Arginine Methylation of Proteins with RGG/RG Motifs in Diseases

by

Ting Cai

Department of Biochemistry

McGill University, Montréal, Québec, Canada

October 2021

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy.

ABSTRACT

Protein arginine methylation, catalyzed by protein arginine methyltransferases (PRMTs), is an important post-translational modification that diversely regulates biological processes. RGG/RG motifs, which include RGG and RG repeats interspersed with spacers of different amino acids, are preferred substrates of PRMTs. RGG/RG motifs containing proteins have been identified in a wide range of life forms (e.g., bacterial, fungi, plant, and animal) and viruses. A large portion of the RNA binding proteins (RBPs) harbor RGG/RG motifs. These proteins are involved in the key process of RNA metabolism and their dysregulation lead to developmental diseases, cancer, and increase susceptibility to viral infections. My thesis defines the role of arginine methylation of two RGG/RG motif containing proteins: RBMX and SARS-CoV-2 Nucleocapsid (N) protein.

RBMX, also known as heterogeneous nuclear ribonucleoprotein (hnRNP) G, is an RBP that contains an RGG/RG motif at its C-terminus. A genetic deletion within the *RBMX* gene is associated with the Shashi X-linked intellectual disability (XLID) syndrome. This deletion introduces a frameshift in the coding sequence and is predicted to truncate the RBMX protein, deleting 38 of the C-terminal amino acids including the RGG/RG motif. To define whether the *RBMX* genetic deletion is sufficient to cause the Shashi-XLID, we generated human induced pluripotent stem cells (hiPSCs) containing the RBMX RGG/RG motif deletion (RBMX-ΔRGG) using CRISPR/Cas9. These iPSCs showed aberrant activation of the p53 pathway and displayed defects in their ability to differentiate into mature cortical neurons, exhibiting excessive apoptosis and altered neuron morphology. Molecularly, we show that the RGG/RG motif of RBMX is symmetrical dimethylated by PRMT5 and it regulates pre-mRNA splicing of many genes including the *MDM4* pre-mRNA. RBMX interacts with the splicing enhancer SRSF1 and regulates its association to the *MDM4* pre-mRNA. Interestingly, immunofluorescence analysis

showed that RBMX and SRSF1 form nuclear puncta that are sensitive to 1,6-hexanediol treatment or PRMT5 depletion. Taken together, we propose that the Shashi-XLID is a developmental disorder caused by RBMX mediated aberrant alternative splicing and p53 activation in neurons.

The second part of my thesis focuses on the Nucleocapsid (N) protein of SARS-CoV-2, another RGG motif-containing protein that plays an essential role in the packaging and assembly of the virus. We show that PRMT1 methylates SARS-CoV-2 N R95 and R177 residues within its RGG/RG sequences. Arginine methylation of N was also confirmed by immunoblotting viral proteins extracted from SARS-CoV-2 virions. We demonstrate that the arginine methylation of N is required for its binding to the viral genomic RNA (gRNA). Treatment with type I PRMT inhibitor (MS023) or substitution of R95K or R177K reduced the ability of N to interact with the 5'-UTR (untranslated region) of SARS-CoV-2 gRNA. Proteomic analysis identified host stress granule nucleation factors G3BP1 and G3BP2 as binding partners of N protein. Ectopic expression of SARS-CoV-2 N protein dissolved arsenite-induced G3BP1 stress granules and this was reverted when cells were treated with MS023 or when R95K substitution was introduced. Finally, pre-treatment of VeroE6 cells with MS023 significantly reduced SARS-CoV-2 replication. Thus, our findings demonstrate arginine methylation is an important post-translational modification to regulate N protein function and identify PRMT1 as a druggable target against SARS-CoV-2 virion production.

In sum, my thesis identifies RBMX and its RGG/RG motif involved in PRMT5/MDM4/p53 pathway regulation, and arginine methylation is required for the proper function of SARS-CoV-2 N protein. These findings contribute to the understanding of the molecular basis of developmental disease and viral infection regulated by arginine methylation and PRMTs.

SOMMAIRE

La méthylation de l'arginine est une modification post-traductionnelle (PTM) qui régule divers processus biologiques. Elle est catalysée par les protéines arginine méthyltransférases (PRMTs) dont les substrats préférés contiennent des motifs RGG/RG, identifiés dans un large éventail des êtres vivants, notamment bactérien, fongique, végétal, animal et viral. Une grande partie des protéines de liaison à l'ARN contient des motifs RGG/RG. Ces protéines sont impliquées dans le métabolisme de l'ARN, contribuant aux maladies du développement, aux cancers et aux infections virales. Ma thèse vise à caractériser la fonction de la méthylation de l'arginine de RBMX et de la protéine de Nucléocapside (N) du SARS-CoV-2.

RBMX (hnRNPG), est une protéine de liaison à l'ARN qui contient un motif RGG/RG à son extrémité C-terminale. Une délétion génétique au sein du gène RBMX est associée au syndrome de déficience intellectuelle liée à l'X de Shashi (XLID). Elle induit un décalage du cadre de lecture dans les séquences codantes et aboutit à la formation d'une forme tronquée de RBMX, supprimant ainsi 38 acides aminés de la partie C-terminale, y compris le motif RGG/RG. Pour comprendre si cette délétion est suffisante pour provoquer le syndrome de Shashi-XLID, nous avons généré des cellules souches pluripotentes induites humaines (hiPSC) contenant la délétion du motif RBMX RGG/RG (RBMX-ΔRGG) par CRISPR/Cas9. Ces iPSCs montrent une activation aberrante de la voie p53 et des défauts de différenciation en neurones corticaux matures, provoquant ainsi une induction d'apoptose et une altération de la morphologie neuronale. D'un point de vue moléculaire, nous montrons que le motif RGG/RG de RBMX est méthylé par PRMT5 et régule l'épissage des ARN pré-messagers de nombreux gènes dont MDM4. RBMX interagit avec SRSF1 et régule son association avec le pré-ARNm MDM4. En effet, l'analyse par immunofluorescence a montré que RBMX et SRSF1 forment des points lacrymaux nucléaires

sensibles au traitement au 1,6-hexanediol ou à la délétion de PRMT5. Par conséquent, nous proposons que le Shashi-XLID est un trouble du développement causé par l'épissage alternatif aberrant médié par RBMX et l'activation de p53.

La deuxième partie de ma thèse porte sur la protéine de Nucléocapside (N) du SARS-CoV-2, une autre protéine avec le motif RGG jouant un rôle essentiel dans l'empaquetage et l'assemblage du virus. Nos résultats montrent que PRMT1 méthyle les résidus R95 et R177 de la séquence RGG/RG de la protéine N. Cette méthylation a été confirmée par immunotransfert de protéines virales extraites de virions du SARS-CoV-2. Nous démontrons que la méthylation de l'arginine de N est nécessaire pour sa liaison à l'ARN génomique viral (ARNg). Le traitement avec MS023, inhibiteur de PRMTs de classe I ou la substitution de R95/K ou R177/K montre une réduction de la capacité de N à interagir avec la région 5'-UTR de l'ARNg du virus. L'analyse protéomique a identifié que les facteurs de nucléation des granules de stress G3BP1 et G3BP2 sont liés à la protéine N. L'expression ectopique de N diminue la formation des granules de stress G3BP1 en présence de l'arsénite. Un phénotype opposé est observé lorsque les cellules sont traitées avec MS023 ou lorsque la substitution R95K a été introduite. Enfin, le prétraitement des cellules VeroE6 avec MS023 réduit la réplication du SARS-CoV-2. Ainsi, nos résultats démontrent que la méthylation de l'arginine est une PTM importante pour réguler la fonction de la protéine N et identifient PRMT1 comme une cible thérapeutique contre la production de virions du SARS-CoV-2.

Dans l'ensemble, ma thèse identifie RBMX et son motif RGG/RG, impliqué dans la régulation de la voie PRMT5/MDM4/p53, et la méthylation de l'arginine, comme étant un facteur essentiel pour le bon fonctionnement de la protéine N du SARS-CoV-2. Ces résultats contribuent à la compréhension des maladies du développement et des infections virales régulées par les PRMTs.

PREFACE

As a manuscript-based thesis, I will include the text and figures of two original research manuscripts that have been published. Each of these chapters (Chapters 2 and 3) 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 4.

Papers included in this Thesis

Chapter 2	**Cai, T**., Cinkornpumin, J. K., Yu, Z., Villarreal, O. D., Pastor, W. A., & Richard, S. (2021). Deletion of RBMX RGG/RG motif in Shashi-XLID syndrome leads to aberrant p53 activation and neuronal differentiation defects. ***Cell Reports***, 36(2): 109337 doi: 10.1016/j.celrep.2021.109337.
Chapter 3	**Cai, T.**, Yu, Z., Wang, Z., Liang, C., & Richard, S. (2021). Arginine methylation of SARS-Cov-2 nucleocapsid protein regulates RNA binding, its ability to suppress stress granule formation, and viral replication. ***Journal of Biological Chemistry***, 297(1):100821 doi: 10.1016/j.jbc.2021.100821.

Contribution of the Authors:

All studies in this thesis were performed under the guidance of Dr. Stéphane Richard. The candidate performed the majority of the research presented in this thesis and wrote the manuscripts with support, assistance and guidance from Dr. Stéphane Richard. The contribution of other authors to this work is as follows:

In Chapter 2, Dr. Zhenbao Yu, Research Associate in our laboratory, performed the mass spectrometry for RBMX methylation in Figure 2.1B. Dr. Oscar Villarreal, a post-doctoral fellow in our laboratory, bioinformatically analyzed the RNA-sequencing data from DRGG1 and control iPS cells and NPCs, generating the Figures 2.1B, 2.3B, 2.6A and 2.6C, Supplemental Figure S2.1B, S2.3B, S2.5C and sashimi plots in Supplemental Figure S2.6. Flag-RBMX constructs used

in the study were generated by Dr. Zhenbao Yu. Jessica Cinkornpumin, a Ph.D. student in Dr. William Pastor's laboratory, and I worked together to generate, maintain and differentiate the iPS cells. I collected samples, performed the downstream experiment and analysis. Dr. William Pastor and Jessica Cinkornpumin provided expertise in stem cell culture and differentiation. Drs. William Pastor and Zhenbao Yu also contributed to the critical reading of the manuscript.

In Chapter 3, Dr. Zhenbao Yu performed the *in vitro* methylation assay (Figure 3.1 A-F) and the cell survival assay (Figure 3.6 A). Dr. Zhen Wang performed the SARS-CoV-2 infection experiments in BL3 lab at the Research Institute of McGill University Health Centre (Figure 3.6B, C). Drs. Zhenbao Yu and Chen Liang contributed to writing and critical reading of the manuscript.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor Dr. Stéphane Richard for providing me the opportunity to explore these brilliant projects and for training me to think and work independently as a scientist. It was his forward-looking and enthusiasm for science that guided me through all the difficulties in the projects. I am especially grateful for his guidance with presentation skills and manuscript writing and editing.

I am thankful for the invaluable feedback and encouragement from Dr. Sonia Victoria Del Rincon and Dr. Sidong Huang, members of my Research Advisory Committee.

I would also like to express my sincere gratitude to all the members in Richard Lab. I would particularly like to thank Dr. Zhenbao Yu for his helping with the projects and thoughtful comments over the past few years. Special thanks to Claudia Dominici, Jeesan Lee and Nivine Srour for all their support. It is a great pleasure to work with them as a team where we talked about science and helped each other make progress.

I acknowledge the contributions from our collaborators listed in the manuscripts who made the studies feasible. I would also like to express my deep appreciation to Dr. William Pastor for his insightful thoughts and willingness to share his expertise through the exciting collaboration. Sincere thanks to Jessica Cinkornpumin, who worked with me for a whole year to move the project forward.

I appreciate the financial supports from Fonds de recherche du Québec en Santé, TD bank scholarship, and Lady Davis Institute and the Department of Biochemistry for the various travel awards.

Finally, I am extremely grateful to my parents Zubing Cai and Xiaofang Ouyang and my boyfriend Xianbing Zhu. It is their selfless love and continuous support that made this thesis possible.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

1,6-HD	1,6-hexanediol
2′OMe	2′O methylation
53BP1	tumor suppressor p53-binding protein 1
3'ss	3’ splice site
5'ss	5’ splice site
ActD	actinomycin D
aDMA	asymmetric dimethylarginine
AKT	protein kinase B
ALS	Amyotrophic lateral sclerosis
AML	acute myeloid leukemia
APPsw	Swedish mutant form of human Aβ precursor protein
AS	alternative splicing
ATR	rad3-related protein
BCR	B cell receptor
cGAS	cyclic GMP-AMP synthase
ChIP	chromatin immunoprecipitation
CHK1	checkpoint kinase 1
CNBP	cellular nucleic acid-binding protein
CNS	central nervous system
COPI	coat protein complex I
CoV	Coronavirus
CSNK1A1	casein kinase 1 alpha 1
CTCF	CCCTC-binding factor
CTD	C-terminal domain
DDAH1	dimethylarginine dimethylaminohydrolase 1
DHX9	DExH-box helicase 9
DLBCL	diffuse large B cell lymphoma
DOT1L	disruptor of telomeric silencing 1-like
DSB	double-stranded break
E protein	envelope protein
EAE	experimental autoimmune encephalomyelitis
EGFR	epidermal growth factor receptor
eIF4G	eukaryotic translation initiation factor 4 G
EMT	epithelial–mesenchymal transition
EWS	Ewing's sarcoma
EZH2	zeste homolog 2
FDR	false discovery rate
FOXP1	forkhead box P1
FTD	frontotemporal dementia

FUS	fused in sarcoma
G3BP1	Ras GTPase-activating protein-binding protein 1
GFI1	growth factor independence 1
HCV	hepatitis C virus
hnRNP	heterogeneous nuclear ribonucleoprotein
HOXA9	homeobox A9
HR	homologous recombination
HSV	herpes simplex virus
IBV	infectious bronchitis virus
ICM	inner cell mass
IDR	intrinsically disordered region
IFN	type I interferon
iPSC	induced pluripotent stem cell
IR	γ-irradiation
IRES	internal ribosome entry site
ITAF	IRES transacting factor
KLF	Krüppel-like factor
LLPS	liquid-liquid Phase separation
lncRNA	long noncoding RNAs
LTR	long terminal repeat
M protein	membrane protein
m^5C	5-methylcytidine
m^6A	N^6-methyladenosine
MEL	murine erythroleukemia
MEF	mouse embryonic fibroblast
MERS	middle east respiratory syndrome
mESC	mouse embryonic stem cell
MHV	mouse hepatitis virus
MLL	mixed-lineage leukemia
MMA	monomethylarginine
MPN	myeloproliferative neoplasms
MRE11	meiotic recombination 11
MS	mass spectrometry
N protein	nucleocapsid protein
NBS1	nijmegen breakage syndrome 1
NHEJ	non-homologous DNA end joining
NMR	nuclear magnetic resonance
NPCs	neural progenitor cells
NTD	N-terminal domain
OPC	oligodendrocyte precursor cell
PAD4	peptidyl arginine deiminase

PAR	poly (ADP-ribose)
PAR-CLIP	photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation
PFV	prototype foamy virus
PI3K	phosphoinositide 3-kinases
PKM	pyruvate kinase muscle isozyme
PML	promyelocytic leukemia
poly(GR)	glycine-arginine repeat
PR	progesterone receptor
PRC2	polycomb repressive complex 2
pre-mRNA	precursor messenger RNA
PRMT	protein arginine methyltransferase
PSD	postsynaptic density
PTBP1	polypyrimidine tract-binding protein 1
PTM	post-translational modifications
RAN translation	repeat-associated non-ATG translation
RBMX	RNA binding motif protein X-linked
RBP	RNA binding protein
rDNA	ribosomal targeted RNA gene
RIP	RNA immunoprecipitations
RPA	replication protein A
S protein	Spike protein
SAM	S-adenosyl methionine
SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
sDMA	symmetric dimethylarginine
SE	splicing enhancer
SELEX	systematic evolution of ligands by exponential enrichment
SET1	SET domain containing 1
SG	stress granule
SILAC	stable isotope labeling by amino acids in cell culture
Sm	small nuclear ribonucleoprotein
SMA	spinal muscular atrophy
smFISH	single molecule fluorescence in situ hybridization
SMN	survival motor neuron
snRNP	small nuclear ribonucleoprotein
SR-rich LKR	serine/arginine-rich linker region
SRSF1	serine/arginine-rich splicing factor 1
SS	splicing silencer
STAT	signal transducer and activator of transcription
TDRD3	Tudor domain-containing protein 3
TFF1	Trefoil factor 1
TGEV	transmissible gastroenteritis virus

TOP1	topoisomerase I
TOP3B	topoisomerase IIIB
TRIM25	tripartite motif protein 25
TRS	transcriptional regulatory sequence
TSA	Trichostatin A
TSS	transcription start site
uORF	upstream open reading frame
vRNP	viral ribonucleoprotein
WDR5	WD Repeat Domain 5
XLID	X-linked-intellectual-disability
ZNF326	zinc Finger Protein 326

CHAPTER 1 Introduction

1.1 Protein arginine methylation

Post-translational modifications (PTM) are chemical modifications that play key roles in protein metabolism. Different types of PTMs have been identified, including phosphorylation, acetylation, glycosylation, methylation, ubiquitylation, sumoylation, myristoylation, and S-nitrosylation. PTMs can affect protein folding, conformation change, localization and catalytic activity, leading to diverse regulation of biological processes, including signaling transduction and epigenetic alteration (Millar et al., 2019).

Arginine is an α-amino acid containing an α-carboxylic acid group, an α-amino group, and a side chain with the guanidino group. The pK_a value of arginine is ~13.8, the highest among all the amino acids. At the physiological pH, arginine is a positively charged, aliphatic amino acid. Arginine is often found to extend beyond the surfaces of globular proteins and interacts with other proteins and nucleic acids by forming hydrogen bonds and van der Waals interactions (Guccione and Richard, 2019).

Arginine methylation is an important PTM catalyzed by protein arginine methyltransferases (PRMTs). Different types of PRMTs have been identified in eukaryotes, but they are absent from prokaryotes. Arginine can be methylated in three different ways: ω-N^G-monomethylarginine (MMA), ω-N^G,N^G-asymmetric dimethylarginine (aDMA), and ω-N^G,N'^G-symmetric dimethylarginine (sDMA). Importantly, methylation does not affect the p*Ka* of guanidinium moiety, but alters the shape and the charge distribution, subsequently increasing its hydrophobicity (Evich et al., 2016).

1.1.1 Protein arginine methyltransferases (PRMTs)

In mammalian cells, PRMTs are a family of nine enzymes transferring a methyl group from S-adenosyl methionine (SAM) to arginine guanidinium nitrogen. PRMTs share a conserved methyltransferase domain with pockets for SAM and substrate binding. PRMTs can be classified into three groups: type I enzymes catalyze asymmetric aDMA formation including PRMT1, 2, 3, 4 (CARM1), 6, and 8; type II enzymes catalyze symmetric sDMA formation, including PRMT5 and 9; type III enzyme catalyzes MMA formation and includes one member PRMT7. In fact, type I and II enzymes catalyze MMA formation as a biosynthetic intermediate before completing to dimethylation. Here, I'll discuss the major type I enzyme PRMT1 (Tang et al., 2000), the major type II enzyme PRMT5 (Branscombe et al., 2001) and type III enzyme PRMT7 (Zurita-Lopez et al., 2012) (Figure 1.1).

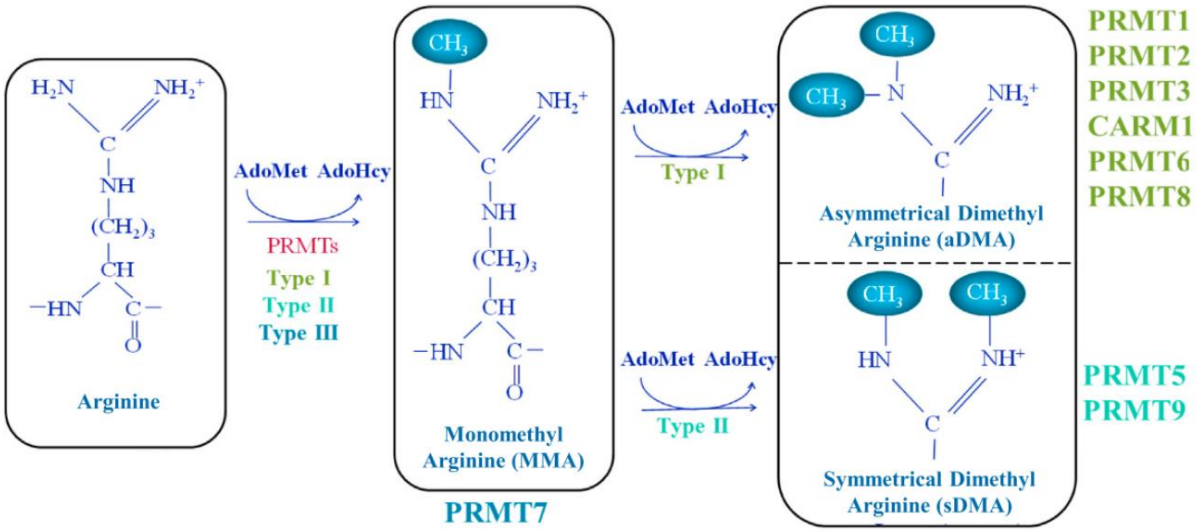


Figure 1. 1 Arginine methylation by PRMTs. Adapted from (Blanc and Richard, 2017)**.**

PRMT1 is responsible for ~85% of the total arginine methylation in RAT1 fibroblast cells and mouse liver (Tang et al., 2000). Seven PRMT1 splicing isoforms vary in their N-terminal domain and they have been reported to have different enzymatic activity and substrate specificity. The

PRMT1v1 and PRMT1v2 are the major isoforms in human. PRMT1v1 localizes in the nucleus and the cytoplasm, whereas PRMT1v2 predominately localizes in the cytoplasm (Goulet et al., 2007). Functional studies in the field mainly focus on PRMT1v1, the most highly expressed among all the isoforms.

PRMT1 forms a ring-like dimer essential for SAM binding and enzymatic activity. The monomeric structure of PRMT1 can be divided into four groups: N-terminal, SAM binding, β barrel, and dimerization arm (Zhang and Cheng, 2003). The SAM binding domain has the consensus fold conserved in other SAM-dependent methyltransferases. The N-terminal region facilitates SAM binding by preventing its exposure to the solvent. The substrate-binding surface of PRMT1 is acidic, which is consistent with the fact that PRMT1 substrates contain multiple arginines (see 1.2). Three disconnected acidic grooves in the interface between SAM binding and β barrel (R1 and R2) and the outer surface of β barrel (R3) form binding sites for protein substrate with three or more RGG repeats (Zhang and Cheng, 2003). The active site pocket harboring targeted arginine locates between the SAM binding pocket and the β barrel. The dimerization arm and the outer surface of SAM binding site engage in a hydrophobic dimer interface, so that the dimer interaction is required for proper SAM binding. Moreover, a ring-like dimer could allow the product of the first methylation reaction, monomethylarginine, to enter the active site of the second molecule of the dimer, without releasing the substrate from the ring or replenishing the methyl donor. It is worth noting that, by gel filtration, PRMT1 is present in a large complex with molecular weight arranging from 275 kDa to 450 kDa, likely caused by oligomerization or interaction with other proteins (Zhang and Cheng, 2003).

PRMT1 generally methylates proteins with RGG/RG repeats (see 1.2.1). Although PRMT1 has a variety of substrates implicated in diverse cellular processes, the regulation of PRMT1

remains largely unknown. Mass spectrometry identified casein kinase 1 alpha 1 (CSNK1A1), a key factor for epidermal progenitor maintenance, as an essential PRMT1-interacting protein. CSNK1A1 phosphorylates PRMT1 and modulates its chromatin localization (Bao et al., 2017).

PRMT5 is the major type II enzyme. Rather than forming a homodimer, PRMT5 complexes with a variety of partner proteins that regulate its function and specificity. These complexes share a core component MEP50. PRMT5:MEP50 forms hetero-octamer where the PRMT5 subunits structure a core tetramer and four MEP50 subunits arranged peripherally (Antonysamy et al., 2012). Unlike PRMT1 dimerizing in a head-to-tail orientation, PRMT5 monomers interact with each other at the center of the oligomer through TIM barrel domain. The TIM barrel also extensively interacts with MEP50, which mediates interaction with PRMT5 substrates. Although MEP50 is away from the catalytic domain of PRMT5, PRMT5 has a moderate enzyme activity without MEP50. One of the hypotheses is that MEP50 interaction introducing a conformation change favors the PRMT5 tetramers formation (Antonysamy et al., 2012).

Like PRMT1, PRMT5 substrates are RGG/RG motif-containing proteins. These proteins are associated with numerous physiological processes such as RNA transcription, precursor messenger RNA (pre-mRNA) splicing, translation, apoptosis and DNA damage signaling, indicating the crucial role of these two PRMTs in biological systems.

PRMT7 is a type III arginine methyltransferase that solely monomethylates arginines (Jain and Clarke, 2019). Unlike other PRMTs, PRMT7 does not form dimers or oligomers. However, PRMT7 contains two tandem methyltransferase domains that resemble the configuration of a homodimer. The N-terminal methyltransferase domain is conserved as in other PRMTs, while the C-terminal domain lacks essential residues involved in SAM and substrate binding, suggesting only the N-terminal domain is functional (Cura et al., 2014). Although the C-terminal domain is

catalytically inactive, depletion of this region attenuates the activity of PRMT7 (Miranda et al., 2004b). The catalytic core module, similar to other PRMTs, can be divided into three parts: SAM binding, β barrel domain, and dimerization arm (Hasegawa et al., 2014; Miranda et al., 2004b).

PRMT7 has a substrate preference for arginine within an RXR motif (Feng et al., 2013). Not many substrates have been identified, owing to the low catalytic activity of PRMT7. *In vitro* methylation experiments identified histone H2B as a substrate of PRMT7 (Feng et al., 2013). Two other non-histone substrates, HSP70 and eIF2a, were reported as PRMT7 substrates recently (Haghandish et al., 2019; Szewczyk et al., 2020). Nevertheless, PRMT7 depletion causes dysregulation of skeletal muscles (Blanc et al., 2016), adipocytes (Leem et al., 2019) and immune cells (Ying et al., 2015), indicating the involvement of potential PRMT7 substrates. Loss-of-function of PRMT7 causes intellectual disability with dysmorphism, including syndrome of short stature, brachydactyly, impaired intellectual development, and seizures (SBIDSS) (Agolini et al., 2018; Kernohan et al., 2017; Poquerusse et al., 2021).

1.1.2 Methylarginine readers

The influence of PTMs is generally carried out by the proteins recognizing this modification, which often called a ‘Reader’. There are ~30 Tudor domain-containing proteins in mammalian cells, with one or multiple tandem Tudor domains. These proteins are classified into methylarginine-binding and methyl lysine-binding proteins. Tudor domain-containing proteins are the major group of methylarginine readers. Interestingly, most methylarginine-binding Tudor domain proteins are involved in RNA-related processes (Chen et al., 2011).

One of the well-characterized Tudor domain-containing proteins is survival motor neuron (SMN) (Selenko et al., 2001). Homologous deletion or mutation of *SMN1* causes neuromuscular disease spinal muscular atrophy (SMA) due to lacking corresponding SMN protein (Lefebvre et

al., 1995). An aromatic cage in the Tudor domain of SMN interacts with symmetrical dimethylarginine through cation-π interactions (Tripsianes et al., 2011). Point mutations in the Tudor domain were found in SMA patients with one copy of *SMN1* (Cusco et al., 2004), indicating the disrupted interaction with symmetrically methylated arginine can be a causative factor.

Tudor Domain Containing 3 (TDRD3) is another methylarginine reader. TDRD3 is recruited to the estrogen receptor-dependent promoters, and acts as a transcriptional co-activator through its Tudor domain binding to the histone activation marks H3R17me2a and H4R3me2a (Yang et al., 2010). The Tudor domain of TDRD3 can also bind to the asymmetrically methylated CTD domain of RNA polymerase II (Sims et al., 2011). This interaction facilitates TDRD3 binding to the *c-MYC* promoter, allowing its binding partner topoisomerase IIIB (TOP3B) to reduce negative supercoiling, resolve the RNA:DNA hybrids or R-loops and promote transcription (Yang et al., 2014). TOP3B itself is also methylated by PRMT1, 3, and 6. The arginine methylation promotes TOP3B topoisomerase activity and enhances its interaction with TDRD3 (Huang et al., 2018). Moreover, TDRD3 directly interacts with DExH-box helicase 9 (DHX9) via its Tudor domain, stimulating DHX9 helicase activity and functioning together with TOP3B to suppress promoter-associated R-loops (Yuan et al., 2021).

Because PRMT substrates significantly outnumber the currently characterized Tudor domain readers, expanding our knowledge of other Tudor Domain-containing proteins is required. Besides, it is intriguing to understand the plasticity of the Tudor Domain regarding its binding to diverse substrates (Chen et al., 2011).

1.1.3 Reversibility of arginine methylation

Whether arginine methylation is a permanent modification (Bedford and Clarke, 2009) or reversible is unclear. Stable isotope labeling by amino acids in cell culture (SILAC) analysis showed that several MMA sites were down-regulated after treatment of transcriptional inhibitor actinomycin D (ActD), but not the corresponding di-methylation or the protein expression levels, indicating MMA may be a dynamically regulated PTM (Sylvestersen et al., 2014).

It is an unsolved question to identify the 'arginine demethylases' that can directly remove methyl groups from arginine. Initially, peptidyl arginine deiminase 4 (PAD4) was reported to convert methylarginines on histone H3 and H4 to citrulline and release methylamine (Wang et al., 2004). However, a later study showed that methylaginine is a poor substrate to PAD4, compared to unmethylated arginines (Kearney et al., 2005). Lysine demethylases KDM3A, KDM4E, KDM5C and JMJD6 were reported to have some activity towards dimethylated arginines *in vitro*, however, the *in vivo* effect was neglectable (Liu et al., 2013; Walport et al., 2016).

1.1.4 PRMT scavenging

Type I and type II PRMTs have shared substrates. Asymmetrical methylation by PRMT1 is the major arginine methylation in cells. The absence of PRMT1 in mouse embryonic fibroblasts (MEFs) leads to increased MMA and sDMA, a phenomenon termed substrate scavenging (Dhar et al., 2013). However, PRMT5 knockdown does not significantly increase the global level of aDMA and MMA because it is not the major enzyme methylating RGG/RG motifs in cells. Small nuclear ribonucleoprotein SmB/B', a well-known substrate of PRMT5 (Dhar et al., 2013) is the only significantly changed MMA protein after PRMT5 depletion. Furthermore, different types of methylation would result in the differential regulation of cellular processes, because MMA, aDMA, and sDMA can be recognized by different Tudor domain reader proteins.

1.1.5 PRMT inhibitors

Several PRMTs inhibitors have been developed and widely used to study PRMTs associated molecular mechanisms and diseases, especially carcinogenesis (see 1.5.2).

Type I PRMT inhibitors include MS023 and GSK3368715 (Eram et al., 2016) (Fedoriw et al., 2019); both are SAM uncompetitive inhibitors. These inhibitors bind to the substrate-binding site, decreasing the global level of aDMA and increasing the MMA and sDMA in proteins, as also observed in PRMT1-depleted cells (Eram et al., 2016; Fedoriw et al., 2019). Importantly, both inhibitors are reversible, providing a temporary regulation tool for controlling the catalytical activity of type I PRMTs.

The well-characterized type II PRMT inhibitor, EPZ015666 (GSK3235025), is a selective inhibitor of PRMT5 (Chan-Penebre et al., 2015). EPZ015666 shows an unambiguous pattern of competitive inhibition with respect to the peptide substrate, and uncompetitive inhibition with respect to SAM. Crystal structure studies reveal that EPZ015666 binds in the peptide-binding site. In addition, the compound forms a cation-π interaction with SAM, which presumably contributes to its selectivity against other PRMTs (Chan-Penebre et al., 2015).

A recent study identified SGC3027, a prodrug form of SGC8158, selectively inhibits PRMT7 though occupying the adenosine pocket in the SAM-binding site (Szewczyk et al., 2020).

Extensive studies have identified the application of PRMT inhibitors as potent cancer therapy, and these studies are discussed in section 1.5.2.

1.2 RGG/RG motif-containing proteins

RGG/RG motifs are repeats of RGG or RG clustering close together and they often cause sequences to be disordered i.e the absence of α-helices or β-sheets (Thandapani et al., 2013). RGG/RG motifs containing proteins have been identified in a wide range of life forms, including archaea, bacteria, fungi, plants, animals, and viruses (Corley and Gready, 2008). More than 1000 human proteins, including greater than 75% of the RNA binding proteins (RBPs), harbor RGG/RG motifs (Thandapani et al., 2013). Many proteins contain several neighboring RGG/RG motifs. For example, fused in sarcoma (FUS) is a DNA/RNA binding protein harboring several RGG/RG motifs function together as a higher-ordered motif (Thandapani et al., 2013).

1.2.1 Function of RGG/RG motif

Arginine methylation of the RGG/RG motif affects nucleic acid binding, protein-protein interactions, and PTM crosstalk.

Nucleic acid binding

Arginine interacts with nucleotide bases via hydrogen bonding and π-stacking; thus, nucleotide conformation that provides access to the bases enhances opportunities for arginine interactions (Chong et al., 2018; Luscombe et al., 2001). Methylation of the arginine does not affect its charge but affects the shape and the charge distribution.

RGG/RG motif is an RNA binding motif. Fusion of hnRNP U RGG/RG motif to a heterologous bacterial protein converted the protein into an RNA binding protein (Kiledjian and Dreyfuss, 1992). The RGG/RG motif was reported to interact with G-quadruplexes RNAs. Structural studies revealed that the FMRP RGGGGR sequence specifically binds to the duplex-quadruplex junction, the mixed tetrad, and the duplex region of the G-quadruplexes RNA through shape complementarity, cation-π interactions, and multiple hydrogen bonds (Vasilyev et al.,

2015). Mutagenesis of the RGGGGR peptide disrupts the affinity and specificity of the RGG-RNA complex formation (Vasilyev et al., 2015). In addition, the RGG/RG motif of Aven binds with G-quadruplexes in the coding regions of *mixed-lineage leukemia* (*MLL1)* and *MLL4* mRNAs(Thandapani et al., 2015).

The RGG/RG motif can also bind to DNA but with a lower affinity (Kiledjian and Dreyfuss, 1992). Engineered RGG/RG motifs (RGGY and RGGF repeats) from FUS/TLS are able to bind and stabilize G-quadruplex DNA (Takahama et al., 2015). Moreover, RGG/RG motifs were reported to bind with nucleic acid-mimicking biopolymer poly (ADP-ribose) (PAR) by electrostatic interactions (Altmeyer et al., 2015).

Protein-protein interaction

RGG/RG motifs are important protein-protein interfaces. Methylated RGG/RG motifs are the interface of interaction with Tudor domain-containing proteins, as discussed earlier. There also exists Tudor domain-independent interactions mediated by RGG/RG motifs. One of the well-studied examples are a group of RGG/RG motifs containing proteins that bind with eukaryotic translation initiation factor 4 G (eIF4G) to regulate translation through their RGG/RG motifs. These proteins include SCD6 (Rajyaguru et al., 2012), SBP1 (Rajyaguru et al., 2012), NPL3 (Rajyaguru et al., 2012) and DED1 (Hilliker et al., 2011), interacting with eIF4G in a simultaneously, or competitively, or context-dependent manner (Rajyaguru and Parker, 2012). Moreover, RGG/RG motifs are implicated in virus-host interaction. Herpes simplex virus (HSV) 1 protein ICP27 is an RGG/RG motif-containing protein, and its interactions with host TBK1 and STING depends on ICP27 RGG/RG motif and TBK1 activity (Christensen et al., 2016). HnRNP A1, a host protein, interacts with human immunodeficiency virus type 1 (HIV-1) Rev protein through its RGG/RG motif (Hadian et al., 2009). Moreover, RGG/RG motifs promote protein

self-associating, evidenced by Ewing's sarcoma (EWS) protein (Shaw et al., 2010) and Scd6 (Poornima et al., 2019).

Liquid-liquid Phase separation (LLPS)

LLPS is a thermodynamic process in which multicomponent cellular mixtures form coexisting liquid phases to minimize free energy in a macromolecule concentration-dependent manner (Espinosa et al., 2020). LLPS results in the formation of condensates or membraneless assemblies of proteins and/or nucleic acids that regulates a wide range of cellular functions (Gomes and Shorter, 2019). Extensive studies revealed that intrinsic disorder region (IDRs), with high multivalency, play key roles in LLPS by mediating protein-protein or protein-nucleic acids interaction. These interactions include electrostatic interactions, cation–π, π–π and hydrogen-bonding interactions. Moreover, IDRs are more accessible to PTMs providing an additional level of regulation of LLPS (Dignon et al., 2020).

IDRs are commonly enriched in charged amino acids which promote LLPS through, for example, cationic and anionic amino acids interaction (Nott et al., 2015). The charge can be modified by PTMs such as phosphorylation (Wegmann et al., 2018) or acetylation (Saito et al., 2019), and the polarity of the environment (Kroschwald et al., 2018). Cation-π interactions are important in phase separation, driven by the interaction between aromatic rings (e.g tyrosine, phenylalanine) and positively charged amino acids (e.g., arginine). In addition, hydrogen bonds are prevalent in the densely packed proteinaceous condensates. A study of FUS LLPS using nuclear magnetic resonance (NMR) and Raman spectroscopy showed that glutamine residues have a high propensity for forming hydrogen bond interaction and modulating partitioning (Murthy et al., 2019). Notably, hydrogen bonding is also a major factor for nucleic acid interaction, and incorporating RNA and DNA into condensates (Zagrovic et al., 2018). Hydrophobic effects,

which refer to the tendency of nonpolar substances to aggregate in an aqueous solution, stabilize the protein elastin aggregation status but do not entail the formation of aggregation core (Rauscher and Pomes, 2017).

The purified RGG/RG motif was shown to be necessary and sufficient for phase separation of proteins *in vitro* (Chong et al., 2018). Because RGG/RG motifs have a high affinity to nucleic acid, the *in vivo* phase separation mediated by RGG/RG motifs possibly involves RNA or DNA. Arginine, as discussed above, can contribute to LLPS by forming electrostatic interactions, π–π interactions, cation–π interactions and hydrophobic interactions.

PTM crosstalk

Eukaryotic proteins are modified by different types of PTMs, and emerging evidence showed that PTM crosstalk is common. PRMTs introduced methylation influences other nearby PTMs, such as phosphorylation and ubiquitination. *Saccharomyces cerevisiae* protein Nop1p has several repeats of SRGG sequence within IDR (Smith et al., 2020). Mass spectrometry analysis showed that the arginine could be methylated by yeast arginine methyltransferase Hmt1p, and the adjacent serine can be phosphorylated by Sky1p. The phosphorylation in the SRGG repeats blocks Hmt1p binding, reduces arginine methylation and further affects the nuclear localization of the Nop1p (Smith et al., 2020). Krüppel-like factor 4 (KLF4), a transcription factor, harbors an RGRR site methylated by PRMT5. This methylation alters the affinity of KLF4 for the E3 ligase VHL, resulting in the reduction of KLF4 ubiquitination and enhancing transcription of KLF4 target genes (Hu et al., 2015). The crosstalk of histone PTM is more sophisticated and is further discussed in section 1.4.2.

1.2.2 Function of RGG/RG motif-containing proteins and arginine methylation

Proteins containing RGG/RG motifs are implicated in excessive cellular processes, including transcription, pre-mRNA splicing, mRNA export and phase separation, affecting biological processes such as synaptic plasticity, the DNA damage response, germ-line development, and nuclear-cytoplasmic shuttling (Table 1.1). We will discuss below the function of RGG/RG motif-containing proteins and their arginine methylation by PRMTs.

Post-transcriptional modification

More than a hundred different post-transcriptional modifications are identified among different RNA species (Nachtergaele and He, 2017). Abundant mRNA modification includes N^6-methyladenosine (m^6A), 5-methylcytidine (m^5C), and 2′O methylation (2′OMe). Modification of mRNA affects its secondary structure, protein binding, splicing, and translation.

m^6A modification is found in mRNA and long noncoding RNAs (lncRNA). Modification of m^6A is carried out by m(6)A methyltransferase METTL3/METTL14 heterodimer. METTL14 harbors a C-terminal RGG/RG motif. A recent study showed that this RGG/RG motif is methylated by PRMT1 (Wang et al., 2021c). The methylation of METTL14 C-terminal RGG/RG motif promotes its interaction with RNA, regulates its RNA methylation activity, and is required for its interaction with RNA polymerase II (RNAPII). Moreover, expression of the METTL14 (R to K) mutant in mouse embryonic stem cells (mESCs) reduced global m^6A level, sensitizing mESCs to DNA damage (Wang et al., 2021c). These findings provide the first evidence of protein arginine methylation-mediated RNA methylation.

Pre-mRNA splicing

Pre-mRNA splicing is orchestrated by a coordinated effort of the splicing factors (See 1.3). A large number of splicing factors harbor RGG/RG motifs methylated by PRMTs. Depletion of

PRMT1 or PRMT5 disrupts the splicing pattern in cells, especially cancer cells, leading to dysregulation of pre-mRNA splicing and subsequent cell death. Conditional depletion PRMT5 in mouse neural progenitor cells leads to postnatal death, resulting from aberrant pre-mRNA alternative splicing, including *Mdm4*, an E3 ubiquitin ligase of p53. This altered splicing event causes downregulation of MDM4 protein, leading to p53 pathway activation and excessive cell death (Bezzi et al., 2013). The MDM4-p53 pathway is also regulated by PRMT1. Epicardial-specific deletion of PRMT1 alters the splicing of *Mdm4*, leading to the accumulation of p53 that blocks epithelial–mesenchymal transition (EMT) and disrupts the formation of epicardial-derived lineages (Jackson-Weaver et al., 2020). Studies using type I PRMT or PRMT5 inhibitors induced splicing dysregulation in cancer cells, providing promising cancer therapeutical strategy (Fong et al., 2019; Sachamitr et al., 2021).

PRMTs regulate alternative splicing through the methylation of splicing factors. Early studies identified a role for PRMT5 in the biogenesis of snRNPs (See 1.3) (Brahms et al., 2001; Meister et al., 2001). Besides, several other RGG/RG motifs containing splicing factors were identified as PRMT5 substrates for splicing regulation. PRMT5 methylome profiling uncovers RGG/RG motif of serine/arginine-rich splicing factor 1 (SRSF1) was methylated. Depletion of PRMT5 attenuates SRSF1 binding with RNA (Radzisheuskaya et al., 2019). Zinc Finger Protein 326 (ZNF326) was identified as another substrate of PRMT5. Loss of ZNF326, like PRMT5 depletion, causes the inclusion of A-T-rich exons in target genes (Rengasamy et al., 2017).

The molecular mechanism of how PRMT5 regulates alternative splicing remains to be understood. A large number of splicing factors harbor RGG/RG motif, whether PRMTs regulate splicing events through a particular splicing factor or by affecting the microenvironment of the splicing process, i.e., LLPS, remain elusive.

DNA damage response

Mammalian cells have two major double-stranded breaks (DSB) repair pathways: homologous recombination (HR), which requires a homologous DNA template, thereby occurring exclusively in the S and G_2 phases, and non-homologous DNA end joining (NHEJ), which repair DSB with up to 4 bp of 'microhomology' can take place throughout the cell cycle (Scully et al., 2019). The DSB repair pathway is a sophisticated process requiring orchestrating numerous protein complexes and PTM regulations.

Meiotic recombination 11 (MRE11), complexing with RAD51 and Nijmegen breakage syndrome 1 (NBS1), forming MRN complex, is one of the first sensors recruited to the DSB site after inducing DNA damage, and triggers the HR pathway (Bian et al., 2019). Proteomic analysis using methylarginine-specific antibodies identified that the RGG/RG motif of MRE11 is asymmetrically methylated (Boisvert et al., 2003). A later study showed this methylation is carried out by PRMT1. MEFs with the expression of the MRE11 (R to K) mutant exhibited rad3-related protein (ATR) and checkpoint kinase 1 (CHK1) signaling defects and impairment in the recruitment of replication protein A (RPA) and RAD51 to the damaged sites, resulting from exonuclease and DNA-binding defects of M(RK)RN complex. Moreover, the Mre11 (R to K) mice are hypersensitive to γ-irradiation (IR), and the cells from these mice display cell cycle checkpoint defects and chromosome instability (Yu et al., 2012).

Tumor suppressor p53-binding protein 1(53BP1) plays an essential role in the repair pathway by promoting NHEJ-mediated DSB repair and preventing HR (Bouwman et al., 2010). 53BP1 contains RGG/RG motif in the kinetochore-binding domain, which is methylated by PRMT1. Fibroblast cells treated with methyltransferase inhibitors form fewer γ-H2AX foci and fail to relocalize 53BP1 to the DNA damage sites. 53BP1 (R to A) mutant is unable to bind to single

and double-stranded DNA, suggesting the RGG/RG motif is required for 53BP1 DNA binding activity (Boisvert et al., 2005).

The methylation of MRE11 and 53BP1 is enhanced in T cells by the transcription factor growth factor independence 1 (GFI1). GFI1 facilitates PRMT1 interacting with and methylating MRE11 and 53BP1. In addition, depletion of GFI1 in T cells decreases asymmetrical methylation of MRE11 and 53BP1 and sensitizes cells to ionizing radiation (Vadnais et al., 2018).

PRMT5 regulates HR by targeting TIP60, a histone acetyltransferase, for methylation. The methylated TIP60 facilitates mobilization of 53BP1 from DNA breaks to promote HR (Clarke et al., 2017). Another study showed that depletion of PRMT5 or treatment of PRMT5 inhibitor alters the splicing pattern of TIP60, which selectively affects TIP60 lysine acetyltransferase activity and leads to impaired HR (Hamard et al., 2018). In addition, PRMT5 also directly methylates 53BP1. Depleting PRMT5 or treatment with PRMT5 inhibitor reduces 53BP1 protein level by destabilizing 53BP1. However, the underlying mechanism is unknown (Hwang et al., 2020).

R-loop formation

R-loop is a three-stranded nucleic acid structure composed of an RNA: DNA hybrid and a displaced single-stranded DNA. R-loops occur at the preferred regions throughout the genome, which is an important source of genome instability and hallmarks of cancer (Santos-Pereira and Aguilera, 2015). A recent study from our lab showed that the resolving of R-loops is mediated by DDX5, an RGG/RG motif-containing DEAD-box family RNA helicase (Mersaoui et al., 2019). DDX5 resolves R-loop and associates with XRN2 for R-loop degradation. The interaction between DDX5 and XRN2 depends on the symmetrical methylation of DDX5 RGG/RG motif by PRMT5. Depletion of DDX5 or PRMT5 leads to the accumulation of cellular R-loops (Mersaoui et al., 2019).

Translation

Many RBPs involved in translation regulation are RGG/RG motifs containing protein.

eIF4G1, a key regulator of translation initiation, harbors an RGG/RG motif methylated by PRMT1. This methylation contributes to eIF4G1 protein stability and subsequently regulates translation initiation complex assembly and activity (Hsu et al., 2017). In another example, PRMT1 catalyzes methylation of DDX3, leading to its binding to a non-AUG upstream open reading frame (uORF) in the 5'UTR of γ-globin mRNA, and subsequent suppression of fetal γ-globin translation. Expression of DDX3 (R to K) mutant or treatment with DDX3 inhibitors, similar to PRMT1 depletion, bypass the μORF and causes γ-globin mRNA translation increase (Wang et al., 2021b). Aven regulates the translation of G-quadruplex RNA by increasing the polysome association. RGG/RG motif of Aven mediates the interaction with both G-quadruplex RNA and polysome, and its methylation by PRMT1 is required for the association with polysomes and RNA helicase DHX36. Depletion of PRMT1 or DHX36 blocks Aven mediated translation of G-quadruplex RNA. (Thandapani et al., 2015).

PRMT5 regulates the translation of a subset of internal ribosome entry site (IRES)-containing genes. hnRNP A1 is an IRES transacting factor (ITAF), and its methylation facilitates the interaction of hnRNP A1 with IRES RNA to promote IRES-dependent translation (Gao et al., 2017).

Membraneless organelles

LLPS is the physical basis of the membraneless organelles formed in cytoplasm and nuclear. These membraneless organelles include stress granules (SGs), P-bodies, germ granules, nucleoli, Cajal bodies, nuclear speckles, paraspeckles, histone–locus bodies, nuclear gems, and

promyelocytic leukemia (PML) 2 bodies (Gomes and Shorter, 2019). Many proteins implicated in these membraneless organelles are RGG/RG motif-containing proteins.

FUS is a well-characterized protein for phase separation (Patel et al., 2015). FUS mutation implicated in subtypes of Amyotrophic lateral sclerosis (ALS) tends to form toxic liquid-liquid demixing in the cytoplasm (Kwiatkowski et al., 2009; Vance et al., 2009). FUS LLPS is driven by multivalent cation-π interactions that occur physiologically between multiple arginines in the C-terminal RGG/RG motifs and multiple tyrosines in its N terminus (Qamar et al., 2018). Moreover, arginine methylation, which affects the cation-π interactions, reduces FUS LLPS and increases droplet dynamics (Hofweber et al., 2018; Qamar et al., 2018).

SGs are large clusters of proteins and RNA in the cytoplasm that form under stress, such as high heat or under starvation. A comprehensive study demonstrated the stress granule assembly through phase separation (Yang et al., 2020). G3BP1 is the core protein of the stress granule that triggers the RNA-dependent LLPS. Three IDRs contribute to the G3BP1 LLPs. In particular, the positive charged IDR3 with RGG/RG motif interacts with the negative charged IDR1 and increasing RNA concentration permits G3BP1 homodimerization and initiating LLPS. Within IDR3, substitution of arginine to alanine, but not lysine disrupts the LLPS formation indicating the positive charge is essential for IDR3 function (Yang et al., 2020). In another study, hypomethylation of G3BP1 promotes stress granule formation (Tsai et al., 2016).

Lots of RBPs have RNA binding domains termed RRM (RNA Recognition Motif) domains and IDRs with RGG/RG motifs: a structure prone to form LLPS. There are many studies that show the involvement of RBPs in phase separation and the role of RGG/RG motifs in the assembly of these membraneless organelles.

Table 1. 1 Known substrates of PRMT1 and PRMT5

Substrate	Function*	Reference/PMID	Enzyme
Fibrillarin	component of snRNP, participate in processing rRNA	*4084504*	PRMT1
Nucleolin	major nucleolar protein	*4084504*	PRMT1
hnRNP A1, B, D, E, G, H, J, K, P, Q, R, U	pre-mRNA splicing, transport of mRNA	*7739561*	PRMT1
TIS21	participate in mRNA deadenylation	*9739451*	PRMT1
hnRNP A2	associates with nascent pre-mRNAs, packaging them into hnRNP particles, involved in pre-mRNA processing, RNA nuclear export, subcellular location, mRNA translation and stability of mature mRNAs	*10772824*	PRMT1
EWS	transcriptional repressor	*20211855*	PRMT1
FUS	transcription regulation, RNA splicing, RNA transport, DNA repair and damage response	*12183049*	PRMT1
TAFII-68	part of the TET family of proteins that can bind RNA and single-stranded DNA	*12183049*	PRMT1
CAPRIN1	regulate the transport and translation of mRNAs of proteins	*12183049*	PRMT1
Sam68	regulation of alternative splicing and influences mRNA splice site selection and exon inclusion	*12529443*	PRMT1
QKI-5	pre-mRNA splicing, mRNA export and protein translation	*12529443*	PRMT1
SLM-1	regulation of alternative splicing and influences mRNA splice site selection and exon inclusion	*12529443*	PRMT1
SLM-2	regulation of alternative splicing and influences mRNA splice site selection and exon inclusion	*12529443*	PRMT1
FGF-2	a ligand for FGF receptor, plays an important role in the regulation of cell survival, cell division, cell differentiation and cell migration	*10652299*	PRMT1
TOP3B	releases the supercoiling and torsional tension of DNA introduced during the DNA replication and transcription by transiently cleaving and rejoining one strand of the DNA duplex	*29471495*	PRMT1
ERa	nuclear hormone receptor	*18657504*	PRMT1
EZH2	histone methyltransferase and EMT inducer	*33092789, 32895488*	PRMT1
PR	transcription factor that affects cellular proliferation and differentiation in target tissues	*32563156*	PRMT1
MRE11	component of the MRN complex, which plays a central role in DSB repair, DNA recombination, maintenance of telomere integrity and meiosis	*21826105*	PRMT1

53BP1	DSB repair protein involved in response to DNA damage, telomere dynamics and class-switch recombination during antibody genesis	*16294045*	PRMT1
SCYL1	maintenance of the Golgi apparatus morphology	*32583741*	PRMT1
SPT5	component of the DRB sensitivity-inducing factor complex, which regulates mRNA processing and transcription elongation by RNA polymerase II	*12718890*	PRMT1, PRMT5
DHX9	unwinds DNA and RNA in a 3' to 5' direction and that plays important roles in many processes, such as DNA replication, transcriptional activation, post-transcriptional RNA regulation, mRNA translation and RNA-mediated gene silencing	*15084609*	PRMT1
METTL14	N6-methyltransferase complex that methylates adenosine residues at the N6 position of some mRNAs and regulates the circadian clock, differentiation of embryonic stem cells and cortical neurogenesis	*33459381*	PRMT1
eIF4G1	component of the protein complex eIF4F, which is involved in the recognition of the mRNA cap, ATP-dependent unwinding of 5'-terminal secondary structure and recruitment of mRNA to the ribosome.	*28655788*	PRMT1
Aven	interacts with G-quadruplexes RNA	*26267306*	PRMT1
SRSF1	plays a role in preventing exon skipping, ensuring the accuracy of splicing and regulating alternative splicing	*20308322, 31611688*	PRMT1, PRMT5
G3BP1	stress granule formation, unwinds DNA/DNA, RNA/DNA, and RNA/RNA substrates with comparable efficiency	*12642610, 9889278*	PRMT1, PRMT5
Npl3p	component of the RNA exosome complex which has 3'->5' exoribonuclease activity and participates in a multitude of cellular RNA processing and degradation events	*8942987*	HMT1
Hrp1p	involved in polyadenylation-dependent pre-mRNA 3'-end formation and cooperates with the cleavage factor CFIA complex and the cleavage and polyadenylation factor complex.	*9499403*	HMT1
Nab2p	binds to polyadenylated RNA and single-stranded DNA	*11779864*	HMT1
Gar1p	catalytic subunit of H/ACA small nucleolar ribonucleoprotein complex, which catalyzes pseudouridylation of rRNA	*12756332*	HMT1
Yra1	pre-mRNA metabolism and component the TREX complex, which operates in coupling transcription elongation to mRNA export.	*15314027*	HMT1

Sm B, B', D1, D3	plays a role in pre-mRNA splicing as a core component of the spliceosomal U1, U2, U4 and U5 ssnRNPs, the building blocks of the spliceosome	*10747894*	PRMT5
LSM4	Plays role in pre-mRNA splicing as component of the U4/U6-U5 tri-snRNP complex that is involved in spliceosome assembly	*11720283*	PRMT5
p80 Coilin	modification and assembly of nucleoplasmic snRNPs	*12486110*	PRMT5
ZNF326	component of the DBIRD complex, a multiprotein complex that acts at the interface between core mRNP particles and RNAPII and integrates transcript elongation with the regulation of alternative splicing	*28977470*	PRMT5
DDX5	pre-mRNA splicing, R-loops resolving	*31267554*	PRMT5
PIWI proteins	piRNA metabolic process	*19377467*	PRMT5
KLF4	transcription factor; can act both as activator and as repressor	*26420673*	PRMT5
KLF5	Transcription factor that binds to GC box promoter elements. Activates the transcription of these genes.	*33972717*	PRMT5
TIP60	Catalytic subunit of the NuA4 histone acetyltransferase complex which is involved in transcriptional activation of select genes principally by acetylation of nucleosomal histones H4 and H2A	*28238654*	PRMT5
cGAS	senses the cytoplasmic double-stranded DNA and activates the antiviral response	*33762328*	PRMT5

** Function summarized by Uniprot database.*

1.2.3 RBMX

RBMX, also known as hnRNP G, is a ubiquitously expressed nuclear protein encoded by a gene on the X chromosome. A homolog is found on the mammalian Y chromosome called RBMY (Ma et al., 1993). Unlike RBMX, RBMY is only expressed in germ cells. Several RBMX pseudogenes were also reported, including RBMXL1, an RBMX retrogene on chromosome 1. The protein sequence of hRBMXL1 is largely identical to hRBMX except for the C-terminal RGG/RG motif. In contrast, the sequence of RBMY varies greatly compared to RBMX.

RBMX contains two RGG/RG motifs and an RRM domain. The N-terminal 1-88 amino acids form a folded RRM domain, while the rest of the 303 amino acid forms a non-structured long IDR (RGG/RG motif: 109-126aa and 372-378aa). A 23bp deletion in the last exon of the *RBMX*

gene has been identified in a family of seven males with a subtype of X-linked intellectual disability called Shashi syndrome (Shashi et al., 2015). This deletion mutation was predicted to generate a truncated RBMX protein with 38 aa deletions at its C-terminus, including the C-terminal RGG/RG motif.

RBMX is a multifunctional protein that has diverse effects on mRNA splicing, DNA damage response, and heterochromatin maintenance. The first characterized function of RBMX is splicing regulation. In an early study, overexpression of RBMX regulates several alternative splicing events. Minigene assay showed that both SMN2 and Tau splicing events depend on RBMX protein concentration, but in an RRM motif-independent manner. Moreover, the RRM motif of RBMX selectively binds to CCA/CCC-rich RNA oligonucleotide by Systematic evolution of ligands by exponential enrichment (SELEX) (Heinrich et al., 2009). Another study elucidated the mechanism of RBMX mediated *SMN2* exon 7 inclusion. They showed that, in addition to the C-terminal RGG/RG motif, both RRM domain and the middle RGG/RG motif contribute to the alternative splicing regulation. Using NMR, they demonstrated that the RRM domain interacts with a 5′-AAN-3′ motif, and RBMX recruits Tra2-β1 and SRSF9 to regulate the *SMN2* exon 7 inclusion (Moursy et al., 2014). Recently, RBMX was reported to interact with m^6A modified mRNA. m^6A modification increases the accessibility of its surrounding RNA sequence and facilitates RBMX interaction (Liu et al., 2017). The follow-up study showed that RBMX interacts with RNA polymerase II and regulates co-transcriptional splicing. m^6A modification near splice sites in nascent pre-mRNA midiates RBMX binding and promotes exon inclusion (Zhou et al., 2019). These studies suggest that the C-terminal RGG/RG motif is essential for RBMX function, and the IDR plays an important role in its interaction with other proteins.

RBMX plays an important role in the DNA damage response. A genome-wide HR screen identified RBMX as a component of the DNA-damage response (Adamson et al., 2012). RBMX accumulated at DNA lesions through multiple domains in a poly (ADP-ribose) polymerase 1-dependent manner. Knockdown of RBMX reduced expression of BRCA2, contributing to the HR deficiency (Adamson et al., 2012). A recent study showed that RBMX regulates genomic stability by interacting with lncRNA *NORAD* and assembling an RNP complex incorporating suppressors of genomic instability topoisomerase I (TOP1), ALYREF and the PRPF19-CDC5L complex (Munschauer et al., 2018). However, the study was challenged as the expression of *NORAD* mutant with deletion of RBMX binding sites rescue the aneuploidy phenotype in *NORAD*$^{-/-}$ HCT116 cells, suggesting RBMX binding is not required for regulation of genomic stability by *NORAD* (Elguindy et al., 2019).

RBMX may play an important role in chromatin maintenance by interacting with DNA. Proteomics study using sonication resistant chromatin fraction identified RBMX and RBMXL1 as regulators for heterochromatin. Moreover, RBMX depletion reduces global H3K9me3 level and elicits sonication sensitivity at the target heterochromatic genes, similar to the effect with SUV39H1 depletion (Becker et al., 2017). The study provided insight into modulation of histone modification by RBMX, although the molecular mechanism was unknown. Notably, a study using gel shift assay showed that RBMX and its RRM domain are directly bound to DNA *in vitro* (Shin et al., 2007).

The earliest phenotypic study of RBMX used *Danio rerio* or zebrafish (Tsend-Ayush et al., 2005). *rbmx* is maternally expressed in zebrafish embryos and then widely expressed in up to 24 hours post-fertilization. During development, *rbmx* transcripts localized predominantly in the brain, branchial arches, and liver primordium. Depletion of *rbmx* using antisense morpholino

leads to underdeveloped head and eyes, reduced body size, defective somite patterning, and absence of jaws. Particularly, the expression of specific markers for the forebrain and hindbrain (*otx2, krox20*) was severely reduced (Tsend-Ayush et al., 2005). Similarly, depletion of rbmx in *Xenopus laevis* causes defects in anterior neural patterning, neural differentiation, and axial muscle segmentation (Dichmann et al., 2008). Interestingly, *rbmx* seems to be a nonessential gene in mice (Dickinson et al., 2016). Proteomic analyses of the postsynaptic density (PSD) components of cortical neurons identified several hnRNPs at synapses. Depletion of RBMX using shRNAs causes altered numbers of dendritic spines, suggesting that RBMX may play an essential role in spine density (Zhang et al., 2012). These phenotypic studies suggest that RBMX may be involved in brain development.

Besides functioning in development, recent studies also explored the link between RBMX and carcinogenesis. A low expression level of RBMX is significantly associated with a favorable overall survival rate among patients with head and neck cancer (Guo et al., 2020). In bladder cancer, RBMX inhibits cell proliferation, colony formation, migration, and invasion *in vitro* and suppresses tumor growth and metastasis *in vivo.* The underlying mechanism showed that RBMX competes with hnRNP A1 to regulate pyruvate kinase muscle isozyme (PKM) alternative splicing, leading to lower PKM2 and higher PKM1 protein levels attenuated the tumorigenicity and progression of bladder cancer (Yan et al., 2021).

1.2.4 SARS-CoV-2 Nucleocapsid protein

Coronaviruses (CoVs) are RNA viruses that cause diseases in human and animal hosts. CoVs can be classified into four genera, namely *Alphacoronaviruses, Betacoronaviruses, Gammacoronaviruses and Deltacoronaviruses*. (Lefkowitz et al., 2018). Seven strains of human CoVs are known (King, 2020), including three β-CoVs, identified in the past two decades,

produce severe symptoms: severe acute respiratory syndrome (SARS)-CoV-1, middle east respiratory syndrome (MERS)-CoV, and SARS-CoV-2.

The COVID-19 pandemic is caused by SARS-CoV-2, a virus that belongs to the family *coronaviridae* of genus *betacoronavirus* and has a positive sense strand RNA genome of ~30 kb (Kim et al., 2020a). Same as other β-CoV, it contains two large overlapping open reading frames (ORF1a and ORF1b) and encodes four structural proteins, namely spike (S), envelope (E), membrane (M) and nucleocapsid (N) proteins as well as nine accessory proteins (Kim et al., 2020a). ORF1a and ORF1b are further processed to generate 16 non-structural proteins (Nsp1-16). N protein is the most abundant in the virions among the viral proteins and is expressed at the highest levels in infected cells (McBride et al., 2014). Thus, its abundance, essential roles during infection, and immunogenic nature, make the SARS-CoV-2 N protein a valuable target for developing new strategies to combat the COVID-19 pandemic (Gordon et al., 2020; Kang et al., 2020; Ray et al., 2020).

SARS-CoV-2 N protein has two conserved and independently folded structural domains, called the N-terminal domain (NTD) and C-terminal domain (CTD) (Kang et al., 2020; Lin et al., 2020; Ye et al., 2020). The two domains are separated by IDRs at the N-terminal arm, central serine/arginine-rich (SR-rich) linker region (LKR) and C-terminal tail. NTD is responsible for RNA binding and CTD for both RNA binding and dimerization (Dinesh et al., 2020; Iserman et al., 2020; Luo et al., 2006). The IDRs are involved in modulating the RNA binding activity (Dinesh et al., 2020). So far, only the structure of NTD has been solved. SARS-CoV-2 N-NTD crystal structure shares similar right-handed (loops)-(β-sheet core)-(loops) sandwiched fold, as conserved among the CoVs N-NTD (Kang et al., 2020). The β-sheet core consists of five antiparallel β-strands with a single short helix just before strand β2, and a protruding β-hairpin

between strands β2 and β5 (Figure 1.2). The β-hairpin is functionally essential for coronavirus infectious bronchitis virus (IBV) N-NTD, implicated in mutational analysis of amino acid residues for RNA binding (Tan et al., 2006). The SARS-CoV-2 N-NTD is enriched in aromatic and basic residues, folding into a right-hand shape, which resembles a protruded basic finger, a basic palm, and an acidic wrist (Figure 1.2). NMR-based titration experiment using 10 mer ssRNA showed that the significantly perturbed residues (L56, G60, K61, K65, F66, A90, R93, I94, R95, K102, D103, L104, T165, T166, G175 and R177) formed a U-shaped binding epitope on the N-NTD surface circumventing the base of the positively charged finger (Dinesh et al., 2020). Mutational analysis of the N-NTD RNA interaction showed R92 and R107 are essential to forming hydrogen bonds network with RNA.

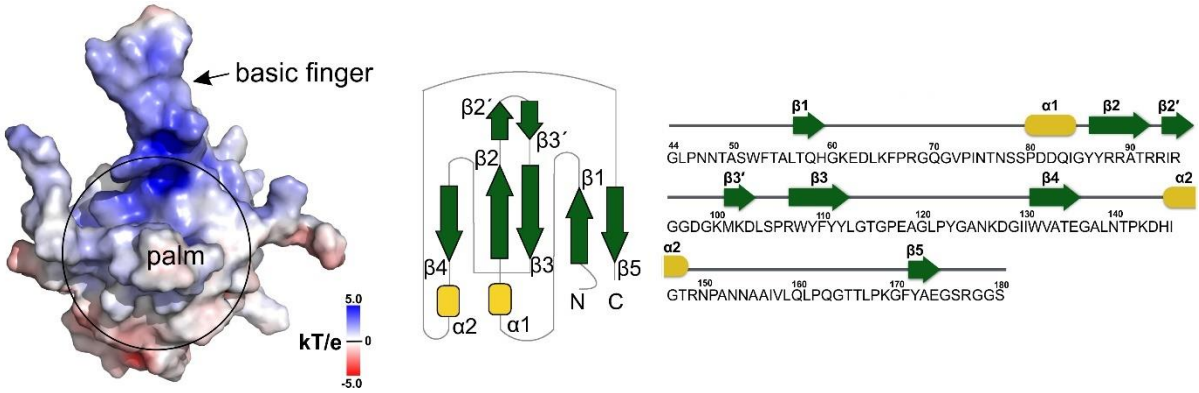


Figure 1. 2 structure of the SARS-CoV-2 N-NTD RNA binding domain (Dinesh et al., 2020).

N protein is involved in several aspects of the virus life cycle (Nikolakaki and Giannakouros, 2020). It binds to viral gRNA to package the genome into the viral ribonucleoprotein (vRNP) complex and mediate vRNP packaging into virions (Chang et al., 2013; Cong et al., 2020). Besides assembly, N protein facilitates viral mRNA transcription, replication, and translation by recruiting host factors. N protein is also involved in host-pathogen interactions by regulating apoptosis, the progression of the host cell cycle, and the stress response (Chen et al., 2020b;

Nakagawa et al., 2018). An important function of NTD is to melt dsRNA, which has been shown to play a role in RNA packaging or other steps in the viral life cycle where RNA remodeling is required. For example, N protein from mouse hepatitis virus (MHV), a virus most closely related to SARS-CoV, binds the transcriptional regulatory sequence (TRS) and melts TRS-cTRS (TRS and its complement) RNA duplexes (Grossoehme et al., 2009). The CTD domain of SARS-CoV-1 has been resolved, which consists of five short α-helices, one 310 helices, and two β-strands, resembling the letter C (Yu et al., 2006). The CTD domain plays an important role in N protein dimerization and oligomerization and is also involved in binding with putative packing signals within the viral RNA (Hsieh et al., 2005).

The packing process is regulated by the interaction of N protein and RNA through phase separation (Carlson et al., 2020; Chen et al., 2020a; Iserman et al., 2020; Lu et al., 2021; Perdikari et al., 2020; Savastano et al., 2020; Wu et al., 2021). The proposed model is that phase separation drives large, micron-sized cytoplasmic condensates of N protein and RNA, which may act as molecular factories that concentrate the components for pre-capsid assembly (Cubuk et al., 2021). A ~40 residue region in the central IDR is essential for RNA-driven phase separation. Quantitative cross-linking mass spectrometry revealed that both the IDR and the CTD show strong intermolecular interactions in RNA-driven condensates (Lu et al., 2021). The central IDR is a serin and arginine (SR) rich region, which can be phosphorylated by GSK-3, a combination of GSK-3 and Cdk1and SRPK1 (Carlson et al., 2020). Phosphorylation of the SR-rich region favors dynamic, liquid-like condensates, while unmodified N protein forms gel-like condensates (Carlson et al., 2020). The computational model suggested phase separation gives rise to a single large spherical cluster with multiple polymers (RNAs) and binders (proteins), rather than multi-genome condensates, which would reflect packaging signals in viral genomes (Cubuk et al., 2021).

The viral assembly process also requires N protein interaction with other viral proteins. Mixtures of N + M + RNA form annular structures where the M protein coats the outside of an N + RNA condensate (Lu et al., 2021). It is worth noting that SARS-CoV-2 gRNA forms a secondary structure and interacts with many host RBPs (Flynn et al., 2021; Schmidt et al., 2021). Depletion of host RBPs such as cellular nucleic acid-binding protein (CNBP) and La-related protein 1 (LARP1) restrict SARS-CoV-2 replication in infected cells (Schmidt et al., 2021). Further studies are required to understand the interplay between host proteins and viral proteins during the virus life cycle.

The interactome of N protein identified several host SG proteins suggests that N protein can function independently of virus gRNA to regulate host stress response during infection (Gordon et al., 2020). Several studies revealed that N protein interacts with G3BP1/2, core components of SG, suppresses SGs formation (Luo et al., 2021; Nabeel-Shah et al., 2020). Overexpression of G3BP1 in N-expressed cells can restore the SGs formation. Moreover, N protein binds to 3'UTR of mRNA and regulates gene expression (Nabeel-Shah et al., 2020). However, whether RNA was required for N protein and G3BP1 interaction needs to be further investigated. In fact, SGs are regulatory components that respond to virus infection, which generally are detrimental to virus replication. Therefore, viruses from diverse families modulate SGs in infected cells by interaction with SG effecter proteins. For example, during early poliovirus infection, SG is induced, but at the later stage, the cells lose the ability to form SGs. This is because the poliovirus express 3C proteinase cleaves G3BP and favors virus replication. Expression of a cleavage-resistant G3BP can rescue the phenotype (White et al., 2007). Hepatitis C virus (HCV) RNA-dependent RNA polymerase Nsp5 interacts with G3BP, and knockdown of G3BP in cells facilitated the virus production (Yi et al., 2006).

Because the N protein of CoVs has a conserved structure, studies with these N proteins would provide insights into the unexplored functions of SARS-CoV-2 N protein. Hereafter, we introduce the previous studies of other CoV N proteins. SARS-CoV-1 and transmissible gastroenteritis virus (TGEV) N protein were shown to be RNA chaperones, enhancing hammerhead ribozyme self-cleavage and nucleic acid annealing (Nabeel-Shah et al., 2020). N protein also plays a role in regulating cellular translation, which N protein was found to interact with human elongation factor 1-alpha (EF1α), an essential component of the translational machinery. The interaction causes aggregation of EF1α inhibiting protein translation (Zhou et al., 2008). Another important function of N protein is to regulate cellular innate immune response (Hu et al., 2017). N protein inhibits type I interferon (IFN) production via interaction with tripartite motif protein 25 (TRIM25) E3 ubiquitin ligase. This regulation is through RLR pathway, where the viral RNA is recognized by RIG-I and MDA5. N protein interferes TRIM25 ubiquitination and activation of RIG-I, thus, suppressing the interferon pathway (Hu et al., 2017).

Different types of vaccines were developed to against SARS-CoV-2. The vaccines with high efficacy against SARS-CoV-2 variants are mRNA-based vaccines provided by Pfizer-BioNTech and Moderna (Jackson et al., 2020). These vaccines introduce mRNA encoding SARS-CoV-2 S(2P) antigen (Corbett et al., 2020), a stabilized S protein variant, stimulating the immune responses in the recipient, and preventing the virus infection. Although the vaccines provide strong protection, the medical treatment for patients with SARS-CoV-2 is still needed. An FDA-approved kinase inhibitor, Alectinib, which prevents N protein phosphorylation by SRPK1, was shown to prevent the virus infection in primary type II pneumocyte cell culturing (Yaron et al., 2020). Screening of 12,000 clinical-stage or FDA-approved small molecules identified thirteen of them harbor effective concentrations commensurate with probable achievable therapeutic

doses in patients. These include PIKfyve kinase inhibitor apilimod and several cysteine protease inhibitors (Riva et al., 2020).

1.3 pre-mRNA splicing

pre-mRNA splicing is a common process in eukaryotic organisms, including constitutive splicing and alternative splicing. 95% of multiexon genes undergo alternative splicing, producing at least two splicing isoforms (Pan et al., 2008). A large number of alternative splicing events generate mRNA transcripts with premature stop codons that selectively to be degraded through nonsense-mediated decay pathway. Moreover, the mRNA sequence alteration affects RNA stability and mRNA localization. There are different types of alternative splicing. The most common type of alternative splicing, accounting for about 40% of events, is exon skipping that generates transcripts with one or several cassette exons spliced out. Other types of alternative splicing are alternative 3'- or 5'- splice sites, mutually exclusive exons, intron retention, alternative promoters and altered polyadenylation. Alternative splicing is dynamically regulated in the development process as well as various diseases such as cancers.

1.3.1 Splicing regulation

The constitutive splicing process involves two steps: 1) cleavage of the 5′ exon-intron phosphodiester bond and formation of a new 2′-5′ bond with the branch site resulting in an intron "lariat", and 2) ligation of the two exons. Introns are removed by a large and highly dynamic RNP, named spliceosome, an assembly of five snRNP particles, U1, U2, U4/U6, and U5 (Figure 1.3 A). Both *cis*-acting regulatory sequences and *trans*-acting factors are crucial for splicing. *cis*-acting regulatory sequences determine recognition by splicing factors to set up and maintain the splicing pattern. For instance, the human 5' splice site (5'ss) consensus sequence is MAG|GURAGU (M

is A or C; R is purine), spanning from position −3 to position +6 relative to the exon-intron junction. With a U →C transition at the position +6, the wildtype 5'ss changes to a weak 5'ss that reduces the binding by U1 snRNP (Buratti et al., 2007). Besides the spliceosome recognition site, there are splicing enhancers (SEs) and silencers (SSs) that to be bound by splicing factors for alternative splicing regulation (Figure 1.3 B).

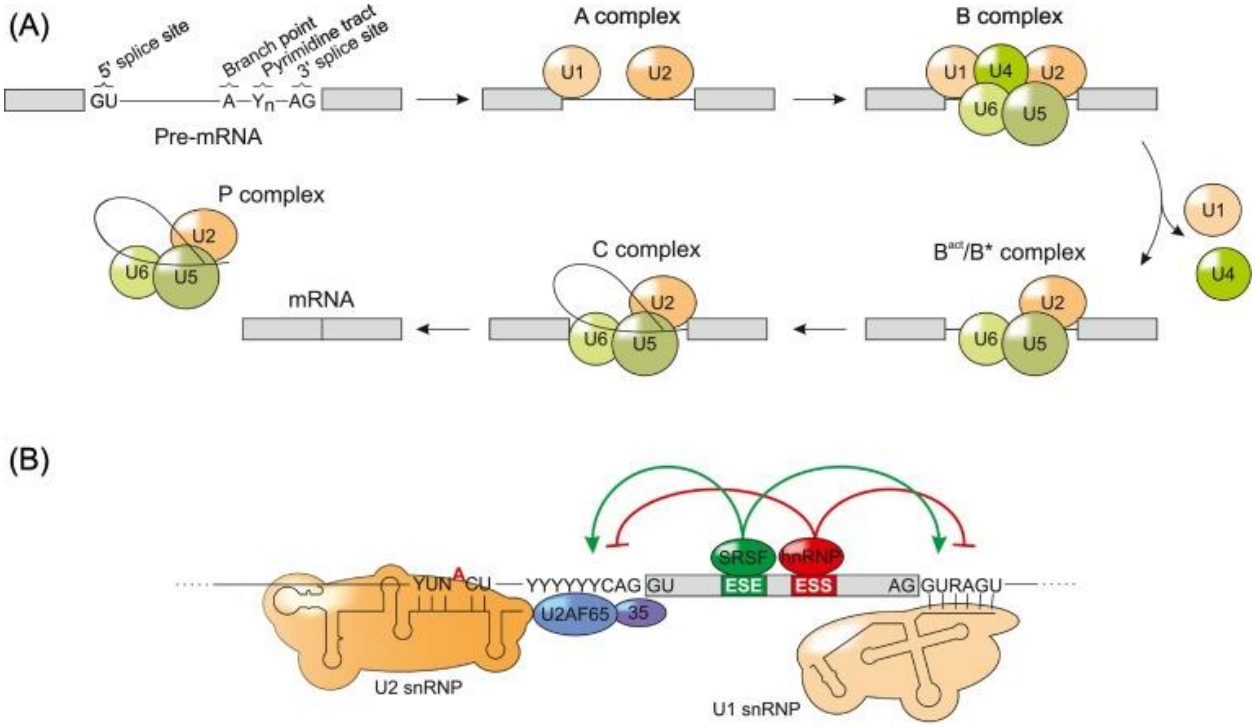


Figure 1. 3 Mechanism of Splicing and Its Regulation (Gehring and Roignant, 2021)**.**

snRNPs consist of specific U-rich snRNAs that base-pairs with the pre-mRNA sequence and Sm core proteins B or B', D1/D2/D3, E, F and G that assemble as a seven-member ring structure. The biogenesis of snRNP starts with snRNA transcription by RNA polymerase II. The pre-snRNAs are then exported into the cytoplasm for further processing, because the Sm core site of these snRNAs (AUUUUUG, AUUUGUG, or similar sequences) are present throughout the transcriptome that are rather promiscuous for Sm protein binding (Pellizzoni et al., 2002; Raker

et al., 1999). The Sm core assembly occurs in the cytoplasm, starting with three Sm heteromeric subcore complexes. SmD1/D2 and SmF/E/G bind to the Sm site of snRNA, which then recruits and are stabilized by SmB/D3 (Raker et al., 1999). snRNAs are further hypermethylated at their 5′-end into trimethylguanosine-cap and trimmed at the 3′ end by exonucleases (Mouaikel et al., 2002). Finally, snRNPs are reimported to the nucleus, mediated by Snurportin-1 and SMN with importin β, where they function in pre-mRNA processing (Nigg et al., 1991). It is worthwhile to emphasize that in addition to snRNPs, DExD/H-box ATPases facilitate structural rearrangement in the spliceosome.

Non-spliceosomal proteins can be classified into three groups: the hnRNP family proteins, SR family proteins and tissue-specific RNA-binding proteins such as RBFOX family proteins and muscleblind/CELF family proteins. These splicing factors bind SEs or SSs to regulate exon inclusion or exclusion. hnRNPs family consists of 20 major types of hnRNP proteins, containing two or more of these protein domains for RNA binding: the RRM, RGG/RG motifs and KH domain (Geuens et al., 2016). The hnRNP I, also known as polypyrimidine tract-binding protein 1 (PTBP1) represses proto-oncogene tyrosine-protein kinase c-Src N1 exon splicing. PTBP1 binds to the polypyrimidine tract, by directly interacting with U1 snRNP at the N1 exon 5'ss, suppressing assembly of spliceosomal components, leading to repression of N1 inclusion (Sharma et al., 2011). Similarly, hnRNP L and hnRNP A1 were reported to repress CD45 exon 4 inclusion (Chiou et al., 2013). However, the mechanism that PTBP1 facilitates exon inclusion has not been completely elucidated. In some cases, PTBP1 binds to splicing enhancers to induce exon inclusion (Hollander et al., 2016). RBMX also promotes SMN exon 7 inclusion by directly interacting with TRA2B, an SR protein (Hofmann and Wirth, 2002).

SR proteins are a family of twelve RBPs with one or two RRM domains followed by serine-arginine repeats. SRSF1 is one of the well-studied SR family proteins that preferentially bind to exonic sequences to promote exon inclusion (Sun et al., 1993). SRSF1 not only recruits as splicing enhancer, but also engages in 5'SS site selection. The phosphorylation of SRSF1 is required for pre-spliceosome to spliceosome transition (Krainer et al., 1990; Roscigno and Garcia-Blanco, 1995). Besides, SRSF9 and SRSF11 were reported to promote exon skipping by binding to splicing silencers (Simard and Chabot, 2002; Wu et al., 2006). Splicing factors also affect alternative splicing by their expression level. Many splicing factors such as SRSF1 are over-expressed in cancer cells and help to form the cancer-specific splicing pattern.

Notably, emerging studies suggest that splicing is regulated by phase separation of splicing factors that harbor intrinsic disorder regions. Indeed, hnRNP family proteins with RRM domain and IDRs prone to form multivalent assemblies with proteins and RNA. Transcripts encoding IDRs are highly regulated by alternative splicing. IDR of hnRNP A and D family proteins are regulated by splicing, and exon skipping of the IDR disrupts multivalent hnRNP assemblies, that in turn, globally regulate splicing pattern (Gueroussov et al., 2017). A recent study proposed a phase separation model that alternative splicing occurs at the interface between SR proteins and hnRNPs, where SR proteins drive and form the core of phase separation and hnRNPs surround the nuclear speckle (Liao and Regev, 2021).

1.3.2 PRMTs and pre-mRNA splicing

Splicing factors are substrates of PRMTs. Arginine methylation modulates the function of these proteins that may contribute to their splicing activity.

It is well-established that PRMT5 regulates snRNP assembly by symmetrically methylating Sm B/B', D1 and D3 (Brahms et al., 2001; Miranda et al., 2004a) (Figure 1.4). The process starts

with Sm protein heterooligomers D1/D2 and D3/B interacting with chaperone protein pICln to form pICln-Sm protein complexes. Then pICln-SmD1/D2 and pICln-SmD3/B are symmetrically methylated by PRMT5/MEP50 complex. And PRMT5/MEP50 complex further recruits SmF/E/G, leading to formation of the 6S comple (Neuenkirchen et al., 2015).

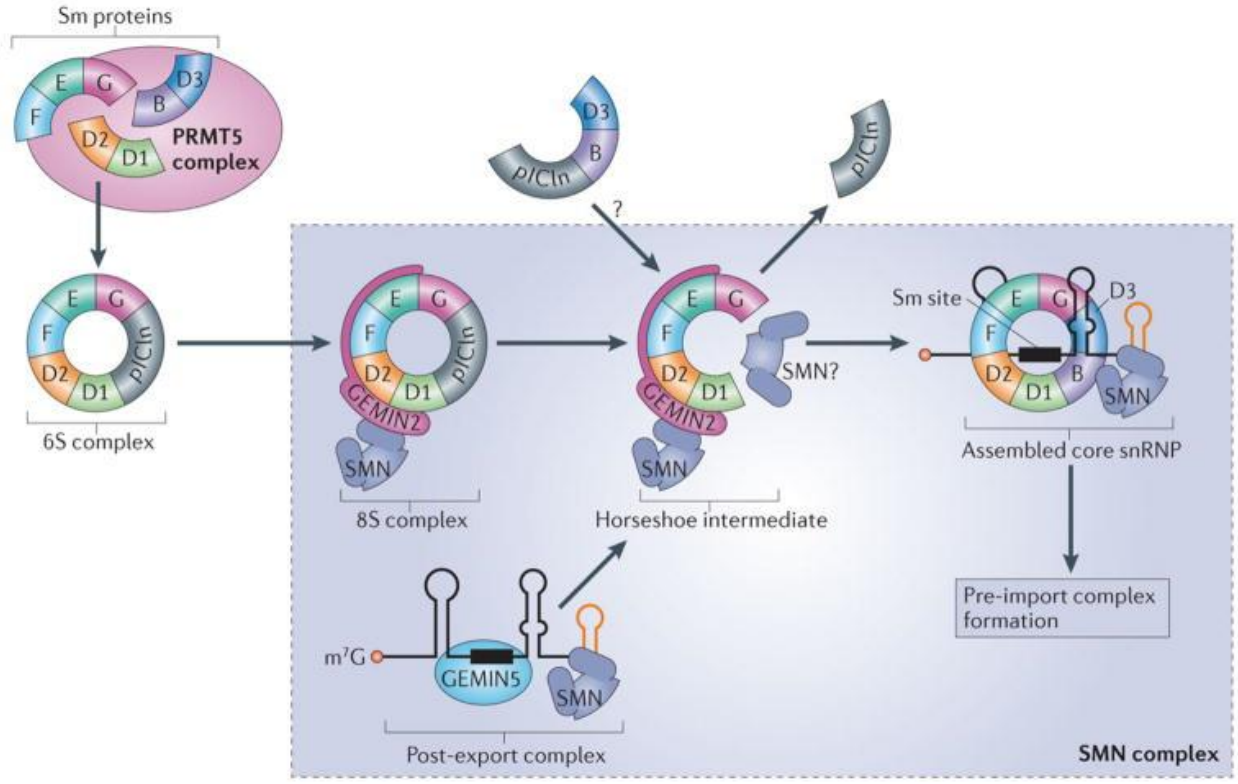


Figure 1. 4 PRMT5 mediates the assembly of human U snRNPs (Matera and Wang, 2014)**.**

hnRNPs are mono- and di- methylated mainly by PRMT1, including hnRNP A, B, D, E, G, H, J, K, P, Q and U (Liu and Dreyfuss, 1995). Arginine methylation affects hnRNPs function in different ways. Arginine methylation of RGG/RG motif in hnRNP Q3 affects its cellular localization. Methylated hnRNP Q3 localizes in the nucleocytoplasm. When cells were treated with type I PRMT inhibitor, hnRNP Q3 localization shifted from the nuclear to the cytoplasm (Passos et al., 2006). Besides, asymmetrical methylation of the hnRNPA2 RGG/RG motif reduces LLPS by disrupting arginine-aromatic interaction (Ryan et al., 2018).

SRSF1 is methylated by both PRMT1 and PRMT5 (Radzisheuskaya et al., 2019; Sinha et al., 2010). Asymmetrical methylation regulates SRSF1 subcellular localization, as an SRSF1 (R to A) mutant localizes to the cytoplasm. Reduced nuclear SRSF1 level leads to aberrant alternative splicing of its target, a similar phenotype also observed with hypophosphorylated SRSF1 (Sinha et al., 2010). Nevertheless, the crosstalk between arginine methylation and phosphorylation of SRSF1 remains to be investigated.

1.4 Histone arginine methylation

The nucleosome is the fundamental repeating unit of chromatin, consists of a histone octamer with two copies of four histone proteins, H2A, H2B, H3 and H4. PTMs in histones encompass methylation, phosphorylation, acetylation, ubiquitylation, SUMOylation, and ADP-ribosylation. Multiple sites of H2A, H3 and H4 can be mono-, symmetrically or asymmetrically di-methylated by PRMTs. These histone marks are implicated in the regulation of transcription, crosstalk with other histone modification markers, and serve as an interface for transcriptional regulators.

1.4.1 Histone arginine methylation catalyzed by PRMTs

PRMT1

PRMT1 catalyzes H4R3 asymmetrical methylation (Wang et al., 2001). H4R3me2a is an active mark that positively regulates gene transcription. *MLL* gene is commonly rearranged by chromosomal translocation, generating the fusion protein that act as oncogenic transcription factors in human malignancies. MLL-EEN transfusion is a major family of MLL fusion proteins (So et al., 1997). Fusion with EEN protein brings in PRMT1 as a component of the MLL complex, causing specific H4R3 asymmetrical methylation and H4 acetylation on oncogenes, such as *Hoxa9* (Cheung et al., 2007). Pull-down assays with H4R3me2a modified H4 tail identified SMARCA4 as a binding partner. H4R3me2a recruited SMARCA4 to the promoter of key factors in the epidermal growth factor receptor (EGFR) signaling pathway including *EGFR* and *TNS4*, enhancing the EGFR signaling and colorectal cancer cell proliferation (Yao et al., 2021). Similarly, in β-cells, the H4R3me2a mark is associated with increased DNA accessibility at the promoter binding sites of CCCTC-binding factor (CTCF) and β-cell transcription factors (Kim et al., 2020b). Switch of transcriptional regulators was observed depending on altered H4R3me2a level. A study of Th17 cells differentiation in the experimental autoimmune encephalomyelitis

(EAE) model showed that PRMT1 is required for controlling the reciprocal recruitment of signal transducer and activator of transcription (STAT3) and STAT5 (Sen et al., 2018). Inhibition of PRMT1 reduces IL-17 level, resulting from lower H4R3me2a level at the IL-17 promoter. The decreased H4R3me2a fails to recruit STAT3 to IL-17 promoter but facilitates STAT5 binding, leading to a subsequent reduction of IL-7 expression (Sen et al., 2018). H4R3me2a marks were observed on many other gene promotors to enhance transcription, including *PPARγ* (Tikhanovich et al., 2017), *Nrf2* (Huang et al., 2013), and *β-globin* (Huang et al., 2005).

PRMT5

PRMT5 deposits histone marks that are important for both gene activation and silencing. H4R3me2s modification is a repressive mark. PRMT5 was identified in the E2F complex that suppresses the expression of cyclin E1. Chromatin immunoprecipitation (ChIP) assay showed that PRMT5 binding to the promoter region of cyclin E1, symmetrically methylating H4R3 and reducing the expression of cyclin E1. Overexpression of PRMT5, thus, negatively regulates cell proliferation (Fabbrizio et al., 2002). Another repressive mark deposited by PRMT5 is H2AR3me2s, which is associated with transcription repression. H2AR3me2s was downregulated in the PRMT5 knockdown ES cells. Expression of methylation-deficient mutant H2A ($H2A^{R3A}$) histones in ES cells, similar to PRMT5 depletion, resulted in the upregulation of genes involved in differentiation (Tee et al., 2010). PRMT5 regulates RNA polymerase I-directed ribosomal targeted RNA gene (rDNA) expression depending on the DNA methylation status. H4R3me2s and H3R8me2s suppress the methylated rDNA expression, but promote the activation of unmethylated rDNA (Majumder et al., 2010). H3R8me2s is also an active mark. PRMT5 is recruited to the myogenin promoter during skeletal muscle differentiation. Depletion of PRMT5 reduces H3R8me2s at the myogenin locus, attenuating the binding of transcriptional factor MyoD

and Brg1-dependent chromatin remodeling complex, leading to loss of expression of myogenin (Dacwag et al., 2007). Besides, H3R2me2s catalyzed by PRMT5 is an active mark. H3R2me2s presented in euchromatic regions excludes binding of RBBP7, a central component of co-repressor complexes, and enhances binding of WD Repeat Domain 5 (WDR5), a common component of the many co-activator complexes (Migliori et al., 2012).

PRMT4(CARM1)

CARM1 is a transcriptional co-activator that catalyzes H3R17 and H3R26 methylation (El Messaoudi et al., 2006; Schurter et al., 2001). The function of these histone marks was studied in mouse embryo development. H3R4, H3R16 and H3R26 methylation are significantly different between cells at the four-cell-stage blastomeres. Cells with high levels of H3 methylation contribute to the inner cell mass (ICM) and polar trophectoderm, while cells with lower-level H3 methylation contribute more to the mural trophectoderm. Consistently, overexpression of CARM1 in individual blastomeres directs their progeny to the ICM and results in a dramatic upregulation of pluripotent cell markers *Nanog* and *Sox2* (Purcell et al., 2007).

1.4.2 Crosstalk between histone modification

Crosstalk between PTMs on histone tails forms the basis for the 'histone code', affecting protein binding and downstream events (Strahl and Allis, 2000). Histone arginine methylation through PTMs crosstalk is likely to be an important way of contributing to histone regulation, as the dynamic regulation of arginine methylation remains unclear. Additionally, several arginines in histone tails can be symmetrically and asymmetrically methylated, which adds another level of complexity level in histone regulation.

In vitro methylation showed that H4R3 methylation by PRMT1 facilitates its subsequent acetylation of H4K8 and H4K12 by p300 (Wang et al., 2001). However, when treated cells with

deacetylase inhibitor Trichostatin A (TSA), the hyperacetylation reduced methylation of H4R3me2a. This was confirmed by synthetic H4 tail methylation assay, when acetylated on K5, K6, K8 or K12 inhibited H4R3 methylation by PRMT1 (Wang et al., 2001). PRMT1 knockdown in a chick erythroid cell line results in loss of acetylation on both histones H3 and H4, including H3K9, H3K14, H4K5, H4K12, and H4K8, while increasing the methylation of H3K9 and H3K27 (Huang et al., 2005). A similar result was also observed in murine erythroleukemia (MEL) cell lines, where depletion of PRMT1 results in loss of H4R3me2a and H3K9/K14ac modification on the β-globin locus (Li et al., 2010).

H4R3me2s also links with DNA methylation. H4R3me2s marks are bound by the DNA methyltransferase, DNMT3A (Zhao et al., 2009). PRMT5 knockdown cells had reduced DNMT3A binding at the *γ-globin* loci, leading to loss of DNA methylation and gene activation (Zhao et al., 2009). Moreover, PRMT5 induces additional repressive epigenetic marks through regulating a repressor complex containing the histone lysine methyltransferase SUV420H1 (Rank et al., 2010). Expression of PRMT5 kinase-dead mutant or depletion of SUV420H1 resulted in the loss of complex binding to the γ-promoter, and reversal of both DNA methylation and histone repressive H4K20me3 mark with increased γ-gene expression (Rank et al., 2010). Acting as a transcription activation mark, H3R2me2s on the transcription start site (TSS) prevents H3K27 methylation by polycomb repressive complex 2 (PRC2). Both symmetrical dimethylation and monomethylation of H3R2 by PRMT5 has also been shown to recruit WDR5 and cause concomitant trimethylation of H3K4 (Chen et al., 2017; Migliori et al., 2012).

CARM1 catalyzes H3R17 and H3R26 dimethylation that involves in crosstalk with H3 lysine acetylation. Following estrogen stimulation, CBP/p300 acetylates H3K18 and K23, leading to R17 methylated by CARM1 on the Trefoil factor 1 (TFF1) encoding gene promoter. CARM1

mediated H3R17 methylation requires the acetylation on H3, as an acetyltransferase deficient CBP mutant is unable to stimulate this event (Daujat et al., 2002). The concomitant deposition of the histone modifications is required for proper transcriptional activation. Similarly, H3R26 methylation trigged by H3K27 acetylation was shown to be a defense mechanism to suppress HIV-1 long terminal repeat (LTR)-mediated cellular transcription (Zhang et al., 2017).

1.5 PRMTs and their substrates are implicated in diseases

Both PRMT1 and PRMT5 knockout cause embryonic lethality at the E6.5 mice (Kim and Ronai, 2020). Studies with the condition and temporary knockout of PRMTs in diverse tissues further established the function of PRMTs at the tissue level. Moreover, RGG/RG motif proteins are involved in the diverse cellular process and mutations can be causative factors for developmental diseases. PRMTs are found to be upregulated in different cancer types, which are linked to poor patient prognosis. Therefore, inhibition of PRMTs is a valuable cancer therapeutics. Indeed, certain PRMT inhibitors are in clinical trials. The substrates contributing to diseases progression are listed in Table 1.2. Finally, both viral proteins and host interactors harbor RGG/RG motifs. It is not surprising that PRMTs are also involved in host-viral interaction and are required for viral replication.

1.5.1 Development and neurologic diseases

The whole-body knockout of either Prmt1 and Prmt5 is embryonic lethal, demonstrating their essential function during mouse development.

Embryos homozygous Prmt1 mutant mice showed development arrest, whereas cell lines derived from these embryos are viable with a significantly low level of Prmt1 and global asymmetrical methylation (Pawlak et al., 2000). The homozygous mutant embryos die close to the time point that PRMT1 is induced at the E7.5 examined by X-Gal-staining and the signal of PRMT1 peaked in the brain at E8.5-E9.5 (Pawlak et al., 2000). This result is consistent with the previous finding in rat brain development. PRMT1 and arginine methylation level are high in the fetal brain and decreases immediately after birth (Paik et al., 1972). PRMT1 expresses in all major central nervous system (CNS) cell types, including neurons, astrocytes, oligodendrocytes, and microglia (Zhang et al., 2014). Conditional knockout of PRMT1 using *Nestin-Cre* mice causes

postnatal growth retardation and most of these conditional knockout mice died within two weeks after birth. Myelin sheath was absent due to lack of mature oligodendrocytes (Hashimoto et al., 2016). PRMT1 regulates axon and dendrite morphogenesis in neurons by methylating SCYL1 and facilitates forming of coat protein complex I (COPI) vesicles that modulate Golgi morphology (Amano et al., 2020). Moreover, increasing the ADMA is associated with *Caenorhabditis elegans* β-amyloid (Aβ) overexpression model, and in human SH-SY5Y neuroblastoma cells overexpressing the Swedish mutant form of human Aβ precursor protein (APPsw). Overexpression of dimethylarginine dimethylaminohydrolase 1 (DDAH1), that hydrolyzes ADMA and MMA, reduces oxidative stress and Aβ secretion in APPsw cells (Luo et al., 2015).

Prmt5 $^{-/-}$ mice embryos die at E6.5 without a discernable embryonic structure (Tee et al., 2010). In zygotes, PRMT5 is maternally inherited in the oocyte cytoplasm and activation of the embryonic Prmt5 starts at two- to four-cell stage (Tee et al., 2010). During 8- to 16- cell transition, PRMT5 translocates to the nucleus (Tee et al., 2010), the onset of global DNA demethylation (Smith et al., 2012). In the nucleus, PRMT5 catalyzes H2A/H4R3me2s on transposon elements (TE) locus, including LINE1 and IAP through an unknown mechanism (Kim et al., 2014). Consistently, loss of PRMT5 or maternally inherited PRMT5 leads to upregulation of these TEs (Kim et al., 2014). These findings suggest that PRMT5 is involved in genome defense during DNA hypomethylation. At the ~E11.5, PRMT5 relocates to the cytoplasm (Kim et al., 2014), where it interacts with Piwi protein for regulating a specific class of small noncoding RNAs, piwi-interacting RNAs (Vagin et al., 2009). Both Drosophila piwi proteins Pivi, Ago3 and Aub and murine Piwi proteins MIW1 and MIWI2 interact with and are methylated by PRMT5 (Kirino et al., 2009; Vagin et al., 2009). Loss of drosophila PRMT5 reduces of piRNAs, Ago3 and Aub proteins and subsequent accumulation of TEs (Vagin et al., 2009). Depletion of PRMT5 in

mESCs leads to upregulation of genes involved in embryonic differentiation and downregulation of pluripotent transcription factors, for instance, *Nanog* and *Sox2* (Tee et al., 2010). However, in human ES cells, depletion of PRMT5 has a limited effect on the expression of *OCT4*, *NANOG* or *SOX2* and did not prevent teratoma formation. Instead, PRMT5 regulates hES cells proliferation through G_1 cell cycle inhibitor p57 (Gkountela et al., 2014). Conditional knockout *Prmt5* with *Nestin-Cre* leads to postnatal death with a defect in brain development, caused by aberrant activation of the p53 pathway (Bezzi et al., 2013). Oligodendrocyte precursor cells (OPCs) with PRMT5 depletion fail to differentiate into mature myelinating oligodendrocytes. This is caused by the destabilization of platelet-derived growth factor receptors (PDGFRα),which hypomethylation lead to its degradation by E3 ligase Cbl (Calabretta et al., 2018).

Studies also revealed that RGG/RG motif-containing proteins play an important role in neurologic diseases. An example is the repeat expansion in the *C9orf72* gene, a frequent cause of ALS and frontotemporal dementia (FTD) (DeJesus-Hernandez et al., 2011). The expansion of G4C2 repeat, although in the noncoding region of *C9orf72* gene, leads to adding a glycine-arginine repeat (poly(GR)) sequence through a non-canonical mechanism known as repeat-associated non-ATG (RAN) translation (Gendron et al., 2013; Mori et al., 2013). The poly(GR) causes mislocalization of nucleocytoplasmic transport factors and nuclear pore complex protein, sequesters TDP-43 and induces cytoplasmic TDP-43 inclusion (Cook et al., 2020). Moreover, poly(GR) has a high affinity to unstructured tubulin tails of microtubules. This interaction impedes cargo transport and microtubule-based motility, and is critical for neuronal survival (Fumagalli et al., 2021). It is not surprising that poly(GR) is asymmetrically methylated by PRMT1, and inhibition of type I PRMT protects against poly(GR) toxicity (Premasiri et al., 2020).

1.5.2 Cancer

Cancer is a leading cause of death worldwide. The hallmarks of cancer comprise six biological capabilities acquired: sustaining proliferative signaling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, and activating invasion and metastasis (Hanahan and Weinberg, 2011). Dysregulation of PRMTs was reported to be a driving force for cancer evolution, including leukemia, lymphoma, breast, lung, colon, bladder, and prostate cancer. Many studies elucidate the contribution of PRMTs for the development and progression of cancer.

Estrogen receptor alpha (ERα) is expressed in 80% of breast cancer and is a well-established target (Johnston, 2010). PRMT1 is overexpressed in breast cancer (Mathioudaki et al., 2011). The ERα was identified as a PRMT1 substrate, and the methylation is required for the interaction with phosphoinositide 3-kinases (PI3K) and Src, and the activation of the downstream effector protein kinase B (AKT) (Le Romancer et al., 2008; Poulard et al., 2012). Moreover, insulin-like growth factor 1 (IGF-1) treatment rapidly induces ERα methylation, facilitating ERα interaction with IGF-1R (Choucair et al., 2019). Induced overexpression of PRMT1, CARM1 and PRMT6 in mice epithelial tissue promotes cell proliferation and transformation, resulting from an up-regulation of the PI3K-AKT pathway (Bao et al., 2019). progesterone receptor (PR), a transcription factor that contributed to breast cancer tumorigenesis, was shown to be methylated by PRMT1. Depleting PRMT1 increasesPR stability and decreases PR recruitment to EGFR and STAT5A locus (Malbeteau et al., 2020). Moreover, PRMT1 affects breast cancer cell EMT, invasion and metastasis by regulating zeste homolog 2 (EZH2) (Li et al., 2020b), a histone lysine methyltransferase that is overexpressed in many types of cancers (Kim and Roberts, 2016). Asymmetrical arginine methylation of EZH2 interferes with the nearby phosphorylation site that

regulated by CDK-1, and subsequently prevents ubiquitylation (Li et al., 2020b).In addition, tumor-associated macrophages secrete interleukin-6 (IL-6) cytokine, causing PRMT1 mediated EZH2 stabilization and promoting breast cancer metastasis (Li et al., 2020c).

PRMT5 is also highly expressed in breast cancer and implicated in tumorigenesis. PRMT5 regulates AKT1 by directly methylation. Arginine methylation of AKT1 promotes its kinase activity and conformation change, leading to a positive feedback loop of PI3K-dependent membrane translocation and enhanced AKT1 activation (Yin et al., 2021). Among MCF7 breast cancer cells, PRMT5 rather than PRMT1 is upregulated in the cancer stem cell population ($ESA^{+}CD24^{low}CD44^{+}$). PRMT5 regulates forkhead box P1 (FOXP1) expression, which is a transcription factor for breast cancer stem cell function. H3R2me2s marks in the *FOXP1* promoter region recruit SET domain containing 1 (SET1), tri-methylating H3K4, and promote *FOXP1* expression (Chiang et al., 2017). KLF5 is an oncogenic factor that promotes breast cancer cells proliferation, migration and stemness. KLF5 was reported to be a substrate for PRMT5, and the methylation stabilizes KLF5 by reducing phosphorylation and ubiquitination (Wang et al., 2021a).

Both PRMT1 and PRMT5 are implicated in hematologic malignancies, including leukemia, lymphoma, and multiple myeloma. Type I PRMT and PRMT5 inhibitors are in clinical trials for treating hematologic malignancies (Li et al., 2019). PRMT1 plays an important role in MLL. MLL fusion protein recruits PRMT1 for oncogene activation (See 1.4.1). In β-catenin ^{high} MML leukemic stem cells, β-catenin and homeobox A9 (HOXA9) co-regulated PRMT1 expression, and treatment with type I PRMT inhibitors can be an alternative avenue for targeting β-catenin ^{high} MLL (Siriboonpiputtana et al., 2017). Through a similar mechanism, fusion protein AML1-ETO, which is related to leukemia development, recruits PRMT1 as a co-activator for its downstream target gene expression (Shia et al., 2012). However, PRMT1 is necessary but not

sufficient for leukemic transformation. KDM4C, an H3K9 demethylase, has been shown to co-regulate the histone code with PRMT1 in acute myeloid leukemia (AML) (Cheung et al., 2016). It is not surprising, as the acetylation marks could affect H4R3me2a deposition (Wang et al., 2001). Other than histone methylation, PRMT1 regulated non-histone protein methylation is also implicated in AML. AML patients with FMS-like receptor tyrosine kinase-3 internal tandem duplication (FLT3-ITD) mutations have poor clinical outcomes. PRMT1 methylates FLT3-ITD protein, which facilitates the recruitment of adaptor proteins to FLT3. Combined treatment of type I PRMT inhibitors and tyrosine kinase inhibitors that block autophosphorylation of FLT3 eliminates leukemia cells growth (He et al., 2019; Zhu et al., 2019b).

PRMT5 acts as an epigenetic regulator that regulates proleukemic gene expression. As a transcriptional suppressor, PRMT5 interacts with Sp1 in a transcription repressor complex, methylates H4R3, leading to the silence of miR-29b. The reduced miR-29b increases its proleukemic target, such as FLT3, a receptor-type tyrosine kinase (Tarighat et al., 2016). In another study, PRMT5 acts as a transcriptional activator in AML cells. Methylation of histone H3 by PRMT5 abrogates its subsequent methylation by PRC2, leading to gene activation (Liu et al., 2020). PRMT5 regulates NF-κB signaling through direct interaction with TRIM21. Inhibition of PRMT5 abrogated NF-κB signaling, resulting in significant cell growth reduction in multiple myeloma cells (Gulla et al., 2018). Synergistic effects are observed with PRMT5 inhibitors and other small molecules. For example, in JAK mutant myeloproliferative neoplasms (MPN), dual PRMT5 and JAK1/2 inhibitors are superior to monotherapy. The mechanism may contribute partially through PRMT5 methylation of E2F1, a transcription factor that mediates the expression of genes involved in the cell cycle and DNA damage repair (Pastore et al., 2020). PRMT5 depletion or inhibition causes aberrant splicing of the multifunctional histone-modifying and

DNA-repair factor TIP60/KAT5, leading to HR defect. A combination of PARP and PRMT5 showed a synergistic effect on AML cells (Hamard et al., 2018). In the MLL model, inhibition of histone methyltransferase disruptor of telomeric silencing 1-like (DOT1L) is one specific therapeutic in clinical trials, and supplement with PRMT5 inhibitors sensitizing cells to chemotherapy (Secker et al., 2019). Several studies explore the mechanism that regulates PRMT5 in leukemia. PRMT5 is positively regulated by transcriptional factors PAFc, MLL1, HOXA9 and STAT5 in leukemic cells. B cell receptor (BCR) regulates PRMT5 expression through PI3K-AKT-MYC signaling in diffuse large B cell lymphoma (DLBCL). Forming a positive feedback loop, PRMT5 promotes cell cycle progression and enhances PI3K-AKT signaling. Thus, dual inhibiting PRMT5 and AKT causes synergistic cell death (Zhu et al., 2019a).

Notably, combination of PRMT5 and type I PRMT inhibitions result in the synergistic killing of cancer cells, including leukemia. Inhibition of type I PRMT induces substrate scavenging, and the addition of PRMT5 inhibitor attenuates the global level of arginine methylation, resulting in hypomethylation of splicing factors and profound effects on alternative splicing that distinct from those observed with either single inhibitor (Fedoriw et al., 2019). Deletion of *MTAP*, a gene encoding methylthioadenosine phosphorylase, was identified in a large portion of various types of cancers (Kryukov et al., 2016; Mavrakis et al., 2016). MTAP loss causes the accumulation of MTA, which is a natural inhibitor of PRMT5. Thus, cancer cells with MTAP deletion have a decreased level of SDMA and are sensitive to type I PRMT inhibitors (Fedoriw et al., 2019). However, a recent study showed that extra care should be taken to design the study and interpret the results, as *in vivo* model showed that homozygous MTAP-deleted primary glioblastoma tumors do not significantly accumulate MTA, which result from the MTAP expressing by stroma cells (Barekatain et al., 2021). On the other hand, because many splicing factors are PRMTs

substrates and change-of-function mutations in RNA splicing factors commonly occur in leukemia, dual inhibition of type I PRMT and PRMT5 further perturbation splicing process in leukemia cells with splicing factors mutant (Fong et al., 2019).

Table 1. 2 PRMTs and their substrates are implicated in a variety of diseases.

PRMT substrate	Diseases	Arginine methylation status	Proposed mechanism	reference/ PMID
EZH2	breast cancer	increase methylation by PRMT1	EZH2 is asymmetrically dimethylated at R342 by PRMT1, resulting in a decrease in EZH2 target gene expression, but an increase in breast cancer cell EMT, invasion and metastasis.	32895488
ERα	breast cancer	increase methylation by PRMT1	ERa is methylated by PRMT1 at R260. The methylation is required for the interaction with PI3K and Src, and the activation of the downstream effector AKT.	18657504
PR	breast cancer	increase methylation by PRMT1	PRMT1 methylates PR at the arginine 637 and reduces the stability of the receptor, thereby accelerating its transcriptional activity.	32563156
FLT3	leukemia	increase methylation by PRMT1	PRMT1 catalyzed FLT3-ITD protein methylation at arginine 972/973, and PRMT1 promoted leukemia cell growth in an FLT3 methylation-dependent manner.	31217189
RUNX1	leukemia	increase methylation by PRMT1	RUNX1 is arginine-methylated by PRMT1, and PRMT1 serves as a transcriptional coactivator for RUNX1 function.	18316480
p14ARF	pancreatic cancer	increase methylation by PRMT1	PRMT1 methylates several arginine residues in p14ARF, which is crucial for nucleolar localization of p14ARF. The arginine methylation is important for p14ARF 's tumor-suppressive function.	33999432
KLF5	breast cancer	increase methylation by PRMT5	PRMT5 interacts with KLF5 and catalyzes the di-methylation of KLF5 at Arginine 57 (R57). The methylation stabilized KLF5 and promoted breast cancer stem cell maintenance.	33972717
AKT	breast cancer	increase methylation by PRMT5	AKT1-R391 methylation is a critical step for PI3K-mediated AKT activation and oncogenic function. This methylation event is enhanced when AKT is in an open state, suggesting a positive feedback loop between AKT-R391 methylation and AKT activation.	34103528
E2F1	myeloproliferative neoplasm	increase methylation by PRMT5	PRMT5 inhibition reduced the expression of E2F targets and altered the methylation status of E2F1, leading to attenuated DNA damage repair, cell-cycle arrest, and increased apoptosis.	32669286
p53	leukemia/lymphoma	increase methylation by PRMT5	Arginine methylation of p53 selectively suppresses the expression of crucial proapoptotic and antiproliferative target genes, thereby sustaining tumor cell self-renewal and proliferation and bypassing the need to acquire inactivating p53 mutations.	25582697
C9orf72	ALS	increase methylation by PRMT1	The poly(GR) causes mislocalization of nucleocytoplasmic transport factors and nuclear pore complex protein, sequesters TDP-43, and induces cytoplasmic TDP-43 inclusion.	32878979

1.5.3 Virus infection

A long list of virus proteins contains RGG/RG motifs. These RGG/RG motifs have been shown to have important functions for the viral life cycle and are the substrate of host PRMTs. For example, Gag protein of prototype foamy virus (PFV) harbors RGG/RG motifs in its C-terminus. Deletion and mutation of the RGG/RG motif disrupted RNA encapsidation, genome reverse transcription, viral particle morphology and virion infectivity (Mullers et al., 2011). The adenovirus type 5 (Ad5) late region 4 (L4) protein (L4-100K) is another example. PRMT1 asymmetrically methylates L4-100K RGG/RG motif, which promotes capsid assembly of the structural protein and requires for efficient Ad replication (Koyuncu and Dobner, 2009). These studies suggest arginine methylation of virus protein is required for their function. Thus, targeting methylation modification could be a strategy to reduce virus production.

On the other hand, recent studies also revealed that PRMTs play an essential role in regulating immune response-related pathways. PRMT5 directly methylates cyclic GMP-AMP synthase (cGAS), a cytoplasmic double-stranded DNA sensor that activates antiviral response (Ma et al., 2021). Methylation of cGAS attenuates its DNA binding ability and antiviral immune response. Moreover, the treatment of PRMT5 inhibitors activates the cGAS pathway and protects mice from HSV-1 infection (Ma et al., 2021). PRMT1 is also involved in the pathway of immune response regulation such TNFα/NFκB (Reintjes et al., 2016), yet studies are further needed to elucidate the comprehensive function of PRMT1 during virus infection.

1.6 The rationale and hypothesis of the research

As described above, more than 75% RBPs are substrates of PRMTs. Elevated expression of PRMTs is implicated in various types of cancer, resulting in altered arginine methylation of proteins. The loss-of-function of PRMT7 causes severe developmental disabilities in human, but

the contributing substrates remain unknown. Amino acid substitutions in SMA patients disrupt the SMN Tudor domain function, indicating its interaction with methylated arginine proteins is important for SMN function. Given that considerable evidence points to the cause of diseases by the dysregulation of protein arginine methylation, we hypothesize that the genetic deletion of an RGG/RG motif would lead to pathogenesis.

CHAPTER 2 Deletion of RBMX RGG/RG motif in Shashi XLID syndrome leads to aberrant p53 activation and neuronal differentiation defects

Ting Cai[1,2], Jessica K. Cinkornpumin[2], Zhenbao Yu[1], Oscar D. Villarreal[1], William A. Pastor[2] and Stéphane Richard[1]*

[1]Segal Cancer Center, Lady Davis Institute for Medical Research and Gerald Bronfman Department of Oncology and Departments of Biochemistry, Human Genetics and Medicine, McGill University, Montréal, Québec, H3T 1E2, Canada

[2]Department of Biochemistry, McGill University, Montreal, QC, H3G 1Y6, Canada and the Rosalind & Morris Goodman Cancer Research Centre, McGill University, Montreal, QC, H3A 1A3, Canada.

2.1 Preface

Previously, Richard lab identified over 1000 human proteins with RGG/RG motifs by genome-wide analysis. RBMX, a hnRNP family protein, is one of the proteins that contain RGG/RG motif. It has been reported that RBMX RGG/RG motif deletion may link to a subtype of XLID, called Shashi syndrome. Herein, I characterized the mechanism underlying this genetic disease.

2.2 Abstract

RNA binding protein RBMX has a C-terminal RGG/RG motif of unknown function that is deleted in patients with Shashi-X-linked-intellectual-disability (XLID) syndrome. Herein, we show RBMX-depleted cells of its RGG/RG motif exhibited aberrant activation of the p53 pathway. We show that the RBMX RGG/RG motif is methylated by PRMT5 and this regulates assembly with the SRSF1 splicing factor into higher-order-complexes. Depletion of RBMX or disruption of RBMX/SRSF1 complex in PRMT5-depleted cells reduced SRSF1 binding to the *MDM4* pre-mRNA leading to exon 6 exclusion and lower MDM4 protein levels. Transcriptomic analysis of isogenic Shashi-XLID induced pluripotent stem cells generated using CRISPR-Cas9 revealed a dysregulation of MDM4 splicing and aberrant p53 upregulation. Shashi-XLID neural progenitor cells (NPCs) displayed differentiation and morphological abnormalities accompanied with excessive apoptosis. Our findings identify RBMX as a regulator of SRSF1 activity and the p53 pathway, suggesting that the loss-of-function of the RBMX RGG/RG motif is the cause of the Shashi-XLID syndrome.

Keywords: RBMX, SRSF1, RGG/RG motif, p53, PRMT5, alternative splicing, membraneless organelles, iPSCs, and neuronal differentiation.

2.3 Introduction

RNA binding proteins (RBPs) are implicated in diverse human disorders including X-linked-intellectual-disability (XLID) accounting for ~16% of all cases of intellectual disability (Ellison et al., 2013). Many RBPs encoded by XLID genes and the underlying mechanism for the pathogenesis are remains to be characterized. More than 75% of RBPs are arginine methylated, which affects RBPs function including alternative splicing regulation (Thandapani et al., 2013a). Alternative splicing is highly involved in brain development, modulating each step of lineage commitment, neurogenesis and neuron function (Su et al., 2018). RNA sequencing comparing developing human brain derived NPC and neuron samples revealed alternative splicing governs cell fate and splicing factors are dynamically regulated during development (Zhang et al., 2016).

RBMX or hnRNP G is an RBP of the heterogeneous nuclear ribonucleoprotein (hnRNP) family. RBMX was proposed as a candidate gene for the XLID subtype termed the Shashi-XLID syndrome for a family pedigree with seven affected males (Shashi et al., 2015). The genetic lesion deletes 23 base pairs of the RBMX coding region, predicted to cause a frameshift and generate a premature termination codon, deleting a low complexity region termed the RGG/RG motif. RBMX is a multifunctional protein with essential roles in pre-mRNA splicing (Elliott et al., 1998; Hofmann and Wirth, 2002; Wang et al., 2011; Zhou et al., 2019) (Heinrich et al., 2009; Munschauer et al., 2018), DNA damage response (Adamson et al., 2012), heterochromatin maintenance (Becker et al., 2017) and genomic stability (Munschauer et al., 2018). Depletion of RBMX in Xenopus laevis (Dichmann et al., 2008) or zebrafish embryos (Tsend-Ayush et al., 2005) leads to impaired brain development and dysregulated somite patterning. Moreover, RBMX knockdown in rat primary hippocampal neurons causes a decrease in dendritic spine density, indicating its important role in the regulation of synaptic activity (Zhang et al., 2012).

RGG/RG motifs are frequently found in RBPs and are the preferred substrates of protein arginine methyltransferases (PRMTs) (Thandapani et al., 2013a). Arginine methylation of RBPs affects their protein stability, cellular localization, RNA binding affinity, interactome and capacity to promote liquid-liquid phase separation (LLPS) (Thandapani et al., 2013a). One of the best characterized XLID gene is FMR1 (Pieretti et al., 1991), which encodes for an RGG/RG motif-containing RBP. Arginine methylation of FMRP1 affects its RNA binding capacity (Blackwell et al., 2010). Remarkably, hypomethylation of the FMRP1 RGG/RG motif promotes its LLPS, leading to mRNA translation activation within synaptic regions (Tsang et al., 2019).

PRMT5 is the major type II enzyme catalyzing mono- and symmetrical di-methylation of arginines (Guccione and Richard, 2019). RGG/RG motif-containing RBPs such as Sm proteins (Meister et al., 2001), ribosomal protein S10 (Ren et al., 2010), and RNA helicase DDX5 (Mersaoui et al., 2019; Villarreal et al., 2020) are known substrates of PRMT5. Conditional knockout of PRMT5 in mouse neural progenitor cells (NPCs) leads to postnatal death, caused by alteration of *MDM4* pre-mRNA alternative splicing and p53 activation (Bezzi et al., 2013).

Herein, we show that methylation of the RGG/RG motif of RBMX by PRMT5 regulates *MDM4* alternative splicing and the p53 pathway. Depletion of RBMX or PRMT5-inhibition diminished SRSF1 binding to the *MDM4* pre-mRNA reducing MDM4 protein expression resulting in the activation of the p53 pathway. Using RBMX-ΔRGG mutant human induced pluripotent stem cells (hiPSCs), as observed in the Shashi-XLID syndrome, we observed dysregulation of the p53 pathway, increased apoptosis and defects in neuronal differentiation. Our findings define a role of the RBMX RGG/RG motif and its implication in the Shashi-XLID syndrome.

2.4 Results

2.4.1 Loss of the RBMX RGG/RG motif in Shashi-XLID syndrome increases the p53 response

The 23bp genetic lesion observed in the Shashi-XLID is located in the last exon of RBMX, predicted to generate a truncated protein causing a frameshift deleting the last C-terminal 38 amino acids encompassing the RGG/RG motif (Shashi et al., 2015). To characterize this genomic lesion, we obtained an induced pluripotent stem cell (iPSC) line derived from a healthy human male and truncated the C-terminus of RBMX using CRISPR-CAS9 to replicate what is observed in patients with Shashi-XLID. Two iPSC lines named DRGG1 and DRGG2, for "deleted RGG" and a non-edited clone (CTRL) were generated and selected for further characterization (Figure S2.1A). Both DRGG1 and DRGG2 harbored C-terminal truncations deleting the RGG/RG motif (Figure S2.1A) and the RBMX-ΔRGG proteins expressed in the DRGG clones migrated with ~4 kDa decrease in mass, as expected (Figure 2.1A). Comparative genomic hybridization of CTRL and DRGG1 iPSCs confirmed no gross karyotypic abnormalities (data not shown).

We next performed RNA sequencing (RNA-seq) of the CTRL and DRGG1 iPSCs to define transcriptomic difference between the cell lines. We identified 847 and 1067 genes which were significantly upregulated and downregulated in DRGG1 iPSCs, respectively. Using gene set enrichment analysis (GSEA) with hallmark gene sets, we identified a few pathways to be dysregulated in DRGG1 cells including the epithelial mesenchymal transition (EMT) pathway, KRAS signaling pathway, estrogen response, hypoxia and the p53 pathway (Figure S2.1B). Since many neurological developmental syndromes have constitutive activation of the p53 pathway (Bowen et al., 2019), we asked whether the Shashi-XLID syndrome could also be classified as such. Detailed analysis identified many consensus p53 transcriptional targets (Fischer, 2017) that

were significantly upregulated in DRGG iPSCs (Figure 2.1B). This included cell cycle and apoptosis targets CDKN1A, TP53INP1, IKBIP, DRAM1 and ZMAT3 (Figure S1C) and CDKN1A and BBC3 were confirmed by RT-qPCR (Figure 2.1C). We next examined p53 activation induced apoptosis in CTRL and DRGG iPSCs by co-immunostaining p53 and the apoptosis marker cleaved caspase 3 (CC3, Figure 2.1D). We observed a significant increase in cells with nuclear p53 in the DRGG1 (19.74%) and DRGG2 cells (19.93%), compared with CTRL (8.67%) (Figure 2.1E) and these nuclear p53 expressing cells frequently co-stained with the cleaved caspase 3 (2.9% in CTRL, 4.8% in DRGG1 and 6.06% in DRGG2; Figure 2.1F). Our findings suggest that the loss of the RBMX RGG/RG motif in Shashi-XLID syndrome causes an upregulation of the p53 pathway with functional consequences.

To define the molecular mechanism of RBMX regulation, we turned to the p53 wild type U2OS cancer cell line. Consistent with the observation in iPSCs, RBMX knockdown in U2OS cells led to a robust increase in p53 target genes at the mRNA and protein levels including CDKN1A/p21, BBC3, BAX, and MDM2, comparing with siLuciferase (siCTRL) (Figure 2.1G, 2.1H). The TP53 mRNA level was unaffected, while the p53 protein level was strikingly elevated in RBMX deficient U2OS cells (Figure 2.1G, 2.1H). Taken together these data suggest RBMX regulates the p53 pathway in various cell types.

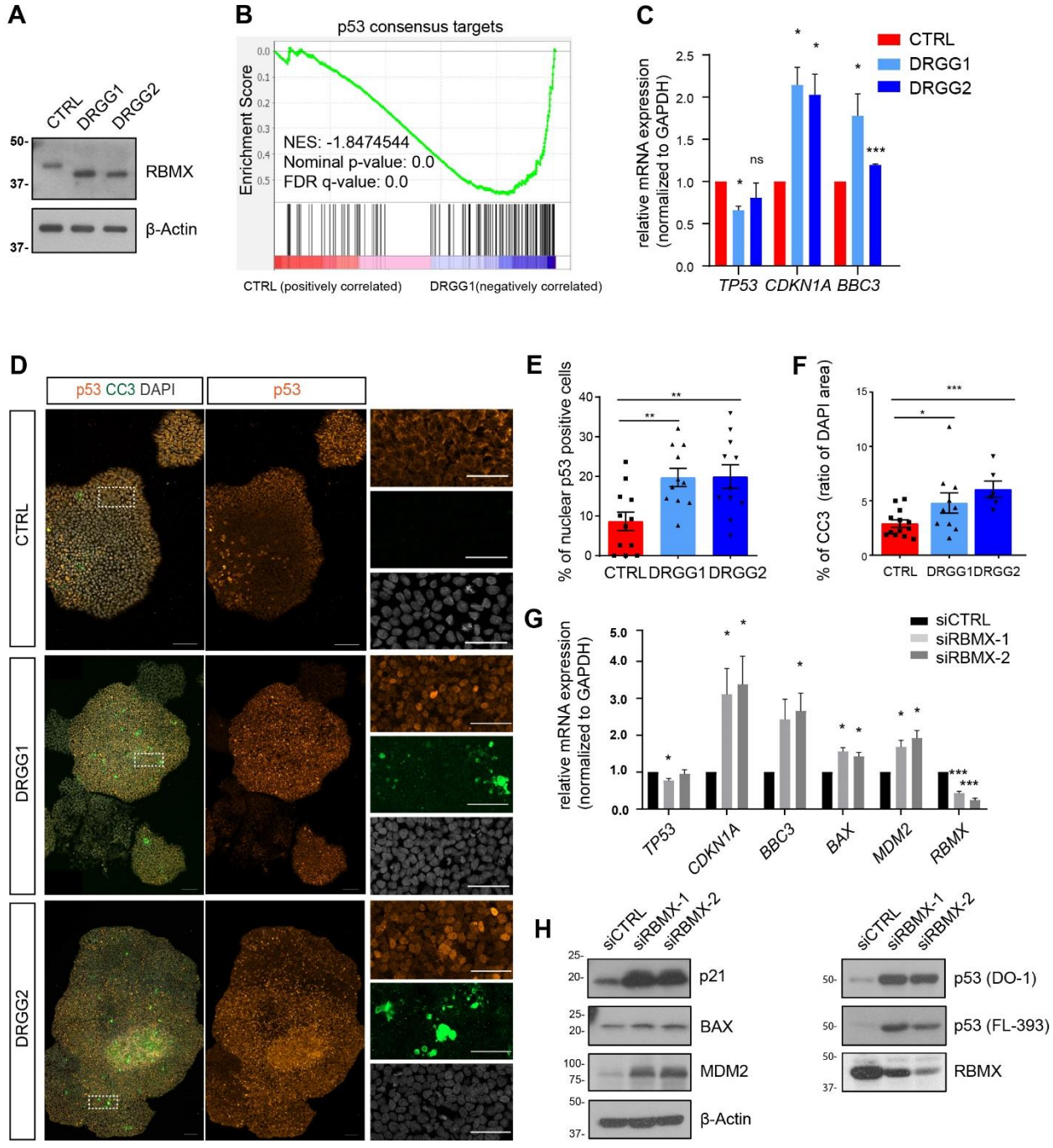


Figure 2. 1 Activation of the p53 pathway in Shashi-XLID iPSCs and U2OS cells.
(A) Whole cell extracts from control (CTRL) and RBMX-ΔRGG (clones DRGG1, DRGG2) iPSCs were immunoblotted with anti-RBMX and -β-actin antibodies. Molecular mass markers are shown on the left in kDa.
(B) GSEA analysis of p53 consensus targets (Fischer, 2017) using FPKM.
(C) RT-qPCR assays were performed for *p53*, *CDKN1A* and *BBC3* in CTRL, DRGG1 and DRGG2 iPSCs. The data were normalized to GAPDH. n = 2 independent experiments. **p* <0.05; ***p* <0.01; n.s. non-significant; unpaired t test. mean ± SEM.

(D) Immunofluorescence staining of hiPSC colonies with anti-p53 and anti-cleaved caspase 3 (CC3) antibodies in CTRL and DRGG1 cells counterstained with DAPI. Scale bar, 100 μm. Insert scale bar, 50 μm.
(E, F) Quantification of percentage of p53 positive cells and percentage of cleaved caspase3 positive area (ratio=cleaved caspase3 positive area/DAPI positive area) in RBMX-ΔRGG and control CTRL. n= 2 independent experiments. * p <0.05; ***p <0.001; unpaired t test comparing each sample to the CTRL. mean ± SEM.
(G) RT-qPCR assays were performed for the expression of indicated mRNAs in U2OS cells transfected with siLuciferase (siCTRL) and two siRNAs targeting RBMX (siRBMX-1 and siRBMX-2). The data were normalized to GAPDH. n =3 independent experiments. * p <0.05; ***p <0.001; unpaired t test. mean ± SEM
(H) U2OS cells transfected with indicated siRNAs were subject to immunoblotting with indicated antibodies.

2.4.2 Requirement of the methylation of RBMX by PRMT5 for p53 protein stability

RGG/RG motifs are short linear motifs that are preferred sites of methylation by protein arginine methyltransferases (PRMTs) (Thandapani et al., 2013a). Human RBMX harbors two RGG/RG motifs; one residing from residues 102 to 127 and another at the C-terminus from residues 369 to 387. In vitro methylation assays were performed with recombinant PRMT5, an enzyme with preference for RGG/RG motifs (Bedford and Clarke, 2009) and methylation of the GST-RBMX:166-391 protein was observed, but not the GST-RBMX:1-166 protein, suggesting only the C-terminal RGG/RG motif is methylated by PRMT5 (Figure 2.2A). Additional C-terminal GST-RBMX truncation proteins were generated and the minimal region methylated in vitro by PRMT5 was mapped to RBMX residues 366 to 391 (Figure 2.2B). To identify methylated arginines in vivo, we performed mass spectrometry analysis with immunoprecipitated FLAG-RBMX. We identified 55 unique RBMX peptides, covering 81% of the protein including peptides containing methylated R369 and R373 located within the C-terminal RGG/RG motif (Figure 2.2C). Non-methylated peptides containing R369 and R373 were not identified, suggesting that R369 and R373 are significantly methylated in cells. In total, we identified 5 dimethylarginines and 9 monomethylarginines within RBMX (Table S2.1). As the Shashi-XLID genetic deletion

involves in the C-terminus of RBMX, we focused on the methylation of the C-terminal RGG/RG motif. To further characterize methylation of RBMX, we generated FLAG-RBMX with a deleted C-terminal RGG/RG motif (FLAG-RBMX-ΔRGG), as predicted for patients with Shashi XLID, and FLAG-RBMX-RK, where the arginines within the C-terminal RGG/RG motif were replaced with lysines (Figure 2.2D). Immunoblotting using an anti-symmetrically dimethylated arginine antibody (SYM10), confirmed that FLAG-RBMX was methylated in U2OS cells (Figure 2.2E, lane 2). Recognition of both FLAG-RBMX-ΔRGG and FLAG-RBMX-RK was significantly less with SYM10 than RBMX (Figure 2.2F, compare lanes 3 and 4 with lane 2), confirming the methylation of C-terminal RGG/RG motif represents the major RBMX epitope for SYM10.

We next defined the functional region of RBMX required for increased p53 protein levels. Cell lines stably expressing siRNA resistant FLAG-RBMX, FLAG-RBMX-ΔN (deletion of the N-terminal RNA recognition motif (RRM), Figure 2.2D) and FLAG-RBMX-RK were generated and transfected with either siCTRL or siRBMX. Cells expressing FLAG-RBMX or FLAG-RBMX-ΔN reversed the induction of p53 in siRBMX-transfected cells, while RBMX-RK did not (Figure 2.2F). These findings suggest that arginine methylation of the RGG/RG motif is required for RBMX function in the regulation of p53 activation. Next, siCTRL and siRBMX cells were treated with the protein synthesis inhibitor cycloheximide and the decay of p53 protein was assessed by immunoblotting for up to 120 min. In siCTRL U2OS cells, the half-life of p53 was in the range of 15 to 30 min, and in siRBMX cells it was > 120 min (Figure 2.2G). FLAG-RBMX-ΔN expression destabilized p53 in RBMX-depleted cells, while RBMX-RK did not (Figure 2.2G). These finding show that RBMX loss-of-function in U2OS cells increased p53 protein stability, as observed in Shashi-XLID iPSCs (Figure 2.1).

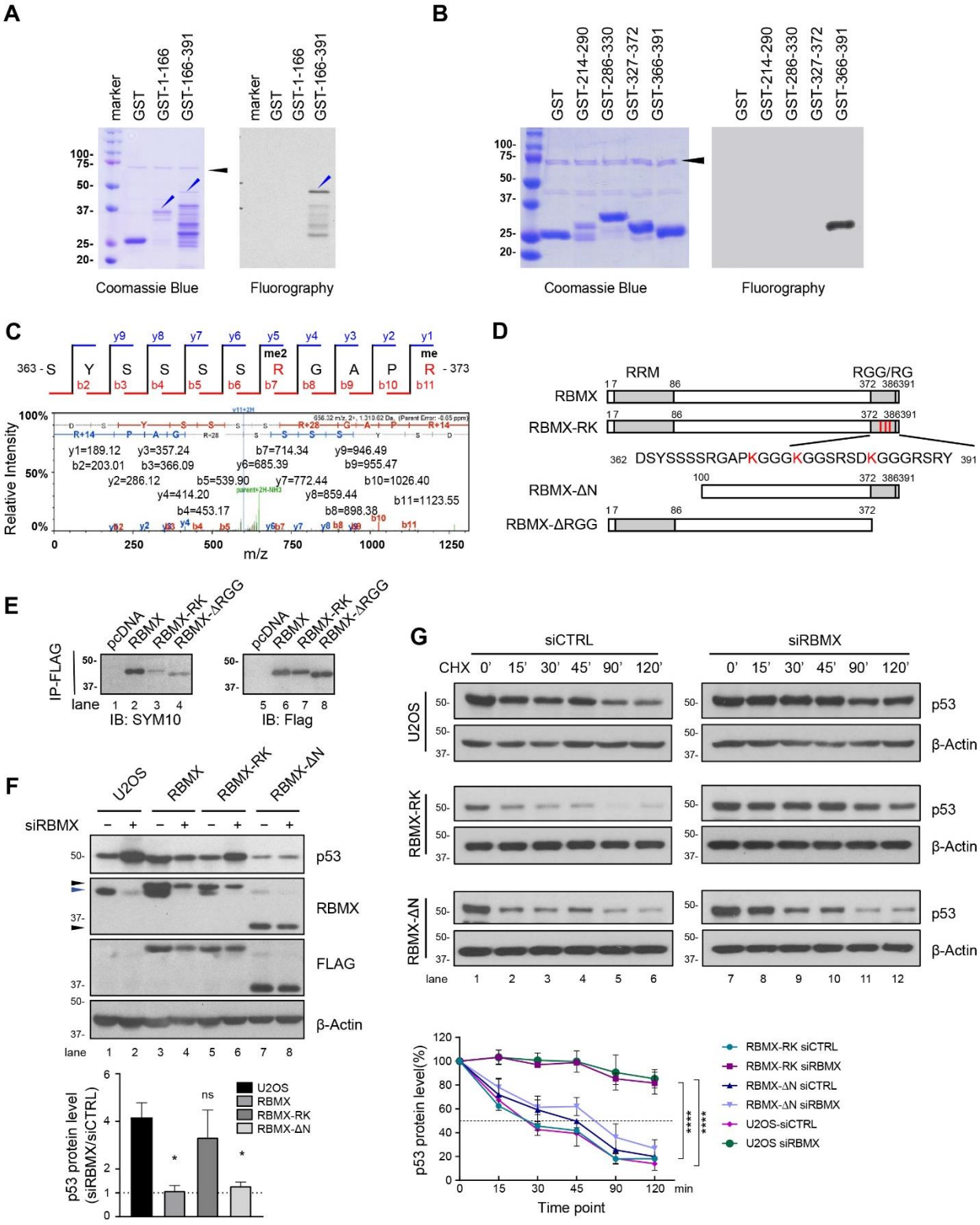


Figure 2. 2 Regulation of p53 protein stability depends on the arginine methylation of the RBMX RGG/RG motif.

(A, B) GST-RBMX fragments were subjected to *in vitro* methylation assays with recombinant PRMT5/MEP50 in the presence of (^{3}H)-S-adenosyl-methionine. The proteins were separated by SDS-PAGE and stained with Coomassie Blue (left panel) and methylation detected by fluorography. The arrowhead indicates the migration of PRMT5 and the molecular mass markers are on the left in kDa.
(C) Identification of RBMX methylarginine by mass spectrometry. FLAG-RBMX expressing HEK293 cells were lysed and anti-FLAG immunoprecipitations followed by mass spectrometry performed. The methylation sites are identified and marked.
(D) Schematic diagram of RBMX with RRM domain and its RGG/RG motif. RBMX-RK harbors RGG-KGG amino acid substitutions, RBMX-ΔN lacks the RRM domain and RBMX-ΔRGG lacks the RGG/RG motif.
(E) HEK293 cells were transiently transfected with control vector pcDNA3.1, FLAG-RBMX, FLAG-RBMX-RK and FLAG-RBMX-ΔRGG were lysed under denaturing conditions and the immunoprecipitated proteins were analyzed for methylarginines using SYM10. The membranes were re-probed with anti-FLAG antibody to confirm equal expression.
(F) Immunoblotting of p53, RBMX, FLAG, and β-actin in U2OS stable cell lines expressing siRNA-resistant full-length mouse FLAG-RBMX, FLAG-RBMX-RK, and FLAG-RBMX-ΔN transiently transfected with siCTRL or siRBMX to deplete the endogenous RBMX. The migration of endogenous RBMX is indicated with a blue arrowhead on the left, while the migration of FLAG-RBMX proteins is indicated with black arrowheads. Three repeats were performed (n=3) and the ratio of p53 protein level expressed in siRBMX/siCTRL is shown as mean ± SEM. *p <0.05; unpaired t test comparing each sample to the U2OS.
(G) U2OS and RBMX-RK and RBMX-ΔN stably transfected U2OS were transfected with siCTRL or siRBMX to deplete endogenous RBMX. The cells were subsequently incubated with 30 μg/ml cycloheximide (CHX) for indicated time period and subjected to immunoblotting with anti-p53 (DO-1) and β-actin antibodies. Three repeats were performed (n=3) and the p53 protein level is shown in percentage expressed as the mean ± SEM. The dashed line represents 50%. Two-way ANOVA comparing siRBMX to siLuc. ****p <0.0001.

2.4.3 RBMX regulates the alternative splicing of *MDM4*

p53 protein stability is known to be negatively regulated at the protein level by the balance of E3 ubiquitin ligases MDM2 and MDM4 (Lenos et al., 2012; Toledo and Wahl, 2006). Gene expression analysis did not show significant changes in MDM2 and MDM4 mRNA expression in DRGG1 iPSCs compared to the CTRL cells. MDM4 expression is known to be regulated at the level of alternative splicing (Rallapalli et al., 1999), where the *MDM4-FL* transcript generates full-length MDM4 protein, and the *MDM4-S* transcript undergoes non-sense mediated decay due to the presence of a premature termination codon (PTC) created by the skipping of exon 6 (Figure 2.3A). Thus, we analyzed our iPSCs RNA-seq data for changes in alternative splicing using rMATs (Shen et al., 2014) and we identified 369 differentially splicing events in DRGG iPSCs.

This included 247 events of skipped exons (SE), 16 events of alternative 3' splice sites (A3SS), 11 events of alternative 5' splice site (A5SS), 93 events of mutually exclusive exons (MXE) and 2 events of retained introns (RI). *MDM4* exon 6 skipping was one of the significant splicing events in DRGG1 iPSCs (Figure 2.3B) and this splicing imbalance was further confirmed by RT-PCR, where exon 6 skipping increased to approximately 41.5% in the DRGG1 and 36.5% DRGG2 iPSCs compared to 10.8% in the CTRL iPSCs (Figure 2.3C). Moreover, *MDM4* exon 6 exclusion was also observed in siRBMX and siPRMT5 transfected U2OS cells (Figure 2.3D) and it correlated with reduced MDM4 protein expression (Figure 2.3E). These data suggest RBMX activates the p53 pathway by dysregulating *MDM4* alternative splicing.

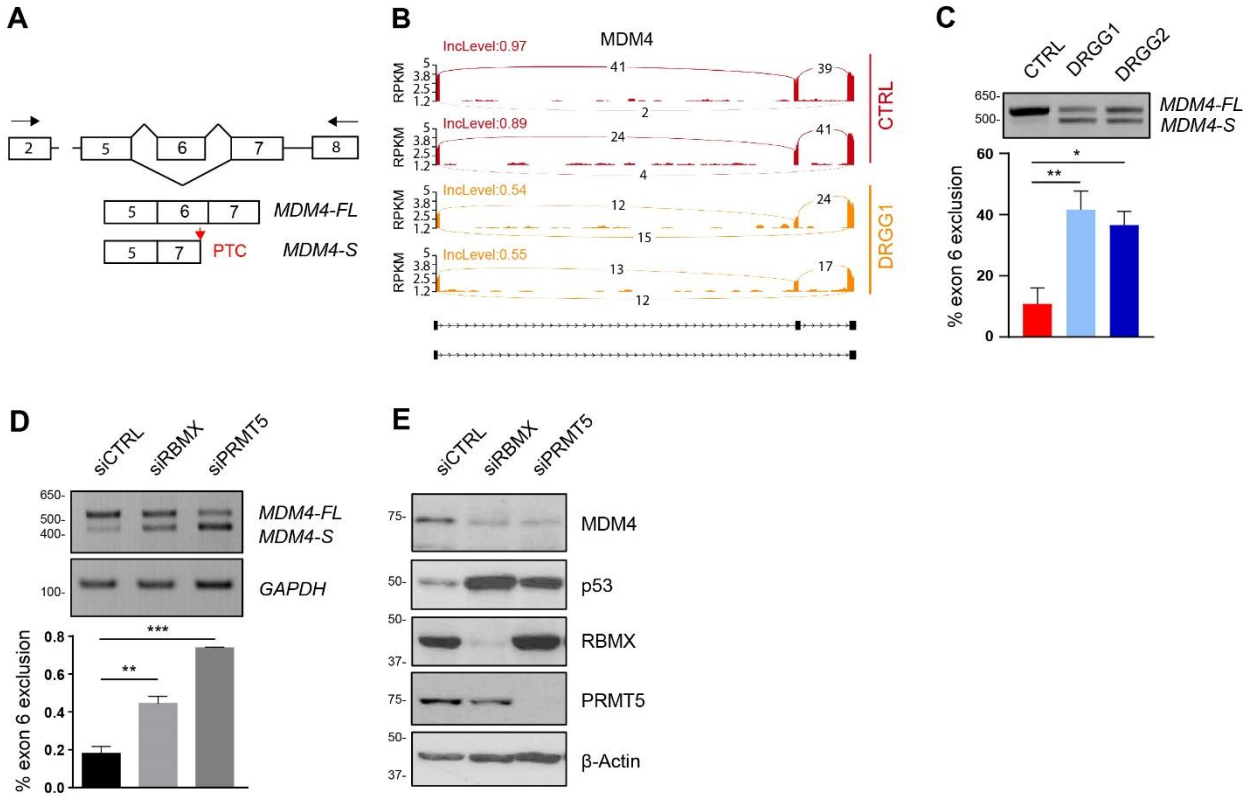


Figure 2. 3 RBMX regulates *MDM4* alternative splicing through its RGG/RG motif.
(A) Schematic representation of *MDM4* exon-intron boundaries and the *MDM4-FL* or *MDM4-S* isoforms by regulation of exon 6 inclusion or exclusion. Primers used to detect *MDM4* alternative splicing were indicated with black arrows. The premature termination codon generated by exon 6 exclusion is indicated with a red arrow.
(B) Sashimi plots of *MDM4* alternative splicing events between CTRL and DRGG1 iPSC.

(C) Semi-quantitative RT-PCR of *MDM4* transcripts in CTRL, DRGG1 and DRGG2. Four repeats were performed (n=4) and the exon 6 exclusion ratio is shown in percentage expressed as the mean ± SEM. **p* <0.05; ***p* <0.01; unpaired t test comparing each sample to the CTRL.
(D) RT-PCR for examining *MDM4* transcripts in siCTRL, siRBMX, and siPRMT5 cells. Three repeats were performed (n=3) and the exon 6 exclusion ratio is shown in percentage expressed as the mean ± SEM. ***p* <0.01; ****p* <0.001; unpaired t test comparing each sample to siCTRL.
(E) Immunoblotting of U2OS transfected with siCTRL, siRBMX and siPRMT5 with indicated antibodies.

2.4.4 RBMX functions in a complex with SRSF1 to regulate alternative splicing

To define higher-order complexes in which RBMX functions to regulate alternative splicing, we performed quantitative stable isotope labeling by amino acids in cell culture (SILAC) mass spectrometry. U2OS cells were either transiently transfected with vector alone (pcDNA3.1; light media) or pcDNA3.1 expressing FLAG-RBMX (heavy media, Figure 2.4A) and subjected to SILAC analysis. We identified 176 RBMX binding partners (FC(H/L)> 2) within two biological replicates (r=0.71 and p< 0.0001; Figure 2.4B). Gene Ontology (GO) analysis using STRING revealed that the RBMX interactome consisted of proteins involved in metabolic processes, RNA processing and splicing (Figure 2.4C). These interactors consisted of the hnRNP repertoire consistent with RBMX being hnRNP G (Liu and Dreyfuss, 1995); SR family proteins SRSF1, SRSF3, SRSF6, SRSF7, SRSF9 and TRA2β (SRSF10), and other known RBMX interactors such as Sam68 and T-STAR/SLM-2 or KHDRBS3 (Moursy et al., 2014).

We next examined the sequence of *cis*-elements near *MDM4* exon 6, and we identified several SRSF1 binding sites and one Sam68 binding site near the 5' weak splice site (Figure S2.2A). These findings suggested that RBMX complexes with SRSF1 and Sam68 to potentially regulate *MDM4* alternative splicing. We confirmed the RBMX interactions with Sam68 and SRSF1 using co-immunoprecipitations (Figure 2.4D, lanes 2 and 5) and the RBMX RGG/RG motif was required for optimal interaction, as RBMX-ΔRGG coimmunoprecipitated reduced Sam68 and SRSF1 levels (Figure 2.4D, lanes 3 and 6). To examine whether the interactions were direct or

mediated by bridging RNAs, cell lysates were treated with RNase I before the co-immunoprecipitation. The RNase I treatment abolished the RBMX/Sam68 interaction and had no effect on the RBMX/SRSF1 interaction (Figure 2.4E, compare lanes 5 and 6), suggesting RNA is bridging the RBMX/Sam68 interaction, but the RBMX/SRSF1 is a direct protein-protein interaction.

We next examined the involvement of Sam68 and SRSF1 in *MDM4* exon 6 exclusion. U2OS cells were transfected with siRNAs against RBMX, Sam68 and SRSF1 and the inclusion/exclusion of *MDM4* exon 6 was monitored by RT-PCR. *MDM4* exon 6 exclusion increased significantly in siRBMX (47.4%), siSam68 (42.1%), and siSRSF1 (65.9%) compared to siCTRL cells (23.2%, Figure 2.4F, 2.4G). Next, we searched ENCODE, an available eCLIP database (Van Nostrand et al., 2016), and identified SRSF1 binding reads, but not Sam68 (KHDRBS1) within *MDM4* exon 6 (Figure S2.2B). Therefore, we decided to focus on the RBMX/SRSF1 interaction. First, RNA immunoprecipitations (RIPs) were performed to confirm SRSF1 and RBMX-binding to *MDM4* exon 6. U2OS cells were transfected with expression vectors encoding GFP-SRSF1, FLAG-RBMX and a minigene pMDM4 expressing a genomic fragment spanning *MDM4* exons 5 to 7. The cells were crosslinked with formaldehyde and processed for immunoprecipitations with control (IgG), anti-GFP and anti-RBMX antibodies. RT-qPCR analysis revealed a significant enrichment of the *MDM4* mRNA associated with GFP-SRSF1 and FLAG-RBMX immunoprecipitations compared to control IgG (Figure 2.4H). To examine whether RBMX directly bound the *MDM4* exon 6, we performed PAR-CLIP (photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation) analysis. We confirmed that GFP-SRSF1 was preferentially enriched at exon 6 (position b) and at nearby 3' and 5' splice sites (positions a and c) rather than flanking consecutive exons (positions e and f;

Figure S2.2C). This specificity was specific for SRSF1, and not RBMX (Figure S2.2C), suggesting RBMX is not directly bound to *MDM4* exon 6, consistent with the absence of RBMX PAR-CLIP reads from previous studies at this location (Liu et al., 2017; Munschauer et al., 2018). Importantly, depletion of RBMX or PRMT5 reduced GFP-SRSF1 binding to the *MDM4* mRNA (Figure 2.4I). These findings suggest that the methylation of RBMX by PRMT5 facilitates access of SRSF1 to the *MDM4* mRNA (Figure 2.4J).

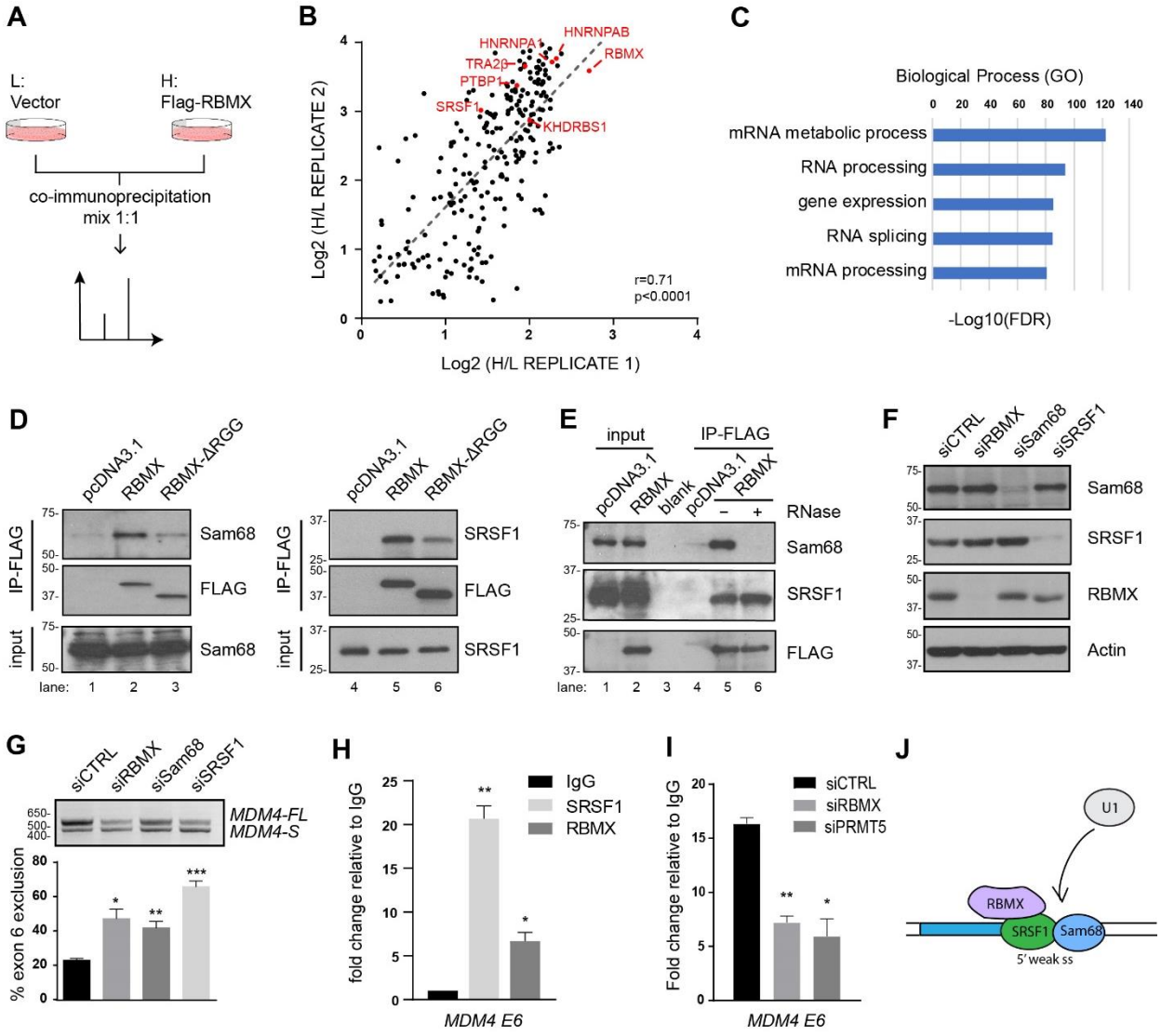


Figure 2. 4 RBMX RGG/RG motif mediates interaction with Sam68 and SRSF1 for RNA binding. (A) Schematic representation of SILAC-based RBMX-interactome analysis.

(B) Correlation analysis of two biological repeats of SILAC-based RBMX-interactome. U2OS cultured in Light (Lys 0; Arg 0) or Heavy (Lys 4, Arg 6) medium were transfected with pcDNA3.1 or FLAG-RBMX. Immunoprecipitation was performed with equal amount of FLAG-M2 beads and mixed at the 1:1 ratio for MS/MS analysis.
(C) GO analysis of potential RBMX binding proteins by STRING. Bar plot presents the top five enriched Biological Process and the significance are presented with -Log10(FDR).
(D) HEK293 were transiently transfected with pcDNA3.1, FLAG-RBMX and FLAG-RBMX-ΔRGG. Extracts were subjected to co-immunoprecipitation analysis using anti-FLAG antibody and the bound proteins immunoblotted after SDS-PAGE with anti-Sam68, -SRSF1 or -FLAG antibodies as indicated.
(E) Same as in panel (D) except the lysates were either incubated or not with RNase I before the anti-FLAG immunoprecipitation step.
(F) U2OS cells were transfected with siRNAs and analyzed by immunoblotting with indicated antibodies.
(G) RT-PCR was performed to examine *MDM4* transcripts in siCTRL, siRBMX, siSam68 and siSRSF1 U2OS cells. Histogram denotes % exclusion. n =3 independent experiments. * p <0.05; ** p <0.01; *** p <0.001. Unpaired t-test comparing each sample to the siCTRL. mean ± SEM.
(H) U2OS cells transfected with GFP-SRSF1, FLAG-RBMX and the minigene pMDM4 were cross-linked with 1% formaldehyde and RIP performed using IgG, anti-GFP and anti-FLAG antibodies respectively. RT-qPCR was used to assess the bound *MDM4* RNA expressed as fold-enrichment over IgG. n=2 independent experiments. * p <0.05; ** p <0.01; unpaired t-test. mean ± SEM.
(I) U2OS cells transfected with indicated siRNA were further transfected with GFP-SRSF1 and pMDM4. RIP was performed and RT-qPCR was used to assess *MDM4* recovery in each RIP sample, expressed as fold-enrichment over pull-down with IgG. n=3 independent experiments. * p <0.05, ** p <0.01, unpaired t-test comparing each sample to the siCTRL. mean ± SEM.
(J) RBMX facilitates Sam68 and SRSF1 binding to the 5' splice site of *MDM4* exon 6. The binding of Sam68 and SRSF1 recruits U1 snRNP and facilitate exon 6 inclusion.

2.4.5 RBMX forms membraneless organelles and colocalizes with SRSF1

Cation-π interaction between arginine and aromatic rings is known to form LLPS and influence membraneless organelle dynamics (Hofweber et al., 2018). First, we examined whether RBMX forms condensates, as the low complexity region of RBMX contains a significant number of arginine and tyrosine residues. Indeed, RBMX formed distinct nuclear compartments in mouse hippocampal neurons (Figure S2.3A). Similar foci were observed in U2OS cells stably expressed FLAG-RBMX, but not with FLAG-RBMX-ΔRGG (Figure S2.3B). Notably, RBMX co-localized with SRSF1 in these nuclear structures (Figure 2.5A). We reasoned that these RBMX membraneless structures would be sensitive to 1,6-hexanediol (1,6-HD) (Kroschwald et al., 2015), which disrupts hydrophobic interactions, but not its inactive isomer 2,5-hexanediol (2,5-HD). Cells treated with 1,6-HD for 10 min significantly reduced the size of the RBMX foci by 43.8%

(Figure 2.5B) and their intensity by 62.48% (Figure 2.5C). Moreover, the colocalization of RBMX and SRSF1 was reduced by ~25% (Figure 2.5D). The 1,6-HD effects observed where not due to the expression level changes of RBMX and SRSF1, as confirmed by immunoblotting (Figure 2.5E). To determine if the modulation of SRSF1 alternative splicing capacity was regulated by multivalent interactions, we performed RNA immunoprecipitations (RIPs) of SRSF1 and monitored its association with *MDM4* exon 6 in the presence of 1,6-HD. Cells treated with 1,6-HD reduced the binding capacity of SRSF1 to the *MDM4* mRNA by ~50% compared with cells treated with 2,5-HD (Figure 2.5F). These data suggest that the SRSF1/RBMX co-localization to membraneless structures is necessary for the association of SRSF1 with the *MDM4* mRNA.

As arginine methylation regulates membraneless structures (Guccione and Richard, 2019), we investigated whether it was involved in the formation of RBMX membraneless structures. Interestingly, RBMX foci in PRMT5-deficient cells were reduced in size by ~33% (Figure 2.5G, 2.5H) and their intensity was reduced by ~35% (Figure 2.5I). Furthermore, the colocalization of RBMX and SRSF1 was diminished in siPRMT5 cells by ~42 % (Figure 2.5J). Furthermore, to examine whether RBMX membraneless structures are implicated in *MDM4* alternative splicing, we performed colocalization studies between the *MDM4* nascent pre-mRNA and RBMX. U2OS cells transfected with pMDM4 and an expression plasmid encoding FLAG-RBMX were treated with α-amanitin. Cells were stained with single molecule fluorescence in situ hybridization (smFISH) probes spanning 2 kb flanking exon 6 of the *MDM4* nascent mRNA and anti-FLAG antibodies. Indeed, we observed colocalization of the *MDM4* nascent mRNA and FLAG-RBMX in punctate foci within the nuclear compartment (Figure 2.5K). These observations suggest nascent mRNAs are present within RBMX membraneless structures, which may serve as the basis

for alternative splicing regulation. Taken together, these findings suggest methylation by PRMT5 plays an important role in localizing the RBMX/SRSF1 complex in 1,6-HD-sensitive structures necessary for *MDM4* splicing function.

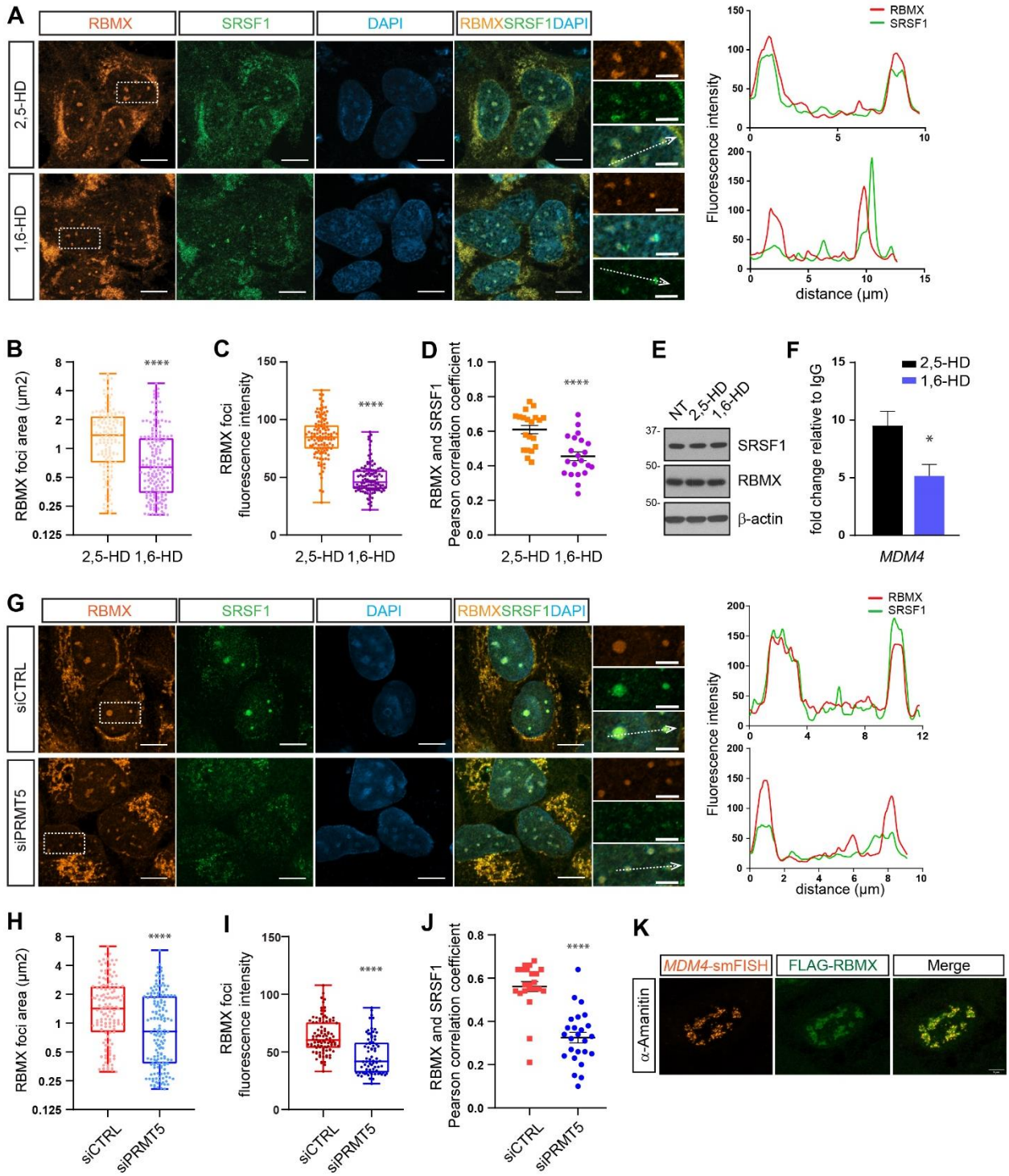


Figure 2. 5 RBMX colocalizes with SRSF1 in nuclear granules and 1,6-HD treatment or PRMT5 depletion disrupts the colocalization.
(A) Immunostaining of RBMX and SRSF1 in U2OS cells fixed with 100% methanol. Scale bar,10 μm. Insert scale bar, 4 μm. Histogram shows the colocalization of RBMX and SRSF1 within the RIO (Region of Interesting) indicated by dash arrow.
(B, C, D) Quantification of RBMX foci size (n=142 foci; B), RBMX foci intensity (mean intensity per foci, n=114 foci; C), and RBMX and SRSF1 Pearson correlation coefficient (n=20 cells; D) were analyzed. **** p <0.0001; unpaired t-test. mean ± SEM.
(E) U2OS cells were treated with 0.3M 2,5-hexanediol (HD) or 1,6-HD for 10 min and analyzed by immunoblotting with indicated antibodies.
(F) U2OS cells were transfected with GFP-SRSF1 and pMDM4 constructs for 48 h. Then cells were treated with 0.3M 2,5-hexanediol (HD) or 1,6-HD for 10 min. RIP were performed and *MDM4* recovery in each RIP sample was expressed as fold-enrichment over pull-down with IgG. n=3 independent experiments.* p <0.05, unpaired t-test. mean ± SEM.
(G) Immunostaining of RBMX and SRSF1 in U2OS cells fixed with 100% methanol. Scale bar,10 μm. Insert, scale bar, 4 μm. Histogram shows the colocalization of RBMX and SRSF1 within the RIO indicated by dash arrow.
(H, I, J) Quantification of RBMX foci. RBMX foci size (n=147 foci; G), RBMX foci intensity (mean intensity per foci- background intensity, n=76 foci; H), and RBMX and SRSF1 Pearson correlation coefficient (n=22 cells, I). **** p <0.0001. unpaired t-test. mean ± SEM.
(K) U2OS cells were transfected with pMDM4 and FLAG-RBMX. 24 h after transfection, the cells were treated with 20 μg/ml α-amanitin for another 9 h and fixed with 4% paraformaldehyde. smFISH was performed using probes labeled with Cy3 to the *MDM4* exon 6 and the flanking introns. Anti-FLAG antibody was used to detect the localization of FLAG-RBMX. The images were taken using a Zeiss LSM800 confocal microscope. The scale bars denote 5 μm.

2.4.6 Shashi-XLID iPSC-derived neurons have increased p53 activation

To explore the neuronal defects of patients with Shashi-XLID, we differentiated the CTRL and DRGG iPSCs into NPCs using the dual SMAD-inhibition protocol (Chambers et al., 2009). The CTRL and DRGG iPSCs were efficiently differentiated into NPCs with >90% SOX1+, SOX2+ or PAX6+ expression (Figure S2.4A, B). Transcriptome analysis revealed 258 and 15 genes to be downregulated and upregulated, respectively, in DRGG1 NPCs compared to CTRL NPCs. GO analysis in the differentially regulated genes identified embryo development, embryo morphogenesis, sensory organ development and central nervous system (CNS) development as major biological processes (Figure S2.4C). Heatmaps for genes involved in CNS development identified transcription factors essential for neurogenesis such as FOXG1, TBR1 and EMX1, and the vesicular glutamate transporter gene SLC17A7 to be downregulated in DRGG1 NPCs (Figure

2.6A). The downregulation of FOXG1, TBR1 and SLC17A7 was confirmed in DRGG1 and DRGG2 by RT-qPCR (Figure 2.6B), suggesting DRGG NPCs should have profound defects in neurogenesis. Alternative splicing analysis identified 111 significant altered splicing events in DRGG NPCs, including 89 skipped exons (SE), 3 events of alternative 3' splice sites (A3SS), 11 events of alternative 5' splice sites (A5SS), 3 events of mutually exclusive exons (MXE) and 3 events of retained introns (RI). Nineteen altered splicing events were identified in both DRGG1 iPSCs and NPCs. Among them, 17 events were SE and the essential neuronal developmental factors PTPRZ1, CHCHD7, RTN4, and RAI14 confirmed by RT-PCR (Figure S2.5).

Interestingly, the differential p53 activation observed in iPSCs was lost in NPCs maintained in media with growth factors, as the p53 pathway was elevated in both CTRL and DRGG NPCs (Figure 2.6C), as seen by non-changing levels of CDKN1A for example (Figure 2.6D, left panel). However, the 14-day differentiation of NPCs to cortical neurons re-established a p53 pathway differential with CDKN1A and BAX being significantly higher in DRGG cultures (Figure 2.6D, right panel). Furthermore, *MDM4* exon 6 exclusion was significantly increased in DRGG cultures of 14-day differentiated neurons, but not in growth factor rich NPCs (Figure 2.6E, left and middle panels). An increased in p53 levels was confirmed in the DRGG differentiated neurons by immunoblotting and β-actin served as a loading control (Figure 2.6E, right panel). We next examined whether the activated p53 pathway in 14-day differentiated neurons triggered cell death, as measured by cleaved caspase 3/7-positive cells. Indeed, the activated p53 pathway observed in DRGG cultures resulted in increased active caspases, suggestive of apoptosis (Figure 2.6F).

To examine whether the loss of arginine methylation of the RGG/RG motif of RBMX contributed to the defects observed in DRGG neurogenesis, we differentiated CTRL NPCs into cortical neurons in the presence of EPZ015666 (Figure S2.6A), a PRMT5-specific inhibitor

(Chan-Penebre et al., 2015). Interestingly, the p53 pathway was activated in EPZ015666-treated CTRL cultures (Figure S2.6B) with an increase in cleaved caspase 3/7 activity (Figure S2.6C), suggesting increase apoptosis. Moreover, FOXG1 and TBR1 were down-regulated by EPZ015666 (Figure S2.6B), as observed in DRGG cultures. These findings suggest that EPZ015666 treated neurons, similar to DRGG cultures, exhibit p53 pathway activation, leading to increased apoptosis with deregulation of FOXG1 and TBR1 expression.

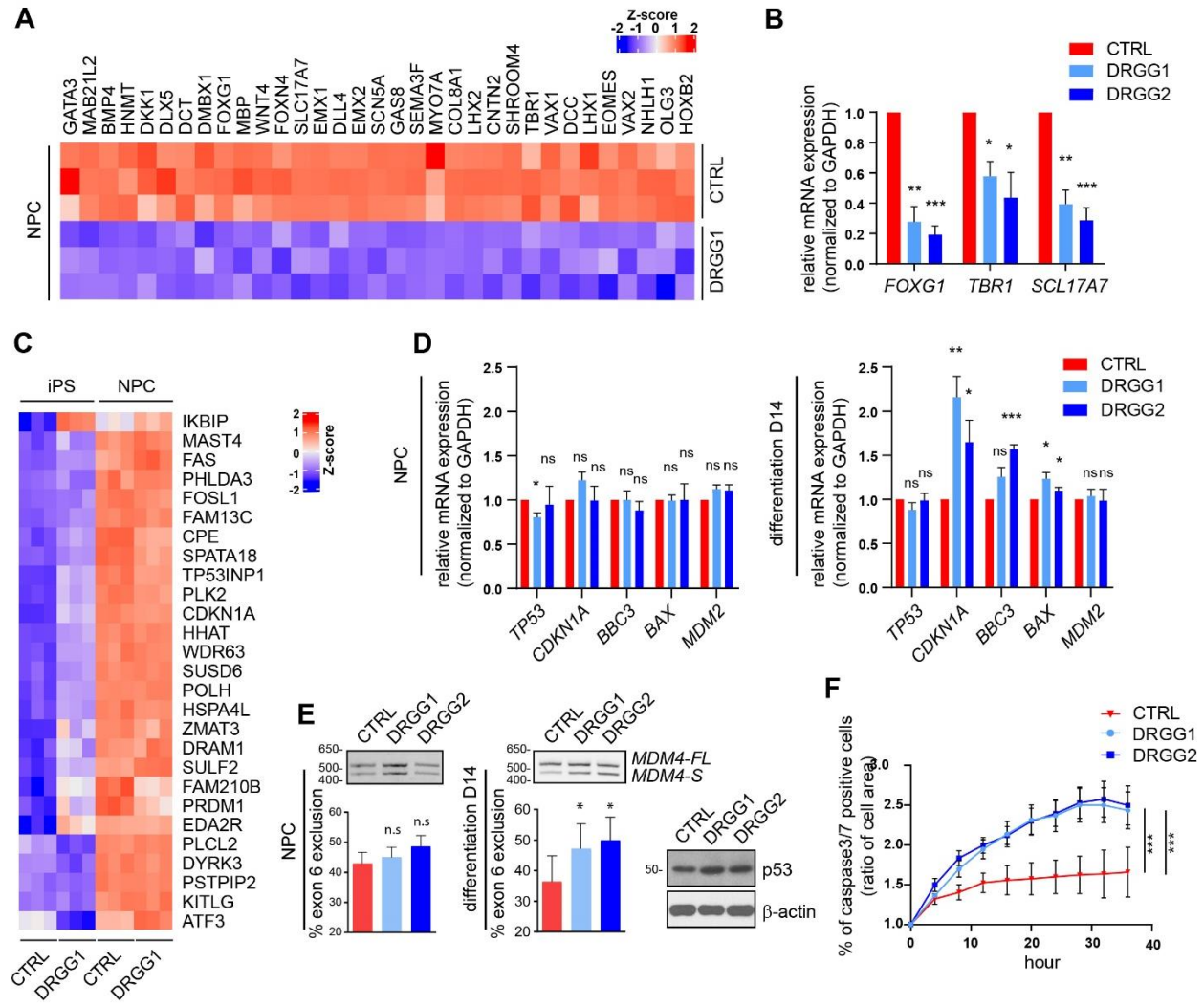


Figure 2. 6 Aberrant regulation of neuronal genes in NPCs and dysregulation of p53 pathway during neurogenesis.

(A) Heatmap displays the expression change of genes listed in the "central nervous system development" in CTRL and DRGG1 NPCs. Neuronal genes with 2-fold change in DRGG1 versus CTRL are presented. Row Z-scores are shown by color. Blue, repressed; red, induced.

(B) RT-qPCR of *FOXG1*, *TBR1* and *SCL17A7* expression in CTRL, DRGG1 and DRGG2 at early differentiation stage (differentiation D7). Normalized to GAPDH. n= 3 independent experiments. * p <0.05; ***p <0.001; ****p <0.0001, unpaired t test. mean ± SEM.
(C) Heatmap displays p53 target genes expression change in DRGG1 and CTRL hiPSC and NPC. p53 consensus targets with 1.5-fold change in IPS versus NPC are presented. Row Z-scores are shown by color. Blue, repressed; red, induced.
(D) RT-qPCR assays were performed for *p53*, *CDKN1A*, *BBC3*, *BAX* and *MDM2* in RBMX- DRGG1 and CTRL at the NPC stage and after 14 days differentiation. The data were normalized to GAPDH. n= 3 independent experiments. * p <0.05; **p <0.01; ***p <0.001, unpaired t test. mean ± SEM.
(E) RT-PCR for *MDM4* exon 6 inclusion/exclusion was performed for CTRL, DRGG1 and DRGG2 NPCs (left panel) and 14 days differentiated neurons (middle panel). Histogram presents the percentage of *MDM4* exon 6 exclusion. n= 2 independent experiments. * p <0.05, paired t test. mean ± SEM. Cell lysates of 14 days differentiated neurons was subjected to immunoblotting for p53 and β-actin (right panel).
(F) The percentage of cleaved caspase 3 (CC-3/7) positive cells within 36 h culture was measured by Incucyte. The graph represents data from two independent experiments. *** p <0.001; paired t test. mean ± SEM.

2.4.7 RBMX-ΔRGG deficiency perturbs neuronal differentiation

We next differentiated NPCs for four weeks into forebrain lineages known to generate GABAergic and glutamatergic cortical interneurons (Bardy et al., 2015; Maroof et al., 2013). We stained the CTRL and DRGG cultures for TUJ1, a neuron-specific class III β-tubulin antibody, to confirm their neuronal identity and to assess their overall morphology. We noted that the number of TUJ1+ neurons was reduced in the DRGG cultures versus CTRL cultures, suggesting decreased neurogenesis (Figure 2.7A). Furthermore, cell morphology as assessed by Sholl analysis revealed a significant increase in unipolar and dipolar neurons (CTRL, 60.44%; DRGG1, 76.52%; DRGG2, 78.74%) with a decrease of multipolar neurons in DRGG cultures (CTRL, 39.56%; DRGG1, 23.48%; DRGG2, 21.26%) (Figure 2.7B). Furthermore, the multipolar neurons had less branchings and shorter neurites (Figure 2.7C, Figure S2.7A). RNA-seq data showed the vesicular glutamate transporter genes SLC17A7 and SLC17A6 were down-regulated in DRGG NPC cultures. Furthermore, we were unable to detect VGLUT1-positive neurons in 4-week differentiated RBMX-ΔRGG cultures, while ~15% of the neurons were VGLUT1-positive in the CTRL cultures (Figure 2.7D). Next, we examined GABAergic neurons, but no significant

difference was observed between DRGG and CTRL cultures (Figure S2.7B). These data show that the genetic lesion in Shashi-XLID, deleting the RGG/RG motif of RBMX, leads to neuronal morphology defects with reduced ability to mature into glutamatergic VGLUT1-positive neurons.

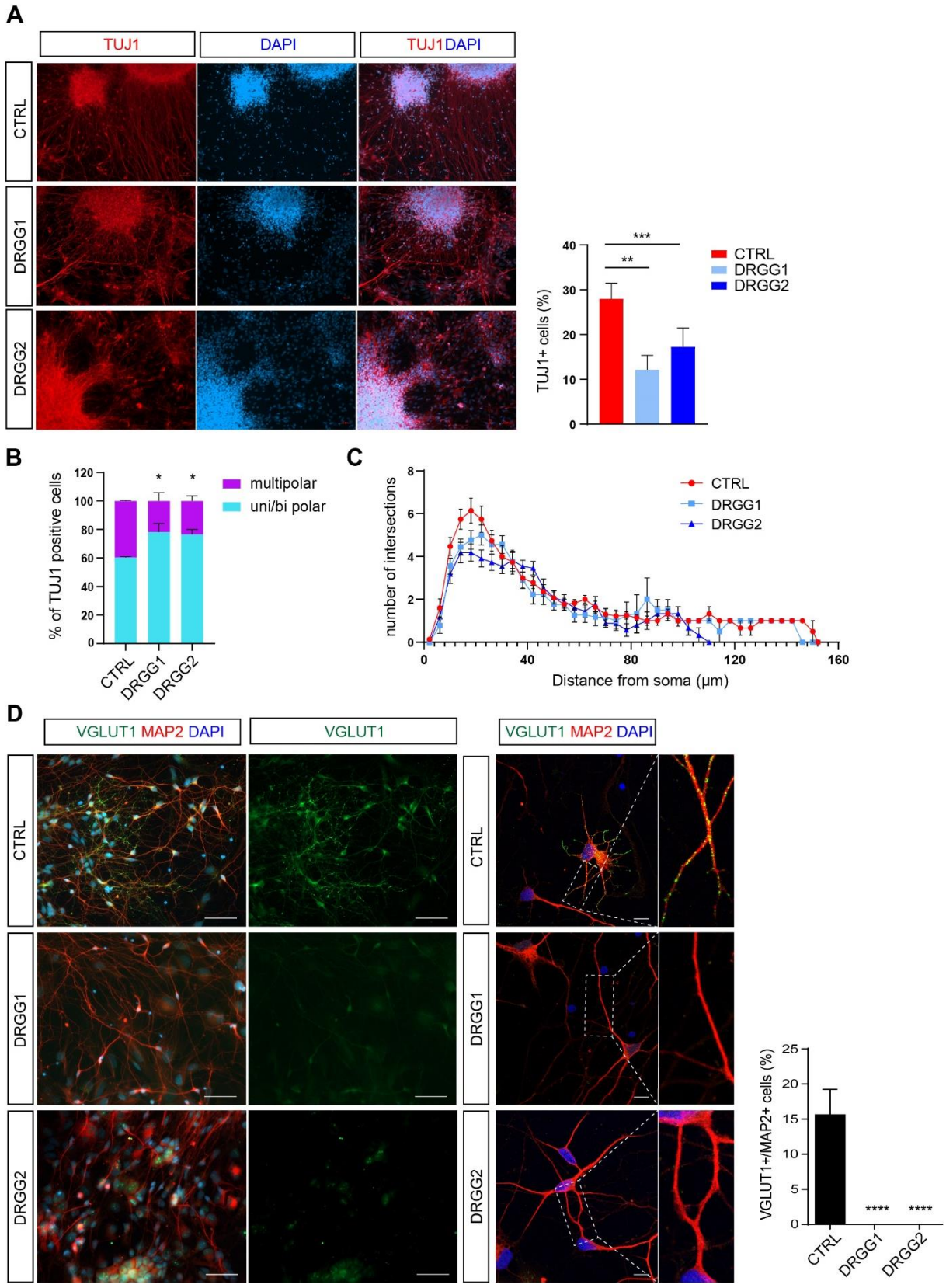


Figure 2. 7 RBMX-ΔRGG hiPSC-derived neurons display altered morphology.

(A) TUJ1 staining of morphologies of CTRL and RBMX-ΔRGG hiPSC-derived neurons. Histogram presents quantification of TUJ1 positive cells over total cell number in CTRL, DRGG1 and DRGG2. ** p <0.01; *** p <0.001; unpaired t test. mean ± SEM.
(B) Quantification of uni/bipolar and multipolar neurons within TUJ1 positive population in CTRL, DRGG1 and DRGG2. * p <0.05; unpaired t test. mean ± SEM.
(C) Sholl analysis of multipolar neurons in CTRL and RBMX-ΔRGG, n = 10-20 neurons were analyzed and presented as mean ± SEM.
(D) Immunostaining of VGLUT1 and MAP2 positive neurons in CTRL and RBMX-ΔRGG cells (left panel scale bar 50 μm, middle panel scale bar 10 μm). Quantification of VGLUT1 positive neurons in CTRL and RBMX-ΔRGG (right panel). **** p <0.0001; unpaired t test. mean ± SEM

2.5 Discussion

We present evidence that the Shashi-XLID syndrome is a splicing disorder with activation of the p53 pathway due to the loss-of-function of the RGG/RG motif of RBMX. Splicing regulation by RBMX was RGG/RG motif-dependent, as it facilitated higher-order complex assembly with other splicing factors such as SRSF1 into membraneless organelles. Moreover, differentiation of RBMX-DRGG NPCs into forebrain lineage neurons showed defects in VGLUT1-positive glutamatergic neurogenesis potentially explaining the molecular basis for Shashi-XLID.

As alternative splicing is required for brain development and function (Naro et al., 2021), it is not surprising that a large number of genes involved in XLID syndromes are RBPs. In the RBMX-DRGG iPSCs we generated alternative splicing defects in axon guidance and neurogenesis genes were observed. For instance, RTN4 and PTPRZ1, which were alternatively spliced in both RBMX-DRGG iPSC and NPCs, are known to be responsible for neuronal branching and the formation of perineuronal nets (Fawcett et al., 2019; Pollen et al., 2015). Our data suggest that RBMX regulates alternative splicing complexed with SRSF1 and Sam68 splicing factors. In fact, SRSF1 and Sam68 play key roles in the regulation of neuronal functions. SRSF1 is expressed in dorsal root ganglia sensory neurons and contributes to the development of neuropathic pain (Hulse et al., 2016). Mutations in SRSF1 binding sites within genes such as ColQ causes congenital myasthenic syndromes (Rahman et al., 2015). Moreover, Sam68 modulates neuronal

activity by regulating the alternative splicing of synaptic receptor Nrxn1 pre-mRNA (Iijima et al., 2011). A recent study showed that RBMX bound to m6A modified exonic regions and regulated their alternative splicing (Zhou et al., 2019). The C-terminus RGG/RG motif was required for RBMX binding to the m6A modified RNA (Liu et al., 2017). Analysis of the RNA-seq data by Liu et al., 2017 showed that depletion of RBMX in HEK293T cells also leads to *MDM4* exon 6 exclusion, consistent with our observations herein. The RBMX alternative splicing of *MDM4* exon 6 likely occurs in the absence of m6A modification, however, as *MDM4* exon 6 was not observed to contain this modification (Liu et al., 2017; Zhou et al., 2019).

We showed that the RBMX C-terminal RGG/RG motif was the site of arginine methylation by PRMT5. Depletion of PRMT5 in hESCs causes proliferation defects and prevents self-renewal, but not pluripotency (Gkountela et al., 2014). Conditional knockdown of Prmt5 using Nestin-Cre leads to increased apoptosis, smaller brain size and postnatal death, as a result of splicing defects especially at the *Mdm4* locus (Bezzi et al., 2013). Similar to PRMT5 depletion, RBMX-deficiency caused dysregulation of *MDM4* splicing and activation of the p53 pathway, leading to increased caspase 3 activation.

Accumulating evidence indicates that alternative splicing is likely regulated by LLPS (Ying et al., 2017). RBMX regulates the function of SRSF1 by modulating its ability to bind certain of its RNA binding elements, for example near *MDM4* exon 6. Several RNPs containing low complexity regions were reported to form membraneless organelles and the most characterized of which is Fused in sarcoma (FUS) (Patel and Lee, 2015). Liquid phase condensation of FUS causes neurotoxicity in amyotrophic lateral sclerosis (Qamar et al., 2018). RGG/RG motif within FUS low complexity region is required for it to phase separate and its hypomethylation was observed in FUS associated frontotemporal lobar degeneration (Qamar et al., 2018). We observed

significant dissociation of SRSF1 from RBMX foci when PRMT5 was depleted by siRNA. Since PRMT5 has been shown to regulate many RBPs (Guccione and Richard, 2019), it is formally possible that PRMT5 deregulating SRSF1/RBMX foci maybe contributed by the activity of other RBPs or the whole spliceosome machinery. Moreover, post-translational modification of SRSF1 by phosphorylation has been shown to alter its subnuclear distribution altering splicing defects patterns (Leva et al., 2012). We showed that treatment of α-amanitin induced RBMX colocalization with the *MDM4* pre-mRNA exon 6 region, indicating nascent mRNAs are present in the RBMX membraneless structures. Interestingly, recent studies of early transcription steps revealed that low levels of RNA at regulatory elements promote condensate formation, whereas high levels of RNA during elongation stimulates condensate dissolution (Henninger et al., 2021). Thus, further studies will be required to understand the details of LLPS separation of RBMX and SRSF1 and alternative splicing of nascent RNAs.

Our data suggest that p53 activation is a consequence of alternative splicing defects observed as a result of subnuclear compartment disruption. p53 activation causes genes involved in the neuronal development to be dysregulated in the DRGG NPCs including essential transcriptional factors that control forebrain development including FOXG1 which promotes telencephalic neuroepithelial cells proliferation (Hanashima et al., 2002) and TBR1 which directs radial glia to postmitotic projection neurons (Englund et al., 2005). Indeed, previous work using a p53 activated mice model to study relevant developmental phenotypes revealed that activation of the p53 pathway in neural crest cells (NCCs) increased apoptosis and downregulated genes involved in neurogenesis (Bowen et al., 2019). This led to dramatically decreased number of NCCs, cells that contribute to the dorsal root ganglia, the enteric nervous system, and the sympathetic ganglia (Bowen et al., 2019).

Aberrant neurite outgrowth and disruption of the balance between excitatory and inhibitory synapses is known to be affected in XLID syndromes including Fragile X syndrome (FXS), and XLID associated by disruptions in KIAA2022, IL1RAPL1, and ZDHHC9 genes (Montani et al., 2017; Shimell et al., 2019; Van Maldergem et al., 2013). iPSCs derived from patients with FXS show defects in initial neurite outgrowth (Doers et al., 2014), and exaggerated type I mGluR activation, a marker of excitatory synapses (Achuta et al., 2017). Neurite outgrowth is a fundamental process during neuronal differentiation conducted by complex intracellular signaling (Khodosevich and Monyer, 2010). Aberrant neurite outgrowth in DRGG neuron cultures may result from splicing disorders and gene expression dysregulation. Vesicular glutamate transporter (VGLUT1) activity and expression has been implicated in neurological and psychiatric diseases, and their loss causes severe cognitive malfunctions and lethality (Wojcik et al., 2004). We observed a downregulation of SLC17A7 mRNA in DRGG NPCs and the loss of the VGLUT1 subpopulation in the DRGG differentiated neuronal cultures.

In sum, our work demonstrates how deletion of the RBMX RGG/RG motif disturbs alternative splicing of specific isoforms, leading to activation of the p53 pathway and causing neuronal VGLUT1 differentiation defects. These findings contribute to understand the molecular function of RBMX and establish the causal links between splicing dysregulation and Shashi-XLID etiology.

2.6 Methods

Cell lines

U2OS and HEK293 cells were purchased from ATCC. Cells were cultured in Dulbecco's modified Eagle's medium supplement with 10% v/v fetal bovine serum (FBS) at 37°C with 5% CO2.

hiPSCs were derived from peripheral blood mononuclear cells obtained from a 42 year-old healthy male donor. Reprograming was performed by Axol biosciences and Episomal vectors were used to generate virus-free induced pluripotent stem cells. The cell line was cultured under feeder-free condition in TeSR-E8 medium (05990, Stem cell technologies). CRISPR/CAS 9 deletion of RBMX RGG/RG motif was achieved by using electroporation with the following guide RNA 5'-CUU CCC CCU UCU AUG GAA AG-3' (synthesis by Synthego). Karyotyping of CTRL and DRGG1 was carried out by SickKids using Affymetrix CytoScan HD array.

siRNAs, plasmids and transfection

All siRNAs were purchased from Dharmacon. siRNA sequences are as follows:

siRBMX#1:5'-UGG CAA AUA UGG ACG AAU AdTdT-3',

siRBMX#2: 5'-UCA AGA GGA UAU AGC GAU AdTdT-3',

siPRMT5: 5'-UGG CAC AAC UUC CGG ACU UdTdT-3'.

The siRNA 5'-CGU ACG CGG AAU ACU UCG AdTdT-3', targeting the firefly luciferase (GL2) was used as control (siCTRL). 20 nM siRNA was used for all knockdown transfections. The N-terminal FLAG-tagged mouse RBMX plasmid was constructed by inserting a FLAG-coding sequence into the pcDNA3.1 (+) vector. The FLAG-RBMX-ΔN, FLAG-RBMX-RK and FLAG-RBMX-ΔRGG were constructed by PCR and the murine RBMX constructs were resistant to the RBMX #1 and #2 siRNAs. Stable cell lines were generated by transfecting U2OS with FLAG-RBMX-ΔN, FLAG-RBMX-RK and FLAG-RBMX-ΔRGG. For pMDM4 minigene construct, *MDM4* gene sequence spanned exon 5 to exon 6 was amplified from U2OS genomic DNA.

In vitro methylation assay

GST-tagged RBMX fragments and mutants were purified from bacteria. 10 μg of each GST-tagged fragments was incubated with 1 μl of (methyl-^{3}H) S-adenosyl-L-methionine solution (15

Ci/mmol stock solution, 0.55 μM final concentration, Perkin-Elmer) and 1 μg of PRMT5:MEP50 active complex (SRP0145, Sigma-Aldrich) in methylation buffer (HEPES pH 7.4 50mM, NP40 0.01%, DTT 1 mM, PMSF 1 mM) for 2h at 37°C. The reaction was visualized by fluorography.

Cell lysis and immunoprecipitation

For co-immunoprecipitation experiments, cells were lysed with a lysis buffer containing 50 mM HEPES, pH 7.4, 150 mM NaCl, 1% Triton X-100 and a cocktail of protease inhibitors and phosphatase inhibitors. Cell lysates were sheared with 26G needles followed with high-speed centrifugation to remove cell debris, then the supernatant was precleared with Protein A Sepharose and incubated with FLAG M2 beads for another hour at 4ºC. For immunoprecipitation treated with RNase I, cells lysate was divided equally and incubated with or without 50 U/ ml RNase I (Ambion, AM2294) for 30 min at 37 ºC followed by FLAG M2 beads for 1 hour at 4ºC. The beads were washed four times with lysis buffer for western blot analysis or four times with lysis buffer and then twice with PBS for mass spectrometry analysis. For immunoprecipitation under denaturing conditions, cells were lysed with denaturing lysis buffer containing 1% SDS and 5 mM EDTA. Cell lysates were sonicated and diluted with Triton X-100 lysis buffer before proceeding to beads incubation.

SILAC (stable isotope labeling with amino acids in cell culture) and MS/MS mass analysis

U2OS were grown in standard light medium (Lys0/Arg0) and heavy (Lys4/Arg6) SILAC medium for 5 passages and then transfected with pcDNA3.1 and FLAG-RBMX respectively. The cells were lysed and subjected to immunoprecipitation with the anti-FLAG M2 beads. The beads with bound proteins were proceed for MS/MS mass analysis as described previously (Thandapani et al., 2013b).

Neural induction

iPSC colonies were dissociated and plated for monolayer culture at a final concentration of 2 × 10^5 cells/cm^2. Neuronal induction was performed using STEMdiff SMADi Neuronal Induction Kit (08581, Stem cell technologies). Then STEMdiff Neural Progenitor Medium (05833, Stem cell technologies) were used to maintain the NPCs and subculture every 3 to 5 days.

Neuronal differentiation

NPCs of passages 2 to 4 only were differentiated into cortical neurons according to the reported protocol (Maroof et al., 2013). Briefly, NPCs were plated on sterile coverslips or 24-well plates coated with poly-L-ornithine (P4957, Sigma-Aldrich) and 5 μg/ml laminin (L2020, Sigma-Aldrich) solution. The next day, the medium was changed to neuron differentiation medium based on DMEM/F12 with N2 supplement (1×, Gibco) and B27 supplement (1×, Gibco). After 1 week culture, the medium was replaced with neuron mature medium (Bardy et al. 2015) based on Neural Basal media (Gibco) containing N2 supplement (1×, Gibco) and B27 supplement (1×, Gibco), GDNF (20 ng/ml, R&D), BDNF (20 ng/ml R&D), and dibutyryl cyclic-AMP (500 μg/ml, Sigma Aldrich).

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

RNA immunoprecipitation (RIP) was performed as previously described with minor modification (Rinn et al., 2007). Briefly, cells were cross-linked with a final concentration of 1% formaldehyde, washed twice with ice cold PBS and resuspended in RIP buffer (150mM KCl, 25mM Tris, pH 7.4, 5mM EDTA, 0.5mM DTT, 0.5% NP40, and 100U/mL RNase inhibitor). Chromatin was sheared by sonication, and DNA fragments were digested with TURBO™ DNase (Thermo fisher, AM2238) at the 37°C for 30 min. Cell debris were pelleted by centrifugation. For each immunoprecipitation, 1 μg of antibody was added, and the samples were incubated for 4 h at 4°C

with rotation. Protein A Sepharose beads (Sigma) were added, and the samples were incubated at 4°C for 1h with rotation. The beads were pelleted by centrifugation, resuspended, and washed in RIP buffer for five times. RNA was eluted in Proteinase K buffer (50 mM Tris, pH 7.5, 75 mM NaCl, 6.5 mM EDTA, and 1% SDS) supplemented with proteinase K and incubated at 50°C for 30 min. RNA was recovered by using 5 volumes TRIzol™ LS Reagent (Thermo Fisher). Equal volume of RNA from each sample was used for the reverse transcription. RT-qPCR was performed using primers targeting *MDM4* exon 6.

PAR-CLIP was performed as previously described with minor modification (Hafner et al., 2010). pMDM4 and FLAG-RBMX or GFP-SRSF1 were co-transfected in HEK293 and 24 h post-transfection, 100 μM 4-thiouridine was added for 16 h and then cross-linked with 0.15 J/cm^2 365 nm UV. Cells were then washed twice with ice cold PBS and resuspended in lysis buffer (150mM KCl, 25mM Tris, pH 7.4, 5mM EDTA, 0.5mM DTT, 0.5% NP40, and 100U/mL RNase inhibitor). After incubation for 20 min the cell debris was cleared by centrifugation. Cell lysates were incubated with 1 U/ml RNase I at 37 °C for 3 min. For each immunoprecipitation, 40 U RNase inhibitor and 2 μg of antibody was added, and the samples were incubated for 2 h at 4°C with rotation. Protein A Sepharose beads (Sigma) were then added, and the samples incubated at 4°C for 1h with rotation. The beads were pelleted by centrifugation, resuspended, and washed in high salt wash buffer three times and once with lysis buffer. After removing the final wash buffer, DNA fragments were digested with 2U TURBO™ DNase (Thermo fisher, AM2238) at the 37°C for 4 min. RNA was eluted in Proteinase K buffer (50 mM Tris, pH 7.5, 75 mM NaCl, 6.5 mM EDTA, and 1% SDS) supplemented with proteinase K and incubated at 50°C for 30 min. The RNA was recovered by using 5 volumes of TRIzol™ LS Reagent (Thermo Fisher). Equal volume of RNA from each sample was used for the reverse transcription.

smFISH

Custom DNA probe sets labeled with Cy3 were designed using Stellaris Probe Designer, synthetized by Biosearch Technologies. The sequence of the probes is listed in Supplementary Table 4. Hybridization buffer, wash buffer A and wash buffer B were purchased from Biosearch Technologies. And smFISH was performed according to the manufacturer's protocol. Cells were transfected with pMDM4 and FLAG-RBMX and 24 h later, the cells were treated with 20 μg/ml α-amanitin for another 9 h. Then the cells were fixed with 4% paraformaldehyde, and washed once with wash buffer A. The cells were hybridized with 50 pmol probes and co-stained with anti-FLAG M2 antibody for 6 h at 37°C with hybridization buffer. Samples were then washed with wash buffer A once, and nuclei stained for DAPI and secondary antibody with wash buffer A for 30 min. Finally, samples were washed with wash buffer B and mounted with ProLong Diamond antifade reagent and visualized by fluorescence microscopy.

RNA sequencing and data analysis

iPSC and NPC RNA samples were purified using GenElute™ Mammalian Total RNA Miniprep Kit (RTN70, Sigma Aldrich). RNA sequencing libraries were generated using TruSeq Stranded mRNA Sample Prep Kit with TruSeq Unique Dual Indexes (Illumina, San Diego, CA). Resulting libraries were multiplexed and sequenced with 100 base pair (bp) paired-end reads (PE100) to a depth of approximately 60 million reads per sample on an Illumina HiSeq 4000. Samples were demultiplexed using bcl2fastq v2.20 Conversion Software (Illumina, San Diego, CA). Illumina adapter sequences were first removed from the paired-end reads of length 101 through the Trimmomatic v0.39 (Bolger et al., 2014) software. The reads were then aligned to the human genome (hg38/GRCh38) with STAR v2.7.1a (Dobin et al., 2013) and the gene expression was quantified for all samples through HOMER v4.11.1 (Heinz et al., 2010). Pathway enrichment

analysis of genes with significant differential expression was carried out with GSEA (Subramanian et al., 2005) (Mootha et al., 2003)and visualized as heatmaps with the ComplexHeatmap R package (Gu et al., 2016).

Alternative splicing analysis

Alternative Splicing (AS) analysis of skipped exons (SE), mutually exclusive exons (MXE), retained introns (RI), alternative 5'-end splice sites (A5SS) and alternative 3'-end splice sites (A3SS) was performed through the rMATS v4.0.2 software with default settings (Shen et al., 2014) using Gencode v19 gene annotations (Frankish et al., 2019). Significant AS events were called at a false discovery rate (FDR) cut-off of 0.05. Sashimi plots illustrating the inclusion level difference between conditions were generated through the rmats2sashimiplot program (https://github.com/Xinglab/rmats2sashimiplot).

Quantification and statistical analysis

Experiments presented here were repeated at least three times, except specified otherwise. All data are presented as mean ± SEM. Significance of comparison between two groups was assessed either by the unpaired or paired Student-t test. The use of the specific tests has been reported in each figure legend. Statistically significant results were defined as follows: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$.

2.7 References

Achuta, V.S., Grym, H., Putkonen, N., Louhivuori, V., Karkkainen, V., Koistinaho, J., Roybon, L., and Castren, M.L. (2017). Metabotropic glutamate receptor 5 responses dictate differentiation of neural progenitors to NMDA-responsive cells in fragile X syndrome. Dev Neurobiol 77, 438-453.

Adamson, B., Smogorzewska, A., Sigoillot, F.D., King, R.W., and Elledge, S.J. (2012). A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response. Nat Cell Biol 14, 318-328.

Bardy, C., van den Hurk, M., Eames, T., Marchand, C., Hernandez, R.V., Kellogg, M., Gorris, M., Galet, B., Palomares, V., Brown, J., et al. (2015). Neuronal medium that supports basic

synaptic functions and activity of human neurons in vitro. Proc Natl Acad Sci U S A 112, E2725-2734.
Becker, J.S., McCarthy, R.L., Sidoli, S., Donahue, G., Kaeding, K.E., He, Z., Lin, S., Garcia, B.A., and Zaret, K.S. (2017). Genomic and Proteomic Resolution of Heterochromatin and Its Restriction of Alternate Fate Genes. Mol Cell 68, 1023-1037 e1015.
Bedford, M.T., and Clarke, S.G. (2009). Protein arginine methylation in mammals: who, what, and why. Mol Cell 33, 1-13.
Bezzi, M., Teo, S.X., Muller, J., Mok, W.C., Sahu, S.K., Vardy, L.A., Bonday, Z.Q., and Guccione, E. (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, 1903-1916.
Blackwell, E., Zhang, X., and Ceman, S. (2010). Arginines of the RGG box regulate FMRP association with polyribosomes and mRNA. Hum Mol Genet 19, 1314-1323.
Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114-2120.
Bowen, M.E., McClendon, J., Long, H.K., Sorayya, A., Van Nostrand, J.L., Wysocka, J., and Attardi, L.D. (2019). The Spatiotemporal Pattern and Intensity of p53 Activation Dictates Phenotypic Diversity in p53-Driven Developmental Syndromes. Dev Cell 50, 212-228 e216.
Chambers, S.M., Fasano, C.A., Papapetrou, E.P., Tomishima, M., Sadelain, M., and Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. Nat Biotechnol 27, 275-280.
Chan-Penebre, E., Kuplast, K.G., Majer, C.R., Boriack-Sjodin, P.A., Wigle, T.J., Johnston, L.D., Rioux, N., Munchhof, M.J., Jin, L., Jacques, S.L., et al. (2015). A selective inhibitor of PRMT5 with in vivo and in vitro potency in MCL models. Nat Chem Biol 11, 432-437.
Dichmann, D.S., Fletcher, R.B., and Harland, R.M. (2008). Expression cloning in Xenopus identifies RNA-binding proteins as regulators of embryogenesis and Rbmx as necessary for neural and muscle development. Dev Dyn 237, 1755-1766.
Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15-21.
Doers, M.E., Musser, M.T., Nichol, R., Berndt, E.R., Baker, M., Gomez, T.M., Zhang, S.C., Abbeduto, L., and Bhattacharyya, A. (2014). iPSC-derived forebrain neurons from FXS individuals show defects in initial neurite outgrowth. Stem Cells Dev 23, 1777-1787.
Elliott, D.J., Oghene, K., Makarov, G., Makarova, O., Hargreave, T.B., Chandley, A.C., Eperon, I.C., and Cooke, H.J. (1998). Dynamic changes in the subnuclear organisation of pre-mRNA splicing proteins and RBM during human germ cell development. J Cell Sci 111 (Pt 9), 1255-1265.
Ellison, J.W., Rosenfeld, J.A., and Shaffer, L.G. (2013). Genetic basis of intellectual disability. Annu Rev Med 64, 441-450.
Englund, C., Fink, A., Lau, C., Pham, D., Daza, R.A., Bulfone, A., Kowalczyk, T., and Hevner, R.F. (2005). Pax6, Tbr2, and Tbr1 are expressed sequentially by radial glia, intermediate progenitor cells, and postmitotic neurons in developing neocortex. J Neurosci 25, 247-251.
Fawcett, J.W., Oohashi, T., and Pizzorusso, T. (2019). The roles of perineuronal nets and the perinodal extracellular matrix in neuronal function. Nat Rev Neurosci 20, 451-465.
Fischer, M. (2017). Census and evaluation of p53 target genes. Oncogene 36, 3943-3956.

Frankish, A., Diekhans, M., Ferreira, A.M., Johnson, R., Jungreis, I., Loveland, J., Mudge, J.M., Sisu, C., Wright, J., Armstrong, J., et al. (2019). GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res 47, D766-D773.
Gkountela, S., Li, Z., Chin, C.J., Lee, S.A., and Clark, A.T. (2014). PRMT5 is required for human embryonic stem cell proliferation but not pluripotency. Stem Cell Rev Rep 10, 230-239.
Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32, 2847-2849.
Guccione, E., and Richard, S. (2019). The regulation, functions and clinical relevance of arginine methylation. Nat Rev Mol Cell Biol 20, 642-657.
Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jr., Jungkamp, A.C., Munschauer, M., et al. (2010). Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell 141, 129-141.
Hanashima, C., Shen, L., Li, S.C., and Lai, E. (2002). Brain factor-1 controls the proliferation and differentiation of neocortical progenitor cells through independent mechanisms. J Neurosci 22, 6526-6536.
Heinrich, B., Zhang, Z., Raitskin, O., Hiller, M., Benderska, N., Hartmann, A.M., Bracco, L., Elliott, D., Ben-Ari, S., Soreq, H., et al. (2009). Heterogeneous nuclear ribonucleoprotein G regulates splice site selection by binding to CC(A/C)-rich regions in pre-mRNA. J Biol Chem 284, 14303-14315.
Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38, 576-589.
Henninger, J.E., Oksuz, O., Shrinivas, K., Sagi, I., LeRoy, G., Zheng, M.M., Andrews, J.O., Zamudio, A.V., Lazaris, C., Hannett, N.M., et al. (2021). RNA-Mediated Feedback Control of Transcriptional Condensates. Cell 184, 207-225 e224.
Hofmann, Y., and Wirth, B. (2002). hnRNP-G promotes exon 7 inclusion of survival motor neuron (SMN) via direct interaction with Htra2-B1. Human Molecular Genetics 11, 2037-2049.
Hofweber, M., Hutten, S., Bourgeois, B., Spreitzer, E., Niedner-Boblenz, A., Schifferer, M., Ruepp, M.D., Simons, M., Niessing, D., Madl, T., et al. (2018). Phase Separation of FUS Is Suppressed by Its Nuclear Import Receptor and Arginine Methylation. Cell 173, 706-719.
Hulse, R.P., Drake, R.A., Bates, D.O., and Donaldson, L.F. (2016). The control of alternative splicing by SRSF1 in myelinated afferents contributes to the development of neuropathic pain. Neurobiol Dis 96, 186-200.
Iijima, T., Wu, K., Witte, H., Hanno-Iijima, Y., Glatter, T., Richard, S., and Scheiffele, P. (2011). SAM68 Regulates Neuronal Activity-Dependent Alternative Splicing of Neurexin-1. Cell 147, 1601-1614.
Khodosevich, K., and Monyer, H. (2010). Signaling involved in neurite outgrowth of postnatally born subventricular zone neurons in vitro. BMC Neurosci 11, 18.
Kroschwald, S., Maharana, S., Mateju, D., Malinovska, L., Nuske, E., Poser, I., Richter, D., and Alberti, S. (2015). Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules. Elife 4, e06807.

Lenos, K., Grawenda, A.M., Lodder, K., Kuijjer, M.L., Teunisse, A.F., Repapi, E., Grochola, L.F., Bartel, F., Hogendoorn, P.C., Wuerl, P., et al. (2012). Alternate splicing of the p53 inhibitor HDMX offers a superior prognostic biomarker than p53 mutation in human cancer. Cancer Res 72, 4074-4084.
Leva, V., Giuliano, S., Bardoni, A., Camerini, S., Crescenzi, M., Lisa, A., Biamonti, G., and Montecucco, A. (2012). Phosphorylation of SRSF1 is modulated by replicational stress. Nucleic Acids Res 40, 1106-1117.
Liu, N., Zhou, K.I., Parisien, M., Dai, Q., Diatchenko, L., and Pan, T. (2017). N6-methyladenosine alters RNA structure to regulate binding of a low-complexity protein. Nucleic Acids Res 45, 6051-6063.
Liu, Q., and Dreyfuss, G. (1995). In vivo and in vitro arginine methylation of RNA binding proteins. Mol Cell Biol 15, 2800-2808.
Maroof, A.M., Keros, S., Tyson, J.A., Ying, S.W., Ganat, Y.M., Merkle, F.T., Liu, B., Goulburn, A., Stanley, E.G., Elefanty, A.G., et al. (2013). Directed differentiation and functional maturation of cortical interneurons from human embryonic stem cells. Cell Stem Cell 12, 559-572.
Meister, G., Eggert, C., Buhler, D., Brahms, H., Kambach, C., and Fischer, U. (2001). Methylation of Sm proteins by a complex containing PRMT5 and the putative U snRNP assembly factor pICln. Curr Biol 11, 1990-1994.
Mersaoui, S.Y., Yu, Z., Coulombe, Y., Karam, M., Busatto, F.F., Masson, J.Y., and Richard, S. (2019). Arginine methylation of the DDX5 helicase RGG/RG motif by PRMT5 regulates resolution of RNA:DNA hybrids. EMBO J *38*, e100986.
Montani, C., Ramos-Brossier, M., Ponzoni, L., Gritti, L., Cwetsch, A.W., Braida, D., Saillour, Y., Terragni, B., Mantegazza, M., Sala, M., et al. (2017). The X-Linked Intellectual Disability Protein IL1RAPL1 Regulates Dendrite Complexity. J Neurosci 37, 6606-6627.
Mootha, V.K., Lindgren, C.M., Eriksson, K.F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstrale, M., Laurila, E., et al. (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 34, 267-273.
Moursy, A., Allain, F.H., and Clery, A. (2014). Characterization of the RNA recognition mode of hnRNP G extends its role in SMN2 splicing regulation. Nucleic Acids Res 42, 6659-6672.
Munschauer, M., Nguyen, C.T., Sirokman, K., Hartigan, C.R., Hogstrom, L., Engreitz, J.M., Ulirsch, J.C., Fulco, C.P., Subramanian, V., Chen, J., et al. (2018). The NORAD lncRNA assembles a topoisomerase complex critical for genome stability. Nature 561, 132-136.
Naro, C., Cesari, E., and Sette, C. (2021). Splicing regulation in brain and testis: common themes for highly specialized organs. Cell Cycle 20, 480-489.
Patel, A., and Lee, H.O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M.Y., Stoy- nov, S., Mahamid, J., Saha, S., Franzmann, T.M., et al. (2015). A Liquid-to- Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. . Cell 162, 1066-1077.
Pieretti, M., Zhang, F.P., Fu, Y.H., Warren, S.T., Oostra, B.A., Caskey, C.T., and Nelson, D.L. (1991). Absence of expression of the FMR-1 gene in fragile X syndrome. Cell 66, 817-822.
Pollen, A.A., Nowakowski, T.J., Chen, J., Retallack, H., Sandoval-Espinosa, C., Nicholas, C.R., Shuga, J., Liu, S.J., Oldham, M.C., Diaz, A., et al. (2015). Molecular identity of human outer radial glia during cortical development. Cell 163, 55-67.
Qamar, S., Wang, G., Randle, S.J., Ruggeri, F.S., Varela, J.A., Lin, J.Q., Phillips, E.C., Miyashita, A., Williams, D., Strohl, F., et al. (2018). FUS Phase Separation Is Modulated by a

Molecular Chaperone and Methylation of Arginine Cation-pi Interactions. Cell 173, 720-734 e715.
Rahman, M.A., Azuma, Y., Nasrin, F., Takeda, J., Nazim, M., Bin Ahsan, K., Masuda, A., Engel, A.G., and Ohno, K. (2015). SRSF1 and hnRNP H antagonistically regulate splicing of COLQ exon 16 in a congenital myasthenic syndrome. Sci Rep 5, 13208.
Rallapalli, R., Strachan, G., Cho, B., Mercer, W.E., and Hall, D.J. (1999). A novel MDMX transcript expressed in a variety of transformed cell lines encodes a truncated protein with potent p53 repressive activity. J Biol Chem 274, 8299-8308.
Ren, J., Wang, Y., Liang, Y., Zhang, Y., Bao, S., and Xu, Z. (2010). Methylation of ribosomal protein S10 by protein-arginine methyltransferase 5 regulates ribosome biogenesis. J Biol Chem 285, 12695-12705.
Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Brugmann, S.A., Goodnough, L.H., Helms, J.A., Farnham, P.J., Segal, E., et al. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 129, 1311-1323.
Shashi, V., Xie, P., Schoch, K., Goldstein, D.B., Howard, T.D., Berry, M.N., Schwartz, C.E., Cronin, K., Sliwa, S., Allen, A., et al. (2015). The RBMX gene as a candidate for the Shashi X-linked intellectual disability syndrome. Clin Genet 88, 386-390.
Shen, S., Park, J.W., Lu, Z.X., Lin, L., Henry, M.D., Wu, Y.N., Zhou, Q., and Xing, Y. (2014). rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. Proc Natl Acad Sci U S A 111, E5593-5601.
Shimell, J.J., Shah, B.S., Cain, S.M., Thouta, S., Kuhlmann, N., Tatarnikov, I., Jovellar, D.B., Brigidi, G.S., Kass, J., Milnerwood, A.J., et al. (2019). The X-Linked Intellectual Disability Gene Zdhhc9 Is Essential for Dendrite Outgrowth and Inhibitory Synapse Formation. Cell Rep 29, 2422-2437 e2428.
Su, C.H., D, D., and Tarn, W.Y. (2018). Alternative Splicing in Neurogenesis and Brain Development. Front Mol Biosci 5, 12.
Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102, 15545-15550.
Thandapani, P., O'Connor, T.R., Bailey, T.L., and Richard, S. (2013a). Defining the RGG/RG motif. Mol Cell 50, 613-623.
Thandapani, P., O'Connor, T.R., Bailey, T.L., and Richard, S. (2013b). Defining the RGG/RG motif. Mol Cell 50, 613-623.
Toledo, F., and Wahl, G.M. (2006). Regulating the p53 pathway: in vitro hypotheses, in vivo veritas. Nat Rev Cancer 6, 909-923.
Tsang, B., Arsenault, J., Vernon, R.M., Lin, H., Sonenberg, N., Wang, L.Y., Bah, A., and Forman-Kay, J.D. (2019). Phosphoregulated FMRP phase separation models activity-dependent translation through bidirectional control of mRNA granule formation. Proc Natl Acad Sci U S A 116, 4218-4227.
Tsend-Ayush, E., O'Sullivan, L.A., Grutzner, F.S., Onnebo, S.M., Lewis, R.S., Delbridge, M.L., Marshall Graves, J.A., and Ward, A.C. (2005). RBMX gene is essential for brain development in zebrafish. Dev Dyn 234, 682-688.
Van Maldergem, L., Hou, Q., Kalscheuer, V.M., Rio, M., Doco-Fenzy, M., Medeira, A., de Brouwer, A.P., Cabrol, C., Haas, S.A., Cacciagli, P., et al. (2013). Loss of function of

KIAA2022 causes mild to severe intellectual disability with an autism spectrum disorder and impairs neurite outgrowth. Hum Mol Genet 22, 3306-3314.
Van Nostrand, E.L., Pratt, G.A., Shishkin, A.A., Gelboin-Burkhart, C., Fang, M.Y., Sundararaman, B., Blue, S.M., Nguyen, T.B., Surka, C., Elkins, K., et al. (2016). Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat Methods 13, 508-514.
Villarreal, O.D., Mersaoui, S.Y., Yu, Z., Masson, J.Y., and Richard, S. (2020). Genome-wide R-loop analysis defines unique roles for DDX5, XRN2, and PRMT5 in DNA/RNA hybrid resolution. Life Sci Alliance 3.
Wang, Y., Wang, J., Gao, L., Stamm, S., and Andreadis, A. (2011). An SRp75/hnRNPG complex interacting with hnRNPE2 regulates the 5' splice site of tau exon 10, whose misregulation causes frontotemporal dementia. Gene 485, 130-138.
Wojcik, S.M., Rhee, J.S., Herzog, E., Sigler, A., Jahn, R., Takamori, S., Brose, N., and Rosenmund, C. (2004). An essential role for vesicular glutamate transporter 1 (VGLUT1) in postnatal development and control of quantal size. Proc Natl Acad Sci U S A 101, 7158-7163.
Ying, Y., Wang, X.J., Vuong, C.K., Lin, C.H., Damianov, A., and Black, D.L. (2017). Splicing Activation by Rbfox Requires Self-Aggregation through Its Tyrosine-Rich Domain. Cell 170, 312-323 e310.
Zhang, G., Neubert, T.A., and Jordan, B.A. (2012). RNA binding proteins accumulate at the postsynaptic density with synaptic activity. J Neurosci 32, 599-609.
Zhang, X., Chen, M.H., Wu, X., Kodani, A., Fan, J., Doan, R., Ozawa, M., Ma, J., Yoshida, N., Reiter, J.F., et al. (2016). Cell-Type-Specific Alternative Splicing Governs Cell Fate in the Developing Cerebral Cortex. Cell 166, 1147-1162 e1115.
Zhou, K.I., Shi, H., Lyu, R., Wylder, A.C., Matuszek, Z., Pan, J.N., He, C., Parisien, M., and Pan, T. (2019). Regulation of Co-transcriptional Pre-mRNA Splicing by m(6)A through the Low-Complexity Protein hnRNPG. Mol Cell 76, 70-81 e79.

2.8 Supplemental information

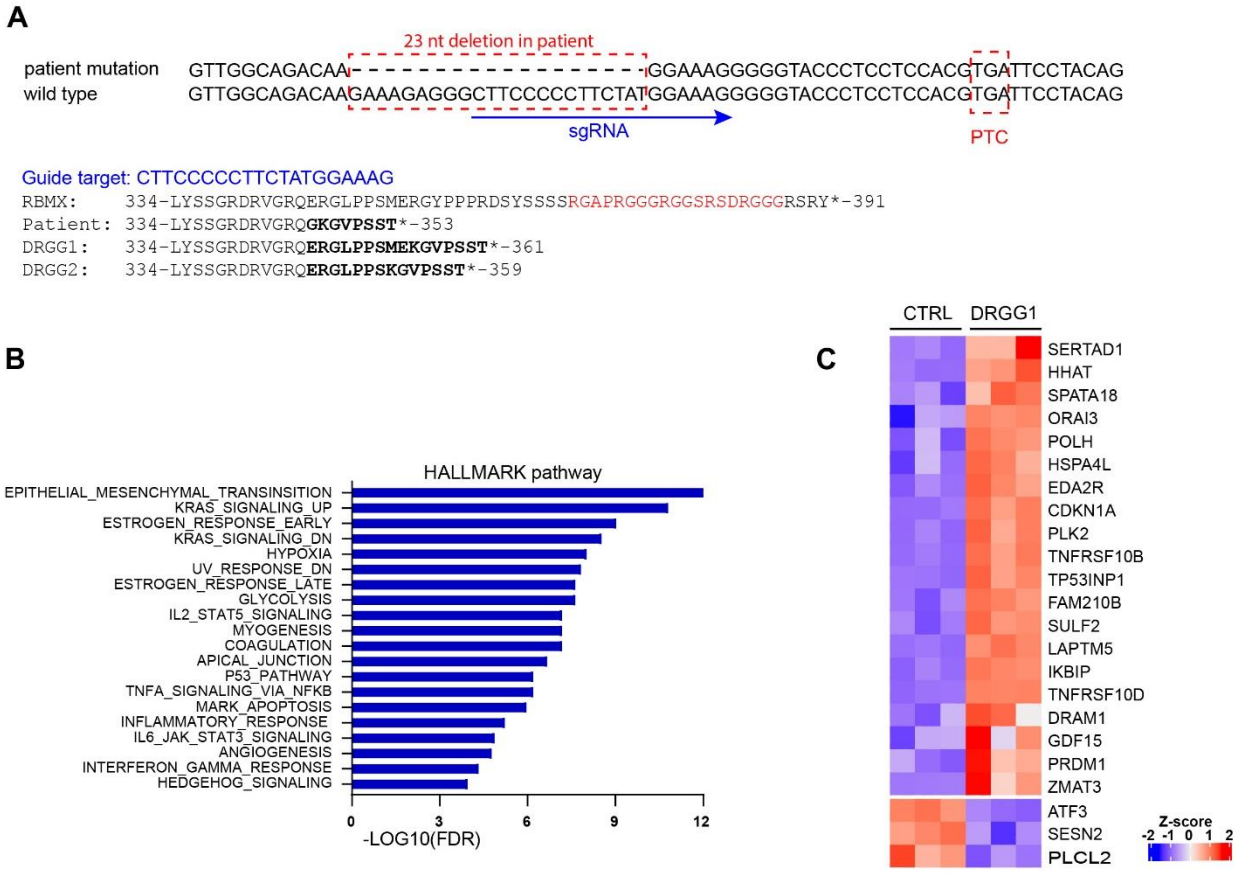
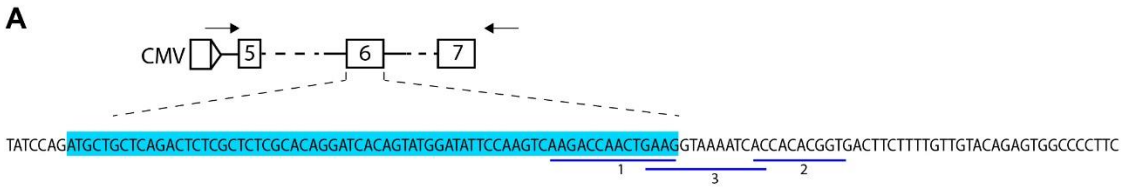
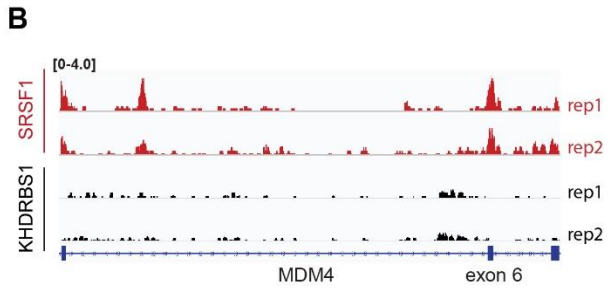


Figure S2. 1 Generation of RBMX-ΔRGG hiPSCs by CRISPR-CAS9.
(A) RBMX DNA sequences of the patients with Shashi-XLID. Its 23 nt deletion is shown as dash line and the wild type sequence is shown below. sgRNA targeted sequence is represented by a blue arrow. RBMX C-terminal (334-391aa) and patient mutation protein sequences are presented.
(B) Hallmark pathway analysis of significantly differential expressed genes(|FC|>1.5) in DRGG1 by GSEA. Bar plot presents the top twenty enriched Hallmark pathways and the significance are presented with -Log10(FDR).
(C) Heatmap of p53 target genes in DRGG1and CTRL. p53 consensus targets with 1.5 fold change in DRGG1 versus CTRL are shown. Row Z-score values are shown by color. Blue, repressed; Red, induced.



	***trans* factors**	***cis* elements**
site 1	SRSF1, SRSF5	AAG ACCAACTGAAG
site 2	SRSF1, SRSF5, SRSF6, SRSF9	ACCACACGGTG
site 3	U1 snRNP, Sam68	GAAGGTAAAATCA



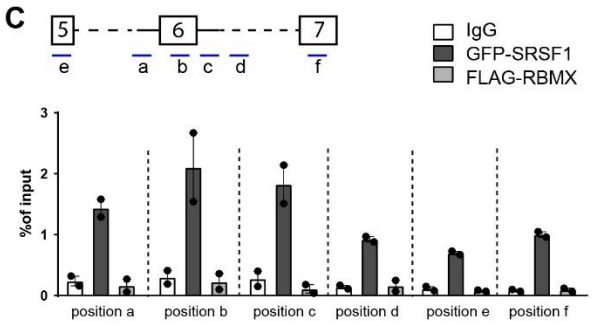


Figure S2. 2 SRSF1 binds to *MDM4* pre-mRNA.
(A) Schematic representation of MDM4 minigene. Primers designed for detecting the minigene transcripts are shown with arrows. The *MDM4* exon 6 sequence is highlighted in blue. The predicted SRSF1 and Sam68 binding sites (1-3) are shown.
(B) Snapshot of eClip data of SRSF1 and KDHRBS1 binding sites flanking *MDM4* exon 6 from IGV browser.
(C) U2OS cells were transfected with pMDM4 and GFP-SRSF1 or FLAG-RBMX for 48 h. PAR-CLIP were performed as described in Materials & Methods. The recovered *MDM4* RNA fragment (labelled position a-f) in each sample was expressed as percentage of input. n=2 independent experiments. mean ± SEM

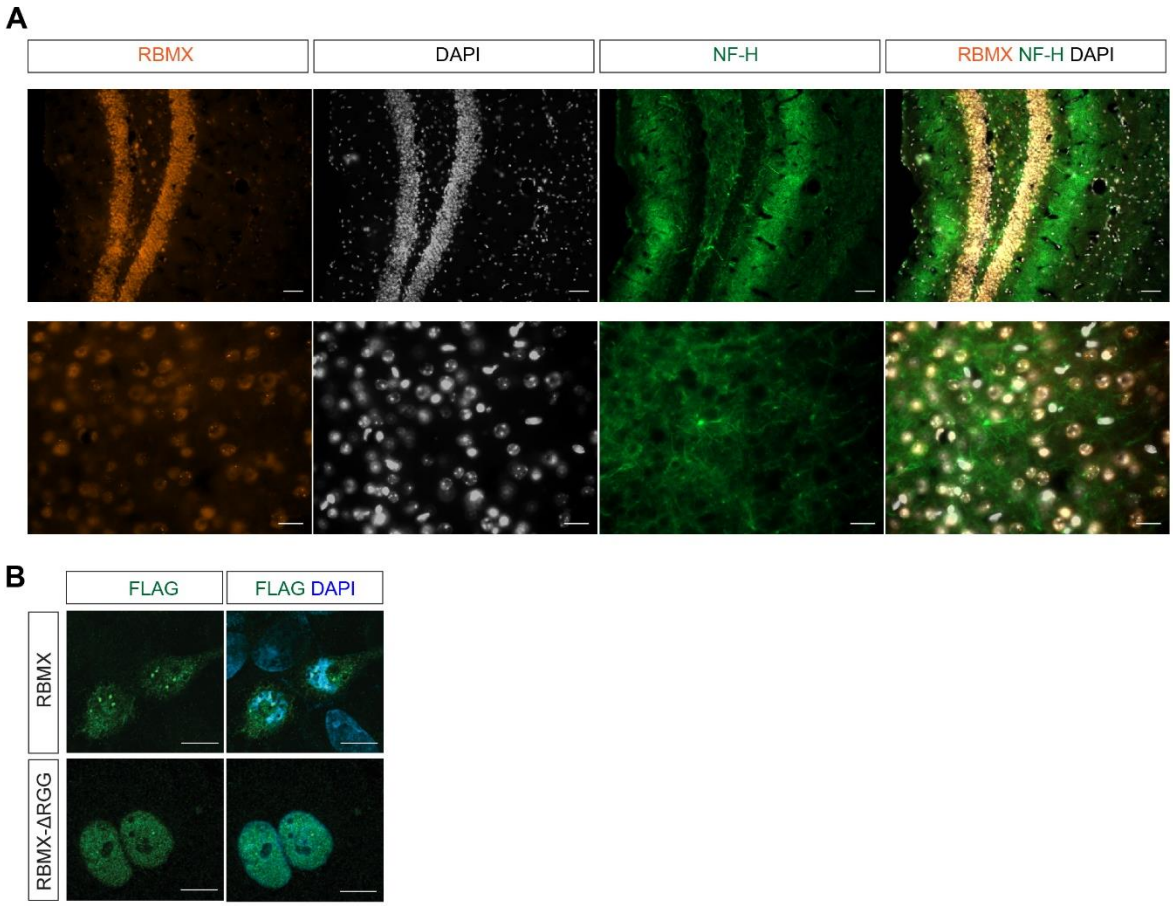


Figure S2. 3 Deletion of RGG/RG motif disrupts RBMX membraneless organelles formation.
(A) Immunostaining was performed on 8-week-old wild type mouse brain slices with anti-RBMX and anti-neurofilament heavy chain (NF-H) antibodies. The hippocampus is shown and the scale bars represent 50 μm (upper panels) and 20 μm (bottom panels).
(B) U2OS cell lines stably expressing FLAG-epitope tagged RBMX or RBMX-ΔRGG were methanol (100%) fixed followed by anti-FLAG M2 antibody staining. Scale bars denote 10 μm.

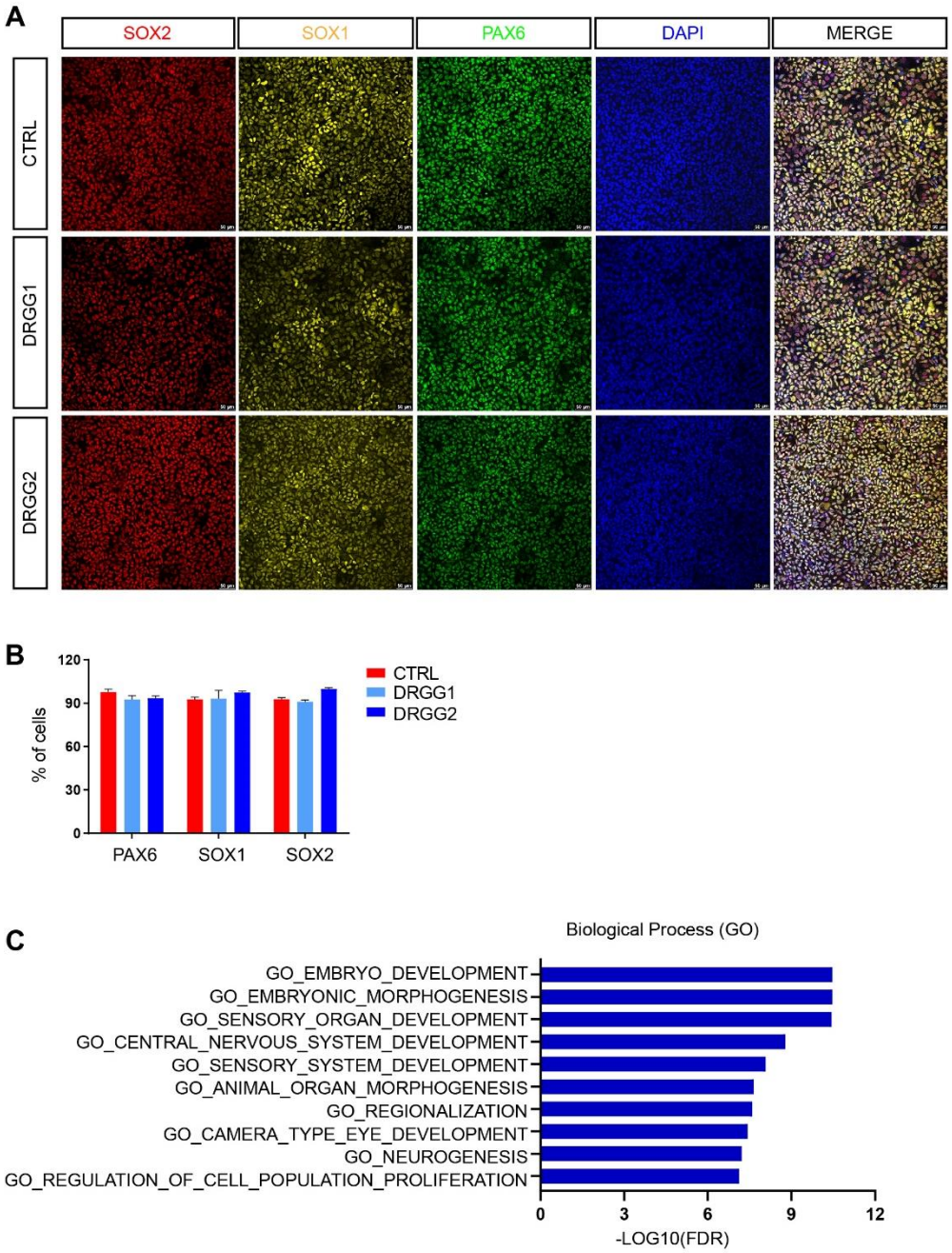


Figure S2. 4 Characterization of RBMX-ΔRGG NPCs.

(A) Immunofluorescence staining of NPC markers using anti-SOX2, -SOX1, and -PAX6 antibodies. Scale bar, 50 μm.

(B) Quantification of SOX1, SOX2 or PAX6 positive cells versus total cell number in CTRL, DRGG1 and DRGG2 from panel A.

(C) GO analysis of significantly down-regulated genes (FC< −2) in RBMX-ΔRGG NPCs by GSEA. Shown are the top 10 enriched biological processes and their significance as -log10(FDR).

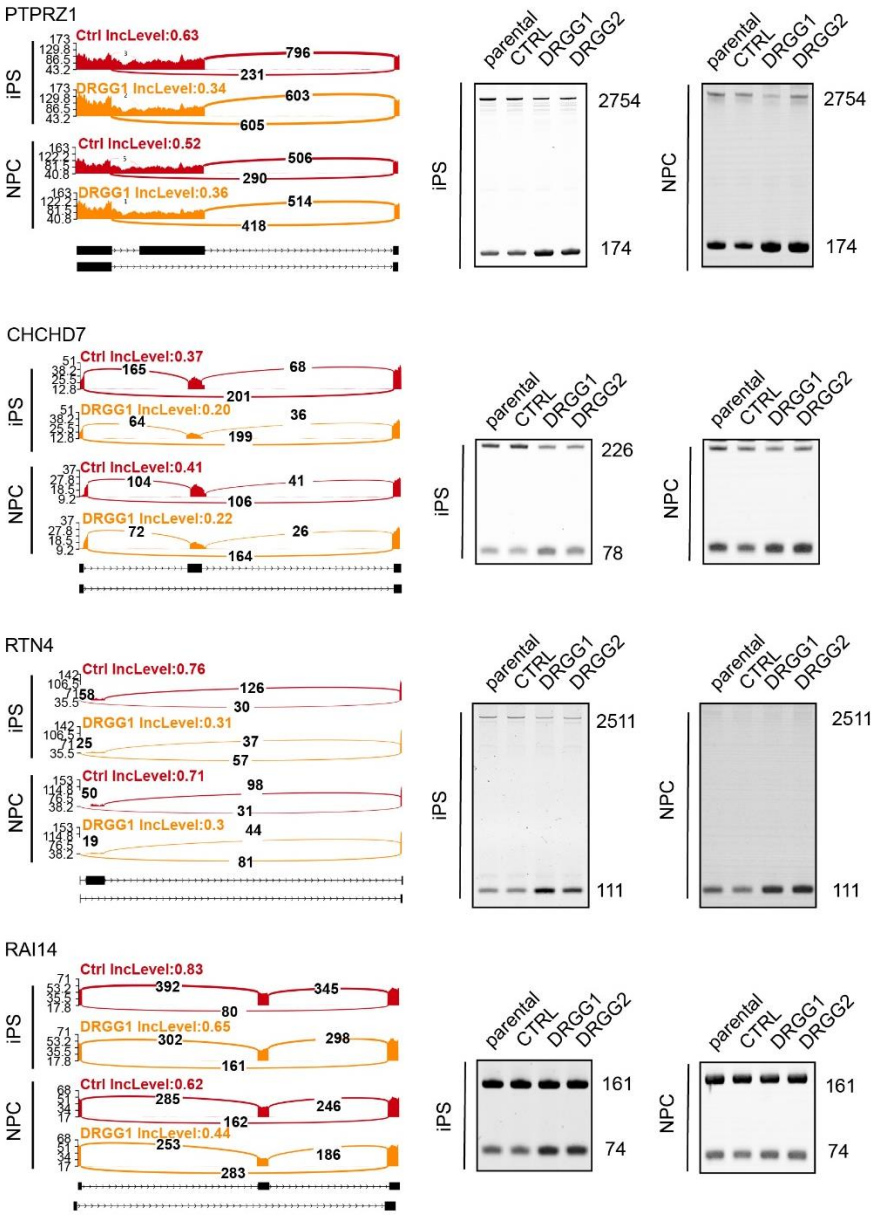


Figure S2. 5 Analysis of splicing dysregulation in RBMX-ΔRGG iPSC and NPC.
Selected alternative splicing events of PTPRZ1, CHCHD7, RTN4, RAII14 are shown in Sashimi plot format (left) and validated by RT-PCR in iPSCs (middle) and NPCs (right) from parental cells (cells used to generate CTRL and DRGGs), CTRL, DRGG1 and DRGG2. The length of expected the DNA fragments for exon inclusion (upper) and exon exclusion (lower) is indicated on the right in bp.

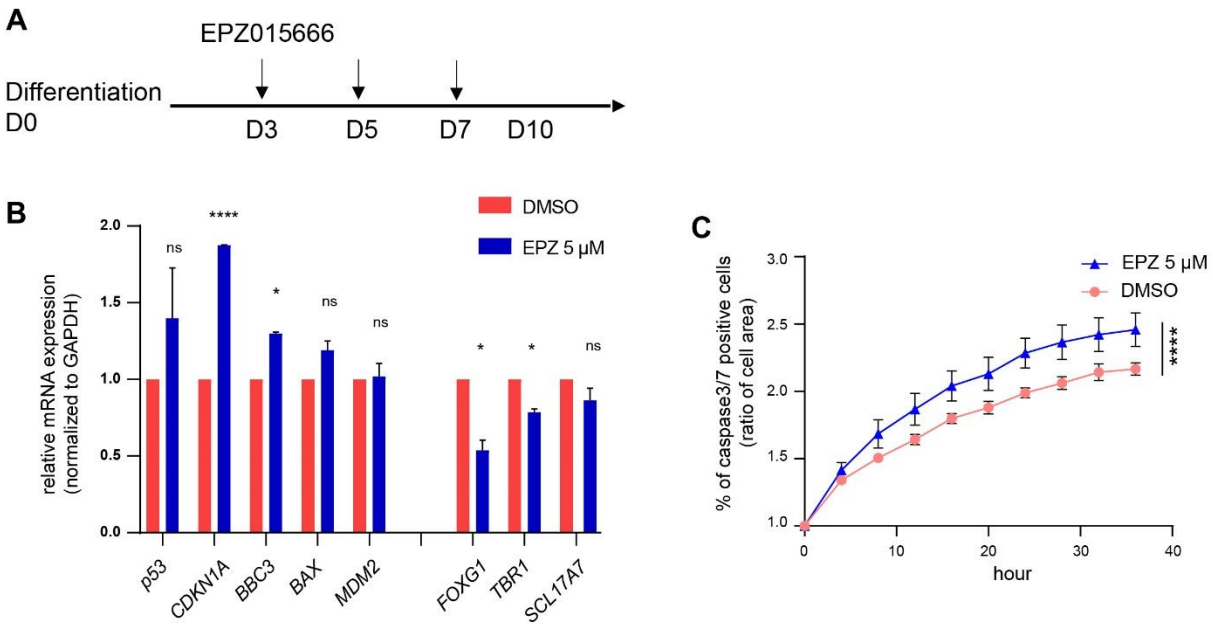


Figure S2. 6 PRMT5 inhibitor treatment in differentiating neurons activates the p53 pathway.
(A) Schematic representation of 5 µM EPZ15666 treatments at days 3, 5, and 7 in differentiating neurons. The cells were harvested for analysis at the differentiation day 10 (D10).
(B) RT-qPCR analysis for p53 and target genes (CDKN1A, BBC3, BAX and MDM2) and FOXG1, TBR1,and SCL17A7 on differentiation D10 with or without treatment with EPZ15666. n= 2 independent experiments. * $p < 0.05$; **** $p < 0.0001$; ns, not significant. unpaired t test. mean ± SEM.
(C) The percentage of cleaved casepase 3/7 positive cells as a ratio of cell area in 36h cultures was measured by Incucyte. n= 2 independent experiments. **** $p < 0.0001$; paired t test. mean ± SEM.

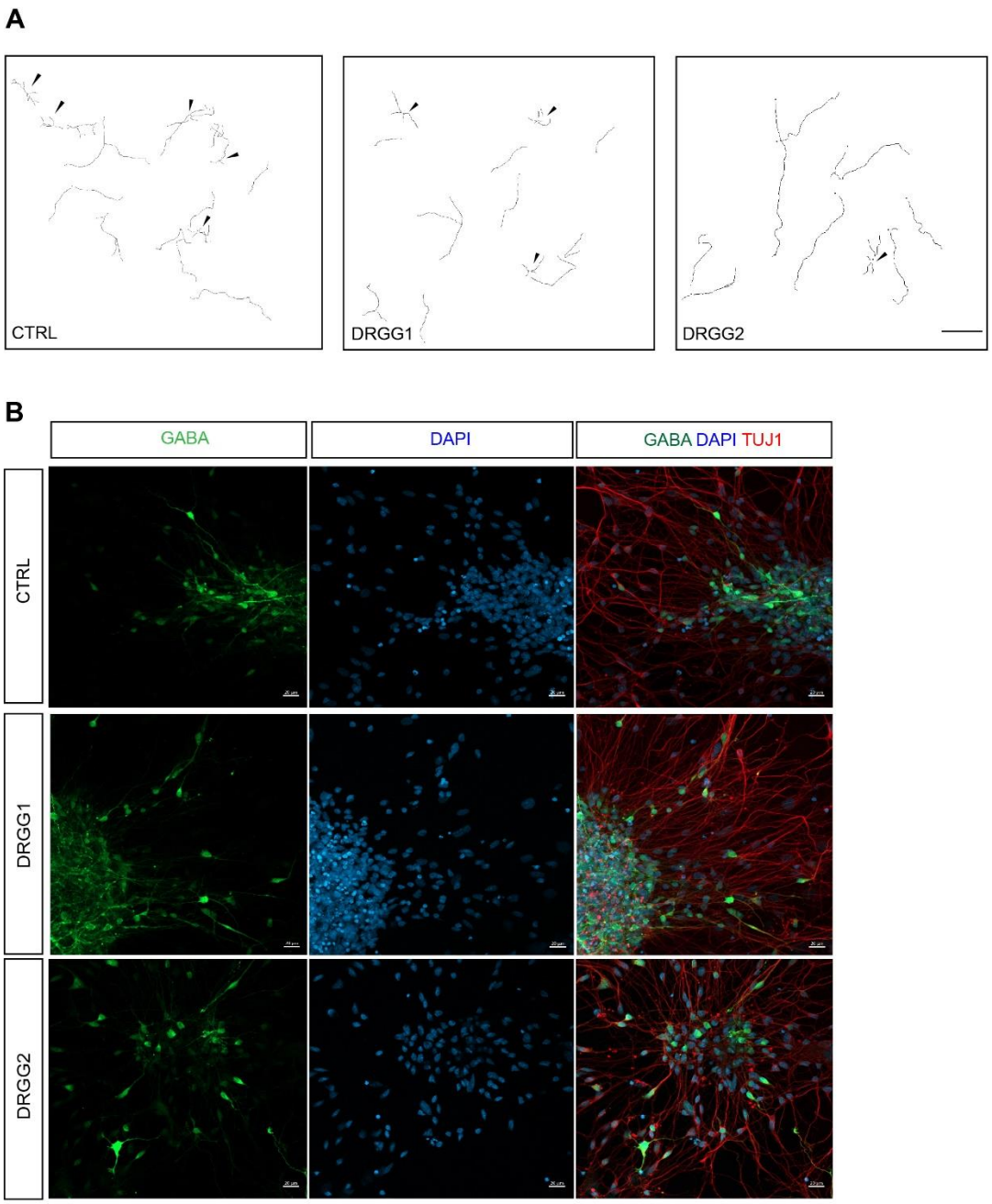


Figure S2. 7 Less branching and shorter neurites in DRGG differentiated neuronal cultures.
(A) Representative traces of CTRL, DRGG1 and DRGG2 iPSC-derived cortical TUJ1 positive neurons at 4 weeks of differentiation. Dots indicate the cell bodies and lines represent the individual neurites. Black arrowheads denote multipolar neurons. Scale bar, 100 μm.
(B) Immunostaining of GABA and TUJ1 positive neurons in CTRL, DRGG1 and DRGG1 cultures. DAPI staining was used to visualize the nuclei. Scale bar, 20 μm.

Table S2. 1 Methylated peptide of RBMX

Peptide	Start site	Stop site
DSYSSSSme2RGAPmeR	362	373
DSYSSSSmeRGAPmeR	362	373
DSYSSSSmeRGAPR	362	373
YDDYSSSmeRDGYGGSR	310	324
SSSGLGGmeRAPVSme2RGR	173	187
SAPSGPVmeRSSSGLGGR	165	180
meRSAPSGPVR	164	172
GGHMDDGGYSMNFTMSSSmeR	126	144
GGHMDDGGYSMNFTMSSSme2R	126	144
GPPSme2RGGHMDDGGYSMNFTMSSSR	121	144
GPPSmeRGGHMDDGGYSMNFTMSSSR	121	144
me3RGLPPPPR	94	101
meRGLPPPPR	94	101
meRGLPPPPmeRSR	94	103
GFAFVTFESPADAKDAAme2R	50	67
VEADme3RPGKLFIGGLNTETNEK	2	22

CHAPTER 3 Arginine Methylation of SARS-Cov-2 Nucleocapsid Protein Regulates RNA Binding, Its Ability to Suppress Stress Granule Formation, and Viral Replication

Ting Cai[1], Zhenbao Yu[1], Zhen Wang[2], Chen Liang[2], and Stéphane Richard[1]*

[1]Segal Cancer Center, Lady Davis Institute for Medical Research and Gerald Bronfman Department of Oncology and Departments of Biochemistry, Human Genetics and Medicine, McGill University, Montréal, Québec, Canada H3T 1E2

[2]McGill Centre for Viral Diseases, Lady Davis Institute for Medical Research and Department of Medicine, Department of Microbiology and Immunology, McGill University, Montréal, Québec, Canada H3T 1E2

3.1 Preface

COVID-19 pandemic has caused millions of deaths and had a profound impact on every aspect of our lives. Arginine methylation is an important PTM regulating virus replication. For example, Richard lab showed the HIV transactivator protein, Tat, is a substrate of host PRMT6, and arginine methylation of Tat regulates its transactivation activity and virus replication. SARS-CoV-2 N protein harbors RGG/RG motif that may be regulated by host PRMTs. In this chapter, we showed that RGG/RG motif and arginine methylation play important roles in N protein function.

3.2 Abstract

Viral proteins are known to be methylated by host protein arginine methyltransferases (PRMTs) playing critical roles during viral infections. Herein, we show that PRMT1 methylates SARS-CoV-2 nucleocapsid (N) protein at residues R95 and R177 within RGG/RG sequences. Arginine methylation of N protein was confirmed by immunoblotting viral proteins extracted from SARS-CoV-2 virions isolated by cell culture. We demonstrate that arginine methylation of N protein is required for its RNA binding capacity, since treatment with a type I PRMT inhibitor (MS023) or substitution of R95K or R177K inhibited interaction with the 5'-UTR of the SARS-CoV-2 genomic RNA. We defined the N interactome in HEK293 cells with or without MS023 treatment and identified PRMT1 and many of its RGG/RG substrates including the known interactor, G3BP1, and other components of stress granules (SG). Methylation of N protein at R95 regulates another function namely its property to suppress the formation of SGs. MS023 treatment or R95K substitution blocked N-mediated suppression of SGs. Also, the co-expression of methylarginine reader TDRD3 quenched N-mediated suppression of SGs in a dose-dependent manner. Finally, pre-treatment of VeroE6 cells with MS023 significantly reduced SARS-CoV-2 replication. With

type I PRMT inhibitors being in clinical trials for cancer treatment, inhibiting arginine methylation to target the later stages of the viral life cycle such as viral genome packaging and assembly of virions may be an additional therapeutic application of these drugs.

Keywords: SARS-CoV-2, type I PRMT inhibitor, nucleocapsid (N) protein, arginine methylation, PRMT1, RGG/RG motif, RNA binding, condensate, stress granules

3.3 Introduction

The COVID-19 pandemic is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), a virus that belongs to the family Coronaviridae of genus Betacoronavirus and has a positive-sense strand RNA genome of ~30 kb (1). It contains two large overlapping ORFs (ORF1a and ORF1b) and encodes four structural proteins, namely spike (S), envelope, membrane, and nucleocapsid (N) proteins as well as nine accessory proteins (1). ORF1a and ORF1b are further processed to generate 16 nonstructural proteins (Nsp1–16). Among the viral proteins, N protein is the most abundant in the virions and is expressed at the highest levels in infected cells (2). Thus, its abundance, essential roles during infection, and immunogenic nature make the SARS-CoV-2 N protein a valuable target for developing new strategies to combat the COVID-19 pandemic (3, 4, 5).

N protein regulates different steps of the coronavirus life cycle (2). The primary role of betacoronavirus N protein is the packaging of the viral genome into helical ribonucleoprotein complexes (6). It is also involved in RNA synthesis with components of the replicase at early stages of infection (7, 8). Betacoronavirus N protein has two conserved and independently folded structural domains, called the N-terminal RNA-binding domain (NTD) and C-terminal dimerization domain (CTD) (4, 9, 10), separated by flexible intrinsically disordered regions at the N terminus, central serine/arginine-rich (SR) linker region, and C-terminal tail, respectively. The

crystal structure of the SARS-CoV-2 NTD RNA-binding domain depicts a U-shaped β-platform containing five short β-strands and an extended hairpin, forming a palm and finger-like structure with a highly positively charged cleft (4, 11).

After viral infection, host cells generate stress granules (SGs) as an antiviral response to inhibit protein synthesis and induce innate immune signaling (12, 13). SARS-CoV N protein plays an important role in host–virus interaction and localizes to cytoplasmic SGs (14). The SG nucleating factor G3BP1 (15, 16) and other SG components were identified in the SARS-CoV-2 N protein interactome (5, 17), suggesting SARS-CoV-2 like SARS-CoV regulates SGs mainly through N protein. The SARS-CoV-2 N protein is able to form condensates with RNA in vitro (18, 19, 20, 21, 22, 23, 24, 25), in the cytoplasm of cells (20, 24) and partially colocalizes within arsenite-induced SGs (26). Several studies have shown that N protein sequesters G3BP1 and disassembles SGs (21, 26, 27, 28), likely as a means to suppress the host immune response to favor virus replication. Recent studies show that intrinsically disordered region 1 and NTD regulate N protein condensates affecting nucleic acid annealing and potentially implicated in viral packaging and assembly (20, 21, 24). The SR linker region is phosphorylated by SRSF protein kinase (SRPK) (29), glycogen synthase kinase 3 (GSK-3) (30), and cyclin-dependent kinase 1–GSK3 (18), influencing N protein condensates (18, 21). Besides phosphorylation, post-translational modifications that regulate N function are not known.

We identify that SARS-CoV-2 N protein contains five undefined and uncharacterized RGG/RG motifs. RGG/RG motifs are prevalent in RNA binding proteins (RBPs) and play key roles in mediating protein–protein and protein–RNA interactions (31, 32). The arginine residues located within the RGG/RG motifs are the preferred sites of methylation by protein arginine methyltransferases (PRMTs) (33). In mammals, there are nine PRMTs (PRMT1–9) that are

classified into three types based on the methyl marks they produce: NG-monomethylarginine, asymmetric NG, NG-dimethylarginine, and symmetric NG, N'G dimethylarginine (33). Methylarginines are bound by Tudor domains that are methylarginine readers (34). Arginine methylation regulates protein–protein interactions and protein–nucleic acid interactions to influence basic cellular processes, including transcription, RNA processing including pre-mRNA splicing, mRNA export, and mRNA translation, signaling transduction, and liquid–liquid phase separation (35, 36). Unlike lysine demethylation, dedicated arginine demethylases have not been identified (36). Many specific small-molecule inhibitors of PRMTs have been generated for cancer therapeutics (37, 38, 39, 40, 41, 42), and a few have entered clinical trials (for review, see (36)).

Arginine methylation is known to methylate host and viral proteins necessary for viral replication. For example, the arginine methylation of HIV Tat protein decreases its transactivation function (43). The inhibition of PRMT5 prevents host heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1) RGG/RG motif methylation and inhibits HIV-1 and human T-cell lymphotropic virus type-1 internal ribosome entry sites function (44). Moreover, PRMT5 methylates hepatitis B virus core protein within its C-terminal arginine-rich domain to regulate its cellular localization (45). Arginine methylation of prototype foamy virus Gag in its glycine–arginine–rich box by PRMT1 regulates its nucleolar localization during replication (46).

In the present study, we report that PRMT1 methylates SARS-CoV-2 N protein within its RGG/RG motifs to regulate the RNA-binding activity of the N protein toward its 5’-UTR genomic RNA. Moreover, arginine methylation modulates the role of N protein to inhibit SG formation. Our findings show for the first time that inhibition of type I PRMTs decreased SARS-CoV N methylation within virions and that arginine methylation is required for viral production.

3.4 Results

3.4.1 SARS-CoV-2 N protein is methylated by PRMT1

We noted that the SARS-CoV-2 N protein harbors two RGG (Fig. 3.1A) and three RG motifs like SARS-CoV, but unlike Middle East respiratory syndrome coronavirus (MERS-CoV). As RGG/RG motifs are preferred sites of PRMT1, PRMT5, and PRMT6 (31), we tested whether they could be methylated. We first expressed and purified glutathione S-transferase (GST)-fusion proteins of the SARS-CoV-2 N protein fragments (GST-N 1–150, GST-N 150–262, and GST-N 263–419, Fig. 3.1A) and performed in vitro arginine methylation assays. Both the N-terminal fragment (amino acid residues 1–150) and the central region (amino acid residues 150–262) were arginine-methylated by PRMT1 (Fig. 3.1B). By contrast, the N protein fragments were not methylated by PRMT5 or PRMT6 (Fig. 3.1, C and D). We then substituted arginines in the RGG/RG motifs to lysines to identify the methylated sites and maintain the charge. Mutation of arginine 68 (R68K) in the N-terminal fragment had no significant effect on arginine methylation, whereas mutation of arginine 95 (R95K) completely abolished PRMT1 methylation (Fig. 3.1E), suggesting that R95 was the methylated residue in the GST-N 1 to 150 fragment. Similarly, mutation analysis identified R177 as the methylated residue in the central fragment (Fig. 3.31F). Taken together, R95 and R177 within the RGG/RG motifs of the SARS-CoV-2 N were methylated in vitro by PRMT1.

We then determined whether the SARS-CoV-2 N protein was methylated in cells. HEK293 cells were transfected with a plasmid expressing Flag-epitope N protein (Flag-N). The cells were lysed, and the N protein was immunoprecipitated with anti-Flag antibodies and its methylation detected by Western blotting using the ASYM26 antibody, which specifically recognizes asymmetrically dimethylated arginine residues within RGG/RG motifs. Importantly, the

asymmetrical dimethylarginine methylation of the Flag-N (N-me2) was significantly reduced by treatment of the cells with the type I PRMT inhibitor MS023 (Fig. 3.1G) and transfection with siPRMT1 (Fig. 3.1H).

We next monitored patient data to identify modulation of PRMT1 expression during SARS-CoV-2 infection. Single-cell RNA sequencing analysis of nasopharyngeal and bronchial samples from 19 clinically well-characterized SARS-CoV-2 patients and five healthy controls was performed (47). Importantly, analysis of their data showed that PRMT1 was significantly upregulated in infected patients (Fig. S3.1). These data suggest PRMT1 may play a role during the SARS-CoV-2 life cycle.

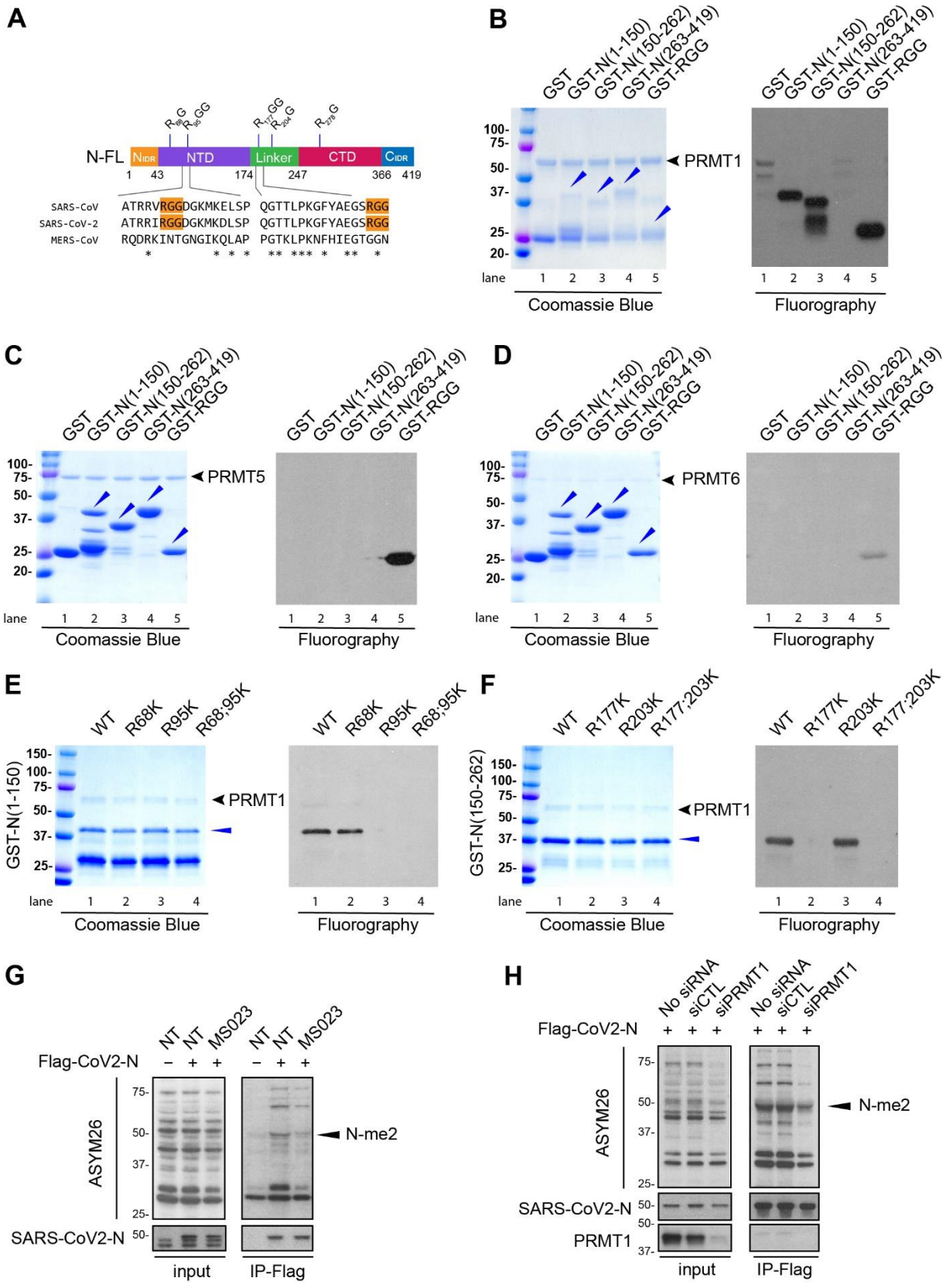

A
N-FL
NIDR
NTD
Linker
CTD
CIDR
1
43
174
247
366
419
SARS-CoV ATRRVRGGDGKMKELSP QGTTLPKGFYAEGSRGG
SARS-CoV-2 ATRRIRGGDGKMKDLSP QGTTLPKGFYAEGSRGG
MERS-CoV RQDRKINTGNGIKQLAP PGTKLPKNFHIEGTGGN
B
GST
GST-N(1-150)
GST-N(150-262)
GST-N(263-419)
GST-RGG
PRMT1
lane
Coomassie Blue
Fluorography
C
PRMT5
D
PRMT6
E
WT
R68K
R95K
R68;95K
GST-N(1-150)
F
R177K
R203K
R177;203K
GST-N(150-262)
G
NT
MS023
Flag-CoV2-N
ASYM26
N-me2
SARS-CoV2-N
input
IP-Flag
H
No siRNA
siCTL
siPRMT1
PRMT1

Figure 3. 1 R95 and R177 within SARS-CoV-2 N RGG/RG motifs are methylated by PRMT1.
(A) Schematic diagram of N protein with its N-terminal and C-terminal domains (NTD, CTD) and its N_{IDR} and C_{IDR} for N- and C-terminal intrinsic disordered regions and lastly the linker region between NTD and CTD known for its SR-rich sequences. Note R95GG and R177GG are conserved in SARS-CoV and SARS-CoV-2, but not MERS-CoV.
(B, C, D) Recombinant GST-N protein fragments were subjected to *in vitro* methylation assays with recombinant B) GST-PRMT1, C) PRMT5/MEP50, and D) GST-PRMT6. Coomassie Blue staining and fluorography images are presented. GST alone and GST-RGG were used as negative and positive controls, respectively. Blue arrowheads indicate the migration of the GST-N protein fragments. The migration of GST-PRMT1, PRMT5 and GST-PRMT6 is shown on the right with a black arrow. The molecular mass markers are shown in kDa on the left.
(E, F) GST-N protein fragments with arginine to lysine substitution were subjected to *in vitro* methylation assays. Coomassie Blue staining and fluorography images are presented.
(G) HEK293 cells were transfected with control (-) or Flag-N (+) expression vectors for 24h and incubated with or without (NT) 1 μM MS023 for another 24h. The cell lysates were subjected to immunoprecipitation with anti-Flag-M2 beads and immunoblotting with anti-asymmetrical dimethylarginine antibody ASYM26 (upper panels) and anti-SARS-CoV-2 N protein antibody (lower panels). The band of the asymmetrically dimethylated N protein (N-me2) is marked by black arrowhead on the right. The molecular mass markers are shown in kDa on the left.
(H) HEK293 cells were transfected with siRNA targeting firefly luciferase (siCTL) or siPRMT1 for 24h and subsequently transfected with Flag-N vector for another 24h. The cell lysates were subjected to immunoprecipitation with anti-Flag-M2 beads and then immunoblotting with anti-SARS-CoV-2 N protein and anti-ASYM26 antibodies. The migration of the methylated N protein is indicated.

3.4.2 The SARS-CoV-2 N interactome defines a complex of RGG/RG proteins and PRMT1

We then performed MS analysis to identify SARS-CoV-2 N-interacting proteins in the absence or presence of MS023. Flag-tagged SARS-CoV-2 N protein was expressed in HEK293 cells and a pull-down performed using anti-Flag affinity resin. Co-purified cellular proteins were subsequently analyzed by affinity-purification (AP) LC-MS/MS. We identified 119 cellular proteins interacting with SARS-CoV-2 N protein (peptide count >2, fold change >2 between Flag-N and empty vector transfected, 0.1% false discovery rate, Fig. 3.2, A and B). Importantly, we identified several protein components of SGs such as G3BP1 and G3BP2 (Ras-GTPase–activating protein SH3 domain-binding protein 1 and 2) (48), and CAPRIN1 (Fig. 3.2A), in line with previous published AP-MS/MS studies (5, 17). Moreover, our MS analysis identified SRPK1 and GSK-3, known to phosphorylate N protein and regulate its localization to SGs (18, 21, 29).

We also identified TRIM25 as a top hit in the N protein interactome. As a K63-linked ubiquitin ligase, TRIM25 mediates retinoic acid–inducible gene 1 ubiquitination and activates TLR/RLR signaling pathway in response to RNA virus infection. It is known that SARS-CoV N protein interacts with TRIM25 and inhibits TRIM25/retinoic acid–inducible gene 1 association (49), suggesting that SARS-CoV-2 N protein may play a similar role in antagonizing the host immune response.

We then performed biological process (Gene Ontology) analysis using the identified interaction partners to assess major cellular pathways. The top ten pathways enriched consisted of RNA metabolic processes (Fig. S3.2). Interestingly, many SARS-CoV-2 N-interacting proteins (27 of 119, Fig. 3.2B) contained multiple RGG/RG motifs including DEAD/DExH family of RNA helicases DDX21, DDX54, DHX30, DHX57, and hnRNPA1, A3, D, DL, G (RBMX), R, U, UL1, and UL2 (Table 3.1). Many of these N-interacting proteins, such as G3BP1 (50), FAM98A (51), FXR1 (52), hnRNPA1 (53), hnRNPUL1 (54), SYNCRIP (55), ILF3 (56), and SERBP1 (57) are known PRMT1 substrates.

N protein interactome changed significantly by two proteins with MS023 treatment. Interaction between N protein and HMGB2 was lost, and interaction with PRMT1 was gained with MS023 treatment (Fig. 3.2C). HMGB2 is a paralog of HMBG1 shown to play a critical role in SARS-CoV-2 replication (58). Interestingly, the AP-MS/MS analysis identified 11 PRMT1 peptides covering 30% of the protein sequence in the Flag-N immunoprecipitation from MS023-treated cells and none in nontreated cells (Fig. 3.2C). These data are consistent with MS023 being noncompetitive with type I PRMT substrates (37). To confirm these interactions, HEK293 cells were transfected with Flag-N and treated with dimethylsulfoxide (DMSO) control or MS023. Cell extracts were prepared and immunoprecipitations performed with anti-Flag antibodies followed

by SDS-PAGE and immunoblotted with anti-PRMT1, anti-G3BP1, or anti-SARS-CoV2 N protein antibodies. Indeed, Flag-N immunoprecipitations showed increased PRMT1 association with MS023 treatment and G3BP1 was observed in immunoprecipitations from treated and nontreated cells (Fig. 3.2D, compare lanes 5 and 6). These data confirm interactions between the N protein and G3BP1 and PRMT1.

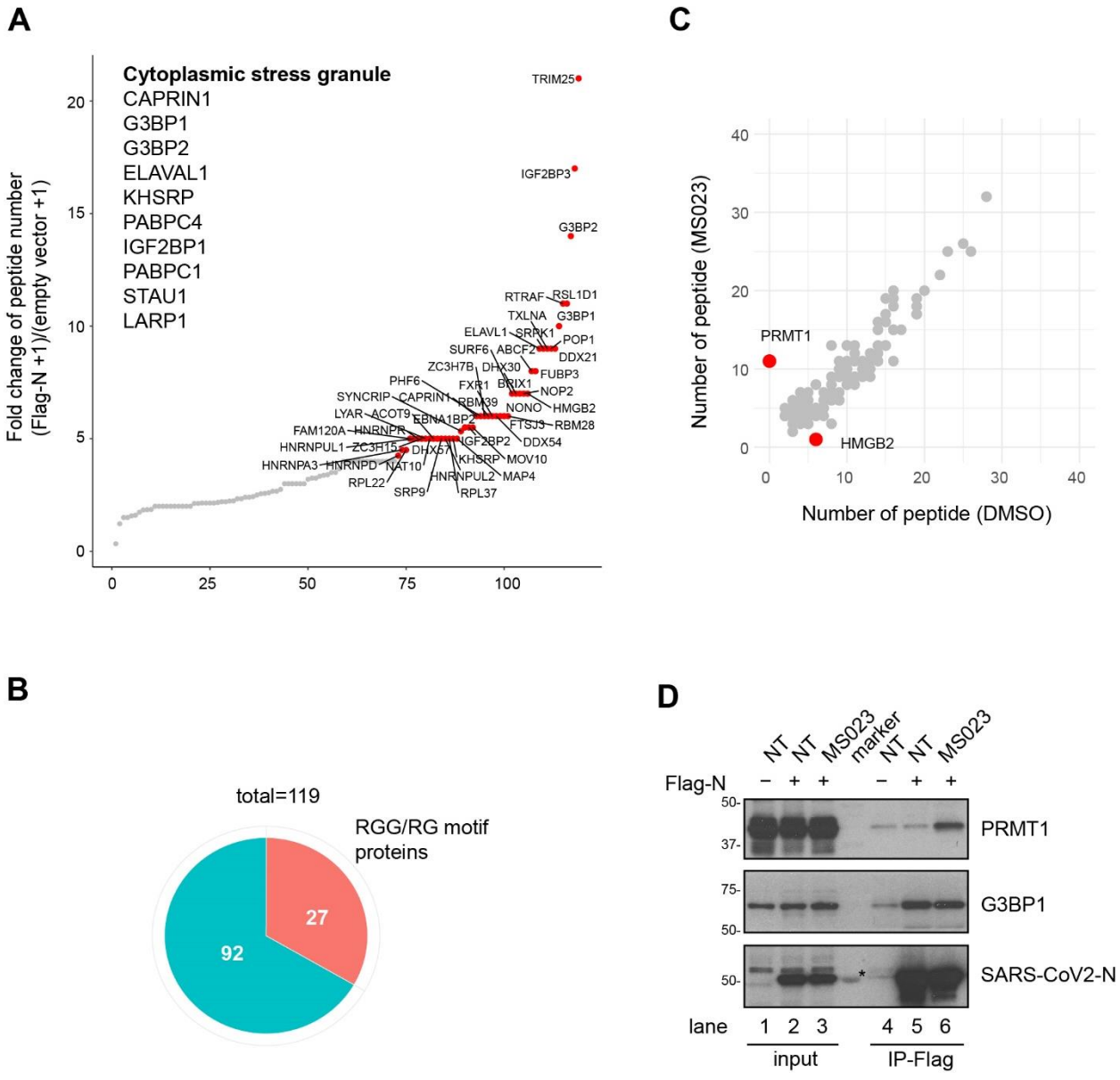


Figure 3. 2 N protein interactome with and without MS023: association with many RGG/RG proteins and PRMT1.

HEK293 cells were transfected with control or Flag-N and the next day Flag-N-transfected cells were subsequently treated with or without 1 μM MS023 for 24 h. Cell lysates were subjected to immunoprecipitation using anti-Flag-M2 beads. The bound proteins were identified by mass spectrometry (A-C).
(A) Interactors were ranked by fold change of unique peptides detected from Flag-N-transfected cells and control plasmid-transfected cells (Flag-N + 1)/(empty vector + 1). Proteins with FC > 4 are highlighted in red. Immunoprecipitated proteins known to be localized in stress granule are listed.
(B) Correlation analysis between MS023 and DMSO treated N protein interactome is shown. Proteins with a significant fold change (> 3 or < 3^{-1}) after MS023 treatment are highlighted in red.
(C) Pie chart represents the number of RGG/RG motif containing proteins among N protein interactors.
(D) HEK293 cells were transfected with control (-) or Flag-N (+) and subsequentially treated with or without (NT) 1 μM MS023 for 24h. Cell lysates were immunoprecipitated with anti-Flag antibodies and the associated proteins separated by SDS-PAGE and immunoblotted with anti-PRMT1, anti-G3BP1 and anti-SARS-CoV-2 N antibodies. The asterisk denotes non-specific recognition of a molecular mass marker protein.

3.4.3 SARS-CoV-2 N prevents SG formation in an arginine methylation–dependent manner

SGs are frequently observed upon infection with DNA or RNA viruses, serving an antiviral function (12, 13). Recent studies reveal that SARS-CoV-2 N protein is associated with SGs and regulates their dynamics (21, 26, 27, 28). To investigate whether arginine methylation regulates the property of SARS-CoV-2 N protein to suppress SG formation and dynamics, we monitored SG formation using anti-G3BP1 antibodies in the hepatoma Huh-7, a cell line frequently used in the study of SARS-CoV-2. Huh-7 cells transfected with an empty plasmid (pcDNA) or a plasmid expressing Flag-N were treated with a mild dose of oxidative stress (0.5 mM sodium arsenite for 1 h) or a harsh dose (1 mM sodium arsenite for 2 h). At 1-μM sodium arsenite, we observed Flag-N colocalizing with G3BP1 in SGs (open arrowheads), and in some Flag-N–expressing cells, there was a reduction or absence of G3BP1 SGs (white arrowheads) (Fig. 3.3A), as reported recently (21, 26, 27, 28). Interestingly, at the mild doses of 0.5 mM sodium arsenite for 1 h, 25.89 ± 2.56% of Flag-N-transfected cells had G3BP1 SGs compared with 70.08 ± 1.93% in the pcDNA-transfected cells (Fig. 3.3B), suggesting that SARS-CoV-2 N protein suppresses G3BP1 SG formation. As regulating SGs formation is an important function for viral replication and host

cell immune response (13), we focused our study on how arginine methylation was implicated in N protein–mediated SG suppression. Thus, all subsequent studies were performed with 0.5 mM sodium arsenite for 1 h to study N protein inhibition of G3BP1 SGs. Huh-7 cells transfected with Flag-N were treated with type I PRMT inhibitor MS023 or control DMSO before induction of SGs with sodium arsenite. Methyltransferase inhibition using MS023 significantly increased the presence of Flag-N–expressing cells with G3BP1 SGs (45.48 ± 4.79% versus 30.99 ± 3.92%), while no significant change was observed in the nontransfected cells with over >70% of the cells with SGs (Fig. 3.3C). PRMT1 inhibition in HeLa cells is known to increase the number of SGs per cell via RGG/RG motif methylation of G3BP1 (50) and UBAP2L (59). To demonstrate the role of arginine methylation suppression of G3BP1 SGs was due to N protein methylation, per se, we transfected Huh-7 cells with WT and R-K Flag-N and monitored SGs. Cells with N protein with R95K or R95K/R177K substitution showed increased SG formation in comparison with those transfected with Flag-N or Flag-N R177K (Fig. 3.4A, R95K 35.82 ± 3.03%, R177K 26.74 ± 2.52%, R95K/R177K 36.62 ± 2.78% versus WT N 25.30 ± 2.62%). These findings show that the methylation of N protein at R95, but not R177, is required for the SARS-CoV-2 N to suppress G3BP1-positive SGs.

We next examined whether the ectopic expression of a methylarginine reader could also interfere with N protein–mediated regulation of G3BP1 SGs. The Tudor domain is a known reader of methylated arginine residues (34). We first examined whether Flag-N associated with ectopically expressed Tudor domain–containing proteins survival of motor neuron (SMN) and Tudor domain-containing protein 3 (TDRD3) by coimmunoprecipitation assays. Indeed, a strong interaction between Flag-N and TDRD3 was observed, whereas the interaction with SMN was weaker, as visualized by immunoblotting (Fig. 3.4B). Next, we tested whether TDRD3 influenced

N protein–mediated SG regulation, as it is known that TDRD3 localizes to SGs (60, 61). We cotransfected increasing amounts of expression plasmid encoding TDRD3 together with Flag-N and visualized the presence of arsenite-induced G3BP1 SGs. We observed an increase in G3BP1 SGs with increased expression of TDRD3 with N protein (Fig. 3.4C). These findings suggest that increasing the methyl reader TDRD3 expression could be a means to quench the effects of N protein on SG regulation.

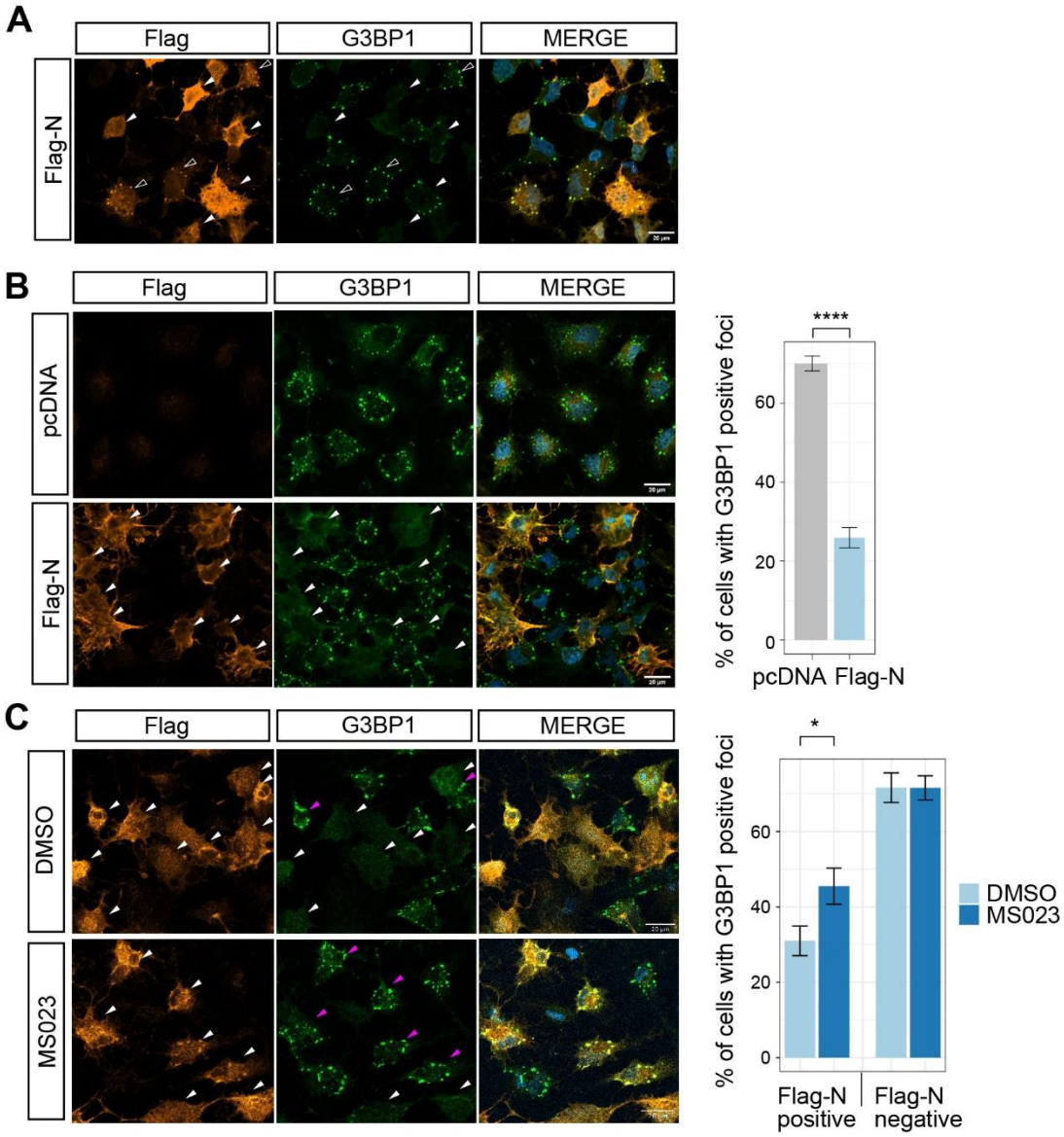


Figure 3. 3 SARS-CoV-2 N protein regulates G3BP1 stress granule formation in an argininemethylation-dependent manner.

(A) Huh-7 cells were transfected with control vector or Flag-N for 24h and subsequently incubated with 1 mM sodium arsenite for 2h. Cells were fixed with 4% PFA and co-immunostained with anti-Flag and anti-G3BP1 antibodies. A typical image is shown. Scale bar represents 20 μm. Arrowheads indicate FLAG-N-transfected cells, and the empty arrowheads indicate cells with FLAG-N and G3BP1 colocalization.
(B) Huh-7 cells were transfected with control vector or Flag-N for 24h and subsequently incubated with 0.5 mM sodium arsenite for 1h. Cells were fixed with 4% PFA and co-immunostained with anti-Flag and anti-G3BP1 antibodies. A typical image is shown. Scale bar represents 20 μm. Arrowheads indicate FLAG-N-transfected cells, and transfected cells with SGs (> 5 G3BP1 foci) were highlighted in magenta. The percentage of cells harboring SGs (> 5 G3BP1 foci) were quantified and shown in the bar plot on the right. n =15 fields from 3 independent experiments are shown. Welch's t test. ****$p < 0.0001$.
(C) Huh-7 cells were transfected with Flag-N overnight and treated with or without 5 μM MS023 for another 24h. Then the cells were incubated with 0.5 mM sodium arsenite for another hour and fixed with 4% PFA. Cells were co-immunostained with anti-Flag and anti-G3BP1 antibodies. A typical image is shown in the left panel. Scale bar, 20 μm. Arrowheads indicate Flag-N-transfected cells, and transfected cells with SGs (> 5 G3BP1 foci) were highlighted in magenta. Percentage of cells harboring SGs (> 5 G3BP1 foci) were quantified in the transfected cells (Flag-N positive) and non-transfected cells (Flag-N negative) respectively and shown in the bar plot on the right. n =20 fields from 2 independent experiments are shown. Welch's t test. *$p < 0.05$.

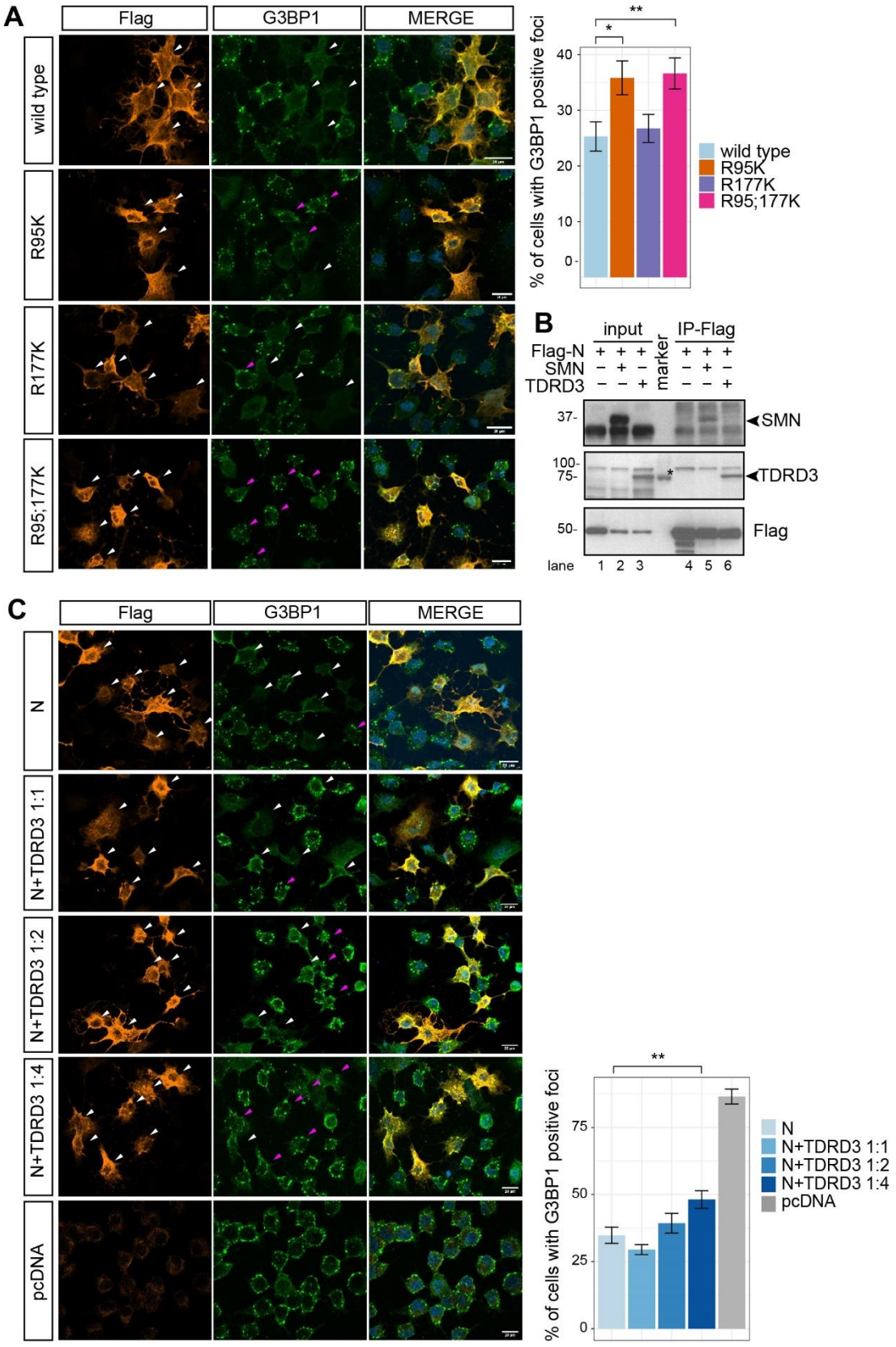

A
Flag
G3BP1
MERGE
wild type
R95K
R177K
R95;177K
**
*
% of cells with G3BP1 positive foci
40
30
20
10
0
wild type
R95K
R177K
R95;177K
B
input
IP-Flag
Flag-N + + + + + +
SMN − + − − + −
TDRD3 − − + − − +
marker
37-
SMN
100-
75-
TDRD3
50-
Flag
lane 1 2 3 4 5 6
C
Flag
G3BP1
MERGE
N
N+TDRD3 1:1
N+TDRD3 1:2
N+TDRD3 1:4
pcDNA
**
% of cells with G3BP1 positive foci
75
50
25
0
N
N+TDRD3 1:1
N+TDRD3 1:2
N+TDRD3 1:4
pcDNA

Figure 3. 4 R95 is required for N protein regulating G3BP1 stress granule and methylarginine reader protein TDRD3 is involved in this process.
(A) Huh-7 cells were transfected with Flag-N (wild type) and its mutants for 24h and subsequently incubated with 0.5 mM sodium arsenite for another hour. Cells were fixed with 4% PFA and co-immunostained with anti-Flag and anti-G3BP1 antibodies. A typical image is shown in the left panel. Scale bar, 20 μm. Arrowheads indicate transfected cells, and transfected cells with SGs (> 5 G3BP1 foci) were highlighted in magenta. Percentage of cells harboring SGs (> 5 G3BP1 foci) in the Flag-N-positive cell group were quantified and shown in the bar plot on the right. n =15 fields from 3 independent experiments are shown. Welch's t test. $*p < 0.05$, $**p < 0.01$.
(B) HEK293 cells were co-transfected with Flag-N and myc-SMN or myc-TDRD3. 24h later, cell lysates were immunoprecipitated with anti-Flag antibodies and the associated proteins separated by SDS-PAGE and immunoblotted with anti-SMN, anti-TDRD3 and anti-Flag antibodies. The asterisk denotes non-specific recognition of a molecular mass marker protein.
(C) Huh-7 cells were co-transfected with Flag-N and myc-TDRD3 with indicated plasmid ratios. 24h later, cells were incubated with 0.5mM sodium arsenite for another hour. Cells were fixed with 4% PFA and co-immunostained with anti-Flag and G3BP1 antibodies. A typical image is shown on the left. Scale bar, 20 μm. Arrowheads indicate transfected cells, and transfected cells with SGs (> 5 G3BP1 foci) were highlighted in magenta. Percentage of the transfected cells (Flag-N positive) harboring SGs (> 5 G3BP1 foci) were quantified and shown in the bar plot on the right. n >15 fields from 2 independent experiments are shown. Welch's t test. $**p < 0.01$.

3.4.4 Arginine methylation of R95 and R177 is required for N protein binding to the SARS-CoV-2 5'-UTR RNA

N is an RNA-binding protein that binds the 5'-UTR of its viral gRNA for viral ribonucleoprotein (vRNP) formation and packaging into virions (62, 63). The SARS-CoV-2 N protein R95 and R177 are located in the NTD and at NTD–SR linker boundary, respectively. Actually, R95 and R177 are within the RNA-binding site of the NTD with R177 being predicted to be implicated in N protein RNA binding (4, 11, 64). Therefore, we reasoned that these arginines and their methylation were likely involved in the RNA-binding activity of N protein. HEK293 cells were cotransfected with Flag-N and an expression vector transcribing ~400 bp of the 5'-UTR RNA sequence of SARS-CoV-2 (p5'-UTR:CoV-2; Fig. 3.5A). Initially, we performed RNA immunoprecipitation (RIP) to monitor N protein RNA-binding activity. We observed a >5-fold enrichment with anti-Flag antibodies in the DMSO-treated cells versus MS023-treated cells (Fig. 3.5A). These data suggest that inhibition of type I PRMTs prevents the binding of Flag-N to the 5'-UTR of SARS-CoV-2 RNA. Next, we wished to confirm that this N protein/RNA interaction

was direct by performing a photoactivatable ribonucleoside–enhanced crosslinking and immunoprecipitation (PAR-CLIP) assay (65). Cells transfected with Flag-N or Flag-N harboring R95K, R177K, R95K/R177K, and p5'-UTR:CoV-2 were labeled with 4-thiouridine and UV cross-linked. The cells were lysed and immunoprecipitated with anti-Flag antibodies after a 'clipping' step with RNase A. RNA was purified and analyzed by RT-qPCR with two sets of primers against the 5'-UTR of the SARS-CoV-2 RNA (positions #1 and #2). Using this strategy, we showed that Flag-N directly binds to the 5'-UTR of SARS-CoV-2 RNA (Fig. 3.5B). Importantly, both the single R95K and R177K or the double R95K/R177K substitution of N protein abolished RNA-binding activity (Fig. 3.5B). Immunoblotting was performed to confirm an equal expression of WT N protein and the R-K proteins immunoprecipitated of the four replicates (Fig. 3.5C). Taken together, these results suggest that arginine methylation of both R95 and R177 of the N protein is a requirement for association with its viral RNA *in cellulo*.

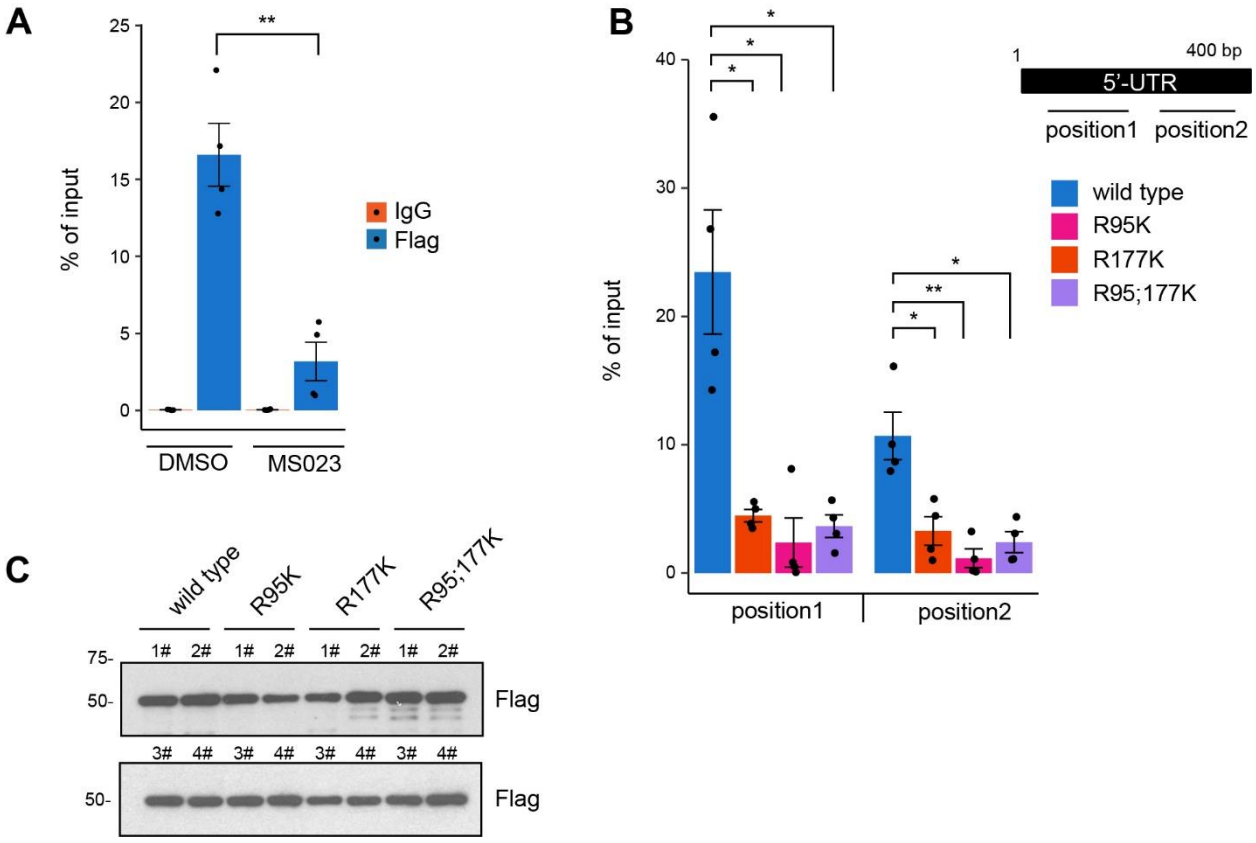


Figure 3. 5 Arginine methylation of N R95 and R177 is a requirement for SARS-CoV-2 N binding to the 5'-UTR of its genomic RNA.
(A) HEK293 cells were co-transfected with the plasmid expressing Flag-N protein and the plasmid expressing a 400 bp RNA fragment of the SARS-CoV-2 5'-UTR region (p5'-UTR:CoV-2, NC_045512: 1-400 bp). The cells were incubated with 5 μM MS023 or DMSO for 24 h. Cells were cross-linked with 1% formaldehyde and subjected to RIP using IgG or anti-Flag antibodies. The immunoprecipitated RNA was extracted and RT-qPCR was used to assess the bound RNA. Data is shown as percentage of input from 2 independent experiments. Welch's t test. $^{**}p < 0.01$.
(B and C) HEK293 cells were co-transfected with plasmids expressing wild type and RK N proteins along with the plasmid expressing p5'-UTR:CoV-2 RNA. Then the cells were treated with 4-thiouridine for 16h and subjected to PAR-CLIP analysis (B). RT-qPCR with primers targeting consecutive regions, shown in the diagram, were used to assess the bound RNA. Data is shown as percentage of input from 2 independent experiments. Welch's t test. $^{*}p < 0.05$, $^{**}p < 0.01$. Immunoblotting for expression of indicated proteins from 2 independent experiments is shown respectively in the top and bottom panels (C).

3.4.5 Methylation of N protein is required for SARS-CoV-2 production

We then studied the effect of type I PRMT inhibitor MS0233 on SARS-CoV-2 replication. First, we performed MS023 toxicity assays with VeroE6 cells, a cell line highly susceptible to SARS-CoV-2 infection. We confirmed that cell proliferation was not affected at concentrations

of MS023 up to 30 μM in complete medium (Fig. 3.6A). Next, in a certified SARS-CoV-2 BL3 laboratory, we then treated VeroE6 cells with 10 μM or 20 μM MS023, versus DMSO control, for 24 h and proceeded with SARS-CoV-2 infection at a low multiplicity of infection (0.1). The infected cells were kept in MS023-containing medium for another 2 days. Supernatant from each, infected well was collected, and the virus inactivated with TRIzol to assay SARS-CoV-2 titer by TaqMan real-time PCR assay. We observed a significant reduction of viral titer when the cells were incubated with 20 μM MS023, and an intermediate viral titer was observed with 10 μM MS023 (Fig. 3.6B).

From the TRIzol organic phase, we extracted the viral proteins and separated them by SDS-PAGE followed by immunoblotting with anti-SARS-CoV-2 S, and N protein or anti-methylarginine ASYM26 antibodies. We observed a slight decrease (∼20–30%) in the total amounts of S and N proteins with MS023 treatment (Fig. 3.6C), consistent with overall decrease in virions (Fig. 3.6B). Notably, the anti-ASYM26 antibodies revealed a ∼50% reduction in N protein arginine methylation, when normalized to N protein levels (Fig. 3.6C, middle panel). In contrast, we could not detect methylation of S with ASYM26 antibody (Fig. 3.6C), likely because of the lack of RGG/RG motifs in S protein sequence. These findings show that the SARS-CoV-2 production is reduced in the presence of type I PRMT inhibitors and that the SARS-CoV-2 N protein is arginine-methylated within the virions.

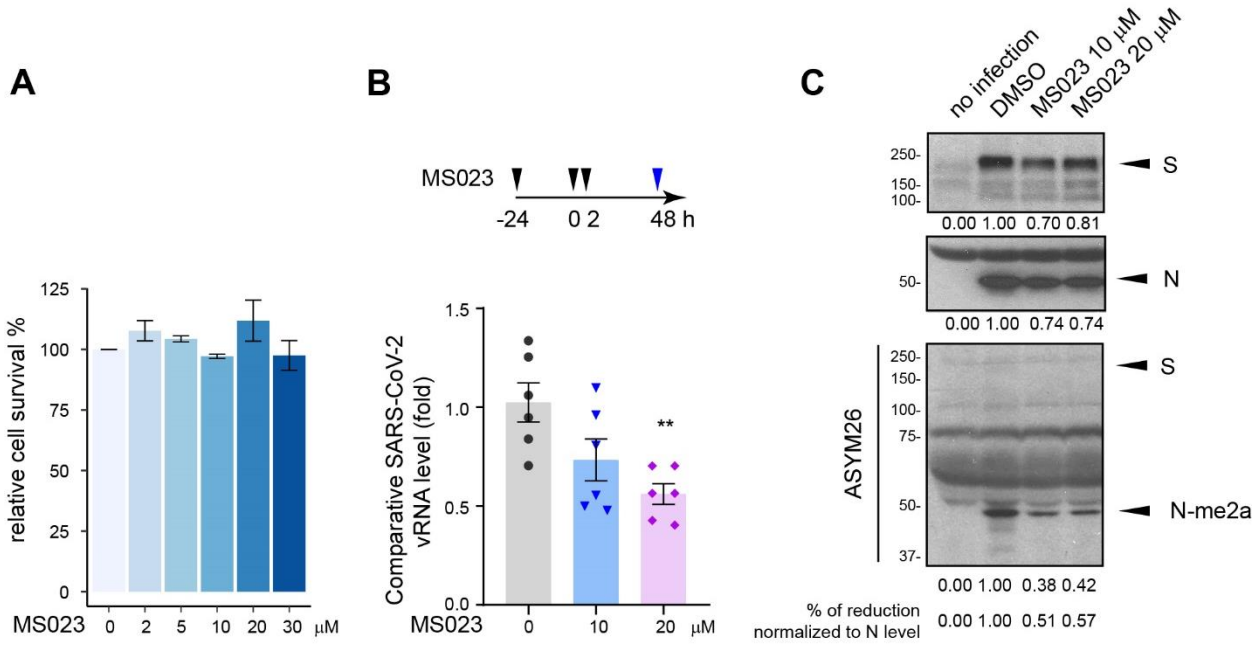


Figure 3. 6 SARS-CoV-2 replication is impaired by type I PRMT inhibitor MS023.
(A) VeroE6 cells were treated with the indicated dose of MS023 for 3 days as described in the "Materials and Methods" section and cell viability was determined by MTT assay. Data represents percentage of survival compared to control (DMSO alone) from 3 independent experiments.
(B) VeroE6 cells were pretreated with MS023 or vehicle DMSO as control for 24 h. Then the cells were infected with SARS-CoV-2 at an MOI of 0.1. Two hours later, the medium was refreshed, and cells were incubated with the same concentration of MS023 or DMSO, respectively. Forty-eight hours later, the cell medium containing the released SARS-CoV-2 virions was collected and lysed in TRIzol immediately. Viral RNA was isolated and Taq-man probe base RT-qPCR was used to assess the viral load. Data represents the fold change against control samples from 2 independent experiments. Welch's t test. $^{**}p < 0.01$.
(C) Viral proteins were extracted from the organic phase of samples in (B) and immunoblotted with anti-S, anti-N, and anti-ASYM26 antibodies. The density of the protein bands was calculated using ImageJ.

3.5 Discussion

In the present article, we identify SARS-CoV-2 N protein to be arginine-methylated within virions. We show that methylation of the N protein is mediated by PRMT1 at R95 within the NTD and R177 at the NTD/SR linker boundary. Both residues are within RGG/RG motifs conserved between SARS-CoV and SARS-CoV-2. Amino acid substitutions R95K or R177K inhibited N protein RNA binding in cellulo to SARS-CoV-2 5'-UTR genomic RNA using a CLIP assay in HEK293 cells. The ectopic expression of N protein in Huh-7 cells was localized to

cytoplasmic granules and inhibition of G3BP1 SG formation was observed. Notably, arginine methylation of N protein at R95 by PRMT1 was necessary for this function. The N protein interactome was defined to contain many known PRMT1 substrates with RGG/RG motifs. Treatment with the type I PRMT inhibitor MS023 did not influence the overall N interactome, but PRMT1 was identified as a new interactor, consistent with a substrate–enzyme interaction in the presence of the noncompetitive inhibitor MS023 (37). Importantly, inhibition of arginine methylation with MS023 significantly reduced SARS-CoV-2 replication in VeroE6 cells. Our findings define arginine methylation as new mode of interfering with N protein regulation of SGs and define PRMT1 as a requirement for the SARS-CoV-2 life cycle. As type I PRMT inhibitors are in clinical trials for cancer treatment (41), these compounds may also be useful to target SARS-CoV-2 replication.

Host proteins and enzymes are needed for the replication of SARS-CoV-2, and many of these were identified using CRISPR screens (58, 66, 67, 68, 69). Factors required for early viral entry and fusion, and components of the endosome and cholesterol pathways were identified (66, 67, 68). Few components, however, were identified that target the later phases of the viral life cycle, such as vRNP formation, phase separation for genome packaging, encapsidation, and assembly of virions. A recent study identified ∼50 host RBPs binding SARS-CoV-2 genomic RNA and the knockdown of some of these RBPs inhibited viral replication (70). As SARS-CoV-2 N protein and cellular RBPs are substrates of PRMT1 (31), our findings suggest that type I PRMT inhibitors or increasing methylarginine readers may regulate N function in vRNP packaging and assembly into virions. Notably, the increased expression of PRMT1 in nasopharyngeal and bronchial samples of SARS-CoV-2-infected patients (71) is consistent with PRMT1 being a host factor required for the virus.

SARS-CoV-2 viral protein extracts immunoblotted with ASYM26 revealed the arginine methylation of the N protein. It is known that the N protein is highly immunogenic and anti-N antibodies are among the first to appear in infected individuals (72). It is likely that SARS-CoV-2 N protein is fully methylated in infected cells, thus epitopes from the methylated RGG/RG peptides of N protein likely contribute to its elevated immunogenicity. Notably, R95- and R177-derived SARS-CoV-2 N peptides are predicted to be B cell epitopes (73, 74). Moreover, R95 and R177 peptides from the N protein were used to design SARS-CoV-2 vaccines in India (73), and we propose that the incorporation of asymmetrical dimethylarginines would increase their immunogenicity. Although our MS significantly covered 74% of the N protein, the R95 was not covered and R177 was covered, but the peptides were unmethylated. Interestingly, the only variant in N protein is R203K;G204R in the B.1.17 strain and T205I in B.1.351 strain, suggesting the RGT sequence is an important regulatory site. The absence of methylation of the S protein observed using anti-ASYM26 is likely due to the lack of RGG/RG motifs. However, the S protein harbors an RRXR sequence within its furin cleavage site, a potential methylation sequence for PRMT7 (75). Thus, MS analysis and various anti-methylarginine antibodies targeting NG-monomethylarginine and symmetric NG, N'G dimethylarginine will be needed to further define the sites of arginine methylation within the SARS-CoV-2 viral proteins.

Coronaviruses, like other viruses, have devised strategies to destabilize and inhibit SG formation to ensure optimal viral replication. The infectious bronchitis coronavirus uses endoribonuclease Nsp15 for SG interference (76). For MERS-CoV, it is the 4a accessory protein that inhibits SG formation (77, 78). The N protein from SARS-CoV and SARS-CoV-2 were shown to localize to SGs (14, 21, 26, 27, 28). Thus, targeting the N protein function in G3BP1 SG regulation represents a new and valid strategy to fight COVID-19. It is known that the SR

linker region when phosphorylated renders the N protein condensate more liquid-like, and inhibition of N protein phosphorylation favors its translocation to SGs (14, 18, 24, 29, 30, 79, 80). We now show that arginine methylation of N protein is a post-translational modification that tunes N protein regulation of G3BP1 SGs. Information about how SGs are formed and regulated is emerging and represents a combination of multivalent interactions of protein–protein, protein–RNA, and RNA–RNA interactions (81). How does arginine methylation of N protein regulate SG formation? Both R95K and R177K were defective in RNA binding to the SARS-CoV-2 5'-UTR RNA, and yet, only R95K within the NTD was required for SG regulation. Our observation is in agreement with a recent report that residues 1 to 175 of N protein are sufficient to disrupt SGs (26). Consistent with R95 and R177 being part of the U-shaped β-platform of the RNA-binding domain (4, 11), we show that both residues are needed to bind the 5'-UTR where the putative viral packaging signal of the genomic RNA resides. It is possible that N protein R177 is not needed to associate nonspecifically with host mRNAs, and hence, its substitution to lysine does not influence SGs. Thus, we propose that RGG/RG motif methylation of the N protein affects SGs by modulating interaction with methyl readers and RNA.

The N interactome was not significantly altered in the presence of MS023, but there was an increase in PRMT1 association. We noted that increasing the concentration of a methylarginine reader TDRD3 blocked N protein from suppressing SG formation. Therefore, we propose that MS023 besides inhibiting the activity of PRMT1, also inactivates N protein by increasing its interaction with PRMT1, therefore allowing PRMT1 to become an enzyme-inactive RGG/RG motif reader. A role for arginine methylation of the RGG/RG motif in RBP phase separation is known (32, 36). For example, the RGG/RG motif protein FUS, dysregulated in ALS, undergoes liquid–liquid phase separation in the absence of PRMT1 (82, 83). Thus, arginine methylation is a

key regulator of ribonucleoprotein condensation. As type I PRMT GSK3368715 inhibitor is in phase I clinical trials for diffuse large B-cell lymphoma (41), our data suggest that these inhibitors may be an effective strategy to interfere with N protein condensation and influence late stages of the SARS-CoV-2 life cycle.

Arginine methylation influences nearby phosphorylation sites often being antagonistic (36). For example, arginine methylation of the FOXO1 transcription factor at R248 and R250 by PRMT1 prevents AKT phosphorylation at S253, blocking nuclear exclusion of FOXO1 (84). Arginine methylation of cytoplasmic tail of the epidermal growth factor receptor at R1175 by PRMT5 enhances its trans-autophosphorylation at Y1173 (85). N protein RNA-binding activity is known to be regulated by phosphorylation. In vitro studies showed that hypophosphorylation of the N protein facilitated interaction with RNA (21). We show that methylation of N protein R95 and R177 is needed for RNA binding. As S176, S180, S183, and S184 are reported to be phosphorylated by SRPK1 (29) and GSK3–cyclin-dependent kinase 1 (18, 30, 79), it is likely that there will be an interplay between phosphorylation and methylation especially near R177 for binding to the 5'-UTR of the SARS-CoV-2 RNA. It is likely that optimal binding of N protein to the 5'-UTR of the SARS-CoV-2 RNA will require a balance of methylarginines and phosphoserines. This interplay may also influence interactions with methyl readers including TDRD3 and phosphoreaders such as 14-3-3 proteins, the latter shown to bind the N protein (86). In sum, our findings suggest that arginine methylation is required for N protein function, and PRMT1 is an essential regulator implicated in SARS-CoV2 life cycle. As PRMT inhibitors are in clinical trials (36), they may have applications for COVID-19, in addition to them being promising cancer drug candidates.

3.6 Experimental procedures

Reagents and antibodies

Immunoblotting was performed using the following antibodies: mouse anti-Flag monoclonal antibody (F1804, Sigma Aldrich, 1:2000); rabbit anti-SARS-CoV-2 N antibody (9103, Prosci, 1:2000); rabbit anti-SARS-CoV-2 S antibody (PA5-81795, Invitrogen, 1:1000); rabbit anti-TDRD3 antibody (Bethyl Laboratory, 1:1000); mouse anti-SMN antibody (610646, BD Biosciences, 1:2000); rabbit anti-G3BP1 antibody (1F1, Rhône-Poulenc Rorer, 1:1000, a kind gift from Dr. Imed Gallouzi at McGill University) (86); rabbit anti-ASYM26 (13-0011, Epicypher, 1:1000). Immunofluorescence were performed with the following antibodies: mouse anti-Flag monoclonal antibody (F1804, Sigma Aldrich, 1:500); rabbit anti-G3BP1 antibody (1F1, 1:500). Alexa Fluor-conjugated goat anti-rabbit, goat anti-mouse secondary antibodies were from Invitrogen. Protease inhibitor cocktail and protein phosphatase inhibitor cocktail were from Roche. MS023 (SML1555), sodium arsenite (S7400), Protein A-Sepharose (P3391) and PRMT5:MEP50 active complex (SRP0145) were purchased from Sigma Aldrich. Protein co-immunoprecipitation was performed using ANTI-FLAG® M2 Affinity Gel (A2220, Sigma Aldrich).

Cell culture and transfection

HEK293 and VeroE6 cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and grown at 37°C with 5% CO2. Huh-7 cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and Non-Essential Amino Acid (NEAA, Gibco) and grown at 37°C with 5% CO2. Cells were transfected with 20 nM siRNA oligonucleotides using Lipofectamine RNAiMAX (Invitrogen) according to the manufacturer's instructions. HEK293 and Huh-7 cells

were transfected with plasmid DNAs by standard calcium phosphate precipitation and Lipofectamine 3000 (Invitrogen), respectively.

Plasmids and siRNAs

The N-terminal Flag-tagged SARS-CoV-2 N plasmid was constructed by inserting a Flag-coding sequence into the pcDNA3.1 (+) vector at the Hind III and Bam HI sites to generate pcDNA3.1-Flag and then the PCR-amplified cDNA of SARS-CoV-2 N coding region at Bam HI and Xho I sites of pcDNA3.1-Flag vector. The PCR template DNA is a plasmid with insertion of synthesized DNA expressing SARS-CoV-2 N protein provided by Dr. Shan Cen based on the SARS-CoV-2 Wuhan-Hu-1 isolate (GenBank: MN_908947). The plasmids for expressing GST fusion proteins of SARS-CoV-2 N truncated fragments were constructed by inserting PCR-amplified SARS-CoV-2 N cDNA fragments in pGEX-6P1 vector at Bam HI and Sal I sites. The GST-RGG construct was generated by inserting a DNA fragment expressing the mouse RBMX C-terminal RGG/RG motif in pGEX-6P1 vector at Bam HI and Sal I sites. The mutants with replacement of the arginine residues with lysine at the RGG/RG motif were constructed by a two-step PCR strategy. p5’-UTR:CoV-2 was constructed using synthesized DNA fragment with the 1-400 bp of 5’-UTR of SARS-CoV-2 gRNA (NC_045512) and the DNA fragment was insert into pcDNA3.1 vector at Bam HI and Xba I sites. The myc-tagged SMN and TDRD3 plasmids were generated in previous studies (87,88). All siRNAs were purchased from Dharmacon. siRNA sequences are as follows: siPRMT1, 5'-CGT CAA AGC CAA CAA GTT AUU-3'. The siRNA 5′ -CGU ACG CGG AAU ACU UCG AdTdT-3′, targeting the firefly luciferase (GL2) was used as control. 20 nM siRNA was used for transfection.

Protein purification and in vitro methylation assay

Expression of GST fusion proteins in bacteria was induced with 1mM isopropyl-□-D-thiogalactopyranoside (IPTG) at room temperature for 16 h. All steps of the purification after growth of bacteria were performed at 4°C. Cells were lysed by sonication in PBS buffer containing a mixture of protease inhibitors. The lysate was clarified by centrifugation and the supernatant was incubated with glutathione-Sepharose 4B beads for 2 h. The resin was washed four times with PBS buffer and then twice with 50 mM Tris-HCl pH7.4 buffer. Protein was eluted with 10 mM reduced glutathione in 50 mM Tris-HCl pH 7.4 buffer. Approximately 10 μg of each GST fusion protein was incubated with 1 μl of (methyl-3H) S-adenosyl-L-methionine solution (15Ci/mmol stock solution, 0.55 μM final concentration, Perkin-Elmer) and 1-2 μg of PRMTs in methylation buffer (50mM Tris-HCl pH 7.4, 1 mM DTT) for 1 to 4 h at 25°C or 37°C. Samples were separated by SDS-PAGE and stained with Coomassie Blue. After de-staining, the gel was then incubated for 1 h in EN3HANCE (Perkin Elmer) according to manufacturer's instructions and the reaction was visualized by fluorography.

Cell lysis, immunoprecipitation, immunoblotting and LC-MS/MS analysis

For co-immunoprecipitation experiments, cells were lysed in lysis buffer (50 mM HEPES, pH 7.4, 150 mM NaCl, 1% Triton X-100 and a cocktail of protease inhibitors and phosphatase inhibitors). Cell lysates were cleared with high-speed centrifugation to remove cell debris, then the supernatant was incubated with anti-Flag M2 beads for 1.5 h at 4ºC. Samples were washed with 1 ml of lysis buffer for 3 times and eluted with 2× SDS loading buffer for western blot analysis. For LC-MS/MS, the beads were further washed with PBS buffer twice. The beads together with the bound proteins were subjected to LC-MS/MS performed at the mass spectrometry facility at IRIC, Université de Montréal. The data were analyzed using Scaffold Proteome Software.

MS023 toxicity assay

Prior to performing viral infections, the MS023 inhibitor was examined for toxicity to the VeroE6 cells. VeroE6 cells (2,500 cells per well) were seeded in 96-well plate and cultured at a condition similar to that in the viral infection analysis. The type I PRMT inhibitor MS023 was dissolved in DMSO and diluted in complete DMEM containing 10% FBS. The medium was added to cells with a final concentration of 2 to 30 μM MS023 and 0.2% DMSO as indicated. Twenty-four hours later, the cell culture medium was replaced with 2% FBS/DMEM medium containing the same concentration of MS023 and DMSO in the corresponding wells. After 48 h of further incubation, the cell viability was analyzed using the MTT assay kit (Abcam, ab211091) according to manufacturer's instructions. Briefly, the media was carefully removed, and both 50 μl of serum-free medium and 50 μl of MTT reagent were added to each well. Following 3 h of incubation at 37°C, the MTT reagent was removed and 150 μl of MTT solvent was added to each well and the plate was incubated at room temperature on an orbital shaker for 30 min prior to reading absorbance at OD590 nm. The absorbance of MS023-treated wells was divided by the absorbance of the DMSO-treated wells to normalize cell survival.

SARS-CoV-2 infection and purification of genomic SARS CoV-2 RNA (gRNA) and proteins

Within the certified BL3 containment facility of the McGill University Health Centre, VeroE6 cells were seeded in 24-well plates (10^5/well in 0.5 ml) and incubated in complete DMEM containing 10% FBS in the presence of PRMT1 inhibitor MS023 or DMSO control for 24 h prior to infection. The cells were then infected with SARS-CoV-2 isolate RIM-1 (GenBank accession number: MW599736) at multiplicity of infection (MOI=0.1) at 37°C for 2h. The virus inoculum was removed, and the cells were washed once with 2% FBS/DMEM medium and then incubated for an additional 48 h in 1 ml of 2% FBS/DMEM containing the PRMT inhibitor at indicated

concentrations or same amount of DMSO as control at 37°C. After the infection was complete, 250 μl cell supernatant was lysed in 0.75 ml TRIzol LS (Invitrogen) and transported out of the BL3 facility. The viral RNA was then extracted from the TRIzol using chloroform extraction following manufacturer's instructions. One step RT-qPCR was performed using TaqMan™ Fast Virus 1-Step Master Mix following the manufacturer's instructions. Viral gRNA was detected using primers (Fw: 5'-ATG AGC TTA GTC CTG TTG-3', Rv: 5'-CTC CCT TTG TTG TGT TGT-3'), and probe (5'Hex-AGA TGT CTT GTG CTG CCG GTA-BHQ-1-3'), specifically targeting RdRp gene as described (89). In addition, viral proteins were extracted from the organic phase of TRIzol solution according to the manufacturer's instruction. Briefly, after removing the aqueous phase, 0.3 ml 100% ethanol was added to the organic phase. Genomic RNA (gRNA) from infected cells was pelleted by centrifugation at 2000g for 5 min. 0.75 ml supernatant was moved to a new tube and mixed with 1.5 ml isopropanol. Proteins were collected by centrifugation at 12000 rpm for 15 min, following by two times of washing with 0.3 M guanidine hydrochloride and 95% ethanal. Liquid was removed and the pellet was air-dried. The dried proteins were dissolved in 1XSDS loading buffer and proceed with western blot analysis.

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

PAR-CLIP (photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation) was performed as previously described with minor modification (90). p5'-UTR:CoV-2 and pFlag-N were co-transfected to HEK293 cells with a 1:1 ratio. 24 h post-transfection, the cells were treated with 100 μM 4-thiouridine for 16 h and cross-linked with 0.15 J/cm2 365 nm UV. Cells were then washed twice with ice cold PBS and resuspended in lysis buffer (150mM KCl, 25mM Tris, pH 7.4, 5mM EDTA, 0.5mM DTT, 0.5% NP40, and 100U/mL RNase inhibitor). After

incubation for 20 min with rotation at 400 rpm and cell debris were cleared by centrifugation. Cell lysates were incubated with 1 U/ml RNase I at 37 °C for 3 min. For each immunoprecipitation, 40 U RNase inhibitor and 2 μg of antibody was added, and the samples were incubated for 2 h at 4°C with rotation. Protein A Sepharose beads (Sigma) were then added and the samples were incubated at 4°C for another hour with rotation. The beads were pelleted by centrifugation, resuspended, and washed in high salt wash buffer for 3 times and lysis buffer once. After removing the final wash buffer, DNA fragments were digested with 2U TURBO™ DNase (Thermo fisher, AM2238) at the 37°C for 4 min. RNA was eluted in Proteinase K buffer (50 mM Tris, pH 7.5, 75 mM NaCl, 6.5 mM EDTA, and 1% SDS) supplemented with proteinase K and incubated at 50°C for 30 min. RNA was recovered by using 5 volumes of TRIzol™ LS Reagent (Thermo Fisher). Equal volume of RNA from each sample was used for the reverse transcription. qPCR was performed using primers targeting gRNA 5'-UTR. Primer sequences used in the experiment are as follows: position 1: Fw: 5'-TCG TTG ACA GGA CAC GAG TA-3', Rv: 5'-CCC GGA CGA AAC CTA GAT GT-3'; position 2: Fw: 5'- CCT TGT CCC TGG TTT CAA CG-3', Rv: 5'- CAC GTC GCG AAC CTG TAA AA-3'.

RNA immunoprecipitation (RIP) was performed as previously described with minor modifications (91). Briefly, cells were cross-linked with a final concentration of 1% formaldehyde, washed twice with ice cold PBS and resuspended in RIP buffer (150mM KCl, 25mM Tris, pH 7.4, 5mM EDTA, 0.5mM DTT, 0.5% NP40, and 100U/mL RNase inhibitor). Chromatin was sheared by sonication, and DNA fragments were digested with TURBO™ DNase (Thermo fisher, AM2238) at the 37°C for 15 min. Cell lysate were proceeded with immunoprecipitation with antibody and wash with RIP buffer. The RNA was recovered from the

precipitate as described above. qPCR was performed using primers: Fw: 5'-TCG TCT ATC TTC TGC AGG CT-3', Rv: 5'-ACG TCG CGA ACC TGT AAA AC-3'.

Arsenite treatment and immunofluorescence

Cells growing on glass coverslips were treated with 0.5 mM arsenite for 1h and fixed for 10 min with 4% paraformaldehyde (PFA). After three washes with PBS the cells were permeabilized for 5 min with 0.25% Triton X-100 in PBS. Coverslips were incubated with blocking buffer containing 5% FBS for 1h, and then incubated with primary antibodies diluted in PBS containing 5% FBS for 2h. After three washes, the coverslips were incubated with corresponding fluorescent secondary antibodies for another hour in PBS buffer containing 5% FBS. After rinsing, the coverslips were mounted with IMMU-MOUNT (Thermo Scientific) mounting medium containing 1µg/ml of 4′,6-diamidino-2-phenylindole (DAPI). Images were taken using a Zeiss LSM800 confocal microscope.

Statistical analysis

All data are expressed as mean ± S.E.M. and compared between groups using the Welch's t test. p value <0.05 was considered statistically significant. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

3.7 References

1. Kim, D., Lee, J. Y., Yang, J. S., Kim, J. W., Kim, V. N., and Chang, H. (2020) The Architecture of SARS-CoV-2 Transcriptome. Cell 181, 914-921
2. McBride, R., van Zyl, M., and Fielding, B. C. (2014) The coronavirus nucleocapsid is a multifunctional protein. Viruses 6, 2991-3018
3. Ray, M., Sarkar, S., and Rath, S. N. (2020) Druggability for COVID-19: in silico discovery of potential drug compounds against nucleocapsid (N) protein of SARS-CoV-2. Genomics Inform 18, e43
4. Kang, S., Yang, M., Hong, Z., Zhang, L., Huang, Z., Chen, X., He, S., Zhou, Z., Zhou, Z., Chen, Q., Yan, Y., Zhang, C., Shan, H., and Chen, S. (2020) Crystal structure of SARS-CoV-2 nucleocapsid protein RNA binding domain reveals potential unique drug targeting sites. Acta Pharm Sin B 10, 1228-1238
5. Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., White, K. M., O'Meara, M. J., Rezelj, V. V., Guo, J. Z., Swaney, D. L., Tummino, T. A., Huttenhain, R., Kaake, R. M., Richards, A. L., Tutuncuoglu, B., Foussard, H., Batra, J., Haas, K., Modak, M., Kim, M.,

Haas, P., Polacco, B. J., Braberg, H., Fabius, J. M., Eckhardt, M., Soucheray, M., Bennett, M. J., Cakir, M., McGregor, M. J., Li, Q., Meyer, B., Roesch, F., Vallet, T., Mac Kain, A., Miorin, L., Moreno, E., Naing, Z. Z. C., Zhou, Y., Peng, S., Shi, Y., Zhang, Z., Shen, W., Kirby, I. T., Melnyk, J. E., Chorba, J. S., Lou, K., Dai, S. A., Barrio-Hernandez, I., Memon, D., Hernandez-Armenta, C., Lyu, J., Mathy, C. J. P., Perica, T., Pilla, K. B., Ganesan, S. J., Saltzberg, D. J., Rakesh, R., Liu, X., Rosenthal, S. B., Calviello, L., Venkataramanan, S., Liboy-Lugo, J., Lin, Y., Huang, X. P., Liu, Y., Wankowicz, S. A., Bohn, M., Safari, M., Ugur, F. S., Koh, C., Savar, N. S., Tran, Q. D., Shengjuler, D., Fletcher, S. J., O'Neal, M. C., Cai, Y., Chang, J. C. J., Broadhurst, D. J., Klippsten, S., Sharp, P. P., Wenzell, N. A., Kuzuoglu-Ozturk, D., Wang, H. Y., Trenker, R., Young, J. M., Cavero, D. A., Hiatt, J., Roth, T. L., Rathore, U., Subramanian, A., Noack, J., Hubert, M., Stroud, R. M., Frankel, A. D., Rosenberg, O. S., Verba, K. A., Agard, D. A., Ott, M., Emerman, M., Jura, N., von Zastrow, M., Verdin, E., Ashworth, A., Schwartz, O., d'Enfert, C., Mukherjee, S., Jacobson, M., Malik, H. S., Fujimori, D. G., Ideker, T., Craik, C. S., Floor, S. N., Fraser, J. S., Gross, J. D., Sali, A., Roth, B. L., Ruggero, D., Taunton, J., Kortemme, T., Beltrao, P., Vignuzzi, M., Garcia-Sastre, A., Shokat, K. M., Shoichet, B. K., and Krogan, N. J. (2020) A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature 583, 459-468

6. Yao, H., Song, Y., Chen, Y., Wu, N., Xu, J., Sun, C., Zhang, J., Weng, T., Zhang, Z., Wu, Z., Cheng, L., Shi, D., Lu, X., Lei, J., Crispin, M., Shi, Y., Li, L., and Li, S. (2020) Molecular Architecture of the SARS-CoV-2 Virus. Cell 183, 730-738 e713
7. Denison, M. R., Spaan, W. J., van der Meer, Y., Gibson, C. A., Sims, A. C., Prentice, E., and Lu, X. T. (1999) The putative helicase of the coronavirus mouse hepatitis virus is processed from the replicase gene polyprotein and localizes in complexes that are active in viral RNA synthesis. J Virol 73, 6862-6871
8. Stertz, S., Reichelt, M., Spiegel, M., Kuri, T., Martinez-Sobrido, L., Garcia-Sastre, A., Weber, F., and Kochs, G. (2007) The intracellular sites of early replication and budding of SARS-coronavirus. Virology 361, 304-315
9. Ye, Q., West, A. M. V., Silletti, S., and Corbett, K. D. (2020) Architecture and self-assembly of the SARS-CoV-2 nucleocapsid protein. Protein Sci 29, 1890-1901
10. Lin, S. M., Lin, S. C., Hsu, J. N., Chang, C. K., Chien, C. M., Wang, Y. S., Wu, H. Y., Jeng, U. S., Kehn-Hall, K., and Hou, M. H. (2020) Structure-Based Stabilization of Non-native Protein-Protein Interactions of Coronavirus Nucleocapsid Proteins in Antiviral Drug Design. J Med Chem 63, 3131-3141
11. Dinesh, D. C., Chalupska, D., Silhan, J., Koutna, E., Nencka, R., Veverka, V., and Boura, E. (2020) Structural basis of RNA recognition by the SARS-CoV-2 nucleocapsid phosphoprotein. PLoS Pathog 16, e1009100
12. White, J. P., and Lloyd, R. E. (2012) Regulation of stress granules in virus systems. Trends Microbiol 20, 175-183
13. McCormick, C., and Khaperskyy, D. A. (2017) Translation inhibition and stress granules in the antiviral immune response. Nat Rev Immunol 17, 647-660
14. Peng, T. Y., Lee, K. R., and Tarn, W. Y. (2008) Phosphorylation of the arginine/serine dipeptide-rich motif of the severe acute respiratory syndrome coronavirus nucleocapsid protein modulates its multimerization, translation inhibitory activity and cellular localization. FEBS J. 275, 4152-4163

15. Yang, P., Mathieu, C., Kolaitis, R. M., Zhang, P., Messing, J., Yurtsever, U., and al, e. (2020) G3BP1 Is a Tunable Switch that Triggers Phase Separation to Assemble Stress Granules. Cell 181, 325-345
16. Guillén-Boixet, J., Kopach, A., Holehouse, A. S., Wittmann, S., Jahnel, M., and al, e. (2020) RNA-Induced Conformational Switching and Clustering of G3BP Drive Stress Granule Assembly by Condensation. Cell 181, 346-361
17. Li, J., Guo, M., Tian, X., Wang, X., Yang, X., Wu, P., Liu, C., Xiao, Z., Qu, Y., Yin, Y., Wang, C., Zhang, Y., Zhu, Z., Liu, Z., Peng, C., Zhu, T., and Liang, Q. (2020) Virus-Host Interactome and Proteomic Survey Reveal Potential Virulence Factors Influencing SARS-CoV-2 Pathogenesis. Med (N Y) 2, 99-112
18. Carlson, C. R., Asfaha, J. B., Ghent, C. M., Howard, C. J., Hartooni, N., Safari, M., Frankel, A. D., and Morgan, D. O. (2020) Phosphoregulation of Phase Separation by the SARS-CoV-2 N Protein Suggests a Biophysical Basis for its Dual Functions. Mol Cell 80, 1092-1103 e1094
19. Chen, H., Cui, Y., Han, X., Hu, W., Sun, M., Zhang, Y., Wang, P. H., Song, G., Chen, W., and Lou, J. (2020) Liquid-liquid phase separation by SARS-CoV-2 nucleocapsid protein and RNA. Cell Res Sep 8, 1-3
20. Cubuk, J., Alston, J.J., Incicco, J.J., Singh, S., Stuchell-Brereton, M.D., Ward,, and M.D., Z., M.I., Vithani, N., Griffith, D., Wagoner, J.A., et al. (2020) The SARS-CoV-2 nucleocapsid protein is dynamic, disordered, and phase separates with RNA. bioRxiv doi: 10.1101/2020.06.17.158121
21. Iserman, C., Roden, C. A., Boerneke, M. A., Sealfon, R. S. G., McLaughlin, G. A., Jungreis, I., Fritch, E. J., Hou, Y. J., Ekena, J., Weidmann, C. A., Theesfeld, C. L., Kellis, M., Troyanskaya, O. G., Baric, R. S., Sheahan, T. P., Weeks, K. M., and Gladfelter, A. S. (2020) Genomic RNA Elements Drive Phase Separation of the SARS-CoV-2 Nucleocapsid. Mol Cell 80, 1078-1091 e1076
22. Savastano, A., Ibanez de Opakua, A., Rankovic, M., and Zweckstetter, M. (2020) Nucleocapsid protein of SARS-CoV-2 phase separates into RNA-rich polymerase-containing condensates. Nat Commun 11, 6041
23. Jack, A., Ferro, L. S., Trnka, M. J., Wehri, E., Nadgir, A., Costa, K., Schaletzky, J., and Yildiz, A. (2020) SARS CoV-2 nucleocapsid protein forms condensates with viral genomic RNA. bioRxiv doi: 10.1101/2020.09.14.295824
24. Perdikari, T. M., Murthy, A. C., Ryan, V. H., Watters, S., Naik, M. T., and Fawzi, N. L. (2020) SARS-CoV-2 nucleocapsid protein phase-separates with RNA and with human hnRNPs. EMBO J 39, e106478
25. Lu, S., Ye, Q., Singh, D., Cao, Y., Diedrich, J. K., Yates, J. R. r., Villa, E., Cleveland, D. W., and Corbett, K. D. (2021) The SARS-CoV-2 nucleocapsid phosphoprotein forms mutually exclusive condensates with RNA and the membrane-associated M protein. Nat Commun 12, 502
26. Luo, L., Li, Z., Zhao, T., Ju, X., Ma, P., Jin, B., Zhou, Y., He, S., Huang, J., Xu, X., Zou, Y., Li, P., Liang, A., Liu, J., Chi, T., Huang, X., Ding, Q., Jin, Z., Huang, C., and Zhang, Y. (2021) SARS-CoV-2 nucleocapsid protein phase separates with G3BPs to disassemble stress granules and facilitate viral production. Sci Bull (Beijing) Jan 19
27. Nabeel-Shah, S., Lee, H., Ahmed, N., Marcon, E., Farhangmehr, S., Pu, S., Burke, G. L., Ashraf, K., Wei, H., Zhong, G., Tang, H., Yang, J., Blencowe, B., Zhang, Z., and Greenblatt, J. F. (2020) SARS-CoV-2 Nucleocapsid protein attenuates stress granule formation and alters

gene expression via direct interaction with host mRNAs. bioRxiv doi.org/10.1101/2020.10.23.342113
28. Wang, J., Shi, C., Xu, Q., and Yin, H. (2021) SARS-CoV-2 nucleocapsid protein undergoes liquid-liquid phase separation into stress granules through its N-terminal intrinsically disordered region. Cell Discov 7, 5
29. Heaton, B. E., Trimarco, J. D., Hamele, C. E., Harding, A. T., Tata, A., Zhu, X., Tata, P. R., Smith, C. M., and Heaton, N. S. (2020) SRSF protein kinases 1 and 2 are essential host factors for human coronaviruses including SARS-CoV-2. bioRxiv doi: 10.1101/2020.08.14.251207
30. Wu, C. H., Yeh, S. H., Tsay, Y. G., Shieh, Y. H., Kao, C. L., Chen, Y. S., Wang, S. H., Kuo, T. J., Chen, D. S., and Chen, P. J. (2009) Glycogen synthase kinase-3 regulates the phosphorylation of severe acute respiratory syndrome coronavirus nucleocapsid protein and viral replication. J Biol Chem 284, 5229–5239
31. Thandapani, P., O'Connor, T. R., Bailey, T. L., and Richard, S. (2013) Defining the RGG/RG motif. Mol Cell 50, 613-623
32. Chong, P. A., Vernon, R. M., and Forman-Kay, J. D. (2018) RGG/RG Motif Regions in RNA Binding and Phase Separation. J Mol Biol 430, 4650-4665
33. Bedford, M. T., and Clarke, S. G. (2009) Protein Arginine Methylation in Mammals: Who, What, and Why. Mol Cell 33, 1-13
34. Gayatri, S., and Bedford, M. T. (2014) Readers of histone methylarginine marks. Biochim Biophys Acta 1839, 702-710
35. Jarrold, J., and Davies, C. C. (2019) PRMTs and Arginine Methylation: Cancer's Best-Kept Secret? Trends Mol Med 25, 993-1009
36. Guccione, E., and Richard, S. (2019) The regulation, functions and clinical relevance of arginine methylation. Nat Rev Mol Cell Biol 20, 642-657
37. Eram, M. S., Shen, Y., Szewczyk, M. M., Wu, H., Senisterra, G., Li, F., Butler, K. V., Kaniskan, H. Ü., Speed, B. A., dela Seña, C., Dong, A., Zeng, H., Schapira, M., Brown, P. J., Arrowsmith, C. H., Barsyte-Lovejoy, D., Liu, J., Vedadi, M., and Jin, J. (2016) A Potent, Selective, and Cell-Active Inhibitor of Human Type I Protein Arginine Methyltransferases. ACS Chem Biol 11, 772-781
38. Chan-Penebre, E., Kuplast, K. G., Majer, C. R., Boriack-Sjodin, P. A., Wigle, T. J., Johnston, L. D., Rioux, N., Munchhof, M. J., Jin, L., Jacques, S. L., West, K. A., Lingaraj, T., Stickland, K., Ribich, S. A., Raimondi, A., Scott, M. P., Waters, N. J., Pollock, R. M., Smith, J. J., Barbash, O., Pappalardi, M., Ho, T. F., Nurse, K., Oza, K. P., Gallagher, K. T., Kruger, R., Moyer, M. P., Copeland, R. A., Chesworth, R., and Duncan, K. W. (2015) A selective inhibitor of PRMT5 with in vivo and in vitro potency in MCL models. Nat Chem Biol 11, 432-437
39. Nakayama, K., Szewczyk, M. M., Dela Sena, C., Wu, H., Dong, A., and al., e. (2018) TP-064, a potent and selective small molecule inhibitor of PRMT4 for multiple myeloma. Oncotarget 9, 18480-18493
40. Szewczyk, M. M., Ishikawa, Y., Organ, S., Sakai, N., Li, F., Ackloo, S., Eram, M. S., dilworth, D., Fukushi, H., Harding, R., dela Sena, C. C., Sugo, T., Hayashi, K., McLeod, D., Zepeda, C., Aman, A., Sanchez-Osuna, M., Bonneil, E., Takagi, S., Rina, A., Tyers, M., Richard, S., Takizawa, M., Gingras, A. C., Arrowsmith, C. H., Vedadi, M., Brown, P. J., Nara, H., and Barsyte-Lovejoy, D. (2020) Pharmacological inhibition of PRMT7 links arginine monomethylation to the cellular stress response. Nat Commun 11, 2396

41. Fedoriw, A., Rajapurkar, S. R., O'Brien, S., Gerhart, S. V., Mitchell, L. H., Adams, N. D., and al., e. (2019) Anti-tumor Activity of the Type I PRMT Inhibitor, GSK3368715, Synergizes with PRMT5 Inhibition through MTAP Loss. Cancer Cell 36, 100-114
42. Shen, Y., Gao, G., Yu, X., Kim, H., Wang, L., Xie, L., Schwarz, M., Chen, X., Guccione, E., Liu, J., Bedford, M. T., and Jin, J. (2020) Discovery of First-in-Class Protein Arginine Methyltransferase 5 (PRMT5) Degraders. J Med Chem 63, 9977-9989
43. Boulanger, M. C., Liang, C., Russell, R. S., Lin, R., Bedford, M. T., Wainberg, M. A., and Richard, S. (2005) Methylation of Tat by PRMT6 regulates human immunodeficiency virus type 1 gene expression. J. Virol. 79, 124-131
44. Barrera, A., Ramos, H., Vera-Otarola, J., Fernández-García, L., Angulo, J., Olguín, V., Pino, K., Mouland, A. J., and López-Lastra, M. (2020) Post-translational modifications of hnRNP A1 differentially modulate retroviral IRES-mediated translation initiation. Nucleic Acids Res 48, 10479-10499
45. Lubyova, B., Hodek, J., Zabransky, A., Prouzova, H., Hubalek, M., Hirsch, I., and Weber, J. (2017) PRMT5: A novel regulator of Hepatitis B virus replication and an arginine methylase of HBV core. PLoS One 12, e0186982
46. Paris, J., Tobaly-Tapiero, J., Giron, M. L., Burlaud-Gaillard, J., Buseyne, F., Roingeard, P., Lesage, P., Zamborlini, A., and Saïb, A. (2018) The invariant arginine within the chromatin-binding motif regulates both nucleolar localization and chromatin binding of Foamy virus Gag. Retroviralogy 15, 48
47. Ramirez Hernandez, E., Hernandez-Zimbron, L. F., Martinez Zuniga, N., Leal-Garcia, J. J., Ignacio Hernandez, V., Ucharima-Corona, L. E., Perez Campos, E., and Zenteno, E. (2021) The Role of the SARS-CoV-2 S-Protein Glycosylation in the Interaction of SARS-CoV-2/ACE2 and Immunological Responses. Viral Immunol doi.org/10.1089/vim.2020.0174
48. Tourrière, H., Chebli, K., Zekri, L., Courselaud, B., Blanchard, J. M., Bertrand, E., and Tazi, J. (2003) The RasGAP-associated endoribonuclease G3BP assembles stress granules. J Cell Biol 160, 823-831
49. Hu, Y., Li, W., Gao, T., Cui, Y., Jin, Y., Li, P., Ma, Q., Liu, X., and Cao, C. (2017) The Severe Acute Respiratory Syndrome Coronavirus Nucleocapsid Inhibits Type I Interferon Production by Interfering with TRIM25-Mediated RIG-I Ubiquitination. J Virol 91
50. Tsai, W. C., Gayatri, S., Reineke, L. C., Sbardella, G., Bedford, M. T., and Lloyd, R. E. (2016) Arginine Demethylation of G3BP1 Promotes Stress Granule Assembly. J Biol Chem 291, 22671-22685
51. Akter, K. A., Mansour, M. A., Hyodo, T., Ito, S., Hamaguchi, M., and Senga, T. (2016) FAM98A is a novel substrate of PRMT1 required for tumor cell migration, invasion, and colony formation. Tumour Biol 37, 4531-4539
52. Stetler, A., Winograd, C., Sayegh, J., Cheever, A., Patton, E., Zhang, X., Clarke, S., and Ceman, S. (2006) Identification and characterization of the methyl arginines in the fragile X mental retardation protein Fmrp. Hum Mol Genet 15, 87-96
53. Wall, M. L., and Lewis, S. M. (2017) Methylarginines within the RGG-Motif Region of hnRNP A1 Affect Its IRES Trans-Acting Factor Activity and Are Required for hnRNP A1 Stress Granule Localization and Formation. J Mol Biol 429, 295-307
54. Gurunathan, G., Yu, Z., Coulombe, Y., Masson, J. Y., and Richard, S. (2015) Arginine methylation of hnRNPUL1 regulates interaction with NBS1 and recruitment to sites of DNA damage. Sci Rep 5, 10475

55. Woodsmith, J., Casado-Medrano, V., Benlasfer, N., Eccles, R. L., Hutten, S., Heine, C. L., Thormann, V., Abou-Ajram, C., Rocks, O., Dormann, D., and Stelzl, U. (2018) Interaction modulation through arrays of clustered methyl-arginine protein modifications. Life Sci Alliance 1, e201800178
56. Tang, J., Kao, P. N., and Herschman, H. R. (2000) Protein-arginine methyltransferase I, the predominant protein-arginine methyltransferase in cells, interacts with and is regulated by interleukin enhancer-binding factor 3. J Biol Chem 275, 19866-19876
57. Lee, Y.-J., Hsieh, W.-Y., Chen, L.-Y., and Li, C. (2012) Protein arginine methylation of SERBP1 by protein arginine methyltransferase 1 affects cytoplasmic/nuclear distribution. J Cell Biochem 113, 2721-2728
58. Wei, J., Alfajaro, M. M., Hanna, R. E., DeWeirdt, P. C., Strine, M. S., and al, e. (2020) Genome-wide CRISPR screen reveals host genes that regulate SARS-CoV-2 infection. bioRxiv doi: 10.1101/2020.06.16.155101
59. Huang, C., Chen, Y., Dai, H., Zhang, H., Xie, M., Zhang, H., Chen, F., Kang, X., Bai, X., and Chen, Z. (2020) UBAP2L arginine methylation by PRMT1 modulates stress granule assembly. Cell Death Differ 27, 227-241
60. Goulet, I., Boisvenue, S., Mokas, S., Mazroui, R., and Cote, J. (2008) TDRD3, a novel Tudor domain-containing protein, localizes to cytoplasmic stress granules. Hum Mol Genet 17, 3055-3074
61. Linder, B., Plöttner, O., Kroiss, M., Hartmann, E., Laggerbauer, B., Meister, G., Keidel, E., and Fischer, U. (2008) Tdrd3 is a novel stress granule-associated protein interacting with the Fragile-X syndrome protein FMRP. Human Molecular Genetics 17, 3236-3246
62. Stohlman, S. A., Baric, R. S., Nelson, G. N., Soe, L. H., Welter, L. M., and Deans, R. J. (1988) Specific interaction between coronavirus leader RNA and nucleocapsid protein. J Virol 62, 4288-4295
63. Chen, S. C., and Olsthoorn, R. C. (2010) Group-specific structural features of the 5'-proximal sequences of coronavirus genomic RNAs. Virology 401, 29-41
64. Khan, S., Hussain, Z., Safdar, M., Khan, A., and Wei, D. Q. (2021) Targeting the N-terminal domain of the RNA-binding protein of the SARS-CoV-2 with high affinity natural compounds to abrogate the protein-RNA interaction: a amolecular dynamics study. J Biomol Struct Dyn Feb 8, 1-9
65. Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jr., Jungkamp, A. C., Munschauer, M., Ulrich, A., Wardle, G. S., Dewell, S., Zavolan, M., and Tuschl, T. (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell 141, 129-141
66. Wang, R., Simoneau, C. R., Kulsuptrakul, J., Bouhaddou, M., Travisano, K. A., Hayashi, J. M., Carlson-Stevermer, J., Zengel, J. R., Richards, C. M., Fozouni, P., Oki, J., Rodriguez, L., Joehnk, B., Walcott, K., Holden, K., Sil, A., Carette, J. E., Krogan, N. J., Ott, M., and Puschnik, A. S. (2021) Genetic Screens Identify Host Factors for SARS-CoV-2 and Common Cold Coronaviruses. Cell 184, 106-119 e114
67. Wei, J., Alfajaro, M. M., DeWeirdt, P. C., Hanna, R. E., Lu-Culligan, W. J., Cai, W. L., Strine, M. S., Zhang, S. M., Graziano, V. R., Schmitz, C. O., Chen, J. S., Mankowski, M. C., Filler, R. B., Ravindra, N. G., Gasque, V., de Miguel, F. J., Patil, A., Chen, H., Oguntuyo, K. Y., Abriola, L., Surovtseva, Y. V., Orchard, R. C., Lee, B., Lindenbach, B. D., Politi, K., van Dijk, D., Kadoch, C., Simon, M. D., Yan, Q., Doench, J. G., and Wilen, C. B. (2021)

Genome-wide CRISPR Screens Reveal Host Factors Critical for SARS-CoV-2 Infection. Cell 184, 76-91 e13
68. Schneider, W. M., Luna, J. M., Hoffmann, H. H., Sanchez-Rivera, F. J., Leal, A. A., Ashbrook, A. W., Le Pen, J., Ricardo-Lax, I., Michailidis, E., Peace, A., Stenzel, A. F., Lowe, S. W., MacDonald, M. R., Rice, C. M., and Poirier, J. T. (2021) Genome-Scale Identification of SARS-CoV-2 and Pan-coronavirus Host Factor Networks. Cell 184, 120-132 e114
69. Zhu, Y., Feng, F., Hu, G., Wang, Y., Yu, Y., Zhu, Y., and al, e. (2021) A genome-wide CRISPR screen identifies host factors that regulate SARS-CoV-2 entry. Nat Commun 12, 961
70. Sun, L., Li, P., Ju, X., Rao, J., Huang, W., Ren, L., Zhang, S., Xiong, T., Xu, K., Zhou, X., Gong, M., Miska, E., Ding, Q., Wang, J., and Zhang, Q. C. (2021) In vivo structural characterization of the SARS-CoV-2 RNA genome identifies host proteins vulnerable to repurposed drugs. Cell 184, 1-19
71. Burbelo, P. D., Riedo, F. X., Morishima, C., Rawlings, S., Smith, D., Das, S., Strich, J. R., Chertow, D. S., Davey, R. T., and Cohen, J. I. (2020) Sensitivity in Detection of Antibodies to Nucleocapsid and Spike Proteins of Severe Acute Respiratory Syndrome Coronavirus 2 in Patients With Coronavirus Disease 2019. J Infect Dis 222, 206-213
72. Noorimotlagh, Z., Karami, C., Mirzaee, S. A., Kaffashian, M., Mami, S., and Azizi, M. (2020) Immune and bioinformatics identification of T cell and B cell epitopes in the protein structure of SARS-CoV-2: A systematic review. Int Immunopharmacol 86, 106738
73. Sarkar, B., Ullah, M. A., Johora, F. T., Taniya, M. A., and Araf, Y. (2020) The Essential Facts of Wuhan Novel Coronavirus Outbreak in China and Epitope-based Vaccine Designing against 2019-nCoV. bioRxiv doi.org/10.1101/2020.02.05.935072
74. Feng, Y., Maity, R., Whitelegge, J. P., Hadjikyriacou, A., Li, Z., Zurita-Lopez, C., Al-Hadid, Q., Clark, A. T., Bedford, M. T., Masson, J. Y., and Clarke, S. G. (2013) Mammalian protein arginine methyltransferase 7 (PRMT7) specifically targets RXR sites in lysine- and arginine-rich regions. J Biol Chem 288, 37010-37025
75. Gao, B., Gong, X., Fang, S., Weng, W., Wang, H., Chu, H., Sun, Y., Meng, C., Tan, L., Song, C., Qiu, X., Liu, W., Forlenza, M., Ding, C., and Liao, Y. (2021) Inhibition of anti-viral stress granule formation by coronavirus endoribonuclease nsp15 ensures efficient virus replication. PLoS Pathog 17, e1008690
76. Rabouw, H. H., Langereis, M. A., Knaap, R. C., Dalebout, T. J., Canton, J., Sola, I., Enjuanes, L., Bredenbeek, P. J., Kikkert, M., de Groot, R. J., and van Kuppeveld, F. J. (2016) Middle East Respiratory Coronavirus Accessory Protein 4a Inhibits PKR-Mediated Antiviral Stress Responses. PLoS Pathog 12, e1005982
77. Nakagawa, K., Narayanan, K., Wada, M., and Makino, S. (2018) Inhibition of Stress Granule Formation by Middle East Respiratory Syndrome Coronavirus 4a Accessory Protein Facilitates Viral Translation, Leading to Efficient Virus Replication. J Virol 92, e00902-00918
78. Wu, C. H., Chen, P. J., and Yeh, S. H. (2014) Nucleocapsid phosphorylation and RNA helicase DDX1 recruitment enables coronavirus transition from discontinuous to continuous transcription. Cell Host Microbe 16, 462-472
79. Nikolakaki, E., and Giannakouros, T. (2020) SR/RS Motifs as Critical Determinants of Coronavirus Life Cycle. Front Mol Biosci 7, 219
80. Corbet, G. A., and Parker, R. (2019) RNP Granule Formation: Lessons from P-Bodies and Stress Granules. Cold Spring Harb Symp Quant Biol 84, 203-215

81. Qamar, S., Wang, G., Randle, S. J., Ruggeri, F. S., Varela, J. A., Lin, J. Q., Phillips, E. C., Miyashita, A., Williams, D., Strohl, F., Meadows, W., Ferry, R., Dardov, V. J., Tartaglia, G. G., Farrer, L. A., Kaminski Schierle, G. S., Kaminski, C. F., Holt, C. E., Fraser, P. E., Schmitt-Ulms, G., Klenerman, D., Knowles, T., Vendruscolo, M., and St George-Hyslop, P. (2018) FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation-π Interactions. Cell 173, 720-734
82. Hofweber, M., Hutten, S., Bourgeois, B., Spreitzer, E., Niedner-Boblenz, A., Schifferer, M., Ruepp, M. D., Simons, M., Niessing, D., Madl, T., and Dormann, D. (2018) Phase Separation of FUS Is Suppressed by Its Nuclear Import Receptor and Arginine Methylation. Cell 173, 706-719
83. Yamagata, K., Daitoku, H., Takahashi, Y., Namiki, K., Hisatake, K., Kako, K., Mukai, H., Kasuya, Y., and Fukamizu, A. (2008) Arginine methylation of FOXO transcription factors inhibits their phosphorylation by Akt. Mol Cell 32, 221-231
84. Hsu, J. M., Chen, C. T., Chou, C. K., Kuo, H. P., Li, L. Y., Lin, C. Y., Lee, H. J., Wang, Y. N., Liu, M., Liao, H. W., Shi, B., Lai, C. C., Bedford, M. T., Tsai, C. H., and Hung, M. C. (2011) Crosstalk between Arg 1175 methylation and Tyr 1173 phosphorylation negatively modulates EGFR-mediated ERK activation. Nat Cell Biol 13, 174-181
85. Tugaeva, K. V., Hawkins, D. E., Smith, J. L., Bayfield, O. W., Ker, D. S., Sysoev, A. A., Klychnikov, O. I., Antson, A. A., and Sluchanko, N. N. (2021) The Mechanism of SARS-CoV-2 Nucleocapsid Protein Recognition by the Human 14-3-3 Proteins. J Mol Biol 433, 166875
86. Omer, A., Barrera, M. C., Moran, J. L., Lian, X. J., Di Marco, S., Beausejour, C., and Gallouzi, I. E. (2020) G3BP1 controls the senescence-associated secretome and its impact on cancer progression. Nat Commun 11, 4979
87. Boisvert, F. M., Cote, J., Boulanger, M. C., Cleroux, P., Bachand, F., Autexier, C., and Richard, S. (2002) Symmetrical dimethylarginine methylation is required for the localization of SMN in Cajal bodies and pre-mRNA splicing. J Cell Biol 159, 957-969
88. Yang, Y., McBride, K. M., Hensley, S., Lu, Y., Chedin, F., and Bedford, M. T. (2014) Arginine methylation facilitates the recruitment of TOP3B to chromatin to prevent R loop accumulation. Mol Cell 53, 484-497
89. Corman, V. M., Landt, O., Kaiser, M., and al, e. (2020) Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. . Euro Surveill 25, 2000045.
90. Song, J., and Richard, S. (2015) Sam68 Regulates S6K1 Alternative Splicing during Adipogenesis. Mol Cell Biol 35, 1926-1939
91. Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., and al, e. (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 129, 1311-1323

3.8 Supplemental information

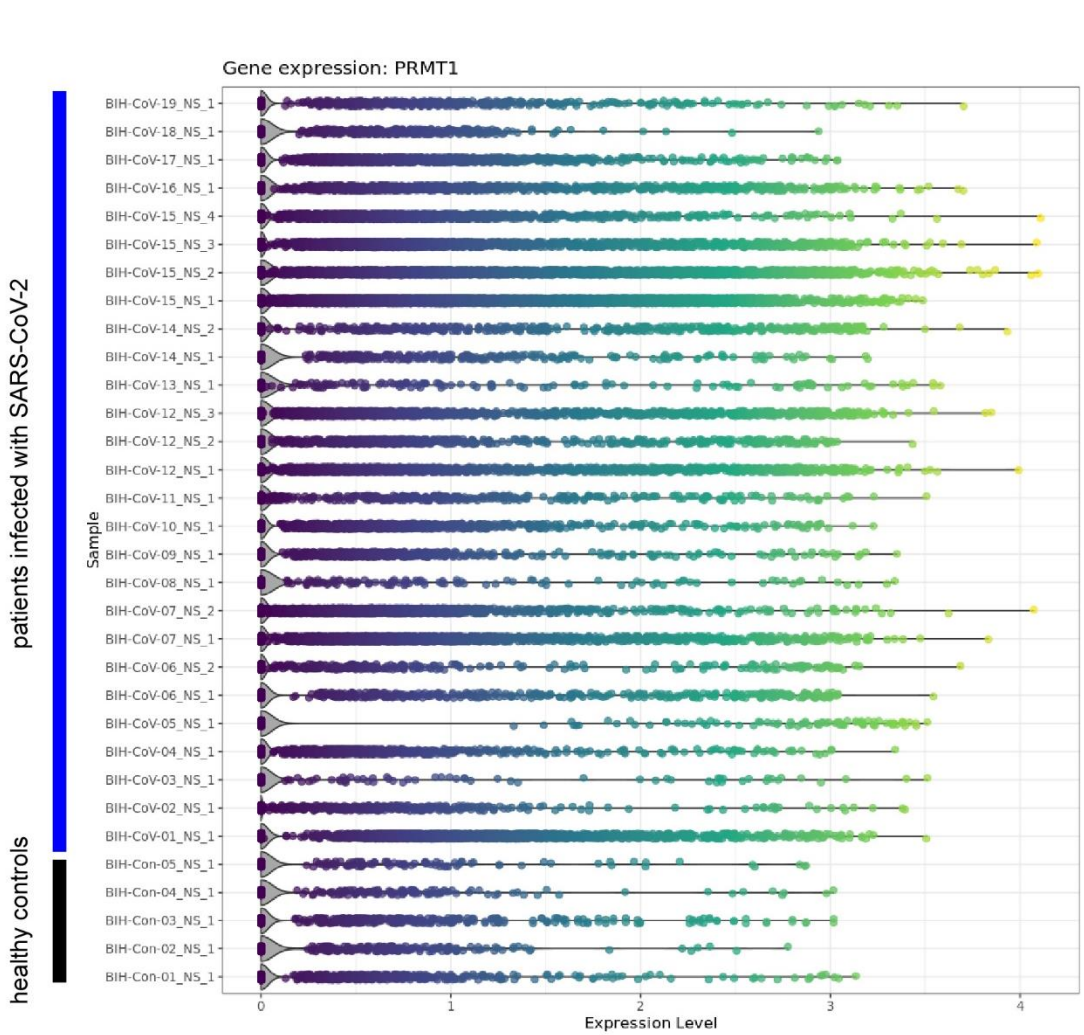


Figure S3. 1 Increased PRMT1 expression in SARS-CoV-2 infected individuals.
Expression of PRMT1 in single cell RNA seq data generated with nasopharyngeal and bronchial samples from 19 clinically well-characterized patients and five healthy controls (https://digital.bihealth.org).

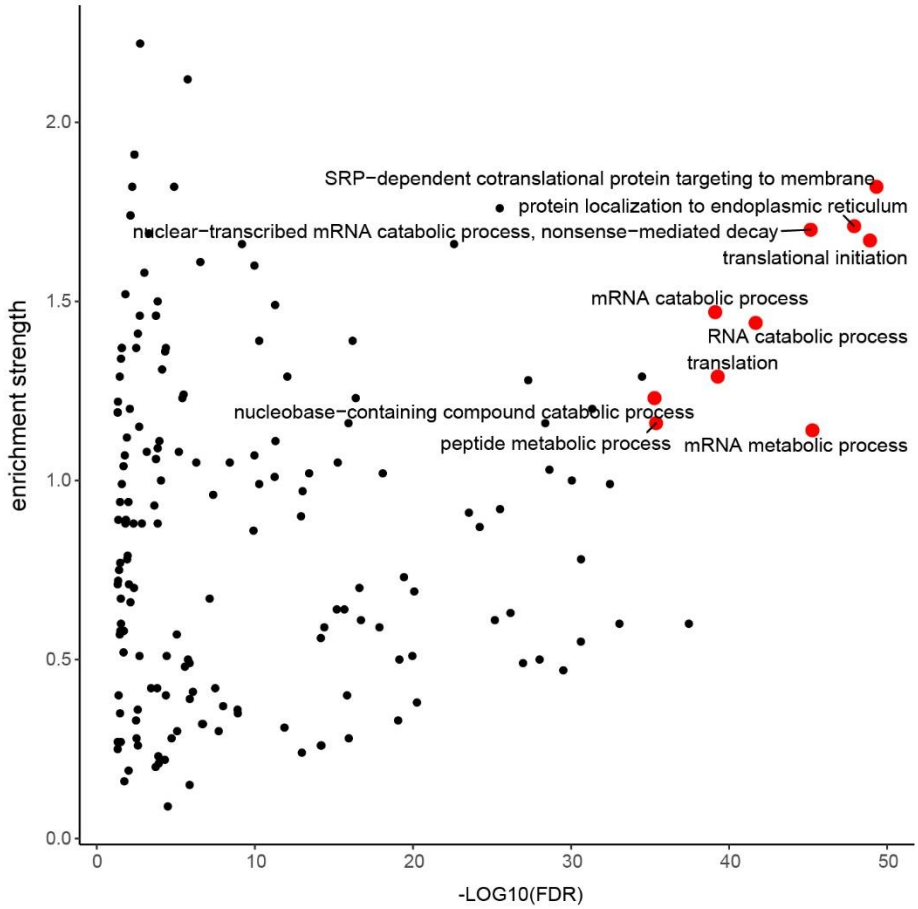


Figure S3. 2 GO analysis of SARS-CoV-2 N protein interactome.
Biological process (Gene Ontology) was performed using STRING analysis. Significant enriched pathway is highlighted in red and defined as -LOG_{10}(FDR) >35 and strength >1.

Table 3. 1 RGG/RG motif containing proteins within SARS-CoV-2 N interactome.

ID	Name	RNA binding
Q14444	CAPRIN1	Yes
Q8TDD1	DDX54	Yes
Q6P158	DHX57	Yes
Q08211	DHX9	Yes
Q01844	EWSR1	Yes
Q9NZB2	FAM120A	Yes
Q8NCA5	FAM98A	Yes
P22087	FBL	Yes
P51114	FXR1	Yes
Q13283	G3BP1	Yes
Q9UN86	G3BP2	Yes
O14979	HNRNPDL	Yes
Q9BUJ2	HNRNPUL1	Yes

Q1KMD3	HNRNPUL2	Yes
Q14103	HNRNPD	Yes
O60506	SYNCRIP	Yes
O43390	HNRNPR	Yes
Q00839	HNRNPU	Yes
Q9NZI8	IGF2BP1	Yes
Q12905	ILF2	Yes
Q12906	ILF3	Yes
Q8NC51	SERBP1	Yes
Q99575	POP1	Yes
P38159	RBMX	Yes
P09651	HNRNPA1	Yes
P51991	HNRNPA3	Yes
P46783	RPS10	Yes

CHAPTER 4 General Discussion

RBMX and SARS-CoV-2 N protein contain RGG/RG motifs that are methylated by PRMTs. In this thesis, I have characterized the role of these RGG/RG motifs and their implication in X-LID and COVID-19, respectively.

The RBMX C-terminus RGG/RG motif is methylated by PRMT5 and may be involved in liquid-liquid phase separation (LLPS).

Intrinsically disordered regions (IDRs) containing proteins are important for LLPS due to their ability to form multivalent interactions with proteins and RNA (Uversky et al., 2015). For example, IDRs from DDX4, hnrnpA1 and FUS are sufficient to initiate phase separation *in vitro* (Molliex et al., 2015; Nott et al., 2015; Patel et al., 2015). Arginines within the RGG/RG motifs can form cation-π interaction with tyrosine, which was observed in FUS LLPS (Hofweber et al., 2018; Qamar et al., 2018). RBMX has an IDR of 303 amino acids, with two RGG/RG motifs and a long stretch of repeated tyrosine-glycine and tyrosine-serine (YG/YS)-rich sequences, implicating that RBMX potentially forms higher-ordered assembly through its RGG/RG motifs interacting with th YG/YS repeats. Indeed, RBMX forms nuclear puncta and the treatment of the cells with 1,6-hexanediol (1,6-HD), an alcohol known to dissolve the LLPS assembly, significantly reduced the size and the fluorescence intensity of RBMX nuclear puncta. These findings suggest that LLPS is likely involved in RBMX nuclear speckle formation.

Our mass spectrometry analysis identified SRSF1 as an interactor for RBMX. Treatment of 1,6-HD dissociated SRSF1 from RBMX nuclear puncta and reduced SRSF1 RNA binding. It has been suggested that splicing factors harboring IDRs contribute to splicing regulation through LLPS (Guo et al., 2019). For example, the splicing factor RBFOX forms higher-order assembly through multiple tyrosines in its IDR. Mutation of tyrosines responsible for π-π interaction causes

disruption of RBFOX assembly, leading to defective splicing of its target genes (Damianov et al., 2016). Another nuclear protein, AKAP95, also forms liquid-like condensates, creating a micro-compartment with splicing factors at a high local concentration to ensure optimal splicing activity (Li et al., 2020a).

In vitro phase separation assays are required to further characterize RBMX LLPS. Multiple arginines to lysines substitutions can be used to understand the contribution of the methylation to RBMX LLPS which may be influenced by the RRM domain and the length of IDR. In fact, it was suggested that the length of an IDR correlates with the ability of proteins to phase separate (Quiroz and Chilkoti, 2015). Moreover, protein domains can largely impact the dynamics of LLPS (Riback et al., 2017). Additionally, multivalent protein-protein and protein-RNA interactions can be drivers of biological phase separation (Dignon et al., 2020). To investigate the contribution of mixed composition to dynamic interaction, *in vitro* phase separation assays could be performed, as in previous studies (Kaur et al., 2021; Yang et al., 2020), using a range of RBMX/RNA ratios, with different SR-rich family proteins. Fluorescence recovery after photobleaching (FRAP) is routinely applied to assess the dynamic of LLPS *in vivo* (Alberti et al., 2019). As the size of the RRM domain is smaller than fluorescent tags such as green fluorescent protein (GFP), a tag that does not affect the RBMX LLPS should be considered for live imaging.

MDM4 alternative splicing regulation

PRMT5 mediated *MDM4* alternative splicing is known to regulate the p53 pathway (Bezzi et al., 2013). Depletion or inhibition of PRMT5 causes aberrant alternative splicing in neurons (Bezzi et al., 2013), which is the underlying mechanism of PRMT5 inhibitor-based cancer therapy (Fong et al., 2019; Gerhart et al., 2018). PRMT5 inhibitors cause splicing change of *MDM4*, generating a short isoform with a premature stop codon, and subsequently reducing MDM4

protein level. Increased p53 protein level and apoptosis were observed in hundreds of cancer cell lines treated with PRMT5 inhibitors, including the p53 mutated cells (Gerhart et al., 2018). However, the mechanism of PRMT5 mediated *MDM4* alternative splicing remains elusive. It is well-established that PRMT5 regulates snRNPs assembly by symmetrical methylation of arginines on Sm proteins. snRNPs participate in all splicing events, not just alternative splicing, therefore why are certain splicing event more sensitive to PRMT5 inhibition? We found that SRSF1 regulates *MDM4* exon 6 inclusion, which was confirmed by PAR-CLIP and mutagenesis assay (Appendix I). A recent study demonstrated that SRSF1 is a PRMT5 substrate, and PRMT5 knockdown affects SRSF1 binding to the target RNA (Radzisheuskaya et al., 2019). However, whether the arginine methylation sites in SRSF1 directly contribute to its RNA binding was not determined. In our study, we show that RBMX associates with SRSF1 to facilitate its interaction with RNA. Depletion of RBMX, similar to PRMT5 knockdown, reduced SRSF1 binding to the *MDM4* pre-mRNA.

We also identified a splicing silencer that controls exon 6 inclusion. Mutation of the predicted Tra2β RNA binding site upstream from 3'ss exon 6 led to exon inclusion (Appendix I). The precise mechanism of how this splicing silencer suppresses exon 6 inclusion will require further investigation.

Together, we provide the molecular basis for an antisense oligonucleotide targeting strategy where targeting the splicing silencer will restore MDM4 protein level. Furthermore, the targeting of the SRSF1 RNA binding site will reduce MDM4 protein levels which will activate the p53 pathway.

RBMX regulates the p53 pathway and neural differentiation

We identified a functional link between RBMX and the p53 pathway. Excessive apoptosis caused by loss-of-function of RBMX in U2OS and iPSCs was also observed in leukemia (Prieto et al., 2021). Fine-tuned regulation of the p53 pathway is essential for normal brain development. Abnormal activation of p53 and excessive apoptosis are considered as a common secondary outcome of genetic mutations that cause brain developmental defects (Bowen and Attardi, 2019). Interestingly, in RBMX-ΔRGG and RBMX knockdown NPCs, the p53 pathway was not activated. This was due to the high basal expression of p53 in the NPCs in comparison to iPSCs. Indeed, it has been reported that loss of p53 activity facilitates the reprogramming of somatic cells to hiPSCs (Zhao et al., 2008). Consistently, the transcriptome profile from the developing human brain showed that p53 is upregulated during cortical neuron differentiation and decreases afterward, indicating that p53 is required in early neuron development (Hawrylycz et al., 2012). Although RBMX-ΔRGG iPSCs showed aberrant p53 pathway activation and subsequent apoptosis, genes involved in iPSC pluripotency were not significantly affected, including *Nanog*, *Sox2*, and *Oct4*. These RBMX-ΔRGG iPSCs were able to differentiate into NPCs, expressing major differentiation markers. However, during NPC to neuron transition, RBMX-ΔRGG NPCs exhibited differentiation deficiency, showing reduced total number of neurons and aberrant neuron morphology with shorter neurite outgrowth and decreased neurite branches. These outcomes may involve cellular processes other than MDM4/p53 mediated apoptosis.

Another way to study the function of RGG/RG motif in brain development is to generate the *Rbmx*$^{\Delta RGG}$ mice model. As we observed the absence of vGLUT1 positive excitatory neurons in cortical lineage neuron differentiation, single-cell RNA-sequencing will be a useful tool to enable unbiased discovery and comprehensive characterization of neuronal cell types in *Rbmx*$^{\Delta RGG}$ mice.

Furthermore, using brain slice electrophysiology will allow the study of features of synapses and neural circuits.

Notably, human RBMXL1 (hRBMXL1) lacks the C-terminal RGG/RG motif comparing to hRBMX, whereas the mouse RBMXL1 (mRBMXL1) contains the C-terminal RGG/RG motif. In *Rbmx* knockout murine leukemia cells, mRBMXL1 can regulate its mRNA level and restore the RBMX protein level(Prieto et al., 2021). Therefore, depletion of the sequence encoding RGG/RG motif of m*RBMXL1* in mouse cells is required to recapitulate the disease phenotype in human.

SARS-CoV-2 N protein is methylated by PRMT1

Host-mediated PTMs affect virus replication by regulating viral proteins antigenicity and virulence properties (Kumar et al., 2020). PTMs, including phosphorylation, arginine methylation, ubiquitination, SUMOylation, ISGylation and Glycosylation were found in virus proteins, playing diverse roles in eliminating the virus from the host or promoting virus survival (Kumar et al., 2020). SARS-CoV-2 N protein was reported to be phosphorylated by SRPK and GSK-3 (Carlson et al., 2020). While examining the N protein sequences, we identified several RG and RGG repeats within its N-terminal domain RNA binding domain(NTD) , C-terminal dimerization domain (CTD) and SR rich link region. We demonstrated that R95 and R177 can be methylated *in vitro* and *vivo*, located within NTD and SR-rich link regions. Both R95 and R177 are located in the U shape RNA binding epitope on the N-NTD surface, circumventing the base of the positively charged finger (Dinesh et al., 2020). Our mass spectrometry data revealed a nearby serine S178 that can be phosphorylated. It is not clear whether R177 methylation and S178 phosphorylation have an impact on one another.

We demonstrated that N protein is a substrate of PRMT1 and is further indicated by direct interaction between the two proteins while treating with type I PRMT inhibitor MS023. This observation is also important to better understand the mechanism of type I inhibitors. Further experiments are required to investigate whether this phenotype can be recapitulated in other type I PRMT inhibitors or in PRMT1 catalytic domain mutation-expressed cells.

The methylated viral epitopes may initiate autoimmune disease in the COVID-19 infected patient. We used a site specific anti-methylarginine MRE11 antibody (ASYM26) to detect SARS-CoV-2 N protein methylation and indeed N protein found in virions was recognized by ASYM26 confirming N protein methylation. This cross-reactivity was possible, as N protein and nuclear protein MRE11 share a common epitope 'SRGGS'. Thus, viral proteins can provide epitopes resembling self-antigens (molecular mimicry) when post-translationally modified that can trigger virus-induced autoimmunity (Yang et al., 2018). It is known that sera from patients with systemic lupus erythematosus (SLE) are known to stain Cajal bodies in the nucleus and splicing speckles (Andrade et al., 1991). The epitopes were mapped to methyl-p80 coilin (Hebert et al., 2002) and methy-SmD1 and SmD3 RGG/RG motifs (Brahms et al., 2000). Therefore, detection of autoantibodies and self-reacting T cells within SARS-CoV-2 infected patients' sera would provide valuable information for preventing or defining the risk factors for future autoimmune diseases post-infection.

RGG/RG motif affects N protein function

The primary function of N protein is to bind with the SARS-CoV-2 gRNA. N protein specifically binds to the 5' and 3' of its gRNA promoting LLPS, while forms non-specific electrostatic interactions with other gRNA sequences to facilitate fluidity and solubilization, eliminating entanglement of the large gRNA molecule (Iserman et al., 2020). We showed that

both R95 and R177 contribute to N protein interaction with 5' of gRNA, as N protein (R to K) mutant fails to bind with 5'UTR of gRNA. Therefore, the two arginines probably contribute to sequence-specific interaction. Whether they also contribute to the non-specific binding of gRNA through LLPS remains to be investigated.

Repressing the stress granules (SGs) formation is important to evade the host innate immune response adapted by several viruses (Miller, 2011). By interfering with SGs formation, virus releases the mRNA translation initiation factors and ribosomal subunits that can be utilized by virus mRNA for protein synthesis. Moreover, virus replication can benefit from suppressing of SGs induced apoptosis. Unveiled by other researchers, N protein has a strong affinity towards SGs core protein G3BP1 and G3BP2 (Gordon et al., 2020; Luo et al., 2021; Nabeel-Shah et al., 2020). We confirmed that overexpression of N protein suppresses arsenite induced SGs formation, and arginine methylation is required for N protein to exert this function. We showed that the interaction between N protein and G3BP1 is mediated by RNA. Note that gRNA was absent in our model, and N protein was possibly sequestered host G3BP1 binding RNA. Indeed, mutating amino acids essential for NTD RNA binding function attenuate the inhibitory effect of N protein on G3BP1-mediated LLPS (Huang et al., 2021). The fact that N protein and G3BP1 have similar RNA-binding properties may explain why N protein is able to dissolve SGs. Substitution of NTD with cytoplasmic activation proliferation-associated protein-1 (Caprin-1) RRM domain loses the capacity to dissolve SGs, while a similar RRM from human antigen R (HuR) maintains the capacity (Huang et al., 2021). Moreover, it was suggested that N protein regulates host transcription process through interacting with the mRNA (Nabeel-Shah et al., 2020).

A limitation of my results is that the mechanisms described herein may not accurately reflect the actual SARS-CoV-2 infection. Further investigation is required to investigate the arginine

methylation of SARS-CoV-2 N protein in an animal infection model and test the type I PRMT inhibitors in these models. Moreover, a biosafety level-2 cell culture system to produce transcription and replication-competent SARS-CoV-2 virus-like-particles (trVLP) was recently developed (Ju et al., 2021), and it will be a valuable system to study arginine methylation in the viral life cycle.

RGG/RG motif-containing proteins function together for cellular process regulation

Both the RBMX interactome and the N protein interactome consist of many other RGG/RG motif-containing proteins. RBMX interacts with splicing factors such as SRSF1, Sam68 and hnRNP family proteins that are known substrates of PRMTs; SGs components that interact with N protein also contain multiple RGG/RG motifs, including G3BP1/2, cell cycle-associated protein 1 (CAPRIN1), and family with sequence similarity 98 member A (FAM98A). It is intriguing to understand how RGG/RG motif proteins interact with each other, as the motif is positively charged (Thandapani et al., 2013). A possibility is that these proteins interact with RNA, forming multivalent assembly. Indeed, both RBMX/Sam68 and N protein/G3BP1 interaction are bridged by RNA. Therefore, performing mass spectrometry assay with RNase A treated lysate could help to test the hypothesis. Besides, these RGG/RG motifs containing proteins adopt aromatic amino acids that form π-π by inter- and intra-molecule interactions. Indeed, purified RGG/RG motif can undergo LLPS without RNA (Elbaum-Garfinkle et al., 2015). Finally, as these RGG/RG motif-containing proteins are methylated and function together, PRMTs likely regulate cellular processes through affecting protein assemblies.

Contribution to Original Knowledge

The studies presented in the thesis focused on defining the roles of RGG/RG motif and arginine methylation of two disease related RBPs, RBMX and SARS-CoV-2 N protein. The functional significance of the RGG/RG motifs has been addressed. The major contributions to original knowledge are summarized below:

1. RBMX-ΔRGG iPS cell lines were generated to study the Shashi-X-linked intellectual disability.

2. RBMX is arginine methylated at the multiple sites and the C-terminal RGG/RG motif is methylated by PRMT5.

3. We report the first RBMX interactome data, as well as gene profiling data and alternative splicing data related to C- terminal RGG/RG motif deletion.

4. p53 pathway activation and excessive apoptosis were observed in RBMX knockdown immortalized cell lines and RBMX-ΔRGG iPSCs.

5. PRMT5 regulates RBMX colocalization with SRSF1 in nuclear puncta and RBMX facilitates SRSF1 binding to *MDM4* pre-mRNA

6. We characterized splicing factors that contribute to *MDM4* alternative splicing regulation.

7. Neuronal defects are observed in cortical neural differentiation using RBMX-ΔRGG iPSCs.

8. R95 and R177 of SARS-CoV-2 N protein are methylated by PRMT1.

9. Virus isolated from the supernatant of infected cell culture confirmed asymmetrical methylation by host PRMT1.

10. The methylation of N protein is required for N protein binding to gRNA.

11. N protein interacts with SGs components and represses SGs formation. This function was diminished by substitution of arginine to lysine or treatment with PRMT1 inhibitor.

12. PRMT1 inhibitor is contributing to reduce SARS-CoV-2 viral replication in infected veroE6 cells.

References

Adamson, B., Smogorzewska, A., Sigoillot, F.D., King, R.W., and Elledge, S.J. (2012). A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response. Nat Cell Biol *14*, 318-328.

Agolini, E., Dentici, M.L., Bellacchio, E., Alesi, V., Radio, F.C., Torella, A., Musacchia, F., Tartaglia, M., Dallapiccola, B., Nigro, V., *et al.* (2018). Expanding the clinical and molecular spectrum of PRMT7 mutations: 3 additional patients and review. Clin Genet *93*, 675-681.

Alberti, S., Gladfelter, A., and Mittag, T. (2019). Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. Cell *176*, 419-434.

Altmeyer, M., Neelsen, K.J., Teloni, F., Pozdnyakova, I., Pellegrino, S., Grofte, M., Rask, M.D., Streicher, W., Jungmichel, S., Nielsen, M.L., *et al.* (2015). Liquid demixing of intrinsically disordered proteins is seeded by poly(ADP-ribose). Nat Commun *6*, 8088.

Amano, G., Matsuzaki, S., Mori, Y., Miyoshi, K., Han, S., Shikada, S., Takamura, H., Yoshimura, T., and Katayama, T. (2020). SCYL1 arginine methylation by PRMT1 is essential for neurite outgrowth via Golgi morphogenesis. Mol Biol Cell *31*, 1963-1973.

Andrade, L.E., Chan, E.K., Raska, I., Peebles, C.L., Roos, G., and Tan, E. (1991). Human autoantibody to a novel protein of the nuclear coiled body: immunological characterization and cDNA cloning of p80-coilin. J Exp Med *173*, 1407-1419.

Antonysamy, S., Bonday, Z., Campbell, R.M., Doyle, B., Druzina, Z., Gheyi, T., Han, B., Jungheim, L.N., Qian, Y., Rauch, C., *et al.* (2012). Crystal structure of the human PRMT5:MEP50 complex. Proc Natl Acad Sci U S A *109*, 17960-17965.

Bao, J., Di Lorenzo, A., Lin, K., Lu, Y., Zhong, Y., Sebastian, M.M., Muller, W.J., Yang, Y., and Bedford, M.T. (2019). Mouse Models of Overexpression Reveal Distinct Oncogenic Roles for Different Type I Protein Arginine Methyltransferases. Cancer Res *79*, 21-32.

Bao, X., Siprashvili, Z., Zarnegar, B.J., Shenoy, R.M., Rios, E.J., Nady, N., Qu, K., Mah, A., Webster, D.E., Rubin, A.J., *et al.* (2017). CSNK1a1 Regulates PRMT1 to Maintain the Progenitor State in Self-Renewing Somatic Tissue. Dev Cell *43*, 227-239 e225.

Barekatain, Y., Ackroyd, J.J., Yan, V.C., Khadka, S., Wang, L., Chen, K.C., Poral, A.H., Tran, T., Georgiou, D.K., Arthur, K., *et al.* (2021). Homozygous MTAP deletion in primary human glioblastoma is not associated with elevation of methylthioadenosine. Nat Commun *12*, 4228.

Becker, J.S., McCarthy, R.L., Sidoli, S., Donahue, G., Kaeding, K.E., He, Z., Lin, S., Garcia, B.A., and Zaret, K.S. (2017). Genomic and Proteomic Resolution of Heterochromatin and Its Restriction of Alternate Fate Genes. Mol Cell *68*, 1023-1037 e1015.

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

Bezzi, M., Teo, S.X., Muller, J., Mok, W.C., Sahu, S.K., Vardy, L.A., Bonday, Z.Q., and Guccione, E. (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*, 1903-1916.

Bian, L., Meng, Y., Zhang, M., and Li, D. (2019). MRE11-RAD50-NBS1 complex alterations and DNA damage response: implications for cancer treatment. Mol Cancer *18*, 169.

Blanc, R.S., and Richard, S. (2017). Arginine Methylation: The Coming of Age. Mol Cell *65*, 8-24.

Blanc, R.S., Vogel, G., Chen, T., Crist, C., and Richard, S. (2016). PRMT7 Preserves Satellite Cell Regenerative Capacity. Cell Rep *14*, 1528-1539.

Boisvert, F.M., Cote, J., Boulanger, M.C., and Richard, S. (2003). A proteomic analysis of arginine-methylated protein complexes. Mol Cell Proteomics *2*, 1319-1330.
Boisvert, F.M., Rhie, A., Richard, S., and Doherty, A.J. (2005). The GAR motif of 53BP1 is arginine methylated by PRMT1 and is necessary for 53BP1 DNA binding activity. Cell Cycle *4*, 1834-1841.
Bouwman, P., Aly, A., Escandell, J.M., Pieterse, M., Bartkova, J., van der Gulden, H., Hiddingh, S., Thanasoula, M., Kulkarni, A., Yang, Q., *et al.* (2010). 53BP1 loss rescues BRCA1 deficiency and is associated with triple-negative and BRCA-mutated breast cancers. Nat Struct Mol Biol *17*, 688-695.
Bowen, M.E., and Attardi, L.D. (2019). The role of p53 in developmental syndromes. J Mol Cell Biol *11*, 200-211.
Brahms, H., Meheus, L., de Brabandere, V., Fischer, U., and Luhrmann, R. (2001). Symmetrical dimethylation of arginine residues in spliceosomal Sm protein B/B' and the Sm-like protein LSm4, and their interaction with the SMN protein. RNA *7*, 1531-1542.
Brahms, H., Raymackers, J., Union, A., de Keyser, F., Meheus, L., and Luhrmann, R. (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*, 17122-17129.
Branscombe, T.L., Frankel, A., Lee, J.H., Cook, J.R., Yang, Z., Pestka, S., and Clarke, S. (2001). PRMT5 (Janus kinase-binding protein 1) catalyzes the formation of symmetric dimethylarginine residues in proteins. J Biol Chem *276*, 32971-32976.
Buratti, E., Chivers, M., Kralovicova, J., Romano, M., Baralle, M., Krainer, A.R., and Vorechovsky, I. (2007). Aberrant 5' splice sites in human disease genes: mutation pattern, nucleotide structure and comparison of computational tools that predict their utilization. Nucleic Acids Res *35*, 4250-4263.
Calabretta, S., Vogel, G., Yu, Z., Choquet, K., Darbelli, L., Nicholson, T.B., Kleinman, C.L., and Richard, S. (2018). Loss of PRMT5 Promotes PDGFRalpha Degradation during Oligodendrocyte Differentiation and Myelination. Dev Cell *46*, 426-440 e425.
Carlson, C.R., Asfaha, J.B., Ghent, C.M., Howard, C.J., Hartooni, N., Safari, M., Frankel, A.D., and Morgan, D.O. (2020). Phosphoregulation of Phase Separation by the SARS-CoV-2 N Protein Suggests a Biophysical Basis for its Dual Functions. Mol Cell *80*, 1092-1103 e1094.
Chan-Penebre, E., Kuplast, K.G., Majer, C.R., Boriack-Sjodin, P.A., Wigle, T.J., Johnston, L.D., Rioux, N., Munchhof, M.J., Jin, L., Jacques, S.L., *et al.* (2015). A selective inhibitor of PRMT5 with in vivo and in vitro potency in MCL models. Nat Chem Biol *11*, 432-437.
Chang, C.K., Chen, C.M., Chiang, M.H., Hsu, Y.L., and Huang, T.H. (2013). Transient oligomerization of the SARS-CoV N protein--implication for virus ribonucleoprotein packaging. PLoS One *8*, e65045.
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, H., Cui, Y., Han, X., Hu, W., Sun, M., Zhang, Y., Wang, P.H., Song, G., Chen, W., and Lou, J. (2020a). Liquid-liquid phase separation by SARS-CoV-2 nucleocapsid protein and RNA. Cell Res *30*, 1143-1145.
Chen, H., Lorton, B., Gupta, V., and Shechter, D. (2017). A TGFbeta-PRMT5-MEP50 axis regulates cancer cell invasion through histone H3 and H4 arginine methylation coupled transcriptional activation and repression. Oncogene *36*, 373-386.

Chen, K., Xiao, F., Hu, D., Ge, W., Tian, M., Wang, W., Pan, P., Wu, K., and Wu, J. (2020b). SARS-CoV-2 Nucleocapsid Protein Interacts with RIG-I and Represses RIG-Mediated IFN-beta Production. Viruses *13*.

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.

Cheung, N., Fung, T.K., Zeisig, B.B., Holmes, K., Rane, J.K., Mowen, K.A., Finn, M.G., Lenhard, B., Chan, L.C., and So, C.W. (2016). Targeting Aberrant Epigenetic Networks Mediated by PRMT1 and KDM4C in Acute Myeloid Leukemia. Cancer Cell *29*, 32-48.

Chiang, K., Zielinska, A.E., Shaaban, A.M., Sanchez-Bailon, M.P., Jarrold, J., Clarke, T.L., Zhang, J., Francis, A., Jones, L.J., Smith, S., *et al.* (2017). PRMT5 Is a Critical Regulator of Breast Cancer Stem Cell Function via Histone Methylation and FOXP1 Expression. Cell Rep *21*, 3498-3513.

Chiou, N.T., Shankarling, G., and Lynch, K.W. (2013). hnRNP L and hnRNP A1 induce extended U1 snRNA interactions with an exon to repress spliceosome assembly. Mol Cell *49*, 972-982.

Chong, P.A., Vernon, R.M., and Forman-Kay, J.D. (2018). RGG/RG Motif Regions in RNA Binding and Phase Separation. J Mol Biol *430*, 4650-4665.

Choucair, A., Pham, T.H., Omarjee, S., Jacquemetton, J., Kassem, L., Tredan, O., Rambaud, J., Marangoni, E., Corbo, L., Treilleux, I., *et al.* (2019). The arginine methyltransferase PRMT1 regulates IGF-1 signaling in breast cancer. Oncogene *38*, 4015-4027.

Christensen, M.H., Jensen, S.B., Miettinen, J.J., Luecke, S., Prabakaran, T., Reinert, L.S., Mettenleiter, T., Chen, Z.J., Knipe, D.M., Sandri-Goldin, R.M., *et al.* (2016). HSV-1 ICP27 targets the TBK1-activated STING signalsome to inhibit virus-induced type I IFN expression. EMBO J *35*, 1385-1399.

Clarke, T.L., Sanchez-Bailon, M.P., Chiang, K., Reynolds, J.J., Herrero-Ruiz, J., Bandeiras, T.M., Matias, P.M., Maslen, S.L., Skehel, J.M., Stewart, G.S., *et al.* (2017). PRMT5-Dependent Methylation of the TIP60 Coactivator RUVBL1 Is a Key Regulator of Homologous Recombination. Mol Cell *65*, 900-916 e907.

Cong, Y., Ulasli, M., Schepers, H., Mauthe, M., V'Kovski, P., Kriegenburg, F., Thiel, V., de Haan, C.A.M., and Reggiori, F. (2020). Nucleocapsid Protein Recruitment to Replication-Transcription Complexes Plays a Crucial Role in Coronaviral Life Cycle. J Virol *94*.

Cook, C.N., Wu, Y., Odeh, H.M., Gendron, T.F., Jansen-West, K., Del Rosso, G., Yue, M., Jiang, P., Gomes, E., Tong, J., *et al.* (2020). C9orf72 poly(GR) aggregation induces TDP-43 proteinopathy. Sci Transl Med *12*.

Corbett, K.S., Edwards, D.K., Leist, S.R., Abiona, O.M., Boyoglu-Barnum, S., Gillespie, R.A., Himansu, S., Schafer, A., Ziwawo, C.T., DiPiazza, A.T., *et al.* (2020). SARS-CoV-2 mRNA vaccine design enabled by prototype pathogen preparedness. Nature *586*, 567-571.

Corley, S.M., and Gready, J.E. (2008). Identification of the RGG box motif in Shadoo: RNA-binding and signaling roles? Bioinform Biol Insights *2*, 383-400.

Cubuk, J., Alston, J.J., Incicco, J.J., Singh, S., Stuchell-Brereton, M.D., Ward, M.D., Zimmerman, M.I., Vithani, N., Griffith, D., Wagoner, J.A., *et al.* (2021). The SARS-CoV-2 nucleocapsid protein is dynamic, disordered, and phase separates with RNA. Nat Commun *12*, 1936.

Cura, V., Troffer-Charlier, N., Wurtz, J.M., Bonnefond, L., and Cavarelli, J. (2014). Structural insight into arginine methylation by the mouse protein arginine methyltransferase 7: a zinc

finger freezes the mimic of the dimeric state into a single active site. Acta Crystallogr D Biol Crystallogr *70*, 2401-2412.
Cusco, I., Barcelo, M.J., del Rio, E., Baiget, M., and Tizzano, E.F. (2004). Detection of novel mutations in the SMN Tudor domain in type I SMA patients. Neurology *63*, 146-149.
Dacwag, C.S., Ohkawa, Y., Pal, S., Sif, S., and Imbalzano, A.N. (2007). The protein arginine methyltransferase Prmt5 is required for myogenesis because it facilitates ATP-dependent chromatin remodeling. Mol Cell Biol *27*, 384-394.
Damianov, A., Ying, Y., Lin, C.H., Lee, J.A., Tran, D., Vashisht, A.A., Bahrami-Samani, E., Xing, Y., Martin, K.C., Wohlschlegel, J.A., *et al.* (2016). Rbfox Proteins Regulate Splicing as Part of a Large Multiprotein Complex LASR. Cell *165*, 606-619.
Daujat, S., Bauer, U.M., Shah, V., Turner, B., Berger, S., and Kouzarides, T. (2002). Crosstalk between CARM1 methylation and CBP acetylation on histone H3. Curr Biol *12*, 2090-2097.
DeJesus-Hernandez, M., Mackenzie, I.R., Boeve, B.F., Boxer, A.L., Baker, M., Rutherford, N.J., Nicholson, A.M., Finch, N.A., Flynn, H., Adamson, J., *et al.* (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron *72*, 245-256.
Dhar, S., Vemulapalli, V., Patananan, A.N., Huang, G.L., Di Lorenzo, A., Richard, S., Comb, M.J., Guo, A., Clarke, S.G., and Bedford, M.T. (2013). Loss of the major Type I arginine methyltransferase PRMT1 causes substrate scavenging by other PRMTs. Sci Rep *3*, 1311.
Dichmann, D.S., Fletcher, R.B., and Harland, R.M. (2008). Expression cloning in Xenopus identifies RNA-binding proteins as regulators of embryogenesis and Rbmx as necessary for neural and muscle development. Dev Dyn *237*, 1755-1766.
Dickinson, M.E., Flenniken, A.M., Ji, X., Teboul, L., Wong, M.D., White, J.K., Meehan, T.F., Weninger, W.J., Westerberg, H., Adissu, H., *et al.* (2016). High-throughput discovery of novel developmental phenotypes. Nature *537*, 508-514.
Dignon, G.L., Best, R.B., and Mittal, J. (2020). Biomolecular Phase Separation: From Molecular Driving Forces to Macroscopic Properties. Annu Rev Phys Chem *71*, 53-75.
Dinesh, D.C., Chalupska, D., Silhan, J., Koutna, E., Nencka, R., Veverka, V., and Boura, E. (2020). Structural basis of RNA recognition by the SARS-CoV-2 nucleocapsid phosphoprotein. PLoS Pathog *16*, e1009100.
El Messaoudi, S., Fabbrizio, E., Rodriguez, C., Chuchana, P., Fauquier, L., Cheng, D., Theillet, C., Vandel, L., Bedford, M.T., and Sardet, C. (2006). Coactivator-associated arginine methyltransferase 1 (CARM1) is a positive regulator of the Cyclin E1 gene. Proc Natl Acad Sci U S A *103*, 13351-13356.
Elbaum-Garfinkle, S., Kim, Y., Szczepaniak, K., Chen, C.C., Eckmann, C.R., Myong, S., and Brangwynne, C.P. (2015). The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. Proc Natl Acad Sci U S A *112*, 7189-7194.
Elguindy, M.M., Kopp, F., Goodarzi, M., Rehfeld, F., Thomas, A., Chang, T.C., and Mendell, J.T. (2019). PUMILIO, but not RBMX, binding is required for regulation of genomic stability by noncoding RNA NORAD. Elife *8*.
Eram, M.S., Shen, Y., Szewczyk, M., Wu, H., Senisterra, G., Li, F., Butler, K.V., Kaniskan, H.U., Speed, B.A., Dela Sena, C., *et al.* (2016). A Potent, Selective, and Cell-Active Inhibitor of Human Type I Protein Arginine Methyltransferases. ACS Chem Biol *11*, 772-781.
Espinosa, J.R., Joseph, J.A., Sanchez-Burgos, I., Garaizar, A., Frenkel, D., and Collepardo-Guevara, R. (2020). Liquid network connectivity regulates the stability and composition of

biomolecular condensates with many components. Proc Natl Acad Sci U S A *117*, 13238-13247.
Evich, M., Stroeva, E., Zheng, Y.G., and Germann, M.W. (2016). Effect of methylation on the side-chain pKa value of arginine. Protein Sci *25*, 479-486.
Fabbrizio, E., El Messaoudi, S., Polanowska, J., Paul, C., Cook, J.R., Lee, J.H., Negre, V., Rousset, M., Pestka, S., Le Cam, A., *et al.* (2002). Negative regulation of transcription by the type II arginine methyltransferase PRMT5. EMBO Rep *3*, 641-645.
Fedoriw, A., Rajapurkar, S.R., O'Brien, S., Gerhart, S.V., Mitchell, L.H., Adams, N.D., Rioux, N., Lingaraj, T., Ribich, S.A., Pappalardi, M.B., *et al.* (2019). Anti-tumor Activity of the Type I PRMT Inhibitor, GSK3368715, Synergizes with PRMT5 Inhibition through MTAP Loss. Cancer Cell *36*, 100-114 e125.
Feng, Y., Maity, R., Whitelegge, J.P., Hadjikyriacou, A., Li, Z., Zurita-Lopez, C., Al-Hadid, Q., Clark, A.T., Bedford, M.T., Masson, J.Y., *et al.* (2013). Mammalian protein arginine methyltransferase 7 (PRMT7) specifically targets RXR sites in lysine- and arginine-rich regions. J Biol Chem *288*, 37010-37025.
Fischer, M. (2017). Census and evaluation of p53 target genes. Oncogene *36*, 3943-3956.
Flynn, R.A., Belk, J.A., Qi, Y., Yasumoto, Y., Wei, J., Alfajaro, M.M., Shi, Q., Mumbach, M.R., Limaye, A., DeWeirdt, P.C., *et al.* (2021). Discovery and functional interrogation of SARS-CoV-2 RNA-host protein interactions. Cell *184*, 2394-2411 e2316.
Fong, J.Y., Pignata, L., Goy, P.A., Kawabata, K.C., Lee, S.C., Koh, C.M., Musiani, D., Massignani, E., Kotini, A.G., Penson, A., *et al.* (2019). Therapeutic Targeting of RNA Splicing Catalysis through Inhibition of Protein Arginine Methylation. Cancer Cell *36*, 194-209 e199.
Fumagalli, L., Young, F.L., Boeynaems, S., De Decker, M., Mehta, A.R., Swijsen, A., Fazal, R., Guo, W., Moisse, M., Beckers, J., *et al.* (2021). C9orf72-derived arginine-containing dipeptide repeats associate with axonal transport machinery and impede microtubule-based motility. Sci Adv *7*.
Gao, G., Dhar, S., and Bedford, M.T. (2017). PRMT5 regulates IRES-dependent translation via methylation of hnRNP A1. Nucleic Acids Res *45*, 4359-4369.
Gehring, N.H., and Roignant, J.Y. (2021). Anything but Ordinary - Emerging Splicing Mechanisms in Eukaryotic Gene Regulation. Trends Genet *37*, 355-372.
Gendron, T.F., Bieniek, K.F., Zhang, Y.J., Jansen-West, K., Ash, P.E., Caulfield, T., Daughrity, L., Dunmore, J.H., Castanedes-Casey, M., Chew, J., *et al.* (2013). Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. Acta Neuropathol *126*, 829-844.
Gerhart, S.V., Kellner, W.A., Thompson, C., Pappalardi, M.B., Zhang, X.P., Montes de Oca, R., Penebre, E., Duncan, K., Boriack-Sjodin, A., Le, B., *et al.* (2018). Activation of the p53-MDM4 regulatory axis defines the anti-tumour response to PRMT5 inhibition through its role in regulating cellular splicing. Sci Rep *8*, 9711.
Geuens, T., Bouhy, D., and Timmerman, V. (2016). The hnRNP family: insights into their role in health and disease. Hum Genet *135*, 851-867.
Gkountela, S., Li, Z., Chin, C.J., Lee, S.A., and Clark, A.T. (2014). PRMT5 is required for human embryonic stem cell proliferation but not pluripotency. Stem Cell Rev Rep *10*, 230-239.
Gomes, E., and Shorter, J. (2019). The molecular language of membraneless organelles. J Biol Chem *294*, 7115-7127.

Gordon, D.E., Jang, G.M., Bouhaddou, M., Xu, J., Obernier, K., White, K.M., O'Meara, M.J., Rezelj, V.V., Guo, J.Z., Swaney, D.L., *et al.* (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature *583*, 459-468.
Goulet, I., Gauvin, G., Boisvenue, S., and Cote, J. (2007). Alternative splicing yields protein arginine methyltransferase 1 isoforms with distinct activity, substrate specificity, and subcellular localization. J Biol Chem *282*, 33009-33021.
Grossoehme, N.E., Li, L., Keane, S.C., Liu, P., Dann, C.E., 3rd, Leibowitz, J.L., and Giedroc, D.P. (2009). Coronavirus N protein N-terminal domain (NTD) specifically binds the transcriptional regulatory sequence (TRS) and melts TRS-cTRS RNA duplexes. J Mol Biol *394*, 544-557.
Guccione, E., and Richard, S. (2019). The regulation, functions and clinical relevance of arginine methylation. Nat Rev Mol Cell Biol *20*, 642-657.
Gueroussov, S., Weatheritt, R.J., O'Hanlon, D., Lin, Z.Y., Narula, A., Gingras, A.C., and Blencowe, B.J. (2017). Regulatory Expansion in Mammals of Multivalent hnRNP Assemblies that Globally Control Alternative Splicing. Cell *170*, 324-339 e323.
Gulla, A., Hideshima, T., Bianchi, G., Fulciniti, M., Kemal Samur, M., Qi, J., Tai, Y.T., Harada, T., Morelli, E., Amodio, N., *et al.* (2018). Protein arginine methyltransferase 5 has prognostic relevance and is a druggable target in multiple myeloma. Leukemia *32*, 996-1002.
Guo, J., Wang, X., Jia, J., and Jia, R. (2020). Underexpression of SRSF3 and its target gene RBMX predicts good prognosis in patients with head and neck cancer. J Oral Sci *62*, 175-179.
Guo, Y.E., Manteiga, J.C., Henninger, J.E., Sabari, B.R., Dall'Agnese, A., Hannett, N.M., Spille, J.H., Afeyan, L.K., Zamudio, A.V., Shrinivas, K., *et al.* (2019). Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. Nature *572*, 543-548.
Hadian, K., Vincendeau, M., Mausbacher, N., Nagel, D., Hauck, S.M., Ueffing, M., Loyter, A., Werner, T., Wolff, H., and Brack-Werner, R. (2009). Identification of a heterogeneous nuclear ribonucleoprotein-recognition region in the HIV Rev protein. J Biol Chem *284*, 33384-33391.
Haghandish, N., Baldwin, R.M., Morettin, A., Dawit, H.T., Adhikary, H., Masson, J.Y., Mazroui, R., Trinkle-Mulcahy, L., and Cote, J. (2019). PRMT7 methylates eukaryotic translation initiation factor 2alpha and regulates its role in stress granule formation. Mol Biol Cell *30*, 778-793.
Hamard, P.J., Santiago, G.E., Liu, F., Karl, D.L., Martinez, C., Man, N., Mookhtiar, A.K., Duffort, S., Greenblatt, S., Verdun, R.E., *et al.* (2018). PRMT5 Regulates DNA Repair by Controlling the Alternative Splicing of Histone-Modifying Enzymes. Cell Rep *24*, 2643-2657.
Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646-674.
Hasegawa, M., Toma-Fukai, S., Kim, J.D., Fukamizu, A., and Shimizu, T. (2014). Protein arginine methyltransferase 7 has a novel homodimer-like structure formed by tandem repeats. FEBS Lett *588*, 1942-1948.
Hashimoto, M., Murata, K., Ishida, J., Kanou, A., Kasuya, Y., and Fukamizu, A. (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*, 2237-2245.
Hawrylycz, M.J., Lein, E.S., Guillozet-Bongaarts, A.L., Shen, E.H., Ng, L., Miller, J.A., van de Lagemaat, L.N., Smith, K.A., Ebbert, A., Riley, Z.L., *et al.* (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. Nature *489*, 391-399.

He, X., Zhu, Y., Lin, Y.C., Li, M., Du, J., Dong, H., Sun, J., Zhu, L., Wang, H., Ding, Z., *et al.* (2019). PRMT1-mediated FLT3 arginine methylation promotes maintenance of FLT3-ITD(+) acute myeloid leukemia. Blood *134*, 548-560.

Hebert, M., Shpargel, K., Ospina, J., Tucker, K., and AG, M. (2002). Coilin methylation regulates nuclear body formation. Developmental Cell *3*, 329-337.

Heinrich, B., Zhang, Z., Raitskin, O., Hiller, M., Benderska, N., Hartmann, A.M., Bracco, L., Elliott, D., Ben-Ari, S., Soreq, H., *et al.* (2009). Heterogeneous nuclear ribonucleoprotein G regulates splice site selection by binding to CC(A/C)-rich regions in pre-mRNA. J Biol Chem *284*, 14303-14315.

Hilliker, A., Gao, Z., Jankowsky, E., and Parker, R. (2011). The DEAD-box protein Ded1 modulates translation by the formation and resolution of an eIF4F-mRNA complex. Mol Cell *43*, 962-972.

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

Hofweber, M., Hutten, S., Bourgeois, B., Spreitzer, E., Niedner-Boblenz, A., Schifferer, M., Ruepp, M.D., Simons, M., Niessing, D., Madl, T., *et al.* (2018). Phase Separation of FUS Is Suppressed by Its Nuclear Import Receptor and Arginine Methylation. Cell *173*, 706-719 e713.

Hollander, D., Donyo, M., Atias, N., Mekahel, K., Melamed, Z., Yannai, S., Lev-Maor, G., Shilo, A., Schwartz, S., Barshack, I., *et al.* (2016). A network-based analysis of colon cancer splicing changes reveals a tumorigenesis-favoring regulatory pathway emanating from ELK1. Genome Res *26*, 541-553.

Hsieh, P.K., Chang, S.C., Huang, C.C., Lee, T.T., Hsiao, C.W., Kou, Y.H., Chen, I.Y., Chang, C.K., Huang, T.H., and Chang, M.F. (2005). Assembly of severe acute respiratory syndrome coronavirus RNA packaging signal into virus-like particles is nucleocapsid dependent. J Virol *79*, 13848-13855.

Hsu, J.H., Hubbell-Engler, B., Adelmant, G., Huang, J., Joyce, C.E., Vazquez, F., Weir, B.A., Montgomery, P., Tsherniak, A., Giacomelli, A.O., *et al.* (2017). PRMT1-Mediated Translation Regulation Is a Crucial Vulnerability of Cancer. Cancer Res *77*, 4613-4625.

Hu, D., Gur, M., Zhou, Z., Gamper, A., Hung, M.C., Fujita, N., Lan, L., Bahar, I., and Wan, Y. (2015). Interplay between arginine methylation and ubiquitylation regulates KLF4-mediated genome stability and carcinogenesis. Nat Commun *6*, 8419.

Hu, Y., Li, W., Gao, T., Cui, Y., Jin, Y., Li, P., Ma, Q., Liu, X., and Cao, C. (2017). The Severe Acute Respiratory Syndrome Coronavirus Nucleocapsid Inhibits Type I Interferon Production by Interfering with TRIM25-Mediated RIG-I Ubiquitination. J Virol *91*.

Huang, B.W., Ray, P.D., Iwasaki, K., and Tsuji, Y. (2013). Transcriptional regulation of the human ferritin gene by coordinated regulation of Nrf2 and protein arginine methyltransferases PRMT1 and PRMT4. FASEB J *27*, 3763-3774.

Huang, L., Wang, Z., Narayanan, N., and Yang, Y. (2018). Arginine methylation of the C-terminus RGG motif promotes TOP3B topoisomerase activity and stress granule localization. Nucleic Acids Res *46*, 3061-3074.

Huang, S., Litt, M., and Felsenfeld, G. (2005). Methylation of histone H4 by arginine methyltransferase PRMT1 is essential in vivo for many subsequent histone modifications. Genes Dev *19*, 1885-1893.

Huang, W., Ju, X., Tian, M., Li, X., Yu, Y., Sun, Q., Ding, Q., and Jia, D. (2021). Molecular determinants for regulation of G3BP1/2 phase separation by the SARS-CoV-2 nucleocapsid protein. Cell Discov *7*, 69.

Hwang, J.W., Kim, S.N., Myung, N., Song, D., Han, G., Bae, G.U., Bedford, M.T., and Kim, Y.K. (2020). PRMT5 promotes DNA repair through methylation of 53BP1 and is regulated by Src-mediated phosphorylation. Commun Biol *3*, 428.

Iserman, C., Roden, C.A., Boerneke, M.A., Sealfon, R.S.G., McLaughlin, G.A., Jungreis, I., Fritch, E.J., Hou, Y.J., Ekena, J., Weidmann, C.A., *et al.* (2020). Genomic RNA Elements Drive Phase Separation of the SARS-CoV-2 Nucleocapsid. Mol Cell *80*, 1078-1091 e1076.

Jackson-Weaver, O., Ungvijanpunya, N., Yuan, Y., Qian, J., Gou, Y., Wu, J., Shen, H., Chen, Y., Li, M., Richard, S., *et al.* (2020). PRMT1-p53 Pathway Controls Epicardial EMT and Invasion. Cell Rep *31*, 107739.

Jackson, L.A., Anderson, E.J., Rouphael, N.G., Roberts, P.C., Makhene, M., Coler, R.N., McCullough, M.P., Chappell, J.D., Denison, M.R., Stevens, L.J., *et al.* (2020). An mRNA Vaccine against SARS-CoV-2 - Preliminary Report. N Engl J Med *383*, 1920-1931.

Jain, K., and Clarke, S.G. (2019). PRMT7 as a unique member of the protein arginine methyltransferase family: A review. Arch Biochem Biophys *665*, 36-45.

Johnston, S.R. (2010). New strategies in estrogen receptor-positive breast cancer. Clin Cancer Res *16*, 1979-1987.

Ju, X., Zhu, Y., Wang, Y., Li, J., Zhang, J., Gong, M., Ren, W., Li, S., Zhong, J., Zhang, L., *et al.* (2021). A novel cell culture system modeling the SARS-CoV-2 life cycle. PLoS Pathog *17*, e1009439.

Kang, S., Yang, M., Hong, Z., Zhang, L., Huang, Z., Chen, X., He, S., Zhou, Z., Zhou, Z., Chen, Q., *et al.* (2020). Crystal structure of SARS-CoV-2 nucleocapsid protein RNA binding domain reveals potential unique drug targeting sites. Acta Pharm Sin B *10*, 1228-1238.

Kaur, T., Raju, M., Alshareedah, I., Davis, R.B., Potoyan, D.A., and Banerjee, P.R. (2021). Sequence-encoded and composition-dependent protein-RNA interactions control multiphasic condensate morphologies. Nat Commun *12*, 872.

Kearney, P.L., Bhatia, M., Jones, N.G., Yuan, L., Glascock, M.C., Catchings, K.L., Yamada, M., and Thompson, P.R. (2005). Kinetic characterization of protein arginine deiminase 4: a transcriptional corepressor implicated in the onset and progression of rheumatoid arthritis. Biochemistry *44*, 10570-10582.

Kernohan, K.D., McBride, A., Xi, Y., Martin, N., Schwartzentruber, J., Dyment, D.A., Majewski, J., Blaser, S., Care4Rare Canada, C., Boycott, K.M., *et al.* (2017). Loss of the arginine methyltranserase PRMT7 causes syndromic intellectual disability with microcephaly and brachydactyly. Clin Genet *91*, 708-716.

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.

Kim, D., Lee, J.Y., Yang, J.S., Kim, J.W., Kim, V.N., and Chang, H. (2020a). The Architecture of SARS-CoV-2 Transcriptome. Cell *181*, 914-921.

Kim, H., and Ronai, Z.A. (2020). PRMT5 function and targeting in cancer. Cell Stress *4*, 199-215.

Kim, H., Yoon, B.H., Oh, C.M., Lee, J., Lee, K., Song, H., Kim, E., Yi, K., Kim, M.Y., Kim, H., *et al.* (2020b). PRMT1 Is Required for the Maintenance of Mature beta-Cell Identity. Diabetes *69*, 355-368.

Kim, K.H., and Roberts, C.W. (2016). Targeting EZH2 in cancer. Nat Med *22*, 128-134.

Kim, S., Gunesdogan, U., Zylicz, J.J., Hackett, J.A., Cougot, D., Bao, S., Lee, C., Dietmann, S., Allen, G.E., Sengupta, R., *et al.* (2014). PRMT5 protects genomic integrity during global DNA demethylation in primordial germ cells and preimplantation embryos. Mol Cell *56*, 564-579.

King, A. (2020). An uncommon cold. New Sci *246*, 32-35.

Kirino, Y., Kim, N., de Planell-Saguer, M., Khandros, E., Chiorean, S., Klein, P.S., Rigoutsos, I., Jongens, T.A., and Mourelatos, Z. (2009). Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability. Nat Cell Biol *11*, 652-658.

Koyuncu, O.O., and Dobner, T. (2009). Arginine methylation of human adenovirus type 5 L4 100-kilodalton protein is required for efficient virus production. J Virol *83*, 4778-4790.

Krainer, A.R., Conway, G.C., and Kozak, D. (1990). The essential pre-mRNA splicing factor SF2 influences 5' splice site selection by activating proximal sites. Cell *62*, 35-42.

Kroschwald, S., Munder, M.C., Maharana, S., Franzmann, T.M., Richter, D., Ruer, M., Hyman, A.A., and Alberti, S. (2018). Different Material States of Pub1 Condensates Define Distinct Modes of Stress Adaptation and Recovery. Cell Rep *23*, 3327-3339.

Kryukov, G.V., Wilson, F.H., Ruth, J.R., Paulk, J., Tsherniak, A., Marlow, S.E., Vazquez, F., Weir, B.A., Fitzgerald, M.E., Tanaka, M., *et al.* (2016). MTAP deletion confers enhanced dependency on the PRMT5 arginine methyltransferase in cancer cells. Science *351*, 1214-1218.

Kumar, R., Mehta, D., Mishra, N., Nayak, D., and Sunil, S. (2020). Role of Host-Mediated Post-Translational Modifications (PTMs) in RNA Virus Pathogenesis. Int J Mol Sci *22*.

Kwiatkowski, T.J., Jr., Bosco, D.A., Leclerc, A.L., Tamrazian, E., Vanderburg, C.R., Russ, C., Davis, A., Gilchrist, J., Kasarskis, E.J., Munsat, T., *et al.* (2009). Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. Science *323*, 1205-1208.

Le Romancer, M., Treilleux, I., Leconte, N., Robin-Lespinasse, Y., Sentis, S., Bouchekioua-Bouzaghou, K., Goddard, S., Gobert-Gosse, S., and Corbo, L. (2008). Regulation of estrogen rapid signaling through arginine methylation by PRMT1. Mol Cell *31*, 212-221.

Leem, Y.E., Bae, J.H., Jeong, H.J., and Kang, J.S. (2019). PRMT7 deficiency enhances adipogenesis through modulation of C/EBP-beta. Biochem Biophys Res Commun *517*, 484-490.

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

Lefkowitz, E.J., Dempsey, D.M., Hendrickson, R.C., Orton, R.J., Siddell, S.G., and Smith, D.B. (2018). Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV). Nucleic Acids Res *46*, D708-D717.

Li, W., Hu, J., Shi, B., Palomba, F., Digman, M.A., Gratton, E., and Jiang, H. (2020a). Biophysical properties of AKAP95 protein condensates regulate splicing and tumorigenesis. Nat Cell Biol *22*, 960-972.

Li, X., Hu, X., Patel, B., Zhou, Z., Liang, S., Ybarra, R., Qiu, Y., Felsenfeld, G., Bungert, J., and Huang, S. (2010). H4R3 methylation facilitates beta-globin transcription by regulating histone acetyltransferase binding and H3 acetylation. Blood *115*, 2028-2037.

Li, X., Wang, C., Jiang, H., and Luo, C. (2019). A patent review of arginine methyltransferase inhibitors (2010-2018). Expert Opin Ther Pat *29*, 97-114.

Li, Z., Wang, D., Lu, J., Huang, B., Wang, Y., Dong, M., Fan, D., Li, H., Gao, Y., Hou, P., *et al.* (2020b). Methylation of EZH2 by PRMT1 regulates its stability and promotes breast cancer metastasis. Cell Death Differ *27*, 3226-3242.

Li, Z., Wang, D., Wang, W., Chen, X., Tang, A., Hou, P., Li, M., Zheng, J., and Bai, J. (2020c). Macrophages-stimulated PRMT1-mediated EZH2 methylation promotes breast cancer metastasis. Biochem Biophys Res Commun *533*, 679-684.

Liao, S.E., and Regev, O. (2021). Splicing at the phase-separated nuclear speckle interface: a model. Nucleic Acids Res *49*, 636-645.

Lin, S.M., Lin, S.C., Hsu, J.N., Chang, C.K., Chien, C.M., Wang, Y.S., Wu, H.Y., Jeng, U.S., Kehn-Hall, K., and Hou, M.H. (2020). Structure-Based Stabilization of Non-native Protein-Protein Interactions of Coronavirus Nucleocapsid Proteins in Antiviral Drug Design. J Med Chem *63*, 3131-3141.

Liu, F., Xu, Y., Lu, X., Hamard, P.J., Karl, D.L., Man, N., Mookhtiar, A.K., Martinez, C., Lossos, I.S., Sun, J., *et al.* (2020). PRMT5-mediated histone arginine methylation antagonizes transcriptional repression by polycomb complex PRC2. Nucleic Acids Res *48*, 2956-2968.

Liu, N., Zhou, K.I., Parisien, M., Dai, Q., Diatchenko, L., and Pan, T. (2017). N6-methyladenosine alters RNA structure to regulate binding of a low-complexity protein. Nucleic Acids Res *45*, 6051-6063.

Liu, Q., and Dreyfuss, G. (1995). In vivo and in vitro arginine methylation of RNA-binding proteins. Mol Cell Biol *15*, 2800-2808.

Liu, W., Ma, Q., Wong, K., Li, W., Ohgi, K., Zhang, J., Aggarwal, A., and Rosenfeld, M.G. (2013). Brd4 and JMJD6-associated anti-pause enhancers in regulation of transcriptional pause release. Cell *155*, 1581-1595.

Lu, S., Ye, Q., Singh, D., Cao, Y., Diedrich, J.K., Yates, J.R., 3rd, Villa, E., Cleveland, D.W., and Corbett, K.D. (2021). The SARS-CoV-2 nucleocapsid phosphoprotein forms mutually exclusive condensates with RNA and the membrane-associated M protein. Nat Commun *12*, 502.

Luo, H., Chen, J., Chen, K., Shen, X., and Jiang, H. (2006). Carboxyl terminus of severe acute respiratory syndrome coronavirus nucleocapsid protein: self-association analysis and nucleic acid binding characterization. Biochemistry *45*, 11827-11835.

Luo, L., Li, Z., Zhao, T., Ju, X., Ma, P., Jin, B., Zhou, Y., He, S., Huang, J., Xu, X., *et al.* (2021). SARS-CoV-2 nucleocapsid protein phase separates with G3BPs to disassemble stress granules and facilitate viral production. Sci Bull (Beijing) *66*, 1194-1204.

Luo, Y., Yue, W., Quan, X., Wang, Y., Zhao, B., and Lu, Z. (2015). Asymmetric dimethylarginine exacerbates Abeta-induced toxicity and oxidative stress in human cell and Caenorhabditis elegans models of Alzheimer disease. Free Radic Biol Med *79*, 117-126.

Luscombe, N.M., Laskowski, R.A., and Thornton, J.M. (2001). Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. Nucleic Acids Res *29*, 2860-2874.

Ma, D., Yang, M., Wang, Q., Sun, C., Shi, H., Jing, W., Bi, Y., Shen, X., Ma, X., Qin, Z., *et al.* (2021). Arginine methyltransferase PRMT5 negatively regulates cGAS-mediated antiviral immune response. Sci Adv *7*.

Ma, K., Inglis, J.D., Sharkey, A., Bickmore, W.A., Hill, R.E., Prosser, E.J., Speed, R.M., Thomson, E.J., Jobling, M., Taylor, K., *et al.* (1993). A Y chromosome gene family with RNA-binding protein homology: candidates for the azoospermia factor AZF controlling human spermatogenesis. Cell *75*, 1287-1295.

Majumder, S., Alinari, L., Roy, S., Miller, T., Datta, J., Sif, S., Baiocchi, R., and Jacob, S.T. (2010). Methylation of histone H3 and H4 by PRMT5 regulates ribosomal RNA gene transcription. J Cell Biochem *109*, 553-563.

Malbeteau, L., Poulard, C., Languilaire, C., Mikaelian, I., Flamant, F., Le Romancer, M., and Corbo, L. (2020). PRMT1 Is Critical for the Transcriptional Activity and the Stability of the Progesterone Receptor. iScience *23*, 101236.

Matera, A.G., and Wang, Z. (2014). A day in the life of the spliceosome. Nat Rev Mol Cell Biol *15*, 108-121.

Mathioudaki, K., Scorilas, A., Ardavanis, A., Lymberi, P., Tsiambas, E., Devetzi, M., Apostolaki, A., and Talieri, M. (2011). Clinical evaluation of PRMT1 gene expression in breast cancer. Tumour Biol *32*, 575-582.

Mavrakis, K.J., McDonald, E.R., 3rd, Schlabach, M.R., Billy, E., Hoffman, G.R., deWeck, A., Ruddy, D.A., Venkatesan, K., Yu, J., McAllister, G., *et al.* (2016). Disordered methionine metabolism in MTAP/CDKN2A-deleted cancers leads to dependence on PRMT5. Science *351*, 1208-1213.

McBride, R., van Zyl, M., and Fielding, B.C. (2014). The coronavirus nucleocapsid is a multifunctional protein. Viruses *6*, 2991-3018.

Meister, G., Eggert, C., Buhler, D., Brahms, H., Kambach, C., and Fischer, U. (2001). Methylation of Sm proteins by a complex containing PRMT5 and the putative U snRNP assembly factor pICln. Curr Biol *11*, 1990-1994.

Mersaoui, S.Y., Yu, Z., Coulombe, Y., Karam, M., Busatto, F.F., Masson, J.Y., and Richard, S. (2019). Arginine methylation of the DDX5 helicase RGG/RG motif by PRMT5 regulates resolution of RNA:DNA hybrids. EMBO J *38*, e100986.

Migliori, V., Muller, J., Phalke, S., Low, D., Bezzi, M., Mok, W.C., Sahu, S.K., Gunaratne, J., Capasso, P., Bassi, C., *et al.* (2012). Symmetric dimethylation of H3R2 is a newly identified histone mark that supports euchromatin maintenance. Nat Struct Mol Biol *19*, 136-144.

Millar, A.H., Heazlewood, J.L., Giglione, C., Holdsworth, M.J., Bachmair, A., and Schulze, W.X. (2019). The Scope, Functions, and Dynamics of Posttranslational Protein Modifications. Annu Rev Plant Biol *70*, 119-151.

Miller, C.L. (2011). Stress Granules and Virus Replication. Future Virol *6*, 1329-1338.

Miranda, T.B., Khusial, P., Cook, J.R., Lee, J.H., Gunderson, S.I., Pestka, S., Zieve, G.W., and Clarke, S. (2004a). Spliceosome Sm proteins D1, D3, and B/B' are asymmetrically dimethylated at arginine residues in the nucleus. Biochem Biophys Res Commun *323*, 382-387.

Miranda, T.B., Miranda, M., Frankel, A., and Clarke, S. (2004b). PRMT7 is a member of the protein arginine methyltransferase family with a distinct substrate specificity. J Biol Chem *279*, 22902-22907.

Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A.P., Kim, H.J., Mittag, T., and Taylor, J.P. (2015). Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. Cell *163*, 123-133.

Mori, K., Arzberger, T., Grasser, F.A., Gijselinck, I., May, S., Rentzsch, K., Weng, S.M., Schludi, M.H., van der Zee, J., Cruts, M., *et al.* (2013). Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. Acta Neuropathol *126*, 881-893.

Mouaikel, J., Verheggen, C., Bertrand, E., Tazi, J., and Bordonne, R. (2002). Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus. Mol Cell *9*, 891-901.
Moursy, A., Allain, F.H., and Clery, A. (2014). Characterization of the RNA recognition mode of hnRNP G extends its role in SMN2 splicing regulation. Nucleic Acids Res *42*, 6659-6672.
Mullers, E., Uhlig, T., Stirnnagel, K., Fiebig, U., Zentgraf, H., and Lindemann, D. (2011). Novel functions of prototype foamy virus Gag glycine- arginine-rich boxes in reverse transcription and particle morphogenesis. J Virol *85*, 1452-1463.
Munschauer, M., Nguyen, C.T., Sirokman, K., Hartigan, C.R., Hogstrom, L., Engreitz, J.M., Ulirsch, J.C., Fulco, C.P., Subramanian, V., Chen, J., *et al.* (2018). The NORAD lncRNA assembles a topoisomerase complex critical for genome stability. Nature *561*, 132-136.
Murthy, A.C., Dignon, G.L., Kan, Y., Zerze, G.H., Parekh, S.H., Mittal, J., and Fawzi, N.L. (2019). Molecular interactions underlying liquid-liquid phase separation of the FUS low-complexity domain. Nat Struct Mol Biol *26*, 637-648.
Nabeel-Shah, S., Lee, H., Ahmed, N., Marcon, E., Farhangmehr, S., Pu, S., Burke, G.L., Ashraf, K., Wei, H., Zhong, G., *et al.* (2020). SARS-CoV-2 Nucleocapsid protein attenuates stress granule formation and alters gene expression via direct interaction with host mRNAs. bioRxiv, 2020.2010.2023.342113.
Nachtergaele, S., and He, C. (2017). The emerging biology of RNA post-transcriptional modifications. RNA Biol *14*, 156-163.
Nakagawa, K., Narayanan, K., Wada, M., and Makino, S. (2018). Inhibition of Stress Granule Formation by Middle East Respiratory Syndrome Coronavirus 4a Accessory Protein Facilitates Viral Translation, Leading to Efficient Virus Replication. J Virol *92*.
Neuenkirchen, N., Englbrecht, C., Ohmer, J., Ziegenhals, T., Chari, A., and Fischer, U. (2015). Reconstitution of the human U snRNP assembly machinery reveals stepwise Sm protein organization. EMBO J *34*, 1925-1941.
Nigg, E.A., Baeuerle, P.A., and Luhrmann, R. (1991). Nuclear import-export: in search of signals and mechanisms. Cell *66*, 15-22.
Nikolakaki, E., and Giannakouros, T. (2020). SR/RS Motifs as Critical Determinants of Coronavirus Life Cycle. Front Mol Biosci *7*, 219.
Nott, T.J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowietz, A., Craggs, T.D., Bazett-Jones, D.P., Pawson, T., Forman-Kay, J.D., *et al.* (2015). Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. Mol Cell *57*, 936-947.
Paik, W.K., Kim, S., and Lee, H.W. (1972). Protein methylation during the development of rat brain. Biochem Biophys Res Commun *46*, 933-941.
Pan, Q., Shai, O., Lee, L.J., Frey, B.J., and Blencowe, B.J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet *40*, 1413-1415.
Passos, D.O., Quaresma, A.J., and Kobarg, J. (2006). The methylation of the C-terminal region of hnRNPQ (NSAP1) is important for its nuclear localization. Biochem Biophys Res Commun *346*, 517-525.
Pastore, F., Bhagwat, N., Pastore, A., Radzisheuskaya, A., Karzai, A., Krishnan, A., Li, B., Bowman, R.L., Xiao, W., Viny, A.D., *et al.* (2020). PRMT5 Inhibition Modulates E2F1 Methylation and Gene-Regulatory Networks Leading to Therapeutic Efficacy in JAK2(V617F)-Mutant MPN. Cancer Discov *10*, 1742-1757.

Patel, A., Lee, H.O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M.Y., Stoynov, S., Mahamid, J., Saha, S., Franzmann, T.M., *et al.* (2015). A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. Cell *162*, 1066-1077.

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.

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

Perdikari, T.M., Murthy, A.C., Ryan, V.H., Watters, S., Naik, M.T., and Fawzi, N.L. (2020). SARS-CoV-2 nucleocapsid protein phase-separates with RNA and with human hnRNPs. EMBO J *39*, e106478.

Poornima, G., Mythili, R., Nag, P., Parbin, S., Verma, P.K., Hussain, T., and Rajyaguru, P.I. (2019). RGG-motif self-association regulates eIF4G-binding translation repressor protein Scd6. RNA Biol *16*, 1215-1227.

Poquerusse, J., Whitford, W., Taylor, J., Alburaiky, S., Snell, R.G., Lehnert, K., and Jacobsen, J.C. (2021). Novel PRMT7 mutation in a rare case of dysmorphism and intellectual disability. J Hum Genet.

Poulard, C., Treilleux, I., Lavergne, E., Bouchekioua-Bouzaghou, K., Goddard-Leon, S., Chabaud, S., Tredan, O., Corbo, L., and Le Romancer, M. (2012). Activation of rapid oestrogen signalling in aggressive human breast cancers. EMBO Mol Med *4*, 1200-1213.

Premasiri, A.S., Gill, A.L., and Vieira, F.G. (2020). Type I PRMT Inhibition Protects Against C9ORF72 Arginine-Rich Dipeptide Repeat Toxicity. Front Pharmacol *11*, 569661.

Prieto, C., Nguyen, D.T.T., Liu, Z., Wheat, J., Perez, A., Gourkanti, S., Chou, T., Barin, E., Velleca, A., Rohwetter, T., *et al.* (2021). Transcriptional control of CBX5 by the RNA binding proteins RBMX and RBMXL1 maintains chromatin state in myeloid leukemia. Nat Cancer *2*, 741-757.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., De Bakker, P.I., and Daly, M.J. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. The American journal of human genetics *81*, 559-575.

Qamar, S., Wang, G., Randle, S.J., Ruggeri, F.S., Varela, J.A., Lin, J.Q., Phillips, E.C., Miyashita, A., Williams, D., Strohl, F., *et al.* (2018). FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation-pi Interactions. Cell *173*, 720-734 e715.

Quiroz, F.G., and Chilkoti, A. (2015). Sequence heuristics to encode phase behaviour in intrinsically disordered protein polymers. Nat Mater *14*, 1164-1171.

Radzisheuskaya, A., Shliaha, P.V., Grinev, V., Lorenzini, E., Kovalchuk, S., Shlyueva, D., Gorshkov, V., Hendrickson, R.C., Jensen, O.N., and Helin, K. (2019). PRMT5 methylome profiling uncovers a direct link to splicing regulation in acute myeloid leukemia. Nat Struct Mol Biol *26*, 999-1012.

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.

Raker, V.A., Hartmuth, K., Kastner, B., and Luhrmann, R. (1999). Spliceosomal U snRNP core assembly: Sm proteins assemble onto an Sm site RNA nonanucleotide in a specific and thermodynamically stable manner. Mol Cell Biol *19*, 6554-6565.

Rank, G., Cerruti, L., Simpson, R.J., Moritz, R.L., Jane, S.M., and Zhao, Q. (2010). Identification of a PRMT5-dependent repressor complex linked to silencing of human fetal globin gene expression. Blood *116*, 1585-1592.

Rauscher, S., and Pomes, R. (2017). The liquid structure of elastin. Elife *6*.

Ray, M., Sarkar, S., and Rath, S.N. (2020). Druggability for COVID-19: in silico discovery of potential drug compounds against nucleocapsid (N) protein of SARS-CoV-2. Genomics Inform *18*, e43.

Reintjes, A., Fuchs, J.E., Kremser, L., Lindner, H.H., Liedl, K.R., Huber, L.A., and Valovka, T. (2016). Asymmetric arginine dimethylation of RelA provides a repressive mark to modulate TNFalpha/NF-kappaB response. Proc Natl Acad Sci U S A *113*, 4326-4331.

Rengasamy, M., Zhang, F., Vashisht, A., Song, W.M., Aguilo, F., Sun, Y., Li, S., Zhang, W., Zhang, B., Wohlschlegel, J.A., *et al.* (2017). The PRMT5/WDR77 complex regulates alternative splicing through ZNF326 in breast cancer. Nucleic Acids Res *45*, 11106-11120.

Riback, J.A., Katanski, C.D., Kear-Scott, J.L., Pilipenko, E.V., Rojek, A.E., Sosnick, T.R., and Drummond, D.A. (2017). Stress-Triggered Phase Separation Is an Adaptive, Evolutionarily Tuned Response. Cell *168*, 1028-1040 e1019.

Riva, L., Yuan, S., Yin, X., Martin-Sancho, L., Matsunaga, N., Pache, L., Burgstaller-Muehlbacher, S., De Jesus, P.D., Teriete, P., Hull, M.V., *et al.* (2020). Discovery of SARS-CoV-2 antiviral drugs through large-scale compound repurposing. Nature *586*, 113-119.

Roscigno, R.F., and Garcia-Blanco, M.A. (1995). SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome. RNA *1*, 692-706.

Ryan, V.H., Dignon, G.L., Zerze, G.H., Chabata, C.V., Silva, R., Conicella, A.E., Amaya, J., Burke, K.A., Mittal, J., and Fawzi, N.L. (2018). Mechanistic View of hnRNPA2 Low-Complexity Domain Structure, Interactions, and Phase Separation Altered by Mutation and Arginine Methylation. Mol Cell *69*, 465-479 e467.

Sachamitr, P., Ho, J.C., Ciamponi, F.E., Ba-Alawi, W., Coutinho, F.J., Guilhamon, P., Kushida, M.M., Cavalli, F.M.G., Lee, L., Rastegar, N., *et al.* (2021). PRMT5 inhibition disrupts splicing and stemness in glioblastoma. Nat Commun *12*, 979.

Saito, M., Hess, D., Eglinger, J., Fritsch, A.W., Kreysing, M., Weinert, B.T., Choudhary, C., and Matthias, P. (2019). Acetylation of intrinsically disordered regions regulates phase separation. Nat Chem Biol *15*, 51-61.

Santos-Pereira, J.M., and Aguilera, A. (2015). R loops: new modulators of genome dynamics and function. Nat Rev Genet *16*, 583-597.

Savastano, A., Ibanez de Opakua, A., Rankovic, M., and Zweckstetter, M. (2020). Nucleocapsid protein of SARS-CoV-2 phase separates into RNA-rich polymerase-containing condensates. Nat Commun *11*, 6041.

Schmidt, N., Lareau, C.A., Keshishian, H., Ganskih, S., Schneider, C., Hennig, T., Melanson, R., Werner, S., Wei, Y., Zimmer, M., *et al.* (2021). The SARS-CoV-2 RNA-protein interactome in infected human cells. Nat Microbiol *6*, 339-353.

Schurter, B.T., Koh, S.S., Chen, D., Bunick, G.J., Harp, J.M., Hanson, B.L., Henschen-Edman, A., Mackay, D.R., Stallcup, M.R., and Aswad, D.W. (2001). Methylation of histone H3 by coactivator-associated arginine methyltransferase 1. Biochemistry *40*, 5747-5756.

Scully, R., Panday, A., Elango, R., and Willis, N.A. (2019). DNA double-strand break repair-pathway choice in somatic mammalian cells. Nat Rev Mol Cell Biol *20*, 698-714.
Secker, K.A., Keppeler, H., Duerr-Stoerzer, S., Schmid, H., Schneidawind, D., Hentrich, T., Schulze-Hentrich, J.M., Mankel, B., Fend, F., and Schneidawind, C. (2019). Inhibition of DOT1L and PRMT5 promote synergistic anti-tumor activity in a human MLL leukemia model induced by CRISPR/Cas9. Oncogene *38*, 7181-7195.
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.
Sen, S., He, Z., Ghosh, S., Dery, K.J., Yang, L., Zhang, J., and Sun, Z. (2018). PRMT1 Plays a Critical Role in Th17 Differentiation by Regulating Reciprocal Recruitment of STAT3 and STAT5. J Immunol *201*, 440-450.
Sharma, S., Maris, C., Allain, F.H., and Black, D.L. (2011). U1 snRNA directly interacts with polypyrimidine tract-binding protein during splicing repression. Mol Cell *41*, 579-588.
Shashi, V., Xie, P., Schoch, K., Goldstein, D.B., Howard, T.D., Berry, M.N., Schwartz, C.E., Cronin, K., Sliwa, S., Allen, A., *et al.* (2015). The RBMX gene as a candidate for the Shashi X-linked intellectual disability syndrome. Clin Genet *88*, 386-390.
Shaw, D.J., Morse, R., Todd, A.G., Eggleton, P., Lorson, C.L., and Young, P.J. (2010). Identification of a self-association domain in the Ewing's sarcoma protein: a novel function for arginine-glycine-glycine rich motifs? J Biochem *147*, 885-893.
Shia, W.J., Okumura, A.J., Yan, M., Sarkeshik, A., Lo, M.C., Matsuura, S., Komeno, Y., Zhao, X., Nimer, S.D., Yates, J.R., 3rd, *et al.* (2012). PRMT1 interacts with AML1-ETO to promote its transcriptional activation and progenitor cell proliferative potential. Blood *119*, 4953-4962.
Shin, K.H., Kim, R.H., Kang, M.K., Kim, R.H., Kim, S.G., Lim, P.K., Yochim, J.M., Baluda, M.A., and Park, N.H. (2007). p53 promotes the fidelity of DNA end-joining activity by, in part, enhancing the expression of heterogeneous nuclear ribonucleoprotein G. DNA Repair (Amst) *6*, 830-840.
Simard, M.J., and Chabot, B. (2002). SRp30c is a repressor of 3' splice site utilization. Mol Cell Biol *22*, 4001-4010.
Sims, R.J., 3rd, Rojas, L.A., Beck, D.B., Bonasio, R., Schuller, R., Drury, W.J., 3rd, Eick, D., and Reinberg, D. (2011). The C-terminal domain of RNA polymerase II is modified by site-specific methylation. Science *332*, 99-103.
Sinha, R., Allemand, E., Zhang, Z., Karni, R., Myers, M.P., and Krainer, A.R. (2010). Arginine methylation controls the subcellular localization and functions of the oncoprotein splicing factor SF2/ASF. Mol Cell Biol *30*, 2762-2774.
Siriboonpiputtana, T., Zeisig, B.B., Zarowiecki, M., Fung, T.K., Mallardo, M., Tsai, C.T., Lau, P.N.I., Hoang, Q.C., Veiga, P., Barnes, J., *et al.* (2017). Transcriptional memory of cells of origin overrides beta-catenin requirement of MLL cancer stem cells. EMBO J *36*, 3139-3155.
Smith, D.L., Erce, M.A., Lai, Y.W., Tomasetig, F., Hart-Smith, G., Hamey, J.J., and Wilkins, M.R. (2020). Crosstalk of Phosphorylation and Arginine Methylation in Disordered SRGG Repeats of Saccharomycescerevisiae Fibrillarin and Its Association with Nucleolar Localization. J Mol Biol *432*, 448-466.
Smith, Z.D., Chan, M.M., Mikkelsen, T.S., Gu, H., Gnirke, A., Regev, A., and Meissner, A. (2012). A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature *484*, 339-344.
So, C.W., Caldas, C., Liu, M.M., Chen, S.J., Huang, Q.H., Gu, L.J., Sham, M.H., Wiedemann, L.M., and Chan, L.C. (1997). EEN encodes for a member of a new family of proteins

containing an Src homology 3 domain and is the third gene located on chromosome 19p13 that fuses to MLL in human leukemia. Proc Natl Acad Sci U S A *94*, 2563-2568.
Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.
Sun, Q., Mayeda, A., Hampson, R.K., Krainer, A.R., and Rottman, F.M. (1993). General splicing factor SF2/ASF promotes alternative splicing by binding to an exonic splicing enhancer. Genes Dev *7*, 2598-2608.
Sylvestersen, K.B., Horn, H., Jungmichel, S., Jensen, L.J., and Nielsen, M.L. (2014). Proteomic analysis of arginine methylation sites in human cells reveals dynamic regulation during transcriptional arrest. Mol Cell Proteomics *13*, 2072-2088.
Szewczyk, M.M., Ishikawa, Y., Organ, S., Sakai, N., Li, F., Halabelian, L., Ackloo, S., Couzens, A.L., Eram, M., Dilworth, D., *et al.* (2020). Pharmacological inhibition of PRMT7 links arginine monomethylation to the cellular stress response. Nat Commun *11*, 2396.
Takahama, K., Miyawaki, A., Shitara, T., Mitsuya, K., Morikawa, M., Hagihara, M., Kino, K., Yamamoto, A., and Oyoshi, T. (2015). G-Quadruplex DNA- and RNA-Specific-Binding Proteins Engineered from the RGG Domain of TLS/FUS. ACS Chem Biol *10*, 2564-2569.
Tan, Y.W., Fang, S., Fan, H., Lescar, J., and Liu, D.X. (2006). Amino acid residues critical for RNA-binding in the N-terminal domain of the nucleocapsid protein are essential determinants for the infectivity of coronavirus in cultured cells. Nucleic Acids Res *34*, 4816-4825.
Tang, J., Frankel, A., Cook, R.J., Kim, S., Paik, W.K., Williams, K.R., Clarke, S., and Herschman, H.R. (2000). PRMT1 is the predominant type I protein arginine methyltransferase in mammalian cells. J Biol Chem *275*, 7723-7730.
Tarighat, S.S., Santhanam, R., Frankhouser, D., Radomska, H.S., Lai, H., Anghelina, M., Wang, H., Huang, X., Alinari, L., Walker, A., *et al.* (2016). The dual epigenetic role of PRMT5 in acute myeloid leukemia: gene activation and repression via histone arginine methylation. Leukemia *30*, 789-799.
Tee, W.W., Pardo, M., Theunissen, T.W., Yu, L., Choudhary, J.S., Hajkova, P., and Surani, M.A. (2010). Prmt5 is essential for early mouse development and acts in the cytoplasm to maintain ES cell pluripotency. Genes Dev *24*, 2772-2777.
Thandapani, P., O'Connor, T.R., Bailey, T.L., and Richard, S. (2013). Defining the RGG/RG motif. Mol Cell *50*, 613-623.
Thandapani, P., Song, J., Gandin, V., Cai, Y., Rouleau, S.G., Garant, J.M., Boisvert, F.M., Yu, Z., Perreault, J.P., Topisirovic, I., *et al.* (2015). Aven recognition of RNA G-quadruplexes regulates translation of the mixed lineage leukemia protooncogenes. Elife *4*.
Tikhanovich, I., Zhao, J., Olson, J., Adams, A., Taylor, R., Bridges, B., Marshall, L., Roberts, B., and Weinman, S.A. (2017). Protein arginine methyltransferase 1 modulates innate immune responses through regulation of peroxisome proliferator-activated receptor gamma-dependent macrophage differentiation. J Biol Chem *292*, 6882-6894.
Tripsianes, K., Madl, T., Machyna, M., Fessas, D., Englbrecht, C., Fischer, U., Neugebauer, K.M., and Sattler, M. (2011). Structural basis for dimethylarginine recognition by the Tudor domains of human SMN and SPF30 proteins. Nat Struct Mol Biol *18*, 1414-1420.
Tsai, W.C., Gayatri, S., Reineke, L.C., Sbardella, G., Bedford, M.T., and Lloyd, R.E. (2016). Arginine Demethylation of G3BP1 Promotes Stress Granule Assembly. J Biol Chem *291*, 22671-22685.

Tsend-Ayush, E., O'Sullivan, L.A., Grutzner, F.S., Onnebo, S.M., Lewis, R.S., Delbridge, M.L., Marshall Graves, J.A., and Ward, A.C. (2005). RBMX gene is essential for brain development in zebrafish. Dev Dyn *234*, 682-688.

Uversky, V.N., Kuznetsova, I.M., Turoverov, K.K., and Zaslavsky, B. (2015). Intrinsically disordered proteins as crucial constituents of cellular aqueous two phase systems and coacervates. FEBS Lett *589*, 15-22.

Vadnais, C., Chen, R., Fraszczak, J., Yu, Z., Boulais, J., Pinder, J., Frank, D., Khandanpour, C., Hebert, J., Dellaire, G., *et al.* (2018). GFI1 facilitates efficient DNA repair by regulating PRMT1 dependent methylation of MRE11 and 53BP1. Nat Commun *9*, 1418.

Vagin, V.V., Wohlschlegel, J., Qu, J., Jonsson, Z., Huang, X., Chuma, S., Girard, A., Sachidanandam, R., Hannon, G.J., and Aravin, A.A. (2009). Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members. Genes Dev *23*, 1749-1762.

Vance, C., Rogelj, B., Hortobagyi, T., De Vos, K.J., Nishimura, A.L., Sreedharan, J., Hu, X., Smith, B., Ruddy, D., Wright, P., *et al.* (2009). Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. Science *323*, 1208-1211.

Vasilyev, N., Polonskaia, A., Darnell, J.C., Darnell, R.B., Patel, D.J., and Serganov, A. (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*, E5391-5400.

Walport, L.J., Hopkinson, R.J., Chowdhury, R., Schiller, R., Ge, W., Kawamura, A., and Schofield, C.J. (2016). Arginine demethylation is catalysed by a subset of JmjC histone lysine demethylases. Nat Commun *7*, 11974.

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 facilitating transcriptional activation by nuclear hormone receptor. Science *293*, 853-857.

Wang, X., Qiu, T., Wu, Y., Yang, C., Li, Y., Du, G., He, Y., Liu, W., Liu, R., Chen, C.H., *et al.* (2021a). Arginine methyltransferase PRMT5 methylates and stabilizes KLF5 via decreasing its phosphorylation and ubiquitination to promote basal-like breast cancer. Cell Death Differ.

Wang, Y., Li, X., Ge, J., Liu, M., Pang, X., Liu, J., Luo, C., Xu, Y., and Zhao, Q. (2021b). The methyltransferase PRMT1 regulates gamma-globin translation. J Biol Chem *296*, 100417.

Wang, Y., Wysocka, J., Sayegh, J., Lee, Y.H., Perlin, J.R., Leonelli, L., Sonbuchner, L.S., McDonald, C.H., Cook, R.G., Dou, Y., *et al.* (2004). Human PAD4 regulates histone arginine methylation levels via demethylimination. Science *306*, 279-283.

Wang, Z., Pan, Z., Adhikari, S., Harada, B.T., Shen, L., Yuan, W., Abeywardana, T., Al-Hadid, Q., Stark, J.M., He, C., *et al.* (2021c). m(6) A deposition is regulated by PRMT1-mediated arginine methylation of METTL14 in its disordered C-terminal region. EMBO J *40*, e106309.

Wegmann, S., Eftekharzadeh, B., Tepper, K., Zoltowska, K.M., Bennett, R.E., Dujardin, S., Laskowski, P.R., MacKenzie, D., Kamath, T., Commins, C., *et al.* (2018). Tau protein liquid-liquid phase separation can initiate tau aggregation. EMBO J *37*.

White, J.P., Cardenas, A.M., Marissen, W.E., and Lloyd, R.E. (2007). Inhibition of cytoplasmic mRNA stress granule formation by a viral proteinase. Cell Host Microbe *2*, 295-305.

Wu, C., Qavi, A.J., Hachim, A., Kavian, N., Cole, A.R., Moyle, A.B., Wagner, N.D., Sweeney-Gibbons, J., Rohrs, H.W., Gross, M.L., *et al.* (2021). Characterization of SARS-CoV-2 nucleocapsid protein reveals multiple functional consequences of the C-terminal domain. iScience *24*, 102681.

Wu, J.Y., Kar, A., Kuo, D., Yu, B., and Havlioglu, N. (2006). SRp54 (SFRS11), a regulator for tau exon 10 alternative splicing identified by an expression cloning strategy. Mol Cell Biol *26*, 6739-6747.

Yan, Q., Zeng, P., Zhou, X., Zhao, X., Chen, R., Qiao, J., Feng, L., Zhu, Z., Zhang, G., and Chen, C. (2021). RBMX suppresses tumorigenicity and progression of bladder cancer by interacting with the hnRNP A1 protein to regulate PKM alternative splicing. Oncogene *40*, 2635-2650.

Yang, M.L., Doyle, H.A., Clarke, S.G., Herold, K.C., and Mamula, M.J. (2018). Oxidative Modifications in Tissue Pathology and Autoimmune Disease. Antioxid Redox Signal *29*, 1415-1431.

Yang, P., Mathieu, C., Kolaitis, R.M., Zhang, P., Messing, J., Yurtsever, U., Yang, Z., Wu, J., Li, Y., Pan, Q., *et al.* (2020). G3BP1 Is a Tunable Switch that Triggers Phase Separation to Assemble Stress Granules. Cell *181*, 325-345 e328.

Yang, Y., Lu, Y., Espejo, A., Wu, J., Xu, W., Liang, S., and Bedford, M.T. (2010). TDRD3 is an effector molecule for arginine-methylated histone marks. Mol Cell *40*, 1016-1023.

Yang, Y., McBride, K.M., Hensley, S., Lu, Y., Chedin, F., and Bedford, M.T. (2014). Arginine methylation facilitates the recruitment of TOP3B to chromatin to prevent R loop accumulation. Mol Cell *53*, 484-497.

Yao, B., Gui, T., Zeng, X., Deng, Y., Wang, Z., Wang, Y., Yang, D., Li, Q., Xu, P., Hu, R., *et al.* (2021). PRMT1-mediated H4R3me2a recruits SMARCA4 to promote colorectal cancer progression by enhancing EGFR signaling. Genome Med *13*, 58.

Yaron, T.M., Heaton, B.E., Levy, T.M., Johnson, J.L., Jordan, T.X., Cohen, B.M., Kerelsky, A., Lin, T.-Y., Liberatore, K.M., Bulaon, D.K., *et al.* (2020). The FDA-approved drug Alectinib compromises SARS-CoV-2 nucleocapsid phosphorylation and inhibits viral infection in vitro. bioRxiv, 2020.2008.2014.251207.

Ye, Q., West, A.M.V., Silletti, S., and Corbett, K.D. (2020). Architecture and self-assembly of the SARS-CoV-2 nucleocapsid protein. Protein Sci *29*, 1890-1901.

Yi, Z., Fang, C., Pan, T., Wang, J., Yang, P., and Yuan, Z. (2006). Subproteomic study of hepatitis C virus replicon reveals Ras-GTPase-activating protein binding protein 1 as potential HCV RC component. Biochem Biophys Res Commun *350*, 174-178.

Yin, S., Liu, L., Brobbey, C., Palanisamy, V., Ball, L.E., Olsen, S.K., Ostrowski, M.C., and Gan, W. (2021). PRMT5-mediated arginine methylation activates AKT kinase to govern tumorigenesis. Nat Commun *12*, 3444.

Ying, Z., Mei, M., Zhang, P., Liu, C., He, H., Gao, F., and Bao, S. (2015). Histone Arginine Methylation by PRMT7 Controls Germinal Center Formation via Regulating Bcl6 Transcription. J Immunol *195*, 1538-1547.

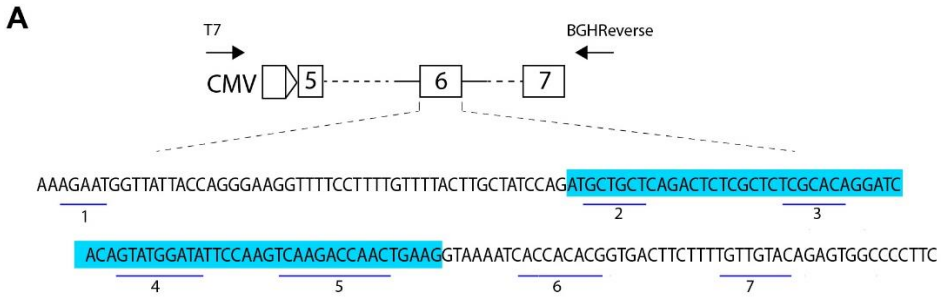
Yu, I.M., Oldham, M.L., Zhang, J., and Chen, J. (2006). Crystal structure of the severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein dimerization domain reveals evolutionary linkage between corona- and arteriviridae. J Biol Chem *281*, 17134-17139.

Yu, Z., Vogel, G., Coulombe, Y., Dubeau, D., Spehalski, E., Hebert, J., Ferguson, D.O., Masson, J.Y., and Richard, S. (2012). The MRE11 GAR motif regulates DNA double-strand break processing and ATR activation. Cell Res *22*, 305-320.

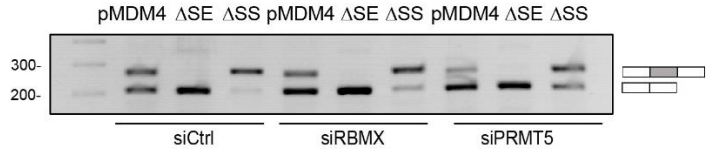
Yuan, W., Al-Hadid, Q., Wang, Z., Shen, L., Cho, H., Wu, X., and Yang, Y. (2021). TDRD3 promotes DHX9 chromatin recruitment and R-loop resolution. Nucleic Acids Res *49*, 8573-8591.

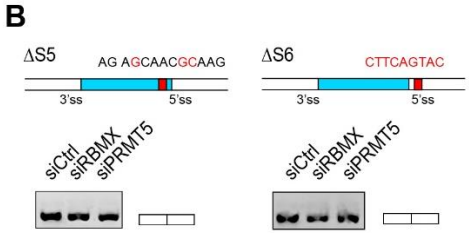
Zagrovic, B., Bartonek, L., and Polyansky, A.A. (2018). RNA-protein interactions in an unstructured context. FEBS Lett *592*, 2901-2916.
Zhang, G., Neubert, T.A., and Jordan, B.A. (2012). RNA binding proteins accumulate at the postsynaptic density with synaptic activity. J Neurosci *32*, 599-609.
Zhang, X., and Cheng, X. (2003). Structure of the predominant protein arginine methyltransferase PRMT1 and analysis of its binding to substrate peptides. Structure *11*, 509-520.
Zhang, Y., Chen, K., Sloan, S.A., Bennett, M.L., Scholze, A.R., O'Keeffe, S., Phatnani, H.P., Guarnieri, P., Caneda, C., Ruderisch, N., *et al.* (2014). An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. J Neurosci *34*, 11929-11947.
Zhang, Z., Nikolai, B.C., Gates, L.A., Jung, S.Y., Siwak, E.B., He, B., Rice, A.P., O'Malley, B.W., and Feng, Q. (2017). Crosstalk between histone modifications indicates that inhibition of arginine methyltransferase CARM1 activity reverses HIV latency. Nucleic Acids Res *45*, 9348-9360.
Zhao, Q., Rank, G., Tan, Y.T., Li, H., Moritz, R.L., Simpson, R.J., Cerruti, L., Curtis, D.J., Patel, D.J., Allis, C.D., *et al.* (2009). PRMT5-mediated methylation of histone H4R3 recruits DNMT3A, coupling histone and DNA methylation in gene silencing. Nat Struct Mol Biol *16*, 304-311.
Zhao, Y., Yin, X., Qin, H., Zhu, F., Liu, H., Yang, W., Zhang, Q., Xiang, C., Hou, P., Song, Z., *et al.* (2008). Two supporting factors greatly improve the efficiency of human iPSC generation. Cell Stem Cell *3*, 475-479.
Zhou, B., Liu, J., Wang, Q., Liu, X., Li, X., Li, P., Ma, Q., and Cao, C. (2008). The nucleocapsid protein of severe acute respiratory syndrome coronavirus inhibits cell cytokinesis and proliferation by interacting with translation elongation factor 1alpha. J Virol *82*, 6962-6971.
Zhou, K.I., Shi, H., Lyu, R., Wylder, A.C., Matuszek, Z., Pan, J.N., He, C., Parisien, M., and Pan, T. (2019). Regulation of Co-transcriptional Pre-mRNA Splicing by m(6)A through the Low-Complexity Protein hnRNPG. Mol Cell *76*, 70-81 e79.
Zhu, F., Guo, H., Bates, P.D., Zhang, S., Zhang, H., Nomie, K.J., Li, Y., Lu, L., Seibold, K.R., Wang, F., *et al.* (2019a). PRMT5 is upregulated by B-cell receptor signaling and forms a positive-feedback loop with PI3K/AKT in lymphoma cells. Leukemia *33*, 2898-2911.
Zhu, Y., He, X., Lin, Y.C., Dong, H., Zhang, L., Chen, X., Wang, Z., Shen, Y., Li, M., Wang, H., *et al.* (2019b). Targeting PRMT1-mediated FLT3 methylation disrupts maintenance of MLL-rearranged acute lymphoblastic leukemia. Blood *134*, 1257-1268.
Zurita-Lopez, C.I., Sandberg, T., Kelly, R., and Clarke, S.G. (2012). Human protein arginine methyltransferase 7 (PRMT7) is a type III enzyme forming omega-NG-monomethylated arginine residues. J Biol Chem *287*, 7859-7870.

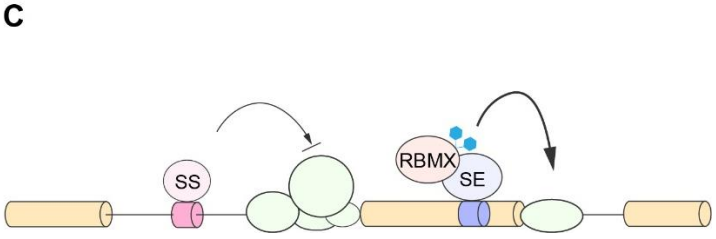
Appendix 1 Essential *cis*-elements are involved in MDM4 pre-mRNA regulation



		trans factors	*cis* elements	mutations
ΔSS	site 1	TRA2B	TGGTTATTA	TGATTTCAA
ΔSE	site 2	SRSF1, SRSF5, SRSF6	CAGATGCTGCTCA	AAGATGAATCTCA
	site 3	SRSF1, SRSF2, SRSF3, SRSF5, SRSF5, SRSF6, SRSF9	GCTCTCGCAC AGGAT	GCGCTTGCGCAAGCT
	site 4	SRSF1, SRSF2, SRSF5, SRSF6, SRSF9	CACAG TATGGATATT	CACAG TGTGCACATT
	site 5	SRSF1, SRSF5	AAG ACCAACTGAAG	CAG AGCAACGCAAG
	site 6	SRSF1, SRSF5, SRSF6, SRSF9	ACCACACGGTG	CTTCAGTACTA
	site 7	SRSF5, SRSF6	TTGTACAGAG	TTAACCAGAG







(A) Schematic representation of MDM4 minigene is presented in the upper panel. Primers designed for detecting the minigene transcripts are shown with arrows. The *MDM4* exon 6 sequence is highlighted in blue. The predicted splicing enhancers (SE) and splicing silencer (SS) sites (1-7) are shown. In the lower panel, U2OS cells transfected with indicated siRNA were further transfected with pMDM4, ΔSE and ΔSS minigene. MDM4 alternative splicing is examined using RT-PCR.
(B) Similar to (A), ΔS5 and ΔS6 minigene were generated with mutation in site 5 and site 6. *MDM4* alternative splicing is examined using RT-PCR.
(C) Schematic representation of *MDM4* alternative splicing regulation. *MDM4* alternative splicing is regulated by a splicing enhancer and a splicing silencer. Methylated RBMX participates in the regulation through affecting a splicing factor that binds to an enhancer.