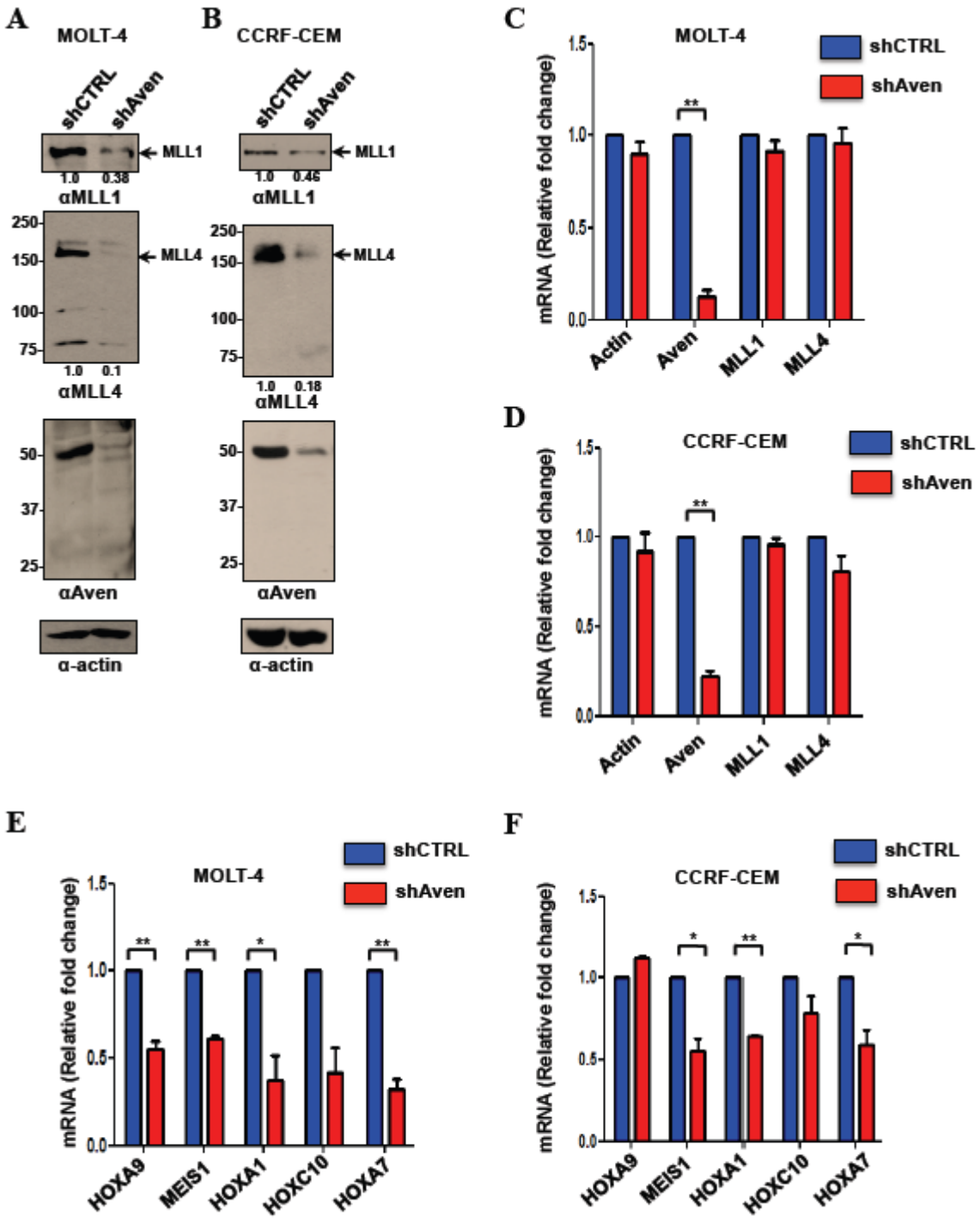
Aven-deficient MOLT4 and CCRF-CEM had decreased proliferation rates consistent with Aven being a survival protein (Figure 4.6G, H). Interestingly, the depletion of PRMT1 with shRNAs phenocopied Aven depletion, i.e. also had reduced proliferation rates (Figure 4.6G, H). These findings show that Aven and PRMT1 are required for the proliferation of T-ALL cells.

4.5.7 Aven and PRMT1 regulate the polysomal association of MLL1 and MLL4 mRNAs

We generated Aven$^{-/-}$ HEK293T cells (clone #2) using CRISPR/Cas9 technology (Mali, Yang et al. 2013) and this was confirmed by immunoblotting (Figure 4.7A). A minor Aven ~34kDa fragment (Figure 4.7A, denoted by asterisk) was observed in HEK293T cells and likely represents a cathepsin D cleaved fragment reported previously (Melzer, Fernández et al. 2012). Aven depletion did not influence the number of ribosomes involved in polysomes, as

compared to control (Figure 4.7B), suggesting that general mRNA translation was not affected by Aven. However, a specific subset of mRNAs may still be particularly sensitive to changes in Aven levels. To investigate this possibility, we monitored the distribution of MLL1, MLL4 and β-actin mRNA in polysomal fractions from Aven-proficient and -deficient cells. Polysomal fractions were isolated and MLL1, MLL4 mRNAs as well as β-actin mRNA, were quantified by RT-qPCR. Loss of Aven expression reduced the amounts of MLL1 mRNA in the heavy polysomal fraction (fractions 12-15), with concomitant increase in lighter polysomal fractions (fractions 6 and 7; Figure 4.7C). Similarly, MLL4 mRNA was reduced in the heavy polysomal fraction (fractions 12-14) and a shift towards the light polysomal and pre-polysomal (fractions 7-8 and 10-11) (fractions 7-8). As a control, we monitored the levels of mRNAs encoding β-actin and depletion of Aven did not have a major effect on its distribution, inasmuch as the most of β-actin mRNA was associated with heavy polysomes (Figure 4.7E). These findings suggest that Aven selectively regulates the polysomal association of MLL1 and MLL4 mRNAs.

Since PRMT1 regulates the ability of Aven to associate with heavy polysomes, we investigated whether PRMT1-depleted cells also had reduced MLL1 and MLL4 mRNAs in heavy polysome fractions. Similarly to Aven, PRMT1 depletion (Figure 4.8A) did not influence the number of ribosomes involved in polysomes, as compared to a control, suggesting that PRMT1 does not affect global mRNA translation (Figure 4.8B). HEK293T cells were transfected with siGFP (siCTRL) or siPRMT1 for 72 h and the distribution of MLL1 and MLL4 mRNAs in polysomal fractions was monitored by RT-qPCR. Comparably to Aven, PRMT1 depletion had a striking effect on the distribution of MLL1 and MLL4 mRNAs in polysomes, as illustrated by their dramatic shift towards ligher fractions as compared to a control (Figure 4.8A, 4.8B). In contrast, depletion of PRMT1 resulted in a modest shift in β-actin mRNA, whereby the majority of β-actin mRNA remained in heavy polysome fractions. (Figure 4.8E, fractions 11-15). These findings suggest that Aven arginine methylation by PRMT1 regulates polysomal association of MLL1 and MLL4, but not β-actin mRNAs.


A
MOLT-4
shCTRL
shAven
MLL1
1.0
0.38
αMLL1
250
150
100
75
MLL4
1.0
0.1
αMLL4
50
37
25
αAven
α-actin
B
CCRF-CEM
shCTRL
shAven
MLL1
1.0
0.46
αMLL1
250
150
MLL4
100
75
1.0
0.18
αMLL4
50
37
25
αAven
α-actin
C
MOLT-4
shCTRL
shAven
mRNA (Relative fold change)
1.5
1.0
0.5
0.0
**
Actin
Aven
MLL1
MLL4
D
CCRF-CEM
shCTRL
shAven
mRNA (Relative fold change)
1.5
1.0
0.5
0.0
**
Actin
Aven
MLL1
MLL4
E
MOLT-4
shCTRL
shAven
mRNA (Relative fold change)
1.5
1.0
0.5
0.0
**
**
*
**
HOXA9
MEIS1
HOXA1
HOXC10
HOXA7
F
CCRF-CEM
shCTRL
shAven
mRNA (Relative fold change)
1.5
1.0
0.5
0.0
*
**
*
HOXA9
MEIS1
HOXA1
HOXC10
HOXA7

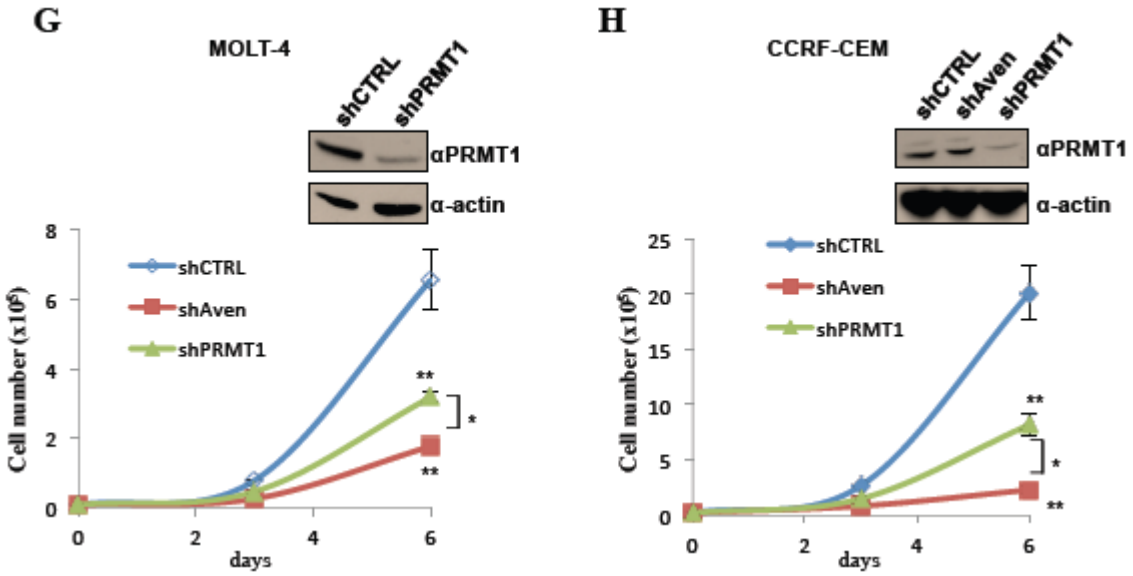


Figure 4.6 Aven regulates MLL1 and MLL4 protein expression required for leukemic cell survival.

A, B) Cellular lysates from MOLT-4 and CCRF-CEM cells stabling expressing shCTRL or shAven were separated by SDS-PAGE and immunoblotted with anti-MLL1, -MLL4, -Aven and β-actin antibodies. n = 3.

C-F) RT-qPCR of the indicated mRNAs was performed from RNA isolated from shCTRL and shAven MOLT-4 and CCRF-CEM cells and expressed as a relative fold change normalized to rpS6 levels. Error bars ± SEM is shown. The significance was measured by the Student t-test and defined as * $p < 0.05$, ** $p < 0.001$, n = 3.

G, H) Proliferation curves for shControl (CTRL), shPRMT1, and shAven MOLT-4 and CCRF-CEM cells are shown. Immunoblots confirm the depletion of PRMT1 in MOLT-4 and CCRF-CEM cells. Error bars ± standard deviation is shown. The data was analysed using ANOVA followed by post hoc comparison using Tukey test. * $p < 0.05$, ** $p < 0.001$, n = 3.

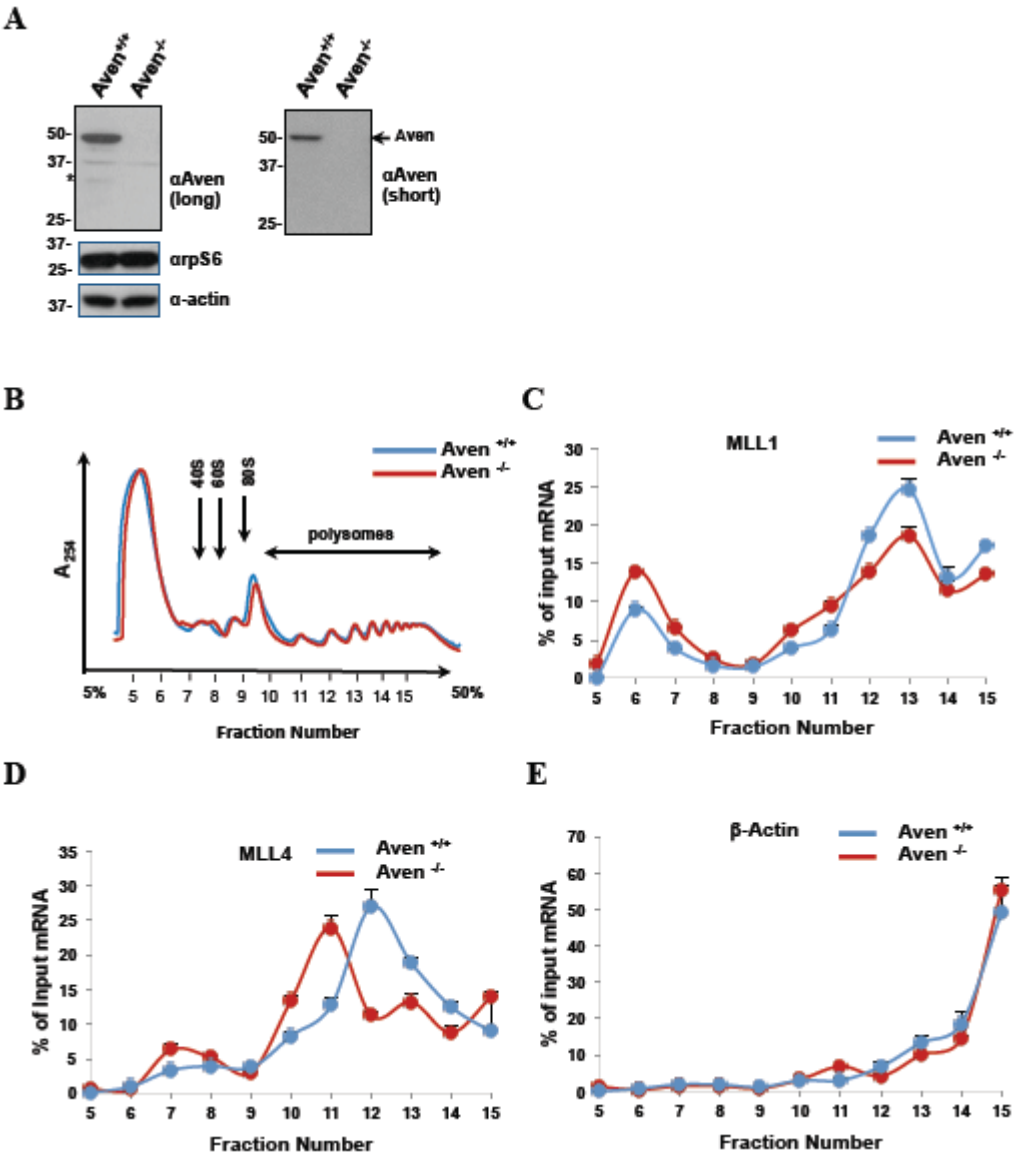


Figure 4.7 The Aven regulates polysomal association of MLL1 and MLL4, but not β-actin mRNA.

A) Aven-deficient HEK293T cells were generated by CRISPR/Cas9. Stable clones were obtained Aven$^{+/+}$ (clone #7) and Aven$^{-/-}$ (clone #2). Anti-Aven, anti-rpS6 and anti-β-actin immunoblots of total cell lysates are shown. The asterisks denotes a minor Aven species of lower molecular mass. The band at ~37kDa is a non-specific band recognized by the anti-Aven antibody. n = 3.

B) Polysome profiles of Aven$^{+/+}$ and Aven$^{-/-}$ HEK293T cells are shown. Cytoplasmic extracts from HEK293T cells were sedimented by centrifugation on a 5 to 50% sucrose gradient, shown as fraction numbers 5-15. Polysome profiles were obtained by continuous monitoring of UV absorbance at 254nm. 40S, 60S and 80S indicate the positions of the respective ribosomal subunits and the monomer on the gradient.

C-E) The indicated polysomal fractions were isolated, the RNA purified and the presence of MLL1, MLL4, or β-actin was quantified by qRT-PCR. mRNAs in each fraction is represented as the percentage of input. Error bars represent ± SEM, n = 5.

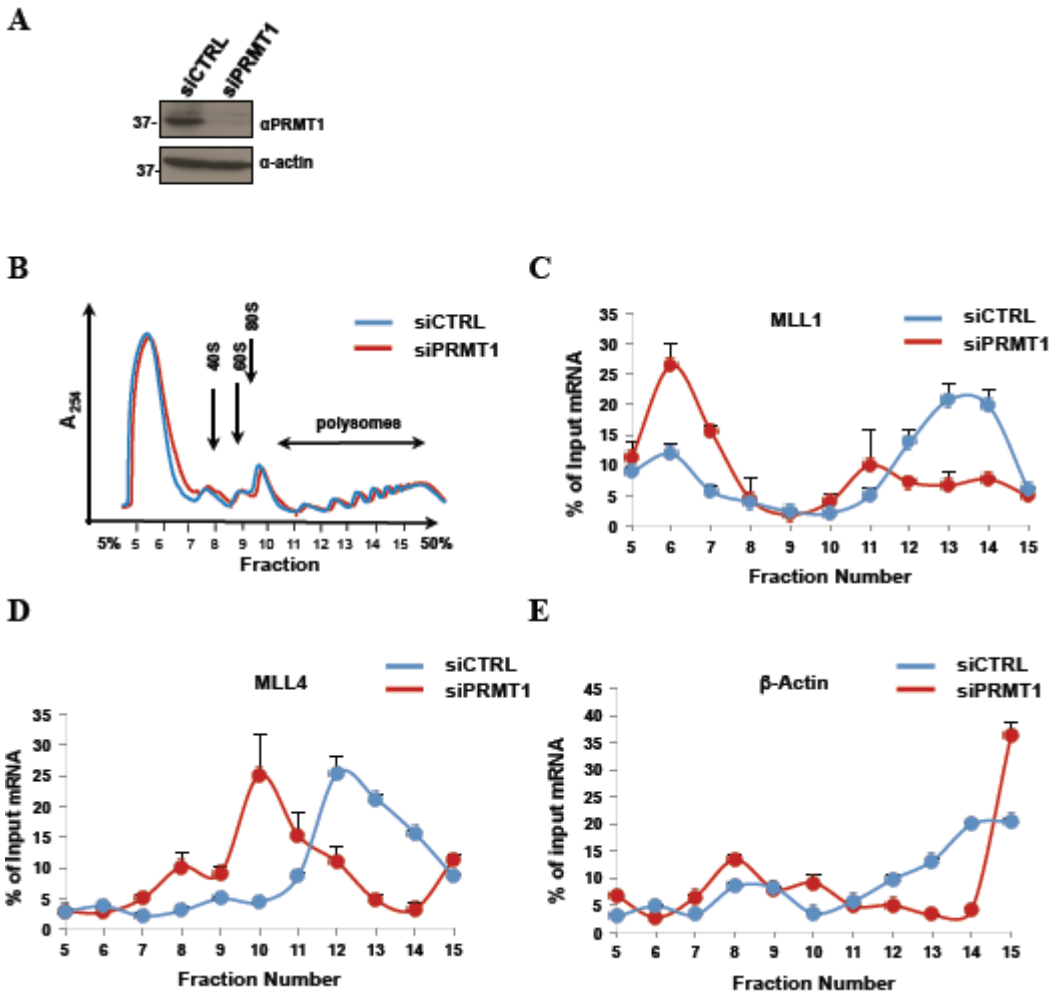


Figure 4.8 PRMT1 is required for the polysomal association of MLL1 and MLL4, but not β-actin mRNA.

A) PRMT1 was depleted by siRNA and cell extracts were immunoblotted with anti-PRMT1 or anti-β-actin antibodies.

B) Polysome profiles of siGFP (siCTRL) or siPRMT1 HEK293T cells are shown. Cytoplasmic extracts from HEK293T cells were sedimented by centrifugation on a 5 to 50% sucrose gradient, shown as fraction numbers 5-15. Polysome profiles were obtained by continuous monitoring of UV absorbance at 254nm. 40S, 60S and 80S indicate the positions of the respective ribosomal subunits and the monomer on the gradient.

C-E) The indicated polysomal fractions were isolated, the RNA purified and the presence of MLL1, MLL4, or β-actin was quantified by qRT-PCR. mRNAs in each fraction is represented as the percentage of input. Error bars represent ± SEM, n = 2.

4.5.8 DHX36 RNA helicase is required for Aven-dependent translation of RNAs

We closely examined the SILAC data for non-ribosomal proteins that are enriched in FLAG-Aven, but not with FLAG-AvenΔRGG immunoprecipitates that may function in unwinding G4 structures. The ATP-dependent RNA helicase DHX36, also known as G4 resolvase I, had an M/H ratio of 2.69. DHX36 has been reported to unwind G4 RNA structures (Creacy, Routh et al. 2008; Lattmann, Giri et al. 2010; Booy, Meier et al. 2012). We postulated that Aven could recruit the RNA helicases DHX36 to resolve G4 structures to facilitate protein synthesis. First, to validate whether the Aven RGG/RG motif is essential for interaction with the RNA helicase DHX36, we performed co-immunoprecipitations from cellular lysates expressing FLAG-Aven or FLAG-AvenΔRGG. The bound proteins were immunoblotted with anti-DHX36 antibodies. Indeed, DHX36 co-immunoprecipitated with FLAG-Aven, but not with FLAG-AvenΔRGG (Figure 4.9A). Moreover, we observed that DHX36 localized in the fast-sedimenting, heavier polysomal fractions with the control rpS6 (Figure 4.9B, fractions 12-15). To examine whether DHX36 influences the polysomal localization of MLL1 and MLL4 mRNAs, we monitored their mRNAs in polysomal fractions. Consistently with the observations for Aven and PRMT1, DHX36 depletion in HEK293T cells did not have a major effect on polysome absorbance profiles, thus indicating that DHX36 does not affect global protein synthesis (Figure 4.9C). However, both MLL1 and MLL4, but not β-actin mRNA shifted toward lighter polysomal fractions in cells depleted of DHX36, as compared to the control (Figure 4.9D-F). These findings suggest that similarly to PRMT1 and Aven, DHX36 regulates translation of MLL1 and MLL4 mRNAs.

4.5.9 MLL4 G4 structure requires the Aven RGG/RG motif and PRMT1 for optimal translation

We next determined whether the polysomal association of MLL4 mediated by Aven requires an intact G4 structure. Reporter mRNAs harboring the G4 of MLL4 or a mutated G4 motif was inserted in-frame with the open reading frame of luciferase (Figure 4.10A). We examined whether Aven could promote the translation of the luciferase reporter protein in a

G4-dependent manner and whether this was rescued by Aven re-expression in Aven-deficient cells. pGL3, pGL3-MLL4, or pGL3-MLL4 G4 mutant were transfected in Aven$^{-/-}$ HEK293T cells with pRenilla, as a control for transfection efficiency, along with either pcDNA3.1, FLAG-Aven, FLAG-AvenΔRGG, or FLAG-AvenR-K. The FLAG-AvenR-K protein was generated where the arginines in the RGG/RG motif were substituted for lysines to maintain the charge of the N-terminus of Aven. Twenty-four hours post-transfection, the cells were harvested for dual luciferase assay. The presence of the MLL4 G4 sequence inhibited the relative luciferase activity by >75% in Aven$^{-/-}$ HEK293T (Figure 4.10B, compare white and black bars labelled pcDNA3.1). The inhibition caused by the presence of the MLL4 G4 sequence was relieved by the transfection of FLAG-Aven, but not FLAG-AvenΔRGG nor FLAG-AvenR-K (Figure 4.10B, black bars). The presence of Aven did not have any significant effects on the luciferase expressed from pGL3 or pGL3-MLL4:G4mutant (Figure 4.10B). We next examined arginine methylation by PRMT1 and DHX36 were required for the stimulation of translation by FLAG-Aven. HEK293T cells were transfected with FLAG-Aven and pGL3-MLL4:G4 in the presence of siCTRL, siPRMT1 or siDHX36. The absence of PRMT1 or DHX36 blocked FLAG-Aven from stimulating translation (Figure 4.10C, 10D), suggesting that both arginine methylation of Aven and DHX36 helicase activity are required to regulate the translation of mRNAs with G4 structures within their ORFs.

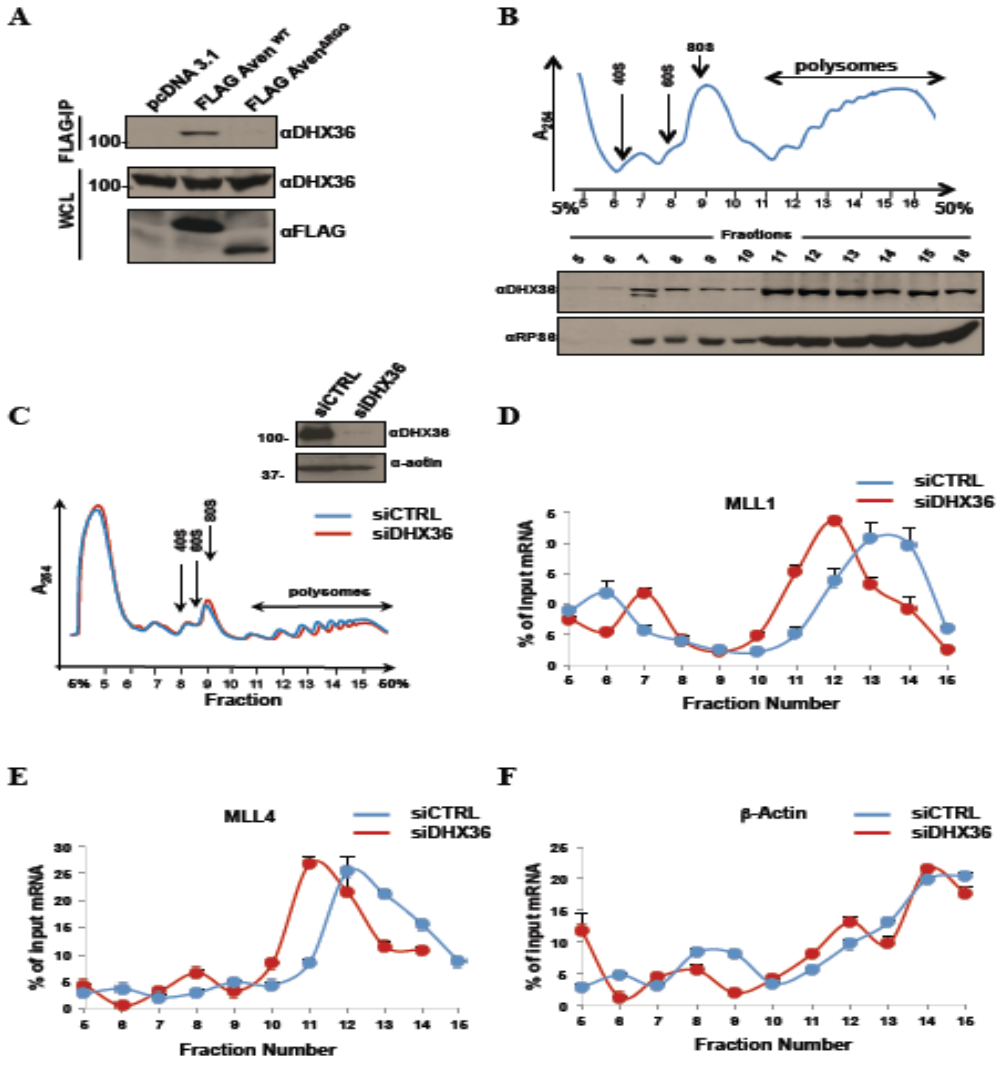


Figure 4.9 DHX36 is required for MLL1 and MLL4 mRNA polysomal association.

A) HEK293T cells expressing FLAG-Aven and the FLAG-AvenΔRGG were immunoprecipitated with anti-FLAG agarose beads and the bound proteins were immunoblotted with anti-DHX36 antibodies. TCL were immunoblotted with anti-DHX36 and anti-FLAG antibodies as indicated.

B) Proteins from the polysomal fractions isolated from HEK293T cells were TCA precipitated, separated by SDS-PAGE and immunoblotted with anti-DHX36 and anti-rpS6 antibodies. The experiment was performed n = 4 times and a typical polysomal profile is shown.

C) Immunoblots of total cell lysates from siGFP (CTRL) and siDHX36 transfected HEK293T cells is shown. Polysome profiles siCTRL and siDHX36 transfected HEK293T cells. Cytoplasmic extracts from HEK293T cells were sedimented by centrifugation on a 5 to 50% sucrose gradient, shown as fraction numbers 5 to 15. Polysome profiles were obtained by continuous monitoring of UV absorbance at 254nm. 40S, 60S and 80S indicate the positions of the respective ribosomal subunits and the monomer on the gradient.

D-F) The indicated polysomal fractions were isolated, total RNA isolated and the presence of MLL1, MLL4 or β-actin was quantified by RT-qPCR in triplicates. mRNAs in each fraction is represented as the percentage of input. Error bars represent ± SEM, n=2.

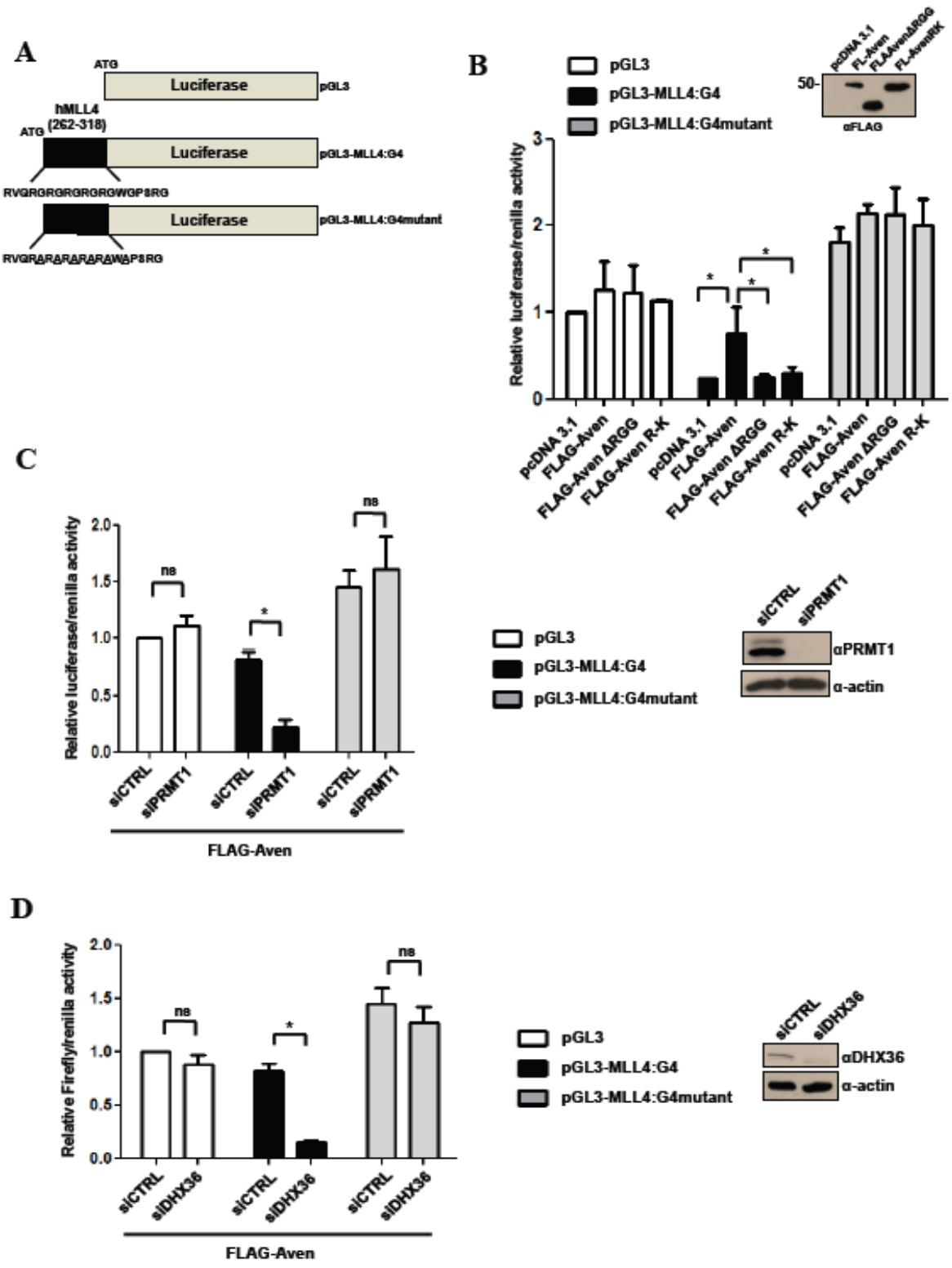


Figure 4.10 PRMT1 and Aven RGG/RG motif required for optimal translation of MLL4 G4 sequence.

A) Schematic of the luciferase reporter plasmid pGL3, as well as the chimeric pGL3-MLL4-G4 and pGL3-MLL4-G4mutant. pGL3-MLL4-G4 harbors the human MLL4 G4 sequence nucleotide 262 to 318 inserted in-frame at the N-terminus of luciferase, while pGL3-MLL4-G4mutant contains glycine to alanine mutations that disrupts the G4

structure.
B) Aven$^{-/-}$ HEK293T cells were transfected with the following reporter genes pGL3, pGL3-MLL4-G4, or pGL3-MLL4-G4mutant and pRenilla as well as pcDNA3.1, FLAG-Aven, FLAG-AvenΔRGG, or FLAG-AvenR-K. The cells were harvested 24 h post-transfection and dual luciferase assays were performed. The relative luciferase/Renilla ratio was normalized to 1.0 in pGL3 pcDNA3.1 transfected cells. Extracts were immunoblotted with anti-FLAG antibodies to confirm Aven, AvenΔRGG, or AvenR-K expression. Error bars represent standard deviation values. The experiments were performed three independent times (n = 3) and each independent experiment was performed in technical triplicates. The significance was measured by ANOVA followed by post hoc comparison using Tukey test. *$p < 0.05$.
C) HEK293T cells were co-transfected with FLAG-Aven and either siGFP (siCTRL) or siPRMT1 along with the following reporter genes pGL3, pGL3-MLL4-G4, or pGL3-MLL4-G4mutant and pRenilla. The cells were harvested 24 h post-transfection and dual luciferase assays were performed. The relative luciferase/Renilla ratio was normalized to 1.0 in pGL3 siCTRL transfected cells. Extracts were immunoblotted with anti-PRMT1 or anti-□-actin antibodies, as indicated. Error bars represent standard deviation values. The experiments were performed three independent times (n = 3) and each independent experiment was performed in technical triplicates. The significance was measured by ANOVA followed by post hoc comparison using Tukey test. *$p < 0.05$, n.s. non-significant.
D) HEK293T cells were co-transfected with FLAG-Aven and either siGFP (siCTRL) or siDHX36 along with the following reporter genes pGL3, pGL3-MLL4-G4, or pGL3-MLL4-G4mutant and pRenilla. The cells were harvested 24 h post-transfection and dual luciferase assays were performed. The relative luciferase/Renilla ratio was normalized to 1.0 in pGL3 siCTRL transfected cells. Extracts were immunoblotted with anti-DHX36 or anti-β-actin antibodies, as indicated. The error bars represent ± the standard deviation. Experiments were performed three times (n = 3) and each experiment was analyzed in triplicates. Significance was measured by the Student t-test. *$p < 0.05$, n.s. non-significant.

4.6 Discussion

The mixed lineage leukemia (*MLL*) proto-oncogene encodes a histone methyltransferase implicated in epigenetic modifications, regulating gene expression for embryonic development and hematopoiesis (Liedtke and Cleary 2009; Smith, Lin et al. 2011). *MLL* is a recurrent site of DNA translocations resulting in an MLL fusion protein where the N-terminus of the *MLL* is fused to a variety of proteins (Liedtke and Cleary 2009; Smith, Lin et al. 2011). In the present manuscript, we identify, within the mRNA coding regions of MLL1 and MLL4, RNA elements that regulate its polysomal association and protein synthesis. These RNA elements are located between 200 and 300 nucleotides downstream of the initiator methionine ATG, and encode protein sequences rich in glycines and arginine-glycine repeats in MLL1 and MLL4. The function of these N-terminal repeats are unknown. We show that Aven binds the MLL1 and MLL4 G4 RNA structures *in vitro* and *in vivo* with its RGG/RG motif. Aven was required for the translational regulation of MLL1 and MLL4, as Aven-deficient T-ALL cells exhibited decreased MLL1 and MLL4 protein expression and consequently decreased the expression of their downstream targets including, the *HOX* genes. The association of Aven with polysomes required the methylation of its RGG/RG motif by PRMT1 and interaction with methyl-binding proteins, TDRD3 and SMN. The Aven interaction with TDRD3 and SMN may require other protein or RNA components in the complex for enhanced association. Deficiency of Aven or PRMT1 in acute leukemic cell lines led to decreased cell proliferation. Taken together, our studies suggest that Aven regulates the translation of MLL1 and MLL4 required cancer survival and that targeting this pathway may have therapeutic potential.

RGG/RG motifs have the biochemical properties to bind both RNA and proteins to fulfill their emerging roles in assembly of RNP complexes and translational control (Rajyaguru and Parker 2012; Thandapani, O'Connor et al. 2013). The RGG/RG motif of yeast proteins Scd6, Npl3 and Sbp1 was shown to interact with the translational initiation factor eIF4G and repress translation by preventing the formation of pre-initiation complex (Rajyaguru, She et al. 2012). In trypanosomes, the RGG/RG motif of SCD6 is involved in regulating the type and number of RNP granules (Krüger, Hofweber et al. 2013). Amyloid-like fibers were

formed when the RGG/RG motif of FUS was incubated with RNA (Schwartz, Wang et al. 2013). These fibers are characterised by the reversible transformation from soluble to polymeric state (Han, Kato et al. 2012; Kato, Han et al. 2012). Although many proteins have an RGG/RG motif, the Aven RGG/RG motif may be more accessible, as it is located at the N-terminus and may protrude outwards. Since the Aven RGG/RG motif is not required for self-association, this suggests that an Aven dimer has 2 protruding RGG/RG motifs that can each mediate their own interactions. Therefore, we speculate that Aven functions as a scaffolding protein to assemble translationally competent RNPs for certain mRNAs containing G4 motifs (Figure 4.11).

DHX36 is a DEAH (aspartic acid, glutamic acid, alanine, histidine)-box helicase and it is the only G4 RNA resolvase known and is a major DNA G4 resolvase (Creacy, Routh et al. 2008; Lattmann, Giri et al. 2010). Aven associated with DHX36 to regulate translation of mRNAs with G4 structures. DHX36 knockdown increased the expression of PITX1 protein without changes in mRNA, suggesting that it functions in translational control (Booy, Howard et al. 2014). Ribosomal footprinting studies have led to the proposal that elongating ribosomes likely use accessory RNA helicases (Rouskin, Zubradt et al. 2014), and our data suggest that DHX36 may be such an accessory helicase. DHX36 null mice are embryonic lethal and deletion in the hematopoietic system using *Vav1-Cre* causes hemolytic anemia and defects at the proerythroblast stage with deregulation of genes with G4 motifs in their promoters (Lai, Ponti et al. 2012), however, a role DHX36 in translational control was not examined.

Many PRMT1 substrates are RBPs with RGG/RG motifs (Bedford and Clarke 2009) and some have been shown to associate with RNAs with G4 motifs such as Nucleolin, FUS, EWS, and FMRP (Thandapani, O'Connor et al. 2013). This suggests that several RBPs likely function in a similar manner to Aven in regulating accessibility of mRNPs with polysomes. It has been shown that the RGG/RG motif of FMRP is required for its polysomal association (Blackwell, Zhang et al. 2010), however, whether arginine methylation by PRMT1 regulates association is unknown. Our findings show for the first time that arginine methylation by PRMT1 regulates translational control. It is known, however, that the yeast homolog of

PRMT3 (RMT3) methylates rpS2, regulating the balance between the small and large ribosomal subunits (Bachand and Silver 2004). However, mammalian PRMT3 did not influence ribosomal assembly or polysomal formation (Swiercz, Person et al. 2004).

It is known that secondary RNA structure including G-quadruplex structures within mRNAs hinder mRNA translation (Koromilas, Lazaris-Karatzas et al. 1992; Sonenberg and Hinnebusch 2009). Stable RNA secondary structures within the 5′-UTRs of mRNAs reduce cap-dependent translation by preventing assembly of the translational initiation machinery at the 5′-cap and also impair the scanning of the start site AUG by the initiation complex (Beaudoin and Perreault 2010; Bugaut and Balasubramanian 2012). Secondary structure in the 5'UTRs including G4 motifs have been shown to require eIF4A for optimal translation output (Wolfe, Singh et al. 2014). The 5′-UTR of NRAS and Zic-1, which harbour G4 structures, reduce translation of a reporter luciferase (Kumari, Bugaut et al. 2007; Arora, Dutkiewicz et al. 2008). G-quadruplex structures within ORFs of the virally encoded EBNA1 transcript were shown to hinder translational elongation by either ribosomal pausing or ribosomal dissociation (Murat, Zhong et al. 2014). We now extend these observations and identify a mechanism regulated by arginine methylation that leads to the positive regulation of mRNAs with G4 structures within their coding region.

Aven is overexpressed in acute leukemia and was proposed to be a prognostic factor in acute childhood lymphoblastic leukemia for poor outcome (Choi, Hwang et al. 2006). Aven is a well-established cell survival protein or inhibitor of apoptosis that prevents apoptosis by stabilizing pro-survival protein Bcl-x_L and inhibiting the function of pro-apoptotic protein Apaf-1 (Chau, Cheng et al. 2000; Kutuk, Temel et al. 2010). It was reported that an N-terminal deleted fragment of Aven cleaved by cathepsin D harbors its anti-apoptotic function (Melzer, Fernández et al. 2012), however, such a ~30kDa Aven species was not visible in MOLT4 and CCRF-CEM cells (Figure 4.6A, 6B), and was faintly observed in HEK293T (Figure 4.7A) as previously described (Melzer, Fernández et al. 2012). Thus cathepsin D mediated cleavage of Aven is unlikely involved in the regulation of translational control described herein. In addition to its pro-survival functions, Aven was identified to be essential for progression of acute leukemia in mice (Eismann, Melzer et al. 2013). The

regions required for association with Bcl-x_L and Apaf-1 reside C-terminal of the RGG/RG motif. Taken together with our findings, this suggests that Aven uses several mechanisms to increase cell survival, 1) preventing apoptosis via Bcl-x_L and Apaf-1, and 2) favouring the translation of mRNAs, including those encoding MLL1 and MLL4 required for cell survival.

PRMT1 was shown to be essential for mixed lineage leukemia by the *MLL-EEN* gene fusion protein (Cheung, Chan et al. 2007). The *EEN* fusion partner leads to the recruitment of PRMT1 to methylate histones and lead to gene activation (Cheung, Chan et al. 2007). Our findings identify a new role for PRMT1 in the cytoplasm that is required for cancer cell survival. This pathway is amenable to therapeutic intervention with future PRMT1 inhibitors and specific RNA G-quadruplex ligands.

4.7 Aknowledgements

We thank Janane Maheswaran and André Galarneau for their contributions on the initial observations to the Aven project. We acknowledge Koren Mann and Josée Hébert for critically reading the manuscript and helpful discussions. This work was supported by a grant from the Canadian Institute of Health Canada (MOP-67070) to S. R. and the Chaire de recherche de l'Université de Sherbrooke en strucuture de l'ARN et génomiquie to J.P.P. P. T. was funded, in part, by a James O. and Maria Meadows studentship and with funds from the Fonds de recherche du Québec - Santé (FRQS) to S.R. S. G. R. holds an NSERC Fellowship.

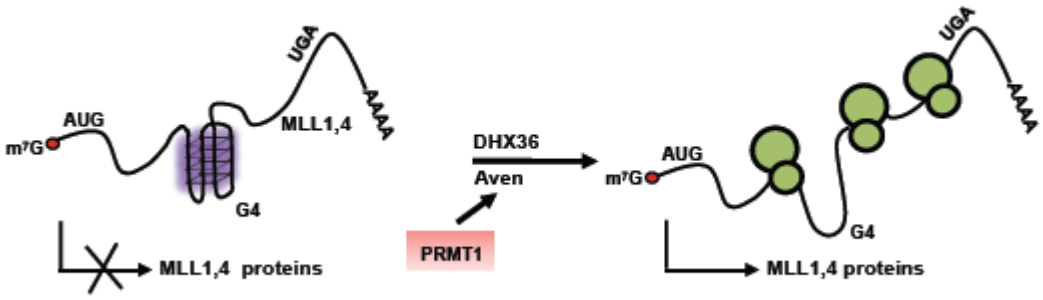


Figure 4.11 Model denoting the role of arginine methylated Aven by PRMT1 and DHX36 in the translation of G4 harboring MLL1 and MLL4 proteins.

4.8 References

An, W., Kim, J., and Roeder, R.G. (2004). Ordered cooperative functions of PRMT1, p300, and CARM1 in transcriptional activation by p53. Cell 117, 735-748.
Arora, A., Dutkiewicz, M., Scaria, V., Hariharan, M., Maiti, S., and Kurreck, J. (2008). Inhibition of translation in living eukaryotic cells by an RNA G-quadruplex motif. RNA (New York, NY) 14, 1290-1296.
Bachand, F., and Silver, P.A. (2004). PRMT3 is a ribosomal protein methyltransferase that affects the cellular levels of ribosomal subunits. EMBO J 23, 2641-2650.
Baltz, A.G., Munschauer, M., Schwanhäusser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M., et al. (2012). The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol Cell 46, 674-690.
Baranski, Z., Booij, T.H., Cleton-Jansen, A.M., Price, L.S., van de Water, B., Bovée, J.V., Hogendoorn, P.C., and Danen, E.H. (2015). Aven-mediated checkpoint kinase control regulates proliferation and resistance to chemotherapy in conventional osteosarcoma. J Pathol Mar 9. doi: 10.1002/path.4528.
Bates, P., Mergny, J.L., and Yang, D. (2007). Quartets in G-major. The First International Meeting on Quadruplex DNA. EMBO Rep 8, 1003-1010.
Beaudoin, J.D., Jodoin, R., and Perreault, J.P. (2013). In-line probing of RNA G-quadruplexes. Methods 64, 79-87.
Beaudoin, J.D., Jodoin, R., and Perreault, J.P. (2014). New scoring system to identify RNA G-quadruplex folding. Nucleic acids research 42, 1209-1223.
Beaudoin, J.D., and Perreault, J.P. (2010). 5'-UTR G-quadruplex structures acting as translational repressors. Nucleic acids research 38, 7022-7036.
Beaudoin, J.D., and Perreault, J.P. (2013). Exploring mRNA 3'-UTR G-quadruplexes: evidence of roles in both alternative polyadenylation and mRNA shortening. Nucleic acids research 41, 5898-5911.
Bedford, M.T., and Clarke, S.G. (2009). Protein arginine methylation in mammals: who, what, and why. Mol Cell 33, 1-13.
Bedford, M.T., and Richard, S. (2005). Arginine methylation an emerging regulator of protein function. Mol Cell 18, 263-272.
Blackwell, E., Zhang, X., and Ceman, S. (2010). Arginines of the RGG box regulate FMRP association with polyribosomes and mRNA. Human Molecular Genetics 19, 1314-1323.
Blagoev, B., Kratchmarova, I., Ong, S.E., Nielsen, M., Foster, L.J., and Mann, M. (2003). A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. Nat Biotechnol 21, 315-318.
Blobel, G., and Sabatini, D. (1971). Dissociation of mammalian polyribosomes into subunits by puromycin. Proc Natl Acad Sci U S A 68, 390-394.
Boisvert, F.-M., Déry, U., Masson, J.-Y., and Richard, S. (2005). Arginine methylation of MRE11 by PRMT1 is required for DNA damage checkpoint control. Genes & Dev 19, 671-676.
Boisvert, F.M., Ahmad, Y., Gierliński, M., Charrière, F., Lamont, D., Scott, M., Barton, G.,

and Lamond, A.I. (2012). A quantitative spatial proteomics analysis of proteome turnover in human cells. Mol Cell Proteomics 22, M111.011429.
Booy, E.P., Howard, R., Marushchak, O., Ariyo, E.O., Meier, M., Novakowski, S.K., Deo, S.R., Dzananovic, E., Stetefeld, J., and McKenna, S. (2014). The RNA helicase RHAU (DHX36) suppresses expression of the transcription factor PITX1. Nucleic acids research 42, 3346-3361.
Booy, E.P., Meier, M., Okun, N., Novakowski, S.K., Xiong, S., Stetefeld, J., and McKenna, S.A. (2012). The RNA helicase RHAU (DHX36) unwinds a G4-quadruplex in human telomerase RNA and promotes the formation of the P1 helix template boundary. Nucleic acids research 40, 4110-4124.
Bugaut, A., and Balasubramanian, S. (2012). 5'-UTR RNA G-quadruplexes: translation regulation and targeting. Nucleic acids research 40, 4727-4741.
Burge, S., Parkinson, G.N., Hazel, P., Todd, A.K., and Neidle, S. (2006). Quadruplex DNA: sequence, topology and structure. Nucleic acids research 34, 5402-5415.
Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M., et al. (2012). Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell 149, 1393-1406.
Castello, A., Fischer, B., Hentze, M.W., and Preiss, T. (2013). RNA-binding proteins in Mendelian disease. Trends Genet 29, 318-327.
Chau, B.N., Cheng, E.H., Kerr, D.A., and Hardwick, J.M. (2000). Aven, a novel inhibitor of caspase activation, binds Bcl-xL and Apaf-1. Mol Cell 6, 31-40.
Chen, C., Nott, T.J., Jin, J., and Pawson, T. (2011). Deciphering arginine methylation: Tudor tells the tale. Nat Rev Mol Cell Biol 12, 629-642.
Chen, Y., and Varani, G. (2013). Engineering RNA-binding proteins for biology. FEBS J 280, 3734-3754.
Cheung, N., Chan, L.C., Thompson, A., Cleary, M.L., and So, C.W. (2007). Protein arginine-methyltransferase-dependent oncogenesis. Nat Cell Biol 9, 1208-1215.
Choi, J., Hwang, Y.K., Sung, K.W., Kim, D.H., Yoo, K.H., Jung, H.L., and Koo, H.H. (2006). Aven overexpression: association with poor prognosis in childhood acute lymphoblastic leukemia. Leukemia research 30, 1019-1025.
Cooper, T.A., Wan, L., and Dreyfuss, G. (2009). RNA and disease. Cell 136, 777-793.
Cote, J., and Richard, S. (2005). Tudor domains bind symmetrical dimethylated arginines. J Biol Chem 280, 28476-28483.
Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26, 1367-1372.
Creacy, S.D., Routh, E.D., Iwamoto, F., Nagamine, Y., Akman, S.A., and Vaughn, J.P. (2008). G4 resolvase 1 binds both DNA and RNA tetramolecular quadruplex with high affinity and is the major source of tetramolecular quadruplex G4-DNA and G4-RNA resolving activity in HeLa cell lysates. J Biol Chem 283, 34626-34634.
Darnell, J.C., Jensen, K.B., Jin, P., Brown, V., Warren, S.T., and Darnell, R.B. (2001). Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function. Cell 107, 489-499.
Eismann, M., Melzer, I.M., Fernandez, S.B.M., Michel, G., Hrabe de Angelis, M., Hoefler, G.,

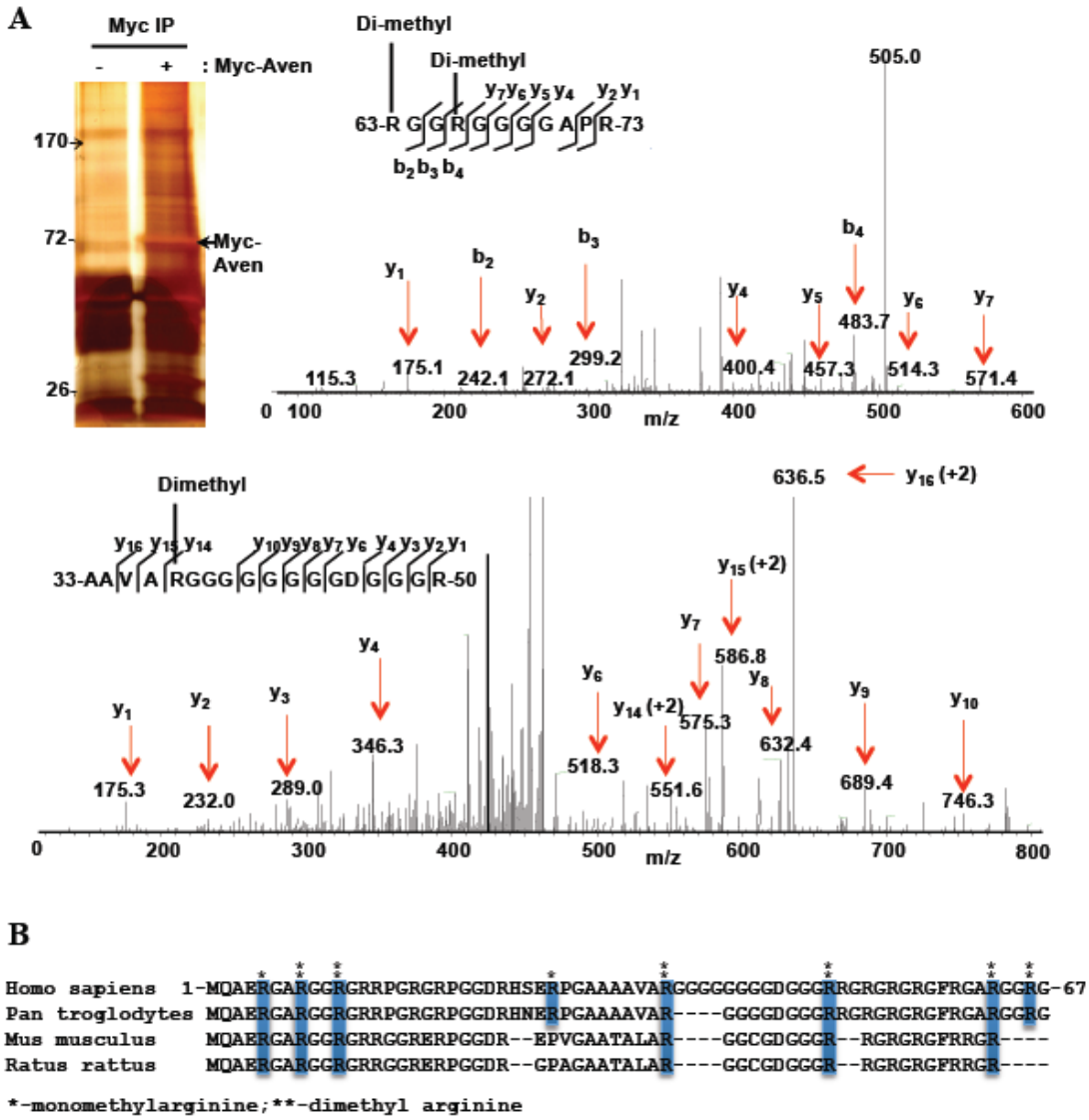
Finkenwirth, P., Jauch, A., Schoell, B., Grez, M., et al. (2013). Overexpression of the anti-apoptotic protein AVEN contributes to increased malignancy in hematopoietic neoplasms. Oncogene 32, 2586-2591.
Esmaili, A.M., Johnson, E.L., Thaivalappil, S.S., Kuhn, H.M., Kornbluth, S., and Irusta, P. (2010). Regulation of the ATM-activator protein Aven by CRM1-dependent nuclear export. Cell Cycle 9, 3913-3920.
Gandin, V., Gutierrez, G.J., Brill, L., Varsano, T., Feng, Y., Aza-Blanc, P., Au, Q., McLaughlan, S., Ferreira, T.A., Alain, T., et al. (2013). Degradation of newly synthesized polypeptides by ribosome-associated RACK1/c-Jun N-terminal kinase/eukaryotic elongation factor 1A2 complex. Mol Cell Biol 33, 2510-2526.
Gandin, V., Sikström, K., Alain, T., Morita, M., McLaughlan, S., Larsson, O., and Topisirovic, I. (2014). Polysome fractionation and analysis of mammalian translatomes on a genome-wide scale. J Vis Exp May 17 (87), doi: 10.3791/51455.
Glisovic, T., Bachorik, J.L., Yong, J., and Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. FEBS Lett 582, 1977-1986.
Goulet, I., Boisvenue, S., Mokas, S., Mazroui, R., and Côté, J. (2008). TDRD3, a novel Tudor domain-containing protein, localizes to cytoplasmic stress granules. Human Molecular Genetics 17, 3055-3074.
Guo, J.Y., Yamada, A., Kajino, T., Wu, J.Q., Tang, W., Freel, C.D., Feng, J., Chau, B.N., Wang, M.Z., Margolis, S.S., et al. (2008). Aven-Dependent Activation of ATM Following DNA Damage. Current Biology 18, 933-942.
Haeusler, A.R., Donnelly, C.J., Periz, G., Simko, E.A., Shaw, P.G., Kim, M.S., Maragakis, N.J., Troncoso, J.C., Pandey, A., Sattler, R., et al. (2014). C9orf72 nucleotide repeat structures initiate molecular cascades of disease. Nature 507, 195-200.
Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jungkamp, A.C., Munschauer, M., et al. (2010). PAR-CliP--a method to identify transcriptome-wide the binding sites of RNA-binding proteins. J Vis Exp 2, 2034.
Han, T.W., Kato, M., Xie, S., Wu, L.C., Mirzaei, H., Pei, J., Chen, M., Xie, Y., Allen, J., Xiao, G., et al. (2012). Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies. Cell 149, 768-779.
Hawley, R.G., Chen, Y., Riz, I., and Zeng, C. (2012). An Integrated Bioinformatics and Computational Biology Approach Identifies New BH3-Only Protein Candidates. Open Biol J 5, 6-16.
Huppertz, I., Attig, J., D'Ambrogio, A., Easton, L.E., Sibley, C.R., Sugimoto, Y., Tajnik, M., König, J., and Ule, J. (2014). iCLIP: protein-RNA interactions at nucleotide resolution. Methods 65, 274-287.
Jodoin, R., Bauer, L., Garant, J.M., Mahdi Laaref, A., Phaneuf, F., and Perreault, J.P. (2014). The folding of 5'-UTR human G-quadruplexes possessing a long central loop. RNA (New York, NY) 20, 1129-1141.
Kato, M., Han, T.W., Xie, S., Shi, K., Du, X., Wu, L.C., Mirzaei, H., Goldsmith, E.J., Longgood, J., Pei, J., et al. (2012). Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. Cell 149, 753-767.
Kiledjian, M., and Dreyfuss, G. (1992). Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box. EMBO J 11, 2655-2664.

Koromilas, A.E., Lazaris-Karatzas, A., and 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.
Krivtsov, A.V., and Armstrong, S.A. (2007). MLL translocations, histone modifications and leukaemia stem-cell development. Nat Rev Cancer 7, 823-833.
Krüger, T., Hofweber, M., and Kramer, S. (2013). SCD6 induces ribonucleoprotein granule formation in trypanosomes in a translation-independent manner, regulated by its Lsm and RGG domains. Mol Biol Cell 24, 2098-2111.
Kumari, S., Bugaut, A., Huppert, J.L., and Balasubramanian, S. (2007). An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation. Nature chemical biology 3, 218-221.
Kutuk, O., Temel, S.G., Tolunay, S., and Basaga, H. (2010). Aven blocks DNA damage-induced apoptosis by stabilising Bcl-xL. Eur J Cancer 46, 2494-2505.
Kwon, S.C., Yi, H., Eichelbaum, K., Föhr, S., Fischer, B., You, K.T., Castello, A., Krijgsveld, J., Hentze, M.W., and Kim, V.N. (2013). The RNA-binding protein repertoire of embryonic stem cells. Nat Struct Mol Biol 20, 1122-1130.
Lai, J.C., Ponti, S., Pan, D., Kohler, H., Skoda, R.C., Matthias, P., and Nagamine, Y. (2012). The DEAH-box helicase RHAU is an essential gene and critical for mouse hematopoiesis. Blood 119, 4291-4300.
Lattmann, S., Giri, B., Vaughn, J.P., Akman, S.A., and Nagamine, Y. (2010). Role of the amino terminal RHAU-specific motif in the recognition and resolution of guanine quadruplex-RNA by the DEAH-box RNA helicase RHAU. Nucleic acids research 38, 6219-6233.
Liedtke, M., and Cleary, M.L. (2009). Therapeutic targeting of MLL. Blood 113, 6061-6068.
Lorenz, R., Bernhart, S.H., Qin, J., Höner zu Siederdissen, C., Tanzer, A., Amman, F., Hofacker, I.L., and Stadler, P.F. (2013). 2D meets 4G: G-quadruplexes in RNA secondary structure prediction. IEEE/ACM Trans Comput Biol Bioinform 10, 832-844.
Lukong, K.E., Chang, K.W., Khandjian, E.W., and Richard, S. (2008). RNA-binding proteins in human genetic disease. Trends Genet 24, 416-425.
Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science 339, 823-826.
Melzer, I.M., Fernández, S.B., Bösser, S., Lohrig, K., Lewandrowski, U., Wolters, D., Kehrloesser, S., Brezniceanu, M.L., Theos, A.C., Irusta, P.M., et al. (2012). The Apaf-1-binding protein Aven is cleaved by Cathepsin D to unleash its anti-apoptotic potential. Cell Death Differ 19, 1435-1445.
Mitchell, S.F., and Parker, R. (2014). Principles and properties of eukaryotic mRNPs. Mol Cell 54, 547-558.
Modelska, A., Turro, E., Russell, R., Beaton, J., Sbarrato, T., Spriggs, K., Miller, J., Gräf, S., Provenzano, E., Blows, F., et al. (2015). The malignant phenotype in breast cancer is driven by eIF4A1-mediated changes in the translational landscape. Cell Death Dis 6, e1603.
Mukundan, V.T., and Phan, A.T. (2013). Bulges in G-quadruplexes: broadening the definition of G-quadruplex-forming sequences. Journal of the American Chemical Society 135, 5017-5028.
Murat, P., Zhong, J., Lekieffre, L., Cowieson, N.P., Clancy, J.L., Preiss, T., Balasubramanian,

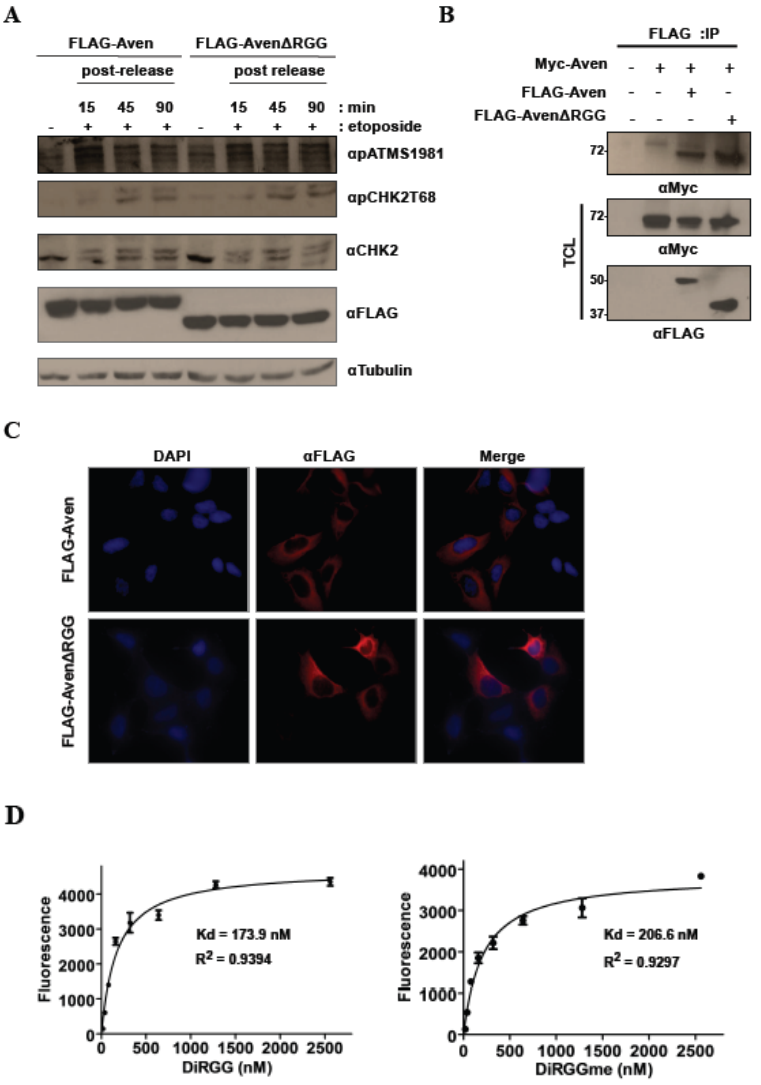
S., Khanna, R., and Tellam, J. (2014). G-quadruplexes regulate Epstein-Barr virus-encoded nuclear antigen 1 mRNA translation. Nature chemical biology 10, 358-364.
Pawlak, M.R., Scherer, C.A., Chen, J., Roshon, M.J., and Ruley, H.E. (2000). Arginine N-Methyltransferase 1 Is Required for Early Postimplantation Mouse Development, but Cells Deficient in the Enzyme Are Viable. Mol Cell Biol 20, 4859-4869.
Paydas, S., Tanriverdi, K., Yavuz, S., Disel, U., Sahin, B., and Burgut, R. (2003). Survivin and aven: two distinct antiapoptotic signals in acute leukemias. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO 14, 1045-1050.
Phan, A.T., Kuryavyi, V., Darnell, J.C., Serganov, A., Majumdar, A., Ilin, S., Raslin, T., Polonskaia, A., Chen, C., Clain, D., et al. (2011). Structure-function studies of FMRP RGG peptide recognition of an RNA duplex-quadruplex junction. Nat Struct Mol Biol 18, 796-804.
Piques, M., Schulze, W.X., Höhne, M., Usadel, B., Gibon, Y., Rohwer, J., and Stitt, M. (2009). Ribosome and transcript copy numbers, polysome occupancy and enzyme dynamics in Arabidopsis. Mol Syst Biol 5, 314.
Rajyaguru, P., and Parker, R. (2012). RGG motif proteins: Modulators of mRNA functional states. Cell Cycle 11, 2594-2599.
Rajyaguru, P., She, M., and Parker, R. (2012). Scd6 Targets eIF4G to Repress Translation: RGG Motif Proteins as a Class of eIF4G-Binding Proteins. Mol Cell 45, 244-254.
Ramaswami, M., Taylor, J.P., and Parker, R. (2013). Altered ribostasis: RNA-protein granules in degenerative disorders. Cell 154, 727-736.
Rouskin, S., Zubradt, M., Washietl, S., Kellis, M., and Weissman, J.S. (2014). Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. Nature 505, 701-705.
Rubio, C.A., Weisburd, B., Holderfield, M., Arias, C., Fang, E., DeRisi, J.L., and Fanidi, A. (2014). Transcriptome-wide characterization of the eIF4A signature highlights plasticity in translation regulation. Genome Biol 15, 476.
Sanchez, G., Dury, A.Y., Murray, L.M., Biondi, O., Tadesse, H., El Fatimy, R., Kothary, R., Charbonnier, F., Khandjian, E.W., and Côté, J. (2013). A novel function for the survival motoneuron protein as a translational regulator. Hum Mol Genet 22, 668-684.
Schwartz, J.C., Wang, X., Podell, E.R., and Cech, T.R. (2013). RNA seeds higher-order assembly of FUS protein. Cell Rep 5, 918-925.
Selenko, P., Sprangers, R., Stier, G., Buhler, D., Fischer, U., and Sattler, M. (2001). SMN tudor domain structure and its interaction with the Sm proteins. Nat Struct Biol 8, 27-31.
Smith, E., Lin, C., and Shilatifard, A. (2011). The super elongation complex (SEC) and MLL in development and disease. Genes Dev 25, 661-672.
Sonenberg, N., and Hinnebusch, A.G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. 136, 731-745.
Strahl, B.D., Briggs, S.D., Brame, C.J., Caldwell, J.A., Koh, S.S., Ma, H., Cook, R.G., Shabanowitz, J., Hunt, D.F., Stallcup, M.R., et al. (2001). Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1. Curr Biol 11, 996-1000.
Swiercz, R., Person, M.D., and Bedford, M.T. (2004). Ribosomal protein S2 is a substrate for mammalian protein arginine methyltransferase 3 (PRMT3). Biochem J Oct 8; Epub ahead of print.

Takahama, K., Kino, K., Arai, S., Kurokawa, R., and Oyoshi, T. (2011). Identification of Ewing's sarcoma protein as a G-quadruplex DNA- and RNA-binding protein. FEBS Journal 278, 988-998.
Takahama, K., Takada, A., Tada, S., Shimizu, M., Sayama, K., Kurokawa, R., and Oyoshi, T. (2013). Regulation of telomere length by G-quadruplex telomere DNA- and TERRA-binding protein TLS/FUS. Chemistry & biology 20, 341-350.
Thandapani, P., O'Connor, T.R., Bailey, T.L., and Richard, S. (2013). Defining the RGG/RG motif. Mol Cell 50, 613-623.
Wang, H., Huang, Z.-Q., Xia, L., Feng, Q., Erdjument-Bromage, H., Strahl, B.D., Briggs, S.D., Allis, C.D., Wong, J., Tempst, P., et al. (2001). Methylation of histone H4 at arginine 3 facilitates transcriptional activation by nuclear hormone receptor. Science 293, 853-857.
Wolfe, A.L., Singh, K., Zhong, Y., Drewe, P., Rajasekhar, V.K., Sanghvi, V.R., Mavrakis, K.J., Jiang, M., Roderick, J.E., Van der Meulen, J., et al. (2014). RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer. Nature Jul 27, doi: 10.1038/nature13485.
Xu, J., Wang, A.H., Oses-Prieto, J., Makhijani, K., Katsuno, Y., Pei, M., Yan, L., Zheng, Y.G., Burlingame, A., Brückner, K., et al. (2013). Arginine Methylation Initiates BMP-Induced Smad Signaling. Mol Cell 51, 5-19.
Yu, Z., Chen, T., Hébert, J., Li, E., and Richard, S. (2009). A mouse PRMT1 null allele defines an essential role for arginine methylation in genome maintenance and cell proliferation. Mol Cell Biol 29, 2982-2996.

4.9 Supplementary Figures



Supplementary Figure 4.1 Aven harbors dimethylarginines within its RGG/RG motif.
A) Immunoprecipitates of Myc-Aven transfected HEK293T cells resolved by SDS-PAGE. The molecular mass markers are indicated in kDa. Mass spectrometry profile of Aven. LC-MS/MS analysis of the excised Myc-Aven band. The sequence of Aven from residues 63 to 73 is shown. LC-MS/MS analysis revealed the presence of a modified peptide RGGRGGGGAPR containing dimethylated R63 and R66. Analysis of the Aven peptide from residues 33 to 50 is shown as well as the dimethylation of R37. Similar analysis identified R8, R50 and R11 to be dimethylated (not shown).
B) The alignment of Aven N terminus from various eukaryotic species. The mono- (*) and di-methylated (**) arginine residues identified by LC- MS/MS analysis and conserved across various eukaryotic species are indicated with blue boxes.



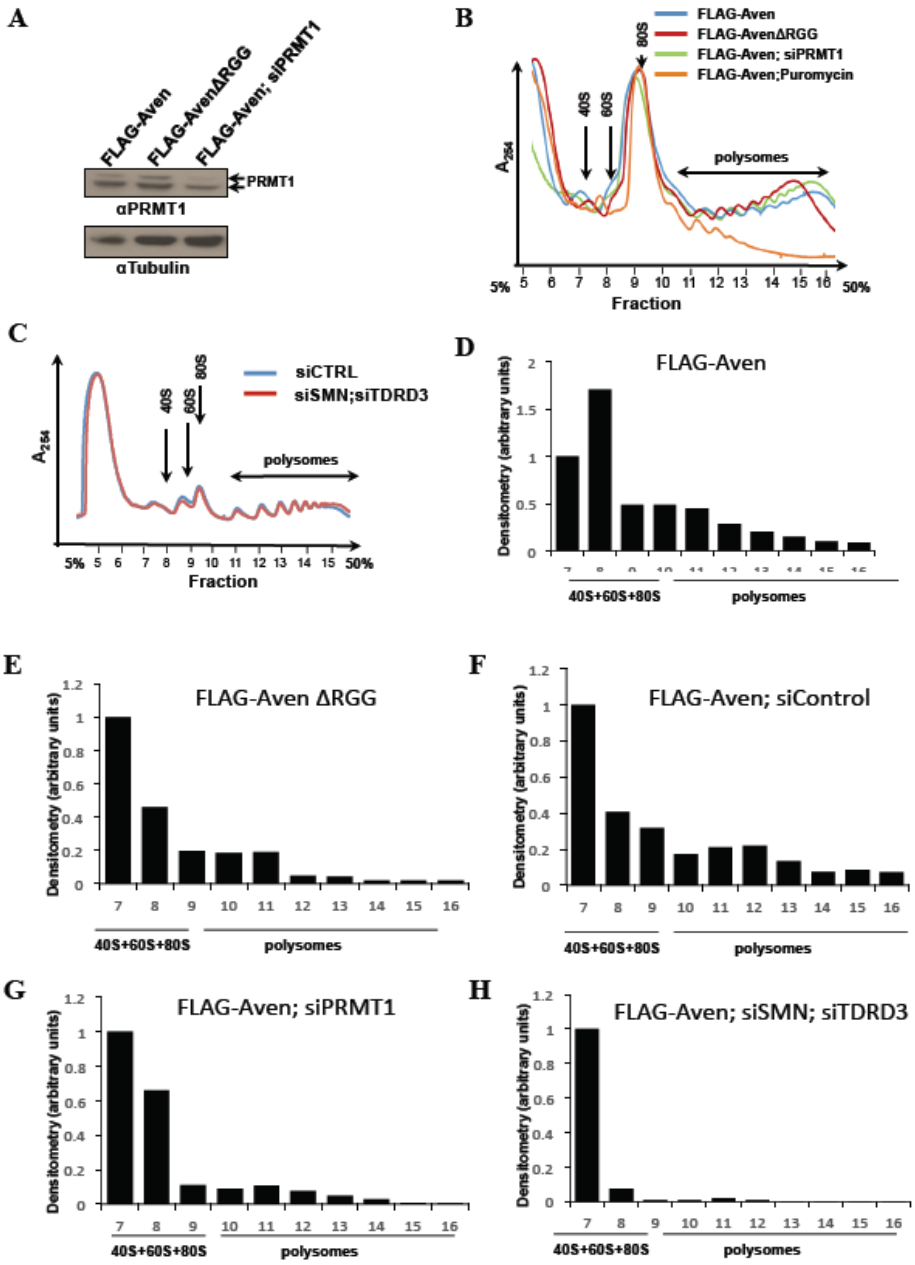
Supplementary Figure 4.2 Aven RGG/RG motif binds RNA and does not regulate ATM activation, nor Aven cellular localization.

A) To investigate the role of RGG/RG motif in ATM activation, U2OS cells were transfected with FLAG-Aven and FLAG-AvenΔRGG. Transfected cells were treated with etoposide (50 ng/ml) for 30 min. Lysates collected at various time points post-treatment were separated by SDS-PAGE and immunoblotted with anti-pATM S1981, anti-FLAG, anti-pCHK2T68, anti- CHK2 and anti-tubulin antibodies. n = 3

B) U2OS cells were co-transfected with Myc-Aven and either FLAG-Aven or FLAG-AvenΔRGG. Immunoprecipitation was performed with anti-FLAG agarose beads and the membranes were immunoblotted with anti-FLAG and anti-Myc antibodies. Ten percent of the lysates were shown in the bottom panel to confirm the expression of the transfected constructs. n = 2.

C) U2OS cells were transfected with FLAG-Aven and FLAG-AvenΔRGG. The cells were fixed and labeled for immunofluorescence with anti-FLAG antibodies. n = 3

D) Biotinylated methylated and unmethylated Aven DiRGG peptides were pre-bound on Streptavidin plates and were incubated with fluorescein-labeled *sc1* G4 RNA. The bound RNA was quantified by measuring absorbance at 515nm.



Supplementary Figure 4.3 Polysomal profiles of siRNA treated cells and quantification of FLAG-Aven and FLAG-AVENΔRGG in polysomal fractions.

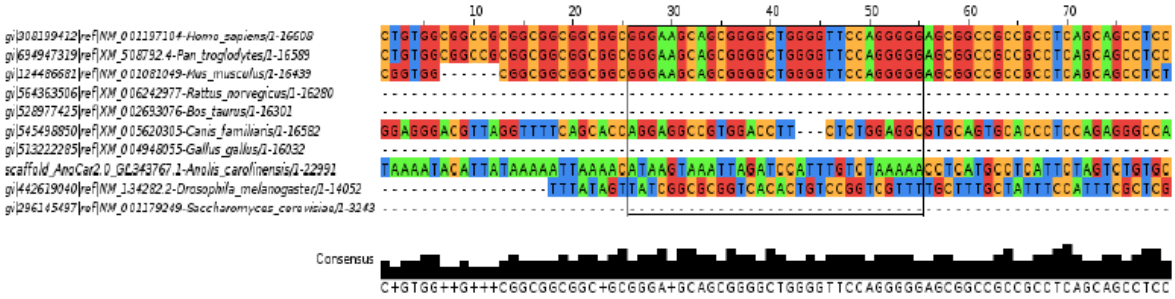
A) Immunoblots confirming the knockdown of PRMT1 and □-Tubulin was used as a loading control. The knockdown was ~2.7-fold, as assessed by densitometric scanning.

B) Overlap of polysome profiles of cells overexpressing FLAG-Aven, FLAG-AvenΔRGG, FLAG-AVEN; siPRMT1, FLAG-AVEN; puromycin treated. Cytoplasmic extracts from the indicated cells were sedimented by centrifugation on a 5 to 50% sucrose gradient, shown as fraction numbers 5 to 16. Polysome profiles were obtained by continuous monitoring of UV absorbance at 254nm. 40S, 60S and 80S indicate the positions of the respective ribosomal subunits and the monomer on the gradient.

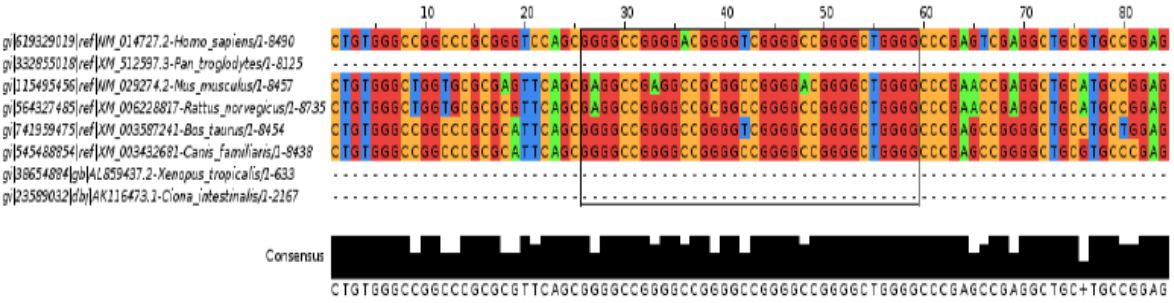
C) Overlap of polysome profiles of cells overexpressing FLAG-Aven; siGFP (siCTRL) or FLAG-AVEN; siSMNsiTDRD3.

D-H) Quantification of FLAG-Aven in each fraction from Figure 4 using densitometry.

MLL1



MLL4



Supplementary Figure 4.4 Sequence conservation of the MLL1 and MLL4 PG4 sequences.

Chapter 5

Polysome-associated Aven/BRISC complex protects against oxidative stress-induced cell death by selectively modulating K63-linked ubiquitination and mRNA translation

5.1 Preface

In the previous chapter, we show that Aven plays a role in MLL1 and MLL4 mRNA translation in human cancer cells. Herein, we would like to investigate the function of Aven under normal and stress conditions. In this chapter, we show that Aven protects against cell death in response to oxidative stress. We report that by binding to the BRISC complex in the polysomes, Aven regulates the translation of a subset of stress-responsive genes.

5.2 Abstract

Lysine 63-linked ubiquitination is a post-translational modification which plays a pivotal role in several cellular processes including endocytosis and DNA repair. In contrast, less is known about the role of lysine 63-linked ubiquitination in the control of mRNA translation. The BRCC36-isopeptidase complex (BRISC) induces lysine 63–linked ubiquitin hydrolysis. Herein, we show that Aven interacts with the components of the BRISC complex, the deubiquitinating enzyme BRCC36 and its adaptor Abro1, whereby a fraction of Aven/BRISC complexes is associated with polysomes. Oxidative stress induces dissociation of Abro1 from the polysome-associated Aven/BRISC complex which is paralleled by modulation of translation of mRNAs encoding stress responsive factors required for cell survival and elevated polysomal K63-linked ubiquitination. These findings identify the Aven/BRISC complex as survival-promoting stress-induced regulator of mRNA translation.

5.3 Introduction

Oxidative stress is a prevalent challenge to cellular homeostasis and can be triggered by diverse biological and environmental radical generating conditions (Apel and Hirt 2004; Klaunig and Kamendulis 2004; Herrero, Ros et al. 2008; Grant 2011). Oxidative stress causes cellular damage affecting cell viability and has been shown to contribute to various human diseases (Finkel and Holbrook 2000; Droge 2002; Klaunig and Kamendulis 2004). To protect from oxidative stress, eukaryotic cells harbor antioxidants, as well as defensive systems to repair oxidative-induced damage (Grant 2011). Oxidative-stress counteracting mechanisms include up-regulation of genes encoding antioxidant specific factors, and increase in protein degradation as well as inhibition of global mRNA translation (Silva, Finley et al. 2015).

K48-linked protein ubiquitination is an essential response to oxidative stress triggering degradation of oxidized proteins by proteasome pathway (Shringarpure, Grune et al. 2003; Medicherla and Goldberg 2008). K63-linked ubiquitination, however is a signaling modification that typically does not result in protein degradation, but rather modifies function of target proteins that play a role in endocytosis, the DNA damage response, and T-cell receptor signaling (Deng, Wang et al. 2000; Zhou, Wertz et al. 2004). It has been shown that the transient exposure to oxidative stress results in K63-linked polyubiquitination accumulation (Silva, Finley et al. 2015; Shringarpure, Grune et al. 2003). Although, this has been characterized in yeast, little is known of the enzymes that lead to this change in mammalian cells.

BRCC36 is a member of a small family of deubiquitinating enzymes (DUBs) called JAMM/MPN$^+$ proteins and it specifically cleaves K63-linked polyubiquitin chains *in vitro* (Cooper, Boeke et al. 2010). BRCC36 is one of the components of BRCC36 isopeptidase containing complex (BRISC), which also comprises RAP80, ABRAXAS (CCDC98), BRE (BRCC45), MERIT40, and BRCA1 (Feng, Wang et al. 2010). This complex localizes in the nucleus and plays a role in DNA damage response (Sobhian, Shao et al. 2007). BRCC36 also associates with Abro1 (KIAA0157), which shares high sequence homology with ABRAXAS (Zheng, Gupta et al. 2013). Compared to ABRAXAS which is predominantly nuclear, Abro1 is mainly cytoplasmic and functions as a scaffold protein (Feng, Wang et al. 2010). Recently,

it has been shown that the DUB activity of BRCC36 is dependent on the presence of Zn^{2+} and the protein-protein interaction with either Abro1 or ABRAXAS (Zeqiraj, Tian et al. 2015). As compared to its roles in the nucleus, the function of the cytosolic BRISC complex remains poorly understood. It has been reported that serine hydroxymethyltransferase (SHMT) directs BRISC activity at K63 ubiquitin chains conjugated to the type I interferon receptor 1 (IFNAR1), which mediates type I interferon signaling (Zheng, Gupta et al. 2013).

Aven is an RGG/RG motif-containing protein predominantly localized in the cytoplasm (Chau, Cheng et al. 2000; Thandapani, O'Connor et al. 2013). Aven has been shown to function as an inhibitor of apoptosis by interacting and stabilizing the anti-apoptotic Bcl-2 family member Bcl-xl and caspase suppressor Apaf-1 (Chau, Cheng et al. 2000). Aven has been shown to play a role in ATM activation, response to DNA damage, and to act as a critical regulator of G_2/M checkpoint (Guo, Yamada et al. 2008; Kutuk, Temel et al. 2010; Baranski, Booij et al. 2015). Aven is also an RNA-binding protein with preference for G-quadruplex (G4) structures which modulates translation of certain mRNAs encoding the mixed lineage leukemia (MLL) proteins, thereby promoting the survival of leukemic cells (Thandapani, Song et al. 2015). Forced expression of Aven in T cells in a transgenic mouse model accelerates lymphomagenesis (Eissmann, Melzer et al. 2013). Overexpression of Aven is also thought to contribute to genesis and progression of acute myeloid and lymphoblastic leukemia, whereby high Aven expression appears to correlate with poor prognosis in clinic (Paydas, Tanriverdi et al. 2003; Choi, Hwang et al. 2006; Eissmann, Melzer et al. 2013).

Herein, we define an Aven/BRISC complex that functions to promote survival by modulating K63-linked ubiquitination and the polysomal association of a subset of stress responsive mRNAs required for cell survival.

5.4 Materials and Methods

5.4.1 Cells, Reagents and Antibodies

HEK293T cells were from the American Type Culture Collection (Manassas, VA). Hydrogen peroxide (H_2O_2) solution 30% (w/w), Protein A/G-Sepharose, anti-FLAG (M2) antibody-coupled agarose beads, mouse anti-FLAG (M2), anti-Myc and anti-β-actin were purchased from Sigma-Aldrich (St. Louis, MO). Anti-GFP was purchased from Roche Life Sciences (#11814460001). Rabbit anti-Aven (ProSci 2413, ProScience) was used for immunoblotting. Mouse anti-rpS6 was from Santa Cruz Biotechnology (Santa Cruz, CA). Rabbit anti-BRCC36 (ab62075), anti-Abro1(ab83860) and anti-uibiquitination (linkage specific to K63 (ab179434) were from Abcam (Cambridge,MA). Anti-BRCA1 (07-434) was purchased from EMD Millipore (Billerica, MA). Rabbit monoclonal against phosphorylated eIF2α at S51 and mouse monoclonal against eIF2α antibodies are from Cell Signaling Technology (Beverly, MA).

5.4.2 DNA Constructs

pMyc-Aven and pFLAG-Aven were previously described (Thandapani et al. 2015). pOZ-HA-tagged Abro1 and pOZ-HA-tagged BRCC36 were a kind gift from Roger Greenberg (Philadelphia, PA). The GFP-tagged Abro1 wild type, N-terminus (1-266) and C-terminus (266-415) plasmid constructs were generated by inserting PCR products amplified using the primer pairs of F1/R2, F1/R1 and F2/R2 respectively, into the pEGFP-C1 vector at *Xho*I and *Bgl*II sites (F1, 5’-GGG TCT CGA GCG GCG TCC ATT TCG GGC TAC-3’, containing a *Xho*I site; R1, 5’-GGG AGA TCT TTA CAT CTG TCT GCT TAA CAC TGC-3’, containing a *Bgl*II site; F2, 5’-GGG TCT CGA GAA AAG GAA CAA GAA AGA AGA- 3’, containing a *Xho*I site; and R2, 5’-GGG AGA TCT TTA AAT CTG GGA GGT CTG AGT GTT -3’, containing a *Bgl*II) pHA-Ub and pHA-UbK63 were subcloned in pcDNA3.1.

5.4.3 Polysome Profiling

Polysome profiling has been performed as described (Gandin, Sikström et al. 2014). Briefly, HEK293T cells in 150 mm plates were transfected with the indicated small interfering RNAs using Lipofectamine RNAiMAX. After 48 h post transfection, the approximately 70% confluent cells were treated with 1mM H_2O_2 for 15 min and subsequently with 100 µg/ml cycloheximide for 5 min to 'freeze' mRNA translation. The cells were washed twice with ice cold-PBS and lysed in hypotonic lysis buffer containing 50 mM Tris-HCl (pH 7.5), 2.5 mM $MgCl_2$, 1.5 mM KCl, 100 µg/ml cycloheximide, 2 mM DTT, 0.5% Triton X-100, and 0.5% sodium deoxycholate. The lysates were spun at 13,000 rpm for 10 min at 4ºC and layered onto a 5% to 50% sucrose gradient as previously described (Gandin, Gutierrez et al. 2013). The gradients were formed using a SW40 rotor (Beckman) at 36,000 rpm for 2 h at 4°C. One ml fractions were collected by upward displacement with 60% sucrose and absorbance was continuously recorded at 254 nm using ISCO fractionator (Teledyne, ISCO). Collected fractions were precipitated with 10% TCA, separated by SDS-PAGE and proteins visualized by immunoblotting. For quantitive RNA analysis, 800 µl TRIzol® reagent (Invitrogen) was added to the 1 ml fractions and RNA was isolated using standard procedures. Isolated RNA was quantified using RT-qPCR. The cDNA samples were serially diluted and the efficiency and *Cq* values were used to generate a standard curve (Piques, Schulze et al. 2009). One standard curve was generated for each primer pair. All standard curves had R^2 value higher than 0.99, with a slope between -3.58 and -3.10. Each data point for each fraction was plotted against the standard curve to calculate the percentage of input.

5.4.4 RT-qPCR primers

Gene	Primer	Sequence (5'->3')
β-actin	Forward	ACCACACCTTCTACAATGAGC
	Reverse	GATAGCACAGCCTGGATAGC
ERCC1	Forward	TCCTGACCACATTTGGATCTCTG
	Reverse	TTCAAGAAGGGCTCGTGCAG

ERCC5	Forward	TAGAAGTTGTCGGGGTCCGC
	Reverse	CCAAATGCTAATATCAACAGCCAGG
Bcl-2	Forward	GGATAACGGAGGCTGGGATG
	Reverse	TGACTTCACTTGTGGCCCAG
ATF4	Forward	CTTGATGTCCCCCTTCGACC
	Reverse	GAAGGCATCCTCCTTGCTGT
GAPDH	Forward	ACCACAGTCCATGCCATCAC
	Reverse	TCCACCACCCTGTTGCTGTA

5.4.5 siRNA Transfections

Small interfering RNAs (siRNAs; Dharmacon Inc.) were transfected using Lipofectamine RNAi MAX (Invitrogen) as per the manufacturer's protocol. The final concentration of the siRNA was 50 nM and the cells were lysed 48h post-transfection. The siRNA target sequences for Aven were siAven 5′-GAG GAG AAA GAA UGG GAU AUU-3′. For Abro1 and BRCC36 siRNAs, SMARTpools were purchased from Dharmacon Inc.

5.4.6 Immunoprecipitations and Immunoblotting

HEK293T cells were transfected using Lipofectamine 2000 (Invitrogen) as per manufacturer's instructions. After 48 h, the cells were lysed with cell lysis buffer (20 mM Tris pH 7.4, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% Triton X-100). For immunoprecipitations, cell lysates were incubated with the primary antibody for 2 h at 4ºC. Then 25 µl of 50% protein A-Sepharose slurry was added and incubated at 4ºC for 45 min with constant end-over-end mixing. The beads were then washed three times with cell lysis buffer and once with 1x PBS. The samples were then boiled with 25 µl of 2x Laemmli buffer, resolved by SDS-PAGE, transferred to nitrocellulose membranes and the proteins visualized by immunoblotting.

5.4.7 Flow cytometry analysis with Propidium Iodide staining

Cells ($2x10^5$ cells/ well) were treated with 0.5 mM H_2O_2 or vehicle control (1x PBS) for

various time points up to 24 h prior to analysis. Cells were harvested by trypsinization and collected along with initial culture medium to ensure inclusion of detached cells. Cells were pelleted by centrifugation (300 g, 5 min) and washed by 1xPBS and subsequently fixed in 70% ethanol for at least 2 h. The pellets were washed by 1xPBS and suspended in 20 µg/ml propidium iodide (Sigma-Aldrich) containing 20 µg/ml RNase A for 15 min at room temperature in the dark. Subsequently, cells were analysed by flow cytometry, using a FACSCalibur (BD Biosciences) operated by CellQuest software and at least 10,000 events were collected per sample. Data was analyzed using FlowJo software (Tree Star Inc., Ashland, OR, USA).

5.5 Results

5.5.1 Association of Aven with BRISC regulates oxidative stress induced K63 ubiquitination

To identify interactors of Aven, we previously performed stable isotope labeling by amino acids (SILAC) in U2OS cells transfected with empty vector (pcDNA3.1) or vector encoding FLAG-Aven (Thandapani, Song et al. 2015). Among the 146 proteins interacting with FLAG-Aven, we identified BRCC36 and Abro1, members of the deubiquitinating BRISC (BRCC36 isopeptidase complex) complex (Cooper, Boeke et al. 2010). We further confirmed that endogenous BRCC36 and Abro1 co-immunoprecipitated with FLAG-tagged and endogenous Aven (Figure 5.1A, 5.1B). The lack of BRCA1 in Aven immunoprecipitates revealed that the nuclear BRCC36/ABRAXAS complex (Feng, Wang et al. 2010) does not associate with Aven (Figure 5.1A). These findings demonstrate that Aven associates with cytoplasmic BRISC.

We next defined the region of Abro1 protein required for its association with Aven. Abro1 and ABRAXAS are 39% identical at the N-terminal region which contains a MPN-like domain and a coil-coiled domain (Wang, Matsuoka et al. 2007; Wang, Hurov et al. 2009). It has been shown that the MPN-like domain is responsible for interaction with BRE (Brain and Reproductive organ-Expressed), while the coiled-coil domain binds to BRCC36 (Feng,

Huang et al. 2009; Shao, Patterson-Fortin et al. 2009). BRE bridges the BRCA1 and BRISC interaction in the nucleus (Cooper, Boeke et al. 2010). ABRAXAS also has a pSXXF domain at its C terminus, which mediates the interaction with BRCA1, while Abro1 lacks this motif (Kim, Huang et al. 2007). HEK293T cells were co-transfected with Myc-Aven and either GFP-Abro1, GFP-Abro1:1-266, or GFP-Abro1:266-415. GFP-Abro1 and GFP-Abro1:1-266, but not GFP-Abro1:266-415 co-immunoprecipitated with Myc-Aven (Figure 5.1C), suggesting that the N terminus of Abro1 is required for its association with Aven. We next investigated whether Aven influences Abro1/BRCC36 interaction by transfecting $Aven^{+/+}$ and $Aven^{-/-}$ cells with GFP-Abro1:1-266. Whereas BRCC36 co-immunoprecipitated with GFP-Abro1:1-266 in $Aven^{+/+}$ cells, this association was severely impaired in $Aven^{-/-}$ cells (Figure 5.1D). Since, it has been reported that Abro1 is essential for DUB activity of BRCC36 (Zeqiraj, Tian et al. 2015), these observations suggest that Aven likely regulates BRISC activity.

In yeast, K63-linked ubiquitination is thought to regulate mRNA translation during oxidative stress whereby ribosomal proteins represent a major target for K63 ubiquitination (Silva, Finley et al. 2015). To assess the impact of oxidative stress on K63 ubiquitination in the mammalian cells, we transfected HEK293T cells with empty vector or vectors encoding HA-tagged ubiquitin (HA-Ub) or K63-specific ubiquitin with other lysines mutated to arginines (HA-K63Ub), treated them with hydrogen peroxide (H_2O_2) for 15 min and monitored ubiquitination using an anti-K63 antibody (Figure 5.1E). The transfection of HA-K63Ub alone caused a dramatic increase in anti-K63 levels, as expected and H_2O_2 treatment led to a further increase in K63 polyubiquitination in cells expressing HA-Ub (Figure 5.1E). Whereas there was no significant change in the cells expressing HA-K63Ub after H_2O_2 treatment, considering that the K63 ubiquitination level was saturated in these cells (Figure 5.1E). These findings show that similarly to yeast, H_2O_2 increases K63 ubiquitination in mammalian cells.

Aven associates with polysomes to regulate mRNA translation of specific mRNAs harboring G4 or G4-like structures (Thandapani, Song et al. 2015). We examined whether Abro1 and BRCC36 were associated with polysomes. Indeed Abro1 and BRCC36

co-sedimented with polysomes along with FLAG-Aven and rpS6 (Supplementary Figure 5.1A). Treatment with puromycin, the antibiotic that disrupts polyribosomes, shifted distribution of FLAG-Aven, Abro1, BRCC36 and rpS6 towards the top of the gradient (Supplementary Figure 5.1A). This indicates that Aven, Abro1 and BRCC36 associate with polysomes but not other cellular structures that possess a similar sedimentation coefficient as polysomes.

Then we determined whether polysome associated Aven and/or BRCC36 complexes modulate K63 ubiquitination in polysome fractions. $Aven^{+/+}$ and $Aven^{-/-}$ cells and BRCC36 depleted cells were treated with H_2O_2. Total cellular lysates (TCL) of HEK293T cells were collected and immunoblotted with anti-Aven, and -BRCC36 antibodies to confirm the depletions (Supplementary Figure 5.2A). Polysomal fractionation of these cells was performed by sedimenting cytoplasmic extracts on 5% to 50% sucrose gradients by ultracentrifugation (Gandin, Sikstrom et al. 2014). Levels of K63-ubiquitinated proteins in each fraction were determined by Western blotting (Figure 5.1F). We observed that H_2O_2 drastically repressed mRNA translation as illustrated by a decrease in the number of ribosomes involved in polysomes compared to non-treated cells (Supplementary Figure 5.2B), as previously reported in yeast (Silva, Finley et al. 2015). Considering that 15 min H_2O_2 treatment only partially disrupted polysomes, we chose this time point for the subsequent experiments (Supplementary Figure 5.2B). Neither Aven nor BRCC36 depletion exerted dramatic effects on the number of ribosomes engaged in polysomes (Supplementary Figure 5.2C), thereby demonstrating that Aven or BRCC36 do not dramatically influence global mRNA translation. Aven- and BRCC36- depleted cells exhibited a modest increase in K63-linked ubiquitination under baseline conditions in total cell lysates (TCLs), and the polyribosome fractions, as compared to a control (Figure 5.1F, upper panels). Strikingly, H_2O_2 increased K63-linked ubiquitination in polysome fractions and this effect was further potentiated by either depletion of Aven or BRCC36 (Figure 5.1F, lower panels). These results suggest that Aven and BRCC36 regulate levels of K63-ubiquitinated proteins in the polysomes in response to oxidative stress.

5.5.2 H_2O_2 induces partial dissociation of Aven/BRISC complex in the polysomes

We next set out to establish the role of Aven and BRISC in regulating K63 ubiquitin levels in polysome fractions in response to oxidative stress. Abro1, BRCC36 and Aven distribution across polysome gradients of control, Aven- or BRCC36-depleted cells were monitored in the presence or absence of H_2O_2. H_2O_2 induced dramatic decrease in co-sedimentation of Abro1, but not Aven or BRCC36 with the polysomes (compare Figure 5.2A with 5.2B). Depletion of Aven did not affect the polysomal association of Abro1 and BRCC36 (Figure 5.2A and 5.2B, middle panels), while silencing of BRCC36 did not significantly influence distribution of Aven and Abro1 across the polysome gradient under basal conditions or oxidative stress (Figure 5.2A and 5.2B, lower panels). We next tested whether H_2O_2 treatment affects Aven association with Abro1 and BRCC36. As shown in Figure 2C, H_2O_2 treatment significantly reduced FLAG-Aven association with Abro1 and thus reduced the association with BRCC36. Taken together, these results suggest that H_2O_2 treatment led to Abro1 dissociation from polysomes (Figure 5.2B). The Abro1 dissociation did not affect the occupation of Aven or BRCC36 on the polysomes, but led to significant decrease of the association between BRCC36 and Abro1, thereby decreasing DUB activity of BRISC in polysomes (Figure 5.2D).

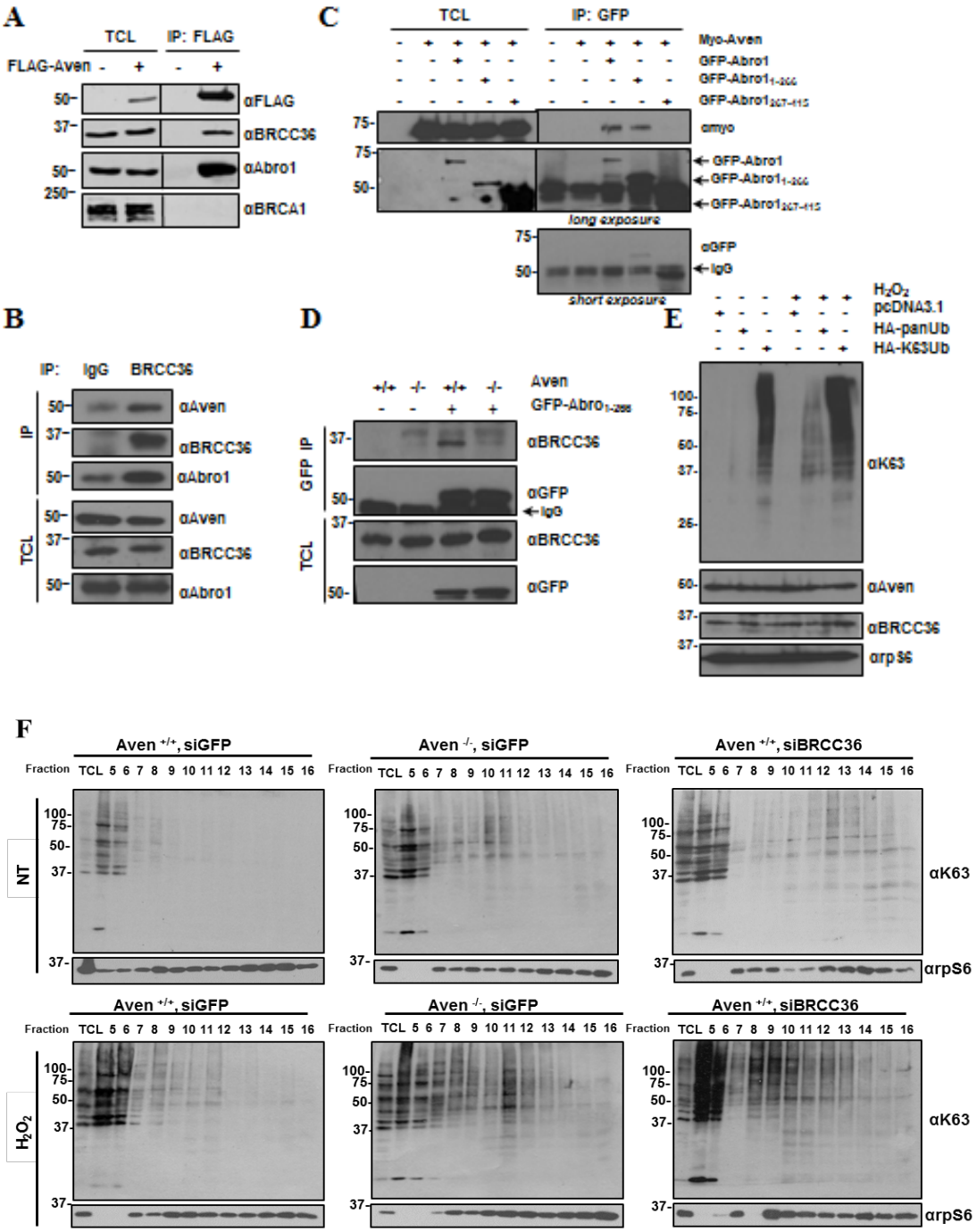


Figure 5.1 Aven associates with BRCC36 and Abro1 to regulate K63 ubiquitination in response to H_2O_2

A) HEK293T cells transfected with empty vector pcDNA3.1 or an expression vector encoding FLAG-Aven were lysed and immunoprecipitated (IP) with anti-FLAG antibody. The total cell lysates (TCL) and the bound proteins were resolved by SDS-PAGE and immunoblotted with anti-Abro1, -BRCC36, -BRCA1 or -FLAG antibodies.

B) HEK293T cells were lysed and immunoprecipitated with anti-IgG or anti-BRCC36 antibody. The bound proteins were resolved by SDS-PAGE and immunoblotted with anti-Abro1, -Aven, or -BRCC36 antibodies.

C) HEK293T cells transfected with pcDNA3.1, pMyc-Aven, pGFP-Abro1,

pGFP-Abro1:1-266, or pGFP-Abro1:267-415 were lysed after 48 h transfection and immunoprecipitated with anti-GFP antibodies. The bound proteins were detected by immunoblotting.
D) Aven $^{+/+}$ or Aven$^{-/-}$ cells were transfected with empty vector pcDNA3.1 and GFP-Abro1:1-266. After 48 h post transfection, the cells were lysed and immunoprecipitated with anti-GFP antibodies. The total cell lysates and the bound proteins were resolved by SDS-PAGE and immunoblotted with anti-BRCC36 and -GFP antibodies.
E) HEK293T cells were transfected with empty vector (pcDNA3.1), HA-Ub or HA-UbK63. After 48 h transfection, the cells were treated with 1 mM H_2O_2 for 15 min or left untreated and subsequently lysed and immunoblotted with anti-K63, -Aven and -BRCC36, and -rpS6 antibodies.
F) Aven $^{+/+}$ or Aven$^{-/-}$ cells were either transfected with siGFP or siBRCC36. After 48 h, the cells were treated with 1 mM H_2O_2 for 15 min or left untreated. Cytoplasmic extracts were sedimented on a 5 to 50% sucrose gradient. The distribution of K63 polyubiquitination across the gradient was monitored by immunoblotting with anti-K63 antibody. Each polysomal profile experiment was performed independently three times.

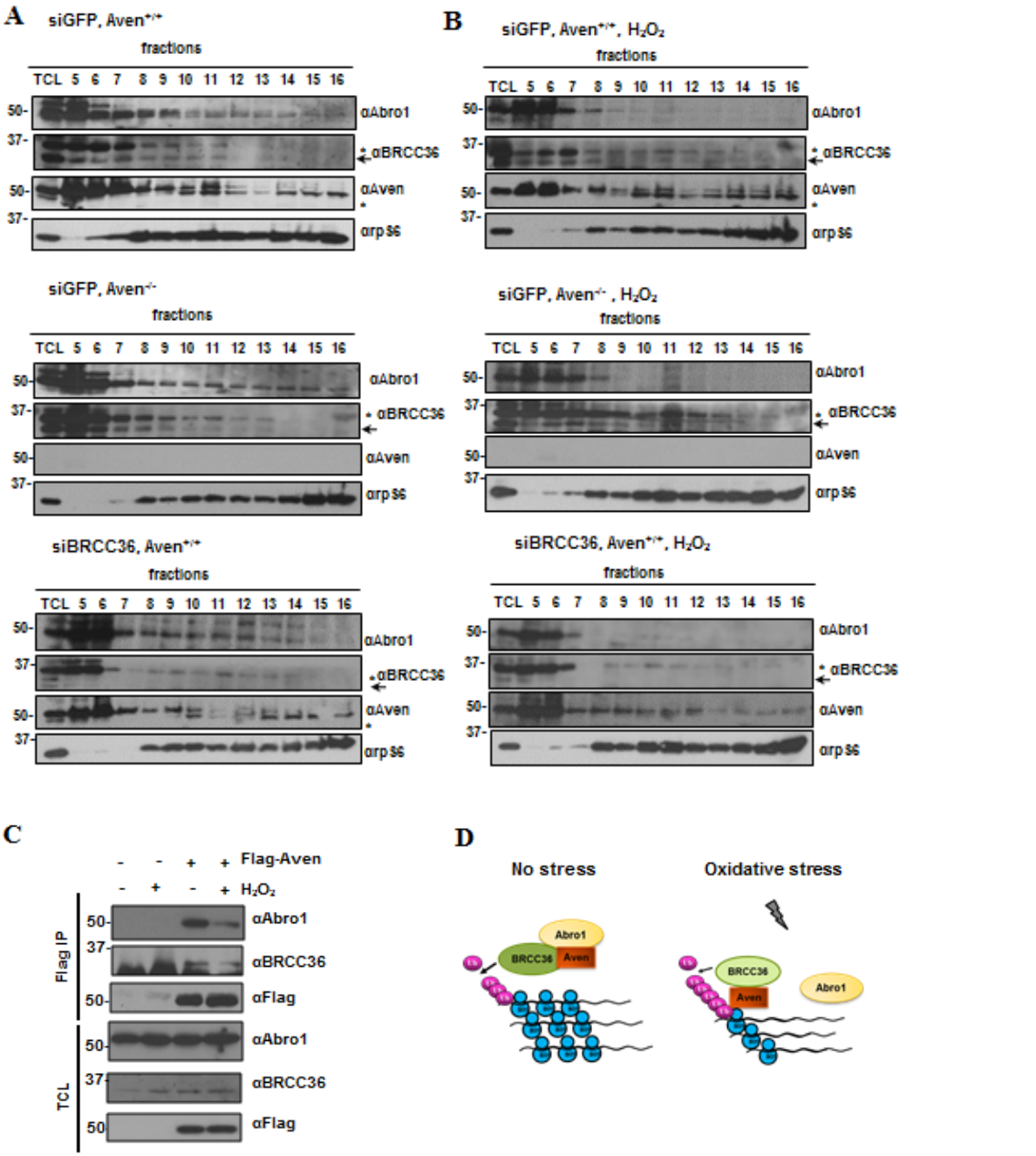


Figure 5.2 Aven/BRISC localize to polysomes and Abro1 is regulated by oxidative stress

A-B) Aven $^{+/+}$ or Aven$^{-/-}$ cells were transfected with siGFP or siBRCC36, treated with 1 mM H2O2 for 15 min or left untreated 48 h post transfection, and the cytoplasmic extracts were collected and submitted to sedimentation by centrifugation on a 5 to 50% sucrose gradient. The distribution of Aven, Abro1 and BRCC36 across the gradient was monitored by immunoblotting with indicated antibodies. rpS6 was used as a loading control. The exposure time was determined using a standard curve with increasing amounts of lysates immunoblotted with anti-Abro1 and anti-BRCC36 antibodies for various times. Each polysomal profile experiment was performed independently three times. The unspecific band of BRCC36 is denoted as *.

C) HEK293T cells were transfected with empty vector pcDNA3.1, FLAG-Aven and treated with 1 mM H_2O_2 for 15 min or left untreated 48 h post transfection. The cell extracts were lysed and immunoprecipitated (IP) with anti-FLAG antibodies. TCLs and the bound proteins were resolved by SDS-PAGE and immunoblotted with anti-Abro1, -BRCC36 and -FLAG antibodies. The experiment was performed independently three times.
D) Model of the interaction of BRISC/Aven complex during the translation cycle in normal condition and under oxidative stress. K63, K63 polyubiquitination. 80S, 80S ribosomal subunits.

5.5.3 Aven and BRCC36 regulate mRNA polysomal association of oxidative stress responsive genes

Exposure to ultraviolet or oxidative stress inhibits global translation and selective mRNAs required for cell survival are recruited to polysomes, encoding ATF4, DNA repair proteins and others (Harding, Zhang et al. 2003; Powley, Kondrashov et al. 2009). Depletion of Aven and BRCC36 did not disrupt polysomes (Supplementary Figure 5.2C), suggesting that global translation is not significantly inhibited. Therefore, we investigated whether Aven and BRCC36 regulate translation of a specific set of mRNAs upon oxidative stress. Using the Genbank database, we tested over 30 genes which are normally induced under oxidative stress. $Aven^{+/+}$ and $Aven^{-/-}$ cells and siBRCC36-depleted cells were treated with H_2O_2 and total RNA analyzed by RT-qPCR. We noticed that oxidative stress induced an increase of total mRNAs including TP53, Bax, DDB1, Mcl-2, Bad, FOXA1, apopt1, IRF3, TRAF3 and MAPK13 (data not shown). Another subset of stress response genes remained unchanged at the mRNA level including ERCC1, ERCC5, Bcl-2 and ATF4 (Figure 5.3A) and PIAS1, MNAT1, TMBIM1, FMO4, SOD1, PML, PINK1, HSPB1, TRIM69, NFE2L2, Akt1, BIRC5, TXNRD and BLVRA (data not shown). Since these genes were not regulated at the steady-state mRNA level, we next examined whether they are regulated at the level of translation. Interestingly, H_2O_2 increased the translation of ERCC1, ERCC5, Bcl-2 and ATF4 mRNAs as illustrated by their shift to the heavy polysomal fractions, as compared to a control, and this effect was reversed by depletion of Aven or BRCC36 (Figure 5.3B). In contrast, neither H_2O_2 nor silencing of Aven or BRCC36 affected the distribution of β-actin mRNA which remained associated with heavy polysomes under normal and stress conditions (Figure 5.3B). These findings suggest that Aven and BRCC36 selectively stimulate translation of stress-induced mRNAs in response to oxidative stress. These mRNAs encode proteins with roles in adaptive responses to oxidative stress, including ERCC1/ERCC5 in nucleotide excision repair (NER) pathway (Christmann and Kaina 2013), Bcl-2 in anti-apoptotic pathway (Susnow, Zeng et al. 2009), and ATF4 as an anti-oxidative response (Harding, Zhang et al. 2003). Since ATF4 is known as a transcription factor induced by eIF2α phosphorylation and reduced ternary complex recycling (Harding, Zhang et al. 2003; Baird and Wek 2012),

we tested whether depletion of Aven or BRCC36 leads to decrease of eIF2α phosphorylation (Supplementary Figure 5.3). We observed that H_2O_2 induced eIF2α phosphorylation in both control cells and the Aven- or BRCC36- depleted cells (Supplementary Figure 5.3). Therefore, we conclude that Aven and BRCC36 regulate stress-induced mRNAs translation in an eIF2α phosphorylation independent manner.

5.5.4 Aven and BRCC36 promotes survival under oxidative stress

To determine the functional consequences of the Aven/BRISC association on the polysomes, cells depleted of Aven, or both Abro1 and BRCC36 (BRISC) were challenged with H_2O_2 to induce apoptosis. We observed that compared to non-treated cells, H_2O_2 induced a significant increase in apoptosis in cells in siGFP control cells which was further exacerbated by depletion of Aven (siAven) or the BRISC complex (siBRISC, Figure 5.4A). These findings are consistent with Aven and BRISC being survival factors (Chau, Cheng et al. 2000; Chai, Wang et al. 2014). We next examined whether ectopic expression of the BRISC complex could attenuate H_2O_2-induced apoptosis. Indeed, we observed that overexpression of the HA-tagged Abro1 and BRCC36 together, but not vector control (pcDNA3.1) attenuated the H_2O_2 induced apoptosis (Figure 5.4B). Taken together, these results indicate that Aven/BRISC complex regulates the translation of stress-induced cellular protective mRNAs, thereby modulating the cellular sensitivity to oxidative stress.

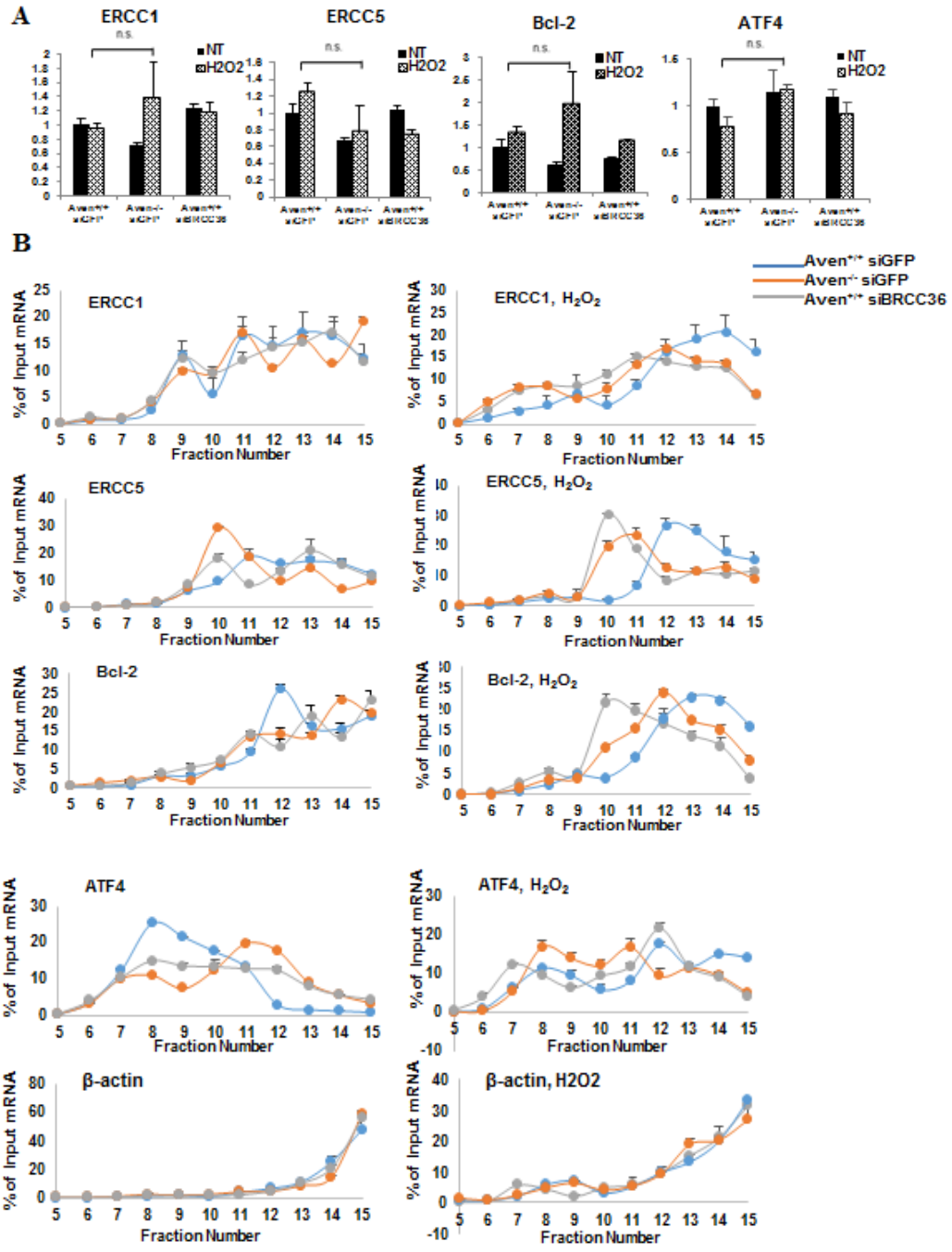


Figure 5.3 Aven regulates the polysomal association of mRNAs encoding stress response factors

A) RT-qPCR of the indicated mRNAs was performed from cytosolic RNA isolated Aven $^{+/+}$ or Aven$^{-/-}$ cells transfected with siGFP or siBRCC36, treated with 1 mM H_2O_2 for 15 min or left untreated 48 h post transfection, and expressed as a relative fold change normalized to β-actin levels. Error bars ± standard deviation is shown. The significance was measured by the Student t-test and defined as n.s. (not significant) n = 3.

B) Aven $^{+/+}$ or Aven$^{-/-}$ cells were transfected with siGFP or siBRCC36, treated with 1 mM H_2O_2 for 15 min or left untreated 48 h post transfection, and the cytoplasmic extracts were collected and submitted to sedimentation by centrifugation on a 5 to 50% sucrose gradient. The cytosolic RNA and the polysomal associated RNA were isolated from the indicated

polysomal fractions, and the presence of mRNAs encoding ERCC1, ERCC5, Bcl-2, ATF4 or β-actin was quantified by qRT-PCR. mRNAs in each fraction is represented as the percentage of input. Error bars represent ± SEM, n = 3.

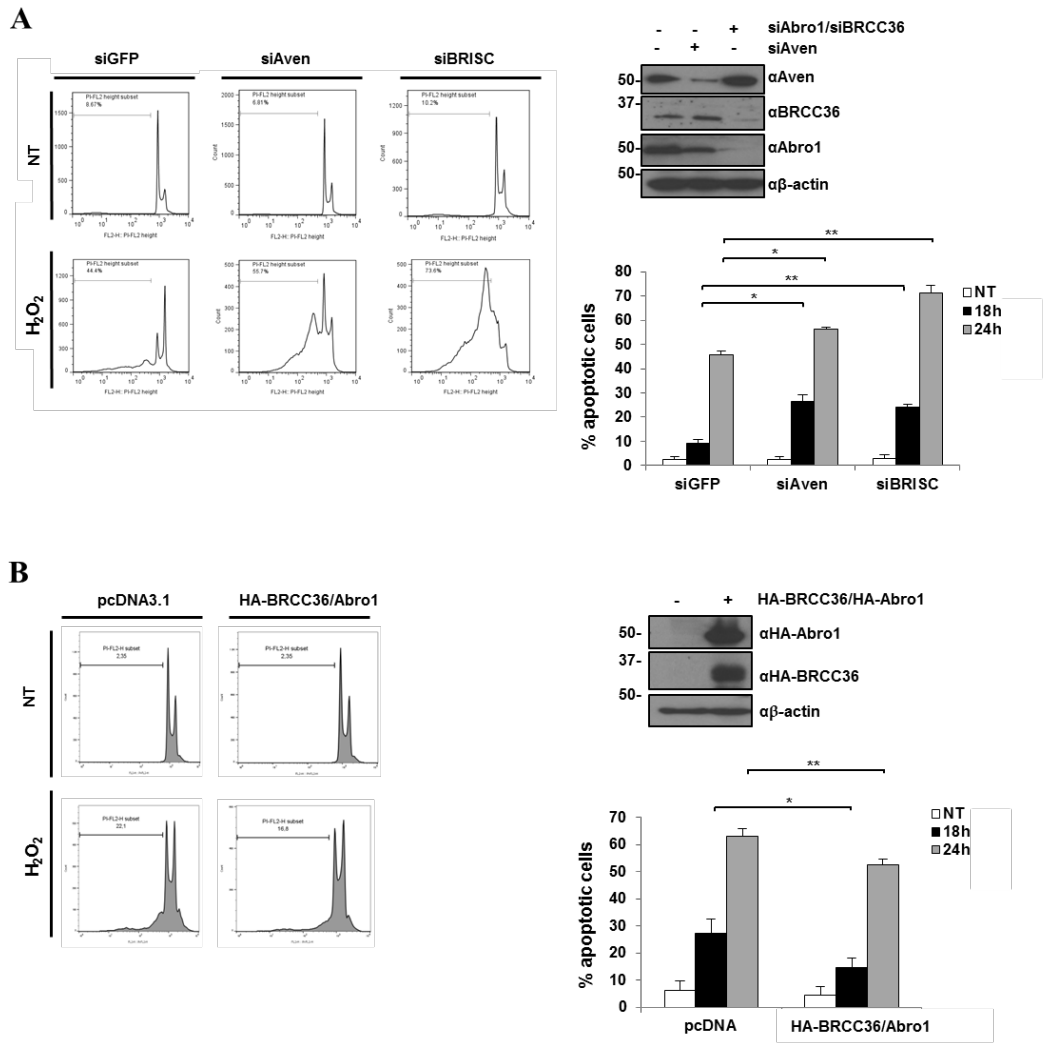


Figure 5.4 Aven/BRISC modulates cell death induced by H_2O_2

A) HEK293T cells were transfected with siGFP, siAven or siBRCC36 and siAbro1. After 48 h transfection, cells were left untreated or treated with 0.5 mM H_2O_2 for 18 h and 24 h, respectively. Ten percent of the cells were lysed and immunoblotted with indicated antibodies. Ninety percent of the cells were collected, fixed in 70% ethanol and submitted to FACS analysis. The percentage of apoptotic cells in the population was monitored by FACS Calibur and shown as sub-G_1 phases. The left panel is the representative of the three individual experiments. The right-bottom panel is the quantification from three individual experiments. Error bars represent ± SEM, * $p < 0.05$, ** $p < 0.01$, n = 3.

B) HEK293T cells were transfected with pcDNA3.1, HA-BRCC36 and HA-Abro1. After 48 h transfection, the cells were treated and analyzed as in (A). The left panel is the representative of the three individual experiments, while the right-bottom panel is the quantification from three individual experiments. Error bars represent ± SEM, * $p < 0.05$, ** $p < 0.01$, n = 3.

5.6 Discussion

In the present study, we characterize the functional interaction between Aven and the components of the BRISC DUB complex, Abro1 and BRCC36. We confirm the association by co-immunoprecipitation experiments and identify Abro1 and BRCC36 to be associated with polyribosomes. Furthermore, we show that oxidative stress by H_2O_2 treatment leads to an accumulation of proteins harboring K63-linked ubiquitination, consistent with previous findings (Silva, Finley et al. 2015). K63 ubiquitinated proteins, are accumulated in response to H_2O_2 which is further potentiated by Aven- or BRCC36-depletion. BRCC36 deubiquitinating activity was modulated by Aven that functions as a regulator of the BRISC complex *in vivo*, while H_2O_2 triggered the dissociation of the Abro1 from the Aven/BRISC complex in polysomes. These findings are consistent with the reported translocation of Abro1 from the cytoplasm to the nucleus in response to stress (Cilenti, Balakrishnan et al. 2011; Ambivero, Cilenti et al. 2012; Zhang, Cao et al. 2014). To our knowledge, this is the first time that H_2O_2 induced K63 ubiquitination is reported in polyribosomes in mammalian cells, although in yeast this is well-established (Silva, Finley et al. 2015). It has been shown that DUBs of USP family are generally inhibited in response to oxidative stress (Shenton, Smirnova et al. 2006; Lee, Baek et al. 2013); whereas for the first time it is shown that BRCC36, the JAMM-MPN^+ member of DUB which is generally considered to be an important regulator in the DNA damage response (Sobhian, Shao et al. 2007; Ng, Wei et al. 2016), regulates K63-linked ubiquitination in polysomes, and this DUB activity is modulated by protein association in response to H_2O_2. It is known that E2 ubiquitin-conjugating enzyme, Ubc13, associates with p53 in the polysomes and mediates the K63-linked ubiquitination of newly synthesized p53, thereby regulating its localization and transcriptional activation (Topisirovic, Gutierrez et al. 2009). Taken together, these findings suggest that Aven/BRISC contributes to the K63-linked homeostasis in polysomes.

Oxidative stress is known to arrest global translation and to favor the translation of specific mRNAs which play a major role in stress response (Holcik and Sonenberg 2005; Shenton, Smirnova et al. 2006). In addition to transcription regulation which has been widely studied, accumulating evidence indicates post-transcriptional control is of greater significance than it

has been assumed previously. Several studies have demonstrated the selective mRNAs are elevated or down-regulated after various kinds of stress (Shenton, Smirnova et al. 2006; Powley, Kondrashov et al. 2009; Leprivier, Rotblat et al. 2015). Consistently, we observed a subset of genes with significant changes at their transcription level (data not shown), while another subset with no dramatic change (Figure 5.3A). We show that although global translation is severely inhibited in response to H_2O_2, Aven/BRISC complex appears to stimulate translation of specific mRNA under these conditions (Figure 5.3B). Mechanistically, it has been shown that cap-dependent translation initiation utilizes eukaryotic initiation factor-4F (eIF4F), which consists of eukaryotic translation initiation factor 4G, 4E and the RNA helicase 4A (eIF4G, eIF4E, eIF4A) (Holcik and Sonenberg 2005; Komar and Hatzoglou 2011). Genome-wide ribosome profiling analysis using eIF4A inhibitors which block the helicase ability to unwind complex RNA structures in 5' UTR, inhibit malignancy-related mRNA translation and led to apoptosis. These data show that 5'UTR complexity, including structured 5'UTR and G quadruplexes, determines translation sensitivity to eIF4A or its subunits (Rubio, Weisburd et al. 2014; Wolfe, Singh et al. 2014; Gandin, Masvidal et al. 2016). Whereas under stress conditions, when eIF4F activity is compromised and protein synthesis is severely inhibited, phosphorylation of the eIF2 α subunit induces selective translation of mRNAs containing upstream open reading frames (uORFs) by reinitiation, as well as the internal translation initiation (IRES) in a cap-independent manner (Holcik and Sonenberg 2005; Komar and Hatzoglou 2011). For both cases, these selective mRNAs harbor typically long, GC-rich, highly structured 5'-UTR (Komar and Hatzoglou 2011). Similarly, we observe ERCC1, ERCC5, Bcl-2, and ATF4, known to have structured 5' UTR or uORFs (Lee, Cevallos et al. 2009; Shahid, Bugaut et al. 2010; Somers, Wilson et al. 2015), to be regulated by Aven and BRCC36 in an eIF2α independent manner. Furthermore, Shahid et al. 2010, have shown that human Bcl-2 forms G4 structures in the 42 nucleotides upstream of its translation start site, which is consistent with previous finding that Aven regulates specific mRNA translation with G4 structures (Shahid, Bugaut et al. 2010; Thandapani, Song et al. 2015). Lastly, it is of interest to investigate how accumulated K63 chain affects the translation of specific subset of

mRNAs. It has been illustrated that abnormal elongated K63 chain in the yeast polysome recruits the proteasome which leads to the ribosome dissociation involved quality control system (Saito, Horikawa et al. 2015). Based on our and other studies, we propose that proper length of K63 polyubiquitin chains is fundamental for efficient translation of mRNAs under stress condition, and the disruption of K63-linked ubiquitination homeostasis leads to aberrant translation and thus pro-survival adaptation events are compromised.

5.7 Acknowledgements

This work was supported by CIHR grant MOP-67070 to S.R.; CIHR/FRSQ training grant in cancer research of the McGill Integrated Cancer Research Training Program (FRN53888 to J.S.).

5.8 References

Ambivero, C. T., L. Cilenti, et al. (2012). "ATF4 interacts with Abro1/KIAA0157 scaffold protein and participates in a cytoprotective pathway." Biochim Biophys Acta 1823(12): 2149-2156.

Apel, K. and H. Hirt (2004). "Reactive oxygen species: metabolism, oxidative stress, and signal transduction." Annu Rev Plant Biol 55: 373-399.

Baird, T. D. and R. C. Wek (2012). "Eukaryotic initiation factor 2 phosphorylation and translational control in metabolism." Adv Nutr 3(3): 307-321.

Baranski, Z., T. H. Booij, et al. (2015). "Aven-mediated checkpoint kinase control regulates proliferation and resistance to chemotherapy in conventional osteosarcoma." J Pathol 236(3): 348-359.

Chai, K. M., C. Y. Wang, et al. (2014). "Downregulation of BRCA1-BRCA2-containing complex subunit 3 sensitizes glioma cells to temozolomide." Oncotarget 5(21): 10901-10915.

Chau, B. N., E. H. Cheng, et al. (2000). "Aven, a novel inhibitor of caspase activation, binds Bcl-xL and Apaf-1." Mol Cell 6(1): 31-40.

Choi, J., Y. K. Hwang, et al. (2006). "Aven overexpression: association with poor prognosis in childhood acute lymphoblastic leukemia." Leuk Res 30(8): 1019-1025.

Christmann, M. and B. Kaina (2013). "Transcriptional regulation of human DNA repair genes following genotoxic stress: trigger mechanisms, inducible responses and genotoxic adaptation." Nucleic Acids Res 41(18): 8403-8420.

Cilenti, L., M. P. Balakrishnan, et al. (2011). "Regulation of Abro1/KIAA0157 during myocardial infarction and cell death reveals a novel cardioprotective mechanism for Lys63-specific deubiquitination." J Mol Cell Cardiol 50(4): 652-661.

Cooper, E. M., J. D. Boeke, et al. (2010). "Specificity of the BRISC deubiquitinating enzyme is not due to selective binding to Lys63-linked polyubiquitin." J Biol Chem 285(14): 10344-10352.

Deng, L., C. Wang, et al. (2000). "Activation of the IkappaB kinase complex by TRAF6 requires a dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain." Cell 103(2): 351-361.

Droge, W. (2002). "Free radicals in the physiological control of cell function." Physiol Rev 82(1): 47-95.

Eissmann, M., I. M. Melzer, et al. (2013). "Overexpression of the anti-apoptotic protein AVEN contributes to increased malignancy in hematopoietic neoplasms." Oncogene 32(20): 2586-2591.

Feng, L., J. Huang, et al. (2009). "MERIT40 facilitates BRCA1 localization and DNA damage repair." Genes Dev 23(6): 719-728.

Feng, L., J. Wang, et al. (2010). "The Lys63-specific deubiquitinating enzyme BRCC36 is regulated by two scaffold proteins localizing in different subcellular compartments." J Biol Chem 285(40): 30982-30988.

Finkel, T. and N. J. Holbrook (2000). "Oxidants, oxidative stress and the biology of ageing." Nature 408(6809): 239-247.

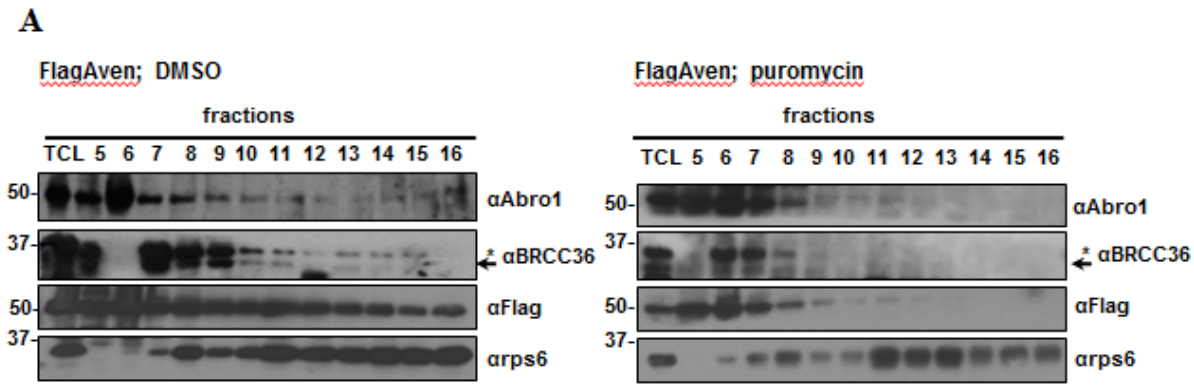
Gandin, V., G. J. Gutierrez, et al. (2013). "Degradation of newly synthesized polypeptides by

ribosome-associated RACK1/c-Jun N-terminal kinase/eukaryotic elongation factor 1A2 complex." Mol Cell Biol 33: 2510-2526.
Gandin, V., L. Masvidal, et al. (2016). "nanoCAGE reveals 5' UTR features that define specific modes of translation of functionally related MTOR-sensitive mRNAs." Genome Res 26(5): 636-648.
Gandin, V., K. Sikström, et al. (2014). "Polysome fractionation and analysis of mammalian translatomes on a genome-wide scale." J Vis Exp May 17 (87): doi: 10.3791/51455.
Gandin, V., K. Sikstrom, et al. (2014). "Polysome fractionation and analysis of mammalian translatomes on a genome-wide scale." J Vis Exp(87).
Grant, C. M. (2011). "Regulation of translation by hydrogen peroxide." Antioxid Redox Signal 15(1): 191-203.
Guo, J. Y., A. Yamada, et al. (2008). "Aven-dependent activation of ATM following DNA damage." Curr Biol 18(13): 933-942.
Harding, H. P., Y. Zhang, et al. (2003). "An integrated stress response regulates amino acid metabolism and resistance to oxidative stress." Mol Cell 11(3): 619-633.
Herrero, E., J. Ros, et al. (2008). "Redox control and oxidative stress in yeast cells." Biochim Biophys Acta 1780(11): 1217-1235.
Holcik, M. and N. Sonenberg (2005). "Translational control in stress and apoptosis." Nat Rev Mol Cell Biol 6(4): 318-327.
Kim, H., J. Huang, et al. (2007). "CCDC98 is a BRCA1-BRCT domain-binding protein involved in the DNA damage response." Nat Struct Mol Biol 14(8): 710-715.
Klaunig, J. E. and L. M. Kamendulis (2004). "The role of oxidative stress in carcinogenesis." Annu Rev Pharmacol Toxicol 44: 239-267.
Komar, A. A. and M. Hatzoglou (2011). "Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states." Cell Cycle 10(2): 229-240.
Kutuk, O., S. G. Temel, et al. (2010). "Aven blocks DNA damage-induced apoptosis by stabilising Bcl-xL." Eur J Cancer 46(13): 2494-2505.
Lee, J. G., K. Baek, et al. (2013). "Reversible inactivation of deubiquitinases by reactive oxygen species in vitro and in cells." Nat Commun 4: 1568.
Lee, Y. Y., R. C. Cevallos, et al. (2009). "An upstream open reading frame regulates translation of GADD34 during cellular stresses that induce eIF2alpha phosphorylation." J Biol Chem 284(11): 6661-6673.
Leprivier, G., B. Rotblat, et al. (2015). "Stress-mediated translational control in cancer cells." Biochim Biophys Acta 1849(7): 845-860.
Medicherla, B. and A. L. Goldberg (2008). "Heat shock and oxygen radicals stimulate ubiquitin-dependent degradation mainly of newly synthesized proteins." J Cell Biol 182(4): 663-673.
Ng, H. M., L. Wei, et al. (2016). "The K63-Deubiquitylating Enzyme BRCC36 Limits DNA Break Processing and Repair." J Biol Chem.
Paydas, S., K. Tanriverdi, et al. (2003). "Survivin and aven: two distinct antiapoptotic signals in acute leukemias." Ann Oncol 14(7): 1045-1050.
Piques, M., W. X. Schulze, et al. (2009). "Ribosome and transcript copy numbers, polysome occupancy and enzyme dynamics in Arabidopsis." Mol Syst Biol 5: 314.
Powley, I. R., A. Kondrashov, et al. (2009). "Translational reprogramming following UVB

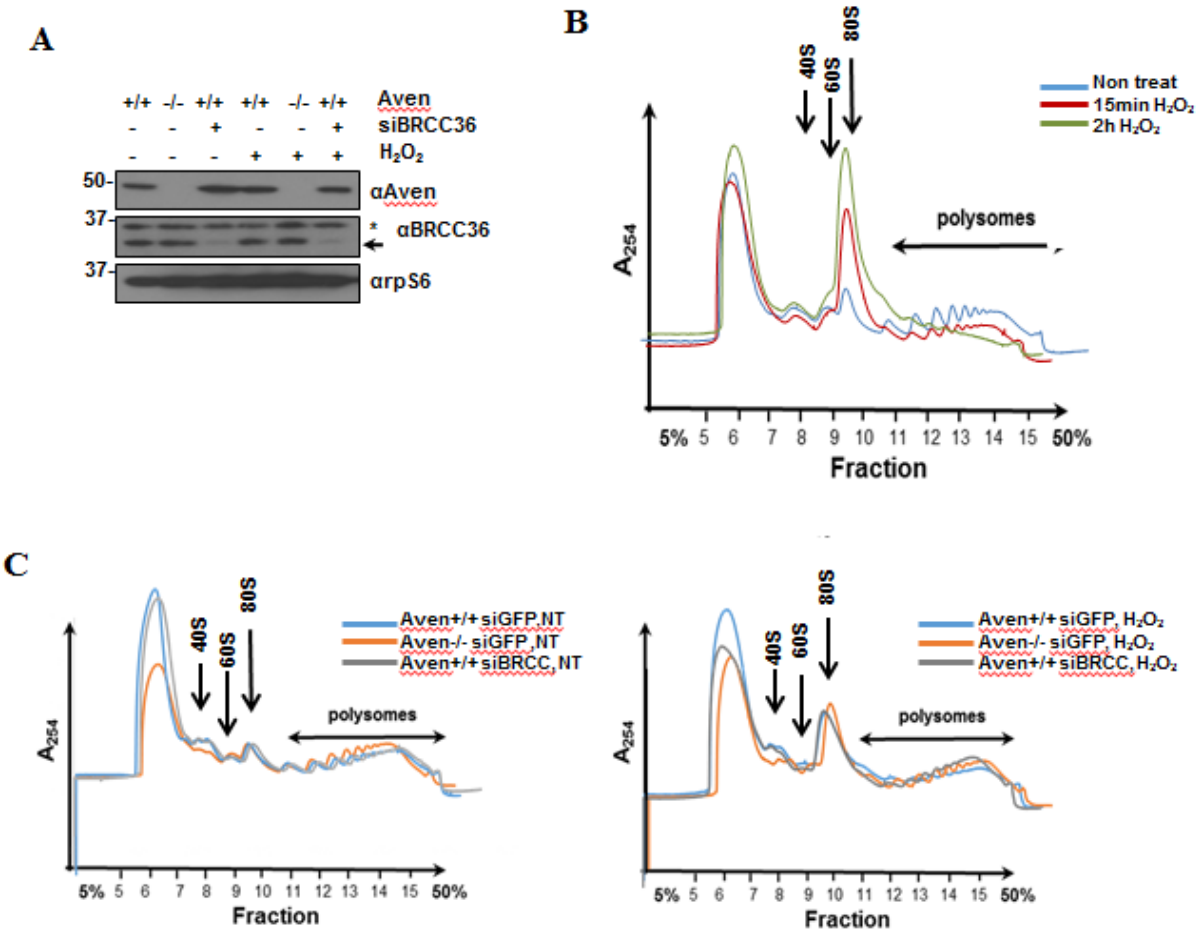
irradiation is mediated by DNA-PKcs and allows selective recruitment to the polysomes of mRNAs encoding DNA repair enzymes." Genes Dev 23(10): 1207-1220.
Rubio, C. A., B. Weisburd, et al. (2014). "Transcriptome-wide characterization of the eIF4A signature highlights plasticity in translation regulation." Genome Biol 15(10): 476.
Saito, K., W. Horikawa, et al. (2015). "Inhibiting K63 polyubiquitination abolishes no-go type stalled translation surveillance in Saccharomyces cerevisiae." PLoS Genet 11(4): e1005197.
Shahid, R., A. Bugaut, et al. (2010). "The BCL-2 5' untranslated region contains an RNA G-quadruplex-forming motif that modulates protein expression." Biochemistry 49(38): 8300-8306.
Shao, G., J. Patterson-Fortin, et al. (2009). "MERIT40 controls BRCA1-Rap80 complex integrity and recruitment to DNA double-strand breaks." Genes Dev 23(6): 740-754.
Shenton, D., J. B. Smirnova, et al. (2006). "Global translational responses to oxidative stress impact upon multiple levels of protein synthesis." J Biol Chem 281(39): 29011-29021.
Shringarpure, R., T. Grune, et al. (2003). "Ubiquitin conjugation is not required for the degradation of oxidized proteins by proteasome." J Biol Chem 278(1): 311-318.
Silva, G. M., D. Finley, et al. (2015). "K63 polyubiquitination is a new modulator of the oxidative stress response." Nat Struct Mol Biol 22(2): 116-123.
Sobhian, B., G. Shao, et al. (2007). "RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites." Science 316(5828): 1198-1202.
Somers, J., L. A. Wilson, et al. (2015). "A common polymorphism in the 5' UTR of ERCC5 creates an upstream ORF that confers resistance to platinum-based chemotherapy." Genes Dev 29(18): 1891-1896.
Susnow, N., L. Zeng, et al. (2009). "Bcl-2 family proteins as regulators of oxidative stress." Semin Cancer Biol 19(1): 42-49.
Thandapani, P., T. R. O'Connor, et al. (2013). "Defining the RGG/RG motif." Mol Cell 50(5): 613-623.
Thandapani, P., J. Song, et al. (2015). "Aven recognition of RNA G-quadruplexes regulates translation of the mixed lineage leukemia protooncogenes." Elife 4.
Topisirovic, I., G. J. Gutierrez, et al. (2009). "Control of p53 multimerization by Ubc13 is JNK-regulated." Proc Natl Acad Sci U S A 106(31): 12676-12681.
Wang, B., K. Hurov, et al. (2009). "NBA1, a new player in the Brca1 A complex, is required for DNA damage resistance and checkpoint control." Genes Dev 23(6): 729-739.
Wang, B., S. Matsuoka, et al. (2007). "Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response." Science 316(5828): 1194-1198.
Wolfe, A. L., K. Singh, et al. (2014). "RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer." Nature 513(7516): 65-70.
Zeqiraj, E., L. Tian, et al. (2015). "Higher-Order Assembly of BRCC36-KIAA0157 Is Required for DUB Activity and Biological Function." Mol Cell 59(6): 970-983.
Zhang, J., M. Cao, et al. (2014). "ABRO1 suppresses tumourigenesis and regulates the DNA damage response by stabilizing p53." Nat Commun 5: 5059.
Zheng, H., V. Gupta, et al. (2013). "A BRISC-SHMT complex deubiquitinates IFNAR1 and regulates interferon responses." Cell Rep 5(1): 180-193.
Zhou, H., I. Wertz, et al. (2004). "Bcl10 activates the NF-kappaB pathway through

ubiquitination of NEMO." Nature 427(6970): 167-171.

5.9 Supplementary Figures



Supplementary Figure 5.1 Aven and BRISC are in the polysomes
A) HEK293T cells overexpressing FLAG-Aven were treated with DMSO or puromycin after 48 h transfection. The distribution of FLAG-Aven, Abro1 and BRCC36 across the gradient was monitored by immunoblotting. The exposure time was determined using a standard curve with increasing amounts of lysates expressing FLAG-Aven immunoblotted with anti-FLAG antibodies for various times. The unspecific band of BRCC36 is denoted as *.

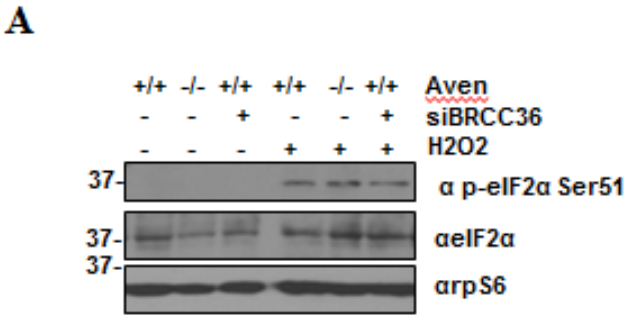


Supplementary Figure 5.2 Aven and BRCC36 do not affect global translation

A) Aven $^{+/+}$ or Aven$^{-/-}$ cells were either transfected with siGFP or siBRCC36. After 48 h, the cells were treated with 1 mM H_2O_2 for 15 min or left untreated and subsequently collected before submitting to polysomal fractionation and immunoblotting with anti-Aven, -BRCC36, and -rpS6 antibodies. The rpS6 served as a loading control. The unspecific band of BRCC36 is denoted as *.

B) HEK293T cells were treated H_2O_2 for 15 min, 2 h or left untreated. Polysome profiles were obtained by continuous monitoring of UV absorbance at 254 nm. 40S, 60S and 80S indicate the positions of the respective ribosomal subunits and monomers on the gradient.

C) Cytoplasmic extracts from (A) were sedimented by centrifugation on a 5 to 50% sucrose gradient. Polysome profiles were obtained by continuous monitoring of UV absorbance at 254nm. 40S, 60S and 80S indicate the positions of the respective ribosomal subunits and the monomer on the gradient.



Supplementary Figure 5.3 Aven and BRCC36 do not regulate eIF2α phosphorylation
A) Aven $^{+/+}$ or Aven$^{-/-}$ cells were either transfected with siGFP or siBRCC36. After 48 h, the cells were treated with 1 mM H_2O_2 for 15 min or left untreated and subsequently collected before submitting to polysomal fractionation and immunoblotting with anti-pSer51 eIF2α, -eIF2α, and -rpS6 antibodies. The rpS6 served as a loading control.

Chapter 6

General discussion

Overall, my work demonstrates new important roles of two RGG/RG motif-containing proteins, Sam68 and Aven, in pre-RNA splicing and mRNA translation. Firstly, my work reveals a new alternative splicing event regulated by Sam68. I show that Sam68 represses the expression of a small isoform of S6K1 called p31, by directly binding to the intron of *ribosomal S6 kinase (Rps6kb1)* pre-mRNA to regulate its alternative splicing. p31 is a nuclear protein whose overexpression leads to the adipogenesis defect. These findings demonstrate that Sam68 regulates the adipogenesis by modulating the alternative splicing of *Rps6kb1*. In the second part of my work, we identify that Aven, which has been shown to be an anti-apoptotic protein, has pro-survival activity. We report that Aven is methylated by PRMT1, allowing it to associate with TDRD3 and SMN in polysomes. We demonstrate that Aven regulates the mRNA translation of MLL1 and MLL4, thereby playing a pivotal role in leukemia cell survival. Additionally, we identify a novel role of Aven as a component of the BRISC complex. We report that Aven and BRCC36 modulate K63 ubiquitination homeostasis in polysomes under oxidative stress, and regulate the mRNA translation of stress-responsive genes. We also show that Aven and BRCC36 are survival factors under oxidative stress.

6.1 A proposed role of Sam68 in alternative splicing

Sam68 has been generally shown to function like an SR protein that regulates alternative splicing and participates in different cellular processes. In my work, we report that Sam68 exerts a suppressive effect on the alternative splicing of *Rps6kb1* in adipocytes. Sam68 binds a “UAAUUAAA” sequence in intron 6 of *Rps6kb1* pre-mRNA, 46 nucleotides downstream of the 5’ splice site as shown *in vitro* with purified components as well as *in vivo* using CLIP and a minigene assay. By associating with this intronic Sam68 binding site, Sam68 promotes the skipping of *Rps6kb1* exons 6a, 6b and 6c located in intron 6, thus preventing the expression of *Rps6kb1-002* mRNA which contains the three alternative exons. Next to the

Sam68 binding site, a binding site of SRSF1 in the exon 6 of *Rps6kb1* was also observed upstream of the 5' splice site. It has been shown that SRSF1 regulates alternative splicing of *Rps6kb1,* generating the *Rps6kb1* isoform-2 (*Rps6kb1-002*) (Ben-Hur, Denichenko et al. 2013), but its *Rps6kb1* binding site(s) and how it regulates the production of *Rps6kb1-002* remains unknown. I show that SRSF1 displays reduced binding to its *Rps6kb1* exon 6 "GAAAGAGAGGGAA" site in the presence of Sam68 by CLIP assay. Additionally, I observe that Sam68 regulates the binding of SRSF1 to its RNA binding site by directly interacting with SRSF1. Thus, by binding to SRSF1 and competing with its positive effect on S6K1 splicing, Sam68 prevents the production of *Rps6kb1* isoform-2. Mechanistically, several possible mechanisms have been proposed to explain how RBPs regulate splicing (Witten and Ule 2011; Das and Krainer 2014; Fu and Ares 2014). RBPs, mainly SR proteins and hnRNPs, bind to splicing enhancers or silencers to activate or inhibit splicing at nearby splice sites, by regulating the overall ability of spliceosome to recognize 5' splice sites. Thus, the fate of an mRNA is decided by the antagonism between RBPs (Erkelenz, Mueller et al. 2013). In sum, I have shown that Sam68 and SRSF1 compete for binding to the regulatory elements of *Rps6kb1* pre-mRNA.

It was previously demonstrated that Sam68 binds an AU-rich sequence in the intron 5 of *mTOR* to regulate its alternative splicing. Depletion of Sam68 leads to the retention of intron 5 and expression of a small isoform of *mTOR* (*mTORi5*) (Huot, Vogel et al. 2012). Intriguingly, for both *mTOR* and *Rps6kb1*, Sam68 binds the intron near the 5' splice and promotes the expression of the canonical long isoform. Thus, it is very likely that Sam68 facilitates the usage of the stronger polyadenylation site (i.e the distal polyadenylation site rather than the proximal one) to make longer isoforms. Consistently, important studies show that U1 snRNP, a spliceosome component that binds the 5' splice site, determines the usage of polyadenylation sites and promotes the production of long canonical isoforms (Kaida, Berg et al. 2010; Berg, Singh et al. 2012). Moreover, SR proteins are important for U1 snRNP recruitment, and SRSF1 interacting with U1 is required for 5' splice site recognition and ternary complex (pre-mRNA/U1 snRNP) formation (Jamison, Pasman et al. 1995; Eperon, Makarova et al. 2000). Thus, it is possible that Sam68 influences alternative polyadenylation

by directly recruiting U1 snRNP and/or by associating with SRSF1 (Figure 6.1). Thus, a general model denoting the role of Sam68 in gene regulation is proposed, though future work is required to verify this model.

Future directions will be focused on performing Sam68 genome-wide RNA-seq on long/short isoform expression, including polyA seq, TAIL-seq, and SAGE (Serial analysis of gene expression), to characterize the role of Sam68 in alternative polyadenylation. The 3' RACE assay could also be performed to examine the usage of polyadenylation signals (PAS) of targets identified from the RNA-seq. Next, it would be interesting to study the detailed mechanism of how Sam68 regulates alternative polyadenylation. To address this, immunoprecipitation could be done to determine the association between Sam68 and U1 snRNP. Alternatively, a siRNA library, CRISPR/Cas9 mediated mutagenesis or candidate approach can also be carried out to identify termination factors involved in Sam68-regulated alternative polyadenylation. Lastly, since alternative polyadenylation has been intensively linked to tumorigenesis and diseases (Batra, Charizanis et al. 2014), it would also be interesting to study this regulation in tumor tissues compared to normal tissues.

6.2 Alternative splicing of S6K1 plays a major role in adipogenesis

Sam68 null mice are lean, with decreased mTOR protein level and downstream signaling (Huot, Vogel et al. 2012). In my thesis, we show that *Rps6kb1-002* and its encoded protein, p31S6K1, are present in Sam68-depleted pre-adipocytes and mouse white adipose tissue of Sam68 null mice. Using an anti-S6K1 antibody that recognizes all the isoforms sharing the common N-terminus, we observed that Sam68 deficiency leads to increased p31S6K1 expression, without a detectable reduction in expression of p70S6K1 and p85S6K1, the two canonical isoforms of S6K1. *S6K1*$^{-/-}$ mice have decreased adipose tissue mass, increased energy expenditure, and are resistant to dietary-induced obesity (Um, Frigerio et al. 2004). p70/p85S6K1 participates in the up-regulation of transcription factors during the commitment phase of adipogenesis (Carnevalli, Masuda et al. 2010). Adipocytes normally express p70/p85S6K1, but not p31S6K1. Unlike p70/p85S6K1, p31S6K1 is a nuclear protein that lacks kinase activity due to its truncated kinase domain (Ben-Hur, Denichenko et al. 2013),

suggesting that it has independent properties relative to p70/p85S6K1. Indeed, we show that mouse p31 is a potent suppressor of adipogenesis, since overexpression of p31 dramatically inhibits adipogenesis and depletion of p31 rescues the adipogenesis defect of the Sam68-null adipocytes. Although the mechanism of action by which it represses adipogenesis was not established in my work, we propose that it likely acts as a co-regulator of transcription factors to facilitate the transcription of adipogenesis markers.

Further studies are required to investigate the biochemical property of p31 and its functions. For instance, Bio-ID analysis can be performed to identify p31S6K1 associated proteins. Moreover, DNA pull-down and luciferase reporter assays can also be carried out to examine if it harbors DNA binding ability.

6.3 A novel G4 RNA-binding protein and its role in translation

G4 structures are prevalently found at 5'-and/or 3'-UTR, where they affect RNA processing such as alternative splicing, polyadenylation, mRNA targeting and translation (Bugaut and Balasubramanian 2012; Millevoi, Moine et al. 2012). G4 structures in UTRs and ORFs influence different stages of mRNA translation. G4 structures within the 5′-UTRs of mRNAs reduce cap-dependent translation by preventing assembly of the translational initiation machinery at the 5′-cap, and also impair the scanning of the start site AUG by the initiation complex (Beaudoin and Perreault 2010; Bugaut and Balasubramanian 2012). G4 motifs in long, structured 5' UTRs have been shown to require eIF4A for optimal translation output (Wolfe, Singh et al. 2014). Notably, a recent genome-wide rG4 seq analysis illustrated that G4 structures are enriched in UTRs and near PAS sequences, indicating their role in mRNA processing and stability (Kwok, Marsico et al. 2016). G-quadruplex structures within ORFs of the virally encoded EBNA1 transcript were shown to hinder translational elongation by either ribosomal pausing or ribosomal dissociation (Murat, Zhong et al. 2014). In my work, we extend these studies to show that ORFs of mRNAs also contain G4 sequences (~1600 pG4s in human ORFs) that encode low complexity sequences and function to regulate translation as well. Specifically, we identify G4 structures that regulate its polysomal association and

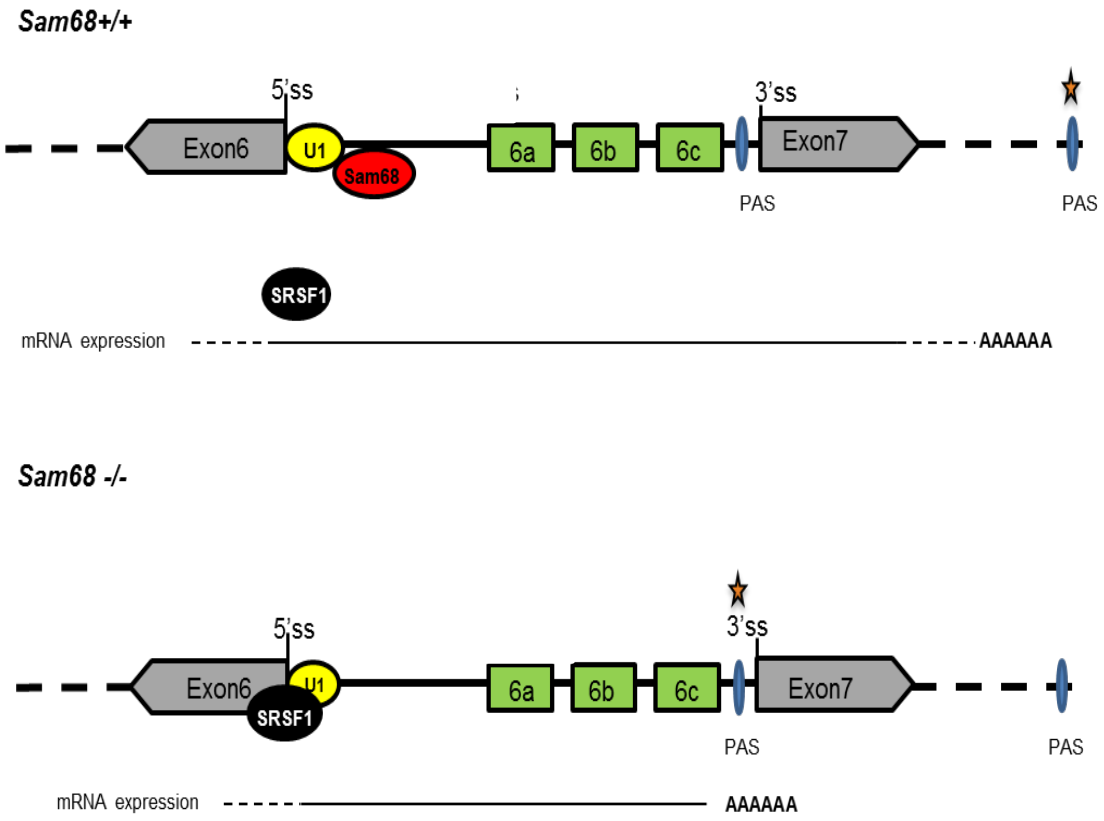


Figure 6.1 Proposed model denoting the role of Sam68 in the regulation of S6K1 alternative polyadenylation.

Nucleotide sequences of exon6, intron 6 and exon 7 from S6K1 pre-mRNA are deciphered and the expressed mRNAs are shown below. When Sam68 is abundant, Sam68 binds the U1snRNP, promoting the usage of the distal PAS and the production of longer isoforms. When Sam68 is depleted, SRSF1 is in close proximity with U1 snRNP, which stimulates the usage of proximal PAS and generates the shorter isoforms. Grey boxes represent exon6 and exon7 while green boxes indicate alternative exons (exon 6a, 6b and 6c). U1 snRNP, Sam68 and SRSF1 are shown. 5'ss, 5' splice site. 3'ss, 3' splice site. PAS, polyadenylation signal. The star denotes the used polyadenylation signal.

protein synthesis within the mRNA coding regions of MLL1 and MLL4. These RNA elements are located between 200 and 300 nucleotides downstream of the initiator methionine ATG, and encode protein sequences rich in glycines and arginine-glycine repeats in MLL1 and MLL4. Mechanistically, we speculate that these G-quadruplexes in ORFs may act as 'roadblocks', playing a significant role in protein synthesis by inhibiting ribosomal progression during elongation, as proposed by Endoh & Sugimoto (Endoh and Sugimoto 2016).

Proteomic and genome-wide studies demonstrated that G4 RNAs are bound by several ribosomal proteins, splicing factors and hnRNPs that contain RGG/RG motifs (von Hacht, Seifert et al. 2014; Anderson, Chopra et al. 2016). FMRP represents a well-studied example as a G4 binding protein. In addition to binding to the G4s, the RGG/RG motif of FMRP also modulates its association with polysomes (Blackwell, Zhang et al. 2010), being consistent with its role in translational regulation (Corbin, Bouillon et al. 1997; Richter, Bassell et al. 2015). In this thesis, we show that a novel RGG/RG motif-containing protein, Aven, binds to G-quadruplexes of MLL1 and MLL4 mRNA ORFs *in vitro* and *in cellulo* via its RGG/RG motif. Similarly to FMRP, Aven is associated with polysomes via its RGG/RG motif in a PRMT1/arginine methylation-dependent manner. Aven regulates the translation of MLL1 and MLL4 in the polysomes, as shown by luciferase reporter assay and polysome fractionation-qPCR. Thus, we provide new insights that G4 structures in ORFs regulate translation, and Aven is the novel modulator that binds G4s.

Future direction will be focused on the detailed mechanism of how Aven functions in translational regulation. Firstly, it would be of great interest to identify other targets of Aven to examine how Aven achieves specificity for G4 binding. This could be addressed by performing genome-wide polysomal RNA-seq, to isolate mRNAs that are highly translated in WT or Aven-depleted cells. Ribosomal footprinting would be an interesting approach as well to study how Aven affects the translation elongation. Additionally, Aven CLIP-seq would be useful to validate the targets of Aven and investigate its consensus binding sequences/structures. Finally, CRISPR would also be interesting to specifically delete the G4 sequence, for studying the function of G4s.

6.4 Arginine methylation of Aven and implications in leukemia

Aven is a survival factor, since it prevents apoptosis by stabilizing pro-survival protein Bcl-xl and inhibiting Apaf-1 assembly within the apoptosome (Chau, Cheng et al. 2000). Moreover, Aven is overexpressed in acute leukemia patients and has been associated with poor prognosis in acute childhood myeloid leukemia (Choi, Hwang et al. 2006). Aven was also identified as being fundamental for the stimulation and progression of acute leukemia (Eissmann, Melzer et al. 2013). We report a new mechanism whereby Aven regulates the translation of MLL1 and MLL4 mRNA and modulates their protein levels in leukemia cell lines, which is impeded by PRMT1 depletion. Thus, Aven promotes leukemic cell proliferation and growth in an arginine methylation-dependent manner.

Mixed lineage family proteins (MLL1-MLL5) are positive transcription regulators, belonging to the evolutionary conserved family of trithorax group (trxG) proteins. They possess the H3K4 methyltransferase activity and are the main components of multiprotein complex containing WDR5, RbBP5 and ASH2L (Muntean and Hess 2012). The N-terminus of MLL1 harboring the G4 structures can fuse with different partners, a translocation which is frequently observed in infant acute myeloid and lymphoid leukemias. MLL1 fusion proteins increase the expression of Hox genes and thus lead to haematopoietic malignancies (Muntean and Hess 2012). We observed that knockdown of Aven decreases *Hox* gene expression, resulting in cell death of CCRF-CEM and MOLT-4 cells (human T-leukemia cell lines).

We show that PRMT1 also plays an essential role in Aven-mediated translation reprogramming in leukemic development. Consistent with our observation, PRMT1 was shown to be involved in development of hematopoietic cells harboring MLL-EEN translocation. MLL-EEN recruits Sam68 via the SH3 domain of EEN, which then associates with PRMT1. PRMT1 then modulates MLL activity on histone acetylation and H4R3 methylation, regulating MLL downstream targets (Cheung, Chan et al. 2007). Our results provide evidence that PRMT1 functions in leukemic development and define a new mechanism of how PRMT1 regulates mRNA translation. These findings shed light on the potential applications of PRMT1 inhibitors and G4 stabilizing ligands for treating patients

with T-ALL.

Further studies will be focused on targeting G4 structures with synthesized small molecules to see whether they can modify oncogene expression and cell proliferation in T-ALL cells. Additionally, *Aven*-null mice could be generated to study its physiological functions.

6.5 A novel role of BRISC-Aven complex under oxidative stress

In my thesis, I characterize the functional interaction between Aven and the components of the BRISC DUB complex, namely the subunits called Abro1 and BRCC36. We confirm their association by co-immunoprecipitation experiments and identify Abro1 and BRCC36 to be associated with polyribosomes in the cytoplasm. BRISC catalyzing K63-linked polyubiquitin plays distinct roles in the nucleus and cytoplasm by associating with different adaptor proteins, ABRAXAS and Abro1 (Feng, Wang et al. 2010). In the nucleus, this complex plays a role in DNA damage response (Sobhian, Shao et al. 2007). It is well-known that RNF8 and RNF168 trigger a cascade of regulatory processes of ubiquitination events at DSB sites (Panier and Durocher 2013), leading to the recruitment of repair factors such as 53BP1 and RAP80/BRCA1 (Thorslund, Ripplinger et al. 2015). This ubiquitination process is counteracted by BRISC for subsequent ubiquitin chain editing (Panier and Durocher 2013).

However, the role of this complex in the cytoplasm is poorly elucidated. It has been reported that serine hydroxymethyltransferase (SHMT) directs BRISC activity at K63 ubiquitin chains conjugated to the type I interferon receptor 1 (IFNAR1), which mediates type I interferon signaling (Zheng, Gupta et al. 2013). We show that oxidative stress by H_2O_2 treatment leads to an accumulation of proteins harboring K63-linked ubiquitination, consistent with previous findings (Silva, Finley et al. 2015). K63 ubiquitinated proteins are accumulated in response to H_2O_2 which is further potentiated by Aven- or BRCC36-depletion. BRCC36 deubiquitinating activity, which is modulated by Aven, functions as a regulator of the BRISC complex *in vivo*, while H_2O_2 triggered the dissociation of the Abro1 from the Aven/BRISC complex in polysomes. For the first time, we show that the BRISC complex is associated with polysomes, and we provide a novel function of cytosolic BRISC complex

under oxidative stress.

Future studies would be focused on studying the dynamic interaction between Aven and BRISC complex. Firstly, the 3D structure analysis would be informative to see how they interact and the biochemical properties of their binding. Also, a DUB activity assay would be necessary to examine whether Aven regulates BRISC DUB activity. Lastly, mass spectrometry, SILAC and TUBE-mass spectrometry are required to see which proteins in the polysome are K63-modified.

6.6 Polysomal K63 ubiquitination is a translational modulator

In our current work, we show that K63 is accumulated in polysomes in human HEK293 cells, thereby affecting the translation of a subset of mRNAs under oxidative stress. Thus we provide insights that K63 is accumulated in polysomes in response to stress, and we elucidate the correlation between K63 and translational regulation in human cells. Polysomal K63 has been well characterized in yeast since 2000, when Spence et al. established that the large ribosomeal subunit L28 is K63-ubiquitinated during G1 of the yeast cell cycle. This modification seems to be evolutionary conserved, since the human ortholog is also ubiquitinated (Spence, Gali et al. 2000). They also observed that K63R mutant is sensitive to translation inhibitors, indicating that polysomal K63 could be a modulator of translation (Spence, Gali et al. 2000). Moreover, this was reanalyzed 15 years later, when Silva et al. discovered that polysomal K63 accumulated in response to H_2O_2 in yeast. SILAC-mass spectrometry analysis identified over 100 new K63 ubiquitinated targets, of which approximately 30% are cytoplasmic translation factors (Silva, Finley et al. 2015).

Consistently, we show that although global translation is severely inhibited in response to H_2O_2, the Aven/BRISC complex appears to stimulate translation of specific mRNA under these conditions, possibly by regulating K63 ubiquitination. We observe ERCC1, ERCC5, Bcl-2, and ATF4, all known to have structured 5’ UTR or uORFs (Lee, Cevallos et al. 2009; Shahid, Bugaut et al. 2010; Somers, Wilson et al. 2015), to be regulated by Aven and BRCC36 in an eIF2α–independent manner. Mechanistically, it has been illustrated that abnormally elongated K63 chains in the yeast polysomes recruit the proteasome to dissociate

the ribosome, which is implicated in the quality control system (Saito, Horikawa et al. 2015). Based on our and other studies, we propose that the proper length of K63 polyubiquitin chains is fundamental for efficient translation of mRNAs under stress conditions, and that the disruption of K63-linked ubiquitination homeostasis leads to aberrant translation causing pro-survival adaptation events to be compromised. Additionally, we propose that under stress conditions, when eIF4F activity is compromised and protein synthesis is severely inhibited, phosphorylation of the eIF2α subunit induces selective translation of mRNAs containing upstream open reading frames (uORFs) by reinitiation, as well as the internal translation initiation (IRES) in a cap-independent manner (Holcik and Sonenberg 2005; Komar and Hatzoglou 2011). For both cases, these selective mRNAs harbor typically long, GC-rich, highly structured 5'-UTR (Komar and Hatzoglou 2011). Similarly, we observe ERCC1, ERCC5, Bcl-2, and ATF4, known to have structured 5' UTR or uORFs (Lee, Cevallos et al. 2009; Shahid, Bugaut et al. 2010; Somers, Wilson et al. 2015), to be regulated by Aven and BRCC36 in an eIF2α independent manner. Furthermore, Shahid et al. 2010, have shown that human Bcl-2 forms G4 structures in the 42 nucleotides upstream of its translation start site, which is consistent with previous finding that Aven regulates specific mRNA translation with G4 structures (Shahid, Bugaut et al. 2010; Thandapani, Song et al. 2015).

In the future, it would be worthwhile to investigate how accumulated K63 chains affect the translation of a specific subset of mRNAs. Moreover, it is also of interest to examine which proteins and which lysines in polysomes (i.e. ribosomal proteins) are K63-modified, and whether mutation of these lysines abolishes Aven/BRCC effect on translation and cell survival under stress.

6.7 Concluding remarks

The knowledge of RGG/RG motif-containing proteins has been expanded over the last several years. In addition to Sam68, increasing amounts of other RGG/RG motif-containing proteins including Aven have been shown to bind RNAs. These proteins have been implicated in a wide range of cellular processes. In the future, more work needs to be done to elucidate

the functions of RGG/RG motifs in various aspects of RNA metabolism. A comprehensive characterization of their consensus binding structures or nucleic acid sequences is needed to further understand the role of arginine methylation. Moreover, the detailed mechanisms involved in these regulatory processes will be essential for investigating whether misregulation is linked to cancer or other diseases.

CONTRIBUTIONS TO ORIGINAL KNOWLEDGE

The present work focused on the identification of new important roles of two RGG/RG motif-containing proteins, Sam68 and Aven, in pre-RNA splicing and mRNA translation. The functional significance of these methylation events has been addressed. The major contributions of this work to original knowledge are summarized below:

1. The characterization of a new mechanism of Sam68 in alternative splicing regulation.
2. The identification of p31S6K1 as a novel potent adipogenesis regulator.
3. A bioinformatic search identified ~1600 G4s in ORFs in human transcriptome.
4. The identification and characterization of the RGG/RG motif of the cell death regulator protein Aven.
5. The characterization of the role of arginine methylation in regulating Aven mediated translation of leukemic genes.
6. The identification of Aven as a component/regulator of cytosolic BRISC.
7. The characterization of BRCC36 and Aven to be essential for the translation of specific mRNAs and for cell survival in response to oxidative stress.

REFERENCES

Abdelmohsen, K. and M. Gorospe (2012). "RNA-binding protein nucleolin in disease." RNA Biol **9**(6): 799-808.

Abdelmohsen, K., K. Tominaga, et al. (2011). "Enhanced translation by Nucleolin via G-rich elements in coding and non-coding regions of target mRNAs." Nucleic Acids Res **39**(19): 8513-8530.

Agarwala, P., S. Pandey, et al. (2013). "The G-quadruplex augments translation in the 5' untranslated region of transforming growth factor beta2." Biochemistry **52**(9): 1528-1538.

Akawi, N., J. McRae, et al. (2015). "Discovery of four recessive developmental disorders using probabilistic genotype and phenotype matching among 4,125 families." Nat Genet **47**(11): 1363-1369.

Al-Dhaheri, M., J. Wu, et al. (2011). "CARM1 is an important determinant of ERalpha-dependent breast cancer cell differentiation and proliferation in breast cancer cells." Cancer Res **71**(6): 2118-2128.

Alinari, L., K. V. Mahasenan, et al. (2015). "Selective inhibition of protein arginine methyltransferase 5 blocks initiation and maintenance of B-cell transformation." Blood **125**(16): 2530-2543.

Allison, L. A., M. Moyle, et al. (1985). "Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases." Cell **42**(2): 599-610.

Altan, B., T. Yokobori, et al. (2016). "Nuclear PRMT1 expression is associated with poor prognosis and chemosensitivity in gastric cancer patients." Gastric Cancer **19**(3): 789-797.

Altschuler, L., J. O. Wook, et al. (1999). "Involvement of receptor-bound protein methyltransferase PRMT1 in antiviral and antiproliferative effects of type I interferons." J Interferon Cytokine Res **19**(2): 189-195.

Ambivero, C. T., L. Cilenti, et al. (2012). "ATF4 interacts with Abro1/KIAA0157 scaffold protein and participates in a cytoprotective pathway." Biochim Biophys Acta **1823**(12): 2149-2156.

An, W., J. Kim, et al. (2004). "Ordered cooperative functions of PRMT1, p300, and CARM1 in transcriptional activation by p53." Cell **117**: 735-748.

Anderson, B. R., P. Chopra, et al. (2016). "Identification of consensus binding sites clarifies FMRP binding determinants." Nucleic Acids Res **44**(14): 6649-6659.

Anderson, P. and N. Kedersha (2008). "Stress granules: the Tao of RNA triage." Trends Biochem Sci **33**(3): 141-150.

Anderson, P. and N. Kedersha (2009). "RNA granules: post-transcriptional and epigenetic modulators of gene expression." Nat Rev Mol Cell Biol **10**(6): 430-436.

Anko, M. L., M. Muller-McNicoll, et al. (2012). "The RNA-binding landscapes of two SR proteins reveal unique functions and binding to diverse RNA classes." Genome Biol **13**(3): R17.

Antonysamy, S., Z. Bonday, et al. (2012). "Crystal structure of the human PRMT5:MEP50 complex." Proc Natl Acad Sci U S A **109**(44): 17960-17965.

Aoto, J., D. C. Martinelli, et al. (2013). "Presynaptic neurexin-3 alternative splicing trans-synaptically controls postsynaptic AMPA receptor trafficking." Cell **154**(1): 75-88.

Apel, K. and H. Hirt (2004). "Reactive oxygen species: metabolism, oxidative stress, and signal transduction." Annu Rev Plant Biol **55**: 373-399.

Arora, A., M. Dutkiewicz, et al. (2008). "Inhibition of translation in living eukaryotic cells by an RNA G-quadruplex motif." RNA **14**(7): 1290-1296.

Arora, A. and B. Suess (2011). "An RNA G-quadruplex in the 3' UTR of the proto-oncogene PIM1 represses translation." RNA Biol **8**(5): 802-805.

Avasarala, S., M. Van Scoyk, et al. (2015). "PRMT1 Is a Novel Regulator of Epithelial-Mesenchymal-Transition in Non-small Cell Lung Cancer." J Biol Chem **290**(21): 13479-13489.

Ayroldi, E., M. G. Petrillo, et al. (2015). "L-GILZ binds p53 and MDM2 and suppresses tumor growth through p53 activation in human cancer cells." Cell Death Differ **22**(1): 118-130.

Babic, I., E. Cherry, et al. (2006). "SUMO modification of Sam68 enhances its ability to repress cyclin D1 expression and inhibits its ability to induce apoptosis." Oncogene **25**(36): 4955-4964.

Babic, I., A. Jakymiw, et al. (2004). "The RNA-binding protein Sam68 is acetylated in tumor cell lines, and its acetylation correlates with enhanced RNA binding activity." Oncogene **23**(21): 3781-3789.

Bachand, F. and P. A. Silver (2004). "PRMT3 is a ribosomal protein methyltransferase that affects the cellular levels of ribosomal subunits." EMBO J. **23**: 2641-2650.

Baird, S. D., M. Turcotte, et al. (2006). "Searching for IRES." RNA **12**(10): 1755-1785.

Baird, T. D. and R. C. Wek (2012). "Eukaryotic initiation factor 2 phosphorylation and translational control in metabolism." Adv Nutr **3**(3): 307-321.

Balasubramanian, S., L. H. Hurley, et al. (2011). "Targeting G-quadruplexes in gene promoters: a novel anticancer strategy?" Nat Rev Drug Discov **10**(4): 261-275.

Baldwin, R. M., M. Bejide, et al. (2015). "Identification of the PRMT1v1 and PRMT1v2 specific interactomes by quantitative mass spectrometry in breast cancer cells." Proteomics **15**(13): 2187-2197.

Balkwill, G. D., K. Derecka, et al. (2009). "Repression of translation of human estrogen receptor alpha by G-quadruplex formation." Biochemistry **48**(48): 11487-11495.

Baltz, A. G., M. Munschauer, et al. (2012). "The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts." Mol Cell **46**: 674-690.

Bang, I. (1910). "Untersuchungen über die Guanylsäure

Biochemische Zeitschrift." **26**: 293-311.

Baranski, Z., T. H. Booij, et al. (2015). "Aven-mediated checkpoint kinase control regulates proliferation and resistance to chemotherapy in conventional osteosarcoma." J Pathol **236**(3): 348-359.

Barbosa-Morais, N. L., M. Irimia, et al. (2012). "The evolutionary landscape of alternative splicing in vertebrate species." Science **338**(6114): 1587-1593.

Bartkowiak, B., P. Liu, et al. (2010). "CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1." Genes Dev **24**(20): 2303-2316.

Bates, P., J. L. Mergny, et al. (2007). "Quartets in G-major. The First International Meeting on Quadruplex DNA." EMBO Rep **8**: 1003-1010.

Batra, R., K. Charizanis, et al. (2014). "Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease." Mol Cell **56**(2): 311-322.

Batsche, E., M. Yaniv, et al. (2006). "The human SWI/SNF subunit Brm is a regulator of alternative splicing." Nat Struct Mol Biol **13**(1): 22-29.

Beaudoin, J. D., R. Jodoin, et al. (2013). "In-line probing of RNA G-quadruplexes." Methods **64**: 79-87.

Beaudoin, J. D., R. Jodoin, et al. (2014). "New scoring system to identify RNA G-quadruplex folding." Nucleic Acids Res **42**(2): 1209-1223.

Beaudoin, J. D. and J. P. Perreault (2010). "5'-UTR G-quadruplex structures acting as translational repressors." Nucleic Acids Res **38**(20): 7022-7036.

Beaudoin, J. D. and J. P. Perreault (2013). "Exploring mRNA 3'-UTR G-quadruplexes: evidence of roles in both alternative polyadenylation and mRNA shortening." Nucleic Acids Res **41**(11): 5898-5911.

Bechara, E. G., M. C. Didiot, et al. (2009). "A novel function for fragile X mental retardation protein in translational activation." PLoS Biol **7**(1): e16.

Bedford, M. T. and S. G. Clarke (2009). "Protein arginine methylation in mammals: who, what, and why." Mol Cell **33**(1): 1-13.

Bedford, M. T., A. Frankel, et al. (2000). "Arginine methylation inhibits the binding of proline-rich ligands to Src homology 3, but not WW, domains." J Biol Chem **275**(21): 16030-16036.

Bedford, M. T. and S. Richard (2005). "Arginine methylation an emerging regulator of protein function." Mol Cell **18**(3): 263-272.

Ben-Hur, V., P. Denichenko, et al. (2013). "S6K1 alternative splicing modulates its oncogenic activity and regulates mTORC1." Cell Rep **3**(1): 103-115.

Bentley, D. L. (2015). "The union of transcription and mRNA processing: 20 years of coupling." RNA **21**(4): 569-570.

Berg, M. G., L. N. Singh, et al. (2012). "U1 snRNP determines mRNA length and regulates isoform expression." Cell **150**(1): 53-64.

Bergeman, J., A. Caillier, et al. (2016). "Localized translation regulates cell adhesion and transendothelial migration." J Cell Sci.

Besnard, E., A. Babled, et al. (2012). "Unraveling cell type-specific and reprogrammable human replication origin signatures associated with G-quadruplex consensus motifs." Nat Struct Mol Biol **19**(8): 837-844.

Bezzi, M., S. X. Teo, et al. (2013). "Regulation of constitutive and alternative splicing by PRMT5 reveals a role for Mdm4 pre-mRNA in sensing defects in the spliceosomal machinery." Genes Dev **27**(17): 1903-1916.

Bhat, M., N. Robichaud, et al. (2015). "Targeting the translation machinery in cancer." Nat Rev Drug Discov **14**(4): 261-278.

Bhattacharya, S. and M. K. Ghosh (2014). "Cell death and deubiquitinases: perspectives in cancer." Biomed Res Int **2014**: 435197.

Bianchi, E., F. Barbagallo, et al. (2010). "Ablation of the Sam68 gene impairs female fertility and gonadotropin-dependent follicle development." Hum Mol Genet **19**(24): 4886-4894.

Bidzinska, J., G. Cimino-Reale, et al. (2013). "G-quadruplex structures in the human genome as novel therapeutic targets." Molecules **18**(10): 12368-12395.

Bielli, P., R. Busà, et al. (2014). "The transcription factor FBI-1 inhibits SAM68-mediated BCL-X alternative splicing and apoptosis." EMBO Rep **15**: 419-427.

Bielli, P., R. Busà, et al. (2011). "The RNA-binding protein Sam68 is a multifunctional player in human cancer." Endocr Relat Cancer **18**: R91-102.

Biffi, G., M. Di Antonio, et al. (2014). "Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells." Nat Chem **6**(1): 75-80.

Bird, J. G., Y. Zhang, et al. (2016). "The mechanism of RNA 5' capping with NAD+, NADH and desphospho-CoA." Nature **535**(7612): 444-447.

Blackwell, E., X. Zhang, et al. (2010). "Arginines of the RGG box regulate FMRP association with polyribosomes and mRNA." Hum Mol Genet **19**(7): 1314-1323.

Blagoev, B., I. Kratchmarova, et al. (2003). "A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling." Nat Biotechnol **21**: 315-318.

Blanc, R. S., G. Vogel, et al. (2016). "PRMT7 Preserves Satellite Cell Regenerative Capacity." Cell Rep **14**(6): 1528-1539.

Blice-Baum, A. C. and M. R. Mihailescu (2014). "Biophysical characterization of G-quadruplex forming FMR1 mRNA and of its interactions with different fragile X mental retardation protein isoforms." RNA **20**(1): 103-114.

Blobel, G. and D. Sabatini (1971). "Dissociation of mammalian polyribosomes into subunits by puromycin." Proc Natl Acad Sci U S A. **68**: 390-394.

Bochman, M. L., K. Paeschke, et al. (2012). "DNA secondary structures: stability and function of G-quadruplex

structures." Nat Rev Genet **13**(11): 770-780.
Boisvert, F.-M., U. Déry, et al. (2005). "Arginine methylation of MRE11 by PRMT1 is required for DNA damage checkpoint control." Genes & Dev **19**(6): 671-676.
Boisvert, F. M., Y. Ahmad, et al. (2012). "A quantitative spatial proteomics analysis of proteome turnover in human cells." Mol Cell Proteomics **22**: M111.011429.
Boisvert, F. M., M. J. Hendzel, et al. (2005). "Methylation of MRE11 regulates its nuclear compartmentalization." Cell Cycle **4**(7): 981-989.
Boisvert, F. M., A. Rhie, et al. (2005). "The GAR motif of 53BP1 is arginine methylated by PRMT1 and is necessary for 53BP1 DNA binding activity." Cell Cycle **4**(12): 1834-1841.
Bonnal, S., C. Schaeffer, et al. (2003). "A single internal ribosome entry site containing a G quartet RNA structure drives fibroblast growth factor 2 gene expression at four alternative translation initiation codons." J Biol Chem **278**(41): 39330-39336.
Booy, E. P., R. Howard, et al. (2014). "The RNA helicase RHAU (DHX36) suppresses expression of the transcription factor PITX1." Nucleic Acids Res **42**: 3346-3361.
Booy, E. P., E. K. McRae, et al. (2015). "Biochemical characterization of G4 quadruplex telomerase RNA unwinding by the RNA helicase RHAU." Methods Mol Biol **1259**: 125-135.
Booy, E. P., M. Meier, et al. (2012). "The RNA helicase RHAU (DHX36) unwinds a G4-quadruplex in human telomerase RNA and promotes the formation of the P1 helix template boundary." Nucleic Acids Res **40**(9): 4110-4124.
Bordeleau, M. E., F. Robert, et al. (2008). "Therapeutic suppression of translation initiation modulates chemosensitivity in a mouse lymphoma model." J Clin Invest **118**(7): 2651-2660.
Botuyan, M. V., J. Lee, et al. (2006). "Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA repair." Cell **127**(7): 1361-1373.
Brahms, H., J. Raymackers, et al. (2000). "The C-terminal RG dipeptide repeats of the spliceosomal Sm proteins D1 and D3 contain symmetrical dimethylarginines, which form a major B-cell epitope for anti-Sm autoantibodies." J Biol Chem **275**(22): 17122-17129.
Bremang, M., A. Cuomo, et al. (2013). "Mass spectrometry-based identification and characterisation of lysine and arginine methylation in the human proteome." Mol Biosyst **9**(9): 2231-2247.
Brooks, T. A., S. Kendrick, et al. (2010). "Making sense of G-quadruplex and i-motif functions in oncogene promoters." FEBS J **277**(17): 3459-3469.
Brown, V., P. Jin, et al. (2001). "Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome." Cell **107**(4): 477-487.
Bugaut, A. and S. Balasubramanian (2012). "5'-UTR RNA G-quadruplexes: translation regulation and targeting." Nucleic Acids Res **40**(11): 4727-4741.
Bugaut, A., R. Rodriguez, et al. (2010). "Small molecule-mediated inhibition of translation by targeting a native RNA G-quadruplex." Org Biomol Chem **8**(12): 2771-2776.
Burge, S., G. N. Parkinson, et al. (2006). "Quadruplex DNA: sequence, topology and structure." Nucleic Acids Res **34**: 5402-5415.
Busa, R., R. Geremia, et al. (2010). "Genotoxic stress causes the accumulation of the splicing regulator Sam68 in nuclear foci of transcriptionally active chromatin." Nucleic Acids Res **38**(9): 3005-3018.
Busa, R., M. P. Paronetto, et al. (2007). "The RNA-binding protein Sam68 contributes to proliferation and survival of human prostate cancer cells." Oncogene **26**(30): 4372-4382.
Butler, J. S., C. I. Zurita-Lopez, et al. (2011). "Protein-arginine methyltransferase 1 (PRMT1) methylates Ash2L, a shared component of mammalian histone H3K4 methyltransferase complexes." J Biol Chem **286**(14):

12234-12244.

Calabretta, S. and S. Richard (2015). "Emerging Roles of Disordered Sequences in RNA-Binding Proteins." Trends Biochem Sci **40**(11): 662-672.

Cammas, A., A. Dubrac, et al. (2015). "Stabilization of the G-quadruplex at the VEGF IRES represses cap-independent translation." RNA Biol **12**(3): 320-329.

Cappellari, M., P. Bielli, et al. (2014). "The transcriptional co-activator SND1 is a novel regulator of alternative splicing in prostate cancer cells." Oncogene **33**(29): 3794-3802.

Carnevalli, L. S., K. Masuda, et al. (2010). "S6K1 plays a critical role in early adipocyte differentiation." Dev Cell **18**: 763-774.

Carpenter, S., E. P. Ricci, et al. (2014). "Post-transcriptional regulation of gene expression in innate immunity." Nat Rev Immunol **14**(6): 361-376.

Castello, A., B. Fischer, et al. (2012). "Insights into RNA biology from an atlas of mammalian mRNA-binding proteins." Cell **149**(6): 1393-1406.

Castello, A., B. Fischer, et al. (2013). "RNA-binding proteins in Mendelian disease." Trends Genet **29**: 318-327.

Castets, M., C. Schaeffer, et al. (2005). "FMRP interferes with the Rac1 pathway and controls actin cytoskeleton dynamics in murine fibroblasts." Hum Mol Genet **14**(6): 835-844.

Cayrou, C., P. Coulombe, et al. (2012). "New insights into replication origin characteristics in metazoans." Cell Cycle **11**(4): 658-667.

Cayrou, C., P. Coulombe, et al. (2011). "Genome-scale analysis of metazoan replication origins reveals their organization in specific but flexible sites defined by conserved features." Genome Res **21**(9): 1438-1449.

Celik, A., S. Kervestin, et al. (2015). "NMD: At the crossroads between translation termination and ribosome recycling." Biochimie **114**: 2-9.

Chai, K. M., C. Y. Wang, et al. (2014). "Downregulation of BRCA1-BRCA2-containing complex subunit 3 sensitizes glioma cells to temozolomide." Oncotarget **5**(21): 10901-10915.

Chakraborty, P. and F. Grosse (2011). "Human DHX9 helicase preferentially unwinds RNA-containing displacement loops (R-loops) and G-quadruplexes." DNA Repair (Amst) **10**(6): 654-665.

Chalupnikova, K., S. Lattmann, et al. (2008). "Recruitment of the RNA helicase RHAU to stress granules via a unique RNA-binding domain." J Biol Chem **283**(50): 35186-35198.

Chambers, V. S., G. Marsico, et al. (2015). "High-throughput sequencing of DNA G-quadruplex structures in the human genome." Nat Biotechnol **33**(8): 877-881.

Chau, B. N., E. H. Cheng, et al. (2000). "Aven, a novel inhibitor of caspase activation, binds Bcl-xL and Apaf-1." Mol. Cell **6**: 31-40.

Chawla, G., C. H. Lin, et al. (2009). "Sam68 regulates a set of alternatively spliced exons during neurogenesis." Mol Cell Biol **29**: 201-213.

Chen, C., T. J. Nott, et al. (2011). "Deciphering arginine methylation: Tudor tells the tale." Nat Rev Mol Cell Biol **12**: 629-642.

Chen, E., M. R. Sharma, et al. (2014). "Fragile X mental retardation protein regulates translation by binding directly to the ribosome." Mol Cell **54**(3): 407-417.

Chen, M. C., P. Murat, et al. (2015). "Insights into the mechanism of a G-quadruplex-unwinding DEAH-box helicase." Nucleic Acids Res **43**(4): 2223-2231.

Chen, T., F. M. Boisvert, et al. (1999). "A role for the GSG domain in localizing Sam68 to novel nuclear structures in cancer cell lines." Mol Biol Cell **10**(9): 3015-3033.

Chen, X., L. Liu, et al. (2016). "Expression of Sam68 Associates with Neuronal Apoptosis and Reactive Astrocytes

After Spinal Cord Injury." Cell Mol Neurobiol.
Chen, Y. and G. Varani (2013). "Engineering RNA-binding proteins for biology." FEBS J **280**: 3734-3754.
Cheng, C. and P. A. Sharp (2006). "Regulation of CD44 alternative splicing by SRm160 and its potential role in tumor cell invasion." Mol Cell Biol **26**(1): 362-370.
Cheng, D., J. Cote, et al. (2007). "The arginine methyltransferase CARM1 regulates the coupling of transcription and mRNA processing." Mol Cell **25**(1): 71-83.
Cheung, N., L. C. Chan, et al. (2007). "Protein arginine-methyltransferase-dependent oncogenesis." Nat Cell Biol **9**(10): 1208-1215.
Cheung, N., T. K. Fung, et al. (2016). "Targeting Aberrant Epigenetic Networks Mediated by PRMT1 and KDM4C in Acute Myeloid Leukemia." Cancer Cell **29**(1): 32-48.
Chiang, C. and K. Ayyanathan (2013). "Snail/Gfi-1 (SNAG) family zinc finger proteins in transcription regulation, chromatin dynamics, cell signaling, development, and disease." Cytokine Growth Factor Rev **24**(2): 123-131.
Choi, J., Y. K. Hwang, et al. (2006). "Aven overexpression: association with poor prognosis in childhood acute lymphoblastic leukemia." Leuk Res **30**(8): 1019-1025.
Christmann, M. and B. Kaina (2013). "Transcriptional regulation of human DNA repair genes following genotoxic stress: trigger mechanisms, inducible responses and genotoxic adaptation." Nucleic Acids Res **41**(18): 8403-8420.
Ciccia, A. and S. J. Elledge (2010). "The DNA damage response: making it safe to play with knives." Mol Cell **40**(2): 179-204.
Cieply, B. and R. P. Carstens (2015). "Functional roles of alternative splicing factors in human disease." Wiley Interdiscip Rev RNA **6**(3): 311-326.
Cilenti, L., M. P. Balakrishnan, et al. (2011). "Regulation of Abro1/KIAA0157 during myocardial infarction and cell death reveals a novel cardioprotective mechanism for Lys63-specific deubiquitination." J Mol Cell Cardiol **50**(4): 652-661.
Clague, M. J., I. Barsukov, et al. (2013). "Deubiquitylases from genes to organism." Physiol Rev **93**(3): 1289-1315.
Clague, M. J., H. Liu, et al. (2012). "Governance of endocytic trafficking and signaling by reversible ubiquitylation." Dev Cell **23**(3): 457-467.
Close, P., P. East, et al. (2012). "DBIRD complex integrates alternative mRNA splicing with RNA polymerase II transcript elongation." Nature **484**(7394): 386-389.
Colak, D., N. Zaninovic, et al. (2014). "Promoter-bound trinucleotide repeat mRNA drives epigenetic silencing in fragile X syndrome." Science **343**(6174): 1002-1005.
Colgan, D. F. and J. L. Manley (1997). "Mechanism and regulation of mRNA polyadenylation." Genes Dev **11**(21): 2755-2766.
Cook-Andersen, H. and M. F. Wilkinson (2015). "Molecular biology: Splicing does the two-step." Nature **521**(7552): 300-301.
Cooper, E. M., J. D. Boeke, et al. (2010). "Specificity of the BRISC deubiquitinating enzyme is not due to selective binding to Lys63-linked polyubiquitin." J Biol Chem **285**(14): 10344-10352.
Cooper, T. A., L. Wan, et al. (2009). "RNA and disease." Cell **136**: 777-793.
Corbin, F., M. Bouillon, et al. (1997). "The fragile X mental retardation protein is associated with poly(A)+ mRNA in actively translating polyribosomes." Hum Mol Genet **6**(9): 1465-1472.
Corden, J. L. (1990). "Tails of RNA polymerase II." Trends Biochem Sci **15**(10): 383-387.
Corley, S. M. and J. E. Gready (2008). "Identification of the RGG box motif in Shadoo: RNA-binding and signaling

roles?" Bioinform Biol Insights **2**: 383-400.

Cote, J., F. M. Boisvert, et al. (2003). "Sam68 RNA-binding protein is an in vivo substrate for protein arginine N-methyltransferase 1." Mol Biol Cell **14**(1): 274-287.

Cote, J. and S. Richard (2005). "Tudor domains bind symmetrical dimethylated arginines." J Biol Chem **280**(31): 28476-28483.

Cox, J. and M. Mann (2008). "MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification." Nat Biotechnol **26**: 1367-1372.

Creacy, S. D., E. D. Routh, et al. (2008). "G4 resolvase 1 binds both DNA and RNA tetramolecular quadruplex with high affinity and is the major source of tetramolecular quadruplex G4-DNA and G4-RNA resolving activity in HeLa cell lysates." J Biol Chem **283**(50): 34626-34634.

Crees, Z., J. Girard, et al. (2014). "Oligonucleotides and G-quadruplex stabilizers: targeting telomeres and telomerase in cancer therapy." Curr Pharm Des **20**(41): 6422-6437.

Crenshaw, E., B. P. Leung, et al. (2015). "Amyloid Precursor Protein Translation Is Regulated by a 3'UTR Guanine Quadruplex." PLoS One **10**(11): e0143160.

Curtin, N. J. (2012). "DNA repair dysregulation from cancer driver to therapeutic target." Nat Rev Cancer **12**(12): 801-817.

Dai, C. Q., T. T. Luo, et al. (2016). "p53 and mitochondrial dysfunction: novel insight of neurodegenerative diseases." J Bioenerg Biomembr **48**(4): 337-347.

Dai, J., Z. Q. Liu, et al. (2015). "Discovery of Small Molecules for Up-Regulating the Translation of Antiamyloidogenic Secretase, a Disintegrin and Metalloproteinase 10 (ADAM10), by Binding to the G-Quadruplex-Forming Sequence in the 5' Untranslated Region (UTR) of Its mRNA." J Med Chem **58**(9): 3875-3891.

Dann, S. G., A. Selvaraj, et al. (2007). "mTOR Complex1-S6K1 signaling: at the crossroads of obesity, diabetes and cancer." Trends Mol Med **13**: 252-259.

Darnell, J. C., K. B. Jensen, et al. (2001). "Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function." Cell **107**(4): 489-499.

Darnell, J. C. and E. Klann (2013). "The translation of translational control by FMRP: therapeutic targets for FXS." Nat Neurosci **16**(11): 1530-1536.

Darnell, J. C., S. J. Van Driesche, et al. (2011). "FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism." Cell **146**(2): 247-261.

Das, S. and A. R. Krainer (2014). "Emerging functions of SRSF1, splicing factor and oncoprotein, in RNA metabolism and cancer." Mol Cancer Res **12**: 1195-1204.

Davidson, L., L. Muniz, et al. (2014). "3' end formation of pre-mRNA and phosphorylation of Ser2 on the RNA polymerase II CTD are reciprocally coupled in human cells." Genes Dev **28**(4): 342-356.

Debler, E. W., K. Jain, et al. (2016). "A glutamate/aspartate switch controls product specificity in a protein arginine methyltransferase." Proc Natl Acad Sci U S A **113**(8): 2068-2073.

Decorsiere, A., A. Cayrel, et al. (2011). "Essential role for the interaction between hnRNP H/F and a G quadruplex in maintaining p53 pre-mRNA 3'-end processing and function during DNA damage." Genes Dev **25**(3): 220-225.

Deigan, K. E., T. W. Li, et al. (2009). "Accurate SHAPE-directed RNA structure determination." Proc Natl Acad Sci U S A **106**(1): 97-102.

DeJesus-Hernandez, M., I. R. Mackenzie, et al. (2011). "Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS." Neuron **72**(2): 245-256.

Deng, L., C. Wang, et al. (2000). "Activation of the IkappaB kinase complex by TRAF6 requires a dimeric

ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain." Cell **103**(2): 351-361.
Derry, J. J., S. Richard, et al. (2000). "Sik (BRK) phosphorylates Sam68 in the nucleus and negatively regulates its RNA binding ability." Mol Cell Biol **20**(16): 6114-6126.
Devaiah, B. N., B. A. Lewis, et al. (2012). "BRD4 is an atypical kinase that phosphorylates serine2 of the RNA polymerase II carboxy-terminal domain." Proc Natl Acad Sci U S A **109**(18): 6927-6932.
Dhar, S., V. Vemulapalli, et al. (2013). "Loss of the major Type I arginine methyltransferase PRMT1 causes substrate scavenging by other PRMTs." Sci Rep **3**: 1311.
Di Lello, P. and S. G. Hymowitz (2016). "Unveiling the Structural and Dynamic Nature of the Ubiquitin Code." Structure **24**(4): 498-499.
Didiot, M. C., Z. Tian, et al. (2008). "The G-quartet containing FMRP binding site in FMR1 mRNA is a potent exonic splicing enhancer." Nucleic Acids Res **36**(15): 4902-4912.
Doma, M. K. and R. Parker (2006). "Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation." Nature **440**(7083): 561-564.
Dong, L., H. Che, et al. (2016). "Sam68 is Overexpressed in Epithelial Ovarian Cancer and Promotes Tumor Cell Proliferation." Med Sci Monit **22**: 3248-3256.
Douglas, A. G. and M. J. Wood (2011). "RNA splicing: disease and therapy." Brief Funct Genomics **10**(3): 151-164.
Droge, W. (2002). "Free radicals in the physiological control of cell function." Physiol Rev **82**(1): 47-95.
Egloff, S., M. Dienstbier, et al. (2012). "Updating the RNA polymerase CTD code: adding gene-specific layers." Trends Genet **28**(7): 333-341.
Eismann, M., I. M. Melzer, et al. (2013). "Overexpression of the anti-apoptotic protein AVEN contributes to increased malignancy in hematopoietic neoplasms." Oncogene **32**: 2586-2591.
El-Andaloussi, N., T. Valovka, et al. (2007). "Methylation of DNA polymerase beta by protein arginine methyltransferase 1 regulates its binding to proliferating cell nuclear antigen." FASEB J **21**(1): 26-34.
El-Andaloussi, N., T. Valovka, et al. (2006). "Arginine methylation regulates DNA polymerase beta." Mol Cell **22**(1): 51-62.
Elakoum, R., G. Gauchotte, et al. (2014). "CARM1 and PRMT1 are dysregulated in lung cancer without hierarchical features." Biochimie **97**: 210-218.
Endoh, T., Y. Kawasaki, et al. (2013). "Translational halt during elongation caused by G-quadruplex formed by mRNA." Methods **64**(1): 73-78.
Endoh, T. and N. Sugimoto (2016). "Mechanical insights into ribosomal progression overcoming RNA G-quadruplex from periodical translation suppression in cells." Sci Rep **6**: 22719.
Eperon, I. C., O. V. Makarova, et al. (2000). "Selection of alternative 5' splice sites: role of U1 snRNP and models for the antagonistic effects of SF2/ASF and hnRNP A1." Mol Cell Biol **20**(22): 8303-8318.
Erkelenz, S., W. F. Mueller, et al. (2013). "Position-dependent splicing activation and repression by SR and hnRNP proteins rely on common mechanisms." RNA **19**(1): 96-102.
Erson-Bensan, A. E. (2016). "Alternative polyadenylation and RNA-binding proteins." J Mol Endocrinol **57**(2): F29-34.
Erson-Bensan, A. E. and T. Can (2016). "Alternative Polyadenylation: Another Foe in Cancer." Mol Cancer Res **14**(6): 507-517.
Esmaili, A. M., E. L. Johnson, et al. (2010). "Regulation of the ATM-activator protein Aven by CRM1-dependent nuclear export." Cell Cycle **9**: 3913-3920.
Faudale, M., S. Cogoi, et al. (2012). "Photoactivated cationic alkyl-substituted porphyrin binding to g4-RNA in the 5'-UTR of KRAS oncogene represses translation." Chem Commun (Camb) **48**(6): 874-876.

Feng, L., J. Huang, et al. (2009). "MERIT40 facilitates BRCA1 localization and DNA damage repair." Genes Dev **23**(6): 719-728.

Feng, L., J. Wang, et al. (2010). "The Lys63-specific deubiquitinating enzyme BRCC36 is regulated by two scaffold proteins localizing in different subcellular compartments." J Biol Chem **285**(40): 30982-30988.

Feng, Y., R. Maity, et al. (2013). "Mammalian protein arginine methyltransferase 7 (PRMT7) specifically targets RXR sites in lysine- and arginine-rich regions." J Biol Chem **288**(52): 37010-37025.

Feracci, M., J. Foot, et al. (2014). "Structural investigations of the RNA-binding properties of STAR proteins." Biochem Soc Trans **42**(4): 1141-1146.

Feracci, M., J. N. Foot, et al. (2016). "Structural basis of RNA recognition and dimerization by the STAR proteins T-STAR and Sam68." Nat Commun **7**: 10355.

Filosa, G., S. M. Barabino, et al. (2013). "Proteomics strategies to identify SUMO targets and acceptor sites: a survey of RNA-binding proteins SUMOylation." Neuromolecular Med **15**(4): 661-676.

Finkel, T. and N. J. Holbrook (2000). "Oxidants, oxidative stress and the biology of ageing." Nature **408**(6809): 239-247.

Finnen, R. L., K. R. Pangka, et al. (2012). "Herpes simplex virus 2 infection impacts stress granule accumulation." J Virol **86**(15): 8119-8130.

Foulk, M. S., J. M. Urban, et al. (2015). "Characterizing and controlling intrinsic biases of lambda exonuclease in nascent strand sequencing reveals phasing between nucleosomes and G-quadruplex motifs around a subset of human replication origins." Genome Res **25**(5): 725-735.

Frees, S., C. Menendez, et al. (2014). "QGRS-Conserve: a computational method for discovering evolutionarily conserved G-quadruplex motifs." Hum Genomics **8**: 8.

Frisone, P., D. Pradella, et al. (2015). "SAM68: Signal Transduction and RNA Metabolism in Human Cancer." Biomed Res Int **2015**: 528954.

Fu, K., X. Sun, et al. (2016). "Sam68/KHDRBS1 is critical for colon tumorigenesis by regulating genotoxic stress-induced NF-kappaB activation." Elife **5**.

Fu, K., X. Sun, et al. (2013). "Sam68 modulates the promoter specificity of NF-kappaB and mediates expression of CD25 in activated T cells." Nat Commun **4**: 1909.

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

Fuhrmann, J. and P. R. Thompson (2016). "Protein Arginine Methylation and Citrullination in Epigenetic Regulation." ACS Chem Biol **11**(3): 654-668.

Fumagalli, S., N. F. Totty, et al. (1994). "A target for Src in mitosis." Nature **368**(6474): 871-874.

Galarneau, A. and S. Richard (2009). "The STAR RNA-binding proteins GLD-1, QKI, SAM68 and SLM-2 bind bipartite RNA motifs." BMC Mol Biol **10**: 47.

Gandin, V., G. J. Gutierrez, et al. (2013). "Degradation of newly synthesized polypeptides by ribosome-associated RACK1/c-Jun N-terminal kinase/eukaryotic elongation factor 1A2 complex." Mol Cell Biol **33**: 2510-2526.

Gandin, V., L. Masvidal, et al. (2016). "nanoCAGE reveals 5' UTR features that define specific modes of translation of functionally related MTOR-sensitive mRNAs." Genome Res **26**(5): 636-648.

Gandin, V., K. Sikström, et al. (2014). "Polysome fractionation and analysis of mammalian translatomes on a genome-wide scale." J Vis Exp **May 17 (87)**: doi: 10.3791/51455.

Gao, Y., Y. Zhao, et al. (2016). "The dual function of PRMT1 in modulating epithelial-mesenchymal transition and cellular senescence in breast cancer cells through regulation of ZEB1." Sci Rep **6**: 19874.

Gayatri, S. and M. T. Bedford (2014). "Readers of histone methylarginine marks." Biochim Biophys Acta **1839**(8):

702-710.

Gellert, M., M. N. Lipsett, et al. (1962). "Helix formation by guanylic acid." Proc Natl Acad Sci U S A **48**: 2013-2018.

Ghildiyal, R. and E. Sen (2015). "Concerted action of histone methyltransferases G9a and PRMT-1 regulates PGC-1alpha-RIG-I axis in IFNgamma treated glioma cells." Cytokine.

Ghosh, A., S. Shuman, et al. (2011). "Structural insights to how mammalian capping enzyme reads the CTD code." Mol Cell **43**(2): 299-310.

Giedroc, D. P. and P. V. Cornish (2009). "Frameshifting RNA pseudoknots: structure and mechanism." Virus Res **139**(2): 193-208.

Gkountela, S., Z. Li, et al. (2014). "PRMT5 is required for human embryonic stem cell proliferation but not pluripotency." Stem Cell Rev **10**(2): 230-239.

Glisovic, T., J. L. Bachorik, et al. (2008). "RNA-binding proteins and post-transcriptional gene regulation." FEBS Lett **582**: 1977-1986.

Gomez, D., A. Guedin, et al. (2010). "A G-quadruplex structure within the 5'-UTR of TRF2 mRNA represses translation in human cells." Nucleic Acids Res **38**(20): 7187-7198.

Gomez, D., T. Lemarteleur, et al. (2004). "Telomerase downregulation induced by the G-quadruplex ligand 12459 in A549 cells is mediated by hTERT RNA alternative splicing." Nucleic Acids Res **32**(1): 371-379.

Goulet, I., S. Boisvenue, et al. (2008). "TDRD3, a novel Tudor domain-containing protein, localizes to cytoplasmic stress granules." Human Molecular Genetics **17**(19): 3055-3074.

Goulet, I., G. Gauvin, et al. (2007). "Alternative splicing yields protein arginine methyltransferase 1 isoforms with distinct activity, substrate specificity, and subcellular localization." J Biol Chem **282**(45): 33009-33021.

Grant, C. M. (2011). "Regulation of translation by hydrogen peroxide." Antioxid Redox Signal **15**(1): 191-203.

Greenblatt, S. M., F. Liu, et al. (2016). "Arginine methyltransferases in normal and malignant hematopoiesis." Exp Hematol **44**(6): 435-441.

Grice, G. L. and J. A. Nathan (2016). "The recognition of ubiquitinated proteins by the proteasome." Cell Mol Life Sci **73**(18): 3497-3506.

Gros, J., A. Guedin, et al. (2008). "G-Quadruplex formation interferes with P1 helix formation in the RNA component of telomerase hTERC." Chembiochem **9**(13): 2075-2079.

Guccione, E., C. Bassi, et al. (2007). "Methylation of histone H3R2 by PRMT6 and H3K4 by an MLL complex are mutually exclusive." Nature **449**(7164): 933-937.

Guderian, G., C. Peter, et al. (2011). "RioK1, a new interactor of protein arginine methyltransferase 5 (PRMT5), competes with pICln for binding and modulates PRMT5 complex composition and substrate specificity." J Biol Chem **286**(3): 1976-1986.

Gui, S., S. Gathiaka, et al. (2014). "A remodeled protein arginine methyltransferase 1 (PRMT1) generates symmetric dimethylarginine." J Biol Chem **289**(13): 9320-9327.

Guo, J. Y., A. Yamada, et al. (2008). "Aven-Dependent Activation of ATM Following DNA Damage." Current Biology **18**(13): 933-942.

Habashy, H. O., E. A. Rakha, et al. (2013). "The oestrogen receptor coactivator CARM1 has an oncogenic effect and is associated with poor prognosis in breast cancer." Breast Cancer Res Treat **140**(2): 307-316.

Haeusler, A. R., C. J. Donnelly, et al. (2014). "C9orf72 nucleotide repeat structures initiate molecular cascades of disease." Nature **507**(7491): 195-200.

Hafner, M., M. Landthaler, et al. (2010). "Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP." Cell **141**(1): 129-141.

Hafner, M., M. Landthaler, et al. (2010). "PAR-CliP--a method to identify transcriptome-wide the binding sites of RNA-binding proteins." J Vis Exp **2**: 2034.

Hall, M. P., R. J. Nagel, et al. (2013). "Quaking and PTB control overlapping splicing regulatory networks during muscle cell differentiation." RNA **19**(5): 627-638.

Han, S., Q. Liu, et al. (2016). "Targeting the SH3 domain of human osteoclast-stimulating factor with rationally designed peptoid inhibitors." J Pept Sci **22**(8): 533-539.

Han, T. W., M. Kato, et al. (2012). "Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies." Cell **149**(4): 768-779.

Han, X., R. Li, et al. (2014). "Expression of PRMT5 correlates with malignant grade in gliomas and plays a pivotal role in tumor growth in vitro." J Neurooncol **118**(1): 61-72.

Hanakahi, L. A., H. Sun, et al. (1999). "High affinity interactions of nucleolin with G-G-paired rDNA." J Biol Chem **274**(22): 15908-15912.

Harding, H. P., Y. Zhang, et al. (2003). "An integrated stress response regulates amino acid metabolism and resistance to oxidative stress." Mol Cell **11**(3): 619-633.

Harigaya, Y. and R. Parker (2010). "No-go decay: a quality control mechanism for RNA in translation." Wiley Interdiscip Rev RNA **1**(1): 132-141.

Hartmann, A. M., O. Nayler, et al. (1999). "The interaction and colocalization of Sam68 with the splicing-associated factor YT521-B in nuclear dots is regulated by the Src family kinase p59(fyn)." Mol Biol Cell **10**(11): 3909-3926.

Hashimoto, M., K. Murata, et al. (2016). "Severe Hypomyelination and Developmental Defects Are Caused in Mice Lacking Protein Arginine Methyltransferase 1 (PRMT1) in the Central Nervous System." J Biol Chem **291**(5): 2237-2245.

Hata, K., N. Yanase, et al. (2016). "Differential regulation of T-cell dependent and T-cell independent antibody responses through arginine methyltransferase PRMT1 in vivo." FEBS Lett **590**(8): 1200-1210.

Hawley, R. G., Y. Chen, et al. (2012). "An Integrated Bioinformatics and Computational Biology Approach Identifies New BH3-Only Protein Candidates." Open Biol J. **5**: 6-16.

Healy, K. C. (1995). "Telomere dynamics and telomerase activation in tumor progression: prospects for prognosis and therapy." Oncol Res **7**(3-4): 121-130.

Heddi, B., V. V. Cheong, et al. (2015). "Insights into G-quadruplex specific recognition by the DEAH-box helicase RHAU: Solution structure of a peptide-quadruplex complex." Proc Natl Acad Sci U S A **112**(31): 9608-9613.

Hellen, C. U. and P. Sarnow (2001). "Internal ribosome entry sites in eukaryotic mRNA molecules." Genes Dev **15**(13): 1593-1612.

Herrero, E., J. Ros, et al. (2008). "Redox control and oxidative stress in yeast cells." Biochim Biophys Acta **1780**(11): 1217-1235.

Hershey, J. W., N. Sonenberg, et al. (2012). "Principles of translational control: an overview." Cold Spring Harb Perspect Biol **4**(12).

Hinnebusch, A. G. (2014). "The scanning mechanism of eukaryotic translation initiation." Annu Rev Biochem **83**: 779-812.

Hinnebusch, A. G., I. P. Ivanov, et al. (2016). "Translational control by 5'-untranslated regions of eukaryotic mRNAs." Science **352**(6292): 1413-1416.

Hocine, S., R. H. Singer, et al. (2010). "RNA processing and export." Cold Spring Harb Perspect Biol **2**(12): a000752.

Holcik, M. and N. Sonenberg (2005). "Translational control in stress and apoptosis." Nat Rev Mol Cell Biol **6**(4):

318-327.

Hong, W., R. J. Resnick, et al. (2002). "Physical and functional interaction between the transcriptional cofactor CBP and the KH domain protein Sam68." Mol Cancer Res **1**(1): 48-55.

Horsburgh, B. C., H. Kollmus, et al. (1996). "Translational recoding induced by G-rich mRNA sequences that form unusual structures." Cell **86**(6): 949-959.

Hou, F., L. Sun, et al. (2011). "MAVS forms functional prion-like aggregates to activate and propagate antiviral innate immune response." Cell **146**(3): 448-461.

Hsin, J. P., A. Sheth, et al. (2011). "RNAP II CTD phosphorylated on threonine-4 is required for histone mRNA 3' end processing." Science **334**(6056): 683-686.

Hu, D., M. Gur, et al. (2015). "Interplay between arginine methylation and ubiquitylation regulates KLF4-mediated genome stability and carcinogenesis." Nat Commun **6**: 8419.

Hu, R. and M. Hochstrasser (2016). "Recent progress in ubiquitin and ubiquitin-like protein (Ubl) signaling." Cell Res **26**(4): 389-390.

Huang, F., X. Zeng, et al. (2013). "Lysine 63-linked polyubiquitination is required for EGF receptor degradation." Proc Natl Acad Sci U S A **110**(39): 15722-15727.

Huang, T. T., S. M. Nijman, et al. (2006). "Regulation of monoubiquitinated PCNA by DUB autocleavage." Nat Cell Biol **8**(4): 339-347.

Huang, X. X., L. N. Zhu, et al. (2014). "Two cationic porphyrin isomers showing different multimeric G-quadruplex recognition specificity against monomeric G-quadruplexes." Nucleic Acids Res **42**(13): 8719-8731.

Huot, M. E., G. Vogel, et al. (2009). "Identification of a Sam68 ribonucleoprotein complex regulated by epidermal growth factor." J Biol Chem **284**(46): 31903-31913.

Huot, M. E., G. Vogel, et al. (2012). "The Sam68 STAR RNA-binding protein regulates mTOR alternative splicing during adipogenesis." Mol Cell **46**(2): 187-199.

Huppert, J. L. and S. Balasubramanian (2005). "Prevalence of quadruplexes in the human genome." Nucleic Acids Res **33**(9): 2908-2916.

Huppert, J. L., A. Bugaut, et al. (2008). "G-quadruplexes: the beginning and end of UTRs." Nucleic Acids Res **36**(19): 6260-6268.

Huppertz, I., J. Attig, et al. (2014). "iCLIP: protein-RNA interactions at nucleotide resolution." Methods **65**: 274-287.

Iijima, T., K. Wu, et al. (2011). "SAM68 regulates neuronal activity-dependent alternative splicing of neurexin-1." Cell **147**(7): 1601-1614.

Inoki, K., T. Zhu, et al. (2003). "TSC2 mediates cellular energy response to control cell growth and survival." Cell **115**: 577-590.

Irimia, M. and S. W. Roy (2014). "Origin of spliceosomal introns and alternative splicing." Cold Spring Harb Perspect Biol **6**(6).

Ivanov, A., T. Mikhailova, et al. (2016). "PABP enhances release factor recruitment and stop codon recognition during translation termination." Nucleic Acids Res **44**(16): 7766-7776.

Iwasaki, H. and T. Yada (2007). "Protein arginine methylation regulates insulin signaling in L6 skeletal muscle cells." Biochem Biophys Res Commun **364**(4): 1015-1021.

Jackson, R. J. (2013). "The current status of vertebrate cellular mRNA IRESs." Cold Spring Harb Perspect Biol **5**(2).

Jackson, R. J., C. U. Hellen, et al. (2010). "The mechanism of eukaryotic translation initiation and principles of its regulation." Nat Rev Mol Cell Biol **11**(2): 113-127.

Jain, K., R. A. Warmack, et al. (2016). "Protein Arginine Methyltransferase Product Specificity Is Mediated by Distinct Active-site Architectures." J Biol Chem **291**(35): 18299-18308.

Jamison, S. F., Z. Pasman, et al. (1995). "U1 snRNP-ASF/SF2 interaction and 5' splice site recognition: characterization of required elements." Nucleic Acids Res **23**(16): 3260-3267.

Jaschke, A., K. Hofer, et al. (2016). "Cap-like structures in bacterial RNA and epitranscriptomic modification." Curr Opin Microbiol **30**: 44-49.

Jelinic, P., J. C. Stehle, et al. (2006). "The testis-specific factor CTCFL cooperates with the protein methyltransferase PRMT7 in H19 imprinting control region methylation." PLoS Biol **4**(11): e355.

Jeong, H. J., H. J. Lee, et al. (2016). "Prmt7 Deficiency Causes Reduced Skeletal Muscle Oxidative Metabolism and Age-Related Obesity." Diabetes **65**(7): 1868-1882.

Jiao, X., J. H. Chang, et al. (2013). "A mammalian pre-mRNA 5' end capping quality control mechanism and an unexpected link of capping to pre-mRNA processing." Mol Cell **50**(1): 104-115.

Jodoin, R., L. Bauer, et al. (2014). "The folding of 5'-UTR human G-quadruplexes possessing a long central loop." RNA **20**(7): 1129-1141.

Kaida, D., M. G. Berg, et al. (2010). "U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation." Nature **468**(7324): 664-668.

Karkhanis, V., L. Wang, et al. (2012). "Protein arginine methyltransferase 7 regulates cellular response to DNA damage by methylating promoter histones H2A and H4 of the polymerase delta catalytic subunit gene, POLD1." J Biol Chem **287**(35): 29801-29814.

Karni, R., E. de Stanchina, et al. (2007). "The gene encoding the splicing factor SF2/ASF is a proto-oncogene." Nat Struct Mol Biol. **14**: 185-193.

Kato, M., T. W. Han, et al. (2012). "Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels." Cell **149**(4): 753-767.

Katsuda, Y., S. Sato, et al. (2016). "A Small Molecule That Represses Translation of G-Quadruplex-Containing mRNA." J Am Chem Soc **138**(29): 9037-9040.

Khoury, G. A., R. C. Baliban, et al. (2011). "Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database." Sci Rep **1**.

Kiledjian, M. and G. Dreyfuss (1992). "Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box." EMBO J **11**(7): 2655-2664.

Kim, H., J. Huang, et al. (2007). "CCDC98 is a BRCA1-BRCT domain-binding protein involved in the DNA damage response." Nat Struct Mol Biol **14**(8): 710-715.

Kim, J. D., K. E. Park, et al. (2015). "PRMT8 as a phospholipase regulates Purkinje cell dendritic arborization and motor coordination." Sci Adv **1**(11): e1500615.

Kim, S., U. Gunesdogan, et al. (2014). "PRMT5 protects genomic integrity during global DNA demethylation in primordial germ cells and preimplantation embryos." Mol Cell **56**(4): 564-579.

Kim, W., E. J. Bennett, et al. (2011). "Systematic and quantitative assessment of the ubiquitin-modified proteome." Mol Cell **44**(2): 325-340.

Klaunig, J. E. and L. M. Kamendulis (2004). "The role of oxidative stress in carcinogenesis." Annu Rev Pharmacol Toxicol **44**: 239-267.

Klein, M. E., T. J. Younts, et al. (2013). "RNA-binding protein Sam68 controls synapse number and local beta-actin mRNA metabolism in dendrites." Proc Natl Acad Sci U S A **110**(8): 3125-3130.

Koh, C. M., M. Bezzi, et al. (2015). "MYC regulates the core pre-mRNA splicing machinery as an essential step in lymphomagenesis." Nature **523**(7558): 96-100.

Komander, D. and M. Rape (2012). "The ubiquitin code." Annu Rev Biochem **81**: 203-229.

Komar, A. A. (2009). "A pause for thought along the co-translational folding pathway." Trends Biochem Sci **34**(1): 16-24.

Komar, A. A. and M. Hatzoglou (2011). "Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states." Cell Cycle **10**(2): 229-240.

Konig, J., K. Zarnack, et al. (2010). "iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution." Nat Struct Mol Biol **17**(7): 909-915.

Konig, S. L., A. C. Evans, et al. (2010). "Seven essential questions on G-quadruplexes." Biomol Concepts **1**(2): 197-213.

Kornblihtt, A. R., I. E. Schor, et al. (2013). "Alternative splicing: a pivotal step between eukaryotic transcription and translation." Nat Rev Mol Cell Biol **14**(3): 153-165.

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

Krüger, T., M. Hofweber, et al. (2013). "SCD6 induces ribonucleoprotein granule formation in trypanosomes in a translation-independent manner, regulated by its Lsm and RGG domains." Mol Biol Cell **24**: 2098-2111.

Krivtsov, A. V. and S. A. Armstrong (2007). "MLL translocations, histone modifications and leukaemia stem-cell development." Nat Rev Cancer **7**: 823-833.

Kryukov, G. V., F. H. Wilson, et al. (2016). "MTAP deletion confers enhanced dependency on the PRMT5 arginine methyltransferase in cancer cells." Science **351**(6278): 1214-1218.

Kumari, S., A. Bugaut, et al. (2007). "An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation." Nat Chem Biol **3**(4): 218-221.

Kunkel, G. T. and X. Wang (2011). "Sam68 guest STARs in TNF-alpha signaling." Mol Cell **43**(2): 157-158.

Kutuk, O., S. G. Temel, et al. (2010). "Aven blocks DNA damage-induced apoptosis by stabilising Bcl-xL." Eur J Cancer **46**(13): 2494-2505.

Kwok, C. K., G. Marsico, et al. (2016). "rG4-seq reveals widespread formation of G-quadruplex structures in the human transcriptome." Nat Methods.

Kwon, S. C., H. Yi, et al. (2013). "The RNA-binding protein repertoire of embryonic stem cells." Nat Struct Mol Biol **20**: 1122-1130.

Lacroix, M., S. El Messaoudi, et al. (2008). "The histone-binding protein COPR5 is required for nuclear functions of the protein arginine methyltransferase PRMT5." EMBO Rep **9**(5): 452-458.

Lai, J. C., S. Ponti, et al. (2012). "The DEAH-box helicase RHAU is an essential gene and critical for mouse hematopoiesis." Blood **119**: 4291-4300.

Lam, E. Y., D. Beraldi, et al. (2013). "G-quadruplex structures are stable and detectable in human genomic DNA." Nat Commun **4**: 1796.

Lammich, S., F. Kamp, et al. (2011). "Translational repression of the disintegrin and metalloprotease ADAM10 by a stable G-quadruplex secondary structure in its 5'-untranslated region." J Biol Chem **286**(52): 45063-45072.

Langereis, M. A., Q. Feng, et al. (2013). "MDA5 localizes to stress granules, but this localization is not required for the induction of type I interferon." J Virol **87**(11): 6314-6325.

Laplante, M. and D. M. Sabatini (2012). "mTOR signaling in growth control and disease." Cell **149**: 274-293.

Larsen, S. C., K. B. Sylvestersen, et al. (2016). "Proteome-wide analysis of arginine monomethylation reveals widespread occurrence in human cells." Sci Signal **9**(443): rs9.

Lattmann, S., B. Giri, et al. (2010). "Role of the amino terminal RHAU-specific motif in the recognition and

resolution of guanine quadruplex-RNA by the DEAH-box RNA helicase RHAU." Nucleic Acids Res **38**(18): 6219-6233.

Lattmann, S., M. B. Stadler, et al. (2011). "The DEAH-box RNA helicase RHAU binds an intramolecular RNA G-quadruplex in TERC and associates with telomerase holoenzyme." Nucleic Acids Res **39**(21): 9390-9404.

Le Romancer, M., I. Treilleux, et al. (2008). "Regulation of estrogen rapid signaling through arginine methylation by PRMT1." Mol Cell **31**(2): 212-221.

Lee, J., J. Sayegh, et al. (2005). "PRMT8, a new membrane-bound tissue-specific member of the protein arginine methyltransferase family." J Biol Chem **280**(38): 32890-32896.

Lee, J. G., K. Baek, et al. (2013). "Reversible inactivation of deubiquitinases by reactive oxygen species in vitro and in cells." Nat Commun **4**: 1568.

Lee, Y. J., W. Y. Hsieh, et al. (2012). "Protein arginine methylation of SERBP1 by protein arginine methyltransferase 1 affects cytoplasmic/nuclear distribution." J Cell Biochem **113**(8): 2721-2728.

Lee, Y. J., H. M. Wei, et al. (2014). "Localization of SERBP1 in stress granules and nucleoli." FEBS J **281**(1): 352-364.

Lee, Y. Y., R. C. Cevallos, et al. (2009). "An upstream open reading frame regulates translation of GADD34 during cellular stresses that induce eIF2alpha phosphorylation." J Biol Chem **284**(11): 6661-6673.

Lefebvre, S., L. Burglen, et al. (1995). "Identification and characterization of a spinal muscular atrophy-determining gene." Cell **80**(1): 155-165.

Leprivier, G., B. Rotblat, et al. (2015). "Stress-mediated translational control in cancer cells." Biochim Biophys Acta **1849**(7): 845-860.

Li, B., L. Liu, et al. (2015). "miR-503 suppresses metastasis of hepatocellular carcinoma cell by targeting PRMT1." Biochem Biophys Res Commun **464**(4): 982-987.

Li, N. and S. Richard (2016). "Sam68 functions as a transcriptional coactivator of the p53 tumor suppressor." Nucleic Acids Res.

Li, Y., R. Zhu, et al. (2015). "Arginine Methyltransferase 1 in the Nucleus Accumbens Regulates Behavioral Effects of Cocaine." J Neurosci **35**(37): 12890-12902.

Li, Z., C. P. Yu, et al. (2012). "Sam68 expression and cytoplasmic localization is correlated with lymph node metastasis as well as prognosis in patients with early-stage cervical cancer." Ann Oncol **23**(3): 638-646.

Liang, Q., T. S. Dexheimer, et al. (2014). "A selective USP1-UAF1 inhibitor links deubiquitination to DNA damage responses." Nat Chem Biol **10**(4): 298-304.

Liao, H. W., J. M. Hsu, et al. (2015). "PRMT1-mediated methylation of the EGF receptor regulates signaling and cetuximab response." J Clin Invest **125**(12): 4529-4543.

Liao, W. T., J. L. Liu, et al. (2013). "High expression level and nuclear localization of Sam68 are associated with progression and poor prognosis in colorectal cancer." BMC Gastroenterol **13**: 126.

Liedtke, M. and M. L. Cleary (2009). "Therapeutic targeting of MLL." Blood **113**: 6061-6068.

Lin, Q., S. J. Taylor, et al. (1997). "Specificity and determinants of Sam68 RNA binding. Implications for the biological function of K homology domains." J Biol Chem **272**(43): 27274-27280.

Lipps, H. J., W. Gruissem, et al. (1982). "Higher order DNA structure in macronuclear chromatin of the hypotrichous ciliate Oxytricha nova." Proc Natl Acad Sci U S A **79**(8): 2495-2499.

Lischwe, M. A., R. G. Cook, et al. (1985). "Clustering of glycine and NG,NG-dimethylarginine in nucleolar protein C23." Biochemistry **24**(22): 6025-6028.

Liu, F., G. Cheng, et al. (2015). "Arginine methyltransferase PRMT5 is essential for sustaining normal adult hematopoiesis." J Clin Invest **125**(9): 3532-3544.

Locatelli, A., K. A. Lofgren, et al. (2012). "Mechanisms of HGF/Met signaling to Brk and Sam68 in breast cancer progression." Horm Cancer **3**(1-2): 14-25.

Lopez de Silanes, I., S. Galban, et al. (2005). "Identification and functional outcome of mRNAs associated with RNA-binding protein TIA-1." Mol Cell Biol **25**(21): 9520-9531.

Lorenz, R., S. H. Bernhart, et al. (2013). "2D meets 4G: G-quadruplexes in RNA secondary structure prediction." IEEE/ACM Trans Comput Biol Bioinform. **10**: 832-844.

Lu, R., H. Wang, et al. (2004). "The fragile X protein controls microtubule-associated protein 1B translation and microtubule stability in brain neuron development." Proc Natl Acad Sci U S A **101**(42): 15201-15206.

Lu, Y., B. H. Lee, et al. (2015). "Substrate degradation by the proteasome: a single-molecule kinetic analysis." Science **348**(6231): 1250834.

Lukong, K. E., K. W. Chang, et al. (2008). "RNA-binding proteins in human genetic disease." Trends Genet **24**(8): 416-425.

Lukong, K. E., D. Larocque, et al. (2005). "Tyrosine phosphorylation of sam68 by breast tumor kinase regulates intranuclear localization and cell cycle progression." J Biol Chem **280**(46): 38639-38647.

Lukong, K. E. and S. Richard (2003). "Sam68, the KH domain-containing superSTAR." Biochim Biophys Acta **1653**(2): 73-86.

Lukong, K. E. and S. Richard (2008). "Motor coordination defects in mice deficient for the Sam68 RNA-binding protein." Behav Brain Res **189**(2): 357-363.

Lv, L., H. Chen, et al. (2015). "PRMT1 promotes glucose toxicity-induced beta cell dysfunction by regulating the nucleo-cytoplasmic trafficking of PDX-1 in a FOXO1-dependent manner in INS-1 cells." Endocrine **49**(3): 669-682.

Maizels, N. and L. T. Gray (2013). "The G4 genome." PLoS Genet **9**(4): e1003468.

Mali, P., L. Yang, et al. (2013). "RNA-guided human genome engineering via Cas9." Science **339**: 823-826.

Mannen, T., S. Yamashita, et al. (2016). "The Sam68 nuclear body is composed of two RNase-sensitive substructures joined by the adaptor HNRNPL." J Cell Biol **214**(1): 45-59.

Mantri, M., T. Krojer, et al. (2010). "Crystal structure of the 2-oxoglutarate- and Fe(II)-dependent lysyl hydroxylase JMJD6." J Mol Biol **401**(2): 211-222.

Marcel, V., P. L. Tran, et al. (2011). "G-quadruplex structures in TP53 intron 3: role in alternative splicing and in production of p53 mRNA isoforms." Carcinogenesis **32**(3): 271-278.

Marjon, K., M. J. Cameron, et al. (2016). "MTAP Deletions in Cancer Create Vulnerability to Targeting of the MAT2A/PRMT5/RIOK1 Axis." Cell Rep **15**(3): 574-587.

Marquez, Y., J. W. Brown, et al. (2012). "Transcriptome survey reveals increased complexity of the alternative splicing landscape in Arabidopsis." Genome Res **22**(6): 1184-1195.

Martadinata, H. and A. T. Phan (2013). "Structure of human telomeric RNA (TERRA): stacking of two G-quadruplex blocks in K(+) solution." Biochemistry **52**(13): 2176-2183.

Martadinata, H. and A. T. Phan (2014). "Formation of a stacked dimeric G-quadruplex containing bulges by the 5'-terminal region of human telomerase RNA (hTERC)." Biochemistry **53**(10): 1595-1600.

Martinez-Rucobo, F. W., R. Kohler, et al. (2015). "Molecular Basis of Transcription-Coupled Pre-mRNA Capping." Mol Cell **58**(6): 1079-1089.

Masamha, C. P., Z. Xia, et al. (2014). "CFIm25 links alternative polyadenylation to glioblastoma tumour suppression." Nature **510**(7505): 412-416.

Matter, N., P. Herrlich, et al. (2002). "Signal-dependent regulation of splicing via phosphorylation of Sam68." Nature **420**(6916): 691-695.

Mavrakis, K. J., E. R. McDonald, 3rd, et al. (2016). "Disordered methionine metabolism in

MTAP/CDKN2A-deleted cancers leads to dependence on PRMT5." Science **351**(6278): 1208-1213.
McDowell, G. S. and A. Philpott (2013). "Non-canonical ubiquitylation: mechanisms and consequences." Int J Biochem Cell Biol **45**(8): 1833-1842.
McMahon, A. C., R. Rahman, et al. (2016). "TRIBE: Hijacking an RNA-Editing Enzyme to Identify Cell-Specific Targets of RNA-Binding Proteins." Cell **165**(3): 742-753.
Medicherla, B. and A. L. Goldberg (2008). "Heat shock and oxygen radicals stimulate ubiquitin-dependent degradation mainly of newly synthesized proteins." J Cell Biol **182**(4): 663-673.
Meister, G., C. Eggert, et al. (2001). "Methylation of Sm proteins by a complex containing PRMT5 and the putative U snRNP assembly factor pICln." Curr Biol **11**(24): 1990-1994.
Mele, M., P. G. Ferreira, et al. (2015). "Human genomics. The human transcriptome across tissues and individuals." Science **348**(6235): 660-665.
Melzer, I. M., S. B. Fernández, et al. (2012). "The Apaf-1-binding protein Aven is cleaved by Cathepsin D to unleash its anti-apoptotic potential." Cell Death Differ **19**: 1435-1445.
Mendoza, O., A. Bourdoncle, et al. (2016). "G-quadruplexes and helicases." Nucleic Acids Res **44**(5): 1989-2006.
Metzger, M. B., V. A. Hristova, et al. (2012). "HECT and RING finger families of E3 ubiquitin ligases at a glance." J Cell Sci **125**(Pt 3): 531-537.
Meyer, H. J. and M. Rape (2014). "Enhanced protein degradation by branched ubiquitin chains." Cell **157**(4): 910-921.
Meyer, N. H., K. Tripsianes, et al. (2010). "Structural basis for homodimerization of the Src-associated during mitosis, 68-kDa protein (Sam68) Qua1 domain." J Biol Chem **285**(37): 28893-28901.
Millevoi, S., H. Moine, et al. (2012). "G-quadruplexes in RNA biology." Wiley Interdiscip Rev RNA **3**(4): 495-507.
Millevoi, S. and S. Vagner (2010). "Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation." Nucleic Acids Res **38**(9): 2757-2774.
Min, M. and C. Lindon (2012). "Substrate targeting by the ubiquitin-proteasome system in mitosis." Semin Cell Dev Biol **23**(5): 482-491.
Miranda, T. B., M. Miranda, et al. (2004). "PRMT7 is a member of the protein arginine methyltransferase family with a distinct substrate specificity." J Biol Chem **279**(22): 22902-22907.
Misra, A. and M. R. Green (2016). "From polyadenylation to splicing: Dual role for mRNA 3' end formation factors." RNA Biol **13**(3): 259-264.
Mitchell, S. F. and R. Parker (2014). "Principles and properties of eukaryotic mRNPs." Mol Cell **54**: 547-558.
Modelska, A., E. Turro, et al. (2015). "The malignant phenotype in breast cancer is driven by eIF4A1-mediated changes in the translational landscape." Cell Death Dis **6**: e1603.
Moore, M. J., C. Zhang, et al. (2014). "Mapping Argonaute and conventional RNA-binding protein interactions with RNA at single-nucleotide resolution using HITS-CLIP and CIMS analysis." Nat Protoc **9**(2): 263-293.
Morales, Y., D. V. Nitzel, et al. (2015). "Redox Control of Protein Arginine Methyltransferase 1 (PRMT1) Activity." J Biol Chem **290**(24): 14915-14926.
Morettin, A., R. M. Baldwin, et al. (2015). "Arginine methyltransferases as novel therapeutic targets for breast cancer." Mutagenesis **30**(2): 177-189.
Morgan, R. K., H. Batra, et al. (2016). "Identification and characterization of a new G-quadruplex forming region within the kRAS promoter as a transcriptional regulator." Biochim Biophys Acta **1859**(2): 235-245.
Morris, M. J. and S. Basu (2009). "An unusually stable G-quadruplex within the 5'-UTR of the MT3 matrix metalloproteinase mRNA represses translation in eukaryotic cells." Biochemistry **48**(23): 5313-5319.
Morris, M. J., Y. Negishi, et al. (2010). "An RNA G-quadruplex is essential for cap-independent translation initiation in human VEGF IRES." J Am Chem Soc **132**(50): 17831-17839.

Morris, M. J., K. L. Wingate, et al. (2012). "The porphyrin TmPyP4 unfolds the extremely stable G-quadruplex in MT3-MMP mRNA and alleviates its repressive effect to enhance translation in eukaryotic cells." Nucleic Acids Res **40**(9): 4137-4145.

Moye, A. L., K. C. Porter, et al. (2015). "Telomeric G-quadruplexes are a substrate and site of localization for human telomerase." Nat Commun **6**: 7643.

Muhlemann, O. (2016). "Spermatogenesis Studies Reveal a Distinct Nonsense-Mediated mRNA Decay (NMD) Mechanism for mRNAs with Long 3'UTRs." PLoS Genet **12**(5): e1005979.

Mukundan, V. T. and A. T. Phan (2013). "Bulges in G-quadruplexes: broadening the definition of G-quadruplex-forming sequences." J Am Chem Soc **135**(13): 5017-5028.

Muller-McNicoll, M. and K. M. Neugebauer (2013). "How cells get the message: dynamic assembly and function of mRNA-protein complexes." Nat Rev Genet **14**(4): 275-287.

Munding, E. M., L. Shiue, et al. (2013). "Competition between pre-mRNAs for the splicing machinery drives global regulation of splicing." Mol Cell **51**(3): 338-348.

Muntean, A. G. and J. L. Hess (2012). "The pathogenesis of mixed-lineage leukemia." Annu Rev Pathol **7**: 283-301.

Murat, P. and S. Balasubramanian (2014). "Existence and consequences of G-quadruplex structures in DNA." Curr Opin Genet Dev **25**: 22-29.

Murat, P., J. Zhong, et al. (2014). "G-quadruplexes regulate Epstein-Barr virus-encoded nuclear antigen 1 mRNA translation." Nat Chem Biol **10**(5): 358-364.

Nachbur, U. and J. Silke (2016). "Killing Lymphoma with Smac-Mimetics: As Easy as ABC?" Cancer Cell **29**(4): 425-427.

Nakka, K. K., N. Chaudhary, et al. (2015). "Nuclear matrix-associated protein SMAR1 regulates alternative splicing via HDAC6-mediated deacetylation of Sam68." Proc Natl Acad Sci U S A **112**(26): E3374-3383.

Napoli, I., V. Mercaldo, et al. (2008). "The fragile X syndrome protein represses activity-dependent translation through CYFIP1, a new 4E-BP." Cell **134**(6): 1042-1054.

Naro, C. and C. Sette (2013). "Phosphorylation-mediated regulation of alternative splicing in cancer." Int J Cell Biol **2013**: 151839.

Natarajan, A. T. and F. Palitti (2008). "DNA repair and chromosomal alterations." Mutat Res **657**(1): 3-7.

Neault, M., F. A. Mallette, et al. (2012). "Ablation of PRMT6 reveals a role as a negative transcriptional regulator of the p53 tumor suppressor." Nucleic Acids Res **40**(19): 9513-9521.

Neuenkirchen, N., C. Englbrecht, et al. (2015). "Reconstitution of the human U snRNP assembly machinery reveals stepwise Sm protein organization." EMBO J **34**(14): 1925-1941.

Ng, H. M., L. Wei, et al. (2016). "The K63-Deubiquitylating Enzyme BRCC36 Limits DNA Break Processing and Repair." J Biol Chem.

Nicholas, C., J. Yang, et al. (2013). "PRMT5 is upregulated in malignant and metastatic melanoma and regulates expression of MITF and p27(Kip1.)." PLoS One **8**(9): e74710.

O'Brien, E. P., M. Vendruscolo, et al. (2012). "Prediction of variable translation rate effects on cotranslational protein folding." Nat Commun **3**: 868.

Okamura, M., H. Inose, et al. (2015). "RNA Export through the NPC in Eukaryotes." Genes (Basel) **6**(1): 124-149.

Onel, B., M. Carver, et al. (2016). "A New G-Quadruplex with Hairpin Loop Immediately Upstream of the Human BCL2 P1 Promoter Modulates Transcription." J Am Chem Soc **138**(8): 2563-2570.

Oshiumi, H., M. Miyashita, et al. (2013). "A distinct role of Riplet-mediated K63-Linked polyubiquitination of the RIG-I repressor domain in human antiviral innate immune responses." PLoS Pathog **9**(8): e1003533.

Pagliarini, V., L. Pelosi, et al. (2015). "SAM68 is a physiological regulator of SMN2 splicing in spinal muscular atrophy." J Cell Biol **211**(1): 77-90.

Pahlich, S., K. Bschir, et al. (2005). "Different methylation characteristics of protein arginine methyltransferase 1 and 3 toward the Ewing Sarcoma protein and a peptide." Proteins **61**(1): 164-175.

Pal, A., M. A. Young, et al. (2014). "Emerging potential of therapeutic targeting of ubiquitin-specific proteases in the treatment of cancer." Cancer Res **74**(18): 4955-4966.

Pal, S., S. N. Vishwanath, et al. (2004). "Human SWI/SNF-associated PRMT5 methylates histone H3 arginine 8 and negatively regulates expression of ST7 and NM23 tumor suppressor genes." Mol Cell Biol **24**(21): 9630-9645.

Pandit, S., Y. Zhou, et al. (2013). "Genome-wide analysis reveals SR protein cooperation and competition in regulated splicing." Mol Cell **50**(2): 223-235.

Panier, S. and D. Durocher (2013). "Push back to respond better: regulatory inhibition of the DNA double-strand break response." Nat Rev Mol Cell Biol **14**(10): 661-672.

Park, J. E., H. Yi, et al. (2016). "Regulation of Poly(A) Tail and Translation during the Somatic Cell Cycle." Mol Cell **62**(3): 462-471.

Park, J. H., M. Szemes, et al. (2015). "Protein arginine methyltransferase 5 is a key regulator of the MYCN oncoprotein in neuroblastoma cells." Mol Oncol **9**(3): 617-627.

Paronetto, M. P., T. Achsel, et al. (2007). "The RNA-binding protein Sam68 modulates the alternative splicing of Bcl-x." J Cell Biol **176**(7): 929-939.

Paronetto, M. P., M. Cappellari, et al. (2010). "Alternative splicing of the cyclin D1 proto-oncogene is regulated by the RNA-binding protein Sam68." Cancer Res **70**(1): 229-239.

Paronetto, M. P., V. Messina, et al. (2011). "Sam68 marks the transcriptionally active stages of spermatogenesis and modulates alternative splicing in male germ cells." Nucleic Acids Res **39**(12): 4961-4974.

Paronetto, M. P., V. Messina, et al. (2009). "Sam68 regulates translation of target mRNAs in male germ cells, necessary for mouse spermatogenesis." J Cell Biol **185**: 235-249.

Pawlak, M. R., C. A. Scherer, et al. (2000). "Arginine N-methyltransferase 1 is required for early postimplantation mouse development, but cells deficient in the enzyme are viable." Mol Cell Biol **20**(13): 4859-4869.

Paydas, S., K. Tanriverdi, et al. (2003). "Survivin and aven: two distinct antiapoptotic signals in acute leukemias." Ann Oncol **14**(7): 1045-1050.

Pedrotti, S., P. Bielli, et al. (2010). "The splicing regulator Sam68 binds to a novel exonic splicing silencer and functions in SMN2 alternative splicing in spinal muscular atrophy." EMBO J **29**(7): 1235-1247.

Pedrotti, S. and C. Sette (2010). "Spinal muscular atrophy: a new player joins the battle for SMN2 exon 7 splicing." Cell Cycle **9**(19): 3874-3879.

Pek, J. W., A. Anand, et al. (2012). "Tudor domain proteins in development." Development **139**(13): 2255-2266.

Pelletier, J. and N. Sonenberg (1985). "Insertion mutagenesis to increase secondary structure within the 5' noncoding region of a eukaryotic mRNA reduces translational efficiency." Cell **40**(3): 515-526.

Perez-Perez, A., F. Sanchez-Jimenez, et al. (2016). "Sam68 Mediates the Activation of Insulin and Leptin Signalling in Breast Cancer Cells." PLoS One **11**(7): e0158218.

Pfoh, R., I. K. Lacdao, et al. (2015). "Deubiquitinases and the new therapeutic opportunities offered to cancer." Endocr Relat Cancer **22**(1): T35-54.

Phan, A. T., V. Kuryavyi, et al. (2011). "Structure-function studies of FMRP RGG peptide recognition of an RNA duplex-quadruplex junction." Nat Struct Mol Biol **18**(7): 796-804.

Piques, M., W. X. Schulze, et al. (2009). "Ribosome and transcript copy numbers, polysome occupancy and

enzyme dynamics in Arabidopsis." Mol Syst Biol **5**: 314.

Polo, S. E. and S. P. Jackson (2011). "Dynamics of DNA damage response proteins at DNA breaks: a focus on protein modifications." Genes Dev **25**(5): 409-433.

Poornima, G., S. Shah, et al. (2016). "Arginine methylation promotes translation repression activity of eIF4G-binding protein, Scd6." Nucleic Acids Res.

Powley, I. R., A. Kondrashov, et al. (2009). "Translational reprogramming following UVB irradiation is mediated by DNA-PKcs and allows selective recruitment to the polysomes of mRNAs encoding DNA repair enzymes." Genes Dev **23**(10): 1207-1220.

Prochazka, L., R. Tesarik, et al. (2014). "Regulation of alternative splicing of CD44 in cancer." Cell Signal **26**(10): 2234-2239.

Proudfoot, N. (2004). "New perspectives on connecting messenger RNA 3' end formation to transcription." Curr Opin Cell Biol **16**(3): 272-278.

Proudfoot, N. J. (2011). "Ending the message: poly(A) signals then and now." Genes Dev **25**(17): 1770-1782.

Rajan, P., C. Dalgliesh, et al. (2009). "Proteomic identification of heterogeneous nuclear ribonucleoprotein L as a novel component of SLM/Sam68 Nuclear Bodies." BMC Cell Biol **10**: 82.

Rajan, P., L. Gaughan, et al. (2008). "The RNA-binding and adaptor protein Sam68 modulates signal-dependent splicing and transcriptional activity of the androgen receptor." J Pathol **215**(1): 67-77.

Rajyaguru, P. and R. Parker (2012). "RGG motif proteins: Modulators of mRNA functional states." Cell Cycle **11**(14): 2594-2599.

Rajyaguru, P., M. She, et al. (2012). "Scd6 Targets eIF4G to Repress Translation: RGG Motif Proteins as a Class of eIF4G-Binding Proteins." Mol Cell **45**(2): 244-254.

Ramaswami, M., J. P. Taylor, et al. (2013). "Altered ribostasis: RNA-protein granules in degenerative disorders." Cell **154**: 727-736.

Ray, D., H. Kazan, et al. (2013). "A compendium of RNA-binding motifs for decoding gene regulation." Nature **499**(7457): 172-177.

Reddy, K., B. Zamiri, et al. (2013). "The disease-associated r(GGGGCC)n repeat from the C9orf72 gene forms tract length-dependent uni- and multimolecular RNA G-quadruplex structures." J Biol Chem **288**(14): 9860-9866.

Reintjes, A., J. E. Fuchs, et al. (2016). "Asymmetric arginine dimethylation of RelA provides a repressive mark to modulate TNFalpha/NF-kappaB response." Proc Natl Acad Sci U S A **113**(16): 4326-4331.

Rhodes, D. and H. J. Lipps (2015). "G-quadruplexes and their regulatory roles in biology." Nucleic Acids Res **43**(18): 8627-8637.

Richard, S. (2010). "Reaching for the stars: Linking RNA-binding proteins to diseases." Adv Exp Med Biol **693**: 142-157.

Richard, S., N. Torabi, et al. (2005). "Ablation of the Sam68 RNA-binding protein protects mice from age-related bone loss." PLoS Genet **1**(6): e74.

Richard, S., G. Vogel, et al. (2008). "Sam68 haploinsufficiency delays onset of mammary tumorigenesis and metastasis." Oncogene **27**(4): 548-556.

Richter, J. D., G. J. Bassell, et al. (2015). "Dysregulation and restoration of translational homeostasis in fragile X syndrome." Nat Rev Neurosci **16**(10): 595-605.

Rodriguez, R. and K. M. Miller (2014). "Unravelling the genomic targets of small molecules using high-throughput sequencing." Nat Rev Genet **15**(12): 783-796.

Rosner, M. and M. Hengstschläger (2011). "Nucleocytoplasmic localization of p70 S6K1, but not of its isoforms p85 and p31, is regulated by TSC2/mTOR." Oncogene **30**: 4509-4522.

Rouleau, S. G., J. D. Beaudoin, et al. (2015). "Small antisense oligonucleotides against G-quadruplexes: specific mRNA translational switches." Nucleic Acids Res **43**(1): 595-606.

Rouskin, S., M. Zubradt, et al. (2014). "Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo." Nature **505**: 701-705.

Rubio, C. A., B. Weisburd, et al. (2014). "Transcriptome-wide characterization of the eIF4A signature highlights plasticity in translation regulation." Genome Biol **15**(10): 476.

Saito, K., W. Horikawa, et al. (2015). "Inhibiting K63 polyubiquitination abolishes no-go type stalled translation surveillance in Saccharomyces cerevisiae." PLoS Genet **11**(4): e1005197.

Sanchez-Jimenez, F. and V. Sanchez-Margalet (2013). "Role of Sam68 in post-transcriptional gene regulation." Int J Mol Sci **14**(12): 23402-23419.

Sanchez-Margalet, V., C. Gonzalez-Yanes, et al. (2003). "The expression of Sam68, a protein involved in insulin signal transduction, is enhanced by insulin stimulation." Cell Mol Life Sci **60**(4): 751-758.

Sanchez, G., A. Y. Dury, et al. (2013). "A novel function for the survival motoneuron protein as a translational regulator." Hum Mol Genet **22**: 668-684.

Saucedo, L. J., X. Gao, et al. (2003). "Rheb promotes cell growth as a component of the insulin/TOR signalling network." Nat Cell Biol **5**: 566-571.

Sayegh, J., K. Webb, et al. (2007). "Regulation of protein arginine methyltransferase 8 (PRMT8) activity by its N-terminal domain." J Biol Chem **282**(50): 36444-36453.

Schaeffer, C., B. Bardoni, et al. (2001). "The fragile X mental retardation protein binds specifically to its mRNA via a purine quartet motif." EMBO J **20**(17): 4803-4813.

Schalm, S. S. and J. Blenis (2002). "Identification of a conserved motif required for mTOR signaling." Curr Biol **12**: 632-639.

Schatz, J. H., E. Oricchio, et al. (2011). "Targeting cap-dependent translation blocks converging survival signals by AKT and PIM kinases in lymphoma." J Exp Med **208**(9): 1799-1807.

Schwartz, J. C., X. Wang, et al. (2013). "RNA seeds higher-order assembly of FUS protein." Cell Rep **5**: 918-925.

Selenko, P., R. Sprangers, et al. (2001). "SMN tudor domain structure and its interaction with the Sm proteins." Nat Struct Biol **8**(1): 27-31.

Sellier, C., F. Rau, et al. (2010). "Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients." EMBO J **29**(7): 1248-1261.

Sexton, A. N. and K. Collins (2011). "The 5' guanosine tracts of human telomerase RNA are recognized by the G-quadruplex binding domain of the RNA helicase DHX36 and function to increase RNA accumulation." Mol Cell Biol **31**(4): 736-743.

Shahid, R., A. Bugaut, et al. (2010). "The BCL-2 5' untranslated region contains an RNA G-quadruplex-forming motif that modulates protein expression." Biochemistry **49**(38): 8300-8306.

Shalaby, T., G. Fiaschetti, et al. (2013). "G-quadruplexes as potential therapeutic targets for embryonal tumors." Molecules **18**(10): 12500-12537.

Shao, G., J. Patterson-Fortin, et al. (2009). "MERIT40 controls BRCA1-Rap80 complex integrity and recruitment to DNA double-strand breaks." Genes Dev **23**(6): 740-754.

Shenton, D., J. B. Smirnova, et al. (2006). "Global translational responses to oxidative stress impact upon multiple levels of protein synthesis." J Biol Chem **281**(39): 29011-29021.

Shi, Y., D. C. Di Giammartino, et al. (2009). "Molecular architecture of the human pre-mRNA 3' processing complex." Mol Cell **33**(3): 365-376.

Shia, W. J., A. J. Okumura, et al. (2012). "PRMT1 interacts with AML1-ETO to promote its transcriptional activation and progenitor cell proliferative potential." Blood **119**(21): 4953-4962.

Shimobayashi, M. and M. Hall (2014). "Making new contacts: the mTOR network in metabolism and signalling crosstalk." Nat Rev Mol Cell Biol **15**: 155-162.

Shringarpure, R., T. Grune, et al. (2003). "Ubiquitin conjugation is not required for the degradation of oxidized proteins by proteasome." J Biol Chem **278**(1): 311-318.

Shuman, S. (2001). "Structure, mechanism, and evolution of the mRNA capping apparatus." Prog Nucleic Acid Res Mol Biol **66**: 1-40.

Shuman, S. (2015). "RNA capping: progress and prospects." RNA **21**(4): 735-737.

Siegfried, N. A., S. Busan, et al. (2014). "RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP)." Nat Methods **11**(9): 959-965.

Silva, G. M., D. Finley, et al. (2015). "K63 polyubiquitination is a new modulator of the oxidative stress response." Nat Struct Mol Biol **22**(2): 116-123.

Simarro, M., D. Mauger, et al. (2007). "Fas-activated serine/threonine phosphoprotein (FAST) is a regulator of alternative splicing." Proc Natl Acad Sci U S A **104**(27): 11370-11375.

Simone, R., P. Fratta, et al. (2015). "G-quadruplexes: Emerging roles in neurodegenerative diseases and the non-coding transcriptome." FEBS Lett **589**(14): 1653-1668.

Sims, R. J., 3rd, L. A. Rojas, et al. (2011). "The C-terminal domain of RNA polymerase II is modified by site-specific methylation." Science **332**(6025): 99-103.

Singh, G., A. Kucukural, et al. (2012). "The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus." Cell **151**(4): 750-764.

Sinvani, H., O. Haimov, et al. (2015). "Translational tolerance of mitochondrial genes to metabolic energy stress involves TISU and eIF1-eIF4GI cooperation in start codon selection." Cell Metab **21**(3): 479-492.

Sissi, C., B. Gatto, et al. (2011). "The evolving world of protein-G-quadruplex recognition: a medicinal chemist's perspective." Biochimie **93**(8): 1219-1230.

Smith, E., C. Lin, et al. (2011). "The super elongation complex (SEC) and MLL in development and disease." Genes Dev **25**: 661-672.

Sobhian, B., G. Shao, et al. (2007). "RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites." Science **316**(5828): 1198-1202.

Somers, J., L. A. Wilson, et al. (2015). "A common polymorphism in the 5' UTR of ERCC5 creates an upstream ORF that confers resistance to platinum-based chemotherapy." Genes Dev **29**(18): 1891-1896.

Sonenberg, N. and A. G. Hinnebusch (2009). "Regulation of translation initiation in eukaryotes: mechanisms and biological targets." Cell **136**(4): 731-745.

Song, E. J., S. L. Werner, et al. (2010). "The Prp19 complex and the Usp4Sart3 deubiquitinating enzyme control reversible ubiquitination at the spliceosome." Genes Dev **24**(13): 1434-1447.

Song, L., L. Wang, et al. (2010). "Sam68 up-regulation correlates with, and its down-regulation inhibits, proliferation and tumourigenicity of breast cancer cells." J Pathol **222**(3): 227-237.

Spence, J., R. R. Gali, et al. (2000). "Cell cycle-regulated modification of the ribosome by a variant multiubiquitin chain." Cell **102**(1): 67-76.

Stalder, L. and O. Muhlemann (2008). "The meaning of nonsense." Trends Cell Biol **18**(7): 315-321.

Stark, H. and R. Luhrmann (2006). "Cryo-electron microscopy of spliceosomal components." Annu Rev Biophys Biomol Struct **35**: 435-457.

Strahl, B. D., S. D. Briggs, et al. (2001). "Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1." Curr Biol **11**(12): 996-1000.

Subramanian, M., F. Rage, et al. (2011). "G-quadruplex RNA structure as a signal for neurite mRNA targeting." EMBO Rep **12**(7): 697-704.

Suchankova, J., S. Legartova, et al. (2014). "PRMT1 arginine methyltransferase accumulates in cytoplasmic bodies that respond to selective inhibition and DNA damage." Eur J Histochem **58**(2): 2389.

Sun, Y., Y. C. Ma, et al. (2005). "Requirement of SRC-family tyrosine kinases in fat accumulation." Biochemistry **44**: 14455-14462.

Sun, Z. Z., T. Zhang, et al. (2016). "B7-H3 upregulates BRCC3 expression, antagonizing DNA damage caused by 5-Fu." Oncol Rep **36**(1): 231-238.

Sundquist, W. I. and S. Heaphy (1993). "Evidence for interstrand quadruplex formation in the dimerization of human immunodeficiency virus 1 genomic RNA." Proc Natl Acad Sci U S A **90**(8): 3393-3397.

Sundquist, W. I. and A. Klug (1989). "Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops." Nature **342**(6251): 825-829.

Susnow, N., L. Zeng, et al. (2009). "Bcl-2 family proteins as regulators of oxidative stress." Semin Cancer Biol **19**(1): 42-49.

Swiercz, R., M. D. Person, et al. (2004). "Ribosomal protein S2 is a substrate for mammalian protein arginine methyltransferase 3 (PRMT3)." Biochem J **Oct 8; Epub ahead of print**.

Sylvestersen, K. B., H. Horn, et al. (2014). "Proteomic analysis of arginine methylation sites in human cells reveals dynamic regulation during transcriptional arrest." Mol Cell Proteomics **13**(8): 2072-2088.

Taha, M. S., K. Nouri, et al. (2014). "Subcellular fractionation and localization studies reveal a direct interaction of the fragile X mental retardation protein (FMRP) with nucleolin." PLoS One **9**(3): e91465.

Takahama, K., K. Kino, et al. (2011). "Identification of Ewing's sarcoma protein as a G-quadruplex DNA- and RNA-binding protein." FEBS Journal **278**(6): 988-998.

Takahama, K., A. Takada, et al. (2013). "Regulation of telomere length by G-quadruplex telomere DNA- and TERRA-binding protein TLS/FUS." Chem Biol **20**(3): 341-350.

Tang, J., A. Frankel, et al. (2000). "PRMT1 is the predominant type I protein arginine methyltransferase in mammalian cells." J Biol Chem **275**(11): 7723-7730.

Tee, A. R., B. D. Manning, et al. (2003). "Tuberous sclerosis complex gene products, Tuberin and Hamartin, control mTOR signaling by acting as a GTPase-activating protein complex toward Rheb." Curr Biol **13**: 1259-1268.

Tee, W. W., M. Pardo, et al. (2010). "Prmt5 is essential for early mouse development and acts in the cytoplasm to maintain ES cell pluripotency." Genes Dev **24**(24): 2772-2777.

Thandapani, P., T. R. O'Connor, et al. (2013). "Defining the RGG/RG motif." Mol Cell **50**(5): 613-623.

Thandapani, P., J. Song, et al. (2015). "Aven recognition of RNA G-quadruplexes regulates translation of the mixed lineage leukemia protooncogenes." Elife **4**.

Thedieck, K., P. Polak, et al. (2007). "PRAS40 and PRR5-like protein are new mTOR interactors that regulate apoptosis." PLoS One **2**: e1217.

Thorslund, T., A. Ripplinger, et al. (2015). "Histone H1 couples initiation and amplification of ubiquitin signalling after DNA damage." Nature **527**(7578): 389-393.

Tisserant, A. and H. Konig (2008). "Signal-regulated Pre-mRNA occupancy by the general splicing factor U2AF." PLoS One **3**(1): e1418.

Todd, A. K., M. Johnston, et al. (2005). "Highly prevalent putative quadruplex sequence motifs in human DNA." Nucleic Acids Res **33**(9): 2901-2907.

Tome, J. M., A. Ozer, et al. (2014). "Comprehensive analysis of RNA-protein interactions by high-throughput sequencing-RNA affinity profiling." Nat Methods **11**(6): 683-688.

Topisirovic, I., G. J. Gutierrez, et al. (2009). "Control of p53 multimerization by Ubc13 is JNK-regulated." Proc Natl Acad Sci U S A **106**(31): 12676-12681.

Topisirovic, I., Y. V. Svitkin, et al. (2011). "Cap and cap-binding proteins in the control of gene expression." Wiley Interdiscip Rev RNA **2**(2): 277-298.

Tsai, W. C., S. Gayatri, et al. (2016). "Arginine demethylation of G3BP1 promotes stress granule assembly." J Biol Chem.

Tuesuwan, B., J. T. Kern, et al. (2008). "Simian virus 40 large T-antigen G-quadruplex DNA helicase inhibition by G-quadruplex DNA-interactive agents." Biochemistry **47**(7): 1896-1909.

Turner, M. R., O. Hardiman, et al. (2013). "Controversies and priorities in amyotrophic lateral sclerosis." Lancet Neurol **12**(3): 310-322.

Ulke-Lemee, A., L. Trinkle-Mulcahy, et al. (2007). "The nuclear PP1 interacting protein ZAP3 (ZAP) is a putative nucleoside kinase that complexes with SAM68, CIA, NF110/45, and HNRNP-G." Biochim Biophys Acta **1774**(10): 1339-1350.

Um, S. H., F. Frigerio, et al. (2004). "Absence of S6K1 protects against age- and diet-induced obesity while enhancing insulin sensitivity." Nature **431**: 200-205.

Valacca, C., S. Bonomi, et al. (2010). "Sam68 regulates EMT through alternative splicing-activated nonsense-mediated mRNA decay of the SF2/ASF proto-oncogene." J Cell Biol **191**(1): 87-99.

Valton, A. L., V. Hassan-Zadeh, et al. (2014). "G4 motifs affect origin positioning and efficiency in two vertebrate replicators." EMBO J **33**(7): 732-746.

Van Nostrand, E. L., G. A. Pratt, et al. (2016). "Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP)." Nat Methods **13**(6): 508-514.

Vander Haar, E., S. I. Lee, et al. (2007). "Insulin signalling to mTOR mediated by the Akt/PKB substrate PRAS40." Nat Cell Biol **9**: 316-323.

Vasilyev, N., A. Polonskaia, et al. (2015). "Crystal structure reveals specific recognition of a G-quadruplex RNA by a beta-turn in the RGG motif of FMRP." Proc Natl Acad Sci U S A **112**(39): E5391-5400.

Vaughn, J. P., S. D. Creacy, et al. (2005). "The DEXH protein product of the DHX36 gene is the major source of tetramolecular quadruplex G4-DNA resolving activity in HeLa cell lysates." J Biol Chem **280**(46): 38117-38120.

Venables, J. P., C. Vernet, et al. (1999). "T-STAR/ETOILE: a novel relative of SAM68 that interacts with an RNA-binding protein implicated in spermatogenesis." Hum Mol Genet **8**(6): 959-969.

Vernet, C. and K. Artzt (1997). "STAR, a gene family involved in signal transduction and activation of RNA." Trends Genet **13**(12): 479-484.

Vogel, G. and S. Richard (2012). "Emerging roles for Sam68 in adipogenesis and neuronal development." RNA Biol **9**(9): 1129-1133.

Vogel, P., M. F. Schneider, et al. (2014). "Improving site-directed RNA editing in vitro and in cell culture by chemical modification of the guideRNA." Angew Chem Int Ed Engl **53**(24): 6267-6271.

von Hacht, A., O. Seifert, et al. (2014). "Identification and characterization of RNA guanine-quadruplex binding proteins." Nucleic Acids Res **42**(10): 6630-6644.

Waldmann, T., A. Izzo, et al. (2011). "Methylation of H2AR29 is a novel repressive PRMT6 target." Epigenetics Chromatin **4**: 11.

Walport, L. J., R. J. Hopkinson, et al. (2016). "Arginine demethylation is catalysed by a subset of JmjC histone lysine demethylases." Nat Commun **7**: 11974.

Wang, B., K. Hurov, et al. (2009). "NBA1, a new player in the Brca1 A complex, is required for DNA damage resistance and checkpoint control." Genes Dev **23**(6): 729-739.

Wang, B., S. Matsuoka, et al. (2007). "Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response." Science **316**(5828): 1194-1198.

Wang, H., Z.-Q. Huang, et al. (2001). "Methylation of histone H4 at arginine 3 facilitates transcriptional activation by nuclear hormone receptor." Science **293**: 853-857.

Wang, L., T. E. Harris, et al. (2007). "PRAS40 regulates mTORC1 kinase activity by functioning as a direct inhibitor of substrate binding." J Biol Chem **282**: 20036-20044.

Wang, L., S. Pal, et al. (2008). "Protein arginine methyltransferase 5 suppresses the transcription of the RB family of tumor suppressors in leukemia and lymphoma cells." Mol Cell Biol **28**(20): 6262-6277.

Wang, L., H. Zeng, et al. (2015). "MED12 methylation by CARM1 sensitizes human breast cancer cells to chemotherapy drugs." Sci Adv **1**(9): e1500463.

Wang, L., Z. Zhao, et al. (2014). "CARM1 methylates chromatin remodeling factor BAF155 to enhance tumor progression and metastasis." Cancer Cell **25**(1): 21-36.

Wang, M., J. Fuhrmann, et al. (2014). "Protein arginine methyltransferase 5 catalyzes substrate dimethylation in a distributive fashion." Biochemistry **53**(50): 7884-7892.

Wang, Q., Y. Li, et al. (2016). "Sam68 affects cell proliferation and apoptosis of human adult T-acute lymphoblastic leukemia cells via AKT/mTOR signal pathway." Leuk Res **46**: 1-9.

Wang, Q., J. Q. Liu, et al. (2011). "G-quadruplex formation at the 3' end of telomere DNA inhibits its extension by telomerase, polymerase and unwinding by helicase." Nucleic Acids Res **39**(14): 6229-6237.

Wang, S., X. Tan, et al. (2012). "The role of protein arginine-methyltransferase 1 in gliomagenesis." BMB Rep **45**(8): 470-475.

Wei, H., R. Mundade, et al. (2014). "Protein arginine methylation of non-histone proteins and its role in diseases." Cell Cycle **13**(1): 32-41.

Wei, H. M., H. H. Hu, et al. (2014). "Arginine methylation of the cellular nucleic acid binding protein does not affect its subcellular localization but impedes RNA binding." FEBS Lett **588**(9): 1542-1548.

Wei, J., Y. Yang, et al. (2015). "Escape, or Vanish: Control the Fate of p53 through MDM2-Mediated Ubiquitination." Anticancer Agents Med Chem **16**(2): 174-189.

Westmark, C. J. and J. S. Malter (2007). "FMRP mediates mGluR5-dependent translation of amyloid precursor protein." PLoS Biol **5**(3): e52.

Weyn-Vanhentenryck, S. M., A. Mele, et al. (2014). "HITS-CLIP and integrative modeling define the Rbfox splicing-regulatory network linked to brain development and autism." Cell Rep **6**(6): 1139-1152.

Witten, J. T. and J. Ule (2011). "Understanding splicing regulation through RNA splicing maps." Trends Genet **27**: 89-97.

Woelk, T., S. Sigismund, et al. (2007). "The ubiquitination code: a signalling problem." Cell Div **2**: 11.

Wolfe, A. L., K. Singh, et al. (2014). "RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer." Nature **513**(7516): 65-70.

Wolfe, A. L., K. Singh, et al. (2014). "RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer." Nature **Jul 27**: doi: 10.1038/nature13485.

Workman, E., S. J. Kolb, et al. (2012). "Spliceosomal small nuclear ribonucleoprotein biogenesis defects and motor neuron selectivity in spinal muscular atrophy." Brain Res **1462**: 93-99.

Wu, F. H., L. Q. Luo, et al. (2014). "Cyclin D1b splice variant promotes alphavbeta3-mediated adhesion and invasive migration of breast cancer cells." Cancer Lett **355**(1): 159-167.

Wu, J. and W. Xu (2012). "Histone H3R17me2a mark recruits human RNA polymerase-associated factor 1 complex to activate transcription." Proc Natl Acad Sci U S A **109**(15): 5675-5680.

Wu, Y., X. Xu, et al. (2015). "Sam68 regulates cell proliferation and cell adhesion-mediated drug resistance (CAM-DR) via the AKT pathway in non-Hodgkin's lymphoma." Cell Prolif **48**(6): 682-690.

Xia, Z., L. A. Donehower, et al. (2014). "Dynamic analyses of alternative polyadenylation from RNA-seq reveal a

3'-UTR landscape across seven tumour types." Nat Commun **5**: 5274.

Xiao, R., P. Tang, et al. (2012). "Nuclear matrix factor hnRNP U/SAF-A exerts a global control of alternative splicing by regulating U2 snRNP maturation." Mol Cell **45**(5): 656-668.

Xu, J., A. H. Wang, et al. (2013). "Arginine Methylation Initiates BMP-Induced Smad Signaling." Mol Cell **51**: 5-19.

Xu, M., B. Skaug, et al. (2009). "A ubiquitin replacement strategy in human cells reveals distinct mechanisms of IKK activation by TNFalpha and IL-1beta." Mol Cell **36**(2): 302-314.

Xu, W., F. Bolduc, et al. (2012). "The use of a combination of computer-assisted structure prediction and SHAPE probing to elucidate the secondary structures of five viroids." Mol Plant Pathol **13**(7): 666-676.

Xu, Y., K. Kaminaga, et al. (2008). "Human telomeric RNA in G-quadruplex structure." Nucleic Acids Symp Ser (Oxf)(52): 175-176.

Xu, Y., Y. Suzuki, et al. (2010). "Telomeric repeat-containing RNA structure in living cells." Proc Natl Acad Sci U S A **107**(33): 14579-14584.

Yadav, N., D. Cheng, et al. (2008). "CARM1 promotes adipocyte differentiation by coactivating PPARgamma." EMBO Rep **9**(2): 193-198.

Yamagata, K., H. Daitoku, et al. (2008). "Arginine methylation of FOXO transcription factors inhibits their phosphorylation by Akt." Mol Cell **32**(2): 221-231.

Yang, X. J. and E. Seto (2008). "Lysine acetylation: codified crosstalk with other posttranslational modifications." Mol Cell **31**(4): 449-461.

Yang, Y. and M. T. Bedford (2013). "Protein arginine methyltransferases and cancer." Nat Rev Cancer **13**(1): 37-50.

Yang, Y., A. Hadjikyriacou, et al. (2015). "PRMT9 is a type II methyltransferase that methylates the splicing factor SAP145." Nat Commun **6**: 6428.

Yang, Y., Y. Lu, et al. (2010). "TDRD3 is an effector molecule for arginine-methylated histone marks." Mol Cell **40**(6): 1016-1023.

Ying, Z., M. Mei, et al. (2015). "Histone Arginine Methylation by PRMT7 Controls Germinal Center Formation via Regulating Bcl6 Transcription." J Immunol **195**(4): 1538-1547.

Yoshida, K., M. Sanada, et al. (2011). "Frequent pathway mutations of splicing machinery in myelodysplasia." Nature **478**(7367): 64-69.

Yost, J. M., I. Korboukh, et al. (2011). "Targets in epigenetics: inhibiting the methyl writers of the histone code." Curr Chem Genomics **5**(Suppl 1): 72-84.

Yu, Z., T. Chen, et al. (2009). "A mouse PRMT1 null allele defines an essential role for arginine methylation in genome maintenance and cell proliferation." Mol Cell Biol **29**: 2982-2996.

Yu, Z., G. Vogel, et al. (2012). "The MRE11 GAR motif regulates DNA double-strand break processing and ATR activation." Cell Res **22**(2): 305-320.

Zarnack, K., J. Konig, et al. (2013). "Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements." Cell **152**(3): 453-466.

Zeqiraj, E., L. Tian, et al. (2015). "Higher-Order Assembly of BRCC36-KIAA0157 Is Required for DUB Activity and Biological Function." Mol Cell **59**(6): 970-983.

Zhang, G., M. Hubalewska, et al. (2009). "Transient ribosomal attenuation coordinates protein synthesis and co-translational folding." Nat Struct Mol Biol **16**(3): 274-280.

Zhang, H., N. Chen, et al. (2016). "The nuclear protein Sam68 is recruited to the cytoplasmic stress granules during enterovirus 71 infection." Microb Pathog **96**: 58-66.

Zhang, J., M. Cao, et al. (2014). "ABRO1 suppresses tumourigenesis and regulates the DNA damage response by

stabilizing p53." Nat Commun **5**: 5059.
Zhang, L., N. T. Tran, et al. (2015). "Cross-talk between PRMT1-mediated methylation and ubiquitylation on RBM15 controls RNA splicing." Elife **4**.
Zhang, X. and X. Cheng (2003). "Structure of the predominant protein arginine methyltransferase PRMT1 and analysis of its binding to substrate peptides." Structure **11**(5): 509-520.
Zhang, Y., C. M. Gaetano, et al. (2014). "FMRP interacts with G-quadruplex structures in the 3'-UTR of its dendritic target Shank1 mRNA." RNA Biol **11**(11): 1364-1374.
Zhang, Y. and M. S. Sachs (2015). "Control of mRNA Stability in Fungi by NMD, EJC and CBC Factors Through 3'UTR Introns." Genetics **200**(4): 1133-1148.
Zhang, Z., J. Li, et al. (2009). "Expression and cytoplasmic localization of SAM68 is a significant and independent prognostic marker for renal cell carcinoma." Cancer Epidemiol Biomarkers Prev **18**(10): 2685-2693.
Zhang, Z., Y. Xu, et al. (2014). "High Sam68 expression predicts poor prognosis in non-small cell lung cancer." Clin Transl Oncol **16**(10): 886-891.
Zhao, D. Y., G. Gish, et al. (2016). "SMN and symmetric arginine dimethylation of RNA polymerase II C-terminal domain control termination." Nature **529**(7584): 48-53.
Zhao, Q., G. Rank, et al. (2009). "PRMT5-mediated methylation of histone H4R3 recruits DNMT3A, coupling histone and DNA methylation in gene silencing." Nat Struct Mol Biol **16**(3): 304-311.
Zhao, X., V. Jankovic, et al. (2008). "Methylation of RUNX1 by PRMT1 abrogates SIN3A binding and potentiates its transcriptional activity." Genes Dev **22**(5): 640-653.
Zheng, H., V. Gupta, et al. (2013). "A BRISC-SHMT complex deubiquitinates IFNAR1 and regulates interferon responses." Cell Rep **5**(1): 180-193.
Zheng, K. W., S. Xiao, et al. (2013). "Co-transcriptional formation of DNA:RNA hybrid G-quadruplex and potential function as constitutional cis element for transcription control." Nucleic Acids Res **41**(10): 5533-5541.
Zheng, S., J. Moehlenbrink, et al. (2013). "Arginine methylation-dependent reader-writer interplay governs growth control by E2F-1." Mol Cell **52**(1): 37-51.
Zhou, H., I. Wertz, et al. (2004). "Bcl10 activates the NF-kappaB pathway through ubiquitination of NEMO." Nature **427**(6970): 167-171.
Zhou, J., M. Cheng, et al. (2015). "Inhibition of Sam68 triggers adipose tissue browning." J Endocrinol **225**(3): 181-189.
Zhou, Q., T. Li, et al. (2012). "RNA polymerase II elongation control." Annu Rev Biochem **81**: 119-143.
Zorio, D. A. and D. L. Bentley (2004). "The link between mRNA processing and transcription: communication works both ways." Exp Cell Res **296**(1): 91-97.
Zou, L., H. Zhang, et al. (2012). "Correlation of SRSF1 and PRMT1 expression with clinical status of pediatric acute lymphoblastic leukemia." J Hematol Oncol **5**: 42.